Progress In Oceanography
January 2023, Volume 210 Pages 102948 (31p.)
https://doi.org/10.1016/j.pocean.2022.102948
https://archimer.ifremer.fr/doc/00928/104016/

Archimer
https://archimer.ifremer.fr

A high-resolution ocean bottom temperature product for the northeast US continental shelf marine ecosystem

du Pontavice Hubert [1, 2, *], Chen Zhuomin [3], Saba Vincent S. [2]

[1] Princeton Univ, Atmospher & Ocean Sci Program, 300 Forrestal Rd,Sayre Hall, Princeton, NJ 08540 ,USA.
[2] Princeton Univ, Northeast Fisheries Sci Ctr, Natl Marine Fisheries Serv, NOAA,Geophys Fluid Dynam Lab, 201 Forrestal Rd, Princeton, NJ 08540 ,USA.
[3] Univ Connecticut, Dept Marine Sci, Groton, CT 06340, USA.

* Corresponding author : Hubert du Pontavice, email address : hubert.dupontavice@princeton.edu

Abstract :

The northeast U.S. continental shelf is a highly productive and socio-economically important marine ecosystem in which annual and seasonal variations of bottom temperature play a major role in the distribution, phenology, and productivity of its predominately demersal marine taxa. However, bottom temperature mea-surements are limited spatially and temporally and thus do not provide the required resolution to assess sub-seasonal variability and trends. Here we combined three ocean products, a regional ocean model (ROMS) and two global ocean data assimilated models (GLORYS12v1 and PSY4V3R1) to build a high-resolution, long-term bottom temperature product for the northeast U.S. continental shelf between 1959 and 2021. We bias-corrected ROMS using monthly decadal climatologies from ocean observations and analyzed long-term changes in the combined time series. Model skill was assessed using a large number of in situ observations. The combined bottom temperature product showed a long-term warming of the northeast U.S. continental shelf of + 0.36 degrees C decade-1 over the past 63 years, with notable variations among seasons and regions. The strongest long-term warming occurred during the summer months and in the Gulf of Maine. Although biases were observed, the bottom temperature product exhibited good performance reproducing seasonal and annual variability in observed temperature. This high-resolution product could be used in a wide range of applications from local to regional spatial scales, from long-term to near-term time scales, and from fisheries to marine ecology.

Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site.

Highlights

► A bottom temperature product for the northeast U.S. continental shelf. ► Long-term high-resolution bottom temperature between 1959 and 2021. ► Three ocean products were combined. ► Warming of the northeast U.S. shelf is estimated at + 0.36 °C decade^{-1}. ► Large warming variations among seasons and regions. ► Potential use for a wide range of applications from fisheries to marine ecology.

Please note that this is an author-produced PDF of an article accepted for publication following peer review. The definitive publisher-authenticated version is available on the publisher Web site.

1 Introduction

The northeast U.S. continental shelf marine ecosystem (NEUS) is a highly productive and socio-economically important ecosystem that spans from Cape Hatteras, North Carolina to the Gulf of Maine (Fig. 1). The NEUS is a western boundary confluence zone and its physical environment is shaped by two major water masses, the equatorward Labrador Current from the north and the poleward Gulf Stream from the south (Chapman and Beardsley, 1989; Chen et al., 2021; Gawarkiewicz et al., 2018; Mountain, 2012; Richaud et al., 2016).

While the temporal and spatial variability of sea surface temperature in the NEUS have been studied using satellite derived and *in situ* datasets (Chen et al., 2020; Richaud et al., 2016), exploring subsurface temperature variations is generally more challenging due to limited observations. Therefore, high-resolution numerical models and reanalyses as well as *in situ* observations are widely used to explore temporal and spatial variability of temperature in the deeper layer of the continental shelf (Chen et al., 2021; Chen and Curchitser, 2020; Friedland et al., 2020; Kavanaugh et al., 2017). On the NEUS, subsurface temperature variations are strongly influenced by local oceanic processes and can be decoupled from variations of surface temperature for a part of the year due to local circulation, stratification, or tidal forcing (Chen et al., 2021, 2018; Chen and Curchitser, 2020; Franks and Chen, 1996; Richaud et al., 2016).

Decadal, annual, and seasonal variations of bottom temperature play a major role in distribution, phenology and productivity of demersal marine biota in the region. Observed and projected shifts in groundfish and invertebrate distributions have been linked to warming bottom temperature (Friedland et al., 2021; Kleisner et al., 2017; Mazur et al., 2020; Tanaka et al., 2020). Phenology of groundfish can also be affected by changes in bottom temperature through warming induced shifts in spawning (Fuchs et al., 2020). Furthermore, previous studies in the

NEUS have identified relationships between seasonal benthic thermal environment and population processes associated with stock productivity. Recruitment variations of several NEUS groundfish including yellowtail flounder (*Limanda ferruginea*; du Pontavice et al., 2022; T. J. Miller et al., 2016), winter flounder (*Pseudopleuronectes americanus*; Bell et al., 2018), and black sea bass (*Centropristis striata*; A. S. Miller et al., 2016) are closely related to interannual and seasonal variations of bottom temperature. Recruitment of American lobster stocks in the Gulf of Maine and Georges Bank (GB) is also associated with spring benthic thermal condition, which impacts lobster habitat suitability (Tanaka et al., 2019). Moreover, other population processes such as growth, mortality, and maturity of Atlantic cod (*Gadus morhua*) and summer flounder (*Paralichthys dentatus*) are associated with interannual and seasonal variations in bottom temperature (Miller et al., 2018; O'Leary et al., 2019). Therefore, a comprehensive understanding of interannual variations and long-term changes in bottom temperature is essential to efficiently manage the NEUS ecosystem, fisheries, and protected species in a changing climate.

Over the last few decades, the NEUS has experienced significant ocean warming affecting the entire water column from the surface to bottom (Friedland et al., 2020; Pershing et al., 2015). Based on *in situ* and remote sensing data, Kavanaugh et al. (2017) highlighted benthic warming between 1982 and 2004 ranging from 0.1 to 0.4°C decade^{-1}. They also showed spatio-temporal variations with higher warming rates in inshore and nearshore regions (*e.g.,* Chesapeake Bay and GB and Gulf of Maine). More recently, Friedland et al. (2020) developed an interpolation method using ship-based measurements to produce spring and fall gridded bottom temperature estimates. They showed that bottom temperature has significantly increased at rates of 0.18°C decade^{-1} in spring and 0.31°C decade^{-1} in fall between 1968 and 2018 with a warming acceleration from 2008 onwards in fall. These studies based on *in situ* and remote sensing data were major steps toward the understanding of spatio-temporal patterns and long-term changes of bottom temperature.

The continuous temperature estimates across space and time provided by ocean models are required to analyze the sub-seasonal variations in bottom temperature. On the NEUS, the long-term (1958–2007) high-resolution (~7km) numerical simulation of the northwest Atlantic in the Regional Ocean Modelling System (hereafter called ROMS-NWA) resolves the spatiotemporal variability of shelf temperature on seasonal timescales (Chen et al., 2018; Chen and Curchitser, 2020). However, previous studies highlight a consistent warm bias in NWA-ROMS bottom temperature that is amplified during the stratified seasons of summer and early fall (Chen et al., 2018; Chen and Curchitser, 2020) and varies spatially within the continental shelf (Chang et al., 2021). The high-resolution global ocean reanalysis product GLobal Ocean ReanalYsis and Simulation (GLORYS12v1) simulates global ocean conditions between 1993 and 2019 (Lellouche et al., 2021). A recent study found that bottom temperature estimates from GLORYS12v1 are highly representative of observational bottom temperature on the NEUS (Chen et al., 2021). Moreover, another ocean product - Operational Mercator global ocean analysis and forecast system (PSY4V3R1) - including an observed data assimilation process and based on the same ocean model as GLORYS12v1 simulates 10 days of bottom temperature forecasts updated daily on a two-full-year time series sliding window (Lellouche et al., 2018). Therefore, PSY4V3R1 allows to extend the time series to the most recent years and can provide bottom temperature variations in near real time.

Here we combined three ocean products (ROMS-NWA, GLORYS12v1, PSY4V3R1) to build a high-resolution (1/12°) long-term bottom temperature product for the NEUS between

1959 and 2021. We first bias-corrected bottom temperature from ROMS-NWA using an observed climatology as in du Pontavice et al. (2022). Then, we analyzed the spatial and temporal patterns and long-term changes in bottom temperature using the combination of the three ocean products. Finally, we assessed the skill of the bottom temperature product over time and space using a large number of observations collected on the NEUS over the past 63 years.

2 Materials and Methods

2.1 Study area

Our study focused on the NEUS, an area of four Ecological Production Units (EPUs) defined by NOAA's Northeast Fisheries Science Center (https://noaa-edab.github.io/tech-doc/epu.html; Fig. 1): Mid-Atlantic Bight (MAB), GB, WGOM (Western-Central Gulf of Maine) and EGOM (Scotian Shelf-Eastern Gulf of Maine). The EPUs were generated through clustering analyses based on a set of physiographic, oceanographic and biotic variables on the NEUS within the 200-m isobath.

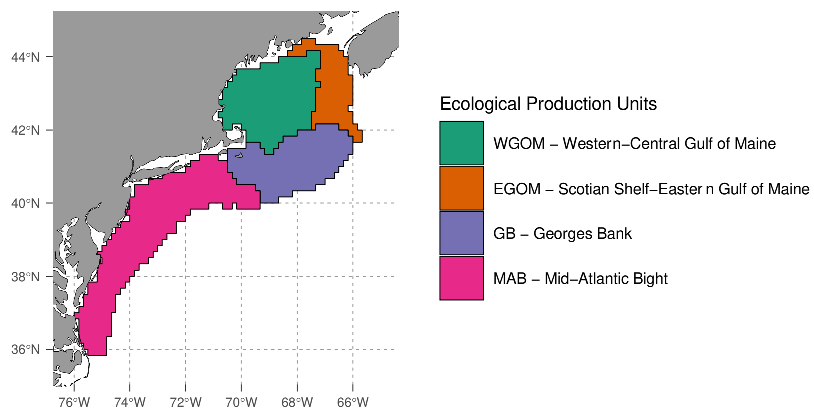


Figure 1. The northeast U.S. shelf and the four Ecological Production Units.

2.2 Ocean model data

Three ocean models were used that simulate high-resolution daily bottom temperature on the NEUS between 1959 and 2021. For the period between 1959 and 1992, we used daily ocean bottom temperature from the long-term (1958–2007) high-resolution numerical simulation of the Northwest Atlantic Ocean in the Regional Ocean Modelling System (ROMS), a split-explicit, free-surface, terrain-following, hydrostatic, primitive equation model (Shchepetkin and McWilliams, 2005). The model domain covers the Northwest Atlantic Ocean with ~7-km horizontal resolution and 40 vertical terrain-following layers. Initial and oceanic boundary forcing are both from Simple Ocean Data Assimilation version 2.1.6 (Carton and Giese, 2008), while atmospheric forcing is from the Coordinated Ocean-ice Reference Experiments datasets. A detailed description of ROMS-NWA can be found in Chen et al. (2018). Previously, ROMS-NWA had been used to study the cold pool dynamics on the MAB (Chen et al., 2018; Chen and Curchitser, 2020), investigate eddy characteristics and kinetic energy in the Gulf Stream region (*e.g.,* Kang et al., 2016; Kang & Curchitser, 2013), analyze range shifts of benthic invertebrates (Fuchs et al., 2020), and incorporate environment in a groundfish stock assessment model (du Pontavice et al., 2022).

For the period between 1992 and 2019, the daily bottom temperature outputs from the GLORYS12v1 ocean reanalysis product were used. GLORYS12v1 is a global ocean, eddy-resolving, and data assimilated hindcast from Mercator Ocean (Fernandez and Lellouche, 2018; Lellouche et al., 2021) with 1/12° horizontal resolution (~7-km to 9-km) and 50 vertical levels. The base ocean model is the Nucleus for European Modelling of the Ocean 3.1 (NEMO 3.1; Madec, 2016) driven at the surface by the European Centre for the Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (Dee et al., 2011). Remotely sensed and *in situ* observations are jointly assimilated by means of a reduced-order Kalman filter. In addition, a 3D-VAR scheme is employed in order to correct the large-scale, slowly-evolving error. The assimilated observations include sea surface temperature (SST) from Reynolds 0.25° Advanced Very-High-Resolution Radiometer-only (Reynolds et al., 2007), reprocessed along-track satellite altimeter missions sea level anomalies (SLA) from Copernicus Marine Environment Monitoring Service (CMEMS), Ifremer/CERSAT sea ice concentration (Ezraty et al., 2007), adjusted mean dynamic topography and in situ temperature and salinity vertical profiles from CMEMS quality controlled CORA database (Cabanes et al., 2013).

For the period between 2020 and 2021, we used daily bottom temperature from the Operational Mercator global ocean analysis and forecast system (PSY4V3R1 as in Lellouche et al., 2018). The PSY4V3R1 is a global ocean, eddy-resolving, monitoring forecasting system (Lellouche et al., 2018; Mercator Ocean International, 2016) with the same ocean model grid (1/12° horizontal resolution and 50 vertical levels) and has many similarities with GLORYS12v1. Using a reduced-order Kalman filter, PSY4V3R1 jointly assimilates remotely sensed and *in situ* observations including SST from the CMEMS Operational Sea Surface Temperature and Ice Analysis (OSTIA) system, CMEMS satellite near-real-time sea ice concentration, CMEMS altimeter satellite SLA, *in situ* vertical profiles from CMEMS database, adjusted mean dynamic topography and temperature and salinity climatology from the World Ocean Atlas (WOA 2013; Levitus et al., 2014). Finally, as in GLORYS12v1, a 3D-VAR scheme provides a correction for the slowly evolving large-scale biases in temperature and salinity.

The similarities between PSY4V3R1 and GLORYS12v1 (and the absence of GLORYS12v1 outputs for the most recent years due to a lag of more than one year to release a new year of data) led us to include PSY4V3R1 in the study to extend our times series until near-real time. However, GLORYS12v1 has three main refinements compared to PSY4V3R1: the use of reanalyzed atmospheric forcing (instead of analyses and forecasts), higher-quality reprocessed observations, improved data assimilation procedures (Lellouche et al., 2021). Hence, GLORYS12v1 was expected to provide better bottom temperature estimates.

Other ocean models simulate the NEUS such as HYCOM (https://hycom.org/hycom; same resolution and period of GLORYS12v1) and DOPPIO (López et al., 2020; covering the period 2007–2017). A recent study assessed the skill of all available global reanalysis products (CFSR93.2, ECCO V5, ORAS5, SODA3.12.2, Bran2020, Glorys12v1, HYCOM3.0, and HYCOM3.1) for the NEUS and showed that bottom temperature from GLORYS12v1 had the best performance when compared to *in situ* observations (Pers. Comm. Alma Carolina Castillo-Trujillo – in review). A regional data-assimilative model reanalysis on the NEUS covering 2007 onwards is under peer review (Anonymous reviewer comment) and it will be interesting to compare its performance with that of GLORYS12v1. However, GLORYS12v1 also demonstrates good performance, and it starts from 1993, which has an overlapping time period with the ROMS-NWA dataset (1958-2007) for calibration and validation. Additionally, GLORYS12v1 can be linked to PSY4V3R1 to provide near-real time bottom temperature since both models are based on the same ocean model and they have the same structure.

2.3 Bottom temperature observations

2.3.1 Decadal monthly climatology

In order to bias-correct ROMS-NWA bottom temperature between 1959 and 1992, we used the monthly climatologies of observed bottom temperature from the Northwest Atlantic Ocean regional climatology (NWARC) over decadal periods from 1955-64 to 1985-94. The NWARC provides monthly high-resolution (1/10° grids) quality-controlled *in situ* ocean temperature based on a large dataset of observed temperature (Seidov, Baranova, Boyer, et al., 2016; Seidov, Baranova, Johnson, et al., 2016; https://www.ncei.noaa.gov/). The decadal climatologies are computed based on profile data from the World Ocean Database (WOD). Although the Northwest Atlantic Ocean has a high volume of available ocean profile temperature, there are a few gaps in data coverage. These spatial and temporal data gaps were filled by interpolation using the objective analysis of irregularly distributed data (Locarnini et al., 2019).

2.3.2 *In situ* bottom temperature observations

In situ bottom temperature observations are discontinuous over space and time but numerous on the NEUS. Multiple sampling programs have monitored the water column temperature between 1959 and 2021. From 1977 onwards, the Northeast Fisheries Science Center (NEFSC) oceanographic database collected and gathered bottom temperature observations from different surveys including the Marine Resources Monitoring Assessment and Prediction program (MARMAP; 1977–1987), the Ecosystem Monitoring program (EcoMon; 1992-present) and the NMFS NEFSC bottom trawl surveys. The temperature profile data from NEFSC oceanographic database were collected using different conductivity, temperature, and depth instruments (CTD) and glass bottle temperatures (included in the type of data Ocean Station Data; OSD). The total number of NEFSC bottom temperature observations was 48,880 with mainly CTD measurements (Table 1).

The second source of bottom temperature observations on the NEUS is the NOAA NCEI's World Ocean Database (Boyer et al., 2018). The WOD is the world's largest collection of quality controlled, publicly available ocean temperature profiles coming from different institutions, agencies, individual researchers, and data recovery initiatives. Temperature was measured using CTD, OSD, Expendable Bathythermograph (XBT) and Mechanical Bathythermograph (MBT). We used 66,507 bottom temperature measurements on the NEUS between 1959 and 2021 from WOD. Since most observations collected by NEFSC after 1977 are also included in WOD, we excluded observations already downloaded from the NEFSC database when extracting observations from WOD to ensure that we avoid duplicates. There are 20,914 measurements from the NEFSC oceanographic database for the period 1977–2021 and 45,593 measurements covering the period 1959–1976. While the NEFSC oceanographic database temperature observations are clearly identified as bottom temperature, it is not the case in WOD. Therefore, we constrained the WOD temperature observations by selecting the deepest measurement within 10 meters of the sea floor.

Table 1. Total number of observations for each type of oceanic instrument extracted from either the Northeast Fisheries Science Center oceanographic database (NEFSC database) or the NOAA NCEI's World Ocean Database (WOD).

	1977–2019		**1959–1976**
	NEFSC database	**WOD**	**WOD**
CTD	44,024	6,355	571
OSD	4,856	3,306	5,423
MBT	0	919	34,905
XBT	0	10,334	7,277
Subtotal	48,880	20,914	
Total	69,311		45,593

The WOD is the only source of observations for the years prior to 1977 and represents a large number of observations up to the late 1990s (Fig. 2). Between 1990 and 2021, the NEFSC oceanographic database is the main source of observations and these measurements are primarily collected using CTD instruments allowing for higher accuracy and higher precision than MBT, XBT and OSD (Fig. 2).

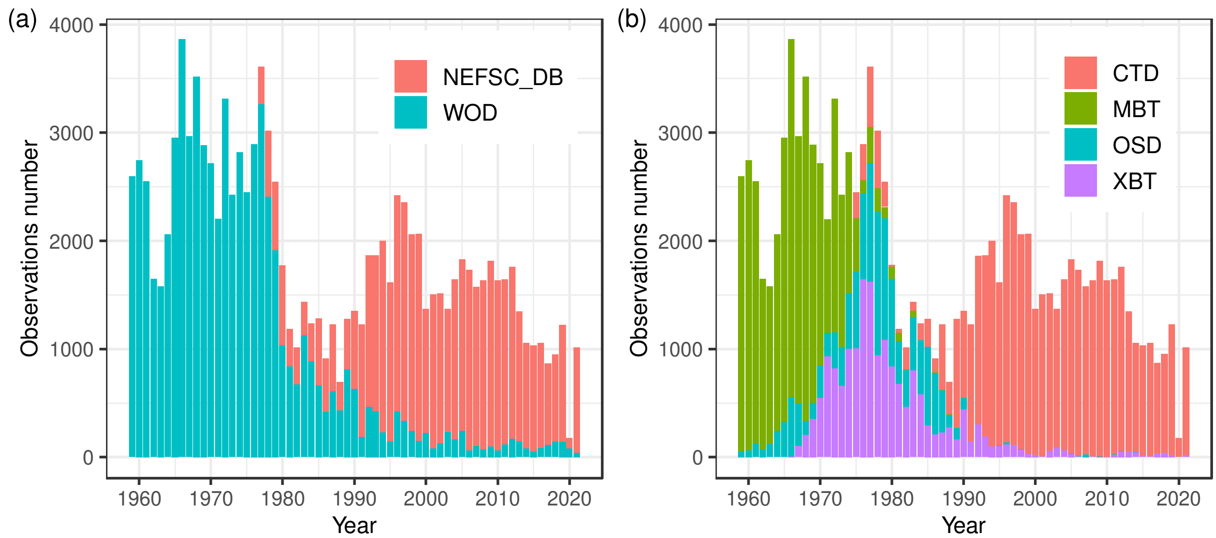


Figure 2. Number of bottom temperature observations per year for each type of oceanic probe extracted from the Northeast Fisheries Science Center oceanographic database (NEFSC_DB) and the NOAA NCEI's World Ocean Database (WOD) (left panel). The right panel indicates the type of probe used to collect the measurements: conductivity, temperature, and depth instruments (CTD), Mechanical Bathythermograph (MBT), Ocean Station Data (OSD) and Expendable Bathythermograph (XBT).

2.4 Bias-correction process of NWA-ROMS

Several studies highlighted that bottom temperature in ROMS-NWA is warmed-biased on the NEUS (Chang et al., 2021; Chen and Curchitser, 2020) and bias-correction methods were developed to reduce the bias (Chang et al., 2021; du Pontavice et al., 2022; Fuchs et al., 2020). In this study, we used the methodology implemented in du Pontavice et al. (2022) based on observed climatology.

The first step was to regrid ROMS-NWA bottom temperature over the same 1/10° horizontal grid as the NWARC using bilinear interpolation. Then, we conducted the bottom temperature bias-correction in the 1/10° NWARC grid using monthly climatologies from NWARC over four decadal periods from 1955 to 1994. A monthly bias was calculated in each

1/10° grid cell and for each decade (1955–1964, 1965–1974, 1975–1984, 1985–1994) (Equation 1).

$$Bias(x_R, y_R, t_{dec}, t_m) = T_{NWARC}(x_R, y_R, t_{dec}, t_m) - \bar{T}_{ROMS}(x_R, y_R, t_{dec}, t_m) \quad (1)$$

where $Bias(x_R,y_R,t_{dec},t_m)$ is the monthly bias, $T_{NWARC}(x_R,y_R,t_{dec},t_m)$ is the monthly bottom temperature from NWARC, $T_{ROMS}(x_R,y_R,t_{dec},t_m)$ is the monthly mean bottom temperature from ROMS-NWA, x_R and y_R are the longitude and latitude of 1/10° NWARC grid cells, t_{dec} is the NWARC decade, and t_m is the month.

Based on this monthly bias, we estimated a daily bias by fitting generalized additive models (GAM) for each decade in each grid cell (see details in Appendix A). Lastly, for each ROMS-NWA grid cell (with coordinates x_{ROMS} and y_{ROMS}) we identified the bias from the closest 1/10° NWARC grid cell (with coordinates x_R and y_R) and subtracted the daily bias, $Bias(x_R,y_R,t_{dec},t_{day})$, to the daily ROMS-NWA bottom temperature estimates $T_{ROMS}(x_R,y_R,t_y,t_{day})$ for all years, t_y, and days t_{day}, of each decade, t_{dec} (Equation 2).

$$T_{ROMS_{cor}}(x_{ROMS}, y_{ROMS}, t_y, t_{day}) = T_{ROMS}(x_{ROMS}, y_{ROMS}, t_y, t_{day}) - Bias(x_R, y_R, t_{dec}, t_{day}) \quad (2)$$

where $T_{ROMScor}(x_{ROMS},y_{ROMS},t_y,t_{day})$ is the daily bias-corrected ROMS-NWA bottom temperature and $T_{ROMS}(x_{ROMS},y_{ROMS},t_y,t_{day})$ is the raw ROMS-NWA bottom temperature.

The spatial and temporal patterns of daily decadal bias estimates are presented in Appendix A.

2.5 Design of the gridded bottom temperature time series

The final bottom temperature product is in a horizontal 1/12° grid between 1959 and 2021 and is made of daily bottom temperature estimates from:

- Bias-corrected ROMS-NWA ($ROMS_{cor}$) between 1959 and 1992 which was regridded in the same 1/12° grid as GLORYS and PSY4V3R1 using bilinear interpolation;
- GLORYS12v1 in its original 1/12° grid between 1993 and 2019;
- PSY4V3R1 in its original 1/12° grid between 2020 and 2021.

2.6 Time series analysis

In order to analyze the time series from the combined bottom temperature product, we conducted a trend analysis based on one of the approaches developed by Hardison et al. (2019). We used a Generalized Least Squares (GLS) model selection (GLS-MS) approach to evaluate bottom temperature trends over the period 1959–2021, as this approach allowed for both linear and quadratic model fits and accounts for potential autocorrelation in the time series. The model selection procedure fitted four models to each time series and selected the best fitting model using the Corrected Akaike's Information Criterion (AIC). The models were, 1) linear trend with uncorrelated residuals, 2) linear trend with correlated residuals, 3) quadratic trend with uncorrelated residuals, and 4) quadratic trend with correlated residuals. Then, the best fit model was tested against the null hypothesis of no trend through a likelihood ratio test ($p < 0.05$).

The identification of bottom temperature time series change points was conducted with a sequential regime shift detection algorithm called STARS (Sequential T-test Analysis of Regime Shifts; Rodionov, 2006, 2004; Stirnimann et al., 2019). This algorithm is based on a sequential t-test analysis in which the hypothesis of a regime shift at each time step is accepted or rejected. A Regime Shift Index was calculated at each time step to determine whether the following values are significantly different from the mean of the previous regime. We set the p-value to 0.05 and the window size (moving window in which the algorithm calculates the probability of a regime shift) to 10 years. Moreover, we used a Huber's weight parameter (h=1) to reduce the effect of outliers.

All analyses were performed using the open-source statistical software R (R Core Team, 2021), the trend analysis models were fit using the R package *nlme* and the STARS algorithm was coded in R and modified from the code developed by Seddon et al., (2014).

2.7 Model bias analysis

We conducted a model bias analysis by comparing the modeled bottom temperature estimates to observations to explore the biases and their annual variations for the entire year, for each season, and for each EPU. This analysis was implemented in the original grid of ROMS-NWA (~7km) for the period 1959–1992, the 1/12° grid of GLORYS12v1, and PSY4V3R1 for the period 1993–2021. We associated each observation to the nearest modeled bottom temperature estimate (based on the grid cell centroid). If more than one observation was assigned to one model estimate (same day/ year/ nearest grid point), we compared the model estimate to the mean of these observations. The difference between modeled and observed bottom temperature (hereafter called "model bias") was calculated for all days and grid cells where observations were available. Moreover, we calculated the root mean square error (RMSE) between model estimates and observations and the Pearson correlations between the modeled and observed bottom temperature times series. We conducted three distinct analyses for $ROMS_{cor}$ (1959–1992), GLORYS12v1 (1993–2019) and PSY4V3R1 (2020–2021).

One of the challenges to study the inter-annual variability of bias is the spatiotemporal heterogeneity of the observations among years, seasons, and EPUs (see Appendices B and C). We set up a procedure to mitigate this issue and allowed us to compare the annual bias despite the difference in spatiotemporal distribution of observations. First, we limited our bias analysis for the EPUs having more than a specific number of observations per year and season to exclude the EPU-season combinations in which too few observations are available. This number depended on the surface areas of the EPUs in order to account for the notable size difference among EPUs. We chose the EPU-season combinations with more than 10, 20, 20, and 35 observations available each year for the EGOM (28,858 km^2), WGOM (61,065 km^2), GB (55,703 km^2) and MAB (99,720 km^2), respectively.

To analyze the interannual variability of annual bias on the NEUS, we selected the years in which more than 12 (out 16) EPU-season combinations met the criteria for the minimum number of observations mentioned above. Then, we calculated the mean biases in all EPU-season combinations and we averaged these mean biases for each year, t_y (Equation 3):

$$MB(t_y) = \frac{\sum_{EPU\ x\ S}\left(\frac{\sum_{c\ x\ i}\left(T_{mod}\left(c,t_{day},t_y\right)-T_{obs}\left(c,t_{day},t_y\right)\right)}{N_{EPU,S}}\right)}{N_{EPU\ x\ S}} \tag{3}$$

where $MB(t_y)$ is the mean model bias for the year t_y, $T_{mod}(c,t_y,t_{day})$ is the modeled bottom temperature estimate in the grid cell c for the day t_{day} of the year t_y, $T_{obs}(c,t_y,t_{day})$ is the observed bottom temperature in the grid cell c for the day t_{day} of the year t_y, $N_{EPU\ x\ s}$ is the number of EPU-season combination for the year t_y, and $N_{EPU,s}$ is the number of cell with available observations for each given EPU-season combination. This method allowed to keep only the years in which EPU-season combinations having a sufficient number of observations and give an equal weight to all the EPUs and seasons.

For the analysis of the interannual variability of seasonal biases, we selected the years in which the number of observations in each EPU met the criteria for the minimum number of observations mentioned above. We calculated the seasonal bias using a similar method as for the annual bias. First, we calculated the bias in each EPU at each season and then averaged these biases for each season (Equation 4):

$$MB(t_y, s) = \frac{\sum_{EPU}\left(\frac{\sum_c \left(T_{mod}\,(c, t_{day}, t_y) - T_{obs}(c, t_{day}, t_y)\right)}{N_{EPU,s}}\right)}{4} \quad (4)$$

where $MB(t_y,s)$ is the mean model bias for the year t_y and the season s, $T_{mod}(c,T_{day},t_y)$ is the modeled bottom temperature estimate in the grid cell c for the day t_{day} of the year t_y, $T_{obs}(c,t_{day},t_y)$ is the observed bottom temperature in the grid cell c for the day t_{day} of the year t_y, 4 is the number of EPUs, and $N_{EPU,s}$ is the number of cell with available observations for each given EPU-season combination.

Finally, to analyze the interannual variability of the bias in each EPU, we selected the years in which the number of observations in each EPU met the criteria for the minimum number of observations available in each season. First, we calculated the bias in each EPU at each season and then averaged these biases for each EPU. We excluded the winter season because of too limited number of observations (or even the absence of observation) for several years between 1959 and 2021.

3 Results

3.1 Spatial and temporal bottom temperature patterns

The mean annual bottom temperature time series showed a long-term significant warming of +0.36±0.06°C decade^{-1} (linear trend with correlated residuals and p-value=4.7e-05) between 1959 and 2021 on the NEUS (Fig. 3). Bottom temperature warmed in each season since 1959 (linear trends with correlated residuals and p-value<1e-03: Fig. 3b) of +0.36 ±0.06°C decade^{-1}, +0.34 ±0.06°C decade^{-1}, +0.41 ±0.06°C decade^{-1}, +0.33 ±0.07°C decade^{-1} in winter, spring, summer, and fall, respectively.

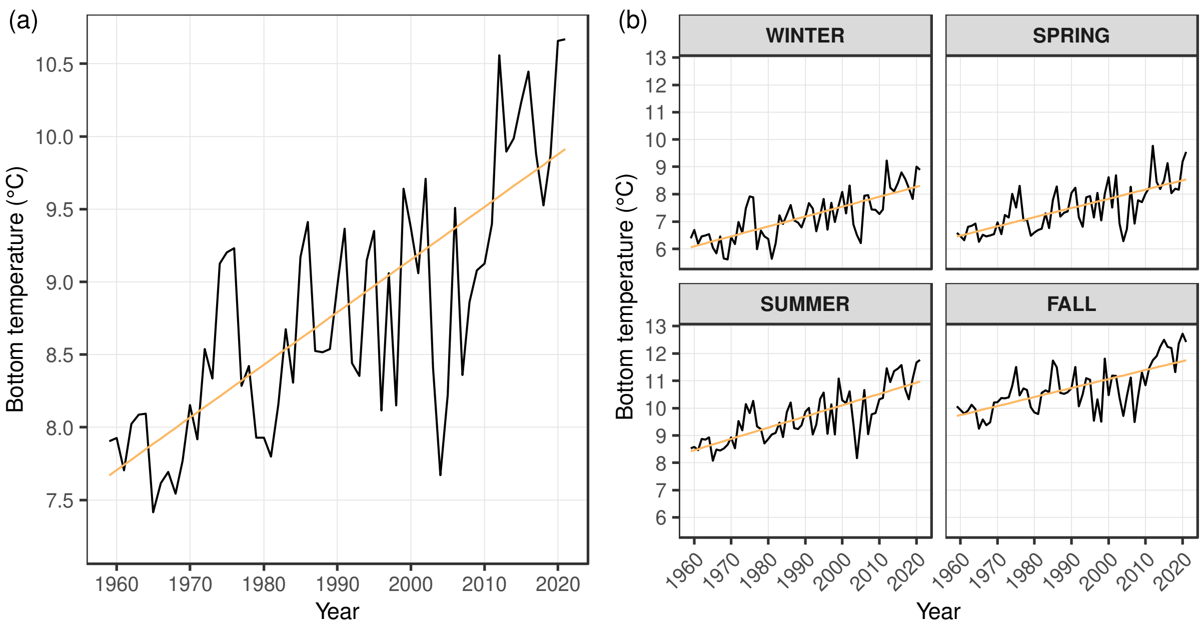


Figure 3. Mean annual (a) and seasonal (b) bottom temperature time series on the northeast U.S. continental shelf between 1959 and 2021 and the linear trends (orange lines) for each season with 95% confidence interval (gray polygons).

In addition to the long-term trend analysis, the change points analysis revealed five abrupt changes in 1972, 1977, 1985, and 2012 (Table 2). These were induced by shifts in bottom temperature in each season (Table 2). The year 1972 was marked by the first abrupt temperature increase of 0.98°C that occurred during the summer. The second and third change points were an increase in bottom temperature of 0.91°C in winter of 1983 and 0.82°C and 0.53°C in the spring and fall of 1985, respectively. The last change points were in the early 2010s (2010, 2011, or 2012) for all the seasons which were characterized by abrupt and large increases in bottom temperature. The comparison between the seasonal mean bottom temperature before and after 2011 revealed an increase of 1.62 °C, 1.48 °C, 1.85 °C, and 1.73 °C in winter, spring, summer and fall, respectively.

Table 2. Change points of NEUS bottom temperature seasonal times series and the mean bottom temperature (Mean BT) of each time period.

Season	Change point	Period	Mean BT
		1959–1982	6.39
Winter	1983	1983–2011	7.30
	2012	2012–2021	8.53
		1959–1984	6.81
Spring	1985	1985–2010	7.63
	2011	2011–2021	8.60
		1959–1971	8.60
Summer	1972	1972–2009	9.58
	2010	2010–2021	11.12
		1959–1984	10.17
Fall	1985	1985–2010	10.70
	2011	2011–2021	12.13
		1959–1971	7.83
	1972	1972–1976	8.91
Entire year	1977	1977–1984	8.19
	1985	1985–2011	8.91
	2012	2012–2021	10.18

Across the NEUS, the trend analysis for the period 1959–2021 showed significant changes in bottom temperature in the majority of the 1/12° grid cells in each season (87%, 80%, 79%, and 77% of the grid cells in winter, spring, summer, and fall, respectively; Fig. 4b). Within the areas that experienced a significant change in bottom temperature, the proportions of the grid cells that exhibit an increase in temperature range between 96% and 100% depending on the season. The combined bottom temperature product resolved the seasonal characteristics of bottom water (*e.g.,* the Cold Pool on MAB; Fig. 4a), and showed that the strongest warming was in the southern MAB (from Chesapeake Bay to the southern boundary of the MAB) and western GB (Nantucket shoals), especially in summer and fall (Fig. 4b). While the WGOM experienced a spatially homogeneous increase in bottom temperature year-round, all other EPUs exhibited several areas wherein the warming was not significant (Fig. 4b). In the MAB, these non-significant areas were numerous in all seasons, even in summer where the decadal trend was the greatest among all regions and seasons.

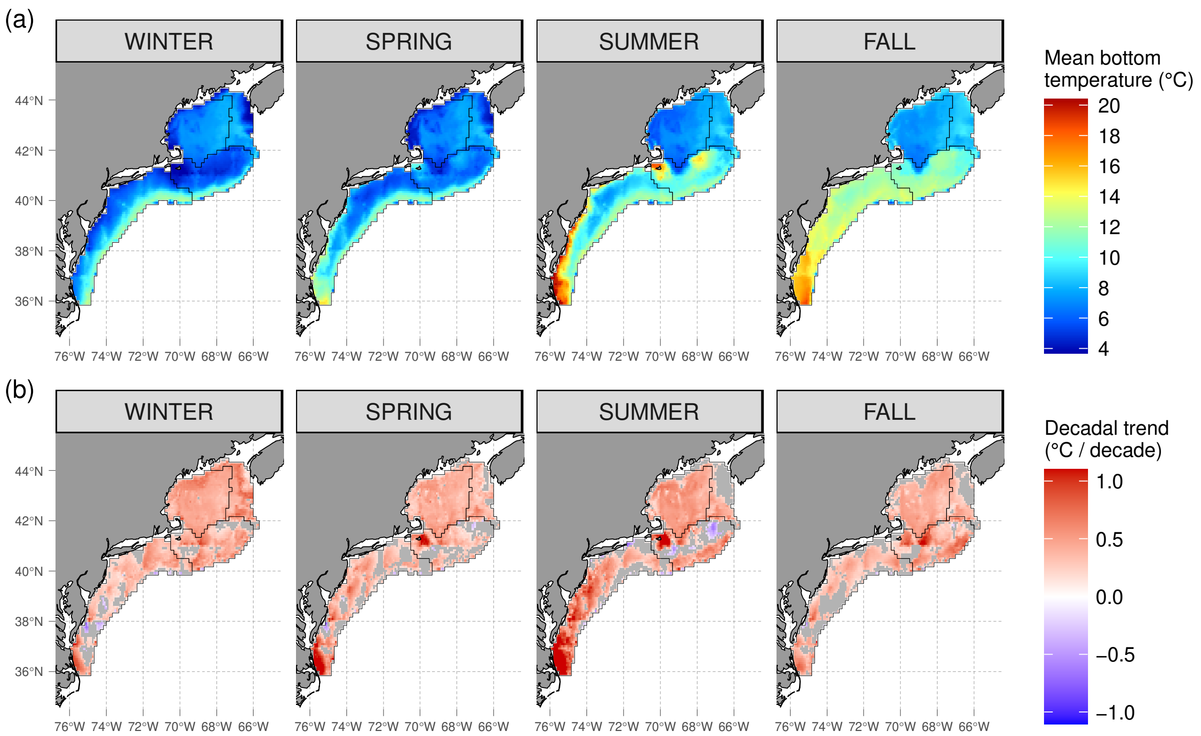


Figure 4. Seasonal mean bottom temperature (a) and decadal trends (b) over the whole time period 1959–2021 in the Northeast U.S. continental shelf. The gray areas in panel b are the grid cells where the decadal trend is not significant (p-value>0.05). The line marks the contour of the Ecological Production Units (see Fig. 1)

At the aggregated spatial scale, the four EPUs have experienced significant warming in bottom temperature in all seasons (Fig. 5). The WGOM experienced the greatest warming in bottom temperature with a decadal trend of +0.43°C decade^{-1} while the decadal trends in the MAB, GB, and EGOM ranged between +0.31 and +0.35°C decade^{-1}. In the WGOM, the warming rate was equal to or greater than 0.39°C decade^{-1} year-round and the decadal trends in bottom temperature varied considerably from season to season. In the EGOM, bottom temperature increased 0.47°C decade^{-1} in winter while the decadal trends in the other seasons did not exceed +0.30°C decade^{-1}. In GB, the two largest increases in bottom temperature were +0.35 and +0.40°C decade^{-1} and occurred in winter and fall. Conversely, the MAB experienced a relatively low increase in bottom temperature in winter and fall, moderate in spring, but very strong in the summer (+0.50°C decade^{-1}).

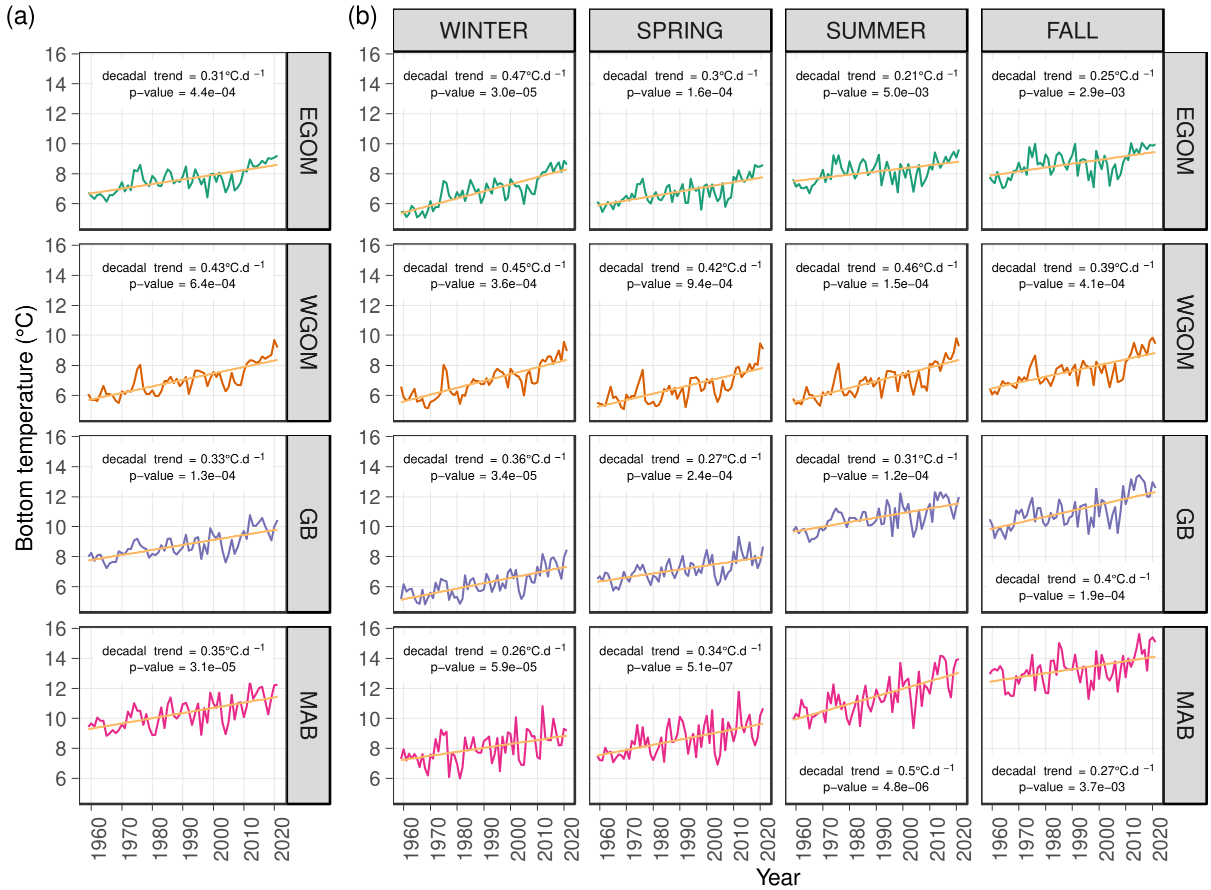


Figure 5. Mean bottom temperature in each Ecological Production Unit (EPU) (a) and in each season and EPU (b) and their linear trends (orange lines) with 95% confidence interval (gray polygons). The decadal trend (decadal trend) and the p-value of the selected GLS model are represented in each panel.

3.2 Model bias exploration using observations

The mean biases on the NEUS over the whole time period are very low with values of 0.03, -0.02 and -0.02 °C for $ROMS_{cor}$, GLORYS12v1, and PSY4V3R1, respectively although the RMSE range between 1.78 and 2.02 (Fig. 6 and Table 3). The mean bottom temperature estimated by the combined product is highly correlated to the observations (r=0.91 and r=0.98 for $ROMS_{cor}$ and GLORYS12v1, respectively, with p-value<0.05) (Fig. 6 and Table 3). The bias correction greatly improved the quality of model estimates (Fig. 6) with a lower model bias and

RMSE and higher correlation (Table 3). The bias analysis also revealed that the mean annual bias was varying over the time series with larger biases and interannual variations during the period covered by ROMS$_{cor}$ (left panel on Fig. 6 and higher RMSE). During the second part of the time series covered by GLORYS12v1 (1993–2019), the mean annual bias is less variable (Fig. 6). We did not calculate the biases for the years 1960, 1961, 1963, 2015, 2016, and 2017 because of the scarcity and spatiotemporal heterogeneity of observations for these years (less than 12 (out 16) EPU-season combinations; see section 2.7). We discuss the quality of the modeled bottom temperature estimates for the years having limited observations with a focus on the years 2015, 2016 and 2017, which are of primary interest for fisheries and ecological applications.

Table 3. Annual mean bias, RMSE and Pearson correlation for the raw ROMS-NWA (1959–1992), the bias-corrected ROMS-NWA (ROMS$_{cor}$; 1959–1992), GLORYS12v1 (1993–2019) and PSY4V3R1 (2020–2021).

	Raw ROMS-NWA	**ROMS$_{cor}$**	**GLORYS12v1**	**PSY4V3R1**
Mean bias	1.00	0.03	-0.02	-0.02
RMSE	2.35	2.02	1.78	1.90
Cor	0.72	0.91	0.98	NA

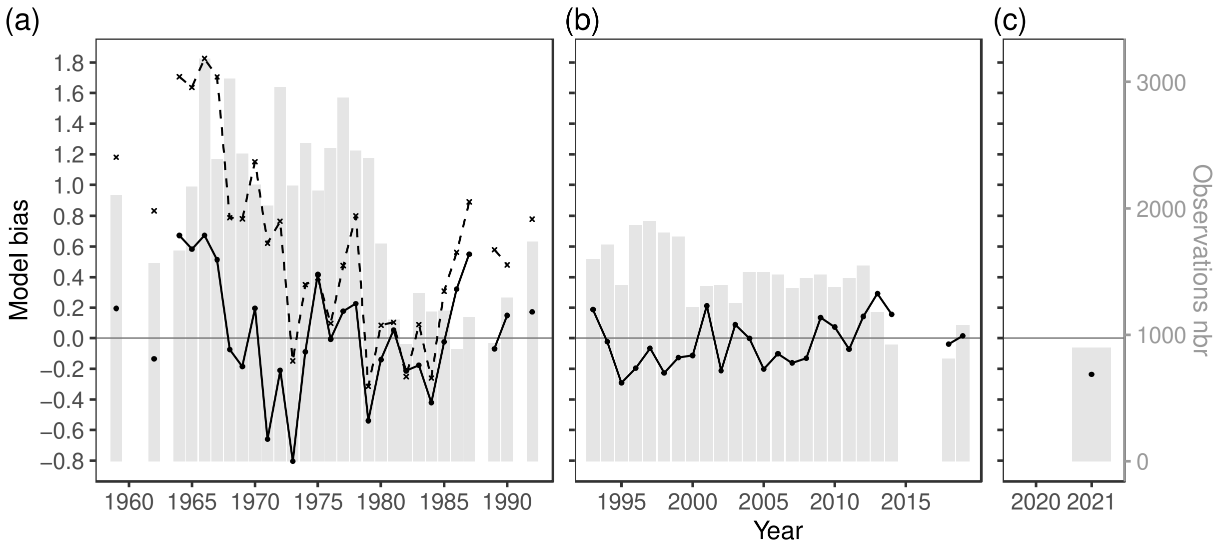


Figure 6. Annual mean model bias in bottom temperature of the raw ROMS-NWA between 1959 and 1992 (dashed line, (a)), the bias-corrected ROMS-NWA between 1959 and 1992 (solid line, (a)), GLORYS12v1 between 1993 and 2019 (b) and PSY4V3R1 in 2020 and 2021 (c). The bar plots represent the number of observations (Observations nbr).

Modeled mean bottom temperature is significantly correlated to the observed time series in each season with higher Pearson correlation coefficients for GLORYS12v1 ranging between 0.93 and 0.99 (Table 4). Correlation coefficients with $ROMS_{cor}$ estimates were lower, especially in the spring and fall (0.82 and 0.84) (Table 4). Furthermore, the mean model bias was low for all three products and for all seasons, especially in spring. The bias correction substantially reduced the bias for all seasons except for winter, where the mean bias and RMSE of the raw ROMS-NWA was slightly lower than $ROMS_{cor}$ (Table 4). However, this result must be interpreted with caution given the scarcity of observations in winter. As for the mean bias, we found larger annual model biases for the period covered by $ROMS_{cor}$ (RMSE of $ROMS_{cor}$ greater than GLORYS12v1 and PSY4V3R1) and in summer for the whole time period (greater RMSE in summer for the three models; Table 4 and Appendix D – Fig. D.1). Moreover, the modeled estimates tended to be slightly warmer than observations in winter, while bottom temperatures of GLORYS12v1 in summer and fall were cooler than observations (Table 4 and Appendix D – Fig. D.1). The bias analysis showed some large mean biases close to 1.0 in the spring of 1965 (+0.93) and 2015 (+0.88), summer of 1971 (-1.44) and 1973 (-0.90), and in fall of 1964 (+0.92), 1966 (+1.01), and 1979 (-1.19). The potential origins of these biases will be discussed.

Table 4. Seasonal mean bias, RMSE and Pearson correlation for the raw NWA-ROMS (1959–1992), the bias-corrected NWA-ROMS ($ROMS_{cor}$; 1959–1992), GLORYS12v1 (1993–2019) and PSY4V3R1 (202-0-2021).

	Season	**Raw ROMS-NWA**	**$ROMS_{cor}$**	**GLORYS12V1**	**PSY4V3R1**
Mean bias	*Winter*	0.22	0.37	0.33	0.17
	Spring	1.03	-0.02	0.03	0.39

	Summer	1.71	-0.27	-0.20	-0.25
	Fall	0.84	0.11	-0.23	-0.52
RMSE	*Winter*	1.54	1.69	1.25	1.22
	Spring	2.14	1.87	1.44	1.50
	Summer	3.13	2.53	2.41	2.64
	Fall	2.12	1.79	1.66	1.68
Cor	*Winter*	0.89	0.92	0.99	NA
	Spring	0.55	0.82	0.93	NA
	Summer	0.76	0.89	0.95	NA
	Fall	0.64	0.84	0.98	NA

Modeled bottom temperature was significantly correlated to observations for the four EPUs with higher correlations in the period covered by GLORYS12v1 and lower correlation in the $ROMS_{cor}$ period (especially in the WGOM, r=0.72) (Table 5). The notable low correlation observed in the WGOM at the beginning of the time series (prior 1975) was mainly due to the modeled estimates which exhibited successively large warm and then cold biases with a major change in 1971 (Appendix D – Fig. D.2). The improvement of mean bias, RMSE and correlation between the raw ROMS-NWA and $ROMS_{cor}$ for all EPUs (except on the EGOM where RMSE was higher for $ROMS_{cor}$) highlighted again the substantial improvements of ROMS-NWA estimates due to the bias correction. Although the mean biases were relatively low for $ROMS_{cor}$, the annual biases were much larger than for the period covered by GLORYS12v1 (Appendix D – Fig. D.2). On the EGOM, the mean bias increased from slightly colder in the 1990s to warmer after 2015 (Appendix D – Fig. D.2). In the WGOM, although modeled bottom temperature is largely biased until 1975, it appeared to be highly consistent with observations between 1994 and 2012. Between 2014 and 2017, we did not analyze the annual biases in the WGOM because there were too few or no observations available in winter and summer in 2014, 2015, 2016 and 2017. In GB, the mean annual modeled bottom temperature was colder than observations (Appendix D – Fig. D.2) with a mean negative bias (Table 5). In the MAB, the low mean bias masked larger annual biases than in the other EPUs with a higher RMSE up to 2.0 for the time periods covered by $ROMS_{cor}$ and GLORYS12v1.

Table 5. Mean bias, RMSE and Pearson correlation in each Ecological Production Unit for the raw NWA-ROMS (1959–1992), the bias-corrected NWA-ROMS ($ROMS_{cor}$; 1959–1992), GLORYS12v1 (1993–2019) and PSY4V3R1 (2020–2021).

	Season	**Raw ROMS-NWA**	**$ROMS_{cor}$**	**GLORYS12V1**	**PSY4V3R1**
Mean bias	*EGOM*	-0.13	0.05	-0.17	-0.26
	WGOM	0.65	0.01	0.19	0.64
	GB	1.22	-0.23	-0.51	-1.01
	MAB	1.56	0.00	0.17	0.17
RMSE	*EGOM*	1.38	1.49	1.28	1.41
	WGOM	1.73	1.25	1.14	1.42
	GB	2.25	1.95	1.99	2.50
	MAB	2.96	2.43	2.34	2.14
Cor	*EGOM*	0.67	0.80	0.95	NA
	WGOM	0.31	0.72	0.95	NA
	GB	0.79	0.91	0.96	NA
	MAB	0.74	0.86	0.93	NA

4 Discussion

4.1 Long-term bottom temperature warming in the NEUS

The bottom temperature estimates from the combination of three models: ROMS, GLORYS12v1, and PSY4V3R1 showed a long-term warming of the NEUS benthic ecosystem between 1959 and 2021. The time series revealed two major shifts in bottom temperature that occurred in the early 1970s and then again in the 2010s, which supports the findings of a recent study that showed an abrupt warming in 2009–2010 and suggested the occurrence of a similar event in the early 1970s (Gonçalves Neto et al., 2021). Previous analyses of NEUS bottom temperature based on ocean models or observed data (*in situ* and/or remote sensing) were focused on shorter time periods and could not detect the first major bottom temperature shift in the early 1970s (Friedland et al., 2020; Fuchs et al., 2020; Kavanaugh et al., 2017; Seidov et al., 2021). The combination of the three ocean products covering a time period of 63 years suggests that bottom temperature in the NEUS started warming as early as the beginning of the 1970s with an increase in temperature of ~1°C in summer between the periods 1959–1971 and 1972–2009. This implies that benthic ecosystems experienced a major change in thermal conditions in the early 1970s that likely impacted marine biota, which has not been explored thus far.

The long-term decadal change in spring and fall is qualitatively consistent with the results of Friedland et al., (2020) but the combined bottom temperature product showed a stronger warming. Based on an interpolation method between 1968 and 2018, the study by Friedland et al. (2020) estimated an increase of +0.18°C decade^{-1} and +0.31°C decade^{-1} in the spring and fall respectively compared to our results of +0.34°C decade^{-1} and +0.36 °C decade^{-1} in the spring and in the fall. The divergence in spring may be partly due to the length of our time series, which is longer plus the adjustment date of their spring time series which is April 3, while we considered that the spring season spans from April to June. Although a direct comparison with the results of Kavanaugh et al., (2017) is not possible because they considered a much shorter time period (1982–2014), we found a similar warming rate in GB of 0.3°C decade^{-1} but a higher warming rate in the Gulf of Maine (WGOM and EGOM) of 0.39°C decade^{-1} compared to 0.2°C decade^{-1} in Kavanaugh et al., (2017). The latter difference may partly be explained by the recent benthic warming in the Gulf of Maine, which cannot be fully captured by the study of Kavanaugh et al., (2017). A recent analysis based on *in situ* data also supports our findings regarding the warming magnitude in the Gulf of Maine showing that bottom temperature rose by almost 1°C between the 1965–1984 and 1995–2017 time periods while the combined product exhibited a 0.94°C increase over the same period (Seidov et al., 2021). Moreover, the cold period observed in the 1960s Gulf of Maine (Loder et al., 2001) was resolved by our bottom temperature product. This is mainly due to the bias-correction process which decreased ROMS bottom temperature during the decades 1955-1964 and 1965-1974 (Appendix A).

The combined bottom temperature product showed that the WGOM experienced the strongest benthic warming in the NEUS. Compared to the other EPUs, the increase in bottom temperature was spatially homogeneous year-round over the past 63 years with a robust warming acceleration in the beginning of the 2010s. The variations in bottom temperature in the Gulf of Maine are primarily influenced by the inflow of slope water through the Northeast Channel (Seidov et al., 2021) except in winter when convective mixing is maximized and when cooler water flowing along the Scotian Shelf enters the region (Mountain and Manning, 1994). The northern shift of the Gulf Stream, which started around 2008-2010, may be one of the primary drivers of the enhanced benthic warming in the WGOM via increasing inflow of Gulf Stream-associated slope water through the Northeast Channel and the reduction of the cooler slope water from the Labrador Current (Gonçalves Neto et al., 2021; Saba et al., 2016; Seidov et al., 2021). A recent modeling study showed that recent anomalous warm events in 2012, 2014, and 2015 observed in the subsurface water in the Gulf of Maine were induced by the interaction between the Gulf Stream and the Labrador Current at the tail end of the Grand Banks (Brickman et al., 2018). This interaction led to anomalous warm and salty water masses penetrating into the Gulf of Maine via deep channels along the shelfbreak. A similar process is likely responsible for the recent bottom temperature warming in the EGOM. An analysis of monthly anomalies in bottom temperature was conducted over the entire time series in order to identify the concurrency of the warming signals (Appendix E). We found several synchronous benthic warming signals in the WGOM and EGOM with a lag of 1–5 months (e.g., 1974, 1994, 1990, 2006, 2010). Each of these warming events occurred first in the WGOM and then in the EGOM which may be due to the relative importance of processes at play affecting each of these areas (Gulf Stream-associated slope water through the Northeast Channel, cooler slope water from the Labrador Current and thermal air-sea interactions). From the 2010s onwards, the main benthic warming signals in the WGOM and the EGOM were stronger but the warming in the EGOM preceded the warming in the WGOM. This suggests that the oceanographic process at play influencing the bottom temperature warming may have shifted since the 2010s. Despite the relative synchrony of warming events on the EGOM, the magnitudes of decadal trends are much lower in spring, summer, and fall than in the WGOM. One of the aspects of the warming on the EGOM is the absence of a long-term, progressive increase in temperature. Instead, bottom temperature abruptly changed twice, first in the 1970s and then in the beginning of the 2010s. This is consistent with a recent study suggesting that two major warming events have affected the entire EGOM in 1968 and 2009 (Gonçalves Neto et al., 2021). Although the NEUS is highly stratified for part of the year, GB is relatively well mixed year-round due to tidal mixing (Franks and Chen, 1996; Richaud et al., 2016). Therefore, interannual bottom temperature variations in GB are associated with both basin scale circulation (as in the EGOM and WGOM; Kavanaugh et al., 2017) and thermal air-sea interactions. The GB followed similar seasonal warming in SST with a stronger warming in the fall and weaker warming in the spring (Kleisner et al., 2017). Furthermore, similarly to SST, the strongest recorded warming in bottom temperature occurred in 2012 (Chen et al., 2020).

In the MAB, bottom temperature has warmed in all seasons but not homogeneously in time and space. During winter, the MAB experienced a significant warming compared to the other seasons. The winter warming is primarily located in southern New England and in the nearshore area from the Chesapeake Bay to the southern boundary of the MAB. The winter warming is consistent with the findings of Kavanaugh et al., (2017), which suggested benthic warming exceeding the surface warming of the MAB in the region between the Delaware and Chesapeake Bays. However, this warming in bottom temperature is unexpected given the stable SST in winter time (Northeast Fisheries Science Center (U.S.), 2022a) and the well-mixed conditions from the end of fall to the beginning of spring. During the summer, the MAB experienced the greatest bottom temperature warming in the NEUS over the last 63 years. The greatest warming occurred in southern MAB in the summer and, to a lesser spatial extent, in the spring which may be due to the northern shift in Gulf Stream (Northeast Fisheries Science Center (U.S.), 2022a, 2022b) where the current is strongest off Cape Hatteras, North Carolina and breaks off of the coastline. In the northern MAB warming may be partly due to the warming of the upstream cold water from the WGOM and GB, which supplies the near-bottom layers of the MAB (Chen et al., 2018; Fairbanks, 1982).

4.2 Performance and limitations of the combined bottom temperature product

While the spatial and temporal patterns of the combined bottom temperature product are overall consistent with other studies based on *in situ* data, we extended the assessment of its level of performance through a model bias analysis using *in situ* observations. The analysis provided valuable information regarding the gap between model estimates and observations but this bias cannot be considered as true error and the outcomes should be analyzed with caution. First, although the model resolutions are high (~7km for ROMS-NWA and 1/12° for GLORYS), the models cannot resolve the fine scale variations in bottom temperature, which are induced by the complex topography of the NEUS (notably in the WGOM) and smoothing of the ROMS-NWA bathymetry that is utilized for numerical stability. Another major issue was related to the sparsity and heterogeneity of observations over time and space (spatial distribution of observations are mapped in Appendix B and number of observations per km^2 in each EPU and season are plotted in Appendix C). To mitigate the bias due to observation heterogeneity, we defined a minimum number of observations in each EPU-season combination. This led us to withdraw observations, but strengthened our model bias analysis and made interannual comparison possible. The lack of observations and especially the drop between 1980 and 1990 and the decreasing number of observations from the 2010s onward hampered the assessment of the combined bottom temperature product for certain years. Beyond our study, the decreasing number of observations over the NEUS should be a major concern for oceanographers who rely on this data to reanalyze, calibrate and test their models (*e.g.,* Chen and Curchitser, 2020; Lellouche et al., 2021) and for managers who use them to inform benthic decadal thermal and seasonal variations (Northeast Fisheries Science Center (U.S.), 2022a, 2022b) or incorporate environmental indices in stock assessment models (*e.g.,* T. J. Miller et al., 2016). Finally, although the quantity of observations has decreased over time, their quality improved significantly with a change in the oceanic instrument used to collect bottom temperature to CTD. The latter is of higher accuracy and precision than MBT and XBT data. This is due to the higher uncertainties with both estimated depths based on XBT/MBT fall rates and temperature (Pers. com. Chris Melrose). Hence, the observations used to evaluate model bias before the 1990s are tainted with large potential measurement uncertainties and the results must be analyzed with caution.

$ROMS_{cor}$ and GLORYS12v1 exhibited good performance to reproduce the annual trend over the NEUS with a significant correlation with observations of 0.91 and 0.98 for each of these models. The bias correction effectively and largely improved the correlations from 0.70 to 0.91 and reduced the mean bias and, to a lesser extent, the RMSE. The reduction of the bias and RMSE and the higher correlation with observation time series were observed at every season and EPU. The improvement of the correlations with observed bottom temperature is especially notable in the WGOM (from 0.31 for raw ROMS-NWA to 0.72 for $ROMS_{cor}$). However, we noticed two exceptions. First, the RMSE was higher for $ROMS_{cor}$ on the EGOM but the correlations with observation time series and mean bias were improved. Second, during winter time, the mean bias and RMSE were slightly higher for $ROMS_{cor}$ but the correlation with observation time series was higher. This was consistent with the monthly bias we calculated using climatology which was close to zero during winter time and reached 1.9°C in August (Appendix A).

The bias correction method improved the model's skill in resolving the bottom temperature trend for the period 1959–1992 even if substantial annual biases persisted. Compared to GLORYS12v1, our results showed that $ROMS_{cor}$ had the largest bias in terms of magnitude and interannual variability. These large biases (*e.g.*, in 1964, 1965, 1966 and in 1971, 1973) are representative of one of the limits of our bias correction process which is based on decadal climatology. We found that the ROMS-NWA bias dramatically changed (Appendix A, Fig. A.1) from 1°C during the decade 1965–1974 to 0.5°C during the decade 1975–1984. So, the large model bias and the inversion of the sign (positive bias in 1964, 1965, 1966 and a negative bias in 1971, 1973) suggested that the decade 1965–1974 was marked by a shift in the ROMS-NWA bias, the period before 1970 being more biased than the period after. The bias correction accounted for the decadal change in bias between the decade 1965–1974 and 1975–1984 but our method could not detect the exact bias shifting point that likely occurred in 1969 (± 1 year). However, ROMS-NWA (raw model) and $ROMS_{cor}$ successfully captured the abrupt bottom temperature changes in the early 1970s (Appendix F). The greater bias during the period covered by $ROMS_{cor}$ may also be induced by the aforementioned lower quality of observations before the 1990s.

The model bias analysis showed that the combined bottom temperature product performed well in all EPUs with low mean bias and high correlation with the observed bottom temperature time series. The RMSE ranged between 1.14°C and 2.64°C for ROM_{cor}, GLORYS12v1, and PSY4V3R1. That means that biases reported for each EPU and season are smoothed over space and time but can be locally and punctually substantial. The aggregated results of the bias analysis masked also differences in the performance among seasons and EPUs. In the WGOM, we found that the combined product performs well with a little biased bottom temperature in all seasons despite some exceptions at the beginning of the time series (Appendix G) which are likely due to the above-mentioned changes in bias magnitude in the late 1960s (Appendix A). Furthermore, a large bias in 2015, 2016, 2017 in spring and fall in the WGOM can be flagged and could induce a light overestimation of warming rates at these seasons (Appendix G). On the EGOM, the shift in bias magnitude at the end of the 1960s could also be seen with large positive bias in 1966, 1967, 1970 followed by a steep decline of the bias. Furthermore, the seasonal bias analysis suggested that bottom temperature on the EGOM could be systematically warm-biased in winter and spring over the whole time series and cold-biased in summer and fall over the period 1993–2019 covered by GLORYS12v1 (Appendix G). This could lead to an underestimation of warming in summer and fall in this region. In GB, the annual bias analysis (Appendix D – Figure D2) showed that bottom temperature may be underestimated. However, the seasonal bias analysis (Appendix G) revealed that this underestimation is driven by a strong cold-bias during summer time which is, on average, of -0.9°C, while the other seasons did not exhibit such large systematic bias (moderate positive bias in winter). The maps of the mean bias on the GB in summer showed that the bias was induced by large differences between the bottom temperature estimates and observations in several shallow areas in the center of the GB and in the Nantucket shoals (Appendix H). In the MAB, the mean bias and high correlation suggested that the bottom temperature patterns were well reproduced but interannual variability is substantial especially for the period covered by $ROMS_{cor}$. The complex oceanographic dynamic, notably linked to the seasonal formation of a cold pool in the MAB (Chen and Curchitser, 2020; Lentz, 2017), makes the modeling of the region challenging and may be one of the reasons for the larger variations in bias. As in the GB and EGOM, the model bias analysis suggested that bottom temperature could be cold biased in winter in the MAB (Appendix G).

Finally, because of the scarcity and spatiotemporal heterogeneity of observations, we were unable to assess the skill of the combined bottom temperature product for some years on the NEUS annually and seasonally in each EPU. We partially filled this gap by examining individually model bias in each season and EPU (Appendix G) where the number of observations met the requirement we set (see section 2.7). Specifically, the limited observations in 2015, 2016, and 2017 did not allow us to evaluate the annual bias for the entire NEUS, in some seasons and EPUs. Based on the EPUs and seasons wherein sufficient observations are available for these three years, we found that bottom temperature may have been greatly overestimated in 2015 in the WGOM in spring and fall, on the EGOM in spring and in the MAB in summer (+2.5°C). Conversely, bottom temperature may have underestimated in summer 2015 on the GB and EGOM and in fall 2016 on the GB and MAB.

In order to explore model bias where observations were limited between 2014 and 2021, we used bottom temperature data from the National Oceanic and Atmospheric Administration (NOAA) electronic Monitoring of Lobster Traps (eMOLT) program (Appendix I). In this program, bottom temperature is measured using sensors attached to lobster pots collected hourly temperature data. The number of observations widely varied seasonally, over the time series and spatially with the majority of observations inshore and on the shelf break (Appendix I – Figure I.1). Model bias based on the eMOLT observations was analyzed separately from the other observations because of the abovementioned high spatiotemporal heterogeneity of observations and the observational uncertainty due to potential monitoring inaccuracies. These inaccuracies cannot be quantified but neither can they be ignored and underestimated (discussed for example in Li et al., 2017). In 2015, 2016 and 2017 where observations are too limited to calculate the model bias (Figure 6), the mean annual bias on the NEUS calculated based on the eMOLT is relatively low ranging between +0.3 in 2015 to -0.3°C in 2017 (Appendix I – Figure I.2). In 2020 where observations are also limited due to the COVID pandemic, the mean bias based on the eMOLT exceeded +0.8°C. That means that bottom temperature could be underestimated in 2020 even though this result must be interpreted with caution.

In the study, we included bottom temperature from PSY4V3R1 that provided near-real-time data. Our model bias analysis suggested that PSY4V3R1 performed well although a large model bias more than -1°C which was found on the GB in 2021. Unfortunately, due to the COVID crisis, only a few observations are available in 2020 to further evaluate the performance of PSY4V3R1. In order to provide some insights, we compared daily estimates from PSY4V3R1 and GLORYS12v1 in 2019, which is the only year they have in common (Appendix J). We found that the daily estimates from the two products were very similar for all EPUs in 2019. The main difference was the late summer warming on the GB that occurred later and was steeper for PSY4V3R1 (Appendix J). In parallel, we compared the estimates of each of the two products with observations in 2019 and found that the mean bias and RMSE were very similar (Appendix J). Even if this finding cannot fully reflect the performance of PSY4V3R1 (compared to GLORYS12v1) because it was focused on one single year, it provides support for the relevance of using PSY4V3R1 to extend the bottom temperature product to near-real time.

4.3 Applications for fisheries and marine ecology

The combined bottom temperature product may be valuable for a range of applications in fisheries, marine ecology, and oceanography in the region. It provides a long-term high-resolution bottom temperature time series to investigate warming impacts from the late 1950s to near-present day. The combined product could also provide key quantitative information regarding the thermal state of benthic environments over multiple decades. Seasonal and annual bottom temperature anomalies, demersal marine heatwaves, and oceanographic features (*e.g.,* the MAB Cold Pool) could be quantified routinely every year over the NEUS. The development of fishery stock assessments that integrate environmental effects was defined as a research priority in the NEUS (Hare et al., 2016). While several fish stocks have identified links between bottom temperature and population processes (*e.g.,* yellowtail flounder, Atlantic cod; see Introduction), challenges persist around data availability and quality to operationalize a stock assessment framework that includes environmental effects. The first prerequisite is to establish and characterize the relationship between temperature and a population process. The spatiotemporal resolution and scale of the combined bottom temperature product would offer the possibility to focus on specific areas, seasons, and time periods associated with a particular population or life history process (*e.g.,* Atlantic American Lobster; Tanaka et al., 2019) and Atlantic sea scallop (Tanaka et al., 2020; Zang et al., 2022). Bottom temperature time series could be incorporated routinely into a stock assessment. Another benefit for environment-linked fisheries assessment models is to provide near-real-time benthic thermal conditions that can be used to update historical time series and, in some cases, improve short-term population forecasts. Specifically, for some population dynamics, the lag between thermal conditions and the population can allow short-term population forecasts using near-real-time bottom temperature. For instance, thermal conditions at the year n can affect the survival and habitat of yellowtail flounder during early life stages that impact, in turn, the recruitment of the year $n+1$ (du Pontavice et al., 2022; T. J. Miller et al., 2016) while winter bottom temperature has an effect on the juvenile survival of the MAB black sea bass (*Centropristis striata*) that likely affects its recruitment (A. S. Miller et al., 2016; Younes et al., 2020).

In addition to the length of the time series, the continuous nature of the combined product across space and time could be critical for fisheries and ecological applications. Although observed data are extensive over the NEUS, their high heterogeneity over time and space (see seasonal and annual observation distribution in Appendix B) can be limiting to study local and/or seasonal ecological processes. For instance, the study of the effects of the MAB Cold Pool could benefit from the combined bottom temperature product. The MAB Cold Pool is a seasonally formed cold water mass that occurs from late spring to early fall (Chen et al., 2018; Chen and Curchitser, 2020; Lentz, 2017). Because of the lack of observations during the Cold Pool season or modeled data over their period of interest, some studies used fall observations bottom trawl survey to quantify the interannual Cold Pool intensity that can lead to a restricted perception of the Cold Pool impacts (Friedland et al., 2022; T. J. Miller et al., 2016). Another limitation that can bias the Cold Pool perception is that the observations fully rely on the NEFSC fall survey that can be limited or even canceled as in 2017 due to a vessel mechanical failure or in 2020 due to the COVID-19 pandemic. Therefore, a combined bottom temperature product based on ROMS-NWA and GLORYS12v1 was developed to study the Cold Pool effects on the yellowtail flounder recruitment and incorporate them into a stock assessment model (du Pontavice et al., 2022). Here we refined and extended a bottom temperature product developed in du Pontavice et al. (2022) and validated it against ocean observations. Our goal was to develop and test a high-resolution time-series that could be used for a wide range of applications in oceanography, marine ecology, and living marine resource management.

Acknowledgments

We thank Laura Nazzaro for providing a code to develop the ROMS-NWA bias-correction method and Chris Melrose for providing observations data from the Northeast Fisheries Science Center (NEFSC) oceanographic database. Funding for Hubert du Pontavice was provided by the NOAA NEFSC's "New England's Groundfish in a Changing Climate" program.

Open Research

The GLORYS12v1 ocean reanalysis and the Operational Mercator global ocean analysis and forecast system (PSY4V3R1) datasets are available at the Copernicus Marine Environment Monitoring Service (CMEMS):

- https://resources.marine.copernicus.eu/product-detail/GLOBAL_MULTIYEAR_PHY_001_030/INFORMATION
- https://resources.marine.copernicus.eu/product-detail/GLOBAL_ANALYSIS_FORECAST_PHY_001_024/INFORMATION.

The bottom temperature ROMS-NWA output will be shared on reasonable request to Zhuomin Chen (zhuomin.chen@uconn.edu).

The Northwest Atlantic Ocean regional climatology is available at the NCEI's (https://www.ncei.noaa.gov/products/northwest-atlantic-regional-climatology).

Bottom temperature from the Northeast Fisheries Science Center (NEFSC) oceanographic database will be shared on reasonable request to Hubert du Pontavice (hubert.dupontavice@princeton.edu).

Bottom temperature from NOAA NCEI's World Ocean Database is available at the NCEI's (https://www.ncei.noaa.gov/products/world-ocean-database).

Appendix A: Bias correction of NWA-ROMS

Conversion of monthly bias into daily bias

Based on the monthly ROMS-NWA bias estimated using the NWARC climatologies, we calculated a daily bias by fitting generalized additive models (GAM) for each decade in each grid cell. The dependent variable was the monthly bias ($Bias(x_R,y_R,t_{dec},t_m)$) and the predictor variable was the month (Each monthly bias was assigned to the middle of the corresponding month such as “0.5=January”, “1.5=February”, …, “11.5=December”). Then, we rescaled the calendar day of each year to range between 0 and 12 (as “Jan 1=(1*12)/365= 0.0329”, “Jan 2=(2*12)/365= 0.0658”, …, “Dec 32 = (365*12)/365 =12”) and predicted the bias for each day and decade using the fitted GAM in each 1/10° NWARC grid cell.

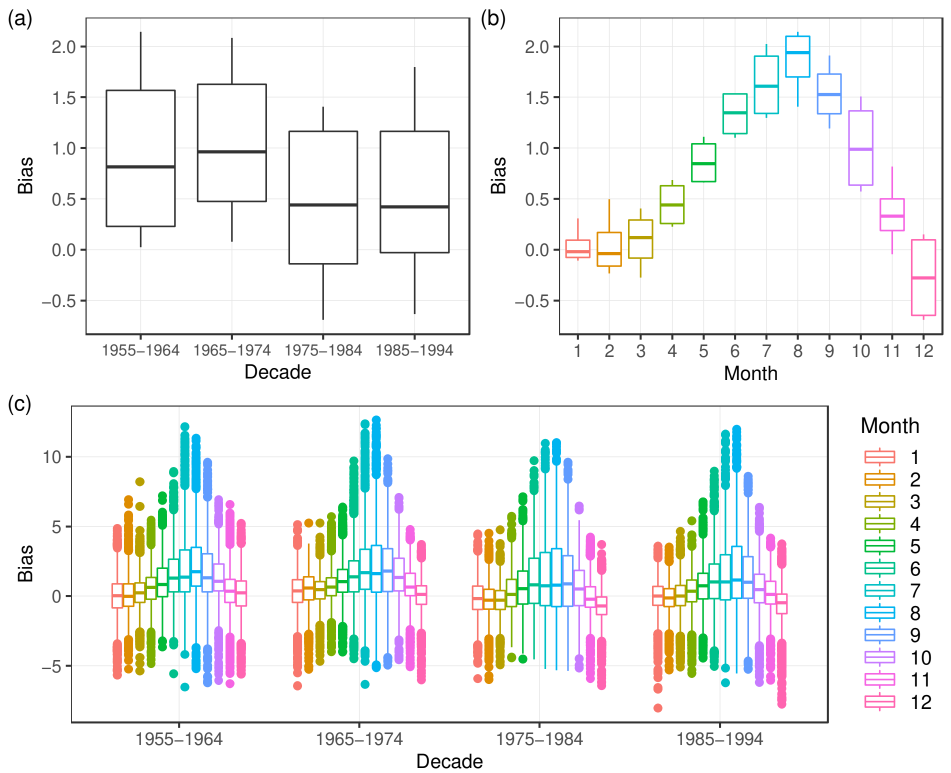


Figure A.1. Decadal bias of bottom temperature in the Northeast U.S. continental shelf between ROMS-NWA and the NWA-climatology. Panels (a) and (b) represent the distribution of the mean decadal (a) and monthly (b) bias estimates while panel (c) represents the raw bias estimates for each month and each decade in each grid cell. The dots in panel c are the outliers within the data distribution for each month and decade.

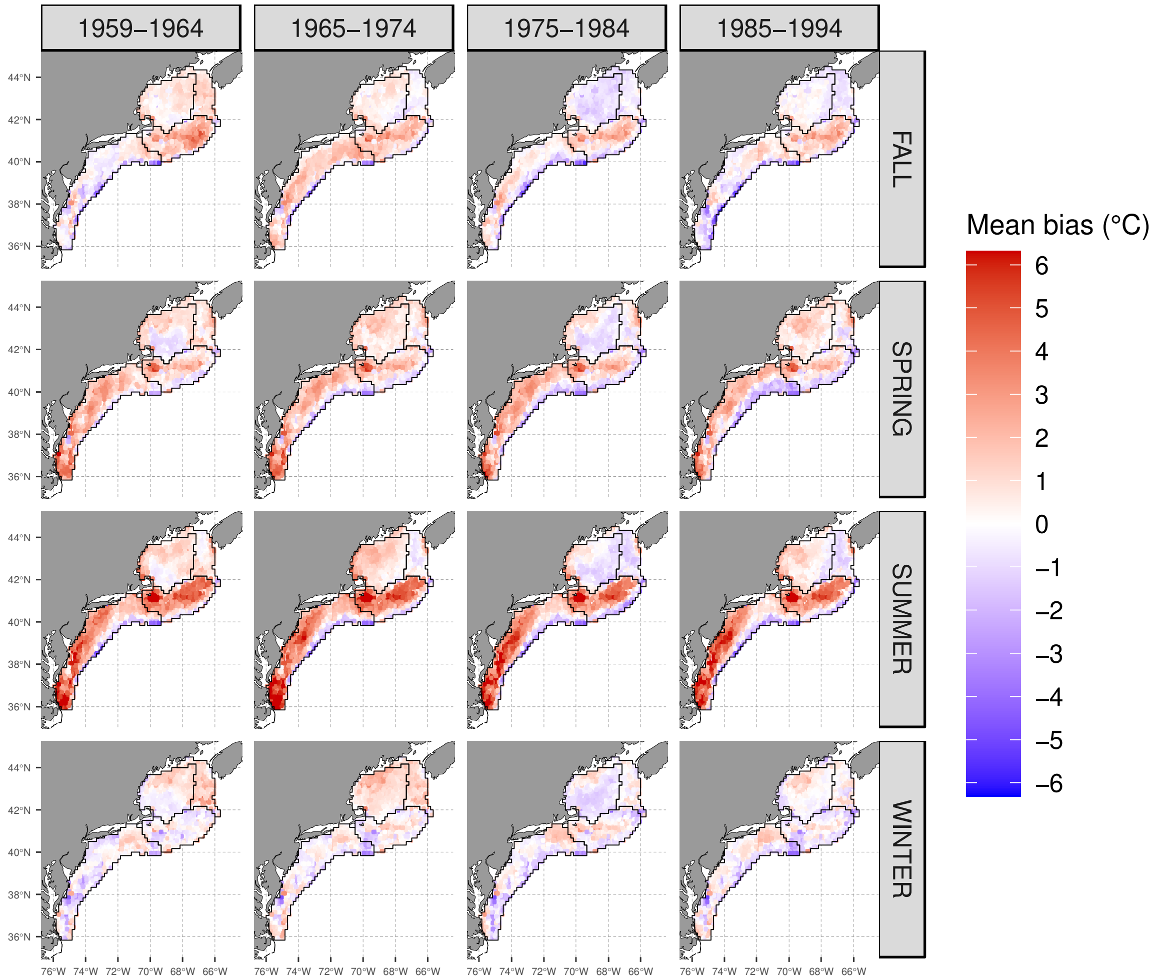


Figure A.2. Maps of the mean seasonal bias of bottom temperature between ROMS-NWA and the NWA-climatology on the Northeast U.S. continental shelf in a 1/10° horizontal grid.

Appendix B: spatial distribution of observations

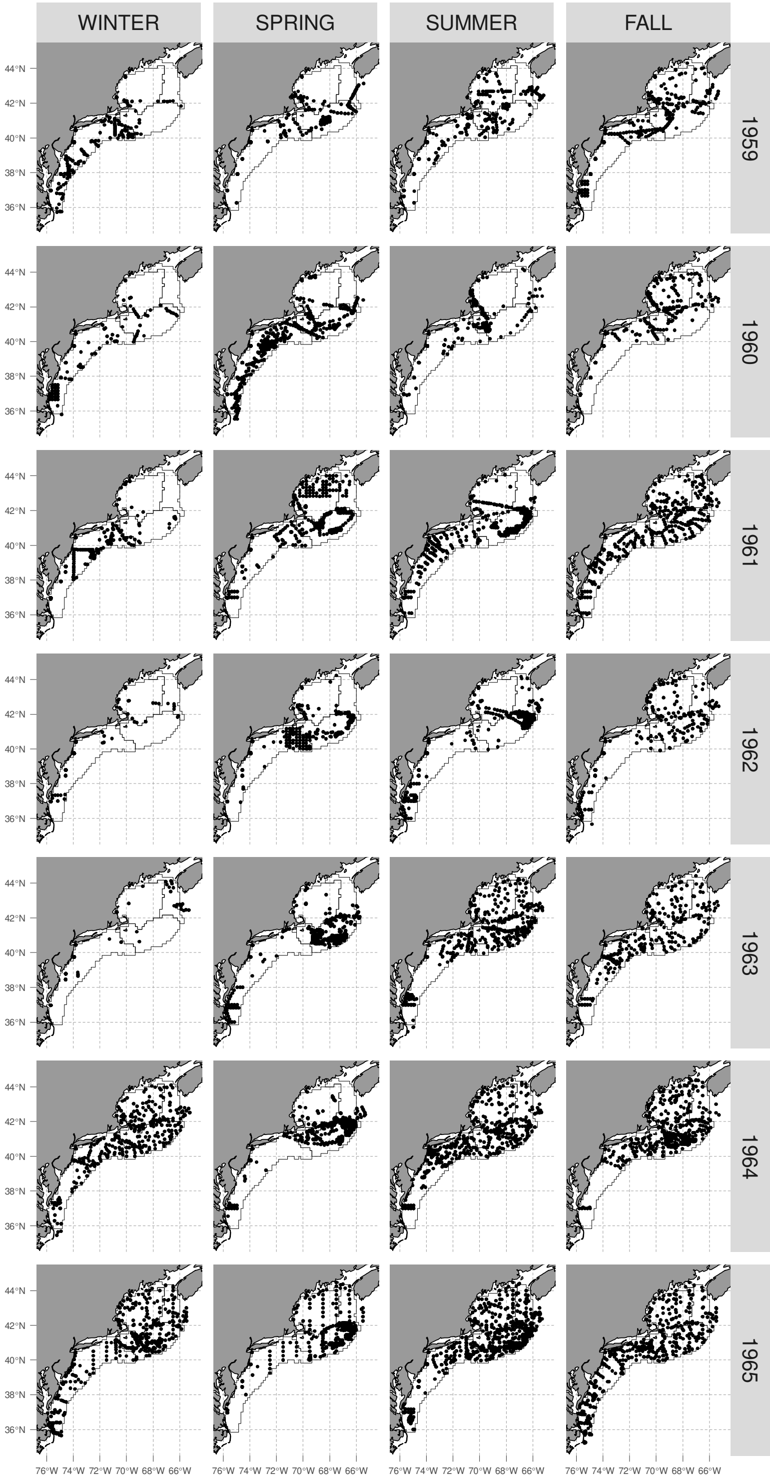
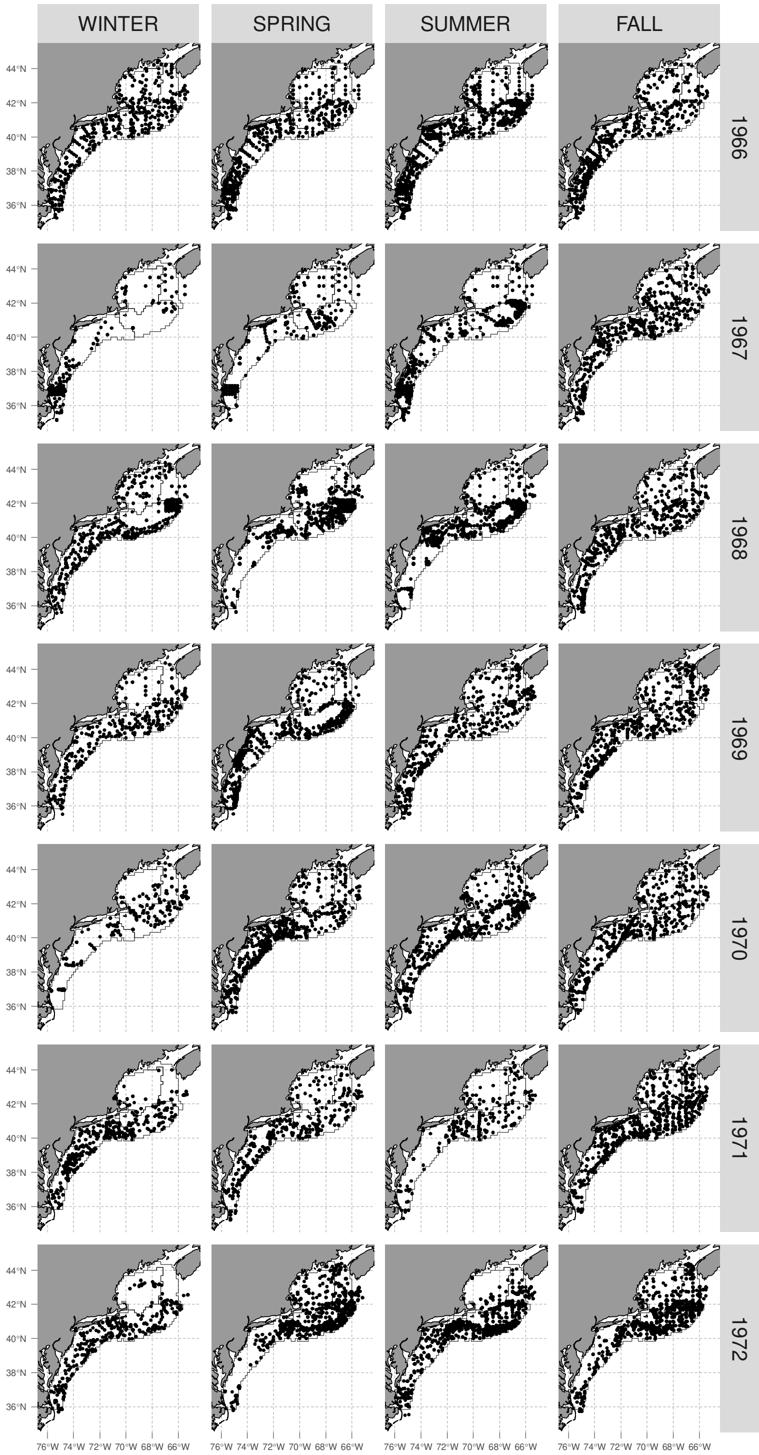
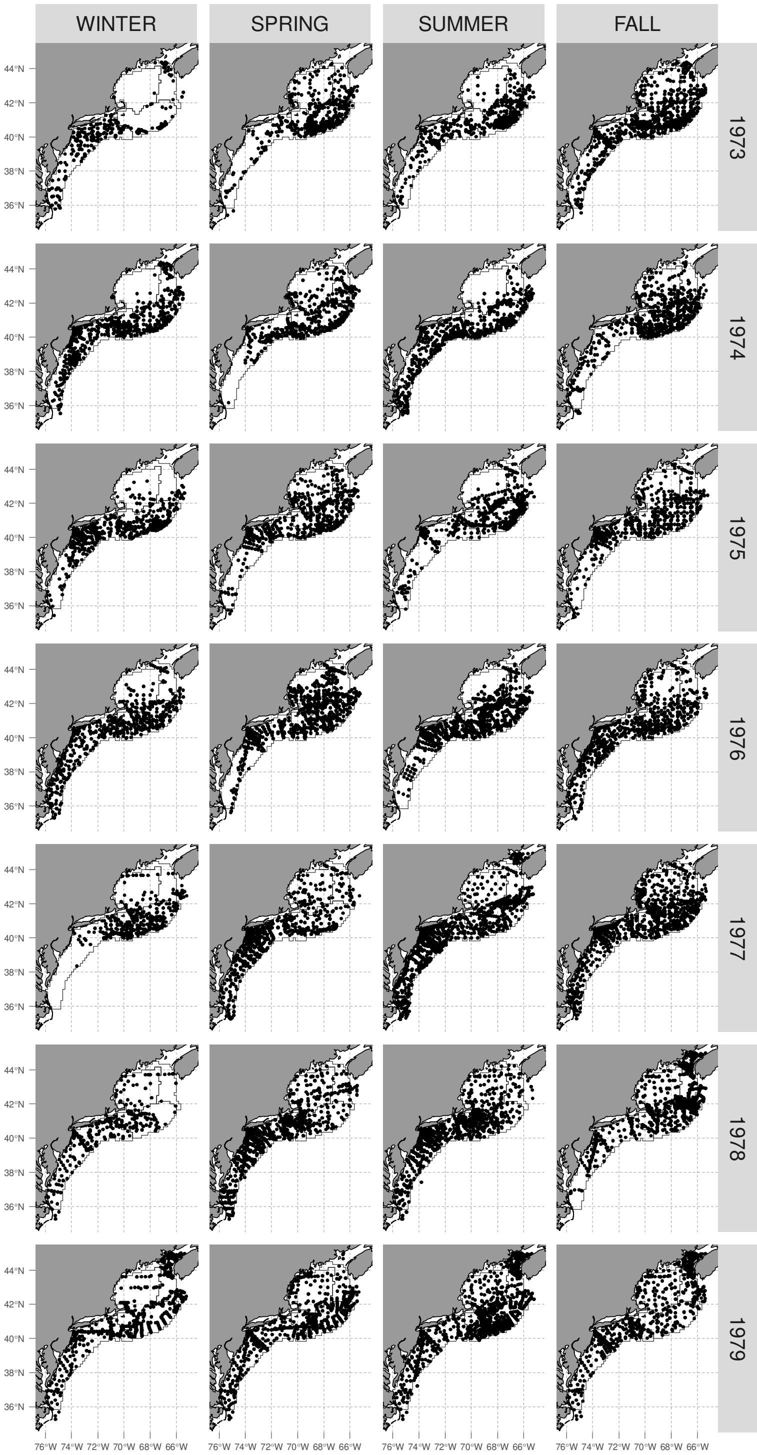


Figure B.1. Spatial distribution of the bottom temperature observations for each season between 1959 and 1965.



796
797 **Figure B.2.** Spatial distribution of the bottom temperature observations for each season between
798 1966 and 1972



799
800 **Figure B.3.** Spatial distribution of the bottom temperature observations for each season between
801 1973 and 1979.

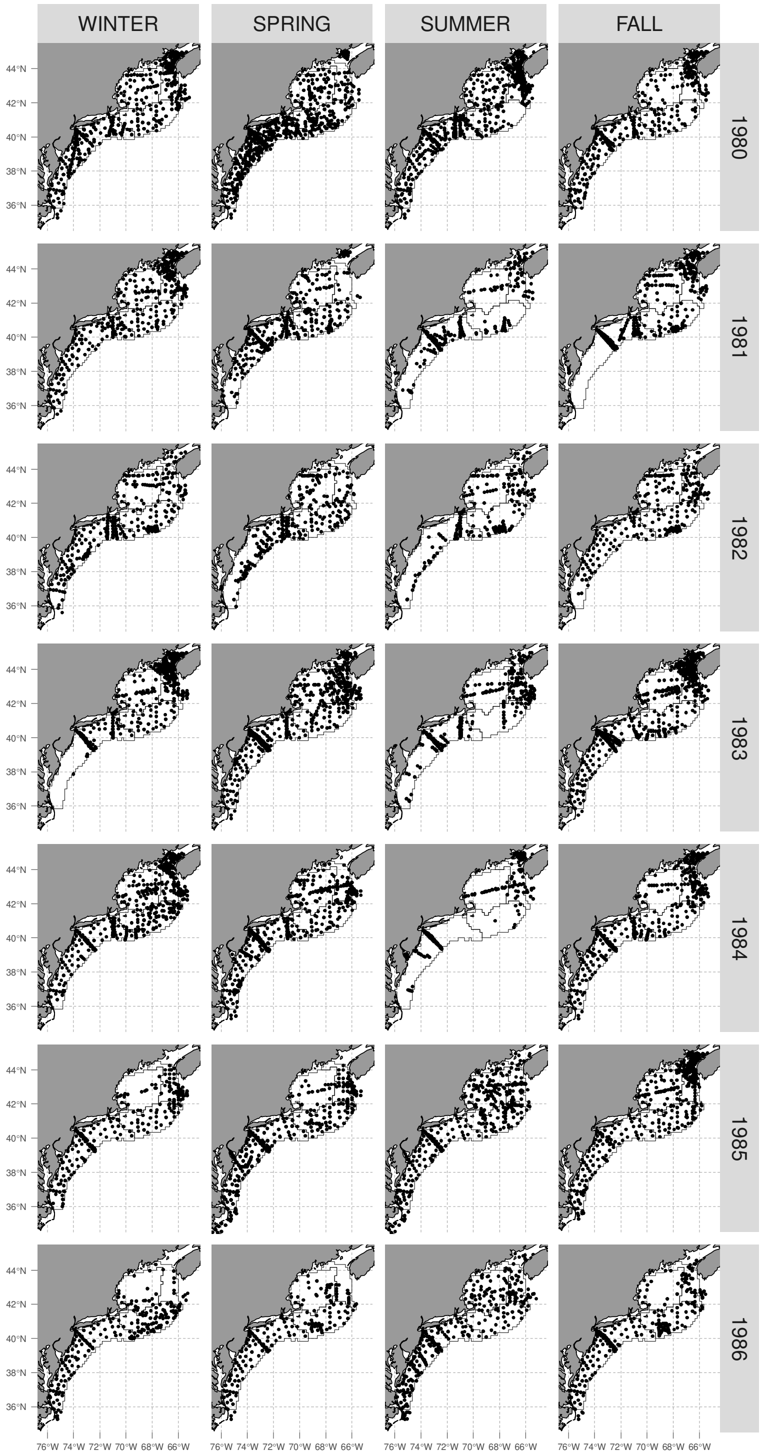


Figure B.4. Spatial distribution of the bottom temperature observations for each season between 1980 and 1986.

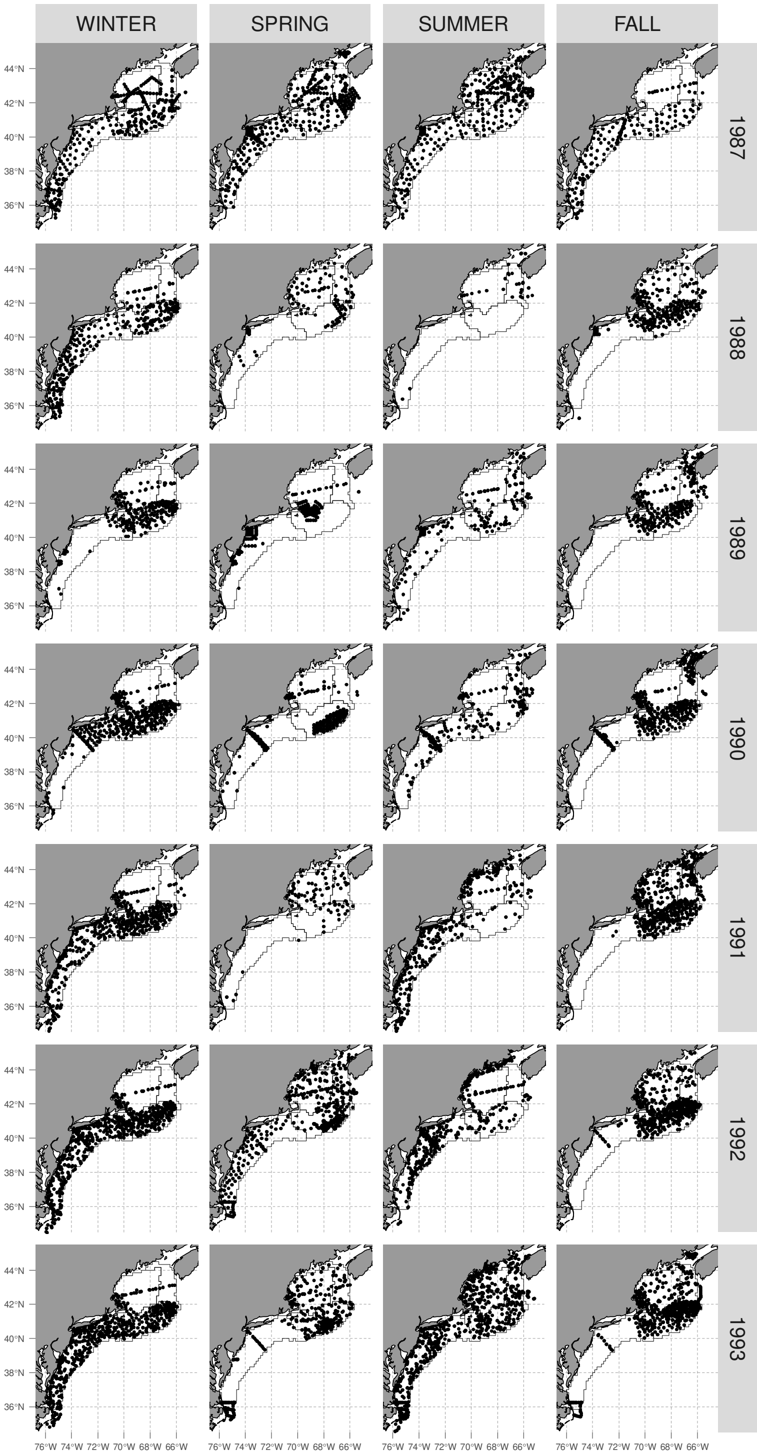


Figure B.5. Spatial distribution of the bottom temperature observations for each season between 1987 and 1993.

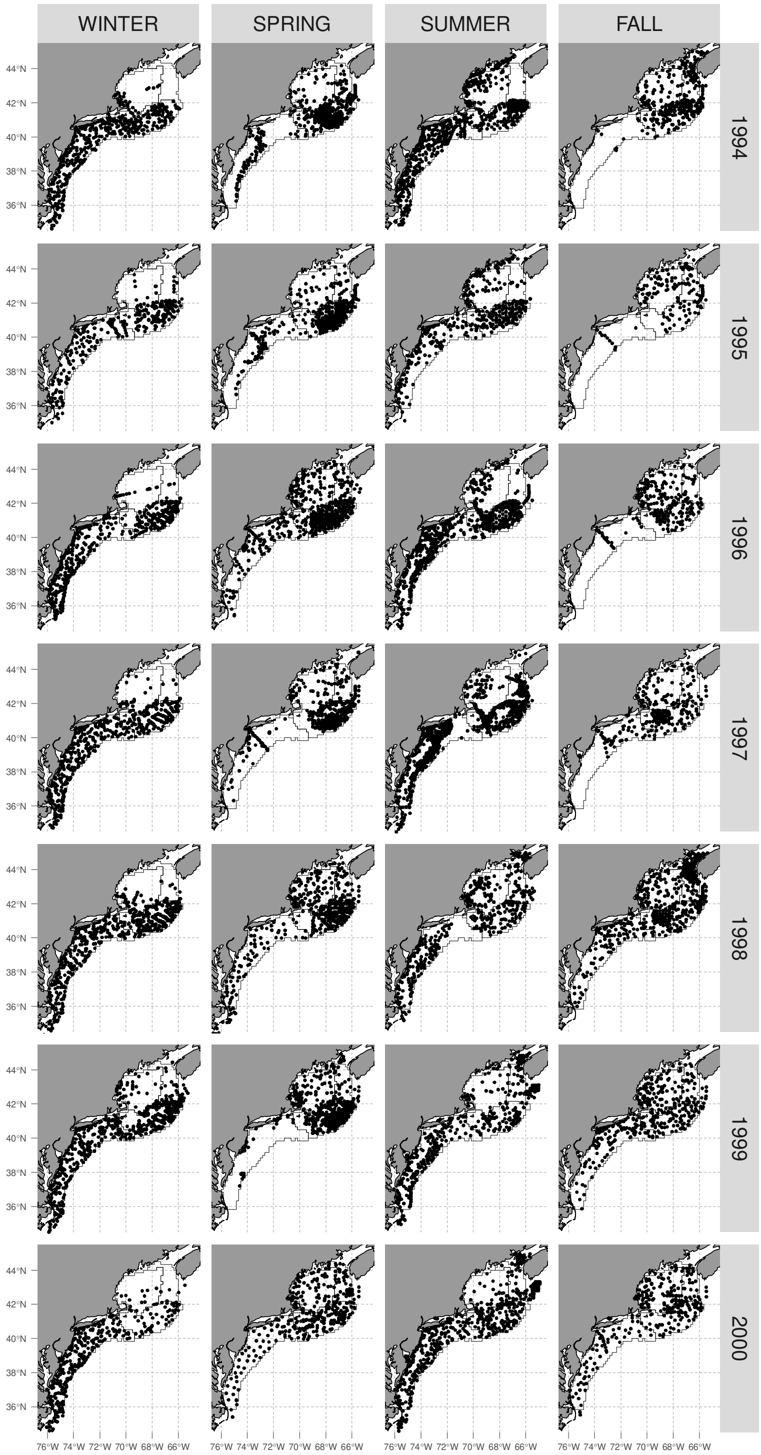


Figure B.6. Spatial distribution of the bottom temperature observations for each season between 1994 and 2000.

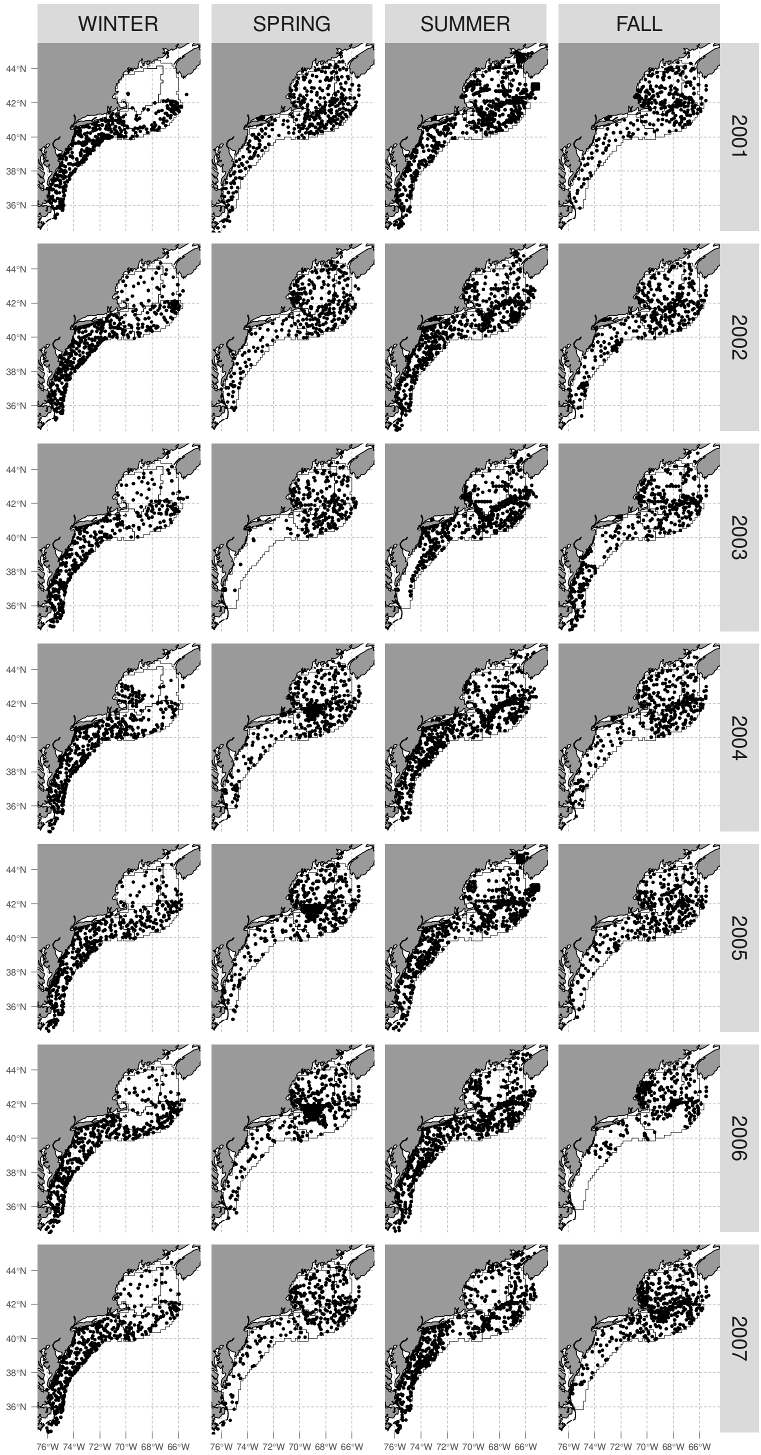
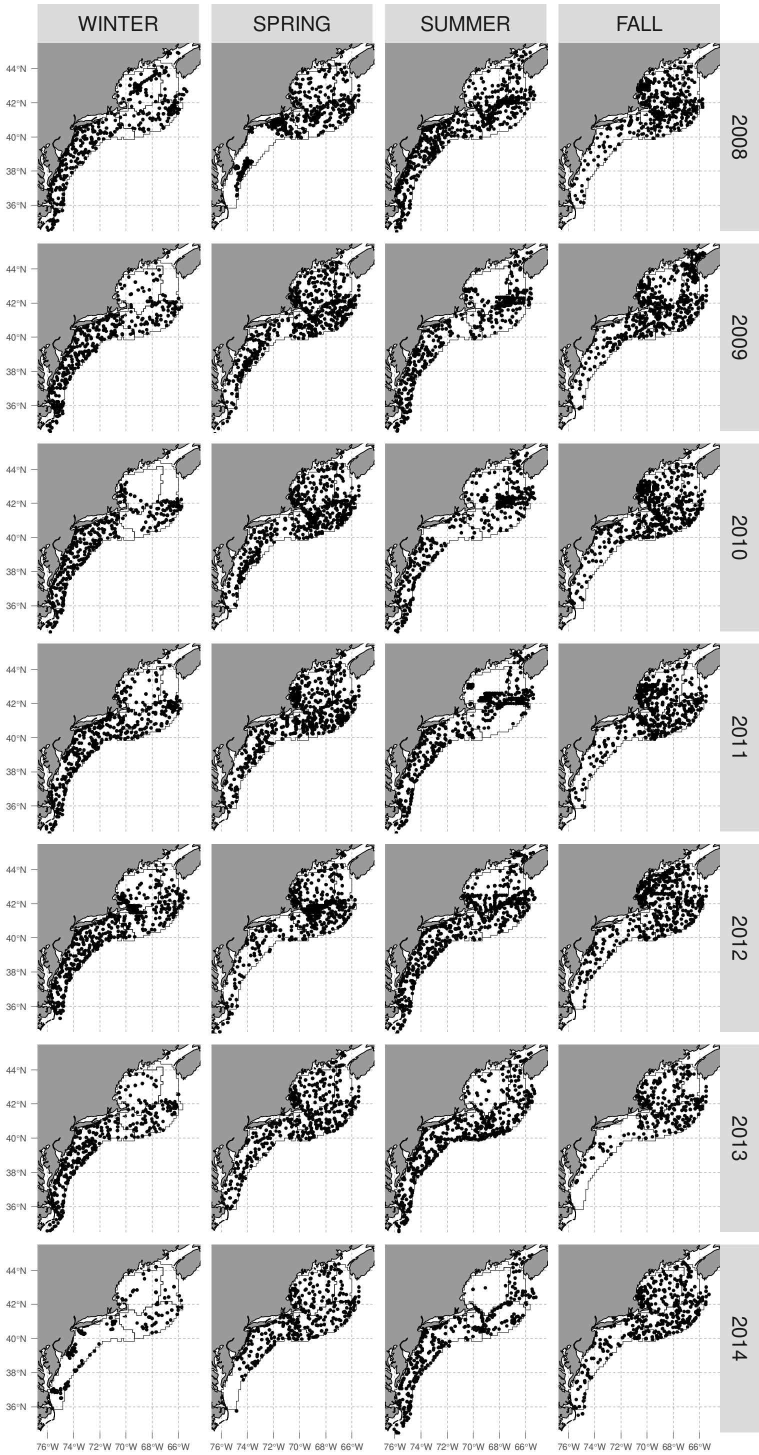


Figure B.7. Spatial distribution of the bottom temperature observations for each season between 2001 and 2007.



814
815 **Figure B.8.** Spatial distribution of the bottom temperature observations for each season between
816 2008 and 2014.

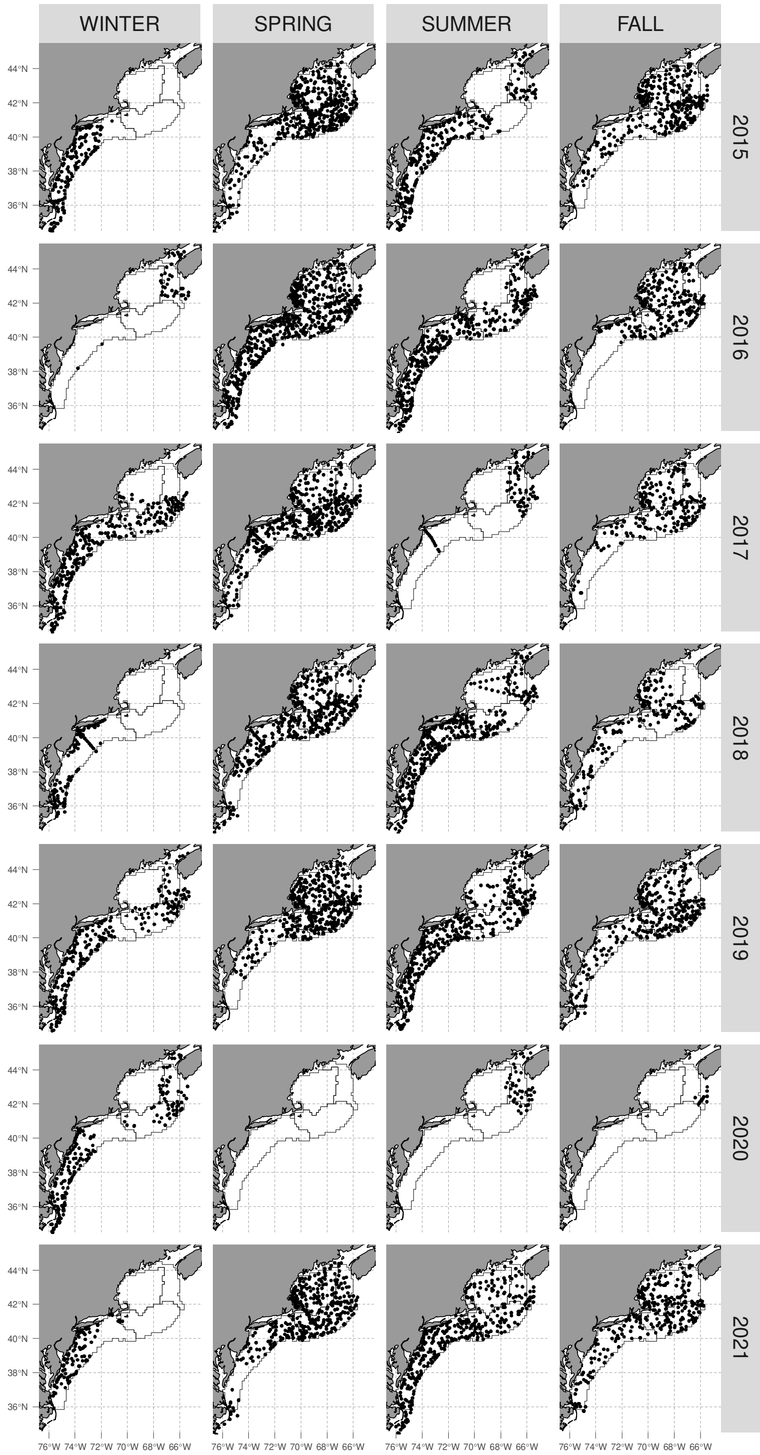


Figure B.9. Spatial distribution of the bottom temperature observations for each season between 2015 and 2021.

Appendix C: Number of observations per km² in each Ecological Production Unit and season

Figure C.1. Times series of the number of *in situ* observations per km² each in Ecological Production Unit and season.

Appendix D: Mean model bias in bottom temperature in each season and Ecological Production Unit between 1959 and 2021

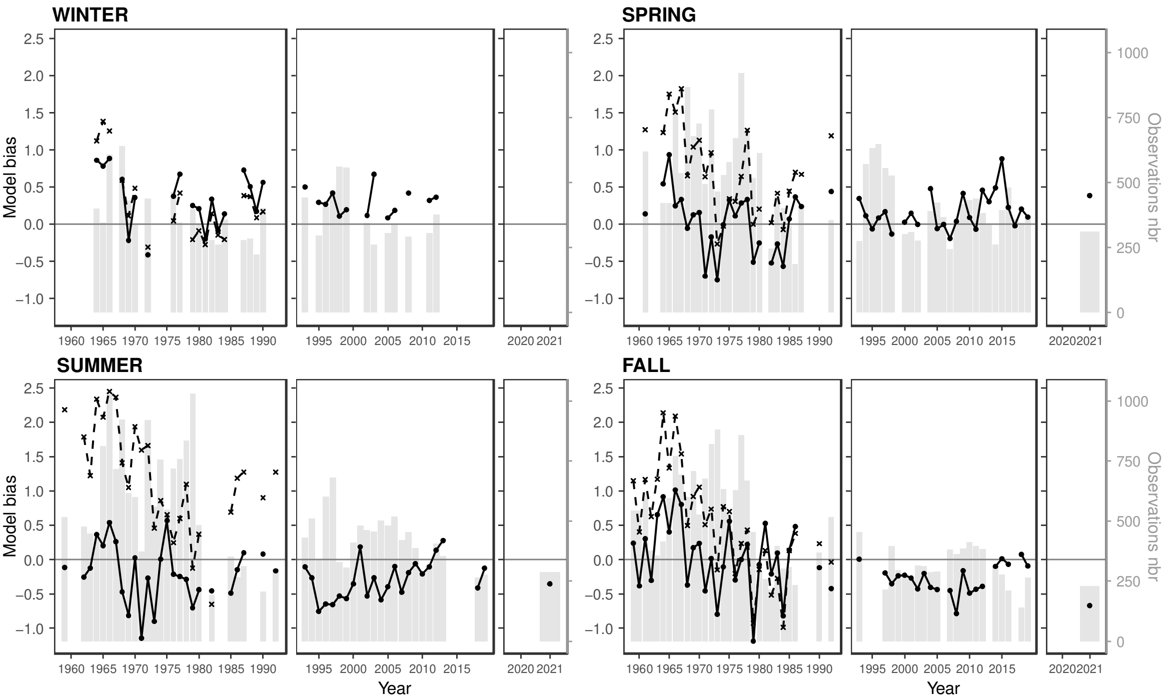


Figure D.1. Seasonal mean model bias in bottom temperature of the raw ROMS-NWA between 1959 and 1992 (dashed line, left panel), the bias-corrected ROMS-NWA between 1959 and 1992 (solid line, left panel), GLORYS12v1 between 1993 and 2019 (centered panel) and PSY4V3R1 in 2020 and 2021 (right panel). The barplots represent the number of observations (Observations nbr).

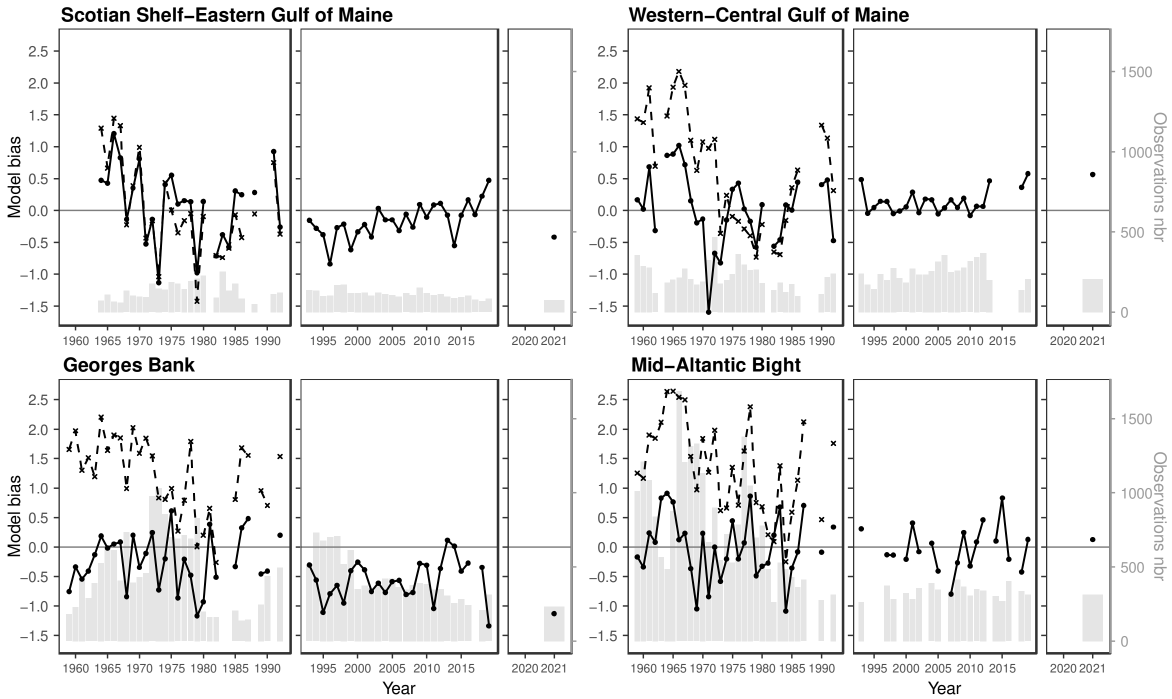


Figure D.2. Mean model bias in bottom temperature in each Ecological Production Unit (EPU) of the raw NWA-ROMS between 1959 and 1992 (dashed line, left panel), the bias-corrected NWA-ROMS between 1959 and 1992 (solid line, left panel), GLORYS12v1 between 1993 and 2019 (centered panel) and PSY4V3R1 in 2020 and 2021 (right panel). The mean annual model bias is calculated for spring, summer and fall. The bar plots represent the number of observations (Observations nbr).

Appendix E: Concurrency of the warming trends

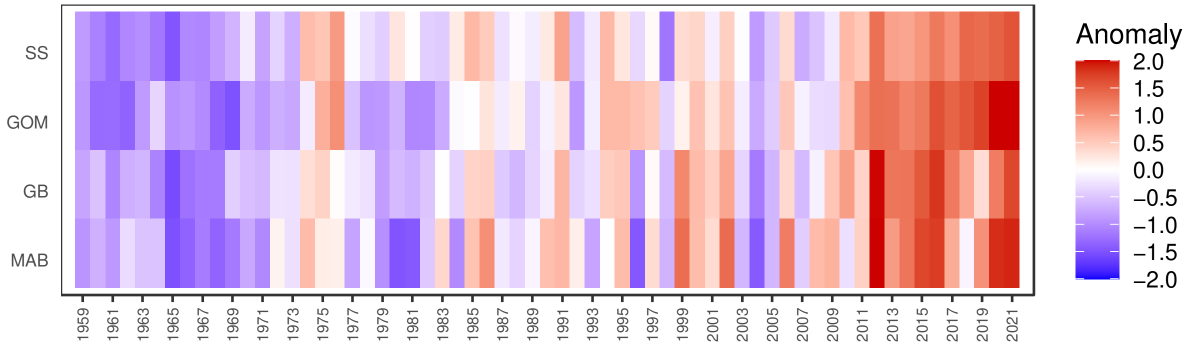


Figure E.1. Annual anomalies in bottom temperature in each Ecological Production Units between 1959 and 2021.

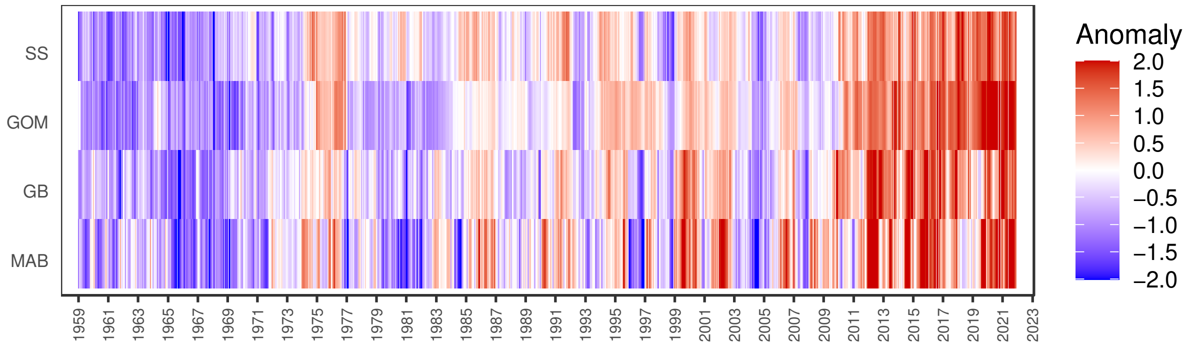


Figure E.2. Monthly anomalies in bottom temperature in each Ecological Production Units between 1959 and 2021.

Appendix F: NWA-ROMS and $ROMS_{cor}$ times series

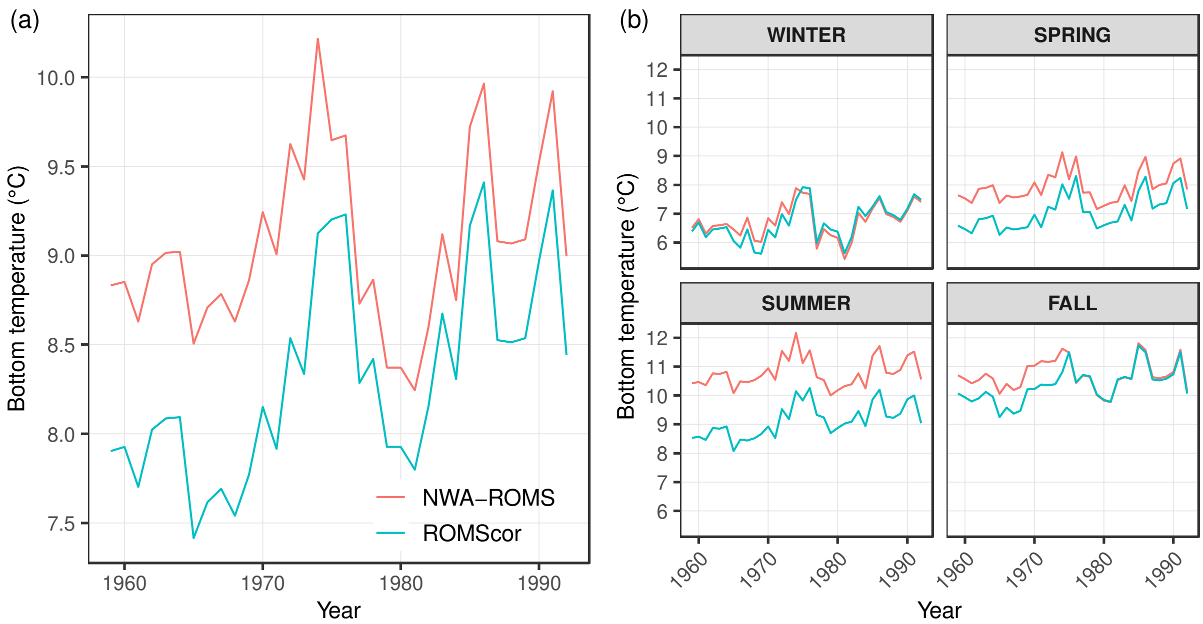
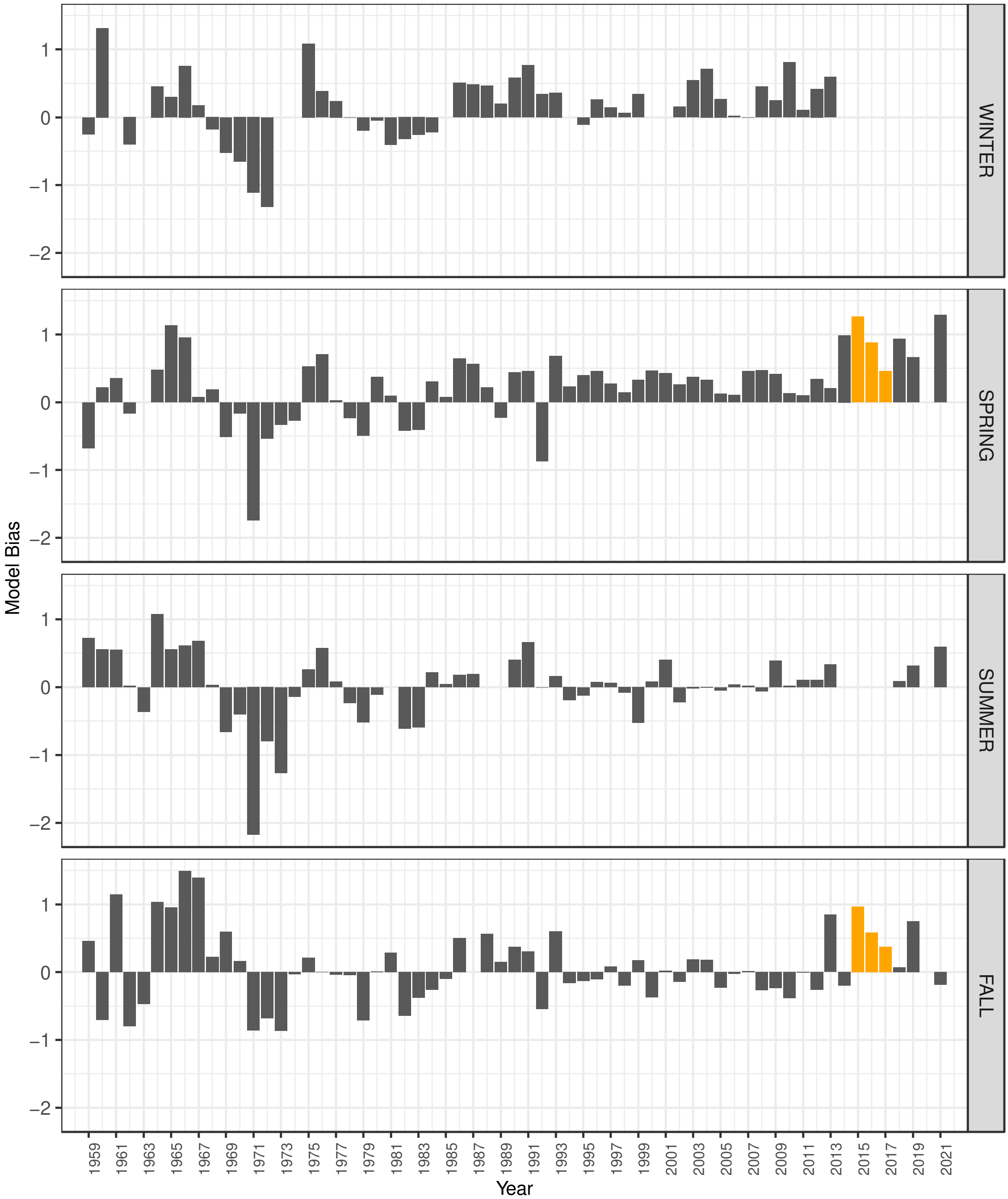


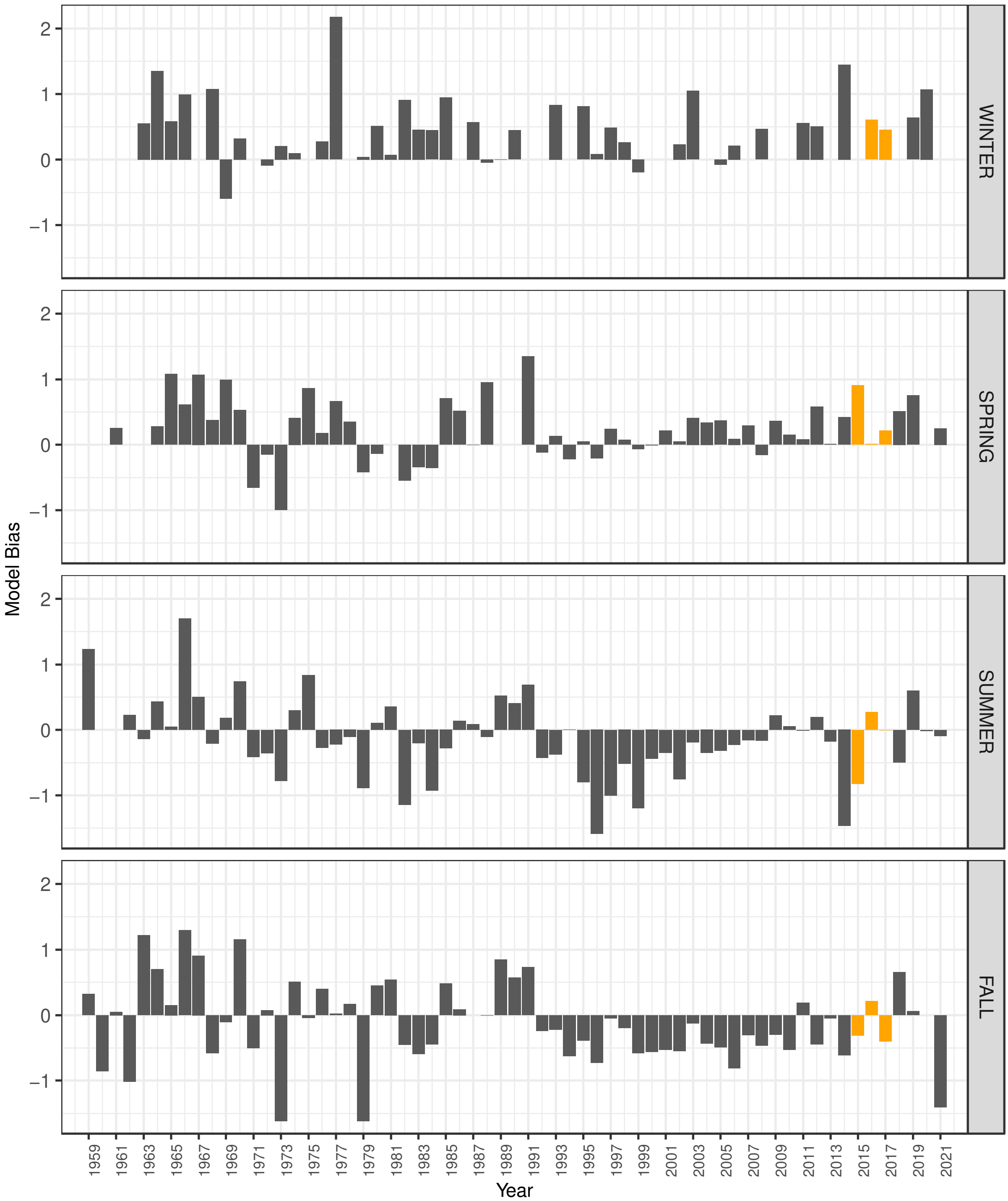
Figure F.1. Mean annual bottom temperature time series on the Northeast U.S. continental shelf between 1959 and 1992 for $ROMS_{cor}$ and NWA-ROMS

855 **Appendix G: Seasonal mean model bias in each EPUs**



856
857 **Figure G.1.** Seasonal mean model bias in the WGOM in bottom temperature between 1959 and
858 2021.

859



860
861 **Figure G.2.** Seasonal mean model bias on the EGOM in bottom temperature between 1959 and
862 2021.

863

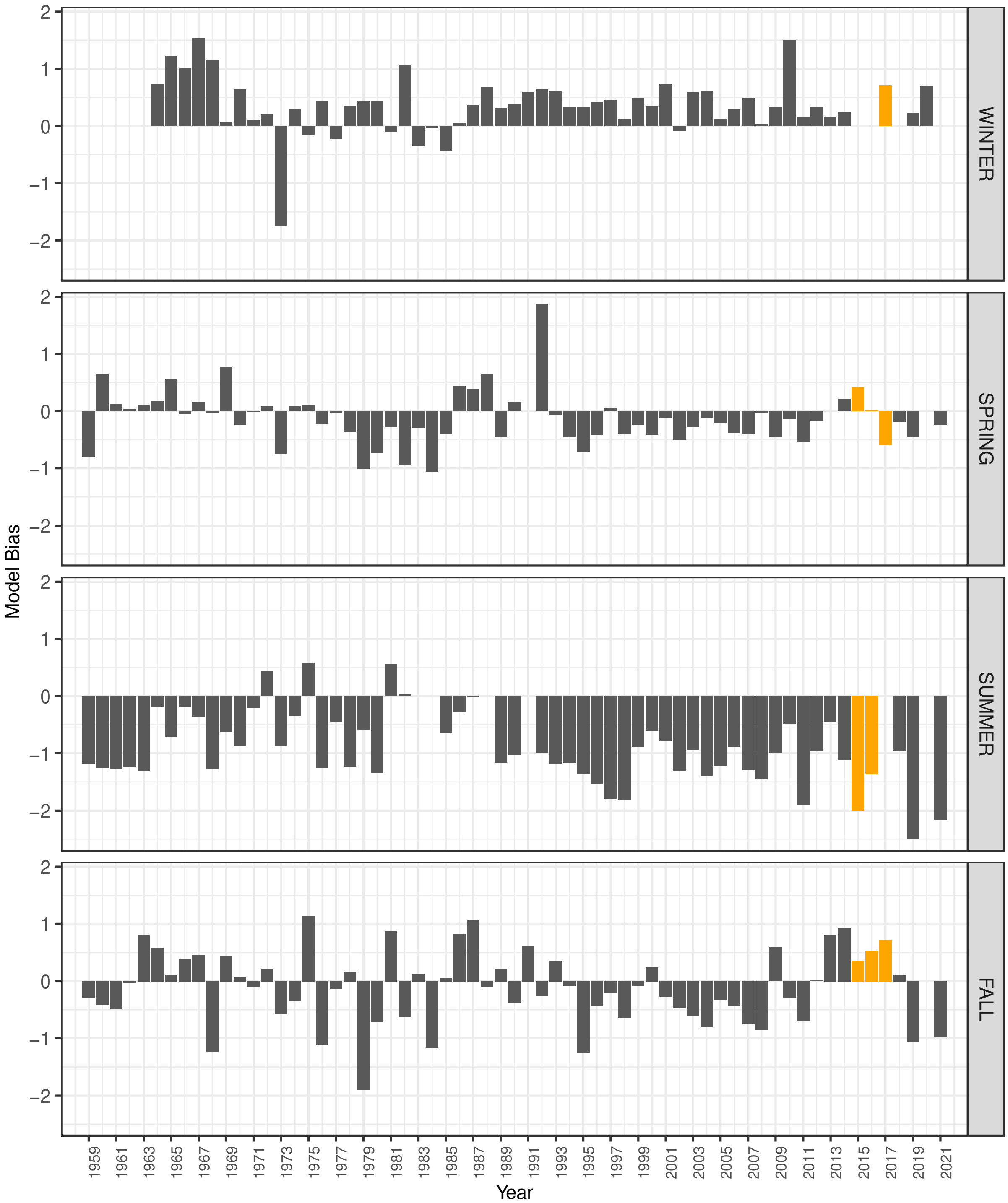


Figure G.3. Seasonal mean model bias on the GB in bottom temperature between 1959 and 2021.

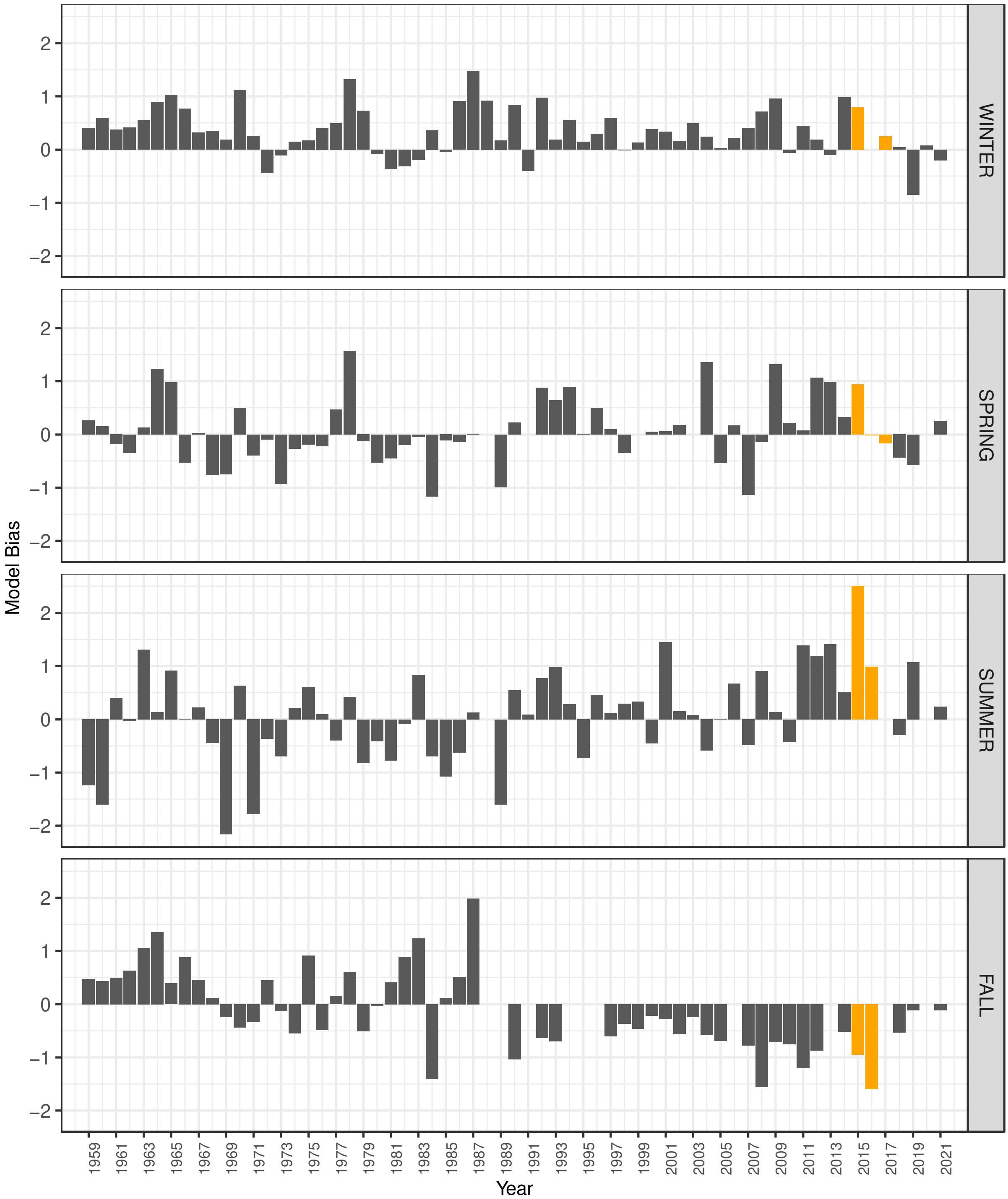
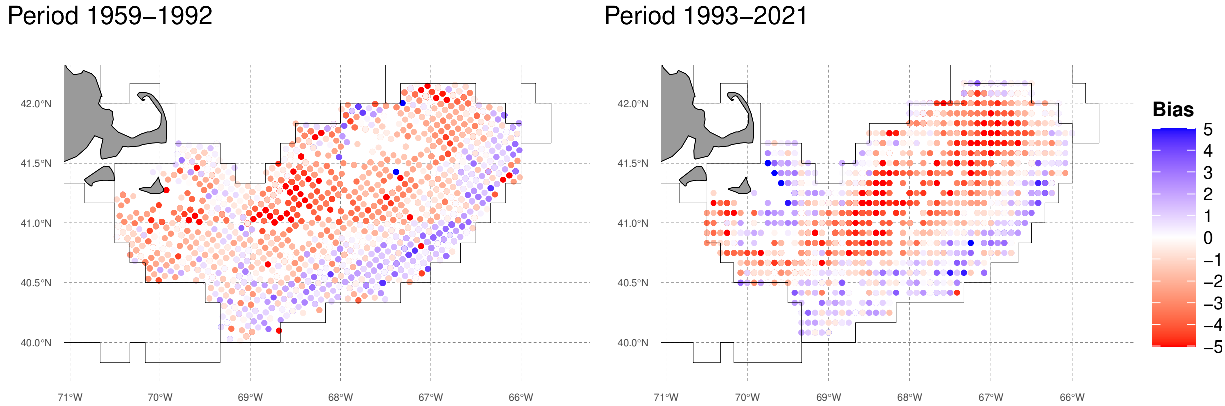


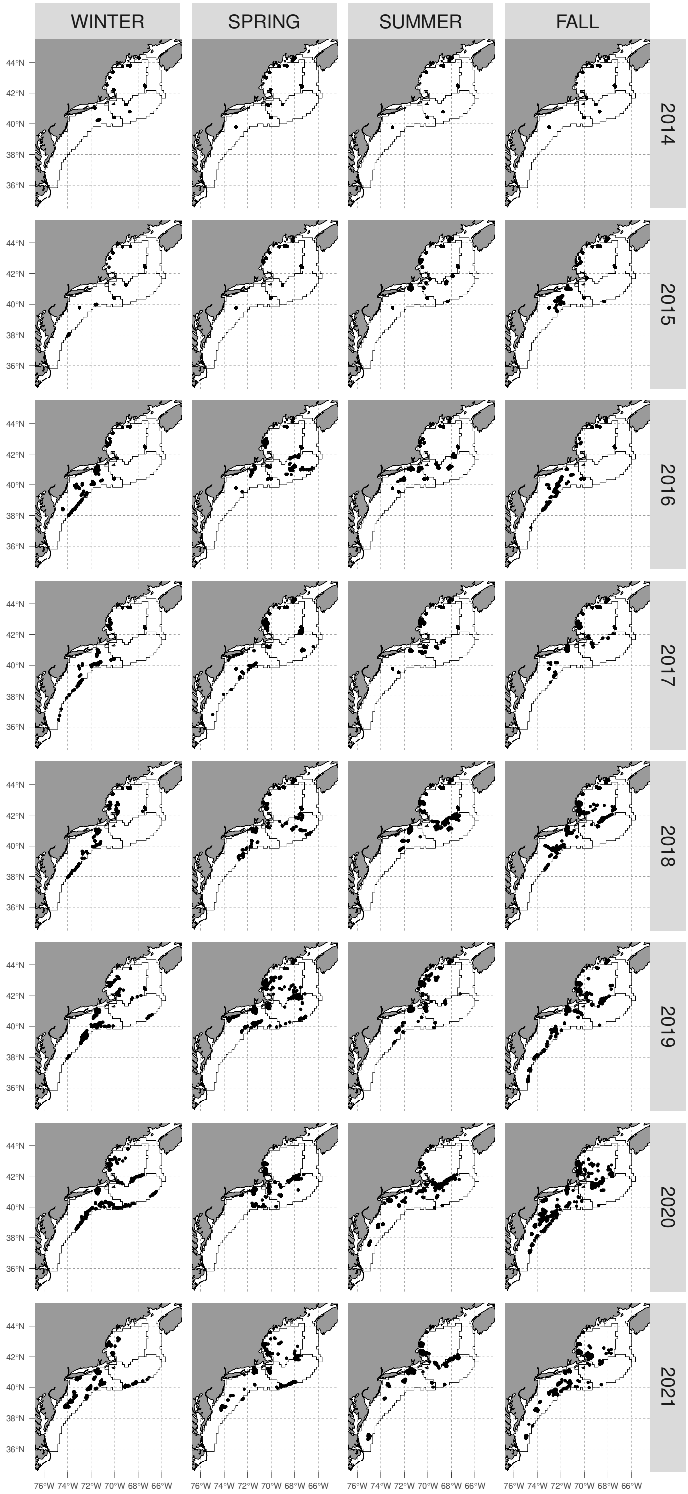
Figure G.4. Seasonal mean model bias in the MAB in bottom temperature between 1959 and 2021.

872 **Appendix H: Maps of the mean bias during summer in the Georges Bank**



873
874 **Figure H.1.** Maps of the mean bias (bottom temperature product vs observation) of bottom
875 temperature in summer on the Georges Bank for the period 1959-1992 covered by ROM_{cor} and
876 for the period 1993-2021 covered by GLORYS12v1 and PSY4V3R1.

877 **Appendix I: Model bias exploration using eMOLT observations**



878

Figure I.1. Spatial distribution of the bottom temperature observations from eMOLT for each season between 2014 and 2021.

Model bias exploration - eMOLT

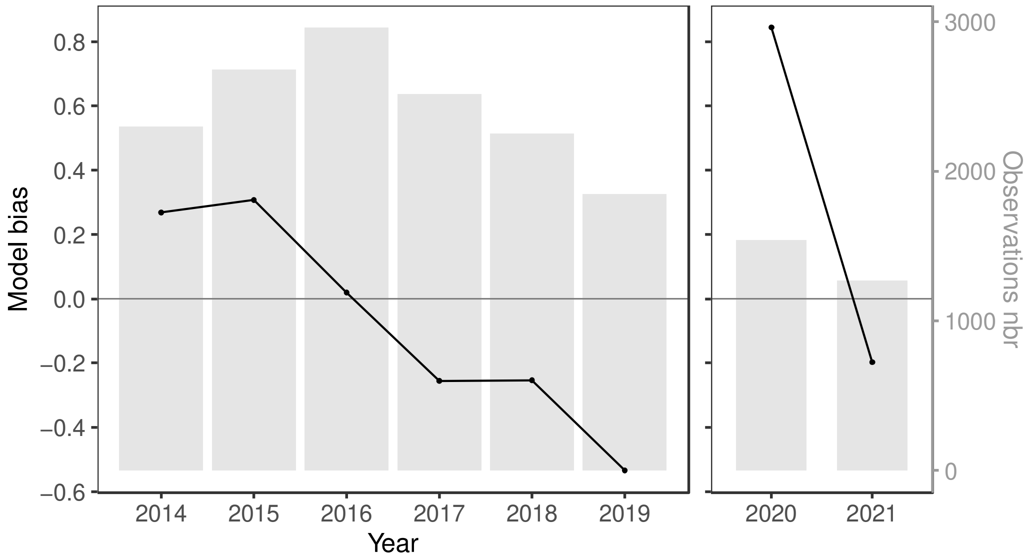


Figure I.2. Annual mean model bias in bottom temperature of GLORYS12v1 between 2014 and 2019 (centered panel) and PSY4V3R1 in 2020 and 2021 (right panel) using EMOLT observations. The bar plots represent the number of observations (Observations nbr).

Table I.1. Annual mean bias and RMSE for GLORYS12v1 (2014–2019) and PSY4V3R1 (2020–2021) using eMOLT observations (eMOLT) and the observations presented in the core of the paper (NEFSC & NCEI).

	GLORYS12v1		**PSY4V3R1**	
	NEFSC & NCEI	eMOLT	NEFSC & NCEI	eMOLT
Mean bias	-0.02	-0.09	-0.02	0.44
RMSE	1.78	1.90	1.90	2.14

Appendix J: Bottom temperature comparison from PSY4V3R1 and GLORYS12v1 in 2019

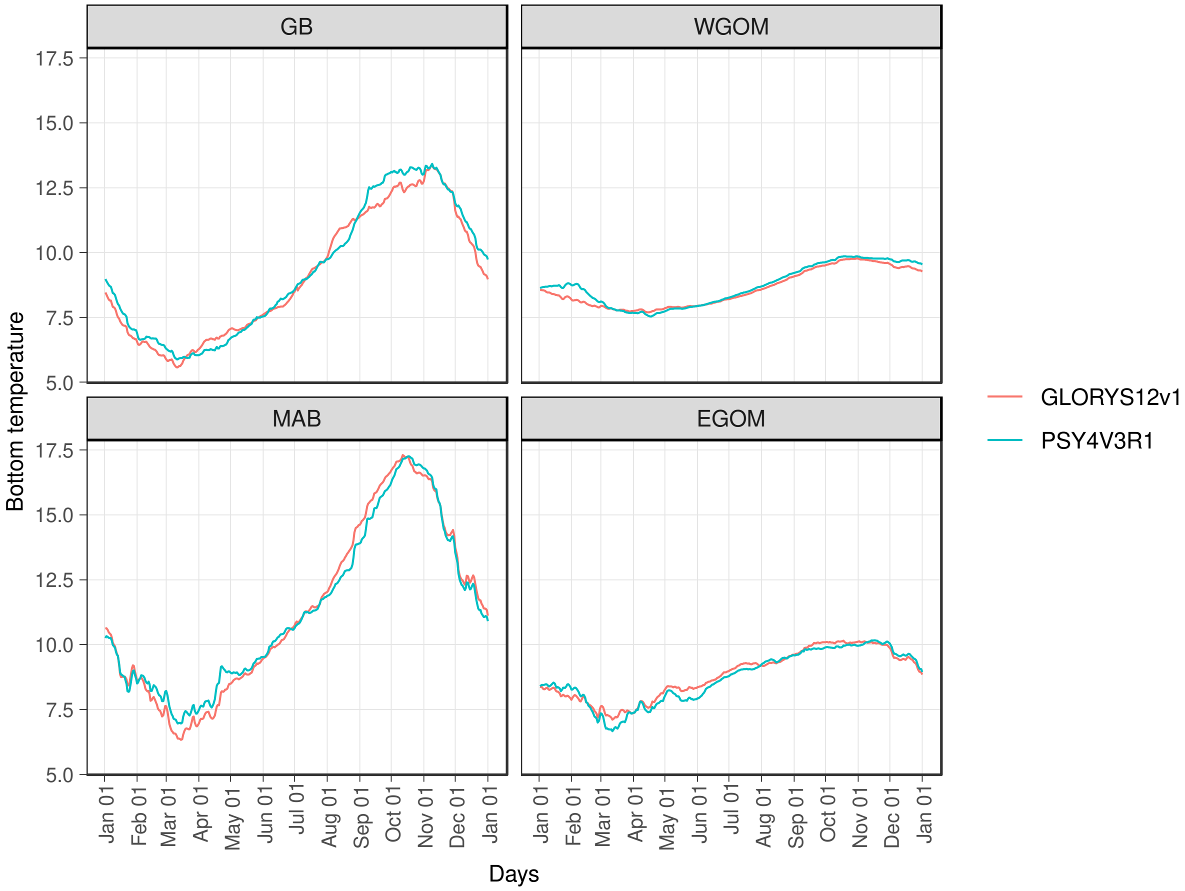


Figure J.1. Daily bottom temperature trends of GLORYS12v1 and PSY4V3R1 in 2019 in each of the four Ecological Production Units on the Northeast U.S. continental shelf.

Table J.1. Mean bias and RMSE in each Ecological Production Unit for GLORYS12v1 and PSY4V3R1 in 2019. The figures in brackets represent the number of observations used to calculate the mean bias and RMSE.

	Season	GLORYS12V1	PSY4V3R1
Mean bias	*EGOM (101)*	0.58	0.34
	WGOM (217)	0.65	0.70
	GB (347)	-0.88	-0.84
	MAB (412)	0.13	0.15
RMSE	*EGOM (101)*	1.39	1.19
	WGOM (217)	1.43	1.43
	GB (347)	2.46	2.44
	MAB (412)	2.54	2.26

Reference

Bell, R.J., Wood, A., Hare, J., Richardson, D., Manderson, J., Miller, T., 2018. Rebuilding in the face of climate change. Can. J. Fish. Aquat. Sci. 75, 1405–1414. https://doi.org/10.1139/cjfas-2017-0085

Boyer, T.P., Baranova, O.K., Coleman, M.A., Garcia, H.E., Grodsky, R.A., Locarnini, R.A., Mishonov, A.V., Paver, C.R., Reagan, J.R., Seidov, D., Smolyar, I.V., Weathers, K.A., Zweng, M.M., 2018. World Ocean Database 2018, A.V. Mishonov, Technical Ed., NOAA Atlas NESDIS 87.

Brickman, D., Hebert, D., Wang, Z., 2018. Mechanism for the recent ocean warming events on the Scotian Shelf of eastern Canada. Continental Shelf Research 156, 11–22. https://doi.org/10.1016/j.csr.2018.01.001

Cabanes, C., Grouazel, A., von Schuckmann, K., Hamon, M., Turpin, V., Coatanoan, C., Paris, F., Guinehut, S., Boone, C., Ferry, N., de Boyer Montégut, C., Carval, T., Reverdin, G., Pouliquen, S., Le Traon, P.-Y., 2013. The CORA dataset: validation and diagnostics of in-situ ocean temperature and salinity measurements. Ocean Sci. 9, 1–18. https://doi.org/10.5194/os-9-1-2013

Carton, J.A., Giese, B.S., 2008. A Reanalysis of Ocean Climate Using Simple Ocean Data Assimilation (SODA). Monthly Weather Review 136, 2999–3017. https://doi.org/10.1175/2007MWR1978.1

Chang, J., Hart, D.R., Munroe, D.M., Curchitser, E.N., 2021. Bias Correction of Ocean Bottom Temperature and Salinity Simulations From a Regional Circulation Model Using Regression Kriging. J. Geophys. Res. Oceans 126. https://doi.org/10.1029/2020JC017140

Chapman, D.C., Beardsley, R.C., 1989. On the Origin of Shelf Water in the Middle Atlantic Bight. J. Phys. Oceanogr. 19, 384–391. https://doi.org/10.1175/1520-0485(1989)019<0384:OTOOSW>2.0.CO;2

Chen, Z., Curchitser, E., Chant, R., Kang, D., 2018. Seasonal Variability of the Cold Pool Over the Mid-Atlantic Bight Continental Shelf. J. Geophys. Res. Oceans 123, 8203–8226. https://doi.org/10.1029/2018JC014148

Chen, Z., Curchitser, E.N., 2020. Interannual Variability of the Mid-Atlantic Bight Cold Pool. J. Geophys. Res. Oceans 125. https://doi.org/10.1029/2020JC016445

Chen, Z., Kwon, Y., Chen, K., Fratantoni, P., Gawarkiewicz, G., Joyce, T.M., 2020. Long-Term SST Variability on the Northwest Atlantic Continental Shelf and Slope. Geophysical Research Letters 47. https://doi.org/10.1029/2019GL085455

Chen, Z., Kwon, Y., Chen, K., Fratantoni, P., Gawarkiewicz, G., Joyce, T.M., Miller, T.J., Nye, J.A., Saba, V.S., Stock, B.C., 2021. Seasonal Prediction of Bottom Temperature on the Northeast U.S. Continental Shelf. J. Geophys. Res. Oceans 126. https://doi.org/10.1029/2021JC017187

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q.J.R. Meteorol. Soc. 137, 553–597. https://doi.org/10.1002/qj.828

du Pontavice, H., Miller, T.J., Stock, B.C., Chen, Z., Saba, V.S., 2022. Ocean model-based covariates improve a marine fish stock assessment when observations are limited. ICES Journal of Marine Science fsac050. https://doi.org/10.1093/icesjms/fsac050

Ezraty, R., Girard-Ardhuin, F., Piollé, J.-F., Kaleschke, L., Heygster, G., 2007. Arctic and Antarctic sea ice concentration and Arctic sea ice drift estimated from Special Sensor Microwave data. Département d'Océanographie Physique et Spatiale, IFREMER, Brest, France and University of Bremen Germany.

Fairbanks, R.G., 1982. The origin of continental shelf and slope water in the New York Bight and Gulf of Maine: Evidence from $H_2{}^{18}O/H_2{}^{16}O$ ratio measurements. J. Geophys. Res. 87, 5796. https://doi.org/10.1029/JC087iC08p05796

Fernandez, E., Lellouche, J.M., 2018. Product user manual for the global ocean physical reanalysis product GLORYS12V1. Copernicus Product User Manual 4, 1–15.

Franks, P.J.S., Chen, C., 1996. Plankton production in tidal fronts: A model of Georges Bank in summer. issn: 0022-2402 54, 631–651. https://doi.org/10.1357/0022240963213718

Friedland, K.D., Miles, T., Goode, A.G., Powell, E.N., Brady, D.C., 2022. The Middle Atlantic Bight Cold Pool is warming and shrinking: Indices from in situ autumn seafloor temperatures. Fisheries Oceanography 31, 217–223. https://doi.org/10.1111/fog.12573

Friedland, K.D., Morse, R.E., Manning, J.P., Melrose, D.C., Miles, T., Goode, A.G., Brady, D.C., Kohut, J.T., Powell, E.N., 2020. Trends and change points in surface and bottom thermal environments of the US Northeast Continental Shelf Ecosystem. Fish Oceanogr 29, 396–414. https://doi.org/10.1111/fog.12485

Friedland, K.D., Smoliński, S., Tanaka, K.R., 2021. Contrasting patterns in the occurrence and biomass centers of gravity among fish and macroinvertebrates in a continental shelf ecosystem. Ecol Evol ece3.7150. https://doi.org/10.1002/ece3.7150

Fuchs, H.L., Chant, R.J., Hunter, E.J., Curchitser, E.N., Gerbi, G.P., Chen, E.Y., 2020. Wrong-way migrations of benthic species driven by ocean warming and larval transport. Nat. Clim. Chang. 10, 1052–1056. https://doi.org/10.1038/s41558-020-0894-x

Gawarkiewicz, G., Todd, R., Zhang, W., Partida, J., Gangopadhyay, A., Monim, M.-U.-H., Fratantoni, P., Malek Mercer, A., Dent, M., 2018. The Changing Nature of Shelf-Break Exchange Revealed by the OOI Pioneer Array. Oceanog 31, 60–70. https://doi.org/10.5670/oceanog.2018.110

Gonçalves Neto, A., Langan, J.A., Palter, J.B., 2021. Changes in the Gulf Stream preceded rapid warming of the Northwest Atlantic Shelf. Commun Earth Environ 2, 74. https://doi.org/10.1038/s43247-021-00143-5

Hardison, S., Perretti, C.T., DePiper, G.S., Beet, A., 2019. A simulation study of trend detection methods for integrated ecosystem assessment. ICES Journal of Marine Science 76, 2060–2069. https://doi.org/10.1093/icesjms/fsz097

Hare, J.A., Borggaard, D.L., Friedland, K.D., Anderson, J., Burns, P., Chu, K., Clay, P.M., Collins, M.J., Cooper, P., Fratantoni, P.S., Johnson, M.R., Manderson, J.P., Milke, L., Miller, T.J., Orphanides, C.D., Saba, V.S., 2016. Northeast Regional Action Plan – NOAA Fisheries Climate Science Strategy (No. NMFS-NE-239). NOAA Fisheries. Northeast Fisheries Science Center, Woods Hole, MA 100.

Kang, D., Curchitser, E.N., 2013. Gulf Stream eddy characteristics in a high-resolution ocean model: Gulf Stream Eddy Characteristics. J. Geophys. Res. Oceans 118, 4474–4487. https://doi.org/10.1002/jgrc.20318

Kang, D., Curchitser, E.N., Rosati, A., 2016. Seasonal Variability of the Gulf Stream Kinetic Energy. Journal of Physical Oceanography 46, 1189–1207. https://doi.org/10.1175/JPO-D-15-0235.1

Kavanaugh, M.T., Rheuban, J.E., Luis, K.M.A., Doney, S.C., 2017. Thirty-Three Years of Ocean Benthic Warming Along the U.S. Northeast Continental Shelf and Slope: Patterns, Drivers, and Ecological Consequences. J. Geophys. Res. Oceans 122, 9399–9414. https://doi.org/10.1002/2017JC012953

Kleisner, K.M., Fogarty, M.J., McGee, S., Hare, J.A., Moret, S., Perretti, C.T., Saba, V.S., 2017. Marine species distribution shifts on the U.S. Northeast Continental Shelf under continued ocean warming. Progress in Oceanography 153, 24–36. https://doi.org/10.1016/j.pocean.2017.04.001

Lellouche, J.-M., Greiner, E., Bourdallé-Badie, R., Garric, G., Melet, A., Drévillon, M., Bricaud, C., Hamon, M., Le Galloudec, O., Regnier, C., Candela, T., Testut, C.-E., Gasparin, F., Ruggiero, G., Benkiran, M., Drillet, Y., Pierre-Yves, L.T., 2021. The Copernicus Global 1/12° Oceanic and Sea Ice GLORYS12 Reanalysis. Front. Earth Sci. 9, 698876. https://doi.org/10.3389/feart.2021.698876

Lellouche, J.-M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C.-E., Bourdalle-Badie, R., Gasparin, F., Hernandez, O., Levier, B., Drillet, Y., Remy, E., Le Traon, P.-Y., 2018. Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1∕12° high-resolution system. Ocean Sci. 14, 1093–1126. https://doi.org/10.5194/os-14-1093-2018

Lentz, S.J., 2017. Seasonal warming of the Middle Atlantic Bight Cold Pool. J. Geophys. Res. Oceans 122, 941–954. https://doi.org/10.1002/2016JC012201

Levitus, S., Boyer, T.P., García, H.E., Locarnini, R.A., Zweng, M.M., Mishonov, A.V., Reagan, J.R., Antonov, J.I., Baranova, O.K., Biddle, M., Hamilton, M., Johnson, D.R., Paver, C.R., Seidov, D., 2014. World Ocean Atlas 2013. https://doi.org/10.7289/V5F769GT

Locarnini, R.A., Mishonov, A.V., Baranova, T.P., Boyer, M.M., Zweng, M.M., Garcia, J.R., Reagan, J.R., Seidov, D., Weathers, C.R., Paver, C.R., Smolyar, I.V., 2019. Volume 1: Temperature, in: Mishonov, A. (Ed.), World Ocean Atlas 2018, NOAA Atlas NESDIS 81. p. 52.

Loder, J.W., Shore, J.A., Hannah, C.G., Petrie, B.D., 2001. Decadal-scale hydrographic and circulation variability in the Scotia–Maine regionSUM. Deep Sea Research Part II: Topical Studies in Oceanography 48, 3–35. https://doi.org/10.1016/S0967-0645(00)00080-1

López, A.G., Wilkin, J.L., Levin, J.C., 2020. Doppio – a ROMS (v3.6)-based circulation model for the Mid-Atlantic Bight and Gulf of Maine: configuration and comparison to integrated coastal observing network observations. Geosci. Model Dev. 13, 3709–3729. https://doi.org/10.5194/gmd-13-3709-2020

Madec, G., 2016. NEMO ocean engine, in: Note Du Pole de Modelisation. p. 396.

Mazur, M.D., Friedland, K.D., McManus, M.C., Goode, A.G., 2020. Dynamic changes in American lobster suitable habitat distribution on the Northeast U.S. Shelf linked to oceanographic conditions. Fish Oceanogr 29, 349–365. https://doi.org/10.1111/fog.12476

Mercator Ocean International, 2016. Global Ocean 1/12° Physics Analysis and Forecast updated Daily. https://doi.org/10.48670/MOI-00016

Miller, A.S., Shepherd, G.R., Fratantoni, P.S., 2016. Correction: Offshore Habitat Preference of Overwintering Juvenile and Adult Black Sea Bass, Centropristis striata, and the

Relationship to Year-Class Success. PLoS ONE 11, e0156355. https://doi.org/10.1371/journal.pone.0156355

Miller, T.J., Hare, J.A., Alade, L.A., 2016. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to southern New England yellowtail flounder. Can. J. Fish. Aquat. Sci. 73, 1261–1270. https://doi.org/10.1139/cjfas-2015-0339

Miller, T.J., O'Brien, L., Fratantoni, P.S., 2018. Temporal and environmental variation in growth and maturity and effects on management reference points of Georges Bank Atlantic cod. Can. J. Fish. Aquat. Sci. 75, 2159–2171. https://doi.org/10.1139/cjfas-2017-0124

Mountain, D.G., 2012. Labrador slope water entering the Gulf of Maine—response to the North Atlantic Oscillation. Continental Shelf Research 47, 150–155. https://doi.org/10.1016/j.csr.2012.07.008

Mountain, D.G., Manning, J.P., 1994. Seasonal and interannual variability in the properties of the surface waters of the Gulf of Maine. Continental Shelf Research 14, 1555–1581. https://doi.org/10.1016/0278-4343(94)90090-6

Northeast Fisheries Science Center (U.S.), 2022a. State of the Ecosystem 2022: Mid-Atlantic. https://doi.org/10.25923/5S5Y-0H81

Northeast Fisheries Science Center (U.S.), 2022b. State of the Ecosystem 2022: New England. https://doi.org/10.25923/YPV2-MW79

O'Leary, C.A., Miller, T.J., Thorson, J.T., Nye, J.A., 2019. Understanding historical summer flounder (*Paralichthys dentatus*) abundance patterns through the incorporation of oceanography-dependent vital rates in Bayesian hierarchical models. Can. J. Fish. Aquat. Sci. 76, 1275–1294. https://doi.org/10.1139/cjfas-2018-0092

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

R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

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

Richaud, B., Kwon, Y.-O., Joyce, T.M., Fratantoni, P.S., Lentz, S.J., 2016. Surface and bottom temperature and salinity climatology along the continental shelf off the Canadian and U.S. East Coasts. Continental Shelf Research 124, 165–181. https://doi.org/10.1016/j.csr.2016.06.005

Rodionov, S.N., 2006. Use of prewhitening in climate regime shift detection. Geophys. Res. Lett. 33, L12707. https://doi.org/10.1029/2006GL025904

Rodionov, S.N., 2004. A sequential algorithm for testing climate regime shifts: ALGORITHM FOR TESTING REGIME SHIFTS. Geophys. Res. Lett. 31, n/a-n/a. https://doi.org/10.1029/2004GL019448

Saba, V.S., Griffies, S.M., Anderson, W.G., Winton, M., Alexander, M.A., Delworth, T.L., Hare, J.A., Harrison, M.J., Rosati, A., Vecchi, G.A., Zhang, R., 2016. Enhanced warming of the Northwest Atlantic Ocean under climate change. J. Geophys. Res. Oceans 121, 118–132. https://doi.org/10.1002/2015JC011346

Seddon, A.W.R., Froyd, C.A., Witkowski, A., Willis, K.J., 2014. A quantitative framework for analysis of regime shifts in a Galápagos coastal lagoon. Ecology 95, 3046–3055. https://doi.org/10.1890/13-1974.1

Seidov, D., Baranova, O.K., Boyer, T., Cross, S.L., Mishonov, A.V., Parsons, A.R., 2016a. Northwest Atlantic regional ocean climatology. 3.2 MB. https://doi.org/10.7289/V5/ATLAS-NESDIS-80

Seidov, D., Baranova, O.K., Johnson, D.R., Boyer, T.P., Mishonov, A.V., Parsons, A.R., 2016b. Northwest Atlantic Regional Climatology (NCEI Accession 0155889). https://doi.org/10.7289/V5RF5S2Q

Seidov, D., Mishonov, A., Parsons, R., 2021. Recent warming and decadal variability of Gulf of Maine and Slope Water. Limnol Oceanogr 66, 3472–3488. https://doi.org/10.1002/lno.11892

Shchepetkin, A.F., McWilliams, J.C., 2005. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Modelling 9, 347–404. https://doi.org/10.1016/j.ocemod.2004.08.002

Stirnimann, L., Conversi, A., Marini, S., 2019. Detection of regime shifts in the environment: testing “STARS” using synthetic and observed time series. ICES Journal of Marine Science 76, 2286–2296. https://doi.org/10.1093/icesjms/fsz148

Tanaka, K.R., Cao, J., Shank, B.V., Truesdell, S.B., Mazur, M.D., Xu, L., Chen, Y., 2019. A model-based approach to incorporate environmental variability into assessment of a commercial fishery: a case study with the American lobster fishery in the Gulf of Maine and Georges Bank. ICES Journal of Marine Science 76, 884–896. https://doi.org/10.1093/icesjms/fsz024

Tanaka, K.R., Torre, M.P., Saba, V.S., Stock, C.A., Chen, Y., 2020. An ensemble high-resolution projection of changes in the future habitat of American lobster and sea scallop in the Northeast US continental shelf. Divers Distrib 26, 987–1001. https://doi.org/10.1111/ddi.13069

Younes, A.F., Cerrato, R.M., Nye, J.A., 2020. Overwintering survivorship and growth of young-of-the-year black sea bass Centropristis striata. PLoS ONE 15, e0236705. https://doi.org/10.1371/journal.pone.0236705

Zang, Z., Ji, R., Hart, D.R., Chen, C., Zhao, L., Davis, C.S., 2022. Modeling Atlantic sea scallop (*Placopecten magellanicus*) scope for growth on the Northeast U.S. Shelf. Fisheries Oceanography 31, 271–290. https://doi.org/10.1111/fog.12577