



Decadal biogeochemical predictions for the changing bottom marine environment of the Gulf of Maine

Vimal Koul^{1,a}, Andrew C. Ross², Charles Stock², Liping Zhang², Andrew T. Wittenberg², and Thomas Delworth²

¹Cooperative Institute for Modeling the Earth System, Princeton University, Princeton, New Jersey, United States

²NOAA/OAR/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, United States

^anow at: Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, New Delhi, India

Correspondence: Vimal Koul (vkoul@iitd.ac.in)

Received: 27 January 2026 – Discussion started: 12 February 2026

Revised: 27 July 2026 – Accepted: 31 August 2026 – Published: 5 October 2026

Abstract. The Gulf of Maine and the surrounding Northeast US Continental Shelf are experiencing rapid marine environmental change arising from complex regional dynamics that challenge near-term (1–10 years) predictive capabilities for valuable living marine resources. Here, using a high-resolution regional ocean model, we demonstrate skillful decadal forecasts of ocean bottom habitat characteristics including bottom temperature, dissolved oxygen (O_2), pH and aragonite saturation state (Ω_{ar}). Bottom temperature and pH predictions show substantial skill driven primarily by radiatively forced change and carbon uptake trends, while bottom O_2 and Ω_{ar} predictions benefit more from initialization due to stronger internal variability. Retrospective forecasts successfully predicted observed historical changes in water masses and environmental properties, including recent cooling/freshening transitions driven by replacement of Warm Slope Water with Labrador Slope Water. This water mass variability also modulates biogeochemical conditions and ocean acidification buffering capacity, with our recent forecasts indicating that benefits from the expected respite from rapid warming might be tempered by challenges posed by rapid acidification. The demonstrated predictability of coupled physical-biogeochemical processes supports developing integrated prediction systems for climate-informed marine resource management.

absorbing approximately 30 % of anthropogenic CO_2 emissions, and over 90 % of the excess heat associated with accumulating greenhouse gases (Sabine et al., 2004; Feely et al., 2023, Friedlingstein et al., 2025, Johnson and Lyman, 2020). This oceanic CO_2 and heat uptake has initiated a cascade of interconnected changes that represent some of the most profound shifts in ocean conditions experienced in the past 50 years. Warming has altered ocean stratification and circulation (Levitus et al., 2012; Abram et al., 2016). Acidification has challenged the basic chemistry upon which marine calcifiers depend (Caldeira and Wickett, 2003; Orr et al., 2005). Declining oxygen concentrations have pressed the metabolic limits of marine organisms (Keeling et al., 2010; Breitburg et al., 2018). Ecosystem and fisheries challenges posed by climate change are compounded by those linked to natural climate variability (Lehodey et al., 2006). A capacity to anticipate physical and biogeochemical changes, whether from climate change or natural variability, may help coastal communities and industries mitigate potential negative impacts, and realize new opportunities that may arise (Clark et al., 2001; Tommasi et al., 2017b).

The Northeast US Continental Shelf (NEUS) – and particularly its highly productive semi-enclosed sea, the Gulf of Maine (GOM) – has been subject to both ocean change trends and pronounced variability. Dynamic warming (Pershing et al., 2015; Saba et al., 2016; Alexander et al., 2018), deoxygenation (Claret et al., 2018) and acidification (Salisbury and Jönsson, 2018; Siedlecki et al., 2021; Stewart et al., 2025) signals have emerged from an interplay of shifting circulation patterns, alterations in the Gulf Stream position, and regional water masses, particularly variations between

1 Introduction

The global ocean has undergone unprecedented biogeochemical transformation since the onset of the industrial era –

cold, fresh Labrador Slope Water (LSW) and Warm Slope Water (WSW) associated with greater Gulf Stream influence, that modulate global warming signals through local oceanographic processes (Chen et al., 2020; Gangopadhyay et al., 2019; Brickman et al., 2018; Gonçalves Neto et al., 2021; Friedland and Hare, 2007; Seidov et al., 2021). In the GOM, oceanic variability from interannual to longer time scales is understood to be largely driven by changes in the contributions of these and other source waters which, in turn, modify the temperature, salinity, and nutrient conditions across the entire ecosystem (Mountain, 2012; Townsend et al., 2023). This water mass variability has also generated complex and sometimes counterintuitive local biogeochemical responses to climate change, including an influx of high-alkalinity Gulf Stream waters driving increased buffering to acidification in certain areas despite rising atmospheric CO₂ concentrations (Stewart et al., 2025; Salisbury and Jönsson, 2018). Such water mass changes may also influence coastal hypoxia risk by altering stratification patterns that control vertical mixing and oxygen dynamics in coastal bottom waters (Scully et al., 2022).

Rapid warming of the GOM has already contributed to the decline of its iconic Atlantic cod fishery and triggered cascading ecosystem changes (Pershing et al., 2015, 2021; Hare et al., 2016). Continued warming, deoxygenation, and acidification pose significant threats to commercially important fisheries, particularly American lobster and shellfish. American lobsters show stage-specific vulnerabilities to acidification and warming that vary significantly between regional populations (Harrington and Hamlin, 2019; Niemisto et al., 2025). Early life stages of shellfish consistently show greater sensitivity to acidification than adults, creating recruitment bottlenecks that may drive population declines (Fay et al., 2017). Impacts extend beyond commercially harvested species to affect protected species such as Atlantic puffins and northern right whales (Meyer-Gutbrod et al., 2021).

Decisions by coastal communities and managers that determine long-term resource and community resilience are made across time horizons and must consider both variability and change. The interannual to decadal time horizon (1–10 years) is particularly notable for its relevance to the setting of annual catch limits, the formulation of stock rebuilding plans, and industry capitalization (Tommasi et al., 2017b). Global Earth System Models (ESMs) have demonstrated multiyear predictability for surface temperature, pH, dissolved oxygen, and net primary production, and enhanced subsurface predictability extending to a decade in the North Atlantic where Atlantic Meridional Overturning Circulation (AMOC) variability strongly influences physical drivers (Frölicher et al., 2020; Krumhardt et al., 2020; Park et al., 2019; Msadek et al., 2014). This ocean predictability enables predictability of the distribution of commercially important open-ocean fish species in the North Atlantic over decadal time scales (Payne et al., 2022).

Fisheries are concentrated on productive and accessible continental shelves, however, where direct applications of global ESMs are limited by coarse spatial resolution (~100–200 km) that cannot adequately represent the complex bathymetry, coastal dynamics, tidal mixing, and mesoscale circulation features that drive ocean and ecosystem variability on the shelf (Tommasi et al., 2017b; Saba et al., 2016; Ross et al., 2023). High-resolution ocean modeling and enhanced observational networks are revealing substantial promise for predicting ocean variability in this highly dynamic region (Gehlen et al., 2015; Ross et al., 2023, 2024; Chen et al., 2021, 2024; Koul et al., 2024).

In the GOM (Fig. 2a), recent observations show cooler, fresher water entering the deeper regions through the Northeast Channel beginning in late 2023, potentially marking the return of cooling LSW influence after over a decade of warm Gulf Stream dominance (Record et al., 2024). Previous work regional ocean prediction studies demonstrated that this physical transition, and others like it, were predictable (Koul et al., 2024). Such shifts, however, also have significant biogeochemical implications, as LSW offers lower buffering against acidification than Gulf Stream water. If the recent cooling represents a sustained return to LSW dominance, the region could face rapid acidification as the reduced buffering will compound ongoing background acidification trends (Salisbury and Jönsson, 2018; Stewart et al., 2025). Such rapid biogeochemical change could challenge marine ecosystems and shellfish industries adapted to a decade of warm yet less acidic conditions.

Here, we build on prior physical ocean prediction studies and present a high-resolution regional physical and biogeochemical prediction system that, when used to downscale global decadal predictions (Delworth et al., 2020), addresses this time-critical challenge by capturing both the physical and biogeochemical dynamics of shifting water masses and persistent change. We furthermore shift the focus of the evaluation to the benthic environment, which is critical for the largest and most lucrative Gulf of Maine fisheries (e.g., lobster, scallops and groundfish, Waller et al., 2023). Through a comprehensive 10-member ensemble approach spanning 60 years of initialized retrospective predictions, accompanied by uninitialized simulations that include only the impacts of external radiative forcings (~6500 years of simulations), we demonstrate skilful decadal (1–10 year) predictions of benthic physical and biogeochemical conditions in the Gulf of Maine. Our experiments allow us to isolate the impacts of radiatively forced ocean warming on prediction skill, and the impacts of predictable water mass variations. Finally, we present probabilistic forecasts for the coming years that may offer fisheries managers actionable insights into the near-term trajectory of the benthic environment in the GOM.

Table 1. Overview of the three types of downscaled experiments presented in this study, focusing on the differences between the experiments.

	NWA12_EN4_NUDG	NWA12_HIND	NWA12_HIST
Simulation type	Continuously running simulation, 1960–2024	10 year predictions	Continuously running simulation, 1960–2034
Purpose	Proxy for actual ocean state, initial conditions for forecasts	Forecast skill assessment	Isolates predicted effect of greenhouse gas forcing
Number of ensemble members	10	10	10
Atmospheric forcing and open boundary conditions	SPEAR_LO global coupled reanalysis	SPEAR_HIND decadal predictions	SPEAR_HIST uninitialized simulations
Initial conditions	SPEAR_LO global coupled reanalysis followed by spin up	NWA12_EN4_NUDG	NWA12_EN4_NUDG
Nudging	EN4 temperature and salinity, waters > 1000 m	SPEAR_HIND near northern boundary	SPEAR_HIST near northern boundary

2 Methods

Several sets of regional ocean simulations were used to study the predictability of physical and biogeochemical changes in the Gulf of Maine. These are summarized here and in Table 1 to provide a broad view of the experiment structure. Further details of each component are then provided in the subsections that follow.

First, a large set of 10-member ensemble retrospective decadal (10-yearlong) forecasts, or “hindcasts”, were generated starting each January from 1965 to 2025 by downscaling ensemble predictions from the coarse resolution Seamless System for Prediction and Earth System Research (SPEAR) decadal prediction system (Delworth et al., 2020; Yang et al., 2021) to the 1/12° Northwest Atlantic (NWA12) regional ocean model of Ross et al. (2023). These predictions, which comprise nearly 6500 years of simulation, are referred to herein as NWA12_HIND, and the global SPEAR hindcasts used to force them are referred to as SPEAR_HIND.

NWA12_HIND hindcasts were initialized with a 10-member ensemble of retrospective NWA12 simulations forced by the same SPEAR ensemble reanalysis used for SPEAR_HIND (Yang et al., 2021). The regional runs were furthermore nudged gently toward the EN4 in waters > 1000 m. These simulations, which will also aid in prediction analysis, are referred to as NWA12_EN4_NUDG.

Finally, an ensemble of simulations of the historical effects of climate change drivers (increasing greenhouse gases, aerosol changes, land use) was generated by downscaling the historical and early projection period of an ensemble of SPEAR climate change simulations. Unlike the predictions in NWA12_HIND and SPEAR_HIND, these climate change runs were not initialized from a data constrained ocean state, but from alternative states from the SPEAR pre-industrial control simulation. The mean of this ensemble is used to estimate the signal associated with climate change, or the

“forced” signal, over natural variations. These runs are referred to as “NWA12_HIST” to indicate that they are derived from the historical portion of a climate change simulation. They are also referred to as “uninitialized” runs to indicate that, unlike predictions that aim to anticipate both climate variations and change, they are not initiated from a data-constrained initial state.

2.1 The 1/12° regional MOM6-COBALT-NWA12 model

The regional downscaled model developed here inherits most of the configuration and parameters from the MOM6-COBALT-NWA12 v1.0 model developed by Ross et al. (2023) and further refined for applications to decadal predictions by Koul et al. (2024). The model comprises coupled components for ocean physics, biogeochemistry, and sea-ice. Compared to the previously published decadal predictions, we add the COBALT biogeochemical model (Stock et al., 2020, 2025) and improve the initial physical state estimates by relaxing the downscaled simulation towards the EN4 dataset (Good et al., 2013) at depth greater than 1000 m.

2.2 The global SPEAR model

We use a 10-member SPEAR-based global simulation (SPEAR_LO; Delworth et al., 2020; Yang et al., 2021) with a low-resolution configuration (1° ocean and sea ice with near-equatorial refinement; and 1° atmosphere and land) to obtain surface forcing and open boundary conditions for the regional model. We use forcing and boundary conditions from the SPEAR_LO global coupled reanalysis to create the NWA12 simulations used to initialize our hindcasts (i.e., NWA12_EN4_NUDG, Table 1), and forcing from SPEAR hindcasts (SPEAR_HIND) to drive NWA12_HIND. For the SPEAR_LO reanalysis, atmospheric winds and temperature were restored toward the 6-hourly Japanese Reanal-

ysis (JRA3Q; Kosaka et al., 2024), and the oceanic SST is restored toward the interpolated monthly-mean observations from the NOAA Extended Reconstructed Sea Surface Temperature version 5 (ERSSTv5) dataset (Huang et al., 2017). The reanalysis ensemble spread originates through alternative deep ocean initializations and is maintained due to the gentle surface-only assimilation approach used by SPEAR_LO (Yang et al., 2021).

2.3 Regional downscaled simulation (NWA12_EN4_NUDG)

SPEAR_LO reanalysis output was used to force the regional MOM6 model at the ocean surface and at the three open ocean boundaries (see Koul et al., 2024) from January 1960 to December 2024. The 10 members of this regional downscaled simulation (NWA12_EN4_NUDG) are derived from the 10 members of the SPEAR_LO: we simply carry forward the member-wise information present in the ocean initial conditions, atmospheric forcing and open ocean boundary conditions.

The interior solution of the NWA12_EN4_NUDG simulation was furthermore relaxed with a 90 d timescale towards the full vertical profiles of ocean temperature and salinity from the EN4 gridded dataset in regions where the ocean bottom is deeper than 1000 m. This moderate relaxation helps keep the model solution closer to the observation-based sub-surface variability, particularly in the late 1990s when the boundary conditions overestimated the propagation of a cool, fresh signal into the domain. Limiting the relaxation to waters deeper than 1000 m helps avoid rapid oceanic re-adjustment in shallower waters upon initialization of retrospective predictions wherein such relaxation towards EN4 is not carried out and is necessary to avoid nudging on the shelf where the 1° EN4 shows low correlation with observed salinity despite maintaining high correlation with observed temperature (Figs. S2 and S3). We also note that the bottom waters of the Gulf of Maine analysed in this study (the deep basins and the Northeast Channel) lie well inshore of the 1000 m isobath and are therefore not directly nudged towards EN4. The relaxation constrains only the off-shelf slope source waters. We do not relax the regional model towards the parent SPEAR model, since SPEAR has coarse ocean resolution and does not assimilate sub-surface ocean data, making it less constrained towards observations than the EN4-gridded dataset.

Open boundary data for biogeochemical tracers were obtained exactly as explained in Ross et al. (2023). The time varying boundary data for alkalinity and dissolved inorganic carbon were estimated from SPEAR_LO monthly average temperature and salinity as input to the algorithm from Carter et al. (2021). Climatological river freshwater discharge, nutrient, carbon and alkalinity data were also used as described in Ross et al. (2023) and Koul et al. (2024). This climatological forcing, which includes seasonal variability but not

interannual variability, was used because coastwide river observations dating back to the 1960s are not available. In the Gulf of Maine, direct river runoff is estimated to be a minor source of freshwater and nutrients (Smith et al., 2001; Townsend, 1998) and most of the interannual variability is driven by changing sources of inflowing water.

2.4 Regional ocean hindcasts (NWA12_HIND)

Following Koul et al. (2024), we use the regional simulation from the NWA12_EN4_NUDG model to start 10-yearlong retrospective ocean physical-biogeochemical predictions (NWA12_HIND) every year on 1 January from 1965 to 2025. The last 10-yearlong simulation, which shifts from hindcast to forecast, is from 1 January 2025 to 31 December 2034. The 10-member NWA12_HIND ensemble are started from ocean initial conditions taken from the corresponding 10 members of NWA12_EN4_NUDG, and integrated forward using atmospheric forcing and open ocean boundary conditions from the 10 members of SPEAR_HIND. To ease the transition between the global and regional model of the complex inflow and outflow along the northern boundary, each member of the 10-yearlong predictions is relaxed towards the corresponding SPEAR_LO prediction north of 45° N. The strength of the restoration corresponds to a relaxation time scale of 90 d at the northern boundary and decreases gradually to 0 at 45° N. No bathymetric mask is applied, as the relaxation drops to zero north of the targeted shallower regions. The latitudinal damping rate r (inverse of relaxation time scale) was prescribed as:

$$r(\phi) = (r_{\max}/2) \times [1 + \cos(\pi(\phi_{\max} - \phi)/(\phi_{\max} - 45))] \quad (1)$$

for ϕ (latitude) from 45° N to $\phi_{\max}$ in the regional model, and zero south of 45° N and $r_{\max} = 1/90 \text{ d}^{-1}$ at $\phi_{\max}$, where $\phi_{\max}$ in the regional model is $\approx 58^\circ \text{ N}$.

2.5 Historical climate change simulations (NWA12_HIST)

The NWA12_HIST is a continuous 75-yearlong simulation started on 1 January 1960 and ending in 31 December 2034. The NWA12_HIST is also a 10-member ensemble forced with atmospheric forcing and open ocean boundary conditions from the 10 members of SPEAR_HIST. The SPEAR_HIST dataset spans from 1850 to 2100 and incorporates future greenhouse gas concentrations based on the SSP585 emissions scenario (Riahi et al., 2017). It should be noted that differences between SSP585 and other emissions scenarios are negligible over the near future (next 1–10 years) examined in this study. Both NWA12_HIST and NWA12_HIND share the response of the model to the common prescribed natural and anthropogenic radiative forcing but the phase of natural variability in NWA12_HIST is not expected to match the observed phase since the impact of 1960 initial conditions is not retained beyond a few years.

2.6 Point observations and observation-based gridded products

Two observation-based gridded data sets were used to assess the quality of model simulations for bottom temperature and salinity: the 1/12° GLORYS12 v1 ocean reanalysis and salinity products (Lellouche et al., 2018) and the 1° EN4 gridded dataset (EN4.2.2, Good et al., 2013) which is used to assess a longer period from 1965–2024 (Fig. S1 in the Supplement, Figs. 3 and 5). For bottom O_2 and bottom Ω_{ar} , we used the Fishing Industry SHared Bottom Oceanographic Timeseries data (https://erddap.ondeckdata.com/erddap/tabledap/fishbot_realtime.html, last access: 29 May 2025), hereafter called “Fishbot”, and vessel (Jiang et al., 2021) and glider-based (IOOS Glider Data Assembly Center: <https://gliders.ioos.us/erddap/index.html>, last access: 29 May 2025) datasets (hereafter called “NOAA OBS”). In addition to these two datasets, we also used the observed bottom temperature and bottom salinity time series (hereafter called “NOAA_SOE”) for the GOM Ecological Production Unit (EPU) from NOAA’s State of the Ecosystem reports (NOAA Fisheries, 2025).

2.7 Preparation of annual mean anomalies

Area-weighted averages were first computed for the model-simulated ensemble mean and the two gridded observational datasets (GLORYS12 and EN4), excluding grid cells with ocean depths less than 30 m. Annual anomalies were then derived by subtracting the 1994–2023 mean from the annual mean time series. Figures S1–S3 show the comparison of annual anomalies from variables across these gridded datasets.

For all Fishbot and NOAA_OBS observations, point data was first linearly interpolated onto a uniform $1/4^\circ \times 1/4^\circ$ grid before area averaging. Monthly anomalies were then computed by subtracting the long-term monthly mean (1994–2023) from each month, provided there were at least three years of data samples. Finally, annual mean anomalies were calculated by averaging monthly anomalies for years with at least four months of data. The uncertainty (error bars in Fig. 2) for that year is $\sigma_{\text{monthly}}/\sqrt{n}$, i.e. the sample standard deviation of that year’s monthly anomalies divided by the square root of the number of valid months (n).

2.8 Gulf Stream index and water mass definitions

The variability of the Gulf Stream’s position was assessed using a method established by Pérez-Hernández and Joyce (2014) and subsequently utilized by Ross et al. (2023). First, the Gulf Stream’s mean position is defined by the latitude where sea surface height (SSH) variance is highest along meridional lines, spaced 1° apart, from 72° to 52° W. Then, to quantify the relative position of the Gulf Stream, the Gulf Stream Index (GSI) was calculated as the annual SSH anomaly, relative to the 1994–2023 average, averaged

over these mean position points. The proportion of the two water masses, Labrador Slope Water and Warm Slope Water, in the Northeast Channel (Fig. 1) analyzed in this study were computed from water mass definitions as described in Ross et al. (2023).

2.9 Skill assessment

We used anomaly correlation coefficients (ACC) to quantify the skill in predicting annual mean anomalies of bottom temperature, bottom O_2 , bottom pH and bottom Ω_{ar} . The ACC for area-averaged quantities was calculated for the Gulf of Maine Ecological Production Units (GOM_EPU, NOAA Fisheries, 2024). For bottom temperature, bottom O_2 , bottom pH and bottom Ω_{ar} the NWA12_EN4_NUDG simulation was used as the observational reference, and the skill was calculated over the period 1965–2024. The lead-dependent climatology (1994–2023 mean) was removed from the initialized hindcasts before calculating the skill metric. This choice of using NWA12_EN4_NUDG as the observational reference was made because continuous spatio-temporal observational records of bottom temperature, O_2 , pH and bottom Ω_{ar} , for any of the EPUs or Large Marine Ecosystems (LMEs) were not available. However, to show that the NWA12_EN4_NUDG was a good proxy for in-situ observations for the skill estimation, we carried out a point-by-point comparison of temporally and spatially co-located model output and available observations (Fig. S7), as well as additional evaluation of interannual variations and covariance with better-constrained physical properties (Figs. S1, 3 and 5). A simple persistence-based skill, defined as the lagged autocorrelation of observed anomalies, was also compared with the hindcast skill.

We assessed the statistical significance of the prediction skill from the initialized NWA12_HIND retrospective predictions. This was done by conducting a two-tailed t-test against a null hypothesis of no correlation, adjusting for the reduced degrees of freedom due to autocorrelation. Additionally, we compared this skill to that of persistence forecasts and the uninitialized (NWA12_HIST) simulation. For this comparison, we employed the Williams (1959) method as described in Steiger (1980) to determine the statistical significance between two dependent correlation coefficients.

2.10 Prediction probabilities and their skill

Probabilistic tercile predictions are developed by post-processing the model forecasts using extended logistic regression (Wilks, 2009). Extended logistic regression fits the probability of the outcome y exceeding a certain threshold q given some data x , and is formulated as

$$P\{y > q|x\} = \sigma(f(x) + g(q)) \quad (2)$$

where the logistic function $\sigma(z) = 1/(1+e^{-z}) = e^z/(1+e^z)$. Here, f is a simple linear function of the ensemble mean pre-

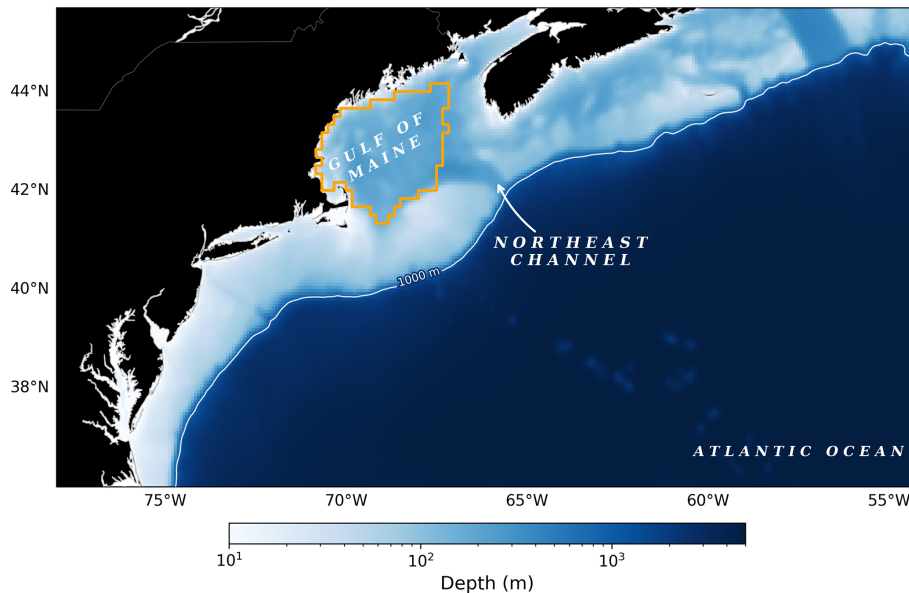


Figure 1. The study region and its characteristics. The locations of the Gulf of Maine Ecological Production Unit (GOM_EPU) and the Northeast Channel along the Northeast US continental shelf are shown along with the model bathymetry and the 1000 m isobath (white line).

diction x : $f(x) = \beta_0 + \beta_1 \cdot x$, and the extension function g is also a simple function of the threshold q : $g(q) = \beta_2 \cdot q$. Maximum likelihood estimation, following Minka (2001) and Pregibon (1981), is used to fit the three beta parameters for each of the four forecast variables in Fig. 7. In this method, the forecast probability is a function of only the ensemble mean, and not the individual ensemble members or the ensemble spread. As is common in seasonal to decadal dynamical model predictions, we find that the raw ensembles tend to be overconfident, with insufficient ensemble spread (Fig. S8). Furthermore, the ensemble spread has low correlation with the actual retrospective forecast error. As a result, extended logistic regression provides a reasonable improvement to the estimates of the forecast uncertainty.

For evaluating probabilistic predictions, we use accuracy as the verification metric for each tercile. The accuracy is calculated as:

$$\text{Accuracy} = \frac{(\text{True Positives} + \text{True Negatives})}{\text{Total Forecasts}}. \quad (3)$$

This metric is computed independently for each tercile category. In this context, a true positive occurs when the tercile with the highest predicted probability correctly matches the observed tercile. A true negative occurs when the model correctly assigns the highest probability to either of the other two terciles (effectively predicting that observations will not fall within the tercile being evaluated).

We also calculated the ranked probability skill score (RPSS) for each variable to assess the accuracy of the predictions across all three categories. Persistence of the previous year's category was used as the reference for

the RPSS, so positive skill scores indicate that the model forecasts have some ability to predict transitions away from recent conditions.

The RPSS is a skill score based on the Ranked Probability Score (RPS), and the Ranked Probability Score (RPS) for a single forecast is defined as

$$\text{RPS} = \frac{1}{K-1} \sum_{k=1}^K \left[\sum_{i=1}^k P_i - \sum_{i=1}^k O_i \right]^2 \quad (4)$$

where K is the number of categories (3 in our case: below, normal, or above), P_i is the cumulative forecast probability for category i , and O_i is the cumulative observed probability (0 or 1) for category i . A perfect forecast, which assigns 100 % probability to the observed category, will have an $\text{RPS} = 0$. Predicting a 100 % chance of the opposite outcome (i.e., below when the outcome is above) gives an RPS of 1, while predicting normal when an outcome is above gives an RPS of 0.5. The RPS is notable in that it evaluates the distribution of the forecast probability around the outcome, penalizing outcomes that are further away. This makes RPS a more sensitive metric than accuracy.

The ranked probability skill score (RPSS):

$$\text{RPSS} = 1 - \text{RPS}(\text{forecast}) / \text{RPS}(\text{persistence}) \quad (5)$$

where $\text{RPS}(\text{forecast})$ is the sum of the ranked probability scores comparing the forecast probabilities with the observed categories, and $\text{RPS}(\text{persistence})$ is the sum of the ranked probability scores comparing the observed categories with the categories observed in the previous year (in other words, we assign 100 % probability to the category observed in the

year before the forecast was initialized, and compare that with the actual category for the next 3 years). By the definition of the skill score, if $RPSS = 1$, the forecasts had perfect accuracy at probabilistically predicting the observations. If $RPSS = 0$, the forecasts had the same overall score as the persistence forecasts of the previous year's state; in other words, the forecasts would have had no skill at predicting transitions.

2.11 Decomposition of the drivers of bottom O_2 alkalinity and aragonite saturation state (Ω_{ar})

Bottom O_2 saturation ($[O_2]$ sat), was computed from the simulated bottom potential temperature and salinity following Garcia and Gordon (1992). DIC and alkalinity were not saved at the ocean bottom and were thus obtained by solving the carbonate system (PyCO2SYS; Humphreys et al., 2022) from the model's bottom pH and Ω_{ar} together with bottom temperature and salinity at a representative pressure of 150 dbar. The result is largely insensitive to this pressure choice – recomputing Ω_{ar} from the recovered DIC and alkalinity reproduces the simulated Ω_{ar} to within 0.01. COBALT additionally applies minor nutrient alkalinity corrections (PO_4 and SiO_4) that are omitted here, with negligible effect on the bottom anomalies.

The contributions of bottom DIC, alkalinity, temperature (T) and salinity (S) to the bottom pH and Ω_{ar} anomalies were quantified by perturbing each variable in turn about the 1994–2023 mean state; the contributions sum to the total to within 0.0045 in pH and 0.008 in Ω_{ar} . To attribute their anomalies, we used a one-factor-at-a-time perturbation about a fixed reference state defined as the 1994–2023 climatological mean of each driver, (DIC_0 , Alk_0 , T_0 , S_0). The reference values pH_0 and Ω_{ar0} were obtained by solving the carbonate system at that state.

Four perturbation time series were constructed to estimate the contribution of each driver to interannual pH and Ω_{ar} anomalies by allowing a single driver to take its full annual-mean value while the other three were held fixed at their climatological mean. The approximate contribution of each driver to the pH (or Ω_{ar}) anomaly in year t is then the departure of its perturbed solution from the reference. For example, $\Delta pH_{DIC}(t) = pH(DIC(t), Alk_0, T_0, S_0) - pH_0$.

Because the carbonate system is non-linear, the four single-driver contributions are not guaranteed to add up exactly to the total anomaly; the difference is the residual cross-term arising from simultaneous variation of two or more drivers. In practice this residual is negligible here: the sum of the DIC, alkalinity, T and S contributions reproduces the full bottom pH and Ω_{ar} anomalies to within 0.0045 pH units and 0.008 in Ω_{ar} over the whole record.

3 Results and discussion

3.1 Bottom temperature and salinity

The regionally downscaled retrospective simulations, NWA12_EN4_NUDG, which were used to initialize retrospective forecasts, have an ensemble mean that matches the complex patterns of bottom warming and cooling in the GOM (Fig. 2a). The period from the 1980s to the mid-1990s shows a gradual warming of about 2°C followed by cooling of the same magnitude until mid-2000s, and intensified warming from the mid-2000s to around the year 2020. This bottom variability in the regional downscaled simulations is consistent with the observed long term surface temperature variations that the Northwest Atlantic Ocean has experienced over the past century, which have been characterized by distinct multi-decadal oscillations and dramatic shifts in water mass properties that have shaped marine ecosystems from the Labrador Sea to Cape Hatteras (Wallace et al., 2018; Townsend et al., 2006; Mountain, 2012). In the late 1990s, after a long period of persistent presence of Warm Slope Water (WSW), cold and fresh Labrador Slope Water (LSW) replaced WSW along the Scotian Shelf and GOM (Drinkwater et al., 1998; Mountain, 2004; Friedland and Hare, 2007). WSW increased in prevalence again during the period from 2004 to 2020 in conjunction with an unprecedented era of accelerated surface and bottom warming in the Northwest Atlantic, making it one of the fastest-warming ocean regions globally (Pershing et al., 2015; Friedland et al., 2020; Kavanaugh et al., 2017; Du Pontavice et al., 2023). Real-time Fishing Industry SHared Bottom Oceanographic Timeseries (Fishbot) data and independent observation-based datasets (Figs. 2a, S1 and S2; see also Record et al., 2024) suggest that the rapid warming has paused since the early 2020s, as predicted by earlier simulations with the downscaled regional model (Koul et al., 2024), and markedly cooler water has returned to the region in 2024.

Simulated bottom salinity closely tracks bottom temperature variability (Figs. 2b and S3; $r = 0.84$); noticeable differences are only seen at the shorter interannual time scales. The correlation between temperature and sparse bottom salinity observations (NOAA_SOE) is weaker but also positive ($r = 0.58$). However, confounding from seasonal and spatial variability limits the inference of annual average salinity anomalies from the sparsely observed salinity field alone. The independent GLORYS12 bottom temperature and salinity estimates, which use a high-resolution model with data assimilation to impute estimates within the spatial and temporal gaps between observations, are highly correlated with our simulation results over the shorter GLORYS record ($r = 0.86$ and $r = 0.70$, respectively) – and have a similarly strong bottom temperature and bottom salinity covariance (Figs. 1b and S3, $r = 0.84$ versus $r = 0.90$).

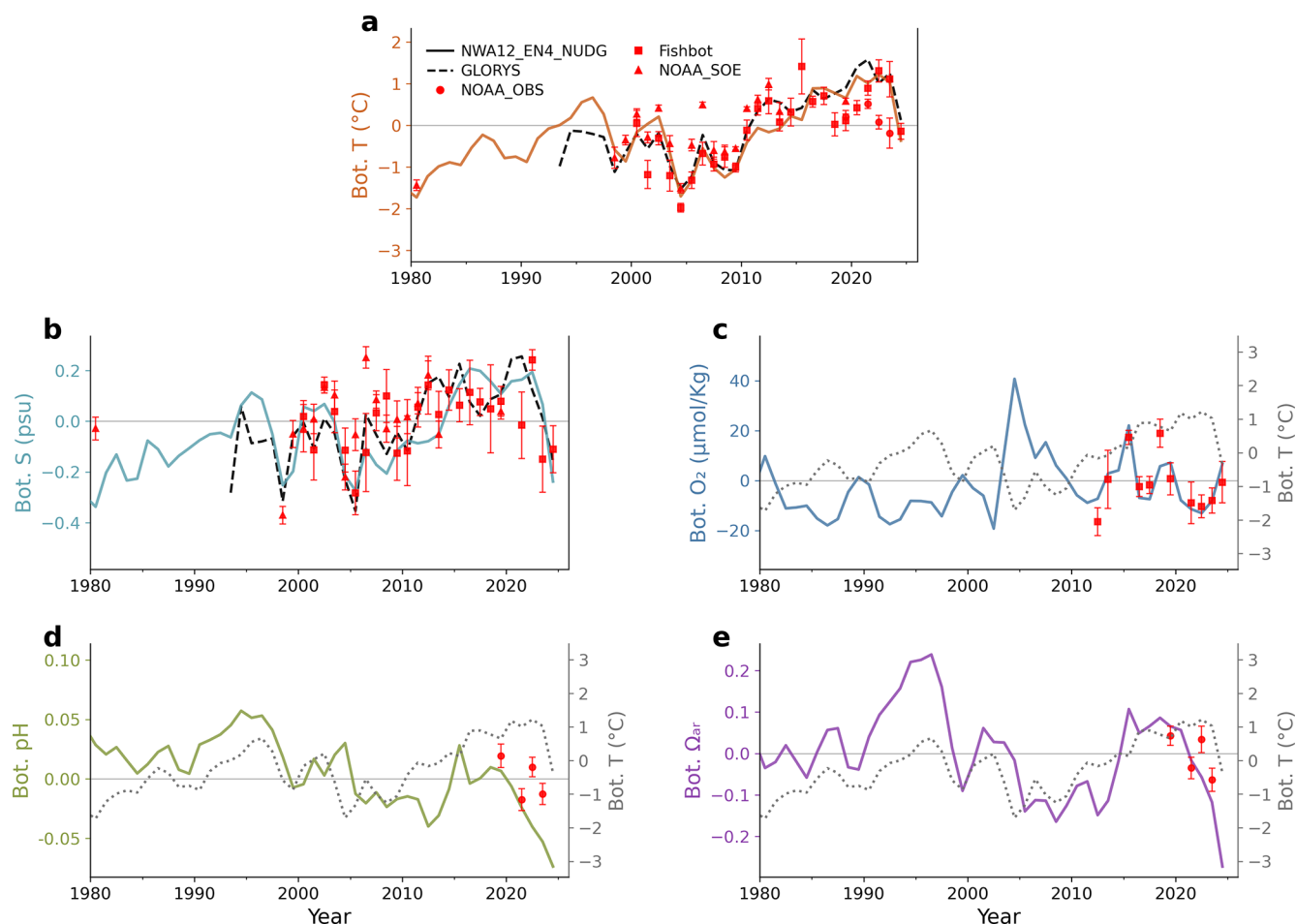


Figure 2. Physical and biogeochemical variability in the GOM. Annual mean anomalies of (a) bottom temperature ($^{\circ}\text{C}$), (b) bottom salinity (psu), (c) bottom oxygen ($\mu\text{mol kg}^{-1}$), (d) bottom pH and (e) bottom aragonite saturation state in the GOM from the ensemble mean of the regional downscaled simulations (NWA12_EN4_NUDG, colored solid lines) and observations (markers with error bars). Observations from the NOAA SOE reports (NOAA_SOE, triangles), the Fishbot bottom data (squares) and the aggregated point observations from vessels and gliders (NOAA_OBS, circles) are also shown for available years. The error bars depict the uncertainty of the annual mean (see Sect. 2). The dashed black lines in (a) and (b) are from the GLORYS12 dataset. All anomalies are calculated with respect to the 1994–2023 mean. To highlight physical and biogeochemical co-variability, bottom temperature lines are shown as dotted grey lines in (c)–(e).

3.2 Bottom oxygen, pH and aragonite saturation ratio

The water mass variations described above are also accompanied by biogeochemical shifts in both the model and observations (Fig. 2c–e). The annual mean bottom oxygen (O_2) anomalies in the GOM show an inverse relationship with bottom temperature and salinity anomalies. The warming and water mass transitions described above would be expected to influence dissolved oxygen concentrations through differences in O_2 concentration in water masses (Claret et al., 2018), solubility effects, and potential changes in biological oxygen demand. The rise in deep oxygen levels from the mid-1990s, peaking in the latter half of 2004 to 2005, corresponds with the cool, fresh, high LSW conditions well documented in the observational record (Mountain, 2012; Feng et al., 2016). LSW is high in oxygen due to high solubility and

rapid oxygen uptake during its formation in the Labrador Sea (Koelling et al., 2023). Since 2005 the bottom oxygen levels have largely dropped, aside from episodic increases in 2015, 2019, and 2024. The relationship between internal (natural) variations in bottom temperature and bottom O_2 confirms that the proportion of WSW/LSW in the Northeast Channel acts as an important driver of this co-variability, though meteorologically linked variations in the winter mixed layer also dictate large episodic variations in bottom O_2 (Figs. 5 and S5). The close correspondence between the observed Fishbot data and the modeled bottom O_2 (Fig. 3c) suggests that our regional downscaled simulation captures the dominant mechanisms controlling bottom oxygen variability.

The bottom pH and aragonite saturation state (Ω_{ar}) anomalies in the GOM show an overall declining trend from 1980–2024 consistent with the uptake of steadily increasing at-

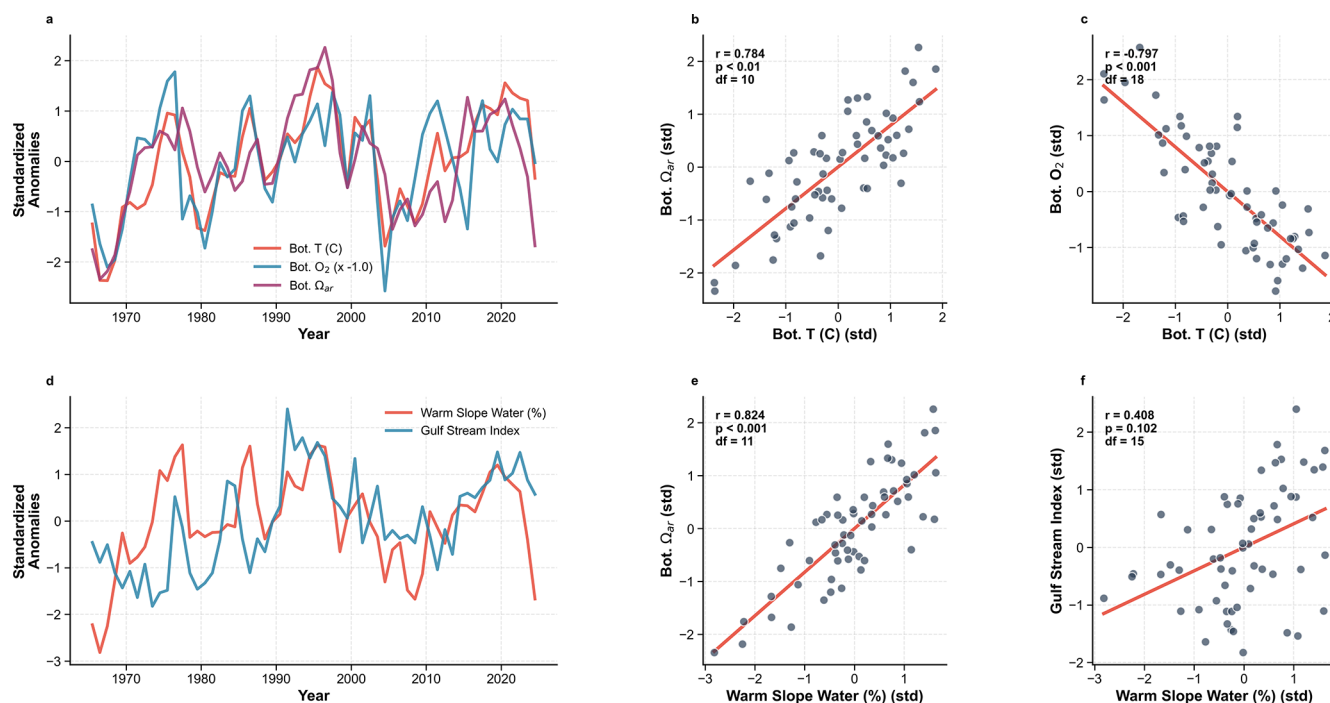


Figure 3. Drivers of internal co-variability in the GOM. Standardized anomalies (with respect to 1994–2023 period) of the internal (natural) component of the annual mean (a) bottom temperature, bottom O_2 and bottom Ω_{ar} in the GOM, (b) percentage of warm slope water (WSW) in the Northeast Channel and the Gulf Stream Index from the regional downscaled simulations. Also shown are scatterplots of internal components of (c) bottom temperature and bottom Ω_{ar} , (d) bottom temperature and bottom O_2 , (e) WSW and bottom Ω_{ar} and (f) WSW and the Gulf Stream Index, along with the linear regression fit (red line). The Pearson correlation coefficient (r), the associated p value and the degrees of freedom (df) are also provided for each pair of variables. The internal component is the difference between the ensemble mean of the downscaled simulation (NWA12_EN4_NUDG) and the ensemble mean of the uninitialized downscaled simulations (NWA12_HIST). The bottom oxygen in (a) is multiplied by -1.0 to show its covariability with bottom temperature.

mospheric CO_2 (i.e., predictable large-scale ocean acidification). The simulated Ω_{ar} time-series, however, also exhibits more significant inter-annual to decadal variability than the bottom pH (Fig. 2d and e), which is consistent with Salisbury and Jönsson (2018). After removing the externally forced ocean acidification signal, Ω_{ar} variations are positively correlated with temperature (Fig. 3, $r = 0.75$). This is consistent with the carbonate chemistry response to warming and increased alkalinity in warm and saline WSW waters (Salisbury and Jönsson, 2018). As in Salisbury and Jönsson's surface analysis, the pronounced warming and shift to WSW from 2010–2020 was associated with increased Ω_{ar} that counteracted the ocean acidification trend seen in pH. The more recent shift toward colder and fresher conditions, however, amplifies the acidification trend and causes a sharp decrease in Ω_{ar} .

3.3 Drivers of bottom biogeochemical variability

To isolate the mechanisms underlying this bottom biogeochemical variability, we decomposed each quantity into its controlling factors (Fig. 4). Bottom O_2 was partitioned into a solubility component, the saturation concentration $[O_2]$

sat (T , S), and a non-solubility residual, $[O_2] - [O_2]_{sat}$, which equals the negative of the apparent oxygen utilisation ($-AOU$) and reflects ventilation, respiration and advection (Fig. 4a). The non-solubility residual accounts for approximately 70 % of the interannual variance of the bottom- O_2 anomaly ($r = 0.89$ with total O_2), compared with about 21 % for the solubility term. The dominant O_2 variations – including the episodic maxima of 2004, 2015 and 2019 – are therefore driven by ventilation of the deep basins through deep winter mixing (Fig. S5) and by the changing oxygen content of the inflowing slope waters, rather than by temperature and salinity-dependent solubility alone. Solubility instead imparts a smaller, smoother decline as the bottom water warms.

For the carbonate system, we examined the bottom dissolved inorganic carbon (DIC) and total alkalinity, which mix conservatively and respond in straightforward ways to gas exchange, respiration and $CaCO_3$ cycling. We also examine alkalinity minus DIC ($Alk - DIC$), a proxy that helps separate the mechanisms driving pH and Ω_{ar} (Xue and Cai, 2020; Fig. 4b). In the model the $Alk - DIC$ anomaly is nearly identical to the bottom Ω_{ar} anomaly ($r = 0.996$; Fig. 4b and d) but tracks pH relatively loosely ($r = 0.82$, because pH also

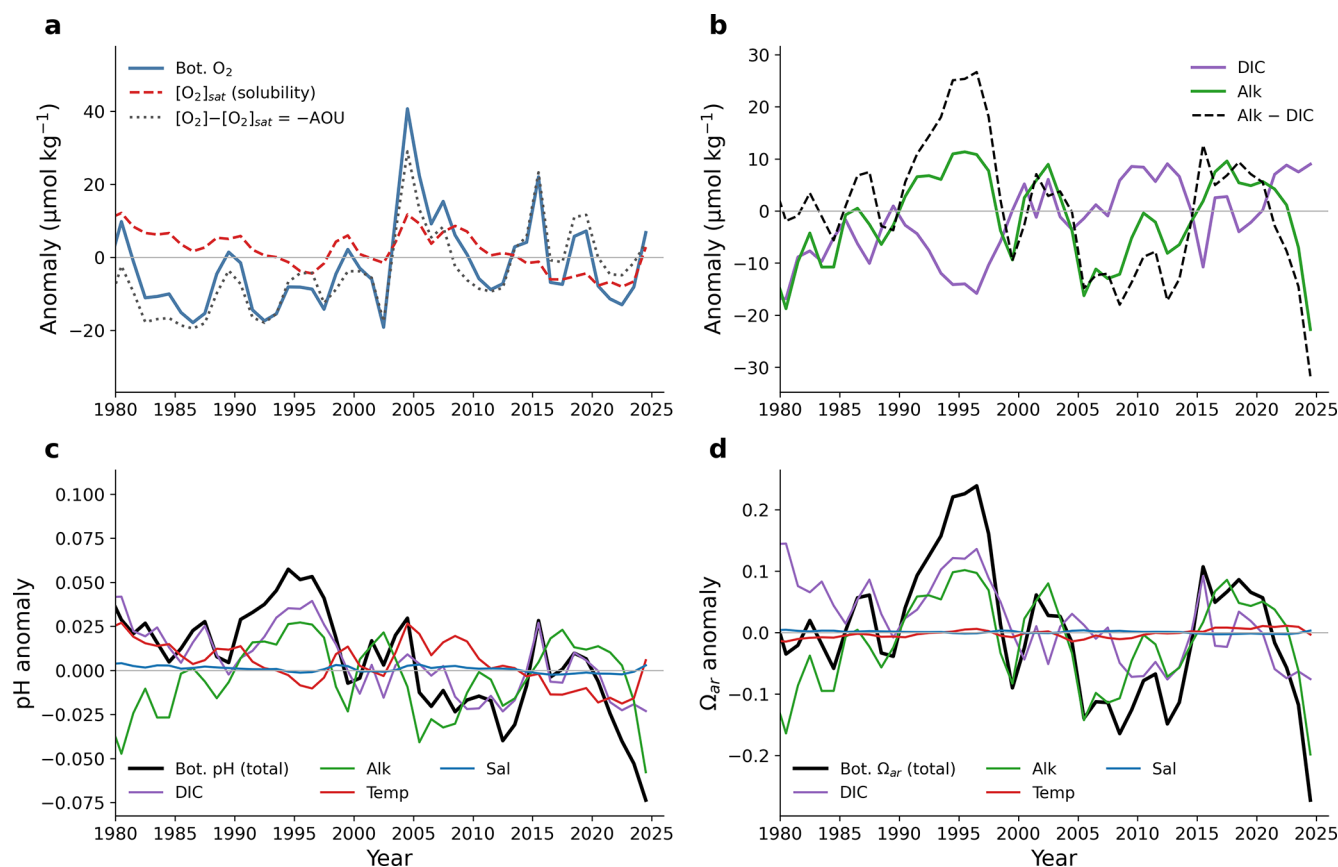


Figure 4. Decomposition of bottom-water biogeochemical variability in the Gulf of Maine. **(a)** Bottom-O₂ anomaly partitioned into the solubility term [O₂]_{sat} and the non-solubility residual [O₂] - [O₂]_{sat} ($\approx$ -AOU) in the NWA12_EN4_NUDG retrospective simulation (ensemble mean anomalies relative to 1994–2023). **(b)** Bottom dissolved inorganic carbon (DIC), total alkalinity (Alk) and Alk - DIC. **(c, d)** Attribution of the bottom **(c)** pH and **(d)** Ω_{ar} anomalies to DIC, Alk, temperature and salinity, obtained by perturbing each driver about the 1994–2023 mean state; the black line is the total anomaly (see Sect. 2).

responds directly to warming). On interannual timescales Alk - DIC (and hence Ω_{ar}) is set by water-mass alkalinity changes, while its long-term decline ($-2.2 \mu\text{mol kg}^{-1}$ per decade), rising DIC outpacing a smaller alkalinity increase, is the anthropogenic acidification trend.

We attributed the bottom pH and Ω_{ar} anomalies to DIC, alkalinity, temperature and salinity (Fig. 4b–d; see Sect. 2). Bottom DIC rises steadily at about $4 \mu\text{mol kg}^{-1}$ per decade, which is the likely fingerprint of anthropogenic carbon uptake. Alkalinity varies primarily with water mass, being higher during the high-alkalinity WSW phases and lower during Labrador Slope Water phases (Fig. 4b). The long-term pH decline (-0.013 per decade) is led by the rising DIC (-0.010 per decade), with warming contributing a further -0.007 per decade, increasing alkalinity partially buffering the trend ($+0.005$ per decade), and salinity making up a small remainder (-0.001 per decade) (Fig. 4c). The decline is thus quantitatively consistent with anthropogenic DIC uptake, modulated by warming and water-mass alkalinity.

The long-term Ω_{ar} decline (-0.015 per decade) is likewise led by rising DIC, whose direct effect (-0.035 per decade) is more than halved by two opposing influences: increasing alkalinity ($+0.016$ per decade) and warming ($+0.004$ per decade) (Fig. 4d). Warming thus raises Ω_{ar} even as it lowers pH, which, together with the larger water-mass alkalinity signal, is why the Ω_{ar} decline is weaker and less monotonic than that of pH. On interannual timescales, Ω_{ar} variability is instead governed primarily by alkalinity (about 57 % of the detrended variance; $r = 0.91$ with the alkalinity contribution), with only a minor direct thermodynamic effect of temperature. The interannual DIC anomaly is small and, because the inflowing WSW in the model does not carry elevated DIC (detrended DIC–alkalinity correlation $r = -0.28$), the DIC signal slightly reinforces rather than opposes the alkalinity-driven variability (Fig. 4d). The correlation between Ω_{ar} and temperature (Fig. 2) therefore arises largely because temperature and alkalinity co-vary through the WSW/LSW water-mass variability, rather than from a direct effect of warming on the carbonate equilibria. Bottom salinity, which co-varies

strongly with bottom temperature through the WSW/LSW water-mass variations (Figs. 2a, b, and S3), is retained as an explicit driver of the pH and Ω_{ar} anomalies in this decomposition (Fig. 4c and d), where its direct contribution is small.

3.4 Prediction skill

Decadal prediction skill is derived from both predictable internal variability of the climate system and from predictable natural and anthropogenic external forcing (Meehl et al., 2009). Capturing the latter requires skillful representation of external drivers (e.g., accumulating CO_2) and its impact on the climate, while capturing the former requires skillful initialization of the climate system and the ability to simulate its natural evolution from that state. The analysis in Figs. 2 and 3 suggested that our retrospective ensemble hindcasts (NWA12_HIND) captured observed variations and covariations across bottom habitat properties, and that the prominence of externally forced and internal variability varies between critical bottom habitat quantities. Comparison of extensive retrospective forecasts against these patterns show that some aspects of both signals were successfully predicted, but others were not (Fig. 5).

Bottom temperature predictions from NWA12_HIND exhibit significant positive anomaly correlations in most of the 10 lead years (Fig. 5a); and in the first lead year, the prediction skill exceeds the skill from the climate change signal (indicated by the diamonds and squares). Both results are consistent with a predictable bottom-temperature signal from climate change and contributions from predictable bottom-temperature (i.e., water mass) variability in the first few lead years. Closer inspection of the prediction skill after linear de-trending further supports the system's capacity to predict aspects of variability around the climate change trend over all the lead years, except at lead years 4 and 5 when the skill drops (Fig. S9). The prediction skill is robust when verified against GLORYS as the observational reference for the period 1993–2024 (Fig. S4).

Initializations at states far from the uninitialized state, the 1980s warming, the rapid 2005–2015 warming and the recent shift to cooler/fresher conditions from 2020 onward, are all correctly predicted by the retrospective predictions (Fig. 6). The 2025 forecast shows a slow reversal of the recent cooling/freshening during the next decade. These results could be interpreted as the model simply relaxing toward the uninitialized state. However, the overall good correlation skill indicates that the model accurately captures the timing and magnitude of such transitions.

The prediction skill for bottom O_2 drops to 0 beyond 2 years and much of the signal for this important habitat factor is linked to predictable internal (natural) variability (Fig. 5b). This is also apparent from the time series (Fig. 2c) which shows that bottom O_2 does not have a uniform long-term trend as was the case for bottom temperature. The rapid drop in prediction skill may also be linked to the pres-

ence of sharp and likely less predictable meteorologically driven variations in the depth of the winter mixed layer. The winter 2003/2004 is a prime example of this phenomenon when large positive winter MLD anomalies in our regional downscaled model drove ventilation of the bottom water and strong positive bottom oxygen anomalies (Figs. 6, and S5). Similar peaks in bottom O_2 coinciding with deep winter mixing and positive MLD anomalies are seen in 2015 and 2019.

The decadal predictions initialized before and after the 2004 peak (in 2002 and 2006 respectively) show that the ensemble mean prediction has difficulties predicting the extent of the rise in bottom O_2 , but the subsequent decay of these elevated bottom O_2 levels is well predicted by the model (Fig. 6b). These characteristics of the past variability in bottom O_2 are consistent with the high skill of the initialized predictions, which outperform the persistence-based skill and the skill of uninitialized simulations in the first two lead years (Fig. 5b), and the significant drop thereafter.

Both bottom pH and Ω_{ar} exhibit a declining trend (Fig. 2e and f), yet they differ in how much the climate change signal influences their prediction skill. Ocean pH is primarily controlled by increasing atmospheric CO_2 , while Ω_{ar} is more strongly influenced by ocean water masses with varying alkalinity (LSW versus WSW). As a result, internal variability contributes more to the prediction skill of Ω_{ar} , whereas the climate change signal has a greater impact on the prediction skill of bottom pH (Fig. 5c and d). Beyond lead year 2, the climate change signal predominantly determines the prediction skill of bottom pH, allowing it to remain high and significantly better than persistence for all lead years (except lead year 3) which is not the case for Ω_{ar} where the skill drops rapidly beyond lead year 2. When the linear trend is removed, the skill of internal variability significantly surpasses the persistence and climate change-driven skill in the first two lead years. The notable decrease in the prediction skill of detrended bottom pH, compared to Ω_{ar} , further demonstrates the additional contribution of internal variability to the prediction skill of Ω_{ar} . Individual 10-year forecasts provide further evidence of the capability of the model to track past deviation of bottom pH and Ω_{ar} from the long-term trend (Fig. 6c and d).

3.5 Probabilistic predictions

The preceding analyses leveraged over 6500 years of retrospective forecasts to assess the capacity of the ensemble mean prediction to anticipate physical and biogeochemical changes on a decadal time horizon. While results suggested significant capability arising from predictable climate variability and change signals, decision-making benefits from actionable information on uncertainties (e.g., Ramos et al., 2013; Roulston et al., 2006). A coastal resource manager or fishing industry representative might ask, “How reliable are forecasts for anomalously warm or cool conditions?”. To quantify this in more accessible terms, we calculated the ac-

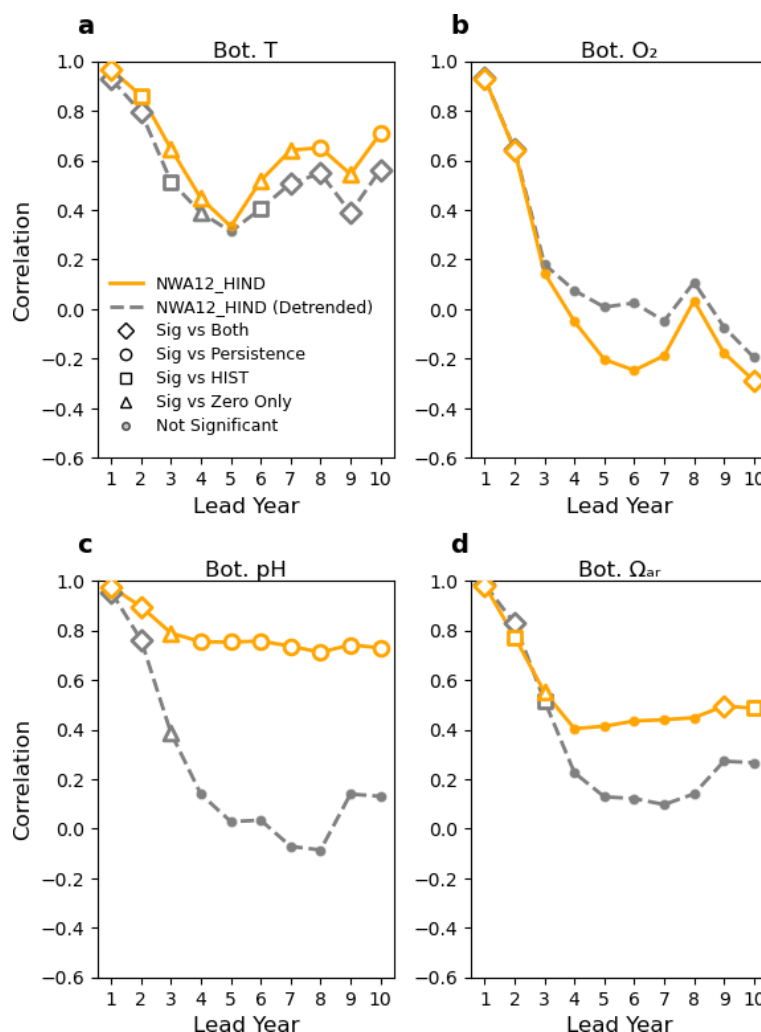


Figure 5. Decadal prediction skill. Anomaly correlation coefficient of the annual mean anomalies from the initialized regional downscaled model (NWA12_HIND) predictions. The symbols indicate statistically significant differences relative to alternative models of no correlation, persistence, and the uninitialized (historical) simulation (see Sect. 2). The correlation is calculated against the regional downscaled simulation (NWA12_EN4_NUDG)-derived annual mean anomalies (ensemble mean) for (a) bottom temperature, (b) bottom O₂, (c) bottom pH and (d) bottom Ω_{ar} over the GOM_EPU. The prediction and persistence skills are calculated for the common 1965–2024 period. The grey dashed lines show skill for the linearly detrended annual mean anomalies.

curacy of probabilistic forecasts derived from the model simulations (Tommasi et al., 2017a, see Sect. 2). We used terciles to define above, normal, and below-normal conditions and focused on predicted conditions averaged over the next 3 years, which corresponds to typical stock assessment time-lines in actively managed fisheries (Tommasi et al., 2017b). Extended logistic regression (Wilks, 2009) was used to calibrate forecast uncertainty based on the forecast error (see Sect. 2).

Our regional downscaled predictions skilfully distinguish between tercile classes. All four variables have positive ranked probability skill scores, and the forecast accuracy for a given tercile exceeds the accuracy of a random forecast (0.56) and the accuracy of consistently negative predic-

tion (0.67) in all but two cases (Fig. 7). The ensemble confidently predicted the observed cooling in the mid-2000s and the subsequent warming phase that extended from 2005 to approximately 2020 (Fig. 7).

The RPSS for all quantities is > 0 , indicating the model has some capacity to anticipate changes (i.e., probabilistic skill beyond the persistence forecast, see Sect. 2). Oxygen has slightly lower skill (RPSS 0.46, accuracy 0.67–0.81) compared to bottom temperature (RPSS 0.47, accuracy 0.72–0.90), reflecting both the weak predictable climate change signals and the influence of less predictable atmospheric variability. Similarly, predictions for bottom pH, which has a very strong predictable climate change signal, have higher

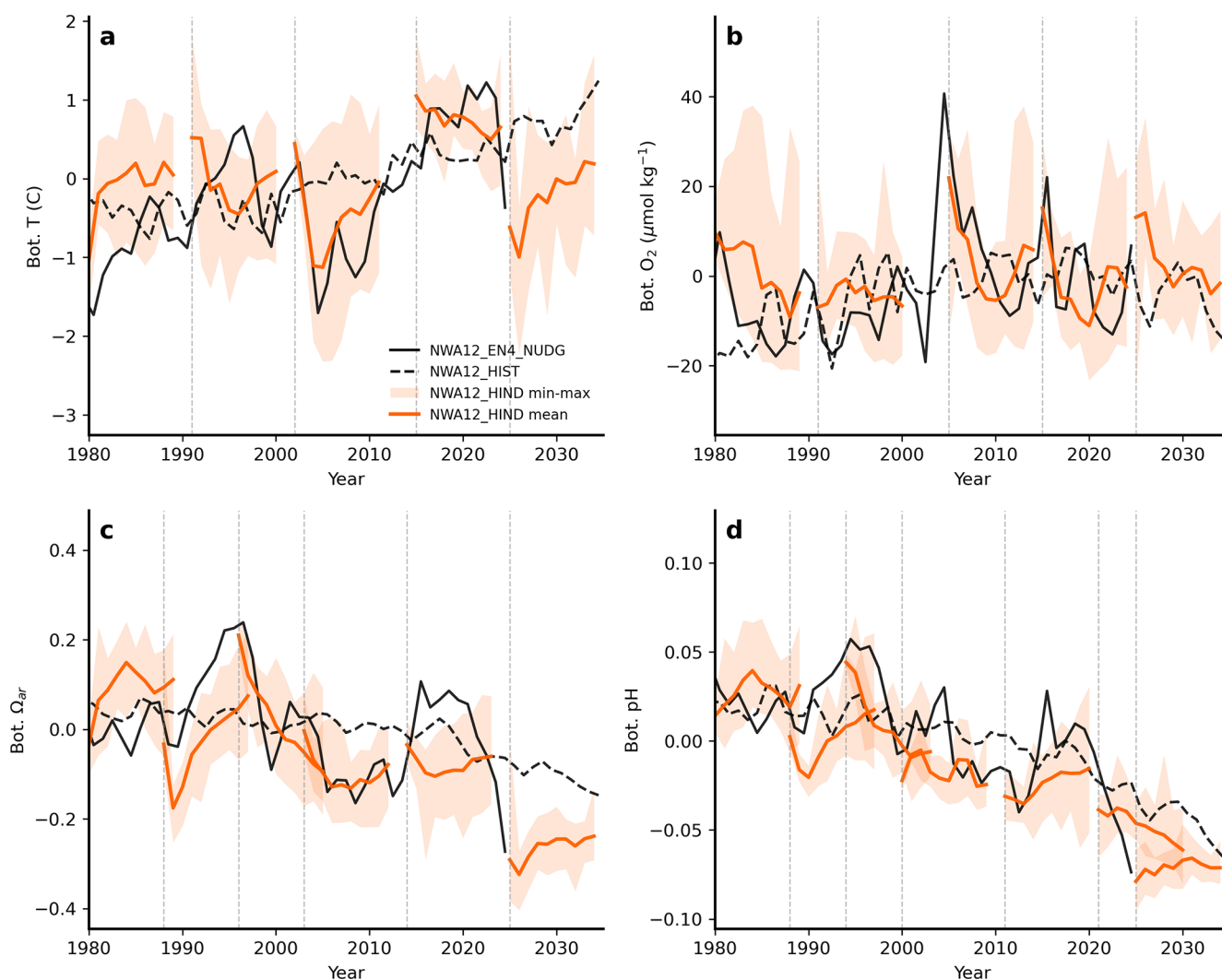


Figure 6. Physical and biogeochemical predictions in the GOM. Annual mean anomalies of (a) bottom temperature ($^{\circ}\text{C}$), (b) bottom oxygen ($\mu\text{mol kg}^{-1}$), (c) bottom aragonite saturation state and (d) bottom pH in the GOM_EPU. The solid black lines show the anomalies from the ensemble mean of regional downscaled simulations (NWA12_EN4_NUDG, as in Fig. 2), the dashed black lines show the ensemble mean of uninitialized downscaled model simulations (NWA12_HIST) carried under historical anthropogenic forcing and SSP585 emissions scenario. The bold orange lines show the 10-yearlong initialized prediction (NWA12_HIND) and the vertical grey lines depict their year of initialization. The bold orange lines are the ensemble means of the respective 10-member ensemble, and the light shading is the min-max range of the ensemble members. A lead-time-dependent climatology was removed from each ensemble member to create anomalies. All anomalies are calculated with respect to the 1994–2023 mean.

skill (RPSS 0.6, accuracy 0.76–0.88) than Ω_{ar} (RPSS 0.51, accuracy 0.62–0.79).

The high accuracy of prediction for bottom temperature, O_2 , pH and Ω_{ar} likely reflects both methodological and mechanistic factors. The bottom temperature, O_2 , pH and Ω_{ar} are validated against the same regional downscaled simulation from which initial conditions were derived (see Sect. 2). While this simulation recreates observed O_2 patterns and covariances with well-constrained physical properties to the extent that they have been observed (Figs. 2, 3 and S7), it should be viewed as an upper limit for prediction skill. Additional autonomous biogeochemical observing platforms ca-

pable of observing shelf environments may help address this limitation in the future (Henson, 2014; Chai et al., 2020).

Looking ahead over the next three years (2025–2027), we forecast below-normal bottom temperatures with high confidence (88.3 % probability), supported by 93.3 % retrospective accuracy when issuing similar high confidence forecasts (14/15 hits). We also predict above-normal bottom O_2 levels with high confidence (92.0 % probability), with an 85.7 % retrospective accuracy at comparable confidence levels (6/7 hits). Additionally, we forecast with very high confidence (> 99 % probability) both below-normal bottom pH and below-normal Ω_{ar} conditions, both of which are sup-

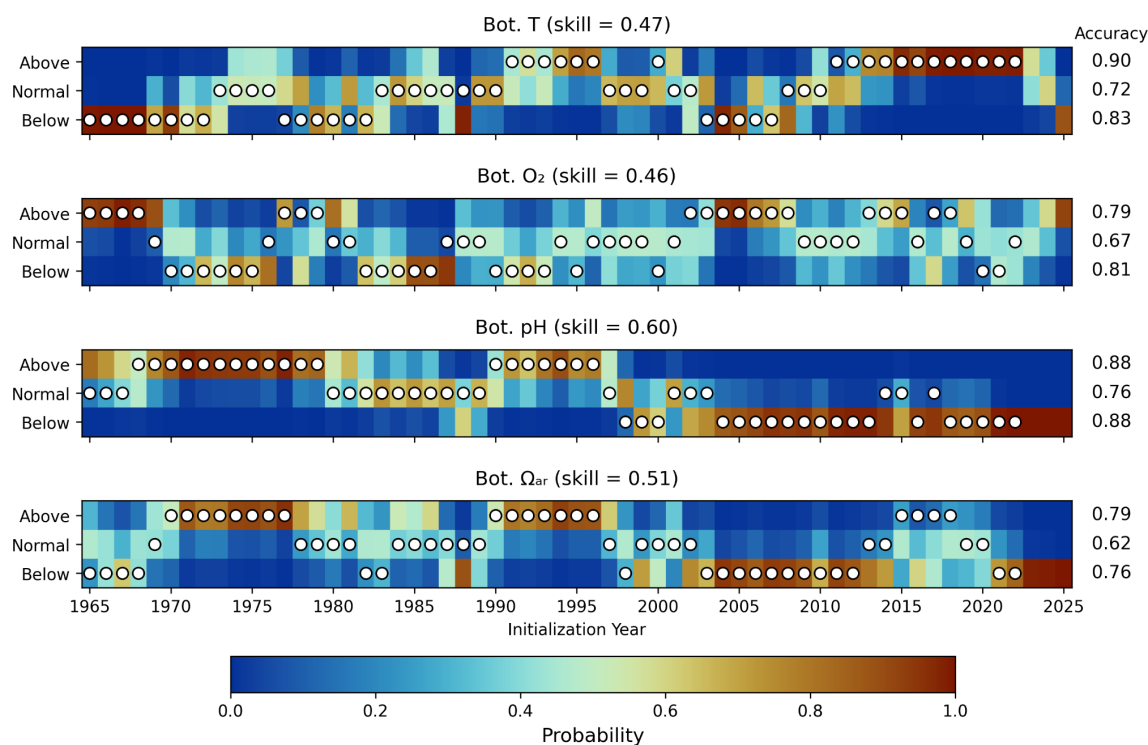


Figure 7. Probabilistic predictions for the GOM. Probability of the predicted lead 1–3 year mean anomalies (with respect to the 1994–2023 mean) of (a) bottom temperature, (b) bottom O₂, (c) bottom pH and (d) bottom Ω_{ar} from the regional downscaled decadal prediction model (NWA12_HIND). The corresponding classifications of the reference anomalies are shown as white circles from the regional down-scaled simulation (NWA12_EN4_NUDG). The tercile boundaries are calculated separately for the predictions and reference/observations. The *x* axis shows the initialization year and the 2025 forecast. The forecast accuracy, (True Positives + True Negatives)/Total forecasts, for each tercile is shown on the right. A random forecast that randomly picks one tercile as most likely for each forecast has an accuracy of 0.56 and the accuracy of a consistently negative (counter) prediction is 0.67. The forecast skill compared to persistence is given in the subtitle of each panel.

ported by perfect retrospective accuracy when issuing very high confidence forecasts, albeit with limited sample size (9/9 hits and 1/1 hit respectively). While these forecasts represent our best estimate, both the confidence levels and the underlying sample sizes should be considered when making decisions based on these predictions, particularly for the forecasts where limited similarly confident historical forecasts constrain our ability to fully validate forecast skill. Nevertheless, the dynamically consistent physical and biogeochemical forecasts provide compelling evidence that the ongoing watermass reorganization will continue to influence bottom water conditions through the 2025–2027 period.

4 Discussion: towards the future of fisheries management

Future projections indicate the entire GOM will experience suboptimal aragonite saturation for most of the year by 2050 under high emissions scenarios (Siedlecki et al., 2021). The region's demonstrated sensitivity to water mass-driven alkalinity variations makes it particularly vulnerable to rapid

acidification once rising CO₂ exceeds the buffering ability of the projected increased Gulf Stream water, potentially triggering swift transitions to persistently corrosive conditions that could impact economically vital shellfish industries (Ekstrom et al., 2015). Although the recent transition from a decade of warm water influence to emerging cold water conditions across portions of the Northeast US Continental Shelf, including the deep GOM (Record et al., 2024), may buy time for fisheries to adapt to the projected long-term warming trend, it also may accelerate or alter the projected biogeochemical shifts.

We assessed the decadal prediction skill for the Gulf of Maine, a semi-enclosed body of water on the Northeast US Continental Shelf. While the patterns of prediction skill that emerged from the interplay of climate change signals and variability were unique to each quantity and the Gulf of Maine, other coastal regions are subject to similar oceanic and atmospheric forcing (Shearman and Lentz, 2010) and may present opportunities for bottom-habitat prediction. The ubiquity of sea surface temperature prediction skill in coastal systems at seasonal and decadal scales provides grounds

for optimism (Stock et al., 2015; Tommasi et al., 2017b). The methodologies employed herein, while computationally intensive, provide a robust means to diagnose the origins of predictive skill and attach probabilities critical to making forecast information actionable for management applications.

More frequent and extensive observations of the ocean, including deep ocean biogeochemistry, are essential for increasing confidence in the accuracy of the forecasts and providing confirmation of predicted changes (Capotondi et al., 2019). Observations are especially important for verifying interannual to decadal scale ocean predictions, where long time series are necessary to verify that the predictions have more skill than simple forecasts of the forced signal or auto-correlated noise. Here, given the scarcity of past deep along-shelf biogeochemical observations, we utilized a numerical model as the reference for forecast verification. This down-scaled simulation was validated using the best available observations (Figs. 2, S1–S3, S6 and S7). Consequently, the reported skill represents potential skill and may be an overestimation. However, the model successfully simulated observed variability in bottom temperature and salinity and captured the expected covariance between these two variables and the bottom biogeochemistry. Koul et al. (2024) also demonstrated forecast skill for surface temperature when compared against actual observations. Furthermore, the model is likely accurate in predicting much of the forced component of skill, which is driven by steadily increasing greenhouse gas concentrations and resulting impacts on ocean warming and chemistry.

Our forecasts of continued cooling alongside below-normal bottom pH, Ω_{ar} and above-normal bottom O_2 levels through 2027 reveal the substantial biogeochemical fingerprint of this water mass transition and potentially offer stakeholders lead time to anticipate ecosystem shifts. Skillful forecasts on these multi-year timescales can meet critical needs of fishers, managers, conservationists, and other stakeholders (Link et al., 2023; Saba et al., 2023). The framework developed here, consisting of a computationally efficient regional model that downscales regularly updated global predictions, is well positioned to meet these needs and is intended to be transitioned towards operational use through NOAA's Changing Ecosystems and Fisheries Initiative. Recognizing that marine resource management is a wicked problem (Jentoft and Chuenpagdee, 2009; Rittel and Webber, 1973), however, careful application-specific analyses of the potential costs, benefits, and tradeoffs of using these forecasts for decision-making will be necessary.

Code availability. The regional MOM6 model's source code can be accessed at <https://github.com/NOAA-GFDL/CEFI-regional-MOM6> (NOAA-GFDL, 2026a). MOM6 operates on an open development framework, with Git repositories at <https://github.com/mom-ocean/MOM6> (mom-ocean, 2026) and

<https://github.com/NOAA-GFDL/MOM6> (NOAA-GFDL 2026b). Source code for other model components is also available at <https://github.com/NOAA-GFDL> (NOAA-GFDL 2026c).

Data availability. The processed model output and the observations used in this research are publicly available via Zenodo (<https://doi.org/10.5281/zenodo.16985350>, Koul, 2025).

Supplement. The supplement related to this article is available online at <https://doi.org/10.5194/os-22-3017-2026-supplement>.

Author contributions. VK, ACR, and CS conceptualized the research and drafted the initial manuscript. VK was responsible for conducting model simulations, performing decadal predictions, post-processing model outputs and observations, and generating the corresponding figures. ACR and CS contributed in designing and discussing relevant prediction skill metrics. TD, AW, and LZ developed the SPEAR global prediction model. All authors contributed to the analysis and interpretation of results, as well as the editing of the paper.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Ocean Science*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Acknowledgements. Computational resources for this work were provided by allocations on NOAA's Research and Development High-Performance Computing System from the Geophysical Fluid Dynamics Laboratory and NOAA's Changing Ecosystems and Fisheries Initiative.

Disclaimer. The statements, findings, conclusions, and recommendations presented in this report are those of the author(s) and do not necessarily reflect the official views of the National Oceanic and Atmospheric Administration or the US Department of Commerce.

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Financial support. Vimal Koul is funded by the NOAA Climate Program Office. Thomas Delworth, Andrew C. Ross, Charles Stock, and Andrew T. Wittenberg are supported by NOAA's Geophysical Fluid Dynamics Laboratory as part of its base activities. Liping Zhang receives support through UCAR under block funding from NOAA/GFDL. This report was prepared by Vimal Koul under

awards NA18OAR4320123 and NA23OAR4320198 from the National Oceanic and Atmospheric Administration, US Department of Commerce.

Review statement. This paper was edited by Elizabeth H. Shadwick and reviewed by two anonymous referees.

References

- Abram, N. J., McGregor, H. V., Tierney, J. E., Evans, M. N., McKay, N. P., and Kaufman, D. S.: Early onset of industrial-era warming across the oceans and continents, *Nature*, 536, 411–418, <https://doi.org/10.1038/nature19082>, 2016.
- Alexander, M. A., Scott, J. D., Friedland, K. D., Mills, K. E., Nye, J. A., Pershing, A. J., and Thomas, A. C.: Projected sea surface temperatures over the 21st century: Changes in the mean, variability and extremes for large marine ecosystem regions of Northern Oceans, *Elem. Sci. Anth.*, 6, 9, <https://doi.org/10.1525/elementa.191>, 2018.
- Breitbart, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V., Gilbert, D., Gutiérrez, D., Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S. W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose, K. A., Seibel, B. A., Telszewski, M., Yasuhara, M., and Zhang, J.: Declining oxygen in the global ocean and coastal waters, *Science*, 359, eaam7240, <https://doi.org/10.1126/science.aam7240>, 2018.
- Brickman, D., Hebert, D., and Wang, Z.: Mechanism for the recent ocean warming events on the Scotian Shelf of eastern Canada, *Cont. Shelf Res.*, 156, 11–22, <https://doi.org/10.1016/j.csr.2018.01.001>, 2018.
- Caldeira, K. and Wickett, M. E.: Anthropogenic carbon and ocean pH, *Nature*, 425, 365–365, <https://doi.org/10.1038/425365a>, 2003.
- Capotondi, A., Jacox, M., Bowler, C., Kavanaugh, M., Lehoudey, P., Barrie, D., Brodie, S., Chaffron, S., Cheng, W., Dias, D. F., Eveillard, D., Guidi, L., Iudicone, D., Lovenduski, N. S., Nye, J. A., Ortiz, I., Pirhalla, D., Pozo Buil, M., Saba, V., Sheridan, S., Siedlecki, S., Subramanian, A., de Vargas, C., Di Lorenzo, E., Doney, S. C., Hermann, A. J., Joyce, T., Merrifield, M., Miller, A. J., Not, F., and Pesant, S.: Observational Needs Supporting Marine Ecosystems Modeling and Forecasting: From the Global Ocean to Regional and Coastal Systems, *Front. Mar. Sci.*, 6, 623, <https://doi.org/10.3389/fmars.2019.00623>, 2019.
- Carter, B. R., Bittig, H. C., Fassbender, A. J., Sharp, J. D., Takeshita, Y., Xu, Y.-Y., Álvarez, M., Wanninkhof, R., Feely, R. A., and Barbero, L.: New and updated global empirical seawater property estimation routines, *Limnol. Oceanogr.-Meth.*, 19, 785–809, <https://doi.org/10.1002/lom3.10461>, 2021.
- Chai, F., Johnson, K. S., Claustre, H., Xing, X., Wang, Y., Boss, E., Riser, S., Fennel, K., Schofield, O., and Sutton, A.: Monitoring ocean biogeochemistry with autonomous platforms, *Nat. Rev. Earth Environ.*, 1, 315–326, <https://doi.org/10.1038/s43017-020-0053-y>, 2020.
- Chen, Z., Kwon, Y.-O., Chen, K., Fratantoni, P., Gawarkiewicz, G., and Joyce, T. M.: Long-Term SST Variability on the Northwest Atlantic Continental Shelf and Slope, *Geophys. Res. Lett.*, 47, e2019GL085455, <https://doi.org/10.1029/2019GL085455>, 2020.
- Chen, Z., Kwon, Y.-O., Chen, K., Fratantoni, P., Gawarkiewicz, G., Joyce, T. M., Miller, T. J., Nye, J. A., Saba, V. S., and Stock, B. C.: Seasonal Prediction of Bottom Temperature on the Northeast U.S. Continental Shelf, *J. Geophys. Res.-Oceans*, 126, e2021JC017187, <https://doi.org/10.1029/2021JC017187>, 2021.
- Chen, Z., Siedlecki, S., Long, M., Petrik, C. M., Stock, C. A., and Deutsch, C. A.: Skillful multiyear prediction of marine habitat shifts jointly constrained by ocean temperature and dissolved oxygen, *Nat. Commun.*, 15, 900, <https://doi.org/10.1038/s41467-024-45016-5>, 2024.
- Claret, M., Galbraith, E. D., Palter, J. B., Bianchi, D., Fennel, K., Gilbert, D., and Dunne, J. P.: Rapid coastal deoxygenation due to ocean circulation shift in the northwest Atlantic, *Nat. Clim. Change*, 8, 868–872, <https://doi.org/10.1038/s41558-018-0263-1>, 2018.
- Clark, J. S., Carpenter, S. R., Barber, M., Collins, S., Dobson, A., Foley, J. A., Lodge, D. M., Pascual, M., Pielke Jr., R., Pizer, W., Pringle, C., Reid, W. V., Rose, K. A., Sala, O., Schlesinger, W. H., Wall, D. H., and Wear, D.: Ecological forecasts: an emerging imperative, *Science*, 293, 657–660, <https://doi.org/10.1126/science.293.5530.657>, 2001.
- Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J.-H., Dunne, K. A., Ginoux, P., Gudgel, R., Hallberg, R. W., Harris, L., Harrison, M. J., Johnson, N., Kapnick, S. B., Lin, S.-J., Lu, F., Malyshev, S., Milly, P. C., Murakami, H., Naik, V., Pascale, S., Paynter, D., Rosati, A., Schwarzkopf, M. D., Shevliakova, E., Underwood, S., Wittenberg, A. T., Xiang, B., Yang, X., Zeng, F., Zhang, H., Zhang, L., and Zhao, M.: SPEAR: The next generation GFDL modeling system for seasonal to multi-decadal prediction and projection, *J. Adv. Model. Earth Syst.*, 12, e2019MS001895, <https://doi.org/10.1029/2019MS001895>, 2020.
- Drinkwater, K. F., Mountain, D. B., and Herman, A.: Recent changes in the hydrography of the Scotian Shelf and Gulf of Maine – a return to conditions of the 1960's?, *Sci. Coun. Res. Doc. NAFO 98*, 16 pp., <https://www.nafo.int/Portals/0/PDFs/sc/1998/scr-98-037.pdf> (last access: 25 September 2026), 1998.
- Du Pontavice, H., Chen, Z., and Saba, V. S.: A high-resolution ocean bottom temperature product for the northeast U.S. continental shelf marine ecosystem, *Prog. Oceanogr.*, 210, 102948, <https://doi.org/10.1016/j.pocean.2022.102948>, 2023.
- Ekstrom, J. A., Suatoni, L., Cooley, S. R., Pendleton, L. H., Waldbusser, G. G., Cinner, J. E., Ritter, J., Langdon, C., van Hooidek, R., Gledhill, D., Wellman, K., Beck, M. W., Brander, L. M., Rittschof, D., Doherty, C., Edwards, P. E. T., and Portela, R.: Vulnerability and adaptation of US shellfisheries to ocean acidification, *Nat. Clim. Change*, 5, 207–214, <https://doi.org/10.1038/nclimate2508>, 2015.
- Fay, G., Link, J. S., and Hare, J. A.: Assessing the effects of ocean acidification in the Northeast US using an end-to-end marine ecosystem model, *Ecol. Model.*, 347, 1–10, <https://doi.org/10.1016/j.ecolmodel.2016.12.016>, 2017.
- Feely, R. A., Jiang, L.-Q., Wanninkhof, R., Carter, B., Alin, S., Bednaršek, N., and Cosca, C.: Acidification of the Global Surface Ocean: What We Have Learned from Observations, *Oceanography*, 36, 120–129, <https://doi.org/10.5670/oceanog.2023.222>, 2023.
- Feng, H., Vandemark, D., and Wilkin, J.: Gulf of Maine salinity variation and its correlation with upstream Scotian Shelf currents

- at seasonal and interannual time scales, *J. Geophys. Res.-Oceans*, 121, 8585–8607, <https://doi.org/10.1002/2016JC012337>, 2016.
- Friedland, K. D. and Hare, J. A.: Long-term trends and regime shifts in sea surface temperature on the continental shelf of the northeast United States, *Cont. Shelf Res.*, 27, 2313–2328, <https://doi.org/10.1016/j.csr.2007.06.001>, 2007.
- Friedland, K. D., Morse, R. E., Manning, J. P., Melrose, D. C., Miles, T., Goode, A. G., Brady, D. C., Kohut, J. T., and Powell, E. N.: Trends and change points in surface and bottom thermal environments of the US Northeast Continental Shelf Ecosystem, *Fish. Oceanogr.*, 29, 396–414, <https://doi.org/10.1111/fog.12485>, 2020.
- Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Arneeth, A., Arora, V., Bates, N. R., Becker, M., Bellouin, N., Berghoff, C. F., Bittig, H. C., Bopp, L., Cadule, P., Campbell, K., Chamberlain, M. A., Chandra, N., Chevallier, F., Chini, L. P., Colligan, T., Decayeux, J., Djeutchouang, L. M., Dou, X., Duran Rojas, C., Enyo, K., Evans, W., Fay, A. R., Feely, R. A., Ford, D. J., Foster, A., Gasser, T., Gehlen, M., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A. K., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Kato, E., Keeling, R. F., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Lan, X., Lauvset, S. K., Lefèvre, N., Liu, Z., Liu, J., Ma, L., Maksyutov, S., Marland, G., Mayot, N., McGuire, P. C., Metzl, N., Monacci, N. M., Morgan, E. J., Nakaoka, S.-I., Neill, C., Niwa, Y., Nützel, T., Olivier, L., Ono, T., Palmer, P. I., Pierrot, D., Qin, Z., Resplandy, L., Roobaert, A., Rosan, T. M., Rödenbeck, C., Schwinger, J., Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Steinhoff, T., Sun, Q., Sutton, A. J., Séférian, R., Takao, S., Tatebe, H., Tian, H., Tilbrook, B., Torres, O., Tourigny, E., Tsujino, H., Tubiello, F., van der Werf, G., Wanninkhof, R., Wang, X., Yang, D., Yang, X., Yu, Z., Yuan, W., Yue, X., Zaehle, S., Zeng, N., and Zeng, J.: Global Carbon Budget 2024, *Earth Syst. Sci. Data*, 17, 965–1039, <https://doi.org/10.5194/essd-17-965-2025>, 2025.
- Frölicher, T. L., Ramseyer, L., Raible, C. C., Rodgers, K. B., and Dunne, J.: Potential predictability of marine ecosystem drivers, *Biogeosciences*, 17, 2061–2083, <https://doi.org/10.5194/bg-17-2061-2020>, 2020.
- Gangopadhyay, A., Gawarkiewicz, G., Silva, E. N. S., Monim, M., and Clark, J.: An Observed Regime Shift in the Formation of Warm Core Rings from the Gulf Stream, *Sci. Rep.*, 9, 12319, <https://doi.org/10.1038/s41598-019-48661-9>, 2019.
- Garcia, H. E. and Gordon, L. I.: Oxygen solubility in seawater: Better fitting equations, *Limnol. Oceanogr.*, 37, 1307–1312, <https://doi.org/10.4319/llo.1992.37.6.1307>, 1992.
- Gehlen, M., Barciela, R., Bertino, L., Brasseur, P., Butenschön, M., Chai, F., Crise, A., Drillet, Y., Ford, D., Lavoie, D., Lehodey, P., Perruche, C., Samuelsen, A., and Simon, E.: Building the capacity for forecasting marine biogeochemistry and ecosystems: Recent advances and future developments, *J. Oper. Oceanogr.*, 8, s168–s187, <https://doi.org/10.1080/1755876X.2015.1022350>, 2015.
- Gonçalves Neto, A., Langan, J. A., and Palter, J. B.: Changes in the Gulf Stream preceded rapid warming of the Northwest Atlantic Shelf, *Commun. Earth Environ.*, 2, 74, <https://doi.org/10.1038/s43247-021-00143-5>, 2021.
- Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, *J. Geophys. Res.-Oceans*, 118, 6704–6716, <https://doi.org/10.1002/2013JC009067>, 2013.
- Hare, J. A., Morrison, W. E., Nelson, M. W., Stachura, M. M., Teeters, E. J., Griffis, R. B., Alexander, M. A., Scott, J. D., Alade, L., Bell, R. J., Chute, A. S., Curti, K. L., Curtis, T. H., Kirchis, D., Kocik, J. F., Lucey, S. M., McCandless, C. T., Milke, L. M., Richardson, D. E., Robillard, E., Walsh, H. J., McManus, M. C., Marancik, K. E., and Griswold, C. A.: A Vulnerability Assessment of Fish and Invertebrates to Climate Change on the Northeast U.S. Continental Shelf, *PLoS ONE*, 11, e0146756, <https://doi.org/10.1371/journal.pone.0146756>, 2016.
- Harrington, A. M. and Hamlin, H. J.: Ocean acidification alters thermal cardiac performance, hemocyte abundance, and hemolymph chemistry in subadult American lobsters *Homarus americanus* H. Milne Edwards, 1837 (Decapoda: Malacostraca: Nephropidae), *J. Crustacean Biol.*, 39, 468–476, <https://doi.org/10.1093/jcbiol/ruz015>, 2019.
- Henson, S. A.: Slow science: the value of long ocean biogeochemistry records, *Philos. T. Roy. Soc. A*, 372, 20130334, <https://doi.org/10.1098/rsta.2013.0334>, 2014.
- Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., and Zhang, H.-M.: Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons, *J. Climate*, 30, 8179–8205, <https://doi.org/10.1175/JCLI-D-16-0836.1>, 2017.
- Humphreys, M. P., Lewis, E. R., Sharp, J. D., and Pierrot, D.: PyCO2SYS v1.8: Marine carbonate system calculations in Python, *Geosci. Model Dev.*, 15, 15–43, <https://doi.org/10.5194/gmd-15-15-2022>, 2022.
- Jentoft, S. and Chuenpagdee, R.: Fisheries and coastal governance as a wicked problem, *Mar. Policy*, 33, 553–560, <https://doi.org/10.1016/j.marpol.2008.12.002>, 2009.
- Jiang, L.-Q., Feely, R. A., Wanninkhof, R., Greeley, D., Barbero, L., Alin, S., Carter, B. R., Pierrot, D., Featherstone, C., Hooper, J., Melrose, C., Monacci, N., Sharp, J. D., Shellito, S., Xu, Y.-Y., Kozyr, A., Byrne, R. H., Cai, W.-J., Cross, J., Johnson, G. C., Hales, B., Langdon, C., Mathis, J., Salisbury, J., and Townsend, D. W.: Coastal Ocean Data Analysis Product in North America (CODAP-NA) – an internally consistent data product for discrete inorganic carbon, oxygen, and nutrients on the North American ocean margins, *Earth Syst. Sci. Data*, 13, 2777–2799, <https://doi.org/10.5194/essd-13-2777-2021>, 2021.
- Johnson, G. C. and Lyman, J. M.: Warming trends increasingly dominate global ocean, *Nat. Clim. Change*, 10, 757–761, <https://doi.org/10.1038/s41558-020-0822-0>, 2020.
- Kavanaugh, M. T., Rheuban, J. E., Luis, K. M. A., and Doney, S. C.: Thirty-Three Years of Ocean Benthic Warming Along the U.S. Northeast Continental Shelf and Slope: Patterns, Drivers, and Ecological Consequences, *J. Geophys. Res.-Oceans*, 122, 9399–9414, <https://doi.org/10.1002/2017JC012953>, 2017.
- Keeling, R. F., Körtzinger, A., and Gruber, N.: Ocean Deoxygenation in a Warming World, *Annu. Rev. Mar. Sci.*, 2, 199–229, <https://doi.org/10.1146/annurev.marine.010908.163855>, 2010.

- Koelling, J., Atamanchuk, D., Wallace, D. W. R., and Karstensen, J.: Decadal variability of oxygen uptake, export, and storage in the Labrador Sea from observations and CMIP6 models, *Front. Mar. Sci.*, 10, 1202299, <https://doi.org/10.3389/fmars.2023.1202299>, 2023.
- Kosaka, Y., Kobayashi, S., Harada, Y., Kobayashi, C., Naoe, H., Yoshimoto, K., Harada, M., Goto, N., Chiba, J., Miyaoka, K., Sekiguchi, R., Deushi, M., Kamahori, H., Nakaegawa, T., Tanaka, T. Y., Tokuhito, T., Sato, Y., Matsushita, Y., and Onogi, K.: The JRA-3Q reanalysis, *J. Meteorol. Soc. Jpn.*, 102, 49–109, <https://doi.org/10.2151/jmsj.2024-004>, 2024.
- Koul, V.: Model output and Data for Decadal biogeochemical predictions for the bottom marine environment of the Northeast U.S. Continental Shelf [Dataset], Zenodo [data set], <https://doi.org/10.5281/zenodo.16985351>, 2025.
- Koul, V., Ross, A. C., Stock, C., Zhang, L., Delworth, T., and Wittenberg, A.: A Predicted Pause in the Rapid Warming of the Northwest Atlantic Shelf in the Coming Decade, *Geophys. Res. Lett.*, 51, e2024GL110946, <https://doi.org/10.1029/2024GL110946>, 2024.
- Krumhardt, K. M., Lovenduski, N. S., Long, M. C., Luo, J. Y., Lindsay, K., Yeager, S., and Harrison, C.: Potential Predictability of Net Primary Production in the Ocean, *Global Biogeochem. Cy.*, 34, e2020GB006531, <https://doi.org/10.1029/2020GB006531>, 2020.
- Lehodey, P., Alheit, J., Barange, M., Baumgartner, T., Beaugrand, G., Drinkwater, K., Fromentin, J.-M., Hare, S. R., Ottersen, G., Perry, R. I., Roy, C., van der Lingen, C. D., and Werner, F.: Climate variability, fish, and fisheries, *J. Climate*, 19, 5009–5030, <https://doi.org/10.1175/JCLI3898.1>, 2006.
- Lellouche, J.-M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C.-E., Bourdalle-Badie, R., Gasparin, F., Hernandez, O., Levier, B., Drillet, Y., Remy, E., and Le Traon, P.-Y.: Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1/12° high-resolution system, *Ocean Sci.*, 14, 1093–1126, <https://doi.org/10.5194/os-14-1093-2018>, 2018.
- Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E., Locarnini, R. A., Mishonov, A. V., Reagan, J. R., Seidov, D., Yarosh, E. S., and Zweng, M. M.: World ocean heat content and thermohaline sea level change (0–2000 m), 1955–2010, *Geophys. Res. Lett.*, 39, L10603, <https://doi.org/10.1029/2012GL051106>, 2012.
- Link, J. S., Thur, S., Matlock, G., and Grasso, M.: Why we need weather forecast analogues for marine ecosystems, *ICES J. Mar. Sci.*, 80, 2087–2098, <https://doi.org/10.1093/icesjms/fsad143>, 2023.
- Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D., and Stockdale, T.: Decadal Prediction: Can It Be Skillful?, *B. Am. Meteorol. Soc.*, 90, 1467–1486, <https://doi.org/10.1175/2009BAMS2778.1>, 2009.
- Meyer-Gutbrod, E. L., Greene, C. H., Davies, K. T. A., and Johns, D. G.: Ocean Regime Shift is Driving Collapse of the North Atlantic Right Whale Population, *Oceanography*, 34, 22–31, <https://doi.org/10.5670/oceanog.2021.308>, 2021.
- Minka, T. P.: Algorithms for maximum-likelihood logistic regression, Tech. Rep. 758, Carnegie Mellon University, Pittsburgh, PA, USA, <https://doi.org/10.1184/R1/6586418>, 2001.
- mom-ocean: MOM6, GitHub [code], <https://github.com/mom-ocean/MOM6> (last access: 25 September 2026), 2026.
- Mountain, D. G.: Variability of the Water Properties in NAFO Sub-areas 5 and 6 During the 1990s, *J. Northwest Atl. Fish. Sci.*, 34, 103–112, <https://doi.org/10.2960/J.v34.m475>, 2004.
- Mountain, D. G.: Labrador slope water entering the Gulf of Maine – Response to the North Atlantic Oscillation, *Cont. Shelf Res.*, 47, 150–155, <https://doi.org/10.1016/j.csr.2012.07.008>, 2012.
- Msadek, R., Delworth, T. L., Rosati, A., Anderson, W., Vecchi, G., Chang, Y.-S., Dixon, K., Gudgel, R. G., Stern, W., Wittenberg, A., Yang, X., Zeng, F., Zhang, R., and Zhang, S.: Predicting a decadal shift in North Atlantic climate variability using the GFDL forecast system, *J. Climate*, 27, 6472–6496, <https://doi.org/10.1175/JCLI-D-13-00476.1>, 2014.
- Niemisto, M., Contreras, A., Wahle, R. A., and Fields, D. M.: Increased temperature and acidification elevate the risk of starvation in American lobster larvae, *ICES J. Mar. Sci.*, 82, fsaf088, <https://doi.org/10.1093/icesjms/fsaf088>, 2025.
- NOAA Fisheries: 2024 State of the Ecosystem: New England, Tech. rep., <https://doi.org/10.25923/f8xc-hj17>, 2024.
- NOAA Fisheries: 2025 State of the Ecosystem: New England, Tech. rep., <https://doi.org/10.25923/zr75-a788>, 2025.
- NOAA-GFDL: CEFI-regional-MOM6, GitHub [code], <https://github.com/NOAA-GFDL/CEFI-regional-MOM6> (last access: 25 September 2026), 2026a.
- NOAA-GFDL: MOM6, GitHub [code], <https://github.com/NOAA-GFDL/MOM6> (last access: 25 September 2026), 2026b.
- NOAA-GFDL: NOAA Geophysical Fluid Dynamics Laboratory (GFDL), GitHub [code], <https://github.com/NOAA-GFDL> (last access: 25 September 2026), 2026c.
- Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., Feely, R. A., Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., Key, R. M., Lindsay, K., Maier-Reimer, E., Matear, R., Monfray, P., Mouchet, A., Najjar, R. G., Plattner, G.-K., Rodgers, K. B., Sabine, C. L., Sarmiento, J. L., Schlitzer, R., Slater, R. D., Totterdell, I. J., Weirig, M.-F., Yamanaka, Y., and Yool, A.: Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms, *Nature*, 437, 681–686, <https://doi.org/10.1038/nature04095>, 2005.
- Park, J.-Y., Stock, C. A., Dunne, J. P., Yang, X., and Rosati, A.: Seasonal to multiannual marine ecosystem prediction with a global Earth system model, *Science*, 365, 284–288, <https://doi.org/10.1126/science.aav6634>, 2019.
- Payne, M. R., Danabasoglu, G., Keenlyside, N., Matei, D., Miesner, A. K., Yang, S., and Yeager, S. G.: Skillful decadal-scale prediction of fish habitat and distribution shifts, *Nat. Commun.*, 13, 2660, <https://doi.org/10.1038/s41467-022-30280-0>, 2022.
- Pérez-Hernández, M. D. and Joyce, T. M.: Two modes of Gulf Stream variability revealed in the last two decades of satellite altimeter data, *J. Phys. Oceanogr.*, 44, 149–163, <https://doi.org/10.1175/JPO-D-13-0136.1>, 2014.
- Pershing, A. J., Alexander, M. A., Hernandez, C. M., Kerr, L. A., Le Bris, A., Mills, K. E., Nye, J. A., Record, N. R., Scannell, H. A., Scott, J. D., Sherwood, G. D., and Thomas, A. C.: Slow adaptation in the face of rapid warming leads to col-

- lapse of the Gulf of Maine cod fishery, *Science*, 350, 809–812, <https://doi.org/10.1126/science.aac9819>, 2015.
- Pershing, A. J., Alexander, M. A., Brady, D. C., Brickman, D., Curchitser, E. N., Diamond, A. W., McClenachan, L., Mills, K. E., Nichols, O. C., Pendleton, D. E., Record, N. R., Scott, J. D., Staudinger, M. D., and Wang, Y.: Climate impacts on the Gulf of Maine ecosystem: A review of observed and expected changes in 2050 from rising temperatures, *Elem. Sci. Anth.*, 9, 00076, <https://doi.org/10.1525/elementa.2020.00076>, 2021.
- Pregibon, D.: Logistic Regression Diagnostics, *Ann. Stat.*, 9, 705–724, <https://doi.org/10.1214/aos/1176345513>, 1981.
- Ramos, M. H., van Andel, S. J., and Pappenberger, F.: Do probabilistic forecasts lead to better decisions?, *Hydrol. Earth Syst. Sci.*, 17, 2219–2232, <https://doi.org/10.5194/hess-17-2219-2013>, 2013.
- Record, N., Pershing, A., and Rasher, D.: Early Warning of a Cold Wave in the Gulf of Maine, *Oceanography*, 37, 6–9, <https://doi.org/10.5670/oceanog.2024.506>, 2024.
- Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humenöder, F., Da Silva, L. A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M.: The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview, *Global Environ. Change*, 42, 153–168, <https://doi.org/10.1016/j.gloenvcha.2016.05.009>, 2017.
- Rittel, H. W. J. and Webber, M. M.: Dilemmas in a general theory of planning, *Policy Sci.*, 4, 155–169, <https://doi.org/10.1007/BF01405730>, 1973.
- Ross, A. C., Stock, C. A., Adcroft, A., Curchitser, E., Hallberg, R., Harrison, M. J., Hedstrom, K., Zadeh, N., Alexander, M., Chen, W., Drenkard, E. J., du Pontavice, H., Dussin, R., Gomez, F., John, J. G., Kang, D., Lavoie, D., Resplandy, L., Roobaert, A., Saba, V., Shin, S.-I., Siedlecki, S., and Simkins, J.: A high-resolution physical–biogeochemical model for marine resource applications in the northwest Atlantic (MOM6-COBALT-NWA12 v1.0), *Geosci. Model Dev.*, 16, 6943–6985, <https://doi.org/10.5194/gmd-16-6943-2023>, 2023.
- Ross, A. C., Stock, C. A., Koul, V., Delworth, T. L., Lu, F., Wittenberg, A., and Alexander, M. A.: Dynamically downscaled seasonal ocean forecasts for North American east coast ecosystems, *Ocean Sci.*, 20, 1631–1656, <https://doi.org/10.5194/os-20-1631-2024>, 2024.
- Roulston, M. S., Bolton, G. E., Kleit, A. N., and Sears-Collins, A. L.: A Laboratory Study of the Benefits of Including Uncertainty Information in Weather Forecasts, *Weather Forecast.*, 21, 116–122, <https://doi.org/10.1175/WAF887.1>, 2006.
- Saba, V., Borggaard, D., Caracappa, J. C., Chambers, R. C., Clay, P. M., Colburn, L. L., Deroba, J., DePiper, G., du Pontavice, H., Fratantoni, P., Ferguson, M., Gaichas, S., Hayes, S., Hyde, K., Johnson, M., Kocik, J., Keane, E., Kircheis, D., Large, S., Lipsky, A., Lucey, S., Mercer, A., Meseck, S., Miller, T. J., Morse, R., Orphanides, C., Reichert-Nguyen, J., Richardson, D., Smith, J., Vogel, R., Vogt, B., and Wikfors, G.: NOAA fisheries research geared towards climate-ready living marine resource management in the northeast United States, *PLOS Clim.*, 2, e0000323, <https://doi.org/10.1371/journal.pclm.0000323>, 2023.
- Saba, V. S., Griffies, S. M., Anderson, W. G., Winton, M., Alexander, M. A., Delworth, T. L., Hare, J. A., Harrison, M. J., Rosati, A., Vecchi, G. A., and Zhang, R.: Enhanced warming of the Northwest Atlantic Ocean under climate change, *J. Geophys. Res.-Oceans*, 121, 118–132, <https://doi.org/10.1002/2015JC011346>, 2016.
- Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T.-H., Kozyr, A., Ono, T., and Rios, A. F.: The Oceanic Sink for Anthropogenic CO₂, *Science*, 305, 367–371, <https://doi.org/10.1126/science.1097403>, 2004.
- Salisbury, J. E. and Jönsson, B. F.: Rapid warming and salinity changes in the Gulf of Maine alter surface ocean carbonate parameters and hide ocean acidification, *Biogeochemistry*, 141, 401–418, <https://doi.org/10.1007/s10533-018-0505-3>, 2018.
- Scully, M. E., Geyer, W. R., Borkman, D., Pugh, T. L., Costa, A., and Nichols, O. C.: Unprecedented summer hypoxia in southern Cape Cod Bay: an ecological response to regional climate change?, *Biogeosciences*, 19, 3523–3536, <https://doi.org/10.5194/bg-19-3523-2022>, 2022.
- Seidov, D., Mishonov, A., and Parsons, R.: Recent warming and decadal variability of Gulf of Maine and Slope Water, *Limnol. Oceanogr.*, 66, 3472–3488, <https://doi.org/10.1002/lno.11892>, 2021.
- Shearman, R. K. and Lentz, S. J.: Long-Term Sea Surface Temperature Variability along the U.S. East Coast, *J. Phys. Oceanogr.*, 40, 1004–1017, <https://doi.org/10.1175/2009JPO4300.1>, 2010.
- Siedlecki, S. A., Salisbury, J., Gledhill, D. K., Bastidas, C., Meseck, S., McGarry, K., Hunt, C. W., Alexander, M., Lavoie, D., Wang, Z. A., Scott, J., Brady, D. C., Mlna, I., Azetsu-Scott, K., Liberti, C. M., Melrose, D. C., White, M. M., Pershing, A., Vandemark, D., Townsend, D. W., Chen, C., Mook, W., and Morrison, R.: Projecting ocean acidification impacts for the Gulf of Maine to 2050: New tools and expectations, *Elem. Sci. Anth.*, 9, 00062, <https://doi.org/10.1525/elementa.2020.00062>, 2021.
- Smith, P. C., Houghton, R. W., Fairbanks, R. G., and Mountain, D. G.: Interannual variability of boundary fluxes and water mass properties in the Gulf of Maine and on Georges Bank: 1993–1997, *Deep-Sea Res. Pt. II*, 48, 37–70, [https://doi.org/10.1016/S0967-0645\(00\)00081-3](https://doi.org/10.1016/S0967-0645(00)00081-3), 2001.
- Steiger, J. H.: Tests for comparing elements of a correlation matrix, *Psychol. Bull.*, 87, 245–251, <https://doi.org/10.1037/0033-2909.87.2.245>, 1980.
- Stewart, J. A., Williams, B., LaVigne, M., Wanamaker, A. D., Strong, A. L., Jellison, B., Whitney, N. M., Thatcher, D. L., Robinson, L. F., Halfar, J., and Adey, W.: Delayed onset of ocean acidification in the Gulf of Maine, *Sci. Rep.*, 15, 2039, <https://doi.org/10.1038/s41598-024-84537-3>, 2025.
- Stock, C. A., Pegion, K., Vecchi, G. A., Alexander, M. A., Tommasi, D., Bond, N. A., Fratantoni, P. S., Gudgel, R. G., Kristiansen, T., O'Brien, T. D., Xue, Y., and Yang, X.: Seasonal sea surface temperature anomaly prediction for coastal ecosystems, *Prog. Oceanogr.*, 137, 219–236, <https://doi.org/10.1016/j.pocean.2015.06.007>, 2015.
- Stock, C. A., Dunne, J. P., Fan, S., Ginoux, P., John, J., Krasting, J. P., Laufkötter, C., Paulot, F., and Zadeh, N.: Ocean biogeo-

- chemistry in GFDL's Earth System Model 4.1 and its response to increasing atmospheric CO₂, *J. Adv. Model. Earth Syst.*, 12, e2019MS002043, <https://doi.org/10.1029/2019MS002043>, 2020.
- Stock, C. A., Dunne, J. P., Luo, J. Y., Ross, A. C., Van Oostende, N., Zadeh, N., Cordero, T. J., Liu, X., and Teng, Y.-C.: Photoacclimation and photoadaptation sensitivity in a global ocean ecosystem model, *J. Adv. Model. Earth Syst.*, 17, e2024MS004701, <https://doi.org/10.1029/2024MS004701>, 2025.
- Tommasi, D., Stock, C. A., Alexander, M. A., Yang, X., Rosati, A., and Vecchi, G. A.: Multi-annual climate predictions for fisheries: an assessment of skill of sea surface temperature forecasts for large marine ecosystems, *Front. Mar. Sci.*, 4, 201, <https://doi.org/10.3389/fmars.2017.00201>, 2017a.
- Tommasi, D., Stock, C. A., Hobday, A. J., Methot, R., Kaplan, I. C., Eveson, J. P., Holsman, K., Miller, T. J., Gaichas, S., Gehlen, M., Pershing, A., Vecchi, G. A., Msadek, R., Delworth, T., Eakin, C. M., Haltuch, M. A., Séférian, R., Spillman, C. M., Hartog, J. R., Siedlecki, S., Samhouri, J. F., Muhling, B., Asch, R. G., Pinsky, M. L., Saba, V. S., Kapnick, S. B., Gaitan, C. F., Rykaczewski, R. R., Alexander, M. A., Xue, Y., Pegion, K. V., Lynch, P., Payne, M. R., Kristiansen, T., Lehodey, P., and Werner, F. E.: Managing living marine resources in a dynamic environment: The role of seasonal to decadal climate forecasts, *Prog. Oceanogr.*, 152, 15–49, <https://doi.org/10.1016/j.pocean.2016.12.011>, 2017b.
- Townsend, D. W.: Sources and cycling of nitrogen in the Gulf of Maine, *J. Mar. Syst.*, 16, 283–295, [https://doi.org/10.1016/S0924-7963\(97\)00024-9](https://doi.org/10.1016/S0924-7963(97)00024-9), 1998.
- Townsend, D. W., Thomas, A. C., Mayer, L. M., Thomas, M. A., and Quinlan, J. A.: Oceanography of the northwest Atlantic continental shelf, *Sea*, 14, 119–168, 2006.
- Townsend, D. W., Pettigrew, N. R., Thomas, M. A., and Moore, S.: Warming waters of the Gulf of Maine: The role of Shelf, Slope and Gulf Stream Water masses, *Prog. Oceanogr.*, 215, 103030, <https://doi.org/10.1016/j.pocean.2023.103030>, 2023.
- Wallace, E. J., Looney, L. B., and Gong, D.: Multi-Decadal Trends and Variability in Temperature and Salinity in the Mid-Atlantic Bight, Georges Bank, and Gulf of Maine, *J. Mar. Res.*, 76, 163–215, <https://doi.org/10.1357/002224018826473281>, 2018.
- Waller, J., Bartlett, J., Bates, E., Bray, H., Brown, M., Cieri, M., Clark, C., DeVoe, W., Donahue, B., Frechette, D., Glon, H., Hunter, M., Huntsberger, C., Kanwit, K., Ledwin, S., Lewis, B., Peters, R., Reardon, K., Russell, R., Smith, M., Uraneck, C., Watts, R., and Wilson, C.: Reflecting on the recent history of coastal Maine fisheries and marine resource monitoring: the value of collaborative research, changing ecosystems, and thoughts on preparing for the future, *ICES J. Mar. Sci.*, 80, 2074–2086, <https://doi.org/10.1093/icesjms/fsad134>, 2023.
- Wilks, D. S.: Extending logistic regression to provide full-probability-distribution MOS forecasts, *Meteorol. Appl.*, 16, 361–368, <https://doi.org/10.1002/met.134>, 2009.
- Williams, E. J.: The Comparison of Regression Variables, *J. Roy. Stat. Soc. B*, 21, 396–399, <https://doi.org/10.1111/j.2517-6161.1959.tb00346.x>, 1959.
- Xue, L. and Cai, W.-J.: Total alkalinity minus dissolved inorganic carbon as a proxy for deciphering ocean acidification mechanisms, *Mar. Chem.*, 222, 103791, <https://doi.org/10.1016/j.marchem.2020.103791>, 2020.
- Yang, X., Delworth, T. L., Zeng, F., Zhang, L., Cooke, W. F., Harrison, M. J., Rosati, A., Underwood, S., Compo, G. P., and McColl, C.: On the development of GFDL's decadal prediction system: Initialization approaches and retrospective forecast assessment, *J. Adv. Model. Earth Syst.*, 13, e2021MS002529, <https://doi.org/10.1029/2021MS002529>, 2021.