Quantifying Dynamical Proxy Potential through Oceanic Teleconnections in the North Atlantic

Nora Loose[1,1,1], Patrick Heimbach[1,1,1], Helen Pillar[1,1,1], and Kerim Nisancioglu[2,2,2]

[1]The University of Texas at Austin
[2]University of Bergen

November 30, 2022

Abstract

Oceanic quantities of interest (QoIs), e.g., ocean heat content or transports, are often inaccessible to direct observation, due to the high cost of instrument deployment and logistical challenges. Therefore, oceanographers seek proxies for undersampled or unobserved QoIs. Conventionally, proxy potential is assessed via statistical correlations, which measure covariability without establishing causality. This paper introduces an alternative method: quantifying dynamical proxy potential. Using an adjoint model, this method unambiguously identifies the physical origins of covariability. A North Atlantic case study illustrates our method within the ECCO (Estimating the Circulation and Climate of the Ocean) state estimation framework. We find that wind forcing along the eastern and northern boundaries of the Atlantic drives a basin-wide response in North Atlantic circulation and temperature. Due to these large-scale teleconnections, a single subsurface temperature observation in the Irminger Sea informs heat transport across the remote Iceland-Scotland ridge (ISR), with a dynamical proxy potential of 19%. Dynamical proxy potential allows two equivalent interpretations: Irminger Sea subsurface temperature (i) shares 19% of its adjustment physics with ISR heat transport; (ii) reduces the uncertainty in ISR heat transport by 19% (independent of the measured temperature value), if the Irminger Sea observation is added without noise to the ECCO state estimate. With its two interpretations, dynamical proxy potential is simultaneously rooted in (i) ocean dynamics and (ii) uncertainty quantification and optimal observing system design, the latter being an emerging branch in computational science. The new method may therefore foster dynamics-based, quantitative ocean observing system design in the coming years.

Posted on 30 Nov 2022 — CC-BY 4.0 — https://doi.org/10.1002/essoar.10502065/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Quantifying Dynamical Proxy Potential through Shared Adjustment Physics in the North Atlantic

N. Loose[1,2], P. Heimbach[1,3,4], H. R. Pillar[1], K. H. Nisancioglu[2,5]

[1]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX, USA
[2]Department of Earth Science, University of Bergen, and Bjerknes Centre for Climate Research, Bergen, Norway
[3]Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA
[4]Institute for Geophysics, The University of Texas at Austin, Austin, TX, USA
[5]Department of Geosciences, University of Oslo, and Centre for Earth Evolution and Dynamics, Oslo, Norway

Key Points:

- Sensitivities of key metrics, representing observed and desirable quantities in the ocean, are computed using an adjoint model
- Similar sensitivity distributions indicate proxy potential through shared ocean adjustment physics
- In contrast to conventional statistical methods, our quantification of proxy potential is rigorously dynamics-based

Corresponding author: Nora Loose, nora.loose@utexas.edu

Abstract

Oceanic quantities of interest (QoIs), e.g., ocean heat content or transports, are often inaccessible to direct observation, due to the high cost of instrument deployment and logistical challenges. Therefore, oceanographers seek proxies for undersampled or unobserved QoIs. Conventionally, proxy potential is assessed via statistical correlations, which measure covariability without establishing causality. This paper introduces an alternative method: quantifying dynamical proxy potential. Using an adjoint model, this method unambiguously identifies the physical origins of covariability. A North Atlantic case study illustrates our method within the ECCO (Estimating the Circulation and Climate of the Ocean) state estimation framework. We find that wind forcing along the eastern and northern boundaries of the Atlantic drives a basin-wide response in North Atlantic circulation and temperature. Due to these large-scale teleconnections, a single subsurface temperature observation in the Irminger Sea informs heat transport across the remote Iceland-Scotland ridge (ISR), with a dynamical proxy potential of 19%. Dynamical proxy potential allows two equivalent interpretations: Irminger Sea subsurface temperature (i) shares 19% of its adjustment physics with ISR heat transport; (ii) reduces the uncertainty in ISR heat transport by 19% (independent of the measured temperature value), if the Irminger Sea observation is added without noise to the ECCO state estimate. With its two interpretations, dynamical proxy potential is simultaneously rooted in (i) ocean dynamics and (ii) uncertainty quantification and optimal observing system design, the latter being an emerging branch in computational science. The new method may therefore foster dynamics-based, quantitative ocean observing system design in the coming years.

Plain Language Summary

To understand the Earth's changing climate, it is important to estimate how much heat the ocean takes up from the atmosphere and how the ocean recirculates the heat around the globe. Directly obtaining these estimates from measurements is complicated because oceanographers cannot measure the ocean everywhere. Ocean measurements taken from ships or freely drifting instruments are expensive and difficult to obtain, especially in regions with ice cover or rough weather conditions. To analyze how existing measurements can be used to estimate unmeasured aspects of the ocean, past studies have used statistical correlations, although it is usually unclear whether correlations have a real, physical origin. This paper introduces a new method: we replace statistical correlations by correlations that have an underlying physical mechanism. As an example, the paper reveals that (A) a subsurface ocean temperature measurement in the Irminger Sea helps to better estimate (B) poleward ocean heat transport across the Iceland-Scotland ridge, hundreds of kilometers away. (A) and (B) are related by physics-based correlation, which is created by a similar dynamical response of (A) and (B) to changes in the near- and far-field wind. The new method can be used to plan effective instrument placements in the future.

1 Introduction

Satellite altimetry and the global array of Argo floats have vastly increased the observational coverage of the world's oceans since the early 1990s (Fu et al., 2018; Riser et al., 2016). Nevertheless, large parts of the ocean remain undersampled in space and time, due to the high cost of instrument deployment, ongoing technical and logistical challenges, and the fact that critically relevant processes occur on a wide range of spatial and temporal scales (e.g., Weller et al., 2019). Therefore, many oceanographic quantities of interest (QoIs) are not directly or continuously measured. Examples are volume, heat, and freshwater transports across many oceanographic passages, straits and latitude bands, in particular, the Atlantic meridional overturning circulation (AMOC). Additional QoIs may be ocean heat and freshwater content in regions not well sampled by Argo floats,

e.g., near the margins or the polar ice sheets. Other examples are strongly related to future societal concerns and key targets for climate predictions. These include future Arctic sea ice cover or regional sea level anomalies. In oceanography, we are therefore on the quest for proxies. That is, it is desirable - and an active part of climate research - to employ observed quantities as proxies for QoIs that are undersampled or unobserved.

Examples of past efforts in this direction include studies which found that anomalies in sea level (available from altimetry) or ocean bottom pressure (available from gravimetry) can serve as a skillful proxy for AMOC variability on interannual time scales (e.g., Bingham & Hughes, 2009; Ezer, 2015; Frajka-Williams, 2015; Landerer et al., 2015; McCarthy et al., 2015). Other studies suggested that, on decadal and longer time scales, North Atlantic surface or subsurface temperature have a characteristic 'fingerprint' associated with changes in AMOC, and that the (better observed) temperature fingerprint can be used as a proxy for (unobserved) AMOC (e.g., Baehr et al., 2007; Caesar et al., 2018; Knight et al., 2005; Latif et al., 2004; Vellinga & Wood, 2004; R. Zhang, 2007, 2008). Consequently, available sea level and (sub)surface temperature records have been used to reconstruct AMOC changes back in time (Ezer, 2015; Frajka-Williams, 2015; Lopez et al., 2017; Ritz et al., 2013; Thornalley et al., 2018; X. Zhang et al., 2015). Moreover, sea surface height, ocean bottom pressure, and hydrographic observations at selected locations have been proposed as a useful observing system to detect AMOC changes in the present ocean and under future climate change scenarios, complementing or substituting current direct North Atlantic trans-basin transport measurements, which are limited in space and time (see Frajka-Williams et al., 2019, for a review).

Proxy potential is typically assessed by means of statistical regression or correlation (e.g., see all AMOC proxy studies referenced in the previous paragraph), including regression using statistical modes of variability obtained e.g., via principal component analysis (von Storch & Zwiers, 1999). Fig. 1(a) sketches the concept of evaluating statistical proxy potential: one assesses covariability between an observable quantity (pink time series) and an unobserved QoI (purple time series), often in model output. This method provides an empirical measure for proxy potential, but does not identify causal relations. Without dynamical underpinnings, reported dependency on model choice, forcing scenario and time period considered (Alexander-Turner et al., 2018; Little et al., 2019; Roberts & Palmer, 2012) complicates robust identification of proxy potential. The goal of this work is to overcome the limitations of statistical proxy potential. Here, we establish a new methodology that quantifies *dynamical*, rather than statistical, *proxy potential*.

Our goal is to unambiguously identify shared dynamical processes and pathways that provide a mechanistic underpinning for what we will refer to as dynamical proxy potential. To do so, we take advantage of the adjoint of an ocean general circulation model (GCM). The adjoint can efficiently uncover the dynamical cause of variations in observed and unobserved ocean quantities, extracted from the equations of motion and conservation laws governing the underlying GCM (Marotzke et al., 1999). For instance, adjoint-derived sensitivities have been used to study the dynamical cause of the following QoIs: Atlantic meridional heat transport (Heimbach et al., 2011; Köhl, 2005; Marotzke et al., 1999), AMOC (Czeschel et al., 2010; Heimbach et al., 2011; Pillar et al., 2016; Smith & Heimbach, 2019), Labrador Sea heat content (Jones et al., 2018), temperature in the east equatorial Pacific (Galanti et al., 2002; Galanti & Tziperman, 2003), and sea level on the Californian coast (Verdy et al., 2013). Building on previous studies, we exploit the adjoint in a novel fashion, as sketched in Fig. 1(b): we identify forcings (green shading) affecting both an observed quantity (e.g., temperature in the pink box) and an unobserved QoI (e.g., heat transport across the purple section). By this approach, we find dynamical causes and controls of covariability between the observed and unobserved quantity.

Moreover, we establish a link between the notion of dynamical proxy potential and two existing quantitative frameworks: representer methods (Bennett, 1985, 2002) and Hessian-based uncertainty quantification (Bui-Thanh et al., 2012; Thacker, 1989). Within

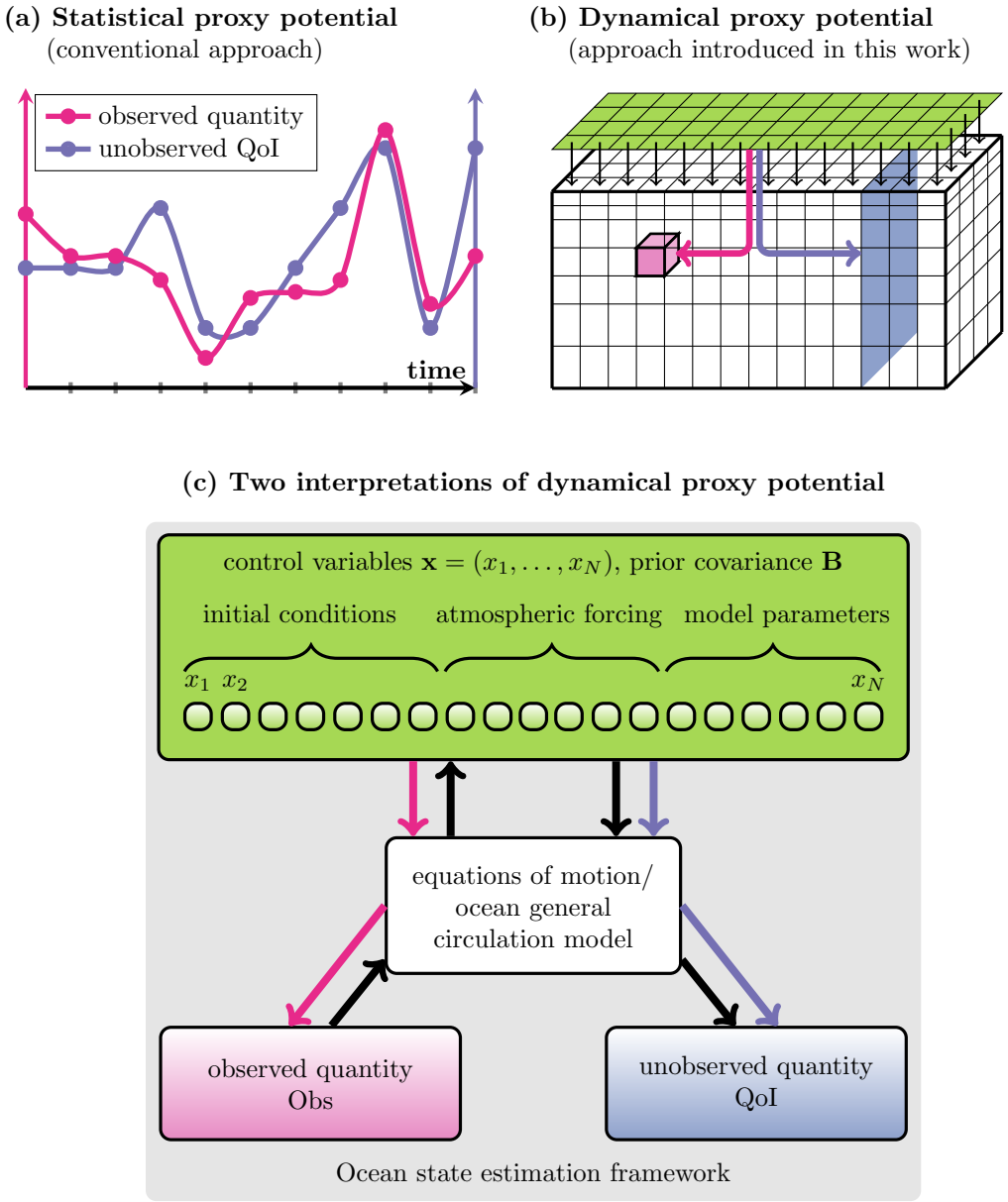

Figure 1. (a),(b) Two approaches to assess proxy potential of an observed quantity (pink) for an unobserved quantity of interest (QoI, purple): (a) statistical proxy potential assesses covariability based on empirical evidence; (b) dynamical proxy potential assesses causes (green shading) of covariability based on dynamical laws. (c) Two equivalent interpretations of dynamical proxy potential (see section 2): via (i) shared ocean adjustment physics (pink & purple arrows) and (ii) uncertainty quantification in ocean state estimation (black arrows).

the latter framework, dynamical proxy potential can be interpreted to measure uncertainty reduction in the QoI, given the new dynamical information provided by the observation, independent of the measurement value. This second interpretation of dynamical proxy potential is further developed in a forthcoming paper.

We illustrate the new concept of dynamical proxy potential for a case study in the North Atlantic, choosing heat transport across the Iceland-Scotland ridge as our exemplary QoI. The Iceland-Scotland ridge (ISR) is the key gateway for poleward heat progression from the North Atlantic toward the Arctic Ocean (Hansen & Østerhus, 2000). Warm Atlantic waters are carried across the ridge by the Norwegian Atlantic Current (NwAC), one of the main branches of the North Atlantic Current (NAC, see Fig. 2). While observational estimates for ISR heat transport since the mid 1990s exist (e.g., Berx et al., 2013; Hansen et al., 2015; Østerhus et al., 2005, 2019), cross-ridge heat transport estimates remain uncertain, due to a sparse array of current meter moorings and the sensitivity to the choice of calculation method (Berx et al., 2013; McCarthy et al., 2019). In contrast, upper ocean temperatures are well constrained throughout the larger part of the North Atlantic basin via remote sensing and in situ platforms. For this reason, we select our representative observed quantities as temperature at the sea surface and at 300 m depth, at two locations in the North Atlantic: in the Irminger Current (IC) and off the Portuguese coast (Fig. 2), monitored by the OSNAP (Lozier et al., 2017, 2019) and OVIDE (Lherminier et al., 2007; Mercier et al., 2015) sections, respectively. These locations are intentionally chosen in two branches of the NAC that are distinct from the branch crossing the ISR (Fig. 2) and are therefore not expected to be ideal placements for monitoring ISR heat transport. We will show that these observations nevertheless provide partial constraints on the QoI through shared adjustment physics, which are uncovered and quantified by dynamical proxy potential.

Here, we work within the global ECCO (Estimating the Circulation and Climate of the Ocean) version 4 state estimation framework (Forget et al., 2015) and focus on monthly to multiannual time scales up to five years, since now approximately five years of continuous OSNAP measurements are available. We note that assessment of dynamical proxy potential does not require actual (here: OSNAP and OVIDE) observational data, since it investigates dynamical relationships in the model equations, rather than observed covariability. This paper is structured as follows. Section 2 introduces our new method that quantifies dynamical proxy potential. Section 3 applies the methodology to our North Atlantic case study. In section 4, we discuss our results as well as limitations and future directions. Section 5 presents the main conclusions.

2 Quantifying Dynamical Proxy Potential

We define proxy potential of an observed quantity, Obs, for a quantity of interest, QoI, as

$$\mathrm{PP}(\mathrm{Obs};\mathrm{QoI}) = \left(\frac{\mathrm{Cov}(\mathrm{Obs},\mathrm{QoI})}{\sigma_{\mathrm{Obs}}\cdot\sigma_{\mathrm{QoI}}}\right)^2 \in [0,1]. \tag{1}$$

Here, the operators $\mathrm{Cov}(\clubsuit,\spadesuit)$ and $\sigma_{\clubsuit}$ denote covariance and standard deviation, respectively. Conventionally, these operators are evaluated statistically, in which case we refer to eq. (1) loosely as 'statistical proxy potential'. A statistical evaluation can be performed, for instance, if time series (e.g., from model output) are available, as sketched in Fig. 1(a); the right hand side of eq. (1) is then equal to the squared Pearson correlation coefficient, often referred to as r^2. To define and assess dynamical proxy potential, we require a dynamics-based evaluation of the operators in eq. (1). For this, we leverage the framework of ocean state estimation and inbuilt adjoint capability, as described in the following.

Ocean state estimation seeks to infer a best estimate from uncertain (and often sparse) ocean observations and an ocean model with uncertain inputs. The uncertain inputs are

also referred to as the control variables, collected in the vector $\mathbf{x} = (x_1, \ldots, x_N)$, and typically consist of spatio-temporal varying atmospheric forcing, initial conditions and certain model parameters (green box in Fig. 1(c), or Forget et al., 2015). An assigned $N \times N$ covariance matrix $\mathbf{B}$ spells out assumptions on the prior uncertainty in the control variables (Tarantola, 2005; Wunsch, 1996). Ocean state estimation then fits the model to the available observations, by adjusting (or 'inverting' for) the control variables within their prescribed uncertainty. In contrast, the model state variables (within the white centered box in Fig. 1(c), e.g., temperature and velocity) adjust freely in response to the adjusted control variables, following the model dynamics to ensure dynamical and kinematic consistency. An implicit assumption in ocean state estimation is that the control variables comprise all possible sources of changes in the ocean state and circulation.

Ocean state estimation offers a comprehensive framework to quantify proxy potential (eq. (1)) and its dynamical origins, where the candidates for proxy origin are formally provided by the control variables. Within this framework, the covariance between Obs and QoI is the scalar

$$\mathrm{Cov}(\mathrm{Obs}, \mathrm{QoI}) = [\nabla_{\mathbf{x}} \mathrm{QoI}]^T \, \mathbf{B} \, \nabla_{\mathbf{x}} \mathrm{Obs}. \tag{2}$$

Here, the gradient $\nabla_{\mathbf{x}}(\clubsuit) = [\partial(\clubsuit)/\partial x_1, \ldots, \partial(\clubsuit)/\partial x_N]^T$ is the vector whose i^{th} component is the linearized sensitivity of the scalar $\clubsuit \in \{\mathrm{QoI}, \mathrm{Obs}\}$ to the control variable x_i, and $[\nabla_{\mathbf{x}}(\clubsuit)]^T$ is its transpose, i.e., the associated row vector. The covariance in eq. (2) can be thought of as being computed from right to left (from the observed quantity via the controls to the QoI), following the black arrows in Fig. 1(c), by means of the adjoint and tangent linear models (Errico, 1997). Importantly, the resulting covariance is consistent with ocean dynamical laws. These dynamical laws are baked into the gradients in eq. (2) via the chain rule, which passes through the equations of motion that are encoded in the underlying ocean GCM (white centered box, Fig. 1(c)). Similarly to eq. (2), the standard deviation of the quantity $\clubsuit \in \{\mathrm{Obs}, \mathrm{QoI}\}$ is

$$\sigma_{\clubsuit} = \sqrt{[\nabla_{\mathbf{x}}(\clubsuit)]^T \, \mathbf{B} \, \nabla_{\mathbf{x}}(\clubsuit)}. \tag{3}$$

The fact that the expressions in eqs. (2),(3) are the dynamics-based analogue of purely statistically derived covariance and standard deviations was established by Bennett (1985, 1990). The values computed in eqs. (2),(3) would be entries in the so-called representer matrix (Bennett, 2002), if both quantities, Obs and QoI, were part of the model state variables.

In this work, we shed new light on the concept of dynamics-based covariances (or 'representers'). To this aim, we introduce the notion of dynamical proxy potential (DPP). Inserting eqs. (2),(3) into eq. (1), and rearranging terms, we obtain the following definition of DPP:

$$\mathrm{DPP}(\mathrm{Obs}; \mathrm{QoI}) = \Big(\underbrace{\left[\sigma_{\mathrm{QoI}}^{-1} \cdot \mathbf{B}^{1/2} \, \nabla_{\mathbf{x}} \mathrm{QoI}\right]}_{=\mathbf{q}} \bullet \underbrace{\left[\sigma_{\mathrm{Obs}}^{-1} \cdot \mathbf{B}^{1/2} \, \nabla_{\mathbf{x}} \mathrm{Obs}\right]}_{=\mathbf{v}} \Big)^2 \in [0, 1], \tag{4}$$

where $\mathbf{B}^{1/2}$ is the $N \times N$ matrix which satisfies $\mathbf{B}^{1/2} \, \mathbf{B}^{1/2} = \mathbf{B}$, and $\bullet$ denotes the dot product in $\mathbb{R}^N$. The bounds in eq. (4) correspond to the cases for which Obs provides *no* information (DPP = 0), and serves as a *perfect* proxy for the QoI (DPP = 1), similar to the statistics-derived r^2. DPP is fully determined by the projection ($\bullet$) of the vectors $\mathbf{q}$ and $\mathbf{v}$ (defined in eq. (4)). $\mathbf{q}$ and $\mathbf{v}$ contain the sensitivities of QoI and Obs to all controls, allowing for two equivalent interpretations: (i) $\mathbf{q}$ and $\mathbf{v}$ reveal the dynamical adjustment mechanisms of QoI and Obs, respectively (sections 3.2 and 3.3); (ii) $\mathbf{q}$ specifies the *information required* to recover the QoI, while $\mathbf{v}$ is the *information captured* by the observed quantity. Interpretation (ii) is rooted in Hessian-based uncertainty quantification (UQ) and optimal observing system design within ocean state estimation. This link to UQ is derived in the supporting information (Text S1) and further explored in a forthcoming paper. In the definition of $\mathbf{q}, \mathbf{v}$ (eq. (4)), multiplication with the matrix

$\mathbf{B}^{1/2}$ (prior-)weights the sensitivity vectors $\nabla_{\mathbf{x}}(\clubsuit)$, and division by the scalar $\sigma_{\clubsuit}$ acts as normalization. Indeed, eq. (3) can be rewritten as the l^2-norm of the weighted sensitivity vector:

$$\sigma_{\clubsuit} = \left\| \mathbf{B}^{1/2} \nabla_{\mathbf{x}}(\clubsuit) \right\|. \tag{5}$$

Just like an entry in the representer matrix (eq. (2) or Bennett, 2002), DPP quantifies the dynamical information content of an observation for a QoI with a single number (eq. (4)). However, DPP places emphasis on *unfolding* this information content (collapsed into a single number via the dot product $\bullet$) and illuminating its dynamical causes (see Figs. 6, 7), through inspection of $\mathbf{q}$ and $\mathbf{v}$.

Fig. 1(c) offers a schematic summary of understanding DPP(Obs; QoI) (eq. (4)) in line with our two equivalent interpretations. Dynamical proxy potential of an observed quantity, Obs (pink box), for an unobserved quantity of interest, QoI (purple box), measures

(i) the similarity between the ocean adjustment physics for the observed vs. unobserved quantity (pink vs. purple arrows) in response to changes in forcing (green box), on a scale from 0% (*no similarity*) to 100% (*identical*);
(ii) the relative uncertainty reduction in the QoI that would be achieved if the observation were to be added without noise to the state estimation framework in Fig. 1(c) (see Text S1). The flow of information and uncertainty reduction within the state estimation framework - from the observation via the controls to the QoI - is delineated by the black arrows in Fig. 1(c).

3 Application to the North Atlantic

This section illustrates our method for a case study in the North Atlantic. Section 3.1 describes the experimental setup, including our choice of QoI and observations. Sections 3.2 and 3.3 present the adjustment mechanisms of the QoI and observations. Section 3.4 assesses the degree to which these adjustments are shared, and quantifies the dynamical proxy potential of the observations for the QoI.

3.1 Experimental Setup

We perform our experiments using the ECCO version 4 release 2 (ECCOv4r2, Forget et al., 2015) solution. The Massachusetts Institute of Technology general circulation model (MITgcm, J. Marshall, Adcroft, et al., 1997; J. Marshall, Hill, et al., 1997), serves as the dynamical core in ECCO and is configured at a nominal horizontal resolution of 1° with 50 vertical levels. The optimized state provides an acceptable fit to most available oceanographic data and has been used extensively for mechanistic investigations of ocean variability, including in the North Atlantic (e.g., Buckley et al., 2014; Jones et al., 2018). We refer the reader to Forget et al. (2015) for details on the model configuration and estimated ocean state.

To quantify dynamical proxy potential and its origins, one requires the linear sensitivities of the QoI and observed quantities to all control variables ($\nabla_{\mathbf{x}}$QoI and $\nabla_{\mathbf{x}}$Obs, eq. (4)). To perform these sensitivity calculations, we take advantage of the flexible ECCOv4 adjoint modeling framework (Forget et al., 2015). Algorithmic differentiation, through source-to-source code transformation with the commercial tool Transformation of Algorithms in Fortran (TAF, Giering & Kaminski, 1998), produces the code for our adjoint models. Ice-covered regions are masked in the sensitivity calculation.

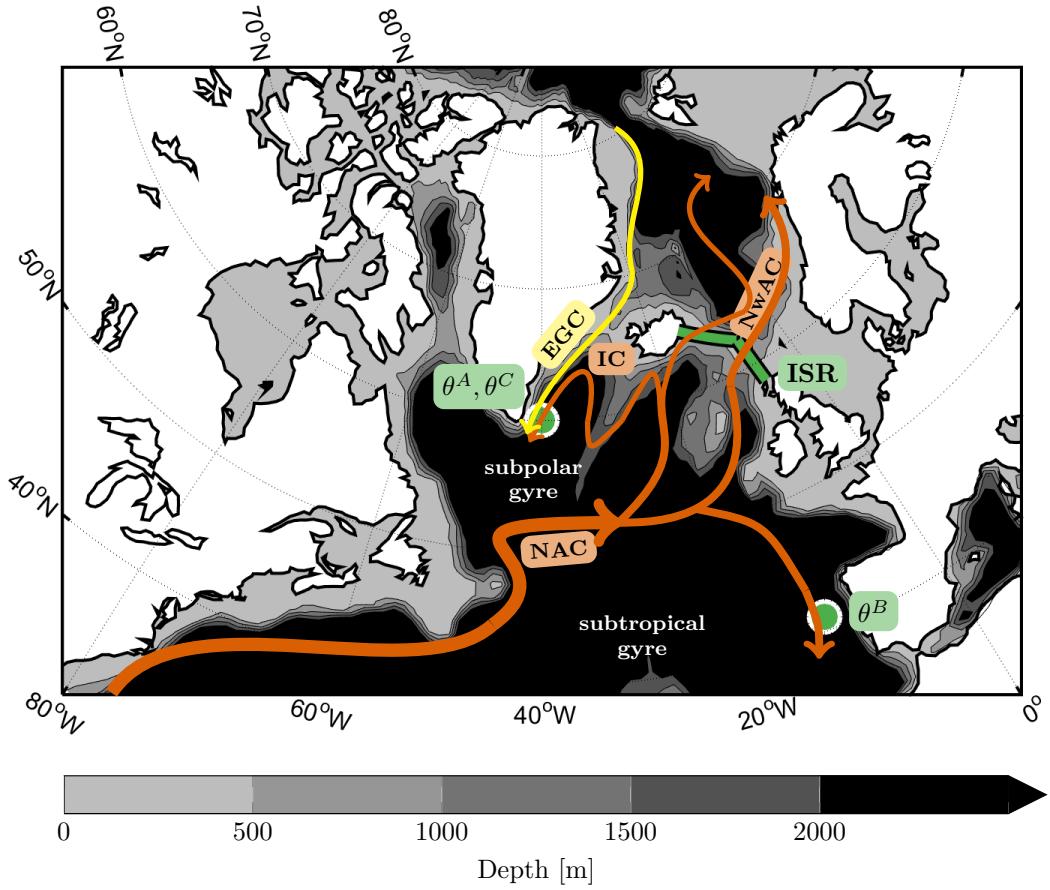

Figure 2. Schematic of the North Atlantic quantities examined in our case study. The quantity of interest (QoI) is heat transport across the Iceland-Scotland ridge (ISR, green line). The temperature observations θ^A, θ^B and θ^C are located inside the green dots. θ^A and θ^B are subsurface (at 300 m depth), θ^C at the sea surface. The arrows represent approximate pathways of major near-surface currents carrying warm, saline Atlantic waters (orange) and cold, fresh Arctic waters (yellow): NAC = North Atlantic Current; NwAC = Norwegian Atlantic Current; IC = Irminger Current; EGC = East Greenland Current.

3.1.1 QoI and Observations

The QoI in our case study is heat transport across the Iceland-Scotland ridge, denoted by $\mathrm{HT_{ISR}}$. We investigate three different temperature observations in the North Atlantic, located inside the green dots in Fig. 2 and labeled by θ^A, θ^B, and θ^C. Observations θ^A and θ^C are located in the Irminger Sea at $(40\,^\circ\mathrm{W}, 60\,^\circ\mathrm{N})$, while observation θ^B is situated in the eastern North Atlantic off the Portuguese coast at $(12\,^\circ\mathrm{W}, 41\,^\circ\mathrm{N})$. θ^A and θ^B are subsurface observations, situated at 300 m depth, and θ^C is a surface observation.

We quantify the dynamical proxy potential of the five-year mean of the observations for the five-year mean of our QoI, for zero lag. Sensitivities of the QoI and observations (eq. (4)) are computed from the final five years (2007-2011) of the ECCOv4r2 state estimate. Dependence on the specific evaluation period and background state is weak, given that $\mathrm{HT_{ISR}}$, θ^A, θ^B, and θ^C depend approximately linearly on the control variables (Appendix A).

The QoI, as simulated by the model, is diagnosed as follows:

$$\mathrm{HT_{ISR}} = \frac{\rho_0\, c_p}{\Delta t} \int_{2007}^{2011} \int_{\mathrm{bottom}}^{\mathrm{top}} \int_L (\theta - \theta_{\mathrm{ref}})\, v^\perp \, dL\, dz\, dt \quad [\mathrm{W}]. \tag{6}$$

L denotes the Iceland-Faroe-Scotland line segment, $\Delta t = \int_{2007}^{2011} dt$ the length of the integration period, $\rho_0 = 1029\,\mathrm{kg/m^3}$ the reference density, and $c_p = 3994\,\mathrm{J/(kg \cdot K)}$ the specific heat capacity of water. θ denotes potential temperature, and $v^\perp$ the velocity perpendicular to the line segment L; sign convention is such that positive $v^\perp$ corresponds to positive north- and eastward velocity. Note that since L is only a partial line segment, rather than a closed boundary, heat transport in eq. (6) has to be defined relative to a reference temperature θ_{ref} (Schauer & Beszczynska-Möller, 2009). Consistent with many observational studies (e.g., Berx et al., 2013; Hansen et al., 2015; Østerhus et al., 2005), we choose $\theta_{\mathrm{ref}} = 0\,^\circ\mathrm{C}$, motivated by the observation that southward flow across the ISR is close to this temperature (Hansen et al., 2003).

For $\star \in \{A, B, C\}$, the observation $\theta^\star$ is diagnosed as the mean potential temperature

$$\theta^\star = \frac{1}{\Delta t \cdot \mathcal{V}^\star} \int_{2007}^{2011} \int_{h_0^\star}^{h_1^\star} \int_{\mathcal{A}^\star} \theta \, dx\, dy\, dz\, dt \quad [^\circ\mathrm{C}]. \tag{7}$$

$(h_0)^\star$, $(h_1)^\star$, and $\mathcal{A}^\star$ denote the lower and upper boundaries, and the horizontal area, of the model grid cell in which the respective observation $\theta^\star$ is located. For the subsurface observations ($\star = A, B$), we have $(h_0)^\star = -325\,\mathrm{m}$ and $(h_1)^\star = -275\,\mathrm{m}$. For the surface observation, we choose the uppermost two model grid cells as a representative depth range, corresponding to $(h_0)^C = -20\,\mathrm{m}$ and $(h_1)^C = 0\,\mathrm{m}$. The area of $\mathcal{A}^A = \mathcal{A}^C$ is approximately $(52\,\mathrm{km})^2$ and the area of $\mathcal{A}^B$ is approximately $(84\,\mathrm{km})^2$. In eq. (7), $\theta(x, y, z, t)$ denotes potential temperature, $\Delta t = \int_{2007}^{2011} dt$ the length of the integration period and $\mathcal{V}^\star = \int_{h_0^\star}^{h_1^\star} \int_{\mathcal{A}^\star} dx\, dy\, dz$ the volume of interest.

3.1.2 Control Variables and Weights

Table 1 lists the set of control variables that is chosen in this work: the spatially-varying forcing fields $F_m(i, j)$ of net upward surface heat flux, $Q_{\mathrm{net},\uparrow}$, net surface freshwater flux, EPR, and zonal and meridional wind stress, τ_x and τ_y, respectively. Consistent with assessing dynamical proxy potential of the five-year mean of the observations for the five-year mean of the QoI, only adjustments to changes in the five-year mean of the forcing fields are considered (fourth column in Table 1). For the sake of a simpler presentation, initial conditions and model parameter fields are omitted, even though they are part of the uncertain inputs in a full ocean state estimation framework (green box, Fig. 1(c)). The four two-dimensional forcing fields are flattened and concatenated into

Table 1. Control Variables and Weights in our Case Study.

m	Forcing $F_m(i,j)$	Symbol	Time average	ΔF_m
1	Net upward surface heat flux	$Q_{\text{net},\uparrow}$	five years	$50\,\text{W/m}^2$
2	Net surface freshwater flux	EPR	five years	$5\cdot10^{-8}\,\text{m/s}$
3	Zonal wind stress	τ_x	five years	$0.05\,\text{N/m}^2$
4	Meridional wind stress	τ_y	five years	$0.05\,\text{N/m}^2$

a long vector, $\mathbf{x}=(x_1,\ldots,x_N)$. The length of the vector, N, is $O(10^6)$, equal to 4 times the number of model surface grid cells covering the global ocean.

For each of the four forcing fields, F_m, we set a spatially constant prior standard deviation, ΔF_m (last column in Table 1). Further, we assume the decorrelation length in the surface forcing to be less than the grid scale ($\sim 1\,^\circ$). Meanwhile, ECCOv4r2 uses spatially varying prior standard deviations, estimated based on the spread between different reanalysis products (Chaudhuri et al., 2013), and sets a decorrelation length of 3 times the grid scale within the same forcing field (but no cross-correlations between distinct forcing fields, Forget et al., 2015). Our choices correspond to a diagonal prior covariance $\mathbf{B}$, and its square root is the diagonal $N\times N$ matrix

$$\mathbf{B}^{1/2}=\text{diag}(\underbrace{\Delta F_1,\ldots,\Delta F_1}_{N/4\text{ times}},\underbrace{\Delta F_2,\ldots,\Delta F_2}_{N/4\text{ times}},\underbrace{\Delta F_3,\ldots,\Delta F_3}_{N/4\text{ times}},\underbrace{\Delta F_4,\ldots,\Delta F_4}_{N/4\text{ times}}). \tag{8}$$

Our assumption of spatially uniform weights ΔF_m and no prior spatial cross-correlations implies that the sensitivity projection (eq. (4)) for each individual forcing field is fully determined by the adjustment physics, and not by the forcing weights. Our simplified choice of forcing covariance therefore adds clarity to the presentation in this paper, whose primary goal is to explain the new concept of dynamical proxy potential.

3.2 Adjustment Mechanisms of the QoI

Figs. 3(a)-(d) show the weighted and normalized sensitivities of the five-year mean heat transport across the Iceland-Scotland ridge (HT_{ISR}):

$$\mathbf{q}_{|F_m(i,j)}=\sigma_{\text{HT}}^{-1}\frac{\partial(\text{HT}_{\text{ISR}})}{\partial F_m(i,j)}\Delta F_m,\qquad m=1,2,3,4. \tag{9}$$

Here, $F_m(i,j)$ are the five-year mean atmospheric forcing fields from Table 1, and ΔF_m their spatially uniform weights. The normalization factor, $\sigma_{\text{HT}}=11\,\text{TW}$, is computed according to eqs. (5),(8), with $\clubsuit=\text{HT}_{\text{ISR}}$. The weighted and normalized sensitivities of HT_{ISR} in eq. (9) (or Figs. 3(a)-(d)) assemble the vector $\mathbf{q}$ (cf. eqs. (4),(8)), which has two equivalent interpretations (section 2): (i) $\mathbf{q}$ reveals all adjustment mechanisms of HT_{ISR}, as will be discussed in the following paragraphs; (ii) $\mathbf{q}$ is the information required to recover HT_{ISR}, our QoI. The bar chart in Fig. 3(e) shows the relative importance of the four forcings F_m for impacting HT_{ISR}. Relative importance is measured by the ratios

$$\left\|\mathbf{q}_{|F_m}\right\|^2=\sigma_{\text{HT}}^{-2}\sum_{i,j}\left(\frac{\partial(\text{HT}_{\text{ISR}})}{\partial F_m(i,j)}\Delta F_m\right)^2,\qquad m=1,2,3,4, \tag{10}$$

equivalent to integrating the sensitivities in Figs. 3(a)-(d) around the globe (in the l^2-norm). Fig. 3(e) demonstrates that the influence of wind stress, τ_x and τ_y, prevails over the influence of buoyancy forcing, $Q_{\text{net},\uparrow}$ and EPR.

The positive sensitivity of HT_{ISR} to τ_y along the western African and European coast (Fig. 3(d)) is consistent with the following dynamical mechanism. An increase in

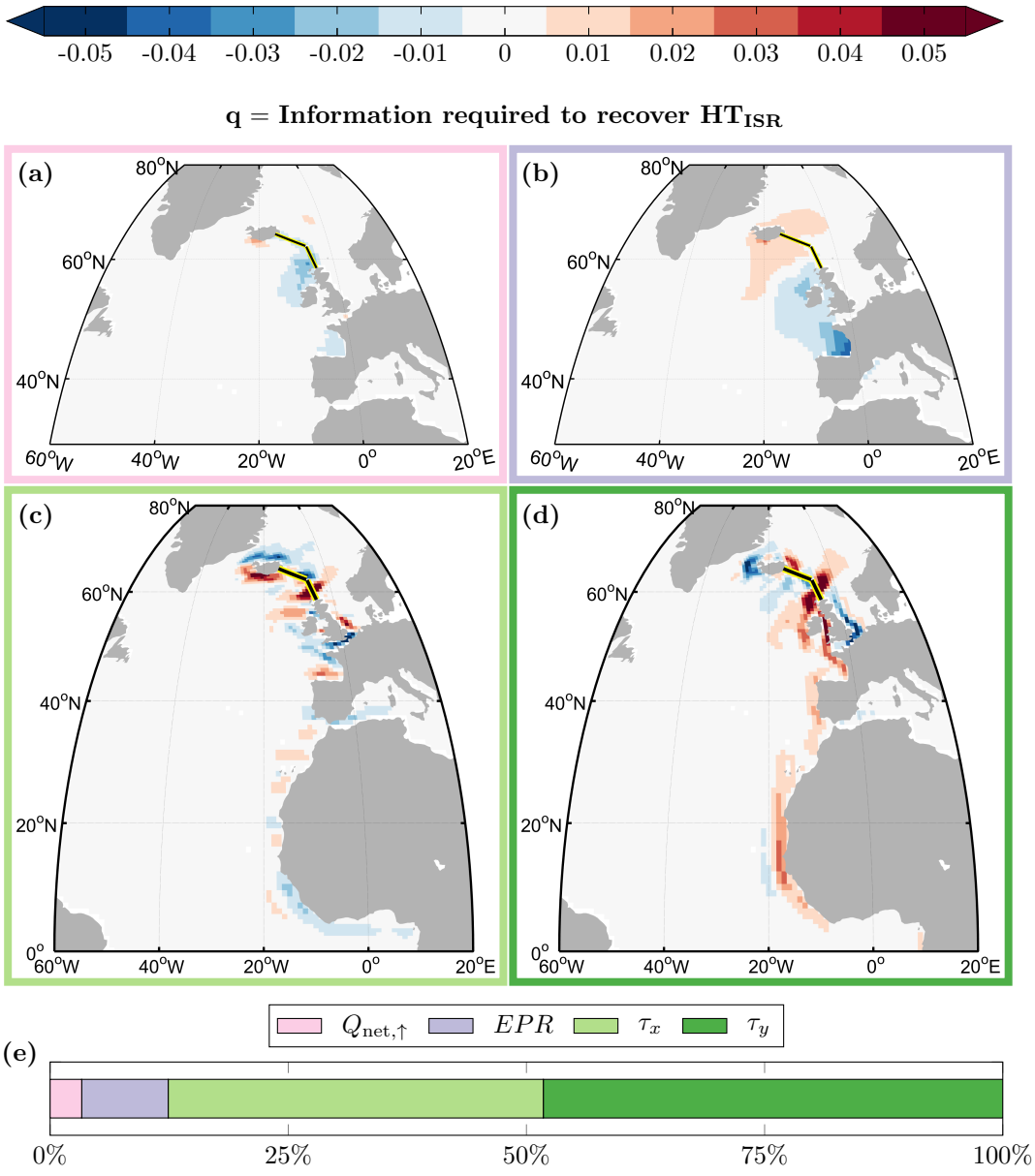

Figure 3. Sensitivities of five-year mean heat transport across the Iceland-Scotland ridge ($\mathrm{HT}_{\mathrm{ISR}}$), to changes in the five-year mean (a) upward surface heat flux $Q_{\mathrm{net},\uparrow}$, (b) surface freshwater flux EPR, (c) zonal wind stress τ_x and (d) meridional wind stress τ_y. The sensitivities are weighted and normalized (thus unitless), and assemble the vector $\mathbf{q}$ (eq. (9)). Red (blue) colors indicate that an increase in (a) heat loss to the atmosphere, (b) surface salinification, (c) eastward wind stress and (d) northward wind stress would lead to a subsequent increase (decrease) in $\mathrm{HT}_{\mathrm{ISR}}$ on a five-year time scale. The solid black-yellow contour in (a)-(d) delineates the ISR. The bar chart in (e) shows the relative contributions of $Q_{\mathrm{net},\uparrow}$, EPR, τ_x and τ_y to $\mathrm{HT}_{\mathrm{ISR}}$ sensitivity, when integrating the sensitivities in (a)-(d) around the globe (eq. (10)).

northward wind stress along the western African and European coast induces Ekman onshore convergence and a positive pressure anomaly along the coast. Boundary waves (e.g., D. P. Marshall & Johnson, 2013) propagate the positive pressure anomaly cyclonically around the North Atlantic basin (Fig. 5(a)). When the positive pressure anomaly reaches the eastern end of the ISR (within a month, see Movie S1), it leads to an increased along-ridge pressure gradient and, by geostrophic balance, a strengthened $\mathrm{HT_{ISR}}$. While the $\mathrm{HT_{ISR}}$ anomaly develops rapidly (after a few months, Fig. A1(b)), it persists for as long as the wind stress perturbation is maintained (here: for 5 years). The sensitivity to zonal wind stress along the western African and European coastline (Fig. 3(c)) is due to the same mechanism. Here, the sensitivity sign alternates because it is determined by the orientation of the coastline.

The wind stress sensitivities of $\mathrm{HT_{ISR}}$ around Iceland and the United Kingdom (UK) that emerge in Figs. 3(c),(d) can be explained similarly. By way of illustration, *negative* sensitivity of $\mathrm{HT_{ISR}}$ to τ_y along the western coast of Iceland (Fig. 3(d)) is consistent with the following mechanism. An increase in northward wind stress along the western Icelandic coast drives Ekman onshore convergence, resulting in a positive pressure anomaly at the Icelandic coast. Through clockwise wave propagation around the Icelandic coastline, the positive pressure anomaly is rapidly communicated to the western end of the ISR (Fig. 5(b), Movie S7). The resulting negative anomaly in the along-ridge pressure gradient leads to a *weakening* of $\mathrm{HT_{ISR}}$ and, consequently, colder temperatures in the Norwegian Sea (Fig. 5(d)). As before, the $\mathrm{HT_{ISR}}$ anomaly develops rapidly, but persists for 5 years (Fig. A1(d)). This mechanism also explains the sign of the τ_x sensitivities around Iceland: increased eastward (westward) wind stress along the southern (northern) coast of Iceland (Fig. 3(c)) drive Ekman offshore divergence, resulting in a negative pressure anomaly at the Icelandic coast and a subsequent increase in $\mathrm{HT_{ISR}}$. Since the UK coastline delivers pressure signals to the eastern (rather than the western) end of the ISR, UK-originated pressure anomalies increase $\mathrm{HT_{ISR}}$ if they are positive (rather than negative). This explains the fact that the sensitivity dipoles around Iceland and the UK are of opposite sign (Figs. 3(c),(d)).

$\mathrm{HT_{ISR}}$ shows positive sensitivity to $Q_{\mathrm{net},\uparrow}$ and EPR to the west of the ISR, around Iceland, and negative sensitivity to the east of the ISR, along the western European coast (Figs. 3(a),(b)). This sensitivity dipole across the core of the NAC is consistent with a strengthening of the cross-ridge geostrophic transport in response to a negative perturbation of the density gradient along the section. The sensitivity of $\mathrm{HT_{ISR}}$ to $Q_{\mathrm{net},\uparrow}$, relative to the remaining forcing fields, is surprisingly small: only 3% (Fig. 3(e)). We note that even if we tripled ΔQ_{net} in Table 1, while keeping the weights for the remaining forcings unchanged, $\mathrm{HT_{ISR}}$ would still be less sensitive to $Q_{\mathrm{net},\uparrow}$ than to any of the remaining three forcing fields in Fig. 3(e). This is consistent with previous observation- and model-based studies, which found that on seasonal to multiannual time scales ISR heat transport variability is predominantly driven by velocity fluctuations, rather than temperature fluctuations (Årthun & Eldevik, 2016; Asbjørnsen et al., 2019; Orvik & Skagseth, 2005).

3.3 Adjustment Mechanisms of the Observations

The weighted and normalized sensitivities of the five-year mean temperature observations $\theta^\star$, $\star = A, B, C$, are given by

$$\mathbf{v}^\star_{|F_m(i,j)} = \sigma_\star^{-1} \frac{\partial \theta^\star}{\partial F_m(i,j)} \Delta F_m, \qquad m = 1,2,3,4, \tag{11}$$

similar to eq. (9). The normalization factors $\sigma_\star$ are computed according to eqs. (5),(8), with $\clubsuit = \theta^\star$, giving $\sigma_A = 0.05\,^\circ\mathrm{C}$, $\sigma_B = 0.06\,^\circ\mathrm{C}$, and $\sigma_C = 0.23\,^\circ\mathrm{C}$. Note that σ_C is much larger than σ_A and σ_B since the surface temperature θ^C is more sensitive to atmospheric forcing than the subsurface temperatures θ^A, θ^B. Figs. 4(a)-(f) show the weighted

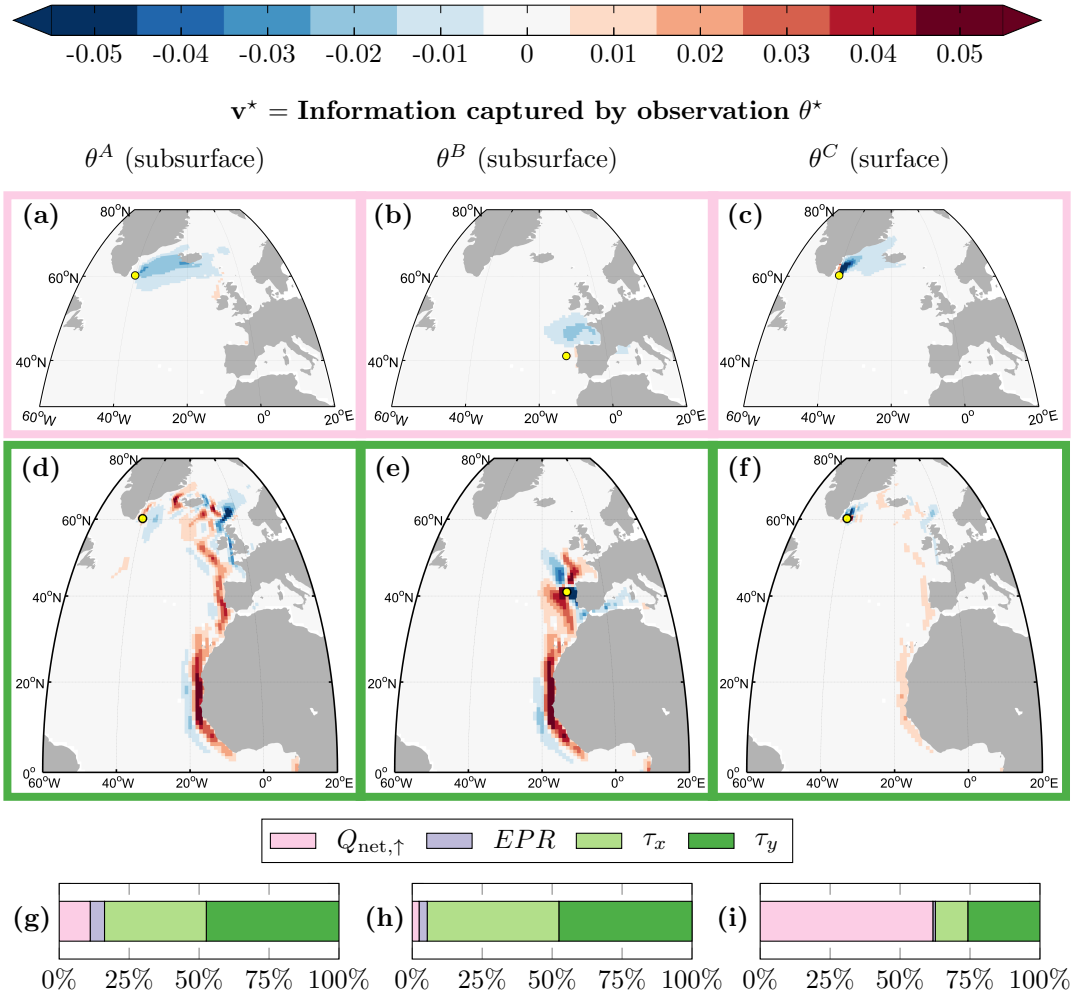

Figure 4. (a)-(c): Sensitivities of five-year mean (a) subsurface temperature in the Irminger Sea (θ^A), (b) subsurface temperature off the Portuguese coast (θ^B), and (c) surface temperature in the Irminger Sea (θ^C) to changes in five-year mean upward surface heat flux $Q_{\text{net},\uparrow}$. (d)-(f): Same as (a)-(c), but sensitivities to meridional wind stress τ_y. The sensitivities are weighted and normalized (thus unitless), and assemble the vector $\mathbf{v}^\star$ (eq. (11)). Red (blue) colors indicate that an increase in (a)-(c) heat loss to the atmosphere and (d)-(f) northward wind stress would lead to a subsequent increase (decrease) in (a),(d) θ^A, (b),(e) θ^B, and (c),(f) θ^C on a five-year time scale. The yellow dots mark the respective locations of the temperature observation. The bar charts in (g)-(i) show the relative contributions of $Q_{\text{net},\uparrow}$, EPR, τ_x and τ_y to (g) θ^A, (h) θ^B, and (i) θ^C sensitivity, computed as in Fig. 3.

and normalized sensitivities (eq. (11)) for two of the four forcings, $F_1 = Q_{\text{net},\uparrow}$ and $F_4 = \tau_y$. The vector $\mathbf{v}^\star$, composed of the weighted and normalized sensitivities in eq. (11), has again two equivalent interpretations: (i) $\mathbf{v}^\star$ reveals the adjustment mechanisms of $\theta^\star$, which will be discussed in the following; (ii) $\mathbf{v}^\star$ is the information captured by the observation $\theta^\star$. The bar charts in Figs. 4(g)-(i) show the relative importance of the four forcings F_m for impacting $\theta^\star$, for $\star = A, B, C$. Relative importance is measured as in Fig. 3(e), by the ratios in eq. (10), where $\mathbf{q}$ is substituted by $\mathbf{v}^\star$.

The relative importance of $Q_{\text{net},\uparrow}$ is high for the surface observation θ^C (Fig. 4(i)), but low for the subsurface observations θ^A and θ^B (Figs. 4(g),(h)). The high sensitivity of θ^C to $Q_{\text{net},\uparrow}$ is concentrated at the observed site (Fig. 4(c)), due to the strong influence of local air-sea heat fluxes on surface temperature. All temperature observations show weak negative $Q_{\text{net},\uparrow}$ sensitivity upstream of the respective observed sites (Figs. 4(a)-(c)), as an increased upward heat flux locally cools surface waters which are then advected (Fig. 2) to the observed locations. For all three temperature observations, the relative importance of EPR is very small (Figs. 4(g)-(i)). Wind stress is important for all three observations (Figs. 4(g)-(i)), and the remainder of this section is devoted to wind stress sensitivities. For the sake of brevity, we focus on τ_y sensitivities, which can be regarded as representative for τ_x sensitivities, too. Indeed, τ_x and τ_y sensitivities emerge along the same pathways (not shown) due to the same wind-driven adjustment mechanisms.

All observations are characterized by a sensitivity dipole local to the observed site, consistent with Ekman dynamics. For instance, at $(12\,^\circ\text{W}, 41\,^\circ\text{N})$, right where θ^B is located, a sensitivity dipole is visible, with positive sensitivities to the west and negative sensitivities to the east (Fig. 4(e)), interrupting the otherwise positive sensitivities along the eastern boundary of the North Atlantic. Here, Ekman theory predicts that a wind stress perturbation matching the sensitivity dipole (i.e., increased northward wind stress to the west and increased southward wind stress to the east) causes Ekman downwelling and pumps warm surface waters down to the subsurface observation, which increases θ^B.

The large-scale wind stress sensitivity patterns of θ^C (Fig. 4(f)) are very similar to the ones of θ^A (Fig. 4(d)), except that they are of much weaker amplitude. The similarity of the patterns suggests that the surface observation θ^C is sensitive to similar remote wind-driven adjustment mechanisms as the subsurface observation θ^A. However, local forcing massively dominates the surface temperature response, as indicated by the strong sensitivities concentrated near $(40\,^\circ\text{W}, 60\,^\circ\text{N})$ in Figs. 4(c),(f).

For all three temperature observations, positive sensitivity to northward wind stress emerges along the western African and European coastline (Figs. 4(d)-(f)), similar to what was seen for HT_{ISR} in section 3.2. To explain the underlying mechanism, we perform a perturbation experiment, in which the final five years of the ECCOv4r2 solution serve as our control simulation. We increase northward wind stress in the region highlighted in Figs. 5(a),(c), along the western African coast, by $0.05\,\text{N/m}^2$, and maintain the perturbation over the full five-year period. Fig. 5(c) shows the response anomalies in subsurface temperature, at a depth of 300 m, time-averaged over the five-year experiment. We see that, in response to the positive northward wind stress anomaly along the western African coast, the northeast Atlantic (north of $25\,^\circ\text{N}$) experiences anomalous high temperatures.

The responsible mechanism operates exactly as demonstrated by Jones et al. (2018), see their Fig. 10. The northward wind stress anomaly creates a positive pressure anomaly along the eastern boundary of the North Atlantic, which, after cyclonic propagation around the basin, sets up an anomalous pressure gradient between the Nordic Seas and the North Atlantic (Fig. 5(a), Movie S1). The basin-scale pressure gradient along the northern boundary of the North Atlantic spins up the subpolar gyre (Movie S2), leading to a warming of the subpolar North Atlantic after 1-2 years (Figs. 5(c), A1(a), Movie S3). The large-scale warming in the subtropical North Atlantic, north of $25\,^\circ\text{N}$ (Fig. 5(c)), is the result

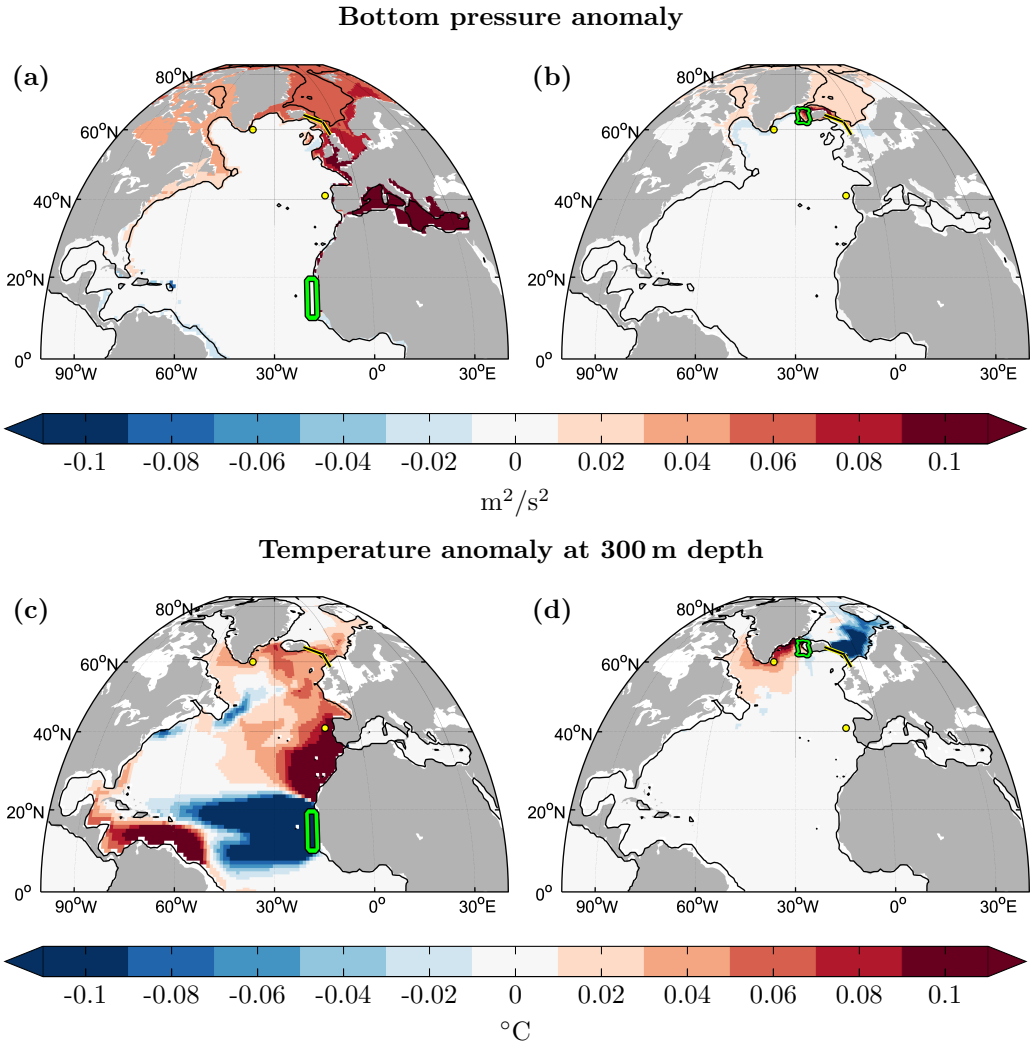

Figure 5. Anomaly in North Atlantic (a),(b) bottom pressure (normalized by density, p/ρ) and (c),(d) potential temperature at 300 m depth, in response to a positive northward wind stress anomaly of amplitude 0.05 $\mathrm{N/m^2}$ along the (a),(c) western African coast and (b),(d) western Icelandic coast. The wind stress perturbations are imposed inside the green contour in (a)-(d), and maintained over five years. The anomalies shown are time-averaged over the same five-year time period. The black line marks the (a),(b) 1000 m, (c),(d) 300 m depth contour. The yellow dots and black-yellow line show the locations of the temperature observations θ^A, θ^B, and the Iceland-Scotland ridge. Movies of the monthly evolution of these anomalies are shown in the supporting information (Movies S1, S7, S3, S9).

of baroclinic Rossby waves propagating a warm temperature anomaly from the eastern boundary westward (Movies S3, S4). The anomalous warming includes the locations of the temperature observations θ^A, θ^B, and θ^C (yellow dots, Fig. 5(c)), explaining the consistently positive sensitivities along the western African coast in Figs. 4(d)-(f). Other accompanying temperature adjustments - most notably, the cooling in the subtropical gyre south of 20 °N (Fig. 5(c), Movies S3, S6) - do not impact the temperature observations (nor the QoI, $\mathrm{HT_{ISR}}$) on a five-year time scale, but may come into play on longer time scales.

The Irminger Sea observations also show sensitivity to wind stress in the northeastern Atlantic, between 50 °N and 70 °N (Figs. 4(d),(f)). In this region, τ_y sensitivities of θ^A, θ^C (Figs. 4(d),(f)) have a similar pattern as τ_y sensitivities of $\mathrm{HT_{ISR}}$ (Fig. 3(d)), except that sensitivities of θ^A, θ^C are of opposite sign to those of $\mathrm{HT_{ISR}}$ (see Figs. 6(c),(d) for a side-by-side comparison). To explain the opposite signs, we perform a second perturbation experiment similar to the one presented in Figs. 5(a),(c). In the second experiment, we increase northward wind stress along the western Icelandic coast, in the region highlighted in Figs. 5(b),(d), where θ^A and θ^C show positive sensitivity (Figs. 4(d),(f)) and $\mathrm{HT_{ISR}}$ shows negative sensitivity (Fig. 3(d)). Fig. 5(d) shows the response anomaly in subsurface temperature, at a depth of 300 m. The Irminger and Labrador Seas experience a warming, while the Norwegian Sea cools.

The underlying mechanism is the following: the northward wind stress anomaly along the western Icelandic coast drives Ekman onshore convergence and a positive pressure anomaly, as discussed in section 3.2. The positive pressure anomaly is rapidly communicated along the entire Icelandic coastline, resulting in an across-bathymetry pressure gradient (Fig. 5(b), Movie S7), which drives an anomalous clockwise barotropic circulation around Iceland (Movie S8). The anomalous clockwise circulation around Iceland weakens the northward transport across the ISR by the NwAC as well as the southward transport through Denmark Strait by the EGC, while strengthening the IC (cf. Fig. 2). The weakened northward transport of warm Atlantic waters across the ISR leads to the anomalous cold temperatures that are seen in the Norwegian Sea in Fig. 5(d) (and Movie S9), and is consistent with a reduced $\mathrm{HT_{ISR}}$, as predicted by the *negative* sensitivities in Fig. 3(d). The weakened southward transport of cold Arctic waters through Denmark Strait, together with the strengthened IC, results in the anomalous warming that is seen in the Irminger and Labrador Seas in Fig. 5(d) (see also Fig. A1(c), Movie S9). The increased temperature in the Irminger Sea is consistent with the *positive* sensitivities along the western Icelandic coast in Figs. 4(d),(f).

The perturbation experiment presented in Figs. 5(b),(d) explains the opposite sign in the sensitivities along the western Icelandic coast in Fig. 3(d) vs. Fig. 4(d). The fact that in Fig. 3(d) vs. Fig. 4(d), sensitivities are *consistently* of opposite sign in the northeast Atlantic between 50 °N and 70 °N can be understood similarly. The sensitivity patterns in this region are characterized by topographically steered bands, which connect to Iceland or the ISR (Figs. 3(d), 4(d)). The sensitivity patterns have opposite sign in Fig. 3(d) vs. Fig. 4(d) because wind stress in this region creates pressure anomalies that are transported to the Icelandic coastline. Once there, the pressure anomalies drive a simultaneous strengthening (or weakening) of the NwAC and EGC, as described before, which results in opposite temperature responses in the Irminger vs. Norwegian Sea, similarly as in Fig. 5(d).

3.4 Assessing Shared Adjustment Mechanisms

This section quantifies the dynamical proxy potential (DPP) of each of the three temperature observations, θ^A, θ^B, and θ^C, for our QoI, heat transport across the Iceland-

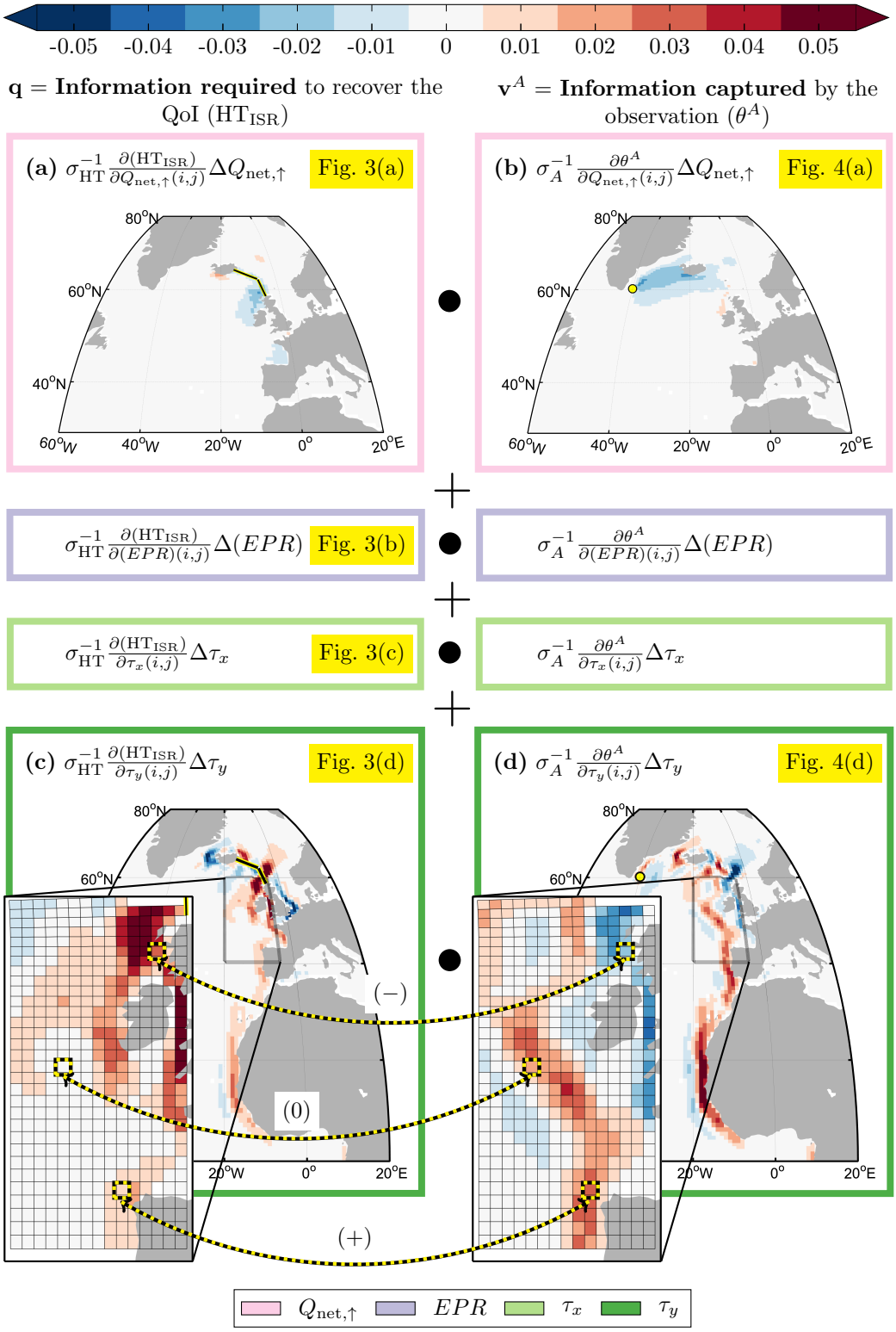

Figure 6. Projection (•) of weighted and normalized sensitivities (eq. (12)) of the QoI, $\mathrm{HT}_{\mathrm{ISR}}$ (**q**, left column), and the observed quantity, θ^A ($\mathbf{v}^A$, right column). All shown sensitivity maps are replots of subpanels in Figs. 3 and 4, as indicated by the yellow labels, and are composed of patterns that are established by the dynamical adjustment mechanisms of $\mathrm{HT}_{\mathrm{ISR}}$ and θ^A, respectively. Shared adjustment physics result in a strong projection (or 'pattern correlation'), elucidating the dynamical origins of proxy potential. The color shading in each of the shown model grid cells (inlets in (c),(d)) corresponds to an entry in either of the two sensitivity vectors, **q** and $\mathbf{v}^A$, associated with the forcing variable τ_y. The three cases (+), (−), and (0), resulting from the elementwise projection, are discussed in the text.

Scotland ridge ($\mathrm{HT_{ISR}}$). Quantification is via the pointwise projection of sensitivities:

$$\mathbf{q} \bullet \mathbf{v}^{\star} = \sum_{m=1}^{4} \sum_{i,j} \left(\sigma_{\mathrm{HT}}^{-1} \frac{\partial(\mathrm{HT_{ISR}})}{\partial F_m(i,j)} \Delta F_m \right) \cdot \left(\sigma_{\star}^{-1} \frac{\partial \theta^{\star}}{\partial F_m(i,j)} \Delta F_m \right), \tag{12}$$

for $\star = A, B, C$, cf. eqs. (4),(9),(11). Shared adjustment physics results in strong projections, elucidating the dynamical origins of proxy potential.

Fig. 6 shows the projection in eq. (12) for the case $\star = A$. Note that the projection can be regarded as a pattern correlation, where the patterns in Fig. 6 are established by dynamical adjustment mechanisms and pathways (sections 3.2, 3.3). We highlight three cases (Figs. 6(c),(d)) resulting from the pointwise projection in eq. (12):

(+) overlapping sensitivities of equal sign, resulting in a *positive* contribution to the projection $\mathbf{q} \bullet \mathbf{v}^{A}$;
(−) overlapping sensitivities of opposite sign, resulting in a *negative* contribution to $\mathbf{q} \bullet \mathbf{v}^{A}$;
(0) non-overlapping sensitivities, resulting in *no* contribution to $\mathbf{q} \bullet \mathbf{v}^{A}$.

Fig. 7(i) is a quantitative summary of Fig. 6, showing total positive (case (+)) and negative (case (−)) contributions to the projection $\mathbf{q} \bullet \mathbf{v}^{A}$, for the four different forcings. The maximum absolute value for the projection is equal to 1, due to normalization by $\sigma_{\mathrm{HT}}, \sigma_{\star}$ (eq. (12)).

For all observations considered, the generation of proxy potential is dominated by existence of wind-driven adjustments that are shared with those for $\mathrm{HT_{ISR}}$ (Figs. 7(i)-(k)). Minor importance of $Q_{\mathrm{net},\uparrow}$ and EPR is not surprising when recalling the fact that $\mathrm{HT_{ISR}}$ is relatively insensitive to $Q_{\mathrm{net},\uparrow}$ and EPR (Fig. 7(a)). Note that even for the surface temperature observation θ^{C}, which is highly sensitive to surface heat fluxes (Fig. 7(d)), the $Q_{\mathrm{net},\uparrow}$ contribution to the projection in eq. (12) is negligible (Fig. 7(k)).

Positive τ_y contributions to $\mathbf{q} \bullet \mathbf{v}^{A}$ arise along the eastern boundary of the subtropical North Atlantic (solid box, Fig. 7(f)), where both $\mathrm{HT_{ISR}}$ and θ^{A} exhibit a band of positive sensitivity along the western African and European coast (Figs. 7(e),(f)), due to the shared pressure adjustment mechanism discussed in sections 3.2, 3.3 and Figs. 5(a),(c). Negative τ_y contributions to $\mathbf{q} \bullet \mathbf{v}^{A}$ arise in the northeast Atlantic (dashed box, Fig. 7(f)), where wind stress sensitivities are of large amplitude and of opposite sign for $\mathrm{HT_{ISR}}$ (Fig. 7(e)) and θ^{A} (Fig. 7(f)), as discussed in section 3.3 and Fig. 5(d). The negative projection in the northeast Atlantic exceeds the positive projection in the eastern Atlantic waveguide (Fig. 7(i)). Total positive and negative contributions sum to $\mathbf{q} \bullet \mathbf{v}^{A} = -0.44$ (Fig. 7(i)). Here, partial cancellation between the positive and negative projections leads to a value reduced by $c^{A} = 0.29$. For $\star = A, B, C$, cancellation is quantified as

$$c^{\star} = |\mathbf{q}| \bullet |\mathbf{v}^{\star}| - |\mathbf{q} \bullet \mathbf{v}^{\star}| \geq 0, \tag{13}$$

where $|\mathbf{w}|$ denotes the vector whose entries are the absolute values of the respective entries of the vector $\mathbf{w}$, for $\mathbf{w} = \mathbf{q}, \mathbf{v}^{\star}$. As an example, if all contributions shown in Fig. 7(i) were either consistently positive or consistently negative (in which cases no cancellation occurred), the absolute value of the projection $\mathbf{q} \bullet \mathbf{v}^{A}$ would be increased by the addition of c^{A}, resulting in an absolute value of $0.44 + 0.29 = 0.73$.

Positive wind stress contributions to $\mathbf{q} \bullet \mathbf{v}^{B}$ (Fig. 7(j)) are of similar magnitude as positive wind stress contributions to $\mathbf{q} \bullet \mathbf{v}^{A}$ (Fig. 7(i)), due to the pressure adjustment mechanism in the eastern Atlantic waveguide, shared among θ^{B}, $\mathrm{HT_{ISR}}$ (and θ^{A}). The total overlap of $\mathrm{HT_{ISR}}$ sensitivity with θ^{B} sensitivity (Fig. 7(j)) is much smaller than with θ^{A} sensitivity (Fig. 7(i)), since θ^{B} does not show any sensitivity north of 55 °N (Fig. 7(g)). For the Irminger Sea surface observation θ^{C}, proxy origins are similar as for the Irminger

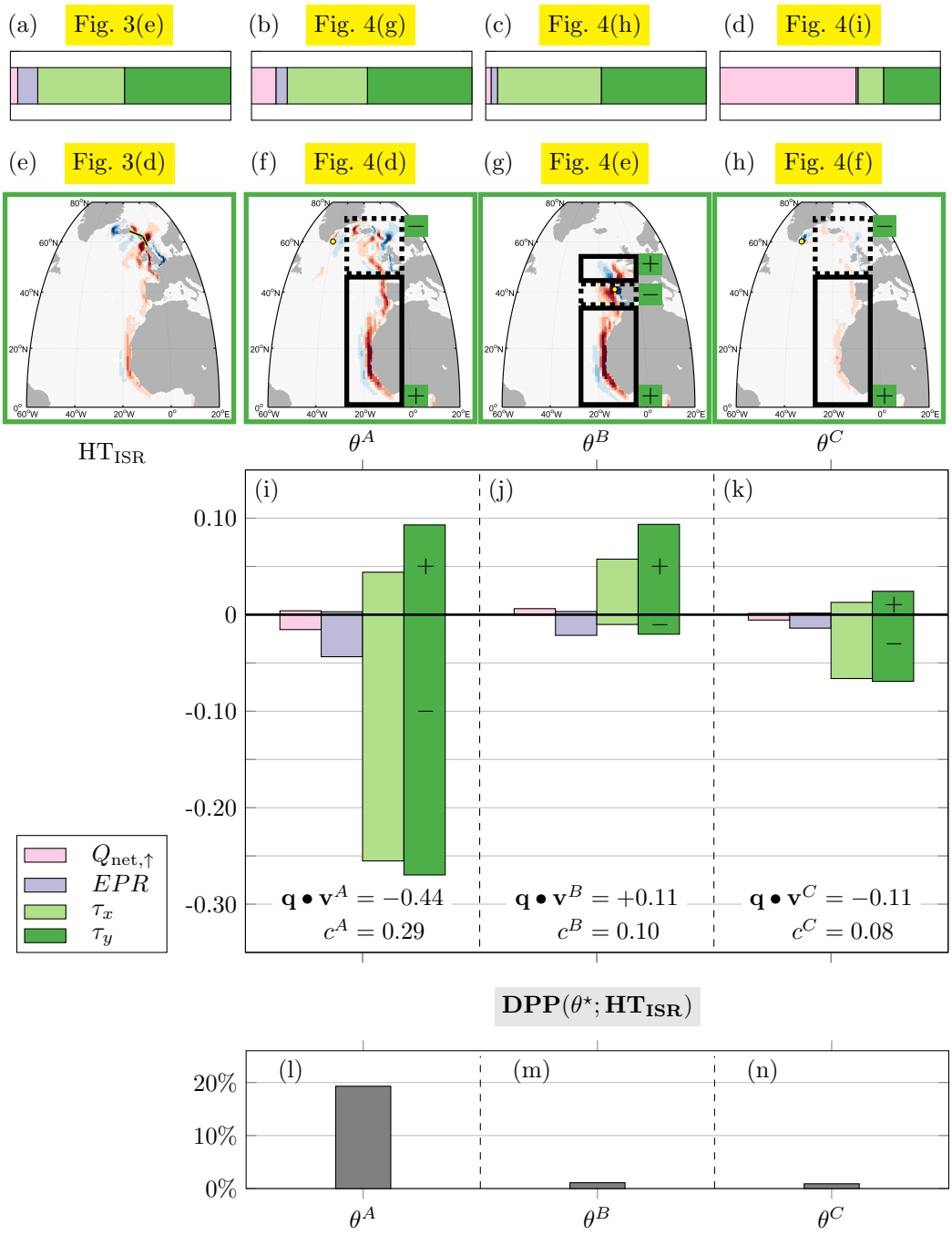

Figure 7. (a)-(h): Replots of subpanels in Figs. 3 and 4, as indicated by the yellow labels. (i)-(k): Contributions from the forcings $F = Q_{\mathrm{net},\uparrow}, EPR, \tau_x, \tau_y$ to the projection $\mathbf{q} \bullet \mathbf{v}^\star$ (eq. (12)), where (i) $\star = A$, (j) $\star = B$, and (k) $\star = C$. The projections are computed as shown in Fig. 6. That is, in each of the subpanels (i)-(k), the τ_y contribution is computed by projecting the sensitivity map $\sigma_{\mathrm{HT}}^{-1} \frac{\partial(\mathrm{HT}_{\mathrm{ISR}})}{\partial \tau_y(i,j)} \Delta\tau_y$ (shown in (e)) onto the respective sensitivity map $\sigma_\star^{-1} \frac{\partial \theta^\star}{\partial \tau_y(i,j)} \Delta\tau_y$ (shown in (f) for $\star = A$, (g) for $\star = B$, (h) for $\star = C$). Positive (negative) τ_y contributions, arise in the Atlantic subregion inside the black solid (dashed) box in (f) for $\star = A$, (g) for $\star = B$, (h) for $\star = C$, inside which sensitivities correlate (anticorrelate) with those in subpanel (e). The value for $\mathbf{q} \bullet \mathbf{v}^\star$ (bottom of subpanels (i)-(k)) is obtained by summing up all upward- and downward pointing bars in the respective subpanel. Here, destructive interference is quantified by $c^\star$ (eq. (13)). (l)-(n): Dynamical proxy potential of (l) θ^A, (m) θ^B, (n) θ^C for $\mathrm{HT}_{\mathrm{ISR}}$, computed by $\mathrm{DPP}(\theta^\star; \mathrm{HT}_{\mathrm{ISR}}) = (\mathbf{q} \bullet \mathbf{v}^\star)^2$ (see eq. (4)).

Sea subsurface observation θ^A (boxes in Figs. 7(f),(h)), but contributions from each forcing are reduced by a factor of about 0.25 (Figs. 7(i),(k)), due to relatively weaker excitation of surface temperature by remote forcing, as discussed in section 3.3.

The dynamical proxy potential (eq. (4)) of $\theta^\star$ for $\mathrm{HT_{ISR}}$ is given by $\mathrm{DPP}(\theta^\star;\mathrm{HT_{ISR}}) = (\mathbf{q} \bullet \mathbf{v}^\star)^2$, i.e., computed by taking the square of the values obtained in Figs. 7(i)-(k). The result is shown in Figs. 7(l)-(n): the DPP of θ^A for $\mathrm{HT_{ISR}}$ is 19%, while the DPPs of θ^B and θ^C for $\mathrm{HT_{ISR}}$ are only 1%. These values can be interpreted either in terms of (i) shared ocean adjustment physics or (ii) uncertainty quantification (cf. Fig. 1(c)). As for (i), $\mathrm{HT_{ISR}}$ shares 19% of its dynamical causes with θ^A, but only 1% with θ^B and θ^C. As a result, θ^A (θ^B, θ^C) captures 19% (1%, 1%) of the variability of $\mathrm{HT_{ISR}}$, taking into account all potential forcing scenarios. As for (ii), DPP predicts that uncertainty in $\mathrm{HT_{ISR}}$ would get reduced by 19% (1%, 1%), if a noise-free measurement value of θ^A (θ^B, θ^C) was added to the state estimation framework that was described in section 3.1.

4 Discussion

The design of effective climate observing systems relies on a both *physical* and *quantitative* understanding of which quantities of interest (QoIs) that capture important aspects of the climate system can be informed by existing or future observations. Toward this goal, we introduced the concept of dynamical proxy potential (DPP) by establishing a parallel between (i) ocean dynamical principles and (ii) Hessian-based uncertainty quantification (UQ). Hessian-based UQ is currently being explored in the computational sciences as a tool for optimal observing system design (Alexanderian et al., 2016; Bui-Thanh et al., 2012, 2013; Flath et al., 2011; Isaac et al., 2015), but has so far not been applied in the context of global ocean state estimation. With its two interpretations, DPP provides a means to optimally design climate observing systems, while giving insight into the governing physical mechanisms. In order to clarify the link between interpretations (i) and (ii), this work considered DPP of a single, noise-free observation. Forthcoming work will generalize this concept to full observing systems and account for observational noise (section 4.4).

In the following, we summarize the results from our North Atlantic case study (section 4.1), discuss our method in the context of related work in oceanography (section 4.2), point out limitations (section 4.3), and provide directions for future work (section 4.4).

4.1 Summary of Shared Pathways

In our case study, we chose heat transport across the Iceland-Scotland ridge ($\mathrm{HT_{ISR}}$) as our exemplary QoI. We explored the potential for three example observed quantities to serve as proxies for this QoI: two temperature observations, θ^A, θ^C, in the Irminger Sea, and one temperature observation, θ^B, off the Portuguese coast. Here, θ^A, θ^B were assumed subsurface, and θ^C at the sea surface. Examination of adjoint-derived sensitivities of QoI and observations revealed the following. On a five-year time scale, $\mathrm{HT_{ISR}}$ and θ^A are most sensitive to changes in wind forcing in two main regions: (I) along the eastern boundary of the subtropical North Atlantic and (II) in the northeast Atlantic and the Nordic Seas. Wind forcing in region (I) excites a pressure adjustment mechanism, which strengthens (or weakens) both the ISR geostrophic transport and the Irminger Current, leading to anomalies in $\mathrm{HT_{ISR}}$ and θ^A of equal sign. Wind forcing in region (II) drives an anomalous barotropic circulation around Iceland which simultaneously strengthens (or weakens) the Norwegian Atlantic and East Greenland Currents, leading to anomalies in θ^A and $\mathrm{HT_{ISR}}$ of opposite sign.

DPP is computed by projecting the sensitivities of the QoI and observation under consideration, by way of eq. (4). This projection measures the degree of shared adjustment physics, taking into account the effects of constructive and destructive interference

of information propagation. Destructive interference of information occurs because wind forcing in region (I) leads to responses in $\mathrm{HT_{ISR}}$ and θ^A of equal sign, while wind forcing in region (II) leads to responses in $\mathrm{HT_{ISR}}$ and θ^A of opposite sign. Considering the five-year mean of the two quantities, we find that the DPP of θ^A for $\mathrm{HT_{ISR}}$ is 19%. DPP allows two equivalent interpretations: θ^A (i) shares 19% of its adjustment physics with $\mathrm{HT_{ISR}}$; (ii) reduces the uncertainty in $\mathrm{HT_{ISR}}$ by 19%, if θ^A is added without noise to the ECCO state estimate.

4.2 Relation to Previous Work

Complete sensitivity information, enabled by the adjoint of an ocean GCM, is the cornerstone of quantifying DPP. In previous work, adjoint-derived sensitivity information has been proposed to support observing system design in a distinct fashion: Marotzke et al. (1999), Köhl and Stammer (2004), and Heimbach et al. (2011) suggest that regions in which a given QoI shows highest sensitivity to hydrographic state variables are to be prioritized when deploying new hydrographic observations. The philosophy of these studies is to discover direct cause and effect relationships between changes in observations and changes in the QoI. DPP looks further: by employing sensitivity information of not only the QoI, but also of the observations, DPP quantifies dynamics-based covariability of the QoI and observations, driven by local or remote forcings. As a result, DPP can exploit that multiple distinct QoIs may be connected by basin-wide dynamical adjustments and thus well constrained by limited instrumentation in this shared adjustment pathway. Moreover, unlike DPP, the adjoint-based studies referenced above do not provide a *quantitative* estimate on how well the QoI is constrained by the suggested observations (and how much information is missing). Instead, the previous studies provide only a relative estimate by suggesting that some observations may be more informative than others.

A similar description to DPP can be obtained through the method of representers (Bennett, 1985, 2002, see also our section 2). In the context of variational data assimilation, a representer assesses the impact of an assimilated observation on the estimated model state (e.g., Bennett, 1985, 1990; Kurapov et al., 2009). Representer-based methods have been used to evaluate observing systems and strategies in regional settings on short (daily to weekly) time scales (e.g., Moore et al., 2017; W. G. Zhang et al., 2010). DPP, as introduced in this work, views representers from a new perspective, with the merit to highlight the important role of dynamical pathways and adjustment processes in establishing and quantifying the information content of an observation, in relation to the information required to recover a QoI. Moreover, DPP operates - for the first time - within global ocean state estimation, focusing on climate observing systems, longer (monthly to multiannual) time scales, and large-scale ocean circulation.

In the context of Arctic observing system design, Kaminski et al. (2015, 2018) utilized a method related to DPP. A key difference is that the authors handle sensitivity information averaged over large regions (e.g., Fig. 2 in Kaminski et al., 2015). While spatially averaged sensitivity enables numerically efficient quantification of the constraint from large-scale data acquisitions (e.g., from satellite or aircraft), it could entail large aggregation errors (Kaminski et al., 2001). Furthermore, we argue that it inhibits clear understanding of proxy origins, by grouping information from dynamically distinct regions. This is especially true for harnessing proxy potential from shared wind-driven adjustments, for which resolution of the coastal waveguides is key.

4.3 Limitations

While the adjoint model provides comprehensive sensitivity information throughout the entire model space-time domain, adjoint-derived sensitivity is limited by the linear approximation and inexactness of the adjoint (Czeschel et al., 2010; Errico, 1997; Hoteit

et al., 2005). Nevertheless, perturbation experiments with the full nonlinear model dynamics in Appendix A show that the adjoint-derived sensitivities in our case study reliably capture the basin-wide adjustment mechanism that are excited by wind stress perturbations in regions (I) and (II). However, estimation errors in the predicted response amplitudes can reach up to 20% - partly due to the linear approximation, and partly due to the inexactness of the adjoint. This emphasizes that we must keep validating the accuracy of adjoint-derived sensitivity information and that improving the exactness of the adjoint would add great value to dynamics-based observing system design.

A second shortcoming of the methodology presented here is that DPP may be dependent on the underlying ocean GCM. Global ocean GCMs are typically too coarse (here: nominally 1° horizontal resolution) to accurately represent important processes including density-driven gravity currents (e.g., across overflow regions in the subpolar North Atlantic), deep convection, and narrow boundary currents. Inability to test model dependency, due to unavailable adjoints for almost all GCMs, is a limiting factor.

As a third drawback, important limiting assumptions entering the calculation of DPP are the choice of control variables and associated prior covariance (or weights). The control variables should include all uncertain elements in the model, i.e., the parts that are not determined by the known governing equations (Fig. 1(c)). The associated weights reflect prior uncertainties in the control variables. In the choice of uncertain control variables and weights, DPP follows the assumptions of ocean state estimation (while in our case study, we simplified controls and weights, for the sake of adding clarity to our presentation, see section 3.1.2). It is important to note that, while the sensitivities, utilized for the computation of DPP, uncover all dynamical adjustment processes (independent of the weights), the relative importance of these mechanisms is determined by the weights.

4.4 Future Directions

In our case study, we investigated how changes in time-mean forcing affect the time-means of temperature observations (e.g., θ^A) and QoI ($\mathrm{HT_{ISR}}$), on five-year time scales. Thereby, we identified key forcings, adjustment pathways, and mechanisms for observations and QoI, but did not disentangle whether the dominating mechanisms operate on weekly, monthly, seasonal, annual, or multiannual ($<$ five years) time scales. Future work should consider time-variable changes in forcing and investigate the variability of observations and QoIs on shorter (e.g., interannual) time scales. This will enable us to not only disentangle the dominant time scales of the operating mechanisms, but also to assess the predictive skill of observations for a QoI, if the QoI's response to certain forcings is lagged behind the observations' response. In view of the recently deployed OSNAP mooring array, it will be exciting to explore the DPP of the OSNAP observations for remote oceanic quantities that are not readily accessible to direct observation but have important climatic repercussions.

An important extension of the work presented here is to account for observational noise as well as data redundancy and complementarity between multiple observations. This objective will be pursued in a forthcoming paper. In our case study, destructive interference of competing adjustment mechanisms in regions (I) and (II) prevents DPP of θ^A for $\mathrm{HT_{ISR}}$ from exceeding 19% (section 4.1). Moreover, we found that θ^B is sensitive to wind forcing in region (I), but entirely insensitive to wind forcing in region (II). We will show that considering θ^A and θ^B in combination, will help to extract some of the information which is lost in destructive interference when viewing θ^A in isolation.

An interesting question is that of an optimal observing strategy, for instance: what is the value of surface (possibly remotely sensed) vs. subsurface (in situ) observations for QoIs that are inaccessible to direct observation? In our case study, we found that the surface temperature observation θ^C is too sensitive to local air-sea heat exchanges. The strong sensitivity to local processes overrides the sensitivity to the large-scale basin-wide

adjustment mechanisms that are relevant for HT_{ISR} - and potentially many other QoIs. The fact that the efficiency of observing systems depends on the targeted QoIs highlights the importance of an ongoing community discussion on which climate QoIs are most important to constrain.

5 Conclusions

The oceanographic community would strongly benefit from a synergistic, quantitative approach to co-design a cost-effective, long-term, and sustained ocean observing system (National Academies of Sciences, Engineering, and Medicine, 2017). Motivated by this objective, we have introduced the concept of dynamical proxy potential (DPP), a dynamics-based alternative to statistical correlation analysis and conventional observing system simulation experiments (OSSEs). Our main conclusion are the following.

(1) In contrast to statistical proxy potential, DPP only accounts for covariability that has a dynamical underpinning. It employs sensitivity information that is encapsulated in the equations of motion and that traces variability back to common causal forcings. The dominant regions and shared wind-driven adjustment mechanisms discussed in this study (section 4.1) have been shown to be key for many distinct oceanic quantities (see adjoint-based investigations referenced in the introduction), reflecting the ubiquitous nature of basin-scale adjustment processes (D. P. Marshall & Johnson, 2013). By identifying these *common dynamical pathways and mechanisms via DPP assessment*, we elucidate the physical cause of observed covariability in the North Atlantic.

(2) Our method accounts for all dynamically feasible pathways between observation and QoI, and with the potential for *constructive & destructive interference* of information propagation. Unraveling constructive and destructive contributions to DPP, as performed here, paves the way for extracting *complementary information* from observations. By targeting information that is complementary to existing observing systems, the notion of DPP can support the design of efficient and effective future observing systems.

(3) In order to evaluate DPP, one does not require actual observational data, since DPP investigates the dynamical relationships between observation and QoI in the model. DPP gains its full power for *observing system design* through the following fact. Independent of the measurement value (potentially taken by a *future* observing system), inclusion of the observation in the underlying state estimation framework will reduce uncertainty in the QoI by the pre-determined value of DPP. While ocean state estimation is a well-established method for optimally combining ocean observations with an ocean GCM (Stammer et al., 2016), characterizing and quantifying the way in which observations constrain climate-related QoIs in the ocean state estimate has remained unexplored. The framework introduced here provides the first clear and dynamical interpretation of uncertainty quantification (UQ) in global ocean state estimation.

(4) Based on endpoint geostrophy, moorings that are to be informative for cross-section transports would be located along the section itself. Here, we demonstrated that *remote hydrographic observations can provide strong constraints* on cross-section transports due to large-scale oceanic teleconnections. This result highlights the importance of further probing the dynamical constraints contained within the existing observational database.

Appendix A Linear Approximation and Inexactness of the Adjoint

The adjoint of an ocean GCM provides comprehensive sensitivity information which is the key ingredient of DPP (eq. (4)). Caveats are that (i) adjoint-derived sensitivities provide only a linear approximation, and that (ii) the adjoint may be inexact, e.g., due to artificially increased viscosity compared to the forward model, which is often a requirement to stabilize the adjoint (Hoteit et al., 2005, 2010; Forget et al., 2015). Here, we verify adjoint-derived sensitivities against perturbation experiments with the nonlinear forward model. Our control simulation covers the final five years of the ECCOv4r2 state estimate. We focus on meridional wind stress perturbations inside the two green regions in Fig. 5, along the western African coast and the western Icelandic coast.

For each of the two chosen regions, we perform two separate perturbation experiments, imposing meridional wind stress anomalies of $\Delta\tau_y = \pm 0.05\,\mathrm{N/m^2}$ inside the region, respectively. We maintain the wind stress perturbation over the full five-year period. For $J \in \{\theta^A, \theta^B, \mathrm{HT_{ISR}}\}$ (eqs. (6),(7)), we then compute the differences

$$\Delta^{\pm}_{\mathrm{fwd}} J := J^{\pm} - J^0. \tag{A1}$$

Here, J^0 denotes the quantity J in the control simulation, and J^+ and J^- the same quantity in the simulation with the positive and negative perturbation, respectively. $\Delta^+_{\mathrm{fwd}} J$ and $-\Delta^-_{\mathrm{fwd}} J$ are identical if J depends linearly on τ_y inside the chosen perturbation region. Even if a resemblance of $\Delta^+_{\mathrm{fwd}} J$ and $-\Delta^-_{\mathrm{fwd}} J$ suggests a linear response, the adjoint-derived anomalies can still deviate from the forward anomalies, due to inexactness of the adjoint. Therefore, we next compare the forward anomalies $\Delta^+_{\mathrm{fwd}} J$ and $-\Delta^-_{\mathrm{fwd}} J$ to the adjoint estimate

$$\Delta^+_{\mathrm{adj}} J := \frac{\partial J}{\partial \tau_y} \cdot \Delta\tau_y, \tag{A2}$$

where $\Delta\tau_y$ now denotes the positive meridional wind stress anomaly ($+0.05\,\mathrm{N/m^2}$ inside the chosen region).

In Fig. A1, we see notable deviations between $\Delta^+_{\mathrm{fwd}} J$ and $-\Delta^-_{\mathrm{fwd}} J$ in two cases: for the anomalies in $J = \theta^B$ in response to a τ_y perturbation along the western African coast (green solid vs. dashed horizontal lines, Fig. A1(a)), and for the anomaly in $J = \theta^A$ in response to a τ_y perturbation along the western Icelandic coast (orange solid vs. dashed horizontal lines, Fig. A1(c)). In both cases, the amplitudes of the time-evolving forward anomalies start to develop an offset after 1-2 years. Since anomalies in θ^A and θ^B are the time-integrated result of ocean transport anomalies, the offset tends to becomes larger over time. In contrast, $\Delta^+_{\mathrm{fwd}} \mathrm{HT_{ISR}}$ and $-\Delta^-_{\mathrm{fwd}} \mathrm{HT_{ISR}}$ coincide (Figs. A1(b),(d)), suggesting that $\mathrm{HT_{ISR}}$ is linear as a function of τ_y forcing in the tested perturbation regions. Note however that, despite the suggested linearity, $\Delta^+_{\mathrm{adj}} \mathrm{HT_{ISR}}$ slightly differs from the forward anomalies in Fig. A1(d), due to an inexact adjoint. A similar situation occurs in Fig. A1(a) for the response anomalies in θ^A. In all cases shown in Fig. A1, the adjoint estimate $\Delta^+_{\mathrm{adj}} J$ predicts the response anomalies in J with the correct sign. Moreover, predicted amplitudes are generally close to those of the forward anomalies, although, in few cases, they can be off by up to 20%.

Acknowledgments

N.L. and K.H.N. were supported by the ice2ice project funded by the European Research Council under the European Community Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement 610055. P.H. and H.R.P. received funding in part from NSF OCE grant #1924456 and the Estimating the Circulation and Climate of the Ocean (ECCO) project via a JPL/Caltech subcontract. The ECCOv4r2 model setup used in this work is available for download on Github (https://github.com/gaelforget/ECCOv4) as an instance of the MIT general circulation model (MITgcm, http://mitgcm.org/). Adjoint code was generated using the TAF software tool, created and maintained by FastOpt GmbH

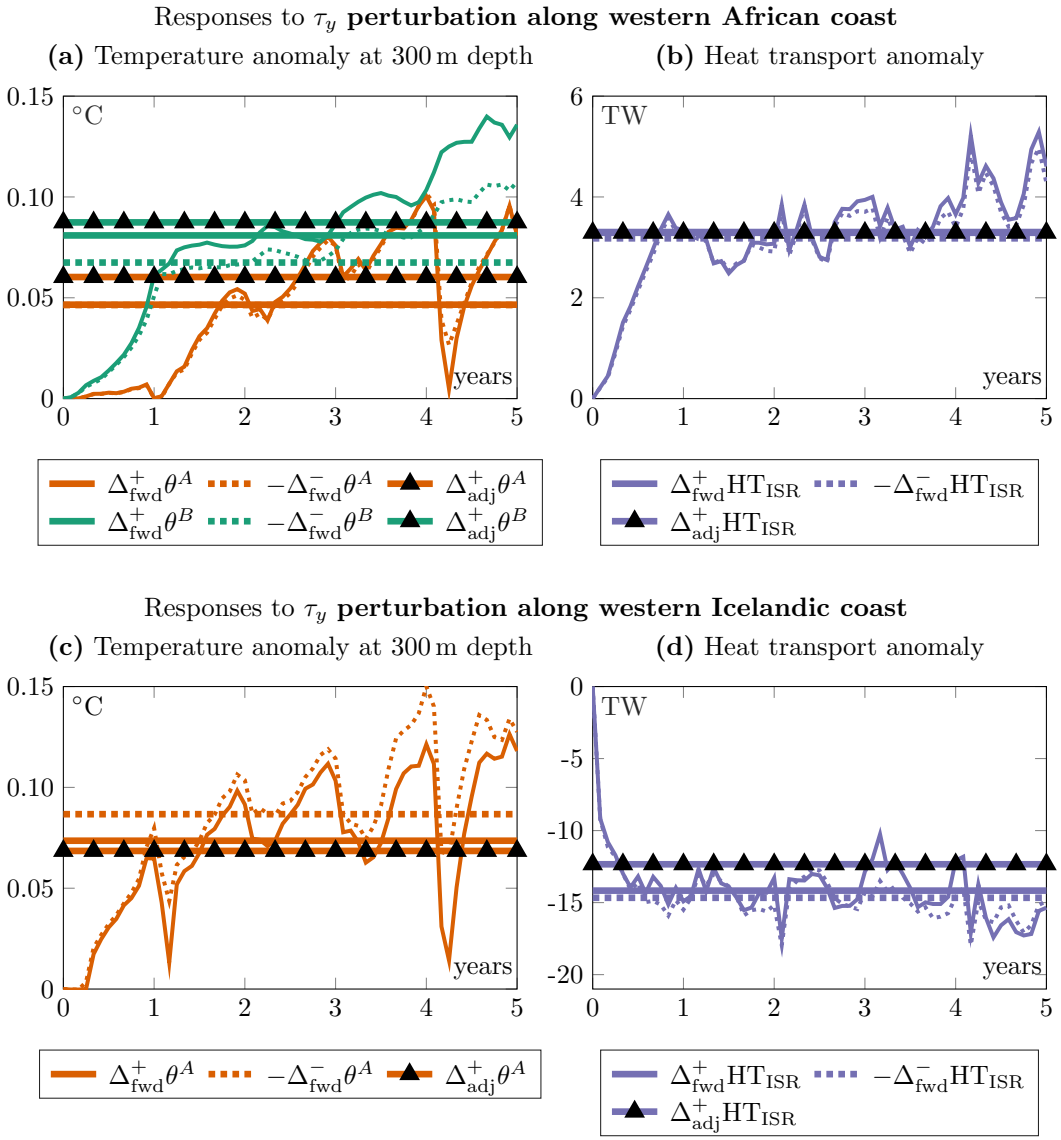

Figure A1. Anomalies in (a),(c) $J = \theta^A, \theta^B$, and (b),(d) $J = \text{HT}_{\text{ISR}}$, in response to meridional wind stress (τ_y) perturbations along the (a),(b) western African coast and (c),(d) western Icelandic coast. The solid vs. dashed, thick, horizontal lines show the five-year mean of the anomalies $\Delta^+_{\text{fwd}}J$ vs. $-\Delta^-_{\text{fwd}}J$ (eq. (A1)), diagnosed from the nonlinear forward perturbation experiments. The corresponding thin lines present the monthly evolution of $\Delta^+_{\text{fwd}}J$ vs. $-\Delta^-_{\text{fwd}}J$, as a function of years since the start of the perturbation. The thick, horizontal lines with black triangles show the adjoint-derived (five-year mean) anomalies $\Delta^+_{\text{adj}}J$ (eq. (A2)). In (c), the response anomaly in θ^B is invisible because it is two orders of magnitude smaller than the response anomaly in θ^A.

(http://www.fastopt.com/). We thank two anonymous reviewers for comments that greatly improved the manuscript. We wish to thank Céline Heuzé, Andrew Moore, Tim Smith, Carl Wunsch, and Daniel Jones for helpful discussions.

References

Alexanderian, A., Petra, N., Stadler, G., & Ghattas, O. (2016). A Fast and Scalable Method for A-Optimal Design of Experiments for Infinite-dimensional Bayesian Nonlinear Inverse Problems. *SIAM Journal on Scientific Computing*, *38*(1), A243–A272. doi: 10.1137/140992564

Alexander-Turner, R., Ortega, P., & Robson, J. I. (2018). How Robust Are the Surface Temperature Fingerprints of the Atlantic Overturning Meridional Circulation on Monthly Time Scales? *Geophysical Research Letters*, *45*(8), 3559–3567. doi: 10.1029/2017GL076759

Årthun, M., & Eldevik, T. (2016). On Anomalous Ocean Heat Transport toward the Arctic and Associated Climate Predictability. *Journal of Climate*, *29*(2), 689–704. doi: 10.1175/JCLI-D-15-0448.1

Asbjørnsen, H., Årthun, M., Skagseth, Ø., & Eldevik, T. (2019). Mechanisms of Ocean Heat Anomalies in the Norwegian Sea. *Journal of Geophysical Research: Oceans*, *124*(4), 2908–2923. doi: 10.1029/2018JC014649

Baehr, J., Haak, H., Alderson, S., Cunningham, S. A., Jungclaus, J. H., & Marotzke, J. (2007). Timely Detection of Changes in the Meridional Overturning Circulation at 26°N in the Atlantic. *Journal of Climate*, *20*(23), 5827–5841. doi: 10.1175/2007JCLI1686.1

Bennett, A. F. (1985). Array design by inverse methods. *Progress in Oceanography*, *15*(2), 129–156. doi: 10.1016/0079-6611(85)90033-3

Bennett, A. F. (1990). Inverse methods for assessing ship-of-opportunity networks and estimating circulation and winds from tropical expendable bathythermograph data. *Journal of Geophysical Research: Oceans*, *95*(C9), 16111–16148. doi: 10.1029/JC095iC09p16111

Bennett, A. F. (2002). *Inverse Modeling of the Ocean and Atmosphere*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511535895

Berx, B., Hansen, B., Østerhus, S., Larsen, K. M., Sherwin, T., & Jochumsen, K. (2013). Combining in situ measurements and altimetry to estimate volume, heat and salt transport variability through the Faroe-Shetland Channel. *Ocean Science*, *9*(4), 639–654. doi: 10.5194/os-9-639-2013

Bingham, R. J., & Hughes, C. W. (2009). Signature of the Atlantic meridional overturning circulation in sea level along the east coast of North America. *Geophysical Research Letters*, *36*(2). doi: 10.1029/2008GL036215

Buckley, M. W., Ponte, R. M., Forget, G., & Heimbach, P. (2014). Low-Frequency SST and Upper-Ocean Heat Content Variability in the North Atlantic. *Journal of Climate*, *27*(13), 4996–5018. doi: 10.1175/JCLI-D-13-00316.1

Bui-Thanh, T., Burstedde, C., Ghattas, O., Martin, J., Stadler, G., & 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* (pp. 1–11).

Bui-Thanh, T., Ghattas, O., Martin, J., & Stadler, G. (2013). A Computational Framework for Infinite-Dimensional Bayesian Inverse Problems Part I: The Linearized Case, with Application to Global Seismic Inversion. *SIAM Journal on Scientific Computing*, *35*(6), A2494–A2523. doi: 10.1137/12089586X

Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G., & Saba, V. (2018). Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature*, *556*(7700), 191–196. doi: 10.1038/s41586-018-0006-5

Chaudhuri, A. H., Ponte, R. M., Forget, G., & Heimbach, P. (2013). 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. doi: 10.1175/JCLI-D-12-00090.1

Czeschel, L., Marshall, D. P., & Johnson, H. L. (2010). Oscillatory sensitivity of Atlantic overturning to high-latitude forcing. *Geophysical Research Letters*, *37*(10), L10601. doi: 10.1029/2010GL043177

Errico, R. M. (1997). What Is an Adjoint Model? *Bulletin of the American Meteorological Society*, *78*(11), 2577–2591. doi: 10.1175/1520-0477(1997)078⟨2577:WIAAM⟩2.0.CO;2

Ezer, T. (2015). Detecting changes in the transport of the Gulf Stream and the Atlantic overturning circulation from coastal sea level data: The extreme decline in 2009-2010 and estimated variations for 1935-2012. *Global and Planetary Change*, *129*, 23–36. doi: 10.1016/j.gloplacha.2015.03.002

Flath, H., Wilcox, L., Akçelik, V., Hill, J., van Bloemen Waanders, B., & 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. doi: 10.1137/090780717

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

Frajka-Williams, E. (2015). Estimating the Atlantic overturning at 26°N using satellite altimetry and cable measurements. *Geophysical Research Letters*, *42*(9), 3458–3464. doi: 10.1002/2015GL063220

Frajka-Williams, E., Ansorge, I. J., Baehr, J., Bryden, H. L., Chidichimo, M. P., Cunningham, S. A., ... Wilson, C. (2019). Atlantic Meridional Overturning Circulation: Observed Transport and Variability. *Frontiers in Marine Science*, *6*. doi: 10.3389/fmars.2019.00260

Fu, L.-L., Lee, T., Liu, W. T., & Kwok, R. (2018, January). 50 Years of Satellite Remote Sensing of the Ocean. *Meteorological Monographs*, *59*, 5.1–5.46. (Publisher: American Meteorological Society) doi: 10.1175/AMSMONOGRAPHS-D-18-0010.1

Galanti, E., & Tziperman, E. (2003). A Midlatitude-ENSO Teleconnection Mechanism via Baroclinically Unstable Long Rossby Waves. *Journal of Physical Oceanography*, *33*(9), 1877–1888. doi: 10.1175/1520-0485(2003)033⟨1877:AMTMVB⟩2.0.CO;2

Galanti, E., Tziperman, E., Harrison, M., Rosati, A., Giering, R., & Sirkes, Z. (2002). The Equatorial Thermocline Outcropping–A Seasonal Control on the Tropical Pacific Ocean-Atmosphere Instability Strength. *Journal of Climate*, *15*(19), 2721–2739. doi: 10.1175/1520-0442(2002)015⟨2721:TETOAS⟩2.0.CO;2

Giering, R., & Kaminski, T. (1998). Recipes for Adjoint Code Construction. *ACM Trans. Math. Softw.*, *24*(4), 437–474. doi: 10.1145/293686.293695

Hansen, B., Larsen, K. M. H., Hátún, H., Kristiansen, R., & Mortensen, S., E. a Østerhus. (2015). Transport of volume, heat, and salt towards the Arctic in the Faroe Current 1993-2013. *Ocean Science*, *11*(5), 743–757. doi: 10.5194/os-11-743-2015

Hansen, B., & Østerhus, S. (2000). North Atlantic–Nordic Seas exchanges. *Progress in Oceanography*, *45*(2), 109–208. doi: 10.1016/S0079-6611(99)00052-X

Hansen, B., Østerhus, S., Hátún, H., Kristiansen, R., & Larsen, K. M. H. (2003). The Iceland-Faroe inflow of Atlantic water to the Nordic Seas. *Progress in Oceanography*, *59*(4), 443–474. doi: 10.1016/j.pocean.2003.10.003

Heimbach, P., Wunsch, C., Ponte, R. M., Forget, G., Hill, C., & 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. doi: 10.1016/j.dsr2.2010.10.065

Hoteit, I., Cornuelle, B., & Heimbach, P. (2010). An eddy-permitting, dynamically consistent adjoint-based assimilation system for the tropical Pacific: Hindcast experiments in 2000. *Journal of Geophysical Research: Oceans*, *115*(C3), C03001. doi: 10.1029/2009JC005437

Hoteit, I., Cornuelle, B., Köhl, A., & Stammer, D. (2005). Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3659–3682. doi: 10.1256/qj.05.97

Isaac, T., Petra, N., Stadler, G., & 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. doi: 10.1016/j.jcp.2015.04.047

Jones, D., Forget, G., Sinha, B., Josey, S., Boland, E., Meijers, A., & 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). doi: 10.1002/2018JC013774

Kaminski, T., Kauker, F., Eicken, H., & Karcher, M. (2015). Exploring the utility of quantitative network design in evaluating Arctic sea ice thickness sampling strategies. *The Cryosphere*, *9*(4), 1721–1733. doi: 10.5194/tc-9-1721-2015

Kaminski, T., Kauker, F., Toudal Pedersen, L., Voßbeck, M., Haak, H., Niederdrenk, L., ... 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. doi: https://doi.org/10.5194/tc-12-2569-2018

Kaminski, T., Rayner, P. J., Heimann, M., & Enting, I. G. (2001). On aggregation errors in atmospheric transport inversions. *Journal of Geophysical Research: Atmospheres*, *106*(D5), 4703–4715. doi: 10.1029/2000JD900581

Knight, J. R., Allan, R. J., Folland, C. K., Vellinga, M., & Mann, M. E. (2005). A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophysical Research Letters*, *32*(20). doi: 10.1029/2005GL024233

Köhl, A. (2005). Anomalies of Meridional Overturning: Mechanisms in the North Atlantic. *Journal of Physical Oceanography*, *35*(8), 1455–1472. doi: 10.1175/JPO2767.1

Köhl, A., & Stammer, D. (2004). Optimal Observations for Variational Data Assimilation. *Journal of Physical Oceanography*, *34*(3), 529–542. doi: 10.1175/2513.1

Kurapov, A. L., Egbert, G. D., Allen, J. S., & Miller, R. N. (2009). Representer-based analyses in the coastal upwelling system. *Dynamics of Atmospheres and Oceans*, *48*(1), 198–218. doi: 10.1016/j.dynatmoce.2008.09.002

Landerer, F. W., Wiese, D. N., Bentel, K., Boening, C., & Watkins, M. M. (2015). North Atlantic meridional overturning circulation variations from GRACE ocean bottom pressure anomalies. *Geophysical Research Letters*, *42*(19), 8114–8121. doi: 10.1002/2015GL065730

Latif, M., Roeckner, E., Botzet, M., Esch, M., Haak, H., Hagemann, S., ... Mitchell, J. (2004). Reconstructing, Monitoring, and Predicting Multidecadal-Scale Changes in the North Atlantic Thermohaline Circulation with Sea Surface Temperature. *Journal of Climate*, *17*(7), 1605–1614. doi: 10.1175/1520-0442(2004)017⟨1605:RMAPMC⟩2.0.CO;2

Lherminier, P., Mercier, H., Gourcuff, C., Alvarez, M., Bacon, S., & Kermabon, C. (2007). Transports across the 2002 Greenland-Portugal Ovide section and comparison with 1997. *Journal of Geophysical Research: Oceans*, *112*(C7). doi: 10.1029/2006JC003716

Little, C. M., Hu, A., Hughes, C. W., McCarthy, G. D., Piecuch, C. G., Ponte, R. M., & Thomas, M. D. (2019). The Relationship Between U.S. East Coast Sea Level and the Atlantic Meridional Overturning Circulation: A Review. *Journal of Geophysical Research: Oceans*, *124*(9), 6435–6458. doi: 10.1029/2019JC015152

Lopez, H., Goni, G., & Dong, S. (2017). A reconstructed South Atlantic Meridional Overturning Circulation time series since 1870. *Geophysical Research Letters*, *44*(7), 3309–3318. doi: 10.1002/2017GL073227

Lozier, M. S., Bacon, S., Bower, A. S., Cunningham, S. A., Femke de Jong, M., de Steur, L., ... 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. doi: 10.1175/BAMS-D-16-0057.1

Lozier, M. S., Li, F., Bacon, S., Bahr, F., Bower, A. S., Cunningham, S. A., ... Zhao, J. (2019). A sea change in our view of overturning in the subpolar North Atlantic. *Science*, *363*(6426), 516–521. doi: 10.1126/science.aau6592

Marotzke, J., Giering, R., Zhang, K. Q., Stammer, D., Hill, C., & Lee, T. (1999). Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity. *Journal of Geophysical Research: Oceans*, *104*(C12), 29529–29547. doi: 10.1029/1999JC900236

Marshall, D. P., & Johnson, H. L. (2013). Propagation of Meridional Circulation Anomalies along Western and Eastern Boundaries. *Journal of Physical Oceanography*, *43*(12), 2699–2717. doi: 10.1175/JPO-D-13-0134.1

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

Marshall, J., Hill, C., Perelman, L., & Adcroft, A. (1997). Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *Journal of Geophysical Research: Oceans*, *102*(C3), 5733–5752. doi: 10.1029/96JC02776

McCarthy, G. D., Brown, P. J., Flagg, C. N., Goni, G., Houpert, L., Hughes, C. W., ... Smeed, D. A. (2019). Sustainable observations of the AMOC: Methodology and Technology. *Reviews of Geophysics*, *58*(1), e2019RG000654. doi: 10.1029/2019RG000654

McCarthy, G. D., Haigh, I. D., Hirschi, J. J.-M., Grist, J. P., & Smeed, D. A. (2015). Ocean impact on decadal Atlantic climate variability revealed by sea-level observations. *Nature*, *521*(7553), 508–510. doi: 10.1038/nature14491

Mercier, H., Lherminier, P., Sarafanov, A., Gaillard, F., Daniault, N., Desbruyères, D., ... Thierry, V. (2015). Variability of the meridional overturning circulation at the Greenland-Portugal OVIDE section from 1993 to 2010. *Progress in Oceanography*, *132*, 250–261. doi: 10.1016/j.pocean.2013.11.001

Moore, A. M., Arango, H. G., & Edwards, C. A. (2017). Reduced-Rank Array Modes of the California Current Observing System. *Journal of Geophysical Research: Oceans*, *123*(1), 452–465. doi: 10.1002/2017JC013172

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. doi: 10.17226/24919

Orvik, K. A., & Skagseth, Ø. (2005). Heat flux variations in the eastern Norwegian Atlantic Current toward the Arctic from moored instruments, 1995-2005. *Geophysical Research Letters*, *32*(14). doi: 10.1029/2005GL023487

Østerhus, S., Turrell, W. R., Jónsson, S., & Hansen, B. (2005). Measured volume, heat, and salt fluxes from the Atlantic to the Arctic Mediterranean. *Geophysical Research Letters*, *32*(7). doi: 10.1029/2004GL022188

Østerhus, S., Woodgate, R., Valdimarsson, H., Turrell, B., Steur, L. d., Quadfasel, D., ... Berx, B. (2019). Arctic Mediterranean exchanges: a consistent vol-

ume budget and trends in transports from two decades of observations. *Ocean Science*, *15*(2), 379–399.

Pillar, H. R., Heimbach, P., Johnson, H. L., & Marshall, D. P. (2016). Dynamical Attribution of Recent Variability in Atlantic Overturning. *Journal of Climate*, *29*(9), 3339–3352. doi: 10.1175/JCLI-D-15-0727.1

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press.

Riser, S. C., Freeland, H. J., Roemmich, D., Wijffels, S., Troisi, A., Belbéoch, M., ... Jayne, S. R. (2016). Fifteen years of ocean observations with the global Argo array. *Nature Climate Change*, *6*(2), 145–153. doi: 10.1038/nclimate2872

Ritz, S. P., Stocker, T. F., Grimalt, J. O., Menviel, L., & Timmermann, A. (2013). Estimated strength of the Atlantic overturning circulation during the last deglaciation. *Nature Geoscience*, *6*(3), 208–212. doi: 10.1038/ngeo1723

Roberts, C. D., & Palmer, M. D. (2012). Detectability of changes to the Atlantic meridional overturning circulation in the Hadley Centre Climate Models. *Climate Dynamics*, *39*(9), 2533–2546. doi: 10.1007/s00382-012-1306-3

Schauer, U., & Beszczynska-Möller, A. (2009). Problems with estimation and interpretation of oceanic heat transport - conceptual remarks for the case of Fram Strait in the Arctic Ocean. *Ocean Science*, *5*(4), 487–494. doi: https://doi.org/10.5194/os-5-487-2009

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

Stammer, D., Balmaseda, M., Heimbach, P., Köhl, A., & Weaver, A. (2016). Ocean Data Assimilation in Support of Climate Applications: Status and Perspectives. *Annual Review of Marine Science*, *8*, 491–518. doi: 10.1146/annurev-marine-122414-034113

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. 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. doi: 10.1029/JC094iC05p06177

Thornalley, D. J. R., Oppo, D. W., Ortega, P., Robson, J. I., Brierley, C. M., Davis, R., ... Keigwin, L. D. (2018). Anomalously weak Labrador Sea convection and Atlantic overturning during the past 150 years. *Nature*, *556*(7700), 227–230. doi: 10.1038/s41586-018-0007-4

Vellinga, M., & Wood, R. A. (2004). Timely detection of anthropogenic change in the Atlantic meridional overturning circulation. *Geophysical Research Letters*, *31*(14). doi: 10.1029/2004GL020306

Verdy, A., Mazloff, M. R., Cornuelle, B. D., & 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. doi: 10.1175/JPO-D-13-018.1

von Storch, H., & Zwiers, F. W. (1999). *Statistical Analysis in Climate Research*. Cambridge University Press.

Weller, R. A., Baker, D. J., Glackin, M. M., Roberts, S. J., Schmitt, R. W., Twigg, E. S., & Vimont, D. J. (2019). The Challenge of Sustaining Ocean Observations. *Frontiers in Marine Science*, *6*. doi: 10.3389/fmars.2019.00105

Wunsch, C. (1996). *The Ocean Circulation Inverse Problem*. Cambridge: Cambridge University Press.

Zhang, R. (2007). Anticorrelated multidecadal variations between surface and subsurface tropical North Atlantic. *Geophysical Research Letters*, *34*(12). doi: 10.1029/2007GL030225

Zhang, R. (2008). Coherent surface-subsurface fingerprint of the Atlantic meridional

overturning circulation. *Geophysical Research Letters*, *35*(20), L20705. doi: 10.1029/2008GL035463

Zhang, W. G., Wilkin, J. L., & Levin, J. C. (2010). Towards an integrated observation and modeling system in the New York Bight using variational methods. Part II: Representer-based observing strategy evaluation. *Ocean Modelling*, *35*(3), 134–145. doi: 10.1016/j.ocemod.2010.06.006

Zhang, X., Prange, M., Merkel, U., & Schulz, M. (2015). Spatial fingerprint and magnitude of changes in the Atlantic meridional overturning circulation during marine isotope stage 3. *Geophysical Research Letters*, *42*(6), 2014GL063003. doi: 10.1002/2014GL063003

Supporting Information for "Quantifying Dynamical Proxy Potential through Shared Adjustment Physics in the North Atlantic"

N. Loose[1,2], P. Heimbach[1,3,4], H. R. Pillar[1], K. H. Nisancioglu[2,5]

[1]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX, USA

[2]Department of Earth Science, University of Bergen, and Bjerknes Centre for Climate Research, Bergen, Norway

[3]Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA

[4]Institute for Geophysics, The University of Texas at Austin, Austin, TX, USA

[5]Department of Geosciences, University of Oslo, and Centre for Earth Evolution and Dynamics, Oslo, Norway

Contents of this file

Additional Supporting Information (Files uploaded separately)

Text S1. Uncertainty Quantification in Ocean State Estimation. In ocean state estimation, one optimizes a vector of control variables, $\mathbf{x} = (x_1, \ldots, x_N)$, such as to minimize a least-squares cost function J (Tarantola, 2005; Wunsch, 1996). For the simple case of a single available observation, J takes the form

$$J(\mathbf{x}) = \underbrace{\frac{1}{2}\left(\frac{y - \mathrm{Obs}(\mathbf{x})}{\varepsilon}\right)^2}_{J_{\mathrm{misfit}}(\mathbf{x})} + \underbrace{\frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{B}\,(\mathbf{x} - \mathbf{x}_0)}_{J_{\mathrm{prior}}(\mathbf{x})}. \tag{S.1}$$

The first term in eq. (S.1), $J_{\text{misfit}}(\mathbf{x})$, measures the misfit between the observation y and the observation counterpart, $\text{Obs}(\mathbf{x})$, simulated by the model. The second term, $J_{\text{prior}}(\mathbf{x})$, penalizes deviations from a first-guess $\mathbf{x}_0$. Observational noise and prior uncertainties are assumed to be Gaussian, with distributions $\mathcal{N}(0, \varepsilon^2)$ and $\mathcal{N}(\mathbf{x}_0, \mathbf{B})$.

The solution of the inverse problem is the minimizer of the cost function (S.1), $\mathbf{x}_{\text{min}} = \min_{\mathbf{x}} J$. The posterior uncertainty in $\mathbf{x}_{\text{min}}$ can be approximated by the Gaussian covariance matrix (Bui-Thanh et al., 2012; Thacker, 1989),

$$\mathbf{P} = \Big(\underbrace{\varepsilon^{-2} \, (\nabla_{\mathbf{x}} \text{Obs}) \, (\nabla_{\mathbf{x}} \text{Obs})^T}_{=\mathbf{H}_{\text{misfit}}} + \mathbf{B}^{-1} \Big)^{-1}, \tag{S.2}$$

with $\nabla_{\mathbf{x}} \text{Obs} = [(\partial(\text{Obs})/\partial x_1)_{|\mathbf{x}_{\text{min}}}, \ldots, (\partial(\text{Obs})/\partial x_N)_{|\mathbf{x}_{\text{min}}}]^T$. The matrix $\mathbf{P}$ in eq. (S.2) is equal to $\mathbf{H}_J^{-1}$, the inverse of the linearized Hessian matrix of J at $\mathbf{x}_{\text{min}}$. The linearized Hessian $\mathbf{H}_J$, in turn, is the sum of two matrices: first, the misfit Hessian, $\mathbf{H}_{\text{misfit}}$, which as the linearized Hessian of the model-data misfit term J_{misfit} (eq. (S.1)) characterizes the observational constraints on the control variables; and second, $\mathbf{B}^{-1}$, which is the Hessian of the regularization term J_{prior} (eq. (S.1)).

By means of the matrix inversion lemma (e.g., Section 2.7.3 in Press et al., 2007), eq. (S.2) can be rewritten as

$$\mathbf{P} = \mathbf{B} - \left(\varepsilon^2 + \sigma^2_{\text{Obs}}\right)^{-1} (\mathbf{B} \, \nabla_{\mathbf{x}} \text{Obs}) \, (\mathbf{B} \, \nabla_{\mathbf{x}} \text{Obs})^T, \tag{S.3}$$

with $\sigma^2_{\text{Obs}} = (\nabla_{\mathbf{x}} \text{Obs})^T \, \mathbf{B} \, (\nabla_{\mathbf{x}} \text{Obs})$. Eq. (S.3) describes uncertainty reduction in all control variables $\mathbf{x}$, which is achieved by the uncertainty propagation via the first two black arrows in Fig. 1(c), from the pink box to the green box. Eq. (S.3) phrases the posterior

uncertainty $\mathbf{P}$ as the prior uncertainty $\mathbf{B}$, less any information obtained from the observation.

To assess uncertainty reduction in a QoI, the uncertainty propagation along the first two black arrows in Fig. 1(c) has to be followed by a subsequent uncertainty propagation along the last two black arrows, from the green box to the purple box. The subsequent propagation is achieved by projecting the prior and posterior error covariance matrices $\mathbf{B}$ and $\mathbf{P}$ onto the QoI, resulting in the prior variance $\sigma^2_{\text{QoI}} = (\nabla_{\mathbf{x}}\text{QoI})^T \mathbf{B} (\nabla_{\mathbf{x}}\text{QoI})$ and posterior variance $(\sigma^{\mathbf{P}}_{\text{QoI}})^2 = (\nabla_{\mathbf{x}}\text{QoI})^T \mathbf{P} (\nabla_{\mathbf{x}}\text{QoI})$. The relative uncertainty reduction is given by

$$\tilde{\Delta}\sigma^2_{\text{QoI}} := \frac{\sigma^2_{\text{QoI}} - (\sigma^{\mathbf{P}}_{\text{QoI}})^2}{\sigma^2_{\text{QoI}}} \in [0,1]. \tag{S.4}$$

Due to the observational information that is propagated through the model dynamics, $(\sigma^{\mathbf{P}}_{\text{QoI}})^2$ is smaller than σ^2_{QoI}, i.e., uncertainty gets reduced. $\tilde{\Delta}\sigma^2_{\text{QoI}} = 0$ represents the case $(\sigma^{\mathbf{P}}_{\text{QoI}})^2 = \sigma^2_{\text{QoI}}$, when the observation does not add any information for the QoI. The other extreme is $\tilde{\Delta}\sigma^2_{\text{QoI}} = 1$, which corresponds to $\sigma^{\mathbf{P}}_{\text{QoI}} = 0$, i.e., a perfectly constrained QoI by the observation. By means of identity (S.3), relative uncertainty reduction in eq. (S.4) can be re-written as

$$\tilde{\Delta}\sigma^2_{\text{QoI}} = \left(\sigma^2_{\text{QoI}} \cdot (\varepsilon^2 + \sigma^2_{\text{Obs}})\right)^{-1} \left(\mathbf{B}^{1/2}\nabla_{\mathbf{x}}\text{QoI} \bullet \mathbf{B}^{1/2}\nabla_{\mathbf{x}}\text{Obs}\right)^2, \tag{S.5}$$

where $\mathbf{B}^{1/2}$ is the square root of the matrix $\mathbf{B}$, and $\bullet$ denotes the dot product of two vectors in $\mathbb{R}^N$. In the limit of vanishing observational noise, $\varepsilon^2 \searrow 0$, relative uncertainty reduction (i.e., the expression in eq. (S.5)) converges to

$$\tilde{\Delta}\sigma^2_{\text{QoI}} \nearrow \Big(\underbrace{\left[\sigma^{-1}_{\text{QoI}} \cdot \mathbf{B}^{1/2}\nabla_{\mathbf{x}}\text{QoI}\right]}_{\mathbf{q}} \bullet \underbrace{\left[\sigma^{-1}_{\text{Obs}} \cdot \mathbf{B}^{1/2}\nabla_{\mathbf{x}}\text{Obs}\right]}_{\mathbf{v}} \Big)^2. \tag{S.6}$$

The limit in eq. (S.6) is equal to the definition of dynamical proxy potential (eq. (4)).

In the limit of vanishing observational noise, relative uncertainty reduction, $\tilde{\Delta}\sigma^2_{\mathrm{QoI}}$, is solely determined by the vectors $\mathbf{q}$ and $\mathbf{v}$ (eq. (S.6)). $\mathbf{q}$ is the direction of interest within the control space, i.e., the *information required* to recover the QoI. On the other hand, $\mathbf{v}$ is the eigenvector of the non-dimensionalized misfit Hessian, $\mathbf{B}^{1/2}\,\mathbf{H}_{\mathrm{misfit}}\,\mathbf{B}^{1/2}$ (cf. eq. (S.2)), and fully characterizes the *information captured* by the observation. (Note that in the case of only one observation, the misfit Hessian and the non-dimensionalized misfit Hessian are matrices of rank one.)

Movie S1. Time-evolving monthly mean anomaly in North Atlantic bottom pressure (normalized by density, p/ρ) in response to a positive northward wind stress anomaly of amplitude $0.05\,\mathrm{N/m^2}$ along the western African coast. The final five years of the ECCOv4r2 solution serve as our control simulation. The wind stress perturbation is imposed inside the green contour, and maintained over the full five-year time period. As for the time label, $t = 0$ corresponds to the (simultaneous) start of the simulation and perturbation. The black lines mark the contours of the 500 m, 1000 m, 1500 m, 2000 m, 3500 m, 4000 m, 4500 m, and 5000 m isobaths. Time averaging of the monthly mean anomalies shown in this movie (over the full five-year time period) generates Fig. 5(a).

Movie S2. As Movie S1, but anomaly in the barotropic stream function. Negative values indicate anomalous counterclockwise rotation.

Movie S3. As Movie S1, but anomaly in North Atlantic potential temperature at 300 m depth. The black line marks the 300 m depth contour. Time averaging of the monthly mean anomalies shown in this movie generates Fig. 5(c).

Movie S4. As Movie S3, but anomaly in potential temperature along the trans-Atlantic section at 41 °N, in longitude-depth space. The green circle marks the location of θ^B. Baroclinic Rossby waves propagate a positive temperature anomaly from the eastern boundary westward.

Movie S5. As Movie S4, but anomaly in vertical velocity along the trans-Atlantic section at 15 °N. The green bar marks the longitudinal extent (20 °W - 17 °W) of the imposed northward wind stress anomaly (whose latitudinal extent is 10 °N - 20 °N). Ekman convergence/pumping occurs to the east of the wind stress anomaly, and Ekman divergence/suction to the west of the wind stress anomaly. Maintenance of the wind stress anomaly locks this vertical velocity dipole in place.

Movie S6. As Movie S5, but anomaly in potential temperature. Ekman pumping to the east of the wind stress anomaly (Movie S5) creates warming, while Ekman suction to the west (Movie S5) creates cooling. The negative thickness anomaly associated with shoaling of the thermocline (and cooling) to the west is propagated into the interior by baroclinic Rossby waves. Note that the positive thickness anomaly associated with deepening of the thermocline (and warming) to the east propagates northward along the North Atlantic boundary and then into the interior, once it is past the latitude band in which the wind

anomaly is applied (Movies S3 and S4).

Movie S7. As Movie S1, but anomaly in response to a positive northward wind stress anomaly along the western Icelandic coast, imposed inside the green contour. Time averaging of the monthly mean anomalies shown in this movie generates Fig. 5(b).

Movie S8. As Movie S2, but anomaly in response to a positive northward wind stress anomaly along the western Icelandic coast, imposed inside the green contour.

Movie S9. As Movie S3, but anomaly in response to a positive northward wind stress anomaly along the western Icelandic coast, imposed inside the green contour. Time averaging of the monthly mean anomalies shown in this movie generates Fig. 5(d).

References

Bui-Thanh, T., Burstedde, C., Ghattas, O., Martin, J., Stadler, G., & 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* (pp. 1–11).

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. 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. doi:

10.1029/JC094iC05p06177

Wunsch, C. (1996). *The Ocean Circulation Inverse Problem*. Cambridge: Cambridge University Press.