Description and evaluation of the CNRM-Cerfacs Climate Prediction System (C3PS)

Emilia Sanchez-Gomez[1], Roland Séférian[2], Lauriane Batté[3], Christophe Cassou[4], Boris Dewitte[5], Marie-Pierre Moine[6], Rym Msadek[7], Chloe Prodhomme[8], Yeray Santana-Falcón[9], Laurent Terray[1], and Aurore Voldoire[10]

[1]CERFACS
[2]CNRM (Météo-France/CNRS)
[3]Météo-France
[4]CNRS-Cerfacs
[5]CEAZA
[6]CECI-CERFACS
[7]CNRS/CERFACS
[8]No Institution
[9]Centre National de Recherches Météorologiques
[10]CNRM / Météo-France

April 05, 2024

Abstract

The CNRM-Cerfacs Climate Prediction System (C3PS) is a new research modeling tool for performing climate reanalyses and seasonal-to-multiannual predictions for a wide array of earth system variables. C3PS is based on the CNRM-ESM2-1 model including interactive aerosols and stratospheric chemistry schemes as well as terrestrial and marine biogeochemistry enabling a comprehensive representation of the global carbon cycle. C3PS operates through a seamless coupled initialization for the atmosphere, land, ocean, sea ice and biogeochemistry components that allows a continuum of predictions across seasonal to interannual time-scales. C3PS has also contributed to the Decadal Climate Prediction Project (DCPP-A) as part of the sixth Coupled Model Intercomparison Project (CMIP6). Here we describe the main characteristics of this novel earth system-based prediction platform, including the methodological steps for obtaining initial states to produce forecasts. We evaluate the entire C3PS initialisation procedure with the most up-to-date observations and reanalysis over 1960-2021, and we discuss the overall performance of the system in the light of the lessons learnt from previous and actual prediction platforms. Regarding the forecast skill, C3PS exhibits comparable seasonal predictive skill to other systems. At the decadal scale, C3PS shows significant predictive skill in surface temperature during the first two years after initialisation in several regions of the world. C3PS also exhibits potential predictive skill in net primary production and carbon fluxes several years in advance. This expands the possibility of applications of forecasting systems, such as the possibility of performing multi-annual predictions of marine ecosystems and carbon cycle.

Hosted file

`Esanchez-Gomez_et_al_C3PS_manuscrit_submitted.docx` available at https://authorea.com/users/676550/articles/721257-description-and-evaluation-of-the-cnrm-cerfacs-climate-prediction-system-c3ps

Hosted file

C3PS_supplementary_figure.docx available at https://authorea.com/users/676550/articles/721257-description-and-evaluation-of-the-cnrm-cerfacs-climate-prediction-system-c3ps

Description and evaluation of the CNRM-Cerfacs Climate Prediction System (C3PS)

E. Sanchez-Gomez[1], R. Séférian[2], L. Batté[3], C. Cassou[1], B. Dewitte[1,4,5], M.P. Moine M[1], R. Msadek[1], C. Prodhomme, Y. Santana-Falcón[2], L. Terray[1], A. Voldoire[2]

[1] CECI, Université de Toulouse, CNRS, Cerfacs, Toulouse, France
[2] CNRM, Université de Toulouse, Météo-France, Toulouse, France
[3] DCSC/ACS, Météo-France, Toulouse, France
[4] Center of Advanced Studies in Arid Zones (CEAZA), Coquimbo, Chile
[5] Center for Ecology and Sustainable Management of Oceanic Islands (ESMOI), Departamento de Biología Marina, Facultad de Ciencias del Mar, Universidad Católica del Norte, Chile

Corresponding author: first and last name (emilia.sanchez@cerfacs.fr)

Key Points:

- This study introduces and assesses C3PS, the new CNRM-Cerfacs Earth System-based prediction platform
- The plateform can provide climate predictions from seasonal to multi-annual timescales for relevant physical and biogeochemical fields
- The most outstanding result is the ability of C3PS to predict the net primary production and carbon fluxes at multi-annual timescales

Abstract

The CNRM-Cerfacs Climate Prediction System (C3PS) is a new research modeling tool for performing climate reanalyses and seasonal-to-multiannual predictions for a wide array of earth system variables. C3PS is based on the CNRM-ESM2-1 model including interactive aerosols and stratospheric chemistry schemes as well as terrestrial and marine biogeochemistry enabling a comprehensive representation of the global carbon cycle. C3PS operates through a seamless coupled initialization for the atmosphere, land, ocean, sea ice and biogeochemistry components that allows a continuum of predictions across seasonal to interannual time-scales. C3PS has also contributed to the Decadal Climate Prediction Project (DCPP-A) as part of the sixth Coupled Model Intercomparison Project (CMIP6).

Here we describe the main characteristics of this novel earth system-based prediction platform, including the methodological steps for obtaining initial states to produce forecasts. We evaluate the entire C3PS initialisation procedure with the most up-to-date observations and reanalysis over 1960-2021, and we discuss the overall performance of the system in the light of the lessons learnt from previous and actual prediction platforms. Regarding the forecast skill, C3PS exhibits comparable seasonal predictive skill to other systems. At the decadal scale, C3PS shows significant predictive skill in surface temperature during the first two years after initialisation in several regions of the world. C3PS also exhibits potential predictive skill in net primary production and carbon fluxes several years in advance. This expands the possibility of applications of forecasting systems, such as the possibility of performing multi-annual predictions of marine ecosystems and carbon cycle.

Plain Language Summary

The study introduces and assesses the new climate prediction platform C3PS developed by the CNRM-Cerfacs modelling group in the framework of the H2020 TRIATLAS project. This prediction system is based on the last version of the CNRM earth system model, CNRM-ESM2.1, and was designed to produce predictions from seasonal to multi-annual scales. C3PS is the result of the joint long-term effort of experts in seasonal and decadal forecasting and modellers of ocean physics and biogeochemistry within the CNRM-Cerfacs research group.

An innovative aspect of our study is that it focuses on validating the initialization procedure, which is not often done in other studies presenting forecasting systems. We believe that the study of the reconstructions created to initialize the climate prediction systems is relevant, and even more so in the context of the new applications offered in the prediction of marine biogeochemistry and carbon fluxes.

Regarding forecast skill, C3PS exhibits comparable seasonal predictive skill to other systems. On a multi-year scale, C3PS shows potential skill not only in physics, but also in net primary

production and carbon fluxes up to three years in advance, which extends the possibilities of application to marine ecosystems and multi-year carbon cycle forecasts.

1 Introduction

The field of near-term climate prediction has grown rapidly since the pioneering studies of Smith et al. (2007), Keenlyside et al. (2008), Pohlmann et al. (2009) and the very first attempt of decadal prediction coordinated experiments as conducted under the umbrella of the Fifth Phase of the Coupled Model Intercomparison Project (CMIP5). The analysis of CMIP5 decadal prediction experiments revealed a wide range of skill for different variables and across various prediction systems (Doblas-Reyes et al., 2013, Garcia-Serrano et al., 2015; Bellucci et al., 2015 amongst others). Recently, CMIP6 has proposed a new decadal prediction coordinated exercise with improvements with respect to CMIP5 (Boer et al., 2016). These improvements include not only model improvements, but also the increase of the number of starting dates and ensemble members in the decadal forecast archive, in order to ensure a robust assessment of decadal predictive skill. Results from CMIP6 show a substantial improvement of the Sea Surface Temperature (SST) prediction in the North Atlantic, in particular over the subpolar gyre (Borchert et al., 2021, Delgado-Torres et al., 2022). Over land, a significant increase in prediction skill of surface air temperature (SAT) is also reported (Monerie et al., 2018 ; Wu et al., 2019 ; Smith et al., 2019). Moreover, by the use of large ensembles, skillful predictions have been achieved for atmospheric patterns such as blocking (Schuster et al., 2019; Athanasiadis et al., 2020) and the North Atlantic Oscillation (Smith et al., 2020).

While the potential for useful applications has been demonstrated, the CMIP5/CMIP6 experiments have also highlighted a number of outstanding research questions and challenges in the climate prediction field (Kirtman et al., 2013; Meehl et al., 2014; Cassou et al., 2018; Bojovic et al., 2019). Previous decadal prediction exercises highlight the need for a better understanding of three key aspects for better exploiting the climate predictive potential and improving estimates of climate predictability at different timescales (Keenlyside and Ba, 2010; Cassou et al., 2018; Verfaillie et al., 2021) : i) the physical mechanisms of climate predictability, ii) initialization, and response to external forcing; iii) and an improvement of the forecast quality evaluation process.

One of the outstanding challenges is to identify the extent to which model prediction skill across a continuum of time-scales may benefit from initialization. Indeed, by establishing a framework for testing the added value of model initialisation, as well as prescribing external forcings, decadal prediction systems have bridged the gap between well-established seasonal prediction and near-term projections (Meehl et al. 2009). In this sense, decadal predictions can provide seamless climate information from one month to several years ahead, offering the opportunity of exploring predictability at different timescales (Choi et Sun, 2023). This is relevant as it provides climate information addressing a growing demand from policy makers and stakeholders in the context of climate risk management.

Moreover, the required reduction of human-induced CO2 emissions and the need for adaptation of several sectors have, over the recent years, widened the range of application of climate predictions, with the inclusion of new Earth System components. Earth System Models (ESMs) have been recently implemented in climate prediction systems, allowing to explore the predictability of marine biogeochemistry and marine ecosystems (Séférian et al., 2014, Park et al., 2019, Yeager et al., 2022), terrestrial carbon fluxes (Seferian et al., 2018) and air-sea carbon fluxes and carbon budgets (Lovenduski et al., 2019, Ilyna et al., 2021). Following this path and in order to provide seamless seasonal to interannual predictions for relevant physical and earth system variables, the CNRM-Cerfacs modelling group has developed a new prediction platform, called C3PS, which is based on the CNRM-ESM2.1 model (Seferian et al., 2019). The birth of C3PS was possible by bringing together the expertise of the CNRM-Cerfacs modelling group in terms of seasonal and multi-annual climate predictions and the latest developments in earth system modelling made in the context of CMIP6 (Séférian et al., 2020).

In the present study, we introduce the C3PS system, by highlighting its main characteristics, and the seamless/coupled initialization method used for the atmosphere, ocean and marine biogeochemistry components. This coupled initialization has been achieved to enable the investigation of predictability across a continuum of time-scales, from seasons to years. Conversely to other studies presenting climate prediction systems, we perform an exhaustive evaluation of the initialization procedure, assessing its strengths and weaknesses. Finally, we also evaluate the performance of the C3PS system based on a variety of diagnostics and metrics.

Section 2 describes the main characteristics of C3PS, the initialization procedure and the experimental protocol used to perform the seasonal to interannual predictions. Section 3 presents the reference datasets and metrics used. Section 4 provides a basic evaluation of the assimilation experiments used in the C3PS initialization. Section 5 assesses the skill of essential physical and biogeochemical fields at different time scales, and the concluding remarks are presented in section 6.

2 Earth system-based prediction platform

2.1 Model description

The backbone of the C3PS platform is CNRM-ESM2-1 which is the Earth System model of second generation developed by CNRM-Cerfacs modelling group for CMIP6 (Séférian et al., 2019).

The atmosphere component of CNRM-ESM2-1 is based on the global spectral model ARPEGE-Climat version 6.3 (Roehrig et al., 2020). ARPEGE-Climat resolves atmospheric dynamics using a T127 linear truncation, the physics is resolved on the corresponding reduced grid which offers a spatial resolution of about 150 km in both longitude and latitude. CNRM-ESM2-1 employs a ‘‘high-top’’ configuration with 91 vertical levels that extend from the surface to 0.01 hPa in the mesosphere; 15 hybrid σ-pressure levels are available below 1500 m.

The atmospheric chemistry scheme of CNRM-ESM2-1 is Reactive Processes Ruling the Ozone Budget in the Stratosphere version 2 (REPROBUS-C_v2). This scheme resolves the spatial distribution of 63 chemistry species but does not represent the low troposphere ozone non-methane hydrocarbon chemistry. CNRM-ESM2-1 also activates an interactive tropospheric aerosol scheme included in the atmospheric component ARPEGE-Climat. This aerosol scheme, named Tropospheric Aerosols for ClimaTe In CNRM (TACTIC_v2), represents the main anthropogenic and natural aerosol species of the troposphere.

The surface state variables and fluxes at the surface-atmosphere interface are simulated by the SURFEX modeling platform version 8.0 over the same grid and with the same time-step as the atmosphere model. Over the land surface, CNRM-ESM2-1 uses the ISBA-CTRIP land surface modeling system to solve energy, carbon and water budgets at the land surface (Decharme et al., 2019 ; Delire et al., 2019). Its physical core explicitly solves the one-dimensional Fourier and Darcy laws throughout the soil, accounting for the hydraulic and thermal properties of soil organic carbon. It uses a 12-layer snow model of intermediate complexity that allows to separate water and energy budgets for the soil and the snowpack. CTRIP is a dynamic river flooding scheme in which floodplains interact with the soil and the atmosphere through free-water evaporation, infiltration and precipitation interception. The ISBA-CTRIP land surface scheme also embeds a two-dimensional diffusive groundwater scheme to represent unconfined aquifers and upward capillarity fluxes into the superficial soil. More details on these physical aspects can be found in (Decharme et al., 2019). ISBA-CTRIP captures the land carbon cycle and vegetation-climate interactions with the representation of plant physiology, carbon allocation and turnover, and carbon cycling through litter and soil. It includes a module for wildfires, land use and land cover changes, and carbon leaching through the soil and transport of dissolved organic carbon to the ocean. A detailed description of the terrestrial carbon cycle can be found in Delire et al. (2019).

The ocean component of CNRM-ESM2-1 is the Nucleus for European Models of the Ocean (NEMO) version 3.6 (Madec et al., 2017) which is coupled to both the Global Experimental Leads and ice for ATmosphere and Ocean (GELATO) sea-ice model (Salas Mélia, 2002) version 6, and also the marine biogeochemical model Pelagic Interaction Scheme for Carbon and Ecosystem Studies version 2-gas (PISCESv2-gas). NEMOv3.6 operates on the eORCA1L75 grid (Mathiot et al., 2017) which offers a nominal resolution of 1° to which a latitudinal grid refinement of 1/3° is added in the tropics; this grid describes 75 ocean vertical layers using a vertical z*-coordinate with partial step bathymetry formulation (Bernard et al., 2006).

The ocean biogeochemical component of CNRM-ESM2-1 uses the Pelagic Interaction Scheme for Carbon and Ecosystem Studies model version 2 coupled with trace gases module (PISCESv2-gas), which derives from PISCESv2 as described in Aumont et al. (2015). PISCESv2-gas simulates the distribution of five nutrients (from macronutrients: nitrate, ammonium, phosphate, and silicate to micronutrient: iron), which regulate the growth of two explicit phytoplankton classes (nanophytoplankton and diatoms). PISCESv2-gas also simulates the ocean carbon cycle with the ocean carbonate chemistry, that is the dissolved inorganic carbon (DIC) and the alkalinity (Alk) and two organic carbon pools. The dissolved oxygen is prognostically simulated using two different oxygen-to-carbon ratios, one when ammonium is converted to or mineralized from organic matter, the other when oxygen is consumed during nitrification. Their values have been set respectively to 131/122 and 32/122. At the ocean surface, PISCESv2-gas exchanges carbon, oxygen, dimethylsulfide (DMS) and nitrous oxide (N_2O) tracers with the atmosphere using the revised air-sea exchange bulk formulation as in Wanninkhof (2014). PISCESv2-gas uses several boundary conditions which represent the supply of nutrients from five different sources: atmospheric deposition, rivers, sediment mobilization, sea-ice and hydrothermal vents.

2.2 Forcings

This section details the CMIP6 external forcing implementation into the C3PS platform. We align as much as possible to requirements of the CMIP6 Decadal Prediction Project (DCPP) protocol (Boer et al. 2016). For all the experiments whose simulated period lies within the historical period as labelled by CMIP6, i.e. from 1850 to 2014, we apply a conservative approach by using the exact set-up that was used for the contribution to the CMIP6/DECK historical experiment (Eyring et al., 2016). Greenhouse gases concentrations, except stratospheric ozone, are implemented as recommended in Meinshausen et al. (2017). The reader is referred to Séférian et al. (2019) and Michou et al. (2020) for details on the implementation of the forcings for CMIP6.

For simulated years after 2014 and in accordance with the DCPP protocol, the Shared Socioeconomic Pathway (SSP) 2-4.5 scenario forcing is prescribed (O’Neill et al., 2016). This is the “middle-of-the-road” scenario of the SSP2 socioeconomic pathways, with an intermediate 4.5 W/m^2 radiative forcing level by 2100 (Gidden et al., 2019).

The major difference between the implementation of the external forcing in the C3PS platform and the usual CMIP6 simulation set-up for CNRM-ESM2-1 is the volcanic forcing. The CMIP6 experimental protocol now requires the use of a stratospheric volcanic background forcing (monthly climatology computed from years 1850–2000 volcanic forcing) during pre-industrial and future eras. However, over the 1850-2014, the volcanic forcing can be lower than the background forcing as used for the future period (beyond 2015). In consequence, we applied a linear ramp-up from the 2014 level to the background level over the 2015-2025 period, as suggested in Gillett et al. (2016).

2.3 Workflow and Data production

Currently, the C3PS platform provides for both, seasonal and multiannual timescales the variables requested in the DCPP/CMIP6 tables (Boer et al., 2016), which are those variables relevant for forecast evaluation against observational datasets. Besides, we have included additional biogeochemical and ocean physics variables that are necessary to force marine ecosystem models. Most of these variables are already requested by the FishMIP initiative (Tittensor et al., 2019). Higher frequency variables, such as daily ocean potential temperature and oxygen are also provided. Concerning the atmosphere, C3PS also provides daily low-level winds (~100m) and solar radiation variables as requested for renewable energy applications.

The C3PS platform follows the DCPP/CMIP6 experimental protocol with regard to the multi-annual predictions, although additional members have been performed to increase the ensemble size from 10 to 15 members.

All the C3PS related simulations were performed on the Belenos supercomputer, hosted at Météo-France site in Toulouse from June 2021 to February 2022. The work-flow is handled by the ECLIS (Environment for CLImate Simulations) package tool that was developed by the CNRM (https://www.umr-cnrm.fr/cm/spip.php?article14).
ECLIS is an ensemble of scripts and tools that allow for setting up and running all the experimental protocols performed by the CNRM-Cerfacs modelling group within CMIP and beyond. In particular, C3PS required additional ECLIS developments, such as dedicated scripts for the perturbation of initial atmospheric conditions and the management and launching of the members for all the starting dates (see section 2.4).

The C3PS diagnostics production is managed by the XIOS output server (Meurdesoif 2018). XIOS has been implemented in all the models developed by the CNRM-Cerfacs group, in particular to facilitate the huge CMIP6 data production. XIOS allows for declaring a priori the requested variables to be saved in the output files for a given experiment. Moreover, XIOS performs online operations on fields, such as spatial and vertical interpolations, vertical, spatial and time averages, vertical level extraction, thus saving a lot of post-processing time. XIOS has also been adapted to produce netCDF “CMOR” (Climate Model Output Rewriter) format files compliant with the CMIP6 Data Request specificities. More information about XIOS functioning can be found in Voldoire et al. (2019).

2.4 Seamless prediction procedure and simulations

Most of the efforts involved in the development of the C3PS platform were oriented to achieve a satisfactory initialization procedure. In this regard, several challenges needed to be tackled. The first challenge was to participate in the DCPP/CMIP6, for which the required hindcast period starts in 1960, when biogeochemical observations required to initialize the biogeochemistry model are practically non-existent. A second challenge is how to robustly initialize a seamless climate prediction platform in which a continuum of timescales need to be considered. For seasonal prediction, atmospheric initialization is relevant for climate prediction (Materia et al., 2014). For longer timescales, atmospheric initialization is less relevant as the predictability mostly lies on the ocean and sea ice persistence and memory. A third challenge is to minimize the climate drifts that occur when the model is initialized from a state away from the climate model attractor. Besides, physical coherence amongst the initial states of all the model components of CNRM-ESM2-1 is necessary in order to avoid incompatibilities that could lead to abrupt initial shocks right after the initialization (Sanchez-Gomez et al., 2016, Pohlmann et al., 2017, Bilbao et al., 2021). Although model drifts in climate prediction systems are partly corrected before skill assessment, it is preferable to minimize them as much as possible to better distinguish the predictable signals (Meehl et al., 2022).

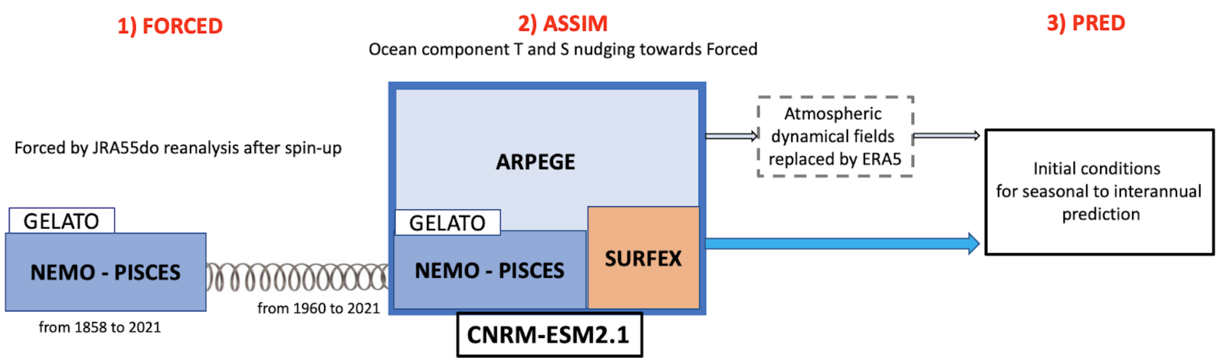

Figure 1. Schematics of the initialization procedure of the C3PS platform.

In order to overcome the three main challenges mentioned above, in the development of C3PS we have implemented an experimental protocol which is carried out in three main steps (Figure 1).

- Step 1: pseudo-observations are obtained through an ocean forced simulation in which the NEMO-PISCESv2gas model is forced by atmospheric fields from the JRA55do reanalysis (Tsujino et al., 2018) over the period 1960-2021 (Figure 1). This simulation (referred to as *FORCED* hereinafter) has been performed under the framework of the Global Carbon Project (GCP) (Hauck et al., 2020; Friedlingstein et al., 2022). The FORCED experiment was launched after a spin-up of 300 years in which the NEMO-PISCESv2gas model was forced by repeated cycles of 5 years corresponding to the 1958-1962 period. The analysis of this spin-up reveals that surface physical fields such as sea surface temperature (SST) and salinity (SSS), or in integrated fields such as ocean heat content (OHC) and Atlantic Meridional Overturning Circulation (AMOC) are almost stabilized after the spin-up.
- Step 2: the 3D potential temperature and salinity fields issued from the FORCED simulation were used to constrain the ocean component of CNRM-ESM2.1 through a Newtonian damping procedure (Figure 1). This nudging simulation is performed over the period 1960-2021, and serves to generate the so-called dcppA-assim experiment according to the DCPP/CMIP6 experiment-id (Boer et al., 2016). The dcppA-assim (referred to ASSIM hereafter) can be considered as an in-house zero-order reanalysis product from which the initial conditions for all the components of CNRM-ESM2.1 are issued. The methodology of the nudging was previously implemented and used in Sanchez-Gomez et al. 2016 for generating initial conditions for the decadal predictions in CMIP5. It was shown to be beneficial to: i) produce initial states physically consistent amongst all the components of CNRM-ESM2.1, ii) to get initial states for the components with non-available observations and iii) to minimize the initial shock and drift in the prediction experiments. Here we use the same nudging strategy which consists in 1) a sea surface restoring of temperature and salinity of the NEMO component towards SST and SSS from the FORCED simulation; 2) a 3D Newtonian damping in temperature and salinity below the mixed layer to constraint the ocean subsurface towards FORCED. The sea surface restoring is applied globally in terms of heat and freshwater fluxes. The values of the restoring coefficients are -40 $Wm^{-2}K^{-1}$ and -864 mmd^{-1} for the heat and freshwater fluxes respectively. Note that the value of the coefficient for freshwater flux significantly differs for those used in previous studies (Servonnat et al., 2015, Sanchez-Gomez et al., 2016, Bilbao et al., 2021). The rationale of this is to have the same restoring time scale for SST and SSS, that is 60 days for a mixed layer of 50m (Barnier et al., 1995). The 3D Newtonian damping is applied as follows: On the vertical, there is no damping above the mixed layer to allow for physical coherence between the mixed layer and the surface processes. Below the thermocline down to 800 m depth, the damping term is set to 10 days and for the deep ocean below, a weak damping is used (~one year). Horizontally, subsurface nudging is only applied *outside* the 15°S–15°N latitudinal band and from 300 km off the coast to avoid spurious vertical currents at the equator and coastal effects respectively (Sanchez-Gomez et al., 2016). A buffer zone of 5^o is

considered between the nudged zones and the rest of the ocean. Similar nudging methodologies are also adopted in Bilbao et al. (2021) in order to obtain initial states for seasonal and decadal predictions. The ASSIM simulation has been duplicated with a set of perturbed parameters in order to obtain an ensemble of 3 members. For this, the ocean and atmosphere diffusivity have been slightly perturbed separately to produce additional ASSIM members (see Figure 1).

- Step 3: the ASSIM ensemble will be used as initial conditions for all the CNRM-ESM2.1 components for both seasonal and interannual predictions. Only the atmospheric restarts provided by ASSIM are modified in order to adapt C3PS to seasonal forecasting. For this purpose, the dynamical fields contained in the restarts of ARPEGE in the ASSIM ensemble are replaced by the dynamical fields provided by the ERA5 reanalysis (Hersbach et al., 2020). Finally, the prediction procedure is performed as follows: For the seasonal timescale, two initializations per year are considered, that is, 1st May and 1st November. For each start date, an ensemble of 30 members is generated. The atmosphere is perturbed by using a small increment of the atmospheric dynamical fields provided by ERA5. This increment, introduced only at the initialization time, is drawn randomly from a set of increments computed during a previous historical atmospheric nudging simulation where the ARPEGE model is weakly constrained towards the ERA5 reanalysis (Batté and Déqué 2012). Ten increments were used for each ASSIM member, thus building a 30-member ensemble. Seasonal predictions starting 1st May are run for 6 months. For the multi-annual timescale, the perturbation procedure is identical to that of the seasonal scale, except that only the forecast starting on 1st November is continued up to 5 years, and with only 15 members. Hereinafter the set of seasonal to multiannual predictions will be referred to as PRED.

3 Datasets and methods to assess C3PS performances

3.1 Reference datasets for verification

Several observational and pseudo-observational products have been used to evaluate the ASSIM reconstruction (section 4) and to compute the forecast skill scores for PRED (section 5).

The physical variables we have considered are: air temperature at 2m (SAT), ocean temperature and salinity, Ocean Heat Content (OHC), Arctic sea ice concentration (SIC) and extent (SIE) and Atlantic Meridional Overturning Circulation (AMOC). To evaluate ASSIM sea surface temperature, we use a blended product consisting of an average of the Hadley Centre Sea Ice and Sea Surface temperature version 1 (HadISST1, Rayner, 2003) and ERSST v5 (Huang et al., 2017) over ice-free sea water. Over land and over sea-ice we average BEST (Muller, Curry, et al., 2013; Muller, Rohde, et al., 2013), CRU-TS4-00 (Harris et al., 2014), and GHCN-CAMS

(Fan and van den Dool, 2008). Note that to compute skill scores for SAT we have used the 2m temperature from JRA55do reanalysis from 1960-onwards (Tsujino et al., 2018).

The latest EN4 objective analysis product is used as a reference for 3D ocean temperature and salinity (Good et al., 2013) and for OHC computation. This is a $1^o x 1^o$ gridded dataset derived from ocean and temperature profiles with quality checks, which runs from 1900 to present. Here we have considered the EN4 analyses with the Gouretski and Reseghetti (2010) bias correction. SST and SIC reference data are issued from Hadley Centre Sea Ice and Sea Surface temperature v4 dataset (HadISSTv4, Kennedy et al., 2017), which combines satellite and in-situ data to provide global picture of the ocean surface over a regular 0.25^o grid for the period 1850 onwards. We use the RAPID time series of the AMOC measured at 26^oN as reference data (Moat et al., 2022), which are available from 2004 to 2022.

To analyse biogeochemistry, we focus on surface chlorophyll, integrated net primary production and global (land and ocean) carbon fluxes. Monthly means of *chlorophyll-a* concentration with a spatial resolution of 1° were issued from the ESA Ocean Colour Climate Change Initiative (ESA-OC-CCIv3.1) project (Valente et al., 2022, https://climate.esa.int/en/projects/ocean-colour/). Net primary production (NPP) was obtained using a spectrally resolved model to simulate changes in photosynthesis as a function of irradiance (Kulk et al., 2020). This model incorporates vertical structure in *chlorophyll-a* concentration from OC-CCIv4.1. NPP data are 1° gridded and are available for the period 1998-2021. Carbon fluxes are evaluated by using the Global Carbon Project (GCP) reconstruction between 1959 and 2021 (Friedlingstein et al., 2022). This reconstruction currently represents the best estimates of the global carbon sink over the industrial era since 1959. For the ocean carbon sink ($f CO_2$) we use the Surface Ocean CO_2 Atlas version 2022 (SOCATv2022; Bakker et al., 2022) for the period 1990–2021.

To evaluate ASSIM reconstruction, besides the observational and analysis products described above, we consider the FORCED simulation, and the historical experiment performed with CNRM-ESM2.1 for CMIP6/DECK (Séférian et al., 2019) and referred here as FREE, which represents the free model (no data assimilation) run.

3.2 Metrics for skill assessment

The pseudo full-field initialization strategy used in C3PS requires to remove the forecast drift that inevitably occurs in any climate prediction system before performing the verification with observations and the skill estimate. We use the standard approach of transforming the raw model data into anomalies relative to the climatological forecast for each lead time.

$$X'_{j,l} = X_{j,l} - X_l \qquad (1)$$

Where $X_{j,l}$ represents the ensemble-mean forecast from starting date *j* at lead time *l* and X_l is the average over these forecasts over all starting dates for a given lead time. This is the so-called *mean drift correction* method, which assumes that forecast drift does not depend on the background climate state, i.e. the drift is not considered to change between two different climate states from the point of view of global warming (Garcia Serrano and Doblas-Reyes 2012, Meehl et al., 2014). Note that for the forecast period 1960-2021, the number of starting dates is 62 x 2 for the seasonal, and 62 for the interannual timescales. In the case of interannual forecasts, starting on 1st November each year, we focus our analysis on the following 5 years beginning in January (2 months after the initialization).

For both seasonal and interannual timescales we use the standard verification framework as outlined in Goddard et al. (2013). We rely on the anomaly correlation coefficient (ACC), root mean square error (RMSE) and the Mean Square Skill Score (MSSS). The MSSS is especially used to assess the added value of the initialization and it is computed following equations 4-6 from Goddard et al. 2013. A MSSS score greater than 0 means that PRED is more accurate than FREE. For the seasonal forecast, persistence scores are used as a benchmark of C3PS scores and a t-test for assessing the statistical significance of the correlation.

According to Goddard et al. (2013), for the skill maps and in order to remove small-scale unpredictable noise, all model and observational data are interpolated to a common 5-degree regular grid using the ESMF patch interpolation included in the NCAR command language – NCL.

We assess the added value of the initialization in C3PS by comparing the hindcasts PRED and the non-initialized historical ensemble (FREE) against FORCED or JRA55do for atmospheric variables. To properly evaluate skill differences between PRED and FREE, either through ACC or MSSS, a non-parametric bootstrap technique ia used to assess the statistical significance of the skill scores (Goddard et al., 2013; Yeager et al., 2018). A block-bootstrap distribution of the scores is constructed at each location (grid point or time series) by resampling (with replacement) pairs of observations and hindcasts across the time dimension, and in addition, the PRED and FREE ensembles across the ensemble member dimension. Following these previous papers, we use a block size of 6 years as a trade-off between autocorrelation of the physical variables and the number of blocks (results are very similar to those based on 5 or 7-year blocks). The derived p-values are estimated as in Yeager et al. (2018).

Finally, the hindcast performance is evaluated by considering the so-called “potential predictability”, which consists in using as reference dataset the ASSIM experiment (Yeager et al., 2022). The skill calculated with respect to ASSIM is the maximum that C3PS can achieve. The notion of “potential predictability” is also interesting to assess forecast performance for biogeochemistry, since observations are available over a short period of time. We will compare "potential predictability" versus "effective predictability", the latter being estimated considering FORCED or JRA55do reanalysis as reference.

4 Basic evaluation of the C3PS initialization procedure

The assessment of the C3PS initialization strategy aims to determine how far the nudging of the ocean physics has affected the performance of CNRM-ESM2.1 at simulating relevant physical and biogeochemistry fields.

Figure 2 shows that FREE exhibits common coupled model biases in the North Atlantic Ocean - the so-called "blue spot" - the southeastern Tropical Atlantic along the Benguela coast and the equatorial Pacific cold tongue. Those biases are reduced in ASSIM, as expected. Over land, ASSIM and FREE do not differ much in terms of biases, though over some regions like North America, Northern Africa temperature biases are slightly reduced. This fact indicates that ocean nudging does not have much impact over the continental areas.

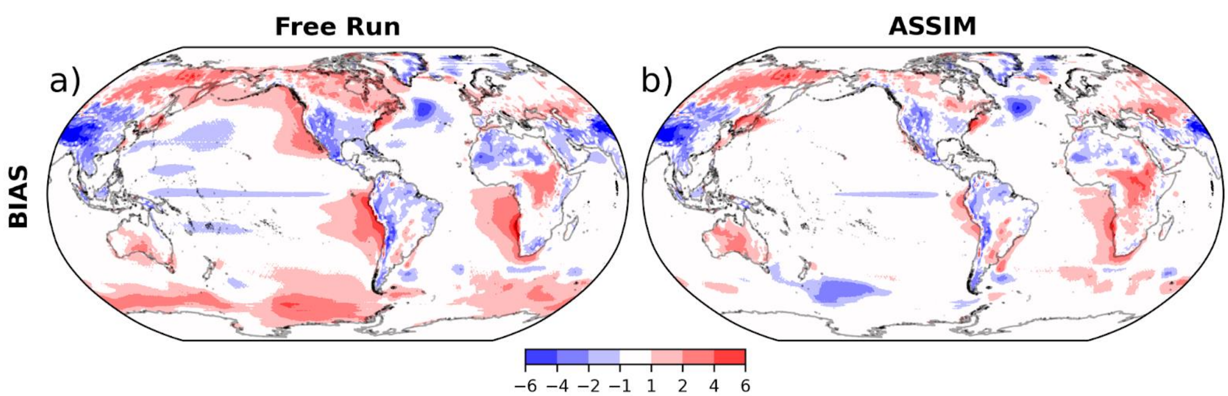

Figure 2. Departure in blended surface temperature of FREE (a) and ASSIM (b) simulations from observations over 1960–2014. Blended surface temperature combines surface-air temperature over land and sea ice and sea surface temperature over ice-free sea water. Observations average several data sets: HadISST1 (Rayner, 2003) and ERSST v5 (Huang et al., 2017) over ice-free sea water; BEST (Muller, Curry, et al., 2013; Muller, Rohde, et al., 2013), CRU-TS4-00 (Harris et al., 2014), and GHCN-CAMS (Fan & van den Dool, 2008) over land and sea ice. Units are in degrees Celsius.

Ocean temperature and salinity fields were used to nudge the ocean component of CNRM-ESM2.1 in order to generate ASSIM as explained above. Therefore, it is essential to evaluate both the performance of FORCED to simulate the mean state of the subsurface ocean as informed by observations, and then to evaluate the bias reduction of the ASSIM simulation with respect to the FREE simulation. Figure 3 displays thus the departure in ocean temperature and salinity at 100 m depth of FORCED with respect to observations, and of FREE and ASSIM with respect to FORCED. The FORCED simulation captures the main distribution of ocean temperature at the subsurface as depicted from observations. Nonetheless, FORCED overestimates temperature in the tropical Atlantic and across the North Pacific and the Southern Oceans, while it underestimates it East of New-Zealand and in the tropical Pacific. In addition, the FORCED simulation strongly underestimates temperature in the North Atlantic where a well-documented "warming hole" has been related to a persistent slowdown of the Atlantic meridional overturning circulation (Drijfhout et al., 2012; Menary et al., 2028; Swingedow et al., 2021). By contrast, salinity is underestimated in this region, over the polar region and most regions of the Pacific Ocean, as reported in Voldoire et al. (2019).

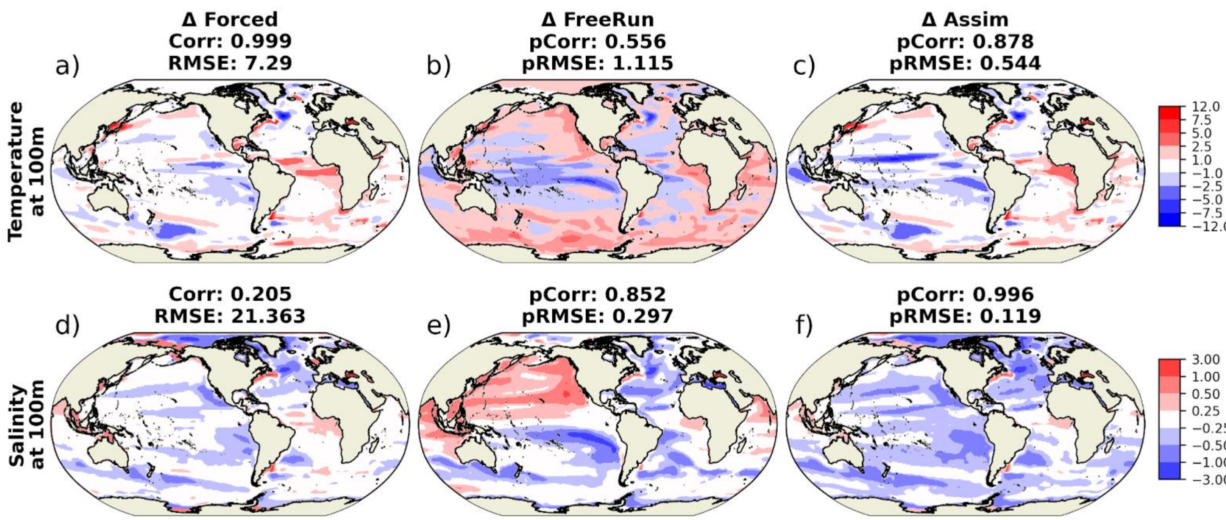

Figure 3. (a, d) Difference of ocean temperature and salinity at 100 m depth between FORCED with respect to EN4 observations over 1960-2014. Differences between FREE (b, e) and ASSIM (c, f) with respect to FORCED over the same period. Spatial correlations and RMSE of the time average over the whole period are shown on the top of each figure. Correlations and RMSE are computed against EN4 for (a, d) and against FORCED simulation for (b, c, e, f). Observations are extracted from the quality-controlled EN4 dataset (Good et al., 2013). Units are in degrees Celsius for temperature and psu for salinity.

As mentioned above, the nudging in subsurface waters is only applied in latitudes higher than ± 15º. Accordingly, as seen in Figure 3, main differences of the ASSIM simulation with respect to the FORCED simulation occur in tropical regions, where the ASSIM tends to underestimate both temperature and salinity. In contrast, the underestimation of temperature expands to the Atlantic and Pacific Oceans in the FREE simulation, while it overestimates temperature in the Southern Ocean and the California Current. The FREE simulation also underestimates salinity across most of the Atlantic and Pacific oceans, while it overestimates it across the North Pacific and the Indian oceans.

In conclusion, as expected, in both FORCED and ASSIM, the biases in surface and subsurface are strongly reduced compared to the FREE run, which confirms the validity of the methodology to generate oceanic initial conditions.

4.1 Drivers of seasonal climate variability

In order to evaluate the realism of ASSIM in accounting for ENSO variability, we focus on how the nudging procedure impacts the ENSO diversity considering that this is a fundamental ENSO property that determines its seasonal evolution and teleconnections (Capotondi et al., 2020). The ENSO diversity or complexity (Timmerman et al., 2018) refers to the existence of warm and cold events with different SST patterns and amplitudes, with the extreme warm events being of Eastern Pacific type, while moderate warm and cold events being of Central Pacific type. Although the CNRM-ESM2.1 model (FREE) has some skill in simulating ENSO feedback strength (Lee et al., 2021), it has difficulty in simulating ENSO amplitude diversity, which manifests as a negative skewness of SST anomaly in the eastern equatorial Pacific.

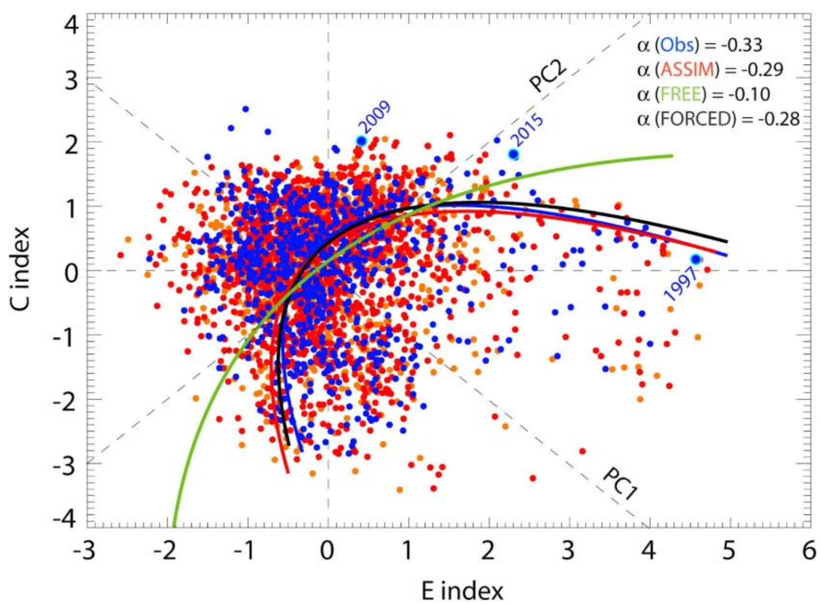

Figure 4. Phase space of the first and second principal components (PC) of monthly SST anomalies in the tropical Pacific (120°E-80°W; 11°S-11°N) with fitted quadratic curves to measure nonlinearity for observations (blue dots, from HadISST 1960-2020) and the ASSIM runs (orange and red dots). Nonlinearity is measured by fitted quadratic curves between PC time series (blue: observations, red: ASSIM, black: FORCED, red: FREE). The PC time series have been rotated by 45° to infer the E and C indices. Three different types of observed El Niño events are highlighted with light blue circles (December): 1997: Extreme Eastern Pacific El Niño, 2009: Central Pacific El Niño and 2015 mixed-type.

Here as a compact measure of ENSO diversity (or nonlinearity), we use the value of the first coefficient of a quadratic fit in the phase plane of the first and second principal components (PCs) of SST anomalies in the tropical Pacific (Karamperidou et al., 2017; Cai et al., 2018), hereafter referred to as α. For HadISSTv4 data, the two branches of this quadratic fit tend to align along axis that correspond to the PC1 and PC2 axes rotated by 45° (Figure 4). The rotation of the PC time series defines the E and C indices, with E=(PC1-PC2)/$\sqrt{2}$ and C=(PC1+PC2)/$\sqrt{2}$, that account for the variability of Eastern Pacific events and Central Pacific events (Figure 5, top), respectively (Takahashi et al., 2011). While α = -0.33 for observations, α = 0.10 ± 0.06 for FREE (the error corresponds to the standard deviation amongst the 10 members), which results from the negative ENSO asymmetry of the CNRM-ESM2.1 model (Lee et al., 2021). ASSIM has a more realistic ENSO non-linearity (α = -0.29), almost identical to FORCED (α = -0.28), indicating that the nudging procedure succeeds in restoring positive ENSO asymmetry to the observed value, along with improving ENSO diversity (see the blue curve paralleling the red curve in Figure 4). Still, ASSIM tends to have a larger ENSO variability than in the observations as evidenced by the larger amplitude of the E and C mode patterns compared to observations (Figure 5 middle), which is due to FORCED overestimating ENSO variability.

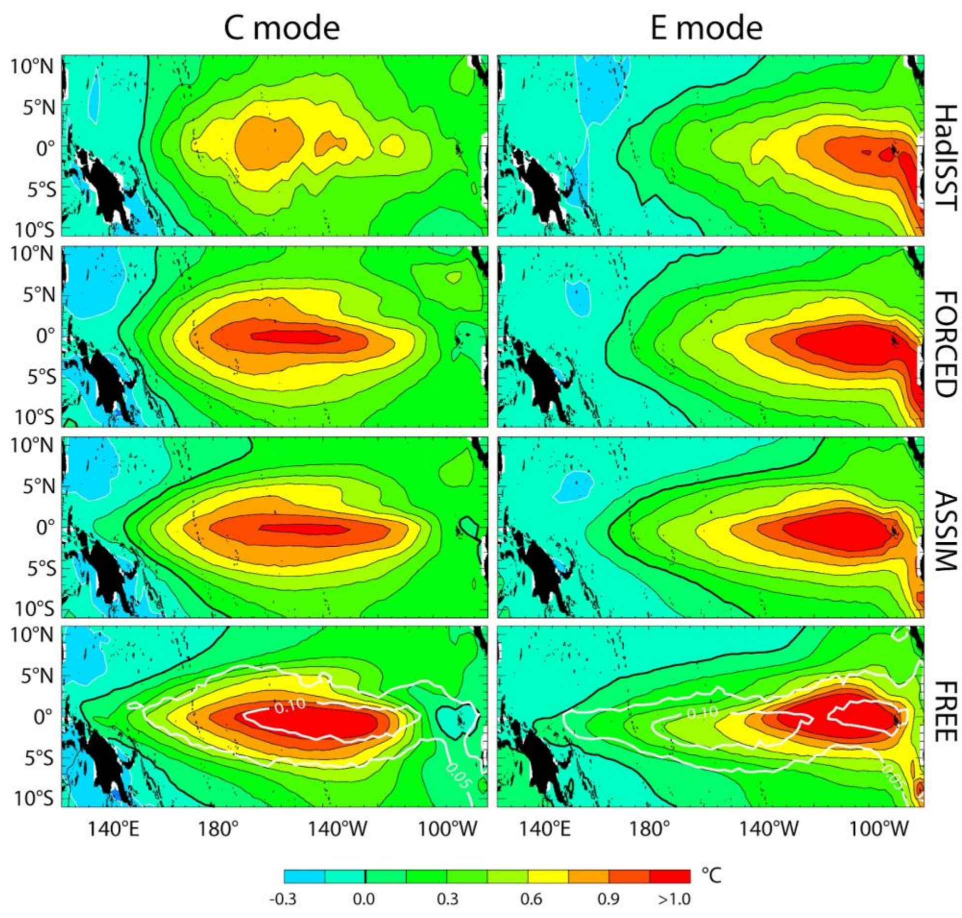

Figure 5. C (left) and E (right) mode patterns for observations (top), FORCED and ASSIM (middle panels) and FREE (bottom). Dispersion (rms amongst the ensemble) is indicated for FREE in white contours.

4.2 Drivers of interannual to decadal climate variability

To examine the drivers of interannual to decadal Pacific variability simulated by ASSIM, we focus on the Tripole Pacific Index (TPI) as defined by Henley et al. (2015). The TPI is a proxy of the Interdecadal Pacific Variability (IPV) and it is based on the difference between the SST anomalies averaged over the central equatorial Pacific minus the average of the SST anomalies in the Northwest and Southwest Pacific (see Henley et al., 2015 and Bilbao et al., 2021 for more details). Here, we do not consider SST anomalies as we are interested not only in the phase of the low frequency variability, but also in the model mean state. ASSIM and FORCED are coherent with HadISSTv4 SSTs evolution (Figure 6a), which is expected due to the sea surface restoring. The temporal correlation of ASSIM ensembles mean and FORCED with respect to HadISSTv4 is 0.92 (see Table 1). Interannual variability of TPI is underestimated by the FREE ensemble as shown by Figure 6a and the variance ratio in Table 1. The smaller amplitude of Pacific decadal variability in the CNRM-Cerfacs models was also reported in Voldoire et al. (2019), which suggests a lack of the ENSO teleconnection at decadal timescales. In terms of RMSE, ASSIM presents an improvement with respect to FREE (Table 1). Interannual variability of OHC integrated over the first 300 meters (OHC300) indicates that ASSIM is quite in phase with EN4, with a correlation value of 0.79 (Figure 6b, Table 1). ASSIM also improves the amplitude of the interannual variability with respect to FREE (see Table 1).

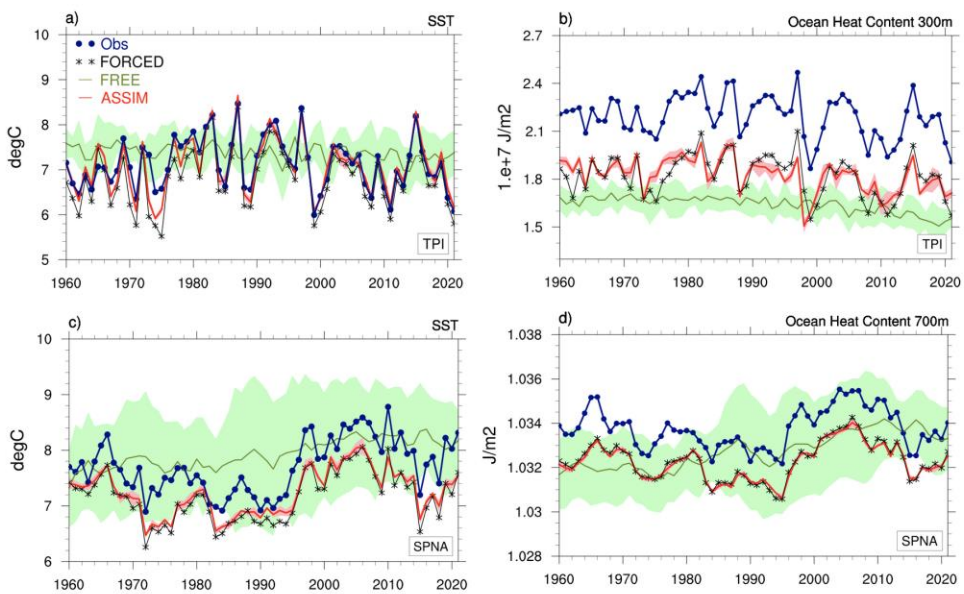

Figure 6. (a) Tripole Pacific Index (TPI) annual time series from 1960 to 2021 for the SST and **(b)** OHC integrated over the first 300m for the FREE ensemble (green), ASSIM ensemble (red), FORCED (black solid) and HadISSTv4/EN4 (blue). **(c)** Subpolar North Atlantic **(**SPNA) index annual time series from 1960 to 2021 for the SST and **(d)** OHC integrated over the first 700m for the same experiments. For the Ocean Heat Content the observational reference is EN4. The TPI index is computed from raw data according to Henley et al. 2015. The SPNA index from raw data is obtained according to Bilbao et al. 2021 (SPNA: 50–65°N, 60–10°W). For FREE and ASSIM the ensemble means (thick line) and plus/minus one standard inter-members deviation is shown (red and green shading).

	TPI (SST)			TPI(OHC300)			SPNA(SST)			SPNA(OHC700)		
	FORCED	*ASSIM*	*FREE*	*FORCED*	*ASSIM*	*FREE*	*FORCED*	*ASSIM*	*FREE*	*FORCED*	*ASSIM*	*FREE*
Correlation	0.92	0.92	0.01	0.95	0.79	0.14	0.93	0.90	0.16	0.93	0.91	0.16
Variance ratio	1.17	1.06	0.83	0.96	0.83	0.71	0.91	0.84	1.30	0.95	0.90	1.57
RMSE	0.39	0.25	0.79	0.37	0.37	0.57	0.58	0.52	0.99	0.01	0.01	0.01

Table 1. Performance metrics (correlation, variance ratio and RMSE) computed with respect to the observational references and the different experiments: FORCED, ASSIM and FREE. The time series used to compute the metrics are displayed in Figure 6. The values shown for ASSIM and FREE are the ensemble mean of the values computed for each individual member.

Regarding the OHC300 mean state, ASSIM and FORCED exhibit a cold bias which is weaker than in FREE (Table 1, RMSE). This cold bias of CNRM-ESM2.1 is also present in the coupled ocean-atmosphere climate model CNRM-CM6.1 (Voldoire et al. 2019). In general, like most coupled models, CNRM-Cerfacs models show a cold temperature bias in the Pacific Ocean from the surface to around 300m depth. This cold bias is suggested to be caused by too strong surface winds curl exerting a pronounced wind curl into the ocean (see also Figure 3c). From Figure 6b, ASSIM mean state lies in between the reference data EN4 and FREE, indicating that initializing the ocean component of CNRM-ESM2.1 from ASSIM could potentially reduce the model drift in the predictions, which is actually the scope of our initialization procedure.

Another driver of interannual to decadal ocean variability is the Atlantic Multidecadal Variability (AMV). It was shown that CNRM-Cerfacs models simulate quite well the AMV spatial pattern with regards to observations (Voldoire et al., 2019). Here we analyze the Subpolar Gyre in the North Atlantic (SPNA), which is closely correlated to the AMV. The SPNA SST time series (Figure 6c) exhibits a high temporal correlation in ASSIM and FORCED versus HadISSTv4, which is 0.90 and 0.93 respectively (see Table 1). Moreover, observations lie within the FREE multi-member spread, indicating that in terms of mean state, the free model performs quite well for this area. Note that the members of FREE show a pronounced variability, as also indicated in the variance ratio in Table 1. Indeed, the models CNRM-CM6.1 and CNRM-ESM2.1 are characterized by a large SST variance over the SPNA at decadal timescales, which is strongly correlated to AMOC variations, Arctic freshwater flux balance and northward salt transports from the tropical area (Voldoire et al., 2019). The marked decadal variability in FREE is also visible in the OHC integrated over the first 700m (Figure 6d and Table 1). Once again, the correlation of ASSIM with regards to EN4 (0.91) indicates a good temporal coherency in the ocean subsurface.

As mentioned above, AMOC variations simulated by CNRM-ESM2.1 are highly correlated to decadal variability over the SPNA and Northern Seas (Voldoire et al., 2019). Time series of maximum AMOC at 26^{o}N show a large low frequency variability in the members of FREE (Figure 7a), previously documented in Séférian et al. (2019) and Waldman et al. (2021). The mean AMOC value at 26^{o}N of FREE is 16.4 ± 2.3 Sv for the period 1960-2014. The uncertainty in the latter value is estimated by considering one standard deviation amongst the members of the ensemble. The FREE ensemble AMOC is in good agreement with the RAPID mean value of 16.8 Sv for the observed period. Moreover, the depth of maximum observed AMOC is well simulated by FREE (Figure 7b). FORCED and ASSIM show a weaker AMOC (Figure 7a-b), with mean values of 12.2 ± 0.7 and 12.6 ± 0.8 respectively. The GCP experimental protocol used to perform FORCED is quite similar to those proposed in OMIP2/CMIP6 (Tsujino et al., 2020). The latter study documents that, in general, the forced ocean simulations show a lower AMOC intensity compared to RAPID. This underestimation of the AMOC is even more pronounced in the NEMO3.6/GELATO forced model configurations, suggesting that coupling with the atmosphere plays an important role in this high variability and intensity of the AMOC in the CNRM-Cerfacs models. The nudging of the temperature and salinity constraints to some extent the AMOC in ASSIM, whose correlation is 0.60 with respect to FORCED.

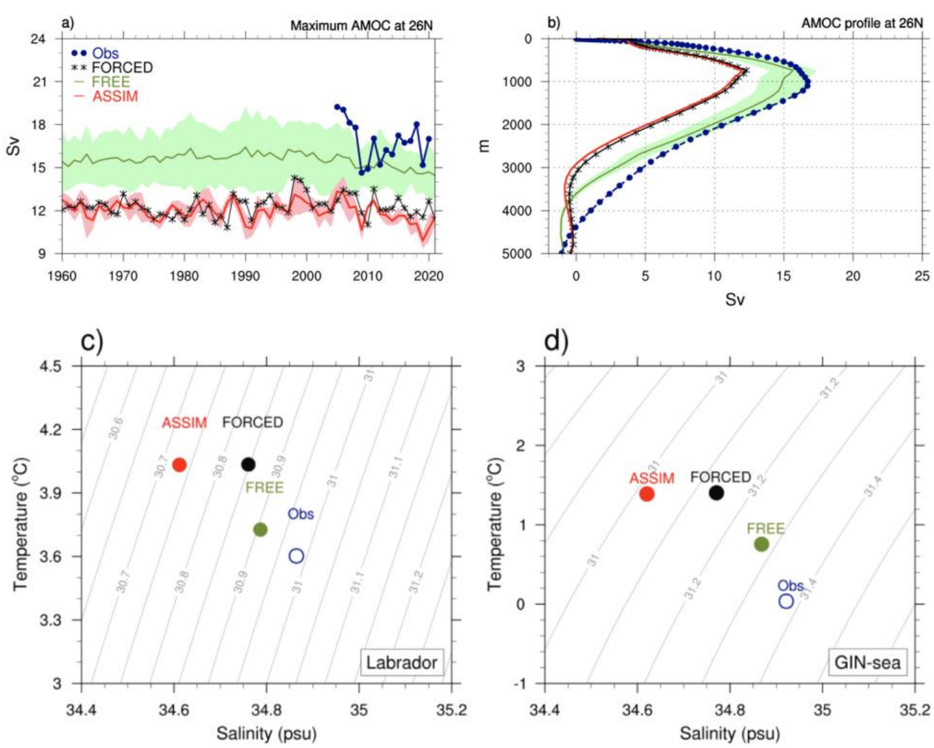

Figure 7. (a) Time series of the maximum AMOC at 26^{o}N for the FREE ensemble (green), ASSIM ensemble (red), FORCED (black solid) and RAPID data (blue). Units in Sv. **(b)** Vertical profile of AMOC at 26awa^{o}N for the FREE ensemble (green), ASSIM ensemble (red), FORCED (black solid) and RAPID data (blue). Units in Sv. For (a) and (b) the FREE and ASSIM the ensemble means are shown (thick line) together with plus/minus one standard inter-members deviation (shading). **(c)** Temperature-Salinity diagram over the Labrador Sea area (70^{o}W-45^{o}W, 50^{o}N-

68°N) at 700m depth for the FREE ensemble (green), ASSIM ensemble (red), FORCED (black solid) and EN4 data (blue). (d) The same as (c) but for the GIN-Sea area (25°W-10°E, 65°N-80°N). Only ensemble means are shown for FREE and ASSIM. Potential density is computed from the NCL function "rho_mwjf".

The reason of the AMOC underestimation of ASSIM and FORCED can be partially explained by less dense subsurface waters of ASSIM and FORCED compared with FREE over the deep convection areas, i.e. Labrador and GIN seas (Figure 7cd). These differences of density are mainly explained by warmer and less salty waters in FORCED and ASSIM, which are less realistic than those of FREE. The impact of the T/S nudging of CNRM-ESM2.1 towards FORCED seems to affect freshwater fluxes over the Labrador and GIN-Sea regions, since ASSIM is less salty than FORCED. The AMOC and related ocean deep convection characteristics in the ASSIM simulations are consistent with regional features of the Arctic Sea Ice Concentration (SIC) climatology (not shown). Indeed, FREE presents a more extended sea ice area with regards ASSIM over the marginal seas in winter, which is coherent with colder and saltier waters over the Labrador and GIN-Seas.

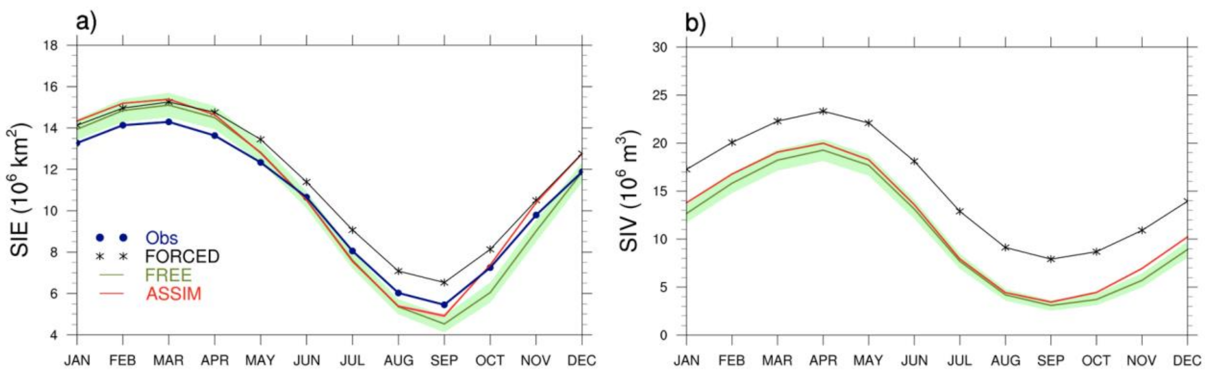

Figure 8. Seasonal cycles of the Arctic Sea Ice Extension (SIE) (a) and Volume (b) computed in the period 1960 to 2021 for the FREE (green), ASSIM (red), FORCED (black solid) and HadISSTv4 (blue). For FREE and ASSIM the ensemble means (thick line) and plus/minus one standard inter-members deviation is shown (shading).

Annual cycle of SIE and SIV shows that ASSIM is comparable to the FREE ensemble (Figure 8ab), except from October to December where ASSIM performs better than FREE. In general, FORCED, ASSIM and FREE overestimates the maximum Arctic SIE which is connected to a too cold mean state with respect to HadISSTv4 (Figure 8a). The SIV simulated by ASSIM overlaps the FREE climatology, indicating a weak control of the nudging on the volume. The correlations between ASSIM and FORCED interannual time-series of SIE are 0.82 and 0.36 for March and September respectively (not shown). The high correlation indicates that the nudging largely constrains the SIE in the ASSIM ensemble. Less control of nudging is shown for SIV, as indicated by the correlation coefficient of 0.45 and 0.26 between ASSIM and FORCED for the climatological maximum and minimum. Our results show that CNRM-ESM2.1 will be initialized from sea ice conditions close to FREE, which could be beneficial for sea ice drift, which may exert detrimental effects on predictability in the SPNA zone (Huang et al., 2015, Bilbao et al., 2021).

4.3 Biogeochemistry

The biases of both ocean surface chlorophyll maximum and minimum show that the FORCED simulation has difficulties in representing surface chlorophyll patterns (Figure 9a-d). In general, the FORCED simulation underestimates the maximum chlorophyll values in the North Atlantic and North Pacific oceans, while it overestimates both maximum and minimum chlorophyll observations in both the Pacific and Southern oceans. The FORCED simulation also overestimates observations in the North Atlantic. The FREE biases with respect to the FORCED simulation are stronger over the western boundaries in the northern oceans and over the Southern Ocean for chlorophyll minimum (Figure 9b-e). The difficulties of CNRM-ESM2.1 to represent surface chlorophyll over the Southern Ocean were documented in Séférian et al. (2019), and related to erroneous phytoplankton growth representation over the high-nutrients areas. Coastal chlorophyll biases were explained by deficiencies in remote-sensing products to represent coastal concentrations of surface chlorophyll (e.g. Gregg and Casey, 2004).

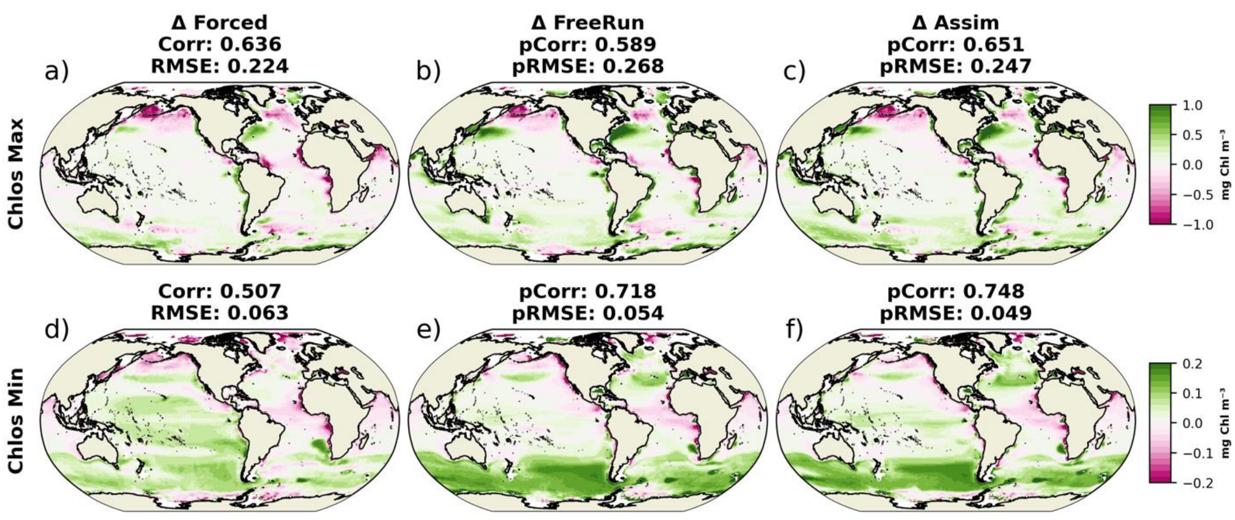

Figure 9. (a, d) Difference of ocean surface chlorophyll maximum (top panels) and minimum (bottom panels) between FORCED with respect to ESA-OC-CCIv3.1 observations over the period 1998 to 2017. Differences between FREE (b, e) and ASSIM (c, f) with respect to FORCED over the same period. Global average spatial correlations and RMSE are shown on the top of each figure. Correlations and RMSE are computed against WOA2018 for (a, d) and against FORCED simulation for (b, c, e, f). Surface chlorophyll maximum corresponds to the average over the months March, April, and May. Surface chlorophyll maximum corresponds to the average over

the months August, September, and October. Observations correspond to monthly climatological data extracted from the quality-controlled 1º resolution ESA-OC-CCIv3.1 dataset (Valente et al., 2022).

ASSIM biases with respect to the FORCED simulation are still similar to those shown for the FREE simulation (Figure 9c-f). A lower RMSE and higher pattern correlation quantitatively indicate that ASSIM deviations from FORCED are smaller, pointing at a marginal impact of the sea surface restoring and 3D nudging. However, the nudging applied to temperature and salinity does not improve chlorophyll concentrations because it fails at improving distribution of nutrients in most regions. Indeed, an analysis on the biases of both surface nitrate (NO3) concentrations and mixed layer depth (MLD) between FREE and ASSIM with respect to WOA2018 climatology (Supplementary Figure S1), suggests that the biases in NO3 are too strong to be compensated by the nudging on ocean physics. Moreover, in both the Southern Ocean and the North Atlantic, an underestimation of the MLD together with an overestimation of NO3 may explain the consistent overestimation of surface chlorophyll minimum in those regions. Nutrient-rich waters that concentrate within a shallow MLD will strengthen the excessive development of phytoplankton. Phytoplankton growth in these regions will become limited by light availability, which explains why the overestimation of surface chlorophyll maximum is not as high as for surface chlorophyll minimum, especially over the North Atlantic.

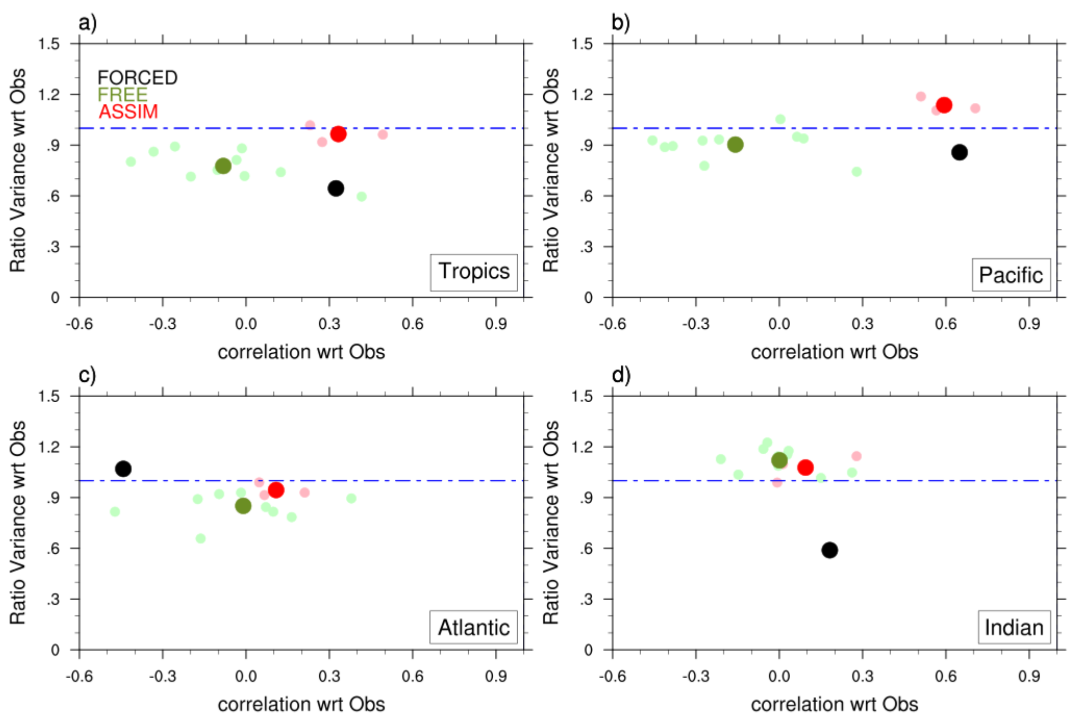

Figure 10. Scatterplot of the variance ratio experiment/observation versus temporal correlation (experiment/observations) for the integrated primary productivity averaged over the (a) whole tropical regions (30S-30N), (b) the tropical Pacific, (c) the tropical Atlantic and (d) the Indian Ocean for the FREE ensemble (green), ASSIM ensemble (red), FORCED (black solid). Blue dashed line indicates a perfect match (=1) for the variance ratio. The observational reference is issued from Kulk et al., 2016 dataset for the available period 1998-2018.

The impact of the nudging on NPP is diagnosed over the Tropical oceans in terms of interannual variability and temporal coherence with respect to observational estimates. Figure 10a shows that the nudging leads to an improvement of simulated interannual variance and temporal coherence with observations for ASSIM with respect to FREE in the tropical band. This improvement comes from the Tropical Pacific (figure 10b), for which the correlation between ASSIM and observations is high (around 0.6). The fact that the sea surface restoring improves the phasing of the NPP interannual variability of CNRM-ESM2.1 with respect to NPP observational estimates over the Pacific Ocean was also documented by Séférian et al. (2014). The SST restoring induces an improvement of SST gradients and short-term dynamical adjustment of winds, which combined with a good representation of nutrients over the area by CNRM-ESM2.1 can lead to a better simulated NPP. Contrary to the Pacific, oceanic nudging does not induce a clear impact in the NPP representation in the other tropical basins. In the Atlantic and the Indian oceans, ASSIM and FREE results are very similar.

The impact of the nudging on the ocean carbon sink is assessed in terms of trends and variability in Figure 11. Figure 11a shows that FREE and ASSIM simulations capture the long-term increase of the global carbon sink as shown by the Global Carbon Project (GCP) reconstruction between 1990 and 2021 (Friedlingstein et al., 2022). Interestingly, both FREE and ASSIM capture the strengthening of the ocean carbon sink over the recent years, whereas GCP models do not. Nonetheless, it is difficult to identify an impact of the nudging on the simulated trends in ocean carbon sink in ASSIM with respect to FREE. In particular, the nudging does not improve the representation of the decadal swing of the ocean carbon sink observed before and after the 2000s. Indeed, all models' configurations fail at capturing the slowdown of the ocean carbon sink in the 2000s, including ASSIM. Yet, models display a better agreement with the data-product displaying a weaker variability.

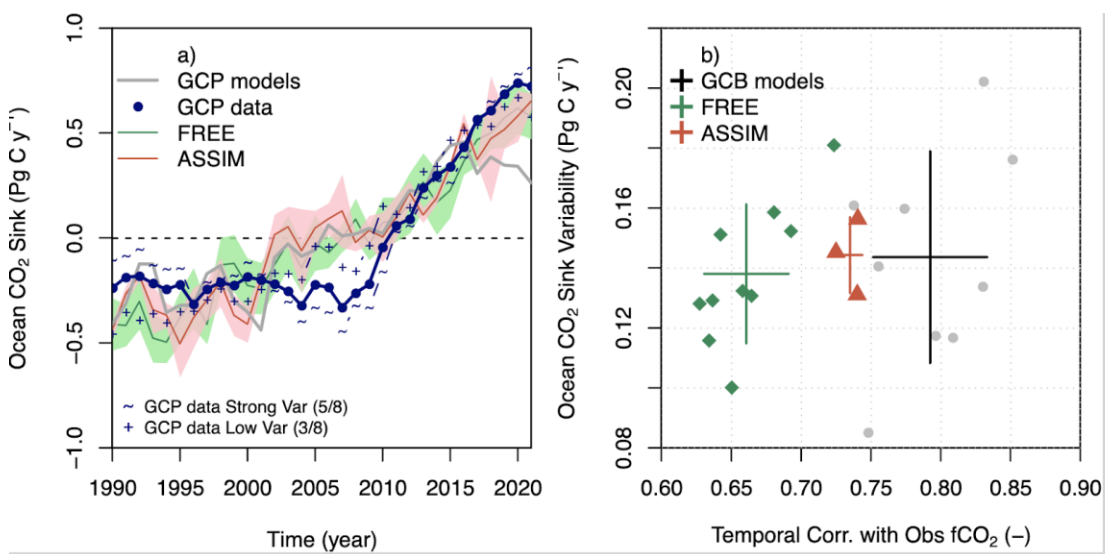

Figure 11. a) Annual time-series of the ocean carbon sink from 1990 to 2021 for FREE and ASSIM ensembles and GCP data product (Friedlingstein et al. 2022). The ocean carbon sink is represented in anomaly with respect to the long-term mean over the 1990-2021 period.. The ensemble mean of available GCP ocean biogeochemical models and observational data products are given in gray and dark blue. For the sake of discussion, the ensemble of the 8 available data-products is splitted in two sub-ensemble characterized by either a stronger (GCP data Strong Var, +) and lower (GCP data Low Var, ~) variability than the ensemble mean. b) Scatter plot comparing model properties in terms of variability of the ocean carbon sink (y-axis) and the chronology of the ocean CO2 fugacity (fCO2) over the 1990-2021 period is provided for individual realization of FREE (green), ASSIM (red) and GCP models (gray). The ensemble average is given by the green, red and black crosses for FREE, ASSIM and GCP models.

Figure 11b helps to identify the added value of the nudging by scrutinizing its impact on the simulated variability in terms of magnitude and chronology. The nudging improves the consistency between modelled and observed chronology in ocean fugacity, and slightly reinforces the magnitude of the ocean carbon sink variability. This improvement is due to the fact that fCO_2 is driven by changes in temperature and salinity in the ocean, which are directly impacted by the nudging approach. Although small, the improvement in the modelled chronology of the ocean carbon sink variability has the potential to improve the capability of the model to predict year-to-year variation in ocean carbon sink.

5 Skill assessment of key climate and biogeochemical fields

5.1 Seasonal timescale

ENSO diversity is considered to assess forecast performance of C3PS considering that central and eastern equatorial Pacific variability modes convey different tropical teleconnections outside the tropical Pacific. For that, the forecast members are projected on the spatial patterns of the two ENSO modes shown in Figure 5 to obtain the E and C indices. As a reminder the E and C indices are uncorrelated by construction. ACC values for the start date of 1st November show very high and significant scores for all leadtimes (Figure 12a,b). C3PS performs better than persistence for leadtimes greater than 6 months (i.e summer after the initialization) for the E-mode and for all leadtimes for the C-mode. C3PS is more skillful at predicting central Pacific ENSO variability than eastern Pacific ENSO variability, which results from difficulty in predicting strong El Nino events that are of E type, a common feature of seasonal prediction systems (L'Heureux et al., 2020). In general, the central equatorial Pacific is more predictable than the eastern edge, where ENSO-related phenomena involve a sharp change in convective regime and non-linear oceanic processes, resulting in a strong positive skewness of the E index (Takahashi et al., 2011). Most coupled models, like CNRM-ESM2.1, have also a warm bias in the far eastern Pacific that is influential on the forecasts (L'Heureux et al., 2022).

C3PS is rather effective since the predictive skill levels remain high for almost one year after initialization. Potential predictability is slightly higher, the difference with ACC computed from observations increases at longer leadtimes. We have checked that the C3PS performances at predicting ENSO during the period 1960-2021 are comparable to the current seasonal predictions systems such as SEAS5-20C (Weisheimer et al., 2021; Sharmila et al., 2022) (not shown). RMSE scores, which take into account the prediction of the ENSO amplitude, beat persistence scores for longer leadtimes (Figure 12cd). Again, potential predictability is above, in particular for ENSO-C.

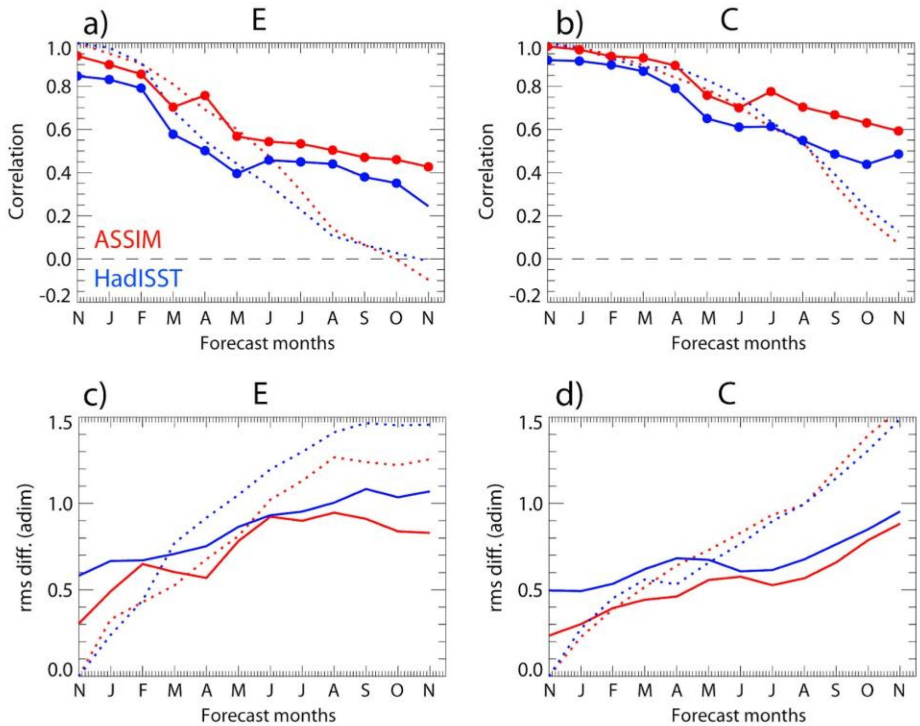

Figure 12. ENSO seasonal forecast skill: (a, b) ACC skill and (c, d) rms error for the ensemble-mean as a function of leadtime for the E and C indices over the period 1960-2021 for the initialization in 1st November and compared to persistence forecasts (dotted line). Red is for ASSIM as the benchmark data (i.e. potential predictability) and the blue is for HadISST data as the reference. Dots indicate where the correlation is significant at the 95% level based on a t test.

ENSO skill for 1st May starting date presents ACC values above persistence from 2 months onwards leadtimes. Again C3PS achieves better performance for ENSO-C mode. Since boreal spring (the season of initialization) corresponds to that of the ENSO onset and the usually enhanced Madden-Julian oscillation variance, we may expect that the system has also good performance in predicting the tropical Pacific teleconnection at that season. Potential predictability is higher than “effective” predictability by about 0.1 correlation value on average for all leadtimes.

5.2 Multi-annual timescale

Skill maps of ACC computed between PRED and FORCED for surface temperature show high and significant skill over large portions of the globe for lead times Y1, Y2 and Y1-5 (Figure 14, left column). ACC scores are usually higher in the tropics than in the extra-tropics. For Y2, the skill rapidly decreases over the Eastern Pacific and the Southern Ocean but remains high and significant over the North Atlantic and Indian Oceans, Europe, Northern Asia, Northern Africa, North America and some areas in South America. Considering the temporal average over the five forecast times, the ACC skill is considerably high and significant over a great portion of the Northern Hemisphere and the Indian Ocean. Potential predictability, measured by the ACC between PRED and ASSIM, is clearly higher (middle column) in many regions of the globe, including most of the continental areas, except India.

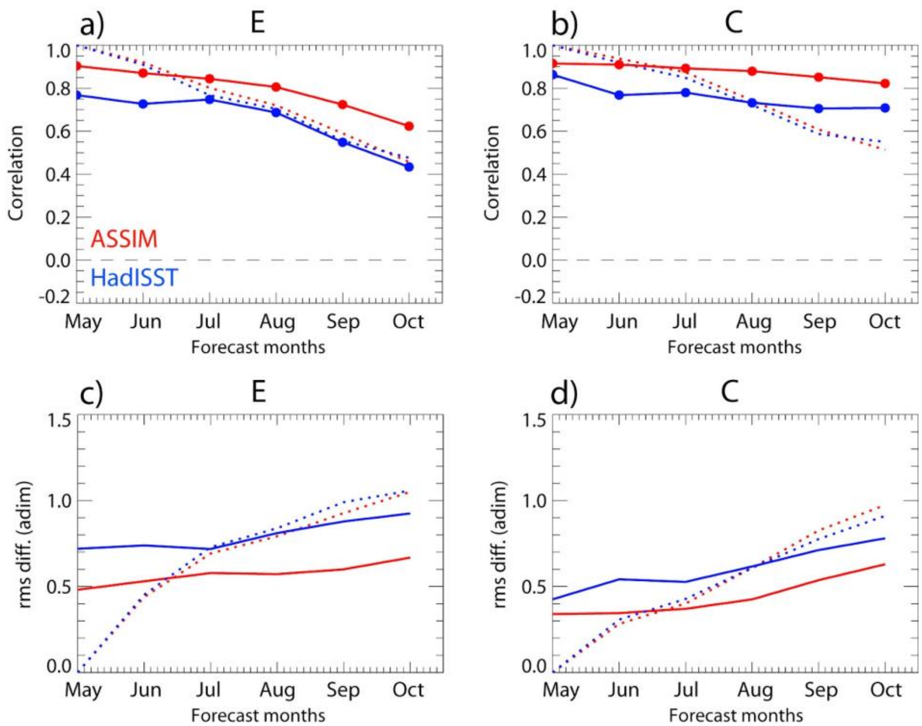

Figure 13. Same as Figure 12 but for the initialization in May (30 members, period: 1960-2018).

When we compare PRED and FREE skills in terms of potential predictability (Figure 14, right column), results show that some regions exhibit larger skill scores in PRED at Y1, indicating that initialization largely improves ACC scores in most of the Pacific Ocean, SPNA, western tropical Atlantic and northern South-America, central Indian Ocean and eastern Australia. In general, from Y2 onwards much of the skill is provided by the large externally-forced trend as shown by the similarity between PRED and FREE skill scores. The regions where initialization still plays an important role are the SPNA, Equatorial Pacific, Southern Pacific and Indian Oceans, as well as over North America and Brazil. At longer timescales, the added value of initialisation remains over the SPNA and Southern Pacific. The fact that one of the areas of clear benefits of the ocean initialisation is the SPNA is consistent with the results reported by current decadal prediction systems (IPCC, 2023).

The fact that the predictive skill is high in the mid-latitudes over the Pacific at Y1 may indicate a good predictability of the IPV mode. Indeed, as indicated by MSSS scores of TPI, PRED is more accurate than FREE at Y1 (Figure 15a), with both “effective” and potential predictability showing similar scores. Comparison with observational estimates are also shown in Figure 15. After Y1, PRED and FREE performances become indifferentiable (Figure 14, Figure 15a). Similarly, focusing on the SST skill over the SPNA area, MSSS indicates that PRED performs better than FREE (Figure 15b) for leadtimes up to 3 years. If we focus on potential predictability, PRED is always more accurate than FREE up to Y4 over the SPNA (Figure 14 and Figure 15c).

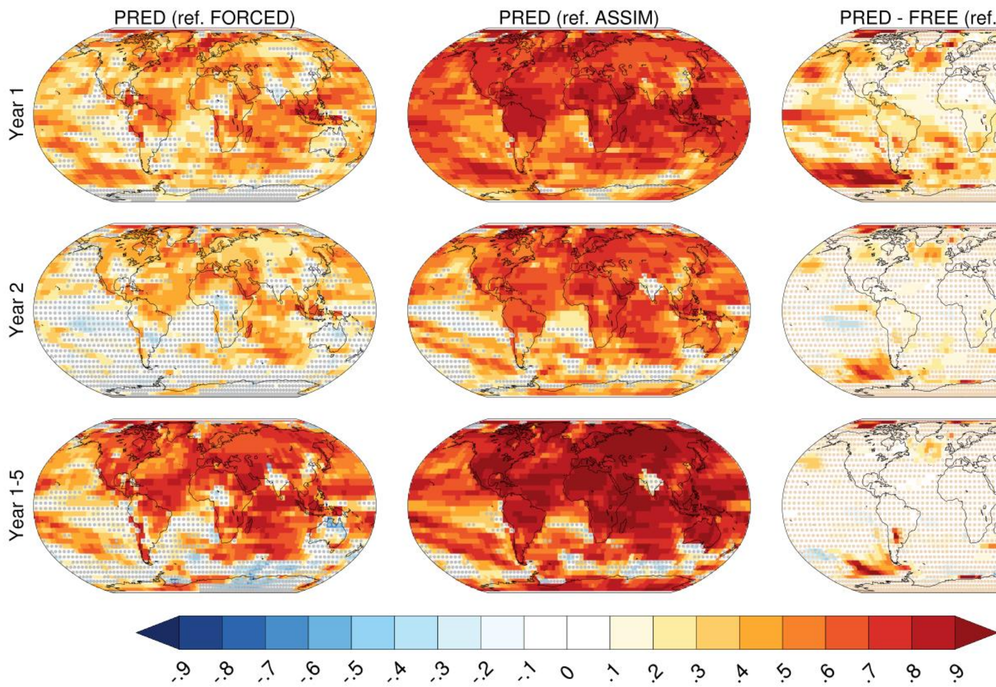

Figure 14. Left column: ACC skill scores for SST over the ocean and SAT over land computed between PRED and JRA55do reanalysis for lead times of 1 year (top), 2 years (middle) and 1–5 years (bottom). Middle column: the same but the ACC is computed between PRED and ASSIM (potential predictability). Right column: Differences between the ACC the PRED versus FREE when ASSIM is used as reference. All the data were interpolated to a regular 5-degree grid before the analysis. Stippling with gray dots indicates skill scores that are not significant at the 10% level based on block-bootstrapping as explained in the text. Stippling with light brown dots indicates ACC differences that are not significant at the 10% level based on block-bootstrapping as explained in the text.

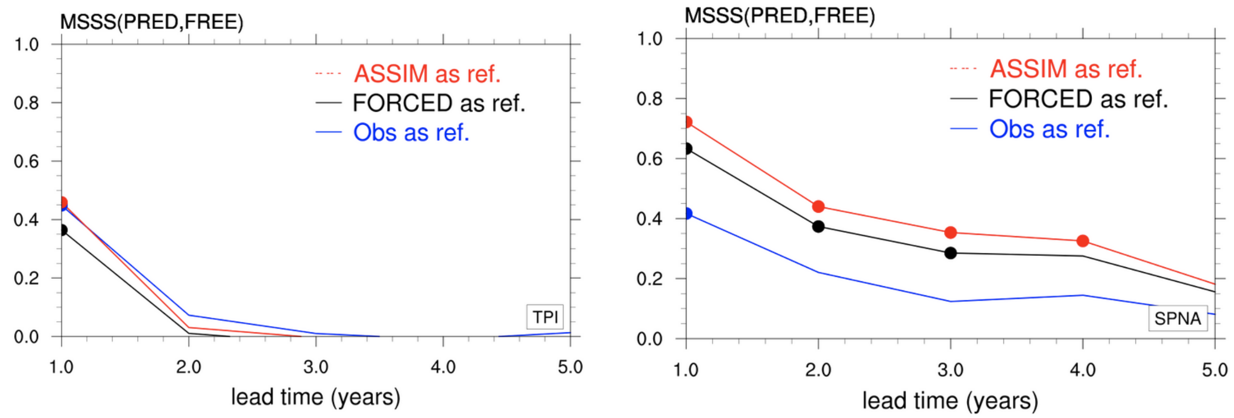

Figure 15. (a) MSSS skill scores for the SSTs for the TPI. To compute MSSS PRED and FREE are compared with respect to the same reference: observations (HadISSTv4, blue line), FORCED (black line) and ASSIM (red line). Positive MSSS indicates that PRED performs better than FREE. The dots indicate where MSSS is statistically significant at the 10% level based on block-bootstrapping as explained in the text. **(b)** The same as (a) but for SSTs over the SPNA box. The TPI index is computed from raw data according to Henley et al. 2015. The SPNA index from raw data is obtained according to Bilbao et al. 2022 (SPNA: 50–65°N, 60–10°W).

Forecasting skills of C3PS for biogeochemical variables such as NPP and ocean carbon fluxes are also assessed using the concept of "effective" and potential predictability. NPP skill scores show in general a high level of predictability over midlatitudes at Y1, Y2 and the average Y1-5 (Figure 16, left column). In contrast, the predictability of NPP in most of the tropics is very low and even the skill can be even negative when the variability of the NPP is opposite in phase with that of the target (FORCED or ASSIM). Such global features of the C3PS predictive skill for NPP contrast with the results of Séférian et al. (2014) using the IPSL-CM5A-LR model and SST anomaly initialization scheme but are in line with the findings of Frölicher et al. (2020) using GFDL-ESM2-M.

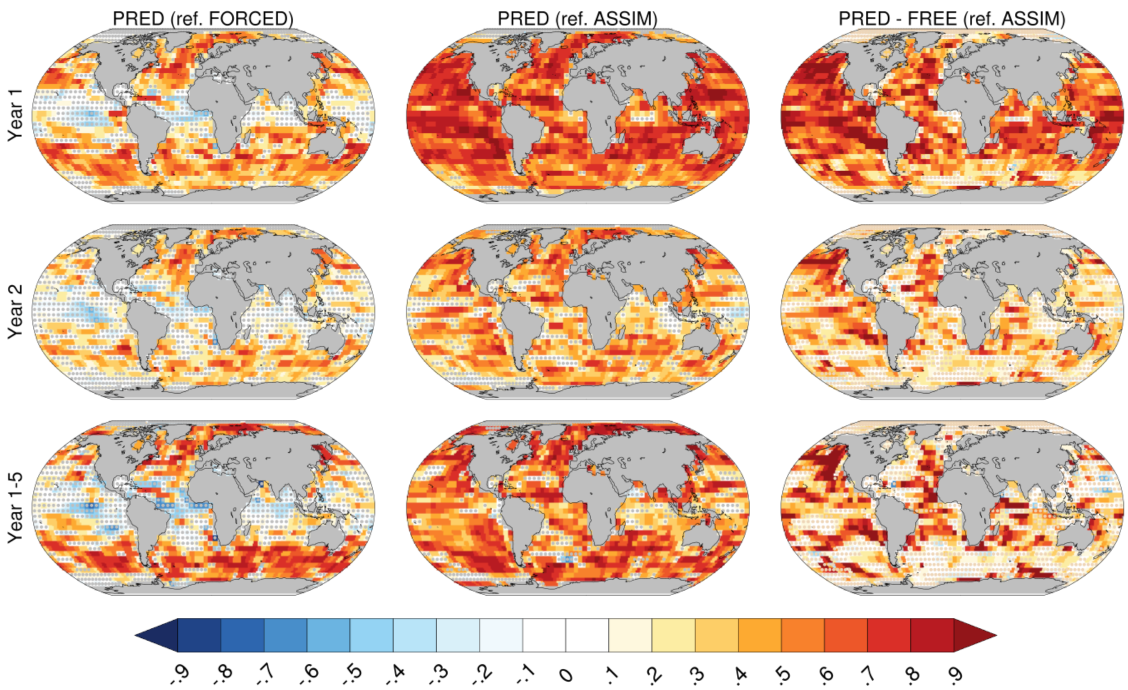

Figure 16. Left column: ACC skill scores for NPP computed between PRED and FORCED for lead times of 1 year (top), 2 years (middle) and 1–5 years (bottom). Middle column: the same but the ACC is computed between PRED and ASSIM (potential predictability). Right column: Differences between the ACC the PRED versus FREE when ASSIM is used as reference. All the data were interpolated to a regular 5-degree grid before the analysis. Stippling with gray dots indicates skill scores that are not significant at the 10% level based on block-bootstrapping as explained in the text. Stippling with light brown dots indicates ACC differences that are not significant at the 10% level based on block-bootstrapping as explained in the text.

Potential predictability of NPP shows good skill scores worldwide for Y1 (Figure 17, middle column). At Y2 NPP skill decreases over some areas in the Equatorial Pacific, western Atlantic, Northern Indian and Southern Oceans, but in general it remains high and statistically significant over most of the ocean for Y1-5. Most importantly, NPP skill is high in the areas of highest marine productivity, such as the equatorial and eastern boundary upwelling systems, in particular the Canary Upwelling System. The impact of model initialisation is more important on the NPP than on the SST beyond the second year of forecasting, indicating that the initialisation of the BGC undoubtedly leads to benefits in predictive ability. PRED performs better than FREE practically everywhere at Y1 (Figure 17, right column). At longer horizons, ACC differences show that PRED is more accurate than FREE over the Eastern North Atlantic, including the Canary Upwelling area, Tropical Atlantic, North Pacific and Central Equatorial Pacific and most of the Indian Ocean.

ACC skill of ocean carbon fluxes with FORCED as the reference is relatively high and significant over the tropical band and Southern Oceans during the first 2 years after the initialization (Figure 17, first and second column). Such result is consistent with the first multi-model assessment of the ocean carbon sink prediction skills (Ilyina et al. 2021).

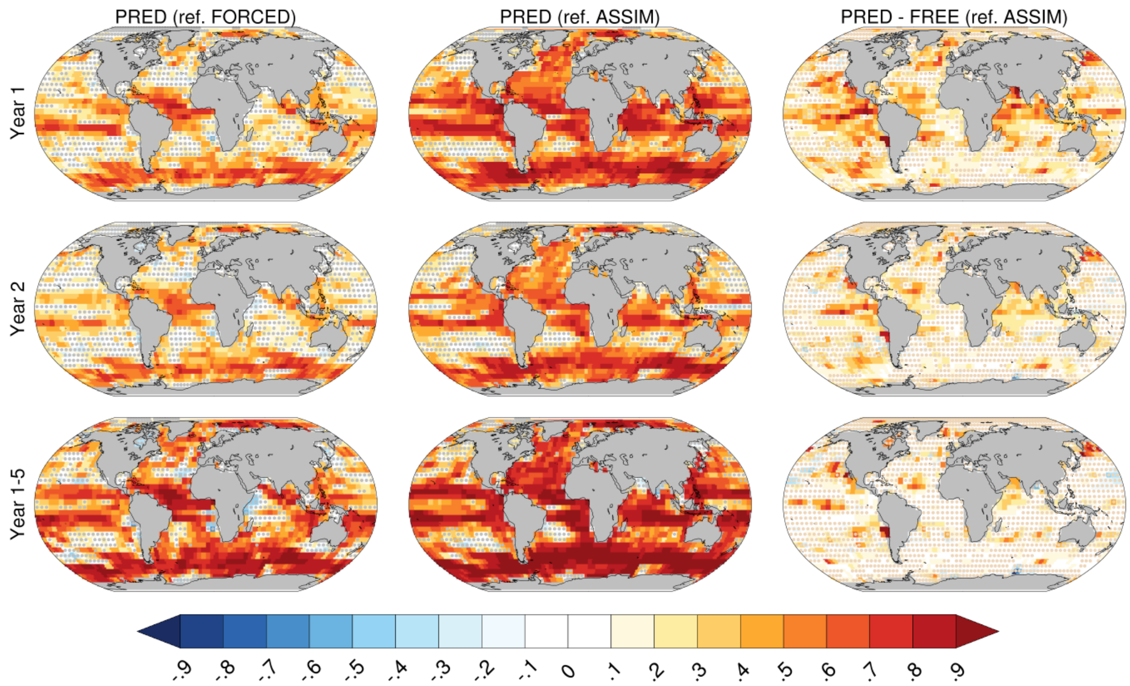

Figure 17. Left column: ACC skill scores for ocean carbon fluxes computed between PRED and FORCED for lead times of 1 year (top), 2 years (middle) and 1–5 years (bottom). Middle column: the same but the ACC is computed between PRED and ASSIM (potential predictability). Right column: Differences between the ACC the PRED versus FREE when ASSIM is used as reference. All the data were interpolated to a regular 5-degree grid before the analysis. Stippling with red dots indicates skill scores that are not significant at the 10% level based on block-bootstrapping as explained in the text. Stippling with gray dots indicates skill scores that are not significant at the 10% level based on block-bootstrapping as explained in the text. Stippling with light brown dots indicates ACC differences that are not significant at the 10% level based on block-bootstrapping as explained in the text.

C3PS provides skillful predictions of ocean carbon uptake at multiannual scale over the high latitude oceans and the tropics. Potential predictability is even higher and indicates the important fact that ocean carbon fluxes can be predictable several years in advance over the areas of large carbon uptake variability such as North Atlantic and North Pacific oceans and Southern Ocean. These results support previous predictability studies based on perfect model frameworks or decadal predictions with ESMs (Lovenduski et al., 2019, Séférian et al., 2019). More importantly, this potential predictability exceeds that inferred by the knowledge of the external forcing for the first two years after the initialization (Figure 17, right column). After that time horizon most of the predictive skill comes from the increase of atmospheric CO2 as the primary driver of the ocean carbon sink. Within the lead years 1, 2 and 1-5, the predictable fraction of the ocean carbon sink is 37%, 19%, 16%. At Y1, predictable regions include the North Atlantic and the Southern Ocean, the two major ocean carbon sink locations, as well as the Equatorial Pacific. After Y2, only the Southern ocean carbon sink remains predictable as well as a smaller fraction of the Equatorial Pacific domain. This result is in line with previous work made with other modelling prediction platforms (Lovenduski et al., 2019, Séférian et al., 2019).

6 Conclusions

In this study, the new climate prediction prototype of the CNRM-Cerfacs modelling group, C3PS is presented and evaluated. The two main novelties are that C3PS is based on an earth system model, CNRM-ESM2.1, and has been designed to produce predictions from seasonal to multiannual scales. C3PS is the result of the joint work of experts in seasonal and decadal forecasting and modellers of ocean physics and biogeochemistry within the CNRM-Cerfacs research group. In addition, for interannual predictions, C3PS has participated in the international DCPP-A exercise, and a subset of the variables produced are published in the ESGF.

The initialisation procedure of C3PS consists of a full-field initialisation in which all the model components are initialised from an in-home reanalysis product obtained in two steps. The first step is a forced experiment in which ocean and biogeochemistry models are driven by JRA55do reanalysis following the GCP protocol. In the second step, the T and S of this forced experiment from step1 are used to constrain only the ocean physics of CNRM-ESM2.1 through sea surface restoring and a Newtonian damping in the ocean subsurface, as described in Sanchez-Gomez et al. (2016). This method has been implemented in other climate prediction systems as in Bilbao et al. (2021). The reconstruction obtained is called dcppA-assim according to the nomenclature used in the DCPP protocol.

In this paper we have performed a basic validation of the dcppA-assim (ASSIM) experiment, which is not often done in other studies presenting forecasting systems. For us it is important to evaluate and to document the quality of our initial conditions and to investigate how the nudging of T and S affects the behaviour of other variables, such as AMOC and biogeochemistry. We believe that the study of the reconstructions created to initialize the climate prediction systems is relevant, and even more so in the context of the new applications offered in the prediction of marine biogeochemistry and carbon fluxes.

ASSIM shows improvements with respect to the historical ensemble FREE in the modes of variability at the seasonal and decadal scales. The improvements are notable in the Pacific, with better representation of ENSO diversity by ASSIM and of Pacific decadal variability associated with the IPV. For other variables and other regions such as in the SPNA, ASSIM shows consistency with the time phase of observations in both ocean surface and heat content.

Regarding the initialisation of biogeochemistry, we found an interesting result. The nudging of T and S is not sufficient to constrain the biogeochemistry, as seen in the biases presented by ASSIM in chlorophyll. We suggest that biases in nutrients, such as NO3, and an underestimation of MLD can consistently explain the misrepresentation of chlorophyll in ASSIM. This result offers perspectives for improving the reconstruction of biogeochemical variables, indicating that we should pay special attention to nutrients, which leads us to think of a nutrient nudging complementary to the nudging of physical variables.

Nevertheless we show that the T/S nudging leads to a significant improvement in the amplitude of the variability and temporal chronology of the NPP in the Tropical Pacific, coherent with previous studies (Seferian et al. 2014). Moreover, our results also show an added value of nudging in representing carbon sink variability in terms of magnitude and timing. This improvement is due to the fact that the fugacity is controlled by changes in T and S in the ocean, which are directly affected by the nudging.

In terms of skill at seasonal scale, C3PS shows a very similar ENSO prediction skill to other seasonal forecasting systems. Considering the diversity of ENSO, the C-ENSO mode exhibits higher and significant skill levels compared to the E-ENSO mode. This is somewhat expected since the E-ENSO mode is associated with the variability of extreme El Niño events which onsets are difficult to predict due to their nonlinear dynamics. Seasonal prediction systems also exhibit a persistent mean temperature bias in the far eastern Pacific, which alters key ENSO processes (e.g. thermocline feedback and atmospheric convection) in this region explaining the lower skill in terms of the E-ENSO mode. However the C3PS skill at seasonal timescales in the tropical Pacific is encouraging for addressing seasonal forecast skill over other regions assuming a realistic simulation of ENSO atmospheric teleconnections. Such an estimation may however suffer from the limitation of using only 30 members for the first prototype of C3PS. We have considered extending the ensemble size for future applications and evaluation.

On an interannual scale, the C3PS results are consistent with those found in other decadal forecasting systems (IPCC 2023), i.e. C3PS shows a clear added value of ocean initialisation in the prediction of SST and SAT in the first two years. The novelty is a significant prediction skill of SSTs in the equatorial Pacific at Y1. On longer time scales, the added value of initialisation is only detectable in the SPNA area.

The most innovative aspect of the C3PS results is the potential predictive skill displayed for NPP and carbon fluxes at different leadtimes. The high levels of NPP potential predictability at multi-annual timescales were already reported in Seferian et al. (2014) and recently addressed in Yeager et al. 2022. These results corroborate previous findings and confirm the potential benefits for marine ecosystem prediction based on integrated physical-biogeochemical forecasting platforms such as C3PS (Tommasi et al. 2017). The fact that the evolution of carbon fluxes is potentially predictable over the regions of major carbon sink locations is also promising for improving our estimations of the future global carbon budget in the climate system.

To finish, although C3PS is designed with an improved initialization scheme, the C3PS multi-annual predictions still suffer from initial shocks and drifts after the initialisation. In particular the ENSO drift documented in Sanchez-Gomez et al. (2016) is still present in the C3PS predictions. As shown in this study, the first year after the initialization is characterized by a quasi-systematic excitation of ENSO warm events that trigger teleconnection patterns over the midlatitudes, potentially polluting the signals to be predicted. The drift problem is one of the major challenges in decadal prediction. Although some progress has been achieved since the early 2000s, drifts are still present in decadal prediction systems. In addition to improving climate models to reduce errors, another essential aspect is the improvement of the data assimilation technique to obtain initial states compatible with the climate model that will be used to make the prediction. On this line, some decadal forecasting centers are opting for “in-home reanalysis” built from the coupled models used to make the forecasts to maintain physical consistency amongst all model components. This idea involves using more complex data assimilation methods, such as the use of the Ensemble Kalman Filter or the particle filter approaches (Counillon et al., 2014; Zunz et al., 2015; Dai et al., 2020) which have been successfully applied in the context of decadal prediction. This offers interesting perspectives for improving initialization and for these reasons the implementation of a particle filter in C3PS is one of the perspectives to improve the initialisation procedure.

Acknowledgments

The C3PS platform has been developed in the framework of the EU-H2020 TRIATLAS project framework (https://triatlas.w.uib.no/), under the grant agreement 817578.
The authors would especially like to thank the entire CNRM-Cerfacs team, whose work in the development of the modelling platform and all the climate models has helped enormously in the realisation of this work. We also thank the Météo-France HPC team for allowing us enough computing hours and resources to complete all the experiments in the Belenos supercomputer and Hendrix server. Boris Dewitte acknowledges support from Agencia Nacional de Investigación y Desarrollo (Concurso de Fortalecimiento al Desarrollo Científico de Centros Regionales 2020-R20F0008-CEAZA, Anillo Eclipse ACT210071, Fondecyt Regular 1231174) and the EU H2020 FutureMares project (Theme LC-CLA-06-2019, Grant agreement No 869300). Finally we would also like to thank the anonymous reviewers for their comments which will help to improve this article.

Open Research

Observational data used in this study:

- The merged SST/TAS dataset would be available upon request.
- The HadISST data are available on https://www.metoffice.gov.uk/hadobs/hadisst/.
- The EN4 ocean temperature data are available on https://www.metoffice.gov.uk/hadobs/en4/.
- The RAPID array data are available on https://rapid.ac.uk/data.php.
- The ESA-OC-CC data are available on https://climate.esa.int/en/projects/ocean-colour/.
- Others datasets, NPP, fCO2 are available upon request.

Software Availability Statement: All of the CNRM-ESM2-1 model outputs are available for download on ESGF under CMIP6 projects. The SURFEX-CTRIP code is available (Open-SURFEX) using a CECILL-C Licence (http://www.cecill.info/licences/Licence_CeCILL-C_V1-en.txt) at the SURFEX website (http://www.umr-cnrm.fr/surfex). NEMO-GELATO-PISCESv2-gas is also available at https://opensource.umr-cnrm.fr/; the access to the Git repository is granted upon request to the corresponding author. OASIS3-MCT can be downloaded at this website (https://verc.enes.org/oasis/download). XIOS can be downloaded at the XIOS website (https://forge.ipsl.jussieu.fr/ioserver). For the ARPEGE-Climat_v6.3 code and exact version applied to each component, please contact the authors. Finally, a number of analyzing tools developed at CNRM, or in collaboration with CNRM scientists, is available on as Open Source code (see https://opensource.cnrm-game-meteo.fr/).

References

Athanasiadis, P.J. et al., (2020) : Decadal predictability of North Atlantic blocking and the NAO. npj Climate and Atmospheric Science, 3(1), 20, doi:10.1038/s41612-020-0120

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

Bakker, D. C. E., Alin, S. R., Becker, M., Bittig, H. C., Castaño-Primo, R., Feely, R. A., Gkritzalis, T. Kadono, K., Kozyr, A., Lauvset, S. K., Metzl, N., Munro, D. R., Nakaoka, S.-I., Nojiri, Y., O'Brien, K. M., Olsen, A., Pfeil, B., Pierrot, D., Steinhoff, T., Sullivan, K. F., Sutton, A. J., Sweeney, C., Tilbrook, B., Wada, C., Wanninkhof, R., Willstrand Wranne, A., Akl, J., Apelthun, L. B., Bates, N., Beatty, C. M., Burger, E. F., Cai, W.-J., Cosca, C. E., Corredor, J. E., Cronin, M., Cross, J. N., De Carlo, E. H., DeGrandpre, M. D., Emerson, S., Enright, M. P., Enyo, K., Evans, W., Frangoulis, C., Fransson, A., García-Ibáñez, M. I., Gehrung, M., Giannoudi, L., Glockzin, M., Hales, B., Howden, S. D., Hunt, C. W., Ibánhez, J. S. P., Jones, S. D., Kamb, L., Körtzinger, A., Landa, C. S., Landschützer, P., Lefèvre, N., Lo

Monaco, C., Macovei, V. A., Maenner Jones, S., Meinig, C., Millero, F. J., Monacci, N. M., Mordy, C., Morell, J. M., Murata, A., Musielewicz, S., Neill, C., Newberger, T., Nomura, D., Ohman, M., Ono, T., Passmore, A., Petersen, W., Petihakis, G., Perivoliotis, L., Plueddemann, A. J., Rehder, G., Reynaud, T., Rodriguez, C., Ross, A., Rutgersson, A., Sabine, C. L., Salisbury, J. E., Schlitzer, R., Send, U., Skjelvan, I., Stamataki, N., Sutherland, S. C., Sweeney, C., Tadokoro, K., Tanhua, T., Telszewski, M., Trull, T., Vandemark, D., van Ooijen, E., Voynova, Y. G., Wang, H., Weller, R. A., Whitehead, C., and Wilson, D. (2022): Surface Ocean CO_2 Atlas Database Version 2022 (SOCATv2022) (NCEI Accession 0253659), NOAA National Centers for Environmental Information [data set], https://doi.org/10.25921/1h9f-nb73

Barnier, B. L. Siefridt, P. Marchesiello, (1995): Thermal forcing for a global ocean circulation model using a three-year climatology of ECMWF analyses, Journal of Marine Systems, Volume 6, Issue 4, Pages 363-380.

Batté and Déqué, 2012 : A stochastic method for improving seasonal predictions *Geophysical Research Letters, vol. 39, L09707*, DOI 10.1029/2012GL051406

Bellucci, A., Haarsma, R., Gualdi, S., Athanasiadis, P. J., Caian, M., Cassou, C., ... & Yang, S. (2015) : An assessment of a multi-model ensemble of decadal climate predictions. Climate Dynamics, 44(9), 2787-2806

Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., et al. (2021) : The ERA5 global reanalysis: Preliminary extension to 1950. Q J R Meteorol Soc, 147(741, 4186–4227. Available from: https://doi.org/10.1002/qj.4174

Bilbao, R., Wild, S., Ortega, P., Acosta-Navarro, J., Arsouze, T., Bretonnière, P.-A., Caron, L.-P., Castrillo, M., Cruz-García, R., Cvijanovic, I., Doblas-Reyes, F. J., Donat, M., Dutra, E., Echevarría, P., Ho, A.-C., Loosveldt-Tomas, S., Moreno-Chamarro, E., Pérez-Zanon, N., Ramos, A., Ruprich-Robert, Y., Sicardi, V., Tourigny, E., and Vegas-Regidor, J. (2021): Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth, Earth Syst. Dynam., 12, 173–196, https://doi.org/10.5194/esd-12-173-2021.

Boyer, Tim P.; Garcia, Hernan E.; Locarnini, Ricardo A.; Zweng, Melissa M.; Mishonov, Alexey V.; Reagan, James R.; Weathers, Katharine A.; Baranova, Olga K.; Seidov, Dan; Smolyar, Igor V. (2018) : World Ocean Atlas 2018. [indicate subset used]. NOAA National Centers for Environmental Information. Dataset. https://www.ncei.noaa.gov/archive/accession/NCEI-WOA18. Accessed [September 2023].

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danaba- soglu, G., Kirtman, B., Kushnir, Y., Kimoto, M., Meehl, G. A., Msadek, R., Mueller, W. A., Taylor, K. E., Zwiers, F., Rixen, M., Ruprich-Robert, Y., and Eade, R. (2016): The Decadal Climate Predic- tion Project (DCPP) contribution to CMIP6, Geosci. Model Dev., 9, 3751–3777, https://doi.org/10.5194/gmd-9-3751-2016.

Bojovic, D., and Coauthors, (2019) : The Biggest Unknowns Related to Decadal Prediction: What 50 Experts Think Are the 5 Major Knowledge Gaps. Bull. Amer. Meteor. Soc., 100, ES255–ES259, https://doi.org/10.1175/BAMS-D-19-0190.1.

Borchert, L.F. et al., (2021): Improved Decadal Predictions of North Atlantic Subpolar Gyre SST in CMIP6. Geophysical Research Letters, 48(3), e2020GL091307, doi:10.1029/2020gl091307.

Cai, W., Wang, G., Dewitte, B., Wu, L., Santoso, A., Takahashi, K., ... & McPhaden, M. J. (2018). Increased variability of eastern Pacific El Niño under greenhouse warming. Nature, 564(7735), 201-206. DOI: https://doi.org/10.1038/s41586-018-0776-9

Capotondi, A., Wittenberg, A., Kug, J. S., Takahashi, K., and McPhaden, M. (2020). "ENSO diversity," in El Niño Southern Oscillation in a Changing Climate, eds. M.J. McPhaden, A. Santoso,W. Cai (Washington DC: AGU), 65–86. doi: 10.1002/9781119548164.ch4

Cassou, C., Y. Kushnir, E. Hawkins, A. Pirani, F. Kucharski, I. Kang, and N. Caltabiano (2018): Decadal Climate Variability and Predictability: Challenges and Opportunities. Bull. Amer. Meteor. Soc., 99, 479–490, https://doi.org/10.1175/BAMS-D-16-0286.1.

Choi, J., Son, SW, (2022): Seasonal-to-decadal prediction of El Niño–Southern Oscillation and Pacific Decadal Oscillation. npj Clim Atmos Sci 5, 29, https://doi.org/10.1038/s41612-022-00251-9

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J.-P., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., & Voldoire, A. (2019). Recent changes in the ISBA-CTRIP land surface system for use in the CNRM-CM6 climate model and in global off-line hydrological applications. *Journal of Advances in Modeling Earth Systems*, 11, 1207–1252. https://doi.org/10.1029/2018MS001545

Delire, C., Foley, J. A., and Thompson, S. (2003), Evaluating the carbon cycle of a coupled atmosphere-biosphere model, *Global Biogeochem. Cycles*, 17, 1012, doi:10.1029/2002GB001870, 1.

Delgado-Torres, C., and Coauthors, 2022: Multi-Model Forecast Quality Assessment of CMIP6 Decadal Predictions. *J. Climate*, **35**, 4363–4382, https://doi.org/10.1175/JCLI-D-21-0811.1.

Doblas-Reyes, F., Andreu-Burillo, I., Chikamoto, Y. et al. (2013): Initialized near-term regional climate change prediction. Nat Commun 4, 1715. https://doi.org/10.1038/ncomms2704

Drijfhout, S., G. J. van Oldenborgh, and A. Cimatoribus, 2012: Is a Decline of AMOC Causing the Warming Hole above the North Atlantic in Observed and Modeled Warming Patterns?. *J. Climate*, **25**, 8373–8379, https://doi.org/10.1175/JCLI-D-12-00490.1.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016): Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimen- tal design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Fan, Y., & van den Dool, H. (2008). A global monthly land surface air temperature analysis for 1948-present. *Journal of Geophysical Research*, 113, D01103.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., Arneth, A., Arora, V. K., Bates, N. R., Becker, M., Bellouin, N., Bittig, H. C., Bopp, L., Chevallier, F., Chini, L. P., Cronin, M., Evans, W., Falk, S., Feely, R. A., Gasser, T., Gehlen, M., Gkritzalis, T., Gloege, L., Grassi, G., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jain, A. K., Jersild, A., Kadono, K., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lindsay, K., Liu, J., Liu, Z., Marland, G., Mayot, N., McGrath, M. J., Metzl, N., Monacci, N. M., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pan, N., Pierrot, D., Pocock, K., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Rodriguez, C., Rosan, T. M., Schwinger, J., Séférian, R., Shutler, J. D., Skjelvan, I., Steinhoff, T., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tanhua, T., Tans, P. P., Tian, X., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., Walker, A. P., Wanninkhof, R., Whitehead, C., Willstrand Wranne, A., Wright, R., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: (2022): Global Carbon Budget 2022, Earth Syst. Sci. Data, 14, 4811–4900, https://doi.org/10.5194/essd-14-4811-2022.

Frölicher, T. L., Ramseyer, L., Raible, C. C., Rodgers, K. B., and Dunne, J. (2020): Potential predictability of marine ecosystem drivers, Biogeosciences, 17, 2061–2083, https://doi.org/10.5194/bg-17-2061-2020.

García-Serrano, J., Guemas, V., and Doblas-Reyes, F. J. (2015): Added-value from initialization in predictions of Atlantic multi-decadal variability, Clim. Dynam., 44, 2539–2555, https://doi.org/10.1007/s00382-014-2370-7.

Gillett, N. P., Shiogama, H., Funke, B., Hegerl, G., Knutti, R., Matthes, K., Santer, B. D., Stone, D., & Tebaldi, C. (2016). The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6. *Geoscientific Model Development*, 9(10), 3685–3697. https://doi.org/10.5194/gmd-9-3685-2016

Goddard, L., and Coauthors, (2013): A verification framework for interannual-to-decadal predictions experiments. Climate Dynamics, 40, 245–272, https://doi.org/10.1007/s00382-012-1481-2.

Gregg WW & N.W. Casey (2004): Global and regional evaluation of the SeaWiFS chlorophyll data set Remote Sens. Environ., 93, pp. 463-479

Hauck Judith, Zeising Moritz, Le Quéré Corinne, Gruber Nicolas, Bakker Dorothee C. E., Bopp Laurent, Chau Thi Tuyet Trang, Gürses Özgür, Ilyina Tatiana, Landschützer Peter, Lenton Andrew, Resplandy Laure, Rödenbeck
Christian, Schwinger Jörg, Séférian Roland (2020): Consistency and Challenges in the Ocean Carbon Sink Estimate for the Global Carbon Budget, Frontiers in Marine Science, vol 7, https://www.frontiersin.org/articles/10.3389/fmars.2020.571720.

Hersbach, H, Bell, B, Berrisford, P, et al. (2020) : The ERA5 global reanalysis. Q J R Meteorol Soc., 146: 1999–2049. https://doi.org/10.1002/qj.3803

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., & Zhang, H. M. (2017). Extended reconstructed sea surface temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *Journal of Climate*, 30(20), 8179–8205. https://doi.org/10.1175/JCLI-D-16-0836.1

Ilyina, T., Li, H., Spring, A., Müller, W. A., Bopp, L., Chikamoto, M. O., et al. (2021): Predictable variations of the carbon sinks and atmospheric CO2 growth in a multi-model framework. Geophysical Research Letters, 48, e2020GL090695. https://doi.org/10.1029/2020GL090695

IPCC, (2023): Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 35-115, doi: 10.59327/IPCC/AR6-9789291691647

Karamperidou, C., Jin, F. F., and Conroy, J. L. (2017). The importance of ENSO nonlinearities in tropical pacific response to external forcing. Clim. Dyn. 49, 2695–2704. doi: 10.1007/s00382-016-3475-y

Keenlyside, N.S. and Ba, J. (2010): Prospects for Decadal Climate Prediction. Wiley Interdisciplinary Reviews: ClimateChange,1,627-635, https://doi.org/10.1002/wcc.69

Kirtman, B., Power, S. B., Adedoyin, J. A., Boer, G. J., Bojariu, R., Camilloni, I., Doblas-Reyes, F. J., Fiore, A. M., Kimoto, M., Meehl, G. A., Prather, M., Sarr, A., Schär, C., Sutton, R., van Oldenborgh, G. J., Vecchi, G., and Wang, H. J.: Near-term Cli- mate Change: Projections and Predictability, in: Climate Change (2013): The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmen- tal Panel on Climate Change, edtied by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Kulk, Gemma & Platt, Trevor & Dingle, James & Jackson, Thomas & Jönsson, Bror & Bouman, Heather & Babin, Marcel & Brewin, Bob & Doblin, Martina & Estrada, Marta & Figueiras, F.G. & Furuya, Ken & González-Benítez, Natalia & Gudfinnsson, Hafsteinn & Gudmundsson, Kristinn & Huang, Bangqin & Isada, Tomonori & Kovač, Žarko & Lutz, Vivian & Sathyendranath, Shubha. (2020): Primary Production, an Index of Climate Change in the Ocean: Satellite-Based Estimates over Two Decades. Remote Sensing. 12. 826. 10.3390/rs12050826.

Lee, J., Planton, Y. Y., Gleckler, P. J., Sperber, K. R., Guilyardi, E., Wittenberg, A. T., ... & Pallotta, G. (2021). Robust evaluation of ENSO in climate models: How many ensemble members are needed?. Geophysical Research Letters, 48(20), e2021GL095041.

L'Heureux, M. L., Levine, A. F., Newman, M., Ganter, C., Luo, J. J., Tippett, M. K., & Stockdale, T. N. (2020): ENSO prediction. El Niño Southern Oscillation in a changing climate, 227-246.

L'Heureux Michelle L., Tippett Michael K., Wang Wanqiu (2022): Prediction Challenges From Errors in Tropical Pacific Sea Surface Temperature Trends, Frontiers in Climate
, https://www.frontiersin.org/articles/10.3389/fclim.2022.837483

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., et al. (2018): Global carbon budget 2018. Earth System Science Data, 10(4), 2141–2194. https://doi.org/10.5194/essd-10-2141-2018

Lovenduski, N. S., Yeager, S. G., Lindsay, K., and Long, M. C.: Predicting near-term variability in ocean carbon uptake, Earth Syst. Dynam., 10, 45–57, https://doi.org/10.5194/esd-10-45-2019, 2019.

Luo Y, Keenan T F and Smith M (2015): Predictability of the terrestrial carbon cycle Glob. Change Biol. 21 1737–51

Madec, G., Bourdallé-Badie, R., Bouttier, P-A., Bricaud, C., Bruciaferri, D., Calvert, D., Chanut, J., Clementi, E., Coward, A., Delrosso, D., Ethé, C., Flavoni, S., Graham, T., Harle, J., Iovino, D., Lea, D., Lévy, C., Lovato, T., Martin, N., Masson, S., Mocavero, S., Paul, J., Rousset, C., Storkey, D., Storto, A., Vancoppenolle, M. (2017). NEMO ocean engine. https://doi.org/10.5281/ZENODO.1472492

Materia, S., A. Borrelli, A. Bellucci, A. Alessandri, P. Di Pietro, P. Athanasiadis, A. Navarra, and S. Gualdi, (2014): Impact of Atmosphere and Land Surface Initial Conditions on Seasonal Forecasts of Global Surface Temperature. J. Climate, 27, 9253–9271, https://doi.org/10.1175/JCLI-D-14-00163.1.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Ki- moto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D., and Stockdale, T. (2009): Decadal Prediction: Can It Be Skillful?, B. Am. Meteorol. Soc., 90, 1467–1486, https://doi.org/10.1175/2009BAMS2778.1.

Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Bransta- tor, G., Cassou, C., Corti, S., Danabasoglu, G., Doblas- Reyes, F., Hawkins, E., Karspeck, A., Kimoto, M., Ku- mar, A., Matei, D., Mignot, J., Msadek, R., Navarra, A., Pohlmann, H., Rienecker, M., Rosati, T., Schneider, E., Smith, D., Sutton, R., Teng, H., van Oldenborgh, G. J., Vec- chi, G., and Yeager, S. (2014): Decadal Climate Prediction: An Up- date from the Trenches, B. Am. Meteorol. Soc., 95, 243–267, https://doi.org/10.1175/BAMS-D-12-00241.1.

Meehl, G.A., Teng, H., Smith, D. et al. The effects of bias, drift, and trends in calculating anomalies for evaluating skill of seasonal-to-decadal initialized climate predictions. Clim Dyn 59, 3373–3389 (2022). https://doi.org/10.1007/s00382-022-06272-7

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., & Weiss, R. (2017). Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geoscientific Model Development*, 10(5), 2057–2116. https://doi.org/10.5194/gmd-10-2057-2017

Menary, M. B., Kuhlbrodt, T., Ridley, J., Andrews, M. B., Dimdore-Miles, O. B., Deshayes, J., et al. (2018). Preindustrial control simulations with HadGEM3-GC3.1 for CMIP6. *Journal of Advances in Modeling Earth Systems*, 10, 3049–3075. https://doi.org/10.1029/2018MS001495

Meurdesoif Y (2018): Xios fortran reference guide. IPSL, http://forge.ipsl.jussieu.fr/ioserver/svn/XIOS/trunk/doc/XIOS_reference_guide.pdf

Moat, B., Frajka-Williams, E., Smeed, D., Rayner, D., Johns, W. E., Baringer, M. O., et al. (2022). Atlantic Meridional Overturning Circulation observed by the RAPID-MOCHA-WBTS (RAPID-Meridional Overturning Circulation and Heatflux array-western boundary time series) array at 26N from 2004 to 2020 (v2020.1) [Dataset]. British Oceanographic Data Centre - Natural Environment Research Council, UK. https://doi.org/10.5285/e91b10af-6f0a-7fa7-e053-6c86abc05a09

Monerie, P.-A., J. Robson, B. Dong, and N. Dunstone, (2018): A role of the Atlantic Ocean in predicting summer surface air temperature over North East Asia? Climate Dynamics, 51(1), 473–491, doi:10.1007/s00382-017-3935-z.

Muller, R. A., Curry, J., Groom, D., Jacobsen, R., Perlmutter, S., Rohde, R., Rosenfeld, A., Wickham, C., & Wurtele, J. (2013). Decadal variations in the global atmospheric land temperatures. *Journal of Geophysical Research: Atmospheres*, 118, 5280–5286. https://doi.org/10.1002/jgrd.50458

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., & Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, 9(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Pohlmann, H., Kröger, J., Greatbatch, R.J. *et al.* Initialization shock in decadal hindcasts due to errors in wind stress over the tropical Pacific. *Clim Dyn* **49**, 2685–2693 (2017). https://doi.org/10.1007/s00382-016-3486-8

Rayner, N. A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, 108(D14), 4407. https://doi.org/10.1029/2002JD002670

Roehrig, R., Beau, I., Saint-Martin, D., Alias, A., Decharme, B., Guérémy, J.-F., et al. (2020). The CNRM global atmosphere model ARPEGE-Climat 6.3: Description and evaluation. *Journal of Advances in Modeling Earth Systems*, 12, e2020MS002075. https://doi.org/10.1029/2020MS002075

Sanchez-Gomez, E., Cassou, C., Ruprich-Robert, Y., Fernandez, E., and Terray, L. (2016): Drift dynamics in a coupled model initialized for decadal forecasts, Clim. Dynam., 46, 1819–1840, https://doi.org/10.1007/s00382-015-2678-y.

Salas Mélia, D. (2002). A global coupled sea ice–ocean model. *Ocean Modelling*, 4(2), 137–172. https://doi.org/10.1016/S1463-5003(01)00015-4

Sathyendranath, S, Brewin, RJW, Brockmann, C, Brotas, V, Calton, B, Chuprin, A, Cipollini, P, Couto, AB, Dingle, J, Doerffer, R, Donlon, C, Dowell, M, Farman, A, Grant, M, Groom, S, Horseman, A, Jackson, T, Krasemann, H, Lavender, S, Martinez-Vicente, V, Mazeran, C, Mélin, F, Moore, TS, Müller, D, Regner, P, Roy, S, Steele, CJ, Steinmetz, F, Swinton, J, Taberner, M, Thompson, A, Valente, A, Zühlke, M, Brando, VE, Feng, H, Feldman, G, Franz, BA, Frouin, R, Gould, Jr., RW, Hooker, SB, Kahru, M, Kratzer, S, Mitchell, BG, Muller-Karger, F, Sosik, HM, Voss, KJ, Werdell, J, and Platt, T (2019) : An ocean-colour time series for use in climate studies: the experience of the Ocean-Colour Climate Change Initiative (OC-CCI). Sensors: 19, 4285. doi:10.3390/s19194285

Schuster, M. et al., 2019: Improvement in the decadal prediction skill of the North Atlantic extratropical winter circulation through increased model resolution. Earth System Dynamics, 10(4), 901–917, doi:10.5194/esd-10- 901-2019.

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., et al (2019). Evaluation of CNRM Earth-System model, CNRM-ESM2-1: role of Earth system processes in present-day and future climate. Journal of Advances in Modeling Earth Systems, https://doi.org/10.1029/ 2019MS001791

Séférian, R., Berthet, S., Yool, A. et al. (2020): Tracking Improvement in Simulated Marine Biogeochemistry Between CMIP5 and CMIP6. Curr Clim Change Rep 6, 95–119, https://doi.org/10.1007/s40641-020-00160-0

Servonnat, J., Mignot, J., Guilyardi, E., Swingedouw, D., Séférian, R., and Labetoulle, S. (2015): Reconstructing the sub- surface ocean decadal variability using surface nudging in a perfect model framework, Clim. Dynam., 44, 315–338, https://doi.org/10.1007/s00382-014-2184-7.

Séférian, R., Berthet, S., and Chevallier, M.: Assessing the Decadal Predictability of Land and Ocean Carbon Uptake, Geophys. Res. Lett., 45, 2455–2466, https://doi.org/10.1002/2017GL076092, 2018. a, b

Sharmila, S., H. Hendon, O. Alves, A. Weisheimer, and M. Balmaseda, 2023: Contrasting El Niño–La Niña Predictability and Prediction Skill in 2-Year Reforecasts of the Twentieth Century. J. Climate, 36, 1269–1285, https://doi.org/10.1175/JCLI-D-22-0028.1.

Smith, D.M. et al., 2019: Robust skill of decadal climate predictions. npj Climate and Atmospheric Science, 2(1), 13, doi:10.1038/s41612-019-0071-y.

Smith, D.M. et al., 2020: North Atlantic climate far more predictable than models imply. Nature, 583(7818), 796–800, doi:10.1038/s41586-020-2525-0.

Swingedouw D., Bily A., Esquerdo C., Borchert L.F., Sgubin G., Mignot J., Menary M. (2021) On the risk of abrupt changes in the North Atlantic subpolar gyre in CMIP6 models, *Annals of the New York Academy of Sciences*, 15(4), 187-201.

Tommasi, D., Stock, C. A., Hobday, A. J., Methot, R., Kaplan, I. C., Eveson, J. P., et al. (2017). Managing living marine resources in a dynamic environment: the role of seasonal to decadal climate forecasts. *Prog. Oceanogr.* 152, 15–49. doi: 10.1016/j.pocean.2016.12.011

Tsujino, H., Urakawa, L. S., Griffies, S. M., Danabasoglu, G., Adcroft, A. J., Amaral, A. E., Arsouze, T., Bentsen, M., Bernardello, R., Böning, C. W., Bozec, A., Chassignet, E. P., Danilov, S., Dussin, R., Exarchou, E., Fogli, P. G., Fox-Kemper, B., Guo, C., Ilicak, M., Iovino, D., Kim, W. M., Koldunov, N., Lapin, V., Li, Y., Lin, P., Lindsay, K., Liu, H., Long, M. C., Komuro, Y., Marsland, S. J., Masina, S., Nummelin, A., Rieck, J. K., Ruprich-Robert, Y., Scheinert, M., Sicardi, V., Sidorenko, D., Suzuki, T., Tatebe, H., Wang, Q., Yeager, S. G., and Yu, Z. (2020): Evaluation of global ocean–sea-ice model simulations based on the experimental protocols of the Ocean Model Intercomparison Project phase 2 (OMIP-2), Geosci. Model Dev., 13, 3643–3708, https://doi.org/10.5194/gmd-13-3643-2020.

Takahashi, K., Montecinos, A., Goubanova, K., & Dewitte, B. (2011). ENSO regimes: Reinterpreting the canonical and Modoki El Niño. Geophysical research letters, 38(10). DOI: https://doi.org/10.1029/2011GL047364

Timmermann, A., An, S. I., Kug, J. S., Jin, F. F., Cai, W., Capotondi, A., ... & Zhang, X. (2018). El Niño–southern oscillation complexity. Nature, 559(7715), 535-545. DOI: 10.1038/s41586-018-0252-6

Valente, A., Sathyendranath, S., Brotas, V., Groom, S., Grant, M., Jackson, T., Chuprin, A., Taberner, M., Airs, R., Antoine, D., Arnone, R., Balch, W. M., Barker, K., Barlow, R., Bélanger, S., Berthon, J.-F., Beşiktepe, Ş., Borsheim, Y., Bracher, A., Brando, V., Brewin, R. J. W., Canuti, E., Chavez, F. P., Cianca, A., Claustre, H., Clementson, L., Crout, R., Ferreira, A., Freeman, S., Frouin, R., García-Soto, C., Gibb, S. W., Goericke, R., Gould, R., Guillocheau, N., Hooker, S. B., Hu, C., Kahru, M., Kampel, M., Klein, H., Kratzer, S., Kudela, R., Le2esma, J., Lohrenz, S., Loisel, H., Mannino, A., Martinez-Vicente, V., Matrai, P., McKee, D., Mitchell, B. G., Moisan, T., Montes, E., Muller-Karger, F., Neeley, A., Novak, M., O'Dowd, L., Ondrusek, M., Platt, T., Poulton, A. J., Repecaud, M., Röttgers, R., Schroeder, T., Smyth, T., Smythe-Wright, D., Sosik, H. M., Thomas, C., Thomas, R., Tilstone, G., Tracana, A., Twardowski, M., Vellucci, V., Voss, K., Werdell, J., Wernand, M., Wojtasiewicz, B., Wright, S., and Zibordi, G.: A compilation of global bio-optical in situ data for ocean colour satellite applications – version three, Earth Syst. Sci. Data, 14, 5737–5770, https://doi.org/10.5194/essd-14-5737-2022, 2022.

Verfaillie D., F.J. Doblas-Reyes, M.G. Donat, N. Pérez-Zanón, B. Solaraju-Murali, V. Torralba, and S. Wild, 2020: How reliable are decadal climate predictions of near-surface air temperature? Journal of Climate, 34(2), 697-713, doi: 10.1175/JCLI-D-20-0138.1.

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., et al. (2019). Evaluation of CMIP6 DECK experiments with CNRM-CM6-1. Journal of Advances in Modeling Earth Systems, 11, 2177–2213. https://doi.org/10.1029/2019MS001683

Waldman, R., J. Hirschi, A. Voldoire, C. Cassou, and R. Msadek, (2021) : Clarifying the Relation between AMOC and Thermal Wind: Application to the Centennial Variability in a Coupled Climate Model. *J. Phys. Oceanogr.*, **51**, 343–364, https://doi.org/10.1175/JPO-D-19-0284.1.

Wanninkhof, R. (2014). Relationship between wind speed and gas exchange over the ocean revisited. *Limnology and Oceanography: Methods*, 12(6), 351–362. https://doi.org/10.4319/lom.2014.12.351

Weisheimer, A., M. Balmaseda, T. Stockdale, M. Mayer, E. de Boisseson, R. Senan, and S. Johnson, 2021: Retrospective two-year ENSO predictions during the 20th century. ECMWF Newsletter, No. 169, Reading, United Kingdom, 7–8, https://www.ecmwfint/sites/default/files/elibrary/2021/20225-newsletter-no-169-autumn-2021.pdf.

Wu, B., T. Zhou, C. Li, W.A. Müller, and J. Lin, 2019: Improved decadal prediction of Northern-Hemisphere summer land temperature. Climate Dynamics, 53(3), 1357–1369, doi:10.1007/s00382-019-04658-8.

Yeager, S. G., Danabasoglu, G., Rosenbloom, N. A., Strand, W., Bates, S. C., Meehl, G. A., ... & Lovenduski, N. S., 2018: Predicting near-term changes in the Earth System: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model. Bulletin of the American Meteorological Society, 99(9), 1867-1886

Yeager, S. G., Rosenbloom, N., Glanville, A. A., Wu, X., Simpson, I., Li, H., et al. (2022). The Seasonal-to-Multiyear Large ensemble (SMYLE) prediction system using the Community Earth System Model version 2. *Geoscientific Model Development*, 15(16), 6451–6493. https://doi.org/10.5194/gmd-15-6451-2022

Zunz V., H. Goosse, S. Dubinkina, 2015. Impact of the initialization on the predictability of Southern Ocean sea ice at interannual to multi-decadal timescales Climate Dynamics DOI 10.1007/s00382-014-2344-9.