The Ocean’s Role in Air-Sea Interaction and Climate Predictability

Jacob Cohen

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2025

Reading Committee:
LuAnne Thompson, Chair
Alison Gray
Wei Cheng

Program Authorized to Offer Degree:
School of Oceanography

University of Washington

Abstract

The Ocean's Role in Air-Sea Interaction and Climate Predictability

Jacob Cohen

Chair of the Supervisory Committee:
LuAnne Thompson
School of Oceanography

Ocean dynamics drive the memory of the climate system. Long-term predictions, therefore, rely on an understanding of the ocean processes that control climate variability, and the usefulness of these predictions requires an understanding of current models' predictive skill. Across three chapters, we study the influence of upper ocean dynamics influence on sea surface temperature variability, predictions of spatially coherent marine heatwave events, and predictions of seasonal-to-decadal ocean heat transport and sea ice.

We first study the turbulent surface heat flux (Q) response to sea surface temperature (SST) anomalies and to mixed-layer heat content (HC) anomalies to understand where and when ocean processes control SST variability and air-sea interaction. From observational data of SST, HC, and Q, we use lagged covariances to define the feedback of SST to Q and the feedback of HC to Q. The feedback sensitivity, defined as the difference between SST-Q feedback and HC-Q feedback, illustrates the relative importance of ocean processes to atmospheric processes in controlling SST and Q variability. The regional and seasonal patterns of the sensitivity feedback demonstrate the varying pathways by which the ocean influences the surface heat flux feedback. Determining these patterns of ocean-dominated variability improves predictive understanding of the climate.

We then evaluate how well climate models detect and predict spatially connected extreme SST events known as marine heatwaves (MHWs). Here, we evaluate a method of detecting and predicting spatially connected MHW objects. We apply object-based forecast verification to the CESM2 Seasonal-to-Multiyear Large Ensemble (SMYLE) experiment, a set of initialized hindcasts with 20-member ensembles of 24-month simulations initialized quarterly from 1970–2019. We demonstrate that SMYLE predicts MHWs that occur near observed MHWs with high skill at long lead times, but with errors in location, area, and intensity that grow with lead time. SMYLE exhibits improved skill in predicting the intensity of MHWs in December and January, and worse skill from August to October. This work illustrates the capacity to forecast connected MHW objects and to quantify the uncertainty in those forecasts with potential applications for future community use.

The final chapter examines the prediction skill of Arctic sea ice and ocean heat transport (OHT) through the Pacific and Atlantic regions. Using the SMYLE dataset and a decadal prediction hindcast dataset (DP), we interrogate how the prediction skill of sea ice is related to the prediction skill of OHT across regions and timescales. We examine the co-occurrence of high OHT and sea ice prediction skill and evaluate the skill of DP in predicting the short-term tendency of the Arctic climate. We find high seasonal-to-decadal prediction skill and a strong relationship between the predictability of OHT and sea ice. This chapter demonstrates how ocean variability influences Arctic sea ice predictability. Together, the chapters in this dissertation contribute to our understanding of the ocean's role in climate variability and predictability.

Table of Contents

Introduction

Ocean dynamics drive the memory of the climate system. Long-term predictions, therefore, rely on an understanding of the ocean processes that control climate variability, and the usefulness of these predictions requires an understanding of current models' predictive skill. Because the ocean's heat capacity is orders of magnitude larger than the atmosphere's, the ocean can serve as a source of long-term predictability (Shi et al., 2022). As a result, determining the role of ocean dynamics in climate variability can lead to improvements in our predictive capabilities.

Interactions between the ocean and the atmosphere influence sea surface temperature (SST) variability as well as weather and climate, from storms and hurricanes to the El Niño-Southern Oscillation (ENSO) (Cronin et al., 2019). In addition to affecting SST variability, oceanic and atmospheric processes drive extreme SST events: marine heatwaves (MHWs). The most widely used definition for MHWs identifies MHWs as "a discrete prolonged anomalously warm water event in a particular location," quantifying events as SST anomalies relative to a long-term climatology which exceed a the local 90th percentile threshold and which last for at least five days with no gaps exceeding two days (Hobday et al., 2016). For monthly data, SST anomalies above a 90th percentile threshold for any duration of at least one month are considered MHWs (Capotondi et al., 2024; Scannell et al., 2020). MHWs affect ocean biogeochemistry (Burger et al., 2022; Kohlman et al., 2025; Le Grix et al., 2022), leading to negative consequences for marine ecosystems (Smale et al., 2019; Smith et al., 2023).

Better understanding the ocean's role in these interactions will improve our understanding and predictions of these events and phenomena. In Chapter 1, I explore the ocean's role in driving air-sea interaction and the exchange of heat between the ocean and the atmosphere. The ocean and the atmosphere drive variability and predictability in each other across timescales (Bach et al., 2019), so understanding how upper ocean processes interact with the atmosphere and control the upper ocean heat budget will improve our modeling and prediction capabilities.

Understanding the physical processes that contribute to accurate climate predictions allows us to trust those predictions. Sources of predictability on various timescales exist across the climate system, from soil moisture to decadal climate modes (Conil et al., 2009; Y. Liu et al., 2023; Richter et al., 2024). On seasonal to interannual timescales, the ocean is an important source of climate predictability due to its long-term memory (De Coëtlogon & Frankignoul, 2003; Deser et al., 2003; Shi et al., 2022). The ocean and the atmosphere interact through a variety of processes, from local, small-scale mixing to remote, large-scale atmospheric teleconnections (Seo et al., 2023). The ocean also interacts with sea ice on similar scales, influencing both variability and predictability (Bushuk et al., 2022; Y. Li et al., 2024). Large-scale modes of climate variability linked to air-sea interaction can further influence Arctic sea ice variability on annual to multi-decadal timescales. Multidecadal modes of variability can have a strong influence on sea ice (He et al., 2024), and tropical variability is linked to Arctic sea ice via teleconnections from El Niño and the Madden-Julien Oscillation (MJO) (Clancy et al., 2021; Yuan et al., 2018). Understanding sources of predictability leads to developments and improvements of future climate projections. Process-level knowledge also makes predictions interpretable and trustworthy (Spillman et al., 2025). When we know why predictions are skillful, we can better communicate forecasts and potential impacts to the user community.

In chapters 2 and 3, I assess predictive skill and sources of predictability by analyzing initialized predictions from a hindcast prediction system. A hindcast system is a set of forecasts (sometimes referred to as reforecasts) that are initialized regularly from past climate states. The Seasonal-to-Multiyear Large Ensemble (SMYLE) is a set of hindcast simulations run using the Community Earth Systems Model (CESM2) with a nominal 1° horizontal resolution for each component model (Yeager, 2022; Yeager et al., 2022). I also use the CESM2 Decadal Prediction (DP) system, which extends the November initializations from SMYLE into 10-year predictions.

I assess prediction skill from initialized climate models by comparing the forecast variable of interest to verification data. The verification data can come either from observations, which tells us about how well our forecasts predict real-world conditions (actual skill), or from the model output that forms the initial conditions, which tells us about how well our forecasts predict the model evolution forced by real conditions (potential skill). In this dissertation I evaluate both potential and actual skill. The initial conditions and observations available to compare against predictions are of first order importance in creating skillful climate forecasts (Edward Blanchard-Wrigglesworth, Bitz, et al., 2011; Conil et al., 2009). Consequently, improving our capabilities to collect accurate and high-resolution observations will improve our forecasting capabilities.

Across three studies, I examine how upper ocean dynamics influence air-sea interaction, how object-based predictions of extreme sea surface temperatures compare to observed events, and how Arctic sea ice and ocean heat transport are predicted on seasonal-to-decadal timescales. Each chapter within this dissertation addresses the overarching theme of the ocean’s role in climate variability and predictability. The proposed research aims to contribute to our understanding of two fundamental questions in climate research:

1. How do upper ocean processes influence air-sea interaction?
2. How predictable is the ocean in initialized climate models and why?

Chapter 1 primarily addresses the first question by examining the relationships between upper ocean dynamics, sea surface temperature, and surface heat fluxes, and chapters 2 and 3 address the second question through analysis of marine heatwaves, ocean heat transport, and Arctic sea ice in SMYLE and CESM2-DP. Each chapter and its specific sub-questions are defined here:

Chapter 1: The Role of Ocean Processes in Driving the Surface Heat Flux Feedback

A. How does the surface heat flux feedback depend on location and season?
B. In which regions and in which seasons do upper ocean processes most strongly influence the surface heat flux feedback?

Chapter 2: Object-Based Evaluation of Seasonal-to-Multiyear Marine Heatwave Predictions

A. How accurate are predictions of spatially connected marine heatwave events?
B. What are the uncertainties associated with predicted marine heatwave location, area, and intensity?

Chapter 3: Seasonal-to-Decadal Prediction of Regional Arctic Sea Ice and Ocean Heat Transport

A. How accurate are predictions of seasonal anomalies and decadal trends of regional Arctic sea ice extent and ocean heat transport?
B. How do ocean–sea-ice interactions influence Arctic predictability?

1. The Role of Ocean Processes in Driving the Surface Heat Flux Feedback

This chapter is reproduced from a manuscript submitted as Cohen, J. T. & Thompson, L. (2025). The role of ocean processes in driving the surface heat flux feedback. Journal of Geophysical Research: Oceans.

1.1. Introduction

Turbulent exchange of heat at the surface of the ocean through latent and sensible heat flux varies significantly by location and season. Turbulent heat fluxes (Q) depend directly on the sea surface temperature (SST); however, the evolution of SST and upper ocean heat content is determined by both surface fluxes as well as oceanic processes. Quantifying whether oceanic processes (such as advection and mixing within the ocean) or atmospheric processes (such as changes in wind stress and air temperature/humidity) dominate SST and Q variability relies on observations or model output of drivers of atmospheric and oceanic variability. The relative importance of variations in the ocean and atmosphere in air-sea heat exchange can be directly calculated in models using saved heat budgets, but in situ ocean observations are too sparse to reliably calculate the role of specific processes in driving air-sea interaction.

A stochastic model of air-sea interaction was introduced by Frankignoul and Hasselman (1977) that allows a quantification of the strength of the coupling between the ocean and the atmosphere. Frankignoul et al. (1998) used the stochastic framework to show that if atmospheric variability is stochastic, then an air-sea feedback parameter can be defined that represents the air-sea coupling through turbulent heat flux (sensible plus latent heat flux), where the atmosphere drives changes in the turbulent heat fluxes which in turn damp SST anomalies. This process acts on time scales of months to years (Frankignoul & Hasselmann, 1977).

Barsugli and Battisti (1998) introduced a coupled stochastic model of an atmospheric boundary layer and an ocean mixed layer. They added a parameter that reduced the strength of the coupling and, as a result, the SST response and the heat flux feedback revert to those in the Hasselman model (Frankignoul et al., 1998). These simple models are to first order valid on monthly time scales in the midlatitudes and in regions where ocean currents are weak. This paradigm, however, does not hold for all timescales and regions.

The relative sizes of the oceanic and the atmospheric processes controlling upper ocean heat content variations and air-sea interaction depend on both spatial and temporal scales (Bishop et al., 2017; Hausmann et al., 2017; Laurindo et al., 2022; Roberts et al., 2017; Small et al., 2008, 2019, 2020; X. Sun & Wu, 2021). Outside of the tropics, surface heat flux (Q) drives monthly and seasonal mixed layer temperature variability, but ocean heat flux convergence drives interannual variability (Roberts et al., 2017). Ocean heat transport can dominate on monthly timescales as well as interannual timescales, and on space scales less than 500-700 km (Bishop et al., 2017; Small et al., 2019).

In regions of strong currents, high SST and turbulent heat flux variability, and deep mixed layer depths such as in the Gulf Stream, the Kuroshio Extension, and the Antarctic Circumpolar Current (ACC) (Fig. 1.1), ocean heat flux convergence from processes such as advection, turbulent mixing, and Ekman transport can drive upper ocean heat content variability as well as turbulent heat flux variability (Bishop et al., 2017; Roberts et al., 2017; Small et al., 2019, 2020; Wu et al., 2006). In high-resolution coupled climate models, which resolve small spatial scales such as eddies and ocean fronts, positive local correlations between SST anomalies and surface turbulent heat flux indicate that large ocean heat transport convergence plays a dominant role in these regions (Kirtman et al., 2012). These high-resolution model results agree with observational results (X. Sun & Wu, 2021). Low-resolution climate models, however, fail to capture the strength of ocean heat transport and thus turbulent heat flux variability where currents are strong (Small et al., 2019). Mesoscale and submesoscale ocean processes can also drive air-sea interactions through eddies and frontogenesis (Jing et al., 2020; Moreton et al., 2021; Song et al., 2022; Strobach et al., 2022).

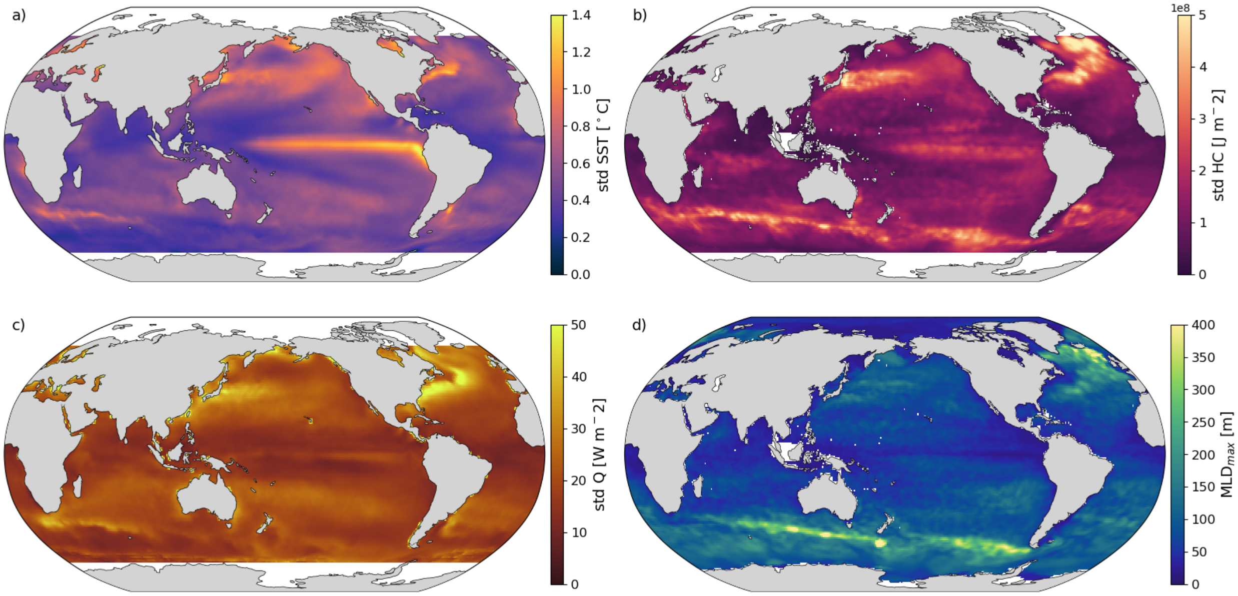


Figure 1.1. Standard deviation of SST (OISST; a), HC (RFROM; b), and Q (OAFLUX; c) anomalies, and $\mathrm{MLD}_{\mathrm{max}}$ (MIMOC; d). The details of each data product can be found in Sec. 1.2.3.

Here we use a simplified anomalous heat budget (Frankignoul et al., 1998; Frankignoul & Hasselmann, 1977) to examine the ocean's role in driving SST and Q variability. The stochastic heat budget describes the time evolution of SST anomalies as the result of stochastic forcing (F') and a linear feedback to SST anomalies T' ($\lambda T'$) due to turbulent fluxes:

$$\frac{\partial T'}{\partial t} = F' - \lambda T'. \tag{1.1}$$

In this framework, the atmosphere drives sea surface temperature (SST) variability that is then damped by turbulent surface heat fluxes. The rate at which the SST anomalies are damped by turbulent fluxes is given by the surface heat flux feedback λ.

This formulation of the anomalous heat budget can be decomposed to separate atmospheric and oceanic components of the forcing and damping (Patrizio & Thompson, 2022b):

$$\frac{\partial T'}{\partial t} = F_a' + F_o' - (\lambda_a + \lambda_o)T'. \tag{1.2}$$

Here, the atmospheric forcing F_a' refers to heat flux driven by intrinsic atmospheric variability while λ_a represents damping via the turbulent heat flux feedback. The ocean forcing F_o' is attributed to dynamical processes such as wind-driven Ekman advection, while processes that cause ocean damping (λ_o) include radiative damping (Laurindo et al., 2022) and advection by strong currents and entrainment (Patrizio & Thompson, 2022b; Siqueira et al., 2024).

Recent studies have examined the roles of oceanic and atmospheric processes within the framework of these stochastic models of air-sea interaction (e.g., Gunnarson et al., 2024; Liu et al., 2023; Patrizio & Thompson, 2022). By examining the damping of SST variability by surface heat flux feedback and by ocean damping separately, Patrizio and Thompson (2021) disentangled the processes that contribute to observed SST variability. While these studies directly quantify the ocean's influence on SST variability, here we consider the ocean's influence on the strength of atmospheric damping as given by the surface heat flux feedback (λ_a).

1.2. Methods and Data

1.2.1. The Turbulent Heat Flux Feedback

The rate at which turbulent heat fluxes cool the surface ocean is defined by the turbulent heat flux feedback. Here, we follow the method from Frankignoul and Kestenare (2002), which has been used to quantify the turbulent heat flux feedback across models and observations (Hausmann et al., 2016; S. Park et al., 2005; X. Sun & Wu, 2021). We calculate the heat flux feedback for SST anomalies (as in Frankignoul and Kestenare, 2002) and for mixed layer heat content (HC) anomalies. HC is given by $\rho C_p H T$ where ρ is the density of seawater, C_p is the specific heat capacity of seawater, H is the climatological mixed layer depth, and T is the mixed layer temperature. We refer to these two air-sea feedbacks as the SST-Q feedback (λ_{SST}) and the HC-Q feedback (λ_{HC}):

$$\lambda_{SST} = \frac{R_{Q'T'}(\tau)}{R_{T'T'}(\tau)}, \tag{1.3}$$

and

$$\lambda_{HC} = \rho c_p \mathrm{MLD}_{\mathrm{max}} \frac{R_{Q'HC'}(\tau)}{R_{HC'HC'}(\tau)}, \tag{1.4}$$

where $R_{Q'T'}$ ($R_{Q'HC'}$) is the cross-covariance at lag τ between SST (HC) anomalies and Q anomalies and $R_{T'T'}$ ($R_{HC'HC'}$) is the auto-covariance of SST (HC) at lag τ. The SST-Q feedback has units of watts per meter squared per degree Celsius, representing the magnitude of turbulent heat flux that results from a 1°C SST anomaly. In Eq. 1.4, we multiply the HC-Q cross-covariance by $\rho C_p H$ where $H = \mathrm{MLD}_{\max}$ such that λ_{HC} has the same units as λ_{SST} and corresponds to a 1°C temperature anomaly within the integrated HC.

We calculate the SST-Q feedback and HC-Q feedback first for year-round fields and then separately evaluate boreal winter and boreal summer fields. For the year-round calculation, the feedback is calculated as the mean feedback over $\tau = -1, -2, -3$ (SST or HC leading Q by one, two, and three months as in Park et al., 2005). Averaging over three lags removes the effect of atmospheric persistence (Frankignoul & Kestenare, 2002; S. Park et al., 2005). In this calculation, we average over the values only at lags where the SST (HC) lagged autocorrelation differs from zero at the 95% confidence level. If the autocorrelation is insignificant at all lags, the SST (HC) anomalies act as a white noise signal and the feedback cannot be estimated. In this case, we set the feedback value to zero. We consider the feedback to be statistically significant if the correlation between SST or HC and Q differs from zero at the 95% confidence level.

For the seasonal and monthly feedback, we calculate the covariance at $\tau = -1$, where SST and HC lead Q by one month. To calculate seasonal covariances, we define winter as December to March for Q and as November to February for SST and HC, and summer as June to September for Q and as May to August for SST and HC. To calculate monthly covariances, the target month and the following month are used for Q and the target month and preceding month are used for SST and HC.

We then calculate the feedback separately for months with positive and for months with negative SST anomalies. We first mask each field to only include locations and times where SST anomalies are positive and again where SST anomalies are negative. We then calculate covariances at these locations using Q at the target month and the following month and using SST and HC at the target month and the preceding month.

1.2.2. Quantifying the Ocean's Influence on λ

While both λ_{SST} and λ_{HC} describe the atmospheric response to changes in the ocean, further analysis of the stochastic model heat budget demonstrates how these quantities are related to each other. The SST-Q feedback describes the atmospheric response to the ocean surface, only accounting for the processes that directly affect SST variability at a given time. The HC-Q feedback, however, integrates all ocean processes that occur within the maximum climatological mixed layer, including entrainment and seasonal reemergence of temperature anomalies that are seasonally subducted below the mixed layer. This reemergence has been shown to be an important driver of SST variability (De Coëtlogon & Frankignoul, 2003; Deser et al., 2003).

We infer the effective mixed layer depth ($\mathrm{MLD}_{\mathrm{eff}}$) from the atmospheric response to SST and HC anomalies defined by the ratio of λ_{SST} to λ_{HC} scaled by $\mathrm{MLD}_{\mathrm{max}}$. $\mathrm{MLD}_{\mathrm{eff}}$ represents the equivalent depth of the water column that would be changed by 1 °C through the turbulent heat flux in the absence of ocean advection and entrainment processes:

$$\mathrm{MLD}_{\mathrm{eff}} = \mathrm{MLD}_{\mathrm{max}} \frac{\lambda_{SST}}{\lambda_{HC}}. \tag{1.5}$$

The magnitude of $\mathrm{MLD}_{\mathrm{eff}}$ represents the turbulent heat flux response to all ocean processes (i.e., advection, entrainment, and mixing) within $\mathrm{MLD}_{\mathrm{max}}$. If the turbulent heat flux response to SST anomalies is less than or equal to the response to equivalent mixed layer HC anomalies, then $\mathrm{MLD}_{\mathrm{eff}}$ will be less than or equal to $\mathrm{MLD}_{\mathrm{max}}$, indicating that mixed layer processes damp SST anomalies and reduce surface heat flux variability. On the other hand, if the turbulent heat flux response to mixed layer HC is greater than the response to SST anomalies ($\text{MLD}_{\text{eff}} > \text{MLD}_{\text{max}}$), then the ocean contributes to driving SST and Q variability.

We also analyze the ocean's role in driving the heat flux feedback by reformulating the stochastic model (Eq. 1.2) in terms of the mixed layer heat content, allowing the depth h to vary in time. Observations of HC are integrated to the maximum climatological mixed layer depth (MLD_{max}) at each location, but the depth of water that participates in air-sea heat exchange and the surface heat flux feedback changes. We can rewrite Eq. 1.2 in terms of HC to arrive at a stochastic model for the HC tendency (see Appendix A):

$$\frac{\partial \mathrm{HC}'}{\partial t} = \rho C_p h(F_a' + F_o') - \left(\lambda_a + \lambda_o + \frac{1}{h}\nabla \cdot h\boldsymbol{u}\right)\mathrm{HC}', \tag{1.6}$$

where $\boldsymbol{u}$ is the horizontal velocity in the mixed layer and $\partial h/\partial t = -\nabla \cdot (h\boldsymbol{u})$ represents heat transport due to MLD changes.

In Eq. 1.6, heat transport convergence in the mixed layer reduces the effective damping of heat anomalies by the atmosphere, thus reducing the effective heat flux feedback to HC anomalies. We calculate λ_{SST} and λ_{HC} according to Eqs. 1.3 and 1.4 and take the difference between them to give the feedback sensitivity:

$$\lambda_{SST} - \lambda_{HC} \propto -\frac{1}{h}\,\nabla \cdot h\boldsymbol{u}. \tag{1.7}$$

The feedback sensitivity represents the net effect of heat convergence within the mixed layer—via entrainment and reemergence processes (subduction or obduction)—on the heat flux feedback. Where the feedback sensitivity is positive, ocean heat content convergence increases the sensitivity of the atmosphere to ocean heat anomalies. A negative feedback sensitivity represents reduced atmospheric sensitivity to SST anomalies due to ocean processes.

1.2.3. Data

To examine the spatial distribution of λ_{SST}, λ_{HC}, $\mathrm{MLD_{eff}}$, and the feedback sensitivity we use observational datasets of Q, SST, HC, and $\mathrm{MLD_{max}}$. Monthly 1° x 1° estimates of turbulent surface heat flux (Q; positive out of the ocean), sea surface temperature (SST) anomalies, and mixed layer heat content (HC) anomalies are derived from observational data products spanning 30 years of satellite observations from 1993 to 2022. The turbulent heat flux is calculated from the sum of latent and sensible heat flux products from the Woods Hole Oceanographic Institution's Objectively Analyzed air-sea Fluxes Project (OAFlux) (Yu & Weller, 2007), which are derived from a synthesis of satellite observations and atmospheric reanalysis and are provided on monthly 1° x 1° grids. For SST, we use the NOAA Optimum Interpolation Sea Surface Temperature (OISST) dataset accessed from OAFLUX (Reynolds et al., 2007; Yu & Weller, 2007), which integrates observations from satellites, ships, buoys, and Argo floats.

To obtain an estimate of the mixed layer HC, we use the weekly 1/4° ocean heat content product from Random Forest Regression Ocean Maps (RFROM) (Lyman & Johnson, 2023). RFROM uses a random forest algorithm with time, latitude, longitude, sea surface height, and SST as predictors for in situ measurements to estimate heat content for 10 layers over the upper 2000 m of the global ocean. We calculate monthly mixed layer heat content by resampling the upper ocean heat content from RFROM to a monthly timescale and integrating the heat content from the surface to $\mathrm{MLD_{max}}$ at each grid cell. We define $\mathrm{MLD_{max}}$ from a 12-month mixed-layer depth climatology from the Monthly Isopycnal and Mixed-layer Ocean Climatology (MIMOC)

(Schmidtko et al., 2013). These mixed-layer depths are based on Argo CTD data and shipboard observations and are provided at a 0.5° horizontal resolution, which we interpolate to a 1° spatial grid.

We interpolate all data to a common 1° x 1° grid and remove the linear trend and the monthly climatology from the Q, SST, and HC fields. Finally, we use a 200 km full width at half maximum Gaussian filter to smooth each variable. This filtering process reduces noise and allows us to focus the interpretation of the results to mesoscale and large-scale ocean circulation features.

1.3. Results

1.3.1. The Seasonal Cycle of λ

The heat flux feedback is first calculated from year-round fields using Eqs. 1.3 and 1.4. We find that λ_{SST} and λ_{HC} is positive over most of the global ocean indicating that the surface heat flux dampens both SST and HC variability on monthly timescales (Fig. 2). This result is consistent with previous studies that use similar methods to calculate the SST-Q feedback from both observational data and model output (Frankignoul & Kestenare, 2002; Hausmann et al., 2016, 2017). The large-scale patterns of λ_{SST} and λ_{HC} are similar, with the highest values occurring along frontal regions such as the Gulf Stream, the Kuroshio Current, and the Agulhas Return Current (ARC). These regions of strong ocean currents have both large ocean heat transport convergence as well as large air-sea temperature differences as continental air encounters ocean water.

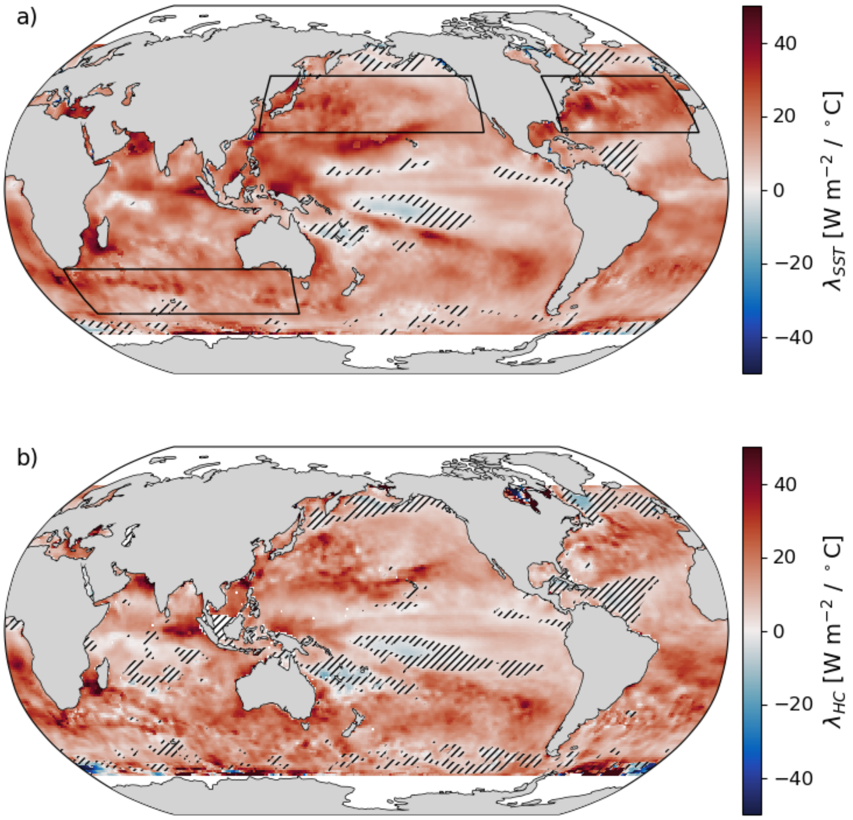


Figure 1.2. Year-round SST-Q feedback (a) and HC-Q feedback (b). Positive values imply that Q anomalies damp SST or HC anomalies and negative values imply that Q anomalies amplify SST or HC anomalies. Hash marks indicate regions with statistically insignificant SST-Q or HC-Q covariance, respectively. The boxes in panel (a) outline the western North Atlantic, North Pacific, and Southern Ocean regions.

There are also large seasonal variations in the magnitude of both λ_{SST} and λ_{HC}, though the spatial patterns are similar between the two quantities and both remain positive across all seasons except in isolated regions in the tropics and high latitudes (Fig. 1.3). Both λ_{SST} and λ_{HC} are largest in the winter and smallest in the summer (Fig. 1.3) due to strong wintertime winds and cold, dry air parcels that cause large turbulent heat fluxes leading to larger atmospheric damping from turbulent fluxes (Frankignoul et al., 2002; Park et al., 2005). Weaker summertime winds lead to smaller feedbacks, indicating weaker coupling between the ocean and the atmosphere. The highest feedbacks again occur in western boundary current regions. Feedbacks are also large in the autumn (SON in the Northern Hemisphere and MAM in the Southern Hemisphere) as the mixed layer deepens.

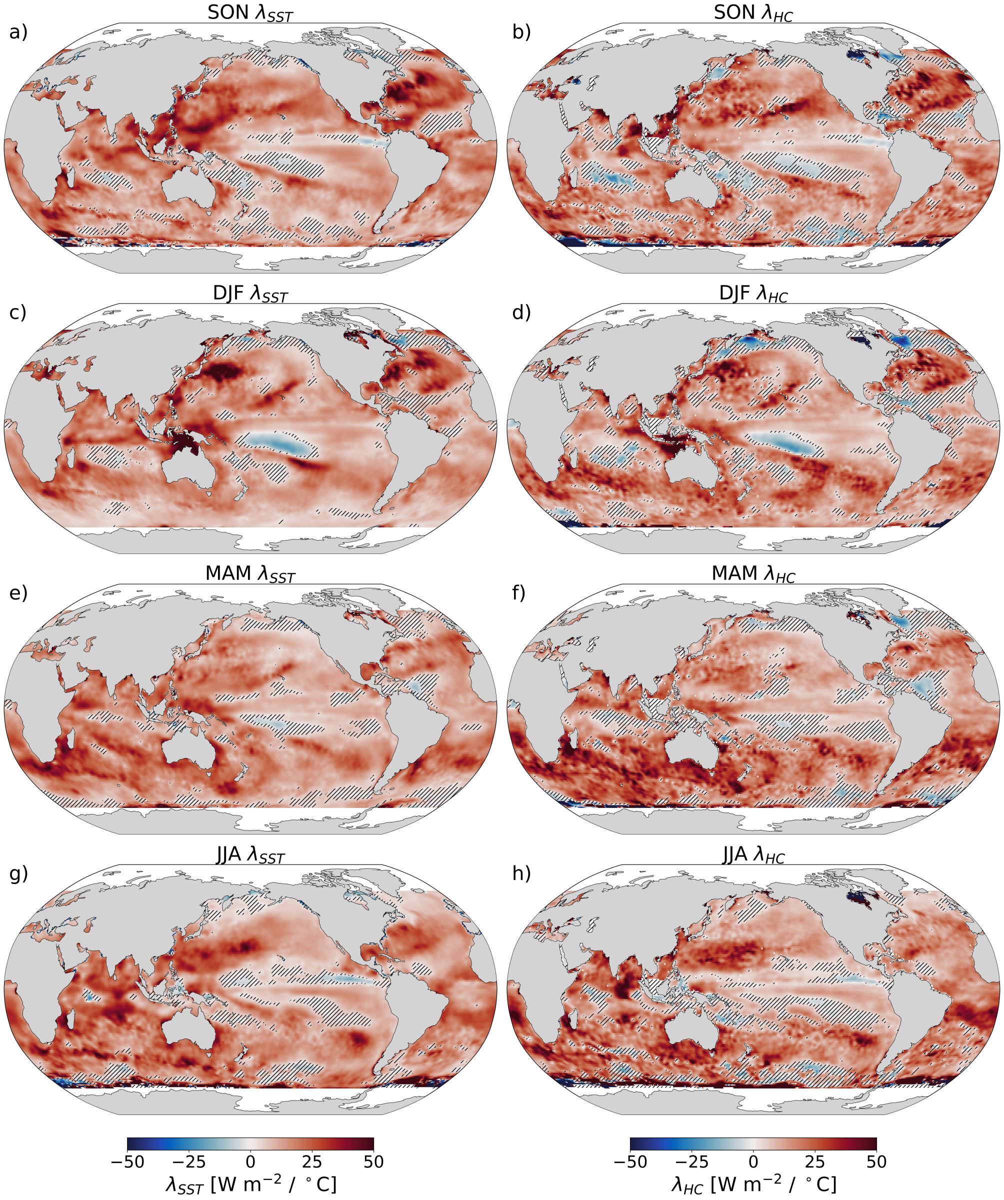


Figure 1.3. Seasonal SST-Q feedback (left column) and HC-Q feedback (right column) for SON

(a, b), DJF (c, d), MAM (e, f), and JJA (g, h). Hash marks indicate regions with statistically insignificant SST-Q or HC-Q covariance, respectively.

λ_{SST} and λ_{HC} also depend on the month of the year, showing distinct seasonal cycles in the North Atlantic Ocean, the North Pacific Ocean, the Southern Ocean (Fig. 1.4 for regions defined in Fig. 1.2a). All three regions include strong western boundary currents: the Gulf Stream, the Kuroshio Current, and the ARC, respectively. In autumn, both λ_{SST} and λ_{HC} increase in all three regions before peaking in November in the Northern Hemisphere and May in the Southern Hemisphere. The basin-wide feedbacks then decrease in the winter, reaching their lowest values in late spring or summer.

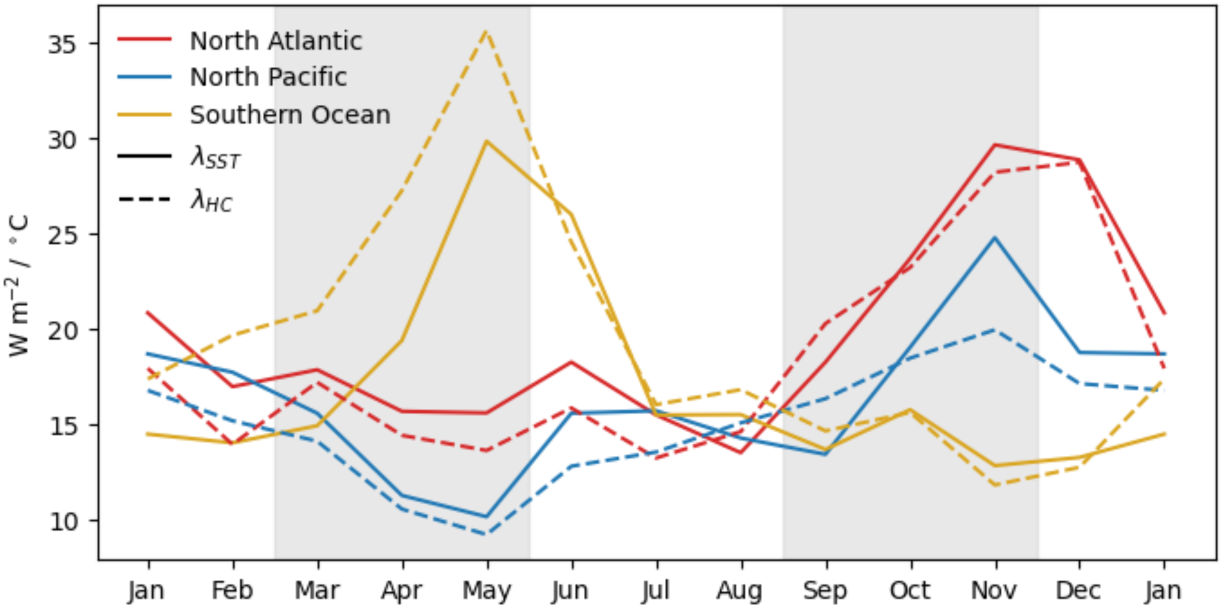


Figure 1.4. Monthly SST-Q feedback (solid) and HC-Q feedback (dashed) in the midlatitude ocean basins outlined in Fig. 1: North Atlantic, red; North Pacific, blue; Southern Ocean, yellow. Horizontal shading separates the seasons used in Fig. 1.3.

Ocean-atmosphere feedbacks (both λ_{SST} and λ_{HC}) vary substantially throughout the year, with peak winter values more than double those found in summer. This seasonal cycle is apparent for both λ_{SST} and λ_{HC}. Previous studies examine turbulent heat flux feedback calculated over the entire year to draw conclusions about the relative roles of the ocean and atmosphere in driving SST variability, but often do not analyze the seasonality of results (e.g., Patrizio & Thompson, 2022b). The results presented here suggest that along with λ_{SST} and λ_{HC}, the roles of the ocean and the atmosphere vary throughout the year.

1.3.2. The Impact of the Ocean Mixed Layer on λ

The MLD_{eff} is used to quantify the relative role of the ocean and atmosphere in driving air-sea turbulent flux exchange. The MLD_{eff} is inferred from the relative magnitude of λ_{SST} to λ_{HC} (Eq. 1.5). In the annual case and in each season, MLD_{eff} is largest in the ACC, the Gulf Stream, and the subpolar North Atlantic (Fig. 1.5b–f). These regions have strong currents and the deepest mixed layers (Fig. 1.5a). Both MLD_{max} and MLD_{eff} have their largest values in the winter, although the interior North Atlantic shows large values of MLD_{eff} in the summer as well (Fig. 1.5d and 1.5f). Deep effective mixed layers suggest that ocean processes both within and below the climatological mixed layer drive heat flux variability. Thus, even in the summer, heat anomalies trapped below the shallow mixed layer still contribute to driving heat flux anomalies. This results in a deeper MLD_{eff}. These results match results found in previous studies that inferred Northern Hemisphere mixed layer depths from turbulent heat fluxes (Cayan, 1992; Deser et al., 2003).

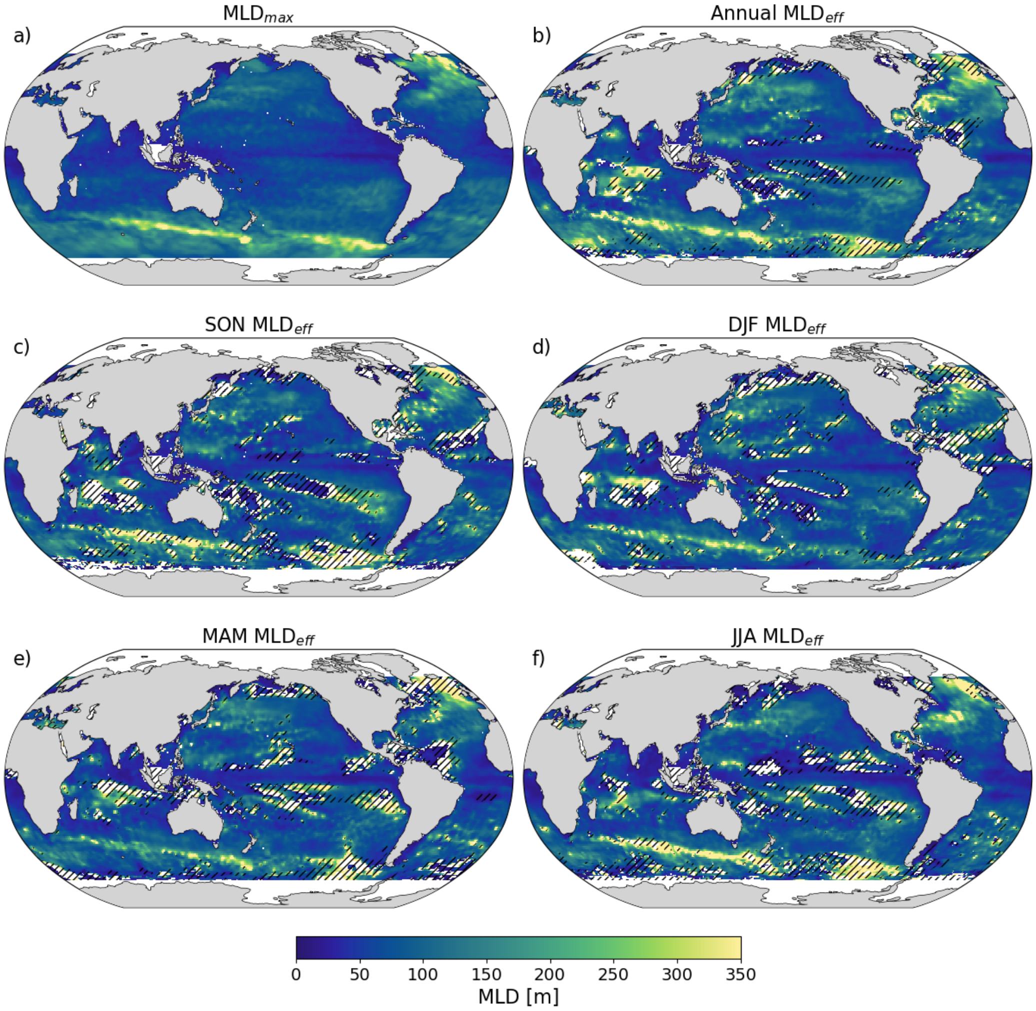


Figure 1.5. The maximum mixed layer depth (MLD$_{\text{max}}$; a), the year-round effective mixed layer depth (MLD$_{\text{eff}}$; b), and the seasonal MLD$_{\text{eff}}$ in SON (c), DJF (d), MAM (e), and JJA (f). Hash marks in panels b–f indicate regions with statistically insignificant SST-Q or HC-Q covariance.

The relative importance of ocean processes in driving the heat flux feedback can be estimated by defining the feedback sensitivity ($\lambda_{SST} - \lambda_{HC}$), which quantifies the strength of processes within the ocean that contribute to the creation of ocean heat content anomalies within the mixed layer.

The feedback sensitivity is positive in many regions, including in the North Atlantic Ocean, the western North Pacific Ocean, and the western Indian Ocean (Fig. 1.6a). The largest feedback sensitivities can be found in the GS, KC, and ARC regions, indicating that processes within the mixed layer in these strong currents drive both SST and Q anomalies.

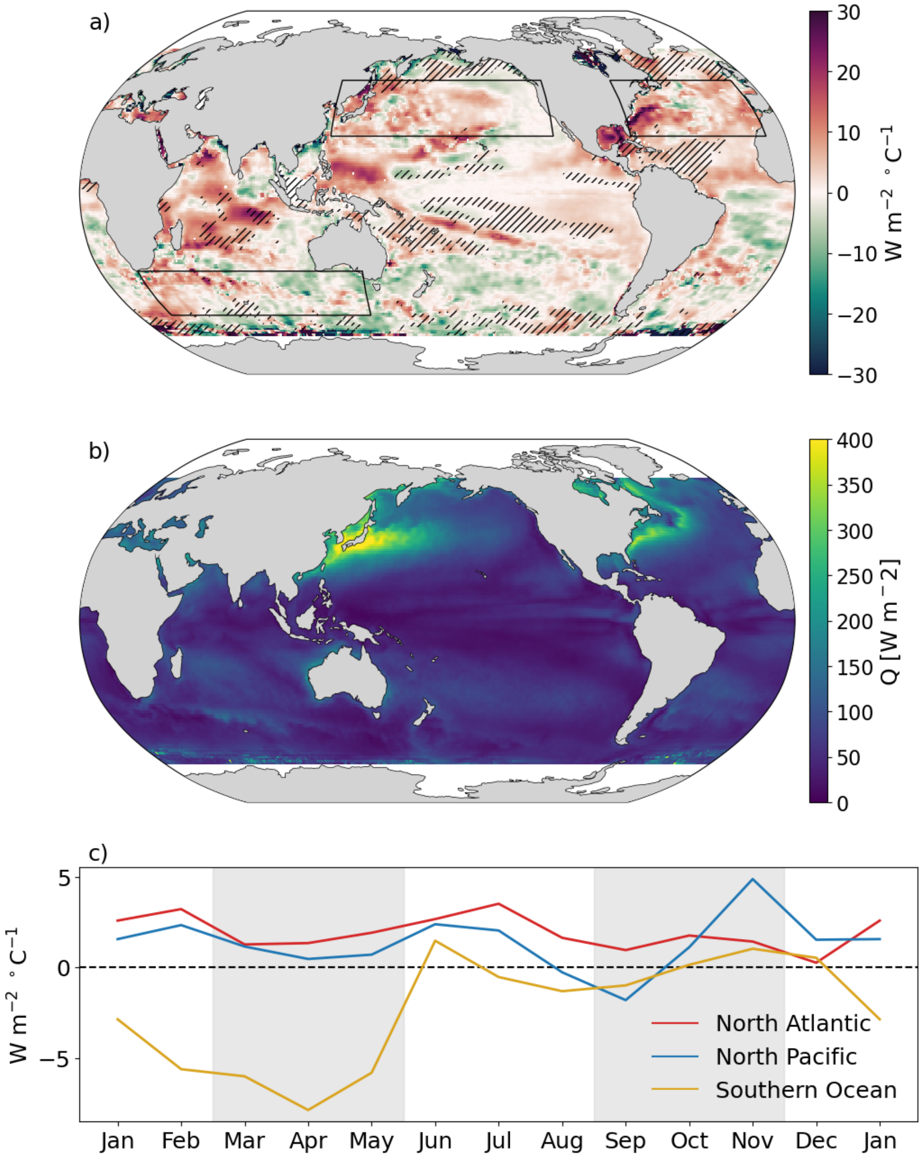


Figure 1.6. The annual feedback sensitivity (a), the seasonal amplitude of turbulent surface heat flux Q (b), and monthly feedback sensitivity (c) for midlatitude ocean basins outlined in panel (a). Hash marks in panel (a) indicate regions with statistically insignificant SST-Q or HC-Q covariance.

In all three midlatitude ocean basins (outlined in Fig. 1.6a) the impact of ocean processes on the heat flux feedback varies throughout the year. In the Northern Hemisphere regions, ocean processes control the heat flux feedback for most of the year. The sensitivity peaks in July in the North Atlantic and in November in the North Pacific (Fig. 1.6c). In the Southern Ocean, however, λ_{HC} effectively exceeds λ_{SST} for much of the year. The feedback sensitivity peaks in both June and November but remains negative from January through May. Southern Ocean mixed layers are deeper than those in the Northern Hemisphere (Fig. 1.1d), but the amplitude of the seasonal cycle of Q (climatological maximum minus climatological minimum) is far smaller in the Southern Ocean than in the Northern Hemisphere, owing to the limited influence of the continental seasonal cycle (Fig. 6b; Whitt et al., 2019). As a result, entrainment through mixed layer deepening and shoaling has a larger effect on the surface heat flux in the North Atlantic and North Pacific regions than in the Southern Ocean, supporting the results in Fig. 1.6c.

The spatial averaging used to calculate the results in Fig. 1.6c obscures the detailed spatial structure of the ocean's effect on the heat flux feedback and the dependence of the spatial patterns on seasonality (Fig. 1.7). The western ocean basins in the Northern Hemisphere show positive feedback sensitivity in the fall and winter when strong currents drive turbulent heat fluxes (Figs. 1.7a and 1.7b), while the central ocean basins have positive feedback sensitivity in the spring and summer when mixed layers are extremely shallow and even small changes in HC can drive heat fluxes (Figs. 1.7c and 1.7d).

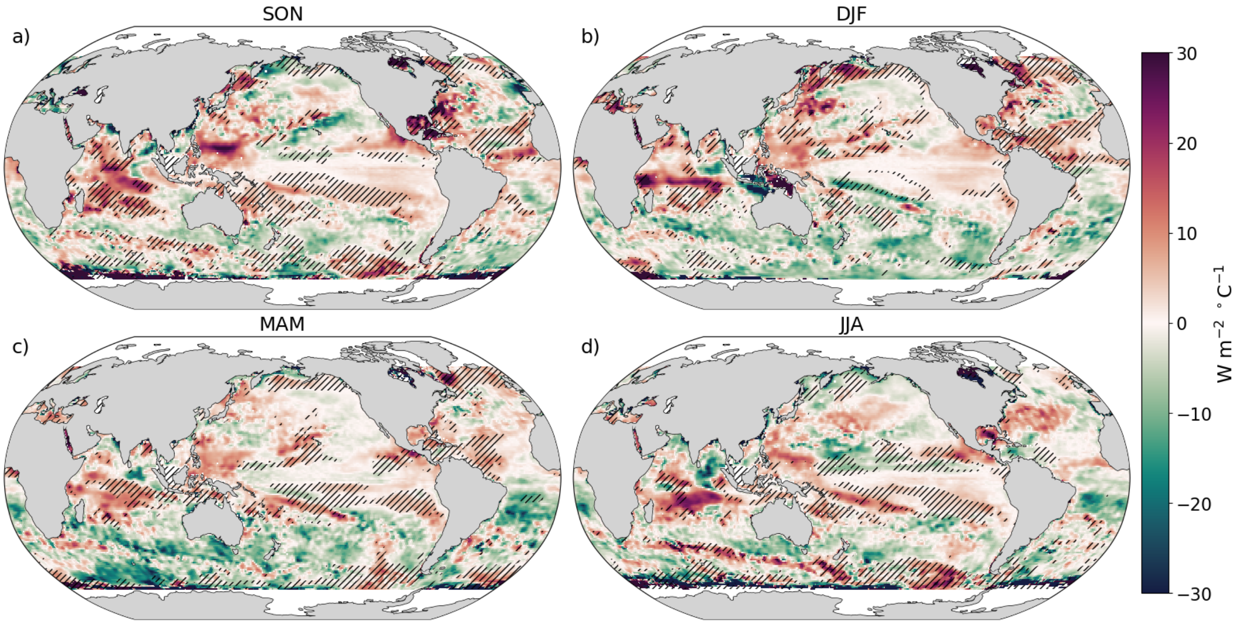


Figure 1.7. Seasonal feedback sensitivity for SON (a), DJF (b), MAM (c), and JJA (d). Hash marks indicate regions with statistically insignificant SST-Q or HC-Q covariance.

In the Southern Ocean, the feedback sensitivity is negative over broad regions in the summer and fall (Figs. 1.7b and 1.7c). Although the summer mixed layer is shallow here as well, smaller turbulent heat flux variability leads to a reduced impact of ocean processes when compared to the midlatitude Northern Hemisphere. In all seasons, the Gulf stream, the Kuroshio Current, and the ARC have positive feedback sensitivity, highlighting the important role of ocean processes in driving heat anomalies. The feedback sensitivity's magnitude and spatial extent in these regions is highest in wintertime and lowest in summertime (Figs. 1.7b and 1.7d).

1.3.3. Dependence on Spatial Scale

The role of ocean processes in driving SST and surface heat flux variability also depends on the spatial scale. At small scales, variability is primarily driven by the ocean, while the atmosphere dominates at larger scales (Laurindo et al., 2022; Roberts et al., 2017). To examine spatial scale dependence, we smooth the SST, HC, and Q data with Gaussian filters that have full-width at half maximum values of 100 km, 200 km, 300 km, 500 km, 700km, and 1000 km, and use those fields to calculate the annual feedback difference to understand how the effect of ocean processes changes with spatial scale.

The ocean's role in driving the heat flux feedback diminishes as the smoothing scale increases, in agreement with previous results on the space-scale dependence of air-sea interaction. As we increase the smoothing distance, both λ_{SST} and λ_{HC} decrease in magnitude. The difference between λ_{SST} and λ_{HC} also decreases, shown via the feedback sensitivity (Fig. 1.8).

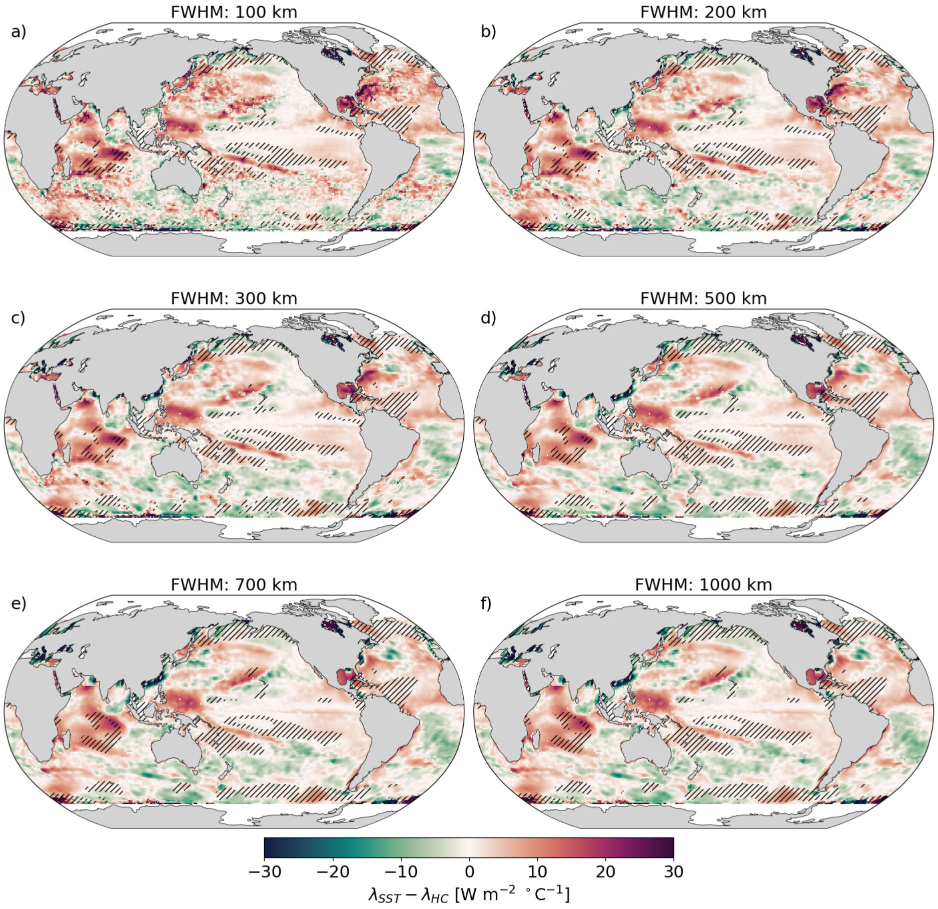


Figure 1.8. Feedback sensitivity by smoothing filter distance. λ_{SST} and λ_{HC} are calculated from SST, HC, and Q data smoothed at 100 km (a), 200 km (b), 300 km (c), 500 km (d), 700 km (e),

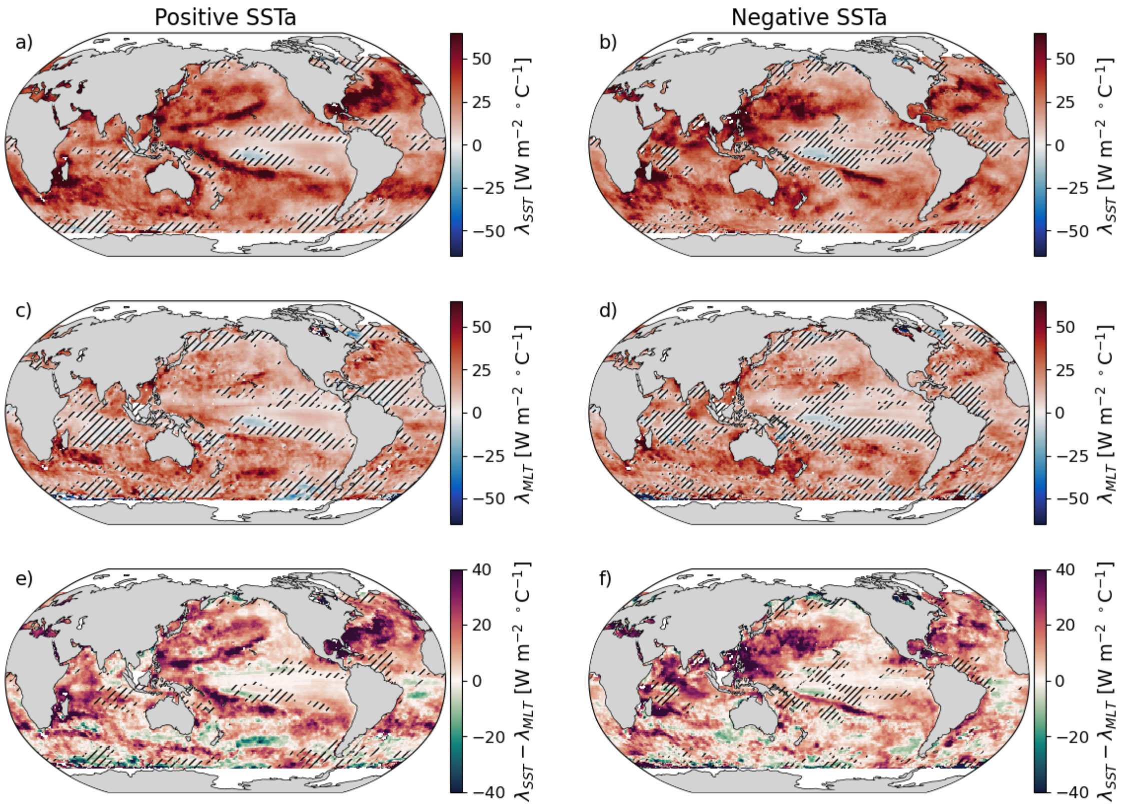
and 1000 km (f). Hash marks indicate regions with statistically insignificant SST-Q or HC-Q covariance.

Smoothing the input data with a 200 km Gaussian filter preserves the mesoscale features in the feedback sensitivity while reducing the noise present in the unfiltered data. Using a larger-scale filter removes some of the mesoscale variability and reduces the effective impact of ocean processes on SST variability and the turbulent heat flux feedback (Fig. **1.**8). The magnitude of the feedback sensitivity diminishes as the full width at half maximum distance of the Gaussian filter increases up to ~500 km; the 700-km- and 1000-km-smoothed feedback sensitivities look similar to the 500-km-smoothed sensitivity. This suggests that the ocean's influence on air-sea interaction is strongest at scales less than 500 km, in agreement with past results (Bishop et al., 2017).

1.3.4. Dependence on the Sign of SST Anomalies

The heat flux feedback to SST and HC anomalies and the ocean's contribution to the heat flux feedback also depends on whether the monthly SST anomalies are positive or negative. Because SST and HC anomalies are linked through the mixed layer depth, there is a nonlinear response between Q and SST or HC (Tomita & Nonaka, 2006). As a result, the role of ocean mixed layer processes on heat flux anomalies depends on the sign of the SST anomalies. Tomita and Nonaka (2006) find that negative March SST anomalies in the North Pacific Ocean induce a larger response in atmospheric forcing than positive SST anomalies. They conclude that more heat flux is required to change negative SST anomalies because negative SST anomalies are dynamically linked to anomalously deep mixed layers. They also conclude that the sensitivity of negative SST anomalies to surface forcing most closely resembles the year-round case and that separating the cases by the sign of SST anomalies reveals a larger estimate of the thermal impact of the ocean on the atmosphere (Tomita & Nonaka, 2006).

To evaluate the ocean-state-dependence of our results, we calculate λ_{SST} and λ_{HC} separately for all locations and months with positive SST anomalies and for those with negative SST anomalies (Fig. 1.9a–d). In agreement with the results from Tomita and Nonaka (2006), λ_{SST} and λ_{HC} resemble the year-round case when calculated from just the negative SST anomalies (Figs. 1.2a, 1.9b, and 1.9d). The feedback sensitivity in the year-round case also most closely resembles the negative SST anomaly case, but in some regions resembles a combination of the positive and negative SST anomaly cases (Figs. 1.6a, 1.7e, and 1.7f).


Positive SSTa

Negative SSTa

Figure 1.9. SST-Q feedback (a, b), HC-Q feedback (c, d), and the feedback sensitivity (e, f) for

positive SST anomalies (left column) and negative SST anomalies (right column). Hash marks indicate regions with statistically insignificant SST-Q or HC-Q covariance.

These results demonstrate that the air-sea feedback and its relationship to ocean processes depends on the sign of SST anomalies. The magnitudes of λ_{SST}, λ_{HC}, and the feedback sensitivity are larger when we calculate them separately for positive and negative SST anomaly cases (Fig. 1.9). This dependence likely results from the different response of SST to turbulent heat fluxes when the mixed layer is anomalously deep or deepening versus when it is anomalously shallow or shoaling. Thus, the impact of averaging feedbacks for both positive and negative SST can reduce the estimates of λ_{SST}, λ_{HC}, and the feedback sensitivity. Selecting months based on positive and negative HC anomalies and negative and positive Q anomalies yields qualitatively similar results.

1.4. Conclusions

The relationships between ocean heat content, air-sea heat exchange, and physical processes within the ocean mixed layer govern climate variability across spatiotemporal scales (Bishop et al., 2017; Laurindo et al., 2022; Small et al., 2020). Uncovering the mechanisms that control and drive these relationships can contribute to the broad understanding of the dual roles of the ocean and atmosphere in driving and damping ocean heat variability.

Here, we show that the relationships between observed turbulent heat fluxes, SST, and HC vary in time and space. The seasonal analysis of λ_{SST} and λ_{HC} (Figs. 1.3 and 1.4) demonstrate that air-sea coupling and feedback are strongest in regions with strong currents and in winter, when cold continental air interacts with the warm ocean water. Seasonal evaluation of MLD_{eff} and the feedback sensitivity further support the conclusion that ocean heat transport convergence controls SST and turbulent heat flux variability in the winter in these strong currents and in the summer when the ocean mixed layer is at its shallowest (Figs. 1.5–1.7). Finally, the turbulent heat flux feedback and the role of upper ocean heat convergence depend on the spatial scale of analysis and on the sign of SST anomalies (Figs. 1.8 and 1.9).

The results presented here vary among ocean basins and support findings from previous studies. Upper ocean heat content variability in the Gulf Stream and the Kuroshio Current from monthly to interannual timescales is primarily controlled by interior ocean processes, and in turn drives turbulent surface heat flux variability (Buckley et al., 2014; Kelly et al., 2010). The Kuroshio Current winter SST variability has been shown to be driven by ocean dynamics such as advection and reemergence (Pak et al., 2017; Qiu, 2002; Yang et al., 2021). Alternatively, as our results suggest, SST and Q variability in the North Atlantic and North Pacific outside the western boundary currents and in the summer is driven by local atmospheric forcing and winds (Buckley et al., 2014; Kelly et al., 2010; Nonaka & Xie, 2003; Small et al., 2019; Wang et al., 2012).

Our results demonstrate that the Southern Ocean SST variability has a different seasonal cycle than the North Atlantic and the North Pacific Oceans (Fig. 1.6). We find that ocean processes are important for winter and fall SST variability but have a minor role in driving SST variability in the summer. This likely occurs because summer mixed layer depths are extremely shallow but turbulent heat fluxes remain relatively constant over the year (Fig. 1.6b). Analyses of the roles of ocean advection and air-sea heat exchange in controlling Southern Ocean heat budgets similarly show that heat advection is more important than surface fluxes except in the summer (Dong et al., 2007; Gao et al., 2022).

One difference between this analysis and past studies (Patrizio & Thompson, 2021, 2022a; Roberts et al., 2017) is that we do not attempt to directly calculate the forcing and damping by ocean processes. Instead, we infer the strength of ocean processes by comparing the turbulent heat flux response to changes in SST versus changes in HC, which are directly impacted by mixed layer processes. We primarily assess the oceanic contribution to SST variability through advection, entrainment, and reemergence processes that interact with the turbulent heat flux. Nonetheless, our results agree well with the results from previous studies and support their conclusions, suggesting that the relationships between SST, HC, and Q capture the effect of ocean heat transport convergence on observed SST and Q variability.

Explaining the conditions, locations, and seasons where ocean processes drive air-sea interaction has important implications for understanding climate variability and predictability. Interactions between the ocean and atmosphere have been shown to enhance midlatitude ocean predictability (Bach et al., 2019). Furthermore, air-sea coupling and feedback parameterization in climate models is an ongoing challenge, as model results depend on resolution, parameterizations, and subgrid-scale physics (Agarwal et al., 2023; Busecke et al., 2025; Torres et al., 2019). Continued research on the drivers of air-sea feedback processes can potentially improve the fidelity of both ocean and coupled models.

Appendix 1A: Deriving the stochastic model for heat content tendency

We first take the tendency of the heat content anomaly, $\mathrm{HC}' = \rho C_p h \mathrm{T}'$,

$$\frac{\partial \mathrm{HC}'}{\partial t} = \rho C_p h \frac{\partial \mathrm{T}'}{\partial t} + \rho C_p \mathrm{T}' \frac{\partial h}{\partial t}. \qquad (1.A1)$$

We then substitute $^{\partial \mathrm{T}'}/_{\partial t}$ with Eq. 1.2, giving,

$$\frac{\partial \mathrm{HC}'}{\partial t} = \rho C_p h\left(F_a' + F_o' - (\lambda_a + \lambda_o)\mathrm{T}' + \frac{1}{h}\mathrm{T}'\right) + \rho C_p \mathrm{T}' \frac{\partial h}{\partial t}. \tag{1.A2}$$

We then substitute in the equation for mass convergence that results from a change in the height of the water column, $\partial h/\partial t = -\nabla \cdot (hu)$, giving:

$$\frac{\partial \mathrm{HC}'}{\partial t} = \rho C_p h\left(F_a' + F_o' - (\lambda_a + \lambda_o)\mathrm{T}' + \frac{1}{h}\mathrm{T}'\big(-\nabla \cdot (hu)\big)\right). \tag{1.A3}$$

Equation 1.A3 simplifies to,

$$\frac{\partial \mathrm{HC}'}{\partial t} = \rho C_p h\left(F_a' + F_o' - \left(\lambda_a + \lambda_o + \frac{1}{h}\nabla \cdot hu\right)\mathrm{T}'\right). \tag{1.A4}$$

Using the definition of the HC anomaly $\mathrm{HC}' = \rho C_p h \mathrm{T}'$, we arrive at Eq 6:

$$\frac{\partial \mathrm{HC}'}{\partial t} = \rho C_p h(F_a' + F_o') - \left(\lambda_a + \lambda_o + \frac{1}{h}\nabla \cdot hu\right)\mathrm{HC}'. \tag{1.A5}$$

We use Eq. 1.A5/Eq. 1.6 similarly to Eq. 1.2 to define the HC-Q feedback.

2. Object-Based Evaluation of Seasonal-to-Multiyear Marine Heatwave Predictions

This chapter is reproduced from Cohen, J. T., Thompson, L., Maroon, E., Deppenmeier, A.-L., Cai, C. (2025). Object-based evaluation of seasonal-to-multiyear marine heatwave predictions. Geophysical Research Letters, 52, e2025GL115021. https://doi.org/10.1029/2025GL115021

2.1. Introduction

Emissions over the past century have led to irreversible climate change. The combination of long-term warming and extreme sea surface temperature (SST) events, known as marine heatwaves (MHWs), is pushing ocean temperatures to increasingly extreme highs (Deser et al., 2024; Frölicher et al., 2018; Oliver et al., 2018). These events cause devastating effects on biological communities and marine ecosystems already experiencing stress from pollutants and microplastics (Jacox et al., 2020; Mills et al., 2013; Smale et al., 2019; Smith et al., 2023). MHWs also have socioeconomic impacts including fish farm mortality and the closure of commercial fisheries, with economic costs that can exceed US$800 million (Smith et al., 2021). Accurate MHW forecasts could allow local decision makers and industries to plan for and respond to these events (Hartog et al., 2023; Hobday et al., 2024).

Understanding the mechanisms that drive MHWs is necessary to improve predictive capabilities. Global analyses reveal that MHWs are driven by local atmospheric and oceanic variability, including changes in shortwave radiation, surface fluxes, advection, and Ekman transport (Bian et al., 2023, 2024; Marin et al., 2022).These drivers are often associated with large-scale modes of climate variability like the Indian Ocean Dipole, the Pacific Decadal Oscillation, and El Niño–

Southern Oscillation (ENSO) (Capotondi et al., 2024; Gregory et al., 2024; Holbrook et al., 2019; Sen Gupta et al., 2020). Focused studies identify drivers of individual MHWs, including those in the Northeastern Pacific (Amaya et al., 2020; Fewings & Brown, 2019; Scannell et al., 2020; Schmeisser et al., 2019), the Southwestern Atlantic (Manta et al., 2018), the Tasman Sea (Kajtar et al., 2022; Oliver et al., 2017), the Indian Ocean (Qi et al., 2022), and off the coast of Western Australia (Benthuysen et al., 2014; Pearce & Feng, 2013).

MHWs are commonly defined as SST anomalies relative to a long-term climatology that exceed the local 90th percentile threshold of SST variability (Hobday et al., 2016). For monthly data, SST anomalies above a 90th percentile threshold for any duration are considered MHWs (Capotondi et al., 2024; Scannell et al., 2020). While useful for standardizing MHW identification, this definition fails to capture the spatial coherence of MHWs, which are spatially connected regions, or objects, that can change shape and move across ocean basins. The spatiotemporal evolution of MHWs has been the focus of numerous studies (e.g., Bonino et al. 2023; Scannell et al. 2024; Sun et al. 2023). Treating MHWs as objects instead of as points introduces new ways to quantify prediction skill and assess sources of predictability.

Analyses of ensemble hindcast systems from the North American Multi-Model Ensemble (NMME) and the European Centre for Medium-Range Weather Forecast (ECMWF) demonstrate long-lasting skill for MHW predictions globally (De Boisséson & Balmaseda, 2024; Jacox et al., 2022). In these studies, a correct prediction only counts when the predicted MHW occurs at the precise location of the observed MHW. If the MHW is predicted in the incorrect location, the event is counted as an error twice: once as a missed prediction where the MHW was observed and again as a false positive where it was predicted. To address this “double penalty effect” (Rossa et al., 2008), the numerical weather prediction community developed forecast verification techniques that go beyond pointwise metrics (Dorninger et al., 2018; Gilleland et al., 2009, 2010). Here we apply an object-based verification technique to hindcast predictions of MHWs to evaluate how well the model predicts the location, overlap, area, and intensity of spatially connected extreme SSTs.

2.2. Data and Methods

2.2.1. Data

We analyze the Seasonal-to-Multiyear Large Ensemble (SMYLE), a set of hindcast simulations run using the Community Earth System Model (CESM2) with a nominal 1° horizontal resolution for models of each component (Danabasoglu et al., 2020; Yeager et al., 2022). The initial conditions for SMYLE come from a forced-ocean-sea-ice model (SMYLE FOSI). SMYLE FOSI is run from 1958–2020 using the Parallel Ocean Program version 2 (POP2) and the Community Ice CodE 5.1.2 (CICE5) forced by historical atmospheric conditions given by the Japanese 55-year Reanalysis (JRA-55-do; Tsujino et al., 2018). SMYLE consists of 24-month long forecasts initialized on the first of every February, May, August, and November from 1970-2019. Each forecast has 20 ensemble members. These data have been used to study predictability across the climate system, including MHWs and ocean acidification (Mogen et al., 2024), ecosystem stressors (Mogen et al., 2023), and the North Atlantic Oscillation (Dunstone et al., 2023).

The initialization month is defined as the month in which the forecast was initialized and the lead month or lead time as the time since initialization. Following Jacox et al. (2022), lead month 0.5 refers to the predicted monthly mean of the initialization month. The target month is the month being predicted, which varies for different initialization months and lead months.

We examine monthly means of the top layer (5 m depth) ocean temperature from SMYLE, which we refer to as SST. We compare the SMYLE forecast data to observations from NOAA's Optimum Interpolation Sea Surface Temperature v2.1 (OISST) product (Huang et al., 2021), which provides daily SST values at 0.25° resolution. Following Jacox et al. (2022), we resample the data to monthly frequency to match the SMYLE forecast resolution. We regrid the SST from both SMYLE and OISST to a common 1° x 1° grid and mask data poleward of 65° to avoid small grid cells at high latitudes. We also exclude the Black Sea, the Baltic Sea, the Red Sea, and the Hudson Bay.

2.2.2. Defining Marine Heatwaves in Time and Space

We define the point locations of MHWs for each month before defining spatially connected MHW objects (Fig. 2.1). In OISST, we remove the monthly climatology and linear trend calculated from the 30-year reference period of 1989–2018 to define SST anomalies. We then calculate a 90th percentile threshold for each month and location, above which any SST anomaly is considered an MHW. In SMYLE, SST anomalies (Fig. 2.1b) are calculated by removing a lead-dependent (24-month) climatology and linear trend for each initialization month based on the 1989–2018 reference period (Jacox et al., 2022). We then calculate a lead-dependent 90th percentile threshold to define the MHWs at each grid cell (Fig. 2.1c).

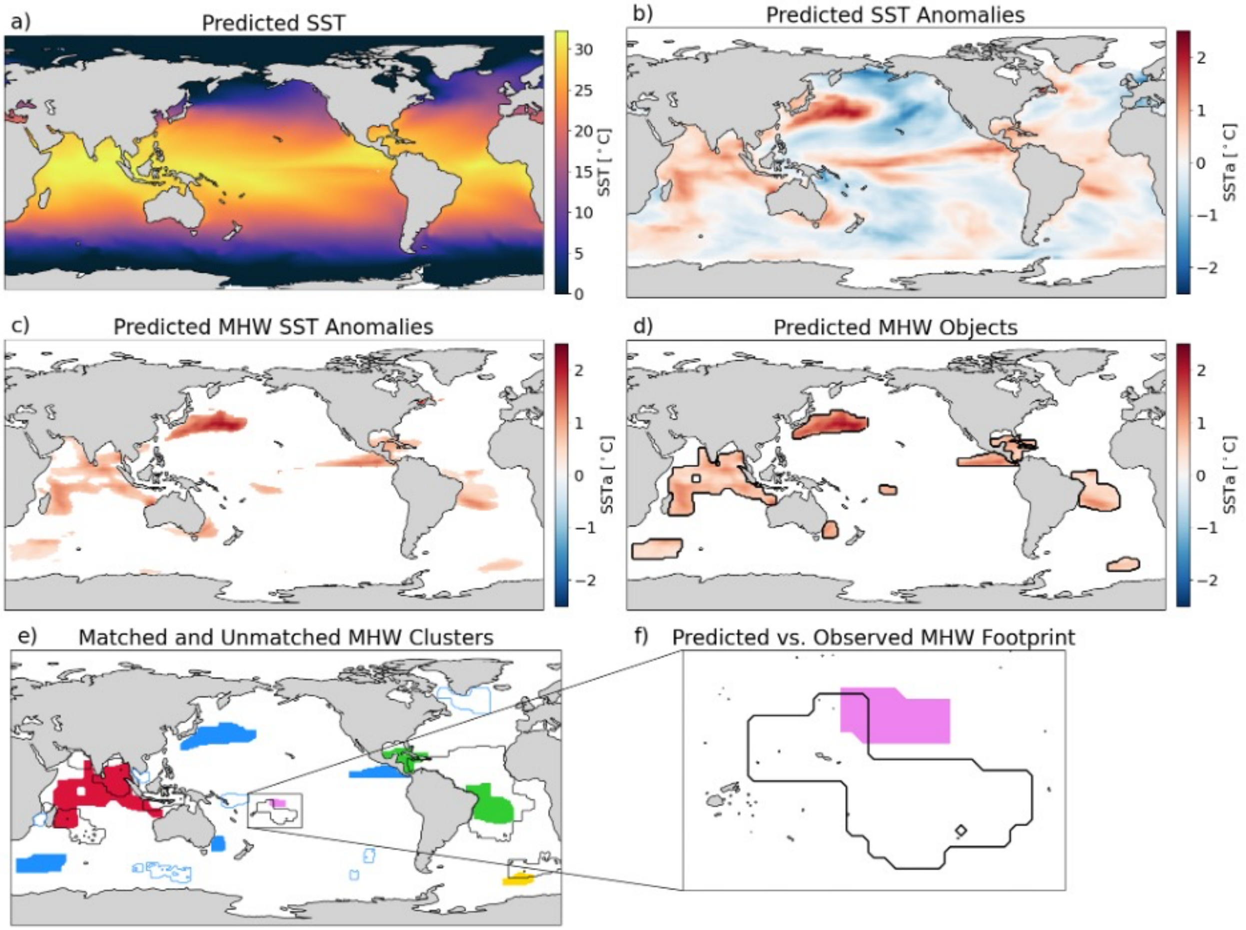


Figure 2.1. Identifying and matching MHW objects in an example forecast. In the first ensemble member prediction of May 2010 initialized on February 1, 2010 (3.5-month lead), we remove the climatology and trend from the (a) predicted SST to calculate (b) the anomalies. We apply the 90th percentile threshold to define (c) MHW point locations and use Ocetrac to define (d) the spatially connected predicted objects. (e) We match forecast clusters (red, pink, green, and yellow regions) with observed clusters (outlined contours) where an overlap exists. Where no overlap exists, forecast objects (blue regions) and observed objects (blue outlines) are unmatched. (f) For each matched forecast cluster, we consider the full forecast footprint (pink region) to be a correct prediction.

After defining MHWs at each ocean grid cell, we use Ocetrac, an MHW tracking tool that uses image processing techniques to smooth and label connected MHW grid cells (Scannell et al., 2024), to identify MHW objects from the MHW points in SMYLE and OISST (Fig. 2.1d; Text 2.S1).

2.2.3. Verifying Marine Heatwave Forecasts

Object-based forecast verification requires a process to match objects between the forecast field (SMYLE) and the verification, or observation, field (OISST) to compare objects. We employ the Method for Object-based Diagnostic Evaluation (MODE; Davis et al. 2006a,b) to perform object-based verification of MHW predictions. MODE has been used to analyze forecast skill of precipitation (Clark et al., 2014; L. Li et al., 2020), atmospheric rivers (DeHaan et al., 2021), drought (Abatan et al., 2018), cloud cover (Mittermaier & Bullock, 2013), and chlorophyll (Mittermaier et al., 2021).

MODE identifies, merges, and matches objects between the forecast and verification fields, and calculates attributes on simple objects, clusters, and cluster pairs (Text 2.S1). MODE defines simple objects as individual objects in either field, clusters as groups of objects in one field that are paired to a cluster in the other field, and cluster pairs as sets of matched clusters in each field. We use the SST anomalies within the MHW objects identified by Ocetrac as the input field for MODE (Fig. 2.1d) and match a forecast object (SMYLE) and an observed object (OISST) at each timestep if they overlap in space (Fig. 2.1e).

2.2.4. Evaluating Forecast Skill

We evaluate forecast skill by assessing how often forecast MHWs are matched to observed MHWs and by examining the similarity of matched MHW clusters. For pointwise forecasts, contingency tables are used to measure how often a forecast is a hit (forecast yes, observed yes), a false alarm (forecast yes, observed no), a miss (forecast no, observed yes), or a correct negative (forecast no, observed no). In the object-based framework, we relax the constraint that a forecast MHW must occur at the same location as the observed MHW. Instead, we define the whole forecast MHW cluster to be a hit when the forecast cluster overlaps with an observed cluster (Fig. 2.1f). Thus, we define locations within matched forecast clusters as hits, locations within unmatched forecast objects as false alarms, locations within observed objects or clusters but not within matched forecast clusters as misses, and locations with no forecast or observed objects as correct negatives.

In the example in Figure 2.1f, a pointwise contingency table would determine that the pink region within the black contour is a hit, the pink region outside the black contour is a false alarm, the contoured white region is a miss, and everywhere else is a correct negative. The object-based contingency table, however, defines the entire pink region, whether a point is within the black contour or not, as a hit because it belongs to a matched forecast object.

We use the contingency table statistics to evaluate two metrics: the False Alarm Ratio (FAR) and the deterministic limit (T_{DL}). FAR quantifies the fraction of predicted events that do not occur and is calculated separately for each lead month:

$$\mathrm{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}}. \tag{2.1}$$

Thus, higher values of FAR indicate weaker predictive skill. T_{DL} indicates the timescale at which MHWs can be reliably predicted. It is calculated as the latest lead time at which the number of hits (H) is greater than or equal to the number of misses and false alarms (X):

$$T_{DL} = \text{Lead[H=X]}. \tag{2.2}$$

We evaluate T_{DL} with both pointwise and object-based contingency table statistics.

These skill metrics convey how often SMYLE predicts MHWs but do not characterize how accurate the predicted MHWs are. We quantify this accuracy by calculating four pair attributes for each cluster pair: centroid distance, intersection over union, area ratio, and intensity ratio. The centroid distance of a cluster pair is the distance between the centroids of the forecast cluster and the observed cluster. The intersection over union is the overlapping area between the forecast cluster and the observed cluster divided by the area of the union of both clusters.

The area and intensity ratios are defined as the ratio of the areas or intensities of the forecast and observed MHW clusters belonging to a cluster pair, with the higher value being divided by the smaller value (Text 2.S2, Fig. 2.S1). For example, the area of the larger MHW cluster is divided by the area of the smaller MHW cluster regardless of whether the forecast MHW is larger or smaller than the observed MHW. The intensity ratio is similarly defined as the ratio of the median SST anomaly of the warmer MHW to the median SST anomaly of the cooler MHW. All statistics are calculated for the whole object cluster, not for individual objects within the cluster.

Perfect forecast skill is indicated by a centroid distance of zero (lower bound), an intersection over union of one (upper bound), an area ratio of one (lower bound), and an intensity ratio of one (lower bound). We average each attribute over every MHW cluster for each initialization month and lead time. Then, for each matched forecast MHW, we assign the value of its attributes to its spatial footprint and average over each initialization month and lead time to create spatial maps of each attribute. Mean values of ratio quantities are calculated as the geometric mean (Text 2.S2), and all lead-dependent mean quantities are averaged over objects, not grid cells, so they are not area-weighted. All skill metrics are evaluated against the skill of a random forecast (Text 2.S3). While Mogen et al. (2024) select the 97.5th percentile of skill scores from 1000 random forecasts, the computational expense of MODE limits our evaluation to a single random forecast. While this benchmark is less statistically robust, we expect that a larger sample size of random forecasts would yield qualitatively similar results.

2.3. Results

We evaluate the skill in predicting the presence of MHWs by evaluating the object-based FAR and by comparing the pointwise T_{DL} to the object-based T_{DL}. FAR is low at short lead times and increases globally with lead time, remaining lower (more skillful) than the random forecast FAR at most locations (Fig. 2.2a–d). At all lead times, FAR is lowest in the tropical Pacific Ocean, where the ENSO pattern of extreme SST anomalies is predicted better than the rest of the globe. Studies of pointwise MHW predictions also show higher skill in the tropical Pacific (Jacox et al., 2022; Mogen et al., 2024).

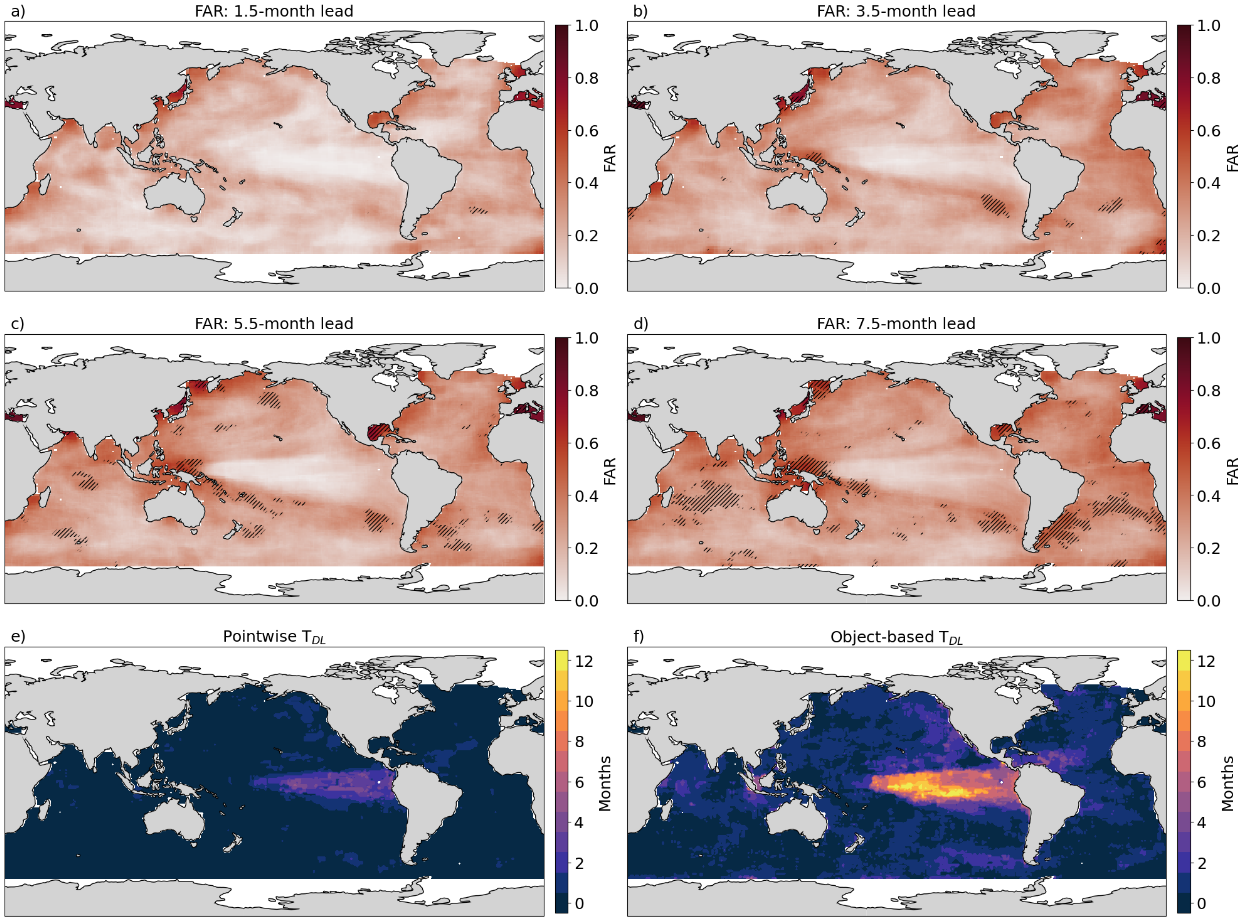


Figure 2.2. The false alarm ratio (FAR) and the deterministic limit (TDL). (a–d) The 1.5-, 3.5-, 5.5-, and 7.5-month lead object-based FAR demonstrate that FAR increases globally with lead time. Hatched regions indicate FAR values greater than the random forecast's FAR values. While the pointwise TDL (e) shows predictive skill past one month only in the tropical Pacific, the object-based TDL (f) indicates multi-month skill globally.

The pointwise T_{DL} (Fig. 2.2e) indicates that SMYLE reliably predicts pointwise MHWs only in the first lead month, except in the tropical Pacific. The object-based T_{DL} (Fig. 2.2f), however, shows that MHW events are predicted to occur near observed events for one to two seasons globally, and up to a year in the tropical Pacific.

We next quantify the errors in forecast MHWs' location, overlap, area, and intensity to evaluate how well the predicted MHWs represent the observed MHWs to which they are matched. The mean forecast errors for each skill metric have similar overall spatial patterns at 3.5 lead months (Figs 2.3b, d, f, h). The lowest forecast errors occur in the tropical Pacific Ocean, suggesting that SMYLE not only predicts the presence of MHWs in this region (Fig. 2.2), but also predicts their locations, areas, and intensities well. Skill values are only averaged over matched forecast clusters, so interpretation of object-based skill metrics must incorporate the FAR and T_{DL} metrics. Because the verification data include a limited sample size of observed MHWs, other regions with especially high or low skill may be influenced by individual events that are especially well or poorly predicted. Evaluating drivers of individual MHWs could help determine why certain events are predicted better than others.

The centroid distance, intersection over union, and area ratio depend primarily on lead time (Fig. 2.3a, c, e). The intensity ratio depends on lead time for the first few months, after which it depends on the target month, indicating the importance of seasonal variability (Fig. 2.3g). The magnitude of each metric varies for MHWs of different sizes (Fig. 2.S2) and depends slightly on the choice of Ocetrac radius (Fig. 2.S3). The average matched forecast MHW has a centroid within 1400 km of the observed MHW's centroid, a 15–20 percent overlap with the observed MHW, an area less than 3 times smaller or larger than the observed MHW's area, and a median SST anomaly 1.25–1.35 times warmer or cooler than the observed MHW's median SST anomaly (Fig. 2.3a, c, e. g). All metrics exhibit high skill at short lead times and exceed the skill of random forecasts for the first 12 lead months.

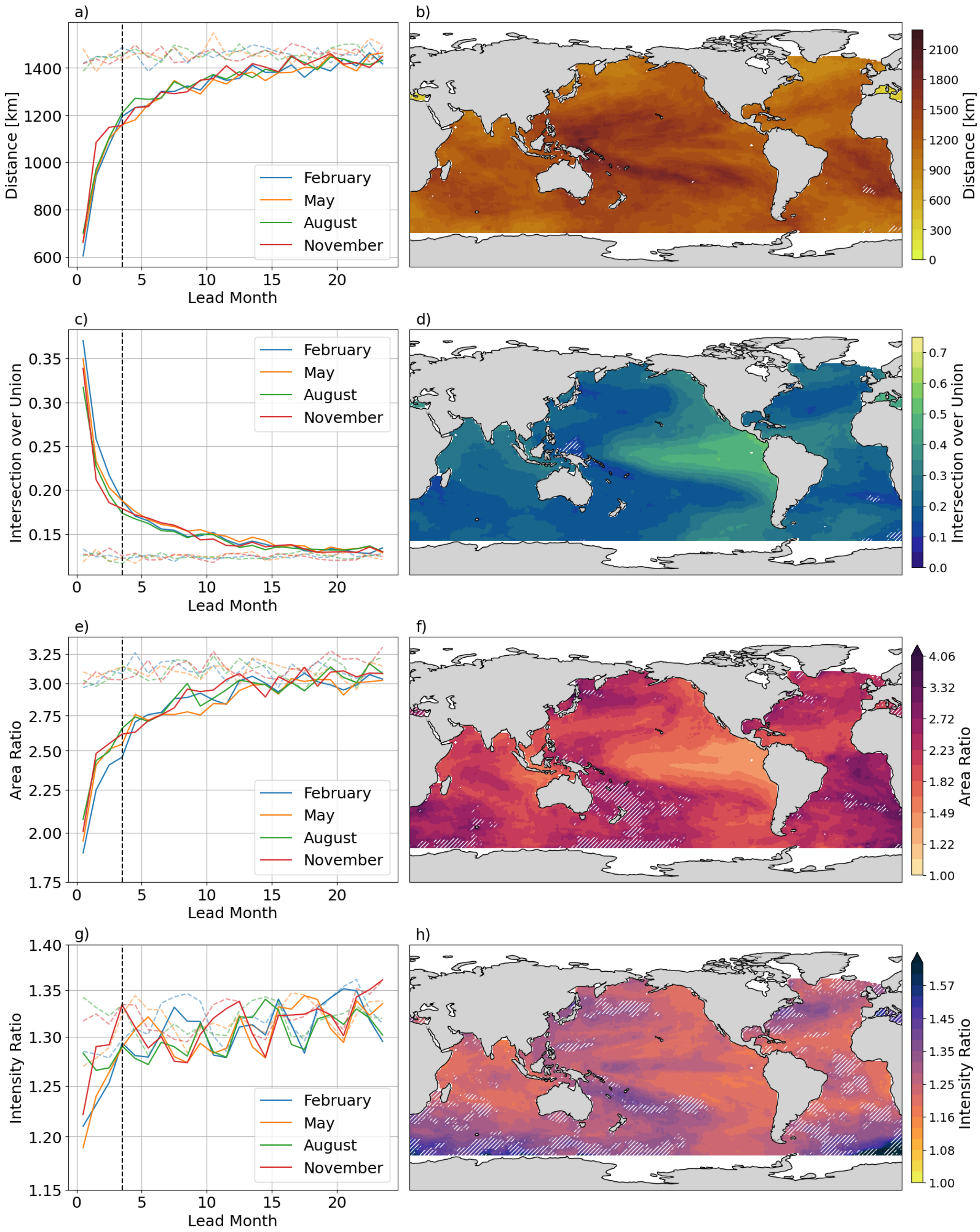


Figure 3. Mean forecast errors in centroid distance (a, b), intersection over union (c, d), area ratio (e, f), and intensity ratio (g, h). The left column (a, c, e, g) shows the mean error of each

metric averaged over all matched cluster pairs for each initialization month (colored lines) and lead time, while the right column (b, d, f, h) shows the mean errors for the 3.5-month lead predictions over all objects and initialization months, where lighter hues signify better skill. Dashed lines in the left column show the mean forecast errors of a random forecast, and hatched regions in the right column indicate skill worse than the random forecast.

We examine the intensity ratio by lead and target month to quantify its target-month dependence and seasonal variability (Fig. 2.4a). We then calculate the intensity ratio anomaly factor (Text 2.S2), the factor by which the intensity ratio is greater than or less than the mean intensity ratio for a given lead month (Fig. 2.4b). Anomaly factors less than one (blue) mean the intensity ratio is more skillful and anomaly factors greater than one (red) mean the intensity ratio less skillful.

The intensity ratio for a given lead time is least skillful when the target month is between August and October, and most skillful when the target month is December or January. At lead times less than 4.5 months, initialization month plays an important role (Figs. 2.3g, 2.4b), after which the intensity ratio depends on the target month. The random forecast exhibits the same target-month improvements as the SMYLE predictions (crosses in Fig. 2.4b), indicating that the seasonal skill improvements are likely due to SMYLE's prediction skill of the overall SST distribution. Nevertheless, SMYLE's predictions of MHW intensity are more skillful than the random forecast's for the first 12 months (Fig. 2.4a). The target-month intensity ratio improvements are dominated by seasonal changes in the tropical and North Pacific, which display better skill in the boreal winter (when ENSO peaks) and worse skill in August–October (Fig. 2.S4). This spatial skill pattern is consistent with previous results on ENSO-related prediction skill (e.g., Jacox et al. 2022; Shin and Newman 2021).

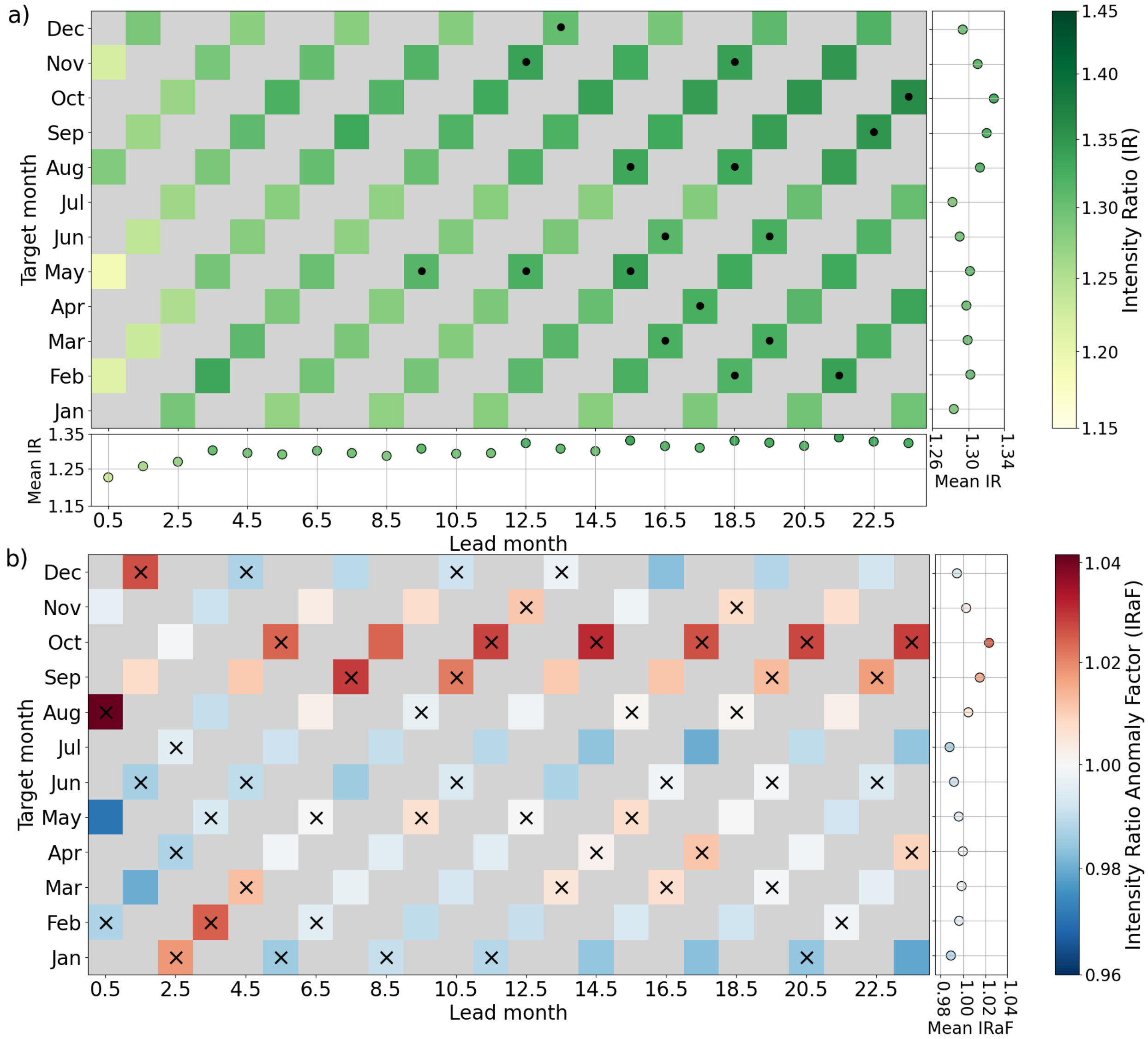


Figure 4. Dependence of the intensity ratio on lead month and target month. (a) The mean intensity ratio for each target month and lead month, with averages for each lead month shown along the bottom x-axis and averages for each target month shown along the right y-axis. Dots indicate intensity ratios greater than the random forecast's. (b) The same as in (a) but showing the intensity ratio anomaly factor relative to the lead-dependent mean. A negative anomaly factor (blue) represents better intensity ratios (closer to one) and a positive anomaly factor (red) represents worse intensity ratios (further from one). Crosses (X) indicate anomaly factors greater than the random forecast's. Panel (b) has no bottom scatter plot because the lead-dependent

mean intensity ratio anomaly factor is one at all lead months. Gray boxes in both panels are lead/target combinations with no data (for example, there is no 0.5-month prediction for June).

2.4. Conclusions

We demonstrate a novel evaluation method for verifying MHW predictions and quantify the skill in predicting the occurrence of MHWs and their associated attributes. By combining methods from both MHW detection and weather forecast verification we evaluate how well CESM2 SMYLE simulates and predicts MHWs. The results indicate that SMYLE forecasts effectively represent observed MHWs from OISST, extending previous evaluations of MHW predictions (De Boisséson & Balmaseda, 2024; Jacox et al., 2022) by predicting MHWs as objects instead of points. The method introduced here avoids incurring double penalties for the misplacement of predicted events and provides information about the types of errors that occur in MHW forecasts.

SMYLE demonstrates global seasonal to annual MHW prediction skill (Fig. 2.2). The error metrics examined here—centroid distance, intersection over union, area ratio, and intensity ratio—depend primarily on lead time and outperform random forecast skill for at least 12 lead months. Predictions of MHWs in December and January best predict MHW intensity, while predictions of MHW intensity between August and October perform worse, likely due to ENSO-related skill in the tropical and North Pacific.

A prediction with low error refers to a predicted MHW that has similar attributes to the observed MHW it is matched with. MHW predictions in locations and lead months with low error but high FAR values still exhibit weak predictive skill overall because the predicted MHW is unlikely to occur in the first place. Thus, forecast error metrics (Fig. 2.3) must be interpreted along with FAR (Fig. 2.2).

Both MODE and Ocetrac are designed to run on evenly spaced grids, unlike the native CESM grid or the interpolated 1° x 1° grid used here. This makes the presented approach problematic at high latitudes, so we constrained our analysis to latitudes less than 65°. MODE was also designed for regional grids instead of global grids, so it does not allow for periodic boundary conditions. This can cause some MHW events that span the southern coast of Africa or the Mediterranean Sea to be split into two events, which leads to more misses and less accurate predictions at the edges of the map (Figs. 2.2a, 2.3). We choose this longitude edge to minimize the impact of this aspect of MODE. Finally, we use permissive criteria for matching forecast and observed events: we require only that the objects overlap in space but do not require the objects to have similar areas or intensities. Moreover, while we allow for the separation of MHWs in space, we evaluate only the temporal co-occurrence of MHWs following other recent studies (e.g., Jacox et al., 2022; Mogen et al., 2024). A double penalty in time, when an MHW is predicted in the right place but the wrong time, could be addressed with temporal aggregation or object tracking (Clark et al., 2014).

Spatial forecasts have several potential applications for community use. Object-based definitions of MHWs and other ocean extremes could be used to improve our process-based understanding of the drivers and the predictability of extreme and compound events (Gruber et al., 2021; Mogen et al., 2024). Object-based seasonal forecasts may also be useful for open-ocean predictions relevant for dynamic ocean management (Maxwell et al., 2015) and living marine resource management, including monitoring and closures and annual catch limits (Tommasi et al., 2017). Spatial uncertainties and errors like those presented here must be incorporated for these operational applications. Spillman et al. (2025) identify the need for useful and useable MHW forecasts that present skillful information at relevant scales and with interpretable metrics.

Object-based predictions contribute to this goal by providing novel and intuitive ways to communicate MHW forecasts both to end users and to a general audience.

2.5. Supplementary Information

2.5.1. Text 2.S1. Extended Methods: Ocetrac and The Method for Object-based Diagnostic Evaluation

We use two software packages, Ocetrac and MODE, to identify MHW objects and match them between the forecast (SMYLE) and observation (SMYLE FOSI) datasets. Here, we explain our choices for the user-defined parameters for each method.

Ocetrac uses image processing techniques to smooth and label connected MHW grid cells in space and time. User-defined parameters include the convolutional radius, which specifies the distance over which Ocetrac applies the smoothing filter, and the minimum area threshold, which allows the user to discard a fraction of the objects with the smallest area. Scannell et al. (2024) includes a robust sensitivity analysis to determine how these parameters affect MHW detection. We set the minimum area threshold to 0 so that no objects are discarded and set the radius to 3 grid cells (3°).

The choice of smoothing radius affects the number and shape of the identified MHW objects as well as the magnitudes of the forecast skill metrics. Smaller radii cause MODE to identify more simple objects at each time step which exponentially increases computational cost, while larger radii can connect sparse grid cells and distort the shape and size of MHW events (Fig. 2.S5). SMYLE and OISST have different native resolutions and capture MHWs at different scales (even after regridding), so the smoothing radius also enables the comparison of objects of similar sizes. The centroid distance and intensity ratio are more skillful for smaller radii, while the intersection over union and area ratio are more skillful for larger radii (Fig. 2.S3). A radius of 3 grid cells is a middle ground that balances the tradeoffs demonstrated here. Future users, however, can adjust the radius to best fit individual use cases.

MODE was developed for applications like precipitation forecasts which use constant thresholds (e.g., integrated precipitation > 5 mm). Because MHWs are defined relative to a local, seasonally varying threshold, MODE's object identification process is incompatible with MHWs. Instead, we use Ocetrac to define MHW objects from the SST field and use the SST anomalies within the MHW objects identified by Ocetrac as the input field for MODE. Since these objects have already been spatially connected, we set the convolutional radius in MODE to 1 (no smoothing) and the threshold to 0 (identifies any anomaly over 0°C). As a result, the objects identified by MODE are identical to the input data, which are the objects identified by Ocetrac (Fig. 2.1d).

MODE identifies, merges, and matches objects between the forecast and observed fields, and calculates attributes on simple objects, object clusters, and cluster pairs, defined for each timestep separately. While Ocetrac labels objects connected in space and time, we do not consider the temporal connectivity of MHW events in this study.

After identifying simple objects in the forecast and verification fields, MODE calculates attributes for each simple object and for every pair of simple objects between the two fields. A full list of the single and pair attributes can be found in the NCAR technical note on MODE (Bullock et al., 2016). Objects are then matched between the forecast and verification fields based on a fuzzy logic framework using the pair attributes. The user defines how important each pair attribute is in matching objects by defining weights, and MODE uses these weights to calculate an interest score—a measure of how similar two objects are—between zero and one for each possible pair of forecast and observed objects. If that score is above a user-defined threshold score (commonly set to 0.7), then the objects are considered a match. In this framework, more than one simple object in each field can be matched to a simple object in the other, creating object clusters in either field. As a result, MODE outputs a dataset of matched forecast clusters, matched observed clusters, unmatched forecast objects, and unmatched observed objects. After matching, MODE recalculates all individual and pair attributes for each matched cluster pair.

2.5.2. Text 2.S2. Extended Methods: Intensity Ratio Averages and Anomalies

The area and intensity ratios are defined in the main text as the ratio of the areas or intensities of the forecast and observed MHW clusters belonging to a matched cluster pair, with the larger value being divided by the smaller value. Here, we define this ratio as the absolute ratio, as it is the absolute factor by which the area or intensity is incorrect. We also define the relative ratio, the factor by which the forecast cluster's area or intensity is greater than the observed cluster's area or intensity. Figure 2.S1 demonstrates the difference between these ratios. We choose to use the absolute ratios in the main text because the mean relative area ratio is not strongly dependent on lead time: the forecast object's area is not usually greater or usually less than the observed object's area on average.

For all calculations performed on the ratio quantities, we use the logarithm of the ratio quantity, or the log ratio. First, this simplifies defining the absolute and relative ratios—the absolute log ratio (higher value divided by smaller value) is the absolute value of the relative log ratio, and the absolute ratio itself is the exponential of the absolute log ratio:

$$\text{Intensity Ratio} = \exp\left|\log\frac{\text{Intensity}_{\text{FCST}}}{\text{Intensity}_{\text{OBS}}}\right|. \qquad (2.\text{S}1)$$

Second, to calculate the mean of a set of ratios, we calculate the geometric mean instead of the arithmetic mean. Practically, the geometric mean is the arithmetic mean of the log ratios:

$$\text{Mean Intensity Ratio} = \exp\text{mean}\left(\left|\log\frac{\text{Intensity}_{\text{FCST}}}{\text{Intensity}_{\text{OBS}}}\right|\right). \qquad (2.\text{S}2)$$

Equations 2.S1 and 2.S2 show the absolute ratios—the relative ratios would be calculated similarly but without taking the absolute value of the log ratios.

The importance of using the geometric mean for these ratios can be illustrated with an example using the relative area ratio: if one predicted MHW is twice the size of the observed MHW and another MHW is half the size of the observed MHW, then the geometric mean is one, indicating that on average the predicted MHW is neither larger nor smaller than the observed MHW. While this example does not apply exactly to the absolute ratios, the need to use the geometric mean remains. In addition to calculating the geometric mean, we present plots of mean ratios with log-scaled axes and color maps so that the reader's perception matches the values' multiplicative differences as opposed to additive differences.

To calculate the intensity ratio anomaly factor used in Fig. 2.4b, we start with the mean log intensity ratio for each lead month and target month. We then calculate the lead-dependent mean log intensity ratio (the log of the values in the bottom panel of Fig. 2.4a) and subtract that from the mean log intensity ratios (the log of the values in the main grid in Fig. 2.4a). The values in Fig. 2.4b are then calculated as the exponential of these anomalies. This is equivalent to dividing the mean intensity ratios (main panel of Fig. 2.4a) by the lead-dependent mean intensity ratios (bottom panel of Fig. 2.4a):

$$\text{Intensity Ratio Anomaly Factor}_{\text{Lead,Target}} = \frac{\text{Mean Intensity Ratio}_{\text{Lead,Target}}}{\text{Mean Intensity Ratio}_{\text{Lead}}}, \qquad (2.\text{S}3)$$

where the numerator is calculated as in Eq. 2.S2 for each lead and target month, and the denominator is the calculated as in Eq. 2.S2 for each lead month.

2.5.3. Text 2.S3. Extended Methods: Random Forecast Benchmark

To provide a benchmark against which to evaluate the skill from the SMYLE forecasts, we generate random forecasts by selecting random initialization years and ensemble members with replacement for each initialization month and lead time. We then match each individual forecast field to a random observation timestep of the correct target month. This creates a dataset of equal size as the SMYLE forecast dataset, with the same number of forecasts for each initialization month and lead month. We perform the same postprocessing to calculate the false alarm ratio (Fig. 2.2) and all object-based forecast skill metrics by location, initialization month, and lead time (Figs. 2.3, 2.4). We consider results to be better than random if they are more skillful than the results from this random forecast.

Mogen et al. (2024) use a similar approach to assign statistical significance to their results, creating a set of 1000 random forecast datasets and selecting the 97.5th percentile of skill scores as the significance benchmark. We were limited to evaluating a single random forecast dataset instead of 1000 due to the computational cost of running MODE. While our benchmark is less statistically robust, it still provides a good standard for our results.

2.5.4. Supplementary Figures 2.S1–2.S5

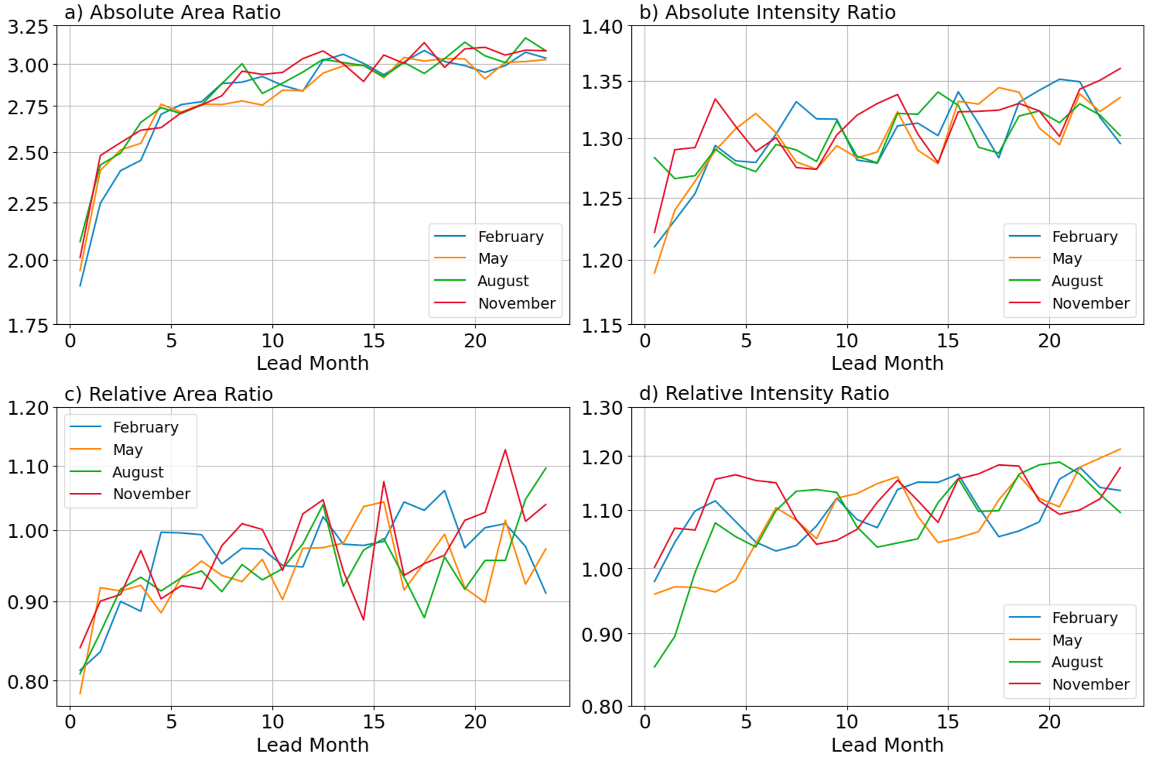


Figure 2.S1. Comparing two methods to calculate the area ratio and the intensity ratio. Panels a and b show the *absolute* area and intensity ratios from Fig. 2.3e and Fig. 2.3g, calculated as the ratio of the bigger/more intense MHW to the smaller/less intense MHW, and panels c and d show the *relative* ratios of the forecast MHW's area/intensity divided by the associated observed MHW's area/intensity.

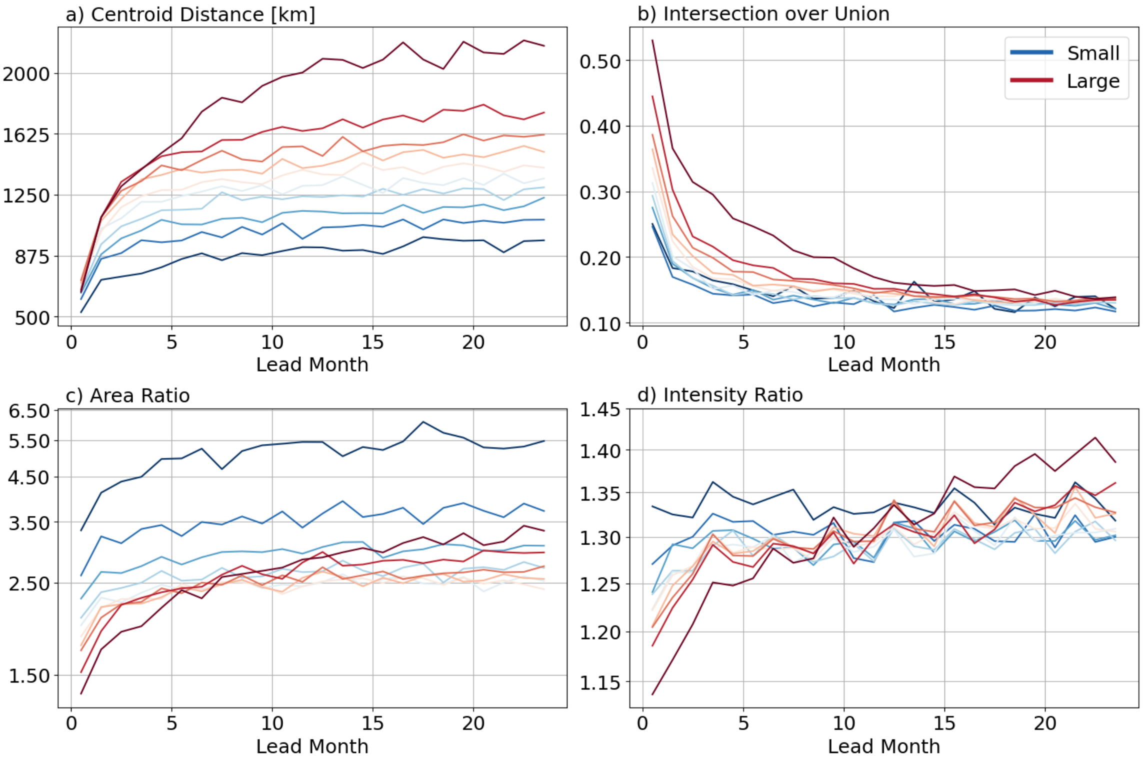


Figure 2.S2. Mean forecast errors by forecast MHW size. Matched forecast clusters are decomposed into deciles by area for each lead month. Plots show the mean error of each metric averaged over all matched cluster pairs for each area decile (colored lines) and lead time. Blue lines indicate smaller MHWs and red lines indicate larger MHWs.

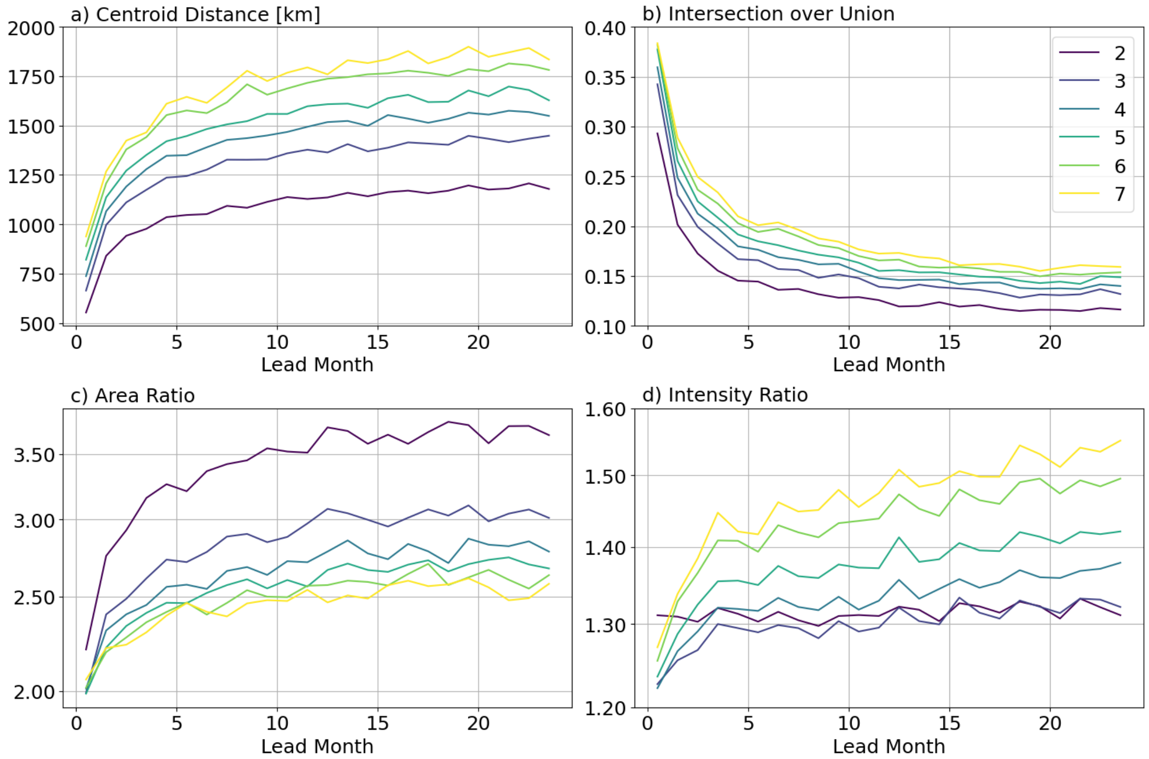


Figure 2.S3. Mean forecast errors by Ocetrac radius. We define MHW objects in SMYLE using Ocetrac radii of 2–7 grid cells and run MODE on predictions from the first 10 ensemble members for each radius. We do not calculate metrics for a radius of 1 grid cell because of computational constraints. Plots show the mean error of each metric averaged over all matched cluster pairs for each radius (colored lines; larger radii have lighter colors) and lead time.

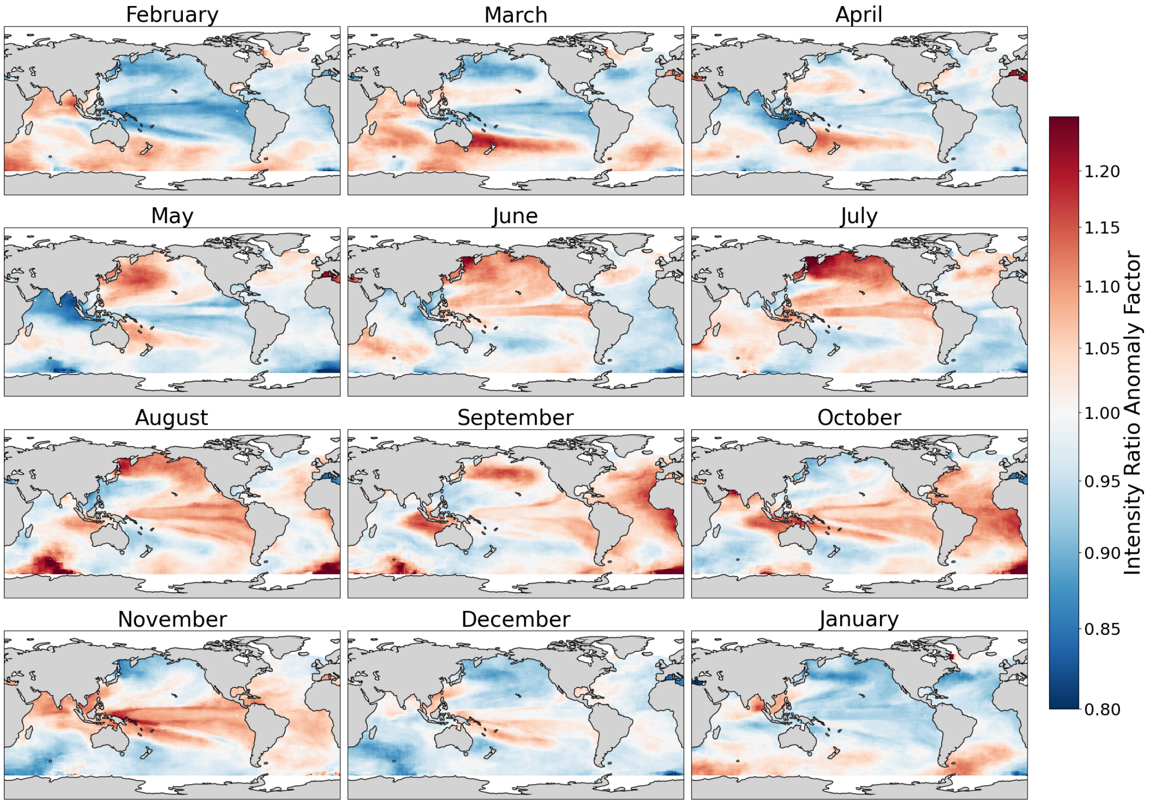
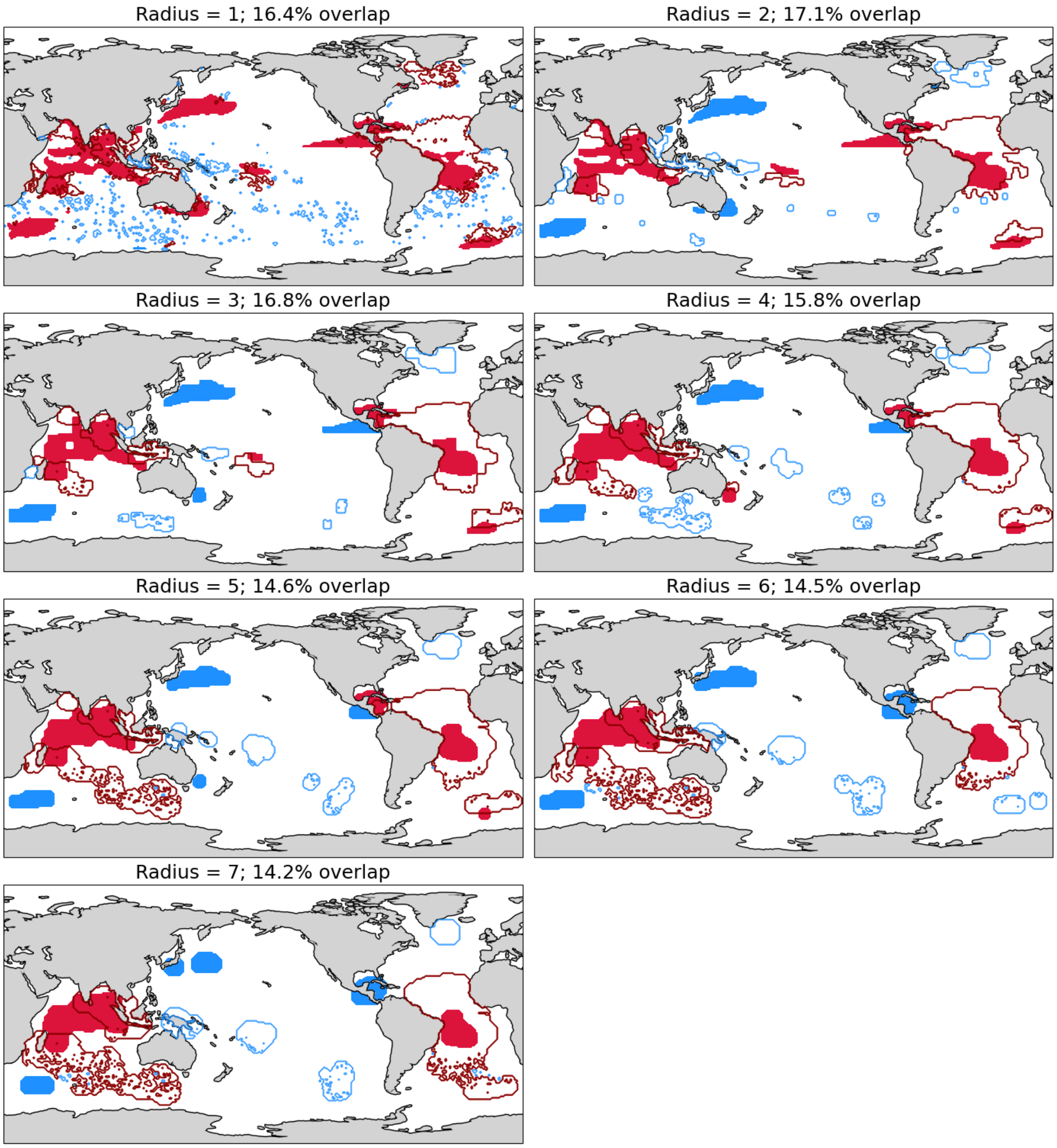


Figure 2.S4. The lead-dependent intensity ratio anomaly factor for each target month, where anomaly factors less than one (blue) represent the multiplicative factor by which the intensity ratio is closer to one (more skillful), and anomaly factors greater than one (red) represent the factor by which the intensity ratio is further from one (less skillful). The left column is the mean anomaly over lead months 0.5, 3.5, 6.5, etc., the middle column is the mean anomaly over lead months 1.5, 4.5, 7.5, etc., and the right column is the mean anomaly over lead months 2.5, 5.5, 8.5, etc.



Radius = 1; 16.4% overlap

Radius = 2; 17.1% overlap

Radius = 3; 16.8% overlap

Radius = 4; 15.8% overlap

Radius = 5; 14.6% overlap

Radius = 6; 14.5% overlap

Radius = 7; 14.2% overlap

Figure 2.S5. Matched (red) and unmatched (blue) forecast (solid regions) and observed (outlined regions) MHW clusters for different Ocetrac radii (in grid cell units) in the first ensemble member prediction of May 2010 initialized on February 1, 2010 (3.5-month lead). A radius of 1 indicates no smoothing with higher radii corresponding to larger smoothing scales. The percent overlap in each subplot title is the percent of total MHW grid cells in both the forecast field and the observation field that overlap (similar to the intersection over union metric).

3. Seasonal-to-Decadal Prediction of Regional Arctic Sea Ice and Ocean Heat Transport

3.1. Introduction

Arctic sea ice extent has declined over the past several decades with global warming (Stroeve et al., 2007), though a slowdown in sea ice loss has occurred since 2007 (England et al., 2025; Stern, 2025). Prior to the slowdown, some climate predictions indicated that ice-free summers would occur before 2020 (Maslowski et al., 2012), while others correctly predicted the continued reduction in sea ice loss (Yeager et al., 2015). While Arctic sea ice loss has slowed since 2007, long-term projections indicate its continued reduction into the future (Årthun et al., 2021). The potential future decline of Arctic sea ice would lead to more area and time with ice-free conditions and more access to the Arctic (Huntington et al., 2022). Understanding the prediction skill of Arctic conditions and the sources of predictability that allow us to trust and improve the model predictions is therefore becoming increasingly important.

Arctic sea ice variability is driven by various atmospheric and oceanic processes. The ocean affects sea ice primarily through circulation and heat transport in the Atlantic (Frankignoul et al., 2024; Oldenburg et al., 2024), while the atmosphere influences sea ice through atmospheric circulation and winds (Ding et al., 2017, 2022; Oldenburg et al., 2024; D.-S. R. Park et al., 2015; Roach & Blanchard-Wrigglesworth, 2022). Most of the ocean's influence on the Arctic comes from water flowing from the Pacific Ocean through the Bering Strait and from the Atlantic Ocean through the Fram Strait and the Barents Sea Opening (Fig. 3.1). Bering Strait volume and heat transports into the Chukchi Sea have increased over recent decades, bringing more heat into the Arctic primarily in the summer and fall (Woodgate, 2018; Woodgate et al., 2012). Atlantic-

Arctic inflow is dominated by ocean heat transport variability through the Barents Sea Opening, which has been linked to sea ice variability in the Barents Sea (Årthun et al., 2012). The same atmospheric and oceanic processes that influence Arctic sea ice variability can act as sources of predictability (Virginie Guemas et al., 2016). Furthermore, seasonal persistence of sea surface temperature (SST) anomalies and year-to-year persistence of sea ice thickness anomalies can serve as mechanisms for the reemergence of sea ice area anomalies (Edward Blanchard-Wrigglesworth, Armour, et al., 2011). Initialized predictions demonstrate that this persistence is a source of predictability and suggest that accurate initializations of sea ice thickness anomalies improve predictions (J. J. Day, Hawkins, et al., 2014).

Arctic sea ice predictability depends strongly on the season and region of prediction (W. Cheng et al., 2016). The seasonality and initial conditions affect prediction skill of seasonal sea ice extent, sea ice volume, and sea ice loss across dynamical prediction systems (Edward Blanchard-Wrigglesworth, Bitz, et al., 2011; Bushuk et al., 2022; J. J. Day, Tietsche, et al., 2014; Yeager et al., 2015). The influence of these processes on sea ice, however, vary substantially by Arctic region, with stark differences between the Atlantic-Arctic and the Pacific-Arctic sectors (Bushuk et al., 2022, 2024).

As the Arctic changes in a warming climate, Arctic sea ice predictability may change as well (Holland et al., 2019). Arctic Ocean changes have already led to stronger ocean-sea-ice interactions in Arctic regional seas (Bianco et al., 2024). Additionally, sources of ocean volume and heat transport are evolving (Årthun et al., 2025), and with them evolves the influence from Atlantic ocean heat transport (OHT) and Pacific OHT on Arctic sea ice (K. Cheng et al., 2025; Dörr et al., 2024). Characterizing sources of Arctic sea ice predictability, such as ocean heat transport variability, has the potential to improve future Arctic predictions. Here, we evaluate predictions of ocean heat transport and Arctic sea ice to evaluate the current skill in predicting seasonal-to-decadal Arctic change and to explore sources of predictability in the ocean–sea-ice system.

3.2. Data and Methods

3.2.1. The CESM2 Seasonal-to-Multiyear Large Ensemble and Decadal Prediction Dataset

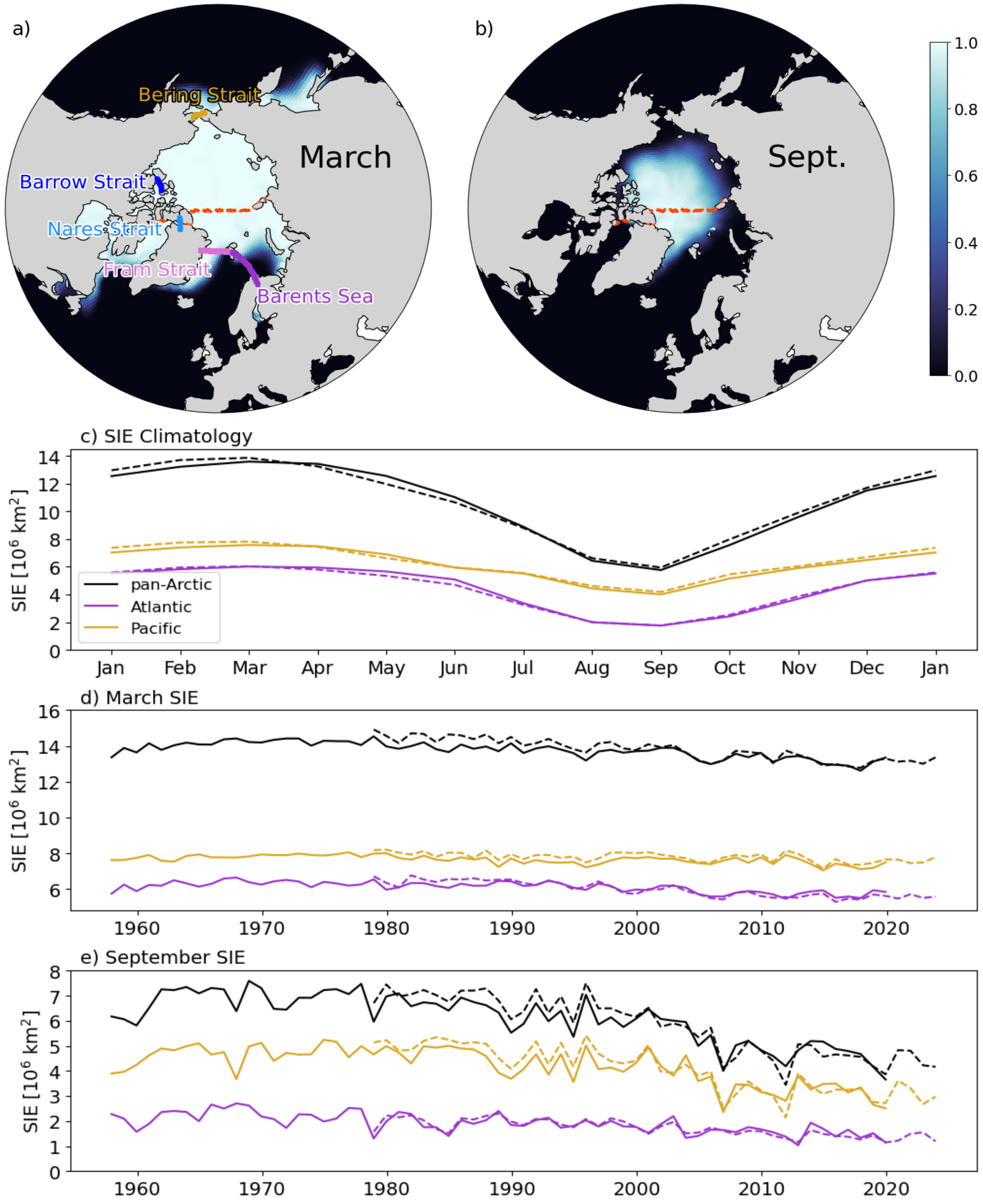
We analyze predictions from the Seasonal-to-Multiyear Large Ensemble (SMYLE) and its decadal extension, the CESM2 Decadal Prediction dataset (DP) (Yeager et al., 2022). SMYLE and DP are hindcast simulations run with the Community Earth Systems Model (CESM2) with a nominal 1° horizontal resolution for models of each component (Danabasoglu et al., 2020). Ocean and sea ice initial conditions for SMYLE come from a forced-ocean-sea-ice model (FOSI), which is run from 1958–2020 using the Parallel Ocean Program version 2 (POP2) and the Community Ice CodE 5.1.2 (CICE5). SMYLE FOSI is forced by historical atmospheric conditions from the Japanese 55-year Reanalysis (JRA-55-do; Tsujino et al., 2018). Atmospheric and land initial conditions for SMYLE also come from JRA-55-do.

SMYLE consists of 24-month long forecasts initialized on the first of every February, May, August, and November from 1970-2019. Each forecast has 20 ensemble members. The November-initialized predictions from SMYLE were then extended from 24 lead months to 122 lead months to comprise DP. Each 10-year prediction again contains 20 ensemble members. We define the initialization month as the month in which the forecast was initialized, and the lead month/year as the time since initialization, with lead month/year 0 being the same as the initialization month/year.

3.2.2. Arctic Sea Ice

To address the predictability of Arctic sea ice, we examine local sea ice concentration (SIC) and integrated sea ice extent (SIE) and sea ice volume (SIV). We compare the forecasts from SMYLE and DP against verification data from FOSI for all fields. Because all ocean and sea ice fields come from POP2 and CICE5, which are calculated on the same spatial grid, we can directly compare gridded and integrated quantities between the prediction and verification data. We also compare SIC and SIE predictions to monthly observations from the National Snow and Ice Data Center (NSIDC; Walt Meier et al., 2024). The NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 5 provides monthly SIC observations from November 1978 through December 2024 on a 25 km x 25 km grid. We use bilinear interpolation to regrid these data to the POP2 grid to enable direct comparison between observed and modeled SIC and SIE.

The SIE is the sum of the areas of all grid cells with SIC greater than or equal to 15% and the SIV is the integrated sea ice thickness over all grid cells. We calculate SIE and SIV over the pan-Arctic region, the Atlantic-Arctic region, and the Pacific-Arctic region (Figs. 3.1a–b and 3.A1). See Appendix 3.A and Figure 3.A1 for the region definitions based on the NSIDC region mask.



c) SIE Climatology

d) March SIE

e) September SIE

Figure 3.1. Sea ice concentration (SIC) and extent (SIE). Climatological SIC from FOSI in March (a) and September (b), the seasonal cycle of SIE in each region (c), and timeseries of March (d) and September (e) SIE. Arctic OHT gateways are indicated with solid lines (a), and

the Atlantic-Arctic and Pacific-Arctic regions are separated by the dashed red line (a, b). Panels c–e, show sea ice data from FOSI (solid) and NSIDC observations (dashed).

The Pacific region comprises most of the Arctic sea ice in all months of the year, making up more of the pan-Arctic sea ice in the summer than in the winter (Fig. 3.1c–3.1e). The Pacific SIE also reflects the downward trend in pan-Arctic September SIE more than the Atlantic SIE (Fig. 3.1e). FOSI agrees well with the observed magnitude of SIE in all seasons and across both Arctic regions. Notably, though, FOSI exhibits a slightly weaker negative trend than the NSIDC observations do (Fig. 3.1d–e).

3.2.3. Arctic Ocean Heat Transport

Heat transport into the Arctic has been connected dynamically to sea ice variability in observations and climate models (Lenetsky et al., 2021; Woodgate et al., 2010). Studies evaluating the role of OHT on sea ice predictability often calculate heat transport across a latitude level (e.g., Li et al. 2024). Here, we replicate the method from Oldenburg et al. (2024) to calculate OHT across five Arctic gateways: the Fram Strait, the Barents Sea, the Bering Strait, the Nares Strait, and the Barrow Strait (Fig. 3.1a). These gateways create a closed contour around the Arctic so that the sum of OHT through all gateways is equal to the total Arctic OHT.

We calculate OHT through each section as the Eulerian-mean ocean heat transport (Oldenburg et al., 2024):

$$\mathrm{OHT}(t) = \rho c_p \int_{x1}^{x2} \int_{z_{\mathrm{bot}}}^{0} \bar{v}(\bar{\theta} - \theta_{ref})\, dz\, dx\,, \tag{3.1}$$

where ρ is the density of seawater, c_p is the heat capacity of seawater, $\bar{v}$ is the monthly velocity, $\bar{\theta}$ is the monthly potential temperature and θ_{ref} is the reference temperature. Values of ρ and c_p are taken from Oldenburg et al. (2024). We integrate $\bar{v}\bar{\theta}$ over the horizontal coordinate x between x_1 and x_2, which represent the spatial boundaries of each section, and over the vertical coordinate z between the surface and z_{bot}, the bottom depth in the ocean.

While Oldenburg et al. (2024) calculate the total mean OHT as the sum of the Eulerian-mean OHT and sub-monthly OHT, we focus only on the Eulerian-mean OHT, since we do not expect sub-monthly OHT to contribute substantially to seasonal-to-decadal prediction skill. To calculate OHT across each Arctic gateway, we use the Physical Analysis of the Gridded Ocean (PAGO) package (Deshayes et al., 2014; Oldenburg et al., 2024). Following similar studies of Arctic OHT, we set the reference temperature for all OHT calculations to 0°C (Dörr et al., 2021; Oldenburg et al., 2024).

After calculating each gateway's monthly OHT, we calculate Atlantic-Arctic OHT as the sum of the OHT through the Fram Strait and Barents Sea, and Pacific-Arctic OHT as the OHT through the Bering Strait. OHT through the Nares Strait and the Barrow Strait is included in the total OHT but not the regional OHT. The seasonal cycle of total, Atlantic-Arctic, and Pacific-Arctic OHT demonstrate the varying strengths of Atlantic and Pacific heat flows into the Arctic Ocean, and time series of interannual OHT show the long-term trend associated with Arctic warming (Fig. 3.2). Negative Pacific-Arctic OHT values are caused by the volume transport of water below 0 °C, at or near the freezing point of Bering Sea seawater (Woodgate, 2018; Woodgate et al., 2005). These heat transport anomalies are thus related primarily to anomalous volume transport and not to anomalous temperatures. Previous studies suggest that CESM2 simulations of Arctic OHT through major pathways agree with observed quantities cite (Dörr et al., 2021; Oldenburg et al., 2024).

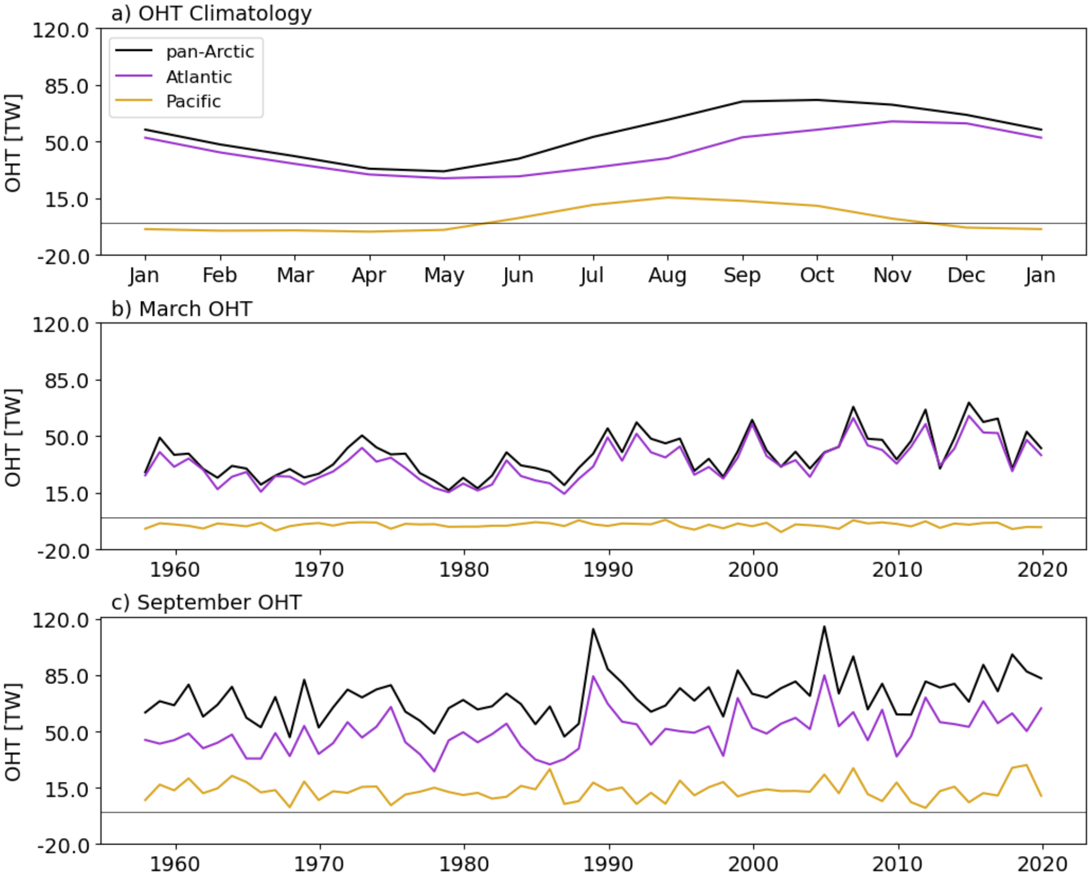


Figure 3.2. The seasonal cycle of OHT in FOSI in each region (a) and timeseries of March (d) and September (e) OHT. In each panel, total Arctic OHT is in black, Atlantic-Arctic OHT is in purple, and Pacific-Arctic OHT is in gold.

We compare OHT forecasts from SMYLE and DP to FOSI, as we do not have reliable observations of monthly OHT. While we could use a reanalysis product such as ORAS5 for monthly estimates of OHT, Dörr et al. (2024) find that intermodel spread in OHT and sea ice co-variability depends on model representation of Atlantic inflow and Arctic ocean volume transport. Evaluating predictions against verification data calculated from the same climate model, as we do here, reduces the effect of model biases on our results.

3.2.4. Forecast Skill Evaluation

Forecast skill is evaluated by comparing anomalies in SMYLE and DP to anomalies in FOSI and observations. Anomalies in FOSI and observations are calculated by removing the monthly climatology and linear trend from all months of data. Both the climatology and the trend are based on the 1979–2019 climatology, as this is the longest common period of reference for all data.

For all variables in SMYLE, we calculate anomalies by subtracting the lead-dependent ensemble mean and the lead-dependent linear trend for each initialization month separately, again relative to data between 1979–2019 (Meehl et al., 2022). To calculate annual mean predictions in DP, which is initialized each November, we calculate the lead-time mean of predictions in the target months January to December. Thus, lead year 1 corresponds to the mean of lead months 2–13, lead year 2 corresponds to the mean of lead months 14–25, etc. From these annual means, we calculate annual anomalies by removing the lead-dependent 1979–2019 climatology. We do not remove the linear trend from the DP predictions nor from the verification data to which we compare DP. We evaluate prediction skill for annual mean values, for winter (January, February, March; JFM) and summer (July, August, September; JAS) mean values, and for monthly mean values.

Preliminary analysis of the skill in predicting seasonal pan-Arctic SIE and SIV was performed in the paper documenting the SMYLE experimental design (Figs. 12 and 13 in Yeager et al., 2022). These results indicate that seasonal SIE prediction skill is higher for JAS than for JFM and that pan-Arctic SIE prediction skill in SMYLE is similar to the skill in other seasonal prediction models (Bushuk et al., 2017, 2021, 2024). As in Yeager et al. (2022), we evaluate forecast skill using the anomaly correlation coefficient (ACC) and the normalized root mean square error (nRMSE):

$$\mathrm{ACC} = \frac{\sum_{i=1}^{N} f_i' \, o_i'}{\sum_{i=1}^{N} f_i'^2 \, \sum_{i=1}^{N} o_i'^2}, \tag{3.2}$$

and

$$\mathrm{nRMSE} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i' - o_i')^2}}{\sigma_o}, \tag{3.3}$$

where f_i' and o_i' are ensemble mean forecast and observed anomalies at the verification time i, N is the sample size (the number of years), and σ_o is the observed standard deviation. Statistical significance of ACC values is determined based on the t-statistic and the effective sample size calculated using the autocorrelations of forecast and observed anomalies (Bretherton et al., 1999). Correlations are considered significantly different from zero when $p < 0.1$ (90% confidence level) (Yeager et al., 2022).

We also compare prediction skill between the best-performing ensemble members and the worst-performing ensemble members to evaluate the effect of OHT prediction skill on SIE prediction skill. For each individual forecast, we compare the predicted JFM OHT from SMYLE to the observed JFM OHT from FOSI and rank the ensemble members for each individual forecast by their mean square error (MSE):

$$\mathrm{MSE}_{i,M} = \left(f_{i,M}' - o_{i,M}'\right)^2, \tag{3.4}$$

where $\mathrm{MSE}_{i,M}$ is the squared difference between the predicted and observed JFM OHT for an individual ensemble member of an individual forecast. We repeat the same process for JAS OHT as well. Then, for each individual forecast we label the ten ensemble members with the lowest MSE scores as "good OHT predictions" and the ten members with the highest MSE scores as "bad OHT predictions". Thus, we create two 10-member ensembles: one that represents the best possible 10-member prediction of OHT and one that represents the worst possible 10-member prediction of OHT. We then evaluate the predictions of SIE in each 10-member ensemble separately.

We calculate the ACC for both sets of predictions and take the difference:

$$\Delta\mathrm{ACC} = \mathrm{ACC}_{\mathrm{M1-M10}} - \mathrm{ACC}_{\mathrm{M11-M20}}, \tag{3.5}$$

where M1–M20 are the ensemble members ranked in order of OHT prediction accuracy. We interpret ΔACC as the relative ACC skill associated with correctly predicting OHT anomalies. If OHT and SIE prediction skill is related, then we expect ensemble members that predict OHT well to also predict SIE well, leading to high values of ΔACC. We note, however, that high values of ΔACC do not on their own indicate a causal relationship between the predictability of OHT and the predictability of sea ice.

3.3. Linear Relationships in FOSI

We first investigate Arctic predictability by examining lagged correlations between variables in FOSI. Linear correlations between the initial conditions of a prediction system and the observed future conditions can explain the role of initial conditions in prediction skill. For Arctic SIE predictions, the initial conditions of sea ice concentration, thickness, and other variables have a demonstrated impact on improving prediction skill (Edward Blanchard-Wrigglesworth, Bitz, et al., 2011; Bushuk et al., 2022; J. J. Day, Hawkins, et al., 2014; J. J. Day, Tietsche, et al., 2014; V. Guemas et al., 2016). Bushuk et al. (2022) demonstrate that SIE prediction skill in dynamical models is largely explained by correlations between SIE and the initial conditions of SIE, SIV, and ocean heat content.

Here, we examine the correlations of SIE with SIE, SIV, and OHT at lags from 0–11 months (Fig. 3.3). SIE persistence is strongest in the Atlantic-Arctic region (Fig. 3.3b). In the Pacific-Arctic and pan-Arctic regions, lagged correlations of SIE anomalies between December and July decrease sharply at increasing lags before showing positive correlations between SIE anomalies and those from the prior year (Fig. 3.3c). This pattern reflects the reemergence of sea ice anomalies in the Pacific-Arctic (Bushuk et al., 2014, 2015; W. Cheng et al., 2016).

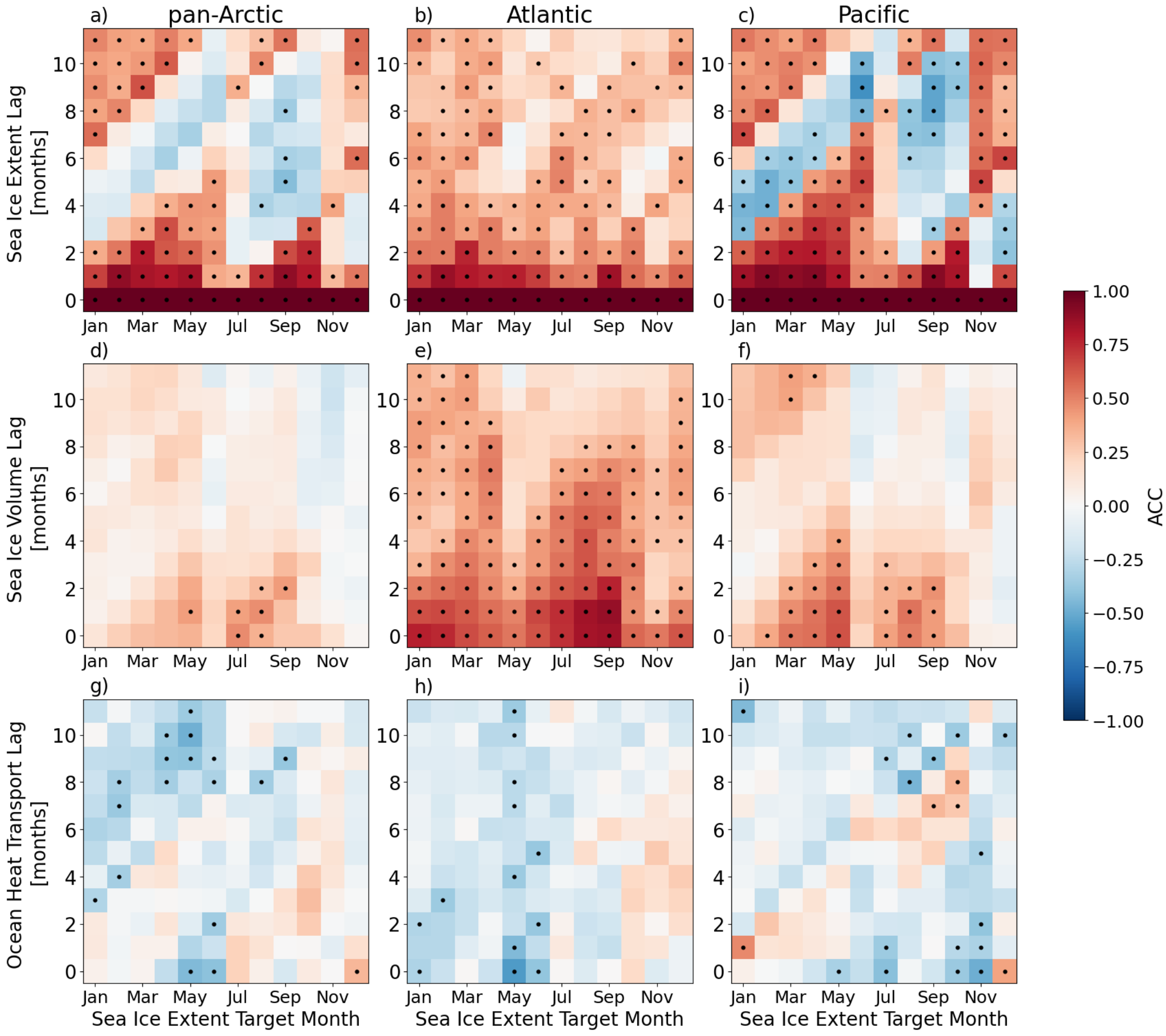


Figure 3.3. Lagged correlations between monthly FOSI sea ice extent and the preceding months' sea ice extent (a–c), sea ice volume (d–f), and ocean heat transport (g–i) in the pan-Arctic region (left column), the Atlantic-Arctic region (middle column), and the Pacific-Arctic region (right column). Dots indicate statistical significance at the 90% confidence level.

SIV anomalies persist similarly to SIE anomalies (Fig. 3.3d–f). SIV lagged correlations with SIE are generally weaker than the SIE-SIE lagged correlations but still show robust relationships that indicate the important role of SIV initialization on seasonal SIE predictions. Further results focus only on SIC and SIE as representative metrics for Arctic sea ice.

Correlations between pan-Arctic, Atlantic-Arctic, and Pacific-Arctic SIE and total, Atlantic-Arctic, and Pacific-Arctic OHT are weakly negative (positive heat transport anomalies lead to negative SIE anomalies) in most months of the year (Fig. 3.3g–i). In the Atlantic-Arctic region, SIE in May and June is strongly anti-correlated with OHT even at long lead times (Fig. 3.3h). In the Pacific-Arctic region (Fig. 3.3i), the relationship between lagged OHT anomalies and SIE is strongest for OHT anomalies that occur between June and November when Bering Strait transport moves heat into the Arctic (Fig. 3.2a).

We calculate the correlation between OHT and annual mean SIC to further evaluate where seasonal OHT is influencing Arctic sea ice. In the annual mean Atlantic-Arctic OHT anomalies are primarily associated with sea ice loss in the Barents Sea, while Pacific-Arctic OHT anomalies are associated with sea ice loss over broad regions of the Chukchi Sea, the East Siberian Sea, and the Laptev Sea (Fig. 3.4a–b). For Atlantic OHT, the JFM and JAS patterns are similar to each other and to the annual mean case, with the strongest relationship between OHT and SIC occurring in the Barents Sea. Significant correlations are also seen in the Pacific sector (Fig. 3.4c and 3.4e).

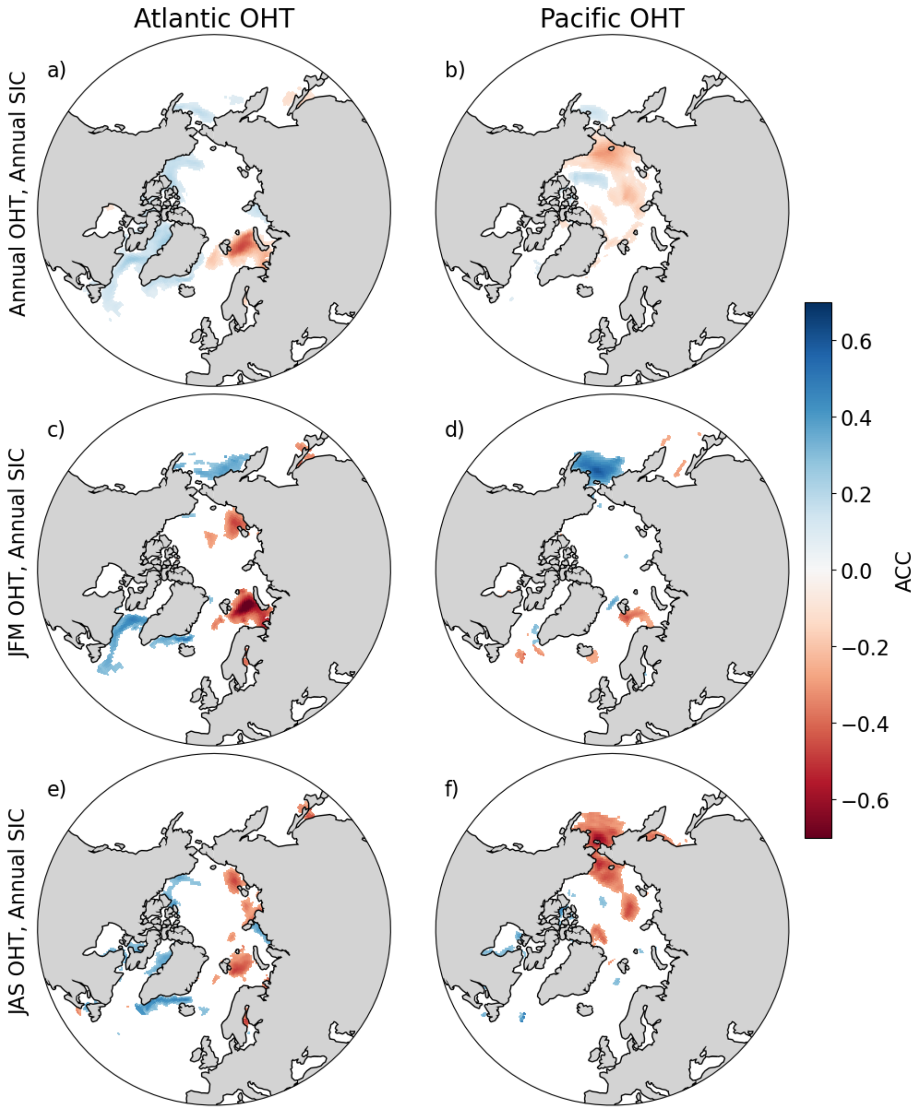


Figure 3.4. Pointwise correlations between annual mean SIC and annual mean OHT (a, b), JFM OHT (c, d), and JAS OHT (e, f). The left column uses Atlantic-Arctic OHT (Fram Strait plus Barents Sea) and the right column uses Pacific-Arctic OHT (Bering Strait). The color bar is flipped such that negative correlations (red) indicate that positive OHT anomalies into the Arctic are associated with reduced SIC. Only correlations that are significant at the 90% confidence level are shown.

Correlations between Pacific-Arctic OHT and annual mean SIC have a strong seasonal dependence, likely due to the seasonal cycle of ocean temperatures in the Bering Strait. Winter OHT anomalies due to anomalous volume transport are primarily associated with positive SIC anomalies in the Bering Sea (Fig. 3.4d), while summer OHT anomalies are associated with significant negative SIC anomalies in the Bering Sea and the Chukchi Sea (Fig. 3.4f). Correlations in FOSI between SIE, SIV, and OHT demonstrate similar patterns of co-variability as seen in other models (Dörr et al., 2021; Muramatsu et al., 2025) and suggest that similar relationships may lead to predictable signals in SMYLE.

3.4. Seasonal Prediction Skill

3.4.1. OHT Prediction Skill

While lagged correlations in FOSI show the co-variability of OHT and sea ice, we evaluate seasonal predictions from SMYLE to analyze prediction skill and predictability. We first evaluate the skill of predicting monthly anomalies of total, Atlantic-Arctic, and Pacific-Arctic OHT. SMYLE demonstrates long-term prediction skill of total Arctic and Atlantic-Artic OHT (Fig. 3.5a and 3.5b). Prediction skill for total OHT and Atlantic-Arctic OHT is similar because Bering Strait OHT is the largest contributor to total OHT and drives the total variability on monthly timescales. Atlantic-Arctic OHT prediction skill depends primarily on lead time and is predictable in all months of the year (Fig. 3.5b).

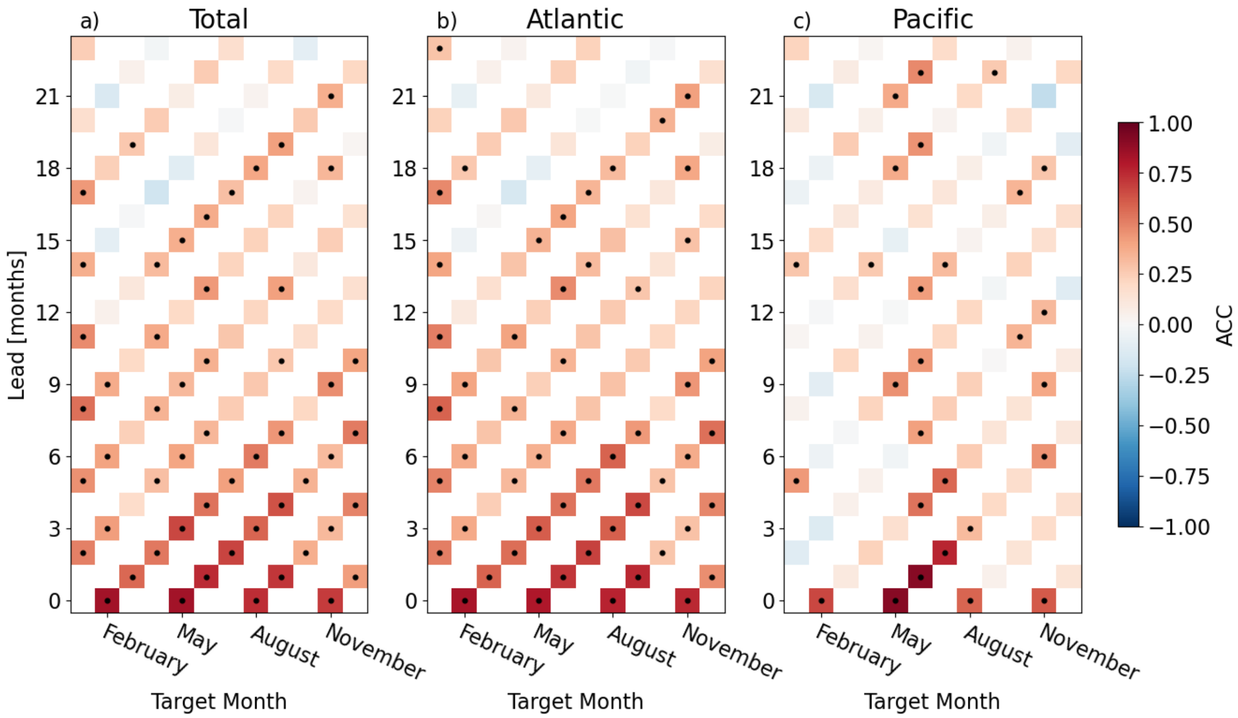


Figure 3.5. ACC by target month and lead time for SMYLE predictions of total OHT anomalies (a), Atlantic-Arctic OHT anomalies (b), and Pacific-Arctic OHT anomalies (c). Dots indicate statistical significance at the 90% confidence level.

Pacific-Arctic OHT depends largely on the target month instead of on lead time, with a strong, long-term predictable signal in May, June, and November (Fig. 3.5c). Prediction skill for Pacific-Arctic OHT reduces between June and November during the late melt season, as increased stratification limits OHT–sea-ice interactions, and increases again in November as sea ice grows (Lenetsky et al., 2021; Lu et al., 2020). In other months of the year, Pacific-Arctic OHT variability is weaker and SMYLE predictions have limited skill.

3.4.2. Sea Ice Prediction Skill

We then examine the ACC for SIE predictions by target month and lead month (Fig. 3.6), expanding on the analysis from the SMYLE experimental design manuscript (Yeager et al., 2022). At short lead times, pan-Arctic, Atlantic-Arctic, and Pacific-Arctic predictions are all statistically significant and indicate that SMYLE has skill similar to other statistical and dynamical prediction systems (Bushuk et al., 2024).

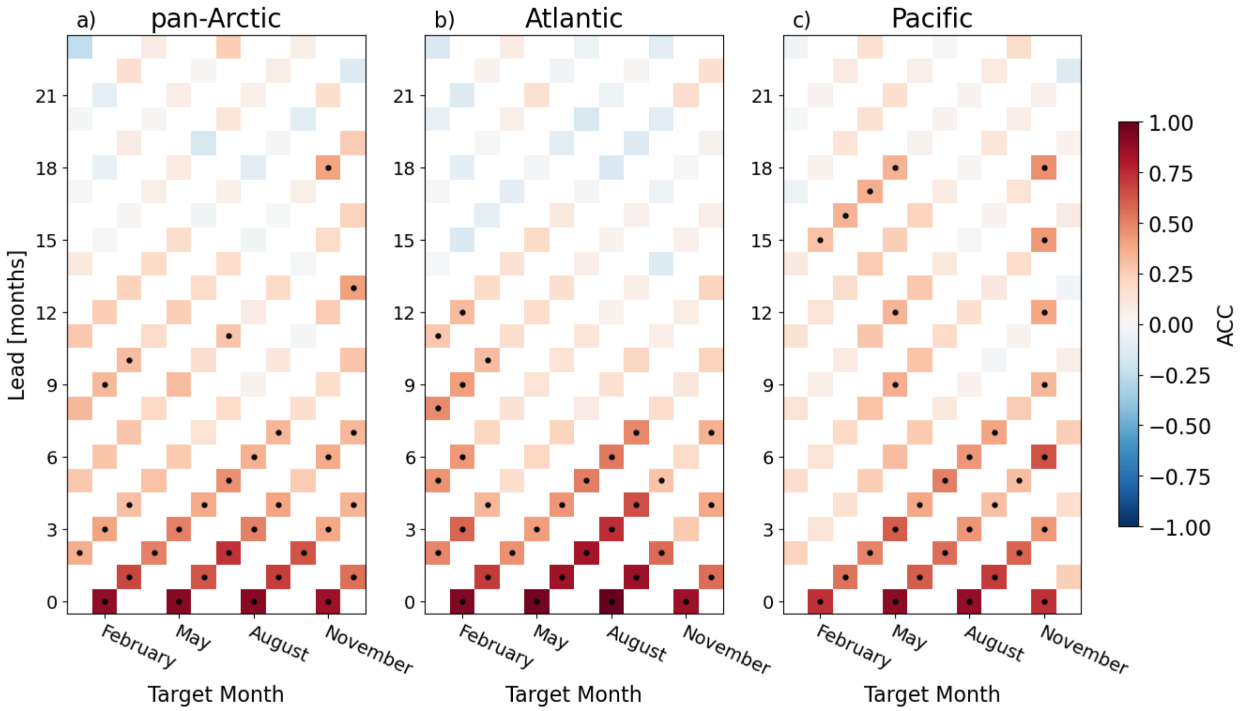


Figure 3.6. ACC by target month and lead time for SMYLE predictions of pan-Arctic SIE anomalies (a), Atlantic-Arctic SIE anomalies (b), and Pacific-Arctic SIE anomalies (c).

Pacific-Arctic SIE is best predicted in late spring and late fall (Fig. 3.6c), matching the seasons where Pacific-Arctic OHT is best predicted. This co-occurrence of high prediction skill between Bering Strait OHT and Pacific-Arctic SIE suggests that the ocean could drive sea ice predictability, sea ice could drive ocean predictability, or related mechanisms in the atmosphere or the coupled ocean-atmosphere-ice system drives predictability in both. Lenetsky et al. (2021) and Bushuk et al. (2022) show that June, July, and November SIE in the Chukchi and Bering Seas can be predicted by upper ocean heat content anomalies many months ahead that are related to the impact of Bering Strait heat transport (Serreze et al., 2016). Between July and November, the Pacific sector sea ice edge is far from the geographical influence of Bering Strait OHT, so the ocean's influence on SIE is limited (Bushuk et al., 2022). Docquier et al., (2022) find that OHT, sea ice area, and 2-meter surface temperature each drive variability in each other, suggesting that air-sea interaction can also serve as a mechanism of predictability.

3.4.3. Co-predictability of OHT and Sea Ice

We further investigate the co-predictability of OHT and sea ice in SMYLE by evaluating SIC predictions in ensemble members with good skill in predicting OHT separated from ensemble members with bad skill in predicting OHT. Following the method described in Sec. 3.2.5, we calculate ΔACC to analyze the SIC prediction skill associated with correctly predicting JFM or JAS OHT.

In winter, the full 20-member SMYLE ensemble has significant prediction skill around the sea ice edge for predictions initialized in November (3-month lead) and May (9-month lead) (Fig. 3.7a and 3.7d). Correctly predicting JFM Atlantic-Arctic OHT leads to associated improvements in prediction skill primarily in the Barents Sea, though limited improvements are also seen in the Eastern Bering Sea and the Labrador Sea (Fig. 3.7b and 3.7e). While the ΔACC is smaller for the 9-month lead than for the 3-month lead, the spatial pattern of skill improvement is the same.

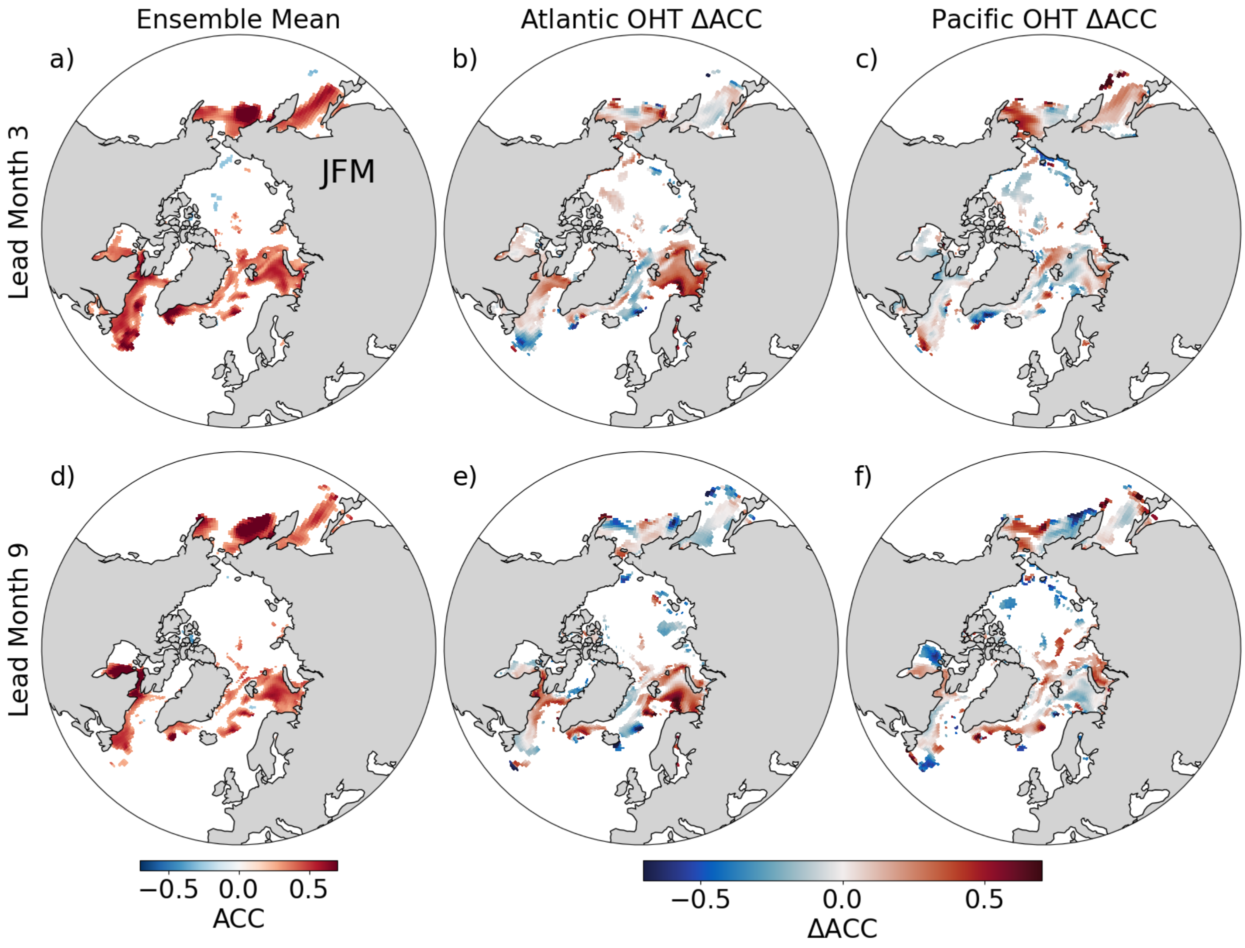


Figure 3.7. JFM SIC skill associated with JFM OHT. JFM SIC skill from the full 20-member SMYLE ensemble at lead month 3 (November initialization; a) and lead month 9 (May initialization; d); ΔACC associated with JFM Atlantic-Arctic OHT prediction skill at lead month 3 (b) and 9 (e); and ΔACC associated with JFM Pacific-Arctic OHT prediction skill at lead month 3 (c) and 9 (f). ΔACC is only locations shown where $\mathrm{ACC}_{\mathrm{M1-M10}}$ or $\mathrm{ACC}_{\mathrm{M11-M20}}$ is significant at the 90% confidence interval.

The SIC prediction skill associated with correct predictions of Pacific-Arctic OHT occurs with an opposite spatial pattern (Fig. 3.7c and 3.7f). Pacific OHT prediction skill is associated with a degradation of SIC prediction skill in the Barents Sea and an improvement of SIC prediction skill in the Bering Sea at both 3- and 9-month leads.

Prediction skill of JAS SIC in SMYLE is widespread and significant at a lead time of 3 months (May initialization) but is limited and insignificant at a lead time of 9 months (November initialization) (Fig. 3.8a and 3.8d). The ΔACC associated with Atlantic-Arctic OHT is significant over most of the Arctic, indicating a large influence of OHT on summer sea ice at both 3 and 9 lead months (Fig. 3.8b and 3.8e). The influence of Pacific-Arctic OHT is seen mostly near the summer sea ice edge in the Pacific sector at a lead time of 3 months (Fig. 3.8c). At a lead time of 9 months, Pacific-Arctic OHT prediction skill is associated with SIC skill mostly in the Central Arctic and the Chukchi Sea (Fig. 3.8f).

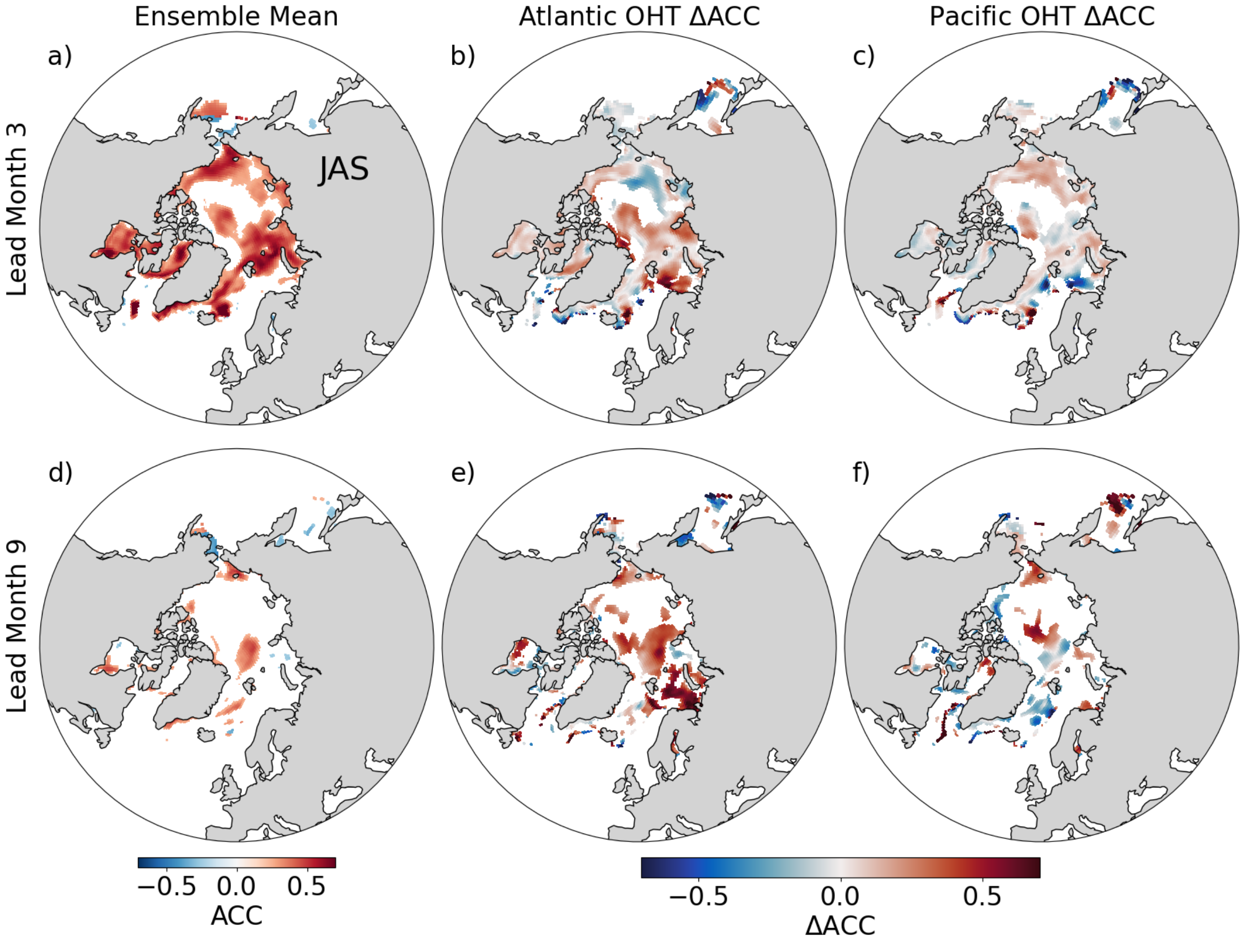


Figure 3.8. As in Figure 3.7, but for JAS SIC skill associated with JAS OHT.

The clear relationship between the prediction skill of regional OHT and SIC in both winter and summer and in both the Atlantic-Arctic and the Pacific-Arctic suggests that Arctic OHT and sea ice predictability are linked throughout the year in different parts of the region. We demonstrate that an improvement of OHT prediction skill co-occurs with improvements of SIC across seasons and regions in the Arctic. This result can be dynamically linked to ocean-driven sea ice variability in both the Atlantic-Arctic and the Pacific-Arctic (Bushuk et al., 2017, 2022; Docquier et al., 2022; Lenetsky et al., 2021; Serreze et al., 2016).

3.5. Decadal Prediction Skill

We next turn to interannual-to-decadal predictions of Arctic OHT and SIE. Decadal variability of Arctic sea ice has been shown to be dependent on variations in both Atlantic OHT (Yeager et al., 2012, 2015) and Pacific OHT (Y. Li et al., 2024). While the forced component of the climate system can explain much of the recent decadal variability in SIE, internal climate variability still plays a key role in observed SIE trends (Dörr et al., 2023; England et al., 2025).

Yeager et al. (2015) demonstrate that decadal trends of pan-Arctic and Atlantic-Arctic winter SIE are linked to oceanic and atmospheric conditions in the subpolar North Atlantic. Using an old version of a CESM decadal prediction system, they correctly predicted the recent continued hiatus in JFM Arctic sea ice loss. Here, we revisit their analysis and expand it to include Pacific-Arctic and summer SIE. We evaluate 5–7-year mean predictions of OHT and SIE before comparing the predicted 10-year linear OHT and SIE trends in DP to trends in FOSI.

The 5–7-year ensemble mean OHT anomalies from DP reflect the long-term upward trends of FOSI OHT (Fig. 3.9). High correlation values between DP and FOSI come mostly from the long-term trend, however. While the ensemble mean fails to capture much of the interannual variability around the trend in any region or season, the ensemble spread captures this internal variability in most years.

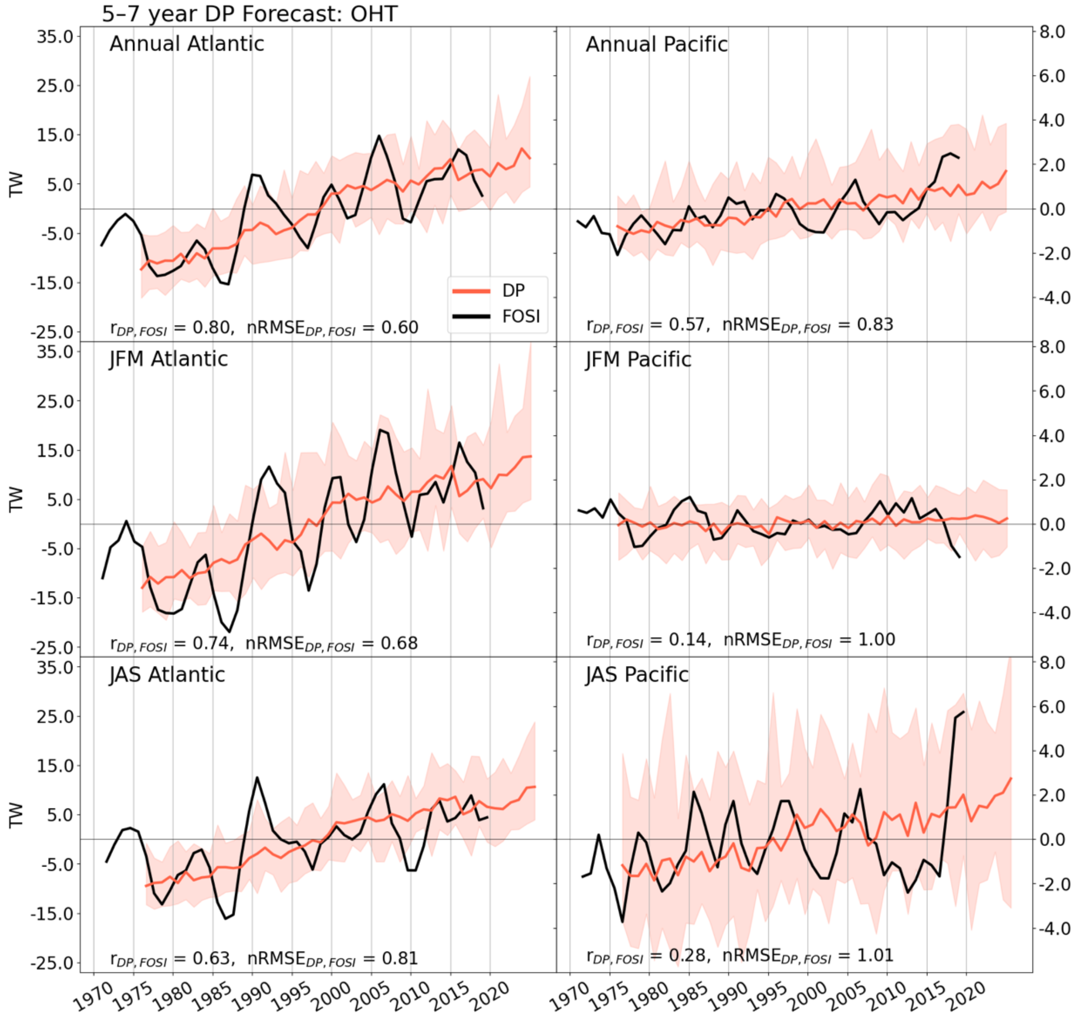


Figure 3.9. Interannual variability in Atlantic-Arctic (left) and Pacific-Arctic (right) annual mean OHT (top), JFM OHT (middle), and JAS OHT (bottom). Black lines show the 3-year running mean OHT from FOSI, the red line shows the 5–7-year ensemble mean prediction from DP, and the red shading shows the DP ensemble spread. Correlation and nRMSE values are shown for each panel.

The 5–7-year ensemble mean SIE anomalies from DP also reflect the long-term trends of SIE from both FOSI and observations and better capture some of the interannual variability in SIE (Fig. 3.10). Again, correlation scores are high due to the strong linear trend in SIE.

Figure 3.10. Interannual variability in pan-Arctic (left), Atlantic-Arctic (middle) and Pacific-Arctic (right) annual mean SIE (top), JFM SIE (middle), and JAS SIE (bottom). Black lines show the 3-year running mean SIE from FOSI, the blue dotted lines show the 3-year running mean SIE from observations, the red line shows the 5–7-year ensemble mean prediction from DP, and the red shading shows the DP ensemble spread. Correlation and nRMSE values between DP and FOSI and between DP and observations are shown for each panel.

Interannual variability in both SIE and OHT is dominated by their linear trends, which 5–7-year DP predictions capture relatively well. The ensemble mean SIE predictions, however, do not capture the 1997–2007 decline in SIE quite as well as the decadal prediction system used in Yeager et al. (2015) did. Nonetheless, the ensemble mean Atlantic-Arctic OHT predictions show some skill in capturing the upward trend between 1997 and 2007 (Fig. 3.9).

As DP shows some skill in producing decadal predictions of SIE and OHT across seasons and Arctic regions, we go on to analyze the skill with which individual DP ensembles predict the short-term trend. To do so, we calculate the 10-year linear trend for each DP initialization following the method in Yeager et al. (2015). The predicted 10-year OHT trends in DP are highly correlated with the observed trends calculated in FOSI in both the Atlantic and the Pacific in all seasons, with the exception of annual mean Pacific OHT (Fig. 3.11).

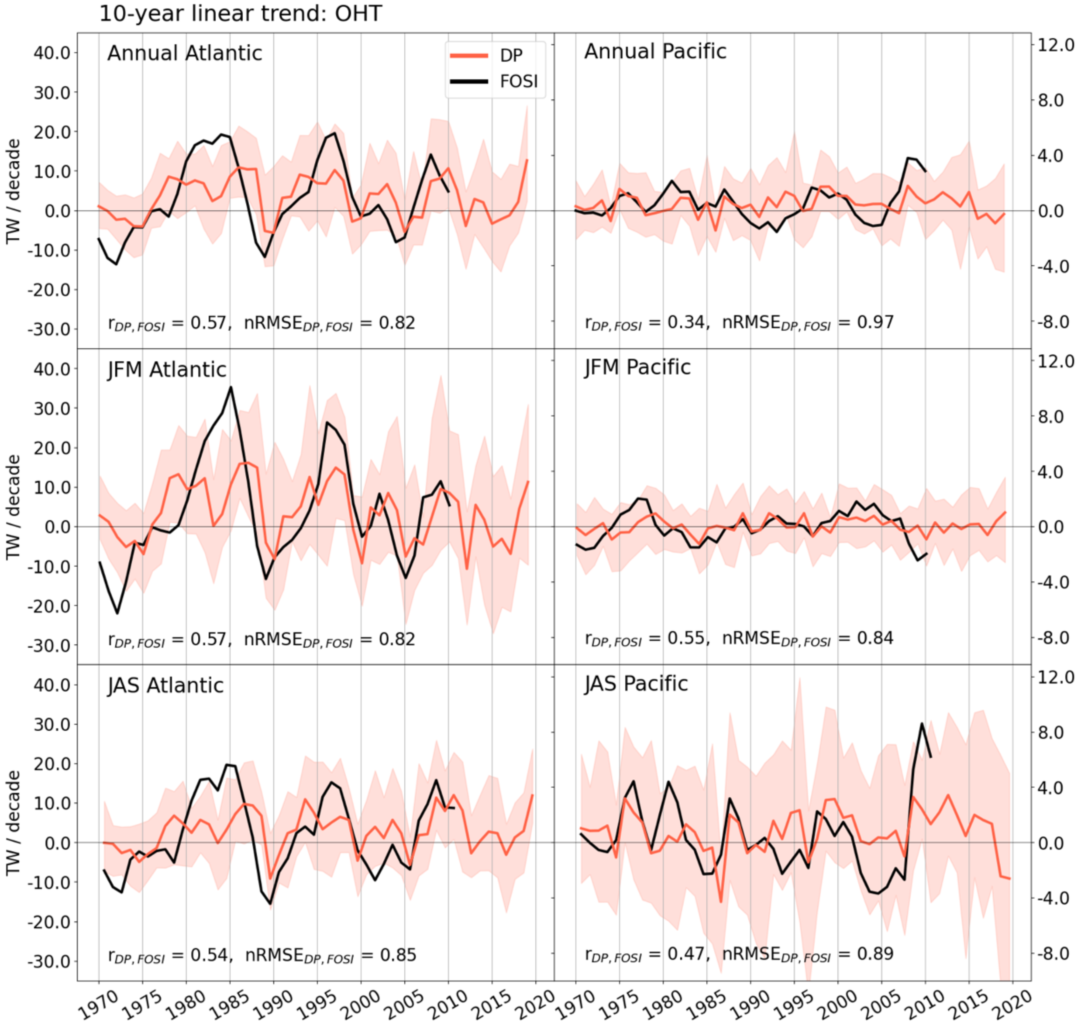


Figure 3.11. Decadal trends in Arctic OHT. 10-year linear trends are calculated for Atlantic-Arctic (left) and Pacific-Arctic (right) annual mean OHT (top), JFM OHT (middle), and JAS OHT (bottom). Black lines show the OHT trend from FOSI, red lines show the linear trend from DP, and red shading shows the DP trend ensemble spread. Correlation and nRMSE values are shown for each panel.

Similarly, the predicted 10-year SIE trends in DP are highly correlated with trends from both FOSI and observations (Fig. 3.12). The ensemble spread of DP trends captures the decadal variability quite well in the annual case and in winter and summer (Fig. 3.12, left column). The strong sea ice decline from 1997–2007 is well reflected by DP trend predictions even though the ensemble mean DP prediction of SIE itself did not capture the extent of the sea ice decline (Fig. 3.10).

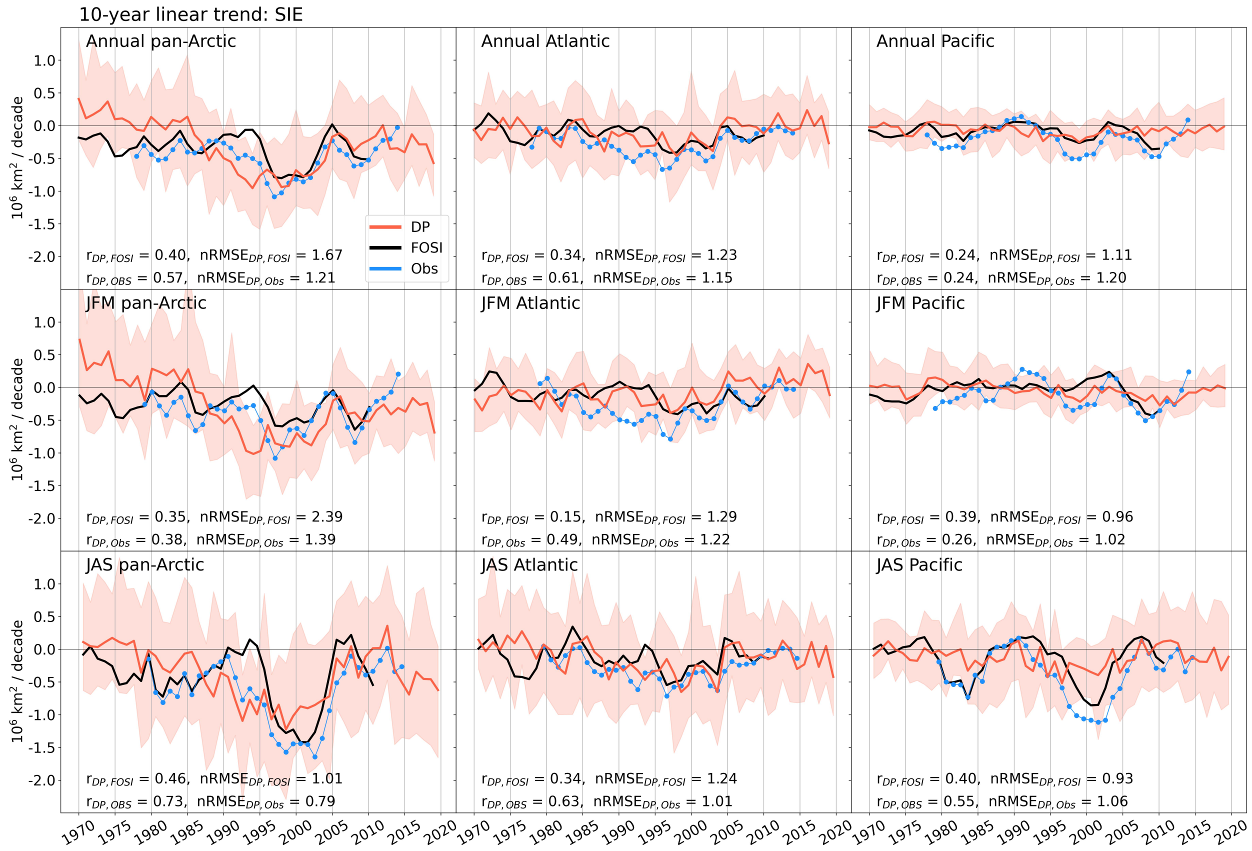


Figure 3.12. Decadal trends in Arctic SIE. 10-year linear trends are calculated for pan-Arctic (left), Atlantic-Arctic (middle) and Pacific-Arctic (right) annual mean SIE (top), JFM SIE (middle), and JAS SIE (bottom). Black lines show the trend from FOSI, the blue dotted lines show the trend from observations, the red line shows the trend from DP, and the red shading shows the DP ensemble spread. Correlation and nRMSE values between DP and FOSI and between DP and observations are shown for each panel.

Looking past the end of the 11-year periods in the FOSI record (2009–2019) and the observational record (2014–2024), we see that recently initialized decadal predictions show a potential decrease in SIE trends, indicating a potential resumption of Arctic sea ice loss. This pattern is mostly seen in the pan-Arctic (Fig. 3.12, left column) and Atlantic-Arctic (Fig. 3.12, middle column) regions, and corresponds to predictions of positive Atlantic-Arctic OHT trends (Fig. 3.11). High skill for 10-year trends of pan-Arctic SIE and regional OHT indicate that DP is able to predict the tendency of the Arctic potentially due to the tendency of OHT through the Atlantic-Arctic and Pacific-Arctic gateways. Consequently, coherent predictions of a positive decadal OHT trend and a negative decadal SIE trend could indicate a return to the long-term trends related to a warming Arctic. At the end of the observational record, however, the observed decadal trend in annual and JFM pan-Arctic sea ice already diverges from the predicted decadal trend in DP. Further analysis is needed in order to confidently predict the potential future scenarios of Arctic sea ice decline.

3.6. Conclusions

Seasonal-to-decadal forecasts of Arctic sea ice and ocean heat transport are important in a dynamic Arctic climate. We evaluate relationships between annual and seasonal Arctic sea ice and OHT with emphases on the role of seasonality and the difference between the Atlantic-Arctic and the Pacific-Arctic regions. On seasonal timescales, predictions of regional SIE and OHT in all months are statistically significant for several months across the year. We demonstrate that the prediction skill of OHT and sea ice are linked in different regions and seasons, likely through ocean–sea-ice interactions. Ensemble members that correctly predict OHT anomalies tend to more accurately predict SIC anomalies in these areas. Finally, we show that decadal predictions of OHT and SIE capture the long-term trends in the Arctic. Evaluation of predictions of 10-year linear trends indicates that DP has significant skill in predicting the decadal tendency of Arctic OHT and SIE.

The results and conclusions presented here add to a growing body of research on Arctic predictability. SMYLE and DP join the ranks of dynamical and statistical prediction models that contribute to seasonal and interannual forecasts of Arctic sea ice (Blanchard-Wrigglesworth et al., 2023; E. Blanchard-Wrigglesworth & Bushuk, 2019; Bushuk et al., 2024; Jonathan J. Day et al., 2016; Stroeve et al., 2014). Along with increasingly robust multi-model ensemble prediction systems, research on the mechanisms that act as sources of predictability continues to be important for assessing the reliability and fidelity of Arctic forecasts.

In this study, we explore the relationships between the predictability of OHT and sea ice. Correlations between OHT and SIE, however, do not necessarily indicate causality and could be related causally to other variables. Nonetheless, robust correlation patterns still suggest that improved modeling of ocean processes and OHT can in turn improve SIC and SIE predictions. One limitation of the current study is the use of a November-initialized decadal prediction system. SMYLE demonstrates that November-initialized predictions of SIE underperform relative to predictions of the same lead times initialized from February, May, and August (Figs. 3.5 and 3.6). This could suggest that the DP predictions, at least in the shorter term (lead years 1–2), provide a lower bound on potential skill from initialized decadal prediction systems.

The regions of Arctic winter sea ice that co-vary with Atlantic and Pacific OHT are expected to change as the sea ice footprint changes in the future (Dörr et al., 2021, 2024). The influence of Bering Strait OHT on winter sea ice is expected to increase, having more influence on sea ice in the Chukchi Sea, the East Siberian Sea, and the Central Arctic. The influence of Barents Sea OHT, on the other hand, will decrease overall and transition toward the Kara and Laptev Seas (Dörr et al., 2024). As winter sea ice variability trends toward depending more on Pacific-Arctic

OHT than on Atlantic-Arctic OHT, improved understanding of Bering Strait OHT is paramount for understanding the ocean's role in Arctic predictability.

Improving model representation of ocean heat transport has implications for improving predictions of seasonal-to-annual Arctic sea ice variability as well as decadal sea ice trends. High-resolution models have been shown to capture relationships between Bering Strait OHT and sea ice variability that low-resolution models like SMYLE and DP miss (Y. Li et al., 2024). Thus, improving decadal OHT predictions will likely improve our predictive capabilities for Arctic sea ice as well. Future work evaluating high-resolution decadal predictions could elucidate the role of model resolution on Arctic OHT and sea ice prediction skill and could yield more confident projections of decadal sea ice change.

Appendix 3A: Arctic Region Definition

To create masks for the Atlantic-Arctic and Pacific-Arctic regions, we regrid the National Snow and Ice Data Center (NSIDC) region mask (Walter Meier & Stewart, 2023) to the POP grid (Fig. 3.A1). The Atlantic-Arctic region includes the Kara Sea, the Barents Sea, the East Greenland Sea, Baffin Bay and the Labrador Seas, the Gulf of St. Lawrence, Hudson Bay, the Baltic Sea and the Central Arctic between -90°E and 90°E (Fig. 3.A1, purple). The Pacific-Arctic region includes the Beaufort Sea, the Chukchi Sea, the East Siberian Sea, the Laptev Sea, the Canadian Archipelago, the Bering Sea, the Sea of Okhotsk, the Sea of Japan, the Bohai and Yellow Seas, the Gulf of Alaska, and the Central Arctic between 90°E and 270°E (Fig. 3.A1, gold). The pan-Arctic region includes all Arctic regions.

Figure 3.A1. Map of the Atlantic-Arctic (purple) and the Pacific-Arctic (gold) regions. Regions were determined by partitioning the NSIDC region mask.

Conclusion

Understanding the variability and predictability of ocean dynamics will improve our knowledge of climate variability. The first chapter of this dissertation improves our understanding of the processes that contribute to air-sea interaction and upper ocean heat content variability across seasons and regions. The regional and seasonal patterns of the feedback sensitivity demonstrate the varying pathways by which the ocean influences the surface heat flux feedback. My results suggest that the region in which the ocean exerts the most influence over the SST and turbulent heat flux variability changes throughout the year as the oceanic processes driving the upper ocean heat budget change. Determining these patterns of ocean-dominated variability improves predictive understanding of the climate.

The second chapter gives us a new perspective on marine heatwave predictions and has the potential to improve our ability to forecast extreme SST events. By applying novel tools to seasonal ocean predictions, I demonstrate that SMYLE predicts marine heatwaves up to a year in advance, but with errors in location, size, and temperature which get worse as events are predicted further in advance. Spatial marine heatwave predictions could be used to inform marine resource management efforts and to communicate uncertainty in operational forecasts. This work could be extended to apply uncertainty parameters to the output of future climate forecasts, improving the distribution of useful, useable, and used marine heatwave forecasts (Spillman et al., 2025).

The third chapter enhances our understanding of sources of predictability for Arctic sea ice forecasts. I demonstrate that CESM2 skillfully predicts the seasonal-to-decadal variability of ocean heat transport and sea ice across different seasons and regions of the Arctic. I also show that CESM2-DP predicts decadal trends in observed and modeled sea ice extent with high skill, potentially allowing us to predict Arctic changes in the next 10 years. Determining the relationship between the ocean heat transport and Arctic sea ice predictability also can potentially improve model representations of Arctic ocean–sea-ice dynamics.

Broadly, this dissertation enhances our understanding of the ocean's role in climate variability and predictability on a range of temporal and spatial scales and across components of the Earth system. As the long-term memory holder of the climate system, the ocean holds the key to climate predictability. Improving our understanding of its internal dynamics and its interactions with the atmosphere and with sea ice will allow us to continue developing skillful prediction systems to accurately forecast future climate states. The three chapters of this dissertation each contribute to this goal of better understanding the role of the ocean in air-sea interaction and climate predictability.

Acknowledgements

This dissertation could not have been completed without the tremendous support of many individuals and institutions. Throughout my Ph.D., my work was funded by the University of Washington Program on Climate Change (UW PCC), the National Science Foundation (NSF), the National Aeronautics and Space Administration (NASA), the National Oceanic and Atmospheric Administration (NOAA), and the Cooperative Institute for Climate, Ocean, & Ecosystem Studies (CICOES). I also received funding for travel support from US CLIVAR and the National Center for Atmospheric Research (NCAR). I am extremely grateful for the federal research funding I received and sincerely hope that the federal government continues to fund basic climate research so that future graduate students have the same opportunities that I did to explore research and, of course, to travel to new places.

None of that would have been possible without my advisor, LuAnne Thompson, who encouraged me to chase new opportunities and allowed me to pivot research topics more than once. I am also thankful for the advisory support I received from collaborators. Elizabeth Maroon and Anna-Lena Deppenmeier helped me with research, writing, career opportunities, and more. Wei Cheng and Stephen Yeager provided the funding, support, and insightful ideas that made Chapter 3 possible. I'd also like to thank my advisory committee members (Alison Gray, Mark Warner, Wei Cheng, and CC Bitz) and past members (Maike Sonnewald and Kyla Drushka) for guiding me through grad school. Lastly, I am so grateful for the leadership opportunities and the community I found through the PCC, which shaped the first several years of my Ph.D..

I owe a lot of my success to my cohort (Kitty, Maleen, Will, Josh, Zoe, Haila, Zoe, and Mary Margaret) and to the Thompson Lab Group (Noah, Cassia, Katie, and Carlyn). I completed most

of the research in this dissertation by working out loud in the presence of many of these people, whether or not they were listening. My Seattle community of biking, ultimate frisbee, and non-sporting friends got me through the non-academic parts of the Ph.D., and I thank them all for making these six years fly by.

Finally, my siblings, parents, and extended family have supported me so much throughout this process. Regardless of what was happening with my work, I knew they were rooting for me. Eli, Michael, and Sophie, thank you for both encouraging me and distracting me whenever I needed either.

References

Abatan, A. A., Gutowski, W. J., Ammann, C. M., Kaatz, L., Brown, B. G., Buja, L., et al. (2018). Statistics of multi-year droughts from the method for object-based diagnostic evaluation. *International Journal of Climatology*, *38*(8), 3405–3420. https://doi.org/10.1002/joc.5512

Agarwal, N., Small, R. J., Bryan, F. O., Grooms, I., & Pegion, P. J. (2023). Impact of Stochastic Ocean Density Corrections on Air-Sea Flux Variability. *Geophysical Research Letters*, *50*(13), e2023GL104248. https://doi.org/10.1029/2023GL104248

Amaya, D. J., Miller, A. J., Xie, S.-P., & Kosaka, Y. (2020). Physical drivers of the summer 2019 North Pacific marine heatwave. *Nature Communications*, *11*(1), 1903. https://doi.org/10.1038/s41467-020-15820-w

Årthun, M., Eldevik, T., Smedsrud, L. H., Skagseth, Ø., & Ingvaldsen, R. B. (2012). Quantifying the Influence of Atlantic Heat on Barents Sea Ice Variability and Retreat*. *Journal of Climate*, *25*(13), 4736–4743. https://doi.org/10.1175/JCLI-D-11-00466.1

Årthun, M., Onarheim, I. H., Dörr, J., & Eldevik, T. (2021). The Seasonal and Regional Transition to an Ice-Free Arctic. *Geophysical Research Letters*, *48*(1), e2020GL090825. https://doi.org/10.1029/2020GL090825

Årthun, M., Brakstad, A., Dörr, J., Johnson, H. L., Mans, C., Semper, S., & Våge, K. (2025). Atlantification drives recent strengthening of the Arctic overturning circulation. *Science Advances*, *11*(28), eadu1794. https://doi.org/10.1126/sciadv.adu1794

Bach, E., Motesharrei, S., Kalnay, E., & Ruiz-Barradas, A. (2019). Local Atmosphere–Ocean Predictability: Dynamical Origins, Lead Times, and Seasonality. *Journal of Climate*, *32*(21), 7507–7519. https://doi.org/10.1175/JCLI-D-18-0817.1

Barsugli, J. J., & Battisti, D. S. (1998). The Basic Effects of Atmosphere–Ocean Thermal Coupling on Midlatitude Variability*. *Journal of the Atmospheric Sciences*, *55*(4), 477–493. https://doi.org/10.1175/1520-0469(1998)055<0477:TBEOAO>2.0.CO;2

Benthuysen, J., Feng, M., & Zhong, L. (2014). Spatial patterns of warming off Western Australia during the 2011 Ningaloo Niño: Quantifying impacts of remote and local forcing. *Continental Shelf Research*, *91*, 232–246. https://doi.org/10.1016/j.csr.2014.09.014

Bian, C., Jing, Z., Wang, H., Wu, L., Chen, Z., Gan, B., & Yang, H. (2023). Oceanic mesoscale eddies as crucial drivers of global marine heatwaves. *Nature Communications*, *14*(1), 2970. https://doi.org/10.1038/s41467-023-38811-z

Bian, C., Jing, Z., Wang, H., & Wu, L. (2024). Scale-Dependent Drivers of Marine Heatwaves Globally. *Geophysical Research Letters*, *51*(3), e2023GL107306. https://doi.org/10.1029/2023GL107306

Bianco, E., Iovino, D., Masina, S., Materia, S., & Ruggieri, P. (2024). The role of upper-ocean heat content in the regional variability of Arctic sea ice at sub-seasonal timescales. *The Cryosphere*, *18*(5), 2357–2379. https://doi.org/10.5194/tc-18-2357-2024

Bishop, S. P., Small, R. J., Bryan, F. O., & Tomas, R. A. (2017). Scale Dependence of Midlatitude Air–Sea Interaction. *Journal of Climate*, *30*(20), 8207–8221. https://doi.org/10.1175/JCLI-D-17-0159.1

Blanchard-Wrigglesworth, E., & Bushuk, M. (2019). Robustness of Arctic sea-ice predictability in GCMs. *Climate Dynamics*, *52*(9–10), 5555–5566. https://doi.org/10.1007/s00382-018-4461-3

Blanchard-Wrigglesworth, E., Bushuk, M., Massonnet, F., Hamilton, L. C., Bitz, C. M., Meier, W. N., & Bhatt, U. S. (2023). Forecast Skill of the Arctic Sea Ice Outlook 2008–2022.

Geophysical Research Letters, *50*(6), e2022GL102531. https://doi.org/10.1029/2022GL102531

Blanchard-Wrigglesworth, Edward, Bitz, C. M., & Holland, M. M. (2011). Influence of initial conditions and climate forcing on predicting Arctic sea ice. *Geophysical Research Letters*, *38*(18), n/a-n/a. https://doi.org/10.1029/2011GL048807

Blanchard-Wrigglesworth, Edward, Armour, K. C., Bitz, C. M., & DeWeaver, E. (2011). Persistence and Inherent Predictability of Arctic Sea Ice in a GCM Ensemble and Observations. *Journal of Climate*, *24*(1), 231–250. https://doi.org/10.1175/2010JCLI3775.1

Bonino, G., Masina, S., Galimberti, G., & Moretti, M. (2023). Southern Europe and western Asian marine heatwaves (SEWA-MHWs): a dataset based on macroevents. *Earth System Science Data*, *15*(3), 1269–1285. https://doi.org/10.5194/essd-15-1269-2023

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., & Bladé, I. (1999). The Effective Number of Spatial Degrees of Freedom of a Time-Varying Field. *Journal of Climate*, *12*(7), 1990–2009. https://doi.org/10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2

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. https://doi.org/10.1175/JCLI-D-13-00316.1

Bullock, R., Brown, B., & Fowler, T. (2016). Method for object-based diagnostic evaluation. *NCAR Technical Note*.

Burger, F. A., Terhaar, J., & Frölicher, T. L. (2022). Compound marine heatwaves and ocean acidity extremes. *Nature Communications*, *13*(1), 4722. https://doi.org/10.1038/s41467-022-32120-7

Busecke, J. J. M., Balwada, D., Martin, P. E., Nicholas, T. E. G., Johnson, Z. C. P., Nalluri, P., et al. (2025). The Impact of Sub-Grid Heterogeneity on Air-Sea Turbulent Heat Flux in Coupled Climate Models. *Geophysical Research Letters*, *52*(13), e2025GL114951. https://doi.org/10.1029/2025GL114951

Bushuk, M., Giannakis, D., & Majda, A. J. (2014). Reemergence Mechanisms for North Pacific Sea Ice Revealed through Nonlinear Laplacian Spectral Analysis*. *Journal of Climate*, *27*(16), 6265–6287. https://doi.org/10.1175/JCLI-D-13-00256.1

Bushuk, M., Giannakis, D., & Majda, A. J. (2015). Arctic Sea Ice Reemergence: The Role of Large-Scale Oceanic and Atmospheric Variability*. *Journal of Climate*, *28*(14), 5477–5509. https://doi.org/10.1175/JCLI-D-14-00354.1

Bushuk, M., Msadek, R., Winton, M., Vecchi, G. A., Gudgel, R., Rosati, A., & Yang, X. (2017). Skillful regional prediction of Arctic sea ice on seasonal timescales. *Geophysical Research Letters*, *44*, 4953–4964. https://doi.org/10.1002/2017GL073155

Bushuk, M., Winton, M., Haumann, F. A., Delworth, T., Lu, F., Zhang, Y., et al. (2021). Seasonal Prediction and Predictability of Regional Antarctic Sea Ice. *Journal of Climate*, *34*(15), 6207–6233. https://doi.org/10.1175/JCLI-D-20-0965.1

Bushuk, M., Zhang, Y., Winton, M., Hurlin, B., Delworth, T., Lu, F., et al. (2022). Mechanisms of Regional Arctic Sea Ice Predictability in Two Dynamical Seasonal Forecast Systems. *Journal of Climate*, *35*(13), 4207–4231. https://doi.org/10.1175/JCLI-D-21-0544.1

Bushuk, M., Ali, S., Bailey, D. A., Bao, Q., Batté, L., Bhatt, U. S., et al. (2024). Predicting September Arctic Sea Ice: A Multimodel Seasonal Skill Comparison. *Bulletin of the American Meteorological Society*, *105*(7), E1170–E1203. https://doi.org/10.1175/BAMS-D-23-0163.1

Capotondi, A., Rodrigues, R. R., Sen Gupta, A., Benthuysen, J. A., Deser, C., Frölicher, T. L., et al. (2024). A global overview of marine heatwaves in a changing climate. *Communications Earth & Environment*, *5*(1), 701. https://doi.org/10.1038/s43247-024-01806-9

Cayan, D. R. (1992). Latent and Sensible Heat Flux Anomalies over the Northern Oceans: Driving the Sea Surface Temperature. *Journal of Physical Oceanography*, *22*(8), 859–881. https://doi.org/10.1175/1520-0485(1992)022<0859:LASHFA>2.0.CO;2

Cheng, K., Shu, Q., Wang, Q., Song, Z., He, Y., Wang, S., et al. (2025). Distinct Impacts of Increased Atlantic and Pacific Ocean Heat Transport on Arctic Ocean Warming and Sea Ice Decline. *Journal of Geophysical Research: Oceans*, *130*(3), e2024JC021178. https://doi.org/10.1029/2024JC021178

Cheng, W., Blanchard-Wrigglesworth, E., Bitz, C. M., Ladd, C., & Stabeno, P. J. (2016). Diagnostic sea ice predictability in the pan-Arctic and U.S. Arctic regional seas. *Geophysical Research Letters*, *43*(22). https://doi.org/10.1002/2016GL070735

Clancy, R., Bitz, C., & Blanchard-Wrigglesworth, E. (2021). The influence of ENSO on Arctic sea ice in large ensembles and observations. *Journal of Climate*, 1–50. https://doi.org/10.1175/JCLI-D-20-0958.1

Clark, A. J., Bullock, R. G., Jensen, T. L., Xue, M., & Kong, F. (2014). Application of Object-Based Time-Domain Diagnostics for Tracking Precipitation Systems in Convection-

Allowing Models. *Weather and Forecasting*, *29*(3), 517–542. https://doi.org/10.1175/WAF-D-13-00098.1

Conil, S., Douville, H., & Tyteca, S. (2009). Contribution of realistic soil moisture initial conditions to boreal summer climate predictability. *Climate Dynamics*, *32*(1), 75–93. https://doi.org/10.1007/s00382-008-0375-9

Cronin, M. F., Gentemann, C. L., Edson, J., Ueki, I., Bourassa, M., Brown, S., et al. (2019). Air-Sea Fluxes With a Focus on Heat and Momentum. *Frontiers in Marine Science*, *6*, 430. https://doi.org/10.3389/fmars.2019.00430

Danabasoglu, G., Lamarque, J. -F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019MS001916

Davis, C., Brown, B., & Bullock, R. (2006a). Object-Based Verification of Precipitation Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas. *Monthly Weather Review*, *134*(7), 1772–1784. https://doi.org/10.1175/MWR3145.1

Davis, C., Brown, B., & Bullock, R. (2006b). Object-Based Verification of Precipitation Forecasts. Part II: Application to Convective Rain Systems. *Monthly Weather Review*, *134*(7), 1785–1795. https://doi.org/10.1175/MWR3146.1

Day, J. J., Tietsche, S., & Hawkins, E. (2014). Pan-Arctic and Regional Sea Ice Predictability: Initialization Month Dependence. *Journal of Climate*, *27*(12), 4371–4390. https://doi.org/10.1175/JCLI-D-13-00614.1

Day, J. J., Hawkins, E., & Tietsche, S. (2014). Will Arctic sea ice thickness initialization improve seasonal forecast skill? *Geophysical Research Letters*, *41*(21), 7566–7575. https://doi.org/10.1002/2014GL061694

Day, Jonathan J., Tietsche, S., Collins, M., Goessling, H. F., Guemas, V., Guillory, A., et al. (2016). The Arctic Predictability and Prediction on Seasonal-to-Interannual TimEscales (APPOSITE) data set version 1. *Geoscientific Model Development*, *9*(6), 2255–2270. https://doi.org/10.5194/gmd-9-2255-2016

De Boisséson, E., & Balmaseda, M. A. (2024). Predictability of marine heatwaves: assessment based on the ECMWF seasonal forecast system. *Ocean Science*, *20*(1), 265–278. https://doi.org/10.5194/os-20-265-2024

De Coëtlogon, G., & Frankignoul, C. (2003). The Persistence of Winter Sea Surface Temperature in the North Atlantic. *Journal of Climate*, *16*(9), 1364–1377. https://doi.org/10.1175/1520-0442-16.9.1364

DeHaan, L. L., Martin, A. C., Weihs, R. R., Delle Monache, L., & Ralph, F. M. (2021). Object-Based Verification of Atmospheric River Predictions in the Northeast Pacific. *Weather and Forecasting*, *36*(4), 1575–1587. https://doi.org/10.1175/WAF-D-20-0236.1

Deser, C., Alexander, M. A., & Timlin, M. S. (2003). Understanding the Persistence of Sea Surface Temperature Anomalies in Midlatitudes. *Journal of Climate*, *16*(1), 57–72. https://doi.org/10.1175/1520-0442(2003)016<0057:UTPOSS>2.0.CO;2

Deser, C., Phillips, A. S., Alexander, M. A., Amaya, D. J., Capotondi, A., Jacox, M. G., & Scott, J. D. (2024). Future changes in the Intensity and Duration of Marine Heat and Cold Waves: Insights from Coupled Model Initial-Condition Large Ensembles. *Journal of Climate*. https://doi.org/10.1175/JCLI-D-23-0278.1

Deshayes, J., Curry, R., & Msadek, R. (2014). CMIP5 Model Intercomparison of Freshwater Budget and Circulation in the North Atlantic. *Journal of Climate*, *27*(9), 3298–3317. https://doi.org/10.1175/JCLI-D-12-00700.1

Ding, Q., Schweiger, A., L'Heureux, M., Battisti, D. S., Po-Chedley, S., Johnson, N. C., et al. (2017). Influence of high-latitude atmospheric circulation changes on summertime Arctic sea ice. *Nature Climate Change*, *7*(4), 289–295. https://doi.org/10.1038/nclimate3241

Ding, Q., Schweiger, A., & Baxter, I. (2022). Nudging Observed Winds in the Arctic to Quantify Associated Sea Ice Loss from 1979 to 2020. *Journal of Climate*, *35*(20), 3197–3213. https://doi.org/10.1175/JCLI-D-21-0893.1

Docquier, D., Vannitsem, S., Ragone, F., Wyser, K., & Liang, X. S. (2022). Causal Links Between Arctic Sea Ice and Its Potential Drivers Based on the Rate of Information Transfer. *Geophysical Research Letters*, *49*(9), e2021GL095892. https://doi.org/10.1029/2021GL095892

Dong, S., Gille, S. T., & Sprintall, J. (2007). An Assessment of the Southern Ocean Mixed Layer Heat Budget. *Journal of Climate*, *20*(17), 4425–4442. https://doi.org/10.1175/JCLI4259.1

Dorninger, M., Gilleland, E., Casati, B., Mittermaier, M. P., Ebert, E. E., Brown, B. G., & Wilson, L. J. (2018). The Setup of the MesoVICT Project. *Bulletin of the American Meteorological Society*, *99*(9), 1887–1906. https://doi.org/10.1175/BAMS-D-17-0164.1

Dörr, J., Årthun, M., Eldevik, T., & Madonna, E. (2021). Mechanisms of Regional Winter Sea-Ice Variability in a Warming Arctic. *Journal of Climate*, *34*(21), 8635–8653. https://doi.org/10.1175/JCLI-D-21-0149.1

Dörr, J., Bonan, D. B., Årthun, M., Svendsen, L., & Wills, R. C. J. (2023). Forced and internal components of observed Arctic sea-ice changes. *The Cryosphere*, *17*(9), 4133–4153. https://doi.org/10.5194/tc-17-4133-2023

Dörr, J., Årthun, M., Eldevik, T., & Sandø, A. B. (2024). Expanding Influence of Atlantic and Pacific Ocean Heat Transport on Winter Sea-Ice Variability in a Warming Arctic. *Journal of Geophysical Research: Oceans*, *129*(2), e2023JC019900. https://doi.org/10.1029/2023JC019900

Dunstone, N., Smith, D. M., Hardiman, S. C., Hermanson, L., Ineson, S., Kay, G., et al. (2023). Skilful predictions of the Summer North Atlantic Oscillation. *Communications Earth & Environment*, *4*(1), 409. https://doi.org/10.1038/s43247-023-01063-2

England, M. R., Polvani, L. M., Screen, J., & Chan, A. C. (2025). Minimal Arctic Sea Ice Loss in the Last 20 Years, Consistent With Internal Climate Variability. *Geophysical Research Letters*, *52*(15), e2025GL116175. https://doi.org/10.1029/2025GL116175

Fewings, M. R., & Brown, K. S. (2019). Regional Structure in the Marine Heat Wave of Summer 2015 Off the Western United States. *Frontiers in Marine Science*, *6*, 564. https://doi.org/10.3389/fmars.2019.00564

Frankignoul, C., & Hasselmann, K. (1977). Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability. *Tellus*, *29*(4), 289–305. https://doi.org/10.1111/j.2153-3490.1977.tb00740.x

Frankignoul, C., & Kestenare, E. (2002). The surface heat flux feedback. Part I: estimates from observations in the Atlantic and the North Pacific. *Climate Dynamics*, *19*(8), 633–647. https://doi.org/10.1007/s00382-002-0252-x

Frankignoul, C., Czaja, A., & L'Heveder, B. (1998). Air–Sea Feedback in the North Atlantic and Surface Boundary Conditions for Ocean Models. *Journal of Climate*, *11*(9), 2310–2324. https://doi.org/10.1175/1520-0442(1998)011<2310:ASFITN>2.0.CO;2

Frankignoul, C., Raillard, L., Ferster, B., & Kwon, Y.-O. (2024). Arctic September sea ice concentration biases in CMIP6 models and their relationships with other model variables. *Journal of Climate*. https://doi.org/10.1175/JCLI-D-23-0452.1

Frölicher, T. L., Fischer, E. M., & Gruber, N. (2018). Marine heatwaves under global warming. *Nature*, *560*(7718), 360–364. https://doi.org/10.1038/s41586-018-0383-9

Gao, Y., Kamenkovich, I., Perlin, N., & Kirtman, B. (2022). Oceanic Advection Controls Mesoscale Mixed Layer Heat Budget and Air–Sea Heat Exchange in the Southern Ocean. *Journal of Physical Oceanography*, *52*(4), 537–555. https://doi.org/10.1175/JPO-D-21-0063.1

Gilleland, E., Ahijevych, D., Brown, B. G., Casati, B., & Ebert, E. E. (2009). Intercomparison of Spatial Forecast Verification Methods. *Weather and Forecasting*, *24*(5), 1416–1430. https://doi.org/10.1175/2009WAF2222269.1

Gilleland, E., Ahijevych, D. A., Brown, B. G., & Ebert, E. E. (2010). Verifying Forecasts Spatially. *Bulletin of the American Meteorological Society*, *91*(10), 1365–1376. https://doi.org/10.1175/2010BAMS2819.1

Gregory, C. H., Artana, C., Lama, S., León-FonFay, D., Sala, J., Xiao, F., et al. (2024). Global Marine Heatwaves Under Different Flavors of ENSO. *Geophysical Research Letters*, *51*(20), e2024GL110399. https://doi.org/10.1029/2024GL110399

Gruber, N., Boyd, P. W., Frölicher, T. L., & Vogt, M. (2021). Biogeochemical extremes and compound events in the ocean. *Nature*, *600*(7889), 395–407. https://doi.org/10.1038/s41586-021-03981-7

Guemas, V., Chevallier, M., Déqué, M., Bellprat, O., & Doblas-Reyes, F. (2016). Impact of sea ice initialization on sea ice and atmosphere prediction skill on seasonal timescales. *Geophysical Research Letters*, *43*(8), 3889–3896. https://doi.org/10.1002/2015GL066626

Guemas, Virginie, Blanchard-Wrigglesworth, E., Chevallier, M., Day, J. J., Déqué, M., Doblas-Reyes, F. J., et al. (2016). A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, *142*(695), 546–561. https://doi.org/10.1002/qj.2401

Gunnarson, J. L., Stuecker, M. F., & Zhao, S. (2024). Drivers of future extratropical sea surface temperature variability changes in the North Pacific. *Npj Climate and Atmospheric Science*, *7*(1), 164. https://doi.org/10.1038/s41612-024-00702-5

Hartog, J. R., Spillman, C. M., Smith, G., & Hobday, A. J. (2023). Forecasts of marine heatwaves for marine industries: Reducing risk, building resilience and enhancing management responses. *Deep Sea Research Part II: Topical Studies in Oceanography*, *209*, 105276. https://doi.org/10.1016/j.dsr2.2023.105276

Hausmann, U., Czaja, A., & Marshall, J. (2016). Estimates of Air–Sea Feedbacks on Sea Surface Temperature Anomalies in the Southern Ocean. *Journal of Climate*, *29*(2), 439–454. https://doi.org/10.1175/JCLI-D-15-0015.1

Hausmann, U., Czaja, A., & Marshall, J. (2017). Mechanisms controlling the SST air-sea heat flux feedback and its dependence on spatial scale. *Climate Dynamics*, *48*(3–4), 1297–1307. https://doi.org/10.1007/s00382-016-3142-3

He, Z., Dai, A., Rose, B. E. J., & Vuille, M. (2024). Influence of the Atlantic and Pacific Multidecadal Variability on Arctic Sea Ice in Pacemaker Simulations during 1920–2013. *Journal of Climate*, *37*(17), 4481–4506. https://doi.org/10.1175/JCLI-D-23-0520.1

Hobday, A. J., Alexander, L. V., Perkins, S. E., Smale, D. A., Straub, S. C., Oliver, E. C. J., et al. (2016). A hierarchical approach to defining marine heatwaves. *Progress in Oceanography*, *141*, 227–238. https://doi.org/10.1016/j.pocean.2015.12.014

Hobday, A. J., Spillman, C. M., Allnutt, J., Coleman, M., Bailleul, F., Blamey, L., et al. (2024). Forecasting a Summer of Extremes: Building Stakeholder Response Capacity to Marine Heatwaves. *Oceanography*. https://doi.org/10.5670/oceanog.2024.508

Holbrook, N. J., Scannell, H. A., Sen Gupta, A., Benthuysen, J. A., Feng, M., Oliver, E. C. J., et al. (2019). A global assessment of marine heatwaves and their drivers. *Nature Communications*, *10*(1), 2624. https://doi.org/10.1038/s41467-019-10206-z

Holland, M. M., Landrum, L., Bailey, D., & Vavrus, S. (2019). Changing Seasonal Predictability of Arctic Summer Sea Ice Area in a Warming Climate. *Journal of Climate*, *32*(16), 4963–4979. https://doi.org/10.1175/JCLI-D-19-0034.1

Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., et al. (2021). Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1. *Journal of Climate*, *34*(8), 2923–2939. https://doi.org/10.1175/JCLI-D-20-0166.1

Huntington, H. P., Zagorsky, A., Kaltenborn, B. P., Shin, H. C., Dawson, J., Lukin, M., et al. (2022). Societal implications of a changing Arctic Ocean. *Ambio*, *51*(2), 298–306. https://doi.org/10.1007/s13280-021-01601-2

Jacox, M. G., Alexander, M. A., Bograd, S. J., & Scott, J. D. (2020). Thermal displacement by marine heatwaves. *Nature*, *584*(7819), 82–86. https://doi.org/10.1038/s41586-020-2534-z

Jacox, M. G., Alexander, M. A., Amaya, D., Becker, E., Bograd, S. J., Brodie, S., et al. (2022). Global seasonal forecasts of marine heatwaves. *Nature*, *604*(7906), 486–490. https://doi.org/10.1038/s41586-022-04573-9

Jing, Z., Wang, S., Wu, L., Chang, P., Zhang, Q., Sun, B., et al. (2020). Maintenance of mid-latitude oceanic fronts by mesoscale eddies. *Science Advances*, *6*(31), eaba7880. https://doi.org/10.1126/sciadv.aba7880

Kajtar, J. B., Bachman, S. D., Holbrook, N. J., & Pilo, G. S. (2022). Drivers, Dynamics, and Persistence of the 2017/2018 Tasman Sea Marine Heatwave. *Journal of Geophysical Research: Oceans*, *127*(8), e2022JC018931. https://doi.org/10.1029/2022JC018931

Kelly, K. A., Small, R. J., Samelson, R. M., Qiu, B., Joyce, T. M., Kwon, Y.-O., & Cronin, M. F. (2010). Western Boundary Currents and Frontal Air–Sea Interaction: Gulf Stream and Kuroshio Extension. *Journal of Climate*, *23*(21), 5644–5667. https://doi.org/10.1175/2010JCLI3346.1

Kirtman, B. P., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., et al. (2012). Impact of ocean model resolution on CCSM climate simulations. *Climate Dynamics*, *39*(6), 1303–1328. https://doi.org/10.1007/s00382-012-1500-3

Kohlman, C., Mogen, S., Cohen, J. T., Araujo, J., Beaudin, E., Bonino, G., et al. (2025). Surface and Subsurface Biogeochemical Impacts of the 2019 Northeast Pacific Marine Heatwave. *Geophysical Research Letters*, *52*(16), e2025GL115230. https://doi.org/10.1029/2025GL115230

Laurindo, L. C., Small, R. J., Thompson, L., Siqueira, L., Bryan, F. O., Chang, P., et al. (2022). Role of Ocean and Atmosphere Variability in Scale-Dependent Thermodynamic Air-Sea Interactions. *Journal of Geophysical Research: Oceans*, *127*(7), e2021JC018340. https://doi.org/10.1029/2021JC018340

Le Grix, N., Zscheischler, J., Rodgers, K. B., Yamaguchi, R., & Frölicher, T. L. (2022). Hotspots and drivers of compound marine heatwaves and low net primary production extremes. *Biogeosciences*, *19*(24), 5807–5835. https://doi.org/10.5194/bg-19-5807-2022

Lenetsky, J. E., Tremblay, B., Brunette, C., & Meneghello, G. (2021). Subseasonal Predictability of Arctic Ocean Sea Ice Conditions: Bering Strait and Ekman-Driven Ocean Heat Transport. *Journal of Climate*, *34*(11), 4449–4462. https://doi.org/10.1175/JCLI-D-20-0544.1

Li, L., Li, Y., & Li, Z. (2020). Object-based tracking of precipitation systems in western Canada: the importance of temporal resolution of source data. *Climate Dynamics*, *55*(9–10), 2421–2437. https://doi.org/10.1007/s00382-020-05388-y

Li, Y., Weijer, W., Kurtakoti, P., Veneziani, M., & Chang, P. (2024). Bering Strait Ocean Heat Transport Drives Decadal Arctic Variability in a High-Resolution Climate Model. *Geophysical Research Letters*, *51*(12), e2024GL108828. https://doi.org/10.1029/2024GL108828

Liu, G., Kwon, Y.-O., Frankignoul, C., & Lu, J. (2023). Understanding the Drivers of Atlantic Multidecadal Variability Using a Stochastic Model Hierarchy. *Journal of Climate*, *36*(4), 1043–1058. https://doi.org/10.1175/JCLI-D-22-0309.1

Liu, Y., Donat, Markus. G., England, Matthew. H., Alexander, Lisa. V., Hirsch, A. L., & Delgado-Torres, C. (2023). Enhanced multi-year predictability after El Niño and La Niña events. *Nature Communications*, *14*(1), 6387. https://doi.org/10.1038/s41467-023-42113-9

Lu, K., Danielson, S., Hedstrom, K., & Weingartner, T. (2020). Assessing the role of oceanic heat fluxes on ice ablation of the central Chukchi Sea Shelf. *Progress in Oceanography*, *184*, 102313. https://doi.org/10.1016/j.pocean.2020.102313

Lyman, J. M., & Johnson, G. C. (2023). Global High-Resolution Random Forest Regression Maps of Ocean Heat Content Anomalies Using In Situ and Satellite Data. *Journal of Atmospheric and Oceanic Technology*, *40*(5), 575–586. https://doi.org/10.1175/JTECH-D-22-0058.1

Manta, G., De Mello, S., Trinchin, R., Badagian, J., & Barreiro, M. (2018). The 2017 Record Marine Heatwave in the Southwestern Atlantic Shelf. *Geophysical Research Letters*, *45*(22). https://doi.org/10.1029/2018GL081070

Marin, M., Feng, M., Bindoff, N. L., & Phillips, H. E. (2022). Local Drivers of Extreme Upper Ocean Marine Heatwaves Assessed Using a Global Ocean Circulation Model. *Frontiers in Climate*, *4*, 788390. https://doi.org/10.3389/fclim.2022.788390

Maslowski, W., Clement Kinney, J., Higgins, M., & Roberts, A. (2012). The Future of Arctic Sea Ice. *Annual Review of Earth and Planetary Sciences*, *40*(1), 625–654. https://doi.org/10.1146/annurev-earth-042711-105345

Maxwell, S. M., Hazen, E. L., Lewison, R. L., Dunn, D. C., Bailey, H., Bograd, S. J., et al. (2015). Dynamic ocean management: Defining and conceptualizing real-time management of the ocean. *Marine Policy*, *58*, 42–50. https://doi.org/10.1016/j.marpol.2015.03.014

Meehl, G. A., Teng, H., Smith, D., Yeager, S., Merryfield, W., Doblas-Reyes, F., & Glanville, A. A. (2022). The effects of bias, drift, and trends in calculating anomalies for evaluating skill of seasonal-to-decadal initialized climate predictions. *Climate Dynamics*, *59*(11–12), 3373–3389. https://doi.org/10.1007/s00382-022-06272-7

Meier, Walt, Fetterer, F., Windnagel, A., Stewart, J. S., & Stafford, T. (2024). NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 5 [Data set]. National Snow and Ice Data Center. https://doi.org/10.7265/RJZB-PF78

Meier, Walter, & Stewart, J. S. (2023). Arctic and Antarctic Regional Masks for Sea Ice and Related Data Products [Data set]. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/CYW3O8ZUNIWC

Mills, K., Pershing, A., Brown, C., Chen, Y., Chiang, F.-S., Holland, D., et al. (2013). Fisheries Management in a Changing Climate: Lessons From the 2012 Ocean Heat Wave in the Northwest Atlantic. *Oceanography*, *26*(2). https://doi.org/10.5670/oceanog.2013.27

Mittermaier, M., & Bullock, R. (2013). Using MODE to explore the spatial and temporal characteristics of cloud cover forecasts from high-resolution NWP models. *Meteorological Applications*, *20*(2), 187–196. https://doi.org/10.1002/met.1393

Mittermaier, M., North, R., Maksymczuk, J., Pequignet, C., & Ford, D. (2021). Using feature-based verification methods to explore the spatial and temporal characteristics of the 2019 chlorophyll- *a* bloom season in a model of the European Northwest Shelf. *Ocean Science*, *17*(6), 1527–1543. https://doi.org/10.5194/os-17-1527-2021

Mogen, S. C., Lovenduski, N. S., Yeager, S., Keppler, L., Sharp, J., Bograd, S. J., et al. (2023). Skillful Multi-Month Predictions of Ecosystem Stressors in the Surface and Subsurface Ocean. *Earth's Future*, *11*(11), e2023EF003605. https://doi.org/10.1029/2023EF003605

Mogen, S. C., Lovenduski, N. S., Yeager, S. G., Capotondi, A., Jacox, M. G., Bograd, S., et al. (2024). Multi-month forecasts of marine heatwaves and ocean acidification extremes. *Nature Geoscience*. https://doi.org/10.1038/s41561-024-01593-0

Moreton, S., Ferreira, D., Roberts, M., & Hewitt, H. (2021). Air-Sea Turbulent Heat Flux Feedback Over Mesoscale Eddies. *Geophysical Research Letters*, *48*(20). https://doi.org/10.1029/2021GL095407

Muramatsu, M., Watanabe, E., Itoh, M., Onodera, J., Mizobata, K., & Ueno, H. (2025). Subsurface warming associated with Pacific Summer Water transport toward the Chukchi Borderland in the Arctic Ocean. *Scientific Reports*, *15*(1), 24. https://doi.org/10.1038/s41598-024-81994-8

Nonaka, M., & Xie, S.-P. (2003). Covariations of Sea Surface Temperature and Wind over the Kuroshio and Its Extension: Evidence for Ocean-to-Atmosphere Feedback*. *Journal of Climate*, *16*(9), 1404–1413. https://doi.org/10.1175/1520-0442(2003)16<1404:COSSTA>2.0.CO;2

Oldenburg, D., Kwon, Y.-O., Frankignoul, C., Danabasoglu, G., Yeager, S., & Kim, W. M. (2024). The Respective Roles of Ocean Heat Transport and Surface Heat Fluxes in Driving Arctic Ocean Warming and Sea Ice Decline. *Journal of Climate*, *37*(4), 1431–1448. https://doi.org/10.1175/JCLI-D-23-0399.1

Oliver, E. C. J., Benthuysen, J. A., Bindoff, N. L., Hobday, A. J., Holbrook, N. J., Mundy, C. N., & Perkins-Kirkpatrick, S. E. (2017). The unprecedented 2015/16 Tasman Sea marine heatwave. *Nature Communications*, *8*(1), 16101. https://doi.org/10.1038/ncomms16101

Oliver, E. C. J., Donat, M. G., Burrows, M. T., Moore, P. J., Smale, D. A., Alexander, L. V., et al. (2018). Longer and more frequent marine heatwaves over the past century. *Nature Communications*, *9*(1), 1324. https://doi.org/10.1038/s41467-018-03732-9

Pak, G., Park, Y.-H., Vivier, F., Bourdallé-Badie, R., Garric, G., & Chang, K.-I. (2017). Upper-ocean thermal variability controlled by ocean dynamics in the Kuroshio-Oyashio Extension region: KOE THERMAL VARIABILITY. *Journal of Geophysical Research: Oceans*, *122*(2), 1154–1176. https://doi.org/10.1002/2016JC012076

Park, D.-S. R., Lee, S., & Feldstein, S. B. (2015). Attribution of the Recent Winter Sea Ice Decline over the Atlantic Sector of the Arctic Ocean*. *Journal of Climate*, *28*(10), 4027–4033. https://doi.org/10.1175/JCLI-D-15-0042.1

Park, S., Deser, C., & Alexander, M. A. (2005). Estimation of the Surface Heat Flux Response to Sea Surface Temperature Anomalies over the Global Oceans. *Journal of Climate*, *18*(21), 4582–4599. https://doi.org/10.1175/JCLI3521.1

Patrizio, C. R., & Thompson, D. W. J. (2021). Quantifying the Role of Ocean Dynamics in Ocean Mixed Layer Temperature Variability. *Journal of Climate*, *34*(7), 2567–2589. https://doi.org/10.1175/JCLI-D-20-0476.1

Patrizio, C. R., & Thompson, D. W. J. (2022a). Observed Linkages Between the Atmospheric Circulation and Oceanic-Forced Sea-Surface Temperature Variability in the Western North Pacific. *Geophysical Research Letters*, *49*(8), e2021GL095172. https://doi.org/10.1029/2021GL095172

Patrizio, C. R., & Thompson, D. W. J. (2022b). Understanding the Role of Ocean Dynamics in Midlatitude Sea Surface Temperature Variability Using a Simple Stochastic Climate

Model. *Journal of Climate*, *35*(11), 3313–3333. https://doi.org/10.1175/JCLI-D-21-0184.1

Pearce, A. F., & Feng, M. (2013). The rise and fall of the “marine heat wave” off Western Australia during the summer of 2010/2011. *Journal of Marine Systems*, *111–112*, 139–156. https://doi.org/10.1016/j.jmarsys.2012.10.009

Qi, R., Zhang, Y., Du, Y., & Feng, M. (2022). Characteristics and Drivers of Marine Heatwaves in the Western Equatorial Indian Ocean. *Journal of Geophysical Research: Oceans*, *127*(10), e2022JC018732. https://doi.org/10.1029/2022JC018732

Qiu, B. (2002). The Kuroshio Extension System: Its Large-Scale Variability and Role in the Midlatitude Ocean-Atmosphere Interaction. *Journal of Oceanography*, *58*(1), 57–75. https://doi.org/10.1023/A:1015824717293

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *Journal of Climate*, *20*(22), 5473–5496. https://doi.org/10.1175/2007JCLI1824.1

Richter, J. H., Glanville, A. A., King, T., Kumar, S., Yeager, S. G., Davis, N. A., et al. (2024). Quantifying sources of subseasonal prediction skill in CESM2. *Npj Climate and Atmospheric Science*, *7*(1), 59. https://doi.org/10.1038/s41612-024-00595-4

Roach, L. A., & Blanchard-Wrigglesworth, E. (2022). Observed Winds Crucial for September Arctic Sea Ice Loss. *Geophysical Research Letters*, *49*(6), e2022GL097884. https://doi.org/10.1029/2022GL097884

Roberts, C. D., Palmer, M. D., Allan, R. P., Desbruyeres, D. G., Hyder, P., Liu, C., & Smith, D. (2017). Surface Flux and Ocean Heat Transport Convergence Contributions to Seasonal and Interannual Variations of Ocean Heat Content: Drivers of Heat Content Variability.

Journal of Geophysical Research: Oceans, *122*(1), 726–744. https://doi.org/10.1002/2016JC012278

Rossa, A., Nurmi, P., & Ebert, E. (2008). Overview of methods for the verification of quantitative precipitation forecasts. In S. Michaelides (Ed.), *Precipitation: Advances in Measurement, Estimation and Prediction* (pp. 419–452). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77655-0_16

Scannell, H. A., Johnson, G. C., Thompson, L., Lyman, J. M., & Riser, S. C. (2020). Subsurface Evolution and Persistence of Marine Heatwaves in the Northeast Pacific. *Geophysical Research Letters*, *47*(23), e2020GL090548. https://doi.org/10.1029/2020GL090548

Scannell, H. A., Cai, C., Thompson, L., Whitt, D. B., Gagne, D. J., & Abernathey, R. P. (2024). Spatiotemporal Evolution of Marine Heatwaves Globally. *Journal of Atmospheric and Oceanic Technology*. https://doi.org/10.1175/JTECH-D-23-0126.1

Schmeisser, L., Bond, N. A., Siedlecki, S. A., & Ackerman, T. P. (2019). The Role of Clouds and Surface Heat Fluxes in the Maintenance of the 2013–2016 Northeast Pacific Marine Heatwave. *Journal of Geophysical Research: Atmospheres*, *124*(20), 10772–10783. https://doi.org/10.1029/2019JD030780

Schmidtko, S., Johnson, G. C., & Lyman, J. M. (2013). MIMOC: A global monthly isopycnal upper-ocean climatology with mixed layers: MIMOC. *Journal of Geophysical Research: Oceans*, *118*(4), 1658–1672. https://doi.org/10.1002/jgrc.20122

Sen Gupta, A., Thomsen, M., Benthuysen, J. A., Hobday, A. J., Oliver, E., Alexander, L. V., et al. (2020). Drivers and impacts of the most extreme marine heatwave events. *Scientific Reports*, *10*(1), 19359. https://doi.org/10.1038/s41598-020-75445-3

Seo, H., O’Neill, L. W., Bourassa, M. A., Czaja, A., Drushka, K., Edson, J. B., et al. (2023). Ocean Mesoscale and Frontal-Scale Ocean–Atmosphere Interactions and Influence on Large-Scale Climate: A Review. *Journal of Climate*, *36*(7), 1981–2013. https://doi.org/10.1175/JCLI-D-21-0982.1

Serreze, M. C., Crawford, A. D., Stroeve, J. C., Barrett, A. P., & Woodgate, R. A. (2016). Variability, trends, and predictability of seasonal sea ice retreat and advance in the C hukchi S ea. *Journal of Geophysical Research: Oceans*, *121*(10), 7308–7325. https://doi.org/10.1002/2016JC011977

Shi, H., Jin, F.-F., Wills, R. C. J., Jacox, M. G., Amaya, D. J., Black, B. A., et al. (2022). Global decline in ocean memory over the 21st century. *Science Advances*, *8*(18), eabm3468. https://doi.org/10.1126/sciadv.abm3468

Shin, S., & Newman, M. (2021). Seasonal Predictability of Global and North American Coastal Sea Surface Temperature and Height Anomalies. *Geophysical Research Letters*, *48*(10), e2020GL091886. https://doi.org/10.1029/2020GL091886

Siqueira, L., Kirtman, B. P., Laurindo, L. C., Fasullo, J. T., & Hu, A. (2024). Quantifying the Role of Ocean Dynamics in SST Variability across GCMs and Observations. *Journal of Climate*, *37*(22), 5721–5737. https://doi.org/10.1175/JCLI-D-23-0686.1

Smale, D. A., Wernberg, T., Oliver, E. C. J., Thomsen, M., Harvey, B. P., Straub, S. C., et al. (2019). Marine heatwaves threaten global biodiversity and the provision of ecosystem services. *Nature Climate Change*, *9*(4), 306–312. https://doi.org/10.1038/s41558-019-0412-1

Small, R. J., deSzoeke, S. P., Xie, S. P., O'Neill, L., Seo, H., Song, Q., et al. (2008). Air–sea interaction over ocean fronts and eddies. *Dynamics of Atmospheres and Oceans*, *45*(3–4), 274–319. https://doi.org/10.1016/j.dynatmoce.2008.01.001

Small, R. J., Bryan, F. O., Bishop, S. P., & Tomas, R. A. (2019). Air–Sea Turbulent Heat Fluxes in Climate Models and Observational Analyses: What Drives Their Variability? *Journal of Climate*, *32*(8), 2397–2421. https://doi.org/10.1175/JCLI-D-18-0576.1

Small, R. J., Bryan, F. O., Bishop, S. P., Larson, S., & Tomas, R. A. (2020). What Drives Upper-Ocean Temperature Variability in Coupled Climate Models and Observations? *Journal of Climate*, *33*(2), 577–596. https://doi.org/10.1175/JCLI-D-19-0295.1

Smith, K. E., Burrows, M. T., Hobday, A. J., Sen Gupta, A., Moore, P. J., Thomsen, M., et al. (2021). Socioeconomic impacts of marine heatwaves: Global issues and opportunities. *Science*, *374*(6566), eabj3593. https://doi.org/10.1126/science.abj3593

Smith, K. E., Burrows, M. T., Hobday, A. J., King, N. G., Moore, P. J., Sen Gupta, A., et al. (2023). Biological Impacts of Marine Heatwaves. *Annual Review of Marine Science*, *15*(1), 119–145. https://doi.org/10.1146/annurev-marine-032122-121437

Song, X., Xie, X., Qiu, B., Cao, H., Xie, S.-P., Chen, Z., & Yu, W. (2022). Air-Sea Latent Heat Flux Anomalies Induced by Oceanic Submesoscale Processes: An Observational Case Study. *Frontiers in Marine Science*, *9*, 850207. https://doi.org/10.3389/fmars.2022.850207

Spillman, C. M., Hobday, A. J., Behrens, E., Feng, M., Capotondi, A., Cravatte, S., et al. (2025). What makes a marine heatwave forecast useable, useful and used? *Progress in Oceanography*, *234*, 103464. https://doi.org/10.1016/j.pocean.2025.103464

Stern, H. L. (2025). Regime Shift in Arctic Ocean Sea-Ice Extent. *Geophysical Research Letters*, *52*(8), e2024GL114546. https://doi.org/10.1029/2024GL114546

Strobach, E., Klein, P., Molod, A., Fahad, A. A., Trayanov, A., Menemenlis, D., & Torres, H. (2022). Local Air-Sea Interactions at Ocean Mesoscale and Submesoscale in a Western Boundary Current. *Geophysical Research Letters*, *49*(7), e2021GL097003. https://doi.org/10.1029/2021GL097003

Stroeve, J., Holland, M. M., Meier, W., Scambos, T., & Serreze, M. (2007). Arctic sea ice decline: Faster than forecast. *Geophysical Research Letters*, *34*(9), 2007GL029703. https://doi.org/10.1029/2007GL029703

Stroeve, J., Hamilton, L. C., Bitz, C. M., & Blanchard-Wrigglesworth, E. (2014). Predicting September sea ice: Ensemble skill of the SEARCH Sea Ice Outlook 2008-2013: Stroeve et al.: Predicting September sea ice. *Geophysical Research Letters*, *41*(7), 2411–2418. https://doi.org/10.1002/2014GL059388

Sun, D., Jing, Z., Li, F., & Wu, L. (2023). Characterizing global marine heatwaves under a spatio-temporal framework. *Progress in Oceanography*, *211*, 102947. https://doi.org/10.1016/j.pocean.2022.102947

Sun, X., & Wu, R. (2021). Seasonality and time scale dependence of the relationship between turbulent surface heat flux and SST. *Climate Dynamics*, *56*(9–10), 3173–3186. https://doi.org/10.1007/s00382-021-05631-0

Tomita, T., & Nonaka, M. (2006). Upper-Ocean Mixed Layer and Wintertime Sea Surface Temperature Anomalies in the North Pacific. *Journal of Climate*, *19*(2), 300–307. https://doi.org/10.1175/JCLI3616.1

Tommasi, D., Stock, C. A., Hobday, A. J., Methot, R., Kaplan, I. C., Eveson, J. P., et al. (2017). Managing living marine resources in a dynamic environment: The role of seasonal to decadal climate forecasts. *Progress in Oceanography*, *152*, 15–49. https://doi.org/10.1016/j.pocean.2016.12.011

Torres, O., Braconnot, P., Hourdin, F., Roehrig, R., Marti, O., Belamari, S., & Lefebvre, M. (2019). Competition Between Atmospheric and Surface Parameterizations for the Control of Air-Sea Latent Heat Fluxes in Two Single-Column Models. *Geophysical Research Letters*, *46*(13), 7780–7789. https://doi.org/10.1029/2019GL082720

Wang, L., Li, T., & Zhou, T. (2012). Intraseasonal SST Variability and Air–Sea Interaction over the Kuroshio Extension Region during Boreal Summer. *Journal of Climate*, *25*(5), 1619–1634. https://doi.org/10.1175/JCLI-D-11-00109.1

Whitt, D. B., Nicholson, S. A., & Carranza, M. M. (2019). Global Impacts of Subseasonal (<60 Day) Wind Variability on Ocean Surface Stress, Buoyancy Flux, and Mixed Layer Depth. *Journal of Geophysical Research: Oceans*, *124*(12), 8798–8831. https://doi.org/10.1029/2019JC015166

Woodgate, R. A. (2018). Increases in the Pacific inflow to the Arctic from 1990 to 2015, and insights into seasonal trends and driving mechanisms from year-round Bering Strait mooring data. *Progress in Oceanography*, *160*, 124–154. https://doi.org/10.1016/j.pocean.2017.12.007

Woodgate, R. A., Aagaard, K., & Weingartner, T. J. (2005). Monthly temperature, salinity, and transport variability of the Bering Strait through flow. *Geophysical Research Letters*, *32*(4), 2004GL021880. https://doi.org/10.1029/2004GL021880

Woodgate, R. A., Weingartner, T., & Lindsay, R. (2010). The 2007 Bering Strait oceanic heat flux and anomalous Arctic sea-ice retreat. *Geophysical Research Letters*, *37*(1), 2009GL041621. https://doi.org/10.1029/2009GL041621

Woodgate, R. A., Weingartner, T. J., & Lindsay, R. (2012). Observed increases in Bering Strait oceanic fluxes from the Pacific to the Arctic from 2001 to 2011 and their impacts on the Arctic Ocean water column. *Geophysical Research Letters*, *39*(24), 2012GL054092. https://doi.org/10.1029/2012GL054092

Wu, R., Kirtman, B. P., & Pegion, K. (2006). Local Air–Sea Relationship in Observations and Model Simulations. *Journal of Climate*, *19*(19), 4914–4932. https://doi.org/10.1175/JCLI3904.1

Yang, P., Jing, Z., Sun, B., Wu, L., Qiu, B., Chang, P., et al. (2021). On the Upper-Ocean Vertical Eddy Heat Transport in the Kuroshio Extension. Part II: Effects of Air-Sea Interactions. *Journal of Physical Oceanography*. https://doi.org/10.1175/JPO-D-21-0013.1

Yeager, S. G. (2022). Seasonal-to-Multiyear Large Ensemble (SMYLE) Experiment [Data set]. UCAR/NCAR - CISL - CDP. https://doi.org/10.26024/PWMA-RE41

Yeager, S. G., Karspeck, A. R., Danabasoglu, G., Tribbia, J., & Teng, H. (2012). A Decadal Prediction Case Study: Late Twentieth-Century North Atlantic Ocean Heat Content. *Journal of Climate*, *25*(15), 5173–5189. https://doi.org/10.1175/JCLI-D-11-00595.1

Yeager, S. G., Karspeck, A. R., & Danabasoglu, G. (2015). Predicted slowdown in the rate of Atlantic sea ice loss. *Geophysical Research Letters*, *42*(24). https://doi.org/10.1002/2015GL065364

Yeager, S. G., Rosenbloom, N., Glanville, A. A., Wu, X., Simpson, I., Li, H., et al. (2022). The Seasonal-to-Multiyear Large Ensemble (SMYLE) prediction system using the Community Earth System Model version 2. *Geoscientific Model Development*, *15*(16), 6451–6493. https://doi.org/10.5194/gmd-15-6451-2022

Yu, L., & Weller, R. A. (2007). Objectively Analyzed Air–Sea Heat Fluxes for the Global Ice-Free Oceans (1981–2005). *Bulletin of the American Meteorological Society*, *88*(4), 527–540. https://doi.org/10.1175/BAMS-88-4-527

Yuan, X., Kaplan, M. R., & Cane, M. A. (2018). The Interconnected Global Climate System—A Review of Tropical–Polar Teleconnections. *Journal of Climate*, *31*(15), 5765–5792. https://doi.org/10.1175/JCLI-D-16-0637.1