
# Assessment of predictability of the C cycle in C-driven simulations

## *Deliverable 2.3*

**Authors: Acosta Navarro J. C., Bernardello R., Bopp L., Ilyina T., Li H., Mignot J., and Tourigny E.**

This project received funding from the Horizon 2020 programme under the grant agreement No. 821003.

# Document Information

| | |
|---|---|
| **GRANT AGREEMENT** | 821003 |
| **PROJECT TITLE** | Climate Carbon Interactions in the Current Century |
| **PROJECT ACRONYM** | 4C |
| **PROJECT START DATE** | 1/6/2019 |
| **RELATED WORK PACKAGE** | WP2 |
| **RELATED TASK(S)** | T2.3.1 and T2.3.2 |
| **LEAD ORGANIZATION** | MPG |
| **AUTHORS** | Acosta Navarro J. C., Bernardello R., Bopp L., Ilyina T., Li H., Mignot J., and Tourigny E. |
| **SUBMISSION DATE** | 2022-01-31 |
| **DISSEMINATION LEVEL** | PU |

# History

| DATE | SUBMITTED BY | REVIEWED BY | VISION (NOTES) |
|---|---|---|---|
| 31/01/2022 | H. Li | 4C Executive Board | |
| | | | |
| | | | |

**Please cite this report as: Acosta Navarro J. C., Bernardello R., Bopp L., Ilyina T., Li H., Mignot J. and Tourigny E. (2022)**, Assessment of predictability of the C cycle in C-driven simulations, D2.3 of the 4C project

# Table of Contents



# List of tables

# List of figures

# About 4C

**Climate-Carbon Interactions in the Coming Century** (4C) is an EU-funded H2020 project that addresses the crucial knowledge gap in the climate sensitivity to carbon dioxide emissions, by reducing the uncertainty in our quantitative understanding of carbon-climate interactions and feedbacks. This will be achieved through innovative integration of models and observations, providing new constraints on modelled carbon-climate interactions and climate projections, and supporting Intergovernmental Panel on Climate Change (IPCC) assessments and policy objectives.

# Executive Summary

This deliverable aims for assessing the predictability of global carbon cycle using 4C decadal prediction systems (EC-Earth3-CC, IPSL-ESM, MPI-ESM) in a multi-model framework together with other available systems initialized by the observed states of the physical climate. We demonstrate a predictive skill for the global ocean carbon sink of up to 6 years for some models with longer regional predictability horizons across single models. On land, a predictive skill of up to 2 years is primarily maintained in the tropics and extra-tropics enabled by the initialization of the physical climate. We further show that anomalies of atmospheric $CO_2$ growth rate inferred from natural variations of the land and ocean carbon sinks are predictable at lead time of 2 years and the skill is limited by the land carbon sink predictability horizon. An increased ensemble size of predictions is found to enhance the significance of predictive skill. Our results of predictability are crucial for understanding the variations and further guiding future predictions of the variable global carbon cycle.

## Keywords

$CO_2$ concentration-driven simulations, reconstruction, carbon cycle predictions

We have used eight available C-driven prediction systems including three 4C models, i.e., IPSL-CM6A-LR, MPI-ESM-LR, and MPI-ESM-HR, to investigate the predictability of global carbon cycle (Ilyina et al. 2021). This deliverable combines the results in Ilyina et al. (2021) and the newly available EC-Earth3-CC predictions to make a multi-model overview of predictability in global carbon cycle.

# 1 Introduction of C-driven prediction systems

## 1.1 EC-Earth3-CC

EC-Earth3-CC is the carbon cycle version of the EC-Earth3 Earth System Model (Döscher et al., 2021). The EC-Earth3-CC GCM (Global Climate Model) version 3.3 comprises three major components: the atmospheric model IFS (Integrated Forecasting System) CY36R4, the ocean model NEMO 3.6 (Madec, 2015) which also includes the LIM3 sea-ice model (Rousset et al., 2015) and OASIS3-MCT (Craig et al., 2017) that couples the main components. IFS is an operational global meteorological forecasting model developed and maintained by the European Centre of Medium-Range Weather Forecasts (ECMWF). NEMO is a state-of-the-art modelling framework for the ocean used for oceanographic research, operational oceanography, seasonal forecasting and climate research studies. EC-Earth3-CC includes additional components to represent the carbon cycle, also coupled via OASIS3-MCT: the LPJ-GUESS dynamic vegetation model (Lindeskog et al., 2013; Smith et al., 201), the PISCES ocean biogeochemistry model (Aumont et al., 2015) and the TM5 global atmospheric transport model (van Noije et al., 2014). LPJ-GUESS is used to simulate the evolution of the land vegetation, agricultural production and carbon fluxes as a function of climate and land-use, PISCES is used to simulate ocean biogeochemistry and $CO_2$ fluxes with the atmosphere and TM5 is used for atmospheric chemistry and transport of trace gases such as $CO_2$ and track surfaces fluxes of $CO_2$ between ocean and land/ocean as well as anthropogenic $CO_2$ emissions. In this project we use the T255-ORCA1 configuration, which corresponds to a spatial resolution of 80 km in the atmosphere/land and 100 km in the ocean, and 3x2 degrees with 10 vertical levels, $CO_2$-only configuration for TM5.

EC-Earth3-CC Initial conditions (ICs) for retrospective predictions with EC-Earth3-CC are obtained from three different procedures for ocean, land vegetation and atmosphere. Atmospheric forcings and ICs are obtained from the latest ECMWF reanalysis datasets (ERA-20C before 1950, ERA5-BE from 1950 to 1978 and ERA5 from 1979 onward). The ocean reconstructions to generate the initial conditions are performed with the same ocean models used in EC-Earth3-CC (NEMO3.6-PISCES-LIM3) by restoring SST and SSS, as well as 3D temperature and salinity below the mixed layer. We use the restoring timescale distribution of Sanchez-Gomez et al. (2016). At the surface, SST is restored using a feedback coefficient of -200 W/m2/K while the feedback parameter for freshwater fluxes is set at -750 mm/day. Below the mixed layer, the restoring timescale

varies between 30 days (up to 800m) and 3650 days (below 800m). Sub-surface nudging applied in the tropical band (15S-15N) is 10 times weaker to avoid spurious effects (Sanchez-Gomez et al. 2016). The reference dataset used for 3D nudging is EN4 (Good et al, 2013) and for surface restoring is the ECMWF Ocean Reanalysis System 5 (ORAS5; Zuo et al. 2019). Sea Ice, as well as ocean biogeochemistry, are let free to evolve in response to the constrained ocean physics. The atmospheric initial conditions are obtained from the ECMWF reanalysis datasets, which have been interpolated to the EC-Earth3 Standard Resolution configuration of T255/N128 in the horizontal and 91 vertical levels, using full-pos from OpenIFS cycle 40r1. The land vegetation initial conditions are obtained from a reconstruction using the EC-Earth3 LSM (Land Surface Model) forced by ECMWF reanalysis surface forcings. The EC-Earth3 LSM couples the OSM (Offline Surface Model), an offline version of H-TESSEL, which is the land surface component of the IFS (Integrated Forecast System), to the LPJ-Guess dynamic vegetation model and uses the same CMIP6 forcings for land-use, nitrogen and greenhouse gases as the coupled ESM version EC-Earth3-CC, thus allowing it to perform PI (Pre-industrial) spinups and historical simulations up to the present day. For more details please see the 4C project Deliverable 2.2 (Bernardello R. et al., 2020).

The EC-Earth-3CC uninitialized historical simulations comprise 10 members at time of writing. The initial conditions are currently available for the 1980-2021 period. The decadal prediction hindcast period is 1981-2020 with an ensemble size of 15 members, with each prediction initialized on Nov. 1$^{st}$ of the previous year spanning 7 years.

## 1.2 IPSL-CM6A-LR

The IPSL (Institut Pierre Simon Laplace) decadal prediction system used here is based on the IPSL-CM6A-LR version of the climate model described extensively in (Boucher et al., 2020). The component models include: LMDZ6 atmosphere (average 157km, 79 levels), NEMOv3.6STABLE ocean on the ORCA1 grid (nominal 1 with 75 vertical levels), LIM3 sea ice (on the same grid as the ocean) and ORCHIDEE (Cheruy et al., 2019) land (same grid as the atmosphere). The ocean biogeochemistry model used in IPSL-CM6A-LR is based on PISCESv2 (Aumont et al., 2015).

The uninitialized historical simulations comprise 32 members. The hindcasts are initialized from a global century long simulation in which anomalies of global EN4 sea surface temperature and Atlantic sea surface salinity presented by Reverdin et al. (2019) have been nudged into the climate model. The nudging procedure is described by Estella-Perez et al. (2020). There is no assimilation of subsurface ocean, sea ice or atmospheric observations. Hindcasts start each December during 1961-2014 and integrated for 10 years; 10 members are launched for each start date.

## 1.3 MPI-ESM

### 1.3.1 MPI-ESM-LR

MPI-ESM-LR is a low-resolution configuration of the Max Planck Institute Earth System Model (MPI-ESM1.1; Giorgetta et al. (2013)), on which the coupled model inter-comparison project phase 5 (CMIP5) simulations are based. The resolution of the ocean model MPIOM is about 150km with 40 vertical levels. The resolution of the atmosphere model ECHAM is T63 (200km) with 47 vertical levels. The ocean biogeochemistry component of MPI-ESM is represented by HAMOCC (Ilyina et al., 2013), and the land and vegetation component is represented by JSBACH. The decadal prediction system comprises 3 set of simulations, i.e., i) an ensemble of 10-member uninitialized historical simulations extended to the RCP4.5 scenario; ii) assimilation run by nudging the ocean 3-D temperature and salinity anomalies from the ECMWF ocean reanalysis system 4 (ORAS4) (Balmaseda et al., 2013) and the atmospheric 3-D full-field temperature, vorticity, divergence, and surface pressure ECMWF Re-Analysis ERA40 (Uppala et al., 2005) during the period 1960-1989 and ERA-Interim (Dee et al., 2011) during the period 1990-2014; iii) An ensemble of 10-member retrospective prediction simulations initialized from the assimilation which is close to the observations, the initialized prediction simulations run for 10 years starting annually from January 1st for the period 1961-2014. There is no direct assimilation of ocean biogeochemical data due to the limit of available data.

### 1.3.2 MPI-ESM-HR

MPI-ESM1.2-HR is based on a latest MPI-ESM model version 1.2 (Müller et al., 2018; Mauritsen et al., 2019), which is used for CMIP6 simulations. The major model development in the physical climate components relative to the CMIP5 model versions is the new radiation and aerosol parameterizations (Mauritsen et al., 2019). The representation of the land vegetation component is extended by including wild fires, multi-layer soil hydrology scheme, and nitrogen cycle. A major development to the ocean biogeochemistry is the implementation of cyanobacteria as additional phytoplankton specie for prognostic representation of nitrogen fixation, improved detritus settling and a number of other refinements. MPI-ESM1.2-HR is configured with grid spacings of 40 km in the ocean and T127 (100 km) in the atmosphere, with 40 ocean vertical levels and 95 atmospheric vertical levels, respectively. The assimilation in the MPI-ESM1.2-HR decadal prediction system is in general the same as in the MPI-ESM-LR prediction system for the atmosphere and the ocean, the difference for nudging is the inclusion of assimilation of sea-ice concentration from the National Snow and Ice Data Center (NSIDC) satellite observations (as described in Bunzel et al. (2016)). In addition, we run a pre-assimilation to spinup the ocean biogeochemistry for about 50 years before the assimilation so that the ocean biogeochemical processes slowly adjust to the new assimilated physical climate states (Li et al., 2019). The ensemble member for the initialized predictions and uninitialized historical of MPI-ESM1.2-HR simulations is 10. Note that the initialized 10-year long predictions of MPI-ESM1.2-HR system start annually from November 1st for the period 1960-2018.

**Table 1. Summary of $CO_2$ concentration-driven prediction systems based on 4C Earth system models (ESMs), i.e., EC-Earth3-CC, IPSL-CM6A-LR, MPI-ESM-LR, and MPI-ESM-HR.**

| ESMs | EC-Earth3-CC | IPSL-CM6A-LR | MPI-ESM-LR | MPI-ESM-HR |
|---|---|---|---|---|
| Resolution Atmosphere | T255, 91 levels | 2.5° x1.3°, 79 levels | T63, 47 levels | T127, 95 levels |
| Resolution Ocean | 1°, 75 levels | 1°, 75 levels | 1.5°, 40 levels | 0.4°, 40 levels |
| Initialization Atmosphere | ERA5 Full field | Indirect | ERA-40 and ERA-Interim: full field vorticity, divergence,log(p), T | ERA-40 and ERA-Interim: full field vorticity, divergence,log(p), T |
| Initialization Ocean | Offline NEMO-PISCES-LIM reconstruction, nudging full field EN4 subsurface T-S, and full field ORAS5 SST and SSS. | EN4 SST and Atlantic SSS | ORAS4 3D T-S anomalies | ORAS4 3D T-S anomalies, sea-ice concentration anomalies from NSIDC |
| Initialization Land | Indirect | Offline land reconstruction with ERA5 atmospheric forcings | Indirect | Indirect |
| Ensemble size | 15 (10 used in analysis) | 10 | 10 | 10 |
| Period of reconstruction | 1980-2021 | 1961-2014 | 1961-2014 | 1960-2018 |
| Retrospective predictions | Yearly from 1st Nov. for 7 years | Yearly from 1st Jan. for 10 years | Yearly from 1st Jan. for 10 years | Yearly from 1st Nov. for 10 years |
| External forcings | CMIP6 | CMIP6 | CMIP5 | CMIP6 |
| References | Döscher et al., (2021) | Boucher et al. (2020) | Giorgetta et al. (2013) | Mauritsen et al. (2019) |

# 2 Predictability of the carbon cycle using C-driven ESMs and carbon reconstructions

## 2.1 Air-sea and air-land $CO_2$ fluxes

To elaborate on the data independence of our predictability assessment, we first estimate the predictive skill with the respective reconstruction simulations to provide evidence towards potential predictability. In Figure 1 we quantify the predictive skill of the $CO_2$ fluxes and atmospheric $CO_2$ growth rate relative to the best-performing reconstructions.

The predictive skill of air-sea $CO_2$ flux (Figure 1A) is up to 7 years in NorCPM1, 5, 4 and 3 years of predictive skill is found in MPI-ESM1.2-HR, CanESM5/CESM-DPLE and IPSL-CM6A-LR. EC-Earth3-CC show high correlation until lead time of 6 years, but the significance of improved predictive skill relative to uninitialized run is not persistent. The ensemble size and model resolution matter for the predictive skill. For MPI-ESM1.2-HR, the predictive skill of 5 years with 10 ensemble members used in this study is higher than that of 3 years based on 5 available ensemble members presented in Li et al. (2019). Moreover, the high resolution version MPI-ESM1.2-HR has longer predictive skill of air-sea $CO_2$ flux than the low resolution version MPI-ESM-LR.

For air-land $CO_2$ flux (Figure 1B) the predictive skill is up to 3 years in EC-Earth3-CC. Both CanESM5 and NorCPM1 present a predictive skill of 2 years for the air-land $CO_2$ flux. MPI-ESM-LR and MIROC-ES2L show lower predictive skill of 1 year. This suggests potential extension of predictability beyond the limit of predictability in ENSO, which is on seasonal scale and largely regulates the air-land $CO_2$ flux variations.

## 2.2 Atmospheric $CO_2$ growth rate

We next examine effects of the global land and ocean carbon sink variations on the inferred variability and predictability of atmospheric $CO_2$ growth rate (Figure 1C). Note that all prediction systems available to us are forced with prescribed evolution of atmospheric concentrations of $CO_2$ and so the atmospheric compartments of those models do not respond to land or ocean $CO_2$ fluxes. Here, the detrended sum of the global land and ocean carbon fluxes serves as a diagnostic of variations in the temporal evolution of the atmospheric $CO_2$ growth driven by climate modulated variability of carbon sinks. We find predictability of variations in atmospheric $CO_2$ growth at lead times of up to 2 years in CanESM5, EC-Earth3-CC and NorCPM1 models, as indicated by higher correlations with reconstruction of the initialized simulations in comparison to the uninitialized ones. Given the higher amplitude of interannual air-land $CO_2$ flux variability, atmospheric carbon growth rate anomalies predominantly follow the land carbon sink evolution, and the ocean carbon sink acts to dampen the land modulated interannual variations of atmospheric $CO_2$.

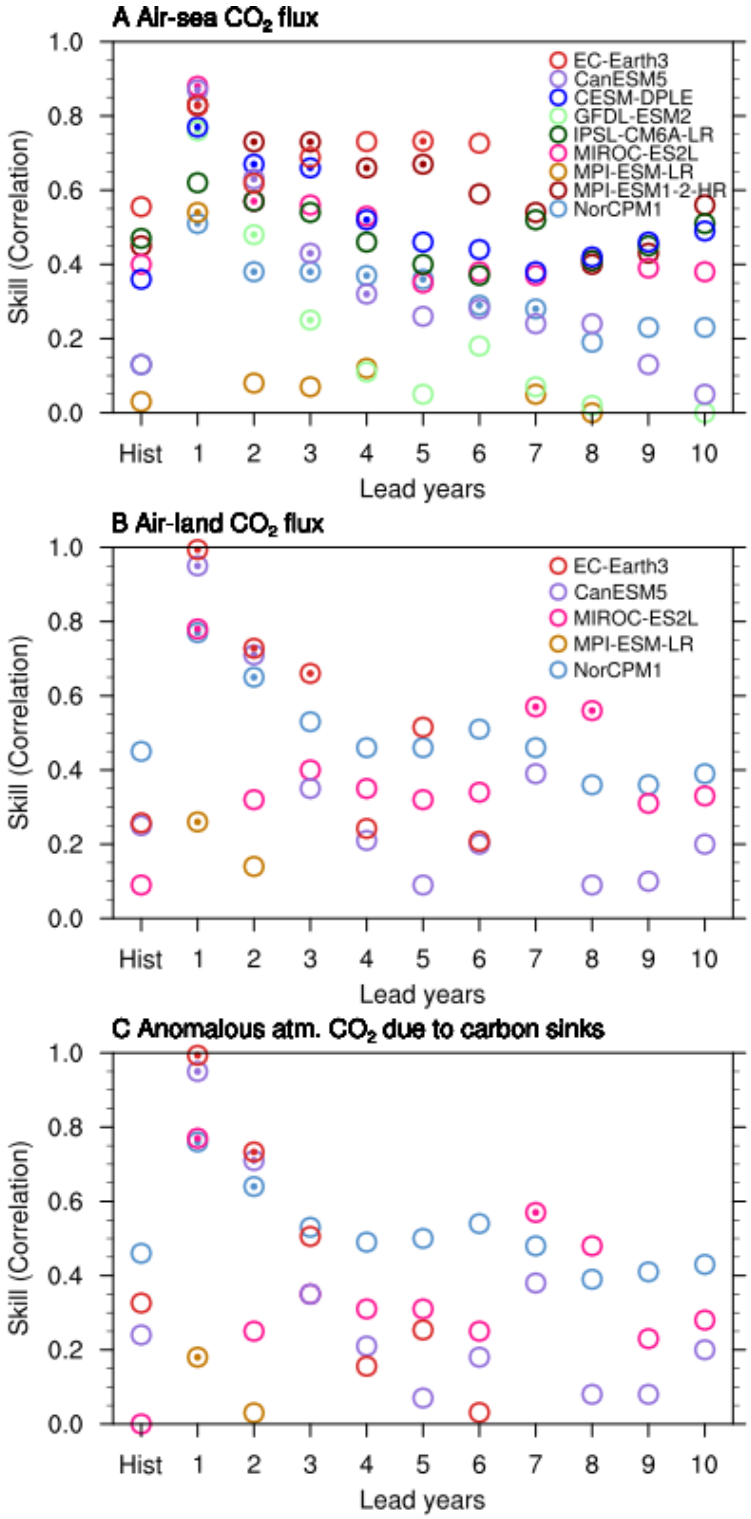


**Figure 1. Predictive skill of the detrended CO2 flux into the ocean (A), CO2 flux into the land (B), and variations in the inferred atmospheric CO2 growth (C). Predictive skill is quantified as anomaly correlation coefficients of the model simulations with respective model**

reconstructions. Note less model reconstructions for air-land CO2 flux are available as shown in B and C. Significantly improved predictive skill at 95% level for initialized over uninitialized simulations are marked with filled dots. Hist represents the uninitialized simulations.

## 2.3 Spatial distribution of predictability relative to reconstruction

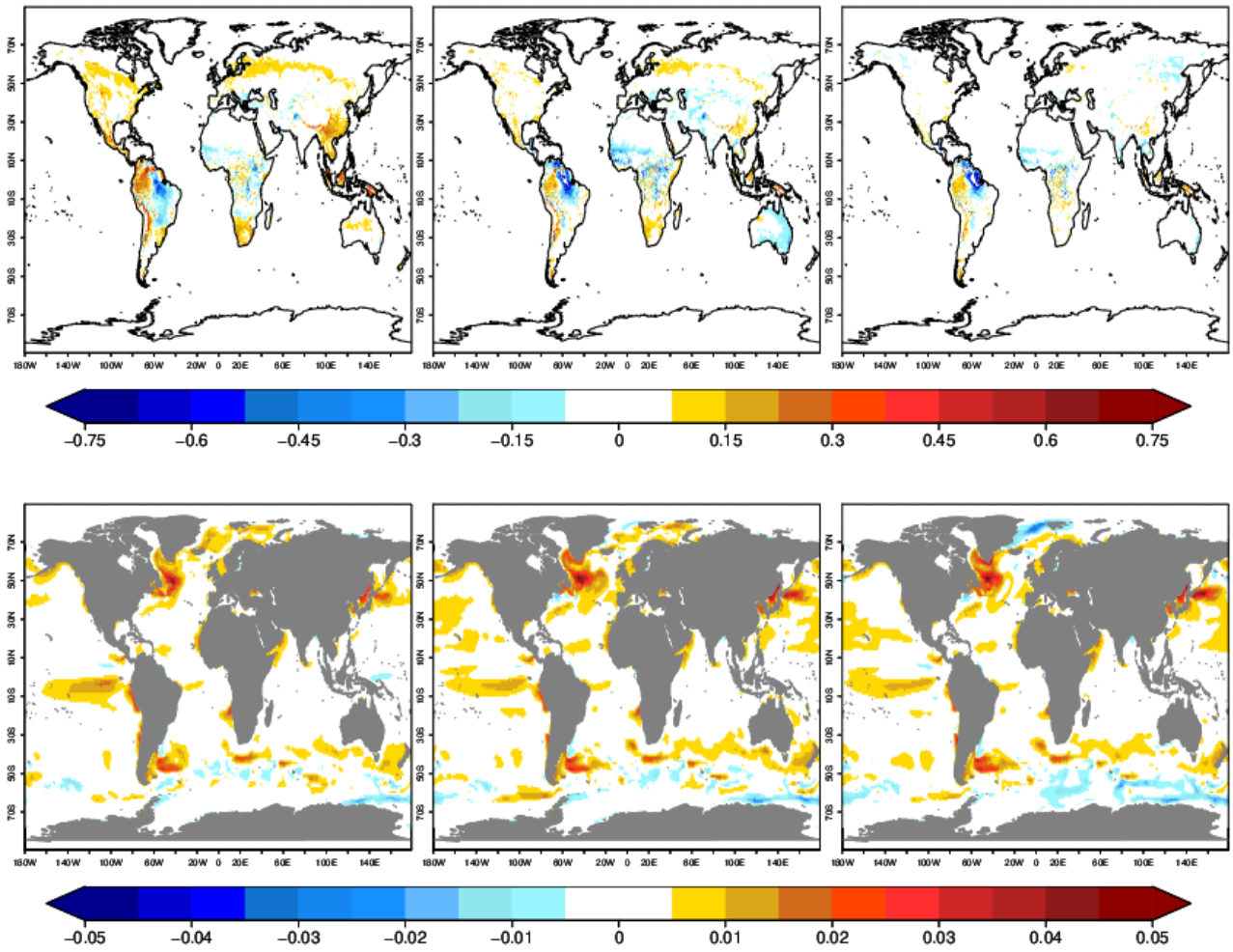


**Figure 2. Average bias of EC-Earth3-CC $CO_2$ flux [kg/m2/s C] of retrospective predictions relative to land/ocean reconstructions. Top panel: land, bottom panel: ocean, left: year 1, center: year 2, right: years 2-6. Dots indicate statistically non-significant values at the 95% confidence.**

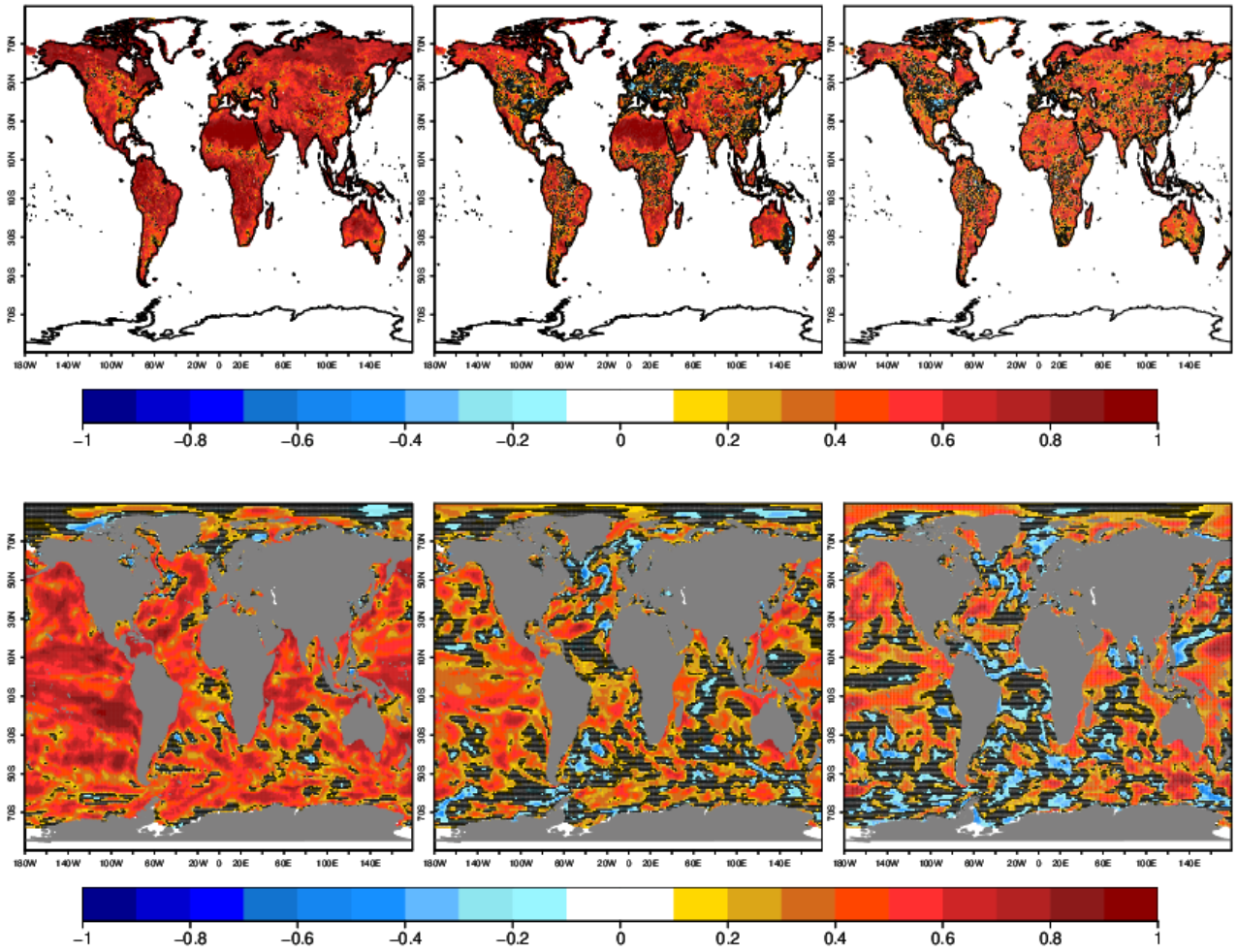


**Figure 3. ACC skill score of detrended EC-Earth3-CC $CO_2$ fluxes of retrospective predictions relative to land/ocean reconstructions. Top panel: land, bottom panel: ocean, left: year 1, center: year 2, right: years 2-6.**

In Figure 2 we show the mean biases of the EC-Earth3-CC hindcasts, compared to the land/ocean reconstructions. For the land, a strong negative bias appears over portions of the Amazon and subtropical Nothern Africa, and is likely related to a southward displacement of the ITCZ (due to widespread negative temperature anomalies over land of the Northern Hemisphere) leading to decreased precipitation (not shown) over these regions. Conversely a positive bias in precipitation in Southern Africa leads to a positive anomaly in air-land $CO_2$ flux. Positive air-land $CO_2$ flux anomalies over the Andes region are associated with a positive bias in precipitation. Positive biases in air-sea $CO_2$ flux are seen in the Northern Atlantic, Northwest Pacific and Southwest Atlantic likely due to negative surface temperature biases in those regions (not shown), that increase the carbon dioxide solubility with respect to the ocean reconstruction.

Spatial maps of ACC skill score of detrended $CO_2$ fluxes from EC-Earth3-CC hindcasts compared to reconstructions are shown in Figure 3. Air-land skill is globally positive for Year 1, but becomes negative from

year 2 over many areas of the Northern Hemisphere as well as the areas with a negative bias in South America and Africa. For the air-ocean fluxes, skill is positive over most areas for Year 1 but becomes more patchy from Year 2, although still predominantly positive.

# 3 Predictability of the carbon cycle using C-driven ESMs and carbon observations

For ocean carbon uptake, we use the Self-Organizing Map-Feed-Forward Network (SOM-FFN; Landschützer et al., 2015) observationally based product to quantify the predictability. For land carbon uptake, direct observational estimates capturing the regional and global temporal variability are not available, hence we use the Global Carbon Budget 2019 (GCB; Friedlingstein et al., 2019) carbon sinks estimates as a benchmark. Undergoing annual updates, GCB offers a comprehensive and temporally consistent time-series of stand-alone land and ocean carbon cycle model simulations forced with observed climate data or climate reanalysis and additional observational products (atmospheric $CO_2$, land cover change, etc.). We use a model mean from all available GCB models.

## 3.1 Air-sea and air-land $CO_2$ fluxes

We further assess predictability horizons of the global ocean and land carbon sinks, as well as of the diagnosed changes in atmospheric $CO_2$ growth represented by the lead years with improved predictive skill due to initialization (Figure 4). Predictive skill of the ocean carbon sink significantly improves with initialization up to lead year 5 against the SOM-FFN data product in MPI-ESM1.2-HR and EC-Earth3 and up to lead year 6 in CESM-DPLE and NorCPM1. A larger ensemble size of CESM-DPLE and NorCPM1 relative to the outputs from the other prediction systems maintains the predictive skill significance. Our previous study (Li & Ilyina, 2018) suggests that a large ensemble size is needed to capture decadal variations in the ocean carbon sink. Therefore, an increased ensemble prediction size could enhance the predictive skill of global carbon fluxes in other prediction systems, as well as in a multi-model ensemble.

Predictive skill due to initialization up to lead year 2 for land carbon sink verified against GCB estimates is found in CanESM5, EC-Earth3, IPSL-CM6A-LR, MPI-ESM, and NorCPM1 (Figure 4B). This skill, supported by higher coherence between GCB estimates and initialized simulations at lead time of 2 years in most models, goes well beyond a seasonal skill attainable in previous studies.

## 3.2 Atmospheric $CO_2$ growth rate

The atmospheric $CO_2$ growth rate changes induced by land and ocean carbon sink variations show predictive skill to lead year 2 (Figure 4) in the same models which have significant 2 years predictive horizons for the land carbon sink (i.e. in CanESM5, EC-Earth3, IPSL-CM6A-LR, MPI-ESM, and NorCPM1). Given the longer predictive horizons of the ocean carbon sink (in the models which provided output from both the ocean and the land biogeochemistry components), our results indicate that predictability of the atmospheric $CO_2$ growth in these models is limited by the land carbon sink predictability. Analogously, a previous study, based on a perfect model framework (Spring & Ilyina, 2020), demonstrates that the predictive skill of atmospheric $CO_2$ concentration of 3 years is dampened by land.

The predictive skill of $CO_2$ fluxes and atmospheric carbon growth relative to data (Figure 4) is in general similar but in some models and variables higher than the predictive skill relative to the respective reconstruction (Figure 3). This can partially refer to the 'signal-to-noise paradox' (Scaife and Smith, 2018), i.e., ensemble predictions generally show higher correlation with observed variability than with their own model simulations. In the meanwhile, the data products show more smoothed variations as in SOM-FFN data, and in the GCB2019 data much high frequency variabilities are canceled out by conducting multi-model mean.

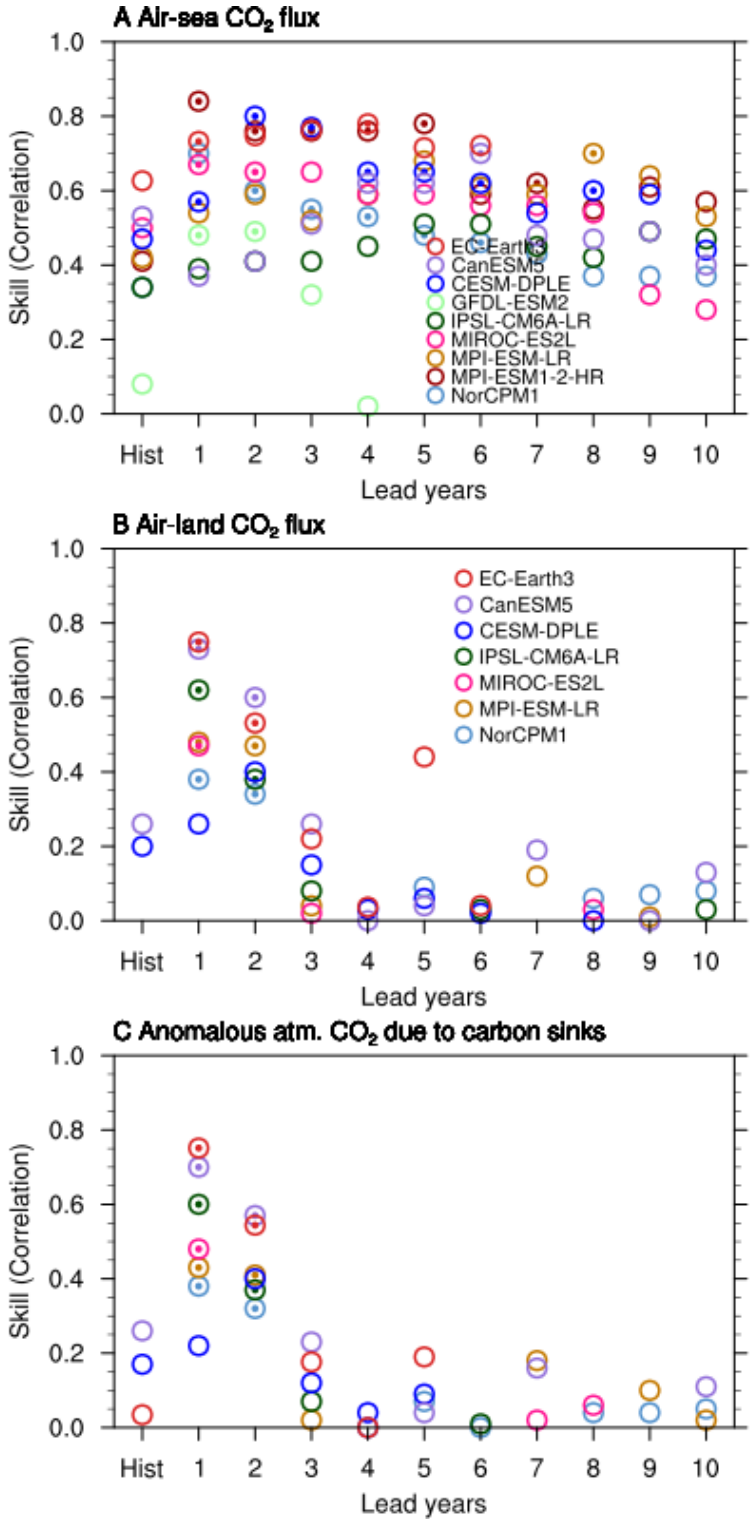


**Figure 4. Predictive skill of the detrended CO2 flux into the ocean (A), CO2 flux into the land (B), and variations in the inferred atmospheric CO2 growth (C). Predictive skill is quantified as anomaly correlation coefficients of the model simulations with the SOM-FFN observation-**

**based product for the air-sea CO2 fluxes (a), and with GCB2019 for the air-land CO2 flux and anomalous atmospheric CO2 due to carbon sinks. Significantly improved predictive skill at 95% level for initialized over uninitialized simulations are marked with filled dots. Hist represents the uninitialized simulations. Note that MIROC-ES2L and EC-Earth3 hindcasts start earliest from year 1980, so from lead year 4 the time period is shorter than 1982–2013. GCB, Global Carbon Budget; SOM-FFN, Self-Organizing Map-Feed-Forward Network. (Figure is redrawn based on Ilyina et al. 2021 and EC-Earth3 model 10 ensemble member outputs.)**

## 3.3 Spatial patterns of predictability horizons of $CO_2$ fluxes

To investigate if the multi-model prediction systems establish robust spatial predictability horizons in the carbon cycle despite using different initialization/assimilation methods, we examine predictability horizons due to initialization. This is represented by the lead time of years when correlations of the initialized simulations are larger than those in the uninitialized simulations. Despite different pattern among models (here we only show the 4C model results in Figure 5), we find rather consistent $CO_2$ flux predictability horizons established due to initialization in the different prediction systems in the hotspots of ocean carbon sink in the Southern Ocean, North Atlantic, and North Pacific.

In these ocean regions acting as carbon sink hotspots, the improved skill is retained for up to 9–10 years, thereby going beyond the predictability horizons of the physical climate variables, e.g. SST (Li et al., 2016; Séférian et al., 2014). Our previous findings (Li et al., 2019) suggest that temperature variations largely control shorter-term (<3 years) predictability of the ocean carbon sink, while longer-term (>3 years) predictability is associated with nonthermal drivers.

For air-land $CO_2$ fluxes, statistically significant improvements due to initialization are found in regions of the tropics (e.g. Amazon, West Africa) and extra-tropics (e.g. Middle East, US Great Plains, Eastern Russia). These prediction systems represent land carbon fluxes improved due to initialization at lead time of 2 years.

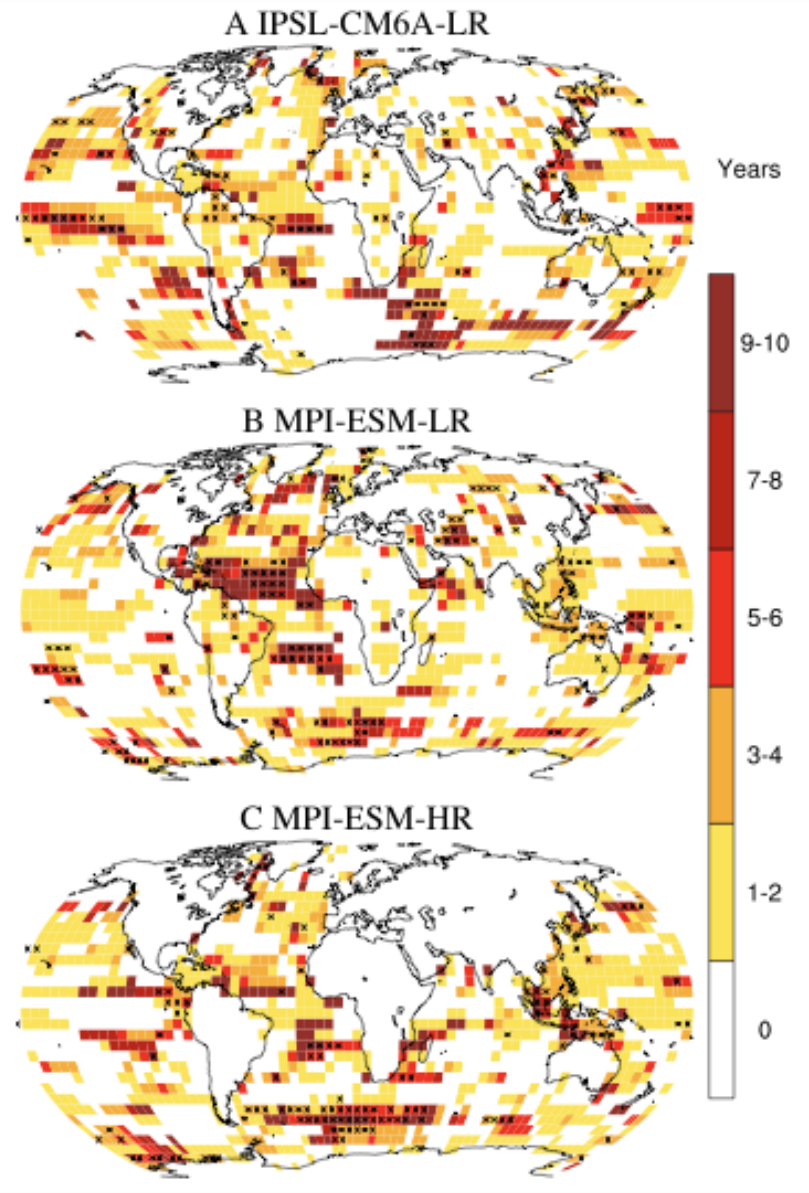


**Figure 5. Predictability horizon of the detrended CO2 flux into the ocean and the land in individual models, represented by the lead years with improved predictive skill due to initialization, i.e., when correlations in the initialized simulations are larger than 0 and also larger than those in the uninitialized simulations. Skill is quantified with anomaly correlation coefficient for the period 1982–2013. Predictive skill of the air-sea CO2 flux gained due to initialization is assessed against SOM-FFN, whereas for the air-land CO2 flux it is assessed against GCB. Crosses show significance at 95% level for the first 2 years. GCB, Global Carbon Budget; SOM-FFN, Self-Organizing Map-Feed-Forward Network. (Figure is redrawn based on Ilyina et al. 2021)**

# 4 Conclusions

We provide a first multi-model assessment of carbon cycle predictions with the ESM-based prediction systems initialized by the observed state of the physical climate, which is an important step toward skillful near-term predictions of the evolution of the land and ocean carbon sinks and the resulting variations in atmospheric $CO_2$ growth in response to climate variability and changes in anthropogenic carbon emissions.

We find improved predictive skill due to initialization in both ocean and land carbon sinks. Predictive skill due to initialization for the global air-sea $CO_2$ flux is up to 6 years in some models with indication of even higher regional skill in some regions. Representation of air-land $CO_2$ flux improved due to initialization in all models considered in this study. We demonstrate predictive horizons of up to 2 years in five out of the seven models considered in this study. As year-to-year variations in atmospheric $CO_2$ are largely determined by variations of the land carbon sink, the predictability horizon of 2 years found for the atmospheric $CO_2$ growth rate is maintained by predictability of air-land $CO_2$ flux.

An increased ensemble prediction size could enhance the predictive skill significance of global carbon fluxes in ESM prediction systems. Predictability of the global carbon cycle relative to reconstruction is in general in line with the predictability relative to observations and data products. In some models, the real predictability relative to observations is higher than that relative to the respective reconstruction, this suggest the potential to improve predictability by model itself together with quality data products to initialize models and to evaluate the predictive skill.

# 5 Publication

Ilyina, T., Li, H., Spring, A., Müller, W., Bopp, L., Chikamoto, M., Danabasoglu, G., Dobrynin, M., Dunne, J., Fransner, F., Friedlingstein, P., Lee, W.-S., Lovenduski, N., Merryfield, W., Mignot, J., Park, J.-Y., Séférian, R., Sospedra-Alfonso, R., Watanabe, M. & Yeager, S. (2021). Predictable variations of the carbon sinks and atmospheric CO2 growth in a multi-model framework. Geophysical Research Letters, 48: e2020GL090695. doi:10.1029/2020GL090695

# 6 References

Aumont, O., Ethe, C., Tagliabue, A., Bopp, L., & Gehlen, M. (2015). PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. Geoscientic Model Development Discussions, 8 (2).

Balmaseda, M. A., Mogensen, K., & Weaver, A. T. (2013). Evaluation of the ecmwf ocean reanalysis system oras4. Quarterly Journal of the Royal Meteorological Society, 139 (674), 1132-1161.

Bernardello R. et al. (2020), Initial conditions for retrospective predictions, D2.2 of the 4C project.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., . . . others (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. Journal of Advances in Modeling Earth Systems, 12 (7).

Bunzel, F., Notz, D., Baehr, J., Müller, W. A., & Fröhlich, K. (2016). Seasonal climate forecasts significantly affected by observational uncertainty of arctic sea ice concentration. Geophysical Research Letters, 43 (2), 852-859.

Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, E., Gastineau, G., . . . others (2019). Improved near surface continental climate in IPSL-CM6A-LR by combined evolutions of atmospheric and land surface physics. Journal of Advances in Modeling Earth Systems, e2019MS002005.

Craig, A., Valcke, S., and Coquart, L.: Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3. 0. GEOSCI MODEL DEV, 10(9). 2017.

Dee, D. P., Uppala, S. M., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., . . . others (2011). The era-interim reanalysis: Con guration and performance of the data assimilation system. Quarterly Journal of the royal meteorological society, 137 (656), 553-597.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arneth, A., Arsouze, T., . . . others (2021). The EC-Earth3 Earth System Model for the Climate Model Intercomparison Project 6.

Estella-Perez, V., Mignot, J., Guilyardi, E., Swingedouw, D., & Reverdin, G. (2020). Advances in reconstructing the AMOC using sea surface observations of salinity. Climate Dynamics, 1-18.

Friedlingstein, P., Jones, M. W., O'sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., ... & Zaehle, S. (2019). Global carbon budget 2019. Earth System Science Data, 11(4), 1783-1838.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M.,. . . others (2013). Climate and carbon cycle changes from 1850 to 2100 in MPIESM simulations for the Coupled Model Intercomparison Project phase 5. Journal of Advances in Modeling Earth Systems, 5 (3), 572-597.

Good, S. A., Matthew J. M., and Rayner, N. A. (2013). EN4: Quality Controlled Ocean Temperature and Salinity Profiles and Monthly Objective Analyses with Uncertainty Estimates. Journal of Geophysical Research: Oceans 118, no. 12 (2013): 6704–16. doi: 10.1002/2013JC009067.

Ilyina, T., Li, H., Spring, A., Müller, W., Bopp, L., Chikamoto, M., Danabasoglu, G., Dobrynin, M., Dunne, J., Fransner, F., Friedlingstein, P., Lee, W.-S., Lovenduski, N., Merryfield, W., Mignot, J., Park, J.-Y., Séférian, R., Sospedra-Alfonso, R., Watanabe, M. & Yeager, S. (2021). Predictable variations of the carbon sinks and

atmospheric CO2 growth in a multi-model framework. Geophysical Research Letters, 48: e2020GL090695. doi:10.1029/2020GL090695

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., & Nunez-Riboni, I. (2013). Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations. Journal of Advances in Modeling Earth Systems, 5 (2), 287-315.

Landschützer, P., Gruber, N., Haumann, F. A., Rödenbeck, C., Bakker, D. C., Van Heuven, S., ... & Wanninkhof, R. (2015). The reinvigoration of the Southern Ocean carbon sink. Science, 349(6253), 1221-1224.

Li, H., Ilyina, T., Müller, W. A., & Sienz, F. (2016). Decadal predictions of the north Atlantic $CO_2$ uptake. Nature Communications, 7(1), 1–7.

Li, H., Ilyina, T., Müller, W. A., & Landschützer, P. (2019). Predicting the variable ocean carbon sink. Science advances, 5 (4), eaav6471.

Lindeskog, M., Arneth, A., Bondeau, A., Waha, K., Seaquist, J., Olin, S., Smith, B. (2013). Implications of accounting for land use in simulations of ecosystem carbon cycling in Africa. Earth System Dynamics 4 (2), 385–407.

Madec, G. (2015). NEMO ocean engine. (chapter 3), France.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., . . . others (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1. 2) and its response to increasing co2. Journal of Advances in Modeling Earth Systems, 11 (4), 998-1038.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., . . .others (2018). A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM1. 2-HR). Journal of Advances in Modeling Earth Systems, 10 (7), 1383-1413.

Pohlmann, H., Smith, D. M., Balmaseda, M. A., Keenlyside, N. S., Masina, S., Matei, D., . . . Rogel, P. (2013). Predictability of the mid-latitude atlantic meridional overturning circulation in a multi-model system. Climate dynamics, 41 (3-4), 775-785.

Reverdin, G., Friedman, A. R., Cha k, L., Holliday, N. P., Szekely, T., Valdimarsson, H., & Yashayaev, I. (2019). North atlantic extratropical and subpolar gyre variability during the last 120 years: a gridded dataset of surface temperature, salinity, and density. part 1: dataset validation and RMS variability. Ocean Dynamics, 69 (3), 385-403.

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An Improved In Situ and Satellite SST Analysis for Climate. Journal of Climate, 15 (13), 1609-1625.

Rousset, C., Vancoppenolle, M., Madec, G., Fichefet, T., Flavoni, S., Barthélemy, A., Benshila, R., Chanut, J., Levy, C., Masson, S., Vivier, F. (2015). The Louvain-La-Neuve sea ice model LIM3.6: global and regional capabilities. Geoscientific Model Development 8 (10), 2991–3005.

Scaife, A. A., & Smith, D. (2018). A signal-to-noise paradox in climate science. npj Climate and Atmospheric Science, 1(1), 1-8.

Séférian, R., Bopp, L., Gehlen, M., Swingedouw, D., Mignot, J., Guilyardi, E., & Servonnat, J. (2014). Multiyear predictability of tropical marine productivity. Proceedings of the National Academy of Sciences, 111(32), 11646–11651.

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. Biogeosciences 11 (7), 2027–2054.

Uppala, S. M., Kallberg, P., Simmons, A., Andrae, U., Bechtold, V. D. C., Fiorino, M., . . . others (2005). The era-40 re-analysis. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 131 (612), 2961-3012.

van Noije, T.P.C., Le Sager, P., Segers, A.J., van Velthoven, P.F.J., Krol, M.C., Hazeleger, W., Williams, A.G., Chambers, S.D. (2014). Simulation of tropospheric chemistry and aerosols with the climate model EC-Earth. Geoscientific Model Development 7 (5), 2435–2475.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M. (2019). The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment, Ocean Sci., 15, 779–808, doi:10.5194/os-15-779-2019.