ARTICLE   OPEN

The impact of seasonality on the annual air-sea carbon flux and its interannual variability

Paridhi Rustogi [1,2,5 ✉], Peter Landschützer [3,4], Sebastian Brune[2] and Johanna Baehr[2]

Interannual variability of the ocean carbon sink is often assessed using annual air–sea carbon fluxes, but the drivers of the variability may instead arise from seasonal processes that are neglected in the annual average. The seasonal cycle largely modulates air–sea carbon exchange, hence understanding seasonal mechanisms and their link to interannual variability is necessary to determine long-term changes in the ocean carbon sink. We contrast carbon fluxes from an Earth System Model large ensemble and an observation-based ensemble to assess the representation of annual and seasonal carbon fluxes in two distinct ocean regions—the North Atlantic basin and the Southern Ocean and investigate if seasonal variability can help diagnose interannual variability. Both ensembles show strong agreement in their annual mean fluxes. However, discrepancies between the two ensembles are one to two times greater for the seasonal fluxes than the annual fluxes in the North Atlantic basin and three to four times greater in the Southern Ocean. These seasonal discrepancies compensate in the annual mean, obscuring significant seasonal mismatches between the ensembles, particularly in the Southern Ocean. A solid understanding of seasonal variability can be leveraged to diagnose interannual variability of carbon fluxes where necessary observational constraints have been built, for example, in the North Atlantic basin, where boreal winter and spring drive the interannual variability. However, in a data-sparse region like the Southern Ocean, both ensembles disagree substantially in their representations of seasonal carbon fluxes and variability, and currently, seasonal variability is of limited use in diagnosing the interannual variability of carbon fluxes in the Southern Ocean.

INTRODUCTION

Synthesis reports like the Global Carbon Budget characterise changes in the ocean carbon sink by analysing the annual air–sea carbon flux[1]. However, year-to-year changes or the interannual variability of carbon fluxes may be driven by seasonal processes that are neglected in the annual average[2]. For most parts of the global ocean, the seasonal cycle is the most prominent mode of variability in air–sea carbon exchange[3]. As a result, seasonal mechanisms and their link to interannual variability have received considerable attention[4–7]. Despite this, inferring the interannual variability of the ocean carbon sink remains a major scientific challenge[8]. Future emissions will lead to seasonally divergent signals[9,10], hence understanding long-term temporal changes of the ocean carbon sink hinges on resolving short-term seasonal signals that drive their variability[4,11].

Seasonal changes in the carbon sink are largely driven by internal processes, such as a shift in temperature-dependent solubility or biological processes that reduce the mean partial pressure of carbon dioxide (pCO_2) in the ocean and increase the amplitude of the annual cycle[12]. In contrast, interannual changes in the carbon sink are driven by external forcing, influenced by climate phenomena like El Niño Southern Oscillation, or in response to increasing atmospheric carbon concentrations[13]. On decadal timescales, feedbacks rein in the interannual variability so that its magnitude is much smaller than that of seasonal variability. Assessing the weaker interannual variability signal requires long-term measurements at the same location and season, an observational constraint seldom satisfied in the global ocean. Consequently, a combination of models and observation-based methods have been employed to understand the

mechanisms that connect changes in the seasonal cycle to interannual variability and to acquire reliable state estimates of the oceanic carbon inventory.

Earth System Models (ESM) are internally consistent tools well suited to represent global annual carbon fluxes and their long-term trends. However, there remains a significant mismatch between ESMs and observation-based data on seasonal timescales and global trends[14,15]. Further, considerable differences exist even among models of the same generation and complexity in regional carbon flux estimates, with disagreements on the phasing of the seasonal cycle in the high latitudes[11,16].

However, both models and observation-based products face limitations[14,17–19]. On the one hand, global ocean biogeochemical models that are forced by the atmosphere have been historically used to close the Global Carbon Budget[14,20], but limited observational data makes it challenging to test and potentially reduce model biases and uncertainty when ocean and atmosphere models are coupled. On the other hand, observation-based carbon flux reconstructions rely on sparse measurements of pCO_2 at the air–sea interface and some form of interpolation technique to statistically fill spatial and temporal gaps[21–26]. As a result, despite being measurement-driven and sourced from the same observational database, there is a lack of consistency across observation-based products[14,18,27], and measurement uncertainties may propagate to the individual products.

In this study, we investigate the seasonality of carbon fluxes in context with their respective annual fluxes over three decades (1990–2020) for which observational constraints are available. We contrast two tools used to assess carbon fluxes—an ESM large ensemble—the Max Planck Institute Grand Ensemble (MPI-GE,

[1]School of Integrated Climate System Sciences, University of Hamburg, Hamburg, Germany. [2]Institute of Oceanography, Center for Earth System Research and Sustainability (CEN), Universität Hamburg, Hamburg, Germany. [3]Max Planck Institute for Meteorology, Hamburg, Germany. [4]Flanders Marine Institute (VLIZ), Ostend, Belgium. [5]Present address: Department of Geosciences, Princeton University, New Jersey, USA. ✉email: prustogi@princeton.edu

$n = 100$) and an observation-based CO_2 flux ensemble (using products from SOCOM ($n = 6$). Details in Data and Methods). The mean of the model ensemble represents the forced signal across members, and the spread between the individual members represents its internal variations. In contrast to the model, the differences between the observation-based products do not stem from differences in internal (or climatic) variations but rather from measurement uncertainties and random errors related to the extrapolation method. All observation-based products rely on significant averaging or interpolation using satellite-derived observations of sea surface temperature (SST), chlorophyll (Chl a) or climatologies of mixed layer depth (MLD) to make basin-scale estimates. Additionally, differences in the choice of the kinetic gas transfer formulation across the air–sea interface and the wind product contribute to the spread, though previous studies have shown that harmonising the pCO2 product ensemble using a common gas transfer formulation can significantly decrease the spread[28]. Here, we follow the approach of the Global Carbon Budget[1] and consider the disharmonised gas transfer as part of the ensemble spread, and hence the overall uncertainty. Regardless of the phasing of internal (or climatic) variations, we treat both ensembles equally, not differentiating between them and how they were built. Contrasting the model and observation-based ensembles, we investigate how seasonal and annual carbon fluxes compare and ascertain if a link exists between seasonal variability that helps diagnose the interannual variability of air–sea carbon fluxes in either ensemble.

We focus on the North Atlantic basin and the Southern Ocean—two distinct ocean regions that show considerable carbon flux variability on seasonal through interannual timescales but differ in the level of convergence between model and observation-based data[18,29–31]. The first region, the North Atlantic basin (10–90°W, 15–60°N), has the most intense carbon sink per unit area[32]. It has historically been one of the most regularly observed ocean regions with year-round coverage and nearly equal distributions of seasonal measurements since the Ship of Opportunity programme started in the 1990s[32,33]. The observed variability in carbon uptake on interannual timescales and the mechanisms driving this variability can be explained across a broad range of ESMs so that there is overall good agreement between models and observation-based products on seasonal timescales[34].

Competing mechanisms direct the phasing of the seasonal cycle of the carbon flux in the North Atlantic. The flux signal can be deconstructed into thermal and nonthermal components, allowing for a more mechanistic description of key driving processes. Temperature variations have a known effect (approximately a 4% increase per unit temperature increase) on the seasonal cycle of the carbon flux[35]. The onset of phytoplankton blooms in boreal spring as part of the ocean's biological pump and the deepening of the mixed layer in boreal winter dampen or counteract thermal-driven variations. Externally forced thermal variations largely determine the subtropical seasonal cycle, while nonthermal variations drive the subpolar seasonal cycle[3,10,31,36]. As a result, the subtropical and subpolar North Atlantic have distinctly different seasonal patterns that are out of phase over the year. In the North Atlantic, particularly in the subpolar region, the North Atlantic Oscillation is considered the dominant mode of variability influencing SST patterns[37]; however, it is suggested that it only accounts for roughly 30% of the variability[38].

The second region, the Southern Ocean (0–360°, 35–65°S), given its vast areal domain, accounts for 40% of the global carbon uptake[30,39–41]. It remains under-sampled to surface ocean pCO2 measurements and satellite-derived observations that suffer from seasonal cloud coverage, e.g., Chl (used to upscale pCO2 measurements) and wind (used to estimate the gas transfer velocity), that are leveraged to estimate the carbon flux. Except for the Drake Passage, where sea surface pCO2 is sampled in four locations parallel to the mean flow of the Antarctic Circumpolar Current[42], year-round observations of pCO2 in the Southern Ocean are sparse[23]. Due to its remoteness and harsh weather conditions, observations remain limited to a few months of the year, mainly in the austral summer[43]. Additionally, in austral winter, when wind speeds and air–sea gas exchange rates are highest, first estimates of the carbon flux from biogeochemical Argo floats reveal a seasonal discrepancy compared to shipboard-based fluxes, which are yet to be resolved[44–46].

In the Southern Ocean, interannual variability of carbon fluxes has been linked to the decadal variability of the Southern Annular Mode, changes in which have been used to explain the weakening of the carbon sink in the 1990s and subsequent reinvigoration in the 2000s[5]. Seasonal dynamics in the Southern Ocean are complex with contrasting extremes—with deep mixing and entrainment processes characterising winter fluxes with carbon outgassing into the atmosphere and net primary productivity characterising summer fluxes with carbon uptake by the ocean[4]. However, the Southern Ocean is meridionally heterogeneous in space, with large-scale drivers including wind stress, surface heating, and mesoscale ocean dynamics influencing carbon exchange. As a result, the physical processes and drivers contributing to the variability of the carbon flux in the Southern Ocean are still debated[47,48], and their representation in models remains weakly constrained[16].

Previous studies have demonstrated that the Max Planck Institute ESM (MPI-ESM) is in limited agreement with other ESMs on the drivers of the seasonal cycle of carbon fluxes in the Southern Ocean[15,49]. Most other models part of the fifth phase of the Coupled Model Intercomparison Project (CMIP5[50]) show an exaggeration of the seasonal rates of change of SST in austral autumn and spring, which modifies the solubility of pCO2 and tips the control of the seasonal cycle toward SST. MPI-ESM compensates for the solubility bias with overly exaggerated primary productivity. Therefore, biologically driven changes in dissolved inorganic carbon (DIC) regulate the seasonal cycle of the flux[49], leading to large discrepancies in the magnitude of seasonal fluxes between MPI-ESM and observation-based products. For this reason, the MPI-ESM is perceived as anomalous in the Southern Ocean. Nevertheless, MPI-ESM shows a coherent seasonal cycle and thus allows us to ascertain whether seasonal variability can help diagnose interannual variability in our study regions or if seasonal discrepancies between the two ensembles, if present, corrupt the signal.

By contrasting a model ensemble and an observation-based ensemble in two regions that span the range of observational coverage and process understanding, our study tests the credibility of model simulations, where across ocean basins, seasonal fluxes are inconsistently represented[29,49]. Further, we re-evaluate the robustness of seasonal observation-based data, which, so far, has been considered well-constrained[16,17].

RESULTS

How do carbon fluxes compare between ensembles?

In the North Atlantic basin and the Southern Ocean, annual carbon fluxes in the model ensemble lie within the range of values in the observation-based ensemble. However, discrepancies between the two ensembles are one to two times greater for the seasonal fluxes than the annual fluxes in the North Atlantic basin and three to four times greater in the Southern Ocean (Figs. 1, 2).

For the North Atlantic basin, the model and observation-based ensembles generally agree on carbon uptake and outgassing regions and the seasonality of fluxes. This agreement indicates that the model captures key processes despite differences in magnitude and spatial inconsistencies in their representation of seasonal fluxes (Fig. 1a–j). In the observation-based ensemble, the
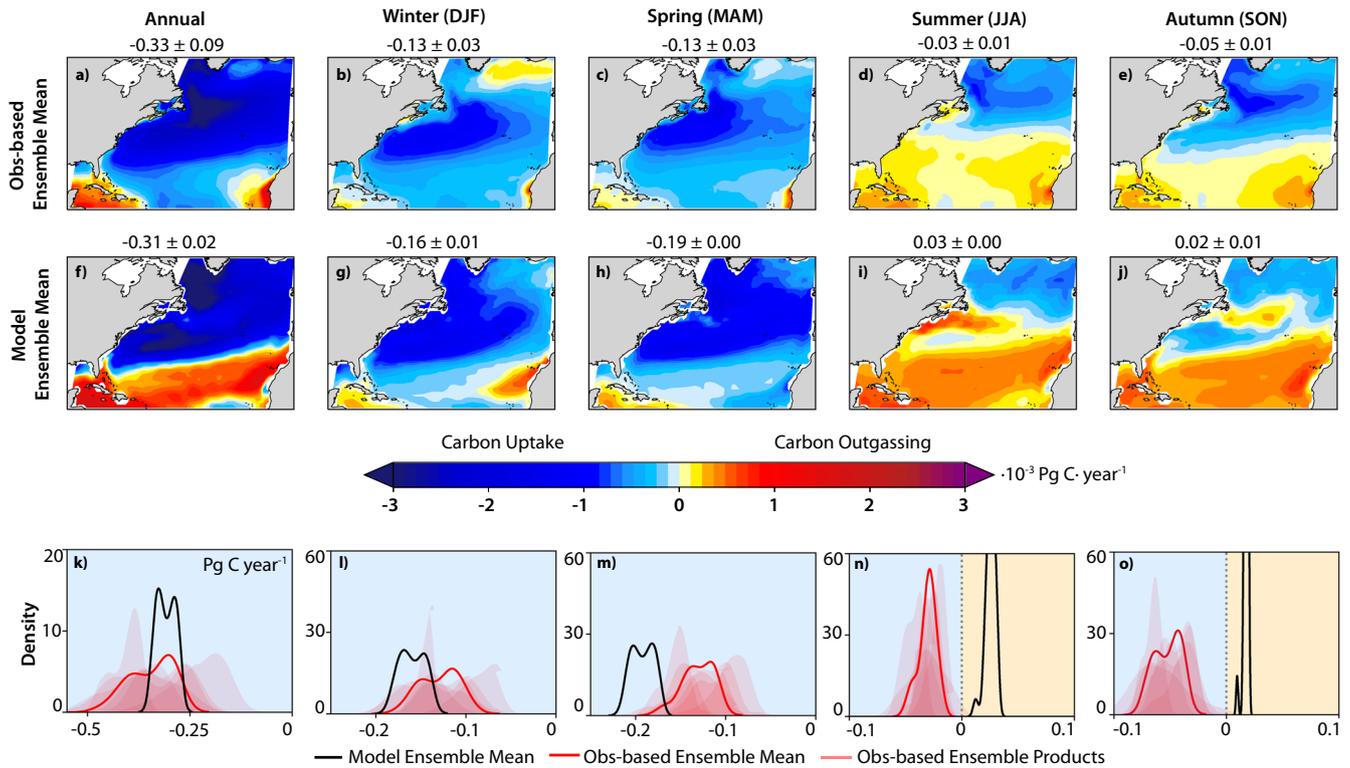
Fig. 1   Air-sea carbon fluxes in the North Atlantic basin (10–90°W, 15–60°N). Annual and seasonal (Pg C·year^{-1}) carbon fluxes averaged over 1990–2020 in the observation-based (**a–e**) and model (**f–j**) ensemble mean with the corresponding kernel density estimation plots in the observation-based (red) and model (black) ensembles (**k–o**). The red shadings represent individual products in the observation-based ensemble. Background colours indicate uptake by the ocean (blue) and outgassing into the atmosphere (orange).

subpolar North Atlantic supports a robust carbon sink in all seasons except winter, where deep convection can lead to outgassing (Fig. 1b–e). In spring and summer, phytoplankton blooms boost the intense carbon sink[51], atop a background state of surface waters travelling poleward that cool and subduct, creating a conduit for carbon to invade the ocean's interior[16]. The model ensemble fails to capture the outgassing associated with deep wintertime convection but represents the subpolar North Atlantic's carbon sink well in other seasons and the annual mean (Fig. 1g–j).

In the observation-based ensemble, the subtropical North Atlantic highlights a contemporary carbon sink region in response to increasing atmospheric carbon forcing[31] (Fig. 1a). In contrast, the model ensemble replicates the effect of temperature variations that drive the seasonality in the subtropical North Atlantic but shows larger outgassing than shown by the observation-based ensemble. Seasonal biological production has a much more substantial impact in the subtropics than most models indicate, which is widely thought to cause the lower seasonal amplitude in observations versus models[52] (Fig. 1f).

Integrated over the subpolar and subtropical regions of the North Atlantic basin, the model ensemble simulates enhanced uptake in half the year (boreal winter and spring) and enhanced outgassing in the other half (boreal summer and autumn) compared to the observation-based ensemble (Fig. 1k–o). The observation-based ensemble shows uptake in summer and autumn, so not only does the model ensemble show enhanced outgassing, but it also misses the direction of the net flux. The regional and seasonal discrepancies between the ensembles compensate in the annual mean, as evidenced by the density distributions of the annual flux, where the annual flux values of the model ensemble are within the distribution of the observation-based ensemble (Fig. 1k). However, at a seasonal

timescale, the distributions of the fluxes barely overlap, except in boreal winter. Further, it is noteworthy that agreement among the observation-based products is limited on an annual and seasonal timescale for the North Atlantic basin despite pooling from the same observational dataset (SOCAT[39]) and being a consistently well-observed ocean region. The disagreement in the observation-based products is consistent with larger differences observed in higher latitudes of the North Atlantic, likely linked to the complex interplay of physical and biological processes, which can be highly variable in the North Atlantic subregions[53], in addition to the effect of diverse choices of observational driver data used in the individual products.

For the Southern Ocean, integrated annual fluxes are close between the model and observation-based ensembles. Still, they disagree substantially on the magnitude and spatial structure of seasonal fluxes (Fig. 2a, f). While the kinetic transfer across the air–sea interface contributes to the mismatch between the two ensembles, we find that a large part of the observed difference in the air–sea carbon exchange stems from their different sea surface pCO2 seasonalities (Supplementary Fig. 1). In the observation-based ensemble, the Southern Ocean is a consistent and moderate CO$_2$ sink in all seasons (Fig. 2b–e), with outgassing focused south of the polar front, where upwelling of old DIC-rich waters raises the surface ocean pCO$_2$, except in the summer when biological productivity linked to the retreat of marginal ice favours carbon uptake. Contrarily, the model ensemble has stronger seasonal signals, simulating a very strong sink for half the year (austral spring and summer) and a moderate source in the other half (austral autumn and winter) (Fig. 2g–j). Previous studies link the spring and summer seasonal discrepancies in MPI-ESM with its deep entrainment bias, favouring biological production and enhancing uptake[16].
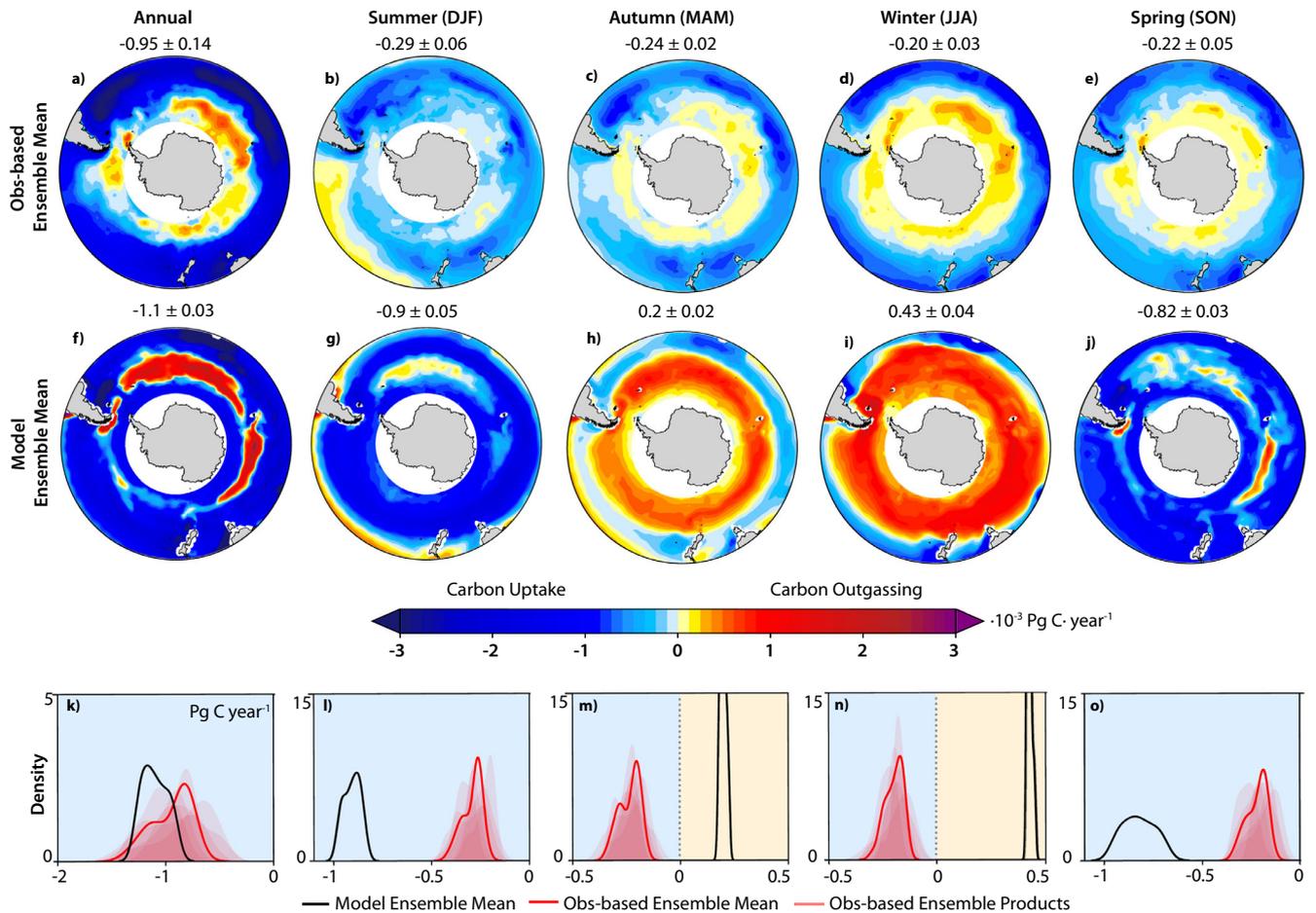
Fig. 2 Air-sea carbon fluxes in the Southern Ocean (0–360°, 35–65°S). Annual and seasonal (Pg C·year^{-1}) carbon fluxes averaged over 1990–2020 in the observation-based (**a**–**e**) and model (**f**–**j**) ensemble mean with the corresponding kernel density estimation plots in the observation-based (red) and model (black) ensembles (**k**–**o**). The red shadings represent individual products in the observation-based ensemble. Background colours indicate uptake by the ocean (blue) and outgassing into the atmosphere (orange).

Like the North Atlantic basin, discrepancies between the ensembles in the Southern Ocean compensate in the annual mean, despite no consistency between the model and observation-based ensembles on seasonal timescales (Fig. 2k–o). While there is a lack of skill in representing key processes that remain poorly understood and biased in the model, the incongruence between the ensembles in the Southern Ocean can also be attributed to the limitations of the observation-based ensemble. In particular, the seemingly strong agreement among the observation-based products could result from limited temporally biased data that reinforce weak signals over periods without observations. Therefore, even though the annual flux might suggest coherence between the model and observation-based ensembles, they compare poorly on the seasonal timescale, necessitating cautious assessment and interpretation of the annual carbon flux in the Southern Ocean in both the model and observation-based products.

Linking seasonal and interannual carbon flux variability

In a large model ensemble like the MPI-GE, the ensemble mean can be used to study the forced response[54]. However, the mean of the ensemble is limited in quantifying variability since averaging across the ensemble space suppresses the internal variability. However, each model ensemble member in MPI-GE has different variability since it is initialised with distinct starting conditions. We leverage this feature of the grand ensemble members ($n = 100$) to

explore the distribution of variability across seasonal and interannual timescales (Fig. 3).

Boreal winter and spring have the largest interannual variability in the North Atlantic basin in both ensembles and the largest range across individual members (Fig. 3a). The underestimation of variability in model members compared to observation-based products is not a new finding[55] and is possibly due to the representation of wintertime convective mixing and biological productivity in the subpolar region. The variability in both ensembles in boreal summer and autumn is low and is dominated by temperature-driven carbon exchange in the subtropical region. The narrow bounds in the model members hint that this process is relatively stable and well-captured in the MPI-ESM. In the North Atlantic basin, both ensembles agree on the seasons that contribute the most to interannual variability, with winter and spring fluxes showing the highest variability.

However, both ensembles diverge in their representation of interannual variability in the Southern Ocean, with a much greater magnitude of interannual variability than the North Atlantic basin across temporal scales (Fig. 3b). No season appears to dominate the variability in the observation-based products with nearly constant interannual variability across seasons. Observational coverage is biased to austral spring and summer when conditions are favourable, so the lack of variability in other seasons in the observation-based products could be an artefact of limited sampling and not a robust signal. In contrast, the model members capture a range of possible states, with the greatest variability in
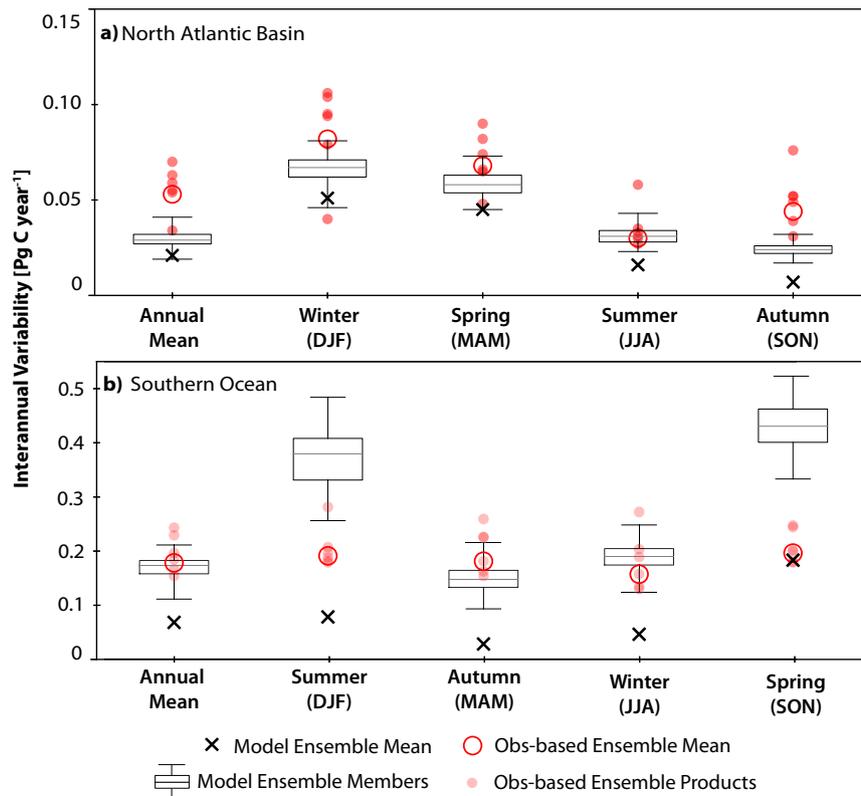
Fig. 3 Interannual variability (IAV) of air-sea carbon fluxes. IAV [Pg C·year^{-1}] of annual and seasonal carbon fluxes in observation-based products (red dots) overlaid on a box plot of IAV in model members for (**a**) the North Atlantic basin and (**b**) the Southern Ocean. The centre line, edges, and whiskers of the box are the median, upper and lower quartiles, and the range of values in the model ensemble space. The black crosses represent the IAV of the model ensemble mean, which is lower than that of the members since averaging dampens variability, and the large red circle is the IAV of the observation-based ensemble mean.

austral summer and spring when observations can be considered reliable. Both ensembles agree in the data-scarce seasons of austral autumn and winter that would be considered the least robust in the observation-based products. Current limitations in validating either the model or the observation-based ensemble make it challenging to deem this a trustworthy signal. Therefore, while seasonal variability can be useful in diagnosing interannual variability in the North Atlantic basin, it is currently of limited use in the Southern Ocean.

DISCUSSION

Process understanding of ocean carbon dynamics on seasonal and interannual timescales varies across the oceans due to hetero-geneous observations and complex underlying mechanisms. However, there is substantially improved historical data coverage and research in the Northern Hemisphere[18,39]. Therefore, seasonal fluxes are expected to be better represented in the North Atlantic basin than in the Southern Ocean. However, both the subtropical and the subpolar regions suffer from limitations in the model representation, failing to capture key seasonal processes linked to biology and mixing, respectively.

Nevertheless, both model and observation-based ensembles agree on the seasons that contribute the most to the interannual variability in the North Atlantic basin. Boreal winter and spring fluxes dominate interannual variability, in agreement with the current understanding of the role of MLD and SST anomalies[56]. Given the coherence between the ensembles, seasonal variability can be useful in diagnosing interannual variability in the North Atlantic and can provide direction to focus future observational campaigns.

In the Southern Ocean, the phasing of the seasonal cycle matches between the ensembles (most outgassing in austral winter and most uptake in austral summer, despite the seasonal cycle amplitude being larger in the model ensemble than the observation-based ensemble. However, both ensembles diverge in their representation of interannual variability. No season appears to dominate the variability in the observation-based products, with nearly constant interannual variability across seasons. This result is not particularly surprising given the lack of long-term observations in the Southern Ocean, except in the Drake Passage, and agrees with the findings of studies adding biogeochemical floats[44].

In contrast, the model members show stronger seasonal signals, with the greatest variability in austral summer and spring. Furthermore, model members cover a range of variability exceeding that of observation-based products, suggesting that observations in the Southern Ocean may be too limited to observe any robust signals in the carbon sink variability on seasonal timescales. That is, until autonomous measurement efforts, such as biogeochemical Argo floats, are deployed long enough to fill this gap[44–46]. Provided the data sparsity in the Southern Ocean, the lack of a dominant season driving the interannual variability could also result from observation-based products extrapolating summer signals, when observations are available, across the year. Without verifying such seasonal aliasing in the observation-based products and the lack of coherence across the model and observation-based ensemble, seasonal variability is currently of limited use in diagnosing interannual variability in the Southern Ocean, with a need for increased observational coverage throughout the year.

We compared annual and seasonal fluxes between an ESM large ensemble and an observation-based ensemble. Both ensembles show coherence in the magnitude and the large-scale spatial patterns of the annual carbon flux in the North Atlantic basin and the Southern Ocean. However, there are much larger discrepancies in seasonal fluxes than in annual fluxes between the two ensembles, one to two times greater in the North Atlantic basin and three to four times greater in the Southern Ocean. These seasonal discrepancies between the ensembles fortuitously compensate for the annual mean, which obscures large seasonal differences, particularly in the Southern Ocean. Therefore, considering only annual mean fluxes when comparing models and observation-based products misses a significant aspect of the comparison, which is only evident at a seasonal timescale.

We conducted this comparison with only one Earth System Model to reduce the uncertainties introduced in analysing a multi-model ensemble with varying physics, biogeochemistry, and configurations. Other ESMs (e.g., CESM-LE) show closer correspondence with observation-based products in the Southern Ocean (Supplementary Figs. 2–4). However, given the lack of a strong observational constraint in the Southern Ocean, it remains questionable whether the current representation of seasonal fluxes in any observation-based product can serve as a tool to benchmark the seasonal fluxes in any ESM[43]. Nevertheless, our analysis reveals that where necessary observational constraints have been built in the past, e.g., in the North Atlantic basin, a solid understanding of the seasonal variability can aid in understanding the interannual variability of carbon fluxes.

Long-term observation strategies can help improve process understanding and close measurement gaps in regions like the Southern Ocean, where models and observations disagree on seasonal carbon fluxes. However, maintaining existing networks can still provide insight into improving relatively well-understood regions like the North Atlantic basin, where fluxes remain prone to observational and model uncertainty[31]. However, data scarcity and its limiting distribution are shortcomings that will not be resolved globally in the short term. In the interim, models offer the ability to study regions where observations are inherently limiting[40]. Therefore, recognising the limitations and advantages of both models and observation-based products is of fundamental importance while interpreting the evolution of the ocean carbon sink.

METHODS

Observation-based data products

We use a set of regularly updated observation-based products where the air–sea carbon flux is computed (1990–2020). The products used in this study, LSCE-FFNN[21], CSIR-ML6[26], MPI-SOMFFN[57], Jena-MLS[23], JMA-MLR[25], and NIES-NN[24] were initially collected by the Surface Ocean pCO2 Mapping (SOCOM[17]) intercomparison project and later extended, for example, within the Global Carbon Budget[20]. The observation-based products are interpolated to the same monthly 1°×1° grid and averaged by treating individual products as an ensemble. All pCO2-based products use measurements from the 2020 release of the Surface Ocean CO2 Atlas (SOCAT[39]) database and a data-based interpolation approach to fill measurement gaps. Most interpolation techniques rely on machine-learning approaches that reconstruct the non-linear relationship between well-observed driver quantities (e.g., sea surface temperature (SST), sea surface salinity (SSS), sea surface height (SSH), mixed layer depth (MLD), chlorophyll a (Chl a), etc.) and sea surface pCO2 measurements. The JENA-MLS method is one exception, which uses a mixed layer scheme and is not dependent on auxiliary driver data. The products make subjective choices for selecting the gas transfer formulation and

wind product and use a quadratic wind-based gas transfer relationship to estimate the flux (more details in Table A3, GCP 2022[1] and refs. [21,23–26,57]).

LSCE-FFNN is an ensemble of neural network models trained on 100 subsampled pCO2 datasets. It uses SSS, SST, SSH, MLD, atmospheric CO_2 mole fraction, Chl a, pCO2 climatology, latitude, and longitude as predictors. CSIR-ML6 is an ensemble average of six machine-learning estimates of ocean pCO2 using two cluster-regression approaches (a) K-mean clustering and b) Fay and McKinley's (2014) CO_2 biomes) and three regression algorithms (a) gradient-boosted decision trees, b) a feed-forward neural network, and c) support vector regression), the product of which results in six estimates, with an ensemble mean and associated spread. MPI-SOMFFN uses a two-step neural network method clustering the ocean into 16 biogeochemical provinces using a self-organising map. The method leverages a non-linear relationship between pCO2 measurements and environmental predictor data (SST, SSS, MLD, Chl a, atmospheric CO2) using a feed-forward network. The established relationship is then used to fill existing data gaps. NIES-NN is a feed-forward neural network that uses SST, SSS, Chl a, MLD and monthly SST anomalies as predictor data. JMA-MLD uses fields of total alkalinity estimated using a multiple linear regression method based on data from GLODAP[58] and satellite data of SST, SSS, SSH, MLD and Chl a to evaluate DIC in the ocean surface. A more detailed description of the individual products is available in their respective core reference, and their intercomparison in part can be found in Rödenbeck et al. (2015) and Fay and Gregor (2021).

Max Planck Institute for Meteorology Grand Ensemble

We contrast the observation-based ensemble mean to the Max Planck Institute for Meteorology (MPI-M) Grand Ensemble (MPI-GE[59]) mean over its 100 members, based on the Max Planck Institute Earth System Model (MPI-ESM) version 1.1 in low-resolution configuration. The ocean (MPIOM[60]) and atmosphere (ECHAM6[61]) components are coupled daily without flux corrections[62]. The carbon flux is computed using a gas exchange parameterisation (based on 10 m above sea level wind speed) and the partial pressure difference at the air–sea interface[63]. MPI-GE is a large ensemble of a single state-of-the-art comprehensive climate model with ensemble members generated in 1850 from different starting conditions randomly selected from a preindustrial control simulation. The historical simulation (1850–2005) is not constrained by observations of the state of the atmosphere and ocean, except for forcing with greenhouse gases and volcanic aerosols, according to CMIP5[50]. For 2005–2020, we used a moderate radiative forcing scenario (RCP 4.5) to extend the MPI-GE until 2020; however, one could also choose a more extreme forcing pathway (e.g., RCP 8.5). The results are insensitive to the RCP choice since MPI-GE does not diverge significantly over the study period (Supplementary Tables 1, 2) but reflects changes between the two periods in both ensembles (Supplementary Tables 3–6). Using a large ensemble, we study the forced response of carbon fluxes due to anthropogenic warming separate from internal variability[54] and analyse the variability across the model ensemble space.

Methods

Annual and seasonal air–sea carbon flux. Mean annual air–sea carbon fluxes for both ensembles are computed from monthly values. We follow the convention that carbon uptake by the ocean (carbon outgassing into the atmosphere) is negative (positive). We define four distinct seasons as 3-month averages. For the North Atlantic basin (10–90°W, 15–60°N), we designate winter (DJF), spring (MAM), summer (JJA) and autumn (SON). These seasons are antiphase in the Southern Ocean (0–360°, 35–65°S), where we exclude the marginal ice zone by limiting the geographical scope

at 65°S and use the 35°S latitude to mark the boundary of the subtropical front.

Probability density estimation. For each year in the study period, we plot the kernel density estimate for the annual and seasonal integrated flux for the ensembles. The distributions of the individual observation-based products show the variability among the products. The closer the values are clustered, the less variable they are in time.

Interannual variability. Summed over the study region, we obtain a time series with the mean annual and seasonal magnitude of the air-sea carbon flux. After averaging over all grid points, we calculate the standard deviation about the mean of a particular season over 1990–2020 to determine which portion of the annual cycle has the largest interannual variability.

CODE AVAILABILITY
The code to reproduce the results of this study is available in the Zenodo archives (https://doi.org/10.5281/zenodo.7631425). The observation-based ensemble members can be accessed here: http://www.bgc-jena.mpg.de/SOCOM/. The MPI-GE members can be accessed here: https://esgf-data.dkrz.de/projects/mpi-ge/.

REFERENCES
1. Friedlingstein, P. et al. Global carbon budget 2021. *Earth Syst. Sci. Data* **14**, 1917–2005 (2022).
2. Rodgers, K. B. et al. A wintertime uptake window for anthropogenic CO_2 in the North Pacific. *Glob. Biogeochem. Cycles* **22**, 1–16 (2008).
3. Takahashi, T. et al. Global sea–air CO_2 flux based on climatological surface ocean pCO2, and seasonal biological and temperature effects. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **49**, 1601–1622 (2002).
4. Gregor, L., Kok, S. & Monteiro, P. M. S. Interannual drivers of the seasonal cycle of CO_2 in the Southern Ocean. *Biogeosciences* **15**, 2361–2378 (2018).
5. Landschützer, P., Gruber, N. & Bakker, D. C. E. Decadal variations and trends of the global ocean carbon sink. *Glob. Biogeochem. Cycles* **30**, 1396–1417 (2016).
6. Prodhomme, C. et al. How does the seasonal cycle control equatorial Atlantic interannual variability? *Geophys. Res. Lett.* **46**, 916–922 (2019).
7. Rödenbeck, C. et al. Interannual sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme. *Biogeosciences* **11**, 4599–4613 (2014).
8. Li, H. & Ilyina, T. Current and future decadal trends in the oceanic carbon uptake are dominated by internal variability. *Geophys. Res. Lett.* **45**, 916–925 (2018).
9. Kwiatkowski, L. & Orr, J. C. Diverging seasonal extremes for ocean acidification during the twenty-first century. *Nat. Clim. Change* **8**, 141–145 (2018).
10. Landschützer, P., Gruber, N., Bakker, D. C. E., Stemmler, I. & Six, K. D. Strengthening seasonal marine CO_2 variations due to increasing atmospheric CO_2. *Nat. Clim. Change* **8**, 146–150 (2018).
11. Nevison, C. D. et al. Evaluating CMIP5 ocean biogeochemistry and Southern Ocean carbon uptake using atmospheric potential oxygen: Present-day performance and future projection. *Geophys. Res. Lett.* **43**, 2077–2085 (2016).
12. Lefèvre, N. & Taylor, A. Estimating pCO2 from sea surface temperatures in the Atlantic gyres. *Deep Sea Res. Part Oceanogr. Res. Pap.* **49**, 539–554 (2002).
13. McKinley, G. A., Fay, A. R., Lovenduski, N. S. & Pilcher, D. J. Natural variability and anthropogenic trends in the ocean carbon sink. *Annu. Rev. Mar. Sci.* **9**, 125–150 (2017).
14. Hauck, J. et al. Consistency and challenges in the ocean carbon sink estimate for the global carbon budget. *Front. Mar. Sci.* **7**, 571720 (2020).
15. Fay, A. R. & McKinley, G. A. Observed regional fluxes to constrain modeled estimates of the ocean carbon sink. *Geophys. Res. Lett.* **48**, e2021GL095325 (2021).
16. Mongwe, N. P., Chang, N. & Monteiro, P. M. S. The seasonal cycle as a mode to diagnose biases in modelled CO_2 fluxes in the Southern Ocean. *Ocean Model.* **106**, 90–103 (2016).
17. Rödenbeck, C. et al. Data-based estimates of the ocean carbon sink variability first results of the Surface Ocean pCO2 Mapping intercomparison (SOCOM). *Biogeosciences* **12**, 7251–7278 (2015).
18. Watson, A. J. et al. Revised estimates of ocean-atmosphere CO_2 flux are consistent with ocean carbon inventory. *Nat. Commun.* **11**, 4422 (2020).
19. Olivarez, H. C. et al. Alternate histories: synthetic large ensembles of sea-air CO_2 flux. *Glob. Biogeochem. Cycles* **36**, e2021GB007174 (2022).
20. Friedlingstein, P. et al. Global carbon budget 2021. https://doi.org/10.5194/essd-2021-386 (2021).
21. Denvil-Sommer, A., Gehlen, M., Vrac, M. & Mejia, C. LSCE-FFNN-v1: a two-step neural network model for the reconstruction of surface ocean pCO2 over the global ocean. *Geosci. Model Dev.* **12**, 2091–2105 (2019).
22. Landschützer, P. et al. A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences* **10**, 7793–7815 (2013).
23. Rödenbeck, C. et al. Global surface-ocean pCO2 and sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme. *Ocean Sci.* **9**, 193–216 (2013).
24. Zeng, J., Nojiri, Y., Landschützer, P., Telszewski, M. & Nakaoka, S. A global surface ocean fCO2 climatology based on a feed-forward neural network. *J. Atmos. Ocean. Technol.* **31**, 1838–1849 (2014).
25. Iida, Y. et al. Trends in pCO2 and sea–air CO2 flux over the global open oceans for the last two decades. *J. Oceanogr.* **71**, 637–661 (2015).
26. Gregor, L., Lebehot, A. D., Kok, S. & Scheel Monteiro, P. M. A comparative assessment of the uncertainties of global surface ocean CO_2 estimates using a machine-learning ensemble (CSIR-ML6 version 2019a) – have we hit the wall? *Geosci. Model Dev.* **12**, 5113–5136 (2019).
27. Ritter, R. et al. Observation-based trends of the Southern Ocean carbon sink. *Geophys. Res. Lett.* **44**, 12,339–12,348 (2017).
28. Fay, A. R. et al. SeaFlux: harmonisation of air–sea CO_2 fluxes from surface pCO2 data products using a standardised approach. *Earth Syst. Sci. Data* **13**, 4693–4710 (2021).
29. Lenton, A. et al. Sea-air CO_2 fluxes in the Southern Ocean for the period 1990–2009. *Biogeosciences* **10**, 4037–4054 (2013).
30. Gruber, N., Landschützer, P. & Lovenduski, N. S. The variable Southern Ocean carbon sink. *Annu. Rev. Mar. Sci.* **11**, 159–186 (2019).
31. Schuster, U. et al. An assessment of the Atlantic and Arctic sea–air CO_2 fluxes, 1990–2009. *Biogeosciences* **10**, 607–627 (2013).
32. Mikaloff Fletcher, S. E. et al. Inverse estimates of anthropogenic CO_2 uptake, transport, and storage by the ocean. *Glob. Biogeochem. Cycles* **20**, 1–19 (2006).
33. Macovei, V. A. et al. Impact of physical and biological processes on temporal variations of the ocean carbon sink in the mid-latitude North Atlantic (2002–2016). *Prog. Oceanogr.* **180**, 102223 (2020).
34. Halloran, P. R. et al. The mechanisms of North Atlantic CO_2 uptake in a large Earth System Model ensemble. *Biogeosciences* **12**, 4497–4508 (2015).
35. Takahashi, T., Olafsson, J., Goddard, J. G., Chipman, D. W. & Sutherland, S. C. Seasonal variation of CO_2 and nutrients in the high-latitude surface oceans: a comparative study. *Glob. Biogeochem. Cycles* **7**, 843–878 (1993).
36. Ishii, M. et al. Air–sea CO_2 flux in the Pacific Ocean for the period 1990–2009. *Biogeosciences* **11**, 709–734 (2014).
37. Schuster, U. & Watson, A. J. A variable and decreasing sink for atmospheric CO_2 in the North Atlantic. *J. Geophys. Res.* **112**, C11006 (2007).
38. Marshall, J. et al. North Atlantic climate variability: phenomena, impacts and mechanisms. *Int. J. Climatol.* **21**, 1863–1898 (2001).
39. Bakker, D. C. E. et al. A multi-decade record of high-quality fCO2 data in version 3 of the surface ocean CO2 Atlas (SOCAT). *Earth Syst. Sci. Data* **8**, 383–413 (2016).
40. Gloege, L. et al. Quantifying errors in observationally based estimates of ocean carbon sink variability. *Glob. Biogeochem. Cycles* **35**, e2020GB006788 (2021).
41. Frölicher, T. L. et al. Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *J. Clim.* **28**, 862–886 (2015).
42. Munro, D. R. et al. Recent evidence for a strengthening CO_2 sink in the Southern Ocean from carbonate system measurements in the Drake Passage (2002–2015). *Geophys. Res. Lett.* **42**, 7623–7630 (2015).
43. Monteiro, P. M. S. et al. Intraseasonal variability linked to sampling alias in air-sea CO_2 fluxes in the Southern Ocean. *Geophys. Res. Lett.* **42**, 8507–8514 (2015).
44. Bushinsky, S. M. et al. Reassessing Southern Ocean air-sea CO_2 flux estimates with the addition of biogeochemical float observations. *Glob. Biogeochem. Cycles* **33**, 1370–1388 (2019).
45. Gray, A. R. et al. Autonomous biogeochemical floats detect significant carbon dioxide outgassing in the high-latitude Southern Ocean. *Geophys. Res. Lett.* **45**, 9049–9057 (2018).
46. Williams, N. L. et al. Assessment of the carbonate chemistry seasonal cycles in the Southern Ocean from persistent observational platforms. *J. Geophys. Res. Oceans* **123**, 4833–4852 (2018).
47. DeVries, T. et al. Decadal trends in the ocean carbon sink. *Proc. Natl Acad. Sci. USA* **116**, 11646–11651 (2019).
48. Landschützer, P. et al. The reinvigoration of the Southern Ocean carbon sink. *Science* **349**, 1221–1224 (2015).

49. Mongwe, N. P., Vichi, M. & Monteiro, P. M. S. The seasonal cycle of pCO2 and CO_2 fluxes in the Southern Ocean: diagnosing anomalies in CMIP5 Earth system models. *Biogeosciences* **15**, 2851–2872 (2018).

50. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).

51. Watson, A. J. et al. Tracking the variable North Atlantic sink for atmospheric CO_2. *Science* **326**, 1391–1393 (2009).

52. Wu, Z. et al. Modeling photosynthesis and exudation in subtropical oceans. *Glob. Biogeochem. Cycles* **35**, e2021GB006941 (2021).

53. Metzl, N. et al. Recent acceleration of the sea surface fCO_2 growth rate in the North Atlantic subpolar gyre (1993-2008) revealed by winter observations. *Glob. Biogeochem. Cycles* **24** (2010).

54. Milinski, S., Maher, N. & Olonscheck, D. How large does a large ensemble need to be? Earth Syst. *Earth Syst. Dyn.* **11**, 885–901 (2020).

55. Crueger, T., Roeckner, E., Raddatz, T., Schnur, R. & Wetzel, P. Ocean dynamics determine the response of oceanic CO_2 uptake to climate change. *Clim. Dyn.* **31**, 151–168 (2008).

56. Gruber, N., Keeling, C. D. & Bates, N. R. Interannual variability in the North Atlantic Ocean carbon sink. *Science* **298**, 2374–2378 (2002).

57. Landschützer, P., Gruber, N., Bakker, D. C. E. & Schuster, U. Recent variability of the global ocean carbon sink. *Glob. Biogeochem. Cycles* **28**, 927–949 (2014).

58. Lauvset, S. K. et al. A new global interior ocean mapped climatology: the 1∘ × 1∘ GLODAP version 2. *Earth Syst. Sci. Data* **8**, 325–340 (2016).

59. Maher, N. et al. The Max Planck Institute Grand Ensemble: enabling the exploration of climate system variability. J. Adv. Model. *J. Adv. Model Earth Syst.* **11**, 2050–2069 (2019).

60. Jungclaus, J. H. et al. Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *J. Adv. Model. Earth Syst.* **5**, 422–446 (2013).

61. Stevens, B. et al. Atmospheric component of the MPI-M Earth System Model: ECHAM6. *J. Adv. Model Earth Syst.* **5**, 146–172 (2013).

62. Giorgetta, M. A. et al. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5: climate changes in MPI-ESM. *J. Adv. Model Earth Syst.* **5**, 572–597 (2013).

63. Wanninkhof, R. Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res.* **97**, 7373 (1992).

AUTHOR CONTRIBUTIONS

J.B., P.L., S.B., and P.R. initiated the design of the study that led to this paper. P.R. led the analysis and designed the figures. P.R. wrote the article with input from P.L., S.B., and J.B. All authors contributed to the writing at all stages.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-023-00378-3.

Correspondence and requests for materials should be addressed to Paridhi Rustogi.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.