



# Reconstructions and predictions of the global carbon budget with an emission-driven Earth system model

Hongmei Li<sup>1</sup>, Tatiana Ilyina<sup>1</sup>, Tamas Loughran<sup>2,a</sup>, Aaron Spring<sup>1</sup>, and Julia Pongratz<sup>1,2</sup>

<sup>1</sup>Max Planck Institute for Meteorology, Hamburg, Germany

<sup>2</sup>Department of Geography, Ludwig-Maximilians-Universität, Munich, Germany

<sup>a</sup>now at: CSIRO Oceans and Atmosphere, Aspendale, Victoria, Australia

**Correspondence:** Hongmei Li (hongmei.li@mpimet.mpg.de)

Received: 27 July 2022 – Discussion started: 1 August 2022

Revised: 9 December 2022 – Accepted: 13 January 2023 – Published: 1 February 2023

**Abstract.** The global carbon budget (GCB) – including fluxes of CO<sub>2</sub> between the atmosphere, land, and ocean and its atmospheric growth rate – show large interannual to decadal variations. Reconstructing and predicting the variable GCB is essential for tracing the fate of carbon and understanding the global carbon cycle in a changing climate. We use a novel approach to reconstruct and predict the variations in GCB in the next few years based on our decadal prediction system enhanced with an interactive carbon cycle. By assimilating physical atmospheric and oceanic data products into the Max Planck Institute Earth System Model (MPI-ESM), we are able to reproduce the annual mean historical GCB variations from 1970–2018, with high correlations of 0.75, 0.75, and 0.97 for atmospheric CO<sub>2</sub> growth, air–land CO<sub>2</sub> fluxes, and air–sea CO<sub>2</sub> fluxes, respectively, relative to the assessments from the Global Carbon Project (GCP). Such a fully coupled decadal prediction system, with an interactive carbon cycle, enables the representation of the GCB within a closed Earth system and therefore provides an additional line of evidence for the ongoing assessments of the anthropogenic GCB. Retrospective predictions initialized from the simulation in which physical atmospheric and oceanic data products are assimilated show high confidence in predicting the following year’s GCB. The predictive skill is up to 5 years for the air–sea CO<sub>2</sub> fluxes, and 2 years for the air–land CO<sub>2</sub> fluxes and atmospheric carbon growth rate. This is the first study investigating the GCB variations and predictions with an emission-driven prediction system. Such a system also enables the reconstruction of the past and prediction of the evolution of near-future atmospheric CO<sub>2</sub> concentration changes. The Earth system predictions in this study provide valuable inputs for understanding the global carbon cycle and informing climate-relevant policy.

## 1 Introduction

The CO<sub>2</sub> fluxes between the atmosphere and the underlying surface, and therefore the atmospheric carbon growth rate, vary substantially on interannual to decadal time scales (Peters et al., 2017; Friedlingstein et al., 2019; Landschützer et al., 2019; Friedlingstein et al., 2020). These variations reflect the combined effects of the internal variability of the global carbon cycle (Li and Ilyina, 2018; Séférian et al., 2018; Spring et al., 2020; Fransner et al., 2020) and its responses to external forcings (McKinley et al., 2020).

To constrain the past global carbon budget (GCB) and facilitate its prediction and projection into the future, the Global Carbon Project (GCP) (Canadell et al., 2007) assesses the anthropogenic GCB – i.e., CO<sub>2</sub> emissions and their redistribution among the atmosphere, ocean, and land – every year since 2007. The annual updates of the GCB inform both the policy and society at large on the ongoing variations in the carbon cycle. This information will be critical in the decarbonization processes. This assessment is based on anthropogenic CO<sub>2</sub> emissions, observations of the atmospheric CO<sub>2</sub> concentration, and individual standalone model simulations of CO<sub>2</sub> fluxes for the ocean and land. The net air–

land CO<sub>2</sub> fluxes are the sum of the natural fluxes and the land-use-change-induced emissions. Within the framework of the GCP, the land use change emissions are based on a bookkeeping approach (Hansis et al., 2015). The standalone simulations of the land and ocean that produce air–land and air–sea CO<sub>2</sub> fluxes are forced by different observation or re-analysis data and their sum does not provide an estimate of the CO<sub>2</sub> fluxes that are consistent with changes in atmospheric CO<sub>2</sub>. Moreover, the accumulated CO<sub>2</sub> fluxes from these standalone model simulations do not exactly match the observations. Therefore, the GCB is not closed but ends up with a budget imbalance term of up to 2 PgCyr<sup>-1</sup> for some years, although the climatological mean value is small, 0.17 PgCyr<sup>-1</sup> (Friedlingstein et al., 2020), which hinders the full attribution of the global carbon cycle variations. A large part of the budget imbalance could also be attributed to the mismatch of net biome production between the dynamic global vegetation models (DGVMs) used in the GCBs and inversions that match the atmospheric CO<sub>2</sub> growth rate (Bastos et al., 2020).

Reconstruction of the variable GCB within a closed Earth system model (ESM) is of essential value in tracing the fate of carbon. In addition to assessing the GCB variations in the past, the GCP also makes a prediction of the GCB for the next year; however, this prediction is based on statistical approaches and it is not possible to trace the changes in carbon budget back to the physical processes. The decadal prediction systems based on ESMs (Marotzke et al., 2016) show a potential to reconstruct and predict the near-term global carbon cycle (Li et al., 2016; Spring and Ilyina, 2020). By assimilating observational products of physical variables, the decadal prediction systems are able to reproduce the variations in CO<sub>2</sub> fluxes as found in observation-based products. Decadal prediction systems can then use states from an assimilation simulation as initial conditions for further multi-year predictions of the global carbon cycle (Li et al., 2016, 2019; Lovenduski et al., 2019a, b; Ilyina et al., 2021). However, as of now, the state-of-the-art decadal prediction systems are typically forced with a prescribed atmospheric CO<sub>2</sub> concentration without an interactive carbon cycle, i.e., the effect of the changes in CO<sub>2</sub> fluxes are not reflected in the atmospheric CO<sub>2</sub> variations. With this conventional model setup, one can assess the air–land and air–sea CO<sub>2</sub> fluxes but not the resulting variations in atmospheric CO<sub>2</sub> concentration and growth.

Prediction systems have proven their skill in predicting air–sea and air–land CO<sub>2</sub> fluxes (Ilyina et al., 2021). For the first time, we extend our previously concentration-driven prediction system to an emission-driven system. The emission-driven system takes into account the interactive carbon cycle and therefore determines atmospheric CO<sub>2</sub> prognostically and predicts atmospheric CO<sub>2</sub> variations. In this study, we assess the GCB in a simulation with assimilated observational products into the Max Planck Institute Earth System Model (MPI-ESM) and further estimate the predictive skill relative to the GCB from 2019 (GCB2019, Friedlingstein

et al., 2019) for CO<sub>2</sub> fluxes and changes in atmospheric CO<sub>2</sub> (Dlugokencky and Tans, 2020).

The assimilation simulation is designed to reconstruct the evolution of the Earth system in the real world by incorporating essential fields from observational products into the MPI-ESM. The reconstruction from the fully coupled model simulation (henceforth known as simply the assimilation simulation) enables the representation of the GCB within a closed Earth system. Therefore, by construction, this approach avoids the budget imbalance term arising from the need to balance carbon fluxes from standalone models and observations. Our reconstructions of the carbon budget provide an additional and novel estimate. The assimilation simulation's states, which are close to the real world through constraints from observations and data products, are used to start the initialized simulations that predict the GCB changes. These initialized predictions are expected to capture the evolution of climate and carbon cycle more realistically than freely evolving uninitialized simulations, due to their improved initial conditions from reconstruction. In prediction studies, the term “uninitialized” refers to simulations that are not initialized from states constrained by observations or data products. The novel approach of reconstructing and predicting the GCB with an ESM-based prediction system has the ability to provide complementary estimates of the terms of the GCB for use in the assessments by the GCP.

## 2 Materials and methods

### 2.1 Model and simulations

We use the MPI-ESM1.2-LR model (Mauritsen et al., 2019), which is the low-resolution version of the MPI-ESM used for the sixth phase of the Coupled Model Intercomparison Project (CMIP6). The atmospheric horizontal resolution has a spectral truncation at T63 (approximately 200 km or 1.88° grid spacing at the Equator) with 47 vertical levels. The resolution of the ocean model MPIOM (Marsland et al., 2003) is about 150 km with 40 vertical levels. The ocean biogeochemistry component of the MPI-ESM is represented by HAMOCC (Ilyina et al., 2013; Paulsen et al., 2017), and the land and vegetation components are represented by JSBACH (Reick et al., 2021).

Similar to our previous prediction system (Li et al., 2016, 2019), we performed three sets of simulations (see Fig. 1 and Table A1) to investigate the ability of our model to reconstruct and predict the GCB: (i) uninitialized freely evolving historical simulations, (ii) an assimilation simulation (also referred to as reconstruction) performed by assimilating the observational signal of climate variations into the model, and (iii) initialized simulations (also referred to as hindcasts or retrospective predictions) starting from initial states obtained from the assimilation simulation. The assimilation run is needed for the initialized prediction simulations, and the uninitialized simulations provide a reference to com-

pare to and assess the improved predictability due to initialization.

The major difference relative to the previous system (Li et al., 2016, 2019) is that the new prediction system is based on emission-driven simulations, which are forced by CO<sub>2</sub> emissions instead of prescribed atmospheric CO<sub>2</sub> concentration. In this way, the atmospheric CO<sub>2</sub> concentration evolves in response to the magnitude and sign of the air–land and air–sea CO<sub>2</sub> fluxes. We use the CMIP6 (Eyring et al., 2016) historical emissions forcing for our simulations, and for simulations extended to 2099 we use the emissions from the SSP2-4.5 scenario (Jones et al., 2016). While the fossil fuel emissions are prescribed, the land-use-change-induced emissions are simulated interactively in our ESM and driven with the Land-Use Harmonization 2 (LUH2) forcing (Hurtt et al., 2020). We use transient land use transitions rather than land use states and include natural disturbances with dynamic vegetation (Reick et al., 2021). An ensemble of 10 members is run for the uninitialized historical and initialized prediction simulations. The uninitialized ensembles are generated by starting from a different year of the pre-industrial control simulation (the model reached equilibrium as shown in the time series of ocean net primary production and CO<sub>2</sub> fluxes from the control simulation in Fig. A1). The individual members of an initialized ensemble are generated with 1 d lagged initializations from a given branching point of the assimilation simulation, i.e., initialized from 31 October and 1 November–9 November. Note that the initialized 5-year-long predictions start annually from 1 November for the period 1960–2018. Figure 1 illustrates the evolution of the atmospheric carbon growth rate in uninitialized, assimilation, and initialized simulations. More details of the simulations are summarized in Fig. 1 and Table A1.

## 2.2 Assimilation methods

Similar to our previous concentration-driven decadal prediction systems (Li et al., 2019), the assimilation is done by nudging the simulated ocean 3-D temperature and salinity anomalies towards the ECMWF ocean reanalysis system 4 (ORAS4) (Balmaseda et al., 2013). Additionally, we nudge the simulated values towards atmospheric 3-D full-field temperature, vorticity, divergence, and log of surface pressure from ECMWF Re-Analysis ERA40 (Uppala et al., 2005) during the period 1959–1979 and ERA-Interim (Dee et al., 2011) during the period 1980–2018. The sea ice concentration is nudged towards the National Snow and Ice Data Center (NSIDC) satellite observations (as described in Bunzel et al., 2016). The nudging is applied at every model time step but with a different relaxation time, i.e., a relatively longer relaxation time of 10 d is used for the ocean temperature and salinity, and shorter relaxation times of 6, 24, and 48 h are used for the atmospheric vorticity, temperature and pressure, and divergence, respectively. The chosen variables for assimilation and their respective relaxation time are selected

based on previous investigations of decadal climate predictions based on the MPI-ESM (Marotzke et al., 2016). Direct assimilation of the carbon-cycle-related variables is not included because of the limited available data; instead, we found that the global carbon cycle is well captured by assimilating only physical variables (Li et al., 2016, 2019; Lovenduski et al., 2019b, a; Ilyina et al., 2021). Furthermore, a recent study based on a perfect-model framework (i.e., based on simulations in which the model tries to predict itself) revealed that direct assimilation of the global carbon cycle only brings trivial improvement to the predictive skill of the global carbon cycle (Spring et al., 2021). To avoid spurious upwelling in the equatorial region caused by assimilation (Park et al., 2018), we exclude the equatorial band of 5° S–5° N from being nudged towards observation-based ocean data.

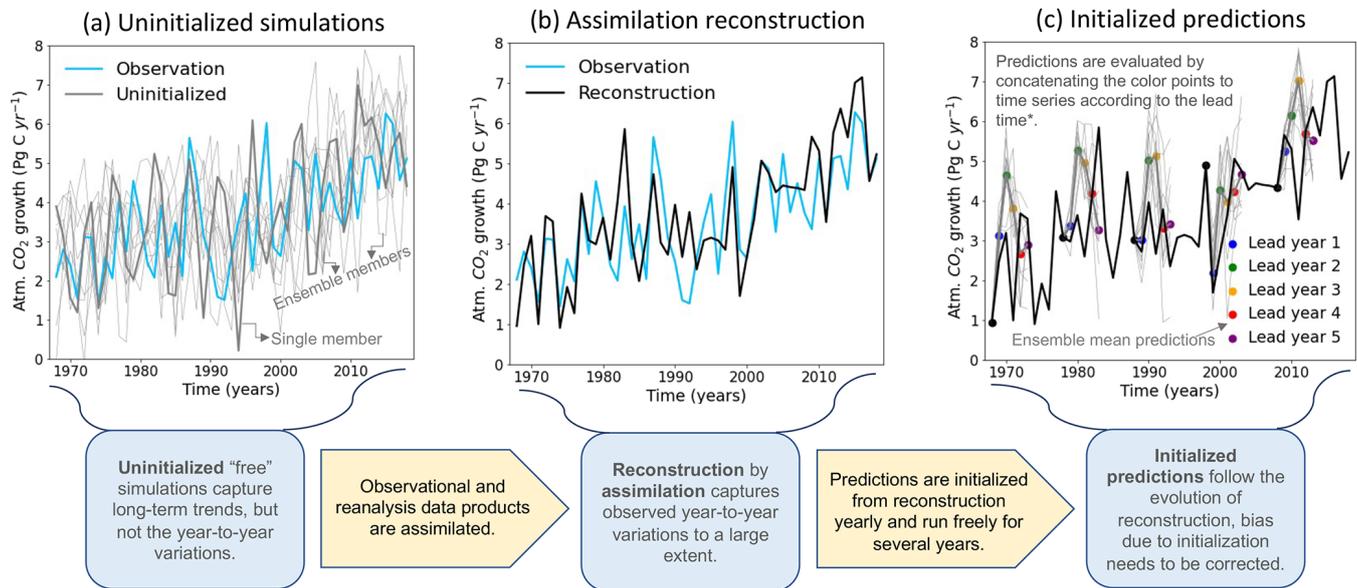
## 2.3 Carbon budget decomposition with CBALONE simulations

The GCB from GCP is decomposed into five terms plus an imbalance term: the two emission terms from fossil fuel and land use changes and the three sink terms for the natural terrestrial sink, ocean sink, and atmospheric growth on annual timescales. The fossil fuel emissions are prescribed as forcing, and the terrestrial and ocean carbon sinks and atmospheric growth terms are simulated and can therefore be directly derived from the ESM. However, only the net land–atmosphere exchange is directly deducible from an ESM, which is the sum of land use change emissions and the natural terrestrial sink. In order to separate the two land-related fluxes, we use a standalone component of JSBACH called CBALONE as a diagnostic for a direct comparison with the land-use emissions term from the GCP (Friedlingstein et al., 2019). CBALONE is forced by the MPI-ESM daily outputs including 2 m air temperature, soil temperature, precipitation, net primary productivity (NPP) per plant functional type (PFT), leaf area index (also per PFT), and maximum wind. We run two parallel simulations, i.e., one with anthropogenic land use changes and another without those changes, differencing the two simulations results in the land-use-change-induced emissions from the land sink. More details on this method of separating the land-use-change-induced emissions can be found in Loughran et al. (2021).

## 2.4 Predictive skill quantification

The focus of this study is on global mean variations in atmospheric CO<sub>2</sub> and globally integrated air–sea and air–land CO<sub>2</sub> fluxes on annual timescales. The initialized simulations are investigated according to their lead time, i.e., for how many model years they have been freely integrated after restarting from the assimilation simulation. The time series of initialized simulations at a lead time of 1 year (2, 3, 4, and 5 years) combine the first year (second, third, fourth, and fifth year) predictions from initialized simulations of all the start-

## Illustration of decadal prediction system based on an Earth system model



**Figure 1.** Illustration of the decadal prediction system based on the MPI Earth system model. The illustrated time series of the atmospheric CO<sub>2</sub> growth rate shows annual means from model simulations plotted together with observations from the Global Carbon Project (GCP). We conduct three sets of simulations (from left to right in sequential order): (a) uninitialized “free” simulations, which are the same as the freely evolving Coupled Model Intercomparison Project (CMIP) historical type simulations, (b) an assimilation simulation to reconstruct the evolution of the climate and carbon cycle towards the real world by nudging the model towards observation and reanalysis data during its integration, and (c) initialized predictions started from reconstruction states produced by the assimilation simulation and integrated freely (i.e., no nudging towards observations) for 5 years. The time series in (a) shows that the uninitialized simulations capture the long-term trend well, but the year-to-year variations are out of phase with the observations. The time series in (b) shows that the assimilation simulation forces the variations in the uninitialized freely run simulation towards the real world and results in a reconstruction that is closer to the observations. Panel (c) shows the reconstruction together with the 5-year-long initialized predictions (i.e., hindcasts). To make the illustration more clear, only predictions with starting years at 10-year intervals are shown.

ing years from 1960–2018. Therefore, the time series at lead time of 1 year (2, 3, 4, and 5 years) corresponds to the period 1961–2019 (1962–2020, 1963–2021, 1964–2022, and 1965–2023). Illustration of how the time series are concatenated is shown in Fig. 1C. The analyses of predictive skill quantification are based on the combined time series. Bias correction is an unavoidable topic for decadal predictions due to an initial shock, which varies with lead time (Boer et al., 2016; Meehl et al., 2021). The decadal prediction studies mostly present anomalies with a focus on variations by removing the climatological mean and/or trend bias due to model drift caused by the initialization of the model based on observations. The anomalies are calculated relative to the respective climatology according to the lead time (Boer et al., 2016; Meehl et al., 2021). To infer predictions of absolute values of the atmospheric CO<sub>2</sub> concentration, the respective anomalies from the predictions are added to the best estimates of climatology and trend from data; here the atmospheric CO<sub>2</sub> observations from NOAA\_GML (Dlugokencky and Tans, 2020) are used.

The predictive skill is quantified by the anomaly correlation coefficient, and the anomalies are calculated by removing the respective climatological mean state. In that sense, the climatological mean bias is removed and the coherence reflects the multi-year variations for which we evaluate the predictions. Here the climatological mean state is based on the ensemble mean of the focus time period: 1970–2018 for Figs. 1–6 and the last 10 years for Figs. 7–8. We exclude the first 12 years, i.e., 1958–1969, from the analyses and focus on the period from 1970–2018 because the assimilation in the first decade is affected by model adjustment. As an example, the spatial pattern of climatological mean ocean net primary production and phosphate nutrient concentration are shown in Fig. A2 in comparison with the respective observations. For the atmospheric CO<sub>2</sub> concentration, which has high correlations close to 1 with observations because of the coherent linear trends, we have also added the root-mean-square error (RMSE) metric to investigate the added value of assimilation and initialization. In this study, the significance of the predictive skill is tested with a nonparametric bootstrap approach (Goddard et al., 2013). The analyses are

based on annual mean data with a focus on the frequency of interannual to multi-year variations.

### 3 Reconstruction of the global carbon budget

By incorporating observation-based information, the assimilation simulation from the decadal prediction system based on the MPI-ESM captures the evolution of the GCB as well as the climate in observations. The time series of carbon fluxes from the MPI-ESM assimilation simulation in comparison to the data and suite of simulations from GCB2019 are shown in Fig. 2.

The CO<sub>2</sub> emissions from fossil fuels and industry are generally consistent with those from GCB2019 but with a slight difference in the 1960–1990s since the assimilation simulation uses the CO<sub>2</sub> emission forcing provided by CMIP6 for historical and SSP2-4.5 simulations. This highlights the uncertainty in the CO<sub>2</sub> forcing, which affects the change in the simulated atmospheric CO<sub>2</sub> concentration as it is a cumulative quantity. The CMIP6 CO<sub>2</sub> emission forcing yields 8.20 PgC higher cumulative emissions than those from the GCB2019, which is equivalent to a difference of atmospheric CO<sub>2</sub> of about 1.93 ppm assuming that 50 % of the emissions stay in the atmosphere (i.e., by dividing 4.10 PgC with a factor of 2.124 PgC ppm<sup>-1</sup>, Ballantyne et al., 2012). This discrepancy in CO<sub>2</sub> emissions might explain to some extent why the simulated atmospheric CO<sub>2</sub> concentration is a few parts per million higher than the NOAA\_GML observations (Dlugokenky and Tans, 2020) (Fig. A3). However, this small difference of a few parts per million in atmospheric CO<sub>2</sub> concentration magnitude does not noticeably affect the interannual variations in CO<sub>2</sub> fluxes and the corresponding atmospheric carbon increment (see Fig. 2d–f).

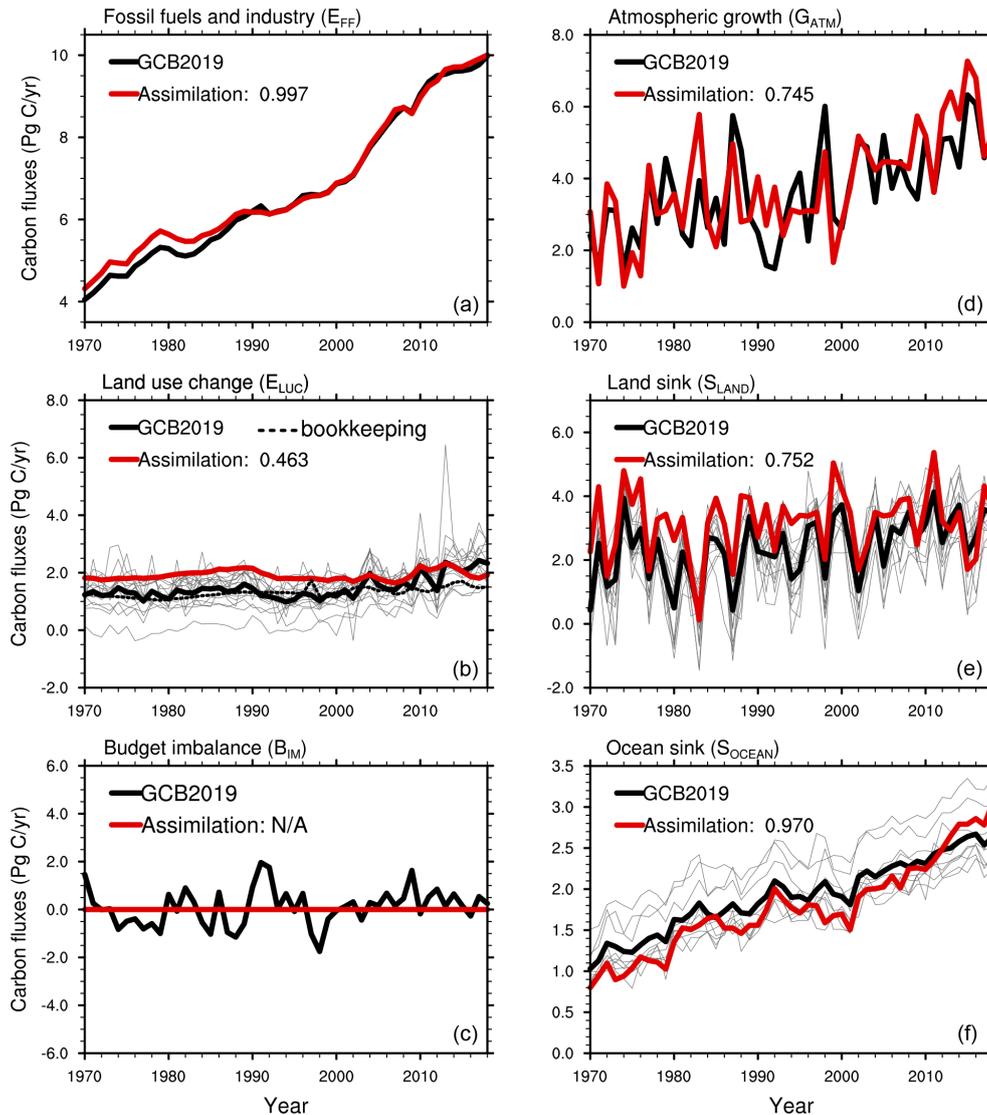
The land-use-change-induced emissions diagnosed by CBALONE are within the range of GCB2019 multi-model (including JSBACH) simulations from Dynamic Global Vegetation Models (DGVMs) (Fig. 2b). The estimates from bookkeeping models show smaller variations than those produced by the DGVMs. Note that the GCP uses the bookkeeping approach for the land-use emissions term. The term bookkeeping implies that carbon fluxes are determined from area changes in vegetation types with different vegetation and soil carbon densities, and specific response curves characterizing the evolution of decay of deforested biomass and recovery of natural vegetation thereafter. Biomass and soil carbon densities may be based on recent observations or models but are generally kept fixed in time, i.e. the effect of changes in environmental conditions are not accounted for. The DGVMs by contrast (which are used to provide only an uncertainty range around the bookkeeping models in the GCBs from the GCP) calculate land use emissions under transient environmental conditions. This implies first that interannual variability in bookkeeping models is only driven by land use change but not by climate variability, which makes the DGVM esti-

mates of land use change (LUC) emissions in general more variable from year to year than the bookkeeping estimates. Second, the DGVM-based land-use emissions estimates include the so-called “loss of additional sink capacity” (Pongratz et al., 2014), which refers to the carbon that could have been stored in forests additionally over the course of history (e.g., due to the “CO<sub>2</sub> fertilization” effect) had these forests not been cleared by the expansion of agriculture and forestry. This loss of additional sink capacity generally increases over time and amounts to about 40 % ( $0.8 \pm 0.3$  PgC yr<sup>-1</sup>) of the emissions calculated by the DGVMs over 2009–2018 (Obermeier et al., 2021). This explains why DGVM estimates in Fig. 2b show higher emissions than bookkeeping estimates in recent decades. The DGVM- and expert-based uncertainty range around the GCP bookkeeping estimates for LUC emissions is large and MPI-ESM-based land use change emission estimates have been found to be at the high end of the GCP estimates for all decades by Loughran et al. (2021), consistent with our findings.

The annual assessment from GCP has a budget imbalance term. This is because the individual budget terms are based on separate measurements, together with ocean and land model simulations, which are not linked to each other in an internally consistent manner (Friedlingstein et al., 2019). In this study, we assimilate atmosphere and ocean data products within a fully coupled ESM that considers their interactions. The assimilation ensures the evolution of the carbon cycle and climate towards the real world, and in contrast to the GCP, the budget is closed within the Earth system, i.e., no budget imbalance occurs by design (Fig. 2c). Therefore, the assimilation simulation based on a fully coupled ESM enables better attribution of the GCB variations than when an imbalance is present. The current method used by the GCP (Friedlingstein et al., 2019) is based on the directly measured atmospheric CO<sub>2</sub> increment and has the advantage of representing the actual evolution of atmospheric CO<sub>2</sub>. Our ESM-based assimilation shows a high correlation of 0.75 with the atmospheric CO<sub>2</sub> measurements but still needs to be improved. Further efforts are required to constrain the atmospheric CO<sub>2</sub> from observations.

Atmospheric carbon growth rate and carbon fluxes are reasonably well reproduced in emission-driven assimilation with prognostic atmospheric CO<sub>2</sub> (Fig. 2d–f). The atmospheric carbon growth and the land carbon sink show more pronounced variations on interannual timescales; however, the ocean carbon sink has more pronounced variations on decadal timescales. These variations are captured in the assimilation, with high correlations between the results from the assimilation simulation and the GCB2019 of 0.75, 0.75, and 0.97 for the atmospheric growth, air–land CO<sub>2</sub> fluxes, and air–sea CO<sub>2</sub> fluxes, respectively.

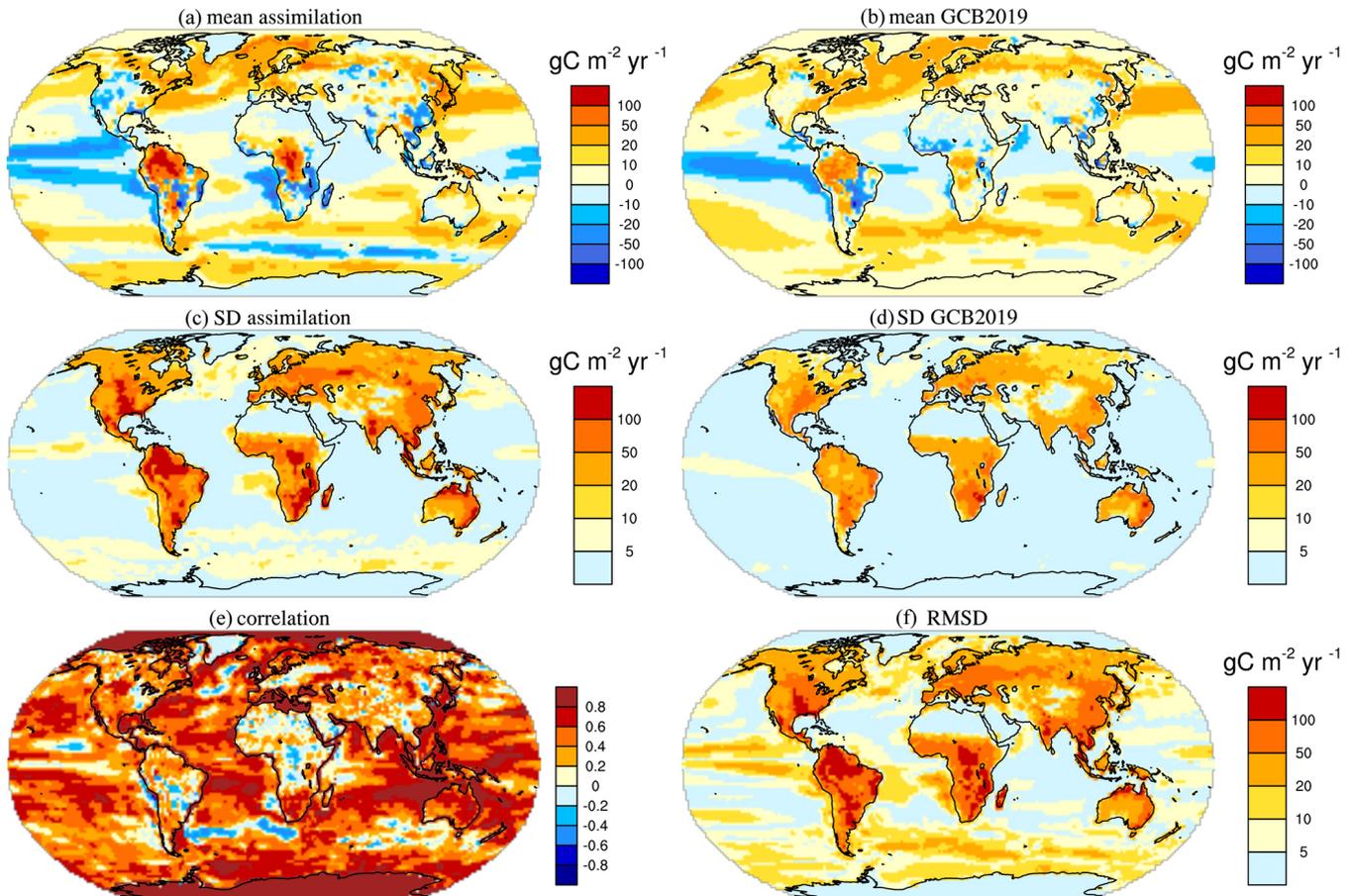
The spatial distribution of climatological mean CO<sub>2</sub> fluxes, their variability expressed as standard deviation, and the comparison in carbon fluxes between GCB2019 and the MPI-ESM assimilation are shown in Fig. 3. The mean states



**Figure 2.** Time series of (a) fossil fuel and industry CO<sub>2</sub> emissions ( $E_{FF}$ ), (b) emissions from land use change ( $E_{LUC}$ ), (c) the budget imbalance ( $B_{IM}$ ) that is not accounted for by the other terms, (d) atmospheric carbon growth rate ( $G_{ATM}$ ), (e) the natural terrestrial carbon fluxes ( $S_{LAND}$ ), and (f) air–sea CO<sub>2</sub> fluxes ( $S_{OCEAN}$ ) from MPI-ESM1.2-LR assimilation in comparison to the Global Carbon Budget (GCB 2019, Friedlingstein et al., 2019). Emissions (a, b) are positive into the atmosphere, while sinks (d, e, f) are positive into their respective compartments. A positive  $B_{IM}$  means a higher sum of emissions than sinks. The thin grey curves in (b), (e), and (f) show individual GCB standalone model results. The numbers in the legend show the correlation coefficients between carbon fluxes from the assimilation simulation and GCB2019.

show a CO<sub>2</sub> influx into the ocean and land in the mid- to high-latitudes, and outgassing into the atmosphere in tropical areas, especially over the tropical Pacific (Fig. 3a–b). The variability of CO<sub>2</sub> over land is larger than that over the ocean; and the magnitude of variability is larger in the assimilation simulation than in the GCB2019 (Fig. 3c–d). This is expected as the GCB2019 is a multi-model mean estimate and therefore smooths out part of the high-frequency variability. The correlation of CO<sub>2</sub> fluxes between the assimilation simulation and GCB2019 is high over the ocean, while the correlation is relatively low over the land (Fig. 3e). The root-mean-

square deviation (RMSD) scales with the magnitude of carbon fluxes, i.e., with larger values on land than over ocean (Fig. 3f). The large RMSD, especially over land, is because of the relatively low coherence of CO<sub>2</sub> fluxes and the larger values of CO<sub>2</sub> fluxes in the MPI-ESM single model simulation than in a smoothed magnitude of fluxes in GCB2019 from the multi-model mean simulations. The difference in the magnitude of fluxes between assimilation and GCB2019 data is more prominent in local areas (Fig. 3a–d) than in the global average (Fig. 2e).



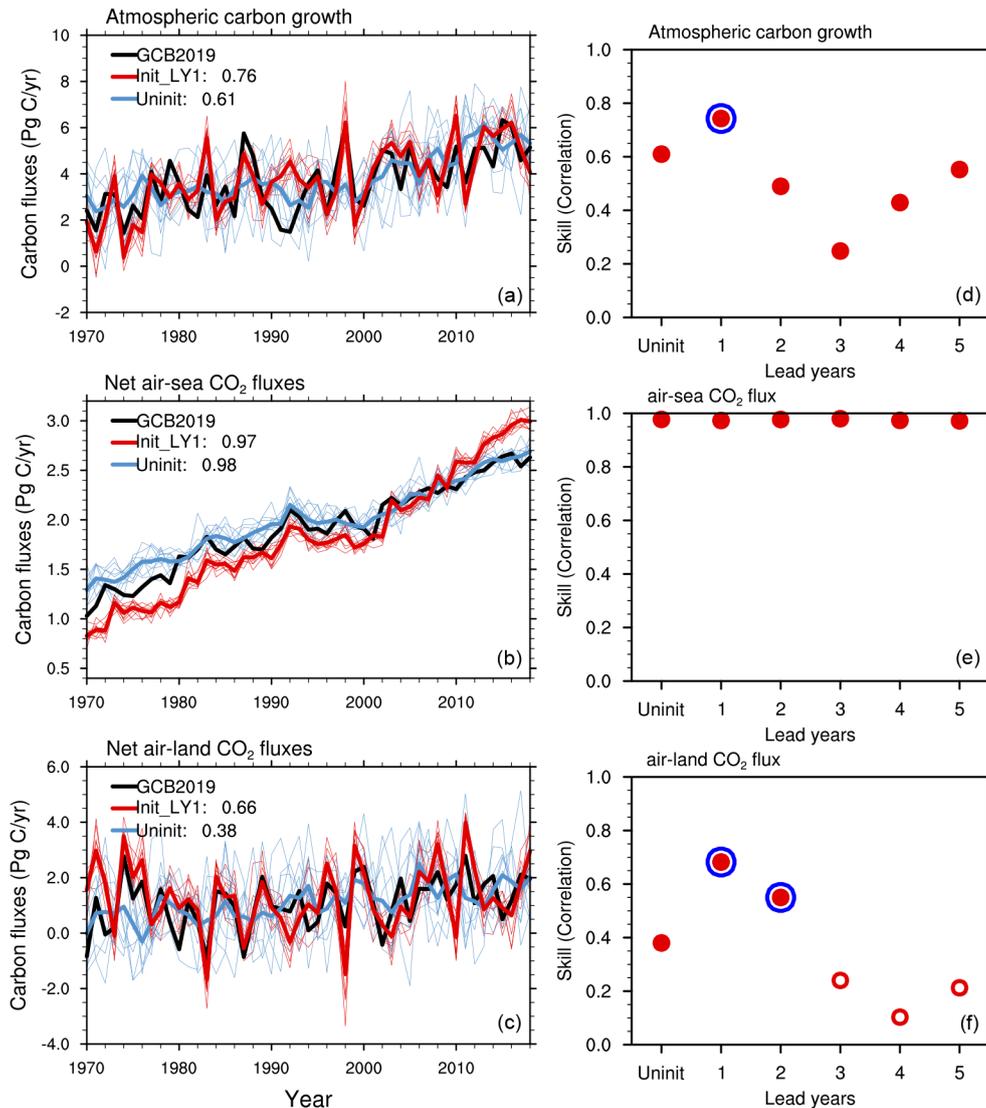
**Figure 3.** Spatial distribution of the CO<sub>2</sub> fluxes from model assimilations compared to GCB2019. Climatological mean CO<sub>2</sub> fluxes into the land and ocean from the atmosphere in assimilation (a) and Global Carbon Budget (GCB 2019, Friedlingstein et al., 2019) estimates (b). Temporal variability, i.e., standard deviation, of CO<sub>2</sub> fluxes in the assimilation simulation (c) and GCB2019 (d). Correlation and root-mean-square deviation (RMSD) between assimilation and GCB2019 are shown in (e) and (f), respectively. The results are based on annual mean data for the time period 1970–2018. Positive values in (a) and (b) refer to CO<sub>2</sub> fluxes into the ocean or land.

In general, the historical GCB is well reproduced by the MPI-ESM when assimilating observational products, which enables a quantification of the GCB within a closed Earth system, showing that prediction systems yield internally consistent estimates of the air–sea and air–land CO<sub>2</sub> fluxes and are able to provide complementary information, in addition to the estimates provided by the GCP, for evaluating annual GCB.

#### 4 Predictability of the global carbon budget

The initialized predictions start from assimilation states which are close to observations. Therefore, information from the observations is incorporated into the prediction system through realistic initial states of the components of the climate system, which enables a more realistic evolution of the global carbon cycle and climate that follows the trajectory of observations until the predictability horizon is reached.

To support the GCP in predicting the next year’s GCB one year in advance, we also investigate the predictability, focusing on model hindcasts at a lead time of 1 year. As shown in Figs. 4 and 5, the initialized simulations at a lead time of 1 year show high correlations with GCB2019. The correlations of global atmospheric CO<sub>2</sub> growth, net air–sea CO<sub>2</sub> fluxes and net air–land CO<sub>2</sub> fluxes are 0.59, 0.52, 0.70 after removing the linear trends (Fig. 5a, c, e); the correlation of the original time series are 0.76, 0.97, and 0.66 (Fig. 4a, c, e). The initialized simulations at a lead time of 2 years still resemble the variations in the GCB2019, with correlations of 0.49 and higher (Fig. 6a, c, e), and the detrended time series also show higher correlations than the detrended uninitialized simulations (Fig. 6b, d, f). This suggests that internal variability can be constrained by initialization. As for atmospheric carbon growth, the initialized simulations at a lead time of 2 years show coherent interannual variations compared to GCB2019, albeit with a smaller correlation (0.49) than that of the historical freely evolving run (0.61), primar-



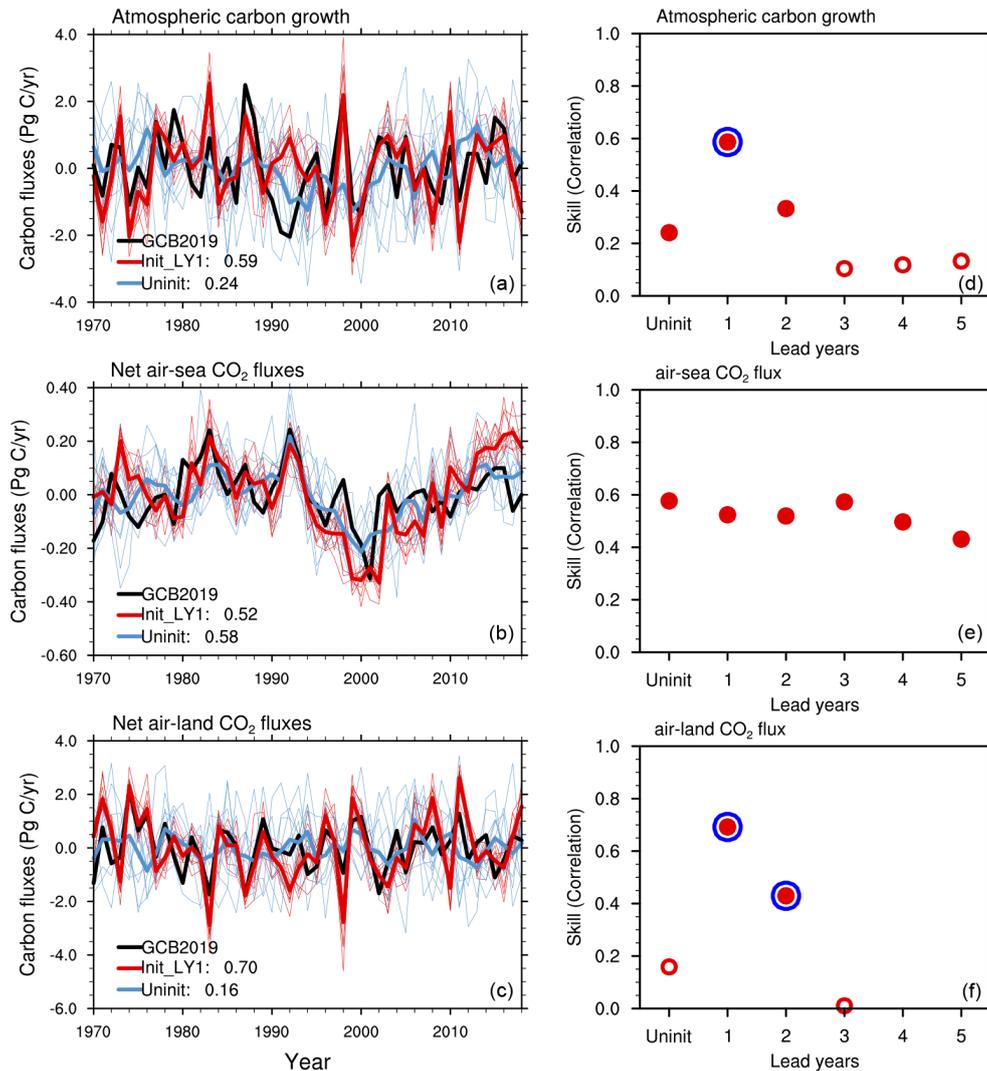
**Figure 4.** Time series of atmospheric carbon growth rate, i.e.,  $G_{ATM}$  (a), net air–sea  $CO_2$  fluxes, i.e.,  $S_{OCEAN}$  (b), and net air–land  $CO_2$  fluxes, i.e.,  $E_{LUC} + S_{LAND}$  (c), from the initialized simulations at a lead time of 1 year together with values from the uninitialized simulations and estimates from the 2019 Global Carbon Budget (GCB 2019, Friedlingstein et al., 2019). Positive values in (b) and (c) indicate  $CO_2$  fluxes into the ocean or land. The numbers in the legend show the correlation coefficients between the simulations and GCB2019, and the ensemble mean data are used for this correlation calculation. Predictive skill of the atmospheric carbon growth rate, i.e.,  $G_{ATM}$  (d), air–sea  $CO_2$  fluxes, i.e.,  $S_{OCEAN}$  (e), and net air–land  $CO_2$  fluxes, i.e.,  $E_{LUC} + S_{LAND}$  (f), in reference to the Global Carbon Budget (GCB 2019, Friedlingstein et al., 2019). The filled red circles on top of the open red circles show that the predictive skill is significant at a 95 % confidence level, and the additional larger blue circles indicate an improved significant predictive skill due to initialization in comparison to the uninitialized simulations. We use a nonparametric bootstrap approach (Goddard et al., 2013) to assess the significance of predictive skill. The results are based on annual mean data for the time period of 1970–2018.

ily due to the trends in atmospheric  $CO_2$  growth rate in the freely evolving run and in GCB2019. After detrending, the correlations are higher in the initialized simulations than in the uninitialized simulations (comparing Fig. 6a, d).

The initialized and uninitialized simulations show a comparably good match to GCB2019 with respect to the net  $CO_2$  flux into the ocean (with a high correlation up to 0.98) (Fig. 4b). The variations in the globally integrated ocean car-

bon sink are driven primarily by external forcing rather than internal variability, as found in McKinley et al. (2020). Figures 4b and 5b show that the ocean carbon sink variations (especially on decadal timescales) in the historical freely evolving uninitialized run are simulated reasonably well.

The net carbon flux into land shows higher correlations for initialized simulations at lead time of 1 year and 2 years than those for uninitialized simulations (Figs. 4c, 5c and 6c, f).

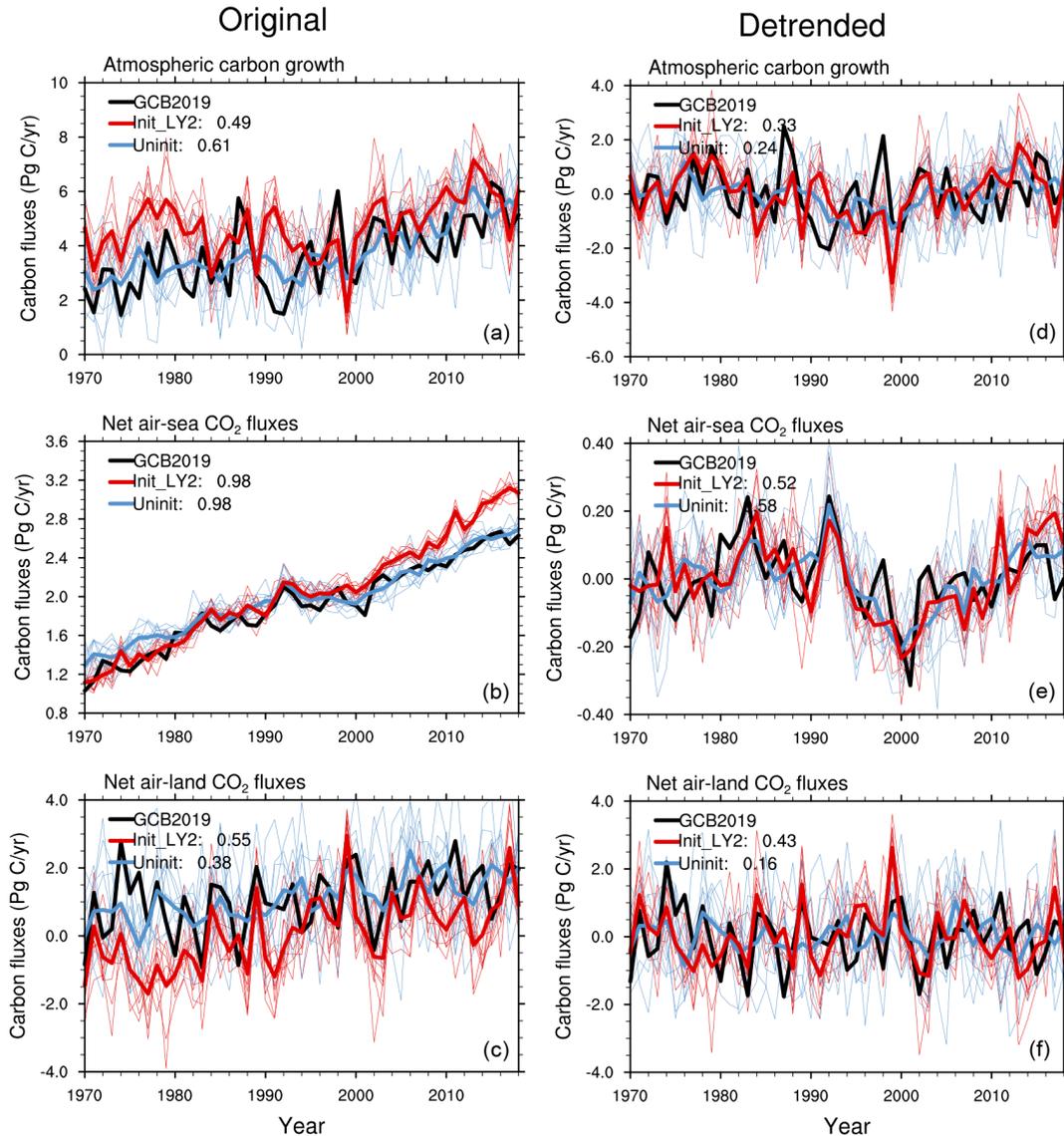


**Figure 5.** The same as Fig. 4 but with linearly detrended time series.

This indicates that the interannual variations are better captured in the initialized model system even after 2 years of free integration. This result implies a predictability of the air–land CO<sub>2</sub> flux for up to 2 years. The air–land CO<sub>2</sub> fluxes are regulated by El Niño–Southern Oscillation (ENSO) variations (Loughran et al., 2021; Dunkl et al., 2021), and the poor skill in predicting ENSO limits the predictability of the air–land CO<sub>2</sub> fluxes. However, the predictive skill of air–land CO<sub>2</sub> of 2 years is beyond the predictability horizon of ENSO, which is limited to a seasonal scale.

We further quantify the predictive skill of the GCB through all the lead times up to 5 years (Fig. 4b, d, f and Fig. 5b, d, f). The correlation skill relative to GCB2019 is significant for the lead time of 5 years in the atmospheric carbon growth and the ocean carbon sink. However, the skill for the air–land CO<sub>2</sub> flux is not statistically significant at the 95 % level after lead time of 2 years (Fig. 4d–f). The im-

proved predictive skill of initialized hindcasts compared to the historical uninitialized run occurs at a lead time of 1 year for atmospheric carbon growth and at a lead time of 2 years for air–land CO<sub>2</sub> flux. The detrended results (Fig. 5d–f) are similar to those from the original time series. The correlation of atmospheric carbon growth at a lead time of 2 years in the initialized hindcasts, compared to the estimates from the GCB2019, is higher than the uninitialized historical run when detrended. This indicates the contribution of a linear trend to the skill of atmospheric carbon growth in uninitialized historical runs as shown in Fig. 4d. Although the improvement in predictive skill in the initialized simulation relative to the uninitialized simulation is not significant for atmospheric CO<sub>2</sub> growth rate, the correlations of both initialized simulations at a lead time of 2 years and the uninitialized simulations are significantly high, as indicated with solid red



**Figure 6.** Time series at a lead time of 2 years for atmospheric carbon growth rate, i.e.,  $G_{ATM}$  (a), net air–sea  $CO_2$  fluxes, i.e.,  $S_{OCEAN}$  (b), and net air–land  $CO_2$  fluxes, i.e.,  $E_{LUC} + S_{LAND}$  (c), from the initialized simulations together with values from the uninitialized simulations and estimates from the 2019 Global Carbon Budget (GCB 2019, Friedlingstein et al., 2019). Panels (d), (e), and (f) are the same as (a), (b), and (c) but show the linearly detrended time series. The time series shown are based on annual mean data for the time period of 1970–2018. Positive values in panels (b), (c), (e), and (f) imply  $CO_2$  fluxes into the ocean or land. The numbers in the legend show the correlation coefficients between the simulations and GCB2019, and the ensemble mean data are used for the calculation.

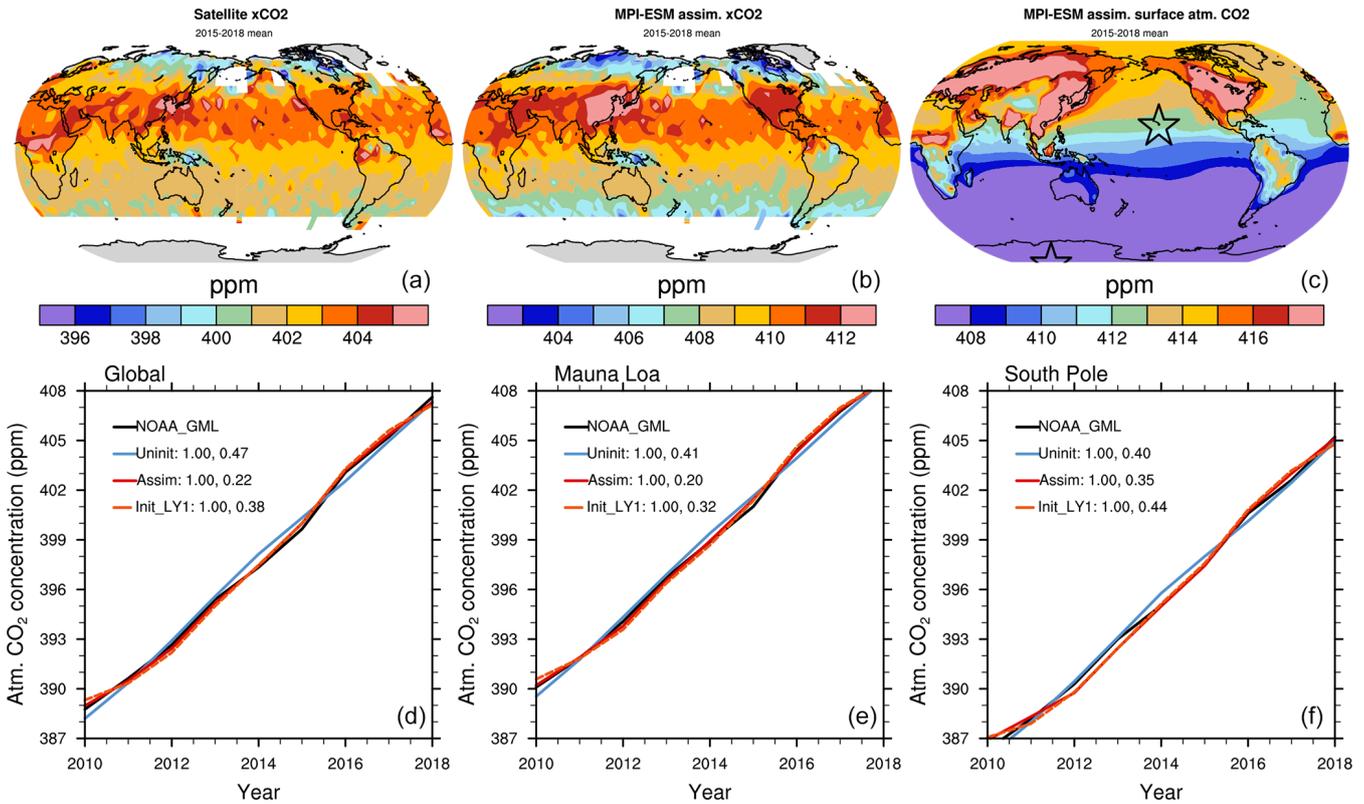
dots. This suggests the predictability of atmospheric carbon growth for up to 2 years.

From our MPI-ESM1.2-LR initialized hindcasts, we find that predictive skill of the air–sea  $CO_2$  flux is relatively high for up to 5 years and that of the air–land  $CO_2$  fluxes is up to 2 years. This is consistent with previous studies without an interactive carbon cycle (Ilyina et al., 2021; Lovenduski et al., 2019a, b). Here we have extended the prediction system for emission-driven simulations, enabling prognostic  $CO_2$  and preserving features of predictability. The prognostic  $CO_2$  from the novel emission-driven decadal prediction sys-

tem suggests a predictive skill of 2 years for the atmospheric  $CO_2$  growth rate.

## 5 Atmospheric $CO_2$ concentration

Figure 7 shows the spatial pattern and time series of atmospheric  $CO_2$  concentration from MPI-ESM simulations, including uninitialized, assimilation, and initialized simulations, together with the satellite  $XCO_2$  (i.e., atmospheric column-average dry-air mole fraction  $CO_2$ ) and NOAA\_GML observations for the period 2015–2018. The

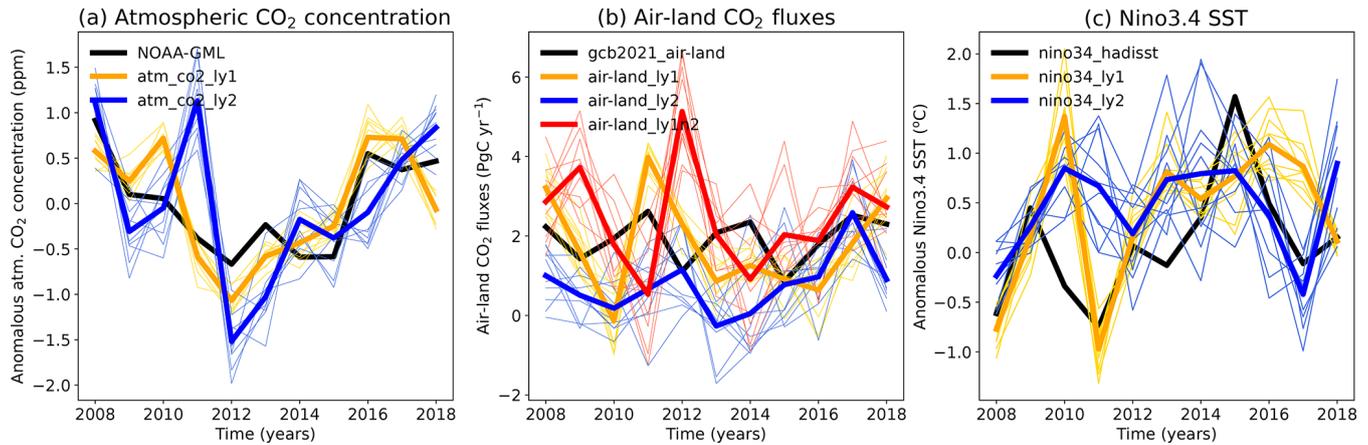


**Figure 7.** Spatial distribution of 2015–2018 mean satellite-based Obs4MIPs XCO<sub>2</sub> (a), model assimilation of XCO<sub>2</sub> (resampled according to satellite data availability) (b), and model assimilation of atmospheric CO<sub>2</sub> concentration at 1000 hPa level (c). A short time period of 2015–2018 is used because of the limited temporal coverage of satellite data. The satellite XCO<sub>2</sub> data product is obtained from the Climate Data Store Copernicus Climate Change Service (Reuter et al., 2013). The conversion of model-simulated CO<sub>2</sub> to XCO<sub>2</sub> is performed according to Gier et al. (2020) (their Appendix A). Atmospheric CO<sub>2</sub> concentration as a global average (d), at Mauna Loa (e), and at the South Pole (f) from the uninitialized (Uninit) and assimilation (Assim) simulations and initialized simulations at a lead time of 1 year (Init\_LY1) compared to observations over the 2010–2018 period. The locations of Mauna Loa and the South Pole are shown in (c). The numbers in the legend show the correlation (left) and root-mean-square error (RMSE, right) of the simulations relative to observational data from NOAA\_GML (Dlugokencky and Tans, 2020). The simulated time series from the MPI-ESM simulations, including uninitialized, assimilation, and initialized simulations, are bias corrected by removing the difference in mean states and the linear trend between observations and simulations according to Boer et al. (2016).

XCO<sub>2</sub> from the assimilation simulation (Fig. 7b) shows the spatial distribution of atmospheric CO<sub>2</sub> concentration, which compares well with the satellite XCO<sub>2</sub> (Fig. 7a). High CO<sub>2</sub> concentrations are found in the tropics to midlatitudes of the Northern Hemisphere. Relatively low CO<sub>2</sub> concentrations are found in the Southern Hemisphere and the polar regions. Note the model simulation is several parts per million higher than the satellite data, and this deviation can be attributed back to the uninitialized historical simulation (see Fig. A3). Additionally, the satellite data do not cover all the seasons in high latitudes, and therefore the sampled values from the assimilation simulation also better represent the summer season's XCO<sub>2</sub> in those regions. The surface level CO<sub>2</sub> shows a more dominant higher concentration in the Northern Hemisphere than in the Southern Hemisphere (Fig. 7c). Here we also compare the surface atmospheric CO<sub>2</sub>

concentration with the measurements at the Mauna Loa and South Pole stations (locations are shown as stars in Fig. 7c).

The atmospheric carbon burden and therefore CO<sub>2</sub> concentration is an accumulative quantity and mainly shows a linear increasing trend in recent decades in response to increasing anthropogenic emissions. Systematically lower or higher simulated carbon uptake by land and ocean compared to the real world therefore accumulates over the time period while the model is integrated. The simulated atmospheric CO<sub>2</sub> concentration can deviate relative to observations. In the MPI-ESM, simulated global mean atmospheric CO<sub>2</sub> concentration is around 8 ppm higher compared to the observations in the 2010s (see Fig. A4). The NOAA\_GML data represent the average of atmospheric CO<sub>2</sub> over marine surface sites (Dlugokencky and Tans, 2020), and these values are slightly lower than the values over land since the anthropogenic CO<sub>2</sub> emissions occur mainly on land. The time series shown in



**Figure 8.** (a) Time series of atmospheric CO<sub>2</sub> concentration anomalies from initialized simulations at a lead time of 1 year and 2 years compared to the NOAA\_GML observations (Dlugokencky and Tans, 2020) over the last 10 years. Anomalies are calculated by detrending the time series and with the climatological mean removed. (b) Time series of CO<sub>2</sub> fluxes from initialized simulations at a lead time of 1 year and 2 years together with estimates from the GCB2019. The red curves present the sum of predictions at a lead time of 2 years and the previous year of prediction at a lead time of 1 year (air-land\_ly1n2). (c) Time series of nino3.4 SST from model simulations and HadISST (Rayner et al., 2002). The time series are original model outputs and concatenated according to the lead time of years.

Fig. 7d–f are bias corrected by removing the difference of mean states and linear trends between observations and simulations according to Boer et al. (2016).

The atmospheric CO<sub>2</sub> concentration from assimilation shown follows the evolution of NOAA\_GML observations well, with a RMSE of 0.22 ppm, which is better than the uninitialized historical run with a RMSE of 0.47 ppm (Fig. 7d). In general, the RMSE increases from a lead time of 1 year to 2 years and decreases until a lead time of 5 years at both the global and observatory sites of Mauna Loa and the South Pole (Figs. A5 and A6). The relatively low predictive skill at a lead time of 2 years in atmospheric CO<sub>2</sub> concentration is because the model failed to predict the neutral ENSO events in 2010 and La Niña in 2011 and instead predicts a strong El Niño in both years (Fig. 8c). The corresponding air–land CO<sub>2</sub> fluxes are reversed, i.e., the land takes up less CO<sub>2</sub> than expected in 2011 (Fig. 8b solid blue curve and solid black curve). As the atmospheric CO<sub>2</sub> concentration is a cumulative quantity, the magnitude of atmospheric CO<sub>2</sub> concentration is affected by the CO<sub>2</sub> fluxes in the current and previous years. We also present the cumulative air–land CO<sub>2</sub> fluxes of the first- and second-year prediction (see the red curves in Fig. 8b), and the variations in cumulative air–land CO<sub>2</sub> fluxes are reversed compared to those in atmospheric CO<sub>2</sub> concentration changes at a lead time of 2 years, as shown in Fig. 8a (blue curves). The results indicate that the air–land CO<sub>2</sub> flux and corresponding atmospheric CO<sub>2</sub> has predictive skill, though the skill at a lead time of 2 years is degraded by the poor predictive skill of ENSO in some starting year predictions.

This retrospective prediction demonstrates the ability of an ESM-based decadal prediction system in reconstructing and predicting the global carbon cycle via only assimilating

the physical atmosphere and ocean fields. As presented in Fig. 5b, d, and f, the hindcasts also show a predictive skill of 5 years for air–sea CO<sub>2</sub> fluxes and 2 years for air–land CO<sub>2</sub> fluxes and atmospheric carbon growth. Hence, the ability of ESMs to predict the next year’s GCB is high.

## 6 Conclusions

For the first time, we have extended a decadal prediction system based on the MPI-ESM to include an interactive carbon cycle driven by fossil fuel emissions that enables prognostic atmospheric CO<sub>2</sub> predictions. The new assimilation and initialized predictions have one more degree of freedom, i.e., prognostic atmospheric CO<sub>2</sub>, and this framework represents the global carbon cycle as it operates in the real world.

The variations in atmospheric carbon growth rate and CO<sub>2</sub> fluxes among the atmosphere, ocean, and land are well reconstructed in our assimilation simulations, with high correlations (0.75, 0.97, and 0.75) compared to the estimates from the GCB2019. This provides confidence in the quantification of the GCB in a closed system within an Earth system model. Reconstructions of the GCB based on ESMs are therefore able to potentially provide additional lines of evidence for quantifying the annual GCB and open new opportunities in assessing the efficiency of carbon sinks. In particular, this approach eliminates the budget imbalance term that arises in GCBs of the GCP due to the combination of various and not fully consistent model and data approaches.

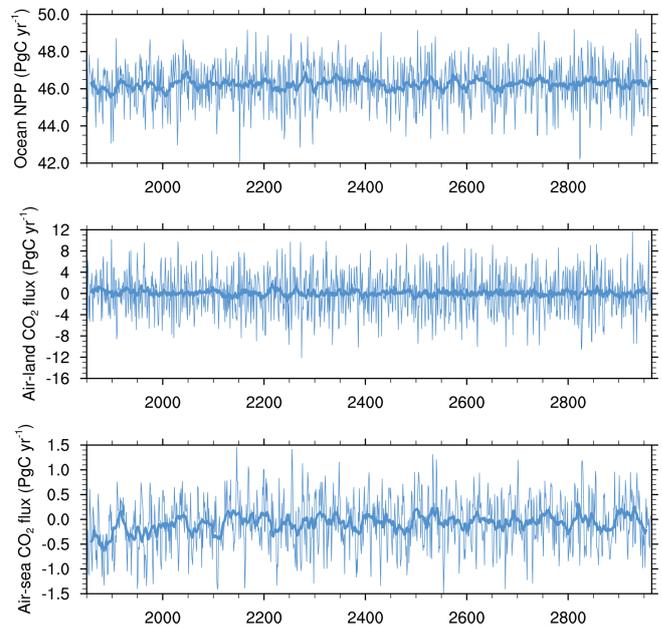
To further support the GCP in predicting next year’s GCB, the focus of the predictability investigations is on the lead time of 1 year. The results show high confidence in predicting the GCB for the next year with the MPI-ESM prediction sys-

tem. We further demonstrate that retrospective predictions of the GCB have a predictive skill for up to 5 years for air–sea  $\text{CO}_2$  fluxes and up to 2 years for air–land fluxes and atmospheric carbon growth rate. This indicates that the variations in atmospheric  $\text{CO}_2$  are better reproduced in the assimilation and retrospective predictions than in the uninitialized freely evolving historical simulations.

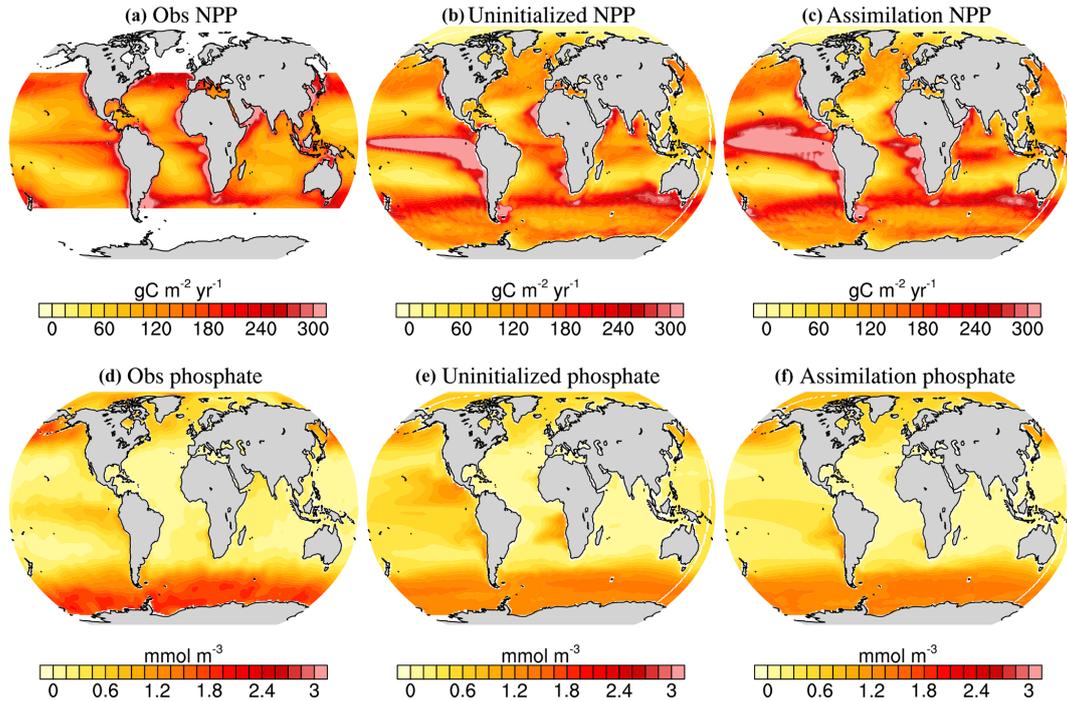
The MPI-ESM decadal prediction framework preserves the high predictive power in an emission-driven configuration, simulating the atmospheric  $\text{CO}_2$  growth rate with reasonable accuracy. In addition, the emission-driven decadal prediction system delivers the huge advantage of simulating the air–land and air–sea  $\text{CO}_2$  fluxes in response to fossil fuel and land use change emissions, including all feedbacks in a consistent framework. Further future efforts that assimilate more observations to initialize ESMs and assess their predictive skill will lead to more reliable reconstructions and predictions of global estimates and spatial distributions of  $\text{CO}_2$  fluxes and atmospheric  $\text{CO}_2$ . This study is based on simulations from a single ESM. Multi-model simulations that adopt a framework similar to that used in this study will allow for the identification of robust changes in the global carbon cycle expected to occur over the next few years.

We have demonstrated that the MPI-ESM-based emission-driven decadal prediction system exhibits the capability to reconstruct and predict the GCB and atmospheric  $\text{CO}_2$  concentration variations. Such ESM-based applications will be a useful tool in supporting global carbon stocktaking in compliance with the goals of the Paris Agreement.

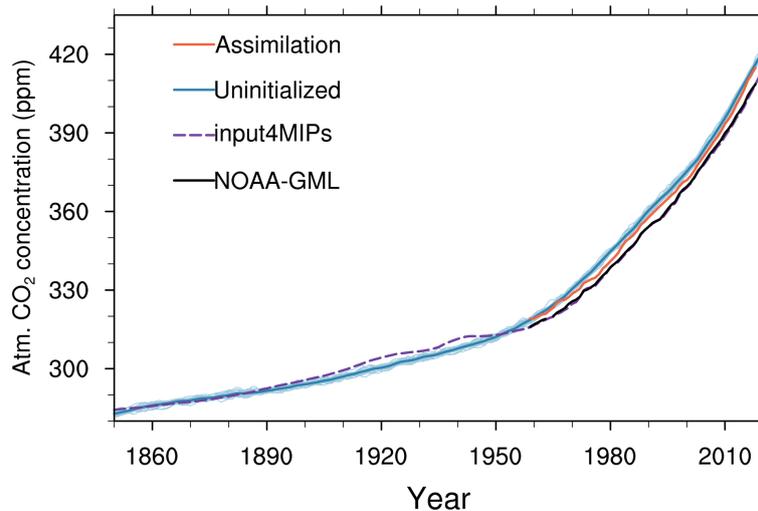
## Appendix A



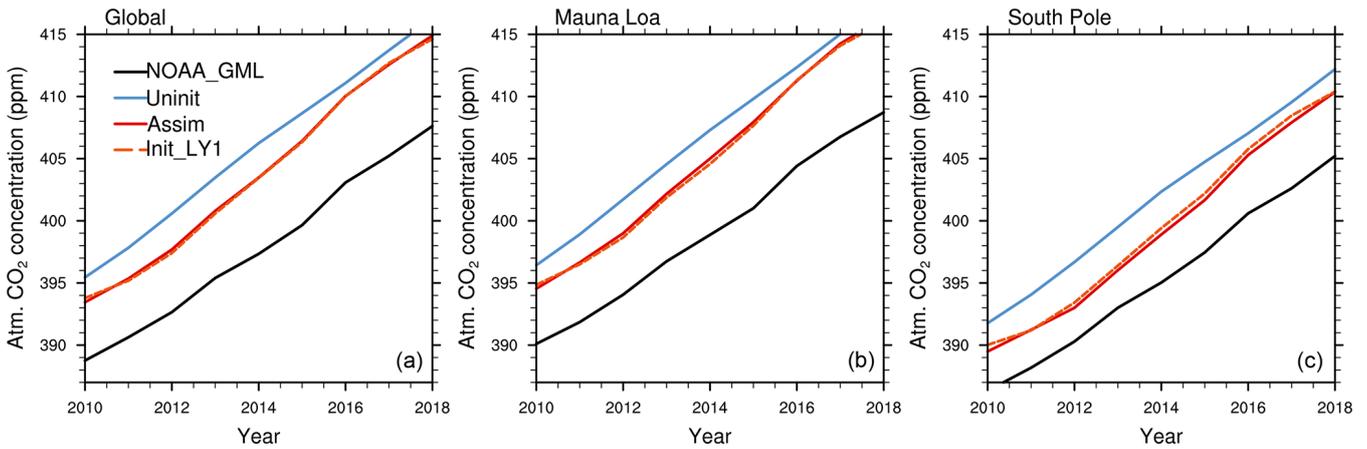
**Figure A1.** Time series of model simulations of ocean net primary production, air–sea  $\text{CO}_2$  flux, and air–land  $\text{CO}_2$  flux in the pre-industrial control run. The thin lines are annual mean time series, and the thick lines are their 20-year running means.



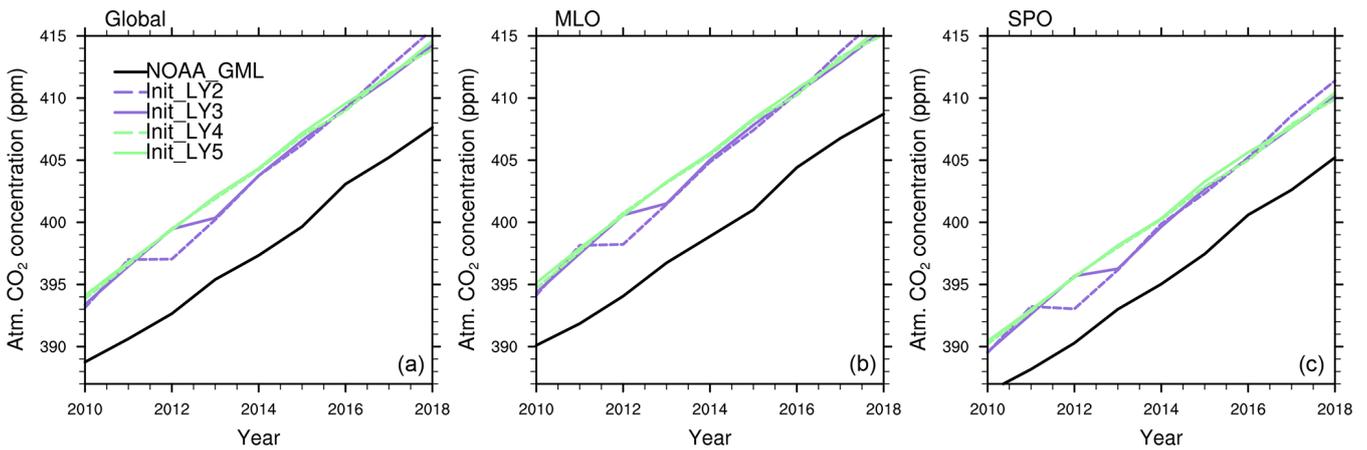
**Figure A2.** Climatological mean of ocean net primary production (NPP, **a–c**) and phosphate concentration (**d–f**) from observations and from model simulations. Observation-based NPP data are estimated from ocean color measurements obtained by the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) instrument of the OrbView-2 satellite for the September 1997–December 2002 period and the Moderate Resolution Imaging Spectroradiometer (MODIS) of the Aqua satellite for the 2003–2014 period (Behrenfeld and Falkowski, 1997, <http://science.oregonstate.edu/ocean.productivity/index.php>, last access: 12 November 2019). Phosphate observations are from the World Ocean Atlas 2018 (Garcia et al., 2019). The corresponding NPP data from model simulations are averaged over the 1998–2017 period, and phosphate data are averaged over the 1970–2018 period according to the availability of the observation data.



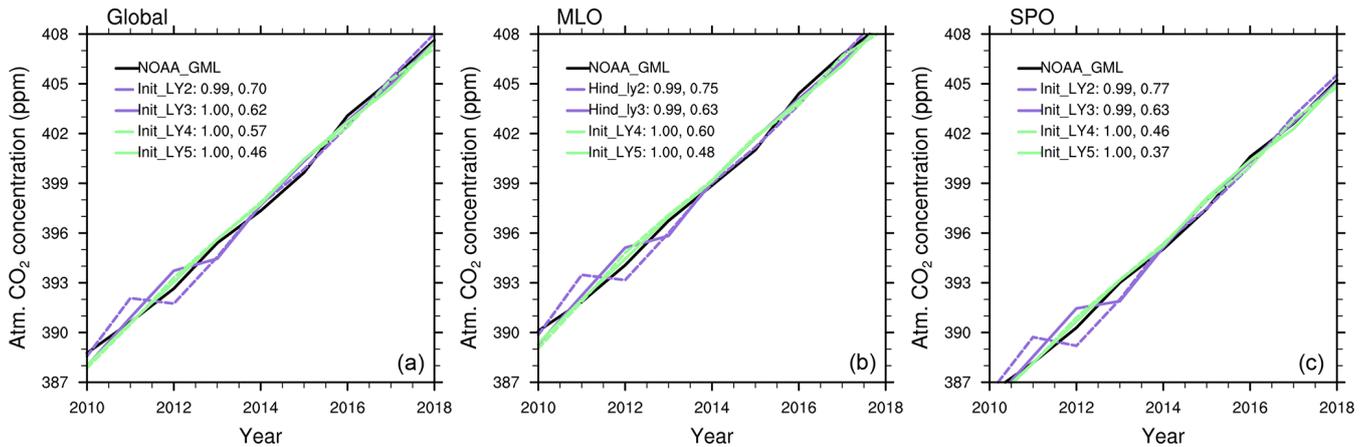
**Figure A3.** Time series of atmospheric  $\text{CO}_2$  concentration from model simulations and observations from 1850 to 2020. The assimilation and uninitialized simulations are shown with solid orange and blue lines, respectively. The CMIP6 input4MIPs atmospheric  $\text{CO}_2$  concentration forcing and the NOAA\_GML observation (Dlugokencky and Tans, 2020) are shown with dashed blue lines and solid black lines, respectively.



**Figure A4.** Atmospheric CO<sub>2</sub> concentration from the assimilation and initialized simulations at a lead time of 1 year together with NOAA\_GML observations (Dlugokencky and Tans, 2020) over the last 10 years. The time series are original model outputs and concatenated according to the lead time of years.



**Figure A5.** Atmospheric CO<sub>2</sub> concentration from initialized simulations at a lead time of 2–5 years together with NOAA\_GML observations (Dlugokencky and Tans, 2020) over the last 10 years. The time series are original model outputs and concatenated according to the lead time of years.



**Figure A6.** The same as Fig. A5 but with bias-corrected mean states and linear trends.

**Table A1.** Simulations based on MPI-ESM1.2-LR. The resolution for the atmosphere is T63L47 and for the ocean is GR15L40. The design of the prediction simulations is according to previous studies (Marotzke et al., 2016; Li et al., 2019). The assimilation starts from the end of the year 1958 in an uninitialized simulation. The nudging is strong; therefore, assimilation starting from a different uninitialized simulation would end up with similar evolution of the climate and carbon cycle. Figure 1 illustrates the simulations with evolution of atmospheric CO<sub>2</sub> growth rate together with observations. The initialized simulations start from the assimilation yearly from October 31st and run freely for 2 months plus 5 years afterwards. We have 59 runs for each ensemble of initialized simulations starting from 1960 to 2018 annually, and these run for 5 years and 2 months each, i.e., November 1960–December 1965 for starting year 1960, November 1961–December 1966 for starting year 1961, and so forth until November 2018–December 2023. The ensembles are generated with lagged 1 d initialization, i.e., the simulations start from 10 consecutive days from 31 October to 9 November. The ensembles for uninitialized simulations (shown as in Fig. A3) are generated by starting from different years of the control simulation (Fig. A1).

Simulations	Ensemble members	Nudging	Initial condition	Time period
Uninitialized	10	n/a	Pre-industrial	1850–2099
Assimilation	1	Atm.: ERA Ocean: ORAS4 anomalies (without 5° N–5° S band) Sea ice: NSIDC	Uninitialized	1959–2018
Initialized	10	n/a	Assimilation	1960–1965 ... 2018–2023

n/a: not applicable.

**Code and data availability.** The primary data and scripts used in the analysis that may be useful in reproducing the authors' work are archived by the Max Planck Institute for Meteorology and can be obtained via the institutional repository <http://hdl.handle.net/21.11116/0000-0009-6B84-A> (Li, 2023).

**Author contributions.** HL and TI conceived the idea. HL designed this study, ran the MPI-ESM simulations, performed the analyses, and drafted the manuscript. TL ran the CBALONE module simulations. TI, TL, AS, and JP contributed to discussing the results and editing the manuscript.

**Competing interests.** The contact author has declared that none of the authors has any competing interests.

**Disclaimer.** Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Acknowledgements.** This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project ID bm1124. We thank Julia Nabel and Jochem Marotzke for internally reviewing this manuscript. We thank Veronika Gayler and Thomas Raddatz for providing script and forcing data for setting up MPI-ESM1.2-LR emission-driven simulations. We thank Yohei Takano for running four members of the uninitialized simulations and Kameswar Modali for helping set up the assimilation and initialized simulations. We thank Pierre Friedlingstein, Judith Hauck, Stephen Sitch, and their co-authors for making the global carbon budget data from Global Carbon Project available. We thank Kelli Johnson for editing the En-

glish writing. We thank the reviewers Wei Li and Vivek Arora for their insightful and constructive comments on the paper. We acknowledge Vivek Arora for his editorial comments and efforts in improving the English writing.

**Financial support.** This research has been supported by the European Union’s Horizon 2020 research and innovation program (4C, grant no. 821003; CRESCENDO, grant no. 641816; COMFORT, grant no. 820989), the Bundesministerium für Bildung und Forschung (the research program MiKlip, grant no. 01LP1517B), and the Deutsche Forschungsgemeinschaft (Germany’s Excellence Strategy – EXC 2037 “CLICCS – Climate, Climatic Change, and Society” – project no. 390683824, contribution to the Center for Earth System Research and Sustainability (CEN) of Universität Hamburg).

The article processing charges for this open-access publication were covered by the Max Planck Society.

**Review statement.** This paper was edited by Vivek Arora and reviewed by Wei Li and Vivek Arora.

## References

- Ballantyne, A., Alden, C., Miller, J., Tans, P., and White, J.: Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years, *Nature*, 488, 70–72, 2012.
- Balmaseda, M. A., Mogensen, K., and Weaver, A. T.: Evaluation of the ECMWF ocean reanalysis system ORAS4, *Q. J. Roy. Meteorol. Soc.*, 139, 1132–1161, 2013.
- Bastos, A., O’Sullivan, M., Ciais, P., Makowski, D., Sitch, S., Friedlingstein, P., Chevallier, F., Rödenbeck, C., Pongratz, J., Lujikx, I. T., Patra, P. K., Peylin, P., Canadell, J. G., Lauerwald, R., Li, W., Smith, N. E., W. Peters, W., Goll, D. S., Jain, A. K., Kato, E., Lienert, S., Lombardozi, D. L., Haverd, V., Nabel, J. E. M. S., Poulter, B., Tian, H., Walker, A. P., and Zaehle, S.: Sources of uncertainty in regional and global terrestrial CO<sub>2</sub> exchange estimates, *Global Biogeochem. Cy.*, 34, e2019GB006393, <https://doi.org/10.1029/2019GB006393>, 2020.
- Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, 42, 1–20, 1997.
- Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, 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.: The Decadal Climate Prediction Project (DCPP) contribution to CMIP6, *Geosci. Model Dev.*, 9, 3751–3777, <https://doi.org/10.5194/gmd-9-3751-2016>, 2016.
- Bunzel, F., Notz, D., Baehr, J., Müller, W. A., and Fröhlich, K.: Seasonal climate forecasts significantly affected by observational uncertainty of Arctic sea ice concentration, *Geophys. Res. Lett.*, 43, 852–859, 2016.
- Canadell, J. G., Le Quééré, C., Raupach, M. R., Field, C. B., Buitenhuis, E. T., Ciais, P., Conway, T. J., Gillett, N. P., Houghton, R., and Marland, G.: Contributions to accelerating atmospheric CO<sub>2</sub> growth from economic activity, carbon intensity, and efficiency of natural sinks, *P. Natl. Acad. Sci. USA*, 104, 18866–18870, 2007.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van deBerg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. Roy. Meteorol. Soc.*, 137, 553–597, 2011.
- Dlugokencky, E. and Tans, P.: Trends in atmospheric carbon dioxide, National Oceanic and Atmospheric Administration, Global Monitoring Laboratory (NOAA\_GML), <https://gml.noaa.gov/ccgg/trends/> (last access: 15 November 2020), 2020.
- Dunkl, I., Spring, A., Friedlingstein, P., and Brovkin, V.: Process-based analysis of terrestrial carbon flux predictability, *Earth Syst. Dynam.*, 12, 1413–1426, <https://doi.org/10.5194/esd-12-1413-2021>, 2021.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- Fransner, F., Counillon, F., Bethke, I., Tjiputra, J., Samuelsen, A., Nummelin, A., and Olsen, A.: Ocean Biogeochemical Predictions-Initialization and Limits of Predictability, *Front. Mar. Sci.*, 7, 386, <https://doi.org/10.3389/fmars.2020.00386>, 2020.
- Friedlingstein, P., Jones, M. W., O’Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quééré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregón, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2019, *Earth Syst. Sci. Data*, 11, 1783–1838, <https://doi.org/10.5194/essd-11-1783-2019>, 2019.
- Friedlingstein, P., O’Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quééré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneeth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V.,

- Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X., and Zaehle, S.: Global Carbon Budget 2020, *Earth Syst. Sci. Data*, 12, 3269–3340, <https://doi.org/10.5194/essd-12-3269-2020>, 2020.
- Garcia, H. E., Weathers, K., Paver, C. R., Smolyar, I., Boyer, T. P., Locarnini, R. A., Zweng, M. M., Mishonov, A. V., Baranova, O. K., Seidov, D., and Reagan, J. R.: World ocean atlas 2018, Vol. 4, Dissolved inorganic nutrients (phosphate, nitrate and nitrate + nitrite, silicate), A. Mishonov Technical Ed., NOAA Atlas NESDIS 84, 35 pp., 2019.
- Gier, B. K., Buchwitz, M., Reuter, M., Cox, P. M., Friedlingstein, P., and Eyring, V.: Spatially resolved evaluation of Earth system models with satellite column-averaged CO<sub>2</sub>, *Biogeosciences*, 17, 6115–6144, <https://doi.org/10.5194/bg-17-6115-2020>, 2020.
- Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., Kharin, V., Merryfield, W., Deser, C., Mason, S. J., Kirtman, B. P., Msadek, R., Sutton, R., Hawkins, E., Fricker, T., Hegerl, G., Ferro, C. A. T., Stephenson, D. B., Meehl, G. A., Stockdale, T., Burgman, R., Greene, A. M., Kushnir, Y., Newman, M., Carton, J., Fukumori, I., and Delworth, T.: A verification framework for interannual-to-decadal predictions experiments, *Clim. Dynam.*, 40, 245–272, 2013.
- Hansis, E., Davis, S. J., and Pongratz, J.: Relevance of methodological choices for accounting of land use change carbon fluxes, *Global Biogeochem. Cy.*, 29, 1230–1246, 2015.
- Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinemann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and Zhang, X.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, *Geosci. Model Dev.*, 13, 5425–5464, <https://doi.org/10.5194/gmd-13-5425-2020>, 2020.
- Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., and Núñez-Riboni, I.: Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations, *J. Adv. Model. Earth Sy.*, 5, 287–315, 2013.
- Ilyina, T., Li, H., Spring, A., Müller, W. A., Bopp, L., Chikamoto, M. O., Danabasoglu, G., Dobrynin, M., Dunne, J., Fransers, F., Friedlingstein, P., Lee, W., Lovenduski, N. S., Merryfield, W. J., Mignot, J., Park, J. Y., Séférian, R., Sospedra-Alfonso, R., Watanabe, M., and Yeager, S.: Predictable variations of the carbon sinks and atmospheric CO<sub>2</sub> growth in a multi-model framework, *Geophys. Res. Lett.*, 48, e2020GL090695, <https://doi.org/10.1029/2020GL090695>, 2021.
- Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J. G., Jung, M., Kawamiya, M., Koven, C., Pongratz, J., Raddatz, T., Randerson, J. T., and Zaehle, S.: C4MIP – The Coupled Climate–Carbon Cycle Model Intercomparison Project: experimental protocol for CMIP6, *Geosci. Model Dev.*, 9, 2853–2880, <https://doi.org/10.5194/gmd-9-2853-2016>, 2016.
- Landschützer, P., Ilyina, T., and Lovenduski, N. S.: Detecting Regional Modes of Variability in Observation-Based Surface Ocean pCO<sub>2</sub>, *Geophys. Res. Lett.*, 46, 2670–2679, 2019.
- Li, H.: Reproducibility repository for “Li, H., Ilyina, T., Loughran, T., Spring, A., and Pongratz, J.: Reconstructions and predictions of the global carbon budget with an emission-driven Earth system model, *Earth System Dynamics*” [data set], <https://hdl.handle.net/21.11116/0000-0009-6B84-A>, last access: 16 January 2023.
- Li, H. and Ilyina, T.: Current and future decadal trends in the oceanic carbon uptake are dominated by internal variability, *Geophys. Res. Lett.*, 45, 916–925, 2018.
- Li, H., Ilyina, T., Müller, W. A., and Sienz, F.: Decadal predictions of the North Atlantic CO<sub>2</sub> uptake, *Nat. Commun.*, 7, 1–7, 2016.
- Li, H., Ilyina, T., Müller, W. A., and Landschützer, P.: Predicting the variable ocean carbon sink, *Sci. Adv.*, 5, eaav6471, <https://doi.org/10.1126/sciadv.aav6471>, 2019.
- Loughran, T. F., Boysen, L., Bastos, A., Hartung, K., Havermann, F., Li, H., Nabel, J. E. M. S., Obermeier, W. A., and Pongratz, J.: Past and future climate variability uncertainties in the global carbon budget using the MPI Grand Ensemble, *Global Biogeochem. Cy.*, 35, e2021GB007019, <https://doi.org/10.1029/2021GB007019>, 2021.
- Lovenduski, N. S., Bonan, G. B., Yeager, S. G., Lindsay, K., and Lombardozzi, D. L.: High predictability of terrestrial carbon fluxes from an initialized decadal prediction system, *Environ. Res. Lett.*, 14, 124074, <https://doi.org/10.1088/1748-9326/ab5c55>, 2019a.
- 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>, 2019b.
- Marotzke, J., Müller, W. A., Vamborg, F. S., Becker, P., Cubasch, U., Feldmann, H., Kaspar, F., Kottmeier, C., Marini, C., Polkova, I., Prömmel, K., Rust, H. W., Stammer, D., Ulbrich, U., Kadow, C., Köhl, A., Kröger, J., Kruschke, T., Pinto, J. G., Pohlmann, H., Reyers, M., Schröder, M., Sienz, F., Timmreck, C., and Ziese, M.: MiKlip: A national research project on decadal climate prediction, *Bull. Am. Meteorol. Soc.*, 97, 2379–2394, 2016.
- Marsland, S. J., Haak, H., Jungclaus, J. H., Latif, M., and Röske, F.: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates, *Ocean Model.*, 5, 91–127, 2003.
- Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenez-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornbluh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six,

- K. D., Stein, L., Stemmler, I., Stevens, B., Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1. 2) and its response to increasing CO<sub>2</sub>, *J. Adv. Model. Earth Sy.*, 11, 998–1038, 2019.
- McKinley, G. A., Fay, A. R., Eddebbbar, Y. A., Gloege, L., and Lovenduski, N. S.: External forcing explains recent decadal variability of the ocean carbon sink, *AGU Adv.*, 1, e2019AV000149, <https://doi.org/10.1029/2019AV000149>, 2020.
- Meehl, G. A., Richter, J. H., Teng, H., Capotondi, A., Cobb, K., Doblus-Reyes, F., Donat, M. G., England, M. H., Fyfe, J. C., Han, W., Kim, H., Kirtman, B. P., Kushnir, Y., Lovenduski, N. S., Mann, M. E., Merryfield, W. J., Nieves, V., Pegion, K., Rosenbloom, N., Sanchez, S. C., Scaife, A. A., Smith, D., Subramanian, A. C., Sun, L., Thompson, D., Ummenhofer, C. C., and Xie, S. P.: Initialized Earth System prediction from subseasonal to decadal timescales, *Nat. Rev. Earth Environ.*, 2, 340–357, 2021.
- Obermeier, W. A., Nabel, J. E. M. S., Loughran, T., Hartung, K., Bastos, A., Havermann, F., Anthoni, P., Arneth, A., Goll, D. S., Lienert, S., Lombardozzi, D., Luysaert, S., McGuire, P. C., Melton, J. R., Poulter, B., Sitch, S., Sullivan, M. O., Tian, H., Walker, A. P., Wiltshire, A. J., Zaehle, S., and Pongratz, J.: Modelled land use and land cover change emissions – a spatio-temporal comparison of different approaches, *Earth Syst. Dynam.*, 12, 635–670, <https://doi.org/10.5194/esd-12-635-2021>, 2021.
- Park, J.-Y., Stock, C. A., Yang, X., Dunne, J. P., Rosati, A., John, J., and Zhang, S.: Modeling global ocean biogeochemistry with physical data assimilation: a pragmatic solution to the equatorial instability, *J. Adv. Model. Earth Sy.*, 10, 891–906, 2018.
- Paulsen, H., Ilyina, T., Six, K. D., and Stemmler, I.: Incorporating a prognostic representation of marine nitrogen fixers into the global ocean biogeochemical model HAMOCC, *J. Adv. Model. Earth Sy.*, 9, 438–464, 2017.
- Peters, G. P., Le Quéré, C., Andrew, R. M., Canadell, J. G., Friedlingstein, P., Ilyina, T., Jackson, R. B., Joos, F., Korsbakken, J. I., McKinley, G. A., Sitch, S., and Tans, P.: Towards real-time verification of CO<sub>2</sub> emissions, *Natu. Clim. Change*, 7, 848–850, 2017.
- Pongratz, J., Reick, C. H., Houghton, R. A., and House, J. I.: Terminology as a key uncertainty in net land use and land cover change carbon flux estimates, *Earth Syst. Dynam.*, 5, 177–195, <https://doi.org/10.5194/esd-5-177-2014>, 2014.
- Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E., Raddatz, T., Roeckner, E., Schnur, R., and Wilkenskjeld, S.: JSBACH 3-The land component of the MPI Earth System Model: documentation of version 3.2, *Berichte zur Erdsystemforschung*, 240, <https://doi.org/10.17617/2.3279802>, 2021.
- Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, 108, 4407, <https://doi.org/10.1029/2002JD002670>, 2002.
- Reuter, M., Bösch, H., Bovensmann, H., Bril, A., Buchwitz, M., Butz, A., Burrows, J. P., O'Dell, C. W., Guerlet, S., Hasekamp, O., Heymann, J., Kikuchi, N., Oshchepkov, S., Parker, R., Pfeifer, S., Schneising, O., Yokota, T., and Yoshida, Y.: A joint effort to deliver satellite retrieved atmospheric CO<sub>2</sub> concentrations for surface flux inversions: the ensemble median algorithm EMMA, *Atmos. Chem. Phys.*, 13, 1771–1780, <https://doi.org/10.5194/acp-13-1771-2013>, 2013.
- Séférian, R., Berthet, S., and Chevallier, M.: Assessing the decadal predictability of land and ocean carbon uptake, *Geophys. Res. Lett.*, 45, 2455–2466, 2018.
- Spring, A. and Ilyina, T.: Predictability Horizons in the Global Carbon Cycle Inferred From a Perfect-Model Framework, *Geophys. Res. Lett.*, 47, e2019GL085311, <https://doi.org/10.1029/2019GL085311>, 2020.
- Spring, A., Ilyina, T., and Marotzke, J.: Inherent uncertainty disguises attribution of reduced atmospheric CO<sub>2</sub> growth to CO<sub>2</sub> emission reductions for up to a decade, *Environ. Res. Lett.*, 15, 114058, <https://doi.org/10.1088/1748-9326/abc443>, 2020.
- Spring, A., Dunkl, I., Li, H., Brovkin, V., and Ilyina, T.: Trivial improvements in predictive skill due to direct reconstruction of the global carbon cycle, *Earth Syst. Dynam.*, 12, 1139–1167, <https://doi.org/10.5194/esd-12-1139-2021>, 2021.
- Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, I., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, Appl. Meteorol. Phys. Oceanogr.*, 131, 2961–3012, 2005.