

# Earth's Future

## RESEARCH ARTICLE

10.1029/2023EF003605

### Key Points:

- Community Earth System Model Seasonal to Multiyear Large Ensemble (SMYLE) forecasts variations in multiple marine stressors up to a year in advance and outperforms statistical persistence
- Novel observation-based products allow for the first skill analysis of subsurface carbon and oxygen
- Analysis of predictability from the SMYLE reconstruction reveals high potential to gain additional temperature and oxygen forecast skill

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

S. C. Mogen,  
[samuel.mogen@colorado.edu](mailto:samuel.mogen@colorado.edu)

### Citation:

Mogen, S. C., Lovenduski, N. S., Yeager, S., Keppler, L., Sharp, J., Bograd, S. J., et al. (2023). Skillful multi-month predictions of ecosystem stressors in the surface and subsurface ocean. *Earth's Future*, 11, e2023EF003605. <https://doi.org/10.1029/2023EF003605>

Received 24 FEB 2023

Accepted 1 OCT 2023

© 2023 The Authors. Earth's Future published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

## Skillful Multi-Month Predictions of Ecosystem Stressors in the Surface and Subsurface Ocean

Samuel C. Mogen<sup>1</sup> , Nicole S. Lovenduski<sup>1</sup> , Stephen Yeager<sup>2</sup>, Lydia Keppler<sup>3</sup> , Jonathan Sharp<sup>4,5</sup> , Steven J. Bograd<sup>6</sup> , Nathali Cordero Quiros<sup>6,7</sup> , Emanuele Di Lorenzo<sup>8</sup> , Elliott L. Hazen<sup>6</sup> , Michael G. Jacob<sup>6,9</sup> , and Mercedes Pozo Buil<sup>6,8</sup> 

<sup>1</sup>Department of Atmospheric and Oceanic Sciences, Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, USA, <sup>2</sup>National Center for Atmospheric Research Climate and Global Dynamics Lab, Boulder, CO, USA, <sup>3</sup>Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA, <sup>4</sup>Cooperative Institute for Climate, Ocean, and Ecosystem Studies, University of Washington, Seattle, WA, USA, <sup>5</sup>National Oceanic and Atmospheric Administration Pacific Marine Environmental Lab, Seattle, WA, USA, <sup>6</sup>National Oceanic and Atmospheric Administration Southwest Fisheries Science Center, Monterey, CA, USA, <sup>7</sup>Institute of Marine Sciences, University of California, Santa Cruz, CA, USA, <sup>8</sup>Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, RI, USA, <sup>9</sup>National Oceanic and Atmospheric Administration Physical Sciences Laboratory, Boulder, CO, USA

**Abstract** Anthropogenic carbon emissions and associated climate change are driving rapid warming, acidification, and deoxygenation in the ocean, which increasingly stress marine ecosystems. On top of long-term trends, short term variability of marine stressors can have major implications for marine ecosystems and their management. As such, there is a growing need for predictions of marine ecosystem stressors on monthly, seasonal, and multi-month timescales. Previous studies have demonstrated the ability to make reliable predictions of the surface ocean physical and biogeochemical state months to years in advance, but few studies have investigated forecast skill of multiple stressors simultaneously or assessed the forecast skill below the surface. Here, we use the Community Earth System Model (CESM) Seasonal to Multiyear Large Ensemble (SMYLE) along with novel observation-based biogeochemical and physical products to quantify the predictive skill of dissolved inorganic carbon (DIC), dissolved oxygen, and temperature in the surface and subsurface ocean. CESM SMYLE demonstrates high physical and biogeochemical predictive skill multiple months in advance in key oceanic regions and frequently outperforms persistence forecasts. We find up to 10 months of skillful forecasts, with particularly high skill in the Northeast Pacific (Gulf of Alaska and California Current Large Marine Ecosystems) for temperature, surface DIC, and subsurface oxygen. Our findings suggest that dynamical marine ecosystem prediction could support actionable advice for decision making.

**Plain Language Summary** Human-driven climate change is rapidly altering the global ocean, with strong warming, increasing acidity, and declining oxygen trends. On top of long-term trends, short term variations can lead to rapid changes that can have major effects on marine ecosystems. There is a growing need to predict these short-term changes in order to better inform marine fisheries managers. In this study, we use a climate model designed to predict changes in the real world months-to-years in advance to better determine our ability to forecast changes. Previous studies with similar goals have been limited by sparse observations of acidity and oxygen. We utilize brand new observational products that estimate acidity and oxygen levels in the subsurface ocean for the first time to analyze subsurface forecasts. Our results demonstrate a high potential to predict warming, acidity, and oxygen levels in key marine ecosystems with this climate model. These results suggest that there is potential for eventual operational forecasts of marine ecosystems to better inform marine managers.

## 1. Introduction

The global ocean is facing growing threats from the accumulation of excess heat and carbon dioxide in the Earth System, leading to ocean warming, acidification, and deoxygenation (Bopp et al., 2013; Doney et al., 2009; Gruber, 2011; Kwiatkowski et al., 2020; Levin, 2018; Whitney, 2022). On top of long-term trends, climate variability and extremes on shorter timescales can rapidly alter temperature and have major effects on regional biogeochemistry (Bednaršek et al., 2018; Di Lorenzo & Mantua, 2016; Mogen et al., 2022, etc.). Since marine

organisms and ecosystems are highly sensitive to changes in their environment across a range of timescales, climate variability and trends will likely alter their health and spatial distribution (Ban et al., 2022; Bednaršek et al., 2016; Cheung et al., 2022; Doney et al., 2009; Pörtner, 2010), which is of great concern for fisheries and aquaculture systems (Cheung et al., 2022; Duarte, 2022; Greene et al., 2017; Mills et al., 2013; C. Moore et al., 2021). Multiple methods are used to make forecasts of marine systems, including statistical forecasts that rely on empirical relationships and dynamical models that simulate fluid dynamics and basic ecosystem processes (Hobday et al., 2016; Jacox et al., 2019, 2020; Tommasi et al., 2017). Dynamical forecasts are much more computationally expensive than statistical forecasts and must outperform statistical forecasts to justify their cost. Persistence forecasts are a type of statistical forecast that propagate anomalies in the ocean state into the future using, for example, autocorrelation and provide regionally valuable predictions that can act as a useful baseline for other forecast methods (Hervieux et al., 2019; Jacox et al., 2019).

New developments in Earth System Model (ESM) forecast systems have enabled short term predictions of variability in ocean physical and biogeochemical state that suggest the possibility of forecast utility for future marine resource management. Accurate forecasts of the ocean state months to years in advance have the potential to inform management practices such as fisheries closures and annual catch limits to rapidly address expected climate variability and change (The State of World Fisheries and Aquaculture 2022, 2022; Tommasi et al., 2017). ESM forecast systems assimilate anomalies in the ocean state and simulate the evolution of these anomalies using a coupled, dynamic model; they have been in use in forecasting studies for multiple decades (Boer, 2004; Brady et al., 2020; Frölicher et al., 2020; Griffies & Bryan, 1997; Lovenduski et al., 2019; Merryfield et al., 2020; Séférian et al., 2014). Recent studies conducted with ESM forecast systems have illustrated high forecast skill months to years in advance for regional sea surface temperature (Jacox et al., 2019; Stock et al., 2015), marine heatwaves (Jacox et al., 2022), subsurface temperature and salinity (Payne et al., 2022), regional surface carbonate chemistry and carbon fluxes (Brady et al., 2019, 2020; Ilyina et al., 2021; Li et al., 2019; Spring et al., 2021), surface chlorophyll (Park et al., 2019), local scale (Fennel et al., 2019; Siedlecki et al., 2016) and near-shore processes (e.g., Brady et al. (2019) in Eastern Boundary Upwelling Systems).

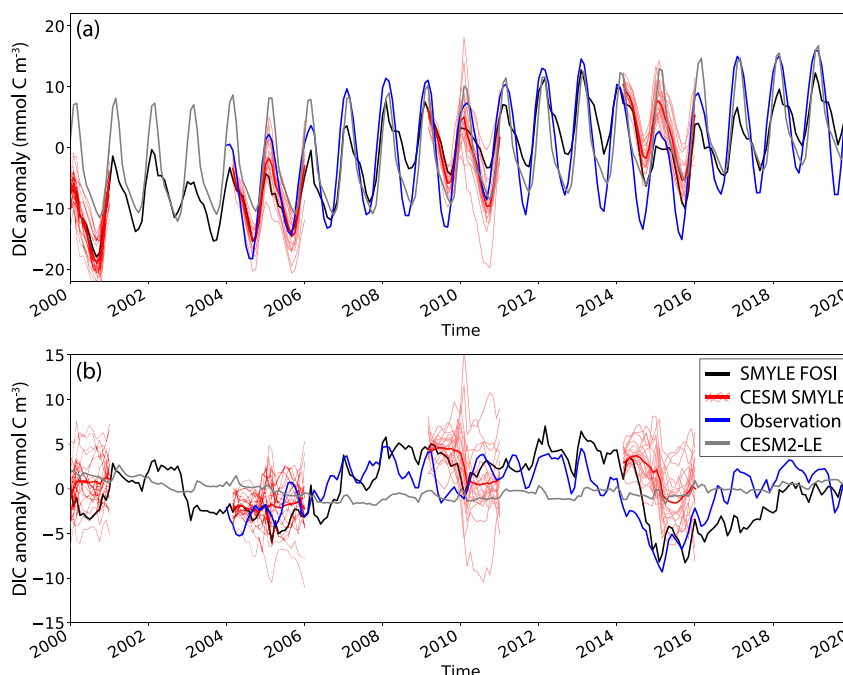
The validation of forecasts of marine stressors is challenged by sparse observations. As such, many short-term prediction studies of ocean biogeochemistry have relied on ESM reconstructions for forecast validation (so-called model predictability) (Frölicher et al., 2020; Krumhardt et al., 2020; Lovenduski et al., 2019; Spring & Ilyina, 2020). Frölicher et al. (2020) investigated predictability horizons of multiple marine stressors in model preindustrial control forecast, and did not use observational products as such. Among those studies that assess true model skill using observations, nearly all focus on forecast validation in the surface ocean, where observations tend to be more plentiful (Brady et al., 2020; Li et al., 2019; Park et al., 2019). Recent advances in ocean biogeochemical observing systems (e.g., Biogeochemical Argo (2023)), together with the widespread use of machine-learning techniques in oceanography, have led to the development of novel global mapped, observation-based products that provide estimates of dissolved oxygen (DO) and dissolved inorganic carbon (DIC) in four dimensions (latitude, longitude, depth, and time) (Keppler et al., 2023b; Sharp et al., 2022b). These new observation-based products facilitate, for the first time, observation-based skill assessment in the interior ocean for short-term forecasts of ocean biogeochemistry.

Here, we use output from a state-of-the-art ESM forecast system and new observation-based products to quantify short-term forecast skill in the physical and biogeochemical state of the surface and subsurface ocean. We focus on the model's ability to make skillful forecasts of marine ecosystem stressors, as these have the greatest potential to inform future management decisions. As we will demonstrate, the ESM generates skillful predictions of temperature, dissolved inorganic carbon, and in the surface and subsurface ocean up to 1 year in advance. We further assess where and when our dynamic model forecasts outperform persistence forecasts and we estimate the as-yet-unrealized forecast skill in marine ecosystem stressors.

## 2. Data and Methods

### 2.1. CESM-SMYLE

Our primary research tool is the Community Earth System Model version 2 (CESM2) Seasonal-to-Multiyear Large Ensemble (SMYLE; (Yeager et al., 2022)). CESM2 simulates the ocean with 60 vertical levels at  $1^\circ \times 1^\circ$  resolution using the Parallel Ocean Program version 2 (POP2) grid (Danabasoglu et al., 2020). CESM2 includes



**Figure 1.** Temporal evolution of monthly surface ocean dissolved inorganic carbon anomalies ( $\text{mmol C m}^{-3}$ ) averaged over the California Current Large Marine Ecosystem from 2000 to 2020 in (red line) four randomly selected CESM SMYLE ensemble forecasts, (black line) SMYLE FOSI, (gray line) CESM2-LE, and (blue line) the observation-based product of Keppler et al. (2023b). DIC anomalies (mean between 2000 and 2020 removed) are plotted (a) with seasonal cycle and long-term trend present, and (b) with seasonal cycle and long-term trend removed.

the Community Atmosphere Model version 6 (CAM6), the Community Land Model version 5 (CLM5), and the CICE version 5.1.2 (sea-ice model; CICE5) (Danabasoglu et al., 2020). CESM2 includes an explicit rendering of marine biogeochemistry from the Marine Biogeochemistry Library (MARBL), which is configured with three explicit phytoplankton functional groups (diatoms, diazotrophs, and picophytoplankton), one implicit group (calcifiers), a single zooplankton type, multi-nutrient co-limitation (N, P, Si, Fe), and prognostic marine carbonate chemistry (Long et al., 2021; J. K. Moore et al., 2001, 2004, 2013). CESM2 is well validated with available ocean observations and reanalysis products, marking an improvement over many structural model aspects as compared to previous generations, with accurate atmospheric and oceanic teleconnections (Danabasoglu et al., 2020). CESM2 is also noted for well represented marine biogeochemistry, apart from the large biases associated with deep North Pacific oxygen ventilation (Long et al., 2021).

As detailed in the SMYLE prediction system description paper (Yeager et al., 2022), CESM SMYLE hindcasts are initialized with physical and biogeochemical output from the Forced Ocean-Sea Ice (FOSI) simulation of CESM2 (SMYLE FOSI). SMYLE FOSI is a simulation of the ocean and sea ice components of CESM2 forced with the Japanese 55-year Reanalysis (JRA-55; (Kobayashi et al., 2015)) momentum, heat, and freshwater fluxes from 1958 to 2019 and atmospheric  $\text{CO}_2$  concentrations (Figure 1a). CESM SMYLE forecasts are initialized quarterly from 1 February 1970 to 1 November 2019 using ocean physical and biogeochemical state variables from SMYLE FOSI. The atmosphere is initialized from JRA-55 output directly interpolated onto the CAM6 grid. The land is initialized from a forced, land-only simulation within CLM5 forced by the merged Climate Research Unit (CRU) and JRA forcing dataset (CRU-JRAv2) applied until equilibrium was achieved. Micro-perturbations of the initial atmospheric temperature state (order  $10^{-14}\text{K}$ ) are applied to each grid cell to generate a 20-member ensemble for each initialization; each ensemble member is integrated for 24 months using the fully coupled CESM2 under historical (1970–2014) and shared socioeconomic pathway (SSP) 3-7.0 scenario (2015–2019) (Figure 1a). CESM SMYLE has previously demonstrated skillful predictions of ENSO indices and surface ocean physical and biogeochemical tracers (Yeager et al., 2022).

We also examine the CESM2 Large Ensemble (CESM2-LE) as an uninitialized model reference for CESM SMYLE. CESM2-LE includes 100 ensemble members integrated over 1850 to 2100, produced to examine the

roles of internal climate variability and external forcing in a changing climate (Rodgers et al., 2021). The historical forcing for SMYLE FOSI is identical to that used in ensemble members 51–100 of CESM2-LE. Members 51–100 have slightly different parameters for ocean deep diffusion, sea ice albedo settings, and MARBL, but still act as a useful benchmark for analysis with CESM SMYLE (Yeager et al., 2022). In contrast to SMYLE FOSI which is forced by reanalysis, CESM2-LE evolves freely with greenhouse gas forcing over two centuries. As the radiative forcing in CESM2-LE is identical to that of CESM2 SMYLE, any difference in behavior stems from initialization. Figure 1a shows the evolution of the CESM2-LE ensemble mean DIC in the California Current surface in comparison to other data products.

## 2.2. Observation-Based Products

We utilize three mapped, global, observation-based products to assess forecast skill: the Roemmich and Gilson (2009) Argo-derived temperature product, the Keppler et al. (2023a) dissolved inorganic carbon product, and the Sharp et al. (2022a) dissolved oxygen product. While DIC is not a direct measure of ocean acidification, ocean acidification (pH) forecasts derived much of their predictability from DIC in Frölicher et al. (2020). The Roemmich-Gilson product provides monthly temperature estimates for the upper ~2,000 m at 1° horizontal resolution over 2004 to present using Argo float data, interpolated to create a mapped observational product. The Roemmich and Gilson (2009) product is well validated, has been used for more than a decade for global ocean analyses, and is regularly updated to include new float data (Roemmich et al., 2015).

Recent work has leveraged combined Argo and Global Ocean Data Analysis Project (GLODAP) climatologies, along with machine learning algorithms to derive gap-filled, gridded, depth-resolved products for both DIC (2004–2019) (Keppler et al., 2023a) and DO (2004–2022) (Sharp et al., 2022a) on monthly timescales. These novel products allow for the first sub-surface model skill assessments at global and regional scales for biogeochemistry. The Mapped Observation-Based Oceanic DIC (MOBO-DIC2004-2019) machine learning approach uses GLODAP cruise data for DIC, along with a series of physical and biogeochemical predictor data to derive a relationship between the tracers and applies this relationship to obtain mapped monthly fields of DIC in the upper 1,500 m of the global ocean ((Keppler et al., 2023b); Figure 1). Similarly, Sharp et al. (2022b) train machine learning algorithms with GLODAP DO and delayed-mode quality-controlled Argo DO data, matched with physical and spatiotemporal predictor data, to derive empirical relationships and create a gap-filled upper ocean DO product - Gridded Ocean Biogeochemistry from Artificial Intelligence—Oxygen (GOBAI-O2 v1.0). Uncertainty estimates in MOBO-DIC and GOBAI-O2 include uncertainties stemming from the measurements, the representation of measurements on a monthly 1° grid, and the prediction uncertainty of the method. The global mean total uncertainty, which accounts for these three sources of uncertainty using standard error propagation is  $18 \mu\text{mol kg}^{-1}$  in MOBO-DIC and  $7.6 \mu\text{mol kg}^{-1}$  in GOBAI-O2 (Keppler et al., 2023b; Sharp et al., 2022b). Keppler et al. (2023b) conducted an in-depth analysis of how MOBO-DIC compares to independent data from time-series stations, BGC-Argo floats, and with synthetic data from an ESM; though GLODAP observations are sparse in time and space and are seasonally biased with few winter observations, MOBO-DIC captures both the seasonal and interannual variability, especially when averaging over large regions. Sharp et al. (2022b) validated GOBAI-O2 using synthetic data from an Earth System Model and demonstrated agreement in local-scale DO seasonality (e.g.,  $R^2 = 0.92 \pm 0.17$  in the upper 200 dbars) and DO interannual variability (e.g.,  $R^2 = 0.66 \pm 0.37$  between 200 and 1,000 dbars).

Observation-based data sets of DIC and DO fill large spatial and temporal gaps. As such, we average over relatively large regions in this analysis (see Section 2.3). As these regions have sufficient training data from GLODAP and/or ARGO, we are confident in the regionally aggregated estimates of DIC and DO. These observation-derived products do not include estimates of DIC or DO for the Arctic Ocean, and MOBO-DIC excludes the Mediterranean Sea.

## 2.3. Statistical Approach

We present forecast skill for three tracers - temperature, DIC, and DO - at two depths - the surface and below the mixed layer at 300 m (as in Frölicher et al. (2020)). Model skill analysis was completed at two spatial scales: large scale U.N. F.A.O. Fisheries Regions (dividing the open ocean) and Large Marine Ecosystems (LMEs, dividing the coastal ocean). We focus our analysis on LMEs with a relatively large number of DIC and DO observational

**Table 1**

*The Number of In Situ Observational Profiles Used in the Creation of MOBO-DIC (From GLODAP Cruise Data) and GOBAI-O<sub>2</sub> (From GLODAP Cruise and ARGO Float Data) Observational Products in Six Select Large Marine Ecosystems (LMEs): The Gulf of Alaska (GOA), California Current (CalCS), Kuroshio Current, Bay of Bengal, Norwegian Sea, and the Greenland Sea*

	MOBO-DIC	GOBAI-O <sub>2</sub>
Gulf of Alaska	195	1,552
California Current	179	944
Kuroshio Current	292	1,131
Bay of Bengal	214	1,190
Norwegian Sea	218	778
Greenland Sea	608	1,062

*Note.* LMEs noted here are shown in Figure S1 in Supporting Information S1.

profiles in GLODAP and BGC Argo, including three LMEs in the North Pacific - the Kuroshio Current LME, the California Current LME, and the Gulf of Alaska LME - and three other regions - the Bay of Bengal LME, the Norwegian Sea LME, and the Greenland Sea LME (Table 1; Figure S1 in Supporting Information S1). We first sorted the LMEs according to the total number of profiles used in the MOBO-DIC/GOBAI-O<sub>2</sub> algorithms and selected LMEs with high observations relative to the rest of this list.

We remove long-term anthropogenic trends (first order, linear) and seasonal climatologies from all data before assessing forecast skill (see, e.g., Figure 1b). We remove model drift from SMYLE forecasts by creating anomalies from model climatology that vary with lead time. Forecast skill at each month after initialization is quantified via the anomaly correlation coefficient (ACC; Pearson's  $r$ -value as shown in Appendix A Equation A1; Figure S2 in Supporting Information S1) and mean absolute error (MAE) of the ensemble mean forecast and the observational product. MAE values for each tracer are normalized by the temporal standard deviation of each tracer at each depth in a given region (see Appendix A Equation A2). The significance (95% Confidence Interval) of ACCs for SMYLE skill assessment is calculated relative

to zero for each lead-time. We generate persistence forecasts by correlating the observed state at initialization with a future state, and multiplying by the autocorrelation function of the observed state at this lag ("damped" persistence). The skill of the uninitialized forecast is assessed via the CESM2-LE ensemble mean state (Figure S2 in Supporting Information S1). Potential predictability in CESM SMYLE is calculated via correlation of CESM SMYLE with SMYLE FOSI. In the interest of brevity, we primarily focus our presentation in this manuscript on the February SMYLE initialization but we include key results from all initializations in the Supplemental.

We assess the unrealized forecast skill, or the difference between the potential predictability of CESM SMYLE and the realized forecast skill as compared to the observational products, using the following approach. First, we count the total number of forecast lead months in the first 13 months for which CESM SMYLE forecast skill ( $r$ ) is larger than 0.5 and exceeds statistical persistence forecasts. Then, we repeat this process for model predictability (wherein SMYLE FOSI is the baseline), and compare the counts. We further quantify the impact of El Nino-Southern Oscillation (ENSO) on background model state by calculating the ACC between the Nino3.4 index and the tracers of interest in SMYLE FOSI.

## 2.4. Model Validation

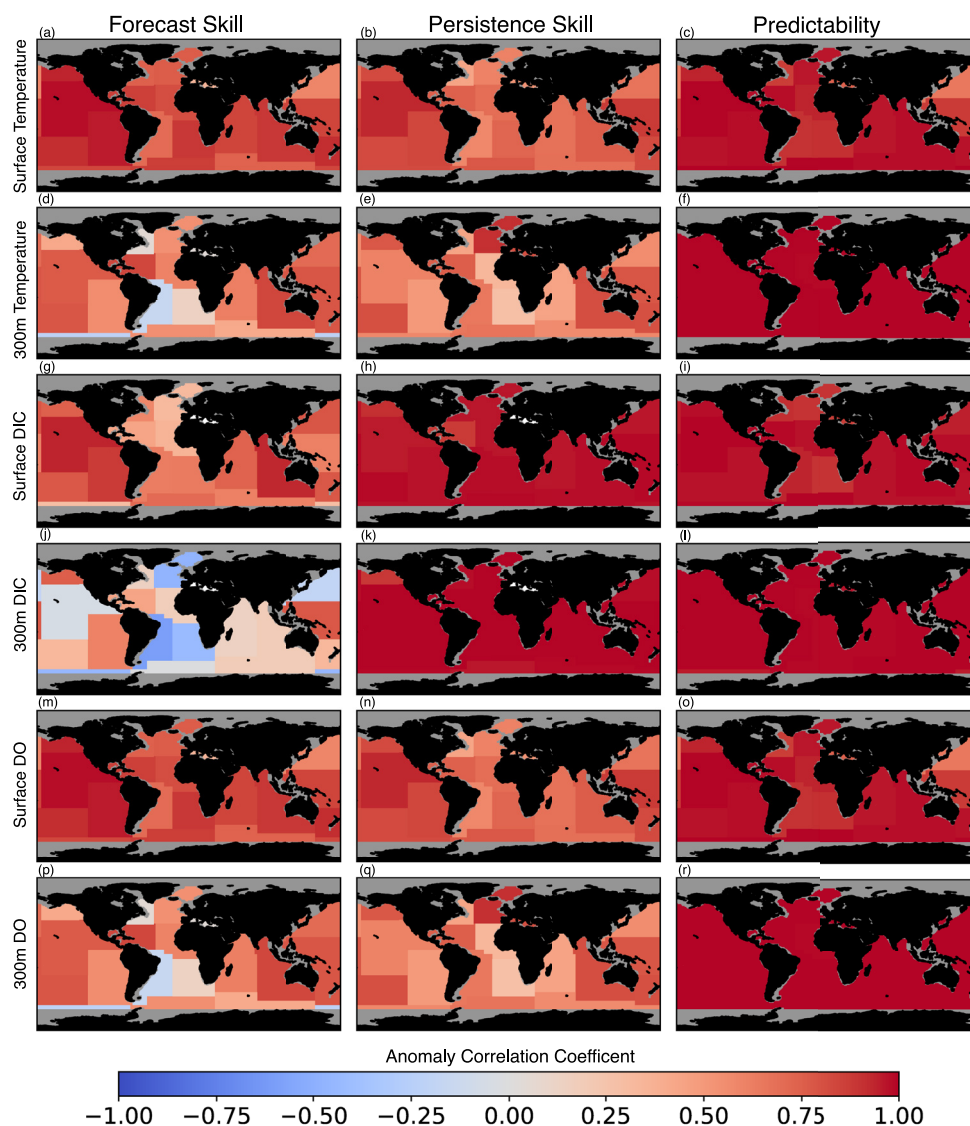
We assess the ability of the SMYLE FOSI to capture spatiotemporal variability in temperature, DIC, and DO by correlating seasonal climatologies and detrended, deseasoned anomalies. In the U.N. F.A.O. Major Fishing Areas, modeled Temperature, DIC, and DO climatologies are well matched with those from the observation-derived products (Figures S3 and S4 in Supporting Information S1). Seasonal climatologies are highly correlated for most of the U.N. fisheries regions (Figure S3 in Supporting Information S1). In certain subsurface regions, such as the southeastern Pacific and eastern equatorial Atlantic (DIC) and the northwest and northeast central Pacific (DO), we find low climatological correlations or anticorrelations, potentially limiting forecast skill (Figure S3 in Supporting Information S1). Anomalies are also highly correlated, with good agreement between observation-derived products and SMYLE FOSI in much of the surface and subsurface ocean for all tracers (Figure S4 in Supporting Information S1), however we note a lack of high correlation in subsurface DIC and DO. These disparities may be related to issues with model representation (e.g., CESM2 deep oxygen ventilation) and observations (relative temporal scarcity of observations). Disparities, regardless of their source, may have contributed to low skill in the noted regions.

## 3. Results

### 3.1. U.N. F.A.O. Fisheries Regions

CESM SMYLE exhibits high forecast skill for surface temperature, DIC, and DO 1 month after initialization in nearly all U.N. fisheries regions (Figure 2, first column). Subsurface DIC forecast skill is notably absent 1 month



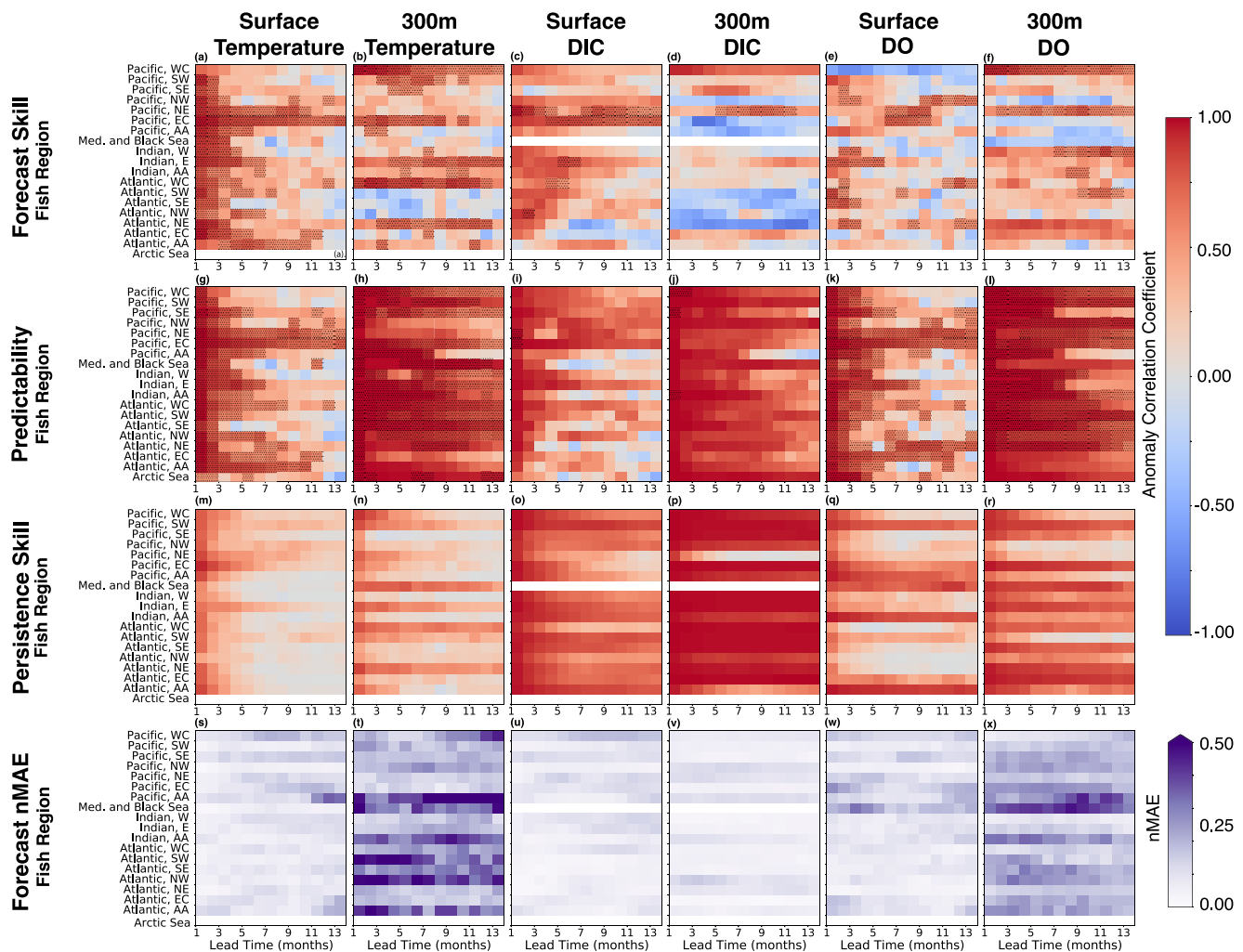


**Figure 2.** Various metrics of forecast skill 1 month following February forecast initialization in the U.N. F.A.O. fisheries regions: (first column) forecast skill, (second column) persistence forecast skill, (third column) predictability. Forecast skill is shown for three variables and two depth levels: (first row) surface temperature, (second row) 300 m temperature, (third row) surface DIC, (fourth row) 300 m DIC, (fifth row) surface DO concentration, and (sixth row) 300 m DO concentration.

after initialization, while temperature and DO exhibit higher subsurface skill in the majority of regions. Although there is a general decline in skill as the fully coupled forecast model evolves further from its initial state (beyond the first lead-time), we find long-lasting (multi-month) skill in many regions that display high skill in the first months after initialization, including in the East-Central and Northeast Pacific (Figures 3a–3f). Corresponding to the decline in skill with forecast lead time, there is also growth in the normalized MAE as the model evolves further from initialization (Figures 3s–3x).

Persistence skill is higher than forecast skill 1 month after initialization for DIC (surface and subsurface) and DO (subsurface) (Figure 2, second column) which likely reflects poor agreement between observations and CESM SMYLE. As with initialized forecast skill, persistence forecasts generally demonstrate a decline in skill with forecast lead time (Figures 3g–3l). Three tracer-depth combinations show particularly high persistence skill: surface and subsurface DIC and subsurface DO, with ACCs close to one with 12 months lead-time.

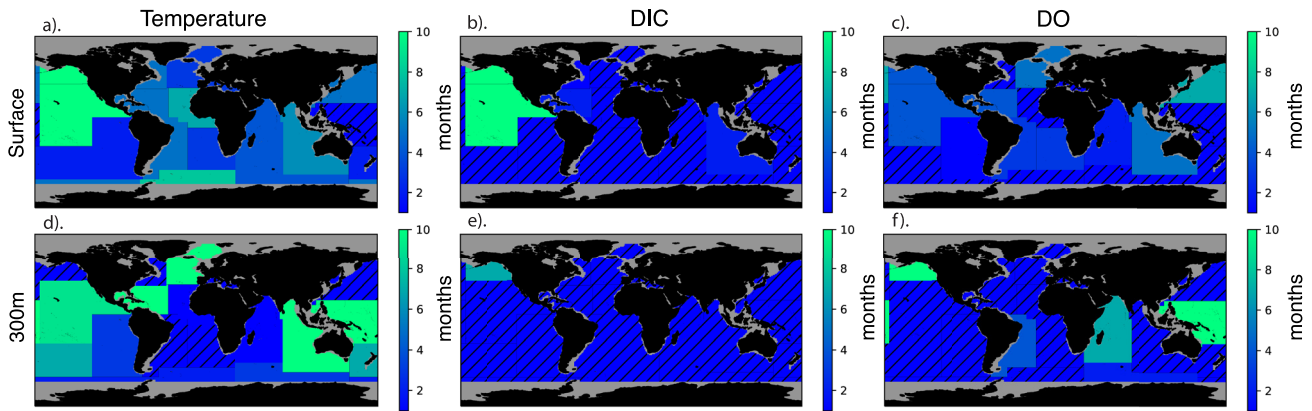
Is the initialized model capable of producing forecasts whose skill exceeds that of statistical persistence? We answer this question by comparing initialized forecast skill to persistence forecast skill across a range of forecast



**Figure 3.** Various metrics of forecast skill over 13 lead months in the U.N. F.A.O. fisheries regions for February forecast initializations: (first row) forecast skill, (second row) predictability forecast, (third row) persistence skill, and (fourth row) normalized mean absolute error of the initialized forecast. Dots in the first and second row indicate skillful forecasts (Model Forecast > Persistence forecast and  $r > 0.5$ ). Forecast skill is shown for three variables and two depth levels: (first column) surface temperature, (second column) 300 m temperature, (third column) surface DIC, (fourth column) 300 m DIC, (fifth column) surface DO concentration, and (sixth column) 300 m DO concentration. The Arctic is not included in observational data products, and thus only appears in model predictability.

lead times (Figure 3). While the persistence skill often matches (or even exceeds) initialized forecast skill in the 1–2 months following initialization (see, e.g., Figure S2 in Supporting Information S1), we find that some U.N. fisheries regions exhibit higher initialized forecast skill than persistence skill after this period, and that this behavior tends to be relatively long-lasting (up to 13 months post-initialization; Figure 3). To draw attention to this multi-month period for which initialized forecast skill exceeds persistence skill, we count the total number of months in which forecast skill is both high ( $r > 0.5$ ) and exceeds persistence for the first 12 months after initialization for each of the U.N. fisheries regions (Figure 4; Figures S5, S6, and S7 in Supporting Information S1). We find up to 10 months of high, persistence skill exceeding forecast skill for most tracer-depth combinations across the U.N. fisheries regions, with consistently high counts in the North Pacific (Figure 4). Temperature in both the surface and subsurface demonstrated high, persistence-exceeding forecast skill in many fisheries regions for up to 10 months, while surface DIC and DO generally display lower counts. Subsurface DIC is nearly devoid of high, persistence-skill-exceeding initialized forecast skill - this is a reflection of low forecast skill combined with particularly high subsurface persistence skill (Figures 3d and 3p). The number of skillful months does not demonstrate large variability with different months of initialization.

We find high, persistence skill exceeding potential predictability (forecast skill quantified using SMYLE FOSI, rather than observations) in surface and subsurface temperature and DO across many of the U.N. fisheries regions



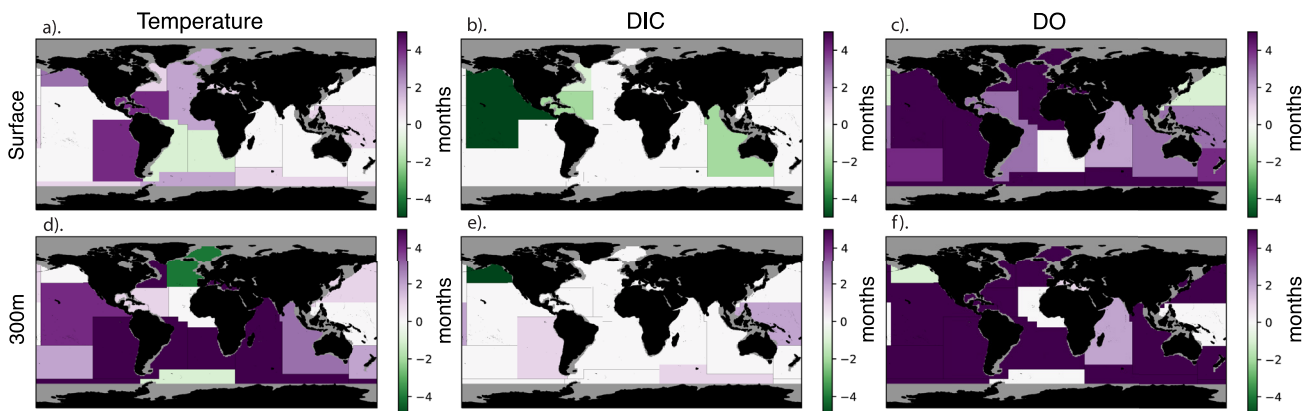
**Figure 4.** The total number of lead-months over which forecast skill is both high ( $r > 0.5$ ), and exceeds persistence skill in the first 12 months following February model initialization in the U.N. F.A.O. fisheries regions for surface (top) and subsurface (bottom) variables. Hatching indicates that there are 0 skillful months.

(Figure 2) for up to 10 months (Figure S8 in Supporting Information S1). In contrast, surface and subsurface DIC potential predictability displays substantially lower lead-month counts (Figure S8 in Supporting Information S1) due to long-lasting persistence forecast skill in SMYLE FOSI.

We find large potential to gain long-lasting forecast skill (Figure 5). We compare the number of skillful (Figure 4) and predictable (Figure S8 in Supporting Information S1) forecast months and find that potential gain is long-lasting in surface and subsurface temperature and DO, whereas we find almost no potential to gain long-lasting forecast skill in surface and subsurface DIC (Figure 5). The lack of potential gain in surface and subsurface DIC is related to the high persistence skill in both observations and model reconstructions, as seen in Figures 3i and 3j, where model predictability is high but falls below persistence forecasts. The consistent potential for gain in model skill relative to model predictability indicates that improvements in both the initialized model and observational products could enhance forecast skill.

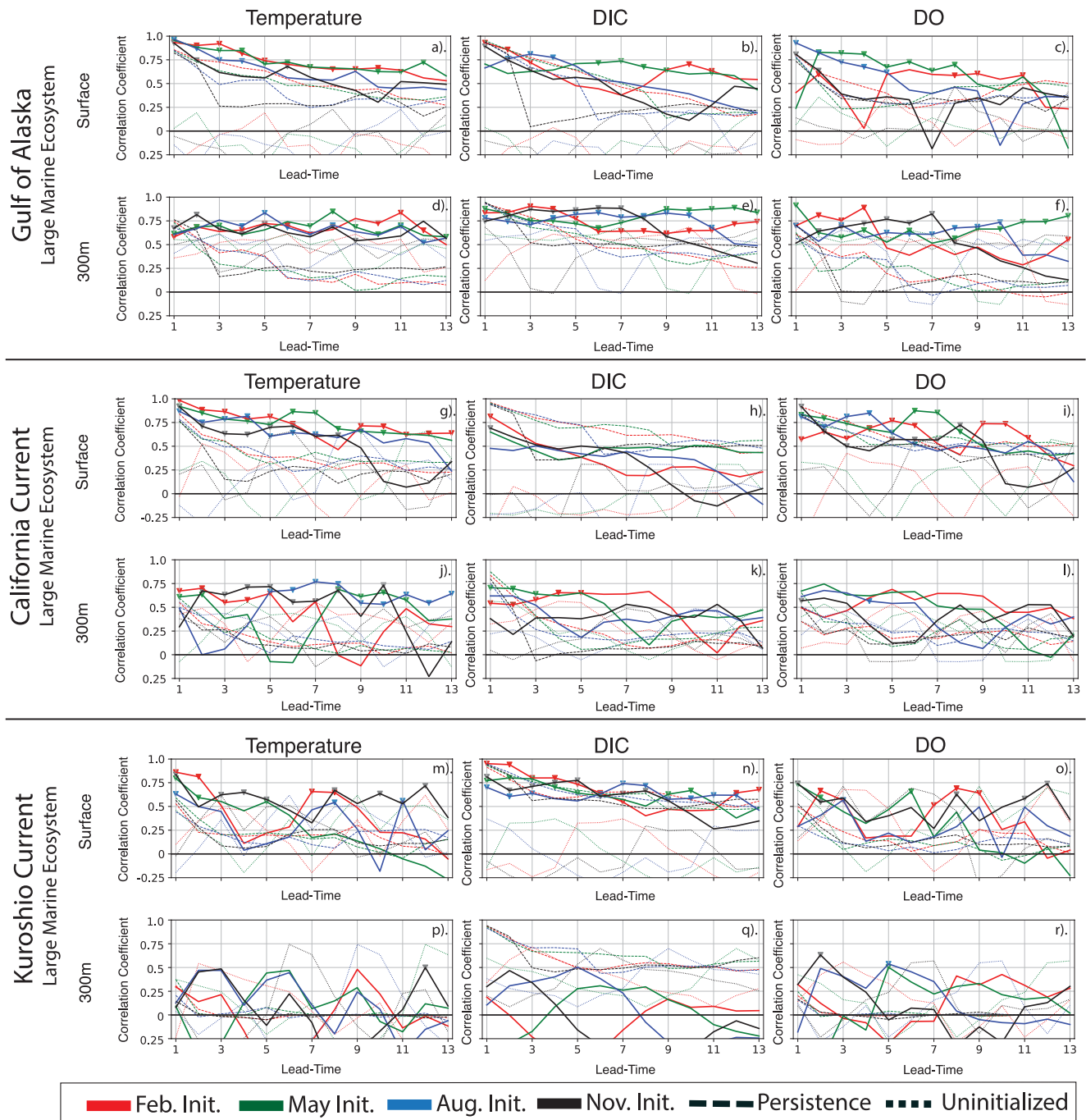
### 3.2. Large Marine Ecosystems

On smaller regional scales, CESM SMYLE displays high, long-lasting forecast skill in three observation-rich North Pacific Large Marine Ecosystems. In Figure 6, we compare temperature, DIC, and DO initialized forecast skill with persistence forecast skill and uninitialized model forecast at the surface and 300 m for the first 13 months post-initialization. The Gulf of Alaska (a–f) displays high skill in both the initialized and persistence forecasts at the surface that decays with forecast lead time. In contrast to the surface, skill is very long-lasting



**Figure 5.** The potential gain in the number of months of forecast skill, estimated as the difference in the number of months for which predictability is high ( $r > 0.5$ ) and exceeds persistence (Figure S8 in Supporting Information S1), and the number of months for which forecast skill is both high and exceeds persistence (Figure 4) in the first 12 months following February model initialization in the U.N. F.A.O. fisheries regions for surface (top) and subsurface (bottom) variables. Positive numbers indicate that predictability exceeds forecast skill.





**Figure 6.** (solid) Forecast skill, (dashed) persistence forecast skill, and (dotted) uninitialized forecast skill in three North Pacific Large Marine Ecosystems for quarterly model initializations over 13 lead-months. Triangles indicate statistically significant skill at the 95% confidence interval.

at depth with high skill ( $r = 0.6$ ) for over a year following initialization at 300 m. We found skillful forecasts in the surface of the Gulf of Alaska for 7 months and ~10 months in the subsurface. The California Current (g–l) has high surface and subsurface skill for temperature with some skill up to 1 year in advance and relatively high surface skill for DO. In contrast, the Kuroshio Current demonstrated low skill, particularly notable in the subsurface. In the Kuroshio Current (m–r), there is high skill ( $r = 0.7$ ) in the surface for DIC, and moderate skill for surface temperature and DO. In contrast, at 300 m, there are no skillful forecasts (excepting DIC persistence). The month of initialization has some impact on forecast skill in North Pacific LMEs, with certain initializations outperforming others; this stands in contrast with the number of skillful months in Fisheries Regions, which was

**Table 2**

*Correlation ( $r$ ) of SMYLE FOSI Surface and 300 m Temperature, DIC, and Oxygen Concentration and the Niño3.4 Index in Observation-Rich Large Marine Ecosystem Regions at Zero-Lag Months*

	GOA	CalCS.	Kuroshio	Bay of Bengal	Greenland Sea	Norwegian Sea
Surface Temp	0.42	0.48	0.08	0.40	0.16	0.21
300 m Temp	0.42	0.48	0.08	0.4	0.16	0.21
Surface DIC	0.44	0.16	0.32	0.17	0.19	0.06
300 m DIC	0.6	0.28	0.17	0.25	0.18	0.13
Surface Oxygen	0.3	0.41	0.04	0.07	0.12	0.26
300 m Oxygen	0.61	0.22	0.16	0.05	0	0.1

relatively insensitive to month of initialization (Figure 4). We found no times/places where forecast skill for a particular monthly initialization is statistically distinct from the other three initializations, although individual pairs are statistically distinct.

We find relatively low forecast skill in three other observation-rich regions outside of the North Pacific (Figure S9 in Supporting Information S1). In the Bay of Bengal, only temperature forecasts display high and long-lasting skill in both the surface and subsurface; subsurface persistence and uninitialized forecasts for DO are also high and long-lasting. In the North Atlantic, neither the Greenland Sea nor Norwegian Sea display consistently high skill. The Greenland Sea is highly persistent in the subsurface for DIC and DO. The Norwegian Sea exhibits skillful surface forecasts for all three tracers, and for subsurface temperature and DO. Although subsurface DIC is highly persistent, we find low initialized forecast skill except for the November model initialization.

Across the observation-rich LMEs, those whose physical and biogeochemical properties are highly correlated with the El Niño-Southern Oscillation (ENSO) are also the LME and tracer combinations that tend to exhibit high forecast skill. We calculated the correlation between the Niño3.4 index and a given variable in each LME of interest for a range of ENSO lead times (Table 2; Table S1 in Supporting Information S1). For example, temperature, and to some extent DIC and DO variations in the Gulf of Alaska and California Current LMEs are highly correlated with ENSO (Table 2, Table S1 in Supporting Information S1); the forecast skill for these tracers/LMEs is also relatively high (Figure 6). In contrast, the other LMEs of interest (aside from surface temperature in the Bay of Bengal) show little relationship with ENSO and also low forecast skill. We speculate that representation of ENSO state at initialization along with good ENSO prediction characteristics of CESM SMYLE in the first 12 forecast months (Yeager et al., 2022) are key contributing factors to high prediction skill for ocean biogeochemical fields in the eastern North Pacific.

#### 4. Conclusions and Discussion

We use an ESM forecast system and new ocean biogeochemical observational products to quantify forecast skill for surface and subsurface ecosystem stressor variations 1–13 months in advance. We find high skill values for three marine stressors—temperature, oxygen, and dissolved inorganic carbon—up to 12 months in advance at two spatial scales. The initialized, dynamic forecast system (CESM SMYLE; (Yeager, 2022)) often produces higher forecast skill than both persistence forecasts and uninitialized forecasts using the same ESM. We also find large potential to gain multi-month forecast skill in temperature and DO, but not for DIC. Ocean biogeochemical forecast skill is somewhat insensitive to the month of initialization, as evidenced by the lack of statistically distinct forecast skill across various monthly initializations in our regions of interest.

In some observation-rich Large Marine Ecosystems, such as the Gulf of Alaska, we find that both the initialized forecast skill and statistical persistence skill tend to be higher in the subsurface than at the surface. As 300 m is well below the dynamic mixed layer and isolated from the atmosphere, we tend to observe long-lasting initialized and persistence forecast skill here. At these depths, statistical persistence skill often outperforms initialized forecast skill. This leads to very few months for which initialized skill is both high ( $r > 0.5$ ) and exceeds the persistence forecast skill. The most notable exception to increasing skill with increasing depth is the lack of model skill in the subsurface Kuroshio Current. This is likely a reflection of the inability of the model to effectively represent the complex dynamics of western boundary currents. As expected, we note that uninitialized forecasts generally

perform worse than initialized and persistence forecasts, except for subsurface DO in the Bay of Bengal (uninitialized outperforms initialized). We also note an unexpected cyclical pattern in uninitialized forecasts in the Bay of Bengal for surface DIC and in the subsurface of some North Pacific LMEs, despite the seasonal climatologies having been removed from all data.

This study builds on decades of prior work utilizing ESMs to forecast a variety of marine stressors months to years in advance. On seasonal to decadal timescales, SST predictive skill has often been closely linked to modes of internal climate variability (Boer, 2004; Griffies & Bryan, 1997; Jacox et al., 2019). While we found some influence of ENSO in our study regions (particularly the northeast Pacific LMEs), this influence could not fully explain forecast skill. Future work is thus needed to identify the role of ENSO in near-term forecast skill.

Our findings align with those reported in another study which quantifies and explores the drivers of surface and subsurface biogeochemical forecasts. Frölicher et al. (2020) investigated the predictability of multiple marine stressors on annual to decadal timescales in both the surface and subsurface using a perfect predictability configuration of a different ESM. Results from their study also indicated a close connection between forecast skill for surface temperature and oxygen (the latter of which is strongly affected by solubility/temperature), with a weaker connection at depth, due to the importance of biological processes in driving oxygen variability. Also similar to our study, Frölicher et al. (2020) noted increasing forecast skill with depth, where waters are less affected by the rapidly varying atmosphere.

As noted above, CESM2 does not accurately represent oxygen in the deep North Pacific Ocean (Long et al., 2021). While the main depths of concern fall below our forecasting interest, it is likely that poor representation of oxygen did impact our results at 300 m. As noted in Frölicher et al. (2020), surface predictions of oxygen are largely driven by solubility/temperature effects, while at depth other factors are at play. For this reason, we are fully confident in our surface results but have some concerns with our results from 300 m as they may be influenced by CESM2 oxygen representation. Further, low correlations of regional inter-monthly variability and seasonal climatology (Figures S3 and S4 in Supporting Information S1) may have impacted DO and DIC forecast skill.

We find that month of initialization has little impact on forecasts in large regions, but can impact smaller scale regions (LMEs). We attribute the sensitivity of smaller scale region predictions to month of initialization to the influence of coastal processes and dynamics on prediction. Coastal regions (e.g., LMEs in this study) exhibit seasonally varying dynamics (e.g., spring upwelling in the California Current System), which might lend skill that varies with the month of initialization or the month that is being forecast. While our model resolution is fairly coarse, LME-scale analysis with similar, coarse-resolution, global forecast systems has demonstrated the ability of such models to accurately represent both physical and biogeochemical properties in LMEs (Brady et al., 2019; Brady et al., 2020; Jacox et al., 2019, 2020, etc.).

While we demonstrate high forecast skill in many cases, comparison with model predictability highlights variables and regions for which skill could be improved further. The potential gain in model skill (i.e., its shortcoming relative to predictability) is apparent, especially for temperature and DO forecasts. In contrast, DIC forecasts show relatively little gain possible relative to predictability, primarily a result of the high skill of statistical persistence forecasts. DIC is relatively insensitive to temperature variability, in contrast with DO which is highly temperature dependent. High memory for DIC indicates that persistence forecasts might continue to be effective in many regions, although it is important to recognize that DIC is not the equivalent of acidity.

Increases in ocean biogeochemical in situ observations (e.g., DO measurements on Biogeochemical Argo floats) will be crucial for realizing potential gains in model skill, by offering new tools for validation and improvements in initialization. The spatial variation in ocean biogeochemical prediction skill likely stems both from limits in the observational record and issues in Earth System Model representation of regional processes. Regions with high observational density are both better understood and simulated. Observational data can be also used to better identify issues with current model structures and lead to improvements. The assimilation of ocean biogeochemical observations into ESMs (Verdy & Mazloff, 2017, e.g.) may be an exciting development for operational biogeochemical forecasting in the future.

The fisheries and aquaculture industry is expected to grow by up to 14% by 2030, with aquaculture projected to overtake capture fisheries as the largest source (The State of World Fisheries and Aquaculture 2022, 2022). As marine ecosystems are increasingly altered by anthropogenic climate change and fisheries adapt, there will be a growing need for accurate forecasts of multiple marine stressors months to years in advance. The CESM

initialized dynamical forecast system can successfully predict variations in biogeochemical stressors in both the surface and subsurface ocean. In addition to previous work focused primarily on physical variables (Jacox et al., 2020; Payne et al., 2017; Tommasi et al., 2017, e.g.), dynamical forecasting of biogeochemistry is thus a promising new direction for the fisheries and aquaculture industry.

Future work should expand on the drivers of skill of physical and biogeochemical stressors within dynamical modeling prediction systems and explore prediction systems using alternate dynamical model structures. In this study, we note that regions of high skill (primarily in the Northeast Pacific) are also regions with large influences from ENSO. ENSO is a predictable phenomena in both linear inverse and dynamical modeling experiments (Barnston et al., 2019; J. Shin et al., 2020; Yeager et al., 2022). While ENSO can be closely connected to marine stressors, as noted by Capotondi et al. (2019), this relationship can only explain up to ~36% of variability. Higher resolution simulations of LMEs within future generations of ESMs or highly tuned regional ocean models (e.g., J-SCOPE; Siedlecki et al., 2016)) may provide further insight into surface and subsurface ocean biogeochemical forecast skill.

In this study, we focus on dynamical forecasts and simple statistical forecasts (persistence), but there exist other statistical approaches for short-term predictions. Statistical Linear Inverse Models (LIMs) have been used to make accurate predictions of El Niño-Southern Oscillation (Capotondi et al., 2019; J. Shin et al., 2020), coastal sea surface height and temperature (Shin & Newman, 2021), and the North Atlantic Oscillation (Albers & Newman, 2021) that are competitive with dynamical forecast systems. Various machine learning (ML) techniques have also been used to create forecasts that are competitive with dynamical modeling systems for predicting tracers such as sea surface temperature and height anomalies (Shao et al., 2021; Wolff et al., 2020), and forecasts of atmospheric events, such as atmospheric rivers (Chapman et al., 2019). Future work may rely on ML techniques to both accelerate model integration and assimilate observations into models (Gettelman et al., 2022), as in Gloege et al. (2022) with marine carbon fluxes. These statistical approaches have also demonstrated promise in predicting different aspects of the Earth System and are often computationally inexpensive when compared with initialized dynamical models, although few studies have directly studied the statistical predictability of marine biogeochemical tracers of interest.

Here, we demonstrate the ability to predict short-term variations in the state of marine stressors, but our results also suggest that there may be potential for forecasting extreme events. Marine heatwaves, for example, can have profound impacts on biogeochemistry throughout the water column (Burger et al., 2020; Mogen et al., 2022). Prior work has demonstrated that dynamical, initialized modeling systems can effectively predict heatwaves (Jacox et al., 2022), but has not examined the potential co-occurring biogeochemical effects which often interact with warm temperatures to redistribute species and affect human uses of the ocean. Future work should focus on using dynamical modeling systems to study the ability to predict ocean biogeochemical extreme events (ocean acidification extremes, deoxygenation events), and their connectivity to physical climate extremes.

CESM SMYLE demonstrates high physical and biogeochemical predictive skill multiple months in advance in key oceanic regions and frequently outperforms persistence forecasts. The three variables that we analyze are important stressors of living marine resources. The regions in which we find high forecast skill, particularly in the North Pacific, hold very large and important fisheries. The continued development of ESM forecasting systems and novel observation-based products may allow for improved marine management in the coming decades.

## Appendix A

### A1. Anomaly Correlation Coefficient (ACC)

Initialized, persistence, and predictability forecasts were assessed with ACC to determined forecast skill from  $-1$  to  $1$ . ACC is a function of initialization month ( $m$ ) and lead-time ( $t$ ) in month (Hervieux et al., 2019).

$$A.C.C. = \frac{\sum_{a=1}^N (F'_a(t, m) \times O'_a(t, m))}{\sqrt{\sum_{a=1}^N (F'_a(t, m)^2 \times O'_a(t, m)^2)}} \quad (A1)$$



where  $F'$  is the forecast anomaly,  $O'$  is the verification field anomaly and ACC is calculated over the period 2004–2020.

## A2. Normalized Mean Absolute Error (nMAE)

nMAE was assessed for initialized forecast as another metric for skill that assesses total error. nMAE, an absolute value, was normalized relative to variance of a given tracer in order to compare across variables and depths.

$$nMAE = \frac{\frac{1}{n} \sum_{i=1}^n |a - b|}{\sigma} \quad (A2)$$

where  $n$  is the sample size,  $a$  and  $b$  are the time series to compare for a given variable, and  $\sigma$  is the standard deviation of the tracer over a given period.

## Data Availability Statement

The CESM Seasonal to Multiyear Large Ensemble and SMYLE FOSI are available at: <https://www.earthsystemgrid.org/dataset/ucar.cgd.cesm2.smyle.html> (Yeager, 2022). The CESM2 Large Ensemble data are available at: <https://www.earthsystemgrid.org/dataset/ucar.cgd.cesm2le.output.html> (IBS Center for Climate Physics et al., 2021).

Argo data were collected and made freely available by the International Argo Program and the national programs that contribute to it (<http://www.argo.ucsd.edu>, <http://argo.jcommops.org>). The Argo Program is part of the Global Ocean Observing System (Argo, 2023). GOBAI-O2 data are available at <https://doi.org/10.25921/z72m-yz67> (Sharp et al., 2022a). The MOBO-DIC summary paper can be found at DOI: <https://doi.org/10.1029/2022GB007677> (Keppler et al., 2023a).

Shapefiles for Large Marine Ecosystems were found at <https://www.sciencebase.gov/catalog/item/55c77722e4b-08400b1fd8244> (FAO, 2020). Shapefiles for U.N. FAO Fisheries Regions were found at <https://www.fao.org/fishery/en/area/search> (Sherman et al., 2017).

## Acknowledgments

SCM and NSL were supported by the National Oceanic and Atmospheric Administration (NA20OAR4310405) and the National Science Foundation (OCE 1752724). SGY acknowledges support from the Regional and Global Model Analysis (RGMA) component of the Earth and Environmental System Modeling Program of the U.S. Department of Energy's Office of Biological & Environmental Research (BER) under Award Number DE-SC0022070. This work also was supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation (NSF) under Cooperative Agreement No. 1852977. We also thank the National Center for Atmospheric Research Earth System Working Group for their development of invaluable software tools used in processing CESM SMYLE: <https://github.com/CESM-ES-PWG/ESP-Lab>. This is CICOES contribution no. 2023-1270 and PMEL contribution no. 5507. We are grateful for helpful feedback from Dillon Amaya and two anonymous reviewers.

## References

- Albers, J. R., & Newman, M. (2021). Subseasonal predictability of the North Atlantic Oscillation. *Environmental Research Letters*, 16(4), 044024. <https://doi.org/10.1088/1748-9326/abe781>
- Argo (2023). Argo float data and metadata from global data assembly centre (Argo GDAC) [Dataset]. SEANO. <https://doi.org/10.17882/42182>
- Ban, Z., Hu, X., & Li, J. (2022). Tipping points of marine phytoplankton to multiple environmental stressors. *Nature Climate Change*, 12(11), 1045–1051. <https://doi.org/10.1038/s41558-022-01489-0>
- Barnston, A. G., Tippet, M. K., Ranganathan, M., & L'Heureux, M. L. (2019). Deterministic skill of ENSO predictions from the North American multimodel ensemble. *Climate Dynamics*, 53(12), 7215–7234. <https://doi.org/10.1007/s00382-017-3603-3>
- Bednaršek, N., Feely, R. A., Beck, M. W., Glippa, O., Kanerva, M., & Engström-Öst, J. (2018). El Niño-Related thermal stress coupled with upwelling-related ocean acidification negatively impacts cellular to population-level responses in pteropods along the California current system with implications for increased bioenergetic costs. *Frontiers in Marine Science*, 5, 486. <https://doi.org/10.3389/fmars.2018.00486>
- Bednaršek, N., Harvey, C. J., Kaplan, I. C., Feely, R. A., & Možina, J. (2016). Pteropods on the edge: Cumulative effects of ocean acidification, warming, and deoxygenation. *Progress in Oceanography*, 145, 1–24. <https://doi.org/10.1016/j.pcean.2016.04.002>
- Boer, G. J. (2004). Long time-scale potential predictability in an ensemble of coupled climate models. *Climate Dynamics*, 23(1), 29–44. <https://doi.org/10.1007/s00382-004-0419-8>
- Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., et al. (2013). Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. *Biogeosciences*, 10(10), 6225–6245. <https://doi.org/10.5194/bg-10-6225-2013>
- Brady, R. X., Lovenduski, N. S., Alexander, M. A., Jacox, M., & Gruber, N. (2019). On the role of climate modes in modulating the air–sea CO<sub>2</sub> fluxes in eastern boundary upwelling systems. *Biogeosciences*, 16(2), 329–346. <https://doi.org/10.5194/bg-16-329-2019>
- Brady, R. X., Lovenduski, N. S., Yeager, S. G., Long, M. C., & Lindsay, K. (2020). Skillful multiyear predictions of ocean acidification in the California current system. *Nature Communications*, 11(1), 2166. <https://doi.org/10.1038/s41467-020-15722-x>
- Burger, F. A., John, J. G., & Frölicher, T. L. (2020). Increase in ocean acidity variability and extremes under increasing atmospheric CO<sub>2</sub>. *Biogeosciences*, 17(18), 4633–4662. <https://doi.org/10.5194/bg-17-4633-2020>
- Capotondi, A., Sardeshmukh, P. D., Di Lorenzo, E., Subramanian, A. C., & Miller, A. J. (2019). Predictability of US west coast ocean temperatures is not solely due to ENSO. *Scientific Reports*, 9(1), 10993. <https://doi.org/10.1038/s41598-019-47400-4>
- Chapman, W. E., Subramanian, A. C., Delle Monache, L., Xie, S. P., & Ralph, F. M. (2019). Improving atmospheric river forecasts with machine learning. *Geophysical Research Letters*, 46(17–18), 10627–10635. <https://doi.org/10.1029/2019GL083662>
- Cheung, W. W. L., Wei, C.-L., & Levin, L. A. (2022). Vulnerability of exploited deep-sea demersal species to ocean warming, deoxygenation, and acidification. *Environmental Biology of Fishes*, 105(10), 1301–1315. <https://doi.org/10.1007/s10641-022-01321-w>
- Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The community earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, 12(2). <https://doi.org/10.1029/2019MS001916>

- Di Lorenzo, E., & Mantua, N. (2016). Multi-year persistence of the 2014/15 North Pacific marine heatwave. *Nature Climate Change*, 6(11), 1042–1047. <https://doi.org/10.1038/nclimate3082>
- Doney, S. C., Fabry, V. J., Feely, R. A., & Kleypas, J. A. (2009). Ocean acidification: The other CO<sub>2</sub> problem. *Annual Review of Marine Science*, 1(1), 169–192. <https://doi.org/10.1146/annurev.marine.010908.163834>
- Duarte, J. A. (2022). Ocean acidification effects on aquaculture of a high resilient calcifier species. A bioeconomic approach. *Aquaculture*, 559, 738426.
- FAO. (2020). FAO statistical areas for fishery purposes [Dataset]. FAO Fisheries and Aquaculture Department. Retrieved from <https://www.fao.org/fishery/area/search/en>
- Fennel, K., Alin, S., Barbero, L., Evans, W., Bourgeois, T., Cooley, S., et al. (2019). Carbon cycling in the North American coastal ocean: A synthesis. *Biogeosciences*, 16(6), 1281–1304. <https://doi.org/10.5194/bg-16-1281-2019>
- Frölicher, T. L., Ramseyer, L., Raible, C. C., Rodgers, K. B., & Dunne, J. (2020). Potential predictability of marine ecosystem drivers. *Biogeosciences*, 17(7), 2061–2083. <https://doi.org/10.5194/bg-17-2061-2020>
- Gettelman, A., Geer, A. J., Forbes, R. M., Carmichael, G. R., Feingold, G., Posselt, D. J., et al. (2022). The future of Earth system prediction: Advances in model-data fusion. *Science Advances*, 8(14), eabn3488. <https://doi.org/10.1126/sciadv.abn3488>
- Gloge, L., Yan, M., Zheng, T., & McKinley, G. A. (2022). Improved quantification of ocean carbon uptake by using machine learning to merge global models and pCO<sub>2</sub> data. *Journal of Advances in Modeling Earth Systems*, 14(2), e2021MS002620. <https://doi.org/10.1029/2021MS002620>
- Greene, C. H., Scott-Buechler, C. M., Hausner, A. L. P., Johnson, Z. I., Lei, X. G., & Huntley, M. E. (2017). Transforming the future of marine aquaculture. *A Circular Economy Approach*, 9.
- Griffies, S. M., & Bryan, K. (1997). Predictability of North Atlantic multidecadal climate variability. *Science*, 275(5297), 181–184. <https://doi.org/10.1126/science.275.5297.181>
- Gruber, N. (2011). Warming up, turning sour, losing breath: Ocean biogeochemistry under global change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 369(1943), 1980–1996. <https://doi.org/10.1098/rsta.2011.0003>
- Hervieux, G., Alexander, M. A., Stock, C. A., Jacox, M. G., Pegion, K., Becker, E., et al. (2019). More reliable coastal SST forecasts from the North American multimodel ensemble. *Climate Dynamics*, 53(12), 7153–7168. <https://doi.org/10.1007/s00382-017-3652-7>
- Hobday, A. J., Spillman, C. M., Paige Eveson, J., & Hartog, J. R. (2016). Seasonal forecasting for decision support in marine fisheries and aquaculture. *Fisheries Oceanography*, 25(S1), 45–56. <https://doi.org/10.1111/fog.12083>
- IBS Center for Climate Physics, Danabasoglu, G., Deser, C., Rodgers, K., & Timmermann, A. (2021). National center for atmospheric research [Dataset]. CESM2 Large Ensemble Community Project, UCAR/NCAR Climate Data Gateway. <https://doi.org/10.26024/KGMP-C556>
- Ilyina, T., Li, H., Spring, A., Müller, W. A., Bopp, L., Chikamoto, M. O., et al. (2021). Predictable variations of the carbon sinks and atmospheric CO<sub>2</sub> growth in a multi-model framework. *Geophysical Research Letters*, 48(6), e2020GL090695. <https://doi.org/10.1029/2020GL090695>
- Jacox, M. G., Alexander, M. A., Amaya, D., Becker, E., Bograd, S. J., Brodie, S., et al. (2022). Global seasonal forecasts of marine heatwaves. *Nature*, 604(7906), 486–490. <https://doi.org/10.1038/s41586-022-04573-9>
- Jacox, M. G., Alexander, M. A., Siedlecki, S., Chen, K., Kwon, Y.-O., Brodie, S., et al. (2020). Seasonal-to-interannual prediction of North American coastal marine ecosystems: Forecast methods, mechanisms of predictability, and priority developments. *Progress in Oceanography*, 183, 102307. <https://doi.org/10.1016/j.pocean.2020.102307>
- Jacox, M. G., Alexander, M. A., Stock, C. A., & Hervieux, G. (2019). On the skill of seasonal sea surface temperature forecasts in the California current system and its connection to ENSO variability. *Climate Dynamics*, 53(12), 7519–7533. <https://doi.org/10.1007/s00382-017-3608-y>
- Keppler, L., Landschützer, P., Lauvset, S. K., & Gruber, N. (2023a). Mapped observation-based oceanic dissolved inorganic carbon monthly fields from 2004 through 2019 (MOBO-DIC2004-2019) (NCEI Accession 0277099) [Dataset]. NOAA National Centers for Environmental Information. <https://doi.org/10.25921/Z31N-3M26>
- Keppler, L., Landschützer, P., Lauvset, S. K., & Gruber, N. (2023b). Recent trends and variability in the oceanic storage of dissolved inorganic carbon. *Global Biogeochemical Cycles*, 37(5), e2022GB007677. <https://doi.org/10.1029/2022GB007677>
- Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 Reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan. Series II*, 93(1), 5–48. <https://doi.org/10.2151/jmsj.2015-001>
- Krumhardt, K. M., Lovenduski, N. S., Long, M. C., Luo, J. Y., Lindsay, K., Yeager, S., & Harrison, C. (2020). Potential predictability of net primary production in the Ocean. *Global Biogeochemical Cycles*, 34(6), e2020GB006531. <https://doi.org/10.1029/2020GB006531>
- Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., et al. (2020). Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. *Biogeosciences*, 17(13), 3439–3470. <https://doi.org/10.5194/bg-17-3439-2020>
- Levin, L. A. (2018). Manifestation, drivers, and emergence of open ocean deoxygenation. *Annual Review of Marine Science*, 10(1), 229–260. <https://doi.org/10.1146/annurev-marine-121916-063359>
- Li, H., Ilyina, T., Müller, W. A., & Landschützer, P. (2019). Predicting the variable ocean carbon sink. *Science Advances*, 5(4), eaav6471. <https://doi.org/10.1126/sciadv.aav6471>
- Long, M. C., Moore, J. K., Lindsay, K., Levy, M., Doney, S. C., Luo, J. Y., et al. (2021). Simulations with the marine biogeochemistry library (MARBL). *Journal of Advances in Modeling Earth Systems*, 13(12), e2021MS002647. <https://doi.org/10.1029/2021MS002647>
- Lovenduski, N. S., Yeager, S. G., Lindsay, K., & Long, M. C. (2019). Predicting near-term variability in ocean carbon uptake. *Earth System Dynamics*, 10(1), 45–57. <https://doi.org/10.5194/esd-10-45-2019>
- Merryfield, W. J., Baehr, J., Batté, L., Becker, E. J., Butler, A. H., Coelho, C. A. S., et al. (2020). Current and emerging developments in subseasonal to decadal prediction. *Bulletin of the American Meteorological Society*, 101(6), E869–E896. <https://doi.org/10.1175/BAMS-D-19-0037.1>
- Mills, K., Pershing, A., Brown, C., Chen, Y., Chiang, F.-S., Holland, D., et al. (2013). Fisheries management in a changing climate: Lessons from the 2012 ocean heat wave in the Northwest Atlantic. *Oceanography*, 26(2). <https://doi.org/10.5670/oceanog.2013.27>
- Mogen, S. C., Lovenduski, N. S., Dallmann, A. R., Gregor, L., Sutton, A. J., Bograd, S. J., et al. (2022). Ocean biogeochemical signatures of the North Pacific Blob. *Geophysical Research Letters*, 49(9), e2021GL096938. <https://doi.org/10.1029/2021GL096938>
- Moore, C., Morley, J. W., Morrison, B., Kolian, M., Horsch, E., Frölicher, T., et al. (2021). Estimating the economic impacts of climate change on 16 major US fisheries. *Climate Change Economics*, 12(01), 2150002. <https://doi.org/10.1142/S2010007821500020>
- Moore, J. K., Doney, S. C., Kleypas, J. A., Glover, D. M., & Fung, I. Y. (2001). An intermediate complexity marine ecosystem model for the global domain. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(1–3), 403–462. [https://doi.org/10.1016/S0967-0645\(01\)00108-4](https://doi.org/10.1016/S0967-0645(01)00108-4)
- Moore, J. K., Doney, S. C., & Lindsay, K. (2004). Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model: Global ecosystem-biogeochemical model. *Global Biogeochemical Cycles*, 18(4), GB4028. <https://doi.org/10.1029/2004GB002220>
- Moore, J. K., Lindsay, K., Doney, S. C., Long, M. C., & Misumi, K. (2013). Marine ecosystem dynamics and biogeochemical cycling in the community earth system model [CESM1(BGC)]: Comparison of the 1990s with the 2090s under the RCP4.5 and RCP8.5 scenarios. *Journal of Climate*, 26(23), 9291–9312. <https://doi.org/10.1175/JCLI-D-12-00566.1>

- Park, J.-Y., Stock, C. A., Dunne, J. P., Yang, X., & Rosati, A. (2019). Seasonal to multiannual marine ecosystem prediction with a global Earth system model. *Science*, 365(6450), 284–288. <https://doi.org/10.1126/science.aav6634>
- Payne, M. R., Danabasoglu, G., Keenlyside, N., Matei, D., Miesner, A. K., Yang, S., & Yeager, S. G. (2022). Skilful decadal-scale prediction of fish habitat and distribution shifts. *Nature Communications*, 13(1), 2660. <https://doi.org/10.1038/s41467-022-30280-0>
- Payne, M. R., Hobday, A. J., MacKenzie, B. R., Tommasi, D., Dempsey, D. P., Fässler, S. M. M., et al. (2017). Lessons from the first generation of marine ecological forecast products. *Frontiers in Marine Science*, 4, 289. <https://doi.org/10.3389/fmars.2017.00289>
- Pörtner, H.-O. (2010). Oxygen- and capacity-limitation of thermal tolerance: A matrix for integrating climate-related stressor effects in marine ecosystems. *Journal of Experimental Biology*, 213(6), 881–893. <https://doi.org/10.1242/jeb.037523>
- Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., et al. (2021). Ubiquity of human-induced changes in climate variability. *Earth System Dynamics*, 12(4), 1393–1411. <https://doi.org/10.5194/esd-12-1393-2021>
- Roemmich, D., Church, J., Gilson, J., Monselesan, D., Sutton, P., & Wijffels, S. (2015). Unabated planetary warming and its ocean structure since 2006. *Nature Climate Change*, 5(3), 240–245. <https://doi.org/10.1038/nclimate2513>
- Roemmich, D., & Gilson, J. (2009). The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. *Progress in Oceanography*, 82(2), 81–100. <https://doi.org/10.1016/j.pocean.2009.03.004>
- Séférian, R., Bopp, L., Gehlen, M., Swingedouw, D., Mignot, J., Guilyardi, E., & Servonnat, J. (2014). Multiyear predictability of tropical marine productivity. *Proceedings of the National Academy of Sciences of the United States of America*, 111(32), 11646–11651. <https://doi.org/10.1073/pnas.1315855111>
- Shao, Q., Li, W., Han, G., Hou, G., Liu, S., Gong, Y., & Qu, P. (2021). A deep learning model for forecasting sea surface height anomalies and temperatures in the South China Sea. *Journal of Geophysical Research: Oceans*, 126(7), e2021JC017515. <https://doi.org/10.1029/2021JC017515>
- Sharp, J. D., Fassbender, A. J., Carter, B. R., Johnson, G. C., Schultz, C., & Dunne, J. P. (2022a). GOBAI-O2: A global gridded monthly dataset of ocean interior dissolved oxygen concentrations based on shipboard and autonomous observations (NCEI accession 0259304) [Dataset]. NOAA National Centers for Environmental Information. <https://doi.org/10.25921/Z72M-YZ67>
- Sharp, J. D., Fassbender, A. J., Carter, B. R., Johnson, G. C., Schultz, C., & Dunne, J. P. (2022b). GOBAI-O2: Temporally and spatially resolved fields of ocean interior dissolved oxygen over nearly two decades (preprint). *ESSD – Ocean/Chemical oceanography*. <https://doi.org/10.5194/essd-2022-308>
- Sherman, K., Shuford, R. L., Adams, S., & Bosch, J. (2017). Large marine ecosystems [Dataset]. USGS. Retrieved from <https://www.sciencebase.gov/catalog/item/55c77722e4b08400b1fd8244>
- Shin, J., Park, S., Shin, S., Newman, M., & Alexander, M. A. (2020). Enhancing ENSO prediction skill by combining model-analog and linear inverse models (MA-LIM). *Geophysical Research Letters*, 47(1), e2019GL085914. <https://doi.org/10.1029/2019GL085914>
- Shin, S., & Newman, M. (2021). Seasonal predictability of global and North American coastal sea surface temperature and height anomalies. *Geophysical Research Letters*, 48(10), e2020GL091886. <https://doi.org/10.1029/2020GL091886>
- Siedlecki, S. A., Kaplan, I. C., Hermann, A. J., Nguyen, T. T., Bond, N. A., Newton, J. A., et al. (2016). Experiments with seasonal forecasts of ocean conditions for the northern region of the California current upwelling system. *Scientific Reports*, 6(1), 27203. <https://doi.org/10.1038/srep27203>
- Spring, A., Dunkl, I., Li, H., Brovkin, V., & Ilyina, T. (2021). Trivial improvements in predictive skill due to direct reconstruction of the global carbon cycle. *Earth System Dynamics*, 12(4), 1139–1167. <https://doi.org/10.5194/esd-12-1139-2021>
- Spring, A., & Ilyina, T. (2020). Predictability horizons in the global carbon cycle inferred from a perfect-model framework. *Geophysical Research Letters*, 47(9), e2019GL085311. <https://doi.org/10.1029/2019GL085311>
- Stock, C. A., Pegion, K., Vecchi, G. A., Alexander, M. A., Tommasi, D., Bond, N. A., et al. (2015). Seasonal sea surface temperature anomaly prediction for coastal ecosystems. *Progress in Oceanography*, 137, 219–236. <https://doi.org/10.1016/j.pocean.2015.06.007>
- The State of World Fisheries and Aquaculture 2022. (2022). FAO. <https://doi.org/10.4060/cc0461en>
- Tommasi, D., Stock, C. A., Hobday, A. J., Methot, R., Kaplan, I. C., Eveson, J. P., et al. (2017). Managing living marine resources in a dynamic environment: The role of seasonal to decadal climate forecasts. *Progress in Oceanography*, 152, 15–49. <https://doi.org/10.1016/j.pocean.2016.12.011>
- Verdy, A., & Mazloff, M. R. (2017). A data assimilating model for estimating Southern Ocean biogeochemistry. *Journal of Geophysical Research: Oceans*, 122(9), 6968–6988. <https://doi.org/10.1002/2016JC012650>
- Whitney, M. M. (2022). Observed and projected global warming pressure on coastal hypoxia. *Biogeosciences*, 19(18), 4479–4497. <https://doi.org/10.5194/bg-19-4479-2022>
- Wolff, S., O'Donncha, F., & Chen, B. (2020). Statistical and machine learning ensemble modelling to forecast sea surface temperature. *Journal of Marine Systems*, 208, 103347. <https://doi.org/10.1016/j.jmarsys.2020.103347>
- Yeager, S. (2022). Seasonal-to-multiyear large ensemble (SMYLE) experiment [Dataset]. UCAR/NCAR - CISL - CDP. <https://doi.org/10.26024/PWMA-RE41>
- Yeager, S. G., Rosenbloom, N., Glanville, A. A., Wu, X., Simpson, I., Li, H., et al. (2022). The seasonal-to-multiyear large ensemble (SMYLE) prediction system using the Community Earth System Model version 2. *Geoscientific Model Development*, 15(16), 6451–6493. <https://doi.org/10.5194/gmd-15-6451-2022>