https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics Discussions
Open Access

Trivial improvements of predictive skill due to direct reconstruction of global carbon cycle

Aaron Spring[1,2], István Dunkl[1,2], Hongmei Li[1], Victor Brovkin[1,3], and Tatiana Ilyina[1]

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

Correspondence: Aaron Spring (aaron.spring@mpimet.mpg.de)

Abstract.

State-of-the-art carbon cycle prediction systems are initialized from reconstruction simulations in which state variables of the climate system are assimilated. While currently only the physical state variables are assimilated, biogeochemical state variables adjust to the state acquired through this assimilation indirectly instead of being assimilated themselves. In the absence of comprehensive biogeochemical reanalysis products, such approach is pragmatic. Here we evaluate a potential advantage of having perfect carbon cycle observational products to be used for direct carbon cycle reconstruction.

Within an idealized perfect-model framework, we define 50 years of a control simulation under pre-industrial CO_2 levels as our target representing observations. We nudge variables from this target onto arbitrary initial conditions 150 years later mimicking an assimilation simulation generating initial conditions for hindcast experiments of prediction systems. We investigate the tracking performance, i.e. bias, correlation and root-mean-square-error between the reconstruction and the target, when nudging an increasing set of atmospheric, oceanic and terrestrial variables with a focus on the global carbon cycle explaining variations in atmospheric CO_2. We compare indirect versus direct carbon cycle reconstruction against a resampled threshold representing internal variability. Afterwards, we use these reconstructions to initialize ensembles to assess how well the target can be predicted after reconstruction. Interested in the ability to reconstruct global atmospheric CO_2, we focus on the global carbon cycle reconstruction and predictive skill.

We find that indirect carbon cycle reconstruction through physical fields reproduces the target variations on a global and regional scale much better than the resampled threshold. While reproducing the large scale variations, nudging introduces systematic regional biases in the physical state variables, on which biogeochemical cycles react very sensitively. Global annual surface oceanic pCO_2 initial conditions are indirectly reconstructed with an anomaly correlation coefficient (ACC) of 0.8 and debiased root mean square error (RMSE) of 0.3 ppm. Direct reconstruction slightly improves initial conditions in ACC by +0.1 and debiased RMSE by -0.1 ppm. Indirect reconstruction of global terrestrial carbon cycle initial conditions for vegetation carbon pools track the target by ACC of 0.5 and debiased RMSE of 0.5 PgC. Direct reconstruction brings negligible improvements for air-land CO_2 flux. Global atmospheric CO_2 is indirectly tracked by ACC of 0.8 and debiased RMSE of 0.4 ppm. Direct reconstruction of the marine and terrestrial carbon cycles improves ACC by 0.1 and debiased RMSE by -0.1 ppm.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

We find improvements in global carbon cycle predictive skill from direct reconstruction compared to indirect reconstruction. After correcting for mean bias, indirect and direct reconstruction both predict the target similarly well and only moderately worse than perfect initialization after the first lead year.

Our perfect-model study shows that indirect carbon cycle reconstruction yields satisfying initial conditions for global CO_2 flux and atmospheric CO_2. Direct carbon cycle reconstruction adds little improvements in the global carbon cycle, because imperfect reconstruction of the physical climate state impedes better biogeochemical reconstruction. These minor improvements in initial conditions yield little improvement in initialized perfect-model predictive skill. We label these minor improvements due to direct carbon cycle reconstruction *trivial*, as mean bias reduction yields similar improvements. As reconstruction biases in real-world prediction systems are even stronger, our results add confidence to the current practice of indirect reconstruction in carbon cycle prediction systems.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

1 Introduction

Predicting variations in weather and climate yields numerous benefits for economic, social, and environmental decision-making (Merryfield et al., 2020). Carbon cycle prediction systems have the ability of predicting the near-term evolution of CO_2 fluxes (Li et al., 2019; Lovenduski et al., 2019a, b) and atmospheric CO_2 (Spring and Ilyina, 2020; Ilyina et al., 2021) to constrain the large internal variability of the global carbon cycle (Spring et al., 2020). Predictions require a forecasting model and initial conditions representing observations, where the forecast is started from. However, due to sparse and temporally incomplete records, there is currently no global biogeochemical reanalysis product to initialize Earth System Models (ESMs). Therefore direct initialization of the carbon cycle is not possible. Practically, all state-of-the-art carbon prediction systems only initialize the physical climate, assuming that carbon cycle follows the initialized climate indirectly. However, this indirect carbon cycle initialization leaves the initial conditions of the carbon cycle unconstrained.

Here, we test how well indirect and direct carbon cycle reconstructions in an ESM initialize the carbon cycle in a perfect-model target reconstruction framework [Tab. 1 presents an overview which variables are reconstructed in which simulation]. We use the term reconstruction to describe methods of initialization of climate and the carbon cycle. Reconstructions aim to reproduce the evolution of the target, like a reanalysis product, in the ESM. Furthermore, we use the term "carbon cycle" to describe the processes exchanging carbon across the surface boundary between land, atmosphere and ocean, represented here by the air-land and air-sea CO_2 fluxes. We ask the following research questions:

- How well can initial conditions be reconstructed in the global carbon cycle?
- Can initialization of the carbon cycle improve predictive skill of the carbon cycle?

In this perfect-model target reconstruction framework, we have perfect knowledge about the ground truth and a perfect model. Literally speaking, this asks how well could perfect observations be reconstructed in an ESM.

Originally, data assimilation is used to align the model state to an observations-based state, generally a reanalysis product. However, here we use the same data assimilation technique to assess how well variables can be reconstructed in an idealized setup. Thus, reconstruction in a climate model interferes with the freely running climate model yielding gains and drawbacks.

The main advantage of climate reconstruction is that the reconstruction forces the climate model to follow the target.

The main handicap associated with reconstruction is that the violation of mass conservation and that the model dynamics and feedbacks are obstructed (Zhu and Kumar, 2018). Consequently, circulation fields may change and this has severe consequences on the biogeochemical tracer distributions in the ocean and carbon pools on land, because they are so sensitive and adapted to the previous climate state. Therefore, reconstructions often lead to biases. A partial solution can be bias removal by post-processing, which is feasible if the bias does not change the climate or ecosystem regime alltogether.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Even if biogeochemical reanalysis products were available, it is unclear whether the reconstruction benefits correct these handicaps.

The lack of reanalysis products available for the reconstruction of carbon cycle initial conditions is often assumed as a weakness of the current predictions systems (Li et al., 2016; Séférian et al., 2018; Lovenduski et al., 2019b, a; Li et al., 2019; Ilyina et al., 2021), but to our knowledge an elaborate assessment is missing. The literature presents two alternative approaches to test the quality of reconstructed initial conditions:

Servonnat et al. (2015) nudge only ocean surface temperature, salinity and sea-ice and assess how well this surface reconstruction penetrates into the subsurface ocean physics, without addressing biogeochemistry in their analysis. This target reconstruction approach in a perfect-model framework allows to directly assess the quality of reconstructed initial conditions, which is useful and practical to know for forecaster issuing a forecast.

In a recent study Fransner et al. (2020) ask whether the initial conditions of ocean biogeochemistry or the initial conditions of ocean physics have a stronger influence on multi-year predictions using perfect-model twin perturbed initial conditions experiments. In the first set of hindcasts, they take identical initial conditions of the ocean physics to ensure identical climate evolution but completely different states from different members for ocean biogeochemistry. In the other set of hindcasts, they slightly perturb the ocean physics to force members on disseminated climate evolutions while keeping the ocean biogeochemistry initial conditions identical. They find that ocean biogeochemistry initial conditions did not affect predictive skill later than the first lead year. Their approach asks the more theoretical question whether initial conditions of ocean biogeochemistry matter compared to ocean physics initial conditions.

We go beyond previous studies by using the methodology of Servonnat et al. (2015), with the aim to understand the quality of initial conditions reconstruction. In contrast to Fransner et al. (2020), we aim to answer the questions about quality of initial conditions produced by different reanalysis approaches. We expand the scope by addressing the global carbon cycle, including the land, ocean and atmospheric compartments and the interactive exchange of CO_2 fluxes between them. We then assess the influence of these previously reconstructed carbon cycle initial conditions for initialized predictions of the natural carbon sinks and atmospheric CO_2. We focus on the global carbon cycle, because atmospheric CO_2 is well-mixed greenhouse gas controlled by carbon cycling on land and in the ocean (Friedlingstein et al., 2020).

After explaining the approach of target reconstruction in section 2, we separate reconstruction and its implication on predictive skill in two parts: We first evaluate physical reconstruction representing indirect biogeochemical reconstruction in sections 3.2.1, 3.3.1 and 3.4.1 for spatial maps and global aggregated values. Furthermore, we test the potential tracking performance increase for direct carbon cycle reconstruction in sections 3.2.2, 3.3.2 and 3.4.2. In part two, we assess the impact of different reconstruction methods on initial conditions predictive skill in section 4. Finally, the main findings and conclusions of this study are summarized in section 5.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

2 Methods

2.1 Model Description

We use the Max Planck Institute ESM (MPI-ESM)(Mauritsen et al., 2019), which was also used in the Coupled Model Intercomparison Project Phase 6 (CMIP6) framework (Eyring et al., 2016). We run the model MPI-ESM1-2-LR, the low resolution configuration with a horizontal resolution of about 1.8° in the atmosphere and on land, and about 1.5° in the ocean with daily coupling of the compartments (Mauritsen et al., 2019). The time steps of the atmosphere/land and the ocean are 600 and 4320s, respectively (Mauritsen et al., 2019). We run the model with prognostic atmospheric CO_2 mixing ratio under pre-industrial conditions (*esm-piControl*).

The marine biogeochemical cycle model HAMOCC (Ilyina et al., 2013) is embedded in the ocean general circulation model MPIOM (Jungclaus et al., 2013). HAMOCC includes carbonate chemistry and an extended NPZD-type cycle including nutrient-light-temperature co-limitation and nitrogen-fixating cyanobacteria (Paulsen et al., 2017). The land carbon cycle model JSBACH includes dynamic vegetation, wildfires, soil carbon decomposition and storage (Schneck et al., 2013). The atmospheric general circulation model ECHAM6 transports the three-dimensional atmospheric prognostic atmospheric CO_2 tracer with a flux-form semi-Lagrangian scheme (Lin and Rood, 1996; Stevens et al., 2013).

2.2 Perfect-Model Target Reconstruction Framework

Simulations in a perfect-model target reconstruction framework aim to reproduce the target climate evolution (Griffies and Bryan, 1997; Servonnat et al., 2015), but are started from an independent initial state, i.e. does not match the initial conditions of the target, while both target and initial conditions share the same climatology. Practically, we choose a 50-year target period from model years 1850 to 1900 and an uncorrelated restart file from model year 2005 from the pre-industrial control simulation (esm-piControl) submitted for the MPI-ESM1-2-LR model for C4MIP (Jones et al., 2016) in CMIP6 (Eyring et al., 2016).

In order to assess how many variables are needed to sufficiently reconstruct climate and biogeochemical cycles, we first perform reconstruction simulations only reconstructing physical state variables in atmosphere and/or ocean [Tab. 1]. In these simulations, carbon cycle is only indirectly effected by the reconstruction of physical variables. In further simulations, we test how much carbon cycle states improve with respect to the target, when also carbon cycle state variables are reconstructed directly.

2.3 Reconstruction Simulations

Newtonian or Haney (1974) relaxation, which is often called *nudging*, is a simple four-dimensional assimilation technique that dynamically reconstructs variables in an ESM. A non-physical relaxation term with relaxation coefficient R (units $1/s$) is added to the prognostic equation to drag the model variable X, which is subject to model forcing F_m, towards its target X_t:

$$\frac{\delta X}{\delta t} = F_m(X) + R(X_t - X) \tag{1}$$

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

	Reconstructed variables for each realm (nudging relaxation time-scale)				
Reconstruction simulations	Atmosphere: temperature (24h) surface pressure (24h) vorticity (6h) divergence (48h)	Ocean (60d): temperature salinity	Sea-ice (60d): concentration thickness	Ocean carbon (60d): DIC alkalinity	Land: all JSBACH (reset restart files Jan 1st)
$\text{indirect}_{\text{ATM only}}$	x				
$\text{indirect}_{\text{OCEAN only}}$		x			
indirect	x	x	x		
direct	x	x	x	x	x

Table 1. Overview over different reconstruction simulations. The first column title marks the labels of the experiments as used in the manuscript. The reconstruction strength as relaxation time-scales is noted in brackets, where h denotes hours and d days. The land carbon cycle is not dynamically reconstructed at each time step, but by a hard reset of restart files each January 1st from the target run. These land restart files include carbon and nitrogen pools, soil physics (moisture, temperature, snow cover), vegetation cover (plant functional types distribution), and canopy (leaf area index).

For reconstruction of the dynamics of the ocean, we reconstruct three-dimensional temperature and salinity as well as sea-ice concentration and thickness [Tab. 1]. We label this reconstruction indirect [Tab. 1] from the carbon cycle's perspective, as the carbon cycle is not reconstructed directly, but rather indirectly follows the reconstructed physical climate. Observational ocean data is often not available at each model time step. Therefore, we interpolate (without adjustments to conserve the temporal mean) monthly model target output to daily frequency as done in previous studies (Pohlmann et al., 2009). We chose a 60-day ocean relaxation time (converted to units $1/s$) like Servonnat et al. (2015) in their perfect-model target reconstruction study. Reconstructions towards observations usually choose a stronger nudging strength (Pohlmann et al., 2009; Keenlyside et al., 2008).

We reconstruct the physics of the atmosphere by nudging temperature, vorticity, divergence and the logarithm of surface pressure (Pohlmann et al., 2019). The high-frequency 6h output serves as the target and is nudged into all 63 spherical harmonics. Temperature and the logarithm of surface pressure is nudged with a relaxation timescale of 24 hours, vorticity is nudged with a relaxation timescale of 6 hours and divergence is nudged with a relaxation timescale of 48 hours. Relaxation coefficients are converted to units $1/s$ and are taken from previously used setups (Rast et al., 2012; Pohlmann et al., 2019; Li et al., 2019). Only nudging the atmosphere with these quite short relaxation times is similar to the forced simulations, such as the Model Intercomparison Projects (MIP) for ocean (OMIP) (Griffies et al., 2016; Orr et al., 2017), land (LMIP) (van den Hurk et al., 2016) and Global Carbon Budget (Friedlingstein et al., 2019) simulations, where (atmospheric) external boundary forcing drives the carbon cycle.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

For reconstructions of oceanic carbon cycle, we use the same nudging approach and strength as for physical ocean reconstruction but on different variables. To reconstruct the components of carbonate system, we nudge three-dimensional dissolved inorganic carbon (DIC) and total alkalinity [Tab. 1].

Unfortunately, there is no nudging module available in the land surface model JSBACH. Here we choose to manually reset the initial conditions every January 1st to the target values instead of the dynamic reconstruction at each time step. We thereby reconstruct land biogeochemistry and land surface physics such as soil moisture by resetting all restart variables every year. In supplementary information section C, we provide several sensitivity analysis by resetting land only every two or five years and resetting the ocean every year in the same way.

2.4 Evaluating Tracking Performance

We assess how well a reconstruction simulation tracks the target using various tracking performance metrics. We compare the target with reconstructions in the various metrics showing different attributes of tracking performance over 10-year windows: bias, anomaly correlation coefficient and root-mean-square-error. For equations please consult the supplementary [sec. A]. We calculate tracking performance over 10-year chunks to capture the variability within tracking performance and reduce the influence of drifts over time.

The non-physical relaxation terms in the prognostic equations can disturb the dynamics in the ESM and introduce biases defined as the differences in the reconstruction compared to the freely running target over time.

The anomaly correlation coefficient skill score (ACC) shows the linear association between the reconstruction and the target over time and therefore measures synchronous evolution while ignoring bias.

The root-mean-square-error (RMSE) takes into account bias and measures the second-order euclidian distance between reconstruction and target simulation over time. Under the assumption that persistent biases can be removed by post-processing, we also assess RMSE after having the mean monthly bias removed.

How do we evaluate that a reconstruction is good enough? While good enough is a subjective judgement, we resample the target simulation along the time dimension with a block length of ten years to check the metric of two randomly compared 10-year chunks. We consider the 95th quantile threshold for ACC and 5th quantile threshold for the remaining distance-based metrics as a baseline of internal variability to be a good enough reconstruction (Efron and Tibshirani, 1993).

2.5 Perfect-Model Predictive skill Framework

In the second part of this study, we perform initialized perfect-model experiments (as in Spring and Ilyina, 2020)]. The simulations in the perfect-model framework are started from the indirect and direct reconstructions as well the target representing perfect initial conditions. We take 19 initialization states chosen every second January 1st between 1860 and 1896, after allowing a 10 years adjustment phase after reconstructions were started. From each of those states from different reconstruction simulations, we fork five ensemble members and simulate three lead years. We generate ensemble members by perturbing the stratospheric horizontal diffusion by 1.0000{member} in the first year. This member generating approach provokes only tiny initial perturbations to the climate system as the ocean and land initial conditions remain identical.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

2.6 Predictive Skill Quantification

We compute predictive skill as the root-mean-square-error (RMSE) between the ensemble member mean and the target as verification (Wilks, 2006; Jolliffe and Stephenson, 2011) [sec. A]. RMSE measures the second-order euclidian distance between forecast members and the verification target. RMSE describes how the individual ensemble members spread over lead time. Please find additional details about the predictive skill metrics and the uninitialized bootstrapping in Spring and Ilyina (2020).

3 Reconstruction in an Earth-System-Model

As the carbon cycle is sensitive to the climate evolution, we first assess how well the physical climate is reconstructed. Therefore, we first evaluate the physical climate state after reconstruction in subsection 3.1]. Afterwards, we assess how these different reconstructions of physical climate indirectly reconstruct the ocean, land and atmospheric carbon cycle in subsections 3.2.1, 3.3.1 and 3.4.1]. The direct reconstructions are shown in subsections 3.2.2, 3.3.2 and 3.4.2].

3.1 Reconstruction of Physical Climate

Reconstructing the ocean and/or the atmosphere systematically disturbs the freely evolving model, which leads to annual mean biases with respect to the original target. The gray stippling in figure 1 shows where this reconstruction bias is larger than the randomly resampled 95% mean absolute error threshold and therefore labeled significantly exceeding internal variability.

All reconstruction yielded identical results for winds and precipitation tracking performance. Reconstructing the ocean and/or the atmosphere introduces biases of up to 0.6 m/s in zonal and 0.9 m/s in meridional 10-m wind speed depicting a southward-shift of the Intertropical Convergence Zone (ITCZ). This bias results in a significant weakening of the equator-ward latitudinal winds, whereas extra-tropical latitudinal intensifies [Fig. 1a]. The intensification and equator-ward shift of the easterly trade winds and weakening of the southern hemisphere westerlies are both not significant [Fig. 1b]. Precipitation is heavily impacted by these biases in atmospheric transport across many regions of the globe. Precipitation significantly shifts southward at the equator with changes of more than 1 mm/day and increases in Western Canada, Western Russia and Southern Australia [Fig. 1c]. Unlike the previously described variables, the 2m-temperature bias depends on whether the ocean is reconstructed or not. Just reconstructing the ocean temperature and salinity ($\text{indirect}_{\text{OCEAN only}}$) leads to small, negative and significant biases in the tropical Atlantic and West Pacific. Also Northern and Southern Africa as well as the Amazon and China is subject to a small cold bias, whereas Saharan Africa and Southeast Asia gets substantially warmer. The polar regions cool significantly [Fig. 1d]. Only reconstructing the atmosphere ($\text{indirect}_{\text{ATM only}}$) leads to a warm bias nearly across the all oceans, but less cold bias over Northern and Southern Africa as well as China [Fig. 1e]. Combining atmosphere and ocean reconstruction (indirect) reduces the overall temperature bias, especially over the oceans [Fig. 1f].

While the above explained biases are liabilities of reconstructions, the linear association measured by the Anomaly Correlation Coefficient (ACC) clearly benefits from reconstruction. Reconstruction clearly recreates climate variability of the target

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics
Discussions

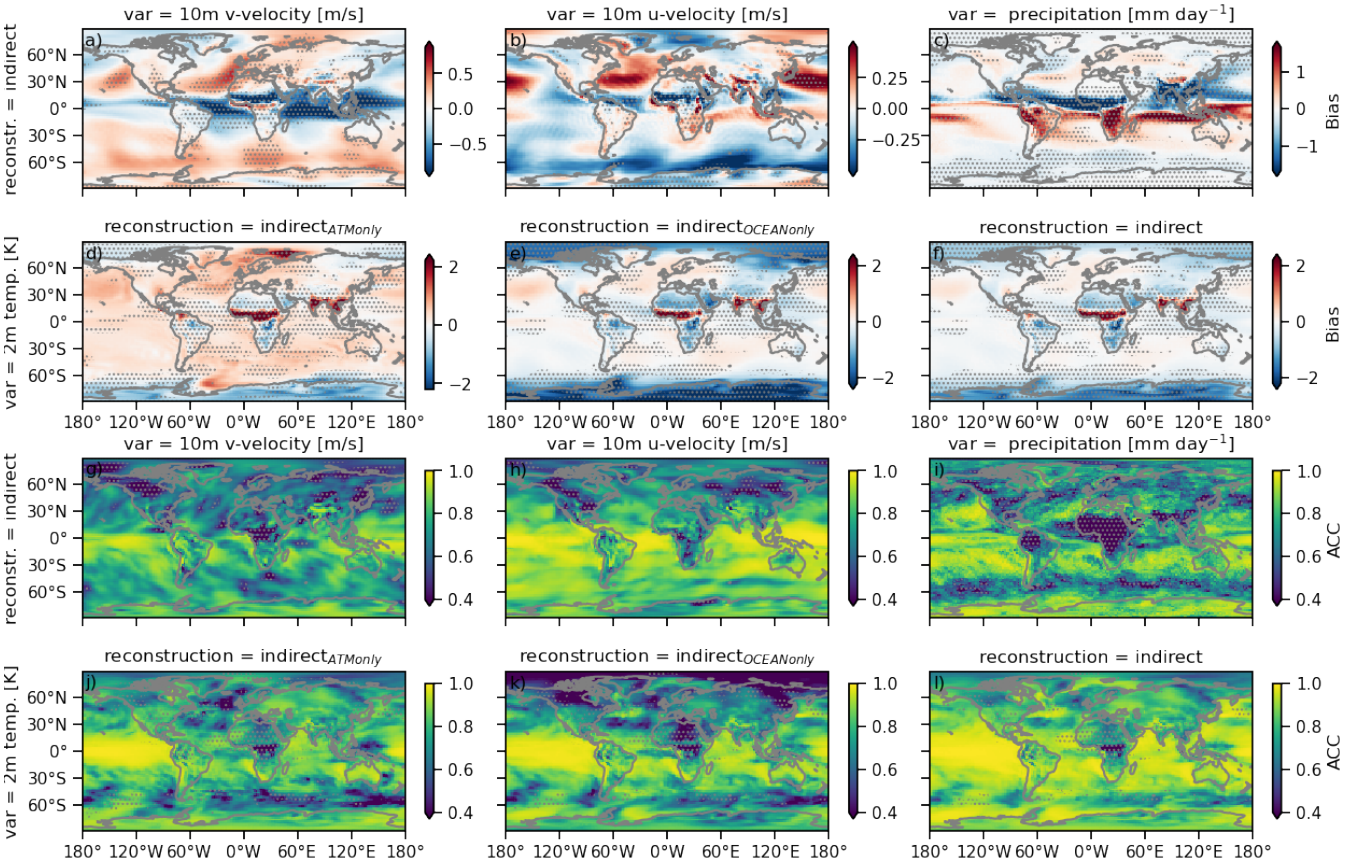

Figure 1. Spatial distribution of the bias (construction - target) (a-f) and anomaly correlation coefficient (ACC) (g-l) of different indirect carbon cycle reconstructions relative to the target over 10-year running windows of annual means [see A]. The reconstruction metrics for 2m temperature are shown for the indirect$_{\text{ATM only}}$ (d,j), indirect$_{\text{OCEAN only}}$ (e,k) and indirect reconstruction (f,l). Because of identical reconstruction skill for all indirect methods, only the indirect reconstruction is shown for other variables zonal westward 10m wind (a,g), and meridional northward 10m wind (b,h), and precipitation (c,i) Gray stippling shows where the metric exceeds the 95th (for a-f) or 5th (for g-l) percentile threshold from random target block resampling, i.e. the reconstruction is statistically significantly exceeding internal variability.

[Fig. 1g-l]. The running 10-year correlation between the target and the reconstruction in atmospheric variables is in most grid cells above 0.4 and significantly better than the randomly resampled threshold. Reconstruction over the oceans is more successful in the tropics than in the extra-tropics, where the Northern and Southern Hemisphere mid-latitude westerlies have low, but still significant correlation. Generally, the atmosphere above the ocean is better reconstructed than above land showing the stabilizing effect of an internally consistent ocean reconstruction on the atmosphere [Fig. 1g-l]. The Southern Hemisphere tropical convergence of winds is well reconstructed, but the meridional winds in central Canada and tropical Africa are not significantly reconstructed [Fig. 1g]. Also zonal winds across North America, Southern Africa and Siberia have low correlation with the target, but the tropical zonal winds are very well reconstructed [Fig. 1h]. Precipitation from the central Atlantic over

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

central Africa is worse reconstructed than the resampled threshold, also the extratropical westerlies have low correlation with the target [Fig. 1i]. Temperature is well reconstructed in the tropical oceans [Fig. 1j-l]. Reconstructing both atmosphere and ocean (indirect) improves 2m temperature correlation better than only reconstructing a single realm. The indirect carbon cycle reconstruction is significantly better than the resampled threshold except in central Africa, where the ITCZ shift changes the climate regime [Fig. 1l].

This physical bias due to reconstruction, especially in the tropics, can be explained by the sensitivity of atmosphere-ocean coupling to perturbation induced by nudging (Milinski et al., 2016). The reconstruction of ocean and atmospheric variables is perfectly aligned with the model climatology into that same model. Hence, the reconstruction error does not arise from inconsistent observations, but from the perturbed interaction of atmospheric and oceanic dynamics. While reconstructing an increasing set of variables shows that nudging can be an efficient way to reconstruct variability (Jeuken et al., 1996), this reconstruction is biasing the climate state in the tropics at the same time (also explained in Zhu and Kumar, 2018). Nudging atmospheric and ocean dynamics including sea-ice all at once (indirect reconstruction), as is often done in state-of-the-art carbon cycle prediction systems, brings large-scale improvements over random resampling and atmosphere-only ($\text{indirect}_{\text{ATM only}}$) reconstruction, but strong regional biases remain [Fig. 1].

3.2 Reconstruction of the Oceanic Carbon Cycle

3.2.1 Indirect Reconstruction of the Oceanic Carbon Cycle

How do these regional physical biases affect the reconstruction of oceanic carbon cycle? In order to assess the tracking performance in the indirect reconstruction of the oceanic carbon cycle, we focus on air-sea CO_2 flux and the state variable of the ocean carbon sink surface ocean pCO_2, which is the oceanic driver of air-sea CO_2 flux (Lovenduski et al., 2019b).

Reconstructing only the atmospheric dynamics ($\text{indirect}_{\text{ATM only}}$) leads to strong positive biases across large parts of the global ocean, which can be reduced by also reconstructing oceanic temperature and salinity (indirect) [Fig. 2a,b,d,e]. The weakening of the Southern hemisphere westerly winds decreases the magnitude of air-sea CO_2 flux, but more importantly reduces the Southern hemisphere overturning circulation and upwelling of carbon-rich waters, which leads to increased Southern Ocean carbon uptake [Fig. 2b,e]. The intensification of easterly trade winds [Fig. 1b] strengthens upwelling and therefore higher pCO_2 in the tropical Atlantic [Fig. 2b] (Lefèvre et al., 2013). The bias pattern of air-sea CO_2 flux is dominated by the bias of pCO_2 (Lovenduski et al., 2019b) [Fig. 2b,e].

The variations in the oceanic carbon cycle, described by the correlation coefficient, are better reconstructed than the resampled threshold. Indirect reconstruction of oceanic and atmospheric dynamics greatly improves tracking performance over atmosphere-only $\text{indirect}_{\text{ATM only}}$ reconstruction and enables largely a correlation above 0.7 [Fig. 3b,e].

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics Discussions
Open Access
EGU

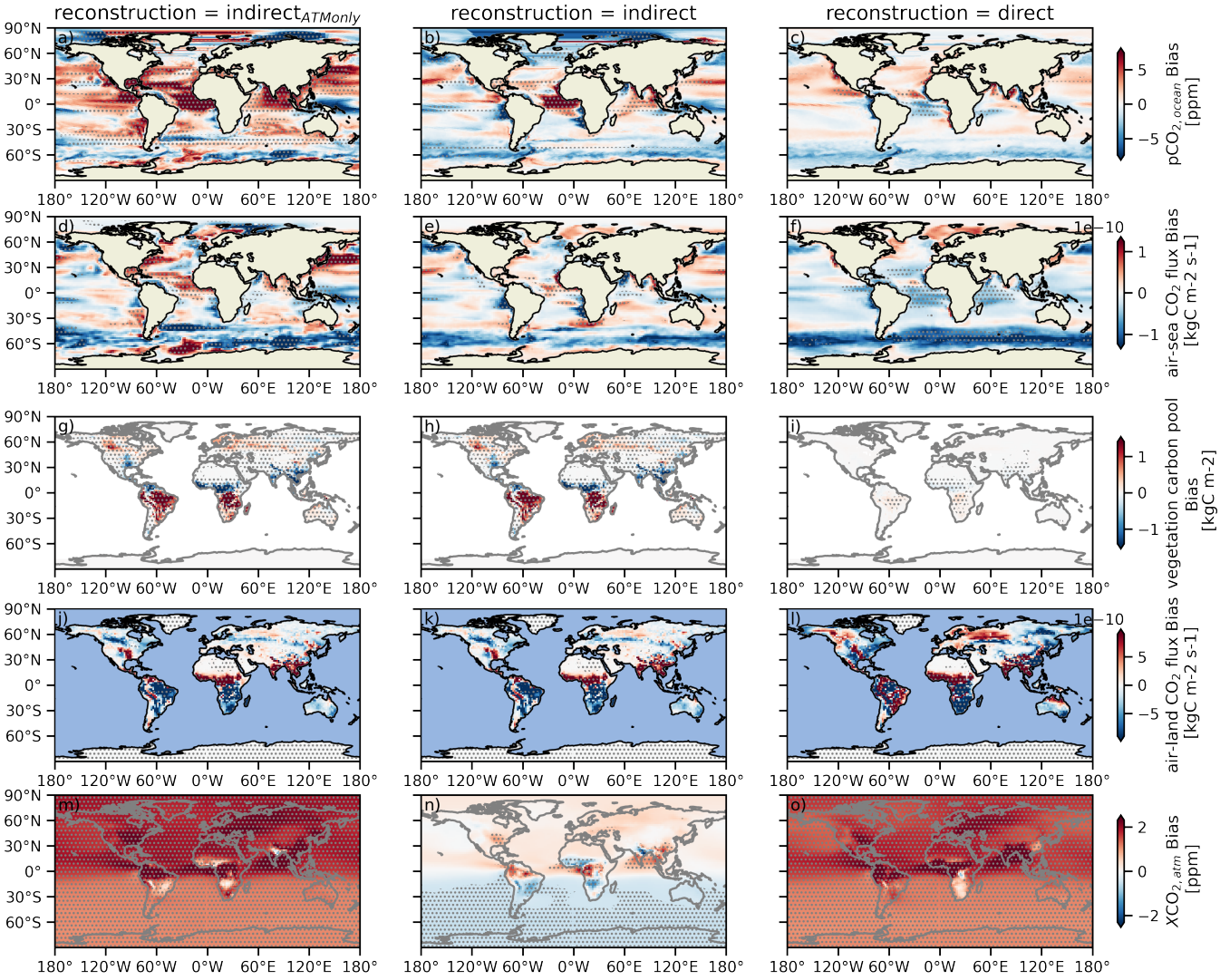

Figure 2. Spatial distribution of the bias between the target and different indirect carbon cycle reconstruction methods over 10-year running windows of annual means [see A]. Columns show the different carbon cycle reconstruction methods [see Tab.1]. Rows show the different variables: the ocean carbon cycle is represented by (a-c) the partial pressure of surface CO_2 in the ocean (pCO_2) and (d-f) surface air-sea CO_2 flux (negative values indicate carbon uptake by the ocean); the land carbon cycle is represented by (g-i) the vegetation carbon pools and (j-l) air-land surface CO_2 flux (negative values indicate carbon uptake by land); and the atmospheric carbon is represented by (m-o) the atmospheric CO_2 mixing ratio (XCO_2). Gray stippling shows where the bias exceeds the 5th percentile mean absolute error threshold from random target block resampling, i.e. the reconstruction is statistically significantly different to internal variability.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics
Discussions

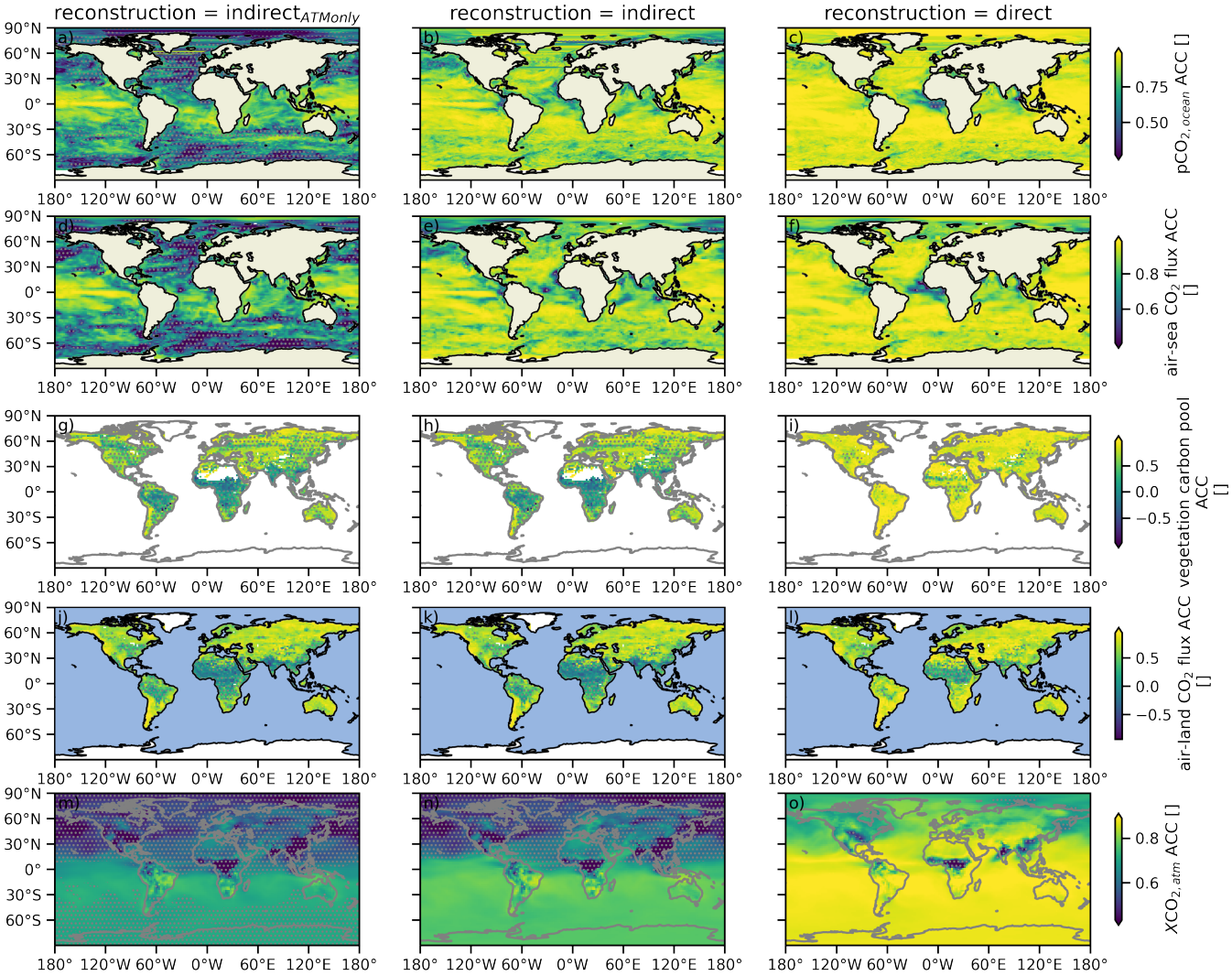

Figure 3. As Fig. 2 but for the anomaly correlation coefficient (ACC). Gray stippling shows where the ACC exceeds the 95th percentile ACC threshold from random target block resampling, i.e. the reconstruction is significantly different to internal variability.

3.2.2 Direct Reconstruction of the Oceanic Carbon Cycle

Next, we compare the previously shown indirect carbon cycle reconstruction with direct carbon cycle reconstruction by nudging dissolved inorganic carbon (DIC) and alkalinity (ALK) towards the target.

While direct oceanic carbon cycle reconstruction reduces the magnitudes of the bias across the ocean, biases are still evident [Fig. 2c,f]. These biases are caused by the physical biases, which the dynamical oceanic carbon cycle model is sensitive to. Hence, the biased ocean physics inhibits additional improvements in tracking performance from direct ocean carbon reconstruction.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics
Discussions
Open Access
EGU

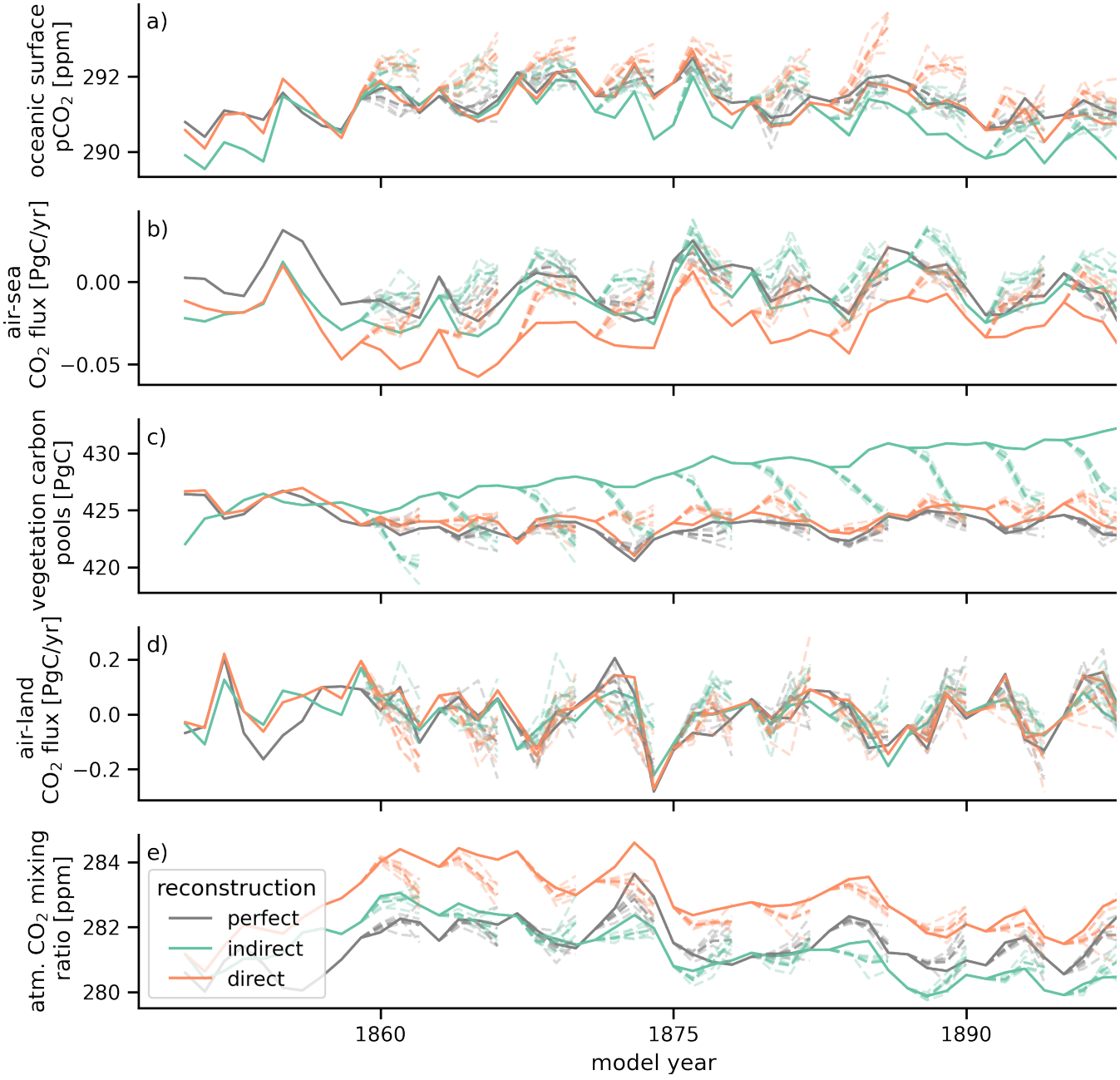

Figure 4. Evolution in global annual mean of (a) surface ocean pCO_2 (b), air-sea surface CO_2 flux (negative values indicate carbon uptake by the ocean) (c), vegetation carbon pools (g-i), air-land surface CO_2 flux (negative values indicate carbon uptake by land) (d) and atmospheric CO_2 mixing ratio (e). The target (gray) is quite well tracked by the indirect (green) and direct (orange) carbon cycle reconstruction. The solid line shows the different reconstruction simulations, the dashed lines show the initialized ensembles started from the different reconstructions.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Direct oceanic carbon cycle reconstruction improves the already high correlations across the oceans [Fig. 3c,f]. The resampled threshold is surpassed nearly everywhere. Only coastal areas, especially in the Eastern tropical Atlantic with strong wind and precipitation biases, have a correlation below 0.7.

Section 3.2 shows how well indirect and direct reconstruction of the ocean carbon cycle work. While the direct reconstruction has larger biases, it also brings even higher correlation.

3.3 Reconstruction of the Land Carbon Cycle

3.3.1 Indirect Reconstruction of Land Carbon Cycle

How do these regional physical biases affect the reconstruction of the land carbon cycle? In order to assess the tracking performance in the best indirect reconstruction of the land carbon cycle, we focus on the state variable cVeg, which represents carbon storage in vegetation (leaves, stems, roots) and drives air-land CO_2 flux and hence the land carbon sink.

For the land carbon cycle, the reconstruction of the ocean temperature and salinity did not matter, when atmospheric temperature was also reconstructed [Figs. 2, 3]. Indirect reconstruction leads to biases compared to the target in carbon storage, and in particular cVeg [Fig. 2g,h], as the land carbon cycle is very sensitive to changes in atmospheric circulation, which are strongest in the tropics due to the ITCZ shift. In the Amazon and Southern Africa, the air-land CO_2 bias increases, most likely caused by the strong positive precipitation bias in these regions [Fig. 1c; 2j,k]. Conversely, the carbon sink in Southeast Asia and central Africa has a carbon release bias due to less precipitation and a warm bias [Fig. 2j,k].

The reconstruction correlations in the land carbon cycle are much lower than for the oceanic carbon cycle. cVeg is well reconstructed in the extratropics, but the biases in the tropics result in correlations with the target lower than the resampled threshold [Fig. 3g,h]. Air-land CO_2 shows the same patterns with lower correlations, which are below the resampled threshold in the tropics [Fig. 3j,k].

3.3.2 Direct Reconstruction of the Land Carbon Cycle

Direct reconstruction of the land carbon cycle, which is here performed by resetting all restart files of the land carbon submodel to the target every Jan 1st, greatly enhances tracking performance of cVeg by simulation design. A sensitivity analysis for less frequent resetting can be found in the supplementary information [section C].

This direct resetting reconstructs cVeg much better than the resampled threshold in the extra-tropics. However, the physical climate biases during the course of a year even introduce cVeg biases stronger than the resampled threshold in the tropics [Fig. 2i]. Also, the biases in the air-land CO_2 flux are not improved [Fig. 2l], which indicates that this hard reset of restart files introduces a shock to the dynamical land model.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

On the other hand, correlations in cVeg and air-land CO_2 flux increased to above 0.5 everywhere expect in the tropics, where the ITCZ shift changes the climate regime [Fig. 3i,l].

Section 3.3 shows the direct land carbon cycle reconstruction yields stronger correlation improvements than ocean direct carbon cycle reconstruction, because the indirect reconstruction of the ocean was already quite good. Direct reconstruction reduces biases in land carbon cycle state variables, but the resulting air-land CO_2 flux biases becomes worse.

3.4 Reconstruction of the Global Carbon Cycle and Atmospheric CO_2

Tracking performance for prognostic atmospheric CO_2 integrates the air-sea and air-land CO_2 fluxes over time (Spring and Ilyina, 2020; Spring et al., 2020). As atmospheric CO_2 mixes fast across the globe, we first examine globally aggregated quantities driving globally averaged atmospheric CO_2 [Fig. 4].

3.4.1 Indirect Reconstruction of the Oceanic and Land Carbon Cycle

We first examine the indirect reconstruction represented by the green error bars in figure 5.

The indirect reconstruction has a negative bias in global pCO_2 in the annual mean, which is negative in boreal winter and positive in boreal summer, indicating that the reconstruction intensifies the seasonal cycle [Fig. D1]. This bias is mostly larger the resampled mean absolute error threshold, which resembles the temporal standard deviation [Fig. 5a]. The global oceanic CO_2 flux is low biased but within the resampled threshold [Fig. 5d].

On the other hand, the variations of the global oceanic carbon cycle measured by ACC are well reconstructed surpassing the resampled threshold [Fig. 5b,e].

The accuracy or distance measured by root mean squared error (RMSE) has strong seasonal errors especially in boreal winter up to 1.3 ppm. When biases are persistent, they can be reduced by a bias reduction procedure, which is often done when applying climate model output to a real-world application. After applying a simple mean bias reduction, RMSE is well below the resampled threshold at below 0.5 ppm [Fig. 5c,f].

The indirect reconstruction also leads to biases in the land carbon cycle. Vegetation carbon pools have a positive bias much larger than the resampled threshold. The bias of global air-land CO_2 flux only surpasses the resampled threshold in August, September and November, reducing the global seasonal cycle by 25% [Fig. 5j]

Global cVeg has a 0.5 correlation with the target, which is lower than the resampled 0.7 correlation. The 10-year running correlation also with a large range of ± 0.3. Global air-land CO_2 is well reconstructed surpassing the resampled threshold [Fig. 5h,k].

Without bias reduction, accuracy measured by RMSE is worse than the resampled cVeg threshold. After bias reduction, cVeg accuracy is still slightly worse than the threshold, but accuracy improved from 5 PgC to 1 PgC. Global air-land CO_2 flux accuracy is below the threshold until June and up to 0.3 PgC/month. But again when applying the bias mean reduction, tracking accuracy is always below the resampled threshold at 0.1 PgC/month [Fig. 5i,l].

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics
Discussions

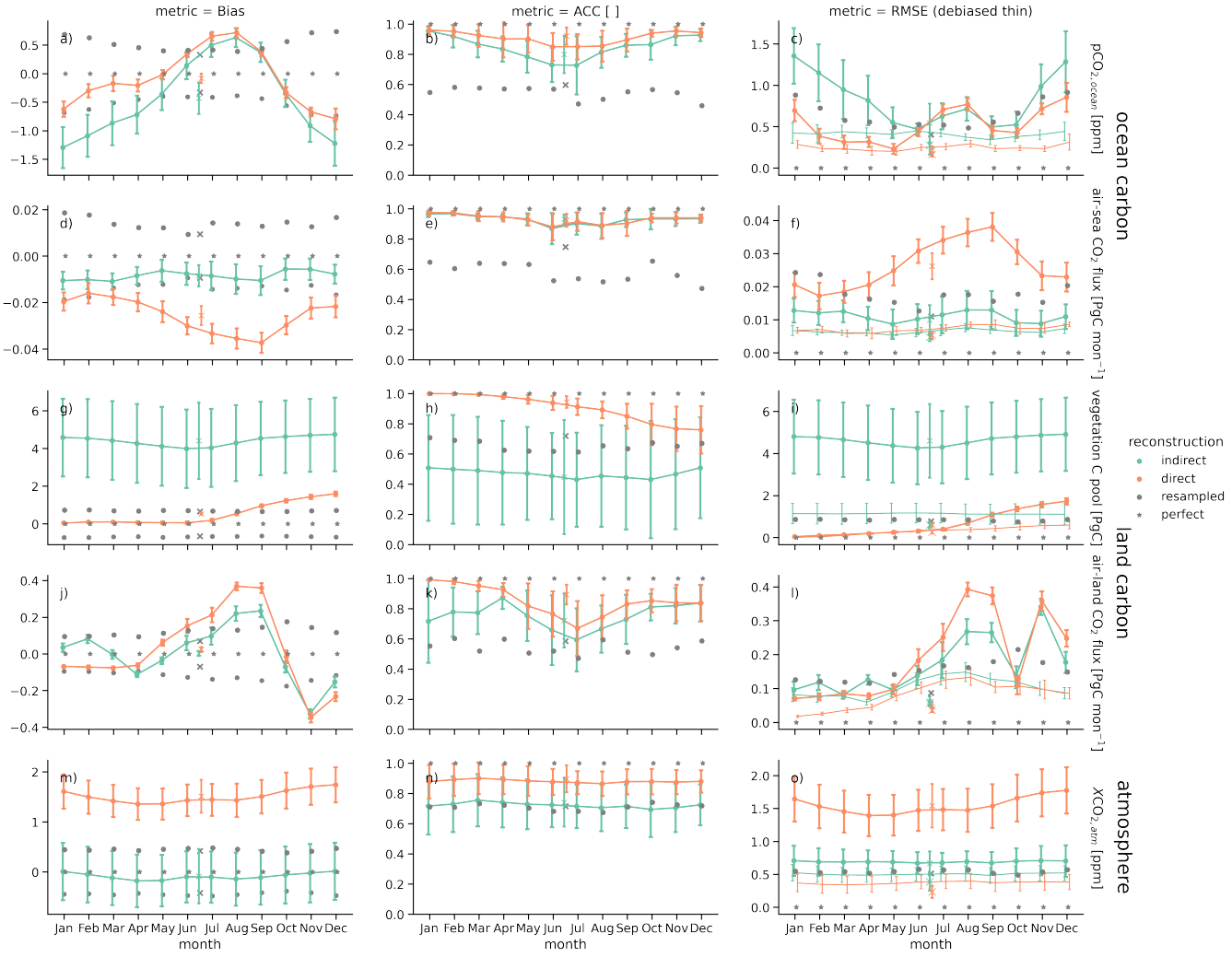

Figure 5. 10-year running mean reconstruction skill per month in bias (left), anomaly correlation coefficient (ACC, middle) and root-mean-square-error (RMSE, right) for global aggregation of carbon cycle variables: (a-c) surface oceanic partial pressure of CO_2, (d-f) air-sea CO_2 flux (negative values indicate carbon uptake by the ocean), (g-i) vegetation carbon pools, (j-l) air-land CO_2 flux (negative values indicate carbon uptake by land) and (m-o) mixing ratio of atmospheric CO_2. Colors show different reconstruction methods: indirect (green) and direct (orange). Gray stars indicate perfect skill. Gray dots mark 95th percentile for ACC and 5th percentile for the remaining distance-based metrics of random reconstruction skill block-bootstrapped from the target control simulation as an unskillful reference skill. Crosses show reconstruction skill of annual mean timeseries. Thin lines show monthly RMSE skill after a mean bias reduction.

Global atmospheric CO_2 has larger variations in reconstruction skill, depending on which 10-year chunk is used to calculate the metric. And the skill has a nearly constant level throughout the year. The mean bias is close to zero [Fig. 5m]. Correlation

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

with the target is above 0.7 and slightly above the resampled threshold [Fig. 5n]. Accuracy is at 0.7 ppm slightly above the threshold, but below the threshold at 0.5 ppm after mean bias reduction [Fig. 5o].

Understanding the tracking performance of the ocean and land carbon cycle, we can now evaluate the spatial distribution of globally averaged atmospheric CO_2. Reconstructing only the atmosphere warmed the globe and also increased atmospheric CO_2 globally [Figs. 1k,2m]. Reconstructing additionally also the ocean keeps the temperature stable, but introduces a less than 1 ppm low bias across the Southern Hemisphere, reflecting the higher uptake of the Southern Ocean carbon sink and the Southern Hemisphere land carbon sink [Fig. 2e,k,n]. The variations in atmospheric CO_2 are well reconstructed with correlation coefficients above 0.6 in the Southern Hemisphere, but across the Northern extra-tropics and the land regions with strong physics biases correlation is at 0.5 below the resampled threshold [Fig. 2m,n].

3.4.2 Direct Reconstruction of the Oceanic and Land Carbon Cycle

Now, we assess the potential improvements in the global carbon cycle due to direct reconstruction of the global carbon cycle variables shown in orange in figure 5.

The global ocean carbon cycle improves after direct DIC and alkalinity reconstruction. Monthly biases remain but are now within the resampled threshold [Fig. 5a]. Correlation improved from 0.8 to above 0.9 in surface pCO_2. Air-sea CO_2 did not improve, but only because of correlations above 0.9 for the indirect reconstruction were already very high. Correlation for boreal winter is above 0.95, indicating that initial conditions in winter are well reconstructable to initialize forecasts with for the oceanic carbon sink [Fig. 5b]. Direct reconstruction improves pCO_2 accuracy to 0.3 ppm [Fig. 5c].

All results for the direct reconstruction of the land carbon cycle must be understood in the context of the method chosen for the direct reconstruction: Because we reset the restart files in Jan 1st to the target, the metrics are near to perfect in January by design. However, then the biogeochemistry is not modified directly for twelve months and only follows the physical climate reconstruction indirectly, so biases triggered by physical biases unaligned with the reset land biogeochemistry pools quickly build up and may approach the metric of the indirect reconstruction. Likewise, there is no bias in global cVeg in January by design. The bias increases with the physical biases, until surpassing the resampled threshold in August increasing until the end of the year [Fig. 5g]. Global air-land CO_2 flux has even stronger bias than the indirect reconstruction [Fig. 5j]. Correlation in the global cVeg is near perfect in January by design and slowly decreases to 0.8 in December while still better than the resampled threshold [Fig. 5h]. Global air-land CO_2 flux maintains a 0.1 higher correlation than indirect reconstruction [Fig. 5k]. Direct reconstruction improves global cVeg accuracy. Accuracy is better than the resampled threshold after mean bias reduction. Direct reconstruction has worse CO_2 flux accuracy than the indirect, but after mean bias reduction the accuracy is slightly better [Fig. 5i,l].

The global CO_2 bias in the direction reconstruction increase to +1.8 ppm [Fig. 5m], but correlation increases from 0.7 to 0.9 [Fig. 5n]. The direct reconstruction has worse accuracy than the indirect due to the new bias, but after mean bias reduction the accuracy is slightly better at 0.4 ppm [Fig. 5o].

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

How does direct carbon cycle reconstruction affect tracking performance in prognostic atmospheric CO_2? Already the time series indicate, that there is a 1-2 ppm atmospheric CO_2 positive bias in the direct reconstruction [Fig. 4e]. This bias is very homogeneous over the oceans [Fig. 2o]. However, correlation strongly increased to 0.9 above the oceans and above 0.7 on land except for central Africa with its persistent biases, where the reconstruction is not better than the resampled threshold.

Section 3.4 shows that atmospheric CO_2 follows the reconstructed land and ocean carbon cycle integrating their respective fluxes over time. The direct carbon cycle reconstruction introduces a large bias in the atmospheric CO_2 distribution that the indirect reconstruction did not suffer from, also after mean bias reduction [Fig. B3]. Globally averaged atmospheric CO_2 after direct reconstruction had a better accuracy tracking performance after the mean bias reduction, showing how global aggregation can balance regional biases. The direct land and ocean carbon cycle reconstructions track target much better than the indirect reconstruction, when measuring correlation.

Hence, in large, this first part showed how direct carbon cycle reconstruction improves linear association between reconstruction and target (measured by ACC), but often increases biases degrading accuracy (measured by RMSE). Only after bias reduction, accuracy improves with respect to the indirect carbon cycle reconstruction.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

4 Impact of Reconstruction on Global Carbon Cycle Predictive skill

The second part of the paper assesses how predictive skill improves due to direct initialization of global carbon cycle variables. Specifically, we verify the RMSE between the five ensemble members initialized from the indirect and direct reconstructions across all initializations based on raw and lead-time dependent bias corrected timeseries [Figs. 4, 6]. We also evaluate predictive skill from a perfectly initialized ensemble, which are started from the perfect initial conditions taken from target simulation, whereas the ensembles from reconstructed initial conditions are biased with respect to the target [Fig. 5]. Acknowledging that our reconstruction simulation developed biases and that biases are commonly reduced by post-processing in predictability research, we also apply a simple lead-time dependent mean bias reduction to the initialized ensembles to show whether skill improvements go beyond what a simple post-processing could deliver. This initialized predictive skill is also compared with uninitialized ensembles randomly generated from the target simulation representing ensembles without common initialization and hence no memory. This uninitialized reference skill is used in predictability research community to assign whether the skill increase stems from initialization.

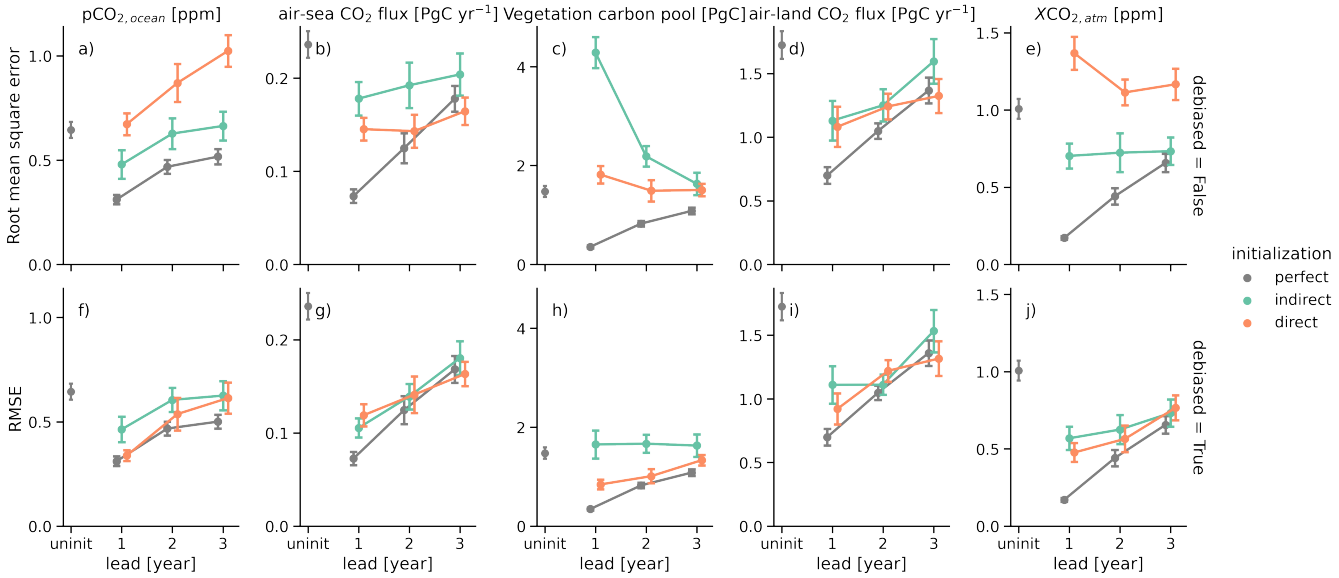

Figure 6. Predictive skill measured by root-mean-square-error (RMSE) between the initialized ensemble mean and the target as a function of lead year for different initialization setups: perfect indicating no reconstruction and hence perfect initial conditions to predict the target (gray), indirect (green) and direct (orange). Columns show global variables: for the ocean carbon cycle (a) oceanic surface pCO_2, (b) air-sea CO_2 flux; for the land carbon cycle (c) total land carbon pools, (d) air-land CO_2 flux and in the atmosphere (e) atmospheric CO_2 mixing ratio. (f-j) show RMSE-based predictive skill as (a-e) after mean bias reduction. Initialized ensembles are resampled with replacement (N=500) along the initialization dimension to account for initialization sampling uncertainty (see Spring and Ilyina, 2020), where errorbars show the resampled initialization skill uncertainty ($\pm 1\sigma$). Uninitialized ensembles, shown at lead 0, are resampled from the target control simulation and show the reference skill without initialization.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

4.1 Oceanic Carbon Cycle

The RMSE between the initialized ensembles and the target simulations in annual globally-averaged pCO_2 continuously increases from lead year one to lead year three as expected. While perfectly and indirectly initialized ensembles stay below the resampled uninitialized threshold for the first two lead years indicating that global pCO_2 is predictable due to initialization [Fig. 6a], the direct initialization has a larger error due to the offsets in global atmospheric CO_2, which pCO_2 tries to equilibrate to [Fig. 4e]. Therefore this persistent bias causes lead year three to be not predictable. A simple mean bias reduction resolves this issue making all three lead years predictable. Direct initialization only beats indirect initialization for lead year one with RMSE of 0.35±0.05 ppm versus 0.45±0.05 ppm [Fig. 6f].

Global air-sea CO_2 is predictable for three years in all initialization methods, which is one year longer than in Spring and Ilyina (2020), possibly because here we use more and more equally distributed initialization dates. Direct initialization is advantageous over the indirect initialization, because the initial lead offset is smaller (0.14±0.01 PgC/year vs 0.18±0.02 PgC/year) [Fig. 6b]. The simple mean bias reduction improves skill of the non-perfect initializations to identical magnitudes [Fig. 6g].

4.2 Land Carbon Cycle

Indirect initialization makes cVeg not predictable. The physical reconstruction biases drive to larger errors in lead year one than in later lead years, also to a lesser extent for the direct reconstruction where some biases are corrected. But both reconstructed initialized ensembles show decreasing distances towards the target, whereas increasing distances are expected for vanishing predictive skill as in the perfectly initialized ensembles [Fig. 6c]. Mean bias reduction eliminates the differences between direct and perfect reconstruction making both predictable unlike the indirect reconstruction [Fig. 6h].

Global air-land CO_2 flux is predictable for three years, again one year longer than found in Spring and Ilyina (2020). Both reconstructed initializations start with a higher error of 1.1±0.2 PgC/year in lead year one compared to perfect initialized 0.7±0.1 PgC/year [Fig. 6d]. Mean bias reduction brings non-perfect initializations within the error bars of the perfect initialization after lead year one [Fig. 6i].

4.3 Atmospheric CO_2

Perfect and indirect initialization atmospheric CO_2 predict the target for three years as found in Spring and Ilyina (2020). While the perfect initialization error grows continuously from zero, the indirect initialization error stays nearly constant at 0.7±0.1 ppm, but below the initialized. The direct initialization error suffers from the bias in the reconstruction simulation [Fig. 6e]. Mean bias reduction improves RMSE, making direct initialization better but still within the margins of the indirect initialization. After lead year one, indirect and direct initializations are similar to perfect-initialization predictive skill at 0.7 ppm [Fig. 6j].

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

These initialized predictive skill results show that indirectly initialized ensembles predict the target quite reasonable. Direct initialization suffers strong shocks in some variables, when reconstruction is started and stopped, but these shocks can be partly reduced by a mean bias reduction. The improvements of direct reconstruction over indirect reconstruction in the global carbon cycle predictive skill after bias reduction are not significant [Fig. 6f-j].

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

5 Summary and Conclusions

In this study, we assess how well the global carbon cycle is reconstructed in an ESM and how well a ground truth target simulation can be predicted by these initializations.

The main limitation of land carbon cycle reconstruction potential is the hard reset of restart files which is fundamentally different to the dynamical nudging applied for ocean and atmospheric physics. Our study represents a first attempt to quantify whether initial conditions reconstruction in land carbon cycle is indeed needed for addressing predictive skill of the global carbon sinks and atmospheric CO_2 concentration. For a real-world application, our direct land carbon reconstruction method cannot be used.

We find that reconstruction, which is an interference into the freely evolving model, leads to biases in physical climate. Because of its sensitivity to physical climate, the global carbon cycle is heavily biased itself by these physical biases. In ESMs, first the atmosphere, then the ocean and only then the carbon cycle is equilibrated and tuned for pre-industrial control conditions. Once reconstruction slightly modifies the mean state in the physical climate, the sensitive carbon cycle deviates from the near-equilibrium state. A previous study reported biases after reconstruction (Zhu and Kumar, 2018). Yet, to our knowledge, we present the first attempt at reconstructing in a perfect-model framework, where no biases due to climatology differences are expectable. Zhu and Kumar (2018) also mention that reconstruction ability likely depends on the model and application area, hence there seems to be no out-of-the-box solution for all ESMs.

We furthermore find that the commonly used indirect reconstruction of carbon cycle, in which only climate physics are reconstruct and the carbon cycle follows indirectly, tracks target reasonably well. A randomly resampled threshold corresponding to internal variability is surpassed across large parts of the globe. Only the areas with strong physical and consequently carbon cycle biases miss that benchmark occasionally. For the ocean carbon cycle, the initialization of the physical ocean fields is critical to reconstruct initial conditions.

Direct reconstruction of ocean and land carbon cycle improves bias, association and accuracy on a grid cell level, but aggregated on the global scale, direct reconstruction does not improve over the indirect reconstruction significantly. Also after a mean bias reduction, which is a common post-processing technique applied to model output for real-word use, accuracy measured in RMSE after direct reconstruction is only slightly better, often still overlapping with indirect reconstruction. Because the advantage of direct reconstruction can similarly be achieved by a simple mean bias reduction, we label these direct reconstruction improvements *trivial* with respect to the indirect method on the global scale.

When the success of atmospheric CO_2 reconstruction is evaluated, caution is needed. Reconstruction of the ocean and land carbon sink can easily introduce offsets from the target, because reconstruction violates conservation of mass by creating or erasing carbon. This can easily lead to offsets in the sinks which quickly accumulate in atmospheric CO_2. If CO_2 reconstruction is the focus, i.e. in reconstructing the transient climate from CO_2 emission, and offsets appear, adjustments of atmospheric CO_2 might be needed to correct for these offsets. However, we find that these offset biases are only of the order of 1-2 ppm in a perfect-model framework, which is small compared to the range of carbon feedbacks seen in atmospheric CO_2 in transient

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

simulations. Hence, these offsets due to the restart files are not in focus. Rather equilibrated land and ocean carbon sinks with reconstructed climate determine realistic reconstructed atmospheric CO_2.

In the second part, we find that predictive skill after indirect initialization is similarly good as after direct initialization. Reconstructed initialized predictive skill is close to perfectly initialized predictive skill after mean bias reduction, especially after lead year one.

Because the improved global predictive skill after direct reconstruction can similarly be achieved by a simple mean bias reduction and predictive skill after both reconstructions mostly overlaps, we label these direct reconstruction predictive skill improvements *trivial*, with respect to the indirect method on the global scale.

We conclude that the indirect carbon cycle reconstruction serves its purpose. Reducing the physical climate bias with its consequences for the carbon cycle holds more potential for improvements in initial conditions and predictive skill than direct carbon cycle initialization (Saito et al., 2011; Lee and Biasutti, 2014; Hua et al., 2019).

Our results add confidence to the current practice of indirect reconstruction in carbon cycle prediction systems (Ilyina et al., 2021).

Code and data availability. Forecast verification was performed with the python package CLIMPRED (Brady and Spring, under review) [https://github.com/pangeo-data/climpred/], which was co-developed with Riley X. Brady from University of Colorado, Boulder. Scripts and data to reproduce this analysis are archived in http://hdl.handle.net/21.11116/0000-0007-A697-3.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Appendix A: Metrics

A1 ACC

The anomaly correlation coefficient (ACC) assesses the synchronous evolution over time of the forecast, here reconstruction $x(t)$ and the reference, here target $\hat{x}(t)$, (Jolliffe and Stephenson, 2011) and is defined as:

$$ACC(x(t),\hat{x}(t)) = \frac{cov(x(t),\hat{x}(t))}{\sqrt{var(x(t))\cdot var(\hat{x}(t))}} = \frac{\frac{1}{T}\sum_{t=1}^{T}(x(t)-\overline{x(t)})(\hat{x}(t)-\overline{\hat{x}(t)})}{\sqrt{\frac{\sum_{t=1}^{T}(x(t)-\overline{x(t)})^2}{T}}\cdot\sqrt{\frac{\sum_{t=1}^{T}(\hat{x}(t)-\overline{\hat{x}(t)})^2}{T}}}. \tag{A1}$$

A2 RMSE

In the initial conditions reconstruction part, the root-mean-square-error (RMSE) measures the second-order distance between forecast $x(t)$, here reconstruction $x(t)$ and the reference, here target $\hat{x}(t)$, (Jolliffe and Stephenson, 2011) and is defined as:

$$RMSE(x(t),\hat{x}(t)) = \sqrt{\frac{\sum_{T=1}^{T}(x(t)-\hat{x}(t))^2}{T}}. \tag{A2}$$

As a predictability metric, the root-mean-square-error (RMSE) measures the second-order distance between forecast $x(t)$ and the target $\hat{x}(t)$ over lead time t (Jolliffe and Stephenson, 2011). RMSE is calculated over all initialisations N and every member M is used as a forecast and verified against the target. RMSE is defined as:

$$RMSE(x(t),\hat{x}(t)) = \sqrt{\frac{\sum_{i,j=1}^{N,M}(x_{i,j}(t)-\hat{x}_j(t))^2}{NM}}. \tag{A3}$$

A3 Bias

We set the target as the ground truth. Therefore any deviation from the reconstructions $x(t)$ to the target $\hat{x}(t)$ is seen as a bias, analogous to the bias between a model simulation (reconstruction) and observations (ground truth).

$$bias(t) = x(t) - \hat{x}(t) \tag{A4}$$

A4 Removing the Bias

After removing the mean bias from reconstruction $\overline{x(t)}$ and target $\overline{\hat{x}(t)}$, the RMSE is also calculated as debiased RMSE.

$$RMSE_{\text{debiased}}(t) = RMSE\left((x(t)-\overline{x(t)},\hat{x}(t)-\overline{\hat{x}(t)}\right) \tag{A5}$$

A5 Running Metric

We calculate the mean tracking performance (mtp) over time for all metrics as a running mean over $s = 10$ years. This reflects that reconstructions are supposed to reconstruct the given climate states within months to a couple of years and the metric should not be prone to long-term trends that are not captured by the reconstruction. We ignore the first $c = 10$ years (out of

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

$t_{max} = 48$ years) of reconstruction, where the model experiences an initial shock after adjusting to the new reconstructed climate (Kröger et al., 2017).

$$tpm(metric) = \frac{1}{t_{max} - s - c} \sum_{t=c}^{t_{max}-s} metric(x_{(t=t..t+s)}, \hat{x}_{(t=t..t+s)}) \tag{A6}$$

A6 Resampled Threshold

To get an estimate of random tracking performance due to internal variability, i.e. how well one 10-year chunk tracks just another random 10-year chunk, we randomly resample 10-year chunks from the target simulation and apply the same tracking metrics. As a baseline skill from this random resampling in the figures, we take the 95% threshold for ACC and the 95% for the remaining distance-based metrics to ensure that the tracking performance from a reconstruction simulation is only worse compared to one out of 20 randomly resampled 10-year chunks.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics Discussions
Open Access
EGU

Appendix B: Reconstruction RMSE Maps

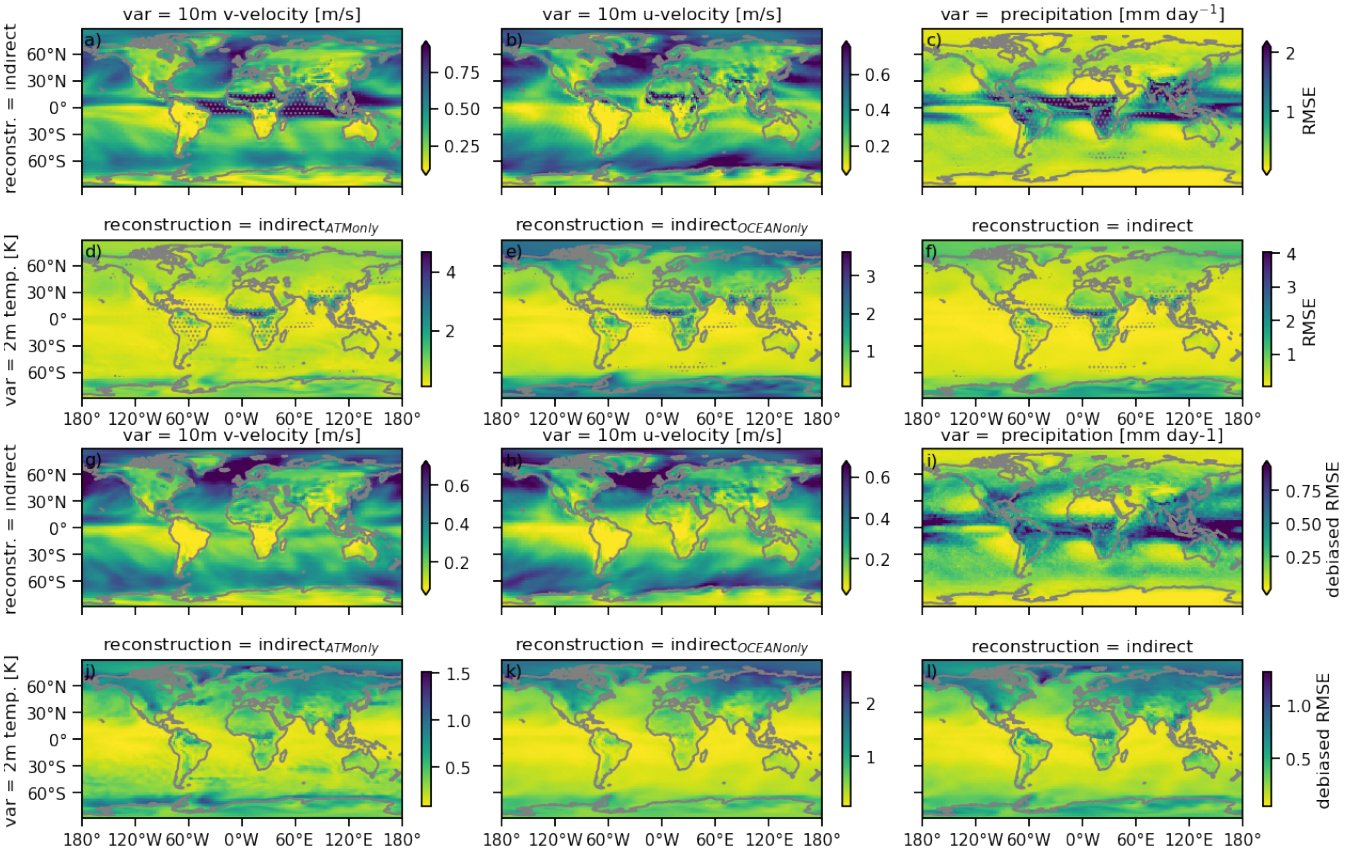

Figure B1. As Fig. 1 but for Root-mean-square-error (RMSE) (a-f) and for RMSE after bias reduction (g-l).

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics
Discussions
Open Access
EGU

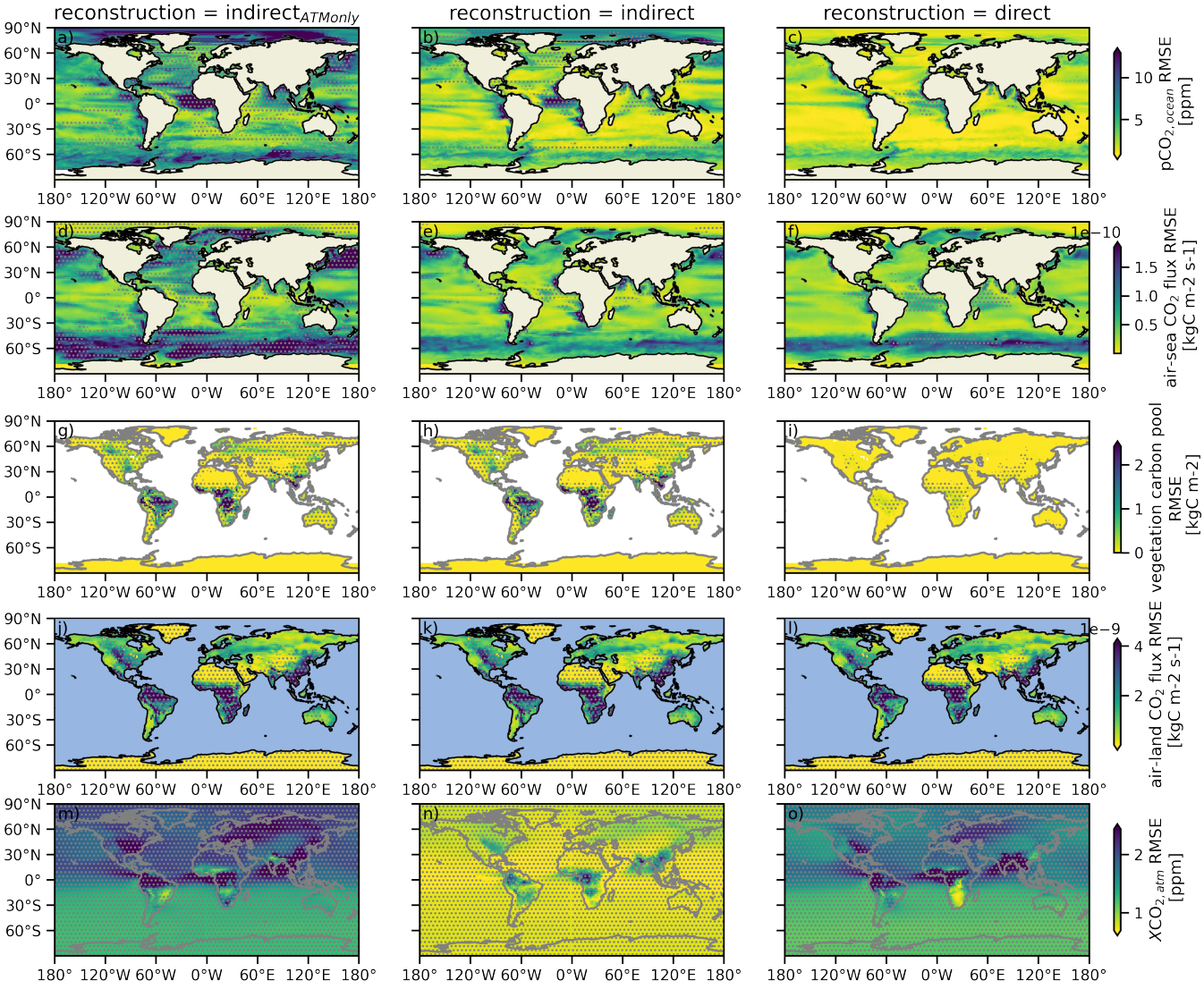

Figure B2. As Fig. 2 but for the Root-mean-square-error (RMSE). Gray stippling shows where the RMSE is worse than the 5th percentile RMSE threshold from random target block resampling, i.e. the reconstruction is not statistically significantly compared to internal variability.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics
Discussions
Open Access
EGU

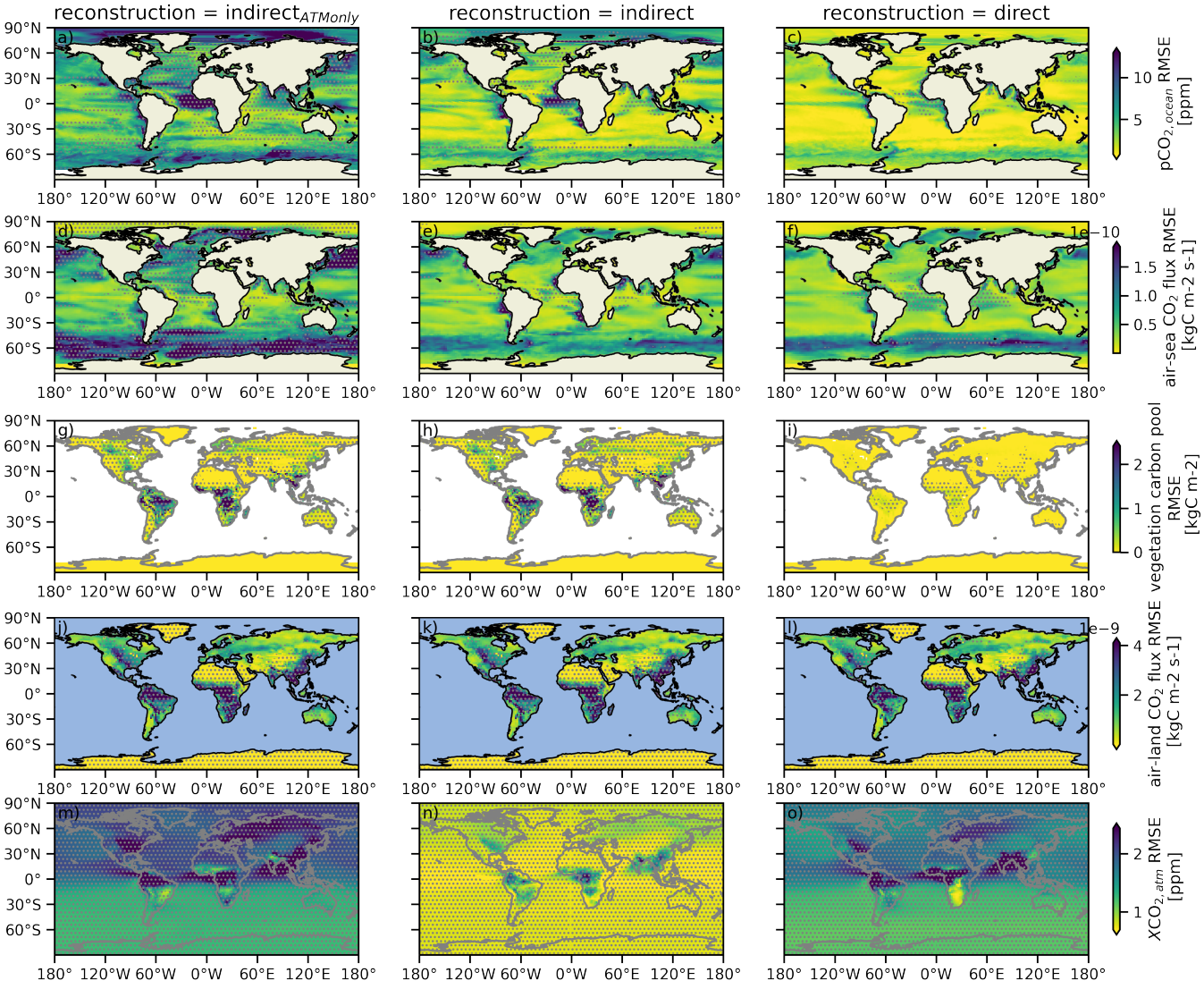

Figure B3. As Fig. B2 but for Root-mean-square-error (RMSE) after bias reduction.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Appendix C: Sensitivity Analysis for Different Reconstruction Timestep in ...

C1 ... on Land Carbon Cycle

We perform sensitivity reconstructions of the land restart file resetting to understand how sensitive this reconstruction method to the frequency of resetting. We performed additional simulations resetting the land model on Jan. 1st every second or every fifth year [orange triangles in fig. C1].

Global cVeg starts by definition with perfect skill in Jan after a reset. When resetting only every second year, the mean January tracking performance is already decreased, and decreases further. The negative correlations for five-year resetting shows the shock to the system if not immediately balanced by further resetting in the every (second) year case.

The global air-land CO_2 flux correlation degrades for less frequent resetting towards the indirect performance, but bias and accuracy improve.

Global atmospheric CO_2 aggregates these results and is also sensitive to biases developing in both sinks. Here, less frequent resetting of the land carbon cycle reduces the bias and therefore accuracy.

The tracking accuracy is of similar magnitude after mean bias reduction.

C2 ... on Ocean Carbon Cycle

We perform the same kind of restart file resetting reconstruction to the ocean model [blue line in fig. C1]. The motivation here is to see whether a resetting of the ocean carbon cycle also yields perfect accuracy (RMSE) skill for January. But the ocean carbon cycle is sensitive to the physical climate and hence the direct ocean carbon cycle resetting accuracy degrades compared to the indirect tracking bias and accuracy, only correlation increases [Fig. C1a-f]. Contrary to resetting restart files in the land model, initial conditions accuracy measured by RMSE does not approach perfect skill of 0, because the physical climate did not experience this hard reset but is nudged dynamically.

In general, this hard reconstruction also seems to work for the ocean carbon cycle, because the tracking performances are not very different from the indirect method [Fig. C1a-f].

The tracking accuracy is of similar magnitude after mean bias reduction.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics
Discussions
Open Access
EGU

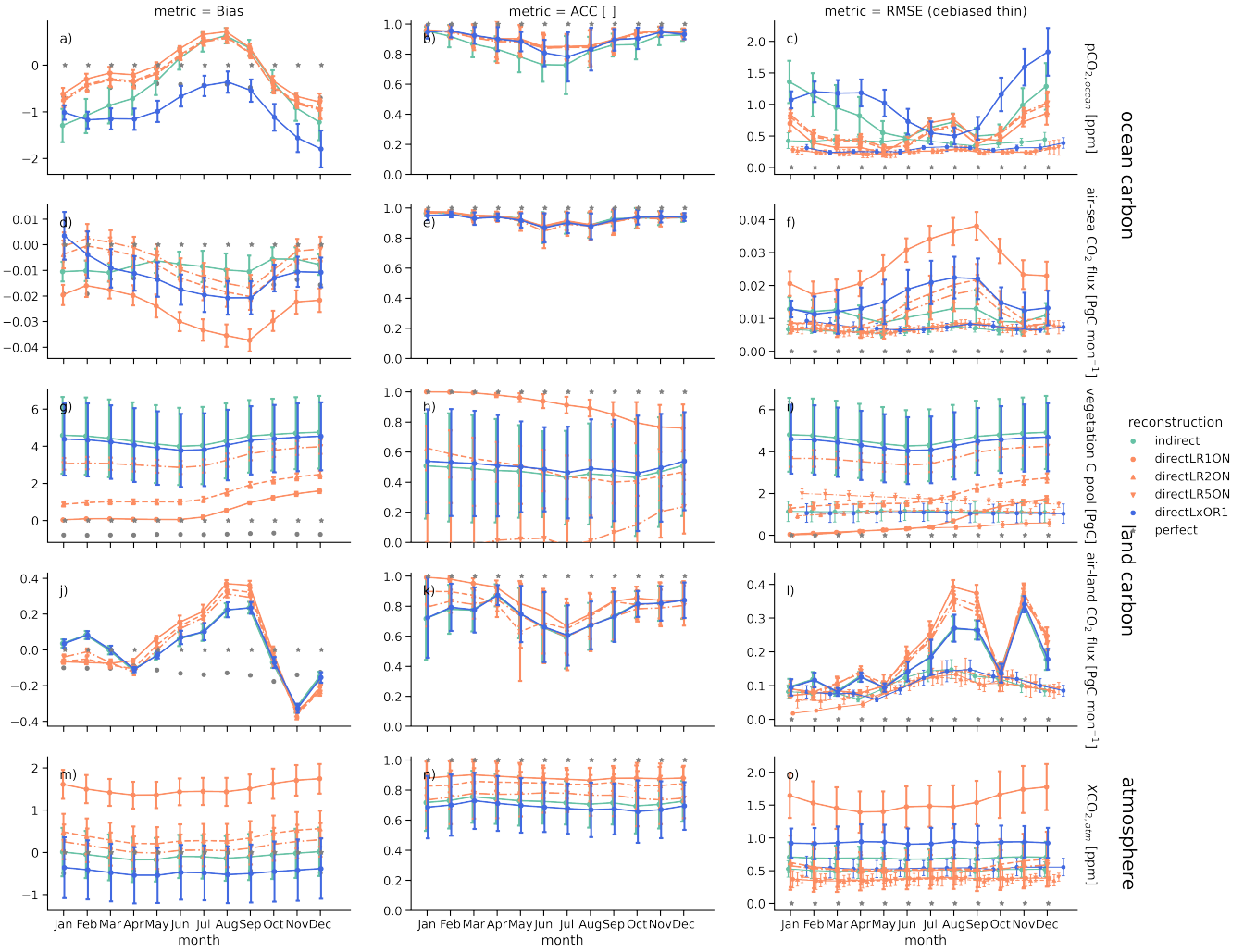

Figure C1. As Fig. 5 but for sensitivity simulations of the restart file resetting reconstruction. In all simulations the physical climate is nudged as in indirect [Tab. 1]. directLR1ON describes land resetting every year and ocean nudging and is the indirect simulation. directLR2ON describes land resetting every second year and ocean nudging. directLR5ON describes land resetting every fifth year and ocean nudging. directLxOR1 describes no land reconstruction and ocean setting every year.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Appendix D: Seasonality

In reference for figure 5 to better understand reconstruction skill in context of target seasonality:

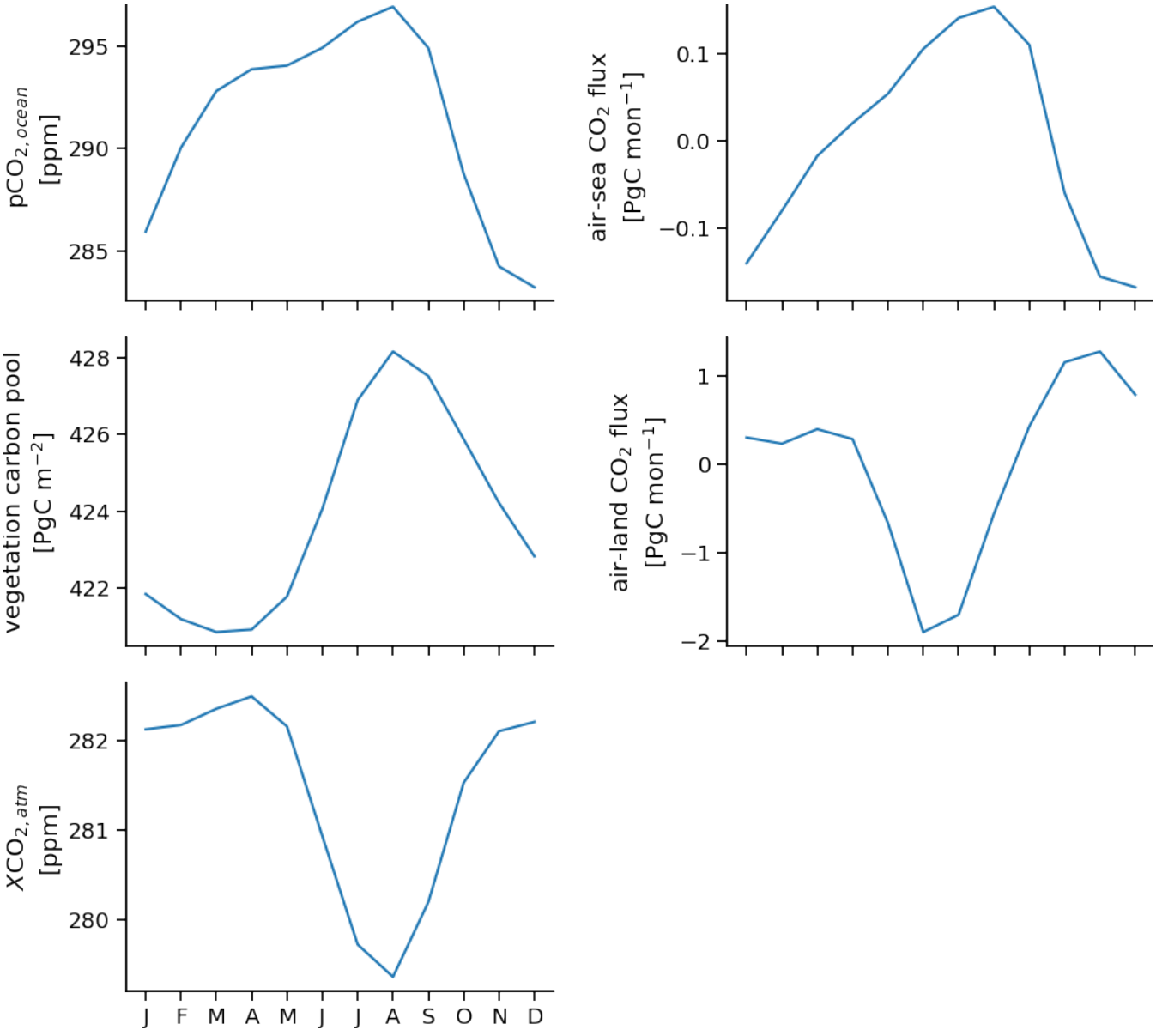

Figure D1. Seasonality of the target simulation for global aggregated carbon cycle variables.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Earth System Dynamics Discussions

Author contributions. A.S. and T.I. conceived the study, A.S. performed the simulations and analysis, created the figures and drafted the manuscript. T.I., I.D., H.L. and V.B. contributed in manuscript editing and provided feedback.

Competing interests. The authors declare no competing interests.

Acknowledgements. We acknowledge funding from European Union's Horizon 2020 research and innovation programme under grant agreement No 821003 "Climate-Carbon Interactions in the Current Century (4C)" , No 820989 "COMFORT" and No 641816 "CRESCENDO". Simulations were performed at the German Climate Computing Center (DKRZ). We thank Jürgen Bader for internal review.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

References

Brady, R. X. and Spring, A.: Climpred: Verification of Weather and Climate Forecasts, Journal of Open Source Software, submitted October 2020, https://joss.theoj.org/papers/246d440e3fcb19025a3b0e56e1af54ef, under review.

Efron, B. and Tibshirani, R. J.: An Introduction to the Bootstrap, Chapman and Hall/CRC, New York, 1st edition edn., https://cds.cern.ch/record/526679/files/0412042312_TOC.pdf, 1993.

(EU), E. U.: European Union (EU) Horizon 2020 Climate-Carbon Interactions in the Coming Century (4C) 821003 Doi: Mydoi, xx.xx.2020.

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/gbpwpt, 2016.

Fransner, F., Counillon, F., Bethke, I., Tjiputra, J., Samuelsen, A., Nummelin, A., and Olsen, A.: Ocean Biogeochemical Predictions—Initialization and Limits of Predictability, Frontiers in Marine Science, 7, https://doi.org/10/gg22rr, 2020.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Quéré, C. L., 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., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, 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., Peregon, 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 System Science Data, 11, 1783–1838, https://doi.org/10/ggd7p4, 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 Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, 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., Tanhua, T., 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 System Science Data, 12, 3269–3340, https://doi.org/10/ghn75s, 2020.

Griffies, S. M. and Bryan, K.: A Predictability Study of Simulated North Atlantic Multidecadal Variability, Climate Dynamics, 13, 459–487, https://doi.org/10/ch4kc4, 1997.

Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., Chassignet, E. P., Curchitser, E., Deshayes, J., Drange, H., Fox-Kemper, B., Gleckler, P. J., Gregory, J. M., Haak, H., Hallberg, R. W., Heimbach, P., Hewitt, H. T., Holland, D. M., Ilyina, T., Jungclaus, J. H., Komuro, Y., Krasting, J. P., Large, W. G., Marsland, S. J., Masina, S., McDougall, T. J., Nurser, A. J. G., Orr, J. C., Pirani, A., Qiao, F., Stouffer, R. J., Taylor, K. E., Treguier, A. M., Tsujino, H., Uotila, P., Valdivieso, M., Wang, Q., Winton, M.,

and Yeager, S. G.: OMIP Contribution to CMIP6: Experimental and Diagnostic Protocol for the Physical Component of the Ocean Model Intercomparison Project, Geoscientific Model Development, 9, 3231–3296, https://doi.org/10/f85pjx, 2016.

Haney, R. L.: A Numerical Study of the Response of an Idealized Ocean to Large-Scale Surface Heat and Momentum Flux, Journal of Physical Oceanography, 4, 145–167, https://doi.org/10/ch3kf3, 1974.

Hua, W., Zhou, L., Nicholson, S. E., Chen, H., and Qin, M.: Assessing Reanalysis Data for Understanding Rainfall Climatology and Variability over Central Equatorial Africa, Climate Dynamics, 53, 651–669, https://doi.org/10/ghthbt, 2019.

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, Journal of Advances in Modeling Earth Systems, 5, 287–315, https://doi.org/10/f5dz39, 2013.

Ilyina, T., Li, H., Spring, A., Müller, W. A., Bopp, L., Chikamoto, M. O., Danabasoglu, G., Dobrynin, M., Dunne, J., Fransner, 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_2 Growth in a Multi-Model Framework, Geophysical Research Letters, p. e2020GL090695, https://doi.org/10/ghsn7h, 2021.

Jeuken, A. B. M., Siegmund, P. C., Heijboer, L. C., Feichter, J., and Bengtsson, L.: On the Potential of Assimilating Meteorological Analyses in a Global Climate Model for the Purpose of Model Validation, Journal of Geophysical Research: Atmospheres, 101, 16 939–16 950, https://doi.org/10/d64x9q, 1996.

Jolliffe, I. T. and Stephenson, D. B.: Forecast Verification: A Practitioner's Guide in Atmospheric Science, John Wiley & Sons, Ltd, Chichester, UK, https://doi.org/10.1002/9781119960003, 2011.

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/f84585, 2016.

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J. S.: 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, 422–446, https://doi.org/10/f5d3x7, 2013.

Keenlyside, N. S., Latif, M., Jungclaus, J., Kornblueh, L., and Roeckner, E.: Advancing Decadal-Scale Climate Prediction in the North Atlantic Sector, Nature, 453, 84–88, https://doi.org/10/ffpmhd, 2008.

Kröger, J., Pohlmann, H., Sienz, F., Marotzke, J., Baehr, J., Köhl, A., Modali, K., Polkova, I., Stammer, D., Vamborg, F. S. E., and Müller, W. A.: Full-Field Initialized Decadal Predictions with the MPI Earth System Model: An Initial Shock in the North Atlantic, Climate Dynamics, https://doi.org/10/gdsnf8, 2017.

Lee, D. E. and Biasutti, M.: Climatology and Variability of Precipitation in the Twentieth-Century Reanalysis, Journal of Climate, 27, 5964–5981, https://doi.org/10/f6b5jn, 2014.

Lefèvre, N., Caniaux, G., Janicot, S., and Gueye, A. K.: Increased CO_2 Outgassing in February-May 2010 in the Tropical Atlantic Following the 2009 Pacific El Niño, Journal of Geophysical Research: Oceans, 118, 1645–1657, https://doi.org/10/ghrngp, 2013.

Li, H., Ilyina, T., Müller, W. A., and Sienz, F.: Decadal Predictions of the North Atlantic CO_2 Uptake, Nature Communications, 7, 11 076, https://doi.org/10/f8wkrs, 2016.

Li, H., Ilyina, T., Müller, W. A., and Landschützer, P.: Predicting the Variable Ocean Carbon Sink, Science Advances, 5, eaav6471, https://doi.org/10/gf4fxm, 2019.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Lin, S.-J. and Rood, R. B.: Multidimensional Flux-Form Semi-Lagrangian Transport Schemes, Monthly Weather Review, 124, 2046–2070, https://doi.org/10/fsh52z, 1996.

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, Environmental Research Letters, 14, 124 074, https://doi.org/10/ggfv2w, 2019a.

Lovenduski, N. S., Yeager, S. G., Lindsay, K., and Long, M. C.: Predicting Near-Term Variability in Ocean Carbon Uptake, Earth System Dynamics, 10, 45–57, https://doi.org/10/gfvxkc, 2019b.

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., Jimenéz-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, 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_2, Journal of Advances in Modeling Earth Systems, 11, 998–1038, https://doi.org/10/gftpps, 2019.

Merryfield, W. J., Baehr, J., Batté, L., Becker, E. J., Butler, A. H., Coelho, C. A. S., Danabasoglu, G., Dirmeyer, P. A., Doblas-Reyes, F. J., Domeisen, D. I. V., Ferranti, L., Ilynia, T., Kumar, A., Müller, W. A., Rixen, M., Robertson, A. W., Smith, D. M., Takaya, Y., Tuma, M., Vitart, F., White, C. J., Alvarez, M. S., Ardilouze, C., Attard, H., Baggett, C., Balmaseda, M. A., Beraki, A. F., Bhattacharjee, P. S., Bilbao, R., de Andrade, F. M., DeFlorio, M. J., Díaz, L. B., Ehsan, M. A., Fragkoulidis, G., Grainger, S., Green, B. W., Hell, M. C., Infanti, J. M., Isensee, K., Kataoka, T., Kirtman, B. P., Klingaman, N. P., Lee, J.-Y., Mayer, K., McKay, R., Mecking, J. V., Miller, D. E., Neddermann, N., Justin Ng, C. H., Ossó, A., Pankatz, K., Peatman, S., Pegion, K., Perlwitz, J., Recalde-Coronel, G. C., Reintges, A., Renkl, C., Solaraju-Murali, B., Spring, A., Stan, C., Sun, Y. Q., Tozer, C. R., Vigaud, N., Woolnough, S., and Yeager, S.: Current and Emerging Developments in Subseasonal to Decadal Prediction, Bulletin of the American Meteorological Society, 101, 869–896, https://doi.org/10/ggvcqv, 2020.

Milinski, S., Bader, J., Haak, H., Siongco, A. C., and Jungclaus, J. H.: High Atmospheric Horizontal Resolution Eliminates the Wind-Driven Coastal Warm Bias in the Southeastern Tropical Atlantic, Geophysical Research Letters, 43, 10,455–10,462, https://doi.org/10/f89qf7, 2016.

Orr, J. C., Najjar, R. G., Aumont, O., Bopp, L., Bullister, J. L., Danabasoglu, G., Doney, S. C., Dunne, J. P., Dutay, J.-C., Graven, H., Griffies, S. M., John, J. G., Joos, F., Levin, I., Lindsay, K., Matear, R. J., McKinley, G. A., Mouchet, A., Oschlies, A., Romanou, A., Schlitzer, R., Tagliabue, A., Tanhua, T., and Yool, A.: Biogeochemical Protocols and Diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP), Geoscientific Model Development, 10, 2169–2199, https://doi.org/10/gcc5zk, 2017.

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, Journal of Advances in Modeling Earth Systems, 9, 438–464, https://doi.org/10/f94tnr, 2017.

Pohlmann, H., Jungclaus, J. H., Köhl, A., Stammer, D., and Marotzke, J.: Initializing Decadal Climate Predictions with the GECCO Oceanic Synthesis: Effects on the North Atlantic, Journal of Climate, 22, 3926–3938, https://doi.org/10/cdvhcr, 2009.

Pohlmann, H., Müller, W. A., Bittner, M., Hettrich, S., Modali, K., Pankatz, K., and Marotzke, J.: Realistic Quasi-Biennial Oscillation Variability in Historical and Decadal Hindcast Simulations Using CMIP6 Forcing, Geophysical Research Letters, 46, 14 118–14 125, https://doi.org/10/ggfrh8, 2019.

https://doi.org/10.5194/esd-2021-4
Preprint. Discussion started: 18 February 2021
© Author(s) 2021. CC BY 4.0 License.

Rast, S., Brokopf, R., Esch, M., Gayler, V., Kirchner, I., Kornblueh, L., Rhodin, A., and Schulzweida, U.: User Manual for ECHAM6, Tech. rep., Hamburg, https://icdc.cen.uni-hamburg.de/fileadmin/user_upload/icdc_Dokumente/ECHAM/echam6_userguide.pdf, 2012.

Saito, M., Ito, A., and Maksyutov, S.: Evaluation of Biases in JRA-25/JCDAS Precipitation and Their Impact on the Global Terrestrial Carbon Balance, Journal of Climate, 24, 4109–4125, https://doi.org/10/fk6wrg, 2011.

Schneck, R., Reick, C. H., and Raddatz, T.: Land Contribution to Natural CO_2 Variability on Time Scales of Centuries, Journal of Advances in Modeling Earth Systems, 5, 354–365, https://doi.org/10/f5dzt5, 2013.

Séférian, R., Berthet, S., and Chevallier, M.: Assessing the Decadal Predictability of Land and Ocean Carbon Uptake, Geophysical Research Letters, https://doi.org/10/gdb424, 2018.

Servonnat, J., Mignot, J., Guilyardi, E., Swingedouw, D., Séférian, R., and Labetoulle, S.: Reconstructing the Subsurface Ocean Decadal Variability Using Surface Nudging in a Perfect Model Framework, Climate Dynamics, 44, 315–338, https://doi.org/10/f6v7kq, 2015.

Spring, A. and Ilyina, T.: Predictability Horizons in the Global Carbon Cycle Inferred From a Perfect-Model Framework, Geophysical Research Letters, 47, e2019GL085 311, https://doi.org/10/ggtbv2, 2020.

Spring, A., Ilyina, T., and Marotzke, J.: Inherent Uncertainty Disguises Attribution of Reduced Atmospheric CO_2 Growth to CO_2 Emission Reductions for up to a Decade, Environmental Research Letters, 15, 114 058, https://doi.org/10/ghmjp8, 2020.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: Atmospheric Component of the MPI-M Earth System Model: ECHAM6, Journal of Advances in Modeling Earth Systems, 5, 146–172, https://doi.org/10/f5dz5h, 2013.

van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J., Flanner, M. G., Zampieri, M., Materia, S., Law, R. M., and Sheffield, J.: LS3MIP (v1.0) Contribution to CMIP6: The Land Surface, Snow and Soil Moisture Model Intercomparison Project – Aims, Setup and Expected Outcome, Geoscientific Model Development, 9, 2809–2832, https://doi.org/10/gcc538, 2016.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, vol. 91 of *International Geophysics Series*, Academic Press, Amsterdam; Boston, 2nd ed edn., http://eamcweb4.usfs.msu.edu/study/2016-GCM/Statistical_Methods_Wilks.pdf, 2006.

Zhu, J. and Kumar, A.: Influence of Surface Nudging on Climatological Mean and ENSO Feedbacks in a Coupled Model, Climate Dynamics, 50, 571–586, https://doi.org/10/gcwrz4, 2018.