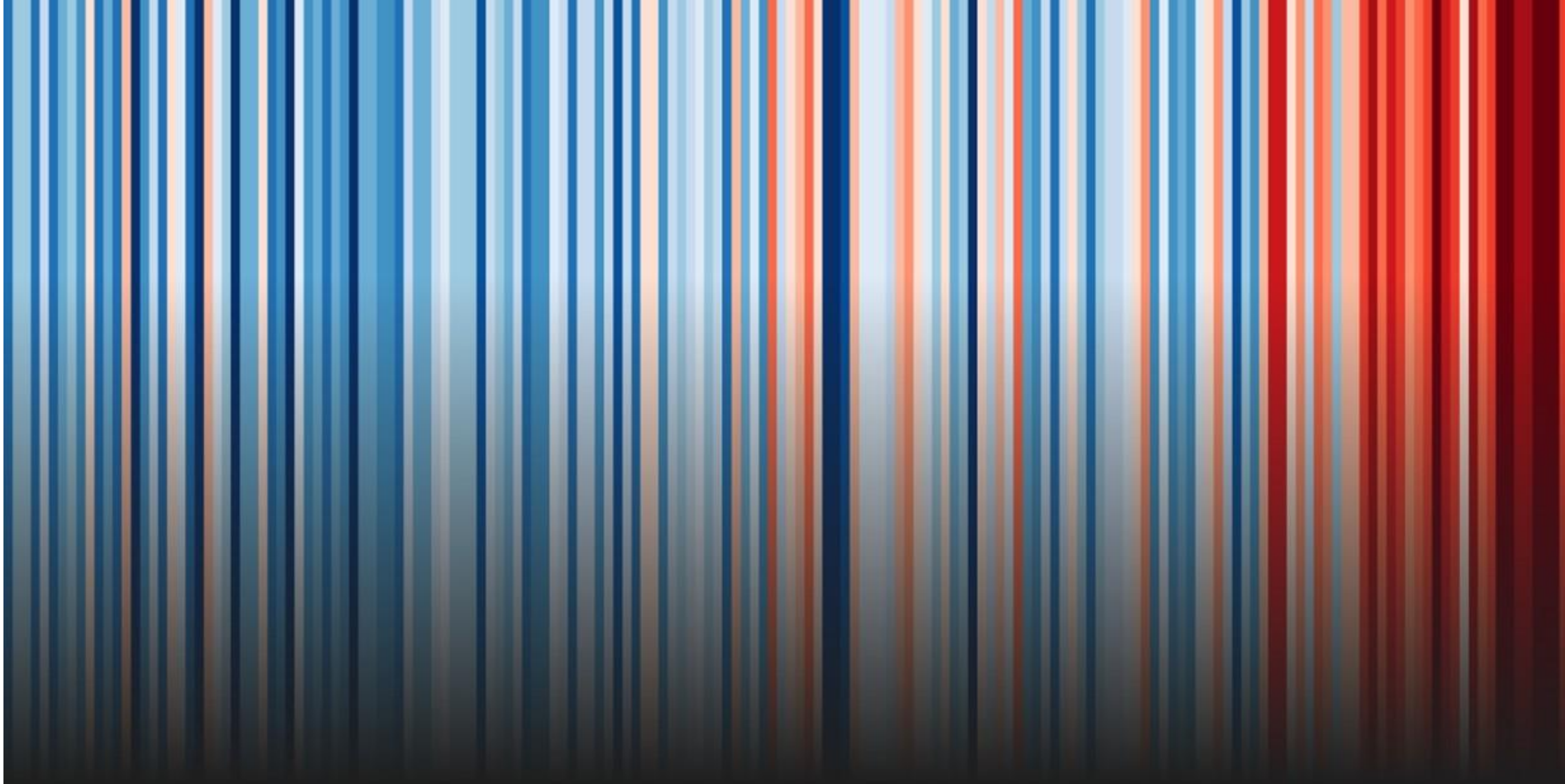


SE4CLIM

Milestone 10 – Reference hindcast runs

Validation Report

Dissemination Level: Public
31/01/26

Funded by
the European Union

DOCUMENT INFORMATION

Work Package Title	WP3 – Climate projections
Milestone	MS10 – Reference hindcasts produced
Version	V1.0
Short Description	In SEACLIM two reference hindcasts, one for each focus region, has been produced for the period 1993-2024. This report contains the description of the runs and the validation of the model output.
Lead Beneficiary	NERSC
Lead Authors	Shuang Gao (NERSC), Lucille Barré (MOi), Annette Samuelsen (NERSC), Karen Guihou (MOi), Meenakshi Sreejith (MOi)
Contributors	Jiping Xie (NERSC), Achref Othmani (NERSC), Stephane Law Chune (MOi)
Due Date	31/01/2026
How to cite this report	

ACKNOWLEDGEMENT
This project has received funding from the European Union’s Horizon programme under grant agreement 101180125.

Table of Content

1. Introduction.......... 8

2. Hindcast configurations.......... 8

  2.1. Model setup and configuration - TOPAZ.......... 8

    2.1.1. The ocean model HYCOM.......... 8

    2.1.2. The sea ice model CICE.......... 8

    2.1.3. The biogeochemical model ECOSMO II(CHL).......... 9

    2.1.4. The coupled system.......... 9

    2.1.5. Model configuration.......... 9

  2.2. Model setup and configuration - NEATL12C.......... 11

    2.2.1. Ocean physics – NEMOv5.0.......... 11

    2.2.2. Ocean biogeochemistry – PISCESv2.......... 12

    2.2.3. Waves – MF-WAM.......... 13

    2.2.4. The coupled system – NEATL12C.......... 13

    2.2.5. Model configuration.......... 14

3. Model validation.......... 15

  3.1. The Northeast Atlantic.......... 15

    3.1.1. Physical validation.......... 16

      3.1.1.1. Temperature.......... 18

      3.1.1.2. Salinity.......... 21

      3.1.1.3. Surface currents.......... 22

      3.1.1.4. Sea surface height.......... 23

      3.1.1.5. Mixed layer depth.......... 24

    3.1.2. Biogeochemical validation.......... 26

      3.1.2.1. Total chlorophyll concentration.......... 29

      3.1.2.2. Euphotic zone depth.......... 32

      3.1.2.3. Nutrients and dissolved oxygen concentration.......... 32

      3.1.2.4. Organisms.......... 36

      3.1.2.5. Primary production.......... 42

      3.1.2.6. Carbonate system.......... 44

    3.1.3. Waves validation.......... 51

  3.2. Arctic.......... 59

    3.2.1. Physical validation.......... 60

    3.2.2. Biogeochemical validation.......... 64

3.2.3. Sea ice validation 73

4. Data access 76

5. References 77

List of Tables

Table 1. List of outputs (only validated ones) 10
Table 2. List of outputs (only validated one, a detailed version is available on EDITO platform), with frequencies and associated files. 14
Table 3. Validated physical variables, with associated figures, metrics and reference datasets. (*: Figure available in appendix). 16
Table 4. Validated biogeochemical variables, with associated figures, metrics and reference datasets. (*: Figures available in appendix). 26
Table 5. Validated physical variables, with associated figures, metrics and reference datasets. 51
Table 6. Estimated net volume transports in Sv of specific water masses through key sections, average fro the year 1993 to 2024 compared to numbers from literature. 63

List of Figures

Figure 1. Model domain and resolution of TOPAZ2. 10
Figure 2. Model domain and bathymetry (m) of NEATL12C. 11
Figure 3. Model domain and bathymetry (m) of MFWAM. 13
Figure 4. Surface temperature (a) from the hindcast simulation and (b) satellite observations (ARMOR3D), for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period. 18
Figure 5. Mooring sites of surface temperature. Black lines represent the domain's limits. 19
Figure 6. Hourly surface temperature at four mooring sites: (a) Gijon, (b) Arcachon, (c) Le Havre and (d) Wadden (see Figure 5 for moorings site locations). The chosen period depends on the availability of data. 19
Figure 7. Hovmoller diagrams of mean temperature (°C, domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) is for the hindcast simulation and (b, e) for the reanalysis (IBIRYSv3), for the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (°C, model – reanalysis) for the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis temperature (°C). 20
Figure 8. Surface salinity (a) from the hindcast simulation and (b) satellite observations (ARMOR3D), for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period. 21
Figure 9. Hovmoller diagrams of mean salinity (domain average). Panels a, b, and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) is for the hindcast simulation and (b, e) reanalysis (IBIRYSv3), for the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (model – reanalysis) for the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis salinity. 22
Figure 10. Surface currents (eastward, i-axis) (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) on the same period. 23
Figure 11. Surface currents (northward, j-axis) (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), on the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) on the same period. 23

Figure 12. Sea surface height anomalies referenced to the mean sea level over the whole period for (a) the Hindcast simulation and (b) IBI36RYS reanalysis. Anomalies are compared by calculating a linear regression (with formulation in the right bottom corner).................. 24
Figure 13. SLA trend calculated for the period 1993 – 2024, for (a) the Hindcast simulation and (b) the Altimetry product. 24
Figure 14. Mixed layer thickness (a) from the hindcast simulation and (b) satellite observations (ARMOR3D), for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period. 25
Figure 15. Mixed layer thickness (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period.. 26
Figure 16. Surface total chlorophyll concentration (a) from the hindcast simulation and (b) satellite observations (CMEMS-ATL), for the period 1998 – 2024 (averaged), in the Atlantic Ocean (Mediterranean excluded). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period. 30
Figure 17. Surface total chlorophyll concentration (a) from the hindcast simulation and (b) satellite observations (CMEMS-MED), for the period 1998 – 2024 (averaged) in the Mediterranean. Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period... 31
Figure 18. Hovmoller diagrams of mean total chlorophyll concentration (mg.m-3, domain average) (a) for the hindcast simulation and (b) reanalysis (IBIRYSv3), on the period 1993 – 2024. (c) Both are compared by calculating the mean bias (mg.m^{-3}, model – reanalysis) on the same period. (d) Mean vertical profiles (domain and period average) of the hindcast and reanalysis total chlorophyll concentration (mg.m^{-3}). Figure focuses on the 0-200 layer........ 31
Figure 19. Euphotic zone depth (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period.. 32
Figure 20. Surface NO_3 concentration (a) from the hindcast simulation and (b) observations (WOA2023), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – observations) for the same period. 33
Figure 21. Surface PO_4 concentration (a) from the hindcast simulation and (b) observations (WOA2023), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – observations) for the same period. 33
Figure 22. Surface Si concentration (a) from the hindcast simulation and (b) observations (WOA2023), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – observations) for the same period. 34
Figure 23. Hovmoller diagrams of mean NO_3 concentration (mmol.m^{-3}, domain average). Panels a, b, and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis NO3 concentration (mmol.m^{-3}). 35
Figure 24. Hovmoller diagrams of mean PO_4 concentration (mmol.m^{-3}, domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias

(mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis PO_4 concentration (mmol.m^{-3}). 35
Figure 25. Hovmoller diagrams of mean Si concentration (mmol.m^{-3}, domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis Si concentration (mmol.m^{-3}). 36
Figure 26. Surface O_2 concentration (a) from the hindcast simulation and (b) observations (WOA2023), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – observations) for the same period. 37
Figure 27. Hovmoller diagrams of mean O_2 concentration (mmol.m^{-3}, domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis O_2 concentration (mmol.m^{-3}). 37
Figure 28. Profiles locations. Black lines represent the area limits. 38
Figure 29. Mean vertical profiles (average of the period 1993 – 2024) at 4 points (3 points in the Atlantic ocean and 1 point in the Med sea, Figure 28 for locations), of (a, e, i, m) hindcast and WOA2023 climatology NO_3 concentration, (b, f, j, n) hindcast and WOA2023 climatology PO_4 concentration, (c, g, k, o) hindcast and WOA2023 climatology Si concentration, and (d, h, l, p) hindcast and WOA2023 climatology O_2 concentration. 39
Figure 30. Surface total phytoplankton concentration in carbon (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period. 40
Figure 31. Surface total zooplankton concentration in carbon (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period. 40
Figure 32. Hovmoller diagrams of mean total phytoplankton (mmol.m^{-3}, domain average) (a) for the hindcast simulation and (b) reanalysis (IBIRYSv3), on the period 1993 – 2024. (c) Both are compared by calculating the mean bias (mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis total phytoplankton concentration (mmol.m^{-3}). Figure focuses on the 0-200m layer. 41
Figure 33. Hovmoller diagrams of mean total zooplankton (mmol.m^{-3}, domain average) (a) for the hindcast simulation and (b) reanalysis (IBIRYSv3), on the period 1993 – 2024. (c) Both are compared by calculating the mean bias (mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis total zooplankton (mmol.m^{-3}). Figure focuses on the 0-200m layer. 41
Figure 34. Depth integrated primary production (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period. 42
Figure 35. Depth integrated primary production (a) from the hindcast simulation and (b) satellite observations (CMEMS-ATL), in the Atlantic Ocean (Mediterranean excluded), for the period 1998 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period. 43

Figure 36. Depth integrated primary production (a) from the hindcast simulation and (b) satellite observations (CMEMS-MED), in the Mediterranean, for the period 1998 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period. 44
Figure 37. Surface pH (Atlantic area) (a) from the hindcast simulation and (b) MOB-TAC product, for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period. 45
Figure 38. Hovmoller diagrams of mean pH (domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis pH. 46
Figure 39. Surface marine pCO_2 (µatm, Atlantic area) (a) from the hindcast simulation and (b) MOB-TAC product, for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (µatm, model – observations) and (d) the mean RMSE for the same period. 47
Figure 40. Air-sea CO_2 fluxes ($mol.m^{-2}.yr^{-1}$, Atlantic area) (a) from the hindcast simulation and (b) MOB-TAC product, for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias ($mol.m^{-2}.yr^{-1}$, model – observations) and (d) the mean RMSE for the same period. 48
Figure 41. Climatology of Significant wave height (VHM0) for the period 1993-2024 from (a) hindcast simulations, (b) reanalysis dataset, IBIRYS and (c) the mean bias (model – observation). 53
Figure 42. Domain averaged hourly Significant wave height (VHM0) for the period 1993-2024 overlayed with monthly mean values, for IBIRYS and SEACLIM-IBI. The statistics represent the comparison of hourly data. 54
Figure 43. Correlation (a) and RMSE (b) obtained from comparison of model and buoy VHM0 at various locations in the IBI region. 54
Figure 44. Climatology of mean wave period (VTM02) for the period 1993-2024 from (a) hindcast simulations, (b) reanalysis dataset, IBIRYS and (c) the mean bias (model – observation). 55
Figure 45. Domain averaged hourly Mean Wave Period (VTM02) for the period 1993-2024 overlayed with monthly mean values, of IBIRYS and SEACLIM-IBI. The statistics represent the comparison of hourly data. 56
Figure 46. Correlation (a) and RMSE (b) obtained from comparison of model and buoy VTM02 at various locations in the IBI region. 56
Figure 47. Climatology of mean wave period (VTPK) for the period 1993-2024 from (a) hindcast simulations, (b) reanalysis dataset, IBIRYS and (c) the mean bias (model – observation). 57
Figure 48. Domain averaged hourly Peak Period (VTPK) for the period 1993-2024 overlayed with monthly mean values, of IBIRYS and SEACLIM-IBI. The statistics represent the comparison of hourly data. 57
Figure 49. Correlation (a) and RMSE (b) obtained from comparison of model and buoy VTPK at various locations in the IBI region. 58
Figure 50. Correlation (a) and RMSE (b) obtained from comparison of model and buoy VMDR at various locations in the IBI region. 59

Figure 51. Hovmuller diagram of temperature from GLORYS12v1 and TOPAZ2, the temperature has been averages over the entire domain at each model depth. 60
Figure 52. Hovmuller diagram of salinity from GLORYS12v1 and TOPAZ2, the temperature has been averages over the entire domain at each model depth. 60
Figure 53. Average surface temperature (upper) and salinity (lower) from 1993 to 2024 from GLORYS12v1 (left), TOPAZ2 (middle) and the bias (right). 61
Figure 54. Average surface temperature (upper) and salinity (lower) from 1993 to 2024 from ARMOR3D (left), TOPAZ2 (middle) and the bias (right). 62
Figure 55. Average sea-level anomaly (upper) and mixed-layer depth (lower) from 1993 to 2024 from ARMOR3D (left), TOPAZ2 (middle) and the bias (right). 62
Figure 56. Average 1993-2024 MLD from TOPAZ2 (left) and from the climatology (middle) and the bias (right) in winter (January to March). 63
Figure 57. Average 1993-2024 MLD from TOPAZ2 (left) and from the climatology (middle) and the bias (right) in summer (July to September). 63
Figure 58. Average surface eastward velocity (upper) and westward velocity (lower) from 1993 to 2024 from GLORYS12v1 (left), TOPAZ2 (middle) and the bias (right). 63
Figure 59. Surface nitrate concentration from the hindcast simulation (left), from observations (WOA2023) (middle) for the period 1993 – 2024 (averaged) and the mean bias (model – observations) for the same period (right). 64
Figure 60. Surface phosphate concentration from the hindcast simulation (left), from observations (WOA2023) (middle) for the period 1993 – 2024 (averaged) and the mean bias (model – observations) for the same period (right). 64
Figure 61. Surface silicate concentration from the hindcast simulation (left), from observations (WOA2023) (middle) for the period 1993 – 2024 (averaged) and the mean bias (model – observations) for the same period (right). 65
Figure 62. Surface oxygen concentration from the hindcast simulation (left), from observations (WOA2023) (middle) for the period 1993 – 2024 (averaged) and the mean bias (model – observations) for the same period (right). 66
Figure 63. Regions used in the validation of biogeochemical variables in the hindcast simulation with TOPAZ2 against in-situ data. 66
Figure 64. Seasonal cycle of nitrate (mmol m-3) from simulated monthly means (blue) against in-situ observations (black) at four depth ranges (0–10, 10–30, 30–50, 50–100 m) across the Lofoten Basin, Norwegian Basin, Barents Sea, and Greenland–Iceland Seas. Shaded areas show standard deviations. Bars indicate the number of in-situ samples. Solid blue lines represent area-averaged model values, while dashed blue lines show averaged model values sampled at the in-situ locations. 67
Figure 65. Seasonal cycle of phosphate (mmol m-3) from simulated monthly means (blue) against in-situ observations (black) at four depth ranges (0–10, 10–30, 30–50, 50–100 m) across the Lofoten Basin, Norwegian Basin, Barents Sea, and Greenland–Iceland Seas. Shaded areas show standard deviations. Bars indicate the number of in-situ samples. Solid blue lines represent area-averaged model values, while dashed blue lines show averaged model values sampled at the in-situ locations. 68
Figure 66. Seasonal cycle of silicate (mmol m-3) from simulated monthly means (blue) against in-situ observations (black) at four depth ranges (0–10, 10–30, 30–50, 50–100 m) across the Lofoten Basin, Norwegian Basin, Barents Sea, and Greenland–Iceland Seas. Shaded areas show standard deviations. Bars indicate the number of in-situ samples. Solid

blue lines represent area-averaged model values, while dashed blue lines show averaged model values sampled at the in-situ locations. 69
Figure 67. Seasonal cycle of oxygen (mmol m-3) from simulated monthly means (blue) against in-situ observations (black) at four depth ranges (0–10, 10–30, 30–50, 50–100 m) across the Lofoten Basin, Norwegian Basin, Barents Sea, and Greenland–Iceland Seas. Shaded areas show standard deviations. Bars indicate the number of in-situ samples. Solid blue lines represent area-averaged model values, while dashed blue lines show averaged model values sampled at the in-situ locations. 71
Figure 68. Time series (1998–2021) of monthly mean surface chlorophyll concentrations (mg Chl m-3) from the model compared with satellite observations in five subregions: Lofoten Basin, Norwegian Basin, Barents Sea, Irminger–Iceland Basin, and Greenland–Iceland Seas. Only co-located data points are included in the comparison. 72
Figure 69. Time series (2002–2023) of monthly mean primary production from the model compared with satellite observations in five subregions: Lofoten Basin, Norwegian Basin, Barents Sea, Irminger–Iceland Basin, and Greenland–Iceland Seas. Only co-located data points are included in the comparison. 73
Figure 70. Sea ice extent from TOPAZ (TP2) and observations (OSI SAF Monthly). 74
Figure 71. Average sea ice thickness in October and from TOPAZ2 and observations of CS2SMOS. The figure show averages from 2011 to 2024. 74

1. Introduction

The SEACLIM milestone, MS10, delivers two reference hindcasts, one covering the North East Atlantic (NEATL12C) and the second covering the Nordic Seas and the Arctic (TOPAZ). These simulations serve two main objectives.

The first is to provide a reference regional free simulation forced by global reanalysis datasets: ERA5 for atmospheric variables Click or tap here to enter text., GLORYS12v1 (Lellouche et al., 2021), FREEBIORYS from the Copernicus Marine Service, used respectively for physical and biogeochemical variables (as lateral boundary and initial conditions).These simulations represent the best possible model performance without data assimilation and will serve as a reference for evaluating downscaled regional ocean decadal predictions and climate projections. For the wave component in the Atlantic, as no spectral boundary data are available for the regional configuration, a free simulation at 1° resolution was performed for this purpose, incorporating surface currents from GLORYS12v1.

The second objective is to provide initial conditions for retrospective decadal predictions (also referred to as hindcasts in WP2).

This report contains a description of the model configurations (Section 2) and the validation of the key model output parameters (Section 3). The output of the reference runs has been uploaded as NetCDF-files on the Edito platform. Section 4 describes how to access the model results.

2. Hindcast configurations

2.1. Model setup and configuration - TOPAZ

2.1.1. The ocean model HYCOM

The Hybrid Coordinate Ocean Model (Bleck, 2002) is a general circulation ocean model that combines isopycnal layers, z-levels, and terrain-following coordinates. This hybrid approach leverages the strong conservation properties of isopycnal coordinates in the ocean interior while high vertical resolution in the well-mixed surface layer and coastal regions. HYCOM has been widely used for both global and regional applications (e.g. Braby et al., 2020; Chassignet et al., 2009; Xie et al., 2020).

The version used in this study builds on previous modelling systems for the North Atlantic and Arctic Oceans (Bertino & Lisæter, 2008; Sakov et al., 2012; Xie et al., 2019). The implementation is based on HYCOM version 2.3 and configured for Arctic applications. The full model code is available in the NERSC Git repository: https://github.com/nansencenter/NERSC-HYCOM-CICE/tree/develop.

2.1.2. The sea ice model CICE

The sea ice component, CICE version 5.1 (Hunke et al., 2015), is coupled to HYCOM using the Earth System Modeling Framework (ESMF). CICE uses the elastic-viscous-plastic (EVP) rheology and represents sea ice using five thickness categories. A “mushy-layer” formulation is applied to simulate variable sea ice salinity and evolving freezing temperatures (Turner & Hunke, 2015).

The shortwave radiation and albedo through snow and sea ice are computed using the delta-Eddington approach (Briegleb & Light, 2007), which implicitly accounts for melt ponds by adjusting albedo based on surface conditions. The model represents a smooth decrease in albedo with decreasing ice thickness. Although CICE includes an optional sea ice biogeochemistry module (Duarte et al., 2015), this feature is not activated in our setup. However, the light transmission module within CICE is used to allow phytoplankton growth beneath sea ice under sufficient light conditions.

2.1.3. The biogeochemical model ECOSMO II(CHL)

We used an updated version of ECOSMO II(CHL) (Yumruktepe et al., 2022), an intermediate-complexity lower trophic level biogeochemical model, to simulate the processes in the water column and the sediment layer in the higher latitudes of the North Atlantic and the Arctic Ocean. ECOSMO II(CHL) resolves nitrate, ammonium, phosphate and silicate as inorganic nutrients, diatoms and flagellates as primary producers, micro- and meso-size classes as herbivorous and omnivorous zooplankton respectively. The model uses the molar Redfield ratio (106 : 16 : 16 : 1) between C : N : Si : P components and simulates oxygen as an explicit state variable. Both dissolved (DOM) and particulate (detritus) organic matter are included in the model. Each phytoplankton has its own chlorophyll *a* (chl-a) component as a state variable that actively resolves the carbon:chl-a ratio which is modelled using the light-dependent formulation by Geider et al. (1998).

2.1.4. The coupled system

The coupled HYCOM-CICE-ECOSMO modeling framework forms the basis for the development of the TOPAZ data assimilation system. The coupling between the HYCOM and CICE uses the Earth System Modeling Framework coupler Earth System Modeling Framework (ESMF, version 8.0.1), while ECOSMO is couple using the Framework for Aquation Biogeochemical models (FABM: Bruggeman & Karsten Bolding, 2011). This coupled ocean–sea ice–biogeochemistry system provides forecasts and reanalyses of physical and biogeochemical ocean conditions and is specifically configured for the North Atlantic and Arctic regions. For this project, we use a relatively coarse-resolution setup (average grid spacing of ~17 km) to produce multi-year hindcast of key physical and biogeochemical variables, including salinity, temperature, nutrients, oxygen, primary production, and zooplankton biomass. Model performance is evaluated against in situ observations, ocean reanalysis (GLORYSv12v1), satellite data, and climatological datasets.

2.1.5. Model configuration

The present configuration focuses on the Nordic Seas and Arctic Ocean, with spatial resolution increasing towards the Arctic. The model is configured on a relatively coarse resolution ranging between 13 and 20 km (Figure 1). Atmospheric forcing data are obtained from the 6-hourly ECMWF ERA5 reanalysis (Hersbach et al., 2020), including 10 m wind speed, 2 m air temperature, 2 m dew-point temperature, atmospheric pressure, total precipitation, and long- and shortwave radiation.

The model has been spun up with ocean physics and sea-ice for ten years starting 1977 with lateral boundary nudging of World Ocean Atlas 2018 (WOA2018) climatology (Locarnini et al., 2019; Zweng et al., 2019). The biogeochemistry tracers, i.e., nutrients and carbon, are initialised with World Ocean Atlas 2013 monthly climatology profiles on 1987-01-01. The hindcast simulation covers the period from November 1992 to 2024. Both ocean physics and biogeochemistry were initialized from a restart file in September 1992. During 1992, temperature, salinity and biogeochemical tracers were relaxed towards the corresponding monthly climatology from WOA2018 and WOA2013 at the lateral boundaries. This relaxation was implemented using zones with a width of twenty grid cells and a relaxation timescale of twenty days.

From 1993 onwards, nesting conditions were applied at the lateral boundaries using data from the GLORYS12 V1 reanalysis (daily outputs at 1/12° horizontal resolution and 50 vertical levels; doi:10.48670/moi-00021) and and the global biogeochemical hindcast (https://doi.org/10.48670/moi-00019) for the biogecochemical varibables. The nesting process used relaxation zones with a width of twenty grid cells and a relaxation timescale of twenty days. This approach was applied to sea surface height (SSH), temperature, salinity, horizontal ocean currents, Montgomery potential, nutrients (nitrate, phosphate and silicate) and oxygen.

Model results were saved as daily means. Restart files were saved every 10 days, with an additional restart file written on November 1st of each year throughout the reference run.

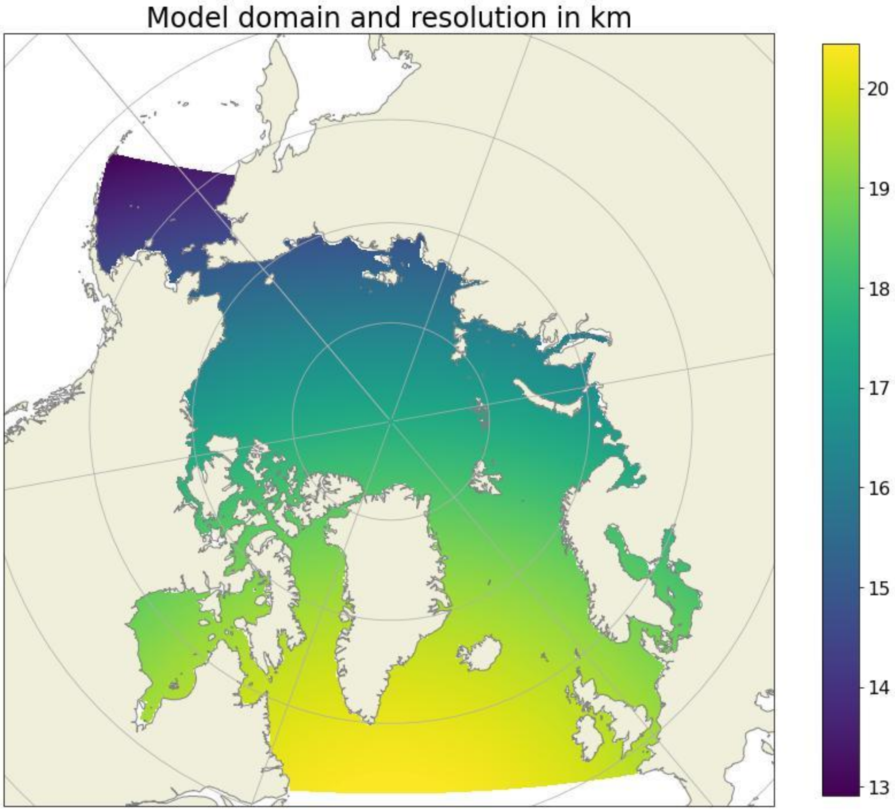


Figure 1. Model domain and resolution of TOPAZ2.

Table 1. List of outputs (only validated ones)

Variable	Unit	Variable name	Frequency
Physical variables 3D			
Ocean Temperature	degC	thetao	monthly
Ocean Salinity	-	so	monthly
Ocean water x velocity	m s-1	uo	monthly
Ocean water y velocity	m s-1	vo	monthly
Physical variables 2D			
Sea surface temperature	degC	sst	daily
Sea surface salinity	-	sss	daily

Sea surface height	m	zos	daily
Ocean mixed layer thickness defined by density	m	mlotst	daily
Biogeochemical variables 3D			
Nitrate	mmol m-3	no3	monthly
Phosphate	mmol m-3	po4	monthly
Silicate	mmol m-3	si	monthly
Oxygen	mmol m-3	o2	monthly
Biogeochemical variables 2D			
Integrated primary production	mgC m-2 -d	npp	daily
Sea-ice variables 2D			
Sea ice thickness	m	hi	daily
Sea ice concentration	-	aice	daily

2.2. Model setup and configuration - NEATL12C

NEATL12C is a regional ocean model for the north-eastern Atlantic and European northwest shelf, including the North Sea and the western Mediterranean Sea (Figure 2). It derives from the IBI reanalysis configuration developed for the Copernicus Marine Service (doi: 10.48670/moi-00029) and is dedicated to climate studies (Chaigneau et al. 2022).

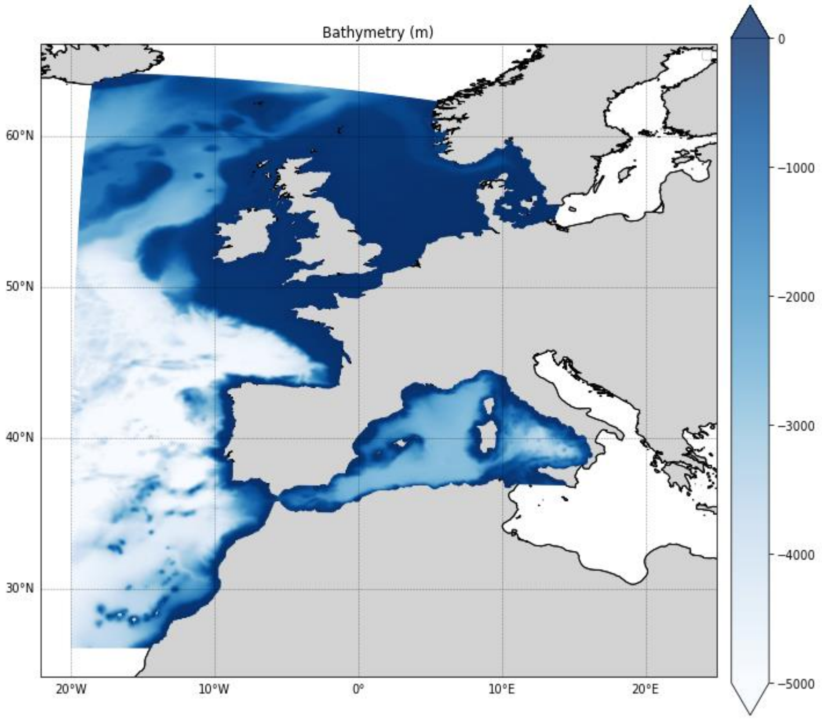


Figure 2. Model domain and bathymetry (m) of NEATL12C.

2.2.1. Ocean physics – NEMOv5.0

The NEATL12C configuration is based on the primitive equation NEMO ocean model (Nucleus for European Modelling on the Ocean, Madec et al., 2024), version 5.0 to represent ocean physics. In the horizontal direction, NEMOv5.0 uses a curvilinear orthogonal grid, with a 1/12° resolution (6-7 km approximately), covering the region from 25°N to 65°N and 21°W to 16°E (Figure 2). The vertical direction is split in 75 levels with a variable thickness (about a meter near the surface and up to 200m near the bottom). To represent the diversity of situations encountered in the ocean on the vertical, NEMOv5.0 uses a space and time dependent vertical coordinate (z*-coordinate) thus considering the variation of the sea surface height by implementing a variable volume (Levier et al., 2007). Additionally, to better fit the bathymetry, NEMOv5.0 uses a partial step method (Bernard et al., 2006). For further details on the Ocean physics representation in the configuration, the reader is referred to Chaigneau et al. (2022).

Initially, NEATL12C was based on NEMO v4.2 but for SEACLIM project, the ocean physics component was updated to NEMO v5.0 (Madec et al., 2024). NEMO v5.0 features two time-stepping schemes: Modified Leap Frog (MLF: Leclair & Madec, 2009; Mesinger & Arakawa, 1976) and Runge Kutta 3rd order (RK3: Wicker & Skamarock, 2002). The RK3 scheme, beside its higher formal accuracy than MLF (3rd order vs less than 2nd order for linear terms), has the advantage of requiring less memory storage and being more computationally efficient. Then, by using XIOS3, activating RK3 and doubling the time step, NEMO v5.0 can be up to 2.5 times faster than NEMO v4. More details about differences between NEMO v5.0 and v4.2 can be found in NEMO user guide: https://sites.nemo-ocean.io/user-guide/changes.html.

2.2.2. Ocean biogeochemistry – PISCESv2

To represent the ocean biogeochemistry, NEATL12C configuration relied on PISCES (Pelagic Interaction System for Carbon and Ecosystem Studies) biogeochemical model, version 2 (Aumont et al., 2015), which is part of the NEMO modelling platform. PISCES-v2 is a model of intermediate complexity, which simulates the lower trophic levels of marine ecosystems and the biogeochemical cycles of carbon, oxygen and the main nutrients controlling the phytoplankton growth. The model features 24 prognostic variables which includes four types of plankton based on size (two zooplankton, micro- (small) and meso- (large) zooplankton; and two phytoplankton, nano-phytoplankton (small) and diatoms (large), detritus, main nutrients (P, N, Fe, and Si), oxygen and a representation of carbonate system.

Both phytoplankton are represented (i.e., a prognostic variable is available) in carbon, iron, silicate (only for diatoms) and chlorophyll with variable Fe/C, Si/C and Chl/C ratios. Both zooplankton are only represented in carbon (no silicified zooplankton considered in PISCES). For all species, C/N/P and O2 ratios are assumed constant, set to values defined in Takahashi et al. (1985) and Körtzinger et al. (2001), respectively. Three types of detritus are defined: small particulate organic carbon, large particulate organic carbon and semi-labile dissolved organic carbon, in which only the iron, silicon and carbonate contents of the particles are computed prognostically. Additionally, PISCESv2 includes a representation of the carbonate system using alkalinity and dissolved inorganic carbon as input variables, then computing pH and marine pCO2.
PISCESv2 was used in various studies (e.g. Bopp et al., 2005; Gehlen et al., 2006, 2007; Schneider et al., 2008; Séférian et al., 2013; Steinacher et al., 2010; Tagliabue et al., 2010) and for products delivered to the Copernicus Marine Service: the global ocean analysis and forecast product at 1/4° (doi: 10.48670/moi-00015), the global ocean non-assimilative hindcast product at 1/4° (doi: 10.48670/moi-00019), the IBI non-assimilative hindcast product at 1/12° (doi: 10.48670/moi-00028) and the IBI analysis and forecast product at 1/36° (doi: 10.48670/moi-00026), for example.

2.2.3. Waves – MF-WAM

The wave component of NEATL12C in SEACLIM is based on the model system currently used for the wave reanalysis in the Copernicus Marine Service for these regions (Toledano et al., 2022) and works with a comparable spatial resolution than the physical system (1/10° native) while angular and spectral resolutions are respectively 24 directions for 30 frequencies. The computing code is based on MFWAM, a third-generation wave model derived from ECWAM (IFS-38R2, ECMWF 2012) that includes specific terms for dissipation by wave breaking and swell damping due to air friction at the sea surface. The white-capping dissipation is ST4-like (Ardhuin et al., 2010) and the model has been enhanced with adjustments to dissipation terms and an improved drag coefficient using a Philips spectrum tail for high-frequency waves. The wind input term uses the one proposed by Bidlot et al. (2007), modified for short waves and a negative swell damping term.

Within the SEACLIM framework, the wave model accounts for ocean currents, but also sea level variations, which are of primary importance in the context of climate change (Chaigneau et al., 2022). For the IBI region, this approach additionally enables the representation of tide–wave interactions, which are particularly significant in the Irish and English Channel waters. Coastal wave breaking is also activated following the formulation of Battjes and Janssen (1978).

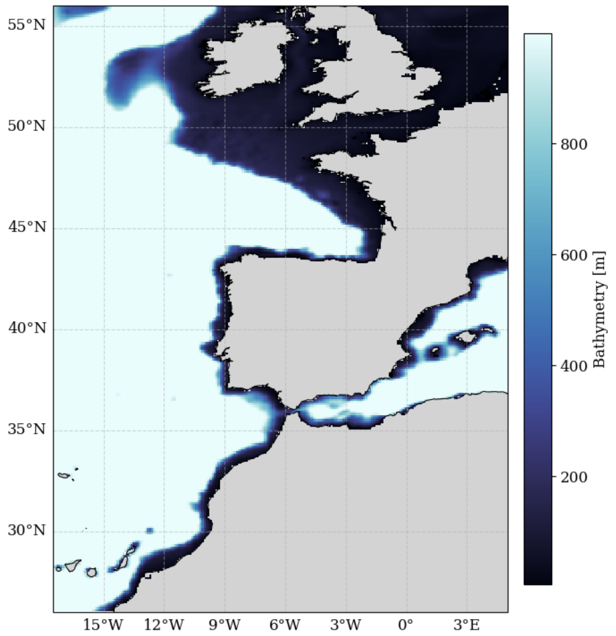


Figure 3. Model domain and bathymetry (m) of MFWAM.

2.2.4. The coupled system – NEATL12C

As a part of the NEMOv5.0 modelling platform, PISCES-v2 is coupled to NEMOv5.0 physics component via the TOP component (tracers component), which manages the source-sink equations associated to PISCES-v2 biogeochemical variables. The coupling is performed online which means that physics and biogeochemistry run in parallel. The coupled configuration horizontal resolution is the same as specified in section 2.2.1: 1/12° (6-7 km approximately).

Waves are not coupled but rather forced by the hourly surface currents and ssh produced by the coupled configuration as forcings. The grid on which they are computed is slightly different than the one used for the coupled configuration. Due to the number of boundaries support by the configuration, the waves model cannot use the extended grid used by the coupled configuration (Figure 2). Additionally, the resolution is 1/10°, which is lower than the one used for the coupled system but remains comparable.

2.2.5. Model configuration

The period covered by the reference hindcast is 1993-2024. The simulation starts by using the ocean reanalyzes GLORYS12v1 and BIORYS4 as initial conditions on January 1st, 1993 (beginning of the reanalysis period). A spin up was performed by looping 5 times over the year 1993 and allowed to reach a satisfying equilibrium state. For the wave component, a three-month spin-up period is applied, which is largely sufficient for wave systems propagating over the global ocean to reach equilibrium.

The hindcast is forced by the same reanalyses as the initial condition at open boundaries: physics variables (temperature, salinity, sea surface height and currents) are provided by GLORYS12v1 for the whole run period; biogeochemical variables are provided by three versions of the biogeochemistry reanalysis: BIORYS (from 1993 to 2019), BIOMER4v2 (from 2019 to 2023) and BIO4GLO12 (from 2023 to 2024), depending on their availability. The wave component uses wave spectra derived from a free global 1° simulation, with boundary forcing provided by ERA5 atmospheric fields and GLORYS12v1 ocean currents.

Results files are provided in netCDF format through XIOS3. Main output variables and associated frequencies are specified in Table 2 (validated variables only, see the EDITO platform for a full dataset: file outputs_hindcast.ods for the description). They were chosen based on deliverable 1.1 (Data management pan, Data summary table, WP3 tab). Considering storage limitations, choices have been made to limit the total size of hindcast results. When more than one frequency was specified, the lower one was chosen for the 3D variables. Moreover, some variables have been grouped in a same results file to avoid writing again dimension variables which are written by default in all files.

Table 2. List of outputs (only validated one, a detailed version is available on EDITO platform), with frequencies and associated files.

Variable	Variable long name	Units	Frequencies	Files suffix
ssh	Sea surface height	m	Hourly 25h average Monthly Yearly	2DT-oce
mlddt02	Mixed layer thickness defined by temperature	m	Hourly Daily Monthly Yearly	
thetao	Temperature	°C	25h average Monthly Yearly	3DT-thetao
so	Salinity	-		3DT-so
uo	Currents (i-axis)	m.s-1		3DU-uo

vo	Currents (j-axis)	m.s-1		3DV-vo
spco2	Surface marine pCO2	µatm	Daily Monthly Yearly	2DT-bgc
Cflx	Air-sea CO2 fluxes	mol.m-2.s-1		
zeu	Euphotic layer depth	m		
npp	Vertically integrated primary production by total phytoplankton	mol.m-2.s-1		
chl	Total chlorophyll concentration	mgChl.m-3	Monthly Yearly	3DT-chl
phyc	Total phytoplankton	mmolC.m-3		3DT-phyc
zooc	Total zooplankton	mmolC.m-3		3DT-zooc
alk	Alkalinity	mmol.m-3		3DT-alk
dic	Dissolved inorganic carbon	mmolC.m-3		3DT-dxc
ph	PH (total scale)	-		3DT-ph
no3	NO3 concentration	mmolN.m-3		3DT-no3
po4	PO4 concentration	mmolP.m-3		3DT-po4
si	Si concentration	mmolSi.m-3		3DT-si
o2	Dissolved oxygen concentration	mmolO.m-3		3DT-o2
VHM0	Spectral Significant Wave Height	m	Hourly	IBI-wav
VTM02	Spectral moments (0,2) wave period (Tm02)	s		
VTPK	Wave period at spectral peak / peak period (Tp)	s		
VMDR	Mean wave direction	degree		

3. Model validation

3.1. The Northeast Atlantic

The hindcast simulation from NEATL12C is evaluated against observations and reanalyses, leveraging on the Copernicus Marine product quality assessment framework. We assessed 2D and 3D variables by using both observations and reanalyses datasets. Used datasets include satellite data (multi-year reprocessed observations from Copernicus Marine/C3S, including based on ESA-CCI processing chains), in situ observations, and IBIRYSv3 (1/12° reanalysis).

More details about these datasets can be found in box 1. In the following, the validation is separated in three subsections: physical validation, biogeochemical validation and waves validation. In each part, a table summarizing the data sets, metrics and type of figure chosen for each validated variable is provided.

3.1.1. Physical validation

In this section, we present the validation of physical variables. The validated variables are specified in Table 3. We also present the used datasets, the metrics and the chosen type of figure for visualization.

Table 3. Validated physical variables, with associated figures, metrics and reference datasets. (: Figure available in appendix).*

Variable	Type of figure	Metrics	Reference dataset
2D variables			
Sea surface temperature	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°) *
		Bias, RMSE	Multi observation global Ocean ARMOR3D L4 reanalysis *[MULTIOBS_GLO_PHY_TSUV_3D_MYNRT_015_012]*
	Time series at moorings sites	None	Global Ocean- In-Situ Near-Real-Time Observations [*INSITU_GLO_PHYBGCWAV_DISCRETE_MYNRT_013_030]*
Sea surface salinity	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°) *
		Bias, RMSE	Multi observation global Ocean ARMOR3D L4 reanalysis *[MULTIOBS_GLO_PHY_TSUV_3D_MYNRT_015_012]*
Surface currents	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°)
Sea surface height	Time series of ssh anomalies (period	Linear regression	Reanalysis: IBI36RYS (1/36°)

	1994 – 2024)		
	2D surface maps of sla trend (period 1993 – 2024)	None	European Seas Gridded L4 Sea Surface Heights And Derived Variables Reprocessed 1993 Ongoing *[SEALEVEL_EUR_PHY_L4_MY_008_068]*
Mixed layer depth	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°)
		Bias, RMSE	Multi observation global Ocean ARMOR3D L4 reanalysis *[MULTIOBS_GLO_PHY_TSUV_3D_MYNRT_015_012]*
3D variables			
Temperature	1) Hovmöller diagrams (period: 1993 – 2024) 2) Vertical profile (average of the period 1993 – 2024)	Bias, Mean bias, Correl	Reanalysis: IBIRYSv3 (1/12°)
Salinity			

Temperature

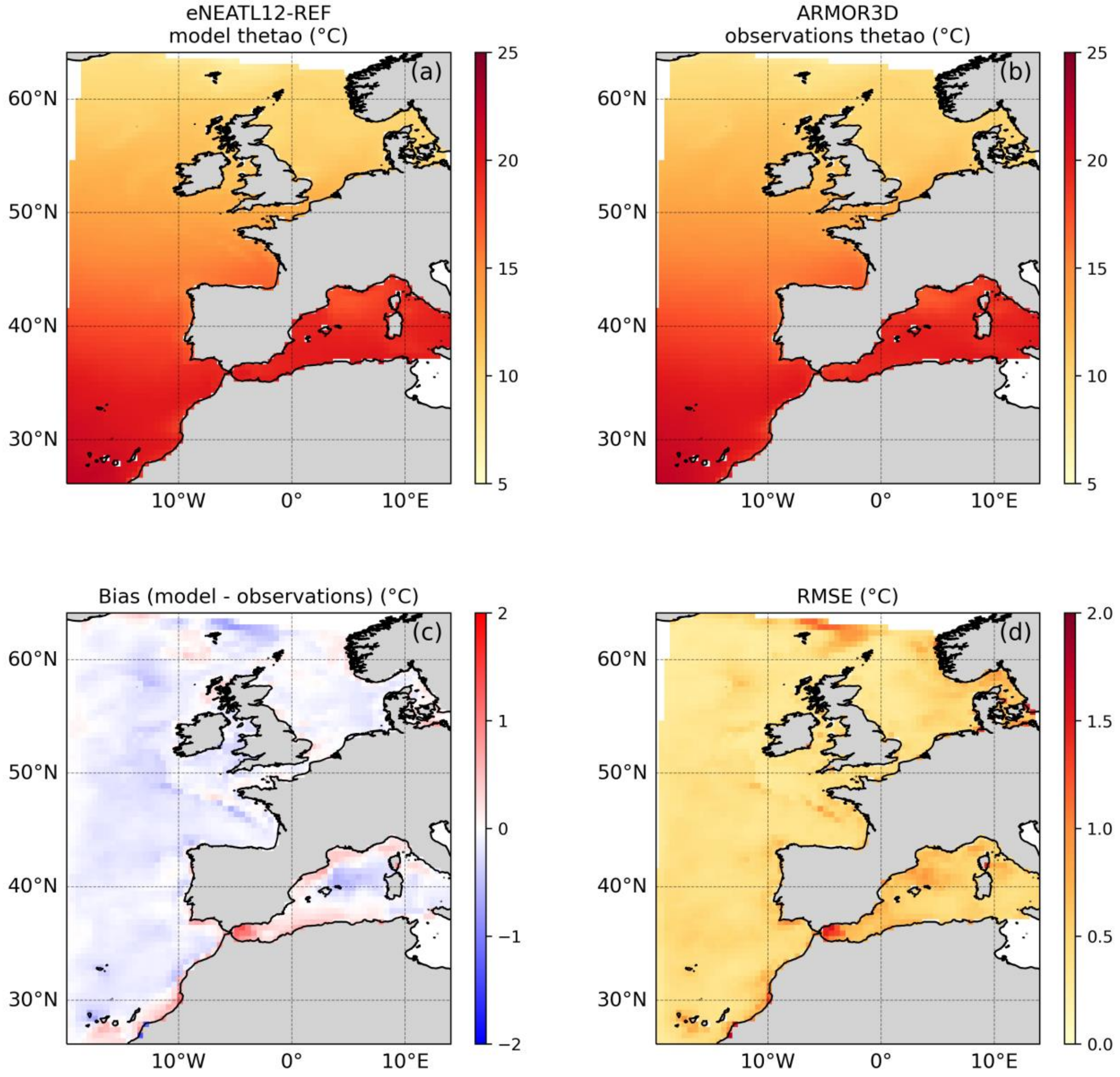


Figure 4. Surface temperature (a) from the hindcast simulation and (b) satellite observations (ARMOR3D), for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period.

Surface ocean temperature is compared to ARMOR3D dataset by calculating the mean bias and RMSE (Figure 4). Another comparison to IBIRYSv3 is also available for the surface in Appendix A. In both cases, comparisons are made over the entire hindcast period (1993 – 2024).

Hindcast surface temperature is generally in good agreements with ARMOR3D temperature: bias (Figure 4c) is mainly included between –0.5°C and 0.5°C and RMSE (Figure 4d) is relatively low (around 0.5°C). The hindcast is almost systematically cooler than the observations, except in two main areas: the Mediterranean coasts (Alboran gyre especially, Figure 4c) and the northwestern African coast. Note that this bias was also observed in GLORYS12v1, which is used to force our hindcast simulation.

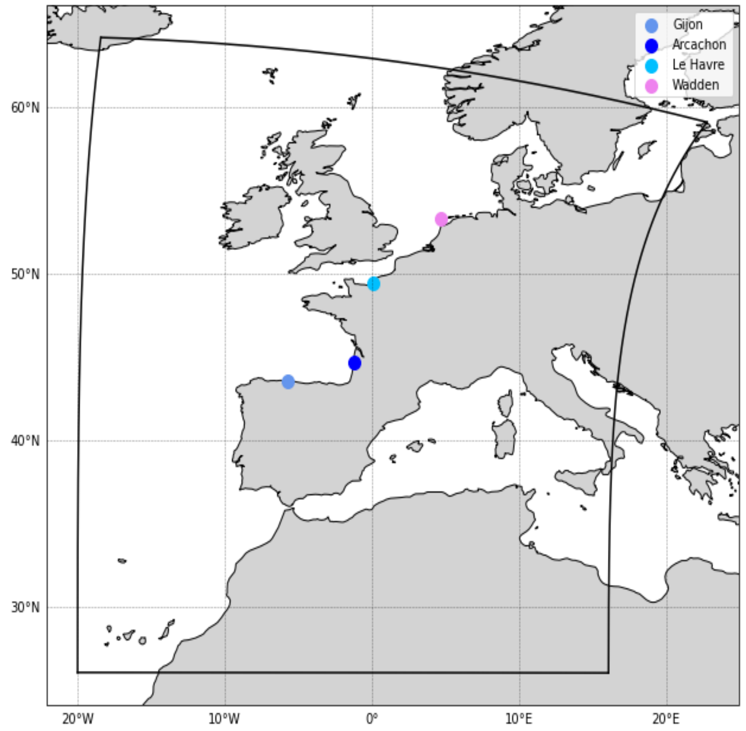


Figure 5. Mooring sites of surface temperature. Black lines represent the domain's limits.

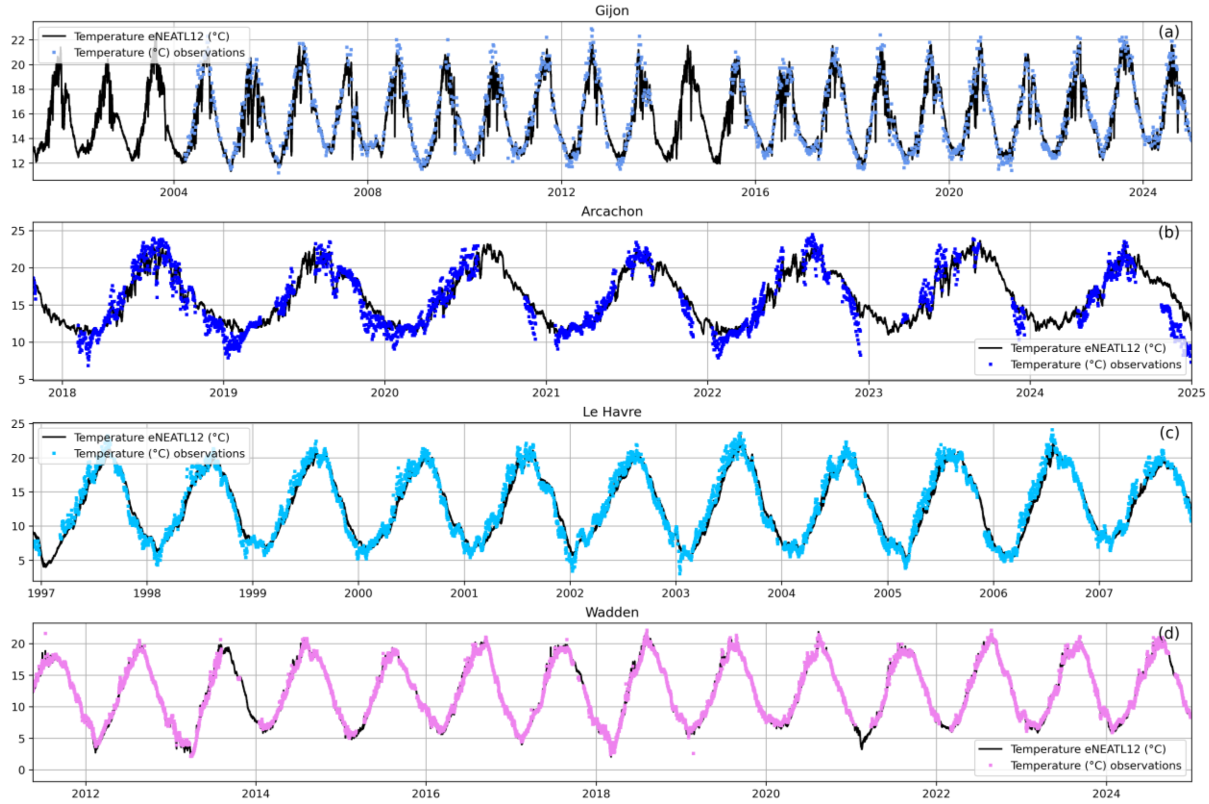


Figure 6. Hourly surface temperature at four mooring sites: (a) Gijon, (b) Arcachon, (c) Le Havre and (d) Wadden (see Figure 5 for moorings site locations). The chosen period depends on the availability of data.

The hindcast surface temperature is also qualitatively compared to mooring sites measurements (hourly measurements). This comparison allows us to test the ability of the model to reproduce surface temperature seasonal cycle at coastal points. Four mooring sites were selected based on the WP5 chosen sites: Gijon, Arcachon, Le Havre and Wadden

(Figure 5). The model manages to reproduce the seasonal cycle at each selected point (Figure 6).

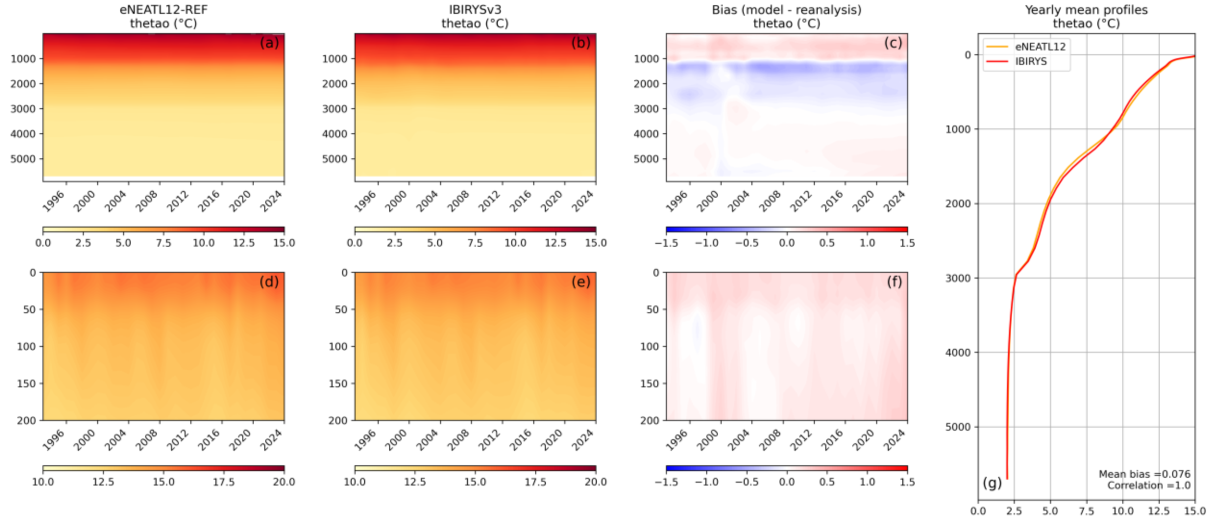


Figure 7. Hovmoller diagrams of mean temperature (°C, domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) is for the hindcast simulation and (b, e) for the reanalysis (IBIRYSv3), for the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (°C, model – reanalysis) for the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis temperature (°C).

Domain averaged temperature of the whole water column is compared to the reanalysis IBIRYSv3 (Figure 7), by calculating the bias and the mean vertical profile. The hindcast temperature is in good agreement with the reanalysis. The bias (Figure 7c) is mainly included between –0.5°C and 0.5°C, showing three different layers based on the bias sign. On the 0-1000m layer, the hindcast temperature is around 0.5°C warmer than the reanalysis. This pattern is even more visible when looking at the Hovmoller diagrams for the 0-200m layer (Figure 7d to f) and surface 2D maps (Fig.A1) where the bias is systematically positive. On the 1000-3000m layer, the hindcast temperature is around 0.5 colder than the reanalysis. Finally, below 3000m the hindcast temperature is slightly higher than the reanalysis (positive bias below 0.5°C). Note that the Hovmoller diagram of bias also shows an event of higher reanalysis bottom temperature for the year 2000.

When looking at the mean vertical profile (Figure 7g), the agreement between the hindcast and reanalysis is even better (correlation of 1 and a slightly positive bias of 0.076), with a good reproduction of the pattern observed for the whole water column (thermocline depth especially).

Salinity

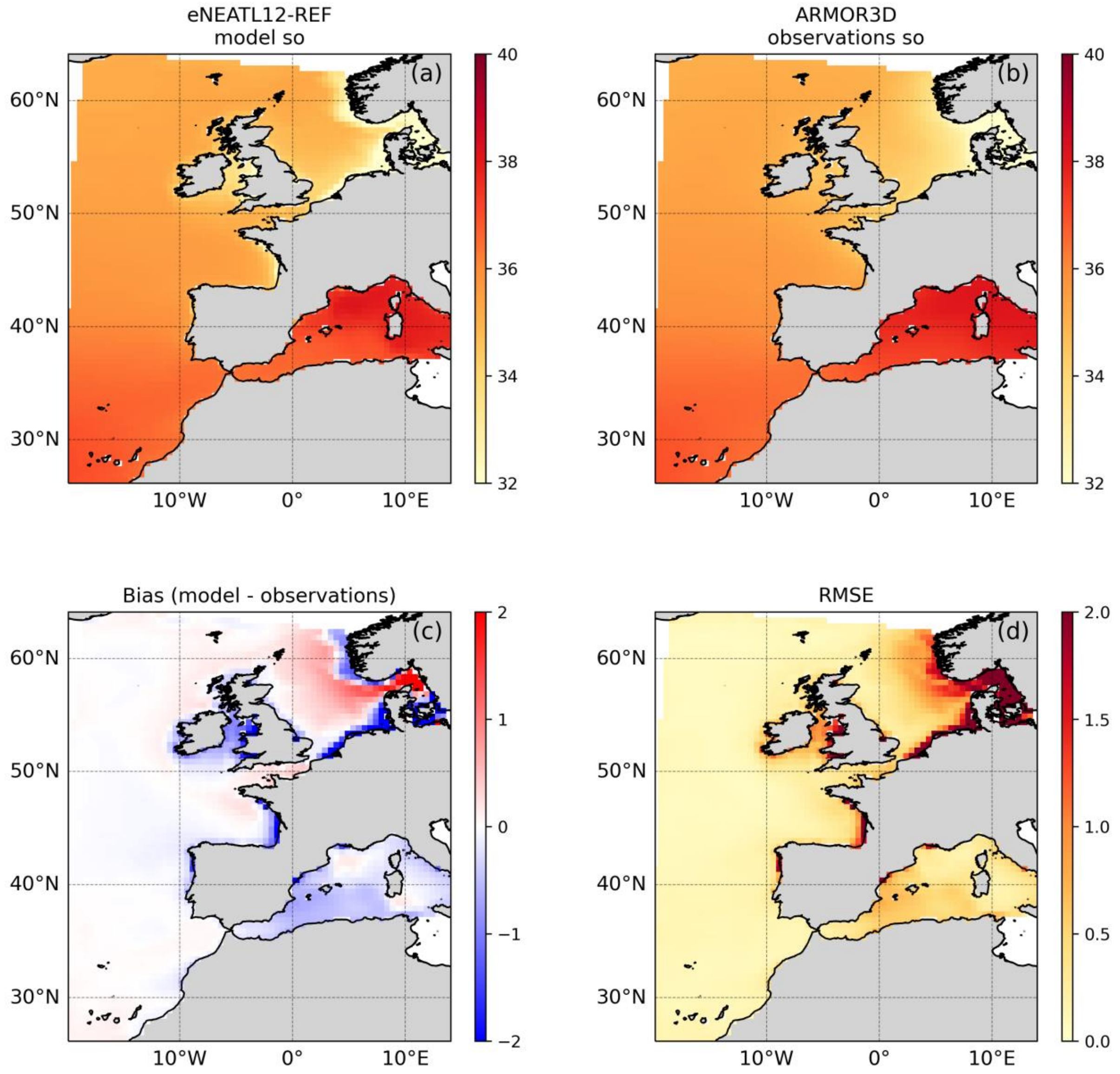


Figure 8. Surface salinity (a) from the hindcast simulation and (b) satellite observations (ARMOR3D), for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period.

Such as surface ocean temperature, surface ocean salinity is first compared to ARMOR3D dataset by calculating the mean bias and RMSE (Figure 8). Another comparison to IBIRYSv3 is available for the surface in Appendix A. In both cases, comparisons are made over the entire hindcast period (1993 – 2024). Offshore, in the Atlantic Ocean open waters, the hindcast's salinity is quite close to the observed one, showing an almost 0 bias and a really close to 0 RMSE. However, when looking at shallower waters, salinity biases and RMSE are higher, especially in the coastal Northeastern area. In these areas, biases are mainly negative (hindcast fresher than the observed one) at the coast, and positive in the North Sea. Salinity near the coast is strongly linked to river forcings, which could explain the obtained bias here. Indeed, we use a climatology (Dai & Trenberth, 2002), which fails to catch the high frequency variability. The Mediterranean Sea also shows higher bias than the Atlantic offshore area, but lower than the one found in the Northern region.

Domain averaged salinity of the whole water column is compared to the reanalysis IBIRYSv3 (Figure 9), by calculating the bias and the mean vertical profile. The three layers (0-1500m, 1500-3000m and below 3000m) reproduced by the reanalyses are well reproduced by our hindcast, same for the decrease of salinity with depth. Bias is generally low (between –0.25 and 0.25). In the surface layer, the bias is mainly negative (hindcast salinity lower than the reanalysis) except around the year 2000 (Figure 9f); in the mid-layer (200m to 1000m) the bias is mainly positive (hindcast salinity higher than the reanalysis) and finally at the bottom layer (below 1000m) the bias is mainly negative (hindcast salinity lower than the reanalysis).

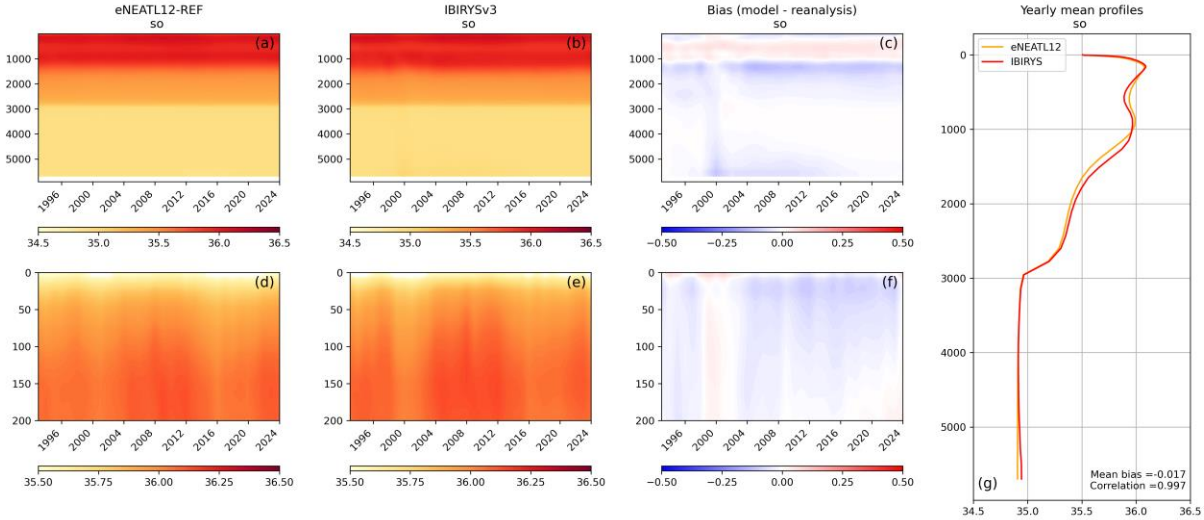


Figure 9. Hovmoller diagrams of mean salinity (domain average). Panels a, b, and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) is for the hindcast simulation and (b, e) reanalysis (IBIRYSv3), for the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (model – reanalysis) for the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis salinity.

When looking at the mean vertical profile (Figure 9g), the agreement between the hindcast and reanalysis is high (correlation of 0.997 and a slightly negative bias of -0.017) with good reproduction of the pattern observed for the whole water column (surface freshening, subsurface increase and decrease from 200m).

Surface currents

Surface currents are compared to IBIRYSv3 over the entire hindcast period (1993 – 2024). At large scale, the hindcast and reanalysis show good agreement (Figure 10, Figure 11). The major currents are visible: one branch of the North Atlantic Drift, the Azores current, the Liguro-Provençal current and the Algerian current.

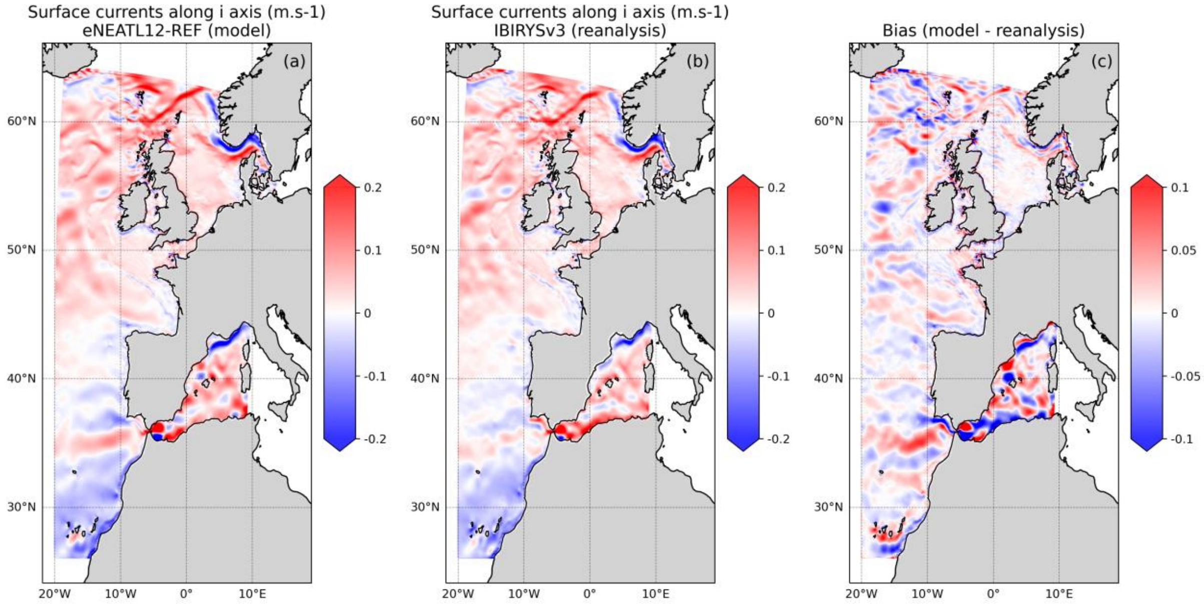


Figure 10. Surface currents (eastward, i-axis) (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) on the same period.

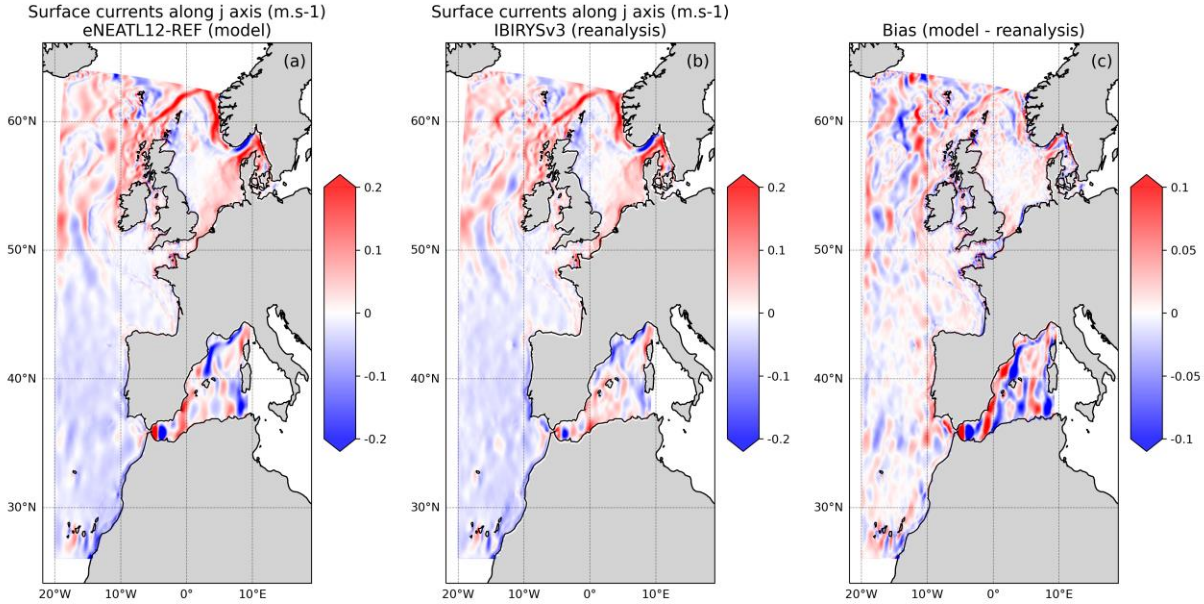


Figure 11. Surface currents (northward, j-axis) (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), on the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) on the same period.

Sea surface height

To validate sea surface height, we compared the hindcast SSH anomalies to IBI36RYS SSH anomalies by calculating a linear regression (Figure 12). The calculation was made by using the 0D variable sshtot (global averaged sea level change) available in the 1/36° reanalysis (not available in the 1/12° reanalysis). Calculated slopes are similar (2.9 mm.yr-1 for the hindcast vs 2.6 mm.yr-1 for the reanalysis). We conclude that the hindcast reproduces well the sea level rise trend depicts by the reanalysis.

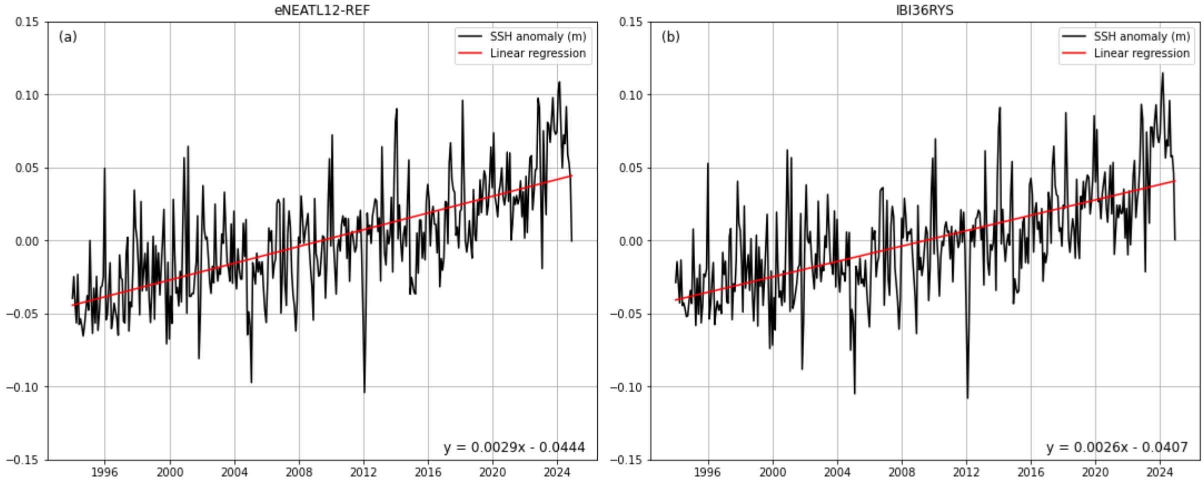


Figure 12. Sea surface height anomalies referenced to the mean sea level over the whole period for (a) the Hindcast simulation and (b) IBI36RYS reanalysis. Anomalies are compared by calculating a linear regression (with formulation in the right bottom corner).

We also compared sea level trend to satellite observations (Figure 13). Overall, the model tends to underestimate the trend described by satellite observations.

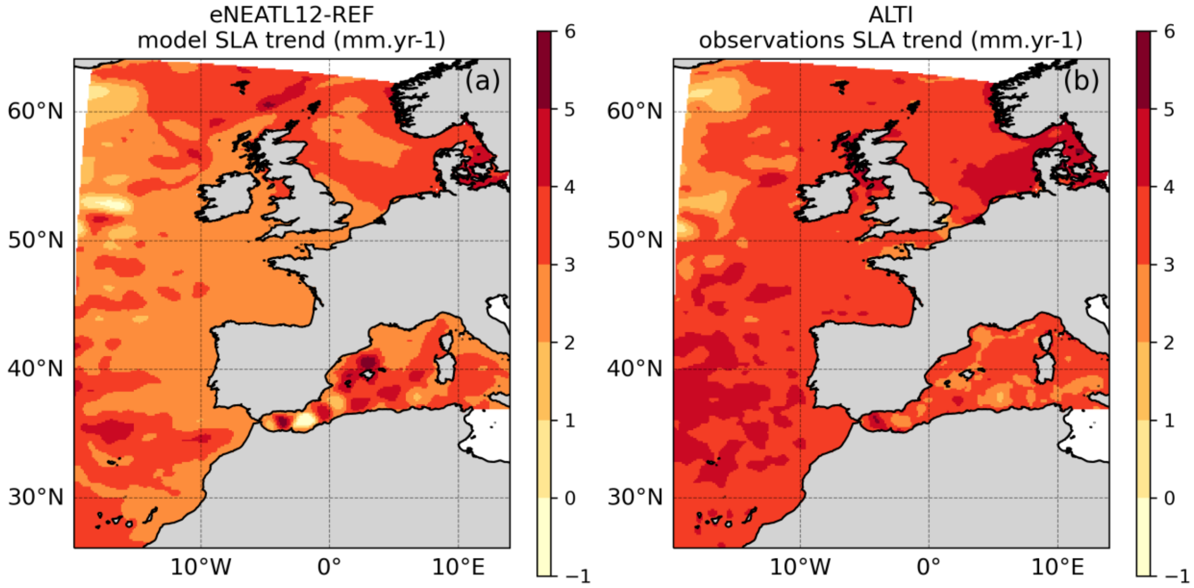


Figure 13. SLA trend calculated for the period 1993 – 2024, for (a) the Hindcast simulation and (b) the Altimetry product.

Mixed layer depth

Mixed layer thickness is compared to ARMOR3D dataset by calculating the mean bias and RMSE (Figure 14) and to IBIRYSv3 by calculating the bias (Figure 15). In both cases, comparisons are made over the entire hindcast period (1993 – 2024). The observation-based product shows the deepest mixed layer depth in the Northwestern corner of the area. The shallowest ones are observed on the continental shelf and in the Mediterranean Sea. This pattern is well reproduced by the hindcast, with, however, systematically higher values (deeper mixed layer depth). The higher bias and RMSE are obtained in the Northwestern area, showing that convective and turbulent processes that affect MLD may be too strong, and that boundaries conditions have a strong impact in this area. High bias and RMSE are also obtained in Gulf of Lion, which is characterised by significant winter deep convection events, not visible when looking at the observations, certainly due to their limited spatial resolution (~ 13km).

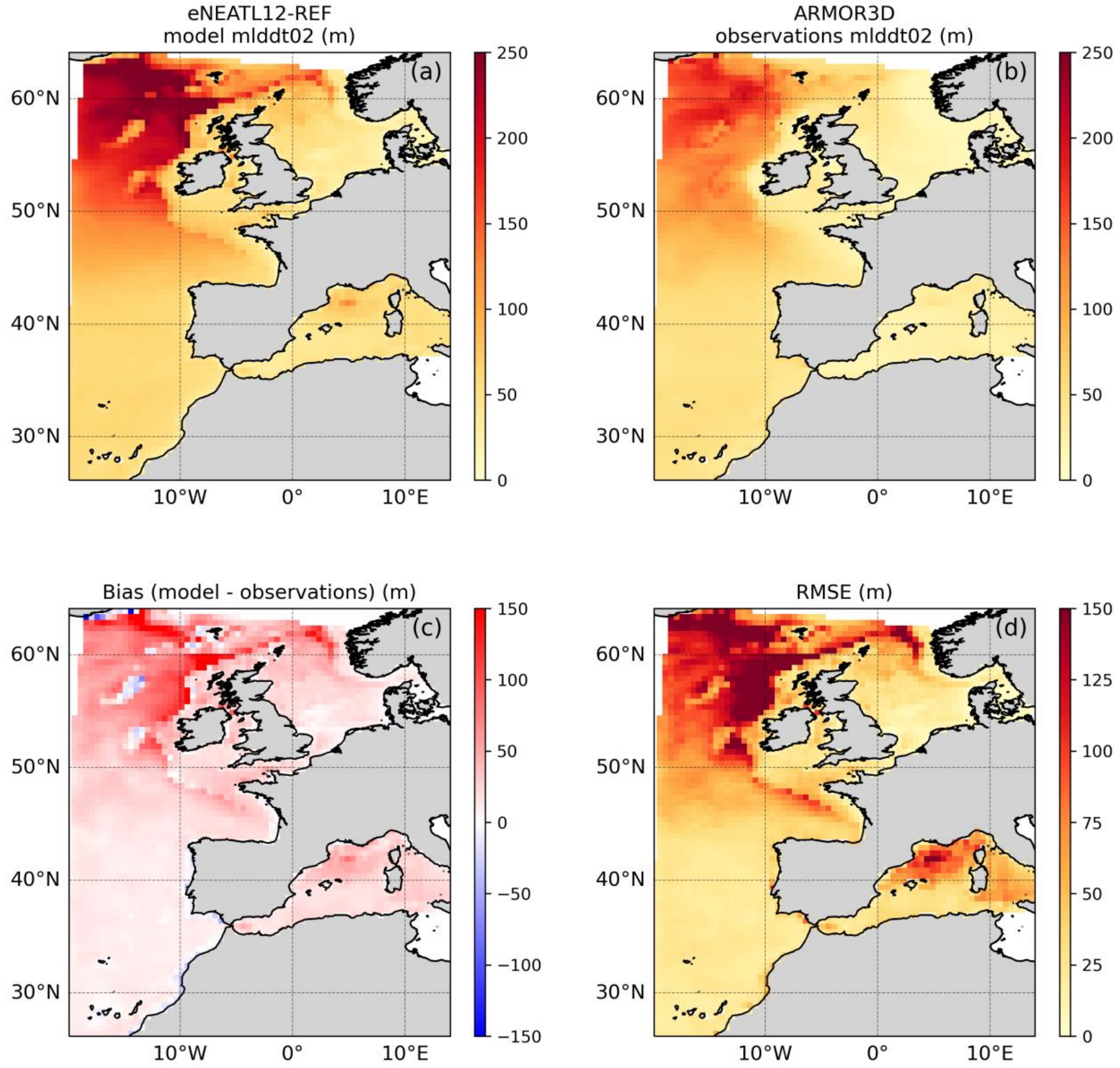


Figure 14. Mixed layer thickness (a) from the hindcast simulation and (b) satellite observations (ARMOR3D), for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period.

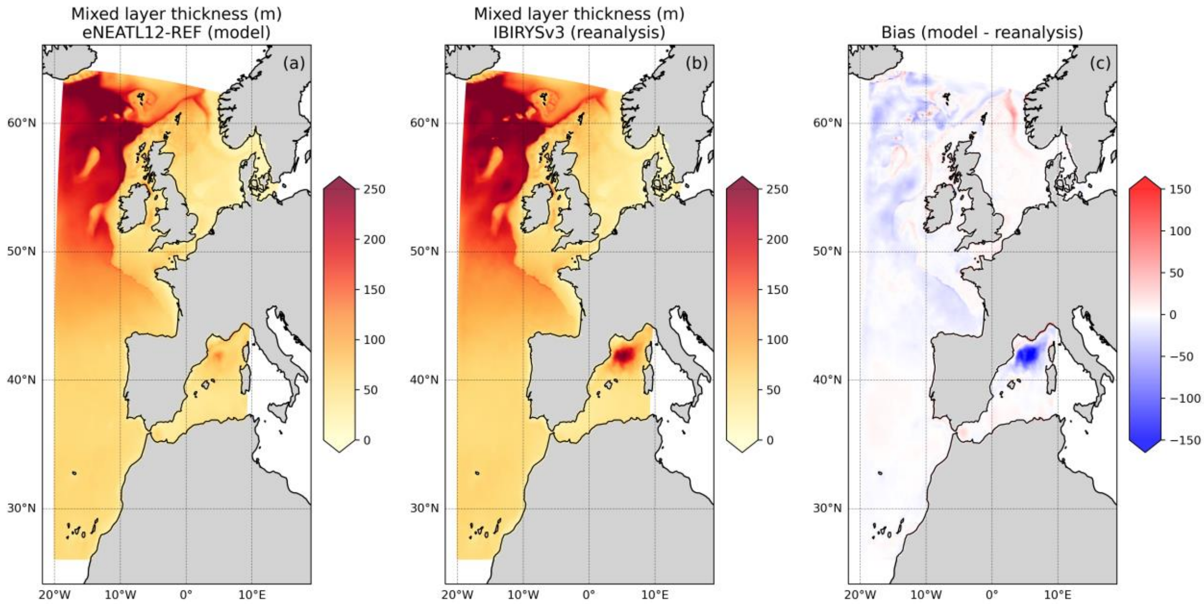


Figure 15. Mixed layer thickness (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period.

Hindcast mixed layer depths are similar to the one obtained in the reanalysis (Figure 15). The most significant bias is obtained in the Gulf of Lion, where the reanalysis shows much deeper mixed layer.

3.1.2. Biogeochemical validation

In this section, we present the validation of biogeochemical variables. The validated variables are specified in Table 4. We also present the used datasets, the metrics and the chosen type of figure for visualization.

Table 4. Validated biogeochemical variables, with associated figures, metrics and reference datasets. (: Figures available in appendix).*

Variable	**Type of figure**	**Metrics**	**Reference dataset**
2D variables			
Surface Chlorophyll	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°) *
	2D surface maps (average of the period 1998 – 2024)	log(bias), RMSE	Mediterranean Sea, Bio-Geo-Chemical, L4, monthly means, daily gapfree and climatology Satellite Observations (1997-ongoing) *[OCEANCOLOUR_MED_BGC_L4_MY_009_144]* Atlantic Ocean Colour (Copernicus-GlobColour), Bio-Geo-Chemical, L4 (daily interpolated) from Satellite Observations (1997-ongoing) *[OCEANCOLOUR_ATL_BGC_L4_MY_009_118]*

Surface total phytoplankton	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°)
Surface total zooplankton	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°)
Surface primary production	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°) *
	2D surface maps (average of the period 1998 – 2024)	Bias, RMSE	Mediterranean Sea, Bio-Geo-Chemical, L4, monthly means, daily gapfree and climatology Satellite Observations (1997-ongoing) *[OCEANCOLOUR_MED_BGC_L4_MY_009_144]* Atlantic Ocean Colour (Copernicus-GlobColour), Bio-Geo-Chemical, L4 (daily interpolated) from Satellite Observations (1997-ongoing) *[OCEANCOLOUR_ATL_BGC_L4_MY_009_118]*
Euphotic zone depth	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°)
Surface nutrients (NO3, PO4, Si) and oxygen	2D surface maps (average of the period	Bias	Reanalysis: IBIRYSv3 (1/12°) *
			WOA2023 climatology (https://www.ncei.noaa.gov/access/world-ocean-atlas-2023/)

	1993 – 2024)		
Carbonate system variables (surface: pH, pCO2 and air-sea CO2 fluxes; Alk*, DIC*)	2D surface maps (average of the period 1993 – 2024)	Bias	Reanalysis: IBIRYSv3 (1/12°) *
		Bias, RMSE	L4 Global Ocean surface carbon product *[MULTIOBS_GLO_BIO_CARBON_SURFACE_MYNRT_015_008]*
3D variables			
Chlorophyll concentration	1) Hovmoller diagrams (period: 1993 – 2024) 2) Vertical profile (average of the period 1993 – 2024)	Bias, Mean bias, Correl	Reanalysis: IBIRYSv3 (1/12°)
Total phytoplankton			
Total zooplankton			
Carbonate system variables (pH, Alk*, DIC*)			
Nutrients (NO3, PO4, Si) and oxygen	1) Hovmoller diagrams (period: 1993 – 2024) 2) Vertical profiles (average of the period 1993 – 2024)	Bias, Mean bias, Correl	Reanalysis: IBIRYSv3 (1/12°)
	Vertical profiles (average	Mean bias,	WOA2023 climatology (https://www.ncei.noaa.gov/access/world-ocean-atlas-2023/)

	of the period 1993 – 2024) at 4 points of the studied area	Absolute bias	

Total chlorophyll concentration

The surface chlorophyll-a concentration (Chl) is first evaluated using satellite ocean colour products for the Atlantic and Mediterranean regions on the period 1998 – 2024 (Figure 16). Another comparison to IBIRYSv3 is also available for the surface on the period 1993 – 2024, in Appendix A.

In the Atlantic area, satellite observations show the following pattern: lower chlorophyll concentrations in the Atlantic subtropical gyre (~ 0.1 $mg.m^{-3}$), these concentrations increase, following the latitude and reach the highest values on the continental shelf (~ 3 $mg.m^{-3}$). Specifically, high values are observed in the shallower parts of the Northern Shelf and northwestern French, Portuguese and Moroccan coast. Productivity is high in these regions due to high nutrients inputs from rivers, or upwelling events.

The hindcast reproduced the lower and higher values in these areas, but the regions cover by these values are generally smaller than the one observed, then leading to important bias and RMSE. On the continental shelf, hindcast chlorophyll concentrations are lower than the observed one. In contrary, offshore, in the Atlantic Ocean, chlorophyll concentrations are higher than the observed ones.

This is especially true along the northern boundary of the domain: high chlorophyll concentrations are modelled but not observed neither in the satellite data nor the reanalysis (Figure 16a and Fig.A3). These values can be explained by phytoplankton and zooplankton biomasses which will be studied in the next parts.

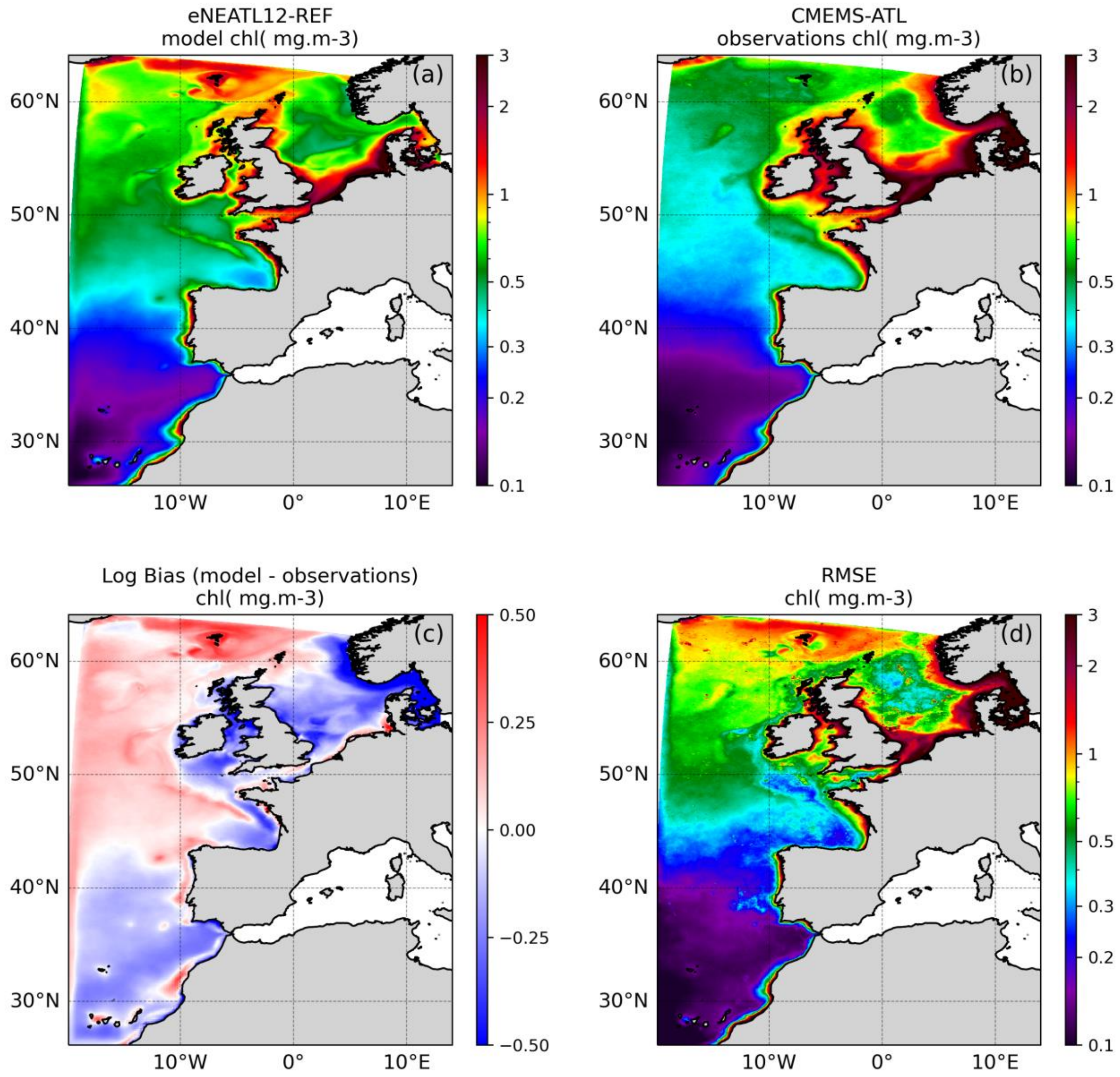


Figure 16. Surface total chlorophyll concentration (a) from the hindcast simulation and (b) satellite observations (CMEMS-ATL), for the period 1998 – 2024 (averaged), in the Atlantic Ocean (Mediterranean excluded). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period.

In the Mediterranean Sea, satellite observations show the higher chlorophyll concentration in the Alboran Sea (near the coasts, gyre excluded) and near the Northern coasts (Figure 17b). The hindcast reproduces the high values obtained in the Alboran Sea and near the Northern coasts. However, it does not manage to reproduce the low concentrations observed in the middle of the Mediterranean Sea, near the Baleares.

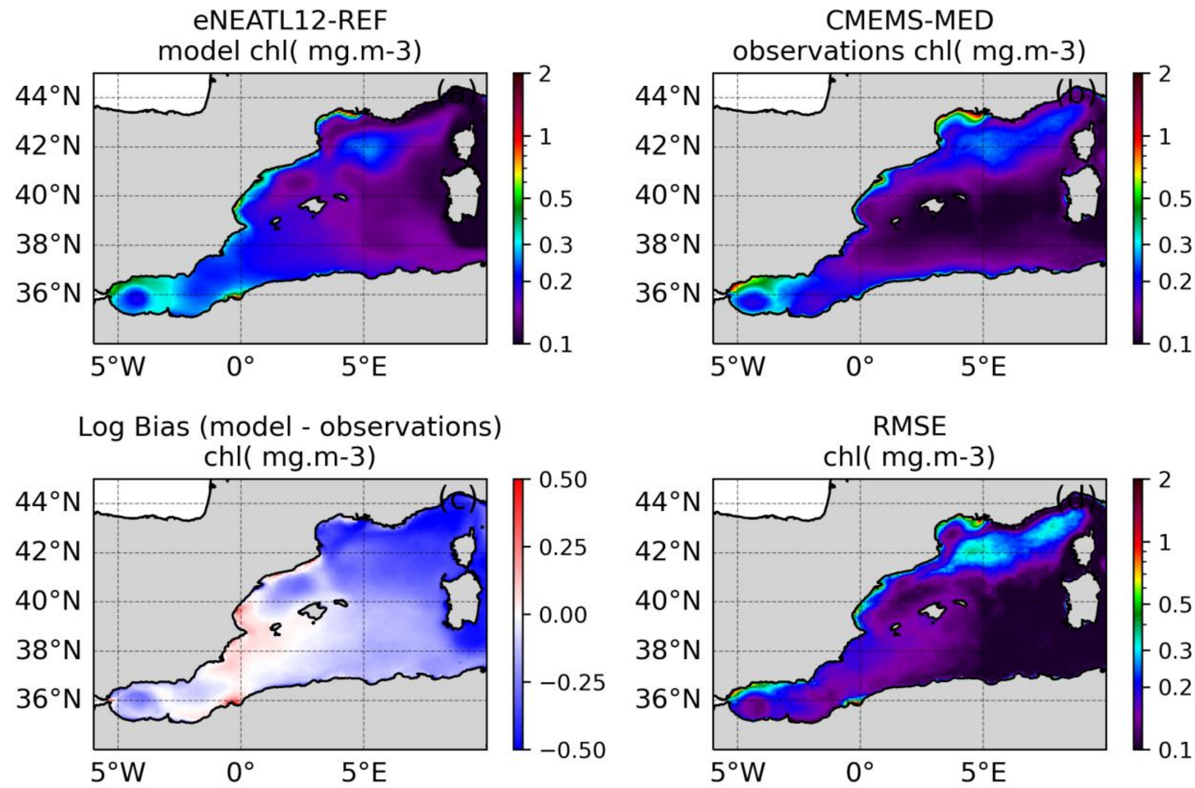


Figure 17. Surface total chlorophyll concentration (a) from the hindcast simulation and (b) satellite observations (CMEMS-MED), for the period 1998 – 2024 (averaged) in the Mediterranean. Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period.

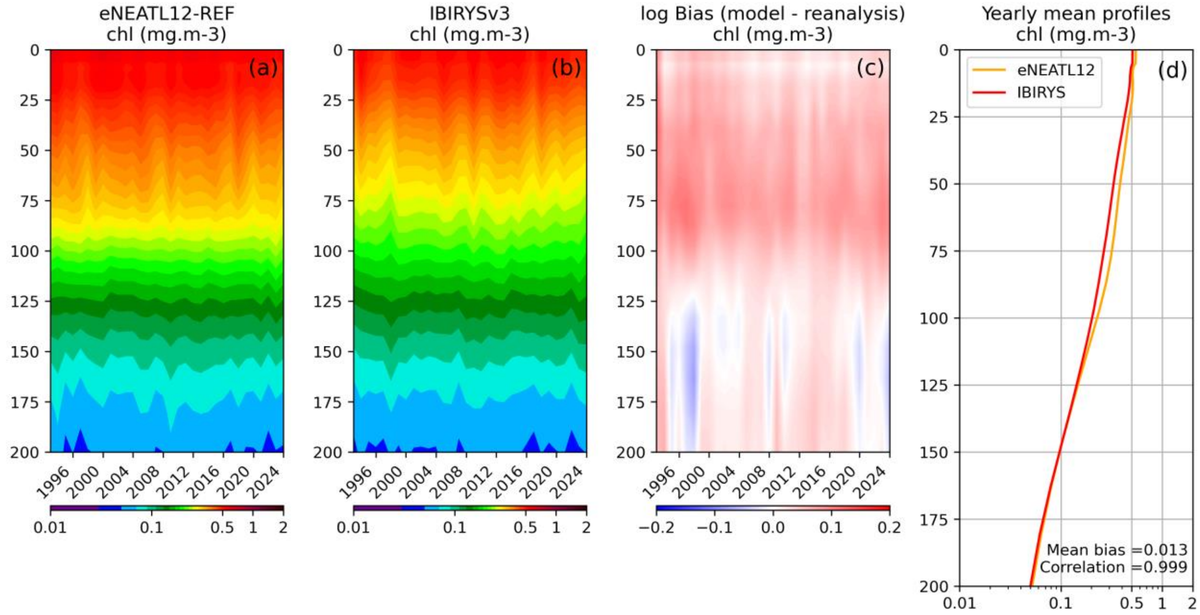


Figure 18. Hovmoller diagrams of mean total chlorophyll concentration (mg.m-3, domain average) (a) for the hindcast simulation and (b) reanalysis (IBIRYSv3), on the period 1993 – 2024. (c) Both are compared by calculating the mean bias (mg.m^{-3}, model – reanalysis) on the same period. (d) Mean vertical profiles (domain and period average) of the hindcast and reanalysis total chlorophyll concentration (mg.m^{-3}). Figure focuses on the 0-200 layer.

Domain averaged chlorophyll of the 0-200m layer is compared to the reanalysis IBIRYSv3 (Figure 18), by calculating the bias and the mean vertical profile. Hindcast chlorophyll concentration is generally higher than the reanalysis one, except from 125m downwards, for specific years. The different layers simulated by the reanalysis are reproduced by the hindcast, but they tend to be deeper as the layer with the highest chlorophyll concentration is wider than the reanalysis one.

Mean vertical profiles exhibit same patterns, showing good agreement between the hindcast and the reanalysis.

Euphotic zone depth

The euphotic zone depth is compared to IBIRYSv3 by calculating the bias (Figure 19). Comparison is made over the entire hindcast period (1993 – 2024). The reanalysis shows the shallowest euphotic zone in the Northeastern area, on the continental shelf and the deepest in the Southwestern corner of the area. This pattern is well reproduced by the hindcast with, however, a shallower euphotic zone depth in the Atlantic and a deeper one in the Mediterranean Sea.

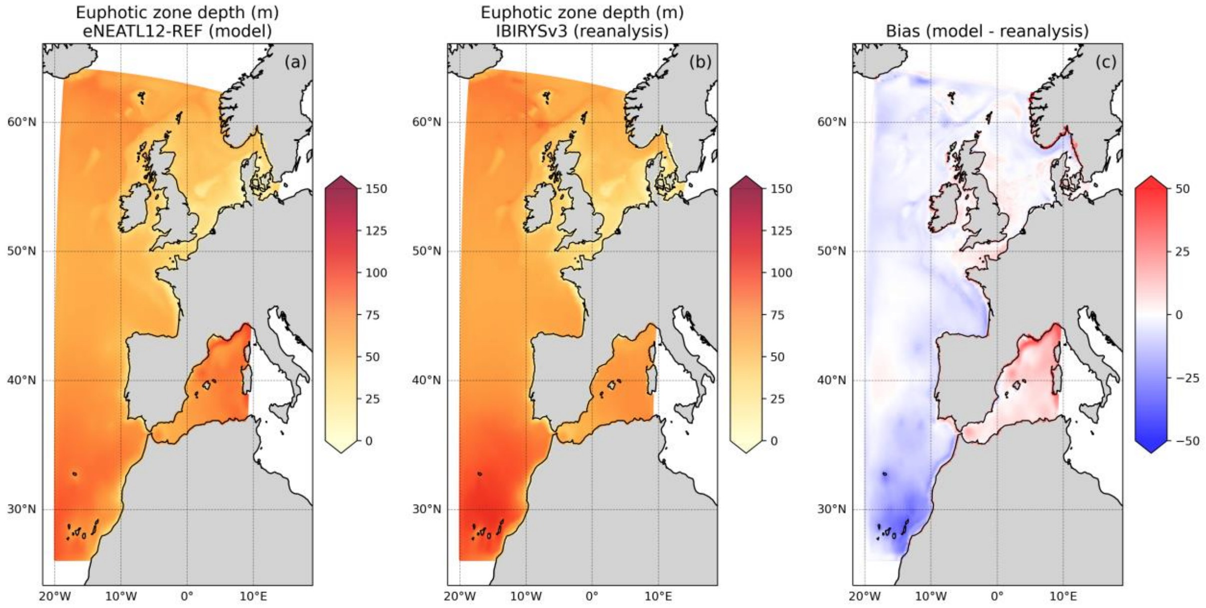


Figure 19. Euphotic zone depth (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period.

The euphotic zone depth pattern is strongly linked to the chlorophyll one. Then, the biases obtained for the chlorophyll can be used to explain the results obtained for the euphotic zone depth. As previously shown in Figure 16, the hindcast shows higher chlorophyll concentration in the southwestern corner of the area, then resulting in shallower euphotic zone. Reasoning in the Mediterranean can be done in the same way: a lower chlorophyll concentration is modelled in the hindcast, then resulting in deeper euphotic zone.

Nutrients and dissolved oxygen concentration

The surface nutrients concentration (including NO_3, PO_4 and Si) is evaluated against WOA2023 climatology. This climatology is compared to the entire hindcast period (1993 – 2024) (Figure 20-Figure 22). Another comparison to IBIRYSv3 is also available for the surface on the period 1993 – 2024, in Appendix A.

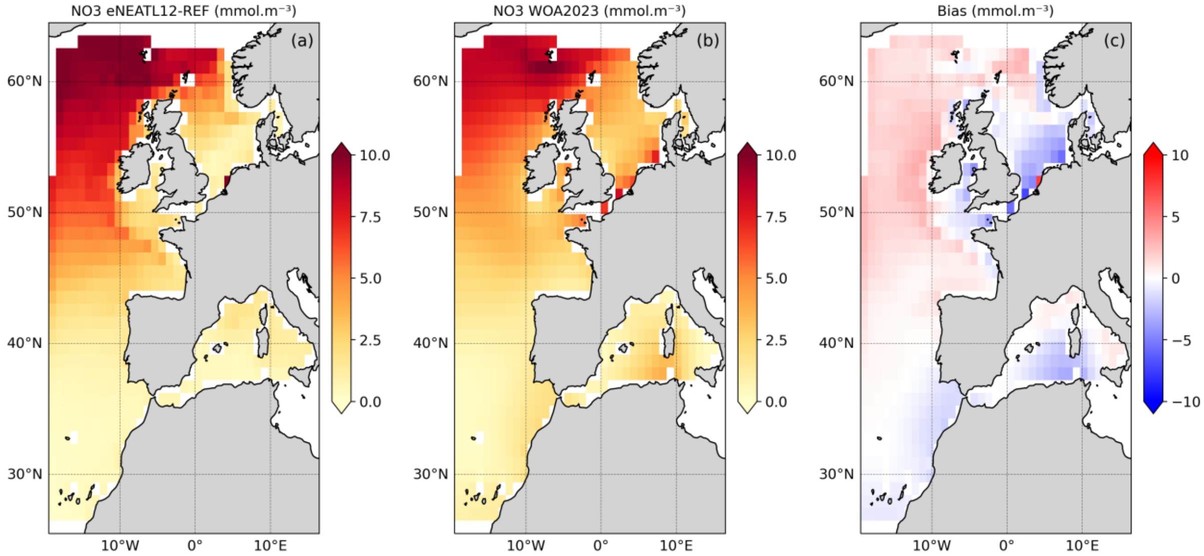


Figure 20. Surface NO_3 concentration (a) from the hindcast simulation and (b) observations (WOA2023), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – observations) for the same period.

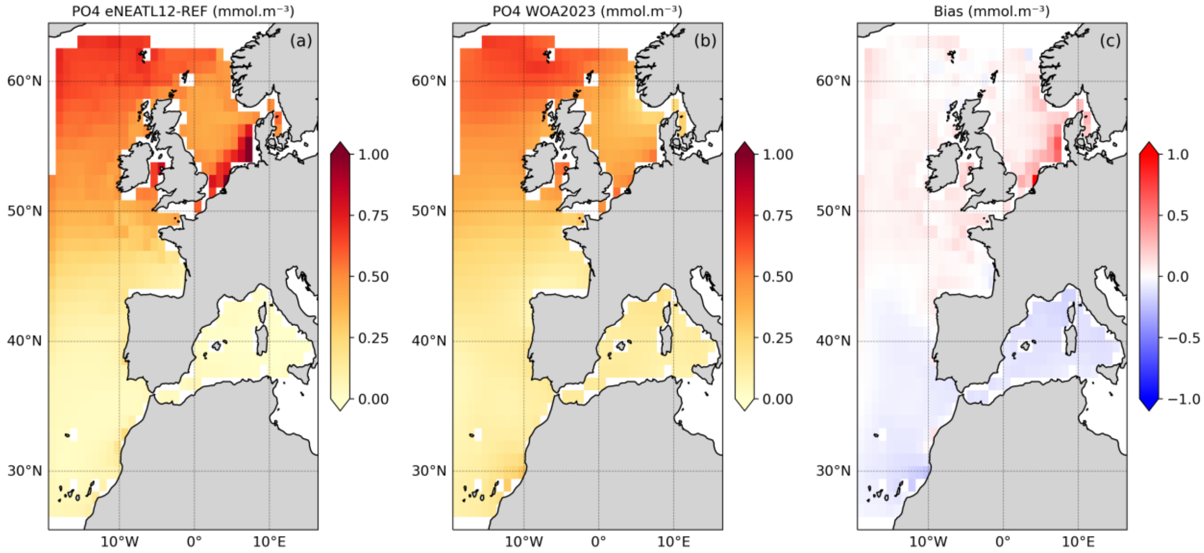


Figure 21. Surface PO_4 concentration (a) from the hindcast simulation and (b) observations (WOA2023), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – observations) for the same period.

Patterns described by observations for NO_3, PO_4 and Si are similar: low NO_3, PO_4 and Si concentrations are observed in the south (Atlantic subtropical gyre) and in the Mediterranean Sea (Figure 20-Figure 22). The concentrations increase to the North with the highest ones observed in the Northwestern area for NO_3 and PO_4 and near the eastern Irelan coast for Si. For NO_3, high values are also observed near the Netherlands coast, in the North Sea.

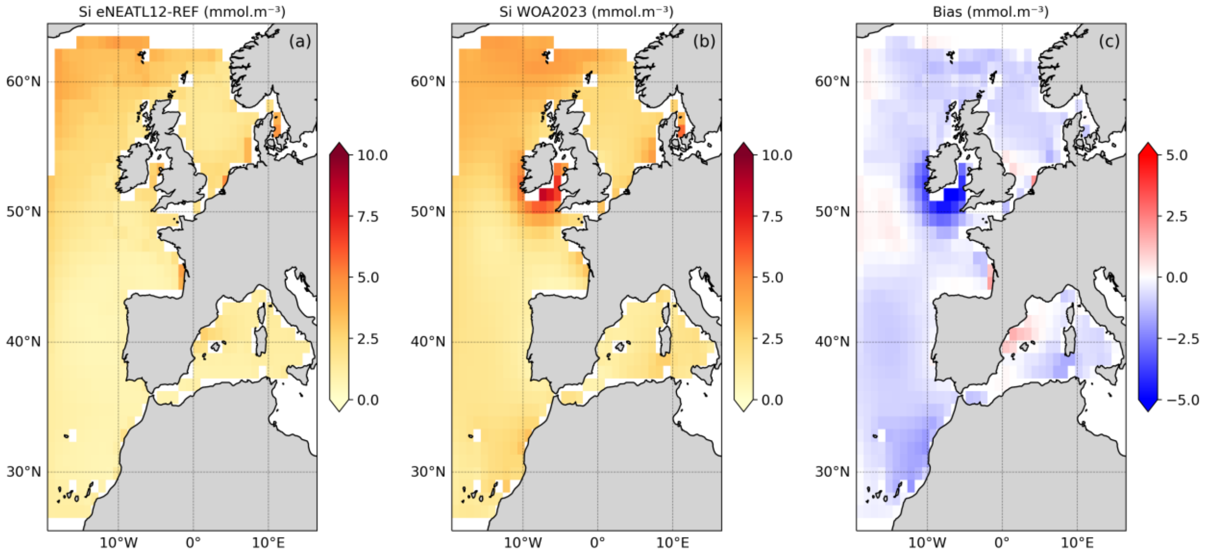


Figure 22. Surface Si concentration (a) from the hindcast simulation and (b) observations (WOA2023), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – observations) for the same period.

Overall, the model manages to reproduce NO_3, PO_4 and Si distributions, however:

- The model shows higher NO_3 concentrations than WOA2023 climatology, in the Northwestern area, where the highest concentrations are observed.
- The model does not manage to reproduce the high NO_3 concentrations observed near the coast of the Netherlands, in the North Sea (Figure 20a).
- NO_3 biases tend to be positive offshore, in the Atlantic (the hindcast concentration is higher than the observed one) and negative in coastal areas and the Mediterranean Sea (the hindcast concentration is lower than the observed one) (Figure 20c).
- The model shows higher PO_4 concentrations than WOA2023 climatology near the coast of the Netherlands, in the North Sea (Figure 21a).
- PO_4 biases tend to be positive in the North of the area and negative in the South (Figure 21c).
- The model does not manage to reproduce the high Si concentrations observed near the eastern Ireland coast (Figure 22).
- Si biases are almost systematically negative, then showing a that surface Si concentration are underestimated by the model (Figure 22c).

Domain averaged nutrients along the whole water column are compared to the reanalysis IBIRYSv3 (Fig.22, 23, 24), by calculating the bias and the mean vertical profile. Multiple patterns are visible on the reanalysis Hovmoller diagrams, especially for NO_3 and Si (Fig.22, 24).

Low NO_3 concentrations are obtained in the surface layer (0-70m, Figure 23e), corresponding to the highest chlorophyll concentrations (Figure 16, Figure 17). The model succeeds in reproducing this pattern but tends to emphasize it: in the 0-70m layer, hindcast concentrations are lower than the one obtained in the reanalysis. Below, hindcast concentrations are higher. These differences emphasize the contrast between the two layers. When looking at the whole water column, NO_3 biases are mainly positive, except in the 1000m layer, showing an overall overestimation of nitrate concentration in the water column. Mean vertical profiles show good agreement between both simulations with increasing bias with depth.

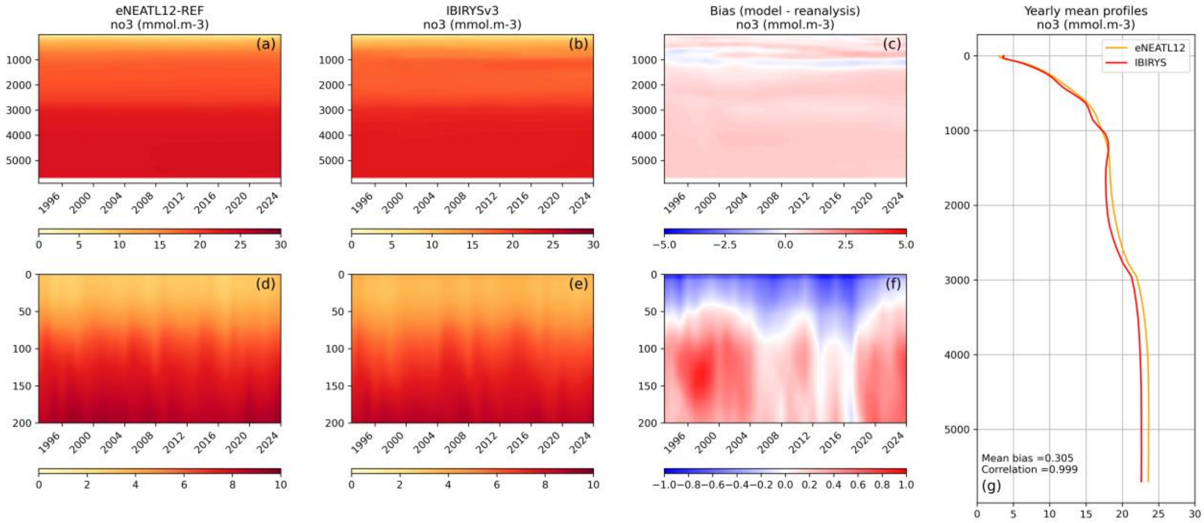


Figure 23. Hovmoller diagrams of mean NO_3 concentration ($mmol.m^{-3}$, domain average). Panels a, b, and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias ($mmol.m^{-3}$, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis NO3 concentration ($mmol.m^{-3}$).

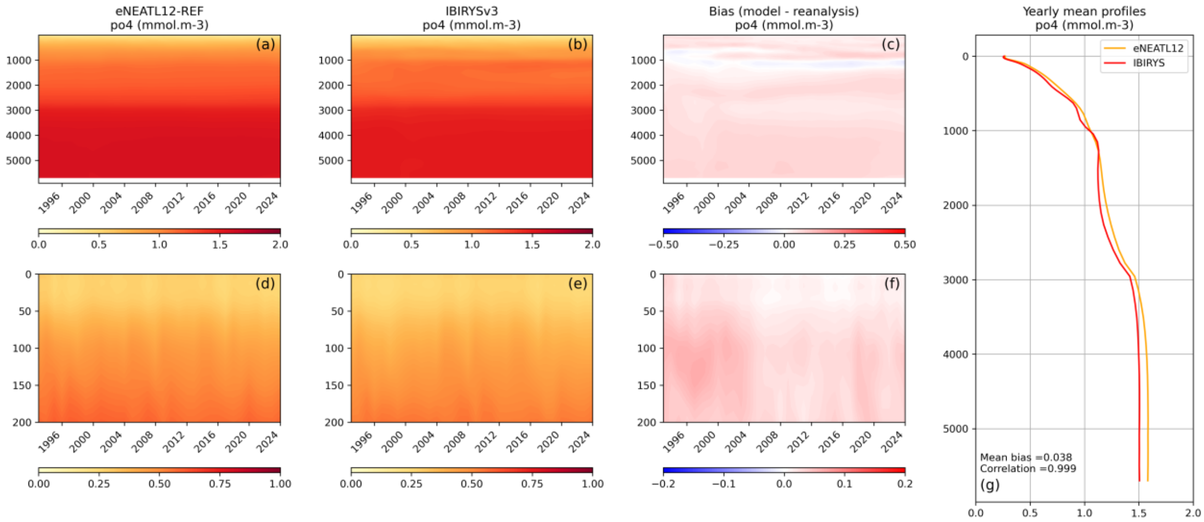


Figure 24. Hovmoller diagrams of mean PO_4 concentration ($mmol.m^{-3}$, domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias ($mmol.m^{-3}$, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis PO_4 concentration ($mmol.m^{-3}$).

PO_4 concentrations exhibit the same patterns than NO_3 with similar biases sign for the whole water column: mainly positive except for the 1000m layer (Figure 24). In the 0-200m layer, contrary to NO_3, the bias sign is homogeneous, showing an overall overestimation of the PO_4 concentration in water column. Mean vertical profiles show good agreement between both simulations with increasing bias with depth.

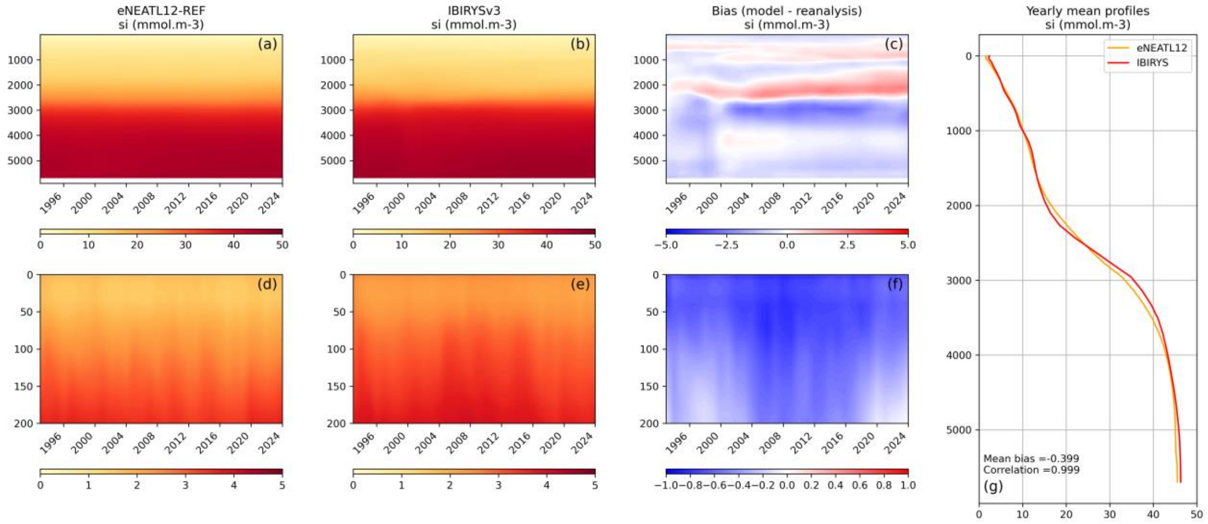


Figure 25. Hovmoller diagrams of mean Si concentration (mmol.m^{-3}, domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis Si concentration (mmol.m^{-3}).

The reanalysis displays low Si concentrations in the 0-2500m layer and higher ones below (Figure 25). The hindcast concentrations reproduce this distribution but are lower in the 0-200m layer, generally higher in the 200-2000m layer and lower below. Mean vertical profiles show good agreement between both simulations with increasing bias with depth.

Regarding the oxygen, the model in general succeeds in reproducing the regional distribution of the surface concentration, that is an increase of the oxygen concentration with the latitude (Figure 26a, b). However, hindcast concentrations are lower than the one from WOA2023, for almost all the regions (negative bias, Figure 26c).

Over the whole water column, we can make the following observations:

- In the surface layer, hindcast oxygen concentrations are lower than the reanalysis and observations ones (Figure 27f and Figure 26c).
- The oxygen minimum layer tends to be closer to the surface and less marked in the hindcast (Figure 26a, b). In this layer, the hindcast shows a positive bias. This difference between both simulations is also exhibited in the mean vertical profiles which show lower oxygen concentrations at 1000m in the reanalysis (Figure 27g).
- In the bottom layer, the bias is mainly positive (hindcast concentrations higher than the reanalysis ones).
- Except the fact that that the oxygen minimum layer tends to be less marked in the hindcast, mean vertical profiles show good agreement between both simulations.

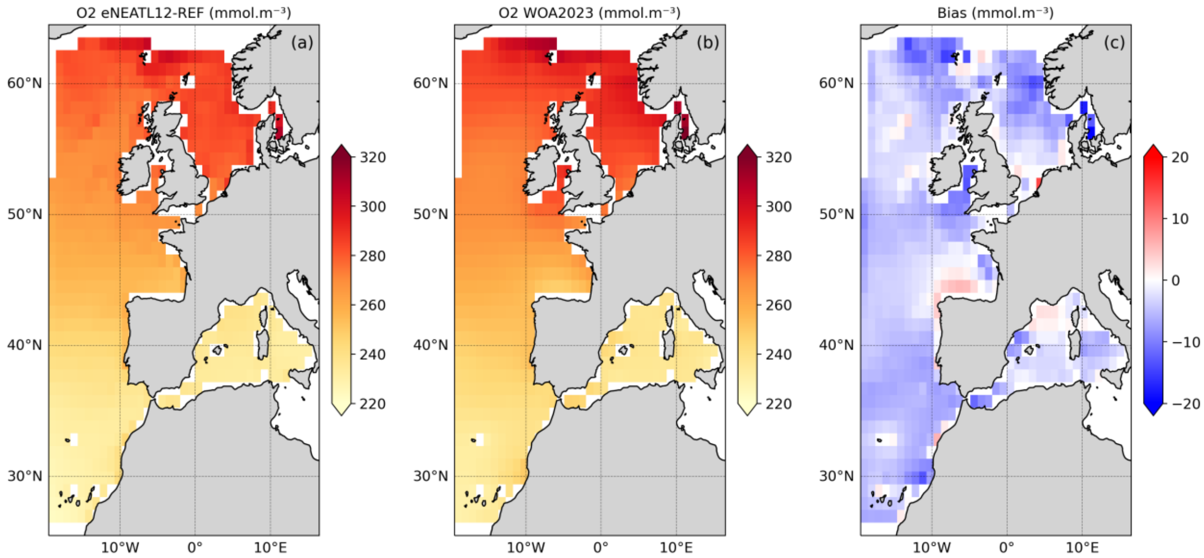


Figure 26. Surface O_2 concentration (a) from the hindcast simulation and (b) observations (WOA2023), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – observations) for the same period.

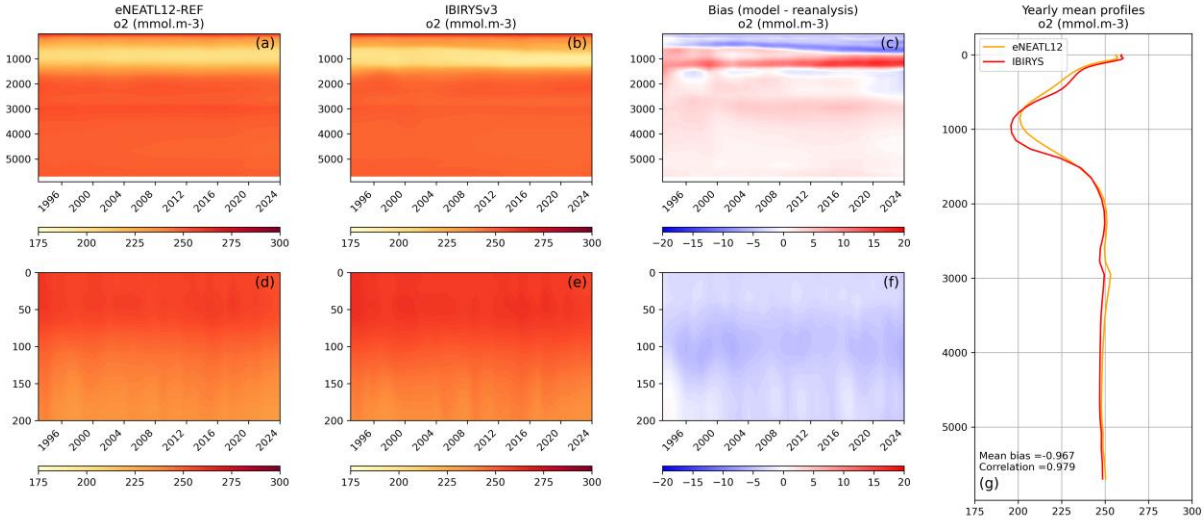


Figure 27. Hovmoller diagrams of mean O_2 concentration (mmol.m^{-3}, domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis O_2 concentration (mmol.m^{-3}).

Finally, a set of vertical profiles was selected to test the ability of the hindcast simulation to reproduce the distribution of nutrients and oxygen in the water column at different locations of the studied area. Four points were chosen to cover representative regions in the area: three in the Atlantic Ocean (North, center and south) and one in the Mediterranean Sea (Figure 28).

The hindcast is compared to WOA2023 climatology by calculating the bias and the absolute bias.

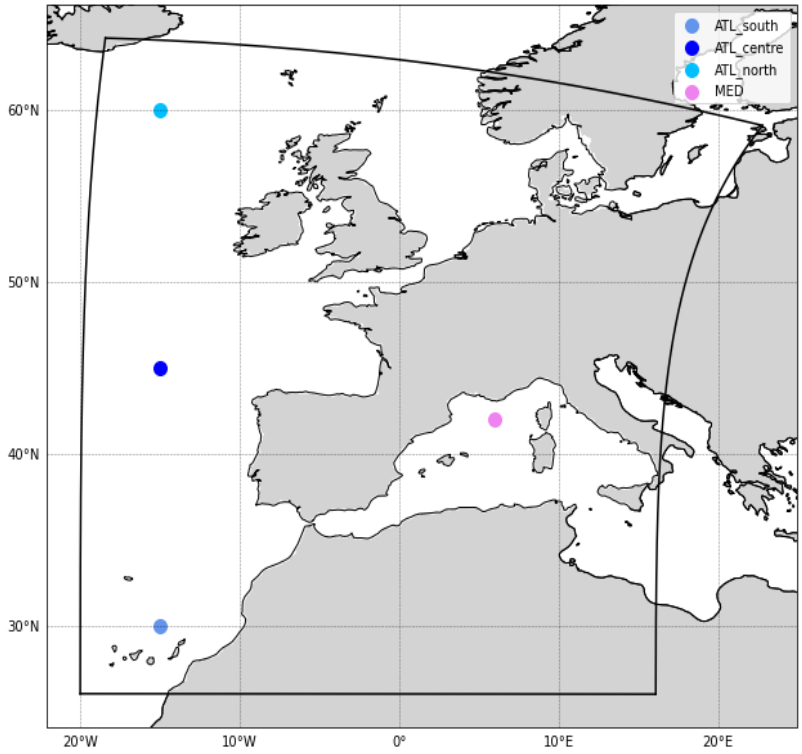


Figure 28. Profiles locations. Black lines represent the area limits.

Mean vertical profiles for South and central Atlantic points show good agreement between model and observations (Figure 29). The model manages to reproduce the vertical distribution of NO_3, PO_4, Si and O_2 but the minimum values of NO_3 (Figure 29a, 500 – 1000m layer) and O_2 (Figure 29d, h, 1000m) tends to be underestimated by the model. At the North Atlantic point, NO_3 concentrations are overall overestimated by the model. PO_4 and Si concentrations were well reproduced except between 600 and 800 meters, where observations show a decrease in these concentrations. O_2 is the least well represented as the hindcast failed at reproduce the observations variation range. Mediterranean profiles show the most differences between model and observations with high absolute bias for each variable.

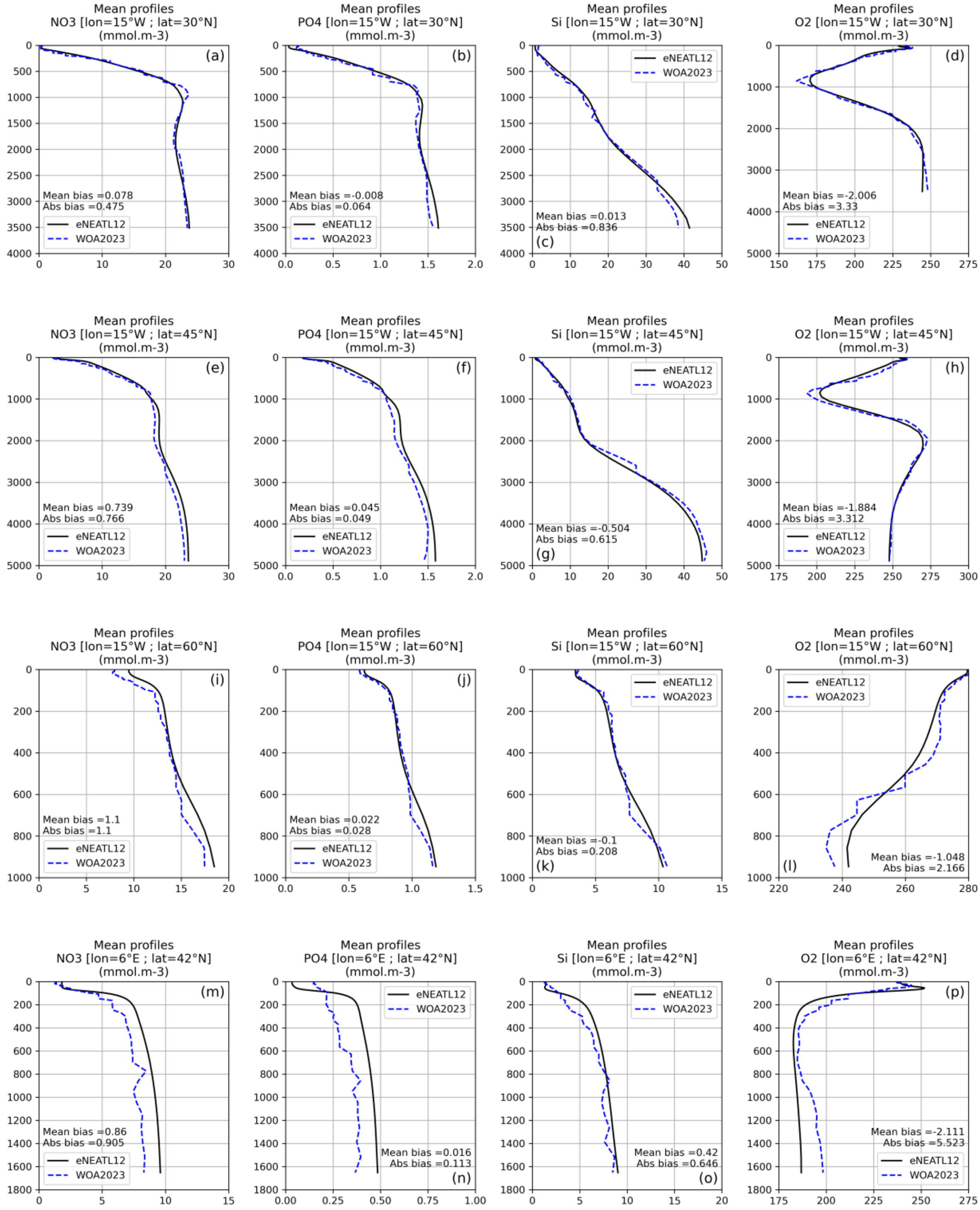


Figure 29. Mean vertical profiles (average of the period 1993 – 2024) at 4 points (3 points in the Atlantic ocean and 1 point in the Med sea, Figure 28 for locations), of (a, e, i, m) hindcast and WOA2023 climatology NO_3 concentration, (b, f, j, n) hindcast and WOA2023 climatology PO_4 concentration, (c, g, k, o) hindcast and WOA2023 climatology Si concentration, and (d, h, l, p) hindcast and WOA2023 climatology O_2 concentration.

Phyto- and zooplankton

The surface total phytoplankton (nano-phytoplankton + diatoms) and zooplankton (micro- + meso-zooplankton) carbon biomass is evaluated using the reanalysis as reference over period (1993 – 2024) (Figure 30, Figure 31).

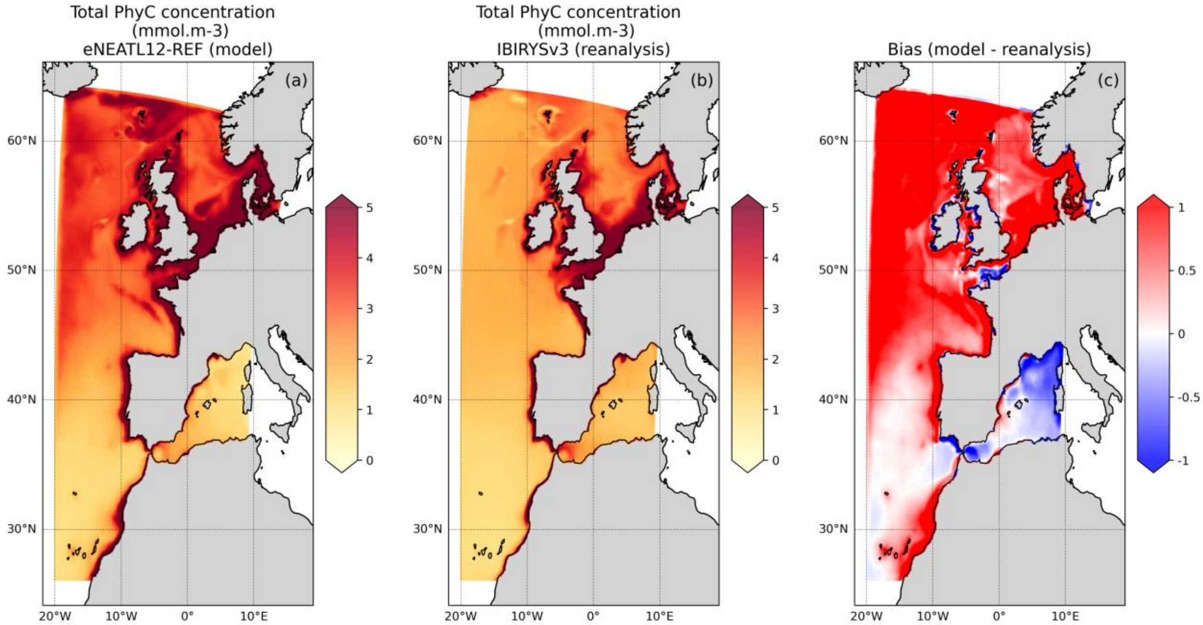


Figure 30. Surface total phytoplankton concentration in carbon (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period.

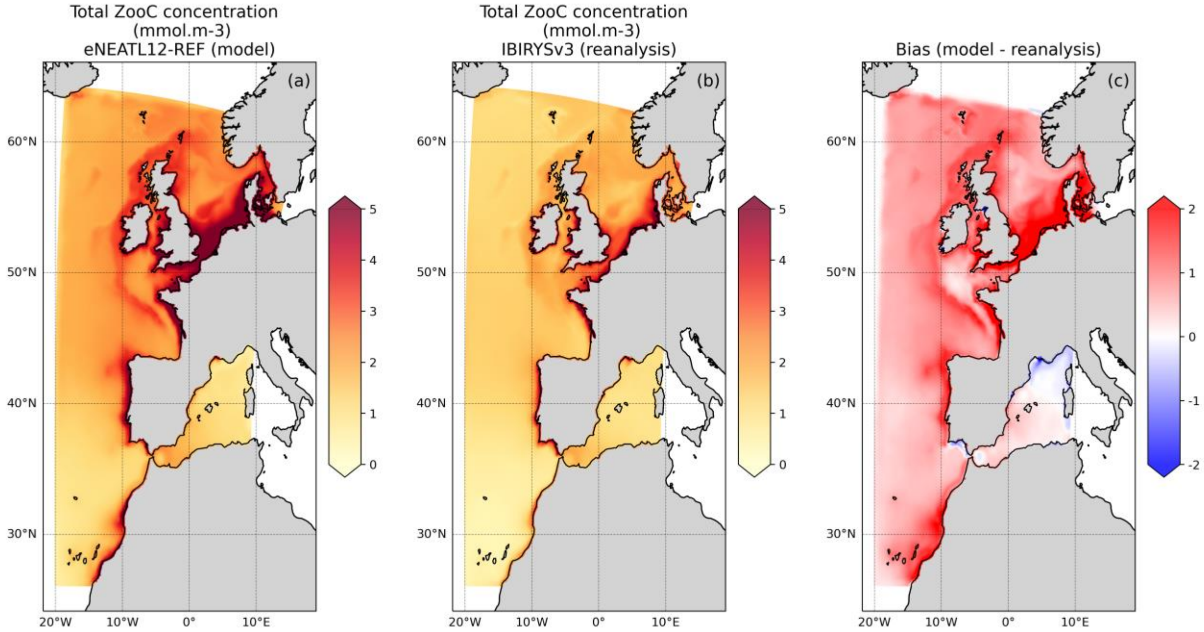


Figure 31. Surface total zooplankton concentration in carbon (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period.

In the reanalysis, high biomasses of phytoplankton and zooplankton are modelled in coastal areas, especially in the English Channel and the North Sea. High phytoplankton biomass covers a larger area than high zooplankton biomass which is concentrated at the coasts. In both cases, the hindcast reproduces this distribution but exhibits much higher biomasses, leading to strong positive biases over almost the entire area (Figure 30, Figure 31).

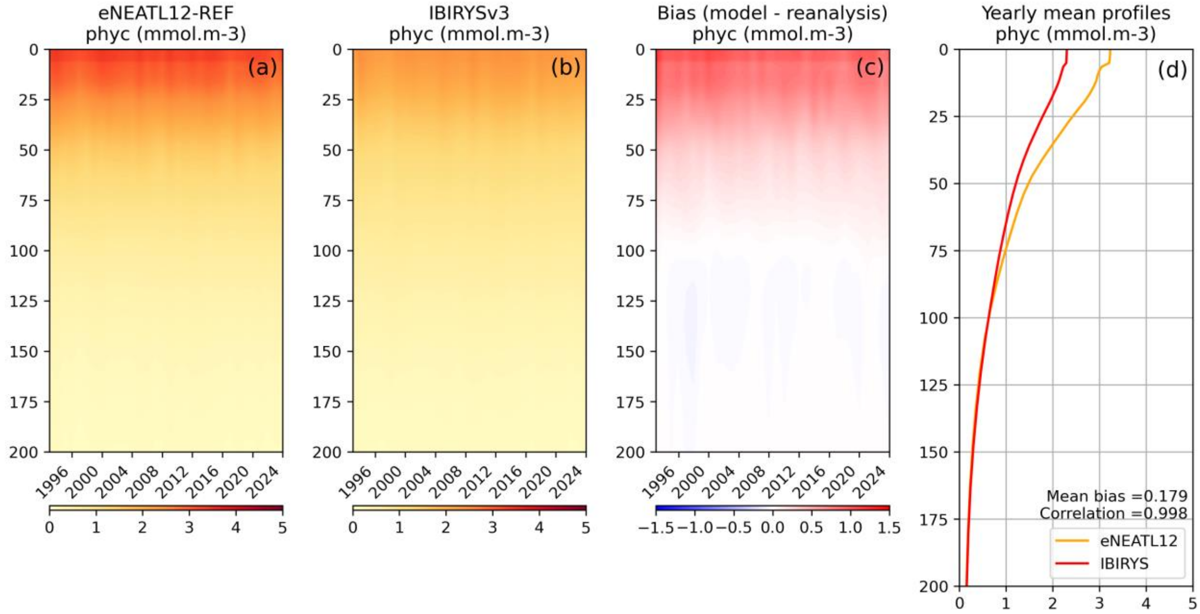


Figure 32. Hovmoller diagrams of mean total phytoplankton (mmol.m^{-3}, domain average) (a) for the hindcast simulation and (b) reanalysis (IBIRYSv3), on the period 1993 – 2024. (c) Both are compared by calculating the mean bias (mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis total phytoplankton concentration (mmol.m^{-3}). Figure focuses on the 0-200m layer.

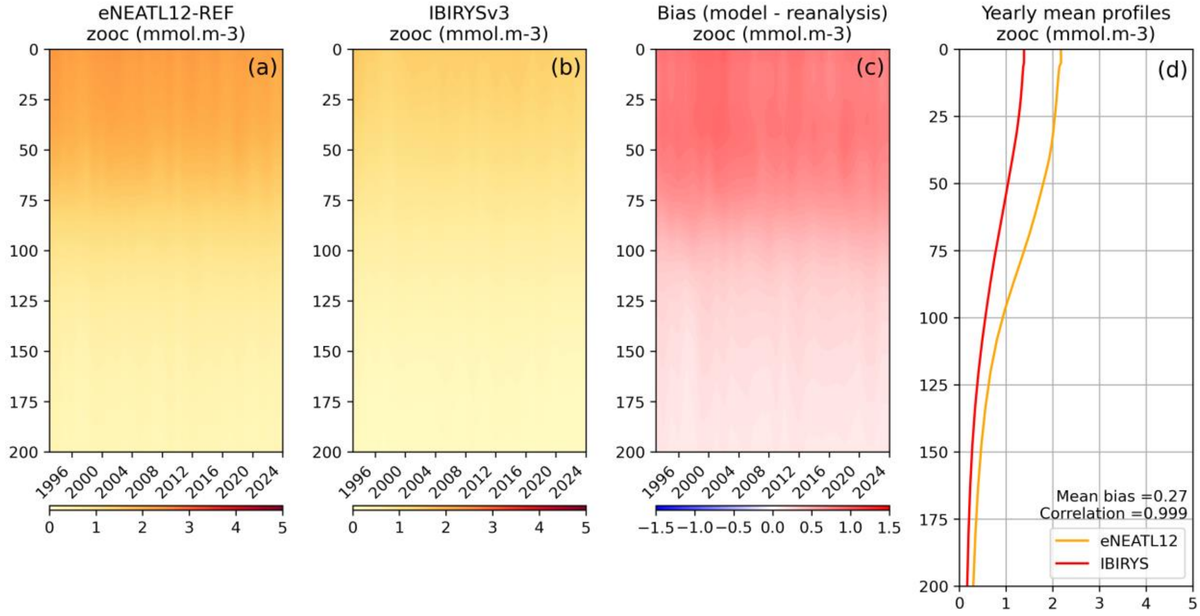


Figure 33. Hovmoller diagrams of mean total zooplankton (mmol.m^{-3}, domain average) (a) for the hindcast simulation and (b) reanalysis (IBIRYSv3), on the period 1993 – 2024. (c) Both are compared by calculating the mean bias (mmol.m^{-3}, model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis total zooplankton (mmol.m^{-3}). Figure focuses on the 0-200m layer.

Domain averaged phytoplankton and zooplankton carbon biomass of the 0-200m layer are compared to the reanalysis IBIRYSv3 (Figure 32, Figure 33), by calculating the bias and the mean vertical profile. As shown with the previous 2D surface maps, biomasses in the surface layer are overestimated. The bias tends to decrease with depth as biomass are close to 0.

Primary production

Depth integrated primary production (PP) is compared to the reanalysis (Figure 34) and satellite derived data (Figure 35, Figure 36). We can make the following observations:

- Reanalysis and observations both show high PP values in coastal areas which are reproduced by the hindcast.
- Both comparisons exhibit high biases and RMSE agreeing on the fact that the hindcast overestimates the PP in the Atlantic.
- Comparison with satellite data also exhibits strong negative biases in the south of the area and in the English Channel and North Sea.
- In the Mediterranean Sea, strong negative biases are obtained in the western area, while strong positive biases are obtained in the eastern area (except for the Alboran gyre), compared to both reanalysis and observations
- Finally, PP is strongly linked to chlorophyll, thus showing similar biases patterns.

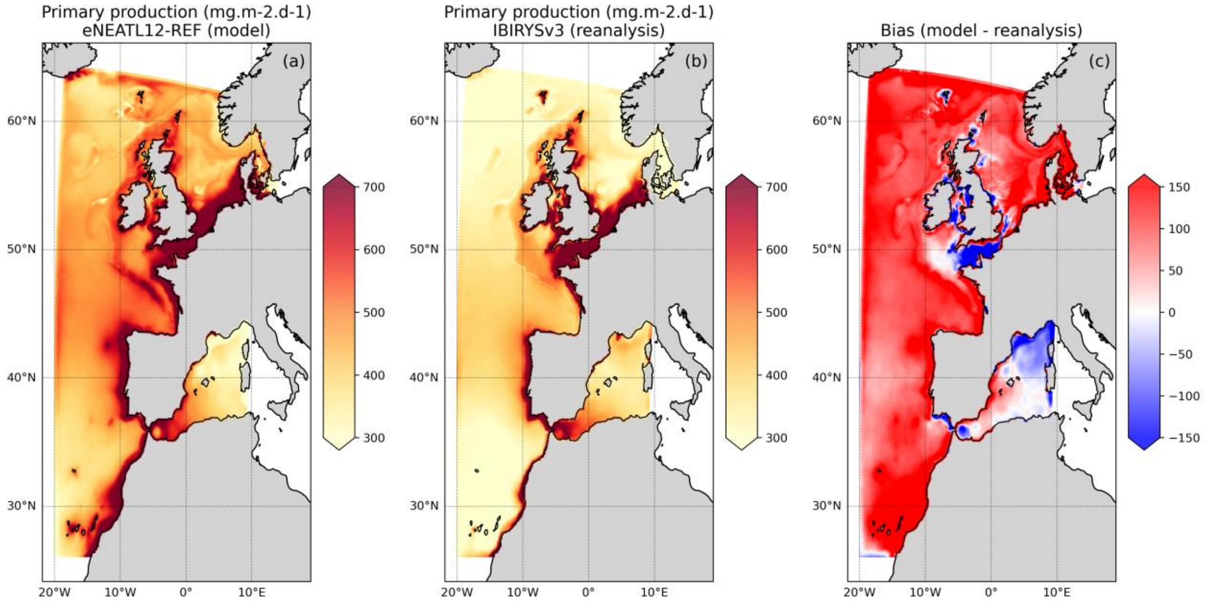


Figure 34. Depth integrated primary production (a) from the hindcast simulation and (b) reanalysis (IBIRYSv3), for the period 1993 – 2024 (averaged). (c) Both are compared by calculating the mean bias (model – reanalysis) for the same period.

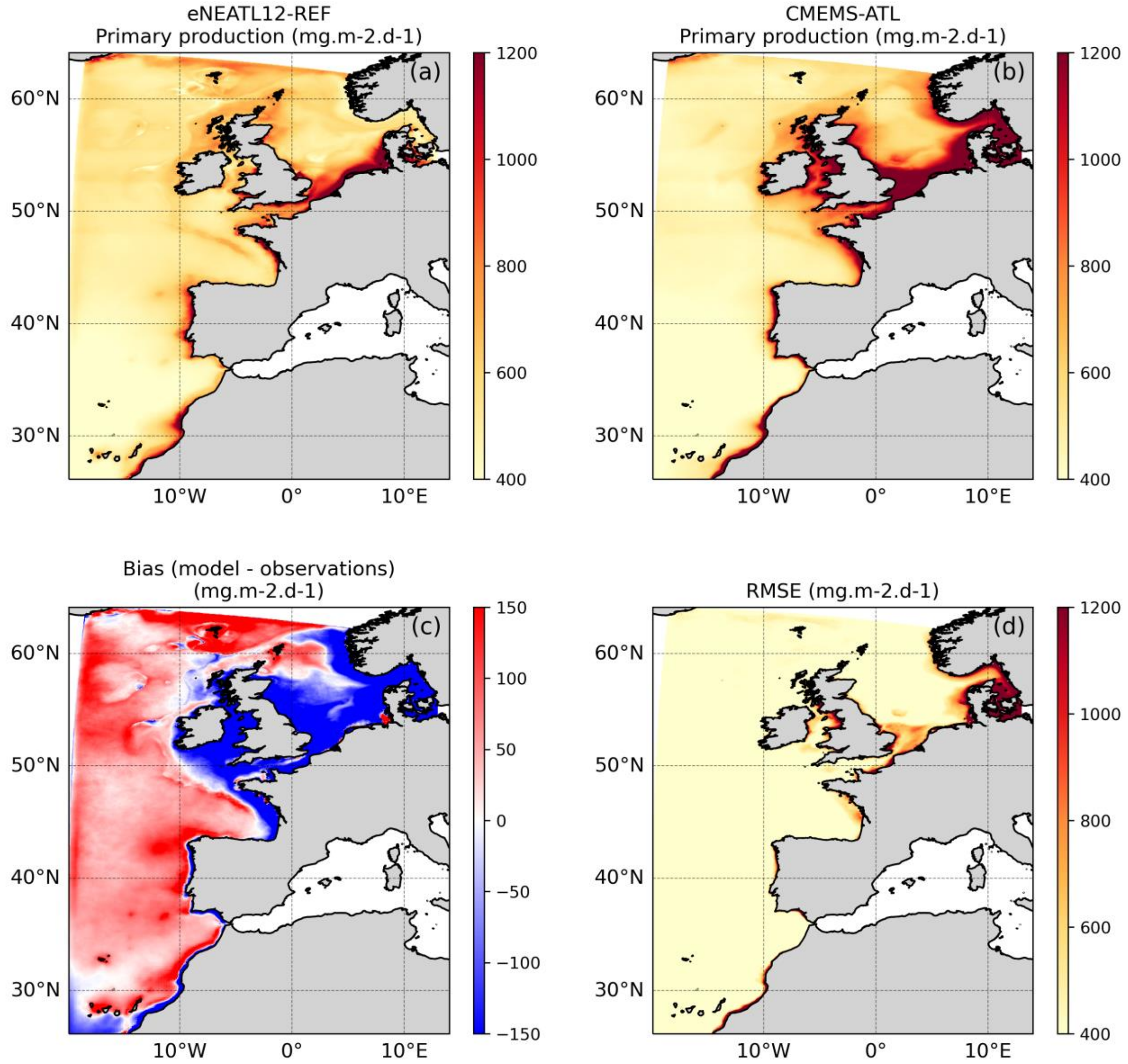


Figure 35. Depth integrated primary production (a) from the hindcast simulation and (b) satellite observations (CMEMS-ATL), in the Atlantic Ocean (Mediterranean excluded), for the period 1998 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period.

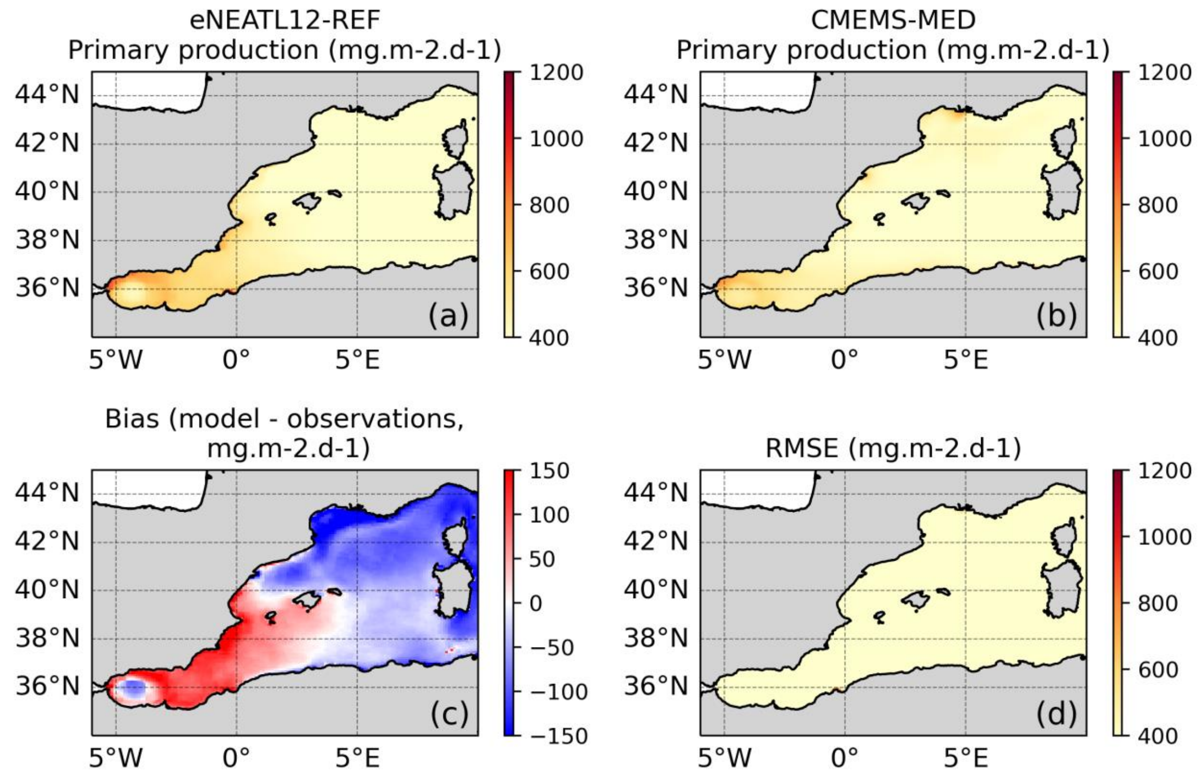


Figure 36. Depth integrated primary production (a) from the hindcast simulation and (b) satellite observations (CMEMS-MED), in the Mediterranean, for the period 1998 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period.

Carbonate system

In this section, pH, surface marine pCO_2 and air-sea CO_2 fluxes ae validated by using the MOB-TAC product as a reference. In the MOB-TAC product, Atlantic values are extrapolated to the Mediterranean Sea then not reflecting the values encounter in this area. So, for this type of comparison, we only represent the Atlantic area. For pH a comparison was also done for the whole water column by using IBIRYSv3 as reference. We also validated alkalinity and DIC (surface and whole water column) using the same products and added the figures in appendix A.

Data show that the lowest surface pH values are observed in the southern part of the Atlantic region below 40°N, while the highest values are found in the northern Atlantic. Model and observations show similar distributions, but the model presents a systematic strong negative bias in all the area (Figure 37).

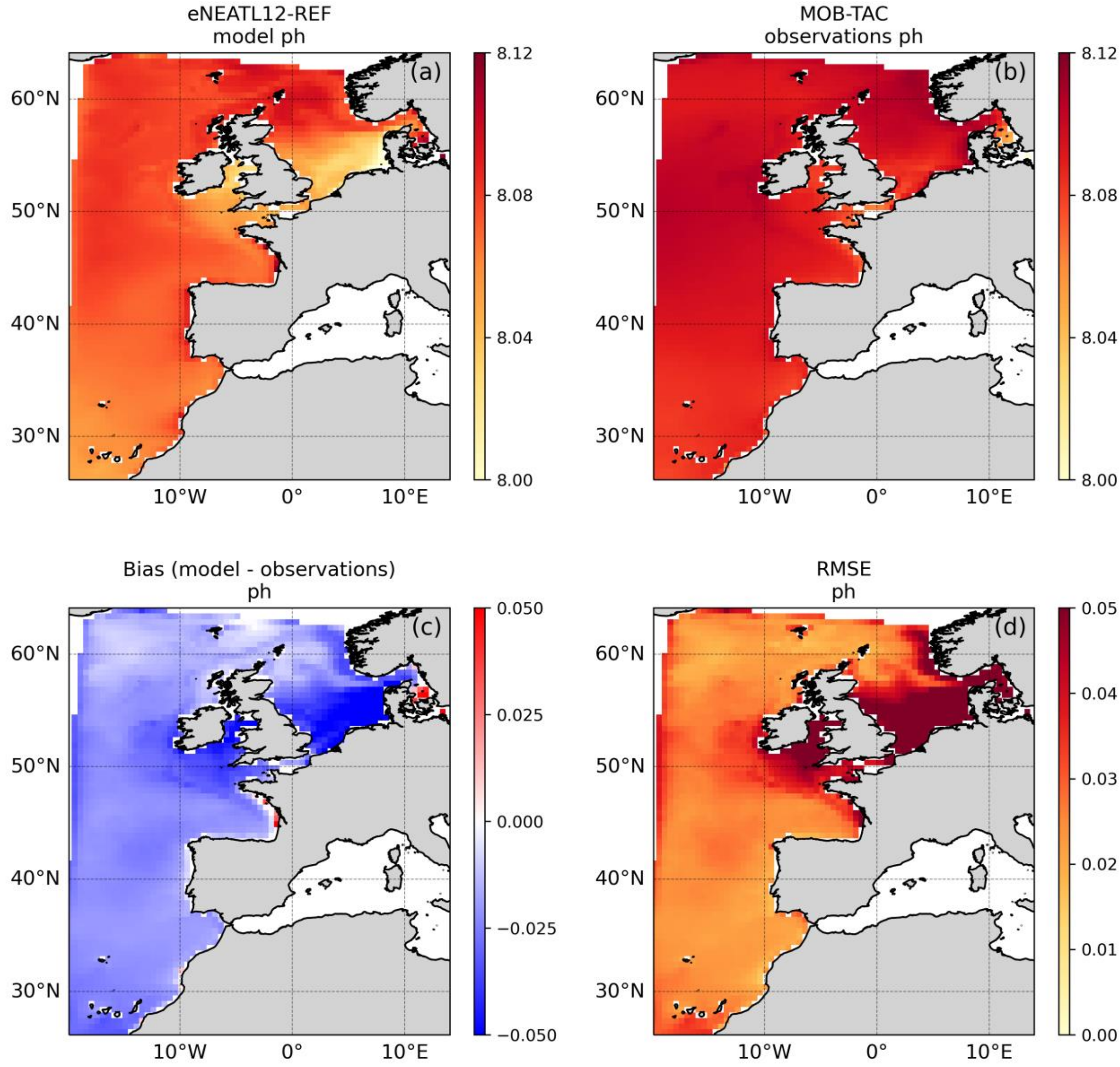


Figure 37. Surface pH (Atlantic area) (a) from the hindcast simulation and (b) MOB-TAC product, for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (model – observations) and (d) the mean RMSE for the same period.

Domain averaged pH of the whole water column is compared to the reanalysis IBIRYSv3 (Figure 38), by calculating the bias and the mean vertical profile. The highest pH values are obtained in the surface layer. Both profiles show a decrease of the pH with the depth. Low positive biases are found for the whole water column up to 2012 in the 0-1000m layer (Figure 38c, f). Then the biases are positive until 2024. The reanalysis mean vertical profile exhibits two abrupt decreases which are not visible on the hindcast mean vertical profile (at 700 m and 1265 m).

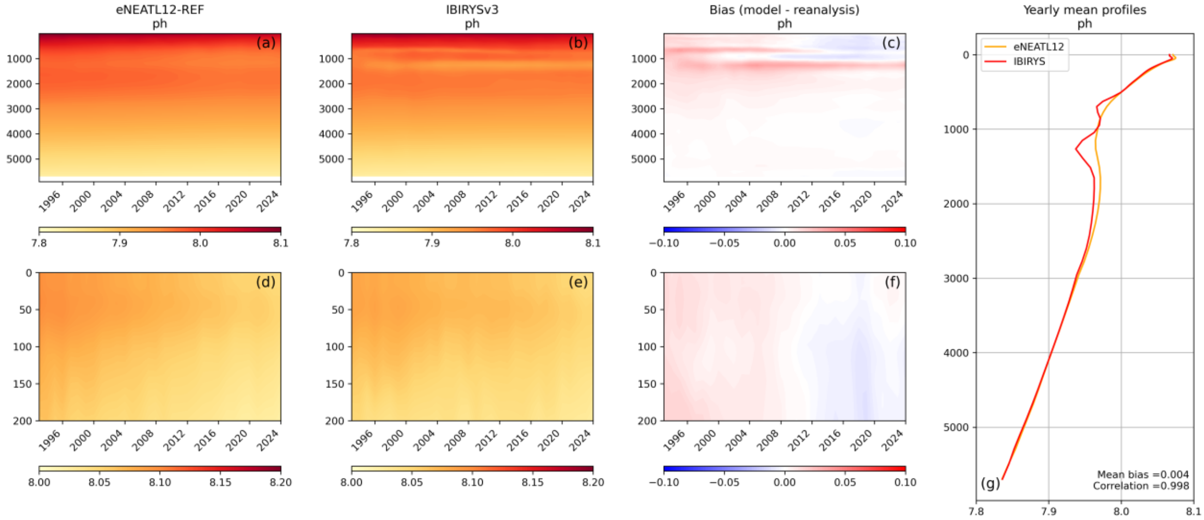


Figure 38. Hovmoller diagrams of mean pH (domain average). Panels a, b and c show the results for the whole water column while d, e and f focus on the 0-200m layer. (a, d) are for the hindcast simulation and (b, e) are for the reanalysis (IBIRYSv3), on the period 1993 – 2024. (c, f) Both are compared by calculating the mean bias (model – reanalysis) on the same period. (g) Mean vertical profiles (domain and period average) of the hindcast and reanalysis pH.

The highest pCO_2 values are observed in the southern Atlantic south of 40°N, and the lowest values are found in the northern Atlantic (Figure 39). The model and data show the same distribution, but the model present negative bias in the northeastern Atlantic (hindcast pCO_2 lower than the observations one) and a positive bias in the northwestern and southern Atlantic (hindcast pCO_2 higher than the observations one).

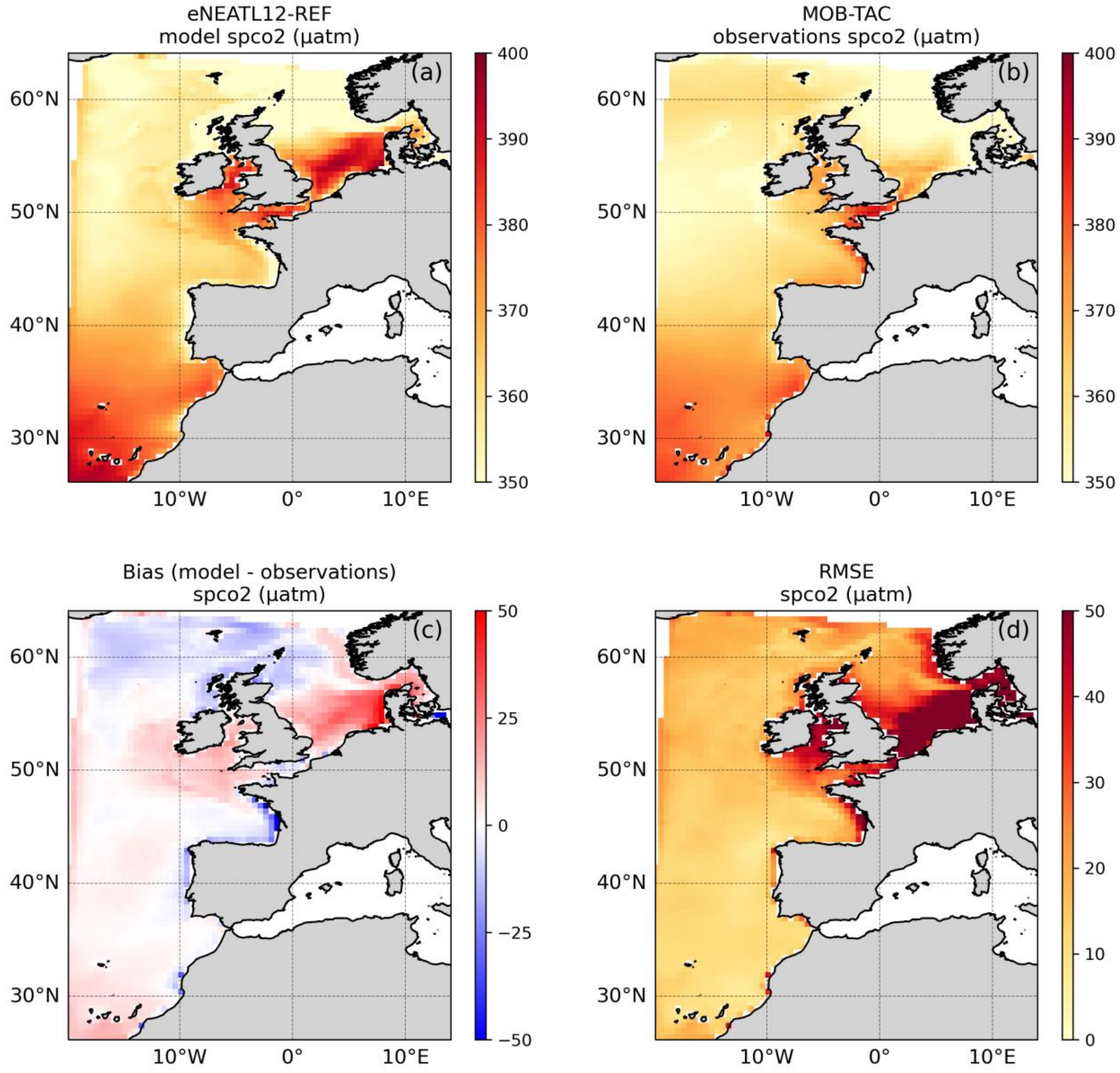


Figure 39. Surface marine pCO₂ (µatm, Atlantic area) (a) from the hindcast simulation and (b) MOB-TAC product, for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias (µatm, model – observations) and (d) the mean RMSE for the same period.

The MOB-TAC product shows only positive air-sea fluxes values while the hindcast shows negative values in the southern Atlantic (Figure 40). Biases are similar to pCO_2 biases when looking at the area but inverse when looking at the sign: the model present positive bias in the northeastern Atlantic and a negative bias in the northwestern and southern Atlantic.

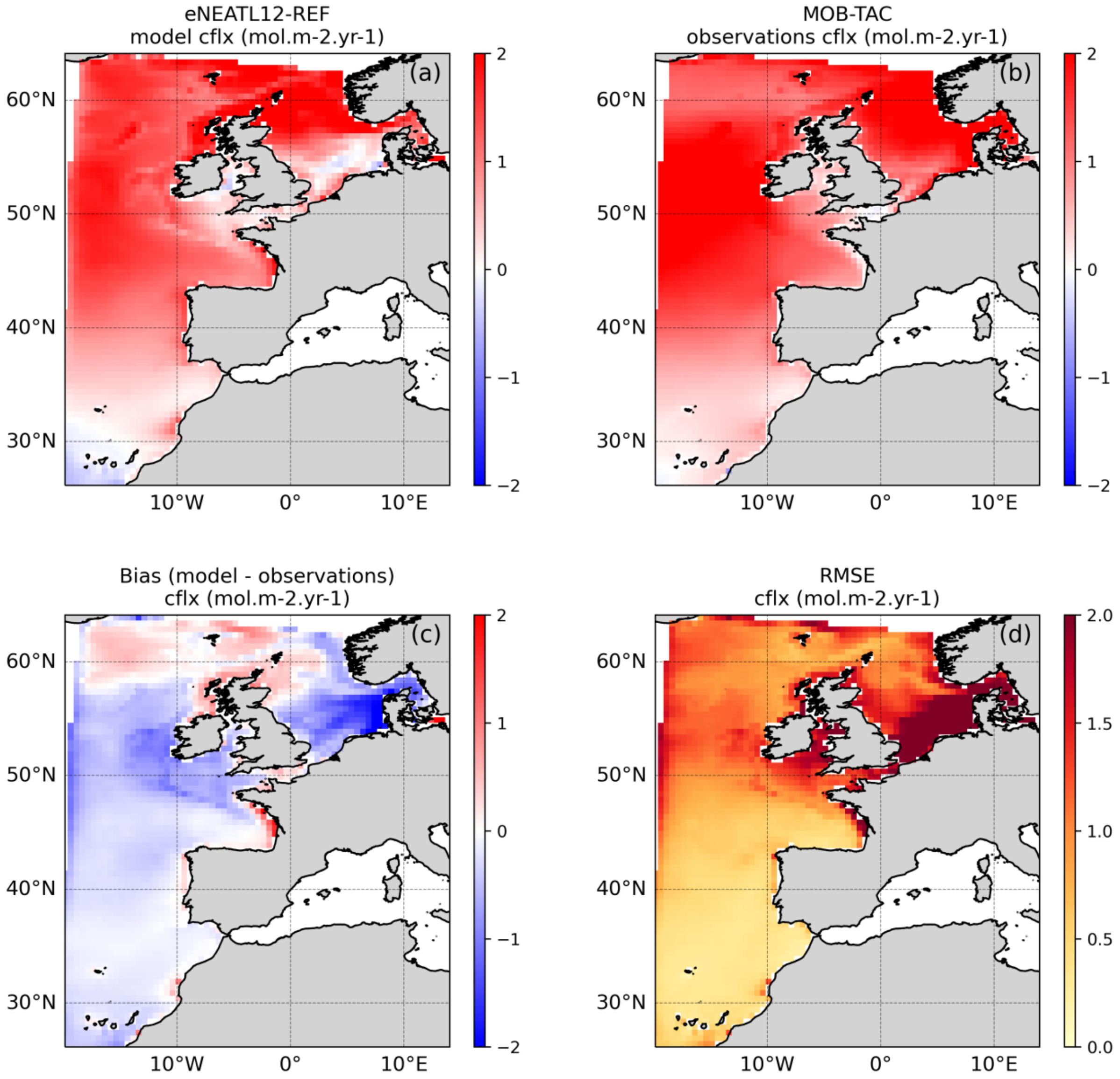


Figure 40. Air-sea CO_2 fluxes ($mol.m^{-2}.yr^{-1}$, Atlantic area) (a) from the hindcast simulation and (b) MOB-TAC product, for the period 1993 – 2024 (averaged). Both are compared by calculating (c) the mean bias ($mol.m^{-2}.yr^{-1}$, model – observations) and (d) the mean RMSE for the same period.

Box 1: Validation dataset

This box summarizes the main characteristics of the datasets used for model validation.

1/ Multi observation global Ocean ARMOR3D L4 reanalysis
Product ID: MULTIOBS_GLO_PHY_TSUV_3D_MYNRT_015_012
Spatial extent: Global Ocean - Lat -82.19° to 89.94° - Lon -179.94° to 179.94°
Spatial resolution: 0.125° × 0.125°
Period: Jan 1993 to Jan 2026
Available frequencies: Daily and monthly
Used for: Temperature, salinity and mixed layer thickness validation
See:

https://data.marine.copernicus.eu/product/MULTIOBS_GLO_PHY_TSUV_3D_MYNRT_015_012/description ; for details.

2/ Global Ocean- In-Situ Near-Real-Time Observations
Product ID: INSITU_GLO_PHYBGCWAV_DISCRETE_MYNRT_013_030
Source: In situ observations
Spatial extent: Global Ocean: Lat -90° to 90°Lon -180° to 180°
Spatial resolution: /
Period: Depending on the chosen site
Available frequencies: Several-time-daily – *daily, several times per day*
Used for: Surface temperature validation
See:
https://data.marine.copernicus.eu/product/INSITU_GLO_PHYBGCWAV_DISCRETE_MYNRT_013_030/description ; for details.

3/ Atlantic Ocean Colour (Copernicus-GlobColour), Bio-Geo-Chemical, L4 (daily interpolated) from Satellite Observations (1997-ongoing)
Product ID: OCEANCOLOUR_ATL_BGC_L4_MY_009_118
Source: Satellite observations
Spatial extent: Global Ocean Atlantic: North - Lat 20.01° to 65.99° - Lon -45.99° to 12.99°
Spatial resolution: 1 × 1 km
Period: Sep 1997 to Jan 2026
Available frequencies: Daily and monthly
Used for: Surface chlorophyll and PP validation
See:
https://data.marine.copernicus.eu/product/OCEANCOLOUR_ATL_BGC_L4_MY_009_118/description ; for details.

4/ Mediterranean Sea, Bio-Geo-Chemical, L4, monthly means, daily gapfree and climatology Satellite Observations (1997-ongoing)
Product ID: OCEANCOLOUR_MED_BGC_L4_MY_009_144
Source: Satellite observations
Spatial extent: Mediterranean Sea: Lat 30° to 46° - Lon -6° to 36.5°
Spatial resolution: 1 × 1 km
Period: Sep 1997 to Jan 2026
Available frequencies: Daily and monthly
Used for: Surface chlorophyll and PP validation
See:
https://data.marine.copernicus.eu/product/OCEANCOLOUR_MED_BGC_L4_MY_009_144/description ; for details.

5/ L4 Global Ocean surface carbon product
Product ID: MULTIOBS_GLO_BIO_CARBON_SURFACE_MYNRT_015_008
Source: In situ observations
Spatial extent: Global Ocean: Lat -89.87° to 89.88° - Lon -179.87° to 179.88°

Spatial resolution: 0.25° x 0.25°
Period: Jan 1985 to Dec 2025
Available frequencies: Monthly
Used for: Carbonate system variables validation
See:
https://data.marine.copernicus.eu/product/MULTIOBS_GLO_BIO_CARBON_SURFACE_MYNRT_015_008/description ; for details.

6/ European Seas Gridded L 4 Sea Surface Heights And Derived Variables Reprocessed 1993 Ongoing
Product ID: SEALEVEL_EUR_PHY_L4_MY_008_068
Source: Satellite observations
Spatial extent: Lat 19.97° to 66.03° - Lon -30.03° to 42.03°
Spatial resolution: 0.0625° x 0.0625°
Period: Jan 1993 to Aug 2025
Available frequencies: Daily, Monthly
Used for: ssh validation
See:
https://data.marine.copernicus.eu/product/SEALEVEL_EUR_PHY_L4_MY_008_068/description ; for details

7/ WOA2023
Source: In situ observations
Spatial extent: Global Ocean
Spatial resolution: 1° x 1°
Period: Climatology 1965 – 2022
Used for: Nutrients and oxygen validation.
See: https://www.ncei.noaa.gov/access/world-ocean-atlas-2023/ ; for details.

8/ IBIRYSv3 Reanalysis
Source: Numerical model
Spatial extent: Iberia-Biscay-Ireland: Lat 26.17° to 56.08°Lon -19.08° to 5.08°
Spatial resolution: 0.083° x 0.083°
Period: Jan 1993 to Jan 2025
Available frequencies: Hourly, daily, monthly and yearly (depending on variables)
Used for all physical and biogeochemical variables validation.
See: ; for details.

9/ IBI36RYS Reanalysis
Source: Numerical model
Spatial extent: Iberia-Biscay-Ireland: Lat 26.17° to 56.08°Lon -19.08° to 5.08°
Spatial resolution: 0.028° x 0.028°
Period: Jan 1993 to Jan 2025
Available frequencies: Hourly, daily, monthly and yearly (depending on variables)

Used for ssh validation.
See: https://data.marine.copernicus.eu/product/IBI_MULTIYEAR_PHY_005_002/description
; for details.

3.1.3. Waves validation

The wave hindcast simulations for the period 1993-2024 is validated against in-situ datasets and reanalysis products leveraging on the Copernicus Marine product quality assessment framework. The table below provides details on the figure list and the reference dataset used to produce these figures.

Table 5. Validated physical variables, with associated figures, metrics and reference datasets.

Variable	**Type of Figure**	**Metrics**	**Reference Dataset**
VHM0	Surface maps (mean of total period 1993-2024)	Bias	Reanalysis: IBI_MULTIYEAR_WAV_005_006 (00.027° × 0.027°) Doi: https://doi.org/10.48670/moi-00030
	Box averaged time series plot (full period from 193-2024)	Hourly data overlayed by monthly mean	
	Buoy validation at several locations in the IBI region	Correlation and RMSE	In-situ observation: INSITU_GLO_WAV_DISCRETE_MY_013_045 Doi: https://doi.org/10.17882/70345
VTM02	Surface maps (mean of total period 1993-2024)	Bias	Reanalysis: IBI_MULTIYEAR_WAV_005_006 (00.027° × 0.027°) Doi: https://doi.org/10.48670/moi-00030

	Box averaged time series plot (full period from 193-2024)	Hourly data overlayed by monthly mean	
	Buoy validation at several locations in the IBI region	Correlation and RMSE	In-situ observation: INSITU_GLO_WAV_DISCRETE_MY_013_045 Doi: https://doi.org/10.17882/70345
VTPK	Surface maps (mean of total period 1993-2024)	Bias	Reanalysis: IBI_MULTIYEAR_WAV_005_006 (00.027° × 0.027°) Doi: https://doi.org/10.48670/moi-00030
	Box averaged time series plot (full period from 193-2024)	Hourly data overlayed by monthly mean	
	Buoy validation at several locations in the IBI region	Correlation and RMSE	In-situ observation: INSITU_GLO_WAV_DISCRETE_MY_013_045 Doi: https://doi.org/10.17882/70345
VMDR	Surface maps (mean of total period 1993-2024)	Bias	Reanalysis: IBI_MULTIYEAR_WAV_005_006 (00.027° × 0.027°) Doi: https://doi.org/10.48670/moi-00030

	Box averaged time series plot (full period from 193-2024)	Hourly data overlayed by monthly mean	
	Buoy validation at several locations in the IBI region	Correlation and RMSE	In-situ observation: INSITU_GLO_WAV_DISCRETE_MY_013_045 Doi: https://doi.org/10.17882/70345

Significant Wave Height (VHM0)

The mean fields of spectral significant wave height (VHM0) of SEACLIM-IBI and IBIRYS is compared in Figure 41. The primary result is the model is in good agreement with the overall VHM0 patterns in the IBI (Figure 41a), with higher waves corresponding to the North Atlantic swells (>3m) in the northwestern IBI and a gradual decrease in wave height as we move south. The model (Figure 41a) has also captured the low wave heights in the Mediterranean Sea and the North Sea basins (<1.8m). This is conclusive with existing literature that both basins are predominantly wind-sea dominated. The bias plot (Figure 41c) shows a general underestimation of wave heights by the model in majority of the IBI with localized over predictions of VHM0 especially near complex coastal bathymetry and shadow regions of small islands. The overall bias is significantly low making the model a good and reliable choice for long-term projection studies.

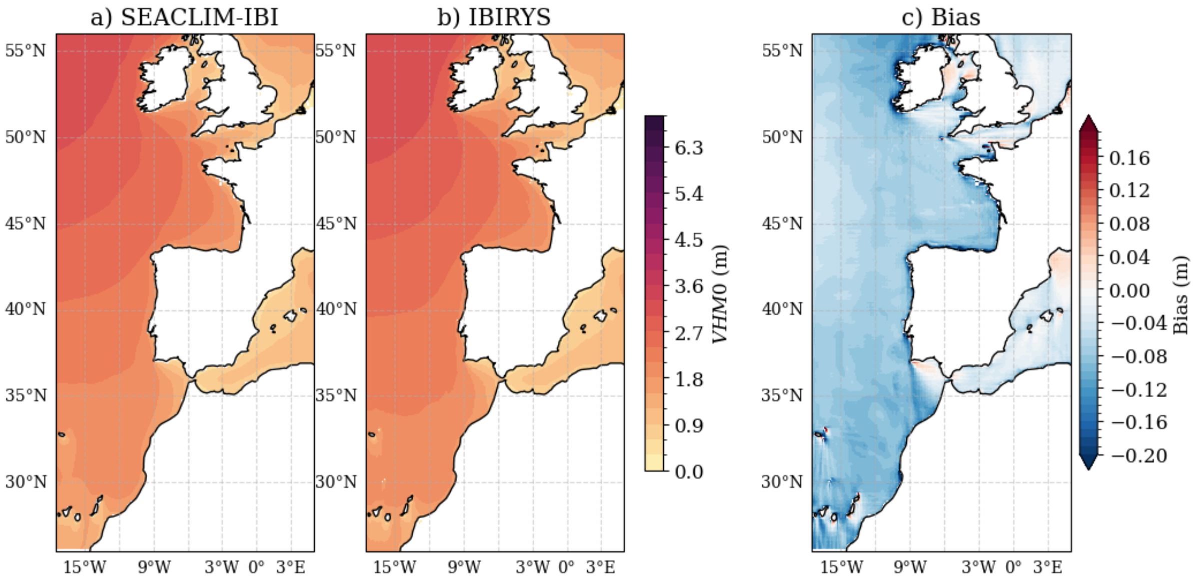


Figure 41. Climatology of Significant wave height (VHM0) for the period 1993-2024 from (a) hindcast simulations, (b) reanalysis dataset, IBIRYS and (c) the mean bias (model –observation).

The domain averaged time series of VHM0 is shown in Figure 42. The statistics is representative of the comparison between hourly SEACLIM-IBI and IBIRYS VHM0. The model has accurately captured the seasonality of VHM0 in the IBI, with winter peaks (>3m; hourly) and summer lows (<1.5m; from hourly). The model was able to track the North Atlantic storm peaks well as depicted by the high correlation and low bias.

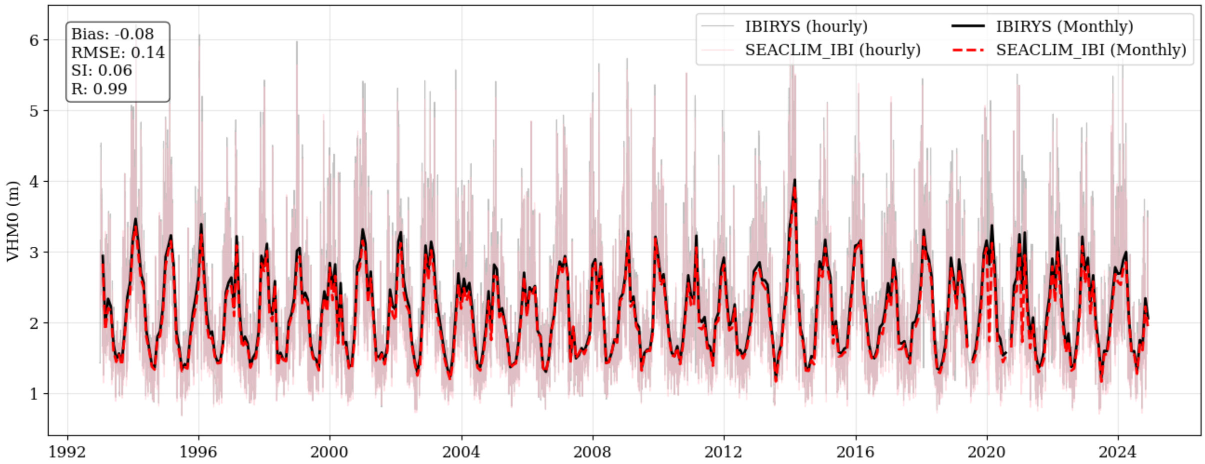


Figure 42. Domain averaged hourly Significant wave height (VHM0) for the period 1993-2024 overlayed with monthly mean values, for IBIRYS and SEACLIM-IBI. The statistics represent the comparison of hourly data.

On comparing the model VHM0 with buoy observations in the IBI (Figure 42), most of the locations show high correlation (Figure 43a; >0.8) indicating good model performance. The RMSE map (Figure 43b), shows higher values along the Brittany coast, Galicia coast and the northwestern IBI, which is attributed to the higher wave heights due to the high energy swells in these regions.

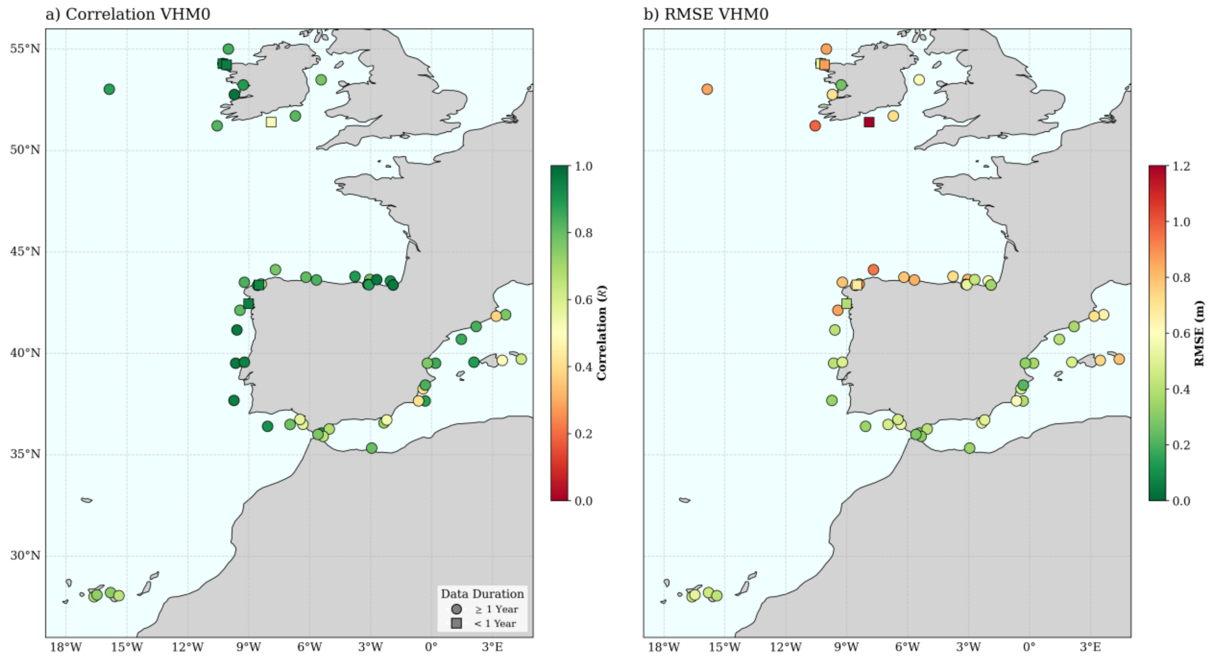


Figure 43. Correlation (a) and RMSE (b) obtained from comparison of model and buoy VHM0 at various locations in the IBI region.

Mean Wave Period (VTM02)

The averaged field of mean wave period (VTM02) for the full period (1993-2024) is compared against the IBIRYS wave period in Figure 44. The overall model performance in producing the wave period patterns is in good agreement with the IBIRYS data. Although, the model fails to capture the swell signatures in the northwestern IBI. There is a general underestimation of wave periods by the model in the open ocean and contradictory overestimations in shadow regions of islands illustrating the challenge of resolving areas of complex bathymetry as depicted in the bias plot. The short period waves in the North Sea and the Mediterranean Sea are captured well by the model as corroborated by the low biases in these basins since both SEACLIM-IBI and IBIRYS models were forced with ERA5 dataset therefore reducing the wind-driven errors. The model is reliable for the open ocean, but the coarser grid puts constraints in capturing the nuanced bathymetry and smaller islands well.

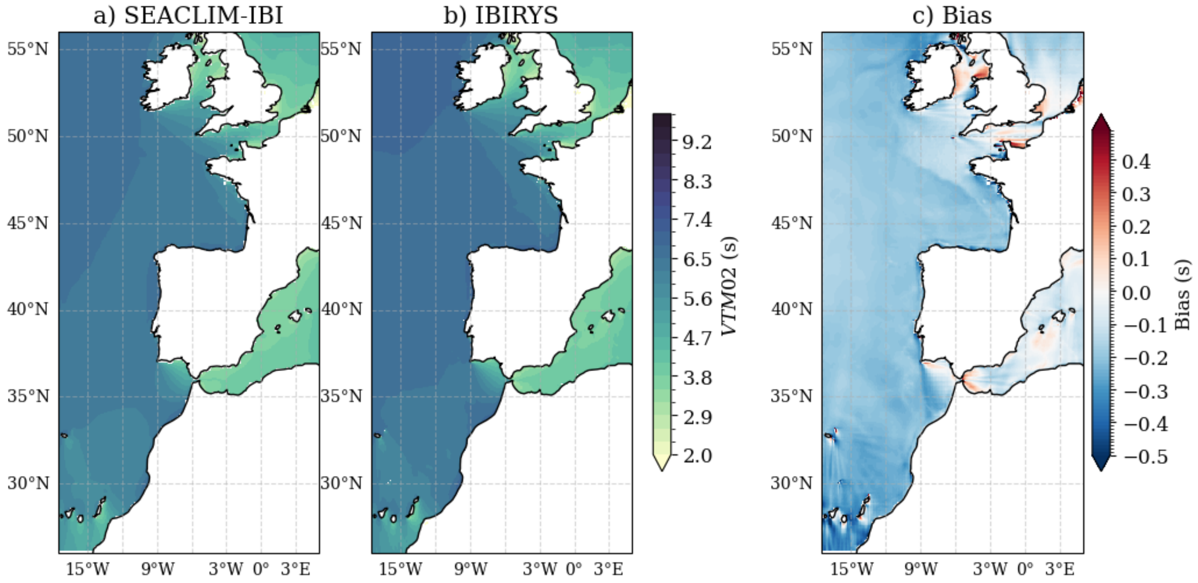


Figure 44. Climatology of mean wave period (VTM02) for the period 1993-2024 from (a) hindcast simulations, (b) reanalysis dataset, IBIRYS and (c) the mean bias (model –observation).

The domain averaged time series of VTM02 similar to that of VHM0, is depicted in Figure 45. The statistics convey good performance of the model in capturing the wave periods of the IBI. As explained earlier, there is a general underestimation (bias: -0.21s) in the wave periods as seen in the timeseries and spatial validation. While this bias is quite low, the IBIRYS product is assimilated with Sentinel-1 (SAR based wave spectra) and CFOSAT since 2016 and 2018, respectively, which introduces a better swell representation in the IBIRYS dataset. This could act unfavorably for the SEACLIM-IBI as it is a free run without assimilation and hence unable to fully capture the North Atlantic swell energy in the region, contributing to the low bias. The high correlation shows that the model performs well in capturing the seasonality of the wave patterns.

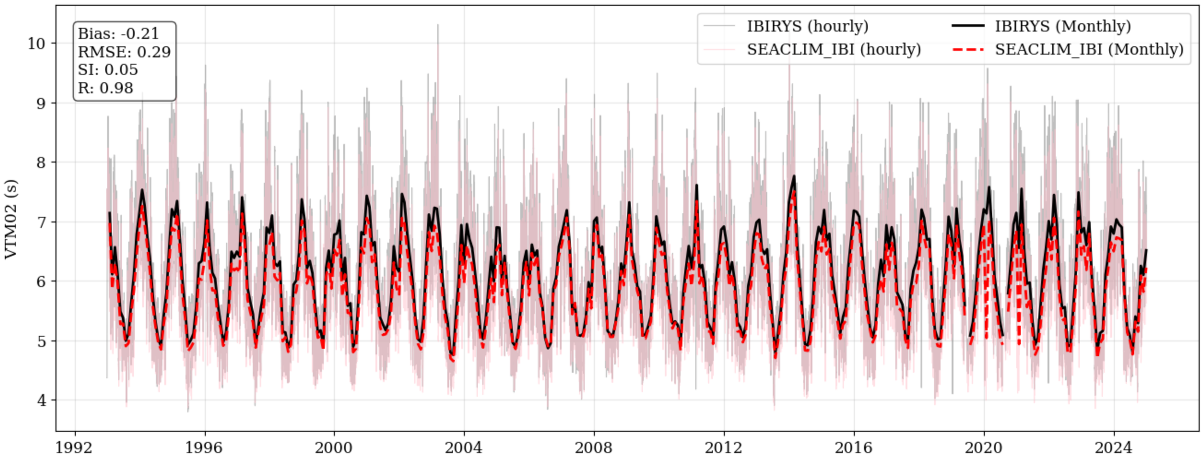


Figure 45. Domain averaged hourly Mean Wave Period (VTM02) for the period 1993-2024 overlayed with monthly mean values, of IBIRYS and SEACLIM-IBI. The statistics represent the comparison of hourly data.

The buoy validation of VTM02 in the IBI is represented in Figure 46. High correlation (>0.8) at majority of the locations is conclusive of the model's ability to capture the wave periods well. The RMSE is comparable at most of the buoy locations, with a few coastal transition zones (low correlation points as well) having high values. These discrepancies could be due to the poor resolution of coastal bathymetry by the coarser SEACLIM-IBI model grid (although it is forced with sea levels unlike IBIRYS) where as IBIRYS is a high resolution (1/36°) product. This is true especially near enclosed or sheltered coastlines and the strait of Gibraltar.

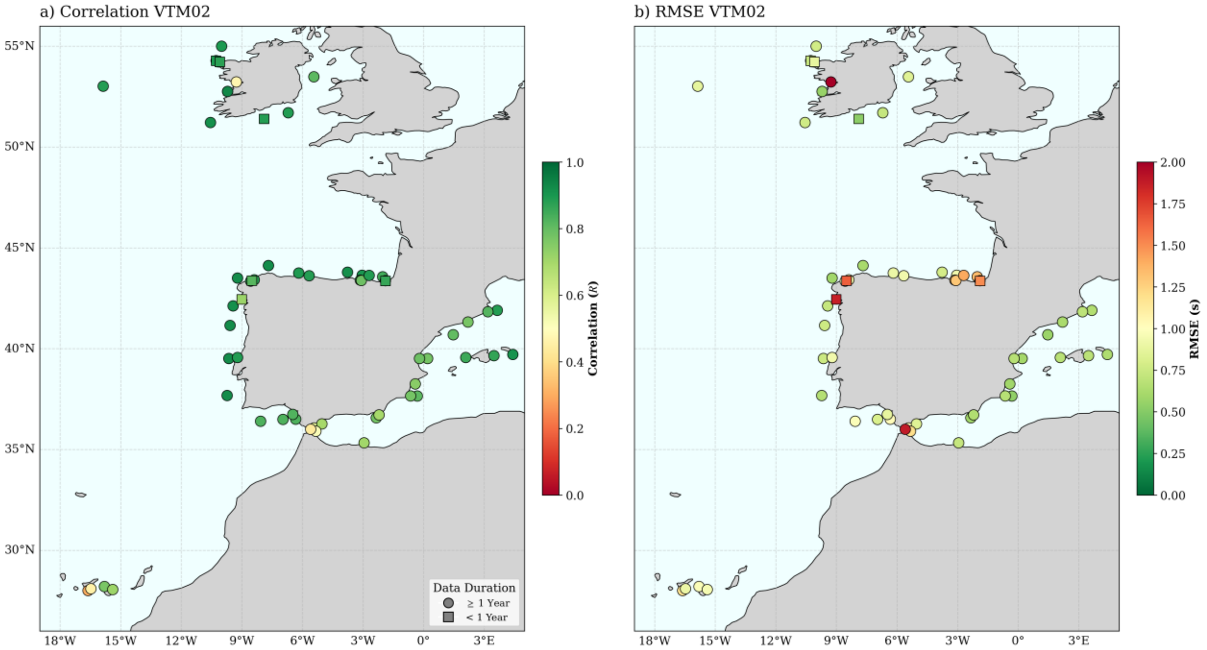


Figure 46. Correlation (a) and RMSE (b) obtained from comparison of model and buoy VTM02 at various locations in the IBI region.

Peak Period (VTPK)

The mean fields of spectral peak period (VTPK) is shown in Figure 47. The swell signatures (expected values >9s) in the North Atlantic is reproduced well by the model (Figure 47a). The shorter waves in the Mediterranean Sea and North sea (both wind-sea dominated basins) have also been captured well by the model. The bias (Figure 47c) is prominent than for VTM02, clearly showcasing the shortcoming in resolving swells. It also shows how well the data assimilation in IBIRYS has helped to capture swell fields. The larger biases also indicate while the model could capture the averaged picture reasonably, it needs more care to represent the more mature and energetic swells in the wave spectrum.

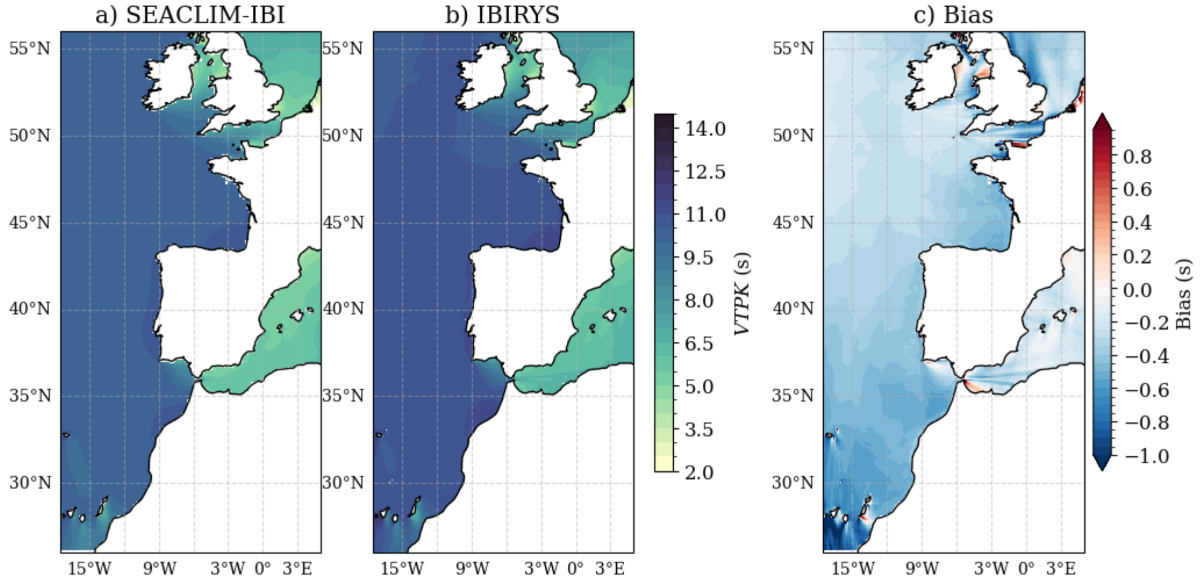


Figure 47. Climatology of mean wave period (VTPK) for the period 1993-2024 from (a) hindcast simulations, (b) reanalysis dataset, IBIRYS and (c) the mean bias (model –observation).

The domain averaged seasonality of VTPK (Figure 48) shows larger bias of -0.42s for the hourly timeseries indicating its caveat in properly resolving the long-period swells. This systematic bias is mostly due to the difference in assimilated IBIRYS and the free run SEACLIM-IBI. The correlation is high and a low scatter index depicts the model's ability to capture the signals well and stability.

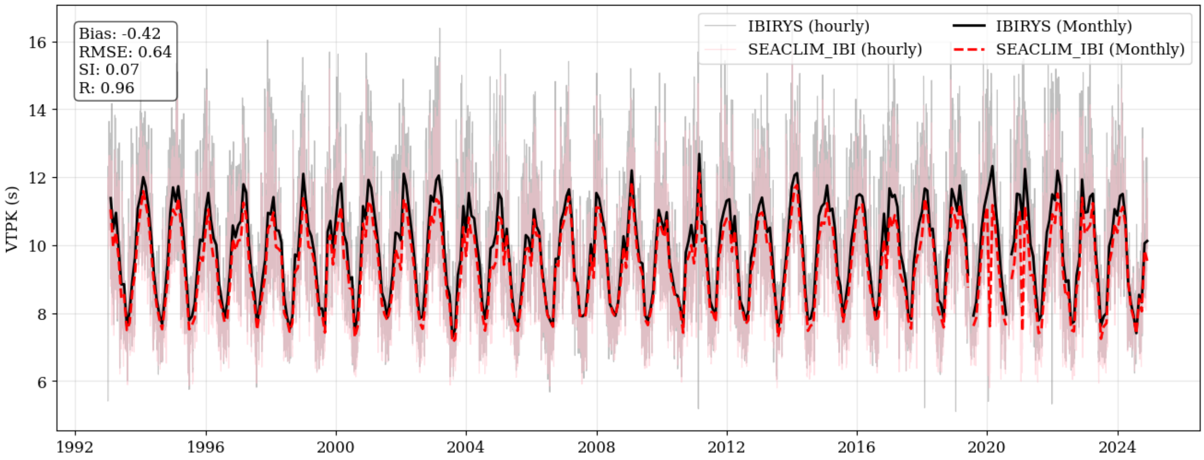


Figure 48. Domain averaged hourly Peak Period (VTPK) for the period 1993-2024 overlayed with monthly mean values, of IBIRYS and SEACLIM-IBI. The statistics represent the comparison of hourly data.

The comparison of VTPK (Figure 49) against buoy locations in the IBI shows poor statistics compared to the other parameters. The low correlation (Figure 49a) and high RMSE (Figure 49b) could only be explained by the difference in computation of both datasets, for example, the temporal window or the bin size etc. A slight shift in the spectral peak can offset the results and thereby give poor skill scores in comparison to the earlier explained continuous mean parameters like VHM0 or VTM02. This can also be due to the previously discussed shortcoming of the model in capturing the fully matured spectrum.

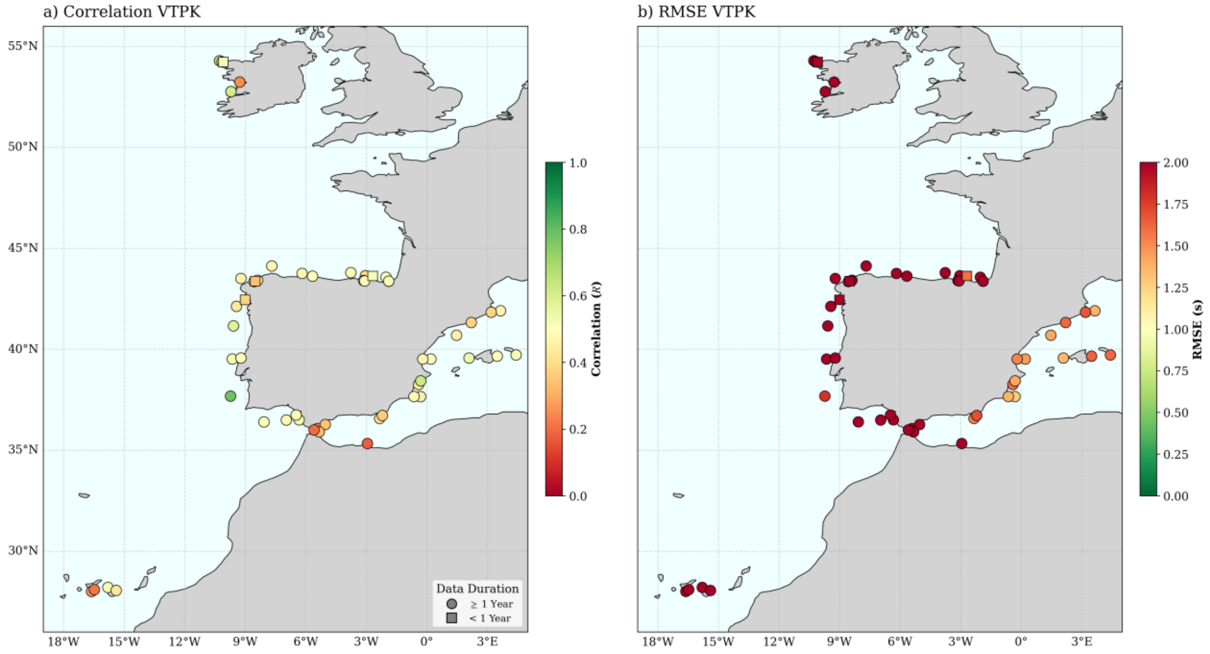


Figure 49. Correlation (a) and RMSE (b) obtained from comparison of model and buoy VTPK at various locations in the IBI region.

Mean Wave Direction (VMDR)

Figure 50 shows the comparison of mean wave direction (VMDR) a several buoy locations in the IBI. The correlation is high (>0.8) at most of the locations where the model is in good agreement with buoy wave direction. The RMSE values are also low (<20 °) at most locations with better model performance in regions prone to the North Atlantic swells. The few locations with low correlation and high RMSE (> 35 °) can be attributed to the enclosed nature of these locations. These locations may be subject to local diffractive and refractive processes which the coarser model was unable to capture.

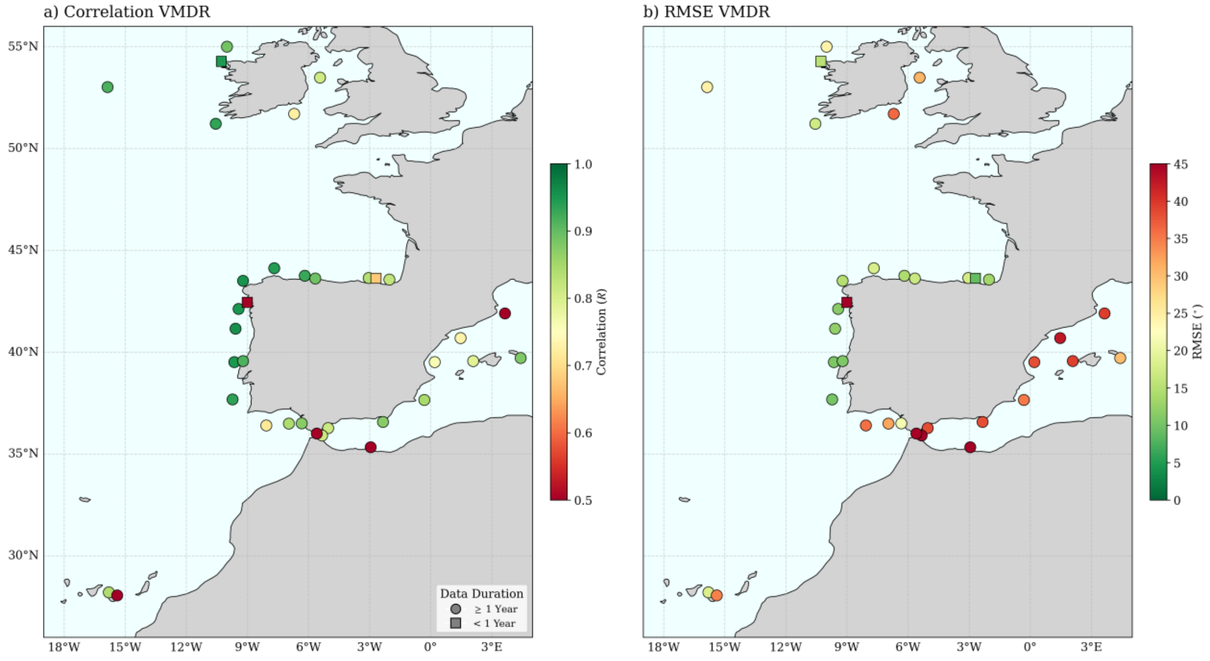


Figure 50. Correlation (a) and RMSE (b) obtained from comparison of model and buoy VMDR at various locations in the IBI region.

Box 2: Validation datasets

A detailed summary of the datasets used for model validation is provided below.

1. Atlantic -Iberian Biscay Irish- Ocean Wave Reanalysis
Product ID: IBI_MULTIYEAR_WAV_005_006
Spatial Extent: Atlantic: Iberia-Biscay-Ireland; Atlantic: North; Lat 26° to 56° Lon -19° to 5°
Spatial Resolution: 0.027° × 0.027°
Period: 1 Jan 1980 to 8 Oct 2025
Available frequency: Hourly; Multi-yearly
Variables used: VHM0, VTM02, VTPK and VMDR
Reference:
https://data.marine.copernicus.eu/product/IBI_MULTIYEAR_WAV_005_006/description

2. Global Ocean - Delayed Mode Wave product
Product ID: INSITU_GLO_WAV_DISCRETE_MY_013_045
Spatial Extent: Atlantic: Global Ocean; Lat -90° to 74.62° Lon -180° to 180°
Spatial Resolution: point data
Period: 27 Apr 1970 to 1 Jan 2025
Available frequency: Hourly
Variables used: VHM0, VTM02, VTPK and VMDR
Reference:
https://data.marine.copernicus.eu/product/INSITU_GLO_WAV_DISCRETE_MY_013_045/description

3.2. Arctic

The hindcast simulations from TOPAZ2 are evaluated against the GLORYS12v1 that is also used as boundary conditions for the physical variables. The hindcast is also compared to ARMOR3D for selected variables and the gridded climatology of MLD based on the method in de Boyer Montégut et al. (2004). The biogeochemical variables were compared to climatologies from WOA2023 and satellite fields of chlorophyll. In addition, nutrients and chlorophyll were compared to a database of in-situ data that cover the Norwegian and Barents Sea. This dataset is regularly used to evaluate the biogeochemical operational forecast model. We focus the validation on the large-scale patterns and trends. We do see some noise an gradient in the boundary regions of the model and for that reason it is advisable to use the model results only away from the lateral boundary regions, at least 20-grid cells.

3.2.1. Physical validation

For the physical validation, we focus on hydrography, current and mixed-layer depth. We have also computed transports though key regions and compared them to estimates from literature.

The temporal evolution of the mean fields of temperature and salinity (Figure 51 and Figure 52) show that TOPAZ has a warm bias compared to GLORYS12v1 below 300 meters and a cold bias in the surface layers. There is more temporal variability at depth in TOPAZ2 and some instances of deep convection that bring warmer waters into the deeper layers, this is absent in GLORYS12v1. The variability in the temperature is reflected in the heat content of TOPAZ2 which is both higher and more variable in TOPAZ2 than GLORYS12v1. There is a drift in the surface salinity. On explanation can be that reference simulation does not use restoration to surface salinity climatology in open waters, which is routinely being done in the operational model version. There seems to be a gradual diffusion of saline water towards the surface in the Arctic, i.e. the model is not able to maintain the thin, fresh surface layer and this leads to a large surface salinity bias in the Arctic (Figure 53). For temperature it is generally slightly colder in the Arctic region with a large warm bias compared to GLORYS12v1 in the vicinity of the southern tip of Greenland (Figure 53).

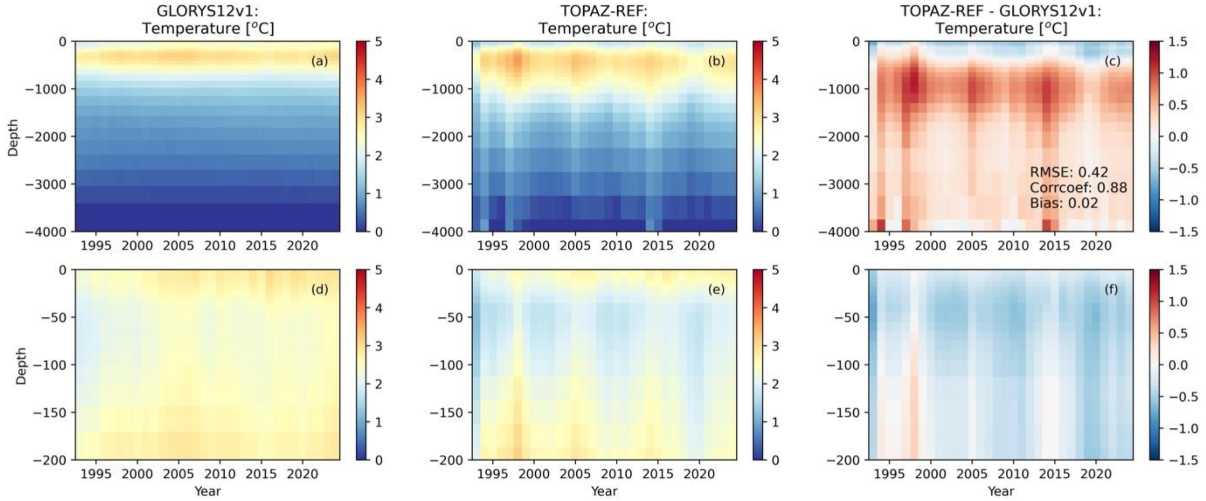


Figure 51. Hovmuller diagram of temperature from GLORYS12v1 and TOPAZ2, the temperature has been averaged over the entire domain at each model depth.

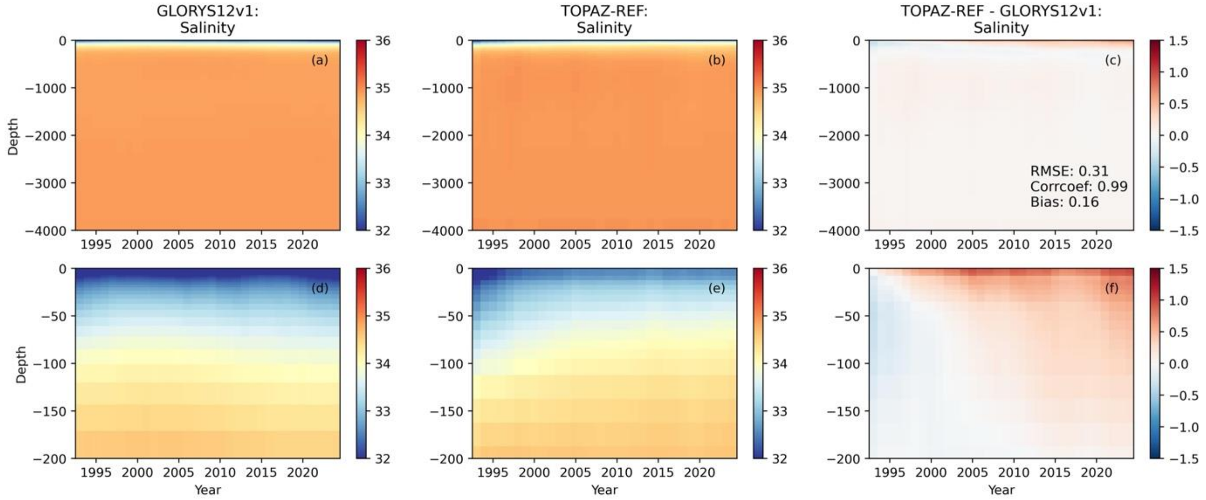


Figure 52. Hovmuller diagram of salinity from GLORYS12v1 and TOPAZ2, the temperature has been averaged over the entire domain at each model depth.

The mixed layer is generally deeper in the Greenland and Labrador Sea and west of Svalbard than the observed values from ARMOR3D (Figure 55). The mixed-layer is generally much deeper than observations during the winter season and shallower during the summer season when compared to the climatology (Figure 56 and Figure 57), this is consistent with previous studies (Ali et al., 2019). HYCOM, being isopycnal, can sometimes produce quite deep mixed-layer depths of more than 1000 meters in the Greenland Sea, this may be the reason why we see intrusions of warmer water to deeper layers in the model. It should be mentioned that observations may be biased to summer and fair-weather conditions and therefore may be underestimating the mixed-layer depth.

The patterns of currents are broadly similar in TOPAZ2 and GLORYS12v1, though GLORYS12v1, being higher resolution, have currents that are narrower than the ones in TOPAZ2 (Figure 58). Some noise is visible in the boundary regions, this indicates that the model results should primarily be used away from the nested boundary zones, which are 20 model grid cells wide. The flow though the Bering Strait appears stronger in GLORYS12v1, this is confirmed by the mean transport in TOPAZ2 that is about half the estimated 0.8 Sv (Table 6). The transports though the Barents Sea Opening and the Svinøy section have reasonable estimates compared to estimates from literature, while the transport through the Fram Strait is much lower than the observational estimates. This current is also visibly weaker in TOPAZ2 than GLORYS12v1 (Figure 58). Both the weak transport in the Bering Strait and Fram Strait may be issues of resolution. The TOPAZ-model with twice the resolution of TOPAZ2, has more realistic transport estimated though both of these straits.

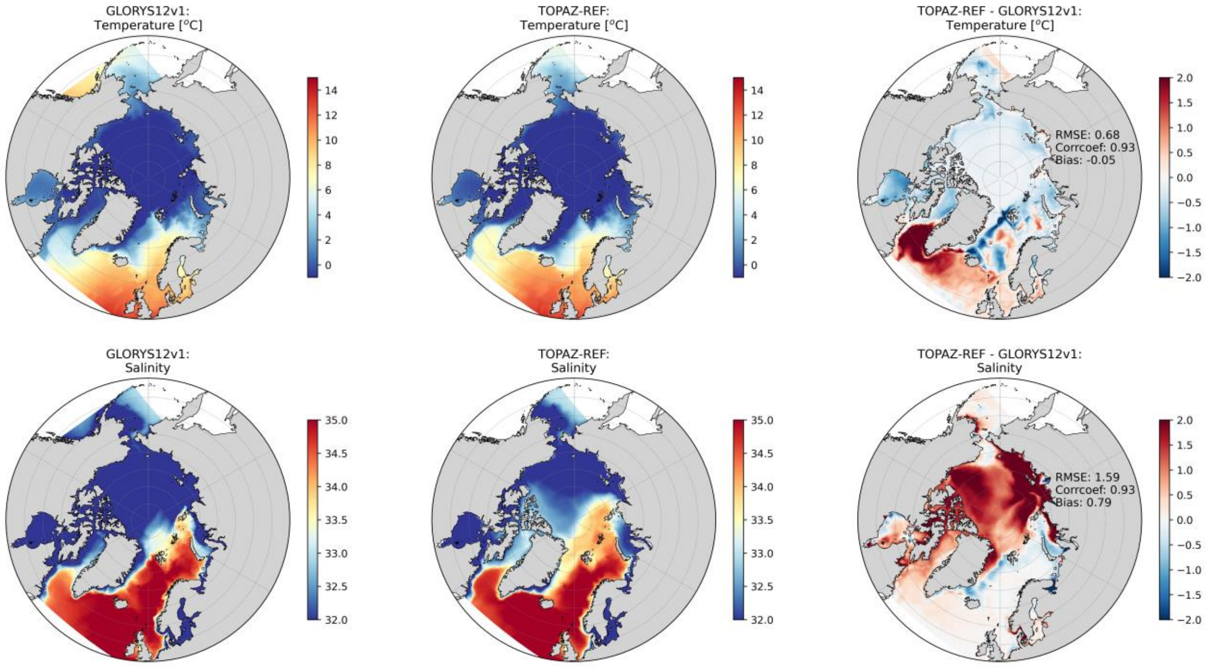


Figure 53. Average surface temperature (upper) and salinity (lower) from 1993 to 2024 from GLORYS12v1 (left), TOPAZ2 (middle) and the bias (right).

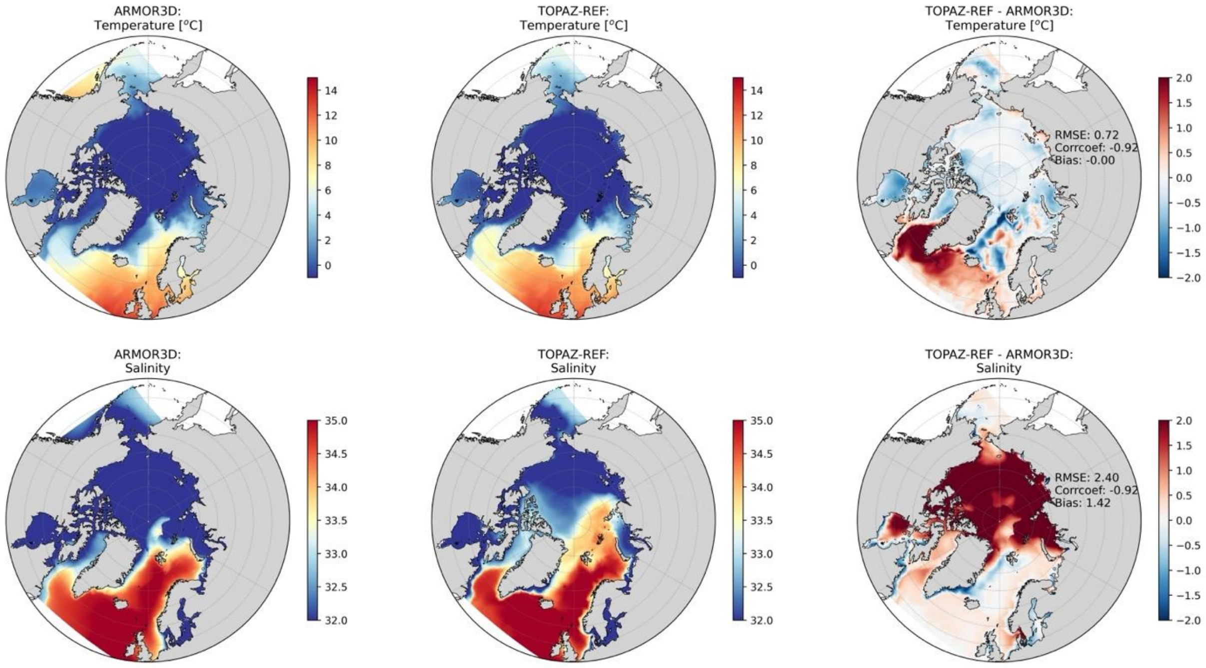


Figure 54. Average surface temperature (upper) and salinity (lower) from 1993 to 2024 from ARMOR3D (left), TOPAZ2 (middle) and the bias (right).

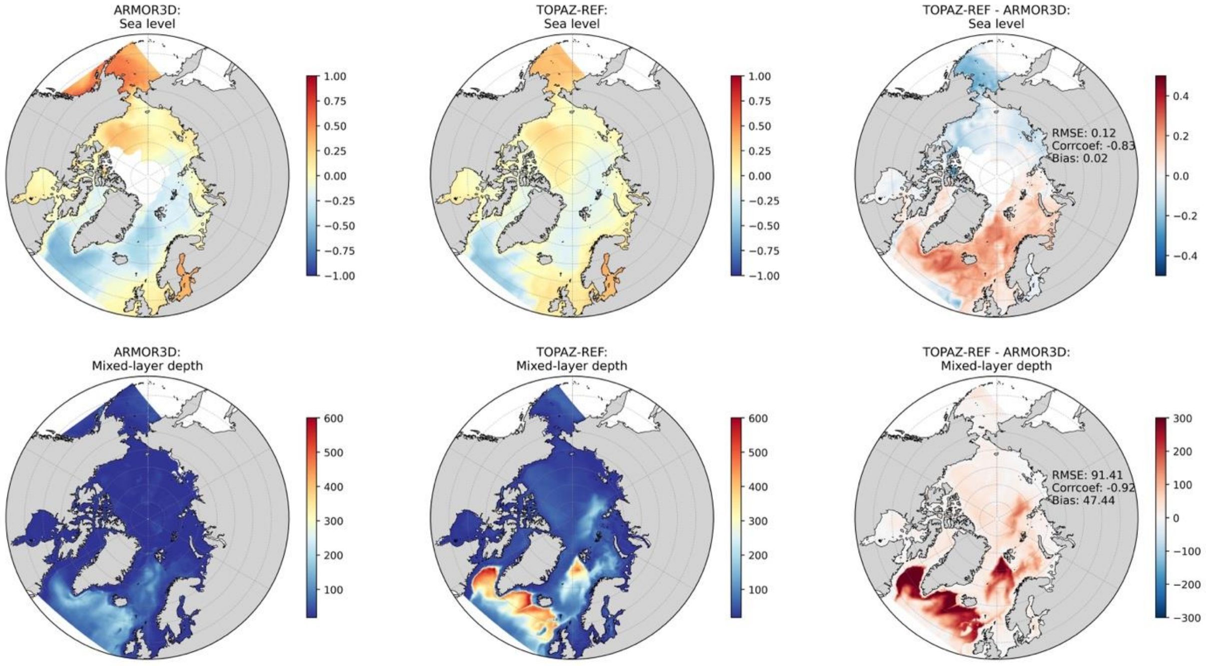


Figure 55. Average sea-level anomaly (upper) and mixed-layer depth (lower) from 1993 to 2024 from ARMOR3D (left), TOPAZ2 (middle) and the bias (right).

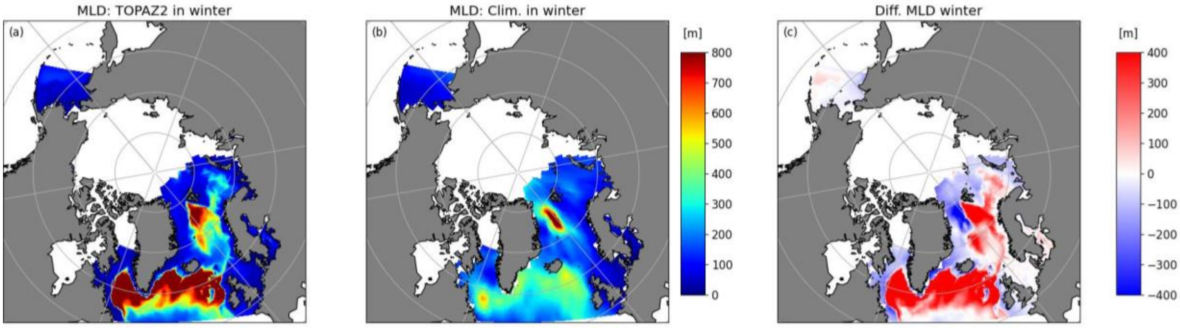


Figure 56. Average 1993-2024 MLD from TOPAZ2 (left) and from the climatology (middle) and the bias (right) in winter (January to March).

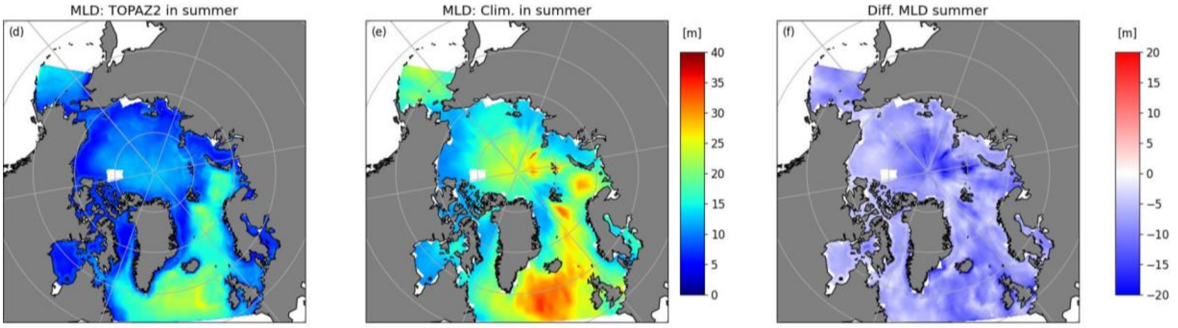


Figure 57. Average 1993-2024 MLD from TOPAZ2 (left) and from the climatology (middle) and the bias (right) in summer (July to September).

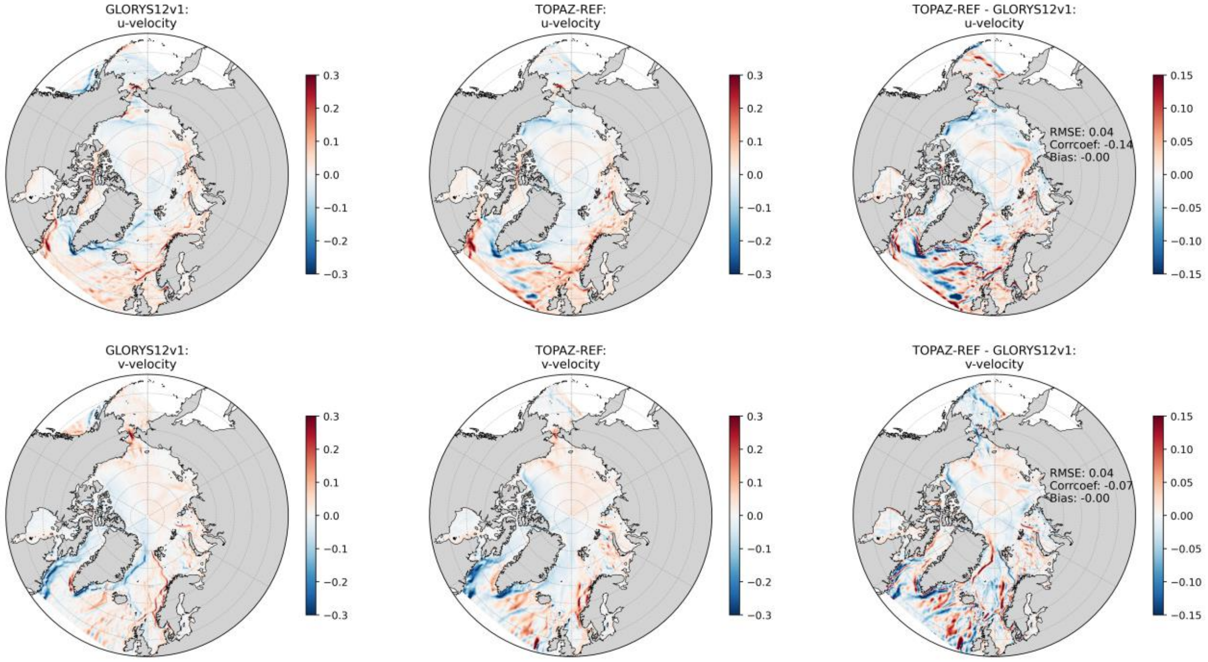


Figure 58. Average surface eastward velocity (upper) and westward velocity (lower) from 1993 to 2024 from GLORYS12v1 (left), TOPAZ2 (middle) and the bias (right).

Table 6. Estimated net volume transports in Sv of specific water masses through key sections, average for the year 1993 to 2024 compared to numbers from literature.

	TOPAZ5	Literature

Bering strait (All waters)	0.43 northward	0.8 (Woodgate et al., 2006) 0.7 (Smedsrud et al., 2022)
Fram strait (AW, T > 2˚C)	0.59 northward	3 ± 0.2 (Beszczynska Möller et al., 2012)
BSO (AW, T > 3˚C; S > 35.0)	1.8 northward	2.0 ± 1.0 (Lien et al., 2016)
Gimsøy (AW, T > 5˚C; S > 35.0)	1.25 northward	-
Svinøy (AW, T > 5˚C; S > 35.0)	4.8 northward	4.4 ± 1.0 (Lien et al., 2016)
FSC (AW, T > 5˚C; S > 35.0)	2.8 northward	2.7 ± 1.0 (Lien et al., 2016)

3.2.2. Biogeochemical validation

3.2.2.1 Nutrients and dissolved oxygen concentration

The surface nutrient concentrations (NO_3, PO_4, and Si) are evaluated against the WOA2023 climatology. The climatological fields are compared with the full hindcast period (1993–2024) to assess overall model performance.

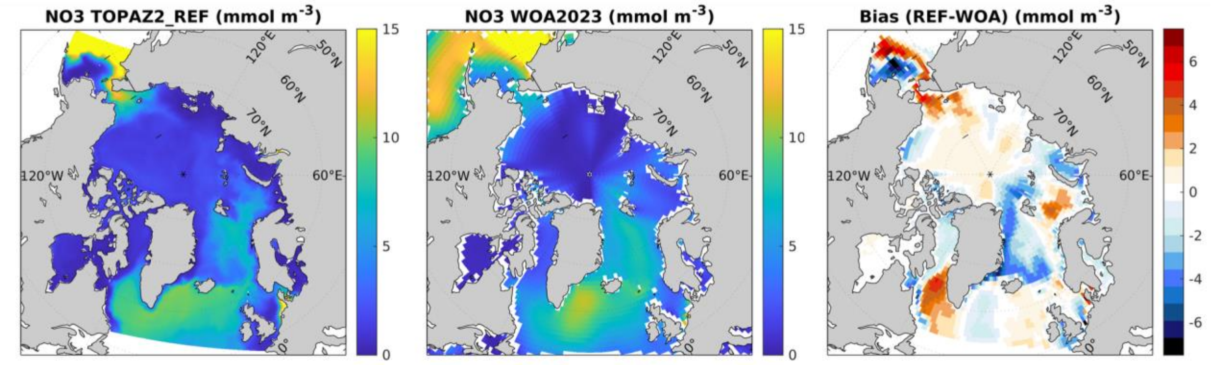


Figure 59. Surface nitrate concentration from the hindcast simulation (left), from observations (WOA2023) (middle) for the period 1993 – 2024 (averaged) and the mean bias (model – observations) for the same period (right).

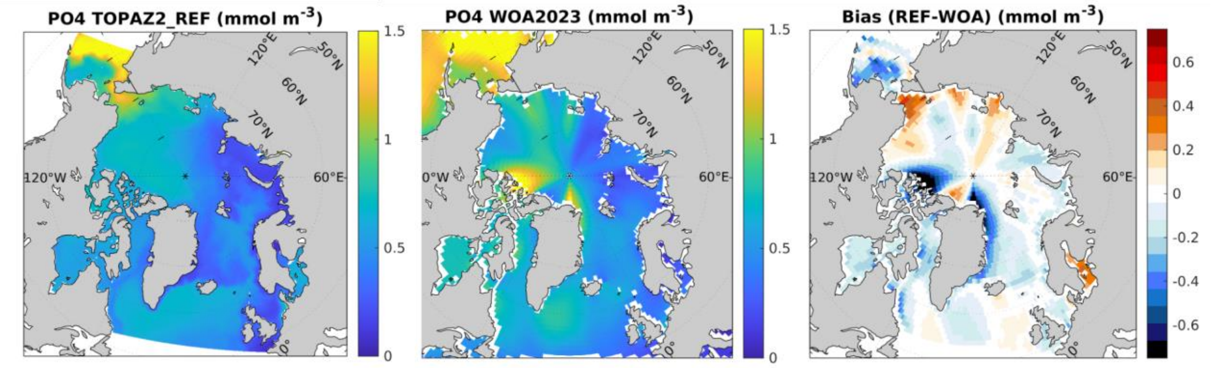


Figure 60. Surface phosphate concentration from the hindcast simulation (left), from observations (WOA2023) (middle) for the period 1993 – 2024 (averaged) and the mean bias (model – observations) for the same period (right).

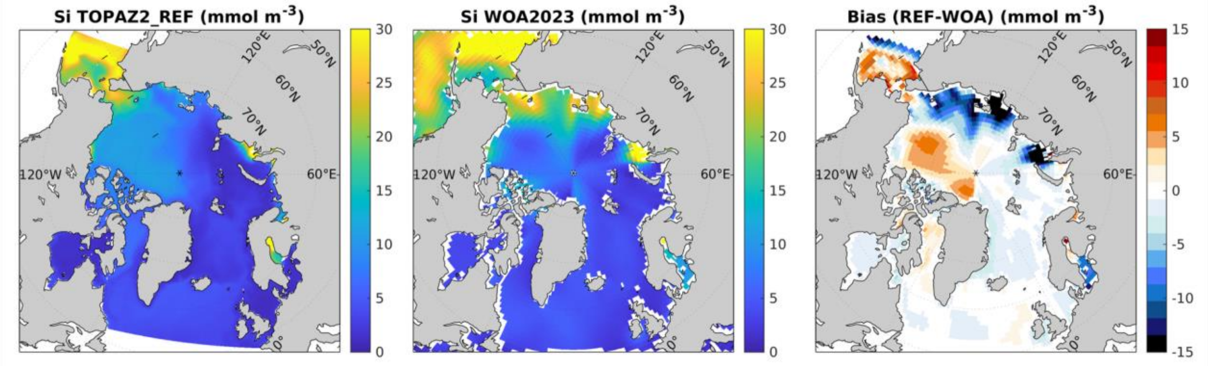


Figure 61. Surface silicate concentration from the hindcast simulation (left), from observations (WOA2023) (middle) for the period 1993 – 2024 (averaged) and the mean bias (model – observations) for the same period (right).

The model reproduces the large-scale spatial distribution of surface nutrients (NO_3, PO_4, and Si) with good agreement when compared to the WOA2023 climatology. As expected, nutrient concentrations are lowest in the central Arctic Ocean, moderately elevated in the North Atlantic sector, and highest in the Pacific-influenced regions of the Arctic, particularly south of the Bering Strait. These patterns reflect the dominant circulation and biogeochemical regimes of the Arctic and surrounding basins.

Despite the overall agreement, several regional discrepancies are evident:

- East Greenland shelf: The model underestimates both nitrate and phosphate concentrations along the eastern Greenland shelf. This may indicate insufficient vertical mixing or underestimation of nutrient-rich Atlantic inflow in this region.
- Bering Strait transition zone: North of the Bering Strait, the model simulates elevated nitrate and phosphate concentrations relative to WOA2023, while south of the strait, both nutrients are underestimated. This sharp gradient suggests potential biases in the representation of Pacific inflow and shelf–basin exchange processes.
- Silicon distribution: The silicon fields show concentrated deviations in the Pacific-influenced sector. The model tends to underestimate Si concentrations along the Siberian shelves, while overestimating them in the Beaufort Sea and Lincoln Sea. These biases may stem from differences in riverine input, and/or diatom uptake dynamics under ice.

Overall, the model captures the dominant nutrient gradients across the Arctic Ocean but exhibits regional mismatches that highlight areas for further refinement. These include shelf dynamics, boundary inflows, and biogeochemical parameterizations. Addressing these discrepancies is essential for improving predictive skill in Arctic ecosystem modeling and for understanding nutrient-driven productivity shifts under ongoing climate change.

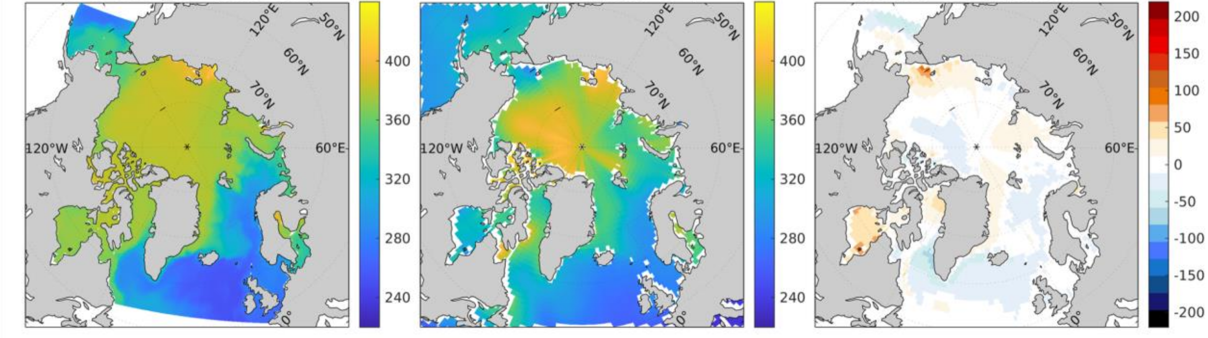


Figure 62. Surface oxygen concentration from the hindcast simulation (left), from observations (WOA2023) (middle) for the period 1993 – 2024 (averaged) and the mean bias (model – observations) for the same period (right).

The model captures the broad spatial distribution of surface oxygen concentrations across the Arctic Ocean, with values generally ranging from ~200 to over 400 mmol m^{-3}. Consistent with climatological data, oxygen levels are highest in the cold central Arctic, and lower in regions influenced by relatively warm North Atlantic Current.

Surface oxygen concentrations are largely driven by saturation dynamics, which are directly related to temperature. Because oxygen solubility decreases with increasing temperature, surface oxygen typically exhibits a negative correlation with sea surface temperature. As demonstrated in the previous section, the model overestimates surface temperature in the North Atlantic sector. Consequently, it is not surprising that comparison with the WOA2023 climatology reveals an underestimation of oxygen concentrations in this region.

A synthesis of in situ hydrographic and biogeochemical data is created for model evaluation. The sources of the synthesis dataset are GLODAPv2.2023 (Lauvset et al., 2024 https://doi.org/10.5194/essd-16-2047-2024), the hydrochemistry dataset from the ICES Marine Data Centre (ICES Data Portal, Dataset, 2025), Arctic Ocean bottle-sampled hydrochemistry data from the CCHDO Hydrographic Data Office (2023 Delayed Mode Biogeochemical product (https://doi.org/10.17882/86207). The https://doi.org/10.6075/J0CCHAM8), and the Copernicus Marine Service Global Ocean – parameters included in the synthesis hydrochemistry dataset are temperature, salinity, dissolved oxygen, nitrate, silicate, phosphate, and chlorophyll-a concentrations. Data duplication among the four source datasets was removed based on the priority ranking table, 1. GLODAPv2, 2. 3. Copernicus Marine Service, ICES, 4. CCHDO, reflecting the relative level of quality control applied to each product. The dataset has a temporal coverage from 1997 to 2021.

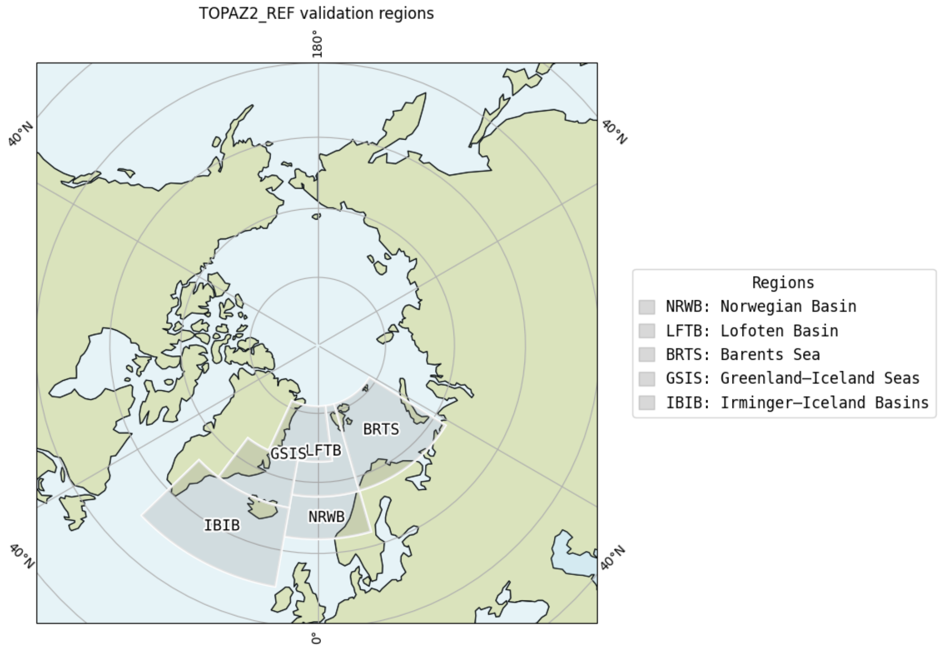
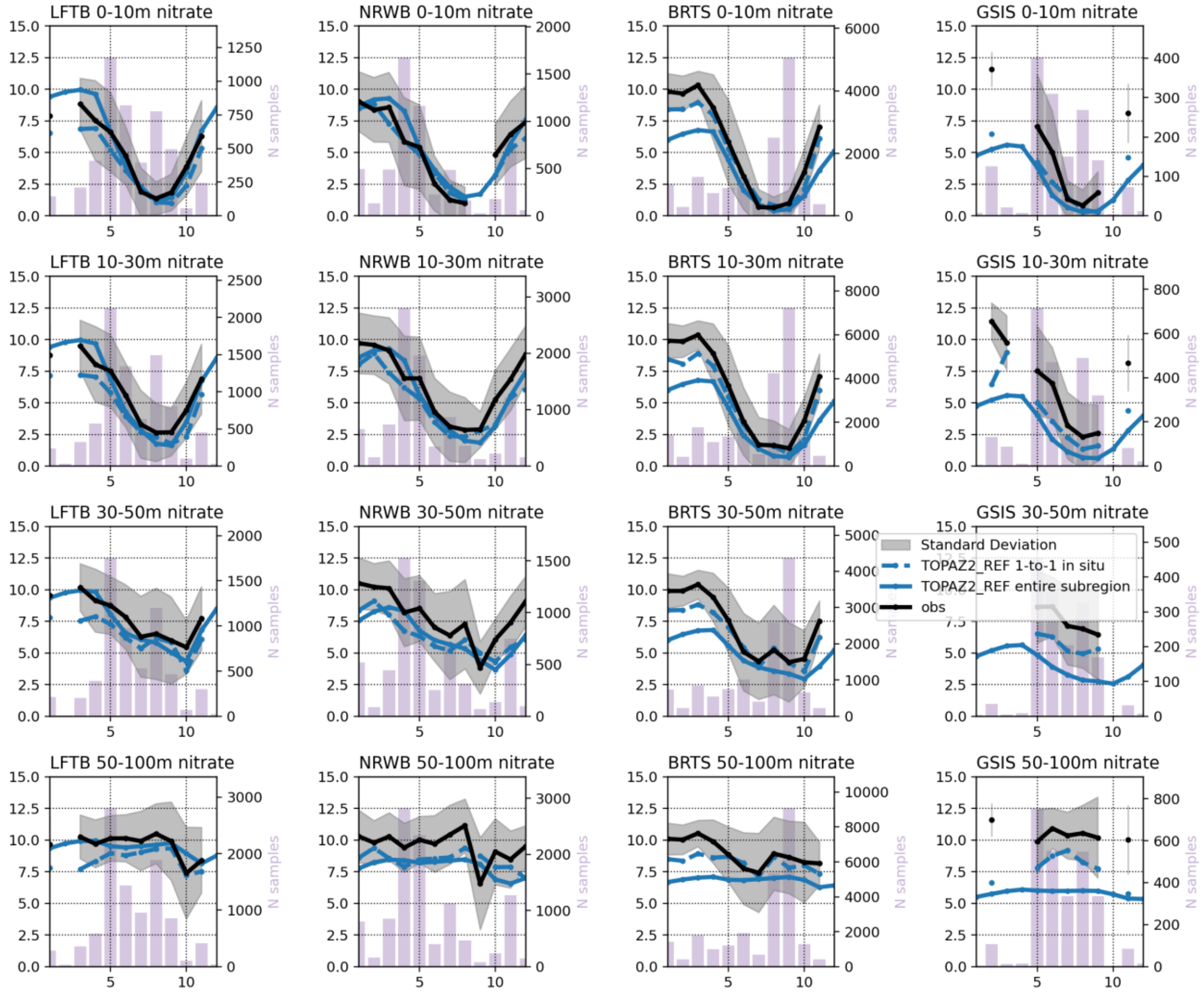
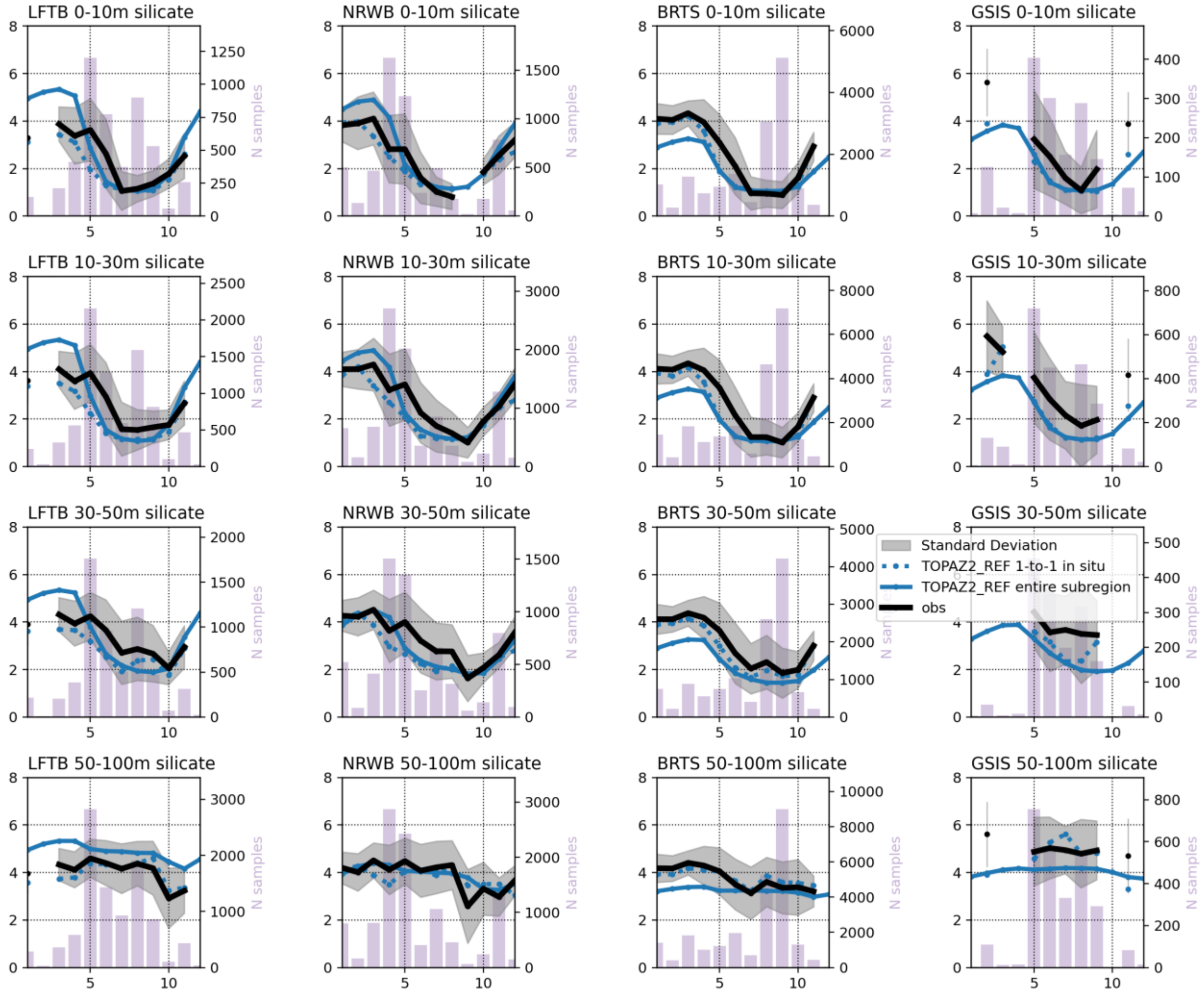

TOPAZ2_REF validation regions

Figure 63. Regions used in the validation of biogeochemical variables in the hindcast simulation with TOPAZ2 against in-situ data.


LFTB 0-10m nitrate
NRWB 0-10m nitrate
BRTS 0-10m nitrate
GSIS 0-10m nitrate
LFTB 10-30m nitrate
NRWB 10-30m nitrate
BRTS 10-30m nitrate
GSIS 10-30m nitrate
LFTB 30-50m nitrate
NRWB 30-50m nitrate
BRTS 30-50m nitrate
GSIS 30-50m nitrate
LFTB 50-100m nitrate
NRWB 50-100m nitrate
BRTS 50-100m nitrate
GSIS 50-100m nitrate
Standard Deviation
TOPAZ2_REF 1-to-1 in situ
TOPAZ2_REF entire subregion
obs
N samples
0.0
2.5
5.0
7.5
10.0
12.5
15.0
5
10

-



present a validation of the simulated seasonal nutrient cycle against in-situ observations across four Arctic subregions (LFTB, NRWB, BRTS, GSIS) and four depth layers (0–10 m, 10–30 m, 30–50 m, 50–100 m) in the period of 1997–2021.

Overall, the model captures the broad seasonal pattern, with winter maxima and summer minima evident across most regions and depths. However, several systematic differences emerge:

Depth Biases:

- At 30–50 m and 50–100 m, the model tends to underestimate nutrient concentrations. This may suggest insufficient vertical supply or overly strong biological drawdown in the simulation.

Regional Biases:

- The Barents Sea (BRTS) and Greenland–Iceland Seas (GSIS) consistently show lower modeled nutrient values than observed, especially for nitrate and phosphate. These discrepancies may reflect limitations in representing shelf–basin exchange, remineralization, or vertical mixing processes.

Taken together, the comparison highlights the model’s ability to reproduce seasonal variability in nutrient concentrations, while also revealing depth- and regional dependent biases given the sampling limitations. These biases highlight areas where the model may require refinement in vertical mixing, nutrient regeneration, or regional circulation to better capture subsurface nutrient dynamics.

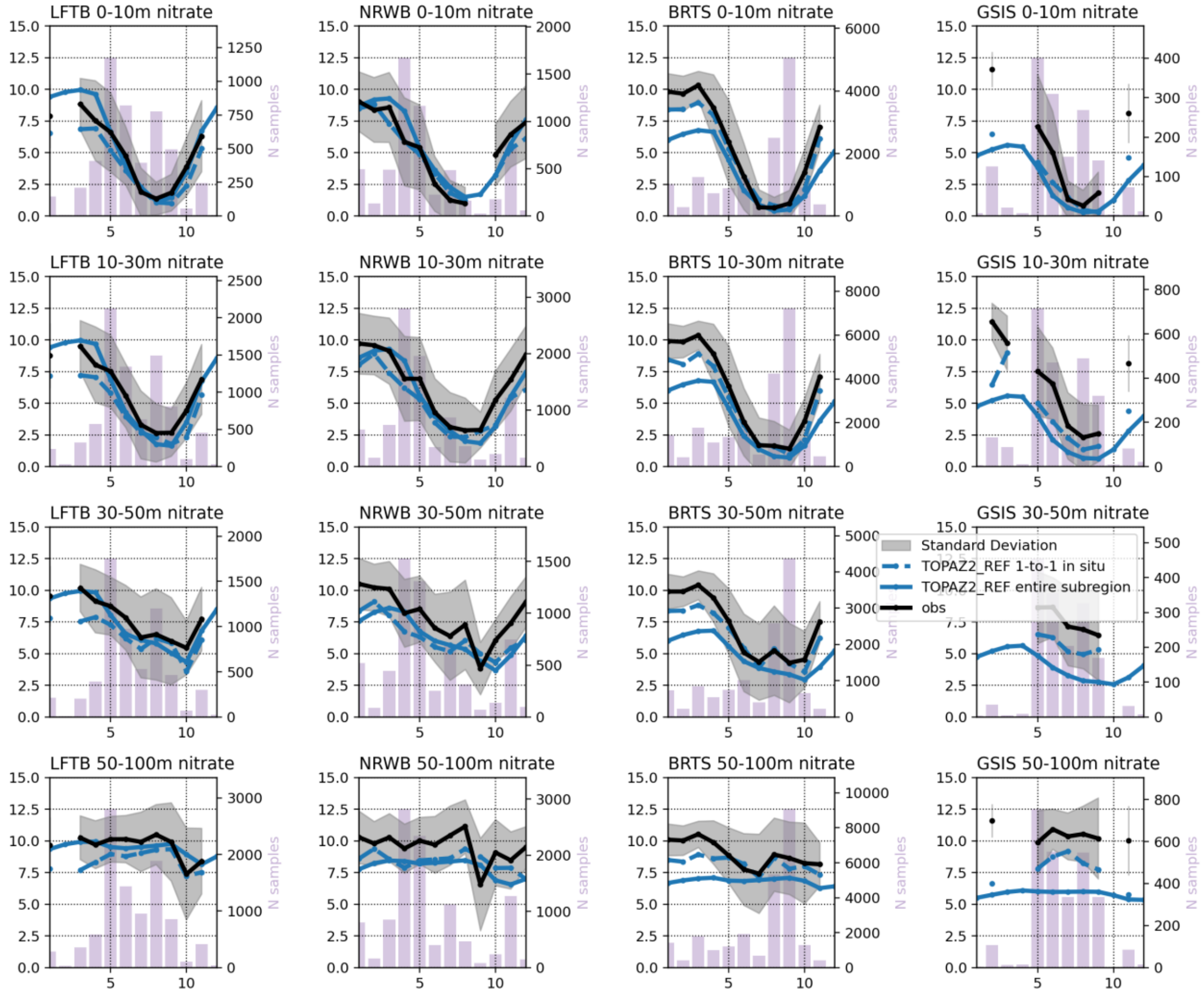


Figure 64. Seasonal cycle of nitrate (mmol m-3) from simulated monthly means (blue) against in-situ observations (black) at four depth ranges (0–10, 10–30, 30–50, 50–100 m) across the Lofoten Basin, Norwegian Basin, Barents Sea, and Greenland–Iceland Seas. Shaded areas show standard deviations. Bars indicate the number of in-situ samples. Solid blue lines represent area-averaged model values, while dashed blue lines show averaged model values sampled at the in-situ locations.

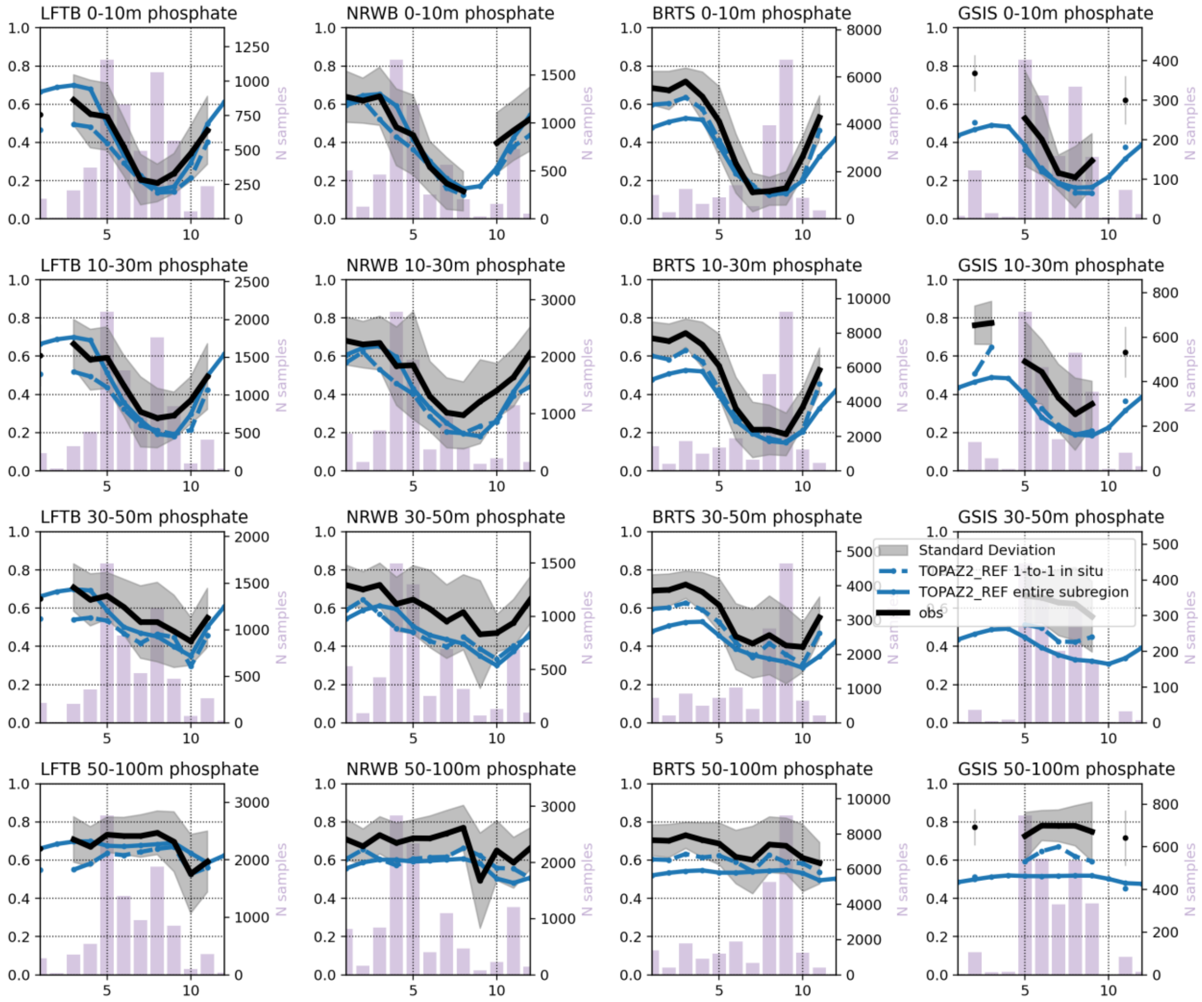


Figure 65. Seasonal cycle of phosphate (mmol m-3) from simulated monthly means (blue) against in-situ observations (black) at four depth ranges (0–10, 10–30, 30–50, 50–100 m) across the Lofoten Basin, Norwegian Basin, Barents Sea, and Greenland–Iceland Seas. Shaded areas show standard deviations. Bars indicate the number of in-situ samples. Solid blue lines represent area-averaged model values, while dashed blue lines show averaged model values sampled at the in-situ locations.

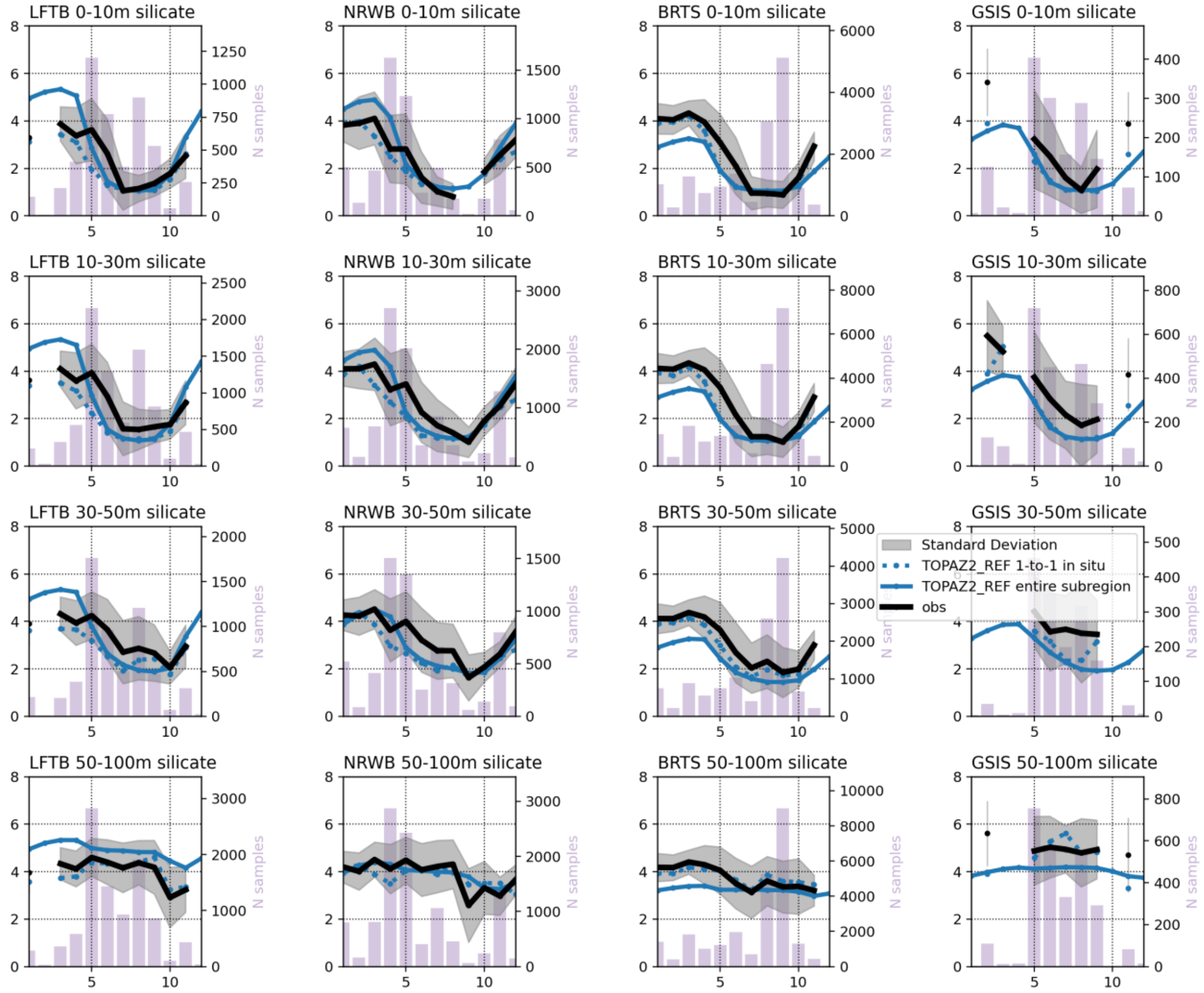
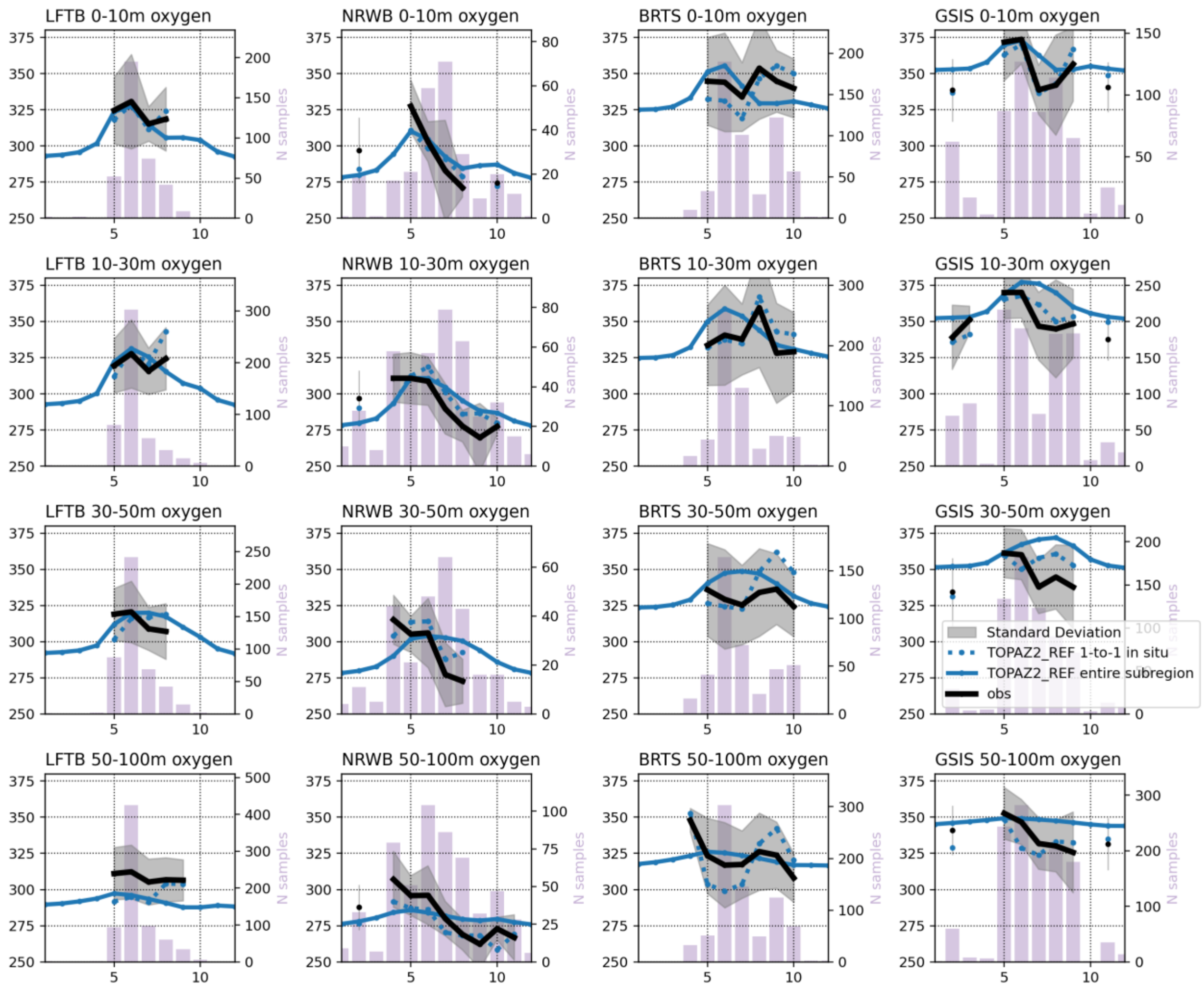


Figure 66. Seasonal cycle of silicate (mmol m-3) from simulated monthly means (blue) against in-situ observations (black) at four depth ranges (0–10, 10–30, 30–50, 50–100 m) across the Lofoten Basin, Norwegian Basin, Barents Sea, and Greenland–Iceland Seas. Shaded areas show standard deviations. Bars indicate the number of in-situ samples. Solid blue lines represent area-averaged model values, while dashed blue lines show averaged model values sampled at the in-situ locations.

The model reproduces the seasonal cycle and regional gradients of oxygen reasonably well across the four Arctic subregions (LFTB, NRWB, BRTS, GSIS) and four depth layers (0–10 m, 10–30 m, 30–50 m, 50–100 m), especially with model-observation one-to-one comparison (dashed blue lines) (



). The subregion-wide averages show larger discrepancies mainly due to limited in-situ sampling numbers.

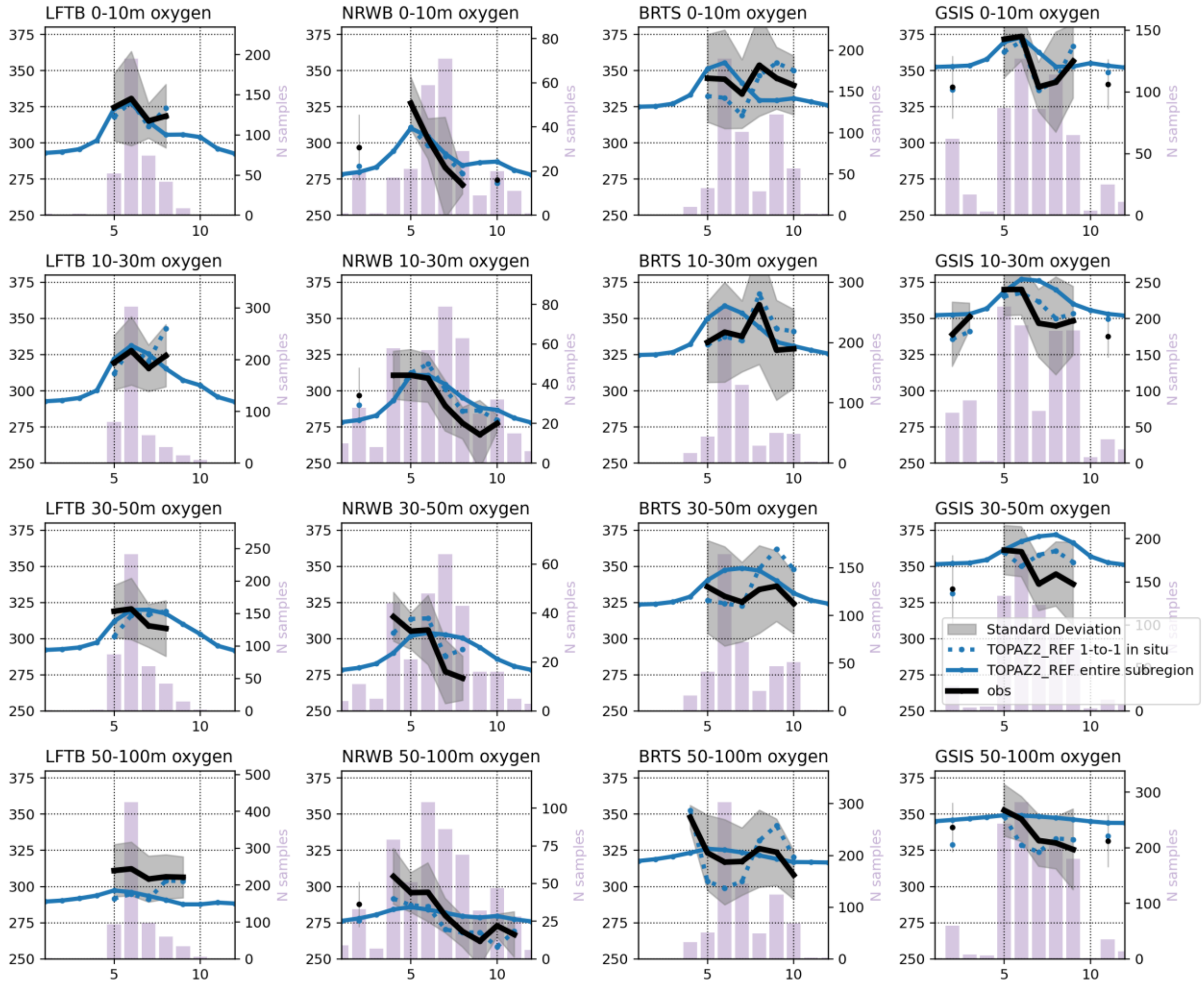


Figure 67. Seasonal cycle of oxygen (mmol m-3) from simulated monthly means (blue) against in-situ observations (black) at four depth ranges (0–10, 10–30, 30–50, 50–100 m) across the Lofoten Basin, Norwegian Basin, Barents Sea, and Greenland–Iceland Seas. Shaded areas show standard deviations. Bars indicate the number of in-situ samples. Solid blue lines represent area-averaged model values, while dashed blue lines show averaged model values sampled at the in-situ locations.

3.2.2.2 Chlorophyll

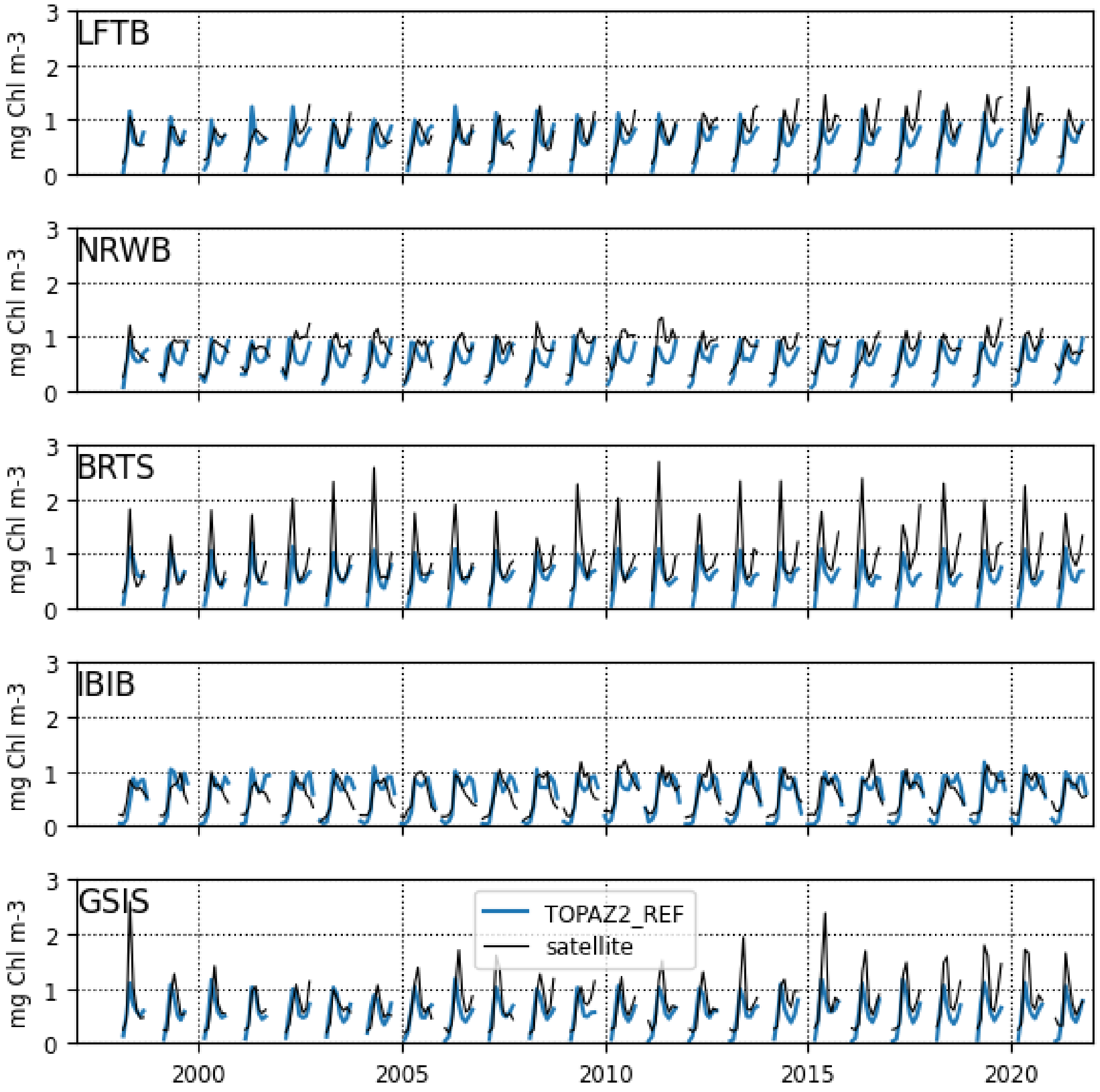


Figure 68. Time series (1998–2021) of monthly mean surface chlorophyll concentrations (mg Chl m-3) from the model compared with satellite observations in five subregions: Lofoten Basin, Norwegian Basin, Barents Sea, Irminger–Iceland Basin, and Greenland–Iceland Seas. Only co-located data points are included in the comparison.

Overall, chlorophyll concentrations are well represented in the model when compared with satellite observations. The main exception is the Barents Sea, where the model persistently underestimates chlorophyll throughout the entire time series.

3.2.2.3 Integrated primary production

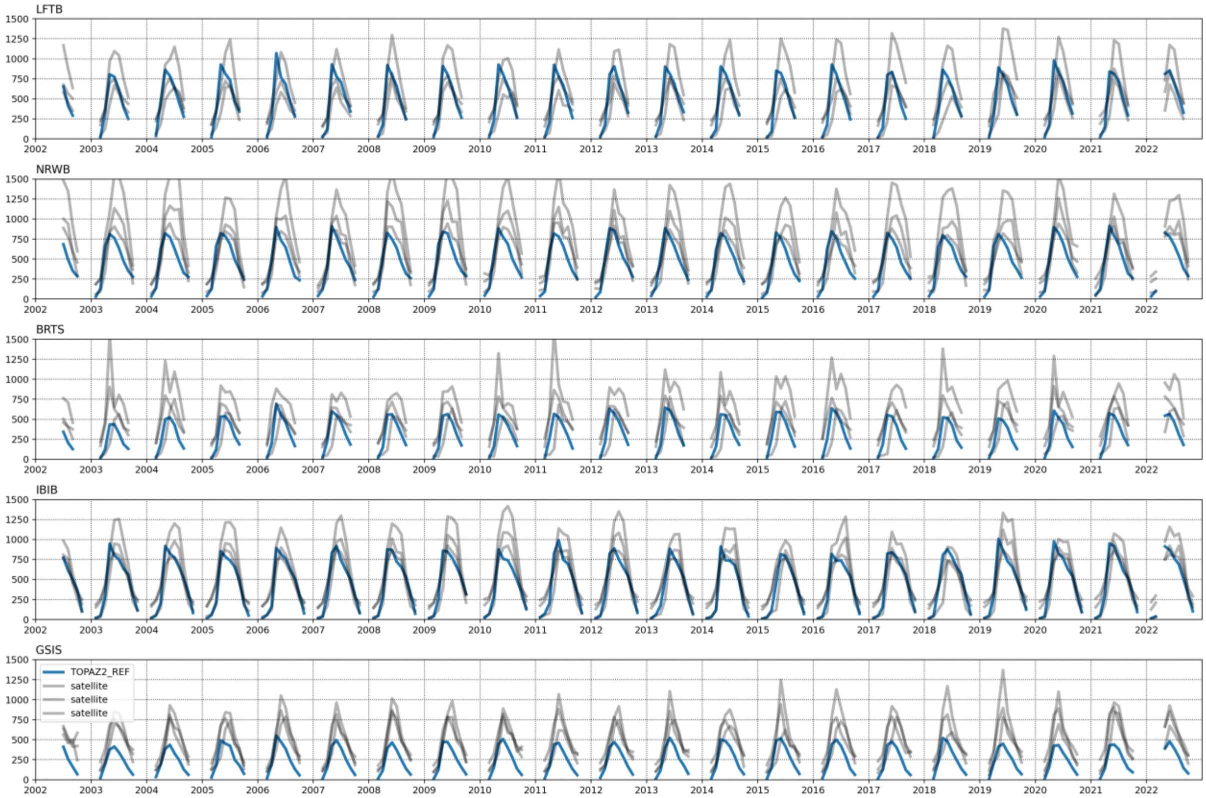


Figure 69. Time series (2002–2023) of monthly mean primary production from the model compared with satellite observations in five subregions: Lofoten Basin, Norwegian Basin, Barents Sea, Irminger–Iceland Basin, and Greenland–Iceland Seas. Only co-located data points are included in the comparison.

The model generally aligns more closely with the lower estimates among the three satellite products. Overall, it tends to underestimate primary production relative to satellite observations, likely in coastal regions. This discrepancy is consistent with known limitations of satellite algorithms, which often overestimate primary production near the coast due to optical complexities and algorithm assumptions.

3.2.3. Sea ice validation

The model reproduced the annual cycle of sea ice extent, but has a positive bias throughout the year (Figure 70). This is consistent with the surface temperatures in the Arctic being generally colder that both reanalysis on ARMOR3D. We do however see the same sea ice decline and interannual variability with years of extraordinary sea ice loss, such as 2007 and 2020, in both TOPAZ2 and OSI SAF. During the low sea-ice year, 2012, however, the minimu is not as visible in TOPAZ2 as in OSISAF, in this year much of the sea ice loss was on the Bering strait side of the Arctic where we can see that TOPAZ2 has a positive asea ice thickness bias (Figure 71). Sea ice thickness is also overall larger than observed from CS2SMOS, except in the central Arctic where it is thinner. The thicker sea ice in TOPAZ is shifted towards the Bering Strait, whereas it is more shifted towards the central Arctic in the observations.

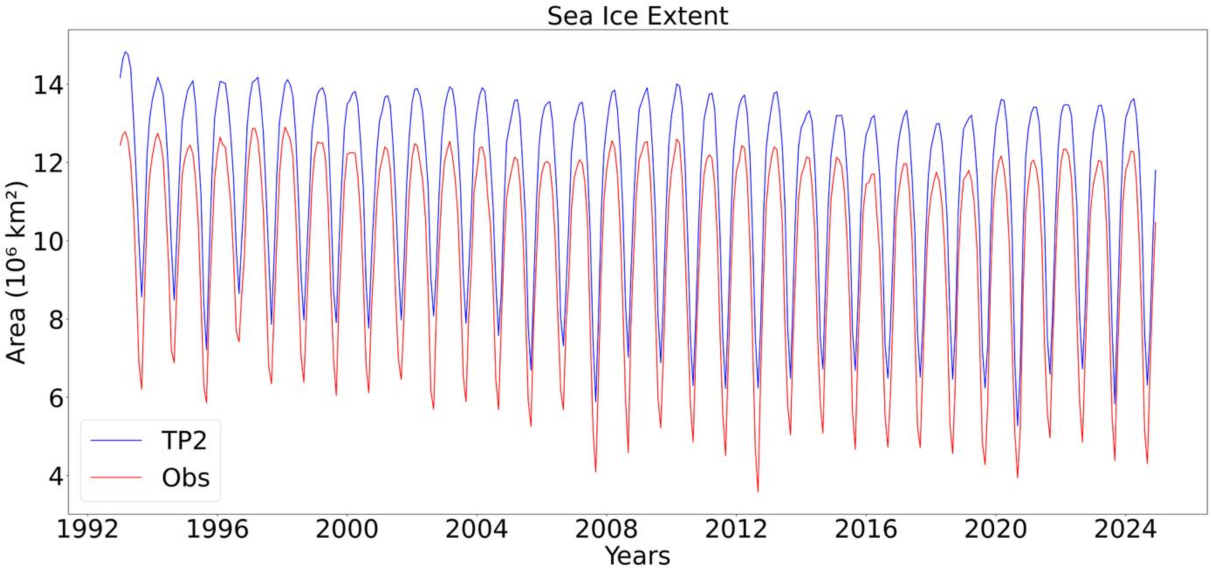
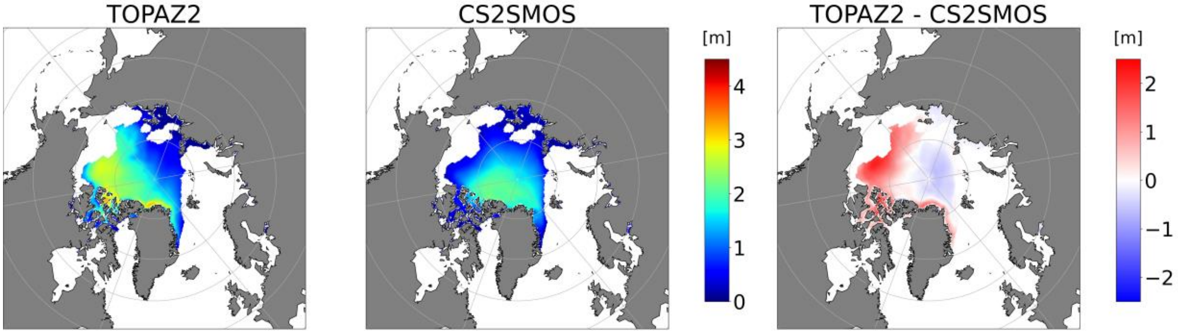


Figure 70. Sea ice extent from TOPAZ (TP2) and observations (OSI SAF Monthly).

SIT: October 2011-2024



SIT: March 2011-2024

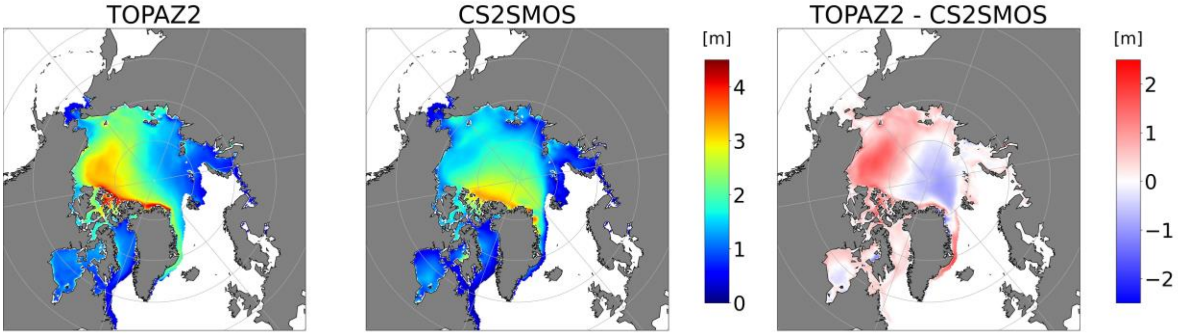


Figure 71. Average sea ice thickness in October and from TOPAZ2 and observations of CS2SMOS. The figure show averages from 2011 to 2024.

Box 3: Validation dataset

This box summarizes the main characteristics of the datasets used for model validation of TOPAZ2.

1/ Global Ocean Physics Reanalysis

Product ID: GLOBAL_MULTIYEAR_PHY_001_030

Spatial extent: Global Ocean Lat -80° to 90°Lon -180° to 179.92°
Spatial resolution: 0.083° × 0.083°
Period: 1 Jan 1993 to 23 Dec 2025
Available frequencies: Daily and monthly
DOI: https://doi.org/10.48670/moi-00021

2/ ARMOR3D
Product ID: MULTIOBS_GLO_PHY_TSUV_3D_MYNRT_015_012
Spatial extent: Global OceanLat -82.19° to 89.94°Lon -179.94° to 179.94°
Spatial resolution: 0.125° × 0.125°
Period: 1 Jan 1993 to 24 Jan 2026
Available frequencies: Daily and monthly
DOI: https://doi.org/10.48670/moi-00052

3/ MLD climatology -Level 3 gridded data
Spatial extent: Global Ocean
Spatial resolution: 2° × 2°
Website: https://mld.ifremer.fr/Data_Description.php

4/ Sea ice extent from OSI SAF
Product ID: SEAICE_GLO_SEAICE_L4_REP_OBSERVATIONS_011_009
Spatial extent: Global Ocean Lat -90.5° to 90°Lon -180° to 179.9°
Spatial resolution: 25 × 25 km
Period: 25 Oct 1978 to 29 Sep 2025
Available frequencies: Daily
DOI: https://doi.org/10.48670/moi-00136

5/ CryoSat-2/Sentinel-3/SMOS Merged Sea Ice Thickness
Spatial extent: 16.6 S to 90 N and 180 W to 180 E
Spatial resolution: 25 × 25 km
Period: 2010-11-15 to present
Available frequencies: Weekly
DOI: https://doi.org/10.57780/sm1-4f787c3

6/ WOA2023
Source: In situ observations
Spatial extent: Global Ocean
Spatial resolution: 1° x 1°
Period: Climatology 1965 – 2022
Used for: Nutrients and oxygen validation.
See: https://www.ncei.noaa.gov/access/world-ocean-atlas-2023/ ; for details.

7/ GLODAPv2.2023: the global interior ocean biogeochemical data product
A uniformly calibrated open ocean data product on inorganic carbon and carbon-relevant variables.
https://doi.org/10.5194/essd-16-2047-2024.

8/ ICES Data Portal, Dataset on Ocean Hydrochemistry, 2025.
https://data.ices.dk/view-map

9/CCHDO Hydrographic Data Archive
CCHDO contains hydrographic data collected from research vessels globally.
https://doi.org/10.6075/J0CCHAM8

10/Net Primary Production (NPP) Suite
The dataset includes Vertically Generalized Production Model (VGPM), Carbon-based productivity Mdoel (CbPM), Carbon, Absorption, and Fluorescence Euphotic-resolving (CAFE) and EPPLEY algorithms.
https://oceancolor.gsfc.nasa.gov/resources/atbd/npp/#sec_1

4. Data access

Hindcast files for both configurations are available on the EDITO platform (https://datalab.dive.edito.eu/file-explorer/seaclim/reference_simulations/hindcast/) in NetCDF format. To access the files, it is necessary to choose seaclim_project in the project frame (top left corner of the tab). Files are organizing by areas (Arctic or IBI Northwestern shelf), types (biogeochemistry, physics or waves), frequency and variable (Fig. X). All available variables are listed in the file list_of_outputs.xls, located in the area repertory. This file lists the outputs of the configuration with associated units, frequencies and files in which the variable can be found.
Currently the SEACLIM area on the EDITO platform is only open to members of SEACLIM, to access the SEACLIM project area, contact the EDITO support: support@edito.eu.

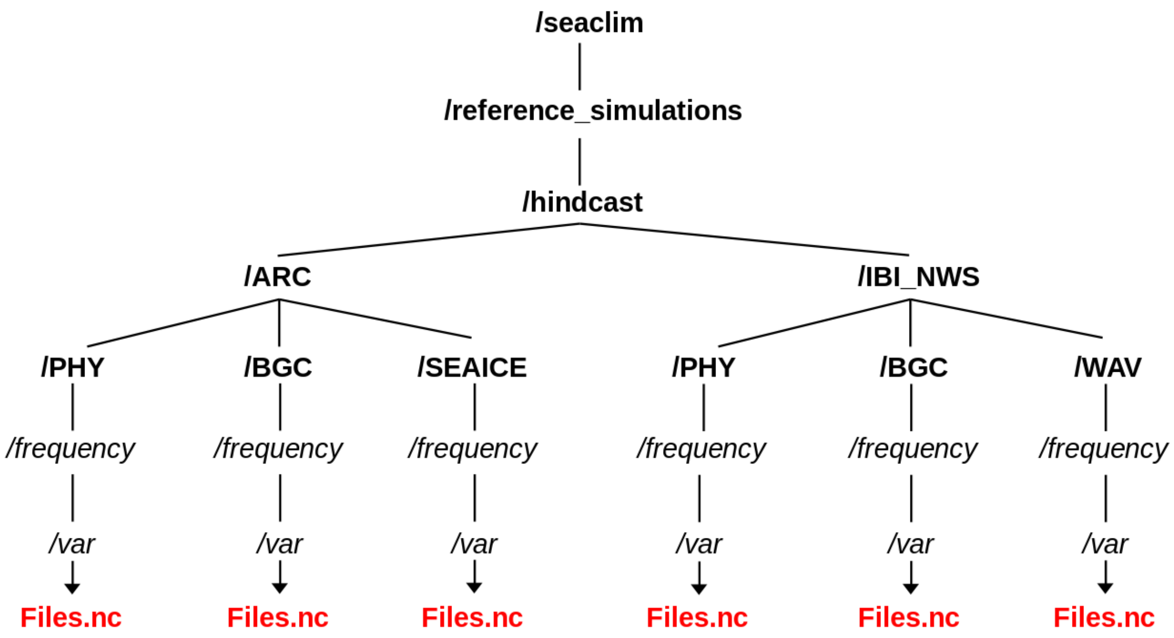


Fig. X: Files organization on EDITO (available frequencies: hourly, daily, 25h, monthly, yearly; var: variable).

5. References

Ali, A., Christensen, K. H., Breivik, Ø., Malila, M., Raj, R. P., Bertino, L., Chassignet, E. P., & Bakhoday-Paskyabi, M. (2019). A comparison of Langmuir turbulence parameterizations and key wave effects in a numerical model of the North Atlantic and Arctic Oceans. *Ocean Modelling*, *137*, 76–97. https://doi.org/10.1016/j.ocemod.2019.02.005

Ardhuin, F., Rogers, E., Babanin, A. V., Filipot, J.-F., Magne, R., Roland, A., van der Westhuysen, A., Queffeulou, P., Lefevre, J.-M., Aouf, L., & Collard, F. (2010). Semiempirical Dissipation Source Functions for Ocean Waves. Part I: Definition, Calibration, and Validation. *Journal of Physical Oceanography*, *40*(9), 1917–1941. https://doi.org/10.1175/2010JPO4324.1

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., & Gehlen, M. (2015). PISCES-v2: An ocean biogeochemical model for carbon and ecosystem studies. *Geoscientific Model Development*, *8*(8), 2465–2513. https://doi.org/10.5194/gmd-8-2465-2015

Bernard, B., Madec, G., Penduff, T., Molines, J.-M., Treguier, A.-M., Le Sommer, J., Beckmann, A., Biastoch, A., Böning, C., Dengg, J., Derval, C., Durand, E., Gulev, S., Remy, E., Talandier, C., Theetten, S., Maltrud, M., McClean, J., & De Cuevas, B. (2006). Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. *Ocean Dynamics*, *56*(5–6), 543–567. https://doi.org/10.1007/s10236-006-0082-1

Bertino, L., & Lisæter, K. A. (2008). The TOPAZ monitoring and prediction system for the Atlantic and Arctic Oceans. *Journal of Operational Oceanography*, *1*(2), 15–18.

http://www.ingentaconnect.com/content/imarest/joo/2008/00002008/00000002/art00002

Bidlot, J.-R., Janssen, P. A. E. M., & Abdalla, S. (2007). *A Revised Formulation of Ocean Wave Dissipation and its Model Impact. ECMWF Technical Memorandum 509.*

Bleck, R. (2002). An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modelling*, *4*(1), 55–88. https://doi.org/https://doi.org/10.1016/S1463-5003(01)00012-9

Bopp, L., Aumont, O., Cadule, P., Alvain, S., & Gehlen, M. (2005). Response of diatoms distribution to global warming and potential implications: A global model study. *Geophysical Research Letters*, *32*(19). https://doi.org/10.1029/2005GL023653

Braby, L., Backeberg, B., Krug, M., & Reason, C. (2020). Quantifying the Impact of Wind-Current Feedback on Mesoscale Variability in Forced Simulation Experiments of the Agulhas Current Using an Eddy-Tracking Algorithm. *Journal of Geophysical Research: Oceans*, *125*(1), 1–23. https://doi.org/10.1029/2019JC015365

Briegleb, B. P., & Light, B. (2007). *A Delta-Eddington Mutiple Scattering Parameterization for Solar Radiation in the Sea Ice Component of the Community Climate System Model*.

Bruggeman, J., & Karsten Bolding. (2011). *Framework for Aquatic Biogeochemical Models (FABM)*. (January).

Chaigneau, A. A., Reffray, G., Voldoire, A., & Melet, A. (2022). IBI-CCS: A regional high-resolution model to simulate sea level in western Europe. *Geoscientific Model Development*, *15*(5), 2035–2062. https://doi.org/10.5194/gmd-15-2035-2022

Chassignet, E., Hurlburt, H., Metzger, E. J., Smedstad, O., Cummings, J., Halliwell, G., Bleck, R., Baraille, R., Wallcraft, A., Lozano, C., Tolman, H., Srinivasan, A., Hankin, S., Cornillon, P., Weisberg, R., Barth, A., He, R., Werner, F., & Wilkin, J. (2009). US GODAE: Global Ocean Prediction with the HYbrid Coordinate Ocean Model (HYCOM). *Oceanography*, *22*(2), 64–75. https://doi.org/10.5670/oceanog.2009.39

Dai, A., & Trenberth, K. E. (2002). Estimates of Freshwater Discharge from Continents: Latitudinal and Seasonal Variations. *Journal of Hydrometeorology*, *3*(6), 660–668.

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research C: Oceans*, *109*(12), 1–20. https://doi.org/10.1029/2004JC002378

Duarte, P., Assmy, P., Hop, H., Spreen, G., Gerland, S., & Hudson, S. R. (2015). The importance of vertical resolution in sea ice algae production models. *Journal of Marine Systems*, *145*, 69–90. https://doi.org/10.1016/j.jmarsys.2014.12.004

Gehlen, M., Bopp, L., Ernprin, N., Aumont, O., Heinze, C., & Raguencau, O. (2006). Reconciling surface ocean productivity, export fluxes and sediment composition in a global biogeochemical ocean model. *Biogeosciences*, *3*(4), 521–537. //000243785300010

Gehlen, M., Gangstø, R., Schneider, B., Bopp, L., Aumont, O., & Ethe, C. (2007). The fate of pelagic CaCO3 production in a high CO2 ocean: a model study. *Biogeosciences*, *4*, 505–519. https://doi.org/10.5194/bg-4-505-2007

Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1998). A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. *Limnol. Oceanogr.*, *43*(4), 679–694. https://doi.org/10.4319/lo.1998.43.4.0679

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/https://doi.org/10.1002/qj.3803

Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., & Elliott, S. (2015). *CICE : the Los Alamos Sea Ice Model Documentation and Software User ' s Manual LA-CC-06-012*. 115.

Körtzinger, A., Hedges, J. I., & Quay, P. D. (2001). Redfield ratios revisited: Removing the biasing effect of anthropogenic CO $_2$. *Limnology and Oceanography*, *46*(4), 964–970. https://doi.org/10.4319/lo.2001.46.4.0964

Leclair, M., & Madec, G. (2009). A conservative leapfrog time stepping method. *Ocean Modelling*, *30*(2–3), 88–94. https://doi.org/10.1016/j.ocemod.2009.06.006

Lellouche, J.-M., Greiner, E., Bourdalle-Badie, R., Garric, G., Melet, A., Drevillon, M., Bricaud, A., Hamon, M., Le Galloudec, O., Reginer, C., Candela, T., Testut, C.-E., Gasparin, F., Ruggiero, G., Benkiran, M., Drillet, Y., & Le Traon, P.-Y. (2021). The Copernicus Global 1/12° Oceanic and Sea Ice GLORYS12 Reanalysis. *Frontiers in Earth Science*, *9*. https://doi.org/10.3389/FEART.2021.698876/FULL

Levier, B., Tréguier, A.-M., Madec, G., & Garnier, V. (2007). Free surface and variable volume in the nemo code. *Zenodo*.

Locarnini, R. A., Mishonov, A. V, Baranova, O. K., Boyer, T. P., Zweng, M. M., Garcia, H. E., Reagan, J. R., Seidov, D., Weathers, K. W., Paver, C. R., & Smolyar, I. V. (2019). *WORLD OCEAN ATLAS 2018 Volume 1: Temperature*. http://www.nodc.noaa.gov/

Madec, G. and the N. S. T. (2024). NEMO Ocean Engine Reference Manual. *Zenodo*.

Mesinger, F., & Arakawa, A. (1976). *Numerical methods used in atmospheric models.*

Sakov, P., Counillon, F., Bertino, L., Lisæter, K. A., Oke, P. R., & Korablev, A. (2012). TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic. *Ocean Science*, *8*(4), 633–656. https://doi.org/10.5194/os-8-633-2012

Schneider, B., Bopp, L., Gehlen, M., Segschneider, J., Frölicher, T. L., Cadule, P., Friedlingstein, P., Doney, S. C., Behrenfeld, M. J., & Joos, F. (2008). Climate-induced interannual variability of marine primary and export production in three global coupled climate carbon cycle models. *Biogeosciences*, *5*(2), 597–614. https://doi.org/10.5194/bg-5-597-2008

Séférian, R., Bopp, L., Gehlen, M., Orr, J. C., Ethé, C., Cadule, P., Aumont, O., Salas y Mélia, D., Voldoire, A., & Madec, G. (2013). Skill assessment of three earth system models with common marine biogeochemistry. *Climate Dynamics*, *40*(9–10), 2549–2573. https://doi.org/10.1007/s00382-012-1362-8

Steinacher, M., Joos, F., Frölicher, T. L., Bopp, L., Cadule, P., Doney, S. C., Gehlen, M., Schneider, B., & Segschneider, J. (2010). Projected 21st century decrease in marine productivity: a multi-model analysis. *Biogeosciences*, *7*, 979–1005. https://doi.org/10.5194/bgd-6-7933-2009

Tagliabue, A., Bopp, L., Dutay, J.-C., Bowie, A. R., Chever, F., Jean-Baptiste, P., Bucciarelli, E., Lannuzel, D., Remenyi, T., Sarthou, G., Aumont, O., Gehlen, M., & Jeandel, C. (2010). Hydrothermal contribution to the oceanic dissolved iron inventory. *Nature Geoscience*, *3*(4), 252–256. https://doi.org/10.1038/ngeo818

Takahashi, T., Broecker, W. S., & Langer, S. (1985). Redfield ratio based on chemical data from isopycnal surfaces. *Journal of Geophysical Research: Oceans*, *90*(C4), 6907–6924. https://doi.org/10.1029/JC090iC04p06907

Toledano, C., Ghantous, M., Lorente, P., Dalphinet, A., Aouf, L., & Sotillo, M. G. (2022). Impacts of an Altimetric Wave Data Assimilation Scheme and Currents-Wave Coupling in an Operational Wave System: The New Copernicus Marine IBI Wave Forecast Service. *Journal of Marine Science and Engineering*, *10*(4). https://doi.org/10.3390/JMSE10040457

Turner, A. K., & Hunke, E. C. (2015). Impacts of a mushy-layer thermodynamic approach in global sea-ice simulations using the <scp>CICE</scp> sea-ice model. *Journal of Geophysical Research: Oceans*, *120*(2), 1253–1275. https://doi.org/10.1002/2014JC010358

Wicker, L. J., & Skamarock, W. C. (2002). Time-Splitting Methods for Elastic Models Using Forward Time Schemes. In *Monthly Weather Review*.

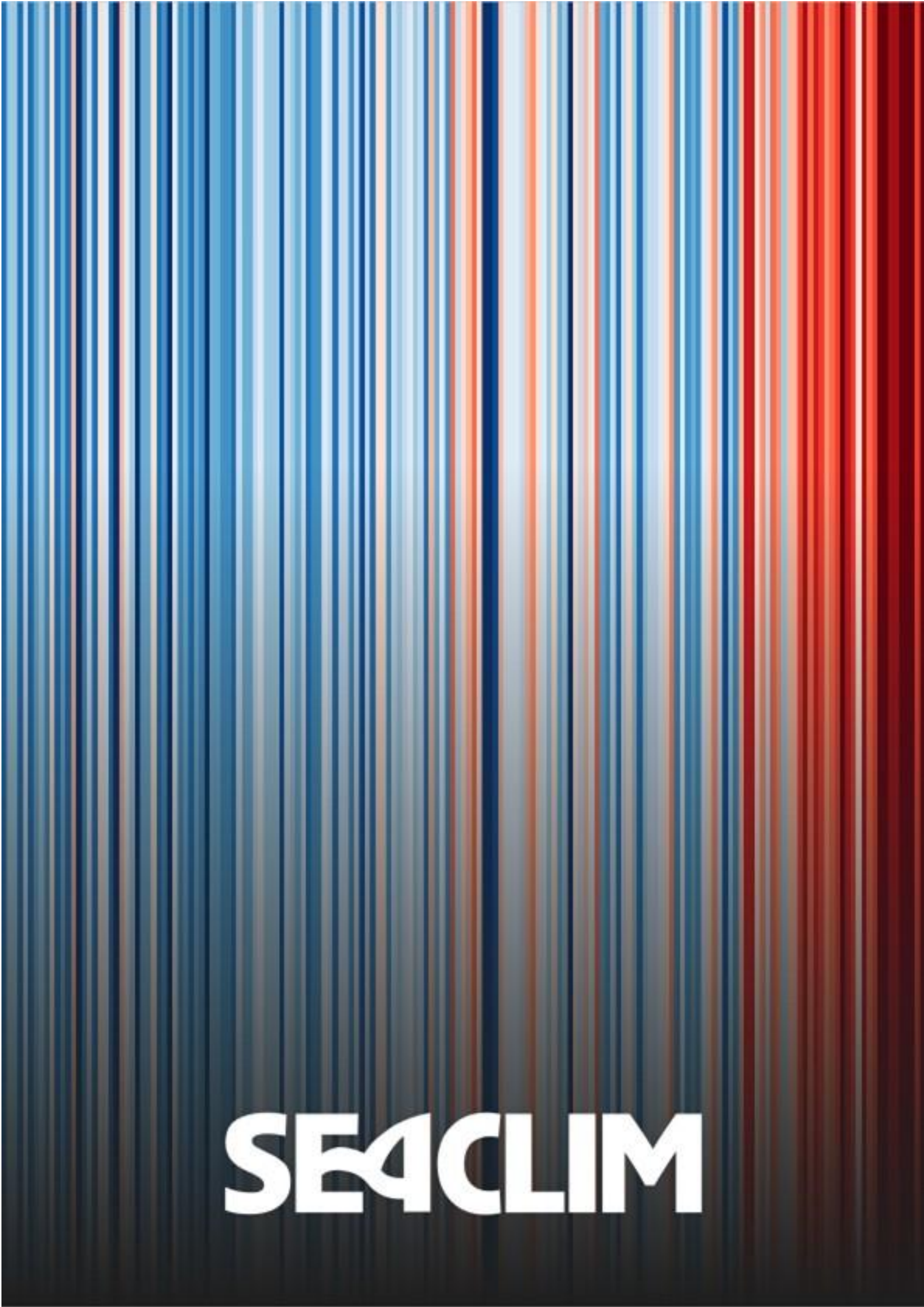
Xie, J., De Vos, M., Bertino, L., Zhu, J., & Counillon, F. (2020). Impact of assimilating altimeter data on eddy characteristics in the South China Sea. *Ocean Modelling*, *155*, 101704. https://doi.org/10.1016/j.ocemod.2020.101704

Xie, J., Raj, R. P. R. P., Bertino, L., Samuelsen, A., & Wakamatsu, T. (2019). Evaluation of Arctic Ocean surface salinities from the Soil Moisture and Ocean Salinity (SMOS) mission against a regional reanalysis and in situ data. *Ocean Science*, *15*(5), 1191–1206. https://doi.org/10.5194/os-15-1191-2019

Yumruktepe, V. Ç., Samuelsen, A., & Daewel, U. (2022). ECOSMO II(CHL): a marine biogeochemical model for the North Atlantic and the Arctic. *Geoscientific Model Development*, *15*(9), 3901–3921. https://doi.org/10.5194/gmd-15-3901-2022

Zweng, M. M., Reagan, J. R., Seidov, D., Boyer, T. P., Locarnini, R. A., Garcia, H. E., Mishonov, A. V, Baranova, O. K., Weathers, K. W., Paver, C. R., Smolyar, I. V, Ross, W. L., Jacobs, N., Volz, S., & Administrator, A. (2019). *WORLD OCEAN ATLAS 2018 Volume 2: Salinity*.

Appendix A: Northeast Atlantic validation (supplementary figures)


SE4CLIM