

Adjoint Modeling and Observing System Design in the Subpolar North Atlantic



Nora Loose

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2019

UNIVERSITY OF BERGEN



Adjoint Modeling and Observing System Design in the Subpolar North Atlantic

Nora Loose



Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 30.08.2019

© Copyright Nora Loose

The material in this publication is covered by the provisions of the Copyright Act.

Year: 2019

Title: Adjoint Modeling and Observing System Design in the Subpolar North Atlantic

Name: Nora Loose

Print: Skipnes Kommunikasjon / University of Bergen

“It is absolutely necessary, for progress in science, to have uncertainty as a fundamental part of your inner nature. [...] And as you develop more information in the sciences, it is not that you are finding out the truth, but that you are finding out that this or that is more or less likely. That is, if we investigate further, we find that the statements of science are not of what is true and what is not true, but statements of what is known to different degrees of certainty. [...] Every one of the concepts of science is on a scale graduated somewhere between, but at neither end of, absolute falsity or absolute truth.”

— Richard Feynman, *Science vs. Religion and Why Uncertainty Is Central to Morality*

Abstract

The near-surface ocean currents of the subpolar North Atlantic transport large amounts of heat from the subtropics to higher latitudes, affecting Arctic sea ice extent, the melting of the Greenland Ice Sheet, and the climate in western Europe and North America. Moreover, deep water formation in the subpolar North Atlantic actively shapes the Atlantic meridional overturning circulation, which connects the surface with the deep ocean and the northern with the southern hemisphere. The recently acquired data from the OSNAP (Overturning in the Subpolar North Atlantic Program) mooring array challenges our understanding of the processes that govern circulation and deep water formation in the subpolar North Atlantic. However, only long-term and sustained ocean observations can provide the much-needed benchmark to evaluate climate model simulations, to advance our understanding of key mechanisms, and to predict the role of the North Atlantic in future climate changes and anthropogenic carbon uptake. Unfortunately, most observational efforts rely on short-term funding periods.

Given the cost of deploying and maintaining ocean observing systems, these systems have to be designed carefully. Key questions are: What information is contained in already existing observation networks? What do existing networks, such as the OSNAP array, tell us about hydrographic and circulation quantities in remote oceanic regions with few observations? In this thesis, a novel approach to ocean observing system design is explored that is able to address these questions. The approach makes use of adjoint modeling and Hessian-based Uncertainty Quantification (UQ) within a global oceanographic inverse problem.

Adjoint-derived sensitivities reveal that the eastern boundary of the North Atlantic and the coasts of Iceland and Greenland are important pathways for communicating wind-driven pressure anomalies around the entire subpolar North Atlantic and the Nordic Seas. Consequently, the OSNAP observing array shares many dynamical pathways and mechanisms with oceanic quantities that are remote from the array. The OSNAP array has therefore potential to inform these unobserved - or unobservable - quantities: for instance, ocean heat content in the Nordic Seas or close to Greenland's margins. In this thesis, this potential is quantified within the state-of-the-art ECCO (Estimating the Circulation and Climate of the Ocean) state estimation framework, by combining physical relationships in the model with prior information and data uncertainties.

The effectiveness of an observing system is determined by how well it captures climate-relevant signals and important dynamical adjustment mechanisms. A second important factor, however, is how strongly the monitored signals are masked by noise. All factors combined, heat transport measurements across the OSNAP-West transect, extending from Labrador to South Greenland, impose an overall much stronger constraint on the ECCO state estimate than heat transport measurements across the OSNAP-East transect, extending from South Greenland to Scotland. This is largely explained by the

fact that climate signals detected by OSNAP-West are less noisy compared to climate signals detected by OSNAP-East. As a result, transport and hydrographic quantities - even in the Nordic Seas - are constrained more efficiently by OSNAP-West than OSNAP-East observations, contrary to recent findings. This suggests that OSNAP-West is important for informing remote climate signals.

This thesis explores the physical mechanisms that link the subpolar North Atlantic and the Nordic Seas, translates the mathematical concepts that underlie Hessian-based UQ to dynamical concepts, and discusses benefits, shortcomings, and future challenges for designing an effective, long-term Atlantic observing system by means of UQ within ocean state estimation.

Acknowledgements

First of all, I thank my supervisors, Kerim H. Nisancioglu and Patrick Heimbach, for all their support. I am grateful to Kerim for giving me the freedom to explore my own research interests, while providing useful mentorship, particularly in the writing and presentation of this work. Kerim's everlasting enthusiasm and encouragement has helped me to never give up. I would also like to thank Kerim for field work opportunities and to support many fun summer schools, conferences, and workshops. I am grateful to Patrick for many inspiring discussions, for valuable input whenever needed, and for generously supporting me to finish this thesis from UT Austin. I also thank Patrick for hosting earlier stays at MIT and UT Austin and for opening the door to networking events.

I owe huge thanks to Helen Pillar and Tim Smith for tremendously helpful discussions and technical support. I am indebted to Helen for helping me to get started with the MITgcm and the adjoint and for giving me invaluable and thorough comments on most of this thesis. I would also like to thank Helen for excellent discussions and ideas on how to make computational tools more accessible and exciting to physical oceanographers. I hope we are on the right track! I am grateful to Tim for sharing his code and experience with the MITgcm and for all his technical assistance. I also thank Tim for giving me feedback on many chapters of this thesis, for in-depth discussions on computing and science, and for all his support outside of this thesis.

I am grateful to Marius Årthun for discussing and proofreading part of this thesis. I owe thanks to Dan Amrhein and Alex Kalmikov for sharing code and to Joel Pedro, An Nguyen, Chuncheng Guo and Mari F. Jensen for sharing data. Thanks are also due to Gael Forget, Ou Wang and Jean-Michel Campin for answering ECCO-related questions. I would like to thank Dan Goldberg, Yavor Kostov, Jake Gebbie, Dan Jones, Helen Johnson, Peter Huybers, Carl Wunsch, and Bruce Cornuelle for inspiring discussions at various stages of my PhD.

I feel privileged to have been part of two excellent research groups in Bergen and Austin, as well as the Ice2Ice community. These groups have provided a wonderful research environment and introduced me to many exciting applications in climate, physical oceanography, paleoceanography, glaciology, and computational science. Finally, I would like to thank my PhD colleagues, office mates, friends and family. Special thanks are due to my friends in Bergen for many skiing, climbing, and running trips in the mountains that kept me sane during this PhD.

Contents

Abstract	iii
Acknowledgements	v
1 Introduction	1
1.1 The subpolar North Atlantic and Nordic Seas	1
1.1.1 Near-surface circulation	1
1.1.2 Shaping the AMOC	3
1.1.3 Interaction with Greenland’s marine-terminating glaciers	4
1.1.4 A player in paleoclimate shifts	4
1.1.5 The OSNAP observing system	7
1.1.6 Open questions and observational needs	8
1.2 Adjoint models	9
1.2.1 State estimation	13
1.2.2 Sensitivity analysis	14
1.3 Observing system design	16
1.3.1 Frameworks	16
1.3.2 OS S Es	17
1.3.3 Adjoint-based methods	19
1.3.4 OSSEs vs. adjoint-based methods	26
1.4 Thesis objectives and outline	27
2 Drivers of Upper-Ocean Heat Content Anomalies in the Nordic Seas	31
2.1 Introduction	31
2.2 Experimental setup	33
2.2.1 Model description and base state	33
2.2.2 Quantities of interest	35
2.2.3 Adjoint model and sensitivities	37
2.3 Identifying adjustment mechanisms and pathways	39
2.3.1 Sensitivity to buoyancy forcing	39
2.3.2 Sensitivity to surface momentum fluxes	49
2.4 Relative importance of forcings and regions of origin	55
2.4.1 Relative importance of forcings	57
2.4.2 Seasonality	59
2.4.3 Relative importance of regions of atmospheric origins	59
2.5 Discussion	60
2.6 Conclusions	67

3	Dynamics-based Ocean Observing System Design	69
3.1	Introduction	69
3.2	Uncertainty Quantification in inverse problems	73
3.3	Computation of data-informed directions and curvatures	77
3.3.1	A single observation	78
3.3.2	Multiple observations	79
3.3.3	Testing various prior and noise matrices	81
3.4	Future observations	82
3.5	Proxy potential	83
3.5.1	Relative uncertainty reduction	84
3.5.2	Hypothetical proxy potential	86
3.5.3	Noise masking	87
3.6	Discussion	89
3.6.1	Key insights	89
3.6.2	Limitations	94
3.7	Conclusions	98
4	Proxy Potential of the OSNAP Array	101
4.1	Introduction	101
4.2	Inverse modeling framework	104
4.2.1	GCM	104
4.2.2	The OSNAP array and data uncertainties	104
4.2.3	Quantities of interest	109
4.2.4	Controls and prior uncertainties	111
4.2.5	Adjoint models	116
4.3	Results	118
4.3.1	Dynamical adjustment mechanisms for OSNAP	118
4.3.2	Dynamical constraints of OSNAP	124
4.3.3	Dynamical adjustment mechanisms for the unobserved QoIs	133
4.3.4	Proxy potential	137
4.4	Discussion	153
4.5	Conclusions	163
5	Testing the Linearity Assumption in the Subpolar North Atlantic	165
5.1	Introduction	165
5.2	Linearity checks	166
5.2.1	Methodology	166
5.2.2	Local heat flux perturbations	167
5.2.3	Buoyancy flux perturbations in the NAC and Gulf Stream	170
5.2.4	Wind stress perturbations along the intergyre boundary	172
5.2.5	Perturbations in the subpolar gyre and Arctic Ocean	174
5.3	Discussion	177
5.3.1	Non-linearity	178
5.3.2	Inexactness of the adjoint	178
5.3.3	Implications	179
5.4	Conclusions	181

6	Summary and Discussion	183
6.1	Mechanisms in the subpolar North Atlantic and Nordic Seas	183
6.1.1	Thermally vs. wind-driven mechanisms	183
6.1.2	Sensitivity of ocean heat content vs. ocean transports	187
6.2	Dynamics-based assessment of observing systems	188
6.2.1	Key insights	189
6.2.2	Full-fledged frameworks	195
6.2.3	Limitations	198
6.3	Broader implications	199
6.3.1	Towards the design of an optimized Atlantic observing system	200
6.3.2	Are state estimation frameworks ready for formal observing system design?	201
7	Conclusions	211
8	Future Work and Applications	215
8.1	Subsurface ocean temperature at Greenland's margins	215
8.1.1	Impact of warm ocean waters on Greenland melt	215
8.1.2	Dynamical proxy potential of remote ocean observing arrays	216
8.1.3	Outlook	221
8.2	Dynamical proxy potential of paleoceanographic observations	221
8.2.1	Combining paleo proxy data with dynamical models	221
8.2.2	Mechanisms contributing to D-O temperature variability	222
8.2.3	Constraints of proxy data on past ocean circulation	230
A	Nordic Seas: Sensitivity vs. Response to Atmospheric Forcing	237
A.1	Regions of high sensitivity	237
A.2	Typical forcing anomalies	238
A.3	High-sensitivity regions vs. high-impact regions	241
A.4	Seasonality	242
A.4.1	Seasonality of unweighted sensitivities	242
A.4.2	Seasonality of forcing anomalies	242
A.4.3	Seasonality of weighted sensitivities	243
B	Uncertainty Quantification	245
B.1	Bayesian formulation of inverse problem	245
B.2	Gaussian approximation of the posterior	246
B.3	Forward uncertainty propagation	246
B.4	Linearized vs. full Hessians	247
C	Prior information in ECCOv4r3	251
C.1	Time-variable atmospheric controls	251
C.2	Smoothing	253
D	Spatial Scales of Atmospheric Adjustment in Ocean State Estimation	259
D.1	Introduction	259
D.2	Spectral transformations	260
D.2.1	Spherical harmonics	261

D.2.2	Gaussian grids	263
D.3	Spectral analysis of ERA-Interim near-surface fields	264
D.3.1	Data and grid	264
D.3.2	Methodology	265
D.3.3	Results	265
D.4	Discussion	271
D.4.1	Spatial scales of ocean surface forcing	271
D.4.2	Regularization in ocean state estimation	273
D.4.3	Control space reduction	274
D.5	Conclusions	276

Chapter 1

Introduction

This thesis is an effort to (i) contribute to our understanding of the physical mechanisms in the subpolar North Atlantic and Nordic Seas, and (ii) explore a novel approach to dynamics-based observing system design for the subpolar North Atlantic. In Section 1.1, I provide a brief overview of the circulation and climate impacts of the subpolar North Atlantic and Nordic Seas, a recently launched subpolar North Atlantic observing system, and open science questions. The following two sections introduce the methods that are used in this thesis: In Section 1.2, I describe the concept and applications of adjoint models, and in Section 1.3, I review multiple techniques for evaluating and designing observing systems. Finally, I outline the main thesis objectives and goals in Section 1.4.

1.1 The subpolar North Atlantic and Nordic Seas

1.1.1 Near-surface circulation

The subpolar North Atlantic comprises roughly the cyclonic ocean gyre north of about 50°N (Fig. 1.1). The major current which brings warm and salty waters from the subtropics to the subpolar North Atlantic is the North Atlantic Current (NAC), which crosses the Atlantic near the surface as the north-eastward extension of the Gulf Stream. The NAC then splits into several branches. The north-eastern branches flow through Rockall Trough and the Iceland Basin across the Iceland-Scotland ridge, transporting warm and salty waters into the Nordic Seas (the Greenland, Iceland, and Norwegian Seas) [Hansen and Østerhus, 2000]. The two branches of the Norwegian Atlantic Current carry the warm and salty Atlantic waters further toward the Arctic Ocean [Orvik and Niiler, 2002]. Another branch of the NAC recirculates in the Iceland Basin, and follows the cyclonic boundary current of the subpolar gyre: The recirculated branch flows southward along the eastern flank of the Reykjanes Ridge, crosses the ridge, and continues northward as the Irminger Current (IC) along the western flank of the ridge. A portion of the IC flows through Denmark Strait into the Nordic Seas, while another portion recirculates and flows around the southern tip of Greenland into the Labrador Sea [Holliday et al., 2009]. Here, the current circulates around the margins of the Labrador Sea, while another portion flows northward through Davis Strait [Cuny et al., 2002].

The various NAC branches transport large amounts of heat from the subtropics to higher latitudes. While being carried cyclonically around the subpolar North Atlantic, a

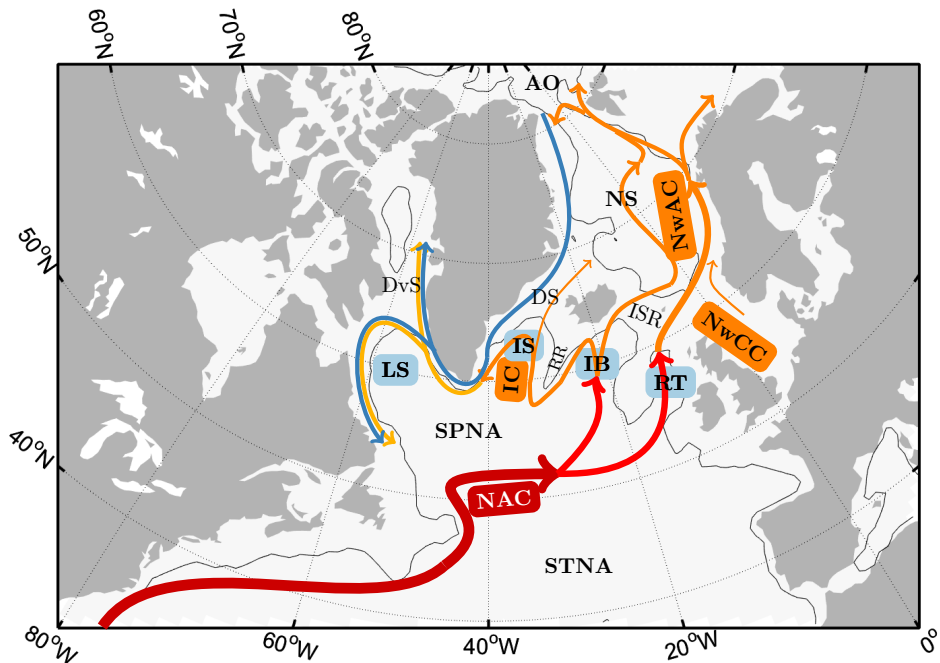


Figure 1.1: Schematic of the major near-surface currents in the subpolar North Atlantic (SPNA) and Nordic Seas (NS). Warm Atlantic-origin water pathways are shown as red to yellow arrows, cold Arctic-origin water pathways as blue arrows. The major basins of the SPNA carry light blue labels. The thin contour lines mark the isobath drawn at 1500 m. NAC = North Atlantic Current; IC = Irminger Current; NwAC = Norwegian Atlantic Current; NwCC = Norwegian Coastal Current; SPNA = subpolar North Atlantic; STNA = subtropical North Atlantic; RT = Rockall Trough; IB = Iceland Basin; IS = Irminger Sea; LS = Labrador Sea; NS = Nordic Seas; AO = Arctic Ocean; ISR = Iceland-Scotland Ridge; RR = Reykjanes Ridge; DS = Denmark Strait; DvS = Davis Strait.

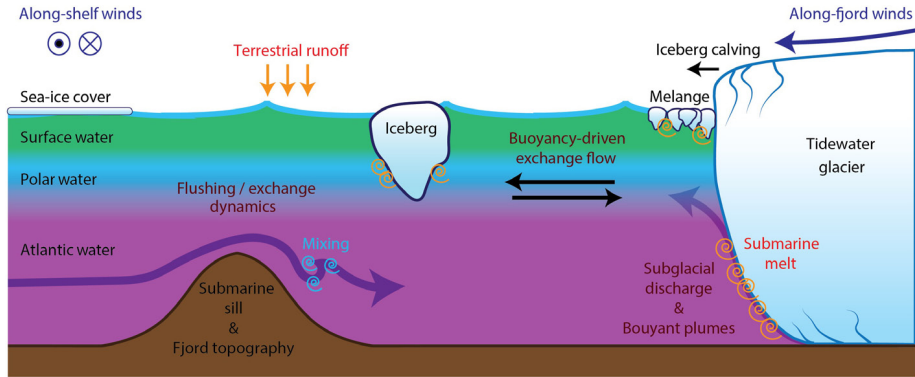


Figure 1.2: Schematic of a glacial fjord in Greenland. Warm Atlantic waters (purple shading) that cross the continental shelf and submarine sill (brown rise in bottom topography) reach the glacier at depth and drive glacier submarine melting (red swirls). Figure from [Straneo et al. \[2019\]](#).

portion of this heat is released back to the atmosphere, thereby having an effect on many climate and weather phenomena in western Europe and North America [e.g., [Sutton and Hodson, 2005](#); [Sutton and Dong, 2012](#)]. Moreover, variability in the poleward progression of ocean heat from the subpolar North Atlantic towards the Arctic Ocean has been linked to Arctic sea-ice extent [[Carmack et al., 2015](#); [Zhang, 2015](#); [Polyakov et al., 2017](#)].

1.1.2 Shaping the AMOC

The northward flow of warm, salty waters in the near-surface layers of the Atlantic is often described as the upper limb of the Atlantic meridional overturning circulation (AMOC) [e.g., [Buckley and Marshall, 2016](#)]. As the warm waters are carried from the Gulf Stream cyclonically around the subpolar North Atlantic by the various NAC branches, they gradually cool along their path [[McCartney and Talley, 1982](#); [Brambilla and Talley, 2008](#)]. The gradual transformation causes these surface waters to become denser - and eventually sink to great depth in the Nordic Seas and Labrador Sea, where heat loss to the atmosphere destabilizes the water column in winter [e.g., [Marshall and Schott, 1999](#)]. These cold, dense waters that are formed in the Nordic Seas and Labrador Sea constitute the components of North Atlantic Deep Water, which are exported southward at depth and feed the lower limb of the AMOC. The subpolar North Atlantic is therefore a region where the strength and structure of the AMOC is actively shaped. The AMOC connects the Northern with the Southern hemisphere, as well as the surface with the deep ocean. It is therefore a key component of the global climate system, for instance, through its cross-equatorial ocean heat and freshwater transport and its role in the global carbon cycle [[Buckley and Marshall, 2016](#); [Khatiwala et al., 2013](#)].

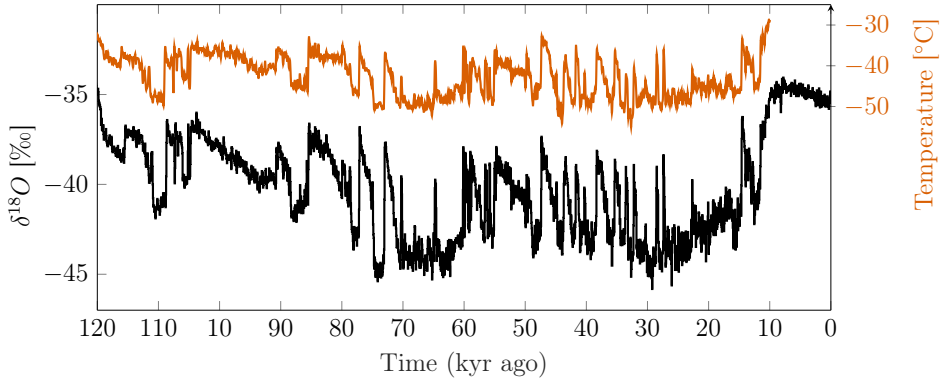


Figure 1.3: Oxygen isotope measurements (black, [North Greenland Ice Core Project members, 2004]) and temperature reconstruction (orange, [Kindler et al., 2014]) from the Greenland ice core NGRIP.

1.1.3 Interaction with Greenland’s marine-terminating glaciers

The cyclonic boundary currents of the subpolar North Atlantic bring warm and salty Atlantic waters around the continental slopes of Greenland. Closer to the coast, cold and fresh waters of Arctic origin flow around Greenland’s shallow (200-300 m deep) continental shelves, partially buffering Greenland’s coast from the warm, Atlantic waters (Figs. 1.1 and 1.2). In the mid-1990s, Greenland’s marine-terminating glaciers started to retreat and accelerate, at a time when the subpolar North Atlantic experienced a rapid warming [Straneo and Heimbach, 2013]. Observational studies show evidence that, concurrent with the warming of the subpolar North Atlantic, the layer of Atlantic water thickened and the shelf waters warmed [e.g., Holland et al., 2008]. Coming in contact with the glaciers, warm ocean waters trigger increased submarine melting (see Fig. 1.2). It is therefore suggested that the subpolar North Atlantic has an important impact on the melting and acceleration of Greenland’s glaciers [Holland et al., 2008; Straneo et al., 2010, 2012; Vieli and Nick, 2011; Joughin et al., 2012; Straneo and Heimbach, 2013; Rainsley et al., 2018]. There are two main implications of increased ice loss from Greenland for the climate system. First, sea-level change [e.g., Stammer, 2008]; and second, ocean surface freshening, which may impact the AMOC, when reaching the deep water formation sites [e.g., Böning et al., 2016; Luo et al., 2016; Yang et al., 2016].

1.1.4 A player in paleoclimate shifts

During the last glacial period, North Atlantic climate was characterized by large millennial-scale variability and a number of abrupt climate shifts. Fig. 1.3 shows oxygen isotope measurements and temperature reconstructions from the Greenland ice core NGRIP. The record shows about 25 large positive spikes, reflecting sudden warmings by 5 – 10 °C within at most a few decades. These sudden changes from cold Greenland stadial conditions to warmer Greenland interstadial conditions are referred to as the Dansgaard-Oeschger (D-O) events [Dansgaard et al., 1993; North Greenland Ice Core Project members, 2004]. The schematics in Fig. 1.4 show land and sea-ice conditions in

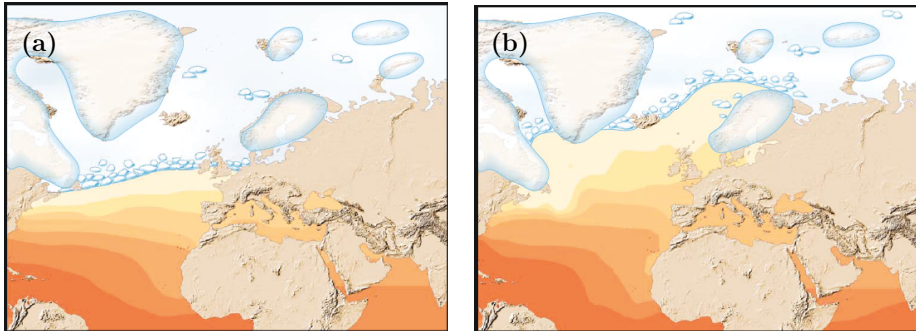


Figure 1.4: Schematics showing land and sea-ice conditions in the North Atlantic and the Nordic Seas during (a) cold Greenland stadials and (b) warmer Greenland interstadials, as suggested by [Dokken et al. \[2013\]](#). Figure from [Dokken et al. \[2013\]](#).

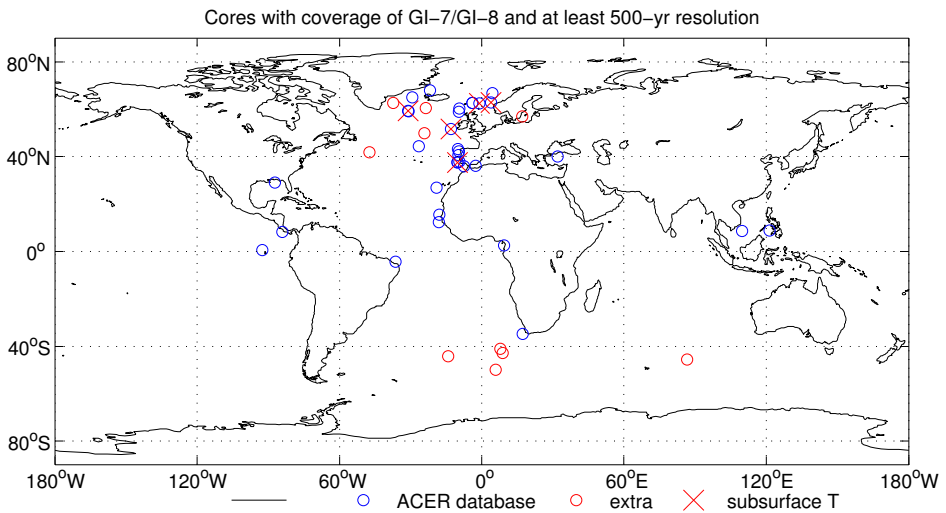


Figure 1.5: Globally available SST proxy data from marine sediment cores that cover GI-7/GI-8 and have at least 500-yr resolution. Figure courtesy Joel Pedro. Cf. with data coverage of modern observations, e.g., Argo floats in [Fig. 1.10](#).

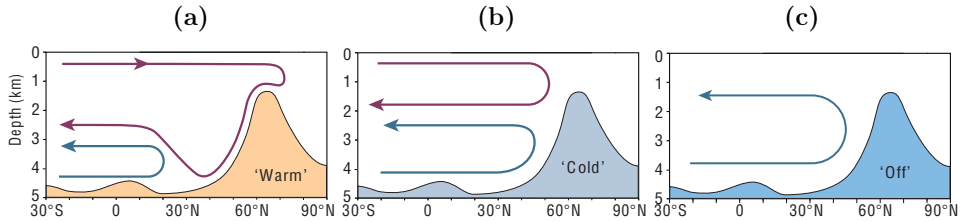


Figure 1.6: Schematic of three suggested modes of ocean circulation that prevailed during different times of the last glacial period. Shown is a section along the Atlantic. The rise in bottom topography symbolizes the shallow sill between Greenland and Scotland. North Atlantic overturning is shown by the red line, Antarctic bottom water by the blue line. Figure and (adapted) caption from [Rahmstorf \[2002\]](#).

the North Atlantic and the Nordic Seas during Greenland stadials and interstadials, as suggested by [Dokken et al. \[2013\]](#).

The D-O climate fluctuations are not only visible in Greenland ice core records, but have also been discovered in numerous marine records from the North Atlantic, the tropics and the Southern Hemisphere (see [Voelker, 2002](#), for a comprehensive review). Reconstructions of past ocean temperatures can be inferred from proxies such as planktic (near-surface dwelling) foraminifera assemblages, from organic geochemistry using alkenone unsaturation indices, from Mg/Ca-concentrations in corals or calcite shells, or from oxygen isotopes. The relationship of proxy observations to physical variables is often tenuous, leading to large uncertainties, and the observations are very sparse in space and time. For instance, Fig. 1.5 shows the sparse spatial coverage of marine sediment cores which provide near-surface proxy data that cover the period between Greenland interstadials 7 and 8 (ca. 35 - 38 kyr ago) with at least 500-year resolution. In contrast, modern ocean observations have much smaller uncertainties and are much more dense in space and time. The Argo floats in Fig. 1.10 comprise only a fraction of the modern ocean observation. Each Argo float samples the temperature and salinity profile of the full upper 2000 m of the water column, while a marine sediment core provides proxy information about the conditions near the surface or the ocean bottom, while the exact depth depends on the species' habitat and is often difficult to estimate.

The mechanism behind the millennial-scale climate variability and the abrupt D-O events is still unclear. The classical view is that the D-O events are caused by changes in ocean circulation [[Broecker et al., 1985](#)], as shown in the schematics in Fig. 1.6. Warm interstadial periods are linked to a vigorous AMOC (Fig. 1.6(a)), which effectively transports heat to the high-latitude North Atlantic. The addition of freshwater into the North Atlantic through the melting of large continental ice sheets is believed to have reduced North Atlantic Deep Water formation, by inducing stratification and inhibiting convection in the North Atlantic. The subsequent slowdown of the AMOC leads to a cooling in the Northern Hemisphere (Fig. 1.6(b)) and cold stadial conditions. In the 'off mode' (Fig. 1.6(c)), the AMOC is not only slowed down but shut down, which leads to

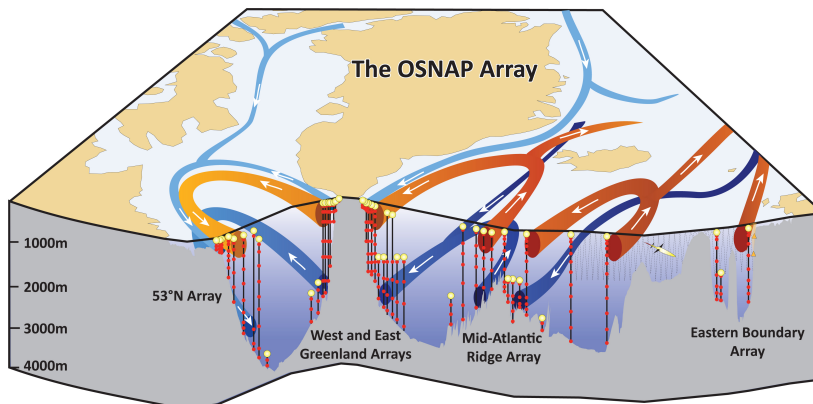


Figure 1.7: The observing array of the Overturning in the Subpolar North Atlantic Program (OSNAP), deployed in the summer 2014 along a section from Labrador to Greenland to Scotland. The OSNAP array is a composition of moorings and gliders. Figure from [Lozier et al. \[2017\]](#). ©American Meteorological Society. Used with permission.

particularly dramatic cooling, referred to as Heinrich events. Another hypothesis is that the D-O events are linked to rapid changes in sea-ice cover in the Nordic Seas and North Atlantic [[Gildor and Tziperman, 2003](#); [Li et al., 2005](#); [Dokken et al., 2013](#); [Petersen et al., 2013](#); [Sadatzki et al., 2019](#)]. Changes in sea-ice influence the atmospheric temperature, by altering heat release from the ocean. In a recent study, [Sadatzki et al. \[2019\]](#) reconstruct sea-ice variability during several D-O cycles. The authors suggest that the changes in sea-ice shaped regime shifts between surface stratification and deep convection in the Nordic Seas and played a persistent role in triggering the abrupt stadial-interstadial climate transitions. A third hypothesis is that the D-O variability is caused by shifts of the atmospheric jet, owing to interactions with the Laurentide and Fennoscandian ice sheets [e.g., [Jackson, 2000](#); [Roe and Lindzen, 2001](#)]. Using a simplified ice sheet-stationary wave model, [Roe and Lindzen \[2001\]](#) show that the mean structure of the westerly wind system is quite different from its modern value, when massive ice sheets are encountered. More than one equilibrium is possible, with the prevailing winds shifting north and south of the Laurentide and Fennoscandian ice sheets.

1.1.5 The OSNAP observing system

In the summer of 2014, OSNAP (the Overturning in the Subpolar North Atlantic Program) deployed an observing system to better understand the subpolar AMOC and gyre circulation [www.o-snap.org; [Lozier et al., 2017](#)]. Being composed of densely spaced moorings and gliders, the OSNAP array measures temperature, salinity, and velocity along a coast-to-coast section in the subpolar North Atlantic (Fig. 1.7). The OSNAP array consists of two sections: OSNAP-West, extending from Labrador to the southwestern tip of Greenland, and OSNAP-East, extending from the southeastern tip of Greenland to Scotland. The moorings are placed in the boundary currents of the four major basins of

the subpolar North Atlantic - the Labrador Sea, Irminger Sea, Iceland Basin, and Rockall Trough (see Fig. 1.1) - and the gliders take measurements in the regions between.

The first 21 months of OSNAP data suggest that the Labrador Sea plays only a minor role in the strength and variability of the overturning and heat transport in the subpolar basin [Lozier et al., 2019]. Instead, over the course of the OSNAP record, subpolar overturning and heat transport are dominated by water mass transformation north of the OSNAP-East leg, between Greenland and Europe. The dominance of deep water formation in the Irminger and Iceland basins, rather than in the Labrador Sea, departs from the prevailing view that changes in Labrador Sea deep water formation are the key process for AMOC variability. This classical view is attributable to the fact that in most ocean-ice models, the multi-annual to decadal variability of the AMOC is set by changes in Labrador Sea deep water formation [Danabasoglu et al., 2016; Xu et al., 2018]. While the 21-month long OSNAP record presented by Lozier et al. [2019] does not resolve these timescales, it challenges our understanding of the processes that govern circulation and deep water formation in the subpolar North Atlantic.

1.1.6 Open questions and observational needs

There is an ongoing debate on the role of the subpolar North Atlantic in observed variability and predictability of Atlantic climate. Some of the debated questions as well as observational needs are listed here:

1. **Causes of observed variability.** What drives variability in Atlantic ocean circulation, e.g., overturning, heat and freshwater transports, upper-ocean heat content, and SST, on monthly to millennial timescales? How much of the observed variability can be attributed to external vs. internal forcing? How much of the external variability is controlled by wind vs. buoyancy forcing? What is the relative importance of local vs. remote forcing? What is the role of North Atlantic deep water formation for generating variability in Atlantic climate?
2. **Predictability.** Which are the mechanisms that govern Atlantic SST predictability on seasonal to decadal timescales? Are processes in the subpolar North Atlantic, such as the formation of North Atlantic deep water, a primary source of decadal climate predictability? Can the monitoring of certain regions or oceanic processes provide an “early warning system” for future changes in Atlantic climate? If so, what are the predictability horizons, and what metrics are most predictable?
3. **Model spread and observational needs.** Due to the complex nature of the oceanic circulation and processes in the subpolar North Atlantic [e.g., Holliday et al., 2018], ocean models show a wide spread in the time-mean and time-varying simulated ocean state in the subpolar North Atlantic [Menary et al., 2015; Marzocchi et al., 2015] and in basin-wide AMOC estimates [Danabasoglu et al., 2014, 2016]. Moreover, climate models from the Fifth Coupled Model Intercomparison Project exhibit varying capability in simulating the stratification [Sgubin et al., 2017] and deep convection [Heuzé, 2017] in the subpolar North Atlantic. However, understanding and predicting the role of the ocean in future climate and anthropogenic carbon uptake/storage in the North Atlantic, will require climate models

that realistically simulate the mechanisms setting the strength, the variability, and the trends of Atlantic ocean circulation. In-situ ocean observations, such as the recently acquired OSNAP data, are beginning to provide the foundation for assessing model performance. However, most observational efforts, including the OSNAP array, rely on short-term funding periods, while only long-term continuous time series can provide the much-needed benchmark to evaluate and validate climate model simulations [e.g., [National Academies of Sciences, Engineering, and Medicine, 2017](#)].

Like in the modern ocean, paleoceanography depends crucially on observations. However, data types and coverage are (and always will be) very limited, and proxy data entails large uncertainties. For paleoclimate reconstructions, model-data synthesis seems promising - if not necessary - since it permits using dynamical principles to constrain interpretations of proxy observations, to quantify the inferential power of proxy data, and to compute observable and unobservable climate aspects. Model-data synthesis for paleoceanographic applications is a research topic that is still in its infancy - with many challenges yet to be tackled.

4. **Design of an effective Atlantic observing system.** Given the cost of deploying and maintaining ocean observing systems, these systems have to be designed carefully. Key questions are: What is the (optimal) instrument configuration for a long-term sustained Atlantic observing system, which is both cost-efficient and capable to monitor Atlantic key processes and ocean variability? Which information is contained in already existing observation networks, e.g., the ones part of the Global Ocean Observing System (GOOS, <http://www.goosocan.org/>)? What do such existing networks tell us about ocean circulation in remote regions away from the instruments? Is there any redundancy in existing configurations?

Some of these questions and needs will be addressed in this thesis. The thesis goals and objectives are outlined in Section 1.4. To achieve these goals, I will use a range of computational tools that are based on the adjoint of an ocean general circulation model. These tools comprise comprehensive sensitivity analysis, and dynamics-based observing system design within an ocean inverse modeling framework. In the next two sections, I will explain these tools further.

1.2 Adjoint models

Ocean General Circulation Models (GCMs) are numerical tools to simulate the non-linear dynamics of the ocean, based on a spatio-temporal discretization of the conservation equations for mass/volume, momentum, heat and salt. This is illustrated within box B of Fig. 1.8. For every model grid cell, the GCM computes the prognostic (or ocean state) variables, i.e., temperature (T), salinity (S), and the three velocity components (u,v,w). Here, the grid cells are represented by the dashed boxes, and labelled from 1 to K. Starting at initial time t_0 , the GCM numerically integrates the discretized equations of fluid motion, and simulates the ocean state variables in each grid cell at the next time t_1 . Iterating this process many times (delineated by the arrows from left to right within box B), one obtains the simulated ocean state at final time t_T . In the course of these numerical computations, the GCM requires input variables x_1, \dots, x_N , shown in

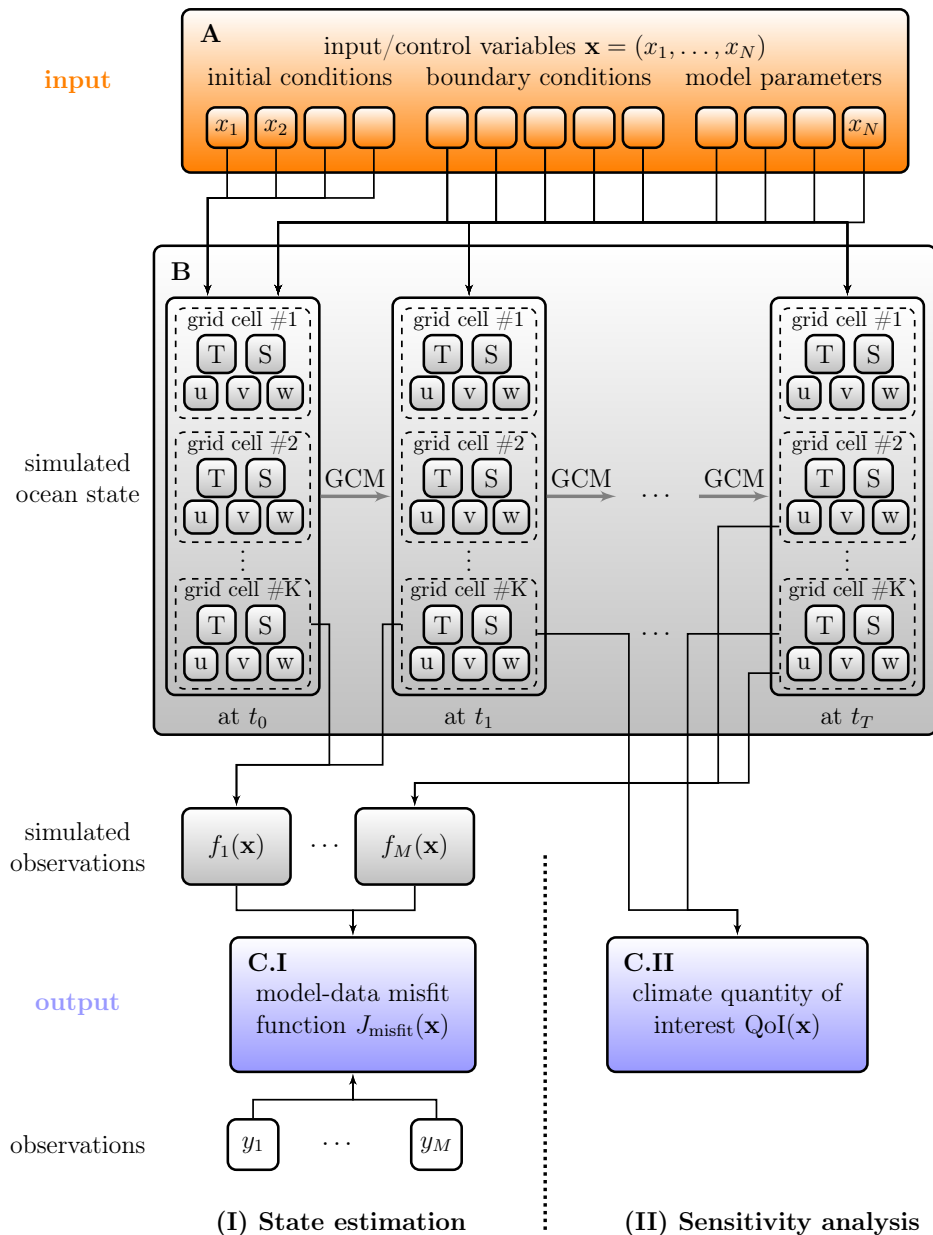


Figure 1.8: Workflow for a GCM (boxes A and B), and for the two main adjoint applications (I) State estimation (left lower branch) and (II) Sensitivity analysis (right lower branch). Box A contains the model input variables, box B the prognostic variables, and boxes C.I and C.II the scalar output variable for the respective adjoint application. The output variable usually depends in a complex way on the input and prognostic variables, indicated through the network diagram.

box A in Fig. 1.8. These input variables consist of (a) initial conditions of the ocean state, e.g., initial T, S, u, v, and w at every model grid cell, (b) (time-evolving) boundary conditions, e.g., atmospheric temperature, wind, and precipitation at every surface grid cell and at every time t_0, t_1, \dots, t_T , and (c) uncertain model parameters, e.g., parameters used for unresolved subgrid-scale processes associated with advection and mixing by eddies. The number of input variables (N) is enormous: depending on the model resolution and length of the GCM simulation, it is on the order of millions to billions (or even trillions).

Many questions investigated by oceanographers and climate dynamicists concern sensitivity of the climate system, or certain characteristics of it. The adjoint of a GCM - also called the associated ‘adjoint model’ - is a powerful tool to compute linearized sensitivities of a scalar GCM output variable with respect to all GCM input variables x_1, \dots, x_N . In ocean and climate research, adjoint models are mainly used for two types of applications, represented by the two branches (I) and (II) in the lower half of Fig. 1.8:

- (I) **State estimation** (left branch in Fig. 1.8). The GCM output variable is chosen to be a *scalar-valued model-data misfit function* $J_{\text{misfit}}(\mathbf{x})$ (box C.I), or a regularized version of it. $J_{\text{misfit}}(\mathbf{x})$ sums up all the (weighted) misfits of available ocean observations y_1, \dots, y_M and the observations’ counterparts $f_1(\mathbf{x}), \dots, f_M(\mathbf{x})$ simulated by the model. The simulated observations $f_1(\mathbf{x}), \dots, f_M(\mathbf{x})$ are derived from the simulated ocean state (box B). For instance, $f_1(\mathbf{x})$ might be the temperature in a certain grid cell averaged over multiple time steps, and $f_M(\mathbf{x})$ might be horizontal velocity averaged over multiple grid cells (see Fig. 1.8). Obtaining the gradient $\nabla_{\mathbf{x}} J_{\text{misfit}} = \left(\frac{\partial J_{\text{misfit}}}{\partial x_1}, \dots, \frac{\partial J_{\text{misfit}}}{\partial x_N} \right)^T$, i.e., the linearized sensitivities of $J_{\text{misfit}}(\mathbf{x})$ to all input variables x_1, \dots, x_N , enables gradient-based algorithms to find the optimal set of input variables which minimizes the model-data misfit function $J_{\text{misfit}}(\mathbf{x})$ [Wunsch and Heimbach, 2007]. The associated simulated ocean state is a data-constrained dynamically consistent GCM simulation. More details are given in Section 1.2.1.
- (II) **Sensitivity analysis** (right branch in Fig. 1.8). The GCM output variable is chosen to be a *scalar-valued climate quantity of interest* $\text{QoI}(\mathbf{x})$ (box C.II), e.g., heat content of a specific ocean basin in winter 2018/2019, or any other diagnostic (or prognostic) scalar variable that can be derived from the time-evolving simulated ocean state (box B). The linearized sensitivities of $\text{QoI}(\mathbf{x})$ to all input variables, i.e., $\frac{\partial(\text{QoI})}{\partial x_1}, \dots, \frac{\partial(\text{QoI})}{\partial x_N}$, reveal all possible dynamical mechanisms and pathways via which small-amplitude perturbations in the input variables, e.g., surface forcing over the globe for all forcing lead times, impact the QoI. Thereby, physical teleconnections and causal chains in the model equations can be examined [e.g., Marotzke et al., 1999; Heimbach et al., 2011], and rigorous dynamics-based attribution studies can be performed [Pillar et al., 2016; Smith and Heimbach, 2019]. More details are explained in Section 1.2.2.

From a computational point of view, these two applications have the following in common: One is interested in the sensitivity of *one* scalar-valued GCM *output variable* (box C.I or C.II in Fig. 1.8) to *many* input variables x_1, \dots, x_N (the many small orange boxes within box A). This is exactly the framework in which an adjoint model is

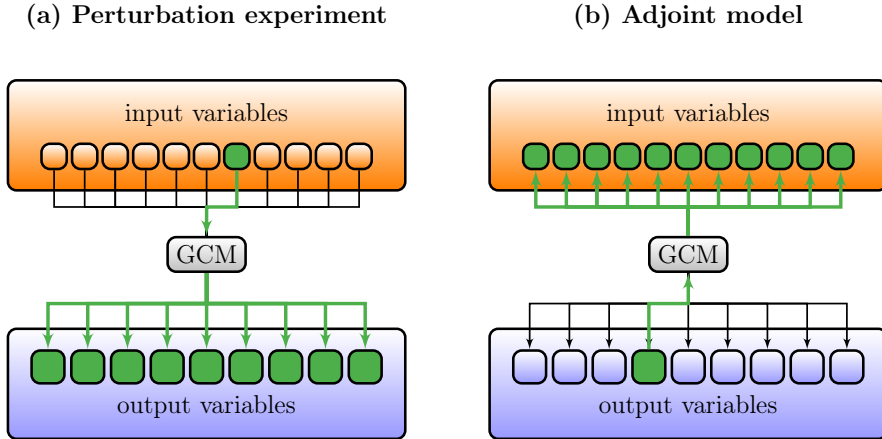


Figure 1.9: Schematic contrasting (a) a perturbation experiment and (b) an adjoint model. A traditional perturbation experiment computes the nonlinear response of all output variables to a single input perturbation. (b) The adjoint model computes the linear sensitivity of one output to all input variables (in a single run). Computing the effect of all input variables on an output variable by means of perturbation experiments is computationally prohibitive for a GCM with millions to billions of input variables. For typical adjoint applications, the output variable of interest is a scalar model-data misfit function (C.I in Fig. 1.8) or climate quantity of interest (C.II in Fig. 1.8).

particularly efficient: it computes the linearized sensitivity of one output to all input variables in a single model run (Fig. 1.9(b)). In contrast, running a traditional perturbation experiment assesses the non-linear response of all model outputs to a single input perturbation (Fig. 1.9(a)). To determine the effect of all input variables on the output variable of interest by means of perturbation experiments would require N (on the order of millions to billions) simulations with an ocean GCM, each representing a perturbation experiment for a different input variable. This is computationally prohibitive.

To generate an adjoint model from its parent non-linear GCM, one requires an algorithmic differentiation tool that generates line-by-line derivatives of the model code [Giering and Kaminski, 1998]. To obtain the derivative of the entire code, the line-by-line derivatives have to be connected by the chain rule. The adjoint model applies the chain rule from bottom to top, and is therefore sometimes thought of as operating backwards (cf. the upward-pointing green arrows in Fig. 1.9(b)). Closely linked to the adjoint model is the tangent linear model. The latter evaluates the chain rule from top to bottom, thereby operating in the forward sense. Conceptually, a tangent linear model corresponds therefore to the schematic in Fig. 1.9(a): it computes linearized sensitivities of all output variables to one input variable (while a perturbation experiment evaluates non-linear responses). While applying the chain rule from bottom to top, the adjoint model traces the dependencies (arrows in Fig. 1.8) of the output variable under consideration backwards to all input variables - and, on the fly, to all prognostic variables (the small gray boxes in Fig. 1.8). These dependencies can be of complicated nature because the input variables enter the GCM calculations at different times, the simulated ocean

state at any time depends on the previous ocean states, and the output variable may depend on the simulated ocean state in an involved manner (see the network diagram in Fig. 1.8).

1.2.1 State estimation

Ocean state estimation combines information contained in (a) ocean observations and (b) physical laws, by “fitting” an ocean GCM optimally to the available ocean observations. The fitting is done by adjusting the uncertain input variables $\mathbf{x} = (x_1, \dots, x_N)$ in box A of Fig. 1.8, also called the ‘control variables’. Typically, control variables are adjusted on a grid point basis of the underlying GCM, i.e., initial conditions and model parameters are adjusted at every model grid cell, and the time-evolving atmospheric forcing fields at every surface ocean grid cell. This results in a tremendously large control space dimension (N), typically on the order of $O(10^6) - O(10^8)$ [Forget et al., 2015]. The optimal set of control variables is given by \mathbf{x}_{\min} , the minimizer of the regularized model-data misfit function

$$J(\mathbf{x}) = J_{\text{misfit}}(\mathbf{x}) + J_{\text{reg}}(\mathbf{x}). \quad (1.1)$$

The minimizer \mathbf{x}_{\min} can be found through an iterative procedure, where the gradient information $\nabla_{\mathbf{x}}J$ provided by the adjoint elucidates where to place a next better guess - until convergence is obtained. The first term in (1.1), $J_{\text{misfit}}(\mathbf{x})$, measures the misfit between the observations and the observations’ counterparts simulated by the model (see left branch in Fig. 1.8). The second term in (1.1), $J_{\text{reg}}(\mathbf{x})$, penalizes deviations from a first-guess. The penalization term is added to ‘regularize’ the minimization problem, which otherwise would be ill-posed, i.e., have many possible solutions [Wunsch, 1996; Tarantola, 2005].

Using the adjoint method to minimize (1.1) for the purpose of state estimation is also referred to as *solving a non-linear inverse problem*. The term ‘inverse problem’ can be explained as follows. Given the available observations at the end of the flowchart of Fig. 1.8, one adjusts the control variables at the start of the flow chart such as to minimize the discrepancy between the model and observations. This procedure can be thought of as ‘inverting’ for the optimal set of control variables [Wunsch, 1996; Tarantola, 2005]. The solution to a *non-linear inverse problem*, or an ocean state estimate, is a freely running forward model. Nonlinear inverse problems therefore guarantee solutions that are dynamically and kinematically self-consistent, a desirable property for studying the ocean’s climate and variability [Wunsch and Heimbach, 2007; Stammer et al., 2016].

Ocean state estimation is for example pursued by the ECCO (Estimating the Circulation and Climate of the Ocean) project [<http://www.ecco-group.org/>]. The ECCO version 4 (ECCOv4) product is the dynamically consistent solution to a non-linear inverse problem, a hindcast of the evolving ocean state from 1992 to 2015, covering the era of satellite altimetry [Forget et al., 2015; Fukumori et al., 2017]. ECCOv4 synthesizes the MIT general circulation model (MITgcm; <http://mitgcm.org/>; [Marshall et al., 1997a,b]) with several hundred million diverse oceanographic observations, including satellite altimetry data, GRACE ocean bottom pressure, temperature and salinity

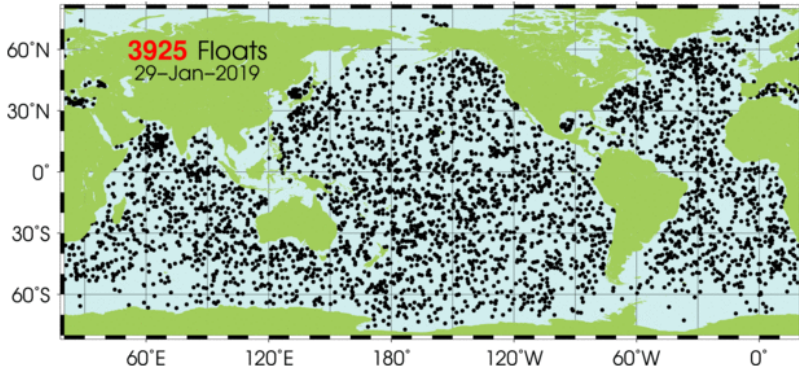


Figure 1.10: Spatial coverage of 3925 Argo floats on January 29, 2019, retrieved from argo.ucsd.edu/About_Argo.html. The temperature and salinity data from the Argo profiling floats are used as data constraints in the ECCOv4 state estimate, alongside with other observations including satellite altimetry data, GRACE ocean bottom pressure, sea-ice measurements, and temperature and salinity profiles from other sources such as CTDs, XBTs, moorings, and seals [Forget et al., 2015; Fukumori et al., 2017].

profiles from Argo, CTDs, XBTs, moorings, seals, and sea-ice measurements. Fig. 1.10 shows a snapshot of the spatial coverage of Argo floats on January 29, 2019.

1.2.2 Sensitivity analysis

Adjoint-based sensitivity analysis reveals causal chains and dynamical relationships among physical variables encoded in the model. Given a chosen scalar quantity of interest, the adjoint model provides the linearized sensitivity to model parameters and forcing fields throughout the model domain, as a function of forcing lead time. For instance, adjoint sensitivities have been used to study the dynamical cause of Atlantic meridional heat transport [Marotzke et al., 1999; Heimbach et al., 2011], the Atlantic meridional overturning circulation [Czeschel et al., 2010; Heimbach et al., 2011; Pillar et al., 2016; Smith and Heimbach, 2019], Florida Current transport [Czeschel et al., 2012], sea level on the Californian coast [Verdy et al., 2013], ocean bottom pressure in the Arctic Mediterranean [Fukumori et al., 2015], and Labrador Sea heat content [Jones et al., 2018]. Moreover, projection of historic forcing anomalies (e.g., from atmospheric reanalyses) onto the corresponding adjoint sensitivities, provides a quantitative reconstruction of wind- and buoyancy-driven variations in the studied oceanic quantity [Pillar et al., 2016; Smith and Heimbach, 2019].

The approach of using adjoint sensitivities stands in sharp contrast to statistical methods, such as correlations or regression. Adjoint sensitivities identify physical relationships and causal chains contained in the model equations (which in turn only approximately represent the real ocean). On the other hand, correlations provide an empirical measure of how two variables vary concurrently or/and at some lead/lag, regardless of whether or not they are causally related. To demonstrate that correlation

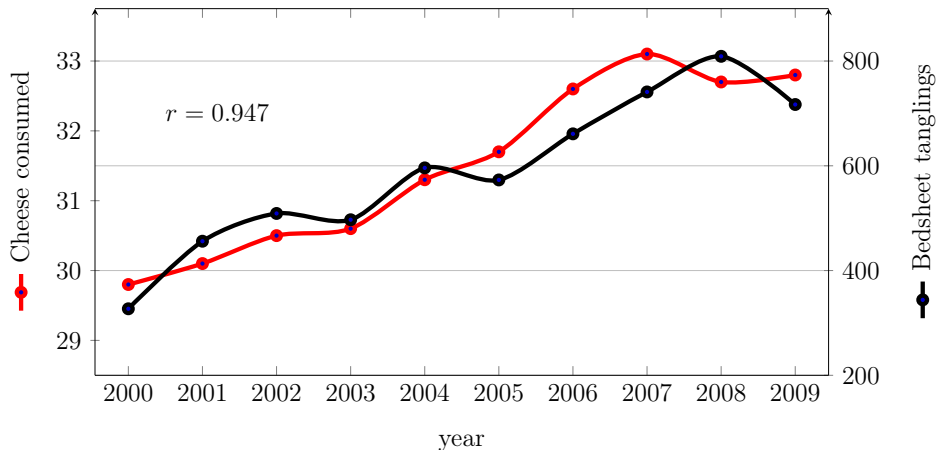


Figure 1.11: The per capita cheese consumption in the US correlates with the number of people in the US who die by becoming tangled in their bedsheets, with a correlation factor of $r = 0.947$ (replotted from [Vigen \[2019\]](#), who makes a case for “correlation does not guarantee causation”). While correlations can be spurious, adjoint sensitivities identify physical relationships contained in the equations of the underlying model. The adjoint of an accurate model would (hopefully!) find no link between consumed cheese and bedsheet tanglings.

does not guarantee causation, [Vigen \[2019\]](#) points out that the per capita cheese consumption in the US correlates with the number of people in the US who die by becoming tangled in their bedsheets (Fig. 1.11).

The adjoint model approach is a powerful tool which allows to identify key regions and physical mechanisms impacting a quantity of interest - in a way that is not feasible with traditional forward perturbation experiments (see Fig. 1.9). However, a caveat of the adjoint approach is that adjoint-derived sensitivities are a linearized approximation to the non-linear responses, and may become less and less accurate with larger perturbations and longer model integrations. Moreover, discontinuous processes, e.g., due to regime transition, can cause instability in the adjoint model and often have to be substituted by approximate adjoints [e.g., [Forget et al., 2015](#)]. Due to these limitations, the linearized approximation and inexactness of the adjoint, the adjoint-derived sensitivities are to be evaluated against (probe) non-linear perturbation experiments in order to assess their applicability. The acceptable size of the input perturbation and considered timescale will depend on the studied quantity of interest and the model setup. For instance, [Czeschel et al. \[2010\]](#) find that in their MITgcm configuration at 1° horizontal resolution, the adjoint and non-linear forward models show broadly consistent results for the evolution of the AMOC at 27°N over the first 15-20 years, after applying basin-wide heat flux perturbations of amplitude 15 W/m^2 over the subpolar gyre.

1.3 Observing system design

Ocean observing systems provide information required to understand ocean circulation and the ocean's impact on climate and climate change, as well as to assess climate model performance. Due to the ocean's long memory, ocean observations need to be sustained for decades (or even centuries) in order to substantially affect our present and future knowledge of global ocean circulation [e.g., [National Academies of Sciences, Engineering, and Medicine, 2017](#)]. However, observing systems are expensive to deploy and maintain. It is thus an urgent task to quantitatively assess the effectiveness of existing and potentially new observing systems.

In Section 1.3.1, I present frameworks within which observing systems can be evaluated and designed. Section 1.3.2 explains the concept of Observing System [Simulation] Experiments (OS[S]Es), the currently most popular approach to Observing System Design. In Section 1.3.3, I give an overview of adjoint-based methods for Observing System Design and whether they have been used for climate applications yet. Finally, in Section 1.3.4, I briefly compare OSSEs and adjoint-based methods—with a view to ocean climate research.

1.3.1 Frameworks

An ocean observation contains information on a much larger part of the ocean state than the local measurement itself. This is because many regions of the ocean are connected over long distances and time scales, e.g., through advective and diffusive processes, or barotropic and baroclinic waves. Observing systems therefore have proxy potential for other, possibly remote and unobserved, climate quantities of interest. *Ocean data assimilation systems* account for this potential by systematically combining observations with dynamical principles, expressed through a numerical model. In a data assimilation system, the model acts as a dynamical interpolator, which propagates the assimilated data from the observing system through space and time. Data assimilation systems are therefore useful frameworks to evaluate ocean observing systems and their potential to monitor non-measured quantities.

However, data assimilation systems differ in the techniques they apply [e.g., [Law et al., 2015](#), [Stammer et al., 2016](#); [Carrassi et al., 2018](#)]:

- **Filters** estimate the ocean state in a temporal *sequential* fashion. Typically, updates of the model forecast are performed every few hours, and for each update information from the new observations since the most recent update are used. Filter-based methods are commonly used for numerical weather prediction and short-term operational ocean forecasts, but also for many ocean reanalyses products [see [Balmaseda et al., 2015](#), for a review and intercomparison]. During the sequential updates of the model forecast (the analysis step), artificial sources and sinks of basic properties are introduced. While filter-based methods acknowledge that the dynamic model is not perfect, their major drawback is that the simulated trajectories of the state variables are no longer continuous and exhibit jumps at each analysis step. An advantage of filter-based methods is that they often compute an uncertainty estimate as part of the solution algorithm. Examples for filter

methods are the Kalman filter, the Ensemble Kalman Filter, or the 3-dimensional variational (3D-Var) method.

- **Smoothers** propagate the information contained in observations both forward and backward in time to constrain ocean circulation in a retrospective analysis. Moreover, smoother-based solutions produce estimates of the ocean state with smooth trajectories. An example for a smoother-based approach is solving a *non-linear inverse problem*, as for instance done by the ECCO project [Forget et al., 2015]. Non-linear inverse problems generate solutions that are freely running forward models simulating hindcasts of the time-evolving ocean state. By means of the adjoint method, the model’s initial and boundary conditions, and sometimes uncertain parameters, are adjusted such as to minimize the discrepancy between the model and observations (Section 1.2.1). Non-linear inverse problems therefore guarantee solutions that obey the ocean dynamics as embedded in the underlying GCM, and are dynamically and kinematically self-consistent over the full estimation period (years to decades). On the downside, uncertainty estimates, which could be derived from the Hessian of the cost function, are not directly part of the solution algorithm, and are computationally demanding (see Section 1.3.3). The 4-dimensional variational (4D-Var) method, in its ‘strong-constraint’ formalism, is the analogue of non-linear inverse problems in numerical weather prediction. Note, however, that the length of the 4D-Var assimilation windows is often chosen on the order of hours [e.g., Dee et al., 2011], instead of the full estimation period, due to the chaotic character of the atmosphere [Pires et al., 1996]. As a result, the continuity of the physical system is broken in between the assimilation windows. An approximation of the 4D-Var method by a sequence of quadratic minimization problems (‘incremental 4D-Var’) is also used within the Regional Ocean Modeling System (ROMS), where the length of the 4D-Var assimilation windows is chosen on the order of a few days [Moore et al., 2011a,b]. Another way to obtain smooth trajectories is the Kalman smoother, a generalization of the Kalman filter that includes not only past values but also future ones. While the Kalman smoother provides uncertainty estimates, the solution does not obey the model equations.

Dynamically consistent solutions, where conservation laws are respected, global ocean budgets are closed, and oceanic transports are physically meaningful over long timescales, are crucial to understand ocean variability on climate timescales. Solving a non-linear inverse problem is therefore the preferred data assimilation technique in support of ocean climate applications [Wunsch and Heimbach, 2007; Stammer et al., 2016; Pilo et al., 2018]. Moreover, the “unbroken” physics in the solution of an inverse problem ensures that ocean climate signals are propagated in a dynamically consistent way over long distances and time scales, such that proxy potentials of observing systems can be detected - based on dynamical principles. The framework of a non-linear inverse problem is therefore the appropriate setting to evaluate and design observing systems that are targeted at climate monitoring.

1.3.2 OS[S]Es

The most popular approaches for evaluating observing systems are observing system experiments (OSEs) or observing system simulation experiments (OSSEs), which are

performed within a data assimilation system.

OSEs In an OSE (sometimes referred to as observation withholding or denial experiment), one withholds or adds a specific component of an observing system in the data assimilation system. One then compares the simulation in which the full observing system was assimilated with the simulation in which the component was withheld [e.g., Oke et al., 2010; Fujii et al., 2019]. The difference in certain ocean fields or diagnostics, caused by the added component, is then interpreted as an error reduction. However, the main disadvantage of an OSE is that the error reduction cannot be estimated accurately because the true state is not known. Another drawback is that OSEs can only be used to evaluate preexisting, not future, observing systems.

OSSEs The two highlighted flaws of an OSE are fixed in an OSSE by using a so-called ‘nature run’. The nature run is a free-running high-resolution simulation with the best model available, and is meant to mimic the true ocean [e.g., Hoffman and Atlas, 2015]. Preferably, the nature run has a different model base than the data assimilation system used. Synthetic observations are sampled from the nature run, and suitable errors are added. Next, the synthetic observations are assimilated into two experiments, where one experiment *does*, and the other experiment *does not* use a specific component of the observing system. The impact of the added component of the observing system is the error reduction, measured by comparing both assimilation experiments to the nature run. The main criticism of OSSEs is that nature runs are often not good enough to realistically model the true ocean and the phenomena of interest [Hoffman and Atlas, 2015]. Another problem is “incestuousness” between the model bases for the nature run and the data assimilation system (personal communication with Nikki Privé, July 2018). A third issue is that results may be system-dependent. Moreover, there is the risk that results may only apply within the used OSSE system, but are not connected to the real world [Fujii et al., 2019].

Applications of OSSEs OSSEs are widely performed in the numerical weather prediction community, to test and accelerate the operational use of future observing systems, [e.g., Masutani et al., 2010], but also to investigate and improve the performance of data assimilation systems [e.g., Privé and Errico, 2013]. However, OSSEs are not restricted to operational and forecast systems, but can in principle be applied to any data assimilation system, including non-linear inverse problems. OSSEs require heavy computer and human resources, due to the requirement of

- (i) a high-resolution nature run, to be validated against the real world;
- (ii) multiple assimilation experiments within a separate state-of-the-art data assimilation system (one control experiment that assimilates all existing observations, and one additional experiment for each tested observing system, which assimilates all existing plus the new observations).

In two recent studies, Li et al. [2017] and Wang et al. [2017] use OSSEs in a simplified form - here, referred to as *OSSE-type methods* - for oceanographic applications. Both studies entail a nature run (OSSE-component (i)), but *not* a data assimilation system (OSSE-component (ii)). The function of a data assimilation system in conventional OSSEs is

to combine the (synthetic) observations with ocean dynamics in order to derive target quantities that are not directly measured (Section 1.3.1), which are then compared to the true values in the nature run. Both Li et al. [2017] and Wang et al. [2017] substitute the data assimilation system by offline dynamical and statistical calculations. For instance, Li et al. [2017] use a statistical estimation method (Objective Analysis) to interpolate the irregularly spaced temperature, salinity and density (pseudo) data from the OSNAP moorings to the full OSNAP section. In addition, dynamical calculations based on first principles, such as geostrophic balance and Ekman transports, are made to infer velocities across the section. Statistical and dynamical calculations combined give an estimate of volume, heat and salt transport across the OSNAP section. The estimated transports inferred from the pseudo observations are then compared to the true transports in the nature run in order to evaluate whether the OSNAP array configuration and the used calculation methods allow for accurate transport estimates. Similarly, Wang et al., 2017 make offline calculations, based on hydrostatic balance, to infer steric height from pseudo temperature and salinity measurements, sampled at potential future instrument locations. In a second step, they derive sea surface height (SSH) at the SWOT scales (the target quantity) from steric height, by applying a statistical filter to remove large-scale signals. The estimated SSH is then compared to the true SSH in the nature run to assess the efficacy of the tested observing systems together with the calculation methods.

One may argue that the OSSE-type methods applied by Li et al. [2017] and Wang et al. [2017] use (to first order) the same dynamical laws encoded in a GCM that underlies a data assimilation system, such that the latter, i.e., OSSE-component (ii), can as well be skipped. However, the unphysical *statistical* calculations in OSSE-type approaches would be unnecessary, and replaced by dynamical information, if a data assimilation system was used, at least when working within the framework of a non-linear inverse problem, rather than a filter-based system (see Section 1.3.1). More disadvantages of OSSE-type methods that do not use a data assimilation system are that the calculation methods (i) might not involve all existing observations (while a large fraction of them might be embedded in a state-of-the-art data assimilation system) and (ii) are developed only for specific target quantities (e.g., transports across the OSNAP section in Li et al., 2017 or sea surface height in Wang et al., 2017), such that the information gain for other, possibly remote, quantities cannot be assessed.

1.3.3 Adjoint-based methods

Adjoint-based methods are a powerful tool for observing system design because an adjoint model enables the detection of teleconnections, physical relationships and causal chains (Section 1.2) that connect the observed quantities to the rest of the global ocean. Nevertheless, the full power of adjoint-based methods has so far not been realized for ocean observing system design. Adjoint-based methods for observing system design differ in their complexity, and not all of them are (yet) used in ocean climate research. In the following, I will briefly discuss these methods and their fields of applications, grouping the methods according to their level of complexity, from Level 1 (simple) to Level 4 (complex). The methods are listed in Table 1.1. Fig. 1.12 sketches the ability of the methods to evaluate observing systems in an increasingly comprehensive fashion.

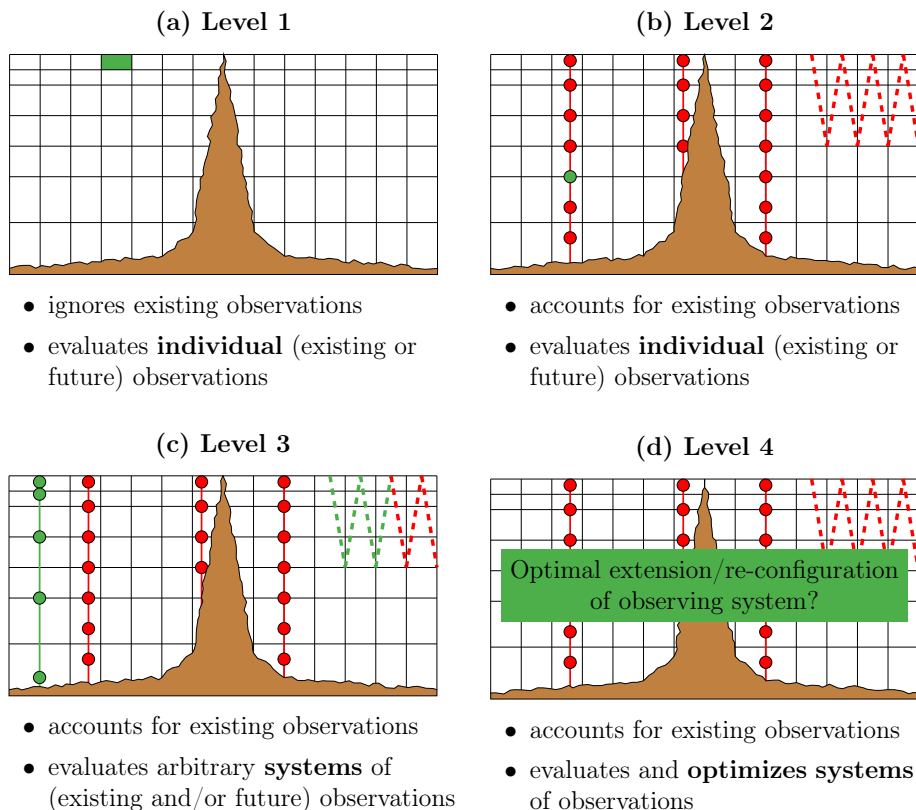


Figure 1.12: Schematics that sketch the ability of the adjoint-based methods in Table 1.1 to evaluate observing systems in an increasingly comprehensive fashion, from (a) to (d). In (a)-(d), red features indicate existing observations that are taken into account, and green features show examples for the observation(s) that can be evaluated by the respective method. The black lines represent the model grid and the brown ridge delineates bottom topography. In this example, the observing system consists of instruments (colored dots) at different depths on several moorings (colored vertical lines), and a glider (colored saw-tooth lines).

Level	Method	References
1	Adjoint sensitivities & Optimal excitations	Marotzke et al. [1999], Heimbach et al. [2011], Zanna et al. [2010]
2	Observation sensitivity	Baker and Daley [2000] (NWP) , Daescu [2008] (NWP), Moore et al. [2011c, 2017b], Köhl and Stammer [2004]
3	Hessian-based Uncertainty Quantification	Bui-Thanh et al. [2012] (CS), Kalmikov and Heimbach [2014, 2018], Moore et al. [2011b, 2017a], Kaminski et al. [2015, 2018]
4	Hessian-based Optimal Experimental Design	Alexanderian et al. [2016] (CS)

Table 1.1: Adjoint-based methods for Observing System Design, ordered according to their level of complexity from simple (Level 1) to complicated (Level 4). Not all methods are yet used in ocean climate research. References from other fields carry the discipline in parantheses (NWP: Numerical Weather Prediction, CS: Computational Science).

Level 1: Adjoint sensitivities

At the lowest level, initial insights into desirable observations can be inferred from performing an adjoint sensitivity analysis for one or multiple given climate quantities of interest (QoIs), as described in Section 1.2. Examples of QoIs are heat content of an ocean basin, or volume or heat transport across a certain section. The adjoint is used to compute $\nabla_{\mathbf{x}}\text{QoI}$, pursuing branch (II) in Chart 1.8. One thereby obtains the sensitivities of the QoI to control and state variables, e.g., to atmospheric forcing or temperature in the green model grid cell in Fig. 1.12(a). It has been suggested that regions that show high sensitivity [e.g., Marotzke et al., 1999; Heimbach et al., 2011], or optimal excitability [e.g., Zanna et al., 2010, 2012; Sévellec et al., 2007, 2017, 2008] are to be prioritized when designing observing systems. However, this approach has the following limitations:

- (a) It evaluates only observations that can be identified directly with adjoint control or state variables, which usually limits the pool of potential observations to hydrographic or altimetric measurements located at the model grid cells and time steps.
- (b) It is not embedded into an inverse modeling framework. It therefore neglects data uncertainties and prior knowledge on ocean circulation; these are crucial ingredients for a non-linear inverse problem (and any data assimilation system), which offers the framework for a comprehensive evaluation of observing systems (see Section 1.3.1).
- (c) It does not take into account already deployed instruments. It can therefore not evaluate the information gain/redundancy in the context of existing observations.

- (d) It can only assess the value of individual (existing or future) observations at a single point in time, but not systems that consist of multiple observations in space and time.
- (e) It does not quantify how well ocean circulation or a certain QoI are being observed, but only how sensitive a QoI is to an observation. High sensitivity to one observation (compared to other candidate observations) does not guarantee high observing skill. For instance, if all candidate observations cannot inform about certain circulation features, even the best observation - identified by highest sensitivity - will have poor observing skill.

While conventional adjoint sensitivity studies give valuable first insights, the limitations listed above inhibit their usage for quantitative observing system design.

Level 2: Observation sensitivity

Out of the five limitations listed in the last paragraph, (a), (b) and (c) can be overcome by computing the sensitivity of a QoI to observations (rather than control and state variables) within an inverse modeling framework. That is, while the “Level 1” method computes $\nabla_{\mathbf{x}}\text{QoI}$ following branch (II) in Fig. 1.8, the “Level 2” method computes $\nabla_{\mathbf{y}}\text{QoI}$, by connecting branches (II) and (I) in Fig. 1.8 via the chain rule. This approach originated in the numerical weather prediction (NWP) community, known as computing the ‘observation sensitivity’ within a 3D-Var [Baker and Daley, 2000; Langland and Baker, 2004; Zhu and Gelaro, 2008; Gelaro and Zhu, 2009] or 4D-Var [Daescu, 2008; Cioaca et al., 2013] data assimilation system. Since in the NWP community the QoI is usually chosen to be the forecast error, the approach is also referred to as assessing the ‘Forecast System Observation Impact’ (FSOI). Computationally, this approach is much more involved than the “Level 1” method, since it requires three additional steps:

- (i) accounting for existing observations and prior knowledge on model input variables;
- (ii) projecting onto the observation space;
- (iii) weighting by observational errors.

Accounting for existing observations through step (i) overcomes limitation (c) of the “Level 1” method; in Fig. 1.12(b), this is visualized by the consideration of the existing (red) observations, while they are neglected in in Fig. 1.12(a). At the same time, step (i) is the computationally most demanding part since it requires left multiplication of $\nabla_{\mathbf{x}}\text{QoI}$ (the “Level 1” sensitivity vector) by the inverse of the Hessian matrix of the cost function (1.1) (corresponding to stages S3 and S4 in [Section 3 of Daescu, 2008]). Step (ii) is accomplished by one tangent linear model run (first half of stage S5 in [Section 3 of Daescu, 2008]); it projects sensitivities from the space of the model control variables, which are allocated on the spatio-temporal model grid, onto the space of observations. This has the effect that observation types, locations and frequencies can be mirrored more accurately in the model (even though their representation will always be limited by model resolution), remedying limitation (a). Finally, accounting for observational errors through step (iii) and prior knowledge on the control variables through step (i) overcomes limitation (b).

The “Level 2” method is a widely used approach in the NWP community, but has also been employed in oceanography within the ROMS 4D-Var data assimilation system by Moore et al. [2011c, 2017b]. It is noted that in another oceanographic application, Köhl and Stammer [2004] use a reduced version of “Level 2”: the authors simplify step (i) by not accounting for existing observations (and thus bypass the matrix-vector computation involving the inverse Hessian). The approach in Köhl and Stammer [2004] simplifies therefore to weighting the “Level 1” adjoint sensitivities by prior and data uncertainties and projecting them onto the observation space, which overcomes limitations (a) and (b) - but not (c) - of the “Level 1” method.

Regardless of whether using the full or reduced “Level 2” method, the method cannot overcome limitation (d). That is, the “Level 2” method can evaluate the impact of an *individual* (existing or future) observation, but not systems consisting of multiple observations in space and time. For instance, one could assess the impact of one temperature or one velocity measurement by a moored instrument at a certain depth and at a certain time, visualized by the green dot in Fig. 1.12(b). However, one could not evaluate the combined impact of the full mooring, i.e., with its instruments at several depths and measurements being taken at many times. Similarly, the “Level 2” method still has limitation (e). Consequently, the method does not inform about how much closer new or existing observations take us to a well-observed ocean or QoI.

Level 3: Hessian-based Uncertainty Quantification

Uncertainty Quantification As opposed to the “Level 2” method, tools from Uncertainty Quantification (UQ) in inverse problems overcome *all* five limitations (a)-(e). By means of the adjoint and tangent linear models of the underlying GCM (Section 1.2), observational information and uncertainty are propagated throughout the global ocean via the linearized model equations. UQ is useful for observing system design because one can measure the effectiveness of a set of observations by the uncertainty reduction that is achieved by including these observations in the inverse modeling framework. Uncertainty reduction can be assessed either for the control variables or for unobserved QoIs, possibly remote from the observations. The strength of the UQ method is that the computed uncertainty reduction is based purely on dynamical information that is propagated via the model equations - i.e., equations of motions representing basic conservation and constitutive laws rendered on a computer - as opposed to techniques that are based on statistical inference, e.g., EOFs. Limitation (d) is overcome because one can assess the value of arbitrary combinations of multiple (existing and/or new) observations of different types in the context of all existing observations. In Fig. 1.12(c), for instance, one could evaluate the combined information gain achieved by a full new mooring, providing measurements at several depths (green dots) and frequencies, together with a re-configured glider (green dashed saw-tooth line). If there are several candidate observation networks, one can perform UQ for each of them and identify the network with the highest information gain. Moreover, the UQ framework provides a way to quantify how well the ocean is observed - with and without the observing system under evaluation. This overcomes limitation (e).

Hessian-based approximation In the Gaussian approximation, posterior uncertainty is captured by the curvature of the regularized model-data misfit function (1.1), expressed through the *Hessian matrix* of J . The Hessian matrix contains the second derivatives of the cost function. An adjoint model, in turn, is able to compute derivatives within the code (Section 1.2). The Hessian-based approximation makes UQ therefore an adjoint-based method for observing system design; here, I classify its complexity as “Level 3”.

Dealing with the (inverse) Hessian While the adjoint model could, in principle, be used to compute the entries of the Hessian matrix, the number of entries is way too large to make the computation and inversion of the full Hessian feasible for a full-fledged oceanographic inverse problem. Since, in a non-linear inverse problem, the control variables are typically adjusted on a grid point basis of the underlying GCM, the control space dimensions is on the order of $O(10^6) - O(10^8)$ [Forget et al., 2015]. The corresponding Hessian therefore consists of $O(10^{12}) - O(10^{16})$ elements, which would require months to years of heavy computer resources for the calculation and inversion, and terabytes to petabytes of storage [cf. Kalmikov and Heimbach, 2018]. However, there are two approaches to deal with the computational challenge:

- **A-priori-reduced control spaces:** The control space is reduced by an *a priori* choice of basis functions. It is then feasible to explicitly compute the full Hessian. This approach has been used in seismology [Hardt and Scherbaum, 1994], for atmospheric transport inversion [Kaminski and Rayner, 2008], in the carbon cycle community, for terrestrial biosphere modeling [see the review by Kaminski and Rayner, 2017], and, recently, for the Arctic sea-ice-ocean system [Kaminski et al., 2015, 2018]. For instance, in their sea-ice-ocean models, Kaminski et al. [2015, 2018] reduce the control space to a total of about 150 control variables in order to compare the efficacy of several hypothetical satellite/airborne measurements for improving sea-ice predictions in the Arctic. It is then feasible to explicitly compute the Hessian (150^2 entries) for model integrations that simulate the sea-ice-ocean state over a period of several weeks, as performed in Kaminski et al. [2015, 2018]. The a-priori-reduction of the control spaces in Kaminski et al. [2015, 2018] is achieved by dividing the Arctic and Nordic Seas into nine regions, over which the control variables are spatially averaged. That is, atmospheric forcing and initial conditions are adjusted uniformly over these nine regions - the ‘*large region approach*’ [Kaminski and Rayner, 2008]) - rather than on a grid point basis of the underlying GCM. Moreover, for the atmospheric forcing fields, only the time-means, not time-variable anomalies, are adjusted. As a result, within any of the nine control regions, changes in neither the spatial patterns of the initial conditions, nor in the spatio-temporal patterns of the atmospheric forcing fields, are resolved. In particular, neglecting time-variable adjustments of the forcing fields is an arguable assumption if the target is forecasting, as in Kaminski et al. [2015, 2018]. In summary, a-priori-reduced control spaces require ad-hoc choices and can lead to large aggregation errors as highlighted in Kaminski et al. [2001].
- **Low-rank approximations of misfit Hessian:** One can exploit the fact that available observations typically inform only a low-dimensional ‘effective’ subspace of the high-dimensional control space. This fact allows to construct low-rank ap-

proximations of the misfit Hessian, i.e., the Hessian of the model-data misfit component J_{misfit} in equation (1.1) [e.g., Bui-Thanh et al., 2012]. Finding the leading eigenvectors of the misfit Hessian characterizes the effective subspace, i.e., which components of the control space are well or poorly determined by the observations. It is noted that the identified effective subspace constrained by the data can be used for an *a-posteriori-reduction* of control spaces, a common tool for data-informed model reduction [e.g., Bashir et al., 2008; Lieberman et al., 2013]. Low-rank approximations of the misfit Hessian via randomized algorithms or Lanczos methods are the state-of-the-art approach in the computational sciences [e.g., Bui-Thanh et al., 2012]. For instance, Isaac et al. [2015] perform Hessian-based UQ based on low-rank approximations in a steady-state ice-flow model for a control space of the dimension $O(10^5) - O(10^6)$, to quantify the effect of satellite measurements for estimating the Antarctic ice mass flux to the ocean. While the ice-flow model in Isaac et al. [2015] is formulated time-independent, oceanographic inverse problems are *time-dependent*, which results in higher-dimensional control spaces and more expensive forward models. Therefore, UQ has not been performed for full-fledged oceanographic inverse problems on climate-relevant time windows. However, recent progress has been made within the ECCO framework: for a simplified version of ECCO, Kalmikov and Heimbach [2014, 2018] use Lanczos methods for low-rank approximations of the misfit Hessian in order to quantify the impact of satellite altimetry data for estimating the Drake Passage transport. Kalmikov and Heimbach [2014, 2018] use a global idealized barotropic configuration, a 3-month time window, and a control space with $O(10^5)$ variables. Low-rank approximations of the posterior error covariance have also been used within the ROMS 4D-Var data assimilation system, a regional model system where the length of the assimilation windows is chosen on the order of a few days [Moore et al., 2011a]. To explore the information content of the California Current Observing System, Moore et al. [2011b, 2017a] make direct use of the Lanczos vectors that are computed as part of the inner quadratic minimization loops within the observation space (which has a dimension of $O(10^4) - O(10^5)$).

Level 4: Hessian-based Optimal Experimental Design

The most complex of all adjoint-based methods presented here is Hessian-based Optimal Experimental Design (OED). OED goes one step further than UQ (the “Level 3” method). While UQ can evaluate information gain via uncertainty reduction for several candidate observing networks, the goal of an OED problem is to find the *optimal* observing network so as to *minimize* uncertainty - either in the control variables or in a specific target quantity. At the same time, one can incorporate penalty functions that are associated with the cost of each instrument [e.g., Alexanderian et al., 2016]. For instance, the goal of an OED problem could be to optimally extend or re-configure the existing (red) observing system in Fig. 1.12(d), in order to minimize uncertainty in the heat flux estimates across the shown section, given a certain budget. As one might expect, OED is technically and computationally even more demanding than UQ (“Level 3”). OED is implemented as a bilevel optimization problem, and so far (to my knowledge) only performed in the computational sciences [e.g., Alexanderian et al., 2016]. Research and applications in the coming years will shed more light on the feasibility - and bring to bear the full potential

- of OED within oceanographic inverse problems of practical relevance.

1.3.4 OSSEs vs. adjoint-based methods

In Sections 1.3.2 and 1.3.3, I presented OSSEs and adjoint-based methods as different approaches to observing system design. These approaches are now contrasted - with a view of ocean climate research.

Goals OSSEs (Section 1.3.2) have proven as an invaluable tool for observing system design when the primary goal is forecast *improvement* - and are therefore widely used in numerical weather prediction and operational oceanography. In ocean climate research, on the other hand, observing systems are not only meant to improve the estimate of the ocean state and its variability; a second aim is to *understand why* the improvement is achieved, which larger climate signals are captured by the observing system, and by which ocean processes remote signals are propagated to the instruments. To this aim, adjoint-based methods (Section 1.3.3) are a powerful tool because they use the adjoint's ability to detect physical relationships and causal chains that connect the observed quantities to the rest of the global ocean (Section 1.2).

Philosophy Both the UQ (i.e., the adjoint-based “Level 3” method presented in Section 1.3.3) and the OSSE methodologies evaluate future or existing observing systems by measuring the error reduction caused by adding the observations to the data assimilation framework. However, the methodologies differ in their underlying philosophies. OSSEs evaluate against a truth given by the nature run. In contrast, the UQ approach sticks to a probabilistic framework, in which truth is unknown, and information gain from observations is measured by the achieved uncertainty reduction, after the information contained in the observations is dynamically propagated via the model equations. It is noted that the probabilistic UQ framework offers a mathematical foundation for exploring *optimal* instrument placements (described as the adjoint-based “Level 4” method in Section 1.3.3), which is not possible within an OSSE.

Feasibility The method of using adjoint sensitivities (adjoint-based “Level 1” method) is based on a single adjoint model integration - a computationally very effective way to evaluate the impact of observations on oceanic quantities of interest (see Section 1.2). Similarly, the method of calculating observation sensitivity is computationally effective, but only if the method is used in a simplified fashion (adjoint-based “Level 2” method, simplified), where one ignores existing observations in order to avoid dealing with the inverse Hessian (Section 1.3.3). However, performing Hessian-based observation sensitivity studies, Hessian-based UQ, and Hessian-based OED (adjoint-based “Level 2”, “Level 3” and “Level 4” methods) is computationally demanding (in increasing order) and has so far not been performed for full-fledged oceanographic inverse problems (Section 1.3.3). However, recent progress has been made within the ECCO framework [Kalmikov and Heimbach, 2014, 2018]. Moreover, it is noted that “Level 2” and “Level 3” have been performed within the ROMS 4D-Var data assimilation system [Moore et al., 2011c, 2017b], a regional model system where the length of the assimilation windows is chosen on the order of a few days. Research and applications in the coming years will shed more light on

the feasibility of Hessian-based methods for dynamically consistent ocean state estimates over climate-relevant time windows.

OSSEs come at a heavy computational cost and are therefore usually only run to assess observing systems that target monitoring on shorter timescales (Section 1.3.2). Recently, OSSEs have been used in more simplified and, thus, cost-efficient forms for oceanographic applications, e.g., in the study by Li et al. [2017] who target ocean monitoring on climate timescales. The OSSE-type methods use a nature run (the first component of an OSSE), but not a data assimilation system (the second component of an OSSE). The data assimilation system, which in conventional OSSEs acts as a dynamical interpolator that propagates the assimilated data from the observing system through space and time, is partly substituted by offline statistical calculations. While such OSSE-type methods are cost-efficient, the full dynamical information contained in model and observations is not extracted (see Section 1.3.2).

Limitations OSSEs rely on the development of a realistic nature run, with a model base that is “independent” from the model base of the data assimilation system used. On the other hand, the probabilistic UQ framework relies on the specification of appropriate prior and noise covariance matrices in order to provide useful uncertainty estimates. Both are extremely difficult tasks [e.g., Hoffman and Atlas, 2015 and Forget et al., 2015]. A limitation of Hessian-based UQ is that the posterior probability distribution is locally approximated by a Gaussian. The accuracy of the approximation depends on how well the linearized model about the cost function minimizer (the most likely state) can reflect the full response of the non-linear system. In contrast, OSSEs operate in fully nonlinear models. For both the OSSE and adjoint-based approaches, results have to be interpreted carefully due to their potential model- and system dependency, and insufficient representation of sub-grid scale processes in the model. A more detailed discussion on the limitations of Hessian-based UQ can be found in Section 3.6.2.

1.4 Thesis objectives and outline

In view of the open questions and observational needs listed in Section 1.1.6, the main goals of this thesis are to

- (i) contribute to our understanding of the physical mechanisms linking the subpolar North Atlantic and Nordic Seas;
- (ii) explore a novel approach to dynamics-based observing system design for the subpolar North Atlantic.

The first 21-month record from the OSNAP array suggest that, over the course of the record, the Nordic Seas are the dominant player in setting the strength and variability of the overturning and heat transport in the subpolar basin [Lozier et al., 2019]. Moreover, variability in the poleward progression of ocean heat across the Nordic Seas, from the subpolar North Atlantic towards the Arctic Ocean, has been linked to Arctic sea ice extent [Carmack et al., 2015; Zhang, 2015; Polyakov et al., 2017]. However, the physical mechanisms driving Nordic Seas heat content variability are not well understood, and especially the relative roles of ocean dynamical processes and local atmospheric forcing

are debated [e.g., [Furevik and Nilsen, 2005](#), and references therein]. Chapter 2 targets goal (i) and presents adjoint sensitivity analyses (see Section 1.2.2) for the Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge on a ten-year timescale. Adjoint-derived sensitivities to atmospheric forcing anomalies are used to disentangle the dynamical mechanisms responsible for generating temperature anomalies in the Nordic Seas, as well as contributions from wind and buoyancy forcing, including time and space origins of these forcings.

Chapters 3 and 4 explore Hessian-based UQ, the adjoint-based “Level 3” method for observing system design (see Table 1.1), from a *dynamical viewpoint*. In Sections 1.3.3 and 1.3.4, I presented an overview of why Hessian-based UQ may be a powerful tool for observing system design. It provides a rigorous mathematical (probabilistic) framework for assessing the impact of existing or future observing systems, in the context of all existing observations. Moreover, it is favorable for climate monitoring, because the dynamical relationships and budgets are respected and treated consistently over long timescales (years to decades). These assets qualify UQ as a promising tool for building and sustaining a cost-effective, long-term Atlantic observing system, while maximizing the information extracted from the observations. The design of an effective Atlantic observing system was identified as an urgent need in Section 1.1.6. However, the tool ‘Hessian-based UQ’ is rather complex and might be deemed a computational black box which lacks simple interpretations. To advance dynamical understanding of this tool, the formal notion of UQ is translated to dynamical concepts in Chapter 3. It is furthermore demonstrated how the previously used method of adjoint sensitivities (“Level 1”) fits into the UQ framework (“Level 3”). Given the novelty of the “Level 3” approach for oceanographic applications, Chapter 3 also clarifies some aspects of the machinery, e.g., the effects of the necessary assumptions on prior information and observational noise. While Chapter 3 focuses on establishing the methodological framework, Chapter 4 applies the framework and the dynamical insights from Chapter 3 to the OSNAP array. Hessian-based UQ within the state-of-the-art ECCOv4 inverse modeling framework is used to explore the constraints of the OSNAP observations on remote and unobserved QoIs. Thus, Chapter 4 undertakes the novel task to put the brand-new OSNAP observations into a broader spatial and temporal context.

The results in Chapters 2 and 4 rely on adjoint-derived sensitivities that are computed within the ECCOv4 inverse modeling framework. Limitations of adjoint-derived sensitivities are that (i) they reflect only a linearized approximation to the non-linear responses and (ii) the adjoint is inexact for certain physical processes (see Section 1.2). Therefore adjoint-derived sensitivities are evaluated against non-linear perturbation experiments in Chapter 5. The linearity checks assess the applicability of adjoint-derived sensitivities in the subpolar North Atlantic - a region known for its complicated and non-linear dynamics - on multi-annual to decadal timescales for atmospheric forcing perturbations of expected amplitude.

Chapter 6 presents a summary and discussion of the main results of this thesis. Chapter 7 provides the main conclusions.

Chapter 8 suggests avenues for future work and presents first results for two further

applications. The first application is motivated by the fact that the interaction of warm subpolar North Atlantic ocean waters with Greenland's marine-terminating glaciers increases submarine melting and has - very likely - an impact on the glaciers' retreat and acceleration (see Section 1.1.3). Since it is logistically challenging to directly measure oceanic heat transport to the ice margin in Greenland's glacial fjords, Chapter 8 asks whether remote ocean observing systems, such as the OSNAP array, can constrain subsurface temperature at Greenland's margins. The second application is to use adjoint models for learning more about past climates. To this aim, Chapter 8 presents first results on dynamical mechanisms that may have contributed to observed past climate variability (see Section 1.1.4) at marine sediment core locations. Furthermore, the framework developed in Chapter 3 is used to explore how good/poor the constraints of proxy data on past ocean circulation are.

Chapter 2

Drivers of Upper-Ocean Heat Content Anomalies in the Nordic Seas

2.1 Introduction

The first 21-month record from the OSNAP array suggest that, over the course of the record, the Nordic Seas are the dominant player in setting the strength and variability of the overturning and heat transport in the subpolar basin [Lozier et al., 2019]. Moreover, variability in the poleward progression of ocean heat across the Nordic Seas, from the subpolar North Atlantic towards the Arctic Ocean, has been linked to Arctic sea ice extent [Carmack et al., 2015; Zhang, 2015; Polyakov et al., 2017] and mass loss from the Greenland Ice Sheet [Holland et al., 2008; Straneo et al., 2010; Rainsley et al., 2018]. Furthermore, upper-ocean heat content variability in the Nordic Seas is reflected in northwestern European climate, and is a skillful metric in decadal climate predictions [Arthun et al., 2017; Yeager and Robson, 2017]. However, the physical mechanisms driving Nordic Seas heat content variability are not well understood, and especially the relative roles of ocean dynamical processes and local atmospheric forcing are debated [e.g., Furevik and Nilsen, 2005, and references therein]. The aim of this chapter is therefore to disentangle local and remote processes that generate upper-ocean heat content anomalies in the Nordic Seas on timescales up to a decade, and identify how atmospheric forcing anomalies in remote regions can generate upper-ocean heat content anomalies in the Nordic Seas. The approach is to compute adjoint sensitivities of Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge to local and remote atmospheric forcing within the ECCOv4 ocean state estimation framework.

Observational, modeling and theoretical studies over the last few decades have elucidated the physical processes that are candidates for driving upper-ocean heat content variability, even if their relative contributions in the Nordic Seas remain unclear. These processes are traditionally divided into two groups: locally and remotely driven processes. Locally driven processes have been illuminated by Hasselmann's theory of stochastic climate models, which explains how local fluctuations in air-sea heat fluxes are temporally integrated by a passive ocean and thereby lead to lower-frequency sea surface temperature (SST) variations [Frankignoul and Hasselmann, 1977; Cayan, 1992; Battisti et al., 1995]. Together with wind-driven Ekman transport and pumping, these local one-dimensional processes explain much of the upper-ocean heat content variabil-

ity at mid-latitudes on interannual timescales [Frankignoul, 1985; Alexander and Deser, 1995]. On the other hand, remotely driven processes require active ocean dynamics, and often longer timescales, to cause upper-ocean heat content anomalies in the region under consideration. Governing ocean dynamics can for instance involve advective transport of temperature anomalies by strong ocean currents [Sutton and Allen, 1997], or changes in the strength and/or pathways of these currents [Grötzner et al., 1998; Dong and Sutton, 2001; Kwon and Frankignoul, 2014].

The relative importance of the above processes depends strongly on the region considered [Buckley et al., 2014], and it is debated how these processes interact in generating anomalies in Nordic Seas upper-ocean heat content. To investigate the relative contributions of local air-sea fluxes and advective ocean processes, many authors have analyzed lagged correlations in either observational datasets or GCM output [e.g., Holliday et al., 2008; Furevik, 2000; Dong and Sutton, 2001; Carton et al., 2011; Mork et al., 2014; Årthun and Eldevik, 2016; Årthun et al., 2017]. The studies differ in their conclusions, although the authors generally agree that advective processes play a prominent role for setting upper-ocean heat content variability in the Nordic Seas. However, remotely triggered perturbations are usually not further disentangled according to their forcings and regions of origin. This is partly due to the fact that the employed technique of using correlations does not allow identification of physical causation, only mutual variability. In contrast, adjoint-based sensitivity studies reveal causal chains and dynamical relationships among physical variables encoded in the model, and shall be the approach of this work.

Adjoint-derived sensitivities have been used to study the dynamical cause of variability of various ocean diagnostics, such as Atlantic meridional heat transport [Marotzke et al., 1999; Heimbach et al., 2011], the Atlantic meridional overturning circulation [Czeschel et al., 2010; Heimbach et al., 2011; Pillar et al., 2016; Smith and Heimbach, 2019], North Atlantic SST on seasonal timescales [Junge and Haine, 2001], Florida Current transport [Czeschel et al., 2012], sea level on the Californian coast [Verdy et al., 2013], and ocean bottom pressure in the Arctic Mediterranean [Fukumori et al., 2015]. In a recent study, Jones et al. [2018] employed adjoint-derived sensitivities to investigate influences on the heat content of the Labrador Sea on a ten-year timescale, within the ECCOv4 adjoint modeling framework. This chapter performs a similar study to Jones et al. [2018], but addresses mechanisms and pathways via which global atmospheric forcing influences heat content in the Nordic Seas, rather than the Labrador Sea, on timescales up to a decade.

The adjoint sensitivity analysis in this chapter classifies the contribution of anomalous atmospheric forcing in different regions across the globe to variability in Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge. There are multiple possible ways in which atmospheric forcing in a remote region can affect Nordic Seas upper-ocean heat content on decadal timescales. Prominent candidates are anomalous buoyancy forcing in the convective regions, which can lead to changes in the thermohaline overturning circulation [Latif et al., 2006; Yeager and Robson, 2017], and anomalous geostrophic heat transport by the Gulf Stream. Changes in geostrophic heat transport can in turn have two types of underlying causes. First, changes in temperature,

caused by ocean-atmosphere heat exchange anomalies in distant regions and advected by the Gulf Stream to the Nordic Seas [Sutton and Allen, 1997; Furevik, 2001], and second, dynamic processes which change the baroclinic structure of the upper ocean and lead to shifts of the Gulf Stream path [Marshall et al., 2001] or modulation of the Gulf Stream or gyre strength [Häkkinen and Rhines, 2009]. Another remotely driven process is ocean wave dynamics [Johnson and Marshall, 2002], with associated fast-propagating barotropic transports [Fukumori et al., 2015], slower-propagating baroclinic transports [Jones et al., 2018], or a combination of the two [Orvik and Skagseth, 2003]. Understanding which of the above mechanisms and which remote regions are relevant for generating Nordic Seas upper-ocean heat content anomalies will be helpful - and necessary - for the design of ocean and climate observing systems.

2.2 Experimental setup

2.2.1 Model description and base state

The experiments in this chapter are performed with the MIT general circulation model (MITgcm; <http://mitgcm.org/>; [Marshall et al., 1997a,b]). The model configuration is adopted from the ECCO version 4 release 2 (ECCOv4r2, <http://www.ecco-group.org/>) framework, described in Forget et al. [2015]. This configuration includes a global model setup at a horizontal resolution of 1° with 50 vertical layers. The domain of version 4 spans the globe and - in contrast to previous ECCO releases - encompasses the Arctic, providing a more realistic setting for the high-latitude case study in this chapter. Unresolved processes associated with advection and mixing by eddies are parameterized following Gent and McWilliams [1990] and Redi [1982]. For the turbulent transport parameters, this study uses the optimized values from the ECCOv4r2 state estimate, which were obtained in a joint parameter and state estimation with a large set of altimetry and hydrographic observations serving as data constraints. Vertical mixing is parametrized using the turbulent kinetic energy scheme of Gaspar et al. [1990]. Convection is treated using simple convective adjustment.

To drive the ocean model, ECCOv4r2 uses 6-hourly ERA-Interim [Dee et al., 2011] near-surface fields of temperature, humidity, downward radiation, precipitation, and wind stress. These first-guess atmospheric input fields are subsequently corrected within the optimization process, to make them consistent with almost all available ocean observations [Forget et al., 2015]. For the adjoint sensitivity analysis in this chapter, the forward ocean model is initialized with the optimized ECCOv4r2 ocean state at 1 January, 2007, 00:00, and driven by an annually repeating cycle of atmospheric forcing. This annually repeating cycle consists of the optimized ECCOv4r2 (6-hourly) atmospheric forcing fields covering the year 2007, with minor modifications made to the precipitation fields, in order to avoid model drift. The year 2007 was chosen because it required less freshwater flux modifications than other years within the 20-year period covered by ECCOv4r2. Altering the freshwater flux is preferred to the frequently used alternative of sea surface salinity relaxation, because the latter would not only distort the forward model, but also adjoint sensitivities to surface freshwater fluxes in the model [Bugnion et al., 2006]. As in ECCOv4r2, the 6-hourly fluxes of heat and freshwater are deter-

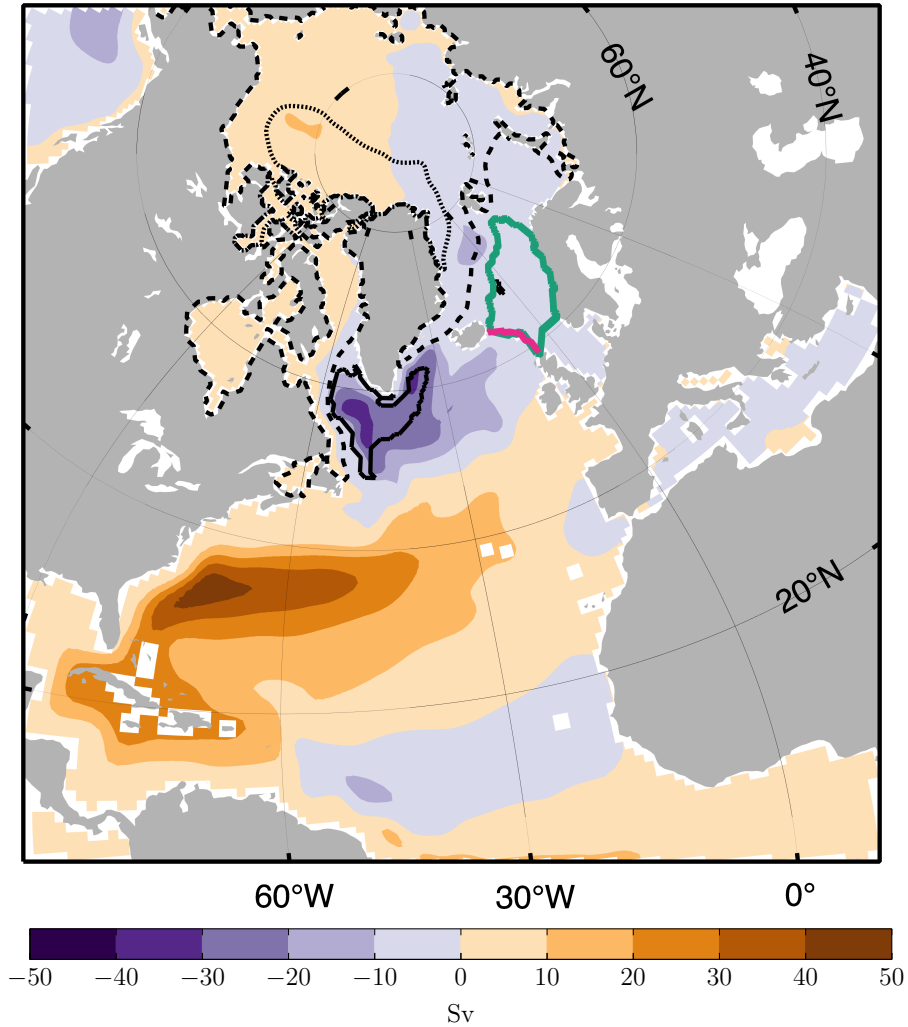


Figure 2.1: Mean barotropic stream function (color shading) for the equilibrated model simulation in this chapter. The black solid contour line bounds the regions wherein the March mixed layer depth exceeds 1000 m. The black dashed and dotted lines depict the sea-ice extent in March and September, respectively. The region inside the green contour line is defined as the Nordic Seas domain, and the pink line marks the section of the Iceland-Scotland ridge.

mined by the bulk formulae of Large and Yeager [2004], computed at every time step from the near-surface atmospheric input fields of temperature, humidity, downward radiation, and precipitation. Freshwater flux is treated as real, that is, it is exchanged through the free surface, such that surface freshwater fluxes change the ocean volume. Fig. 2.1 shows the resulting mean barotropic stream function for the model simulation in this chapter. It furthermore delineates the 1000 m March mixed layer depth contour, as well as March and September sea-ice extents.

The motivation for simplifying - or “equilibrating” - the ECCOv4r2 forward state for the adjoint sensitivity study in this chapter is to separate (i) perturbations of the ocean circulation that arise from oceanic adjustment mechanisms in response to surface forcing and (ii) “perturbations” of the ocean circulation that arise from strong interannual variability in the forward background state. Indeed, (ii) is eliminated in the equilibrated model simulation, as it is visible in Fig. 2.2. Each subpanel in Fig. 2.2 shows the evolution of an ocean quantity of interest, in the ECCOv4r2 (black) and equilibrated model simulation (red). Due to the removed interannual variability in the equilibrated simulation, sensitivities (e.g., to prior surface forcing) can be assumed to be stationary. For instance, the sensitivity of Nordic Seas upper-ocean temperature (Fig. 2.2(a)), defined in March of the final year 20 will be almost identical to the sensitivity of Nordic Seas upper-ocean temperature, defined in March of year 15.

2.2.2 Quantities of interest

The region of interest in this chapter is the eastern Nordic Seas, bounded by the green contour line in Fig. 2.1. This region - for simplicity referred to as the Nordic Seas (NS) for the remainder of this thesis - is the main gateway for poleward ocean heat transport between the North Atlantic and the Arctic Ocean (see Fig. 1.1). The domain is bounded in the south by the Iceland-Scotland ridge (pink line in Fig. 2.1) and the gateway between Scotland and Norway, in the east by the Norwegian coast, and in the north by the Barents Sea Opening. The western limit is chosen as the maximal occurring March sea-ice extent over the period 1992-2011 in the ECCOv4r2 state estimate. The mean climatological March mixed layer depth in the ECCOv4r2 state estimate, averaged over the Nordic Seas domain, is 275 m. In this chapter, upper-ocean heat content in the Nordic Seas is therefore defined as the mean temperature over the upper 275 m in the domain bounded by the green contour line in Fig. 2.1. The black line in Fig. 2.2(a) shows the evolution of Nordic Seas upper-ocean heat content in the ECCOv4r2 solution. Besides the seasonal cycle (black thin line), with summer temperatures that are about 2 K warmer than winter temperatures, one observes an interannual and decadal variability with temperature fluctuations on the order of 1 K (black thick line). This chapter investigates how local and remote atmospheric forcing anomalies contribute to this variability. In this case study, influences on *March* Nordic Seas upper-ocean heat content is studied, corresponding to the annual minima in Fig. 2.2(a), when the mixed layer depth is at its deepest. The quantity of interest is defined as

$$\text{UOHC}_{\text{NS}} = \frac{1}{\Delta t \cdot V_0} \int_{\text{March}} \int_{z=275 \text{ m}}^{\text{top}} \int_{\text{NS}} \theta \, dA \, dz \, dt, \quad (2.1)$$

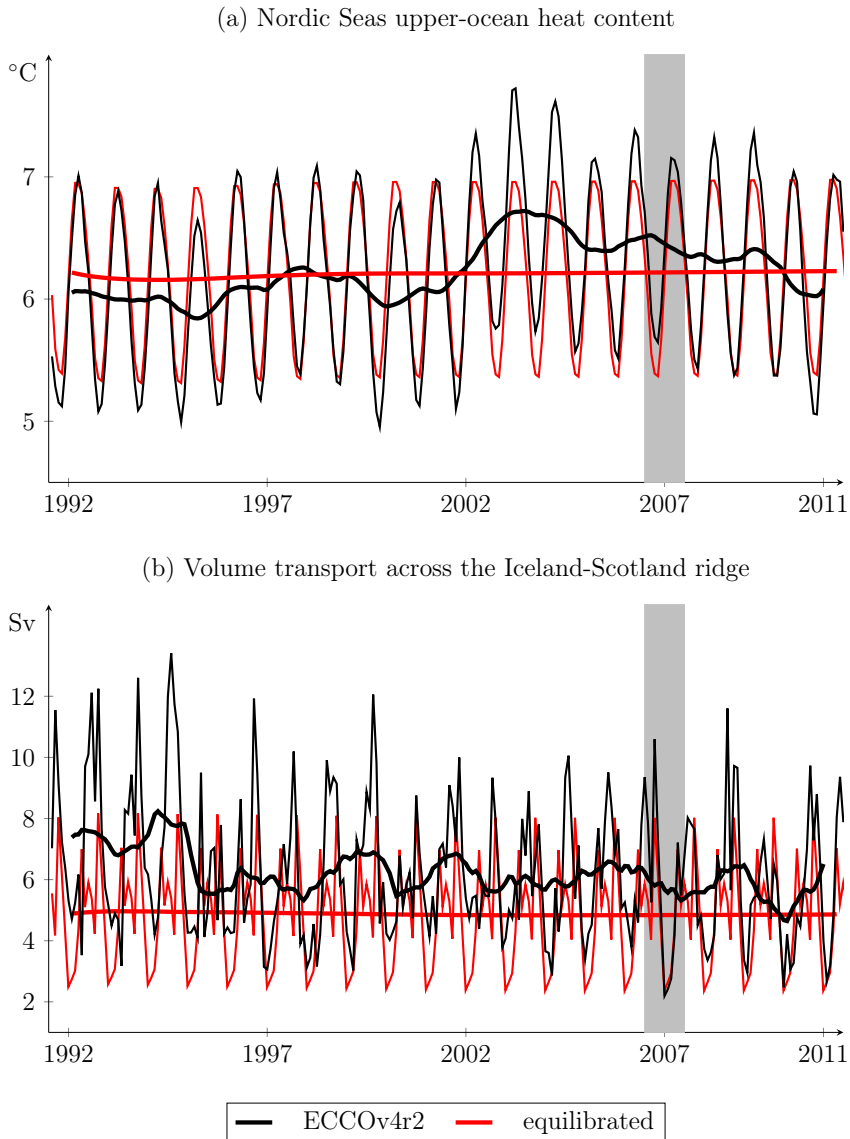


Figure 2.2: Time series of (a) Nordic Seas upper-ocean heat content and (b) volume transport across the Iceland-Scotland ridge, in the 20-year ECCOv4r2 state estimate from 1992 to 2011 (black), and in the equilibrated simulation without interannual variability (red). The optimized ECCOv4r2 forcing of the year 2007 (indicated by the gray shading) is used as an annually repeating cycle of atmospheric forcing to generate the equilibrated simulation. The quantities of interest are computed as (a) the mean temperature over the upper 275 m in the domain bounded by the green contour line in Fig. 2.2 (cf. equation (2.1)) and (b) north-eastward transport across the pink line in Fig. 2.2 (cf. equation (2.2)). The black and red *thin* lines show the respective monthly means, and the black and red *thick* lines the respective running annual means.

where θ is the potential temperature, $\Delta t = \int_{\text{March}} dt$ the time interval of interest, and $V_0 = \int_{z=275\text{ m}}^{\eta} \int_{\text{NS}} dA dz$ the volume of the region of interest.

Note that the quantity defined in (2.1) is not strictly upper-ocean heat content, but mean upper-ocean temperature. However, upper-ocean heat content and mean upper-ocean temperature differ only by a constant factor, namely by $\rho_0 c_p V_0$, the product of a reference density ρ_0 , the heat capacity c_p and the target volume V_0 . The results are chosen to be presented for mean upper-ocean temperature, as defined in (2.1), because units of changes in mean temperature (K) over the volume V_0 are found to be more intuitive than changes in total heat content (J).

The key gateway for heat inflow into the Nordic Seas is the Iceland-Scotland ridge (Fig. 1.1). Consequently, one expects the cross-ridge volume transport, the second quantity of interest, to be closely linked to Nordic Seas upper-ocean heat content. Volume transport across the ridge is defined as the integral over the entire water column and again taken as the March monthly mean:

$$T_{\text{IS}} = \frac{1}{\Delta t} \int_{\text{March}} \int_{\text{bottom}}^{\text{top}} \int_{\text{IS}} v^{\perp} dS dz dt. \quad (2.2)$$

Here, IS denotes the Iceland-Scotland model transect (pink line in Fig. 2.1) and v^{\perp} the velocity perpendicular to the Iceland-Scotland ridge. Sign convention is such that positive v^{\perp} corresponds to positive northward and eastward velocity. Therefore, positive T_{IS} indicates (positive) volume transport from the North Atlantic into the Nordic Seas basin.

2.2.3 Adjoint model and sensitivities

To examine local and remote oceanic processes that generate anomalies in the two quantities of interest that were defined in (2.1) and (2.2), this chapter performs two adjoint sensitivity analyses. Section 1.2 gives an overview of adjoint sensitivity analyses, and how these studies differ from traditional forward perturbation experiments (see Fig. 1.9). The method of computing adjoint sensitivities can be thought of as follows. One perturbs the model target metric, e.g., UOHC_{NS} , and then employs the adjoint model. The adjoint model integrates the linearized ocean physics backward in time, to compute where and when past surface heat, freshwater or momentum flux anomalies are capable of causing such a perturbation in UOHC_{NS} . This is distinct to a forward perturbation experiment, which would assess where and when the entire model state responds to a specific imposed surface flux perturbation. Running an ensemble of forward perturbation experiments is the conventional approach to model sensitivity analyses. However, to determine the effect of global surface buoyancy and momentum fluxes on UOHC_{NS} and T_{IS} in the same detail as with the adjoint approach would require millions of simulations with an ocean general circulation model, each representing a forward perturbation experiment with a surface flux perturbation at a different model grid point and time step. This is computationally prohibitive. In contrast, for each of the two quantities of interest, UOHC_{NS} and T_{IS} , the linearized adjoint sensitivities to surface flux anomalies at *all* model grid points and lead times can be computed in a single model integration.

The sensitivity analyses in this chapter are performed by employing the same adjoint modeling framework that is described by [Forget et al. \[2015\]](#) and has been used to obtain the ECCOV4 state estimate. Algorithmic differentiation, through source-to-source code transformation with the algorithmic differentiation tool transformation of algorithms in Fortran (TAF; [Giering and Kaminski, 1998](#)), produces the code for the adjoint models. The adjoint model computes the linearized sensitivities of $UOHC_{NS}$ and T_{IS} to past anomalies in surface buoyancy and momentum fluxes at every model grid point and any time step. In mathematical terms, the adjoint sensitivities are the partial derivatives

$$\frac{\partial UOHC_{NS}}{\partial Q_{net}(x, y, -t)}, \quad \frac{\partial UOHC_{NS}}{\partial (E-P-R)(x, y, -t)}, \quad \frac{\partial UOHC_{NS}}{\partial \tau_x(x, y, -t)}, \quad \frac{\partial UOHC_{NS}}{\partial \tau_y(x, y, -t)}, \quad (2.3)$$

$$\frac{\partial T_{IS}}{\partial Q_{net}(x, y, -t)}, \quad \frac{\partial T_{IS}}{\partial (E-P-R)(x, y, -t)}, \quad \frac{\partial T_{IS}}{\partial \tau_x(x, y, -t)} \quad \text{and} \quad \frac{\partial T_{IS}}{\partial \tau_y(x, y, -t)}, \quad (2.4)$$

around the time-dependent equilibrated model simulation, described in Section 2.2.1. Here, $Q_{net}(x, y, -t)$, $E-P-R(x, y, -t)$, $\tau_x(x, y, -t)$ and $\tau_y(x, y, -t)$ denote surface heat flux, surface freshwater flux, zonal and meridional wind stress, respectively, at model grid point (x, y) and lead time t , where t is discretized into months. Q_{net} includes short-wave, longwave, sensible, and latent heat fluxes.

Recall from Section 2.2.1 that heat and freshwater fluxes are computed by bulk formulae, rather than being applied as input forcings directly. As a result, the variables Q_{net} and $E-P-R$ are not independent. For example, a heat flux anomaly from the atmosphere into the ocean warms the ocean locally, and consequently, the bulk formulae will determine an increase in evaporation and thus a change in freshwater flux. Another effect of employing bulk formulae for the computation of buoyancy fluxes is that heat flux-driven SST anomalies are damped almost immediately because the bulk formulae compute heat exchange based on the difference of the atmospheric and the (altered) ocean temperature. On the contrary, this effect is not present for freshwater flux. Even though the usage of bulk formulae encompasses a more complex interplay between variables, it is preferred over the less realistic formulations of boundary flux forcing or surface restoring. For a more detailed discussion on how the formulation of surface forcing terms in ocean models impacts adjoint sensitivities, the reader is referred to [Bugnion et al. \[2006\]](#) and [Kostov et al. \[2019\]](#).

Examining how the adjoint sensitivities (2.3) and (2.4) evolve through time and space reveals physical mechanisms and key regions that impact the quantities of interest at various timescales. As already stressed in Section 1.2, the approach of using adjoint sensitivities stands in sharp contrast to statistical methods, such as correlations or regression. Adjoint sensitivities identify physical relationships and causal chains contained in the model equations (which in turn only approximately represent the real ocean). On the other hand, correlations provide an empirical measure of how two variables vary concurrently and/or at some lead/lag, regardless of whether or not they are causally related (see Fig. 1.11).

A caveat of the adjoint approach is that adjoint-derived sensitivities are a linearized approximation to the non-linear responses, and may become less and less accurate with larger perturbations and longer model integrations. Moreover, discontinuous processes,

e.g., due to regime transition, can cause instability in the adjoint model and often have to be substituted by approximate adjoints [e.g., Forget et al., 2015]. Due to these limitations, the linearized approximation and inexactness of the adjoint, the adjoint-derived sensitivities are to be evaluated against (probe) non-linear perturbation experiments in order to assess their applicability. The acceptable size of the input perturbation and considered timescale will depend on the studied quantity of interest and the model setup. For instance, Czeschel et al. [2010] find that in their MITgcm configuration at 1° horizontal resolution, the adjoint and non-linear forward models show broadly consistent results for the evolution of the AMOC at 27°N over the first 15-20 years, after applying basin-wide heat flux perturbations of amplitude 15 W/m^2 over the subpolar gyre. In Chapter 5, the validity of the linearity assumption for the case study in this chapter will be assessed, by evaluating the adjoint-derived sensitivities against forward model perturbation experiments.

2.3 Identifying adjustment mechanisms and pathways

This section presents sensitivity fields $\frac{\partial \text{UOHC}_{\text{NS}}}{\partial F}(x, y, -t)$ and $\frac{\partial \text{T}_{\text{IS}}}{\partial F}(x, y, -t)$. UOHC_{NS} and T_{IS} are March Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge, defined in (2.1) and (2.2), respectively. Sensitivities are computed to surface buoyancy and momentum fluxes F at lead times t of 1 month, 1 year, 3 years and 8 years. A lead time of 1 month refers to monthly mean flux anomalies that have taken place in the previous month, i.e., in February, a lead time of 1 year to monthly mean flux anomalies that have occurred in March in the previous year, and so on. The sensitivities are normalized to a forcing anomaly applied over an area of 1 m^2 for one month. All fluxes are defined to act at the ocean surface. That is, for sea-ice covered areas the flux is defined at the ice-ocean interface.

The distribution of positive and negative sensitivities for increasing lead times reveals all mechanisms by which small-amplitude perturbations in local or remote surface forcing at the considered lead times may impact March Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge. Note that for each quantity of interest, sensitivity amplitudes *cannot* be compared across different forcings, since they are computed per forcing unit. In Section 2.4 sensitivities are weighted by typical buoyancy and momentum flux anomalies, which will make a comparison of UOHC_{NS} variability driven by different external forcings possible. The aim of the current section, however, is the identification of physical mechanisms and teleconnections, which can be performed for each forcing separately.

2.3.1 Sensitivity to buoyancy forcing

First, sensitivities of Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge to surface buoyancy fluxes are investigated. By convention, the buoyancy fluxes are directed upwards, that is, positive surface heat flux Q_{net} corresponds to ocean cooling, and positive freshwater flux E-P-R to increased evaporation and thus ocean salinification.

Nordic Seas upper-ocean heat content

Figs. 2.3 and 2.4 present snapshots of Nordic Seas upper-ocean heat content sensitivity to net surface heat flux Q_{net} and freshwater flux E-P-R, respectively. The logarithmic contour intervals are useful to detect mechanisms acting on different timescales, which drive anomalies in Nordic Seas upper-ocean heat content of different amplitude. While two logarithmic decades are used for sensitivities to freshwater flux (Fig. 2.4), three logarithmic decades are shown for sensitivities to heat flux (Fig. 2.3) to illustrate the rapid loss of sensitivity to heat flux. The sensitivities are normalized to a forcing anomaly applied over an area of 1 m^2 for one month. For example in Fig. 2.3(a), the strong sensitivity over the Nordic Seas basin is approximately uniform of order -10^{-15} K/W . A uniform month-long increase in upward surface heat flux of 1 W/m^2 across the $1.5 \cdot 10^{12} \text{ m}^2$ -sized Nordic Seas basin would therefore result in a monthly anomaly in UOHC_{NS} of $1.5 \cdot 10^{-3} \text{ K}$ in the next month.

- 1 month At a lead time of 1 month, a signature of strong negative sensitivity to upward surface heat flux is visible over the Nordic Seas as well as upstream in the NAC (Fig. 2.3(a)). This indicates that an anomalous surface heat flux directed out of the ocean over these regions results in negative upper-ocean heat content anomalies in the Nordic Seas 1 month later, as one would expect. This mechanism reflects the dominant effect of *local air-sea heat fluxes* on short timescales. In contrast, the amplitude of sensitivities to freshwater flux at a lead time of 1 month (Fig. 2.4(a)) is small - an order of magnitude less than at longer lead times of 1 or 3 years (Figs. 2.4(b),(c)). The weak negative sensitivities to freshwater flux in the Nordic Seas region (Fig. 2.4(a)) are due to the fact that, in the model, freshwater fluxes carry the local SST. An increased freshwater flux, directed out of the ocean, therefore removes waters that are relatively warmer than the temperature mean taken over the upper 275 m in the Nordic Seas basin. Positive sensitivities to freshwater flux that emerge around Iceland are visible at all lead times, and will be discussed below.

- 1 year Negative sensitivities to heat flux in the Nordic Seas region at a lead time of 1 year (Fig. 2.3(b)) show that local air-sea fluxes can impact the Nordic Seas upper-ocean heat content 1 year later, although the effect is of an order of magnitude less than the immediate response (Fig. 2.3(a)). The fact that heat flux anomalies from the previous winter still have an impact is due to the *reemergence mechanism* [Alexander and Deser, 1995], which stores local temperature anomalies at depth and then reentrains them into the mixed layer the following winter.

The sensitivity dipole along the main flow line of the NAC, upstream of the Iceland-Scotland ridge, in the map of sensitivity to freshwater forcing at a lead time of 1 year (Fig. 2.4(b)) can be explained by the dynamics of *thermal wind balance* altering the geostrophic current (u_g, v_g) via

$$\frac{\partial u_g}{\partial z} = \frac{g}{\rho_0 f} \frac{\partial \rho}{\partial y}, \quad \frac{\partial v_g}{\partial z} = -\frac{g}{\rho_0 f} \frac{\partial \rho}{\partial x}. \quad (2.5)$$

A freshwater flux perturbation matching the displayed distribution in Fig. 2.4(b), i.e., net

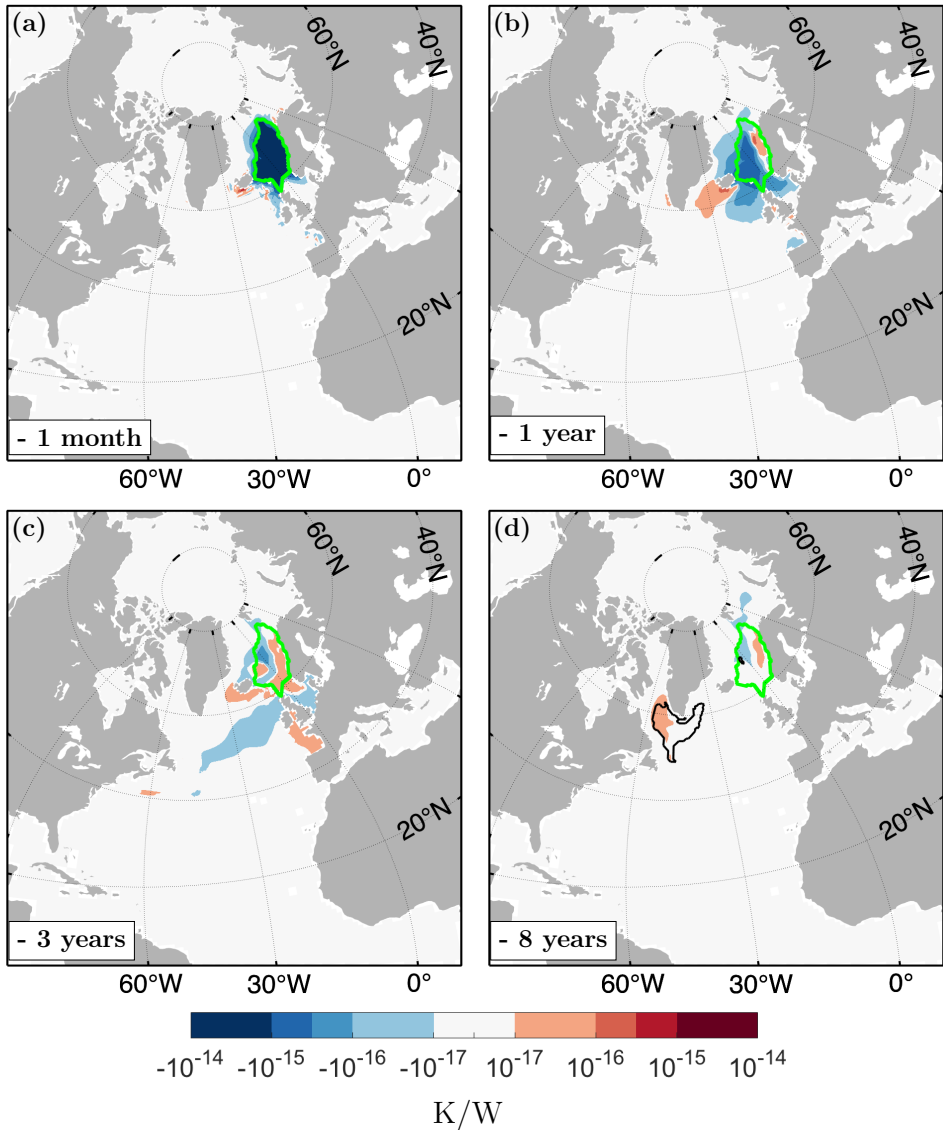


Figure 2.3: Linearized sensitivity $\frac{\partial \text{UOHC}_{\text{NS}}}{\partial Q_{\text{net}}(x,y,-t)}$ of March Nordic Seas upper-ocean heat content UOHC_{NS} to surface heat flux Q_{net} for increasing lead times t of (a) 1 month, (b) 1 year, (c) 3 years, and (d) 8 years. Positive sensitivity indicates that heat loss to the atmosphere at the indicated lead time causes higher Nordic Seas upper-ocean heat content. The green contour line marks the border of the Nordic Seas basin. The black contour line in (d) bounds the regions wherein the March mixed layer depth exceeds 1000 m.

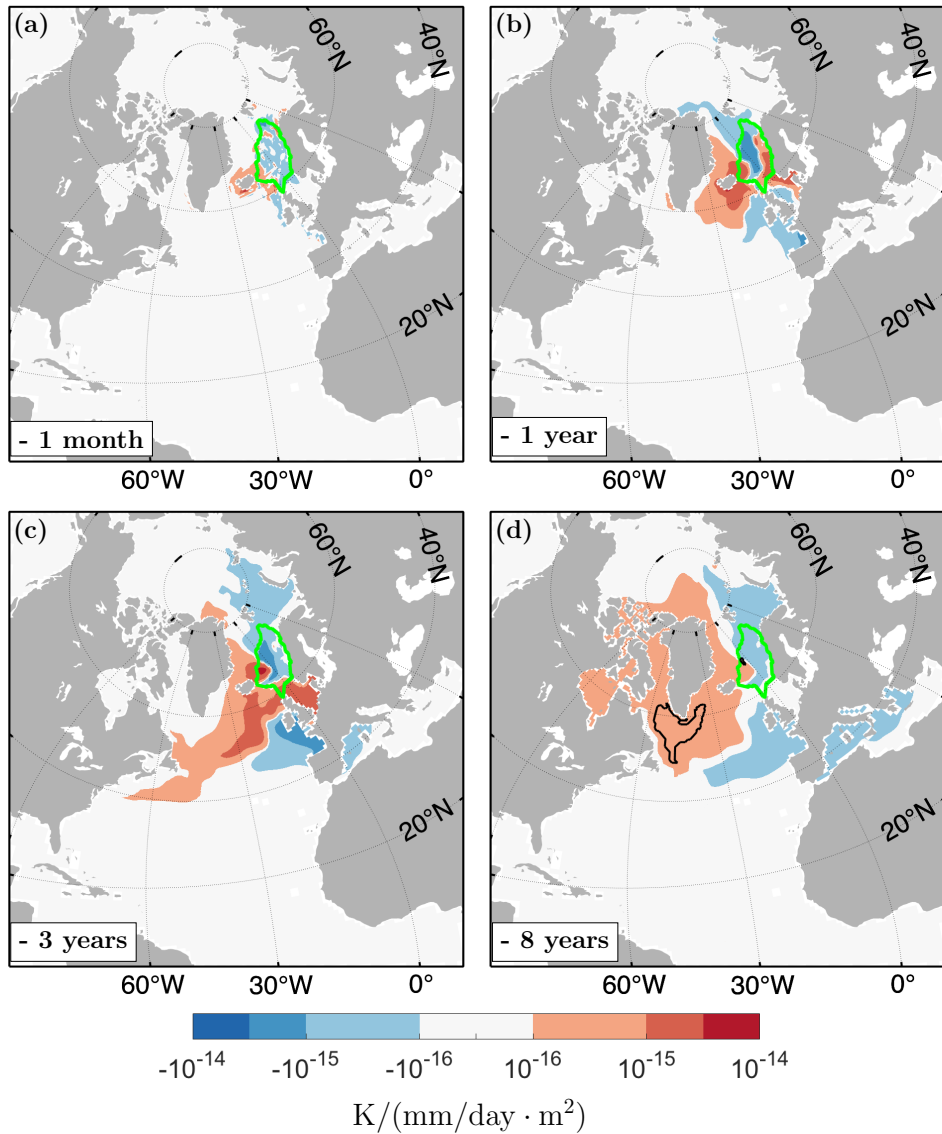


Figure 2.4: As in Fig. 2.3 but linearized sensitivity $\frac{\partial \text{UOHC}_{\text{NS}}}{\partial (\text{E-P-R})(x,y,-t)}$ of March **Nordic Seas upper-ocean heat content** to **surface freshwater flux E-P-R** for increasing lead times of (a) 1 month, (b) 1 year, (c) 3 years, and (d) 8 years. Positive sensitivity indicates that an increase in monthly mean evaporation at the indicated lead time causes higher Nordic Seas upper-ocean heat content. The green and black contour lines are the same as in Fig. 2.3.

surface salinification to the left and net surface freshening to the right of the main flow line of the NAC, generates a negative anomaly in the upper-ocean density gradient across the NAC. According to equation (2.5), this negative density gradient anomaly strengthens the geostrophic current along the NAC path, with more heat being transported to the Nordic Seas. The same mechanism is responsible for the positive sensitivities to surface heat flux south of Iceland (Fig. 2.3(b)) because a positive upward heat flux anomaly (corresponding to an anomalous heat loss to the atmosphere) south of Iceland would lead to an ocean densification in that region and again create an anomalous negative zonal density gradient across the NAC.

Positive sensitivities to upward heat flux along the Norwegian coast (Figs. 2.3(b)-(d)) seem counter-intuitive because this feature indicates that a local heat loss from the ocean to the atmosphere increases upper-ocean heat content of the Nordic Seas 1-8 years later. However, the same characteristic is seen in the freshwater flux sensitivity map at a lead time of 1 year (Fig. 2.4(b)), which hints toward a dynamic process changing the baroclinic structure of the upper ocean. Buoyancy loss in the region of positive sensitivities off the Norwegian coast (Figs. 2.3(b) and 2.4(b)) creates a positive anomaly in the upper-ocean density gradient across the Norwegian Coastal Current (NwCC), which is a second source water pathway for the Nordic Seas (Fig. 1.1) and carries slightly colder water than the NAC. By thermal wind balance (2.5), the positive density gradient decreases the strength of the NwCC, and thus the relative contribution of warmer NAC waters brought to the Nordic Seas increases. This interpretation is supported by the fact that the negative-positive dipole straddles the NwCC all the way back into the North Sea in the sensitivity map to freshwater flux (Fig. 2.4(b)). However, the extension of the dipole into the North Sea is not seen in the sensitivity pattern to heat flux (Fig. 2.3(b)) because air-sea heat fluxes act not only as a forcing for the described dynamic thermal wind balance mechanism but also via the one-dimensional process of pumping heat into the ocean, where the latter seems to win the competition. The warm temperature anomalies, caused by the one-dimensional process, are then advected with the NwCC from the North Sea into the Nordic Seas, which is why negative sensitivities, corresponding to ocean heat gain from the atmosphere, are visible in the North Sea (Fig. 2.3(b)).

- 3 years With increasing forcing lead time, Nordic Seas upper-ocean heat content sensitivity emerges continuously farther upstream along the NAC/Gulf Stream path. At a lead time of about 2.5 - 3 years, the sensitivities have reached the American east coast near Cape Hatteras (Figs. 2.3(c) and 2.4(c)). Negative sensitivities to heat flux extend into the NAC and eastern part of the Gulf Stream (Fig. 2.3(c)), and indicate that an anomalous heat flux from the atmosphere into these ocean currents trigger positive anomalies in upper-ocean heat content in the Nordic Seas 3 years later. The responsible mechanism is advection carrying heat flux-driven positive temperature anomalies via the Gulf Stream and NAC into the Nordic Seas.

The sensitivity dipole along the main flow line of the NAC in the freshwater flux sensitivity map (Fig. 2.4(c)) describes a buoyancy-forced increase in baroclinicity and in the associated sheared transport across the core of the NAC. The stronger NAC subsequently advects more warm waters into the Nordic Seas. In fact, the positive-negative dipole extends all the way across the Fram Strait into the *Arctic basin*. This

suggests that thermal wind balance can be an effective mechanism even downstream of the Nordic Seas, by “sucking” more Atlantic water into the Nordic Seas. In contrast to freshwater flux anomalies, the Arctic Ocean has no notable sensitivity to surface heat flux anomalies that occur at a lead time of 3 years (Fig. 2.3(c)).

- 8 years At a forcing lead time of 8 years, sensitivity distributions to heat and freshwater flux show a fingerprint of the Atlantic *thermohaline circulation* (THC). Positive sensitivities to heat flux in the Labrador Sea (Fig. 2.3(d)) and to freshwater flux in the entire subpolar gyre (Fig. 2.4(d)) indicate that feeding the deep convection regions with denser waters - either by removing freshwater or by ocean cooling - intensifies the THC, which leads to an increased heat transport to the Nordic Seas and thus higher Nordic Seas upper-ocean heat content 8 years later.

The sensitivity to freshwater flux persists in straddling the main flow line of the NAC (Fig. 2.4(d)), indicating that the *thermal wind balance* mechanism is still a player in remotely driving upper-ocean heat content anomalies in the Nordic Seas on longer timescales. Moreover, the sensitivities to freshwater flux in the Arctic basin are still of notable amplitude at a lead time of 8 years (Fig. 2.4(d)). The sensitivity distribution in the *Arctic* is striking. The Eurasian basin shows negative sensitivities, whereas the Canadian Basin exhibits positive sensitivities. Positive sensitivities in the Canadian Basin can be seen as an expansion of positive sensitivities in the entire subpolar gyre basin: positive salinity anomalies will eventually be exported from the Arctic Ocean to the deep convection regions, passing the Fram Strait and Denmark Strait. As opposed to the positive sensitivities to freshwater flux that populate the entire subpolar gyre and its upstream regions in the Arctic (Fig. 2.4(d)), positive sensitivity to surface heat flux is concentrated at the deep convection region in the Labrador Sea (Fig. 2.3(d)). This remarkable discrepancy is due to the fact that SST anomalies are damped by air-sea fluxes, whereas sea surface salinity (SSS) anomalies are not subject to a similar damping effect. SSS anomalies can therefore persist over longer timescales.

Volume transport across the Iceland-Scotland ridge

Figs. 2.5 and 2.6 display sensitivities of volume transport across the Iceland-Scotland ridge to net surface heat flux Q_{net} and freshwater flux E-P-R, respectively.

- 1 month The sensitivity patterns to heat and freshwater flux at a forcing lead time of 1 month show two striking features.

The first feature is a concentration of high sensitivity around the Iceland-Scotland ridge (Figs. 2.5(a) and 2.6(a)), with positive sensitivity to the west of the ridge (around Iceland), negative sensitivity to the east and south of the ridge (along the west coast of Great Britain), and positive-negative oscillating sensitivities in between. The responsible mechanism is *thermal wind balance* via equation (2.5). Heat and freshwater flux perturbations matching the sensitivity distribution at the end points of the ridge generate an anomalous upper ocean negative density gradient along the Iceland-Scotland ridge, which is then accompanied by a strengthening of the north-eastward transport. The complex positive-negative structure between the end points of the Iceland-Scotland

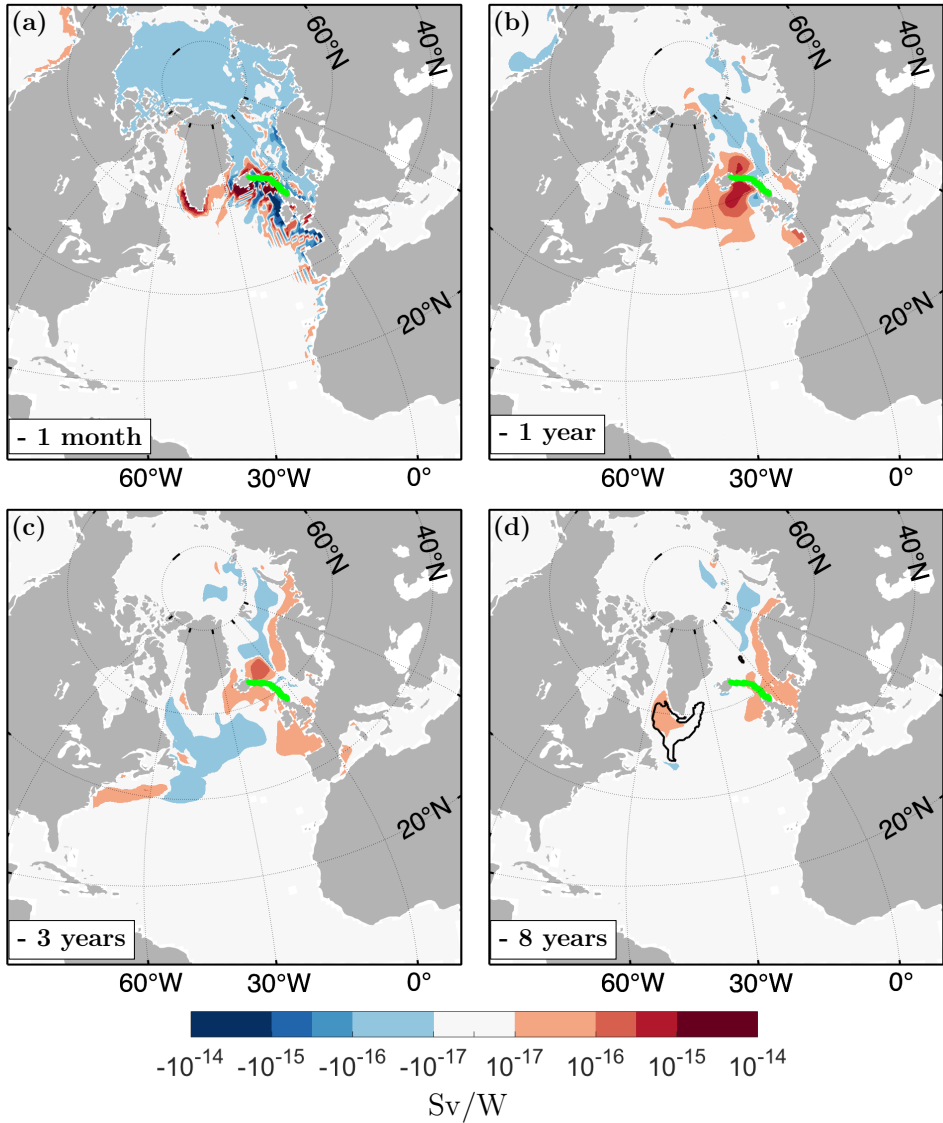


Figure 2.5: As Fig. 2.3 but linearized sensitivity $\frac{\partial T_{\text{IS}}}{\partial Q_{\text{net}}(x,y,-t)}$ of March **volume transport across the Iceland-Scotland ridge** to **surface heat flux Q_{net}** for increasing lead times of (a) 1 month, (b) 1 year, (c) 3 years, and (d) 8 years. Positive sensitivity indicates that heat loss to the atmosphere at the indicated lead time causes larger north-eastward volume transport across the ridge. The green contour line marks the section of the Iceland-Scotland ridge, and the black contour line in (d) is the same as in Fig. 2.3(d).

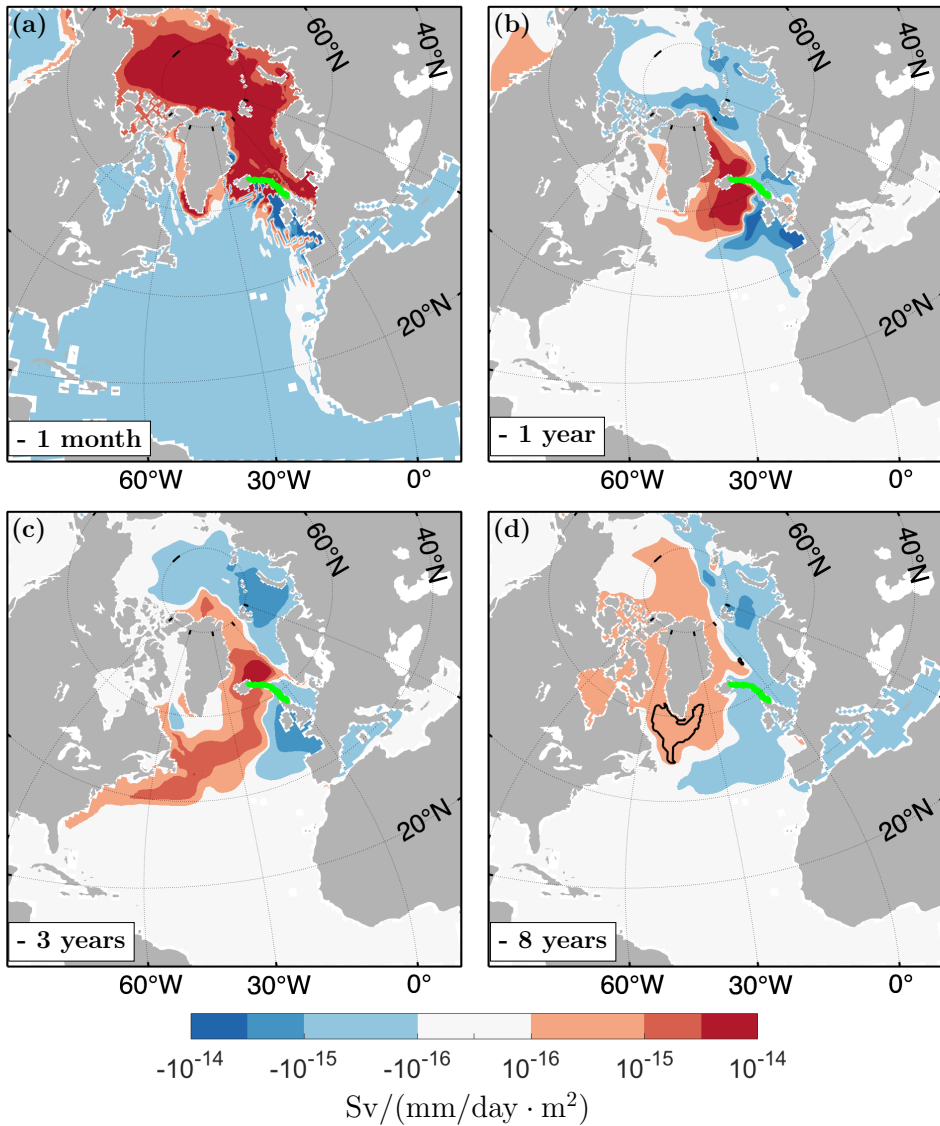


Figure 2.6: As Fig. 2.5 but linearized sensitivity $\frac{\partial T_{IS}}{\partial (E-P-R)(x,y,-t)}$ of March **volume transport across the Iceland-Scotland ridge** to **surface freshwater flux E-P-R** for increasing lead times of (a) 1 month, (b) 1 year, (c) 3 years, and (d) 8 years. Positive sensitivity indicates that an increase in monthly mean evaporation at the indicated lead time causes larger north-eastward volume transport across the ridge. The green and black contour lines are the same as in Fig. 2.5.

ridge is due to the fact that the ridge is not a straight, zonally aligned section - neither in the real ocean nor in the model configuration used. Oscillatory sensitivities extend also to the south of the ridge in Figs. 2.5(a) and 2.6(a) and are topographically steered, due to the communication of barotropic signals along contours of constant f/H .

The second feature is of global nature. Positive sensitivities to freshwater flux populate the Greenland, Iceland, Norwegian (GIN) Seas and the Arctic, and negative sensitivities to freshwater flux populate the remainder of the global ocean (Fig. 2.6(a)). Similarly, basin-wide sensitivities to heat flux emerge in the GIN Seas and the Arctic, but, as opposed to sensitivities to freshwater flux, they carry a negative sign (Fig. 2.5(a)). The signature of basin-wide sensitivities to freshwater flux can be explained as follows. Freshwater flux perturbations lead to a change in volume. Barotropic waves communicated this signal rapidly through the interior, into the coastal waveguides. Subsequently, Kelvin waves propagate the signal in the wave guides cyclonically around the basin, to the Iceland-Scotland ridge. In the Northern Hemisphere, Kelvin waves travel with the coast to their right. Thus, when volume is removed from the Arctic or GIN Seas by increased evaporation, this signal arrives at the western end of the Iceland-Scotland ridge section, which generates a positive perturbation in the zonal pressure gradient along the Iceland-Scotland ridge section. By *geostrophic balance*, this leads to a strengthening of the northward geostrophic transport v_g :

$$v_g = \frac{1}{\rho_0 f} \frac{\partial p}{\partial x}. \quad (2.6)$$

This explains why sensitivity in the Arctic and GIN Seas is of uniformly positive sign. In contrast, if volume is removed from the rest of the ocean, this signal arrives at the eastern end of the ridge, which generates a negative perturbation in the zonal pressure gradient along the ridge. According to equation (2.6), this weakens the northward geostrophic transport v_g , which explains why sensitivity in ocean regions south of the ridge is of uniformly negative sign.

As opposed to freshwater flux perturbations, heat flux perturbations do not have a direct effect on ocean volume in the model. (Volume changes due to thermal expansion driven by surface heat fluxes are negligible in comparison to volume changes due to freshwater input [Greatbatch, 1994]. Thermal expansion due to surface heat fluxes is therefore neglected in the model.) Instead, the basin-wide negative sensitivities in the Arctic and GIN Seas are due to an indirect effect, which is linked to the interplay of heat and freshwater fluxes that was explained in Section 2.2.3: negative heat flux anomalies entail positive freshwater flux anomalies because a warmer ocean leads to increased evaporation. The negative sensitivity to heat flux in the Arctic and GIN Seas is therefore imprinted by the positive sensitivity to freshwater flux in the same region. Due to the *indirect* effect, the sensitivity to heat flux caused by fast barotropic adjustment is of much smaller amplitude than sensitivity to freshwater flux (e.g., when seen relative to local sensitivity close to the ridge). Sensitivity to heat flux does therefore not emerge south of the section under the colorbar used.

- 1 year Sensitivities to heat and freshwater flux at a lead time of 1 year (Figs. 2.5(b) and 2.6(b)) highly resemble each other. Moreover, a very similar pattern has been

found in the sensitivity map of UOHC_{NS} to freshwater flux for the same lead time (Fig. 2.4(b)). The resemblance of these three sensitivity maps supports the previous interpretation for freshwater flux sensitivities of UOHC_{NS} : a strengthening of the surface geostrophic transport in the NAC advects more heat to the Nordic Seas. The only noticeable discrepancy between freshwater flux sensitivities of T_{IS} (2.6(b)) and UOHC_{NS} (Fig. 2.4(b)) is that the dipole across the NwCC, detected in the sensitivity pattern of UOHC_{NS} (Fig. 2.4(b)) and attributed to an alteration of the NwCC as a second source water pathway for the Nordic Seas, is absent in the sensitivity of T_{IS} (Fig. 2.6(b)), as expected.

-3 years The sensitivity to freshwater flux at a lead time of 3 years (Fig. 2.6(c)) shows similarities to those seen for UOHC_{NS} at the same lead time (Fig. 2.4(c)): a dipole across the Gulf Stream and NAC, which extends all the way into the GIN Seas and the Arctic basin, impacting the strength of the NAC via thermal wind balance. However, sensitivity to heat flux (Fig. 2.5(c)) shows a different pattern. This might appear counter-intuitive because both heat and freshwater flux can only change volume transport through a change in the density structure, and are therefore expected to have identical sensitivity patterns - as was the case for a lead time of 1 year. I hypothesize that this discrepancy is due to the fact that negative (positive) heat flux anomalies do not only impact density directly, via ocean warming (cooling), but also indirectly, due to increased (decreased) evaporation and thus ocean salinification (freshening), causing density anomalies of opposite sign. Therefore, a surface heat flux anomaly in the Gulf Stream or NAC creates both an SST *and* an SSS anomaly. These anomalies are advected toward the ridge and have competing effects on the baroclinic structure of the current: one anomaly strengthens the current, while the other anomaly weakens it. Along the way, the SST anomaly is damped by air-sea heat fluxes, while the SSS anomaly can persist. Consequently, the effect of the SSS anomaly overrides the effect of the SST anomaly on the long run. This is reflected by the fact that subregions of the NAC close to the ridge show positive sensitivity in Fig. 2.5(c), indicating that density changes due to SST anomalies dominate the response, while regions farther upstream from the ridge show negative sensitivity, indicating that density changes due to SSS anomalies control the response. The full complexity of this process, including its seasonal dependence on the occurrence of the forcing anomaly, is studied in Kostov et al. [2019].

-8 years For a lead time of 8 years, the sensitivities of T_{IS} to heat and freshwater flux are identical to those for UOHC_{NS} . This shows that for a forcing lead time of 8 years both T_{IS} and UOHC_{NS} can be increased most effectively by (i) a strengthened overturning circulation via heat flux-driven ocean cooling in the deep convection regions (Figs. 2.3(d) and 2.5(d)) and freshwater flux-driven positive SSS anomalies in the entire subpolar gyre, Greenland Sea and Arctic Ocean (Figs. 2.4(d) and 2.6(d)), and (ii) a strengthened horizontal circulation induced by an anomalous salinity gradient across the main flow line of the NAC (Figs. 2.4(d) and 2.6(d)).

Further discussion on the link between UOHC_{NS} and T_{IS} sensitivities to surface buoyancy fluxes will be given in Section 2.5.

2.3.2 Sensitivity to surface momentum fluxes

Next, sensitivities of Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge to surface momentum fluxes are investigated. By convention, positive zonal wind stress τ_x corresponds to eastward wind, and positive meridional wind stress τ_y to northward wind.

Nordic Seas Upper-Ocean Temperature

Figs. 2.7 and 2.8 show snapshots of sensitivity of Nordic Seas upper-ocean heat content to zonal and meridional wind stress, τ_x and τ_y , at lead times of 1 month, 1 year, 3 years and 8 years. The structure of the sensitivity patterns exhibits many complex small-scale characteristics, due to the importance of wave dynamics, the impact of topographic boundaries, and local Ekman dynamics. For a clearer discussion, the paragraphs in the following are therefore ordered by characteristics of the sensitivity patterns, rather than by lead time (as done in Section 2.3.1).

Moderate loss of sensitivity Wind stress anomalies at a lead time of 1 year (Figs. 2.7(b) and 2.8(b)) can trigger anomalies in Nordic Seas upper-ocean heat content of the same order of magnitude as wind stress anomalies at a lead time of 1 month (Figs. 2.7(a) and 2.8(a)). Compared to the rapid loss of sensitivity to heat flux by two orders of magnitude over the course of 8 years (Fig. 2.3), the loss of sensitivity to historic wind stress is moderate. For wind stress anomalies at a lead time of 8 years (Figs. 2.7(d) and 2.8(d)), one can still detect sensitivities that are only of an order of magnitude less than sensitivities to wind stress anomalies at a lead time of 1 month (Figs. 2.7(a) and 2.8(a)).

Stationarity Moreover, it is striking that many features in the sensitivity patterns appear stationary, persisting for lead times of months to years. Two examples of such stationary sensitivity features are highlighted by the solid green boxes in Figs. 2.7 and 2.8:

- S.1** The sensitivity feature around Iceland. Sensitivity to zonal wind stress is positive along the southern coast of Iceland, and negative along the northern coast of Iceland. Sensitivity to meridional wind stress is negative along the western coast of Iceland and positive along the eastern coast of Iceland.
- S.2** The sensitivity feature at the eastern edge of the Atlantic. A band of positive sensitivities to meridional wind stress exists along the African and European west coast, crossing the Iceland-Scotland ridge, and extending into the Nordic Seas basin. The sign of sensitivity to zonal wind stress alternates and is determined by the orientation of the West African coastline. NW-SE-aligned coastlines show negative sensitivities adjacent to the coast, and SW-NE-aligned coastlines show positive sensitivities to zonal wind stress.

Sensitivity dipole along intergyre boundary Not all features in the sensitivity patterns to wind stress are stationary. With increasing lead time, a sensitivity dipole emerges which straddles the zero-line of the barotropic stream function. The sensitivity dipole is highlighted by the dashed green box in Figs. 2.7(c),(d) and 2.8(c),(d) and is characterized by

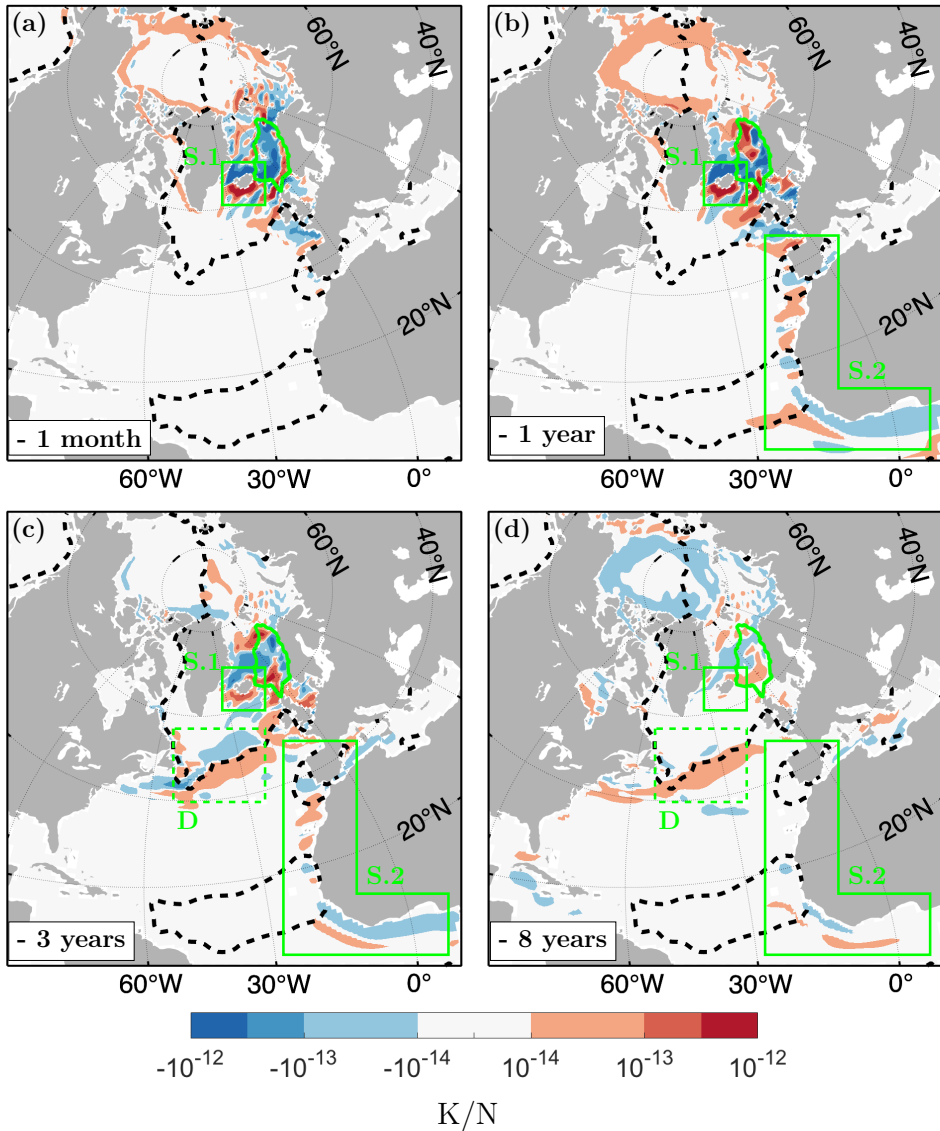


Figure 2.7: As Fig. 2.3 but linearized sensitivity $\frac{\partial \text{UOHC}_{\text{NS}}}{\partial \tau_x(-t, x, y)}$ of March Nordic Seas upper-ocean heat content to zonal wind stress τ_x for increasing lead times of (a) 1 month, (b) 1 year, (c) 3 years, and (d) 8 years. Positive sensitivity indicates that an increase in monthly mean eastward wind at the indicated lead time causes higher Nordic Seas upper-ocean heat content. The green contour line marks the border of the Nordic Seas basin and the black dashed contour line shows the zero-line of the barotropic stream function. The solid green boxes S.1 and S.2 contain stationary sensitivity patterns. The dashed green box D contains a sensitivity dipole along the intergyre boundary.

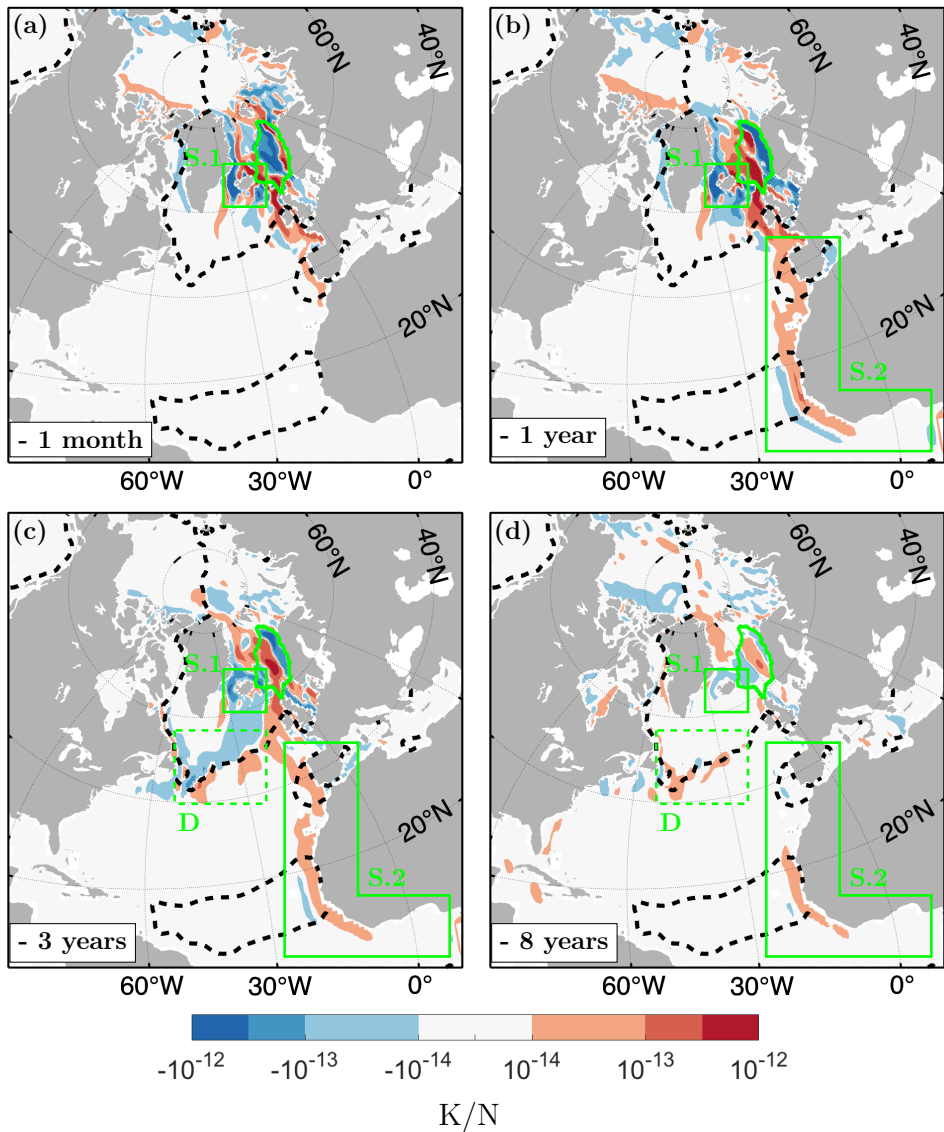


Figure 2.8: As Fig. 2.7 but linearized sensitivity $\frac{\partial \text{UOHC}_{\text{NS}}}{\partial \tau_y(-t, x, y)}$ of March Nordic Seas upper-ocean heat content to meridional wind stress τ_y for increasing lead times of (a) 1 month, (b) 1 year, (c) 3 years, and (d) 8 years. Positive sensitivity indicates that an increase in monthly mean northward wind at the indicated lead time causes higher Nordic Seas upper-ocean heat content. The green and black contour lines and boxes are the same as in Fig. 2.7.

negative sensitivities to zonal and meridional wind stress in the subpolar gyre, and positive sensitivities to zonal and meridional wind stress in the subtropical gyre. Therefore the sign of the dipole is such that a positive anomaly in wind stress curl along the intergyre boundary leads to a positive anomaly in Nordic Seas upper-ocean heat content. For a lead time of 2 years (not shown), the sensitivity dipole populates the NAC eastward of 40°W , and by a lead time of 2.5 - 3 years the dipole has reached the American east coast near Cape Hatteras (Figs. 2.7(c) and 2.8(c)).

The above findings - moderate loss of Nordic Seas upper-ocean heat content sensitivity to prior wind stress anomalies, stationarity, and the sensitivity dipole along the intergyre boundary - will be discussed in more detail in Section 2.5, after investigating these aspects for volume transport across the Iceland-Scotland ridge.

Volume transport across the Iceland-Scotland ridge

Fig. 2.9 shows snapshots of sensitivity of volume transport across the Iceland-Scotland ridge to wind stress.

- 1 month The patterns S.1 and S.2 that were detected in Figs. 2.7 and 2.8 also appear in the sensitivity maps of volume transport across the Iceland-Scotland ridge at a lead time of 1 month (inside the green boxes in Figs. 2.9(a),(b)) and are now investigated further.

Since Kelvin waves travel with the coast to the right in the Northern Hemisphere, the Icelandic coastline is a waveguide for waves propagating clockwise around Iceland, arriving at the western end of the Iceland-Scotland ridge section. Thus, negative pressure anomalies along the entire Icelandic coastline result in a positive perturbation of the zonal pressure gradient along the Iceland-Scotland ridge section, and by geostrophic balance (2.6), a strengthening of the northward geostrophic transport across the ridge. This pathway explains feature S.1 inside the green box around Iceland in Figs. 2.9(a),(b). Wind stress perturbations that match the sign of this sensitivity feature, i.e., an increased northward (eastward, southward, westward) wind stress along the eastern (southern, western, northern) coast of Iceland, drives Ekman offshore divergence, resulting in a negative pressure anomaly at the Icelandic coast, which is communicated to the western end of the Iceland-Scotland section and strengthens the northward geostrophic transport.

To explain feature S.2, it is noted that the African and European west coast are upstream waveguides of the Iceland-Scotland ridge section, where signals arrive at the eastern end of the section. Positive pressure anomalies along the African and European west coast will therefore result in a positive perturbation in the zonal density gradient across the Iceland-Scotland ridge section, and again, by geostrophic balance (2.6), strengthen the northward geostrophic transport across the ridge. Anomalies in meridional wind stress that match the sign of the sensitivity feature in Fig. 2.9(b), i.e., increased northward wind stress along the African and European west coast, drive Ekman onshore convergence, resulting in the required positive pressure anomaly at the coastal boundary. Alternating sensitivities to zonal wind stress along the African and European west coast inside the green box in Fig. 2.9(a) (or increasingly visible for a lead time of

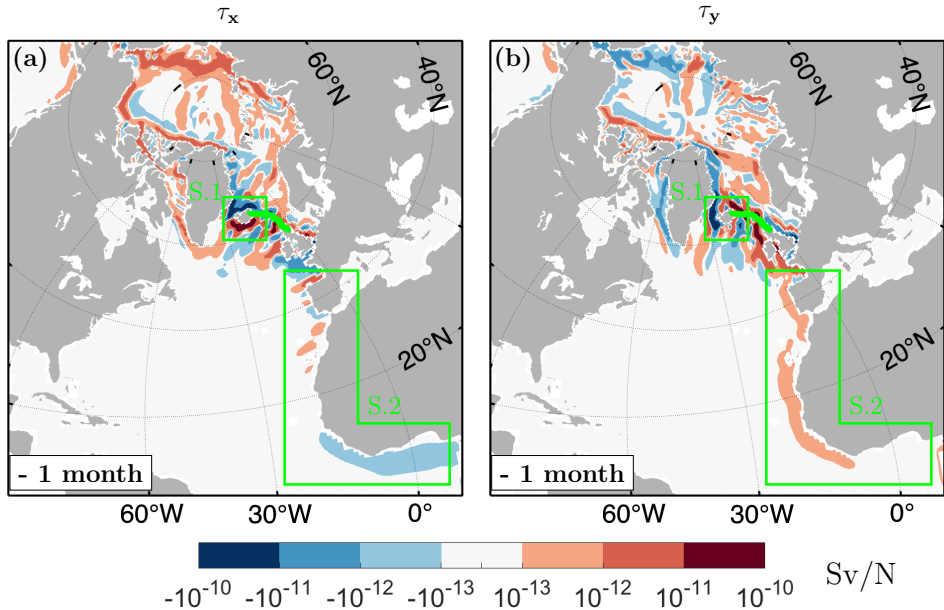


Figure 2.9: Linearized sensitivity $\frac{\partial T_{IS}}{\partial F(-t,x,y)}$ of March **volume transport across the Iceland-Scotland ridge** to (a) **zonal wind stress** ($F = \tau_x$) and (b) **meridional wind stress** ($F = \tau_y$) for a lead time of 1 month. Positive sensitivity indicates that an increase in monthly mean eastward wind in (a) and monthly mean northward wind in (b) causes higher north-eastward volume transport across the ridge (indicated by the green contour line) 1 month later. The green boxes contain the same sensitivity patterns as the respective boxes in Figs. 2.7 and 2.8.

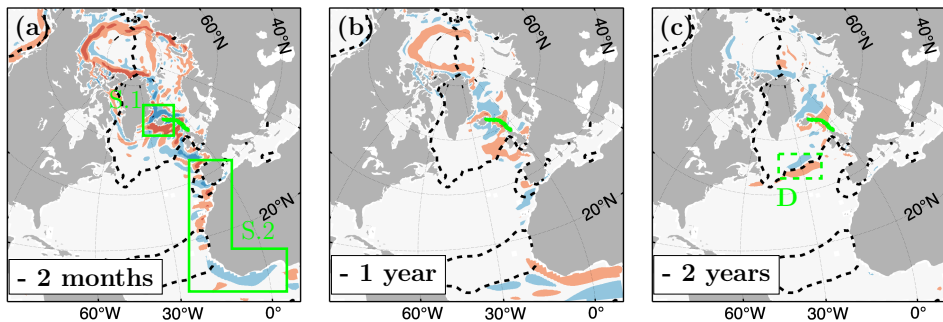


Figure 2.10: Linearized sensitivity $\frac{\partial T_{IS}}{\partial \tau_x(-t,x,y)}$ of March **volume transport across the Iceland-Scotland ridge** to **zonal wind stress** for increasing lead times of (a) 2 months, (b) 1 year, and (c) 2 years. The colorbar is the same as in Fig. 2.9. Positive sensitivity indicates that an increase in monthly mean eastward wind at the indicated lead time causes larger north-eastward volume transport across the ridge. The black and green contour lines and boxes are the same as in Fig. 2.7.

2 months in Fig. 2.10(a)) can be explained similarly. Sensitivity to zonal wind stress is positive along SW-NE-aligned coastlines because here, eastward wind drives onshore Ekman transport and thus positive pressure anomalies at the coastal boundary. Along NW-SE-aligned coastlines, on the other hand, westward wind is required for onshore Ekman transport. Therefore NW-SE-aligned coastlines are populated with negative sensitivity to zonal wind stress.

The adjustment processes that were discussed in the previous paragraphs are due to fast barotropic coastally trapped waves. These waves propagate at a speed of \sqrt{gH} , where g denotes gravitational acceleration and H the depth of the water column. Even for very shallow depths of 250 m, waves propagate with a speed of about 50 m/s (or 4300 km/day). Since the sensitivities are scaled to monthly forcing anomalies, the signature of fast barotropic waves, e.g., via the discussed patterns S.1 and S.2, emerges strongly in the sensitivity maps for lead times of 1-2 months, the shortest lead times that are considered.

Rapid loss of sensitivity Figs. 2.9(a) and 2.10(a)-(c) show sensitivity of volume transport across the Iceland-Scotland ridge to zonal wind stress for lead times of 1 month, 2 months, 1 year and 2 years. For increasing lead time, one observes a rapid loss of sensitivity to zonal wind stress. For instance, sensitivities around Iceland decrease by an order of magnitude between lead times of 1 month (Fig. 2.9(a)) and 2 months (Fig. 2.10(a)), and by another order of magnitude between lead times of 2 months (Fig. 2.10(a)) and 1 year (Fig. 2.10(b)). The loss of sensitivity to meridional wind stress (not shown here) is equally rapid as to zonal wind stress.

No stationarity Stationarity of sensitivity to wind stress, which was found for Nordic Seas upper-ocean heat content, is absent for volume transport across the Iceland-Scotland ridge. For example, the band of negative sensitivities along the African west coast from Cape Verde to the Gulf of Guinea for lead times of 1 and 2 months (Figs. 2.9(a) and 2.10(a)) turns into a band of positive sensitivities for a lead time of 1 year (Fig. 2.10(b)), reflecting a complex temporal evolution of barotropic and baroclinic adjustment processes.

Sensitivity dipole along intergyre boundary Finally, the wind stress sensitivity dipole along the intergyre boundary that was detected in the sensitivity maps of Nordic Seas upper-ocean heat content (Figs. 2.7(c),(d) and 2.8(c),(d)) is also visible in the sensitivity maps of volume transport across the Iceland-Scotland ridge to zonal wind stress (Fig. 2.10(c)) and meridional wind stress (not shown).

To summarize, the findings of this subsection are:

- Fast barotropic adjustment processes alter the volume transport across the Iceland-Scotland ridge on daily to weekly timescales. The same processes are imprinted on the sensitivity maps of Nordic Seas upper-ocean heat content on monthly to yearly timescales.

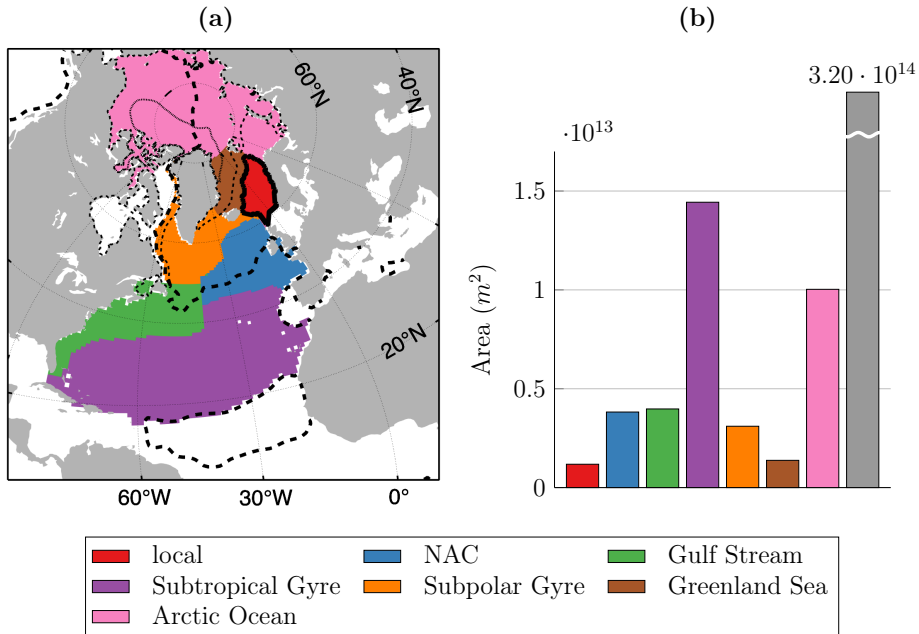


Figure 2.11: (a) Subregions of the global ocean that are considered in this work and (b) a comparison of their geographical area. In (a), the thick solid line marks the boundary of the Nordic Seas region, the thick dashed contour line is the zero-line of the barotropic stream function, and the thin dashed and dotted lines depict the sea-ice extent in March and September, respectively, in the model simulation in this chapter. The gray bar in (b) refers to the rest of the global ocean that is not part of any of the defined regions in (a).

- Cross-ridge volume transport is characterized by a rapid loss of sensitivity to wind stress anomalies, whereas Nordic Seas heat content can be impacted years later by local and remote wind stress anomalies.
- Positive anomalies in wind stress curl along the intergyre boundary cause positive anomalies in north-eastward cross-ridge volume transport and Nordic Seas heat content after 2 years and later.

These results are further discussed in Section 2.5.

2.4 Relative importance of forcings and regions of origin

Definition of regions

To summarize the results from the last section, the globe is divided up into geographical regions, as shown in Fig. 2.11(a). This allows a quantification of the relative importance of these regions on various timescales in generating variability in Nordic Seas upper-ocean heat content. Seven regions, which cover the North Atlantic and the Arctic, are defined. The region called “local” is the Nordic Seas domain itself, which was defined in

Section 2.2.2. The North Atlantic Current (NAC) is defined as the part of the North Atlantic which exhibits annual mean SST between 10 and 16°C in the equilibrated model simulation (described in Section 2.2.1). The boundary with the Gulf Stream is considered to be located at 40°W. The Gulf Stream, in turn, is defined as the region east of Florida, and between 25°N and 50°N, which has annual mean surface speed larger than 0.07 m/s in the equilibrated model simulation. In addition, the region west of 40°W and between 40°N and 50°N is included in the Gulf Stream region, as this region has been suggested as an important “transition zone” for strong water mass transformation in Buckley and Marshall [2016]. In the transition zone, the separated Gulf Stream, the North Atlantic Current, and Labrador Current all interact with each other at the western junction of the subtropical and subpolar gyres. The Subtropical Gyre is defined as the remaining part of the North Atlantic north of 16°N and south of the Gulf Stream and NAC regions where the barotropic stream function exhibits positive values. Similarly, the Subpolar Gyre is determined as the remaining part of the North Atlantic north of the Gulf Stream regions and NAC with negative barotropic stream function values. The Greenland Sea is the subregion of the Greenland, Iceland, Norwegian (GIN) Seas which is not part of the local region. The last region is the Arctic Ocean including the Barents Sea.

Response functions

To study the relative importance for Nordic Seas upper-ocean heat content variability across regions and forcings, the sensitivities are weighted by temporally-spatially varying fields $|\Delta F|$, which reflect the magnitude of typical anomalies in the respective forcing fields. This turns sensitivity fields $\frac{\partial \text{UOHC}_{NS}}{\partial F}$ into response fields $\frac{\partial \text{UOHC}_{NS}}{\partial F} \cdot |\Delta F|$. The absolute values of the response fields are then integrated over each of the defined regions \mathcal{R} , resulting in the *response functions*

$$|\Delta \text{UOHC}_{NS}|^{\mathcal{R},F}(t) = \int_{\mathcal{R}} \left| \frac{\partial \text{UOHC}_{NS}}{\partial F(x,y,-t)} \right| \cdot |\Delta F(x,y,m(-t))| dx dy, \quad (2.7)$$

where the lead time t is discretized into monthly values. Here, F can be either surface heat flux Q_{net} , surface freshwater flux E-P-R, zonal wind stress τ_x or meridional wind stress τ_y .

For each grid point (x, y) , 12 different values for $\Delta F(x, y, m)$ are used, one for each month m . These values are computed as the standard deviation of the departures from the seasonal cycle, using the statistics of the 20-year ECCOv4r2 forcing fields, which have been adjusted from ERA-Interim fields in the state estimation optimization process [Forget et al., 2015]. Since the sensitivities $\frac{\partial \text{UOHC}_{NS}}{\partial F}$ in Section 2.3 were computed with respect to surface fluxes F that act at the ocean surface (i.e., at the ice-ocean interface for sea-ice covered areas), the anomalies ΔF are computed according to the same definition. The function $m(-t)$ in (2.7) evaluates the month which corresponds to the lead time t . Since UOHC_{NS} is computed in March, a lead time of 1, 13, 25, ... months corresponds to the month February, a lead time of 2, 14, 26, ... months to the month January, and so on.

It is noted that the response functions, as defined in (2.7), ignore the signs of sensitivities and forcing anomalies - and more importantly, the product of these signs. For a reconstruction of Nordic Seas upper-ocean heat content anomalies, caused by global atmospheric forcing anomalies in the past, one would have to evaluate $\frac{\partial \text{UOHC}_{NS}}{\partial F} \cdot \Delta F$, taking into account both the sign of the sensitivities, $\frac{\partial \text{UOHC}_{NS}}{\partial F}$, and the sign of the actual occurring forcing anomalies, ΔF . If a forcing anomaly had uniform sign within a certain region which contains both positive and negative sensitivity subregions, this would lead to cancellations in the actual responses. For instance, Figs. 2.4(b),(c),(d), Figs. 2.7(c),(d) and Figs. 2.8(c),(d) showed adjacent positive and negative sensitivities in the form of the positive-negative dipole across the NAC and Gulf Stream. Moreover, the sensitivity fields to wind stress (Figs. 2.7 and Figs. 2.8) were characterized by a small-scale positive-negative structure at many locations. The same cancellation effect occurs if a region shows sensitivities of uniform sign but forcing anomalies exhibit alternating sign within that region. Both cancellation types will certainly occur in reality. Metric (2.7) can therefore be seen as an upper bound for absolute upper-ocean heat content anomalies in the Nordic Seas that are generated by atmospheric forcing anomalies. The magnitude of computed heat content anomalies will only be realized if the anomaly fields have the same positive-negative structure as the sensitivity fields such that no cancellations take place. However, studying response functions, as defined in (2.7), is the clearest way of highlighting regions of - potentially - greatest importance for upper-ocean Nordic Seas heat content.

2.4.1 Relative importance of forcings

The response functions (2.7) are shown in Fig. 2.12, where contributions from atmospheric forcing anomalies in different regions are stacked on top of each other. The response functions all have the same unit (K, the unit of UOHC_{NS}), regardless of the forcing type F . Contributions due to different forcings are therefore now comparable to each other.

Stacked sensitivities in Fig. 2.12(a), reflecting the accumulated effect on Nordic Seas upper-ocean heat content from heat flux anomalies in all regions, show a rapid decrease for increasing lead time. In contrast, sensitivity to freshwater flux does not show a notable decrease in amplitude on a ten-year timescale (Fig. 2.12(b)). As a result, while of minor relative importance on short timescales, freshwater flux anomalies become an important driver of UOHC_{NS} variability on increasingly longer timescales. Figs. 2.12(c),(d) confirm that - after an initial loss of sensitivity over the course of the first two years - sensitivity to wind stress decreases *moderately*, as already discussed in Section 2.3.2. The moderate loss of wind stress sensitivity of UOHC_{NS} stands in contrast to the very rapid loss of wind stress sensitivity of T_{IS} (not shown). Loss of sensitivity with increasing lead time for the different forcing types is sketched by the black arrows in Fig. 2.12.

Counting contributions from all regions together, the impact of anomalies in either of the wind stress components is as large as the impact of heat and freshwater flux anomalies combined, on all considered timescales. For example, monthly flux anomalies occurring at a lead time of one month, when heat flux and wind stress anomalies have the largest impact, can lead to Nordic Seas upper-ocean heat content anomalies of up to 0.09 K

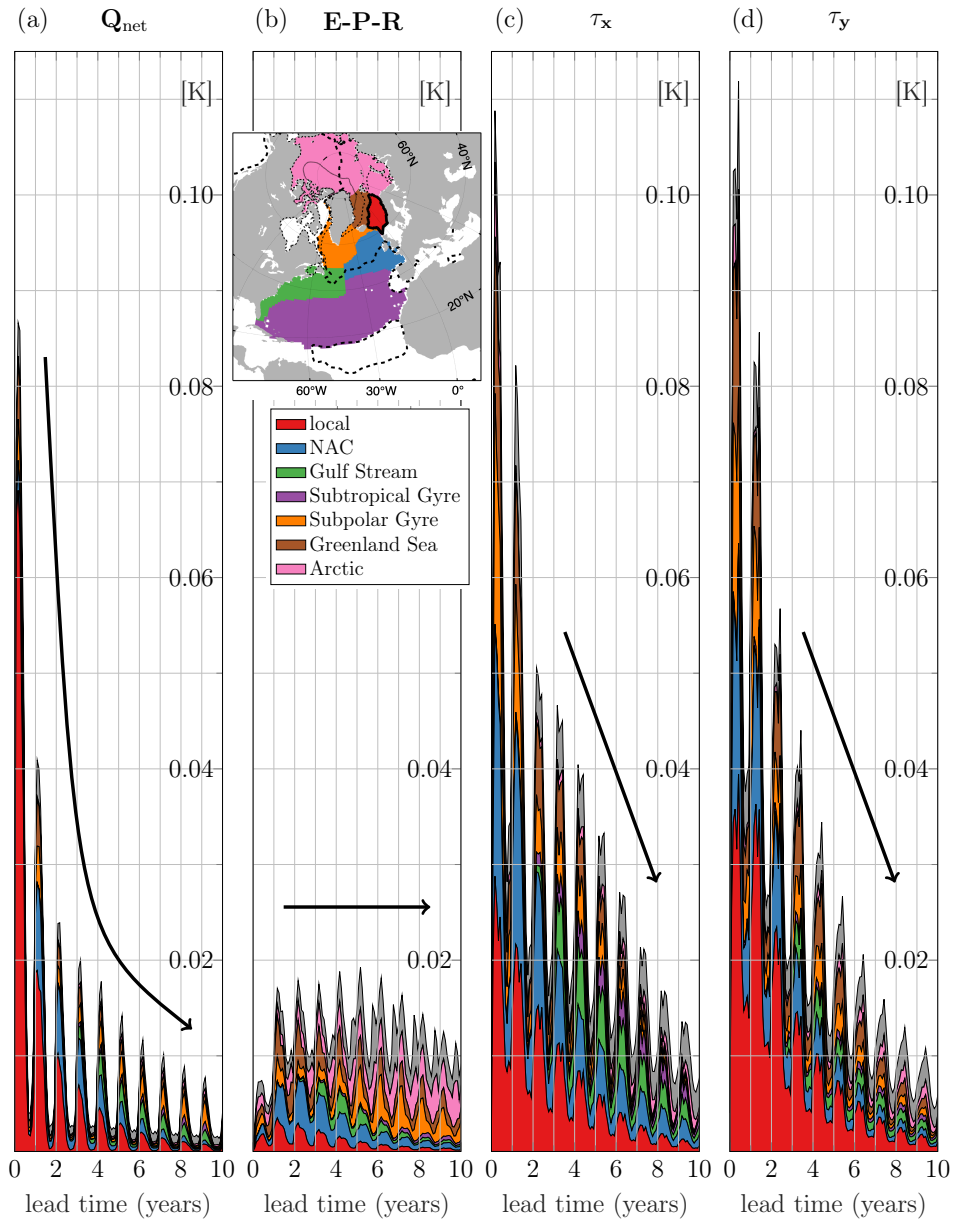


Figure 2.12: Response functions $|\Delta\text{UOHC}_{\text{NS}}|^{\mathcal{R},F}(t)$, defined in (2.7), of March Nordic Seas upper-ocean heat content to (a) surface heat flux Q_{net} , (b) surface freshwater flux E-P-R, (c) zonal wind stress τ_x , and (d) meridional wind stress τ_y . The responses are shown as a function of lead time t (in years) of monthly forcing anomalies. Contributions from atmospheric forcing anomalies in different regions \mathcal{R} are stacked on top of each other. The regions are color-coded according to the legend. Gray shading refers to the rest of the global ocean that is not part of any of the shown regions in the map. The black arrows sketch the characteristic loss of sensitivity to the respective forcing types.

for monthly heat flux anomalies (Fig. 2.12(a)), up to 0.01 K for monthly freshwater flux anomalies (Fig. 2.12(b)), and up to approximately 0.11 K for typical monthly zonal or meridional wind stress anomalies (Figs. 2.12(c),(d)). If typical monthly flux anomalies occur in the previous summer, corresponding to a lead time of about 0.5 years, generated anomalies in March Nordic Seas upper-ocean heat content are much smaller, namely of magnitude less than 0.01 K for heat and freshwater fluxes, respectively, and of magnitude 0.04 K for each of the wind stress components. At a lead time of 3 years (corresponding to the month March 3 years earlier), Nordic Seas upper-ocean heat content can be altered by 0.02 K by typical variations in monthly heat and freshwater flux anomalies, respectively, and by about 0.05 K by monthly anomalies in each of the wind stress components.

The dominant role of wind stress compared to buoyancy fluxes is suggested to be due to (i) a strong reponse of Nordic Seas heat content to barotropic adjustment processes on short timescales (< 1 year), and (ii) the moderate loss of sensitivity to wind stress with increasing lead time. Both aspects have been discussed in Section 2.3.2. It is noted, however, that the controlling influence of wind stress anomalies over buoyancy anomalies could partly be a result of the chosen metric, too. That is, as noted before, the response functions, defined in (2.7) and shown in Fig. 2.12, might favor sensitivities to wind stress. In reality, the small-scale structure in the sensitivity fields, detected in Figs. 2.7 and Figs. 2.8, will be partly cancelled if forcing anomalies with larger-scale structure are projected onto them.

2.4.2 Seasonality

The variation of the responses at annual frequency in Figs. 2.12(a)-(d) shows that the effect of forcing anomalies depends strongly on the season in which they occur. Generally, the impact is highest if forcing anomalies occur in winter and spring when mixed layer depths are at their deepest. A more detailed analysis of the seasonality of the responses is presented in Appendix A.

2.4.3 Relative importance of regions of atmospheric origins

On short timescales (< 1 year) *local* surface heat flux anomalies are most influential (Fig. 2.12(a)) for altering March Nordic Seas upper-ocean heat content. Fig. 2.12(a) confirms that the effect of local heat flux anomalies decreases rapidly with increasing lead time, as discussed in Section 2.3.1. Moreover, it supports the identified reemergence mechanism, reflected in the fact that local heat flux anomalies from several winters back in time still have an impact. Sensitivity to heat flux is also detectable in the neighboring regions of the Nordic Seas, namely the Greenland Sea, the Subpolar Gyre, and the NAC. Sensitivity to heat flux anomalies in the Gulf Stream is visible for lead times of 2.5 years and greater, set by the advective timescale.

On timescales shorter than 5 years, the UOHC_{NS} response to surface freshwater flux anomalies is strongest in the NAC and Greenland Sea (Fig. 2.12(b)). Here, freshwater flux anomalies can alter the geostrophic current via thermal wind balance, as discussed in Section 2.3.1. Fig. 2.12(b) shows that the maximum UOHC_{NS} response to freshwater flux occurs for a forcing lead time of 5 years (as opposed to an immediate response to

local heat fluxes), which is a result of freshwater flux anomalies requiring active ocean dynamics in order to impact Nordic Seas heat content. Similarly as seen for heat fluxes, a notable response to freshwater flux anomalies in the Gulf Stream emerges at lead times of 2.5 years and greater. For forcing lead times greater than 5 years, sensitivity to surface freshwater flux is highest in the Subpolar Gyre, the Greenland Sea, and the Arctic Ocean.

For any forcing lead time between 0 and 10 years, wind stress anomalies in the local region, the NAC, the Subpolar Gyre and the Greenland Sea show all similar contributions (Figs. 2.12(c),(d)). Some of the immediate response to local wind stress anomalies are due to Ekman dynamics. The remaining contributions are attributed to the adjustment mechanisms that are reflected by the stationary wind stress sensitivity patterns in Section 2.3, imprinted by barotropic adjustment processes that alter the volume transport across the Iceland-Scotland ridge. With increasing lead time, from 2 to 10 years, the contributions due to these processes decrease approximately linearly at a similar slope (Figs. 2.12(c) and (d)). On timescales less than 3 years, these mechanisms generate most of the variability in Nordic Seas upper-ocean heat content.

On timescales between 3 and 10 years, strong response to wind stress in the Gulf Stream and NAC in Figs. 2.12(c) and (d) is consistent with sensitivity to wind stress curl anomalies along the intergyre boundary, detected in Section 2.3. UOHC_{NS} response to zonal wind stress anomalies in the Subtropical Gyre emerges together with UOHC_{NS} response to wind stress anomalies in the Gulf Stream, for the first time at a forcing lead time of 2-3 years. This is a signature of Rossby waves, which communicate pressure anomalies at mid-latitude westward across the North Atlantic basin. Anomalies carried by these Rossby waves are fed into the Gulf Stream, which eventually carries them to the Nordic Seas. The Rossby wave effect was too small to be discernible in Fig. 2.7, but would be visible with a colorbar showing the next smaller logarithmic decade. However, as an integrated effect over the large area of the Subtropical Gyre, it constitutes a substantial contribution to Nordic Seas upper-ocean heat content variability.

In Appendix A, it is investigated whether the identified contributions of forcing anomalies in a particular region are due to (i) high sensitivity to this forcing type in the considered region, (ii) high typical forcing anomalies in this region, or (iii) a large region size, which integrates rather small effects to a significant contribution.

2.5 Discussion

The adjoint of a realistic global ocean model has been used to investigate local and remote atmospheric origins of variability in Nordic Seas upper-ocean heat content on a ten-year timescale. Adjoint sensitivity distributions reveal the pathways and the role of physical mechanisms in carrying anomalies generated across the global ocean surface to the Nordic Seas. By comparing adjoint sensitivities of Nordic Seas upper-ocean heat content to adjoint sensitivities of volume transport across the Iceland-Scotland ridge, the important dynamical linkage between these two quantities is elucidated. Furthermore, it is quantified how different regions across the globe and different forcing types contribute to generating Nordic Seas upper-ocean heat content variability on different timescales.

Importance of local air-sea heat fluxes on short timescales

Nordic Seas upper-ocean heat content is found to be most sensitive to anomalies in local heat fluxes which take place in the same year during winter, when the mixed layer depth is at its deepest. Due to reemergence, these heat flux anomalies still show an impact on Nordic Seas upper-ocean heat content in the following winter, although their strength is reduced by 70%. The effect on Nordic Seas heat content in subsequent winters decreases rapidly. Heat flux anomalies taking place in summer are ineffective in altering Nordic Seas upper-ocean heat content on timescales longer than a few months.

Sensitivity of ocean heat content vs. ocean transports

For *remote*, rather than local, atmospheric forcing anomalies, the Iceland-Scotland ridge is the key adjustment pathway for Nordic Seas upper-ocean heat content on both short (< 1 year) and longer timescales.

Volume transport across the Iceland-Scotland ridge shows high sensitivity to wind stress in regions that are upstream waveguides of the Iceland-Scotland ridge section, including subregions of the GIN Seas (in particular the Icelandic coastline), the Arctic Ocean, the NAC, and the northern hemisphere African and European west coast. Pressure anomalies are carried by fast barotropic coastally trapped waves within days from the region of origin of the wind stress anomaly to the Iceland-Scotland ridge section, where they impact the geostrophic transport across the ridge. For increasing lead time, the time-evolving wind sensitivity patterns reflect a complex interplay of barotropic and baroclinic adjustment processes, but their impact on cross-ridge volume transport decreases rapidly: The effect of wind stress anomalies that occurred more than one month earlier is of at least one order of magnitude less in strength than the fastest barotropic signal. The importance of barotropic adjustment processes for cross-ridge volume transport on short timescales and a rapid loss of sensitivity to wind stress are consistent with the results from previous studies that investigated adjoint sensitivities of volume transports across other sections in the Atlantic [Pillar et al., 2016; Smith and Heimbach, 2019].

Analyzing Nordic Seas upper-ocean heat content sensitivities alongside sensitivities of volume transport across the Iceland-Scotland ridge highlights that Nordic Seas upper-ocean heat content is to a large degree the *integrated consequence* of anomalous wind-driven cross-ridge volume transport. This is reflected in the wind stress sensitivity distributions by the following characteristics: The sensitivity patterns of cross-ridge volume transport that are set by the fast barotropic adjustment processes, which dominate the net response in volume transport, are imprinted onto Nordic Seas upper-ocean heat content sensitivities as a stationary feature. This persistence can be explained as follows. While volume transport soon loses its memory to adjustment processes that act on daily timescales, temperature anomalies carried across the ridge by anomalous volume transport are “stored in memory” of Nordic Seas upper-ocean heat content. As a consequence, there is only moderate loss of sensitivity to wind forcing with increasing lead time, and wind stress anomalies in the upstream waveguides of the Iceland-Scotland ridge section can affect Nordic Seas heat content even years later. In short, a large portion of Nordic Seas upper-ocean heat content variability on monthly to yearly timescales is generated by fast barotropic adjustment processes which alter the volume transport across the

Iceland-Scotland ridge.

The large majority of previous adjoint-based studies have focused on metrics of ocean circulation, e.g., heat and volume transports across certain sections [e.g., Heimbach et al., 2011; Pillar et al., 2016; Smith and Heimbach, 2019]. A novelty of this work is that adjoint-derived sensitivities of a kinematic (transport) quantity (T_{IS}) are examined alongside adjoint-derived sensitivities of a closely related thermodynamic quantity ($UOHC_{NS}$). Despite the close association of the quantities, there are surprising differences in their sensitivity to external forcing, e.g., stationary wind stress sensitivities of $UOHC_{NS}$, as discussed in the previous paragraph. These differences in the sensitivities are due to different “residence capacities” of the metrics and have important implications for inferring changes in ocean heat storage from transport metrics.

Wind stress curl anomalies along intergyre boundary affect the Nordic Seas

Apart from barotropic adjustment processes, the adjoint sensitivity distributions reveal another mechanism by which wind stress anomalies can alter Nordic Seas upper-ocean heat content effectively. A positive anomaly in wind stress curl along the intergyre boundary leads to positive anomalies in the north-eastward volume transport across the Iceland-Scotland ridge and in Nordic Seas heat content. For a wind stress curl anomaly occurring at $40^\circ W$ along the intergyre boundary, the response signal for Nordic Seas upper-ocean heat content arrives about 2 years later but persists for many years. Forward perturbation experiments with the full non-linear model validate the detected adjustment mechanism (see Chapter 5). It is noted that the detected teleconnection between wind stress curl at the intergyre boundary and Nordic Seas upper-ocean heat content is particularly relevant because the intergyre boundary is the location where wind stress curl anomalies associated with the NAO are the strongest [e.g., Marshall et al., 2001; Häkkinen et al., 2011].

Many studies have investigated the effect of wind stress curl and the wind-driven gyre circulations on high-latitude ocean temperature. Hátún et al. [2005] and Häkkinen et al. [2011] argue that a contraction of the subpolar gyre leads to the NAC being composed of more warm, saline subtropical waters and less cold, fresher subpolar waters than usual. This anomaly in the NAC composition is then reflected in a warmer and more saline inflow into the Nordic Seas. Indeed, the sensitivity distribution found in this chapter - which indicates that a weakening of the westerlies in the subpolar gyre and strengthening of the westerlies in the subtropical gyre eventually increases Nordic Seas upper-ocean heat content - can be read in a way consistent with Hátún et al. [2005] and Häkkinen et al. [2011]. The modification of the westerlies in the way just described would weaken the cyclonic subpolar gyre and strengthen the anticyclonic subtropical gyre, such that more warm waters flow into the Nordic Seas (although Häkkinen et al. [2011] suggest that a weakening of *both* gyres is even more efficient for the subpolar front to move westward and to open the pathway for more subtropical waters). However, the results in this chapter also show that volume transport across the Iceland-Scotland ridge is sensitive to wind stress curl anomalies along the intergyre boundary in the same way as Nordic Seas upper-ocean heat content. This indicates that the responsible mechanism is not solely a change in the gyre shapes and in the relative contribution of subpolar and subtropical

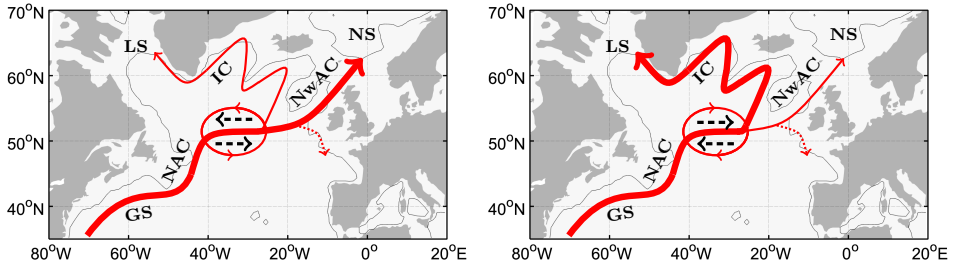


Figure 2.13: Schematic of the effect of wind stress curl along the intergyre boundary on North Atlantic surface currents. (a) shows weakened westerlies in the subpolar gyre and strengthened westerlies in the subtropical gyre (black dashed arrows), corresponding to a *positive anomaly in wind stress curl* along the intergyre boundary. A cyclonic intergyre-gyre is spun up (Marshall et al. [2001]), which steers relatively more water into the NwAC and relatively less water into the IC. (b) shows a *negative anomaly in wind stress curl* along the intergyre boundary, which spins up an anticyclonic intergyre-gyre driving relatively more water into the IC and relatively less water into the NwAC. Red solid arrows represent the Gulf Stream (GS), the North Atlantic Current (NAC), the Irmingier Current (IC), the Norwegian Atlantic Current (NwAC), and the intergyre-gyre. The red dashed arrow represents the southeastern pathway discussed in Häkkinen and Rhines [2009]. Black dashed arrows show anomalies in wind stress. Furthermore, the schematics shows the location of the Nordic Seas (NS) and the Labrador Sea (LS).

waters, as suggested by Hátún et al. [2005] and Häkkinen et al. [2011] - such a change would mostly lead to anomalies in the advected temperature and *not* in the cross-ridge volume transport.

In contrast, Orvik and Skagseth [2003] suggest a baroclinic adjustment mechanism along the Gulf Stream/NAC path, which *does* alter cross-ridge volume transport in response to a change in wind stress curl along the intergyre boundary. The authors propose that anomalous local Ekman upwelling at the intergyre boundary, induced by a positive wind stress curl anomaly, forces a baroclinic Rossby wave, which propagates with the NAC until it hits the shelf edge current along the Irish-Scottish coast. The signal is then carried across the Iceland-Scotland ridge as a barotropic coastally trapped boundary wave. The induced anomaly in the pressure gradient along the Iceland-Scotland ridge section alters the cross-ridge geostrophic transport. Orvik and Skagseth [2003] deem the described adjustment mechanism responsible for the positive correlation that they detect between the zonally integrated wind stress curl at 55°N and the volume transport across the Svinøy section, which is located a few degrees north of the Iceland-Scotland ridge, with a 15-month lag. The positive correlation found by Orvik and Skagseth [2003], including the diagnosed time lag, is confirmed by the findings in this chapter based on adjoint sensitivities, which identify physical relationships and causal chains contained in the model equations, rather than covariability of the two quantities.

Here, it is hypothesized that yet another adjustment mechanism is at play in generating Nordic Seas upper-ocean heat content anomalies in response to wind stress curl

anomalies along the intergyre boundary. The suggested mechanism is illustrated by the schematics in Fig. 2.13 and acts through a shift of the North Atlantic current system. The Norwegian Atlantic Current (NwAC), which transport warm Atlantic waters across the Iceland-Scotland ridge, is not the only extension of the North Atlantic Current (NAC). Another main branch of the NAC is the Irminger Current (IC), which carries a large portion of the warm water masses carried by the NAC to the Irminger and Labrador Seas (Fig. 1.1). The wind stress curl along the main streamline of the NAC has an influence on how the water is distributed between these two branches. Fig. 2.13(a) sketches the effect of a *positive* anomaly in wind stress curl along the intergyre boundary. As explained in Marshall et al. [2001], a cyclonic intergyre-gyre is spun up, which shifts the main streamline of the NAC to its right (as seen from the view of a water parcel traveling with the NAC). The streamline shift drives relatively more waters into the NwAC, which branches to the right, and less waters into the IC, which branches to the left. The result is a positive anomaly in volume transport across the Iceland-Scotland ridge and thus a positive anomaly in Nordic Seas upper-ocean heat content. As shown in Fig. 2.13(b), a negative anomaly in wind stress curl has the opposite effect. The induced anticyclonic intergyre-gyre steers relatively more waters into the IC, and less into the NwAC, which leads to a negative anomaly in volume transport across the Iceland-Scotland ridge and in Nordic Seas upper-ocean heat content.

The mechanism suggested in Fig. 2.13 is supported by adjoint sensitivities of Labrador Sea heat content, which also shows a wind stress sensitivity dipole along the intergyre boundary, but of *reversed sign* [Jones et al., 2018]. Figs. 8(e),(f) in Jones et al. [2018] show sensitivities to zonal wind stress at forcing lead times of 7.9 and 3.9 years, respectively. In these figures, the sensitivity dipole is visible west of 40°W along the main stream line of the Gulf Stream, with positive sensitivity to the north and negative sensitivity to the south. To support this finding, I carried out a sensitivity experiment (not shown) for another quantity of interest in the Labrador Basin, heat transport across the OSNAP-West section, extending from southern Labrador to southwestern Greenland. The obtained sensitivity distributions confirm the location and sign of the sensitivity dipole in the Gulf Stream for a forcing lead time of about 3.5 years, as shown in Fig. 8(f) of Jones et al. [2018]. Moreover, the dipole straddles the NAC at about 30°W for a forcing lead time of 2.5 years. That is, a negative anomaly in wind stress curl eventually leads to a positive anomaly in Labrador Sea temperature. Since the Labrador Sea is fed with warm Atlantic waters by the IC, the highlighted causal chain is consistent with the mechanism sketched in Fig. 2.13(b).

It is noted that Marshall et al. [2001] suggest the opposite effect of a cyclonic intergyre-gyre, which has been spun up by a positive anomaly in wind stress curl (Fig. 2.13(a)). They argue that the resulting southward shift of the NAC reduces poleward heat transport, which would entail a negative anomaly in Nordic Seas upper-ocean heat content. However, the authors use a simple model, in which each gyre is represented by a single box. Their model does therefore not incorporate the detailed current system of the NAC with its various branches which transport waters to different regions at high-latitudes, e.g., the Nordic Seas vs. the Labrador Sea. Häkkinen and Rhines [2009] also suggest a mechanism by which wind stress curl along the intergyre boundary can shift the North Atlantic currents, similar to the mechanism illustrated in Fig. 2.13. How-

ever, the mechanism in [Häkkinen and Rhines \[2009\]](#) determines how much waters from the NAC enter the NwAC vs. a southeastern pathway (indicated by the dashed red arrow in [Fig. 2.13](#)), rather than the IC, as discussed here. A more detailed analysis will have to clarify how dependent the sensitivities and the associated suggested wind stress curl-mechanism are on the model base state and the representation of the North Atlantic current system in the model.

Temperature vs. flow anomalies in the NAC and Gulf Stream

The NAC and Gulf Stream are not only key regions for wind stress anomalies. It is found that also surface buoyancy flux anomalies in these regions can affect Nordic Seas upper-ocean heat content, although less effectively than wind stress anomalies. The response time is set by the advective timescale, similarly as for wind stress anomalies. A buoyancy forcing anomaly occurring in the NAC at 40°W shows its first response in Nordic Seas upper-ocean heat content about 2 years later. After that, a response signal in the Nordic Seas is notable for many years.

Buoyancy forcing in the NAC and Gulf Stream can impact Nordic Seas upper-ocean heat content via geostrophic advection anomalies in two distinct ways: (i) advection of temperature anomalies by the mean current, and (ii) advection of the mean temperature by an anomalous geostrophic current, driven by density changes. Surface freshwater fluxes can contribute to Nordic Seas upper-ocean heat content variability only via (ii), i.e., by changing the geostrophic transport, whereas surface heat fluxes can change geostrophic transport *and* the advected temperature, therefore driving both (i) and (ii). The adjoint sensitivity distributions reveal that freshwater flux anomalies can alter Nordic Seas upper-ocean heat content effectively along the entire NAC and Gulf Stream path by thermal wind balance, a mechanism of type (ii). In contrast, heat flux anomalies in the NAC have the strongest impact on Nordic Seas upper-ocean heat content if they take place just south of the Iceland-Scotland ridge, while they gradually lose their influence on the Nordic Seas the further they occur upstream. This is attributed to the fact that mechanism (i) is counteracted by SST damping. Therefore, the farther away SST anomalies are generated by anomalous heat fluxes, the longer they are subject to a damping by air-sea heat fluxes while being advected to the Nordic Seas. A detailed discussion on the interplay of oceanic geostrophic advection and air-sea heat fluxes in the Gulf Stream and NAC is given in [Dong and Sutton \[2001\]](#), [Dong et al. \[2007\]](#), and [Buckley et al. \[2014\]](#). It is noted that the ocean-only model in this chapter does not have an active atmosphere. Instead, the SST damping effect is captured via the model bulk formulae of [Large and Yeager \[2004\]](#).

Interestingly, the adjoint sensitivity distributions suggest that SST damping does not only effect mechanism (i), as discussed in the last paragraph, but also density-driven changes of the geostrophic current, i.e., mechanisms of type (ii). Surface heat flux anomalies occurring south of 55°N alter the geostrophic transport across the Iceland-Scotland ridge in the opposite direction as expected from the heat flux-induced density change. This can be explained as follows. Surface heat flux anomalies affect density directly, via ocean warming (cooling), but also indirectly, due to increased (decreased) evaporation and thus ocean salinification (freshening). The results in this chapter suggest

that the indirect effect overrides the direct effect if the surface heat flux anomaly occurs far enough upstream, south of 55°N , again due to SST damping. A similar effect is found and thoroughly investigated in [Kostov et al. \[2019\]](#). The identified effect of remote heat flux anomalies on volume transport is a characteristic of the ECCOv4 ocean boundary conditions, which are formulated in such a way that heat and freshwater fluxes can communicate with each other. On the contrary, the effect has not been observed in previous adjoint sensitivity studies, where less realistic boundary conditions were used, e.g., flux boundary conditions together with surface restoring [[Junge and Haine, 2001](#); [Pillar et al., 2016](#)].

Freshwater dominates over heat flux forcing in the subpolar gyre and Arctic Ocean

For forcing lead times of 5 years and longer, Nordic Seas upper-ocean heat content shows increasingly larger sensitivity to buoyancy forcing in the subpolar gyre and the Arctic Ocean. The adjoint sensitivity distributions suggest that negative buoyancy forcing anomalies (making the surface ocean denser) eventually lead to a larger volume transport across the Iceland-Scotland ridge and higher Nordic Seas upper-ocean heat content. This causal chain is consistent with a strengthened overturning circulation, induced by denser waters in the deep convection regions. The stronger overturning brings more warm Atlantic waters across the ridge and into the Nordic Seas.

Heat flux anomalies are only found to be effective if they occur right inside the deep convection region in the Labrador Sea. On the other hand, sensitivity to freshwater flux anomalies is visible in the entire subpolar gyre and Arctic Ocean. This discrepancy is again attributed to the fact that SST anomalies are damped by air-sea heat fluxes while SSS anomalies can persist while being advected to the deep convection regions from elsewhere in the subpolar gyre and Arctic Ocean. These findings are consistent with other studies [e.g., [Deshayes et al., 2014](#)] which conclude that salinity anomalies play a more significant role in AMOC variability on longer timescales due to a less vigorous damping. Moreover, the propagation of salinity anomalies from the Arctic Ocean to the subpolar North Atlantic is supported by observations [[Karcher et al., 2005](#)].

The analysis in this chapter shows that surface freshwater flux anomalies in the *Arctic Ocean* play a dominant role on longer (> 5 years) timescales. This is mainly due to large freshwater flux anomalies in the Arctic, caused by high sea-ice variability. It is noted that freshwater flux anomalies are expected to become even larger in the future due to more sea-ice melt.

Limitations

The results in this chapter are limited by the linearity assumption of the adjoint. Adjoint-derived sensitivities are therefore evaluated against non-linear forward perturbation experiments in Chapter 5. A second limitation is that the sensitivities of volume transport across the Iceland-Scotland ridge were computed in depth coordinates. It has been shown that in the subpolar basin, density coordinates are a more appropriate choice [[Zhang, 2010](#)], due to the predominance of diapycnal mixing and a region-wide sloping of isopycnals.

2.6 Conclusions

The main conclusions of this chapter are:

- The schematic in Fig. 2.14 summarizes the identified atmospheric origins of upper-ocean heat content anomalies in the Nordic Seas, as well as the associated adjustment mechanisms.
 - (A) On short timescales (< 1 year), Nordic Seas upper-ocean heat content shows high sensitivity to *local air-sea heat fluxes* (dark red shading inside the Nordic Seas domain in Fig. 2.14). Due to reemergence, the signal is visible in the following winter, but decreases rapidly afterward.
 - (B) *Wind-driven barotropic adjustment processes* alter the volume transport across the Iceland-Scotland ridge and have an immediate effect on Nordic Seas upper-ocean heat content. Moreover, since Nordic Seas heat content integrates volume transport anomalies, a response signal due to barotropic adjustment processes is visible up to 5 years later. Adjustment pathways are the upstream waveguides of the Iceland-Scotland ridge section, sketched by the dark green arrows in Fig. 2.14.
 - (C) Wind stress curl anomalies along the intergyre boundary (light green shading in Fig. 2.14) can affect Nordic Seas upper-ocean heat content on even longer timescales due to a shifting of the North Atlantic surface currents. A first strong signal is notable after 2 years, but the effect is visible throughout the end of the studied 10 year time frame.
 - (D) Due to SST damping, *advection of heat flux-driven temperature anomalies* into the Nordic Seas is only effective if the heat flux anomalies occur nearby - in the Greenland Sea and in the NAC close to the Iceland-Scotland ridge (dark red shading north of 55°N outside the Nordic Seas in Fig. 2.14). The signal becomes negligible after 2 years.
 - (E) On the other hand, *advection of the mean temperature by anomalous currents* can affect Nordic Seas upper-ocean heat content between 1 to 10 years after the occurrence of surface buoyancy flux anomalies. Freshwater and heat flux anomalies in the Gulf Stream and NAC (dark red shading south of 55°N in Fig. 2.14), as well as freshwater flux anomalies in the Greenland and Barents Seas (dark blue shading in Fig. 2.14) change the baroclinic structure of the upper ocean and thus alter the geostrophic transport across the Iceland-Scotland ridge.
 - (F) Surface heat flux anomalies inside the deep convection regions (light red shading in Fig. 2.14) and freshwater flux anomalies in the entire subpolar gyre and Arctic Ocean (light blue shading in Fig. 2.14) drive changes in the *overturning* circulation with a notable response in Nordic Seas upper-ocean heat content after about 5 years. These processes are expected to play an important role on timescales beyond the studied 10 years.
- Local and remote atmospheric forcing anomalies are most effective in driving the adjustment mechanisms (A)-(F) in winter and spring when mixed layer depths are at their deepest.

- Volume transport across the Iceland-Scotland ridge and Nordic Seas heat content are closely linked oceanic quantities. Nevertheless, there are surprising differences in the sensitivity of these two quantities to remote external forcing, due to different “residence capacities” of the quantities. For instance, volume transport rapidly loses its memory to remote wind perturbations that are communicated to the Iceland-Scotland ridge via mechanism (B). In contrast, temperature anomalies carried across the ridge by this mechanism are “stored in memory” of Nordic Seas heat content. The differences in sensitivity have important implications for inferring changes in ocean heat storage from transport metrics.

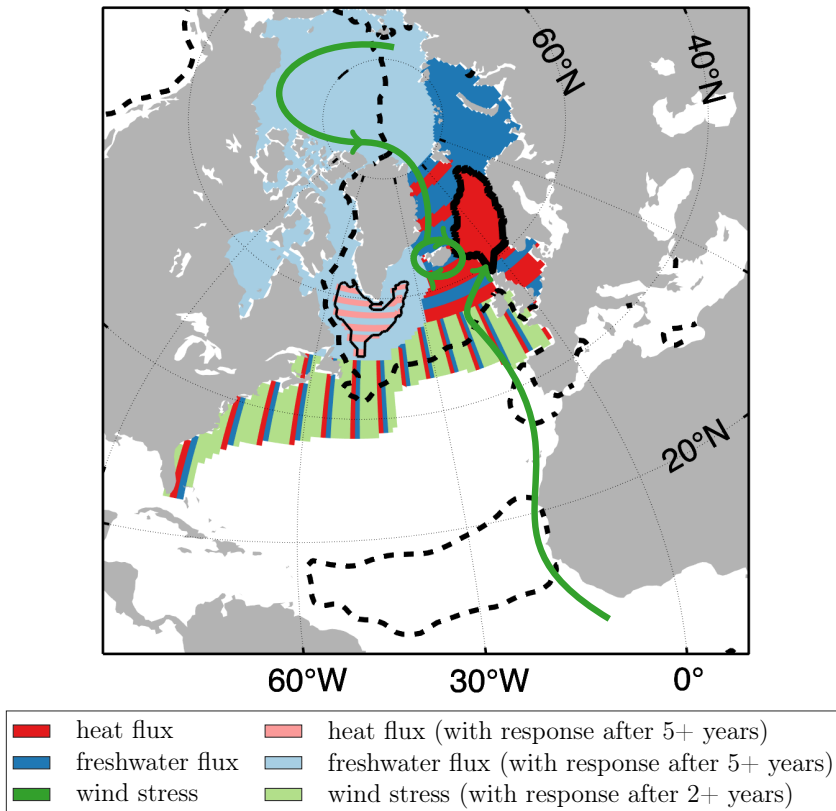


Figure 2.14: Schematic of atmospheric origins of upper-ocean heat content anomalies in the Nordic Seas on 10 year timescales, as identified in this study. The associated mechanisms (A)-(F) are described in the text. The thick solid black contour line bounds the studied Nordic Seas domain, the thick dashed black line shows the zero-line of the barotropic stream function, and the thin solid black contour line bounds the region wherein the March mixed layer depth exceeds 1000 m.

Chapter 3

Dynamics-based Ocean Observing System Design

3.1 Introduction

Long-term and sustained ocean observations are needed to advance our understanding of key mechanisms and ocean variability, to better constrain uncertain processes in models, and to contribute to climate predictions (e.g., [National Academies of Sciences, Engineering, and Medicine, 2017](#) and Section 1.1.6). However, observing systems are extremely expensive to build and maintain. It is thus an urgent task to quantitatively assess the effectiveness of existing and potentially new observing systems. Observing System Simulation experiments (OSSE) are the most popular approach for observing system design and evaluation (see Section 1.3.2). OSSEs have proven as a useful tool for improving forecasts [e.g., [Masutani et al., 2010](#)] - and are therefore widely used in numerical weather prediction and operational oceanography. In ocean climate research, on the other hand, observing systems are not only meant to improve the estimate of the ocean state and its variability; aims include *understanding why* the improvement is achieved, which larger climate signals are captured by the observing system, and by which ocean processes remote signals are propagated to the instruments. For these aims, the adjoint of a GCM is a powerful tool because it can efficiently compute the sensitivity of a model quantity to all other variables encoded in the GCM (Fig. 1.9). Adjoint models can therefore detect physical relationships and causal chains that connect the observed quantities to the rest of the global ocean (Section 1.2).

In ocean and climate research, adjoint models are mainly used for two types of applications, represented by the two branches (I) and (II) in the lower half of Fig. 1.8: state estimation and sensitivity analysis. A state estimate combines observations with a numerical model in a dynamically consistent way, and is obtained by solving a *non-linear inverse problem*, minimizing a regularized model-data misfit function J (Section 1.2.1). In a sensitivity analysis one computes the linearized sensitivity $\frac{\partial(\text{QoI})}{\partial \mathbf{x}}$ of a climate quantity of interest (QoI) to all control and prognostic variables, such as surface forcing and ocean temperature over the globe and at all lead times (Section 1.2.2). Sensitivity analysis has been used to gain insights into observations which are desirable to understand and monitor a QoI better, following the motto “high sensitivity informs high priority” [e.g., [Marotzke et al., 1999](#); [Heimbach et al., 2011](#)]. This approach of using adjoint sen-

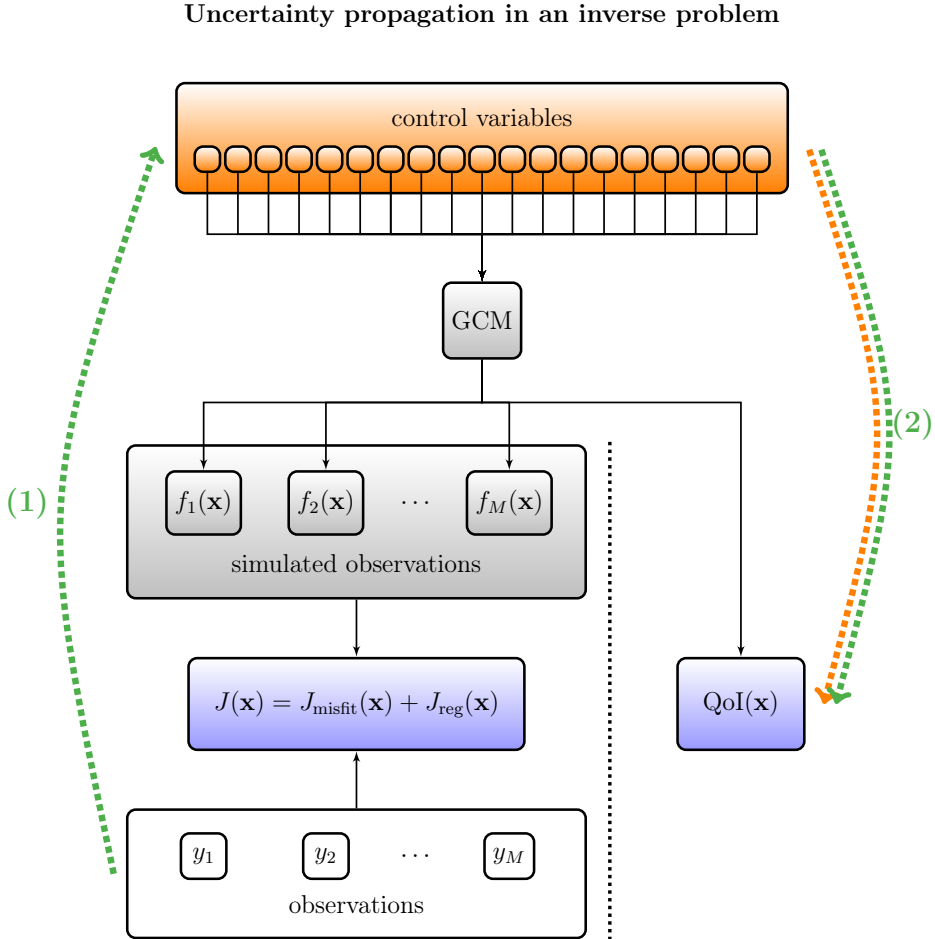


Figure 3.1: An extension of Fig. 1.8, where the workflow of Uncertainty Quantification (UQ) in an inverse problem is added to the chart. By means of the adjoint and tangent linear models of the GCM, observational information and uncertainty are dynamically propagated throughout the GCM. Starting from the observations, an *inverse uncertainty propagation* along path (1) reduces the uncertainty in the control variables. A subsequent *forward uncertainty propagation* along path (2) reduces the uncertainty in a chosen model quantity of interest (QoI). UQ-based observing system design is fundamentally dynamics-based and measures the effectiveness of a set of observations by the uncertainty reduction that is achieved by including these observations in the inverse modeling framework. For example, one can evaluate the difference between the prior and posterior uncertainties in a chosen QoI. Here, one computes the prior uncertainty in the QoI by propagating prior uncertainties (dotted orange arrow) along path (2), and the posterior uncertainty in the QoI by propagating posterior uncertainties (dotted green arrow) along path (2).

sitivities (“Level 1” in Table 1.1) for ocean observing system design has the limitations that (i) it does not operate within a non-linear inverse problem, thereby neglecting data uncertainties and prior knowledge on ocean circulation, (ii) existing observations are not taken into account, and (iii) it can only assess the value of individual observations at a single point in time, but not systems of multiple observations (see Section 1.3.3 and Fig. 1.12(a)).

A more advanced adjoint-based technique for observing system design is Hessian-based Uncertainty Quantification (UQ) in non-linear inverse problems (“Level 3” in Table 1.1). By means of the adjoint and tangent linear models of the underlying GCM (Section 1.2), observational information and uncertainty are propagated throughout the global ocean via the linearized model equations. When more observations are included in the inverse problem, an inverse uncertainty propagation along path (1) in Fig. 3.1 leads to an uncertainty reduction in the control variables of the inverse problem. A subsequent forward uncertainty propagation along path (2) in Fig. 3.1 reduces uncertainty in a chosen climate QoI. UQ is useful for observing system design because one can measure the effectiveness of a set of observations by the uncertainty reduction that is achieved by including these observations in the inverse modeling framework. The strength of the UQ method is that the computed uncertainty reduction is based purely on dynamical information that is propagated via the model equations - i.e., equations of motions representing basic conservation and constitutive laws rendered on a computer - as opposed to techniques that are based on statistical inference, e.g., EOFs. In practical terms, posterior uncertainty is captured by the curvature of the regularized model-data misfit function $J = J_{\text{misfit}} + J_{\text{reg}}$, expressed through the *Hessian matrix* of J . Fig. 3.1 shows that UQ makes crucial use of the left branch, i.e., the non-linear inverse problem (or state estimation), and connects it to the right branch if one is interested in specific QoIs; UQ therefore overcomes limitation (i). The UQ approach also overcomes limitations (ii) and (iii) since it can account for already deployed instruments and evaluate the impact of large existing or future observing systems including a myriad of observation types, taken at different points in space and time (see Section 1.3.3 and Fig. 1.12(c)). Hessian-based UQ is therefore a promising and powerful tool for dynamics-based observing system design.

Hessian-based UQ is a concept that has been around for a long time: for the field of oceanography, it was formulated by Thacker [1989] about 30 years ago. Yet, Hessian-based UQ is a computationally demanding undertaking that has so far not been realized for full-fledged oceanographic inverse problems on climate-relevant time windows. The computational challenge stems from the combination of two facts: oceanographic inverse problems involve (i) high-dimensional control spaces and (ii) complex forward GCMs that have to be integrated over long time windows. Some studies have tried to overcome challenge (i) by an a-priori-reduction of the control space (see Section 1.3.3). For instance, Kaminski and Rayner [2008]; Kaminski et al. [2015, 2018] pursue the ‘large region approach’, where atmospheric forcing and initial conditions are adjusted uniformly over large regions, rather than on a grid point basis of the underlying GCM. However, a-priori-reduced control spaces require ad-hoc choices and can lead to large aggregation errors as highlighted in Kaminski et al. [2001]. A second approach is to take advantage of the fact that the available observations typically inform only a low-dimensional ‘effective’

subspace of the high-dimensional control space. This fact can be exploited to construct low-rank approximations of the ‘misfit Hessian’ (the Hessian of J_{misfit}), by finding the misfit Hessian’s leading eigenvectors and eigenvalues [e.g., Flath et al., 2011; Bui-Thanh et al., 2012; Isaac et al., 2015]. At the same time, the leading eigenvectors provide valuable insights for observing system design because they characterize the data-informed directions within the control space, i.e., the components that are well determined by the observations. Low-rank approximations for Hessian-based UQ is a rapidly growing research topic in the computational sciences [Bui-Thanh et al., 2012]. Using advances in the computational sciences, recent progress has been made in oceanography toward low-rank approximations for Hessian-based UQ within the ECCO (Estimating the Circulation and Climate of the Ocean) inverse modeling framework [Kalmikov and Heimbach, 2014, 2018] and the ROMS (Regional Ocean Modeling System) 4D-Var data assimilation system [Moore et al., 2011a,b, 2017a].

Hessian-based UQ for observing system design is a dynamics-based and mathematically rigorous tool, but the method is rather complex and might be deemed a computational black box which lacks simple interpretations. In this chapter, I revisit the computational tool ‘Hessian-based UQ’ from a *dynamical viewpoint* and explicitly work out how it makes use of dynamical relationships between observations and remote QoIs along the paths (1) and (2) in Fig. 3.1. To this aim, I develop an alternative method to obtain the misfit Hessian’s eigenvectors and eigenvalues. Following the steps of the alternative method clears for instance up how the data-informed directions are linked to adjoint sensitivities of observed quantities. Adjoint sensitivities, in turn, capture dynamical adjustment mechanisms and pathways, and are a concept that the oceanographic community may be more familiar with than Hessian eigenvectors, since adjoint sensitivities have been used for many sensitivity analyses (branch II in Fig. 1.8 and references in Section 1.2.2). Moreover, the technique provides a recipe for how to modify the sensitivity patterns of the observed quantities in order to account for data redundancy. On the downside, the technique is only feasible for *simple observing systems* with few observations. Meanwhile, the technique can handle high-dimensional control spaces and complex forward models. Since the method still gives the *correct* (though intractable) recipe to perform UQ for complex observing systems with tens of millions of observations, the dynamical insights gained here for simple observing systems carry over to observing system design within full-fledged oceanographic inverse problems. Gaining basic understanding is a necessary step before implementing UQ-based observing system design within costly high-performance computing frameworks. Therefore, the technique is targeted at advancing *dynamical understanding*, rather than the numerics, of UQ-based observing system design. The gained insights will contribute to foster dynamics-based ocean observing system design targeted at climate monitoring.

In Section 3.2, I offer a self-contained presentation of UQ in non-linear inverse problems, the Hessian-based approximation, and the formal expression for inverse uncertainty propagation, i.e., path (1) in Fig. 3.1. In Section 3.3, I present the alternative technique to perform UQ, i.e., to compute the misfit Hessian’s eigenvectors and eigenvalues. In Section 3.4, I explain why UQ can be used to evaluate not only existing, but also *future* observations. Section 3.5 is about forward uncertainty propagation, i.e., appending path (2) to (1) in Fig. 3.1. The concept of joining paths (1) and (2) is rephrased as as-

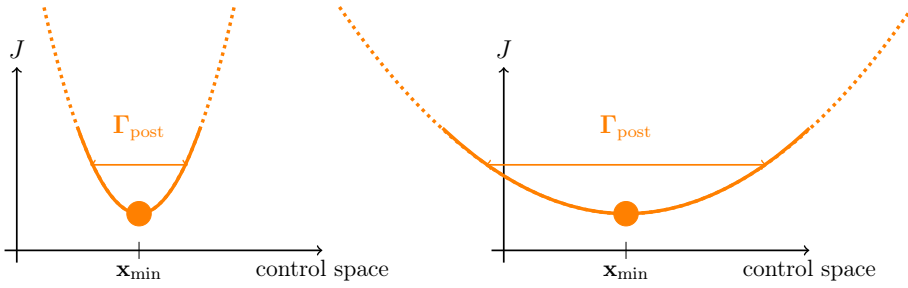


Figure 3.2: The uncertainties after adding data constraints are captured by the posterior covariance matrix Γ_{post} , which is the inverse of the Hessian matrix of the cost function J , evaluated at \mathbf{x}_{min} , the minimizer of (3.1). The Hessian matrix measures the curvature of the cost function J at \mathbf{x}_{min} . If the cost function has high curvature (or a deep valley), corresponding to small posterior uncertainty, the minimum \mathbf{x}_{min} is well-constrained (left). On the contrary, if the cost function has little curvature (or a shallow valley), corresponding to high posterior uncertainty, the minimum \mathbf{x}_{min} is poorly constrained (right).

sessing the *proxy potential* of an observing system for unobserved climate quantities of interest. I furthermore establish and explore two concepts whose combination fully characterizes proxy potentials. These two concepts, easier to grasp than the abstract concept of uncertainty propagation, clarify the roles of ocean dynamics, prior information, and observational uncertainty for determining the effectiveness of an observing system to monitor unobserved, potentially remote, QoIs. In Section 3.6, the key insights and limitations are discussed. Appendix ?? offers auxiliary computations and explanations for the reader who is interested in more technical details. While this chapter focuses on establishing the methodological framework, all methods will be illustrated with applications in Chapter 4.

3.2 Uncertainty Quantification in inverse problems

The solution to an oceanographic inverse problem is found by optimizing a high-dimensional control vector $\mathbf{x} = (\mathbf{x}_0, \mathbf{q}, \boldsymbol{\mu}) \in \mathbb{R}^N$, consisting of initial conditions \mathbf{x}_0 , boundary forcings \mathbf{q} , and model parameters $\boldsymbol{\mu}$, such as to minimize the least-squares cost function

$$J(\mathbf{x}) = \underbrace{\frac{1}{2}(\mathbf{y} - \mathbf{f}(\mathbf{x}))^T \Gamma_{\text{noise}}^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{x}))}_{J_{\text{misfit}}(\mathbf{x})} + \underbrace{\frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \Gamma_{\text{prior}}^{-1} (\mathbf{x} - \mathbf{x}_0)}_{J_{\text{reg}}(\mathbf{x})}, \quad (3.1)$$

see also Section 1.2.1. Here, \mathbf{f} is the non-linear observation operator that maps the control vector \mathbf{x} through a dynamical model onto a vector of observables. Albeit its seemingly simple expression (3.1), the optimization problem is highly complex because \mathbf{f} comprises the non-linear dynamics of a general circulation model, which means to represent the physics in the real ocean (see Fig. 1.8). The first term in (3.1), $J_{\text{misfit}}(\mathbf{x})$, measures the misfit between the vector of actual observations $\mathbf{y} = (y_1, \dots, y_M)$ and the vector of observation counterparts $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_M(\mathbf{x}))$ simulated by the model,

weighted by the matrix $\mathbf{\Gamma}_{\text{noise}}^{-1}$. Here, the vector \mathbf{y} can contain observations of different types, e.g., in situ and satellite observations, and made at different times and places. The matrix $\mathbf{\Gamma}_{\text{noise}}$ contains observational uncertainties, which are due to both measurement errors and representation errors in the model. The second term, $J_{\text{reg}}(\mathbf{x})$, penalizes deviations from a first-guess \mathbf{x}_0 , where the matrix $\mathbf{\Gamma}_{\text{prior}}^{-1}$ acts as a regularization operator [Wunsch, 1996; Tarantola, 2005].

In the geosciences, the traditional approach to inverse problems is a deterministic one, where the minimizer of the cost function (3.1), $\mathbf{x}_{\text{min}} = \min_{\mathbf{x}} J$, is considered as the “best-guess solution” of the time-evolving ocean state, with no formal uncertainty estimate for this solution. In contrast, such uncertainties can be quantified when framing the deterministic inverse problem as one of Bayesian inference. In the Bayesian context, the weight matrices $\mathbf{\Gamma}_{\text{noise}}$ and $\mathbf{\Gamma}_{\text{prior}}$ are interpreted as error covariances, assuming Gaussian observational noise $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_{\text{noise}})$ and a Gaussian prior distribution $\mathcal{N}(\mathbf{x}_0, \mathbf{\Gamma}_{\text{prior}})$. Bayes’ Theorem then computes a posterior probability distribution for the solution of the inverse problem, given by $\pi_{\text{post}}(\mathbf{x}|\mathbf{y}) \propto e^{-J(\mathbf{x})}$ (see Appendix B.1). Hence, the deterministic and Bayesian formulation of the inverse problem are interconnected by the fact that the deterministic least squares cost function J is the negative log-posterior in the Bayesian interpretation, with the deterministic solution \mathbf{x}_{min} being the Maximum a Posteriori (MAP) point, i.e., the most likely solution, in the Bayesian framework [e.g., Tarantola, 2005; Stuart, 2010; Law et al., 2015].

Despite the choice of Gaussian probability distributions for prior and noise, the posterior probability distribution $e^{-J(\mathbf{x})}$ does not need to be Gaussian, due to the non-linearity of \mathbf{f} . The non-Gaussianity of the posterior distribution poses challenges for getting a good grasp of the solution of the inverse problem - which is given by the posterior probability distribution. The method of choice for understanding a possibly complex probability distribution is to sample it by Markov chain Monte Carlo (MCMC), which allows to compute statistics of the distribution. However, for comprehensive oceanographic inverse problems, as the one solved by ECCO v4 (Section 1.2.1), MCMC or other conventional techniques are intractable. This is due to the combination of two facts; first, the high dimensionality N of the control space, typically of the order $O(10^6) - O(10^8)$ [Forget et al., 2015], and, second, the high cost of evaluating \mathbf{f} because each evaluation requires to run a general circulation model. In other words, while one can formulate a framework to quantify uncertainties, trying to execute this framework is not feasible - a common issue for inverse problems in many fields of application [Bui-Thanh et al., 2012].

To make uncertainty quantification tractable, a linearization of the non-linear observation operator \mathbf{f} about the MAP Point \mathbf{x}_{min} is necessary [e.g., Bui-Thanh et al., 2012]. This yields

$$\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\mathbf{x}_{\text{min}}) + \mathbf{A}(\mathbf{x} - \mathbf{x}_{\text{min}}), \quad (3.2)$$

where $\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}|_{\mathbf{x}_{\text{min}}}$ is the Jacobian matrix of the operation operator. This linearization yields a local Gaussian approximation $\mathcal{N}(\mathbf{x}_{\text{min}}, \mathbf{\Gamma}_{\text{post}})$ of the posterior distribution, where the error covariance matrix is given by

$$\mathbf{\Gamma}_{\text{post}} = (\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \mathbf{\Gamma}_{\text{prior}}^{-1})^{-1}, \quad (3.3)$$

(see Appendix B.2). The posterior error covariance matrix (3.3) is the inverse of $\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \mathbf{\Gamma}_{\text{prior}}^{-1}$. The latter matrix is the Hessian of the cost function J , evaluated at \mathbf{x}_{\min} , if one uses the linearized observation operator (3.2). Therefore it is also referred to as the *linearized Hessian*. Linearized vs. full Hessians are briefly discussed in Appendix B.4. The Hessian of J (whether full or linearized) contains information about the curvature of the cost function. Intuitively, if the cost function exhibits a deep valley with *high* curvature, corresponding to *small* posterior uncertainty, the solution \mathbf{x}_{\min} is much better constrained than for a relatively flat cost function (see Figure 3.2).

The Hessian of J , whose inverse emerges in (3.3), is the sum of two matrices: the ‘misfit Hessian’ $\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A}$, which is the Hessian of the model-data misfit term $J_{\text{misfit}}(\mathbf{x})$ in (3.1), and $\mathbf{\Gamma}_{\text{prior}}^{-1}$, the Hessian of the regularization term $J_{\text{reg}}(\mathbf{x})$ in (3.1). Since observations typically inform only a low-dimensional subspace of the high-dimensional control space, the high-dimensional misfit Hessian is of low rank [Bui-Thanh et al., 2012]. This means that $J_{\text{misfit}}(\mathbf{x})$ is only curved in a few directions, namely in the *data-informed directions*. On the other hand, the Hessian of $J_{\text{reg}}(\mathbf{x})$ is given by $\mathbf{\Gamma}_{\text{prior}}^{-1}$, which is chosen to be of full rank. This means that $J_{\text{reg}}(\mathbf{x})$ is curved along *all* directions. Consequently, $J_{\text{reg}}(\mathbf{x})$ curves the cost function $J(\mathbf{x})$ upward along the directions in which it otherwise would have been flat [Thacker, 1989], or, equivalently, $\mathbf{\Gamma}_{\text{prior}}^{-1}$ regularizes the Hessian in (3.3) so it becomes an invertible matrix - therefore its name “regularization term”.

The number of matrix entries of $\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \mathbf{\Gamma}_{\text{prior}}^{-1}$ is equal to N^2 , the number of control variables *squared*, which is on the order of $O(10^{12}) - O(10^{16})$ for a full-fledged oceanographic inverse problem [Forget et al., 2015]. An explicit computation of this matrix and its inverse in (3.3) would therefore require months to years of heavy computer resources, and terabytes to petabytes of storage [cf. Kalmikov and Heimbach, 2018] - clearly, an intractable effort. Nevertheless, one can exploit the low-dimensionality of the subspace that is informed by the data: the prior-preconditioned (or non-dimensionalized) misfit Hessian

$$\tilde{\mathbf{H}}_{\text{misfit}} = \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} \in \mathbb{R}^{N \times N}, \quad (3.4)$$

which is of the same (low) rank as the misfit Hessian, can be re-written as its eigen-decomposition

$$\tilde{\mathbf{H}}_{\text{misfit}} = \sum_{i=1}^{M'} \lambda_i \mathbf{v}_i \mathbf{v}_i^T. \quad (3.5)$$

For typical inverse problems, the eigenvalues of $\tilde{\mathbf{H}}_{\text{misfit}}$ decay rapidly, and only the M' largest (or leading) eigenvalues and corresponding eigenvectors are retained [e.g., Bui-Thanh et al., 2012]. These leading eigenvectors are the most important ones because they characterize the directions within the control space that are best informed by the data. The equality in (3.5) is then an approximation, a so-called *low-rank approximation*. For small sets of observations, as considered in this work, M' is equal to the number of observations, M , (unless a subset of the observations is completely superfluous), and the equality in (3.5) is exact. In either case, M' is small compared to N , the dimension of the control space.

The matrix inversion lemma¹ provides an efficient way to compute the inverse of the sum of a low-rank and a full-rank matrix, and the posterior covariance matrix can then be written as

$$\begin{aligned}\mathbf{\Gamma}_{\text{post}} &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \left(\tilde{\mathbf{H}}_{\text{misfit}} + \mathbf{I} \right)^{-1} \mathbf{\Gamma}_{\text{prior}}^{1/2} \\ &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \left(\mathbf{I} - \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1} \mathbf{v}_i \mathbf{v}_i^T \right) \mathbf{\Gamma}_{\text{prior}}^{1/2} \\ &= \mathbf{\Gamma}_{\text{prior}} - \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1} \left(\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{v}_i \right) \left(\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{v}_i \right)^T.\end{aligned}\quad (3.6)$$

Equation (3.6) describes uncertainty reduction for the control variables, which is achieved by the inverse uncertainty propagation (1) in Fig. 3.1; it phrases the posterior uncertainty as the prior uncertainty $\mathbf{\Gamma}_{\text{prior}}$, less any information obtained from the observations, filtered through the prior.

Computing the eigen-decomposition (3.5) of the *prior-preconditioned* misfit Hessian $\tilde{\mathbf{H}}_{\text{misfit}}$, rather than the misfit Hessian $\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A}$, corresponds to non-dimensionalizing the control problem, i.e., minimizing $J(\mathbf{u})$ with $\mathbf{u} = \mathbf{\Gamma}_{\text{prior}}^{-1/2} \mathbf{x}$, rather than $J(\mathbf{x})$ in (3.1) - as done in practical ocean inverse problems [Forget et al., 2015]. In fact, the prior-preconditioned misfit Hessian $\tilde{\mathbf{H}}_{\text{misfit}}$ (3.4) is the Hessian of the non-dimensionalized model-data misfit term in (3.1), i.e., the Hessian of $J_{\text{misfit}}(\mathbf{u})$. Accordingly, the leading eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_{M'}$ in eigen-decomposition (3.5) are the directions along which $J_{\text{misfit}}(\mathbf{u})$ is the most curved, with respective curvatures $\lambda_1, \dots, \lambda_{M'}$. I therefore refer to the ‘prior-preconditioned misfit Hessian’, a term used in the computational sciences [e.g., Flath et al., 2011; Bui-Thanh et al., 2012], as the *non-dimensionalized misfit Hessian*. Non-dimensionalizing (or prior-preconditioning) the Hessian before computing its eigen-decomposition in (3.5) is a step that is not only done for purely numerical reasons, i.e., for increasing the condition number of the matrix. In fact, in an oceanographic inverse problem, it is a necessary step due to the heterogeneity of the control variables. Extracting the leading eigenvectors and eigenvalues of the misfit Hessian without non-dimensionalizing it would disfavor certain data-informed directions within the control space. For instance, in the ECCO v4 framework, data-informed components of the control vector \mathbf{x} associated with precipitation would be disfavored compared to data-informed components associated with downward radiation. This is because, before non-dimensionalization, adjustments of precipitation are typically on the order of $O(10^{-7})$ [m/s] or less, and therefore of much smaller amplitude than adjustments of downward radiation, which are typically on the order of $O(10^1)$ - $O(10^2)$ [W/m²]. Therefore, a low-rank approximation and analysis of the ‘pure’ misfit Hessian $\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A}$, totally separated from the chosen prior $\mathbf{\Gamma}_{\text{prior}}$, is not meaningful if the control space includes physical variables of different orders of magnitude.

¹ The matrix inversion lemma (or Sherman-Morrison-Woodbury formula) states that

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{U} (\mathbf{C}^{-1} + \mathbf{VA}^{-1} \mathbf{U})^{-1} \mathbf{VA}^{-1}.$$

3.3 Computation of data-informed directions and curvatures

The data-informed directions $\mathbf{v}_1, \dots, \mathbf{v}_{M'}$ and their associated curvatures $\lambda_1, \dots, \lambda_{M'} > 0$ govern uncertainty reduction in (3.6). They are extracted by an eigen-decomposition (3.5) of the non-dimensionalized misfit Hessian. Even though hundreds of millions of ocean observations are included in oceanographic inverse problems such as ECCO v4 (Section 1.2.1), it is typical for inverse problems that the eigenvalues of $\tilde{\mathbf{H}}_{\text{misfit}}$ decay rapidly, such that the data-informed directions are effectively given by only the leading eigenvectors and eigenvalues of $\tilde{\mathbf{H}}_{\text{misfit}}$ [Bui-Thanh et al., 2012]. It is unclear how many eigenvectors have to be retained for oceanographic inverse problems - obtaining these eigenvectors is computationally so demanding for high-dimensional control spaces (large N) and complex forward models that it has not yet been done for full-fledged global oceanographic inverse problems over climate-relevant time windows. Nevertheless, progress has been made within the ECCO framework [Kalmikov and Heimbach, 2014, 2018] (although still for reduced control spaces and shorter time windows), by exploiting matrix-free Lanczos methods, developed in the computational science and engineering community.

In this subsection, I develop an alternative method to compute the eigenvectors and eigenvalues of the non-dimensionalized misfit Hessian in (3.5). This method uses only simple linear algebra operations, under the assumption that \mathbf{A} , the Jacobian matrix of the observation operator \mathbf{f} (evaluated at \mathbf{x}_{\min}), is available. Note that \mathbf{A} is sometimes also called the sensitivity matrix [e.g., Thacker, 1989]. For an observing system with M observations, given by the observation vector $\mathbf{y} = (y_1, \dots, y_M)^T$, the observation operator is given by $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_M(\mathbf{x}))^T$, where $f_i(\mathbf{x})$ is a scalar function and describes the simulated counterpart of observation y_i . Each function $f_i(\mathbf{x})$ describes how a change in the control variables \mathbf{x} affects the simulated observation through the model dynamics (see Fig. 3.1). For instance, $f_i(\mathbf{x})$ might be the temperature in a certain model grid cell averaged over multiple model time steps, or the horizontal velocity at a certain point in time averaged over multiple model grid cells (see also Fig. 1.8). The assumption that \mathbf{A} is available is therefore equivalent to

Assumption* For each $i = 1, \dots, M$, the sensitivity $\left[\frac{\partial f_i}{\partial \mathbf{x}}\right]_{\mathbf{x}_{\min}}^T \in \mathbb{R}^N$ of the observed quantity f_i to the full control vector \mathbf{x} , is available.

For each $i = 1, \dots, M$, the sensitivity vector $\left[\frac{\partial f_i}{\partial \mathbf{x}}\right]_{\mathbf{x}_{\min}}^T \in \mathbb{R}^N$ can be obtained by a single adjoint model integration (see Fig. 1.9), if one defines the scalar quantity of interest in box C.II to be $\text{QoI}(\mathbf{x}) = f_i(\mathbf{x})$, following the right branch in Fig. 1.8. Thus, the assumption is met, if one performs M adjoint sensitivity experiments. The method is therefore feasible for *simple observing systems*, for which M , the number of observations, is small. On the other hand, a complex forward model and a high-dimensional control space are no limiting factors for the developed technique, due to the efficacy of the adjoint, which computes sensitivities to the full control vector in a single run (Fig. 1.9). For complex observing systems with millions of observations, the method still gives the correct recipe to obtain the eigenvectors and eigenvalues in (3.5). However, it would be computationally intractable to meet its assumption, i.e., performing M adjoint sensitivity experiments. Therefore, the method is not targeted at advancing computational

methods, but rather a dynamical *understanding* of UQ for observing system design. For instance, the method will establish an explicit link of the data-informed directions \mathbf{v}_i in (3.5) to adjoint sensitivities of observed quantities. Adjoint sensitivities, in turn, capture dynamical adjustment mechanisms and pathways of the observed quantities, as, e.g., exploited in Chapter 2.

3.3.1 A single observation

To gain insight into what determines the data-informed directions and associated curvatures, it is helpful to first consider the simplest case. I assume $M = 1$, i.e., that only a single observation y_1 is available, with Gaussian noise $\mathcal{N}(0, \varepsilon^2)$. The observation operator is then given by a scalar function $f_1(\mathbf{x})$, and the error covariance matrix $\mathbf{\Gamma}_{\text{noise}}$ is equal to the scalar ε^2 . Equation (3.4) reads as the eigen-decomposition $\tilde{\mathbf{H}}_{\text{misfit}} = \lambda_1 \mathbf{v}_1^T \mathbf{v}_1$, with

$$\mathbf{v}_1 = \frac{\mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial f_1}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}}^T}{\left\| \mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial f_1}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}}^T \right\|} \in \mathbb{R}^N \quad \text{and} \quad \lambda_1 = \frac{1}{\varepsilon^2} \left\| \mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial f_1}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}}^T \right\|^2. \quad (3.7)$$

The vector $\left[\frac{\partial f_1}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}}^T \in \mathbb{R}^N$ in (3.7) is available by **Assumption***; it contains the sensitivity of the observed quantity f_1 to the control vector \mathbf{x} . This vector is *prior-weighted*, i.e., multiplied by the matrix $\mathbf{\Gamma}_{\text{prior}}^{1/2} \in \mathbb{R}^{N \times N}$. The prior-weighted observation sensitivity vector contains the sensitivity of f_1 to the non-dimensionalized control vector $\mathbf{u} = \mathbf{\Gamma}_{\text{prior}}^{-1/2} \mathbf{x} \in \mathbb{R}^N$. If the prior covariance matrix $\mathbf{\Gamma}_{\text{prior}}$ is diagonal, the i^{th} entry of the prior-weighted observation sensitivity vector can be interpreted as the approximated (linearized) response of the observable f_1 to a perturbation of the i^{th} control variable by its prior standard deviation $\sqrt{\mathbf{\Gamma}_{\text{prior}}}_{ii}$. Finally, one obtains \mathbf{v}_1 by dividing the prior-weighted sensitivity vector by its Euclidean norm $\|\cdot\|^2$. \mathbf{v}_1 is unitless and of (Euclidean) norm 1. It describes the only direction within the control space that is constrained by the observation y_1 , i.e., the only direction along which $J_{\text{misfit}}(\mathbf{u})$, the model-data misfit term in the non-dimensionalized cost function, is curved.

The corresponding curvature λ_1 in (3.7) is the ratio of the squared amplitude of the prior-weighted observation sensitivity to the noise variance ε^2 . The numerator, i.e., the squared amplitude of the prior-weighted observation sensitivity, is the prior uncertainty in f_1 . The prior uncertainty in f_1 is obtained by projecting the prior covariance matrix $\mathbf{\Gamma}_{\text{prior}}$ onto f_1 . This projection corresponds to a forward uncertainty propagation along path (2) in Fig. 3.1, where f_1 takes on the role of the quantity of interest (QoI).³ The curvature λ_1 describes therefore a *prior-to-noise ratio* of the observed quantity f_1 , i.e., the quality of prior information on the observed quantity - determining the *relevance* of

² The Euclidean norm of a vector $\mathbf{v} = (v_1, \dots, v_N) \in \mathbb{R}^N$ is given by $\|\mathbf{v}\| := \sqrt{\sum_{i=1}^N v_i^2}$.

³ Forward uncertainty propagation from the control variables to scalar model quantities is discussed in Appendix B.3. The fact that $\left\| \mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial f_1}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}}^T \right\|^2$ is approximately the projection of $\mathbf{\Gamma}_{\text{prior}}$ onto f_1 are equations (B.8)-(B.10), where $f_1(\mathbf{x})$ replaces QoI(\mathbf{x}).

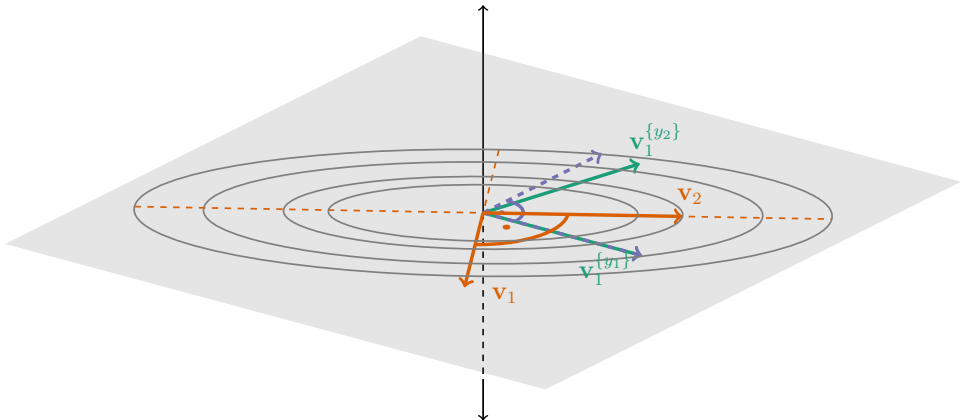


Figure 3.3: Schematics showing the concept of **data-informed directions** for the case of *two observations* $\{y_1, y_2\}$. The gray plane is the data-informed two-dimensional subspace within the control space, and the vertical black line represents its orthogonal complement, i.e., the uninformed directions. The ellipses sketch the orientation of the contour lines of $J_{\text{misfit}}(\mathbf{u})$, the model-data misfit term in the non-dimensionalized cost function. The dashed orange lines indicate the orthogonal directions of maximal curvature λ_1 and minimal curvature λ_2 of $J_{\text{misfit}}(\mathbf{u})$ within the data-informed subspace. The orange unit vectors point into these directions of maximal and minimal curvature; these vectors represent \mathbf{v}_1 and \mathbf{v}_2 , the eigenvectors of the non-dimensionalized misfit Hessian (3.4) (i.e., the Hessian of $J_{\text{misfit}}(\mathbf{u})$), with corresponding eigenvalues λ_1 and λ_2 . The schematic shows how \mathbf{v}_1 and \mathbf{v}_2 can be computed via steps (i) and (ii). The starting point are the two solid green vectors $\mathbf{v}_1^{\{y_1\}}$ and $\mathbf{v}_1^{\{y_2\}}$; these vectors are the directions that are informed by each of the two observations individually, given by the normalized prior-weighted observation sensitivities (3.7). Step (i) removes any redundant information by orthonormalizing the solid green vectors; the result are the two dashed purple vectors. Step (ii) rotates the purple vectors onto the orange orthogonal vectors \mathbf{v}_1 and \mathbf{v}_2 .

the new data - versus the *quality* of the new data.

The curvature λ_1 is large when (a) the observed quantity f_1 has high sensitivity to the control variables, (b) the prior uncertainties in the control variables are large, and (c) the observational noise is small.

Examples and illustrations for the data-informed direction \mathbf{v}_1 and its associated curvature λ_1 in the case of a single observation are shown in Fig. 4.5.

3.3.2 Multiple observations

I now assume that a set of multiple observations is available, corresponding to $\mathbf{y} = (y_1, \dots, y_M)^T$ and associated observation operator $\mathbf{f}(x) = (f_1, \dots, f_M)^T(\mathbf{x})$. The eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_{M'}$ and corresponding eigenvalues $\lambda_1, \dots, \lambda_{M'} > 0$ (where $M' \leq M$) of the non-dimensionalized misfit Hessian (3.4) can be computed by performing the follow-

ing two steps:

- (i) QR decomposition of $\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T$: Find⁴ an orthonormal matrix $\mathbf{Q} \in \mathbb{R}^{N \times M}$ and an upper triangular matrix $\mathbf{R} \in \mathbb{R}^{M \times M}$ such that

$$\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T = \mathbf{Q}\mathbf{R}. \quad (3.8)$$

Note that \mathbf{A}^T , the matrix that contains the sensitivity vectors $\left[\frac{\partial f_i}{\partial \mathbf{x}} \right]_{\mathbf{x}_{\min}}^T$, is available by **Assumption***. Furthermore, left-multiplication of \mathbf{A}^T by $\mathbf{\Gamma}_{\text{prior}}^{1/2}$ is assumed to be easy since, in practice, $\mathbf{\Gamma}_{\text{prior}}$ is chosen of sparse (or even diagonal) structure.

- (ii) Eigenvalue decomposition in \mathbb{R}^M : Find⁵ an orthogonal matrix $\mathbf{O} \in \mathbb{R}^{M \times M}$ and $\lambda_1 \geq \dots \geq \lambda_M \geq 0$ such that

$$\mathbf{R} \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{R}^T = \mathbf{O} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \lambda_M \end{pmatrix} \mathbf{O}^T. \quad (3.9)$$

Then

$$\begin{aligned} \tilde{\mathbf{H}}_{\text{misfit}} &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} = \mathbf{Q} \mathbf{R} \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{R}^T \mathbf{Q}^T \\ &= \mathbf{Q} \mathbf{O} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \lambda_M \end{pmatrix} \mathbf{O}^T \mathbf{Q}^T, \end{aligned}$$

and the i^{th} column of $\mathbf{Q}\mathbf{O}$ contains the i^{th} eigenvector of $\tilde{\mathbf{H}}_{\text{misfit}}$, with corresponding eigenvalue $\lambda_i \geq 0$. The eigenvectors corresponding to non-zero eigenvalues are the data-informed directions $\mathbf{v}_1, \dots, \mathbf{v}_{M'}$. It is noted that the steps (i) and (ii) come at negligible computational cost. Dense matrix algebra is only performed within \mathbb{R}^M , where M , the

⁴ Denoting $\mathbf{b}_i = \mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial f_i}{\partial \mathbf{x}} \right]_{\mathbf{x}_{\min}}^T$, the QR decomposition of $\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T = (\mathbf{b}_1 \mid \dots \mid \mathbf{b}_M)$ can be computed with the Gram-Schmidt process:

- $\tilde{\mathbf{w}}_1 := \mathbf{b}_1, \quad \mathbf{w}_1 = \|\tilde{\mathbf{w}}_1\|^{-1} \cdot \tilde{\mathbf{w}}_1$
- For $j = 2, \dots, M$: $\tilde{\mathbf{w}}_j = \mathbf{b}_j - \sum_{i=1}^{j-1} \langle \mathbf{b}_j, \mathbf{w}_i \rangle \mathbf{w}_i, \quad \mathbf{w}_j = \|\tilde{\mathbf{w}}_j\|^{-1} \cdot \tilde{\mathbf{w}}_j$

Then the matrices

$$\mathbf{Q} := (\mathbf{w}_1 \mid \dots \mid \mathbf{w}_M) \in \mathbb{R}^{N \times M} \quad \text{and} \quad \mathbf{R} = \begin{pmatrix} \|\tilde{\mathbf{w}}_1\| & \langle \mathbf{w}_1, \mathbf{b}_2 \rangle & \langle \mathbf{w}_1, \mathbf{b}_3 \rangle & \dots \\ 0 & \|\tilde{\mathbf{w}}_2\| & \langle \mathbf{w}_2, \mathbf{b}_3 \rangle & \dots \\ 0 & 0 & \|\tilde{\mathbf{w}}_3\| & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \in \mathbb{R}^{M \times M}$$

are as desired in (3.8).

⁵ Since $\mathbf{\Gamma}_{\text{noise}}$ is a covariance matrix, the matrix $\mathbf{R} \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{R}^T$ is symmetric and positive semi-definite. Therefore the decomposition in (3.9) exists. It can be computed with dense matrix algebra since M is assumed to be small here.

number of observations, is assumed to be small.

The steps (i) and (ii) reflect how the directions $\mathbf{v}_1, \dots, \mathbf{v}_{M'}$ that are informed by a combined set $\{y_1, \dots, y_M\}$ of observations are related to the M directions that are informed by each observation $\{y_i\}$ individually. To demonstrate this, let me assume that $M = 2$. Equation (3.7) (applied twice) computes the directions that are informed by $\{y_1\}$ and $\{y_2\}$ individually, and I denote these directions by $\mathbf{v}_1^{\{y_1\}}$ and $\mathbf{v}_1^{\{y_2\}}$, with corresponding eigenvalues $\lambda_1^{\{y_1\}}$ and $\lambda_1^{\{y_2\}}$. The vectors $\mathbf{v}_1^{\{y_1\}}$ and $\mathbf{v}_1^{\{y_2\}}$ span either a one⁶ - or two-dimensional subspace within the control space. Schematic 3.3 sketches the case in which $\mathbf{v}_1^{\{y_1\}}$ and $\mathbf{v}_1^{\{y_2\}}$ (green vectors), span a *two*-dimensional subspace (the gray plane). The directions \mathbf{v}_1 and \mathbf{v}_2 (orange vectors) that are informed by the combined set $\{y_1, y_2\}$ span the same two-dimensional subspace. However, steps (i) and (ii) modify $\mathbf{v}_1^{\{y_1\}}$ and $\mathbf{v}_1^{\{y_2\}}$ in the following way to generate \mathbf{v}_1 and \mathbf{v}_2 : The vector pair $(\mathbf{v}_1^{\{y_1\}}, \mathbf{v}_1^{\{y_2\}})$ is

- (i) orthonormalized, which corresponds to *removal of redundant information*, and
- (ii) rotated and/or reflected, such that \mathbf{v}_1 points into the direction of maximal curvature of $J_{\text{misfit}}(\mathbf{u})$ (the model-data misfit term in the non-dimensionalized cost function).

Examples and illustrations for the data-informed directions $\mathbf{v}_1, \mathbf{v}_2$ and its associated curvatures λ_1, λ_2 in the case of two observations will be given in Fig. 4.7. Moreover, Section 4.3.2 will delineate the close link between the vector pairs $(\mathbf{v}_1^{\{y_1\}}, \mathbf{v}_1^{\{y_2\}})$ and $(\mathbf{v}_1, \mathbf{v}_2)$.

3.3.3 Testing various prior and noise matrices

Table 3.1 provides an overview of the presented method for computing the eigenvectors and eigenvalues $\{(\mathbf{v}_i, \lambda_i)\}$ of the non-dimensionalized misfit Hessian (3.4), which are the key to quantifying uncertainty. The check marks in the last two columns show which steps depend on the choice of the prior and noise covariance matrices $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$. In step 0, M adjoint model integrations are performed to obtain the Jacobian $\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}|_{\mathbf{x}_{\text{min}}}$ of the observation operator. The observation operator $\mathbf{f} = (f_1, \dots, f_M)^T$ contains the observed quantities, as simulated by the model, and does not depend on the choice of the prior and noise matrices. Meanwhile, \mathbf{x}_{min} , the point where the Jacobian \mathbf{A} is evaluated, *does* depend on these matrices because \mathbf{x}_{min} is the minimizer of the cost function (3.1), in which $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ are used as weight matrices. If the observation operator \mathbf{f} was linear, however, its Jacobian \mathbf{A} would not depend on the evaluation point \mathbf{x}_{min} . Therefore the dependence of the computations in Step 0 on $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ is assumed to be minor (indicated by the *parentheses* around the check marks in the first line in

⁶ If $\mathbf{v}_1^{\{y_1\}}$ and $\mathbf{v}_1^{\{y_2\}}$ span a one-dimensional subspace, i.e., $\mathbf{v}_1^{\{y_1\}} = \pm \mathbf{v}_1^{\{y_2\}}$, steps (i) and (ii) result in $\lambda_2 = 0$. This means that the observations $\{y_1, y_2\}$ inform only a one-dimensional subspace, spanned by $\mathbf{v}_1 = \pm \mathbf{v}_1^{\{y_1\}}$. Combining the observations y_1 and y_2 can still lead to an information gain because the curvature λ_1 along \mathbf{v}_1 can be larger compared to the curvatures $\lambda^{\{y_i\}}$. For instance, if y_1 and y_2 are independent observations (i.e., $(\mathbf{\Gamma}_{\text{noise}})_{12} = (\mathbf{\Gamma}_{\text{noise}})_{21} = 0$), the new curvature is given by $\lambda_1 = \lambda^{\{y_1\}} + \lambda^{\{y_2\}}$. This is the sum of the curvatures that would be achieved when including only one of the observations, respectively.

Step	Type	Cost	Computed Quantity	Use $\mathbf{\Gamma}_{\text{prior}}$	Use $\mathbf{\Gamma}_{\text{noise}}$
0	run adjoint model(s)	costly	$\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} _{\mathbf{x}_{\text{min}}}$	(✓)	(✓)
(i)	post-processing	cheap	$\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T$	✓	
(ii)	post-processing	cheap	$\mathbf{R} \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{R}^T$		✓

Table 3.1: Overview of the presented method for computing the eigenvectors and eigenvalues $\{(\mathbf{v}_i, \lambda_i)\}$ of the non-dimensionalized misfit Hessian (3.4) for simple observing systems. The check marks in the last two columns show which steps depend on the choice of the prior and noise covariance matrices $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$. The parantheses around the check marks for Step 0 indicate that the computed quantities in this step are assumed to depend only in a minor way on $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ (see text).

Table 3.1). The dependence of linearized sensitivities on the evaluation point (or base state) is discussed by Heimbach et al. [2011], who conclude that the large-scale features of computed sensitivities are robust across different base states, and even across different model configurations, while sensitivity amplitudes might slightly differ. While the dependence of step 0 on $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ is assumed to be minor, $\mathbf{\Gamma}_{\text{prior}}$ enters the computation crucially in step (i), and $\mathbf{\Gamma}_{\text{noise}}$ is used crucially for step (ii). Steps (i) and (ii) are cheap *offline post-processing steps*, which do not require the integration of a forward or adjoint model. Thus, once step 0 is performed, the obtained Jacobian can be used to cheaply compute $\{(\mathbf{v}_i, \lambda_i)\}$ for various choices of $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$, by re-performing steps (i) and (ii).

Notice that computing $\{(\mathbf{v}_i, \lambda_i)\}$ with other, conventional (e.g., Lanczos), methods does not allow a systematic assessment of the roles of the prior and noise matrices because $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ will be tangled up in the solution algorithm for the computation of the Hessian [e.g., Flath et al., 2011; Moore et al., 2011a,b; Bui-Thanh et al., 2012; Kalmikov and Heimbach, 2014; Isaac et al., 2015]. While the here presented method is only feasible for simple observing systems with few observations, it can handle high-dimensional control spaces and complex forward models, and might therefore provide a useful testbed to gain more insight into the roles of $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ for posterior error estimates.

3.4 Future observations

Observations inject their information content into the UQ framework via inverse uncertainty propagation along path (1) in Fig. 3.1. Inverse uncertainty propagation is formalized by identity (3.6). Inspecting this identity, the ingredients are the prior covariance $\mathbf{\Gamma}_{\text{prior}}$, and $\{(\mathbf{v}_i, \lambda_i)\}$, the eigenvectors and corresponding eigenvalues of the misfit Hessian (3.4). The misfit Hessian and $\{(\mathbf{v}_i, \lambda_i)\}$ depend in turn solely on

- prior information (through the chosen matrix $\mathbf{\Gamma}_{\text{prior}}$),
- the observation locations, times, and types (through the computed matrix $\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}|_{\mathbf{x}_{\text{min}}}$, the linearized sensitivities of the simulated observation counterparts),
- observational noise (through the chosen matrix $\mathbf{\Gamma}_{\text{noise}}$),

but *not on the actual measurement values* of the observations. This fact became also apparent when I derived explicit expressions for $\{(\mathbf{v}_i, \lambda_i)\}$ in Section 3.3: The data-informed directions \mathbf{v}_i are (orthonormalized and rotated) prior-weighted sensitivities of the simulated observation counterparts, which reflect how the observation locations and types are connected to other variables of the model ocean. The eigenvalues λ_i can be understood as prior-to-noise ratios. The actual values $y_1, \dots, y_{M_{\text{obs}}}$ of the observations do not appear in the expressions for $\{(\mathbf{v}_i, \lambda_i)\}$. Seen from a mathematical point of view, the fact that the presented UQ framework is independent of the actual measurement values is due to utilizing the *linearized* Hessian instead of full Hessian. Using linearized Hessians is a popular approach in UQ and observing system design [e.g., Moore et al., 2011a, 2012, 2011b; Bui-Thanh et al., 2012; Kaminski et al., 2015, 2018]. Linearized vs. full Hessians are discussed in detail in Appendix B.4.

Since no actual measurement values are needed, the UQ framework can evaluate not only existing observing systems, but also hypothetical, future observations. Uncertainty quantification in inverse problems can therefore be used as a guiding tool for *observing system design*. Moore et al. [2017a] compare the fact that UQ-based observing system design requires *only* $\{(\mathbf{v}_i, \lambda_i)\}$, i.e., the eigenvectors and eigenvalues of the linearized Hessian, but *not* the actual measurement values, to principles in antenna theory. An antenna array that is subjected to an electric or magnetic field resonates at characteristic mode frequencies, corresponding to $\{(\mathbf{v}_i, \lambda_i)\}$. For the design of wireless devices, the characteristic modes of different antenna configurations can be optimized to provide maximum coverage at specific frequencies. An ocean observing system can be viewed as an antenna array, except that it is stimulated by certain ocean circulation patterns $\{(\mathbf{v}_i, \lambda_i)\}$, rather than electric or magnetic fields.

3.5 Proxy potential

Matrix identity (3.6) encompasses a massive amount of information; the identity describes how the considered observations cause uncertainty reduction for each control variable of the full inverse problem, through inverse uncertainty propagation along path (1) in Fig. 3.1. However, in ocean climate research, one does typically not seek to understand the observational constraints on the model control variables (such as initial and boundary conditions), but rather on unobserved climate quantities of interest (QoIs), which can be computed from the model ocean state. A QoI may for example be the maximum AMOC, ocean heat content in a specific ocean basin, the Niño 3.4 SST Index, or any other diagnostic (or prognostic) scalar variable. A corresponding scalar function $\text{QoI}(\mathbf{x})$ describes how a change in the control variables \mathbf{x} affects the QoI through the model dynamics (see box C.II in Fig. 1.8). Observational constraints on such QoIs can be assessed by an inverse uncertainty propagation along path (1) and a subsequent forward uncertainty propagation along path (2) in Fig. 3.1.

Practically, appending path (2) to path (1) in Fig. 3.1 means that one studies the curvature of the (non-dimensionalized) cost function along *one* target direction, associated with the QoI, rather than along *all* directions within the control space. Here, the

target direction is the direction within the control space that needs to be constrained in order to inform the QoI. It is given by the normalized, prior-weighted sensitivity of the QoI to the full control vector (see Appendix B.3), i.e.,

$$\mathbf{q} := \frac{\mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\text{min}}}^T}{\left\| \mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\text{min}}}^T \right\|}. \quad (3.10)$$

The architecture of the target direction in (3.10) can be thought of as follows. The response of the QoI is a superposition of many signals that are propagated throughout the global ocean by different dynamical mechanisms, e.g., by advective and diffusive processes, or barotropic and baroclinic waves. Therefore, to fully constrain the QoI, all processes that can transport signals to the QoI have to be ‘‘controlled’’. This is why the sensitivities $\frac{\partial(\text{QoI})}{\partial \mathbf{x}} = \left(\frac{\partial(\text{QoI})}{\partial x_1}, \dots, \frac{\partial(\text{QoI})}{\partial x_N} \right)$ compose the target direction (3.10); they capture all possible dynamical mechanisms via which small-amplitude perturbations in the control variables can change the QoI.

3.5.1 Relative uncertainty reduction

Forward propagation of prior uncertainties (dotted orange arrow) and posterior uncertainties (dotted green arrow) along path (2) in Fig. 3.1 is achieved by projecting the prior and posterior error covariance matrices $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{post}}$ onto the QoI (see Appendix B.3). These projections result in the prior and posterior variances of the QoI, denoted by σ_{prior}^2 and σ_{post}^2 , respectively. Their reciprocals $\sigma_{\text{prior}}^{-2}$ and $\sigma_{\text{post}}^{-2}$ are equal to the curvature of the non-dimensionalized cost function along the target direction, before and after adding the considered set of observations, respectively. Due to the observational constraints that are propagated through the model dynamics, σ_{post}^2 is smaller than σ_{prior}^2 . This means that uncertainty (or variance) gets reduced, or, equivalently, curvature gets increased. The relative uncertainty (or variance) reduction

$$\tilde{\Delta}\sigma^2 := \frac{\sigma_{\text{prior}}^2 - \sigma_{\text{post}}^2}{\sigma_{\text{prior}}^2} \in [0, 1] \quad (3.11)$$

is a measure for the effectiveness of the considered observing system in informing the QoI.

I will also refer to the measure (3.11) as the *proxy potential* of the observing system for the QoI. The first term, ‘proxy’, is used because (3.11) targets unobserved QoIs or climate signals, possibly remote from the observations. Therefore, while the QoIs might be non-observable, (3.11) evaluates how the observing system informs the QoIs via dynamical relationships in the model. This may be compared to the concept of proxies in paleoceanography, which are used to infer information about (clearly non-observable) ocean circulation in the past. While inferences from paleo proxies are often based on statistically derived transfer functions, here, inferences made from ‘proxies’ are purely dynamics-based. The second term, *potential*, is chosen because evaluating the change in QoI *uncertainty* as in (3.11), rather than the change in the QoI *estimate* itself, explores the general *capability* or *potential* for the observations to inform the QoIs. This

assessment does not require the actual measurement values, as discussed in Section 3.4. Meanwhile, if proxy potential is identified by an effective uncertainty reduction (3.11), having the actual measurement values available will allow a better constrained *estimate* of the QoI.

$\tilde{\Delta}\sigma^2 = 0$ represents the case $\sigma_{\text{post}}^2 = \sigma_{\text{prior}}^2$, when the considered observing system does not add any information for the QoI, and has therefore *no* proxy potential for the QoI. The other extreme is $\tilde{\Delta}\sigma^2 = 1$, which corresponds to $\sigma_{\text{post}}^2 = 0$, i.e., a perfectly constrained QoI by the observing system. Relative uncertainty reduction, as defined in (3.11), is commonly used in the computational sciences as a scalar UQ measure to evaluate observational constraints on specific QoIs; it is known under the term ‘C-Optimality’ [e.g., Chaloner and Verdinelli, 1995]. It is noted that using the scalar measure does not simplify the UQ problem computationally: the calculation of the Hessian eigen-decomposition is still necessary for moving along path (1) in Fig. 3.1. This is a preliminary step for projecting the posterior covariance onto the QoI, i.e., to continue along path (2) in Fig. 3.1. Meanwhile, being a scalar instead of a high-dimensional matrix, the introduced measure *does* simplify the presentation and interpretation of the results.

I will now investigate the measure for proxy potential further, using the insights from Section 3.3. By means of identity (3.6), the relative uncertainty reduction (3.11) for a QoI can be re-written as

$$\tilde{\Delta}\sigma^2 = \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1} \langle \mathbf{q}, \mathbf{v}_i \rangle^2, \quad (3.12)$$

as derived in Appendix B.3. Here, $\{(\mathbf{v}_i, \lambda_i)\}$ are the eigenvectors and corresponding eigenvalues of the non-dimensionalized misfit Hessian (3.5), \mathbf{q} is the target direction, defined in (3.10), and $\langle \cdot, \cdot \rangle$ the Euclidean inner product in the control space, i.e., \mathbb{R}^N . Note the similarity of expression (3.10), representing the target direction \mathbf{q} , to the expressions that represent the data-informed directions $\{\mathbf{v}_i\}$. For the simple case of a single observation, \mathbf{v}_1 is given by (3.7), the normalized prior-weighted sensitivity of the observation to the full control vector. This is exactly the same expression as (3.10), except that the model operator QoI(\mathbf{x}) is substituted by the observation operator $f_1(\mathbf{x})$. For the case of multiple observations, the data-informed directions $\{\mathbf{v}_i\}$ are normalized prior-weighted observation sensitivities, too, but in an orthogonalized and rotated fashion (Section 3.3). The fact that the data-informed directions $\{\mathbf{v}_i\}$ and the target direction \mathbf{q} are closely related to adjoint sensitivities emphasizes that the measure (3.12) has its roots in dynamical principles.

Relative uncertainty reduction (3.12), reflecting the proxy potential of an observing system for an unobserved QoI, is characterized by a sum of positive numbers, where each summand is determined by two factors:

- F.1** The squared inner product $0 \leq \langle \mathbf{q}, \mathbf{v}_i \rangle^2 \leq 1$, related to the *hypothetical proxy potential*, which is the proxy potential of the observing system for the QoI in the hypothetical case of noise-free observations. The hypothetical proxy potential is

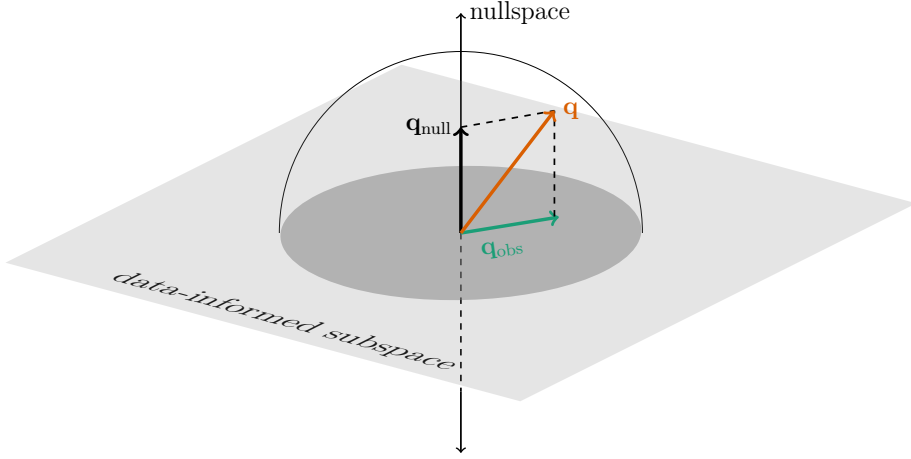


Figure 3.4: Schematic showing the decomposition (3.13) of the target direction \mathbf{q} into the sum of its informed component \mathbf{q}_{obs} and nullspace component \mathbf{q}_{null} . The gray plane is the data-informed subspace within the control space, spanned by the eigenvectors $\{\mathbf{v}_i\}$ of the non-dimensionalized misfit Hessian (3.4). The vertical black line represents the nullspace, i.e., the orthogonal complement of the data-informed subspace. Parts of the unit sphere of the control space is displayed in black. The orange unit vector within the control space shows the target direction \mathbf{q} , and the green and black vectors show the projections of \mathbf{q} onto the data-informed subspace, \mathbf{q}_{obs} , and onto the nullspace, \mathbf{q}_{null} . The longer the vector \mathbf{q}_{obs} inside the dark gray unit circle of the data-informed subspace, the higher the *hypothetical proxy potential* of the considered observing system for the QoI.

determined by the extent to which adjoint sensitivities of the QoI project onto adjoint sensitivities of observed quantities.

F.2 The factor $0 < \frac{\lambda_i}{\lambda_i+1} < 1$, which induces only a fraction (less than 100%) of the hypothetical proxy potential to be transmitted, as a result of *noise masking*.

I will develop and explain the highlighted concepts - the hypothetical proxy potential and masking by noisy observations - in the next two subsections.

3.5.2 Hypothetical proxy potential

The projection of the target direction \mathbf{q} onto the data-informed subspace is given by

$$\mathbf{q}_{\text{obs}} := \sum_{i=1}^{M'} \langle \mathbf{q}, \mathbf{v}_i \rangle \cdot \mathbf{v}_i,$$

where $\{\mathbf{v}_1, \dots, \mathbf{v}_{M'}\}$ is the set of orthonormal data-informed directions, i.e., the set of eigenvectors of the non-dimensionalized misfit Hessian (3.4). The target direction \mathbf{q} can then be written as the orthogonal decomposition

$$\mathbf{q} = \mathbf{q}_{\text{obs}} + \mathbf{q}_{\text{null}}. \quad (3.13)$$

Here, \mathbf{q}_{obs} is the component that lies in the data-informed subspace. The remainder $\mathbf{q}_{\text{null}} := 1 - \mathbf{q}_{\text{obs}}$ belongs to the orthogonal complement of the data-informed subspace, which I refer to as the *nullspace*. The decomposition of the target direction into its informed and nullspace component is sketched in Schematic 3.4. Uncertainty gets only reduced along the data-informed component \mathbf{q}_{obs} , not along the nullspace component \mathbf{q}_{null} . The squared length of the informed component,

$$\|\mathbf{q}_{\text{obs}}\|^2 = \sum_{i=1}^{M'} \langle \mathbf{q}, \mathbf{v}_i \rangle^2 \in [0, 1], \quad (3.14)$$

is an upper bound for the relative uncertainty reduction in (3.12):

$$\tilde{\Delta}\sigma^2 \leq \sum_{i=1}^{M'} \langle \mathbf{q}, \mathbf{v}_i \rangle^2 = \|\mathbf{q}_{\text{obs}}\|^2. \quad (3.15)$$

In (3.15), equality would be attained if all factors $\frac{\lambda_i}{\lambda_i+1}$ in (3.12) were equal to 1. This would be the case in the limit of noise-free observations, as will be discussed in Section 3.5.3. The squared length of the informed component, given by (3.14), is therefore the proxy potential of the examined observing system for the QoI in the hypothetical case of noise-free observations, on a scale from 0 to 1. It will therefore be referred to as the *hypothetical proxy potential*. The remaining fraction $\|\mathbf{q}_{\text{null}}\|^2 = 1 - \|\mathbf{q}_{\text{obs}}\|^2$ will not be informed by the considered set of observations, even if the observations were noise-free.

The hypothetical proxy potential, $\|\mathbf{q}_{\text{obs}}\|^2$, is determined by the squared inner products $0 \leq \langle \mathbf{q}, \mathbf{v}_i \rangle^2 \leq 1$ that appear as summands in (3.14). A value of $\langle \mathbf{q}, \mathbf{v}_i \rangle^2 = 1$ means that the data-informed direction \mathbf{v}_i constrains the QoI perfectly, whereas a value of $\langle \mathbf{q}, \mathbf{v}_i \rangle^2 = 0$ indicates that \mathbf{v}_i does not contribute at all to informing the QoI. Note that the target direction \mathbf{q} and the data-informed directions \mathbf{v}_i are determined by adjoint sensitivities of the QoI and the observed quantities, capturing their dynamical adjustment mechanisms and pathways. The squared projections $\langle \mathbf{q}, \mathbf{v}_i \rangle^2$ evaluate therefore to what degree observations and QoIs have shared adjustment mechanisms and pathways.

Examples and illustrations for the concept of hypothetical proxy potential will be given in Section 4.3.4.

3.5.3 Noise masking

The factors

$$0 < \frac{\lambda_i}{\lambda_i + 1} < 1, \quad (3.16)$$

emerging in (3.12), take values less than 1 and determine what fraction (less than 100%) of the hypothetical proxy potential can be extracted. The factors (3.16) are therefore referred to as *information transfer factors*. These factors are independent of the studied QoI. Instead, they are determined solely by $\{\lambda_i\}$, the eigenvalues of the non-dimensionalized misfit Hessian (3.4). Figure 3.5 shows a plot of the function $\lambda \mapsto \frac{\lambda}{\lambda+1}$. The function evaluates to 0 for $\lambda = 0$, is monotonically increasing in λ , and approaches its upper bound 1 in the limit $\lambda \rightarrow \infty$. That means that large eigenvalues λ_i generate

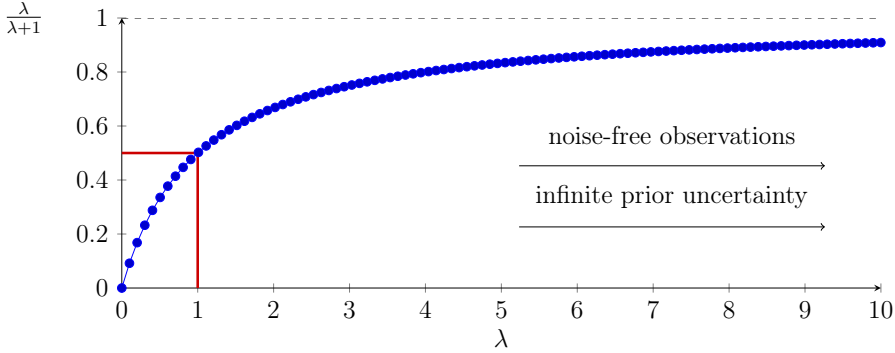


Figure 3.5: The function $\lambda \mapsto \frac{\lambda}{\lambda+1}$, which describes the dependency of the *information transfer factor* in (3.12) on the eigenvalue λ of the non-dimensionalized misfit Hessian (3.4). For an eigenvalue of 1, the information transfer factor evaluates to 0.5 (red lines), i.e., half of the hypothetical proxy potential can be extracted in (3.12).

larger factors $\frac{\lambda_i}{\lambda_i+1}$ than small eigenvalues λ_i do. Therefore large eigenvalues, corresponding to directions along which J_{misfit} is more curved, induce a larger fraction of the hypothetical proxy potential to be transmitted and lead to a higher effective proxy potential.

Recalling the results from Section 3.3 helps to understand what determines the size of the eigenvalues $\{\lambda_i\}$. In the case of only one available observation, the eigenvalue λ_1 is given by the prior-to-noise ratio

$$\lambda_1 = \frac{1}{\varepsilon^2} \left\| \mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial f_1}{\partial \mathbf{x}} \right]_{\mathbf{x}_{\min}}^T \right\|^2,$$

see equation (3.7). Here, f_1 denotes the observed quantity, as simulated by the model, and the observational noise distribution is given by the one-dimensional Gaussian $\mathcal{N}(0, \varepsilon^2)$. This explicit expression for the eigenvalue λ_1 shows that large eigenvalues are a result of (a) high sensitivity of the observations to changes in the model control variables, (b) large prior uncertainty in the control variables, and (c) small observational noise. In the limit of a noise-free observation ($\varepsilon^2 \searrow 0$), the eigenvalue λ_1 tends to infinity, and the factor $\frac{\lambda_1}{\lambda_1+1}$ converges to 1. A scaling argument generalizes this concept to an arbitrary number of observations. If one modifies the noise and prior covariance matrices by a multiplication of scalars $r^2, b^2 > 0$,

$$\mathbf{\Gamma}_{\text{noise}}[r^2] := r^2 \cdot \mathbf{\Gamma}_{\text{noise}} \quad \text{and} \quad \mathbf{\Gamma}_{\text{prior}}[b^2] := b^2 \cdot \mathbf{\Gamma}_{\text{prior}}, \quad (3.17)$$

the eigenvectors of the modified non-dimensionalized misfit Hessian $\tilde{\mathbf{H}}_{\text{misfit}}[r^2, b^2]$ remain unchanged (where $\tilde{\mathbf{H}}_{\text{misfit}}[r^2, b^2]$ is (3.4), with modified noise and prior covariances $\mathbf{\Gamma}_{\text{noise}}[r^2]$, $\mathbf{\Gamma}_{\text{prior}}[b^2]$, but unchanged \mathbf{A}). The corresponding eigenvalues, however, are modified according to⁷

$$\lambda_i[r^2, b^2] = (b/r)^2 \cdot \lambda_i[1, 1]. \quad (3.18)$$

⁷ The matrix $\tilde{\mathbf{H}}_{\text{misfit}}[1, 1]$, with eigenvectors and eigenvalues $\{(\mathbf{v}_i, \lambda_i)\}$, can be written as its singular

The scaling factor in identity (3.18) supports the interpretation of the eigenvalues λ_i as prior-to-noise ratios. The information transfer factors $\frac{\lambda_i}{\lambda_i+1}$ would therefore be equal to 1 if the observations were noise-free ($r^2 \searrow 0$) or the prior uncertainty would be infinitely high ($b^2 \nearrow \infty$). In these hypothetical cases, the proxy potential would be equal to the hypothetical proxy potential from Section 3.5.2. In realistic scenarios, however, a fraction less than 100% of the hypothetical proxy potential will be extractable, due to masking by noisy observations.

Examples and illustrations for the concept of masking by imperfect observations will be given in Section 4.3.4.

3.6 Discussion

Uncertainty Quantification (UQ) in inverse problems is a powerful tool for ocean observing system design targeted at climate monitoring. UQ provides a mathematically rigorous and a fundamentally dynamics-based method to assess the impact of existing or future observing systems, in the context of all existing observations. As opposed to statistical methods, uncertainty reduction (accomplished by observations) is assessed by propagating observational information via the model dynamics of the underlying GCM throughout the global ocean. UQ is founded on a mathematically rigorous framework, which formulates the deterministic inverse problem, i.e., the ocean state and parameter estimation problem, as one of Bayesian inference. The weight matrices $\mathbf{\Gamma}_{\text{noise}}$ and $\mathbf{\Gamma}_{\text{prior}}$ from the deterministic cost function J take on the role of Gaussian covariance matrices for the prior and noise distributions in the Bayesian inverse problem. While the minimizer of the cost function J turns out to be the most likely (MAP point) of all states, posterior uncertainty (as a result of uncertainty propagation) in this state is captured by the curvature of the cost function J , expressed through the Hessian matrix of J . However, calculating the Hessian is so computationally demanding that UQ for full-fledged oceanographic inverse problems on climate-relevant time windows has so far been intractable. While computational progress toward low-rank approximations of the Hessian are underway [Kalmikov and Heimbach, 2014, 2018; Moore et al., 2011a,b, 2017a], the goal of this thesis chapter was to revisit the computational tool 'Hessian-based UQ' from a dynamical viewpoint, in order to advance the *dynamical understanding*, rather than the numerics, of the tool - a necessary step on the pathway toward dynamics-based ocean observing system design.

3.6.1 Key insights

Performing Hessian-based UQ means finding the directions within the control space in which the non-dimensionalized model-data misfit function is the most curved. These directions are found by the eigenvalue decomposition: $\tilde{\mathbf{H}}_{\text{misfit}}[1, 1] = \sum_{i=1}^{M'} \lambda_i \mathbf{v}_i \mathbf{v}_i^T$. From (3.4), it can easily be seen that the modified non-dimensionalized misfit Hessian is the scaled version $\tilde{\mathbf{H}}_{\text{misfit}}[r^2, b^2] = (b/r)^2 \cdot \tilde{\mathbf{H}}_{\text{misfit}}[1, 1]$. Therefore,

$$\tilde{\mathbf{H}}_{\text{misfit}}[r^2, b^2] = \sum_{i=1}^{M'} ((b/r)^2 \cdot \lambda_i) \mathbf{v}_i \mathbf{v}_i^T,$$

with unchanged eigenvectors \mathbf{v}_i , but modified eigenvalues $(b/r)^2 \cdot \lambda_i$.

rections are given by the leading eigenvectors (i.e., the eigenvectors that correspond to the largest eigenvalues) of the non-dimensionalized misfit Hessian. The leading eigenvectors provide valuable insights for observing system design because they characterize the leading *data-informed directions* within the control space, capturing the most potent data constraints on the model by the considered observing system. To shed light on what determines the leading data-informed of an observing system, I presented an alternative technique for computing the eigenvectors and eigenvalues $\{(\mathbf{v}_i, \lambda_i)\}$ of the non-dimensionalized misfit Hessian. Following the steps of the alternative technique leads to the following key insights.

In an inverse problem, an observing system that provides M observations (in space and time) is represented by M scalar functions $f_1(\mathbf{x}), \dots, f_M(\mathbf{x})$, the observation counterparts simulated by the model. Each function $f_i(\mathbf{x})$ describes how a change in the control variables \mathbf{x} affects the simulated observation through the model dynamics (see Fig. 3.6). For an observing system that consists only of a single observation (in space and time), represented by the scalar function $f_1(\mathbf{x})$ in the model, the *only* data-informed direction is given by $\mathbf{\Gamma}_{\text{prior}}^{1/2} [\frac{\partial f_1}{\partial \mathbf{x}}]^T$, i.e., the *linearized sensitivity of the observed quantity* to all control variables, weighted by the prior uncertainties of the control variables. The eigenvector \mathbf{v}_1 of the non-dimensionalized misfit Hessian is the normalized version of this prior-weighted sensitivity vector. The sensitivities $\frac{\partial f_1}{\partial \mathbf{x}} = (\frac{\partial f_1}{\partial x_1}, \dots, \frac{\partial f_1}{\partial x_N})$ of $f_1(\mathbf{x})$ to all control variables $\mathbf{x} = (x_1, \dots, x_N)$ can be obtained by a single adjoint model integration (see Fig. 1.9). Moreover, the presented technique provides a formal recipe for how to obtain the eigenvectors $\{\mathbf{v}_i\}$ of the non-dimensionalized misfit Hessian for an observing system that is composed of M observations, represented by $f_1(\mathbf{x}), \dots, f_M(\mathbf{x})$. The set of data-informed directions by the individual observations, i.e., the set $\{\mathbf{\Gamma}_{\text{prior}}^{1/2} \frac{\partial f_1}{\partial \mathbf{x}}, \dots, \mathbf{\Gamma}_{\text{prior}}^{1/2} \frac{\partial f_M}{\partial \mathbf{x}}\}$ of pre-computed and prior-weighted observation sensitivities, has to be modified by two steps: it is (i) orthonormalized, in order to remove redundancy in the data, and (ii) rotated, in order to find the directions with the highest curvatures of the non-dimensionalized model-data misfit function. In particular, the subspace of the control space that is informed by the observing system is spanned by the vectors $\{\mathbf{\Gamma}_{\text{prior}}^{1/2} \frac{\partial f_1}{\partial \mathbf{x}}, \dots, \mathbf{\Gamma}_{\text{prior}}^{1/2} \frac{\partial f_M}{\partial \mathbf{x}}\}$, since steps (i) and (ii) are transformations within that subspace. This fact establishes an explicit link to dynamical principles: the sensitivities $\frac{\partial f_i}{\partial \mathbf{x}}$ capture all possible dynamical adjustment mechanisms and pathways via which small-amplitude perturbations in the control variables can change the observed quantity. This link reflects the fact that an ocean observation is the superposition of many signals that are propagated to the instrument by different processes, e.g., by advective and diffusive processes, or barotropic and baroclinic waves. Therefore, the observation contains information on a much larger part of the ocean state than the local measurement itself. Examples and illustrations for the data-informed directions $\{\mathbf{v}_i\}$ can be found in Chapter 4.

The eigenvalue λ_i of the non-dimensionalized misfit Hessian is the curvature of the non-dimensionalized model-data misfit function along the data-informed direction \mathbf{v}_i . The curvatures $\{\lambda_i\}$ can be thought of as *prior-to-noise ratios*. This is easiest explained for an observing system with a single observation, represented by the scalar function $f_1(\mathbf{x})$: the curvature λ_1 along the data-informed direction \mathbf{v}_1 is then given by the prior uncertainty in $f_1(\mathbf{x})$, divided by the noise in observation y_1 . The prior uncertainty in

$f_1(\mathbf{x})$ is obtained by propagating the prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$ of the control variables along path (*) in Fig. 3.6 to $f_1(\mathbf{x})$ and computes as $\|\mathbf{\Gamma}_{\text{prior}}^{1/2} \frac{\partial f_1}{\partial \mathbf{x}}\|^2$. The noise in observation y_1 is $\mathbf{\Gamma}_{\text{noise}}$, which, in the simple case of a single observation, is the *scalar* noise variance. Hence, the prior-to-noise ratio λ_1 compares the quality of prior information on the observed quantity - determining the *relevance* of the new data - to the *quality* of the new data. In Fig. 3.6, this prior-to-noise comparison is indicated by the downward-pointing orange and upward-pointing green arrows between $f_1(\mathbf{x})$ and y_1 . In the shown example in Fig. 3.6, the noise in observation y_1 (green error bar) is much larger than the prior uncertainty in $f_1(\mathbf{x})$ (orange error bar), and the prior-to-noise ratio λ_1 would be small. In contrast, if the observing system consisted of the single observation y_M , and model counterpart $f_M(\mathbf{x})$, the prior-to-noise ratio would be large in Fig. 3.6. As a rule of thumb, the prior-to-noise ratio is large when (a) the observed quantity has high sensitivity to the control variables, (b) the prior uncertainties in the control variables are large, and (c) the observational noise is small. Examples and illustrations for the prior-to-noise ratios $\{\lambda_i\}$ can be found in Chapter 4.

For complex (realistic) observing systems the number of observations M is on the order of millions. The presented technique for computing $\{(\mathbf{v}_i, \lambda_i)\}$ is not feasible for such complex observing systems, since it would require to perform M adjoint sensitivity experiments, where each observed quantity would be the objective function of a new experiment. Performing millions of adjoint sensitivity experiments is replaced by extracting the leading eigenvectors of the non-dimensionalized misfit Hessian via Lanczos or randomized SVD methods [e.g., Isaac et al., 2015; Kalmikov and Heimbach, 2014, 2018]. Nevertheless, even for the most complex observing system, the method presented in this chapter still gives the correct (though intractable) recipe to obtain the eigenvectors and eigenvalues of the non-dimensionalized misfit Hessian. Therefore the conclusions and dynamical insights stated in the last two paragraphs carry over to full-fledged observing system design within oceanographic inverse problems - and provide a pathway for dynamics-based ocean observing system design. In short, the data-informed directions $\{\mathbf{v}_i\}$ will be a composition, i.e., a linear combination, of prior-weighted adjoint sensitivities of the observed quantities $f_i(\mathbf{x})$ to all controls. The way this composition is formed takes into account observation uncertainties, prior information, and data redundancy.

In ocean climate research, one does not typically seek to understand the observational constraints on all model control variables \mathbf{x} (consisting of initial conditions, boundary forcings, and model parameters), but rather on unobserved (scalar) climate quantities of interest (QoIs). The QoI is modeled by a scalar function $\text{QoI}(\mathbf{x})$, which describes how a change in the control variables \mathbf{x} affects the QoI through the model dynamics (see Fig. 3.6). In the UQ framework, focusing on a certain QoI, rather than on all control variables, means that one studies the curvature of the non-dimensionalized cost function along *one* target direction, associated with the QoI, rather than along *all* directions within the control space. The target direction is the direction within the control space that needs to be constrained in order to inform the QoI. Just as the data-informed directions are the set of (orthonormalized and rotated) prior-weighted sensitivities of observed quantities, the target direction is the prior-weighted sensitivity of the QoI, given by $\mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]^T$. The sensitivities $\frac{\partial(\text{QoI})}{\partial \mathbf{x}}$ capture all possible dynamical adjustment

mechanisms and pathways which can transport signals to the QoI and which therefore have to be informed by the observations.

The *proxy potential* $\tilde{\Delta}\sigma^2$ of an observing system for an unobserved climate QoI was defined as the relative uncertainty reduction in the QoI, by including the observing system under consideration in the inverse problem. Equivalently, $\tilde{\Delta}\sigma^2$ measures how much the curvature of the cost function gets increased along the QoI target direction, accomplished through the information gain by the observing system. The higher the proxy potential $\tilde{\Delta}\sigma^2 \in [0, 1]$, the better the improvement of the constraints on the QoI by the added observing system. I established and explored two concepts whose combination fully determine the proxy potential $\tilde{\Delta}\sigma^2$. The two concepts are summarized in Fig. 3.6, and illustrated with applications in Chapter 4.

F.1 The *hypothetical proxy potential* of the considered observing system for the QoI, which is characterized by the extent to which the target direction projects onto the data-informed directions. Or, rephrasing this by means of the dynamical insights from above: the hypothetical proxy potential is characterized by the extent to which prior-weighted sensitivities of the QoI(\mathbf{x}) (blue box in the right branch of Fig. 3.6) projects onto the set of (orthonormalized and rotated) prior-weighted sensitivities of the observed quantities $f_1(\mathbf{x}), \dots, f_M(\mathbf{x})$ (blue boxes in the left branch of Fig. 3.6). Therefore, physically speaking, the hypothetical proxy potential is large if the observations are influenced by similar dynamical adjustment mechanisms and pathways as the unobserved QoI. In the hypothetical case of noise-free observations, the effective proxy potential $\tilde{\Delta}\sigma^2$ is given by the hypothetical proxy potential.

F.2 *Masking by noisy observations*, depending on the prior-to-noise ratios of the observations. The information transfer factors $\lambda_i/(\lambda_i + 1) < 1$ determine what fraction of the hypothetical proxy potential of the observations can be extracted, and are independent of the considered QoI. As explained above, the prior-to-noise ratios $\{\lambda_i\}$ compare the quality of prior information on the newly observed quantities to the quality of the new data. In Fig. 3.6, this prior-to-noise comparison is indicated by the downward-pointing orange and upward-pointing green arrows between the simulated observations $f_i(\mathbf{x})$ and actual observations y_i . For *noise-free* observations the quality of the new data would be perfect, and the associated information transfer factors would be equal to 1. In this hypothetical case, the hypothetical proxy potential could be fully retrieved.

Both concepts **F.1** and **F.2** emphasize that the notion of proxy potential (and, more generally, UQ in inverse problems) is based on *dynamical principles*, rather than on statistical inference alone. **F.1** compares dynamical adjustment mechanisms via adjoint sensitivities of observed and unobserved quantities, rather than using statistical correlations between the quantities. Similarly, **F.2** depends on the prior-to-noise ratio, which stands in contrast to the concept of a statistical signal-to-noise ratio. Indeed, the prior-to-noise ratio compares *prior uncertainty* in the observed quantity with the noise in the actual (new) observation. Prior uncertainty in the observed quantity is measured by propagating prior information and uncertainty via the model equations throughout the global ocean (orange arrow along path (*) in Fig. 3.6). In contrast, the statistical

UQ for dynamics-based observing system design - revisited

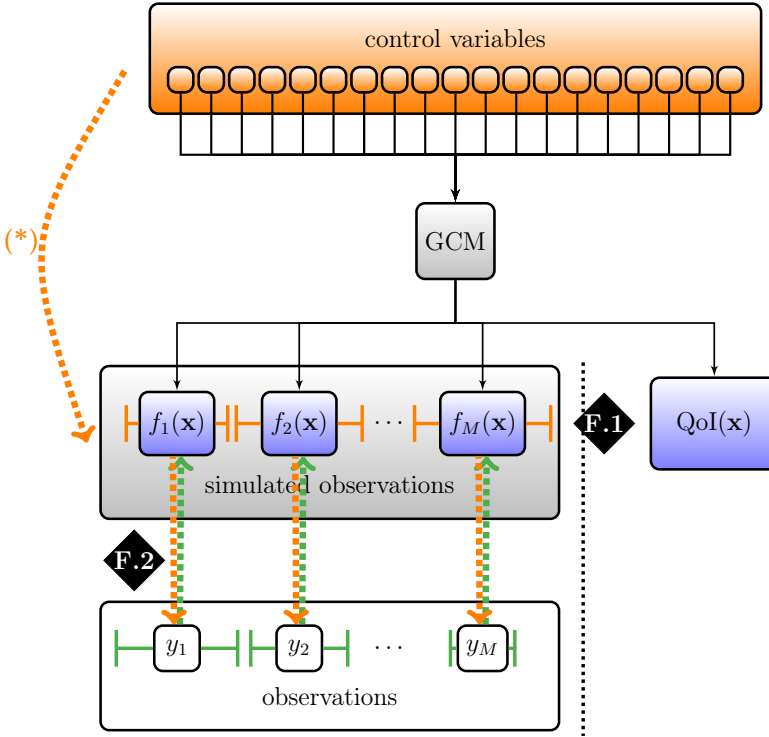


Figure 3.6: Fig. 3.1 revisited, highlighting the two concepts F.1 and F.2 (black diamonds) whose combination fully characterizes the *effective proxy potential* of an observing system for an unobserved quantity of interest (QoI). Concept **F.1** is the *hypothetical proxy potential*, which measures to which degree $\mathbf{\Gamma}_{\text{prior}}^{1/2} \frac{\partial \text{QoI}}{\partial \mathbf{x}}$, the prior-weighted sensitivities of the quantity of interest QoI(x) (blue box in the right branch), project onto the set of (orthonormalized and rotated) $\{\mathbf{\Gamma}_{\text{prior}}^{1/2} \frac{\partial f_i}{\partial \mathbf{x}}\}$, the prior-weighted sensitivities of the observed quantities $f_i(\mathbf{x})$ (blue boxes in the left branch). Concept **F.2** is *masking by noisy observations*. The degree of the masking depends on the prior-to-noise ratios, indicated by the downward-pointing orange and upward-pointing green arrows between the simulated observations $f_i(\mathbf{x})$ and actual observations y_i . The prior-to-noise ratio for the observed quantity $f_i(\mathbf{x})$ is given by the prior uncertainty in $f_i(\mathbf{x})$ (orange error bars), divided by the noise in observation y_i (green error bars). Here, the prior uncertainty in $f_i(\mathbf{x})$ is obtained by propagating the prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$ of the control variables along path (*) to $f_i(\mathbf{x})$. The prior-to-noise ratios compare the quality of prior information on the observed quantities - determining the relevance of the new data - to the quality of the new data.

signal-to-noise ratio compares the *signal* of the observed quantity with the noise in the observation. Here, the signal of the observed quantity is usually assessed by an empirical statistical measure, i.e., the sample variance for a measurement or its counterpart in a model.

Concepts [F.1](#) and [F.2](#) also clear up how UQ-based observing system design generalizes the previously used method of performing an adjoint sensitivity analysis for a certain QoI (“Level 1” in [Table 1.1](#)). The “Level 1” method computes the linearized sensitivity $\frac{\partial(\text{QoI})}{\partial \mathbf{x}}$ of a climate quantity of interest (QoI) to the control variables, and suggests to prioritize observations in regions of high sensitivity. Therefore, seen from a UQ perspective, the “Level 1” method computes only the target direction. In contrast, UQ exploits in addition that high sensitivity regions for the target direction might be part of one or several dynamical adjustment mechanisms which transport signals to the QoI. These dynamical adjustment mechanisms, in turn, might influence also remotely placed observations. Consequently, some of the desired information about the QoI will already be captured by existing observations. Moreover, to acquire the remaining information, UQ suggests that a new observing system can be designed more efficiently than what is qualitatively suggested by the “Level 1” method: that is, by arranging the instruments in such a way that important dynamical adjustment mechanisms are captured ([concept F.1](#)), rather than deploying them at every place that shows high local sensitivity. Moreover, as opposed to the “Level 1” method, UQ-based observing system design takes into account data redundancy (by orthonormalizing and rotating the set of sensitivities for the observed quantities) and observational noise (through [concept F.2](#)).

Finally, [concept F.1](#) offers a clear interpretation of how aggregation errors [[Kaminski et al., 2001](#)] are generated if one were to reduce the control space a-priori. A-priori-reductions of the control space have been used in some past studies to overcome the challenge of computing the Hessian matrix within inverse problems with high-dimensional control spaces. For instance, [Kaminski et al. \[2015, 2018\]](#) pursue the ‘large region approach’, where atmospheric forcing and initial conditions are adjusted uniformly over large regions, rather than on a grid point basis of the underlying GCM. In practice, the ‘large region approach’ means that sensitivities of observed quantities and QoIs are spatially averaged over these large regions. Consequently, the dynamical adjustment mechanisms that are found in the “high-resolution” sensitivity maps might not be properly resolved for the ‘large region approach’ - even if the underlying GCM is of high resolution (see [Fig. 6.10](#) in [Chapter 6](#)). This might corrupt the identified proxy potential, which crucially depends on comparing dynamical adjustment mechanisms in the sensitivity maps of observed quantities and QoIs ([concept F.1](#)). Therefore, information that is propagated from the observations to the QoIs in the real ocean might get lost when assessed within a framework that is based on the ‘large region approach’.

3.6.2 Limitations

Assumptions on prior and noise

The foundation of UQ-based observing system design is uncertainty propagation throughout the global ocean, via dynamical principles encoded in the model equations ([Fig. 3.1](#)).

However, before uncertainty can be propagated, one has to make assumptions on prior information and observational uncertainties in the form of probability distributions for the prior and noise. These probability distributions are generally assumed to be Gaussian such that the assumptions are encapsulated in a prior guess \mathbf{x}_0 , and the covariance error matrices $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ (see Appendix B.1). While the specification of \mathbf{x}_0 , $\mathbf{\Gamma}_{\text{prior}}$, and $\mathbf{\Gamma}_{\text{noise}}$ makes all assumptions explicit and transparent, the result of uncertainty propagation (the inferred posterior uncertainties) and UQ-based observing system design will depend on these assumptions. At the same time, a specification of accurate error covariances $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ is a difficult task, and often done in a somewhat simplistic way - either due to the lack of knowledge or for the sake of numerical efficiency. For instance, the prior ECCO v4 variances (i.e., the diagonal entries of $\mathbf{\Gamma}_{\text{prior}}$) for the atmospheric boundary fields are computed based upon the spread of available atmospheric reanalysis products [Chaudhuri et al., 2012; Forget et al., 2015]. Further, the Weaver and Courtier [2001] smoother is applied to mimic prior cross-correlations between control variables that represent close-by grid points in the same field, while no prior cross-correlations are assumed between distinct fields, such as air temperature and specific humidity [Forget et al., 2015]. Accurate specification of turbulent transport parameters such as diapycnal and isopycnal diffusivities is even more difficult, and often relies on ad-hoc choices [Forget et al., 2015; Fukumori et al., 2017]. Similarly, it is challenging to make accurate assumptions on the observational uncertainties $\mathbf{\Gamma}_{\text{noise}}$. The challenge stems from the fact that observational uncertainties do not only comprise the instrument error, but are often dominated by the representation error in the model [e.g., Wunsch and Heimbach, 2007], and representation errors are difficult to determine. Moreover, due to the lack of knowledge of cross-correlations between observations, off-diagonal entries in $\mathbf{\Gamma}_{\text{noise}}$ are often assumed to be zero [e.g., Forget et al., 2015].

To gain more insight into the roles of the prior and noise matrices $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ (e.g., the importance of off-diagonal entries) for dynamics-based observing system design, the method developed in this chapter may be used as a computationally efficient tool to perform UQ with variable prior and noise matrices (Section 3.3.3). Computational efficiency of the method stems from the fact that different choices of $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ can be tested through cheap offline post-processing steps, which do not require the integration of a forward or adjoint model, as opposed to conventional (e.g., Lanczos) methods. While the method presented here is only feasible for simple observing systems with few observations (thus, small-sized $\mathbf{\Gamma}_{\text{noise}}$), it can handle high-dimensional control spaces (thus, full-sized $\mathbf{\Gamma}_{\text{prior}}$) and complex forward models. The method might therefore provide a useful testbed to gain more insight into the roles of $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$ for UQ-based observing system design - before the latter is implemented within costly high-performance computing frameworks.

An integral aspect of accurate assumptions on prior and noise is that prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$ are “in harmony” with observational uncertainties $\mathbf{\Gamma}_{\text{noise}}$. This can be explained by means of concept F.2. For instance, if the prior uncertainties were specified unrealistically small, the prior uncertainty in the observed quantities (orange error bars in Fig. 3.6, obtained by prior uncertainty propagation along (*)) would be consistently too small. In the prior-to-noise comparison (F.2 in Fig. 3.6), this would lead to the belief that the actual observations are not very relevant because the observational uncertainties

(green error bars in Fig. 3.6) would be comparatively large. To conclude, dynamics-based observing system design requires accurate and “harmonized” matrices $\mathbf{\Gamma}_{\text{prior}}$ and $\mathbf{\Gamma}_{\text{noise}}$, and therefore a close dialogue between the modeling and the observational communities.

Assumption of (weak non-) linearity

A limitation of Hessian-based UQ is that the posterior probability distribution is locally approximated by a Gaussian because the sampling of the full, possibly complex, probability distribution (e.g., by MCMC) is intractable. The accuracy of the Gaussian approximation depends on how well the linearized model about the cost function minimizer (the MAP point) can reflect the full response of the non-linear system. That is, if the observation operator $\mathbf{f}(\mathbf{x})$ was entirely linear, the posterior probability would be Gaussian, and the inverse Hessian of J would provide the exact posterior error covariance. Moreover, for the case of a linear observation operator, the linearized Hessian and full Hessian are identical (see Appendix B.4). Therefore, for a weakly non-linear observation operator, both the linearized Hessian (used in this thesis, as well as in Thacker [1989], Losch and Wunsch [2003], Moore et al. [2011a, 2012, 2011b], Bui-Thanh et al. [2012], and Kaminski et al. [2015, 2018]) and the full Hessian (used in Kalmikov and Heimbach [2014, 2018], Isaac et al. [2015], and Alexanderian et al. [2016]) should provide a good basis for a Gaussian approximation. Moreover, the assumption of weak non-linearity was used at several points in this chapter, specifically when suggesting a computationally efficient method to test various prior and noise matrices (see Section 3.3.3) and when boiling down the notion of proxy potential to the two concepts F.1 and F.2 (see Appendix B.3). That is, at both points it was assumed that the derivatives of either the observation operator $\mathbf{f}(\mathbf{x})$ or the quantity of interest $\text{QoI}(\mathbf{x})$, depend only insignificantly on their evaluation (or base) point \mathbf{x} . Here, this assumption is based on the fact that in the linear case, the derivatives would not depend on the evaluation point at all. Heimbach et al. [2011] discuss the dependence of derivatives on the base state further. On what time windows and in which ocean regions the assumption of “weak non-linearity” is satisfied, is an open question in ocean climate science. An effort to test the linearity assumption in the subpolar North Atlantic will be made in Chapter 5.

If the oceanic observed quantities are dominated by strongly non-linear model dynamics on the considered timescales, the cost function may have several local minima, or, equivalently, the posterior probability function may be multimodal. A Gaussian approximation about the global (or a local) cost function minimizer will then perform poorly when seen as a global approximation. However, in the vicinity of each local minimum (corresponding to one mode of the posterior distribution), a local Gaussian approximation based on the local Hessian should still be reasonably accurate. In such a situation, the main problem is in the first place that, before Hessian-based UQ for the various minima of the cost function can be performed, these minima have to be detected by the optimization algorithm; if not, significant probability for the estimated ocean state might lie in unidentified distant regions.

It is noted that in the case of multi-modal distributions, the approach of solving a non-linear inverse problem together with Hessian-based UQ might still be seen superior to using ensemble-based data assimilation methods. Ensemble-based approaches also

compute Gaussian error covariances. However, both the ensemble mean and covariance matrix might be spurious and unphysical because information from different modes of the true distribution is intermingled in the Gaussian approximation. In contrast, the mean of the Gaussian approximation obtained by solving a non-linear inverse problem is chosen as the MAP point, and represents a dynamically and kinematically consistent ocean state. Moreover, error covariances in Hessian-based UQ are computed by propagating uncertainties from the full control space via the model equations (albeit linearized). In contrast, ensemble-based error covariances miss uncertainties that lie outside the subspace defined by the ensemble forecast, and, thus, often have to be inflated artificially [e.g., [Anderson and Anderson, 1999](#); [Whitaker and Hamill, 2002](#)].

Model and system dependency

Finally, inferences made from UQ-based observing system design have to be interpreted carefully due to their potential model- and system dependency - a problem common to all methods for model-informed observing system design.

Since uncertainty propagation is achieved via the (linearized) model equations of the underlying GCM, dynamics-based observing system design, including proxy potential, depends on how dynamical principles are formulated in the GCM. Since many principles, e.g., equations of motion representing basic conservation and constitutive laws, are robust across GCMs, other features, such as model resolution and parametrization of sub-grid scale processes, are model-dependent. To minimize model-dependency, one can include uncertain process parameters in the control vector [[Forget et al., 2015](#)], such that parametric uncertainty is explicitly taken into account in the UQ framework. Meanwhile, structural model uncertainty cannot be captured by the UQ framework because structural uncertainty is often controlled by discrete modeling choices and switches, rather than by continuous parameters. Examples for structural model uncertainty are the choice of a numerical scheme or the formulation of individual processes. While determining structural model uncertainty is a difficult problem, a first assessment is presented by [Forget et al. \[2015\]](#). The authors conclude that uncertainties in process parameters and initial and boundary conditions, which are included in the ECCO v4 control vector, dominate over structural model uncertainty.

One possibility to account for structural model error in inverse problems would be to relax the strong constraint that the solution (the estimated ocean state) has to follow the model equations to a weak constraint (see strong constraint 4D-Var vs. weak constraint 4D-Var, e.g., [Sasaki, 1970](#); [Moore et al., 2011a](#)). However, a weakly-constrained solution will not respect conservation laws and closed global ocean budgets anymore, and might therefore be deemed less appealing for ocean climate research [[Stammer et al., 2016](#)].

Dynamics-based observing system design and proxy potential are not only model- but also system-dependent, i.e., they depend on how the inverse problem is formulated. Besides the assumptions on prior and noise, which were discussed above, the formulation of the inverse problem requires first and foremost the choice of a set of control variables. The control space should include all uncertain (or unconstrained) elements in the model, and, strictly speaking, it is implicitly assumed that by an adjustment of the control

variables (within the assumed prior uncertainties) the true ocean state can be recovered. While it was already discussed that structural uncertainty cannot be easily included in the control vector, one might choose to neglect additional control variables that, in principle, *could* be included. Such choices are often made for the sake of computational feasibility, most notably when large a-priori-reductions of the control space are performed, as for instance within the 'large region approach' [Kaminski et al., 2015, 2018]. Issues that may arise from such a-priori-reductions of the control space have already been pointed out above, as well as by Kaminski et al. [2001]. In a more subtle fashion, every inverse modeling framework will perform some form of an a-priori-reduction of the control space. For instance, while atmospheric control variables within the ECCO v4 inverse modeling framework are adjusted on a grid point basis, the adjustments are made on a bi-weekly basis, rather than on a 6-hourly basis (as the associated atmospheric forcing fields in the forward run) [Forget et al., 2015]. While, arguably, specific choices and assumptions are reasonable for certain types of applications, it is important to be aware of them and make them carefully, in order to support informative dynamics-based ocean observing system design.

3.7 Conclusions

The main conclusions of this chapter are:

- Uncertainty Quantification (UQ) within inverse modeling frameworks is a computational tool for observing system design that has not yet been applied in support of global ocean climate research - partly due to difficult dynamical interpretations of the method. In this chapter, the mathematical concepts of UQ are “translated” to dynamical concepts:
 - The information captured - and subsequently communicated - by an observing system is described by the eigenvectors and eigenvalues $\{(\mathbf{v}_i, \lambda_i)\}$ of the misfit Hessian.
 - The eigenvectors \mathbf{v}_i are determined by the linear sensitivities of the observed oceanographic quantities to all control variables. These sensitivities capture all possible dynamical adjustment mechanisms and pathways via which perturbations in the control variables can change the observed quantities.
 - Projecting the sensitivity of the observations onto the sensitivity of unobserved and remote quantities of interest (QoIs) gives a quantitative estimate of how much of the information required to constrain the QoI is provided by the observations. The goodness of the projections translates to the question: “How similar are adjustment mechanisms, pathways and relative magnitudes for the observing system and the unobserved QoIs?”
 - The observing system provides an effective information gain on the unobserved QoI if the projection of the sensitivities is large, prior information on the QoI is small, and noise masking is not too strong. The ratio of prior information to noise is quantified by the eigenvalues λ_i .
- Novel aspects of the tool UQ for ocean observing system design are: it

-
- is fundamentally dynamics-based, making use of dynamical relationships in a GCM;
 - is favorable for climate monitoring because the dynamical relationships and budgets are respected and treated consistently over long timescales (years to decades);
 - can assess data redundancy and complementarity in the context of other observations;
 - can identify the constraints of observing systems on climate signals and QoIs that
 - * are unobservable or unobserved, different in type from the observations,
 - * are not spatially collocated with the observations.

Chapter 4

Proxy Potential of the OSNAP Array

4.1 Introduction

In the summer of 2014, OSNAP (Overturning in the Subpolar North Atlantic Program) deployed a trans-basin mooring array in the subpolar North Atlantic, consisting of two legs: OSNAP-West, extending from Labrador to the southwestern tip of Greenland, and OSNAP-East, extending from the southeastern tip of Greenland to Scotland (Fig. 1.7). The OSNAP array provides monthly estimates of the overturning circulation, and the heat and freshwater transports across the two OSNAP sections. Departing from the classical view that changes in Labrador Sea deep convection are the key process for AMOC variability [e.g., Danabasoglu et al., 2016; Xu et al., 2018], the first 21 months of OSNAP data suggest that water mass transformation north of the OSNAP-East section, rather than the OSNAP-West section, dominate the strength and variability of the subpolar overturning and the associated heat transport [Lozier et al., 2019].

The OSNAP array was designed with the principal goal to continuously monitor the subpolar overturning and heat and freshwater transports at the latitudes of the array [Lozier et al., 2017; Li et al., 2017; Lozier et al., 2019]. However, since many regions of the ocean are connected over long distances and timescales, e.g., through advective and diffusive processes, or barotropic and baroclinic waves, the OSNAP observations contain information on a much larger part of the ocean state than just locally at the array. That is, the OSNAP moorings sample signals that might impact unobserved circulation and hydrographic quantities of interest (QoIs) downstream or upstream of the OSNAP array at a later or earlier time. Therefore, the OSNAP array could inform - or could even be used as a proxy for - unobserved climate QoIs that are *remote* from the array. Uncovering the proxy potential of the OSNAP array for unobserved and remote QoIs would put the brand-new OSNAP observations into a broader spatial and temporal context - a task that yet has to be undertaken.

Proxy potential of the OSNAP array arises for such QoIs that are forced by similar oceanic processes, via similar adjustment pathways, as the OSNAP observations. To explore proxy potential, one therefore has to answer three questions:

- (I) What are the dynamical pathways, relative magnitudes, and mechanisms, by which OSNAP observations can be influenced?

- (II) To what extent are mechanisms and pathways identified in (I) shared with those influencing the unobserved QoIs?
- (III) To what extent is the information about shared mechanisms and signals in (II) masked by observational noise?

The solution strategy for answering questions (I) and (II) is to use the adjoint of a GCM. The adjoint of a GCM is a numerically efficient tool that reveals all possible dynamical mechanisms and pathways via which small-amplitude perturbations in the input variables, e.g., surface forcing over the globe for all forcing lead times, impact the observations and the QoIs (Section 1.2).

Answering questions (I), (II), and (III), are the central components of a formal, mathematically rigorous evaluation, used within dynamics-based observing system design, as outlined in Chapter 3. The mathematical foundation is an adjoint-based inverse modeling framework, which is used for ocean state and parameter estimation, also referred to as ‘solving a non-linear inverse problem’. In an inverse modeling framework, a GCM serves as a dynamical interpolator in order to spread information from the included ocean observations through space and time. The solution of an inverse problem is a dynamically and kinematically consistent and data-constrained estimate of the ocean state. Formally, proxy potential of the OSNAP array for a QoI is assessed via uncertainty reduction in the estimated QoI when the OSNAP observations are included in the inverse modeling framework. In Chapter 3 it was shown that the degree of uncertainty reduction in the QoI is determined by the combination of the answers to (II) and (III). The strength of assessing proxy potential via this approach is that uncertainty reduction within an inverse problem is based purely on dynamical information that is propagated via the model equations - i.e., equations of motions representing basic conservation and constitutive laws rendered on a computer - as opposed to techniques that are based on statistical inference, e.g., EOFs.

Using uncertainty reduction within inverse problems as an approach to dynamics-based observing system design and the identification of proxy potential, is a growing research topic in the computational sciences [e.g., Flath et al., 2011; Bui-Thanh et al., 2012; Isaac et al., 2015], but has not yet been applied to full-fledged oceanographic inverse problems on climate-relevant time windows. This chapter uses these novel tools for ocean climate research and explores OSNAP’s proxy potential within the state-of-the-art ECCOv4 (Estimating the Circulation and Climate of the Ocean, version 4) inverse modeling framework [Forget et al., 2015; Fukumori et al., 2017]. While previous studies have performed a-priori-reductions of the control space within their inverse modeling frameworks to overcome computational challenges, e.g., via the ‘large region approach’ [Kaminski and Rayner, 2008; Kaminski et al., 2015, 2018], here, the ECCOv4 control space at its full spatial resolution, i.e., on a grid point basis, is pertained. A control space at the spatial resolution of the underlying GCM ensures that the dynamical adjustment mechanisms for the OSNAP observations and QoIs - the central ingredients for the answers of (I) and (II) - can be represented at the resolution of the GCM (see Section 3.6 and Fig. 6.10).

To keep the demonstration simple, only the proxy potential of *heat transport* across the two OSNAP sections is explored, while the proxy potentials of OSNAP mass and freshwater transports, as well as of direct OSNAP hydrographic and velocity measurements, are left for future work. Another simplification is that I evaluate the proxy potential of the *long-term (5-year) mean* of OSNAP heat transport for the *long-term (5-year) mean* of unobserved QoIs. The 5-year time window is chosen because the linearized model equations generated by the adjoint are found to approximate the response of the non-linear ocean processes in the subpolar North Atlantic to an acceptable degree on a five-year timescale (see Chapter 5). While only 21 months (and not 5 years) of OSNAP data exist to date [Lozier et al., 2019], I exploit the fact that the identification of proxy potential does not require the actual measurement values, only observation locations, times, types, and their uncertainties. The fact that no actual observations are needed is unique to the chosen method, which explores dynamical relationships between the OSNAP observations and remote QoIs in the GCM equations, rather than employing statistical methods. As a proof of concept, I explore the proxy potential of time-mean OSNAP heat transport for the time-mean of two QoIs that have already been studied in Chapter 2: Nordic Seas heat content, and volume transport across the Iceland-Scotland ridge. Despite the vicinity of the two chosen QoIs to the OSNAP-East section, this chapter will show that observations from OSNAP-West are at least as valuable as observations of OSNAP-East for constraining these QoIs. This highlights that OSNAP-West observations might be important for informing *remote* climate signals, even though Lozier et al. [2019] suggest that OSNAP-West plays a minor role for monitoring the *local* OSNAP overturning and heat transport metrics.

This chapter applies the novel concept of proxy potential to the field of ocean climate research. The innovative aspects of this methodology are that (i) it targets unobserved or unobservable climate-relevant signals *remote* from the observations, and (ii) it is purely dynamics-based. Given the novelty of the approach for oceanographic applications, this chapter will also test aspects of the machinery, by elucidating some effects of the necessary assumptions on prior information and observational noise. These insights will lay the ground for employing the tools of dynamics-based proxy potential and observing system design in future and more general work. Dynamics-based proxy potential and observing system design are valuable guiding tools for building and sustaining a cost-effective, long-term Atlantic observing system - while maximizing the information extracted from the observations.

This chapter is structured as follows. Section 4.2 describes the components of the inverse modeling framework. Section 4.3.1 identifies the dynamical adjustment mechanisms and pathways for time-mean heat transport across OSNAP-West and OSNAP-East, answering question (I). Section 4.3.2 links the mechanisms and pathways identified in Section 4.3.1 to formal dynamical constraints that OSNAP provides for the inverse modeling framework. Section 4.3.3 explores the dynamical adjustment mechanisms for the time-mean QoIs, gearing up for the answer to question (II). Section 4.3.4 answers questions (II) and (III), evaluating the proxy potential within the inverse modeling framework. In Section 4.4, the results are discussed.

4.2 Inverse modeling framework

To explore the proxy potential of the OSNAP array, I adopt the inverse modeling framework of the ECCO project [<http://www.ecco-group.org/>], in the version 4 (ECCOv4) edition, described in Forget et al. [2015]; Fukumori et al. [2017]. The general components of an inverse modeling framework are described in Section 3.2 and Fig. 3.1. The choices for the components made in this work are introduced in the following subsections. Section 4.2.1 describes the underlying GCM, Section 4.2.2 the set of OSNAP observations \mathbf{y} , as well as their model counterparts $\mathbf{f}(\mathbf{x})$ and their uncertainties Γ_{noise} , Section 4.2.3 the unobserved quantities of interest $\text{QoI}(\mathbf{x})$, and Section 4.2.4 the specification of control variables \mathbf{x} as well as the prior uncertainties Γ_{prior} . Finally, section 4.2.5 presents the adjoint models of the GCM that are employed to answer questions (I) and (II).

4.2.1 GCM

The model used in this work is the Massachusetts Institute of Technology general circulation model (MITgcm, Marshall et al. [1997a,b]), in the same configuration as used in ECCOv4 release 2 (ECCOv4r2). This model configuration includes a global model setup at a nominal horizontal resolution of 1° with 50 vertical layers. Unresolved processes associated with advection and mixing by eddies are parameterized following Gent and McWilliams [1990] and Redi [1982]. Vertical mixing is parametrized using the turbulent kinetic energy scheme of Gaspar et al. [1990]. Convection is treated using simple convective adjustment. The ocean model is driven by an atmospheric boundary layer scheme, where 6-hourly fluxes of heat and freshwater are determined by the bulk formulae of Large and Yeager [2004], computed at every time step from the atmospheric input fields of air temperature, specific humidity, downward longwave and shortwave radiation, and precipitation. Fluxes of momentum are not computed through bulk formulae; instead, the model is directly forced by input wind stress fields. The reader is referred to Forget et al. [2015] for more details on the ECCOv4r2 model configuration. The model initial conditions, the atmospheric input fields, and some of the parameter fields which determine parametrized turbulent transports are part of the control vector within the inverse modeling framework, as further described in Section 4.2.4. Model simulations in this work cover the period 2007-2011, the final 5 years of the ECCOv4r2 state estimate. This time window is chosen because the adjoint is found to approximate the response of the non-linear processes in the subpolar North Atlantic to an acceptable degree on a five-year timescale (see Chapter 5).

4.2.2 The OSNAP array and data uncertainties

The two sections of the OSNAP (www.o-snap.org) mooring array in the subpolar North Atlantic are sketched in Fig. 4.1: OSNAP-West, extending from Labrador to the southwestern tip of Greenland, and OSNAP-East, extending from the southeastern tip of Greenland to Scotland. OSNAP provides monthly estimates of the full-depth transports of mass, heat and freshwater across the OSNAP-West and the OSNAP-East sections, which are computed from direct measurements of the OSNAP moorings and Argo profiling float data [Li et al., 2017; Lozier et al., 2019]. For instance, for the first 21 months of OSNAP data from August 2014 to April 2016, Lozier et al. [2019] report mean heat

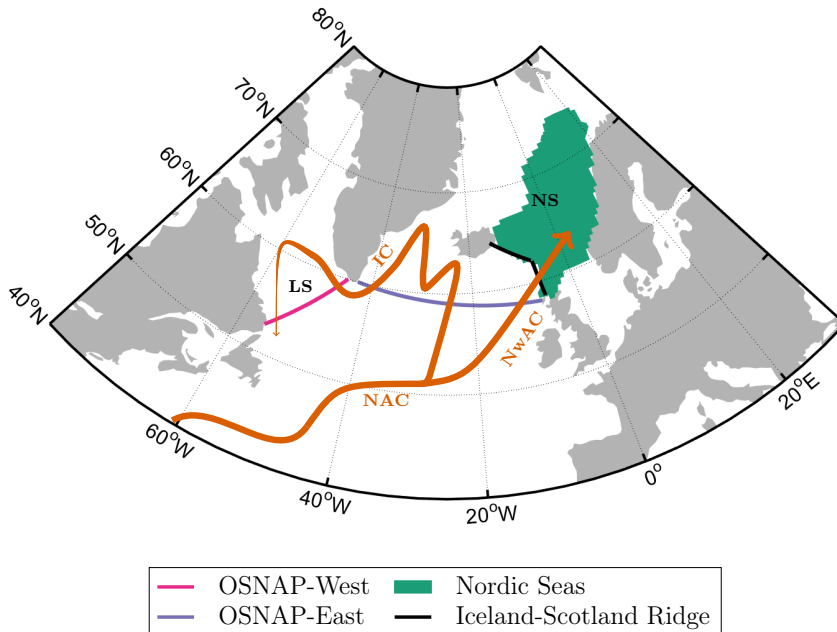


Figure 4.1: Schematic map showing the regions and transects related to the observed and unobserved model QoIs in this study. The investigated observed QoIs are heat transport across the OSNAP-West leg (pink line), and heat transport across the OSNAP-East leg (purple line). The studied unobserved QoIs are ocean heat content in the Nordic Seas region (green domain), and volume transport across the Iceland-Scotland ridge (black line). The orange arrows represent major pathways of surface currents in the subpolar North Atlantic. NAC = North Atlantic Current; NwAC = Norwegian Atlantic Current; IC = Irminger Current; NS = Nordic Seas; LS = Labrador Sea.

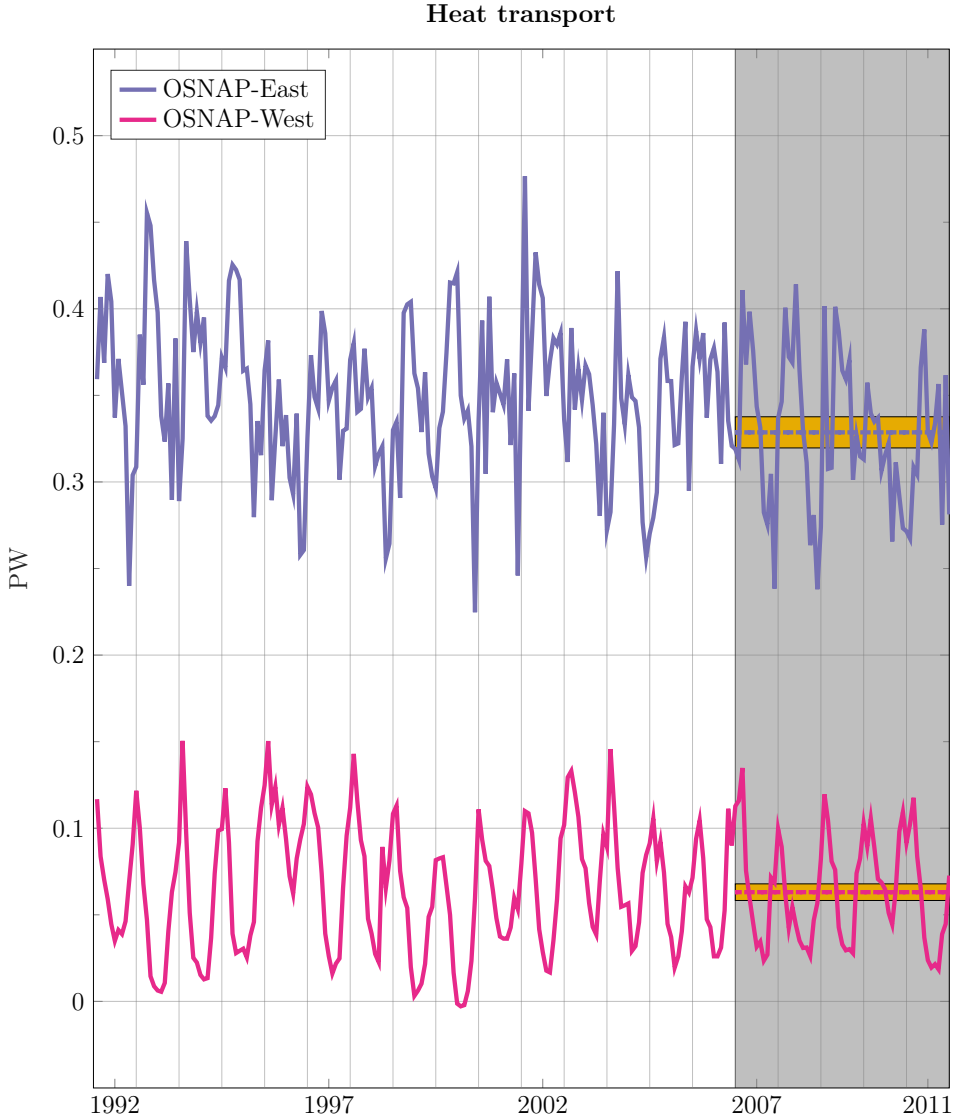


Figure 4.2: Model heat transports across OSNAP-West (solid pink line) and OSNAP-East (solid purple line) in the ECCOv4r2 solution from January 1992 to December 2011. The heat transports are diagnosed across the sections indicated by the pink and purple lines in Fig. 4.1, and integrated over the entire water column. The gray background covers the final 5 years (2007-2011), the representative 5-year time window chosen for this work. The 5-year mean heat transports, averaged over the period 2007-2011, are equal to $MHT_{OSNAP-W}$ and $MHT_{OSNAP-E}$ (defined in (4.2) and (4.3)) and are indicated by the dashed (constant) lines in the respective colors. The two yellow bars show the prior uncertainties for these mean values, computed by propagating the prior uncertainties Γ_{prior} from the control variables onto the respective model quantities. Displayed are two-standard-deviation envelopes, i.e., error bars of amplitude $\pm 2 \cdot \sigma_{\text{prior}}^W$ and $\pm 2 \cdot \sigma_{\text{prior}}^E$, respectively. The values for $(\sigma_{\text{prior}}^\bullet)^2$ are shown in Table 4.3.

transports of

$$0.080 \pm 0.004 \text{ PW} \quad (\text{OSNAP-West}), \quad 0.38 \pm 0.02 \text{ PW} \quad (\text{OSNAP-East}). \quad (4.1)$$

While ECCOv4 has not assimilated the new OSNAP observations yet, the proxy potential of the OSNAP array can yet be assessed. In fact, dynamics-based observing system design and an assessment of proxy potential do not require the actual measurement values of the observations under consideration. Instead, only observation locations, times, types, and their uncertainties are needed, as pointed out in Section 3.4. This can be explained by the fact that observations are connected to remote QoIs via dynamical principles expressed in the underlying GCM, rather than via statistical methods. For the latter approach, one would certainly require the measurement values.

In this work, I investigate OSNAP's proxy potential for unobserved QoIs, at locations remote from the array. Given the novelty of applying the concepts of dynamics-based observing system design and proxy potential for remote unobserved signals within the field of ocean climate research, two simplifying assumptions are made for the purpose of demonstration. First, the observed quantities are assumed to be *heat transport measurements* across OSNAP-West and OSNAP-East, rather than direct hydrographic and velocity measurements from the OSNAP moorings. The second simplification is that I evaluate the proxy potential of the *long-term (5-year) mean* of the observations - on the long-term (5-year) mean of unobserved QoIs. For assessing the full potential of the OSNAP observing system, however, one would have to include the direct measurements (i) from each individual OSNAP instrument, and (ii) at the measurement frequency. This will be further discussed in Section 4.4.

The 5-year time window is chosen because the linearized model equations generated by the adjoint are found to approximate the response of the non-linear ocean processes in the subpolar North Atlantic to an acceptable degree on 5-year timescales (see Chapter 5). The period 2007-2011, the final 5 years of the ECCOv4r2 state estimate, is used as a representative 5-year time window to evaluate the OSNAP's proxy potential. The period is representative for other 5-year time windows if the linearized model equations do not strongly depend on the chosen base state 2007-2011, which is assumed. This assumption is encouraged by the fact that 5-year mean quantities are studied. Note that while choosing the period 2007-2011, I exploit the fact that no actual OSNAP measurements are needed because the OSNAP array was only deployed in 2014.

In the model, the OSNAP-West and OSNAP-East transects are defined as the shortest lines from (57°W, 52°N) to (45°W, 60°N), and from (43°W, 60°N) to (5°W, 58°N), respectively, while following a grid line path as illustrated in Figure C2 in Forget et al. [2015]. The obtained OSNAP-West model transect is denoted by S_W , and the OSNAP-East model transect by S_E . The observations' counterparts simulated by the model are the 5-year mean heat transports across these model sections, i.e.,

$$\text{MHT}_{\text{OSNAP-W}} = 10^{-15} \cdot \frac{\rho_0 c_p}{\Delta t} \cdot \int_{2007}^{2011} \int_{\text{bottom}}^{\text{top}} \int_{S_W} \theta v^\perp dS dz dt, \quad [\text{PW}] \quad (4.2)$$

Observing system	Observation operator [PW]	Noise covariance matrix [(PW) ²]
OSNAP-West	$\mathbf{f}^W = \text{MHT}_{\text{OSNAP-W}}$	$\mathbf{\Gamma}_{\text{noise}}^W = (0.004)^2$
OSNAP-East	$\mathbf{f}^E = \text{MHT}_{\text{OSNAP-E}}$	$\mathbf{\Gamma}_{\text{noise}}^E = (0.02)^2$
OSNAP-Combined	$\mathbf{f} = \begin{pmatrix} \text{MHT}_{\text{OSNAP-W}} \\ \text{MHT}_{\text{OSNAP-E}} \end{pmatrix}$	$\mathbf{\Gamma}_{\text{noise}} = \begin{pmatrix} (0.004)^2 & 0 \\ 0 & (0.2)^2 \end{pmatrix}$

Table 4.1: Definition of the OSNAP observing systems

and

$$\text{MHT}_{\text{OSNAP-E}} = 10^{-15} \cdot \frac{\rho_0 c_p}{\Delta t} \cdot \int_{2007}^{2011} \int_{\text{bottom}}^{\text{top}} \int_{S_E} \theta v^\perp dS dz dt. \quad [\text{PW}] \quad (4.3)$$

Here, $\Delta t = \int_{2007}^{2011} dt$ denotes the length of the integration period, $\rho_0 = 1029 \text{ kg/m}^3$ the reference density, $c_p = 3994 \text{ J/(kg} \cdot \text{K)}$ the heat capacity, θ potential temperature, and v^\perp the velocity perpendicular to the sections S_W and S_E , respectively. Sign convention is such that positive v^\perp corresponds to positive northward velocity. Model heat transports (4.2) and (4.3) are computed in depth coordinates. The solid orange and purple lines in Fig. 4.2 show the model heat transports across the OSNAP-West and OSNAP-East transects, respectively, as simulated by the ECCOv4r2 solution from January 1992 to December 2011. The dashed lines indicate the respective 2007-2011 means, i.e., the values of $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$, computed along the ECCOv4r2 solution.

In dynamics-based observing system design, an observing system is characterized by two pieces of information; the first one is *which quantities* are measured (and at which location and time), and the second one is *how accurate* these quantities are measured. These two pieces of information are formalized by, first, the observation operator \mathbf{f} , describing the observations' counterpart simulated by the model, and, second, the noise covariance matrix $\mathbf{\Gamma}_{\text{noise}}$, containing the assumed observational uncertainties. To illustrate the concepts of data redundancy and complementarity, I study three different observing systems: the OSNAP-West observing system, the OSNAP-East observing system and the OSNAP-Combined observing system. Table 4.1 shows the associated observation operators and noise matrices that characterize these observing systems. The OSNAP-West observing system is assumed to only measure the long-term mean heat transport across OSNAP-West (indicated by \mathbf{f}^W), and the OSNAP-East observing system is assumed to only measure the long-term mean heat transport across OSNAP-East (indicated by \mathbf{f}^E). The OSNAP-Combined observing system is assumed to have both of these measurements available (indicated by \mathbf{f}). The diagonal entries of the noise covariance matrices $\mathbf{\Gamma}_{\text{noise}}^W$, $\mathbf{\Gamma}_{\text{noise}}^E$, and $\mathbf{\Gamma}_{\text{noise}}$ in Table 4.1 are the uncertainty estimates for the respective time-mean heat transports from Lozier et al. [2019], cf. equation (4.1). Note that even though the Lozier et al. [2019] uncertainty estimates are computed for 21-month means, in this work I adopt these OSNAP uncertainty estimates for the respective 5-year means (for which no uncertainty estimates exist). Moreover, it is noted that the matrix $\mathbf{\Gamma}_{\text{noise}}$ is chosen to be diagonal. The missing off-diagonal entries are equivalent to the assumption that time-mean heat transport across OSNAP-West and time-mean heat transport across OSNAP-East are independent (or uncorrelated) observations. This

is a common, though very problematic, assumption in practical applications for complex inverse problems, also within the ECCOv4 framework, since the specification of accurate noise covariance matrices is considered as extremely difficult [Forget et al., 2015].

To study the effects of observational noise on the effective proxy potential, I further define

$$\mathbf{\Gamma}_{\text{noise}}^W[r^2] = r^2 \cdot \mathbf{\Gamma}_{\text{noise}}^W, \quad \mathbf{\Gamma}_{\text{noise}}^E[r^2] = r^2 \cdot \mathbf{\Gamma}_{\text{noise}}^E, \quad \text{and} \quad \mathbf{\Gamma}_{\text{noise}}[r^2] = r^2 \cdot \mathbf{\Gamma}_{\text{noise}}, \quad (4.4)$$

where $r^2 > 0$ is a constant number. Different choices of $r^2 > 0$ will be studied to consider the cases of less noisy ($r^2 < 1$) and more noisy ($r^2 > 1$) OSNAP data, while $r^2 = 1$ corresponds to the matrices shown in Table 4.1. Note that the ratio of the diagonal entries of $\mathbf{\Gamma}_{\text{noise}}[r^2]$ is independent of the scaling factor r^2 , equal to $(1/5)^2$, as in Lozier et al. [2019]. Moreover, the scaling factor r^2 does not affect the diagonal structure of $\mathbf{\Gamma}_{\text{noise}}$.

4.2.3 Quantities of interest

Proxy potential targets unobserved or unobservable quantities of interest (QoIs), remote from the observations. As a proof of concept, OSNAP's proxy potential is explored for the two QoIs that have already been studied in Chapter 2. The first QoI is ocean heat content in the Nordic Seas (OHC_{NS}), and the second one is volume transport across the Iceland-Scotland ridge (T_{IS}). The Nordic Seas basin and the Iceland-Scotland ridge are shown in Fig. 4.1. Consistent with considering the 5-year mean of the OSNAP observations, the proxy potential is evaluated for the 5-year mean of the chosen QoIs, within the same representative 5-year time window 2007-2011 (see Section 4.2.2).

In the model, the Iceland-Scotland ridge is defined as the composition of the shortest lines from Iceland (16°W , 65°N) to Faroe (7°W , 62.5°N), and Faroe (6.5°W , 62.5°N) to Scotland (4°W , 57°N), following grid line paths as described in Section 4.2.2. The obtained Iceland-Scotland model transect is denoted by S . The Nordic Seas domain is defined by the following boundaries: To the south by the S model transect, to the east by the Norwegian coast, to the north by the Barents Sea Opening, and to the west by the climatological 5°C contour of SST in the 20-year ECCOv4r2 solution. The obtained model domain is denoted by D . The model operators corresponding to the two chosen QoIs are given by

$$\text{OHC}_{\text{NS}} = \frac{1}{\Delta t} \rho_0 c_p \int_{2007}^{2011} \int_{\text{bottom}}^{\text{top}} \int_D \theta dA dz dt \quad [J], \quad (4.5)$$

and

$$\text{T}_{\text{IS}} = \frac{1}{\Delta t} \int_{2007}^{2011} \int_{\text{bottom}}^{\text{top}} \int_S v^\perp dS dz dt \quad [\text{m}^3/\text{s}]. \quad (4.6)$$

Here, $\Delta t = \int_{2007}^{2011} dt$ denotes the length of the integration period, $\rho_0 = 1029 \text{ kg/m}^3$ the reference density, $c_p = 3994 \text{ J}/(\text{kg} \cdot \text{K})$ the heat capacity, θ potential temperature, and v^\perp the velocity perpendicular to the section S ; sign convention is such that positive v^\perp

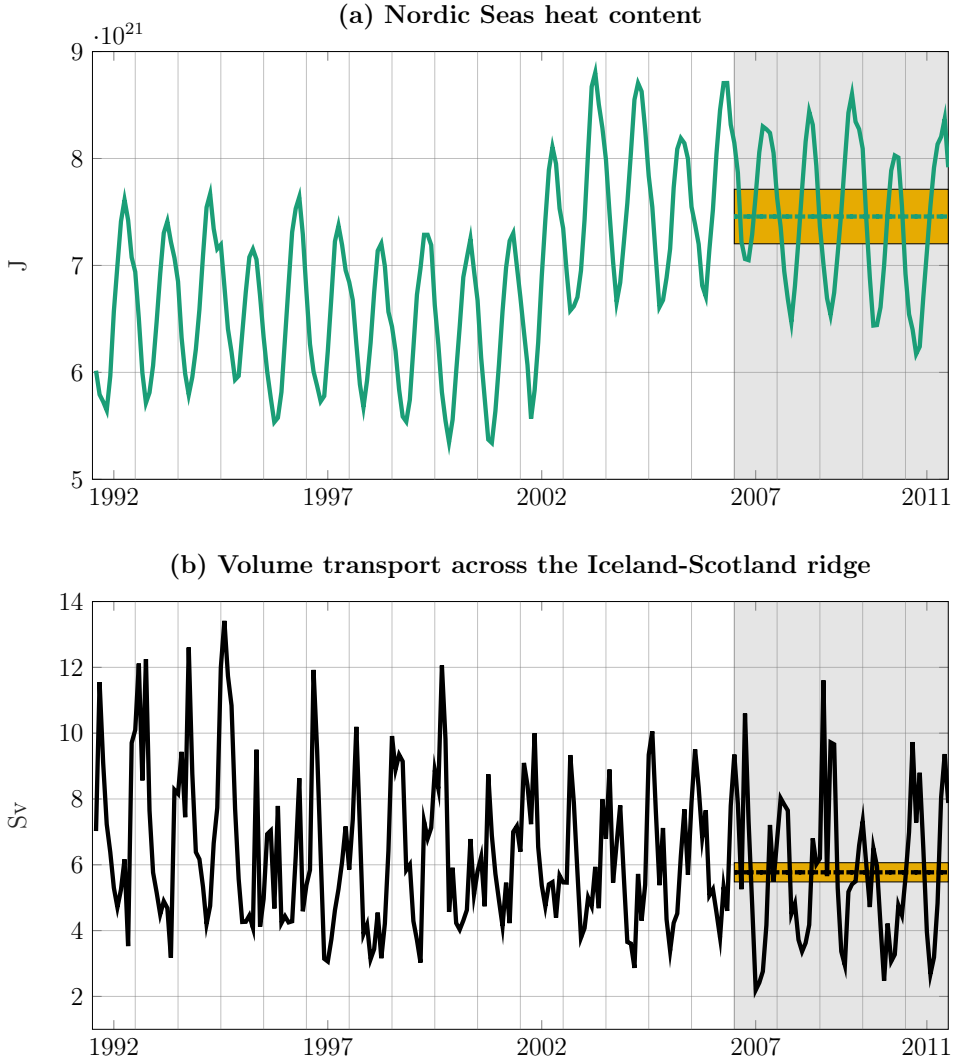


Figure 4.3: (a) Nordic Seas heat content and (b) volume transport across the Iceland-Scotland ridge in the ECCOv4r2 solution from January 1992 to December 2011 (solid lines). Nordic Seas heat content is integrated over the horizontal domain indicated by the green region in Fig. 4.1, and the entire water column. Volume transport across the Iceland-Scotland ridge is diagnosed across the black line in Fig. 4.1, and integrated over the entire water column. The gray background covers the final 5 years (2007-2011), the representative 5-year time window chosen for this work. The 5-year mean quantities, averaged over the period 2007-2011, are equal to OHC_{NS} and T_{IS} (defined in (4.5) and (4.6)) and are indicated by the dashed (constant) lines. The two yellow bars show the prior uncertainties for these mean values, computed by propagating the prior uncertainties Γ_{prior} from the control variables onto the respective model quantities. Displayed are two-standard-deviation envelopes, i.e., error bars of amplitude $\pm 2 \cdot \sigma_{\text{prior}}^A$ and $\pm 2 \cdot \sigma_{\text{prior}}^B$, respectively. The values for $(\sigma_{\text{prior}}^{\bullet})^2$ are shown in Table 4.3.

corresponds to positive north- and eastward velocity. The model volume transport (4.6) is computed in depth coordinates. The solid lines in Fig. 4.3 show Nordic Seas heat content, and volume transport across the Iceland-Scotland ridge, as simulated by the ECCOV4r2 solution from January 1992 to December 2011. The dashed lines indicate the respective 2007-2011 5-year means, i.e., the values of OHC_{NS} and V_{IS} , evaluated along the ECCOV4r2 solution.

The definition of T_{IS} in (4.6) is exactly the same as the definition used in Chapter 2. For OHC_{NS} , the definitions in this chapter and Chapter 2 differ slightly. Here, (4.5) defines Nordic Seas heat content integrated over the *entire water column*, consistent with the full-depth transports (4.2), (4.3), and (4.6) across the OSNAP sections and the Iceland-Scotland ridge, respectively. Meanwhile, Chapter 2 defined OHC_{NS} as *upper-ocean* heat content (over a slightly different Nordic Seas basin). In both chapters, sensitivities of OHC_{NS} with respect to atmospheric forcing play an important role. Even though heat content is defined for different parts of the water column, the sensitivity patterns to atmospheric forcing can be expected to be very similar across the two chapters. This is because temperature at intermediate depth and in the deep ocean is much less sensitive to atmospheric forcing, compared to temperature in the upper ocean. Minor differences are that, in this chapter, (i) local winds gain in importance relative to local heat fluxes because Ekman pumping can affect waters at depth more effectively than local heat fluxes, and (ii) remote heat fluxes gain in importance relative to remote heat fluxes because heat at depth can be most effectively altered via changes in the advected heat transport into the basin, which, in turn, are to a large degree caused by remotely forced changes in the overturning and gyre circulations. The differences (i) and (ii) are however found to be very small, and hardly discernible when comparing sensitivity maps across the chapters.

4.2.4 Controls and prior uncertainties

The vector \mathbf{x} of control variables used in this study is shown in Table 4.2; it includes initial conditions for the global three-dimensional fields of temperature, salinity, and velocity (first block in Table 4.2), the global three-dimensional fields of model parameters (second block in Table 4.2), and boundary conditions for the ocean model in the form of global two-dimensional atmospheric input fields (third block in Table 4.2). The adjustment frequency linked to these fields is reported in the second column of Table 4.2. Initial conditions are only adjusted at the first time step. For parameter fields and atmospheric boundary conditions, the (spatially-varying) time-mean over the complete estimation period is adjusted. The total number of control variables linked to the fields is shown in the third column of Table 4.2. It is equal to the number of ocean grid cells (for three-dimensional fields) and surface ocean grid cells (for two-dimensional fields), since the initial and time-mean fields are adjusted on a grid point basis during the optimization procedure.

While the total length of the control vector is on the order of $O(10^7)$, I arrange the control variables into the five groups *initial*, *param*, *precip*, *thermal*, and *wind*, as indicated by the fourth column of Table 4.2. Control variables linked to the group *thermal* can be thought of altering the surface heat flux through radiative, sensible, and latent

Description	Frequency	Size	Group	Prior Uncertainty
Initial temperature	N/A	$2.4 \cdot 10^6$	initial	ECCOV4r3, no WC01
Initial salinity	N/A	$2.4 \cdot 10^6$	initial	ECCOV4r3, no WC01
Initial zonal velocity	N/A	$2.4 \cdot 10^6$	initial	ECCOV4r3, no WC01
Initial meridional velocity	N/A	$2.4 \cdot 10^6$	initial	ECCOV4r3, no WC01
Diapycnal diffusivity	Time mean	$2.4 \cdot 10^6$	param	ECCOV4r3, no WC01
Isopycnal diffusivity	Time mean	$2.4 \cdot 10^6$	param	ECCOV4r3, no WC01
GM intensity	Time mean	$2.4 \cdot 10^6$	param	ECCOV4r3, no WC01
Precipitation	Time mean	$6.0 \cdot 10^5$	precip	ECCOV4r3, no WC01
Air temperature (2m)	Time mean	$6.0 \cdot 10^5$	thermal	ECCOV4r3, no WC01
Specific humidity (2m)	Time mean	$6.0 \cdot 10^5$	thermal	ECCOV4r3, no WC01
Longwave radiation (\downarrow)	Time mean	$6.0 \cdot 10^5$	thermal	ECCOV4r3, no WC01
Shortwave radiation (\downarrow)	Time mean	$6.0 \cdot 10^5$	thermal	ECCOV4r3, no WC01
Zonal wind stress	Time mean	$6.0 \cdot 10^5$	wind	ECCOV4r3, no WC01
Meridional wind stress	Time mean	$6.0 \cdot 10^5$	wind	ECCOV4r3, no WC01

Table 4.2: Control variables included in the control vector \mathbf{x} , and prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$ used in this study. Column 1 states the names of the physical two- and three-dimensional fields, column 2 the adjustment frequency, column 3 the total number of control variables related to the respective field, column 4 the groups under which the control variables are classified, and column 5 the prior uncertainties used for $\mathbf{\Gamma}_{\text{prior}}$. The setup is adopted from ECCOV4r3 (cf. Fukumori et al. [2017] and Table 7 in Forget et al. [2015]), except that, here, I omit (i) time-variable bi-weekly adjustments of the atmospheric control variables and (ii) Weaver and Courtier [2001] [WC01] smoothing for all prior uncertainties (see text).

heat fluxes, whereas control variables comprising the field of precipitation (*precip*) affect the surface freshwater flux. Specific humidity affects both the net heat flux (via latent heat), and the net freshwater flux (via evaporation), through the air-sea flux formulation from [Large and Yeager \[2004\]](#). I classify the control variables related to specific humidity under the group *thermal* because in this work, the spatial distribution of linearized sensitivity to specific humidity is generally found to resemble the spatial distribution of linearized sensitivity to air temperature and radiation. This suggests a similar effect of specific humidity on the studied model quantities as the purely thermal variables. It is noted that the five groups are defined solely for the sake of a more compact visualization of the results later, without an actual reduction of the control vector.

The choice of control variables, as shown in [Table 4.2](#), matches the set of control variables used in the ECCOv4 release 3 (ECCOv4r3) inverse modeling framework [[Fukumori et al., 2017](#)], but with one important modification. For the atmospheric boundary fields, ECCOv4r3 adjusts not only the time-means, as in my setup, but also time-variable anomalies on a bi-weekly basis. In this chapter, I omit time-variable adjustments of the atmospheric forcing variables for the purpose of a simpler demonstration. Including only the adjustments of the *time-mean* of the atmospheric forcing variables in the control vector is consistent with the simplification to explore the proxy potential of *time-mean* OSNAP observations for *time-mean* QoIs (see [Sections 4.2.2](#) and [4.2.3](#)). In fact, the time-mean OSNAP heat transports ([4.2](#)) and ([4.3](#)), as well as the time-mean unobserved QoIs ([4.5](#)) and ([4.6](#)) are about 9 times more sensitive to changes in the time-mean of the atmospheric forcing variables than to time-variable changes in the atmospheric forcing variables (see [Fig. C.1](#) in [Appendix C](#)). Therefore, in this work, the time-mean atmospheric controls are the dominant driver for dynamical adjustment mechanisms for time-mean OSNAP observations and unobserved QoIs. Since such dynamical adjustment mechanisms are the central ingredients for answering questions (I) and (II) (see [Section 4.1](#)), here, the results are relatively insensitive to the simplification of omitting the time-variable atmospheric controls. This is further discussed in [Appendix C](#).

Another, although minor, modification from ECCOv4r3 is that I do not include initial sea surface height in the control vector. The results are not sensitive to this choice because the sensitivity of the 5-year mean OSNAP observations and unobserved QoIs to initial conditions (first block in [Table 4.2](#)) is negligible compared to the sensitivity to parameter fields (second block in [Table 4.2](#)) and atmospheric controls (third block in [Table 4.2](#)), as [Section 4.3](#) will show. It is noted that initial sea surface height was not part of the control vector in the previous ECCO release (ECCOv4r2), either (cf. [Table 7](#) in [Forget et al. \[2015\]](#)). Another change from ECCOv4 release 2 to release 3 is the inclusion of initial velocity in the control vector, a change that I adopt in my setup. The augmentation of the control vector by parameters that determine parametrized turbulent transports (comprising the second block in [Table 4.2](#)) is a novelty of ECCO version 4, and not done in previous ECCO versions [[Forget et al., 2015](#)].

Assumptions on prior information about the control variables are made in the form of the specification of prior uncertainty, encapsulated in the prior covariance matrix Γ_{prior} . In this work, the prior is specified as a diagonal matrix, where the diagonal entries are adopted from the ECCOv4r3 inverse modeling framework (as reported by

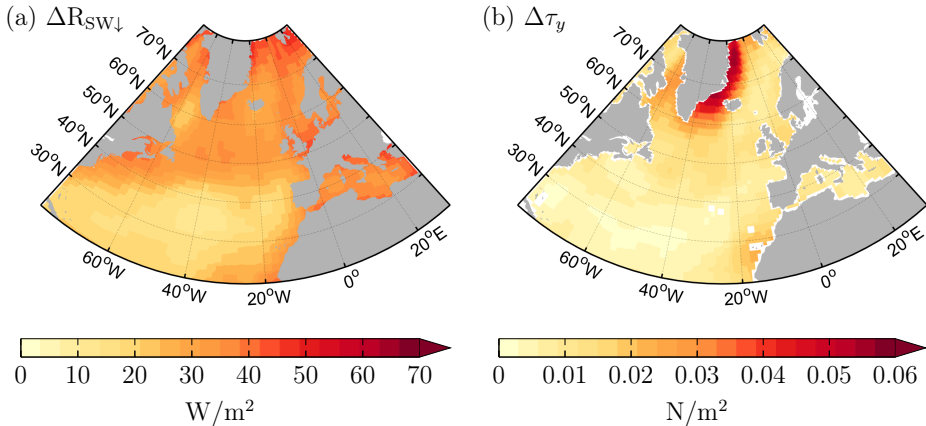


Figure 4.4: ECCOV4r3 prior standard deviations for the *time-mean* adjustments of the atmospheric forcing fields (a) shortwave downward radiation ($R_{SW\downarrow}$), and (b) meridional wind stress (τ_y).

the fifth column of Table 4.2). Since all off-diagonal entries in $\mathbf{\Gamma}_{\text{prior}}$ are zero, no prior cross-correlations are assumed in this work. In the ECCOV4r3 inverse modeling framework, a Weaver and Courtier [2001] smoother is applied to mimic prior cross-correlations between control variables that represent close-by grid points in the same field [Forget et al., 2015]. This smoother is not applied in this work (indicated by “no WC01” in Table 4.2), to simplify the presentation. The effect of including or excluding the Weaver and Courtier [2001] smoother is further discussed in Appendix C. Note that even in the full ECCOV4r3 framework, no prior cross-correlations between different types of control variables are assumed.

For the atmospheric boundary fields, the ECCOV4r3 prior uncertainties are an updated version of the uncertainties presented in Chaudhuri et al. [2012], computed based upon the spread of available atmospheric reanalysis products and satellite-derived observations. As an example, the prior time-mean uncertainty for two atmospheric forcing fields, shortwave downward radiation and meridional wind stress, are shown in Figs. 4.4(a) and (b), corresponding to Figs. 5(c) and 2(c) in Chaudhuri et al. [2012]. Time-mean uncertainty in shortwave downward radiation arises because downward radiation depends heavily on the parametrization of clouds. Therefore, regions of high cloud cover, such as the subpolar North Atlantic and the Nordic Seas shown in Fig. 4.4(a), have large systematic uncertainties in reanalysis products [Chaudhuri et al., 2012]. Time-mean uncertainty in meridional wind stress shows high values east and west of Greenland (Fig. 4.4(b)) because wind speeds are high in these regions. In many regions of strong winds, the spread of reanalysis products and satellite scatterometer measurements is enhanced [Chaudhuri et al., 2012]. Differences between reanalysis products and satellite scatterometer measurements arise because in reanalysis products (i) the impact of SST gradients on wind fields is not well-resolved [e.g., Chelton, 2005], and (ii) derived wind stress fields depend on the specific formulations of drag coefficient used in the bulk aerodynamic formulas [Risien and Chelton, 2008]. The square of the prior standard de-

QoI	Definition	Prior Uncertainty
MHT _{OSNAP-W}	(4.2)	$(\sigma_{\text{prior}}^W)^2 = (0.0024)^2$ [(PW) ²]
MHT _{OSNAP-E}	(4.3)	$(\sigma_{\text{prior}}^E)^2 = (0.0045)^2$ [(PW) ²]
OHC _{NS}	(4.5)	$(\sigma_{\text{prior}}^A)^2 = (1.28 \cdot 10^{20})^2$ [(J) ²]
T _{IS}	(4.6)	$(\sigma_{\text{prior}}^B)^2 = (0.146)^2$ [(Sv) ²]

Table 4.3: Prior uncertainties of the studied 5-year mean observed and unobserved model quantities of interest (QoIs). The prior uncertainties $(\sigma_{\text{prior}})^2$ shown in the third column are obtained by propagating the prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$ in the control variables to the respective QoI (along the orange path (2) in Fig. 3.1). Here, $(\sigma_{\text{prior}})^2$ computes as the projection shown in equation (4.7).

viations shown in Fig. 4.4, at each grid point, is used as a diagonal entry for the diagonal matrix $\mathbf{\Gamma}_{\text{prior}}$. It is noted that, despite the shown restriction to the North Atlantic, control variables in this work are defined on a grid point basis of the underlying *global* ocean model (adopted from ECCOv4r3).

The ECCOv4r3 prior uncertainties for the three-dimensional parameter fields of diapycnal diffusivity, isopycnal diffusivity, and GM intensity are set to the spatially constant values of $5 \cdot 10^{-6} \text{ m}^2/\text{s}^2$, $250 \text{ m}^2/\text{s}^2$ and $250 \text{ m}^2/\text{s}^2$, respectively. It is noted that the values for the spatially constant prior uncertainties of these three parameter fields have been lowered by more than an order of magnitude from ECCOv4 release 2 to release 3 (personal communication with Ou Wang, August 2018).¹ All ECCOv4r3 prior uncertainties are scaled by a (spatially-varying) grid area weighting to make adjustments independent of the chosen model grid [Fukumori et al., 2017]. The penalty weights corresponding to all ECCOv4r3 prior uncertainties are available for download at ftp://ecco.jpl.nasa.gov/Version4/Release3/input_init/error_weight/ctrl_weight/.

In an inverse modeling framework, all uncertain elements of the model are assumed to be captured by the chosen set of control variables together with the associated prior covariance matrix. To obtain the prior uncertainty σ_{prior}^2 for an arbitrary model quantity QoI(\mathbf{x}), one therefore propagates the prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$ in the control variables to the model quantity of interest (along the orange path (2) in Fig. 3.1) by means of the linearized model equations. Due to the construction of $\mathbf{\Gamma}_{\text{prior}}$ as a diagonal matrix, with diagonal entries as in ECCOv4r3 (cf. fifth column of Table 4.2), this forward uncertainty propagation simplifies to the projection

$$\sigma_{\text{prior}}^2 = \sum_{i=1}^N \left(\frac{\partial(\text{QoI})}{\partial x_i} \cdot \Delta x_i \right)^2, \quad (4.7)$$

where Δx_i is the ECCOv4r3 prior standard deviation corresponding to the control variable x_i , and N is the total number of control variables. The specified uncertainties $\mathbf{\Gamma}_{\text{prior}}$ for the control variables can be examined through the lens of projected prior uncertainties (4.7). The projected uncertainties are easier to inspect since they are scalars,

¹ In release 2, the values for diapycnal diffusivity, isopycnal diffusivity, and GM intensity were chosen as $5000 \text{ m}^2/\text{s}^2$, $5000 \text{ m}^2/\text{s}^2$ and $10^{-4} \text{ m}^2/\text{s}^2$, although [Forget et al., 2015] mistakenly report $500 \text{ m}^2/\text{s}^2$ for the first two fields (personal communication with Ou Wang).

rather than a collection of $O(10^7)$ numbers (or more, if the prior covariance matrix has off-diagonal entries). Column 3 of Table 4.3 shows the values for the prior uncertainties that are obtained when projecting $\mathbf{\Gamma}_{\text{prior}}$ via equation 4.7 onto the 5-year mean model quantities $\text{MHT}_{\text{OSNAP-W}}$, $\text{MHT}_{\text{OSNAP-E}}$, OHT_{NS} , and T_{IS} , defined in Sections 4.2.2 and 4.2.3. Furthermore, the yellow bars in Figs. 4.2 and 4.3 visualize the respective computed prior uncertainties. Shown are two-standard-deviation envelopes, i.e., error bars of amplitude $\pm 2 \cdot \sigma_{\text{prior}}$. Inspecting the yellow error bars in Figs. 4.2 and 4.3, which reflect the uncertainties in the respective model quantities *before* adding any data constraints, leaves the impression that the prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$ are chosen too small. It is noted that if time-variable adjustments of the atmospheric control variables were included, such as done by ECCOV4r3 but not in this work, the projected prior uncertainties σ_{prior}^2 in Table 4.3 would only be increased by at most 6%, depending on the model quantity (see Fig. C.1 in Appendix C). Moreover, if the full ECCOV4r3 inverse modeling framework was adopted, by adding not only time-variable atmospheric controls but also Weaver and Courtier [2001] smoothing for all control variables to the setup in this work, the computed prior uncertainties σ_{prior}^2 would be smaller than in Table 4.3 (see Fig. C.1 in Appendix C). The issue of seemingly too small ECCOV4r3 prior uncertainties is further discussed in Section 4.4.

Due to the commonly known difficulty of specifying accurate prior covariance matrices (see Section 3.6 and e.g., Forget et al., 2015), the effects of the chosen prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$ on the effective proxy potential will be assessed. To this aim, I define

$$\mathbf{\Gamma}_{\text{prior}}[b^2] = b^2 \cdot \mathbf{\Gamma}_{\text{prior}}, \quad (4.8)$$

where $b^2 > 0$ is a constant number. Different choices of $b^2 > 0$ will be studied to consider the assumptions of more certain ($b^2 < 1$) and less certain ($b^2 > 1$) prior information on the chosen control variables than in ECCOV4r3. The case $b^2 = 1$ corresponds to choosing $\mathbf{\Gamma}_{\text{prior}}$, adopted from ECCOV4r3 (cf. fifth column of Table 4.2). Note that the relative weighting of the control variables, i.e., the ratio of the diagonal entries of $\mathbf{\Gamma}_{\text{prior}}[b^2]$, is independent of the scaling factor b^2 , and equal to the relative weighting as specified within the ECCOV4r3 inverse modeling framework. Moreover, the scaling factor b^2 does not affect the diagonal structure of $\mathbf{\Gamma}_{\text{prior}}$.

4.2.5 Adjoint models

To answer questions (I) and (II) from Section 4.1, or, formally, to propagate uncertainty via the GCM equations (cf. equation 4.7), one requires the linearized sensitivities

$$\frac{\partial(\text{MHT}_{\text{OSNAP-W}})}{\partial \mathbf{x}}, \quad \frac{\partial(\text{MHT}_{\text{OSNAP-E}})}{\partial \mathbf{x}}, \quad \frac{\partial(\text{OHC}_{\text{NS}})}{\partial \mathbf{x}}, \quad \frac{\partial(\text{T}_{\text{IS}})}{\partial \mathbf{x}} \quad (4.9)$$

with respect to the entire control vector \mathbf{x} . To this aim, I perform four adjoint sensitivity experiments, taking advantage of the flexible ECCOV4 adjoint modeling framework. Algorithmic differentiation, through source-to-source code transformation with the commercial tool transformation of algorithms in Fortran (TAF; [Giering and Kaminski, 1998]), produces the code for the adjoint models. For each of the four experiments, I redefine the least-squares cost function, used in the ECCOV4 optimization problem, to a new scalar (or objective) function. In the first two experiments, the objective

functions are the two components of the observation operator $\mathbf{f}(\mathbf{x})$, $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$, defined in (4.2) and (4.3). In the last two experiments, the objective functions are given by the unobserved quantities of interests OHC_{NS} and T_{IS} , defined in (4.5) and (4.6). All objective functions are defined as the 5-year mean over the period 2007-2011. Adjoint sensitivities are computed along the model trajectory given by the final 5 years (2007-2011) of the ECCOv4r2 estimate. In each of the experiments, the adjoint model computes the linearized sensitivities of the objective function to all control variables in a single model integration.

Since ECCOv4 does not include an adjoint representation of the sea-ice model, linearized sensitivities to the atmospheric boundary fields are corrected by a factor of $1 - \alpha$, where α is the fractional coverage of sea-ice area in each model grid cell. As an example, the sensitivity to atmospheric forcing in a completely ice-covered grid cell ($\alpha = 1$) is set to zero.

It is worth noting that the greatest part of the control vector in this work (second and third block in Table 4.2) consists of adjusting the *time-mean* of a certain control variable. The linearized sensitivity to the time-mean of a control variable corresponds to accumulating sensitivities to changes in that control variable from all lead times. When accumulating sensitivities over time, cancellations occur at the locations where sensitivities have opposite signs for different lead times. Features in the sensitivity maps which carry through the accumulation process are either the ones (i) of strongest amplitude, typically found for short lead times, or (ii) stationary on a five-year timescale.

Finally, it is noted that choosing the 2007-2011 ECCOv4r2 solution as a base state for the linearized sensitivities in (4.9), formally corresponds to “adopting” the 2007-2011 ECCOv4r2 solution as \mathbf{x}_{\min} (cf. Sections 3.3 and 3.5). Forget et al. [2015] obtained the ECCOv4r2 solution by minimizing a least-squares cost function of type (3.1), which combines the MITgcm with several hundred million satellite and in situ ocean observations. Strictly speaking, a separate minimization problem would have to be solved to find the minimizer \mathbf{x}_{\min} corresponding to the least-squares cost function in this work; this cost function does not include the observations used in Forget et al. [2015], but the observations introduced in Section 4.2.2. However, the focus of this work is not set on solving an optimization problem - least of all, an optimization problem which includes only 2 observations but $O(10^7)$ control variables - but on the novel task of exploring proxy potential in an oceanographic context. In other words, this work is not about estimating the exact QoI trajectories in Fig. 4.3, or their 2007-2012 means (dashed lines), but rather about studying their uncertainties (the yellow bars), and how the integration of the OSNAP observations would reduce these uncertainty bars. The uncertainty bars in Fig. 4.3, as well as their reduction by observational information, depend only insignificantly on the QoI base trajectory, if one assumes that the system is “not too non-linear” (see the discussion in Section 3.6). Assessing a change in uncertainty (the yellow bars), rather than the change in the actual estimate (the dashed lines), explores the general *capability* or *potential* for the OSNAP observations to inform the QoIs - therefore the chosen terminology ‘proxy *potential*’. This assessment does not require the actual OSNAP measurement values, as discussed before. Meanwhile, if proxy potential is identified, having the actual OSNAP measurement values available will allow a better constrained estimate

of the QoI (i.e., improve the value for the dashed lines).

4.3 Results

4.3.1 Dynamical adjustment mechanisms for OSNAP

This section tackles question (I) from Section 4.1: the dynamical adjustment mechanisms and pathways for the time-mean heat transports across OSNAP-West ($\text{MHT}_{\text{OSNAP-W}}$) and OSNAP-East ($\text{MHT}_{\text{OSNAP-E}}$) are identified.

To this aim, I employ the linearized sensitivities $\frac{\partial \text{MHT}_{\text{OSNAP-W}}}{\partial \mathbf{x}}$ and $\frac{\partial \text{MHT}_{\text{OSNAP-E}}}{\partial \mathbf{x}}$, which are obtained by the adjoint models described in Section 4.2.5. The maps in Fig. 4.5 show the linearized sensitivities of $\text{MHT}_{\text{OSNAP-W}}$ (in (a),(c)) and $\text{MHT}_{\text{OSNAP-E}}$ (in (b),(d)) to changes in the time-mean of shortwave downward radiation ($R_{\text{SW}\downarrow}$, in (a),(b)) and meridional wind stress (τ_y , in (c),(d)). The linearized sensitivities shown in Fig. 4.5 are “prior-weighted”, i.e., multiplied by the prior ECCOV4r3 standard deviation fields $\Delta R_{\text{SW}\downarrow}$ (in Figs. 4.5(a),(b)) and $\Delta \tau_y$ (in Figs. 4.5(c),(d)). The weights $\Delta R_{\text{SW}\downarrow}$ and $\Delta \tau_y$ are shown in Figs. 4.4(a) and (b), respectively. Moreover, the prior-weighted sensitivities are normalized, i.e., the prior-weighted sensitivities of $\text{MHT}_{\text{OSNAP-W}}$ are divided by the constant value of σ_{prior}^W , and the prior-weighted sensitivities of $\text{MHT}_{\text{OSNAP-E}}$ divided by σ_{prior}^E . The (squared) values for σ_{prior}^W and σ_{prior}^E are reported in Table 4.3. σ_{prior}^W and σ_{prior}^E are the prior uncertainties in $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$, respectively. These prior uncertainties are computed by accumulating prior-weighted sensitivity to *all* control variables, as shown in equation (4.7), where QoI in (4.7) is chosen as $\text{MHT}_{\text{OSNAP-W}}$ or $\text{MHT}_{\text{OSNAP-E}}$, respectively. The sensitivity maps in Fig. 4.5 are unitless, since they are prior-weighted and normalized. Moreover, due to the normalization, all absolute values seen in the maps of Figs. 4.5(a)-(d) are less than 1.

The sensitivity maps in Fig. 4.5 reveal all possible dynamical mechanisms via which small-amplitude perturbations in the time-mean shortwave downward radiation and meridional wind stress can impact the time-mean heat transport across OSNAP-West and OSNAP-East on a five-year timescale. The many different features that compose the maps in Fig. 4.5 reflect the fact that the heat transports across the OSNAP sections are the superposition of many signals that are propagated to the OSNAP sections by different adjustment mechanisms, via different pathways. The sensitivity maps in Fig. 4.5 disentangle these different mechanisms and pathways, which will be further explored in the following.

Sensitivity to thermal forcing

Figs. 4.5(a) and (b) are proportional (up to normalizing constants) to the prior-weighted linearized sensitivities of 5-year mean heat transport across OSNAP-West (Fig. 4.5(a)) and of heat transport across OSNAP-East (Fig. 4.5(b)) to changes in the 5-year mean of shortwave downward radiation $R_{\text{SW}\downarrow}$. The two-dimensional time-mean field of shortwave downward radiation has been classified under the group *thermal* in Table 4.2, together with the two-dimensional time-mean fields of air temperature at 2 m, specific humidity at 2 m, and downward longwave radiation. The pattern in Fig. 4.5(a) can be considered

Dynamical adjustment mechanisms for OSNAP

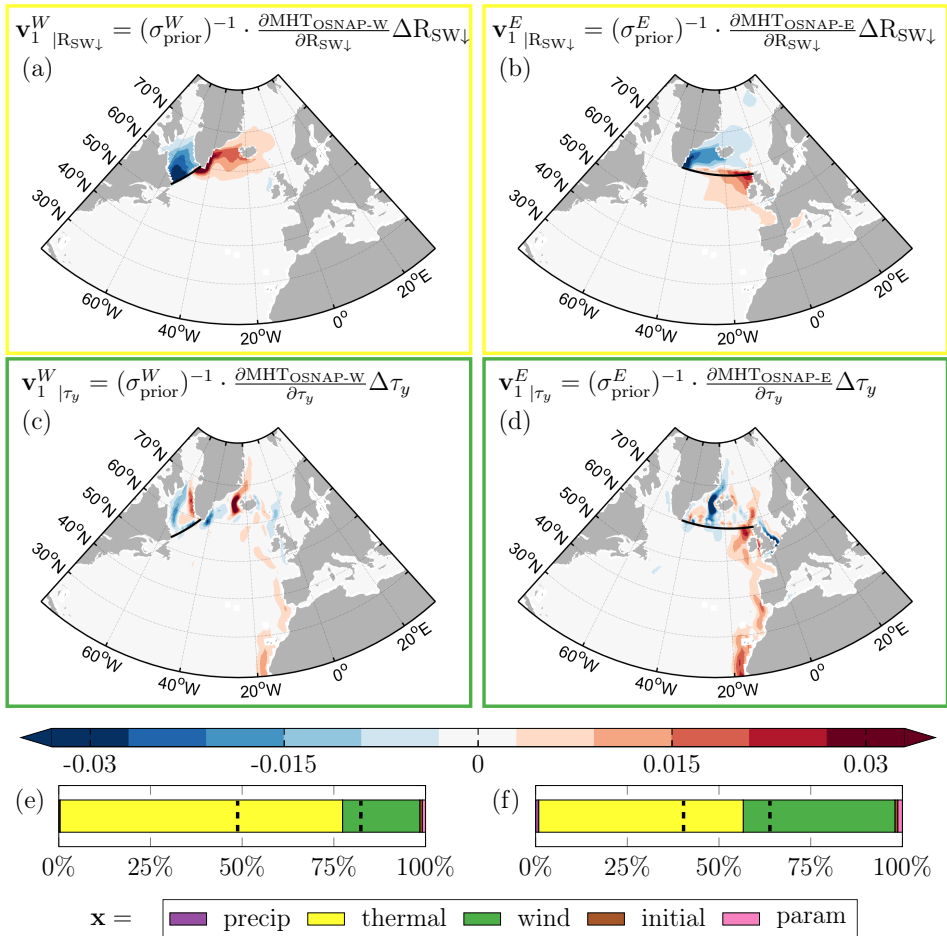


Figure 4.5: Linearized sensitivities of time-mean heat transport across (a),(c) OSNAP-West ($\text{MHT}_{\text{OSNAP-W}}$) and (b),(d) OSNAP-East ($\text{MHT}_{\text{OSNAP-E}}$) to time-mean (a),(b) shortwave downward radiation $R_{SW\downarrow}$ and (c),(d) meridional wind stress τ_y . The sensitivities are prior-weighted and normalized. Red (blue) colors indicate that an increase in (a),(b) downward radiation and (c),(d) northward wind stress would lead to a subsequent increase (decrease) in (a),(c) $\text{MHT}_{\text{OSNAP-W}}$ and (b),(d) $\text{MHT}_{\text{OSNAP-E}}$ on a 5 year timescale. The solid black contours delineate the cross-sectional (a),(c) OSNAP-West and (b),(d) OSNAP-East transects. (a),(c) is a clipped version of the vector \mathbf{v}_1^W , the direction informed by the OSNAP-West observing system (cf. Table 4.1). Similarly, (b),(d) is a clipped version of the vector \mathbf{v}_1^E , the direction informed by the OSNAP-East observing system (cf. Table 4.1). The bar charts show the relative importance (see definition 4.10) of the control variables for impacting (e) $\text{MHT}_{\text{OSNAP-W}}$ and (f) $\text{MHT}_{\text{OSNAP-E}}$. The control variables are grouped according to Table 4.2. In (e) and (f), the fractions right of the vertical dashed line within the yellow bar and within the green bar are the relative importance of shortwave downward radiation and meridional wind stress, respectively.

representative of prior-weighted sensitivities of $\text{MHT}_{\text{OSNAP-W}}$ to all 4 thermal forcing fields because prior-weighted sensitivities of $\text{MHT}_{\text{OSNAP-W}}$ to these 4 fields all show the same qualitative picture. This is because the thermal forcing fields all drive changes in surface heat flux, which results in similar changes in $\text{MHT}_{\text{OSNAP-W}}$ through similar oceanic dynamical adjustment mechanisms. For a discussion of these physical mechanisms, it is therefore enough to focus on sensitivities to shortwave downward radiation. Similarly, Fig. 4.5(b) can be seen as representative of sensitivities of $\text{MHT}_{\text{OSNAP-E}}$ to all thermal forcing fields. The yellow frames around Figs. 4.5(a) and (b) highlight the figures' representative function for the entire group *thermal*.

Red (blue) colors in Figs. 4.5(a) and (b) indicate the locations where an increase (decrease) in downward radiation would lead to a subsequent increase in northward heat transport across OSNAP-West (Fig. 4.5(a)) and OSNAP-East (Fig. 4.5(b)) on a five-year timescale. The maps in Figs. 4.5(a) and (b) show a similar pattern. Positive sensitivity is found at the respective eastern ends of the OSNAP-West section (in (a)), and the OSNAP-East section (in (b)), and negative sensitivities at the respective western ends of OSNAP-West section (in (a)) and the OSNAP-East section (in (b)). Moreover, in Figs. 4.5(a) and (b), positive sensitivity extends southward of the respective OSNAP section, and upstream the surface current which leads up to the respective section (cf. Fig. 4.1). Specifically, positive sensitivity can be found upstream around southern Greenland and along the Irminger Current in Fig. 4.5(a), and upstream the North Atlantic Current south of the OSNAP-East transect in Fig. 4.5(b). On the other hand, negative sensitivity extends northward of the sections in Figs. 4.5(a) and (b), following again the surface current upstream: negative sensitivity populates the Labrador Sea in Fig. 4.5(a), and the Irminger Current in Fig. 4.5(b).

The described sensitivity pattern, shared by Figs. 4.5(a) and (b), can be explained by the combination of two adjustment mechanisms, which alter the heat transport across the OSNAP-West and OSNAP-East section, respectively. The first mechanism is *advection* of temperature anomalies across the sections. Increased (decreased) downward radiation induces warm (cold) ocean temperature anomalies. A perturbation in downward radiation matching the distributions in Figs. 4.5(a) and (b) induces warm anomalies to the south of the sections, and cold anomalies to the north of the sections. This distribution of temperature anomalies leads to a strengthened northward heat transport, when the temperature anomalies are advected across the sections by the surface currents that lead up to the sections. Advection of temperature anomalies goes hand in hand with a second mechanism, namely the dynamics of *thermal wind balance*. Warm anomalies at the eastern end of the sections, and cold anomalies at the western end of the sections generate a negative anomaly in the upper-ocean density gradient across the sections. According to thermal wind balance (equation (2.5)), this negative density gradient anomaly strengthens the northward geostrophic current across the sections.

Sensitivity to wind forcing

Figs. 4.5(c) and (d) are proportional to the prior-weighted linearized sensitivities of $\text{MHT}_{\text{OSNAP-W}}$ (Fig. 4.5(c)) and $\text{MHT}_{\text{OSNAP-E}}$ (Fig. 4.5(d)) to meridional wind stress τ_y . The two-dimensional time-mean fields of zonal and meridional wind stress comprise the

group *wind* in Table 4.2. Even though sensitivity patterns for zonal and meridional wind stress differ by nature, they reflect similar wind-driven mechanisms that can alter heat transport across the OSNAP array. In this work, zonal wind stress sensitivities are generally weaker than meridional wind stress sensitivities, and I therefore choose to present and discuss the latter. The green frames around Figs. 4.5(c) and (d) highlight the figures' representative function for the group *wind*.

The most noticeable features in Fig. 4.5(c) are (i) negative sensitivities concentrated along the OSNAP-West transect, (ii) a sensitivity dipole in the Labrador Sea, which continues northward into the Davis Strait, (iii) negative sensitivities along the southeastern Greenlandic coast, and (iv) a band of strong positive sensitivities along the western Icelandic coast. Fig. 4.5(d) strikes with (v) a band of strong negative sensitivities along the western Icelandic coast. A shared feature of Figs. 4.5(c) and (d) is (vi) a band of positive sensitivity along the western European and African coastline. This band continues northwestward toward the Irminger Sea in Fig. 4.5(c) and northward into the Nordic Seas in Fig. 4.5(d).

Feature (i) in Fig. 4.5(c), together with negative *zonal* wind stress sensitivities of $MHT_{OSNAP-W}$ along the OSNAP-West transect (not shown), can be explained by local *Ekman dynamics*. A wind stress anomaly that points to the southwest along the OSNAP-West section induces an anomalous Ekman transport across OSNAP-West in northwestern direction. Feature (i) has no counterpart in Fig. 4.5(d), i.e., a sensitivity band along the OSNAP-East array is absent, because the OSNAP-East leg is much less tilted in north-south direction than the OSNAP-West leg. Instead, a band of negative sensitivities along the OSNAP-East section can be detected in the sensitivity map of $MHT_{OSNAP-E}$ to *zonal* wind stress (not shown).

Feature (ii), the sensitivity dipole in Fig. 4.5(c), is due to *geostrophic balance* (2.6). A meridional wind stress perturbation that matches the sensitivity dipole in the Labrador Sea and the Davis Strait drives Ekman onshore convergence along the west Greenlandic coast, and Ekman offshore divergence along the coasts of Newfoundland, Labrador and Baffin Island. This results in a positive anomaly of the zonal pressure gradient, and a stronger northward geostrophic transport, and thus, an increase in $MHT_{OSNAP-W}$. Feature (iii) in Fig. 4.5(c) is closely linked to feature (ii). A southward wind stress anomaly along the southeastern Greenlandic coast drives Ekman onshore convergence and a positive pressure anomaly. This positive pressure anomaly can be communicated by Kelvin waves around the southern tip of Greenland to the eastern end of the OSNAP-West section, where it contributes to a strengthening of the northward geostrophic transport across the OSNAP-West transect.

Feature (vi) in Figs. 4.5(c) and (d) was already identified in Chapter 2, and can be explained by a similar wave mechanism as discussed in the last paragraph. Northward wind stress perturbations along the western European and African coastline drive Ekman onshore convergence. The resulting positive pressure anomalies are carried northward by coastally-trapped Kelvin waves. The induced positive zonal pressure anomaly induces a strengthened northward geostrophic transport which increases both $MHT_{OSNAP-W}$ and $MHT_{OSNAP-E}$. At about 55°N , however, the bands in Figs. 4.5(c) and (d) follow sep-

arate branches of the North Atlantic Current: The extension of the sensitivity band is directed northwestward toward the Irminger Sea in Fig. 4.5(c) and northward toward the Nordic Seas in Fig. 4.5(d). Both extensions follow contour lines of f/H , and can therefore be explained by a pressure gradient adjustment mechanism, similarly as described above. This mechanism strengthens the Irminger current and $\text{MHT}_{\text{OSNAP-W}}$ in Fig. 4.5(c), and the Norwegian Atlantic Current and $\text{MHT}_{\text{OSNAP-E}}$ in Fig. 4.5(d).

Finally, features (iv) and (v) highlight the Icelandic coast as a key location where wind stress perturbations can remotely alter changes in subpolar ocean circulation. The Icelandic coastline has already been found highly sensitive to wind stress perturbations for the quantities that were studied in Chapter 2. Here, high sensitivities to meridional wind stress along the western coast of Iceland are amplified by high prior ECCOv4r3 meridional wind stress uncertainties $\Delta\tau_y$ at this location (see Fig. 4.4(b)). These two factors combined lead to very high amplitudes for the prior-weighted sensitivities along the western Icelandic coast in Figs. 4.5(c) and (d), with values falling out of the shown colorbar range: normalized prior-weighted sensitivities to meridional wind stress take values up to +0.08 in Fig. 4.5(c) and down to -0.12 in Fig. 4.5(d). The positive sensitivities of $\text{MHT}_{\text{OSNAP-W}}$ (in Fig. 4.5(c)) and the negative sensitivities of $\text{MHT}_{\text{OSNAP-E}}$ (in Fig. 4.5(d)) to meridional wind stress along the western Icelandic coast are due to a similar physical mechanism as the one discussed in Chapter 2. A southward wind stress anomaly along the western Icelandic coast, matching the negative sensitivities in Fig. 4.5(d), drives Ekman offshore divergence, resulting in a negative pressure anomaly. The negative pressure anomaly is carried by a Kelvin wave clockwise around Iceland, to the eastern Icelandic coast. The induced positive zonal pressure anomaly along the section between Iceland and Scotland strengthens the geostrophic current of the Norwegian Atlantic Current (NwAC, cf. Fig. 4.1). A strengthened NwAC causes a stronger northward transport across OSNAP-East, and an increase in $\text{MHT}_{\text{OSNAP-E}}$. In contrast, a northward wind stress perturbation along the western Icelandic coast, matching the positive sensitivities in Fig. 4.5(c), *weakens* the NwAC, through the same chain of events, but involving a pressure anomaly of opposite sign. A weakened NwAC leads to a relatively stronger Irminger Current (IC), which, besides the NwAC, is a second branch of the North Atlantic Current (Fig. 4.1). A strengthened IC drives a stronger northward transport across OSNAP-West, and thus, an increase in $\text{MHT}_{\text{OSNAP-W}}$.

Relative importance of control variables

The relative importance of the control variables for impacting the observed quantities $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$ are shown by the bar charts in Figs. 4.5(e) and (f). Here, the control variables are grouped according to Table 4.2. Specifically, the bar charts show the ratios

$$\frac{\sum_{i|x_i \in \text{group}} \left(\frac{\partial \text{MHT}_{\text{OSNAP-W}}}{\partial x_i} \cdot \Delta x_i \right)^2}{\sum_{i=1}^N \left(\frac{\partial \text{MHT}_{\text{OSNAP-W}}}{\partial x_i} \cdot \Delta x_i \right)^2} \quad \text{and} \quad \frac{\sum_{i|x_i \in \text{group}} \left(\frac{\partial \text{MHT}_{\text{OSNAP-E}}}{\partial x_i} \cdot \Delta x_i \right)^2}{\sum_{i=1}^N \left(\frac{\partial \text{MHT}_{\text{OSNAP-E}}}{\partial x_i} \cdot \Delta x_i \right)^2} \quad (4.10)$$

in Fig. 4.5(e) and Fig. 4.5(f), respectively. Here, Δx_i is the ECCOv4r3 prior standard deviation corresponding to the control variable x_i , and N is the total number of control variables. Note that the denominators in (4.10) are the prior uncertainties $(\sigma_{\text{prior}}^W)^2$ and

$(\sigma_{\text{prior}}^E)^2$, computed by accumulating prior-weighted sensitivity to *all* control variables (cf. equation (4.7)). Meanwhile, the enumerators in (4.10) are the prior uncertainties due to the control variables in the specific group, computed by accumulating prior-weighted sensitivity only to the control variables in that group.

The bar charts in Fig. 4.5(e) and (f) are useful to obtain a first impression of which of the control variables can be informed by the OSNAP-West and OSNAP-East observations. For example, in Fig. 4.5(e), the large yellow and green fractions indicate that control variables belonging to the group *thermal* or *wind* have high potential to be informed by observed heat transport across OSNAP-West, whereas control variables belonging to the groups *precip*, *initial*, and *param*, which $\text{MHT}_{\text{OSNAP-W}}$ is relatively insensitive to, will stay uninformed by the OSNAP-West.

Both observed quantities $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$ are most sensitive to the group of thermal forcing fields. The relative importance of control variables classified under the group *thermal* is 77% for $\text{MHT}_{\text{OSNAP-W}}$ (Fig. 4.5(e)), and 56% for $\text{MHT}_{\text{OSNAP-E}}$ (Fig. 4.5(f)). The group with the second largest influence is *wind*; the relative contribution by wind-driven processes, represented by the color green, is 21% to $\text{MHT}_{\text{OSNAP-W}}$ (Fig. 4.5(e)) and 41% to $\text{MHT}_{\text{OSNAP-E}}$ (Fig. 4.5(f)). While thermally-driven processes (integrated throughout the globe) are most important for both $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$, wind-driven processes contribute to a higher degree to $\text{MHT}_{\text{OSNAP-E}}$ than $\text{MHT}_{\text{OSNAP-W}}$. While further investigation is required to pinpoint the underlying reason, I suggest that the following two facts play a role: First, the identified non-local wind-driven pressure adjustment mechanisms (iv), (v), and (vi), shared between OSNAP-West and OSNAP-East, have their source and “center of action” closer in space to OSNAP-East than OSNAP-West. Until the signal reaches OSNAP-West via the Irminger Current, some of the signal’s strength might have been lost underway. Second, $\text{MHT}_{\text{OSNAP-W}}$ might be more sensitive to changes in thermal forcing than $\text{MHT}_{\text{OSNAP-E}}$ because the OSNAP-West section cuts across the Labrador Sea, which features the deepest mixed layers in the Northern hemisphere. Therefore, thermal surface forcing that is applied in the vicinity of the OSNAP-West transect can impact the temperature of a deeper water column, and, thus, alter heat transport more effectively, than thermal surface forcing close to the OSNAP-East transect. It is noted that stronger local prior-weighted thermal sensitivities for OSNAP-West are *not* due to the prior ECCOv4 uncertainties. In fact, prior ECCOv4 uncertainties for shortwave downward radiation in the vicinity of OSNAP-West are smaller (rather than larger) than in the vicinity of OSNAP-East (cf. Fig. 4.4(a)).

The groups *precip*, *initial*, and *param* have minimal impact. The insignificant contributions by the groups *precip* and *initial* may be partly explained by the fact that prior-weighted sensitivities on a *five-year* timescale are studied. Precipitation is expected to play a more important role on timescales exceeding 5 years, through changes in the thermohaline circulation, e.g., driven by freshwater flux anomalies in the deep convection regions (cf. Fig. 2.12 or Fig. 2.14 in Chapter 2, which show that freshwater flux anomalies in the deep convection regions affect subpolar gyre circulation on lead times of >5 years). In contrast, initial conditions are expected to be more influential on shorter (daily to weekly) timescales. However, the main underlying reason for the neg-

ligible influence of the groups *initial* and *param* is suggested to be an underestimation of the prior uncertainties for the groups *initial* and *param* in ECCOV4r3 (and prior uncertainties are adopted from ECCOV4r3 in this work). In fact, the ECCOV4r3 standard deviations for *param* are about an order of magnitude smaller than the ECCOV4r2 standard deviations (see Section 4.2.4). Moreover, *effective* prior uncertainties for *initial* and *param* in ECCOV4r3 are hypothesized to be much larger than the reported ECCOV4r3 regularization weights, because prior uncertainties are “lost in (pre-)optimization” steps, as further discusses in Section 4.4.

The two dashed vertical lines in each of the bar charts of Figs. 4.5(e) and (f) establish a link to Figs. 4.5(a)-(d). The chunk right of the dashed vertical line within the yellow bar represents a relative contribution of 29% to $MHT_{OSNAP-W}$ (in Fig. 4.5(e)) and 16% to $MHT_{OSNAP-E}$ (in Fig. 4.5(f)). For each of the two observed OSNAP quantities, this fraction indicates the relative importance of shortwave downward radiation, the field which was discussed in representation of the thermal forcing fields. The link to Figs. 4.5(a) and (b) is the following: The squared entries in the sensitivity maps of Figs. 4.5(a) and (b) add up to 0.29 and 0.16, respectively, if values on a model grid-scale basis, i.e., at a nominal resolution of 1° , were distinguishable in the figures. Values on a grid-scale basis will be visualized in Fig. 4.10. Similarly, the relative importance of meridional wind stress is represented by the chunk right of the dashed vertical line within the green bar, and is 16% for $MHT_{OSNAP-W}$ (Fig. 4.5(e)), and 34% for $MHT_{OSNAP-E}$ (Fig. 4.5(f)). Accordingly, the squared entries in the sensitivity maps of Figs. 4.5(c) and (d) add up to 0.16 and 0.34, respectively.

4.3.2 Dynamical constraints of OSNAP

In this section, the mechanisms and pathways identified in Section 4.3.1 are linked to the dynamical constraints that OSNAP provides for the inverse modeling framework. These dynamical constraints will be further used in Section 4.3.4 to assess the proxy potential of the OSNAP array.

Formally, the dynamical constraints of a set of observations are assessed via uncertainty reduction that is achieved by including the observations in the inverse modeling framework. Here, uncertainty reduction is based purely on dynamical information that is propagated via the linearized model equations. In practical terms, uncertainty reduction is captured by the curvature of the (non-dimensionalized) model-data misfit function

$$J_{\text{misfit}}(\mathbf{u}) = \frac{1}{2} \left[\left(\mathbf{y} - \mathbf{f}(\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{u}) \right)^T \mathbf{\Gamma}_{\text{noise}}^{-1} \left(\mathbf{y} - \mathbf{f}(\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{u}) \right) \right], \quad (4.11)$$

where \mathbf{y} the set of considered observations, \mathbf{f} the corresponding observation operator, $\mathbf{\Gamma}_{\text{noise}}$ the chosen noise covariance matrix, and $\mathbf{u} = \mathbf{\Gamma}_{\text{prior}}^{-1/2} \mathbf{x}$ the non-dimensionalized control vector. As explained in Section 4.2.2, \mathbf{f} and $\mathbf{\Gamma}_{\text{noise}}$ are the two pieces of information about an observing system that are required for assessing the dynamical constraints and proxy potential of the observations. In contrast, the measurement values \mathbf{y} are not needed because the curvature of (4.11) does (to first order) not depend on \mathbf{y} (see Section B.4).

The directions within the control space in which the model-data misfit function (4.11) is curved are the *data-informed directions*, characterizing the components that are most informed by the considered observations and along which uncertainty is reduced. The data-informed directions and associated curvatures are equal to the eigenvectors and eigenvalues $\{(\mathbf{v}_i, \lambda_i)\}$ of the Hessian matrix of (4.11).

In the following, I will study the data-informed directions and associated curvatures that correspond to the OSNAP observing systems, defined in Table 4.1. These data-informed directions and curvatures characterize the dynamical constraints of the OSNAP array.

Individual constraints of OSNAP-West and OSNAP-East

First, the OSNAP-West and the OSNAP-East observing systems are considered separately, assuming that one measures time-mean heat transport only across OSNAP-West *or* across OSNAP-East (Table 4.1). In each of two cases, either the pair $(\mathbf{f}^W, \mathbf{\Gamma}_{\text{noise}}^W)$ or the pair $(\mathbf{f}^E, \mathbf{\Gamma}_{\text{noise}}^E)$ from Table 4.1 takes on the role of $(\mathbf{f}, \mathbf{\Gamma}_{\text{noise}})$ in (4.11). Each of the two examples corresponds to the first case discussed in Section 3.3, in which only a single observation is available. The control space direction informed by the OSNAP-West observing system is denoted by \mathbf{v}_1^W , with associated curvature λ_1^W , and the control space direction informed by the OSNAP-East observing system by \mathbf{v}_1^E , with associated curvature λ_1^E .

The OSNAP-West-informed direction \mathbf{v}_1^W and the OSNAP-East-informed direction \mathbf{v}_1^E are closely linked to the dynamical adjustment mechanisms and pathways identified in Section 4.3.1. Indeed, \mathbf{v}_1^W and \mathbf{v}_1^E are given by the normalized prior-weighted linearized sensitivities of $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$, respectively, to the full control vector. While the two vectors \mathbf{v}_1^W and \mathbf{v}_1^E are directions within the high-dimensional control space (see third column of Table 4.2), Figs. 4.5(a),(c) (left column) show the restriction of \mathbf{v}_1^W to *parts* of the control variables, namely to the two-dimensional fields of shortwave downward radiation and meridional wind stress. Similarly, Figs. 4.5(b),(d) (right column) show the restriction of \mathbf{v}_1^E to the same subset of control variables.

As explained in Section 3.3, the eigenvalues (or curvatures) λ_1^W and λ_1^E corresponding to the data-informed directions \mathbf{v}_1^W and \mathbf{v}_1^E reflect the prior-to-noise ratios of time-mean heat transport across OSNAP-West and OSNAP-East, respectively. They are computed as

$$\lambda_1^W = \frac{(\sigma_{\text{prior}}^W)^2}{\mathbf{\Gamma}_{\text{noise}}^W} \quad \text{and} \quad \lambda_1^E = \frac{(\sigma_{\text{prior}}^E)^2}{\mathbf{\Gamma}_{\text{noise}}^E}. \quad (4.12)$$

The numerators $(\sigma_{\text{prior}}^W)^2$ and $(\sigma_{\text{prior}}^E)^2$ are the prior uncertainties for $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$, computed via (4.7), i.e., by propagating the prior information $\mathbf{\Gamma}_{\text{prior}}$ from the control variables to the respective model quantities $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$. In other words, $(\sigma_{\text{prior}}^W)^2$ ($(\sigma_{\text{prior}}^E)^2$) reflects the quality of prior information on the model quantity $\text{MHT}_{\text{OSNAP-W}}$ ($\text{MHT}_{\text{OSNAP-E}}$), and determines the *relevance* of the actual OSNAP-West (OSNAP-East) measurement values. On the other hand, the denominators $\mathbf{\Gamma}_{\text{noise}}^W$ and $\mathbf{\Gamma}_{\text{noise}}^E$ are the assumed noise variances for observed time-mean OSNAP-West and OSNAP-East heat transports, respectively. In other words, $\mathbf{\Gamma}_{\text{noise}}^W$ ($\mathbf{\Gamma}_{\text{noise}}^E$) reflects the

quality of the OSNAP-West (OSNAP-East) data. Therefore, the prior-to-noise ratio λ_1^W (λ_1^E) compares the relevance to the quality of OSNAP-West (OSNAP-East) data. Using the values for the computed prior uncertainties and assumed noise variances, shown in Tables 4.3 and 4.1, the prior-to-noise ratios in (4.12) evaluate to

$$\lambda_1^W \approx (0.60)^2 \quad \text{and} \quad \lambda_1^E \approx (0.23)^2. \quad (4.13)$$

The fact that the eigenvalue (or curvature) λ_1^W is higher than λ_1^E means that OSNAP-West provides relatively more new reliable information for the inverse modeling framework than OSNAP-East. In fact, the ratio of the two eigenvalues is

$$\frac{\lambda_1^W}{\lambda_1^E} \approx (2.5)^2. \quad (4.14)$$

The ratio (4.14) can be explained by the combination of two facts. First, the ratio of $\sqrt{\Gamma_{\text{noise}}^W}$ and $\sqrt{\Gamma_{\text{noise}}^E}$ is equal to 1/5 (see Table 4.1 or equation (4.1)), i.e., the assumed noise standard deviation for time-mean heat transport across OSNAP-East is assumed to be 5 times higher than for time-mean heat transport across OSNAP-West. Second, the prior standard deviation σ_{prior}^E in $\text{MHT}_{\text{OSNAP-W}}$ is about twice as large as σ_{prior}^W (see Table 4.3), i.e., OSNAP-East data is about twice as “relevant” as OSNAP-West data. However, OSNAP-East data being more “relevant” by (only) a factor of 2 cannot make up for OSNAP-East data being more noisy by a factor of 5. This yields the ratio (4.14), which reflects a stronger constraint by OSNAP-West on the inverse modeling framework.

Note that the values for “relevance”, i.e., $(\sigma_{\text{prior}}^W)^2$ and $(\sigma_{\text{prior}}^E)^2$, are computed by propagating the prior uncertainties Γ_{prior} from the control variables onto $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$, respectively (via equation 4.7). Since the same prior uncertainties Γ_{prior} are used for both propagations, the contrasting values for $(\sigma_{\text{prior}}^W)^2$ and $(\sigma_{\text{prior}}^E)^2$ are purely set by the differences in the linearized sensitivities of $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$ to the control variables (cf. equation 4.7). Roughly speaking, $\text{MHT}_{\text{OSNAP-E}}$ is “on average” twice as sensitive to changes in the uncertain control variables as $\text{MHT}_{\text{OSNAP-W}}$ (where the average can be seen as a prior-weighted average in the l^2 norm). OSNAP-East *sensitivity* being larger than OSNAP-West sensitivity by a factor of (only) 2, stands in contrast to the OSNAP-East *signal* being larger than the OSNAP-West signal by about a factor of 5. The factor-5-relation in the time-mean signals of OSNAP-East vs. OSNAP-West can be detected in the OSNAP measurement values from August 2014 to April 2016 (see equation 4.1), but also the ECCOv4r2 state estimate from 1992 to 2011 (Fig. 4.2). Therefore, if one were to follow the statistical (rather than the dynamical) approach, and substituted the prior-to-noise ratios in (4.12) by the signal-to-noise ratios, one would obtain a similar value for OSNAP-West as for OSNAP-East in (4.13). For the statistical approach, the resulting ratio in (4.14) would therefore be approximately 1, suggesting similarly strong constraints by OSNAP-West and OSNAP-East. In contrast, the dynamical approach suggests the ratio of $(2.5)^2$ in (4.14), obtained by using the model dynamics and ocean connectivity in order to propagate Γ_{prior} from all uncertain model input variables (initial conditions, model parameters, and atmospheric forcing; see Table 4.2) throughout the global ocean. Here, Γ_{prior} is adapted from ECCOv4r3, as described in Section 4.2.2, and uses for example information from reanalysis data-sets (and their uncertainties) for the atmospheric forcing fields.

Choosing different values of $r^2 > 0$ and $b^2 > 0$ in (4.4) and (4.8), respectively, reflects the assumptions of less ($r^2 < 1$) or more ($r^2 > 1$) noisy OSNAP observations and/or more ($b^2 < 1$) or less ($b^2 > 1$) certain prior information. Varying $\mathbf{\Gamma}_{\text{noise}}^W[r_W^2] = r_W^2 \cdot \mathbf{\Gamma}_{\text{noise}}^W$, $\mathbf{\Gamma}_{\text{noise}}^E[r_E^2] = r_E^2 \cdot \mathbf{\Gamma}_{\text{noise}}^E$, and $\mathbf{\Gamma}_{\text{prior}}[b^2] = b^2 \cdot \mathbf{\Gamma}_{\text{prior}}$ in that manner has the following effect on the individual dynamical constraints of OSNAP-West and OSNAP-East on the inverse modeling framework (see Section 3.5.3):

- (i)-ind The eigenvectors \mathbf{v}_1^W and \mathbf{v}_1^E remain the same, as visualized in Fig. 4.5, i.e., they are independent of the choices of $r_W^2, r_E^2, b^2 > 0$.
- (ii)-ind The eigenvalues λ_1^W and λ_1^E change according to $\lambda_1^W[r_W^2, b^2] = (b/r_W)^2 \cdot \lambda_1^W[1, 1]$, and $\lambda_1^E[r_E^2, b^2] = (b/r_E)^2 \cdot \lambda_1^E[1, 1]$, where $\lambda_1^W[1, 1]$ and $\lambda_1^E[1, 1]$ are equal to the values of λ_1^W and λ_1^E in (4.13).
- (iii)-ind If r_W^2 and r_E^2 coincide, taking the value $r^2 > 0$, the ratio of the OSNAP-West vs. the OSNAP-East prior-to-noise (or relevance-to-quality) ratios $\lambda_1^W[r^2, b^2]$ and $\lambda_1^E[r^2, b^2]$ remains the same as in (4.14), independently of the choice of $r^2, b^2 > 0$.

Constraints by combining information from OSNAP-West and OSNAP-East

Next, OSNAP-West and OSNAP-East are combined to a joint observing system, the OSNAP-Combined observing system (Table 4.1). That is, it is assumed that time-mean heat transport measurements across both OSNAP-West *and* OSNAP-East are available and the pair $(\mathbf{f}, \mathbf{\Gamma}_{\text{noise}})$ as specified in Table 4.1 is used in the model-data misfit function J_{misfit} (equation (4.11)). The dynamical constraints of the OSNAP-Combined observing system are characterized by the two eigenvectors \mathbf{v}_1 and \mathbf{v}_2 of the Hessian of J_{misfit} , with associated positive curvatures λ_1 and λ_2 .

\mathbf{v}_1 and \mathbf{v}_2 span the two-dimensional subspace within the high-dimensional control space (defined in Table 4.2) that is informed by the OSNAP-Combined observing system. The displayed plane in Figure 4.6 represents this OSNAP-informed two-dimensional subspace, while all directions orthogonal to this plane are not informed by the data. The ellipses in Figure 4.6 sketch the orientation of the contour lines of J_{misfit} . The dashed black line shows the one-dimensional subspace spanned by \mathbf{v}_1 , and indicates the direction of maximal curvature of J_{misfit} , i.e., the direction along which uncertainty is reduced most efficiently. The maximal curvature is given by the eigenvalue λ_1 . The dashed gray line indicates the subspace spanned by \mathbf{v}_2 , orthogonal to the black line. The curvature in the direction of \mathbf{v}_2 is given by the eigenvalue λ_2 . The curvature in all directions orthogonal to the shown plane is zero. The black and gray vectors, associated to the black and gray lines, are scaled by the inverse of $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$, respectively. Thus, the fact that the black vector is shorter than the gray vector reflects J_{misfit} being more curved—and uncertainty being reduced more efficiently—along \mathbf{v}_1 than along \mathbf{v}_2 .

\mathbf{v}_1 and \mathbf{v}_2 are closely related to \mathbf{v}_1^W and \mathbf{v}_1^E , the directions informed by the individual OSNAP-West and OSNAP-East observing systems. The pair $(\mathbf{v}_1^W, \mathbf{v}_1^E)$ spans the same two-dimensional subspace within the control space as $(\mathbf{v}_1, \mathbf{v}_2)$, and can therefore be displayed in the plane of Fig. 4.6. The one-dimensional subspaces spanned by \mathbf{v}_1^W and \mathbf{v}_1^E

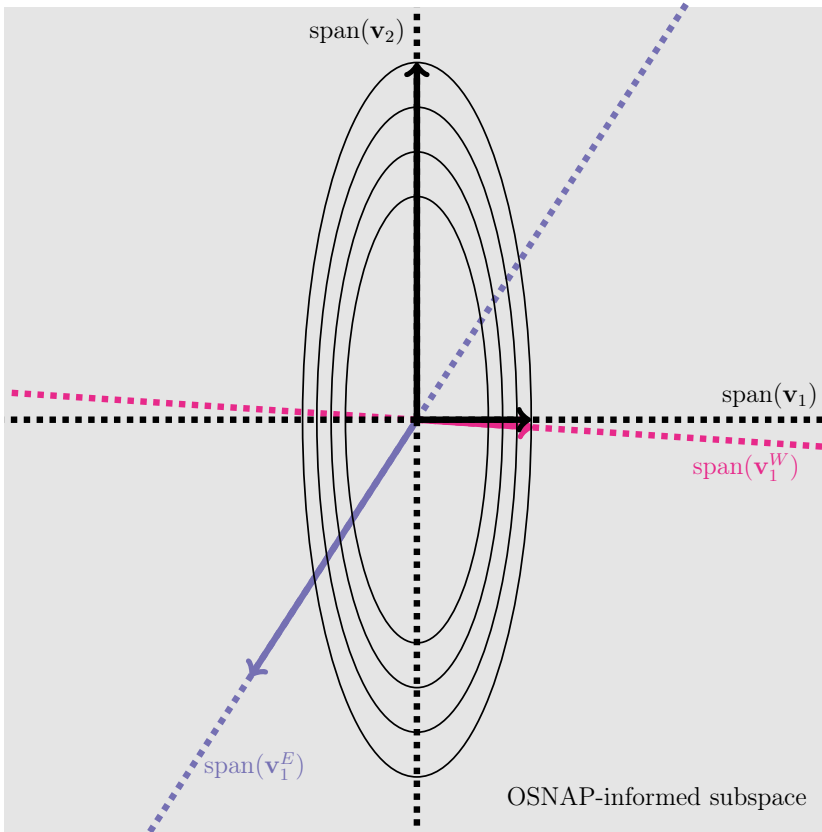


Figure 4.6: The two-dimensional subspace informed by heat transport measurements across OSNAP-West and OSNAP-East. All directions orthogonal to the displayed gray plane are not informed by the data. The ellipses show the orientation of the contour lines of J_{misfit} , the data-misfit part of the cost function that is associated with the observation operator $\mathbf{f} = (\text{MHT}_{\text{OSNAP-W}}, \text{MHT}_{\text{OSNAP-E}})^T$ and noise covariance $\mathbf{\Gamma}_{\text{noise}}$, as defined in Table 4.1. The dashed horizontal black line shows the one-dimensional subspace spanned by \mathbf{v}_1 , and indicates the direction of maximal curvature of J_{misfit} . The dashed vertical black line indicates the subspace spanned by \mathbf{v}_2 (orthogonal to the dashed horizontal black line). The associated black vectors are scaled by the inverse of $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$, respectively. The dashed pink and purple lines show the one-dimensional subspaces informed by the individual OSNAP-West and OSNAP-East observations, spanned by \mathbf{v}_1^W and \mathbf{v}_1^E , respectively. The associated pink and purple vectors are scaled by the inverse of $\sqrt{\lambda_1^W}$ and $\sqrt{\lambda_1^E}$, respectively.

Dynamical constraints of OSNAP-Combined

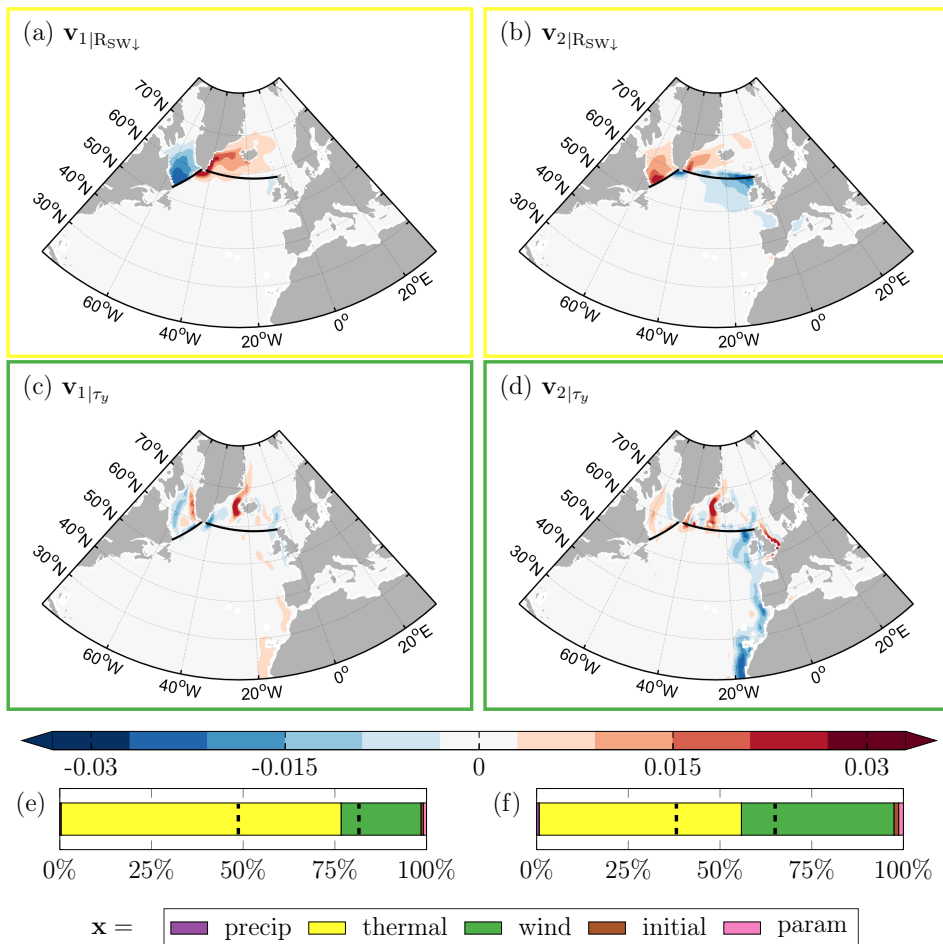


Figure 4.7: (a),(c) and (b),(d) show clipped versions of the vectors \mathbf{v}_1 and \mathbf{v}_2 , respectively. \mathbf{v}_1 and \mathbf{v}_2 characterize the dynamical constraints of the OSNAP-Combined observing system (cf. Table 4.1); they are the eigenvectors of the Hessian of the model-data misfit function (4.11). The maps show the restrictions of the respective vectors to (a),(b) shortwave downward radiation $R_{SW\downarrow}$ and (c),(d) meridional wind stress τ_y . The vectors \mathbf{v}_1 and \mathbf{v}_2 span the same data-informed subspace within the control space as \mathbf{v}_1^W and \mathbf{v}_1^E in Fig. 4.5, but sensitivity patterns are “re-distributed” across the eigenvectors, by a procedure that takes into account data redundancy and prior-to-noise ratios. The solid black contours in (a)-(d) delineate the cross-sectional OSNAP-West and OSNAP-East transects. The bar charts show the relative importance (see definition 4.18) of the control variables for the vectors (e) \mathbf{v}_1 and (f) \mathbf{v}_2 . The control variables are grouped according to Table 4.2. The fractions right of the vertical dashed lines in (e) and (f) indicate relative importance of shortwave downward radiation and meridional wind stress, similarly as in Fig. 4.5.

are shown by the dashed pink and purple lines in Fig. 4.6, respectively. The pink and purple lines are exactly those directions in which J_{misfit}^W and J_{misfit}^E are curved, where J_{misfit}^W and J_{misfit}^E are the model-data misfit functions (4.11) associated with the OSNAP-West and OSNAP-East observing systems, respectively. The associated pink and purple arrows are scaled by the inverse of $\sqrt{\lambda_1^W}$ and $\sqrt{\lambda_1^E}$, respectively. The pink vector is about 2.5 times shorter than the purple vector, reflecting identity (4.14). Thus, the shorter pink vector indicates that J_{misfit}^W has higher curvature than J_{misfit}^E , corresponding to a more effective uncertainty reduction, along the respective informed direction.

In Section 3.3 it was explained how to obtain $(\mathbf{v}_1, \lambda_1)$ and $(\mathbf{v}_2, \lambda_2)$ from $(\mathbf{v}_1^W, \lambda_1^W)$ and $(\mathbf{v}_1^E, \lambda_1^E)$. First, \mathbf{v}_1^W and \mathbf{v}_1^E are orthonormalized; this corresponds to removing redundant information. Second, the obtained vector pair is rotated and reflected within the informed subspace in such a way that \mathbf{v}_1 points into the direction of maximal curvature of the cost function. Due to the high prior-to-noise ratio λ_1^W of OSNAP-West, compared to the prior-to-noise ratio λ_1^E of OSNAP-East, \mathbf{v}_1 deviates only very little from \mathbf{v}_1^W . The exact representations of \mathbf{v}_1 and \mathbf{v}_2 as linear combinations of \mathbf{v}_1^W and \mathbf{v}_1^E , computed by steps (i) and (ii) in Section 3.3, are

$$\mathbf{v}_1 = 0.96 \cdot \mathbf{v}_1^W - 0.07 \cdot \mathbf{v}_1^E, \quad (4.15)$$

$$\mathbf{v}_2 = -0.72 \cdot \mathbf{v}_1^W - 1.2 \cdot \mathbf{v}_1^E. \quad (4.16)$$

The eigenvalues compute as

$$\lambda_1 = 0.362 \quad \text{and} \quad \lambda_2 = 0.037. \quad (4.17)$$

As already seen for the individual constraints of OSNAP-West and OSNAP-East, the assumptions on observational noise and prior information will have an impact on the inferred dynamical constraints of OSNAP-Combined. Assuming less ($r^2 < 1$) or more ($r^2 > 1$) noisy OSNAP observations via the choice of $\mathbf{\Gamma}_{\text{noise}}[r^2] = r^2 \cdot \mathbf{\Gamma}_{\text{noise}}$, and/or more ($b^2 < 1$) or less ($b^2 > 1$) certain prior information via choosing $\mathbf{\Gamma}_{\text{prior}}[b^2] = b^2 \cdot \mathbf{\Gamma}_{\text{prior}}$, has the following effect on the dynamical constraints of OSNAP-Combined on the inverse modeling framework (see Section 3.5.3 and cf. (i)-ind-(iii)-ind):

- (i)-comb The eigenvectors \mathbf{v}_1 and \mathbf{v}_2 remain the same, as visualized in Fig. 4.7, i.e., they are independent of the choices of $r^2, b^2 > 0$.
- (ii)-comb The eigenvalues λ_1 and λ_2 change according to $\lambda_1[r^2, b^2] = (b/r)^2 \cdot \lambda_1[1, 1]$, and $\lambda_2[r^2, b^2] = (b/r)^2 \cdot \lambda_2[1, 1]$, where $\lambda_1[1, 1]$ and $\lambda_2[1, 1]$ are equal to the values of λ_1 and λ_2 in (4.17).

Since the directions \mathbf{v}_1 and \mathbf{v}_2 are closely related to \mathbf{v}_1^W and \mathbf{v}_1^E (via equations (4.15) and (4.16)), \mathbf{v}_1 and \mathbf{v}_2 can be presented as linear combinations of the prior-weighted linearized sensitivities shown in Fig. 4.5. This is done in Fig. 4.7, where Figs. 4.7(a),(c) (left column) show a clipped version of \mathbf{v}_1 , and Figs. 4.7(b),(d) (right column) a clipped version of \mathbf{v}_2 . For both vectors, restrictions to the fields of shortwave downward radiation (Figs. 4.7(a),(b)) and meridional wind stress (Figs. 4.7(c),(d)) are shown. The left columns of Fig. 4.7 and Fig. 4.5, i.e., subfigures (a),(c), are almost identical because \mathbf{v}_1 almost coincides with \mathbf{v}_1^W , as seen in equation (4.15) and Fig. 4.6. On the other hand, \mathbf{v}_2

is a linear combination of \mathbf{v}_1^W and \mathbf{v}_1^E , with significant contributions from both of these vectors, visible from equation (4.16) and Fig. 4.6. This is reflected by the fact that the sensitivity map in Fig. 4.7(b) has features from both Figs. 4.5(a) and (b) - with flipped signs, due to the negative multipliers of \mathbf{v}_1^W and \mathbf{v}_1^E in equation (4.16). For instance, positive sensitivities in the Labrador Sea (Fig. 4.7(b)) are due to heat advection across OSNAP-West, and negative sensitivities west of Great Britain (Fig. 4.7(b)) due to heat advection across OSNAP-East; these features are inherited from Figs. 4.5(a) and (b), respectively, with flipped signs.

The region close to the Southeast Greenlandic coast is a location where sensitivities of $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$ overlap. For instance, $\text{R}_{\text{SW} \downarrow}$ sensitivities along the Southeast Greenlandic coast were found to be positive for heat transport across OSNAP-West (Fig. 4.5(a)), and negative for heat transport across OSNAP-East (Fig. 4.5(b)). Similarly, τ_y sensitivities that populate the Denmark Strait are of opposite sign for heat transport across OSNAP-West (positive, Fig. 4.5(c)) and heat transport across OSNAP-East (negative, Fig. 4.5(d)). The opposite signs lead to partial cancellations in these regions when computing \mathbf{v}_2 via equation (4.16). As a result, $\text{R}_{\text{SW} \downarrow}$ sensitivities along the Southeast Greenlandic coast are of weaker (absolute) amplitude in Fig. 4.7(b) than in Figs. 4.5(a) and (b), and τ_y sensitivities along the Denmark Strait of weaker (absolute) amplitude in Fig. 4.7(d) than in Figs. 4.5(c) and (d).

Figs. 4.7(a)-(d) show that the process of combining information from OSNAP-West and OSNAP-East is reflected in combining sensitivity patterns of OSNAP-West and OSNAP-East. Sensitivity patterns are “re-distributed” across the eigenvectors, by a procedure that takes into account data redundancy and prior-to-noise ratios. This insight generalizes to observing systems with many more observations than just two (as considered here): All patterns that can be detected in the eigenvectors, –the data-informed directions within the control space–, originate from sensitivity maps associated to the individual observations that are part of the observing system.

The relative importance of the control variables for the eigenvectors are shown in Figs. 4.7(e) and (f). For each group, the relative importance is measured by

$$\sum_{i|x_i \in \text{group}} v_i^2, \quad (4.18)$$

where (v_1, \dots, v_N) denote the vector components of \mathbf{v}_1 (in Fig. 4.7(e)) and \mathbf{v}_2 (in Fig. 4.7(f)), and (x_1, \dots, x_N) are the associated control variables. Note that the definition in (4.18), used for the bar charts in Figs. 4.7(e) and (f), corresponds one-to-one to the definition in (4.10), used for the bar charts in Figs. 4.5(e) and (f), since the eigenvectors \mathbf{v}_1 and \mathbf{v}_2 are normalized, i.e., $\sum_{i=1}^N v_i^2 = 1$. The dominance of the groups *thermal* and *wind* in Figs. 4.7(e) and (f) is inherited from Figs. 4.5(e) and (f), i.e., from the fact that the studied observations, heat transport across OSNAP-West and OSNAP-East, was found to be most sensitive to these two groups.

Dynamical adjustment mechanisms for the unobserved QoIs

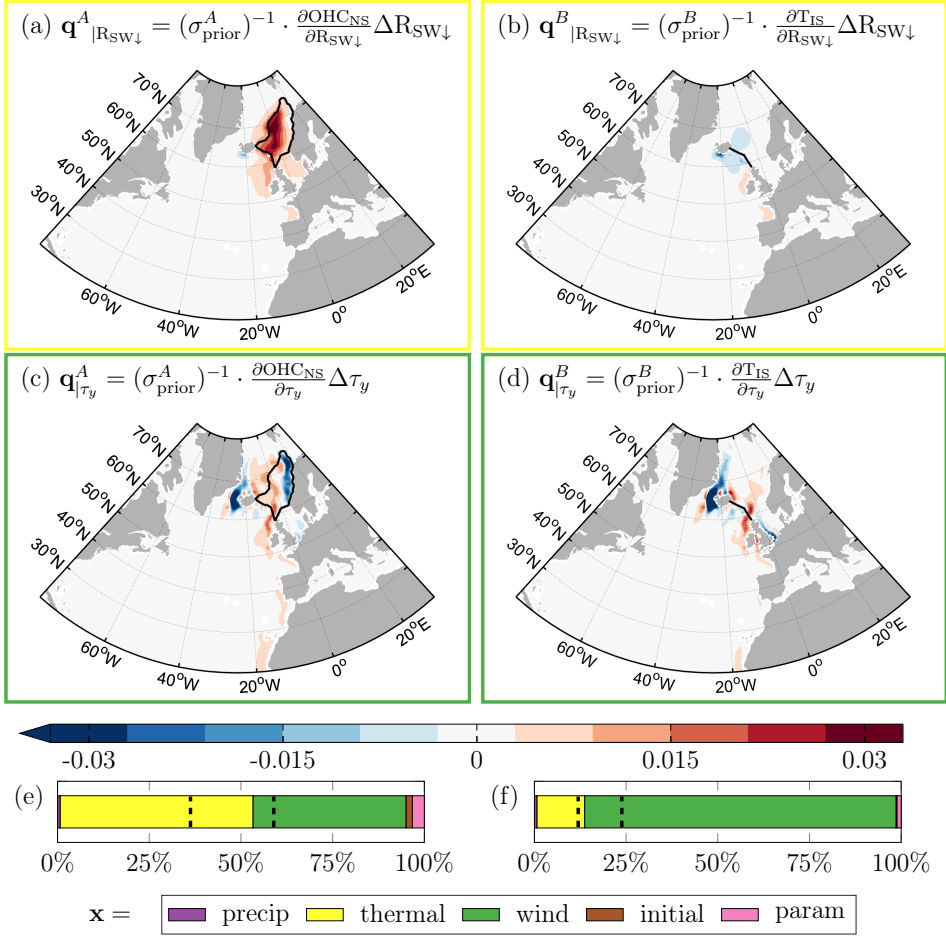


Figure 4.8: Linearized sensitivities of the unobserved quantities of interests (QoIs) time-mean (a),(c) Nordic Seas heat content (OHC_{NS}) and (b),(d) volume transport across the Iceland-Scotland Ridge (T_{IS}) to time-mean (a),(b) shortwave downward radiation $R_{SW\downarrow}$ and (c),(d) meridional wind stress τ_y . The sensitivities are prior-weighted and normalized. Red (blue) colors indicate that an increase in (a),(b) downward radiation and (c),(d) northward wind stress would lead to a subsequent increase (decrease) in (a),(c) OHC_{NS} and (b),(d) T_{IS} on a 5 year timescale. The solid black contours delineate the (a),(c) horizontal area and (b),(d) cross-sectional transect for the respective calculations of (a),(c) OHC_{NS} and (b),(d) T_{IS} . (a),(c) is a clipped version of the vector \mathbf{q}^A , the OHC_{NS} target direction that needs to be informed by the observations. Similarly, (b),(d) is a clipped version of the vector \mathbf{q}^B , the T_{IS} target direction. The bar charts show the relative importance (see definition 4.19) of the control variables for impacting (e) OHC_{NS} and (f) T_{IS} . The control variables are grouped according to Table 4.2. The fractions right of the vertical dashed lines in (e) and (f) indicate relative importance of shortwave downward radiation and meridional wind stress, similarly as in Fig. 4.5.

4.3.3 Dynamical adjustment mechanisms for the unobserved QoIs

This section identifies the dynamical adjustment mechanisms and pathways for the two chosen unobserved quantities of interest (QoIs): time-mean Nordic Seas heat content (OHC_{NS}) and time-mean volume transport across the Iceland-Scotland ridge (T_{IS}).

To this aim, I employ the linearized sensitivities $\frac{\partial \text{OHC}_{\text{NS}}}{\partial \mathbf{x}}$ and $\frac{\partial \text{T}_{\text{IS}}}{\partial \mathbf{x}}$, which are obtained by the adjoint models described in Section 4.2.5. Fig. 4.8 shows the normalized prior-weighted sensitivities of OHC_{NS} (in (a),(c)) and T_{IS} (in (b),(d)) to changes in the time-mean of R_{SW↓} (in (a),(b)) and τ_y (in (c),(d)). Note that Fig. 4.8 is the analogue of Fig. 4.5, but now, prior-weighted sensitivities of the unobserved QoIs, rather than the observed OSNAP quantities are shown. Prior-weighting in Fig. 4.8 is done through multiplication by the prior ECCOV4r3 standard deviations fields $\Delta R_{\text{SW}\downarrow}$ and $\Delta \tau_y$ from Figs. 4.4, exactly as in Fig. 4.5. The normalizing is achieved by dividing the prior-weighted sensitivities of OHC_{NS} by the constant value of σ_{prior}^A , and the prior-weighted sensitivities T_{IS} by the constant value of σ_{prior}^B . σ_{prior}^A and σ_{prior}^B are the prior uncertainties in OHC_{NS} and T_{IS}, respectively, and their (squared) values are reported in Table 4.3. These prior uncertainties are computed by accumulating prior-weighted sensitivity to *all* control variables, as shown in equation (4.7), where QoI in (4.7) is chosen as OHC_{NS} or T_{IS}, respectively.

The sensitivity maps in Fig. 4.8 reveal all possible dynamical mechanisms via which small-amplitude perturbations in the time-mean shortwave downward radiation and meridional wind stress can impact the time-mean Nordic Seas heat content and volume transport across the Iceland-Scotland ridge on a five-year timescale. The many different features that compose the maps in Fig. 4.8 reflect the fact that the QoIs are the superposition of many signals that are propagated to the Nordic Seas or the Iceland-Scotland ridge by different adjustment mechanisms. All these different signals are to be informed by the remote observations. The sensitivity maps in Fig. 4.8 disentangle the different adjustment mechanisms and pathways for the QoIs, which will be further explored in the following.

Sensitivity to thermal forcing

Figs. 4.8(a) and (b) are proportional (up to normalizing constants) to the prior-weighted linearized sensitivities of 5-year-mean Nordic Seas heat content (Fig. 4.8(a)) and of volume transport across the Iceland-Scotland ridge (Fig. 4.8(b)) to changes in the 5-year mean of shortwave downward radiation R_{SW↓}. Similarly as in Fig. 4.5, the pattern in Figs. 4.8(a) and (b) can be considered representative of prior-weighted sensitivities to all forcing fields that belong to the group *thermal*. This is indicated by the yellow figure frames.

There is a close link between Figs. 4.8(a) and (b) and the sensitivities to surface heat flux in Chapter 2. While I investigated non-weighted March sensitivities to heat flux perturbations at different lead times in Section 2.3, sensitivities in Figs. 4.8(a) and (b) are (i) prior-weighted, i.e., multiplied by prior uncertainties $\Delta R_{\text{SW}\downarrow}$, and (ii) *accumulated* over 5 years, corresponding to time-mean adjustments of the control variable shortwave

downward radiation (cf. Table 4.2). When accumulating sensitivities over time, cancellations occur at the locations where sensitivities have opposite signs for different lead times. The patterns visible in Figs. 4.8(a) and (b) are the ones from Section 2.3 that carry through the accumulation process, e.g., patterns from Section 2.3 that are either (i) of strongest amplitude, typically found for short lead times, or (ii) stationary on a five-year timescale. Another difference to the sensitivities from Section 2.3 is that the sign of the sensitivities in Figs. 4.8(a) and (b) is reversed because sensitivities are computed to *downward* instead of upward surface fluxes. Red (blue) colors in Figs. 4.8(a) and (b) indicate the locations where an increase (decrease) in downward radiation would lead to a subsequent increase in Nordic Seas heat content (Fig. 4.8(a)) and northward volume transport across the Iceland-Scotland ridge (Fig. 4.8(b)) on a five-year timescale.

In Fig. 4.8(a), strong positive sensitivities populate the Nordic Seas domain. This pattern is due to locally driven, one-dimensional air-sea processes, for which no active ocean dynamics are required. Increased shortwave downward radiation leads to higher local ocean temperature throughout the mixed layer, and thus increases Nordic Seas ocean heat content (OHC_{NS}). Weaker positive sensitivities are found in the upstream regions of the Nordic Seas domain, south of the Iceland-Scotland ridge. These sensitivities are due to advection of temperature anomalies across the ridge into the Nordic Seas. The sensitivity distribution in Fig. 4.8(a) is consistent with the results from Chapter 2, where local air-sea heat fluxes were found to be a very efficient mechanism for altering Nordic Seas heat content. This mechanism was identified as one that acts on short timescales; the response of ocean heat content is strongest within the same year of the local heat flux perturbation.

Volume transport across the Iceland-Scotland ridge (Fig. 4.8(b)) is much less sensitive to shortwave downward radiation than Nordic Seas heat content (Fig. 4.8(a)). The faint sensitivities visible in Fig. 4.8(b) are due to the dynamics of thermal wind balance, a mechanism that was discussed both in Chapter 2 and for OSNAP heat transports in Section 4.3.1. $R_{\text{SW} \downarrow}$ anomalies matching the sensitivity distribution in Fig. 4.8(b), i.e., anomalies of negative sign around Iceland, and of positive sign along the western European coast, generate a positive perturbation in the zonal upper-ocean density gradient. By thermal wind balance, this perturbation in the zonal density gradient increases the northward geostrophic transport across the Iceland-Scotland ridge (T_{IS}).

Sensitivity to wind forcing

Figs. 4.8(c) and (d) are proportional to the prior-weighted linearized sensitivities of OHC_{NS} (in (c)) and T_{IS} (in (d)) to meridional wind stress τ_y . The green frames highlight that the figures represent sensitivities to wind forcing, i.e., control variables classified under the group *wind*. As for shortwave downward radiation, the time-mean sensitivities in Figs. 4.8(c) and (d) are closely linked to the sensitivities to meridional wind stress presented for different lead times in Section 2.3.

Fig. 4.8(c) and (d) show a band of strong negative sensitivity along the western coast of Iceland, where high sensitivities to meridional wind stress are amplified by

high prior ECCOv4r3 meridional wind stress uncertainties $\Delta\tau_y$ (see Fig. 4.4(b)). This leads to strongly negative values that fall out of the shown colorbar range: normalized prior-weighted sensitivities to meridional wind stress along the western coast of Iceland take values down to -0.13 in Fig. 4.8(c) and down to -0.22 in Fig. 4.8(d). The band of negative sensitivities along the western coast of Iceland is a key pattern that was already discovered both in Section 2.3 and in Fig. 4.5(d). Kelvin waves transport negative pressure anomalies, induced by Ekman offshore divergence, to the western end of the section between Iceland and Scotland, and geostrophic balance strengthens the northward geostrophic transport across the Iceland-Scotland ridge. This increases both quantities T_{IS} and OHC_{NS} .

Moreover, both Figs. 4.8(c) and (d) show a band of positive sensitivities along the western European coast, which extends northward across the ridge into the Nordic Seas. This pattern was also found in Fig. 4.5(d), and is again due to geostrophic balance. Positive pressure anomalies, induced by Ekman onshore convergence are transported by northward traveling Kelvin waves, and strengthen the northward geostrophic transport across the ridge. This leads to an increase of T_{IS} and OHC_{NS} .

Fig. 4.8(c) shows furthermore negative sensitivities along the Norwegian coast, a pattern that is due to local processes in the Nordic Seas and therefore not visible in Fig. 4.8(d). The negative sensitivities along the Norwegian coast are part of a sensitivity dipole, where the positive counterpart is found further offshore in the Nordic Seas. A wind stress anomaly which matches the sensitivity distribution of the dipole induces Ekman downwelling. As a result, warm surface waters are pumped to deeper ocean depths, which increases OHC_{NS} .

Relative importance of control variables

The relative importance of the control variables for the unobserved quantities OHC_{NS} and T_{IS} are shown by the bar charts in Figs. 4.8(e) and (f). Here, the control variables are again grouped according to Table 4.2. Relative importance is measured similarly as in (4.10) for Figs. 4.5(e) and (f): here, the bar charts show the ratio

$$\frac{\sum_{i|x_i \in \text{group}} \left(\frac{\partial OHC_{NS}}{\partial x_i} \cdot \Delta x_i \right)^2}{\sum_{i=1}^N \left(\frac{\partial OHC_{NS}}{\partial x_i} \cdot \Delta x_i \right)^2} \quad \text{and} \quad \frac{\sum_{i|x_i \in \text{group}} \left(\frac{\partial T_{IS}}{\partial x_i} \cdot \Delta x_i \right)^2}{\sum_{i=1}^N \left(\frac{\partial T_{IS}}{\partial x_i} \cdot \Delta x_i \right)^2} \quad (4.19)$$

in Fig. 4.8(e) and Fig. 4.8(f), respectively. As before, the denominators in (4.19) are the prior uncertainties $(\sigma_{\text{prior}}^A)^2$ and $(\sigma_{\text{prior}}^B)^2$, computed by accumulating prior-weighted sensitivity to *all* control variables (cf. equation (4.7)). Meanwhile, the enumerators in (4.19) are the prior uncertainties due to the control variables in the specific group, computed by accumulating prior-weighted sensitivity only to the control variables in that group.

The bar charts for the QoIs are useful to obtain a first impression of which of the control variables are to be informed by (remote) observations. For example, in Fig. 4.8(e), the large yellow and green fractions indicate that, for informing Nordic Seas heat content, it is more important to inform control variables belonging to the group *thermal* and *wind*

than control variables belonging to the groups *precip*, *initial*, and *param*, which OHC_{NS} is almost insensitive to.

The most important group for OHC_{NS} is the *thermal* forcing fields (Fig. 4.8(e)), with a relative contribution of 53% to prior uncertainty in OHC_{NS}. This is consistent with the strong local sensitivities in Fig. 4.8(a), due to the importance of local air-sea heat fluxes. *Wind* is the second most influential group for OHC_{NS}, with a relative contribution of 42%. In contrast, the group *thermal* has a 13% share only of prior uncertainty in T_{IS} (Fig. 4.8(f)), consistent with the weak sensitivities in Fig. 4.8(b). *Wind*, on the other hand, has a large influence, as to be expected: its relative contribution to prior uncertainty in T_{IS} is 85%. The contributions by the groups *precip*, *initial*, and *param* are negligible for both OHC_{NS} and T_{IS}, which may be explained exactly as for Fig. 4.5: investigating a five-year timescale lets *precip* play a minor role, and the small impact of the groups *initial* and *param* might be due to (too) small specified ECCOv4r3 prior uncertainties for initial conditions and model parameters.

Target directions in inverse modeling framework

Section 4.3.2 investigated the dynamical constraints of OSNAP on all model control variables \mathbf{x} , consisting of initial conditions, boundary forcings, and model parameters. Formally, this investigation was done by studying curvature of the model-data misfit function (4.11) along all directions within the control space. The directions along which the model-data misfit function is curved are the ones along which OSNAP reduces uncertainty. In the following, the dynamical constraints of OSNAP on the two chosen QoIs, rather than on all model control variables shall be studied. In the inverse modeling framework, focusing on a certain QoI, rather than on all control variables, means that one studies the curvature of (4.11), or, equivalently, uncertainty reduction, along *one* target direction, associated with the QoI, rather than along *all* directions within the control space.

The target direction associated with QoI^A = OHC_{NS} is denoted by \mathbf{q}^A , and the target direction associated with QoI^B = T_{IS} by \mathbf{q}^B . The target directions \mathbf{q}^A and \mathbf{q}^B are closely linked to the dynamical adjustment mechanisms and pathways for the QoIs, identified in the previous paragraphs. Indeed, \mathbf{q}^A and \mathbf{q}^B are given by the normalized prior-weighted linearized sensitivities of OHC_{NS} and T_{IS}, respectively, to the full control vector. While the vectors \mathbf{q}^A and \mathbf{q}^B are directions within the high-dimensional control space (see third column of Table 4.2), Figs. 4.8(a),(c) (left column) show the restriction of \mathbf{q}^A to *parts* of the control variables, namely to the two-dimensional fields of short-wave downward radiation and meridional wind stress. Similarly, Figs. 4.8(b),(d) (right column) show the restriction of \mathbf{q}^B to the same subset of control variables.

Choosing different values for $b^2 > 0$ in $\mathbf{\Gamma}_{\text{prior}}[b^2] = b^2 \cdot \mathbf{\Gamma}_{\text{prior}}$ reflects the assumption of more ($b^2 < 1$) or less ($b^2 > 1$) certain prior information. This has the following effect on the target directions:

- (i)-**qoi** Due to normalization, the target directions \mathbf{q}^A and \mathbf{q}^B are independent of $b^2 > 0$, equal to the vectors shown in Fig. 4.8. (Obviously, \mathbf{q}^A and \mathbf{q}^B

Observing System	QoI	OHC _{NS}	T _{IS}
OSNAP-West		3%	17%
OSNAP-East		13%	39%
OSNAP-Combined		13%	40%

Table 4.4: Hypothetical proxy potential of the OSNAP observing systems for the QoIs Nordic Seas heat content (OHC_{NS}, see (4.5)) and volume transport across the Iceland-Scotland ridge (T_{IS}, see (4.6)). The OSNAP observing systems are defined in Table 4.1.

are also independent of the choice of $r^2 > 0$ in (4.4): assumptions on data uncertainties are always irrelevant for the QoI target directions.)

4.3.4 Proxy potential

This section evaluates OSNAP’s proxy potential for the two chosen unobserved QoIs, Nordic Seas heat content and volume transport across the Iceland-Scotland ridge. Formally, the proxy potential is assessed via the dynamical constraints of OSNAP on the QoIs within the inverse modeling framework, measured by the relative uncertainty reduction $\tilde{\Delta}\sigma^2$ for each of the QoIs, induced by OSNAP. Here, uncertainty reduction is caused by dynamical information that is propagated from the OSNAP observations to the QoIs via the linearized model equations. In Section 3.5 it was explained how $\tilde{\Delta}\sigma^2$ is composed of two aspects, hypothetical proxy potential (F.1), and masking by noisy observations (F.2). The first aspect, hypothetical proxy potential, is the answer to question (II). That is, it evaluates to what degree dynamical adjustment mechanisms and pathways for the OSNAP observations and the QoIs are shared. If the OSNAP observations were noise-free, the effective proxy potential would be equal to the hypothetical proxy potential. The second aspect, masking by imperfect observations, deals with question (III), i.e., studies to what degree observational noise inhibits the extraction of information from the OSNAP observations.

Hypothetical proxy potential

Here, I first study F.1, the hypothetical proxy potential of noise-free OSNAP observations for the QoIs. It measures the potential of perfect (noise-free) OSNAP observations for informing the respective QoI, on a scale from 0 (no constraints) to 1 (perfect constraints), based on dynamical relationships in the GCM. For a given QoI and its associated target direction \mathbf{q} , the hypothetical proxy potential of the three different OSNAP observing systems is computed by the following expressions (see Section 3.5):

$$\text{OSNAP-West : } \langle \mathbf{q}, \mathbf{v}_1^W \rangle^2 \quad (4.20)$$

$$\text{OSNAP-East : } \langle \mathbf{q}, \mathbf{v}_1^E \rangle^2 \quad (4.21)$$

$$\text{OSNAP-Combined : } \langle \mathbf{q}, \mathbf{v}_1 \rangle^2 + \langle \mathbf{q}, \mathbf{v}_2 \rangle^2 \quad (4.22)$$

Here, \mathbf{q} , \mathbf{v}_1^W , \mathbf{v}_1^E , \mathbf{v}_1 , and \mathbf{v}_2 are all vectors within the control space \mathbb{R}^N , and $\langle \cdot, \cdot \rangle$ denotes the inner (or dot) product of two vectors. The vectors \mathbf{v}_1^W , \mathbf{v}_1^E , \mathbf{v}_1 , and \mathbf{v}_2 are OSNAP-informed directions and were presented in Section 4.3.2. Either of the target

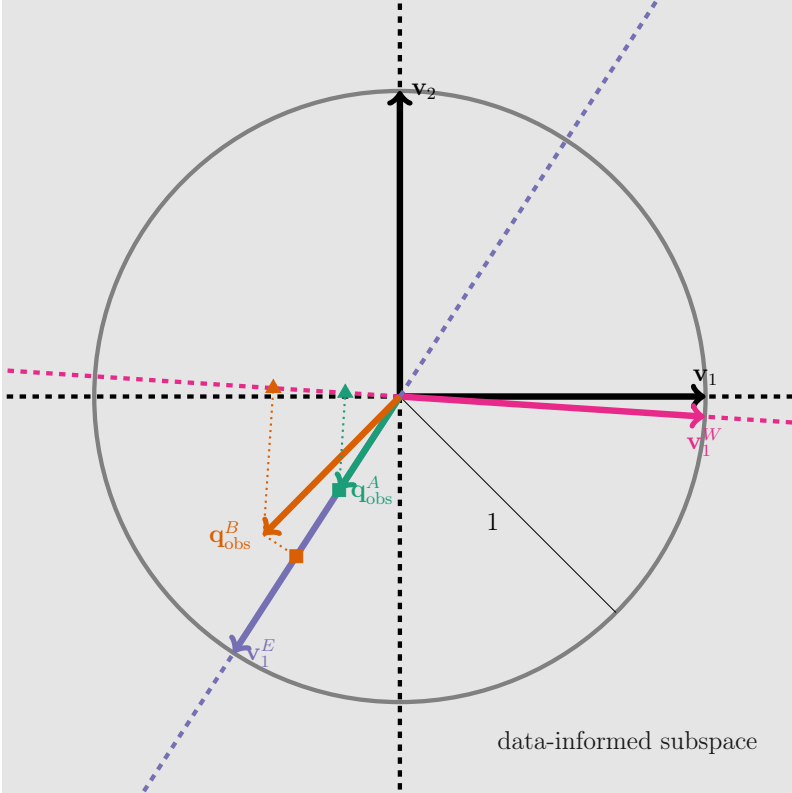


Figure 4.9: The OSNAP-informed directions and subspaces (replotted from Fig. 4.6), and the projections of the target directions q^A and q^B onto these subspaces. The gray plane is the two-dimensional subspace informed by the OSNAP-Combined observing system, and the dashed pink and purple lines show the one-dimensional subspaces informed by the OSNAP-West and OSNAP-East observing systems, respectively. The dark gray circle marks the unit sphere within the gray plane. q_{obs}^A (green vector) and q_{obs}^B (orange vector) are the projections of the respective target directions q^A and q^B onto the gray plane. The orange and green *triangles* mark the projections of q_{obs}^A and q_{obs}^B onto the pink line, while the orange and green *squares* are the projections of q_{obs}^A and q_{obs}^B onto the purple line.

directions \mathbf{q}^A and \mathbf{q}^B presented in Section 4.3.3 takes on the role of the vector \mathbf{q} in (4.20)-(4.22). Due to **(i)-ind**, **(i)-comb**, and **(i)-qoi**, the hypothetical proxy potentials in (4.20)-(4.22) are independent of the choices of $r^2, b^2 > 0$, i.e., varying assumptions on OSNAP noise and prior information via (4.4) and (4.8).

The hypothetical proxy potential of perfect OSNAP observations for the studied QoIs, computed by (4.20)-(4.22), are shown in Table 4.4. Two facts stand out. First, the hypothetical proxy potentials for volume transport across the Iceland-Scotland ridge (T_{IS}) are higher than for Nordic Seas heat content (OHC_{NS}), no matter which of the OSNAP observing system is considered. Second, the OSNAP-East observing system has an almost equally high potential to constrain the QoIs as the OSNAP-Combined observing system: If the observations were noise-free, OSNAP-East would reduce uncertainty in OHC_{NS} by exactly the same amount as the OSNAP-Combined observations, namely by 13%. For T_{IS} , the hypothetical proxy potential of OSNAP-East equals 39%; this value is not much smaller than the hypothetical proxy potential of OSNAP-Combined, which equals 40%. This means that, in the hypothetical case of noise-free observations, OSNAP-West would not add much (for T_{IS}), or even any (for OHC_{NS}), extra value to the information that can be extracted from OSNAP-East. Moreover, the hypothetical proxy potential of OSNAP-West alone is relatively low for both QoIs: 3% for OHC_{NS} and 17% for T_{IS} .

In the following, the two highlighted facts will be investigated further; first, by visualizing (4.20)-(4.22) formally, as projections in the OSNAP-informed subspace within the control space, and second, by establishing the link to sensitivity patterns and the underlying physics.

Fig. 4.9 replots the OSNAP-informed subspaces from Fig. 4.6: the light gray plane shows the two-dimensional subspace informed by the OSNAP-Combined observing system, and the pink and purple dashed lines the one-dimensional subspaces informed by the OSNAP-West and OSNAP-East observing systems, respectively. The green and orange vectors are \mathbf{q}_{obs}^A and \mathbf{q}_{obs}^B , the projections of the target directions associated with $QoI^A = OHC_{NS}$ and $QoI^B = T_{IS}$ onto the gray plane, i.e., the subspace informed by OSNAP-Combined. These vectors are shorter than 1 because the normalized target directions \mathbf{q}^A and \mathbf{q}^B have a component in the nullspace, orthogonal to the OSNAP-Combined-informed plane. Since the green vector is shorter than the orange vector, the nullspace component of OHC_{NS} is larger than the one of T_{IS} . $\|\mathbf{q}_{obs}^\bullet\|^2$, the squared length of the green/orange vector, equals $\langle \mathbf{q}^\bullet, \mathbf{v}_1 \rangle^2 + \langle \mathbf{q}^\bullet, \mathbf{v}_2 \rangle^2$, and is the hypothetical proxy potential for the respective QoI by the OSNAP-Combined observing system (see equation (4.22)). \mathbf{q}_{obs}^A (green vector) has a length of 0.37, which translates into a hypothetical proxy potential of $0.37^2 = 0.13$ for OHC_{NS} in Table 4.4, and \mathbf{q}_{obs}^B (orange vector) has a length of 0.64, such that the hypothetical proxy potential for T_{IS} computes as $0.64^2 = 0.40$. To obtain the hypothetical proxy potentials of the OSNAP-West and OSNAP-East observing systems, one projects \mathbf{q}_{obs}^A and \mathbf{q}_{obs}^B further onto the respective one-dimensional subspaces informed by OSNAP-West and OSNAP-East. (Since the one-dimensional subspaces informed by OSNAP-West and OSNAP-East are subspaces of the gray plane, the projection of \mathbf{q}^\bullet is the same as the projection of \mathbf{q}_{obs}^\bullet .) For OSNAP-

West, these projections result in the orange and green triangles on the pink dashed line in Fig. 4.9, and, for OSNAP-East, in the orange and green squares on the purple dashed line. The hypothetical proxy potentials are the squared distances of the triangles/squares from the origin. Since $\mathbf{q}_{\text{obs}}^A$ lines up perfectly with the purple dashed line, the OSNAP-West hypothetical proxy potential for OHC_{NS} is the same as the OSNAP-Combined hypothetical proxy potential, namely $0.37^2 = 0.13$. As for OHC_{NS} , $\mathbf{q}_{\text{obs}}^B$ projects better onto the purple line than onto the pink line. This results in a higher OSNAP-East hypothetical proxy potential ($0.62^2 = 0.39$) than OSNAP-West hypothetical proxy potential ($(-0.42)^2 = 0.17$) for T_{IS} .

The hypothetical proxy potential is determined by squared inner products of type $\langle \mathbf{q}, \mathbf{v} \rangle^2$ (equations (4.20)-(4.22)). The inner product $\langle \mathbf{q}, \mathbf{v} \rangle$ represents the projection of a QoI target direction \mathbf{q} onto an OSNAP-informed direction \mathbf{v} , as visualized in Fig. 4.9. Data-informed and target directions have been interpreted by means of sensitivity maps and dynamical adjustment mechanisms in Sections 4.3.1, 4.3.2 and 4.3.3. The projections in Fig. 4.9 can therefore be interpreted as *projections of sensitivity patterns* onto each other. Accordingly, the better the sensitivity maps of the target and OSNAP-informed directions project onto each other, the higher is the hypothetical proxy potential of the OSNAP observations for the QoI. As an example, Fig. 4.10 visualizes the way of computing the inner product between \mathbf{q}^A and \mathbf{v}_1^E , which were identified as the normalized prior-weighted sensitivity of the unobserved quantity OHC_{NS} in Fig. 4.8 and the observed quantity $\text{MHT}_{\text{OSNAP-E}}$ in Fig. 4.5, respectively. In the left column of Fig. 4.10, subplots (c),(g) replot Figs. 4.8(a),(c), and their magnifications are shown in subplots (a),(e). Similarly, in the right column of Fig. 4.10, subplots (d),(h) replot Figs. 4.5(b),(d), and their magnifications are shown in subplots (d),(f). The color shading in each of the shown grid cells in the magnified maps in Figs. 4.10 corresponds to an entry of the vectors $\mathbf{q}^A = (q_1, \dots, q_N)$ (in (a),(e)) and $\mathbf{v}_1^E = (v_1, \dots, v_N)$ (in (b),(f)), associated with the control variables $R_{\text{SW}\downarrow}$ (in (a),(b)) and τ_y (in (e),(f)) in the shown subregion of the North-East Atlantic and the Nordic Seas. The dashed arrows sketch the procedure of projecting \mathbf{q}^A onto \mathbf{v}_1^E , computed via the inner product $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle = \sum_{i=1}^N q_i \cdot v_i$. Here, the two columns in Fig. 4.10 present by far not all N vector entries of \mathbf{q}^A and \mathbf{v}_1^E ; for that, one would have to show normalized prior-weighted sensitivities to all global two- and three dimensional input fields that are named in Table 4.2. Note that the details in Figs. 4.10(a),(b),(e),(f), i.e., the similarities and differences in the adjustment mechanisms on a model grid point basis, are only resolved because the full spatial resolution of the ECCOv4 control space was pertained in this work. If, instead, the control space had been reduced a-priori, e.g., via the 'large region approach' [Kaminski and Rayner, 2008; Kaminski et al., 2015, 2018], and one of the 'large regions' had been defined to contain the shown magnified region of the North-East Atlantic and the Nordic Seas, Figs. 4.10(a),(b),(e),(f) would only show one value each, namely the spatial average over all grid cells in that region. This would lead to large aggregation errors in the projection [see also Kaminski et al., 2001].

The similar wind sensitivity patterns in Figs. 4.10(e) and (f) project very well onto each other. Physically speaking, this is because the unobserved quantity OHC_{NS} and the observed quantity $\text{MHT}_{\text{OSNAP-E}}$ have *similar pressure adjustment mechanisms and pathways*. Most prominently, both meridional wind stress sensitivity maps (in

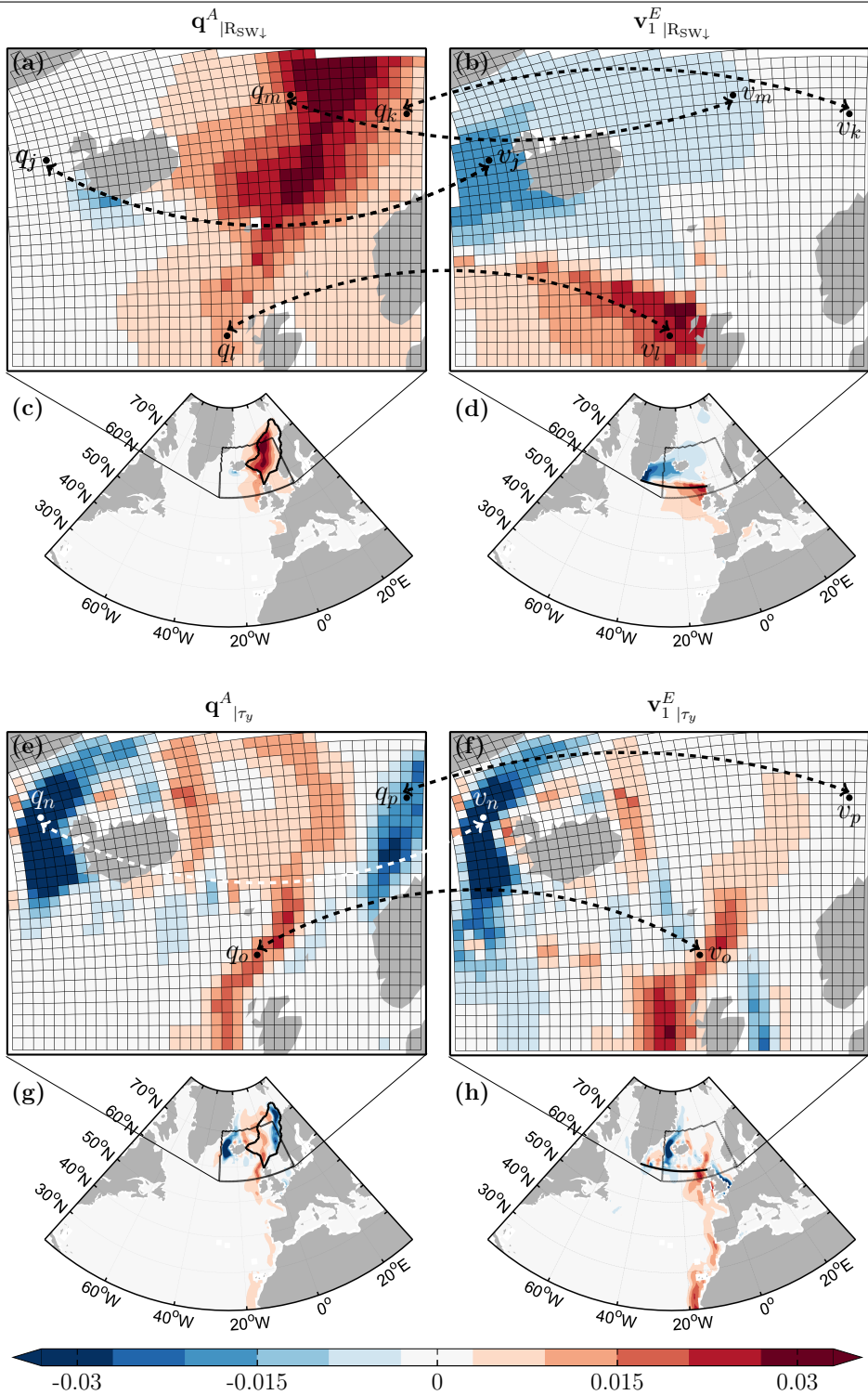


Figure 4.10: Caption appears on following page.

Figure 4.10: (Figure appears on preceding page.) The projection of \mathbf{q}^A , i.e., normalized prior-weighted sensitivities of the unobserved quantity OHC_{NS} , (left column) onto \mathbf{v}_1^E , i.e., normalized prior-weighted sensitivities of the observed quantity $\text{MHT}_{\text{OSNAP-E}}$, (right column). The maps show the restrictions of $\mathbf{q}^A, \mathbf{v}_1^E$ to $\mathbf{R}_{\text{SW}\downarrow}$ ((a)-(d)) and τ_y ((e)-(h)). Subplots (c),(g) replot Figs. 4.8(a),(c), and their magnifications are shown in subplots (a),(e). Subplots (d),(h) replot Figs. 4.5(b),(d), and their magnifications are shown in subplots (d),(f). Wind sensitivities in (e) and (f) project very well onto each other, while the projection of thermal sensitivities in (a) and (b) is poor.

Figs. 4.10(e),(f) exhibit a band of strongly negative sensitivity along the western Icelandic coast (e.g., reflected in the vector entries $q_n < 0, v_n < 0$). This location is a waveguide for the western end of the section between Iceland and Scotland, where pressure anomalies can effectively strengthen or weaken the Norwegian Atlantic Current. Negative sensitivity along the western Icelandic coast is visible in both Figs. 4.10(e) and (f) because changes in the strength of the Norwegian Atlantic Current, driven by wind perturbations along the western Icelandic coast, affect both OHC_{NS} and $\text{MHT}_{\text{OSNAP-E}}$ in a similar way (see Sections 4.3.1 and 4.3.3). Similarly, both wind sensitivity maps in Figs. 4.10(e),(f) exhibit a band of positive sensitivity along the eastern boundary of the North Atlantic (e.g., reflected in $q_o > 0, v_o > 0$), an upstream coastal wave guide for both OHC_{NS} and $\text{MHT}_{\text{OSNAP-E}}$. While OHC_{NS} and $\text{MHT}_{\text{OSNAP-E}}$ share most adjustment mechanisms and pathways, reflected in many shared large-scale features in Figs. 4.10(e) and (f), some sensitivity features are local and unique to only one of the two quantities. An example is the band of negative sensitivity along the Norwegian coast in Fig. 4.10(e) (e.g., $q_p < 0$), due to local Ekman dynamics (see Section 4.3.3), with no counterpart in Fig. 4.10(f) ($v_p = 0$).

Compared to the excellent projection of the wind sensitivities in Figs. 4.10(e),(f), the thermal sensitivities in Figs. 4.10(a),(b) project only poorly onto each other. The poor projection consists partly in non-overlapping sensitivities, e.g., west of Iceland ($v_l < 0$ but $q_l = 0$) or along the Norwegian coast ($q_k > 0$ but $v_k = 0$), due to the importance of *local processes* that affect only one of the two quantities, i.e., either OHC_{NS} or $\text{MHT}_{\text{OSNAP-E}}$. But even regions, where thermal sensitivities overlap, do not necessarily show similar sensitivity features. For instance, between Iceland and Norway, strongly positive sensitivities in Fig. 4.10(a) (e.g., $q_m > 0$), associated with local air-sea processes affecting OHC_{NS} (see Section 4.3.3), stand in contrast to negative sensitivities in Fig. 4.10(b) ($v_m < 0$), associated with the southward advection of temperature anomalies to the OSNAP-East transect (see Section 4.3.1). Meanwhile, positive sensitivities in the North-East Atlantic south of the OSNAP-East transect are a shared feature between Figs. 4.10(a) and (b) (e.g., $q_l > 0, v_l > 0$), and due to the northward advection of temperature anomalies affecting both OHC_{NS} and $\text{MHT}_{\text{OSNAP-E}}$ in a similar way (see Sections 4.3.1 and 4.3.3).

As seen in the last two paragraphs, Fig. 4.10 shows the occurrence of all three of the following cases:

- (a) overlapping sensitivities with positive contributions (e.g., $q_l \cdot v_l > 0, q_n \cdot v_n > 0$,

$q_o \cdot v_o > 0$) to the projection $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle = \sum_{i=1}^N q_i \cdot v_i$;

(b) overlapping sensitivities with negative contributions (e.g., $q_m \cdot v_m < 0$) to $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle$;

(c) non-overlapping sensitivities with no contribution (e.g., $q_j \cdot v_j = 0$, $q_k \cdot v_k = 0$, $q_p \cdot v_p = 0$) to $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle$.

Note that “matching sensitivity patterns”, due to similar adjustment mechanisms for the QoI and observations, do not necessarily have to fall under the case (a), as in the example of Fig. 4.10, but can also be included in case (b). For instance, if one defined the observed quantity to be southward, rather than northward heat transport across OSNAP-East (i.e., $-\text{MHT}_{\text{OSNAP-E}}$ rather than $\text{MHT}_{\text{OSNAP-E}}$), the OSNAP-informed direction would be $-\mathbf{v}_1^E$ rather than \mathbf{v}_1^E , and all sensitivities in the right column of Fig. 4.10 would have flipped signs. Then, for instance, matching large-scale wind sensitivities in the upstream waveguides in Figs. 4.10(e) and (f) would carry opposite signs (rather than equal signs) and would therefore fall under the case (b), rather than under the case (a). Since $\langle \mathbf{q}^A, -\mathbf{v}_1^E \rangle^2 = \langle \mathbf{q}^A, \mathbf{v}_1^E \rangle^2$, the hypothetical proxy potential (see equation 4.21), and thus, the effective proxy potential, would remain the same.

Positive and negative contributions to the projection $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle$, i.e., coexistence of (a) and (b), partially cancel each other out. The more sensitivities overlap (i.e., the less of (c) occurs) and the less cancellations occur (i.e., the less (a) and (b) conflict), the larger is the *squared* inner product $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle^2$, determining the hypothetical proxy potential. Cancellation in information (i.e., coexistence of (a) and (b)), can be understood better by drawing an analogy to the phenomenon of wave interference, e.g., for light or acoustic waves. Both the observed OSNAP-East quantity and the unobserved QoI are the superposition of many different signals. The signals are triggered locally or remotely, and subsequently carried to OSNAP-East or the Nordic Seas by various adjustment mechanisms. Similarly, light waves can be the superposition of many individual light waves that originate from different point sources. The superposition of light waves results in constructive interference when the waves are in phase, or destructive interference when the waves arrive 180° out of phase. While overlapping sensitivities (case (a) and/or (b)) suggest that both OHC_{NS} and $\text{MHT}_{\text{OSNAP-E}}$ respond to similar signals, the superposition of these signals can lead to “constructive interference”, in the case of no cancellation (occurrence of only (a) or (b)), or “destructive interference”, in the case of cancellation (coexistence of (a) and (b)). That is, even if an observed quantity and an unobserved QoI are affected by similar oceanic processes, it is important that these processes affect both quantities with *consistent signs* - otherwise cancellation in information (or destructive interference) will occur. While non-overlapping sensitivities (case (c)) means that no information is transferred at all, cancellation in information (coexistence of (a) and (b)) means that information is in principle communicated from observations to QoI, but is not readily extractable, due to competing dynamical adjustments. Including either more observations or more prior information will help to extract the information (or bring the waves in phase). For instance, if prior information on air-sea heat fluxes in the Nordic Seas was perfect, corresponding to prior uncertainties $\Delta R_{\text{SW}\downarrow} = 0$, the (normalized) prior-weighted sensitivities in Figs. 4.10(a)-(d) would be zero over the Nordic Seas. This would result in $q_m = v_m = 0$ and the cancellation due to the negative contribution

by the pair (q_m, v_m) (case (b)) would disappear.

Fig. 4.11(b) is a quantitative summary of Fig. 4.10: upward-pointing bars in Fig. 4.11(b) show the positive contributions (case (a)) to $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle$, and downward-pointing bars the negative contributions (case (b)). While Fig. 4.10 shows only contributions by subsets of $R_{\text{SW}\downarrow}$ and τ_y , Fig. 4.11(b) shows contributions by *all control variables*, summarized by means of the 5 groups defined in Table 4.2, resulting in 5 upward- and 5 downward-pointing bars. For instance, the contribution $q_l \cdot v_l > 0$ by the pair (q_l, v_l) from Figs. 4.10(a),(b) is contained in the upward-pointing yellow bar in Fig. 4.11(b), and the contribution $q_m \cdot v_m < 0$ by the pair (q_m, v_m) from Figs. 4.10(a),(b) in the downward-pointing yellow bar in Fig. 4.11(b). Contributions by the pairs (q_j, v_j) and (q_k, v_k) are zero, and are therefore not contained in any of the yellow bars in Fig. 4.11(b). Upward- and downward pointing yellow bars of similar amplitude in Fig. 4.11(b) reflects the cancellation effect occurring within the group of thermal control variables, which was discussed in the last few paragraphs. On the contrary, most pairs in Figs. 4.10(e),(f) have a positive contribution (e.g., $q_n \cdot v_n > 0$, $q_o \cdot v_o > 0$), and are therefore contained in the upward-pointing green bar in Fig. 4.11(b). The downward-pointing green bar has negligible amplitude. The number +0.37, shown below Fig. 4.11(b) is the value of $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle$, obtained by summing up all 10 bars in the subpanel. That is, contributions of +0.10 (upward-pointing yellow bar) and -0.06 (downward-pointing yellow bar) by the group *thermal*, +0.31 (upward-pointing green bar) and -0.01 (downward-pointing green bar) by the group *wind*, +0.01 (upward-pointing pink bar) and -0.00 (absent downward-pointing pink bar) by the group *param*, and negligible contributions by the remaining groups *precip* and *initial* sum up to $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle = 0.37$. Note that the maximum absolute value for $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle$ is equal to 1, only attained in the case of a perfect projection. This is because \mathbf{q}^A and \mathbf{v}_1^E are *normalized* vectors with Euclidean norms $\sum_{i=1}^N q_i^2$ and $\sum_{i=1}^N v_i^2$ equal to 1 (cf. colorbar for the vector entries q_i, v_i in Fig. 4.10, while by far not all N vector entries are shown). Note that the value $\langle \mathbf{q}^A, \mathbf{v}_1^E \rangle = 0.37$ is equal to the distance of the dark green square from the origin in Fig. 4.9, where the dark green square marked the projection of \mathbf{q}^A onto the subspace spanned by \mathbf{v}_1^E .

The remaining subpanels in Fig. 4.11 show the positive and negative contributions to the remaining inner products from equations (4.20)-(4.22). Inner products that project the OHC_{NS} target direction \mathbf{q}^A onto an OSNAP-informed direction are shown in Figs. 4.11(a)-(d), and inner products that project the T_{IS} target direction \mathbf{q}^B onto an OSNAP-informed direction are displayed in Figs. 4.11(e)-(h). If in a subpanel of Fig. 4.11 the upward- and downward pointing bars of the *same* color both have significant amplitudes, there is cancellation *within* this group of control variables, when the respective target direction is projected onto the respective OSNAP-informed direction. Bars of *different* colors in a subpanel of Fig. 4.11 that point in opposite directions reflect cancellations *among* groups for the respective projection. As before, the maximal absolute value for the respective inner products, shown below each subpanel, is equal to 1. The hypothetical proxy potential, is determined by the *square* of this value (see (4.20)-(4.22)).

In all subpanels of Fig. 4.11, *wind* is the group with the longest bars. At the same time, there is essentially *no* cancellation effect within the group *wind*, since, in each sub-

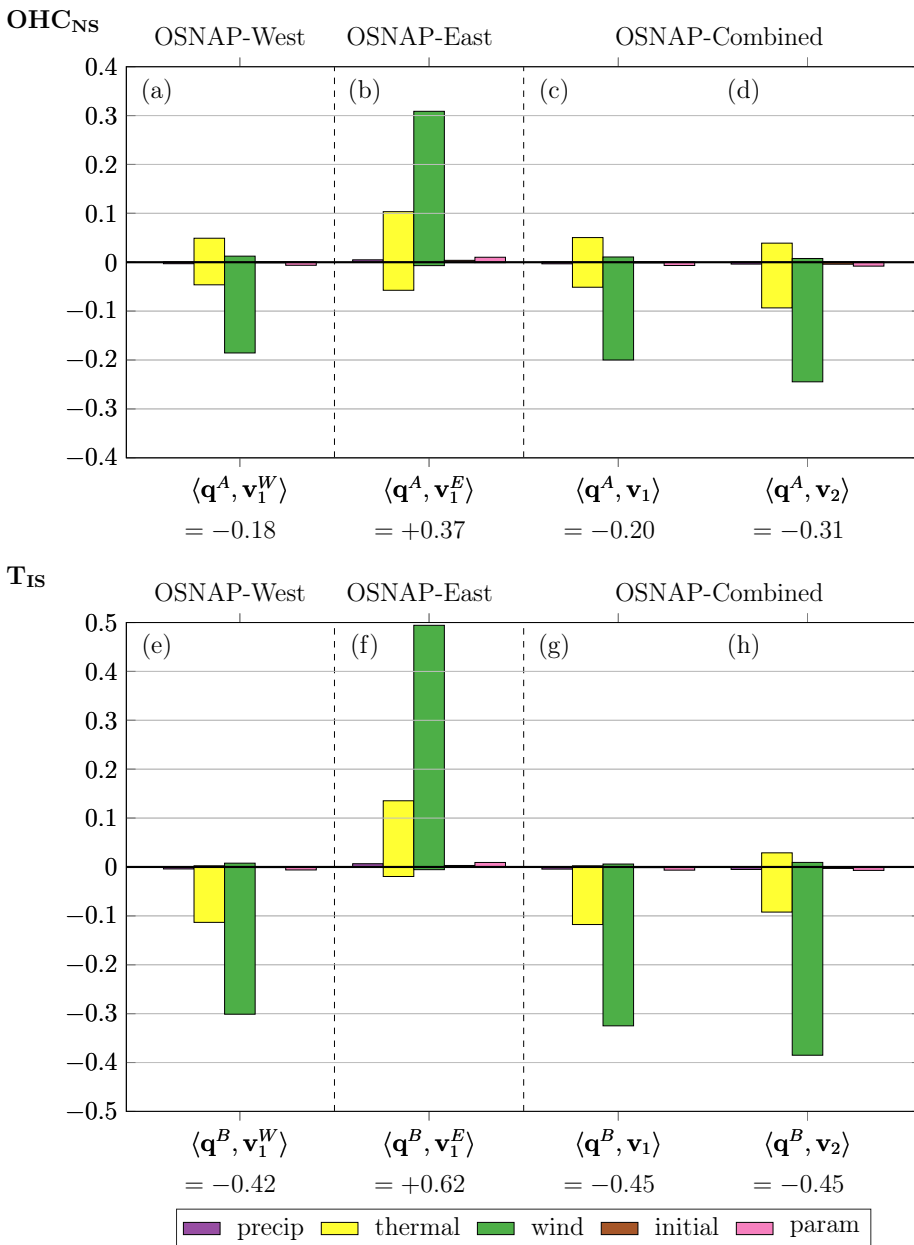


Figure 4.11: Contributions by the 5 groups of control variables, defined in Table 4.2, to the inner products of type $\langle \mathbf{q}, \mathbf{v} \rangle$ that appear in (4.20)-(4.22); these inner products represent projections of normalized prior-weighted sensitivity maps, as shown in Fig. 4.10. Subpanels (a)-(d): projections involving the OHC_{NS} target direction \mathbf{q}^A . Subpanels (e)-(h): projections involving the T_{IS} target direction \mathbf{q}^B . Below each subpanel, the value for the respective inner product is shown; this number is obtained by summing up the 10 bars in the subpanel, where each bar shows either the positive (upward-pointing bars) or negative (downward-pointing bars) contribution by one of the 5 groups of control variables.

panel, either the upward-pointing or the downward-pointing green bar is of negligible amplitude. These two facts combined indicate that for both QoIs, information is gained to a large degree by uncertainty reduction in *wind*, induced by the OSNAP observations. The underlying reason is that the QoIs and the observed quantities have similar pressure adjustment mechanisms and pathways, which was already highlighted in the similar wind sensitivity patterns in Fig. 4.10 (giving rise to Fig. 4.11(b)). Most prominently, all meridional wind stress sensitivity maps in Figs. 4.5(c),(d), Figs. 4.7(c),(d), and Figs. 4.8(c),(d) exhibit the band of strong sensitivity along the western Icelandic coast. This recurrent feature of strong amplitude yields an excellent projection of the respective normalized prior-weighted wind stress sensitivity maps, which results in long green bars in Fig. 4.11. Due to the varying sign of the sensitivity band, the projections onto the OSNAP-West-informed direction (Figs. 4.11(a),(e)), onto \mathbf{v}_1 (Figs. 4.11(c),(g)), and onto \mathbf{v}_2 (Figs. 4.11(d),(h)) relate to long green bars of *negative* amplitude, and the projections onto the OSNAP-East-informed direction (Figs. 4.11(b),(f)) relate to long green bars of *positive* amplitude. The fact from Table 4.4 that the hypothetical proxy potentials of OSNAP-East are higher than the hypothetical proxy potentials of OSNAP-West stems to a large degree from the different strengths of the wind stress sensitivities projections. This can be identified in Fig. 4.11, where the green bars in the subpanels (b) and (f) are longer than in the subpanels (a) and (e), respectively. The underlying reason for the stronger wind stress sensitivity projections involving OSNAP-East is that $\text{MHT}_{\text{OSNAP-E}}$ is relatively more sensitive to wind forcing than $\text{MHT}_{\text{OSNAP-W}}$, as identified by the bar charts in Figs. 4.5(e) and (f). A second highlighted fact from Table 4.4 was that the hypothetical proxy potentials for T_{IS} are higher than for OHC_{NS} . This fact is again reflected in Fig. 4.11, where green bars in the subpanels (e)-(h) are longer than in the subpanels (a)-(d). Similarly as before, the stronger wind stress sensitivity projections involving T_{IS} are originated in very strong relative sensitivity to wind forcing of T_{IS} (bar charts in Figs. 4.8(e),(f)).

Thermal sensitivity patterns project much more weakly onto each other than wind stress sensitivity patterns, reflected by the overall shorter yellow bars in Fig. 4.11. This is despite the fact that all OSNAP-informed directions and, in addition, the OHC_{NS} target direction are to the largest degree controlled by the group of *thermal* forcing fields (Figs. 4.5(e),(f), Figs. 4.7(e),(f), and Fig. 4.8(e)). However, the thermal sensitivity maps have more features that reflect rather local processes than the wind stress sensitivity maps, as already identified in Fig. 4.10, explaining subpanel (b) in Fig. 4.11. While non-overlapping thermal sensitivities for OHC_{NS} and the observed OSNAP quantities are reflected by overall *short* upward- and downward pointing yellow bars in Figs. 4.11(a)-(d), a second notable factor is cancellation, evident by upward- and downward pointing yellow bars of *similar amplitude* in all subpanels of Figs. 4.11(a)-(d). On the other hand, neither the “non-overlap” nor the cancellation effect occurs for the QoI T_{IS} (yellow bars in Figs. 4.11(e)-(h)). The weak thermal sensitivities of T_{IS} are mostly concentrated around Iceland (Fig. 4.8(b)), due to the thermal wind balance mechanism in response to alteration of the zonal density gradient across the section between Iceland and Scotland (see Section 4.3.3). These weak thermal sensitivities around Iceland happen to fall into regions where both $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$ show sensitivity of uniform sign (Figs. 4.5(a),(b)), due to advective processes (see Section 4.3.1), such that no cancellation within the group *thermal* occurs (i.e., no cancellation of yellow upward- by yellow

downward-pointing bars, and vice versa, in any of the subpanels (e)-(h) of Fig. 4.11). Moreover, the superposition of thermal-driven and wind-driven signals affecting T_{IS} is “in phase” with the superposition of thermal-driven and wind-driven signals affecting the OSNAP observations, such that no cancellation between the yellow and the green bars occurs in any of the subpanels (e)-(h) of Fig. 4.11. The non-existence of cancellation effects in information by thermal processes, together with the strong wind stress sensitivity projections discussed in the last paragraph, leads to a strong projection of the T_{IS} target direction onto the OSNAP-informed directions, and, thus, results in a high OSNAP hypothetical proxy potentials for T_{IS} .

For all inner products (Fig. 4.11(a)-(h)), the groups *precip*, *initial*, and *param* show negligible contribution. This characteristic is consistent with the bar charts in Figs. 4.8(e),(f), Figs. 4.7(e),(f), and Figs. 4.5(e),(f); these charts elucidated that the groups *precip*, *initial*, and *param* have tiny relative importance for both QoIs (Figs. 4.8(e),(f)) and the observed OSNAP quantities (Figs. 4.5(e),(f)). As a result, the relative contribution of these groups to any of the inner products $\langle \mathbf{q}, \mathbf{v} \rangle$, and thus, the hypothetical proxy potentials (equations (4.20)-(4.22)), *has* to be small. This means that the groups *precip*, *initial*, and *param* do *not* participate in transferring information from the observations to the QoIs for the studied setup, i.e., on a five-year timescale.

Masking by noise

The second factor that determines OSNAP’s effective proxy potential for the QoIs is **F.2**, which characterizes masking by imperfect OSNAP observations. The information transfer factors regulate what fraction of the OSNAP hypothetical proxy potentials can be extracted, and are independent of the considered QoI. Each information transfer factor is of the form $0 < \frac{\lambda}{\lambda+1} < 1$, depending on the eigenvalue $\lambda = \lambda_1^W, \lambda_1^E, \lambda_1, \lambda_2$ that corresponds to the eigenvector (or OSNAP-informed direction) $\mathbf{v} = \mathbf{v}_1^W, \mathbf{v}_1^E, \mathbf{v}_1, \mathbf{v}_2$. λ_1^W and λ_1^E are the prior-to-noise ratios of OSNAP-West and OSNAP-East, respectively. If these prior-to-noise ratios were equal to 1, the prior uncertainties for the OSNAP model quantities would be exactly as high as the observational uncertainties (the noise), and the information transfer factors would take the value of 50%, reflecting a “fair compromise” between trusting prior model information and information injected by the observations.

As opposed to the hypothetical proxy potentials, the eigenvalues and, thus, the information transfer factors *do* depend on the choices of $r^2, b^2 > 0$, the scaling factors used for the noise and prior covariances $\mathbf{\Gamma}_{\text{noise}}[r^2] = r^2 \cdot \mathbf{\Gamma}_{\text{noise}}$ and $\mathbf{\Gamma}_{\text{prior}}[b^2] = b^2 \cdot \mathbf{\Gamma}_{\text{prior}}$. The scaling factors reflect the assumptions of less ($r^2 < 1$) or more ($r^2 > 1$) noisy OSNAP observations and/or more ($b^2 < 1$) or less ($b^2 > 1$) certain prior information. The dependence of the eigenvalues on $r^2, b^2 > 0$ is as follows:

$$\lambda[r^2, b^2] = (b/r)^2 \cdot \lambda[1, 1],$$

as seen in **(ii)-ind** and **(ii)-comb**. The information transfer factors can therefore be written as

$$\frac{\lambda}{\lambda+1}[r^2, b^2] = \frac{(b/r)^2 \cdot \lambda[1, 1]}{(b/r)^2 \cdot \lambda[1, 1] + 1} = \frac{\lambda[1, 1]}{\lambda[1, 1] + (r/b)^2}.$$

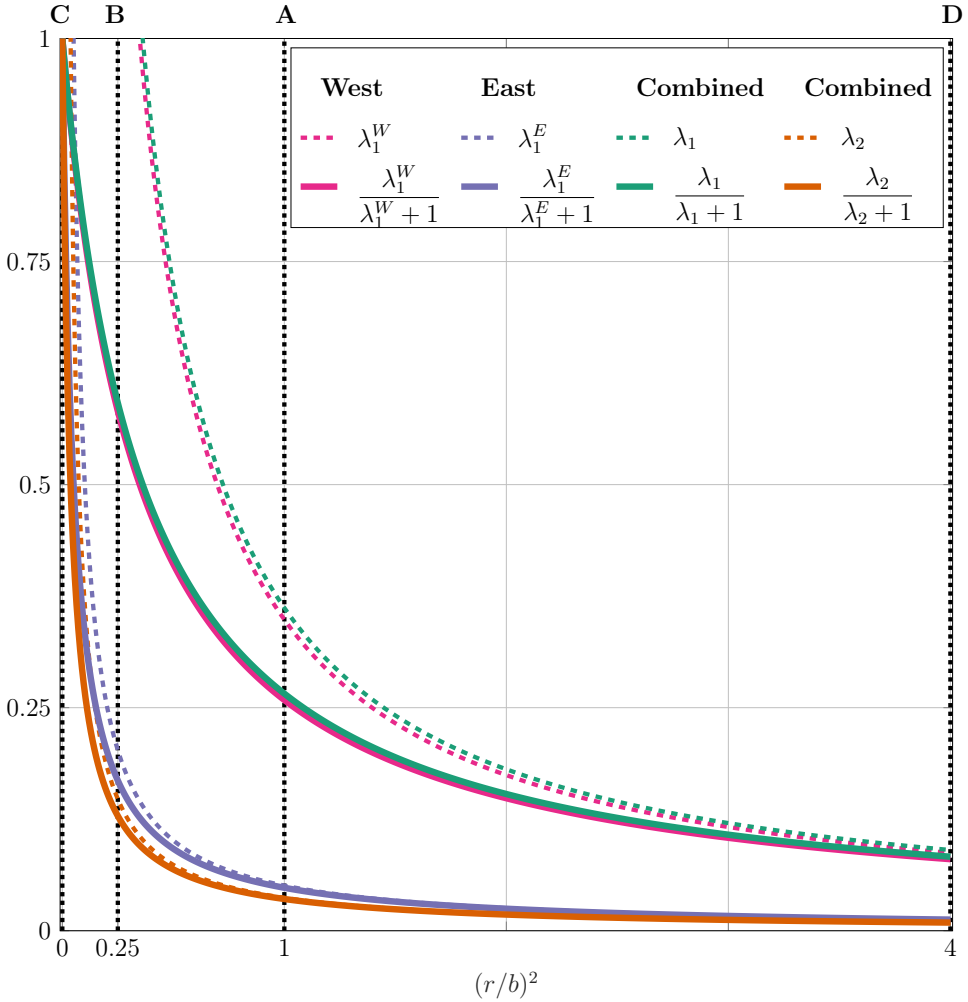


Figure 4.12: Eigenvalues λ (dashed lines) and associated information transfer factors $0 < \frac{\lambda}{\lambda+1} < 1$ (solid lines), associated with the different OSNAP observing systems, defined in Table 4.1. The information transfer factors indicate what fraction of the hypothetical proxy potential (reported in Table 4.4) of the OSNAP observations can be retrieved. Eigenvalues and information transfer factors are shown as a function of the ratio $(r/b)^2$, where $r^2, b^2 > 0$ are scaling factors for the noise and prior covariances $\Gamma_{\text{noise}}[r^2] = r^2 \cdot \Gamma_{\text{noise}}$ and $\Gamma_{\text{prior}}[b^2] = b^2 \cdot \Gamma_{\text{prior}}$ (equations (4.8) and (4.4)), reflecting the assumptions of less ($r^2 < 1$) or more ($r^2 > 1$) noisy OSNAP observations and/or more ($b^2 < 1$) or less ($b^2 > 1$) certain prior information. The vertical dotted black lines mark the values of $(r/b)^2$ for the scenarios C, B, A, and D (from left to right); these scenarios are discussed in the text.

Thus, both the eigenvalues and the information transfer factors are a function of the ratio $(r/b)^2$. These functions are shown in Fig. 4.12. As already identified in relation (4.14), the OSNAP-West prior-to-noise ratio λ_1^W (dashed pink curve) is about $(2.5)^2$ times larger than the OSNAP-East prior-to-noise ratio λ_1^E (dashed purple curve), for any value of $(r/b)^2$ (see (iii)-ind). For the case study in this work, the first eigenvalue λ_1 associated with OSNAP-Combined (dashed green curve) almost coincides with λ_1^W (dashed pink curve), and the second eigenvalue λ_2 associated with OSNAP-Combined (dashed orange curve) almost coincides with λ_1^E (dashed purple curve). Accordingly, the information transfer factors associated with λ_1 and λ_2 (solid green and orange curves) follow the information transfer factors associated with λ_1^W and λ_1^E (solid pink and purple curves).

The dotted vertical line corresponding to scenario A ($(r/b)^2 = 1$) in Fig. 4.12 covers the case $r^2 = b^2 = 1$ in equations (4.4) and (4.8). This case reflects the setup in which the OSNAP data uncertainties are adopted from Lozier et al. [2019] (see Section 4.2.2), and the prior uncertainties are adopted from the ECCOv4r3 inverse modeling framework (Fukumori et al., 2017, see Section 4.2.4). For scenario A, the prior-to-noise ratios λ_1^W and λ_1^E are only about 0.36 (dotted pink curve) and 0.05 (dotted purple curve), respectively (cf. also (4.13)). The fact that λ_1^W and λ_1^E are much smaller than 1 means that the prior uncertainties in the simulated observed OSNAP quantities $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$ are much smaller than the assumed noise in the OSNAP data. In fact, the prior uncertainty in $\text{MHT}_{\text{OSNAP-W}}$ ($\text{MHT}_{\text{OSNAP-E}}$) is only 36% (5%) of the noise for the measured mean OSNAP-West (OSNAP-East) heat transport. As a result, the information transfer factors take the values 0.26 and 0.05 for scenario A, which means that one can extract 26% and 5% of the hypothetical proxy potentials of OSNAP-West and OSNAP-East, respectively. The information transfer factors corresponding to λ_1 and λ_2 “inherit” the values of 27% (solid green curve) and 4% (solid orange curve), respectively.

Scenario B ($(r/b)^2 = 0.25$) includes the cases in which either (i) the OSNAP standard deviations are assumed to be half as large as in Lozier et al. [2019] ($r^2 = 0.25$ in (4.4)), or (ii) the prior standard deviations are assumed to be double as large as in ECCOv4r3 ($b^2 = 4$ in (4.8)). In scenario B, the OSNAP-West prior-to-noise ratio λ_1^W and the OSNAP-East prior-to-noise ratio λ_1^E are much larger than in Scenario A: 1.44 for OSNAP-West (dashed pink curve, falling out of the shown range), and 0.21 for OSNAP-East (dashed purple curve). The fact that λ_1^W is larger than 1 means that OSNAP-West data is qualitatively better than the prior information about its simulated counterpart in the model ($\text{MHT}_{\text{OSNAP-W}}$). This is still not the case for the OSNAP-East data. As a result of the higher prior-to-noise ratios, a higher portion of the OSNAP-West and OSNAP-East hypothetical proxy potentials can be extracted than in Scenario A, namely about 59% (solid pink curve) and 17% (solid purple curve), respectively. Higher retrievable fractions of the OSNAP hypothetical proxy potentials are also reflected in higher values for the information transfer factors corresponding to λ_1 and λ_2 : they evaluate as 59% (solid green curve) and 13% (solid orange curve), respectively.

Scenario C ($(r/b)^2 \searrow 0$) is a hypothetical scenario. It reflects the cases (i) in which the observations are noise-free ($r^2 \searrow 0$), or (ii) the prior uncertainties are infinitely high ($b^2 \nearrow \infty$). All OSNAP-informed directions correspond to infinitely high eigenvalues

(dashed curves), and therefore, all OSNAP observations are of equal (perfect) quality. All information transfer factors take the value 1 (solid curves) and 100% of all OSNAP hypothetical proxy potentials can be retrieved.

Scenario D ($(r/b)^2 = 4$) includes the cases in which either (i) the OSNAP standard deviations are assumed to be double as large as in Lozier et al. [2019] ($r^2 = 4$ in (4.4)), or (ii) the prior standard deviations are assumed to be half as large as in ECCOV4r3 ($b^2 = 0.25$ in (4.8)). For scenario D, the prior-to-noise ratios λ_1^W and λ_1^E are very small: only about 0.09 and 0.01, respectively. This means that the OSNAP data is too noisy compared to prior information that is already contained in the model, such that OSNAP observations cannot introduce much reliable new information. As a result, the information transfer factors take very small values, 8% for OSNAP-West (solid pink curve) and 1% for OSNAP-East (solid purple curve), and only a very small fraction of the OSNAP hypothetical proxy potentials can be retrieved.

Effective proxy potential

Fig. 4.13 shows the effective proxy potential of the OSNAP observing systems for the unobserved QoIs OHC_{NS} (in (a),(b)) and T_{IS} (in (c),(d)). The effective OSNAP proxy potentials are shown as a function of the ratio $(r/b)^2$, where $r^2, b^2 > 0$ are the scaling factors (in equations (4.4) and (4.8)) that reflect the assumptions of less ($r^2 < 1$) or more ($r^2 > 1$) noisy OSNAP observations and/or more ($b^2 < 1$) or less ($b^2 > 1$) certain prior information. The effective proxy potential is measured by relative uncertainty reduction $\tilde{\Delta}\sigma^2$ for each of the QoIs, caused by dynamical information that is propagated from OSNAP-West, OSNAP-East, or OSNAP-Combined to the respective QoIs via the linearized model equations. Relative uncertainty reduction $\tilde{\Delta}\sigma^2$ by OSNAP-West (pink curve in Figs. 4.13 (a) and (c)) is computed as

$$\frac{\lambda_1^W}{\lambda_1^W + 1} [(r/b)^2] \cdot \langle \mathbf{q}, \mathbf{v}_1^W \rangle^2, \quad (4.23)$$

and $\tilde{\Delta}\sigma^2$ by OSNAP-East (purple curve in Figs. 4.13 (a) and (c)) as

$$\frac{\lambda_1^E}{\lambda_1^E + 1} [(r/b)^2] \cdot \langle \mathbf{q}, \mathbf{v}_1^E \rangle^2, \quad (4.24)$$

for the respective target directions $\mathbf{q} = \mathbf{q}^A$ (in Fig. 4.13(a)) and $\mathbf{q} = \mathbf{q}^B$ (in Fig. 4.13(c)). Equations (4.23) and (4.24) highlight that relative uncertainty reduction $\tilde{\Delta}\sigma^2$ is a combination of the two factors discussed in the last two subsections: the hypothetical proxy potential (factor F.1), and masking by noisy observations (factor F.2). In fact, the hypothetical proxy potential and the associated information transfer factor are simply multiplied together.

$\tilde{\Delta}\sigma^2$ by the OSNAP-Combined observing system (black curve in Figs. 4.13 (a) and (c)) is computed as

$$\frac{\lambda_1}{\lambda_1 + 1} [(r/b)^2] \cdot \langle \mathbf{q}, \mathbf{v}_1 \rangle^2 + \frac{\lambda_2}{\lambda_2 + 1} [\rho/\beta] \cdot \langle \mathbf{q}, \mathbf{v}_2 \rangle^2, \quad (4.25)$$

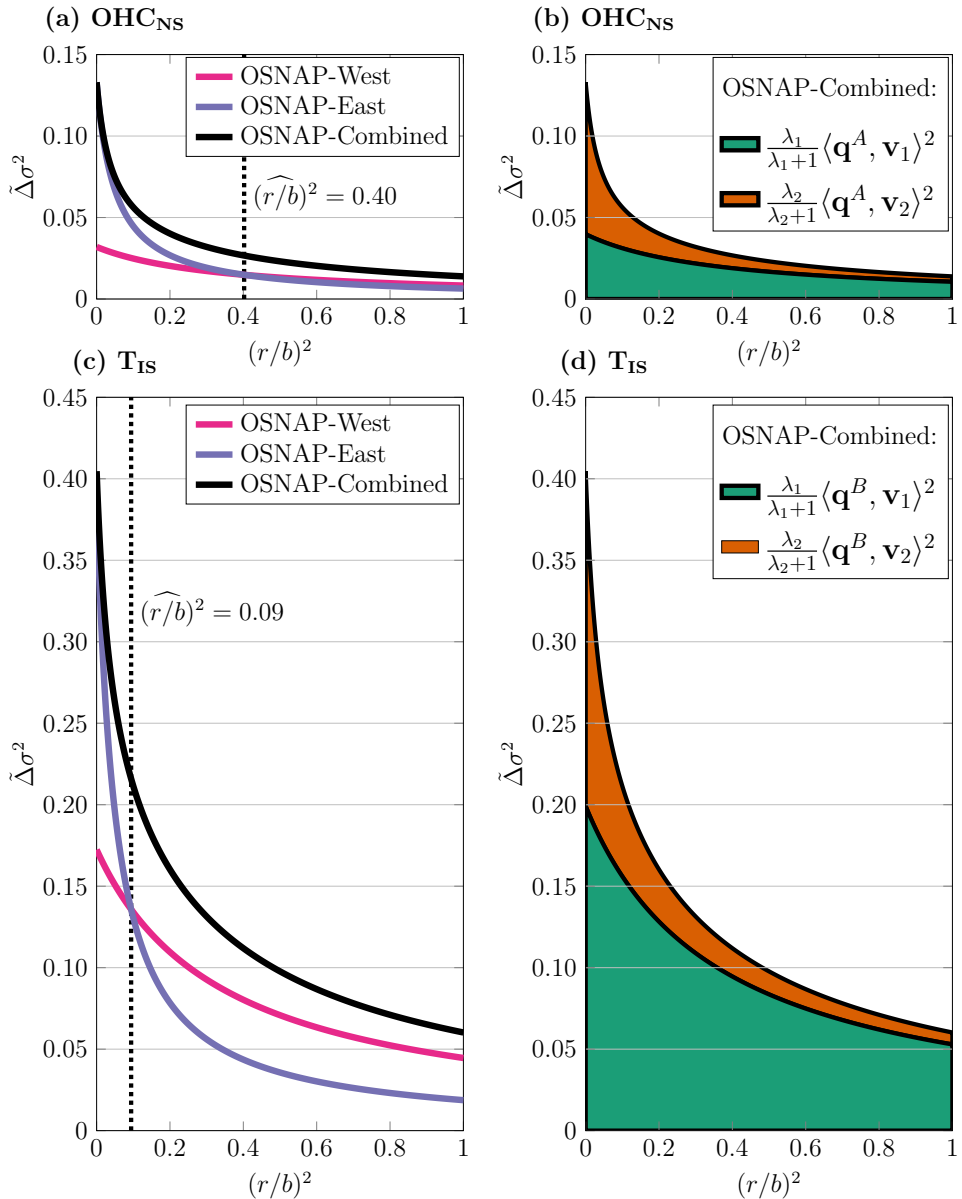


Figure 4.13: Subpanels (a),(c): Effective proxy potential (or relative uncertainty reduction) $\hat{\Delta}\sigma^2$ for the unobserved QoIs (a) OHC_{NS} and (c) T_{IS} by the three OSNAP observing systems defined in Table 4.1, shown as a function of the ratio $(r/b)^2$ (cf. Fig. 4.12). The dashed vertical lines in (a) and (c) are drawn at the QoI-dependent value $(\widehat{r/b})^2$, where the pink curve intersects the purple curve. For scenarios left of the vertical line (including the noise-free case $(r/b)^2 = 0$), OSNAP-East is more informative for the respective QoI, and for scenarios right of the vertical line (including the case $r^2 = b^2 = 1$, adopted from Lozier et al. [2019] and ECCOV4r3) OSNAP-West is more informative. Subpanels (b),(d): Contributions by $(\mathbf{v}_1, \lambda_1)$ (green) and $(\mathbf{v}_2, \lambda_2)$ (orange) to OSNAP-Combined-induced $\hat{\Delta}\sigma^2$ for (b) OHC_{NS} and (d) T_{IS} , adding up to the black line in Figs. 4.13(a) and (c), respectively.

for $\mathbf{q} = \mathbf{q}^A$ (in Fig. 4.13(a)) and $\mathbf{q} = \mathbf{q}^B$ (in Fig. 4.13(c)). The black curves in Figs. 4.13(a) and (c) always exceed the pink and purple curves, reflecting the fact that the OSNAP-Combined observing system is always more informative than each of the individual OSNAP-West or OSNAP-East observing systems. However, summing up the purple and pink curves would result in a curve that exceeds the black curve; this is due to redundancy in the OSNAP-West and the OSNAP-East observing systems. This redundancy gets removed by performing the sum in (4.25), instead of the sum of the pink and purple curve. Each of the two summands in (4.25) is again a product of the factors **F.1** and **F.2**. Fig. 4.13(b) shows the contributions by the first summand (green) and the second summand (orange) to $\tilde{\Delta}\sigma^2$ in (4.25) for OHC_{NS} . That is, the sum of the orange and the green contributions in Fig. 4.13(b) is equal to the black curve in Fig. 4.13(a). Similarly, Fig. 4.13(d) shows the contributions by the two summands to $\tilde{\Delta}\sigma^2$ for T_{IS} , the black curve in Fig. 4.13(c).

For the hypothetical cases of noise-free OSNAP observations or infinitely high prior uncertainties ($(r/b)^2 \searrow 0$), the information transfer factors are equal to 1, as shown in scenario C in Fig. 4.12. For noise-free observations, the effective proxy potential of the OSNAP observing systems for the QoIs would therefore be equal to the hypothetical proxy potentials shown in Table 4.4. For instance, in Fig. 4.13(a), relative uncertainty reduction by both the OSNAP-East (purple curve) and the OSNAP-Combined observing system (black curve) is $\tilde{\Delta}\sigma^2[0] = 0.13$, consistent with hypothetical proxy potentials of 13% for OHC_{NS} (Table 4.4).

However, for more realistic scenarios, in which observations are noisy, the effective information gain by the observations is damped. For the case $(r/b)^2 = 1$ (scenario A in Fig. 4.12), i.e., OSNAP data uncertainties as estimated in [Lozier et al. \[2019\]](#) and prior uncertainties as in [ECCOV4r3](#), the relative uncertainty reduction in OHC_{NS} is $\tilde{\Delta}\sigma^2[1] < 2\%$ for all OSNAP observing systems (pink, purple, and black curves in Fig. 4.13(a)). For T_{IS} , relative uncertainty reduction by OSNAP-West, OSNAP-East, and OSNAP-Combined is $\tilde{\Delta}\sigma^2[1] = 4\%, 2\%, 6\%$, respectively (pink, purple, and black curves in Fig. 4.13(c)). The small values for relative uncertainty reduction $\tilde{\Delta}\sigma^2[1]$ for both QoIs and by all OSNAP observing systems are due to the fact that the prior information in [ECCOV4r3](#) is assumed to be relatively high compared to the OSNAP data quality (see last subsection). Doubling [ECCOV4r3](#) standard deviations would correspond to the case $(r/b)^2 = 0.25$ (scenario B in Fig. 4.12). For OHC_{NS} , relative uncertainty reduction by OSNAP-Combined would still be below $\tilde{\Delta}\sigma^2[0.25] = 4\%$ (black curve in Fig. 4.13(a)), and even lower by the individual OSNAP-West and OSNAP-East observing systems (pink and purple curves in Fig. 4.13(a)). Uncertainty in T_{IS} would be reduced by $\tilde{\Delta}\sigma^2[0.25] = 10\%, 7\%, 14\%$ by OSNAP-West, OSNAP-East, and OSNAP-Combined, respectively (pink, purple, black curves in Fig. 4.13(c)).

Importantly, while OSNAP-East is more informative than OSNAP-West for both QoIs in the hypothetical noise-free scenario, in realistic (noisy) scenarios, OSNAP-West is more informative than OSNAP-East. This is reflected in each of the Figs. 4.13(a),(c) by the fact that the pink OSNAP-West curve intersects the purple OSNAP-East curve for a QoI-dependent value $(\widehat{r/b})^2$, marked by the vertical dotted line. The intersection

occurs at $(\widehat{r/b})^2 = 0.40$ for OHC_{NS} , and at $(\widehat{r/b})^2 = 0.09$ for T_{IS} . For scenarios left of the vertical line, including the noise-free scenario $(r/b)^2 = 0$, the purple OSNAP-East curve lies *above* the pink OSNAP-West curve, which means that the OSNAP-East observing system is more informative for the respective QoI than the OSNAP-West observing system. This is due to higher OSNAP-East hypothetical proxy potentials, which are originated in the fact that the dynamical adjustment mechanisms for the QoIs are more similar to the dynamical adjustment mechanisms for OSNAP-East heat transport than the dynamical adjustment mechanisms for OSNAP-West heat transport. For scenarios right of the vertical line, including the scenario $r^2 = b^2 = 1$ (i.e., OSNAP data uncertainties as estimated in [Lozier et al. \[2019\]](#) and prior uncertainties as in ECCOv4r3), the purple OSNAP-East curve lies *below* the pink OSNAP-West curve, which means that the OSNAP-West observing system is more informative. This is due to the fact that the purple OSNAP-East curve decays much more rapidly than the pink OSNAP-West curve for increasing $(r/b)^2$ and both QoIs (Figs. 4.13(a) and (c)). In fact, the decay rates of the pink OSNAP-West curve and the purple OSNAP-East curve are independent of the QoI. For instance, the decay rates of the purple OSNAP-East curves in Figs. 4.13(a) and (c) are exactly the same - the difference consists rather in the OSNAP-East curve in Fig. 4.13(a) starting at a higher base value $\tilde{\Delta}\sigma^2[0]$ (the hypothetical proxy potential) than the OSNAP-East curve in Fig. 4.13(c). The more rapid *QoI-independent* decay of the OSNAP-East curve is originated in the fact that the OSNAP-East prior-to-noise ratio is $(2.5)^2$ times lower than the OSNAP-West prior-to-noise ratio (independently of the value $(r/b)^2$, see [\(iii\)-ind](#)). The higher OSNAP-West prior-to-noise ratio indicates that OSNAP-West data is generally more relevant for the ECCOv4r3 inverse modeling framework, given the prior information, independently of changes in $(r/b)^2$.

4.4 Discussion

In this chapter, I presented a novel tool for ocean climate research to explore the proxy potential of the OSNAP array. The novel aspects are that proxy potential (i) targets unobserved or unobservable QoIs *remote* from the observations, and (ii) is purely dynamics-based. Exploring proxy potential is founded on a mathematically rigorous procedure, which is a growing research topic in the computational sciences: uncertainty quantification within an inverse modeling framework [e.g., [Flath et al., 2011](#); [Bui-Thanh et al., 2012](#); [Isaac et al., 2015](#)]. In Chapter 3, the formal notion of uncertainty quantification was translated to dynamical concepts and, in this chapter, applied to the OSNAP array within the state-of-the-art ECCOv4 inverse modeling framework [[Forget et al., 2015](#); [Fukumori et al., 2017](#)]. As a proof of concept, the proxy potential of the long-term (5-year) means of heat transport across OSNAP-West and OSNAP-East was explored for the long-term (5-year) means of two QoIs, Nordic Seas heat content and volume transport across the Iceland-Scotland ridge. Within the framework employed here, proxy potential explores dynamical relationships between observations and unobserved QoIs in the GCM equations, rather than using statistical methods. As a result, no actual OSNAP measurements were needed, which, to date, only exist for a period of 21 months [[Lozier et al., 2019](#)], i.e., a much shorter period than 5 years. As opposed to OSNAP measurement values, estimates for the *uncertainty* of OSNAP data are required for assessing OSNAP proxy potential. OSNAP uncertainty estimates were adopted from [Lozier et al.](#)

[2019].

Proxy potential was explored by investigating questions (I), (II), and (III). In the following, the answers to these three questions are summarized and discussed.

(I): What are the dynamical pathways, relative magnitudes, and mechanisms, by which OSNAP observations can be influenced?

A numerical adjoint was used to compute linearized sensitivities of time-mean heat transports across the OSNAP-West and OSNAP-East sections to all ECCOv4 control variables. The prior-weighted linearized sensitivities reveal all possible dynamical mechanisms and pathways via which small-amplitude perturbations in the ECCOv4 control variables, e.g., surface forcing over the globe, impact the observed OSNAP heat transports. On 5-year timescales, the dominant adjustment mechanisms were identified to be either thermally or wind-driven, forced both locally and remote from the OSNAP array.

Thermally driven mechanisms A dominant adjustment mechanism for the OSNAP heat transports on 5-year timescales is the advection of temperature anomalies. The temperature anomalies are forced by air-sea heat fluxes upstream of the OSNAP array, and subsequently carried to the OSNAP sections by surface currents, in particular by the cyclonic boundary currents of the subpolar gyre. Moreover, air-sea heat fluxes close to the western or eastern end of either OSNAP section change the upper-ocean density gradient across the respective OSNAP section, such that, by thermal wind balance, the geostrophic transport across the OSNAP section is altered. The dominant thermally driven adjustment mechanisms for the OSNAP quantities, i.e., advective processes and the thermal wind balance mechanism, are forced rather *locally*, i.e., close to the OSNAP sections.

Wind-driven mechanisms The dominant wind-driven adjustment mechanisms for OSNAP heat transports are pressure adjustment mechanisms, forced by wind anomalies *remote* from the OSNAP array. Most prominent is the effect of wind anomalies that occur along (i) the Icelandic coastline and/or (ii) the western European and African coast. Wind anomalies at these locations drive onshore or offshore Ekman transport, resulting in pressure anomalies. These pressure anomalies are communicated fast and efficiently by barotropic coastally trapped waves (i) around Iceland and/or (ii) along the eastern boundary of the Atlantic and affect the North Atlantic near-surface current system, leading to transport changes across both OSNAP sections. Prior-weighted sensitivities of both OSNAP heat transports are particularly high along the western coast of Iceland, since high sensitivity, due to pressure adjustment mechanism (i), is amplified by large prior ECCOv4r3 meridional wind stress uncertainties at this location. Prior wind stress uncertainties are large east of Greenland because wind speeds are high there, which increases the spread and thus, the uncertainty, in wind stress estimates in reanalysis products [Chaudhuri et al., 2012]. Importantly, (i) and (ii) are shared adjustment mechanisms and pathways for OSNAP-West and OSNAP-East (even though the same wind perturbations along the Icelandic coastline lead to responses of opposite sign in OSNAP-West and OSNAP-East heat transports). Besides the shared dominant mechanisms (i)

and (ii), some of the identified wind-driven processes are unique to either OSNAP-West or OSNAP-East, e.g., locally-forced Ekman transports.

Thermally vs. wind-driven mechanisms For changes in time-mean heat transport across OSNAP-West, the relative importance of thermally driven and wind-driven processes is 77% and 21%, respectively. The remaining 1% are controlled by precipitation and uncertain initial conditions and model parameters. For changes in time-mean heat transport across OSNAP-East, the relative importance of thermally-driven processes is 56%, of wind-driven processes 41%, and of remaining factors 3%. Here, relative importance of thermally (wind-) driven mechanisms is measured by integrating absolute prior-weighted thermal (wind) sensitivities throughout the globe. While thermally-driven processes are relatively most important for both OSNAP-West and OSNAP-East heat transports, wind-driven processes contribute to a higher degree to OSNAP-East than to OSNAP-West. Further investigation is required to pinpoint the underlying reason, but I suggest that the following two facts play a role. First, the identified non-local wind-driven pressure adjustment mechanisms (i) and (ii), shared between OSNAP-West and OSNAP-East, have their source and “center of action” closer in space to OSNAP-East than OSNAP-West. Until the signal reaches OSNAP-West via the Irminger Current, some of the signal’s strength might have been lost underway. Second, OSNAP-West might be relatively more sensitive to changes in thermal forcing than OSNAP-East because the OSNAP-West section cuts across the Labrador Sea, which features the deepest mixed layers in the Northern hemisphere. Therefore, thermal surface forcing that is applied in the vicinity of the OSNAP-West transect can impact the temperature of a deeper water column, and, thus, alter heat transport more effectively, than thermal surface forcing close to the OSNAP-East transect.

(II): To what extent are mechanisms and pathways identified in (I) shared with those influencing the unobserved QoIs?

The dominant adjustment mechanisms for the two chosen unobserved QoIs, time-mean Nordic Seas heat content (OHC_{NS}) and time-mean volume transport across the Iceland-Scotland ridge (T_{IS}), were identified by computing the linearized sensitivities of the QoIs to all ECCOv4 control variables - again, by means of the adjoint. As for the OSNAP quantities, the dominant adjustment mechanisms of the QoIs on 5-year timescales are either thermally or wind-driven. Projecting the prior-weighted sensitivities of the QoIs onto the prior-weighted sensitivities of the OSNAP quantities evaluates to what degree QoIs and OSNAP-West or OSNAP-East observations have shared adjustment mechanisms and pathways. A good projection means that the hypothetical proxy potential of noise-free OSNAP-West or OSNAP-East observations for the unobserved QoIs is high, which, if not considerably masked by noise, leads to a skillful proxy potential. To measure the hypothetical proxy potential of OSNAP-Combined, i.e., the dynamical constraints on the QoIs by *combining* the information extracted from OSNAP-West and OSNAP-East, one projects prior-weighted sensitivities of the QoIs onto modified sensitivity patterns of OSNAP-West and OSNAP-East. The modification of the sensitivity patterns is performed by a formal operation which removes data redundancy. For accurate projections, i.e., an accurate assessment of how (dis)similar dynamical adjustment processes for QoIs and OSNAP are, it was important to pertain the ECCOv4 control

space at its full spatial resolution, i.e., the resolution of the underlying GCM ($\sim 1^\circ$). In contrast, previous studies have “coarsened” the control space spatially via the ‘large region approach’ [Kaminski and Rayner, 2008; Kaminski et al., 2015, 2018], which could lead to large aggregation errors [Kaminski et al., 2001].

Projection of wind sensitivities Prior-weighted wind sensitivities of the unobserved QoIs and observed OSNAP quantities project extremely well onto each other. This is because the QoIs and the OSNAP quantities have *similar pressure adjustment mechanisms and pathways*, most prominently the mechanisms (i) and (ii), described above (while answering (I)) for the OSNAP heat transports. For all QoIs and OSNAP quantities, prior-weighted sensitivities are very strong west of Iceland, a fact which contributes particularly effectively to the goodness of the projections.

Projection of thermal sensitivities In contrast to the good projections of prior-weighted wind sensitivities, prior-weighted thermal sensitivities of the QoIs and the OSNAP quantities project only poorly onto each other. This is despite the fact that 5-year-mean heat transport across OSNAP-West and OSNAP-East, as well as the QoI OHC_{NS} , are to the largest degree controlled by thermally-driven adjustment mechanisms. However, the thermally-driven adjustment mechanisms, captured by the thermal sensitivities, reflect rather local processes that are not shared by the QoIs and the OSNAP observations. For instance, OHC_{NS} is to a large extent controlled by locally driven, one-dimensional air-sea processes. Due to the importance of thermally-driven *local* processes, sensitivities of QoIs and OSNAP quantities do to a large extent not overlap. A second reason for the poor projection of prior-weighted thermal sensitivities are *cancellation effects*. While overlapping sensitivities suggest that both the QoI and the OSNAP quantities respond to similar signals, the superposition of these signals can lead to “destructive interference”, i.e., cancellation in information, if sensitivities project with inconsistent signs. Non-overlapping sensitivities reflect that no information is transferred at all, whereas cancellations for overlapping sensitivities (destructive interference) means that information is in principle communicated from observations to QoI, but is not readily extractable, due to competing dynamical adjustments. Including either more observations or more prior information will help to achieve “constructive interference” so that the information contained in the observations can eventually be extracted for the QoIs.

Hypothetical proxy potential The hypothetical proxy potential measures how well prior-weighted sensitivities of the unobserved QoIs *to all ECCOv4r3 controls* project onto prior-weighted sensitivities of the observed OSNAP quantities *to all ECCOv4r3 controls*, on a scale from 0% (poor) to 100% (perfect). The hypothetical proxy potentials are reported in the last two columns of Table 4.5. The hypothetical proxy potentials for T_{IS} are 17% by OSNAP-West, 39% by OSNAP-East, and 40% by OSNAP-Combined. The hypothetical proxy potentials for OHC_{NS} are 3% by OSNAP-West, 13% by OSNAP-East, and 13% by OSNAP-Combined. The differences in the hypothetical proxy potentials can be explained by the following three facts:

- The hypothetical proxy potential, i.e., the goodness of the projections, is mainly determined by how effectively prior-weighted *wind* sensitivities project onto each

other. The reason is that many wind-driven adjustment mechanisms and pathways are non-local and shared between QoIs and OSNAP, while thermally driven adjustment mechanisms are local and lead to cancellation effects.

- For both QoIs, OSNAP-East has higher hypothetical proxy potential than OSNAP-West. Moreover, the hypothetical proxy potentials of OSNAP-Combined are not much higher than the ones of OSNAP-East, which means that OSNAP-West does not add much information to OSNAP-East about adjustment mechanisms for the QoIs. The underlying reason is the following: while both OSNAP-West and OSNAP-East show sensitivity to the non-local pressure adjustment mechanisms that are shared with the QoIs, OSNAP-East is relatively more sensitive to these wind-driven processes than OSNAP-West - a fact that was discussed above (in (I)).
- The hypothetical proxy potentials for T_{IS} are much higher than the ones for OHC_{NS} , for a similar reason as in the last bullet point: The relative importance of wind-driven processes for T_{IS} is much higher, namely 85%, as opposed to 42% for OHC_{NS} .

(III): To what extent is the information about shared mechanisms and signals in (II) masked by observational noise?

Since observations are noisy, the hypothetical proxy potential, i.e., the information obtained in (II) about shared dynamical mechanisms and signals between QoIs and OSNAP observations, will always be masked. The information transfer $\lambda/(\lambda + 1)$ regulate what fraction (between 0 and 1) of the OSNAP hypothetical proxy potentials can be extracted, and are independent of the considered QoI. The information transfer factors are determined by the prior-to-noise ratios

$$\lambda = \frac{\text{prior uncertainty in modelled observed quantity}}{\text{observational uncertainty (noise)}} \hat{=} \frac{1/\text{data relevance}}{1/\text{data quality}}. \quad (4.26)$$

of the OSNAP observations, which measure how much new reliable information the OSNAP observations can inject into the inverse modeling framework. Specifically, the prior-to-noise ratio λ of OSNAP-West (OSNAP-East) compares the *prior uncertainty* in the model counterpart of the OSNAP-West (OSNAP-East) observation with the *observational uncertainty* in the OSNAP-West (OSNAP-East) data (see equation (4.26)). In the limit of noise-free OSNAP observations, the prior-to-noise ratios λ of OSNAP-West and OSNAP-East would approach infinity and the associated information transfer factors would approach 1, such that 100% of the hypothetical proxy potential could be extracted.

Stronger constraint by OSNAP-West The prior-to-noise ratio of OSNAP-West is larger than the prior-to-noise ratio of OSNAP-East by about a factor of $(2.5)^2$. This means that observing time-mean heat transport across OSNAP-West exhibits a much stronger constraint on the ECCOv4r3 inverse modeling framework than observing time-mean heat transport across OSNAP-East. The factor-of- $(2.5)^2$ difference can be explained by the combination of two facts. First, the noise standard deviation for time-mean heat transport across OSNAP-East is 5 times higher than for time-mean heat transport across OSNAP-West [Lozier et al., 2019]. Second, in ECCOv4r3, the prior standard deviation

for OSNAP-East heat transport is about twice as high as for OSNAP-West heat transport, resulting in OSNAP-East data being about twice as “relevant” as OSNAP-West data. The prior standard deviations are computed by propagating prior information encapsulated in ECCOv4, e.g., information from reanalysis data sets (and their uncertainties), via the model equations throughout the global ocean. The underlying reason for the second fact is that, on average, OSNAP-East heat transport is twice as sensitive to changes in the uncertain ECCOv4r3 control variables as OSNAP-West heat transport. OSNAP-East *sensitivity* being larger than OSNAP-West sensitivity by a factor of (only) 2, stands in contrast to the OSNAP-East *signal* being larger than the OSNAP-West signal by about a factor of 5, both in ECCOv4 [Forget et al., 2015] and real observations [Lozier et al., 2019]. Therefore, if one were to follow the statistical (rather than the dynamical) approach, and substituted the prior-to-noise ratios by the signal-to-noise ratios, one would obtain similar (signal-to-noise) ratios for OSNAP-West and OSNAP-East, suggesting similar constraints by OSNAP-West and OSNAP-East.

Small prior-to-noise ratios The prior-to-noise ratios for OSNAP-West and OSNAP-East time-mean heat transports are 0.36 and 0.05, respectively. If the prior-to-noise ratios were equal to 1, the ECCOv4r3 prior uncertainties for the OSNAP model quantities would be exactly as high as the observational uncertainties (the noise), and the information transfer factors would take the value of 50%, reflecting a “fair compromise” between trusting prior model information and information injected by the observations. The fact that the OSNAP-West (OSNAP-East) prior-to-noise ratio is smaller than 1 means that the ECCOv4r3 prior uncertainty for time-mean heat transport across OSNAP-West (OSNAP-East) is smaller than the assumed noise in the corresponding OSNAP-West (OSNAP-East) data - in fact, the prior uncertainty is only 36% (5%) of the observational noise for OSNAP-West (OSNAP-East). The small prior-to-noise ratios λ lead to small values for $\lambda/(\lambda + 1)$: only 26% of the OSNAP-West hypothetical proxy potentials, and 5% of the OSNAP-East hypothetical proxy potentials can be extracted.

ECCOv4r3 prior uncertainties - too small? The small prior-to-noise ratios for the OSNAP quantities may be partly due to an underestimation of the prior uncertainties in ECCOv4r3. Indeed, examining the ECCOv4r3 prior uncertainties, which are specified for the ECCOv4r3 control variables, through the lens of the induced prior uncertainties for the observed OSNAP quantities and unobserved QoIs, leaves the impression that the prior uncertainties are chosen too small in ECCOv4r3 (cf. Figs. 4.2 and 4.3). Fig. C.1 in Appendix C shows that the small induced prior uncertainties are not due to the two aspects in which the setup in this work deviates from the ECCOv4r3 setup, i.e., the exclusion of the ECCOv4r3 time-variable atmospheric control variables and skipping Weaver and Courtier [2001] smoothing. The small ECCOv4r3 prior uncertainties for OSNAP heat transports - suggesting minor relevance of the OSNAP data for ECCOv4r3 - stand in contradiction to large adjustments to the control variables, and subsequent large adjustments in the OSNAP heat transport estimates, when including hydrography data of (a fraction of) the OSNAP moorings into the ASTER1 framework (personal communication with An Nguyen), a regional version of ECCOv4. In an inverse modeling framework, the specified prior uncertainties for the control variables reflect ranges within which the control variables are to be adjusted. Subsequently, the OSNAP heat trans-

ports should only be adjusted within the ranges of the ECCOv4r3 prior uncertainties for the OSNAP heat transports - but exceed these ranges by far!

ECCOv4r3 prior uncertainties - “lost in optimization”? I suggest that the reason for the contradiction described in the last paragraph is that, in practical ocean state estimation, one often starts with adjustments from - not necessarily documented - pre-optimization steps. For instance, ECCOv4 performs pre-optimization of initial conditions and parameter fields, constrained only by the temporal mean and seasonal variability of Argo data, to generate a relative steady state for the model, substantially reducing artificial drift [see Section 2a in [Smith and Heimbach, 2019](#)]. At the same time, pre-optimization steps entail that *effective* adjustments of control variables are of much bigger amplitude than the reported prior uncertainties. In particular, I suggest that the reported ECCOv4r3 prior standard deviations for initial conditions and parameter fields are too small, due to the described ECCOv4 pre-optimization step. Moreover, when performing pre-optimization steps, the specified prior covariance matrix for the subsequent full optimization problem (as the posterior covariance matrix after pre-optimization) should have non-diagonal entries (i.e., cross-correlations) that go beyond the cross-correlations that are mimiced by the [Weaver and Courtier \[2001\]](#) smoother. To strengthen the argument for *ad hoc* choices of prior uncertainties in ocean state estimation frameworks, it shall be noted that the values for the spatially constant prior uncertainties for the parameter fields of diapycnal diffusivity, isopycnal diffusivity, and GM intensity have been lowered by more than an order of magnitude from ECCOv4 release 2 to release 3 (personal communication with Ou Wang). Other *ad hoc* decisions during the optimization process may consist of temporarily “switching on” only a subset of the control variables, or “cranking” up and down weights of different control variables to make the optimization more sensitive to certain data types and control variables. To sum up, it is not always straight-forward to accurately translate weights from the regularization term in the ECCOv4 (or any other state estimate) optimization problem to covariance matrices in the probabilistic inverse problem. Uncertainties might get “lost in (pre-)optimization” - a problem that has to be resolved before state estimation frameworks can become practical tools for observing system design.

Varying the assumptions on prior uncertainties and/or OSNAP noise It is commonly known that the specification of accurate assumptions in inverse modeling frameworks is difficult [e.g. [Forget et al., 2015](#)] - not only of assumptions on prior uncertainties, as discussed in the last paragraph, but also on observational uncertainties (see Section 3.6). Given the novelty of using an oceanographic inverse modeling framework to explore proxy potential, some effects of these assumptions were elucidated in this chapter. Scaling noise and prior covariances matrices by constant factors, reflecting the assumptions of smaller/larger OSNAP observational uncertainties and/or prior control variable uncertainties, has the following implications:

- The hypothetical proxy potentials are unchanged, equal to the ones discussed in (II).
- The OSNAP-West prior-to-noise ratio is still $(2.5)^2$ times larger than the OSNAP-East prior-to-noise ratio, suggesting that OSNAP-West exhibits a $(2.5)^2$ times stronger constraint on the inverse modeling framework than OSNAP-East.

		A		B		C	
		Lozier et al. [2019] ECCOV4r3		Lozier et al. [2019] 4x ECCOV4r3		noise-free ECCOV4r3	
Obs	QoI	OHC _{NS}	T _{IS}	OHC _{NS}	T _{IS}	OHC _{NS}	T _{IS}
		OSNAP-West	1%	4%	2%	10%	3%
	OSNAP-East	1%	2%	2%	7%	13%	39%
	OSNAP-Combined	1%	6%	4%	14%	13%	40%

Table 4.5: Effective proxy potential of the three OSNAP observing systems for the QoIs Nordic Seas heat content (OHC_{NS}) and volume transport across the Iceland-Scotland ridge (T_{IS}). The effective proxy potential combines the answers to (II) and (III) and is formally measured as the relative uncertainty reduction for the respective QoI, when including the respective OSNAP observing system in the ECCOV4r3 inverse modeling framework. Scenarios A, B, and C show different assumptions on OSNAP data uncertainties and prior uncertainties in the ECCOV4r3 inverse modeling framework. In scenario B, “4x” denotes the multiplying factor for the ECCOV4r3 prior covariance, which corresponds to a doubling of the ECCOV4r3 prior standard deviations.

- The prior-to-noise ratios λ *do* change, by a constant factor. As a result, the associated information transfer factors $\lambda/(\lambda + 1)$ change, too.

These three facts, and in particular the latter, encourage the question: “How do ECCOV4r3 prior control variable uncertainties have to be changed so that model prior uncertainties are larger than OSNAP data uncertainties?”, or, the related question: “How accurate does OSNAP data have to be to exceed prior information in the model?”. If OSNAP standard deviations are assumed to be half as large as in Lozier et al. [2019] or the prior standard deviations are assumed to be double as large as in ECCOV4r3, the prior-to-noise ratios are 1.44 for OSNAP-West, and 0.21 for OSNAP-East. In other words, for these choices, OSNAP-West data would be qualitatively better than the prior information about its simulated counterpart in the model, while this is still not the case for the OSNAP-East data.

Combining the answers to (II) and (III): OSNAP’s proxy potential

The effective proxy potentials of OSNAP for the two QoIs are obtained by combining the answers to (II) and (III). Formally, this combination corresponds to assessing relative uncertainty reduction in the QoIs, when the OSNAP observations are included in the inverse modeling framework. In short, the answers to (II) and (III), discussed in the last few paragraphs, are:

- (II) hypothetical proxy potential: OSNAP-West < OSNAP-East (for each of the QoIs)
- (III) prior-to-noise (quality-to-relevance) ratio: $1 \gg \text{OSNAP-West} > \text{OSNAP-East}$ (independent of QoI)

For the hypothetical case of noise-free OSNAP observations, the answer to (III) can be ignored and the effective proxy potentials coincide with the hypothetical proxy potentials from (II) (Scenario C in Table 4.5). However, for more realistic (noisy) scenarios,

the hypothetical proxy potentials from (II) are masked, as discussed in (III). Adopting the estimates for prior uncertainties from ECCOv4r3 and for OSNAP data uncertainties from Lozier et al. [2019], the hypothetical proxy potentials are so strongly masked (by small information transfer factors discussed in (III)) that the effective proxy potentials are extremely small (Scenario A in Table 4.5): even OSNAP-Combined has an effective proxy potential of only 1% and 6% for OHC_{NS} and T_{IS} , respectively. Doubling the ECCOv4r3 standard deviations while still using the Lozier et al. [2019] estimates for the OSNAP data uncertainties, leads to less masking, as discussed in the last paragraph, and slightly higher effective proxy potentials (Scenario B in Table 4.5). For instance, OSNAP-West and OSNAP-Combined have effective proxy potentials of 10% and 14% for T_{IS} , respectively.

Importantly, while OSNAP-East has higher proxy potential than OSNAP-West for both QoIs in the hypothetical noise-free case (Scenario C in Table 4.5), it is the other way round for realistic, noisy scenarios (Scenarios A and B in Table 4.5). In the noise-free case, the answer to (II), i.e., "OSNAP-West < OSNAP-East", is the only determining factor. As discussed in (II), higher OSNAP-East dynamical proxy potentials are originated in the fact that the dynamical adjustment mechanisms for the QoIs are more similar to the dynamical adjustment mechanisms for OSNAP-East heat transport than the dynamical adjustment mechanisms for OSNAP-West heat transport. In Scenarios A and B, the answer to (III) becomes dominant, i.e., "OSNAP-West > OSNAP-East", caused by OSNAP-West's much higher prior-to-noise ratio, which indicates that OSNAP-West data exhibits a much larger constraint on the ECCOv4r3 inverse modeling framework than OSNAP-East data. While more work has to be done to ensure accurate prior-to-noise ratios, i.e., to remedy or justify " $1 \ll \text{OSNAP-West, OSNAP-East}$ ", the results here highlight the importance of (III), related to the question "Where to observe in order to detect the least noise-masked signals?".

The results in this work point out that, despite the vicinity of the two chosen QoIs to the OSNAP-East section and despite higher OSNAP-East hypothetical proxy potentials, the remote OSNAP-West observations are more effective in informing the QoIs. The recent study by Lozier et al. [2019] suggests a more important role of OSNAP-East compared to OSNAP-West, due to OSNAP-West's minor role for setting strength and variability of the OSNAP overturning during the first 21 months of the OSNAP record. However, the results of this chapter suggest that the much higher quality-to-relevance ratio of the OSNAP-West data leads to an (overall) much stronger constraint by OSNAP-West than OSNAP-East data. The stronger constraint by OSNAP-West can have wide-reaching implications - for climate signals away from OSNAP-West.

This work explored the proxy potential of time-mean OSNAP heat transports on a five-year timescale. The obtained low values for the effective OSNAP proxy potentials may be improved by a more comprehensive (future) study. Indeed, for assessing the full proxy potential of the OSNAP observing system, one should include the direct hydrographic and velocity measurements (i) from each individual OSNAP instrument, and (ii) at the measurement frequency. It is expected that extending this study by (i) and (ii), will increase the inferred effective OSNAP proxy potential - not only because it offers more (non-integrated) information about the OSNAP observations, but also because

uncertainties of the raw (non-interpolated) data products might be smaller. Moreover, combining the OSNAP observations with other observations, for instance Argo profiling floats, will further increase the proxy potential. In particular, there is much potential for improving the dynamical proxy potential: local Argo floats will help to constrain thermally driven processes forced in the vicinity of the QoIs - information that remote OSNAP observations were missing, even in the noise-free case. Meanwhile, the issue of small prior-to-noise ratios - and, thus, strong masking of hypothetical proxy potentials for any QoI - will persist, if the same ECCOv4r3 prior uncertainties are used as in this work. Prior uncertainties in state estimation frameworks, such as ECCOv4, will have to undergo a critical review - not an easy task - before state estimation frameworks can serve as a useful guiding tool for dynamics-based observing system design.

Limitations

Section 3.6.2 offers a detailed discussion on the limitations of the applied method for identifying proxy potential within an inverse modeling framework. The main limitations are (i) the dependence on the assumptions on prior uncertainties and noise, (ii) the linearity assumption of the adjoint, and (iii) model and system dependency. These limitations will be briefly reviewed now.

Assumptions on prior and noise By testing various choices for prior and observational uncertainties, this chapter highlighted that the accuracy of prior information and observational uncertainties has a direct influence on the identification of proxy potential within an inverse modeling framework. For instance, Fig. C.2 in Appendix C shows that including the Weaver and Courtier [2001] smoother in the inverse modeling framework, such as done in the full ECCOv4r3 setup, decreases dynamical proxy potentials, prior-to-noise ratios, and proxy potentials in Table 4.5, due to a diminished role of wind sensitivities. It is noted that an accurate estimation is not only difficult for prior uncertainties Γ_{prior} , as already discussed above, but also for observational uncertainties Γ_{noise} . Indeed, the OSNAP data uncertainties calculated in Lozier et al. [2019] could be an underestimation, because the estimates do not include uncertainties due to the OSNAP calculation method by which heat transports are inferred from the direct OSNAP observations (see supplementary material for Lozier et al. [2019]). A further limitation is that there could be a mismatch between the OSNAP array design in the real world, and how the array is represented in the model - based on discretization in the model world, but also due to the different calculation methods of heat transports across the array: OSNAP uses density coordinates [Lozier et al., 2019], whereas this work uses depth coordinates (clearly a limitation of this work, as already discussed in Section 2.5). The assimilation of OSNAP data into the ECCOv4 framework in future work will help to determine more reliable estimates for prior and observational uncertainties that are “in harmony” with each other (see Section 3.6.2). While proxy potential depends on the assumptions on Γ_{prior} and Γ_{noise} , it is noted that the specification of these matrices are a precise and explicit statement about all prior hypotheses.

Linearity and model (base state) dependency The identification of proxy potential relies on the linearity assumption and exactness of the adjoint (see Section 1.2). This assumption is evaluated against non-linear perturbation experiments in Chapter 5. It will be shown

that the adjoint-derived approximation of the response to non-linear ocean processes in the subpolar North Atlantic is flawed, especially for timescales exceeding 5 years. Therefore, this work investigated OSNAP proxy potential only on 5-year timescales. Moreover, it was argued that the chosen test period 2007-2011 is a representative 5-year time window and that the obtained results for the OSNAP proxy potential carry over to other 5-year time windows. This argument is valid if the linearized model equations do not strongly depend on the chosen base state 2007-2011. The dependence of linearized sensitivities on the base state is discussed by [Heimbach et al. \[2011\]](#), who conclude that the large-scale features of computed sensitivities are robust across different base states, and even across different model configurations, while amplitudes might slightly differ. The assumption of minor dependence on the model base state here is encouraged by the fact that 5-year mean quantities are studied. It is suggested that the limitation of model base state dependency is outweighed by the other mentioned limitations, but this remains to be tested in future work. The probably biggest shortcoming of the methodology of investigating proxy potential is that the results depend on the chosen model and inverse modeling framework - a problem that all methods for model-informed observing system design have in common.

4.5 Conclusions

The main conclusions of this chapter are:

- Heat transports across the OSNAP-West and OSNAP-East sections are sensitive to local and remote atmospheric forcing anomalies, with large contributions from *local air-sea heat flux* and *remote wind stress* anomalies.
- A dominant adjustment mechanism for the OSNAP heat transports on a five-year timescale is the advection of temperature anomalies. The temperature anomalies are forced by air-sea heat fluxes upstream of the OSNAP sections, and subsequently carried to the OSNAP array by the cyclonic boundary currents of the subpolar gyre.
- The dominant wind-driven adjustment mechanisms for the OSNAP heat transports are pressure adjustment mechanisms, forced by near-coastal wind anomalies remote from the OSNAP array. Most prominent is the effect of wind anomalies that occur along the Icelandic and Greenlandic coastline and the western European and African coast.
- The OSNAP observations share adjustment mechanisms and pathways with unobserved - or unobservable - quantities of interest (QoIs), for instance Nordic Seas heat content. Examples are the eastern boundary of the North Atlantic and the coasts of Iceland and Greenland, pathways via which wind-driven pressure anomalies are communicated around the entire subpolar North Atlantic and the Nordic Seas. These shared dynamical pathways and mechanisms imply that OSNAP has potential to inform the remote and unobservable QoIs. Uncertainty Quantification (UQ) within adjoint-based state estimation frameworks serves as a mathematically rigorous tool to assess this potential.

- In addition to the adjustment physics that is shared by the OSNAP observations and remote, unobserved QoIs, each QoI exhibits important, unique sensitivity patterns. For instance, Nordic Seas heat content shows strong sensitivity to local air-sea heat fluxes - processes that are not captured by the OSNAP observations. Argo floats, however, could provide these constraints, when combined with the OSNAP observations. UQ provides the mathematical framework to quantify data redundancy and complementarity of multiple observing systems, for instance the OSNAP mooring array and Argo profiling floats.
- Heat transport measurements across OSNAP-West impose an overall much stronger constraint on the estimated ocean state than heat transport measurements across OSNAP-East. This is largely due to the fact that climate signals detected by OSNAP-West are less noise-masked than climate signals detected by OSNAP-East. The stronger constraint by OSNAP-West has wide-reaching implications - for climate signals away from OSNAP-West, for instance in the Nordic Seas.
- Making accurate assumptions on prior and data uncertainties is a prerequisite for state estimation frameworks to become practical tools for observing system design. Ad-hoc choices and preoptimization steps in ocean state estimation hinder a straightforward translation from weights in the state estimation optimization problem to covariance matrices in the probabilistic inverse problem.

Chapter 5

Testing the Linearity Assumption of the Adjoint in the Subpolar North Atlantic

5.1 Introduction

The adjoint of an ocean GCM solves a set of equations that are linearized around the non-linear time-dependent ocean base state. Therefore adjoint-derived sensitivities, as presented in Chapter 2, are the *linearized approximation* for the non-linear responses of the studied quantities of interest (QoIs). Moreover, in Section 3.6.2, it was discussed that the assessment of observing systems and proxy potential via Hessian-based Uncertainty Quantification (UQ), as performed in Chapter 4, relies on the assumption that the observed and unobserved QoIs are dominated by only weakly non-linear model dynamics on the considered timescale. The linearity assumption, which Chapters 2 and 4 are based on, is expected to become less and less accurate with larger perturbations (Chapter 2) or larger prior uncertainties (Chapter 4) and longer model integrations. The acceptable size of the input perturbation and considered timescale will depend on the studied QoIs and the model setup. For instance, Czeschel et al. [2010] find that in their MITgcm configuration at 1° horizontal resolution, the adjoint and non-linear forward models show broadly consistent results for the evolution of the AMOC at 27°N over the first 15-20 years, after applying basin-wide heat flux perturbations of amplitude 15 W/m^2 over the subpolar gyre. The studies in Chapters 2 and Chapters 4 are performed within the ECCOV4 framework, which uses a MITgcm configuration at a nominal horizontal resolution of 1° , too. However, realistic air-sea flux anomalies (or prior uncertainties) are of higher amplitude than the ones tested by Czeschel et al. [2010]; for instance, the amplitude of expected surface heat flux anomalies in the Gulf Stream might be as large as 60 W/m^2 (exceeding 15 W/m^2 considerably). Moreover, one might expect more non-linearities for the QoIs studied in Chapters 2 and 4 - being located in the subpolar North Atlantic - than for the AMOC at subtropical latitudes (the QoI in Czeschel et al. [2010]).

In addition to being only the linearized approximation, a second problem is that the adjoint is *inexact*. Approximate adjoints have to be introduced for highly non-linear or discontinuous processes, e.g., due to regime transition, which cause instability in the adjoint model. Examples for approximate adjoints in the ECCOV4 adjoint modeling framework are a pseudo-sea-ice adjoint and the omission of the parametric dependency of the diffusivities and viscosities computed in the Gaspar et al. [1990] vertical mixing

parametrization in the adjoint. In contrast, the [Gent and McWilliams \[1990\]](#) and [Redi \[1982\]](#) components are retained in the adjoint. Note that the latter parameterizations were omitted from the adjoint of earlier ECCO releases [[Forget et al., 2015](#)].

Due to these limitations - the tangent linear approximation and inexactness of the adjoint - this chapter is devoted to “linearity checks”. Adjoint-derived sensitivities are evaluated against non-linear forward perturbation experiments, for perturbed surface forcing in the subpolar North Atlantic, the Nordic Seas and the Arctic Ocean. The skill of the adjoint to approximate non-linear responses is analyzed for the QoIs that were studied in [Chapter 2](#): Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge.

5.2 Linearity checks

5.2.1 Methodology

The computed adjoint sensitivities in [Chapter 2](#) provide an estimate of the response of the QoIs, UOHC_{NS} and T_{IS} , to surface buoyancy and momentum flux perturbations at any point in (model) space and time. Complete validation of the full space-time structure of the linearized response is computationally intractable, demanding an enormous ensemble of forward perturbation experiments (cf. [Fig. 1.9](#) in [Section 1.2](#)). Thus, for every atmospheric forcing type F , a few key regions are chosen, based on the findings in [Chapter 2](#). Here, the forcing type F is either (upward) surface heat flux Q_{net} , surface freshwater flux E-P-R, (eastward) zonal wind stress τ_x or (northward) meridional wind stress τ_y . The surface forcing F is perturbed in a small subdomain of each key region and during the season with the highest impact on the studied QoIs. All forcing perturbations are applied for exactly one month. The perturbation amplitudes reflect those of typical monthly forcing anomalies in the chosen month. The amplitudes are chosen equal to those of the temporally-spatially varying forcing anomalies ΔF which were used for weighting the adjoint sensitivities in [equation \(2.7\)](#) in [Chapter 2](#) (but averaged over the small subdomain). The forcing anomalies in [Chapter 2](#) were constructed as follows: For each grid point (x, y) , the anomaly $\Delta F(x, y, m)$ is a timeseries of 12 values, where for each month m , the anomaly is given by the standard deviation from the monthly mean ECCOv4r2 climatology.

Each linearity check consists of two separate perturbation experiments, with perturbations $\pm\Delta F$ to the chosen forcing and region. After integrating the non-linear GCM forward in time, the differences

$$\Delta_{\text{fwd}}^{\pm} := \text{QoI}^{\pm} - \text{QoI}$$

are computed. Here, QoI^+ and QoI^- denote the quantity of interest in the simulation with the positive and negative perturbation, respectively, and QoI the same quantity of interest in the model control run. The quantity of interest is either UOHC_{NS} , Nordic Seas upper-ocean heat content (defined in [equation \(2.1\)](#)) or T_{IS} , volume transport across the Iceland-Scotland ridge (defined in [equation \(2.2\)](#)). Recall that UOHC_{NS} in [Chapter 2](#) was defined as the mean upper-ocean temperature (rather than the actual upper-ocean

heat content) of the Nordic Seas, and units of UOHC_{NS} are K (rather than heat content units J). The model control run is the equilibrated simulation which was derived from the ECCOv4 solution, as described in Section 2.2. Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ are identical if the model response is fully linear. The forward anomalies Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ are then compared to the adjoint estimate

$$\Delta_{\text{adj}} := \frac{\partial \text{QoI}}{\partial F} \cdot \Delta F,$$

where ΔF denotes the positive forcing anomaly. Figs. 5.1 - 5.5 present the comparisons of Δ_{fwd}^+ (solid lines), $-\Delta_{\text{fwd}}^-$ (dashed lines), and Δ_{adj} (dots); the comparisons will be discussed below.

A forward perturbation experiment evaluates the *response* of the QoIs *at all future times* to a perturbation at a single point in time. In contrast, an adjoint experiment gives an estimate of the response of a QoI at a certain point in time to *perturbations at all previous times*. Due to this difference and for the sake of a more useful comparison between forward and adjoint experiments, for each of the two QoIs, a suite of 11 additional adjoint experiments is performed. In these additional adjoint experiments, the respective QoI is not evaluated in March (as in Chapter 2), but in January, February, April etc. In Figs. 5.1 - 5.5, the adjoint-derived responses, Δ_{adj} , are presented by dots. The dots which intersect the vertical lines in Figs. 5.1 - 5.5, corresponding to responses in March after integer (1, 2, etc.) years, are the adjoint-derived responses from the experiments in which the QoI was evaluated in March, i.e., the experiments which have been analyzed in Chapter 2. Some of the additional adjoint experiments have been run for less than 10 years of lead time, which is why for some of the yearly intervals, separated by two vertical lines, less than 12 dots are plotted.

5.2.2 Local heat flux perturbations

In Chapter 2, Nordic Seas upper-ocean heat content was found to be highly sensitive to local air-sea heat fluxes at short lead times. Therefore, the first linearity check is performed for the response of UOHC_{NS} to anomalous local heat flux. To this end, perturbations $\Delta Q_{\text{net}} = \pm 52.0 \text{ W/m}^2$ are applied to March surface heat flux inside the pink region in Fig. 5.1.

Fig. 5.1 shows the anomalies Δ_{fwd}^+ (solid orange line) and $-\Delta_{\text{fwd}}^-$ (dashed orange line) in UOHC_{NS} , respectively. The fact that Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ are almost identical indicates a high degree of linearity. Consistent with the suggested linearity, the adjoint-derived anomalies Δ_{adj} (orange dots) are in very good agreement with the forward anomalies. To conclude, the adjoint-derived estimate captures the response of UOHC_{NS} to local air-sea heat flux anomalies appropriately - and in particular the characteristic that the response of UOHC_{NS} is strongest at short time lags. Note that the relatively small magnitude of the responses in UOHC_{NS} in Fig. 5.1 is due to the small spatial (pink region) and temporal (1 month) extents of the imposed perturbations.

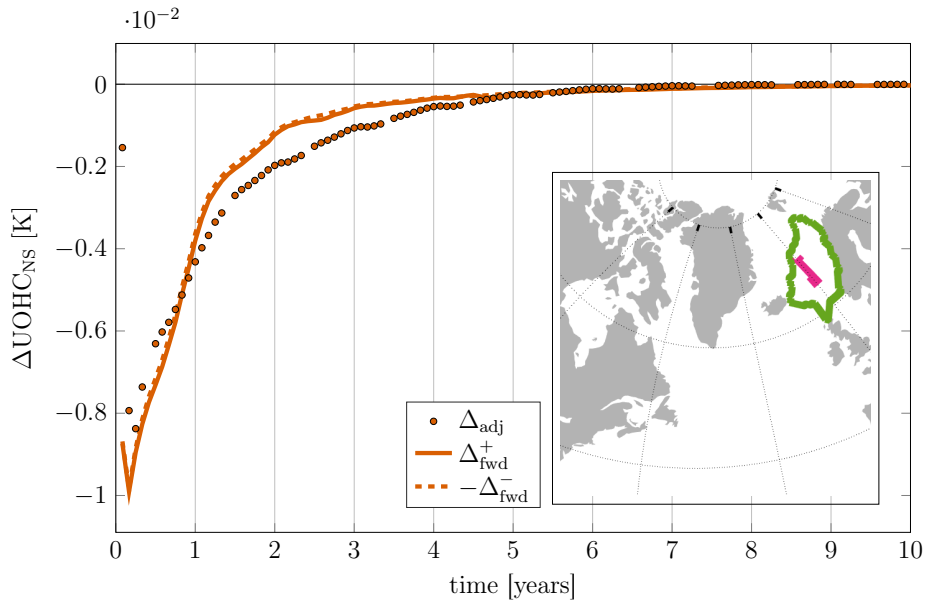


Figure 5.1: Anomalies in Nordic Seas upper-ocean heat content, UOHC_{NS} , caused by a prior **local heat flux perturbation** $\Delta Q_{\text{net}} = \pm 52.0 \text{ W/m}^2$, applied during the entire month **March** inside the pink region. The solid and dashed orange lines show the anomalies Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$, respectively, computed from forward perturbation experiments. The orange dots display the adjoint-derived anomaly Δ_{adj} . The green contour line in the map marks the boundary of the Nordic Seas region.

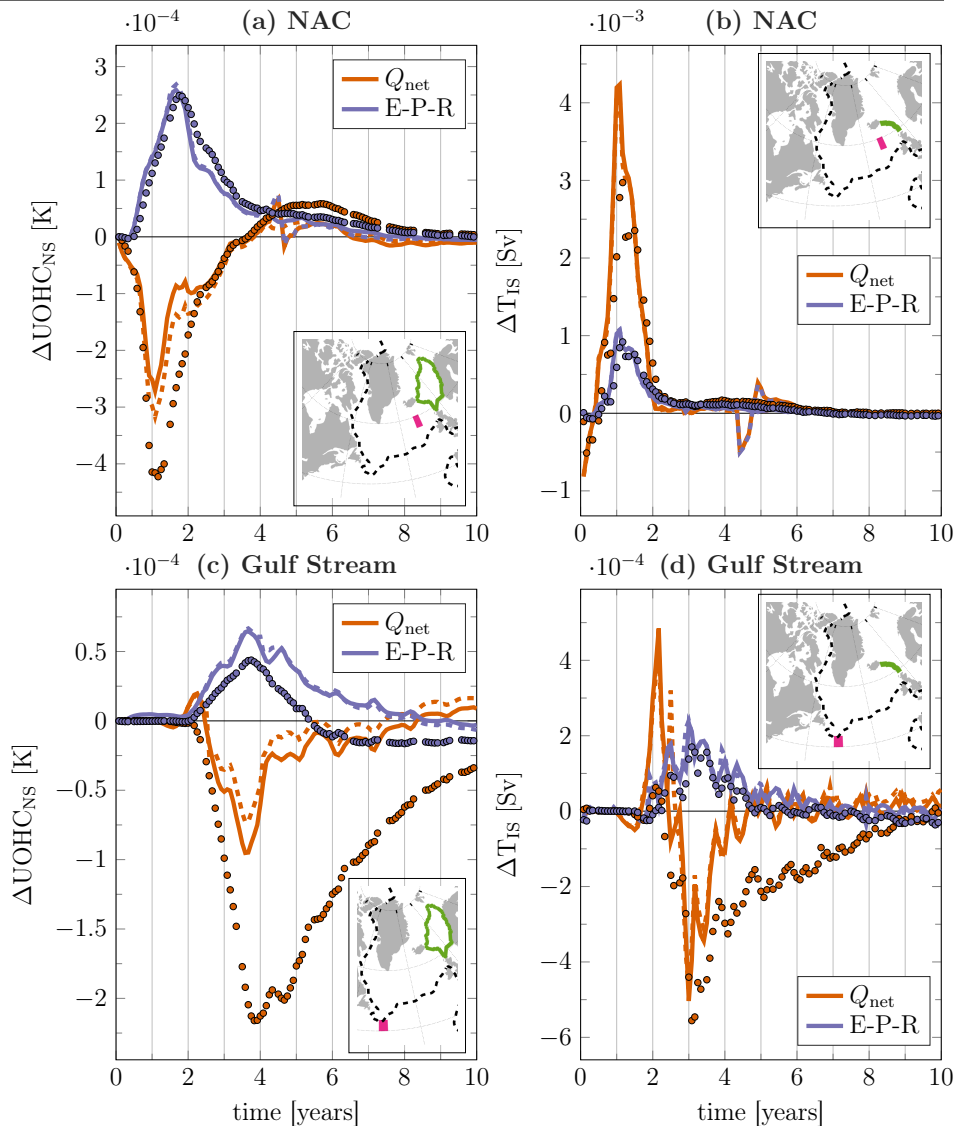


Figure 5.2: Anomalies in (a),(c) Nordic Seas upper-ocean heat content, UOHC_{NS} , and (b),(d) volume transport across the Iceland-Scotland ridge, T_{IS} . These anomalies are caused by prior perturbations in **surface heat flux**, Q_{net} , and **surface freshwater flux**, E-P-R, in a small region (pink box) within (a),(b) the **NAC** and (c),(d) the **Gulf Stream**. The amplitudes of the imposed anomalies are (a),(b) $\Delta Q_{\text{net}} = \pm 40.6 \text{ W/m}^2$ (orange lines) and $\Delta[\text{E-P-R}] = \pm 0.8 \text{ mm/day}$ (purple lines) in the NAC, and (c),(d) $\Delta Q_{\text{net}} = \pm 57.3 \text{ W/m}^2$ (orange lines) and $\Delta \text{E-P-R} = \pm 1.7 \text{ mm/day}$ (purple lines) in the Gulf Stream. The perturbations are applied during the entire month **March**. The solid and dashed lines show the anomalies Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$, respectively, computed from forward perturbation experiments, and the dots display the adjoint-derived anomalies Δ_{adj} . The green lines in the maps mark (a),(c) the boundary of the Nordic Seas region and (b),(d) the section of the Iceland-Scotland ridge. The dashed line in the maps shows the zero-line of the barotropic stream function.

5.2.3 Buoyancy flux perturbations in the NAC and Gulf Stream

In Chapter 2, the adjoint-derived sensitivities revealed that surface buoyancy flux anomalies in the NAC and Gulf Stream affect both UOHC_{NS} and T_{IS} , with the response time being set by the advective timescale. Since one might expect complex and non-linear behavior along the advective pathway of the separated jet, the next suite of linearity checks is performed for surface buoyancy flux perturbations in the NAC and Gulf Stream.

NAC

Two pairs of perturbation experiments are carried out: anomalies are applied to March (i) surface heat flux ($\Delta Q_{\text{net}} = \pm 40.6 \text{ W/m}^2$) and (ii) surface freshwater flux ($\Delta[\text{E-P-R}] = \pm 0.8 \text{ mm/day}$) inside a small subdomain of the NAC, indicated by the pink region in Figs. 5.2(a),(b). For both pairs of perturbation experiments, the response anomalies in UOHC_{NS} are shown in Fig. 5.2(a), while the response anomalies in T_{IS} for the same perturbation experiments are shown in Fig. 5.2(b).

For the heat flux perturbation, Δ_{fwd}^+ shows a rapid fluctuation after 4.5 years (solid orange lines in Figs. 5.2(a),(b)), while $-\Delta_{\text{fwd}}^-$ does not exhibit this fluctuation (dashed orange lines in Figs. 5.2(a),(b)). For the freshwater flux perturbation, $-\Delta_{\text{fwd}}^-$ (dashed purple lines in Figs. 5.2(a),(b)) displays a similar spontaneous fluctuation after 4.5 years, while Δ_{fwd}^+ (solid purple lines in Figs. 5.2(a),(b)) does not. The discrepancies between Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ and the occurrence of rapid fluctuations suggest that buoyancy perturbations of typical size in the chosen subdomain of the NAC can result in non-linearities in the response of UOHC_{NS} and T_{IS} at about 4.5 years later. On shorter timescales, a minor disagreement of Q_{net} -driven Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ in UOHC_{NS} at lags between 1 and 2.5 years (solid and dashed orange lines in Fig. 5.2(a)) is visible, suggesting only a weakly non-linear response of UOHC_{NS} at lags between 1 and 2.5 years after the imposed heat flux perturbation.

Consistent with the non-linearities suggested by the forward perturbation experiments, there is a minor deviation of the adjoint-derived anomaly $\Delta_{\text{adj}} \text{UOHC}_{\text{NS}}$ from the forward anomalies $\Delta_{\text{fwd}}^+ \text{UOHC}_{\text{NS}}$, $-\Delta_{\text{fwd}}^- \text{UOHC}_{\text{NS}}$ in response to the heat flux perturbation at time lags between 1 and 2.5 years, and after 4.5 years and onward (orange lines and dots in Fig. 5.2(a)). Apart from these minor deviations, the adjoint estimates Δ_{adj} for the anomalies in UOHC_{NS} (orange and purple dots in Fig. 5.2(a)) and for the anomalies in T_{IS} (orange and purple dots in Fig. 5.2(b)) show very good agreement with the anomalies Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ computed from the forward perturbation experiments.

In particular, the forward perturbation experiments confirm the timing and sign of the responses in UOHC_{NS} and T_{IS} that were identified by means of the adjoint-derived sensitivities in Figs. 2.3, 2.4, 2.5, and 2.6. Buoyancy forcing anomalies in the NAC show a sudden response in T_{IS} , when the signals are advected across the Iceland-Scotland section: for the perturbations in Fig. 5.2(b), the response in T_{IS} shows a sharp peak at about one year. After the signal is advected across the section, the impact on T_{IS} decays rapidly. In contrast, the response in UOHC_{NS} peaks at a lag of about 1.5 years after the imposed buoyancy flux perturbations in the NAC subdomain (Fig. 5.2(a)), delayed by about half a year compared to the peak response of T_{IS} . Moreover, the

response in UOHC_{NS} is characterized by a more gradual monotonic increase until its peak, followed by a more gradual monotonic decrease after its peak. As a result, a response signal in UOHC_{NS} is notable for many more years than for T_{IS} . The *delayed response* and the *longer impact* for UOHC_{NS} , as compared to T_{IS} , are due to the fact that UOHC_{NS} anomalies that are driven by upstream buoyancy forcing anomalies are to a large degree the integrated consequence of T_{IS} anomalies. This was discussed in Section 2.5. The forward perturbation experiments in Fig. 5.2(a) also confirm that the response of UOHC_{NS} to upward surface heat flux is of *opposite sign* as the response to E-P-R (orange vs. purple lines), as identified in the adjoint sensitivity patterns in Figs. 2.3(b),(c) vs. Figs. 2.4(b),(c) of Chapter 2. Similarly, the forward perturbation experiments in Fig. 5.2(b) verify that the response of T_{IS} to upward surface heat flux and E-P-R in the chosen NAC subdomain are of *equal sign* (orange vs. purple lines), as analyzed in Figs. 2.5(b),(c) and Fig. 2.6(b),(c) of Chapter 2. This difference was attributed to the fact that changes in temperature, advected by the mean current, affect UOHC_{NS} , but not T_{IS} (see Section 2.5).

Gulf Stream

Figs. 5.2(c),(d) show the response anomalies in UOHC_{NS} and T_{IS} for two pairs of perturbation experiments in the Gulf Stream. Forcing anomalies are added to March (i) surface heat flux ($\Delta Q_{\text{net}} = \pm 57.3 \text{ W/m}^2$) and (ii) surface freshwater flux ($\Delta[\text{E-P-R}] = \pm 1.7 \text{ mm/day}$) inside a small subdomain of the Gulf Stream, shown by the pink region in Figs. 5.2(c),(d).

The anomalies Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ associated with either of the quantities UOHC_{NS} and T_{IS} coincide for freshwater flux perturbations (solid vs. dashed purple lines in Figs. 5.2(c),(d)), suggesting a linear response in both quantities. In response to heat flux perturbations, a non-linearity occurs after 2.5 years, after which the curves Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ disagree for UOHC_{NS} (orange lines in Fig. 5.2(c)), but converge again for T_{IS} (orange lines in Fig. 5.2(d)). The suggested different degrees of non-linearity (i) of the two QoIs and (ii) in the two variables heat and freshwater flux is intuitive. As discussed in Sections 2.3 and 2.5, heat flux anomalies affect density directly, via ocean warming (cooling), but also indirectly, due to increased (decreased) evaporation and thus ocean salinification (freshening). The twofold dynamic effect is reflected in (ii), i.e., a more non-linear behavior of the QoIs in response to heat flux anomalies than to freshwater flux anomalies. As for (i), UOHC_{NS} can be altered by changes in cross-ridge volume transport T_{IS} and, in addition, by changes in the advected temperature. Therefore, UOHC_{NS} is a more non-linear function of remote heat flux anomalies than T_{IS} is.

Despite of the complex response in the forward perturbation experiments, the adjoint estimate Δ_{adj} agrees with Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ in response to freshwater flux in the Gulf Stream for T_{IS} (purple dots and lines in Fig. 5.2(d)), and to an acceptable degree for UOHC_{NS} up to time lags of 4.5 years (purple dots and lines in Fig. 5.2(c)). In response to heat flux, however, the adjoint estimates Δ_{adj} (orange dots in Figs. 5.2(c),(d)) deviate from Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$. The adjoint overestimates the amplitude of Q_{net} -driven UOHC_{NS} anomalies by about a factor of 2 (orange dots in Fig. 5.2(c)). For T_{IS} , the amplitude of the negative peak at a lag of 3 years is estimated correctly by the adjoint, but the amplitude

of the positive peak at a lag of 2 years is not (orange dots in Fig. 5.2(d)). Nevertheless, in Figs. 5.2(c),(d), both the *timing* of the maximum signal and the *general shape* of the response as a function of lag time is captured by the adjoint. As discussed for Figs. 5.2(a) vs. (b), the delayed response and longer impact for UOHC_{NS}, as compared to T_{IS}, is also visible in Fig. 5.2(c) vs. (d). Moreover, in Fig. 5.2(d), the sign change in ΔT_{IS} , occurring 2.5 years after the imposed heat flux anomaly in the Gulf Stream, is captured correctly by the adjoint (orange lines and dots). This sign change was discussed in Sections 2.3 and 2.5 and attributed to the twofold dynamic effect of heat flux anomalies on density, with generated density anomalies of opposite signs.

5.2.4 Wind stress perturbations along the intergyre boundary

Next, a suite of wind stress perturbation experiments is performed. For the perturbation locations, two subdomains of the NAC are chosen, situated north and south of the intergyre boundary. The intergyre boundary is defined as the zero-line of the barotropic stream function, as in Section 2.2, and is shown as the dashed black line in the map inlets in Fig. 5.3. The northern domain, shown as the orange area in the map inlets in Fig. 5.3, is referred to as “NAC-N”, and the southern domain, shown as the purple area, is referred to as “NAC-S”. In Section 2.3, sensitivity dipoles were detected, with sensitivities of opposite signs in NAC-N and NAC-S (and farther along the intergyre boundary). The dipole was visible for both UOHC_{NS} and T_{IS} sensitivities, and in sensitivities to both zonal and meridional wind stress (Figs. 2.7(c),(d), Figs. 2.8(c),(d), Fig. 2.9(e)). The forcing anomalies added to zonal wind stress are $\Delta\tau_x = \pm 0.09 \text{ N/m}^2$ in NAC-N and $\Delta\tau_x = \pm 0.07 \text{ N/m}^2$ in NAC-S. The forcing anomalies added to meridional wind stress are $\Delta\tau_y = \pm 0.06 \text{ N/m}^2$ in NAC-N and $\Delta\tau_y = \pm 0.05 \text{ N/m}^2$ in NAC-S. All perturbations are applied during the entire month March.

Figs. 5.3(a),(c) show the response of UOHC_{NS} to zonal and meridional wind stress perturbations, respectively, and Figs. 5.3(b),(d) display the response of T_{IS}. In each subpanel of Fig. 5.3, the curves Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ that correspond to the same perturbation (solid vs. dashed line of the same color) are almost identical. This suggests that the response in both QoIs to zonal and meridional wind stress perturbations along the intergyre boundary is almost linear. Therefore, one would expect the adjoint-derived anomalies Δ_{adj} to be reliable approximations of the forward anomalies. This expectation is proven wrong: the signal that arrives 2 to 4 years after the imposed perturbation is generally overestimated by the adjoint. That is, the absolute amplitude of Δ_{adj} is generally greater than the one of Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ in Figs. 5.3(a)-(d). The facts that Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ agree, whereas Δ_{adj} disagrees with Δ_{fwd}^+ , $-\Delta_{\text{fwd}}^-$, suggest that the error in the adjoint estimate arises from the *inexactness* of the adjoint, rather than non-linearity. This is further discussed in Section 5.3.

As for the surface buoyancy flux perturbations in the NAC and Gulf Stream (Section 5.2.3), the timing and the general shape of the UOHC_{NS} and T_{IS} responses in Fig. 5.3 are captured by the adjoint. An exception is the response in UOHC_{NS} to a zonal wind stress perturbation in NAC-S, for which the adjoint estimate Δ_{adj} (purple dots in Fig. 5.3(a)) deviates considerably after 4.5 years from Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ - not only in amplitude, but also in phase. The underlying cause requires further investi-

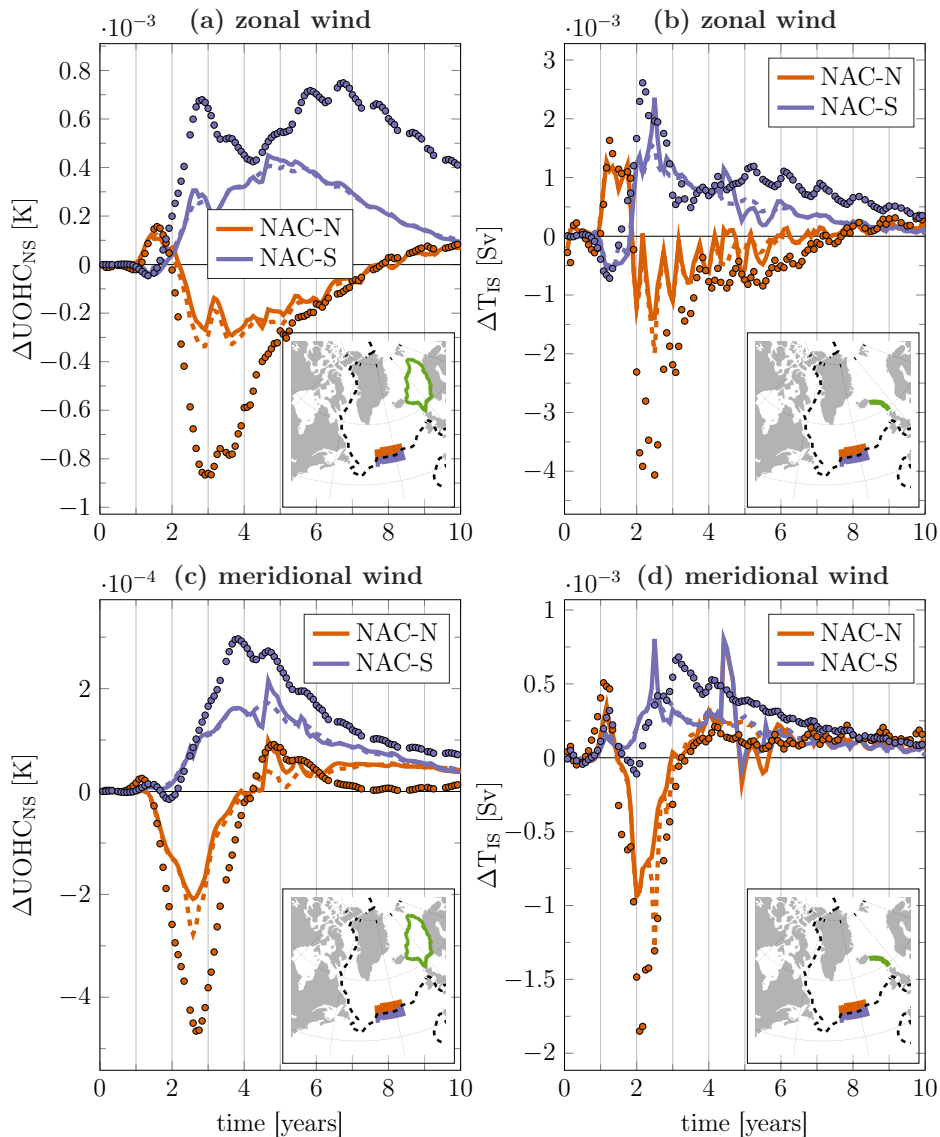


Figure 5.3: Anomalies in (a),(c) Nordic Seas upper-ocean heat content, UOHC_{NS} , and (b),(d) volume transport across the Iceland-Scotland ridge, T_{IS} , caused by prior **wind stress perturbations** inside the domains NAC-N (orange) and NAC-S (purple), located **north and south of the intergyre boundary**, respectively. The intergyre boundary, defined as the zero-line of the barotropic stream function, is shown by the thick dashed line in the maps. (a),(b) show anomalies due to zonal wind stress perturbations $\Delta\tau_x = \pm 0.09 \text{ N/m}^2$ in NAC-N (orange lines) and $\Delta\tau_x = \pm 0.07 \text{ N/m}^2$ in NAC-S (purple lines), and (c),(d) anomalies due to meridional wind stress perturbations $\Delta\tau_y = \pm 0.06 \text{ N/m}^2$ in NAC-N (orange lines) and $\Delta\tau_y = \pm 0.05 \text{ N/m}^2$ in NAC-S (purple lines). All perturbations are applied during the entire month **March**. The solid and dashed lines show anomalies computed from forward perturbation experiments, and the dots display the adjoint-derived anomalies.

gation. It is emphasized that, even though the adjoint estimate does not capture the response amplitudes accurately, it reproduces all sign shifts of Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ *exactly* in Figs. 5.3(a)-(d). In particular, Fig. 5.3 verifies the wind stress sensitivity dipole along the intergyre boundary, which was detected in Section 2.3. Moreover, the forward perturbation experiments confirm that upstream wind stress anomalies in the NAC and Gulf Stream play an at least as important role as upstream buoyancy flux anomalies. For instance, the wind-driven anomalies $\Delta\text{UOHC}_{\text{NS}}$ in Fig. 5.3(a) have comparable amplitudes to the buoyancy-driven anomalies $\Delta\text{UOHC}_{\text{NS}}$ in Fig. 5.2(a) and larger amplitudes than the buoyancy-driven anomalies $\Delta\text{UOHC}_{\text{NS}}$ in Fig. 5.2(c).

5.2.5 Perturbations in the subpolar gyre and Arctic Ocean

The final linearity checks are carried out in two regions where high non-linearities are expected: the subpolar gyre and sea-ice covered Arctic Ocean. The perturbation experiments are performed for freshwater flux because Fig. 2.12 revealed that freshwater flux perturbations are the forcing anomalies with highest impact - in the subpolar gyre for lag times greater than 5 years, and for the Arctic Ocean for any considered lag time.

Subpolar gyre

For the subpolar gyre, freshwater flux perturbations of $\Delta[\text{E-P-R}] = 1.4 \text{ mm/day}$ are imposed during March, when the mixed layer depth is at its deepest, and inside the pink subdomain shown in Fig. 5.4. Responses of UOHC_{NS} are shown in Fig. 5.4(a), and responses of T_{IS} in Fig. 5.4(b). The emerging rapid fluctuation in $-\Delta_{\text{fwd}}^-$ after 4.5 years for both UOHC_{NS} and T_{IS} (dashed orange lines in Figs. 5.4(a),(b)) hints at the occurrence of a strong non-linearity. From this fluctuation onward, Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ differ for UOHC_{NS} (solid vs. dashed orange line in Fig. 5.4(a)), but coincide for T_{IS} (solid vs. dashed orange line in Fig. 5.4(b)), except for another spontaneous fluctuation for T_{IS} at about 7 years (dashed orange line in Fig. 5.4(b)). The adjoint estimate Δ_{adj} is very accurate up to a response time of 4.5 years (orange dots in Figs. 5.4(a),(b)). After that, it approximates Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ for T_{IS} reasonably well (Fig. 5.4(b)), but diverges for UOHC_{NS} (Fig. 5.4(a)). The agreement of Δ_{fwd}^+ , $-\Delta_{\text{fwd}}^-$ and Δ_{adj} for T_{IS} , as opposed to their disagreement after 4.5 years for UOHC_{NS} , is due to an important difference between the two QoIs. If a non-linearity travels through the system, the adjoint estimate Δ_{adj} can regain skill for T_{IS} , after the non-linearity has passed. However, regaining skill is more difficult to achieve for the integrated quantity UOHC_{NS} , which accumulates the non-linearities.

Arctic Ocean

According to the results in Section 2.4, freshwater flux perturbations in the Arctic Ocean have the highest impact if they occur in summer. This is mainly due to the fact that perturbations in summer are of larger amplitude than in winter, due to high summer sea-ice variability. As opposed to all previous perturbation experiments, freshwater flux in the Arctic Ocean is therefore perturbed in *September*, inside the pink subdomain in Fig. 5.5. The chosen subdomain experiences large variations in freshwater flux (according to the ECCOv4r2 climatology), reflected in a large amplitude of $\Delta[\text{E-P-R}] = \pm 2.1 \text{ mm/day}$. In

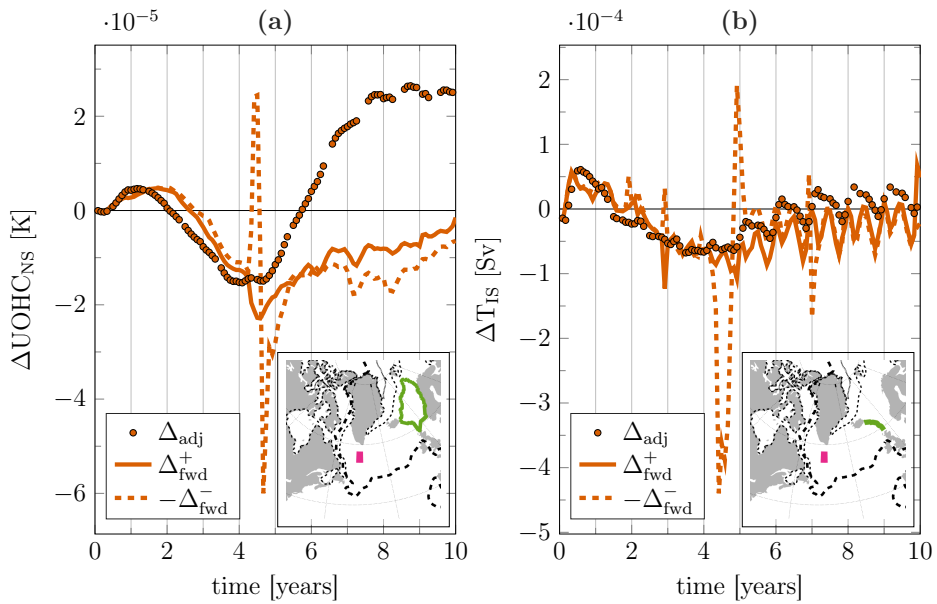


Figure 5.4: Anomalies in (a) Nordic Seas upper-ocean heat content, UOHC_{NS} , and (b) volume transport across the Iceland-Scotland ridge, T_{IS} , caused by prior **freshwater flux perturbations** $\Delta[\text{E-P-R}] = \pm 1.4 \text{ mm/day}$ in the **subpolar gyre**. The perturbations are applied during the entire month **March** inside the pink domain. The solid and dashed lines show anomalies computed from forward perturbation experiments, and the dots display the adjoint-derived anomalies. The green lines in the maps mark (a) the boundary of the Nordic Seas region and (b) the section of the Iceland-Scotland ridge. The black thick dashed line in the maps shows the zero-line of the barotropic stream function, and the black thin dashed and dotted lines mark March and September sea-ice extent, respectively.

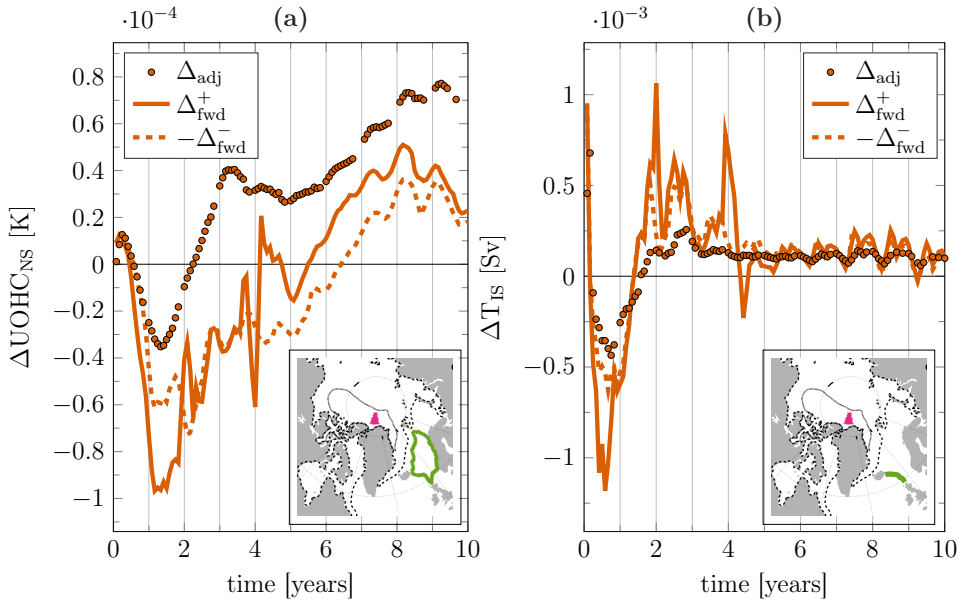


Figure 5.5: Anomalies in (a) Nordic Seas upper-ocean heat content, UOHC_{NS} , and (b) volume transport across the Iceland-Scotland ridge, T_{IS} , caused by prior **freshwater flux perturbations** $\Delta[\text{E-P-R}] = \pm 2.1 \text{ mm/day}$ inside the pink domain in the **Arctic Ocean**. The perturbations are applied during the entire month **September**. The solid and dashed lines show anomalies computed from forward perturbation experiments, and the dots display the adjoint-derived anomalies. The green lines in the maps mark (a) the boundary of the Nordic Seas region and (b) the section of the Iceland-Scotland ridge. The black dashed and dotted lines in the maps show March and September sea-ice extent, respectively. In September, the perturbation domain is partly covered in sea-ice.

the used control simulation, the subdomain is partly covered in September sea-ice.

Fig. 5.5(a) shows the response of UOHC_{NS} , and Fig. 5.5(b) the response of T_{IS} . The adjoint captures the general shape, seasonality, and sign changes of the response in T_{IS} , but underestimates the amplitude of the response (Fig. 5.5(b)). In contrast, the adjoint-derived anomaly $\Delta_{\text{adj}}T_{\text{IS}}$ is an excellent approximation for response times greater than 5 years, throughout the end of the perturbation experiments. For the response in UOHC_{NS} , the adjoint estimate Δ_{adj} follows the curves Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$ with a constant offset (Fig. 5.5(a)) - while an offset is already visible between Δ_{fwd}^+ and $-\Delta_{\text{fwd}}^-$. This constant offset is again attributed to the integrative nature of the quantity UOHC_{NS} : the abrupt non-linear response, occurring at a lag of 1 year, is not captured by the sea-ice adjoint and stays “in memory” of UOHC_{NS} .

It is noted that the forward perturbation experiments presented in Fig. 5.5 confirm the long memory of both T_{IS} and UOHC_{NS} to freshwater flux perturbations in the Arctic. Even though the amplitude of UOHC_{NS} anomalies for response times longer than 5 years is overestimated by the adjoint (Fig. 5.5(a)), the estimated total impact of freshwater flux perturbations in the Arctic on long timescales is not necessarily overestimated. Additional perturbation experiments (not shown) clear up that UOHC_{NS} anomaly amplitudes in response to freshwater perturbations at other locations in the Arctic Ocean can be either over- or underestimated by the adjoint. Due to non-linearity and the inexactness of the sea-ice adjoint, adjoint-derived sensitivities of UOHC_{NS} and T_{IS} to perturbations in the Arctic have to be interpreted carefully. Finally, it is noted that the opposite effect of freshwater flux perturbations in the Canadian and Eurasian Basins, detected in Fig. 2.4, is confirmed by forward perturbation experiments. Freshwater *removal* (corresponding to increased E-P-R) in the Canadian Basin leads to positive anomalies in Nordic Seas upper-ocean heat content after 5 years and later (Fig. 5.5(a)), whereas in the Eurasian basin it is freshwater *input* that causes positive anomalies in Nordic Seas upper-ocean heat content (forward perturbation experiments not shown).

5.3 Discussion

In this chapter, the skill of the adjoint to approximate non-linear responses of QoIs in the subpolar North Atlantic was put to the test. A series of different forward perturbation experiments was presented to provide a more thorough investigation than what has been presented in previous work. Moreover, to disentangle estimation errors by the adjoint (i) due to non-linear processes and (ii) due to inexactness of the adjoint, the linearity checks consisted of two steps. Step (i) is the comparison of anomalies $\Delta_{\text{fwd}}^+ \text{QoI}$ and $-\Delta_{\text{fwd}}^- \text{QoI}$, obtained from two forward perturbation experiments in which perturbations of opposite sign are imposed. This step gives an indication of whether strong non-linearities occur in the response of the QoIs. In step (ii), the forward perturbation experiments are compared to the adjoint estimate. Even if step (i) suggests a linear response, the adjoint estimate can still deviate from the forward anomalies, due to inexactness of the adjoint. Estimation errors due to inexactness of the adjoint are therefore uncovered by step (ii).

5.3.1 Non-linearity

The linearity checks were chosen to target regions which, based on the adjoint-derived sensitivities in Chapter 2, were identified as origins of variability in the subpolar North Atlantic QoIs. Importantly, the perturbations were chosen of expected amplitude, which notably exceed perturbation amplitudes from previously published linearity checks [e.g., Czeschel et al., 2010; Jones et al., 2018], and imposed in the most influential season, albeit only in small spatial domains. Moreover, for the choice of the perturbation domains, special attention was given to regions where non-linearities may be expected to play an important role, such as the Gulf Stream, NAC, subpolar gyre, and Arctic Ocean. As expected, step (i) identifies non-linear responses in the studied QoIs after imposed perturbations in the *subpolar gyre* and *Arctic Ocean*. In response to freshwater flux perturbations in the subpolar gyre, strong non-linear behavior in the QoIs is detected at a lag of about 4.5 years, possibly due to a convection event. In response to freshwater flux perturbations in a partly sea-ice covered subregion of the Arctic Ocean, strong non-linear behavior in the QoIs is visible already after a few months, possibly due to regime shifts related to sea-ice coverage. In contrast, no strong non-linearities are detected in response to surface buoyancy and momentum flux perturbations in the NAC and Gulf Stream regions. For these regions, the anomalies $\Delta_{\text{fwd}}^+ \text{QoI}$ and $-\Delta_{\text{fwd}}^- \text{QoI}$ in the non-linear forward perturbation experiments show remarkable agreement, despite complex fluctuations in the responses, visible in Figs. 5.2 and 5.3. The surprisingly good agreement, with occasional minor deviations in amplitude, suggests a mostly linear response of the QoIs to perturbed surface forcing in the NAC and Gulf Stream.

5.3.2 Inexactness of the adjoint

Despite the suggested linear behavior of the QoIs in response to perturbed surface forcing in the NAC and Gulf Stream, the adjoint-derived anomalies deviate from the forward anomalies in Figs. 5.2(c),(d) and 5.3(a)-(d). While the amplitudes of the adjoint-derived estimates are off, the sign, timing and general shape of the response is captured by the adjoint. The deviation identified by step (ii) is attributed to the inexactness of the adjoint. Pinpointing the cause for the found inexactness needs further investigation. The following aspects might play a role [cf. Forget et al., 2015].

- (a) The parametric dependency of the diffusivities and viscosities that are computed by the GGL [Gaspar et al., 1990] vertical mixing parametrization is excluded from the adjoint.
- (b) Even though the parametric dependency of the GM [Gent and McWilliams, 1990] and Redi [Redi, 1982] components are retained in the adjoint, inexactness might still occur due to the
 - use of simple clipping schemes for large isopycnal slopes;
 - omission of parametric dependency of isopycnal slopes on the ocean density field.
- (c) In the adjoint simulation, horizontal viscosity is increased by a factor of 2 (compared to the forward run), in order to reduce grid-scale noise in vertical velocity that destabilizes the adjoint.

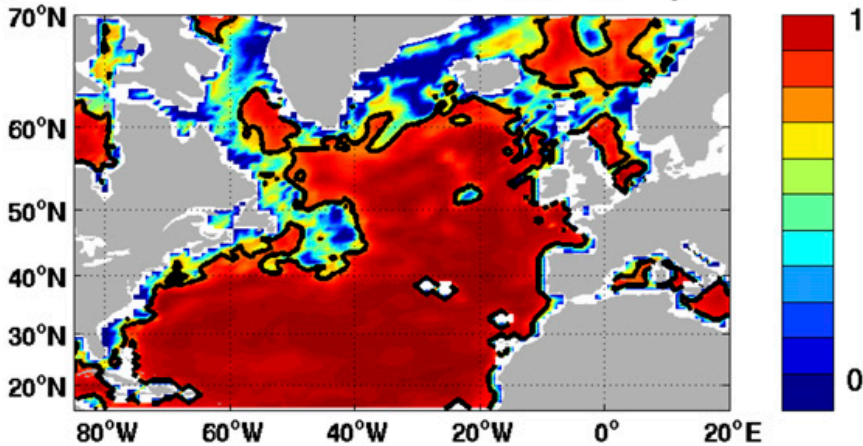


Figure 5.6: Different dynamical regimes of upper-ocean heat content variability, identified by Buckley et al. [2014, 2015] via budget calculations in ECCOv4. Blue shading indicates regions where diffusion and bolus transports are strong. The color shading corresponds to fraction of variance explained by the combined effect of (i) advective Ekman transport convergences, (ii) advective geostrophic transport convergences, and (iii) local air-sea heat fluxes. Figure from Buckley et al. [2015]. ©American Meteorological Society. Used with permission.

As part of a budget analysis for the ECCOv4 state estimate, Buckley et al. [2014, 2015] identify regions where diffusion (impacted by the GGL and Redi schemes) and eddy-driven bolus transports (impacted by the GM scheme) play a large role. These regions are characterized by blue shading in Fig. 5.6. The green-blue shading in the Gulf Stream, in particular in the Mann eddy region centered at (45°N, 40°W), suggests that aspects (a) and (b) could contribute to the identified inexactness related to the buoyancy flux perturbations in Figs. 5.2(c),(d). The perturbation regions in 5.3(a)-(d) are chosen farther downstream the Gulf Stream and NAC, along the intergyre boundary. Here, wind stress curl perturbations lead to Ekman up- or downwelling, and a subsequent shift of the main streamline of the NAC (cf. Fig. 2.13). The exact amplitudes involved in this mechanism might be corrupted by inexactness due to aspects (a) and (c).

5.3.3 Implications

Implied limitations for this thesis

The forward perturbation experiments in this chapter revealed the limitations of the adjoint-derived approximations that were used in Chapters 2 and 4.

Chapter 2 In Chapter 2, adjoint-derived sensitivities were used to disentangle local and remote origins and mechanisms that cause anomalies in the studied QoIs, Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge, on a ten-year timescale. The forward perturbation experiments in this chapter confirmed that the response to local heat flux perturbations and to upstream buoyancy flux per-

turbations that occur nearby are estimated accurately by the adjoint. However, the experiments also showed that for anomalous surface forcing farther upstream the NAC, as well as in the Gulf Stream, the amplitude of the response in the QoIs is not captured accurately by the adjoint, due to an inexact and incomplete adjoint. The QoI anomalies driven by these remote forcings are smaller in amplitude than the locally driven QoI anomalies (e.g., the anomaly amplitudes in Fig. 5.2(c),(d) are an order of magnitude smaller than the ones in Fig. 5.1). However, errors in anomalies of small amplitude could potentially lead to large errors, when accumulated over larger regions and longer times. For forcing anomalies in the subpolar gyre, the adjoint estimates the induced QoI anomalies accurately up to lags of about 5 years. Meanwhile, adjoint-derived anomalies in response to forcing applied in the Arctic Ocean have to be interpreted carefully on all timescales. Unreliable adjoint estimates in the Arctic Ocean are due to active Arctic sea-ice dynamics, which are omitted from current MITgcm adjoint simulations [Forget et al., 2015], due to persisting technical issues. However, the adjoint of the thermodynamic sea-ice model was successfully included in previous work [Heimbach et al., 2010; Fenty and Heimbach, 2012], which motivates future efforts to reinstate the sea-ice adjoint in the MITgcm adjoint modeling framework.

The forward perturbation experiments showed that the adjoint performs better for volume transport across the Iceland-Scotland ridge. In contrast, Nordic Seas upper-ocean heat content is not only a more non-linear quantity, but also an *integrated* quantity, and therefore accumulates the non-linearities, which is reflected in the adjoint estimate. All in all, the forward perturbation experiments confirmed that the adjoint *does* reproduce sign shifts and the general shape of the QoI response anomalies as a function of time. The adjoint is therefore useful to identify adjustment pathways and mechanisms, even if originated at longer lead times and in regions where the adjoint shows inexactness and, consequently, over- or underestimated amplitudes. Identifying adjustment mechanisms and pathways for Nordic Seas upper-ocean heat content was the purpose of Chapter 2. To get a full quantitative picture, the mechanisms detected by the adjoint-derived sensitivities should be further investigated in non-linear forward experiments, both with noneddying ocean-only models (as done here) and eddy-resolving ocean models as well as coupled climate models.

Chapter 4 In Chapter 4, adjoint integrations were performed for a 5-year time window, in order to minimize errors in the adjoint-derived approximation due to non-linear processes in the convective regions. Moreover, the QoIs were chosen as 5-year-mean quantities, where the time-averaging is expected to generate more linear QoIs. On the five-year timescale considered in Chapter 4, the dominant thermally driven mechanisms are forced locally or in the vicinity of the QoIs. These mechanisms are estimated correctly by the adjoint, as shown in this chapter. The dominant wind-driven mechanisms identified in Chapter 4 are either locally driven (by Ekman transports) or remotely driven. In the latter case, wind-driven pressure anomalies are communicated to the QoIs via fast coastally-trapped waves. Additional forward experiments will have to be performed to test the adjoint-derived estimate for anomalies due to the remotely driven pressure adjustment mechanisms. However, the adjoint is expected to estimate these fairly accurately, due to the short (daily to weekly) timescales on which the signals are communicated. It is noted that, despite the fast communication, the signal can have long-lasting impacts on the QoIs; for instance, on Nordic Seas heat content, due to the

ingrative nature of the quantity.

General implications

The ‘linearity assumption’, i.e., the fact that the adjoint provides only a linearized approximation to non-linear ocean physics, is generally viewed as the the greatest limitation of adjoint applications. The ‘linearity checks’ in this chapter showed that in some regions where one might expect the linearity assumption to be broken, such as in the tested Gulf Stream and NAC regions, the issue lies in the inexactness of the adjoint, rather than non-linearity. While some of the presented checks might look discouraging, the encouraging result is that there is potential for improvement, namely by a more exact and complete adjoint (whereas issues due to non-linearity could not be fixed). A detailed assesement of the accuracy of the adjoint in different regions, for different perturbed variables, and for different target quantities, as pursued in this chapter, is a first step in improving the exactness of the adjoint.

5.4 Conclusions

The skill of the adjoint to approximate non-linear responses in the underlying forward model depends on the studied quantities of interest (QoIs) and the model setup. The results from this chapter refer to the QoIs Nordic Seas upper-ocean heat content and volume transport across the Iceland-Scotland ridge, tested within the adjoint modeling framework ECCO version 4 release 2. The main conclusions are:

- The response to local heat flux perturbations and to upstream buoyancy flux perturbations that occur nearby are estimated accurately by the adjoint.
- Non-linear responses in the two QoIs are found after imposing perturbations in the subpolar gyre and Arctic Ocean. For forcing anomalies in the subpolar gyre, the adjoint estimates the induced anomalies in the QoIs accurately up to lags of about 5 years. Meanwhile, adjoint-derived anomalies in response to forcing applied in the Arctic Ocean are unreliable on all timescales, due to the omitted sea-ice adjoint in the current MITgcm adjoint simulations.
- Surprisingly, no strong non-linearities are detected in response to surface buoyancy and momentum flux perturbations in the Gulf Stream and North Atlantic Current. Here, it appears that the issue lies in the inexactness/incompleteness of the adjoint, rather than non-linearity.
- Where the adjoint shows inexactness, it still reproduces the sign, timing, and general shape of the QoI responses. The adjoint is therefore useful to identify adjustment pathways and mechanisms that influence QoIs in the subpolar North Atlantic. The amplitudes, however, may be over- or- underestimated by the adjoint.
- The adjoint performs better for volume transport across the Iceland-Scotland ridge than for Nordic Seas upper-ocean heat content. The latter is not only a more non-linear quantity, but also an *integrated* quantity, and therefore accumulates the errors in the approximation.

- Improving the exactness and completeness of the adjoint would be a valuable advancement toward dynamics-based ocean observing system design on timescales longer than 5 years.

Chapter 6

Summary and Discussion

In this thesis, the adjoint of a GCM was used to

- (i) investigate the underlying physical mechanisms and pathways of ocean circulation in the subpolar North Atlantic and the Nordic Seas (in Chapters 2 and 4);
- (ii) explore a novel approach to dynamics-based observing system design, with application to the subpolar North Atlantic (in Chapters 3 and 4).

Section 6.1 summarizes and discusses the key insights from (i) and Section 6.2 the key insights from (ii). The broader implications of the results are discussed in Section 6.3.

6.1 Mechanisms in the subpolar North Atlantic and Nordic Seas

Chapters 2 and 4 examined the adjoint-derived sensitivities of four subpolar North Atlantic ocean quantities of interest (QoIs, see Fig. 6.1) to atmospheric forcing anomalies. The sensitivities were used to disentangle dynamical drivers of ocean circulation in the subpolar North Atlantic and the Nordic Seas, as well as contributions from wind and buoyancy forcing, and time and space origins of these forcings. In Chapter 5, the adjoint-derived sensitivities were evaluated against non-linear forward perturbation experiments.

6.1.1 Thermally vs. wind-driven mechanisms

Controlling mechanisms of observed variability in the subpolar North Atlantic and the Nordic Seas are a function of timescale. Fig. 2.14 in Section 2.5 offers a summary of locally and remotely driven atmospheric forcing anomalies that generate Nordic Seas upper-ocean heat content anomalies up to 10 years later. While Chapter 2 elucidated space origins of forcing anomalies as a function of forcing lead time, Chapter 4 investigated the time-integrated picture and asked: “What is the effect of anomalies in the 5-year mean of atmospheric forcing (around the globe) on the 5-year mean of the QoIs?”. Fig. 6.1 reports the relative importance of globally integrated thermally vs. wind-driven processes for the studied 4 QoIs, as derived in Chapter 4.

The thermally driven mechanisms and pathways that govern changes in the four QoIs are summarized in Table 6.1. Moreover, the most dominant mechanisms and pathways are identified in Fig. 6.2(a), which shows the linearized sensitivity of Nordic Seas upper-ocean heat content to thermal forcing for a forcing lead time of 1 year. Nordic Seas heat

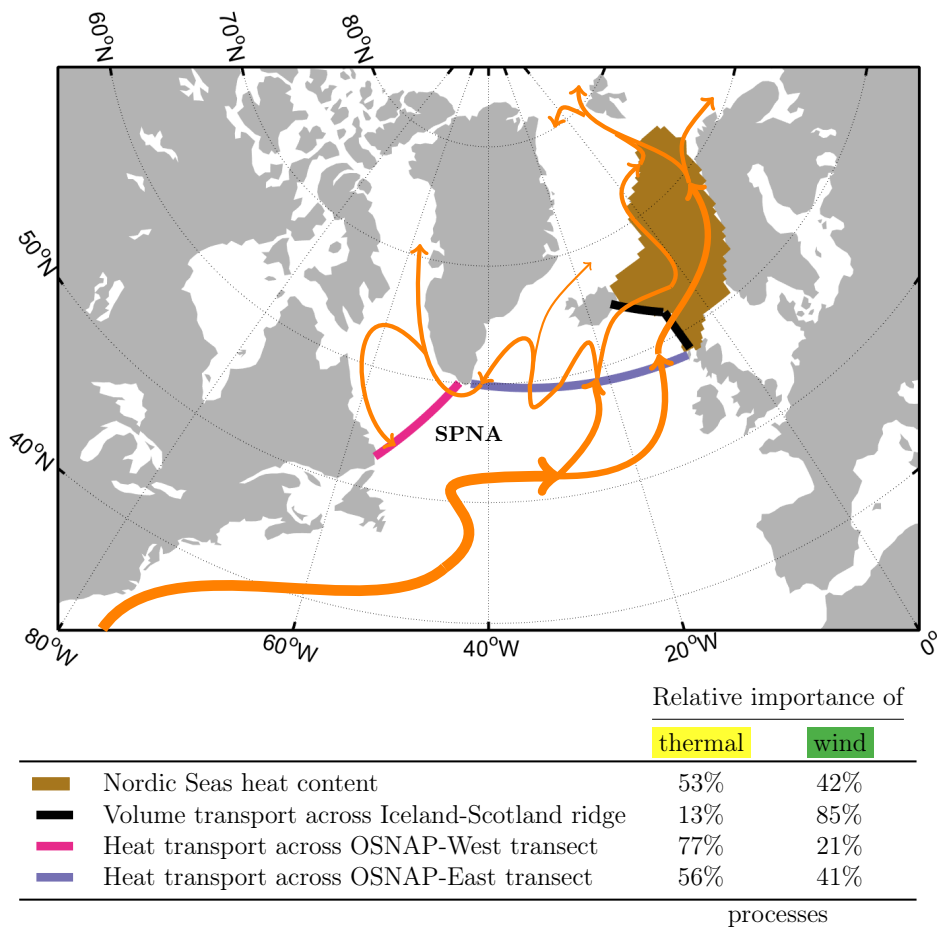


Figure 6.1: The four quantities of interest (QoIs, see legend) that were studied in Chapters 2 and 4, together with the major near-surface currents (orange arrows) in the sub-polar North Atlantic (SPNA) and the Nordic Seas. The last two columns of the table indicate the relative importance of globally integrated thermally driven (third column) and wind-driven (fourth column) processes for the time-means of the respective QoIs on a five-year timescale. The relative importance of thermally (wind-) driven processes is computed within the ECCOV4r3 state estimation framework, as the contribution to the total (prior) uncertainty in the time-mean QoIs by uncertainties in thermal (wind) forcing—equivalent to accumulating prior-weighted thermal (wind) sensitivities from around the globe (see Chapter 4). The remaining uncertainties stem from precipitation, uncertain initial conditions and parameters.

content is most sensitive to one-dimensional air-sea processes (1 in Table 6.1). Anomalies in heat transport across either of the two OSNAP transects is governed by the advection of temperature anomalies (2 in Table 6.1) that are forced upstream, in the vicinity of the transects. Since the dominant thermally driven mechanisms are forced locally or in the vicinity of the QoIs, the emerging adjustment pathways are a characteristic of each QoI (e.g., strong local sensitivities in the Nordic Seas are characteristic for the QoI Nordic Seas heat content, see Fig. 6.2(a)) - and to a large degree not shared among QoIs.

The wind-driven mechanisms and pathways that govern changes in the four QoIs are summarized in Table 6.2. As for thermally driven mechanisms, *locally forced* wind-driven processes play a role, via Ekman transports (1 in Table 6.2 and Fig. 6.2(b)). However, also *remote* wind anomalies can impact the QoIs efficiently: remote wind-driven pressure anomalies are communicated to the QoIs via coastally-trapped waves (2 in Table 6.2 and Fig. 6.2(b)). Pressure anomalies carried by the wave guides along the Icelandic coastline and the eastern boundary of the Atlantic (see Fig. 6.2(b)) have a wide-reaching effect on the North Atlantic current system and therefore impact all four QoIs. The fact that the OSNAP observing array (pink and blue in Fig. 6.1) sample signals that impact unobserved QoIs (such as brown and black in Fig. 6.1) downstream or upstream of the OSNAP array at a later or earlier time, gives rise to proxy potential, to be reviewed in Section 6.2.2 and Fig. 6.4.

Chapter 2 identified another remotely forced wind-driven mechanism that impacts all four studied QoIs, but operating on longer, namely advective, timescales. Wind stress curl perturbations along the intergyre boundary (3 in Table 6.2 and Fig. 6.3) can shift the North Atlantic surface current system, as suggested in Fig. 2.13 in Chapter 2. It is noted that the intergyre boundary is the location where wind stress curl anomalies associated with the NAO are the strongest [e.g., Marshall et al., 2001; Häkkinen et al., 2011]. Therefore, even though the signal that is triggered by mechanism 3 is found to be much smaller in amplitude than signals communicated by fast coastally-trapped waves (via 2), mechanism 3 deserves further investigation.

The adjoint-derived sensitivities that were used to detect the mechanisms in Tables 6.1 and 6.2 are only a linearized approximation for the non-linear responses of the studied quantities of interest (QoIs). Furthermore, the adjoint-derived sensitivities can be corrupted by inexact and incomplete adjoint models (see Chapter 5). Therefore, the adjoint-derived sensitivities were tested against non-linear forward perturbation experiments in Chapter 5. The perturbation experiments confirmed that the mechanisms listed in Tables 6.1 and 6.2 are drivers of variability in the QoIs, and that the timescales and pathways detected by the adjoint are correct. Moreover, mechanisms 1 and 2 are captured with correct amplitude (i.e., the color scale in Fig. 6.2(a) is correct). The response amplitudes driven by mechanisms along the advective pathway of the separated Gulf Stream, such as 3 can be inexact (i.e., the amplitude of the color scale in Fig. 6.3 is inexact), due to inexactness of the adjoint. However, timing and sign of mechanism 3 are captured correctly by the adjoint (i.e., light green lead time labels and red-blue shading in Fig. 6.3 are correct). For forcing anomalies in the subpolar gyre, the adjoint estimates the induced QoI anomalies accurately up to lags of about 5 years. Therefore, the magnitude of adjoint-derived QoI responses due to buoyancy-driven overturning changes

Table 6.1: **Thermally driven** mechanisms and pathways governing the subpolar North Atlantic QoIs in Fig. 6.1.

	Mechanism	Timescale	Origin/Pathway
1	one-dimensional air-sea processes	strongest: < 1 year	local (Fig. 6.2(a))
2	advection of temperature anomalies	advective timescale	upstream of QoI (Fig. 6.2(a)); only effective if forced nearby , due to SST damping
3	advection of mean temperature by anomalous geostrophic current (“thermal wind balance”)	advective timescale	local & upstream of QoI (Fig. 6.2(a))
4	overturning changes due to surface buoyancy forcing	> 5 years	convection regions

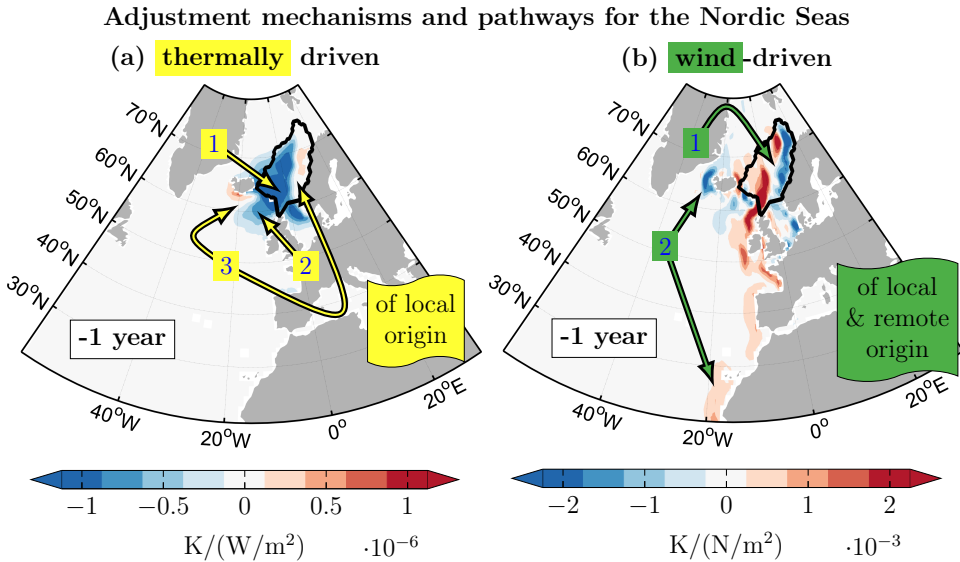




Figure 6.2: Linearized sensitivity of March Nordic Seas upper-ocean (0-275 m) temperature to (a) upward surface heat flux $Q_{\text{net}} \uparrow$ and (b) meridional wind stress τ_y for a forcing lead time of 1 year. Positive sensitivity indicates that heat loss to the atmosphere (in (a)), and increased northward wind (in (b)) causes higher Nordic Seas upper-ocean temperature 1 year later (see Chapter 2). While the governing thermally driven mechanism in (a) are mostly forced locally, the governing wind-driven mechanisms in (b) are forced both locally and remotely. The highlighted mechanisms **1**, **2**, **3**, **1**, **2** are described in Tables 6.1 and 6.2.

Table 6.2: **Wind-driven** mechanisms and pathways governing the subpolar North Atlantic QoIs in Fig. 6.1.

	Mechanism	Timescale	Origin/Pathway
1	Ekman transport	strongest: <1 year	local (Fig. 6.2(b))
2	pressure adjustment mechanisms (via coastally-trapped waves)	strongest: <1 year; → 5 years	remote ; emerging wave guides: Icelandic coastline and eastern boundary of Atlantic (Fig. 6.2(b))
3	Shifting of North Atlantic surface current system	advective timescale	Wind curl perturbations along intergyre boundary (Fig. 6.3)

(mechanism 4) has to be interpreted carefully because the estimate could be corrupted by strong non-linearities in the convective regions.

6.1.2 Sensitivity of ocean heat content vs. ocean transports

Anomalies in Nordic Seas heat content (OHC_{NS} , ) that are driven by upstream forcing anomalies (e.g., mechanisms 2, 3, 2, 3) are to a large degree the integrated consequence of anomalies in volume transport across the Iceland-Scotland ridge (T_{IS} , ). This link has been previously shown via budget analyses [e.g., [Árthun and Eldevik, 2016](#)] and is also visible in the forward perturbation experiments in Chapter 5, reflected by the following two facts:

- (a) The peak response in OHC_{NS} to upstream forcing anomalies is *delayed* as compared to the peak response in T_{IS} .
- (b) Upstream forcing anomalies have a *longer-lasting impact* on OHC_{NS} than on T_{IS} because signals are “stored in memory” of OHC_{NS} , while T_{IS} loses memory to signals that have traveled past the ridge.

Interestingly, in Chapter 2 it was found that the aspects (a) and (b) are reflected in the sensitivity maps by the following characteristics. As for (a), the sensitivities of OHC_{NS} and T_{IS} share many features, but these features emerge at shorter lead times for T_{IS} than for OHC_{NS} . For instance, sensitivity due to mechanism 3 emerges at about 70°W at a lead time of 2 years for heat transport across OSNAP-East (Fig. 6.3(a)), and at a lead time of 3 years for OHC_{NS} (Fig. 6.3(b)). (The time lag shown here, between OSNAP-East and OHC_{NS} , is even more pronounced than the time lag between T_{IS} and OHC_{NS} .) As for (b), the sensitivity patterns that reflect adjustment processes via coastally-trapped waves (i.e., patterns 2 in Fig. 6.2(b)) are only visible at short (daily to monthly) lead times for T_{IS} . For increasing lead time, T_{IS} is characterized by a rapid loss of sensitivity to wind forcing. In contrast, the sensitivity patterns 2 in Fig. 6.2(b) are imprinted as a stationary feature on OHC_{NS} , because, despite the short timescales associated with mechanism 2, the communicated signal can have long-lasting impacts on the OHC_{NS} , due to the integrative nature of the quantity. Consequently, OHC_{NS} is characterized by an only moderate loss of sensitivity to wind forcing with increasing lead time.

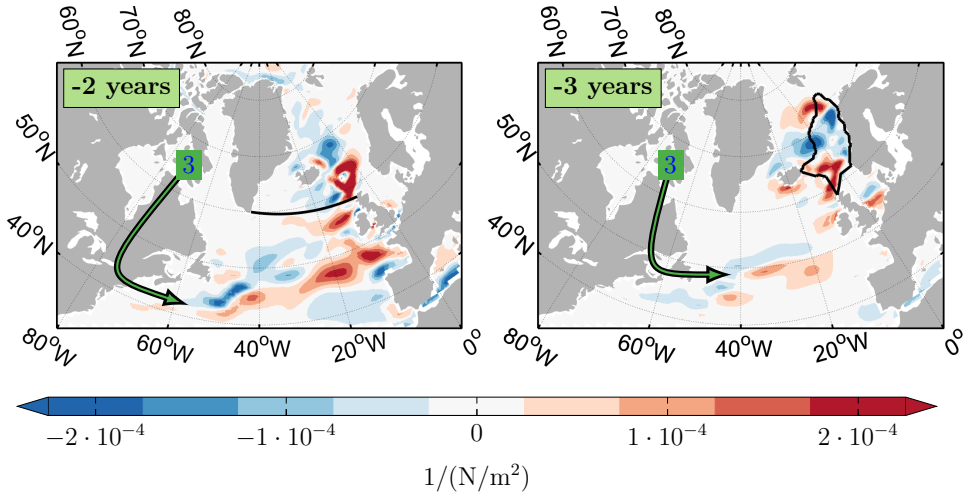


Figure 6.3: Normalized linearized sensitivity of December (a) heat transport across OSNAP-East ($MHT_{\text{OSNAP-E}}$, \blacksquare) and (b) Nordic Seas heat content (OHC_{NS} , \blacksquare) to zonal wind stress τ_x for a forcing lead time of (a) 2 years and (b) 3 years. The sensitivity dipole along the intergyre boundary (mechanism **3** in Table 6.2) indicates that a positive anomaly in wind stress curl along the intergyre boundary causes increased northward heat transport across OSNAP-East 2 years later (in (a)) and increased Nordic Seas heat content 3 years later (in (b)). Normalization is by the December 2011 values of (a) $MHT_{\text{OSNAP-E}}$ and (b) OHC_{NS} .

6.2 Dynamics-based assessment of observing systems

In Chapters 3 and 4, a novel approach to assessing ocean observing systems was presented. The innovative aspects of the approach are that it

- (1) is fundamentally dynamics-based, making use of dynamical relationships in a GCM;
- (2) is favorable for climate monitoring because the dynamical relationships and budgets are respected and treated consistently over long timescales (years to decades);
- (3) can assess data redundancy and complementarity in the context of other observations;
- (4) can identify the constraints of observing systems on climate signals and quantities of interest (QoIs) that
 - are unobservable or unobserved, different in type from the observations,
 - are not spatially collocated with the observations.

The explored approach has a rigorous mathematical foundation: Hessian-based uncertainty quantification (UQ) within an oceanographic inverse modeling (or state estimation) framework. UQ is a growing research topic in the computational sciences [e.g., Flath et al., 2011; Bui-Thanh et al., 2012; Isaac et al., 2015], but has not yet been applied to full-fledged oceanographic inverse problems on climate-relevant time windows.

6.2.1 Key insights

Assessing the constraints of an observing system on unobserved or unobservable climate signals, remote from the observations (aspect (4)), was rephrased as exploring the *proxy potential* of the observing system. The notion ‘proxy’ calls to mind paleoceanographic proxies, which are used to infer unobservable ocean circulation in the past. Here, proxy potential refers to unobserved (or unobservable) modern-day oceanographic QoIs, e.g., Nordic Seas heat content, and is explored via dynamical relationships in a GCM, rather than based on statistically derived transfer functions. Within the framework here, the identification of proxy potential does not require the actual measurement values of an observing system, only observation locations, times, types, and their uncertainties. In Chapter 3, the mathematical concepts of UQ and proxy potential were “translated” to dynamical concepts. In Chapter 4, the dynamical interpretation of UQ and proxy potential was illustrated for the recently installed OSNAP array [Lozier et al., 2017, 2019] within the state-of-the-art ECCOv4 inverse modeling framework [Forget et al., 2015; Fukumori et al., 2017]. In short, proxy potential of the OSNAP array arises

- for such QoIs that are forced by similar oceanic processes, via similar adjustment pathways, as the OSNAP observations;
- if the information contained in the OSNAP observations is not masked too strongly by observational noise.

The adjustment physics of observations and unobserved QoIs can be identified using the adjoint of a GCM, exactly as discussed in Section 6.1.

The translation “UQ concepts \rightarrow dynamical concepts”, as derived in Chapter 3, is summarized in Table 6.3 and will be briefly reviewed in the following, illustrated by the results from Chapter 4.

Information captured The information captured - and subsequently communicated - by an observing system is described by the eigenvectors and eigenvalues $\{(\mathbf{v}_i, \lambda_i)\}$ of the (non-dimensionalized) misfit Hessian (row A in Table 6.3). The eigenvectors \mathbf{v}_i are the *data-informed directions* within the control space of uncertain input variables, along which the model-data misfit term J_{misfit} in the (non-dimensionalized) cost function is curved; the eigenvalues λ_i describe the corresponding curvatures. The leading eigenvectors \mathbf{v}_i provide valuable insights for observing system design because they capture the most potent data constraints on the inverse modeling framework by the considered observing system.

To establish the link between data-constrained directions \mathbf{v}_i and dynamical concepts, Fig. 6.4(a) shows the information that is captured by the OSNAP-East observing system about local and remote meridional wind stress (τ_y) perturbations in ECCOv4. Specifically, Fig. 6.4(a) shows the (τ_y -subspace of the) eigenvector \mathbf{v}_1 , if the OSNAP-East observing system consisted only of measuring long-term mean heat transport across OSNAP-East ($\text{MHT}_{\text{OSNAP-E}}$). In this case, \mathbf{v}_1 computes as the normalized prior-weighted *sensitivity of the observed quantity* $\text{MHT}_{\text{OSNAP-E}}$ to all control variables,

Table 6.3: Summary of the established translation “UQ concepts \rightarrow dynamical concepts”

		UQ terminology [<i>equation in Chapter 3</i>]	Dynamical concept
A	Information captured by observing system	$\{(\mathbf{v}_i, \lambda_i)\}$: eigen-decomposition of misfit Hessian [(3.5)]	Fig. 6.4(b) & Fig. 6.5
B	Information required to constrain the QoI	\mathbf{q} : target direction [(3.10)]	Fig. 6.4(a)
C	Hypothetical proxy potential	$0 \leq \sum_{i=1}^M \langle \mathbf{q}, \mathbf{v}_i \rangle^2 \leq 1$ rel. uncertainty reduction in QoI for noise-free obs. [(3.14)]	How similar are adjustment physics of observations and unobserved QoIs? (Fig. 6.4)
D	Masking of information due to imperfect (noisy) observations	$0 < \frac{\lambda_i}{\lambda_{i+1}} < 1$ [(3.16)]	Masking by noise (Fig. 6.5)
E	Effective proxy potential	$0 \leq \sum_{i=1}^M \frac{\lambda_i}{\lambda_{i+1}} \langle \mathbf{q}, \mathbf{v}_i \rangle^2 < 1$: relative uncertainty reduction in QoI [(3.12)]	as hypothetical proxy potential, but masked by noise (Fig. 6.6)

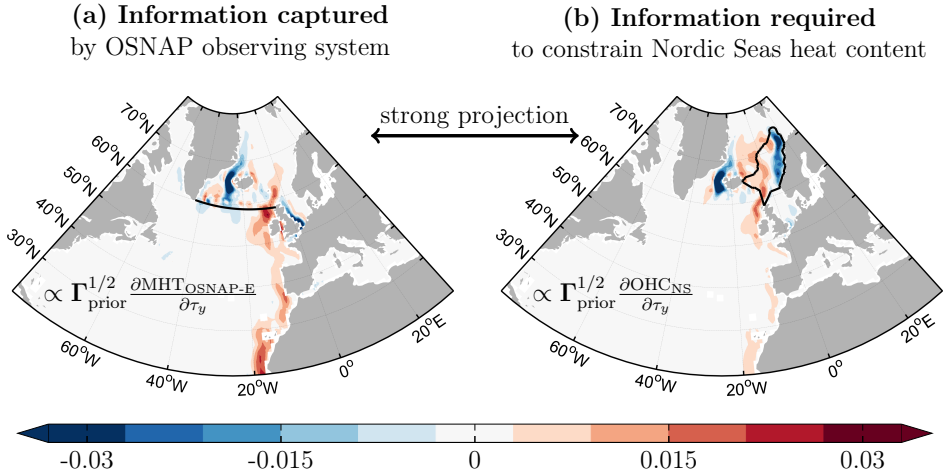


Figure 6.4: Normalized prior-weighted sensitivities of 5-year (2007–2011) mean (a) observed heat transport across the OSNAP-East section ($\text{MHT}_{\text{OSNAP-E}}$) and (b) unobserved Nordic Seas heat content (OHC_{NS}), to 5-year (2007–2011) mean meridional wind stress τ_y , computed within ECCOV4. Both quantities are strongly sensitive to changes in τ_y anywhere in the coastal wave guides along the eastern boundary of the North Atlantic and the Icelandic coastline. The shared adjustment mechanisms and pathways imply a strong projection of (a) and (b), and, thus, strong hypothetical τ_y -constraints of $\text{MHT}_{\text{OSNAP-E}}$ on OHC_{NS} .

capturing all possible dynamical adjustment mechanisms and pathways via which small-amplitude perturbations in the control variables can change the observed OSNAP quantity. Fig. 6.4(a) emphasizes that an observing system contains information on a much larger part of the ocean state than at the local measurements itself: OSNAP-East observations are the superposition of many different signals, e.g., wind stress perturbations in the coastal wave guides along the eastern boundary of the North Atlantic and the Icelandic coastline that are propagated to OSNAP-East by coastally-trapped waves (see also Section 6.1).

For adding more observations to the OSNAP-East observing system, e.g., freshwater and volume transports across OSNAP-East or the direct hydrographic and velocity measurements from the OSNAP-East moorings, the eigenvectors \mathbf{v}_i would be composed of the prior-weighted sensitivities of all observed OSNAP-East quantities (i.e., maps as in Fig. 6.4(a)), but modified by a formal operation which removes data redundancy, which is described in Chapter 3.

Information required The information required to constrain a given unobserved QoI is described by the QoI target direction \mathbf{q} (row B in Table 6.3). For example, Fig. 6.4(b) shows the (τ_y -subspace of the) target direction for unobserved long-term mean Nordic Seas heat content (OHC_{NS}), computed within ECCOv4. Just as the data-informed directions are composed of prior-weighted sensitivities of the observed quantities, the target direction is the normalized prior-weighted *sensitivity of the QoI* to all control variables. The prior-weighted τ_y sensitivities of OHC_{NS} in Fig. 6.4(b) elucidate that remote wind stress perturbations in the coastal wave guides along the eastern boundary of the North Atlantic and the Icelandic coastline affect Nordic Seas heat content, through signals transported by wave mechanisms (as discussed in Section 6.1). The dynamical interpretation of Fig. 6.4(b) is that these remotely driven wave mechanisms, together with all emerging locally driven mechanisms, are required to be informed by the observations.

Hypothetical proxy potential The constraints of an observing system on an unobserved QoI is formally assessed by how much uncertainty in the QoI gets reduced, when adding the observing system to the inverse modeling framework. If the observations were noise-free, the relative uncertainty reduction would be given by $0 \leq \sum_{i=1}^M \langle \mathbf{q}, \mathbf{v}_i \rangle^2 \leq 1$, which is referred to as the *hypothetical proxy potential* (row C in Table 6.3). The hypothetical proxy potential involves the projections of the target direction \mathbf{q} onto all eigenvectors \mathbf{v}_i . In the hypothetical case of noise-free observations, the degree of uncertainty reduction would therefore be solely determined by how well prior-weighted sensitivities of the observed quantities (to all control variables) project onto prior-weighted sensitivities of the unobserved QoIs (to all control variables), see Fig. 6.4. The goodness of the projections translates to the question: “How similar are adjustment mechanisms, pathways and relative magnitudes for the observing system and the unobserved QoIs?”, on a scale from 0% (no similarity) to 100% (identical).

The results from Chapter 4 show that, even though the degree of shared adjustment physics for $\text{MHT}_{\text{OSNAP-E}}$ and OHC_{NS} in Fig. 6.4 is high, the hypothetical proxy potential of the OSNAP observing system for the Nordic Seas within ECCOv4 is relatively

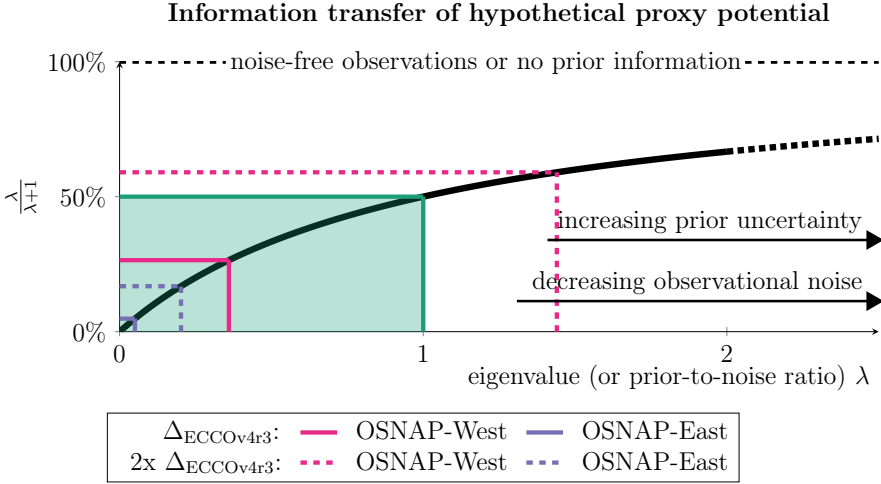


Figure 6.5: OSNAP-West and OSNAP-East eigenvalues λ (row A in Table 6.3) and associated information transfer factors $\lambda/(\lambda+1)$ (row D in Table 6.3) of the hypothetical proxy potential. The eigenvalues λ describe prior-to-noise ratios, computed as in (6.1), using ECCOV4r3 prior standard deviations (Δ_{ECCOV4r3} , solid lines) or doubled ECCOV4r3 prior standard deviations ($2x \Delta_{\text{ECCOV4r3}}$, dashed lines), together with the estimates by Lozier et al. [2019] for OSNAP data uncertainties. An observation that falls into the green rectangle has prior-to-noise ratio < 1 .

small (13%). This is because wind-driven processes contribute to only about half of the prior uncertainty in OHC_{NS} , while the other half is due to thermally driven processes (cf. Table 6.1). The most dominant thermally driven processes are forced locally (see Section 6.1 and Fig. 6.2(a)), and, thus the thermal sensitivities of OSNAP heat transport and Nordic Seas heat content project only poorly onto each other. In contrast, the hypothetical proxy potential for volume transport across the Iceland-Scotland ridge (T_{IS}) is much higher (40%) than for OHC_{NS} , because T_{IS} (as well its prior uncertainty) is dominated by wind-driven processes, while thermally driven processes are less important (cf. Table 6.1).

Masking of information Since observations are noisy, the hypothetical proxy potential, i.e., the information about shared adjustment physics between observing system and unobserved QoIs will always be partially masked. The fraction of the hypothetical proxy potential that is extractable is measured by the information transfer factors $0 < \lambda_i/(\lambda_i+1) < 1$ (row D in Table 6.3). These factors are independent of the considered QoI and are solely determined by the eigenvalues λ_i that characterize the observing system (row A in Table 6.3). The eigenvalue λ corresponding to an individual observation is equal to the *prior-to-noise ratio* (different from the statistical signal-to-noise ratio) of the observed quantity, i.e., the ratio

$$\lambda = \frac{\text{prior uncertainty in modelled observed quantity}}{\text{observational uncertainty (i.e., noise)}} \cong \frac{1/\text{data relevance}}{1/\text{data quality}}. \quad (6.1)$$

Fig. 6.5 shows the eigenvalues λ and associated factors $\lambda/(\lambda + 1)$ for the OSNAP-West and OSNAP-East observing systems, obtained in Chapter 4, using ECCOV4r3 prior uncertainties (for the numerator in (6.1)) and the estimates by Lozier et al. [2019] for OSNAP data uncertainties (for the denominator in (6.1)). The following facts stand out:

- For *noise-free* observations, the quality of the new data would be perfect ($\lambda \nearrow \infty$ in (6.1)). The associated information transfer factors $\lambda/(\lambda + 1)$ would be equal to 100% and the hypothetical proxy potential could be fully retrieved.
- The prior-to-noise ratio λ for OSNAP-West (solid pink line) is much larger than the prior-to-noise ratio λ for OSNAP-East (solid purple line). This means that OSNAP-West exhibits a much stronger constraint on the ECCOV4r3 inverse modeling framework than OSNAP-East.
- The prior-to-noise ratios λ for OSNAP-West and OSNAP-East are much smaller than 1 (purple and pink solid lines), i.e., the ECCOV4r3 prior uncertainty for time-mean heat transport across OSNAP-West (OSNAP-East) is smaller than the assumed noise in the corresponding OSNAP-West (OSNAP-East) data (see (6.1)). The small prior-to-noise ratios λ lead to small values for $\lambda/(\lambda + 1)$, i.e., major masking of the OSNAP hypothetical proxy potentials. The small prior-to-noise ratios for the observed OSNAP quantities may be partly due to an underestimation of the prior uncertainties in ECCOV4r3, which will be further discussed in Section 6.3.
- If prior standard deviations are assumed to be double as large as in ECCOV4r3 (“2x Δ_{ECCOV4r3} ”), the prior-to-noise ratio of OSNAP-West is larger than 1 (dashed pink line), i.e., OSNAP-West data would be qualitatively better than the prior information about its simulated counterpart in the model, while this is still not the case for OSNAP-East data (dashed purple line).

Effective proxy potential The effective proxy potential of an observing system for an unobserved QoI is formally measured by the relative uncertainty reduction $\tilde{\Delta}\sigma^2 \in [0, 1)$ in the QoI, when adding the observing system to the inverse modeling framework. Assessing a change in QoI uncertainty, rather than the change in the actual QoI estimate, explores the general *capability* or *potential* for the OSNAP observations to inform the QoIs - therefore the chosen terminology ‘proxy *potential*’. This assessment does not require the actual measurement values of the observing system, as explained in Section 3.4. Meanwhile, if the effective proxy potential is identified to be high, having the actual measurement values available will allow a better constrained *estimate* of the QoI. The effective proxy potential $\tilde{\Delta}\sigma^2$ is given by $\sum_{i=1}^M \frac{\lambda}{\lambda+1} \langle \mathbf{q}, \mathbf{v}_i \rangle^2$ (row E in Table 6.3), and combines hypothetical proxy potential (row C in Table 6.3) and masking by noise (row D in Table 6.3), i.e., the answers to the two questions:

- C) How similar are adjustment physics of observations and unobserved QoIs?
- D) What is the quality-to-relevance ratio of the observations?

Fig. 6.6 shows the combined answer to C) and D), summarizing the results from Chapter 4 for the effective OSNAP proxy potential. For the hypothetical case of noise-free

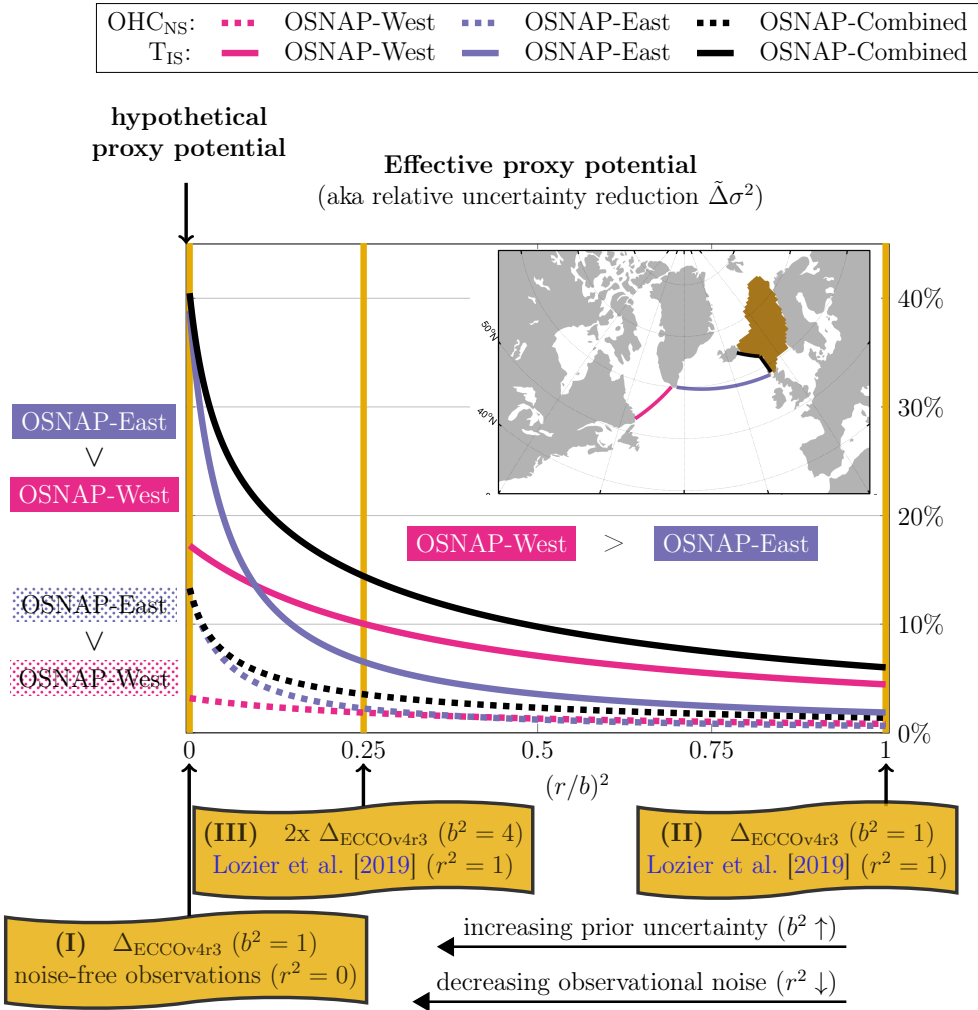


Figure 6.6: Effective proxy potential (aka relative uncertainty reduction $\tilde{\Delta}\sigma^2$) for the unobserved QoIs OHC_{NS} (dashed curves) and T_{IS} (solid curves) by the three observing systems OSNAP-West (pink), OSNAP-East (purple), and the combination OSNAP-West & OSNAP-East (black). Scenario (I) reflects the hypothetical case of noise-free OSNAP observations. Scenario (II) adopts ECCOV43 prior uncertainties and the estimates in Lozier et al. [2019] for the OSNAP data uncertainties. Scenario (III) reflects a doubling of the ECCOV43 standard deviations while still using OSNAP data uncertainties as estimated in Lozier et al. [2019]. In scenario (I), OSNAP-East is more informative than OSNAP-West for each of the QoIs. In scenarios (II) and (III), OSNAP-West is more informative than OSNAP-East.

OSNAP observations, the answer to D) can be ignored and the effective proxy potentials coincide with the hypothetical proxy potentials (Scenario (I) in Fig. 6.6). For each of the QoIs, the hypothetical proxy potential of OSNAP-East is higher than of OSNAP-West ("OSNAP-East > OSNAP-West"). However, for more realistic (noisy) scenarios ((II) and (III) in Fig. 6.6), the hypothetical proxy potentials are masked. The results of Chapter 4 show that for such realistic scenarios, the answer to D) becomes dominant, i.e., "OSNAP-West > OSNAP-East", caused by OSNAP-West's much higher prior-to-noise ratio, which indicates that OSNAP-West data exhibits a much larger constraint on the ECCOv4r3 inverse modeling framework than OSNAP-East data. Fig. 6.6 also shows that combining information from OSNAP-West and OSNAP-East exceeds information from each individual transect ("OSNAP-Combined > OSNAP-West, OSNAP-East"), but is smaller than the sum ("OSNAP-Combined < OSNAP-West + OSNAP-East"), due to data redundancy.

6.2.2 Full-fledged frameworks

Chapter 4 explored the proxy potential of long-term (5-year) mean observed OSNAP-West and OSNAP-East heat transports (gray elongated bars in Schematic 6.7(a)) - as well as of their combination OSNAP-Combined - for long-term (5-year) mean unobserved QoIs (blue elongated bars in Schematic 6.7(a)). For the sake of simplicity, the time-variable adjustments of the ECCOv4r3 atmospheric forcing fields (hatched controls in Schematic 6.7) were omitted from the control vector in Chapter 4. Instead, all remaining ECCOv4r3 control variables, i.e., the two-dimensional fields of time-mean atmospheric controls and the three-dimensional fields of initial conditions and parameter fields (collected in the first line of the legend in Fig. 6.7, cf. Table 4.2) were included. The results for the hypothetical and effective OSNAP proxy potentials obtained in Chapter 4 were *not* affected by the omission of the ECCOv4r3 time-variable atmospheric control variables (see Figs. C.2(a) vs. (b)), because long-term mean observed and unobserved QoIs were found to be relatively insensitive to the time-variable atmospheric control variables, demonstrated by the small fraction of hatched contributions in Fig. 6.8(a).

Schematic 6.7(b) shows a more comprehensive framework to evaluate the OSNAP (or any other) observing system, potentially in the context of other existing observations, e.g., all Argo profiling floats. Here, one takes into account the heterogeneous data streams of many - possibly millions of - observations (short gray bars in Schematic 6.7(b)) of different observation types, locations, and at different measurement frequencies, as well as their uncertainties. For instance, for the OSNAP mooring array, one could include the direct hydrographic and velocity measurements from each individual OSNAP instrument, at a monthly or higher frequency, rather than the derived OSNAP data products of heat transport across OSNAP-West and OSNAP-East. Long-term mean observations (gray elongated bar in Schematic 6.7(b)) are still relevant for the general framework. Indeed, ECCOv4r3 handles climatological data sets, e.g., the World Ocean Atlas [Antonov et al., 2010; Locarnini et al., 2010], in this fashion [Fukumori et al., 2017]. In the general framework, one could investigate the proxy potential of both long-term mean (blue elongated bars in Schematic 6.7(b)) and monthly/daily means of unobserved QoIs (blue short bars in Schematic 6.7(b)). When observed and unobserved quantities that are averaged over short time intervals enter the framework, it is important to include the time-variable at-

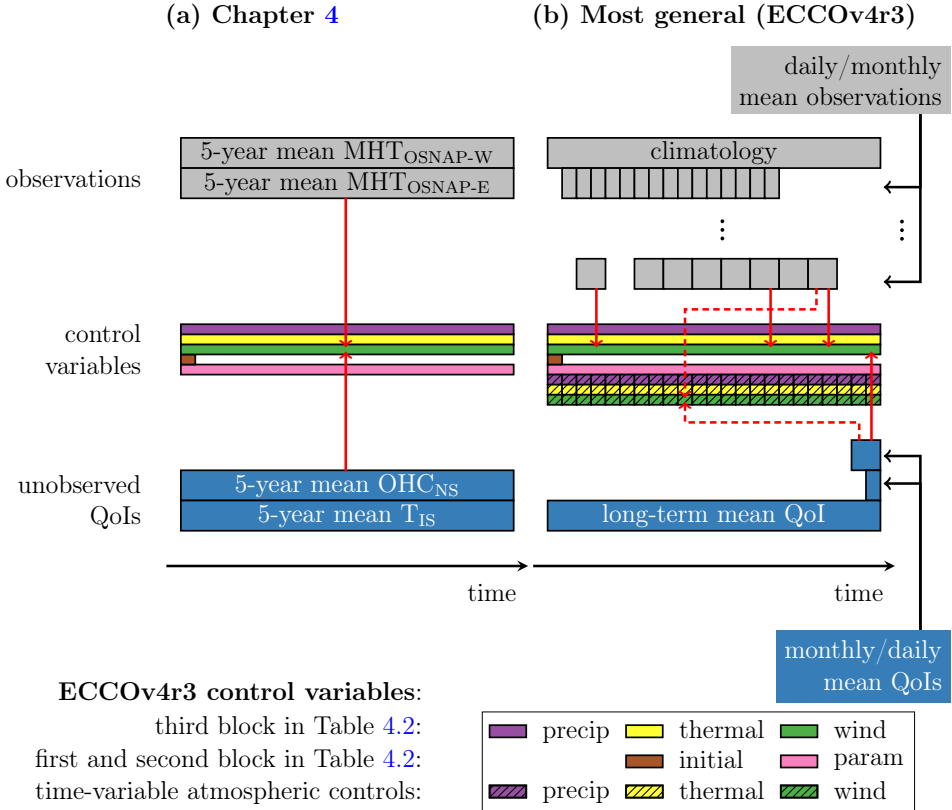


Figure 6.7: The architectures of the inverse modeling frameworks employed (a) in Chapter 4 and (b) in its most general form (ECCOv4r3). The red arrows sketch the computation of linearized sensitivities in order to explore proxy potential (see text).

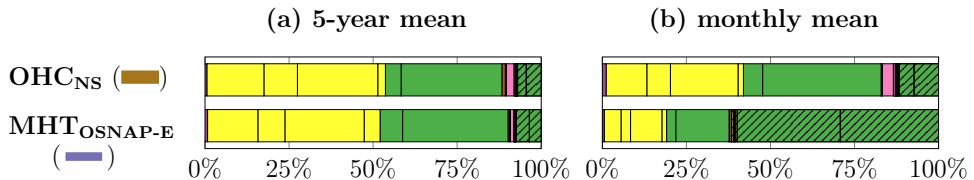


Figure 6.8: Relative contribution of the ECCOv4r3 control variables (legend as in Fig. 6.7) to prior uncertainty in (a) 5-year mean and (b) monthly mean QoIs, evaluated within the final 5 years (2007-2011) of the ECCOv4r2 state estimate. The QoIs in (a) are time-averaged over the full 5-year estimation period (corresponding to the blue elongated bar in Fig. 6.7(b)) and the QoIs in (b) are time-averaged over the *final* month December 2011 (corresponding to a blue short bar in Fig. 6.7(b)). The contribution by each control variable is computed by accumulating prior-weighted sensitivities of the respective QoI to the control variable from around the globe and, for the (hatched) time-variable control variables, additionally, from all times. The vertical lines separate the contributions by the different physical variables within each group, ordered according to Table 4.2. For instance, the leftmost (rightmost) fraction within the yellow bars is due to air temperature (shortwave downward radiation). The shown QoIs are Nordic Seas heat content (OHC_{NS} , upper bars) and heat transport across the OSNAP-East section ($\text{MHT}_{\text{OSNAP-E}}$, lower bars).

ospheric control variables (hatched in Schematic 6.7(b)). For instance, the lower bar in Fig. 6.8(b) shows that the monthly mean of heat transport across OSNAP-East is most sensitive to time-variable adjustments of wind forcing (green hatched contributions), consistent with the general understanding that, on short (intra-annual) timescales, ocean transport variability is to a large degree the response to wind forcing. In contrast, the upper bar in Fig. 6.8(b) indicates that Nordic Seas heat content - as an integrated quantity (see Section 6.1.2) - is also on short timescales more sensitive to changes in the *long-term mean* of atmospheric forcing, which suggests (proxy) potential for prediction. However, other unobserved climate signals and QoIs might be strongly sensitive to time-variable changes in atmospheric forcing. Therefore, time-variable atmospheric controls should be included in the general framework, in order to not neglect uncertainty to time-variable (past, present, and future) atmospheric forcing.

In Chapter 4, the information captured by the OSNAP observing system, i.e., the eigen-decomposition $\{(\mathbf{v}_i, \lambda_i)\}$ of the misfit Hessian (see row A in Table 6.3), was derived from pre-computed adjoint sensitivities of the observed quantities, where each observation required a separate adjoint sensitivity experiment. While this technique was feasible in Chapter 4, where only the two observations $\text{MHT}_{\text{OSNAP-W}}$ and $\text{MHT}_{\text{OSNAP-E}}$ were included (see Fig. 6.7(a)), it becomes intractable for a general framework as depicted in Fig. 6.7(b): the technique would require performing millions of adjoint sensitivity experiments, where each observed quantity would be the objective function of a new experiment. Performing millions of adjoint sensitivity experiments is replaced by extracting the leading eigenvectors $\{\mathbf{v}_i\}$ of the misfit Hessian via Lanczos or randomized SVD methods [e.g., Isaac et al., 2015; Kalmikov and Heimbach, 2014, 2018]. Nevertheless, even for the most complex observing system within Fig. 6.7(b), the method

developed in Chapter 3 and performed in Chapter 4 still gives the correct (though intractable) recipe to obtain $\{(\mathbf{v}_i, \lambda_i)\}$, i.e., the information captured by the considered observing system. Therefore, the insights about dynamics-based assessment of observing systems and proxy potential in this thesis, summarized in Section 6.2.2, carry over to the most general framework, i.e., the one sketched in Fig. 6.7(b). For example:

- The leading (or rather all) data-informed directions $\{\mathbf{v}_i\}$, characterizing the observing system, will be a composition, i.e., a linear combination, of prior-weighted linearized sensitivities of the observed quantities (to all controls). The way this composition is formed takes into account observation uncertainties, prior information, and data redundancy, and is explained in detail in Chapter 3.
- Hypothetical proxy potential is determined by how well prior-weighted sensitivities of the observed quantities project onto prior-weighted sensitivities of the unobserved QoIs. In Chapter 4, where long-term mean quantities were studied, the sensitivities to changes in the long-term mean of atmospheric control variables were the main player in the projections (cf. Fig. 6.4), sketched by the red solid arrows in Fig. 6.7(a). In the general framework, this idea is extended by two facts. First, control variables can be informed by observations from different times within a data stream, indicated by *multiple* red solid arrows pointing to one group of control variables in Fig. 6.7(b). Second, prior-weighted sensitivities to *time-variable* adjustments of the atmospheric forcing fields will be important for the projections, too. Translated to dynamical principles, this means that not only *pathways*, but also the exact *timescales* of shared adjustment mechanisms play a role. As part of the projections, one would also compare prior-weighted sensitivities of observed and unobserved quantities to changes of atmospheric forcing at *different lead times*, as indicated by the two dashed red arrows in Fig. 6.7(b). In practice, the comparison indicated by the two dashed red arrows could for instance be prior-weighted Fig. 6.3(a) vs. prior-weighted Fig. 6.3(b). The good projection of the wind sensitivities in Fig. 6.3, caused by OHC_{NS} “lagging behind” MHT_{OSNAP-E} in response (cf. Section 6.1.2), would imply a strong constraint on the control variable that corresponds to the (green hatched) box in Fig. 6.7(b) which is the target of the two red dashed arrows, with implications for prediction studies.

6.2.3 Limitations

For a detailed discussion of the limitations of dynamics-based assessment of observing systems and proxy potential via Hessian-based uncertainty quantification, the reader is referred to Section 3.6.2. The main limitations are

- (i) dependence upon assumptions on prior information and observational noise;
- (ii) the Gaussian or (“linearity”) approximation;
- (iii) model dependency.

(i) Prior information & observational noise Assumptions on prior information and observational uncertainties are explicit choices that are made within the inverse modeling framework, in the form of covariance error matrices for prior and noise. The accurate specifications of these matrices is a difficult task [e.g., Forget et al., 2015], and often involves ad

hoc choices. As part of this thesis, the dependency of observing system assessment and proxy potential upon prior and noise assumptions was investigated in more detail. For instance, in Fig. 6.6, different choices of $(r/b)^2$ reflect different relative weighting of prior uncertainties vs. observational uncertainties. An integral aspect of accurate assumptions on prior and noise is that prior uncertainties are “in harmony” with observational uncertainties, requiring a close dialogue between the modeling and the observational communities. For instance, if the prior uncertainties are specified unrealistically small, the employed framework deems the actual observations irrelevant, because the data seems much more uncertain than the information that is already known. The results in Chapter 4 suggest that the reported ECCOv4r3 prior uncertainties are unrealistically small. This will be further discussed in Section 6.3.2.

(ii) Gaussian/Linearity assumption Hessian-based uncertainty quantification is a Gaussian approximation of the posterior probability function for the estimated ocean state, given prior information, observations, and ocean physics. The accuracy of the Gaussian approximation depends on how well the linearized model about the cost function minimizer (the most likely ocean state) can reflect the full response of the non-linear system. If the oceanic observed quantities are dominated by strongly non-linear model dynamics on the considered timescales, the posterior probability function may be multimodal or, equivalently, the cost function may have several local minima. In this case, a global Gaussian approximation would be inappropriate. Chapter 5 was an effort to test the entailed linearity assumption in the subpolar North Atlantic. The results in Chapter 5 indicate that for 5-year mean quantities (as considered in Chapter 4), non-linearity is not an obstacle. Chapter 5 furthermore suggests that improving the exactness and completeness of the adjoint would be a valuable advancement toward dynamics-based ocean observing system design on longer timescales.

(iii) Model dependency Dynamics-based assessment of observing systems and proxy potential is achieved via the (linearized) model equations of the underlying GCM. Many principles, e.g., equations of motion representing basic conservation and constitutive laws, are robust across GCMs, but other features, such as model resolution and parametrization of sub-grid scale processes, are model-dependent. Therefore, the results may suffer from model dependency - a problem common to all methods for model-informed observing system design. To minimize the shortcoming of a single model approach, an intercomparison across multiple evaluation frameworks is therefore recommended.

6.3 Broader implications

Long-term and sustained ocean observations are needed to advance our understanding of key mechanisms and ocean variability, to better constrain uncertain processes in models, and to contribute to climate predictions (e.g., [National Academies of Sciences, Engineering, and Medicine, 2017](#) and Section 1.1.6). However, observing systems are extremely expensive to build and maintain, and often rely on short-term funding periods, including the OSNAP array. The assets (1) - (4), listed at the beginning of Section 6.2, qualify Hessian-based uncertainty quantification as a promising guiding tool for

building and sustaining a cost-effective, long-term, and dynamics-based Atlantic observing system, while maximizing the information extracted from the observations. While computational progress is underway for putting the tool into practice for oceanographic applications [Kalmikov and Heimbach, 2014, 2018; Moore et al., 2011a,b, 2017a], this thesis aimed at advancing the dynamical understanding of the tool. The implications of the gained dynamical insights for dynamics-based observing system design are discussed in Section 6.3.1. Moreover, given the novelty of the approach for oceanographic applications, Chapters 3 and 4 also tested aspects of the machinery within the state-of-the-art state estimation framework ECCOv4r3. Section 6.3.2 discusses the insights on current shortcomings that have to be overcome before ocean state estimation frameworks can become practical tools for formal dynamics-based observing system design. Moreover, suggestions for future avenues are made to tackle the identified challenges.

6.3.1 Towards the design of an optimized Atlantic observing system

Remotely observed signals Many of the dynamical pathways and mechanisms by which ocean observations are influenced, have also an important impact on unobserved QoIs. The shared adjustment physics becomes evident when comparing sensitivities of observed and unobserved quantities. The sensitivity patterns often show basin-mode type structure or concentration along the coastal boundaries, where the latter serve as an efficient communicator for signals via wave physics. For instance, Chapter 4 identified the eastern boundary of the North Atlantic and the coasts of Iceland and Greenland as important pathways and wave guides for both observed and unobserved quantities in the subpolar North Atlantic (cf. Fig. 6.4). Shared dynamical pathways and mechanisms imply that observations have (proxy) potential to inform remote and unobserved (or unobservable) QoIs, upstream or downstream from the observations. The detailed mathematical framework underlying this fact was established in this thesis.

Sensitivity information used in a more comprehensive way Acknowledging that observations capture information about unobserved or unobservable remote QoIs, extends a previously used method for adjoint-based observing system design. The previously used method (“Level 1” in Table 1.1 of Section 1.3.3) also uses linearized (adjoint-derived) sensitivities, however only computed for the unobserved QoIs. The method then suggests that instruments are most important at locations where the unobserved QoIs show high sensitivity, following the motto “high sensitivity informs high priority” [e.g., Marotzke et al., 1999; Heimbach et al., 2011]. In contrast, the novel approach via Hessian-based uncertainty quantification (“Level 3” in Table 1.1 of Section 1.3.3) that was presented in this thesis, suggests that observing systems could potentially be designed more efficiently if the instruments are arranged in such a way that important dynamical adjustment mechanisms are reliably captured. That is, the novel approach exploits that high sensitivity regions for the unobserved QoIs might be characteristics of dynamical adjustment mechanisms which also influence remotely or locally placed, existing or future, ocean observations. Consequently, some of the desired information about the QoI will already be captured by existing observations, which is ignored by the previously used method. The already captured information by existing observations can be formally assessed within the mathematical framework, taking into account data redundancy (see Chapter 3).

Target metrics In addition to the adjustment physics that is shared by many observations and unobserved QoIs, each QoI exhibits characteristics that are unique to the QoI, e.g., sensitivity to local processes. Therefore, the efficiency of observing systems will depend on the targeted climate QoIs. There is an ongoing discussion on which are the most important, and potentially most predictable, climate-relevant metrics to be targeted [e.g., Pohlmann et al., 2009; Branstator and Teng, 2014]. To list some of the prominent candidates, target metrics could be meridional or regional transports of heat, freshwater, or other properties, sea surface temperatures, heat or freshwater content, heat/freshwater exchange with the atmosphere, or sea level. The presented framework in this thesis offers a dynamics-based approach to quantifying the degree to which these target metrics can be informed by existing and future observing systems. Further investigation is required to identify combinations of efficient observing systems and target metrics. Integrated quantities, such as ocean heat content, might have a larger prediction potential than transport metrics because they are not sensitive to large transport anomalies driven by the instantaneous wind field (cf. Fig. 6.8(b)). Argo floats, drifting through the upper 2000 m of the ocean, could provide constraints on local air-sea heat fluxes that influence ocean heat content - a constraint, which mooring arrays, such as OSNAP, cannot provide in isolation (see Chapter 4).

Noise masking matters The results in Chapters 3 and 4 highlight that the effectiveness of observing systems is not only determined by the degree to which they capture adjustment physics of unobserved QoIs. A second important factor is to what degree the information about shared adjustment physics is masked by noise. In Chapter 4, it was found that, despite the vicinity of the two chosen QoIs to the OSNAP-East section and despite higher OSNAP-East hypothetical proxy potentials, the remote OSNAP-West observations are more effective in informing the QoIs. The recent study by Lozier et al. [2019] suggests a more important role of OSNAP-East compared to OSNAP-West, due to OSNAP-West’s minor role for setting strength and variability of the OSNAP overturning during the first 21 months of the OSNAP record. However, the results of Chapter 4 suggest that the much higher quality-to-relevance ratio of the OSNAP-West data leads to an (overall) much stronger constraint by OSNAP-West than OSNAP-East data. The stronger constraint by OSNAP-West can have wide-reaching implications - for climate signals away from OSNAP-West, for instance in the Nordic Seas. These results highlight that the *combination* of the two following questions is important: “What is your target metric?” and “Where to observe in order to detect the least noise-masked signals?”.

6.3.2 Are state estimation frameworks ready for formal observing system design?

Prior uncertainties: “Lost in optimization”?

Fig. 6.9 shows the model heat transport across OSNAP-East ($MHT_{OSNAP-E}$) in the EC-COv4 solution from 1992 to 2011. The associated green, orange, and yellow error bars represent two standard deviations of uncertainty for the respective monthly mean (green and orange) and 5-year-mean (yellow) *prior estimates* for $MHT_{OSNAP-E}$ within the EC-COv4r3 framework, i.e., the estimates *before* adding any data constraints. The prior uncertainties are computed by projecting the ECCOv4r3 prior uncertainties from *all*

Heat Transport across OSNAP-East in ECCOV4

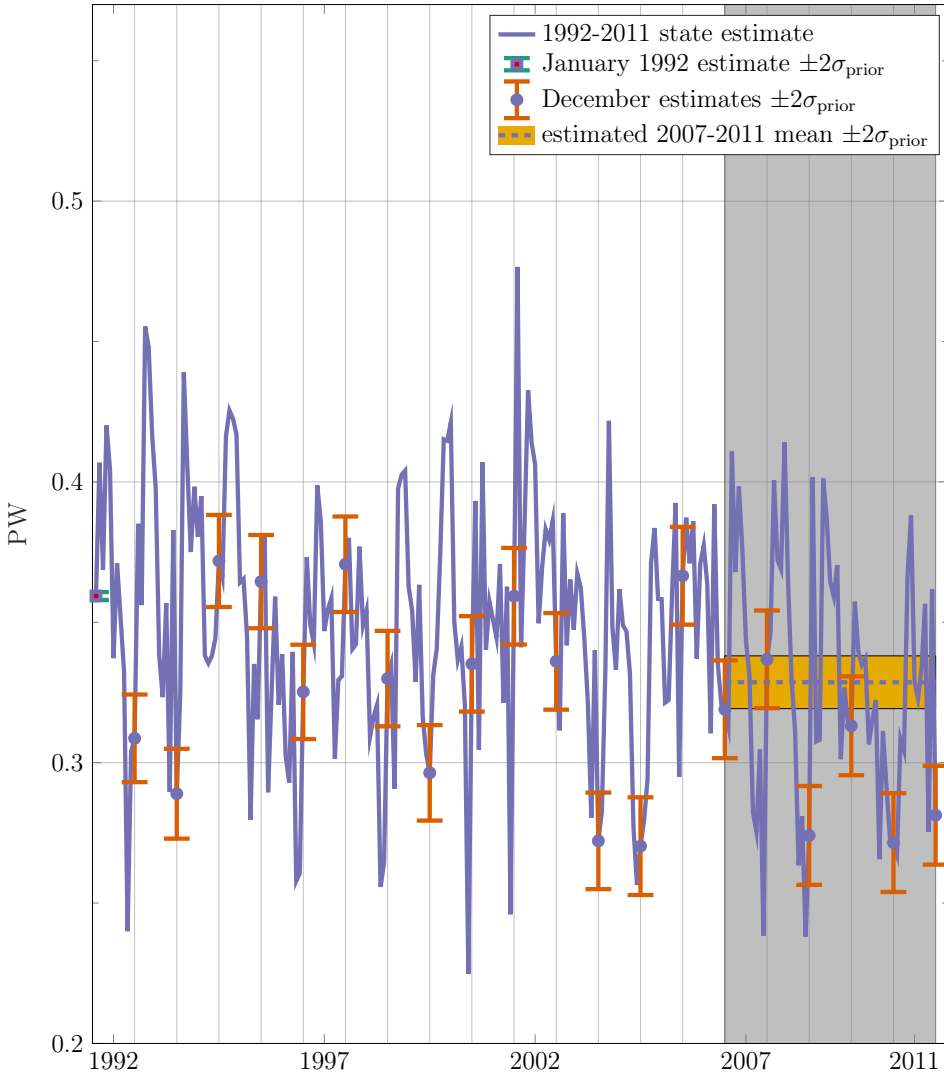


Figure 6.9: Prior ECCOV4r3 uncertainties examined through the lens of the induced prior uncertainties for heat transport across OSNAP-East ($MHT_{OSNAP-E}$). The purple line shows the ECCOV4 solution for $MHT_{OSNAP-E}$ from 1992 to 2011. The associated green, orange, and yellow error bars represent two standard deviations $\pm 2\sigma_{\text{prior}}$ of uncertainty for the respective monthly mean (green and orange) and 5-year-mean (yellow) *prior estimates* for $MHT_{OSNAP-E}$ within the ECCOV4r3 framework, i.e., the estimates *before* adding any data constraints. The prior uncertainties are computed by projecting the ECCOV4r3 prior uncertainties from *all* ECCOV4r3 control variables (but without Weaver and Courtier [2001] smoothing) onto $MHT_{OSNAP-E}$ for the respective time intervals (via (4.7) in Section 4.2.4).

ECCOV4r3 control variables (but without Weaver and Courtier [2001] smoothing¹) onto $MHT_{OSNAP-E}$ for the respective time intervals (via (4.7) in Section 4.2.4). Examining the ECCOV4r3 choice of the prior uncertainties for the control variables through the lens of the induced prior uncertainties for $MHT_{OSNAP-E}$ in Fig. 6.9, leaves the impression that the reported ECCOV4r3 prior uncertainties (in the form of weights for the control variables) are unrealistically small. In an inverse modeling framework, the specified prior uncertainties for the control variables reflect ranges within which the control variables are to be adjusted. The small reported ECCOV4r3 prior uncertainties stand in contradiction to large adjustments to the control variables, and a subsequent large adjustment of the $MHT_{OSNAP-E}$ estimate, when data constraints are included (personal communication with An Nguyen).

I suggest that the reason for this contradiction is that, in practical ocean state estimation, one often starts with adjustments from - not necessarily documented - pre-optimization steps. For instance, ECCOV4 performs pre-optimization of initial conditions and parameter fields (see the discussion in Section 4.4). Pre-optimization steps entail that *effective* adjustments of control variables are of much bigger amplitude than the reported prior uncertainties. In particular, the ECCOV4r3 prior standard deviations for initial conditions and parameter fields may be reported too small, due to the ECCOV4 pre-optimization step. This is supported by the short green error bar in Fig. 6.9, which - reflecting prior uncertainty in the initial month of the state estimate - should be much larger due to effectively larger prior uncertainty in initial conditions. To conclude, it is not always straight-forward to accurately translate weights from the regularization term in the ECCOV4 (or any other state estimate) optimization problem to covariance matrices in the probabilistic inverse problem. Uncertainties might get “lost in (pre-) optimization”. Within the Hessian-based observing system design framework presented in this thesis, unrealistically small chosen (or reported) prior uncertainties imply that observations are generally deemed less relevant than they should be, which is reflected in small prior-to-noise ratios, cf. Figs. 6.5 and 6.6. Erroneous prior uncertainties lead therefore to flaws when assessing observing systems and proxy potential (see Section 6.2.3). The issue of inaccurately specified (or reported) prior uncertainties has to be resolved before state estimation frameworks can become practical tools for observing system design.

An important part of ocean state estimation, as well as an accompanying uncertainty quantification, is to gain an understanding of how “good” the choice of the prior and observational uncertainties within the inverse modeling framework actually was. Finding the “right” uncertainties is therefore part of the science problem. This problem was addressed in Chapter 4 by scaling the noise and prior covariances matrices by constant factors, reflecting the assumptions of smaller/larger OSNAP observational uncertainties and/or smaller/larger prior control variable uncertainties. Introducing the scaling factors was a first approach to investigating the question “How do ECCOV4r3 prior control variable uncertainties have to be changed so that model prior uncertainties are larger than OSNAP data uncertainties?”, or, the related question: “How accurate does OSNAP

¹ Including Weaver and Courtier [2001] smoothing, as in ECCOV4, leads to even smaller respective prior uncertainties than the green, orange, and yellow error bars in Fig. 6.9. The loss of uncertainty is due to the large smearing effect of Weaver and Courtier [2001] smoothing on wind stress sensitivities, as will be discussed below.

data have to be to exceed prior information in the model?”. The answer is provided by Fig. 6.5. A further future direction will be to question the assumption that all control variables follow a Gaussian prior distribution. The ECCOv4 parameter control fields, such as diapycnal diffusivity, can vary over several orders of magnitude, with according uncertainties, and modeling them as lognormals might be beneficial. Lognormals may also be a better choice for some of the remaining control variables, e.g., for positive variables [Bocquet et al., 2010], such as precipitation, downwelling radiation, and specific humidity.

Smoothing

The diffusion operator approach of Weaver and Courtier [2001] (WC01) is a widely employed technique in ocean state estimation and data assimilation, used to model prior error covariance matrices [e.g., Moore et al., 2011a; Forget et al., 2015]. The application of the WC01 diffusion (or “smoothing”) operator to a diagonal prior error covariance matrix introduces prior cross-correlations between control variables that represent close-by grid points in the same field. In practice, employing WC01 smoothing implies that throughout the optimization process, control variables cannot be corrected at spatial scales smaller than the used WC01 diffusion scale. For instance, in ECCOv4 the WC01 diffusion scale is set to 3 times the grid scale of the underlying GCM [Forget et al., 2015]. Therefore, due to WC01 smoothing, observations can only inform control variable adjustments (e.g., adjustments of the atmospheric forcing fields) that are of larger scale than the scales that are potentially informable by the observations, leading to a potential “information loss”. A potential information loss is supported by the results in Appendix C, where information loss is caused by WC01-induced “smearing” of the constraints on wind stress forcing. This WC01-induced smearing effect is visible in Figs. 6.10(a) vs. (b).

In particular, it is found in Appendix C that the WC01-induced information loss on thermal forcing constraints is much less pronounced than the information loss on wind stress constraints, due to the fact that wind sensitivities are generally characterized by smaller-scale features than thermal sensitivities. The dissimilar effect of WC01 smoothing on thermal vs. wind sensitivities is found to have a significant impact on the assessment of observing systems and proxy potential within the framework presented in Chapters 3 and 4 (see Figs. C.1 and C.2 in Appendix C). It is suggested that the method of applying WC01 smoothing to thermal forcing fields and wind stress fields alike, with the same diffusion scales, may have to be revisited. This is supported by the spectral analysis performed in Appendix D, which shows that wind stress fields in reanalyses products are characterized by features of much smaller scales than thermal forcing fields.

Effect of computational simplifications

Hessian-based uncertainty quantification as a tool for dynamics-based ocean observing system design is a computationally demanding undertaking, due to the high-dimensional control space - a characteristic of oceanographic state estimation frameworks - comprised of about $O(10^6)$ – $O(10^8)$ control variables. Several previous studies have overcome the challenge of computing the high-dimensional Hessian matrix (whose dimension is the

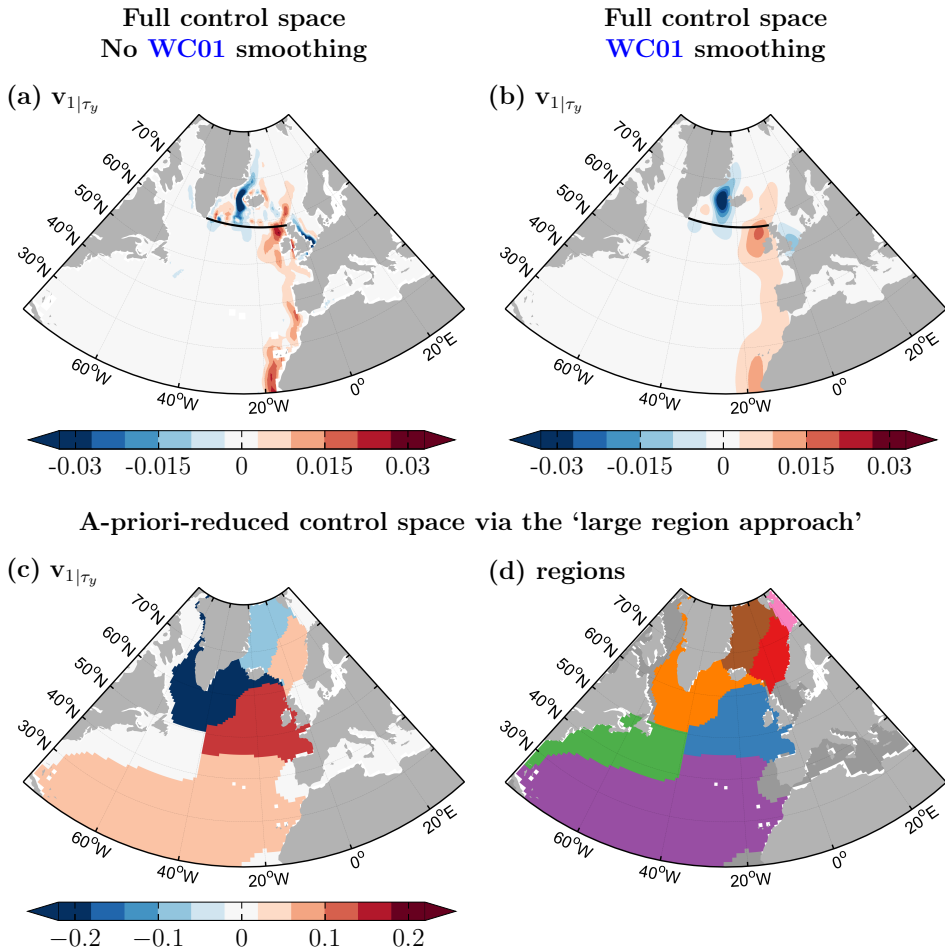


Figure 6.10: Information on meridional wind stress captured by the OSNAP-East observing system, expressed via $\mathbf{v}_1|_{\tau_y}$, the τ_y -subspace of the eigenvector \mathbf{v}_1 (cf. row A in Table 6.3). (a),(b),(c) show $\mathbf{v}_1|_{\tau_y}$, as computed for different choices of the control space. (a),(b) use the full ECCOv4r3 control space, but (a) without and (b) with Weaver and Courtier [2001] (WC01) smoothing. (c) uses an a-priori-reduced control space via the 'large region approach', where the large regions are defined according to the color coding in (d). (a) vs. (b): WC01 smoothing in (b) has a smearing effect on the sensitivity information from (a), potentially leading to information loss. (a) vs. (c): The patterns in (c) are the sensitivities from (a), but averaged over the large regions that are defined by the color coding in (d).

number of control variables *squared*) via two simplifications:

- (a) *A-priori-reductions* of the control space, for instance via the ‘large region approach’.
- (b) *Omitting* the computation of *the off-diagonal terms* of the Hessian.

The insights gained in this thesis elucidate that either of these two computational simplifications may lead to serious flaws in the assessments of observing systems and proxy potential.

(a) A-priori-reductions: In the ‘large region approach’, control variables are adjusted uniformly over large regions, rather than on a grid point basis of the underlying GCM [Kaminski and Rayner, 2008; Kaminski et al., 2015, 2018]. For instance, for an Arctic sea-ice prediction study, Kaminski et al. [2015, 2018] use the ‘large region approach’ to reduce the control space of their ocean-sea-ice model to only about 150 control variables, which makes it feasible to explicitly compute the full Hessian (150^2 entries). Fig. 6.10 elucidates the potential effects of the ‘large region approach’, if not used carefully. Both Figs. 6.10(a) and (c) show the information captured by the OSNAP-East observing system, expressed via the eigenvector \mathbf{v}_1 of the misfit Hessian (cf. row A in Table 6.3). Fig. 6.10(a), a replot of Fig. 6.4(a), shows (the τ_y -subspace of) \mathbf{v}_1 as computed in Chapter 4, which used the ECCOv4r3 control space at its full spatial resolution. In contrast, Fig. 6.10(c) is (the τ_y -subspace of) \mathbf{v}_1 computed via the ‘large region approach’, if the regions were defined according to the color coding in Fig. 6.10(d): sensitivities (aka information potential) get spatially averaged over these large regions. For instance, sensitivity concentrated along the eastern boundary of the Atlantic, visible in Fig. 6.10(a), would get distributed basin-wide over the North Atlantic in Fig. 6.10(c). The concentration of sensitivity along the boundaries in Fig. 6.10(a) is as a result of the coastal boundaries being a key pathway for signals that are observed by OSNAP-East (see Sections 6.1 and 6.2). Fig. 6.10 clarifies that, while this dynamical pathway is captured by a “high-resolution” control space, it may not be resolved when using the ‘large region approach’ - even if the underlying GCM is of high resolution. Unresolved dynamical pathways and mechanisms for observed and unobserved QoIs lead to errors when assessing observing systems and proxy potential (see Section 6.2).

Appendix D explores an alternative to the ‘large region approach’ for a-priori-reductions of the control space. It is investigated whether a truncated series of spherical harmonics could provide a useful basis for atmospheric adjustment in ocean state estimation. While this question needs further investigation, first insights suggest that a series of spherical harmonic basis functions, truncated at relatively low wavenumbers, may be an option for the adjustment of some of the thermal forcing fields. Given this fact can be confirmed by a more thorough analysis, computational and storage requirements could be reduced for high-resolution ocean state estimates, either by implementing spherical harmonic basis functions or performing atmospheric adjustment on a lower-resolution grid by means of multi-scale adjoints. It should also be explored whether other (e.g., physics-informed) decompositions serve as a useful and efficient low-rank basis.

(b) Skipping the off-diagonal terms: The black curves in Fig. 6.11, equal to the black curves in Fig. 6.6, show the relative uncertainty reduction (aka effective proxy potential) in the two QoIs studied in Chapter 4 by the combined OSNAP observing system. These black

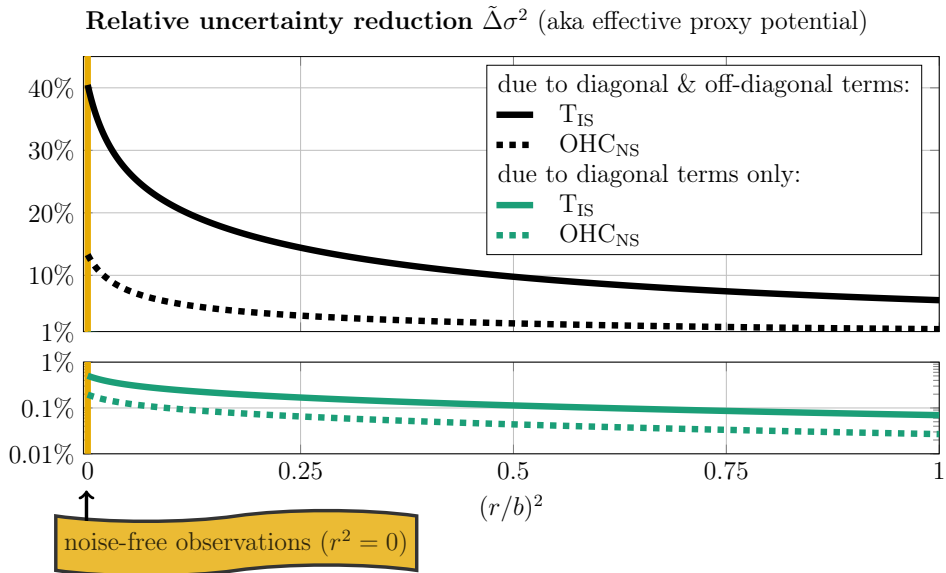


Figure 6.11: Relative uncertainty reduction, $\tilde{\Delta}\sigma^2$, for the unobserved QoIs Nordic Seas heat content (OHC_{NS} , dashed curves) and volume transport across the Iceland-Scotland ridge (T_{IS} , solid curves) by the combined OSNAP observing system, cf. Fig. 6.6. The black curves show $\tilde{\Delta}\sigma^2$ when taking into account the full information captured by OSNAP, encoded in diagonal *and* off-diagonal terms of the misfit Hessian, and coincide with the ones shown in Fig. 6.6. The green curves show $\tilde{\Delta}\sigma^2$, computed from only the diagonal terms of the misfit Hessian.

Relative variance reduction (only diagonal terms) in the control variables

(a) shortwave downward radiation $R_{SW\downarrow}$ (b) meridional wind stress τ_y

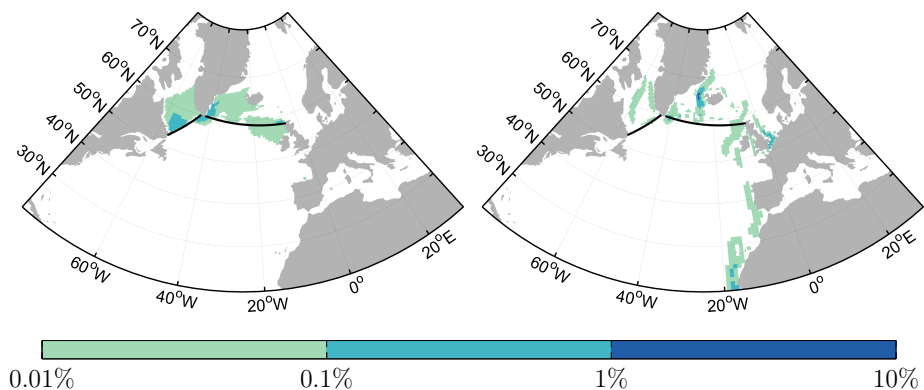


Figure 6.12: Relative variance reduction in the control fields of (a) shortwave downward radiation and (b) meridional wind stress, computed from the diagonal terms in the misfit Hessian, under the assumption of *noise-free* OSNAP observations. Note the logarithmic scale.

curves show the result of taking into account the full information captured by the observations, encoded in diagonal *and* off-diagonal terms of the misfit Hessian (cf. row A in Table 6.3). In contrast, the green curves in Fig. 6.11 show the uncertainty reduction in the same QoIs by the same observing system, but computed from only the diagonal terms of the misfit Hessian. If only the diagonal terms are taken into account, uncertainty in the QoIs gets reduced by only less than 1%, even for the hypothetical case of noise-free observations (green curves at $r^2 = 0$, indicated by the yellow vertical line in Fig. 6.11). In other words, the diagonal terms in the misfit Hessian do essentially not contribute to uncertainty reduction and proxy potential, whereas the largest part of the information contained in the observations is captured by the off-diagonal terms. This result is consistent with the findings by Kalmikov and Heimbach [2014, 2018], who identify the Hessian diagonal terms to be the major player for uncertainty reduction in Drake Passage transport by satellite altimetry data. Physically speaking, off-diagonal entries in the misfit Hessian reflect dynamical coupling between different physical variables as well as spatio-temporal dynamical teleconnections within the same physical field. The dynamical coupling and teleconnections, visible in the off-diagonal terms, are introduced by adding information on ocean dynamical principles to the observations, achieved by the data-model synthesis within an ocean state estimate.

Neglecting off-diagonal terms in the misfit Hessian corresponds to analyzing variance (rather than the full covariance) reduction in the control variables. Relative variance reduction in the control fields of shortwave downward radiation and meridional wind stress is shown in Fig. 6.12. Consistent with the negligible relative uncertainty reduction in the QoIs achieved by the diagonal entries in Fig. 6.11, relative variance reduction in Fig. 6.12 exceeds 1% only for few control variables, namely meridional wind stress in a few surface grid cells between Greenland and Iceland (Fig. 6.12(b)). Comparing the maps in Fig. 6.12 to the eigenvectors of the misfit Hessian corresponding to the combined OSNAP observing system, clarifies that variance reduction is achieved wherever sensitivity of the observed quantities is detectable. However variance reduction neglects information on the spatial distribution of positive vs. negative sensitivities - information that is closely linked to dynamical mechanisms and teleconnections among the control variables.

It is noted that analyzing only the diagonal terms of the Hessian, while ignoring its off-diagonal terms, is a commonly used method for observing system design² in many fields of applications, e.g., used for parameter estimation in porous medium flow problems, magnetotelluric methods, and image processing [e.g., Haber et al., 2011, Alexandrian et al., 2016], but has also been suggested for regional ocean state estimation [Moore et al., 2012, 2011b]. As suggested by Figs. 6.11 and 6.12, assessing variance reduction only, while neglecting all off-diagonal entries of the Hessian, can potentially be an extremely poor approximation for ocean state estimation problems. These facts pose the question what distinguishes ocean state estimation from state and parameter estimation in other fields of applications. One reason could be that in ocean state estimation, ocean observations show sensitivity to a wide range of control variables, spread throughout space, time, and across different physical control variables. This spreading is due to ocean dynamical coupling as well as the facts that ocean state estimation is a time-dependent problem and makes adjustments to various control fields that have a similar

² Targeting the minimization of the trace of the posterior covariance matrix (i.e., the inverse Hessian) is an alternative to targeting the minimization of uncertainty in a QoI (often referred to as A-optimality vs. C-optimality in Bayesian Experimental Design, see Chaloner and Verdinelli [1995]).

effect on observations (e.g., downward longwave radiation, downward shortwave radiation and air temperature). In other applications, the observation operator, i.e., the dependency of observations on all control variables in the model (denoted by $\mathbf{f}(\mathbf{x})$ in Chapters 3 and 4), might be more direct, causing the sensitivities of observations to be more localized within the control space. An example for a potentially more direct observation operator may be the one studied by Isaac et al. [2015]: in a time-independent problem, the basal sliding for the Antarctic ice-sheet - a *single* control parameter field - is informed by satellite measurements of surface ice flow velocity.

Way forward: Low-rank approximations The examples in Figs. 6.10, 6.11, and 6.12 suggest that the computational simplifications (a) and (b) do not seem to be a favorable approach in general, as they run the risk of substantially degrading the assessment of observing systems and proxy potential. Therefore, the computational challenge posed by high-dimensional control spaces may require a different solution strategy. Although the dimension of the control space is formally of tremendous order, one can suspect that only a low-dimensional subspace is “active” or “effective”, even if millions of observations are included in the ocean state estimate - or any other geophysical inverse problem [Constantine, 2015; Bui-Thanh et al., 2012]. To find this active subspace, the correct approach seems to “let the data speak through the lens of the model”, rather than making subjective a-priori decisions via approach (a). The data-informed approach is put into practice via low-rank approximations of the misfit Hessian, constructed by finding the misfit Hessian’s leading eigenvectors and eigenvalues, which characterize the components of the control space that are best determined by the data [e.g., Flath et al., 2011; Bui-Thanh et al., 2012; Isaac et al., 2015]. The leading eigenvectors and eigenvalues of the misfit Hessian are the components that matter the most in the assessment of observing systems and proxy potential. The leading eigenvectors can be further used for an *a-posteriori-reduction* of the control space, a common tool for data-informed model reduction [e.g., Bashir et al., 2008; Lieberman et al., 2013]. Low-rank approximations of the misfit Hessian can be achieved via Lanczos or randomized methods. Kalmikov and Heimbach [2014, 2018] have already employed Lanczos methods within a simplified version of the ECCO framework, using a control space with $O(10^5)$ variables, over a 3-month time window. For a tractable extension to the full ECCO control space and over climate-relevant time windows, it will be worth investigating the applicability of the ‘incremental 4D-Var’ algorithm [Courtier et al., 1994; Trémolet, 2007]. As part of the (lower-resolution) inner loop quadratic minimization problem, the ‘incremental 4D-Var’ algorithm computes the leading eigenvectors and eigenvalues of the Hessian, evaluated at the current estimation point of the outer loop, via a Lanczos process. ‘Incremental 4D-Var’ is for instance implemented within the ROMS (Regional Ocean Modeling System; Moore et al., 2011a) and the ECMWF (European Centre for Medium-Range Weather Forecasts; Rabier et al., 2000; Mahfouf and Rabier, 2000; Klinker et al., 2000; Trémolet, 2007] data assimilation systems, albeit in both frameworks for shorter time windows. In the computational sciences, low-rank approximations via Lanczos or randomized SVD methods are a rapidly growing research topic [e.g., Flath et al., 2011; Bui-Thanh et al., 2012; Isaac et al., 2015]. Research and applications in the coming years will shed more light on the feasibility of using low-rank approximations for oceanographic inverse problems.

Chapter 7

Conclusions

The main conclusions of this thesis are:

- The controlling mechanisms of variability in the subpolar North Atlantic and the Nordic Seas are a function of the timescale (Chapter 2).
- Anomalies in wind and surface buoyancy forcing generate seasonal to interannual variability in Nordic Seas heat content and in heat transport across the recently installed OSNAP-West and OSNAP-East transects. Both local and remote atmospheric forcing anomalies are important (Chapters 2 and 4).
- Surface heat flux anomalies influence hydrographic and transport quantities most efficiently when they occur locally (Chapters 2 and 4). For instance:
 - Nordic Seas heat content is highly sensitive to local air-sea heat fluxes.
 - Due to SST damping, advection of heat flux-driven temperature anomalies across the OSNAP transects shows a strong response in the OSNAP heat transports only if the heat flux anomalies occur in the vicinity of the array.
- Remote wind anomalies can impact Nordic Seas heat content and the OSNAP transports more efficiently than remote buoyancy flux anomalies. Pressure anomalies, forced by remote near-coastal wind perturbations, are communicated via coastally-trapped waves to the Nordic Seas and the OSNAP sections (Chapters 2 and 4).
- Wind stress curl anomalies along the main flow line of strong currents impact downstream hydrographic and transport metrics, via shifts in the current path. For instance, positive wind stress curl anomalies along the North Atlantic Current increase Nordic Seas heat content on advective timescales (Chapter 2).
- Even though Nordic Seas heat content and volume transport across the Iceland-Scotland ridge are closely linked quantities, they show surprising differences in their sensitivity to remote external forcing, due to different “residence capacities” of the quantities. For instance, volume transport rapidly loses its memory to remote wind-driven pressure anomalies that are communicated across the Iceland-Scotland ridge via coastally-trapped waves. In contrast, temperature anomalies carried across the ridge by this mechanism are “stored in memory” of Nordic Seas heat content (Chapter 2).

- The eastern boundary of the North Atlantic and the coasts of Iceland and Greenland are important pathways for communicating wind-driven pressure anomalies around the entire subpolar North Atlantic and the Nordic Seas. Consequently, the OSNAP observing array shares many dynamical pathways and mechanisms with quantities of interest (QoIs) that are remote from the array. The OSNAP array has therefore potential to inform these unobserved - or unobservable - QoIs, for instance ocean heat content in the Nordic Seas (Chapter 4) or close to Greenland's margins (Chapter 8). Hessian-based Uncertainty Quantification (UQ) within ocean state estimation frameworks serves as a mathematically rigorous tool to assess this potential (Chapter 3).
- In addition to the adjustment physics that is shared by the OSNAP observations and remote, unobserved QoIs, each QoI exhibits important, unique sensitivity patterns. For instance, Nordic Seas heat content shows strong sensitivity to local air-sea heat fluxes - processes that are not captured by the OSNAP observations (Chapter 4). Argo floats, however, could provide these constraints, when combined with the OSNAP observations. UQ provides the mathematical framework to quantify data redundancy and complementarity of multiple observing systems, for instance the OSNAP mooring array and Argo profiling floats.
- The effectiveness of an observing system is determined by how well it captures climate-relevant signals and important dynamical adjustment mechanisms. The captured signals are characterized by the Hessian eigenvectors. A second important factor, however, is how strongly the monitored signals are masked by noise. This is quantified by the Hessian eigenvalues (Chapter 3).
- Heat transport measurements across OSNAP-West impose a much stronger constraint on the overall ocean state estimated by ECCO version 4 than heat transport measurements across OSNAP-East, largely due to smaller OSNAP-West data uncertainties. As a result, even transport and hydrographic quantities in the Nordic Seas are constrained more efficiently by OSNAP-West than OSNAP-East observations, despite their vicinity to the OSNAP-East section (Chapter 4).
- The effectiveness of an observing system depends on the targeted climate QoIs.
- UQ within ocean state estimation may be a valuable tool for building and sustaining a cost-effective, long-term Atlantic observing system - while maximizing the information extracted from the observations. Novel aspects of the tool are: it
 - is fundamentally dynamics-based, making use of physical relationships in a general circulation model;
 - is favorable for climate monitoring because the dynamical relationships and budgets are respected and treated consistently over long timescales;
 - can assess data redundancy and complementarity in the context of other observations;
 - can identify the constraints of observing systems on climate signals and QoIs that are unobservable or unobserved, remote from the observations.

- Assumptions on prior information and data uncertainties are crucial ingredients of ocean state estimation, made explicit through the specification of error covariance matrices. Making accurate assumptions and modeling prior information appropriately is a prerequisite for state estimation frameworks to become practical tools for observing system design (Chapter 4).
- Improving the exactness and completeness of the adjoint would be a valuable advancement toward dynamics-based ocean observing system design on climate-relevant timescales (Chapter 5).
- Thermal atmospheric forcing fields are characterized by a large-scale spatial structure, while wind stress fields show spatial variability on very small scales (Appendix D). This suggests that computational and storage requirements for ocean state estimation may be reduced by adjusting the thermal atmospheric control variables on coarser grids than currently performed.

Chapter 8

Future Work and Applications

This chapter suggests avenues for future work and presents first results for two further applications. Section 8.1 explores whether remote ocean observing systems, such as the OSNAP array, can constrain subsurface temperature at Greenland's margins. The second application is to use adjoint models for learning more about past climates. Section 8.2 presents dynamical mechanisms that may have contributed to observed past climate variability at marine sediment core locations. Moreover, it is investigated how good/poor the constraints of proxy data on past ocean circulation are.

8.1 Subsurface ocean temperature at Greenland's margins

8.1.1 Impact of warm ocean waters on Greenland melt

The interaction of warm subpolar North Atlantic ocean waters with Greenland's marine-terminating glaciers has been suggested as a dominant trigger for the glaciers' recent retreat and acceleration, through increased submarine melting (see Fig. 1.2 in Section 1.1.3 and [Holland et al., 2008](#); [Straneo et al., 2010, 2012](#); [Vieli and Nick, 2011](#); [Joughin et al., 2012](#); [Straneo and Heimbach, 2013](#); [Rainsley et al., 2018](#)). Helheim Glacier and Jakobshavn Isbræ are two examples of marine-terminating glaciers in Greenland that have gained much attention recently, as they have undergone rapid changes since the 1990s [e.g., [Straneo et al., 2010](#); [Holland et al., 2008](#)].

Helheim Glacier is a major outlet of the ice sheet in southeast Greenland, marked by the yellow circle in Fig. 8.1(a). The glacier terminates in Sermilik Fjord, which connects Helheim Glacier with the Irminger Sea. In the early 2000s, the terminus of Helheim Glacier retreated about 8 km and its flow speed almost doubled [[Howat et al., 2007](#); [Stearns and Hamilton, 2007](#)]. Jakobshavn Isbræ is a marine-terminating glacier located on Greenland's west coast, marked by the yellow circle in Fig. 8.2(a). It drains 6.5% of the Greenland Ice Sheet [[Joughin et al., 2004](#)] into the Ilulissat Icefjord, which connects Jakobshavn Isbræ with Disko Bay. Warm waters carried by the Irminger Current can reach the glacial fjords of Helheim Glacier and Jakobshavn Isbræ at subsurface, and hence the two glaciers, after crossing the continental shelf (cf. Figs. 1.1 and 1.2 in Section 1.1).

8.1.2 Dynamical proxy potential of remote ocean observing arrays

It is logistically challenging to directly measure oceanic heat transport to the ice margin in Greenland's glacial fjords. However, existing *remote* ocean observing systems may constrain - and be able to reveal - subsurface temperature at Greenland's margins effectively because large-scale ocean dynamics are responsible for the delivery of warm subsurface waters to Greenland's ice front. The constraints - or proxy potential - of remote observing systems on subsurface temperature close to Greenland, based on dynamical principles, can be assessed by the approach described in this thesis (as summarized in Section 6.2). This approach uses an ocean state estimation framework as a dynamical interpolator from observed ocean quantities to unobserved ocean quantities.

In the following, OSNAP's dynamical proxy potential for unobserved subsurface temperature close to Helheim Glacier (subT_{HG}) and Jakobshavn Isbræ (subT_{JI}) is assessed and first results are presented. Here, the same framework as in Chapter 4 is employed, which assesses proxy potential of the *long-term* (5-year) mean of OSNAP heat transport for the *long-term* (5-year) mean of the unobserved quantities, within the ECCOv4r3 ocean state estimation framework [Forget et al., 2015; Fukumori et al., 2017]. For more details, the reader is referred to Section 4.2.

Helheim Glacier

Subsurface temperature close to Helheim glacier (subT_{HG}) is defined as the mean temperature averaged (i) over the horizontal area delineated by the yellow-black contour in Fig. 8.1(a) and (ii) a depth range of 150-590 m. Focusing on temperature below 150 m is motivated by the fact that, close to the Greenlandic coast, warm Atlantic waters flow below a layer of cold and fresh waters of Arctic origin, where the latter occupies the upper ~ 150 m of the water column (based on evidence from observations [Straneo et al., 2011, 2012] and the ECCOv4 state estimate [Forget et al., 2015]). The lower bound of the depth range is set based on the depth of Sermilik Fjord, which is estimated as 600-900 m deep [Schjøth et al., 2012]. The red-blue color shading in Fig. 8.1(a) shows the linearized sensitivity of the 5-year mean of subT_{HG} to the (spatially varying) 5-year mean of meridional wind stress, where the sensitivities are prior-weighted and normalized as explained in Section 4.3. Red (blue) colors indicate that an increase in northward wind stress would lead to a subsequent increase (decrease) in subT_{HG} . Positive-negative sensitivity dipoles straddle the main flow lines of the near-surface currents (green arrows), indicating that a change in wind stress over these currents can efficiently modulate the subpolar North Atlantic current system, which subsequently impacts subT_{HG} .

The remotely driven dynamical mechanisms that impact subT_{HG} (e.g., driven by wind stress anomalies throughout the subpolar North Atlantic, as visible in Fig. 8.1(a)) have potential to be informed by remote observations, if the observations are affected by similar dynamical mechanisms and pathways as subT_{HG} . Fig. 8.1(b) elucidates some of the dynamical mechanisms and pathways that affect observed heat transport across the OSNAP-East section ($\text{MHT}_{\text{OSNAP-E}}$, as defined in Section 4.2.2; marked by the yellow section in Fig. 8.1(b)). The red-blue shading in Fig. 8.1(b) shows normalized and prior-weighted sensitivity of $\text{MHT}_{\text{OSNAP-E}}$ to meridional wind stress. Red (blue) colors indicate that an increase in northward wind stress would lead to a subsequent increase

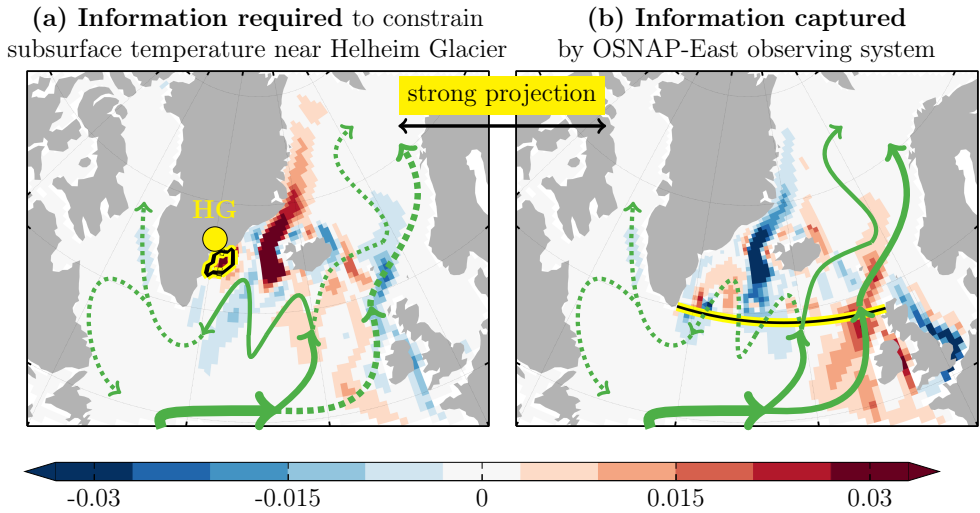


Figure 8.1: Normalized prior-weighted sensitivities of 5-year mean (a) unobserved subsurface ocean temperature (150-590 m) close to Helheim Glacier ($\text{sub}T_{\text{HG}}$) and (b) observed heat transport across the OSNAP-East section ($\text{MHT}_{\text{OSNAP-E}}$), to 5-year mean meridional wind stress, computed within ECCOV4. Red (blue) colors indicate that an increase in northward wind stress would lead to a subsequent increase (decrease) in (a) $\text{sub}T_{\text{HG}}$ and (b) $\text{MHT}_{\text{OSNAP-E}}$. In (a) and (b), sensitivity dipoles (red-blue shading) straddle the main flow lines of the near-surface currents (solid and dashed green arrows). Those surface currents that are *strengthened* by wind stress perturbations matching the sensitivity distributions in (a) and (b) are shown as *solid* green arrows in the respective maps. The yellow-black contour delineates the (a) horizontal area and (b) cross-sectional OSNAP-East transect for the respective calculations of (a) mean temperature and (b) heat transport.

(decrease) in northward heat transport across OSNAP-East. As in Fig. 8.1(a), sensitivity of $MHT_{OSNAP-E}$ is concentrated along the major near-surface currents of the subpolar North Atlantic, reflecting the impact of wind stress on the subpolar North Atlantic current system, which, if shifted, can effectively alter heat transport across OSNAP-East. Comparing the wind sensitivity maps in Figs. 8.1(a) and (b), one observes that most of the signs of the sensitivity dipoles in (a) vs. (b) are flipped. Flipped signs are due to the fact that $subT_{HG}$ can be increased efficiently by a wind-driven strengthening of the *Irminger current* (indicated by the solid green arrows in Fig. 8.1(a)), whereas $MHT_{OSNAP-E}$ is most efficiently increased by a wind-driven strengthening of the *eastern NAC branches* (indicated by the solid green arrows in Fig. 8.1(b)). The Irminger Current crosses the OSNAP-East transect several times but eventually flows southward around the southern tip of Greenland, which strongly affects $subT_{HG}$, but leads only to a small *net* northward heat transport across OSNAP-East. In contrast, a relative strengthening of the eastern NAC branches that flow into the Nordic Seas and Arctic, rather than a relative strengthening of the Irminger Current, leads to an efficient increase in $MHT_{OSNAP-E}$.

The extremely similar structure of the sensitivity maps in Fig. 8.1(a) and (b) - of *consistently* opposite sign - leads to a strong projection of the two maps. The excellent projection implies that OSNAP-East has high proxy potential for $subT_{HG}$, if three further requirements are fulfilled:

- 1) OSNAP-East also captures dynamical mechanisms driven by other controls than wind stress, e.g., mechanisms driven by surface heat fluxes.
- 2) No cancellation in information occurs, due to competing dynamical adjustments, e.g., in response to thermally vs. wind-driven mechanisms (see Sections 4.3.4 and 4.4 for further explanations).
- 3) The masking by noise in the OSNAP observations is not too strong.

If requirement 3) is ignored, i.e., OSNAP observations are assumed to be noise-free, the hypothetical proxy potential of the OSNAP array for $subT_{HG}$ is equal to 45% (measured as relative uncertainty reduction, see Section 6.2). The missing 55% are mostly due to lack of information on thermally driven mechanisms, i.e., requirement 1) is not satisfied. Many of the thermally driven processes that affect Helheim Glacier are caused by surface heat fluxes that occur locally or in the advective influence region of Helheim Glacier (not shown). These are not unambiguously captured by the OSNAP array. Other observations, such as Argo profiling floats, may be able to constrain these thermally driven processes to some degree, which is to be assessed in future work. It is noted that missing information on thermally driven mechanisms means that thermally driven mechanisms do not compete with wind-driven mechanisms in the information transfer from observed $MHT_{OSNAP-E}$ to unobserved $subT_{HG}$. Therefore, no information is cancelled, i.e., requirement 2) is fulfilled.

Jakobshavn Isbræ

Subsurface temperature close to Jakobshavn Isbræ ($subT_{JI}$) is defined as the mean temperature averaged (i) over the horizontal area delineated by the yellow-black contour in Fig. 8.2(a) and (ii) a depth range of 150-380 m. The lower bound of the depth range is

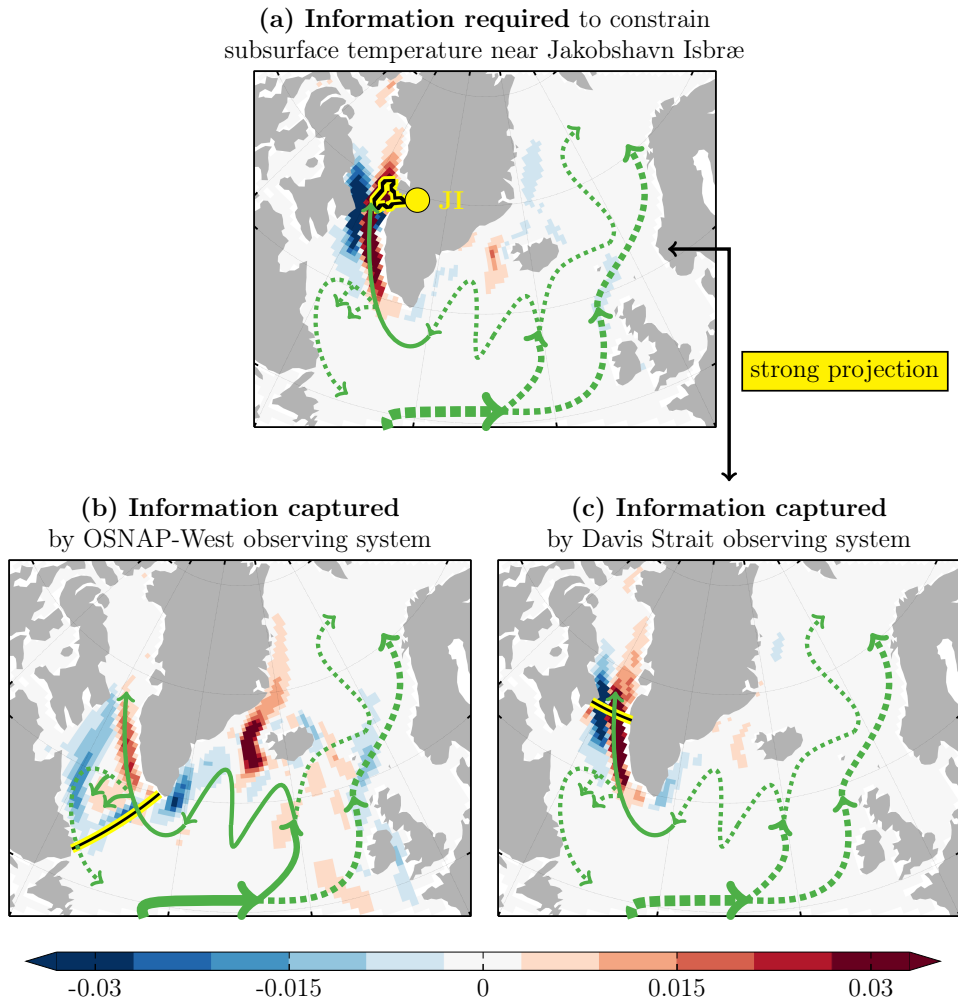


Figure 8.2: Normalized prior-weighted sensitivities of 5-year mean (a) unobserved subsurface ocean temperature (150–380 m) close to Jakobshavn Isbræ (subT_{JI}), (b) observed heat transport across the OSNAP-West section ($\text{MHT}_{\text{OSNAP-W}}$), and (c) observed heat transport across the Davis Strait Array ($\text{MHT}_{\text{Davis}}$), to 5-year mean meridional wind stress, computed within ECCOV4. Red (blue) colors indicate that an increase in northward wind stress would lead to a subsequent increase (decrease) in (a) subT_{JI} , (b) $\text{MHT}_{\text{OSNAP-W}}$, and (c) $\text{MHT}_{\text{Davis}}$. Those surface currents that are *strengthened* by wind stress perturbations matching the sensitivity distributions in (a), (b), and (c) are shown as *solid* green arrows in the respective maps. The yellow-black contour delineates the (a) horizontal area, (b) OSNAP-East section, and (c) Davis Strait for the respective calculations of (a) mean temperature and (b),(c) heat transport.

chosen shallower than for Helheim Glacier because warm subsurface waters that enter Ilulissat Icefjord must pass over a relatively shallow entrance sill, at depths between 150 and 450 m [Schumann et al., 2012]. Fig. 8.2(a) shows the normalized and prior-weighted sensitivity of the 5-year mean of subT_{JI} to the 5-year mean of meridional wind stress. Red (blue) colors indicate that an increase in northward wind stress would lead to a subsequent increase (decrease) in subT_{JI} . Compared to the spatially wide-spread wind sensitivities of subT_{HG} in Fig. 8.1(a), the wind sensitivities of subT_{JI} in Fig. 8.2(a) are concentrated in the Labrador Sea and Baffin Bay, close to Jakobshavn Isbræ. A wind stress perturbation matching the sensitivity dipole between Baffin Island and the west Greenlandic coast in Fig. 8.2(a) leads to a strengthening of the branch of the Irminger Current that continues northward into Baffin Bay and toward Jakobshavn, branching off the current that circles the Labrador Sea (solid arrow in Fig. 8.2(a)). The local wind sensitivities of subT_{JI} , contrasting the spatially wide-spread wind sensitivities of subT_{HG} , can be explained as follows: while Helheim Glacier lies right in the middle of the subpolar North Atlantic current system, with the Irminger Current passing nearby, Jakobshavn Isbræ is “tucked away” in Baffin Bay. Therefore, subT_{JI} shows highest sensitivity to wind-driven processes that cause warm Irminger waters to flow northward into Baffin Bay (solid arrow in Fig. 8.2(a)), rather than recirculating around the Labrador basin or ending up in the central Labrador Sea (dotted arrows in Fig. 8.2(a)).

The OSNAP-West transect extends from Labrador to the southwestern tip of Greenland, shown by the yellow section in Fig. 8.2(b). Fig. 8.2(b) elucidates the wind-driven dynamical mechanisms and pathways that can affect observed heat transport across the OSNAP-West section ($\text{MHT}_{\text{OSNAP-W}}$, as defined in Section 4.2.2). Fig. 8.2(b) shows normalized and prior-weighted sensitivity of $\text{MHT}_{\text{OSNAP-W}}$ to meridional wind stress. Red (blue) colors indicate that an increase in northward wind stress would lead to a subsequent increase (decrease) in northward heat transport across OSNAP-West. Similarly as for OSNAP-East (Fig. 8.1(b)), $\text{MHT}_{\text{OSNAP-W}}$ shows sensitivity to wind stress spread across the entire subpolar North Atlantic surface system. Compared to OSNAP-East (Fig. 8.1(b)), sensitivity has opposite sign, favoring wind stress directions that strengthen the Irminger Current (solid green arrows in Fig. 8.2(b)), rather than the eastern branches of the NAC (dotted arrows in Fig. 8.2(b)). Comparing the dynamical mechanisms and pathways that impact $\text{MHT}_{\text{OSNAP-W}}$ (Fig. 8.2(b)) with those that impact subT_{JI} (Fig. 8.2(a)) elucidates that OSNAP-West can capture some of the adjustment physics for subT_{JI} , but is “distracted” by too many other signals that are triggered elsewhere: the hypothetical proxy potential for noise-free OSNAP observations for subT_{JI} computes to only 6%.

The Davis Strait Array is a sparse array of subsurface moorings [Curry et al., 2010], which are placed across Davis Strait, marked by the yellow section in Fig. 8.2(c). Fig. 8.2(c) shows normalized and prior-weighted sensitivity of heat transport across Davis Strait ($\text{MHT}_{\text{Davis}}$) to meridional wind stress. Red (blue) colors indicate that an increase in northward wind stress would lead to a subsequent increase (decrease) in northward heat transport across the Davis Strait. The similarity of the wind sensitivity maps in Figs. 8.2(a) and (c) demonstrates that subT_{JI} and $\text{MHT}_{\text{Davis}}$ have shared wind-driven adjustment physics. This leads to a strong projection of Figs. 8.2(a) and (c) and 35% hypothetical proxy potential of observed $\text{MHT}_{\text{Davis}}$ for unobserved subT_{JI} ,

observing system	unobserved subsurface temperature close to	
	Helheim Glacier	Jakobshavn Isbræ
OSNAP Array	45%	6%
Davis Strait Array	<1%	35%

Table 8.1: Proxy potential of the OSNAP and Davis Strait Arrays for unobserved subsurface temperature at Greenland’s margins, in the hypothetical case of noise-free observations.

assuming noise-free observations. The missing 65% are due to differences in the adjustment physics of subT_{JI} and $\text{MHT}_{\text{Davis}}$ to other controlling forces, e.g., by surface heat fluxes.

8.1.3 Outlook

The preliminary results for the dynamical proxy potential of the OSNAP and Davis Strait Arrays for subsurface temperature close to Helheim Glacier and Jakobshavn Isbræ are summarized in Table 8.1. The values assume that both observing systems deliver noise-free measurements of heat transport across the respective arrays. In reality, the effective proxy potentials will be smaller than the ones reported in Table 8.1, because the information on shared dynamical mechanisms and pathways will be masked by noise. Future work will account for noisy observations. Moreover, the study presented in this section will be extended by: (i) assessing proxy potential for *summer vs. winter* subsurface temperature close to the glaciers, (ii) including the array observations at their full spatio-temporal information, by taking into account instrument measurements from the moorings, and (iii) employing ECCOv4 (or a regional spin-off) at a higher resolution, to ensure better resolved dynamics close to the Greenlandic coast.

8.2 Dynamical proxy potential of paleoceanographic observations

8.2.1 Combining paleo proxy data with dynamical models

Paleoclimate archives provide information on the evolution of past climates and can improve our understanding of the processes that underlie low-frequency variability in the climate system. Furthermore, abrupt climate shifts in the past may provide analogs for future climate change. Inference of past environmental conditions is complicated by the fact that they can only be assessed on the basis of *indirect* proxy observations. The relationship of proxy observations to physical variables is often tenuous, leading to large uncertainties, and the observations are very sparse in space and time (see Section 1.1.4).

For paleoclimate reconstructions, model-data synthesis is promising since it permits using dynamical principles to constrain interpretations of proxy observations, to quantify the inferential power of proxy data, and to compute observable and unobservable climate aspects. To this aim, many authors have undertaken efforts to combine models and data into paleoceanographic state estimates (see Huybers and Wunsch [2010]

for a review). For instance [LeGrand and Wunsch \[1995\]](#), [Huybers et al. \[2007\]](#), [Marchal and Curry \[2008\]](#), [Gebbie and Huybers \[2006, 2011a,b\]](#) use box inverse methods to combine observational information with dynamical principles and generate estimates for either water mass distributions or oceanic circulation and transports. By means of the adjoint method, [Winguth et al. \[1999\]](#), [Dail and Wunsch \[2013\]](#), [Kurahashi-Nakamura et al. \[2017\]](#), [Amrhein et al. \[2018\]](#) combine a full ocean GCM with paleoceanographic data from the Last Glacial Maximum (roughly 19,000 - 23,000 years ago), inverting for a seasonally varying equilibrium solution. Such paleoceanographic state estimates use the assumption that the utilized proxy data is representative of the annual or seasonal mean from a single *time-slice*.

Adjoint-derived paleoceanographic state estimates, entailing a time-slice assumption, cannot resolve abrupt climate changes, such as rapid transitions from cold Greenland stadials to warm Greenland interstadials (see Section 1.1.4). However, adjoint models may still be valuable to answer scientific questions in paleoceanography, as long as they concern gradual, rather than sudden, very non-linear, changes. Note that the use of adjoint models for paleoceanographic applications is limited by the fact that the linearity assumption becomes less and less accurate over long timescales (Chapter 5). In this section, two questions will be investigated by means of an adjoint model. Section 8.2.2 studies dynamical mechanisms that may have contributed to observed D-O variability in the proxy records of various sediment cores, with a focus on MD99-2284 (light green circle in Fig. 8.3) in the southern Norwegian Sea. Section 8.2.3 is an attempt to quantify how efficient the constraints by proxy data from marine sediment cores on ocean state estimates are. These constraints are compared to those by modern-day ocean observations.

8.2.2 Mechanisms contributing to D-O temperature variability

Marine core site MD99-2284

The marine sediment core MD99-2284 (light green circle in Fig. 8.3) is located in the northward inflow of Atlantic waters into the southern Norwegian Sea. Based on proxy records from this core, [Dokken et al. \[2013\]](#) and [Sadatzki et al. \[2019\]](#) suggest that the abrupt warming of Greenland air temperature, i.e., the shifts from cold Greenland stadials to warm Greenland interstadials during the last glacial (see Section 1.1.4), are triggered by a sea-ice reduction in the Nordic Seas. The sea-ice break up, in turn, is suggested to be preceded by a heat buildup below the sea-ice, driven by enhanced northward heat transport of Atlantic waters into the Nordic Seas. However, the mechanism that induced the enhanced northward heat transport is debated. Similarly, the mechanism that caused the steadily increasing sea-ice cover in the southern Norwegian Sea during the latter half of the warm Greenland interstadials, as reconstructed by [Sadatzki et al. \[2019\]](#), is unclear. The authors hypothesize that the driver is internal ocean dynamics, such as reduced northward advection of warm Atlantic waters or local surface ocean freshening.

To explore origins and mechanisms that can drive variability in temperature - and consequently sea-ice - at core site MD99-2284, adjoint sensitivities of SST in the model

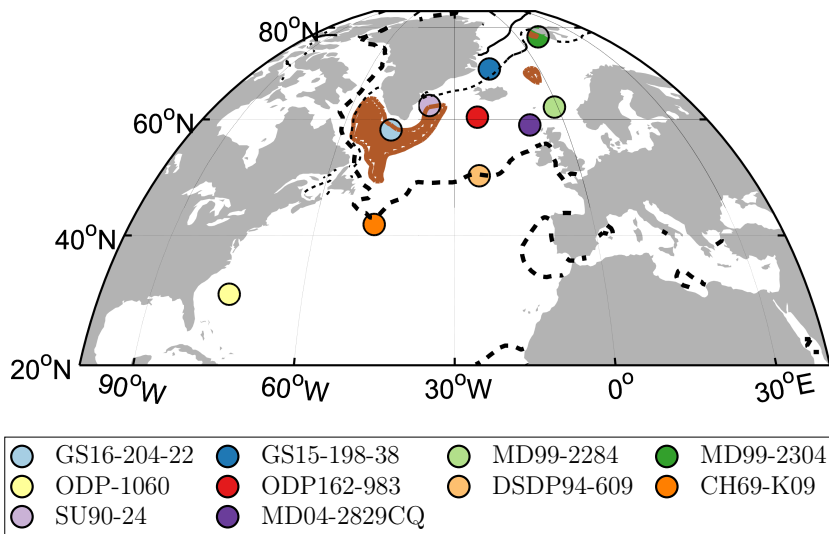


Figure 8.3: Map showing locations of the study sites (colored circles). The sites and the associated names in the legend relate to marine sediment cores with SST proxy data from marine isotope state 3 (29-60 kyr ago) during the last glacial, which shows a number of abrupt climate change events (Section 1.1.4). The thick black dashed line delineates the zero-line of the barotropic stream function. The thin black dashed and solid lines show March and September sea-ice extent, respectively. The brown shading marks regions with a March mixed layer depth greater than 1000 m. Barotropic stream function, sea ice extent and mixed layer depth are all computed for the modern reference simulation.

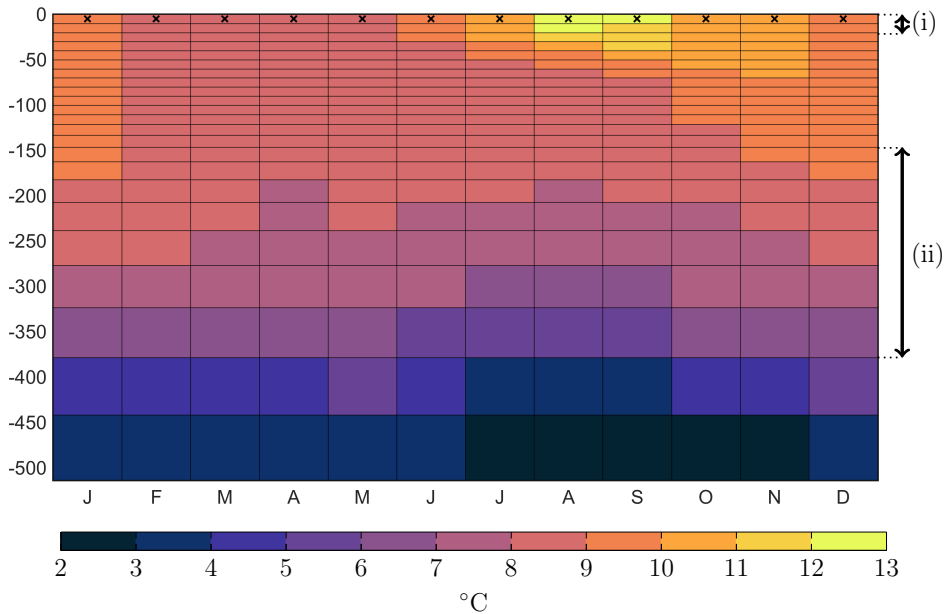


Figure 8.4: Seasonal cycle of temperature in the 500 m upper ocean at core location MD99-2284 (light green circle in Fig. 8.3) in the modern reference simulation. The 12 letters enumerate the 12 months January to December. For each month, the uppermost depth level (0-5 m) is the model grid cell with the warmest temperature throughout the water column (indicated by the black crosses). The two depth ranges, marked on the right, indicate the chosen depth levels for computing (i) surface and (ii) subsurface temperature at site MD99-2284.

grid cell nearest to MD99-2284 are presented. The sensitivities are computed around the equilibrated *modern* reference simulation that was described in Section 2.2. This equilibrated simulation exhibits no interannual variability, yet a seasonal cycle. As an example, the seasonal cycle of the ocean temperature distribution throughout the upper 500 m at core location MD99-2284 is shown in Fig. 8.4. In the modern reference simulation, core location MD99-2284 shows the warmest temperature at the surface (indicated by the black crosses), throughout the entire year. However, in the winter and spring months, temperature is approximately uniform over the upper 200-300 m of the water column.

Figs. 8.5(a) and (b) show linearized sensitivities of SST (0-10 m) at core location MD99-2284 to local and remote surface buoyancy forcing. For sea-ice covered areas, the flux is defined at the ice-ocean interface, below the sea-ice. The sensitivities in Fig. 8.5 are accumulated for forcing lead times up to 15 years. Mechanisms that operate on longer timescales are not included in this presentation, due to the inaccuracy of the adjoint for longer time integrations (see Chapter 5). The sensitivities are prior-weighted and smoothed, as described in Sections 4.2.4 and Appendix C. In Fig. 8.5, the prior uncertainties for air-sea heat and freshwater fluxes are chosen equal to the constants $\Delta Q_{\text{net}} = 40 \text{ W/m}^2$ and $\Delta(\text{E-P-R}) = 10^{-7} \text{ m/s}$. Spatially constant prior uncertainties are a common choice in paleoceanographic state estimates [e.g., Dail and Wunsch, 2013, Kurahashi-Nakamura et al., 2017, Amrhein et al., 2018], due to the lacking knowledge of the statistics of air-sea fluxes.

SST at MD99-2284 shows localized sensitivity to surface heat flux (Fig. 8.5(a)), due to the direct impact of air-sea heat fluxes on surface temperature. In contrast, sensitivity to surface freshwater flux emerges also remotely (Fig. 8.5(b)), throughout the entire North Atlantic north of 40°N , as well as in Baffin Bay, west of Greenland, and in Hudson Bay. Positive sensitivities in the subpolar North Atlantic, Baffin Bay, and Hudson Bay indicate that enhanced E-P-R (e.g., by less ice melt) in these regions will lead to higher SST at MD99-2284 on the considered timescale (15 years). The sensitivity dipole across the main flow line of the NAC indicates that freshwater flux can alter baroclinicity and, by thermal wind balance, the associated sheared transport across the core of the separated jet. A stronger Gulf Stream and NAC carries more warm waters to core location MD99-2284. This highlights the possibility that melting ice-sheets may not only have lowered temperature at MD99-2284 by inhibiting deep convection, but also by weakening the NAC.

A next step will be to investigate how the modern-day sensitivities presented in Figs. 8.5(a),(b) change if they are instead computed for a glacial background state, e.g., when the Nordic Seas and a large portion of the North Atlantic are covered in sea-ice (see Fig. 1.4(a)). Since the laws of ocean physics, i.e., equations of motions representing basic conservation and constitutive laws, do not change over time and sensitivities in Figs. 8.5(a),(b) are computed to fluxes at the ice-ocean interface (below the sea-ice), glacial sensitivity patterns are expected to be similar as those for the modern ocean in Figs. 8.5(a),(b). For instance, SST at MD99-2283 will still show strong local sensitivity to surface heat flux (as in Fig. 8.5(a)) - even if surface heat fluxes might have come to a halt if the core site was covered in sea-ice. Similarly, freshwater flux sensitivities are

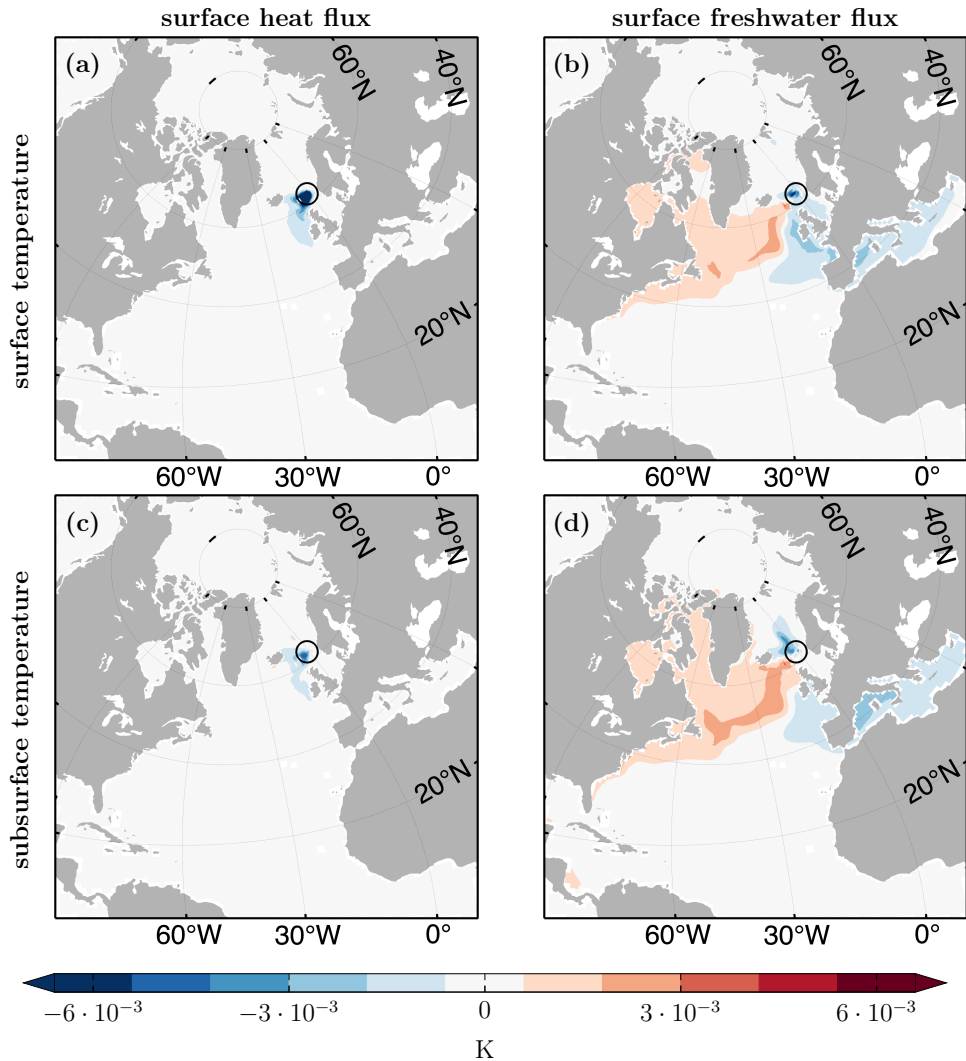


Figure 8.5: Linearized sensitivities of (a),(b) annual mean SST and (c),(d) winter (JFM) subsurface temperature at core location MD99-2284 (black circle) to (a),(c) surface heat flux Q_{net} and (b),(d) surface freshwater flux E-P-R. Positive sensitivity indicates that increased (a),(c) heat loss to the atmosphere and (b),(d) freshwater removal causes higher temperature at MD99-2284. Sensitivities are computed around a *modern* reference simulation, accumulated for forcing lead times up to 15 years, prior-weighted, and smoothed. The prior uncertainties (or weights) for the air-sea surface fluxes are chosen spatially constant, equal to $\Delta Q_{\text{net}} = 40 \text{ W/m}^2$ and $\Delta(\text{E-P-R}) = 10^{-7} \text{ m/s}$.

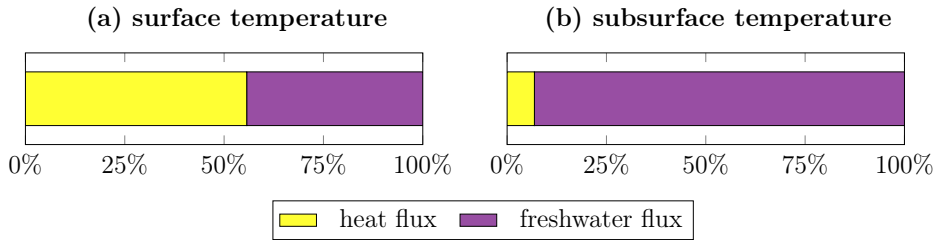


Figure 8.6: Relative sensitivity of (a) annual mean surface (0-10 m) and (b) winter subsurface (150-380 m) temperature at core location MD99-2284 to surface heat and freshwater flux, integrated around the globe.

expected to show roughly the same pattern as in Fig. 8.5, with a sensitivity dipole along the main flow line of the NAC. The center of the dipole, however, might be shifted, depending on the position of the surface currents in past climates.

Surface vs. subsurface temperature proxies

Paleoceanographic SST reconstructions are commonly obtained from planktonic (near-surface dwelling) foraminifera assemblages, via a transfer function approach [e.g., CLIMAP Project Members, 1976]. Transfer functions are trained on calibration data sets of the same organisms in the modern ocean. Calibration techniques usually correlate planktonic foraminifer counts to *surface* temperature. However, many species of foraminifera can potentially live at *subsurface*, between 0 and 300 m water depths, depending on the species-specific environmental preferences and hydrographical conditions [e.g., Table 3 in Staines-Urías et al., 2013]. For instance, Carstens et al. [1997] provide evidence that planktonic foraminifera in (modern) Arctic environments prefer the subsurface Atlantic waters over low-salinity polar surface waters. Motivated by the uncertainty in water depth which paleoceanographic temperature proxies are representative of, the sensitivity analysis presented in Figs. 8.5(a),(b) is repeated for winter (JFM) *subsurface* (150-380 m) temperature, at the same core location (MD99-2284). Here, winter temperature, rather than annual mean temperature, is chosen because the winter mixed layer depth at MD99-2284 concurs approximately with the lower bound of the studied subsurface range (380 m), while the summer mixed layer depth is shallow (Fig. 8.4). At the same time, some species of planktonic foraminifera are thought to reflect winter (rather than annual mean) conditions [e.g., Waelbroeck et al., 2002]. The chosen depth ranges for (i) surface and (ii) subsurface are marked in Fig. 8.4.

Figs. 8.5(c) and (d) show the sensitivity of *subsurface* temperature at core site MD99-2284 to surface buoyancy fluxes. The emerging sensitivity patterns are very similar to the respective ones for surface temperature in Figs. 8.5(a) and (b), since surface and subsurface temperature are impacted by the same buoyancy-driven dynamical mechanisms (as discussed above). However, sensitivities of surface vs. subsurface temperature show large differences in *amplitude*. Subsurface temperature (Fig. 8.5(c)) is much less sensitive to local air-sea heat fluxes than surface temperature (Fig. 8.5(a)), since only a small fraction of air-sea heat fluxes entrain the subsurface. In contrast, sensitivity to

freshwater flux is slightly higher for subsurface temperature (Fig. 8.5(d)) than for surface temperature (Fig. 8.5(b)). This is explained by the fact that temperature changes, induced by freshwater-driven circulation changes, are maintained more efficiently at subsurface, where they are isolated from the atmosphere, than at the surface, where SST damping is at play.

The differences in the sensitivity amplitudes for surface vs. subsurface temperature, identified in Fig. 8.5, is confirmed by Fig. 8.6. Fig. 8.6(a) shows the relative magnitudes of the prior-weighted sensitivities in Figs. 8.5(a) and (b), when integrating them around the globe (in the l^2 -norm, cf. equation 4.10). For the chosen weights, the relative sensitivity of SST at MD99-2284 to surface heat flux is 56% (yellow bar in Fig. 8.6(a)). Solely *local* heat flux sensitivities are responsible for this fraction (as visible in Fig. 8.5(a)), while one would hope that proxy data at one study site provides information on larger-scale climate forcings and states. The remaining 44% are due to freshwater flux around the globe (purple bar in Fig. 8.6(a)). In contrast, if one integrates prior-weighted sensitivities of *subsurface* temperature at MD99-2284 (Figs. 8.5(c),(d)) around the globe, the relative sensitivity to surface heat fluxes is only 7% (yellow bar in Fig. 8.6(b)). As a result, the relative sensitivity to freshwater flux increases to 93% (purple bar in Fig. 8.6(b)). To conclude, subsurface temperature is overall more sensitive to remotely driven processes than surface temperature. Remotely driven processes, in turn, may not only be of local, but potentially of larger-scale climatic importance. Therefore, species of planktonic foraminifera that are thought to reflect subsurface, rather than surface, conditions may be more indicative of the large-scale, rather than local, climate in the past.

Freshwater flux sensitivity of core locations north vs. south of the ridge

In today's climate, the regions that experience the largest variability in freshwater forcing are the Arctic Ocean, the Greenland Sea, and parts of the supolar gyre (see Fig. A.2). These are the regions that are characterized by high sea-ice variability. In a glacial climate, the spatial extent and amplitude of freshwater forcing may have been very different, due to the presence of large continental ice-sheets and a more extensive sea-ice cover, possibly expanding into the Nordic Seas and parts of the North Atlantic (see Fig. 1.4(a)). Therefore, in the next few paragraphs, special attention is given to *freshwater flux* sensitivities. The following question is investigated: How does temperature at different core sites respond to freshwater input across the North Atlantic, the Nordic Seas, and the Arctic Ocean? Here, first results are shown for the two core sites MD99-2283 and ODP162-983, located north and south of the Iceland-Scotland ridge.

Figs. 8.7(a) and (b) show freshwater flux sensitivity of SST (0-10 m) at MD99-2284 and ODP162-983, respectively. Sensitivities are computed around the modern reference simulation and accumulated for forcing lead times up to 15 years, as in Fig. 8.5. Fig. 8.7(a) is the same as Fig. 8.5(b), except that now sensitivities are not weighted by forcing anomalies (such that the unit is K/(mm/day), rather than K). Moreover, due to the logarithmic color scale in Fig. 8.7(a), sensitivities of smaller amplitude, which were masked in Fig. 8.5(b), are now visible. As a result, one observes that the sensitivity dipole along the main flow line of the NAC extends all the way across the Fram Strait into the Arctic basin. This suggests that thermal wind balance can act even downstream

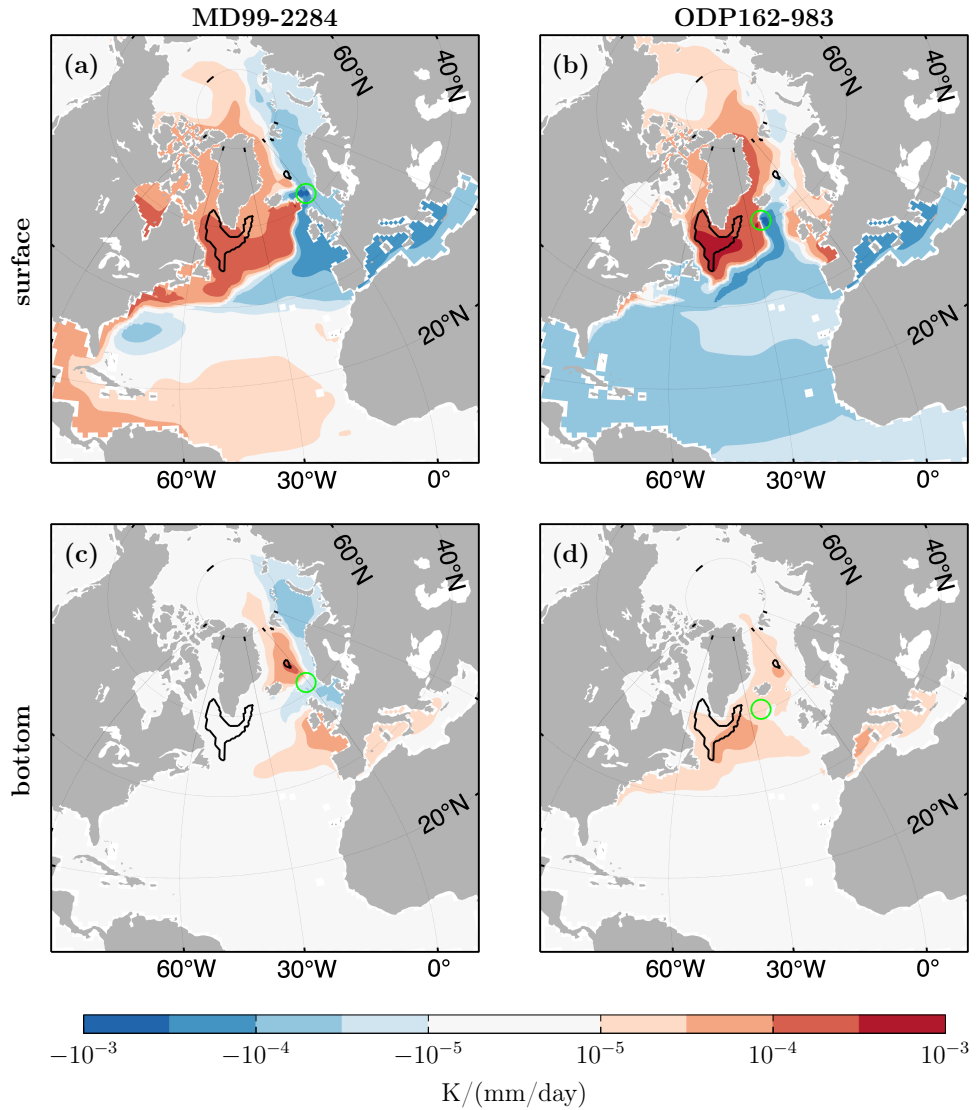


Figure 8.7: Linearized sensitivity of annual mean (a),(b) *surface* temperature and (c),(d) *bottom* temperature at core sites (a),(c) MD99-2284 and (b),(d) ODP162-983 to surface freshwater flux. The core sites in (a)-(d) are marked by the respective green circles. Positive sensitivity indicates that increased evaporation (corresponding to increased ocean salinification) causes higher temperature at the respective core location and water depth. Sensitivities are computed around a *modern* reference simulation and accumulated for forcing lead times up to 15 years. The black contour line bounds the regions wherein the March mixed layer depth exceeds 1000 m. Note the logarithmic color scale.

of the Nordic Seas, by “sucking” more Atlantic water into the Nordic Seas. Similarly as for SST at MD99-2284, a sensitivity dipole emerges for SST at ODP162-983 in the North Atlantic and the Nordic Seas (Fig. 8.7(b)). However, south of the Iceland-Scotland ridge, the center of the dipole is shifted to the west such that it follows the Irminger Current (Fig. 8.7(b)), rather than the Norwegian Atlantic Current (Fig. 8.7(a)). This is explained as follows: a freshwater flux perturbation that matches the sensitivity dipole in Fig. 8.7(b) strengthens the Irminger Current, which brings more warm surface waters to core site ODP162-983. Note that the negative sensitivities in the eastern North Atlantic extend all the way into the subtropical gyre (Fig. 8.7(b)).

Marine sediment cores provide not only proxy information about the ocean conditions near the surface, inferred from planktonic (near-surface-dwelling) species, but also about the ocean bottom, inferred from benthic (bottom-dwelling) species. Therefore, freshwater flux sensitivities of ocean *bottom* temperature at MD99-2284 and ODP162-983 are shown in Figs. 8.7(c) and (d), respectively. Ocean bottom temperature is defined as the mean temperature in the deepest model grid cell which contains the respective core site. The depth ranges are 960-1030 m for MD99-2284 and 2170-2220 m for ODP162-983, respectively. As before, sensitivities are accumulated for forcing lead times up to 15 years - timescales that are clearly too short to assess the full impact of surface forcing onto temperature at the ocean bottom. This fact partly explains the weak sensitivities seen in Figs. 8.7(c),(d), with amplitudes that are about one order of magnitude less than the ones for surface temperature in Figs. 8.7(a),(b). Nevertheless, a comparison of Figs. 8.7(c),(d) and Figs. 8.7(a),(b) is interesting. The sensitivity patterns for bottom temperature (Figs. 8.7(c),(d)) show a signature of changes in overturning circulation. Strongest sensitivities are found in the deep convection regions in the Nordic Seas for MD99-2283 (Fig. 8.7(c)) and in the deep convection regions in the subpolar North Atlantic and the Nordic Seas for ODP162-983 (Fig. 8.7(d)). This supports the hypothesis that the deep convection regions are the sweet spots for freshwater input to obtain a response in the deep ocean on decadal - and likely longer - timescales. This hypothesis does not prove true for a response in surface temperature: as discussed before, the sensitivity patterns for surface temperature (Figs. 8.7(a),(b)) show a signature of horizontal circulation changes in the subpolar North Atlantic, rather than changes in overturning. It remains to be investigated if changes in horizontal circulation still outweigh changes in overturning circulation if longer timescales are taken into account.

8.2.3 Constraints of proxy data on past ocean circulation

Decomposing the Hessian of the regularized model-data misfit function into its eigenvectors and eigenvalues elucidates which components of ocean circulation are well/poorly constrained by the available data (A in Table 6.3). The eigenvectors tell about which components are constrained - in the hypothetical case of noise-free data - and are composed of sensitivity maps as in Figs. 8.5 and 8.7. The eigenvalues, in turn, inform about the strength of these constraints. This section explores how strong the constraints by modern vs. proxy SST observations are on ocean circulation.

For a single observation, e.g., an SST data point from a single core site, the eigenvalue

		(a)	(b)	(c)	(d)	(e)	
		Paleo A	Paleo B	Paleo C	Paleo D	Modern	Unit
Prior uncertainty	atemp	4	4	4	1	ECCOV4r3	°C
	lwdown	0	0	20	0	ECCOV4r3	W/m ²
	swdown	20	20	20	10	Fig. C.3(a)	W/m ²
	aqh	$2 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	10^{-3}	ECCOV4r3	-
	precip	$4 \cdot 10^{-8}$	$4 \cdot 10^{-8}$	$4 \cdot 10^{-8}$	10^{-8}	ECCOV4r3	m/s
	ustress	0.06*	0.12*	0.1	0.01*	ECCOV4r3	N/m ²
	vstress	0.06*	0.12*	0.1	0.01*	Fig. C.3(c)	N/m ²

Figure 8.8: (a)-(d): Spatially constant prior uncertainties in air temperature at 2 m (atemp), downward longwave radiation (lwdown), downward shortwave radiation (swdown), specific humidity at 2 m (aqh), precipitation (precip), zonal wind stress (ustress), and meridional wind stress (vstress), as chosen in 4 paleo state estimates in the literature. Paleo A: [Dail and Wunsch \[2013\]](#) (“LGM_S3”); Paleo B: [Dail and Wunsch \[2013\]](#) (“LGM_S4”); Paleo C: [Amrhein et al. \[2018\]](#); Paleo D: [Kurahashi-Nakamura et al. \[2017\]](#). Prior uncertainties for ustress and vstress that are marked with ‘*’ have been converted from the respective original prior uncertainties for zonal and meridional wind at 10 m (see text). (e): Modern prior uncertainties are adopted from ECCOV4r3 [[Fukumori et al., 2017](#)], and are *not* spatially constant.

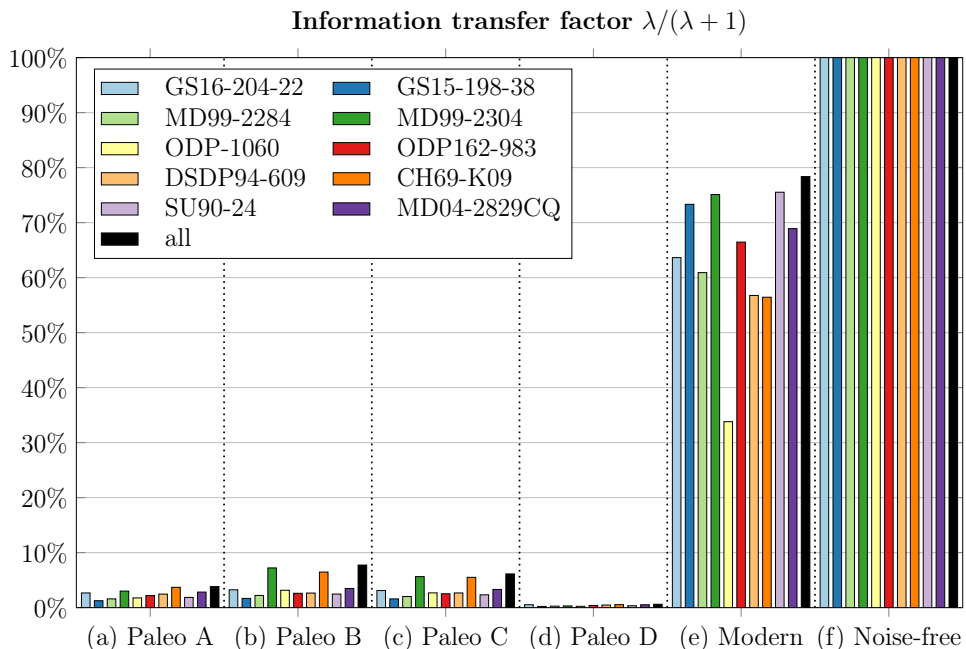


Figure 8.9: Overall quality of SST data constraints measured by the fraction $\lambda/(\lambda + 1)$. Here, λ is the largest eigenvalue of the misfit Hessian, if including SST data from the 10 core sites in Fig. 8.3 individually (colored bars) and combined (black bar). Subpanels (a)-(f) show the data constraints for different state estimation frameworks (see text).

λ is given by

$$\lambda = \frac{(\text{total sensitivity of observation to controls}) \cdot (\text{prior uncertainty})}{\text{data uncertainty}}, \quad (8.1)$$

see Section 3.3. λ quantifies the strength of the constraint (on the component of ocean circulation that is informable by the observation under consideration) and is high if

- (i) the observed quantity has *high sensitivity* to the unknown control variables (e.g., to atmospheric forcing);
- (ii) *prior uncertainty* is *high*;
- (iii) the *data uncertainty* is *small*.

How does a “typical” λ compare in a paleoceanographic vs. a modern-day state estimate? Prior uncertainty is higher for past climates than for the modern-day climate. Therefore, according to factor (ii), there is potential for high constraints by proxy data, since potentially more new information can be gained. On the other hand, the data uncertainty of proxy observations is very high. According to factor (iii), this will weaken the constraints by proxy data on ocean circulation. As for (i), the sensitivities of observations (e.g., SST) will depend on the background state, e.g., the modern-day vs. paleo sea-ice distribution. While differences between modern-day and paleo sensitivities are expected, the details are likely to be subtle.

To get an estimate of the constraints of paleo vs. modern-day SST observations via (8.1), six different state estimation frameworks are considered: four paleoceanographic frameworks that have previously been used by Dail and Wunsch [2013], Amrhein et al. [2018], and Kurahashi-Nakamura et al. [2017], one modern framework, and one hypothetical framework with noise-free data.

Every state estimation framework has to make an assumption on factor (ii), i.e., the prior uncertainties for the unknown control variables that are to be adjusted by means of the data constraints. Columns (a)-(d) in Table 8.8 represent 4 different sets of choices for spatially constant prior uncertainties in atmospheric control variables (atmospheric temperature, downward radiation, specific humidity, precipitation, and wind stress) that have been made in the studies by Dail and Wunsch [2013] (Paleo A+B), Amrhein et al. [2018] (Paleo C), and Kurahashi-Nakamura et al. [2017] (Paleo D), for paleoceanographic state estimates. Dail and Wunsch [2013] and Kurahashi-Nakamura et al. [2017] specify prior uncertainty for wind at 10 m, rather than a wind *stress* forcing, as considered here. Nevertheless, for the sake of consistency, all six frameworks are assumed to use the same set of control variables. Therefore, in this section the prior uncertainties for wind from Dail and Wunsch [2013] and Kurahashi-Nakamura et al. [2017] are converted to prior uncertainties for wind stress (marked with ‘*’ in Table 8.8), using estimates based on the bulk formulae of Large and Yeager [2004]. The prior uncertainties for the modern state estimation framework (column (e) in Table 8.8) are adopted from ECCOv4r3 [Fukumori et al., 2017], chosen equal to the uncertainties for the time-mean adjustments of the atmospheric control variables (cf. Section 4.2.4). In contrast to the frameworks Paleo A-D, the prior uncertainties in ECCOv4r3 are *not* spatially constant. As an example,

Figs. C.3(a),(c) show the spatially varying prior uncertainties for shortwave downward radiation and meridional wind stress.

Besides the choice of prior uncertainties for the control variables, a state estimation framework requires the choice of factor (iii), i.e., the data uncertainties for all observations used in the state estimate. Here, in the frameworks Paleo A-D, it is assumed that the SST proxy data uncertainties at all 10 core sites in Fig. 8.3 are 1°C . This choice is independent of the studies by Dail and Wunsch [2013], Amrhein et al. [2018], and Kurahashi-Nakamura et al. [2017], who investigate different data points from the ones shown in Fig. 8.3. It is noted that specifying a uniform SST uncertainty of 1°C for each of the core locations is a simplified - and optimistic - choice. While uncertainties for proxy data are difficult to assess and not always provided, the MARGO project estimates uncertainties for a global compilation of SST proxy observations from the Last Glacial Maximum [Waelbroeck et al., 2009]. Across different sediment cores, the MARGO-defined uncertainties range from 0.8°C to 4.9°C with a mean of 2.3°C . For modern-day temperature observations, the uncertainty is much smaller. For instance, Argo profiling floats have a measurement accuracy of 0.002°C [Argo, 2019]. Sensor drift introduces additional uncertainty, where amplitudes are largest in regions of strong currents. However, the largest addition to data uncertainty in a state estimation framework is the representation error, which arises due to sub-grid scale processes that affect the observations but cannot be resolved in the model. The representation error depends on model resolution and data location. Here, for simplicity, an uncertainty of 0.1°C is specified uniformly for all SST observations at each of the ten locations shown in Fig. 8.3.

As for factor (i) in equation (8.1), the sensitivity of annual mean SST at each of the marine sediment core sites in Fig. 8.3 is computed by means of the adjoint, in the equilibrated *modern* reference simulation that was described in Section 2.2. Even though sensitivities are expected to strongly depend on the background state, as discussed above, they are assumed to be the same in the modern-day and past ocean circulations. A next step will be to investigate how sensitivities change for a glacial background state, e.g., when the Nordic Seas and a large portion of the North Atlantic are covered in sea-ice (see Fig. 1.4(a)). In contrast to Section 8.2.2, sensitivities are accumulated much longer here, over 50 years. This choice is an attempt to follow previous studies on paleoceanographic state estimation and neglects the fact that the linearity assumption of the adjoint will be broken on a 50-year timescale (Chapter 5).

Each colored bar in Fig. 8.9 shows the efficiency of the constraint by an individual SST observation, located at one of the 10 locations shown in Fig. 8.3. Subpanels (a)-(d) represent the four paleoceanographic frameworks Paleo A-D (described in the previous paragraphs), subpanel (e) the modern framework, and subpanel (f) the hypothetical framework of noise-free observations. All subpanels show the factor $\lambda/(\lambda + 1)$, computed from (8.1). The factor $\lambda/(\lambda + 1)$ is equal to the fraction of the dynamical information that can be extracted from the SST observation under consideration, as compared to the same observation in the noise-free case (see D in Table 6.3). The factor $\lambda/(\lambda + 1)$ serves therefore as a measure for the overall strength of the data constraint, while it does not specify *what* part of the circulation is constrained. Assuming noise-free data, one can extract 100% of the dynamical information that is captured by each core location

(Fig. 8.9(f)). In all other - noisy - frameworks (Figs. 8.9(a)-(e)), the factor $\lambda/(\lambda + 1)$ takes values less than 100% and is large if λ in (8.1) is large. Moreover, in each subpanel of Fig. 8.9, the black bar shows the fraction $\lambda_1/(\lambda_1 + 1)$ for the largest eigenvalue λ_1 , when combining all 10 SST observations. The black bar is always larger than all individual (colored) bars and its magnitude is set by the largest individual bar.

Comparing Figs. 8.9(a)-(d) with Fig. 8.9(e) suggests that the constraints by proxy data on past ocean circulation are much weaker than the constraints by the corresponding modern data on the modern-day ocean circulation. For all SST observations, only a small fraction of the dynamical information can be extracted in the paleoceanographic frameworks: less than 5% in Paleo A (Fig. 8.9(a)), less than 8% in Paleo B (Fig. 8.9(a)), less than 6% in Paleo C (Fig. 8.9(a)), and less than 1% in Paleo D (Fig. 8.9(a)). In contrast, the extractable dynamical information from the corresponding modern SST observations range from 33% to 76% (Fig. 8.9(e)), depending on the location.

The large difference between the data constraints in Paleo A-D (Figs. 8.9(a)-(d)) and the data constraints in the modern framework (Fig. 8.9(e)) is due to the choice of prior uncertainties (factor (ii) in (8.1)) and data uncertainties (factor (iii) in (8.1)), since sensitivities (factor (i) in (8.1)) only differ among core locations, but not across the frameworks in Figs. 8.9(a)-(f). Data uncertainties are assumed to be 10 times higher for SST proxy observations than for modern SST observations. Therefore, to obtain data constraints in Paleo A-D that are as strong as those in the modern framework, prior uncertainties in Paleo A-D would have to be chosen 10 times higher than in the modern framework. Such a choice is not performed in practice. For example, Fig. 4.4(a) elucidates that in the subpolar North Atlantic and the Nordic Seas, modern prior standard deviations range from 30 to 60 W/m² for shortwave downward radiation (swdown) while the spatially constant values for the paleo prior standard deviations are chosen as 20 W/m² in Paleo A-C and as 10 W/m² in Paleo D (Table 8.8), i.e., even *smaller* than the modern prior uncertainties (rather than 10 times larger). For meridional wind stress, modern prior standard deviations range from about 0.01 to 0.06 N/m² in the subpolar North Atlantic and the Nordic Seas (Fig. 4.4(b)). The corresponding choice in the frameworks Paleo A-D range from spatially constant values of 0.01 N/m² (Paleo D) to 0.12 N/m² (Paleo B). Modern prior standard deviations times a factor of 10 are therefore only realized in the frameworks Paleo B and C - at few locations. As a result, the data constraints in Paleo A-D (Figs. 8.9(a)-(d)) are much weaker than in the modern framework (Fig. 8.9(e)), as outlined in the previous paragraph. The extremely weak data constraints in Fig. 8.9(d) are due to the assumption of very small prior uncertainties in Paleo D, compared to Paleo A-C (Table 8.8). Smaller prior uncertainties in Paleo D correspond to assuming more prior knowledge about the past climate state, compared to Paleo A-C. If the climate and ocean state is already well-known based on prior information, as assumed in Paleo D, less information can be added by uncertain data (cf. equation 8.1).

Within each of the paleoceanographic frameworks (Figs. 8.9(a)-(d)), prior uncertainties and data uncertainties do not differ across the core locations. Therefore, within each subpanel (a)-(d), the notable differences in $\lambda/(\lambda + 1)$ across the 10 locations is due to varying *sensitivity* of SST at the individual core sites. According to equation (8.1), the

higher the sensitivity, the larger λ and the data constraint $\lambda/(\lambda + 1)$. For instance, in each subpanel (a)-(d), MD99-2304 (green bar in Figs. 8.9(a)-(d)) and CH69-K09 (orange bar Figs. 8.9(a)-(d)) impose generally larger constraints, because the sensitivity of SST at these two locations is higher than at the remaining core locations. SST at MD99-2304 shows high sensitivity to air-sea fluxes in the subpolar North Atlantic and the Norwegian Sea, whereas SST at most other core sites shows only sensitivity in the subpolar North Atlantic (not shown). This is because the Norwegian Sea is an upstream region for MD99-2304, but a downstream region for most other core locations (Fig. 8.3). CH69-K09 appears to show high sensitivity because the core is located at the western junction of the subtropical and subpolar gyres, where the separated Gulf Stream, the North Atlantic Current, and Labrador Current all interact with each other (Fig. 8.3). Within the modern framework (Fig. 8.9(e)), prior uncertainties are spatially varying. Therefore, sensitivity to air-sea fluxes is not weighted uniformly across the globe. For instance, prior uncertainty in wind stress is very high east of Greenland (Fig. 4.4(b)) and, thus, wind stress sensitivity east of Greenland is weighted more strongly than wind stress sensitivity elsewhere. As a result, core locations close to Greenland's east coast, e.g., SU90-24 and GS15-198-38 (light purple and dark blue bars in Fig. 8.9), impose relatively large data constraints in the modern framework. Fig. 8.9 emphasizes that data constraints in modern and paleoceanographic state estimates are not only highly dependent on data uncertainties, but also on prior uncertainties - both affect the weighting of the data.

To conclude, Figs. 8.9(a)-(d) suggest that the data constraints on past ocean and climate states are very weak, due to high data uncertainties and, potentially, too small prior uncertainties. Note again that the constraints in Figs. 8.9(a)-(e) are measured against the constraints of hypothetical noise-free SST data (Fig. 8.9(f)). Therefore, Fig. 8.9 does not explain which components of ocean circulation the considered (noisy or noise-free) SST data would constrain - this information would be captured by the eigenvectors, rather than the eigenvalues. In fact, the bottom line of Fig. 8.9 is that even if noise-free SST data were able to constrain certain circulation or hydrographic quantities, proxy uncertainties are too large to extract this information. It should be noted that even noise-free SST data would not be able to constrain ocean circulation rates, unless an assumption of a reference velocity is made, such that transport rates can be inferred from thermal wind balance. This is discussed in detail by [Huybers and Wunsch \[2010\]](#).

The finding that proxy data cannot provide effective constraints in paleoceanographic state estimation is consistent with previous work [e.g., [Huybers et al., 2007](#); [Kurahashi-Nakamura et al., 2014](#)]. While the approach of most previous studies has been to perform identical twin experiments, this section supports the result by means of a different method. Computing the eigenvalues of the Hessian, as performed here, provides a systematic and mathematically rigorous way to quantifying data constraints in ocean state estimation. However, the approach is limited by a local Gaussian approximation of the posterior probability distribution. As discussed in Section 3.6.2, this assumption is problematic if the cost function has multiple minima, corresponding to multiple possible states that are consistent with the data. Multiple minima may be expected in paleoceanographic frameworks, when prior uncertainties and data uncertainties are chosen large. A second limitation is that sensitivities in Fig. 8.9 are accumulated only for 50

years. While the results from Chapter 5 suggest that even for this time window, the linearized approximation for the non-linear response of the adjoint is not accurate, 50 years represent a short time window for paleoceanographic timescales. Under the assumption that the linearity assumption holds, the constraints in Figs. 8.9(a)-(d) would become more effective, if the total sensitivity increases considerably when sensitivities are accumulated over centuries and beyond (rather than 50 years). This is not guaranteed: when accumulating sensitivities over time, cancellations occur at the locations where sensitivities have opposite signs for different lead times. Moreover, highest sensitivities are usually found for shorter lead times, which are already included in Figs. 8.9(a)-(d). Therefore, the results from this section, together with those from previous studies, suggest that it will remain challenging to use paleoceanographic proxy data in order to place firm bounds on past ocean states.

Appendix A

Nordic Seas: Sensitivity vs. Response to Atmospheric Forcing

Section 2.4 analyzes sensitivities of Nordic Seas upper-ocean heat content that are weighted by typical forcing anomalies. Here, two aspects of the weighted sensitivities - or responses - are investigated. The first goal is to disentangle whether a high response to forcing anomalies in a particular region is due to (i) high sensitivity to this forcing type in the considered region, (ii) high typical forcing anomalies in this region, or (iii) a large region size, which integrates rather small effects to a significant contribution. The second goal is to analyze the seasonality of the responses, i.e., the effect of local and remote forcing anomalies on Nordic Seas upper-ocean heat content, depending on the season in which the forcing anomalies occur. Seasonality of the responses can be caused by seasonality of (unweighted) sensitivities, or seasonality of typical forcing anomalies, or by a combined effect.

First, regions of high sensitivity are identified, regardless of whether the region experiences high forcing anomalies or not. This is an interesting metric to look at because typical forcing anomalies could change in the future. For example, freshwater flux anomalies in the Arctic could become larger due to more sea-ice melt. The second step is an analysis of the typical forcing anomalies which sensitivities have been weighted with.

A.1 Regions of high sensitivity

The average value of the absolute sensitivity of Nordic Seas upper-ocean heat content to air-sea fluxes in each of the regions \mathcal{R} defined in Section 2.4 is computed as a function of lead time t , where t is discretized into monthly values:

$$\left| \frac{\partial \text{UOHC}_{\text{NS}}}{\partial F} \right|^{\mathcal{R}}(t) = \frac{\int_{\mathcal{R}} \left| \frac{\partial \text{UOHC}_{\text{NS}}}{\partial F(x,y,-t)} \right| dx dy}{\int_{\mathcal{R}} dx dy}. \quad (\text{A.1})$$

Here, F can be either surface heat flux Q_{net} , surface freshwater flux E-P-R, zonal wind stress τ_x or meridional wind stress τ_y . One could also define (A.1) without taking the absolute values in the numerator, as for example done in Czeschel et al. [2010]. However, this would lead to cancellations of positive-negative dipoles and small-scale structure in the sensitivity patterns that emerge within one region, e.g., the sensitivity dipoles in

Figs. 2.4(b),(c),(d), Figs. 2.7(c),(d) and Figs. 2.8(c),(d) that straddle the main flow line of the Gulf Stream and the NAC. Therefore the metric (A.1) is chosen, which is a measure for whether a region exhibits high or low absolute sensitivity, regardless of whether sensitivities are positive or negative.

The four panels in the first row of Fig. A.1 show the average absolute sensitivities (A.1) to each of the four different forcings, where contributions from the different regions are stacked. Note that each panel has to be studied separately since average sensitivities are measured as upper-ocean heat content anomaly per flux and area unit, and are thus not comparable across different forcings. For each forcing type, sensitivity of Nordic Seas upper-ocean heat content is highest to forcing anomalies that are applied locally, on all considered timescales. For heat flux anomalies, other regions are hardly discernible in Fig. A.1(a), due to the dominance of sensitivity to locally occurring anomalies. For freshwater flux and wind stress anomalies, sensitivity is also visible in the NAC, the Gulf Stream, the Subpolar Gyre, and the Greenland Sea (Figs. A.1(b)-(d)). For all forcing fields, sensitivity in the Arctic Ocean and Subtropical Gyre is relatively small compared to the other regions. This can partly be an artifact of their large areas, over which (A.1) computes the average sensitivity. If sensitivity is only strong in subdomains of these regions, average sensitivities will still remain small.

A.2 Typical forcing anomalies

Figs. A.2(a)-(d) show the forcing anomalies $|\Delta F|$, which were used as a weighting in (2.7). The forcing anomalies are shown as the spatial average over each region \mathcal{R} , and for each month m :

$$\overline{\Delta F(m)}^{\mathcal{R}} = \frac{\int_{\mathcal{R}} |\Delta F(x, y, m)| dx dy}{\int_{\mathcal{R}} dx dy}. \quad (\text{A.2})$$

It is recalled that for each month m , the spatially varying $\Delta F(x, y, m)$ are the standard deviations of the departures from the seasonal cycle, using the statistics of the 20-year ECCOv4r2 forcing fields. In the Subpolar Gyre, the Greenland Sea and the Arctic Ocean, a considerable portion of the forcing anomalies is due to high sea-ice variability. To see this, 12 masks are created, where each mask identifies the area of high sea-ice variability in the respective month. A surface grid cell is considered as one of high sea-ice variability if the standard deviation of the sea-ice covered area fraction, seen as departure from the seasonal cycle, exceeds 10%. The statistics are again based on the 20-year ECCOv4r2 state estimate [Forget et al., 2015], which covers the years 1992 to 2011. As an example, Figs. A.2(e),(f) show standard deviations of the sea-ice covered area fraction and inferred masks for the months March and September. Part of the forcing anomalies can now be attributed to forcing anomalies due to high sea-ice variability, which is indicated by the dashed subareas in Figs. A.2(a)-(d). Forcing anomalies due to high sea-ice variability are visible within the Subpolar Gyre, the Greenland Sea, and the Arctic Ocean, (and moreover, a non-significant fraction of the Gulf Stream). As expected, sea-ice variability impacts freshwater flux variability the most. This is due to the direct effect that melting and freezing sea-ice has on surface freshwater flux, in addition to the effect of shielding the surface ocean from the atmosphere, which impacts variability of all air-sea fluxes.

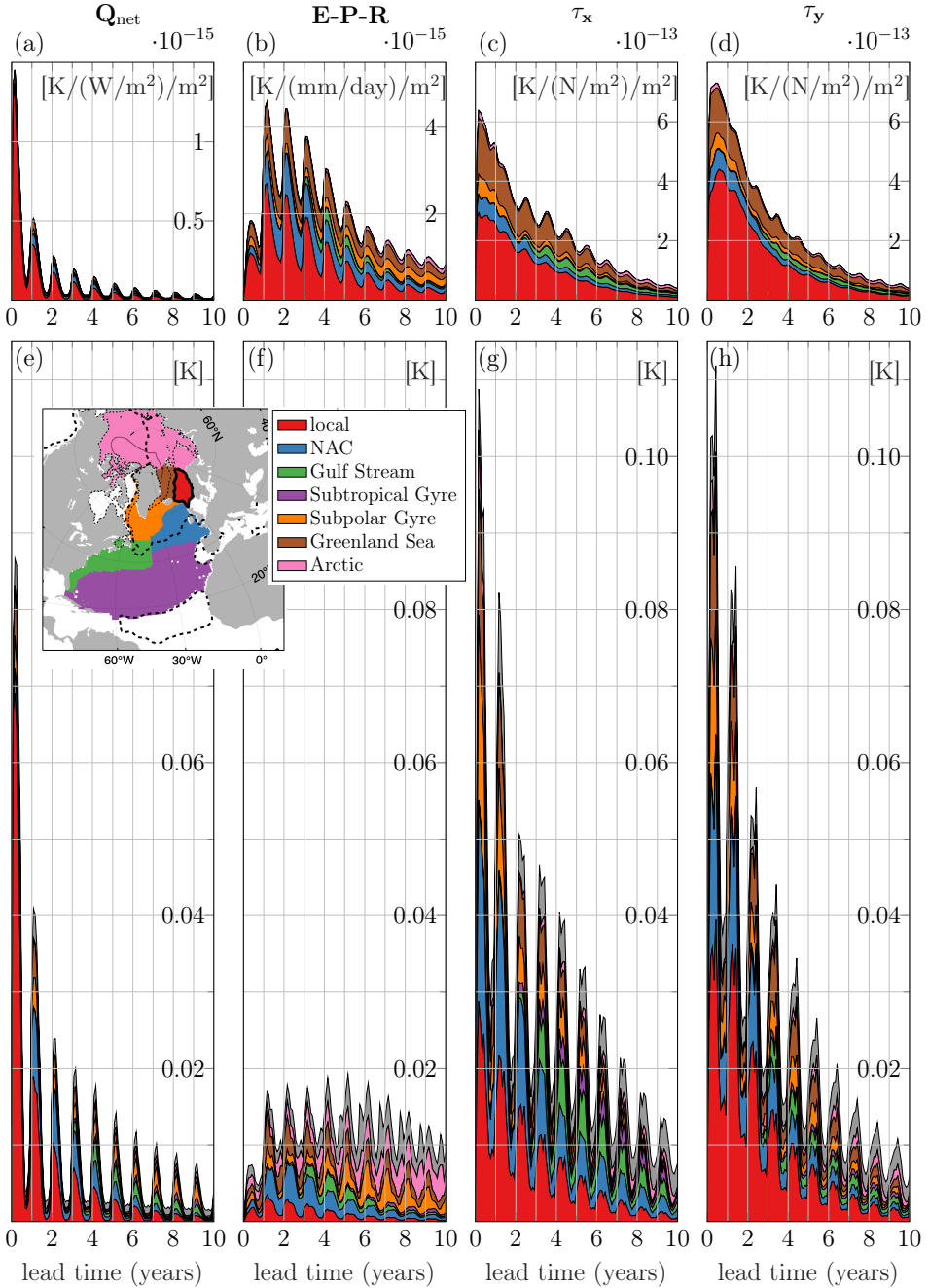


Figure A.1: (a)-(d) Average (unweighted) absolute sensitivity $\left| \frac{\partial \text{UOHCNS}}{\partial F} \right|^{\mathcal{R}}(t)$, defined in (A.1), and (e)-(h) weighted absolute sensitivity $|\Delta \text{UOHCNS}|^{\mathcal{R},F}(t)$, defined in (2.7) (and replotted from Fig. 2.12), of March Nordic Seas upper-ocean temperature to (a),(e) surface heat flux Q_{net} , (b),(f) surface freshwater flux E-P-R, (c),(g) zonal wind stress τ_x , and (d),(h) meridional wind stress τ_y .

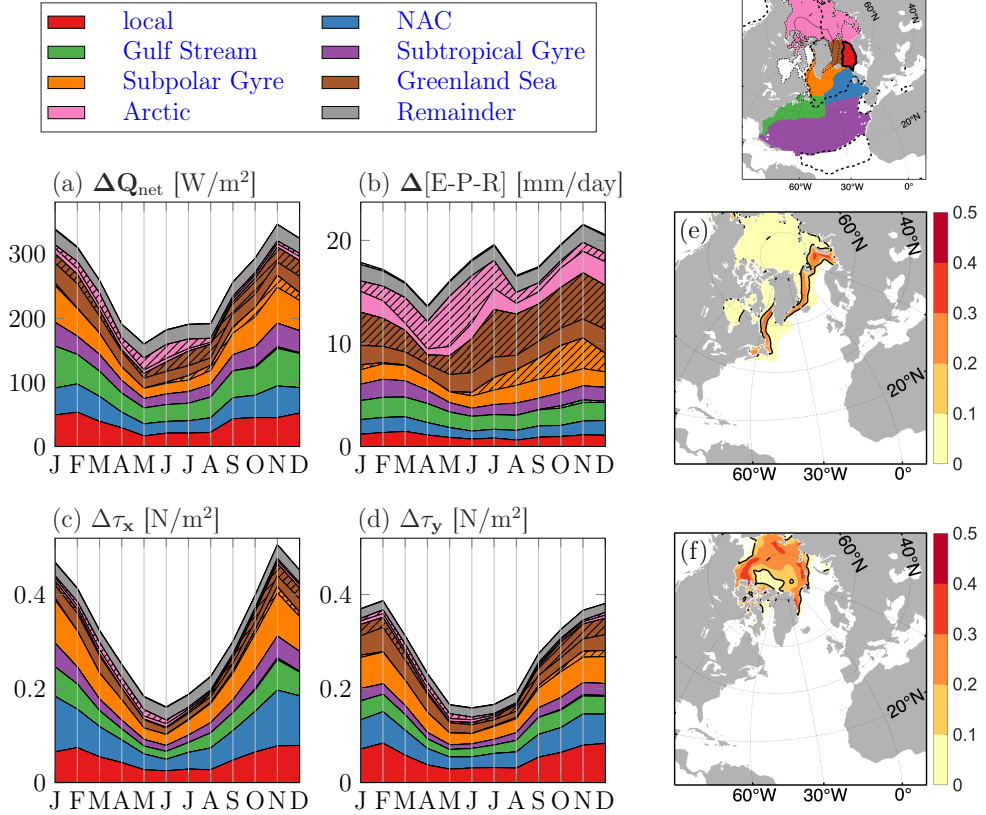


Figure A.2: **Left:** Typical forcing anomalies $|\Delta F|$ that are used as weights in (2.7) and Figs. A.1(e)-(h) (or Fig. 2.12). For each month m , the spatially varying $|\Delta F(x, y, m)|$ are the standard deviations of the departures from the seasonal cycle, using the statistics of the 20-year ECCOv4r2 adjusted forcing fields. (a)-(d) show the region-averages $\overline{\Delta F(m)}^{\mathcal{R}}$, defined in (A.2), for each region \mathcal{R} in the legend. The dashed areas show forcing anomalies due to high sea-ice variability. **Right:** Standard deviations of the sea-ice covered area fraction, seen as departures from the seasonal cycle, for the months (e) March and (f) September. High values indicate regions of high sea-ice variability in the respective months. The black lines mark the 10% contour lines. For each of the 12 months, the respective 10% contour line bounds the mask, which is applied to $|\Delta F(x, y, m)|$ in (a)-(d) to identify the dashed areas of high sea-ice variability.

A.3 High-sensitivity regions vs. high-impact regions

While the first row of Fig. A.1 shows unweighted sensitivities and therefore identifies “high-sensitivity regions”, the second row of Fig. A.1 replots the weighted sensitivities from Fig. 2.12, displaying “high-impact regions”. Unweighted and weighted sensitivities are now compared for each of the forcing components. A gain in relative importance of a region, when shifting from “sensitivity” to “impact”, can have two different causes: first, typical forcing anomalies being higher in that region compared to other regions, and second, a larger region size, which can lead to an integration of rather small effects to a significant contribution. The first aspect can be examined by means of Figs. A.2(a)-(d), which present forcing anomalies, and the second aspect by means of Fig. 2.11(b), which shows the geographical area of each region. Comparing the pairs (a) and (e), (b) and (f), (c) and (g), and (d) and (h) in Fig. A.1, four characteristics highlighting the topic “high/low sensitivity” vs. “high/low impact” are identified and attributed to one of the two possible causes.

- The *NAC and Gulf Stream* regions gain in importance relative to the local Nordic Seas region, when moving from “sensitivity” to “impact”, a fact seen for all 4 forcings. This characteristic is reflected by a “stretching” of the NAC and Gulf Stream region contributions compared to the local Nordic Seas contribution when passing from (a) to (e), from (b) to (f), from (c) to (g), and from (d) to (h) in Fig. A.1. The weighted sensitivities in Figs. A.1(e),(f),(g),(h) show that contributions from the three regions are of approximately the same magnitude (although there are exceptions, e.g., the local contribution for heat flux is still the dominating one on short time scales). Figs. A.2(a)-(d) reveal that this is *not* due to larger forcing anomalies in the Gulf Stream and NAC, but rather due to the fact that the Gulf Stream and NAC region are more than twice as big in size as the local Nordic Seas region (Fig. 2.11(b)). Therefore, the effect of the Nordic Seas being a high sensitivity region is partially counteracted by the fact that this region is of relatively small size.
- The *Subpolar Gyre and Greenland Sea* regions experience a large gain in relative importance for freshwater flux, when shifting from “sensitivity” to “impact”, i.e. from Fig. A.1(b) to Fig. A.1(f). This is due to much larger freshwater anomalies in these regions compared to regions that are not affected by high sea-ice variability (Fig. A.2(b)). For the other forcing components, a relative gain in importance of the Subpolar Gyre and the Greenland Sea is seen as well, although of minor magnitude, and is due to a combination of slightly larger forcing anomalies and the relatively large size of the Subpolar Gyre.
- The *Arctic* region strongly gains in relative importance, when passing from “sensitivity” to “impact”, for freshwater flux (from Fig. A.1(b) to Fig. A.1(f)) and wind stresses (from Fig. A.1(c) to Fig. A.1(g) and from Fig. A.1(d) to Fig. A.1(h)). This is due to the large size of the Arctic region (Fig. 2.11(b)), and for freshwater flux, in addition, due to large forcing anomalies in the Arctic (Fig. A.2(b)). Note that large freshwater flux anomalies make the Arctic a dominant player on increasingly longer timescales (Fig. A.1(f)).

- While sensitivities in the *Subtropical Gyre* are not discernible in the unweighted sensitivities (Figs. A.1(a)-(d)), weighting gives the Subtropical Gyre some prominence for zonal wind stress anomalies, visible in Fig. A.1(g). This is due to the large size of the Subtropical Gyre, which accumulates the rather small sensitivity to Rossby wave dynamics.

A.4 Seasonality

A.4.1 Seasonality of unweighted sensitivities

Unweighted sensitivities show a seasonal cycle (Figs. A.1(a)-(d)), which reflects the fact that the sensitivity of Nordic Seas upper-ocean heat content to local and remote forcing anomalies depends on the season in which the forcing anomalies occur. For heat and freshwater fluxes, sensitivity peaks in the winter months, roughly corresponding to lead times of integer (0, 1, 2, etc.) years, and displays minima in the summer months, corresponding to lead times of approximately 0.5, 1.5, etc. years (Figs. A.1(a),(b)). This shows that surface buoyancy flux anomalies are most effective when occurring in winter, when mixed layer depths are at their deepest and buoyancy anomalies can entrain deeper into the water column. Compared to sensitivity to buoyancy fluxes (Figs. A.1(a),(b)), sensitivity to momentum fluxes (Figs. A.1(c),(d)) shows a much smaller seasonality, with peaks in the summer months. A seasonal cycle with summer maxima for absolute sensitivity to wind stress is most pronounced in the local Nordic Seas, the Greenland Sea and the Gulf Stream regions. More analysis is required to fully understand the mechanism that underlies this feature.

A.4.2 Seasonality of forcing anomalies

Fig. A.2 shows the seasonality of typical forcing anomalies, where the latter were computed as the standard deviations of the departures from the seasonal cycle. Average heat flux and wind stress standard deviations show a qualitatively similar picture. The maximum is taken roughly between November and February, the northern hemisphere winter months, and the minimum approximately between May and August, the northern hemisphere summer months (Figs. A.2(a),(c),(d)). Since the defined regions are located in the Northern Hemisphere, the observed timing of maxima and minima reflects the fact that forcing anomalies are more variable in the winter than in the summer months (while the gray remainder term contains contributions from both hemispheres and is in fact constant throughout the year). The described seasonal cycle of the forcing standard deviations follows essentially the seasonal cycle of the (absolute) forcing mean values (not shown). Surface heat fluxes and wind stresses are of highest magnitude in winter when the temperature difference between the (warm) ocean and the (cold) atmosphere is greatest - and therefore the strongest air-sea heat fluxes are exchanged - and when the westerlies are most intense. The seasonal cycle of freshwater flux anomalies shows a different pattern. Two peaks throughout the annual cycle are visible in the Subpolar Gyre, the Greenland Sea, and the Arctic Ocean (Fig. A.2(b)). The dashed areas in Fig. A.2(b) demonstrate that the peak in summer is due to high sea-ice variability in summer, when melting sea-ice releases freshwater to the ocean and therefore

creates highly variable freshwater fluxes. The regions not affected by high sea-ice variability, that is, all non-dashed contributions, display weak seasonality of freshwater flux anomalies, with a maximum in March, similarly as seen for heat flux and wind stress anomalies.

A.4.3 Seasonality of weighted sensitivities

Seasonality of the weighted sensitivities, shown in Figs. A.1(e)-(h) (or Fig. 2.12) can now be explained. Weighted sensitivities to freshwater flux and wind stress display two peaks throughout the annual cycle (Figs. A.1(f)-(h)), while weighted sensitivity to heat flux shows only one peak, in winter time (Fig. A.1(e)). The double-peaks within the seasonal cycle of weighted sensitivities to freshwater flux (Fig. A.1(f)) are imprinted by the double-peaks within the seasonal cycle of freshwater flux anomalies in the Subpolar Gyre, the Greenland Sea, and the Arctic Ocean. In fact, the winter peaks in the seasonal cycles of unweighted sensitivities to freshwater flux (Fig. A.1(b)) and of freshwater flux anomalies (Fig. A.2(b)) are slightly offset, which is why a minor double-peak appears even for the regions not affected by high sea-ice variability. The double-peaks within the seasonal cycle of weighted sensitivities to wind stress (Figs. A.1(g),(h)) are due to the half-year-offset of summer maxima in the seasonal cycle of unweighted sensitivities to wind stress (Figs. A.1(c),(d)) and winter maxima in the seasonal cycle of wind stress anomalies (Figs. A.2(c),(d)). For heat flux, maxima in the seasonal cycles of unweighted sensitivities (Fig. A.1(a)) and forcing anomalies (Fig. A.2(a)) are attained at the same time, in winter, resulting in a single peak within the seasonal cycle of weighted sensitivities to heat flux (Fig. A.1(e)).

Appendix B

Uncertainty Quantification

B.1 Bayesian formulation of inverse problem

The Bayesian approach states an inverse problem as one of Bayesian inference over the space of unknown control variables, which are to be inferred from the data and the numerical model. The solution is given by a posterior probability density function $\pi_{\text{post}}(\mathbf{x}|\mathbf{y})$, computed by Bayes' theorem

$$\pi_{\text{post}}(\mathbf{x}|\mathbf{y}) \propto \pi_{\text{like}}(\mathbf{y}|\mathbf{x}) \pi_{\text{prior}}(\mathbf{x}). \quad (\text{B.1})$$

The posterior probability distribution assigns to any multivariate value \mathbf{x} of the control space a probability $\pi_{\text{post}}(\mathbf{x}|\mathbf{y})$ that \mathbf{x} is the true value that gave rise to the observed data \mathbf{y} .

The regularization term in the deterministic inverse problem (3.1) takes on the role of the prior distribution, which is chosen as the Gaussian $\mathcal{N}(\mathbf{x}_0, \mathbf{\Gamma}_{\text{prior}})$, i.e.,

$$\pi_{\text{prior}}(\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{\Gamma}_{\text{prior}}^{-1} (\mathbf{x} - \mathbf{x}_0)\right). \quad (\text{B.2})$$

One furthermore assumes that the noise

$$\mathbf{e} = \mathbf{y} - \mathbf{f}(\mathbf{x}),$$

which is the difference between the actual observations \mathbf{y} and the observation counterparts $\mathbf{f}(\mathbf{x})$ simulated by the model, follows a Gaussian distribution $\mathcal{N}(0, \mathbf{\Gamma}_{\text{noise}})$. The weight matrix $\mathbf{\Gamma}_{\text{noise}}$ for the model-data misfit term in the deterministic inverse problem (3.1) therefore takes on the role of the noise distribution. Here, it shall be noted that the observational noise is not only due to measurement errors, but also due to representation errors in the model, and model errors. Lumping measurement errors and representation/model errors together can be misleading because they are of very different character. For example, a model bias does *not* correspond to zero mean, and depends on \mathbf{x} , while observation errors do not share these properties [Chapter 2 of [Stuart, 2010](#)].

The likelihood can then be modeled as

$$\pi_{\text{like}}(\mathbf{y}|\mathbf{x}) = \pi(\mathbf{y} - \mathbf{f}(\mathbf{x})) \propto \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{f}(\mathbf{x}))^T \mathbf{\Gamma}_{\text{noise}}^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{x}))\right). \quad (\text{B.3})$$

Combining (B.1), (B.2) and (B.3), the posterior density function computes as

$$\pi_{\text{post}}(\mathbf{x}|\mathbf{y}) \propto \exp \left(-\frac{1}{2} \left[(\mathbf{y} - \mathbf{f}(\mathbf{x}))^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{x})) + (\mathbf{x} - \mathbf{x}_0)^T \boldsymbol{\Gamma}_{\text{prior}}^{-1} (\mathbf{x} - \mathbf{x}_0) \right] \right). \quad (\text{B.4})$$

A comparison with the deterministic formulation of the inverse problem via the least-squares cost function (3.1) shows that $\pi_{\text{post}}(\mathbf{x}|\mathbf{y}) \propto e^{-J(\mathbf{x})}$. Hence, the deterministic cost function J is the negative log-posterior in the Bayesian formulation. Furthermore, the deterministic solution \mathbf{x}_{min} is the Maximum a Posteriori (MAP) point, i.e., the most likely solution in the Bayesian framework. For more details, the reader is referred to the books by Tarantola [2005], Stuart [2010], or Law et al. [2015].

B.2 Gaussian approximation of the posterior

To make the computation of the posterior probability density function (B.4) computationally tractable, it is a common approach to linearize the non-linear observation operator \mathbf{f} about \mathbf{x}_{min} [e.g., Bui-Thanh et al., 2012]. This yields

$$\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\mathbf{x}_{\text{min}}) + \mathbf{A}(\mathbf{x} - \mathbf{x}_{\text{min}}),$$

where $\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}|_{\mathbf{x}_{\text{min}}}$ is the Jacobian matrix. The posterior distribution (B.4) then becomes

$$\pi_{\text{post}}(\mathbf{x}|\mathbf{y}) \propto C \cdot \exp \left(-\frac{1}{2} \left[(\mathbf{x} - \mathbf{x}_{\text{min}})^T (\mathbf{A}^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \boldsymbol{\Gamma}_{\text{prior}}^{-1})^{-1} (\mathbf{x} - \mathbf{x}_{\text{min}}) \right] \right), \quad (\text{B.5})$$

where C is a constant factor, given by

$$C = \exp \left(-\frac{1}{2} \left[(\mathbf{y} - \mathbf{f}(\mathbf{x}_{\text{min}}))^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{x}_{\text{min}})) + (\mathbf{x}_{\text{min}} - \mathbf{x}_0)^T \boldsymbol{\Gamma}_{\text{prior}}^{-1} (\mathbf{x}_{\text{min}} - \mathbf{x}_0) \right] \right),$$

and can therefore be absorbed by the proportionality \propto . The right hand side of (B.5) describes a Gaussian $\mathcal{N}(\mathbf{x}_{\text{min}}, \boldsymbol{\Gamma}_{\text{post}})$ with mean \mathbf{x}_{min} and covariance matrix $\boldsymbol{\Gamma}_{\text{post}} = (\mathbf{A}^T \boldsymbol{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \boldsymbol{\Gamma}_{\text{prior}}^{-1})^{-1}$.

B.3 Forward uncertainty propagation

Let $\text{QoI}(\mathbf{x})$ denote the scalar function that describes how a given quantity of interest depends on the control variables \mathbf{x} through the model dynamics. Consistent with the linearization of the observation operator \mathbf{f} about \mathbf{x}_{min} in (3.2), the function $\text{QoI}(\mathbf{x})$ is linearized about \mathbf{x}_{min} :

$$\text{QoI}(\mathbf{x}) \approx \text{QoI}(\mathbf{x}_{\text{min}}) + \frac{\partial(\text{QoI})}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{\text{min}}} (\mathbf{x} - \mathbf{x}_{\text{min}}). \quad (\text{B.6})$$

The posterior distribution of the Bayesian solution of the inverse problem, $\pi_{\text{post}}(\mathbf{x}|\mathbf{y})$, is approximately Gaussian, given by $\mathcal{N}(\mathbf{x}_{\text{min}}, \boldsymbol{\Gamma}_{\text{post}})$. A forward propagation of the posterior uncertainty (green arrow) along path (2) in Fig. 3.1 leads to a posterior Gaussian

distribution for $\text{QoI}(\mathbf{x})$, too, since (B.6) describes an affine transformation. The distribution is given by $\mathcal{N}(\text{QoI}(\mathbf{x}_{\min}), \sigma_{\text{post}}^2)$, where σ_{post}^2 is the (scalar) variance, given by the projection

$$\sigma_{\text{post}}^2 = \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}}^T \mathbf{\Gamma}_{\text{post}} \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}}. \quad (\text{B.7})$$

Similarly, the prior distribution of $\text{QoI}(\mathbf{x})$ is obtained by a forward uncertainty propagation of the Gaussian prior $\mathcal{N}(\mathbf{x}_0, \mathbf{\Gamma}_{\text{prior}})$ (orange arrow) along path (2) in Fig. 3.1. This leads to a prior Gaussian distribution for $\text{QoI}(\mathbf{x})$, given by $\mathcal{N}(\text{QoI}(\mathbf{x}_0), \sigma_{\text{prior}}^2)$ with

$$\sigma_{\text{prior}}^2 = \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_0}^T \mathbf{\Gamma}_{\text{prior}} \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_0} \quad (\text{B.8})$$

$$\approx \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}}^T \mathbf{\Gamma}_{\text{prior}} \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}} \quad (\text{B.9})$$

$$= \left\| \mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}} \right\|^2. \quad (\text{B.10})$$

The approximation in (B.9) is legitimate if $\text{QoI}(\mathbf{x})$ depends approximately linear on \mathbf{x} . Indeed, if $\text{QoI}(\mathbf{x})$ was a linear function of \mathbf{x} , the sensitivity vector $\frac{\partial(\text{QoI})}{\partial \mathbf{x}}$ would not depend on its point of evaluation, i.e., it would be the same at \mathbf{x}_0 and \mathbf{x}_{\min} . Avoiding the computation of the sensitivity of the QoI twice leads to the following compact formula for the relative uncertainty reduction:

$$\frac{\sigma_{\text{prior}}^2 - \sigma_{\text{post}}^2}{\sigma_{\text{prior}}^2} = \frac{1}{\sigma_{\text{prior}}^2} \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1} \left\langle \mathbf{\Gamma}_{\text{prior}}^{1/2} \left[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \right]_{|\mathbf{x}_{\min}}^T, \mathbf{v}_i \right\rangle^2 = \sum_{i=1}^{M'} \frac{\lambda_i}{\lambda_i + 1} \langle \mathbf{q}, \mathbf{v}_i \rangle^2, \quad (\text{B.11})$$

where $\{(\mathbf{v}_i, \lambda_i)\}$ denote the eigenvectors and eigenvalues of the non-dimensionalized misfit Hessian (3.5), \mathbf{q} the QoI target direction (3.10), and $\langle \cdot, \cdot \rangle$ the Euclidean inner product in the control space, i.e., \mathbb{R}^N . The first equality in (B.11) makes use of (B.9), (B.7), and (3.6), and the second equality uses (B.10).

B.4 Linearized vs. full Hessians

Denoting $\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}|_{\mathbf{x}_{\min}}$, the matrix

$$\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \mathbf{\Gamma}_{\text{prior}}^{-1} \quad (\text{B.12})$$

is the Hessian of the cost function J (3.1), evaluated at \mathbf{x}_{\min} , if one uses the linearization (3.2) of the observation operator \mathbf{f} . Therefore, (B.12) is also referred to as the *linearized Hessian*, the ‘first-order Hessian’, or the ‘Gauss-Newton Hessian’ [e.g., Chen, 2011]. The *full Hessian* (or ‘second-order Hessian’) of the cost function J , evaluated at \mathbf{x}_{\min} , is given by

$$\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2}|_{\mathbf{x}_{\min}} \mathbf{\Gamma}_{\text{noise}}^{-1} [\mathbf{f}(\mathbf{x}_{\min}) - \mathbf{y}] + \mathbf{\Gamma}_{\text{prior}}^{-1}, \quad (\text{B.13})$$

with an additional term that contains the second derivatives $\frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2}|_{\mathbf{x}_{\min}}$ of the observation operator. The first two terms in (B.13) comprise the full misfit Hessian, i.e., the full Hessian of J_{misfit} in (3.1), evaluated at \mathbf{x}_{\min} . Note that the linearized and the full Hessian are identical if (i) the observation operator $\mathbf{f}(\mathbf{x})$ is linear, or (ii) the residual misfit between the simulated observations $\mathbf{f}(\mathbf{x}_{\min})$ and the actual observations \mathbf{y} is zero.

In this thesis, the posterior probability density function $\pi_{\text{post}}(\mathbf{x}|\mathbf{y}) \propto e^{-J(\mathbf{x})}$ is locally approximated by the Gaussian $\mathcal{N}(\mathbf{x}_{\min}, \mathbf{\Gamma}_{\text{post}})$, with mean equal to the MAP point \mathbf{x}_{\min} , and $\mathbf{\Gamma}_{\text{post}}$ given by the inverse of the linearized Hessian (B.12) of J . Note that, alternatively, one could use the local Gaussian approximation $\mathcal{N}(\mathbf{x}_{\min}, \hat{\mathbf{\Gamma}}_{\text{post}})$, where $\hat{\mathbf{\Gamma}}_{\text{post}}$ is given by the inverse of the full Hessian (B.13) of J . Both versions have been used in past UQ-related work. For instance, Thacker [1989], Losch and Wunsch [2003], Moore et al. [2011a, 2012, 2011b], Bui-Thanh et al. [2012], and Kaminski et al. [2015, 2018] use the linearized Hessian. On the other hand, Kalmikov and Heimbach [2014, 2018], Isaac et al. [2015], and Alexanderian et al. [2016] use the full Hessian.

The linearized Hessian has the advantage that it is a positive-semidefinite matrix, while the full Hessian does not always have this property in practice (e.g., when \mathbf{x}_{\min} is not quite a local minimum). Positive-semidefiniteness is a desirable property not only for optimization, but also for UQ-based observing system design because it guarantees that adding new observations can never increase uncertainties, only decrease them (or leave them unchanged). Another convenient property of the linearized Hessian is the one discussed in Section 3.4, i.e., that the matrix only depends on the observation locations, times, types and uncertainties, but not the actual measurement values. This has important implications for observing system design: *future* observing systems can easily be evaluated, before the measurements are available. In contrast, the full Hessian needs the actual measurement values (since the vector \mathbf{y} appears in (B.13)). It is noted that, nevertheless, there are strategies to assess future observing systems by means of the full Hessian - but they need to be more involved [e.g., Alexanderian et al., 2016].

Seemingly, the method of using the full Hessian with second derivatives of the observation operator should provide a better approximation than the linearized Hessian. Nevertheless, it has also been argued that the numerical noise is increased so substantially by including the second derivatives of the observation operator that it outweighs the increased sophistication in the Hessian representation, at least for pre-conditioning purposes in the minimization process [see references in Chen, 2011]. One could also argue that forward uncertainty propagation onto a scalar quantity of interest only uses *first* derivatives of the scalar function $\text{QoI}(\mathbf{x})$ (see equations (B.6) - (B.10) in Appendix B.3) - so why “bother” and use second derivatives of the observation operator $\mathbf{f}(\mathbf{x})$? The few studies that have compared the performance of linearized and full Hessians for practical oceanographic applications work within idealized settings. For instance, for a small-size shallow-water equations model, Cioaca et al. [2013] find that the linearized Hessian provides results of acceptable accuracy with respect to the full Hessian.

Until future research and applications shed more light on linearized vs. full Hessians, I conclude that - as long as the observation operator is not too non-linear - either the

linearized or the full Hessian should provide a good basis for a Gaussian approximation (with covariance equal to either $\mathbf{\Gamma}_{\text{post}}$ or $\hat{\mathbf{\Gamma}}_{\text{post}}$). Whether to use the linearized or the full Hessian may be a matter of whichever Hessian is more convenient to compute in the model framework under consideration. Linearized Hessians may be preferred if only first-order adjoints are available, or if one even relies on finite difference approximations of the Hessian (as for instance done in Kaminski et al. [2018]). If one uses automatic differentiation tools that can generate second-order adjoints, as for instance the commercial tool TAF [Giering and Kaminski, 1998], the full Hessian might be the Hessian of choice. If the system is strongly non-linear and the posterior probability distribution is multimodal, *both* Gaussian approximations (no matter whether based on the linearized or the full Hessian) will perform poorly when seen as a global approximation. However, in the vicinity of the local minima (corresponding to the different modes of the distribution), a local Gaussian approximation based on either the linearized or full Hessian should still be useful.

Appendix C

Prior information in ECCOv4r3

For the purpose of a simpler presentation, the choice of the control vector \mathbf{x} and corresponding prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$ in this work deviates in to major aspects from the ECCOv4r3 setup: I omit (i) time-variable bi-weekly adjustments of the atmospheric control variables and (ii) [Weaver and Courtier \[2001\]](#) smoothing for all prior uncertainties (cf. Table 4.2). Omitting (ii) results in a diagonal prior covariance matrix $\mathbf{\Gamma}_{\text{prior}}$, where the diagonal entries are adopted from ECCOv4r3. The impacts of these two departures from the ECCOv4r3 setup are discussed in the following, by re-investigating the results for three different choices for the inverse modeling framework. Setup (a) is equal to the one chosen in this work. Setup (b) includes *all* ECCOv4r3 control variables, i.e. adds the time-variable atmospheric controls, but still uses a diagonal prior. Setup (c) uses the full ECCOv4r3 setup, i.e., setup (b), but with [Weaver and Courtier \[2001\]](#) smoothing applied to all control variables.

C.1 Time-variable atmospheric controls

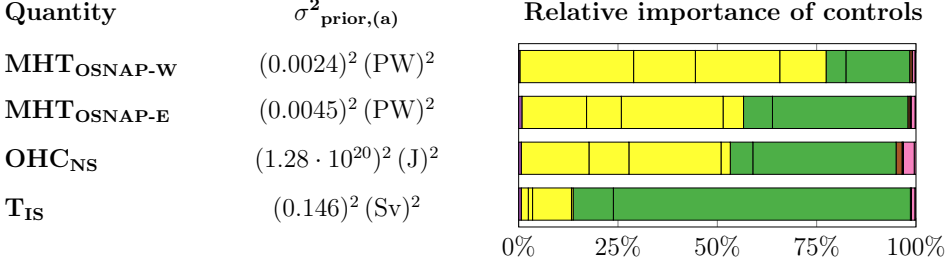
Excluding the time-variable atmospheric forcing variables from the control vector has a very minor impact on the results in this work. For instance, the minor impact is visible in Figs. C.1(a) and (b), which compares prior uncertainties for the setups (a) and (b). The second column in Fig. C.1 shows the values for the prior uncertainties σ_{prior}^2 in the 5-year-mean observed and unobserved QoIs $\text{MHT}_{\text{OSNAP-W}}$, $\text{MHT}_{\text{OSNAP-E}}$, OHC_{NS} , and T_{IS} (defined in (4.2), (4.3), (4.5), and (4.6)). The values in Fig. C.1(a) coincide with the values in Table 4.3. Each prior uncertainty σ_{prior}^2 is computed by projecting the prior covariance matrix $\mathbf{\Gamma}_{\text{prior}}$ onto the respective QoI (see (B.8)), which, in the case of a diagonal prior, as in the setups (a) and (b), simplifies to the projection (4.7). For each QoI, the prior uncertainty $\sigma_{\text{prior,(b)}}^2$ in setup (b) is larger than the prior uncertainty $\sigma_{\text{prior,(a)}}^2$ in setup (a), due to the additional contributions by the time-variable atmospheric control variables:

$$\sigma_{\text{prior,(b)}}^2 = \sigma_{\text{prior,(a)}}^2 + \sum_{\substack{i|x_i \in \text{time-variable} \\ \text{atmospheric ctrls}}} \left(\frac{\partial(\text{QoI})}{\partial x_i} \cdot \Delta x_i \right)^2.$$

However, for each QoI, $\sigma_{\text{prior,(b)}}^2$ is only <10% higher than $\sigma_{\text{prior,(a)}}^2$ (second column of Figs. C.1(a),(b)). The underlying reason is that the 4 studied *time-mean* QoIs are much

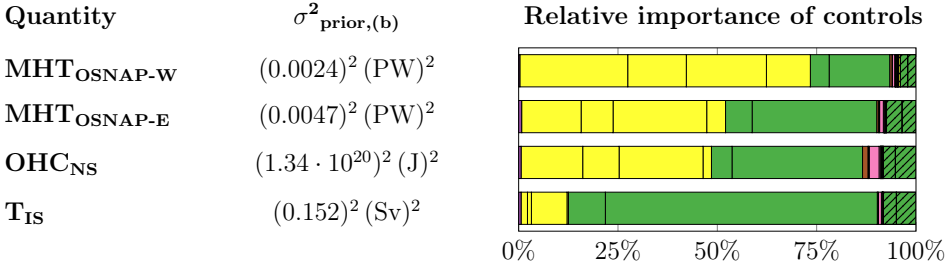
(a) Ctrls: ECCOv4r3 without time-variable atmospheric controls

Prior: Diagonal ECCOv4r3 (i.e., ECCOv4r3, but no WC01 smoothing)



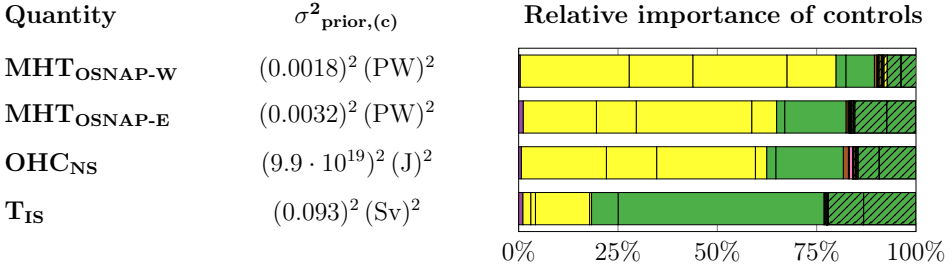
(b) Ctrls: All ECCOv4r3 controls

Prior: Diagonal ECCOv4r3 (i.e., ECCOv4r3, but no WC01 smoothing)



(c) Ctrls: All ECCOv4r3 controls

Prior: Full ECCOv4r3 (incl. WC01 smoothing)



third block in Table 4.2:
 first and second block in Table 4.2:
 time-variable atmospheric controls:

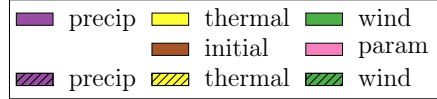


Figure C.1: (a),(b),(c) present three different choices for the control vector \mathbf{x} and corresponding prior uncertainties $\mathbf{\Gamma}_{\text{prior}}$, examined through the lens of projected scalar prior uncertainties σ_{prior}^2 (second column) for the four shown *time-mean* model quantities (named in the first column). The third column shows the relative importance of the control variables in contributing to prior uncertainty σ_{prior}^2 . (a) reflects the setup chosen in this work, with \mathbf{x} and diagonal $\mathbf{\Gamma}_{\text{prior}}$ as shown in Table 4.2. (b) includes *all* ECCOv4r3 control variables, i.e., adds the time-variable atmospheric control variables (hatched in third column) to \mathbf{x} , but still uses a diagonal prior $\mathbf{\Gamma}_{\text{prior}}$, composed of the ECCOv4r3 prior uncertainties. (c) uses the full ECCOv4r3 setup, i.e., setup (b), but with Weaver and Courtier [2001] (WC01) smoothing applied to all control variables.

more sensitive to changes in the time-mean of the atmospheric forcing variables than to time-variable changes. The dominant role of time-mean atmospheric controls for time-mean QoIs can also be seen in the third column of Fig. C.1, which shows the relative importance of the control variables in contributing to the prior uncertainties $\sigma_{\text{prior,(a)}}^2$ (Fig. C.1(a)) and $\sigma_{\text{prior,(b)}}^2$ (Fig. C.1(b)). The contributions by the time-variable atmospheric control variables, hatched in Fig. C.1(b), are much smaller than the contributions by the time-mean atmospheric control variables. (Fig. 6.8 in Chapter 6 shows that time-variable atmospheric controls may play a much larger role for QoIs that are averaged over shorter time periods than 5 years, e.g., one month.) Note that that the 4 bar charts in Fig. C.1(a) are the same as in Figs. 4.5(e),(f) and Figs. 4.8(e),(f), respectively.

As a result of the minor importance of the time-variable atmospheric control variables in contributing to the prior uncertainties σ_{prior}^2 , the effective proxy potentials (or relative uncertainty reductions) $\tilde{\Delta}\sigma^2$ for the QoIs by the OSNAP observations are essentially the same for setups (a) and (b) (Figs. C.2(a) vs. (b)). In Fig. C.2, the curves for $\tilde{\Delta}\sigma^2[(r/b)^2]$ are determined by two aspects: the hypothetical proxy potentials (the values for $\tilde{\Delta}\sigma^2[0]$) and the prior-to-noise ratios (the decay of the curves). Both aspects, hypothetical proxy potentials and prior-to-noise ratios, essentially coincide for setups (a) and (b). Consequently, the results for the effective OSNAP proxy potential obtained in this work are *not* affected by the omission of ECCOv4r3 time-variable atmospheric control variables.

C.2 Smoothing

In the ECCOv4r3 inverse modeling framework, Weaver and Courtier [2001] smoothing is used for the prior covariance matrix to mimic prior cross-correlations between control variables that represent close-by grid points in the same field [Forget et al., 2015]. Fig. C.3 shows a subset of the prior standard deviation fields used in ECCOv4r3. Constructing the ECCOv4r3 prior covariance matrix consists of two steps. First, inserting the squares of the prior standard deviations at each model grid point as diagonal entries into an $N \times N$ matrix, where N denotes the length of the control vector, results in a diagonal matrix $\mathbf{\Gamma}_{\text{prior}}$. This diagonal matrix $\mathbf{\Gamma}_{\text{prior}}$ is used in setups (a) and (b) (where setup (a) only uses standard deviations from Figs. C.3(a),(c)). To obtain the full ECCOv4r3 prior covariance matrix $\hat{\mathbf{\Gamma}}_{\text{prior}}$, a second step consists of modifying the diagonal matrix $\mathbf{\Gamma}_{\text{prior}}$ by the Weaver and Courtier [2001] smoother (or diffusion operator), with smoothing scale set to three grid points [Forget et al., 2015]. The smoother introduces off-diagonal entries, which represent covariances between control variables that represent close-by grid points in the same field.

It is noted that the high-dimensional prior matrix $\hat{\mathbf{\Gamma}}_{\text{prior}}$ is never constructed in practice. The prior enters both the optimization process and the identification of proxy potential (via uncertainty quantification) only in the form of matrix-vector multiplications. Left-multiplication of a sensitivity vector $[\frac{\partial(\text{QoI})}{\partial\mathbf{x}}]^T$ with the matrix $\mathbf{\Gamma}_{\text{prior}}^{1/2}$ (or $\hat{\mathbf{\Gamma}}_{\text{prior}}^{1/2}$) is referred to as *prior-weighting* in Sections 4.3.1 and 4.3.3 (and in Section 3.3 for the general case). In practice, the two steps described above translate to the follow-

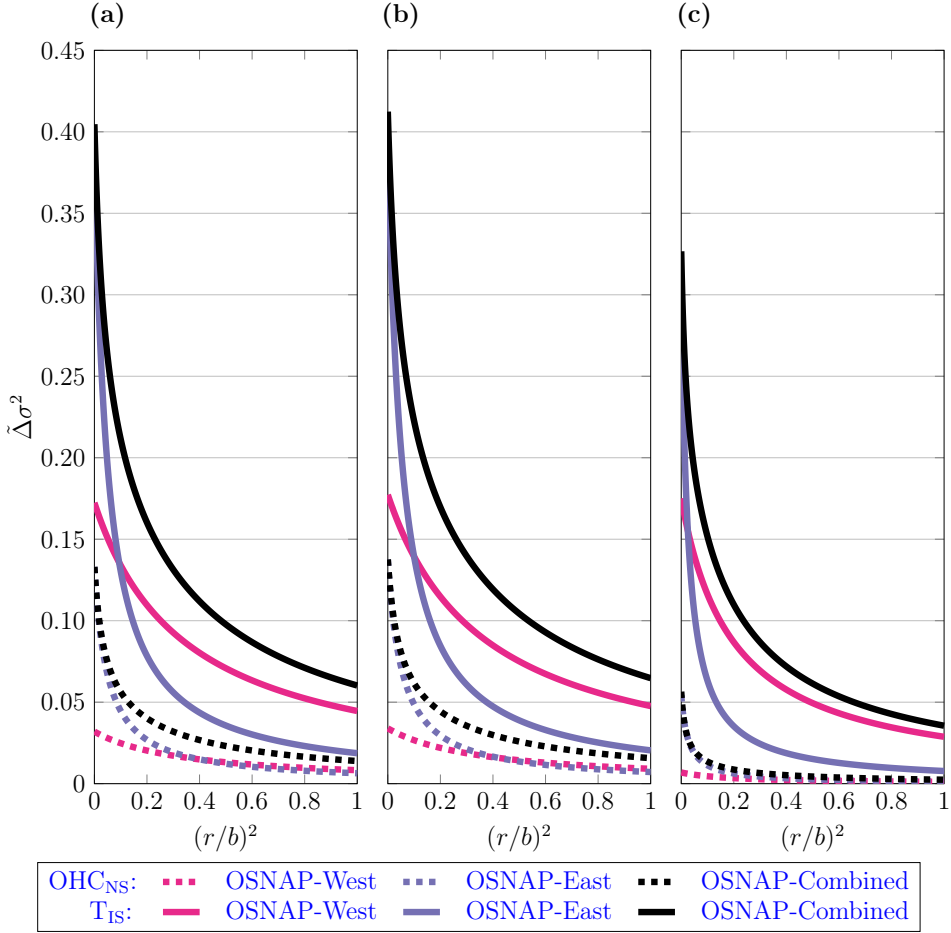


Figure C.2: Effective proxy potential (or relative uncertainty reduction) $\tilde{\Delta}\sigma^2$ for the unobserved QoIs OHC_{NS} (dashed curves) and T_{IS} (solid curves) by the three OSNAP observing systems, shown as a function of the ratio $(r/b)^2$. The subpanels (a),(b),(c) present the three different choices for \mathbf{x} and $\mathbf{\Gamma}_{\text{prior}}$ which were also shown in Fig. C.1. Subpanel (a) reflects the setup chosen in this work and replots all curves from Figs. 4.13(a),(c). Subpanel (b) includes *all* ECCOv4r3 control variables, i.e., adds the time-variable atmospheric control variables to \mathbf{x} , but still uses a diagonal prior $\mathbf{\Gamma}_{\text{prior}}$, composed of the ECCOv4r3 prior uncertainties. Subpanel (c) uses the full ECCOv4r3 setup, i.e., setup (b), but with Weaver and Courtier [2001] smoothing applied to all control variables.

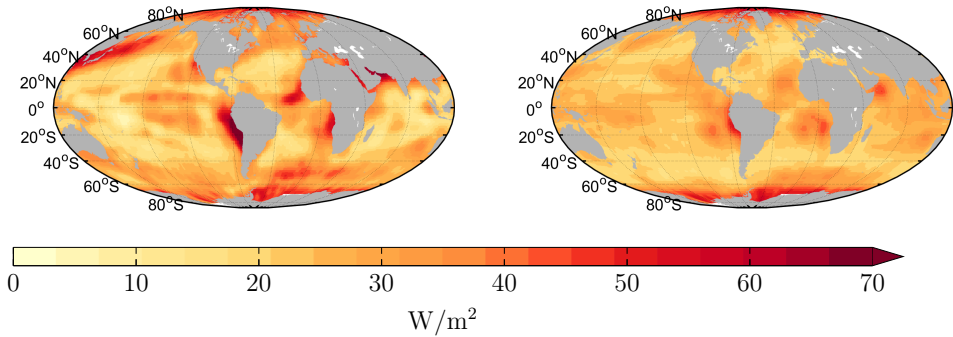
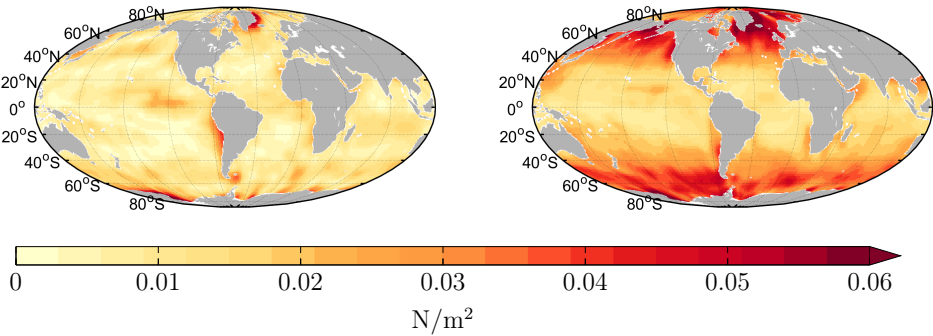
(a) $\Delta R_{SW\downarrow}$ time-mean(b) $\Delta R_{SW\downarrow}$ time-variable(c) $\Delta \tau_y$ time-mean(d) $\Delta \tau_y$ time-variable

Figure C.3: ECCOv4r3 prior standard deviations for the atmospheric forcing fields (a),(b) shortwave downward radiation ($R_{SW\downarrow}$), and (c),(d) meridional wind stress (τ_y). The shown prior standard deviations are for (a),(c) time-mean adjustments, and (b),(d) time-variable adjustments of the respective atmospheric fields, and computed based upon the spread of available atmospheric reanalysis products [Chaudhuri et al., 2012]. Maps (a) and (c) are the global versions of Figs. 4.4(a) and (b), respectively.

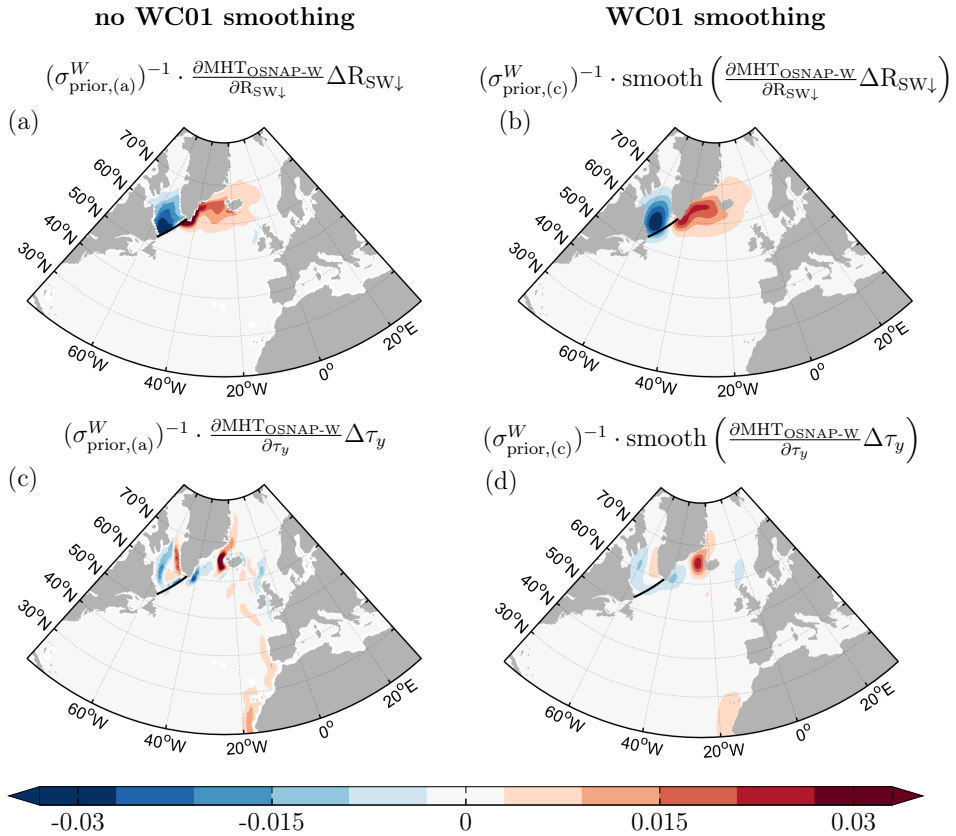


Figure C.4: Effect of the [Weaver and Courtier \[2001\]](#) (WC01) smoothing. The maps show normalized prior-weighted sensitivities of time-mean heat transport across OSNAP-West ($\text{MHT}_{\text{OSNAP-W}}$) to (a),(b) shortwave downward radiation ($R_{\text{SW}\downarrow}$) and (c),(d) meridional wind stress (τ_y). Prior-weighting and normalization are performed using (a),(c) all ECCOv4r3 controls but *no* WC01 smoothing and (b),(d) the full ECCOv4r3 setup including WC01 smoothing.

ing two steps for prior-weighting. First, the i^{th} entry of the given vector is multiplied by the prior standard deviation Δx_i that corresponds to the i^{th} control variable, resulting in the vector $[\frac{\partial(\text{QoI})}{\partial \mathbf{x}} \cdot \Delta \mathbf{x}]^T$. This vector is the prior-weighted sensitivity vector used for Figs. 4.5 and 4.8 in Sections 4.3.1 and 4.3.3. The second step consists in applying the Weaver and Courtier [2001] smoother to the vector obtained after the first step, with smoothing scale set to three grid points. The “smoothing” or “smearing” effect of the second step is visible in Fig. C.4: shown are normalized prior-weighted sensitivities of $\text{MHT}_{\text{OSNAP-W}}$, where prior-weighting and normalization are done using setup (a) (in Figs. 4.5(a),(c)) and setup (c) (in Figs. 4.5(b),(d)). Note that Figs. C.4(a),(c) are a replot of Figs. 4.5(a),(c).

For setup (c), the projection of the prior uncertainties $\hat{\Gamma}_{\text{prior}}$, from the control variables onto a chosen QoI computes as

$$\sigma_{\text{prior,(c)}}^2 = \sum_{i=1}^N \left[\text{smooth} \left(\frac{\partial(\text{QoI})}{\partial x_j} \cdot \Delta x_j \right) \right]_i^2. \quad (\text{C.1})$$

For the 4 time-mean QoIs studied in this work, the values for $\sigma_{\text{prior,(c)}}^2$ are shown in the second column of Fig. C.1(c). The prior uncertainties $\sigma_{\text{prior,(c)}}^2$ are considerably smaller than the respective prior uncertainties for setups (a) and (b) (Figs. C.1(a),(b)), i.e., the setups without Weaver and Courtier [2001] smoothing. This is because smoothing eliminates a significant portion of the prior-weighted wind sensitivities, supported by the third column of Fig. C.1(c), which shows that time-mean wind stress has lost a significant portion of its relative importance in contributing to $\sigma_{\text{prior,(c)}}^2$ (compared to Figs. C.1(a),(b)). Smoothing sensitivities leads to cancellations where non-smoothed sensitivities show a sensitivity dipole. Sensitivity dipoles are a typical feature of *wind* sensitivities because they characterize wind-driven Ekman pumping, e.g., in the Labrador Sea in Fig. C.4(c). The sensitivity dipole is much weaker in Fig. C.4(d), after Weaver and Courtier [2001] smoothing is applied, due to partial cancellation of close-by positive and negative sensitivities. A second reason for the partial elimination of wind sensitivities is that wind sensitivity patterns are often characterized by narrow sensitivity bands of strong sensitivity amplitude, such as between Greenland and Iceland in Fig. C.4(c). Concentrated sensitivity of strong amplitude contributes particularly effectively to prior uncertainty, i.e., the (squared) Euclidean norm in (C.1). Applying the Weaver and Courtier [2001] smoother to the left column of Fig. C.4 causes sensitivity to spread out spatially. While the spatial spreading leads to a considerable diffusion of wind sensitivities (Fig. C.4(d)) and, thus, diffusion of prior uncertainty caused by wind stress, the diffusion of prior uncertainties caused by thermal forcing fields is less pronounced (Fig. C.4(b)). This is because thermal sensitivities are characterized by spatially widespread patterns, such as in the Irminger and Labrador Seas in Fig. C.4(a), such that diffusion mostly (re-)distributes sensitivity *within* the sensitivity regions (rather than across the boundaries), due to the relatively shorter boundaries of the thermal sensitivity patterns than of the wind sensitivity patterns.

The relatively weaker wind stress sensitivities in setup (c) lead to smaller values for the hypothetical proxy potentials, reflected by smaller values for $\tilde{\Delta}\sigma^2[0]$ in Fig. C.2(c),

compared to Figs. C.2(a),(b). Indeed, wind sensitivities were shown to be the major player in causing effective projections of (normalized) prior-weighted sensitivities of QoIs onto (normalized) prior-weighted sensitivities of OSNAP quantities. Moreover, the smaller overall values for the prior uncertainties $\sigma_{\text{prior,(c)}}^2$ lead to smaller prior-to-noise ratios for setup (c), reflected in more rapidly decaying curves in Fig. C.2(c) than in Figs. C.2(a),(b). Smaller hypothetical proxy potentials and prior-to-noise ratios cause a less effective proxy potential (or relative uncertainty reduction) $\tilde{\Delta}\sigma^2$ for setup (c) than for the setups (a) and (b) (Fig. C.2).

The findings that [Weaver and Courtier \[2001\]](#) smoothing affects

- proxy potential significantly,
- thermal sensitivities vs. wind sensitivities in very dissimilar ways,

are the motivation for Chapter D. In Chapter D, it is investigated whether [Weaver and Courtier \[2001\]](#) smoothing in ocean state estimation is a favorable choice, and if yes, whether it should be applied to all atmospheric forcing fields in the same fashion.

Appendix D

Spatial Scales of Atmospheric Adjustment in Ocean State Estimation

D.1 Introduction

The diffusion operator approach of [Weaver and Courtier \[2001\]](#) (hereafter WC01) is a widely employed technique in ocean state estimation and data assimilation, used to model prior error covariance matrices [e.g., [Moore et al., 2011a](#); [Forget et al., 2015](#)]. The application of the WC01 diffusion (or “smoothing”) operator to a diagonal prior error covariance matrix introduces prior cross-correlations between control variables that represent nearby grid points in the same field. At the same time, WC01 smoothing is a type of regularization, which avoids overfitting to noisy observations. For instance, for all ECCOv4 control variables, the scale of the WC01 spatial correlation model is set to 3 times the grid scale of the underlying GCM, where the nominal horizontal resolution of the latter is 1° [[Forget et al., 2015](#)]. Appendix C provides more details on how the WC01 smoother is implemented and used in practice.

Atmospheric forcing fields are part of the control variables in ocean state estimation [e.g., [Forget et al., 2015](#)]. It shall be noted that WC01 smoothing is not used for the first-guess of the time-varying atmospheric forcing, which is imposed at the horizontal resolution of the underlying GCM (usually, after being interpolated from the grid of an atmospheric model). Rather, ocean state estimation uses WC01 smoothing for the atmospheric *adjustments*. This implies that throughout the optimization process, atmospheric forcing cannot be corrected at spatial scales smaller than the used WC01 diffusion scale. In other words, due to WC01 smoothing, observations can only inform atmospheric adjustments that are of larger scale than the scales that are potentially informable by the observations, leading to an “information loss”. Loss of information, due to WC01 smoothing, is supported by the results in Chapter 4, which explored the information captured by observing systems. It was shown that the information potential is given by prior-weighted linearized sensitivities of the observed quantities to all control variables. Fig. C.4 shows the sensitivities of (or captured information by) an observed OSNAP quantity, with and without WC01 smoothing. A comparison of the two columns in Fig. C.4 highlights the potential “smearing” effect that WC01 smoothing has on information captured by observations.

In particular, it was found in Appendix C that the WC01 smearing has a dissimilar effect on thermal sensitivities vs. wind sensitivities, due to the fact that wind sensitivities are generally characterized by smaller-scale features than thermal sensitivities. That is, in Fig. C.4, WC01 smoothing diffuses (and partially eliminates) wind sensitivities more effectively than thermal sensitivities, even though in both cases the same diffusion scale is used. The dissimilar effect on thermal vs. wind sensitivities was found to have a significant impact on the assessment of observing systems and proxy potential within the framework presented in Chapters 3 and 4 (see Figs. C.1 and C.2 in Section C). The findings in Appendix C pose the following questions:

- (i) On what scales does spatial variability occur in atmospheric forcing fields?
- (ii) Is WC01 smoothing a favorable choice for atmospheric adjustment in ocean state estimation? Should it be applied to all atmospheric forcing fields with the same diffusion scale?

This chapter is an effort to gain first insights into (i) and (ii). The approach is to perform a spectral analysis of the ERA-Interim [Dee et al., 2011] near-surface fields that are adjusted during the ECCOv4 optimization process. The spectral analysis consists of projecting the ERA-Interim near-surface fields onto the spherical harmonics, where the latter form a frequency-space basis on the sphere.

Truncated series of spherical harmonics serve as basis functions in many atmospheric models. The same spectral methods are usually not employed in ocean models since the global ocean surface is not a sphere, as it is interrupted by land. Nevertheless, it is worth investigating whether spherical harmonics are a useful alternative to the currently used bases for atmospheric adjustment in ocean state estimation. For the currently used bases, each basis function corresponds to one atmospheric forcing variable at one surface model grid point. The choice of the current bases entails heavy computer resources and storage requirements, in particular, for high-resolution ocean state estimates. Despite the fact that the discretized model grid has a high dimensionality, the control variables usually have an intrinsically lower-dimensional structure which could be accessed through the spherical harmonic basis functions. Therefore, this chapter will explore a third question:

- (iii) Does a truncated set of spherical harmonics provide a useful lower-dimensional basis for atmospheric adjustment in ocean state estimation?

Section D.2 provides an introduction to spectral methods, spherical harmonics and Gaussian grids. In Section D.3, a spectral analysis of ERA-Interim near-surface fields is performed. In Section D.4, the results and insights into questions (i), (ii), and (iii) are discussed.

D.2 Spectral transformations

Many atmospheric models are based on spectral methods, which represent geophysical fields in horizontal direction in terms of a frequency-space basis over the sphere. This basis is formed by a truncated series of the spherical harmonic functions. The spherical harmonics owe their name to the fact that they are the eigenfunctions of the Laplacian on

the sphere. In spectral models, where the spectral space representation of fields is based on the spherical harmonic basis functions, the nature of the underlying physical latitude-longitude transform grid, also called a *Gaussian grid*, is tightly coupled to the parameters of the spherical harmonic expansion itself. This section provides a short introduction to spherical harmonics and Gaussian grids, closely following [Cushman-Roisin and Beckers \[2011\]](#), Chapter 19.7.

D.2.1 Spherical harmonics

The spherical harmonics are of the form

$$Y_{m,n}(\lambda, \sin \phi) = P_{m,n}(\sin \phi) e^{im\lambda},$$

where λ and ϕ denote longitude and latitude, respectively, and $P_{m,n}$ are Legendre functions¹. The basis functions $Y_{m,n}$ are indexed according to two integers, the order (or wavenumber), n , and the degree, m . These satisfy the constraints that $n \geq 0$ and $|m| \leq n$. Thus, there are $2n + 1$ basis functions of order n . The spherical harmonics $\{Y_{m,n}\}$ form a complete set of orthogonal basis functions on the sphere. Fig. D.1 shows the real part of the spherical harmonics $Y_{m,n}$, up to wavenumber $n = 3$. Spherical harmonics of degree $m = 0$ are zonal, i.e., they do not depend upon longitude. Spherical harmonics with $n = |m|$ are sectoral, i.e., there are no zero crossings in latitudinal direction. For increasing n , the spherical harmonics $Y_{m,n}$ represent smaller and smaller scales (or higher and higher frequencies) on the sphere.

Using the orthogonality of the spherical harmonics $Y_{m,n}$, a (time-dependent) field $u(\lambda, \phi, t)$ on the sphere can be expanded as

$$u(\lambda, \phi, t) = \sum_{m=-\infty}^{\infty} \sum_{n \geq |m|} \alpha_{m,n}(t) Y_{m,n}(\lambda, \sin \phi). \quad (\text{D.1})$$

The coefficient $\alpha_{m,n}$, the “intensity” of the mode $Y_{m,n}$, can be computed by using the orthogonality property of the spherical harmonics:

$$\alpha_{m,n}(t) = \int_{-1}^1 \left(\int_0^{2\pi} u(\lambda, \xi, t) e^{-im\lambda} d\lambda \right) P_{m,n}(\xi) d\xi, \quad (\text{D.2})$$

with $\xi = \sin(\phi)$. In numerical schemes, the sum in (D.1) is truncated, and one obtains an approximation of u :

$$\tilde{u} = \sum_{m=-M}^M \sum_{n=|m|}^{L(m)} \alpha_{m,n} Y_{m,n}. \quad (\text{D.3})$$

¹ The Legendre functions $P_{m,n}$ are defined in terms of Legendre polynomials P_n :

$$P_{m,n}(x) = \begin{cases} \sqrt{\frac{(2n+1)(n-m)!}{2(n+m)!}} (1-x^2)^m \frac{d^m}{dx^m} P_n(x) & m \geq 0 \\ (-1)^m P_{-m,n}(x) & m < 0 \end{cases}$$

with

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n].$$

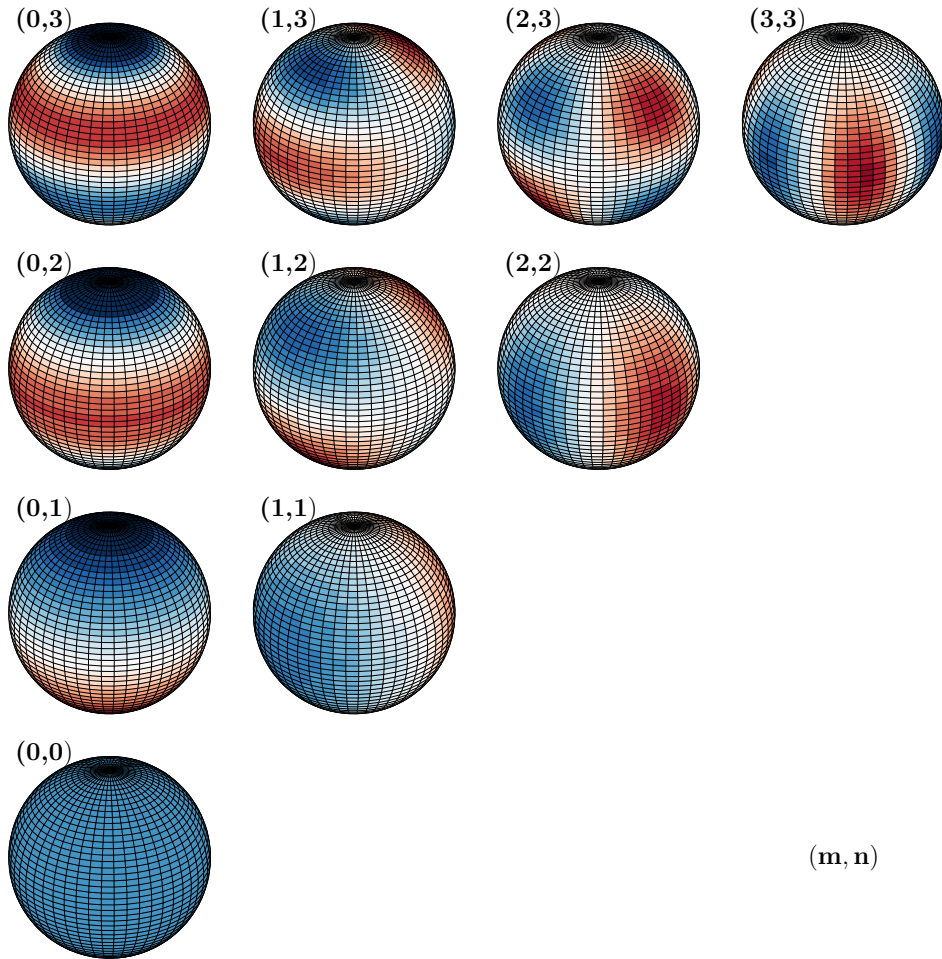


Figure D.1: Real part of the spherical harmonics $Y_{m,n}$, shown for wavenumbers up to $n = 3$. The shown integer pairs are the respective values for (m, n) . The white color marks the separation between positive and negative values. Since the real part of $Y_{m,n}$ and $Y_{-m,n}$ are either equal (for even m) or differ by a factor of -1 (for odd m), only spherical harmonics for $m \geq 0$ are shown. Spherical harmonics of degree $m = 0$ are zonal, i.e., they do not depend upon longitude. Spherical harmonics with $n = |m|$ are sectoral, i.e., there are no zero crossings in latitudinal direction. For increasing n , the spherical harmonics $Y_{m,n}$ represent smaller and smaller scales (or higher and higher frequencies) on the sphere.

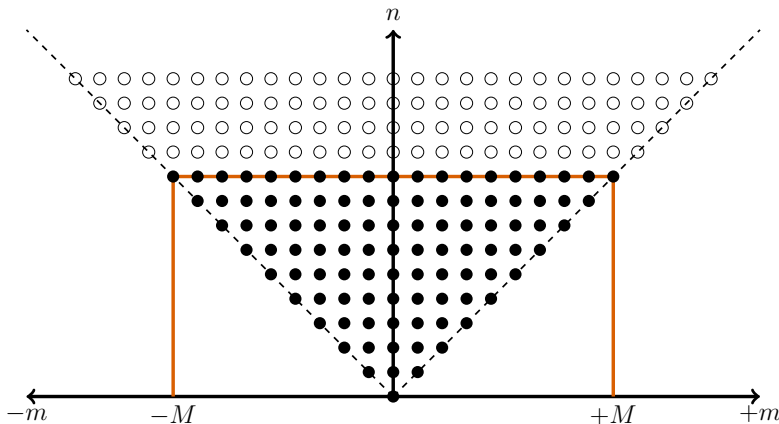


Figure D.2: Triangular truncation of the spherical harmonics, by choosing $L(m) = M$ in (D.3). The single constant spectral truncation parameter M prescribes a triangular region of spherical harmonics in the (m, n) wavenumber space, as indicated by the orange lines and filled circles.

A common choice for the truncation is $L(m) = M$, which corresponds to a triangular truncation (see Fig. D.2), leading to a uniform and isentropic resolution on the sphere. The triangular truncation is for instance employed in the ECMWF (European Centre for Medium-Range Weather Forecasts) model, used for producing the ERA-Interim re-analysis product [Dee et al., 2011].

D.2.2 Gaussian grids

With discrete versions of integrals, orthogonality of the $Y_{m,n}$ basis functions is no longer guaranteed. A forward transformation followed by an inverse transformation (i.e., (D.2) followed by (D.1)) does not ensure a perfect return to the original function. In particular, given a discretized field $\{u(\lambda_i, \phi_j)\}_{i,j}$ on an arbitrary lon-lat grid $\{(\lambda_i, \phi_j)\}_{i,j}$, the computation of the coefficients $\alpha_{m,n}$ via (D.2) results in numerical integration errors. A Gaussian grid $\{(\lambda_i, \phi_j)\}_{i,j}$ is constructed in such a way that the numerical integration errors are minimized. Gaussian grids are therefore tightly coupled to the parameters of the spherical harmonic expansion itself, as described in the following.

Inner integral: FFT The discrete Fourier transform maintains orthogonality. Therefore, the inner integral of (D.2) can be evaluated through a Fast Fourier Transform (FFT) and the corresponding inverse transform by an inverse FFT (see e.g., Appendix C.5 in Cushman-Roisin and Beckers [2011]). That is, FFT is performed at a series of given latitudes ϕ_j , $j = 1, \dots, J$, in zonal direction, which returns the Fourier coefficients

$$\beta_m(\xi_j, t) = \int_0^{2\pi} u(\lambda, \xi_j, t) e^{-im\lambda} d\lambda, \quad (\text{D.4})$$

where $\xi_j = \sin(\phi_j)$. The (exact) discrete Fourier transform operates by sampling at equidistant points $\{\lambda_i\}$. Therefore, the only constraint on the longitudes $\{\lambda_i\}$ in a

Gaussian grid is that they are chosen as equally spaced. In contrast, the Gaussian latitudes $\{\phi_j\}$ will be unequally spaced, since they are defined by Gaussian quadrature used for the outer integral, as described next.

Outer integral: Numerical quadrature There exists no numerical tool similar to the FFT for the outer integral in (D.2), which involves the Legendre functions $P_{m,n}$, and one has to use Gaussian quadrature in meridional direction:

$$\alpha_{m,n}(t) = \sum_{j=1}^J w_j \beta_m(\xi_j, t) P_{m,n}(\xi_j). \quad (\text{D.5})$$

Here, the weights $\{w_j\}$ and locations $\{\xi_j\}$ can be defined so as to reduce integration errors. It is beneficial to choose the J Gaussian latitudes $\{\phi_j\}$ in such a way that the corresponding locations $\{\xi_j = \sin(\phi_j)\}$ are the zeros of the Legendre polynomial of degree J (which are all real), i.e., $P_J(\xi_j) = 0$ for $j = 1, \dots, J$. The corresponding weights are defined as

$$w_j = \frac{2}{(1 - \xi_j^2) \left[\frac{dP_J}{d\xi}(\xi_j) \right]^2}.$$

Regular and reduced Gaussian grids A regular Gaussian grid is defined by the J Gaussian latitudes $\{\phi_j\}$ and $2J$ equally spaced longitudes $\{\lambda_i\}$. In particular, the number of gridpoints λ_i along the longitudes is constant across the chosen latitudes ϕ_j (namely equal to $2J$). In a reduced (or thinned) Gaussian grid, the number of gridpoints per latitude ϕ_j decreases towards the poles, which keeps the gridpoint separation approximately constant across the sphere [Hortal and Simmons, 1990].

D.3 Spectral analysis of ERA-Interim near-surface fields

In the ECCOv4 state estimate [Forget et al., 2015; Fukumori et al., 2017], the first-guess atmospheric forcing is adopted from the ECMWF (European Centre for Medium-Range Weather Forecasts) ERA-Interim reanalysis product [Dee et al., 2011]. ECCOv4 uses 6-hourly ERA-Interim near-surface fields of air temperature at 2 m (atemp), downward longwave radiation (lwdown), downward shortwave radiation (swdown), specific humidity at 2 m (aqh), precipitation (precip), and zonal and meridional wind stress (ustress & vstress). The model heat and freshwater fluxes are determined by bulk formulae of Large and Yeager [2004], computed from atemp, lwdown, swdown, aqh, and precip. The first-guess ERA-Interim forcing fields are subsequently adjusted by the ECCOv4 optimization procedure. Spatially varying adjustments to the ERA-Interim fields are made to both the long-term mean (over the full ECCOv4r3 integration period 1992-2015) and on a bi-weekly basis [Fukumori et al., 2017]. To explore spatial scales of importance for the atmospheric adjustment in ocean state estimation, this section performs a spectral analysis of the ERA-Interim near-surface fields.

D.3.1 Data and grid

The spectral analysis in this chapter uses ERA-Interim reanalysis data from January 1979 until February 2016 for the near-surface fields atemp, lwdown, swdown, aqh, precip,

ustress, and vstress. The ERA-Interim reanalysis product is archived on the reduced Gaussian grid N128 (corresponding to a spectral T255 grid) as well as on lower-resolution Gaussian grids. On the reduced N128 grid, modes up to the truncation parameter $M = 2 \cdot 128 = 256$ are resolved, corresponding to approx. 80 km. The analysis in this chapter is performed for ERA-Interim data retrieved from the coarser reduced Gaussian grid N48, which resolves modes up to $M = 2 \cdot 48 = 96$, corresponding to approx. 210 km. The spectral analysis for the coarse-resolution N48 data will be sufficient to provide first insights to questions (i), (ii), and (iii) and set the stage for future spectral analyses performed for higher-resolution data. (When the study was designed, the choice of analyzing data on the coarse N48 grid was motivated by the main objective to investigate the potential for model reduction in paleoceanographic state estimation.)

D.3.2 Methodology

Long-term means First, the long-term monthly means are computed from the ERA-Interim near-surface fields (using data from January 1979 to February 2016). For each of the fields atemp, lwdown, swdown, aqh, precip, ustress and vstress, this results in 12 fields, i.e., one for each month. Investigating long-term means, rather than averages over short time intervals, is motivated by the fact that ECCOV4r3 makes (spatially varying) adjustments to the long-term means of these atmospheric forcing fields, as described above. Since ECCOV4r3 additionally makes spatio-temporal varying adjustments to the bi-weekly means of the fields, the analysis should be repeated for ERA-Interim fields averaged over shorter time windows, which are expected to be characterized by smaller scales than the long-term means.

Projection A spectral analysis expands a two-dimensional field in terms of the spherical harmonic basis functions $Y_{m,n}$, which represent different spatial scales, as explained in Section D.2. For each of the seven atmospheric forcing fields and each of the 12 months, a spectral analysis is performed for the long-term monthly mean field. The spectral analyses are carried out by projecting the respective long-term monthly mean fields onto the spherical harmonics, via (D.2). The projections return the coefficients $\alpha_{m,n}$, which reflect the intensity of the respective modes $Y_{m,n}$ for the respective atmospheric forcing fields. The projections are performed via an FFT for the inner integral in (D.2), and Gaussian quadrature for the outer integral in (D.2), as described in Section D.2.2. Since the ERA-Interim data was retrieved on a Gaussian grid, following the technique in Section D.2.2 returns accurate coefficients $\alpha_{m,n}$.

D.3.3 Results

Figs. D.3(a)-(g) show the long-term January means of the seven investigated ERA-Interim fields. Accordingly, Figs. D.4(a)-(g) present the result of the spectral analyses of the respective long-term January mean fields, i.e., the projections of Figs. D.3(a)-(g) onto the spherical harmonics. The projections are only shown for the long-term *January* mean fields, since for each of the seven investigated fields, the projections exhibit very similar structure across the 12 months. In Fig. D.4, the result of the spectral analyses is shown in terms of the ratios $|\alpha_{m,n}|/(\max_{m,n} |\alpha_{m,n}|)$, i.e., the relative intensities of the spectral modes $Y_{m,n}$ to the most intense mode. Here, for the fields of atemp, lwdown,

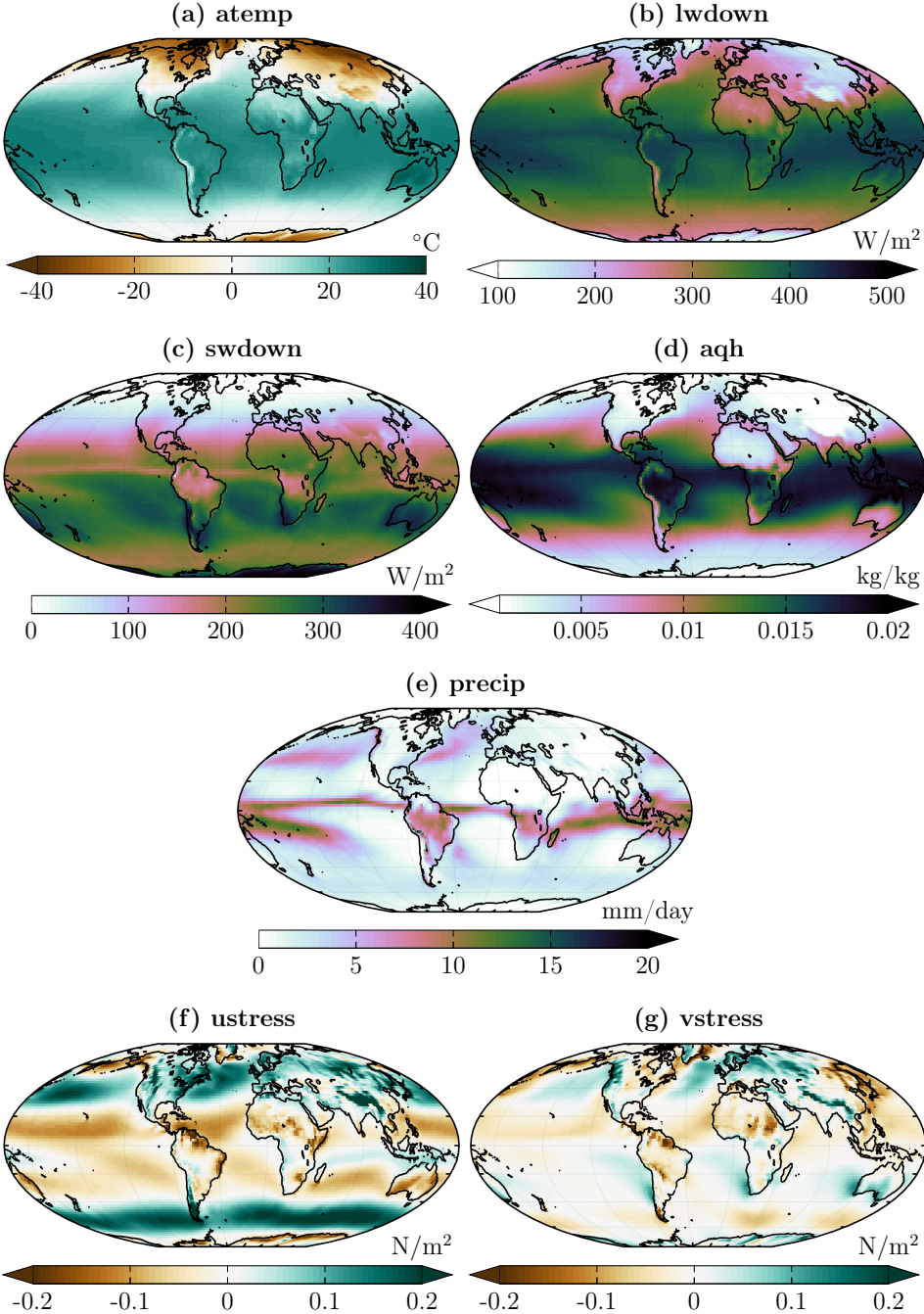


Figure D.3: ERA-Interim long-term January means for the near-surface fields of (a) air temperature at 2 m, (b) downward longwave radiation, (c) downward shortwave radiation, (d) specific humidity at 2 m, (e) precipitation, (f) eastward zonal wind stress, and (g) northward meridional wind stress.

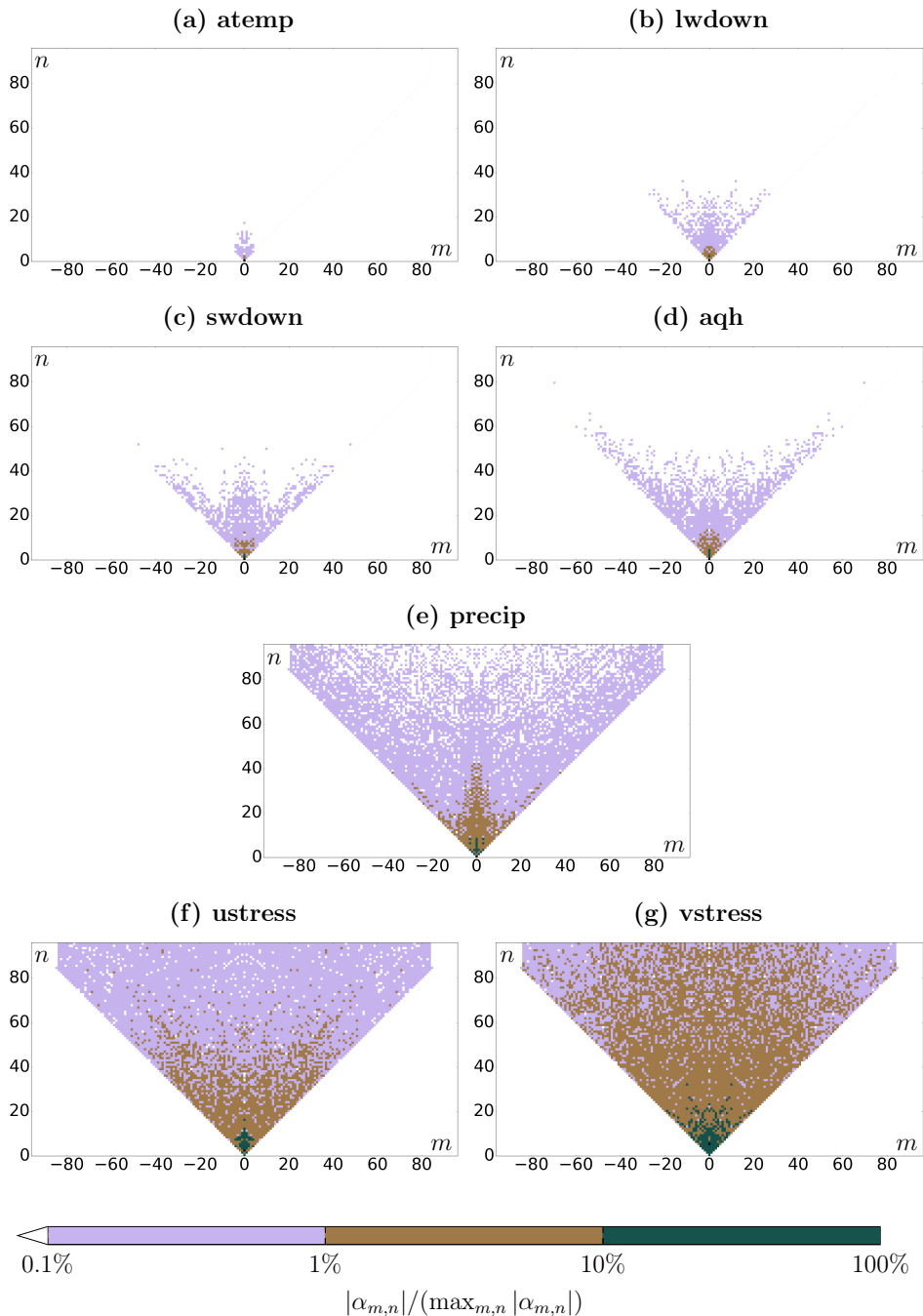


Figure D.4: Relative intensity of the spectral modes $Y_{m,n}$, shown as the ratio $|\alpha_{m,n}| / (\max_{m,n} |\alpha_{m,n}|)$, for the ERA-Interim long-term January means of (a) air temperature at 2 m, (b) downward longwave radiation, (c) downward shortwave radiation, (d) specific humidity at 2 m, (e) precipitation, (f) zonal wind stress, and (g) meridional wind stress.

swwdown, aqh, and precip, $\max_{m,n} |\alpha_{m,n}|$ is equal to $|\alpha_{0,0}|$. That is, for these fields, the constant spherical harmonic $Y_{0,0}$ is found to have highest intensity, which is also true for the remaining 11 months. The coefficient $\alpha_{0,0}$, corresponding to the constant spherical harmonic $Y_{0,0}$, is real and equal to the spatial mean of the respective field. For the fields of ustress and vstress, $\max_{m,n} |\alpha_{m,n}|$ is attained by the absolute value of coefficients corresponding to various modes $Y_{m,n}$ throughout the months, namely by $|\alpha_{0,2}|$, $|\alpha_{0,4}|$, and $|\alpha_{0,5}|$. Note that the coefficients $\alpha_{0,2}$, $\alpha_{0,4}$, $\alpha_{0,5}$ are real and correspond to either of the zonal spherical harmonics $Y_{0,2}$, $Y_{0,4}$, or $Y_{0,5}$ (cf. Fig. D.1).

The relative intensities shown in Figs. D.4(a)-(g) are symmetric about the axis $m = 0$ because for real input data (as the case here), $\alpha_{-m,n}$ is equal to the positive (for even m) or negative (for odd m) complex conjugate of $\alpha_{m,n}$ (cf. equation (D.2)). The color scale in Fig. D.4 is logarithmic, and relative intensities below 0.1% are not shown. Small scales are important for those of the atmospheric fields that show active modes $Y_{m,n}$ up to large wavenumbers n . For each of the seven atmospheric fields, Table D.1 shows the maximal wavenumber $n = n_{\max}$ of the active modes $Y_{m,n}$ that are needed in the spherical harmonic expansion to represent the respective field, where the maximum is taken over all 12 months. Moreover, Table D.1 shows the corresponding smallest scale Δx resolved by the spherical harmonics $Y_{m,n_{\max}}$ of highest wavenumber n_{\max} , computed by $\Delta x = \pi R/n_{\max}$, where R denotes the radius of Earth, and rounded to the nearest ten. The three columns in Table D.1 show three different cases, which, from left to right, take modes $Y_{m,n}$ of increasingly smaller relative intensity $|\alpha_{m,n}|/(\max_{m,n} |\alpha_{m,n}|)$ into account.












	$ \alpha_{m,n} /(\max_{m,n} \alpha_{m,n}) \geq 10\%$ 		$ \alpha_{m,n} /(\max_{m,n} \alpha_{m,n}) \geq 1\%$ 		$ \alpha_{m,n} /(\max_{m,n} \alpha_{m,n}) \geq 0.1\%$   	
	n_{\max}	Δx	n_{\max}	Δx	n_{\max}	Δx
atemp	0	∞	4	5 000 km	18	1 130 km
lwdown	2	10 000 km	8	2 500 km	36	560 km
swwdown	6	3 340 km	18	1 130 km	66	300 km
aqh	4	5 000 km	16	1 250 km	88	230 km
precip	12	1 670 km	51	390 km	96	<210 km
ustress	19	1 050 km	96	<210 km	96	<210 km
vstress	32	630 km	96	<210 km	96	<210 km

Table D.1: Largest wavenumber $n = n_{\max}$ that is required when expanding the ERA-Interim long-term monthly means of air temperature at 2 m (atemp), downward longwave radiation (lwdown), downward shortwave radiation (swwdown), specific humidity at 2 m (aqh), precipitation (precip), zonal wind stress (ustress), and meridional wind stress (vstress) in terms of the spherical harmonics $Y_{m,n}$. Δx indicates the smallest scale (rounded to the nearest ten) resolved by the spherical harmonics $Y_{m,n_{\max}}$ of highest wavenumber n_{\max} . The three columns show three different cases, which, from left to right, take modes $Y_{m,n}$ of increasingly smaller relative intensity $|\alpha_{m,n}|/(\max_{m,n} |\alpha_{m,n}|)$ into account. Including increasingly more, and less intense, modes (from left to right) corresponds to taking into account modes that are shaded by  (first column),   (second column), and    (third column) in Fig. D.4.

Thermal forcing fields

In ECCOV4, the fields of air temperature, downwelling radiation, and specific humidity, shown in subpanels (a)-(d) of Figs. D.3 and D.4, alter the surface heat flux through radiative, sensible, and latent heat fluxes, and are therefore referred to as thermal forcing fields. It is noted that specific humidity does not only affect the net heat flux (via latent heat), but also the net freshwater flux (via evaporation), through the air-sea flux formulation from Large and Yeager [2004].

The most notable feature in the thermal forcing fields shown in Figs. D.3(a)-(d) is a large-scale structure, composed of broad zonal bands. The large-scale structure of the fields is confirmed by Figs. D.4(a)-(d), where modes of relative intensity $\geq 10\%$ (■) are only visible for modes up to wavenumber $n_{\max} = 0$ for atemp (Fig. D.4(a)), $n_{\max} = 2$ for lwdown and sdown (Figs. D.4(b),(c)), and $n_{\max} = 4$ for aqh (Fig. D.4(d)), cf. the first column of Table D.1 for the respective maximal wavenumbers n_{\max} over *all* months. The broad zonal bands visible in Figs. D.3(a)-(d) translate to the fact that all modes of dark green shading (■) in Figs. D.4(a)-(d) are zonal spherical harmonics, i.e., of degree $m = 0$ (cf. D.1).

If one takes into account modes of relative intensity $\geq 1\%$ (■ ■ in Fig. D.4 and second column of Table D.1), the maximal wavenumbers n_{\max} for the thermal forcing fields are still very low. If maximized over all 12 months, the maximal required wavenumber is $n_{\max} = 4$ for atemp, which corresponds to resolved scales $\Delta x > 5000$ km (second column in Table D.1). For lwdown, the maximal required wavenumber is $n_{\max} = 8$ and sdown and aqh show maximal wavenumbers still below $n_{\max} < 20$, corresponding to resolved scales $\Delta x > 1000$ km. For relative intensity $\geq 1\%$, lwdown, sdown, and aqh, show active modes that are not zonal (■ shading for modes of degree $m \neq 0$ in Figs. D.4(b),(c),(d)). This corresponds to a deviation from a zonal structure in the physical long-term mean fields, still plainly visible in Figs. D.3(b)-(d), and is due to the non-zonal structure of cloud coverage.

Modes of high frequency n are absent in the thermal forcing fields, even if accounting for modes of relative intensity $\geq 0.1\%$ (■ ■ ■ in Fig. D.4 and third column of Table D.1). Modes of the highest wavenumbers are mostly zonal for air temperature (spherical harmonics of ■ shading are concentrated around $m = 0$ in Fig. D.4(a)). In contrast, modes of the highest wavenumbers for specific humidity tend to be rather sectoral, reflected by the populated diagonals (of ■ shading) toward n_{\max} in Fig. D.4(d).

Precipitation

In the long-term January mean field of precipitation (Fig. D.3(e)), the narrow zonal band over the ITCZ stands out. The narrow zonal band translates to zonal spherical harmonics ($m = 0$) in Fig. D.4(e) being active at high intensity ($\geq 10\%$, ■ shading) and up to wavenumber $n_{\max} = 12$ (cf. first column of Table D.1). For slightly weaker modes, with relative intensity $\geq 1\%$ (■ ■), spherical harmonics up to wavenumber $n_{\max} = 51$ are active (second column of Table D.1). Here, the emerging modes of the highest wavenumbers ($n \in [35, 50]$) are characterized by a zonal-type structure, and for slightly smaller wavenumbers ($n < 35$) also by a sectoral-type structure, reflected by the

■ shading being mostly concentrated around $m = 0$ and the diagonals in Fig. D.4(e). The maximal wavenumber $n_{\max} = 51$ corresponds to resolved scales $\Delta x > 390$ km. If one accounts for modes of relative intensity $\geq 0.1\%$ (■ ■ ■), spherical harmonics up to the highest wavenumbers resolved in the investigated data are active, i.e., up to $n_{\max} = 96$, corresponding to $\Delta x = 210$ km. To identify the smallest scales of relative intensity $\geq 0.1\%$, data on a finer grid than the considered N48 grid would have to be analyzed.

Wind stress

The long-term January mean wind stress fields in Figs. D.3(f) and (g) show the major well-known large-scale features, such as the subtropical anticyclonic gyres in Pacific and Atlantic Oceans in the Northern and Southern hemispheres, and the strong westerly winds over the Southern Ocean. Inspecting the corresponding spectral decompositions in Figs. D.4(f) and (g) elucidates that besides the visible large-scale structure in Figs. D.3(f) and (g), many small-scale features seem to play a role, too. The most intense modes, of relative intensity $\geq 10\%$ (■ shading), populate wavenumbers up to $n_{\max} = 19$ for u_{stress} and up to $n_{\max} = 32$ for v_{stress} (first column in Table D.1), i.e., much higher wavenumbers than needed for the remaining atmospheric fields. For slightly less intense modes of relative intensity $\geq 1\%$ (■ ■ shading), the active modes already comprise *all* modes up to the highest wavenumbers resolved in the investigated ERA-Interim data, corresponding to scales $\Delta x = 210$ km (second column in Table D.1).

In summary, Fig. D.4 and Table D.1 suggest that wind stress fields are composed of much smaller spatial scales than the thermal forcing fields, and even of considerably smaller spatial scales than precipitation. Reviewing the physical wind stress fields in Figs. D.3(f) and (g) leads to the impression that small-scale features emerge mainly over land, especially over steep orography as the Andes in South America, the Rocky Mountains in North America, or the Himalayas in Asia. In a spectral analysis, there is no natural way to “mask out” spatial variability over land because the spherical harmonics are global basis functions on the sphere. Therefore, the spectral analysis does not distinguish whether the identified small-scale structure in the wind stress fields arises mostly from high-frequency spatial variability over land, e.g., due to orography, or also from high-frequency spatial variability over the ocean. This question needs to be investigated further. Meanwhile, it is noted that the literature reveals many examples for topographically induced features in the wind stress fields along *land-ocean boundaries* that are important for ocean dynamical processes, such as the Greenland tip jet, causing deep convection events in the Irminger Sea [Doyle and Shapiro, 1999; Moore, 2003; Pickart et al., 2003], or other island corner accelerations and gap wind features [Bourassa et al., 1999; Chelton et al., 2000a,b, 2004; Xie et al., 2001, 2005]. Steep gradients in the wind stress fields across several land-ocean boundaries are visible in Figs. D.3(f) and (g), e.g., across the Greenlandic coastline, or the west coasts of South America and South Africa (Fig. D.3(g)). Gradients in the wind stress fields at the boundaries of ocean basins are of particular dynamical importance for ocean models, as will be discussed further in Section D.4.2.

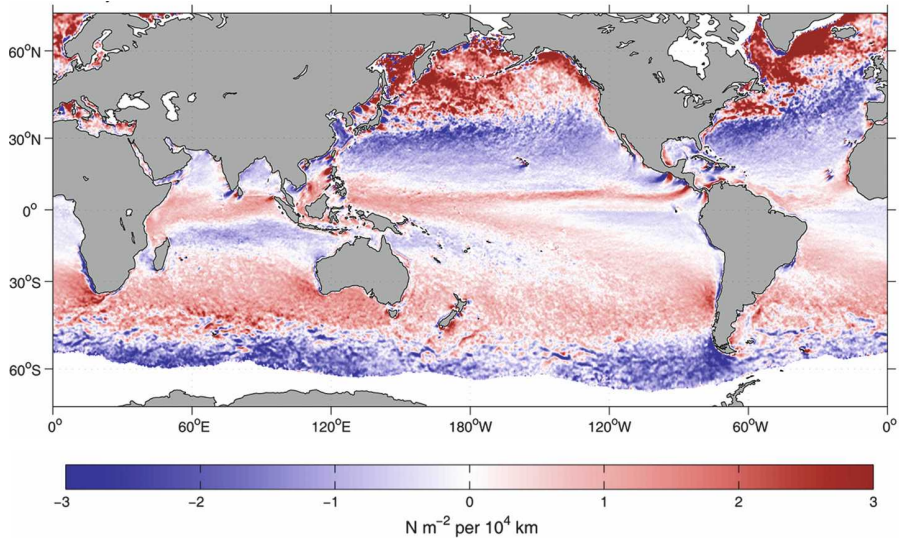


Figure D.5: January mean wind stress curl, as estimated by the Scatterometer Climatology of Ocean Winds (SCOW, [Risien and Chelton, 2008]), plotted on a $0.25^\circ \times 0.25^\circ$ grid. SCOW is estimated from the 8-year record (1999-2007) of wind measurements by the NASA Quick Scatterometer (QuikSCAT). Figure from Risien and Chelton [2008]. ©American Meteorological Society. Used with permission.

D.4 Discussion

This chapter explored spatial scales of atmospheric forcing fields that are adjusted as part of the control vector in ocean state estimation. To this aim, the ERA-Interim long-term monthly mean fields of air temperature, downward longwave radiation, downward shortwave radiation, specific humidity, precipitation, zonal wind stress, and meridional wind stress were expanded into the spherical harmonic basis functions, which form a frequency-space basis on the sphere. The spectral analysis was performed for ERA-Interim data retrieved on the relatively coarse grid N48, which truncates the spherical harmonics at wavenumber $M = 96$ and therefore resolves only length scales larger than approx. 210 km. While a future avenue will be to extend this study to data archived on finer grids (e.g., ERA-Interim data on the N128 grid, corresponding to resolved length scales that are larger than 80 km) and averaged over shorter time intervals, the results in this chapter provided initial insights into the questions (i), (ii), and (iii) that were formulated in Section D.1. These first insights will be discussed in the following.

D.4.1 Spatial scales of ocean surface forcing

Question (i), posed in Section D.1, inquires about the scales of spatial variability in atmospheric forcing fields. Table D.1 provides the answer that was obtained for the long-term monthly ERA-Interim fields investigated in this chapter. To summarize, Table D.1 suggests that the seven investigated atmospheric forcing fields can be divided into three groups. The fields in the first group, comprised of the thermal forcing fields

of air temperature, specific humidity, and downward shortwave and longwave radiation, are characterized by a large-scale spatial structure. Indeed, if one neglects modes of relative intensity $< 1\%$, all spatial variability in the investigated monthly long-term mean thermal forcing fields occurs on scales $\Delta x > 1100$ km, or even $\Delta x > 5000$ km for air temperature. Precipitation, constituting the second group, shows smaller-scale features than the thermal forcing fields. Still, all spatial variability in the investigated monthly long-term mean precipitation fields occurs on scales $\Delta x > 400$ km, if modes of relative intensity $< 1\%$ are neglected. Here, the spectral analysis revealed that the smallest scales $\Delta x \approx 400$ km are induced by narrow zonal bands, linked to the tropical rain belt in the ITCZ. The third group is composed of zonal and meridional wind stress fields. All modes considered in this study, including the highest wavenumbers (or smallest spatial scales), are necessary to expand the wind stress fields in terms of the spherical harmonics. As a result, spatial variability in the investigated wind stress fields occurs on scales $\Delta x < 210$ km. Therefore, the ERA-Interim data on the coarse N48 grid utilized in this study was sufficient to provide an answer to question (i) for the first and second group, but not for the third group.

To obtain further insights into the scales of spatial variability for the third group, comprised of wind stress fields, Fig. D.5 presents a plot from Risien and Chelton [2008]. The plot shows the January mean wind stress curl, as estimated by the Scatterometer Climatology of Ocean Winds (SCOW, [Risien and Chelton, 2008]), which is provided on a $0.25^\circ \times 0.25^\circ$ grid. SCOW is estimated from the 8-year record (1999-2007) of wind measurements by the NASA Quick Scatterometer (QuikSCAT). The wind stress curl field in Fig. D.5 unravels gradients in wind stress more efficiently than the plain wind stress fields, and shows that SCOW wind stress is characterized by features of very small spatial scale. The small-scale structure in Fig. D.5 offers the suggestion that, in order to answer question (i), the spectral analysis performed in this chapter should be repeated for wind stress data on a much finer grid than the utilized N48 ERA-Interim data. Moreover, results from previous studies suggest that spectral analyses performed for (1) reanalysis products such as ERA-Interim or NCEP-NCAR and (2) products directly derived from satellite scatterometer measurements such as SCOW, may give two very different answers. Indeed, Milliff et al. [2004], Chelton [2005]; Chelton et al. [2006, 2004], Maloney and Chelton [2006], and Risien and Chelton [2008] show that reanalysis products from ECMWF (European Centre for Medium-Range Weather Forecasts) and NCEP (National Centers for Environmental Prediction) suffer from an overly smooth structure in the global wind stress fields, when compared to QuikSCAT observations, despite the fact that QuikSCAT observations are assimilated in the reanalysis products. Chelton [2005], Chelton and Wentz [2005], and Maloney and Chelton [2006] argue that the underestimation of the small-scale structure in the reanalysis wind and wind stress fields, even in relatively high-resolution reanalysis products, is to a large degree due to the low-resolution SST boundary condition that is employed in the reanalysis data assimilation frameworks. As a result of poorly resolved SST, the influence of SST gradients on the stability of the marine atmospheric boundary layer and, thus, the surface wind stress fields is not well represented in reanalysis products.

D.4.2 Regularization in ocean state estimation

Question (ii), formulated in Section D.1, concerns WC01 smoothing for atmospheric adjustments in ocean state estimation. WC01 smoothing is a form of regularization and, at the same time, a technique to model prior assumptions on cross-correlations between control variables that represent nearby grid points in the same physical field. In practice, WC01 smoothing implies that throughout the optimization process, first-guess atmospheric forcing cannot be corrected at spatial scales smaller than the chosen WC01 diffusion scale. As an example, ECCOv4 uses ERA-Interim atmospheric forcing fields as a first guess and sets the WC01 diffusion scale to 3 times the underlying LLC90 grid scale, where the latter corresponds to a nominal horizontal resolution of 1° , with refinements at the equator.

The results in this thesis, combined with the insights from previous studies discussed in Section D.4.1, challenge the strategy to apply smoothing to adjustments of wind stress forcing, as performed in global ocean state estimation such as ECCOv4. For instance, Chelton et al. [2006] find that the global operational ECMWF and NCEP models considerably underestimate the spatial variability of wind fields on scales smaller than about 1000 km (see their Fig. 1), even though the horizontal grid resolution of the tested operational models is about 40-50 km. At the same time, small-scale features in the wind stress fields are of dynamical importance for ocean processes, supported by the small-scale structure in the wind sensitivities of various oceanic quantities identified in Chapters 2 and 4. Here, particularly high sensitivity to small-scale features in the wind stress fields was found (a) along the boundaries of the ocean basins, due to wind-driven onshore/offshore Ekman transport and subsequent coastally-trapped waves, and (b) in strong currents such as the Gulf Stream, the NAC, and the various NAC branches, which can be effectively shifted by wind stress curl perturbations. The two facts combined, (1) observed *small-scale spatial variability* in wind stress fields and (2) *high sensitivity* of ocean processes to *small-scale features* in wind stress, suggest that it may be favorable for ECCOv4 to abandon WC01 smoothing for wind stress adjustments and resort to alternative forms of regularization. A promising alternative may be total variation regularization [Rudin et al., 1992], often used for image processing. While WC01 smoothing and Tikhonov regularization are isotropic operators, penalizing sharp features in all directions, total variation regularization is implemented as an anisotropic operator which only penalizes adjustments in directions tangent to the interface. Therefore, total variation regularization would preserve sharp boundaries in the wind stress fields that are informed by the data.

While being a debatable choice for wind stress adjustments, WC01 smoothing may be a suitable strategy for handling adjustments in the thermal forcing fields and precipitation in global ocean state estimation. This is supported by two facts. First, Chapters 2 and 4 showed that, in contrast to wind sensitivities, thermal and precipitation sensitivities (on weekly to decadal timescales) are generally characterized by a smoother structure, due to advective processes. Second, the larger-scale structure is visible not only in the thermal and precipitation *sensitivity fields*, but also in the associated *observed forcing fields*. Indeed, the second column in Table D.1 shows that spatial variability in thermal forcing and precipitation fields occurs on scales $\Delta x > 400$ km, i.e., on scales

larger than the WC01 diffusion scale employed in ECCOv4. Table D.1 also suggests that one may define different WC01 diffusion scales for different atmospheric fields. For instance, the diffusion scale for air temperature can be chosen larger than the diffusion scales for downwelling radiation and specific humidity. Moreover, precipitation requires a smaller diffusion scale than the thermal forcing fields. The spatial scales in Table D.1 were derived from ERA-Interim long-term monthly mean fields and may therefore provide insights for how to choose the diffusion scales for the long-term mean adjustments in ECCOv4. For the ECCOv4 time-variable adjustments, smaller scales than the ones reported in Table D.1 are expected to be important, too. Further investigation is required to determine accurate diffusion scales for time-mean and time-variable atmospheric adjustments. The investigation should involve atmospheric forcing data from various reanalysis products as well as satellite-derived observations, due to different biases in the single products. For instance, reanalysis products typically exhibit biases in the cloud cover [e.g., Free et al., 2016] or in certain regions as the Arctic [Jakobson et al., 2012].

D.4.3 Control space reduction

The spectral analysis performed in this chapter also provides first insights into question (iii), i.e., whether a truncated series of spherical harmonics is a useful alternative basis for atmospheric adjustment in ocean state estimation. This question is motivated by the computational challenges related to the tremendous dimension of the control space, as a result of atmospheric adjustments being made on a grid-point basis. For instance, ECCOv4 operates on the Lat-Lon-Cap (LLC) 90 grid, where the grid spacing varies as a function of latitude (and towards the poles, also as a function of longitude, see Figs. 2 and 3 in Forget et al. [2015]). With its grid spacing ranging between approx. 10 and 100 km, the LLC90 grid has about 60,600 surface grid cells. This results in about 60,600 basis functions per atmospheric forcing field for each long-term mean and bi-weekly adjustment in the ECCOv4 state estimate. Moreover, adjusting atmospheric control variables on a grid-point basis entails a control space dimension that depends quadratically on the chosen horizontal model resolution. Due to limitations in computational and storage resources, the quadratic dependence poses a challenge to high-resolution state estimates, such as ECCO's regional Arctic spin-off ASTE (Arctic and Subpolar gyre sTate Estimate release 1, [Nguyen et al., 2011]) at nominal horizontal resolution $1/3^\circ$.

In contrast to the physical grids that are used by ocean models, many atmospheric models use spectral methods which allocate the grid in the wavenumber space on the sphere, as shown in Fig. D.2. In order to generate fields of real numbers, with imaginary part equal to zero, the coefficients $\alpha_{-m,n}$ have to satisfy the constraint of being equal to the positive (for even m) or negative (for odd m) complex conjugate of $\alpha_{m,n}$. Therefore, if ocean state estimation were to calculate atmospheric adjustments on a grid as in Fig. D.2, i.e., by adjusting coefficients in a spherical harmonic expansion with a triangular truncation, only the coefficients $\alpha_{m,n}$ for $m \geq 0$ would have to be computed. For truncation parameter M (cf. Fig. D.2), this would result in computing a total of $(M + 1)^2/2$ coefficients per atmospheric forcing field and adjustment time interval. For instance, choosing a Gaussian N128 grid, as utilized for the ERA-Interim reanalysis, would result in computing $(2 \cdot 128 + 1)^2/2 \approx 33\,000$ coefficients, about half the num-

ber of coefficients currently computed by ECCOv4. It is noted that the Gaussian N128 grid resolves scales larger than 80 km, which at most locations is coarser than ECCOv4's LLC90 grid.

The spectral analysis in this chapter provides first insights into the question whether for atmospheric adjustment in ocean state estimation it would be sufficient to compute even fewer coefficients than in the spectral N128 grid - corresponding to choosing a smaller truncation parameter M . Based on the results in this chapter, one could truncate at relatively low wavenumbers for the adjustments of the thermal forcing fields. Indeed, if modes of relative intensity $< 1\%$ are neglected (second column in Table D.1), the maximal required wavenumber n_{\max} is smaller than 19 for all thermal forcing fields. Choosing the truncation parameter $M = 19$ would result in having to compute only $(19+1)^2/2 = 200$ coefficients, rather than the 60,600 coefficients that are currently computed by ECCOv4. For precipitation, one could truncate at $M = 51$, which results in the computation of about 1350 coefficients - still less than 3% of the number of coefficients computed by ECCOv4. For the wind stress fields, the investigated N48 data in this chapter is too coarse to determine an appropriate truncation parameter. All that can be inferred from the results in this chapter is that more than $(96+1)^2/2 \approx 4700$ coefficients will be needed for the wind stress fields. It is noted that the truncation parameters as suggested by Table D.1 can be seen as a legitimate first guidance regarding the adjustments of the *long-term mean* of the atmospheric control variables. For the *time-variable* atmospheric adjustments, further insights will be gained by repeating the spectral analysis for atmospheric forcing data averaged over shorter time windows. Moreover, atmospheric forcing data from various reanalysis and satellite-derived products should be considered, due to different biases in the products, as already discussed in Section D.4.2. At the same time, one should keep in mind that the adjustments to the control variables not only represent physics-based corrections, as assumed in this chapter, but also account for model biases. This is due to the fact that model biases are not explicitly taken into account in the cost function but lumped together with the observational errors (see Section B.1).

To conclude, the spectral analysis in this chapter suggests that one might be able to reduce computational and storage requirements for ocean state estimation, by making atmospheric adjustments on coarser grids than the underlying model grid. This could be put into practice by (1) implementing a truncated series of spherical harmonics as new basis functions or (2) using multi-scale adjoints that perform atmospheric adjustment on a lower-resolution grid. A shortcoming of (1) is that ocean models are usually not defined on Gaussian grids, and orthogonality of the spherical harmonic basis functions is no longer ensured. However, option (1) seems an attractive approach, when compared to other techniques that reduce the control space a-priori. The choice of the truncation parameter, corresponding to the choice of the finest resolved scale, is the only assumption that is made a-priori about the unknown adjustments. In contrast, defining the basis functions as a set of EOFs that are derived from a certain data set, e.g., model output, predetermines atmospheric adjustments to mirror characteristic patterns in the analyzed data set. Another method to reduce the control space a-priori is the 'large region approach' [Kaminski and Rayner, 2008; Kaminski et al., 2015, 2018], which does not seem a favorable approach in general, as elucidated by Fig. 6.10 in Section 6.3.2.

D.5 Conclusions

The main conclusions of this chapter are:

- The scales of observed spatial variability differ across various atmospheric forcing types. Thermal forcing fields (air temperature, specific humidity, and downward shortwave and longwave radiation) are characterized by a large-scale spatial structure. Precipitation shows smaller-scale features than thermal forcing fields. Wind stress fields are characterized by spatial variability on the smallest scales.
- Ocean hydrographic and transport quantities show high *sensitivity* to small-scale features in the wind stress fields (i) along the boundaries of the ocean basins, due to wind-driven onshore/offshore Ekman transport and subsequent coastally-trapped waves, and (ii) in strong currents such as the Gulf Stream and the North Atlantic Current, which can be effectively shifted by wind stress curl perturbations. In contrast, sensitivity to thermal forcing fields (on weekly to decadal timescales) is generally characterized by a smoother structure, due to advective processes.
- [Weaver and Courtier \[2001\]](#) (WC01) regularization is a widely employed technique in ocean state estimation, mimicking cross-correlations between nearby model grid points at the chosen diffusion scale. For instance, in ECCOv4, the WC01 diffusion scale is set to 3 times the grid scale of the underlying GCM, for all control variables alike. The fact that thermal forcing and wind stress show contrary behavior in spatial variability in *both* forcing *and* sensitivity fields (see previous two bullet points) suggests that WC01 diffusion scales should be customized.
- Computational and storage requirements for ocean state estimation may be reduced, by making atmospheric adjustments on coarser grids than the underlying model grid, at least for thermal forcing fields. This could be put into practice by (i) implementing a truncated series of spherical harmonics as new basis functions or (ii) using multi-scale adjoints that perform atmospheric adjustment on a lower-resolution grid.

Bibliography

- Alexander, M. A. and Deser, C. (1995). A Mechanism for the Recurrence of Wintertime Midlatitude SST Anomalies. *Journal of Physical Oceanography*, 25(1):122–137.
- Alexanderian, A., Petra, N., Stadler, G., and Ghattas, O. (2016). A Fast and Scalable Method for A-Optimal Design of Experiments for Infinite-dimensional Bayesian Non-linear Inverse Problems. *SIAM Journal on Scientific Computing*, 38(1):A243–A272.
- Amrhein, D. E., Wunsch, C., Marchal, O., and Forget, G. (2018). A Global Glacial Ocean State Estimate Constrained by Upper-Ocean Temperature Proxies. *Journal of Climate*, 31(19):8059–8079.
- Anderson, J. L. and Anderson, S. L. (1999). A Monte Carlo Implementation of the Non-linear Filtering Problem to Produce Ensemble Assimilations and Forecasts. *Monthly Weather Review*, 127(12):2741–2758.
- Antonov, J., Seidov, D., Boyer, T., Locarnini, R., Mishonov, A., Garcia, H., Baranova, O., Zweng, M., and Johnson, D. (2010). World Ocean Atlas 2009, Volume 2: Salinity. *S. Levitus, Ed. NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 pp.*
- Argo (2019). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). *SEANOE*.
- Årthun, M. and Eldevik, T. (2016). On Anomalous Ocean Heat Transport toward the Arctic and Associated Climate Predictability. *Journal of Climate*, 29(2):689–704.
- Årthun, M., Eldevik, T., Viste, E., Drange, H., Furevik, T., Johnson, H. L., and Keenlyside, N. S. (2017). Skillful prediction of northern climate provided by the ocean. *Nature Communications*, 8:15875.
- Baker, N. L. and Daley, R. (2000). Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Quarterly Journal of the Royal Meteorological Society*, 126(565):1431–1454.
- Balmaseda, M. A., Hernandez, F., Storto, A., Palmer, M. D., Alves, O., Shi, L., Smith, G. C., Toyoda, T., Valdivieso, M., Barnier, B., Behringer, D., Boyer, T., Chang, Y.-S., Chepurin, G. A., Ferry, N., Forget, G., Fujii, Y., Good, S., Guinehut, S., Haines, K., Ishikawa, Y., Keeley, S., Köhl, A., Lee, T., Martin, M. J., Masina, S., Masuda, S., Meyssignac, B., Mogensén, K., Parent, L., Peterson, K. A., Tang, Y. M., Yin, Y., Vernieres, G., Wang, X., Waters, J., Wedd, R., Wang, O., Xue, Y., Chevallier, M., Lemieux, J.-F., Dupont, F., Kuragano, T., Kamachi, M., Awaji, T., Caltabiano, A.,

- Wilmer-Becker, K., and Gaillard, F. (2015). The Ocean Reanalyses Intercomparison Project (ORA-IP). *Journal of Operational Oceanography*, 8(sup1):s80–s97.
- Bashir, O., Willcox, K., Ghattas, O., Waanders, B. v. B., and Hill, J. (2008). Hessian-based model reduction for large-scale systems with initial-condition inputs. *International Journal for Numerical Methods in Engineering*, 73(6):844–868.
- Battisti, D. S., Bhatt, U. S., and Alexander, M. A. (1995). A Modeling Study of the Interannual Variability in the Wintertime North Atlantic Ocean. *Journal of Climate*, 8(12):3067–3083.
- Bocquet, M., Pires, C. A., and Wu, L. (2010). Beyond Gaussian Statistical Modeling in Geophysical Data Assimilation. *Monthly Weather Review*, 138(8):2997–3023.
- Böning, C. W., Behrens, E., Biastoch, A., Getzlaff, K., and Bamber, J. L. (2016). Emerging impact of Greenland meltwater on deepwater formation in the North Atlantic Ocean. *Nature Geoscience*, 9(7):523–527.
- Bourassa, M. A., Zamudio, L., and O’Brien, J. J. (1999). Noninertial flow in NSCAT observations of Tehuantepec winds. *Journal of Geophysical Research: Oceans*, 104(C5):11311–11319.
- Brambilla, E. and Talley, L. D. (2008). Subpolar Mode Water in the northeastern Atlantic: 1. Averaged properties and mean circulation. *Journal of Geophysical Research: Oceans*, 113(C4).
- Branstator, G. and Teng, H. (2014). Is AMOC More Predictable than North Atlantic Heat Content? *Journal of Climate*, 27(10):3537–3550.
- Broecker, W. S., Peteet, D. M., and Rind, D. (1985). Does the ocean-atmosphere system have more than one stable mode of operation? *Nature*, 315(6014):21–26.
- Buckley, M. W. and Marshall, J. (2016). Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review. *Reviews of Geophysics*, 54(1):2015RG000493.
- Buckley, M. W., Ponte, R. M., Forget, G., and Heimbach, P. (2014). Low-Frequency SST and Upper-Ocean Heat Content Variability in the North Atlantic. *Journal of Climate*, 27(13):4996–5018.
- Buckley, M. W., Ponte, R. M., Forget, G., and Heimbach, P. (2015). Determining the Origins of Advective Heat Transport Convergence Variability in the North Atlantic. *Journal of Climate*, 28(10):3943–3956.
- Bugnion, V., Hill, C., and Stone, P. H. (2006). An Adjoint Analysis of the Meridional Overturning Circulation in an Ocean Model. *Journal of Climate*, 19(15):3732–3750.
- Bui-Thanh, T., Burstedde, C., Ghattas, O., Martin, J., Stadler, G., and Wilcox, L. C. (2012). Extreme-scale UQ for Bayesian Inverse Problems Governed by PDEs. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC ’12*, pages 1–11.

- Carmack, E., Polyakov, I., Padman, L., Fer, I., Hunke, E., Hutchings, J., Jackson, J., Kelley, D., Kwok, R., Layton, C., Melling, H., Perovich, D., Persson, O., Ruddick, B., Timmermans, M.-L., Toole, J., Ross, T., Vavrus, S., and Winsor, P. (2015). Toward Quantifying the Increasing Role of Oceanic Heat in Sea Ice Loss in the New Arctic. *Bulletin of the American Meteorological Society*, 96(12):2079–2105.
- Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G. (2018). Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5):e535.
- Carstens, J., Hebbeln, D., and Wefer, G. (1997). Distribution of planktic foraminifera at the ice margin in the Arctic (Fram Strait). *Marine Micropaleontology*, 29(3):257–269.
- Carton, J. A., Chepurin, G. A., Reagan, J., and Häkkinen, S. (2011). Interannual to decadal variability of Atlantic Water in the Nordic and adjacent seas. *Journal of Geophysical Research: Oceans*, 116(C11):C11035.
- Cayan, D. R. (1992). Latent and Sensible Heat Flux Anomalies over the Northern Oceans: Driving the Sea Surface Temperature. *Journal of Physical Oceanography*, 22(8):859–881.
- Chaloner, K. and Verdinelli, I. (1995). Bayesian Experimental Design: A Review. *Statistical Science*, 10(3):273–304.
- Chaudhuri, A. H., Ponte, R. M., Forget, G., and Heimbach, P. (2012). A Comparison of Atmospheric Reanalysis Surface Products over the Ocean and Implications for Uncertainties in Air–Sea Boundary Forcing. *Journal of Climate*, 26(1):153–170.
- Chelton, D. B. (2005). The Impact of SST Specification on ECMWF Surface Wind Stress Fields in the Eastern Tropical Pacific. *Journal of Climate*, 18(4):530–550.
- Chelton, D. B., Freilich, M. H., and Esbensen, S. K. (2000a). Satellite Observations of the Wind Jets off the Pacific Coast of Central America. Part I: Case Studies and Statistical Characteristics. *Monthly Weather Review*, 128(7):1993–2018.
- Chelton, D. B., Freilich, M. H., and Esbensen, S. K. (2000b). Satellite Observations of the Wind Jets off the Pacific Coast of Central America. Part II: Regional Relationships and Dynamical Considerations. *Monthly Weather Review*, 128(7):2019–2043.
- Chelton, D. B., Freilich, M. H., Sienkiewicz, J. M., and Von Ahn, J. M. (2006). On the Use of QuikSCAT Scatterometer Measurements of Surface Winds for Marine Weather Prediction. *Monthly Weather Review*, 134(8):2055–2071.
- Chelton, D. B., Schlax, M. G., Freilich, M. H., and Milliff, R. F. (2004). Satellite Measurements Reveal Persistent Small-Scale Features in Ocean Winds. *Science*, 303(5660):978–983.
- Chelton, D. B. and Wentz, F. J. (2005). Global Microwave Satellite Observations of Sea Surface Temperature for Numerical Weather Prediction and Climate Research. *Bulletin of the American Meteorological Society*, 86(8):1097–1116.

- Chen, P. (2011). Hessian Matrix vs. Gauss–Newton Hessian Matrix. *SIAM Journal on Numerical Analysis*, 49(4):1417–1435.
- Cioaca, A., Sandu, A., and de Sturler, E. (2013). Efficient methods for computing observation impact in 4d-Var data assimilation. *Computational Geosciences*, 17(6):975–990.
- CLIMAP Project Members (1976). The Surface of the Ice-Age Earth. *Science*, 191(4232):1131–1137.
- Constantine, P. (2015). *Active Subspaces*. SIAM Spotlights. Society for Industrial and Applied Mathematics.
- Courtier, P., Thépaut, J.-N., and Hollingsworth, A. (1994). A strategy for operational implementation of 4d-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519):1367–1387.
- Cuny, J., Rhines, P. B., Niiler, P. P., and Bacon, S. (2002). Labrador Sea Boundary Currents and the Fate of the Irminger Sea Water. *Journal of Physical Oceanography*, 32(2):627–647.
- Curry, B., Lee, C. M., and Petrie, B. (2010). Volume, Freshwater, and Heat Fluxes through Davis Strait, 2004–05. *Journal of Physical Oceanography*, 41(3):429–436.
- Cushman-Roisin, B. and Beckers, J.-M. (2011). *Introduction to Geophysical Fluid Dynamics: Physical and Numerical Aspects*. Academic Press.
- Czeschel, L., Eden, C., and Greatbatch, R. J. (2012). On the Driving Mechanism of the Annual Cycle of the Florida Current Transport. *Journal of Physical Oceanography*, 42(5):824–839.
- Czeschel, L., Marshall, D. P., and Johnson, H. L. (2010). Oscillatory sensitivity of Atlantic overturning to high-latitude forcing. *Geophysical Research Letters*, 37(10):L10601.
- Daescu, D. N. (2008). On the Sensitivity Equations of Four-Dimensional Variational (4d-Var) Data Assimilation. *Monthly Weather Review*, 136(8):3050–3065.
- Dail, H. and Wunsch, C. (2013). Dynamical Reconstruction of Upper-Ocean Conditions in the Last Glacial Maximum Atlantic. *Journal of Climate*, 27(2):807–823.
- Danabasoglu, G., Yeager, S. G., Bailey, D., Behrens, E., Bentsen, M., Bi, D., Biastoch, A., Böning, C., Bozec, A., Canuto, V. M., Cassou, C., Chassignet, E., Coward, A. C., Danilov, S., Diansky, N., Drange, H., Farneti, R., Fernandez, E., Fogli, P. G., Forget, G., Fujii, Y., Griffies, S. M., Gusev, A., Heimbach, P., Howard, A., Jung, T., Kelley, M., Large, W. G., Leboissetier, A., Lu, J., Madec, G., Marsland, S. J., Masina, S., Navarra, A., George Nurser, A. J., Pirani, A., y Méliá, D. S., Samuels, B. L., Scheinert, M., Sidorenko, D., Treguier, A.-M., Tsujino, H., Uotila, P., Valcke, S., Voldoire, A., and Wang, Q. (2014). North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part I: Mean states. *Ocean Modelling*, 73:76–107.

- Danabasoglu, G., Yeager, S. G., Kim, W. M., Behrens, E., Bentsen, M., Bi, D., Biastoch, A., Bleck, R., Böning, C., Bozec, A., Canuto, V. M., Cassou, C., Chassignet, E., Coward, A. C., Danilov, S., Diansky, N., Drange, H., Farneti, R., Fernandez, E., Fogli, P. G., Forget, G., Fujii, Y., Griffies, S. M., Gusev, A., Heimbach, P., Howard, A., Ilicak, M., Jung, T., Karspeck, A. R., Kelley, M., Large, W. G., Leboissetier, A., Lu, J., Madec, G., Marsland, S. J., Masina, S., Navarra, A., Nurser, A. J. G., Pirani, A., Romanou, A., Salas y Mélia, D., Samuels, B. L., Scheinert, M., Sidorenko, D., Sun, S., Treguier, A.-M., Tsujino, H., Uotila, P., Valcke, S., Voldoire, A., Wang, Q., and Yashayaev, I. (2016). North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part II: Inter-annual to decadal variability. *Ocean Modelling*, 97:65–90.
- Dansgaard, W., Johnsen, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., Hvidberg, C. S., Steffensen, J. P., Sveinbjörnsdóttir, A. E., Jouzel, J., and Bond, G. (1993). Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, 364(6434):218.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597.
- Deshayes, J., Curry, R., and Msadek, R. (2014). CMIP5 Model Intercomparison of Freshwater Budget and Circulation in the North Atlantic. *Journal of Climate*, 27(9):3298–3317.
- Dokken, T. M., Nisancioglu, K. H., Li, C., Battisti, D. S., and Kissel, C. (2013). Dansgaard-Oeschger cycles: Interactions between ocean and sea ice intrinsic to the Nordic seas. *Paleoceanography*, 28(3):491–502.
- Dong, B.-W. and Sutton, R. T. (2001). The dominant mechanisms of variability in Atlantic Ocean Heat Transport in a Coupled Ocean-Atmosphere GCM. *Geophysical Research Letters*, 28(12):2445–2448.
- Dong, S., Hautala, S. L., and Kelly, K. A. (2007). Interannual Variations in Upper-Ocean Heat Content and Heat Transport Convergence in the Western North Atlantic. *Journal of Physical Oceanography*, 37(11):2682–2697.
- Doyle, J. D. and Shapiro, M. A. (1999). Flow response to large-scale topography: the Greenland tip jet. *Tellus A*, 51(5):728–748.
- Fenty, I. and Heimbach, P. (2012). Coupled Sea Ice–Ocean-State Estimation in the Labrador Sea and Baffin Bay. *Journal of Physical Oceanography*, 43(5):884–904.

- Flath, H., Wilcox, L., Akçelik, V., Hill, J., van Bloemen Waanders, B., and Ghattas, O. (2011). Fast Algorithms for Bayesian Uncertainty Quantification in Large-Scale Linear Inverse Problems Based on Low-Rank Partial Hessian Approximations. *SIAM Journal on Scientific Computing*, 33(1):407–432.
- Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., and Wunsch, C. (2015). ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geosci. Model Dev.*, 8(10):3071–3104.
- Frankignoul, C. (1985). Sea surface temperature anomalies, planetary waves, and air-sea feedback in the middle latitudes. *Reviews of Geophysics*, 23(4):357–390.
- Frankignoul, C. and Hasselmann, K. (1977). Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability. *Tellus*, 29(4):289–305.
- Free, M., Sun, B., and Yoo, H. L. (2016). Comparison between Total Cloud Cover in Four Reanalysis Products and Cloud Measured by Visual Observations at U.S. Weather Stations. *Journal of Climate*, 29(6):2015–2021.
- Fujii, Y., Rémy, E., Zuo, H., Oke, P., Halliwell, G., Gasparin, F., Benkiran, M., Loose, N., Cummings, J., Xie, J., Xue, Y., Masuda, S., Smith, G. C., Balmaseda, M., Germaineaud, C., J., L. D., Larnicol, G., Bertino, Laurent Bonaduce, A., Pierre, B., Donlon, C., Heimbach, P., Kim, Y., Kourafalou, V., Yves Le Traon, P., Martin, M., Paturi, S., Tranchant, B., and Usui, N. (2019). Observing system evaluation based on ocean data assimilation and prediction systems: On-going challenges and future vision for designing/supporting ocean observational networks. Submitted to *Frontiers in Marine Science*.
- Fukumori, I., Wang, O., Fenty, I., Forget, G., Heimbach, P., and Ponte, R. M. (2017). ECCO Version 4 Release 3. Retrieved June 28, 2018, from <http://hdl.handle.net/1721.1/110380>.
- Fukumori, I., Wang, O., Llovel, W., Fenty, I., and Forget, G. (2015). A near-uniform fluctuation of ocean bottom pressure and sea level across the deep ocean basins of the Arctic Ocean and the Nordic Seas. *Progress in Oceanography*, 134:152–172.
- Furevik, T. (2000). On Anomalous Sea Surface Temperatures in the Nordic Seas. *Journal of Climate*, 13(5):1044–1053.
- Furevik, T. (2001). Annual and interannual variability of Atlantic Water temperatures in the Norwegian and Barents Seas: 1980–1996. *Deep Sea Research Part I: Oceanographic Research Papers*, 48(2):383–404.
- Furevik, T. and Nilsen, J. E. (2005). Large-Scale Atmospheric Circulation Variability and its Impacts on the Nordic Seas Ocean Climate—A Review. In *The Nordic Seas: An Integrated Perspective*, pages 105–136. American Geophysical Union.
- Gaspar, P., Grégoris, Y., and Lefevre, J.-M. (1990). A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: Tests at station Papa and long-term

- upper ocean study site. *Journal of Geophysical Research: Oceans*, 95(C9):16179–16193.
- Gebbie, G. and Huybers, P. (2006). Meridional circulation during the Last Glacial Maximum explored through a combination of South Atlantic $\delta^{18}\text{O}$ observations and a geostrophic inverse model. *Geochemistry, Geophysics, Geosystems*, 7(11).
- Gebbie, G. and Huybers, P. (2011a). How is the ocean filled? *Geophysical Research Letters*, 38(6):L06604.
- Gebbie, G. and Huybers, P. (2011b). The Mean Age of Ocean Waters Inferred from Radiocarbon Observations: Sensitivity to Surface Sources and Accounting for Mixing Histories. *Journal of Physical Oceanography*, 42(2):291–305.
- Gelaro, R. and Zhu, Y. (2009). Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus A*, 61(2):179–193.
- Gent, P. R. and McWilliams, J. C. (1990). Isopycnal Mixing in Ocean Circulation Models. *Journal of Physical Oceanography*, 20(1):150–155.
- Giering, R. and Kaminski, T. (1998). Recipes for Adjoint Code Construction. *ACM Trans. Math. Softw.*, 24(4):437–474.
- Gildor, H. and Tziperman, E. (2003). Sea-ice switches and abrupt climate change. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 361(1810):1935–1944.
- Greatbatch, R. J. (1994). A note on the representation of steric sea level in models that conserve volume rather than mass. *Journal of Geophysical Research: Oceans*, 99(C6):12767–12771.
- Grötzner, A., Latif, M., and Barnett, T. P. (1998). A Decadal Climate Cycle in the North Atlantic Ocean as Simulated by the ECHO Coupled GCM. *Journal of Climate*, 11(5):831–847.
- Haber, E., Magnant, Z., Lucero, C., and Tenorio, L. (2011). Numerical methods for A-optimal designs with a sparsity constraint for ill-posed inverse problems. *Computational Optimization and Applications*, 52(1):293–314.
- Häkkinen, S. and Rhines, P. B. (2009). Shifting surface currents in the northern North Atlantic Ocean. *Journal of Geophysical Research: Oceans*, 114(C4):C04005.
- Häkkinen, S., Rhines, P. B., and Worthen, D. L. (2011). Warm and saline events embedded in the meridional circulation of the northern North Atlantic. *Journal of Geophysical Research: Oceans*, 116(C3):C03006.
- Hansen, B. and Østerhus, S. (2000). North Atlantic–Nordic Seas exchanges. *Progress in Oceanography*, 45(2):109–208.
- Hardt, M. and Scherbaum, F. (1994). The design of optimum networks for aftershock recordings. *Geophysical Journal International*, 117(3):716–726.

- Hátún, H., Sandø, A. B., Drange, H., Hansen, B., and Valdimarsson, H. i. (2005). Influence of the Atlantic Subpolar Gyre on the Thermohaline Circulation. *Science*, 309(5742):1841–1844.
- Heimbach, P., Menemenlis, D., Losch, M., Campin, J.-M., and Hill, C. (2010). On the formulation of sea-ice models. Part 2: Lessons from multi-year adjoint sea-ice export sensitivities through the Canadian Arctic Archipelago. *Ocean Modelling*, 33(1–2):145–158.
- Heimbach, P., Wunsch, C., Ponte, R. M., Forget, G., Hill, C., and Utke, J. (2011). 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.
- Heuzé, C. (2017). North Atlantic deep water formation and AMOC in CMIP5 models. *Ocean Science*, 13(4):609–622.
- Hoffman, R. N. and Atlas, R. (2015). Future Observing System Simulation Experiments. *Bulletin of the American Meteorological Society*, 97(9):1601–1616.
- Holland, D. M., Thomas, R. H., Young, B. d., Ribergaard, M. H., and Lyberth, B. (2008). Acceleration of Jakobshavn Isbræ triggered by warm subsurface ocean waters. *Nature Geoscience*, 1(10):ngeo316.
- Holliday, N. P., Bacon, S., Allen, J., and McDonagh, E. L. (2009). Circulation and Transport in the Western Boundary Currents at Cape Farewell, Greenland. *Journal of Physical Oceanography*, 39(8):1854–1870.
- Holliday, N. P., Bacon, S., Cunningham, S. A., Gary, S. F., Karstensen, J., King, B. A., Li, F., and McDonagh, E. L. (2018). Subpolar North Atlantic Overturning and Gyre-Scale Circulation in the Summers of 2014 and 2016. *Journal of Geophysical Research: Oceans*, 123(7):4538–4559.
- Holliday, N. P., Hughes, S. L., Bacon, S., Beszczynska-Möller, A., Hansen, B., Lavín, A., Loeng, H., Mork, K. A., Østerhus, S., Sherwin, T., and Walczowski, W. (2008). Reversal of the 1960s to 1990s freshening trend in the northeast North Atlantic and Nordic Seas. *Geophysical Research Letters*, 35(3):L03614.
- Hortal, M. and Simmons, A. J. (1990). Use of Reduced Gaussian Grids in Spectral Models. *Monthly Weather Review*, 119(4):1057–1074.
- Howat, I. M., Joughin, I., and Scambos, T. A. (2007). Rapid Changes in Ice Discharge from Greenland Outlet Glaciers. *Science*, 315(5818):1559–1561.
- Huybers, P., Gebbie, G., and Marchal, O. (2007). Can Paleooceanographic Tracers Constrain Meridional Circulation Rates? *Journal of Physical Oceanography*, 37(2):394–407.
- Huybers, P. and Wunsch, C. (2010). Paleophysical oceanography with an emphasis on transport rates. *Annual Review of Marine Science*, 2:1–34.

- Isaac, T., Petra, N., Stadler, G., and Ghattas, O. (2015). Scalable and efficient algorithms for the propagation of uncertainty from data through inference to prediction for large-scale problems, with application to flow of the Antarctic ice sheet. *Journal of Computational Physics*, 296:348–368.
- Jackson, C. (2000). Sensitivity of stationary wave amplitude to regional changes in Laurentide ice sheet topography in single-layer models of the atmosphere. *Journal of Geophysical Research: Atmospheres*, 105(D19):24443–24454.
- Jakobson, E., Vihma, T., Palo, T., Jakobson, L., Keernik, H., and Jaagus, J. (2012). Validation of atmospheric reanalyses over the central Arctic Ocean. *Geophysical Research Letters*, 39(10).
- Johnson, H. L. and Marshall, D. P. (2002). A Theory for the Surface Atlantic Response to Thermohaline Variability. *Journal of Physical Oceanography*, 32(4):1121–1132.
- Jones, D., Forget, G., Sinha, B., Josey, S., Boland, E., Meijers, A., and Shuckburgh, E. (2018). Local and Remote Influences on the Heat Content of the Labrador Sea: An Adjoint Sensitivity Study. *Journal of Geophysical Research: Oceans*, 0(0).
- Joughin, I., Abdalati, W., and Fahnestock, M. (2004). Large fluctuations in speed on Greenland’s Jakobshavn Isbræ glacier. *Nature*, 432(7017):608.
- Joughin, I., Alley, R. B., and Holland, D. M. (2012). Ice-Sheet Response to Oceanic Forcing. *Science*, 338(6111):1172–1176.
- Junge, M. M. and Haine, T. W. N. (2001). Mechanisms of North Atlantic Wintertime Sea Surface Temperature Anomalies. *Journal of Climate*, 14(24):4560–4572.
- Kalmikov, A. and Heimbach, P. (2014). A Hessian-Based Method for Uncertainty Quantification in Global Ocean State Estimation. *SIAM Journal on Scientific Computing*, 36(5):S267–S295.
- Kalmikov, A. G. and Heimbach, P. (2018). On Barotropic Mechanisms of Uncertainty Propagation in Estimation of Drake Passage Transport. arXiv: 1804.06033v2.
- Kaminski, T., Kauker, F., Eicken, H., and Karcher, M. (2015). Exploring the utility of quantitative network design in evaluating Arctic sea ice thickness sampling strategies. *The Cryosphere*, 9(4):1721–1733.
- Kaminski, T., Kauker, F., Toudal Pedersen, L., Voßbeck, M., Haak, H., Niederdrenk, L., Hendricks, S., Ricker, R., Karcher, M., Eicken, H., and Gråbak, O. (2018). Arctic Mission Benefit Analysis: impact of sea ice thickness, freeboard, and snow depth products on sea ice forecast performance. *The Cryosphere*, 12(8):2569–2594.
- Kaminski, T. and Rayner, P. J. (2008). Assimilation and Network Design. In *The Continental-Scale Greenhouse Gas Balance of Europe*, Ecological Studies, pages 33–52. Springer, New York, NY.
- Kaminski, T. and Rayner, P. J. (2017). Reviews and syntheses: guiding the evolution of the observing system for the carbon cycle through quantitative network design. *Biogeosciences*, 14(20):4755–4766.

- Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G. (2001). On aggregation errors in atmospheric transport inversions. *Journal of Geophysical Research: Atmospheres*, 106(D5):4703–4715.
- Karcher, M., Gerdes, R., Kauker, F., Köberle, C., and Yashayaev, I. (2005). Arctic Ocean change heralds North Atlantic freshening. *Geophysical Research Letters*, 32(21):L21606.
- Khatiwala, S., Tanhua, T., Mikaloff Fletcher, S., Gerber, M., Doney, S. C., Graven, H. D., Gruber, N., McKinley, G. A., Murata, A., Ríos, A. F., and Sabine, C. L. (2013). Global ocean storage of anthropogenic carbon. *Biogeosciences*, 10(4):2169–2191.
- Kindler, P., Guillevic, M., Baumgartner, M., Schwander, J., Landais, A., and Leuenberger, M. (2014). Temperature reconstruction from 10 to 120 kyr b2k from the NGRIP ice core. *Climate of the Past*, 10(2):887–902.
- Klinker, E., Rabier, F., Kelly, G., and Mahfouf, J.-F. (2000). The ECMWF operational implementation of four-dimensional variational assimilation. III: Experimental results and diagnostics with operational configuration. *Quarterly Journal of the Royal Meteorological Society*, 126(564):1191–1215.
- Köhl, A. and Stammer, D. (2004). Optimal Observations for Variational Data Assimilation. *Journal of Physical Oceanography*, 34(3):529–542.
- Kostov, Y., Johnson, H. L., and Marshall, D. P. (2019). AMOC sensitivity to surface buoyancy fluxes: the role of air-sea feedback mechanisms. *Climate Dynamics*.
- Kurahashi-Nakamura, T., Losch, M., and Paul, A. (2014). Can sparse proxy data constrain the strength of the Atlantic meridional overturning circulation? *Geosci. Model Dev.*, 7(1):419–432.
- Kurahashi-Nakamura, T., Paul, A., and Losch, M. (2017). Dynamical reconstruction of the global ocean state during the Last Glacial Maximum. *Paleoceanography*, 32(4):2016PA003001.
- Kwon, Y.-O. and Frankignoul, C. (2014). Mechanisms of Multidecadal Atlantic Meridional Overturning Circulation Variability Diagnosed in Depth versus Density Space. *Journal of Climate*, 27(24):9359–9376.
- Langland, R. H. and Baker, N. L. (2004). Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus A*, 56(3):189–201.
- Large, G. and Yeager, S. (2004). Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies. *Technical Report TN-460+STR*, NCAR.
- Latif, M., Collins, M., Pohlmann, H., and Keenlyside, N. (2006). A Review of Predictability Studies of Atlantic Sector Climate on Decadal Time Scales. *Journal of Climate*, 19(23):5971–5987.
- Law, K., Stuart, A., and Zygalakis, K. (2015). Data Assimilation - A Mathematical Introduction. *Springer*.

- LeGrand, P. and Wunsch, C. (1995). Constraints from paleotracer data on the North Atlantic circulation during the Last Glacial Maximum. *Paleoceanography*, 10(6):1011–1045.
- Li, C., Battisti, D. S., Schrag, D. P., and Tziperman, E. (2005). Abrupt climate shifts in Greenland due to displacements of the sea ice edge. *Geophysical Research Letters*, 32(19):L19702.
- Li, F., Lozier, M. S., and Johns, W. E. (2017). Calculating the Meridional Volume, Heat, and Freshwater Transports from an Observing System in the Subpolar North Atlantic: Observing System Simulation Experiment. *Journal of Atmospheric and Oceanic Technology*, 34(7):1483–1500.
- Lieberman, C., Fidkowski, K., Willcox, K., and Waanders, B. v. B. (2013). Hessian-based model reduction: large-scale inversion and prediction. *International Journal for Numerical Methods in Fluids*, 71(2):135–150.
- Locarnini, R., Mishonov, A., Antonov, J., Boyer, T., Garcia, H., Baranova, O., Zweng, M., and Johnson, D. (2010). World Ocean Atlas 2009, Volume 1: Temperature. *S. Levitus, Ed. NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 pp.*
- Losch, M. and Wunsch, C. (2003). Bottom Topography as a Control Variable in an Ocean Model. *Journal of Atmospheric and Oceanic Technology*, 20(11):1685–1696.
- Lozier, M. S., Bacon, S., Bower, A. S., Cunningham, S. A., Femke de Jong, M., de Steur, L., deYoung, B., Fischer, J., Gary, S. F., Greenan, B. J. W., Heimbach, P., Holliday, N. P., Houpert, L., Inall, M. E., Johns, W. E., Johnson, H. L., Karstensen, J., Li, F., Lin, X., Mackay, N., Marshall, D. P., Mercier, H., Myers, P. G., Pickart, R. S., Pillar, H. R., Straneo, F., Thierry, V., Weller, R. A., Williams, R. G., Wilson, C., Yang, J., Zhao, J., and Zika, J. D. (2017). Overturning in the Subpolar North Atlantic Program: A New International Ocean Observing System. *Bulletin of the American Meteorological Society*, 98(4):737–752.
- Lozier, M. S., Li, F., Bacon, S., Bahr, F., Bower, A. S., Cunningham, S. A., Jong, M. F. d., Steur, L. d., deYoung, B., Fischer, J., Gary, S. F., Greenan, B. J. W., Holliday, N. P., Houk, A., Houpert, L., Inall, M. E., Johns, W. E., Johnson, H. L., Johnson, C., Karstensen, J., Koman, G., Bras, I. A. L., Lin, X., Mackay, N., Marshall, D. P., Mercier, H., Olmanns, M., Pickart, R. S., Ramsey, A. L., Rayner, D., Straneo, F., Thierry, V., Torres, D. J., Williams, R. G., Wilson, C., Yang, J., Yashayaev, I., and Zhao, J. (2019). A sea change in our view of overturning in the subpolar North Atlantic. *Science*, 363(6426):516–521.
- Luo, H., Castelao, R. M., Rennermalm, A. K., Tedesco, M., Bracco, A., Yager, P. L., and Mote, T. L. (2016). Oceanic transport of surface meltwater from the southern Greenland ice sheet. *Nature Geoscience*, 9(7):528–532.
- Mahfouf, J.-F. and Rabier, F. (2000). The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics. *Quarterly Journal of the Royal Meteorological Society*, 126(564):1171–1190.

- Maloney, E. D. and Chelton, D. B. (2006). An Assessment of the Sea Surface Temperature Influence on Surface Wind Stress in Numerical Weather Prediction and Climate Models. *Journal of Climate*, 19(12):2743–2762.
- Marchal, O. and Curry, W. B. (2008). On the Abyssal Circulation in the Glacial Atlantic. *Journal of Physical Oceanography*, 38(9):2014–2037.
- Marotzke, J., Giering, R., Zhang, K. Q., Stammer, D., Hill, C., and Lee, T. (1999). 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.
- Marshall, J., Adcroft, A., Hill, C., Perelman, L., and Heisey, C. (1997a). A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research: Oceans*, 102(C3):5753–5766.
- Marshall, J., Hill, C., Perelman, L., and Adcroft, A. (1997b). Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *Journal of Geophysical Research: Oceans*, 102(C3):5733–5752.
- Marshall, J., Johnson, H., and Goodman, J. (2001). A Study of the Interaction of the North Atlantic Oscillation with Ocean Circulation. *Journal of Climate*, 14(7):1399–1421.
- Marshall, J. and Schott, F. (1999). Open-ocean convection: Observations, theory, and models. *Reviews of Geophysics*, 37(1):1–64.
- Marzocchi, A., Hirschi, J. J. M., Holliday, N. P., Cunningham, S. A., Blaker, A. T., and Coward, A. C. (2015). The North Atlantic subpolar circulation in an eddy-resolving global ocean model. *Journal of Marine Systems*, 142:126–143.
- Masutani, M., Woollen, J. S., Lord, S. J., Emmitt, G. D., Kleespies, T. J., Wood, S. A., Greco, S., Sun, H., Terry, J., Kapoor, V., Treadon, R., and Campana, K. A. (2010). Observing system simulation experiments at the National Centers for Environmental Prediction. *Journal of Geophysical Research: Atmospheres*, 115(D7).
- McCartney, M. S. and Talley, L. D. (1982). The Subpolar Mode Water of the North Atlantic Ocean. *Journal of Physical Oceanography*, 12(11):1169–1188.
- Menary, M. B., Hodson, D. L. R., Robson, J. I., Sutton, R. T., Wood, R. A., and Hunt, J. A. (2015). Exploring the impact of CMIP5 model biases on the simulation of North Atlantic decadal variability. *Geophysical Research Letters*, 42(14):2015GL064360.
- Milliff, R. F., Morzel, J., Chelton, D. B., and Freilich, M. H. (2004). Wind Stress Curl and Wind Stress Divergence Biases from Rain Effects on QSCAT Surface Wind Retrievals. *Journal of Atmospheric and Oceanic Technology*, 21(8):1216–1231.
- Moore, A. M., Arango, H. G., Broquet, G., Powell, B. S., Weaver, A. T., and Zavala-Garay, J. (2011a). The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part I - System overview and formulation. *Progress in Oceanography*, 91(1):34–49.

- Moore, A. M., Arango, H. G., and Broquet, G. (2012). Estimates of Analysis and Forecast Error Variances Derived from the Adjoint of 4d-Var. *Monthly Weather Review*, 140(10):3183–3203.
- Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B., Foley, D., Doyle, J. D., Costa, D., and Robinson, P. (2011b). The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part II - Performance and application to the California Current System. *Progress in Oceanography*, 91(1):50–73.
- Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B., Foley, D., Doyle, J. D., Costa, D., and Robinson, P. (2011c). The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part III - Observation impact and observation sensitivity in the California Current System. *Progress in Oceanography*, 91(1):74–94.
- Moore, A. M., Arango, H. G., and Edwards, C. A. (2017a). Reduced-Rank Array Modes of the California Current Observing System. *Journal of Geophysical Research: Oceans*, 123(1):452–465.
- Moore, A. M., Jacox, M. G., Crawford, W. J., Laughlin, B., Edwards, C. A., and Fiechter, J. (2017b). The impact of the ocean observing system on estimates of the California current circulation spanning three decades. *Progress in Oceanography*, 156:41–60.
- Moore, G. W. K. (2003). Gale force winds over the Irminger Sea to the east of Cape Farewell, Greenland. *Geophysical Research Letters*, 30(17).
- Mork, K. A., Skagseth, Ø., Ivshin, V., Ozhigin, V., Hughes, S. L., and Valdimarsson, H. (2014). Advective and atmospheric forced changes in heat and fresh water content in the Norwegian Sea, 1951–2010. *Geophysical Research Letters*, 41(17):2014GL061038.
- National Academies of Sciences, Engineering, and Medicine (2017). *Sustaining Ocean Observations to Understand Future Changes in Earth's Climate*. Washington, DC: The National Academies Press.
- Nguyen, A. T., Menemenlis, D., and Kwok, R. (2011). Arctic ice-ocean simulation with optimized model parameters: Approach and assessment. *Journal of Geophysical Research: Oceans*, 116(C4).
- North Greenland Ice Core Project members (2004). High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature*, 431(7005):147.
- Oke, P. R., Balmaseda, M. A., Benkiran, M., Cummings, J., Dombrowsky, E., Fujii, Y., Guinehut, S., Larnicol, G., Traon, P.-Y. L., and Martin, M. J. (2010). Observational Requirements of GODAE Systems. In *Proceedings of OceanObs'19: Sustained Ocean Observations and Information for Society (Vol. 2)*.
- Orvik, K. A. and Niiler, P. (2002). Major pathways of Atlantic water in the northern North Atlantic and Nordic Seas toward Arctic. *Geophysical Research Letters*, 29(19):2–1–2–4.

- Orvik, K. A. and Skagseth, O. (2003). The impact of the wind stress curl in the North Atlantic on the Atlantic inflow to the Norwegian Sea toward the Arctic. *Geophysical Research Letters*, 30(17):1884.
- Petersen, S. V., Schrag, D. P., and Clark, P. U. (2013). A new mechanism for Dansgaard-Oeschger cycles. *Paleoceanography*, 28(1):24–30.
- Pickart, R. S., Spall, M. A., Ribergaard, M. H., Moore, G. W. K., and Milliff, R. F. (2003). Deep convection in the Irminger Sea forced by the Greenland tip jet. *Nature*, 424(6945):152–156.
- Pillar, H. R., Heimbach, P., Johnson, H. L., and Marshall, D. P. (2016). Dynamical Attribution of Recent Variability in Atlantic Overturning. *Journal of Climate*, 29(9):3339–3352.
- Pilo, G. S., Oke, P. R., Coleman, R., Rykova, T., and Ridgway, K. (2018). Impact of data assimilation on vertical velocities in an eddy resolving ocean model. *Ocean Modelling*, 131:71–85.
- Pires, C., Vautard, R., and Talagrand, O. (1996). On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus A*, 48(1):96–121.
- Pohlmann, H., Jungclaus, J. H., Köhl, A., Stammer, D., and Marotzke, J. (2009). Initializing Decadal Climate Predictions with the GECCO Oceanic Synthesis: Effects on the North Atlantic. *Journal of Climate*, 22(14):3926–3938.
- Polyakov, I. V., Pnyushkov, A. V., Alkire, M. B., Ashik, I. M., Baumann, T. M., Carmack, E. C., Goszczko, I., Guthrie, J., Ivanov, V. V., Kanzow, T., Krishfield, R., Kwok, R., Sundfjord, A., Morison, J., Rember, R., and Yulin, A. (2017). Greater role for Atlantic inflows on sea-ice loss in the Eurasian Basin of the Arctic Ocean. *Science*, 356(6335):285–291.
- Privé, N. C. and Errico, R. M. (2013). The role of model and initial condition error in numerical weather forecasting investigated with an observing system simulation experiment. *Tellus A: Dynamic Meteorology and Oceanography*, 65(1):21740.
- Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F., and Simmons, A. (2000). The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quarterly Journal of the Royal Meteorological Society*, 126(564):1143–1170.
- Rahmstorf, S. (2002). Ocean circulation and climate during the past 120,000 years. *Nature*, 419(6903):207–214.
- Rainsley, E., Menviel, L., Fogwill, C. J., Turney, C. S. M., Hughes, A. L. C., and Rood, D. H. (2018). Greenland ice mass loss during the Younger Dryas driven by Atlantic Meridional Overturning Circulation feedbacks. *Scientific Reports*, 8(1):11307.
- Redi, M. H. (1982). Oceanic Isopycnal Mixing by Coordinate Rotation. *Journal of Physical Oceanography*, 12(10):1154–1158.

- Risien, C. M. and Chelton, D. B. (2008). A Global Climatology of Surface Wind and Wind Stress Fields from Eight Years of QuikSCAT Scatterometer Data. *Journal of Physical Oceanography*, 38(11):2379–2413.
- Roe, G. H. and Lindzen, R. S. (2001). The Mutual Interaction between Continental-Scale Ice Sheets and Atmospheric Stationary Waves. *Journal of Climate*, 14(7):1450–1465.
- Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268.
- Sadatzi, H., Dokken, T. M., Berben, S. M. P., Muschitiello, F., Stein, R., Fahl, K., Menviel, L., Timmermann, A., and Jansen, E. (2019). Sea ice variability in the southern Norwegian Sea during glacial Dansgaard-Oeschger climate cycles. *Science Advances*, 5(3):eaau6174.
- Sasaki, Y. (1970). Some basic formalisms in numerical variational analysis. *Monthly Weather Review*, 98(12):875–883.
- Schjøth, F., Andresen, C. S., Straneo, F., Murray, T., Scharrer, K., and Korabely, A. (2012). Campaign to map the bathymetry of a major Greenland fjord. *Eos, Transactions American Geophysical Union*, 93(14):141–142.
- Schumann, K., Völker, D., and Weinrebe, W. R. (2012). Acoustic mapping of the Ilulissat Ice Fjord mouth, West Greenland. *Quaternary Science Reviews*, 40:78–88.
- Sgubin, G., Swingedouw, D., Drijfhout, S., Mary, Y., and Bennabi, A. (2017). Abrupt cooling over the North Atlantic in modern climate models. *Nature Communications*, 8:14375.
- Smith, T. and Heimbach, P. (2019). Atmospheric origins of variability in the South Atlantic meridional overturning circulation. *Journal of Climate*.
- Staines-Urías, F., Kuijpers, A., and Korte, C. (2013). Evolution of subpolar North Atlantic surface circulation since the early Holocene inferred from planktic foraminifera faunal and stable isotope records. *Quaternary Science Reviews*, 76:66–81.
- Stammer, D. (2008). Response of the global ocean to Greenland and Antarctic ice melting. *Journal of Geophysical Research: Oceans*, 113(C6).
- Stammer, D., Balmaseda, M., Heimbach, P., Köhl, A., and Weaver, A. (2016). Ocean Data Assimilation in Support of Climate Applications: Status and Perspectives. *Annual Review of Marine Science*, 8:491–518.
- Stearns, L. A. and Hamilton, G. S. (2007). Rapid volume loss from two East Greenland outlet glaciers quantified using repeat stereo satellite imagery. *Geophysical Research Letters*, 34(5).
- Straneo, F., Curry, R. G., Sutherland, D. A., Hamilton, G. S., Cenedese, C., Våge, K., and Stearns, L. A. (2011). Impact of fjord dynamics and glacial runoff on the circulation near Helheim Glacier. *Nature Geoscience*, 4(5):322–327.

- Straneo, F., Hamilton, G. S., Sutherland, D. A., Stearns, L. A., Davidson, F., Hammill, M. O., Stenson, G. B., and Rosing-Asvid, A. (2010). Rapid circulation of warm subtropical waters in a major glacial fjord in East Greenland. *Nature Geoscience*, 3(3):182–186.
- Straneo, F. and Heimbach, P. (2013). North Atlantic warming and the retreat of Greenland’s outlet glaciers. *Nature*, 504(7478):36–43.
- Straneo, F., Sutherland, D. A., Holland, D., Gladish, C., Hamilton, G. S., Johnson, H. L., Rignot, E., Xu, Y., and Koppes, M. (2012). Characteristics of ocean waters reaching Greenland’s glaciers. *Annals of Glaciology*, 53(60):202–210.
- Straneo, F., Sutherland, D. A., Stearns, L., Catania, G., Heimbach, P., Moon, T., Cape, M. R., Laidre, K. L., Barber, D., Rysgaard, S., Mottram, R., Olsen, S., Hopwood, M. J., and Meire, L. (2019). The Case for a Sustained Greenland Ice Sheet-Ocean Observing System (GrIOOS). *Frontiers in Marine Science*, 6.
- Stuart, A. M. (2010). Inverse problems: A Bayesian perspective. *Acta Numerica*, 19:451–559.
- Sutton, R. T. and Allen, M. R. (1997). Decadal predictability of North Atlantic sea surface temperature and climate. *Nature*, 388(6642):563.
- Sutton, R. T. and Dong, B. (2012). Atlantic Ocean influence on a shift in European climate in the 1990s. *Nature Geoscience*, 5(11):788–792.
- Sutton, R. T. and Hodson, D. L. R. (2005). Atlantic Ocean Forcing of North American and European Summer Climate. *Science*, 309(5731):115–118.
- Sévellec, F., Ben Jelloul, M., and Huck, T. (2007). Optimal Surface Salinity Perturbations Influencing the Thermohaline Circulation. *Journal of Physical Oceanography*, 37(12):2789–2808.
- Sévellec, F., Fedorov, A. V., and Liu, W. (2017). Arctic sea-ice decline weakens the Atlantic Meridional Overturning Circulation. *Nature Climate Change*, 7(8):604–610.
- Sévellec, F., Huck, T., Ben Jelloul, M., Grima, N., Vialard, J., and Weaver, A. (2008). Optimal Surface Salinity Perturbations of the Meridional Overturning and Heat Transport in a Global Ocean General Circulation Model. *Journal of Physical Oceanography*, 38(12):2739–2754.
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics.
- Thacker, W. C. (1989). The role of the Hessian matrix in fitting models to measurements. *Journal of Geophysical Research: Oceans*, 94(C5):6177–6196.
- Trémolet, Y. (2007). Incremental 4d-Var convergence study. *Tellus A*, 59(5):706–718.

- Verdy, A., Mazloff, M. R., Cornuelle, B. D., and Kim, S. Y. (2013). Wind-Driven Sea Level Variability on the California Coast: An Adjoint Sensitivity Analysis. *Journal of Physical Oceanography*, 44(1):297–318.
- Vieli, A. and Nick, F. M. (2011). Understanding and Modelling Rapid Dynamic Changes of Tidewater Outlet Glaciers: Issues and Implications. *Surveys in Geophysics*, 32(4):437–458.
- Vigen, T. (2015 (retrieved January 18, 2019)). Spurious correlations. <http://tylervigen.com>.
- Voelker, A. H. L. (2002). Global distribution of centennial-scale records for Marine Isotope Stage (MIS) 3: a database. *Quaternary Science Reviews*, 21(10):1185–1212.
- Waelbroeck, C., Labeyrie, L., Michel, E., Duplessy, J., McManus, J., Lambeck, K., Balbon, E., and Labracherie, M. (2002). Sea-level and deep water temperature changes derived from benthic foraminifera isotopic records. *Quaternary Science Reviews*, 21(1):295–305.
- Waelbroeck, C., Paul, A., Kucera, M., Rosell-Melé, A., Weinelt, M., Schneider, R., Mix, A. C., Abelmann, A., Armand, L., Bard, E., Barker, S., Barrows, T. T., Benway, H., Cacho, I., Chen, M.-T., Cortijo, E., Crosta, X., Vernal, A. d., Dokken, T., Duprat, J., Elderfield, H., Eynaud, F., Gersonde, R., Hayes, A., Henry, M., Hillaire-Marcel, C., Huang, C.-C., Jansen, E., Juggins, S., Kallel, N., Kiefer, T., Kienast, M., Labeyrie, L., Leclaire, H., Londeix, L., Mangin, S., Matthiessen, J., Marret, F., Meland, M., Morey, A. E., Mulitza, S., Pflaumann, U., Pisias, N. G., Radi, T., Rochon, A., Rohling, E. J., Scaffi, L., Schäfer-Neth, C., Solignac, S., Spero, H., Tachikawa, K., and Turon, J.-L. (2009). Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum. *Nature Geoscience*, 2(2):127–132.
- Wang, J., Fu, L.-L., Qiu, B., Menemenlis, D., Farrar, J. T., Chao, Y., Thompson, A. F., and Flexas, M. M. (2017). An Observing System Simulation Experiment for the Calibration and Validation of the Surface Water Ocean Topography Sea Surface Height Measurement Using In Situ Platforms. *Journal of Atmospheric and Oceanic Technology*, 35(2):281–297.
- Weaver, A. and Courtier, P. (2001). Correlation modelling on the sphere using a generalized diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, 127(575):1815–1846.
- Whitaker, J. S. and Hamill, T. M. (2002). Ensemble Data Assimilation without Perturbed Observations. *Monthly Weather Review*, 130(7):1913–1924.
- Winguth, A. M. E., Archer, D., Duplessy, J.-C., Maier-Reimer, E., and Mikolajewicz, U. (1999). Sensitivity of paleonutrient tracer distributions and deep-sea circulation to glacial boundary conditions. *Paleoceanography*, 14(3):304–323.
- Wunsch, C. (1996). *The Ocean Circulation Inverse Problem*. Cambridge University Press, Cambridge.

- Wunsch, C. and Heimbach, P. (2007). Practical global oceanic state estimation. *Physica D: Nonlinear Phenomena*, 230(1–2):197–208.
- Xie, S.-P., Liu, W. T., Liu, Q., and Nonaka, M. (2001). Far-Reaching Effects of the Hawaiian Islands on the Pacific Ocean-Atmosphere System. *Science*, 292(5524):2057–2060.
- Xie, S.-P., Xu, H., Kessler, W. S., and Nonaka, M. (2005). Air-Sea Interaction over the Eastern Pacific Warm Pool: Gap Winds, Thermocline Dome, and Atmospheric Convection. *Journal of Climate*, 18(1):5–20.
- Xu, X., Chassignet, E. P., and Wang, F. (2018). On the variability of the Atlantic meridional overturning circulation transports in coupled CMIP5 simulations. *Climate Dynamics*.
- Yang, Q., Dixon, T. H., Myers, P. G., Bonin, J., Chambers, D., van den Broeke, M. R., Ribergaard, M. H., and Mortensen, J. (2016). Recent increases in Arctic freshwater flux affects Labrador Sea convection and Atlantic overturning circulation. *Nature Communications*, 7:10525.
- Yeager, S. G. and Robson, J. I. (2017). Recent Progress in Understanding and Predicting Atlantic Decadal Climate Variability. *Current Climate Change Reports*, 3(2):112–127.
- Zanna, L., Heimbach, P., Moore, A. M., and Tziperman, E. (2010). Optimal Excitation of Interannual Atlantic Meridional Overturning Circulation Variability. *Journal of Climate*, 24(2):413–427.
- Zanna, L., Heimbach, P., Moore, A. M., and Tziperman, E. (2012). 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.
- Zhang, R. (2010). Latitudinal dependence of Atlantic meridional overturning circulation (AMOC) variations. *Geophysical Research Letters*, 37(16):L16703.
- Zhang, R. (2015). Mechanisms for low-frequency variability of summer Arctic sea ice extent. *Proceedings of the National Academy of Sciences*, 112(15):4570–4575.
- Zhu, Y. and Gelaro, R. (2008). Observation Sensitivity Calculations Using the Adjoint of the Gridpoint Statistical Interpolation (GSI) Analysis System. *Monthly Weather Review*, 136(1):335–351.



Graphic design: Communication Division, UIB / Print: Skjipes Kommunikasjon AS



uib.no

ISBN: 9788230841334 (print)
9788230864913 (PDF)