Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL085311

Predictability Horizons in the Global Carbon Cycle Inferred From a Perfect-Model Framework

Aaron Spring[1,2] **and Tatiana Ilyina**[1]

[1]Max Planck Institute for Meteorology, Hamburg, Germany, [2]Max Planck Institute for Meteorology, International Max Planck Research School of Earth System Modelling, IMPRS, Hamburg, Germany

Key Points:
- Global annual atmospheric CO_2 variations are potentially predictable up to 3 years in advance
- The global oceanic CO_2 flux is predictable for 2 years with regional predictability up to a decade
- The global land CO_2 flux predictability of 2 years is dominated by the tropical forests

Supporting Information:
- Supporting Information S1

Correspondence to:
A. Spring,
aaron.spring@mpimet.mpg.de

Citation:
Spring, A., Ilyina, T. (2020). Predictability horizons in the global carbon cycle inferred from a perfect-model framework. *Geophysical Research Letters*, *47*, e2019GL085311. https://doi.org/10.1029/2019GL085311

Received 7 SEP 2019
Accepted 9 APR 2020
Accepted article online 23 APR 2020

Abstract On interannual timescales the growth rate of atmospheric CO_2 is largely controlled by the response of the land and ocean carbon sinks to climate variability. Yet, it is unknown to what extent this variability limits the predictability of atmospheric CO_2 variations. Using perfect-model Earth System Model simulations, we show that variations in atmospheric CO_2 are potentially predictable for 3 years. We find a 2-year predictability horizon for global oceanic CO_2 flux with longer regional predictability of up to 7 years. The 2-year predictability horizon of terrestrial CO_2 flux originates in the tropics and midlatitudes. With the predictability of the isolated effects of land and ocean carbon sink on atmospheric CO_2 of 5 and 12 years respectively, land dampens the overall predictability of atmospheric CO_2 variations. Our research shows the potential of Earth System Model-based predictions to forecast multiyear variations in atmospheric CO_2.

Plain Language Summary The amount of anthropogenic carbon emissions absorbed by land and ocean from the atmosphere varies annually due to their sensitivity to climate. Therefore, the atmospheric CO_2 growth rate does not strictly follow the emissions signal. Whether decadal prediction systems can also predict variations of atmospheric CO_2 has not been shown yet but is crucial to inform policy makers about the efficiency of the implementation of the Paris Agreement. Using numerical Earth System simulations in an idealized prediction framework, we show that global atmospheric CO_2 is predictable up to 3 years in advance. The global ocean and land CO_2 fluxes are predictable for 2 years. The isolated effects of the land and ocean carbon sink on atmospheric CO_2 are predictable for 5 and 12 years, respectively. Therefore, the land carbon cycle limits atmospheric CO_2 predictability. Our study demonstrates that simulation-based multiyear forecasts have the potential to predict natural atmospheric CO_2 variations.

1. Introduction

The atmospheric CO_2 mixing ratio rises in response to increasing anthropogenic carbon emissions. The terrestrial and oceanic carbon sinks modulate the atmospheric CO_2 mixing ratio by currently absorbing about 30% and 25% of the anthropogenic carbon emissions, respectively (Friedlingstein et al., 2019). However, the interannual atmospheric CO_2 growth rate does not strictly follow the changes in anthropogenic CO_2 emissions (Bacastow, 1976; Keeling et al., 1995; Peters et al., 2017). This discrepancy is due to internal variability of the oceanic and terrestrial carbon sinks driven by the large-scale modes of variability of the climate system (Doney et al., 2006; Heinze et al., 2015; Li & Ilyina, 2018; McKinley et al., 2017) and volcanic eruptions (Eddebbar et al., 2019; Frölicher et al., 2013; Jones et al., 2001). Indeed, this internal climate variability of the natural carbon sinks might obscure the identification of the anthropogenic emissions signal in atmospheric CO_2 mixing ratio (Peters et al., 2017). This sensitivity of interannual changes in atmospheric CO_2 concentration to natural climate variability is a major uncertainty in our understanding of the near-term evolution of atmospheric CO_2 and the remaining carbon budget. How this sensitivity limits the near-term predictability of the atmospheric CO_2 growth rate has not been addressed before.

The Paris Agreement (UNFCCC, 2015) aims to limit global warming to at most 2 °C by mostly mitigation of fossil fuel emissions. Article 14 demands for an assessment of collective progress periodically every five years "in the light (…) of the best available science." The potential of prediction systems with prognostic CO_2 has to be explored in the context of their ability to predict near-term atm. CO_2 variations in response to the changes of the natural carbon sinks. Given that CO_2 predictions are feasible, changes in emissions can

be earlier separated from internal variability. Therefore, the near-term efficacy of the planned mitigation efforts can be skillfully forecasted in atm. CO_2 in advance until the prediction horizon, rather than being assessed only in the following Global Stocktake. As a first step toward a viable atm. CO_2 prediction system, we here show its potential to be successful in an idealized setup.

Over the last decade, multiyear climate predictions have been established. Decadal prediction systems based on comprehensive Earth System Models (ESM) take advantage of the memory of the climate system. Observational products are assimilated into an ESM and used as initial conditions of the forecast (Meehl et al., 2009). While decadal prediction systems are not perfect, they proved themselves successful for many aspects of the climate system (Marotzke et al., 2016; Meehl et al., 2013; Yeager et al., 2018). Having established prediction skill for many climate variables, which the carbon cycles are sensitive to, prediction systems can be used to assess decadal predictability of the global carbon cycle.

In the emerging predictability studies of terrestrial and ocean carbon cycle components, only physical climate variables such as ocean temperature and salinity are nudged toward observations due to the sparse and short observational period of biogeochemical variables. Li et al. (2016) find predictability skill of 4 to 7 years in the winter North Atlantic air-sea CO_2 uptake when evaluating against an ocean reconstruction with uninitialized biogeochemical components. The global oceanic carbon uptake can be skilfully predicted for up to 2 years evaluated against observations (Li et al., 2019) and ocean reconstruction (Lovenduski, Yeager, et al., 2019). Séférian et al. (2018) estimates the potential predictability horizon of the global carbon sink predictability of 4 to 6 years in idealized perfect-model simulations. Unlike the previously mentioned ESM-based predictions, Betts et al. (2016, 2018) use a statistical model to forecast the growth rate of the atmospheric CO_2 mixing ratio at Mauna Loa. A regression of previous imprints of the El Ninõ–Southern Oscillation (ENSO) on atmospheric CO_2 mixing ratio and anthropogenic emissions predicts the next year atmospheric CO_2. Although ENSO is one of the main drivers of CO_2 flux variability (Jones & Cox, 2005; Zeng et al., 2005), this method lacks the representation of other dynamic processes and a multiyear outlook. Zeng et al. (2008) demonstrate technical feasibility of ESM-based seasonal air-land CO_2 flux predictions. They find a higher prediction skill for Mauna Loa atmospheric CO_2 than a statistical ENSO regression and thereby indicate the added value of ESM-based carbon cycle predictions. A very recent study finds 2-year predictability in terrestrial net ecosystem production predictability, which is closely linked to terrestrial CO_2 flux predictability (Lovenduski, Bonan, et al., 2019).

Previous studies on carbon cycle predictability lack a fully coupled global carbon cycle. Terrestrial carbon cycle based studies (Betts et al., 2018; Zeng et al., 2008) miss the oceanic carbon sink and multiyear lead times. (Oceanic) carbon cycle based predictability studies (Li et al., 2016, 2019; Lovenduski, Yeager, et al., 2019; Lovenduski, Bonan, et al., 2019; Séférian et al., 2018) rest upon *concentration-driven* simulations, in which atmospheric CO_2 mixing ratio is prescribed. In such simulations, air-sea and air-land CO_2 flux do not alter atmospheric CO_2 mixing ratio. We go beyond previous studies by using fully coupled carbon cycle ESM simulations. In our simulations, the atmospheric mixing ratio of CO_2 is a prognostic property and can be physically consistently predicted. We examine the limits of predictability of air-sea and air-land surface CO_2 flux as well as the resulting atmospheric CO_2 mixing ratio. Furthermore, we determine the isolated contributions of the terrestrial and oceanic carbon sinks to the limits of predictability of atmospheric CO_2. We examine the relevance of different regions of the overall land and ocean predictability based on two metrics.

We use a perfect-model framework, in which we assumes that the model can reproduce observed variability (Griffies & Bryan, 1997). As it is free of initialization from reanalysis products, the perfect-model framework cannot perform predictions, but only estimate time scales after which the state of initialization is lost in the ensemble due the chaotic nature of the Earth system. The assimilation of reanalysis products in prediction systems inevitably induces drifts from the reanalysis climatology toward the native model climatology due to biases between reanalysis and ESMs (Kröger et al., 2017). As marine biogeochemical cycles are highly sensitive to changes in ocean circulation, such ocean reconstructions can disrupt biogeochemical tracer distributions (Park et al., 2018; Toggweiler et al., 1989). This perfect-model framework offers the only self-consistent setup to examine processes leading to predictability without methodological artifacts due to initialization from reanalysis. We therefore use it as a first-order approach study to examine the limits of predictability of a fully coupled global carbon cycle.

2. Methods

2.1. Model Description

Our perfect-model framework study is based on perturbed initial conditions ensemble simulations from a pre-industrial control simulation with prognostic atmospheric CO_2 mixing ratio, which is known as *esm-piControl* in the Coupled Model Intercomparison Project Phase 6 framework (Eyring et al., 2016). We use the close to equilibrium spin-up simulation of the Coupled Model Intercomparison Project Phase 6 version of the Max Planck Institute ESM (MPI-ESM) (Mauritsen et al., 2019) taken for the Coupled Climate-Carbon Cycle Model Intercomparison Project (C^4MIP) (Jones et al., 2016) in its low-resolution configuration (horizontal resolution of about 1.8° in the atmosphere and on land, and about 1.5° in the ocean).

The ocean general circulation model Max Planck Institute Ocean Model (Jungclaus et al., 2013) provides the circulation field to the marine biogeochemical cycle model HAMOCC (Ilyina et al., 2013). HAMOCC encompasses carbonate chemistry and an NPZD-type ecosystem including additional nitrogen-fixating cyanobacteria (Paulsen et al., 2017), nutrient-light-temperature colimitation, calcite and opal export treated explicitly and iron as a macronutrient. The terrestrial carbon cycle model JSBACH incorporates wildfires, dynamic vegetation, soil carbon decomposition, and storage (Schneck et al., 2013). The atmospheric general circulation model ECHAM6 relies on a flux-form semi-Lagrangian scheme (Lin & Rood, 1996) to represent the transport of the three-dimensional atmospheric prognostic atmospheric CO_2 tracer (Stevens et al., 2013).

2.2. Perfect-Model Framework

The simulations in the perfect-model framework are started from 12 randomly chosen initialization states. From each of those states of a stable 300-year control simulation, we branch off nine ensemble members and integrate them for 20 years. Ensemble members are generated by perturbing the stratospheric horizontal diffusion by 1.0000{member} in the first year after initialization. This initialization approach induces only tiny perturbations to the climate system as the ocean and land initial conditions remain identical. Therefore, this perfect-model initialization presents an upper limit estimate of predictability. As a result of such initialization approach, the ensemble members first follow a similar evolution and then diverge until reaching the time scale of the predictability horizon, when temporal variability is statically indistinguishable from variability across members (Figure S1a in the supporting information).

2.3. Predictability Skill Assessment

Because only slightly perturbed initial conditions distinguish ensemble members, all ensemble members can equally serve as a verification to compute predictability according to a certain metric. Therefore, we compute predictability skill as the anomaly correlation coefficient skill score (ACC) and root-mean-square-error (RMSE) between the ensemble member mean, with verification member excluded, as forecast and every single ensemble member as verification (Jolliffe & Stephenson, 2011; Wilks, 2006). ACC assesses the linear association of the forecast and the verification and therefore measures synchronous evolution. RMSE measures the second-order distance between forecast and verification. RMSE describes how the individual ensemble members spread over lead time and is therefore a more conservative attribute of predictability skill. The two metrics measure related but different attributes of predictability skill (Murphy, 1988). Find additional details about the metrics in the supporting information.

We assess the significance of initialized forecast predictability by bootstrapping with replacement (Efron & Tibshirani, 1993; Goddard et al., 2013). To compare against random forecasts, we create hypothetical uninitialized ensembles randomly drawn as 20-year chunks from the control simulation. The p value denotes the probability that random forecasts are more skillful than the initialized forecast resampled with replacement over all initializations based on 5,000 iterations (except noted otherwise). This nonparametric approach relies on no further assumptions and is especially free from any normalization choice, which can bias the predictability horizon (Hawkins et al., 2016). We define the limit of predictability, the so-called *predictability horizon* of a metric as the last significant lead year (with p value<0.05).

2.4. Diagnosing Global Atmospheric CO_2 Mixing Ratio

We furthermore disentangle the isolated effects of the global terrestrial and oceanic carbon sinks on the global atmospheric CO_2 mixing ratio. For this we convert zero-mean free-of-trend (denoted by ′) temporally accumulated global surface CO_2 flux into diagnosed atmospheric CO_2 mixing ratio by multiplying with

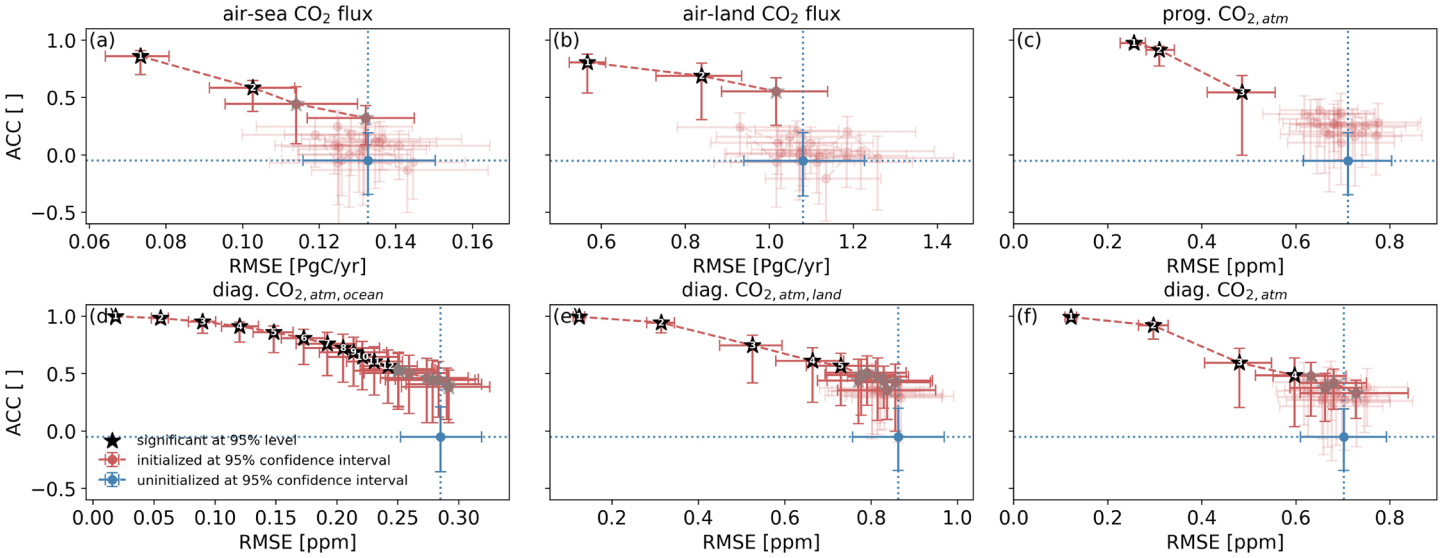

Figure 1. Comparison of the mean potential prediction skill of the initialized ensemble (red) versus random uninitialized ensembles (blue) in global annual surface quantities of the carbon cycle with the anomaly correlation coefficient (ACC) on the *y* axis and root-mean-square-error (RMSE) on the *x* axis for lead years represented as dots: (a) air-sea CO_2 flux, (b) air-land CO_2 flux, (c) prognostic surface atmospheric CO_2, (d) diagnosed atmospheric CO_2 based on oceanic carbon sink, (e) diagnosed atmospheric CO_2 based on the global terrestrial carbon sink, and (f) diagnosed atmospheric CO_2 based on the global oceanic and terrestrial carbon sink (see section 2.4). Error bars show 95% confidence intervals based on bootstrapping with replacement ($N = 5{,}000$). The last lead year with a bootstrapped *p* value (which represents that uninitialized ensembles beat initialized ensembles) lower than 5% marks the predictability horizon. Black stars with white integer denote significant lead years in ACC and RMSE, gray stars if only one metric is significant, and lead years nonsignificant in both metrics are blurred.

$\frac{ppm}{2.12PgC}$ (Ballantyne et al., 2012; Friedlingstein et al., 2019):

$$XCO_{2,\mathrm{atm,diag}}(t) = XCO_{2,atm}(0) + \sum_{t'}^{t} CO_2\mathrm{flux}'(t') \cdot \frac{ppm}{2.124PgC}$$

where$XCO_{2,atm}(0)$ is a free choice parameter of initial diagnosed atmospheric surface CO_2 mixing ratio. This approach assumes the atmosphere as one instantaneously mixed box and implicitly incorporates direct CO_2 flux from land to ocean. Converting the sum of terrestrial and oceanic CO_2 flux into diagnosed atmospheric CO_2 yields a very similar evolution as prognostic atmospheric CO_2.

2.5. Variance-Weighted Mean Period

We infer potentially predictable regions from a single control simulation following Branstator and Teng (2010). Accordingly, the variance-weighted mean period highlights regions with low-frequency variations with longer potentially predictable periods:

$$P_x = \sum_k V(f_k, x) / \sum_k f_k V(f_k, x),$$

where$V(f, x)$ is the variance per unit frequency for frequency f at location x and k are the individual frequencies from the power spectrum.

3. Predictability of Air-Sea CO_2 Flux

We first examine predictability of the global ocean carbon sink. Global annual air-sea CO_2 flux is potentially predictable for 2 and 4 years, when assessed with ACC and RMSE, respectively (Figure 1a). This means that initialized ensembles predict each other significantly better due to initialization than uninitialized ensembles. When assessing predictability with the same metric RMSE and predictability horizon definition as Séférian et al. (2018), global air-sea CO_2 flux predictability horizon in our model MPI-ESM-LR is comparable to CNRM-ESM1 (for details see supporting information section S3.3).

On a regional scale longer predictability horizons are found. For instance, the North and Subtropical Atlantic as well as the North and Subtropical Pacific have a predictability horizon of more than 6 years and thereby

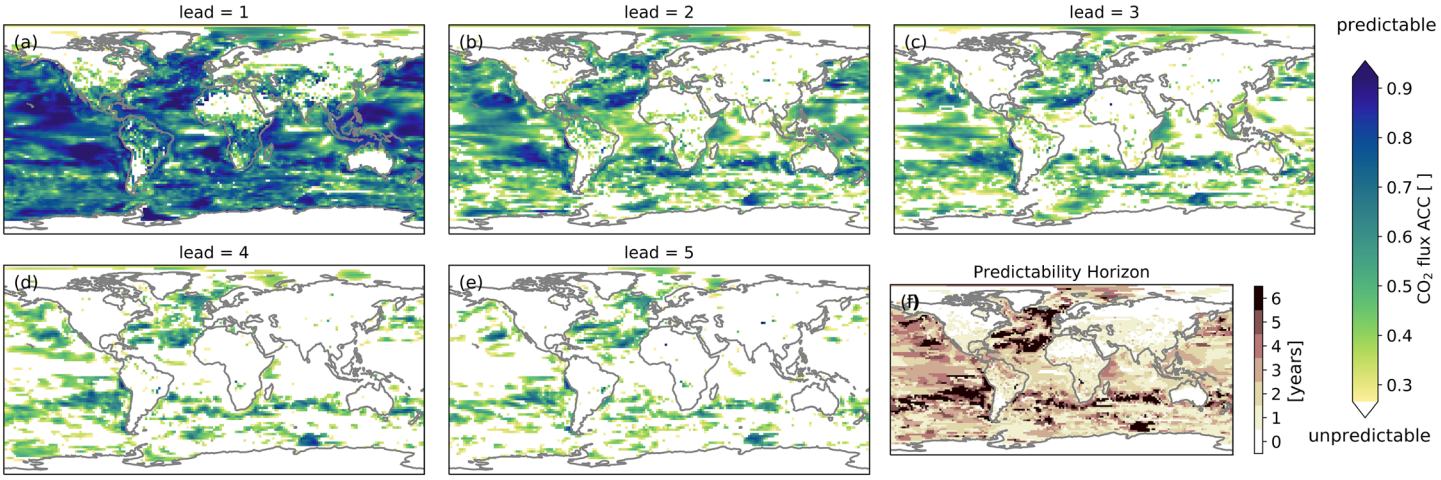

Figure 2. Spatial distribution of ACC-based surface CO_2 flux predictability: (a–e) predictability skill over the first five lead years. White areas indicate unpredictable areas where the uninitialized predictability skill exceeds the initialized skill at 5% probability based on bootstrapping with replacement ($N = 1,000$). (f) The predictability horizon marks the last significant lead year.

exceed the global oceanic predictability horizon (Figures 2 and S2) (Li et al., 2016, 2019; Lovenduski, Yeager, et al., 2019). These are regions of potentially predictable low-frequency air-sea CO_2 flux variations (Figure 3b). However, which areas support the predictability of air-sea CO_2 flux at the global scale? Assessing the magnitude of CO_2 flux spread between ensemble members measured by RMSE shows very little relevance of the subtropical gyres for the global oceanic carbon sink predictability (Figure S2). While ACC shows linear association independent of magnitude, RMSE puts weight on the distance between ensemble members. The subtropical gyres are predictable in ACC up to a decade (Figure 2), but this region is irrelevant for the global oceanic carbon sink predictability as indicated by the low RMSE from this region in contrast to other regions. The larger areas of the Southern Ocean are predictable only for 1 year. We attribute this little memory in CO_2 flux to the wind-driven circulation that sets local oceanic pCO_2 in the Southern Ocean, but atmospheric processes with less memory are harder to predict than oceanic processes (Pohlmann et al., 2004; Kirtman & Power, 2013; Zhang, Delworth, Yang, et al., 2017; Zhang, Delworth, & Jia, 2017). Owing to the large magnitude of Southern Ocean CO_2 flux and short predictability horizon, the Southern Ocean is the main limiting contributor to the predictability of the global air-sea CO_2 flux (Figures S2 and S5b). The imprint of longer regional predictability in the subtropical gyres is erased in the global predictability signal.

4. Predictability of Air-Land CO_2 Flux

Analogous to air-sea CO_2 flux, global air-land CO_2 flux is predictable for 3 years in ACC and for 2 years in RMSE, respectively (Figure 1b). These results are likewise comparable to Séférian et al. (2018).

On the regional scale, terrestrial biogeochemical processes have weaker memory toward atmospheric circulation in contrast to variations of marine biogeochemical processes responding to lower-frequency oceanic circulation (Figure 3a). Also, the imprints of initial conditions over land are faster lost than over the ocean as shown by faster and stronger increasing RMSE (Figures 1a and 1b and S2a–S2e). This results in lower regional predictability horizons on land than over the oceans (Figures 2f and S2f). Terrestrial CO_2 flux RMSE also increases to higher magnitudes due to the larger magnitudes of variability of terrestrial CO_2 flux (Figure S5b).

As for the oceanic CO_2 flux predictability, we assess the relevance of different region for the global terrestrial CO_2 flux predictability signal. Terrestrial high-latitudes and deserts CO_2 flux is not predictable at lead year one in both metrics indicating no interannual memory. The midlatitudes and tropical forests, which are highly variable (Figure S5b), are predictable up to 2 years, and thereby dominate the predictability of global terrestrial CO_2 flux (Figures 2 and S2). Zeng et al. (2008) find also highest ACC-based predictability skill in the tropics and midlatitudes over three seasons, while skill in other regions drops quickly within 3 months.

The high-RMSE midlatitudes and tropics are also strongly variable areas in which net primary production and heterotrophic respiration respond heavily to temperature and precipitation patterns associated with

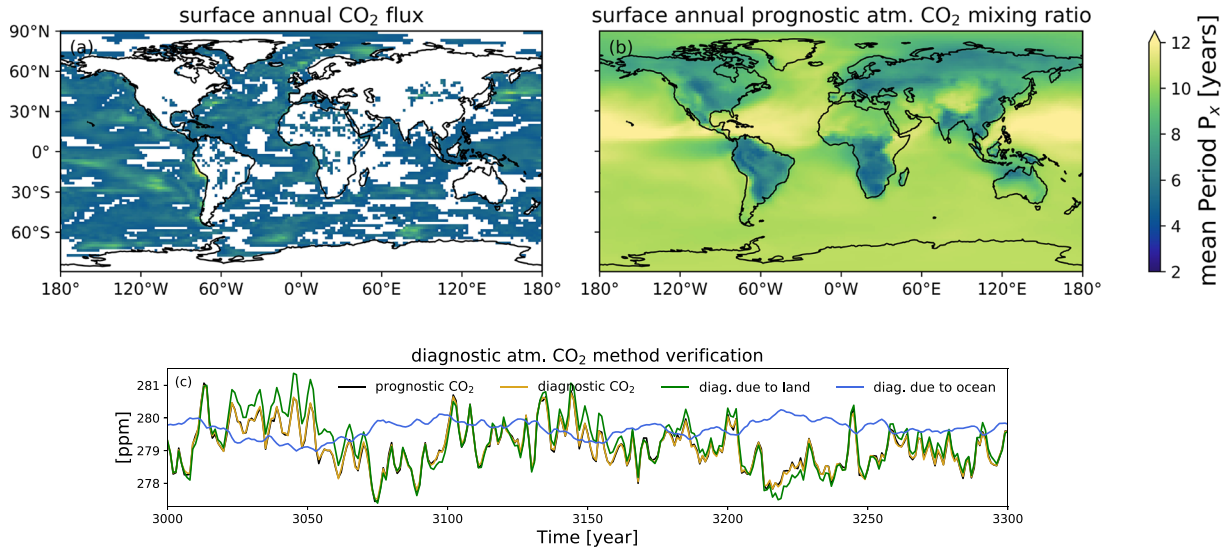

Figure 3. (a, b) Spatial distribution of variance-weighted mean period (Branstator & Teng, 2010) from the 300-year control simulation. In white areas the null hypothesis, that the variable is independent and random, is not rejected at 95% significance level based on$N = 500$ resamplings with replacement: (a) surface annual CO_2 flux and (b) prognostic annual surface atmospheric CO_2 mixing ratio. (c) Verification of diagnostic global annual atmospheric CO_2 mixing ratio (gold) with prognostic atmospheric CO_2 (black) based on temporally accumulated surface CO_2 flux of ocean (blue) and land (green) in the control simulation (see section 2.4).

ENSO (Figures S5, S6, S10, and S11). Our results support ENSO as the limiting process of predictability of air-land CO_2 flux (Betts et al., 2018; Keeling et al., 1995; Lovenduski, Bonan, et al., 2019; Jones et al., 2001; Zeng et al., 2005, 2008). Note that unlike global air-sea CO_2 flux predictability, global air-land CO_2 predictability is determined by regions which also show the longest predictability horizon.

5. Predictability of Atmospheric CO_2 Variations

Having examined the predictability of surface air-sea and air-land CO_2 fluxes, we can now explain their implications for predictability of the mixing ratio of surface atmospheric CO_2. Our results indicate that global atmospheric CO_2 mixing ratio is predictable for 3 years assessed with both metrics (Figure 1c). Yet, if the combined global terrestrial and oceanic CO_2 flux is predictable for 2 years only (Figures 1a and 1b), how can atmospheric CO_2 mixing ratio be predictable longer than the global surface CO_2 flux? As surface CO_2 flux directly changes atmospheric CO_2 content, atmospheric CO_2 mixing ratio can be regarded as an integrator of surface CO_2 flux (Keppel-Aleks et al., 2013). Reliable predictions in CO_2 fluxes up to 2 years result in longer predictability of their imprints on atmospheric CO_2.

To explain the limits of globally averaged atmospheric surface CO_2 predictability, we diagnose the individual contributions of the global terrestrial and oceanic carbon sinks to the atmospheric CO_2 mixing ratio. The predictability evolution of diagnosed atmospheric CO_2 mixing ratio is very similar to that of prognostic atmospheric surface CO_2 (Figures 1c and 1f). We find RMSE-based predictability horizons of 5 and 12 years for the global terrestrial and oceanic carbon sinks, respectively (Figures 1d and 1e). Consequently, the oceanic imprint on atmospheric CO_2 would be predictable for more than a decade, but this atmospheric CO_2 signal is overlaid by the dominant imprint of the terrestrial carbon sink on atmospheric CO_2 mixing ratio. Note that the predictability horizons of the individual contributions of oceanic and terrestrial carbon sinks to atmospheric CO_2 predictability are longer than when combined (Figure 1d–1f). We explain this by a lower combined global carbon sink variability compared to terrestrial CO_2 flux (Figure 3c). This has been also illustrated in a previous study showing that the ocean acts to suppress the atmospheric CO_2 variability. (Doney et al., 2006). Furthermore, the terrestrial and oceanic carbon sinks contributing to the atmospheric CO_2 reservoir are both driven by climate variability but respond on different timescales with different magnitudes of variability.

The spatial distribution of atmospheric CO_2 mixing ratio ACC predictability (Figure 4) is strongly influenced by the subjacent surface CO_2 flux predictability (Figures 2 and S2). Atmospheric surface CO_2 mixing ratio

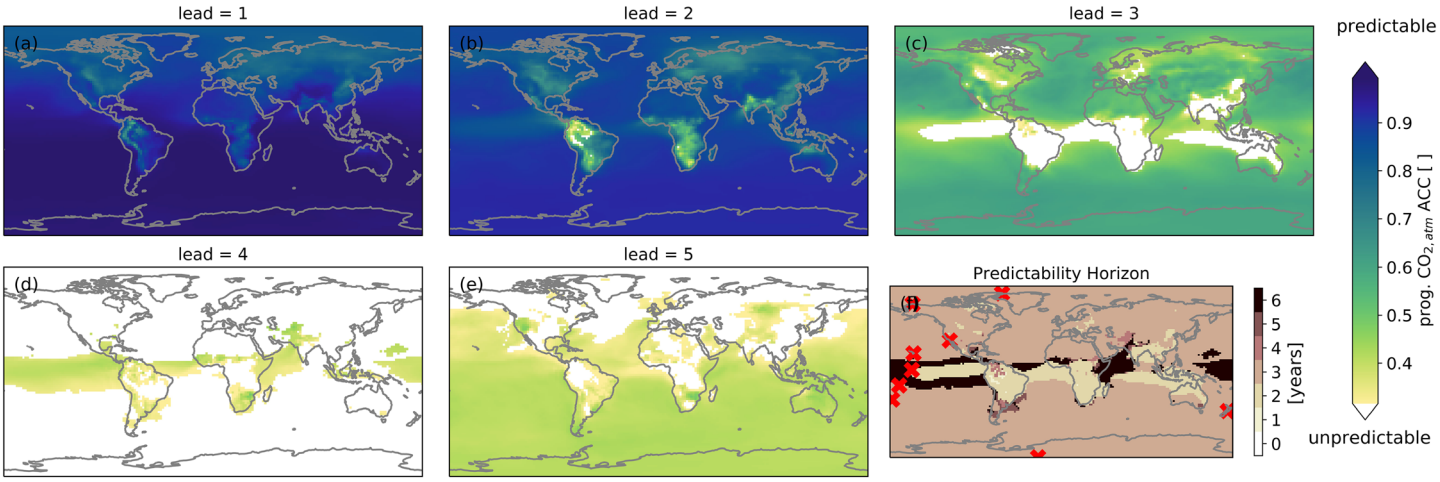

Figure 4. Spatial distribution of ACC-based atmospheric surface CO_2 predictability: (a–e) predictability skill over the first five lead years. White areas indicate unpredictable areas where the uninitialized predictability skill exceeds the initialized skill at 5% probability based on bootstrapping with replacement ($N = 1,000$). (f) The predictability horizon marks the last significant lead year. Red crosses show location of long-standing atmospheric CO_2 mixing ratio measurement stations (Keeling et al., 2005).

predictability persists over the first two lead years over the whole globe. Predictability decreases faster and stronger over areas close to the terrestrial biosphere. We explain this atmospheric CO_2 predictability, which is more skillful than for the surface CO_2 fluxes, with the integrator feature of the atmospheric CO_2 leading to lower-frequency variations (Figures 3a and 3b). Strong high-frequency perturbations from underlying air-land CO_2 flux are subsequently transported zonally on the time scale of months (Figures S5b, S6a, and S6b) and dilute as they spread across the globe on annual time scales (Figure 2).

Interestingly, the atmospheric surface CO_2 ACC predictability, which was lost over the tropics in lead year three, reemerges in the fourth lead year and persists over a few lead years over the tropical oceans and Southern Hemisphere (Figure 4). The low probabilities at which initialized forecasts beat uninitialized forecasts at lead years three to seven in ACC indicate this reemergence also on global scale (Figure 1c). However, this reemergence, while significant, is only slightly better than uninitialized ACC and therefore unlikely to be exploited as a feature of atmospheric CO_2 predictions. The limited number of twelve initializations, in which positive, neutral and negative ENSO initial states are not normally distributed, might also have triggered this reemergence. We explain the weak reemergence of ACC predictability with the oscillatory behavior of ENSO. Strong ENSO events inject a large pulse of CO_2 into the atmosphere in the tropics (Figure S5), which is zonally transported within months and dilutes meridionally over years (Figure S6). When the opposite ENSO phase is reached after the typical 3–4 years, ensemble members follow similar evolutions, which is measured with the ACC metric (Figure S12). As RMSE-based predictability skill lacks reemergence (Figure S4), this illustrates the different attributes of predictability skill and also shows the metric dependence of predictability horizon.

We further examine atmospheric CO_2 predictability in the context of the existing monitoring framework (Keeling, 2008). Due to the dilution and poleward transport of atmospheric CO_2 signals, the measurement station's predictability horizon is quite homogeneous over the ocean but can be higher in the tropics due to reemergence. The predictability horizon decreases in proximity to terrestrial CO_2 sources explainable by high terrestrial CO_2 flux variability (Figure 4f and Table S1). The highest ACC-based predictability horizon from Christmas Island station of 6 years is explained by the reemergence of atmospheric CO_2 predictability skill originating in the tropical forests, which reach this central Pacific island before the initial memory is lost (Figure 4f). The atmospheric CO_2 mixing ratio at the Mauna Loa station, which is an appropriate observational reference for global atmospheric CO_2 variations, is predictable in ACC for 3 years (Figures 4f, S1b, and S1c).

6. Summary and Conclusions

This study estimates decadal potential predictability of the fully coupled global carbon cycle. Using a perfect-model approach, we estimate the upper bound of ESM-based predictability in the context of two metrics. We find the RMSE-based predictability horizon of the global air-sea and air-land CO_2 flux to be 2 years (Li et al., 2019; Lovenduski, Bonan, et al., 2019; Lovenduski, Yeager, et al., 2019). While the ACC-based 4- and 3-year predictability horizons of the air-sea and air-land CO_2 flux signal are longer, RMSE-based predictability illustrates the relevance of individual sinks and regions for the predictability of atmospheric CO_2. Previous estimates of predictability horizon were higher (Séférian et al., 2018), but end up comparable when computed by our proposed bootstrapping methodology, which is backed by the established prediction frameworks (Goddard et al., 2013; Marotzke et al., 2016; Yeager et al., 2018).

We furthermore show that the global oceanic CO_2 flux predictability is mainly controlled by the Southern Ocean. The global terrestrial CO_2 flux predictability is dominated by the tropical forests and midlatitudes affected by ENSO. We find a different regional contribution pattern for land and ocean to global predictability of CO_2 flux. On land, predictability is maintained due to regions with the longest predictability horizon. On the contrary in the ocean, regions with longest predictability horizon are less relevant for the global signal.

The predictability horizon for climate modulated variations in the global atmospheric CO_2 mixing ratio is 3 years. The isolated effect of global terrestrial carbon sink on atmospheric CO_2 of 5 years dominates over the oceanic contribution of 12 years. Consequently, while the oceanic carbon cycle dampens the imprint of the terrestrial carbon cycle on atmospheric CO_2, the predictability of atmospheric CO_2 is limited by the terrestrial carbon sink.

The atmospheric CO_2 predictability horizon is globally quite homogeneous, except over parts of the tropical Pacific, because of a weak but significant re-emergence pattern which might also affect initialized predictions. As the dominant features of terrestrial CO_2 flux predictability in the tropics arise from ENSO, multiyear forecasts performed by initialized prediction systems for atmospheric CO_2 variations are expected to be challenging.

In this study, we demonstrate that ESM-based initialized forecasts of the global carbon cycle may deliver multiyear outlooks on the evolution of the atmospheric CO_2 mixing ratio. Hence, ESM-based predictions have the potential to constrain uncertainty of changes in atmospheric CO_2 due to internal climate variability in the near-term future and can thereby inform the pentadal stocktakes (UNFCCC, 2015).

Acknowledgments
Forecast verification was computed with the python package climpred (https://github.com/bradyrx/climpred/), which was codeveloped with Riley X. Brady from University of Colorado, Boulder. Scripts and primary data to reproduce this analysis are archived in https://hdl.handle.net/21.11116/0000-0004-8276-4 website. We acknowledge funding from European Union's Horizon 2020 research and innovation programme under Grant Agreement 821003 "Climate-fv (4C)" and 641816 (CRESCENDO). T. I. received funding from the Federal Ministry of Education and Research in Germany (BMBF) through the research programme "MiKlipII" (FKZ: 01LP1517B). We thank Leonard Borchert for internal review.

References

Bacastow, R. B. (1976). Modulation of atmospheric carbon dioxide by the Southern Oscillation. *Nature*, *261*(5556), 116–118.
Ballantyne, A. P., Alden, C. B., Miller, J. B., Tans, P. P., & White, J. W. C. (2012). Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years. *Nature*, *488*(7409), 70–72.
Betts, R. A., Jones Chris D., Knight Jeff. R., Keeling Ralph. F., Kennedy John. J., Wiltshire Andrew J., et al. (2018). A successful prediction of the record CO_2 rise associated with the 2015/2016 El Niño. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1760), 20170301.
Betts, R. A., Jones, C. D., Knight, J. R., Keeling, R. F., & Kennedy, J. J. (2016). El Niño and a record CO_2 rise. *Nature Climate Change*, *6*(9), 806–810.
Branstator, G., & Teng, H. (2010). Two limits of initial-value decadal predictability in a CGCM. *Journal of Climate*, *23*(23), 6292–6311.
Doney, S. C., Lindsay, K., Fung, I., & John, J. (2006). Natural variability in a stable, 1000-yr global coupled climate–Carbon cycle simulation. *Journal of Climate*, *19*(13), 3033–3054.
Eddebbar, Y. A., Rodgers, K. B., Long, M. C., Subramanian, A. C., Xie, S.-P., & Keeling, R. F. (2019). El Niño-like physical and biogeochemical ocean response to tropical eruptions. *Journal of Climate*, *32*, 262–2649.
Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap* (1st ed.). New York: Chapman and Hall/CRC.
(EU), European Union (2020). European Union (EU) Horizon 2020 Climate-Carbon Interactions in the Coming entury (4C) 821003. https://doi.org/10.1029/2019GL085311
Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958.
Frölicher, T. L., Joos, F., Raible, C. C., & Sarmiento, J. L. (2013). Atmospheric CO 2 response to volcanic eruptions: The role of ENSO, season, and variability: VOLCANOES AND THE GLOBAL CARBON BUDGET. *Global Biogeochemical Cycles*, *27*(1), 239–251.
Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., et al. (2019). Global carbon budget 2019. *Earth System Science Data*, *11*(4), 1783–1838.
Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., et al. (2013). A verification framework for interannual-to-decadal predictions experiments. *Climate Dynamics*, *40*(1-2), 245–272.
Griffies, S. M., & Bryan, K. (1997). A predictability study of simulated North Atlantic multidecadal variability. *Climate Dynamics*, *13*(7-8), 459–487.
Hawkins, E., Tietsche, S., Day, J. J., Melia, N., Haines, K., & Keeley, S. (2016). Aspects of designing and evaluating seasonal-to-interannual Arctic sea-ice prediction systems. *Quarterly Journal of the Royal Meteorological Society*, *142*(695), 672–683.

Heinze, C., Meyer, S., Goris, N., Anderson, L., Steinfeldt, R., Chang, N., et al. (2015). The ocean carbon sink – impacts, vulnerabilities and challenges. *Earth System Dynamics*, *6*(1), 327–358.

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., & Núñez-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. https://doi.org/10.1029/2012MS000178

Jolliffe, I. T., & Stephenson, D. B. (2011). *Forecast verification: A practitioner's guide in atmospheric science*. Chichester, UK: John Wiley & Sons, Ltd.

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., et al. (2016). C4MIP – The coupled climate–Carbon Cycle Model Intercomparison Project: Experimental protocol for CMIP6. *Geoscientific Model Development*, *9*(8), 2853–2880.

Jones, C. D., Collins, M., Cox, P. M., & Spall, S. A. (2001). The carbon cycle response to ENSO: A coupled climate–carbon cycle model study. *Journal of Climate*, *14*(21), 4113–4129.

Jones, C. D., & Cox, P. M. (2005). On the significance of atmospheric CO_2 growth rate anomalies in 2002–2003. *Geophysical Research Letters*, *32*(14), L14816. https://doi.org/10.1029/2005GL023027

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., et al. (2013). Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *Journal of Advances in Modeling Earth Systems*, *5*(2), 422–446. https://doi.org/10.1002/jame.20023

Keeling, R. F. (2008). Recording Earth's vital signs, *319*(5871), 1771–1772.

Keeling, C. D., Piper, S. C., Bacastow, R. B., Wahlen, M., Whorf, T. P., Heimann, M., & Meijer, H. A. (2005). Atmospheric CO_2 and $13CO_2$ exchange with the terrestrial biosphere and oceans from 1978 to 2000: Observations and carbon cycle implications. In Baldwin, I. T., Caldwell, M. M., Heldmaier, G., Jackson, R. B., Lange, O. L., Mooney, H. A., Schulze, E.-D., Sommer, U., Ehleringer, J. R., Denise Dearing, M., & Cerling, T. E. (Eds.), *A history of atmospheric CO_2 and its effects on plants, animals, and ecosystems*, Ecological Studies. New York, NY: Springer New York, pp. 83–113.

Keeling, C. D., Whorf, T. P., Wahlen, M., & van der Plichtt, J. (1995). Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980. *Nature*, *375*(6533), 666.

Keppel-Aleks, G., Randerson, J. T., Lindsay, K., Stephens, B. B., Keith Moore, J., Doney, S. C., et al. (2013). Atmospheric carbon dioxide variability in the Community Earth System Model: Evaluation and transient dynamics during the twentieth and twenty-first centuries. *Journal of Climate*, *26*(13), 4447–4475.

Kirtman, B., & Power, S. B. (2013). Projections and Predictability. In *Climate Change 2013: The Physical Science Basis Working Group I contribution to the IPCC Fifth Assessment Report* (p. 76).

Kröger, J., Pohlmann, H., Sienz, F., Marotzke, J., Baehr, J., Köhl, A., et al. (2017). Full-field initialized decadal predictions with the MPI earth system model: An initial shock in the North Atlantic. *Climate Dynamics*, *51*, 2593–2608.

Li, H., & Ilyina, T. (2018). Current and future decadal trends in the oceanic carbon uptake are dominated by internal variability. *Geophysical Research Letters*, *45*(2), 916–925. https://doi.org/10.1002/2017GL075370

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

Li, H., Ilyina, T., Müller, W. A., & Sienz, F. (2016). Decadal predictions of the North Atlantic CO_2 uptake. *Nature Communications*, *7*, 11,076.

Lin, S.-J., & Rood, R. B. (1996). Multidimensional flux-form semi-Lagrangian transport schemes. *Monthly Weather Review*, *124*(9), 2046–2070.

Lovenduski, N. S., Bonan, G. B., Yeager, S. G., Lindsay, K., & Lombardozzi, D. L. (2019). High predictability of terrestrial carbon fluxes from an initialized decadal prediction system. *Environmental Research Letters*, *14*(12), 124,074.

Lovenduski, N. S., Yeager, S. G., Lindsay, K., & Long, M. C. (2019). Predicting near-term variability in ocean carbon uptake. *Earth System Dynamics*, *10*(1), 45–57.

Marotzke, J., Müller, W. A., Vamborg, FrejaS. E., Becker, P., Cubasch, U., Feldmann, H., et al. (2016). MiKlip: A National Research Project on decadal climate prediction. *Bulletin of the American Meteorological Society*, *97*(12), 2379–2394.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., et al. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO 2. *Journal of Advances in Modeling Earth Systems*, *11*(4), 998–1038. https://doi.org/10.1029/2018MS001400

McKinley, G. A., Fay, A. R., Lovenduski, N. S., & Pilcher, D. J. (2017). Natural variability and anthropogenic trends in the ocean carbon sink. *Annual Review of Marine Science*, *9*(1), 125–150. https://doi.org/10.1146/annurev-marine-010816-060529

Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., et al. (2013). Decadal climate prediction: An update from the trenches. *Bulletin of the American Meteorological Society*, *95*(2), 243–267.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., et al. (2009). Decadal prediction: Can it be skillful? *Bulletin of the American Meteorological Society*, *90*(10), 1467–1486.

Murphy, A. H. (1988). Skill scores based on the mean square error and their relationships to the correlation coefficient. *Monthly Weather Review*, *116*(12), 2417–2424.

Park, J.-Y., Stock, C. A., Yang, X., Dunne, J. P., Rosati, A., John, J., & Zhang, S. (2018). Modeling global ocean biogeochemistry with physical data assimilation: A pragmatic solution to the equatorial instability. *Journal of Advances in Modeling Earth Systems*, *10*(3), 891–906. https://doi.org/10.1002/2017MS001223

Paulsen, H., Ilyina, T., Six, K. D., & Stemmler, I. (2017). Incorporating a prognostic representation of marine nitrogen fixers into the global ocean biogeochemical model HAMOCC. *Journal of Advances in Modeling Earth Systems*, *9*(1), 438–464. https://doi.org/10.1002/2016MS000737

Peters, G. P., Le Quéré, C., Andrew, R. M., Canadell, J. G., Friedlingstein, P., Ilyina, T., et al. (2017). Towards real-time verification of CO_2 emissions. *Nature Climate Change*, *7*(12), 848–850.

Pohlmann, H., Botzet, M., Latif, M., Roesch, A., Wild, M., & Tschuck, P. (2004). Estimating the decadal predictability of a coupled AOGCM. *Journal of Climate*, *17*(22), 4463–4472.

Séférian, R., Berthet, S., & Chevallier, M. (2018). Assessing the decadal predictability of land and ocean carbon uptake. *Geophysical Research Letters*, *45*, 2455–2466. https://doi.org/10.1002/2017GL076092

Schneck, R., Reick, C. H., & Raddatz, T. (2013). Land contribution to natural CO_2 variability on time scales of centuries. *Journal of Advances in Modeling Earth Systems*, *5*(2), 354–365. https://doi.org/10.1002/jame.20029

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., et al. (2013). Atmospheric component of the MPI-M Earth System Model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, *5*(2), 146–172. https://doi.org/10.1002/jame.20015

Toggweiler, J. R., Dixon, K., & Bryan, K. (1989). Simulations of radiocarbon in a coarse-resolution world ocean model: 1. Steady state prebomb distributions. *Journal of Geophysical Research*, *94*(C6), 8217–8242.

UNFCCC (2015). *Adoption of the Paris Agreement*. United Nations, Paris: United Nations Framework Convention on Climate Change. https://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf

Wilks, D. S. (2006). *Statistical methods in the atmospheric sciences, International Geophysics Series* (2nd, Vol. 91). Amsterdam; Boston: Academic Press.

Yeager, S. G., Danabasoglu, G., Rosenbloom, N. A., Strand, W., Bates, S. C., Meehl, G. A., et al. (2018). Predicting near-term changes in the Earth System: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model. *Bulletin of the American Meteorological Society*, *99*(9), 1867–1886. https://doi.org/10.1175/BAMS-D-17-0098.1

Zeng, N., Mariotti, A., & Wetzel, P. (2005). Terrestrial mechanisms of interannual CO_2 variability. *Global Biogeochemical Cycles*, *19*(1), GB1016. https://doi.org/10.1029/2004GB002273

Zeng, N., Yoon, J.-H., Vintzileos, A., Collatz, G. J., Kalnay, E., Mariotti, A., et al. (2008). Dynamical prediction of terrestrial ecosystems and the global carbon cycle: A 25-year hindcast experiment. *Global Biogeochemical Cycles*, *22*(4), GB4015. https://doi.org/10.1029/2008GB003183

Zhang, L., Delworth, T. L., & Jia, L. (2017). Diagnosis of decadal predictability of Southern Ocean sea surface temperature in the GFDL CM2.1 Model. *Journal of Climate*, *30*(16), 6309–6328.

Zhang, L., Delworth, T. L., Yang, X., Gudgel, R. G., Jia, L., Vecchi, G. A., & Zeng, F. (2017). Estimating decadal predictability for the Southern Ocean using the GFDL CM2.1 Model. *Journal of Climate*, *30*(14), 5187–5203.