Skillful multi-month predictions of ecosystem stressors in the surface and subsurface ocean

Samuel C. Mogen[1], Nicole S. Lovenduski[1], Stephen Yeager [2], Lydia Keppler[3], Jonathan Sharp[4], Steven J. Bograd[5], Nathali Cordero Quiros[5,7], Emanuele Di Lorenzo[8], Elliott L. Hazen[5], Michael G. Jacox[5,6], Mercedes Pozo Buil[5,7]

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

Key Points:

- CESM 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

Corresponding author: Samuel C Mogen, samuel.mogen@colorado.edu

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 ecosystems 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 forecasts 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, 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 driving major alterations within 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; Kwiatkowski et al., 2020; Whitney, 2022; Levin, 2018; Gruber, 2011). On top of long-term trends, climate variability and extremes on shorter timescales can rapidly alter temperature and have major effects on regional biogeochemistry (Di Lorenzo & Mantua, 2016; Bednaršek et al., 2018; Mogen et al., 2022). 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 success and spatial distribution (Cheung et al., 2022; Doney et al., 2009; Ban et al., 2022; Bednaršek et al., 2016; Pörtner, 2010), which is of great concern for fisheries and aquaculture systems (Moore et al., 2021; Cheung et al., 2022; Mills et al., 2013; Duarte, 2022; Greene et al., 2022). 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 (Tommasi et al., 2017; Jacox et al., 2020). Persistence forecasts are a type of statistical forecast that propagate anomalies in the ocean state into the future using, for example, autocorrelation and pro-

vide regionaly valuable predictions that can act as a useful baseline for other forecast methods (Jacox et al., 2019; Hervieux et al., 2019).

New developments in Earth System Model 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 (Tommasi et al., 2017; *The State of World Fisheries and Aquaculture 2022*, 2022). Earth System Model (ESM) forecast systems assimilate anomalies in the ocean state and simulate the evolution of these anomalies using a coupled, dynamic model (Merryfield et al., 2020; Brady et al., 2020; Lovenduski et al., 2019). 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), marine heatwaves (Jacox et al., 2022), subsurface temperature and salinity (Payne et al., 2022), regional surface carbonate chemistry and carbon fluxes (Li et al., 2019; Brady et al., 2020; Ilyina et al., 2021; Spring et al., 2021), surface chlorophyll (Park et al., 2019), and local scale processes (Siedlecki et al., 2016; Fennel et al., 2019).

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) (Lovenduski et al., 2019; Krumhardt et al., 2020; Spring & Ilyina, 2020). 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 (Park et al., 2019; Li et al., 2019; Brady et al., 2020). Recent advances in ocean biogeochemical observing systems (e.g., Biogeochemical Argo (Argo, 2022)), 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) (Sharp et al., 2022; Keppler et al., 2022). 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 innovative 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 one 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 the Community Atmosphere Model version (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 (J. Moore et al., 2001; J. K. Moore et al., 2004, 2013; Long et al., 2021). CESM2 is well validated with available ocean observations and renalysis 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 CO_2 concentrations (Figure 1a). CESM2 SMYLE forecasts are initialized quarterly from February 1, 1970 to November 1, 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}K) are applied 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 the 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.

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. (2022) dissolved inorganic carbon product, and the Sharp et al. (2022) dissolved oxygen product. The Roemmich-Gilson product provides monthly temperature estimates for the upper ∼2000 m at 1° horizontal resolution over 2004 to present using Argo float data. Temperature and salinity data from the global fleet of Argo floats is 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., 2022) and

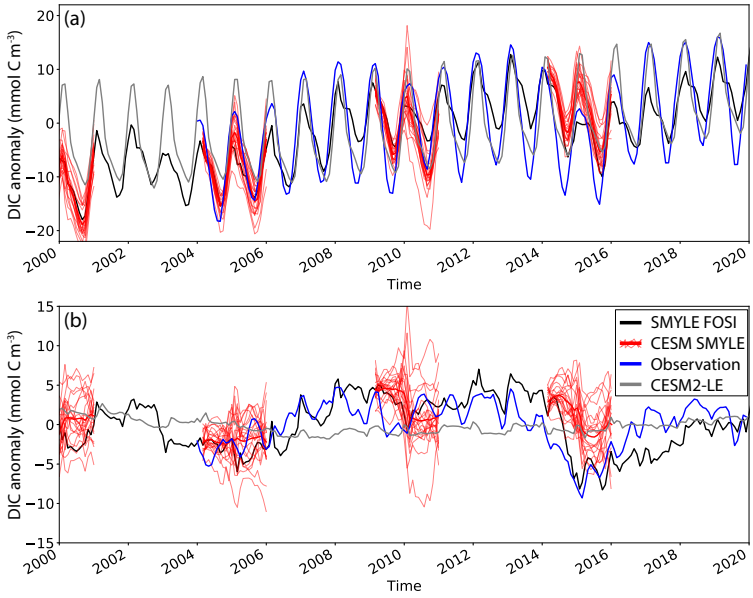

Figure 1. Temporal evolution of monthly surface ocean dissolved inorganic carbon anomalies (mmol C m^{-3}) averaged over the California Current Large Marine Ecosystem from 2000 to 2020 in (red line) four randomly-selected CESM SMYLE ensemble forecasts, (black line) SMYLE FOSI, (gray line) CESM2-LE, and (blue line) the observation-based product of Keppler et al. (2022). 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.

DO (2004-2022) (Sharp et al., 2022) 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-DIC) 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 1500 m of the global ocean ((Keppler et al., 2022); Figure 1). Similarly, Sharp et al. (2022) 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). Uncertainty estimates in MOBO-DIC include uncertainties stemming from the GLODAP measurements, the representation of measurements on a monthly 1° grid, and the prediction uncertainty of the method. The global mean total uncertainty in MOBO-DIC is 18 μmol kg^{-1}. Uncertainties in GOBAI-O2 are assessed in a similar fashion, with a global mean total uncertainty in of about 6 μmol kg^{-1}.

Observation-based datasets 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.

	$MOBO_{GLODAP}$	$GOBAI_{GLODAP}$	$GOBAI_{ARGO}$
Gulf of Alaska	195	232	1020
California Current	179	173	771
Kuroshio Current	292	641	490
Bay of Bengal	214	94	1096
Norwegian Sea	218	98	680
Greenland Sea	608	260	802

Table 1. The number of direct observations used in the creation of MOBO-DIC ($MOBO_{GLODAP}$) and GOBAI-O2 ($GOBAI_{GLODAP}$ and $GOBAI_{ARGO}$) 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. LMEs noted here are shown in Figure S1.

2.3 Statistical approach

We present forecast skill for three tracers - temperature, DIC, and - at two depths - the surface and below the mixed layer at 300 meters. 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 observations 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).

We remove long-term anthropogenic trends (1st 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 (Pearsons r-value; Figure S2) and mean absolute error (MAE) of the ensemble mean forecast and the observational product. MAE values for each tracer are normalized by the standard deviation of each tracer at each depth in a given region. The significance (95% Confidence Interval) of correlation coefficients 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). 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 correlation coefficient between the Nino3.4 index and the tracers of interest in SMYLE FOSI.

2.4 Model validation

We assess the ability of the model reconstruction 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 (Figure S3, S4). Seasonal climatologies are highly correlated for most of the U.N. fisheries regions (Figure S3). 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). 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), however we note a lack of high correlation in subsurface DIC and DO.

3 Results

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

CESM SMYLE exhibits high forecast skill for surface temperature, DIC, and DO one month after initialization in nearly all U.N. fisheries regions (Figure 2, 1^{st} column). Subsurface DIC forecast skill is notably absent one month 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 (Figure 3a-f). 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 (Figure 3s-x).

Persistence forecasts are more consistently skillful than initialized forecasts one month after initialization (Figure 2, 2^{nd} 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 (Figure 3g-l). Three tracer-depth combinations show particularly high persistence forecast skill: surface and subsurface DIC and subsurface DO, with anomaly correlation coefficients 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 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), 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 skill exceeds persistence, 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, S5, S6, S7). We find up to 10 months of high, persistence-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-exceeding forecast skill - this is a reflection of low forecast skill combined with particularly high subsurface persistence skill

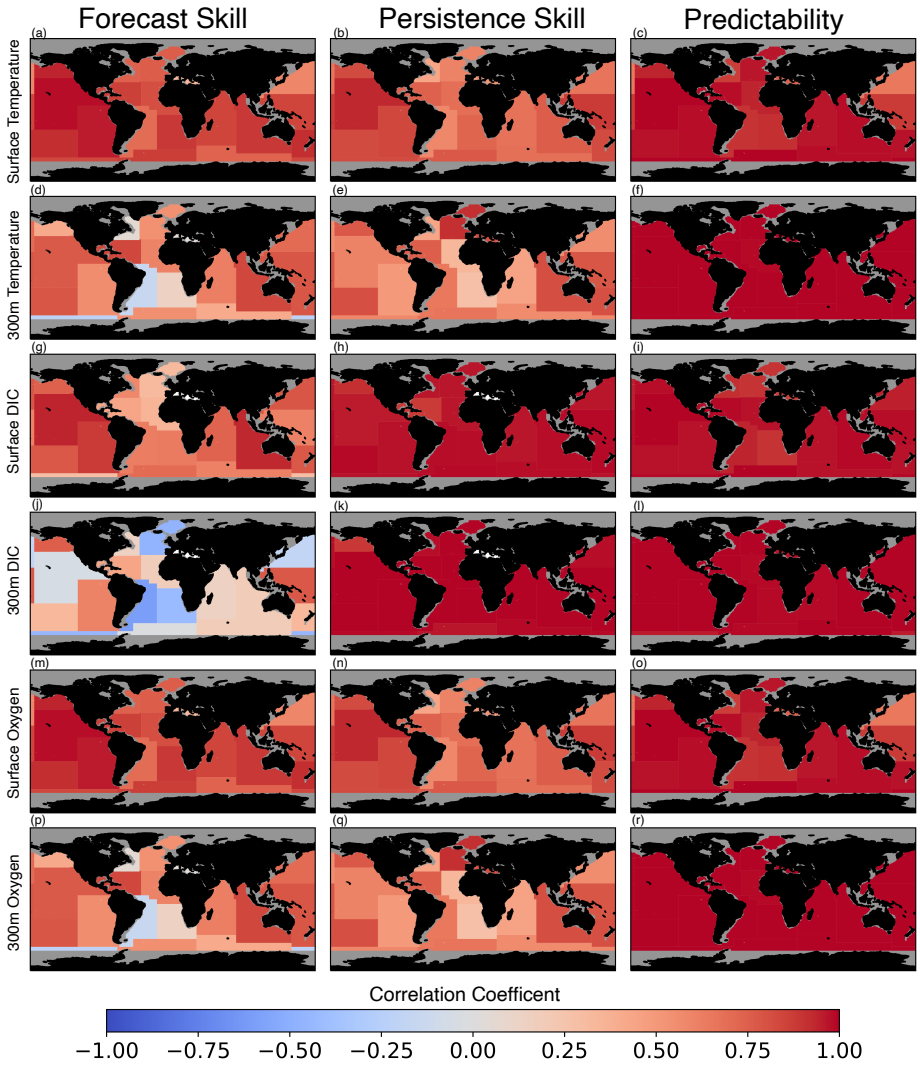

Figure 2. Various metrics of forecast skill one month following February forecast initialization in the U.N. F.A.O. fisheries regions: (1^{st} column) forecast skill, (2^{nd} column) persistence forecast skill, (3^{rd} column) predictability. Forecast skill is shown for three variables and two depth levels: (1^{st} row) surface temperature, (2^{nd} row) 300 m temperature,(3^{rd} row) surface DIC, (4^{th} row) 300 m DIC, (5^{th} row) surface DO concentration, and (6^{th} row) 300 m DO concentration.

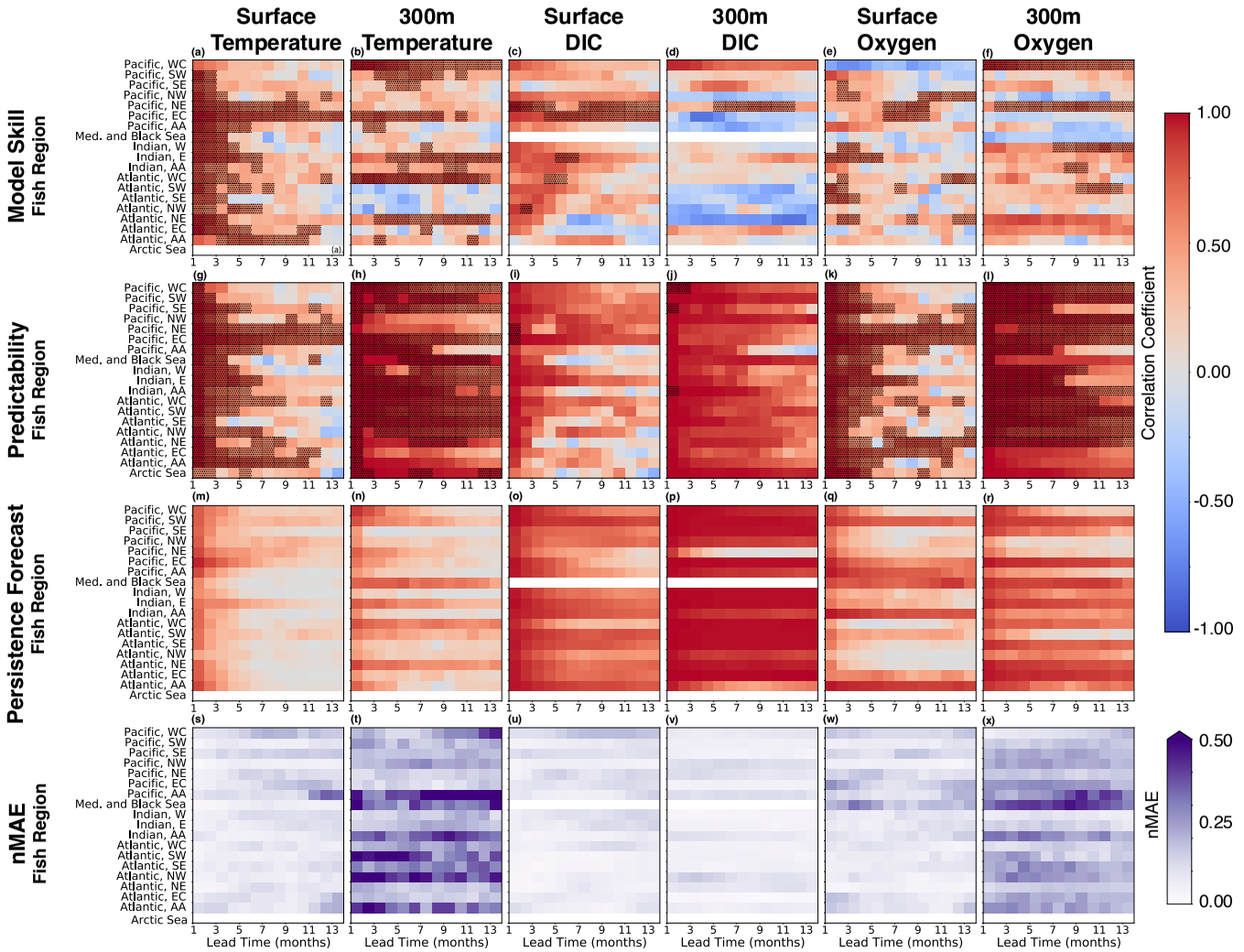

Figure 3. Various metrics of forecast skill over 13 lead months in the U.N. F.A.O. fisheries regions for February forecast initializations: (1^{st} row) forecast skill, (2^{nd} row) predictability forecast, (3^{rd} row) persistence forecast skill, and (4^{th} row) normalized mean absolute error of the forecast. Shading in the 1^{st} and 2^{nd} row indicate skillful forecasts (Model Forecast > Persistence forecast and $r > 0.5$). Forecast skill is shown for three variables and two depth levels: (1^{st} column) surface temperature, (2^{nd} column) 300 m temperature,(3^{rd} column) surface DIC, (4^{th} column) 300 m DIC, (5^{th} column) surface DO concentration, and (6^{th} column) 300 m DO concentration. Dots in the 1^{st}-3^{rd} rows indicated skillful forecast (r>0.5 and exceeds persistence forecast skill). The Arctic is not included in observational data products, and thus only appears in model predictability.

manuscript submitted to *Earth's Future*

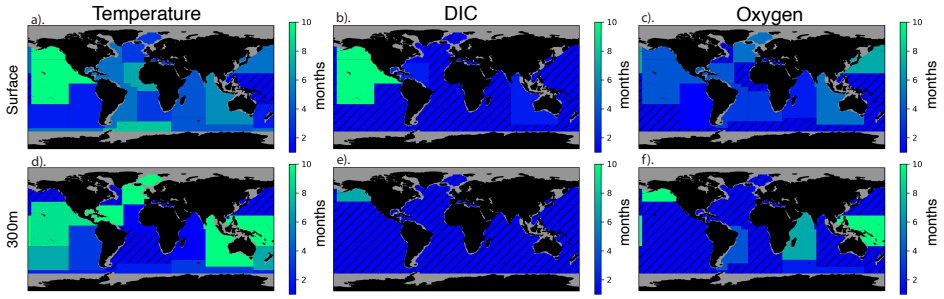

Figure 4. The total number of lead-months over which forecast skill is both high (r>0.5), and exceeds persistence forecast skill in the first 12 months following February model initialization in the U.N. F.A.O. fisheries regions for surface (top) and subsurface (bottom) variables. Hatching indicates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

(Figure 3d, 3p). The number of skillful months does not does demonstrate large variability with different months of initialization.

We find high, persistence-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 2) for up to 10 months (Figure S8). In contrast, surface and subsurface DIC potential predictability displays substantially lower lead-month counts (Figure S8) 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 (Supplemental Figure S8) 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 forecast skill in both observations and model reconstructions, as seen in Figure 3 i-j, where model forecast 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 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 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 300m, there are no skillful forecasts (excepting DIC persistence). The month of initialization has some impact on forecast skill in North Pacific LMEs, with certain

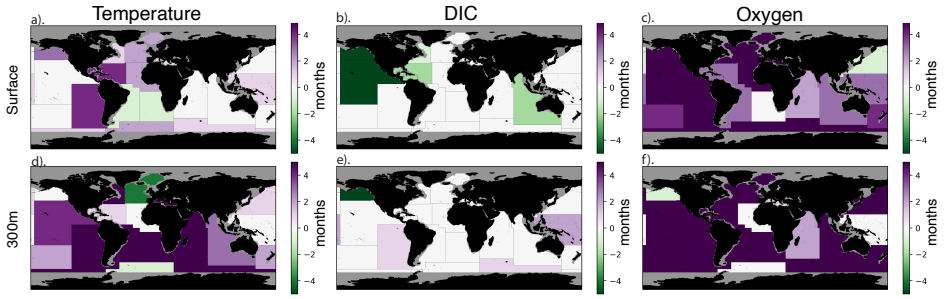

Figure 5. The potential gain in the number of months of forecast skill, estimated as the difference in the number of months for which predictability is high (r>0.5) and exceeds persistence (Supplemental Figure S8), and the number of months for which forecast skill is both high and exceeds persistence (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.

initializations outperforming others; this stands in contrast with the number of skillful months in Fisheries Regions, which was relatively insensitive to month of initialization (Figure 4).

We find relatively low forecast skill in three other observation-rich regions outside of the North Pacific (Figure S9). 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). 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); 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 et al., 2022))

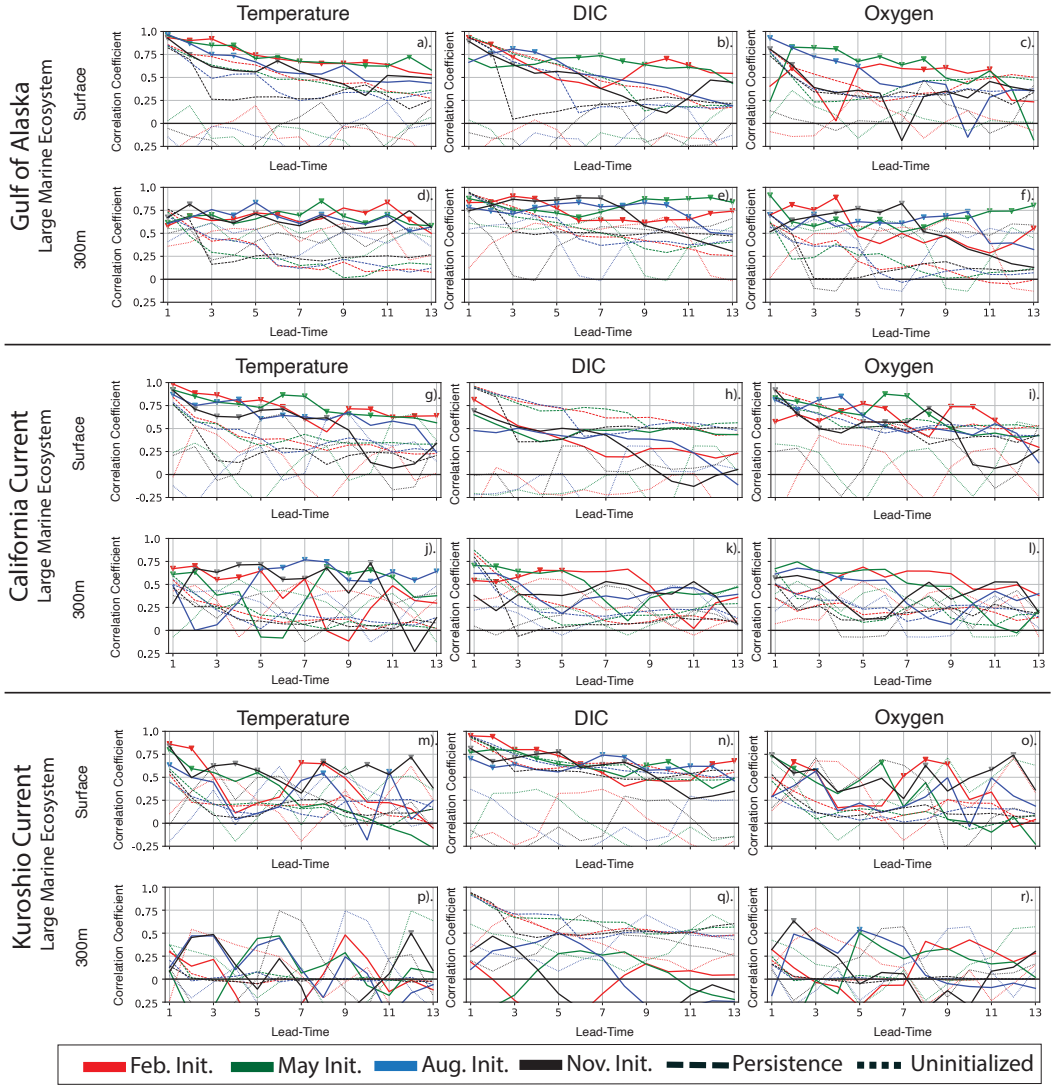

Figure 6. (solid) Forecast skill, (dashed) persistence forecast skill, and (dotted) uninitialized forecast skill in three North Pacific Large Marine Ecosystems for quarterly model intializations over 13 lead-months. Triangles indicate statistically significant skill at the 95% confidence interval.

	GOA	CalCS.	Kuroshio	Bay of Bengal	Greenland Sea	Norwegian Sea
Surface Temp.	0.42	0.48	0.08	0.40	0.16	0.21
300m 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
300m 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
300m Oxygen	0.61	0.22	0.16	0.05	0	0.1

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

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 skill is seemingly insensitive to month of initialization in large regions, SS7), but shows some sensitivity in smaller-scale LMEs.

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.

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.

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., it's 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 earth system odels (Verdy & Mazloff, 2017, e.g.,) will be key for the development of 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, the need for accurate forecasts of multiple marine stressors months to years in advance will grow. 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 (Payne et al., 2017; Tommasi et al., 2017; Jacox et al., 2020), 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 modelling 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 modelling experiments (Barnston et al., 2019; 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 $\sim$36% of variability.

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; 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 will likely 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 modelling systems can effectively predict heatwaves (Jacox et al., 2022), but has not examined the potential co-occuring bio-

geochemical 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. Our results highlight predictive skill in large scale fisheries regions globally, and smaller scale LMEs in the North Pacific. The continued development of ESM forecasting systems and novel observation-based products may allow for improved marine management in the coming decades.

Open Research Section

The CESM Seasonal to Multiyear Large Ensemble and SMYLE FOSI are available at: https://www.earthsystemgrid.org/dataset/ucar.cgd.cesm2.smyle.html. The CESM2 Large Ensemble data are available at: https://www.earthsystemgrid.org/dataset/ucar.cgd.cesm2le.output.html

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. GOBAI-O2 data are available at https://doi.org/10.25921/z72m-yz67. A preprint for the MOBO-DIC summary paper can be found at DOI: essoar.167160635.51342340/v1.

Shapefiles for Large Marine Ecosystems were found at https://www.sciencebase.gov/catalog/item/55c77722e4b08400b1fd8244. Shapefiles for U.N. FAO Fisheries Regions were found at https://www.fao.org/fishery/en/area/search

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-ESPWG/ESP-Lab.

References

Albers, J. R., & Newman, M. (2021, April). Subseasonal predictability of the North Atlantic Oscillation. *Environmental Research Letters*, *16*(4), 044024. Retrieved 2023-02-06, from `https://iopscience.iop.org/article/10.1088/1748-9326/abe781` doi: 10.1088/1748-9326/abe781

Argo. (2022). *Argo float data and metadata from Global Data Assembly Centre (Argo GDAC).* SEANOE. Retrieved 2023-01-31, from `https://www.seanoe.org/data/00311/42182/` (Type: dataset) doi: 10.17882/42182

Ban, Z., Hu, X., & Li, J. (2022, October). Tipping points of marine phytoplankton to multiple environmental stressors. *Nature Climate Change*. Retrieved 2022-10-27, from `https://www.nature.com/articles/s41558-022-01489-0` doi:

10.1038/s41558-022-01489-0

Barnston, A. G., Tippett, M. K., Ranganathan, M., & L'Heureux, M. L. (2019, December). Deterministic skill of ENSO predictions from the North American Multimodel Ensemble. *Climate Dynamics*, *53*(12), 7215–7234. Retrieved 2023-02-03, from http://link.springer.com/10.1007/s00382-017-3603-3 doi: 10.1007/s00382-017-3603-3

Bednaršek, N., Feely, R. A., Beck, M. W., Glippa, O., Kanerva, M., & Engström-Öst, J. (2018, December). 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. Retrieved 2022-10-27, from https://www.frontiersin.org/article/10.3389/fmars.2018.00486/full doi: 10.3389/fmars.2018.00486

Bednaršek, N., Harvey, C. J., Kaplan, I. C., Feely, R. A., & Možina, J. (2016, June). Pteropods on the edge: Cumulative effects of ocean acidification, warming, and deoxygenation. *Progress in Oceanography*, *145*, 1–24. Retrieved 2022-10-27, from https://linkinghub.elsevier.com/retrieve/pii/S0079661115300112 doi: 10.1016/j.pocean.2016.04.002

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., ... Vichi, M. (2013, October). Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. *Biogeosciences*, *10*(10), 6225–6245. Retrieved 2022-10-27, from https://bg.copernicus.org/articles/10/6225/2013/ doi: 10.5194/bg-10-6225-2013

Brady, R. X., Lovenduski, N. S., Yeager, S. G., Long, M. C., & Lindsay, K. (2020, December). Skillful multiyear predictions of ocean acidification in the California Current System. *Nature Communications*, *11*(1), 2166. Retrieved 2022-10-27, from http://www.nature.com/articles/s41467-020-15722-x doi: 10.1038/s41467-020-15722-x

Burger, F. A., John, J. G., & Frölicher, T. L. (2020, September). Increase in ocean acidity variability and extremes under increasing atmospheric CO₂. *Biogeosciences*, *17*(18), 4633–4662. Retrieved 2022-10-27, from https://bg.copernicus.org/articles/17/4633/2020/ doi: 10.5194/bg-17-4633-2020

Capotondi, A., Sardeshmukh, P. D., Di Lorenzo, E., Subramanian, A. C., & Miller, A. J. (2019, December). Predictability of US West Coast Ocean Temperatures is not solely due to ENSO. *Scientific Reports*, *9*(1), 10993. Retrieved 2022-12-14, from http://www.nature.com/articles/s41598-019-47400-4 doi: 10.1038/s41598-019-47400-4

Chapman, W. E., Subramanian, A. C., Delle Monache, L., Xie, S. P., & Ralph, F. M. (2019, September). Improving Atmospheric River Forecasts With Machine Learning. *Geophysical Research Letters*, *46*(17-18), 10627–10635. Retrieved 2023-02-06, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2019GL083662 doi: 10.1029/2019GL083662

Cheung, W. W. L., Wei, C.-L., & Levin, L. A. (2022, October). Vulnerability of exploited deep-sea demersal species to ocean warming, deoxygenation, and acidification. *Environmental Biology of Fishes*, *105*(10), 1301–1315. Retrieved 2022-10-27, from https://link.springer.com/10.1007/s10641-022-01321-w doi: 10.1007/s10641-022-01321-w

Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., ... Strand, W. G. (2020, February). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2). Retrieved 2022-10-27, from https://onlinelibrary.wiley.com/doi/10.1029/2019MS001916 doi: 10.1029/2019MS001916

Di Lorenzo, E., & Mantua, N. (2016, November). Multi-year persistence of the 2014/15 North Pacific marine heatwave. *Nature Climate Change*, *6*(11),

1042–1047. Retrieved 2022-10-27, from http://www.nature.com/articles/nclimate3082 doi: 10.1038/nclimate3082

Doney, S. C., Fabry, V. J., Feely, R. A., & Kleypas, J. A. (2009, January). Ocean Acidification: The Other CO 2 Problem. *Annual Review of Marine Science*, *1*(1), 169–192. Retrieved 2022-10-27, from https://www.annualreviews.org/doi/10.1146/annurev.marine.010908.163834 doi: 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*, 10.

Fennel, K., Alin, S., Barbero, L., Evans, W., Bourgeois, T., Cooley, S., ... Wang, Z. A. (2019, March). Carbon cycling in the North American coastal ocean: a synthesis. *Biogeosciences*, *16*(6), 1281–1304. Retrieved 2023-02-03, from https://bg.copernicus.org/articles/16/1281/2019/ doi: 10.5194/bg-16-1281-2019

Gettelman, A., Geer, A. J., Forbes, R. M., Carmichael, G. R., Feingold, G., Posselt, D. J., ... Zuidema, P. (2022, April). The future of Earth system prediction: Advances in model-data fusion. *Science Advances*, *8*(14), eabn3488. Retrieved 2023-02-06, from https://www.science.org/doi/10.1126/sciadv.abn3488 doi: 10.1126/sciadv.abn3488

Gloege, L., Yan, M., Zheng, T., & McKinley, G. A. (2022, February). Improved Quantification of Ocean Carbon Uptake by Using Machine Learning to Merge Global Models and pCO 2 Data. *Journal of Advances in Modeling Earth Systems*, *14*(2). Retrieved 2023-02-06, from https://onlinelibrary.wiley.com/doi/10.1029/2021MS002620 doi: 10.1029/2021MS002620

Greene, C. H., Scott-Buechler, C. M., Hausner, A. L. P., Johnson, Z. I., Lei, X. G., & Huntley, M. E. (2022). A CIRCULAR ECONOMY APPROACH. *Oceanography*, 9.

Gruber, N. (2011, May). Warming up, turning sour, losing breath: ocean biogeochemistry under global change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1943), 1980–1996. Retrieved 2023-02-21, from https://royalsocietypublishing.org/doi/10.1098/rsta.2011.0003 doi: 10.1098/rsta.2011.0003

Hervieux, G., Alexander, M. A., Stock, C. A., Jacox, M. G., Pegion, K., Becker, E., ... Tommasi, D. (2019, December). More reliable coastal SST forecasts from the North American multimodel ensemble. *Climate Dynamics*, *53*(12), 7153–7168. Retrieved 2022-10-27, from http://link.springer.com/10.1007/s00382-017-3652-7 doi: 10.1007/s00382-017-3652-7

Ilyina, T., Li, H., Spring, A., Müller, W. A., Bopp, L., Chikamoto, M. O., ... Yeager, S. (2021, March). Predictable Variations of the Carbon Sinks and Atmospheric CO 2 Growth in a Multi-Model Framework. *Geophysical Research Letters*, *48*(6). Retrieved 2022-10-27, from https://onlinelibrary.wiley.com/doi/10.1029/2020GL090695 doi: 10.1029/2020GL090695

Jacox, M. G., Alexander, M. A., Amaya, D., Becker, E., Bograd, S. J., Brodie, S., ... Tommasi, D. (2022, April). Global seasonal forecasts of marine heatwaves. *Nature*, *604*(7906), 486–490. Retrieved 2022-12-14, from https://www.nature.com/articles/s41586-022-04573-9 doi: 10.1038/s41586-022-04573-9

Jacox, M. G., Alexander, M. A., Siedlecki, S., Chen, K., Kwon, Y.-O., Brodie, S., ... Rykaczewski, R. (2020, April). Seasonal-to-interannual prediction of North American coastal marine ecosystems: Forecast methods, mechanisms of predictability, and priority developments. *Progress in Oceanography*, *183*, 102307. Retrieved 2022-10-27, from https://linkinghub.elsevier.com/retrieve/pii/S007966112030046X doi: 10.1016/j.pocean.2020.102307

Jacox, M. G., Alexander, M. A., Stock, C. A., & Hervieux, G. (2019, December). 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. Retrieved 2022-10-27, from http://link.springer.com/10.1007/s00382-017-3608-y doi: 10.1007/s00382-017-3608-y

Keppler, L., Landschützer, P., Lauvset, S., & Gruber, N. (2022, June). Recent trends and variability in the oceanic storage of dissolved inorganic carbon. *ESS Open Archive*. doi: 10.22541/essoar.167160635.51342340/v1

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., ... Takahashi, K. (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5–48. Retrieved 2023-01-30, from https://www.jstage.jst.go.jp/article/jmsj/93/1/93_2015-001/_article doi: 10.2151/jmsj.2015-001

Krumhardt, K. M., Lovenduski, N. S., Long, M. C., Luo, J. Y., Lindsay, K., Yeager, S., & Harrison, C. (2020, June). Potential Predictability of Net Primary Production in the Ocean. *Global Biogeochemical Cycles*, *34*(6). Retrieved 2022-12-14, from https://onlinelibrary.wiley.com/doi/10.1029/2020GB006531 doi: 10.1029/2020GB006531

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., ... Ziehn, T. (2020, July). Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. *Biogeosciences*, *17*(13), 3439–3470. Retrieved 2022-10-27, from https://bg.copernicus.org/articles/17/3439/2020/ doi: 10.5194/bg-17-3439-2020

Levin, L. A. (2018, January). Manifestation, Drivers, and Emergence of Open Ocean Deoxygenation. *Annual Review of Marine Science*, *10*(1), 229–260. Retrieved 2022-10-27, from https://www.annualreviews.org/doi/10.1146/annurev-marine-121916-063359 doi: 10.1146/annurev-marine-121916-063359

Li, H., Ilyina, T., Müller, W. A., & Landschützer, P. (2019, April). Predicting the variable ocean carbon sink. *Science Advances*, *5*(4), eaav6471. Retrieved 2022-10-27, from https://www.science.org/doi/10.1126/sciadv.aav6471 doi: 10.1126/sciadv.aav6471

Long, M. C., Moore, J. K., Lindsay, K., Levy, M., Doney, S. C., Luo, J. Y., ... Sylvester, Z. T. (2021, December). Simulations With the Marine Biogeochemistry Library (MARBL). *Journal of Advances in Modeling Earth Systems*, *13*(12). Retrieved 2023-01-30, from https://onlinelibrary.wiley.com/doi/10.1029/2021MS002647 doi: 10.1029/2021MS002647

Lovenduski, N. S., Yeager, S. G., Lindsay, K., & Long, M. C. (2019, January). Predicting near-term variability in ocean carbon uptake. *Earth System Dynamics*, *10*(1), 45–57. Retrieved 2022-10-27, from https://esd.copernicus.org/articles/10/45/2019/ doi: 10.5194/esd-10-45-2019

Merryfield, W. J., Baehr, J., Batté, L., Becker, E. J., Butler, A. H., Coelho, C. A. S., ... Yeager, S. (2020, June). Current and Emerging Developments in Subseasonal to Decadal Prediction. *Bulletin of the American Meteorological Society*, *101*(6), E869–E896. Retrieved 2022-10-27, from https://journals.ametsoc.org/view/journals/bams/101/6/bamsD190037.xml doi: 10.1175/BAMS-D-19-0037.1

Mills, K., Pershing, A., Brown, C., Chen, Y., Chiang, F.-S., Holland, D., ... Wahle, R. (2013). Fisheries Management in a Changing Climate: Lessons From the 2012 Ocean Heat Wave in the Northwest Atlantic. *Oceanography*, *26*(2). Retrieved 2022-10-27, from https://tos.org/oceanography/article/fisheries-management-in-a-changing-climate-lessonsfrom-the-2012-ocean-heat- doi: 10.5670/oceanog.2013.27

Mogen, S. C., Lovenduski, N. S., Dallmann, A. R., Gregor, L., Sutton, A. J., Bograd, S. J., ... Yeager, S. (2022, May). Ocean Biogeochemical Signatures of the North Pacific Blob. *Geophysical Research Letters*, *49*(9). Retrieved 2022-10-27, from https://onlinelibrary.wiley.com/doi/10.1029/2021GL096938

doi: 10.1029/2021GL096938

Moore, Morley, J. W., Morrison, B., Kolian, M., Horsch, E., Frölicher, T., ... Griffis, R. (2021, February). Estimating the economic impacts of climate change on 16 major us fisheries. *Climate Change Economics*, *12*(01), 2150002. Retrieved 2022-10-27, from https://www.worldscientific.com/doi/10.1142/S2010007821500020 doi: 10.1142/S2010007821500020

Moore, J., Doney, S. C., Kleypas, J. A., Glover, D. M., & Fung, I. Y. (2001, January). An intermediate complexity marine ecosystem model for the global domain. *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*(1-3), 403–462. Retrieved 2022-10-27, from https://linkinghub.elsevier.com/retrieve/pii/S0967064501001084 doi: 10.1016/S0967-0645(01)00108-4

Moore, J. K., Doney, S. C., & Lindsay, K. (2004, December). Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model: GLOBAL ECOSYSTEM-BIOGEOCHEMICAL MODEL. *Global Biogeochemical Cycles*, *18*(4), n/a–n/a. Retrieved 2022-10-27, from http://doi.wiley.com/10.1029/2004GB002220 doi: 10.1029/2004GB002220

Moore, J. K., Lindsay, K., Doney, S. C., Long, M. C., & Misumi, K. (2013, December). 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. Retrieved 2022-10-27, from https://journals.ametsoc.org/doi/10.1175/JCLI-D-12-00566.1 doi: 10.1175/JCLI-D-12-00566.1

Park, J.-Y., Stock, C. A., Dunne, J. P., Yang, X., & Rosati, A. (2019, July). Seasonal to multiannual marine ecosystem prediction with a global Earth system model. *Science*, *365*(6450), 284–288. Retrieved 2022-10-27, from https://www.science.org/doi/10.1126/science.aav6634 doi: 10.1126/science.aav6634

Payne, M. R., Danabasoglu, G., Keenlyside, N., Matei, D., Miesner, A. K., Yang, S., & Yeager, S. G. (2022, December). Skilful decadal-scale prediction of fish habitat and distribution shifts. *Nature Communications*, *13*(1), 2660. Retrieved 2022-10-27, from https://www.nature.com/articles/s41467-022-30280-0 doi: 10.1038/s41467-022-30280-0

Payne, M. R., Hobday, A. J., MacKenzie, B. R., Tommasi, D., Dempsey, D. P., Fässler, S. M. M., ... Villarino, E. (2017, September). Lessons from the First Generation of Marine Ecological Forecast Products. *Frontiers in Marine Science*, *4*, 289. Retrieved 2023-02-23, from http://journal.frontiersin.org/article/10.3389/fmars.2017.00289/full doi: 10.3389/fmars.2017.00289

Pörtner, H.-O. (2010, March). 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. Retrieved 2023-02-06, from https://journals.biologists.com/jeb/article/213/6/881/10149/Oxygen-and-capacity-limitation-of-thermal doi: 10.1242/jeb.037523

Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., ... Yeager, S. G. (2021, December). Ubiquity of human-induced changes in climate variability. *Earth System Dynamics*, *12*(4), 1393–1411. Retrieved 2022-12-14, from https://esd.copernicus.org/articles/12/1393/2021/ doi: 10.5194/esd-12-1393-2021

Roemmich, D., Church, J., Gilson, J., Monselesan, D., Sutton, P., & Wijffels, S. (2015, March). Unabated planetary warming and its ocean structure since 2006. *Nature Climate Change*, *5*(3), 240–245. Retrieved 2023-01-31, from http://www.nature.com/articles/nclimate2513 doi: 10.1038/nclimate2513

Roemmich, D., & Gilson, J. (2009, August). 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. Retrieved 2023-01-30, from

https://linkinghub.elsevier.com/retrieve/pii/S0079661109000160 doi: 10.1016/j.pocean.2009.03.004

Shao, Q., Li, W., Han, G., Hou, G., Liu, S., Gong, Y., & Qu, P. (2021, July). A Deep Learning Model for Forecasting Sea Surface Height Anomalies and Temperatures in the South China Sea. *Journal of Geophysical Research: Oceans*, *126*(7). Retrieved 2023-02-06, from https://onlinelibrary.wiley.com/doi/10.1029/2021JC017515 doi: 10.1029/2021JC017515

Sharp, J. D., Fassbender, A. J., Carter, B. R., Johnson, G. C., Schultz, C., & Dunne, J. P. (2022, September). *GOBAI-O₂: temporally and spatially resolved fields of ocean interior dissolved oxygen over nearly two decades* (preprint). ESSD – Ocean/Chemical oceanography. Retrieved 2023-01-20, from https://essd.copernicus.org/preprints/essd-2022-308/ doi: 10.5194/essd-2022-308

Shin, & Newman, M. (2021, May). Seasonal Predictability of Global and North American Coastal Sea Surface Temperature and Height Anomalies. *Geophysical Research Letters*, *48*(10). Retrieved 2023-02-06, from https://onlinelibrary.wiley.com/doi/10.1029/2020GL091886 doi: 10.1029/2020GL091886

Shin, Park, S., Shin, S., Newman, M., & Alexander, M. A. (2020, January). Enhancing ENSO Prediction Skill by Combining Model-Analog and Linear Inverse Models (MA-LIM). *Geophysical Research Letters*, *47*(1). Retrieved 2023-02-03, from https://onlinelibrary.wiley.com/doi/10.1029/2019GL085914 doi: 10.1029/2019GL085914

Siedlecki, S. A., Kaplan, I. C., Hermann, A. J., Nguyen, T. T., Bond, N. A., Newton, J. A., ... Feely, R. A. (2016, June). Experiments with Seasonal Forecasts of ocean conditions for the Northern region of the California Current upwelling system. *Scientific Reports*, *6*(1), 27203. Retrieved 2023-02-03, from https://www.nature.com/articles/srep27203 doi: 10.1038/srep27203

Spring, A., Dunkl, I., Li, H., Brovkin, V., & Ilyina, T. (2021, November). Trivial improvements in predictive skill due to direct reconstruction of the global carbon cycle. *Earth System Dynamics*, *12*(4), 1139–1167. Retrieved 2023-01-27, from https://esd.copernicus.org/articles/12/1139/2021/ doi: 10.5194/esd-12-1139-2021

Spring, A., & Ilyina, T. (2020, May). Predictability Horizons in the Global Carbon Cycle Inferred From a Perfect-Model Framework. *Geophysical Research Letters*, *47*(9). Retrieved 2023-01-27, from https://onlinelibrary.wiley.com/doi/10.1029/2019GL085311 doi: 10.1029/2019GL085311

The State of World Fisheries and Aquaculture 2022. (2022). FAO. Retrieved 2022-10-27, from http://www.fao.org/documents/card/en/c/cc0461en doi: 10.4060/cc0461en

Tommasi, D., Stock, C. A., Hobday, A. J., Methot, R., Kaplan, I. C., Eveson, J. P., ... Werner, F. E. (2017, March). Managing living marine resources in a dynamic environment: The role of seasonal to decadal climate forecasts. *Progress in Oceanography*, *152*, 15–49. Retrieved 2022-12-14, from https://linkinghub.elsevier.com/retrieve/pii/S0079661116301586 doi: 10.1016/j.pocean.2016.12.011

Verdy, A., & Mazloff, M. R. (2017, September). A data assimilating model for estimating southern ocean biogeochemistry. *Journal of Geophysical Research: Oceans*, *122*(9), 6968–6988. Retrieved 2023-02-07, from https://onlinelibrary.wiley.com/doi/10.1002/2016JC012650 doi: 10.1002/2016JC012650

Whitney, M. M. (2022, September). Observed and projected global warming pressure on coastal hypoxia. *Biogeosciences*, *19*(18), 4479–4497. Retrieved 2022-10-27, from https://bg.copernicus.org/articles/19/4479/2022/ doi: 10.5194/bg-19-4479-2022

Wolff, S., O'Donncha, F., & Chen, B. (2020, August). Statistical and machine learning ensemble modelling to forecast sea surface temperature. *Journal of Marine Systems*, *208*, 103347. Retrieved 2023-02-06, from https://linkinghub.elsevier.com/retrieve/pii/S0924796320300439 doi: 10.1016/j.jmarsys.2020.103347

Yeager, S. G., Rosenbloom, N., Glanville, A. A., Wu, X., Simpson, I., Li, H., ... King, T. (2022, August). The Seasonal-to-Multiyear Large Ensemble (SMYLE) prediction system using the Community Earth System Model version 2. *Geoscientific Model Development*, *15*(16), 6451–6493. Retrieved 2022-10-27, from https://gmd.copernicus.org/articles/15/6451/2022/ doi: 10.5194/gmd-15-6451-2022

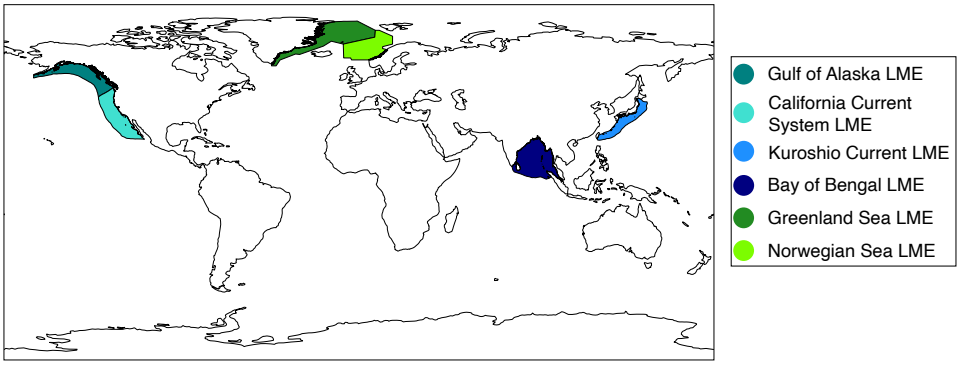

Figure S1. Locations of key Large Marine Ecosystems of interest.

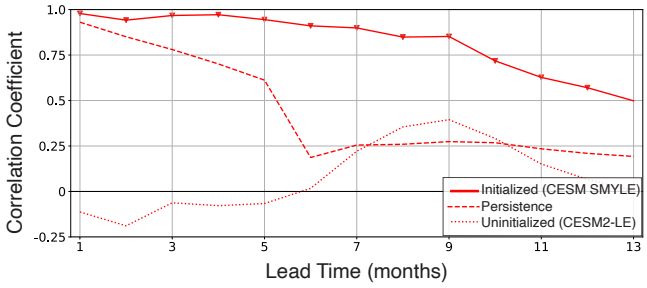

Figure S2. Forecast skill (anomaly correlation coefficient) as a function of forecast lead time (months) for sea surface temperature in the Pacific, Eastern Central U.N. F.A.O. Fisheries region for 13 months following initialization. Triangles indicate statistically significant forecast skill.

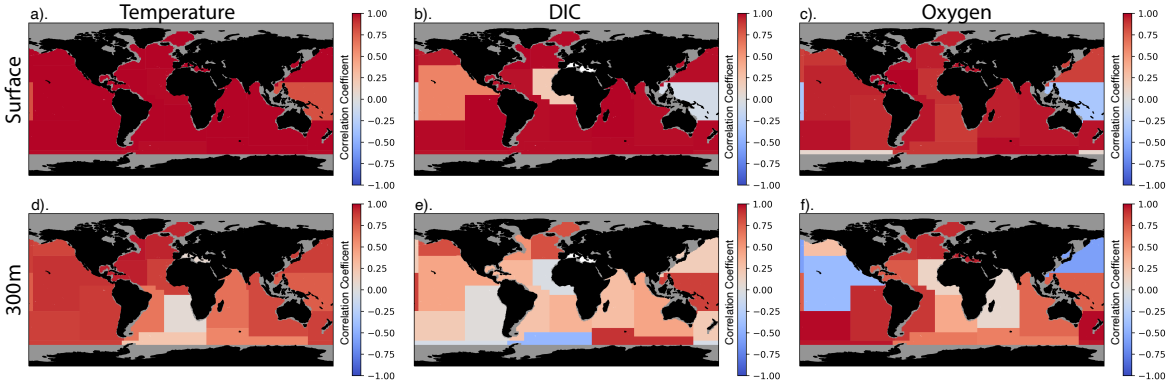

Figure S3. Correlation of the monthly seasonal cycle averaged in the observational products and SMYLE FOSI for (a,d) surface and 300 m temperature (2004-2021), (b,e) surface and 300 m DIC (2004-2019), and (c, f) surface and 300 m oxygen (2004-2021) in the U.N. F.A.O. fisheries regions.

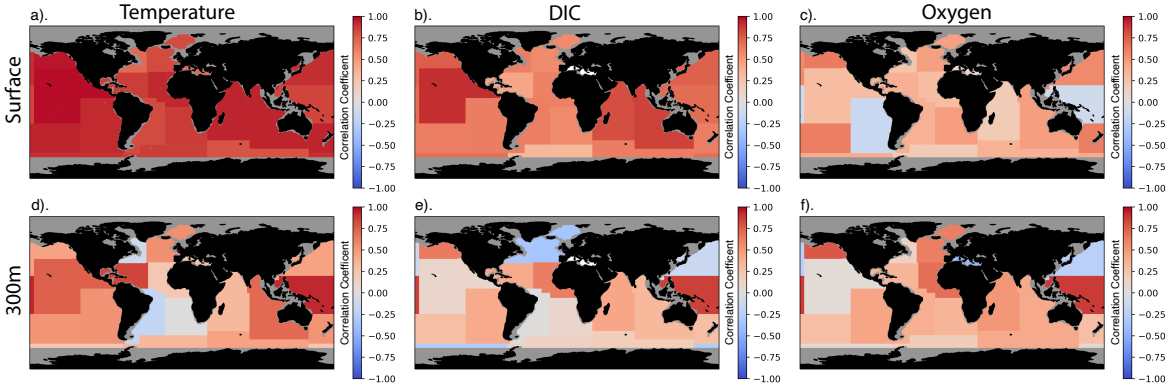

Figure S4. Correlation of the detrended, deseasoned anomalies averaged in the observational products and SMYLE FOSI for (a,d) surface and 300 m temperature (2004-2021), (b,e) surface and 300 m DIC (2004-2019), and (c, f) surface and 300 m oxygen (2004-2021) in the U.N. F.A.O. fisheries regions.

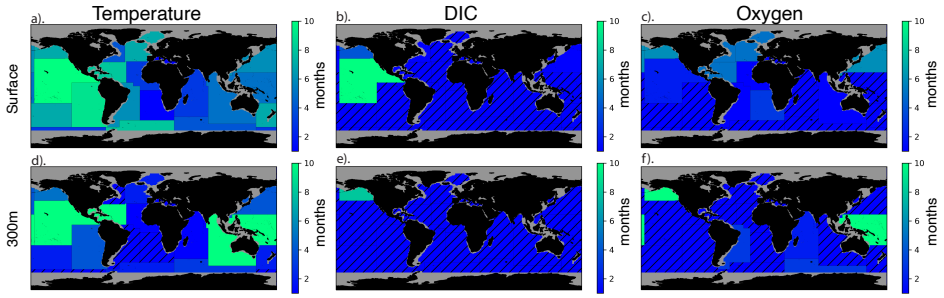

Figure S5. Same as Figure 4, but for May initialization.

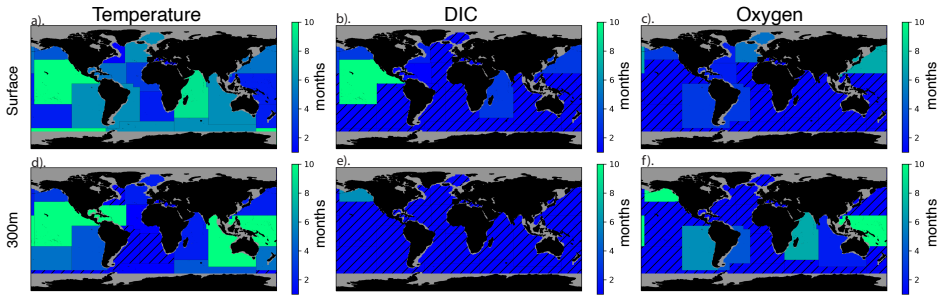

Figure S6. Same as Figure 4, but for August initialization.

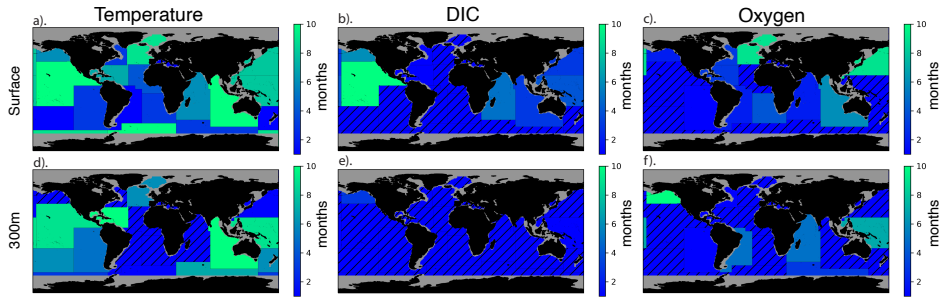

Figure S7. Same as Figure 4, but for November initialization.

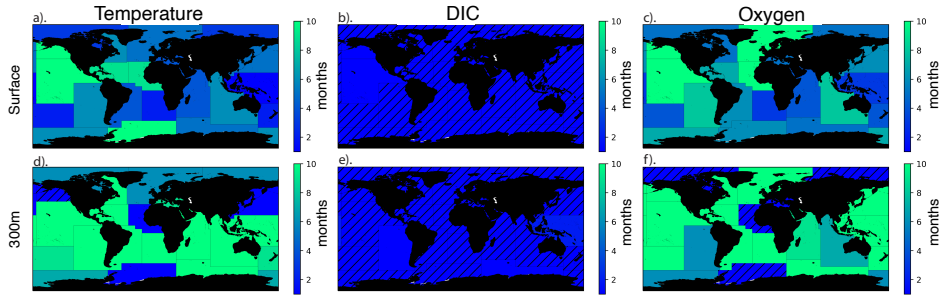

Figure S8. Same as Figure 4, but for predictability.

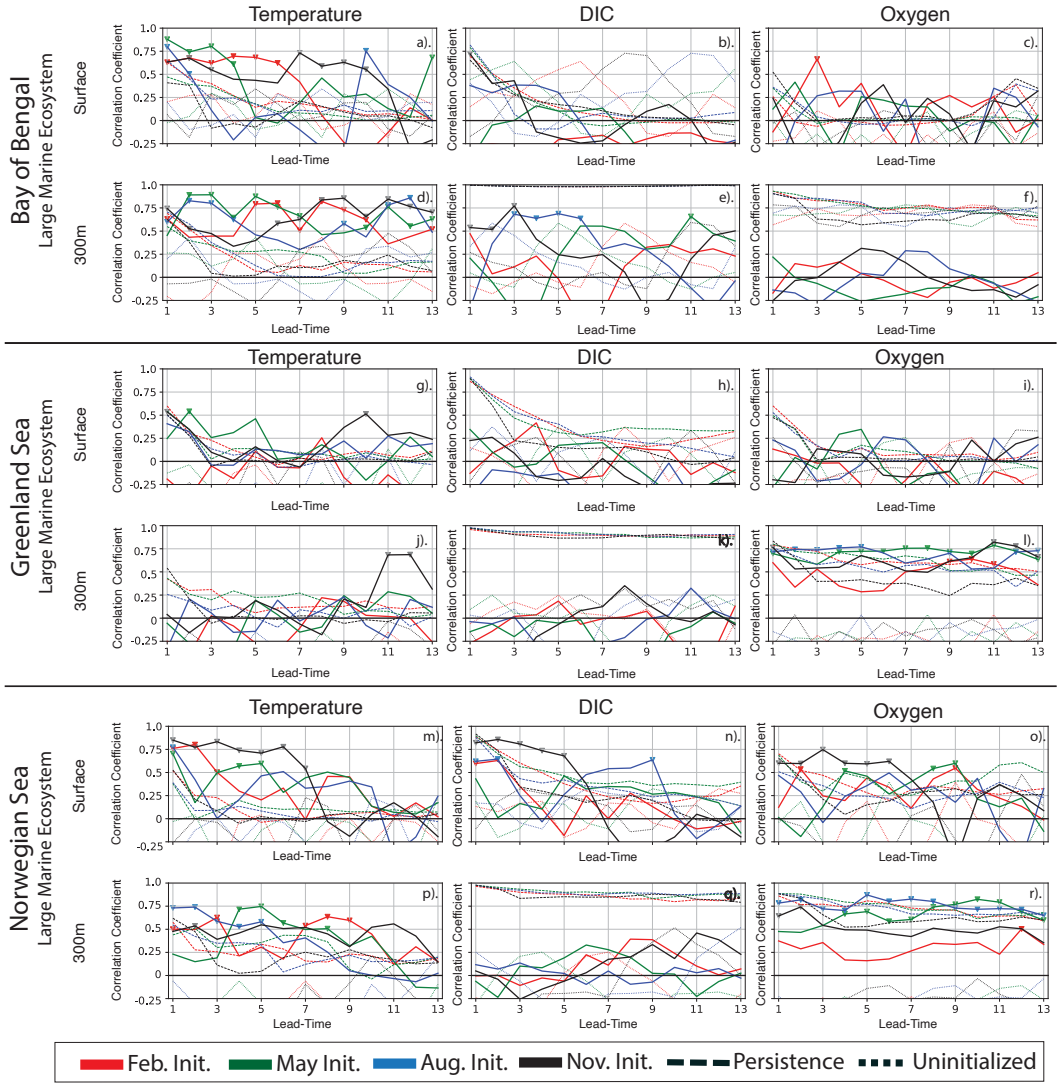

Figure S9. (solid) Forecast skill, (dashed) persistence forecast skill, and (dotted) uninitialized forecast skill in four global Large Marine Ecosystems for quarterly model intializations over 13 lead-months. Triangles indicate statistically significant skill at the 95% confidence interval.

Surface Temp.	GOA	CalCS.	Kuroshio	Bay of Bengal	Greenland Sea	Norwegian Sea
0	0.42	0.48	0.08	0.40	0.16	0.21
1	0.41	0.48	0.04	0.36	0.14	0.21
2	0.41	0.49	0	0.29	0.14	0.21
3	0.4	0.49	0.03	0.21	0.14	0.21
4	0.41	0.5	0.05	0.14	0.14	0.21
5	0.41	0.51	0.08	0.08	0.13	0.19
6	0.42	0.5	0.09	0.04	0.12	0.17
7	0.43	0.49	0.08	0.01	0.09	0.16
300m Temp.	GOA	CalCS.	Kuroshio	Bay of Bengal	Greenland Sea	Norwegian Sea
0	0.42	0.48	0.08	0.4	0.16	0.21
1	0.41	0.48	0.04	0.36	0.14	0.21
2	0.41	0.49	0	0.29	0.14	0.21
3	0.4	0.49	0.03	0.21	0.14	0.21
4	0.41	0.5	0.05	0.14	0.14	0.21
5	0.41	0.51	0.08	0.08	0.13	0.19
6	0.42	0.5	0.09	0.04	0.12	0.17
7	0.43	0.49	0.08	0.01	0.09	0.16
Surface DIC	GOA	CalCS.	Kuroshio	Bay of Bengal	Greenland Sea	Norwegian Sea
0	0.44	0.16	0.32	0.17	0.19	0.06
1	0.43	0.16	0.32	0.14	0.22	0.06
2	0.41	0.16	0.32	0.1	0.25	0.06
3	0.41	0.16	0.32	0.04	0.28	0.06
4	0.4	0.16	0.32	0.03	0.3	0.06
5	0.41	0.17	0.33	0.1	0.31	0.06
6	0.42	0.18	0.32	0.16	0.31	0.07
7	0.43	0.19	0.32	0.2	0.29	0.07
300m DIC	GOA	CalCS.	Kuroshio	Bay of Bengal	Greenland Sea	Norwegian Sea
0	0.6	0.28	0.17	0.25	0.18	0.13
1	0.56	0.27	0.16	0.25	0.18	0.13
2	0.51	0.27	0.14	0.25	0.18	0.14
3	0.46	0.26	0.13	0.25	0.18	0.15
4	0.4	0.24	0.13	0.25	0.18	0.16
5	0.35	0.21	0.12	0.24	0.18	0.17
6	0.31	0.18	0.12	0.24	0.18	0.17
7	0.28	0.15	0.12	0.23	0.18	0.17
Surface Oxygen	GOA	CalCS.	Kuroshio	Bay of Bengal	Greenland Sea	Norwegian Sea
0	0.3	0.41	0.04	0.07	0.12	0.26
1	0.32	0.4	0.01	0.1	0.11	0.25
2	0.34	0.4	0.03	0.13	0.11	0.23
3	0.35	0.39	0.04	0.14	0.13	0.2
4	0.36	0.38	0.06	0.14	0.14	0.17
5	0.36	0.38	0.08	0.12	0.15	0.14
6	0.35	0.36	0.08	0.1	0.14	0.12
7	0.33	0.34	0.06	0.07	0.12	0.1
300m Oxygen	GOA	CalCS.	Kuroshio	Bay of Bengal	Greenland Sea	Norwegian Sea
0	0.61	0.22	0.16	0.05	0	0.1
1	0.57	0.24	0.16	0.02	0	0.11
2	0.51	0.26	0.15	0	0	0.11
3	0.45	0.28	0.16	0.03	0.01	0.11
4	0.39	0.3	0.15	0.06	0.02	0.1
5	0.32	0.31	0.13	0.08	0.02	0.09
6	0.27	0.31	0.12	0.11	0.03	0.08
7	0.22	0.3	0.1	0.14 0.04	0.06	

Table S1. 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, when Niño leads by 0-7 months.