Equatorial Pacific pCO$_2$ Interannual Variability in CMIP6 Models

Suki C. K. Wong, Galen A. McKinley, and Richard Seager

1Department of Earth and Environmental Sciences, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA

Abstract The El Niño-Southern Oscillation (ENSO) in the equatorial Pacific is the dominant mode of global air-sea carbon dioxide (CO$_2$) flux interannual variability (IAV). Air-sea CO$_2$ fluxes are driven by the difference between atmospheric and surface ocean pCO$_2$, with variability of the latter driving flux variability. Previous studies found that models in Coupled Model Intercomparison Project Phase 5 (CMIP5) failed to reproduce the observed ENSO-related pattern of CO$_2$ fluxes and had weak pCO$_2$ IAV, which were explained by both weak upwelling IAV and weak mean vertical dissolved inorganic carbon (DIC) gradients. We assess whether the latest generation of CMIP6 models can reproduce equatorial Pacific pCO$_2$ IAV by validating models against observations-based data products. We decompose pCO$_2$ IAV into thermally and non-thermally driven anomalies to examine the balance between these competing anomalies, which explain the total pCO$_2$ IAV. The majority of CMIP6 models underestimate pCO$_2$ IAV, while they overestimate sea surface temperature IAV. Insufficient compensation of non-thermal pCO$_2$ to thermal pCO$_2$ IAV in models results in weak total pCO$_2$ IAV. We compare the relative strengths of the vertical transport of temperature and DIC and evaluate their contributions to thermal and non-thermal pCO$_2$ anomalies. Model-to-observations-based product comparisons reveal that modeled mean vertical DIC gradients are biased relative to their mean vertical temperature gradients, but upwelling acting on these gradients is insufficient to explain the relative magnitudes of thermal and non-thermal pCO$_2$ anomalies.

Plain Language Summary To date, the global ocean has been responsible for absorbing over a third of carbon dioxide (CO$_2$) emissions, slowing down the growth of atmospheric CO$_2$ levels which drives global warming. Of interest is the equatorial Pacific Ocean, which is the largest oceanic source of CO$_2$ to the atmosphere with large fluctuations that are apparent in the record of global atmospheric CO$_2$. To study the ocean’s ability to absorb future CO$_2$ emissions, we need models of the Earth system that can accurately capture fluctuations in the equatorial Pacific. In this paper, we assess surface ocean CO$_2$ fluctuations in the equatorial Pacific in the latest generation of models and we examine their deviations from observations. Compared to observations, models underestimate surface ocean CO$_2$ fluctuations as a result of excessive cancellation between competing drivers of CO$_2$ change. We find that the vertical gradient of carbon in models is too weak, which through ocean circulation, would contribute to weak surface CO$_2$ fluctuations. However, this does not fully account for underestimations in surface CO$_2$ fluctuations. Other processes have a significant role in excessively canceling surface CO$_2$ concentrations and requires further research.

1. Introduction

Carbon dioxide (CO$_2$) in the atmosphere is the main driver of anthropogenic radiative forcing via the greenhouse effect. Natural sinks in the ocean and land are damping the atmospheric CO$_2$ growth rate. The latest assessment of the global carbon budget averaged over recent decades (1960–2020) estimates the airborne fraction of atmospheric CO$_2$ emissions to be about 45%, with the remainder of emissions partitioned among the ocean (25%) and land (30%) (Friedlingstein et al., 2022). However, uncertainties in quantifying aspects of the global carbon cycle result in an imbalance in the carbon budget, which is largely attributed to errors in land and ocean sink estimates (Friedlingstein et al., 2022). Constraining ocean interannual variability (IAV) will help to reduce uncertainty in land IAV.

The equatorial Pacific is the largest natural oceanic source of CO$_2$ to the atmosphere (Takahashi et al., 2009), as a result of wind-driven upwelling in the region; upwelling brings cool waters that are rich in dissolved inorganic carbon (DIC) to the surface, which increases the partial pressure of CO$_2$ in the surface ocean (pCO$_2$). CO$_2$ outgassing IAV in the equatorial Pacific is dominated by the El Niño-Southern Oscillation (ENSO), and is
the dominant mode of global ocean sink IAV (Rödenbeck et al., 2014). ENSO mechanisms of air-sea CO₂ flux (FCO₂) variability are well understood. During an ENSO warm phase (El Niño), slackening trade winds over the equator reduces upwelling and brings about warm sea surface temperature (SST) anomalies (Bjerknes, 1966). Warm SST anomalies increase pCO₂ via reduced CO₂ solubility. However, it is the reduction in surface DIC due to reduced upwelling that dominates the CO₂ response (reduced CO₂ outgassing) during an El Niño (McKinley et al., 2004). During an ENSO cold phase (La Niña), the opposite happens and CO₂ outgassing is enhanced.

In Coupled Model Intercomparison Project Phase 5 (CMIP5), atmosphere-ocean global climate models were coupled with biogeochemical processes for the first time in CMIP history, allowing for carbon cycling in models (Emori et al., 2016; Taylor et al., 2012). Studies have reported biases in simulated equatorial Pacific pCO₂ and FCO₂ IAV in CMIP5 models, where weak surface DIC variability was found to be a source of bias in some models (Dong et al., 2017; Jin et al., 2019). Given ongoing climate change, there is a need for Earth System Models (ESMs) to make accurate climate projections. The latest generation of ESMs from CMIP6 have demonstrated progress in representing the mean state of ocean biogeochemistry (Séférian et al., 2020). However, as in CMIP5, weak FCO₂ IAV were also found in CMIP6 (Vaittinada Ayar et al., 2022). Identifying sources of model biases in FCO₂ IAV for the contemporary period, where some data constraints exist, is a first step toward model improvements.

Here, we assess equatorial Pacific pCO₂ IAV in 18 CMIP6 models over recent decades, comparing amplitudes and spatial patterns of variability against state-of-the-art observations-based pCO₂ products that span over five decades. We also compare the covariability of ENSO-related variables, such as SSTs, vertical velocity at 50 m (w₅₀), and thermocline depths with pCO₂ anomalies across the CMIP6 subset through lagged correlations. To understand biases in pCO₂ IAV, we decompose pCO₂ IAV into thermally (SST) and non-thermally (DIC, alkalinity and salinity) driven components. Imbalances between these competing components provide insight on biases in the total pCO₂ IAV.

In the equatorial Pacific, surface DIC variability dominates pCO₂ variability (Doney et al., 2009). Though there are several processes that drive DIC variability (FCO₂, freshwater fluxes, biology, vertical and horizontal transport), studies show that variability in the vertical transport of DIC is important to the overall budget of pCO₂ variability in the equatorial Pacific Ocean (Liao et al., 2020). Including temperature-driven pCO₂ variability, Liao et al. (2020) showed that the vertical transport term contributed the largest amount in their full mixed-layer pCO₂ budget decomposition (accounting for about 40% of the pCO₂ response; FCO₂ ~20%; biology ~18%; FW ~11%; horizontal transport ~10%; thermal and residual <1%). This demonstrated importance of the vertical transport of DIC in the equatorial Pacific motivates our investigation of its variability in CMIP6. There is also reason to believe that models are biased in mean vertical gradients (Farneti et al., 2022; Li & Xie, 2012), which through upwelling, could contribute to biases in surface DIC variability.

Our objectives are as follows: (a) compare equatorial Pacific pCO₂ IAV in CMIP6 models against observations-based data products, (b) understand why models underestimate pCO₂ IAV, and (c) identify sources of bias in the vertical transport of DIC in models. Given biases in mean vertical gradients of DIC and temperature, we quantify the degree to which upwelling anomalies (acting on biased gradients) contribute to the relative magnitudes of non-thermal and thermal pCO₂ IAV, respectively. Such assessment is necessary to ground work on how future changes in the variability and mean state of the tropical Pacific atmosphere-ocean system will also impact variability and shifts in air-sea CO₂ fluxes, with potential climate impact.

2. Models, Data and Methods

2.1. Models

Outputs from historical simulations (1959–2014) from 18 CMIP6 models (Table 1) are from the Pangeo cloud (http://pangeo.io), which were originally downloaded from the Earth System Grid Federation (ESGF)'s online archives (http://esgf-node.llnl.gov/projects/cmip6). We apply a data pre-processing Python tool to clean and unify data inconsistencies before any analysis (Busecke & Abernathey, 2020). We assess 18 models which have monthly pCO₂, FCO₂, SST, near-surface wind speeds measured at 10 m (uₙ₀), ocean temperatures (T), w₅₀ and DIC data available. Vertical velocities are calculated using the three-dimensional continuity equation for models that only have horizontal circulation data. For analyses that involve multiple ensemble members, ensemble members are chosen only if they have outputs for all the variables named above. This ensures that the internal variability,
unique to each run of a model (an ensemble member), is conserved across all output variables from a single run. For a list of the members that we use for each model, see Table S1 in Supporting Information S1.

2.2. Observations-Based Data Products

We use five out of the six available monthly gridded observations-based pCO$_2$ products from SeaFlux (Fay et al., 2021) for FCO$_2$ and surface ocean pCO$_2$ estimates. These five products include JENA-MLS, MPI-SOMFFN, CMEMS-FFN, JMA-MLR, and CSIR-ML6. We exclude a sixth product (NIES-FFNN) from our assessment as it was not able to recreate ENSO variability in pCO$_2$, such as the strong 1997–1998 El Niño event seen in the other products. For $u_{10}$ data, we also use the three wind reanalysis products (CCMPv2, JRA55, and ERA5), used in SeaFlux to estimate fluxes. DIC and total alkalinity (Alk) climatologies are from GLODAPv2 (Lauvset et al., 2021). GLODAPv2 is a mapped three-dimensional climatological data product of inorganic and carbon-related ocean variables. Observations of DIC and Alk are distributed in time too scarcely to allow determination of its time variation, so in GLODAPv2, the data have been averaged into a DIC climatology estimate. Monthly estimates of SST, ocean circulation and ocean temperature (1959–2014) are from a reanalysis product, ORASS of the European Center for Medium Range Weather Forecasts (Zuo et al., 2019). Unlike DIC and Alk, time variations can be resolved for SST, ocean circulation and ocean temperature variables in observations-based data products. We calculate vertical velocity using zonal and meridional ocean circulation data from ORASS via the vertical integration of the equation continuity. SST observations from another data set, HadISST (1959–2014; Rayner et al., 2003), are a secondary source for SST comparisons against models.

2.3. Methods

Model outputs are regridded to the same $1° \times 1°$ longitude-latitude grid before any analysis. We define a region of the equatorial Pacific ($5°N$–$5°S$) between $180°E$ and $270°E$, which encompasses the Niño 3 and 3.4 regions, extending $10°$ west of Niño 3.4, and refer to it as the Tropical Pacific Index (TPI) region. The Niño 3 and 3.4 regions are typically used to study the nature of ENSO variability over the equatorial Pacific Ocean, but here, the broader TPI region was chosen such that any longitudinal differences in the ENSO centers of action in models would be captured.

To compare relative amplitudes of IAV across models and other datasets, we use one standard deviation ($\sigma$) of detrended and deseasonalized monthly anomalies. Modeled FCO$_2$, pCO$_2$, and $u_{10}$ IAV are compared against SeaFlux IAV. Note that historical simulations in CMIP6 models generate their own internal climate variability, and will not replicate the timings of historical events unless they are externally forced. Thus, when comparing SeaFlux IAV to model IAV, the temporal evolution is not expected to match. When calculating and comparing multi-year means between CMIP6 models and SeaFlux, data from the same time frame (1990–2014) are compared. This is done since multi-year means are sensitive to anthropogenic trends in CO$_2$; the ocean sink is changing over time in both observations-based data products and historical simulations, such that multi-year means are sensitive to the time frame over which the average is taken. The 1990–2014 time frame is chosen for multi-year means, because temporal coverage begins in 1990 for SeaFlux, and 2014 is the end year for CMIP6 historical simulations. Climatological monthly means taken over the study period are subtracted from monthly timeseries data to obtain deseasonalized monthly anomalies, and then, the data are detrended with the least squares method. In addition to model comparisons against SeaFlux, modeled SST IAV and vertical DIC gradients IAV are also compared against observations-based data products.

Spatial patterns of pCO$_2$ IAV are compared and assessed by calculating its first empirical orthogonal function (EOF) after detrending and deseasonalizing. EOF analyses are done on individual ensemble members that retain
full internal variability, and then averaged across ensemble members. The first principal components (PC1) and associated EOFs are all shown for the La Niña state (i.e., positive pCO₂ anomaly), as determined with reference to the sign of the TPI SST index. Model performances in reproducing IAV are assessed using spatial correlation coefficients (SCCs) between each model and observations-based pCO₂ patterns of IAV.

In order to examine the mechanisms of pCO₂ variability in models, local correlations between pCO₂ and SST anomalies within the tropical Pacific are calculated. Areas of strong correlations indicate regions in models where upwelling dominates pCO₂, which is consistent with the dominant ENSO signal. Lagged temporal correlations between pCO₂, SST, w_50 and thermocline depth (z_therm) anomalies are also done to investigate the co-variability of ENSO-related variables to pCO₂ anomalies. We define the z_therm as the depth of the maximum vertical temperature gradient. Time lags between variables are based on the lags seen in the observations-based data products: pCO₂ and SST are concurrently correlated, while w_50 and z_therm anomalies lead pCO₂ by up to 3 months. Three-month running means of w_50 and z_therm anomalies are taken before correlating them to the pCO₂ of the fourth month (e.g., the January-to-March mean of w_50 anomalies are correlated to April’s pCO₂ anomaly). pCO₂ is long-lived in the ocean, such that the influence of w_50 and z_therm variability on local pCO₂ advects west due to mean currents during the 3 months of lag. To account for some of the westward advection of pCO₂ during the lag period, w_50 and z_therm anomalies are calculated over a region 20° east of the TPI box region before correlating with pCO₂ anomalies over the TPI box region.

### 2.4. Thermal and Non-Thermal pCO₂, IAV

DIC, alkalinity (Alk) and salinity (S) are the non-thermal drivers of pCO₂ variability, while SST variability is the thermal driver. Thermal effects on pCO₂ typically oppose and dampen the non-thermal effects with ENSO (Sutton et al., 2014): for example, a reduction in upwelling brings less DIC to the surface which decreases surface pCO₂; simultaneously, the warmer SST anomalies, as a result of weakened upwelling, drives surface pCO₂ up via reduced solubility. We separate the non-thermally driven pCO₂ (pCO₂,nonT) from the thermally-driven counterpart (pCO₂,therm) in order to explain modeled pCO₂ IAV. For pCO₂,nonT temperature effects are removed by normalizing pCO₂ outputs to a long-term mean SST (Takahashi et al., 2002), following an empirical formulation determined by Takahashi et al. (1993):

\[
pCO₂,\text{nonT} = pCO₂ \times e^{0.0423 \cdot (SST - \overline{SST})},
\]

where \(\overline{SST}\) is the multiyear mean of SST over time. The thermally driven component, pCO₂,therm, is computed using the following equation (Takahashi et al., 2002):

\[
pCO₂,\text{therm} = pCO₂,\overline{\text{T}} \times e^{0.0423 \cdot (\overline{SST} - SST)},
\]

where \(pCO₂,\overline{\text{T}}\) is the multiyear mean of pCO₂ during 1990–2014.

### 2.5. Vertical Transport of Dissolved Inorganic Carbon

Temporal changes in pCO₂ are a function of temporal changes in DIC, Alk, S and T, and can be expressed as the following linearly decomposed time derivative (Le Quéré et al., 2000; Liao et al., 2020; Takahashi et al., 1993):

\[
\partial_t pCO₂ = \partial_{\text{DIC}} pCO₂ \frac{\partial \text{DIC}}{\partial T} + \partial_{\text{Alk}} pCO₂ \frac{\partial \text{Alk}}{\partial T} + \partial_{S} pCO₂ \frac{\partial S}{\partial T} + \partial_{\overline{T}} pCO₂ \frac{\partial \overline{T}}{\partial T},
\]

where we use the notation \(\partial\) to denote a partial derivative with respect to time. Temporal changes in DIC, Alk and S drive pCO₂,nonT, while temporal changes in SST drive pCO₂,therm.

In the tropical Pacific, DIC variability has been found to be the dominant driver of pCO₂ variability, compared to Alk, S and T drivers (Doney et al., 2009; Le Quéré et al., 2000). From model results, Liao et al. (2020) found that alkalinity-driven effects on pCO₂ can exceed DIC-driven effects during El Niño, though DIC effects generally dominate in the eastern equatorial Pacific. Other model studies confirm that DIC is the dominant term in
the region (Jin et al., 2019; Long et al., 2013). The time tendency of surface DIC (\(\partial_t DIC\)) is controlled by several processes including horizontal (\(H\)) and vertical (\(V\)) ocean transport, \(FCO_2\), biological processes (\(Bio\)) and FW:

\[
\partial_t DIC \approx \partial_t DIC_H + \partial_t DIC_V + \partial_t DIC_{FCO_2} + \partial_t DIC_{Bio} + \partial_t DIC_{FW}
\]

(4)

Liao et al. (2020) showed that though other processes are non-negligible, vertical transport contributed the largest effect on \(pCO_2\) change, with important vertical contributions from alkalinity and DIC in the model composite response. They also showed that the other processes are sensitive to changes in vertical transport: an increase in upwelling (increased surface DIC) drives an air-sea flux response, which damps surface DIC; upwelled nutrient-rich waters increase biological activity causing an increased uptake of DIC, which again damps surface DIC; upwelled DIC + \(t\) results in a diverging transport, also damping. For CMIP5 models, Jin et al. (2019) showed that the vertical transport of DIC to the surface ocean overwhelms over other processes in reducing \(pCO_2\) during an El Niño. In this study, we assess only the variability in vertical transport of DIC (\(\partial_t DIC_V\)).

In order to quantify the contribution of the vertical transport of DIC (\(\partial_t DIC_V\)) to \(pCO_2\) variability (\(\partial pCO_2,\text{nonT}\)), we evaluate the former in the same units as the latter—in units of the time tendency of \(pCO_2\) (\(\muatm^{-1}\)—and write \(\partial_t DIC_V\) as \(w_{50}\partial_z DIC\). Using coefficients from Equation 3, we can get both terms into the same units:

\[
\frac{\partial pCO_2}{\partial DIC} w_{50}\partial_z DIC \quad [\text{units} : \muatm^{-1}]
\]

(5)

\[
\partial_t pCO_2,\text{nonT} \quad [\text{units} : \muatm^{-1}]
\]

(6)

The coefficients used for the \(pCO_2\) dependence on DIC are approximated as follows (Lovenduski et al., 2007):

\[
\frac{\partial pCO_2}{\partial DIC} \approx \frac{pCO_2}{DIC} \cdot \frac{3 \times \text{Alk} \times DIC - 2 \times DIC^2}{2 \times DIC - \text{Alk}} \left( \text{Alk - DIC} \right),
\]

(7)

which can be expressed more simply as:

\[
\frac{\partial pCO_2}{\partial DIC} \approx \frac{pCO_2}{DIC} \cdot \gamma_{DIC},
\]

(8)

where \(\gamma_{DIC}\) is the buffer factor (Sarmiento & Gruber, 2006).

2.6. Reynolds' Decomposition

Using Reynolds' decomposition, we can separate the mean and the time-varying component:

\[
w_{50} = \bar{w}_{50} + w_5', \quad \text{and}
\]

(9)

\[
\partial_z DIC = \partial_z DIC + \partial_z DIC'.
\]

(10)

where primes denote detrended monthly anomalies and overbars denote long-term means. The time-varying component of the vertical transport of DIC, \(\frac{\partial pCO_2}{\partial DIC} (w_{50}\partial_z DIC')\), can be decomposed into three Reynolds' terms:

\[
\frac{\partial pCO_2}{\partial DIC} (w_{50}\partial_z DIC') = \frac{\partial pCO_2}{\partial DIC} \left( w_{50}\partial_z DIC - \bar{w}_{50}\partial_z DIC' \right)
\]

\[
= \frac{\partial pCO_2}{\partial DIC} \left( \bar{w}_{50}\partial_z DIC' + w_5'\partial_z DIC' \right)
\]

(11)

\[
\text{1st term}
\]

\[
\text{2nd term}
\]

\[
\text{3rd term}
\]
Note that the left-hand-side of Equation 11 is only the turbulent term of the vertical transport of DIC. For models, we compute all three Reynolds terms. In GLODAPv2, gridded DIC data are only available as a climatology. Since we can’t calculate the time-varying vertical gradient of DIC, \( \partial \), we only compare the second Reynolds term \( A_\left( \frac{\partial^2}{\partial z^2} \langle w' \rangle \frac{\partial}{\partial z} \text{DIC} \right) \) between models and data.

### 3. Results

#### 3.1. FCO\(_2\) and pCO\(_2\) Multiyear Means

A large region of FCO\(_2\) outgassing can be seen in the equatorial Pacific Ocean, with the highest (positive, red) values being in the eastern region in SeaFlux (Figure 1: top-left). Comparing mean fluxes, five models (MIROC-ES2L, CNRM-ESM2-0, UKESM1-0-LL, MRI-ESM2-0, and MPI-ESM-1-2-HAM) are shown in Figure 1. These five models are selected to represent the range of abilities of CMIP6 models to reproduce data. Similar maps for all the CMIP6 models are available in Figure S1 in Supporting Information S1. The models have similar patterns to SeaFlux to the first order, with a basin-wide outgassing feature seen over the equatorial Pacific region, and the largest values lying in the eastern region. Model mean fluxes in the equatorial Pacific are typically weaker than SeaFlux, with the exception of UKESM1-0-LL which has a mean magnitude closer to SeaFlux (the CESM2 family of models also have comparable mean FCO\(_2\) values, Figure S1 in Supporting Information S1). The mean outgassing in the equatorial region is noticeably weaker in MRI-ESM2-0 than the other models, and the MPI models show a narrow band of near-zero flux at the equator in the middle of the broader equatorial outgassing pattern.

Multiyear mean maps of pCO\(_2\), averaged over 1990–2014, are plotted for SeaFlux and the same five CMIP6 models (Figure 2; Figure S2 in Supporting Information S1: all models). A SCC over the TPI region is calculated between each model and SeaFlux to quantify the model skill at reproducing the mean pCO\(_2\) pattern. Note that a high SCC score does not indicate that the magnitude of the mean maps are similar. Generally, the majority of models produce the high pCO\(_2\) equatorial structure seen in SeaFlux, with a third of all 18 models having an SCC score above 0.8 (Figure S2 in Supporting Information S1). The largest pCO\(_2\) values are seen off coastal Peru and Panama in SeaFlux, with exaggerated coastal values seen in some of the models (MRI-ESM2-0 and UKESM1-0-LL). Unlike SeaFlux, the high pCO\(_2\) equatorial structure extends almost all the way across the basin.
in the majority of models, except for the NorESM2 models (Figure S2 in Supporting Information S1). Similar to the FCO\textsubscript{2} multiyear mean maps, the MPI models show mean pCO\textsubscript{2} structures that exhibit an equatorial band of low pCO\textsubscript{2} that splits up the general high pCO\textsubscript{2} structure seen in SeaFlux and the other models.

3.2. Interannual Variability

The outgassing of CO\textsubscript{2} in the equatorial Pacific Ocean (Figure 1) is modulated by ENSO variability, which dominates the variability of global oceanic FCO\textsubscript{2} (Landschützer et al., 2016; McKinley et al., 2004, 2017; Rödenbeck et al., 2014). Amplitudes of FCO\textsubscript{2} IAV (σFCO\textsubscript{2}) in the TPI region in CMIP6 differ from SeaFlux observations-based data products (Figure 3a). The majority of CMIP6 models underestimate FCO\textsubscript{2} IAV relative to SeaFlux over the TPI region with the exception of CESM2, CESM2-FV2, CNRM-ESM2-1 and MIROC-ES2L, which have members with FCO\textsubscript{2} IAV amplitudes that overlap with SeaFlux.

FCO\textsubscript{2} is a function of surface ocean and atmospheric pCO\textsubscript{2}, and in the parameterization used in the models and data products, has a quadratic relationship to near-surface wind speeds, u\textsubscript{10} (Wanninkhof, 2014). To investigate the underestimation of FCO\textsubscript{2} seen in CMIP6, we assess their amplitudes of pCO\textsubscript{2} and u\textsubscript{10} IAV: σpCO\textsubscript{2} and σu\textsubscript{10}, respectively (see Figures 3b and 3c). Similar to the FCO\textsubscript{2} IAV estimates, the majority of CMIP6 models underestimate pCO\textsubscript{2} IAV relative to SeaFlux. Meanwhile, u\textsubscript{10} IAV is overestimated across the majority of models, with the exception of the CanESM5 models, the MPI models, and some smaller underestimation discrepancies from the GFDL-CM4 and MRI-ESM2-0 models, relative to three wind reanalysis data products. The underestimation in modeled pCO\textsubscript{2} IAV appears to be compensated by the overestimation in u\textsubscript{10} IAV. In the MPI models, pCO\textsubscript{2} IAV is within range of data products, but FCO\textsubscript{2} is low due to low u\textsubscript{10} IAV.

ENSO-driven variability has a concomitant effect on SST variability in the equatorial Pacific via the upwelling of cool waters. Figure 3d shows that the majority of CMIP6 models overestimate SST IAV in the TPI region, relative to ORAS5 and HadISST estimates. Models that underestimate pCO\textsubscript{2} IAV also overestimate SST IAV, with the exception of the CanESM5 models which underestimate both SST and pCO\textsubscript{2} IAV. Models also tend to overestimate u\textsubscript{10} variance (Figures 3c and 3d). This is consistent with the coupling of wind speeds and SST variability via the Bjerknes feedback, where they amplify each other’s anomalies.

Figure 2. Tropical Pacific pCO\textsubscript{2} multi-year means from 1990 to 2014 (units: μatm) from the SeaFlux ensemble average (top-left) and five Coupled Model Intercomparison Project Phase 6 (CMIP6) models (other panels). The box in the SeaFlux panel marks the Tropical Pacific Index (TPI) region. The number (r) on the top right of each model’s map is the spatial correlation coefficient between the model and SeaFlux in the TPI region. Model multi-year means are evaluated using a single ensemble member per model. Similar maps for all CMIP6 models are in Figure S2 in Supporting Information S1.
Figure 3. Comparison of IAV amplitudes in models (one standard deviation over the 1959–2014 period), and in observations-based data products (one standard deviation over the 1990–2014 period) in the Tropical Pacific Index region (5°N–5°S, 180°E–270°E). Top-left: $\text{FCO}_2$ IAV (units: mol C m$^{-2}$ yr$^{-1}$); top-right: $\text{pCO}_2$ IAV (units: $\mu$atm); bottom-left: $u_10$ IAV (units: m s$^{-1}$); bottom-right: Sea surface temperature IAV (units: °C). Boxplots represent the spread in IAV amplitudes within a model’s ensemble members. For models where fewer than three members were available, the spread is shown without a boxplot. Observations-based data products are represented as the filled circles and the gray shaded regions indicate the range of IAV amplitudes within the observations-based data products.
3.3. Spatial Patterns of pCO$_2$ IAV

The EOF1 of SeaFlux (Figure 4: top-left) explains 41% of the total variance in pCO$_2$ IAV in the tropical Pacific, with a pattern that resembles that of ENSO variability of FCO$_2$ (McKinley et al., 2004; Resplandy et al., 2015). Its corresponding first PC1 is highly correlated with ORAS5 SST anomalies in the TPI region ($r = -0.82$, see Figure S3 in Supporting Information S1 for PC1 results), which indicates ENSO-driven variability in the tropical Pacific Ocean in the observations-based products.

In CMIP6, few models have an EOF1 that resembles the ENSO pattern seen in SeaFlux (Figure 4; Figure S4 in Supporting Information S1: all models). Further, models that have a realistic spatial pattern have too little variance in the first EOF mode. For example, MIROC-ES2L has an EOF1 pattern most similar to SeaFlux (SCC = 0.79), and explains 30% of the total pCO$_2$ variance. MIROC-ES2L, CNRM-ESM2-1 and UKESM1-0-LL reveal almost two-centers of action—near the coastlines on either side of the tropical Pacific—for pCO$_2$ variance. The weak correlation over the TPI region between SeaFlux and CNRM-ESM2-1 (SCC = -0.04) is because the positive pCO$_2$ variance in the model's EOF1 is shifted slightly south of the equator. MPI models show a “negative” EOF1 pattern, revealing pCO$_2$ variability that is opposite to what is expected from ENSO variability—that is, the pCO$_2$ and SST variability in its TPI region are positively correlated, in contrast to the negative correlation in SeaFlux.

Models that reproduce a realistic multiyear mean pCO$_2$ map (Figure 2), with respect to SeaFlux, do not necessarily have a realistic ENSO pattern of variability (Figure 4). Nevertheless, the relationship between PC1 and TPI SST anomalies do tend to be strong, with a median correlation of $r = -0.73$ (Figure S3 in Supporting Information S1). This is consistent with the ENSO signal where upwelling dominates pCO$_2$ variability (Feely et al., 2006; Sutton et al., 2014).

Figure 5 compares maps of the local correlation coefficient between pCO$_2$ and SST anomalies in models for the tropical Pacific. These correlations reveal the relative magnitude of pCO$_2$$_{2,T}$ and pCO$_2$$_{2,nonT}$ components of pCO$_2$ variability, since the dominance of either component will result in a correlation coefficient that is either positive (thermally dominant) or negative (non-thermally dominant). The strong, negative correlation pattern (blue areas) over the equatorial Pacific, seen in SeaFlux (Figure 5: top-left), indicates variability in upwelling of water that is both cool and DIC-rich with ENSO oscillations. Areas of positive correlations (red areas) indicate pCO$_2$ variability that is thermally driven; warmer SSTs drive higher pCO$_2$ levels. The negative pCO$_2$-SST relationship covers
a broad region in SeaFlux that spans the basin, with the strongest negative correlations at the equator. Compared to SeaFlux, MIROC-ES2L shows a pattern that covers a similar longitudinal span, however, the intensity of the negative correlations are not as strong, and does not extend as far north. MRI-ESM2-0 shows stronger correlations; however, its negative pattern does not cover the same longitudinal span as seen in SeaFlux. The lack of the negative pCO$_2$-to-SST extension to the west, common to most of the CMIP6 models, indicates that the ENSO-CO$_2$ co-variability lies more east in models than in SeaFlux. CNRM-ESM2-1, UKESM1-0-LL, MRI-ESM2-0 and MPI-ESM-1-2-HAM have a positive correlation zone within the Niño 3.4 region; CESM2 also has an anomalous positive correlation zone that lies more toward the east (Figure S5 in Supporting Information S1).

### 3.4. Thermal and Non-Thermal pCO$_2$ IAV

For SeaFlux and the CMIP6 models, detrended pCO$_2$ monthly anomalies decomposed into thermally (pCO$_2,T$) and non-thermally (pCO$_2,nonT$) driven anomalies (Equation 1 and 2) indicate the relative magnitudes of thermally and non-thermally driven pCO$_2$ variability (Figure 6: IPSL-CM6A-LR; S6: other models).

In SeaFlux, pCO$_2,T$ (pCO$_2,nonT$) anomalies are strongly, positively (negatively) correlated with SST anomalies, with correlation coefficients greater than 0.98. The total pCO$_2$ anomaly (Figure 6, bold black line) is strongly negatively correlated ($r = -0.92$) with TPI SST anomalies, due to the non-thermal component being dominant over the thermal component ($\sigma_{pCO_2,nonT} > \sigma_{pCO_2,T}$).

In contrast, in IPSL-CM6A-LR (Figure 6: right), the non-thermal and thermal components have similar amplitudes but opposite sign ($\sigma_{pCO_2,nonT} \sim \sigma_{pCO_2,T}$). This results in the total pCO$_2$ anomaly having almost no correlation ($r = -0.03$) with SST anomalies. pCO$_2,T$ variability almost completely counteracts pCO$_2,nonT$ variability, resulting in a weak total pCO$_2$ anomaly in IPSL-CM6A-LR. pCO$_2$ components in other CMIP6 models are also plotted (Figure S6 in Supporting Information S1). A summary plot of the relative amplitudes of the thermal and non-thermal components is shown in Figure 7.

Figure 7a compares the amplitudes of pCO$_2,T$ and pCO$_2,nonT$ anomalies across CMIP6 models’ ensemble means. $\sigma_{pCO_2,T}$ is 14.4 μatm for SeaFlux-ORAS5, and modeled values range from 11.5 to 25.3 μatm, with the multi-model
Figure 6. Timeseries of thermal (pCO$_2$, red), non-thermal (pCO$_2$,nonT, blue) and total pCO$_2$ anomalies (bold black) from an ensemble average of the SeaFlux products (left) and from a single member of IPSL-CM6A-LR (right); units: μatm. Thin black lines represent Tropical Pacific Index (TPI) sea surface temperature (SST) anomalies (units: °C; the SST y-axis is located on the right-hand-side). The top-left numbers in each panel are the IAV amplitudes (σ) of pCO$_2$,T, pCO$_2$,nonT, and total pCO$_2$ anomalies; the bottom-right numbers are their correlation coefficients (r) with TPI SST anomalies.

Figure 7. (a) Amplitudes of pCO$_2$,T IAV (x-axis) versus pCO$_2$,nonT IAV (y-axis) averaged over the Tropical Pacific Index region (units: μatm). Model ensemble means are represented by the filled circles, while the unfilled diamond represents the observations-based data products. Box plots around the figure show the distribution among models for pCO$_2$,T and pCO$_2$,nonT IAV amplitudes. (b) Ratios of pCO$_2$,nonT/pCO$_2$,T IAV amplitudes in models (circles) and in the observations-based data products (diamond). Each scatter point represents the ensemble average for models and SeaFlux. The overlaid rectangle is magnified to see the models better.
median variance slightly higher than SeaFlux-ORAS5. On the other hand, $\sigma_{pCO_{2,nonT}}$ for SeaFlux-ORAS5 is 23.5 μatm, while modeled $\sigma_{pCO_{2,nonT}}$ ranges from 5.40 to 31.1 μatm with a multi-model median variance lower than that of SeaFlux-ORAS5. Figure 7b compares the ratios of $\sigma_{pCO_{2,nonT}}$ to $\sigma_{pCO_{2,T}}$ in models against the ratio found in the observations-based data products; SeaFlux-ORAS5 has a ratio of 1.63, while the models all have smaller ratios, ranging from 1.44 (ACCESS-ESM1-5) to 0.44 (MPI-ESM1-2-HR). As such, compared to SeaFlux, modeled $\sigma_{pCO_{2,nonT}}$ variability are not appropriately balanced against $\sigma_{pCO_{2,T}}$. Models with a more dominant non-thermal component, that is, $\sigma_{pCO_{2,nonT}}$, $\sigma_{pCO_{2,T}}$ ratios closer to SeaFlux-ORAS5, have total $pCO_{2}$ anomalies that are more negatively correlated with TPI SST anomalies (Figure S6 in Supporting Information S1).

3.5. pCO$_2$ Correlations With Other Variables

We evaluate the co-variability of ENSO-related variables with $pCO_2$ in order to better understand the controls on $pCO_{2,T}$ and $pCO_{2,nonT}$ in models versus observations-based data products. Reduced upwelling brings less cool, DIC-rich water to the surface, resulting in warmer SSTs and reduced surface ocean $pCO_2$. The winds that drive upwelling also force $z_{therm}$ anomalies; $z_{therm}$ anomalies are positive (deeper) in the TPI region when the trades relax and upwelling weakens.

Correlations of SST, $z_{therm}$, and $w_{50}$ anomalies with $pCO_2$ anomalies in the TPI region for SeaFlux are consistent with ENSO-driven variability as described above (Figure 8, clear diamonds; Figure S7 in Supporting Information S1), indicating that the observations-based products have realistic relationships between these variables and $pCO_2$, in particular with SST. For CMIP6, there is a large spread in correlations with $pCO_2$ (Figures 8a and 8b). NorESM2-MM and MIROC-ES2L have correlations similar to those seen in the observations-based data products. Models with incorrect correlation signs imply a lack of realistic relationships between these physical variables and $pCO_2$. For example, IPSL-CM6A-LR, and the MPI models have incorrect correlation signs between $pCO_2$ and the variables considered here.

Models that have more realistic correlation coefficients between $pCO_2$ and $w_{50}$ or $z_{therm}$ anomalies tend to do better at capturing the negative correlation between $pCO_2$ and SST anomalies (Figures 8a and 8b). The same models are better at representing $pCO_{2,nonT}$ variability (Figures 8c and 8d). Models with the weakest $pCO_{2,nonT}$ variances tend to be the same models with (a) weak or wrong-sign correlations (Figures 8a–8d), (b) lower non-thermal/thermal variance ratios (Figure 7; Figures 8c–8f), and (c) tend to do poorly in other areas throughout this assessment (Figures 2, 4, and 5). We leave out models that have incorrect correlation signs (negative) for $pCO_2$ and $w_{50}$ anomalies (Figures 8b, 8d, and 8f) when looking at the vertical transport of DIC, since these models do not have realistic $pCO_2$-upwelling relationships (Table 1).

3.6. Vertical Ocean Transport of DIC

Despite having higher $pCO_{2,nonT}$ variances in the better performing models, the balances between the non-thermal and thermal components of $pCO_2$ variability are still not the same as seen in SeaFlux-ORAS5 (Figure 7b). The balance between these components are such that for a given magnitude of $pCO_{2,T}$ IAV, the relative magnitudes of $pCO_{2,nonT}$ IAV in models are insufficient to overwhelm it and produce the total $pCO_2$ IAV seen in SeaFlux. This motivates the rest of this assessment where we take a closer look at the vertical transport of DIC and its contribution to $pCO_{2,nonT}$ variability.

In Figure 9a, timeseries for each term in the Reynolds’ decomposition (Equation 11) of the vertical transport of DIC in a single model (CESM2) are plotted against the time-tendency of $pCO_{2,nonT}$ ($\partial_t pCO_{2,nonT}$). Figure 9b shows what can be obtained from data, which is just the second Reynolds term involving the climatological vertical DIC gradient and variable upwelling. With Reynolds’ decomposition, we are able to isolate in models the contributions from variability in the vertical DIC gradient (Figure 9a: first panel) and the contributions from upwelling variability (Figure 9a: second panel) to $\partial_t pCO_{2,nonT}$. The non-linear term (Figure 9a: third panel) is small. The fourth panel in Figure 9a compares the total anomaly of the vertical ocean transport of DIC against $\partial_t pCO_{2,nonT}$. In CESM2, the first two Reynolds terms are roughly the same in magnitude, with standard deviations 2.65 and 2.77 times larger than the standard deviation of $\partial_t pCO_{2,nonT}$. The non-linear term is approximately the same magnitude as $\partial_t pCO_{2,nonT}$. The total anomaly (Figure 9a: fourth panel) has a standard deviation five times larger than the standard deviation of $\partial_t pCO_{2,nonT}$, and has a positive correlation of $r = 0.70$. The magnitude of
the total anomaly in vertical DIC transport means that it is important to $pCO_2$, $\text{nonT}$ variability, and also that there must be strong damping terms. A summary of the Reynolds’ terms in other models is in Table S2 in Supporting Information S1. Other models have similar results as CESM2 in that the total anomaly of vertical transport of DIC is significant in magnitude relative to the magnitude of $pCO_2$, $\text{nonT}$ variability. Values of their relative magnitudes,

**Figure 8.** Scatter plot of the correlation coefficients of $pCO_2$-to-$z_{\text{therm}}$ against (a) the correlation coefficients of $pCO_2$-to-sea surface temperature', (c) $pCO_2$, $\text{nonT}$ variances, and (e) non-thermal:thermal variance ratios. Correlation coefficients are calculated over the Tropical Pacific Index region for the observations-based data products (clear diamonds), and the Coupled Model Intercomparison Project Phase 6 models (filled circles). The model correlation coefficients shown are ensemble means. The gray shading indicates the 95% confidence threshold for the correlations. Right-hand-side (b, d, f) is the same as the left-hand-side (a, c, e), but with $pCO_2$-to-$w_{50}$ correlation coefficients along the x-axis instead.
σ ratio, range from 2.94 to 5.55 (Table S2 in Supporting Information S1), which together with strong correlations, means that variability in the vertical transport of DIC is an important source of pCO$_2$, nonT variability. Across the models, the first two Reynolds' terms, $\overline{\partial w \frac{\partial \text{DIC}'}{\partial z}}$ and $\overline{\partial w' \frac{\partial \text{DIC}'}{\partial z}}$, are the largest terms (Table S2 in Supporting Information S1), which suggests that the variability in both upwelling and vertical DIC gradients are similarly important to pCO$_2$, nonT variability. In MIROC-ES2L, the non-linear term is almost the same amplitude as the first two terms. For observations-based data products, the second Reynolds term $\overline{(w' \frac{\partial \text{DIC}'}{\partial z})}$ has a standard deviation four times bigger than the standard deviation of the observations-based $\overline{\partial \text{pCO}_2, \text{nonT}}$, (Figure 9b). Compared to the observations-based data products, the $\overline{\partial w' \frac{\partial \text{DIC}'}{\partial z}}$ term is weak in models (Table S2 in Supporting Information S1, second column), except for UKESM1-0-LL. This could be due to either a weak vertical gradient of climatological DIC, or weak upwelling variability, or a combination of both.

A time-averaged vertical velocity section from ORAS5 (Figure S8a in Supporting Information S1) reveals that the depth at which upwelling occurs is within the upper 100 m, with a maxima between 50 and 75 m at 220°E. We compare upwelling variability in models versus ORAS5 in Figure 10a at 50 m. We find that the range of upwelling variability across models is comparable and inclusive of the upwelling variability seen in ORAS5. In contrast, Figure 10b compares the vertical gradient of climatological DIC at 50 m to GLODAPv2. All the models have weaker gradients. We repeat this comparison at 80 m (Figure S9 in Supporting Information S1) and confirm that it is robust. Modeled vertical gradients of climatological DIC are biased weak, causing the second Reynolds term $\overline{(w' \frac{\partial \text{DIC}'}{\partial z})}$ in models to be weaker than the observations-based estimate (Figure 9; Table S2 in Supporting Information S1). To summarize, the second Reynolds's term $\overline{(w' \frac{\partial \text{DIC}'}{\partial z})}$ is an important term in the overall varia-

Figure 9. (a) Timeseries of the first (blue), second (orange) and third (green) Reynolds' terms from Equation 11, and the full variability is shown in the bottom panel (red) for one member from CESM2 (units: μatm s$^{-1}$). The time-tendency of pCO$_2$, nonT is shown in every panel (black line). $\sigma_{\text{ratio}} = \sigma(\text{Reynolds' term}) / \sigma(\overline{\partial \text{pCO}_2, \text{nonT}})$ is annotated in every panel. The correlation coefficient ($r$) between the timeseries are also shown. For the other models, a summary of this information can be found in Table S2 in Supporting Information S1. (b) timeseries of the second Reynolds' term computed from observations-based data products.
bility of the vertical transport of DIC, which is important to the variability in pCO$_2$, nonT, and thus pCO$_2$ variability. Underestimations in $\Delta w'_50 \frac{\partial}{\partial z}$DIC may result in an underestimation in pCO$_2$ variability. Alongside modeled mean vertical DIC gradients, we plot the mean vertical temperature gradients $\Delta (\frac{\partial}{\partial z} T)$ at 50 m depth to compare the relative strengths of gradients in models, and to identify model biases from observations-based data products (Figure 11a). Vertical temperature gradients are negative since ocean temperatures decrease with depth. The spread in strengths of modeled temperature gradients encompasses that seen in ORAS5, though the majority of models have weaker temperature gradients. The percentage difference between ORAS5 and the models’ median temperature gradient is about 21%. For the vertical gradient of climatological DIC, all models underestimate it compared to GLODAPv2, and the multi-model median has a percentage difference of about 39%. While the models tend to underestimate both the vertical gradients of climatological DIC and temperature, the climatological DIC gradients are more weakly biased, which for a given upwelling will tend to result in weaker pCO$_2$, nonT variability relative to pCO$_2$, T variability. The contributions of $\Delta w'_50 \frac{\partial}{\partial z} T$ and $\Delta w'_50 \frac{\partial}{\partial z}$DIC in units of pCO$_2$ change ($\mu$atm) can be computed using the coefficients from Equation 7. Contributions from $\Delta w'_50 \frac{\partial}{\partial z}$DIC to pCO$_2$ in models are about 6 times greater than the thermal contributions (Figure 11b). The vertical DIC term is much bigger than the vertical T term, but the associated pCO$_2$, nonT variability is not proportionally bigger than pCO$_2$, T variability (Figure 7). Thus, weak vertical gradients cannot fully explain the pCO$_2$, nonT-pCO$_2$, T differences. Monthly pCO$_2$ variability ($\sigma$ pCO$_2$) in the TPI region in SeaFlux range from 8.9 to 13.6 $\mu$atm, and values in models range from 3.7 to 12.9 $\mu$atm (Figure 3). These values of monthly TPI pCO$_2$ variability are on the same order of magnitude as the $\Delta w'_50 \frac{\partial}{\partial z}$DIC contributions to pCO$_2$ (Figure 11b: y-axis). In observations, and in models, $\Delta w'_50 \frac{\partial}{\partial z}$DIC contributions to pCO$_2$ are larger than their monthly pCO$_2$ variability in the TPI region. Despite the large contributions from $\Delta w'_50 \frac{\partial}{\partial z}$DIC to pCO$_2$ variability, significant damping must be happening to reduce overall pCO$_2$ variability and cause its underestimation.

4. Discussion

The majority (15) of the 18 CMIP6 models underestimate pCO$_2$ IAV, while they overestimate SST IAV. FCO$_2$ IAV is also underestimated by the majority of CMIP6 models. Previous studies of historical simulations from the earlier CMIP5 found that FCO$_2$ IAV were also underestimated in models (Dong et al., 2016, 2017). Results from
an another CMIP6 study also find that most models simulate weak FCO$_2$ anomalies while overestimating SST IAV (Vaitinada Ayar et al., 2022).

We find that the correlations between pCO$_2$ and other ENSO-related variables vary. Most models have correlations weaker than observed over the TPI region, a few are consistent with observations, and another few are opposite to observed. Weak ENSO-driven relationships were also noted in previous CMIP5 studies (Dong et al., 2017; Jin et al., 2019). Dong et al. (2016) also found that 12 models out of the 18 CMIP5 subset failed to show ENSO characteristics in FCO$_2$ variability. They also found that models differed among themselves the most in regions with strong vertical movement, such as the tropical Pacific.

Modeled pCO$_2_{2000}$ variance in CMIP6 is not appropriately balanced with pCO$_2_{5T}$ variability. Weak pCO$_2_{2000}$ anomalies are insufficient to counteract the pCO$_2_{5T}$ anomalies resulting in total pCO$_2$ anomalies that are too weak. In the equatorial Pacific, Jin et al. (2019) found pCO$_2$ biases in two CMIP5 models that resulted from weak DIC contributions to pCO$_2$. Weak DIC contributions were found to be mainly caused by weak vertical gradients of climatological DIC and weak upwelling anomalies, which both limit the vertical transport of DIC (Jin et al., 2019). We find that upwelling anomalies in CMIP6 are comparable to ORAS5 (Figure 11b). Together, these findings indicate that weak DIC contributions to pCO$_2$ variability are significantly damped by other processes.

Changes in the vertical transport of DIC affects surface DIC variability, which is known to be the dominant driver of pCO$_2$ variability in the surface equatorial Pacific Ocean (Liao et al., 2020; McKinley et al., 2004). We find model pCO$_2$ anomalies due to variability in the vertical transport of DIC are larger than their pCO$_2_{2000}$ anomalies by a factor of 3 to almost 6 times (see Table S2 in Supporting Information S1: last column) but are positively correlated. This suggests that variability in the vertical transport of DIC is an important source of pCO$_2_{2000}$ variability in models. At the same time, $w_{50}n_5_3$ contributions to pCO$_2$ are comparable in magnitude to daily pCO$_2$ variability in the TPI region (Figure 11b). Together, these findings indicate that $w_{50}n_5_3$ contributions to pCO$_2$ variability are significantly damped by other processes.

Figure 11. (a) The relative strength of vertical mean temperature gradients (x-axis, units: °C m$^{-1}$) against vertical climatological dissolved inorganic carbon (DIC) gradients (y-axis, units: mmol m$^{-4}$) in models (filled circles represent a single ensemble member) and in ORAS5 versus GLODAPv2 data (clear diamond). The boxplots represent the distribution in gradients among models, excluding observations-based data products (clear diamonds). (b) The vertical transport of the climatological vertical temperature gradient versus the DIC gradient, converted from units of μatm s$^{-1}$ into units of μatm month$^{-1}$.
The vertical gradient of climatological DIC is consistently weak across all the models relative to observations-based data products (Figure 10), which is consistent with prior model results from CMIP5 (Jin et al., 2019). Vertical gradients of climatological temperature are not as weak. The imbalance in the relative strengths of these vertical gradients, for a given upwelling anomaly, contributes toward weaker non-thermal pCO$_2$ variability, relative to the thermal.

While the relative strengths of mean vertical gradients, through upwelling, can result in weaker pCO$_{2, nonT}$: pCO$_{2,T}$ ratios, we do not find a linear scaling between the relative strengths in mean vertical gradients and the ratios of pCO$_{2, nonT}$: pCO$_{2,T}$ across the models (Figure S10 in Supporting Information S1). A linear scaling would indicate that biases in the relative strengths of the mean vertical gradients proportionally bias the pCO$_2$ ratios. Thus, we find the relative strengths of mean vertical gradients alone do not determine the imbalance in pCO$_2$ ratios. A more complete assessment that includes the other processes that contribute to pCO$_2$ variability will be necessary to understand the causes of insufficient pCO$_{2, nonT}$ variability.

Other processes that contribute to equatorial Pacific DIC variability that can dampen pCO$_{2, nonT}$ variability, include the horizontal transport of DIC, biological processes, FW and air-sea CO$_2$ fluxes. For example, when DIC is brought to the surface via upwelling, though pCO$_2$ increases, the instantaneous air-sea CO$_2$ flux response dampens surface DIC concentrations (Liao et al., 2020). The biological response also damps surface DIC concentrations; upwelling of nutrient-rich waters enhances biologically-driven uptake of DIC (Chavez et al., 1999). Freshwater fluxes (rainfall) also dilute surface DIC concentrations, and westward horizontal transport along the equator removes DIC from the upwelling region (Doney et al., 2009).

Aside from DIC, other ocean biogeochemical variables influence surface pCO$_{2, nonT}$, such as alkalinity. We repeat our analysis for the vertical ocean transport of DIC with alkalinity data to evaluate how much surface alkalinity variability dampens pCO$_2$ variability. We find that the contributions from $w'_z \partial_z$Alk dampens $w'_z \partial_z$DIC contributions to pCO$_2$ anomalies by roughly 10% across models and observations-based data products (Figure S11 in Supporting Information S1). However, this amount of damping is not enough to explain the insufficient pCO$_{2, nonT}$ variability. Vaittinada Ayar et al. (2022) find that models with strong alkalinity biases have weak surface DIC biases (i.e., weak surface DIC variability), which leads to a reduction in pCO$_{2, nonT}$ variability. They find that for some models (CanESM5, GFDL-CM4 and MRI-ESM2-0), pCO$_{2, nonT}$ variability is weak enough that pCO$_2$ variability can dominate total pCO$_2$ anomalies. However, an alkalinity bias alone does not explain all the models that underestimate pCO$_{2, nonT}$ relative to pCO$_{2,T}$, as we analyze here. For example, Vaittinada Ayar et al. (2022) shows that IPSL-CM6A-LR doesn’t have a strong alkalinity bias, however, we find that its pCO$_{2, nonT}$: pCO$_{2,T}$ variance ratio is weaker than the ratio in MRI-ESM2-0 (Figure 7b), which is a model they show with a strong alkalinity bias.

Vaittinada Ayar et al. (2022) proposed that models without a strong alkalinity bias may be better predictors of future ENSO-CO$_2$ flux dynamics. However, we find that these models underestimate equatorial Pacific pCO$_2$ IAV and ENSO-related covariability. For example, IPSL-CM6A-LR did not have realistic correlations between pCO$_2$ and SST, $\gamma_{nonT}$, $\gamma_{T}$ anomalies (Figure 8b). We propose that a wide range of variables need to be considered when selecting models for analysis of future trends. While this study looks at ENSO-driven pCO$_2$ IAV, it has relevance for trends. Trends in SSTs, thermocline depths and upwelling in response to rising atmospheric CO$_2$ involve many of the same coupled dynamics that drive ENSO variability (Cane et al., 1997; Clement et al., 1996; Seager et al., 2019). CMIP6 models cannot reproduce the observed trends in the tropical Pacific physical state and hence it is possible that they are also misrepresenting the trends in pCO$_2$ and air-sea CO$_2$ fluxes, with potential influence on the airborne fraction of anthropogenic CO$_2$. Validating ENSO-driven pCO$_2$ variability in models is a necessary first step to examining the tropical Pacific’s coupled climate-carbon response to anthropogenic climate change.

5. Conclusions

In the equatorial Pacific, weak ENSO-related pCO$_2$ variability in CMIP6 models is explained by an imbalance between pCO$_{2, nonT}$ and pCO$_{2,T}$ anomalies, whereby pCO$_{2, nonT}$ variability is insufficient to counteract strong pCO$_{2,T}$ variability. Strong pCO$_2$ variability in CMIP6 is driven by excessive SST variance. Variability in the vertical transport of DIC does matter to pCO$_{2, nonT}$ variability in that upwelling anomalies acting on weak vertical DIC gradients can lead to weaker surface DIC variability. However, this alone does not explain the relative magnitudes of pCO$_{2, nonT}$ and pCO$_{2,T}$ anomalies. To guide model development, assessments of other processes that drive DIC
variability will help to identify the causes of significant damping of $pCO_{2,air}$ variability that ultimately leads to weak $pCO_{2}$ variability in models.

Conflict of Interest
The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

References
