

1 **Decadal biogeochemical predictions for the changing bottom marine**
2 **environment of the ~~Northeast U.S. Continental Shelf~~ Gulf of Maine**

3
4 Vimal Koul^{1,2}, Andrew C. Ross³, Charles Stock³, Liping Zhang³,
5 Andrew T. Wittenberg³ and Thomas Delworth³
6

7 ¹Cooperative Institute for Modeling the Earth System, Princeton University,
8 Princeton, New Jersey, United States
9

10 ²Now at Centre for Atmospheric Sciences, Indian Institute of Technology Delhi,
11 New Delhi, India
12

13 ³NOAA/OAR/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey,
14 United States
15

16 **Correspondence:** Vimal Koul
17 **E-mail:** vkoul@iitd.ac.in
18

19 **Abstract**

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

38 **Introduction**

39 The global ocean has undergone unprecedented biogeochemical transformation
40 since the onset of the industrial era — absorbing approximately 30% of
41 anthropogenic CO₂ emissions, and over 90% of the excess heat associated with
42 accumulating greenhouse gases (Sabine et al., 2004; Feely et al., 2023,
43 Friedlingstein et al., 2025, Johnson and Lyman, 2020). This oceanic ~~CO₂~~CO₂ and
44 heat uptake has initiated a cascade of interconnected changes that represent some
45 of the most profound shifts in ocean conditions experienced in the past 50 years.
46 Warming has altered ocean stratification and circulation (Levitus et al., 2012;
47 Abram et al., 2016). Acidification has challenged the basic chemistry upon which
48 marine calcifiers depend (Caldeira and Wickett, 2003; Orr et al., 2005). Declining
49 oxygen concentrations have pressed the metabolic limits of marine organisms
50 (Keeling et al., 2010; Breitburg et al., 2018). Ecosystem and fisheries challenges
51 posed by climate change are compounded by those linked to natural climate
52 variability (Lehodey et al., 2006). A capacity to anticipate physical and
53 biogeochemical changes, whether from climate change or natural variability, may
54 help coastal communities and industries mitigate potential negative impacts, and
55 realize new opportunities that may arise (Clark et al., 2001, Tommasi et al., 2017).

56 The Northeast U.S. Continental Shelf (NEUS) — and particularly its highly
57 productive semi-enclosed sea, the Gulf of Maine (GOM) — has been subject to

58 both ocean change trends and pronounced variability. Dynamic warming (Pershing
59 et al., 2015, Saba et al., 2016; Alexander et al., 2018), deoxygenation (Claret et al.,
60 2018) and acidification (Salisbury and Jönsson, 2018, Siedlecki et al. 2021,
61 Stewart et al, 2025) signals have emerged from an interplay of shifting circulation
62 patterns, alterations in the Gulf Stream position, and regional water masses,
63 particularly variations between cold, fresh Labrador Slope Water (LSW) and
64 Warm Slope Water (WSW) associated with greater Gulf Stream influence, that
65 modulate global warming signals through local oceanographic processes (Chen et
66 al., 2020; Gangopadhyay et al., 2019; Brickman et al., 2018; Gonçalves Neto et al.,
67 2021; Friedland and Hare, 2007, Seidov et al., 2021). In the GOM, oceanic
68 variability from interannual to longer time scales is understood to be largely driven
69 by changes in the contributions of these and other source waters which, in turn,
70 modify the temperature, salinity, and nutrient conditions across the entire
71 ecosystem (Mountain, 2012; Townsend et al., 2023). This water mass variability
72 has also generated complex and sometimes counterintuitive local biogeochemical
73 responses to climate change, including an influx of high-alkalinity Gulf Stream
74 waters driving increased buffering to acidification in certain areas despite rising
75 ~~CO₂~~atmospheric CO₂ concentrations (Stewart et al., 2025; Salisbury and Jönsson,
76 2018). Such water mass changes may also influence coastal hypoxia risk by

77 altering stratification patterns that control vertical mixing and oxygen dynamics in
78 coastal bottom waters (Scully et al., 2022).

79 Rapid warming of the GOM has already contributed to the decline of its iconic
80 Atlantic cod fishery and triggered cascading ecosystem changes (Pershing et al.,
81 2015, 2021; Hare et al., 2016). Continued warming, deoxygenation, and
82 acidification pose significant threats to commercially important fisheries,
83 particularly American lobster and shellfish. American lobsters show stage-specific
84 vulnerabilities to acidification and warming that vary significantly between
85 regional populations (Harrington et al. 2019, Niemisto et al. 2025). Early life
86 stages of shellfish consistently show greater sensitivity to acidification than adults,
87 creating recruitment bottlenecks that may drive population declines (Fay et al.
88 2017). Impacts extend beyond commercially harvested species to affect protected
89 species such as Atlantic puffins and northern right whales (Meyer-Gutbrod et al.,
90 2021).

91 Decisions by coastal communities and managers that determine long-term resource
92 and community resilience are made across time horizons and must consider both
93 variability and change. The interannual to decadal time horizon (1-10 years) is
94 particularly notable for its relevance to the setting of annual catch limits, the
95 formulation of stock rebuilding plans, and industry capitalization (Tommasi et al.,
96 2017). 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., 2017; Saba et al., 2016, Ross et al., 2023). High-resolution ocean modeling and enhanced observational networks are revealing substantial promise for predicting biogeochemical variability in this highly dynamic region (Gehlen et al., 2015, Ross et al., 2023, 2024, Chen et al., 2021, 2024, Koul et al., 2024).~~

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 \(Tommasi et al., 2017; Saba et al., 2016, Ross et al., 2023\). 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. [1a](#),[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). ~~This~~[Previous work regional ocean prediction studies demonstrated that this physical transition, and others like it, were predictable \(Koul et al., 2024\). Such](#) shift ~~has~~, however, also have significant biogeochemical implications, ~~however~~, 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

137 to downscale global [decadal](#) predictions (Delworth et al. 2020), addresses this
138 time-critical challenge by capturing both the physical and biogeochemical
139 dynamics of shifting water masses and persistent change. [We furthermore shift the](#)
140 [focus of the evaluation to the benthic environment, which is critical for the largest](#)
141 [and most lucrative Gulf of Maine fisheries \(e.g., lobster, scallops and groundfish,](#)
142 [Waller et al., 2023\).](#) Through a comprehensive 10-member ensemble approach
143 spanning 60 years of initialized retrospective predictions, accompanied by
144 uninitialized simulations that include only the impacts of external radiative
145 forcings (~~~6000~~[\(~6500](#) years of simulations), we demonstrate skilful decadal (1-10
146 year) predictions of benthic physical and biogeochemical conditions in the Gulf of
147 Maine. Our experiments allow us to isolate the impacts of radiatively forced ocean
148 warming on prediction skill, and the impacts of predictable water mass variations.
149 Finally, we present probabilistic forecasts for the coming years that may offer
150 fisheries managers actionable insights into the near-term trajectory of the benthic
151 environment in the GOM.

Results and Discussion:

Bottom Temperature and Salinity

The regionally downscaled retrospective simulations, 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. 1b). 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, duPontavice et al., 2023). Real time Fishing Industry SHared Bottom Oceanographic Timeseries (Fishbot) data and independent observation-based datasets (Fig. 1a, Fig. S1 and Fig. 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 (Fig 1b, Fig. 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 (Fig 1b, Fig. S3, $r = 0.84$ versus $r = 0.90$).~~

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. 1c-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 covariability, though meteorologically-linked variations in the winter mixed layer also dictate large episodic variations in bottom O_2 (Fig. 2, Fig. S5). The close correspondence between the observed Fishbot data and the modeled bottom O_2 (Fig. 1c) 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 atmospheric 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. 1d, e). After removing the externally-
forced ocean acidification signal, Ω_{ar} variations are positively correlated with
temperature (Fig. 2, $r = 0.75$). This is consistent with the carbonate chemistry
response to warming and increased alkalinity in warm and saline WSW waters
(Salisbury and Jonsson, 2018). As in Salisbury and Jonsson’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, but this variability is
less pronounced for pH. The more recent shift toward colder and fresher
conditions, however, amplifies the acidification trend.~~

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 CO₂) 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. 1–2 suggested that our retrospective ensemble 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. 3).

Bottom temperature predictions exhibit significant positive anomaly correlations in most of the 10 lead years (Fig. 3a); and in the first lead year, the prediction skill exceeds the skill from the climate change signal (indicated by the diamonds and squares). Both of these 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 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).

The model shows encouraging capacity to anticipate shifts in bottom temperature trends, as demonstrated by individual 10-year predictions during periods when the GOM transitioned from warming to cooling (Fig. 4a). While there are important caveats regarding initialization states, the model exhibits skill by predicting anomalies that diverge from the uninitialized state. This is particularly evident in the 2002 initialization (close to the uninitialized state), which accurately captured a cooling trend and its subsequent reversal in the mid-2000s. In contrast, initializations from the 1990s (also close to the uninitialized state) showed poor skill. The strong correlation observed in the initial lead years (Fig. 3a) suggests that most forecasts align well with reference anomalies.

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. 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. 3b). This is also apparent from the time series (Fig. 1c) 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 presence of sharp and likely less predictable meteorologically driven variations in the depth of the winter mixed layer. The winter 2003/04 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 (Fig. 4, Fig. 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 4b). 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 3b), and the significant drop thereafter.

Both bottom pH and Ω_{eff} exhibit a declining trend (Fig 1e, 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_{eff} 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_{eff} , whereas the climate change signal has a greater impact on the prediction skill of bottom pH (Fig 3 c,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_{eff} 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_{eff} , further demonstrates the additional contribution of internal variability to the prediction skill of Ω_{eff} . Individual 10-year forecasts provide further evidence of the capability of the model to track past deviation of bottom pH and Ω_{eff} from the long term trend (Fig 4 c,d).

Probabilistic predictions

The preceding analyses leveraged over 6000 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 demands actionable information on uncertainties. 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 accuracy of anomalously warm or anomalously cold forecasts (Tommasi et al. (2017a), see Methods). 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 timelines in actively managed fisheries (Tommasi et al., 2017b). Accuracy is defined as the ratio of correct forecasts (i.e., those that either correctly forecasted the occurrence of a tercile, or its absence) over the total number of forecasts. Extended logistic regression (Wilks et al., 2009) was used to calibrate forecast uncertainty based on the forecast error (see Methods).

Our regional downscaled predictions skilfully distinguish between tercile classes, providing a more nuanced view of prediction capability. The forecast accuracy for a given tercile generally exceeds the accuracy of a random forecast (0.56) or the accuracy of consistently negative prediction (0.67) in all cases, but differs significantly by variable (Fig. 5). The ensemble confidently predicted the observed cooling in the mid-2000s and the subsequent warming phase that extended from 2005 to approximately 2020 (Fig. 5). Oxygen generally has lower accuracy (0.67–0.81) compared to bottom temperature (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 accuracy (0.76–0.88) than Ω_{eff} (0.62–0.79).~~

~~The high accuracy of prediction for bottom temperature, O_2 , pH and Ω_{eff} likely reflects both methodological and mechanistic factors. The bottom temperature, O_2 , pH and Ω_{eff} are validated against the same regional downscaled simulation from which initial conditions were derived (see Methods). While this simulation recreates observed O_2 patterns and covariances with well-constrained physical properties to the extent that they have been observed (Figs. 1–2, Fig. S7), it should be viewed as an upper limit for prediction skill. Additional autonomous biogeochemical observing platforms capable 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 a 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_{eff} conditions, both of which are supported 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.~~

~~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 U.S. 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 U.S. 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 autocorrelated noise. Here, given the scarcity of past deep alongshelf biogeochemical observations, we utilized a numerical model as the reference for forecast verification. This downscaled simulation was validated using the best available observations (Fig 1, Fig. S1, Fig. S2, Fig. S3, Fig. S6 and Fig. S7). Consequently, the reported skill represents potential skill and may be an overestimation. However, the model successfully predicted known past variations, and Koul et al. (2024) demonstrated skill for surface temperature when compared against actual observations. Furthermore, the model is likely accurate in predicting the forced component of skill.

Our forecasts of continued cooling alongside below normal bottom pH, Ω_{eff} 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

405 ~~Fisheries Initiative. Recognizing that marine resource management is a wicked~~
406 ~~problem (Jentoft and Chuenpagdee 2009; Rittle and Webber, 1973), however,~~
407 ~~careful application-specific analyses of the potential costs, benefits, and tradeoffs~~
408 ~~of using these forecasts for decision-making will be necessary. In the fields such as~~
409 ~~weather and hydrological forecasting, probabilistic predictions are well known to~~
410 ~~enable better decisions to be made (e.g., Ramos et al. 2013, Roulston et al. 2006),~~
411 ~~and the probabilistic decadal predictions developed in this study are an important~~
412 ~~step towards realizing these benefits in the marine resource space (Payne et al.~~
413 ~~2017).~~

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 > 1000m. 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.

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)

453 ~~coupled to the ocean model,~~[2025\)](#) and improve the initial [physical](#) state estimates
454 by relaxing the downscaled simulation towards the EN4 dataset (Good et al.,
455 ~~2013).~~[2013\)](#) at depth greater than 1000m.

456 **The global SPEAR model**

457 We use a 10-member SPEAR-based global simulation (SPEAR_LO; Delworth et
458 al., 2020; Yang et al., 2021) with a low-resolution configuration (1° ocean and sea
459 ice with near-equatorial refinement; and 1° atmosphere and land) to obtain surface
460 forcing and open boundary conditions for the regional model. ~~The~~[We use forcing](#)
461 [and boundary conditions from the](#) SPEAR_LO [global coupled reanalysis to create](#)
462 [the NWA12 simulations used to initialize our hindcasts \(i.e.,](#)
463 [NWA12_EN4_NUDG, Table 1\), and forcing from SPEAR hindcasts](#)
464 [\(SPEAR_HIND\) to drive NWA12_HIND. For the SPEAR_LO reanalysis,](#)
465 atmospheric winds and temperature ~~are~~[were](#) restored toward the 6-hourly Japanese
466 Reanalysis (JRA3Q; Kosaka et al., 2024), and the oceanic SST is restored toward
467 the interpolated monthly-mean observations from the NOAA Extended
468 Reconstructed Sea Surface Temperature version 5 (ERSSTv5) dataset (Huang et
469 al., 2017). [The reanalysis ensemble spread originates through alternative deep](#)
470 [ocean initializations and is maintained due to the gentle surface-only assimilation](#)
471 [approach used by SPEAR_LO \(Yang et al., 2021\).](#)

~~The regional downscaled simulations~~Regional Downscaled Simulation
(NWA12_EN4_NUDG)

~~We use SPEAR_LO output 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. To initialize the decadal predictions, we use the interior solution of the downscaled simulation (named NWA12_EN4_NUDG) which is relaxed with a 90-day 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 meters. Our tests with different choices of relaxation time scales and depth masking showed that the 90-day/1000-meter mask relaxation was (a) sufficient to bring the model solution closer to the observation-based sub-surface variability, particularly in the late 1990s, (b) appropriate to avoid rapid oceanic re-adjustment in shallower waters upon initialization of retrospective predictions wherein such relaxation towards EN4 is not carried out, and (c) necessary to avoid nudging on the shelf where the 1° EN4 shows low correlation with observed salinity despite maintaining high correlation with observed temperature (Fig. S2, Fig. S3). 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. Furthermore, the relaxation is limited to ocean~~

~~temperature and salinity only, and the ocean biogeochemical variability is dictated purely by the regional model solution. 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). The climatological river freshwater discharge, nutrient, carbon and alkalinity data were also used as described in Ross et al., (2023).~~

The regional high resolution decadal predictions

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-day 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 meters. This moderate relaxation helps keep the model solution closer to the observation-based sub-surface variability, particularly in the

511 late 1990s when the boundary conditions overestimated the propagation of a cool,
512 fresh signal into the domain. Limiting the relaxation to waters deeper than 1000 m
513 helps avoid rapid oceanic re-adjustment in shallower waters upon initialization of
514 retrospective predictions wherein such relaxation towards EN4 is not carried out
515 and is necessary to avoid nudging on the shelf where the 1° EN4 shows low
516 correlation with observed salinity despite maintaining high correlation with
517 observed temperature (Fig. S2, Fig. S3). We also note that the bottom waters of the
518 Gulf of Maine analysed in this study (the deep basins and the Northeast Channel)
519 lie well inshore of the 1000 m isobath and are therefore not directly nudged
520 towards EN4. The relaxation constrains only the off-shelf slope source waters. We
521 do not relax the regional model towards the parent SPEAR model, since SPEAR
522 has coarse ocean resolution and does not assimilate sub-surface ocean data, making
523 it less constrained towards observations than the EN4-gridded dataset.

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

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

535 **Regional Ocean Hindcasts (NWA12_HIND)**

536 Following Koul et al 2024, we use the ~~10-member~~ regional simulation from the
537 NWA12_EN4_NUDG model to start ~~10-year-long~~[yearlong](#) retrospective ocean
538 physical-biogeochemical predictions (NWA12_HIND) every year [on January 1st](#)
539 from 1965 to ~~2024~~[2025](#). The last ~~10-yearlong~~ [simulation, which shifts from](#)
540 [hindcast to forecast, is from 01-01-2025 to 31-12-2034. The 10-member](#)
541 [NWA12_HIND ensemble are started from ocean initial conditions taken from the](#)
542 [corresponding 10 members of NWA12_EN4_NUDG, and integrated forward using](#)
543 [atmospheric forcing and open ocean boundary conditions from the 10 members of](#)
544 [SPEAR_HIND](#). To ease the transition between the global and regional model of
545 the complex inflow and outflow along the northern boundary, each member of the
546 ~~10-year-long~~[yearlong](#) predictions is relaxed towards the corresponding
547 SPEAR_LO prediction north of 45°N. The strength of the ~~restoring~~[restoration](#)
548 corresponds to a relaxation time scale of 90 days at the northern boundary and
549 decreases ~~sinusoidally~~[gradually](#) to 0 at 45°N. No bathymetric mask is applied, as
550 the relaxation drops to zero north of the targeted shallower regions. The

~~atmospheric forcing is used directly as available from the SPEAR_LO decadal predictions.~~ latitudinal damping rate r (inverse of relaxation time scale) was prescribed as:

~~The regional high-resolution uninitialized (historical) simulations~~

~~To quantify the signal of forced climate change, we created a 10-member ensemble of uninitialized “historical” ocean simulations (NWA12_HIST). These simulations are linked to specific historical years through the prescribed greenhouse gas concentrations and other radiative forcings, thus providing an estimate of the historical climate change signal. They were produced through dynamic downscaling of the first 10 members from a larger 30-member ensemble of SPEAR_LO historical climate simulations (SPEAR_HIST). The SPEAR_HIST dataset spans from 1850 to 2100 and incorporates future greenhouse gas concentrations based on the SSP585 emissions scenario. It should be noted that differences between SSP585 and other emissions scenarios are negligible over the near-future time period (next 1–10 years) examined in this study.~~

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

for latitudes from 45°N to max(latitude) in the regional model, and zero south of 45°N and

$r_{\text{max}} = 1/90 \text{ day}^{-1}$ at max(latitude), where max(latitude) in the regional model is $\approx 58^\circ\text{N}$

Historical Climate Change Simulations (NWA12_HIST)

The NWA12_HIST is a continuous 75-yearlong simulation started on 01-01-1960 and ending in 31-12-2034. 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.

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 interim analysis products (Lellouche et al., 2018) and the 1° EN4 gridded dataset (EN.4.2.2, Good et al., 2013) which is used to assess a longer period from 1965-2024 (Fig. S1, Fig. [23](#) and Fig. [3-5](#)). For bottom O₂ and bottom

Ω_{ar} , we used the Fishing Industry SHared Bottom Oceanographic Timeseries data (https://erddap.ondeckdata.com/erddap/tabledap/fishbot_realtime.html), hereafter called “Fishbot”, and vessel (Jiang et al., 2021) and glider-based (IOOS Glider Data Assembly Center: <https://gliders.ioos.us/erddap/index.html>) 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).

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 meters. Annual anomalies were then derived by subtracting the 1994-2023 mean from the annual mean time series. Figures Fig. S1, Fig. ~~S2~~[S2](#) and Fig. 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 \(n\).](#)

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°W 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 1a)~~[\(Fig. 1\)](#) analyzed in this study were computed from water mass definitions as described in Ross et al 2023.

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

629 Ecological Production Units (GOM_EPU, NOAA Fisheries, 2024). For bottom
630 temperature, bottom O₂, bottom pH and bottom Ω_{ar} the NWA12_EN4_NUDG
631 simulation was used as the observational ~~reference~~[reference](#), and the skill was
632 calculated over the period 1965-2024. The lead-dependent climatology (1994-2023
633 mean) was removed from the initialized hindcasts before calculating the skill
634 metric. This choice of using NWA12_EN4_NUDG as the observational reference
635 was made because continuous spatio-temporal observational records of bottom
636 temperature, O₂, pH and bottom Ω_{ar} , for any of the EPUs or Large Marine
637 Ecosystems (LMEs) were not available. However, to show that the
638 NWA12_EN4_NUDG was a good proxy for in-situ observations for the skill
639 estimation, we carried out a point-by-point comparison of temporally and spatially
640 ~~collocated~~[co-located](#) model output and available observations (Fig. S7), as well as
641 additional evaluation of interannual variations and covariance with better-
642 constrained physical properties (~~Figs. 1-2,~~[\(Fig. S1,](#) Fig. ~~S1,~~[Fig. 23](#) and Fig. ~~3)-5).~~
643 A simple persistence-based skill, defined as the lagged autocorrelation of observed
644 anomalies, was also compared with the hindcast skill.

645 We assessed the statistical significance of the prediction skill from the initialized
646 NWA12_HIND retrospective predictions. This was done by conducting a two-
647 tailed t-test against a null hypothesis of no correlation, adjusting for the reduced
648 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.

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

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

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 prediction x : $f(x) = \beta_0 + \beta_1 * x$, and the extension function g is also a simple function of the threshold q : $g(q) = \beta_2 * q$. 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 Figure 5.7. 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).

668 Furthermore, the ensemble spread has low correlation with the actual retrospective
669 forecast error. As a result, extended logistic regression provides a reasonable
670 improvement to the estimates of the forecast uncertainty,

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

673
$$\text{Accuracy} = (\text{True Positives} + \text{True Negatives}) / \text{Total Forecasts}$$

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

680 We also calculated the ranked probability skill score (RPSS) for each variable to
681 assess the accuracy of the predictions across all three categories. Persistence of the
682 previous year's category was used as the reference for the RPSS, so positive skill
683 scores indicate that the model forecasts have some ability to predict transitions
684 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

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

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 $RPS = 0$. 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):

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

where $RPS(\text{forecast})$ is the sum of the ranked probability scores comparing the forecast probabilities with the observed categories, and $RPS(\text{persistence})$ 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.

Decomposition of the drivers of bottom O₂, Alkalinity and Aragonite

Saturation State (Ω_{ar})

Bottom O₂ saturation ([O₂] 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 (PO₄ and SiO₄) that are omitted here, with negligible effect on the bottom anomalies.

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

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

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

Results and Discussion:

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, duPontavice et al., 2023](#)). Real-time [Fishing Industry SHared Bottom Oceanographic Timeseries \(Fishbot\)](#) data and independent observation-based datasets (Fig. 2a, Fig. S1 and Fig. 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 (Fig 2b, Fig. 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 (Fig 1b, Fig. S3, $r = 0.84$ versus $r = 0.90$).

Bottom Oxygen, pH and Aragonite Saturation Ratio

779 The water mass variations described above are also accompanied by
780 biogeochemical shifts in both the model and observations (Fig. 2c-e). The annual
781 mean bottom oxygen (O₂) anomalies in the GOM show an inverse relationship
782 with bottom temperature and salinity anomalies. The warming and water mass
783 transitions described above would be expected to influence dissolved oxygen
784 concentrations through differences in O₂ concentration in water masses (Claret et
785 al., 2018), solubility effects, and potential changes in biological oxygen demand.
786 The rise in deep oxygen levels from the mid-1990s, peaking in the latter half of
787 2004 to 2005, corresponds with the cool, fresh, high LSW conditions well
788 documented in the observational record (Mountain, 2012; Feng et al., 2016). LSW
789 is high in oxygen due to high solubility and rapid oxygen uptake during its
790 formation in the Labrador Sea (Koelling et al., 2023). Since 2005 the bottom
791 oxygen levels have largely dropped, aside from episodic increases in 2015, 2019,
792 and 2024. The relationship between internal (natural) variations in bottom
793 temperature and bottom O₂ confirms that the proportion of WSW/LSW in the
794 Northeast Channel acts as an important driver of this co-variability, though
795 meteorologically linked variations in the winter mixed layer also dictate large
796 episodic variations in bottom O₂ (Fig. 5, Fig. S5). The close correspondence
797 between the observed Fishbot data and the modeled bottom O₂ (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 atmospheric CO₂ (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, e), which is consistent with Salisbury and Jonsson (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 Jonsson, 2018). As in Salisbury and Jonsson'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} .

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 O₂ was

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

828 For the carbonate system, we examined the bottom dissolved inorganic carbon
829 (DIC) and total alkalinity, which mix conservatively and respond in
830 straightforward ways to gas exchange, respiration and $CaCO_3$ cycling. We also
831 examine alkalinity minus DIC ($Alk - DIC$), a proxy that helps separate the
832 mechanisms driving pH and Ω_{ar} (Xue and Cai, 2020; Fig. 4b). In the model the Alk
833 $- DIC$ anomaly is nearly identical to the bottom Ω_{ar} anomaly ($r = 0.996$; Fig. 4b, d)
834 but tracks pH relatively loosely ($r = 0.82$, because pH also responds directly to
835 warming). On interannual timescales $Alk - DIC$ (and hence Ω_{ar}) is set by water-
836 mass alkalinity changes, while its long-term decline ($-2.2 \mu\text{mol kg}^{-1} \text{decade}^{-1}$),

837 rising DIC outpacing a smaller alkalinity increase, is the anthropogenic
838 acidification trend.

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

849 The long-term Ω_{ar} decline ($-0.015 \text{ decade}^{-1}$) is likewise led by rising DIC, whose
850 direct effect ($-0.035 \text{ decade}^{-1}$) is more than halved by two opposing influences:
851 increasing alkalinity ($+0.016 \text{ decade}^{-1}$) and warming ($+0.004 \text{ decade}^{-1}$) (Fig. 4d).

852 Warming thus raises Ω_{ar} even as it lowers pH, which, together with the larger
853 water-mass alkalinity signal, is why the Ω_{ar} decline is weaker and less monotonic
854 than that of pH. On interannual timescales, Ω_{ar} variability is instead governed
855 primarily by alkalinity (about 57% of the detrended variance; $r = 0.91$ with the
856 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 (Fig. 2a, b, Fig S3), is retained as an explicit driver of the pH and Ω_{ar} anomalies in this decomposition (Fig. 4c, d), where its direct contribution is small.

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. 2-3 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

877 bottom habitat quantities. Comparison of extensive retrospective forecasts against
878 these patterns show that some aspects of both signals were successfully predicted,
879 but others were not (Fig. 5).

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

891 Initializations at states far from the uninitialized state, the 1980s warming, the
892 rapid 2005-2015 warming and the recent shift to cooler/fresher conditions from
893 2020 onward, are all correctly predicted by the retrospective predictions (Fig. 6).
894 The 2025 forecast shows a slow reversal of the recent cooling/freshening during
895 the next decade. These results could be interpreted as the model simply relaxing

896 toward the uninitialized state. However, the overall good correlation skill indicates
897 that the model accurately captures the timing and magnitude of such transitions.

898 The prediction skill for bottom O₂ drops to 0 beyond 2 years and much of the
899 signal for this important habitat factor is linked to predictable internal (natural)
900 variability (Fig. 5b). This is also apparent from the time series (Fig 2c) which
901 shows that bottom O₂ does not have a uniform long-term trend as was the case for
902 bottom temperature. The rapid drop in prediction skill may also be linked to the
903 presence of sharp and likely less predictable meteorologically driven variations in
904 the depth of the winter mixed layer. The winter 2003/04 is a prime example of this
905 phenomenon when large positive winter MLD anomalies in our regional
906 downscaled model drove ventilation of the bottom water and strong positive
907 bottom oxygen anomalies (Fig. 6, Fig. S5). Similar peaks in bottom O₂ coinciding
908 with deep winter mixing and positive MLD anomalies are seen in 2015 and 2019.

909 The decadal predictions initialized before and after the 2004 peak (in 2002 and
910 2006 respectively) show that the ensemble mean prediction has difficulties
911 predicting the extent of the rise in bottom O₂, but the subsequent decay of these
912 elevated bottom O₂ levels is well predicted by the model (Fig 6b). These
913 characteristics of the past variability in bottom O₂ are consistent with the high skill
914 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, f), yet they differ in how much the climate change signal influences their prediction skill. Ocean pH is primarily controlled by increasing atmospheric CO₂, 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 5 c,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 6 c,d).

Probabilistic predictions

935 The preceding analyses leveraged over 6500 years of retrospective forecasts to
936 assess the capacity of the ensemble mean prediction to anticipate physical and
937 biogeochemical changes on a decadal time horizon. While results suggested
938 significant capability arising from predictable climate variability and change
939 signals, decision-making benefits from actionable information on uncertainties
940 (e.g., Ramos et al. 2013, Roulston et al. 2006). A coastal resource manager or
941 fishing industry representative might ask, “How reliable are forecasts for
942 anomalously warm or cool conditions?” To quantify this in more accessible terms,
943 we calculated the accuracy of probabilistic forecasts derived from the model
944 simulations (Tommasi et al. (2017a), see Methods). We used terciles to define
945 above, normal, and below-normal conditions and focused on predicted conditions
946 averaged over the next 3 years, which corresponds to typical stock assessment
947 timelines in actively managed fisheries (Tommasi et al., 2017b). Extended logistic
948 regression (Wilks et al., 2009) was used to calibrate forecast uncertainty based on
949 the forecast error (see Methods).

950 Our regional downscaled predictions skilfully distinguish between tercile classes.
951 All four variables have positive ranked probability skill scores, and the forecast
952 accuracy for a given tercile exceeds the accuracy of a random forecast (0.56) and
953 the accuracy of consistently negative prediction (0.67) in all but two cases (Fig. 7).
954 The ensemble confidently predicted the observed cooling in the mid-2000s and the

955 subsequent warming phase that extended from 2005 to approximately 2020 (Fig.
956 7).

957 The RPSS for all quantities is > 0 , indicating the model has some capacity to
958 anticipate changes (i.e., probabilistic skill beyond the persistence forecast, see
959 Methods). Oxygen has slightly lower skill (RPSS 0.46, accuracy 0.67-0.81)
960 compared to bottom temperature (RPSS 0.47, accuracy 0.72-0.90), reflecting both
961 the weak predictable climate change signals and the influence of less predictable
962 atmospheric variability. Similarly, predictions for bottom pH, which has a very
963 strong predictable climate change signal, have higher skill (RPSS 0.6, accuracy
964 0.76-0.88) than Ω_{ar} (RPSS 0.51, accuracy 0.62-0.79).

965 The high accuracy of prediction for bottom temperature, O_2 , pH and Ω_{ar} likely
966 reflects both methodological and mechanistic factors. The bottom temperature, O_2 ,
967 pH and Ω_{ar} are validated against the same regional downscaled simulation from
968 which initial conditions were derived (see Methods). While this simulation
969 recreates observed O_2 patterns and covariances with well-constrained physical
970 properties to the extent that they have been observed (Figs. 2-3, Fig. S7), it should
971 be viewed as an upper limit for prediction skill. Additional autonomous
972 biogeochemical observing platforms capable of observing shelf environments may
973 help address this limitation in the future (Henson 2014, Chai et al., 2020).

974 Looking ahead over the next three years (2025-2027), we forecast below-normal
975 bottom temperatures with high confidence (88.3% probability), supported by
976 93.3% retrospective accuracy when issuing similar high confidence forecasts
977 (14/15 hits). We also predict above-normal bottom O₂ levels with high confidence
978 (92.0% probability), with an 85.7% retrospective accuracy at comparable
979 confidence levels (6/7 hits). Additionally, we forecast with very high confidence
980 (>99% probability) both below-normal bottom pH and below-normal Ω_{ar}
981 conditions, both of which are supported by perfect retrospective accuracy when
982 issuing very high confidence forecasts, albeit with limited sample size (9/9 hits and
983 1/1 hit respectively). While these forecasts represent our best estimate, both the
984 confidence levels and the underlying sample sizes should be considered when
985 making decisions based on these predictions, particularly for the forecasts where
986 limited similarly confident historical forecasts constrain our ability to fully validate
987 forecast skill. Nevertheless, the dynamically consistent physical and
988 biogeochemical forecasts provide compelling evidence that the ongoing watermass
989 reorganization will continue to influence bottom water conditions through the
990 2025-2027 period.

991 **Discussion: Towards the Future of Fisheries Management**

992 Future projections indicate the entire GOM will experience suboptimal aragonite
993 saturation for most of the year by 2050 under high emissions scenarios (Siedlecki
994 et al., 2021). The region's demonstrated sensitivity to water mass-driven alkalinity
995 variations makes it particularly vulnerable to rapid acidification once rising CO₂
996 exceeds the buffering ability of the projected increased Gulf Stream water,
997 potentially triggering swift transitions to persistently corrosive conditions that
998 could impact economically vital shellfish industries (Ekstrom et al., 2015).
999 Although the recent transition from a decade of warm water influence to emerging
1000 cold water conditions across portions of the Northeast U.S. Continental Shelf,
1001 including the deep GOM (Record et al., 2024), may buy time for fisheries to adapt
1002 to the projected long-term warming trend, it also may accelerate or alter the
1003 projected biogeochemical shifts.

1004 We assessed the decadal prediction skill for the Gulf of Maine, a semi-enclosed
1005 body of water on the Northeast U.S. Continental Shelf. While the patterns of
1006 prediction skill that emerged from the interplay of climate change signals and
1007 variability were unique to each quantity and the Gulf of Maine, other coastal
1008 regions are subject to similar oceanic and atmospheric forcing (Shearman and
1009 Lentz, 2010) and may present opportunities for bottom-habitat prediction. The
1010 ubiquity of sea surface temperature prediction skill in coastal systems at seasonal
1011 and decadal scales provides grounds for optimism (Stock et al., 2015; Tommasi et

1012 al., 2017b). The methodologies employed herein, while computationally intensive,
1013 provide a robust means to diagnose the origins of predictive skill and attach
1014 probabilities critical to making forecast information actionable for management
1015 applications.

1016 More frequent and extensive observations of the ocean, including deep ocean
1017 biogeochemistry, are essential for increasing confidence in the accuracy of the
1018 forecasts and providing confirmation of predicted changes (Capotondi et al. 2019).

1019 Observations are especially important for verifying interannual to decadal scale
1020 ocean predictions, where long time series are necessary to verify that the

1021 predictions have more skill than simple forecasts of the forced signal or
1022 autocorrelated noise. Here, given the scarcity of past deep alongshelf

1023 biogeochemical observations, we utilized a numerical model as the reference for
1024 forecast verification. This downscaled simulation was validated using the best
1025 available observations (Fig 2, Fig. S1, Fig. S2, Fig. S3, Fig. S6 and Fig. S7).

1026 Consequently, the reported skill represents potential skill and may be an
1027 overestimation. However, the model successfully simulated observed variability in

1028 bottom temperature and salinity and captured the expected covariance between
1029 these two variables and the bottom biogeochemistry. Koul et al. (2024) also

1030 demonstrated forecast skill for surface temperature when compared against actual
1031 observations. Furthermore, the model is likely accurate in predicting much of the

1032 forced component of skill, which is driven by steadily increasing greenhouse gas
1033 concentrations and resulting impacts on ocean warming and chemistry.

1034 Our forecasts of continued cooling alongside below-normal bottom pH, Ω_{ar} and
1035 above-normal bottom O₂ levels through 2027 reveal the substantial biogeochemical
1036 fingerprint of this water mass transition and potentially offer stakeholders lead time
1037 to anticipate ecosystem shifts. Skillful forecasts on these multi-year timescales can
1038 meet critical needs of fishers, managers, conservationists, and other stakeholders
1039 (Link et al., 2023; Saba et al., 2023). The framework developed here, consisting of
1040 a computationally efficient regional model that downscales regularly updated
1041 global predictions, is well positioned to meet these needs and is intended to be
1042 transitioned towards operational use through NOAA's Changing Ecosystems and
1043 Fisheries Initiative. Recognizing that marine resource management is a wicked
1044 problem (Jentoft and Chuenpagdee 2009; Rittle and Webber, 1973), however,
1045 careful application-specific analyses of the potential costs, benefits, and tradeoffs
1046 of using these forecasts for decision-making will be necessary.

1047 **Acknowledgements**

1048 Computational resources for this work were provided by allocations on NOAA's
1049 Research and Development High-Performance Computing System from the
1050 Geophysical Fluid Dynamics Laboratory and NOAA's Changing Ecosystems and
1051 Fisheries Initiative.

1052 **Funding**

1053 V.K. is funded by the NOAA Climate Program Office. T.L.D., A.C.R., C.S., and
1054 A.W. are supported by NOAA's Geophysical Fluid Dynamics Laboratory as part of
1055 its base activities. L.Z. receives support through UCAR under block funding from
1056 NOAA/GFDL. This report was prepared by Vimal Koul under awards
1057 NA18OAR4320123 and NA23OAR4320198 from the National Oceanic and
1058 Atmospheric Administration, U.S. Department of Commerce. The statements,
1059 findings, conclusions, and recommendations presented in this report are those of
1060 the author(s) and do not necessarily reflect the official views of the National
1061 Oceanic and Atmospheric Administration or the U.S. Department of Commerce.

1062 **Data Availability**

1063 The processed model output and the observations used in this research are publicly
1064 available via zenodo (doi: 10.5281/zenodo.16985350)

1065 **Code Availability**

1066 The regional MOM6 model's source code can be accessed at
1067 <https://github.com/NOAA-GFDL/CEFI-regional-MOM6>. MOM6 operates on an
1068 open development framework, with Git repositories at [https://github.com/mom-](https://github.com/mom-ocean/MOM6)
1069 [ocean/MOM6](https://github.com/mom-ocean/MOM6) and <https://github.com/NOAA-GFDL/MOM6>. Source code for other
1070 model components is also available at <https://github.com/NOAA-GFDL>.

1071 **Author contributions**

1072 V.K., A.C.R., and C.S. conceptualized the research and drafted the initial
1073 manuscript. V.K. was responsible for conducting model simulations, performing
1074 decadal predictions, post-processing model outputs and observations, and
1075 generating the corresponding figures. A.C.R. and C.S. contributed in designing and
1076 discussing relevant prediction skill metrics. TD, AW, and LZ developed the
1077 SPEAR global prediction model. All authors contributed to the analysis and
1078 interpretation of results, as well as the editing of the paper.

1079 **Corresponding author**

1080 Correspondence to Vimal Koul.

1081 **Ethics declarations**

1082 **Competing interests**

1083 At least one of the (co-)authors serves as editor for the special issue to which this
1084 paper belongs.

References

- Abram, N. J., McGregor, H. V., Tierney, J. E., Evans, M. N., McKay, N. P., & Kaufman, D. S. (2016). Early onset of industrial-era warming across the oceans and continents. *Nature*, 536(7617), 411–418.
<https://doi.org/10.1038/nature19082>
- Alexander, M. A., Scott, J. D., Friedland, K. D., Mills, K. E., Nye, J. A., Pershing, A. J., & Thomas, A. C. (2018). Projected sea surface temperatures over the 21st century: Changes in the mean, variability and extremes for large marine ecosystem regions of Northern Oceans. *Elementa: Science of the Anthropocene*, 6, 9.
<https://doi.org/10.1525/elementa.191>
- Breitburg, 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., ... Zhang, J. (2018). Declining oxygen in the global ocean and coastal waters. *Science*, 359(6371), eaam7240. <https://doi.org/10.1126/science.aam7240>

1102 Brickman, D., Hebert, D. and Wang, Z., 2018. Mechanism for the recent
 1103 ocean warming events on the Scotian Shelf of eastern Canada. *Continental*
 1104 *Shelf Research*, 156, pp.11-22.

1105 Caldeira, K., & Wickett, M. E. (2003). Anthropogenic carbon and ocean pH.
 1106 *Nature*, 425(6956), 365–365. <https://doi.org/10.1038/425365a>

1107 Capotondi, A., Jacox, M., Bowler, C., Kavanaugh, M., Lehodey, P., Barrie,
 1108 D., Brodie, S., Chaffron, S., Cheng, W., Dias, D. F., Eveillard, D., Guidi,
 1109 L., Iudicone, D., Lovenduski, N. S., Nye, J. A., Ortiz, I., Pirhalla, D., Pozo
 1110 Buil, M., Saba, V., ... Pesant, S. (2019). Observational Needs Supporting
 1111 Marine Ecosystems Modeling and Forecasting: From the Global Ocean to
 1112 Regional and Coastal Systems. *Frontiers in Marine Science*, 6, 623.
 1113 <https://doi.org/10.3389/fmars.2019.00623>

1114 Carter, B.R., Bittig, H.C., Fassbender, A.J., Sharp, J.D., Takeshita, Y., Xu, Y.-
 1115 Y., Álvarez, M., Wanninkhof, R., Feely, R.A. and Barbero, L. (2021),
 1116 New and updated global empirical seawater property estimation routines.
 1117 *Limnol Oceanogr Methods*, 19: 785-809.
 1118 <https://doi.org/10.1002/lom3.10461>

1119 Chai, F., Johnson, K.S., Claustre, H. *et al.* Monitoring ocean biogeochemistry
 1120 with autonomous platforms. *Nat Rev Earth Environ* 1, 315–326 (2020).
 1121 <https://doi.org/10.1038/s43017-020-0053-y>

1122 Chen, Z., Kwon, Y.-O., Chen, K., Fratantoni, P., Gawarkiewicz, G., & Joyce,
 1123 T. M. (2020). Long-Term SST Variability on the Northwest Atlantic
 1124 Continental Shelf and Slope. *Geophysical Research Letters*, 47(1),
 1125 e2019GL085455. <https://doi.org/10.1029/2019GL085455>

1126 Chen, Z., Kwon, Y.-O., Chen, K., Fratantoni, P., Gawarkiewicz, G., Joyce, T.
 1127 M., Miller, T. J., Nye, J. A., Saba, V. S., & Stock, B. C. (2021). Seasonal
 1128 Prediction of Bottom Temperature on the Northeast U.S. Continental
 1129 Shelf. *Journal of Geophysical Research: Oceans*, 126(5), e2021JC017187.
 1130 <https://doi.org/10.1029/2021JC017187>

1131 Chen, Z., Siedlecki, S., Long, M., Petrik, C. M., Stock, C. A., & Deutsch, C.
 1132 A. (2024). Skillful multiyear prediction of marine habitat shifts jointly
 1133 constrained by ocean temperature and dissolved oxygen. *Nature*
 1134 *Communications*, 15(1), 900. <https://doi.org/10.1038/s41467-024-45016-5>

1135 Claret, M., Galbraith, E. D., Palter, J. B., Bianchi, D., Fennel, K., Gilbert, D.,
 1136 & Dunne, J. P. (2018). Rapid coastal deoxygenation due to ocean

1137 circulation shift in the northwest Atlantic. *Nature Climate Change*, 8(10),
 1138 868–872. <https://doi.org/10.1038/s41558-018-0263-1>

1139 Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A.,
 1140 Lodge, D.M., Pascual, M., Pielke Jr, R., Pizer, W. and Pringle, C., 2001.
 1141 Ecological forecasts: an emerging imperative. *science*, 293(5530), pp.657-
 1142 660.

1143 Delworth, T. L., Cooke, W. F., Adcroft, A., Bushuk, M., Chen, J. H., Dunne,
 1144 K. A., ... & Zhao, M. (2020). SPEAR: The next generation GFDL
 1145 modeling system for seasonal to multidecadal prediction and projection.
 1146 *Journal of Advances in Modeling Earth Systems*, 12(3), e2019MS001895.

1147 Drinkwater, K. F., Mountain, D. B., & Herman, A. (1998). Recent changes in
 1148 the hydrography of the Scotian Shelf and Gulf of Maine-- a return to
 1149 conditions of the 1960's?. *Sci. Coun. Res. Doc. NAFO*, (98), 16.

1150 Du Pontavice, H., Chen, Z., & Saba, V. S. (2023). A high-resolution ocean
 1151 bottom temperature product for the northeast U.S. continental shelf marine
 1152 ecosystem. *Progress in Oceanography*, 210, 102948.
 1153 <https://doi.org/10.1016/j.pocean.2022.102948>

1154 Ekstrom, J. A., Suatoni, L., Cooley, S. R., Pendleton, L. H., Waldbusser, G.
 1155 G., Cinner, J. E., Ritter, J., Langdon, C., Van Hooideonk, R., Gledhill, D.,

1156 Wellman, K., Beck, M. W., Brander, L. M., Rittschof, D., Doherty, C.,
 1157 Edwards, P. E. T., & Portela, R. (2015). Vulnerability and adaptation of
 1158 US shellfisheries to ocean acidification. *Nature Climate Change*, 5(3),
 1159 207–214. <https://doi.org/10.1038/nclimate2508>

1160 Fay, G., Link, J. S., & Hare, J. A. (2017). Assessing the effects of ocean
 1161 acidification in the Northeast US using an end-to-end marine ecosystem
 1162 model. *Ecological Modelling*, 347, 1-10.

1163 Feely, Feely, F., Jiang, L.-Q., Wanninkhof, R., Carter, B., Alin, S., Bednaršek,
 1164 N., & Cosca, C. (2023). Acidification of the Global Surface Ocean: What
 1165 We Have Learned from Observations. *Oceanography*.
 1166 <https://doi.org/10.5670/oceanog.2023.222>

1167 Feng, Hui, Doug Vandemark, and John Wilkin. "Gulf of Maine salinity
 1168 variation and its correlation with upstream Scotian Shelf currents at
 1169 seasonal and interannual time scales." *Journal of Geophysical Research:*
 1170 *Oceans* 121, no. 12 (2016): 8585-8607.

1171 Friedland, K. D., & Hare, J. A. (2007). Long-term trends and regime shifts in
 1172 sea surface temperature on the continental shelf of the northeast United
 1173 States. *Continental Shelf Research*, 27(18), 2313–2328.
 1174 <https://doi.org/10.1016/j.csr.2007.06.001>

1175 Friedland, K. D., Morse, R. E., Manning, J. P., Melrose, D. C., Miles, T.,
 1176 Goode, A. G., Brady, D. C., Kohut, J. T., & Powell, E. N. (2020). Trends
 1177 and change points in surface and bottom thermal environments of the US
 1178 Northeast Continental Shelf Ecosystem. *Fisheries Oceanography*, 29(5),
 1179 396–414. <https://doi.org/10.1111/fog.12485>

1180 Friedlingstein, P., O’Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J.,
 1181 Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Olsen, A., Peters, G.
 1182 P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G.,
 1183 Ciais, P., Jackson, R. B., Alin, S. R., ... Zeng, J. (2025). Global Carbon
 1184 Budget 2024. *Earth System Science Data*, 17(3), 965–1039.
 1185 <https://doi.org/10.5194/essd-17-965-2025>

1186 Frölicher, T. L., Ramseyer, L., Raible, C. C., Rodgers, K. B., & Dunne, J.
 1187 (2020). Potential predictability of marine ecosystem drivers.
 1188 *Biogeosciences*, 17(7), 2061–2083. [https://doi.org/10.5194/bg-17-2061-](https://doi.org/10.5194/bg-17-2061-2020)
 1189 [2020](https://doi.org/10.5194/bg-17-2061-2020)

1190 Gangopadhyay, A., Gawarkiewicz, G., Silva, E. N. S., Monim, M., & Clark, J.
 1191 (2019). An Observed Regime Shift in the Formation of Warm Core Rings
 1192 from the Gulf Stream. *Scientific Reports*, 9(1), 12319.
 1193 <https://doi.org/10.1038/s41598-019-48661-9>

1194 [Garcia, H.E. and Gordon, L.I., 1992. Oxygen solubility in seawater: Better](#)
 1195 [fitting equations. *Limnology and oceanography*, 37\(6\), pp.1307-1312.](#)

1196 Gehlen, M., Barciela, R., Bertino, L., Brasseur, P., Butenschön, M., Chai, F.,
 1197 Crise, A., Drillet, Y., Ford, D., Lavoie, D., Lehodey, P., Perruche, C.,
 1198 Samuelsen, A., & Simon, E. (2015). Building the capacity for forecasting
 1199 marine biogeochemistry and ecosystems: Recent advances and future
 1200 developments. *Journal of Operational Oceanography*, 8(sup1), s168–
 1201 s187. <https://doi.org/10.1080/1755876X.2015.1022350>

1202 Good, S. A., Martin, M. J., & Rayner, N. A. (2013). EN4: Quality controlled
 1203 ocean temperature and salinity profiles and monthly objective analyses
 1204 with uncertainty estimates. *Journal of Geophysical Research: Oceans*,
 1205 118(12), 6704-6716.

1206 Gonçalves Neto, A., Langan, J. A., & Palter, J. B. (2021). Changes in the Gulf
 1207 Stream preceded rapid warming of the Northwest Atlantic Shelf.
 1208 *Communications Earth & Environment*, 2(1), 74.
 1209 <https://doi.org/10.1038/s43247-021-00143-5>

1210 Henson Stephanie A. (2014). Slow science: the value of long ocean
 1211 biogeochemistry records. *Phil. Trans. R. Soc. A*.37220130334
 1212 <http://doi.org/10.1098/rsta.2013.0334>

1213 Hare, J. A., Morrison, W. E., Nelson, M. W., Stachura, M. M., Teeters, E. J.,
 1214 Griffis, R. B., Alexander, M. A., Scott, J. D., Alade, L., Bell, R. J., Chute,
 1215 A. S., Curti, K. L., Curtis, T. H., Kircheis, D., Kocik, J. F., Lucey, S. M.,
 1216 McCandless, C. T., Milke, L. M., Richardson, D. E., ... Griswold, C. A.
 1217 (2016). A Vulnerability Assessment of Fish and Invertebrates to Climate
 1218 Change on the Northeast U.S. Continental Shelf. *PLOS ONE*, 11(2),
 1219 e0146756. <https://doi.org/10.1371/journal.pone.0146756>
 1220 Harrington, A. M., & Hamlin, H. J. (2019). Ocean acidification alters thermal
 1221 cardiac performance, hemocyte abundance, and hemolymph chemistry in
 1222 subadult American lobsters *Homarus americanus* H. Milne Edwards, 1837
 1223 (Decapoda: Malacostraca: Nephropidae). *Journal of Crustacean Biology*,
 1224 39(4), 468–476. <https://doi.org/10.1093/jcbiol/ruz015>
 1225 Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore,
 1226 J. H., ... & Zhang, H. M. (2017). Extended reconstructed sea surface
 1227 temperature, version 5 (ERSSTv5): upgrades, validations, and
 1228 intercomparisons. *Journal of Climate*, 30(20), 8179-8205.
 1229 [Humphreys, M.P., Lewis, E.R., Sharp, J.D. and Pierrot, D., 2022. PyCO2SYS](#)
 1230 [v1. 8: Marine carbonate system calculations in Python. *Geoscientific*](#)
 1231 [Model Development, 15\(1\), pp.15-43.](#)

1232 Jentoft, S., & Chuenpagdee, R. (2009). Fisheries and coastal governance as a
 1233 wicked problem. *Marine Policy*, 33(4), 553–560.
 1234 <https://doi.org/10.1016/j.marpol.2008.12.002>

1235 Jiang, L.-Q., Feely, R. A., Wanninkhof, R., Greeley, D., Barbero, L., Alin, S.,
 1236 Carter, B. R., Pierrot, D., Featherstone, C., Hooper, J., Melrose, C.,
 1237 Monacci, N., Sharp, J. D., Shellito, S., Xu, Y.-Y., Kozyr, A., Byrne, R. H.,
 1238 Cai, W.-J., Cross, J., Johnson, G. C., Hales, B., Langdon, C., Mathis, J.,
 1239 Salisbury, J., and Townsend, D. W.: Coastal Ocean Data Analysis Product
 1240 in North America (CODAP-NA) – an internally consistent data product
 1241 for discrete inorganic carbon, oxygen, and nutrients on the North
 1242 American ocean margins, *Earth Syst. Sci. Data*, 13, 2777–2799,
 1243 <https://doi.org/10.5194/essd-13-2777-2021>, 2021.

1244 Johnson, G. C., & Lyman, J. M. (2020). Warming trends increasingly
 1245 dominate global ocean. *Nature Climate Change*, 10(8), 757-761.

1246 Kavanaugh, M. T., Rheuban, J. E., Luis, K. M. A., & Doney, S. C. (2017).
 1247 Thirty-Three Years of Ocean Benthic Warming Along the U.S. Northeast
 1248 Continental Shelf and Slope: Patterns, Drivers, and Ecological
 1249 Consequences. *Journal of Geophysical Research: Oceans*, 122(12), 9399–
 1250 9414. <https://doi.org/10.1002/2017JC012953>

1251 Keeling, R. F., Körtzinger, A., & Gruber, N. (2010). Ocean Deoxygenation in
 1252 a Warming World. *Annual Review of Marine Science*, 2(1), 199–229.
 1253 <https://doi.org/10.1146/annurev.marine.010908.163855>

1254 Koelling, J., Atamanchuk, D., Wallace, D. W., & Karstensen, J. (2023).
 1255 Decadal variability of oxygen uptake, export, and storage in the Labrador
 1256 Sea from observations and CMIP6 models. *Frontiers in Marine Science*,
 1257 10, 1202299.

1258 Kosaka, Y., Kobayashi, S., Harada, Y., Kobayashi, C., Naoe, H., Yoshimoto,
 1259 K., ... & Onogi, K. (2024). The JRA-3Q reanalysis. *Journal of the*
 1260 *Meteorological Society of Japan. Ser. II*, 102(1), 49-109.

1261 Koul, V., Ross, A. C., Stock, C., Zhang, L., Delworth, T., & Wittenberg, A.
 1262 (2024). A Predicted Pause in the Rapid Warming of the Northwest
 1263 Atlantic Shelf in the Coming Decade. *Geophysical Research Letters*,
 1264 51(17), e2024GL110946. <https://doi.org/10.1029/2024GL110946>

1265 Krumhardt, K. M., Lovenduski, N. S., Long, M. C., Luo, J. Y., Lindsay, K.,
 1266 Yeager, S., & Harrison, C. (2020). Potential Predictability of Net Primary
 1267 Production in the Ocean. *Global Biogeochemical Cycles*, 34(6),
 1268 e2020GB006531. <https://doi.org/10.1029/2020GB006531>

1269 Lehodey, P., Alheit, J., Barange, M., Baumgartner, T., Beaugrand, G.,
 1270 Drinkwater, K., ... & Werner, F. (2006). Climate variability, fish, and
 1271 fisheries. *Journal of Climate*, 19(20), 5009-5030.

1272 Levitus, S., Antonov, J. I., Boyer, T. P., Baranova, O. K., Garcia, H. E.,
 1273 Locarnini, R. A., Mishonov, A. V., Reagan, J. R., Seidov, D., Yarosh, E.
 1274 S., & Zweng, M. M. (2012). World ocean heat content and thermosteric
 1275 sea level change (0–2000 m), 1955–2010. *Geophysical Research Letters*,
 1276 39(10). <https://doi.org/10.1029/2012GL051106>

1277 Link, J. S., Thur, S., Matlock, G., & Grasso, M. (2023). Why we need weather
 1278 forecast analogues for marine ecosystems. *ICES Journal of Marine*
 1279 *Science*, 80(8), 2087–2098. <https://doi.org/10.1093/icesjms/fsad143>

1280 Msadek, R., Delworth, T.L., Rosati, A., Anderson, W., Vecchi, G., Chang,
 1281 Y.S., Dixon, K., Gudgel, R.G., Stern, W., Wittenberg, A. and Yang, X.,
 1282 2014. Predicting a decadal shift in North Atlantic climate variability using
 1283 the GFDL forecast system. *Journal of Climate*, 27(17), pp.6472-6496.

1284 Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu,
 1285 G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E., Hegerl, G.,
 1286 Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A.,
 1287 Pulwarty, R., Smith, D., Stammer, D., & Stockdale, T. (2009). Decadal

1288 Prediction: Can It Be Skillful? *Bulletin of the American Meteorological*
 1289 *Society*, 90(10), 1467–1486. <https://doi.org/10.1175/2009BAMS2778.1>

1290 Meyer-Gutbrod, E. L., Greene, C. H., Davies, K. T. A., & Johns, D. G. (2021).
 1291 Ocean Regime Shift is Driving Collapse of the North Atlantic Right
 1292 Whale Population. *Oceanography*, 34(3), 22–31.
 1293 <https://doi.org/10.5670/oceanog.2021.308>

1294 Minka, T. P. (2001). Algorithms for maximum-likelihood logistic regression
 1295 (Tech Report No. 758). Carnegie Mellon University, Pittsburgh, PA, USA.

1296 Mountain, D. G. (2004). Variability of the Water Properties in NAFO
 1297 Subareas 5 and 6 During the 1990s. *Journal of Northwest Atlantic Fishery*
 1298 *Science*, 34, 103–112. <https://doi.org/10.2960/J.v34.m475>

1299 Mountain, D. G. (2012). Labrador slope water entering the Gulf of Maine—
 1300 Response to the North Atlantic Oscillation. *Continental Shelf Research*,
 1301 47, 150–155. <https://doi.org/10.1016/j.csr.2012.07.008>

1302 Niemisto, M., Contreras, A., Wahle, R. A., & Fields, D. M. (2025). Increased
 1303 temperature and acidification elevate the risk of starvation in American
 1304 lobster larvae. *ICES Journal of Marine Science*, 82(6), fsaf088.
 1305 <https://doi.org/10.1093/icesjms/fsaf088>

1306 NOAA Fisheries: 2025 state of the ecosystem: New England. Technical report
 1307 (2025)

1308 Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., Feely, R. A.,
 1309 Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., Key, R. M., Lindsay,
 1310 K., Maier-Reimer, E., Matear, R., Monfray, P., Mouchet, A., Najjar, R. G.,
 1311 Plattner, G.-K., Rodgers, K. B., ... Yool, A. (2005). Anthropogenic ocean
 1312 acidification over the twenty-first century and its impact on calcifying
 1313 organisms. *Nature*, 437(7059), 681–686.
 1314 <https://doi.org/10.1038/nature04095>

1315 Park, J.-Y., Stock, C. A., Dunne, J. P., Yang, X., & Rosati, A. (2019).
 1316 Seasonal to multiannual marine ecosystem prediction with a global Earth
 1317 system model. *Science*, 365(6450), 284–288.
 1318 <https://doi.org/10.1126/science.aav6634>

1319 Payne, M. R., Hobday, A. J., MacKenzie, B. R., Tommasi, D., Dempsey, D.
 1320 P., Fässler, S. M. M., Haynie, A. C., Ji, R., Liu, G., Lynch, P. D., Matei,
 1321 D., Miesner, A. K., Mills, K. E., Strand, K. O., & Villarino, E. (2017).
 1322 Lessons from the First Generation of Marine Ecological Forecast
 1323 Products. *Frontiers in Marine Science*, 4.
 1324 <https://doi.org/10.3389/fmars.2017.00289>

1325 Payne, M. R., Danabasoglu, G., Keenlyside, N., Matei, D., Miesner, A. K.,
 1326 Yang, S., & Yeager, S. G. (2022). Skilful decadal-scale prediction of fish
 1327 habitat and distribution shifts. *Nature Communications*, 13(1), 2660.
 1328 <https://doi.org/10.1038/s41467-022-30280-0>

1329 Pérez-Hernández, M.D. and Joyce, T.M., 2014. Two modes of Gulf Stream
 1330 variability revealed in the last two decades of satellite altimeter data.
 1331 *Journal of Physical Oceanography*, 44(1), pp.149-163.

1332 Pershing, A. J., Alexander, M. A., Brady, D. C., Brickman, D., Curchitser, E.
 1333 N., Diamond, A. W., McClenachan, L., Mills, K. E., Nichols, O. C.,
 1334 Pendleton, D. E., Record, N. R., Scott, J. D., Staudinger, M. D., & Wang,
 1335 Y. (2021). Climate impacts on the Gulf of Maine ecosystem: A review of
 1336 observed and expected changes in 2050 from rising temperatures.
 1337 *Elementa: Science of the Anthropocene*, 9(1), 00076.
 1338 <https://doi.org/10.1525/elementa.2020.00076>

1339 Pershing, A. J., Alexander, M. A., Hernandez, C. M., Kerr, L. A., Le Bris, A.,
 1340 Mills, K. E., Nye, J. A., Record, N. R., Scannell, H. A., Scott, J. D.,
 1341 Sherwood, G. D., & Thomas, A. C. (2015). Slow adaptation in the face of
 1342 rapid warming leads to collapse of the Gulf of Maine cod fishery. *Science*,
 1343 350(6262), 809–812. <https://doi.org/10.1126/science.aac9819>

1344 Pregibon, D. (1981). Logistic Regression Diagnostics. *The Annals of*
 1345 *Statistics*, 9(4), 705–724. doi: 10.1214/aos/1176345513

1346 Ramos, M. H., Van Andel, S. J., & Pappenberger, F. (2013). Do probabilistic
 1347 forecasts lead to better decisions? *Hydrology and Earth System Sciences*,
 1348 17(6), 2219–2232. <https://doi.org/10.5194/hess-17-2219-2013>

1349 Record, N., Pershing, A., & Rasher, D. (2024). Early Warning of a Cold Wave
 1350 in the Gulf of Maine. *Oceanography*, 37(3).
 1351 <https://doi.org/10.5670/oceanog.2024.506>

1352 [Riahi, K., Van Vuuren, D.P., Kriegler, E., Edmonds, J., O’neill, B.C.,](#)
 1353 [Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O. and Lutz, W.,](#)
 1354 [2017. The Shared Socioeconomic Pathways and their energy, land use,](#)
 1355 [and greenhouse gas emissions implications: An overview. *Global*](#)
 1356 [environmental change, 42, pp.153-168.](#)

1357 Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of
 1358 planning. *Policy Sciences*, 4(2), 155–169.
 1359 <https://doi.org/10.1007/BF01405730>

1360 Ross, A. C., Stock, C. A., Adcroft, A., Curchitser, E., Hallberg, R., Harrison,
 1361 M. J., Hedstrom, K., Zadeh, N., Alexander, M., Chen, W., Drenkard, E. J.,
 1362 Du Pontavice, H., Dussin, R., Gomez, F., John, J. G., Kang, D., Lavoie,

1363 D., Resplandy, L., Roobaert, A., ... Simkins, J. (2023). A high-resolution
 1364 physical–biogeochemical model for marine resource applications in the
 1365 northwest Atlantic (MOM6-COBALT-NWA12 v1.0). *Geoscientific Model*
 1366 *Development*, 16(23), 6943–6985. [https://doi.org/10.5194/gmd-16-6943-](https://doi.org/10.5194/gmd-16-6943-2023)
 1367 [2023](https://doi.org/10.5194/gmd-16-6943-2023)

1368 Ross, A. C., Stock, C. A., Koul, V., Delworth, T. L., Lu, F., Wittenberg, A., &
 1369 Alexander, M. A. (2024). Dynamically downscaled seasonal ocean
 1370 forecasts for North American east coast ecosystems. *Ocean Science*, 20(6),
 1371 1631–1656. <https://doi.org/10.5194/os-20-1631-2024>

1372 Roulston, M. S., Bolton, G. E., Kleit, A. N., & Sears-Collins, A. L. (2006). A
 1373 Laboratory Study of the Benefits of Including Uncertainty Information in
 1374 Weather Forecasts. *Weather and Forecasting*, 21(1), 116–122.
 1375 <https://doi.org/10.1175/WAF887.1>

1376 Saba, V., Borggaard, D., Caracappa, J. C., Chambers, R. C., Clay, P. M.,
 1377 Colburn, L. L., Deroba, J., DePiper, G., Pontavice, H. du, Fratantoni, P.,
 1378 Ferguson, M., Gaichas, S., Hayes, S., Hyde, K., Johnson, M., Kocik, J.,
 1379 Keane, E., Kircheis, D., Large, S., ... Wikfors, G. (2023). NOAA fisheries
 1380 research geared towards climate-ready living marine resource

1381 management in the northeast United States. *PLOS Climate*, 2(12),
 1382 e0000323. <https://doi.org/10.1371/journal.pclm.0000323>

1383 Saba, V. S., Griffies, S. M., Anderson, W. G., Winton, M., Alexander, M. A.,
 1384 Delworth, T. L., Hare, J. A., Harrison, M. J., Rosati, A., Vecchi, G. A., &
 1385 Zhang, R. (2016). Enhanced warming of the Northwest Atlantic Ocean
 1386 under climate change. *Journal of Geophysical Research: Oceans*, 121(1),
 1387 118–132. <https://doi.org/10.1002/2015JC011346>

1388 Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L.,
 1389 Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F.
 1390 J., Peng, T.-H., Kozyr, A., Ono, T., & Rios, A. F. (2004). The Oceanic
 1391 Sink for Anthropogenic CO₂. *Science*, 305(5682), 367–371.
 1392 <https://doi.org/10.1126/science.1097403>

1393 Salisbury, J. E., & Jönsson, B. F. (2018). Rapid warming and salinity changes
 1394 in the Gulf of Maine alter surface ocean carbonate parameters and hide
 1395 ocean acidification. *Biogeochemistry*, 141(3), 401–418.
 1396 <https://doi.org/10.1007/s10533-018-0505-3>

1397 Scully, M. E., Geyer, W. R., Borkman, D., Pugh, T. L., Costa, A., & Nichols,
 1398 O. C. (2022). Unprecedented summer hypoxia in southern Cape Cod Bay:

1399 an ecological response to regional climate change?. *Biogeosciences*
1400 *Discussions*, 2022, 1-19.

1401 Seidov, D., Mishonov, A., & Parsons, R. (2021). Recent warming and decadal
1402 variability of Gulf of Maine and Slope Water. *Limnology and*
1403 *Oceanography*, 66(9), 3472–3488. <https://doi.org/10.1002/lno.11892>

1404 Shearman, R. K., & Lentz, S. J. (2010). Long-Term Sea Surface Temperature
1405 Variability along the U.S. East Coast. *Journal of Physical Oceanography*,
1406 40(5), 1004–1017. <https://doi.org/10.1175/2009JPO4300.1>

1407 Siedlecki, S., Salisbury, J., Gledhill, D., Bastidas, C., Meseck, S., McGarry,
1408 K., Hunt, C., Alexander, M., Lavoie, D., Wang, Z., Scott, J., Brady, D.,
1409 Mlsna, I., Azetsu-Scott, K., Liberti, C., Melrose, D., White, M., Pershing,
1410 A., Vandemark, D., ... Morrison, R. (2021). Projecting ocean acidification
1411 impacts for the Gulf of Maine to 2050: New tools and expectations.
1412 *Elementa: Science of the Anthropocene*, 9(1), 00062.
1413 <https://doi.org/10.1525/elementa.2020.00062>

1414 [Smith, P. C., Houghton, R. W., Fairbanks, R. G., and Mountain, D. G. \(2001\).](#)
1415 [Interannual variability of boundary fluxes and water mass properties in the](#)
1416 [Gulf of Maine and on Georges Bank: 1993–1997. *Deep Sea Research Part*](#)
1417 [II: Topical Studies in Oceanography](#), 48(1), 37–70.632

1418 Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix.
 1419 *Psychological bulletin*, 87(2), 245.

1420 Stewart, J. A., Williams, B., LaVigne, M., Wanamaker, A. D., Strong, A. L.,
 1421 Jellison, B., Whitney, N. M., Thatcher, D. L., Robinson, L. F., Halfar, J.,
 1422 & Adey, W. (2025). Delayed onset of ocean acidification in the Gulf of
 1423 Maine. *Scientific Reports*, 15(1), 2039. [https://doi.org/10.1038/s41598-](https://doi.org/10.1038/s41598-024-84537-3)
 1424 [024-84537-3](https://doi.org/10.1038/s41598-024-84537-3)

1425 Stock, C. A., K. Pegion, G. A. Vecchi, M. A. Alexander, D. Tommasi, N. A.
 1426 Bond, P. S. Fratantoni, R. G. Gudgel, T. Kristiansen, T. D. O'Brien, Y.
 1427 Xue, and X. Yang, 2015: Seasonal sea surface temperature anomaly
 1428 prediction for coastal ecosystems. *Prog. Oceanogr.*, **137**, 219-236, doi:
 1429 [10.1016/j.pocean.2015.06.007](https://doi.org/10.1016/j.pocean.2015.06.007).

1430 Stock, C. A., Dunne, J. P., Fan, S., Ginoux, P., John, J., Krasting, J. P., ... &
 1431 Zadeh, N. (2020). Ocean biogeochemistry in GFDL's Earth System Model
 1432 4.1 and its response to increasing atmospheric CO₂. *Journal of Advances*
 1433 *in Modeling Earth Systems*, 12(10), e2019MS002043.

1434 Stock, C. A., Dunne, J. P., Luo, J. Y., Ross, A. C., Van Oostende, N., Zadeh,
 1435 N., ... & Teng, Y. C. (2025). Photoacclimation and photoadaptation

1436 sensitivity in a global ocean ecosystem model. *Journal of Advances in*
 1437 *Modeling Earth Systems*, 17(6), e2024MS004701.

1438 Tommasi, D., Stock, C. A., Hobday, A. J., Methot, R., Kaplan, I. C., Eveson,
 1439 J. P., Holsman, K., Miller, T. J., Gaichas, S., Gehlen, M., Pershing, A.,
 1440 Vecchi, G. A., Msadek, R., Delworth, T., Eakin, C. M., Haltuch, M. A.,
 1441 Séférian, R., Spillman, C. M., Hartog, J. R., ... Werner, F. E. (2017).
 1442 Managing living marine resources in a dynamic environment: The role of
 1443 seasonal to decadal climate forecasts. *Progress in Oceanography*, 152,
 1444 15–49. <https://doi.org/10.1016/j.pocean.2016.12.011>

1445 Tommasi, D., Stock, C. A., Alexander, M. A., Yang, X., Rosati, A., & Vecchi,
 1446 G. A. (2017). Multi-annual climate predictions for fisheries: an assessment
 1447 of skill of sea surface temperature forecasts for large marine ecosystems.
 1448 *Frontiers in Marine Science*, 4, 201.

1449 [Townsend, D. W. \(1998\). Sources and cycling of nitrogen in the Gulf of](#)
 1450 [Maine. *Journal of Marine Systems*, 16\(3\), 283–660.](#)

1451 Townsend, D. W., Pettigrew, N. R., Thomas, M. A., & Moore, S. (2023).
 1452 Warming waters of the Gulf of Maine: The role of Shelf, Slope and Gulf
 1453 Stream Water masses. *Progress in Oceanography*, 215, 103030.
 1454 <https://doi.org/10.1016/j.pocean.2023.103030>

1455 Townsend, D.W., Thomas, A.C., Mayer, L.M., Thomas, M.A. and Quinlan,
 1456 J.A., 2006. Oceanography of the northwest Atlantic continental shelf (1,
 1457 W). *The sea, 14*, pp.119-168.

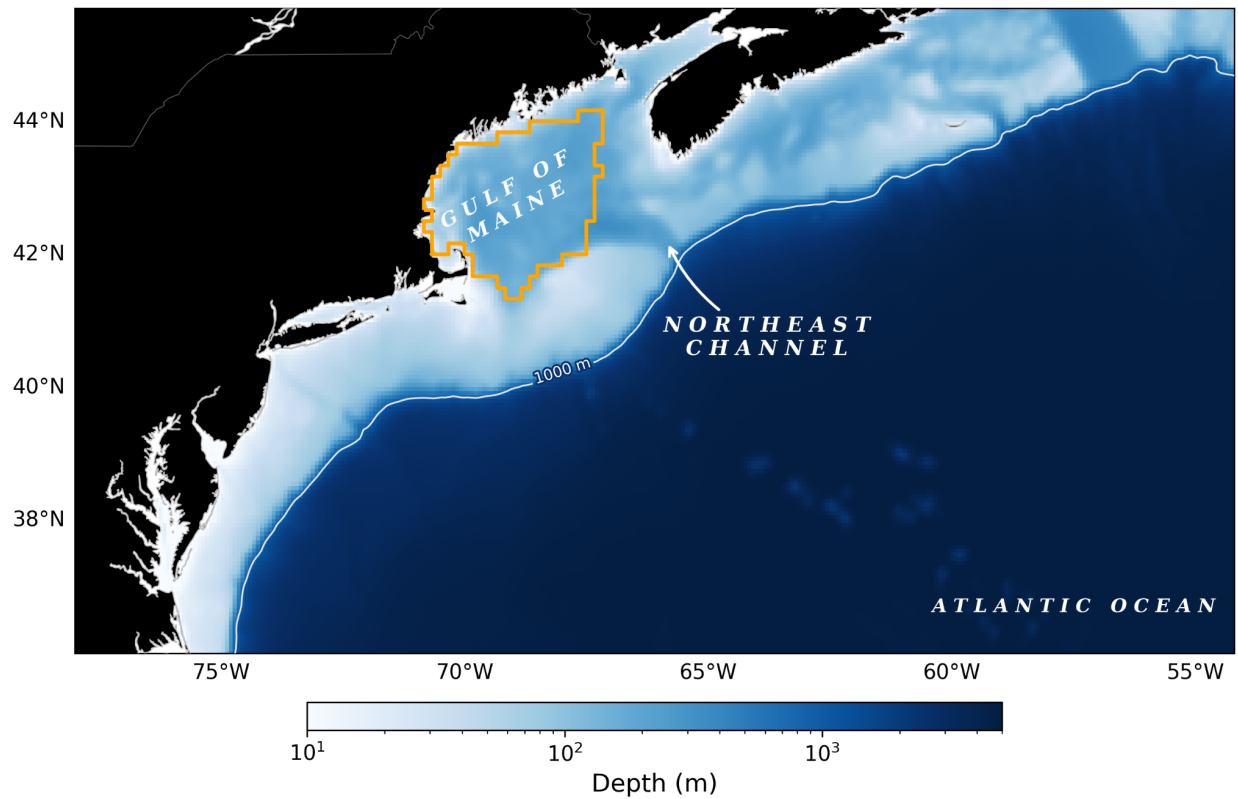
1458 Wallace, E. J., Looney, L. B., & Gong, D. (2018). Multi-Decadal Trends and
 1459 Variability in Temperature and Salinity in the Mid-Atlantic Bight,
 1460 Georges Bank, and Gulf of Maine. *Journal of Marine Research*, 76(5),
 1461 163–215. <https://doi.org/10.1357/002224018826473281>

1462 [Waller, J., Bartlett, J., Bates, E., Bray, H., Brown, M., Cieri, M., Clark, C.,](#)
 1463 [DeVoe, W., Donahue, B., Frechette, D. and Glon, H., 2023. Reflecting on](#)
 1464 [the recent history of coastal Maine fisheries and marine resource](#)
 1465 [monitoring: the value of collaborative research, changing ecosystems, and](#)
 1466 [thoughts on preparing for the future. *ICES Journal of Marine Science*,](#)
 1467 [80\(8\), pp.2074-2086.](#)

1468 Wilks, D. S. (2009). Extending logistic regression to provide full-probability-
 1469 distribution MOS forecasts. *Meteorological Applications*, 16(3), 361–368.
 1470 <https://doi.org/10.1002/met.134>

1471 [Xue, L. and Cai, W.J., 2020. Total alkalinity minus dissolved inorganic carbon](#)
 1472 [as a proxy for deciphering ocean acidification mechanisms. *Marine*](#)
 1473 [Chemistry](#), 222, p.103791.

1474 Yang, X., Delworth, T. L., Zeng, F., Zhang, L., Cooke, W. F., Harrison, M. J.,
1475 ... & McColl, C. (2021). On the development of GFDL's decadal
1476 prediction system: Initialization approaches and retrospective forecast
1477 assessment. *Journal of Advances in Modeling Earth Systems*, 13(11),
1478 e2021MS002529.



1479

1480 **Fig 1: The study region and its characteristics.** The locations of the Gulf of Maine Ecological
 1481 Production Unit (GOM EPU) and the Northeast Channel along the Northeast US continental shelf are
 1482 shown along with the model bathymetry. The names and approximate locations of the Warm Slope Water
 1483 and Labrador Slope Water are also shown. The 1000-meter isobaths is also shown as a white line.

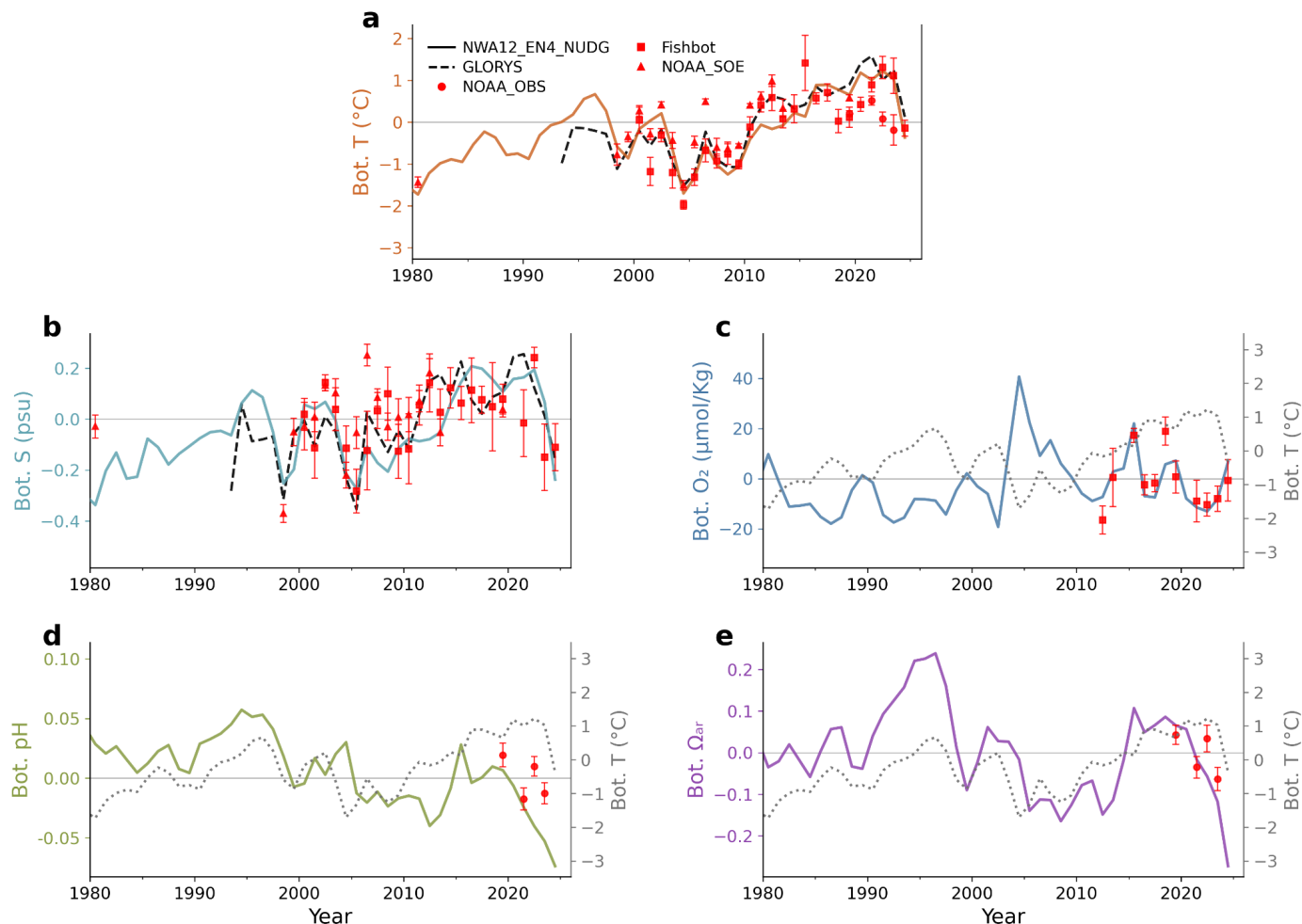


Fig 2: Physical and biogeochemical variability in the GOM. Annual mean anomalies of (a) bottom temperature ($^{\circ}\text{C}$) (b) bottom salinity (psu) (c) bottom oxygen ($\mu\text{mol/kg}$), (d) bottom pH and (e) bottom aragonite saturation state in the GOM from the ensemble mean of the regional downscaled 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 Methods). 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).

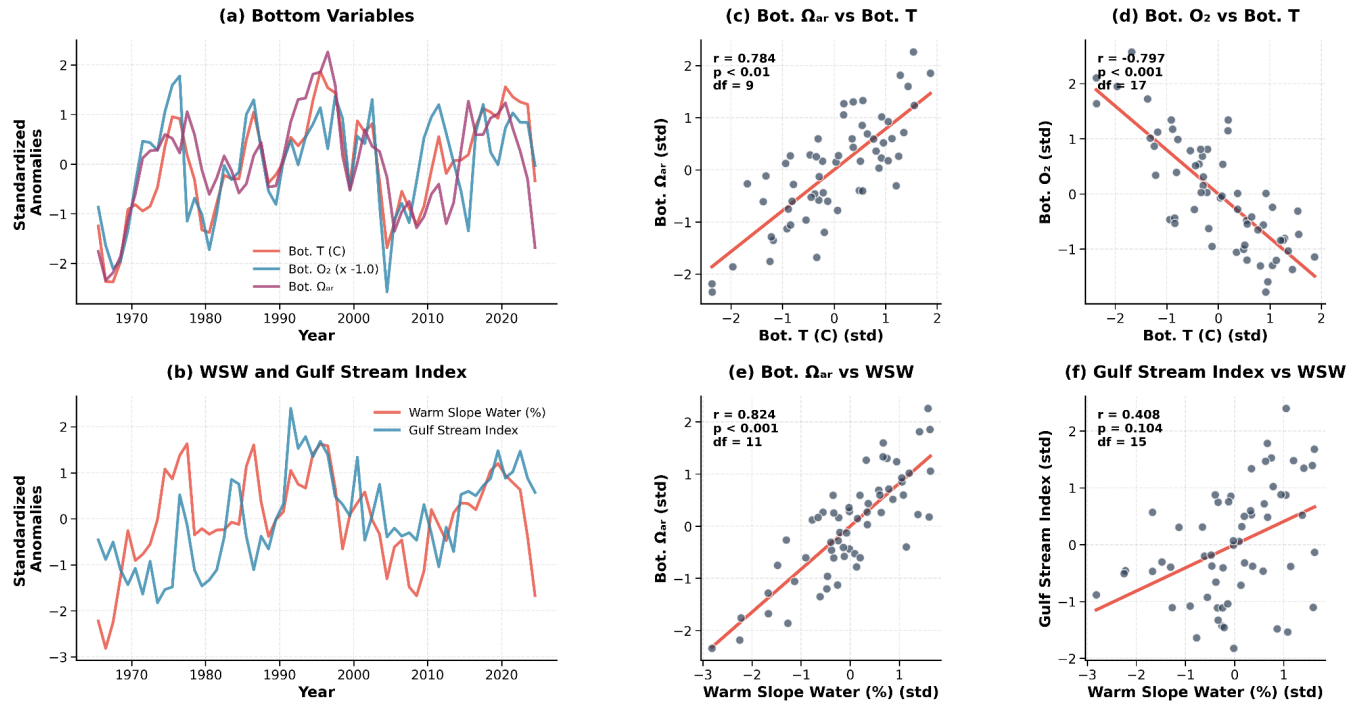


Fig. 3: Drivers of internal co-variability in the GOM. Standardized anomalies (with respect to 1994-2023 period) of the internal (natural) component of the annual mean (a) bottom temperature, bottom O_2 and bottom Ω_{ar} in the GOM (b) percentage of warm slope water (WSW) in the Northeast Channel and 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.

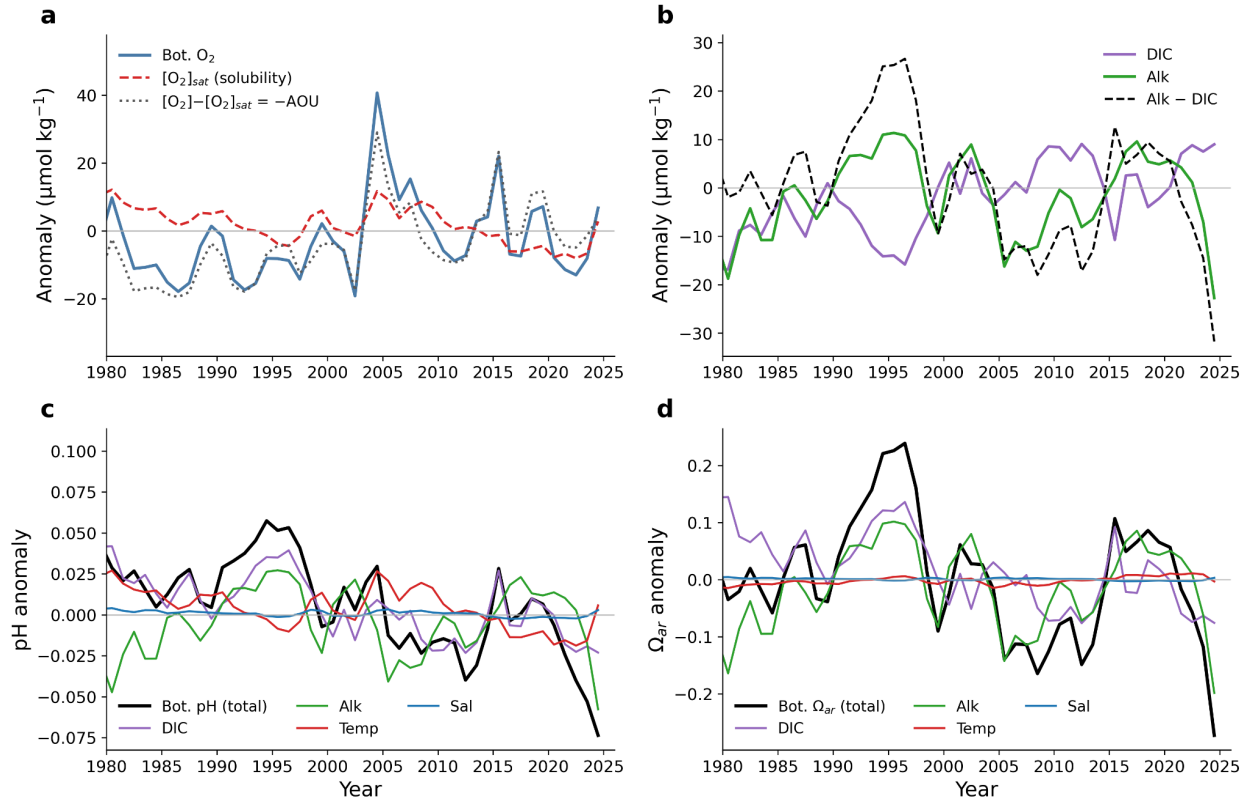


Figure 4: Decomposition of bottom-water biogeochemical variability in the Gulf of Maine. (a) Bottom- O_2 anomaly partitioned into the solubility term $[O_2]_{sat}$ and the non-solubility residual $[O_2] - [O_2]_{sat} (\approx -AOU)$ in the NWA12_EN4_NUDG retrospective simulation (ensemble mean anomalies 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 Methods).

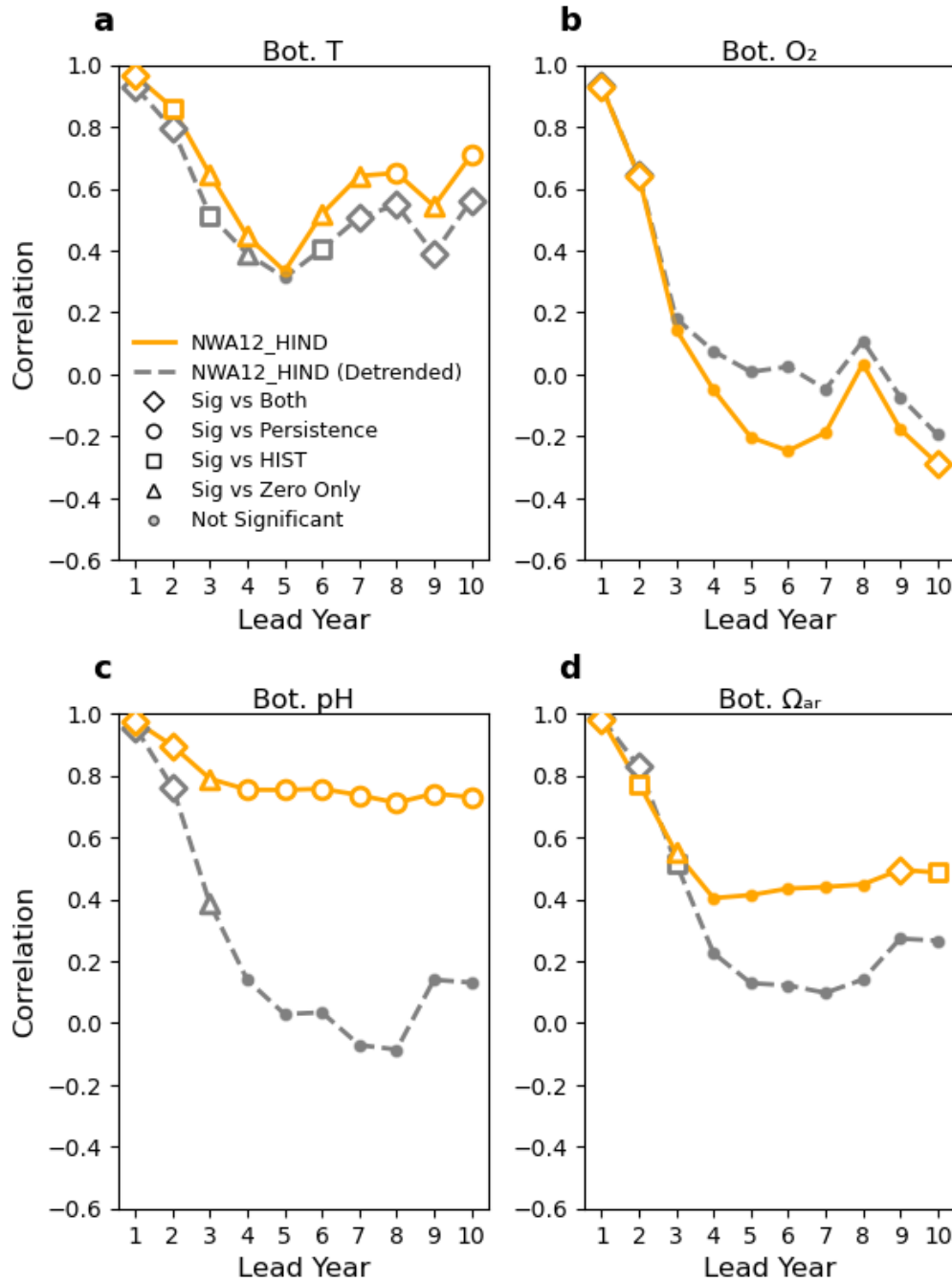


Fig. 5: Decadal prediction skill. Anomaly correlation coefficient of the annual mean anomalies from the initialized regional downscaled model (NWA12_HIND) predictions. The symbols indicate statistically significant differences relative to alternative models of no correlation, persistence, and the uninitialized (historical) simulation (see Methods). 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.

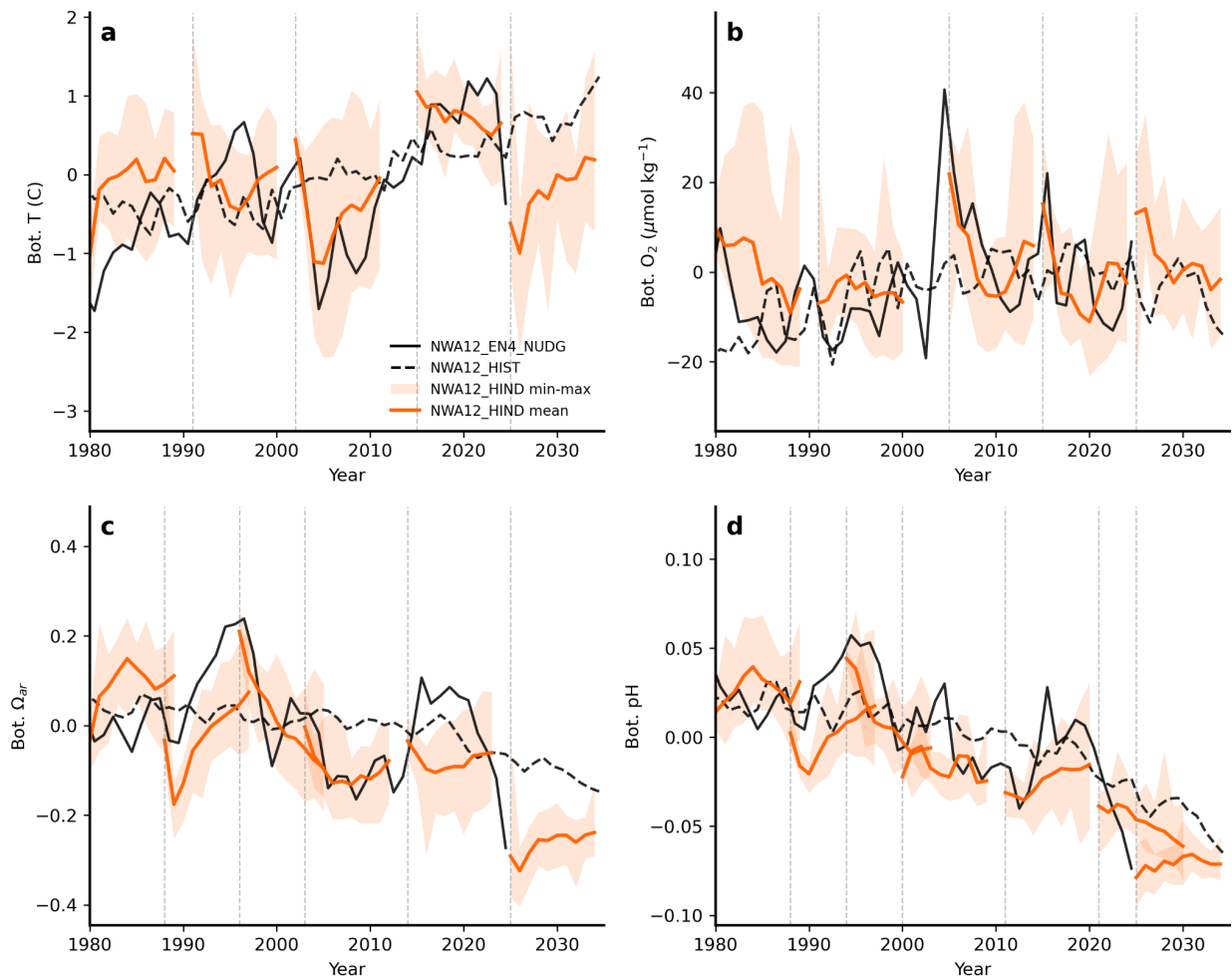


Fig 1: Physical and biogeochemical variability in the GOM. The locations of the Gulf of Maine Ecological Production Unit (GOM_EPU) and the Northeast Channel along the Northeast US continental shelf are shown in (a) along with the model bathymetry, Annual mean anomalies of (b) bottom temperature ($^{\circ}\text{C}$) (c) bottom salinity (psu) (d) bottom oxygen ($\mu\text{mol/kg}$), (e) bottom pH and (f) bottom aragonite saturation state in the GOM from the ensemble mean of the regional downscaled simulations (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 standard error of annual anomalies. The dashed black lines in (b) and (e) are from the GLORYS12 dataset. All anomalies are calculated with respect to the 1994-2023 mean. To highlight physical and biogeochemical covariability, normalized bottom temperature lines are shown as orange lines in (d) (f) and inverted in (d).

Fig. 6: Physical and biogeochemical predictions in the GOM. Annual mean anomalies of (a) bottom temperature ($^{\circ}\text{C}$) (b) bottom oxygen ($\mu\text{mol/kg}$), (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.

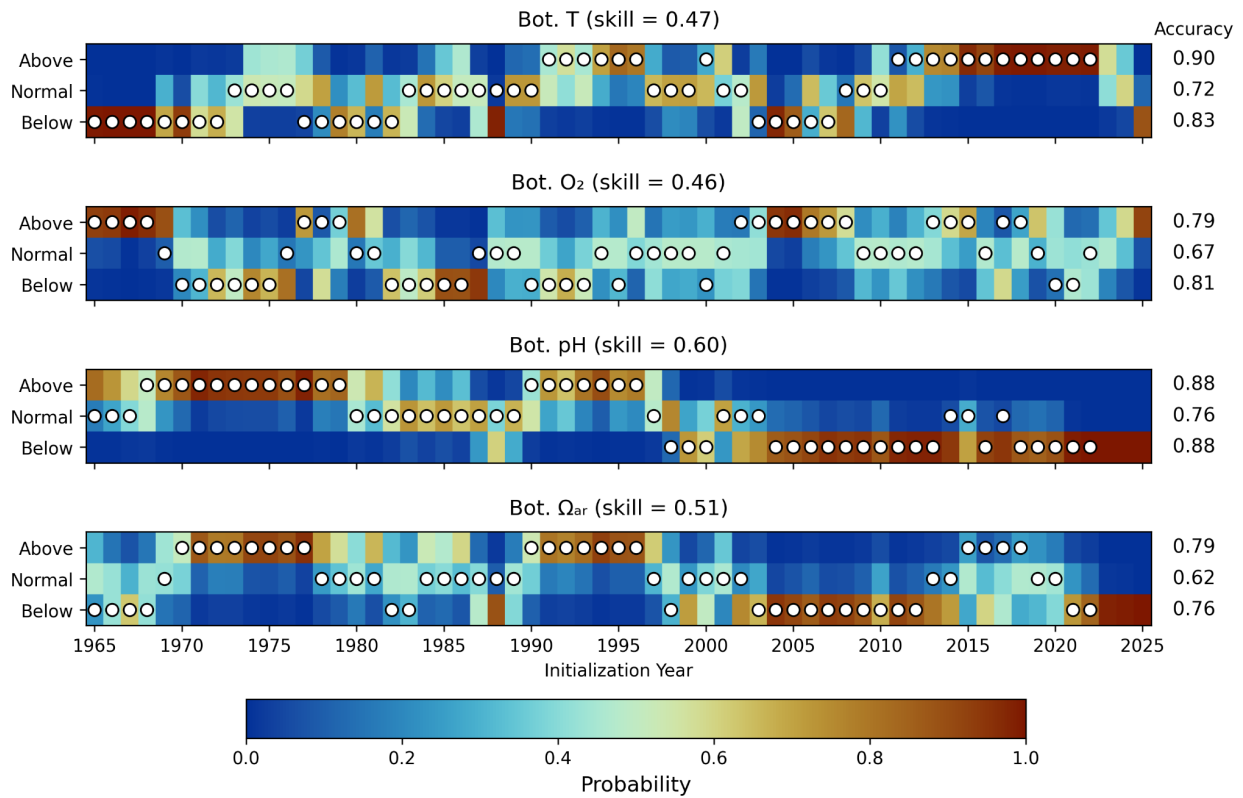


Fig. 2: Drivers of internal co-variability in the GOM. Standardized anomalies (with respect to 1994–2023 period) of the internal (natural) component of the annual mean (a) bottom temperature, bottom O_2 and bottom Ω_{ar} in the GOM (b) percentage of warm slope water (WSW) in the Northeast Channel and 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.

Fig. 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 O_2 (c) bottom pH and (d) bottom Ω_{ar} from the regional downscaled decadal prediction model (NWA12_HIND). The

1568 corresponding classifications of the reference anomalies are shown as white circles from the regional
1569 downscaled simulation (NWA12_EN4_NUDG). The tercile boundaries are calculated separately for the
1570 predictions and reference/observations. The x-axis shows the initialization year and the 2025 forecast.
1571 The forecast accuracy, (True Positives+True Negatives)/Total forecasts, for each tercile is shown on the
1572 right. A random forecast that randomly picks one tercile as most likely for each forecast has an accuracy
1573 of 0.56 and the accuracy of a consistently negative (counter) prediction is 0.67. The forecast skill
1574 compared to persistence is given in the subtitle of each panel.

1575

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

1576

1577

1578

1579

1580

1581

1582

1583

1584

Fig. 3: Decadal prediction skill. Anomaly correlation coefficient of the annual mean anomalies from the initialized regional downscaled model (NWA12_HIND) predictions. The symbols indicate statistical significant differences relative to alternative models of no correlation, persistence, and the uninitialized (historical) simulation (see Methods). The correlation is calculated against the regional downscaled simulation (NWA12_EN4_NUDG) derived annual mean anomalies (ensemble mean) for (a) bottom temperature (b) bottom O_2 (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.

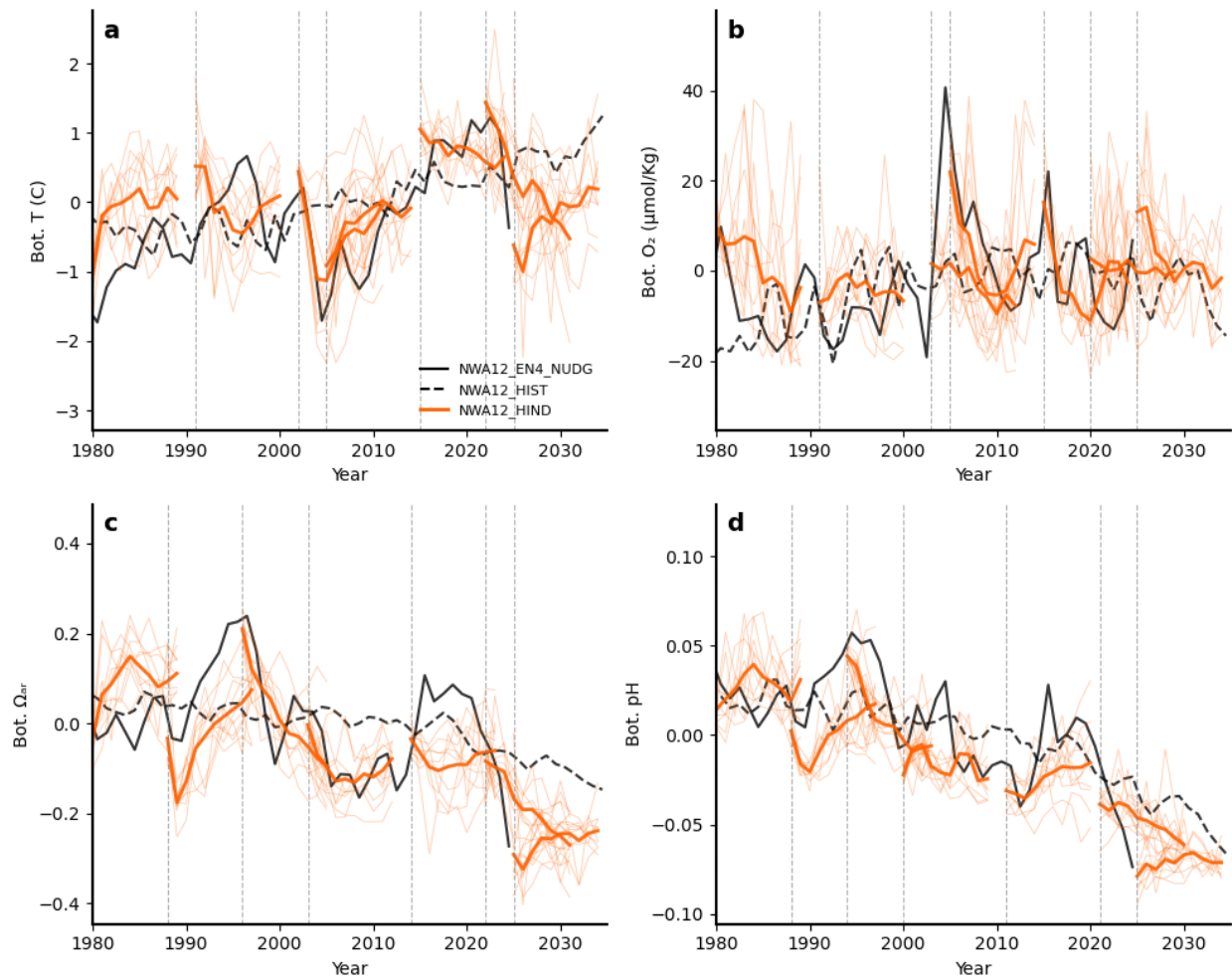


Fig. 4: Physical and biogeochemical predictions in the GOM. Annual mean anomalies of (a) bottom temperature (°C) (b) bottom oxygen (μmol/kg), (c) bottom aragonite saturation state and (d) bottom pH in the GOM_EPU. The solid black lines show the anomalies from the ensemble mean of regional downscaled simulations (NWA12_EN4_NUDG), 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-year long 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 orange lines are the individual ensemble members. A leadtime-dependent climatology was removed from each ensemble member to create anomalies. All anomalies are calculated with respect to the 1994–2023 mean.

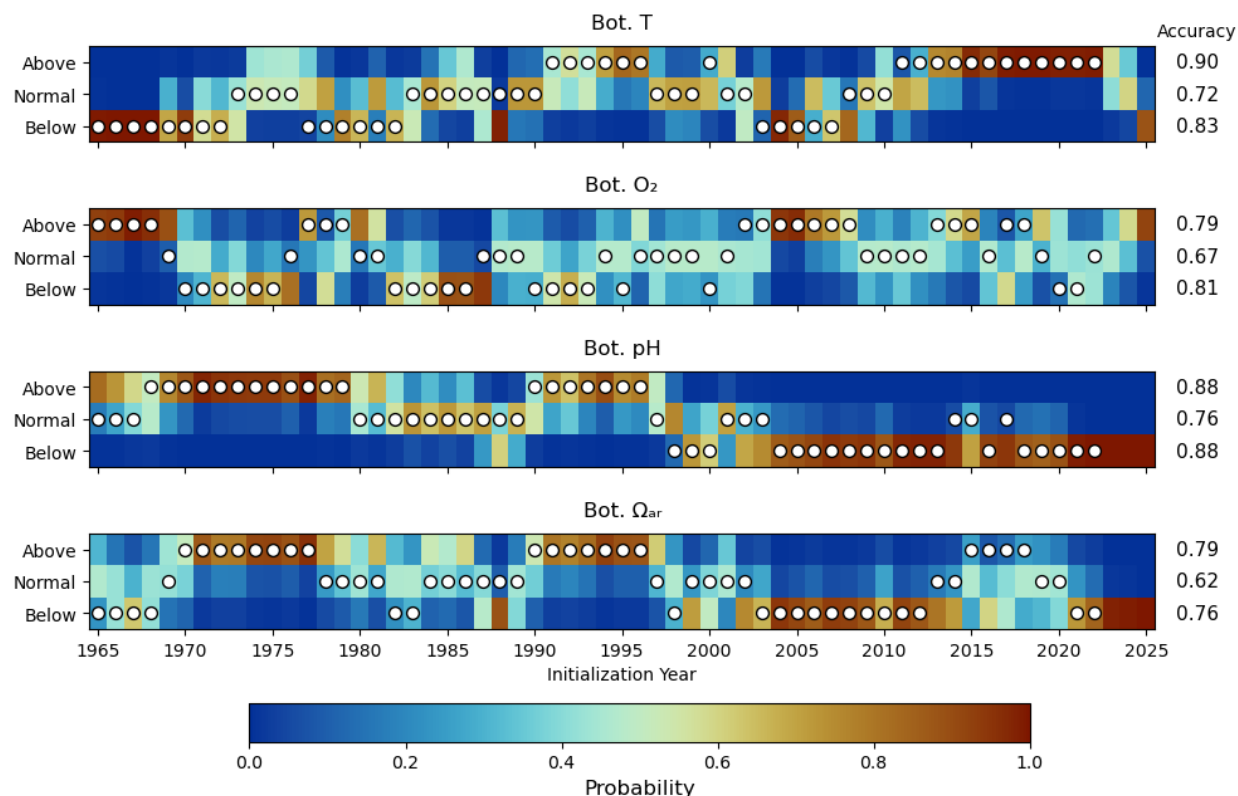


Fig. 5: Probabilistic predictions for the GOM. Probability of the predicted lead 1–3 year mean anomalies (with respect to the 1994–2023 mean) of (a) bottom temperature, (b) bottom O_2 (c) bottom pH and (d) bottom Q_{ar} from the regional downscaled decadal prediction model (NWA12_HIND). The corresponding classifications of the reference anomalies are shown as white circles from the regional downscaled 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, $(\text{True Positives} + \text{True Negatives}) / \text{Total forecasts}$, 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.

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