University of Southampton

Faculty of Environmental and Life Sciences
School of Ocean and Earth Science

Sources and sinks of variability and predictability in the North Atlantic

DOI: 10.1002/0470841559.ch1

by

[Dafydd Petrus Stephenson]

ORCiD: 00000-0001-7358-6899

A thesis for the degree of
Doctor of Philosophy

June 2021

University of Southampton

Abstract

Faculty of Environmental and Life Sciences
School of Ocean and Earth Science

Doctor of Philosophy

Sources and sinks of variability and predictability in the North Atlantic

by [Dafydd Petrus Stephenson]

The North Atlantic has long been recognised to have a unique role in climate, owing to its ability via deep water formation to sequester large quantities of heat and carbon, and due to the associated meridional overturning circulation (MOC), which permits the northward transport of heat (impacting the climate of neighbouring regions such as Europe). In recent decades, a "data revolution" in both computational power and observational coverage and quality have revealed a significant amount of variability in the North Atlantic, but the origins of these variations and their predictability remain open questions. The simplest explanation is that variations in the North Atlantic are the passive response of the atmospheric forcing to the ocean. In this perspective, surface-borne anomalies are carried into the ocean interior along the ventilation pathways of its water masses. These pathways therefore motivate the start point for this thesis, which begins by presenting a new configuration of an ocean general circulation model (OGCM), and its application to their study. The configuration is used to describe the passive origins and fate of two ocean water masses present in the North Atlantic: North Atlantic Deep Water and North Atlantic Subtropical Mode Water. The configuration repurposes the tangent-linear and adjoint model (TAM) framework used for tracking perturbations and sensitivities, such that feedbacks are nullified and transport is passive. Using the forward and backward modes of the TAM in 400 year simulations, spatiotemporal and thermohaline probability distributions of water mass origins and fate are constructed. This highlights a disparity between the forward and backward modes suggesting an important role for water mass transformation (changes in thermohaline properties). Following this, the passive configuration is applied to exploring the passive and active nature of atmospherically forced heat content variability, and the North Atlantic is compared with other major ocean basins. A stochastic representation of atmospheric forcing is diagnosed from a coupled climate model and projected onto heat content sensitivity fields produced by the TAM. It is shown that surface layer heat content variability is primarily explained ($> 92\%$ agreement in all basins) by the behaviour of the passive ocean simulation, but that for full-depth heat content, the North Atlantic is uniquely poorly represented ($\sim 27\%$) by a passive ocean. The spatiotemporal origins of this discrepancy are

explored, revealing signatures of the MOC which slow the development of the active ocean variability. Having established an active role of the large-scale ocean dynamics when stimulated by atmospheric stochastic forcing, the relative role of ocean mesoscale eddy turbulence is then considered in the development of uncertainty in the North Atlantic. A stochastic representation of these internal buoyancy fluxes is diagnosed from an eddy-permitting ocean model and projected onto TAM sensitivity fields of subpolar and subtropical MOC and heat content. This representation is applied alongside the stochastic atmospheric representation, so that their relative contributions can be quantified. It is shown that, in the subtropics, the atmosphere ultimately generates around 60% of MOC uncertainty when averaged over a month or decade (owing primarily to momentum and buoyancy fluxes, respectively). For annually averaged MOC, however, ocean internal noise is responsible for $\sim 60\%$ of the total after 60 years. In the subpolar regions, atmospheric forcing prevails for all metrics and time averages, explaining up to $\sim 90\%$ of the unpredictable variability in the case of monthly MOC. The spatial origins of this variability are diagnosed and compared with the spatial patterns which have been determined to most efficiently stimulate variance in an optimal stochastic framework, and are in general overall agreement. Motivated by the concept that nonlinear ocean fluctuations can drive large variations in the annual average subtropical MOC (such as those recorded by modern observational arrays), a nonlinear optimal perturbation technique is then applied to an eddy-permitting OGCM in an attempt to recreate one such event from the recent observational record ("the 2009 event") via an oceanic perturbation. This perturbation, obtained by an iterative optimisation approach, is able (from an anomaly of size ~ 0.1 K) to push the ocean into a temporary state of reduced overturning (3.2 Sv decline), while a linear context, no decline was able to be induced. The linear and nonlinear optimal perturbations differ in both their magnitude and targeted spatial scales, with the former ineffectively concentrating density anomalies within smaller-scale structures and the latter targeting large-scale patterns. The work presented in this thesis collectively reaffirms the unique position of the North Atlantic, demonstrating the long lifespan of its water masses, the interaction between its circulation and atmospherically forced anomalies, and a role for the ocean in forcing substantial deviations from a mean state.

Contents

List of Figures

List of Tables

List of Abbreviations

AMV	Atlantic Multidecadal Variability
CMIP	Coupled Model Intercomparison Project
CO_2	Carbon dioxide
CPU	Central Processing Unit
DFS	Drakkar Forcing Set
FCT	Flux-Corrected Transport
FTCS	Forward-Time-Centred-Space
FWF	FreshWater Flux
HF	Heat Flux
HPC	High Performance Computing
HRM	Higher-Resolution Model
IPSL	Institute Pierre Simon Laplace
LIM	Linear Inverse Modelling *or* Louvain-la-neuve sea Ice Model
LMDZ	Laboratoire de Météorologie Dynamique Zoom
LoA	Level of Agreement
LOP	Linear Optimal Perturbation
LRM	Lower-Resolution Model
LSW	Labrador Sea Water
MHT	Meridional Heat Transport
MLD	Mixed Layer Depth
MMF	Meridional Momentum Flux
MMW	Madeira Mode Water
[A]MOC	[Atlantic] Meridional Overturning Circulation
MVT	Meridional Volume Transport
[SP/A]-NADW	[SubPolar/Arctic]-North Atlantic Deep Water
NAO	North Atlantic Oscillation
NASMW	North Atlantic Subtropical Mode Water
NEMO	Nucleus for European Modelling of the Ocean
[C]NOP	[Conditional] Nonlinear Optimal Perturbation
NYF	Normal Year Forcing
[O]GCM	[Ocean] General Circulation Model
OHC	Ocean Heat Content

OPA	Océan Parallélisé
OSP	Optimal Stochastic Perturbation
OW	Overflow Water
PDF	Probability Distribution Function
PSD	Power Spectral Density
PV	Potential Vorticity
QDR	Quad Data Rate
RMS[L]E	Root-Mean-Square [Logarithmic] Error
SSH	Sea Surface Height
SSS	Sea Surface Salinity
SST	Sea Surface Temperature
STG	SubTropical Gyre
SVD	Singular Value Decomposition
TAM	Tangent-linear and Adjoint Model
TS	Temperature-Salinity
TVD	Total Variation Diminishing
[D]WBC	[Deep] Western Boundary Current
WGN	White Gaussian Noise
ZMF	Zonal Momentum Flux

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as: Dafydd Stephenson, Simon A Müller, and Florian Sévellec. Tracking water masses using passive-tracer transport in NEMO v3. 4 with NEMOTAM: application to North Atlantic Deep Water and North Atlantic Subtropical Mode Water. *Geoscientific Model Development*, 13 (4):2031–2050, 2020

Signed: Date: 22nd June 2021

Acknowledgements

"I am who I am because of everyone"

(Mark Beaumont, endurance cyclist)

A great number of people are to thank for the existence of this thesis. It exists at its current standard, however, thanks to the efforts of my primary supervisor, Florian Sévellec, who has gone above and beyond his role by rigorously checking and re-checking every line, so that much of this work could also be submitted for publication. Thank you for this painstaking effort, as well as your encouragement to pursue my own ideas, and your help with getting out of the many resulting rabbit holes. All of this has added up to a huge help. Similarly, the standard of my computing has been raised substantially by Simon Müller, who patiently and enthusiastically sat and worked with me as I slowly became comfortable with the many environments that I now use daily. A huge thanks also to Jeff Blundell and Adam Blaker for their ever-present help and encouragement with technical matters.

I thank the other members of my PhD panel, Sybren Drijfhout and Kevin Oliver, for providing helpful feedback and support, and making sure I stayed on track. Kevin in particular, you helped me *onto* this track, as a prospective MSc student with no background in Earth Science: a special thanks for initiating an unforgettable journey.

I owe an enormous debt to Yvonne Firing for showing me a whole other side to oceanography and helping me throw myself into it: guiding me through field work, helping to set up and supervise my project at Scripps, and supporting many other opportunities and experiences. I am similarly thankful to Teresa Chereskin at Scripps for hosting my stay and supervising my work enthusiastically, and to all the other staff there who foster such a wonderfully collaborative environment.

Most importantly, I have been buoyed up by an incredible circle of family and friends. Mum and Dad, you put the foundations of this in place from my earliest memories; Rhi, you've kept it propped up for almost as long. Au sud de la Manche, a number of PhD survivors blazed a trail for me: Heather, I couldn't have seen this through without you in my corner, and Delphine, I have no idea how you stuck this madness out without a Delphine showing you how. Clément, the world's greatest bike thief/ambulance driver, thanks for the tour of your country's emergency services. Charles, Marseille still awaits us: just one more leg to go... North of the Channel, legends Megan, Alice and Rosie can take credit for any remnants of my sanity.

Finally, I am simply unable to convey the inspiration and encouragement endowed by my high school mathematics teacher and earliest academic mentor, Mr. Daniel Mutlow, to whom this thesis is dedicated – thank you.

To Mr. Mutlow

Chapter 1

Introduction

1.1 Historical background

"The first vehicle in history", my father used to tell me, "was a dugout canoe" (B. Stephenson, pers. comm.). There can indeed be little doubt that mankind has possessed an intimate familiarity with the oceans for millennia, considering events in human history such as the traversal of the Pacific during the Austronesian expansion. Nevertheless, it was not until the late modern era that this familiarity became divorced from nautical roots, which had motivated it ever since the first seaworthy dugout. It was during this period that the contemporary field of physical oceanography was first established, with the Atlantic Ocean at its centre.

1.1.1 Piecing together the Atlantic circulation

1.1.1.1 The surface circulation

That the Atlantic was the primary region of study during this era was largely a coincidence of history, given its position as the gateway to European imperialism. Its potentially central role in climate, owing to its uniquely three-dimensional circulation, was not recognised until much later, and early work focused on charting its two-dimensional surface flow to expedite maritime endeavours.

These efforts began with the 1769 Franklin-Folger map of the Gulf Stream (Fig. 1.1b; Franklin, 1786) and continued with painstaking efforts over the following century to chart the surface circulation using techniques we would now describe as "Lagrangian" (or flow-following). These included messages in bottles (Becher, 1852) and reconstructions based on ship drift, recorded in archived captains' logs (Rennell, 1832; Maury, 1855; Richardson, 1985). The result was an admirably recognisable rendition

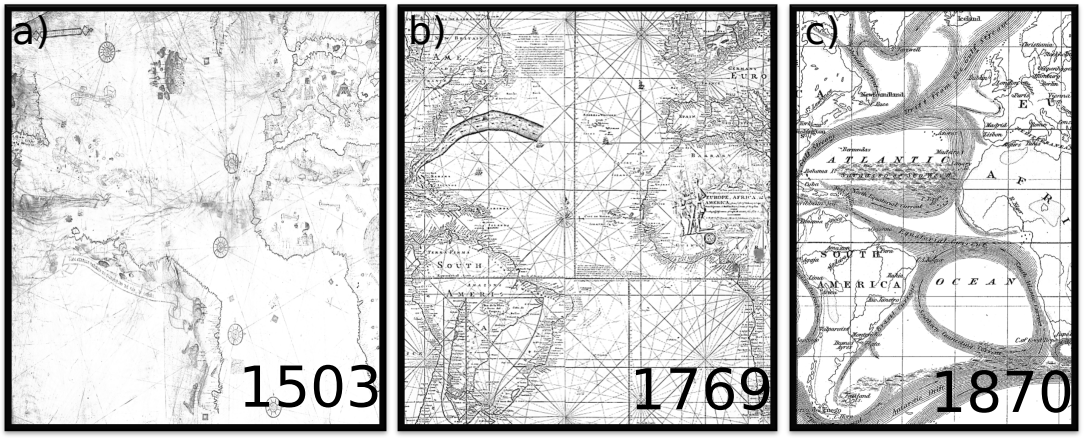


FIGURE 1.1: Historical representations of the Atlantic Ocean. a) Caveri's c. 1503 map (Van Duzer, 2020), the first to show the Atlantic basin as would be recognised today. b) The Franklin-Folger map of the Gulf Stream, the first map of a basin-scale ocean current (Franklin, 1786). In both cases, rhumb lines highlight their primarily nautical motivation. c) Johnson (1870) schematic map of ocean currents, showing a substantial advancement in knowledge over the preceding century. Crude measurement techniques meant actual transport estimates remained some way off.

(e.g. Fig. 1.1c) by the time of the launch of the Challenger expedition of 1872, an event widely described as the start point of modern oceanography (e.g. Macdougall, 2019). Such charts demonstrate knowledge of the two major gyres circulating between Africa and the Americas, the North Atlantic Current continuing the path of the Gulf Stream towards the Norwegian Sea, and a subpolar return flow tracking Iceland, Greenland and Canada.

The mathematical theory describing these flow patterns evolved over the next century. Indeed, it was the same phenomenon of ship drift that led to these sketches which also motivated Ekman (1905) to formulate a simplified model of the effect of wind on near-surface transport. This later evolved to explain the circulation of the gyres (Sverdrup, 1947) and the intense currents such as the Gulf Stream at their western boundaries (Stommel, 1948). Pivotally, Stommel went on to extend this formulation, proposing a return flow at depth that was also locked to the western boundary by similar physical laws (Stommel and Arons, 1959).

1.1.1.2 The three-dimensional circulation

By the time Stommel predicted the existence of the deep western boundary current, a connection between these surface and deep circulations had long been recognised as a phenomenon unique to the Atlantic: the sinking to depth of surface waters upon their reaching of subpolar latitudes.

The discovery of this phenomenon is commonly attributed to slave ship captain Henry Ellis (1752), who observed cold deep water in the tropics, and Benjamin Thompson,

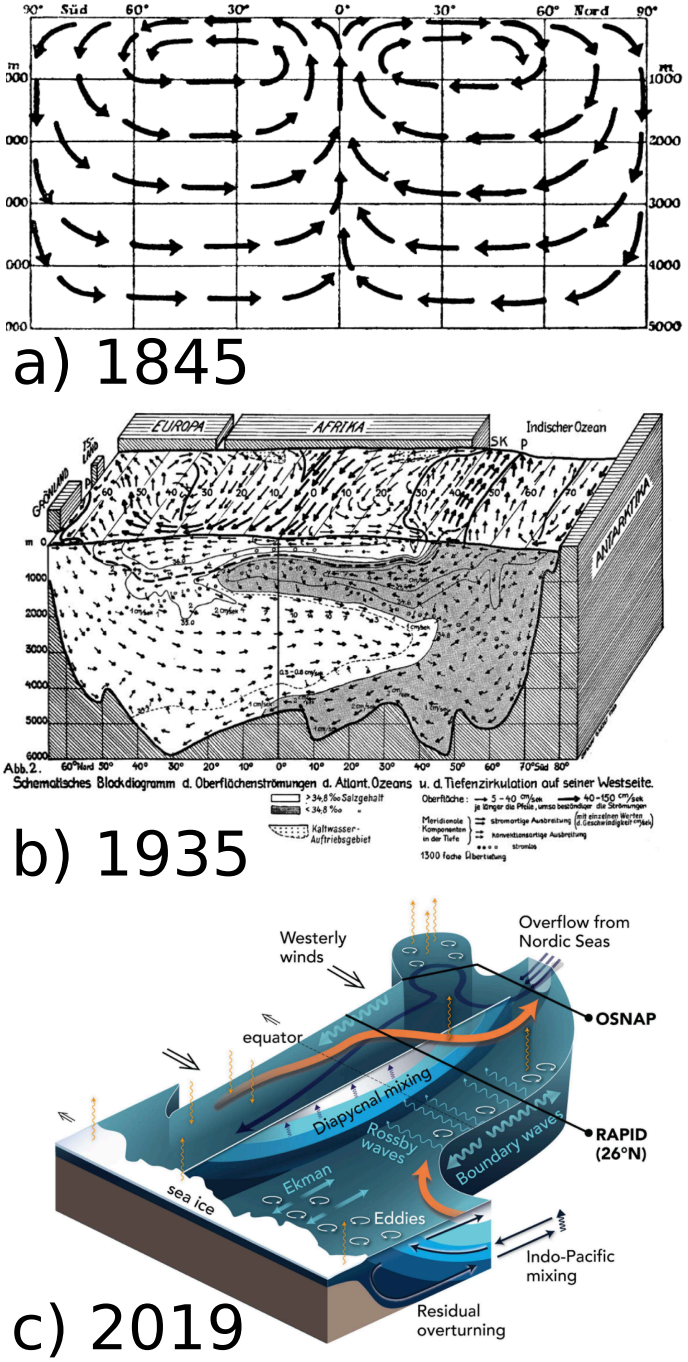


FIGURE 1.2: Advances in schematic representations of the Atlantic MOC. a) diagram by Lenz (1845) portraying two hypothetical overturning cells which upwell at the Equator. b) Wüst (1949) version of a 1935 schematic based on the Meteor expedition (Merz, 1925), showing the observed water mass and flow structure of the Atlantic MOC, as well as surface currents. c) A modern schematic from Johnson et al. (2019) showing the many processes now understood to impact MOC variability. Modern observing systems (black lines) have revealed a large amount of variability in meridional transport.

Count Rumford who "supposed" its polar origin during their subsequent correspondence in 1798 (Thomson Ct. Rumford, 1874). The idea took most of the next two centuries to reach maturity, initially manifesting itself in schematic diagrams (e.g. Fig. 1.2a) which, similarly to Franklin's Gulf Stream before, captured the spirit of the idea, but certainly not its details (Richardson, 2008, provides a fascinating review). It was not until the comprehensive survey of the Atlantic by the Meteor expedition (Merz, 1925) that a latitude-depth "map" of the circulation emerged (Fig. 1.2b; Wüst, 1935) that would be recognised today as the [Atlantic] Meridional Overturning Circulation ([A]MOC): the east-west average of the transport of all of the Atlantic's interconnected flow components. The residual northward transport in the upper ocean, which remains when this average is taken, came to be recognised in the coming decades as an unignorable feature of the climate.

1.1.2 The Atlantic and climate

The entire ocean is of first-order importance for climate, a fact surely recognised not long after the discovery of heat capacity by Joseph Black in 1760 (Black and Robison, 1803): the volumetric heat capacity of water is thousands of times as large as that of air, and so the fundamental role of the ocean in climate is often heuristically rationalised by the logic that the first few metres of the ocean contain more heat energy than the entire atmosphere. The Atlantic Ocean in particular, however, has an additional role owing to its unique circulation, a fact that was not recognised for some two centuries after Black's work.

The first serious consideration of the climatic importance of the AMOC can be attributed to Jung (1952), who estimated its role in the poleward transport of heat. That heat in the climate system must be transported poleward had been elucidated some 250 years earlier, when Halley (1693) had first quantitatively linked the temperature differential between the equator and the poles to the tilt of the Earth and the resultant imbalance of incident solar energy – an imbalance which necessitates a poleward heat transport to be equilibriated. Jung first estimated quantitatively that both the Atlantic and overlying atmosphere could make comparable contributions to this transport. Jung concluded that "it appears extremely important to determine whether or not such net meridional circulations do actually exist", and several further estimates of the meridional heat transport followed (e.g. Bryan, 1962; Vonder Haar and Oort, 1973; Hall and Bryden, 1982), reinstating the importance of the ocean.

Meanwhile, the evolving field of palaeoclimatology proposed a similarly important role for the ocean circulation in large climate shifts observed in ice core records. Within a decade of Jung's paper, a simple overturning model driven by latitudinal density differences was demonstrated by Stommel (1961) to have two possible steady regimes, depending on whether surface temperature or salinity changes had the greater impact

on density. Rooth (1982), who had previously worked with Stommel to expand this model (Stommel and Rooth, 1968), later suggested that the discharge of freshwater associated with past catastrophic ice melting events could have pushed the Atlantic into the salinity-driven regime, shutting down the AMOC. This idea was advanced by Broecker et al. (1985) who captured imaginations beyond the realms of palaeoclimatology with a highly simplified schematic of the ocean "thermohaline circulation" (with thermo- and -haline respectively referring to the roles of heat and salt, but lacking a rigorous definition beyond this[1]). The schematic showed the circulation as a "conveyor belt" (Broecker, 1987) which steadily acted to mediate the climate, but could shut down in response to large changes in atmospheric carbon dioxide (CO_2). In conjunction with the realisation that human activities were pushing atmospheric CO_2 to concentrations unseen in human history (e.g. Charney et al., 1979), a new energy was invested into researching the stability of the circulation.

1.1.3 The end of "a century of undersampling"

Investigations of the stability of the AMOC took place predominantly using maturing general circulation models (GCMs; e.g. Semtner, 1995) which, against the backdrop of rapidly increasing computational power, provided an as-before unparalleled insight into the large-scale ocean circulation. The importance of Southern Ocean winds for balancing deep water formation in the North Atlantic became apparent (e.g. Toggweiler and Samuels, 1995), as did the importance of representing mesoscale eddy fluxes for correctly modelling the locations of deep water formation (e.g. Danabasoglu et al., 1994). This occurred contemporaneously with the first continuous view of these eddies' ubiquity from space (Fu et al., 1994) and the development of satellite-tracked autonomous profiling floats (Davis et al., 1992), firmly signalling the end of a period dubbed by Munk and Wunsch (1982) as a "century of undersampling".

The suggestions of model-based MOC studies were not just that the MOC would decline in response to future anthropogenic scenarios (although not shut down; e.g. Cubasch et al., 2001) but that this decline could potentially be masked for decades by natural variations in the ocean circulation (e.g. Delworth et al., 1993). The breadth of this natural variability has since become apparent. As increasing volumes of data become available (Fig. 1.3) from ever-refining models (Fox-Kemper et al., 2019), a global fleet of autonomous platforms (Riser et al., 2016), evolving satellite technology (Tapley et al., 2004; Alsdorf et al., 2007), and basin-scale observational arrays

[1] A key issue with the notion of the thermohaline circulation was this lack of consistent definition (e.g. Wunsch, 2002) and its apparently implicit assumption that the global circulation was driven by sinking at high latitudes, compensated by micro-scale vertical mixing elsewhere. The Stommel (1961) model included mechanical stirring devices to take care of this, but observed mixing in the real ocean (e.g. Gregg, 1989) fell far short of that required (e.g. Munk, 1966) for balance. Additionally, the role of wind as a driver cannot be ignored. However, it is ultimately the meridional transport itself that is important to climate, not a particular driving mechanism, and so the more rigorously defined MOC is typically what is considered now.

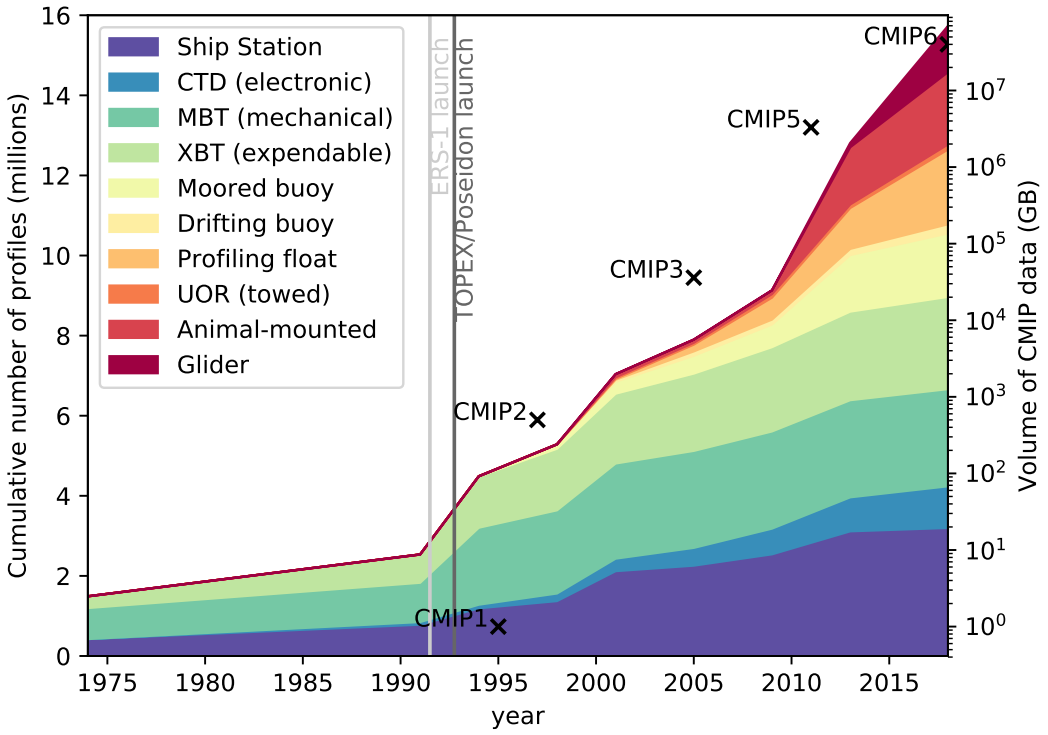


FIGURE 1.3: The data revolution of the last few decades. Area plot shows the number of ocean depth profiles (left axis) in the World Ocean Database at each release (Garcia et al., 2018), with colours illustrating the means by which the profile was obtained, showing changing observational technology. Crosses show the parallel evolution of climate modelling, based on the quantity of data (right axis; McGinnis, 2016) in each generation of the Coupled Model Intercomparison Project (CMIP). The light grey line marks the launch date of the ERS-1 satellite, since which near-continuous high-accuracy global SST measurements have been available. Similarly, the dark grey line marks the launch of TOPEX/Poseidon, the first satellite with a total altimetric error (10 cm) below the threshold for operational use. Subsequent launches have increased both resolution and accuracy.

(Frajka-Williams et al., 2019), variability beyond prior anticipation has revealed itself, throughout the ocean and at all timescales, ending once-and-for-all what Munk (2000) described as the "theology of a steady ocean". A pressing issue is therefore the determination of the origins, magnitude, timescales, and predictability of ocean variations, and their relevance to climate.

1.2 Ocean ventilation by the atmosphere

Climate, when viewed as the long-term average of weather, is by definition an atmospheric concern, and so the role of the ocean ultimately reduces to its interactions with the atmosphere. The two media exchange properties at the ocean surface, and atmospheric quantities are communicated into the ocean interior, ventilating it.

The ventilation of the ocean is arguably the most commonly recognised aspect of its role in climate, due to the ability of the ocean to sequester enormous quantities of anthropogenic tracers such as carbon and heat. The oceans have absorbed 93% of the excess heat energy associated with anthropogenic warming over the last half-century (IPCC, 2013) and about a third of anthropogenically emitted carbon since the pre-industrial era (Sabine et al., 2004; Gruber et al., 2019). However, the spatial distribution of this uptake is not uniform (e.g. Takahashi et al., 2002), as certain atmospheric conditions are preferentially selected, giving rise to ocean water masses, pockets of ventilated water with signature properties.

The well-mixed surface layer of the ocean, vigorously stirred by wind and heat loss (for example via shear instability and convective instability, respectively) allows atmospheric properties to be carried down to its base (the mixed layer depth; MLD). Following winter (when the MLD is greatest due to heat loss) these properties can then become shielded from re-exposure by the warming of the surface, adding an effective "lid" of buoyant waters, sitting atop like oil. Simple theoretical models (e.g. Luyten et al., 1983) suggest that, in the subtropics, the gyre recirculation also acts to shield water, forcing it to "subduct" beneath lower density water during its passage south in order to preserve certain fundamental properties. At high latitudes, there is little vertical structure separating mixed layer and deep waters, such that ventilated water is less obstructed from reaching the homogeneous deep ocean.

The North Atlantic makes a disproportionate contribution to the global ocean carbon sink due to the longevity of its water masses, and has taken up a quarter of the anthropogenic carbon in the ocean, despite representing only 15% of the ocean surface area (Sabine et al., 2004). In the subtropics, North Atlantic Subtropical Mode Water (NASMW; Worthington, 1959) forms on the southern flank of the Gulf Stream as an anomalously large, shielded volume of water with a characteristic temperature of around 18°C. NASMW ventilation acts as an important short-term upper ocean reservoir of carbon (Bates et al., 2002; Bates, 2012; Andersson et al., 2013) while, at higher latitudes, the subpolar waters of the North Atlantic account for the deepest penetration of anthropogenic carbon anywhere in the world ocean (Gruber et al., 2019). This reflects the formation of North Atlantic Deep water (NADW; Hanawa and Talley, 2001), the most dense water mass formed in the northern hemisphere, resulting from both intense heat loss and high salinity relative to other basins at this latitude. It is not just anthropogenic inputs which follow these ventilation pathways, however, but also natural variations of properties such as heat, generated within the atmosphere.

1.3 Atmospherically driven variability and ocean predictability

1.3.1 Predictability of the first and second kind

The future uptake of anthropogenic tracers (such as greenhouse gases) by ocean ventilation is largely dependent on future socioeconomic scenarios. For forecasts of the state of climate a century from now, the breadth of the range of possible scenarios is the dominant source of uncertainty. A further source of uncertainty in the nearer term is the effect of natural fluctuations such as weather (Hawkins and Sutton, 2009).

Our limited ability to forecast the weather is well known due to the extreme sensitivity of the future evolution of the atmosphere to small changes (Lorenz, 1963). The concept has entered the public consciousness through the analogy of the "butterfly effect" (Lorenz, 1972), which asserts that it is impossible to predict the exact state of the atmosphere by combining knowledge of its dynamics with information about its current state: the omission of a detail as seemingly trivial as the flap of a butterfly's wings would push the system onto such a different future path that even differing hurricanes would eventually form. This property places a fundamental upper limit on the predictability of the atmospheric state of a few weeks (Lorenz, 1969), after which the natural atmospheric fluctuations which penetrate the ocean are effectively random. This is the dominant source of uncertainty on short timescales (Hawkins and Sutton, 2009).

The respective problems of uncertainty arising from natural variability and scenario uncertainty were termed by Lorenz (1975) as prediction problems of the first and second kind, equivalent in mathematical terms to initial value and boundary value problems. For climate forecasts on so-called "near term" timescales, of years to decades, both problems can be significant (Fig. 1.4; Meehl et al., 2009, 2014), and understanding natural variability becomes crucial.

1.3.2 Predictability and variability in the North Atlantic

A large body of research has demonstrated enhanced predictability of the ocean state in the North Atlantic (Latif et al., 2006), in some cases up to decades, but the sources of this predictability are contentious owing to model disagreements and the immaturity of observations.

The simplest explanation is that upper ocean temperatures (the ultimate source of ocean influence on the atmosphere) remain predictable for longer in the North Atlantic because its uniquely deep mixed layers take longer to "fill" with random noise from the atmosphere, which slowly dilutes the signal of the known existing temperature

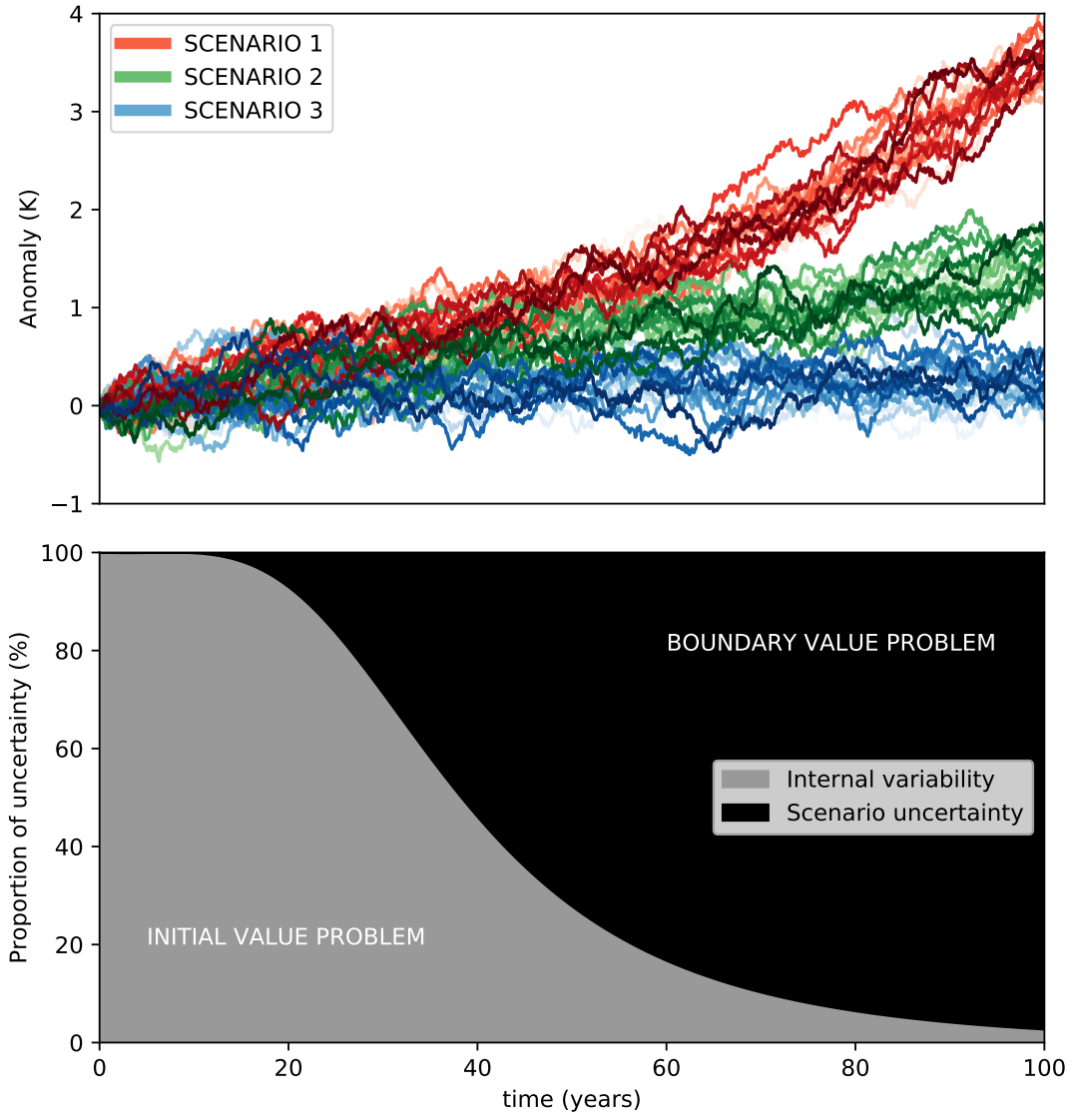


FIGURE 1.4: Schematic highlighting the nature of predictions of the first and second kind. Mock "ensembles" (top panel, with colour brightness denoting different ensemble members) following three mock "scenarios" (with colour hue denoting different scenarios). Early on, internal variability causes the members within each ensemble to diverge in response to initialisation error, while later on the range of scenarios cause the ensembles themselves to diverge. The "internal variability" is a prescribed stochastic function (an Ornstein–Uhlenbeck process, see Chapter 4) while the "scenarios" are prescribed (sigmoid) functions. This mathematically idealised set up allows exact calculation of the relative contribution of each to the uncertainty (lower panel). Function parameters are chosen to visually mimic values reminiscent of IPCC projections, and are purely illustrative.

(Buckley et al., 2019). This gives rise to "damped persistence" forecasting, involving no knowledge of the ocean circulation, and maintaining that the best guess is that future SST is current SST, plus some slowly accumulating error due to weather. As there is no long-term way to forecast the weather, the slowness of this error growth represents the fundamental predictability of SST.

An alternative explanation maintains that a perturbation from the atmosphere can trigger a deterministic, periodic response in the ocean circulation, which is, at least potentially, predictable (Boer, 2004). This would instigate a nonrandom influence on upper ocean temperatures from below, in addition to the random noise from above, potentially restoring some predictability. This is possible because, unlike carbon, which is a passive tracer, heat is an active tracer: warm or cold fluctuations can induce density anomalies, in turn triggering circulation changes through lateral pressure gradients or convective sinking, for example.

These two concepts can be associated with two methods for assessing them: the prognostic (or "classical") and diagnostic (or "potential") approaches (Boer, 2000). The former quantifies predictability through the rate of error growth (but cannot determine its source), while the latter assesses whether there is some source of periodic variability influencing a quantity (but cannot determine its predictability).

A number of prognostic studies have demonstrated predictability of the North Atlantic state reaching years or even decades into the future (Griffies and Bryan, 1997; Grötzner et al., 1999; Collins and Sinha, 2003; Pohlmann et al., 2004; Msadek et al., 2010), typically speculating that this predictability arises through deterministic interactions between surface conditions and the AMOC. Diagnostic studies have supported this conceptually (Boer, 2000; Boer and Lambert, 2008; Boer, 2011), highlighting the existence of nonrandom surface temperature variability in coupled models. Observations also suggest a repeating pattern in observed SST in the North Atlantic (the Atlantic Multidecadal Variability, AMV; Kerr, 2000). This phenomenon has been recognised for decades (Folland et al., 1984), but no observational study has yet been able to demonstrate a robust link to the AMOC. Recent studies have furthermore argued that the statistical patterns of AMV can be reproduced in a "slab" model without an interactive ocean circulation (e.g. Clement et al., 2015, 2016; Cane et al., 2017), precluding any role for the ocean circulation in amplifying the predictability of the North Atlantic (Srivastava and DelSole, 2017)[2].

[2]While particular modes of observed variability such as the AMV lie outside the scope of this thesis, a deeper consolidation of the current knowledge concerning AMV-AMOC interactions can be found in the review of Zhang et al. (2019).

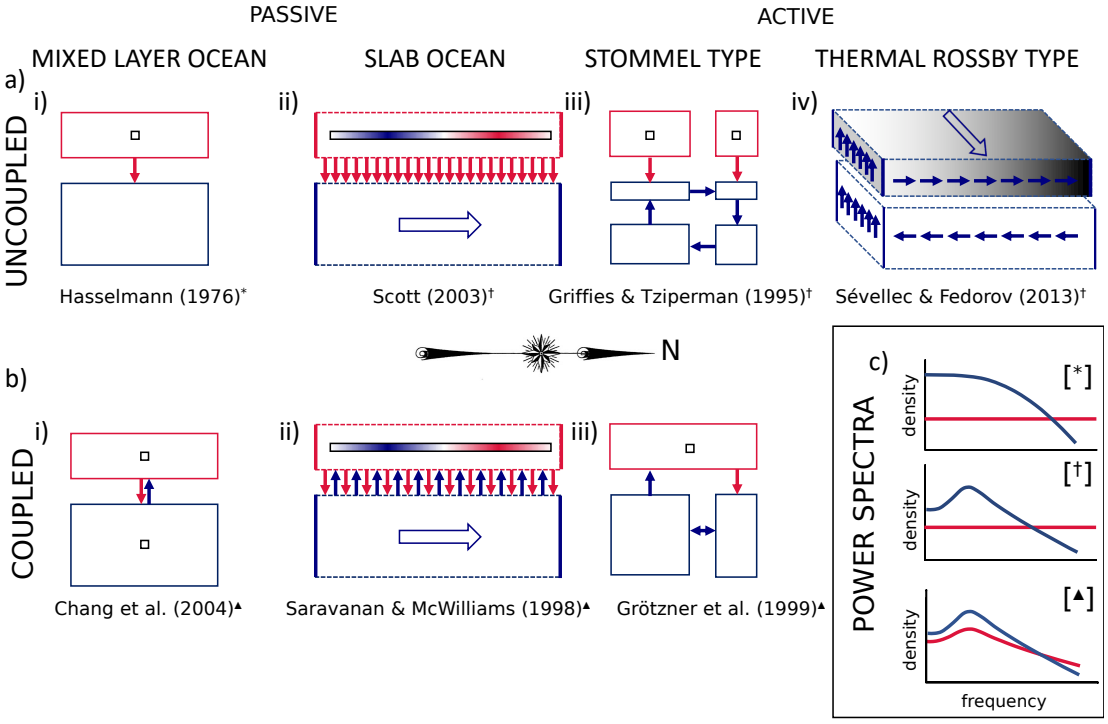


FIGURE 1.5: a,b: Schematics of the idealised/box models considered here. Red boxes show atmosphere, blue show ocean. Red arrows are fluxes originating in the atmosphere, blue arrows are fluxes originating in the ocean. White arrows with blue outlines show mean advection in the ocean. Dotted box edges show a continuous spatial dimension. Black-outlined rectangles show the presence of stochastic noise: for a(ii) and b(ii) the fill colour (red positive, blue negative) shows the spatial pattern of the coherent stochastic noise, which acts with the mean advection term to induce stochastic resonance. For a(iv), the gradient in the upper ocean box signifies the mean density gradient of the thermal Rossby mechanism. c: Schematic power spectral density distributions of the box models (adapted from Drijfhout et al., 2013). (*) shows the damped persistence case, without spectral peaks, (†) shows the case of stochastically forced ocean-only oscillations, (▲) shows the case of a coupled oscillator.

1.3.3 Idealised models of variability and predictability

The complexity of the coupled models used in the prognostic and diagnostic approaches, alongside the inability of the respective methods to disentangle the sources and predictability of variations, presents a barrier to understanding.

An alternative approach begins at the opposite end of the problem, exploring hypothesised mechanisms in minimal models which, in many cases, are analytically soluble, providing a complete mathematical insight into their dynamics. This has been referred to as the "bottom-up" approach to studying variability, complementary to the "top-down" approach of beginning with a realistic model (Frankcombe et al., 2009).

The zero-order "mixed layer" model (Fig. 1.5ai), as described above, is the simplest model of ocean surface temperature variability in response to atmospheric forcing, and was proposed by Hasselmann (1976). The only contribution of the ocean to predictability in this framework is its "memory" of the past, owing to its sluggish response to atmospheric heat fluxes. The corresponding ("red") power spectrum (showing the distribution of oceanic anomalies of different timescales; Fig. 1.5c, top panel), is in this case monotonic: decreasing smoothly towards higher frequencies/shorter timescales as a further reflection of its sluggishness relative to the atmosphere (for which anomalous events at all timescales are much more evenly distributed, leading to a flat, "white" spectrum). Conversely, if a repeating process contributed to SST, as is hypothesised to represent the Atlantic, the power spectrum would steepen and exhibit a peak: a concentration of events around a preferred timescale indicating the presence of a nonrandom dynamical process or ("oscillatory mode") being preferentially stimulated in the system (e.g. Fig. 1.5c, centre panel). Even within the simplified Hasselmann framework, a number of further simplifying assumptions are made: that atmospheric fluxes can be represented by white noise, in which variations ("shocks") instantaneously appear and decay is, of course, impossible, and based on the idea of a large differential between the timescales of evolution of the atmosphere and ocean. In reality, heat flux anomalies decay in days, while surface temperature anomalies decay over months. Regarding the latter, the Hasselmann model treats the response decay time as constant, and the noise as independent of the rest of the system. As highlighted by Sura and Sardeshmukh (2008), observed SST variations on very short timescales are everywhere better represented using a decay time which is itself a white noise, rationalised by its relationship on these timescales to gustiness. This transforms the model from an additive noise system (where totally independent noise accumulates) to a multiplicative noise system (where the stochastic component depends on the current state of the system). The implications and limitations of additive noise are considered further in various climactic contexts by Sardeshmukh and Sura (2009). Nevertheless, for the idealised models of variability considered here, and for the work presented in this thesis, the forcing of the ocean is taken to be sufficiently well-represented by additive noise, considering the timescales of years-to-decades which are of interest here.

Several studies have presented simple extensions to the Hasselmann model in which spectral peaks arise. These include, remarkably, highly simplified "slab" models which represent unchanging ocean transport along a single dimension. The mechanism responsible, "stochastic resonance", was presented in a coupled ocean–atmosphere context (Fig. 1.5bii) by Saravanan and McWilliams (1997, 1998). The ocean component was later isolated (Fig. 1.5aii) by Scott (2003) to examine its predictability via the prognostic error growth approach. Scott demonstrated that certain points in the (one-dimensional) domain were more predictable than others owing to "stochastic cancellation" of the spatially correlated, temporally random atmospheric noise, which

was prescribed as a symmetric sinusoidal pattern: positive atmospheric variations of certain frequencies could synchronise with the transport to be nullified by their negative-signed counterparts downstream. Wang and Chang (2004) reaffirm this, further demonstrating that spectral peaks arise in the model above a certain threshold where advection overpowers local damping (that is, when anomalies can be carried away from their origin faster than they can be "forgotten"). They highlight that spatial differences in predictability cancel out such that, overall, the predictability of the domain-averaged SST does not increase. They also show that passive advection contributes little to the spectral characteristics of a more realistic representation of the North Atlantic, concluding that another mechanism is likely at play in the prognostic GCM studies mentioned above.

Perhaps the simplest extension of the Hasselmann (1976) model to possess a variable ocean circulation is that of Cessi (1994). The model adds a random atmospheric component to the seminal two-box model of Stommel (1961). As described previously, the Stommel model can minimally represent the temperature and salinity properties of the Atlantic using a box representing the (cold, precipitative) high latitude subpolar region connected to a box representing the (hot, evaporative) low latitude tropics. A circulation is sustained by this contrast alone, via a transport term proportional to the density difference between high- and low-latitude waters. The Stommel model does not exhibit any spectral peaks relevant to the current discussion, but a slight extension of it, discerning between the upper and abyssal ocean (Tziperman et al., 1994), does.

Stochastic forcing was introduced to this four-box model (Fig. 1.5aiii) by Griffies and Tziperman (1995) and stimulates an oscillatory response. The mechanism arises due to the model's mixed boundary conditions (its asymmetric representation of ocean-atmosphere interaction concerning temperature and salinity at the surface). As the ocean and atmosphere can exchange heat, but not salt, modelled temperature anomalies at the surface gradually diminish in a restorative representation of this exchange. Conversely, salinity at the surface is controlled by freshwater evaporation and precipitation, mimicked by a constant flux term. The former occurs more rapidly than the latter, such that surface temperature anomalies are less robust over time than surface salinity anomalies.

As Griffies and Tziperman themselves highlight, meridional transport in reality is typically associated with east-west (rather than north-south) pressure gradients due to balance against the Earth's rotation (or geostrophy), which is not represented explicitly in the model. A second group of studies (e.g. the review of Buckley and Marshall, 2016) attributes low-frequency oscillatory variability not to high latitude deep water formation, but to rotation and east-west differences. A minimal model of this was presented by Sévellec and Fedorov (2013a), in a similar two-layer vein to the model of Griffies and Tziperman, but with rotation and continuous in horizontal space (Fig. 1.5aiv). Against the background of an existing meridional density gradient and

mean eastward transport (as in the North Atlantic Current system), oscillations can develop. The interplay between the background meridional gradient and zonal pressure gradients (set up by anomalies in response to this background gradient), forms the basis of an oscillatory mechanism known as the "thermal Rossby mode", for which evidence has been demonstrated in a large number of modern climate models (Muir and Fedorov, 2017).

We lastly consider the impact of coupling on the above systems. While fully coupled physics are beyond the scope of this thesis, they are briefly discussed here for three important reasons. Firstly, as previously mentioned, it is ultimately the detection of predictable oceanic variations in the atmosphere which determines the importance of the ocean for climate prediction. Secondly, ocean-only dynamical mechanisms for oscillatory variability such as that of Griffies and Tziperman must be treated with some caution where their existence stems from the choice of boundary conditions – effectively a parameterised substitute for true coupling. Lastly, coupling alone may offer an alternative means by which spectral peaks can form in very simple systems without ocean dynamics. The former two points are effectively addressed by Grötzner et al. (1999) in the second part of their paper.

The Grötzner et al. conceptual model (Fig. 1.5biii) consists of a Stommel-type two box ocean with a single atmospheric box, where the latter has no memory at all. Both atmosphere and ocean show a spectral peak at decadal timescales (e.g. Fig 1.5c, lower panel), but only the ocean exhibits any predictability unless the coupling strength is extremely high. Grötzner et al. use this conceptual model as a potential explanation for the lack of predictability of atmospheric variables in their coupled GCM ensemble despite the presence of oscillatory behaviour and oceanic predictability.

Chang et al. (2004) explore this further in a purely conceptual paper dedicated to the matter of predictability in simple coupled systems. Their model begins as two parametrically identical Hasselmann models which are coupled to each other (Fig. 1.5bi), causing each to exhibit a spectral peak around the frequency associated with the coupling strength, on the basis of coupling alone. They note that decreasing the coupling strength towards zero pushes this peak towards zero frequency such that when the two variables are uncoupled the monotonic power spectra which result still conceptually correspond to "spectral peaks", just at zero frequency. This is a fundamental result, as it implies that the existence alone of a deterministic oscillatory process such as those described in the above studies does not itself lead to an increase in predictability. This is echoed in the predictability review of DelSole (2017): "the location of a spectral peak need not have anything to do with predictability. Instead, predictability time scale often is related to the *width* of a spectral peak".

Chang et al. (2004) go on to demonstrate that the introduction of an asymmetry between the "memory" timescale of the two coupled components (such as would be

expected between the atmosphere and ocean) is the only means by which the predictability of the simple coupled system is greater than that of the uncoupled components, due to the associated transformation from a "normal" system to a "non-normal" system, as is explored further in the proceeding section.

1.3.4 Non-normal growth and optimal perturbations

The idealised models of the oceanic large scale that we have discussed so far have all either been linear (Hasselmann, 1976; Saravanan and McWilliams, 1998; Scott, 2003; Wang and Chang, 2004; Chang et al., 2004; Sévellec and Fedorov, 2013a) or treated as linear (on the assumption that atmospheric perturbations are small; Stommel, 1961; Cessi, 1994; Griffies and Tziperman, 1995; Grötzner et al., 1999), meaning that steady changes in input lead to steady changes in output. In particular, they have been linearly stable: any small perturbation (such as a random disturbance caused by the weather) may or may not trigger an oscillatory response in the ocean, but any response will decay such that the disturbance is eventually forgotten (consistently with the idea that the oceanic large scale circulation has behaved as a stable system for the most recent ten millennia; Dansgaard et al., 1993; Petit et al., 1999). For a linearly stable ocean, sustained atmospheric forcing is required for sustained oceanic variability. A central property of linear systems is that they possess certain fundamental perturbation patterns (eigenvectors) which will not change shape when applied to the system, only magnitude – they ultimately simply either grow to infinity (if the system is linearly unstable) or shrink to zero (if it is linearly stable).

Any possible state of the system can be represented as a combination of its eigenvector patterns, and so while it typical and intuitive to represent the system as a geographical map of variables such as temperature and salinity, it is more mathematically useful to consider the system not in directions such as north and east, but the "directions" of eigenvector 1, eigenvector 2, etc. If these directions are not perpendicular, they can project onto each other, in the same way that if north and east were not perpendicular (or "normal"), travelling east would appear to result in also changing one's latitude. This dependence of eigenvectors is a central property of non-normal systems, and results in the curious behaviour that, even if every conceivable perturbation ultimately decays (that is, the system is stable), certain perturbations can temporarily grow in their overall size before decaying, purely due to this interference (e.g. Bale and Govindarajan, 2010).

This oceanic transient growth has substantial implications for predictability: classical predictability studies typically assume that the ocean is both stable and perfectly initialised. The latter is, of course, unattainable, and oceanic transient growth suggests initial errors due to missing information in certain regions of the ocean will

temporarily amplify, even in a stable ocean. This motivates targeted ocean observation and model initialisation to minimise error in these regions of fastest growth.

The patterns of fastest growth in a non-normal system are its "singular vectors", close cousins of its eigenvectors (a thorough mathematical treatment is provided by Palmer et al., 1994). Singular vectors have been explored by a number of studies of Atlantic predictability, but the framework presents some challenging obstacles. The process of calculating singular vectors (singular value decomposition, SVD) is very computationally demanding, with the classical problem having cubic complexity (Golub and Van Loan, 1996; Pan and Chen, 1999). That is to say that, even in the simple case of moving from the Stommel box model to the Tziperman et al. (1994) box model, doubling the number of boxes would require $2^3 = 8$ times the compute time for SVD. Realistic ocean models are many millions of times more complex than this. Additionally, the analysis requires a linear representation of the ocean. Any nonlinear conceptual models we have seen so far are almost trivially linearised given their mathematical simplicity, but in the case of a realistic circulation model, this is a notoriously difficult task (Giering and Kaminski, 1998). Additionally, the problem necessarily requires constraint: one needs to define what is meant by the "size" of the perturbation and its impact, not least because the variables of the system have distinct units. How this is defined (the choice of "norm" for the problem) will, fundamentally, affect the solution found.

Early SVD studies overcame the issue of computational cost by using highly idealised representations of the Atlantic, such as box models (Lohmann and Schneider, 1999; Tziperman and Ioannou, 2002) or zonally averaged models (Zanna and Tziperman, 2005, 2008; Alexander and Monahan, 2009). More recent studies have applied the techniques to GCMs, either in simplified configurations (Zanna et al., 2011, 2012) or by constructing simplified linear approximations to the models using the technique of linear inverse modelling (LIM; Tziperman et al., 2008; Hawkins and Sutton, 2009). The LIM approach "compresses" the model (potentially losing a degree of non-normality, Farrell and Ioannou, 2001) into a combination of patterns which capture the most important aspects of its dynamics (Penland and Magorian, 1993). However, this range of formulations gives a wide range of results. The transient amplification timescale for optimal MOC perturbations in the zonally averaged model of Alexander and Monahan (2009), for example, ranged from years to decades depending on model diffusion and boundary conditions, while error amplification times between 7 and 50 years in the LIM study of Tziperman et al. (2008) depended strongly on choices made in model reconstruction. The SVD approach is also highly sensitive to the choice of norm: using a quadratic norm (sum of squared transport anomalies), Zanna and Tziperman (2008) find optimal surface excitation of the MOC to occur after a century, compared with a decade when considering an infinity norm (transport maximum). The common use of such quadratic MOC norms presents a further issue.

Optimal perturbations to temperature and salinity which are exactly compensated in density will have "zero" size when measured in terms of MOC impact, leading to the peculiar outcome that, if density changes following their evolution, the MOC undergoes "infinite" amplification. This singularity can only be avoided by adjusting the outcome such that the initial density anomaly is non-zero.

An alternative approach, linear optimal perturbations (LOPs), was presented by Sévellec et al. (2007), and precludes the need for a norm or costly SVD computation entirely. This is achieved by seeking the optimal impact on a linear function of the ocean state, rather than on a nonlinear norm. Sévellec and Fedorov (2015) used this technique to demonstrate optimal MOC amplification after 9 years triggered by SST and SSS anomalies in the western subpolar North Atlantic, centred over the deepest mixed layers of the model. Sévellec and Fedorov (2013b) similarly found that initialisation errors in the deep ocean are particularly prone to amplification – a potentially significant issue for predictability, given the present sparsity of deep ocean observations. With a similar motivation, Sévellec et al. (2018) applied the LOP method to the hypothetical design of an optimal monitoring system, targeted to reduce error growth in those regions where it is most problematic for a selection of metrics of the North Atlantic. They evaluate the theoretical model response to both atmospheric noise and initial condition uncertainty and seek to reduce the latter to prolong predictive power. They suggest that a fleet of ten gliders would be able to reduce the uncertainty in AMOC predictions by half for up to about 6 years.

1.4 Nonlinearities and oceanically driven variability

In all of our previous examples, we have considered the ocean to be a linearly stable system: any disturbance from its current state will die away over time. In reality, ocean flow on many scales is unstable, nonlinear, and highly turbulent. The relevance and extent of its nonlinearity on large spatial scales and climatic timescales remains in question. If instabilities and nonlinearities predominantly lead to small-scale turbulence, but the large-scale circulation is still linearly stable, then it can be seen to be forced by these small-scale chaotic variations in a similar way to its forcing by atmospheric weather. If, however, the large scale is not linearly stable after all, then small disturbances to it (which would grow to infinity in a linear approximation) will grow until reaching a saturation point due to nonlinear feedbacks.

1.4.1 Idealised models

The latter concept has been explored at length in the context of decadal variability, as an alternative explanation for oscillatory variability. In contrast to the position that

these oscillations are damped and sustained by atmospheric stimulation, many authors have proposed that such oscillations can be self-sustained, and it is the ocean driving them, rather than the atmosphere (e.g. Weaver and Sarachik, 1991; Kushnir, 1994; Greatbatch and Zhang, 1995; Huck and Vallis, 2001). In this perspective, the Atlantic exhibits linearly unstable oscillations which grow until saturating at a certain amplitude at which they are sustained without external input from the atmosphere (e.g. Estella-Perez, 2019).

The instability responsible for feeding the oscillations (e.g. Colin de Verdière and Huck, 1999) has been proposed to be baroclinic instability, the mechanism responsible for ocean mesoscale eddies and atmospheric weather patterns, which liberates potential energy from sloped density surfaces. These formations are typically associated with the smallest spatial scale at which planetary rotational effects become important (the "deformation radius"), and can, in the simplest case, appear in ocean models with fine horizontal resolution, but which vertically discern only between the upper and deep ocean, as a result of interactions between the unified ("barotropic") and separate ("baroclinic") motions of these two distinct layers. As highlighted by Colin de Verdière (1986), similar instability at larger scales than the deformation radius is not supported in this classical model, but is supported when the ocean vertical dimension is divided into larger numbers of layers, such that further distinct combinations of interactions (baroclinic "modes") are possible (Liu, 1999). This has been termed "generalised" baroclinic instability (Te Raa and Dijkstra, 2002), and can act over distances at which large-scale variability can be directly impacted. Extending the two-layer damped oscillatory model of Sévellec and Fedorov (2013a) to three layers so as to support this instability, Sévellec and Huck (2015) investigate this mechanism and demonstrate its theoretical validity. Nevertheless, multiple studies (Colin de Verdière and Huck, 1999; Huck et al., 1999; Huck and Vallis, 2001; Te Raa and Dijkstra, 2002) find that only in very non-dissipative model configurations is enough energy available to sustain the oscillations. Typically, parameterised eddy diffusivity must remain small and, if mixed boundary conditions are used, the surface damping of temperature anomalies must be extremely gradual relative to parameters typically used in OGCMs. Winton (1997) also argues that realistic ocean geometry would act to dampen the oscillations. More thorough investigations of the impacts of topography (e.g. Te Raa et al., 2004; Arzel et al., 2018) and turbulence (e.g. Huck et al., 2015) have followed, with it being apparent that oscillations can survive both. However, it is unclear whether the presence of resolved eddy turbulence acts to stimulate what is in fact a large-scale stable mode (decompositions addressing this are currently restricted to idealised models; e.g. Hochet et al., 2020). Furthermore, strong sensitivity of stability to surface boundary conditions suggests that the roles of ocean and atmosphere cannot be so neatly separated.

1.4.2 Realistic models

Recent advances in computational power have stimulated a renewed effort to diagnose variability generated within the ocean. In this framework, realistic ocean models are run with some prescribed difference and compared. These differences could be between resolution (e.g. Hirschi et al., 2013; Grégorio et al., 2015), forcing (e.g. Sérazin et al., 2015, 2016) or initialisation (a prognostic ensemble approach, such that oceanic instabilities cause members to diverge; e.g. Sérazin et al., 2017; Leroux et al., 2018; Jamet et al., 2019). Sérazin et al. (2017) suggest that the majority (91%) of the world's upper ocean heat content variability is predominantly forced by the atmosphere, with exceptions only in highly turbulent regions of the upper ocean. Conversely, abyssal heat content variability in the Atlantic is shown to be driven primarily by the ocean. Leroux et al. (2018) and Jamet et al. (2019) determine the spatial patterns of intrinsic AMOC variability, finding large-scale coherence for a large portion of the variability. Leroux et al. (2018) show that the dominant such pattern is focused on the South Atlantic, while Jamet et al. (2019) (who use a higher resolution regional model but do not resolve the South Atlantic) find the dominant (local) intrinsic pattern to be largely similar to the dominant atmospherically forced pattern, and centred on the North Atlantic. In an attempt to isolate the relevance of the large and small scales, Sérazin et al. (2015) apply filtering techniques to an eddy resolving simulation and show that large-scale low-frequency sea level variability is predominantly forced by the atmosphere (Fig. 1.6). Grégorio et al. (2015) compare low- and high-resolution simulations to show almost no large-scale intrinsic variability in a low-resolution forced model (consistently with the idea that such large-scale variability would be damped by parameterisations; e.g. Colin de Verdière and Huck, 1999).

1.4.3 Nonlinear optimal perturbations

Even if the large scale circulation may be treated as linearly stable, we have seen that it is almost certainly non-normal, and so small perturbations may grow large enough for nonlinear effects to become important. This motivates the study of optimal perturbations in a nonlinear system, but such perturbations are not easily found. As in the linear case, two approaches have been presented, with one (Conditional Nonlinear Optimal Perturbations; CNOPs) acting as a natural extension of the SVD approach and the other (Nonlinear Optimal Perturbations; NOPs) acting as a natural extension of the LOP approach.

In a demonstration of the importance of nonlinearities in even very simple models, Mu et al. (2004) and Sun et al. (2005) explore CNOPs of the Stommel (1961) box model. They reiterate using a SVD approach that the optimally growing perturbation in a linear approximation to the model is a high latitude salinity perturbation (Lohmann

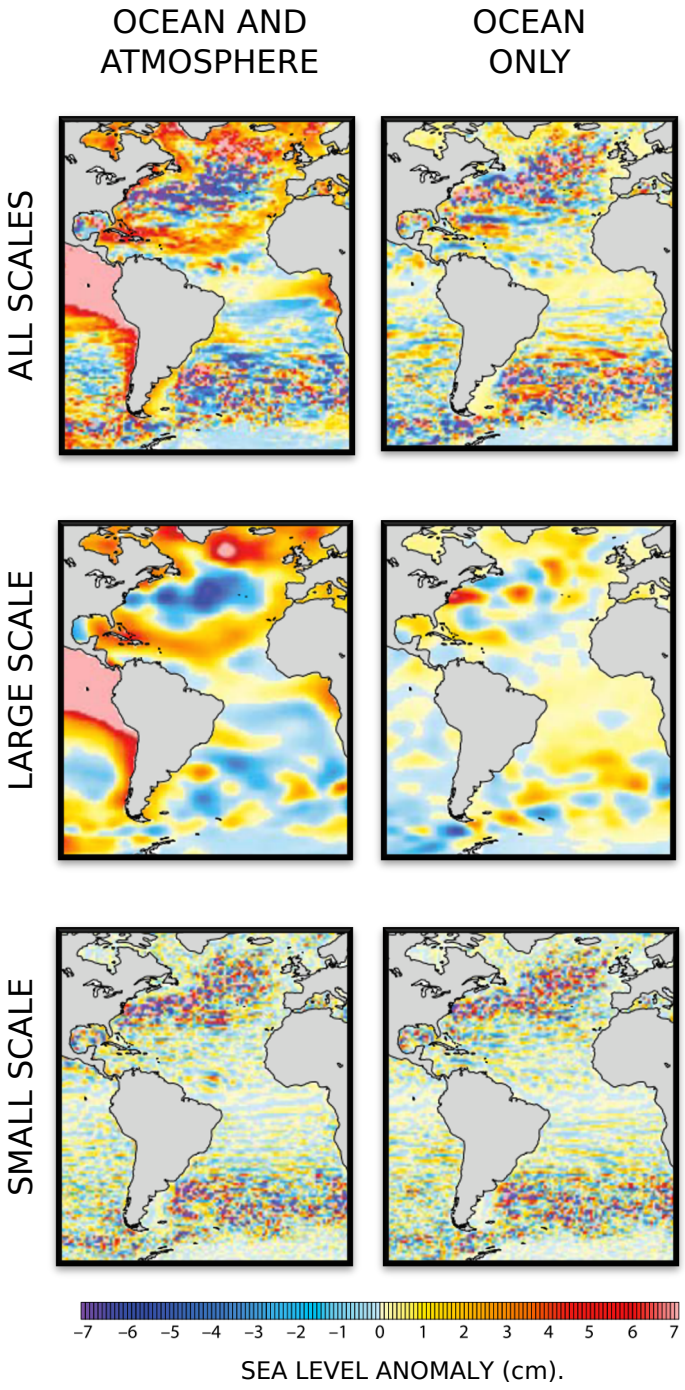


FIGURE 1.6: Origins and spatial scales of chaotic variability in sea level, adapted from Sérazin et al. (2015). The left column shows an anomaly snapshot from an eddy-permitting simulation with historical forcing (including chaotic atmospheric variability), the right column shows output from the same configuration but with climatological forcing (such that any chaotic variability is assumed to derive purely from nonlinearities in the ocean). The output is spatially filtered to remove scales smaller than 12° in the centre row and larger than 6° in the bottom row, highlighting the differences in scale between chaotic variability in the ocean and atmosphere.

and Schneider, 1999; Zanna and Tziperman, 2005). Due to linearity, the response scales with the perturbation magnitude – twice the anomalous salinity results in twice the impact, while a positive and negative perturbation of the same magnitude result in equally anomalous positive and negative responses, respectively. They demonstrate using CNOPs that this symmetry does not exist in the nonlinear regime, and that optimal freshwater perturbations grow more effectively than salt perturbations of the same magnitude. This was reinforced in an idealised three-dimensional model by Zu et al. (2013, 2016). In seeking an optimal SSS perturbation of the MOC, Zu et al. (2016) show (in a "partly linearised" model) that different nonlinear processes control the response to the CNOP for the first three decades, before the perturbation decays in a linear oscillatory manner. In particular, convective feedbacks act to strengthen the MOC for the first two years (compared with direct weakening with a linearised representation of convection), causing a large (around 1 Sv) initial difference between the nonlinear and linear responses. Omission of the nonlinear transport of buoyancy anomalies by (zonal) advection anomalies causes a similarly sized discrepancy about a decade later.

Zu et al. (2016) highlight that, unlike the linear SVD approach, the CNOP cannot be seen as a calculation of the fastest growing error pattern. This is because the pattern itself is dependent on the amplitude of the perturbation: while the optimal linear pattern does not change for larger amplitudes, this will not necessarily induce an optimised response in a nonlinear system, as we have seen. This amplitude dependence suggests that the optimal perturbation of a nonlinear system is better thought of as the optimal precursor to a particular event than the optimally growing error. If we adopt this perspective, it is perhaps arbitrary to constrain the size of the perturbation, rather than the response. The NOP method (Müller and Sévellec, 2020) takes this upended approach, seeking the smallest perturbation which leads to a particular response, as opposed to the perturbation in a particular size range which leads to the largest response.

The method begins by perturbing the nonlinear model with the LOP, thereby creating a new state which is closer to the target, but does not reach it, due to nonlinearities such as those highlighted by Zu et al. (2016). From this new state, a new LOP is calculated, and the process is repeated iteratively until it is considered that the target state has been effectively reached. The approach was first proposed as an extension to the LOP method in Sévellec and Fedorov (2013b), in the context of reducing discrepancies between modelled and observed values, in a manner reminiscent of data assimilation. Assessing the method in this context, Müller and Sévellec (2020) attempt to reproduce a synthetic dataset by applying an optimal perturbation. They demonstrate that in a coarse resolution model about a quarter of this target is achieved using the LOP, with $> 97\%$ agreement after 250 iterations, but that this skill is severely impacted at higher resolution. With eddies permitted in a narrow band

around the equator, the 250$^{\text{th}}$ iteration reaches only 56% of the target, while in a globally eddy permitting model (even with much shorter evaluation times), this falls to just 28%, again raising the importance of small-scale interactions with larger scales.

In a follow-up study, Müller et al. (2020) evaluate optimal precursors to a range of anomalous AMOC fluctuations so as to determine the detectability of so-called "events" (uncharacteristic departures of the AMOC from its mean, such as have been observed; e.g. Bryden et al., 2014) given measurement errors in the global ocean observing system. With this more simplified target, the LOP can reduce over 92% of this transport averaged over a year, even with higher effective resolution in the tropics. This increases to 99.999% after 4 iterations, essentially "shutting down" the transport, with the signature of the perturbation remaining for decades.

1.5 Summary

Over history, we have seen several paradigm shifts in our view of the ocean circulation. Seafaring civilisations have long attempted to chart its features for nautical purposes, but it was not until the late modern era that rigorous explanations for circulation phenomena were sought. This led to a revolution in understanding, culminating in the picture of the global ocean circulation as a connected system of processes acting steadily in concert to mediate the climate. Over the last few decades, another revolution – in both computing power and observational coverage – has revealed the true complexity of the system, leaving many open questions about the role of the ocean in climate at a range of spatiotemporal scales.

The simplest picture of the interaction of the ocean with climate is that it passively internalises the signatures of atmospheric anomalies along consistent ventilation pathways, those followed by its water masses. The implication of this perspective is that anthropogenic heat and carbon will be taken up along the same pathways, as will unpredictable atmospheric weather fluctuations. The latter represent a sink of oceanic predictability on short time scales, with the former becoming the dominant source of uncertainty in the longer term.

Dedicated model studies exploring the uptake of atmospheric noise have demonstrated enhanced decadal predictability in the North Atlantic, but the complexity of models used for such studies makes the origins of this predictability difficult to disentangle. We therefore considered a number of proposed explanations in simplified idealised models of ocean variability. These show that pronounced spectral peaks, corresponding to deterministic mechanisms with a preferred timescale, can occur in very simple representations of the Atlantic, even when the ocean and atmosphere are decoupled and the ocean transports anomalies passively – that is, without the transport changing in response to them. However, we have seen that it is not the presence of these

spectral peaks which leads to enhanced predictability, but non-normality stemming from asymmetries in the system such as its differing response time to temperature and salinity.

Having established the importance of non-normality, we saw how it could be leveraged to impact predictability through transient growth. By targeted model initialisation using observations, it is theoretically possible to align variations in the model and reality for a limited time, prolonging predictive skill. Conversely, errors can also undergo transient amplification, rapidly pushing the modelled ocean away from reality. This motivates the study of optimal perturbations – the patterns which most efficiently grow. We have considered two common approaches to this in linear systems, SVD and LOP, with the former being a useful but computationally expensive tool in the context of forecast error growth and the latter being a computationally efficient way to optimise the response of specific (linear) metrics of the ocean state.

We finally considered an alternative perspective on oscillatory behaviour and perturbation growth – that of linear instability and nonlinear feedbacks. A range of idealised studies have presented the idea that oscillatory variability in the ocean is not the stable response to the atmosphere, but originates from large scale instabilities in the ocean circulation itself. The question of the ocean's role as a driving force has seen a renaissance in light of increasing computational resources, with a raft of realistic model studies demonstrating that a substantial portion of ocean variability is intrinsic.

These recent studies seem to suggest that the large scale circulation is stable in realistic ocean models, but is fed by turbulence at much smaller scales (in addition to the atmosphere), although debate continues to some degree. If the large-scale circulation is in fact stable, nonlinearities at all scales may still have a role to play as the transient growth of perturbations could push them outside of the realms of linearity. We thus considered ways to account for this, comparing optimal perturbation frameworks in nonlinear systems, similarly to the linear case.

1.6 Thesis outline

This thesis presents a selection of case studies concerning the above phenomena in regimes of increasing complexity. The structure is as follows. In Chapter 2, we present a linearised ocean model which has been modified to allow passive tracer propagation, and apply this to tracking the surface origins and internal destinations of North Atlantic water masses following ventilation. In Chapter 3, we compare the passive and active configurations, considering the variance of their response in heat content to a stochastic representation of external atmospheric forcing. Chapter 4 extends the stochastic representation to include mesoscale eddy turbulence, and looks at the relative impact of external atmospheric and internal oceanic forcing on uncertainty in

metrics of the North Atlantic state. The realistic stochastic representation is compared with an idealised stochastic noise derived in an optimisation framework similar to LOP. Motivated by the finding that the ocean internally stimulates a large portion of variability, Chapter 5 examines (using the NOP method) the oceanic precursors to an AMOC event recently recorded by *in situ* monitoring. Conclusions and recommendations for future work are presented in Chapter 6.

Chapter 2

Tracking water masses using passive-tracer transport in NEMO v3.4 with NEMOTAM:

application to North Atlantic Deep Water and North Atlantic Subtropical Mode Water

A version of this chapter has been published in Geoscientific Model Development under the same title.

Abstract

Water mass ventilation provides an important link between the atmosphere and the global ocean circulation. In this study, we present a newly developed, probabilistic tool for offline water mass tracking. In particular, NEMOTAM, the tangent-linear and adjoint counterpart to the NEMO ocean general circulation model, is modified to allow passive-tracer transport. By terminating dynamic feedbacks in NEMOTAM, tagged water can be tracked forward and backwards in time as a passive dye, producing a probability distribution of pathways and origins, respectively. To represent surface (re-)ventilation, we optionally decrease the tracer concentration in the surface layer, and track this concentration removal to produce a ventilation record.

Two test cases are detailed, examining the creation and fate of North Atlantic Subtropical Mode Water (NASMW) and North Atlantic Deep Water (NADW) in a 2° configuration of NEMO run with repeated annual forcing for up to 400 years. Model NASMW is shown to have an expected age of 4.5 years, and is predominantly eradicated by internal processes. A bed of more persistent NASMW is detected below the mixed layer with an expected age of 8.7 years. It is shown that while model NADW has two distinct outcrops (in the Arctic and North Atlantic), its formation primarily takes place in the subpolar Labrador and Irminger seas. Its expected age is 112 years.

Plain language summary

Different water types are created at the sea surface with a signature based on the local conditions of the atmosphere. They then take these conditions with them into the deeper ocean, and so their journey is an important climate process to understand. In this study, we modify and repurpose a specialised model which simulates the ocean forward and backward in time to determine where new ocean water goes, where at the surface existing water comes from, and how old it is, by tracking it as a dye.

2.1 Introduction

The intricate process by which the atmosphere and ocean exchange properties has decisive effects on oceanic and atmospheric circulation, biogeochemistry, and climate. Tracing the pathways of newly ventilated ocean water masses is a multidisciplinary pursuit, for which an entire toolbox of methods has been developed.

Thermocline ventilation and water mass formation is an area with a long history of inquiry. From charts of hydrographic sections, Iselin (1939) inferred communication between isopycnals at depth and their outcrops. This idea was developed theoretically over the following decades in analytical models (Welander, 1959, 1971; Luyten et al., 1983). In this vertically-discretised, laminar framework, potential vorticity (PV) conservation facilitates a bijective relationship between the origin and destination of a water parcel. In reality, however, turbulent mixing means that surface-to-deep pathways have a pronounced probabilistic character.

Observational studies have been able to capture snapshots of this turbulent behaviour locally in passive-tracer dye-release experiments (e.g. Schuert, 1970). However, the effect of turbulence on water mass pathways has only been observed on a large scale since the widespread adoption of cost-effective Lagrangian profiling floats, which have been observed to follow chaotic trajectories (e.g. Fischer and Schott, 2002; Fratantoni et al., 2013; Bower et al., 2019). Despite their number (typically several tens of floats for dedicated pathway tracking studies), these trajectories remain under-sampled.

Simulations of larger numbers of Lagrangian particles in sophisticated eddy-resolving ocean general circulation models (OGCMs, e.g. Blanke and Raynaud, 1997; Gary et al., 2014) are reworking long-held assumptions about the routes taken by newly formed water masses. Despite these new developments, such studies are limited by the large computational expense of eddy-resolving models, and the fact that sub-grid-scale dispersion of Lagrangian particles is typically not parameterised in laminar models. An alternative approach is to use a simulated passive tracer (c.f. England and Maier-Reimer, 2001). Such tracers are spatial distributions of dye concentrations

which undergo diffusion. Lagrangian particles, conversely, are indivisible nodules which may only be advected by the immediate flow. The Eulerian perspective of the passive tracer method makes it inherently probabilistic - the continuous field (tracer concentration) propagated by the model corresponds to a continuous probability distribution. Lagrangian particle trajectories can be seen as discrete samples from this distribution, and so a very large (theoretically infinite) number of them is required to adequately reconstruct it.

This study presents a method for tracking water masses by means of passive-tracer deployment in the tangent-linear and adjoint model (TAM) developed for the NEMO OGCM (Madec, 2012; Vidard et al., 2015). TAMs are typically used for sensitivity studies and data assimilation (Errico, 1997). However, they offer two distinct advantages over online passive-tracer tracking in a nonlinear OGCM. Firstly, as with many Lagrangian methods, the TAM propagates tracer fields offline, meaning only one (more computationally demanding) simulation in the nonlinear model is required to conduct many experiments under similar conditions. Secondly, the TAM propagates probabilities both backwards and forwards in time (detailed further in Section 2.2.1).

The adjoint and tangent-linear of the model can respectively track the origins and fate of passive-tracer-tagged water in this manner.

TAM construction is typically a laborious process (e.g. Giering and Kaminski, 1998). The Jacobian of a highly involved nonlinear model (often with millions of degrees of freedom) must be computed with respect to the ocean state. This provides a linear function mapping small perturbations to the ocean state to their future outcome (the tangent-linear). The adjoint of this linear operator provides the sensitivity of the ocean state to earlier perturbations.

Several studies have bypassed such complications by computing approximations to the true adjoint of tracer transport in an OGCM. An early application of an adjoint approximation to the tracking of water masses was developed by Fukumori et al. (2004), who tracked eastern equatorial Pacific water with the Massachusetts Institute of Technology GCM. Their approximation capitalises on the comparatively simple nature of passive-tracer transport in isolation. For passive tracers, advection and diffusion are respectively skew-symmetric and self-adjoint. They were thus able to map possible past locations of the water mass by reversing the velocity field without changing the diffusion tensor, similarly to a true passive tracer adjoint. The method is further used by Qu et al. (2009), Gao et al. (2011), and Gao et al. (2012) to track Pacific waters and Qu et al. (2013) to track the salinity maximum of the North Atlantic subtropical gyre.

Another pseudo-adjoint passive tracer approach is presented by Khatiwala et al. (2005). Rather than deriving the Jacobian of an OGCM, they derive a tangent-linear approximation empirically. The procedure constructs a “transport matrix” row-by-row

by repeatedly perturbing the nonlinear model along basis vectors of the ocean state. The responses following a single time step can be synthesised to derive the elements of the matrix. While more approximate than the Fukumori et al. (2004) approach, it is also more computationally efficient by design. This allows the transport matrix method to be used for long-term passive tracer experiments, for example simulations of long half-life radioisotopes.

The work presented herein is of a similar nature to the above studies, but repurposes an existing TAM for passive tracer tracking. In this sense its adjoint is the true adjoint of the nonlinear model, rather than a bespoke approximation.

We demonstrate the efficacy of the development through case studies of two climatically important water masses of the North Atlantic, whose formation regions are closely aligned with major components of the Atlantic Meridional Overturning Circulation (AMOC). These are North Atlantic Subtropical Mode Water (NASMW) or Eighteen Degree Water, formed in the vicinity of the Gulf Stream, and North Atlantic Deep Water (NADW), formed in the subpolar North Atlantic.

NASMW is the name given to a homogeneous body of water found in the upper region of the western subtropical gyre (STG). It was first identified as early as the Challenger expedition, where repeat soundings revealed its unusual uniformity (Thomson, 1877, their plate XI). It was later named Eighteen Degree Water for its characteristic temperature by Worthington (1959), who provided a first formal definition of the water mass.

NASMW formation has historically been attributed to surface heat loss during winter (e.g. Worthington, 1972; Speer and Tziperman, 1992) although mechanical forcing has been proposed to have an effect on formation by destroying PV (e.g. Thomas, 2005). Formation rate estimates by varying methods have produced substantially different results (addressed by Marshall et al., 2009), and it has recently been suggested that the nature of formation is location-dependent (Joyce, 2011).

Newly ventilated NASMW follows the subtropical gyre circulation, travelling north with the Gulf Stream (Klein and Hogg, 1996). Data from profiling floats (Kwon and Riser, 2005; Fratantoni et al., 2013) and Lagrangian particle modelling (Gary et al., 2014) suggest that a minority of this water should then be exported to the subpolar gyre.

Multiple studies have sought to quantify or describe the fate of NASMW. Forget et al. (2011) estimate that atmospheric exchange is roughly twice as effective at destroying existing NASMW as ocean internal mixing. Davis et al. (2013) discern between NASMW removal by air–sea exchange and internal mixing as respectively fast and slow processes, which occur at different stages of its seasonal cycle. They argue for the existence of a persistent reservoir of NASMW below the mixed layer, which is shielded

from destruction by a layer of high PV. Gary et al. (2014) use a Lagrangian modelling approach to analyse the final stages of the NASMW life cycle, and show that the nature of NASMW destruction depends on how the water mass is defined - the lack of a universally accepted definition of NASMW commonly leads to such conflicts between studies (Joyce, 2011).

A quite different water mass in nature, NADW is one of the two primary high-density water masses formed in polar waters which act as preconditioners of the thermohaline component of the AMOC (Sévellec and Fedorov, 2011). NADW formation involves a range of processes and source water types. The dominant contributors are Labrador Sea Water (LSW, e.g. Talley and McCartney, 1982) and denser Overflow Waters (OW, e.g. Swift et al., 1980; Hansen et al., 2001), which respectively contribute to lower and upper NADW. LSW is created by cooling and convection of mode waters (McCartney and Talley, 1982) in the Labrador and Irminger Seas (Pickart et al., 2003a,b; Jong and Steur, 2016) during sufficiently severe winters (Clarke and Gascard, 1983). Meanwhile, OW forms as warm North Atlantic water crosses the Greenland-Scotland ridge, cools and sinks (Quadfasel and Käse, 2007). Resulting pressure gradients drive OW back over the ridge into the Atlantic at depth (Hansen and Østerhus, 2000).

Both source waters are exported southward following formation. The classical view from in-situ current measurements (e.g. Leaman and Harris, 1990; Molinari et al., 1992) is that the main southward pathway is the deep western boundary current (DWBC). However, new data from floats and modelled Lagrangian drifters (Bower et al., 2009, 2011) suggest that this is not the complete story. Gary et al. (2011) investigate alternative export pathways in a hierarchy of model resolutions. They find a second, slower pathway driven by a series of deep, eddy-driven anticyclonic gyres. These facilitate the southward transport of water parcels which have become detached from the DWBC, but are only found in models which are of eddy-resolving resolution. At lower resolution, Lagrangian particles tend to follow the DWBC.

Beyond the Equator, NADW mixes laterally and vertically with ambient water masses, eventually entering the Indian and Pacific basins via the Antarctic Circumpolar Current (Reid and Lynn, 1971). In these basins, it retains a thickness of several hundred metres (Johnson, 2008), and reaches an estimated age of several centuries (e.g. Broecker, 1979; Hirst, 1999).

Here, we present a new tool which allows us to track water masses to and from their point of formation, and apply it to NASMW and NADW. We demonstrate the use of this development by considering the life cycles of the above water masses in a new framework. The paper is set out as follows. In Sect. 2.2, we summarise the mathematical theory of the TAM approach. We then describe the model, our modifications to it, and how these may be applied to water mass tracking. The application of this development to NASMW is discussed in Sect. 2.3 with similar

considerations for NADW in Sect. 2.4. We conclude in Sect. 2.5 with a summary of our method and findings, highlighting any recommendations for future work.

2.2 Development of the passive tracer module

2.2.1 Mathematical background

The tangent-linear method describes the evolution of perturbations to the ocean state in a linear framework. Perturbations are described by vectors following the structure of the ocean state vector, containing values for all prognostic variables at all locations. The linear evolution of an initial perturbation $|\boldsymbol{u}_0\rangle$ to the ocean state vector is described by the equation

$$|\boldsymbol{u}(t)\rangle = \boldsymbol{\Psi}(t, t_0)|\boldsymbol{u_0}\rangle \tag{2.1}$$

where $|\boldsymbol{u}(t)\rangle$ provides the condition of the perturbation at time t, $\boldsymbol{\Psi}(t, t_0)$ is known as the propagator matrix and we have used the "bra-ket" notation of Dirac (1939) in which row and column vectors are respectively written as bras ($\langle\boldsymbol{a}|$) and kets ($|\boldsymbol{b}\rangle$) such that closed bra-ket terms become scalar (through the scalar product $\langle\boldsymbol{a}|\boldsymbol{b}\rangle = c$).

The adjoint approach considers the sensitivity of properties of interest to earlier perturbations. For a "cost function" J (mapping the state vector to such a property) which is scalar-valued and linear, that is $J(\boldsymbol{u}) = \langle\mathsf{F}|\boldsymbol{u}\rangle$ (for some F), the sensitivity is given by

$$\frac{\partial J}{\partial|\boldsymbol{u_0}\rangle} = \boldsymbol{\Psi}^\dagger(t_0, t)|\mathsf{F}\rangle \tag{2.2}$$

where $\boldsymbol{\Psi}^\dagger(t_0, t)$ is the adjoint of the propagator matrix, corresponding to its transpose in a Euclidean inner product space. These relationships are derived in full by Errico (1997), who provides a thorough treatment of the linear method and its limitations, beyond the simplified, inherently linear use-case of the present study. Finally, we have

$$\langle\mathsf{F}|\boldsymbol{\Psi}(t, t_0)|\boldsymbol{u_0}\rangle = \langle\boldsymbol{u_0}|\boldsymbol{\Psi}^\dagger(t_0, t)|\mathsf{F}\rangle \tag{2.3}$$

following the definition of the adjoint as a linear operator.

2.2.2 Passive tracer implementation and model description

Perturbations $|\boldsymbol{u}_0\rangle$ such as in Eq. 2.1 typically induce an active, or dynamic response (Marotzke et al., 1999). For temperature or salinity perturbations, this corresponds to a modified density field evoking changes in streamlines by modifying pressure gradients. This occurs in tandem with the passive advection and diffusion of the perturbed tracer. We may represent this by splitting the active (A) model propagator

into dynamical (D) and purely passive (P) components
$\mathbf{\Psi}_{\mathrm{A}}(t,t_0) = \mathbf{\Psi}_{\mathrm{D}}(t,t_0) + \mathbf{\Psi}_{\mathrm{P}}(t,t_0)$. Note that we use the term "purely passive" for $\mathbf{\Psi}_{\mathrm{P}}$ because the dynamical component $\mathbf{\Psi}_{\mathrm{D}}$ encompasses all dynamic interaction between the unperturbed ocean state and the perturbed field. This itself includes a partially passive element (for example passive transport of dynamically modified tracers). The purely passive component $\mathbf{\Psi}_{\mathrm{P}}(t,t_0)$, on the other hand, involves no such interactions.

In the case of propagating a passive tracer, $\mathbf{\Psi}_{\mathrm{D}}(t,t_0)$ vanishes. Our perturbation is now merely an initial injection of dye of a specified concentration, $|\boldsymbol{c}_0\rangle$, and so Eq. 2.1 collapses to

$$|\boldsymbol{c}(t)\rangle = \mathbf{\Psi}_{\mathrm{P}}(t,t_0)|\boldsymbol{c}_0\rangle \tag{2.4}$$

This dye is unable to feed back on the ocean state, and so as a perturbation is trivial. We will hence refer to initial perturbations as "injections" and evolved perturbations as dye concentrations, as is more appropriate for the passive tracer case. The common term "cost function", which relates to optimisation problems, is also a misnomer in the more simple context of propagating passive tracers. In our analysis, $\langle\mathbf{F}|$ acts on the concentration vector $|\boldsymbol{c}(t)\rangle$ to produce a tracer budget. As such we will more suitably refer to cost functions as budget co-vectors, $\langle\mathbf{B}|$.

Hence, in summary, the tangent-linear model $\mathbf{\Psi}_{\mathrm{P}}$ describes the time evolution of dye concentrations $|\boldsymbol{c}(t)\rangle$ in response to initial dye injections $|\boldsymbol{c}_0\rangle$. Accordingly, the adjoint model $\mathbf{\Psi}_{\mathrm{P}}^{\dagger}$ describes the sensitivity of tracer budget co-vectors $\langle\mathbf{B}|$ to earlier dye injections $|\boldsymbol{c}_0\rangle$.

The model used herein is the NEMO 3.4 OGCM (Madec, 2012) with the tangent-linear and adjoint package NEMOTAM (Vidard et al., 2015). In particular, the former (nonlinear model) provides the unperturbed, nonlinear model trajectory. Meanwhile the latter (linear model) provides a model implementation of the propagator and its adjoint. Both models are used in the ORCA2-LIM configuration. This consists of a 2° grid, with refinement of the equatorial and Mediterranean regions (Madec and Imbard, 1996). There are 31 vertical levels, whose height follows a hyperbolic tangent function of depth, ranging from 10 m at the surface to 1000 m. There are two ice layers (Fichefet and Maqueda, 1997) in the background state. The ice model is not linearised or coupled to NEMOTAM, although in the context of this study, ice dynamics are inherently unnafected (as a tracer which is passive cannot affect ice). The model is forced using the single, repeating normal year of the CORE forcing package (Large and Yeager, 2004). We therefore consider our results in a climatological, as opposed to historical, context. In this configuration, model NASMW (see Section 2.3 for more details on the definition and properties of NASMW in the nonlinear model) most persistently outcrops in the neighbourhood of the time-mean North Atlantic barotropic stream function maximum, which is 38.8 Sv (Fig. 2.1, red shading and grey contours). Below the surface, it occupies a narrow latitudinal band at depths of up to 310 m (Fig.

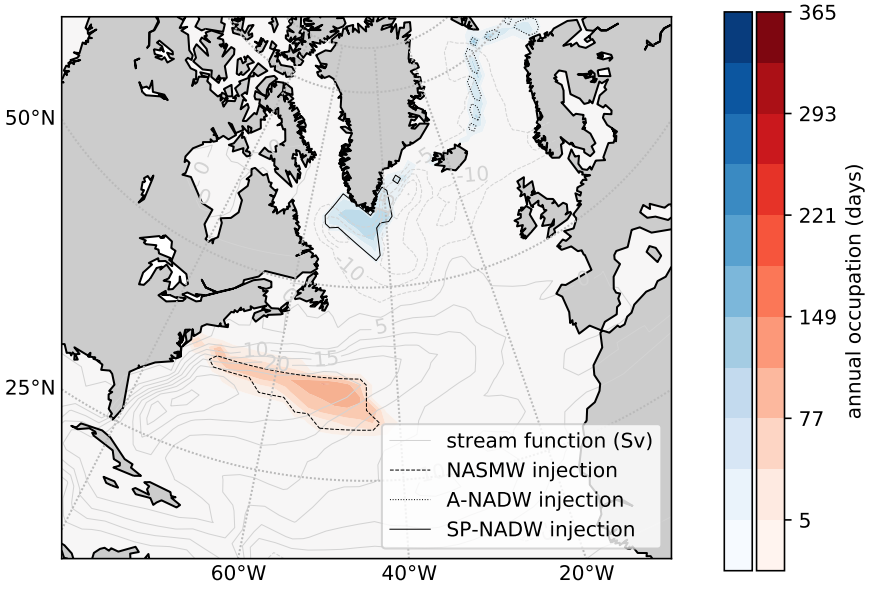
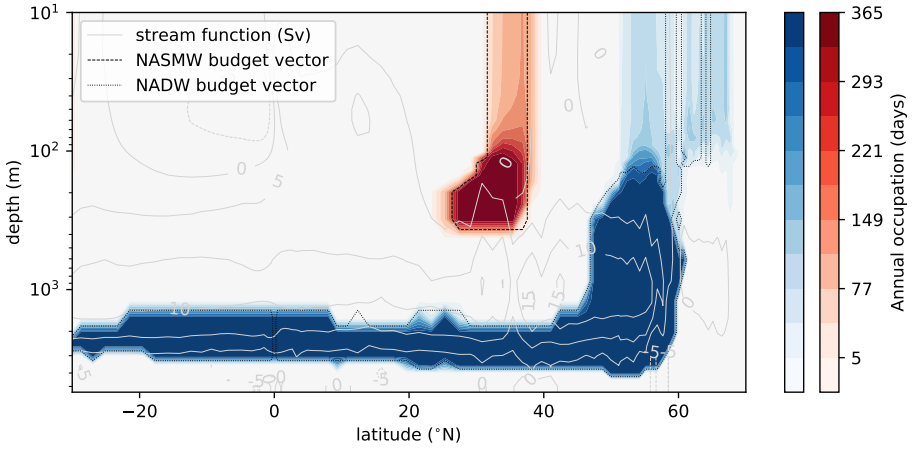


FIGURE 2.1: Outcrop locations of NASMW (red) and NADW (blue) from a 60 year climatology of the nonlinear model, following their respective definitions in Sections 2.3 and 2.4. Shading indicates the number of days of the year when the water mass is found at a given latitude and longitude. Light grey contours show the time-mean barotropic stream function for the North Atlantic. The dashed (solid, dotted) contour shows the distribution of the dye injection which initialises the tangent-linear run for NASMW (SP-NADW,A-NADW) at the annual maximum (see Section 2.3.2 (2.4.2.1,2.4.2.2))

2.2, red shading). Conversely, model NADW (as defined and described in Section 2.4) most often surfaces in the region close to the barotropic stream function minimum of -25.3 Sv (Fig. 2.1, blue shading and grey contours). At most latitudes, NADW occupies depths below the local time-mean meridional stream function maximum (Fig. 2.2, blue shading and grey contours). The overall time-mean North Atlantic meridional stream function maximum is 16.1 Sv, occurring at 42°N at a depth of 870 m.

To isolate the purely passive response detailed in Eq. 2.4, we set velocity and sea surface height (SSH) modifications to zero in the NEMOTAM time stepping procedure (Fig. 2.3). Further feedbacks involving parameterisations such as vertical and eddy-induced mixing are already absent in NEMOTAM due to approximations in linearisation (Vidard et al., 2015) and so did not require further modification. Additionally, eddy-induced advective velocities are computed online by the nonlinear model and not stored explicitly in the trajectory. They may be recomputed from other trajectory variables, but this is not the default behaviour of NEMOTAM. As such, they were not included for the passive-tracer experiments in this study. They have been since been enabled in the passive-tracer source code as an optional online calculation. A preliminary comparative test suggests that errors are minor for our case studies (positive and negative concentration differences tend to cancel out, but even summing absolute differences, we do not exceed $\sim 6\%$ relative to the injection volume).

FIGURE 2.2: Zonally averaged distribution of NASMW (red) and NADW (blue) in a 60 year climatology of the nonlinear model, following their respective definitions in Sections 2.3 and 2.4. Shading indicates the number of days of the year when the water mass is found at a given latitude and depth. Light grey contours show the meridional stream function of the Atlantic. Dashed contours show the distribution at the point of the year when the adjoint run is started



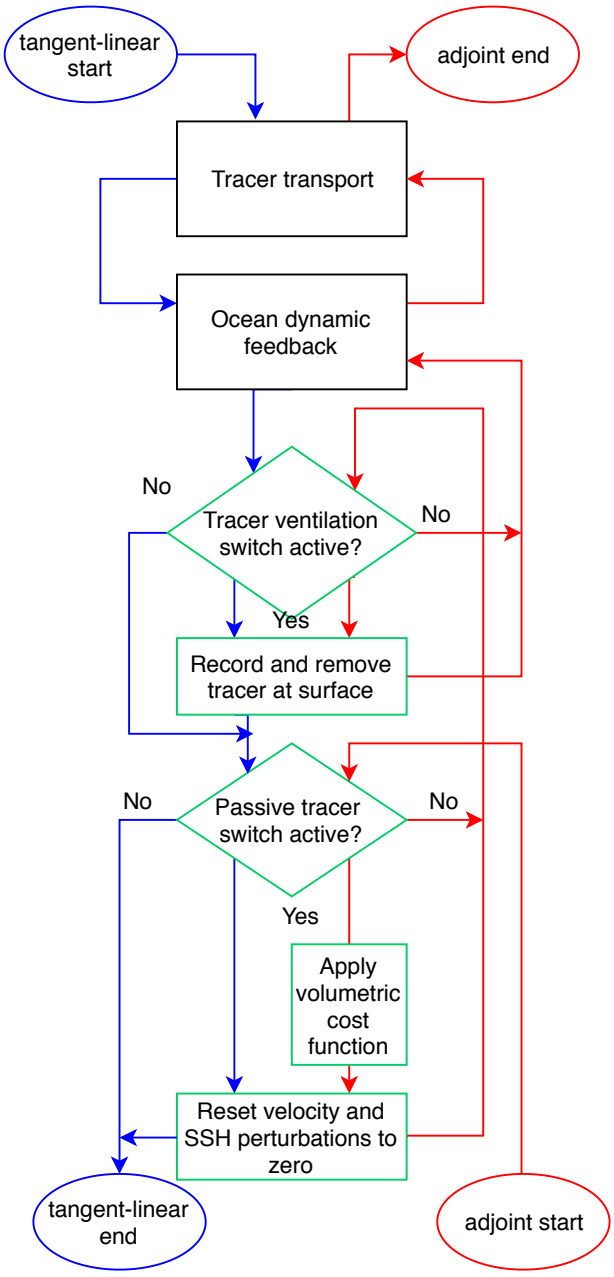
We expect the difference to be greater in steep isopycnal regions such as the Southern Ocean.

Our water mass tracking procedure is as follows. We begin by defining the water which we intend to track, for example using a temperature-salinity (TS) range. To track the evolution of newly formed water of this type, we identify all surface locations where T and S properties meet our definition in the nonlinear model trajectory. We then inject into the tangent-linear model a dye concentration of 1 in these locations, and run the model forward. To identify the origins of existing water using the adjoint model, the procedure is similar. We begin by identifying model grid cells in the nonlinear trajectory (at all depths) which satisfy our TS definition. The budget vector is populated using the volume of these grid cells, with zero values elsewhere. This budget vector is then propagated back in time using the adjoint model to produce a sensitivity of the budget to earlier concentrations. In both cases, we remove any tracer which reaches the surface using the model's surface temperature restoring scheme. This scheme restores concentrations towards zero with a time scale of 60 days for a 50 m mixed layer (Madec, 2012). In the adjoint case, a record of this restoring leads to a spatiotemporal probability distribution of the surface origins of the tracer.

2.2.3 Advection schemes

The default advection scheme of the nonlinear model, used to compute the background state around which the model is linearised, is a total variation diminishing (TVD) scheme. This is a flux-corrected transport scheme (FCT, Lévy et al., 2001), which

FIGURE 2.3: Adaptation of NEMOTAM time-stepping procedure for the tangent-linear (blue) and adjoint (red) cases. Steps outlined in green are new additions to the scheme.



balances an upstream scheme with a centred scheme using a nonlinear weighting parameter described by Zalesak (1979). This scheme is preferable to others available in the model for non-eddying configurations (Lévy et al., 2001), but its nonlinearity means that it is not suitable for a TAM. As such, the standard advection scheme of NEMOTAM is a linearised counterpart in the form of a second-order forward-time-centred-space (FTCS) finite difference scheme. A caveat of such schemes is the presence of a negative artificial diffusion coefficient, which can lead to negative quantities of positive-definite tracers (Owen, 1984). This encourages down-gradient diffusion of otherwise zero concentrations, acting as a source of further negative tracer

when combined with the surface restoring scheme. This is particularly problematic in the case of passive-tracer advection as, unlike in the fully active TAM, negative quantities are unphysical.

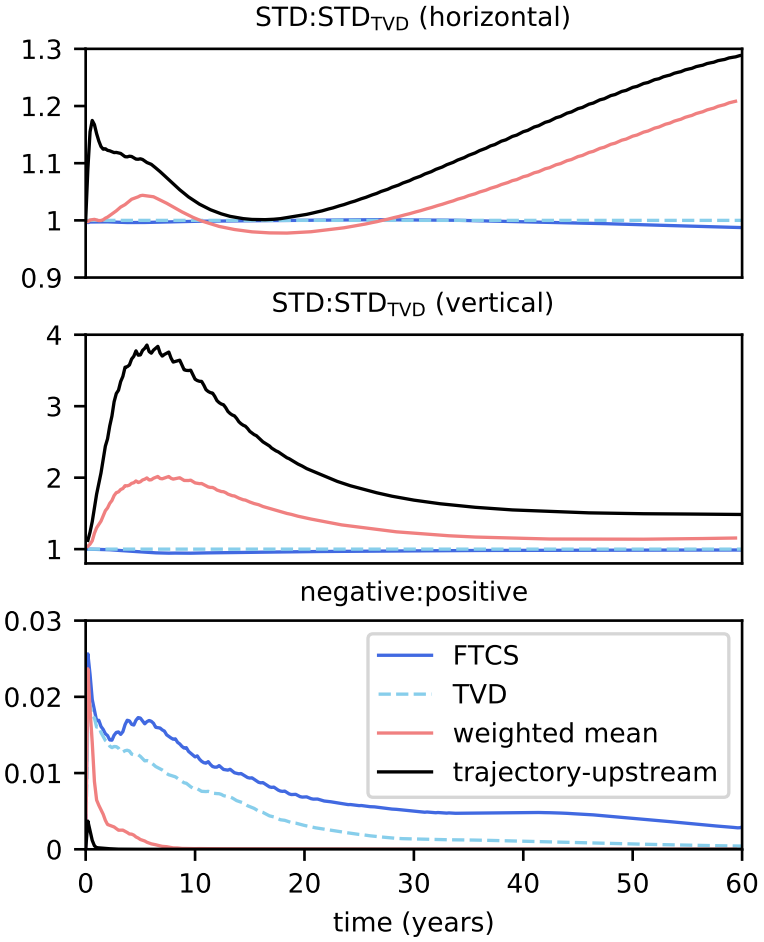
An alternate scheme which is first-order linear (thus compatible with tangent-linear models and added to NEMOTAM for this study) is the trajectory-upstream scheme. This is an approximation to the classical upstream scheme, whereby the upstream direction is inferred from the trajectory velocity field (avoiding nonlinear sign functions). As the velocity field cannot be modified in the passive tracer case, however, the method is everywhere equivalent to the classical scheme. The scheme is forward in both time and space, and offers the advantage of not producing negative tracer values. Nevertheless, this scheme also possesses an artificial (positive-valued) diffusion term, leading to a disproportionate spread of tracer, particularly in the vertical.

To strike a balance between the relative merits and caveats of the linear schemes described above, an approximation to the weighted-mean scheme of Fiadeiro and Veronis (1977) was coded into NEMOTAM. Like the TVD scheme, the true weighted-mean scheme is a linear combination of the classical upstream and FTCS methods. As before, we approximate the classical upstream method by using the trajectory-upstream scheme. The weighting parameter is determined dynamically from the trajectory. Strongly advective regions lean towards upstream advection, while strongly diffusive regions favour the centred formulation. Despite inheriting issues from both FTCS and the trajectory-upstream schemes, it is preferable to either individually. Its negative tracer concentrations are less persistent than those of the TVD and FTCS schemes, and its anomalous vertical spread was found to be less than half that of the upstream scheme (Fig. 2.4). Some issues have been raised with multivariate and time-dependent extrapolations of the weighted-mean scheme (e.g. Gresho and Lee, 1981). However, nonlinear flux-corrected transport schemes are required in order to improve on the weighted-mean scheme (Gerdes et al., 1991), and these are not compatible with the linear model. We thus proceed using our approximation to the weighted-mean scheme (in both the horizontal and vertical) for the water mass case studies detailed in this paper. We remark that it is entirely possible to incorporate a nonlinear advection scheme into NEMOTAM, but that the resulting quasi-linear model would not have a true adjoint, which we require for our study.

2.2.4 Performance

In order to test performance, the model was run for one-hundred days at 2° resolution (the ORCA2 configuration used in our case studies) and five days at 0.25° resolution (the ORCA025 configuration, Madec, 2012). Each configuration was run in four different modes (nonlinear, nonlinear with TAM trajectory production, tangent-linear with passive tracer, adjoint with passive tracer). Trajectory files were written once per

FIGURE 2.4: Comparison of advection schemes in NEMOTAM. To assess the schemes, one of the runs of this study (the tangent-linear run of Sect. 2.3.2) was repeated with four different schemes and the tracer distributions were compared. Top panel: Lateral spread (spatial standard deviation around mean tracer position) as a ratio to the corresponding spread in the nonlinear TVD scheme, Middle panel: vertical spread (as above). Bottom panel: ratio of volumes occupied by negative and positive tracer concentrations, respectively.



day, and the linear advection scheme was the weighted-mean scheme described in Section 2.2.3. Each test was conducted with a range of parallelisation arrangements (16, 32, 64, 128, 256 and 512 CPU cores) and repeated 10 times to account for system variability (Table 2.1). The tests were conducted on a local HPC system, with 72 compute nodes, each offering two eight-core 2.6GHz processors (Intel Xeon E5-2650 v2) and 64GiB, connected by an InfiniBand QDR network.

TABLE 2.1: Results of performance tests for two model configurations (ORCA2, upper section and ORCA025, lower section). Top rows: trajectory storage requirements per output. Lower rows: runtime per model day for four model modes: Nonlinear with trajectory output (NLT), nonlinear without trajectory output (NL), tangent-linear (TL) and adjoint (AD). The standard deviation of ten runs is given as a ± uncertainty. Dashes indicate tests which failed due to insufficient memory.

Cores (nodes)			16 (1)	32 (2)	64 (4)	128 (8)	256 (16)	512 (32)
ORCA2								
Traj. size (MB/d)			43.909	44.418	43.978	44.412	44.848	46.351
Mode	NLT	Run time (s/d)	2.09±0.20	1.54±0.17	1.23±0.17	1.39±0.26	2.53±0.27	4.08±0.35
	NL		1.98±0.17	1.37±0.10	1.10±0.09	0.99±0.08	1.85±0.36	2.18±0.70
	TL		0.78±0.01	0.50±0.01	0.36±0.09	0.36±0.13	0.45±0.18	0.88±0.10
	AD		0.89±0.01	0.61±0.02	0.53±0.04	0.50±11	0.59±0.05	1.99±0.61
ORCA025								
Traj. size (MB/d)				5217.427	5218.495	5220.824	5217.658	5218.448
Mode	NLT	Run time (s/d)		1091.55±69.38	692.98±21.38	428.36±13.76	302.18±30.55	143.60±12.73
	NL		-	1065.48±54.56	656.94±13.74	386.34±13.89	262.72±34.56	131.70±13.28
	TL				315.26±26.17	207.82±14.43	136.42±13.00	64.70±4.20
	AD			-	403.74±36.18	252.36±14.17	181.12±7.70	92.40±6.26

The general order of time efficiency is consistent across all tests: the linear forward model is between 2.5 (16 cores) and 3.1 (64 cores) times as fast as the nonlinear model without trajectory output in ORCA2, and 1.9 (64 cores) to 2.0 (512 cores) times as fast in ORCA025. The adjoint is slightly less efficient. The nonlinear runs which produce trajectory output are the slowest across the board, due to the high level of output. Memory use appears higher in the linear model than the nonlinear, such that the linear model could not be run on two 64GiB compute nodes alone in ORCA025 (column two).

The model shows good scaling in the ORCA025 configuration at all CPU arrangements tested here, but begins to worsen in ORCA2 beyond 128 CPU cores. Further, the added time required for trajectory output in nonlinear ORCA2 runs on 128 CPU cores can be considerable. This is possibly due to the generation of a very large number of files during the model run, as the number of files generated is proportional to the number of CPU cores, the trajectory write frequency, and the run length. The required storage size remains relatively constant for runs with larger number of CPU cores, despite the greater number of files (Table 2.1, first row).

An additional limitation which we discovered during our longest experiments comes from system file-number limits, which can readily be reached for typical systems for very long runs using many cores.

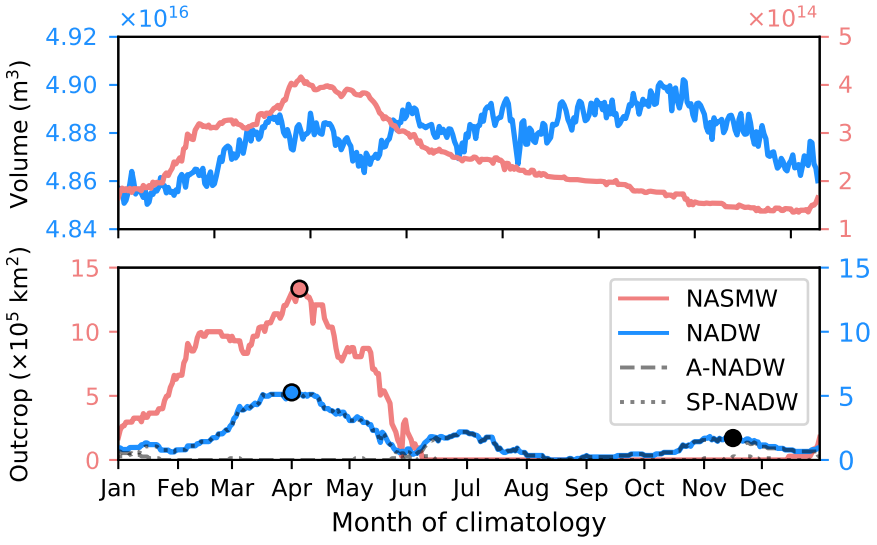
2.3 Application to North Atlantic Subtropical Mode Water

2.3.1 NASMW definition and properties

We now apply the developments of Sect. 2.2 to the problem of tracking the origins and fate of NASMW in the ORCA2-LIM simulation. There is no universally accepted definition of NASMW, with differing definitions leading to conflicting results between studies (Joyce, 2011). While the method allows us to define water masses in terms of any model variable, we choose for simplicity to utilise the common approach of a temperature-salinity range. In particular, we define model NASMW as water lying in the temperature band $[17, 19]^\circ$ C, as in the original definition of Worthington (1959), with salinity in the range of $[36.4, 36.6]$ psu. As in some other studies (e.g. Kwon and Riser, 2004; Gary et al., 2014), we impose a third criterion, in our case that NASMW has a thickness of at least 125 m. This condition ensures homogeneity and isolates the water mass from the adjacent Madeira Mode Water (MMW) to its east (Siedler et al., 1987).

Model NASMW has a consistent outcrop in a single area (Fig. 2.1) and does not extend far south of its outcrop region below the surface (Fig. 2.2). The outcrop area of

FIGURE 2.5: Mean-year time series of NADW (blue) and NASMW (red) volume (top panel) and outcrop surface area (bottom panel) including those of the Arctic (dashed line) and subpolar (dotted line) NADW outcrops from the nonlinear trajectory. Circles indicate the injection time of NASMW (red), NADW (blue) and A-NADW (black) in the forward simulations (see Sections 2.3.2 and 2.4.2)

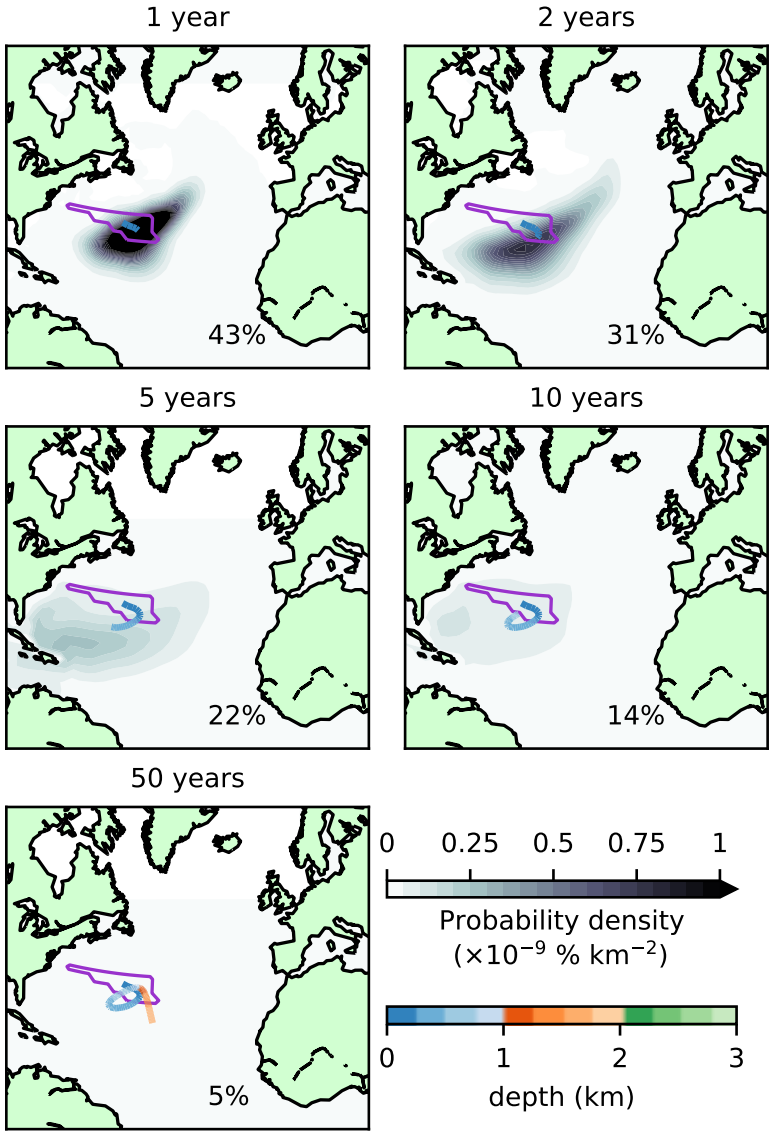


NASMW in the model peaks at the beginning of April at 1.3×10^6 km^2 (Fig. 2.5), coinciding with peaks in volume (4.2×10^5 km^3) and thickness (310 m) of the water mass. These maxima occur a month before the annual maximum local MLD of 430 m. Following this, rapid stratification due to summer warming shoals the mixed layer. The outcrop area concurrently diminishes until the water mass is completely sheltered from air–sea exchange. This period of submergence extends from early June to early December, as the water mass' upper surface progressively deepens to a maximum of 101 m. The volume also decreases, towards 1.3×10^5 km^3 at the annual minimum.

2.3.2 Tangent-linear run

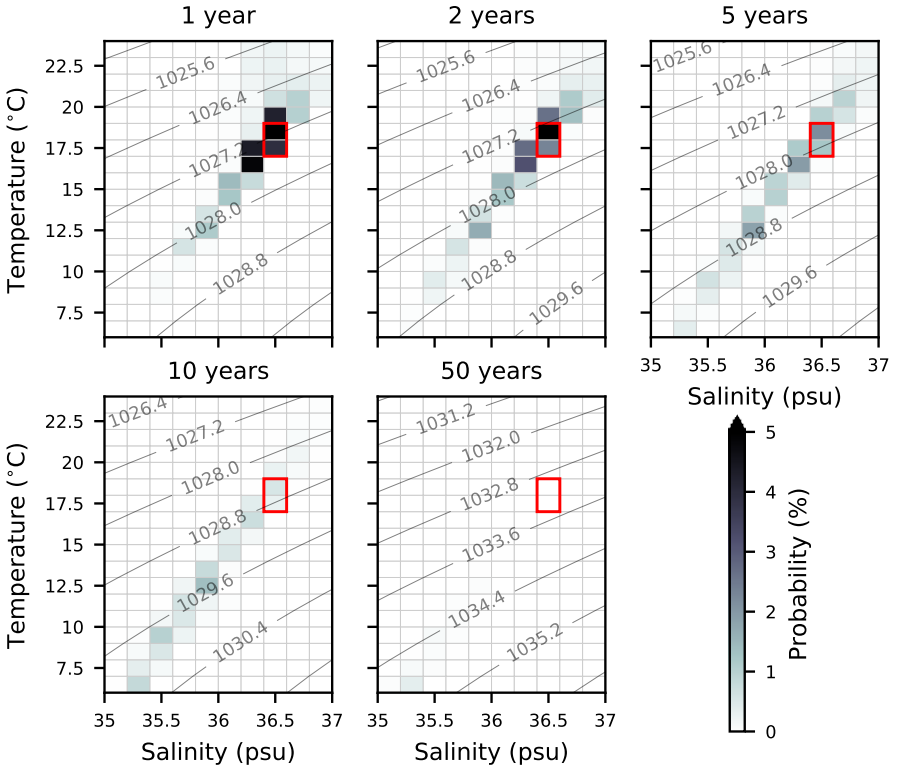
As outlined in Sect. 2.2.2, we begin by identifying all water matching our NASMW definition in the surface layer of the nonlinear model at a given time. We choose this to be the time when the outcrop is maximal (Fig. 2.5). We then inject a concentration of 1 into the tangent-linear model at these locations, and propagate it for 60 years. Newly ventilated model NASMW initially resides close to the surface (Fig. 2.6). Here, its behaviour is governed by surface currents, which fractionate the water mass. At the point of application, 37% of the tracer lies in the Gulf Stream. Most of this water follows the recirculation of the subtropical gyre, while a minority is exported to the subpolar region (2.5% over four years). Our quantification of subpolar export is on the same order of that found under similar conditions by Burkholder and Lozier (2011), who analyse inter-gyre exchange in a purpose-built Lagrangian study.

FIGURE 2.6: Evolution of NASMW-tagged tracer in the tangent-linear model. Black and white shading indicates depth-integrated probability density (corresponding to the likelihood that NASMW is found in that region) at 1, 2, 5, 10, and 50 yr. Inset percentages show the global integral of this field, i.e. the total proportion of tracer not yet re-ventilated at the surface. The coloured line tracks the centre of mass of the tracer from initialisation to its current position, with colours indicating the mean depth of the tracer. The purple contour shows the distribution of the original tracer injection.



Our NASMW is short-lived. Persistent proximity to the surface leads to the re-ventilation of 95% of the initialised tracer over the 60 year run. Of that which remains in the ocean, 90% is transformed and no longer fits our definition of NASMW (Fig. 2.7). As time progresses, because of vertical mixing together with near-surface tracer removal by the restoring scheme, remaining tracer tends to reach deeper, colder, fresher waters. Using the passive tracer approach, it is not possible to establish a full,

FIGURE 2.7: Evolution of model NASMW in TS space. Shading indicates likelihood that the water mass found in a particular TS class after 1, 2, 5, 10 and 50 yr. The red box marks the TS range used to define the water mass in this study. Contours show the density at the average depth level of the tracer.



qualitative description of the fate of NASMW. However, it can be observed that only around 5% of all model NASMW re-ventilation occurs within its surface outcrop. The rest is mixed out of the TS class, either remaining in the system for the remainder of the 60-year simulation, or (more commonly) resurfacing elsewhere.

Due to rapid re-exposure, the e-folding time of our NASMW is just 60 days. This is shorter than the estimations of Gary et al. (2014, 1 yr) and Kwon et al. (2015, 3 yr), but this is likely due to inherent differences in formulation. In the above studies, Lagrangian particles released in NASMW are not removed upon contact with the surface, so as to account for re-ventilation. Accounting for re-exposure, Fratantoni et al. (2013) suggest around 75 days, more closely aligned with our own findings.

2.3.3 Adjoint runs

To track existing model NASMW back to its source locations, we again follow the procedure outlined in Sect. 2.2.2. This consists of taking a budget of NASMW in the final year of a nonlinear simulation and propagating this budget backward using the

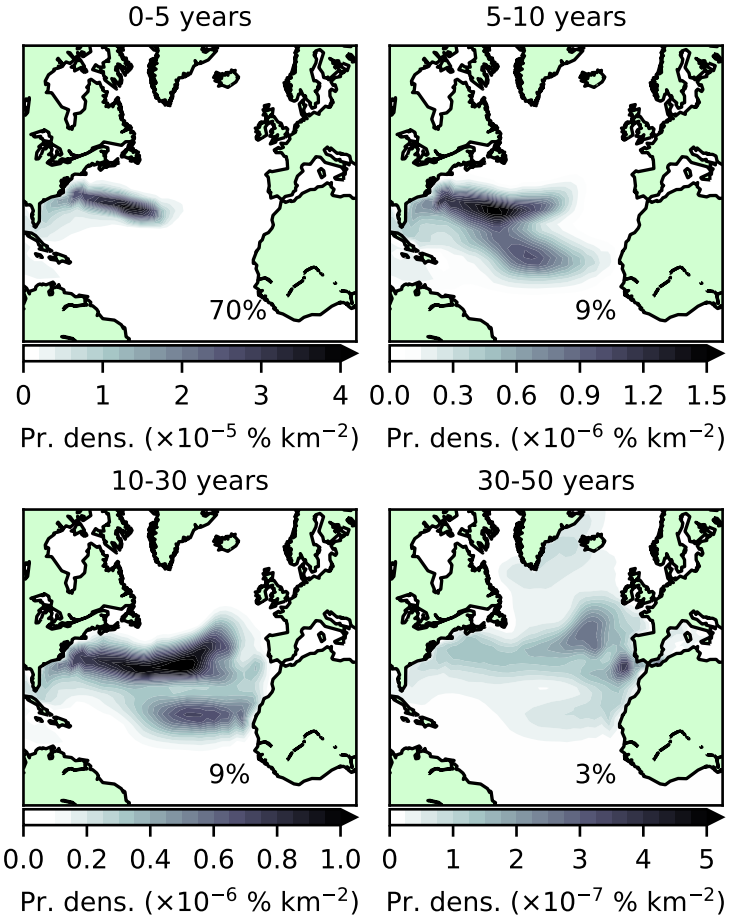
adjoint model. We begin the adjoint run at the same point in the annual cycle as the tangent-linear run, when model NASMW volume is at its maximum (Fig. 2.5). We divide the mode water budget co-vector $\langle \mathbf{B}_{\mathrm{MW}}|$ into tracer above the mixed layer depth ($\langle \mathbf{B}^{\mathrm{U}}_{\mathrm{MW}}|$) and tracer below ($\langle \mathbf{B}^{\mathrm{L}}_{\mathrm{MW}}|$), and propagate each part separately. This allows us to explore the potential for a resilient layer of older NASMW below the MLD (Davis et al., 2013). By linearity, the union of the propagated budget vectors is equivalent to the propagated budget vector of NASMW as a whole:
$\boldsymbol{\Psi}^{\dagger}(t_0,t)(|\mathbf{B}^{\mathrm{U}}_{\mathrm{MW}}|\rangle + |\mathbf{B}^{\mathrm{L}}_{\mathrm{MW}}\rangle) = \boldsymbol{\Psi}^{\dagger}(t_0,t)|\mathbf{B}_{\mathrm{MW}}\rangle$

Most NASMW propagated with the adjoint model reaches the surface quickly; 70% of the tracer-tagged water mass is under 5 years old. During this early stage, ventilation occurs predominantly within the subtropical gyre recirculation, in the neighbourhood of the initial NASMW outcrop (Fig. 2.8). The 60-year mean formation location is also within this region, at 32°N, 58°W. This is almost coincident with the core of NASMW formation determined by Warren (1972), and agrees with the air–sea exchange-based estimate of Worthington (1959).

For water over 5 years old, current patterns begin to have a distinct influence on formation. There is a clear signature of the Gulf Stream on the youngest model NASMW. This evolves backwards as the adjoint propagates, eventually leaving an imprint of the entire subtropical gyre (Fig. 2.8). Simultaneously, newly formed mode water in the eastern North Atlantic (MMW) begins to cross into NASMW within 10 years. The signature of the Mediterranean outflow is particularly strong on the very oldest model NASMW, which also contains contributions from the subpolar region. The culmination of all of these water types throughout the 60-year evolution is evident when the model's sources of NASMW are viewed in TS space (Fig. 2.9). Also apparent is that the primary source of NASMW has a warmer signature, which reflects the advection of warmer waters from the south, which cool and subduct to form the water mass. While the surface distribution suggests that a relatively small neighbourhood dominates the formation of model NASMW (Fig. 2.8), we remind the reader that the seasonal variability of surface properties (particularly the outcrop area) reflects strongly on the thermohaline properties of NASMW at formation (Fig. 2.9). This suggests that the NASMW surface outcrop is not, in fact, the dominant origin of NASMW in the model, contrary to the intuition provided by laminar models of the ventilated thermocline (e.g. Luyten et al., 1983). The implication is that mode water formation is not a passive process, and that ocean dynamics must play a fundamental role.

We also consider the time scales involved with NASMW formation in the model. By recording the time at which tracer is removed from the budget by the surface restoring scheme, we may construct a probability distribution function (PDF) of water mass age (Fig. 2.10). There is a visibly lower skewness for NASMW lying below the MLD (Fig. 2.10, blue bars). Also clear is the seasonal cycle of shielding brought on by summer

FIGURE 2.8: Surface origins of model NASMW as determined by the backtracking budget analysis (adjoint model simulation). Shading indicates probability density. This corresponds to the likelihood that model NASMW is formed in a given region during the time periods [0,5] yr, (5,10] yr, (10,30] yr and (30,50] yr. Inset percentages show the global integral (that is, the total proportion of the budget which is formed during each time period). Note that, contrary to Fig. 2.6, which displays instantaneous fields, here time-integrated fields are displayed. Due to the large variability of its extent over the integration window, the outcrop region is not shown.



stratification, with less NASMW reaching the surface during these periods. Due to the simulation beginning at the annual maximum in NASMW production, peak NASMW formation occurs almost instantly. For tracer below the mixed layer (3.78% of the total) the peak does not occur until the outcrop maximum of the third year. From the PDFs, the expected age of the model NASMW constituents above and below the MLD are, respectively, 4.70 yr and 8.79 yr. It is evident that NASMW located below the MLD is shielded from renewal, with nearly double the expected age of model NASMW as a whole.

We finally address the asymptotic tail of the PDF, which represents the oldest waters found within NASMW. Our findings suggest a high-latitude source makes up a large

FIGURE 2.9: Surface origins of model NASMW as determined by the backtracking budget analysis (adjoint model simulation) in TS space. Shading indicates the proportion of the budget originating from a particular TS class at the surface within 50 yr. The red box marks the TS definition of NASMW used in this study. Contours show surface density.

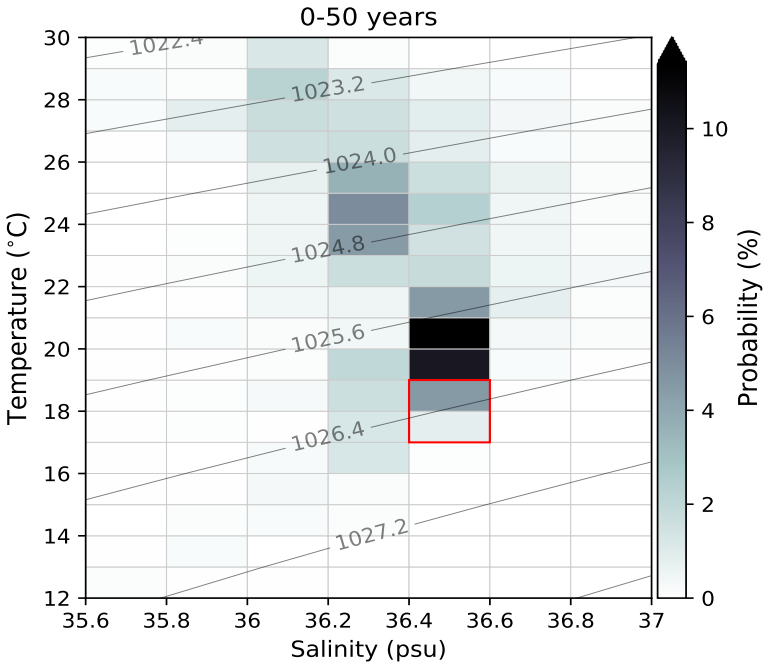
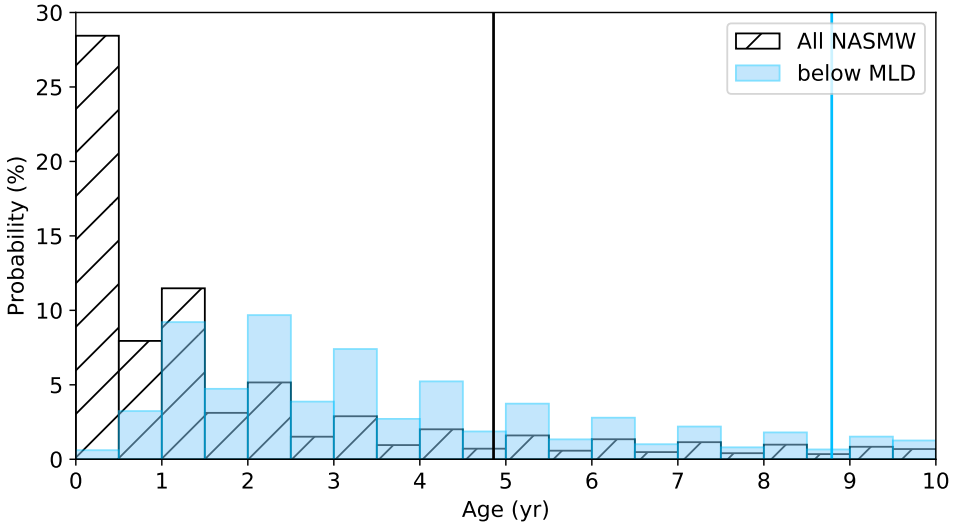


FIGURE 2.10: Probability distribution of model NASMW age (hatched bars) and age of NASMW restricted to be found below the mixed layer depth (blue bars). The percentage of the NASMW budget formed in a given 0.5 yr bin is indicated by its associated bar. The expected value of the distribution is marked by a solid line.



fraction of this water, having followed a pathway from outside of our defined thermohaline range from the surface to eventually contribute to the make-up of NASMW (Fig. 2.8). Indeed, we find that for the 30-50 yr period, some 68.3 % of NASMW originates at latitudes of 45° or higher. The idea of a distant source of the oldest NASMW is concurrent with the few studies which have considered it. Douglass et al. (2013) use an ideal-age tracer to construct a histogram similar to our own, acknowledging a remote source of the very oldest NASMW. Kwon et al. (2015) find that at least 20% of NASMW stems from other regions, and highlight the potential importance of subpolar latitudes.

2.4 Application to North Atlantic Deep Water

2.4.1 NADW definition and properties

As in Sect. 2.3.1, we now consider the properties and behaviour of NADW in the ORCA2-LIM simulation. We define NADW to fall within a temperature-salinity band of $[2, 4]$ °C, $[34.9, 35.0]$ psu, in close alignment with the definition of Worthington and Wright (1970). Analysis of the nonlinear model trajectory (about which the model is linearised) shows that there are two distinct latitudes at which water of this TS signature persistently outcrops in the simulation (Fig. 2.2, blue shading). These correspond to the (subpolar) Labrador-Irminger Sea region, southwest of the Greenland-Scotland ridge, and the (Arctic) Greenland Sea region, northeast of the ridge (Fig. 2.1, blue shading). The outcrop oscillates between the two regions with the seasonal cycle. While observed NADW is known to form only in extreme winters and has a strong interannual signature (Avsic et al., 2006), our use of repeated forcing implies that formation has little year-to-year variation. The Labrador-Irminger outcrop peaks at the end of March, around a month after the annual local mixed layer depth maximum of ~ 1 km. At this peak, water in our NADW TS class occupies 5.3×10^5 km^2 at the surface (Fig. 2.5), before the area diminishes with the shoaling mixed layer. The model's Greenland Sea NADW outcrop peaks in November, with a maximal extent of 2.3×10^5 km^2. At this time, surface NADW is exclusively found northeast of the Greenland-Scotland ridge, and the water mass is shielded from ventilation elsewhere. On average, 66% of the total annual outcrop area is southwest of the ridge, and 34% is to its northeast.

Model NADW volume is almost constant year-round at 4.9×10^7 km^3 (18.7% of the model North Atlantic), deviating by less than 0.5% annually.

2.4.2 Tangent-linear runs

As with NASMW, the dye injection $|c_0\rangle$ coincides with the time step corresponding to the annual maximum outcrop extent, which here falls in April. It should be noted that all of the locations corresponding to this maximum are southwest of the Greenland-Scotland ridge, in the subpolar Labrador-Irminger region. To investigate the nature of the Arctic outcrop, we follow a second injection of dye, in November, when its own outcrop extent is maximal. At this point, the water mass exclusively surfaces within the Arctic circle. We thus refer to these distinct surface waters as SP-NADW (for the southwestern, subpolar outcrop) and A-NADW (for the northeastern, Arctic outcrop). The tangent-linear model was run for 400 years with the SP-NADW dye injection, and 60 years with the A-NADW dye injection. The run lengths were determined by the rate of surface tracer re-ventilation in each case.

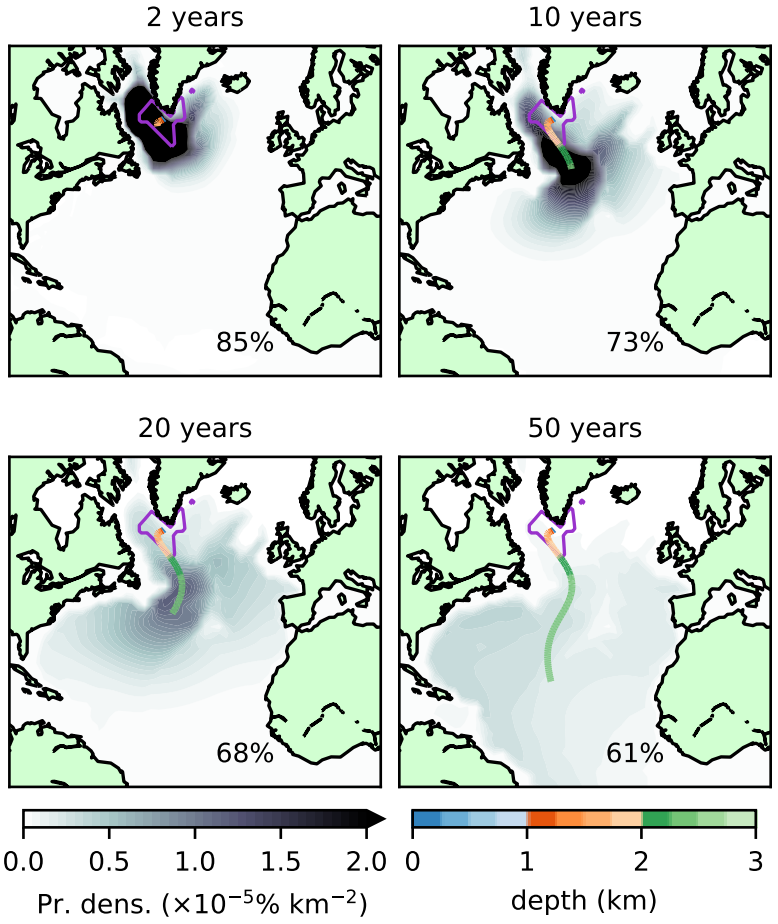
2.4.2.1 SP-NADW

SP-NADW rapidly sinks, with tracer reaching an average depth of 1235 m after 0.2 yr. It initially moves quickly westward, departing the surface region around Cape Farewell. It then spreads throughout the Labrador basin at depth and extends into the Irminger Sea (Fig. 2.11) at a temperature of 3.5° C (Fig. 2.12). This initial behaviour is in broad agreement with temperature data from hydrographic sections (e.g. McCartney and Talley, 1982) and spatial distributions captured by CFC measurements (e.g. Rhein et al., 2002), floats (e.g. Bower et al., 2009) and models (e.g. Bower et al., 2011; Gary et al., 2011).

The tracer patch then moves southward, steadily deepening. Its mean position is at first closely tied to the DWBC. However, beyond the Flemish Cap, it takes a more central course through the basin interior. While interior southward routes generated by deep eddies have been found parallel to the DWBC in recent profiling float data (Lavender et al., 2000; Bower et al., 2009), our model configuration is non-eddying. As such, these pathways are represented by parameterised turbulent diffusion of the tracer, in a manner which would not be captured by Lagrangian drifters in our model (Gary et al., 2011).

The tracer initialised in the model is quickly sequestered, and is thus not vulnerable to re-exposure. Indeed, while 27% of the initial volume of SP-NADW is re-ventilated within the first decade, only a further 24% is removed during the rest of the century (Fig. 2.13). The tracer's homogeneity is also preserved, with limited mixing into neighbouring TS classes on decadal time scales (48% remains in the original TS class after 20 years, Fig. 2.12). There is a tendency over hundreds of years for the remaining water to cool and freshen to temperatures lower than 1.25° C and salinities below 34.65 psu (Fig. 2.14). This follows the well-described behaviour of NADW following

FIGURE 2.11: As in Fig. 2.6 but for subpolar-outcropping model NADW, and times 2, 10, 20 and 50 yr.

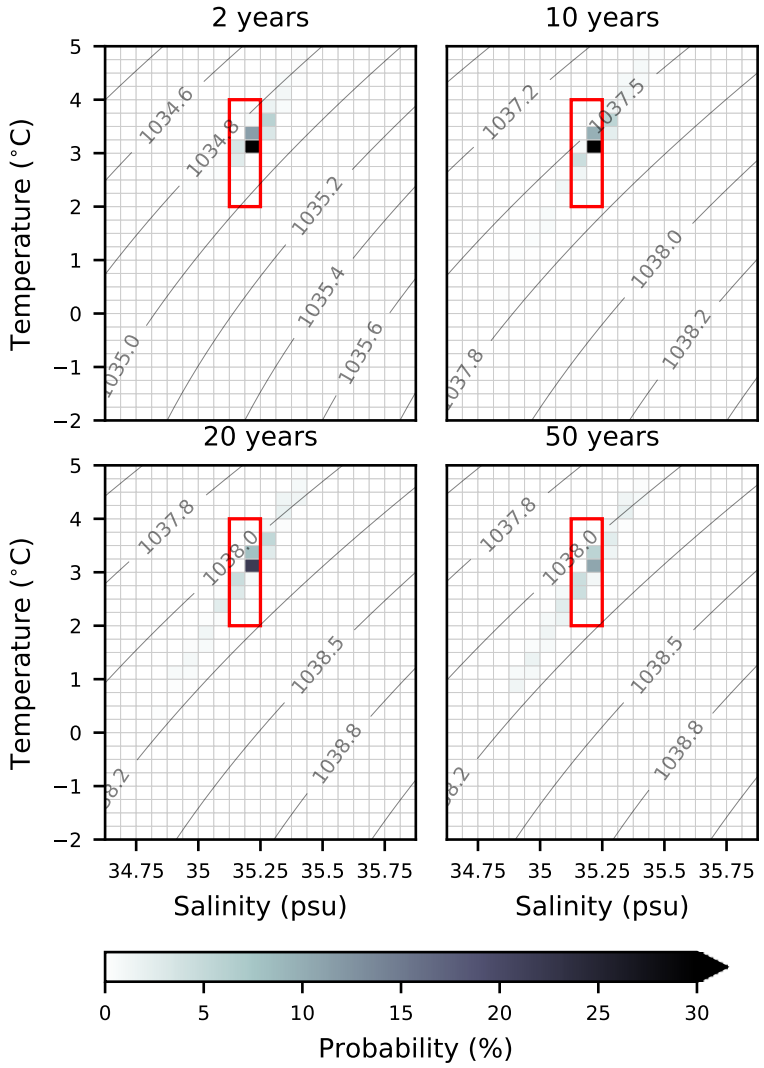


intrusion into the Southern Ocean. Observed NADW is known to mix into Circumpolar Deep Water, eventually transforming into bottom water with similar thermohaline characteristics to these (e.g. Orsi et al., 1999).

2.4.2.2 A-NADW

The forward evolution of A-NADW (Fig. 2.15) is altogether different to SP-NADW, categorised by rapid diffusion and depletion. Unlike SP-NADW, which reaches great depths quickly following formation, A-NADW remains close to the surface on the Greenland and Barents shelves, with an average depth of only 502 m after 1 yr. Due to this surface proximity throughout the run, the majority of tracer is re-exposed to the atmosphere. It is accordingly removed by the model's surface restoring scheme, with 95% of the initial volume removed after 20 yr. That which remains either spreads throughout the Arctic ocean, or spills over the shelf to join its SP-NADW counterpart in the Labrador Sea.

FIGURE 2.12: As in Fig. 2.7, but for subpolar-outcropping model NADW at 2, 10, 20 and 50 yr. Note that the region enclosed by the red box is the same in both subpolar- and Arctic-outcropping (Fig. 2.16) cases.



Of all of the initialised A-NADW, that destined for the Atlantic basin represents just 3.8%. We may use the velocity fields of the nonlinear model to estimate transport pathways of this passive tracer into the basin. Consider an idealised case with two openings into the basin (here taken to represent the Denmark Strait and west of the Reykjanes Ridge), at points y_1 and y_2. The total volume flux into the basin between two times t_0 and t_1 will be approximately

$$\Delta V_{t_0,t_1} = \int_{t_0}^{t_1} ((u_1(\tau)c_1(\tau)\Delta y_1 \Delta z_1) + (u_2(\tau)c_2(\tau)\Delta y_2 \Delta z_2))d\tau \tag{2.5}$$

where c_i is the concentration of tracer at opening i, Δy_i is the width of the opening, Δz_i is the depth, u_i is the velocity normal to $\Delta y_i \Delta z_i$, t_0 is the injection time and t_1 is 50 years later. As all of the above quantities are present in the model output, we may scale this approach up to the model grid and estimate each contribution. In this

FIGURE 2.13: As in Fig. 2.6 but for subpolar-outcropping model NADW, and times 100, 200 and 400 yr.

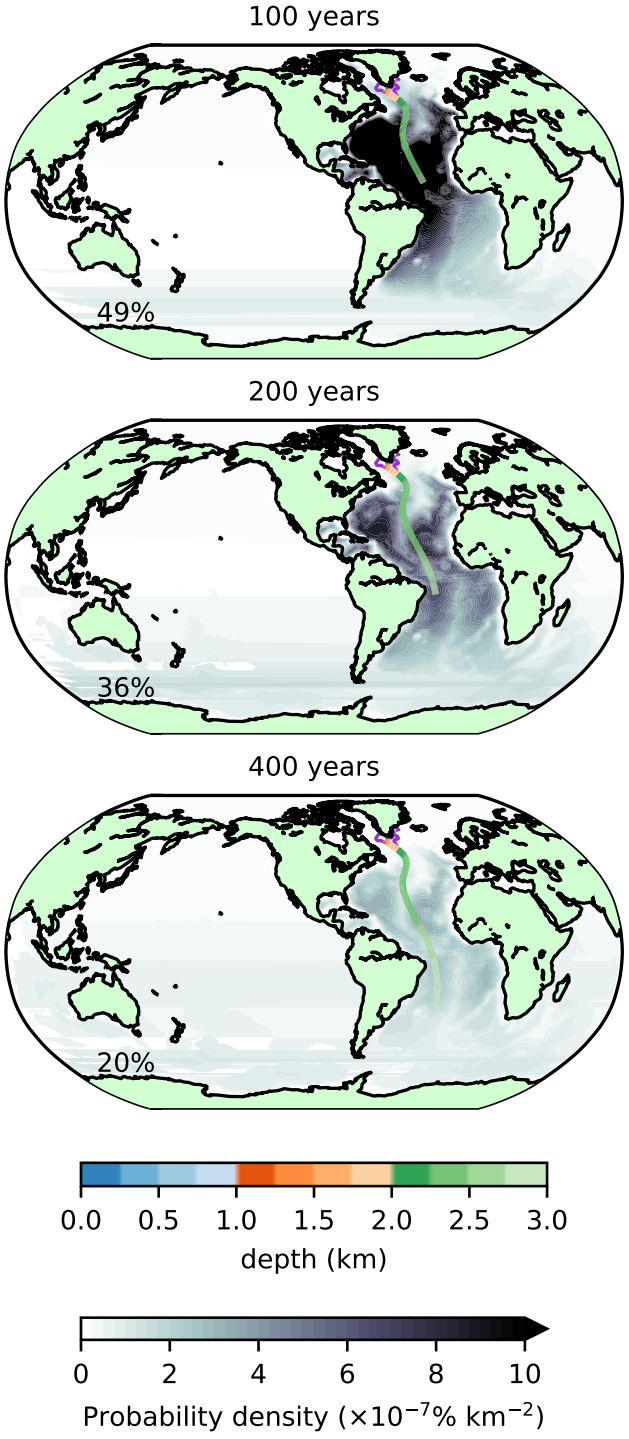


FIGURE 2.14: As in Fig. 2.7, but for subpolar-outcropping model NADW at 100, 200 and 400 yr. Note that the region enclosed by the red box is the same in both the subpolar- and Arctic-outcropping (Fig. 2.16) cases.

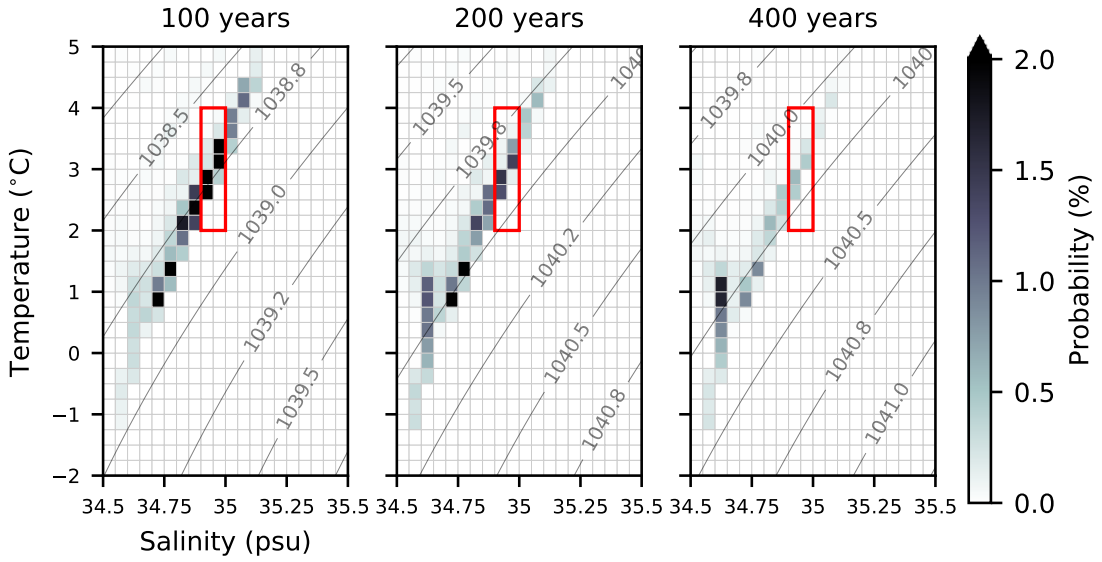


FIGURE 2.15: As in Fig. 2.11, but for Arctic-outcropping model NADW.

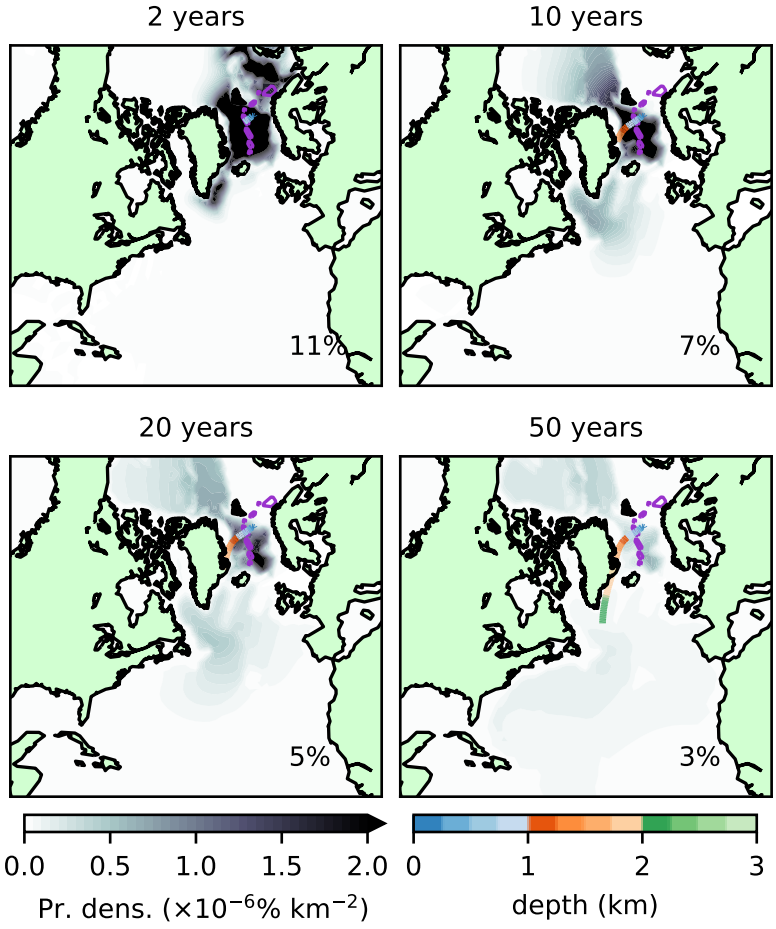
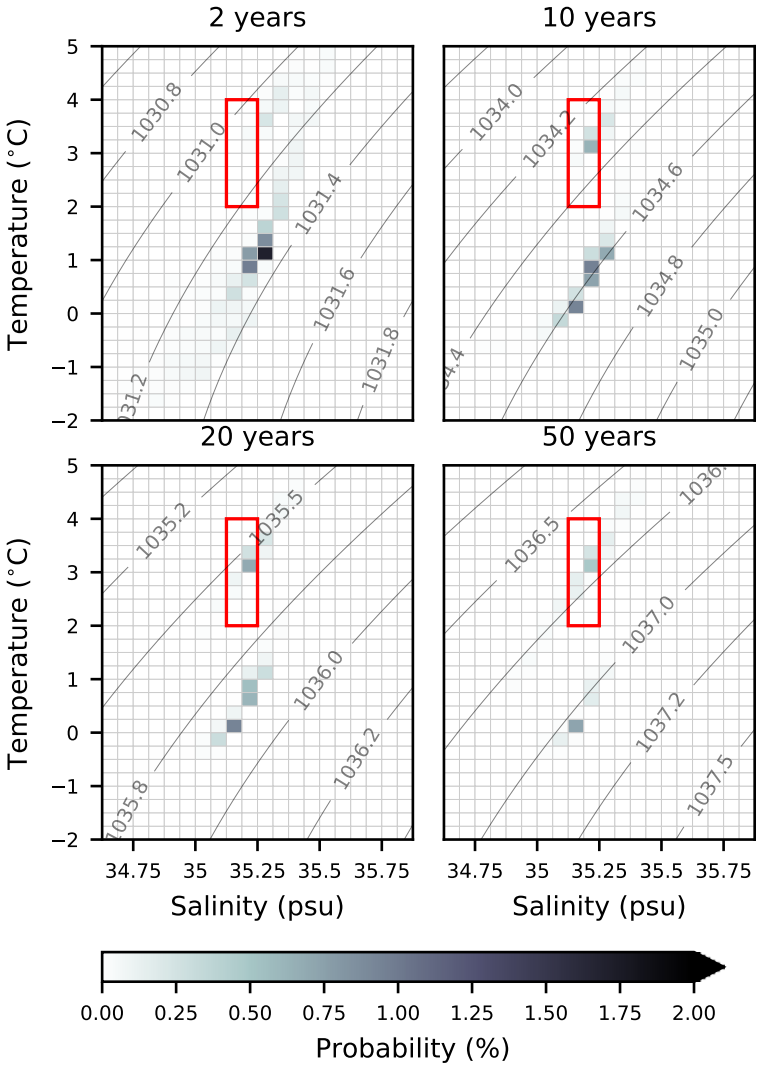


FIGURE 2.16: As in Fig. 2.12, but for Arctic-outcropping NADW. Note that the region enclosed by the red box is the same in both subpolar- (Figs. 2.12 and 2.14) and Arctic-outcropping cases.



manner, we find that the vast majority of North-Atlantic-destined A-NADW (88%) enters through the Denmark Strait.

Passive tracer injected at the A-NADW outcrop rapidly moves through TS space (Fig. 2.16. A colder, fresher water type splinters away from the surface-borne model NADW, and as a result, the tracer occupies two distinct regions in TS space for the remainder of the run. The warmer of these two water types shares most of its TS properties with those of the model NADW at the surface. On the other hand, the colder water type, at $(-0.5, 1)^\circ$ C, bears a similar temperature signature to observed OW (LeBel et al., 2008).

Our simulation shows that surface-borne A-NADW does not proceed to form a substantial part of subsurface NADW in the simulation. Within 0.6 yr, 78% of the tracer has been re-exposed to the atmosphere. Of that which remains, some 98% has

left the NADW TS class. The little tracer which stays in our defined NADW class is persistent, eventually following a similar trajectory to SP-NADW through TS space.

The two tangent-linear experiments (following water from each of the two distinct NADW outcrop regions) suggest that the more northeasterly outcrop contributes quantitatively little to the NADW bulk. Hence, any surface origins of NADW in the model's Arctic are likely found outside of the NADW TS class, and undergo transformation away from the surface.

2.4.3 Adjoint run

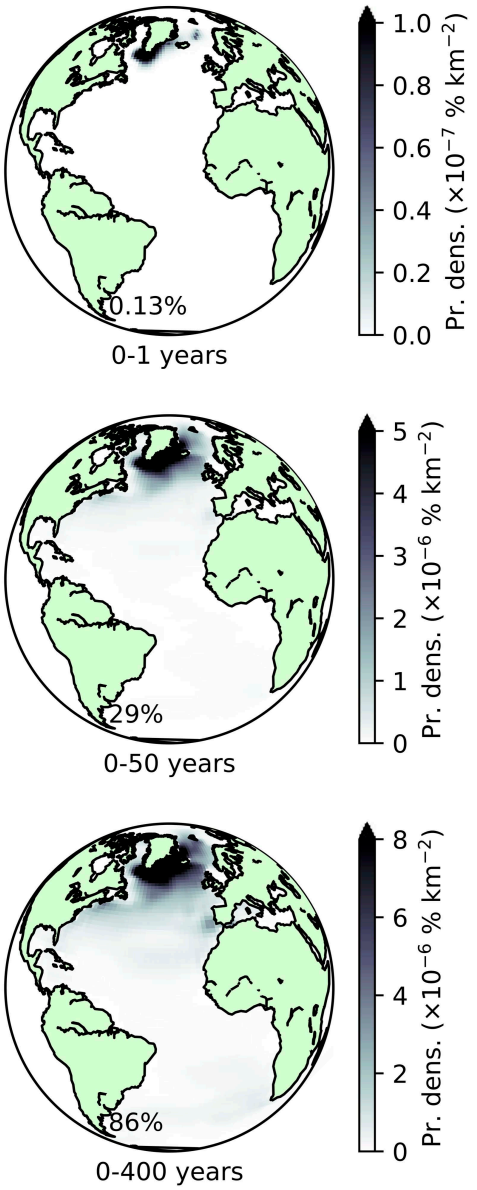
As before, we take a water mass budget at the end of the nonlinear model simulation (in this case for NADW after 400 model years) and provide it to the adjoint model. Using this approach, 86% of the model's NADW budget can be traced to its creation within 400 years (Fig. 2.17), with the rest remaining in the ocean interior.

During the first year, only 0.13% of the total volume of model NADW is tracked back to the surface. The strong presence of shallower NADW during this period leads to a clean signature of the two outcrops (Fig. 2.17, top panel). On multidecadal time scales, tracer from NADW can be traced back to locations spanning most of the northern North Atlantic (Fig. 2.17, centre panel), dominated still by the region surrounding its subpolar outcrop. Of particular interest on these time scales is the signature of the Mediterranean Outflow. It has been proposed from hydrographic data that the northward penetration of Mediterranean Water can intermittently reach the subpolar gyre, influencing LSW (Lozier and Stewart, 2008). However, we may only speculate as to whether this mechanism is present in our simulation.

Modelled NADW spanning the entire 400-year run (Fig. 2.17, bottom panel) has sources throughout the Atlantic basin, with a notable contribution from the eastern boundary of the South Atlantic, local to the Benguela current. The Labrador-Irminger sector remains the primary source of NADW at all considered time scales, however.

The propagated budget vector can be separated into different source regions and signatures. For instance, 31% of the ventilated NADW can be traced back to the model's Irminger Sea, versus just 14% to the Labrador Sea. We may also construct a volumetric census of model NADW source waters in thermohaline space (Fig. 2.18, upper left panel). This may further be partitioned into waters originating from the two climatic regions (i.e. subpolar and Arctic) associated with the distinct outcrops of NADW (Fig. 2.18, upper right and lower left panels, respectively). 45 % of the total NADW formed during the 400-year run may be attributed to the subpolar region. The dominant TS class associated with NADW formation in the model matches that of our definition ($[2, 4]^\circ$C, [34.9, 35.0] psu), though there are contributions from a cluster of subpolar water types in the range $[2, 10]^\circ$C, [34.5, 35.5] psu. Despite the substantial

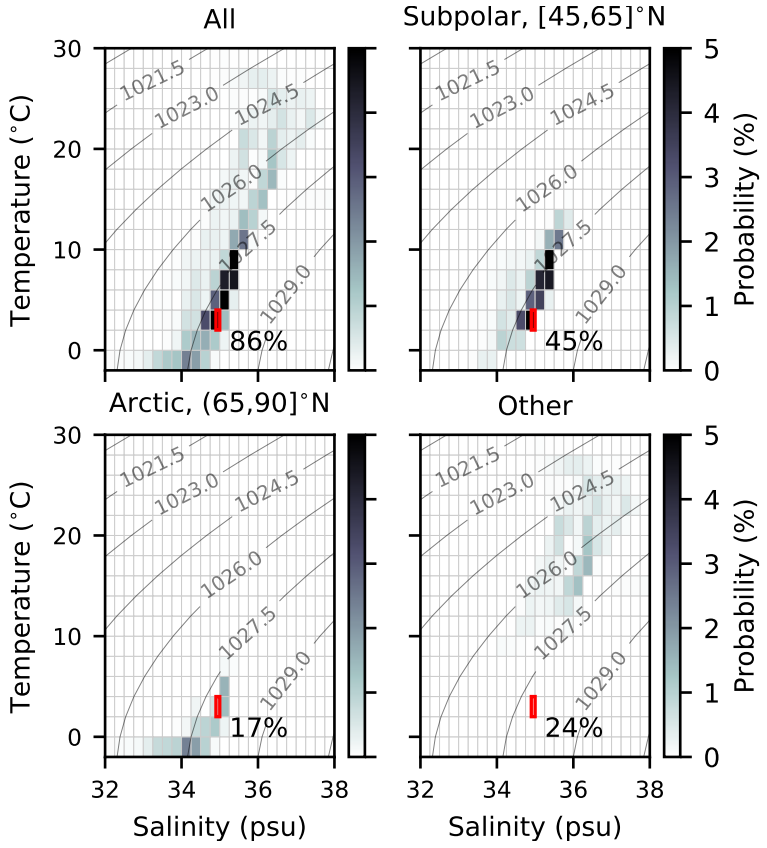
FIGURE 2.17: As in Fig. 2.8, but for NADW and time periods [0,1] yr (top), [0,50] yr (center) and [0,400] yr (bottom)



seasonal surface exposure of NADW in the Arctic (34% of the annual outcrop area, see Sect. 2.4.1), this region ultimately accounts for only 17% of modelled NADW formation in 400 years. This agrees with our findings in the tangent-linear model (Sect. 2.4.2) that surface-borne NADW in the Arctic is subject to rapid re-ventilation and does not contribute to the NADW bulk. It further suggests that there is no other narrowly defined TS band which outcrops in the Arctic from which modelled Arctic NADW originates. As such, NADW from this region generally derives from a broad range of waters colder and fresher than we define NADW to be.

NADW which does not originate from either of these two regions of the North Atlantic

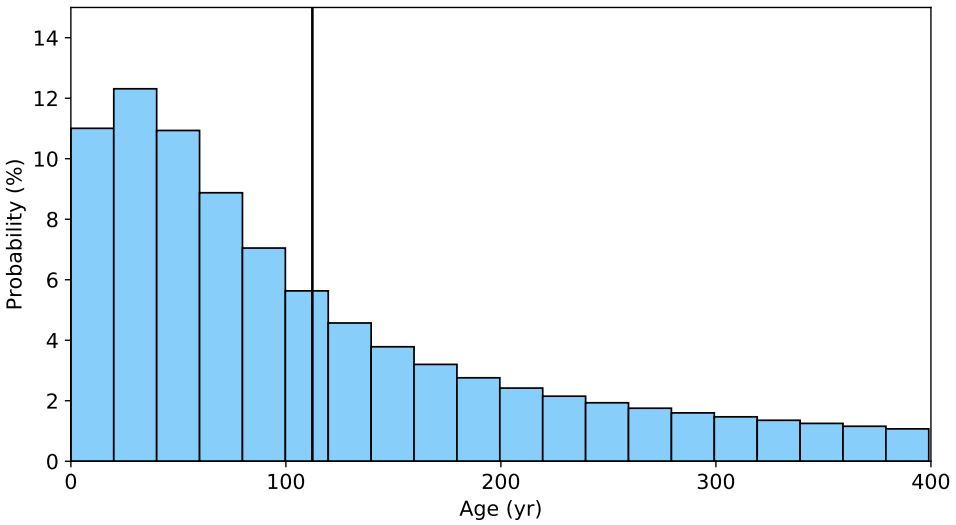
FIGURE 2.18: As in Fig. 2.9, but for NADW and time period [0,400] yr. Top left: NADW total volume. Top right: NADW volume traced to the subpolar North Atlantic. Bottom left: NADW volume traced to the Arctic. Bottom right: volume of NADW originating elsewhere.



makes up a substantial proportion (24%, Fig. 2.18, lower right panel) of the budget. Of this, 4% is formed at high latitudes in the Southern hemisphere. 5% of the total originates from outside the Atlantic.

As in Sect. 2.3, we may also consider the temporal distribution of simulated NADW formation (Fig. 2.19). Age peaks in the 13th year, during which 0.6% of the total budget can be traced back to the surface. The expected age is 112 yr. It should be noted that the mean age of model NADW is likely slightly higher. As with any water mass studied in this manner, a proportion will always remain in the ocean (closing the overall budget). Due to this, surface ventilation does not quite account for 100% of the total budget during our 400-year run. We limit our runs to this length due to technical issues with the HPC system used. In Sect. 2.3.3, we remarked on the unusual formation regions associated with the very oldest North Atlantic Subtropical Mode Water. This was associated with its much shorter life cycle. As such, the corresponding tail of the PDF of NADW age holds fewer surprises. Much like the youngest NADW considered here, the oldest NADW originates predominantly from winter convection in the model's Labrador and Irminger seas.

FIGURE 2.19: As in blue shading of Fig. 2.10, but for NADW. The black line is the expected age.



2.5 Discussion and conclusions

We have presented a newly developed addition to the NEMOTAM tangent-linear and adjoint modelling framework for the NEMO OGCM. This package allows tangent-linear and adjoint tracking of passive-tracer transport. Our framework is rooted in concentrations and probabilities, comparable to the statistical properties of a high-resolution Lagrangian approach with a large (theoretically infinite) number of particles. The development was achieved by deactivating dynamic feedbacks in the time-stepping routine of NEMOTAM, and embedding an alternative advection scheme suitable for passive-tracer transport. This advection scheme, proposed by Fiadeiro and Veronis (1977), is constructed as a linear combination of the upstream and centred schemes, so as to reduce spurious numerical diffusion. Performance tests show good scaling up to 128 cores at 2° resolution and up to at least 512 cores at 0.25° resolution, with much shorter run times (but higher memory demand) in the linear model with passive-tracer transport when compared with the fully-active nonlinear model.

We have exhibited the use of this tool in two case studies concerning the tracking of North Atlantic-borne water masses, North Atlantic Subtropical Mode Water and North Atlantic Deep Water. The method's versatility has been demonstrated through the calculation of several quantities pertaining to these water masses. If a sufficiently long trajectory is used, a near-complete statistical distribution of surface formation can be constructed. This allows diagnosis of expected age and expected origin location, as well as the rate of eradication and re-ventilation of newly ventilated waters.

The linearity of the method ensures that the water being tracked can be partitioned into several components. When these components are propagated separately, the tracking of the whole is equivalent to the tracking of their union.

The results of our case studies show good agreement between the model configuration and common aspects of prior observational and computational studies. We have shown, for example, that, in our simulation, on average, the expected surface origin of tracer initialised within NASMW is 32°N, 58°W (in line with Worthington, 1959; Warren, 1972), that its decay time is around two months (as in Fratantoni et al., 2013), and that a small minority (2.5% in four years) is exported to the subpolar gyre (concurrent with Burkholder and Lozier, 2011). We have estimated the average age of model NASMW as 4.5 yr, also close to observations (e.g. Kwon and Riser, 2004), and shown that a more persistent NASMW subset (with an expected age of over 8.5 yr) underlies the bulk of the water mass (as suggested by Davis et al., 2013).

For simulated NADW, we have shown (in the tangent-linear) that an Arctic outcrop of water with its signature to the northeast of the Greenland-Scotland ridge contributes little to its final form. However, we have found (by backtracking) that Arctic surface waters still make a contribution to simulated NADW formation (17% here), but from a broad range in thermohaline space. It is understood that OW is produced by the cooling and freshening of North Atlantic inflow to the Greenland Sea (e.g. Quadfasel and Käse, 2007; Dickson et al., 2002). The finer details of this transformation, and subsequent resupply of the transformed water into the North Atlantic are not well-captured by the method. The relative importance of pathways into the Atlantic for overflow water are still poorly known (Macrander et al., 2005), but we find using a broad transport estimate that the Denmark Strait is dominant in the model. For NADW formed in the subpolar outcrop of the model's Labrador Sea, our tracer spread reflects well that observed using CFCs (e.g. Rhein et al., 2002). The diffusive behaviour of the passive tracer also means that, despite the non-eddying nature of our model, eddy-driven southward export pathways are better represented than they would be by Lagrangian drifters (e.g. Gary et al., 2011).

Nevertheless, there are several more intricate details of North Atlantic water mass formation and transport which are the subject of ongoing investigation, and as such deserve a dedicated study at higher resolution. For example, the importance of fine-scale bathymetry for an accurate description of overflow-based NADW formation is well-described (Hansen et al., 2001; LeBel et al., 2008; Smethie Jr and Fine, 2001; Dickson et al., 2002). From a selection of models, Chang et al. (2009) find that the contribution of OW to NADW is misrepresented by models coarser than $\frac{1}{12}^\circ$ in their study. In such models, the Faroe Bank Channel is typically unresolved and the Denmark Strait Overflow resultantly dominates, as is the case here. Furthermore, although our tracer qualitatively reproduces southward export of LSW more effectively than a Lagrangian approach would in a laminar model, diffusion is still parameterised.

Similar reasoning applies to the subpolar export of NASMW, for which ocean turbulence is a significant contributor (Brambilla and Talley, 2006). It is unknown how well our parameterisation represents such processes. The extent to which resolved eddies would impact the age distribution presented here is less clear: the Lagrangian NASMW study of Kwon et al. (2015) in an eddy-resolving model concludes that it is the mean flow which sets the timescales of reventilation, for example. Nevertheless, the temporal distribution of water mass origins is likely impacted by the restoring scheme used to describe their ventilation, and a complementary investigation in a coupled model context would thus be valuable.

Despite good broad agreement between the passive-tracer pathways and those noted in previous studies of these water masses, there is an interesting disparity between the forward and backward modes within the TAM itself. This originates from using a TS-based description to inform the initial tracer distribution. For example, the backtracking method suggests that NASMW predominantly originates in slightly warmer surface waters than those of the outcrop used to inform the forward model. Meanwhile, A-NADW, while occupying over a third of the simulated NADW annual outcrop, ultimately contributes almost nothing to the subsurface water mass. These deviations highlight the approximation used by many water-mass-tracking model studies - thermohaline characteristics are not a purely passive tracer. Water parcels experience changes in their thermohaline properties, and so water in a particular TS class at depth is not exclusively related to the same TS class at the surface through a passive advection pathway.

TAM use at high resolution is typically limited, due to baroclinic instability. However, this is not detrimental for passive tracer tracking, due to lack of dynamic feedbacks. As such, our tool may be used in conjunction with higher-resolution configurations of NEMO (e.g. ORCA12, Treguier et al., 2017). The main barrier to higher resolution for users of our development is the necessity of frequent output and storage of trajectory snapshots from the nonlinear model, which are required for NEMOTAM operation (Vidard et al., 2015). Future versions of our tool will allow regional trajectory storage to overcome this barrier.

Although we have presented the development in the context of water mass tracking, there are many potential further applications. Ocean heat uptake pathways and carbon sequestration have been studied by means of modelled passive tracers (e.g. Banks and Gregory, 2006; Xie and Vallis, 2012), and adjoint models (e.g. Hill et al., 2004). A slight modification to ignore vertical velocities read from the trajectory would force positive buoyancy on the tracer. This could allow buoyant anthropogenic pollutants to be tracked, with potential application to ocean plastic tracking and oil spill drift prediction. Water mass tracking may itself be complemented by considering continuous (rather than instantaneous) inputs of tracer at the surface. It is hoped that

an off-the-shelf tool, bolted onto an existing OGCM will stimulate further research in these areas.

Despite TAM use for sensitivity analysis being traceable to the 1940s (Park and Xu, 2013), its application to ocean science is still in its infancy. The TAM approach is highly versatile, and its application to state-of-the-art OGCMs permits a great many new insights into ocean dynamics. Our development demonstrates the ability of a tangent-linear and adjoint model among the suite of existing water mass tracking methods. It is hoped that this novel tool will encourage new users to realise the potential of this powerful branch of ocean modelling.

Chapter 3

The active and passive roles of the ocean in generating basin-scale heat content variability

A version of this chapter is in review for publication in Geophysical Research Letters under the same title.

Abstract

The role of ocean circulation in transforming surface forcing into interannual-to-multidecadal oceanic variability is an area of ongoing debate. Here, a novel method, establishing exact causal links, is used to quantitatively determine the role of ocean active and passive processes in transforming stochastic surface forcing into heat content variability. To this end, we use a global ocean model in which the dynamical response to forcing can be switched on (fully active) or off (purely passive) and consider the resulting effect on heat content variance. While ocean passive processes mainly control the surface variance (over 92%) in all basins, most regions show the importance of active processes at depth. This role is particularly important for full-depth North Atlantic heat content, which we investigate further, highlighting signatures of the meridional overturning circulation.

Plain Language Summary

The ocean's role in climate is fundamental due to its ability to absorb significant amounts of heat relative to the other components of the Earth system. However, changes in heat can modify the ocean currents which transport it. The importance of this feedback effect remains uncertain, and so our study aims to determine how important this process is. We achieve this by alternately switching on and off the ability of simulated ocean currents to respond to changes in heat and salt driven by the atmosphere in a state-of-the art numerical simulation of the ocean. We then compare how variable the heat content of the ocean is in both "on" and "off" cases. We show that ocean circulation changes are unimportant near the surface, but in most regions they play a key role at depth. We look in detail at the North Atlantic, the region where circulation changes have the most important effect.

3.1 Introduction

It is well documented that the oceanic heat reservoir has a crucial role in climate; the ocean has absorbed over 90% of the excess energy associated with anthropogenic warming (Trenberth et al., 2014), for instance. However, this single number obscures the spatiotemporal heterogeneity of ocean heat content change, which is punctuated by hiatuses and surges (e.g. Meehl et al., 2011), geographically differential warming (e.g. Drijfhout et al., 2012), and varying impacts at different depths (e.g. Balmaseda et al., 2013). The mechanisms underlying these variations are in many cases elusive and remain challenging to disentangle due to the complexity of the climate system. This is particularly relevant on interannual-to-multidecadal timescales, where natural variability and external forcing have comparable amplitude (Meehl et al., 2009). Understanding these variations is thus crucial for modelling and predicting them.

The simplest explanation of heat content anomalies in the ocean is that they originate in the atmosphere, either via external forcing or natural, internal fluctuations, are fluxed into the mixed layer, and then passively circulate around the ocean interior along its preferred ventilation pathways. In this paradigm, the anomalous heat can be considered density compensated in that the ocean circulation does not change (e.g. Mauritzen et al., 2012). This approximation is often assumed when modelling the long-term response to anthropogenic forcing (e.g. Marshall et al., 2015; Zanna et al., 2019; Newsom et al., 2020), with anomalous heat fluxes represented by a passive tracer. However, investigations of the validity of this approximation for heat uptake typically flag the North Atlantic as a region to which it is particularly ill-suited (Banks and Gregory, 2006; Xie and Vallis, 2012; Garuba and Klinger, 2016, 2018), due to the Atlantic Meridional Overturning Circulation (AMOC) and its link with heat storage in models (Kostov et al., 2014). The involvement of the AMOC in natural,

interannual-to-multidecadal ocean temperature variations remains a contentious issue, however. Recent studies have argued that the predominant patterns of Atlantic Multidecadal Variability (AMV) in climate simulations featuring realistic ocean general circulation models (OGCMs) can be recreated by coupling a realistic atmosphere to a time-invariant "slab" ocean (Clement et al., 2015, 2016; Cane et al., 2017), suggesting these patterns are purely passive. In this slab ocean case, common features with fully active ocean simulations can only be established statistically. On the other hand, the previously discussed passive tracer approach, by propagating a passive "temperature" tracer initially coincident with the active temperature field in a single simulation and considering their divergence, provides a more thorough decomposition. Nevertheless, statistical slab–OGCM comparisons remain the de facto standard for determining the role of the ocean in near-term regional low-frequency variability (Dommenget and Latif, 2002; Dommenget, 2010; Wang and Dommenget, 2016; Delworth et al., 2017; Zhang, 2017).

In this study, we present an alternative approach to the question of regional heat content variability, using an adjoint model. Unlike a conventional model, which integrates anomalies forward in time, an adjoint model describes the sensitivity of a metric of interest (here heat content) to past changes (here stochastic atmospheric forcing), establishing causes, rather than effects (Errico, 1997). This has been leveraged to attribute the sources of temporal ocean variability in response to historical atmospheric forcing (Pillar et al., 2016; Smith and Heimbach, 2019) and establish the evolution of oceanic variance in response to representative stochastic atmospheric forcing (Sévellec et al., 2018).

We use this approach to isolate the role of the ocean in modelled heat content variability, by projecting a realistic stochastic representation of atmospheric buoyancy and momentum fluxes onto passive and active surface adjoint sensitivity fields. In the passive case, buoyancy anomalies cannot change the circulation.

3.2 Method and diagnostics

To characterise low-frequency ocean variability, Hasselmann (1976) and Frankignoul and Hasselmann (1977) developed an idealised, single-variable stochastic model of ocean surface temperature in response to random heat fluxes. These atmospheric fluxes can be seen as a continuous stream of small disturbances to ocean surface temperature, which accumulate and are slowly "forgotten". This can be represented as

$$u(t) = \int_0^t e^{-\lambda(t-\tau)} L \, \mathrm{d}W(\tau), \tag{3.1}$$

where $u(t)$ is the ocean temperature anomaly at time t ($u(0) = 0$ without loss of generality) and λ is the inverse damping timescale representing the ocean dynamics. W is a standard-normal Wiener process and L^2 describes the intensity of the stochastic fluxes (variance of their temperature impact per unit time).

Remarkably, this principle can be generalised to high-dimensional linear models, featuring multiple interacting variables and locations (represented by a single anomaly state vector, $|\boldsymbol{u}\rangle$) and more involved processes than simple exponential decay (representable by the propagator, $\boldsymbol{\Psi}$, of the ocean model). This reads:

$$|\boldsymbol{u}(t)\rangle = \int_0^t \boldsymbol{\Psi}(t,\tau)\mathbf{L}\,\mathrm{d}\,|\boldsymbol{W}(\tau)\rangle\,, \tag{3.2}$$

where $|\boldsymbol{W}\rangle$ is a vector of independent standard-normal Wiener processes and $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^\dagger$ is a covariance matrix (describing the stochastic flux intensity and spatial coherence). As before, $|\boldsymbol{u}(0)\rangle$ is assumed zero-valued.

From this formula, one can obtain the outcome of a metric of interest $\langle\mathbf{F}|\boldsymbol{u}\rangle$, such as heat content, in a fully active ($\boldsymbol{\Psi}_\mathrm{A}$) or purely passive ocean ($\boldsymbol{\Psi}_\mathrm{P}$) model. While the heat content variation in a fully active model is a classical problem of modern ocean physics, it is important to explicitly describe the routes by which its purely passive component can exhibit heat content variations. The first is the fluxing of heat content anomalies from the atmosphere which then propagate through the ocean by mean advection and diffusion. The second is the introduction of circulation anomalies by the wind. Although buoyancy anomalies cannot modify the circulation in the purely passive case, momentum fluxes may still create an anomalous circulation. This can then create heat content variations by redistributing the underlying mean ocean temperature field.

Given the metric of interest $\langle\mathbf{F}|\boldsymbol{u}\rangle$, one can also compute its variance from (3.2). The covariance (σ_AP) between outcomes in the two configurations (fully active and purely passive) of the model can similarly be calculated to determine their common components. Using the defining property of the adjoint $\langle\boldsymbol{a}|\mathbf{X}|\boldsymbol{b}\rangle = \langle\boldsymbol{b}|\mathbf{X}^\dagger|\boldsymbol{a}\rangle$ (where $|\boldsymbol{a}\rangle$ and $|\boldsymbol{b}\rangle$ are two state vectors, $\mathbf{X}$ and $\mathbf{X}^\dagger$ are an operator and its adjoint, and $\langle\boldsymbol{a}|\boldsymbol{b}\rangle$ is the Euclidean inner product) and following a multi-dimensional generalisation of Itô's isometry (e.g. Section 3.6 of Duan and Wang, 2014), the covariance at time t reads

$$\begin{aligned}
\sigma_\mathrm{AP}(t) &= \mathrm{Cov}(\langle\mathbf{F}|\boldsymbol{u}_\mathrm{A}(t)\rangle\,,\langle\mathbf{F}|\boldsymbol{u}_\mathrm{P}(t)\rangle) \\
&= \mathrm{E}\left[\langle\mathbf{F}|\int_0^t \boldsymbol{\Psi}_\mathrm{A}(t,\zeta)\mathbf{L}\,\mathrm{d}|\boldsymbol{W}(\zeta)\rangle\,\langle\mathbf{F}|\int_0^t \boldsymbol{\Psi}_\mathrm{P}(t,\gamma)\mathbf{L}\,\mathrm{d}|\boldsymbol{W}(\gamma)\rangle\right] \\
&= \int_0^t \langle\mathbf{F}|\boldsymbol{\Psi}_\mathrm{A}(t,\tau)\boldsymbol{\Sigma}\boldsymbol{\Psi}_\mathrm{P}^\dagger(\tau,t)|\mathbf{F}\rangle\;d\tau,
\end{aligned} \tag{3.3}$$

where σ_AP is the covariance between fully active and purely passive version of the model denoted by $\boldsymbol{\Psi}_\mathrm{A}$ and $\boldsymbol{\Psi}_\mathrm{P}$, respectively, and $\mathrm{E}[\cdot]$ is the expectation of a stochastic

Itô process. This leads to expressions for the variance of the fully active and purely passive component at time t:

$$\sigma^2_{\mathrm{A}}(t) = \mathrm{Var}(\langle \mathbf{F}|\boldsymbol{u}_{\mathrm{A}}(t)\rangle) = \int_0^t \langle \mathbf{F}|\boldsymbol{\Psi}_{\mathrm{A}}(t,\tau)\boldsymbol{\Sigma}\boldsymbol{\Psi}^{\dagger}_{\mathrm{A}}(\tau,t)|\mathbf{F}\rangle \; d\tau;$$
$$\sigma^2_{\mathrm{P}}(t) = \mathrm{Var}(\langle \mathbf{F}|\boldsymbol{u}_{\mathrm{P}}(t)\rangle) = \int_0^t \langle \mathbf{F}|\boldsymbol{\Psi}_{\mathrm{P}}(t,\tau)\boldsymbol{\Sigma}\boldsymbol{\Psi}^{\dagger}_{\mathrm{P}}(\tau,t)|\mathbf{F}\rangle \; d\tau. \tag{3.4}$$

These equations describe the level of variance of the ocean heat content obtained after a time t in response to stochastic forcing starting from rest in the fully active (σ^2_{A}) and purely passive (σ^2_{P}) cases. These will asymptotically tend towards their associated climatological heat content variance. The covariance describes how much of this variance is common to both, and can be normalised to give a Level of Agreement (LoA) between the purely passive and fully active cases, which we define as $\mathrm{LoA}(t) = \frac{\sigma_{\mathrm{AP}}(t)}{\sigma_{\mathrm{A}}(t)\sigma_{\mathrm{P}}(t)}$. If the LoA is unity at a given time, it is taken that anomalous heat content variation in the fully active ocean has been entirely controlled by purely passive processes.

These diagnostics have three requirements. Firstly, a linearised ocean general circulation model (OGCM) is needed to provide the propagator $\boldsymbol{\Psi}_{\mathrm{A}}$ and its adjoint $\boldsymbol{\Psi}^{\dagger}_{\mathrm{A}}$. Secondly, this propagator requires an isolated purely passive component $\boldsymbol{\Psi}_{\mathrm{P}}$ and its adjoint $\boldsymbol{\Psi}^{\dagger}_{\mathrm{P}}$. The model, its adjoint, and the purely passive configuration are described in Section 3.3.1. Lastly, we require a stochastic representation $\boldsymbol{\Sigma}$ of surface fluxes. We diagnose this from a coupled climate model (also described in Section 3.3.1). In particular, we assume that buoyancy and momentum flux anomalies from the coupled simulation climatology follow a band-limited (therefore finite power), spatially covarying Gaussian white noise. At each location, the power spectral density (PSD) of the flux anomalies is therefore assumed constant up to a few days, and zero at higher frequency. The cutoff is determined by the e-folding decorrelation timescales of the fluxes (Fig. 3.1, contours). We also have an implicit low-frequency limit imposed by the 20-year length of the coupled simulation. The elements of $\boldsymbol{\Sigma}$ are then given by the (effectively constant) PSD averaged over this band.

It is important to remark on linearity and independence, which allow for further decomposition of the above diagnostics. As the model propagators are linear, we can consider the fully active model $\boldsymbol{\Psi}_{\mathrm{A}}$ to be the sum of its purely passive model $\boldsymbol{\Psi}_{\mathrm{P}}$ and a dynamical-only component $\boldsymbol{\Psi}_{\mathrm{D}}$, encompassing just the feedback terms. Furthermore, the propagation of multiple metrics is equal to the propagation of their sum by linearity: $\boldsymbol{\Psi}^{\dagger}(|\mathbf{F_1}\rangle + |\mathbf{F_2}\rangle) = \boldsymbol{\Psi}^{\dagger}|\mathbf{F_1}\rangle + \boldsymbol{\Psi}^{\dagger}|\mathbf{F_2}\rangle$.

We additionally take surface buoyancy fluxes (described by $\boldsymbol{\Sigma}^{\mathrm{b}}$) and momentum fluxes (described by $\boldsymbol{\Sigma}^{\mathrm{m}}$) to be independent, and so the response to each can be determined separately, with $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\mathrm{b}} + \boldsymbol{\Sigma}^{\mathrm{m}}$. We emphasise that the covariance between the buoyancy components (heat and freshwater fluxes) and between the momentum

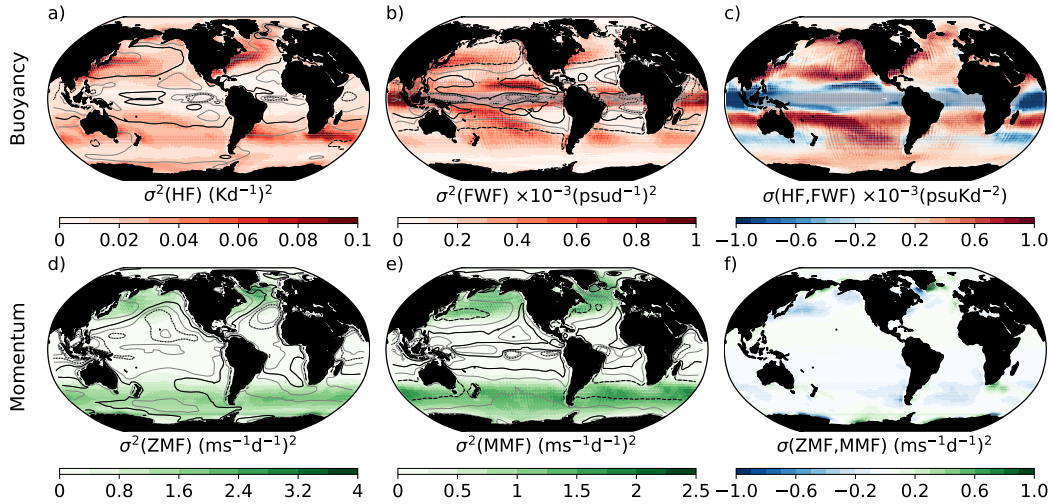


FIGURE 3.1: Local (co)variance (shading) and decorrelation time (λ^{-1}, contours) of surface fluxes in the coupled model. (a) Variance in rate of temperature change due to heat flux (HF). (b) Variance in rate of salinity change due to freshwater flux (FWF). (c) Covariance between rate of temperature and salinity change. (d and e) Variance in rate of zonal and meridional velocity change due to zonal and momentum fluxes (ZMF and MMF), respectively. (f) Covariance between rate of zonal and meridional velocity change. Thick dashed, solid, and dotted black contours indicate decorrelation time (λ^{-1}) of one, two, and three days, respectively. Thin grey contours are intermediate values, separated by half a day.

components (zonal and meridional fluxes) remain fully acknowledged. Using this to calculate σ_{AP} separately in response to buoyancy only and momentum only allows the LoA to be partitioned accordingly, by modifying its numerator while retaining the denominator. Finally, although the diagnostics of σ^2_{A}, σ^2_{P} and σ_{AP} are scalar values, they can be computed elementwise without summation, such that the contribution of each variable at each location to the total can be isolated. Similarly, the time integral can be decomposed to obtain the contribution of any time interval. This permits us to see the surface distribution and timing of sources leading to the resulting (scalar) heat content variance.

3.3 Application to an OGCM

3.3.1 Model description

Our stochastic representation is constructed from thermal, haline, and zonal and meridional momentum fluxes diagnosed from a coupled climate model (Fig. 3.1). Specifically, a twenty year simulation using the IPSL-CM5A-LR coupled model was run in its CMIP5 pre-industrial control configuration (c.f. Dufresne et al., 2013) with daily average output. The model was chosen as its ocean component is NEMO (v3.2)

with its ORCA2 global configuration (2° nominal resolution with 31 vertical levels), similarly to our linearised ocean model (described below). The atmospheric component is the LMDZ5a model, with 3.75°×1.9° horizontal resolution and 39 vertical levels (Hourdin et al., 2013).

The linear ocean model which we use to diagnose oceanic variability in the fully active case is NEMOTAM (Vidard et al., 2015), which is derived from NEMO v3.4 (Madec, 2012) and is used in its ORCA2-LIM configuration. The model configuration is similar to that detailed in Stephenson et al. (2020), which also discusses the implementation of the purely passive configuration in detail. The nonlinear model, which provides the simulation about which NEMOTAM is linearised, is forced by a single representative year (CORE normal year forcing; Large and Yeager, 2004).

3.3.2 Results

We now apply the derivations of Section 3.2 to attribute the generation of heat content variance in the fully active simulation to its different sources. We evaluate heat content to three depth horizons (10 m, 1500 m, and full-depth) which effectively correspond to sea surface temperature, heat content in the upper ocean, and the total heat content, respectively. We also consider both the global ocean and a seven-region partition of it (Fig. 3.2, black lines). These regions are the Arctic Ocean (>70°N), the North ([35,70]°N) and intertropical ([−35,35)°N) Atlantic and Pacific, the Indian Ocean (> −35°N), and the Southern Ocean (< −35°N).

Our analysis reveals that the Level of Agreement between purely passive and fully active heat content variance after 60 years varies significantly depending on the depth extent and geographical region (Fig. 3.2, bars). The LoA is extremely high for sea surface temperature variance in all regions. This ranges from 92.0% in the intertropical Pacific to 99.5% in the Southern Ocean, with a majority stimulated by buoyancy forcing. This implies that the purely passive uptake of heat controls temperature variability at the surface. There is a dramatic reduction in agreement when heat content is computed over a thicker layer. For the upper-1500 m heat content, variance common to both the purely passive and fully active simulations accounts for as little as 30.9% in the case of the Indian Ocean, and just over half (52.0%) globally. The nature of stimulation of the purely passive component also changes over this depth-range, shifting to a primarily wind-driven regime for all regions except the Arctic Ocean.

When heat content is defined over the full depth, it generally follows similar patterns to upper-1500 m heat content, with notable exceptions in the North Atlantic and Arctic oceans. For those basins, another dramatic reduction in correspondence between the purely passive and fully active simulations occurs, with the LoA reducing to 27.3% and 25.0%, respectively. More subtle reductions can be seen elsewhere, and

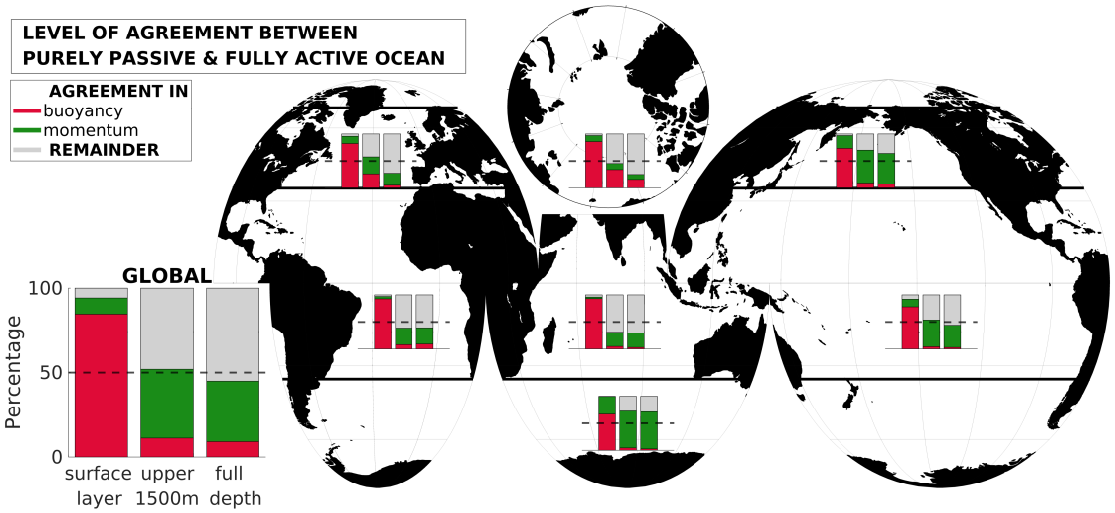


FIGURE 3.2: Level of Agreement (LoA) between purely passive and fully active simulations in generating the accumulated final (60-year) heat content variance due to buoyancy (red) and momentum (green) surface stochastic fluxes, determined by calculating σ_{AP} in response to each. LoA is shown for the three cases (surface layer – corresponding to SST, upper 1500 m, and full-depth heat content). Largest bar plot shows the case for the total global ocean heat content variance, smaller inner plots show regional values. Thinner dashed black lines signify a LoA of 50%. Black solid lines on the map mark the boundaries of the regions in our definitions.

only in the North Pacific and Southern Ocean does the purely passive component still dominate the fully active simulation at full depth. It is worth noting the substantial impact (>50%) of purely passive wind effects in these regions.

While the LoA provides a useful quantification of the ultimate role of the purely passive component of the ocean, it does not describe in detail the differences between the purely passive and fully active simulations (e.g. the timing of the variance growth or its source location). To tackle this question, we consider the time-evolving variance growth for each, along with its components (Fig. 3.3). We focus on the full-depth case, where the differences between these components are greatest. Similar decompositions have been considered for surface (Supp. Fig. 3.5) and upper ocean (Supp. Fig. 3.6) cases, and exhibit similar (but less significant) behaviour.

The temporal evolution of the variability reveals that the purely passive and fully active simulations differ in both magnitude and timing. As discussed in Section 3.2, linearity permits the decomposition of the fully active model into the sum of the purely passive component and a remaining dynamical-only component. The difference between evolving variance in the fully active model and the purely passive model (Fig. 3.3, solid and dotted lines, respectively) can thus be attributed to internal ocean feedbacks within this dynamical-only component, which are not always constructive. Indeed, the variance in the fully active simulation is often weaker than that of its purely passive counterpart. This suggests that certain behaviour is possible only in the purely passive case, and is cancelled out by the dynamical-only term in the fully active

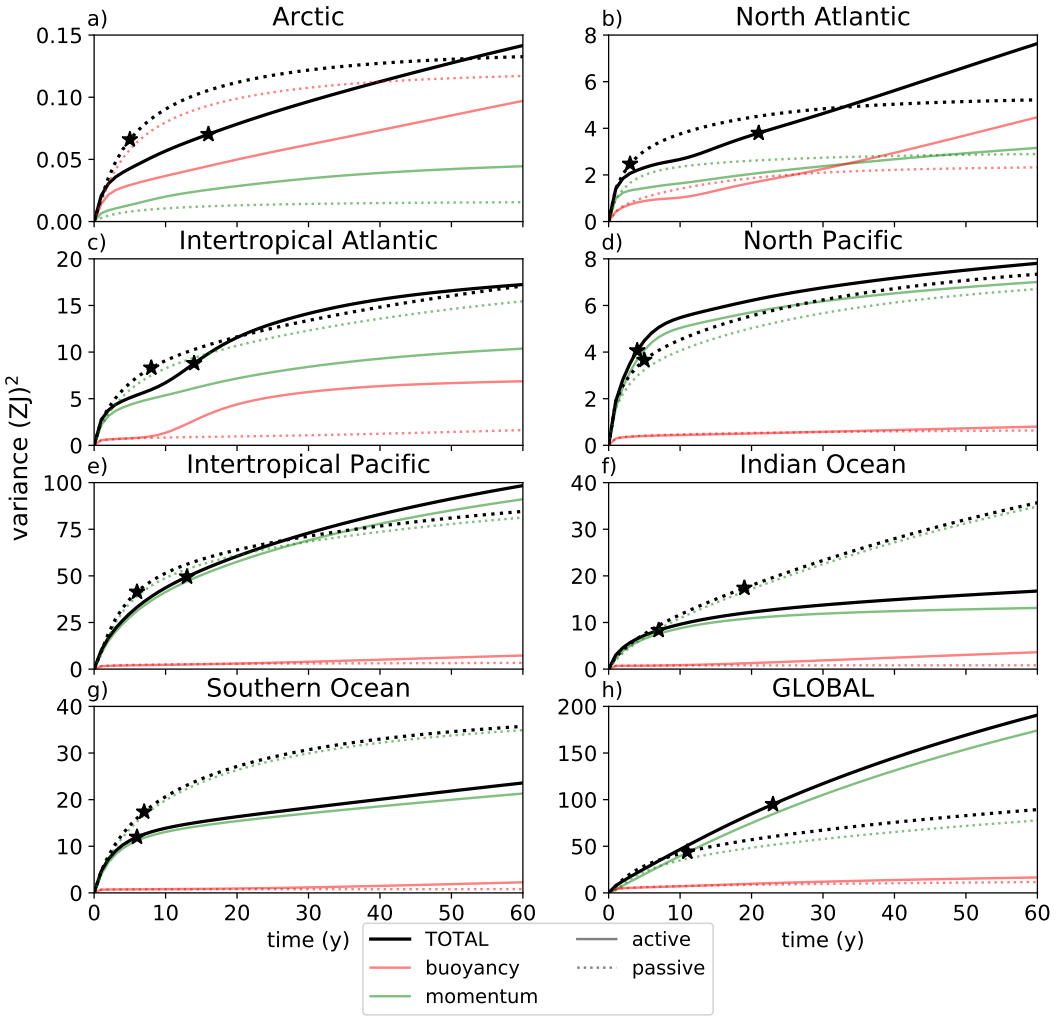


FIGURE 3.3: Evolution of full-depth heat content variance in response to stochastic surface forcing in the purely passive (dotted lines) and fully active (solid lines) simulations. The difference between these lines is linked to the dynamical-only component, which may act destructively (passive>active) or constructively (active>passive). Thinner lines show separately the buoyancy-forced (red) and wind-driven (green) components. Stars mark the point at which 50% of the final (60-year) variance is reached.

simulation. This is particularly visible for heat content variance in the Indian and Southern Oceans (dominated by wind stress). There, after two decades, most of the variance growth of the purely passive component stimulated by wind stress is cancelled by the dynamical-only component. A possible example of such behaviour is provided by Cronin and Tozuka (2016), who demonstrate that Ekman transport is determined not purely by wind stress and latitude (as in the classical analysis of Ekman, 1905), but also local geostrophic shear. In this perspective, Ekman transport has both a purely passive and dynamical-only component, which can act against each other. Conversely, for global ocean heat content (Fig. 3.3h), the passive variance is much smaller than the active, and smaller than that suggested by the individual subdomains, a disparity particularly clear in the wind-forced component. This arises from the disappearance of mid-ocean boundaries (Fig 3.2, black lines) when all of the

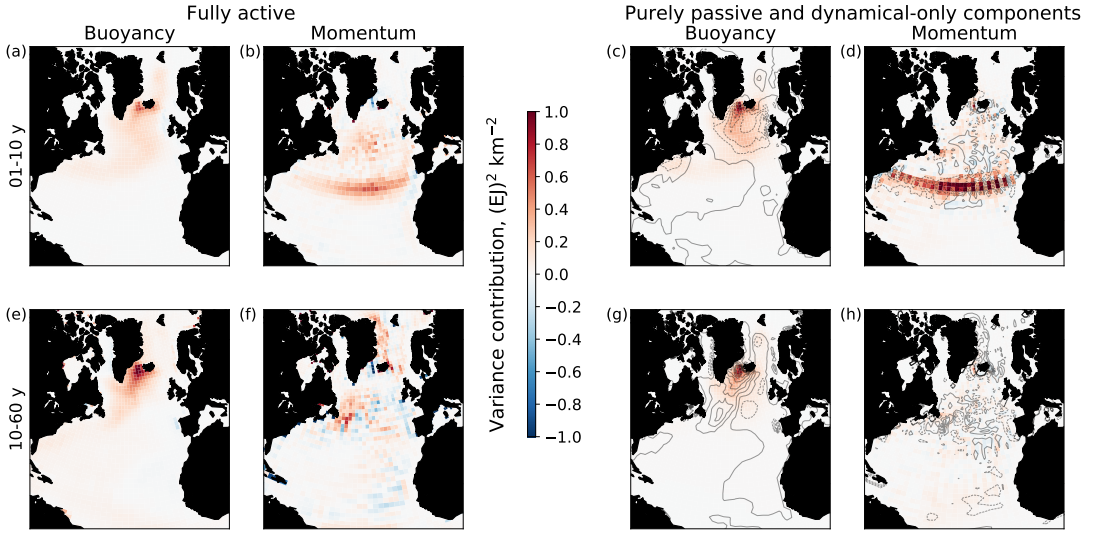


FIGURE 3.4: Surface sources of (a,c,e,g) buoyancy- and (b,d,f,h) wind-stimulated full-depth heat content variance in the North Atlantic, integrated over years 0-10 (upper panels) and 10-60 (lower panels) in the (a,b,e,f) fully active and (shading in c,d,g,h) purely passive simulations, and (contours in c,d,g,h) in the dynamical-only diagnosed component. The global integrals of the fully active and purely passive fields for the upper panels produce the values shown in Fig. 3.3 at 10 years. With the addition of the global integral of the fields from the lower panels, it provides the values shown in Fig. 3.3 at 60 years. The dynamical-only component is defined as the difference between the fully active and purely passive simulations. Solid and dashed contours indicate positive and negative values, respectively, with contour intervals of 0.05 (EJ)^2 km^{-2} for buoyancy, and of 0.2 (EJ)^2 km^{-2} for momentum.

subdomain regions are unified, such that variability can no longer be stimulated by Ekman transport across these boundaries.

As a measure of the rate at which the climatological variance is approached, we consider the time taken for the full-depth variance in each simulation to reach half of its final (60-year) value (Fig. 3.3, stars). Following on from the previous discussion, the dynamical-only momentum component in the Indian Ocean acts to accelerate variance evolution, with $\sigma^2_{\mathrm{A}}(t)$ reaching $0.5\sigma^2_{\mathrm{A}}(60\text{ years})$ in 19 years in the purely passive simulation, as opposed to only 7 years in the fully active simulation. At the opposite extreme, for the Arctic and North Atlantic, the dynamical-only contribution slows the variance evolution substantially. Indeed, in the North Atlantic, half of the final value is reached in only 3 years in the purely passive simulation, compared with 21 years in the fully active case.

The source of this continued growth in the active North Atlantic, even after the purely passive component appears to have saturated, corresponds to a regime change of the fully active simulation in its response to buoyancy stimulation after 10 years. To determine the origin of this, we consider separately the surface distribution of the variance accumulated during the first 10 years (Fig. 3.4a,b,c,d) and from 10 to

60 years (Fig. 3.4e,f,g,h). This is determined from the elementwise computation of the variance, prior to summation, as outlined in Section 3.2.

In the first decade, the passive and active simulations maintain a high Level of Agreement (above 75%) and their spatial patterns are similar. Focusing on buoyancy forcing, the relatively focused region reflects the model's deep water formation site, as described in the passive tracer study of Stephenson et al. (2020). The difference between the fully active and purely passive distributions (contours) is the dynamical-only contribution. This corresponds to a large-scale dipole. The negative peak of the dipole overlies the positive contribution by the purely passive component, having a slight compensating effect (Fig. 3.4c). On decadal timescales, positive contributions to variance growth in both the purely passive and dynamical-only components coincide in location, and so the two components act constructively (Fig. 3.4e,g).

The primary difference between stimulation by wind for the fully active and purely passive components in the first decade is the intensity of the induced variance (Fig. 3.4b,d). Both components are dominated by Ekman transport across a zonal band defining the region's boundary (35°N), but the addition of the dynamical-only component reduces the intensity of this pattern. Also notable in the fully active case is a seemingly persistent (Fig. 3.4b,f) stimulation of variance at the subtropical-subpolar gyre interface, as well as stimulation (both positive and negative) in coastal regions of the eastern North Atlantic and Greenland Sea.

3.4 Discussion and conclusions

We have considered the stimulation of variance in ocean heat content by surface atmospheric noise. We evaluated heat content over a range of different regions and depths in a linearised global ocean model, comparing purely passive and fully active realisations of the ocean model. In the purely passive framework, temperature anomalies either arise due to random surface heat fluxes (and can be passively transported by the mean flow), or due to random surface momentum fluxes (which redistribute existing heat). However, these resulting temperature anomalies are unable to modify the ocean circulation.

In contrast to the established techniques of using a passive tracer (e.g. Banks and Gregory, 2006; Xie and Vallis, 2012; Marshall et al., 2015; Garuba and Klinger, 2016, 2018) or a slab ocean model (e.g. Dommenget and Latif, 2002; Dommenget, 2010; Clement et al., 2015; Wang and Dommenget, 2016) to investigate the role of the ocean, we have utilised a novel adjoint-based approach (Sévellec et al., 2018). The use of an adjoint model has uniquely allowed us to causally attribute heat content variance to different variables, times, and locations at the surface, by projecting onto surface

sensitivity fields a realistic stochastic representation of atmospheric fluxes diagnosed from a coupled climate model.

Our findings for the surface ocean (i.e. sea surface temperature) are that at least 92% of the variance in the fully active simulation is in agreement with its purely passive component. This is consistent with studies which suggest that oceanic dynamics are not needed to generate surface decadal variability (e.g. Clement et al., 2015, 2016; Cane et al., 2017). However, while variance patterns in both simulations may express a high (normalised) Level of Agreement, a purely passive model could greatly over-estimate the amplitude of the variance, as the purely passive component can be partially compensated by the corresponding dynamical-only component in a fully active ocean.

The dynamical redistribution of existing heat by currents arising from buoyancy anomalies has been shown in past studies to substantially impact heat uptake (e.g. Banks and Gregory, 2006; Xie and Vallis, 2012), particularly in the North Atlantic. However, we have shown that the passive redistribution of the existing heat reservoir by wind anomalies is often more important in the context of heat content variability, leading to a driving role for the passive component over several regions and depths. Nevertheless, the deep North Atlantic also stands out here as a region with an important role for ocean feedbacks, with the dynamical-only component acting to slow the growth of heat content variance. We considered the time taken to reach 50% of the variance at the end of the simulation, and found that the fully active model takes 7 times longer (21 years) to reach this point than the purely passive simulation (3 years) in this region. This has potential consequences for climate predictability, as the variance growth can also be seen as the accumulation of error following model initialisation (Sévellec et al., 2018). The time taken to reach half of the climatological variability is often taken as a measure of the upper limit of predictability, beyond which noise dominates the predictable signal (e.g. Griffies and Bryan, 1997; Grötzner et al., 1999). The reason for this delay in the fully active North Atlantic is a regime shift in the response to buoyancy forcing. On sub-decadal timescales, the dynamical-only component slows variance growth, before sustaining it on timescales greater than ten years, resulting in an "S"-shaped growth curve. In exploring the spatial distribution of the components of the fully active simulation, we have observed a basin-scale dipole pattern in the North Atlantic. These patterns echo earlier sensitivity studies of the region in predecessors of our model (e.g. Sévellec and Fedorov, 2017). These studies relate North Atlantic heat content sensitivity to an ocean-only mode of variability in which heat content and AMOC anomalies feed back on each other via basin-scale thermal Rossby wave propagation (Sévellec and Fedorov, 2013a) consistently with observations of the AMV (Frankcombe et al., 2009).

There are a number of considerations which are not accounted for in our approach. Firstly, our conclusions are likely oversimplified by our use of atmospheric variability sources alone in a linear, laminar model. In a recent ensemble study at eddy-permitting

resolution, Sérazin et al. (2017) suggested that a substantial portion of ocean heat content variability is intrinsic, generated by chaotic, nonlinear processes within the ocean. This suggests that we underestimate the role of the dynamical-only component by restricting it to large-scale, laminar feedbacks. This will be addressed in a separate study. In addition, the role of coupling in the stimulation of interdecadal variability is an entire field of research on its own (cf. the review of Liu, 2012). Here, our model uses an uncoupled ocean and a stochastic representation of the atmosphere. This limits the conclusions of our work, in particular for sea surface temperature (where the surface boundary conditions have more impact). Furthermore, our stochastic representation is of limited bandwidth, effectively averaging the power spectrum of a two-decade coupled simulation. The result is a stationary (although globally coherent) white noise representation of daily-to-bidecadal atmospheric variability. We emphasise, however, that these simplifications have allowed us to use an adjoint ocean model to causally attribute the surface sources of heat content variability exactly, and with limited computational expense, an approach which offers several unique advantages of its own.

Supplementary Figures

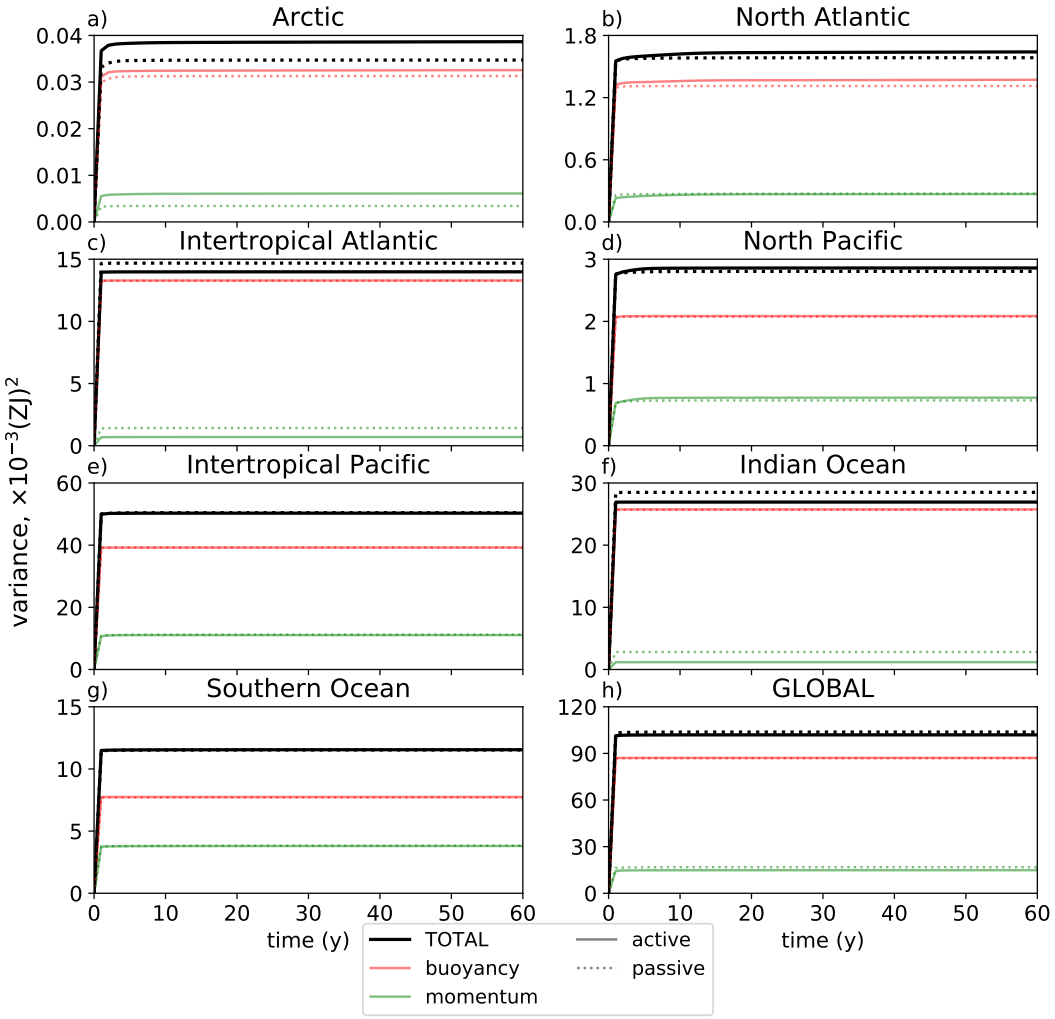


FIGURE 3.5: As in Fig. 3.3, but for surface-layer (0-10 m) ocean heat content

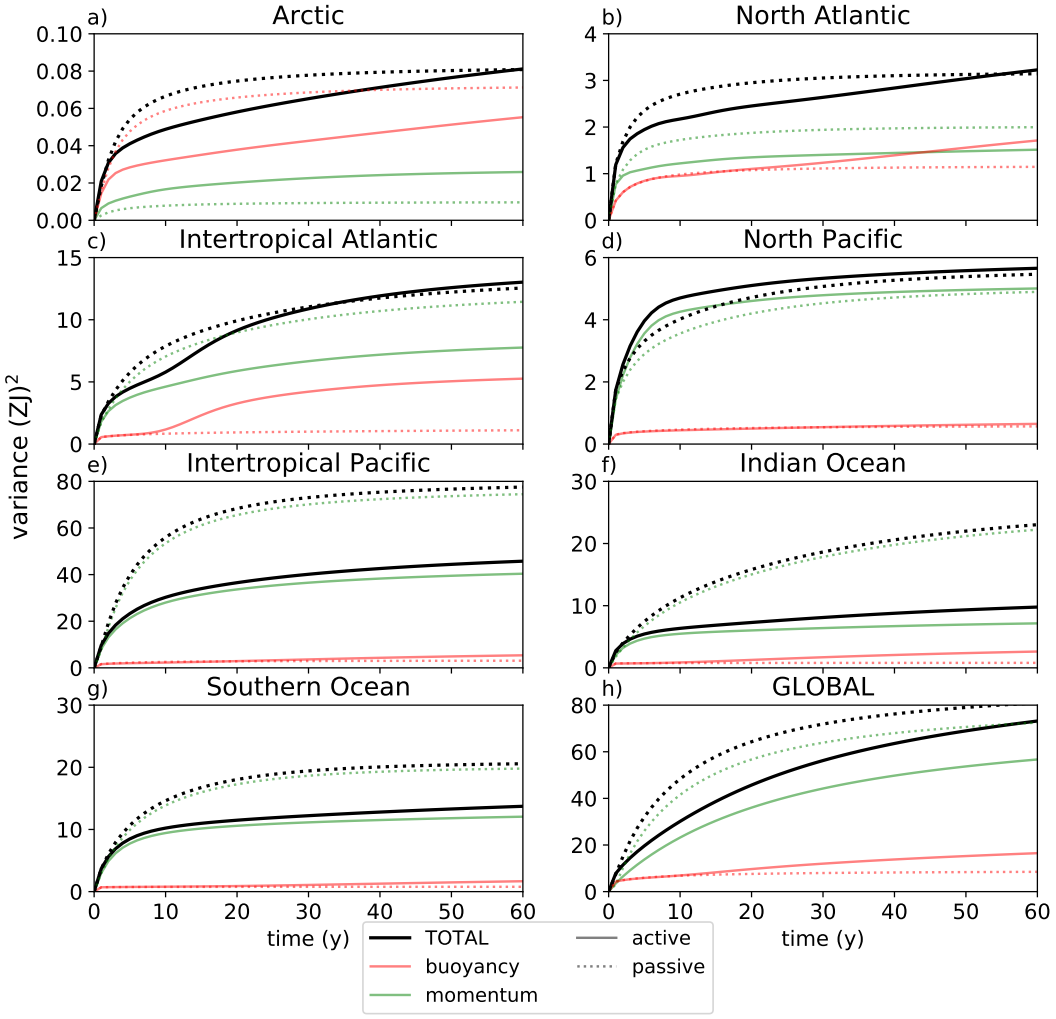


FIGURE 3.6: As in Fig. 3.3, but for upper-ocean (0-1500 m) ocean heat content

Chapter 4

Dynamical attribution of North Atlantic interdecadal predictability to oceanic and atmospheric turbulence

under realistic and optimal stochastic forcing

A version of this chapter is in review for publication in Journal of Climate under the same title.

Abstract

Unpredictable variations in the ocean originate from both external atmospheric forcing and chaotic processes internal to the ocean itself, and are a crucial sink of predictability on interdecadal timescales. In a global ocean model, we present i.) an optimisation framework to compute the most efficient noise patterns to generate uncertainty and ii.) a uniquely inexpensive, dynamical method for attributing sources of ocean uncertainty to internal (mesoscale eddy turbulence) and external (atmospheric) origins, sidestepping the more typical ensemble approach. These two methods are then applied to a range of metrics (heat content, volume transport, and heat transport) and time averages (monthly, yearly, and decadal) in the subtropical and subpolar North Atlantic. We demonstrate that optimal noise patterns target features of the underlying circulation such as the North Atlantic Current and deep water formation regions. We then show that noise forcing in the actual climate system stimulates these patterns with various degrees of efficiency, ultimately leading to the growth of error. We reaffirm the established notion that higher frequency variations are primarily wind driven, while surface buoyancy forcing is the ultimately dominant source of uncertainty at lower frequencies. For year-averaged quantities in the subtropics, it is mesoscale eddies which contribute the most to ocean error, accounting

for up to 60% after 60 years of growth in the case of volume transport at 25°N. The impact of eddies is greatly reduced in the subpolar region, which we suggest may be explained by overall lower sensitivity to small-scale noise there.

Plain language summary

Climate does not change steadily; it naturally fluctuates around a general trend. The prediction of climate several decades to a century ahead depends mostly on the ability to anticipate future human activity, but for the coming years to a few decades ahead (when the future pathway of human activity is not yet fully apparent) natural fluctuations also have an important role. These fluctuations, however, cannot be perfectly predicted for long. The ability to predict them is limited, for example, by the build-up of unwelcome "noise" from erratic processes such as the weather. In this study, we look at the different sources of this noise, how important they are, and how they impact prediction accuracy of climatically important ocean quantities decades in the future. To achieve this, we use a unique computer simulation of the ocean, which works backwards and describes how to most effectively create change. This uncovers the mechanisms by which noise is most effectively amplified by the ocean, and also shows how this compares with the behavior of noise in the real ocean-atmosphere system. We demonstrate that in the climatically important region of the North Atlantic, unpredictable ocean circulation changes in the more southerly tropical region are mostly due to oceanic mesoscale eddies (the oceanic equivalent of atmospheric storms). Further north, however, it is the atmosphere which is primarily responsible for the development of oceanic prediction error.

4.1 Introduction

As the slow component of the climate system, the ocean is key to predicting variations on timescales of seasons or longer. However, the ocean is now known to exhibit substantial variability at all timescales. The predictability of these variations, and their attribution to different sources, is crucial to the understanding and prediction of climate, particularly on so-called "near-term" timescales on which the anthropogenically forced signal is not yet dominant (Meehl et al., 2009).

Variations in the North Atlantic have long been hypothesised to be uniquely predictable due to interactions between its meridional overturning circulation (MOC) and anomalies in upper ocean heat content. In the late 1990s, an increase in computational resources allowed this hypothesis to be tested in state-of-the-art climate models using the prognostic technique of ensemble modelling (e.g. the review of Latif and Keenlyside, 2011). In this framework, each member of a coupled climate model

ensemble is initialised with a slightly perturbed atmospheric state. As the atmosphere has no predictability beyond a few weeks (Lorenz, 1969), the atmospheric components of the ensemble rapidly diverge such that their differences are indistinguishable from stochastic noise. The rate of divergence of the ocean components in response thus quantifies ocean predictability. Early studies using this methodology revealed enhanced predictability, often up to decades, in the North Atlantic sector against a background of strong MOC influence (Griffies and Bryan, 1997; Grötzner et al., 1999; Collins and Sinha, 2003; Msadek et al., 2010; Persechino et al., 2013). The implication that large-scale ocean dynamics slow error growth forced by the atmosphere is promising for near-term prediction in the region, but these studies collectively fail to account for oceanic mesoscale turbulence as an additional source of uncertainty. As ocean components of cutting-edge climate models evolve towards eddying resolution (Haarsma et al., 2016), the relative importance of this source is becoming increasingly scrutinised.

A new generation of studies is now addressing the question of attributing oceanic variability to internal (generated by chaotic oceanic processes) and external (atmospherically forced) origins using the prognostic ensemble approach in high-resolution ocean-only models (e.g. Sérazin et al., 2017; Leroux et al., 2018; Jamet et al., 2019). Each member has a common atmospheric forcing, but differing oceanic initial conditions. As such, the ensemble mean is taken to smooth out any intrinsic oceanic variability, such that its temporal variability is assumed to derive purely from fluctuations in the forcing. Contrarily, the ensemble spread, given their common atmospheric forcing, is assumed to come solely from intrinsic oceanic differences.

In this manner, Sérazin et al. (2017) conclude that ocean intrinsic variability is the dominant contributor to deep-ocean heat content fluctuations in the North Atlantic subtropical gyre and Gulf Stream regions, while Leroux et al. (2018) estimate that intrinsic MOC variability is 60% that of atmospheric at 26°N. In a regional model, Jamet et al. (2019) find that over half of the variability in the annually averaged Atlantic MOC at this latitude is intrinsic. Although oceanic variability forced at the domain boundaries will appear "external" in a regional model, this result agrees closely with the global model results of Grégorio et al. (2015). All studies show a shift in behaviour at subpolar latitudes, where the atmospheric component dominates.

Despite the revolutionary advances in computing which now allow studies such as these to utilise ensembles containing as many as 50 members in a global, eddy-permitting ocean (as in Leroux et al., 2018), such investigations are still prohibitively expensive for routine research. Furthermore, the ensemble approach does not allow a causal description of the translation of internal and external sources of unpredictable variability into expressed oceanic error growth or prediction uncertainty. An alternative framework, allowing dynamical attribution of the large-scale oceanic response to small perturbations (such as those from atmospheric fluxes or the

mesoscale eddy field) is the adjoint method (Errico, 1997). While the ensemble approach begins by applying small changes and then evaluates their impact on oceanic metrics of interest, the adjoint method turns the problem inside out: it begins with an oceanic metric of interest and then describes its sensitivity to small changes.

This method has been applied to attributing Atlantic MOC fluctuations to different surface fluxes in the MITgcm by Pillar et al. (2016), and was used in the OPA model (the oceanic component of the model used herein) by Sévellec et al. (2018) to determine the relative impacts of atmospheric and initial condition uncertainty on the divergence of a theoretical ocean ensemble.

This study builds further on the theoretical ensemble approach of Sévellec et al. (2018). Here, we explore ocean error growth from two perspectives. In the first, we use an adjoint model to determine the most efficient patterns for stimulating ensemble divergence (the optimal stochastic perturbations, or OSPs, Sévellec et al., 2007). In this framework, the model is blind to actual, "real world" sources of chaotic variability, and instead describes how these sources should look in order to have the greatest effect on oceanic uncertainty. In this sense, the outcome describes, for different metrics, the sensitivity of their variance to different sources and locations, highlighting oceanic patterns of efficient error growth.

In the second perspective, we provide the model with realistic, stochastic representations of real-world internal and external turbulent variability sources. This allows us to dynamically attribute ocean uncertainty to these different sources. The realistic sources are diagnosed from more complex models; the external, atmospheric component is calculated from a coupled non-eddying climate model, while the internal, mesoscale-eddy-driven component is calculated from an eddy-permitting ocean model. The attribution method is uniquely inexpensive – a single bidecadal simulation of a coupled climate model and an eddy-permitting ocean model are used to compute the stochastic properties, while the highly efficient adjoint ocean model in a non-eddying (laminar) configuration can recreate a theoretically infinite ensemble with a single simulation (Sévellec and Sinha, 2018).

The study proceeds as follows. In Section 4.2, we outline the mathematical theory of stochastically forced ensembles which underlies our two approaches. This begins with a treatment of the classical, temporally uncorrelated ("white noise") case, which provides the theoretical framework for deriving the OSPs. We then advance to time-correlated stochastic noise, more appropriate for creating a representation of realistic turbulence in the case of oceanic mesoscale eddies. In Section 4.3, we describe how this time-correlated representation is diagnosed, along with the three models used for the study and the configuration of our experiments. Our results are presented for both the optimal and diagnosed forcing cases in Section 4.4 before being discussed along with our conclusions in Section 4.5.

4.2 Theoretical framework: variance of stochastically forced linear systems

4.2.1 Temporally uncorrelated forcing

One of the simplest models of low-frequency variability generation in the ocean is that of Hasselmann (1976). In it, mixed layer temperature changes are assumed to be a purely passive response to random, serially uncorrelated surface heat fluxes. These are absorbed and slowly "forgotten" by the ocean, which tends back toward its unperturbed state. The model is univariate and entirely determined by two parameters: the timescale on which this restoring occurs (parameterising the ocean dynamics as a single memory term) and the volatility of the random fluxes (parameterising the atmospheric forcing). It may be written as the stochastic differential equation

$$\mathrm{d}u = -\lambda u \,\mathrm{d}t + L \,\mathrm{d}W, \tag{4.1}$$

which has solution (for initial condition zero)

$$u(t_0, t_1) = \int_{t_0}^{t_1} e^{-\lambda(t_1 - t)} L \,\mathrm{d}W(t), \tag{4.2}$$

where u is the surface temperature, t_0 and t_1 are the initial and final time, L^2 is the variance of temperature change induced by random surface atmospheric heat fluxes during a time increment $\mathrm{d}t$, λ^{-1} defines the e-folding timescale of the ocean dynamics (i.e. its memory), and $\mathrm{d}W$ is an increment of a standard-normal Wiener process W (akin to the distance of a random walk during the time increment $\mathrm{d}t$). (4.2) is thus an Itô integral (Itô, 1944). It may be noted that the response is an Ornstein-Uhlenbeck process (Uhlenbeck and Ornstein, 1930), such that variability generation follows the autocovariance function:

$$\mathrm{Cov}(u(t_0, t_1), u(t_0, t_2)) = \frac{L^2}{2\lambda} \left(e^{-\lambda |t_2 - t_1|} - e^{-\lambda (t_2 + t_1 - 2t_0)} \right). \tag{4.3}$$

This autocovariance function is weakly stationary in the limit $t_0 \to -\infty$ and so corresponds via the Wiener-Khinchin theorem to the power spectral density (PSD; e.g. Sect. 1.2 of Lindner, 2009) function

$$S(\omega) = \frac{2L^2}{\lambda^2 + (2\pi\omega)^2}, \tag{4.4}$$

where ω is the time frequency and S is the PSD.

In this simple framework, the ocean therefore low-pass filters spectrally constant (white noise) surface heat fluxes, producing a frequency spectrum which is constant (i.e. white noise) in the limit of low frequency ($\omega \ll \lambda$) and follows an inverse square

law (i.e. red noise) in the limit of high frequency ($\omega \gg \lambda$). The transition frequency is determined by the ocean adjustment timescale (i.e. λ). We will return to these classical results concerning Ornstein-Uhlenbeck processes in Section 4.2.3.

Although a useful first-order representation of the evolution of unpredictable surface temperature variability (Frankignoul and Hasselmann, 1977), the model is inherently limited by its treatment of a single forcing and response term, representing a spatial average of a single independent region of the ocean and atmosphere (without accounting for any internal ocean processes, beyond a crude memory term). In a more realistic representation, atmospheric forcing may coherently influence multiple regions of the ocean, which may interact with each other through a range of variables and processes. If the dynamics of these interactions remain linear, (4.1) can be generalised to a non-autonomous linear system of stochastic differential equations:

$$\mathrm{d}\,|\boldsymbol{u}\rangle = \mathbf{A}(t)\,|\boldsymbol{u}\rangle\,\mathrm{d}t + \mathbf{L}\,\mathrm{d}\,|\boldsymbol{W}(t)\rangle\,, \tag{4.5}$$

where $|\boldsymbol{u}\rangle$ is the ocean state vector anomaly, describing the response of each prognostic variable at each location, $|\boldsymbol{W}(t)\rangle$ is a vector of independent standard-normal Wiener processes, $\mathbf{A}(t)$ describes the linear interactions between all ocean variables and locations, and $\mathbf{L}$ is the lower-triangular matrix describing the stochastic atmospheric fluxes through the Cholesky decomposition $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^{\dagger}$ of their covariance matrix. In this decomposition, $\dagger$ represents the adjoint defined by the Euclidean inner product.

Realistic ocean models are not linear, but for small anomalies $|\boldsymbol{u}\rangle$ the complementary equation of (4.5) can provide a first-order description of their anomalous behaviour. Consider a nonlinear system such as a typical ocean general circulation model (GCM):

$$\mathrm{d}\,|\boldsymbol{U}\rangle = \mathcal{N}(|\boldsymbol{U}\rangle\,, t)\,\mathrm{d}t,$$

where $\mathcal{N}$ is a nonlinear operator, t is time and $|\boldsymbol{U}\rangle$ the full state vector. Expansion of the full state vector $|\boldsymbol{U}\rangle = |\bar{\boldsymbol{u}}\rangle + |\boldsymbol{u}\rangle$ (about a mean state $|\bar{\boldsymbol{u}}\rangle$)) yields

$$\mathrm{d}(|\bar{\boldsymbol{u}}\rangle + |\boldsymbol{u}\rangle) = \Big[\mathcal{N}(|\bar{\boldsymbol{u}}\rangle\,, t) + \mathbf{A}(t)\,|\boldsymbol{u}\rangle + \mathcal{O}(|\boldsymbol{u}\rangle^2)\Big]\,\mathrm{d}t, \tag{4.6}$$

Noting that $\mathrm{d}\,|\bar{\boldsymbol{u}}\rangle = \mathcal{N}(|\bar{\boldsymbol{u}}\rangle\,, t)\,\mathrm{d}t$ and neglecting higher order terms leads to the complementary equation of (4.5). In this context, $\mathbf{A}(t)$ is the Jacobian of the nonlinear system with respect to the ocean state:

$$\mathbf{A}(t) = \frac{\partial}{\partial\,|\boldsymbol{U}\rangle}\mathcal{N}(|\bar{\boldsymbol{u}}\rangle\,, t). \tag{4.7}$$

The (zero initial condition) solution to (4.5) is given by

$$|\boldsymbol{u}(t_0, t_1)\rangle = \int_{t_0}^{t_1} \mathbf{\Psi}(t_1, t)\mathbf{L}\,\mathrm{d}\,|\boldsymbol{W}(t)\rangle\,, \tag{4.8}$$

where $\boldsymbol{\Psi}(t_1, t_0)$ is the propagator matrix [the scalar $\Psi(t_1, t_0) = e^{-\lambda(t_1 - t_0)}$ in the univariate case of (4.2)] which describes the linear response of the ocean at time t_1 to changes originating from time t_0.

Beginning from the last formula, we can diagnose the covariance between any two scalar-valued metrics of the ocean state which are linear. These metrics can be defined by the co-vectors $|\boldsymbol{F}_{1,2}\rangle$ where the scalar products $\langle \boldsymbol{F}_{1,2}|\boldsymbol{u}\rangle = \langle \boldsymbol{u}|\boldsymbol{F}_{1,2}\rangle$ are the Euclidean inner products of the co-vectors and the ocean state vector anomaly. We have

$$\mathrm{Cov}(\langle \boldsymbol{F}_1|\boldsymbol{u}(t_0, t_1)\rangle, \langle \boldsymbol{F}_2|\boldsymbol{u}(t_0, t_1)\rangle) = \mathrm{E}\left[\langle \boldsymbol{F}_1| \int_{t_0}^{t_1} \boldsymbol{\Psi}(t_1, t)\mathbf{L}\, \mathrm{d}|\boldsymbol{W}(t)\rangle \, \langle \boldsymbol{F}_2| \int_{t_0}^{t_1} \boldsymbol{\Psi}(t_1, s)\mathbf{L}\, \mathrm{d}|\boldsymbol{W}(s)\rangle\right] \tag{4.9}$$

where s represents time. A multi-dimensional generalisation of Itô's isometry may be applied to this expression (e.g. Section 3.6 of Duan and Wang, 2014). In particular, the Itô integral terms may be written as non-anticipatory (left) Riemann sums such that the right hand side of (4.9) becomes

$$\lim_{K\to\infty} \mathrm{E}\left[\sum_{i=1}^{K}\sum_{j=1}^{K} \langle \boldsymbol{F}_1|\, \boldsymbol{\Psi}(t_1, t_i)\mathbf{L}\, |\boldsymbol{\Delta W}_i\rangle \, \langle \boldsymbol{F}_2|\, \boldsymbol{\Psi}(t_1, t_j)\mathbf{L}\, |\boldsymbol{\Delta W}_j\rangle\right], \tag{4.10}$$

with

$$t_k = t_0 + k\frac{t_1 - t_0}{K}, \quad |\boldsymbol{\Delta W}_k\rangle = (|\boldsymbol{W}(t_{k+1})\rangle - |\boldsymbol{W}(t_k)\rangle),$$

where i, j, k are discrete increment indices, and K is the total number of discrete increments. Applying a transpose and Fubini's theorem:

$$\lim_{K\to\infty} \sum_{i=1}^{K}\sum_{j=1}^{K} \langle \boldsymbol{F}_1|\, \boldsymbol{\Psi}(t_1, t_i)\mathbf{L}\mathrm{E}\left[|\boldsymbol{\Delta W}_i\rangle \, \langle \boldsymbol{\Delta W}_j|\right] \mathbf{L}^\dagger \boldsymbol{\Psi}^\dagger(t_j, t_1)\, |\boldsymbol{F}_2\rangle\,. \tag{4.11}$$

We note that, $\forall\, i \neq j$, the increments of the Wiener processes do not overlap and so are independent by definition, reducing the expression to a single sum

$$\lim_{K\to\infty} \sum_{i=1}^{K} \langle \boldsymbol{F}_1|\, \boldsymbol{\Psi}(t_1, t_i)\mathbf{L}\mathrm{E}\left[|\boldsymbol{\Delta W}_i\rangle \, \langle \boldsymbol{\Delta W}_i|\right] \mathbf{L}^\dagger \boldsymbol{\Psi}^\dagger(t_i, t_1)\, |\boldsymbol{F}_2\rangle\,, \tag{4.12}$$

in which the central outer product corresponds to a diagonal matrix, as the vectors are elementwise independent. As Wiener increments are normally distributed as $W(t_{k+1} - t_k) \sim N(0, t_{k+1} - t_k)$, in their infinitesimal limit the equation becomes

$$\mathrm{Cov}(\langle \boldsymbol{F}_1|\boldsymbol{u}(t_0, t_1)\rangle, \langle \boldsymbol{F}_2|\boldsymbol{u}(t_0, t_1)\rangle) = \int_{t_0}^{t_1} \langle \boldsymbol{F}_1|\boldsymbol{\Psi}(t_1, t)\boldsymbol{\Sigma}\boldsymbol{\Psi}^\dagger(t, t_1)|\boldsymbol{F}_2\rangle\, \mathrm{d}t. \tag{4.13}$$

Note that our solution generalises the result heuristically derived by Sévellec et al. (2018). Similarly to their approach, we remark that while it is standard to diagnose the variance evolution of a metric by propagating many realisations of (4.8) as an

ensemble and considering its spread, (4.13) does not require us to propagate any such realisation. Instead, it describes the response of such an ensemble (in the theoretical limit of large ensemble size) using only the statistical properties ($\boldsymbol{\Sigma}$) of the noise. It further provides a dynamical link between the response of the target metrics $\langle \boldsymbol{F}_{1,2}|$ and the stochastic source of variability represented by $\boldsymbol{\Sigma}$. Where this representation can be linearly partitioned into independent sources (for instance internal and external, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_\mathbf{I} + \boldsymbol{\Sigma}_\mathbf{E}$), the variance can be dynamically attributed to each. The only requirements of the method are that

1. Our metrics of interest $\langle \boldsymbol{F}_{1,2}|$ are linear functions of the ocean state;
2. We have a linear model of ocean dynamics, $\boldsymbol{\Psi}(t_1, t_0)$ [we take a linearised OGCM which following (4.6) is valid for small variations about a trajectory, see Section 4.3];
3. We have a complete statistical description $\boldsymbol{\Sigma}$ of any stochastic sources of variability.

Regarding the latter point, two approaches may be taken: the properties of the stochastic processes may be diagnosed and prescribed (as in Sévellec et al., 2018, for instance), or they may be determined from the linear model itself (in the framework of an optimisation problem, as in Sévellec et al., 2007, 2009, for instance). We begin with the latter approach, which provides insight into the mechanisms by which sources of variability are translated into oceanic variance in a theoretical setting.

4.2.2 Optimal Stochastic Perturbations

As $\boldsymbol{\Sigma}$ can be allowed to take any form in (4.13), the problem of variance estimation can be reformulated as an optimisation question: what form should $\boldsymbol{\Sigma}$ take such that variance $\mathrm{Var}(\langle \boldsymbol{F}|\boldsymbol{u}(t_0,t_1)\rangle) = \int_{t_0}^{t_1} \langle \boldsymbol{F}|\boldsymbol{\Psi}(t_1,t)\boldsymbol{\Sigma}\boldsymbol{\Psi}^\dagger(t,t_1)|\boldsymbol{F}\rangle \,\mathrm{d}t$ is maximal for a given metric $\langle \boldsymbol{F}|$? The solution to the problem, under certain conditions, can be determined dynamically from the linear model itself, allowing insight into the mechanisms behind oceanic uncertainty without explicitly prescribing sources of uncertainty.

To determine the optimal $\boldsymbol{\Sigma}$, we apply two constraints to the optimal variance source: its global average has fixed amplitude, and any two points which are not independent have a correlation of ± 1. The former implies that the stochastic process has finite power (corresponding to band-limited white noise), while the latter assumes that if two points covary, they must do so completely constructively (as would be optimal). We begin by considering the general case, where the stochastic process is partitioned into "N" such regions (where each point in the region is perfectly correlated), before considering the specific cases corresponding to the two limits of N: (i) $N = 1$

corresponding to a fully global correlation (as in Sévellec et al., 2007, 2009) and (ii) $N=n$ (where n is the dimension of the state vector, $|\boldsymbol{u}\rangle$), corresponding to the absence of any correlation.

4.2.2.1 General case

As outlined above, we partition the stochastic process into N regions such that points within the regions are perfectly covarying, but are independent of points in other regions. Equivalently, we separate $\mathbf{\Sigma}\in\mathbb{R}^{n\times n}$ into N local matrices $\mathbf{\Sigma}_i\in\mathbb{R}^{m_i\times m_i}$ (where m_i is the local dimension of the i$^{\text{th}}$ region), and define a binary projection $\mathbf{B}_i\in\mathbb{R}^{n\times m_i}$ such that

$$\mathbf{\Sigma}=\sum_{i=1}^{N}\mathbf{B}_i\mathbf{\Sigma}_i\mathbf{B}_i^{\dagger}. \tag{4.14}$$

Following (4.8), the evolution of the state vector in response to stimulation in the i$^{\text{th}}$ region is

$$|\boldsymbol{u}_i(t_0,t_1)\rangle=\int_{t_0}^{t_1}\mathbf{\Psi}(t_1,t)\mathbf{B}_i\mathbf{L}_i\,\mathrm{d}\,|\boldsymbol{W}_i(t)\rangle\,, \tag{4.15}$$

where $\mathbf{\Sigma}_i=\mathbf{L}_i\mathbf{L}_i^{\dagger}$ is the Cholesky decomposition of the local covariance matrix, equivalently to the global case. Fundamentally, as the region is perfectly correlated, it may be written in terms of a single stochastic process. The vector $\mathbf{L}_i\,\mathrm{d}\,|\boldsymbol{W}_i\rangle$ thus becomes $|\boldsymbol{L}_i\rangle\,\mathrm{d}W_i$, such that $\mathbf{\Sigma}_i$ is the outer product $\mathbf{\Sigma}_i=|\boldsymbol{L}_i\rangle\langle\boldsymbol{L}_i|$. The implication is that in the region, a single Wiener process is "shaped" by a pattern of local amplitudes $|\boldsymbol{L}_i\rangle$.

In order to determine the optimal shape of this pattern, we utilise the method of Lagrange multipliers (consistently with Sévellec et al., 2007). In particular, we wish to maximise the local contribution to the variance

$$\mathrm{Var}(\langle\boldsymbol{F}|\boldsymbol{u}_i\rangle)=\int_{t_0}^{t_1}\langle\boldsymbol{F}|\mathbf{\Psi}(t_1,t)\mathbf{B}_i\mathbf{\Sigma}_i\mathbf{B}_i^{\dagger}\mathbf{\Psi}(t,t_1)|\boldsymbol{F}\rangle\,dt \tag{4.16a}$$

under the constraint that the amplitude ϵ_i of $\mathbf{\Sigma}_i$ follows

$$\mathrm{Tr}(\mathbf{S}_i\mathbf{\Sigma}_i)=\langle\boldsymbol{L}_i|\mathbf{S}_i|\boldsymbol{L}_i\rangle=\epsilon_i^2, \tag{4.16b}$$

where $\mathbf{S}_i\in\mathbb{R}^{m_i\times m_i}$ is a (diagonal) volumetric weighting matrix. The corresponding Lagrange function can be expressed as

$$\mathcal{L}(\gamma_i,|\boldsymbol{L}_i\rangle\,,t_0,t_1)=\int_{t_0}^{t_1}\langle\boldsymbol{F}|\mathbf{\Psi}(t_1,t)\mathbf{B}_i|\boldsymbol{L}_i\rangle^2\,\mathrm{d}t-\gamma_i(\langle\boldsymbol{L}_i|\mathbf{S}_i|\boldsymbol{L}_i\rangle-\epsilon_i^2), \tag{4.17}$$

where the scalar γ_i is the Lagrange multiplier. Maximising the Lagrangian leads to

$$\begin{aligned} \left.\frac{\partial \mathcal{L}}{\partial \left|\boldsymbol{L}_i\right\rangle}\right|_{\{\gamma_i^*,\left|\boldsymbol{L}_i^*\right\rangle\}} &= 0, \\ \int_{t_0}^{t_1} \left(\mathbf{B}_i^\dagger \boldsymbol{\Psi}^\dagger(t,t_1)\left|\boldsymbol{F}\right\rangle\left\langle\boldsymbol{F}\right|\boldsymbol{\Psi}(t_1,t)\mathbf{B}_i\right) \mathrm{d}t \left|\boldsymbol{L}_i^*\right\rangle - \gamma_i^* \mathbf{S}_i \left|\boldsymbol{L}_i^*\right\rangle &= 0, \end{aligned} \tag{4.18}$$

which holds when γ_i^* and $\left|\boldsymbol{L}_i^*\right\rangle$ are an eigenvalue-eigenvector pair of

$$\mathbf{S}_i^{-1} \int_{t_0}^{t_1} \left(\mathbf{B}_i \boldsymbol{\Psi}^\dagger(t,t_1)|\boldsymbol{F}\rangle\langle\boldsymbol{F}|\boldsymbol{\Psi}(t_1,t)\mathbf{B}_i^\dagger\right) \mathrm{d}t, \tag{4.19}$$

since $\mathbf{S}_i$ (as an operator representing a norm) is invertible. Any such eigenpair represents a particular solution to the optimization problem, but of these we seek the solution with the greatest effect. We note that left multiplication of (4.18) by $\langle \boldsymbol{L}_i^* | \mathbf{S}_i$ results in

$$\int_{t_0}^{t_1} \langle \boldsymbol{F}|\boldsymbol{\Psi}(t_1,t)\mathbf{B}_i|\boldsymbol{L}_i^*\rangle^2 \,\mathrm{d}t = \gamma_i^* \left\langle \boldsymbol{L}_i^*|\mathbf{S}_i|\boldsymbol{L}_i^*\right\rangle,$$

or, equivalently, $\mathrm{Var}(\langle \boldsymbol{F}|\boldsymbol{u}_i(t_0,t_1)\rangle) = \gamma_i^* \epsilon_i^2$, so that the Lagrangian multiplier, γ_i, is essentially representing the variance that we wish to maximise. Hence the eigenvector $\left|\boldsymbol{L}_i^{\mathrm{opt}}\right\rangle$ corresponding to the universally optimal solution of (4.18) is that belonging to the leading eigenvalue γ_i^{opt}. Rescaling the outer product of this eigenvector, the optimal covariance matrix with amplitude meeting the constraint (4.16b) in the i$^{\mathrm{th}}$ region is therefore

$$\boldsymbol{\Sigma}_i^{\mathrm{opt}} = \epsilon_i^2 \frac{\left|\boldsymbol{L}_i^{\mathrm{opt}}\right\rangle \left\langle \boldsymbol{L}_i^{\mathrm{opt}}\right|}{\left\langle \boldsymbol{L}_i^{\mathrm{opt}}|\mathbf{S}_i|\boldsymbol{L}_i^{\mathrm{opt}}\right\rangle}. \tag{4.20}$$

Our local magnitude ϵ_i may be chosen arbitrarily, and so, although the N regions correspond to N independent problems, we seek an optimal scaling ϵ_i which maximises their individual contribution to the overall variance, while constraining the total magnitude $\sum_{i=1}^N \epsilon_i^2 = \epsilon^2$. In particular, we note that the total variance $\mathrm{Var}(\langle \boldsymbol{F}|\boldsymbol{u}(t_0,t_1)\rangle) = \sum_{i=1}^N \epsilon_i^2 \gamma_i^{\mathrm{opt}}$ following the above. This may be alternatively rewritten as an inner product $\mathrm{Var}(\langle \boldsymbol{F}|\boldsymbol{u}(t_0,t_1)\rangle) = \langle \boldsymbol{E}|\boldsymbol{\gamma}\rangle$, where $|\boldsymbol{E}\rangle$ and $|\boldsymbol{\gamma}\rangle$ are vectors of dimension N concatenating all the amplitudes (ϵ_i^2) and optimal variances (γ_i^{opt}), respectively, of the local optimal shape ($\left|\boldsymbol{L}_i^{\mathrm{opt}}\right\rangle$) for the N regions. As the inner product is maximal for parallel vectors (i.e. $|\boldsymbol{E}\rangle$ parallel to $|\boldsymbol{\gamma}\rangle$), it follows after some algebra that

$$\epsilon_i^2 = \frac{\epsilon^2 \gamma_i^{\mathrm{opt}}}{\sum_{i=1}^N \gamma_i^{\mathrm{opt}}}. \tag{4.21}$$

Hence, for these choices of ϵ_i, we have

$$\boldsymbol{\Sigma}^{\mathrm{opt}} = \frac{\epsilon^2}{\sum_{i=1}^N \gamma_i^{\mathrm{opt}}} \sum_{i=1}^N \gamma_i^{\mathrm{opt}} \mathbf{B}_i \frac{\left|\boldsymbol{L}_i^{\mathrm{opt}}\right\rangle \left\langle \boldsymbol{L}_i^{\mathrm{opt}}\right|}{\left\langle \boldsymbol{L}_i^{\mathrm{opt}}|\mathbf{S}_i|\boldsymbol{L}_i^{\mathrm{opt}}\right\rangle} \mathbf{B}_i^\dagger, \tag{4.22}$$

where, as described above, $|\boldsymbol{L}_i^{\text{opt}}\rangle$ and γ_i^{opt} is the leading eigenpair of

$$\mathbf{S}_i^{-1} \int_{t_0}^{t_1} \mathbf{B}_i \boldsymbol{\Psi}^\dagger(t, t_1)\, |\boldsymbol{F}\rangle \langle \boldsymbol{F}|\, \boldsymbol{\Psi}(t_1, t) \mathbf{B}_i^\dagger \, \mathrm{d}t.$$

4.2.2.2 Limiting cases

The above derivation applies to the case of N perfectly correlated independent regions, but we may consider two specific cases of this in order to imitate conditions similar to the atmospherically forced and eddy-driven variability felt by the ocean. In particular, we consider the two limiting cases: $N=1$ and $N=\dim(|\boldsymbol{u}\rangle)$. The former case, where the forcing is everywhere perfectly correlated, can be applied to the surface layer as an idealised representation of the large-scale coherent patterns of the atmosphere (Sévellec et al., 2007, 2009). The latter case, where the forcing is uncorrelated between all variables and locations, is taken as an idealised representation of small-scale noise in the ocean (i.e. noise induced by subgrid processes). These cases correspond to solving a single eigenvalue problem vs. solving $\dim(|\boldsymbol{u}\rangle)$ (trivially scalar) eigenvalue problems. In particular, for $N=1$, the sole projection matrix is the identity matrix $\mathbf{B}_1=\mathbf{I}$, while for $N=n=\dim(|\boldsymbol{u}\rangle)$, the projection matrices become the standard basis vectors $\mathbf{B}_i=|\boldsymbol{e}_i\rangle$ (i.e. $|\boldsymbol{e}_i\rangle$ projects a scalar to the i$^{\text{th}}$ location of the full state vector).

In the former (everywhere perfectly covarying) case, (4.22) becomes

$$\boldsymbol{\Sigma}^{\text{opt}} = \epsilon^2 \frac{|\boldsymbol{L}^{\text{opt}}\rangle \langle \boldsymbol{L}^{\text{opt}}|}{\langle \boldsymbol{L}^{\text{opt}}|\mathbf{S}|\boldsymbol{L}^{\text{opt}}\rangle}, \tag{4.23}$$

where $|\boldsymbol{L}^{\text{opt}}\rangle$ is the leading eigenvector of

$$\mathbf{S}^{-1} \int_{t_0}^{t_1} \boldsymbol{\Psi}^\dagger(t, t_1)\, |\boldsymbol{F}\rangle \langle \boldsymbol{F}|\, \boldsymbol{\Psi}(t_1, t)\, \mathrm{d}t.$$

The latter (everywhere uncorrelated) case corresponds to the condition that every point is independent, and $\boldsymbol{\Sigma}^{\text{opt}}$ is diagonal. The associated eigen"vector" problems are scalar, such that the eigenspace is infinite. All terms in (4.22) are now scalars such that $|\boldsymbol{L}_i^{\text{opt}}\rangle$ can be seen to cancel, while the matrices $\mathbf{S}_i$ may be written as S_i. Ultimately,

$$\boldsymbol{\Sigma}^{\text{opt}} = \frac{\epsilon^2}{\sum_{i=1}^{N} \gamma_i^{\text{opt}}} \sum_{i=1}^{N} |\boldsymbol{e}_i\rangle \frac{\gamma_i^{\text{opt}}}{S_i} \langle \boldsymbol{e}_i|, \tag{4.24}$$

where, solving (4.18) with $\mathbf{B}_i=|\boldsymbol{e}_i\rangle$, the eigenvalues γ_i^{opt} are trivially the diagonal elements of

$$\mathbf{S}^{-1} \int_{t_0}^{t_1} \boldsymbol{\Psi}^\dagger(t, t_1)\, |\boldsymbol{F}\rangle \langle \boldsymbol{F}|\, \boldsymbol{\Psi}(t_1, t) \mathrm{d}t.$$

The sum of the eigenvalues is also the trace of this (scaled outer product) matrix, and is thus given by the corresponding inner product. Therefore, from (4.24) the optimal

stochastic covariance matrix in the completely uncorrelated case is

$$\mathbf{\Sigma}^{\mathrm{opt}} = \frac{\epsilon^2}{\int_{t_0}^{t_1} \langle \boldsymbol{F}|\mathbf{\Psi}(t_1,t)\mathbf{S}^{-1}\mathbf{\Psi}^{\dagger}(t,t_1)|\boldsymbol{F}\rangle \,\mathrm{d}t} \mathrm{diag}\left[\mathbf{S}^{-1}\int_{t_0}^{t_1} \mathbf{\Psi}^{\dagger}(t,t_1)\,|\boldsymbol{F}\rangle\,\langle \boldsymbol{F}|\,\mathbf{\Psi}(t_1,t)\mathrm{d}t\right]\mathbf{S}^{-1} \tag{4.25}$$

(where the diag[·] operator corresponds to the diagonal matrix with the same diagonal). We respectively use these two limiting cases to explore theoretical variance linked to idealised atmospheric forcing (assuming perfect correlation everywhere over the surface and zero noise in the interior) and ocean internal subgrid fluxes (assuming noise everywhere, with zero correlation between locations and variables).

A useful metric of the OSP is the ratio of the output variance to the input variance $A_* = \mathrm{Var}(\langle \boldsymbol{F}|\boldsymbol{u}(t_0,t_1)\rangle/\epsilon^2$, which we term the response ratio. Notably, for the globally perfect covariance case, this is simply the associated eigenvalue

$$A_* = \gamma^{\mathrm{opt}}. \tag{4.26a}$$

For the globally decorrelated case,

$$A_* = \int_{t_0}^{t_1} \langle \boldsymbol{F}|\mathbf{\Psi}(t_1,t)\mathbf{S}^{-1}\mathbf{\Psi}^{\dagger}(t,t_1)|\boldsymbol{F}\rangle \,\mathrm{d}t \tag{4.26b}$$

is the sum of the eigenvalues.

4.2.3 Temporally correlated forcing

Our considerations so far have involved stochastic forcing with varying levels of spatial coherence, but which is serially decorrelated (therefore band-limited white noise). While this allows an idealised, theoretical exploration of variance generation mechanisms in the optimal case, it is inadequate for realistically representing turbulent fluxes in the climate system, as we wish to in the diagnosed case. Indeed, while white noise is typically considered an acceptable representation of atmospheric variability (which decays on timescales much shorter than those of the oceanic large scale; Hasselmann, 1976), the ocean mesoscale eddy field evolves much more slowly (e.g. Chelton et al., 2007). To realistically represent this using diagnosed fluxes, we therefore extend our framework to include temporally correlated stochastic forcing. We consider again the Ornstein-Uhlenbeck case, which is a simple example of a temporally correlated stochastic process. As in Williams et al. (2016), we treat the internal time-correlated stationary eddy component as sufficiently well-represented by an additive noise term, based on the generally Gaussian nature of the mesoscale eddy field (Biri et al., 2015).

We begin by modifying (4.5) such that anomalous fluxes are now modelled by a continuous, time-integrable stochastic process (contrary to the former, white noise

case, where they were everywhere discontinuous and representable only in the framework of distribution theory). The equation becomes

$$\mathrm{d}\left|\boldsymbol{u}\right\rangle = \left(\mathbf{A}(t)\left|\boldsymbol{u}\right\rangle + \left|\boldsymbol{X}(t)\right\rangle\right)\mathrm{d}t, \tag{4.27}$$

where, as before, $|\boldsymbol{u}\rangle$ defines the state vector anomaly, $\mathbf{A}$ defines the system's linear dynamics [for instance via the Jacobian of a corresponding nonlinear system, as in (4.7)], and, in contrast to the previous cases, $|\boldsymbol{X}\rangle$ is the forcing from continuous, spatiotemporally correlated stochastic processes. The zero-initial-condition solution is given by

$$\left|\boldsymbol{u}(t_0,t_1)\right\rangle = \int_{t_0}^{t_1} \boldsymbol{\Psi}(t_1,t)\left|\boldsymbol{X}(t)\right\rangle \,\mathrm{d}t \tag{4.28}$$

where the complementary equation and therefore the propagator matrix, $\boldsymbol{\Psi}(t_1,t_0)$, are notably identical to (4.8). As in (4.9), we seek the covariance between two metrics of the state vector, given by

$$\mathrm{Cov}(\langle\boldsymbol{F}_1|\boldsymbol{u}(t_0,t_1\rangle\,,\langle\boldsymbol{F}_2|\boldsymbol{u}(t_0,t_1\rangle) = \int_{t_0}^{t_1}\int_{t_0}^{t_1} \langle\boldsymbol{F}_1|\,\boldsymbol{\Psi}(t_1,t)\mathrm{E}\left[\left|\boldsymbol{X}(t)\right\rangle\left\langle\boldsymbol{X}(s)\right|\right]\boldsymbol{\Psi}^\dagger(s,t_1)\left|\boldsymbol{F}_2\right\rangle \mathrm{d}t\,\mathrm{d}s, \tag{4.29}$$

where the term $\mathrm{E}\left[\left|\boldsymbol{X}(t)\right\rangle\left\langle\boldsymbol{X}(s)\right|\right]$ gives the spatiotemporal covariance matrix of the forcing. In the white noise case, the autocorrelation conceptually corresponds to the Dirac delta function, leading to $\mathrm{E}\left[\left|\boldsymbol{X}(t)\right\rangle\left\langle\boldsymbol{X}(s)\right|\right]=\delta(t-s)\mathbf{L}\mathbf{L}^\dagger$, consistently with (4.13). For a vector $|\boldsymbol{X}\rangle$ of saturated Ornstein-Uhlenbeck processes [such as (4.2) with $t_0\to-\infty$], a multivariate generalisation of (4.3) gives

$$\mathrm{E}\left[\left|\boldsymbol{X}(t)\right\rangle\left\langle\boldsymbol{X}(s)\right|\right] = e^{-\boldsymbol{\lambda}t}\mathbf{L}\mathbf{L}^\dagger e^{-\boldsymbol{\lambda}^\dagger s}, \tag{4.30}$$

where $\boldsymbol{\lambda}$ is a diagonal matrix of reciprocal e-folding times of the anomalous fluxes at each location, and $\mathbf{L}\mathbf{L}^\dagger=\boldsymbol{\Sigma}$ is their spatial covariance matrix. As these quantities can be diagnosed from an appropriate dataset, we can use this formulation to diagnose the variance growth.

In the proceeding section we diagnose (from realistic models) $\boldsymbol{\lambda}$ and $\boldsymbol{\Sigma}$ for the cases of external (atmospheric; $\boldsymbol{\lambda}_\mathrm{E}$, $\boldsymbol{\Sigma}_\mathrm{E}$) and internal (oceanic turbulent mesoscale eddy driven; $\boldsymbol{\lambda}_\mathrm{I}$, $\boldsymbol{\Sigma}_\mathrm{I}$) turbulent fluxes, assessing the appropriateness of the Ornstein-Uhlenbeck representation. We then proceed to attribute the variance of different metrics in response to these sources, which, following (4.29) and assuming independence between the internal and external components is given by

$$\begin{aligned}\mathrm{Var}(\langle\boldsymbol{F}|\boldsymbol{u}(t_0,t_1\rangle) = &\int_{t_0}^{t_1}\int_{t_0}^{t_1} \langle\boldsymbol{F}|\,\boldsymbol{\Psi}(t_1,t)e^{-\boldsymbol{\lambda}_\mathrm{I}t}\boldsymbol{\Sigma}_\mathrm{I}e^{-\boldsymbol{\lambda}_\mathrm{I}^\dagger s}\boldsymbol{\Psi}^\dagger(s,t_1)\left|\boldsymbol{F}\right\rangle \mathrm{d}t\,\mathrm{d}s \\ &+ \int_{t_0}^{t_1}\int_{t_0}^{t_1} \langle\boldsymbol{F}|\,\boldsymbol{\Psi}(t_1,t)e^{-\boldsymbol{\lambda}_\mathrm{E}t}\boldsymbol{\Sigma}_\mathrm{E}e^{-\boldsymbol{\lambda}_\mathrm{E}^\dagger s}\boldsymbol{\Psi}^\dagger(s,t_1)\left|\boldsymbol{F}\right\rangle \mathrm{d}t\,\mathrm{d}s.\end{aligned} \tag{4.31}$$

The variance may be broken down further still, by writing the covariance matrices as the sum of their different components. For example, we are interested in the independent contributions of buoyancy and momentum fluxes to the externally forced component $\mathrm{Var_E}$ of the variance (corresponding to the $\boldsymbol{\lambda}_\mathrm{E}$, $\boldsymbol{\Sigma}_\mathrm{E}$ terms), and, in the latter case, the separate contributions of the covarying zonal and meridional momentum fluxes. The final term of 4.31 can accordingly be split into:

$$\begin{aligned}\mathrm{Var_E}(\langle \boldsymbol{F}|\boldsymbol{u}(t_0,t_1)\rangle) = &\int_{t_0}^{t_1}\int_{t_0}^{t_1} \langle \boldsymbol{F}|\,\boldsymbol{\Psi}(t_1,t)e^{-\boldsymbol{\lambda}_\mathrm{E}^\mathrm{b} t}\boldsymbol{\Sigma}_\mathrm{E}^{b}e^{-\boldsymbol{\lambda}_\mathrm{E}^\mathrm{b}{}^\dagger s}\boldsymbol{\Psi}^\dagger(s,t_1)\,|\boldsymbol{F}\rangle\,\mathrm{d}t\,\mathrm{d}s\\ &+\int_{t_0}^{t_1}\int_{t_0}^{t_1} \langle \boldsymbol{F}|\,\boldsymbol{\Psi}(t_1,t)e^{-\boldsymbol{\lambda}_\mathrm{E}^\mathrm{u} t}\boldsymbol{\Sigma}_\mathrm{E}^{\mathrm{u}}e^{-\boldsymbol{\lambda}_\mathrm{E}^\mathrm{u}{}^\dagger s}\boldsymbol{\Psi}^\dagger(s,t_1)\,|\boldsymbol{F}\rangle\,\mathrm{d}t\,\mathrm{d}s\\ &+\int_{t_0}^{t_1}\int_{t_0}^{t_1} \langle \boldsymbol{F}|\,\boldsymbol{\Psi}(t_1,t)e^{-\boldsymbol{\lambda}_\mathrm{E}^\mathrm{v} t}\boldsymbol{\Sigma}_\mathrm{E}^{\mathrm{v}}e^{-\boldsymbol{\lambda}_\mathrm{E}^\mathrm{v}{}^\dagger s}\boldsymbol{\Psi}^\dagger(s,t_1)\,|\boldsymbol{F}\rangle\,\mathrm{d}t\,\mathrm{d}s\\ &+\int_{t_0}^{t_1}\int_{t_0}^{t_1} \langle \boldsymbol{F}|\,\boldsymbol{\Psi}(t_1,t)e^{-\boldsymbol{\lambda}_\mathrm{E}^\mathrm{u} t}\boldsymbol{\Sigma}_\mathrm{E}^{\mathrm{u,v}}e^{-\boldsymbol{\lambda}_\mathrm{E}^\mathrm{v}{}^\dagger s}\boldsymbol{\Psi}^\dagger(s,t_1)\,|\boldsymbol{F}\rangle\,\mathrm{d}t\,\mathrm{d}s, \qquad (4.32)\end{aligned}$$

where $(\boldsymbol{\lambda}_\mathrm{E}^{\{\mathrm{b,u,v}\}}, \boldsymbol{\Sigma}_\mathrm{E}^{\{\mathrm{b,u,v}\}})$ are the external noise properties for the buoyancy, and zonal and meridional momentum fluxes, respectively, $\boldsymbol{\Sigma}_\mathrm{E}^{\mathrm{u,v}}$ is for the zonal and meridional covariance term.

Finally, in addition to separating the variance into contributions from different variables, we note that we can also isolate contributions from different regions of space. The inner products of (4.31) represent spatial integrals of local contributions to the total variance (integrated over volume in the internal case and over area in the external case). An alternative formulation of (4.31) is therefore

$$\mathrm{Var}(\langle \boldsymbol{F}|\boldsymbol{u}(t_0,t_1)\rangle) = \int_\Omega \mathcal{V}_\mathrm{I}(x,y,z,t_0,t_1)\,d\mathrm{V} + \int_{\Omega_0} \mathcal{V}_\mathrm{E}(x,y,t_0,t_1)\,d\mathrm{A}, \qquad (4.33)$$

where $\mathcal{V}_\mathrm{I}$ and $\mathcal{V}_\mathrm{E}$ are continuous functions representing the respective internal variance contribution per unit volume and external variance contribution per unit area, Ω and $d\mathrm{V}$ represent the ocean interior and a volume increment, respectively, Ω_0 and $d\mathrm{A}$ represent the ocean surface and an area increment, respectively, and x, y and z are the zonal, meridional and vertical coordinates. The corresponding integrands are thus spatial distributions of variance contributions. This can be applied to both (4.31) and (4.32) without loss of generality.

4.3 Model configurations, methods, and experimental design

4.3.1 Linear ocean model configuration

As outlined in Section 4.2, we use a linear ocean model to provide the propagator matrix $\boldsymbol{\Psi}$ which is used to both derive our OSPs [following (4.22)] and evolve our prescribed, diagnosed stochastic processes [following (4.29)]. The model is v3.4 of the NEMO GCM (Madec, 2012) whose routines are linearised in the tangent-linear and adjoint model (TAM) package NEMOTAM (Vidard et al., 2015). The model is run in the nominal 2° ORCA2 configuration with 31 vertical levels in partial-step z-coordinates, subject to repeated CORE normal year forcing (Large and Yeager, 2004). More details can be found in Stephenson et al. (2020). We note that the same ocean model is common to our linear propagator $\boldsymbol{\Psi}$, the coupled climate model (Section 4.3.2) used to diagnose our stochastic external flux representation ($\boldsymbol{\lambda}_{\mathrm{E}}$, $\boldsymbol{\Sigma}_{\mathrm{E}}$) and the high-resolution ocean model (Section 4.3.3) used to diagnose our stochastic internal flux representation ($\boldsymbol{\lambda}_{\mathrm{I}}$, $\boldsymbol{\Sigma}_{\mathrm{I}}$). We therefore consider seasonal variations of the oceanic large scale to be common to all three cases, which are explicitly captured in the trajectory $|\bar{\boldsymbol{u}}\rangle$ of (4.6). In this sense, our stochastic representations are of anomalies [$|\boldsymbol{u}\rangle$ in (4.6)] from this shared climatology, which are unresolved in the low-resolution, ocean-only model. We note that this inter-model approach is subject to the same approximation as other studies which diagnose forcings from higher fidelity models (or indeed observations): the ocean circulation and vertical structure is assumed to be the same in all cases. Projecting from one model to another will thus naturally lead to misrepresentations around technically challenging features (such as the Gulf Stream separation, for example), or areas where isopycnals do not align between models. The impact of such mismatches, similarly to that of non-stationary eddy variations such as seasonal cycles, is taken to be small.

4.3.2 Diagnosis of realistic stochastic atmospheric fluxes

In order to represent the effects of anomalous surface fluxes by an Ornstein-Uhlenbeck process we diagnose the parameters $\boldsymbol{\lambda}_{\mathrm{E}}$, $\boldsymbol{\Sigma}_{\mathrm{E}}$ from the outputs of a coupled climate model. In particular, we use the IPSL-CM5A-LR coupled model, which was run for twenty years in its CMIP5 pre-industrial control configuration (c.f. Dufresne et al., 2013). The ocean component of the model is NEMO-ORCA2 (v3.2) which has the same (ORCA2) configuration as our linearised ocean model. In order to cleanly separate atmospherically forced variability from internally forced turbulent ocean variability (which is diagnosed separately; Section 4.3.3), the ocean component of the chosen climate model is laminar, such that there is no internal turbulent variability

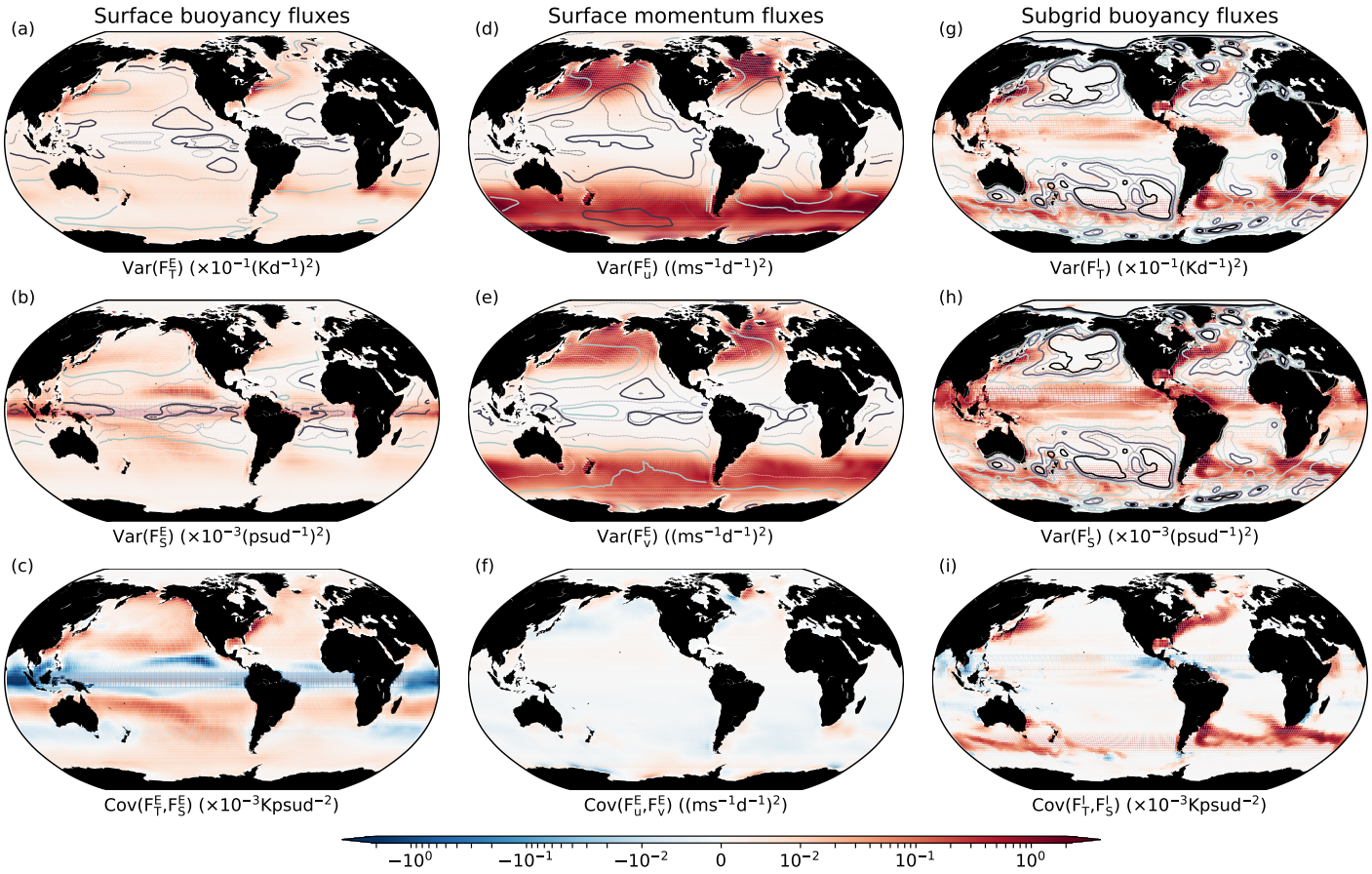


FIGURE 4.1: Leading diagonal of flux covariance matrices (shading) and flux decorrelation times (contours) for external (atmospheric; a-f) and internal (ocean mesoscale eddy; g-h) turbulent fluxes. Contours are separated by half a day and increase in darkness, with thicker, solid contours at 0.5 (lightest), 1.5 and 2.5 (darkest) days. In the latter case, quantities are depth-averaged and contours are separated by ten days with thicker contours at 10 (lightest), 30 and 50 (darkest) days.

(Grégorio et al., 2015). The atmospheric component is the LMDZ5a model, with a horizontal resolution of (3.75×1.9)° and 39 levels in the vertical (Hourdin et al., 2013).

To isolate the impact of external forcing, the twenty year time series of daily averaged surface wind stress, heat and freshwater fluxes produced by the coupled model were considered. As described in Section 4.3.1, the climatologies of these fluxes were taken to be present in the trajectory of the linear model (via its repeated annual forcing) and so were removed. The remaining anomalies were then linearly mapped to a corresponding external-flux-induced rate of change in ocean surface zonal and meridional velocity (F_u^E and F_v^E, respectively), sea surface temperature (SST; F_T^E), and sea surface salinity (SSS; F_S^E). The covariance and e-folding decorrelation time of these time series (Fig. 4.1a-e) were then used to construct the stochastic representation.

The variance of the heat flux term (Fig. 4.1a) is broadly distributed away from the tropics with regions of intense focus such as western boundary currents, while the freshwater flux variance term is conversely highest in the tropics (Fig. 4.1b). Their covariance (Fig. 4.1c) reflects this difference such that decreasing F_T^E corresponds to salinification in these regions of highest variance in F_S^E and freshening in regions of highest F_T^E variance. Both temperature and salinity changes are most persistent at low latitudes (Fig. 4.1a,b, contours). For wind-stress-induced surface velocity changes, zonal and meridional variances show broadly similar spatial patterns, focused at high latitudes (Fig. 4.1d and e, respectively). The zonal component is notably more intense and more persistent (Fig. 4.1d,e, contours).

The matrix $\mathbf{\Sigma}_E$ was populated using the covariances of these time series with the corresponding time series of each dependent variable at every other location (Fig. 4.1 shows the lead diagonal of $\mathbf{\Sigma}_E$). $\boldsymbol{\lambda}_E$ is a diagonal matrix of local e-folding times calculated from the lag-autocorrelation of the time series (shown by contours in Fig. 4.1). Buoyancy and momentum fluxes were assumed independent of each other, but their components (temperature and salinity for the former, meridional and zonal momentum for the latter) are allowed to spatially covary.

To evaluate the goodness of fit of the Ornstein-Uhlenbeck process representation, we compare the PSD of a theoretically perfect process with matching parameters at each location [following (4.4)] with the PSD produced by the time series. To fairly weight all frequencies, we use the root-mean-square logarithmic error (RMSLE) metric, normalised by the mean of the logarithm of the PSD. This effectively corresponds to taking a normalised root-mean-square error, but in logarithmic space, such that all frequencies contribute evenly. For comparison we also evaluate the error in the same way when the more traditional Gaussian white noise representation (i.e. constant PSD) is used to fit the model outputs. This reveals that the Ornstein-Uhlenbeck model is almost everywhere an improvement in representing our diagnosed anomalous fluxes (Fig. 4.2).

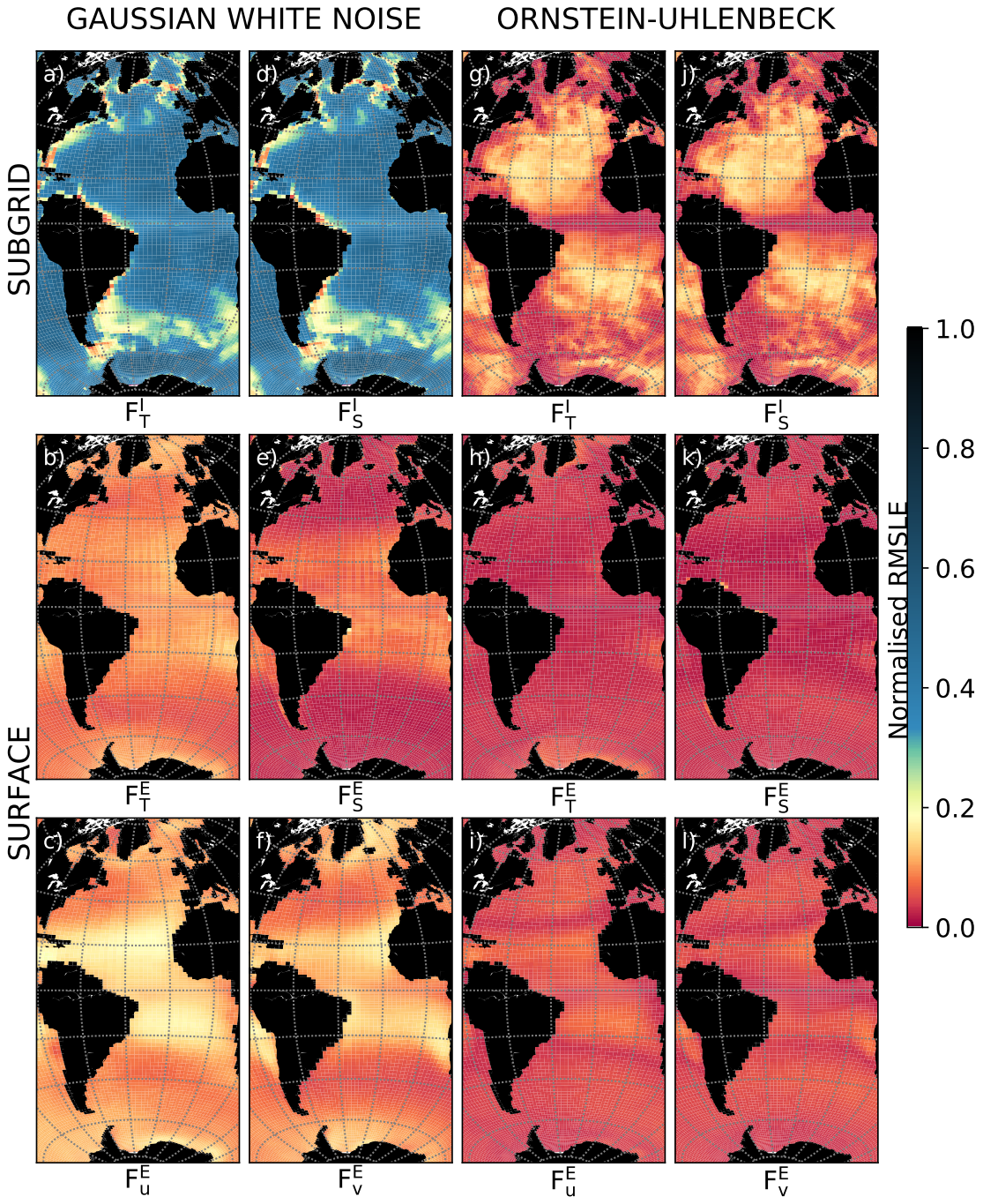


FIGURE 4.2: Map of error in fitting power spectra of internal (top row; depth-averaged values shown) and external (middle and lower rows) turbulent fluxes to the theoretical power spectrum of a white Gaussian noise (a-f) and an Ornstein-Uhlenbeck process (g-l). a,g: internal heat flux; b,h: external heat flux; c,i: external momentum flux (zonal component); d,j: internal salt flux; f,l: external momentum flux (meridional component). The error is given as the ratio of root-mean-square logarithmic error (RMSLE) to the log-mean value of the spectrum (equivalent to the RMSE:mean ratio in log space), such that the whole spectrum is weighted evenly.

4.3.3 Diagnosis of realistic ocean mesoscale eddy fluxes

In addition to the variability driven by turbulent atmospheric processes, processes creating variability exist within the ocean interior which are also unresolved by our laminar ocean-only model, due to the coarseness of its spatial discretisation. To show this, we utilise spatiotemporal Reynolds averaging, in which large-scale temperature variations are potentially impacted by small-scale anomalies in a purely advective transport framework. For the temperature, the associated advection equation (at high Péclet number, such that diffusive processes can be neglected) reads

$$\begin{aligned} \partial_t T + \langle \boldsymbol{V} | \nabla T \rangle &= 0, \\ \partial_t (\widehat{T} + \widetilde{T}) + \langle \widehat{\boldsymbol{V}} + \widetilde{\boldsymbol{V}} | \nabla (\widehat{T} + \widetilde{T}) \rangle &= 0, \end{aligned} \tag{4.34}$$

where T and $\boldsymbol{V}$ are the scalar and tridirectional vector fields of temperature and of velocity, respectively, $\nabla\cdot$ is the tridirectional gradient operator, $\langle\cdot|\cdot\rangle$ is the inner product, $\widehat{\cdot}$ is a tridirectional spatial averaging operator, and $\widetilde{\cdot}$ is its associated spatial fluctuation. This separation is such that the lower-resolution model (LRM) is able to resolve temperatures at the scale of the spatial average (e.g. $\widehat{T}$), while the higher-resolution model (HRM) resolves the sum of the spatial average and its fluctuation (e.g. $T = \widehat{T} + \widetilde{T}$). We are interested in the mean effect of the small scale on the large scale following application of the spatial averaging operator. Applying this operator, the equation reduces to

$$\partial_t \widehat{T} + \langle \widehat{\boldsymbol{V}} | \nabla \widehat{T} \rangle = -\langle \widehat{\widetilde{\boldsymbol{V}} | \nabla \widetilde{T}} \rangle. \tag{4.35}$$

As before, we consider the large-scale climatological cycle to be common to the HRM and the LRM, so we separate (4.35) into a trajectory ($\overline{\cdot}$) and a temporal fluctuation ($\cdot'$):

$$\partial_t \left(\overline{\widehat{T}} + \widehat{T}' \right) + \left\langle \overline{\widehat{\boldsymbol{V}}} + \widehat{\boldsymbol{V}}' \middle| \nabla \left(\overline{\widehat{T}} + \widehat{T}' \right) \right\rangle = - \left[\overline{\langle \widehat{\widetilde{\boldsymbol{V}} | \nabla \widetilde{T}} \rangle} + \langle \widehat{\widetilde{\boldsymbol{V}} | \nabla \widetilde{T}} \rangle' \right], \tag{4.36}$$

where $\partial_t \overline{\widehat{T}} + \langle \overline{\widehat{\boldsymbol{V}}} | \nabla \overline{\widehat{T}} \rangle = -\overline{\langle \widehat{\widetilde{\boldsymbol{V}} | \nabla \widetilde{T}} \rangle}$ is the trajectory component common to both LRM and HRM. The unresolved component in the LRM is therefore

$$\partial_t \widehat{T}' + \langle \overline{\widehat{\boldsymbol{V}}} | \nabla \widehat{T}' \rangle + \langle \widehat{\boldsymbol{V}}' | \nabla \overline{\widehat{T}} \rangle = -\langle \widehat{\widetilde{\boldsymbol{V}} | \nabla \widetilde{T}} \rangle', \tag{4.37}$$

where smaller, second order time-fluctuating terms ($\langle \widehat{\boldsymbol{V}}' | \nabla \widehat{T}' \rangle$) have been neglected. The flux terms on the left-hand side represent large-scale interactions between the temporal mean and its fluctuations, while the right hand side describes the temporal fluctuation of small-scale fluxes. We note that the latter term also contains interactions between the climatology and fluctuations (for instance, seasonal variations in eddy temperature), as can be seen by separating its interior velocity and temperature components into their own time-mean and fluctuating terms. Of those

various flux terms, we are interested in the stationary eddy-driven component (i.e. the transport of small-scale buoyancy fluctuations by small-scale current fluctuations). Thus, neglecting large-scale and seasonal–sub-seasonal interactions, we are left with the (eddy-driven) internal-flux-induced rate of change in temperature:

$$\mathrm{F}_{\mathrm{T}}^{\mathrm{I}} = -\langle \widehat{\widetilde{\boldsymbol{V}}'|\nabla\widetilde{T}'} \rangle' , \tag{4.38}$$

where similar considerations may be made for the internal eddy-driven salt flux ($\mathrm{F}_{\mathrm{S}}^{\mathrm{I}}$).

We apply this approach to determine the turbulent eddy heat and salt fluxes unresolved in our LRM (the linear model of Section 4.3.1) using an eddy permitting ocean model (the HRM). In particular, NEMO (v3.5) was run for twenty years in its $1/4°$, 75-level ORCA025 configuration, with climatological forcing. The configuration effectively mirrors that of Grégorio et al. (2015), who produce the forcing by creating a mean year from the Drakkar Forcing Set (Brodeau et al., 2010). A smoothly forced ocean-only model was chosen to minimise the impact of turbulent atmospheric fluxes (which were determined separately; Section 4.3.2). The spatial averaging of (4.34) was undertaken by averaging all gridpoints in the HRM which fall within a single grid cell of the LRM.

As in the external case, the time series of internal turbulent fluxes [$\mathrm{F}_{\mathrm{T}}^{\mathrm{I}}$ and $\mathrm{F}_{\mathrm{S}}^{\mathrm{I}}$, following (4.38)] were used to determine (via the lag-autocorrelation e-folding time) $\boldsymbol{\lambda}_{\mathrm{I}}$ and (via the covariance with other locations) $\boldsymbol{\Sigma}_{\mathrm{I}}$. Owing to the much greater number of elements in $\boldsymbol{\Sigma}_{\mathrm{I}}$ due to the vertical dimension, technical constraints prohibit a fully global treatment of spatial covariance. Instead, we assume spatial covariance to occur only locally: within a $(3\times2°)^2=(6°)^2$ area (i.e, in a nine-point horizontal neighborhood of each location), and throughout the corresponding vertical. Features larger than this would be resolved by the LRM (e.g. Griffies and Treguier, 2013). This assumption allows us to use a sparse matrix representation of $\boldsymbol{\Sigma}_{\mathrm{I}}$, reducing computational demand to the same order as that of $\boldsymbol{\Sigma}_{\mathrm{E}}$.

The temperature ($\mathrm{F}_{\mathrm{T}}^{\mathrm{I}}$; Fig. 4.1g) and salinity ($\mathrm{F}_{\mathrm{S}}^{\mathrm{I}}$; Fig. 4.1h) components of the subgrid fluxes can be seen to exhibit generally similar variance distributions, with almost indistinguishable decorrelation timescales (Fig. 4.1, contours). Common to both components is the strong imprint of the Gulf Stream, Agulhas, Zapiola gyre, and Kuroshio. Their covariance (Fig. 4.1i) emphasises these common regions and is effectively everywhere positive, while salinity flux variability uniquely shows strong signatures in the Amazon and Niger outflow regions. There is some latitudinal dependence of decay time (as may be expected from the changing deformation radius, e.g. Chelton et al., 1998) but decay times $\boldsymbol{\lambda}_{\mathrm{I}}$ largely reflect the variance itself, $\boldsymbol{\Sigma}_{\mathrm{I}}$. For example, the shortest times (on the order of days), at the Equator, may also be found at much higher latitudes in turbulent regions such as the Gulf Stream. Meanwhile, the gyre interiors show greater persistence, up to many months in the Pacific, and these

are the regions where the fluxes are also weakest. These quiescent, persistent regions are understandably where a constant-spectrum approximation (with instantaneous decay) fits most poorly. Consistently this is where the greatest improvements are seen when moving from a Gaussian white noise representation to an Ornstein-Uhlenbeck process representation (Section 4.3.2; Fig. 4.2).

4.3.4 Experiment design

As described in Section 4.2, we can use our linear model configuration and stochastic approach to analyse the variance evolution of any linear, scalar-valued function of the ocean state, in both a theoretical (optimised stochastic representation) and realistic (diagnosed stochastic representation) context. We choose to focus on a range of climatically relevant metrics: the meridional volume transport (MVT, integrated from the surface to the depth of maximum overturning), full-depth meridional heat transport (MHT), and ocean heat content (OHC, over the present depth range of the majority of the Argo fleet, 0-2000 m). These metrics are calculated for the subtropical (at 25°N for MVT [0-870 m] and MHT [full depth], from 15° to 40°N for OHC [0-2000 m]) and subpolar (at 55°N for MVT [0-1200 m] and MHT [full depth], from 40° to 65°N for OHC [0-2000 m]) North Atlantic. In all cases, monthly, annually, and decadally averaged quantities are considered.

4.4 Results

4.4.1 Subtropical North Atlantic

4.4.1.1 Optimal Stochastic Perturbations

We now consider (using the limiting cases of Section 4.2.2) the spatially correlated external and decorrelated internal OSPs of the metrics of Section 4.3.4 in our linearised ocean model (Section 4.3.1). The sensitivity of the metric to different potential sources of variability is indicated by the response ratio (Table 4.1), following (4.26). For instance, the correlated surface zonal momentum flux OSP of yearly MHT has a response ratio of 0.228×10^{-3} $\mathrm{PW}^2(\mathrm{cms}^{-1})^{-2}\mathrm{s}$. This implies that a stochastic surface heat flux following the correlated OSP which has a magnitude of 1 $(\mathrm{cms}^{-1})^2\mathrm{s}^{-1}$ will induce a response in annual averaged MHT with a variance of 0.282×10^{-3} PW^2 across a large ensemble. This noise amplitude is purely illustrative, as the response ratio is a linear rescaling which is independent of the strength of the noise. The response ratios for MVT and MHT suggest a change in regime when averaging times are increased. For these metrics, sensitivity to large-scale spatially correlated buoyancy fluxes at the surface remains relatively constant at all timescales, producing a response

TABLE 4.1: Normalised response ratios of various OSPs [following (4.26)] for the North Atlantic subtropical region. The unit of the response ratio is that of output variance (given in the row header) divided by that of the normalised noise amplitude (given in the column header). A stochastic forcing with the spatial distribution of the OSP (Figs. 4.3, 4.4, and 4.5) and with unit amplitude will thus stimulate a response in the target metric with the given response variance. Note that, in line with (4.26), the response ratio is independent of the amplitude of the noise in the linear framework. The table shows unit amplitudes for illustrative purposes. Left hand columns correspond to perfectly correlated surface OSPs. Right-hand columns correspond to totally uncorrelated internal OSPs. Note that the amplitude units differ between the correlated and uncorrelated cases.

		Surface (correlated)				Full-depth (uncorrelated)	
	avg. time	T ($\times 10^{-3}$ K^{2}s^{-1})	S ($\times 10^{-4}$ psu^{2}s^{-1})	u ((cms^{-1})^{2}s^{-1})	v ((cms^{-1})^{2}s^{-1})	T ($\times 10^{-3}$ K^{2}d^{-1})	S ($\times 10^{-4}$ psu^{2}d^{-1})
MVT (Sv^2)	30d	0.229	0.599	0.143	0.079	21.740	53.858
	1y	0.215	0.558	0.051	0.038	12.018	30.426
	10y	0.186	0.495	0.040	0.033	2.935	7.812
MHT ($\times 10^{-3}$PW^2)	30d	1.162	2.729	1.876	0.524	82.882	201.714
	1y	1.099	2.616	0.282	0.227	53.946	134.327
	10y	0.886	2.160	0.195	0.145	13.652	37.132
OHC ($\times 10^{-3}$K^2)	30d	2.744	6.602	0.194	0.225	12.886	42.240
	1y	2.766	6.573	0.193	0.224	12.676	41.625
	10y	2.551	6.217	0.182	0.212	10.380	35.564

of similar amplitude. Conversely, sensitivity to internal, spatially decorrelated buoyancy fluxes falls sharply with increasing average time, particularly for MVT. Surface momentum flux sensitivity also sees a sharp decline from monthly to annual timescales for both MVT and MHT. OHC variability exhibits no apparent regime shift of this nature, with a steady sensitivity to changes in all variables across all timescales.

So as to understand the mechanisms of variability generation in the model, we now consider the spatial distribution of the perturbations (for year-averaged quantities) in more detail (Figs 4.3, 4.4, and 4.5). The optimal perturbations for MVT and MHT (Figs 4.3 and 4.4; shading) are broadly similar. In the uncorrelated, internal case (panels a and b) the perturbation can have no large-scale structure and simply reflects the distribution of sensitivity amplitudes. These are greatest in the Gulf Stream, and along the evaluation line of the metrics. The large-scale patterns of the correlated external buoyancy forcing, however, reflect strongly the model mean state. In particular, subtropical meridional transport variability displays a strong sensitivity to subpolar surface buoyancy fluxes, reflected as a large-scale gradient across the northern boundary of the subtropical gyre. Wind sensitivity displays very consistent patterns

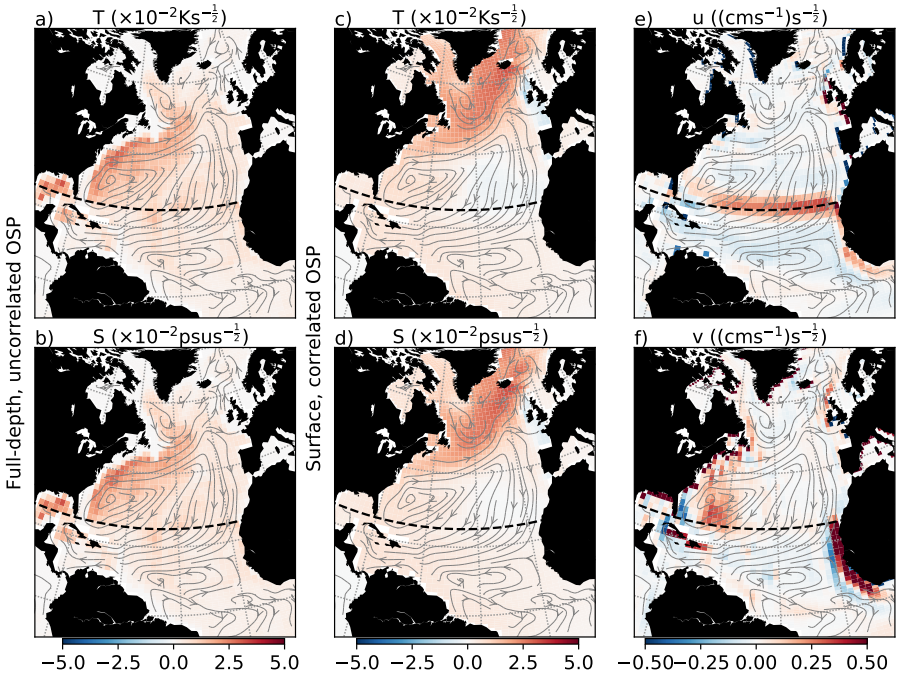


FIGURE 4.3: Optimal stochastic perturbation for year-averaged subtropical meridional volume transport in the fully spatially uncorrelated (a,b, depth-averages shown) and perfectly spatially correlated surface-only (c,d,e,f) cases. Streamlines show time-averaged volume transport over the upper 2 km in the trajectory. Dashed lines show 25°N, the latitude at which the meridional volume transport is evaluated.

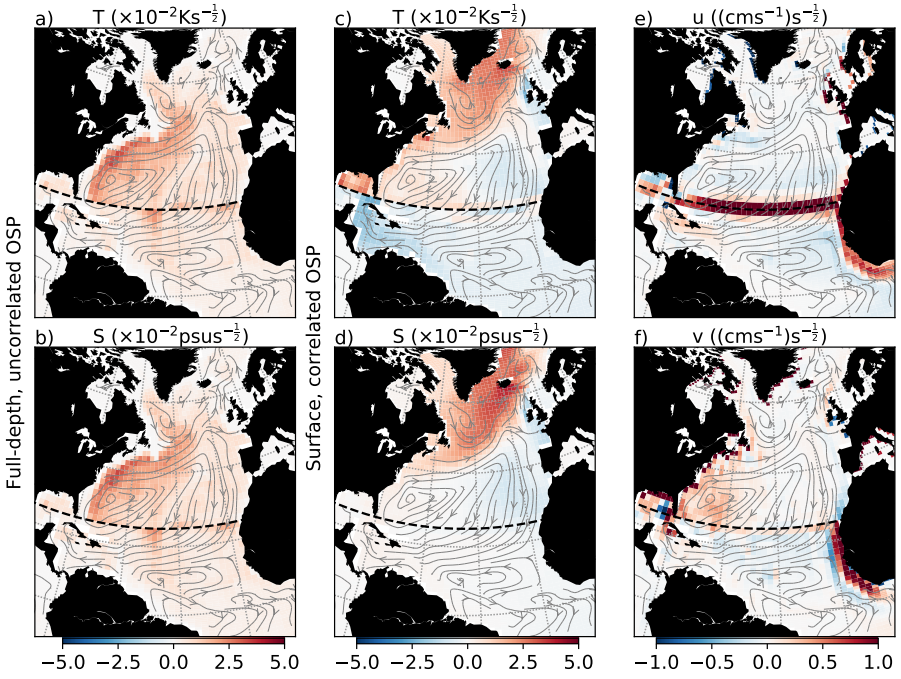


FIGURE 4.4: As in Fig. 4.3, but for subtropical meridional heat transport (evaluated at 25°N, denoted by the dashed line).

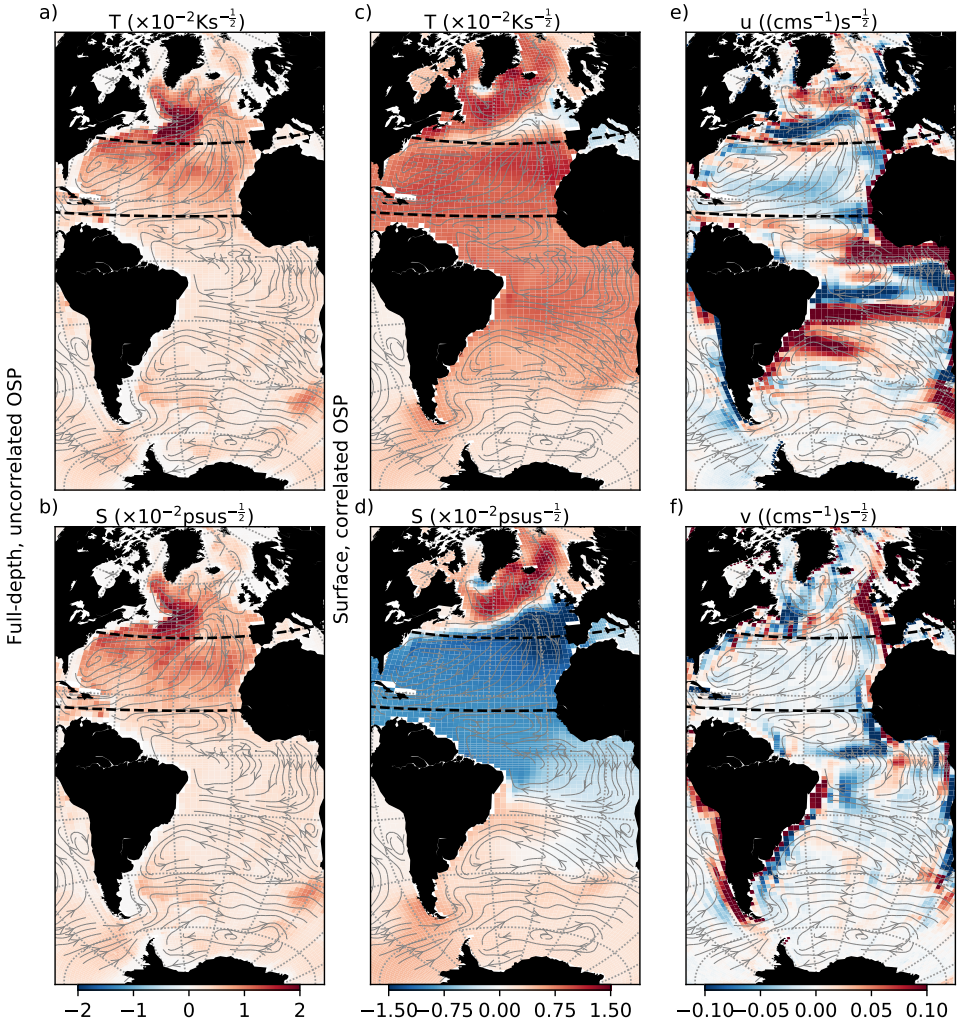


FIGURE 4.5: As in Fig. 4.3, but for subtropical ocean heat content (evaluated between 15°N and 40°N denoted by the two dashed lines).

indicating stimulation of Ekman transport (in the case of zonal wind) and western boundary transport change combined with eastern boundary up- or downwelling (in the case of meridional wind). Upwelling directly impacts the volume transport locally through geostrophy (Hirschi et al., 2007; Kanzow et al., 2010; Polo et al., 2014), but this pattern has also been observed in other sensitivity studies to trigger pressure anomalies which reach great distances along the eastern boundary (Pillar et al., 2016; Jones et al., 2018).

The optimal OHC perturbation in the uncorrelated case (Fig. 4.5a and b) shows sensitivity to buoyancy fluxes throughout the region, but particularly at the subpolar–subtropical gyre interface, which has been highlighted as a key region for variability generation in the Atlantic (Buckley and Marshall, 2016). Also clear, but less pronounced, are local peaks around the Agulhas retroflection and the Zapiola gyre. The correlated surface OSPs are notably different in the cases of temperature and salinity due to the ability of surface temperature fluxes to impact heat content

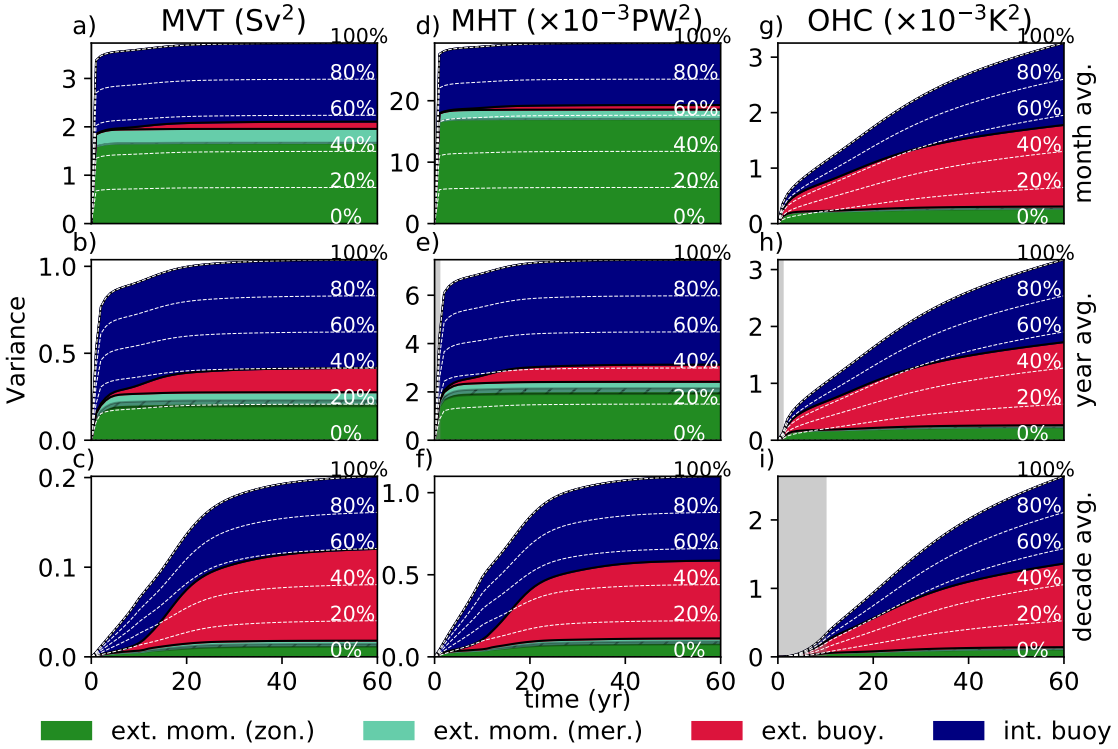


FIGURE 4.6: Attribution of uncertainty following initialisation for the subtropical ocean metrics (MVT: a-c; MHT: d-f; OHC: g-i) over different averaging times (month: a,d,g; year: b,e,h; decade: c,f,i), following (4.31) and (4.32). Green and red shading indicate variance due to external (atmospheric) momentum and buoyancy fluxes, respectively. Blue shading indicates variance due to internal buoyancy fluxes (due to oceanic mesoscale eddy forcing). Dashed white contours show percentages (inset text) of the total variance. Shaded boxes show the averaging window over which the metric is evaluated. Variance due to surface momentum fluxes is partitioned into zonal (dark green) and meridional (light green) components, where shading between them indicates covariance.

variability both directly and indirectly through passive and active mechanisms, which sometimes conflict (Stephenson and Sévellec, 2021a). The active mechanisms are made clear by the correlated salinity OSP, which shows stark gradients across the northern boundary of the North Atlantic and South Atlantic subtropical gyres, as well as a local peak in the deep water formation region of the model (Stephenson et al., 2020). The temperature perturbation echoes this, but with a distribution which is almost everywhere equally signed, so as to passively stimulate heat content. The momentum flux perturbations (Fig. 4.5e and f) are generally more complex but can still be seen to broadly coincide with predominantly zonal streamlines and coastal regions in the zonal and meridional cases, respectively. There is a notable focus along the subpolar–subtropical gyre interface for the zonal momentum flux.

4.4.1.2 Dynamical attribution of subtropical variance

Having explored the patterns and mechanisms by which oceanic variability can be optimally stimulated in our model, we turn our attention to the ways in which it is actually stimulated in the real climate system, as derived in Section 4.2.3. Following (4.31) and (4.32), application of each component of the stochastic forcing separately allows the resultant variance evolution to be partitioned accordingly (Fig. 4.6). There is a substantial difference between the nature of month- and decade-averaged transport metrics, both in the variance amplitude and in the impacts of different sources, as in the OSP case (shown by the amplification factors of Table 4.1). External momentum fluxes are responsible for 52% of month-averaged MVT and for 63% of month-averaged MHT by the end of the 60-yr simulation, but just 9% and 10%, respectively, for decade-averaged MHT. Similarly, the external buoyancy component contributes just 4% to month-averaged MVT variance at 60-yr, but over 50% in the decade-averaged case. For year-averaged MVT and MHT, the ocean internal component is the dominant contributor to the final variance, at 60% for MVT and 58% for MHT. In addition to differences between monthly and decadal metrics in the final (60 yr) variability, a difference in the evolution of this variance is also apparent. Contributions from all sources are fairly steady in time for MVT and MHT for the quickly-saturating month-averaged case. For ten-year average MVT and MHT, there is a more notable shift. Following initialisation, external momentum and internal buoyancy fluxes are the main causes of error growth. However, the contribution of wind peaks abruptly, while the eddy component grows for around 6 years, peaking at nearly 80% of the total uncertainty. On longer timescales, the eddy-turbulence component falls to slightly less than half of the total contribution over the remainder of the simulation. During this stage, it is the more slowly acting external buoyancy component that develops and contributes the remaining variance.

Notably, as in the OSP perspective, the components of the OHC variance after 60 yr are consistent across different time averages, with an almost equal contribution (around 45% each) from external and internal buoyancy fluxes. This follows the slow growth of the internal component, which, at its lowest, contributes only around 25% of the total uncertainty. This is in contrast with the MVT and MHT, where it is the external buoyancy contribution which is the slowest to develop.

Following (4.33), we consider the spatial distributions of these contributions to the 60-yr variance for the annually averaged case, within the transition between the two discussed (month- and decade-average) cases (Fig. 4.7; where the zonal and meridional momentum flux covariance contribution is not shown). There is generally a high level of agreement between the patterns shown in the optimal case (i.e. what the ocean "wants"; Figs 4.3, 4.4, and 4.5) and the realistic case (Fig. 4.7). This is linked to the overall relative constant shape of the realistic forcing (i.e. what the ocean "gets";

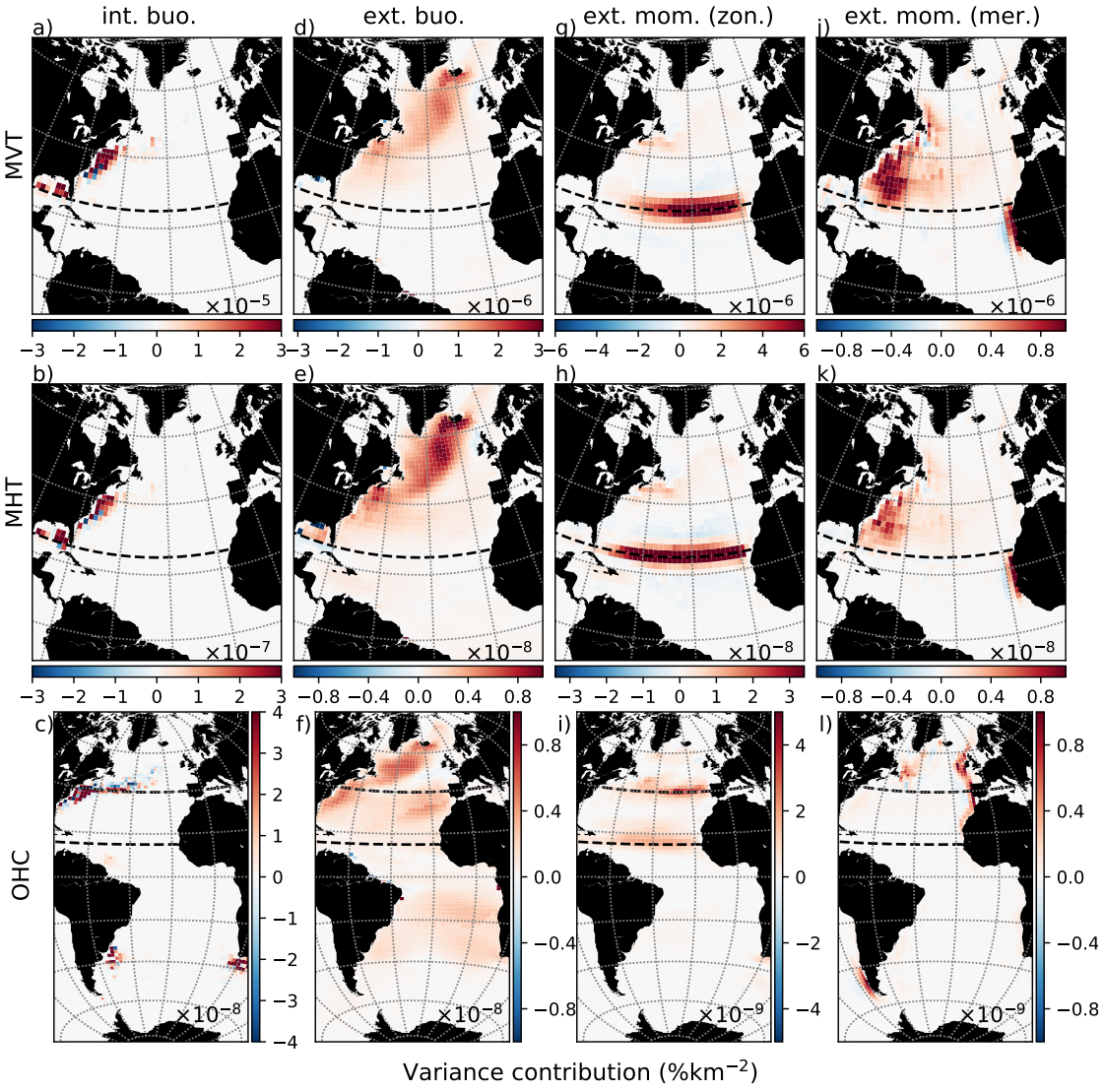


FIGURE 4.7: Spatial distribution of sources of accumulated variance for subtropical ocean metrics (MVT:a,d,g,j; MHT:b,e,h,j; OHC:c,f,i,l) after 60 years of simulation following (4.31), (4.32), and (4.33). Variance per unit volume due to internal buoyancy fluxes is depth integrated to give the water column total contribution per unit area (a-c), variance due to external momentum (zonal component, g-i; meridional component, j-l) and buoyancy (d-f) fluxes are surface distributions of contribution per unit area. Dashed lines show the latitude (MVT, MHT metrics) or region (OHC metric) where the metric is evaluated. Note the differing (sometimes by orders of magnitude) color scales, reflecting the differing contributions shown in Fig. 4.6.

Fig. 4.1). Although we stress the contrast between the two frameworks (i.e. white vs. temporally correlated noise) when making any such comparisons.

In particular, volume and heat transport variability are primarily driven by ocean internal buoyancy fluxes local to the western boundary, and by remote external buoyancy fluxes in the subpolar region. Zonal surface momentum fluxes, consistently with the OSP, almost exclusively stimulate a zonal band along the evaluation line (Fig. 4.7g and h), while in the meridional case a combination of western boundary current and eastern along-shelf stimulation pervade. The agreement between the prescribed (temporally correlated) and optimal (white noise) forcing is less apparent in the case of OHC. Internal buoyancy fluxes affecting heat content variability can be predominantly traced in the prescribed case to highly focused sources in the noisiest regions of the Atlantic (Fig. 4.7c vs. Fig. 4.1g and h), while the optimal white noise perturbation is more evenly distributed throughout the Atlantic with a local peak in the subtropical–subpolar "transition zone" (Buckley and Marshall, 2016). The distribution in the prescribed case also exhibits a selection of locations which make a negative contribution, particularly north of the North Atlantic current. These arise from the covariance of neighbouring points with an otherwise strong contribution gradient, and act as a compensatory "source" of predictability relative to that which would stem from a spatially decorrelated representation. External buoyancy fluxes contribute over a broader area than the internal case, with the most concentrated contributions in the remote subpolar region. The contribution from zonal wind is almost exclusively along the evaluation region's boundaries, whereas in the meridional case (as also seen in the OSP) the western coasts of Europe and South America have the clearest impact.

4.4.2 Subpolar North Atlantic

4.4.2.1 Optimal Stochastic Perturbations

Applying the considerations of Section 4.4.1.1 to the subpolar region, differences emerge in the amplitude of the response to the optimal stochastic forcing (Table 4.2). For subpolar MVT, the correlated surface OSP is much more effective at generating variability than in the subtropics, particularly on annual timescales (for which the response ratio is around four times as large as in the subpolar region). For MHT, the values are similar in both regions. The opposite is apparent in the spatially uncorrelated case, where, for example, the response of monthly MHT to its uncorrelated optimal noise perturbation is over six times as large in the subtropics as in the subpolar region. OHC again shows consistent behaviour across all time averages, but is much more sensitive to external momentum and internal buoyancy changes than in the subtropics.

TABLE 4.2: As in Table 4.1, but for subpolar OSPs, (whose spatial distributions are shown in Figs. 4.8,4.9, and 4.10).

		Surface (correlated)				Full-depth (uncorrelated)	
	avg. time	T ($\times 10^{-3}$ K^{2}s^{-1})	S ($\times 10^{-4}$ psu^{2}s^{-1})	u ((cms^{-1})^{2}s^{-1})	v ((cms^{-1})^{2}s^{-1})	T ($\times 10^{-3}$ K^{2}d^{-1})	S ($\times 10^{-4}$ psu^{2}d^{-1})
MVT (Sv^2)	30d	0.393	1.359	0.264	0.178	7.879	27.599
	1y	0.848	2.086	0.534	0.197	3.793	12.724
	10y	0.296	1.070	0.182	0.115	0.919	2.964
MHT ($\times 10^{-3}$PW^2)	30d	1.705	4.265	0.468	0.384	13.736	36.138
	1y	1.541	4.032	0.409	0.344	11.238	29.837
	10y	0.831	2.249	0.192	0.154	3.692	10.446
OHC ($\times 10^{-3}$K^2)	30d	4.764	12.347	1.990	1.462	48.312	147.220
	1y	4.709	12.295	1.982	1.457	47.853	145.972
	10y	4.327	11.294	1.869	1.392	38.633	121.612

The OSP for meridional volume transport (Fig. 4.8) shows a much more concentrated spatial distribution than its subtropical equivalent. In the uncorrelated ocean interior case, almost all of the weight is focused at the core of the subpolar gyre (panels a and b). For the perfectly correlated surface case, this hotspot, coincident with the surface outcrop of the model's North Atlantic Deep Water (Stephenson et al., 2020), is complemented by a dipole pattern crossing the North Atlantic Current (panels c and d). This dipole resembles the surface sensitivity of the least damped interdecadal mode of variability (corresponding to a large-scale thermal Rossby wave) present in an earlier version of the model (Sévellec and Fedorov, 2013a). As with the subtropical metric, the optimal momentum flux patterns are an east-west band in the zonal case and a predominantly eastern-boundary-following pattern in the meridional case.

While having many common features with that of MVT, the optimal pattern for MHT (Fig. 4.9) is much less focused, neglecting the hotspot of the north-west Atlantic for a more spread out distribution.

The optimal internal perturbation consists of buoyancy fluxes throughout the subpolar gyre, as well as in the subtropical–subpolar intergyre region. In the correlated case, the dipole feature between gyres (already visible for the subtropical case, Fig. 4.4) is more heavily emphasised. In addition to the familiar features of the velocity OSPs, fainter bands encircle the subpolar gyre.

The OSPs of subpolar OHC variance (Fig. 4.10) exhibit many similar behaviours to those described for other metrics. The uncorrelated interior noise favours the subtropical–subpolar gyre boundary, while the correlated surface heat flux pattern

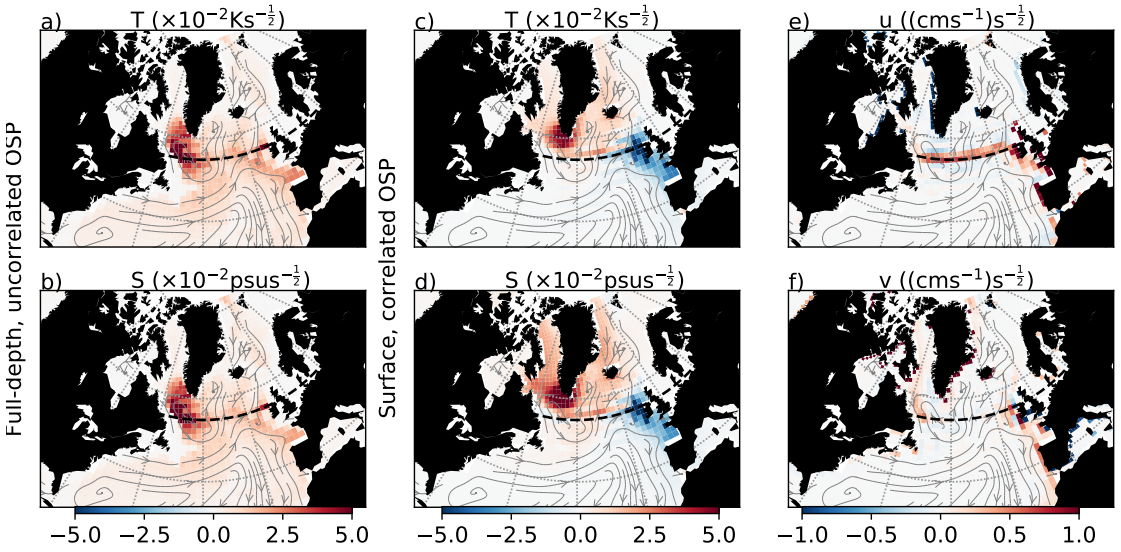


FIGURE 4.8: As in Fig. 4.3, but for subpolar meridional volume transport (evaluated at 55°N denoted by the dashed line).

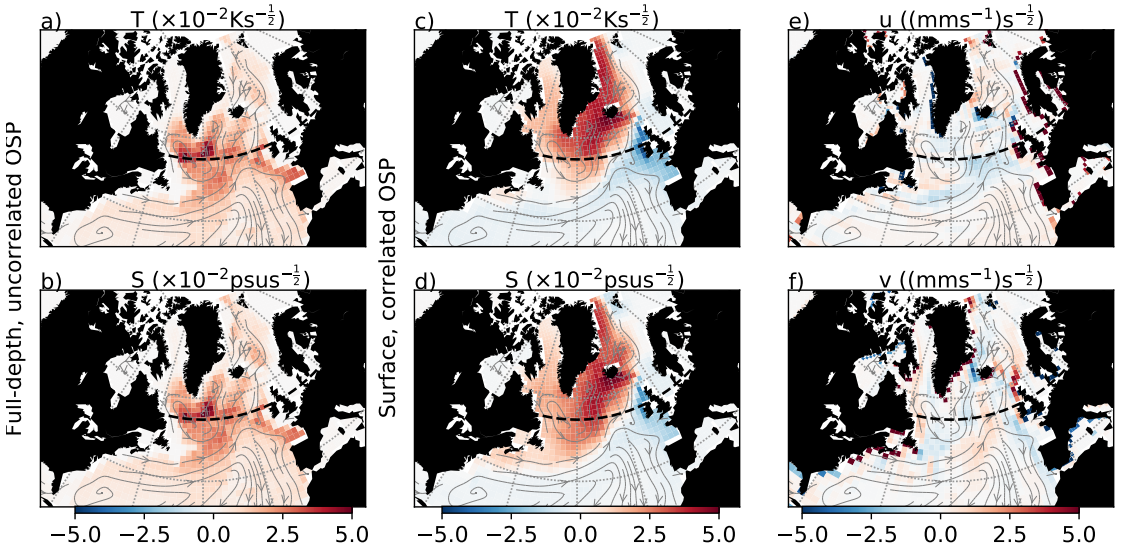


FIGURE 4.9: As in Fig. 4.3, but for subpolar meridional heat transport (evaluated at 55°N denoted by the dashed line).

targets oppositely the deep water outcrop regions and the wider North Atlantic, with a particular focus on the North Atlantic Current. Similarly to the correlated OSP of subtropical heat content, the correlated subpolar zonal velocity OSP displays a complex arrangement of alternating bands which broadly coincide with strongly zonal currents in the trajectory, while the meridional pattern predominantly targets coastal upwelling and downwelling (i.e. alongshore velocity/momentum fluxes) in these same regions.

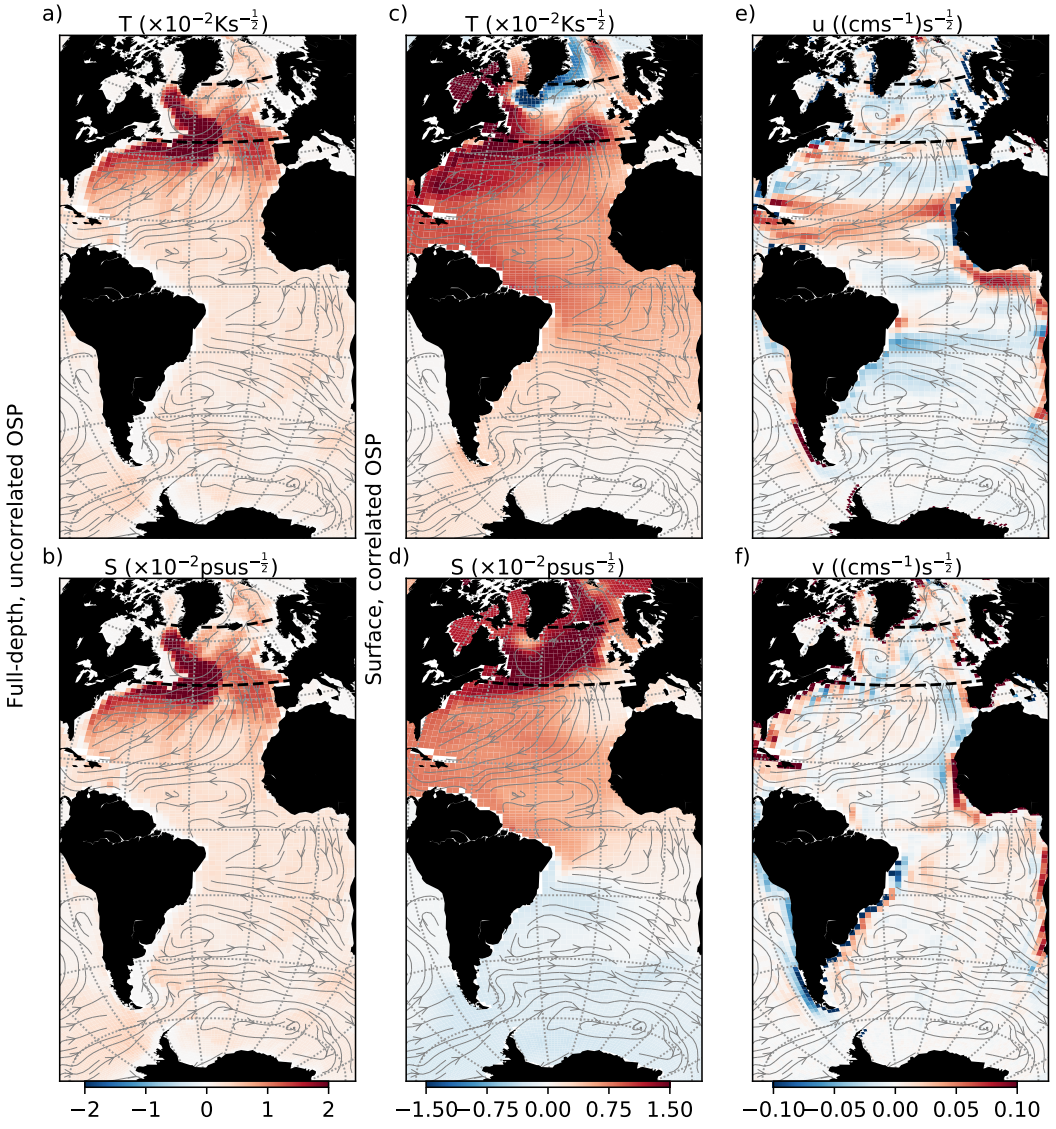


FIGURE 4.10: As in Fig. 4.3, but for subpolar heat content (evaluated between 40°N and 65°N denoted by the two dashed lines).

4.4.2.2 Dynamical attribution of subpolar variance

Under prescribed, realistic sources of variability, the subpolar region is dominated by external forcing (Fig. 4.11), which accounts for up to 94% of the total variance after 60 years in the case of month-averaged heat transport. As in the subtropics, the meridional transport metrics exhibit a regime shift when moving from month-averaged quantities (up to 86% momentum-driven) to decade-averaged quantities (where over 60% of the final variance can be attributed to surface buoyancy fluxes). For all time averages, momentum fluxes contribute most of the early-stage error growth of MVT and MHT following initialisation, but the buoyancy component becomes more established over the first decade. MVT and MHT are much less variable overall than in the subtropics, while heat content variance is slightly higher, again showing

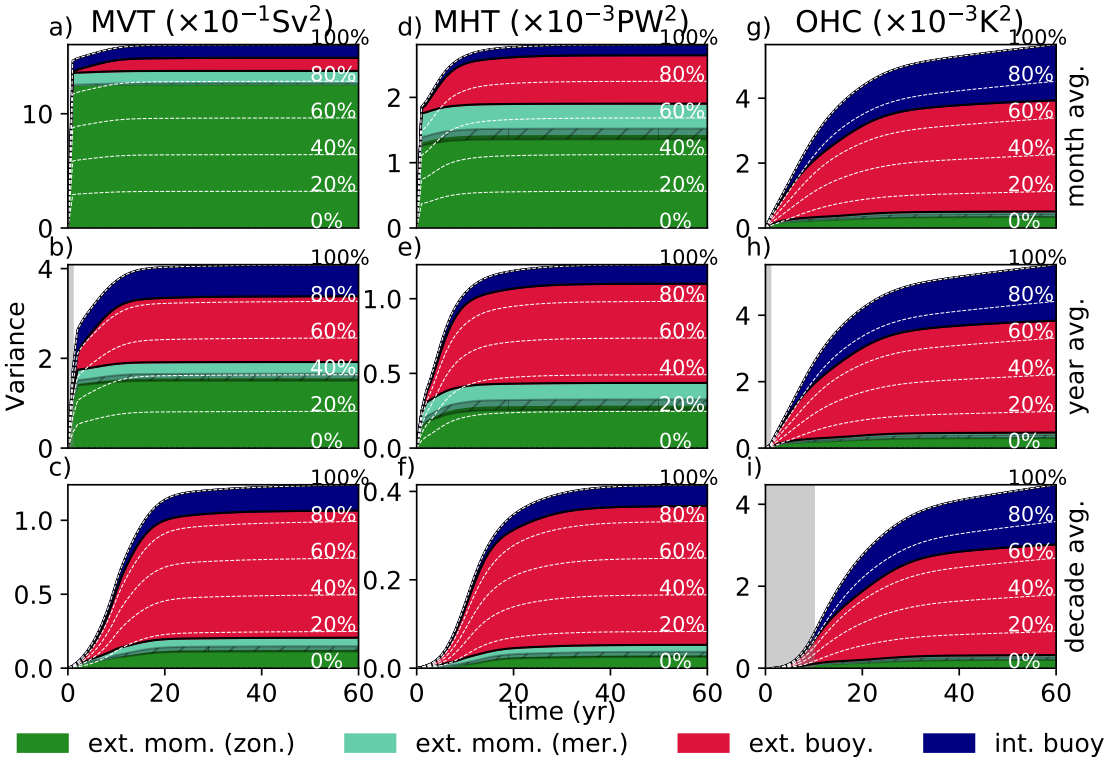


FIGURE 4.11: As in Fig. 4.6, but for subpolar ocean metrics

consistent behaviour across all considered time averages. Also notable is that, despite full convergence not being reached after the 60 years, heat content seemingly shows a higher degree of saturation in the subpolar region than in the subtropical region.

The spatial patterns of subpolar variance origins in response to prescribed fluxes (Fig. 4.12; where the zonal and meridional momentum flux covariance contribution is not shown) are generally less similar to the corresponding optimal perturbations (Section 4.4.2.1) than in the subtropics (Section 4.4.1), although we again treat comparisons between the two frameworks with caution. The differences are particularly clear for internal buoyancy fluxes, which for all metrics share a common maximum at around 40°N, far south of the corresponding peaks in the uncorrelated OSPs. For MVT there is a large contribution on the evaluation line west of Scotland, apparently coincident with a local peak in the uncorrelated OSP, but the most sensitive region in the central subpolar gyre is only weakly stimulated. As in the subtropical region, negative contributions flank the Gulf Stream and its extension, acting as a compensatory "source" of predictability offsetting its covarying sinks. Variance due to (temporally correlated) prescribed external buoyancy fluxes more closely agrees with the (white noise) spatially correlated OSP. In particular, the northern portion of the optimal dipole shape is discernible for MHT, while the deep water outcrop hotspot can be faintly recognised, along with the west-European shelf in the case of MVT. Heat content variability due to external buoyancy fluxes largely coincides with the most

concentrated region of the correlated OSP, in the North Atlantic Current, but shows little agreement elsewhere. The external momentum flux components are qualitatively similar for all three metrics, again stimulating transport across constant latitude lines in the zonal case (where the noise input [Fig. 4.1] constructively stimulates the most sensitive regions [Figs 4.8, 4.9, and 4.10]), while highlighting the coasts for the meridional case. Both zonal (in the case of MVT and OHC) and meridional (in the case of MHT) momentum flux contributions are offset by a weakly negative compensation bordering the regions of strongest positive variance stimulation.

4.5 Discussion and conclusions

The climate system contains a number of sinks of predictability or, equivalently, sources of uncertainty, from which unpredictable noise can grow and eventually overwhelm predictable signal (such as that provided to an initialised forecast). In this study, we have considered the sources (and compensatory sinks) of uncertainty in metrics of the North Atlantic from two perspectives. In the first perspective, a complementary pair of optimal stochastic forcings were calculated, encapsulating the patterns which generate maximum variance in the metric. These are a representation of the sensitivity of the metric to random forcing. The pair differ by their spatial coherence: one being fully spatially uncorrelated, with the other fully correlated over the surface layer. These are the extrema of possible spatial correlation, and respectively mimic, in an idealised sense, the behaviour of stochastic fluxes due to (mesoscale) oceanic turbulence and (synoptic scale) atmospheric turbulence. In the second perspective, the optimal stochastic forcing is instead replaced with a prescribed, realistic stochastic representation of these sources, including spatiotemporal covariance. The properties of the representations are diagnosed from more complex (fully coupled and eddy-permitting) models. This has allowed us to compare the commonalities between the optimal and actual cases (albeit in a limited way, given their differences in spatiotemporal correlation). We have further been able, in the diagnosed case, to dynamically attribute variability to its origins. The latter ability notably forgoes the more typical ensemble attribution approach, which generally necessitates many simulations in a high-complexity model, and cannot ensure causality. The sources determined by these two perspectives can qualitatively be seen as what the ocean “wants” (in order to maximise variability) and what the ocean “gets” (in the real world). Regions where the ocean “gets” what it “wants” offer particularly poor prospects for prediction, as both the sources of noise and their mechanisms of translation into large-scale uncertainty play a role.

Variations on the OSP technique have been utilised in the context of optimal excitation of MOC variability in a number of studies (a thorough review is provided by Monahan et al., 2008). However, due to the complexity of the problem these studies are typically

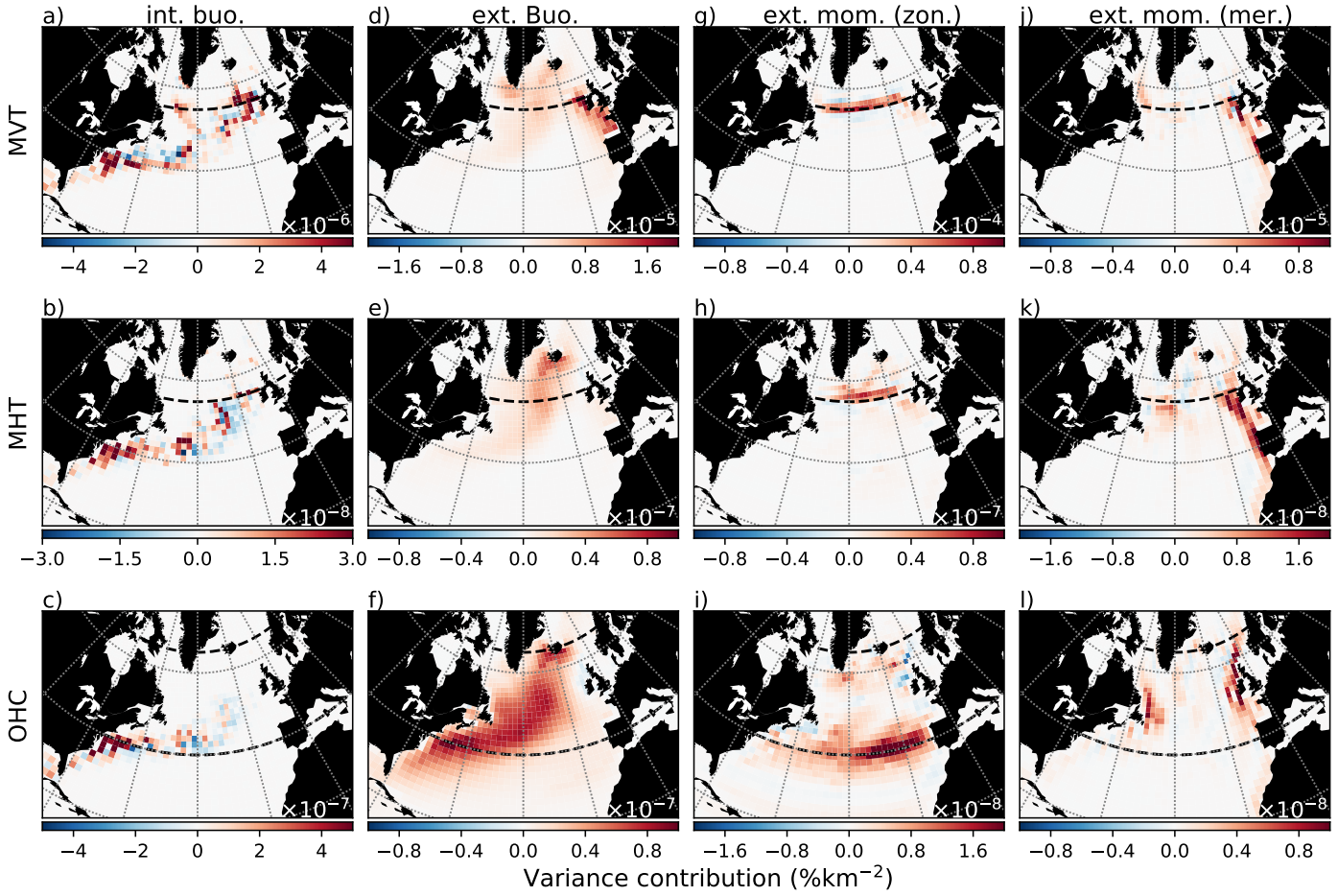


FIGURE 4.12: As in Fig. 4.7, but for subpolar ocean metrics. We note again that the differing contributions (as shown in Fig. 4.11) lead to large differences in the color scales between panels.

undertaken in an idealised context, utilising either box models (e.g. Tziperman and Ioannou, 2002; Zanna and Tziperman, 2008) or idealised ocean models (e.g. Sévellec et al., 2007, 2009). We have adapted the framework to a global OGCM by reducing the covariance matrix to block diagonal form and considering its limiting cases. We note (e.g. Farrell and Ioannou, 1996) the close relationship between optimal stochastic forcings and optimal initial perturbations: the former is in a sense a linear combination of the latter such that the coefficients are determined by the OSP approach. As the linear optimal perturbation of a linear ocean metric is simply a rescaling of the adjoint sensitivity field (Sévellec et al., 2007), we may consider the sources highlighted by the OSP in the context of past adjoint sensitivity studies, where they appear robust across differing models, metrics, and time scales. Recurring mechanisms evident in our study include, for instance, the along-shelf stimulation by meriodional wind and subsequent triggering of coastal pressure anomalies, particularly along the west coast of Africa. This pattern has been stressed by Jones et al. (2018) in an adjoint sensitivity study of Labrador Sea heat content, Loose et al. (2020) regarding heat transport across the Greenland-Scotland ridge, and Pillar et al. (2016) in the context of meridional overturning in the subtropics. The latter study additionally analyses fainter alternating bands of wind stress sensitivity as also seen here, concluding that these communicate pressure anomalies via topographically steered Rossby waves.

Common to the surface thermohaline OSPs of all metrics considered here is a large-scale buoyancy gradient pivoting on the North Atlantic current, which has in dynamical studies been seen to stimulate subtropical (Pillar et al., 2016; Kostov et al., 2019) and subpolar (Sévellec et al., 2017b) volume transport, as well as basin-wide (Sévellec and Fedorov, 2017) and Labrador Sea (Jones et al., 2018) heat content. This is joined by a "hotspot" common to the heat content and subpolar volume transport OSPs in both the correlated and uncorrelated cases which is associated with the passive transport of buoyancy anomalies via deep water pathways (Sévellec and Fedorov, 2015; Stephenson et al., 2020).

To estimate the extent to which these intrinsic ocean sensitivities are exploited by actual sources of stochastic variability, and to quantify the respective contribution of these sources to oceanic uncertainty, we then considered the metrics from the second, prescriptive, perspective. A number of studies have dynamically attributed oceanic changes to prescribed external surface forcings using adjoint methods (Pillar et al., 2016; Sévellec et al., 2018; Smith and Heimbach, 2019) but the relative quantification of internal oceanic mesoscale eddy contributions has thus far been restricted to a resource-intensive ensemble framework (e.g. Bessières et al., 2017). These contributions may present a key sink of predictive skill in high-resolution climate models however, and so are of increasing importance. By incorporating temporal correlation, we have presented a realistic stochastic representation (an Ornstein-Uhlenbeck process) of the slowly evolving ocean mesoscale which can also be projected onto the adjoint

sensitivity fields. This stochastic representation fits the power spectrum of modelled eddy buoyancy fluxes much more closely than Gaussian white noise, which is the more commonly employed framework when considering atmospherically driven low-frequency variability (e.g. the review of Farneti, 2017, and references therein). This has allowed us to bypass the ensemble approach in exchange for the much more numerically efficient dynamical method for both oceanic (internal) and atmospheric (external) sources of error growth.

The diagnosed stochastic forcing approach reveals a regime change in meridional transport variability for longer time averages. In particular, we have shown that surface momentum fluxes dominate for month-averaged transport metrics while surface buoyancy fluxes take over for decade averages. This regime shift is well documented (Dong and Sutton, 2001; Hirschi et al., 2007; Polo et al., 2014) but we find that in the early stages of the error growth, and for annual averages, it is ocean internal buoyancy fluxes, due to mesoscale eddies, which form the greatest contribution in the subtropics. As early-stage growth is when the signal-to-noise ratio diminishes most rapidly, it may be internal sources which present the greatest barrier to subtropical predictability. Our results indicate that these sources ultimately account for up to 60% of annually-averaged volume transport variability at 25°N. This quantification broadly agrees with the varying estimates of ensemble studies (albeit at the higher end; e.g. Grégorio et al., 2015; Jamet et al., 2019), which typically place a local peak in internal oceanic contributions to MVT variability near 25°N (our subtropical metric latitude) with a corresponding trough near 55°N (our subpolar metric latitude) consistently with the decrease we show here. We did not find any such regime shift in the case of ocean heat content, whose variability for all time averages is dominated by external forcing, particularly in the more quiescent subpolar region (consistent with the ensemble study of Sérazin et al., 2017).

When comparing the theoretically deduced (white noise) OSPs with the sources of variability in response to diagnosed (temporally correlated) stochastic forcing, a general overlap was observed in the subtropical region. This suggests efficient stimulation of the preferred mechanisms of the ocean, despite the differing temporal correlation of the two frameworks. This was less true of the subpolar region, which may go some way to explaining the smaller diagnosed variance there relative to the subtropics, despite its higher sensitivity to surface forcing (quantified via response ratios) in the optimal framework. Regarding the subsurface component, it is commonly discussed that the smaller deformation radius at higher latitudes necessitates an ocean model with a fully eddy-resolving resolution in order to faithfully represent the internal contribution. As such, this contribution is likely under-represented in eddy-permitting ensemble studies, which typically portray it as very minor (e.g. Grégorio et al., 2015; Leroux et al., 2018). This lower contribution also impacts our own approach of diagnosing mesoscale eddy fluxes in an eddy-permitting model. However, we reinforce

that even without prescribed forcing, the theoretical OSP framework has allowed us to quantify the subtropical sensitivity to spatially uncorrelated noise as being many times as large as the subpolar region. It is thus apparent that large-scale oceanic metrics are simply less affected by small-scale noise in this region, potentially offering increased benefit from targeted monitoring systems.

Previous studies investigating interactions between the oceanic mesoscale and the low-frequency large scale (such as those considered here) present conflicting behaviour. While some studies show constructive stimulation of low-frequency variability (e.g. Berloff et al., 2007; Arbic et al., 2014), others show its destruction by small-scale noise (e.g. LaCasce and Pedlosky, 2004; Hochet et al., 2020; Sévellec et al., 2020). The framework of our study describes variability from a linear, ensemble perspective in which any divergence in phase space constitutes an irreversible accumulation of error (a source of uncertainty). This framework is not well-suited to isolating such destructive feedbacks, but we have seen that some contributors to the net positive error growth are weakly negative. This slows this growth and restores some predictability. This is particularly apparent along the boundaries of noisy regions such as the North Atlantic current, suggesting a partial compensatory source of predictability within the turbulent internal field.

We finally comment on some other limitations of the approach. While computationally efficient, we have used a linearised model under the assumption of small deviations from a trajectory, alongside a stationary, band-limited stochastic representation of dynamical processes which, in reality, are highly intricate. For example, our internal turbulent buoyancy flux representation cannot encompass coherent inter-basin exchanges, which have been speculated to be an important mechanism of Atlantic MOC variability (e.g. Biastoch et al., 2008). While a coupled climate model was used to determine the surface fluxes, the modelled ocean response is unable to interact with these, precluding the existence of any coupled feedbacks and associated modes, which may have a pronounced impact on interdecadal variability (e.g. Liu, 2012). Despite these drawbacks, the framework offers a uniquely efficient and thorough method for investigating the sources of oceanic variance and associated impacts on predictability. The result is an exact analytical calculation of oceanic uncertainty (otherwise requiring a theoretically infinite ensemble) which can be cleanly partitioned into its sources and locations.

Chapter 5

Optimal thermohaline precursors to the 2009 AMOC slowdown:

use of a nonlinear iterative method in an eddy-permitting model

Abstract

The year from 2009 to 2010 saw an observed 30% reduction in overturning in the subtropical Atlantic, whose thermohaline precursors remain in question. Here, we use an iterative method in an eddy-permitting ocean general circulation model with the help of its adjoint to determine an optimal buoyancy precursor of the 2009 transport decrease. In this context, we particularly focus on the decrease that is not explained by the action of the surface wind field. While the linear optimal perturbation has a negligible effect, the optimal perturbation estimated in a nonlinear context accounts for 93.7% of the outstanding transport decline (that which is not explained by wind) after only 10 iterations of the method. Of note is the favouring by the iterative process of large-scale coherent features over small-scale disturbances, with the presence of the latter in the linearly determined optimal perturbation removed and redistributed among the former in the nonlinear case. Targeted large-scale patterns include remote coastal buoyancy anomalies and local basin-scale zonal density gradients, and are highly efficient, inducing an additional 3.4 Sv decline in the annually averaged volume transport from a perturbation on the order of only 0.1 K in magnitude.

5.1 Introduction

Modern oceanographic monitoring systems have uncovered unforeseen levels of variability in large-scale metrics of the ocean circulation at all observed frequencies. A metric under particular scrutiny is the Atlantic Meridional Overturning Circulation (AMOC), for which several basin-scale mooring arrays have been deployed (Frajka-Williams et al., 2019). These act, in part, as a system to monitor for an AMOC decline, as has been predicted under future anthropogenic emissions pathways (IPCC, 2013). Such a decline could lead via positive feedbacks to a potentially irreversible shutdown (Weijer et al., 2019, provide a thorough review on this topic). Any such downturn could lead to a range of hypothesised climate impacts (Cheng et al., 2013), due in part to the associated decrease in northward heat transport. Confoundingly, the recent observational record has revealed that transient downturns, with varying degrees of persistence, are a systematic feature of the AMOC volume transport, whose range (day-averaged values) is over 150% its mean (Cunningham et al., 2007). The most notable such downturn of the monitoring era was the 2009-2010 event, corresponding to a 30% decline in measured volume transport on average over the period (Bryden et al., 2014). This has been statistically linked to subtropical mixed layer cooling the following year (Cunningham et al., 2013) as well as unprecedented transient sea level rise along the eastern seaboard of North America (Goddard et al., 2015).

The downturn associated with the 2009-2010 event was measured at 26.5°N by the RAPID observational program, which combined a basin-spanning mooring array (measuring bottom pressure and full-water-column temperature and salinity) with western boundary current meters and existing systems to measure Gulf Stream and wind-driven (Ekman) transport. Together, these produced a twice-daily time series of these transport components (upper mid-ocean geostrophic, Gulf Stream, and Ekman transports) at 26.5°N (Kanzow et al., 2007). These components, combined, produced an estimate of the net AMOC transport.

It is understood from these measurements that the event was the result of an uncharacteristic decline in Ekman transport associated with the extreme North Atlantic Oscillation (NAO) state of that winter (Buchan et al., 2014) in combination with a longer-term increase in the (southward) upper mid-ocean component of the AMOC (McCarthy et al., 2012). McCarthy et al. argue that, as the northward transport of the Gulf Stream was largely compensated by the gyre recirculation instead of the deep return flow, the Atlantic temporarily switched from an overturning to a recirculating regime, stymieing northward heat transport. Nevertheless, the mechanisms and importance of the role of the ocean are a subject of contention. Jackson et al. (2016) argue (using an ocean-assimilated eddying model which faithfully reconstructs the RAPID time series without directly assimilating it) that the Ekman decline was superposed onto an underlying downturn associated with lower-frequency

AMOC variability of remote origin. They highlight the importance western boundary density anomalies propagating from the Labrador Sea. These anomalies, forced in part by variations in surface heat fluxes, ultimately reach the subtropics and change the zonal density gradient (as monitored by RAPID), accelerating or decelerating through geostrophy the upper mid-ocean component of the AMOC volume transport – a mechanism thought to be instrumental in low-frequency AMOC fluctuations (Robson et al., 2014). Roberts et al. (2013) compare an initial condition ensemble in the same model as Jackson et al. (2016) with an assimilated ocean to show that the ultimate origins of the AMOC downturn were primarily atmospheric. Zhao and Johns (2014) further argue that even the mid-ocean geostrophic component of the 2009-2010 event was significantly influenced by longer-term adjustment to wind forcing, in addition to the fast response of the locally wind-forced Ekman component. Conversely, an alternate paradigm suggests that such events can occur spontaneously in a fully eddy-resolving ocean model with no atmospheric variability at all (Thomas and Zhai, 2013).

In this study, we consider oceanic precursors to the 2009-2010 AMOC event from both an historical and an optimal perspective. In particular, we first demonstrate the role of buoyancy-forced preconditioning of the ocean in the event by comparing a historically representative simulation with a parallel reference simulation (where the only historic forcing comes from wind), highlighting the differences in their ocean state and transport at 26.5°N in 2009-2010. We then consider the optimality of this historical thermohaline preconditioning by determining the optimal thermohaline precursor to an equivalent downturn in the form of an instantaneous perturbation applied to the reference run.

For the latter, we utilise the method of nonlinear optimal perturbations (NOP), which extends the framework of linear optimal perturbations (LOP, Sévellec et al., 2007) in a nonlinear context, iteratively. The LOP approach determines the perturbation to the ocean state which has greatest linear impact on a metric of interest, and is well-established in the context of AMOC dynamics (e.g. Sévellec et al., 2017a) and variability (e.g. Sévellec and Fedorov, 2015). However, it is typically used in idealised or laminar ocean models where linear dynamics dominate the response to small perturbations. Here we use an eddy-permitting model which reproduces faithfully the observed variability of the AMOC (e.g. Leroux et al., 2018), including the 2009 event (Blaker et al., 2015), but for which anomaly propagation is strongly nonlinear. There may thus be a significant mismatch between the (linear) predicted and (nonlinear) actual response to an LOP, and this mismatch may be reduced iteratively by further LOP calculations. The iterative method has been suggested by Sévellec and Fedorov (2013b) and used by Müller and Sévellec (2020) to minimise model-data misfit, and was applied by Müller et al. (2020) in the context of the predictability of interannual AMOC anomalies. While Müller et al. (2020) examine a range of anomalous strengths

and lead times in a laminar (non-eddying) configuration of a state-of-the-art ocean model so as to ascertain the detectability of such AMOC "events", we specifically target weakening linked to the 2009-2010 event in a global high-resolution model.

The chapter is set out as follows. In Section 5.2, we describe the NOP method. In Section 5.3 we describe the model configuration, the mean state of the historical and reference simulations, and our experiment design. The evolution of the iterative process is described in Section 5.4, before the structure of the resulting approximation to the NOP after 10 iterations, and its impact in the perturbed run are described in Section 5.5. Finally, the results are discussed and recommendations for future work are presented in Section 5.6.

5.2 Theory

The ocean may be represented at a given time t by its state vector $|\boldsymbol{U}(t)\rangle$, which contains the values of all prognostic values at all locations. The future state of the ocean under a given set of boundary conditions is described by a nonlinear system determined by its current state:

$$\frac{\partial}{\partial t}|\boldsymbol{U}(t)\rangle = \mathcal{N}(|\boldsymbol{U}(t_0)\rangle, t), \tag{5.1}$$

where $\mathcal{N}$ is a nonlinear operator and t_0 is the initialisation time. A particular evolution $|\bar{\boldsymbol{u}}(t)\rangle$ will follow this equation, but applying a perturbation $|\boldsymbol{u_0}\rangle$ at time t_0 causes an alternate evolution of the state $|\widetilde{\boldsymbol{u}}(t)\rangle$, described by

$$\frac{\partial}{\partial t}|\widetilde{\boldsymbol{u}}(t)\rangle = \mathcal{N}(|\bar{\boldsymbol{u}}(t_0)\rangle + |\boldsymbol{u_0}\rangle, t). \tag{5.2}$$

This may be expanded about the unperturbed ("reference") state as a Taylor series

$$\frac{\partial}{\partial t}|\bar{\boldsymbol{u}}(t)\rangle + \frac{\partial}{\partial t}|\boldsymbol{u}(t)\rangle = \mathcal{N}(|\bar{\boldsymbol{u}}(t_0)\rangle, t) + \mathbf{A}(t, t_0)|\boldsymbol{u}_0\rangle + \mathcal{O}(|\boldsymbol{u}_0\rangle)^2), \tag{5.3}$$

where $|\boldsymbol{u}(t)\rangle = |\widetilde{\boldsymbol{u}}(t)\rangle - |\bar{\boldsymbol{u}}(t)\rangle$ is the anomalous evolution due to the perturbation and the expansion's linear coefficient $\mathbf{A}$ is the Jacobian

$$\mathbf{A}(t, t_0) = \frac{\partial \mathcal{N}(|\bar{\boldsymbol{u}}(t_0)\rangle, t)}{\partial |\boldsymbol{U}\rangle}.$$

The leftmost terms on each side of (5.3) correspond to (5.1) applied to the unperturbed state $|\bar{\boldsymbol{u}}\rangle$, while for small perturbations the nonlinear terms may be neglected such that the future state of the perturbation in isolation is approximately described by the solution to a linear non-autonomous system of differential equations, which reads

$$\frac{\partial}{\partial t}|\boldsymbol{u}(t)\rangle = \mathbf{A}(t, t_0)|\boldsymbol{u}_0\rangle, \tag{5.4}$$

for which the existence and uniqueness of a solution is assured (Teschl, 2012). This solution may be written as

$$|\boldsymbol{u}(t)\rangle = \boldsymbol{\Psi}(t, t_0)|\boldsymbol{u}_0\rangle, \tag{5.5}$$

where $\boldsymbol{\Psi}(t, t_0)$ is the linear propagator operator (propagating the perturbation $|\boldsymbol{u}_0\rangle$ at t_0 to its state $|\boldsymbol{u}(t)\rangle$ at t).

We are interested in the effect of applying a perturbation with regard to a given scalar-valued ("cost") function $\langle\mathbf{F}|$ of the ocean state vector. With the aim of studying the 2009-2010 year-long slowdown of the AMOC monitored at 26.5°N, $\langle\mathbf{F}|$ is the year-average volume transport at 26.5°N integrated from the surface to the depth of maximum overturning. Hence we aim to efficiently perturb the ocean state at time t_0 in a way which changes the average AMOC strength over the time period $[t_1, t_2]$ to a given target value, γ^*. From (5.5), this corresponds to finding the smallest perturbation $|\boldsymbol{u}_0^*\rangle$ such that

$$\frac{1}{t_2 - t_1}\int_{t_1}^{t_2} \langle\mathbf{F}|\boldsymbol{\Psi}(t, t_0)|\boldsymbol{u}_0^*\rangle dt = \frac{-1}{t_2 - t_1}\int_{t_1}^{t_2} \langle\mathbf{F}|\bar{\boldsymbol{u}}(t)\rangle dt + \gamma^*, \tag{5.6}$$

which can be seen as the evolved perturbation having the effect of subtracting the mean state value and replacing it with the target value.

It can be shown by the method of Lagrange multipliers (Sévellec et al., 2007) that, in a linear context, the perturbation $|\boldsymbol{a}\rangle$ of size $\epsilon^2 = \langle\boldsymbol{a}|\mathbf{S}|\boldsymbol{a}\rangle$ (using a norm defined by the Euclidean inner product $\langle\boldsymbol{a}|\boldsymbol{a}\rangle$ with weighting matrix $\mathbf{S}$) which has the greatest impact on $\langle\mathbf{F}|$ during $[t_1, t_2]$ is

$$|\boldsymbol{a}\rangle = \pm\epsilon \frac{\int_{t_1}^{t_2} \mathbf{S}^{-1}\boldsymbol{\Psi}^\dagger(t_0, t)|\mathbf{F}\rangle dt}{\sqrt{\int_{t_1}^{t_2}\int_{t_1}^{t_2} \langle\mathbf{F}|\boldsymbol{\Psi}(t, t_0)\mathbf{S}^{-1}\boldsymbol{\Psi}^\dagger(t_0, s)|\mathbf{F}\rangle dt\, ds}}. \tag{5.7}$$

This is the linear optimal perturbation (LOP) of magnitude ϵ. In our experiments we are interested in perturbations to the ocean buoyancy field and so define the magnitude in terms of a density-based norm:

$$\langle\boldsymbol{a}|\mathbf{S}|\boldsymbol{a}\rangle = \frac{\int_\Omega [T^2\alpha^2 + S^2\beta^2]\mathrm{d}V}{\int_\Omega \mathrm{d}V}, \tag{5.8}$$

where Ω is a three-dimensional ocean region over which the norm is evaluated (in our case the full depth Atlantic from 30° S and Arctic) and dV is a volume increment. The terms T, S, α, and β in the integrand are three-dimensional fields respectively representing temperature and salinity perturbations and the mean state thermal expansion and haline contraction coefficients. Of note is that a particular anomaly in density may be brought about in infinitely many ways by perturbing temperature and salinity, and that this choice of norm effectively dictates that temperature and salinity contribute equally and everywhere constructively, as would have an optimal impact on

density. Temperature and salinity anomalies in reality are often compensated, and so an alternative approach would acknowledge their local covariance in the weighting matrix. Such a norm would lead to a perturbation more closely aligned with the behaviour of realistic temperature and salinity anomalies, but at the expense of optimality, which was the focus here.

By choosing

$$\epsilon = \frac{-\int_{t_1}^{t_2} \langle \mathbf{F} | \bar{\boldsymbol{u}}(t) \rangle dt + \gamma^*(t_2 - t_1)}{\sqrt{\int_{t_1}^{t_2} \int_{t_1}^{t_2} \langle \mathbf{F} | \boldsymbol{\Psi}(t, t_0) \mathbf{S}^{-1} \boldsymbol{\Psi}^\dagger(t_0, s) | \mathbf{F} \rangle dt\, ds}}, \tag{5.9}$$

the LOP is optimally scaled and (5.6) is satisfied, as required. However, (5.6) is based on linearisation of the ocean dynamics. As the nonlinear terms in (5.3) are non-zero, this perturbation will not induce the target overturning strength exactly.

We instead treat the optimally scaled LOP as a first guess to the true optimal perturbation, from which we may begin an iterative procedure. In particular we perturb the nonlinear model using the optimally scaled LOP, update the linearisation, and repeat the procedure. In sequential terms, combining (5.7) and (5.9):

$$|\boldsymbol{u}_0^{*n+1}\rangle = |\boldsymbol{u}_0^{*n}\rangle + \left(\gamma^*(t_2 - t_1) - \int_{t_1}^{t_2} \langle \mathbf{F} | \widetilde{\boldsymbol{u}}^n(t) \rangle \, dt\right) \frac{\int_{t_1}^{t_2} \mathbf{S}^{-1} \boldsymbol{\Psi}_n^\dagger(t_0, t) | \mathbf{F} \rangle dt}{\int_{t_1}^{t_2} \int_{t_1}^{t_2} \langle \mathbf{F} | \boldsymbol{\Psi}_n(t, t_0) \mathbf{S}^{-1} \boldsymbol{\Psi}_n^\dagger(t_0, s) | \mathbf{F} \rangle dt ds}, \tag{5.10}$$

where $|\boldsymbol{u}_0^{*0}\rangle = |\mathbf{0}\rangle$ corresponds to no perturbation being applied to the reference state, and where

$$|\widetilde{\boldsymbol{u}}^n(t)\rangle = \mathcal{N}\left(|\bar{\boldsymbol{u}}(t_0)\rangle + \sum_{i=1}^{n} |\boldsymbol{u}_0^{*i}\rangle, t\right) \tag{5.11a}$$

is the ocean state in response to the cumulative total perturbation vector following all previous iterations, and $\boldsymbol{\Psi}_n(t, t_0)$ is the approximate propagator solving the linear system

$$\frac{\partial}{\partial t} |\boldsymbol{u}^n(t)\rangle = \mathbf{A}_n(t, t_0) \, |\boldsymbol{u}_0^n\rangle \, , \mathbf{A}_n = \frac{\partial \mathcal{N}(|\widetilde{\boldsymbol{u}}^n(t_0)\rangle, t)}{\partial \, |\boldsymbol{U}\rangle}. \tag{5.11b}$$

This is obtained by linearisation about the perturbed state following n previous iterations. Advancing the procedure, the increments $|\boldsymbol{u}_0^{*n}\rangle - |\boldsymbol{u}_0^{*n-1}\rangle$ are expected, after an unknown number of iterations, to decrease in magnitude such that the nonlinear and linear response correspondingly converge. The process corresponds to the Newton-Raphson root-finding method applied to a scalar-valued function of a vector (the cost function and ocean state vector, respectively).

5.3 Experimental setup

5.3.1 Model description

We now apply the iterative process derived in Section 5.2 to estimate the optimal thermohaline precursor to the 2009-2010 AMOC event using a selection of simulations (historical, reference and perturbed) in an eddy-permitting ocean model. The model used is v3.4 of the Nucleus for European Modelling of the Ocean (NEMO) General Circulation Model (GCM; Madec, 2012) in the global $^1/_4°$ ORCA025 configuration with 75 vertical levels in partial-step z-coordinates (Bernard et al., 2006). Tracer diffusion was handled using a Laplacian scheme with the configuration default diffusivity of 300 m^2s^{-1}, while momentum diffusion used a bilaplacian scheme with the configuration default viscosity parameter of 1.5×10^{11} m^4s^{-1}. Parameterised eddy-induced velocities were inactive. The linearisation of the model which provides the propagator $\boldsymbol{\Psi}$ of (5.5) is NEMOTAM (Vidard et al., 2015).

Three key differences between the nonlinear and linear model setup were implemented for stability. Tracer diffusion was doubled to 600 m^2s^{-1} in the linear model (for tracer-induced sensitivity fields), and the bilaplacian eddy viscosity coefficient was increased by a factor of 6 (to 9×10^{11} m^4s^{-1}). The latter in particular was chosen to suppress the extreme sensitivities triggered by baroclinic instability in the linear adjoint simulations used to compute the LOPs. While in a nonlinear regime this instability grows to a critical point before saturating (producing ocean eddies), the nonlinear feedback mechanisms which induce this "levelling off" are not present in the linear model. The model instead unrealistically reports that more and more potential energy is available for conversion, increasing the sensitivity without bound. Increasing the viscosity in the linear model alone means that this unrealistic sensitivity growth is restricted without impacting the eddy field in the nonlinear response, as is common practice in high-resolution ocean assimilation, for example (e.g. Verdy and Mazloff, 2017; Nguyen et al., 2020). The chosen TAM viscosity value is hard-coded to be increased by a further factor of 4 in the Mindanao Dome region of the western equatorial Pacific, where linear instability is particularly problematic for adjoint models (Hoteit et al., 2005). Sensitivities in this region ultimately do not impact our results significantly, however, due to its remoteness relative to the RAPID array and the short (one-year) timescale of the 2009-2010 event considered here. Thirdly, the configuration has been modified to allow the linear and nonlinear model to run with differing time steps, and the linear model was run with half the time step (540 s) of the nonlinear model (1080 s). This ensures numerical stability in the TAM without compromising efficiency in the nonlinear model where this reduction is not needed.

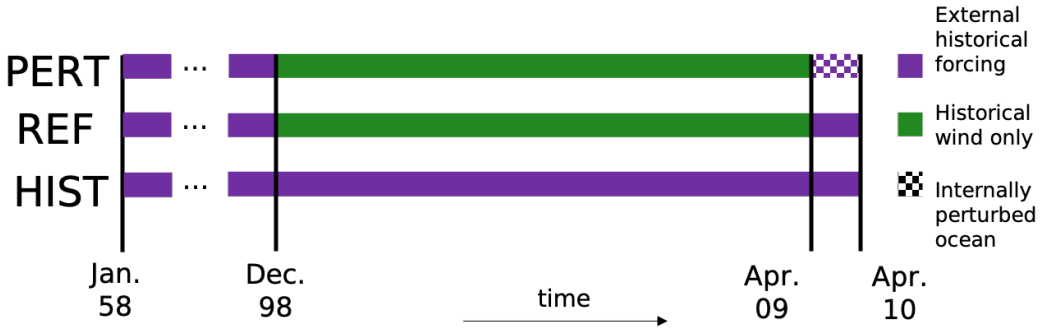


FIGURE 5.1: Schematic highlighting the externally forced and internally perturbed differences between the three simulations considered in our study, "PERT" (top bar), "REF" (middle), and "HIST" (bottom), and the times at which the differences are applied.

5.3.2 Experiment design and model mean state

Our experiment design consists of two base simulations, one recreating the 2009-2010 AMOC event recorded by RAPID as representatively as possible (HIST) and a parallel reference run (REF) where the preconditioning of the event is limited to the wind field, such that there is a lack of realistic surface buoyancy-forced internal thermohaline preconditioning. We aim to perturb the latter run (to induce the full transport reduction of the former) in a further simulation (PERT).

We begin by recreating the 2009-2010 AMOC event in the model by running it subject to historically representative surface buoyancy and wind forcing applied from 1958 to 2010 (Fig. 5.1, "HIST"). The forcing set used is the Drakkar Forcing Set (DFS, v5.2; Brodeau et al., 2010). This simulation exhibits a strong distinction between the subtropical and subpolar regions in key annually averaged climate metrics such as sea surface temperature (SST; Fig. 5.2a) and upper ocean heat content (OHC) distribution (Fig. 5.2b). Here, the OHC distribution is calculated as $\int_{-2000\ m}^{0} \bar{T}(x,y,z) c_p \rho_0 \, \mathrm{d}z$ where $\bar{T}$ is the time-averaged three-dimensional temperature field from 2009-2010, $c_p = 3850$ JK^{-1}kg^{-1} is the specific heat capacity and $\rho_0 = 1027.5$ kgm^{-3} is the reference (volume-weighted mean) density. The area integral of this over the North Atlantic gives a total upper ocean heat content of 2.270×10^{24} J.

The subpolar-subtropical distinction is also apparent in dynamic metrics such as the dynamic height anomaly (relative to the average depth of maximum overturning and calculated as $\int_{-857\ m}^{0} (\rho - \rho_0)\rho_0^{-1} \, \mathrm{d}z$; Fig. 5.2c) and the barotropic stream function (Fig. 5.2d), the cumulative zonally integrated transport from the west (W) of the basin to location x. This is given by $\int_{-H}^{0} \int_{W}^{x} \bar{v}(\zeta,y,z) \mathrm{d}z \, \mathrm{d}\zeta$ where $\bar{v}$ is the 2009-2010 time-averaged meridional velocity, H is the full depth and ζ is a zonal variable of integration. The 2009-2010 time-averaged barotropic stream function takes its North Atlantic maximum value of 54.9 Sv (1 Sv$= 1 \times 10^6$ m^{3}s^{-1}) within the Gulf Stream (36.5°N, 71.6°W). The 2009-2010 year-mean meridional stream function integrated

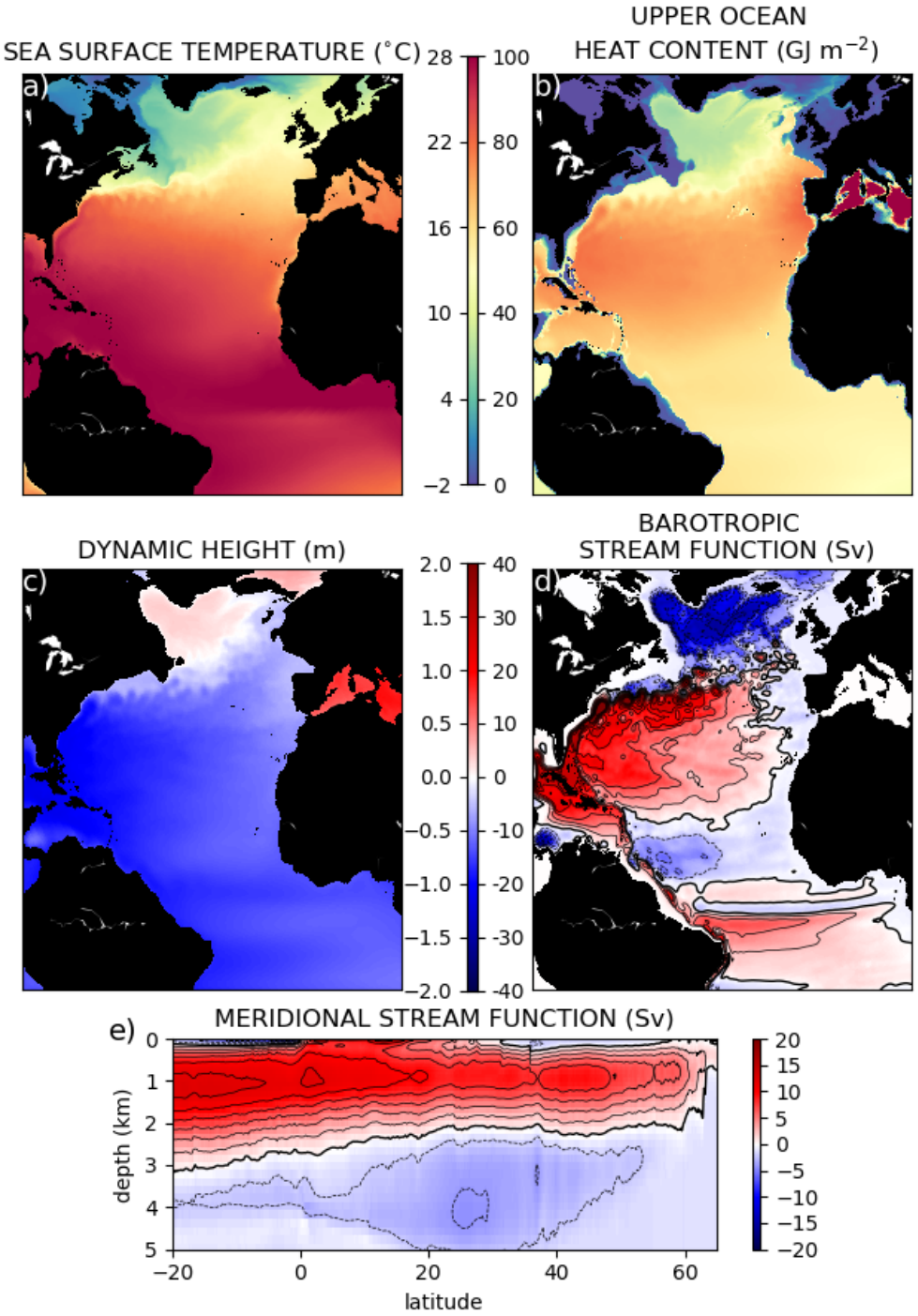


FIGURE 5.2: April 2009 to April 2010 (i.e. RAPID year) average metrics of the representative HIST run. a: sea surface temperature, b: heat content over the top 2 km, c: dynamic height anomaly relative to 857 m, d: Atlantic barotropic stream function integrated from the western boundary (contour intervals are 5 Sv with 0 Sv indicated by a thicker contour), e: Atlantic meridional stream function (contour intervals are 2 Sv with 0 Sv indicated by a thicker contour).

from the surface to depth z (Fig. 5.2e; calculated as $\int_z^0 \int_W^E \bar{v}(x, y, \xi) \mathrm{d}x \, \mathrm{d}\xi$ where ξ is a vertical variable of integration and the zonal integral is to the east [E] of the basin) takes a maximum value of 12.9 Sv.

We are interested in the ocean internal thermohaline precursors to the 2009-2010 event and so, in a parallel run (Fig. 5.1, "REF") only the wind field is taken from the historical DFS, with the surface buoyancy forcing instead taken from the repeat annual CORE "normal year" forcing (NYF; Large and Yeager, 2004) package. By construction, this forcing possesses the climatological properties of decades of realistic surface conditions, as well as typical sub-annual variations and neutral indices of major modes of variability. In the REF run, the surface buoyancy-forced component of the ocean buoyancy field is allowed to drift away from the "true" state from December 1998 until the beginning of the 2009 RAPID year (2009-04-01), when fully historical forcing is reinstated (for both momentum and buoyancy). In this sense, the ocean internal state of the HIST run is fully preconditioned by the forcing at the beginning of the RAPID year, while the REF run possesses very little realistic surface buoyancy-forced thermohaline preconditioning.

Considering the mean state of the REF simulation from April 2009 to April 2010, a southward shift of the North Atlantic Current in HIST is evident, particularly in the SST field of the former (Fig. 5.3a) and resulting in a higher concentration of subpolar heat content in REF (Fig. 5.3b) despite a slightly lower overall integrated North Atlantic heat content (2.25×10^{24} J). The North Atlantic barotropic stream function maximum in REF is found at almost the same location as HIST (36.7°N, 71.1°E; Fig. 5.3d) but takes a much larger value (69.0 Sv), while the upper AMOC cell displays a greater depth extent (Fig. 5.3e). Considering all depths and latitudes, the year-mean meridional stream function of REF has a maximal North Atlantic value of 14.5 Sv. The anomalies of HIST relative to REF are made further clear by considering the differences between these metrics (Fig. 5.4).

As the REF simulation does not possess the buoyancy preconditioning of the HIST simulation, any downturn in AMOC transport in the REF run during the 2009-2010 period is ascribed to either external forcing during the 2009 RAPID year or the delayed effects of wind from previous years. Equivalently, the difference between HIST and REF is solely due to the surface buoyancy flux preconditioning over the period from December 1998 to April 2009. We aim to determine the optimal internal buoyancy perturbation at the beginning of the RAPID year which most efficiently induces the remaining average AMOC decline which is present in HIST but not in REF (Fig. 5.1, "PERT").

The AMOC cost function ($\langle \boldsymbol{F} | \boldsymbol{u}(t) \rangle$) is defined as the one-year-average (from April 2009 to April 2010) volume transport integrated from the surface to the climatological depth of maximum overturning (857 m) and from the western boundary to the eastern

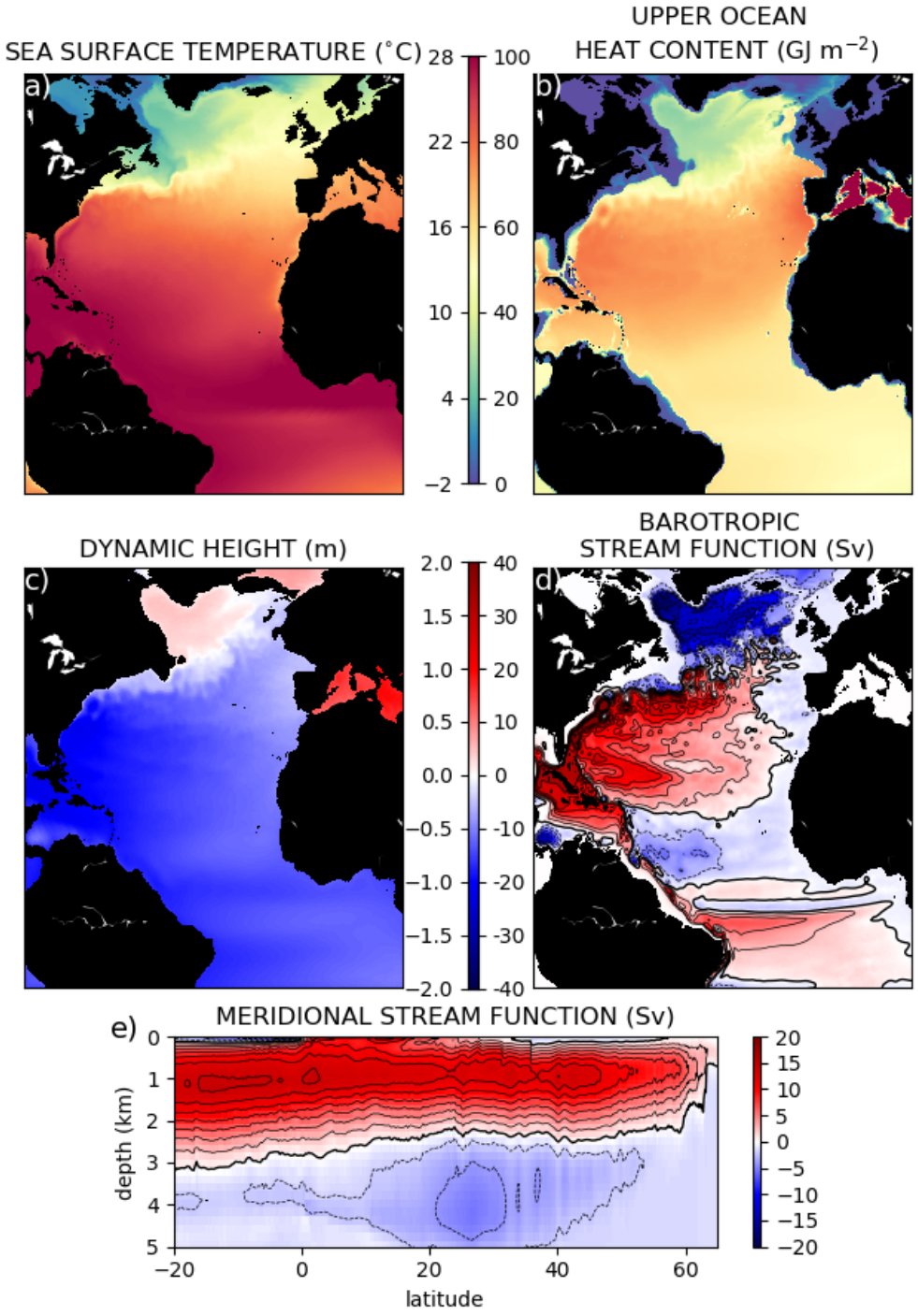


FIGURE 5.3: As in 5.2, but for the REF simulation

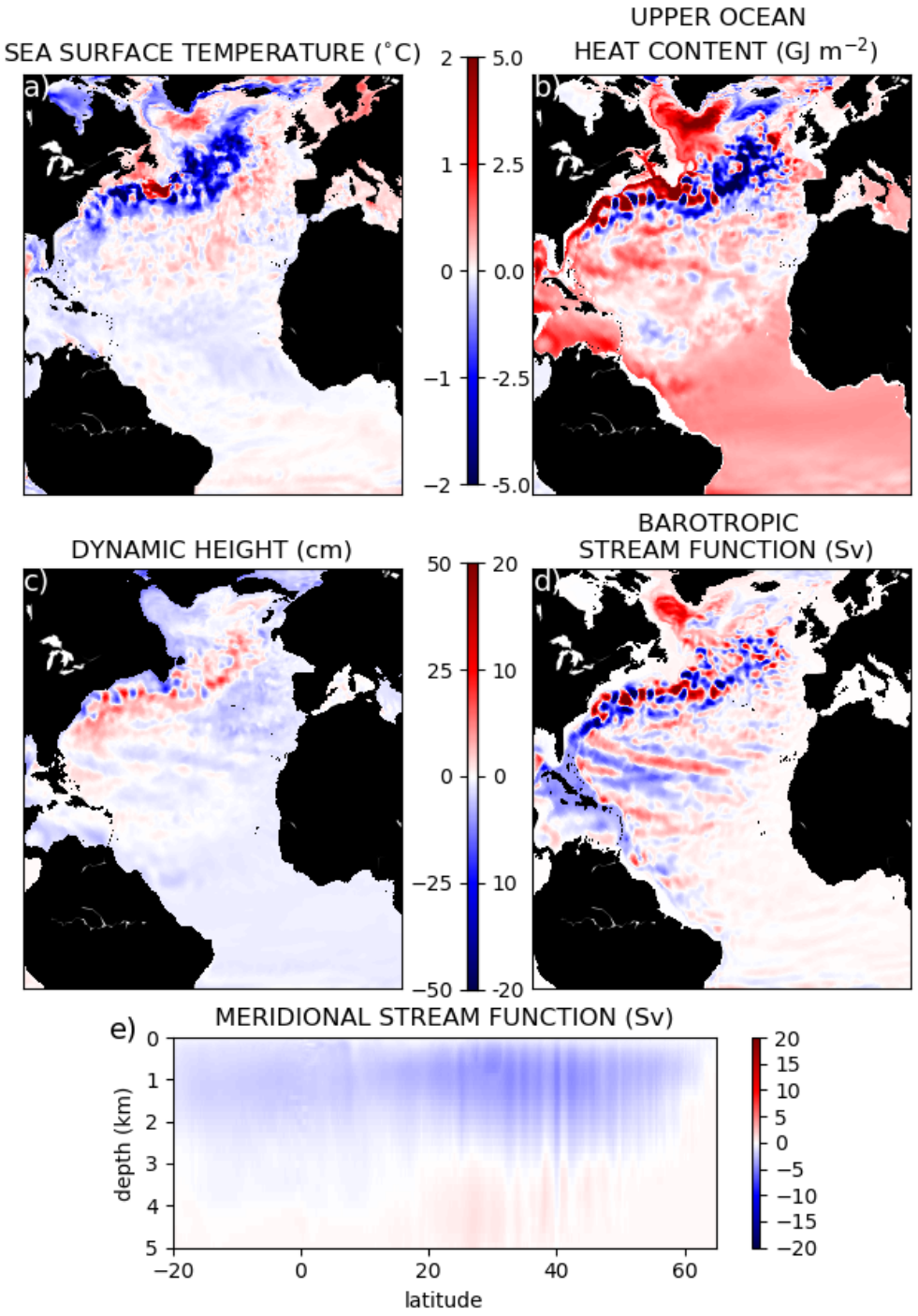


FIGURE 5.4: Differences (HIST-REF) between mean state metrics (as in Figs. 5.2 and 5.3) in the two base simulations during the April 2009 to April 2010 RAPID year. This field can be interpreted as the changes solely due to the the surface buoyancy flux preconditioning over the December 1998 to March 2009 period.

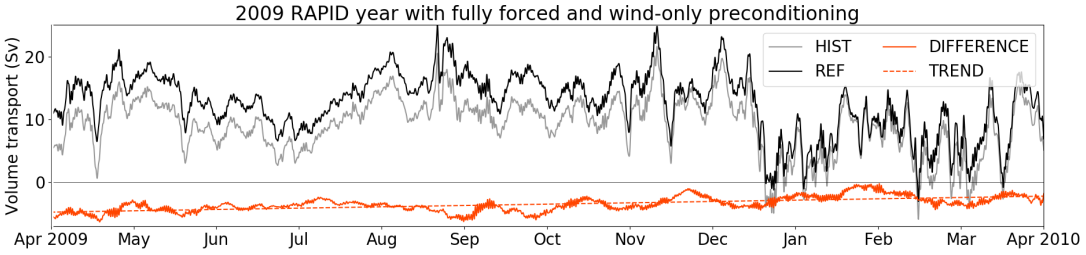


FIGURE 5.5: Day average time series of Atlantic volume transport integrated from the surface to 857 m at 26.5°N in the HIST (grey) and REF (black) runs, along with their difference (HIST-REF, orange). The difference shows the part of the 2009-2010 event induced by the surface-forced thermohaline preconditioning. The different preconditioning but common forcing of the runs causes them to converge during the run, with a linear line of best fit (orange, dashed) shown to represent this.

boundary at the model grid index which is on average closest to 26.5°N ($j = 608$, deviating on average by only 0.01° with 0.25° model resolution). A constant index was used to mitigate the risk of spurious behaviour near index changes.

The configuration was modified such that both the adjoint and nonlinear models received the cost function as input, with the latter returning a time series of the value taken by the cost function at each time step (Fig. 5.5 shows daily averages). This allows (at the highest possible sampling frequency) inexpensive evaluation of the true evolution in the nonlinear model of optimal perturbations determined in the linear model. The cost function (at 26.5°N, to 857 m) takes value 12.5 Sv in the REF run compared with just 9.12 Sv in the HIST run. As they share a common forcing when these values are calculated, this large (3.38 Sv) difference can be attributed to existing differences in their internal buoyancy field, originating from their different preconditioning. Most of the short-term transport variations during the year are common to both runs, with a correlation of 0.97 in the transport time series suggesting that these sub-annual variations are primarily externally forced. There is an apparent convergence of the two runs (linear trend of -2.33 Sv yr^{-1} in their absolute difference) over the year from the tendency of their common external forcing to close the phase differences established by their different preconditioning. For example, while the year-averaged transport difference is 3.38 Sv between REF and HIST, the difference between their month-averaged transport is, in the first month, 4.83 Sv and in the final month is only 2.68 Sv.

In the proceeding section, we apply the approach outlined in Section 5.2 to find the optimal precursor which can also induce a 3.38 Sv year-mean decrease, in the form of an instantaneous internal buoyancy perturbation applied to REF at the beginning of the 2009-2010 event.

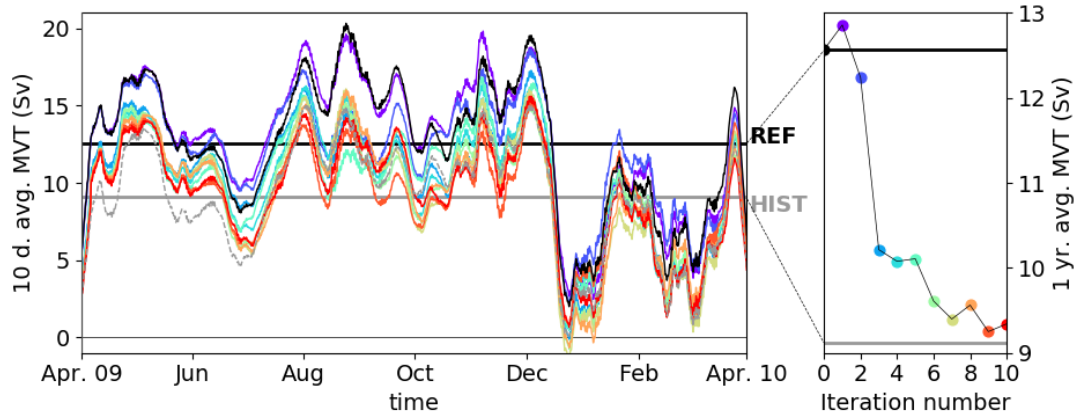


FIGURE 5.6: Response of volume transport at 26.5°N (from 0-857 m) to perturbations produced by the iterative procedure. Left panel shows 10-day average values as a time series over the year of evaluation. Grey dashed line shows the HIST case, black solid line shows the REF case. Thicker grey and black horizontal lines show their respective mean values over the year. Coloured lines show the impact of different perturbations (the perturbation number is given by the corresponding x-axis value in the right panel). Right panel shows the year average (cost function) values after each iteration (circles, whose fill colours also correspond to the iteration number in the left panel), as well as for REF and HIST cases (horizontal black and grey lines, respectively).

5.4 Iterative procedure

As outlined in the preceding sections, we begin by perturbing the REF simulation using its optimally scaled LOP – following (5.7) and (5.9) – in order to reach an equivalent decline to the HIST simulation in terms of year-averaged transport. This procedure is iterated by calculating a new optimally scaled LOP of the perturbed run and adding the result to the existing perturbation, perturbing the original run again, and so on, such that the accumulated LOPs serve as a converging approximation to the NOP of REF.

The LOP, $|\boldsymbol{u}_0^{*1}\rangle$, has a very weak effect which is seemingly unable to overcome natural variations in the model AMOC, leading to a 0.3 Sv increase in the cost function (Fig. 5.6). However, after 3 iterations, the perturbed volume transport value is significantly closer to the target value than to the original reference value, following an 18.7% reduction on the latter (corresponding to a 68.3 % reduction of the mismatch). This increases to 25.7% after 10 iterations, reducing the mismatch by 93.7 %.

While it is ultimately the impact on the average volume transport in the perturbed run which we are interested in, this does not tell us much about the building up of the perturbation itself, from LOP to NOP, due to the iterative procedure. Two properties of interest in this regard concerning the spatial distributions of the LOP and NOP are their correlation and magnitude. At the possible extremes, the LOP and NOP may have identical spatial patterns but different magnitude, or they may have similar

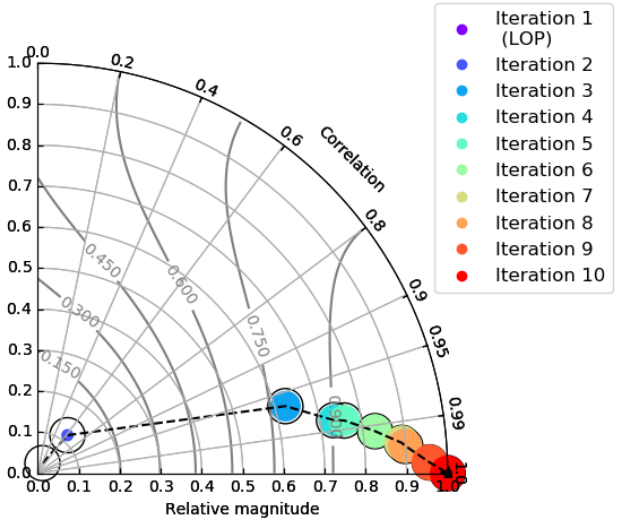


FIGURE 5.7: Taylor diagram showing the evolution of the perturbation over the iterative process. The azimuth shows the correlation with the approximate NOP (i.e. $|\boldsymbol{u}^{*10}\rangle$ obtained after 10 iterations) following (5.12). The radial value indicates the magnitude ratio $\epsilon_i^2/\epsilon_{10}^2$. The proportion of the marker which is filled indicates the proximity to the cost function target (0% for REF, 100% for HIST), and the fill colour indicates the iteration number. Contours describe the "skill score", a weighted evaluation of both coordinates, as described in Taylor (2001).

magnitude but no overlap at all in their shape. This was addressed by Taylor (2001), who proposed a diagrammatic method for representing both aspects of similarity between geospatial fields (Fig. 5.7). In our case the correlation between the i^{th} and j^{th} approximation to the NOP is given by

$$\text{Corr}(|\boldsymbol{u}_0^{*i}\rangle\,,|\boldsymbol{u}_0^{*j}\rangle) = \frac{\langle \boldsymbol{u}_0^{*i}|\mathbf{S}|\boldsymbol{u}_0^{*j}\rangle}{\sqrt{\langle \boldsymbol{u}_0^{*i}|\mathbf{S}|\boldsymbol{u}_0^{*i}\rangle\,\langle \boldsymbol{u}_0^{*j}|\mathbf{S}|\boldsymbol{u}_0^{*j}\rangle}}. \tag{5.12}$$

We are interested in the correlation between the perturbation produced in each iteration and the final iteration of our experiments (which we take as the closest approximation to the true NOP), as well as the ratio of their (squared) magnitudes: $\epsilon_i^2/\epsilon_j^2 = \langle \boldsymbol{u}_0^{*i}|\mathbf{S}|\boldsymbol{u}_0^{*i}\rangle \,/\, \langle \boldsymbol{u}_0^{*j}|\mathbf{S}|\boldsymbol{u}_0^{*j}\rangle$. The LOP ($|\boldsymbol{u}^{*1}\rangle$) exhibits little spatial correlation with the NOP approximation obtained after 10 iterations ($|\boldsymbol{u}^{*10}\rangle$) at 0.42 , and is of only 2.7% the squared magnitude. As in the response time series, these scores change substantially over the next two iterations, with the relative squared magnitude successively increasing five-fold and the correlation increasing by around 150% each time (reaching 62.6% and 0.97, respectively, at the third iteration).

To understand the evolution from the LOP to the closest NOP approximation, we consider their spatial distributions (Figs. 5.8 and 5.9). Certain large-scale structures are evident in both cases, but in the linear case, these are disrupted by noisy small-scale features, particularly in the Gulf Stream and equatorial regions and at depth. The $|\boldsymbol{u}^{*10}\rangle$ distribution exhibits some similar small-scale noise, but largely

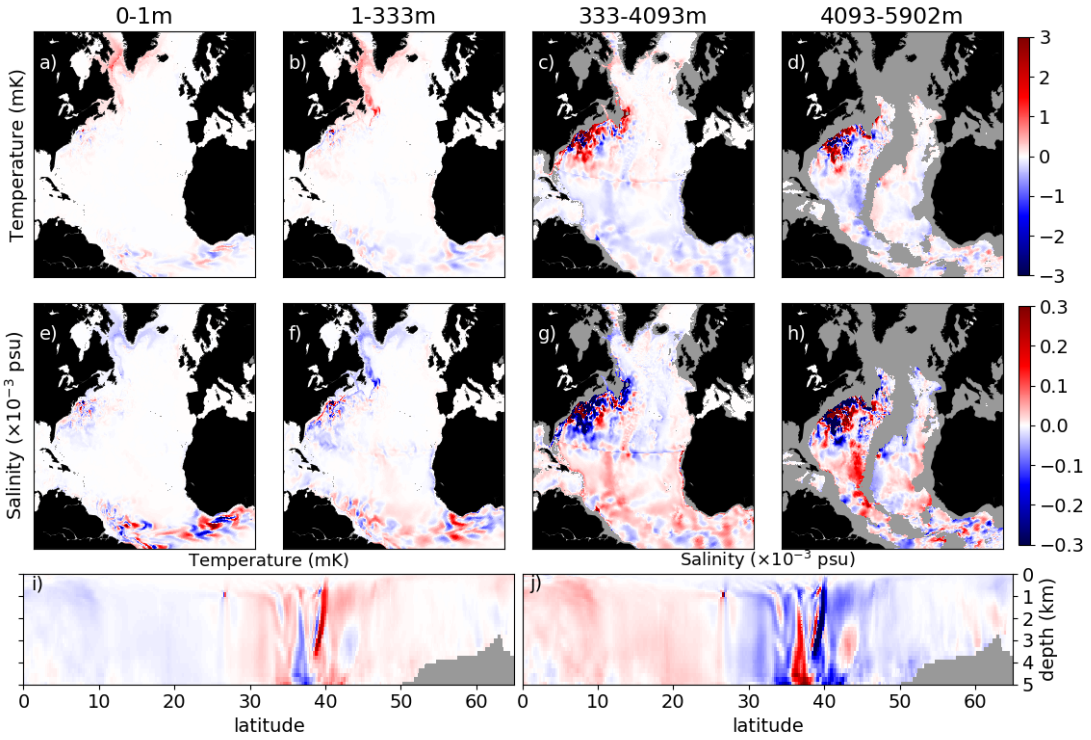


FIGURE 5.8: Spatial distribution of the linear optimal perturbation in temperature (a-d,i) and salinity (e-h,j). a,e: the uppermost level of the model, showing the surface distribution. b,f: depth-weighted average values over the next 332 m. c,g: depth-weighted average values over the next 3760 m. d,h: depth-weighted average values over the bottom 999 m. i,j: zonally averaged values across the Atlantic basin in latitude-depth space. i and j respectively follow the colour scales of a-d and e-h.

redistributes the density anomalies concentrated in small-scale structures in the LOP into larger scale patterns. This behaviour can be isolated by considering the difference in respective contributions between the LOP and NOP to their squared magnitude. By rescaling the LOP such that its magnitude is equal to that of $|\boldsymbol{u}^{*10}\rangle$, then differencing the depth-integrated contributions to their norms, we obtain a field R showing these redistributions (Fig. 5.10):

$$R(x,y,z_1,z_2) = \frac{1}{\int_\Omega \mathrm{d}V}\left[\int_{z_1}^{z_2}(T_{10}^2\alpha^2 + S_{10}^2\beta^2)\mathrm{d}z - \frac{\epsilon_{10}^2}{\epsilon_1^2}\int_{z_1}^{z_2}(T_1^2\alpha^2 + S_1^2\beta^2)\mathrm{d}z\right]. \quad (5.13)$$

The terms in this equation are as in (5.8) but subscripts on T, S indicate the iteration which they belong to, with 1 for the LOP and 10 for $|\boldsymbol{u}^{*10}\rangle$. $\epsilon_1, \epsilon_{10}$ are their respective magnitudes and z_1, z_2 specify a depth range. Note that the global integral $\int_{\Omega_0} R(x,y,0,H)\mathrm{d}A = \epsilon_{10}^2 - \frac{\epsilon_{10}^2}{\epsilon_1^2}\epsilon_1^2 = 0$ by construction (where H is the full ocean depth, Ω_0 represents the surface and $\mathrm{d}A$ is an area increment). R takes negative values where density anomalies in the LOP have been removed in the (approximated) NOP and positive values where these density anomalies have been redistributed, showing the action of the iterative process to modulate the shape. The highly concentrated pockets of negative values in the Gulf Stream and equatorial regions and broad patterns of

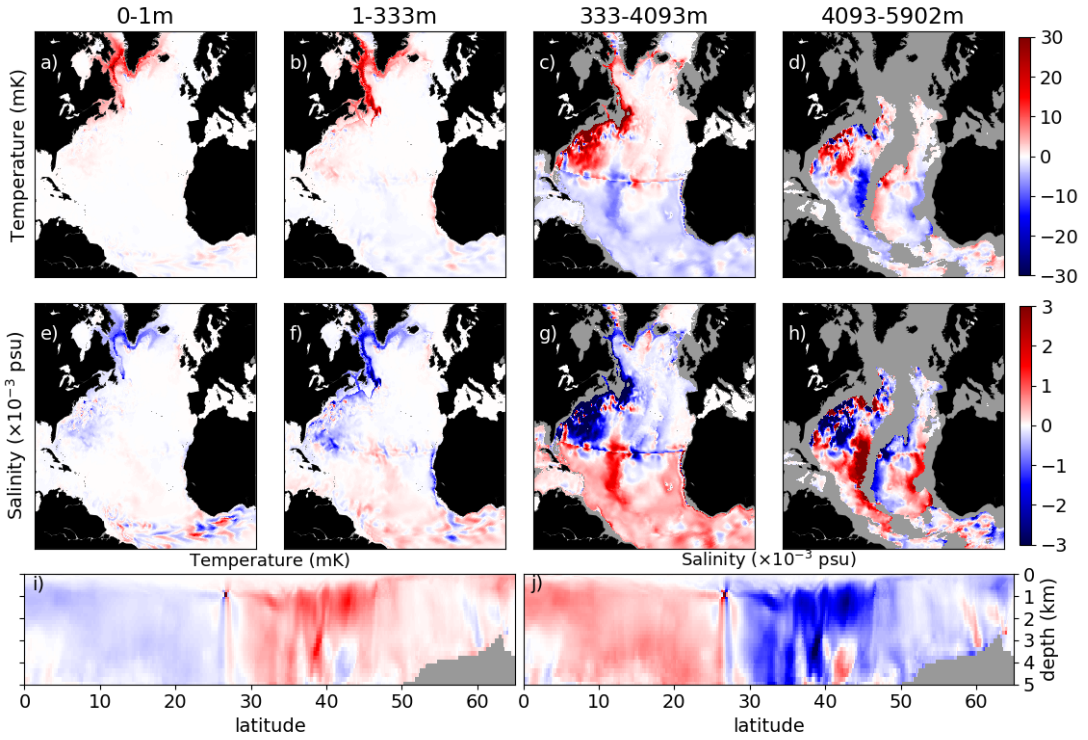


FIGURE 5.9: As in Fig. 5.8, but for the NOP approximation (i.e. $|\boldsymbol{u}^{*10}\rangle$ obtained after 10 iterations). Note in particular the order of magnitude difference in the colour scales compared with the LOP shown in Fig. 5.8.

positive values along the boundaries support the idea that small-scale noise in the LOP (located within these specific regions) is removed and redistributed among large-scale patterns in the NOP.

5.5 Perturbation structure and model response

The NOP pattern, whose large-scale tendency is reinforced by the iterative procedure, varies in nature across different depths and locations. At the surface (Fig. 5.9a,e), there is some stimulation of an equatorial wave pattern, although this has been significantly weakened by the iterative procedure compared with its signature in the LOP. The primary concentration is otherwise a lightening of subpolar waters around the Labrador Sea coast and Baffin Bay, particularly along the Canadian coast. This warming and freshening moves further south with increasing depth along the western boundary, involving greater influence from the subtropics over the next ~ 300 m (Fig. 5.9b,f) with lightening along the western boundary north of 26.5°N and, to a lesser extent, along the eastern boundary to its south. In contrast, there is an overall densification south of the evaluation line such that the zonally averaged optimal perturbation exhibits a dipole shape pivoting on 26.5°N (5.9i,j). This pattern is reinforced in the immediate locality of the evaluation line, with an intensely focused

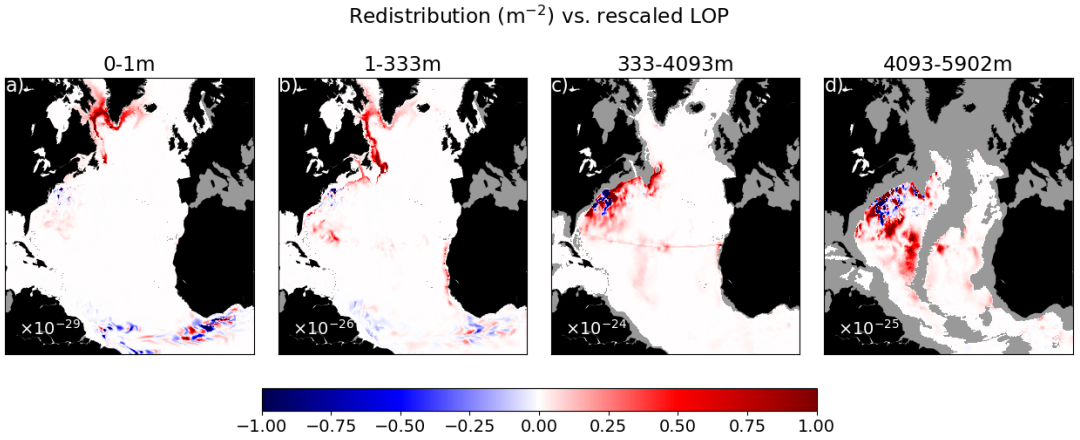


FIGURE 5.10: Redistribution of spatial patterns from the LOP to the approximate NOP (i.e. $|\boldsymbol{u}^{*10}\rangle$ obtained after 10 iterations) and shown as a dimensionless quantity per unit area following (5.13). Negative (blue) values show locations where density anomalies are removed from the LOP pattern, positive (red) values show where they are redistributed in the NOP. The global area integral of the summed four panels is zero-valued and dimensionless.

density gradient across the RAPID line below the depth of maximum overturning and its weaker continuation to the bottom. We take the unrealistic nature of this extreme gradient along the evaluation line with caution, and hypothesise that it might be an artefact of the experimental setup, due to the lack of any lag between the perturbation time and beginning of the evaluation window (the impacts of rapid adjustment processes will be measured and, where they reduce overturning, amplified in subsequent iterations).

The dominant contributions over the next few thousand metres, extending from the depth of the shelf to the abyssal zone, come from the subtropics (Fig. 5.9c,g). The lateral distribution of the meridional density gradient is such that additional heat (freshening) is added along the western boundary north of the RAPID line (within the subtropical gyre recirculation) with cooling (salinification) more strongly concentrated along the west of the Mid-Atlantic Ridge and eastern boundary to its south, such that the NOP simultaneously induces both a meridional and across-basin density gradient in an efficient manner with minimal cross-interference. These bathymetry-following patterns are even more apparent at abyssal depths (Fig. 5.9d,h), where two consecutive rotated gradients develop in the western and eastern North Atlantic basins, between the corresponding continental boundary and the Mid-Atlantic Ridge.

The ocean state of the PERT simulation relative to REF bears little resemblance to the anomaly of HIST relative to REF (compare Figs. 5.11 and 5.4), suggesting that the perturbation does not re-phase the model towards the buoyancy preconditioning of the HIST state, but rather induces a new state whereby an equivalent volume transport reduction is achieved more efficiently. This in turn suggests that the state of the HIST simulation was not the most efficient pathway to reach such an AMOC slowdown. The

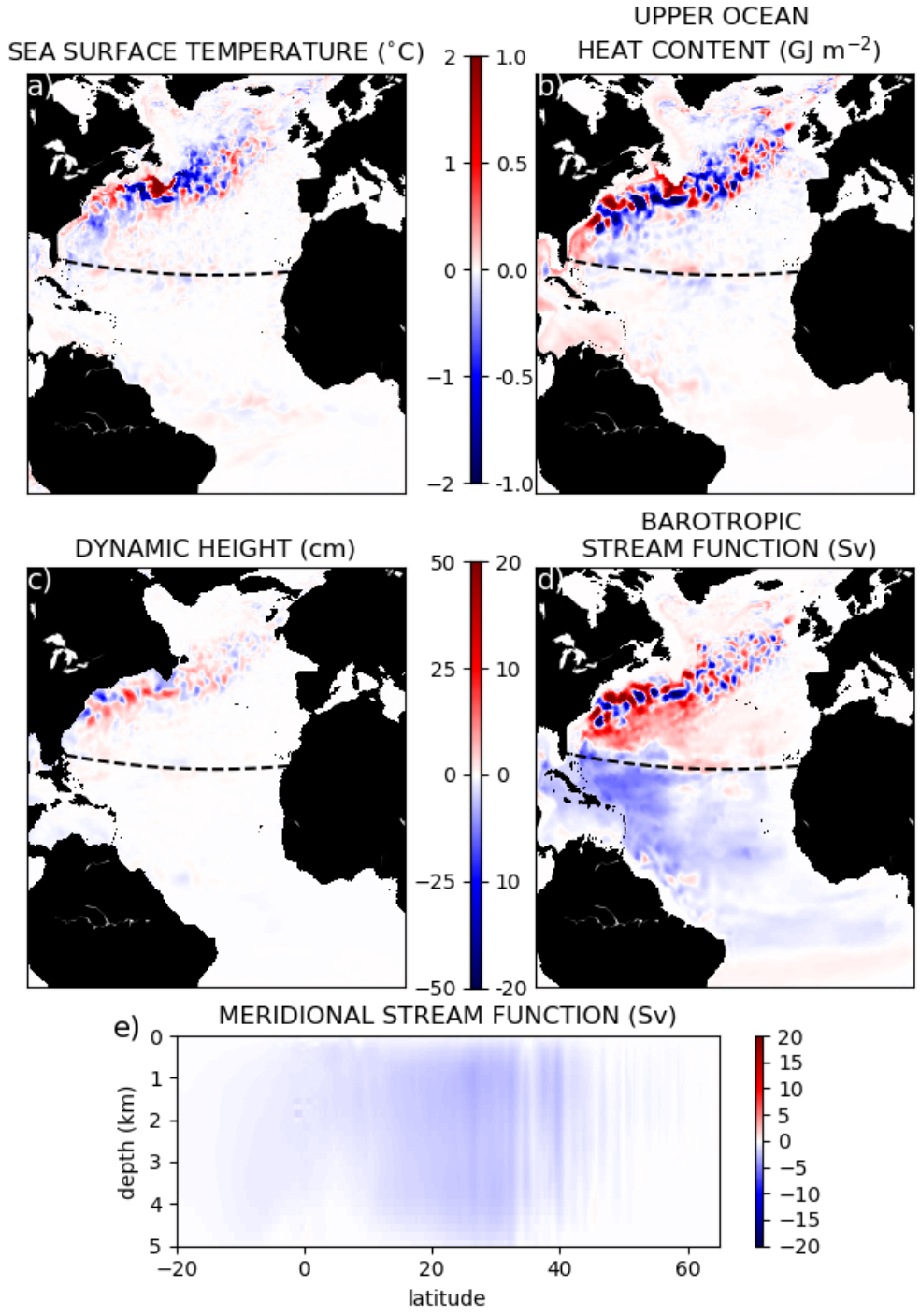


FIGURE 5.11: As in Fig. 5.4 but showing the difference between the perturbed and reference runs (PERT-REF) rather than between the historical and reference runs (HIST-REF). This shows the annually averaged effect (from April 2009 to April 2010) of the approximate NOP (i.e. $|\boldsymbol{u}^{*10}\rangle$ obtained after 10 iterations). Dashed black lines show the location of AMOC evaluation (26.5°N).

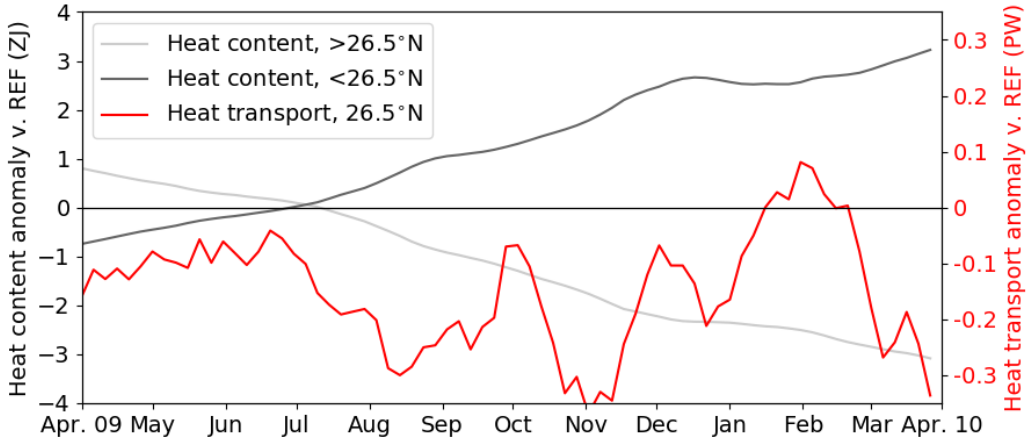


FIGURE 5.12: Time series of impact of the approximate NOP (i.e. $|\boldsymbol{u}^{*10}\rangle$ obtained after 10 iterations) on full-depth Atlantic heat content (PERT-REF) north and south of 26.5°N (light and dark grey, respectively) over the evaluation year (from April 2009 to April 2010). The red curve and right vertical axis show the corresponding changes in full-depth meridional heat transport across 26.5°N. All values are calculated using 5-day averaged model output.

small amplitude of the perturbation relative to typical mesoscale eddies means that the signal of the perturbation is quickly lost when viewing its distribution in space, due to the chaotic divergence of the mesoscale field in response to small perturbations. This is exacerbated by its evaluation window of one year being short relative to a typical eddy lifespan, such that the latter cannot be easily filtered from the output. Despite this, the NOP impact on the overturning circulation remains evident in basin-scale maps (Fig. 5.11d,e), such that the barotropic stream function is coherently increased throughout much of the basin north of the RAPID line, and decreased along it and to its south. It is also of note that the response to the NOP is mainly located on the western half of the basin, without much signature in the eastern half. The impact on the meridional stream function shows that the perturbation has a coherent decreasing effect spanning much of the basin at all depths, which is slightly clearer at the location of evaluation (26.5°N). Bands of interference north of the Gulf Stream separation and along the North Atlantic Current are evident, such that the response becomes relatively weak at subpolar latitudes.

Although the signal of the perturbation is weak in the spatial distribution of the impact on heat content (Figs. 5.11b), its effects are much clearer in the total (full-depth) Atlantic heat content anomaly to the north and south of the evaluation line (Fig. 5.12, light and dark grey lines) and the associated surface-to-bottom heat transport across it (Fig. 5.12, red line). The initial meridional dipole of the perturbation is evident at first with added heat poleward of 26.5N and cooling to the south, inducing a rapid response in heat transport of around -0.1 PW relative to REF (whose mean heat transport over the simulated year is 1.35 PW). This decrease prompts the gradual reversal of the

dipole pattern, with a zero-crossing after around three months preceding eventual warming (cooling) south (north) of the RAPID line over three times as large in magnitude as the initial cooling (warming). The heat transport anomaly, meanwhile, remains negative on average, reaching a peak reduction of 0.37 PW after 7.5 months.

5.6 Discussion, conclusions, and future work

The 2009 RAPID event has become emblematic of the unprecedented variability captured by modern AMOC observation systems. The event itself consisted of a prolonged departure of the AMOC at 26.5°N from its mean value over the 2009-2010 observational period (a decrease of 30% on the measured average; Bryden et al., 2014). Statistical links with North Atlantic cooling the following year (Cunningham et al., 2013), the possible involvement in the devastating European winter which followed (Buchan et al., 2014), and the potential for seasonal prediction of such outcomes in the future (Duchez et al., 2016) have led to the continued discussion of the mechanisms of the event over a decade later (e.g. Jamet et al., 2020). A large decrease in local wind-driven transport only offers a partial explanation, with the remaining component of the downturn evident in the upper-mid-ocean geostrophic transport captured by the basin-spanning density measurements of the RAPID array (McCarthy et al., 2012). These measurements illustrate an increase in the subtropical gyre recirculation strength along with an increased deep return flow associated with the southward transport of subpolar-borne deep water, but explanations for these phenomena vary, ranging from local wind-driven adjustment processes (McCarthy et al., 2012) to remotely sourced buoyancy anomalies (Jackson et al., 2016). In this study we have created an environment in which wind-forced anomalies of all kinds are separate from our results, then considered historical and optimal precursors in temperature and salinity to the 2009-2010 AMOC slowdown event.

We first simulated the 2009-2010 event using an eddy-permitting ocean model fully forced with historical reanalysis fields (the HIST simulation) and which has previously been shown to reliably reproduce the event (Blaker et al., 2015). We then ran the same configuration with interannually varying wind forcing but repeated annual surface buoyancy forcing for a decade prior to the event, imposing an ocean state with minimal surface buoyancy-forced thermohaline preconditioning, which led to only a partial replication of the 2009-2010 event (the REF simulation). We then sought the nonlinear optimal thermohaline perturbation to this state to most efficiently induce the remaining transport decline using an iterative adjoint method (Müller and Sévellec, 2020).

Our iterative process began with a calculation of the optimal perturbation to the volume transport at 26.5°N in a linear context (Sévellec et al., 2007) and an evaluation

of its impact in the fully nonlinear model, which was negligibly smaller than the chaotic variations it induced. After ten iterations of this procedure, 93.7% of the annual-average volume transport mismatch had been removed. We have shown that the iterative process redistributes the perturbation budget from small-scale features to large-scale patterns, amplifying their magnitude to produce the nonlinear optimal perturbation.

At the surface, this perturbation pattern exhibits a lightening primarily along the Labrador Sea coast, consistent with the notion that the dominant mechanism of communication between subpolar density anomalies and subtropical volume transport on these timescales is coastally guided waves, rather than much slower advective processes (e.g. Marshall and Johnson, 2013). The weakened equatorial pattern at the surface may also be an expression of sensitivity to equatorially guided Kelvin wave propagation (e.g. Heimbach et al., 2011). The exact impact of these patterns remains to be studied further, for instance by applying a subset of the perturbation along a narrow neighbourhood of a proposed coastal waveguide such as the western boundary and considering its propagation along the boundary at high temporal resolution.

At depth, we have shown that the optimal pattern corresponds, in the zonal average, to a basin-scale meridional density dipole pivoting on the RAPID line. Perhaps the oldest recognised mechanism of AMOC variability (following AMOC-based interpretations of the generalised study of Stommel, 1961), the role of meridional density gradients in the AMOC has seen renewed interest following demonstrations of their ability to coexist with zonal density gradients and geostrophic transport (e.g. Cimatoribus et al., 2013). Indeed, our NOP distribution is such that across- and along-basin density gradients are stimulated, with lightening in the north focused to the west and densification to the south focused in the east. These across-basin gradients are the primary feature of the abyssal zone, where the basin is separated along the Mid-Atlantic Ridge.

It is perhaps surprising, given the demonstration by many studies that a large portion of annual subtropical volume transport variability is eddy driven (e.g. Thomas and Zhai, 2013; Grégorio et al., 2015; Jamet et al., 2019; Stephenson and Sévellec, 2021b), that an eddy-permitting model actively removes small-scale features in creating a bespoke AMOC perturbation. The ability of the linear model to create effectively targeted new eddies is inhibited by the large suppression of baroclinic instability, creating a mismatch between the forward and adjoint configurations. A simple selection of further experiments in this regard could be devised in which eddy viscosity is equal in both modes, but the evaluation period of the AMOC cost function is much shorter (for example a one month average), such that baroclinic instability does not overwhelm the sensitivity fields. Nevertheless, the small amplitude of the derived perturbation and removal of existing eddy signatures by the iterative process suggests that optimised, large-scale coherent patterns are more efficient in creating variations, while mesoscale turbulence is likely to be, in reality, simply more energetic and

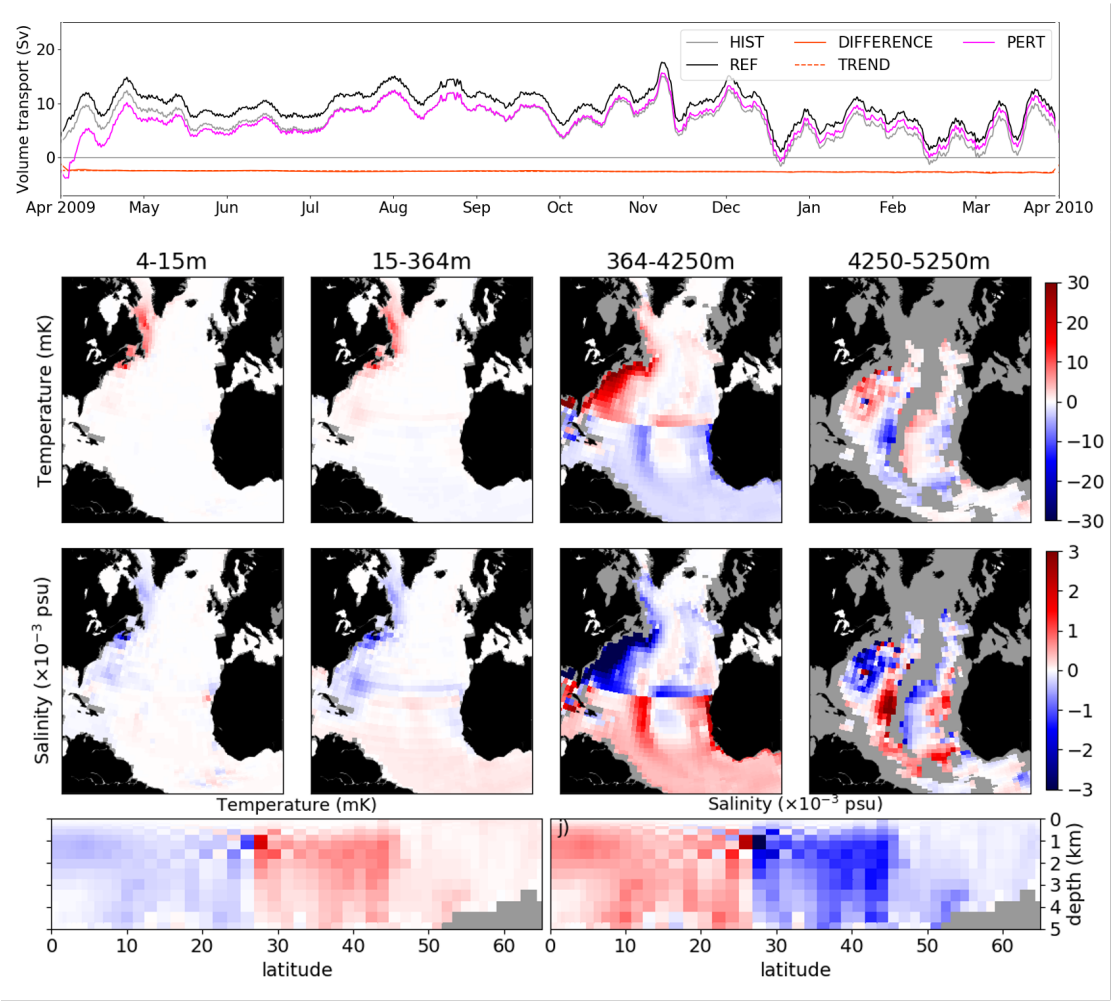


FIGURE 5.13: upper panel: as in Fig. 5.5 but simulated using the lower-resolution ORCA2 configuration (see Chapters 2-4). The response to the NOP after 10 iterations is shown in pink. Lower panels: As in Fig. 5.9, but for the NOP calculated in the ORCA2 configuration. Note the different depth averaging ranges stemming from the coarser vertical discretisation.

ubiquitous. This suggests that an eddy field is not the most optimal pathway for a year-long AMOC slowdown, but we note that it does not mean that eddies cannot induce such an outcome. The favouring of large scales suggests that the immense technical demand of the global high-resolution simulation may be potentially circumvented by combining a lower resolution adjoint model with a high-resolution forward model, as is being explored in data assimilation applications (M. Mazloff, 2019, pers. comm.). Indeed, a preliminary investigation repeating the experimental set-up of this chapter as closely as possible entirely in the ORCA2 configuration (see Chapters 2-4) produces an optimal perturbation with striking visual similarity to that detailed here at very low cost and rapid convergence (98.9% with the LOP, 99.999% within 10 iterations; Fig. 5.13). This is an admittedly quick and crude assessment, but provides a start-point for a more technical exploration of efficiently exploiting similarities between the model configurations. This consideration also suggests that a linear perturbation

approach (as used in Chapters 2-4) is robust at low-resolution, but (as shown in this chapter) breaks down at higher resolution, where a nonlinear framework is needed.

An immediate next step in evaluating the mechanisms highlighted herein is the dissection of the perturbation into apparently distinct regimes (such as subpolar and subtropical) and evaluation of their outcomes individually (along with their union, to assess covariances). The time scales of these mechanisms may be further distinguished by considering different lags between perturbation and evaluation. An ensemble of randomly perturbed HIST, REF, and PERT runs would further allow error margins on the measured impacts of the optimal perturbation to be established. Additionally, a more accurate optimal precursor could perhaps be found by targeting the time-evolving cost function (as shown in Fig. 5.5) rather than an annual mean value. However, given the extremely high level of similarity between the time series of PERT and HIST (Fig. 5.6) it is likely that a generally similar outcome would be reached.

This study has established the motivation for (given the large mismatch between the LOP and NOP) and demonstrated the successful technical implementation of a powerful optimisation method for an era of increasing model complexity. We have applied the method to a pressing question posed by the recent observational record in a global, high-resolution ocean model and its adjoint. Although further work is to be done following the above suggestions, this offers a promising new avenue of investigation of ocean variability, which will be explored extensively in future work.

Chapter 6

Conclusions

6.1 Summary

In the seven decades since Jung (1952) first quantitatively proposed that a significant portion of the Earth's poleward heat transport may be facilitated by the Atlantic circulation, our ability to observe and simulate the circulation and its role in climate has evolved considerably. The ever-increasing fidelity of ocean models and observations seems, however, to uncover as many questions as answers, as the true complexity and variability of the ocean system is revealed, and the separability of spatiotemporal scales becomes less clear cut. Nevertheless, a thorough understanding of the climate system, its natural variability, and its predictable limits is of pressing societal importance as human activities continue to exert an unprecedented influence on the environment (IPCC, 2013).

The signal of human influence has long been known to enter the oceans through the process of ventilation, by which properties are exchanged with the atmosphere at the surface which are then subducted into the ocean interior as homogeneous masses of water. This process has come to be recognised as probabilistic, rather than steady in nature, however (e.g. Bower et al., 2019). Furthermore, these pathways additionally facilitate the intrusion of natural atmospheric fluctuations with extremely limited predictability (Lorenz, 1969), stimulating large-scale oceanic variability (Delworth et al., 2016) and placing fundamental limits on the predictability of oceanic quantities (e.g. Griffies and Bryan, 1997). This uncertainty is further fed by the intense turbulent energy internal to the ocean at small spatial scales (Kirtman et al., 2017), which, in turn, may dissipate energy made available to large-scale variability by larger-scale instabilities (Hochet et al., 2020).

This thesis has presented a series of case studies related to these concepts in regimes of increasing complexity. The first of these investigated ocean ventilation pathways by

presenting a modified tangent-linear and adjoint ocean model which could be used to passively propagate tracers. This was then utilised to diagnose spatiotemporal probability distributions of the surface origins and interior destinations of North Atlantic water masses (NASMW and NADW). While this passive framework was able to describe the ventilation pathways of water masses in the North Atlantic circulation, it remained unknown how well it could describe the surface-to-interior pathways of atmospheric quantities such as heat, which are active in nature and can induce circulation changes. The long-term uptake of heat associated with future climate change scenarios is thought to be poorly represented by passive processes in the North Atlantic for this reason (Banks and Gregory, 2006; Xie and Vallis, 2012; Garuba and Klinger, 2016, 2018). However, recent studies have suggested that, in the context of short-term North Atlantic heat fluctuations driven by stochastic surface forcing, the passive component is dominant (Clement et al., 2015, 2016; Cane et al., 2017). This motivated the second investigation undertaken here, attributing the variability of heat content stimulated by random atmospheric forcing in different ocean regions to passive and active ocean components. This was achieved by comparing the aforementioned passive linear configuration with its fully dynamically active counterpart.

While heat content variability near the surface was primarily found to be due to passive processes, expanding the reservoir to the ocean floor led to poor agreement between the passive and active amplification of variability in the North Atlantic. Recent advances in computing power, allowing studies at higher resolution, have shown that ocean mesoscale eddies additionally stimulate a large portion of ocean heat content variability, suggesting an even greater role for active ocean processes (Penduff et al., 2018). This was investigated in Chapter 4, where the stochastic forcing framework was expanded to account for (slower) ocean mesoscale eddy buoyancy fluxes as well as (faster) atmospheric fluctuations. This allowed the dynamical attribution of prediction uncertainty to atmospheric and oceanic origins for a number of metrics of the North Atlantic circulation. Reformulation as an optimisation problem also allowed a level of comparison with the preferred patterns of uncertainty generation in a theoretical context. Of particular interest was the dominance of mesoscale eddies in stimulating error growth in annually averaged metrics of the subtropical AMOC.

Motivated by these findings, the final study presented in this thesis sought the optimal thermohaline precursor to a realistic AMOC "event", replacing the framework of a stochastic eddy representation (forcing a linear model) with a fully nonlinear, eddy-permitting model configuration. It was demonstrated that, rather than showing a preference for mesoscale eddies, the evolution from the linear to nonlinear optimal precursor actively redistributed density anomalies from small to large scales. This suggests that, if eddies are indeed able to force large annual AMOC fluctuations (as suggested by Chapter 4), it is more a product of their ubiquity than their efficiency.

6.2 Outline of conclusions and implications

6.2.1 Passive pathways of North Atlantic water masses

Utilising classical temperature-salinity based definitions in a 2° model, we showed in Chapter 2 that the surface outcrop of NASMW peaks in April before disappearing from July to December, while a similar April peak in NADW is followed by a smaller peak corresponding to a higher-latitude outcrop of water with the same characteristics in November. Tracer initialised in the (subtropical) NASMW outcrop remained near the surface and followed the gyre circulation, while tracer initialised in the lower-latitude (subpolar) NADW outcrop quickly sank and spread throughout the North Atlantic at depth. Conversely, tracer initialised within the higher-latitude Arctic NADW outcrop rapidly entered denser water and did not ultimately contribute much to the final form of NADW. A backtracking analysis using the model adjoint indeed showed that high-latitude sources of model NADW come from a broad spatial and thermohaline range, rather than a narrow outcrop, and that the lower-latitude subpolar region contributed a much greater quantity of NADW overall. A 400-year run accounting for 86% of the NADW budget gave an expected age of at least 112 years, while a 60-year run accounting for 95% of the NASMW budget gave an expected age of 4.5 years. Thermohaline probability distributions revealed that warmer surface waters than the applied definitions make dominant contributions to water mass formation, such that water mass transformation is not trivial, particularly in the case of NASMW.

6.2.2 Passive and active components of heat content variability

By diagnosing (from a two-decade coupled model simulation) a stochastic white Gaussian noise representation of surface buoyancy and momentum fluxes, Chapter 3 saw the configuration of Chapter 2 adopted to investigate the passive and active nature of heat content variability in intertropical and higher-latitude regions of the global ocean. Projecting the stochastic representation onto both the passive and active adjoint sensitivity fields allowed the calculation of a level of agreement, which reached up to 99.5% for the surface layer (in the Southern Ocean, owing to the particularly high impact of the passive, wind-forced redistribution of the existing heat reservoir there). Extending the depth range to 1500 m, the global heat content variance remained mostly passively driven (52% agreement between passive and active), but the Arctic, Indian Ocean, and intertropical Atlantic switched to an active-dominated regime. At full depth, the amplitude of variability in the North Atlantic increased in the active case (7.6 ZJ^2 at 60 years) relative to the passive case (5.2 ZJ^2), as did the time taken to reach signal-noise equivalence (> 21 years vs. 3 years), a metric often used to quantify the limit of predictability. A decomposition of the surface origins of this variability demonstrated the difference between the active and passive components

as having a dipole structure, conflicting with the passive subpolar pathway described in Chapter 2 to delay the growth of variance in the short term. The findings suggest that, although variability and predictability in the well-monitored surface region are governed by passive processes, the enhanced predictability in the more sparsely observed deep ocean is related to AMOC dynamics.

6.2.3 Internally and externally forced ocean variance growth

While the role of the ocean at large scales as described in Chapter 3 is to slow error growth and enhance predictability, Chapter 4 demonstrated the role of smaller scales in feeding error growth. A time-correlated Ornstein-Uhlenbeck process representation of ocean mesoscale eddies and anomalous atmospheric fluxes was diagnosed (advancing the stochastic theory of Chapter 3) and it was shown that, for subtropical metrics of the AMOC, the former has a decisive influence. At lead times of up to 17 years (23 years), error growth in decadal variations of volume (heat) transport remained dominated by ocean internal mesoscale buoyancy fluctuations, while the atmospherically driven error grew more slowly, ultimately representing up to 60% of the total. Annual-average AMOC uncertainty was found to be controlled at all lead times by eddies, which ultimately accounted for 60% of error. This effect was less dramatic for upper ocean heat content, where the internal component was conversely slowest to grow and represented $< 50\%$ after 60 years. Nevertheless, the implication is that ocean dynamics as a whole (acting on all spatial and temporal scales) represent even more of a contribution to upper ocean heat content variability than that suggested by the large-scale laminar configuration of Chapter 3, and indeed other low-resolution climate model studies. This was less apparent in the subpolar region where internal contributions remained below 20% for measures of the AMOC at all timescales, and reached only 33% for upper ocean heat content. These results suggest a decrease in predictive skill among decadal prediction systems as they mature towards eddying resolution in their ocean components, as may be expected from the increased presence of unpredictable turbulence in the system. From the perspective of our framework, eddies can only feed uncertainty in this manner, due to their treatment as an independent random process. Nevertheless, studies in other frameworks have suggested that eddying models may even demonstrate enhanced predictability in certain regions, either through the damping of otherwise overactive large-scale variability (e.g. Sévellec et al., 2020), or by facilitating surface-depth communication, increasing persistence (e.g. Kravtsov, 2020; Zhang et al., 2021), for example. This has been suggested to redress the apparent ability of lower-resolution climate models to predict reality more skilfully than they can predict their own future states (the "signal-to-noise paradox", Scaife and Smith, 2018) as their ensembles are excessively divergent. While debate as to the origin of this paradox continues (e.g. Zhang and Kirtman, 2019; Mayer et al., 2021), its effect would seemingly worsen following the results shown here.

6.2.4 Preferred oceanic patterns of variability

In addition to our consideration of ocean behaviour under realistic circumstances, Chapters 4 and 5 also considered the optimal patterns of ocean perturbation to induce variations of interest. The OSP framework of Chapter 4 sought the optimal stimulation (by idealised external and internal white noise) of unpredictable variability in the metrics described above. We uncovered a consistent range of locations and patterns across all metrics and timescales which efficiently transform noise forcing into oceanic variability. These include a surface buoyancy dipole pivoting on the North Atlantic Current, buoyancy stimulation of the deep water formation region, remote alongshore meridional wind forcing and local zonal wind stimulation of Ekman transport (across the boundaries of the areas under consideration). The diagnostic of the response ratio was employed to quantify how efficiently these mechanisms act. It was demonstrated that, while the internal OSP of subpolar heat content is > 3.7 times as effective as its subtropical counterpart for all time averages, volume transport in the subtropics is at least 2.8 times as sensitive to optimal internal noise as in the subpolar region.

Regarding the latter, the optimal internal stimulation of subtropical volume transport anomalies was also considered within Chapter 5, in a more realistic configuration. This was carried out in the context of the 2009-2010 AMOC downturn captured by the RAPID observation system. In this case, from a perturbation of just ~ 0.1 K, a simulated 3.4 Sv annual-average transport decline (which was associated with the AMOC event and which could not be explained by wind) was almost completely recovered by the application of a nonlinear optimal perturbation. This perturbation, targeting the optimal thermohaline precursor to the 2009-2010 event, was constructed from 10 repetitions of an iterative procedure. The first iteration, producing the optimal perturbation in a linear context (similarly to Chapter 4) had no negative impact on the volume transport in the eddy-permitting configuration used, due to its inability to account for nonlinear feedbacks on small scale anomalies. The corresponding nonlinear perturbation was over 5 times as large as that obtained in the linear framework, and predominantly targeted larger scales than in the linear case. This large-scale nonlinear pattern consisted of remote coastal increases in buoyancy (especially intensified at the surface in the Labrador Sea) and large-scale local density gradients (particularly at intermediate depths in the subtropical gyre recirculation, and at the bottom along the Mid Atlantic Ridge). The results of Chapters 4 and 5 imply that large-scale spatial coherence of perturbations can result in much more effective translation of small anomalies into substantial oceanic variability, while small-scale noise, such as that which is abundant at the oceanic mesoscale, can instead achieve substantial variations as a result of its large scope and magnitude.

6.3 Recommendations for future work

Much of the work presented here is, at least in part, of a technical or theoretical nature and, while it has been exemplified with relevant and important case studies, its further application offers much potential for investigation, both into these topics and those of the wider field.

Chapter 2 presents two directions for further work. The passive-tracer module developed for the NEMO GCM itself presents many avenues by which it may be further optimised, and this is an area of continuing work. Indeed, at time of writing, a modification has been implemented which allows a significant reduction in trajectory storage, by running the full global model but only storing areas relevant to the region of study for the offline TAM run. For an eddy-permitting five-year simulation tracking NASMW, this corresponded to a 90% reduction in required storage, allowing limited disk space to be used for longer experiments. Moreover, the module itself still propagates the tracer offline as if it were active, before voiding the active component, resulting in the majority of computations being wasted. Removing this effect could prove complicated, but if implemented successfully would allow significant reductions in runtime for higher-resolution configurations. This provides a foundation for another important direction of this work.

Although parameterised diffusion provides an advantage over similar-resolution Lagrangian particle tracking tools lacking this feature, a need for higher resolution has still been demonstrated here, particularly in the case of NADW, where poorly resolved bathymetry precludes the accurate representation of the pathways of water originating in the Nordic Seas. Further work considering the finer scale pathways and physics involved in water mass ventilation is thus pressing.

Applying the passive configuration of Chapter 2, Chapter 3 provided a broad comparison of passive and active processes leading to heat content variability throughout the ocean. This ultimately highlighted the Atlantic sector as uniquely determined by active feedback processes at depth, owing to apparent signatures of the AMOC. This warrants two further directions of inquiry. Firstly, the finding that near-surface heat content variability is primarily passively forced should be investigated under a range of boundary conditions (flux and mixed boundary conditions were considered here, using model default parameters). Secondly, the role and mechanisms of the AMOC deserve further investigation, in particular the suggested link to a known mode of internal variability in the model. A novel approach could be the calculation of surface optimal stochastic perturbations (following Chapter 4) which could be formulated as a linear combination of a passive component and perpendicular remainder (encompassing non-passive processes alone). These perpendicular OSPs would isolate the passive and active mechanisms of variability, and could be compared with diagnosed surface noise to further investigate their relationship.

The established problem of fully correlated optimal stochastic perturbations was reformulated in Chapter 4 in a more general way, such that spatial correlation was considered as a constraint. This remained idealised in the context of ocean internal noise owing to the lack of spatiotemporal coherence. Introducing a time-correlated dimension to the OSP framework would require further constraint and introduce substantial complexity, but is worthy at least of consideration. Meanwhile, in order to represent ocean internal noise, we considered only the case of total spatial decorrelation: increasing the size of lateral coherence would increase computational demand in solving the associated eigenvalue problems, but would provide valuable insight into the impact of broader spatial coherence on noise efficiency. Additionally, while the realistic representation provided an efficient means of dynamical attribution, it was diagnosed from an eddy-permitting model, known to under-represent ocean turbulence at higher latitudes. The now-demonstrated methodology could without complication be applied to a representation diagnosed from a fully eddy-resolving model to address this, which would be expected to increase the proportion of diagnosed internally driven variability in the subpolar region to some extent.

The successful implementation of an iterative method to determine nonlinear optimal perturbations in a global, eddy-permitting model in Chapter 5 is, to the author's knowledge, a first on these timescales. It provides an exciting opportunity to pursue in cutting-edge (global, high-resolution) models many avenues previously restricted to idealised configurations. The work described here concentrates on a particular event, using an instantaneous perturbation (as was appropriate to the use case, rather than by limitation of the method to instantaneous perturbations). The focus on a single possible precursor to a particular event, however, was necessitated by the computational expense of the procedure. Reducing this expense and considering a range of events and precursors, along with their common patterns, could have significant implications for targeted observations and their assimilation into operational forecast systems, laying the foundations for a bespoke "early warning system" for major circulation events such as the 2009 event considered here. A crude reproduction of the results shown here in a lower-resolution model suggested one possible pathway to such an expense reduction. A middle-ground solution, maintaining the realism of the ocean response, would be the combination of a laminar TAM configuration with a high-resolution nonlinear model. This could potentially open further options for combining configurations, such as pairing an ocean-only TAM with a coupled climate model (to determine NOPs in a coupled context). Such options will be investigated thoroughly after the finalisation of the preliminary results presented here.

6.4 Reflection on thesis aims: successes and limitations

This thesis aimed to identify the pathways, origins, magnitude, timescales, and predictability of climate-relevant oceanic variations in the North Atlantic. The scope and complexity of this goal give rise to an abundance of possible approaches to address it, each with advantages and limitations. The work here has presented but a few such approaches which are hoped to fortify the existing body of work on the subject. An advantage of many of the methodologies employed here is their computational efficiency and causal nature relative to typical approaches used to address the same questions. Nevertheless, this efficiency and causality was based on a number of simplifying assumptions. For instance, in Chapter 2 we were able to propagate a continuous probability distribution water mass formation for several centuries without reconstructing it from advected particles in an expensive eddying model, but the effects of turbulent mixing were parameterised. In Chapter 3, we could causally deduce the passive and active origins of heat content variability, but without realistic atmospheric coupling. In Chapter 4, we could inexpensively dynamically attribute the impacts of mesoscale eddy turbulence without explicitly resolving it, but under the assumption that it was independent of large-scale variability. Conversely, for Chapter 5, we could not disentangle these scales, producing a more realistic representation, but at greater expense and without this clean separability. Despite some drawbacks such as these, a major success of the work is its close agreement with many results established in much more complex and realistic configurations by past studies. In this sense, this work can be seen as complementary to them, and by virtue of its efficiency and causality, will allow uninhibited future investigation into the complex and far-reaching subject of North Atlantic variability at low expense.

References

Julie Alexander and Adam H Monahan. Nonnormal perturbation growth of pure thermohaline circulation using a 2D zonally averaged model. *Journal of Physical Oceanography*, 39(2):369–386, 2009.

Douglas E Alsdorf, Ernesto Rodríguez, and Dennis P Lettenmaier. Measuring surface water from space. *Reviews of Geophysics*, 45(2), 2007.

Andreas J Andersson, Lilian A Krug, Nicholas R Bates, and Scott C Doney. Sea–air CO_2 flux in the North Atlantic subtropical gyre: role and influence of Sub-Tropical Mode Water formation. *Deep Sea Research Part II: Topical Studies in Oceanography*, 91:57–70, 2013.

Brian K Arbic, Malte Müller, James G Richman, Jay F Shriver, Andrew J Morten, Robert B Scott, Guillaume Sérazin, and Thierry Penduff. Geostrophic turbulence in the frequency–wavenumber domain: Eddy-driven low-frequency variability. *Journal of Physical Oceanography*, 44(8):2050–2069, 2014.

Olivier Arzel, Thierry Huck, and Alain Colin de Verdière. The internal generation of the Atlantic Ocean interdecadal variability. *Journal of Climate*, 31(16):6411–6432, 2018.

Tom Avsic, Johannes Karstensen, Uwe Send, and Jürgen Fischer. Interannual variability of newly formed Labrador Sea Water from 1994 to 2005. *Geophysical Research Letters*, 33(21), 2006.

Rahul Bale and Rama Govindarajan. Transient growth and why we should care about it. *Resonance*, 15(5):441–457, 2010.

Magdalena A Balmaseda, Kevin E Trenberth, and Erland Källén. Distinctive climate signals in reanalysis of global ocean heat content. *Geophysical Research Letters*, 40 (9):1754–1759, 2013.

Helene T Banks and Jonathan M Gregory. Mechanisms of ocean heat uptake in a coupled climate model and the implications for tracer based predictions of ocean heat uptake. *Geophysical Research Letters*, 33:L07608, 2006.

Nicholas R Bates. Multi-decadal uptake of carbon dioxide into subtropical mode water of the North Atlantic Ocean. *Biogeosciences*, 9(7), 2012.

Nicholas R Bates, A Christine Pequignet, Rodney J Johnson, and Nicolas Gruber. A short-term sink for atmospheric CO_2 in subtropical mode water of the North Atlantic Ocean. *Nature*, 420(6915):489–493, 2002.

AB Becher. The Nautical Magazine and Naval Chronicle for February 1843, and November and December 1852. *Simkin, Marshall, and Co., Stationers' Hall Court*, 1852.

Pavel Berloff, A Mc C Hogg, and W Dewar. The turbulent oscillator: A mechanism of low-frequency variability of the wind-driven ocean gyres. *Journal of Physical Oceanography*, 37(9):2363–2386, 2007.

Barnier Bernard, Gurvan Madec, Thierry Penduff, Jean-Marc Molines, Anne-Marie Treguier, Julien Le Sommer, Aike Beckmann, Arne Biastoch, Claus Böning, Joachim Dengg, et al. Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. *Ocean Dynamics*, 56(5-6): 543–567, 2006.

Laurent Bessières, Stéphanie Leroux, Jean-Michel Brankart, Jean-Marc Molines, Marie-Pierre Moine, Pierre-Antoine Bouttier, Thierry Penduff, Laurent Terray, Bernard Barnier, and Guillaume Sérazin. Development of a probabilistic ocean modelling system based on NEMO 3.5: application at eddying resolution. *Geoscientific Model Development*, 10(3):1091–1106, 2017.

Arne Biastoch, Claus W Böning, and JRE Lutjeharms. Agulhas leakage dynamics affects decadal variability in Atlantic overturning circulation. *Nature*, 456(7221):489, 2008.

Stavroula Biri, Martin G Scharffenberg, and Detlef Stammer. A probabilistic description of the mesoscale eddy field in the ocean. *Journal of Geophysical Research: Oceans*, 120(7):4778–4802, 2015.

Joseph Black and John Robison. *Lectures on the Elements of Chemistry, Delivered in the University of Edinburgh; By the Late Joseph Black, MD Professor of Chemistry in that University, Physician to His Majesty for Scotland; Member of the Royal Society of Edinburgh, of the Royal Academy of Sciences at Paris, and the Imperial Academy of Sciences at St. Petersburgh. Now Published from His Manuscripts, by John Robison, LLD. Profesor of Natural Philosophy in the University of Edinburgh.* Mundell and Son, 1803.

Adam T Blaker, Joël J-M Hirschi, Gerard McCarthy, Bablu Sinha, Sarah Taws, Robert Marsh, Andrew Coward, and Beverly de Cuevas. Historical analogues of the

recent extreme minima observed in the Atlantic meridional overturning circulation at 26° N. *Climate Dynamics*, 44(1-2):457–473, 2015.

Bruno Blanke and Stéphane Raynaud. Kinematics of the Pacific equatorial undercurrent: An Eulerian and Lagrangian approach from GCM results. *Journal of Physical Oceanography*, 27(6):1038–1053, 1997.

George J Boer. A study of atmosphere-ocean predictability on long time scales. *Climate Dynamics*, 16(6):469–477, 2000.

George J Boer. Long time-scale potential predictability in an ensemble of coupled climate models. *Climate Dynamics*, 23(1):29–44, 2004.

George J Boer. Decadal potential predictability of twenty-first century climate. *Climate Dynamics*, 36(5-6):1119–1133, 2011.

George J Boer and SJ Lambert. Multi-model decadal potential predictability of precipitation and temperature. *Geophysical Research Letters*, 35(5), 2008.

Amy S Bower, M Susan Lozier, Stefan F Gary, and Claus W Böning. Interior pathways of the North Atlantic meridional overturning circulation. *Nature*, 459 (7244):243, 2009.

Amy S Bower, Susan Lozier, and Stefan Gary. Export of Labrador Sea water from the subpolar North Atlantic: a Lagrangian perspective. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(17):1798–1818, 2011.

Amy S Bower, S Lozier, A Biastoch, K Drouin, N Foukal, H Furey, M Lankhorst, S Rühs, and S Zou. Lagrangian views of the pathways of the atlantic meridional overturning circulation. *Journal of Geophysical Research: Oceans*, 2019.

Elena Brambilla and Lynne D Talley. Surface drifter exchange between the north atlantic subtropical and subpolar gyres. *Journal of Geophysical Research: Oceans*, 111(C7), 2006.

Laurent Brodeau, Bernard Barnier, Anne-Marie Treguier, Thierry Penduff, and Sergei Gulev. An ERA40-based atmospheric forcing for global ocean circulation models. *Ocean Modelling*, 31(3-4):88–104, 2010.

Wallace S Broecker. A revised estimate for the radiocarbon age of North Atlantic Deep Water. *Journal of Geophysical Research: Oceans*, 84(C6):3218–3226, 1979.

Wallace S Broecker. The biggest chill. In *Global Climate Change Linkages; Acid Rain, Air Quality, and Stratospheric Ozone*, pages 13–22. Elsevier New York, 1987.

Wallace S Broecker, Dorothy M Peteet, and David Rind. Does the ocean–atmosphere system have more than one stable mode of operation? *Nature*, 315(6014):21–26, 1985.

Kirk Bryan. Measurements of meridional heat transport by ocean currents. *Journal of Geophysical Research*, 67(9):3403–3414, 1962.

Harry L Bryden, Brian A King, Gerard D McCarthy, and EL McDonagh. Impact of a 30% reduction in Atlantic meridional overturning during 2009-2010. *Ocean Science*, 10(4):683–691, 2014.

Jian Buchan, Joël J-M Hirschi, Adam T Blaker, and Bablu Sinha. North Atlantic SST anomalies and the cold North European weather events of winter 2009/10 and December 2010. *Monthly Weather Review*, 142(2):922–932, 2014.

Martha W Buckley and John Marshall. Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review. *Reviews of Geophysics*, 54(1):5–63, 2016.

Martha W Buckley, Tim DelSole, M Susan Lozier, and Laifang Li. Predictability of North Atlantic sea surface temperature and upper-ocean heat content. *Journal of Climate*, 32(10):3005–3023, 2019.

Kristin C Burkholder and M Susan Lozier. Subtropical to subpolar pathways in the North Atlantic: Deductions from Lagrangian trajectories. *Journal of Geophysical Research: Oceans*, 116:C07017, 2011.

Mark A Cane, Amy C Clement, Lisa N Murphy, and Katinka Bellomo. Low-pass filtering, heat flux, and Atlantic multidecadal variability. *Journal of Climate*, 30(18): 7529–7553, 2017.

Paola Cessi. A simple box model of stochastically forced thermohaline flow. *Journal of Physical Oceanography*, 24(9):1911–1920, 1994.

Ping Chang, R Saravanan, Timothy DelSole, and Faming Wang. Predictability of linear coupled systems. Part I: Theoretical analyses. *Journal of Climate*, 17(7): 1474–1486, 2004.

Yeon S Chang, Zulema D Garraffo, Hartmut Peters, and Tamay M Özgökmen. Pathways of Nordic Overflows from climate model scale and eddy resolving simulations. *Ocean Modelling*, 29(1):66–84, 2009.

Jule G Charney, Akio Arakawa, D James Baker, Bert Bolin, Robert E Dickinson, Richard M Goody, Cecil E Leith, Henry M Stommel, and Carl I Wunsch. *Carbon dioxide and climate: a scientific assessment.* National Academy of Sciences, Washington, DC, 1979.

Dudley B Chelton, Roland A DeSzoeke, Michael G Schlax, Karim El Naggar, and Nicolas Siwertz. Geographical variability of the first baroclinic Rossby radius of deformation. *Journal of Physical Oceanography*, 28(3):433–460, 1998.

Dudley B Chelton, Michael G Schlax, Roger M Samelson, and Roland A de Szoeke. Global observations of large oceanic eddies. *Geophysical Research Letters*, 34(15), 2007.

Wei Cheng, John CH Chiang, and Dongxiao Zhang. Atlantic meridional overturning circulation (AMOC) in CMIP5 models: RCP and historical simulations. *Journal of Climate*, 26(18):7187–7197, 2013.

Andrea A Cimatoribus, S Drijfhout, and Henk A Dijkstra. Reconciling the north–south density difference scaling for the Meridional Overturning Circulation strength with geostrophy. *Ocean Science Discussions*, 10(6):2461–2479, 2013.

R Allyn Clarke and Jean-Claude Gascard. The formation of Labrador Sea water. Part I: Large-scale processes. *Journal of Physical Oceanography*, 13(10):1764–1778, 1983.

Amy Clement, Katinka Bellomo, Lisa N Murphy, Mark A Cane, Thorsten Mauritsen, Gaby Rädel, and Bjorn Stevens. The Atlantic Multidecadal Oscillation without a role for ocean circulation. *Science*, 350(6258):320–324, 2015.

Amy Clement, Mark A. Cane, Lisa N. Murphy, Katinka Bellomo, Thorsten Mauritsen, and Bjorn Stevens. Response to Comment on "The Atlantic Multidecadal Oscillation without a role for ocean circulation". *Science*, 352(6293):1527–1527, 2016. ISSN 0036-8075. . URL https://science.sciencemag.org/content/352/6293/1527.2.

Alain Colin de Verdière. On mean flow instabilities within the planetary geostrophic equations. *Journal of Physical Oceanography*, 16(11):1981–1984, 1986.

Alain Colin de Verdière and Thierry Huck. Baroclinic instability: An oceanic wavemaker for interdecadal variability. *Journal of Physical Oceanography*, 29(5): 893–910, 1999.

Matthew Collins and Bablu Sinha. Predictability of decadal variations in the thermohaline circulation and climate. *Geophysical Research Letters*, 30(6), 2003.

Meghan F Cronin and Tomoki Tozuka. Steady state ocean response to wind forcing in extratropical frontal regions. *Scientific Reports*, 6:28842, 2016.

Ulrich Cubasch, GA Meehl, GJ Boer, RJ Stouffer, M Dix, A Noda, CA Senior, S Raper, and KS Yap. Projections of future climate change. In *Climate Change 2001: The scientific basis. Contribution of WG1 to the Third Assessment Report of the IPCC (TAR)*, pages 525–582. Cambridge University Press, 2001.

Stuart A Cunningham, Torsten Kanzow, Darren Rayner, Molly O Baringer, William E Johns, Jochem Marotzke, Hannah R Longworth, Elizabeth M Grant, Joël J-M Hirschi, Lisa M Beal, et al. Temporal variability of the Atlantic meridional overturning circulation at 26.5° N. *Science*, 317(5840):935–938, 2007.

Stuart A Cunningham, Christopher D Roberts, Eleanor Frajka-Williams, William E Johns, Will Hobbs, Matthew D Palmer, Darren Rayner, David A Smeed, and Gerard McCarthy. Atlantic Meridional Overturning Circulation slowdown cooled the subtropical ocean. *Geophysical Research Letters*, 40(23):6202–6207, 2013.

Gokhan Danabasoglu, James C McWilliams, and Peter R Gent. The role of mesoscale tracer transports in the global ocean circulation. *Science*, 264(5162):1123–1126, 1994.

Willi Dansgaard, Sigfús J Johnsen, Henrik B Clausen, D Dahl-Jensen, NS Gundestrup, CU Hammer, CS Hvidberg, JP Steffensen, AE Sveinbjörnsdottir, Jean Jouzel, et al. Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, 364(6434):218–220, 1993.

RE Davis, LA Regier, J Dufour, and DC Webb. The Autonomous LAgrangian Circulation Explorer (ALACE). *Journal of Atmospheric and Oceanic Technology*, 9 (3):264–285, 1992.

Xujing Jia Davis, Fiammetta Straneo, Young-Oh Kwon, Kathryn A. Kelly, and John M. Toole. Evolution and formation of North Atlantic Eighteen Degree Water in the Sargasso Sea from moored data. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 91:11–24, 2013. ISSN 09670645. .

Timothy DelSole. Decadal prediction of temperature: Achievements and future prospects. *Current Climate Change Reports*, 3(2):99–111, 2017.

Thomas L Delworth, Syukuru Manabe, and Ronald J Stouffer. Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. *Journal of Climate*, 6(11):1993–2011, 1993.

Thomas L Delworth, Fanrong Zeng, Gabriel A Vecchi, Xiaosong Yang, Liping Zhang, and Rong Zhang. The North Atlantic Oscillation as a driver of rapid climate change in the Northern Hemisphere. *Nature Geoscience*, 9(7):509–512, 2016.

Thomas L Delworth, Fanrong Zeng, Liping Zhang, Rong Zhang, Gabriel A Vecchi, and Xiaosong Yang. The central role of ocean dynamics in connecting the North Atlantic Oscillation to the extratropical component of the Atlantic Multidecadal Oscillation. *Journal of Climate*, 30(10):3789–3805, 2017.

Bob Dickson, Igor Yashayaev, Jens Meincke, Bill Turrell, Stephen Dye, and Juergen Holfort. Rapid freshening of the deep North Atlantic Ocean over the past four decades. *Nature*, 416(6883):832–837, 2002.

Paul Adrien Maurice Dirac. A new notation for quantum mechanics. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 35, pages 416–418. Cambridge Univ Press, 1939.

Dietmar Dommenget. The slab ocean El Niño. *Geophysical Research Letters*, 37(20), 2010.

Dietmar Dommenget and Mojib Latif. Analysis of observed and simulated SST spectra in the midlatitudes. *Climate Dynamics*, 19(3-4):277–288, 2002.

B-W Dong and RT Sutton. The dominant mechanisms of variability in Atlantic ocean heat transport in a coupled ocean-atmosphere GCM. *Geophysical Research Letters*, 28(12):2445–2448, 2001.

Elizabeth M Douglass, Young-Oh Kwon, and Steven R Jayne. A comparison of North Pacific and North Atlantic subtropical mode waters in a climatologically-forced model. *Deep Sea Research Part II: Topical Studies in Oceanography*, 91:139–151, 2013.

Sybren S Drijfhout, Geert Jan Van Oldenborgh, and Andrea Cimatoribus. Is a decline of AMOC causing the warming hole above the North Atlantic in observed and modeled warming patterns? *Journal of Climate*, 25(24):8373–8379, 2012.

Sybren S Drijfhout, David P Marshall, and Henk A Dijkstra. Conceptual models of the wind-driven and thermohaline circulation. In *Ocean Circulation and Climate: A 21*$^{\text{st}}$ *century perspective*, volume 103, pages 257–282. Elsevier, 2013.

Jinqiao Duan and Wei Wang. Stochastic Calculus in Hilbert Space. In *Effective dynamics of stochastic partial differential equations*, chapter 3, pages 21–45. Elsevier, 2014.

Aurélie Duchez, Peggy Courtois, Elizabeth Harris, Simon A Josey, Torsten Kanzow, Robert Marsh, DA Smeed, and JJ-M Hirschi. Potential for seasonal prediction of Atlantic sea surface temperatures using the RAPID array at 26° N. *Climate Dynamics*, 46(9-10):3351–3370, 2016.

J-L Dufresne, M-A Foujols, Sébastien Denvil, Arnaud Caubel, Olivier Marti, Olivier Aumont, Yves Balkanski, Slimane Bekki, Hugo Bellenger, Rachid Benshila, et al. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics*, 40(9-10):2123–2165, 2013.

Vagn Walfrid Ekman. On the influence of the earth's rotation on ocean-currents. *Arkiv for Matematik, Astronomi och Fysik*, 2(11), 1905.

Henry Ellis. XXXII. A letter to the Rev. Dr. Hales, FRS from Captain Henry Ellis, FRS dated Jan. 7, 1750-51, at Cape Monte Africa, ship Earl of Hallifax. *Philosophical Transactions of the Royal Society of London*, 47:211–216, 1752.

Matthew H England and Ernst Maier-Reimer. Using chemical tracers to assess ocean models. *Reviews of Geophysics*, 39(1):29–70, 2001.

Ronald M. Errico. What is an adjoint model? *Bulletin of the American Meteorological Society*, 78(11):2577–2591, 1997. ISSN 00030007. .

Victor Estella-Perez. *Mechanisms of Atlantic variability and sensitivity of the Atlantic to surface initial conditions*. PhD thesis, University of Southampton, 2019.

Riccardo Farneti. Modelling interdecadal climate variability and the role of the ocean. *Wiley Interdisciplinary Reviews: Climate Change*, 8(1):e441, 2017.

Brian F Farrell and Petros J Ioannou. Generalized stability theory. Part I: Autonomous operators. *Journal of the Atmospheric Sciences*, 53(14):2025–2040, 1996.

Brian F Farrell and Petros J Ioannou. Accurate low-dimensional approximation of the linear dynamics of fluid flow. *Journal of the Atmospheric Sciences*, 58:2771–2789, 2001.

Manuel E Fiadeiro and George Veronis. On weighted-mean schemes for the finite-difference approximation to the advection-diffusion equation. *Tellus*, 29(6): 512–522, 1977.

T. Fichefet and M. A. Morales Maqueda. Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. *Journal of Geophysical Research*, 102(C6):12609, 1997. ISSN 0148-0227. .

Jürgen Fischer and Friedrich A Schott. Labrador Sea Water tracked by profiling floats—From the boundary current into the open North Atlantic. *Journal of Physical Oceanography*, 32(2):573–584, 2002.

CK Folland, DE Parker, and FE Kates. Worldwide marine temperature fluctuations 1856–1981. *Nature*, 310(5979):670–673, 1984.

Gaël Forget, Guillaume Maze, Martha Buckley, and John Marshall. Estimated seasonal cycle of North Atlantic eighteen degree water volume. *Journal of Physical Oceanography*, 41(2):269–286, 2011. ISSN 0022-3670. .

Baylor Fox-Kemper, Alistair Adcroft, Claus W Böning, Eric P Chassignet, Enrique Curchitser, Gokhan Danabasoglu, Carsten Eden, Matthew H England, Rüdiger Gerdes, Richard J Greatbatch, et al. Challenges and prospects in ocean circulation models. *Frontiers in Marine Science*, 6:65, 2019.

Eleanor Frajka-Williams, Isabelle J Ansorge, Johanna Baehr, Harry L Bryden, Maria Paz Chidichimo, Stuart A Cunningham, Gokhan Danabasoglu, Shenfu Dong, Kathleen A Donohue, Shane Elipot, et al. Atlantic Meridional Overturning Circulation: Observed Transports and Variability. *Frontiers in Marine Science*, 6: 260, 2019.

Leela M Frankcombe, Henk A Dijkstra, and Anna Von der Heydt. Noise-induced multidecadal variability in the North Atlantic: Excitation of normal modes. *Journal of Physical Oceanography*, 39(1):220–233, 2009.

Claude Frankignoul and Klaus Hasselmann. Stochastic climate models, part ii application to sea-surface temperature anomalies and thermocline variability. *Tellus*, 29(4):289–305, 1977.

Benjamin Franklin. A Letter from Dr. Benjamin Franklin, to Mr. Alphonsus le Roy, Member of Several Academies, at Paris. Containing Sundry Maritime Observations. *Transactions of the American Philosophical Society*, 2:294–329, 1786.

David M. Fratantoni, Young-Oh Kwon, and Benjamin A. Hodges. Direct observation of subtropical mode water circulation in the western North Atlantic Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 91:35–56, 2013. ISSN 09670645. .

Lee-Lueng Fu, Edward J Christensen, Charles A Yamarone, Michel Lefebvre, Yves Menard, Michel Dorrer, and Philippe Escudier. TOPEX/POSEIDON mission overview. *Journal of Geophysical Research: Oceans*, 99(C12):24369–24381, 1994.

Ichiro Fukumori, Tong Lee, Benny Cheng, and Dimitris Menemenlis. The origin, pathway, and destination of Niño-3 water estimated by a simulated passive tracer and its adjoint. *Journal of Physical Oceanography*, 34(3):582–604, 2004.

Shan Gao, Tangdong Qu, and Ichiro Fukumori. Effects of mixing on the subduction of South Pacific waters identified by a simulated passive tracer and its adjoint. *Dynamics of Atmospheres and Oceans*, 51(1-2):45–54, 2011.

Shan Gao, Tangdong Qu, and Dunxin Hu. Origin and pathway of the Luzon Undercurrent identified by a simulated adjoint tracer. *Journal of Geophysical Research: Oceans*, 117(C5), 2012.

HE Garcia, TP Boyer, RA Locarnini, OK Baranova, and MM Zweng. World Ocean Database 2018: User's Manual (prerelease). *AV Mishonov, Technical Ed., NOAA, Silver Spring, MD (Available at https://www. NCEI. noaa. gov/OC5/WOD/pr_wod. html)*, 2018.

Oluwayemi A Garuba and Barry A Klinger. Ocean heat uptake and interbasin transport of the passive and redistributive components of surface heating. *Journal of Climate*, 29(20):7507–7527, 2016.

Oluwayemi A Garuba and Barry A Klinger. The role of individual surface flux components in the passive and active ocean heat uptake. *Journal of Climate*, 31(15): 6157–6173, 2018.

Stefan F Gary, M Susan Lozier, Claus W Böning, and Arne Biastoch. Deciphering the pathways for the deep limb of the Meridional Overturning Circulation. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(17-18):1781–1797, 2011.

Stefan F Gary, M Susan Lozier, Y-O Kwon, and Jong Jin Park. The fate of North Atlantic subtropical mode water in the FLAME model. *Journal of Physical Oceanography*, 44(5):1354–1371, 2014.

Rüdiger Gerdes, Cornelia Köberle, and Jürgen Willebrand. The influence of numerical advection schemes on the results of ocean general circulation models. *Climate Dynamics*, 5(4):211–226, 1991.

Ralf Giering and Thomas Kaminski. Recipes for adjoint code construction. *ACM Transactions on Mathematical Software (TOMS)*, 24(4):437–474, 1998.

Paul B Goddard, Jianjun Yin, Stephen M Griffies, and Shaoqing Zhang. An extreme event of sea-level rise along the Northeast coast of North America in 2009–2010. *Nature Communications*, 6(1):1–9, 2015.

Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, third edition, 1996.

Richard J Greatbatch and Sheng Zhang. An interdecadal oscillation in an idealized ocean basin forced by constant heat flux. *Journal of Climate*, 8(1):81–91, 1995.

MC Gregg. Scaling turbulent dissipation in the thermocline. *Journal of Geophysical Research: Oceans*, 94(C7):9686–9698, 1989.

Sandy Grégorio, Thierry Penduff, Guillaume Sérazin, Jean-Marc Molines, Bernard Barnier, and Joël J-M Hirschi. Intrinsic Variability of the Atlantic Meridional Overturning Circulation at Interannual-to-Multidecadal time scales. *Journal of Physical Oceanography*, 45(7):1929–1946, 2015.

Philip M Gresho and Robert L Lee. Don't suppress the wiggles—they're telling you something! *Computers & Fluids*, 9(2):223–253, 1981.

Stephen M Griffies and Kirk Bryan. Predictability of North Atlantic multidecadal climate variability. *Science*, 275(5297):181–184, 1997.

Stephen M Griffies and Anne Marie Treguier. Ocean circulation models and modeling. In *International Geophysics*, volume 103, pages 521–551. Elsevier, 2013.

Stephen M Griffies and Eli Tziperman. A linear thermohaline oscillator driven by stochastic atmospheric forcing. *Journal of Climate*, 8(10):2440–2453, 1995.

A Grötzner, Mojib Latif, Axel Timmermann, and Reinhard Voss. Interannual to decadal predictability in a coupled ocean–atmosphere general circulation model. *Journal of Climate*, 12(8):2607–2624, 1999.

Nicolas Gruber, Dominic Clement, Brendan R Carter, Richard A Feely, Steven Van Heuven, Mario Hoppema, Masao Ishii, Robert M Key, Alex Kozyr, Siv K Lauvset, et al. The oceanic sink for anthropogenic CO_2 from 1994 to 2007. *Science*, 363(6432):1193–1199, 2019.

Reindert J Haarsma, Malcolm J Roberts, Pier Luigi Vidale, Catherine A Senior, Alessio Bellucci, Qing Bao, Ping Chang, Susanna Corti, Neven S Fučkar, Virginie Guemas, et al. High resolution model intercomparison project (HighResMIP v1. 0) for CMIP6. *Geoscientific Model Development*, 9(11):4185–4208, 2016.

Mindy M Hall and Harry L Bryden. Direct estimates and mechanisms of ocean heat transport. *Deep Sea Research Part I: Oceanographic Research Papers*, 29(3): 339–359, 1982.

Edmond Halley. IX. A discourse concerning the proportional heat of the sun in all latitudes, with the method of collecting the same, as it was read before the Royal Society in one of their late meetings. *Philosophical Transactions of the Royal Society of London*, 17(203):878–885, 1693.

Kimio Hanawa and Lynne D Talley. Mode waters. *International Geophysics Series*, 77: 373–386, 2001.

Bogi Hansen and S Østerhus. North Atlantic–Nordic seas exchanges. *Progress in Oceanography*, 45(2):109–208, 2000.

Bogi Hansen, William R Turrell, and Sven Osterhus. Decreasing overflow from the Nordic seas into the Atlantic Ocean through the Faroe Bank channel since 1950. *Nature*, 411(6840):927, 2001.

Klaus Hasselmann. Stochastic climate models part I. Theory. *Tellus*, 28(6):473–485, 1976.

Ed Hawkins and Rowan Sutton. The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, 90(8): 1095–1108, 2009.

Patrick Heimbach, Carl Wunsch, Rui M Ponte, Gael Forget, Chris Hill, and Jean Utke. Timescales and regions of the sensitivity of Atlantic meridional volume and heat transport: Toward observing system design. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(17-18):1858–1879, 2011.

Chris Hill, Véronique Bugnion, Mick Follows, and John Marshall. Evaluating carbon sequestration efficiency in an ocean circulation model by adjoint sensitivity analysis. *Journal of Geophysical Research: Oceans*, 109(C11), 2004.

Joël J-M Hirschi, Peter D Killworth, and Jeffrey R Blundell. Subannual, seasonal, and interannual variability of the North Atlantic meridional overturning circulation. *Journal of Physical Oceanography*, 37(5):1246–1265, 2007.

Joël J-M Hirschi, Adam T Blaker, Bablu Sinha, Andrew C Coward, Beverly A de Cuevas, Steven G Alderson, and Gurvan Madec. Chaotic variability of the

meridional overturning circulation on subannual to interannual timescales. *Ocean Science*, 9:805–823, 2013.

Anthony C Hirst. Determination of water component age in ocean models: Application to the fate of North Atlantic Deep Water. *Ocean Modelling*, 1(2-4):81–94, 1999.

Antoine Hochet, Thierry Huck, Olivier Arzel, Florian Sévellec, Alain Colin de Verdière, Matthew Mazloff, and Bruce Cornuelle. Direct temporal cascade of temperature variance in eddy-permitting simulations of multidecadal variability. *Journal of Climate*, 33(21):9409–9425, 2020.

I Hoteit, B Cornuelle, A Köhl, and D Stammer. Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3659–3682, 2005.

Frédéric Hourdin, Marie-Alice Foujols, Francis Codron, Virginie Guemas, Jean-Louis Dufresne, Sandrine Bony, Sébastien Denvil, Lionel Guez, François Lott, Josefine Ghattas, et al. Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. *Climate Dynamics*, 40(9-10): 2167–2192, 2013.

Thierry Huck and Geoffrey K Vallis. Linear stability analysis of the three-dimensional thermally-driven ocean circulation: application to interdecadal oscillations. *Tellus A*, 53(4):526–545, 2001.

Thierry Huck, Alain Colin de Verdière, and Andrew J Weaver. Interdecadal variability of the thermohaline circulation in box-ocean models forced by fixed surface fluxes. *Journal of Physical Oceanography*, 29(5):865–892, 1999.

Thierry Huck, Olivier Arzel, and Florian Sévellec. Multidecadal variability of the overturning circulation in presence of eddy turbulence. *Journal of Physical Oceanography*, 45(1):157–173, 2015.

IPCC. Climate change 2013: the physical science basis. Intergovernmental panel on climate change, working group I contribution to the IPCC fifth assessment report (AR5), 2013.

C O Iselin. The influence of vertical and lateral turbulence on the characteristics of the waters at mid-depths. *Eos, Transactions American Geophysical Union*, 20(3): 414–417, 1939.

Kiyosi Itô. Stochastic integral. *Proceedings of the Imperial Academy*, 20(8):519–524, 1944.

Laura C Jackson, K Andrew Peterson, Chris D Roberts, and Richard A Wood. Recent slowing of Atlantic overturning circulation as a recovery from earlier strengthening. *Nature Geoscience*, 9(7):518, 2016.

Quentin Jamet, WK Dewar, N Wienders, and Bruno Deremble. Spatiotemporal Patterns of Chaos in the Atlantic Overturning Circulation. *Geophysical Research Letters*, 46(13):7509–7517, 2019.

Quentin Jamet, William K Dewar, Nicolas Wienders, Bruno Deremble, Sally Close, and Thierry Penduff. Locally and remotely forced subtropical AMOC variability: A matter of time scales. *Journal of Climate*, 33(12):5155–5172, 2020.

Alvin Jewett Johnson. *Johnson's New Illustrated (Steel Plate) Family Atlas, With Descriptions, Geographical, Statistical, And Historical (1870 Ed.)*. Johnson and Ward, 1870.

Gregory C Johnson. Quantifying Antarctic bottom water and North Atlantic deep water volumes. *Journal of Geophysical Research: Oceans*, 113(C5), 2008.

Helen L Johnson, Paola Cessi, David P Marshall, Fabian Schloesser, and Michael A Spall. Recent contributions of theory to our understanding of the Atlantic Meridional Overturning Circulation. *Journal of Geophysical Research: Oceans*, 124 (8):5376–5399, 2019.

Daniel C Jones, Gael Forget, Bablu Sinha, Simon A Josey, Emma JD Boland, Andrew JS Meijers, and Emily Shuckburgh. Local and remote influences on the heat content of the Labrador Sea: An adjoint sensitivity study. *Journal of Geophysical Research: Oceans*, 123(4):2646–2667, 2018.

Marieke Femke Jong and Laura Steur. Strong winter cooling over the Irminger Sea in winter 2014–2015, exceptional deep convection, and the emergence of anomalously low SST. *Geophysical Research Letters*, 43(13):7106–7113, 2016.

Terrence M Joyce. New perspectives on eighteen-degree water formation in the North Atlantic. *Journal of Oceanography*, 68(1):45–52, 2011. ISSN 0916-8370. .

Glenn H Jung. Note on the meridional transport of energy by the oceans. *Journal of Marine Research*, 11(2):139–146, 1952.

Torsten Kanzow, Stuart A Cunningham, Darren Rayner, Joël J-M Hirschi, William E Johns, Molly O Baringer, Harry L Bryden, Lisa M Beal, Christopher S Meinen, and Jochem Marotzke. Observed flow compensation associated with the MOC at 26.5° N in the Atlantic. *Science*, 317(5840):938–941, 2007.

Torsten Kanzow, Stuart A Cunningham, WE Johns, J Hirschi, J Marotzke, MO Baringer, CS Meinen, MP Chidichimo, C Atkinson, LM Beal, et al. Seasonal variability of the Atlantic Meridional Overturning Circulation at 26.5° n. *Journal of Climate*, 23(21):5678–5698, 2010.

Richard A Kerr. A North Atlantic climate pacemaker for the centuries. *Science*, 288 (5473):1984–1985, 2000.

Samar Khatiwala, Martin Visbeck, and Mark A Cane. Accelerated simulation of passive tracers in ocean circulation models. *Ocean Modelling*, 9(1):51–69, 2005.

Ben P Kirtman, Natalie Perlin, and Leo Siqueira. Ocean eddies and climate predictability. *Chaos*, 27(12):126902, 2017.

Birgit Klein and Nelson Hogg. On the variability of 18 degree water formation as observed from moored instruments at 55°W. *Deep Sea Research Part I: Oceanographic Research Papers*, 43(11):1777–1806, 1996.

Yavor Kostov, Kyle C Armour, and John Marshall. Impact of the Atlantic meridional overturning circulation on ocean heat storage and transient climate change. *Geophysical Research Letters*, 41(6):2108–2116, 2014.

Yavor Kostov, Helen L Johnson, and David P Marshall. AMOC sensitivity to surface buoyancy fluxes: the role of air-sea feedback mechanisms. *Climate Dynamics*, 53 (7-8):4521–4537, 2019.

Sergey Kravtsov. Dynamics and predictability of hemispheric-scale multidecadal climate variability in an observationally constrained mechanistic model. *Journal of Climate*, 33(11):4599–4620, 2020.

Yochanan Kushnir. Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *Journal of Climate*, 7(1):141–157, 1994.

Young-Oh Kwon and Stephen C. Riser. North Atlantic subtropical mode water: A history of ocean-atmosphere interaction 1961-2000. *Geophysical Research Letters*, 31: L19307, 2004. ISSN 00948276. .

Young-Oh Kwon and Stephen C. Riser. General circulation of the western subtropical North Atlantic observed using profiling floats. *Journal of Geophysical Research: Oceans*, 110(10):1–22, 2005. ISSN 01480227. .

Young-Oh Kwon, Jong-Jin Park, Stefan F Gary, and M Susan Lozier. Year-to-year reoutcropping of eighteen degree water in an eddy-resolving ocean simulation. *Journal of Physical Oceanography*, 45(4):1189–1204, 2015.

JH LaCasce and J Pedlosky. The instability of Rossby basin modes and the oceanic eddy field. *Journal of physical oceanography*, 34(9):2027–2041, 2004.

William George Large and Stephen G Yeager. Diurnal to decadal global forcing for ocean and sea-ice models: the data sets and flux climatologies. *NCAR Technical Note*, NCAR/TN-460+STR, 2004.

Mojib Latif and Noel S Keenlyside. A perspective on decadal climate variability and predictability. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58 (17-18):1880–1894, 2011.

Mojib Latif, M Collins, Holger Pohlmann, and Noel Keenlyside. A review of predictability studies of atlantic sector climate on decadal time scales. *Journal of Climate*, 19(23):5971–5987, 2006.

Kara L Lavender, Russ E Davis, and W Brechner Owens. Mid-depth recirculation observed in the interior Labrador and Irminger seas by direct velocity measurements. *Nature*, 407(6800):66, 2000.

Kevin D Leaman and Jessie E Harris. On the average absolute transport of the deep western boundary currents east of Abaco Island, the Bahamas. *Journal of Physical Oceanography*, 20(3):467–475, 1990.

Deborah A LeBel, William M Smethie, Monika Rhein, Dagmar Kieke, Rana A Fine, John L Bullister, Dong-Ha Min, Wolfgang Roether, Ray F Weiss, Chantal Andrié, et al. The formation rate of North Atlantic Deep Water and Eighteen Degree Water calculated from CFC-11 inventories observed during WOCE. *Deep Sea Research Part I: Oceanographic Research Papers*, 55(8):891–910, 2008.

Emil Lenz. Bemerkungen über die Temperatur des Weltmeers in verschiedenen Tiefen. *Bulletin de la Classe Physico-Mathématique de l'Académie Impériale des Sciences de Saint-Pétersbourg*, 5:67–74, 1845.

Stephanie Leroux, Thierry Penduff, Laurent Bessières, Jean-Marc Molines, Jean-Michel Brankart, Guillaume Sérazin, Bernard Barnier, and Laurent Terray. Intrinsic and atmospherically forced variability of the AMOC: insights from a large-ensemble ocean hindcast. *Journal of Climate*, 31(3):1183–1203, 2018.

Marina Lévy, Audrey Estublier, and Gurvan Madec. Choice of an advection scheme for biogeochemical models. *Geophysical Research Letters*, 28(19):3725–3728, 2001.

Benjamin Lindner. A brief introduction to some simple stochastic processes. In Carlo Laing and Gabriel J Lord, editors, *Stochastic Methods in Neuroscience*, chapter 1, pages 1–28. Oxford University Press, 2009.

Zhengyu Liu. Planetary wave modes in the thermocline: Non-doppler-shift mode, advective mode and green mode. *Quarterly Journal of the Royal Meteorological Society*, 125(556):1315–1339, 1999.

Zhengyu Liu. Dynamics of interdecadal climate variability: A historical perspective. *Journal of Climate*, 25(6):1963–1995, 2012.

Gerrit Lohmann and Joachim Schneider. Dynamics and predictability of Stommel's box model. A phase-space perspective with implications for decadal climate variability. *Tellus A*, 51(2):326–336, 1999.

Nora Loose, Patrick Heimbach, HR Pillar, and KH Nisancioglu. Quantifying dynamical proxy potential through shared adjustment physics in the North Atlantic. *Journal of Geophysical Research: Oceans*, 125(9):e2020JC016112, 2020.

Edward N Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, 1963.

Edward N Lorenz. The predictability of a flow which possesses many scales of motion. *Tellus*, 21(3):289–307, 1969.

Edward N Lorenz. Does the flap of a butterfly's wings in Brazil set off a tornado in Texas? In *Global Atmospheric Research Program. 139th Meeting of the American Association for the Advancement of Science. Unpublished*, 1972.

Edward N Lorenz. Climatic predictability. *The physical basis of climate and climate modelling*, pages 132–136, 1975.

M Susan Lozier and Nicole M Stewart. On the temporally varying northward penetration of Mediterranean Overflow Water and eastward penetration of Labrador Sea Water. *Journal of Physical Oceanography*, 38(9):2097–2103, 2008.

J. R. Luyten, J. Pedlosky, and H. Stommel. The ventilated thermocline. *Journal of Physical Oceanography*, 13(2):292–309, 1983. .

Doug Macdougall. *Endless Novelties of Extraordinary Interest: The Voyage of HMS Challenger and the Birth of Modern Oceanography*. Yale University Press, 2019.

Andreas Macrander, Uwe Send, H Valdimarsson, S Jónsson, and Rolf H Käse. Interannual changes in the overflow from the Nordic Seas into the Atlantic Ocean through Denmark Strait. *Geophysical Research Letters*, 32(6), 2005.

Gurvan Madec. the NEMO team: NEMO ocean engine–Version 3.4. *Note du Pôle de modélisation. Institut Pierre-Simon Laplace (IPSL), France*, 27, 2012.

Gurvan Madec and Maurice Imbard. A global ocean mesh to overcome the North Pole singularity. *Climate Dynamics*, 12(6):381–388, 1996. ISSN 0930-7575. .

Jochem Marotzke, Ralf Giering, Kate Q Zhang, Detlef Stammer, Chris Hill, and Tong Lee. Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity. *Journal of Geophysical Research: Oceans*, 104(C12):29529–29547, 1999.

David P Marshall and Helen L Johnson. Propagation of meridional circulation anomalies along western and eastern boundaries. *Journal of Physical Oceanography*, 43(12):2699–2717, 2013.

John Marshall, Raffaele Ferrari, G Forget, G Maze, A Andersson, N Bates, W Dewar, S Doney, D Fratantoni, T Joyce, et al. The CLIMODE field campaign: Observing the cycle of convection and restratification over the Gulf Stream. *Bulletin of the American Meteorological Society*, 90(9):1337–1350, 2009.

John Marshall, Jeffery R Scott, Kyle C Armour, J-M Campin, Maxwell Kelley, and Anastasia Romanou. The ocean's role in the transient response of climate to abrupt greenhouse gas forcing. *Climate Dynamics*, 44(7-8):2287–2299, 2015.

C Mauritzen, A Melsom, and Rowan T Sutton. Importance of density-compensated temperature change for deep North Atlantic Ocean heat uptake. *Nature Geoscience*, 5(12):905–910, 2012.

Matthew Fontaine Maury. *The physical geography of the sea, and its meteorology.* Sampson Low, Marston, Low and Searle, 1855.

Björn Mayer, André Düsterhus, and Johanna Baehr. When does the Lorenz 1963 model exhibit the signal-to-noise paradox? *Geophysical Research Letters*, 48(4), 2021.

Gerard McCarthy, Eleanor Frajka-Williams, William E Johns, Molly O Baringer, Christopher S Meinen, HL Bryden, Darren Rayner, A Duchez, C Roberts, and Stuart A Cunningham. Observed interannual variability of the Atlantic meridional overturning circulation at 26.5° N. *Geophysical Research Letters*, 39(19), 2012.

Michael S McCartney and Lynne D Talley. The subpolar mode water of the North Atlantic Ocean. *Journal of Physical Oceanography*, 12(11):1169–1188, 1982.

Seth McGinnis. Big Data Lessons from the Climate Science Community, 2016. URL https://slideplayer.com/slide/9420952/. National Center for Atmospheric Research presentation, 2016-01-14.

Gerald A Meehl, Lisa Goddard, James Murphy, Ronald J Stouffer, George Boer, Gokhan Danabasoglu, Keith Dixon, Marco A Giorgetta, Arthur M Greene, ED Hawkins, et al. Decadal prediction: Can it be skillful? *Bulletin of the American Meteorological Society*, 90(10):1467–1486, 2009.

Gerald A Meehl, Julie M Arblaster, John T Fasullo, Aixue Hu, and Kevin E Trenberth. Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Climate Change*, 1(7):360–364, 2011.

Gerald A Meehl, Lisa Goddard, George Boer, Robert Burgman, Grant Branstator, Christophe Cassou, Susanna Corti, Gokhan Danabasoglu, Francisco Doblas-Reyes, Ed Hawkins, et al. Decadal climate prediction: an update from the trenches. *Bulletin of the American Meteorological Society*, 95(2):243–267, 2014.

Alfred Merz. *Die Deutsche Atlantische Expedition auf dem Vermessungs-und Forschungsschiff" Meteor": I. Bericht.* Akad. d. Wiss., 1925.

Robert L Molinari, Rana A Fine, and Elizabeth Johns. The deep western boundary current in the tropical North Atlantic Ocean. *Deep Sea Research Part I. Oceanographic Research Papers*, 39(11-12):1967–1984, 1992.

Adam H Monahan, Julie Alexander, and Andrew J Weaver. Stochastic models of the meridional overturning circulation: time scales and patterns of variability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1875):2525–2542, 2008.

R Msadek, KW Dixon, TL Delworth, and W Hurlin. Assessing the predictability of the Atlantic meridional overturning circulation and associated fingerprints. *Geophysical Research Letters*, 37(19), 2010.

Mu Mu, Liang Sun, and Henk A Dijkstra. The sensitivity and stability of the ocean's thermohaline circulation to finite-amplitude perturbations. *Journal of Physical Oceanography*, 34(10):2305–2315, 2004.

Les C Muir and Alexey V Fedorov. Evidence of the AMOC interdecadal mode related to westward propagation of temperature anomalies in CMIP5 models. *Climate Dynamics*, 48(5-6):1517–1535, 2017.

Simon Müller and Florian Sévellec. An Iterative Solver for Time-Offset Model-Data Bias Reduction: an implementation for ocean predictability studies using NEMO and NEMOTAM. in prep., 2020.

Simon Müller, Florian Sévellec, and Dafydd Stephenson. Towards a nonlinear perspective on the interannual predictability of the Atlantic meridional overturning circulation. in prep., 2020.

Walter Heinrich Munk. Abyssal recipes. *Deep Sea Research and Oceanographic Abstracts*, 13(4):707–730, 1966.

Walter Heinrich Munk. Achievements in physical oceanography. *50 Years of Ocean Discovery: National Science Foundation 1950-2000*, page 44, 2000.

Walter Heinrich Munk and Carl Isaac Wunsch. Observing the ocean in the 1990s. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 307(1499):439–464, 1982.

Emily Newsom, Laure Zanna, Samar Khatiwala, and Jonathan M Gregory. The influence of warming patterns on passive ocean heat uptake. *Geophysical Research Letters*, page e2020GL088429, 2020.

An T Nguyen, Helen Pillar, Victor Ocaña, Arash Bigdeli, Timothy A Smith, and Patrick Heimbach. The Arctic Subpolar gyre sTate Estimate (ASTE): Description and assessment of a data-constrained, dynamically consistent ocean-sea ice estimate for 2002-2017. *Earth and Space Science Open Archive ESSOAr*, 2020.

AH Orsi, GC Johnson, and JL Bullister. Circulation, mixing, and production of Antarctic Bottom Water. *Progress in Oceanography*, 43(1):55–109, 1999.

A Owen. Artificial diffusion in the numerical modelling of the advective transport of salinity. *Applied Mathematical Modelling*, 8(2):116–120, 1984.

Timothy Noel Palmer, R Buizza, F Molteni, Y-Q Chen, and S Corti. Singular vectors and the predictability of weather and climate. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, 348(1688): 459–475, 1994.

Victor Y Pan and Zhao Q Chen. The complexity of the matrix eigenproblem. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 507–516, 1999.

Seon Ki Park and Liang Xu. *Data assimilation for atmospheric, oceanic and hydrologic applications*, volume 2. Springer Science & Business Media, 2013.

Thierry Penduff, Guillaume Sérazin, Stephanie Leroux, Sally Close, Jean-Marc Molines, Bernard Barnier, Laurent Bessières, Laurent Terray, and Guillaume Maze. Chaotic variability of ocean heat content: climate-relevant features and observational implications. *Oceanography*, 31(2):63–71, 2018.

Cécile Penland and Theresa Magorian. Prediction of Niño 3 sea surface temperatures using linear inverse modeling. *Journal of Climate*, 6(6):1067–1076, 1993.

Aurélie Persechino, Juliette Mignot, Didier Swingedouw, Sonia Labetoulle, and Eacute Guilyardi. Decadal predictability of the Atlantic meridional overturning circulation and climate in the IPSL-CM5A-LR model. *Climate Dynamics*, 40(9-10):2359–2380, 2013.

Jean-Robert Petit, Jean Jouzel, Dominique Raynaud, Narcisse I Barkov, J-M Barnola, Isabelle Basile, Michael Bender, Jérôme Chappellaz, M Davis, Gilles Delaygue, et al. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature*, 399(6735):429, 1999.

Robert S Pickart, Michael A Spall, Mads Hvid Ribergaard, GWK Moore, and Ralph F Milliff. Deep convection in the Irminger Sea forced by the Greenland tip jet. *Nature*, 424(6945):152, 2003a.

Robert S Pickart, Fiammetta Straneo, and GWK Moore. Is Labrador Sea water formed in the Irminger basin? *Deep Sea Research Part I: Oceanographic Research Papers*, 50(1):23–52, 2003b.

Helen R Pillar, Patrick Heimbach, Helen L Johnson, and David P Marshall. Dynamical attribution of recent variability in Atlantic overturning. *Journal of Climate*, 29(9): 3339–3352, 2016.

Holger Pohlmann, Michael Botzet, Mojib Latif, Andreas Roesch, Martin Wild, and Peter Tschuck. Estimating the decadal predictability of a coupled AOGCM. *Journal of Climate*, 17(22):4463–4472, 2004.

Irene Polo, Jon Robson, Rowan Sutton, and Magdalena Alonso Balmaseda. The importance of wind and buoyancy forcing for the boundary density variations and the geostrophic component of the AMOC at 26°N. *Journal of Physical Oceanography*, 44(9):2387–2408, 2014.

Tangdong Qu, Shan Gao, Ichiro Fukumori, Rana A Fine, and Eric J Lindstrom. Origin and pathway of equatorial 13°C water in the Pacific identified by a simulated passive tracer and its adjoint. *Journal of Physical Oceanography*, 39(8):1836–1853, 2009.

Tangdong Qu, Shan Gao, and Ichiro Fukumori. Formation of salinity maximum water and its contribution to the overturning circulation in the North Atlantic as revealed by a global general circulation model. *Journal of Geophysical Research: Oceans*, 118 (4):1982–1994, 2013.

Detlef Quadfasel and Rolf H Käse. Present-day manifestation of the Nordic Seas overflows. In *Natural Gas Hydrates-Occurence, Distribution and Detection (ed. CK Paull and WP Dillon) American Geophysical Union*, volume 173, pages 75–89. American Geophysical Union, 2007.

Joseph L Reid and Ronald J Lynn. On the influence of the Norwegian-Greenland and Weddell seas upon the bottom waters of the Indian and Pacific oceans. *Deep Sea Research*, 18:1063–1088, 1971.

James Rennell. *An investigation of the currents of the Atlantic Ocean: and of those which prevail between the Indian Ocean and the Atlantic*. Lady Rodd, 1832.

Monika Rhein, Jürgen Fischer, WM Smethie, D Smythe-Wright, RF Weiss, Christian Mertens, D-H Min, Ulrich Fleischmann, and A Putzka. Labrador Sea Water: Pathways, CFC inventory, and formation rates. *Journal of Physical Oceanography*, 32(2):648–665, 2002.

Philip L Richardson. Drifting derelicts in the North Atlantic 1883–1902. *Progress in Oceanography*, 14:463–483, 1985.

Philip L Richardson. On the history of meridional overturning circulation schematic diagrams. *Progress in Oceanography*, 76(4):466–486, 2008.

Stephen C Riser, Howard J Freeland, Dean Roemmich, Susan Wijffels, Ariel Troisi, Mathieu Belbéoch, Denis Gilbert, Jianping Xu, Sylvie Pouliquen, Ann Thresher, et al. Fifteen years of ocean observations with the global Argo array. *Nature Climate Change*, 6(2):145–153, 2016.

CD Roberts, J Waters, KA Peterson, MD Palmer, GD McCarthy, E Frajka-Williams, K Haines, DJ Lea, MJ Martin, D Storkey, et al. Atmosphere drives recent interannual variability of the Atlantic meridional overturning circulation at 26.5°N. *Geophysical Research Letters*, 40(19):5164–5170, 2013.

Jon Robson, Dan Hodson, Ed Hawkins, and Rowan Sutton. Atlantic overturning in decline? *Nature Geoscience*, 7(1):2–3, 2014.

Claes Rooth. Hydrology and ocean circulation. *Progress in Oceanography*, 11(2): 131–149, 1982.

Christopher L Sabine, Richard A Feely, Nicolas Gruber, Robert M Key, Kitack Lee, John L Bullister, Rik Wanninkhof, CSL Wong, Douglas WR Wallace, Bronte Tilbrook, et al. The oceanic sink for anthropogenic CO_2. *Science*, 305(5682): 367–371, 2004.

R Saravanan and James C McWilliams. Stochasticity and spatial resonance in interdecadal climate fluctuations. *Journal of Climate*, 10(9):2299–2320, 1997.

R Saravanan and James C McWilliams. Advective ocean–atmosphere interaction: An analytical stochastic model with implications for decadal variability. *Journal of Climate*, 11(2):165–188, 1998.

Prashant D Sardeshmukh and Philip Sura. Reconciling non-gaussian climate statistics with linear dynamics. *Journal of Climate*, 22(5):1193–1207, 2009.

Adam A Scaife and Doug Smith. A signal-to-noise paradox in climate science. *npj Climate and Atmospheric Science*, 1(1):1–8, 2018.

Edward A Schuert. Turbulent diffusion in the intermediate waters of the North Pacific Ocean. *Journal of Geophysical Research*, 75(3):673–682, 1970.

Robert B Scott. Predictability of SST in an idealized, one-dimensional, coupled atmosphere–ocean climate model with stochastic forcing and advection. *Journal of Climate*, 16(2):323–335, 2003.

Albert J Semtner. Modeling ocean circulation. *Science*, 269(5229):1379–1385, 1995.

Guillaume Sérazin, Thierry Penduff, Sandy Grégorio, Bernard Barnier, Jean-Marc Molines, and Laurent Terray. Intrinsic variability of sea level from global ocean simulations: Spatiotemporal scales. *Journal of Climate*, 28(10):4279–4292, 2015.

Guillaume Sérazin, Benoit Meyssignac, Thierry Penduff, Laurent Terray, Bernard Barnier, and Jean-Marc Molines. Quantifying uncertainties on regional sea level change induced by multidecadal intrinsic oceanic variability. *Geophysical Research Letters*, 43(15):8151–8159, 2016.

Guillaume Sérazin, Alexandre Jaymond, Stéphanie Leroux, Thierry Penduff, Laurent Bessières, William Llovel, Bernard Barnier, Jean-Marc Molines, and Laurent Terray. A global probabilistic study of the ocean heat content low-frequency variability: Atmospheric forcing versus oceanic chaos. *Geophysical Research Letters*, 44(11): 5580–5589, 2017.

Florian Sévellec and Alexey V Fedorov. Stability of the Atlantic meridional overturning circulation and stratification in a zonally averaged ocean model: Effects of freshwater flux, Southern Ocean winds, and diapycnal diffusion. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(17-18):1927–1943, 2011.

Florian Sévellec and Alexey V Fedorov. The leading, interdecadal eigenmode of the Atlantic meridional overturning circulation in a realistic ocean model. *Journal of Climate*, 26(7):2160–2183, 2013a.

Florian Sévellec and Alexey V Fedorov. Model bias reduction and the limits of oceanic decadal predictability: importance of the deep ocean. *Journal of Climate*, 26(11): 3688–3707, 2013b.

Florian Sévellec and Alexey V Fedorov. Optimal excitation of AMOC decadal variability: Links to the subpolar ocean. *Progress in Oceanography*, 132:287–304, 2015.

Florian Sévellec and Alexey V Fedorov. Predictability and decadal variability of the North Atlantic ocean state evaluated from a realistic ocean model. *Journal of Climate*, 30(2):477–498, 2017.

Florian Sévellec and Thierry Huck. Theoretical investigation of the Atlantic Multidecadal Oscillation. *Journal of Physical Oceanography*, 45(9):2189–2208, 2015.

Florian Sévellec and Bablu Sinha. Predictability of Decadal Atlantic Meridional Overturning Circulation Variations. In *Oxford Research Encyclopedia of Climate Science*. Oxford University Press, 2018.

Florian Sévellec, Mahdi Ben Jelloul, and Thierry Huck. Optimal surface salinity perturbations influencing the thermohaline circulation. *Journal of Physical Oceanography*, 37(12):2789–2808, 2007.

Florian Sévellec, Thierry Huck, Mahdi Ben Jelloul, and Jérôme Vialard. Nonnormal multidecadal response of the thermohaline circulation induced by optimal surface salinity perturbations. *Journal of Physical Oceanography*, 39(4):852–872, 2009.

Florian Sévellec, Alain Colin de Verdiére, and Michel Ollitrault. Evolution of intermediate water masses based on Argo float displacements. *Journal of Physical Oceanography*, 47(7):1569–1586, 2017a.

Florian Sévellec, Alexey V Fedorov, and Wei Liu. Arctic sea-ice decline weakens the Atlantic meridional overturning circulation. *Nature Climate Change*, 7(8):604–610, 2017b.

Florian Sévellec, Henk A Dijkstra, Sybren S Drijfhout, and Agathe Germe. Dynamical attribution of oceanic prediction uncertainty in the North Atlantic: application to the design of optimal monitoring systems. *Climate Dynamics*, 51(4):1517–1535, 2018.

Florian Sévellec, Alberto C. Naveira Garabato, and Theirry Huck. Damping of climate-scale oceanic variability by mesoscale eddy turbulence. *Journal of Physical Oceanography*, pages doi: 10.1175/JPO–D–20–0141.1, 2020.

Gerold Siedler, Anni Kuhl, and Walter Zenk. The Madeira mode water. *Journal of Physical Oceanography*, 17(10):1561–1570, 1987.

William M Smethie Jr and Rana A Fine. Rates of North Atlantic Deep Water formation calculated from chlorofluorocarbon inventories. *Deep Sea Research Part I: Oceanographic Research Papers*, 48(1):189–215, 2001.

Timothy Smith and Patrick Heimbach. Atmospheric origins of variability in the South Atlantic meridional overturning circulation. *Journal of Climate*, 32(5):1483–1500, 2019.

Kevin Speer and Eli Tziperman. Rates of water mass formation in the North Atlantic Ocean. *Journal of Physical Oceanography*, 22(1):93–104, 1992.

Abhishekh Srivastava and Timothy DelSole. Decadal predictability without ocean dynamics. *Proceedings of the National Academy of Sciences*, 114(9):2177–2182, 2017.

Dafydd Stephenson and Florian Sévellec. The active and passive roles of the ocean in generating regional heat content variability. manuscript submitted to *Geophysical Research Letters*, 2021a.

Dafydd Stephenson and Florian Sévellec. Dynamical attribution of North Atlantic interdecadal predictability to oceanic and atmospheric turbulence under realistic and optimal stochastic forcing. manuscript submitted to *Journal of Climate*, 2021b.

Dafydd Stephenson, Simon A Müller, and Florian Sévellec. Tracking water masses using passive-tracer transport in NEMO v3. 4 with NEMOTAM: application to North Atlantic Deep Water and North Atlantic Subtropical Mode Water. *Geoscientific Model Development*, 13(4):2031–2050, 2020.

Henry Stommel. The westward intensification of wind-driven ocean currents. *Eos, Transactions American Geophysical Union*, 29(2):202–206, 1948.

Henry Stommel. Thermohaline convection with two stable regimes of flow. *Tellus*, 13 (2):224–230, 1961.

Henry Stommel and AB Arons. On the abyssal circulation of the world ocean — I. Stationary planetary flow patterns on a sphere. *Deep Sea Research*, 6:140–154, 1959.

Henry Stommel and Claes Rooth. On the interaction of gravitational and dynamic forcing in simple circulation models. *Deep Sea Research and Oceanographic Abstracts*, 15(2):165–170, 1968.

Liang Sun, Mu Mu, De-Jun Sun, and Xie-Yuan Yin. Passive mechanism of decadal variation of thermohaline circulation. *Journal of Geophysical Research: Oceans*, 110 (C7), 2005.

Philip Sura and Prashant D Sardeshmukh. A global view of non-gaussian sst variability. *Journal of Physical Oceanography*, 38(3):639–647, 2008.

Harald Ulrich Sverdrup. Wind-driven currents in a baroclinic ocean; with application to the equatorial currents of the eastern Pacific. *Proceedings of the National Academy of Sciences of the United States of America*, 33(11):318, 1947.

James H Swift, Knut Aagaard, and Svend-Aage Malmberg. The contribution of the Denmark Strait overflow to the deep North Atlantic. *Deep Sea Research Part I. Oceanographic Research Papers*, 27(1):29–42, 1980.

Taro Takahashi, Stewart C Sutherland, Colm Sweeney, Alain Poisson, Nicolas Metzl, Bronte Tilbrook, Nicolas Bates, Rik Wanninkhof, Richard A Feely, Christopher Sabine, et al. Global sea–air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(9-10):1601–1622, 2002.

LD Talley and MS McCartney. Distribution and circulation of Labrador Sea water. *Journal of Physical Oceanography*, 12(11):1189–1205, 1982.

Byron D Tapley, S Bettadpur, Mo Watkins, and Ch Reigber. The Gravity Recovery and Climate Experiment: Mission overview and early results. *Geophysical Research Letters*, 31(9), 2004.

Karl E Taylor. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7):7183–7192, 2001.

Lianke Te Raa, Jeroen Gerrits, and Henk A Dijkstra. Identification of the mechanism of interdecadal variability in the north atlantic ocean. *Journal of Physical Oceanography*, 34(12):2792–2807, 2004.

Lianke A Te Raa and Henk A Dijkstra. Instability of the thermohaline ocean circulation on interdecadal timescales. *Journal of Physical Oceanography*, 32(1): 138–160, 2002.

Gerald Teschl. *Ordinary differential equations and dynamical systems*, volume 140. American Mathematical Soc., 2012.

Leif N Thomas. Destruction of potential vorticity by winds. *Journal of Physical Oceanography*, 35(12):2457–2466, 2005.

MD Thomas and X Zhai. Eddy-induced variability of the meridional overturning circulation in a model of the North Atlantic. *Geophysical Research Letters*, 40(11): 2742–2747, 2013.

Charles Wyville Thomson. *The Voyage of the "Challenger": The Atlantic; a Preliminary Account of the General Results of the Exploring Voyage of HMS "Challenger" During the Year 1873 and the Early Part of the Year 1876*, volume 1. Macmillan and Company, 1877.

Benjamin Thomson Ct. Rumford. *The Complete Works of Count Rumford: Vol. 3*. American Academy of Arts and Sciences, 1874.

JR Toggweiler and B Samuels. Effect of Drake Passage on the global thermohaline circulation. *Deep Sea Research Part I: Oceanographic Research Papers*, 42(4): 477–500, 1995.

Anne Marie Treguier, Bernard Barnier, Aurélie Albert, Julie Deshayes, Julien leSommer, Camille Lique, Jean-Marc Molines, Thierry Penduff, and Claude Talandier. ORCA12, a global ocean-ice model at $\frac{1}{12}^{\circ}$: successes, shortcomings and their impact on ocean forecasting. In *EGU General Assembly Conference Abstracts*, volume 19, page 17950, 2017.

Kevin E Trenberth, John T Fasullo, and Magdalena A Balmaseda. Earth's energy imbalance. *Journal of Climate*, 27(9):3129–3144, 2014.

Eli Tziperman and Petros J Ioannou. Transient growth and optimal excitation of thermohaline variability. *Journal of Physical Oceanography*, 32(12):3427–3435, 2002.

Eli Tziperman, JR Toggweiler, Kirk Bryan, and Yizhak Feliks. Instability of the thermohaline circulation with respect to mixed boundary conditions: Is it really a problem for realistic models? *Journal of Physical Oceanography*, 24(2):217–232, 1994.

Eli Tziperman, Laure Zanna, and Cecile Penland. Nonnormal thermohaline circulation dynamics in a coupled ocean–atmosphere gcm. *Journal of Physical Oceanography*, 38 (3):588–604, 2008.

George E Uhlenbeck and Leonard S Ornstein. On the theory of the Brownian motion. *Physical Review*, 36(5):823, 1930.

Chet Van Duzer. Introduction to the Carta Marina. In *Martin Waldseemüller's' Carta marina'of 1516*, pages 1–53. Springer, 2020.

Ariane Verdy and Matthew R Mazloff. A data assimilating model for estimating Southern Ocean biogeochemistry. *Journal of Geophysical Research: Oceans*, 122(9): 6968–6988, 2017.

A. Vidard, P. A. Bouttier, and F Vigilant. NEMOTAM: Tangent and adjoint models for the ocean modelling platform NEMO. *Geoscientific Model Development*, 8(4): 1245–1257, 2015. ISSN 19919603. .

Thomas H Vonder Haar and Abraham H Oort. New estimate of annual poleward energy transport by northern hemisphere oceans. *Journal of Physical Oceanography*, 3(2):169–172, 1973.

Faming Wang and Ping Chang. Effect of oceanic advection on the potential predictability of sea surface temperature. *Journal of Climate*, 17(18):3603–3615, 2004.

Gang Wang and Dietmar Dommenget. The leading modes of decadal SST variability in the Southern Ocean in CMIP5 simulations. *Climate Dynamics*, 47(5-6): 1775–1792, 2016.

Bruce A Warren. Insensitivity of subtropical mode water characteristics to meteorological fluctuations. *Deep Sea Research*, 19:1–19, 1972.

Andrew J Weaver and ES Sarachik. Evidence for decadal variability in an ocean general circulation model: An advective mechanism. *Atmosphere-Ocean*, 29(2): 197–231, 1991.

Wilbert Weijer, Wei Cheng, Sybren S Drijfhout, Alexey V Fedorov, Aixue Hu, Laura C Jackson, Wei Liu, Elaine L McDonagh, Jennifer V Mecking, and Jiaxu Zhang. Stability of the Atlantic Meridional Overturning Circulation: A review and synthesis. *Journal of Geophysical Research: Oceans*, 124(8):5336–5375, 2019.

Pierre Welander. An advective model of the ocean thermocline. *Tellus A: Dynamic Meteorology and Oceanography*, 11(3):309–318, 1959.

Pierre Welander. The thermocline problem. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 270(1206): 415–421, 1971.

Paul D Williams, Nicola J Howe, Jonathan M Gregory, Robin S Smith, and Manoj M Joshi. Improved climate simulations through a stochastic parameterization of ocean eddies. *Journal of Climate*, 29(24):8763–8781, 2016.

Michael Winton. The damping effect of bottom topography on internal decadal-scale oscillations of the thermohaline circulation. *Journal of Physical Oceanography*, 27(1): 203–208, 1997.

L V Worthington. The 18° water in the Sargasso Sea. *Deep Sea Research*, 5(2): 297–305, 1959.

L V Worthington. Negative oceanic heat flux as a cause of water-mass formation. *Journal of Physical Oceanography*, 2(3):205–211, 1972.

L V Worthington and W Redwood Wright. North Atlantic Ocean atlas of potential temperature and salinity in the deep water including temperature, salinity and

oxygen profiles from the Erika Dan Cruise of 1962. Technical report, Woods Hole Oceanographic Institution, 1970.

Carl Wunsch. What is the thermohaline circulation? *Science*, 298(5596):1179–1181, 2002.

G Wüst. Schichtung und zirkulation des Atlantischen ozeans: Die Stratosphare des Atlantischen ozeans, Wissenschaftliche Ergebrisse der Deutschen Atlantischen Expedition auf dem forschungs und Vermessungshiff METEOR 1925–1927, Band VI Teil 1, 1935.

G Wüst. Die Kreisläufe der atlantischen Wassermassen, ein neuer Versuch räumlicher Darstellung. *Forschungen und Fortschritte*, 25:285–289, 1949.

Peng Xie and Geoffrey K Vallis. The passive and active nature of ocean heat uptake in idealized climate change experiments. *Climate Dynamics*, 38(3-4):667–684, 2012.

Steven T Zalesak. Fully multidimensional flux-corrected transport algorithms for fluids. *Journal of Computational Physics*, 31(3):335–362, 1979.

Laure Zanna and Eli Tziperman. Nonnormal amplification of the thermohaline circulation. *Journal of Physical Oceanography*, 35(9):1593–1605, 2005.

Laure Zanna and Eli Tziperman. Optimal surface excitation of the thermohaline circulation. *Journal of Physical Oceanography*, 38(8):1820–1830, 2008.

Laure Zanna, Patrick Heimbach, Andrew M Moore, and Eli Tziperman. Optimal excitation of interannual atlantic meridional overturning circulation variability. *Journal of Climate*, 24(2):413–427, 2011.

Laure Zanna, P Heimbach, AM Moore, and Eli Tziperman. Upper-ocean singular vectors of the North Atlantic climate with implications for linear predictability and variability. *Quarterly Journal of the Royal Meteorological Society*, 138(663):500–513, 2012.

Laure Zanna, Samar Khatiwala, Jonathan M Gregory, Jonathan Ison, and Patrick Heimbach. Global reconstruction of historical ocean heat storage and transport. *Proceedings of the National Academy of Sciences*, 116(4):1126–1131, 2019.

Rong Zhang. On the persistence and coherence of subpolar sea surface temperature and salinity anomalies associated with the Atlantic multidecadal variability. *Geophysical Research Letters*, 44(15):7865–7875, 2017.

Rong Zhang, Rowan Sutton, Gokhan Danabasoglu, Young-Oh Kwon, Robert Marsh, Stephen G Yeager, Daniel E Amrhein, and Christopher M Little. A review of the role of the Atlantic meridional overturning circulation in Atlantic multidecadal variability and associated climate impacts. *Reviews of Geophysics*, 57(2):316–375, 2019.

Wei Zhang and Ben Kirtman. Understanding the signal-to-noise paradox with a simple markov model. *Geophysical Research Letters*, 46(22):13308–13317, 2019.

Wei Zhang, Ben Kirtman, Leo Siqueira, Amy Clement, and Junfei Xia. Understanding the signal-to-noise paradox in decadal climate predictability from cmip5 and an eddying global coupled model. *Climate Dynamics*, pages 1–19, 2021.

Jian Zhao and William Johns. Wind-forced interannual variability of the Atlantic Meridional Overturning Circulation at 26.5° N. *Journal of Geophysical Research: Oceans*, 119(4):2403–2419, 2014.

Ziqing Zu, Mu Mu, and Henk A Dijkstra. Optimal nonlinear excitation of decadal variability of the North Atlantic thermohaline circulation. *Chinese Journal of Oceanology and Limnology*, 31(6):1368–1374, 2013.

Ziqing Zu, Mu Mu, and Henk A Dijkstra. Optimal initial excitations of decadal modification of the Atlantic meridional overturning circulation under the prescribed heat and freshwater flux boundary conditions. *Journal of Physical Oceanography*, 46 (7):2029–2047, 2016.