# Leveraging the signature of heterotrophic respiration on atmospheric $CO_2$ for model benchmarking.

Samantha J. Basile[1], Xin Lin[1] William R. Wieder[2,3], Melannie D. Hartman [2,4], Gretchen Keppel-Aleks[1]

[1] Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, 48105, USA
[2] Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, 80305, USA
[3] Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, 80309, USA
[4] Natural Resource Ecology Laboratory, Colorado State University, Fort Collins CO, 80523, USA

*Correspondence to:* Samantha Basile (sjbasile@umich.edu)

**Abstract**

Spatial and temporal variations in atmospheric carbon dioxide ($CO_2$) reflect large-scale net carbon exchange between the atmosphere and terrestrial ecosystems. Soil heterotrophic respiration (HR) is one of the component fluxes that drive this net exchange but, given observational limitations, it is difficult to quantify this flux or to evaluate global-scale model simulations thereof. Here, we show that atmospheric $CO_2$ can provide a useful constraint on large-scale patterns of soil heterotrophic respiration. We analyze three soil model configurations (CASA-CNP, MIMICS and CORPSE) that simulate HR fluxes within a biogeochemical testbed that provides each model with identical net primary productivity (NPP) and climate forcings. We subsequently quantify the effects of variation in simulated terrestrial carbon fluxes (NPP and HR from the three soil testbed models) on atmospheric $CO_2$ distributions using a three-dimensional atmospheric tracer transport model. Our results show that atmospheric $CO_2$ observations can be used to identify deficiencies in model simulations of the seasonal cycle and interannual variability in HR relative to NPP. In particular, the two models that explicitly simulated microbial processes (MIMICS and CORPSE) were more variable than observations at interannual timescales and showed a stronger than observed temperature sensitivity. Our results prompt future research directions to use atmospheric $CO_2$, in combination with additional constraints on terrestrial productivity or soil carbon stocks, for evaluating HR fluxes.

## 1. Introduction

Atmospheric $CO_2$ observations reflect net exchange of carbon between the land and oceans with the atmosphere. Observations of atmospheric $CO_2$ concentration have been collected *in situ* since the late 1950s (Keeling et al., 2011), and global satellite observations have become available within the last decade (Crisp et al., 2017; Yokota et al., 2009). The high precision and accuracy of in situ observations and the fact that these measurements integrate information about ecosystem carbon fluxes over a large spatial footprint make atmospheric $CO_2$ a strong constraint on model predictions of net carbon exchange (Keppel-Aleks et al., 2013). For example, at seasonal timescales, atmospheric $CO_2$ can be used to evaluate the growing-season net flux, especially in the Northern Hemisphere (Yang et al., 2007). At interannual timescales, variations in the atmospheric $CO_2$ growth rate are primarily driven by changes in terrestrial carbon fluxes in response to climate variability (Cox et al., 2013; Humphrey et al., 2018; Keppel-Aleks et al., 2014). Recent studies have hypothesized that soil carbon processes represent one of the key processes in driving these interannual variations (Cox et al., 2013; Wunch et al, 2013). Moreover, soil carbon processes represent one of the largest uncertainties in predicting future carbon-climate feedbacks, in part because non-permafrost soils contain an estimated 1500 to 2400 PgC (Bruhwiler et al., 2018), at least a factor of three larger than the pre-industrial atmospheric carbon reservoir.

Soil heterotrophic respiration (HR), the combination of litter decay and microbial breakdown of organic matter, is the main pathway for $CO_2$ release from soil carbon pools to the atmosphere. Currently, insights on HR rates and controls are mostly derived from local-scale observations. Ecosystem respiration, or the combination of autotrophic and heterotrophic respiration fluxes, can be backed out from eddy covariance net ecosystem exchange observations at spatial scales around 1 $km^2$, but with substantial uncertainty (Baldocchi 2008; Barba et al., 2018; Lavigne et al., 1997). The bulk of ecosystem respiration fluxes comes from soils, but soil respiration fluxes from chamber measurements can exceed ecosystem respiration measurements from flux towers, highlighting uncertainties in integrating spatial and temporal variability in ecosystem and soil respiration measurements (Barba et al. 2018). Further partitioning of soil respiration measurements into autotrophic and heterotrophic components to derive their appropriate environmental sensitivities remains challenging, but critical to determining net ecosystem exchange of $CO_2$ with the atmosphere (Bond-Lamberty et al., 2004, 2011, 2018). Additionally, because fine-scale variations in environmental drivers such as soil type and soil

moisture affect rates of soil respiration, it is difficult to scale local respiration observations to regional or global levels (Zhao et al., 2017). Specifically, soil heterotrophic respiration (HR), the combination of litter decay and microbial breakdown of organic matter, is the main pathway for $CO_2$ release from soil carbon pools to the atmosphere. Currently, insights on HR rates and controls are mostly derived from local-scale observations. Soil chamber observations can be used to measure soil respiration at spatial scales on the order of 100 $cm^2$ (Davidson et al., 2002; Pumpanen et al., 2004; Ryan and Law, 2005).

Local-scale observations reveal that HR is sensitive to numerous climate drivers, including temperature, moisture, and freeze-thaw state (Baldocchi, 2008; Barba et al., 2018; Lavigne et al., 1997). Because of these links to climate, predicting the evolution of HR and soil carbon stocks within coupled Earth system models is necessary for climate predictions. Within prognostic models, heterotrophic respiration has been represented as a first-order decay process based on precipitation, temperature, and a linear relationship with available substrate (Jenkinson et al., 1990; Parton, 1996, Randerson et al., 1996). However, such representations may neglect key processes for the formation of soil and persistence of soil organic carbon (SOC) stocks (Lehmann and Kleber 2015; Schmidt et al., 2011; Rasmussen et al., 2018). More recently, models have begun to explicitly represent microbial processes into global-scale simulations of the formation and turnover of litter and SOC (Sulman et al., 2014; Wieder et al., 2013) as well as to evaluate microbial trait-based signatures on SOC dynamics (Wieder et al., 2015). These advances in the representation of SOC formation and turnover increase capacities to test emerging ideas about soil C persistence and vulnerabilities, but also increase the uncertainties in how to implement and parameterize these theories in models (Bradford et al., 2016; Sulman et al., 2018; Wieder et al., 2018).

Given these uncertainties, developing methods to benchmark model representations of HR fluxes is an important research goal (Bond-Lamberty et al., 2018) as model predictions for soil carbon change over the 21$^{st}$ century are highly uncertain (Schuur et al., 2018; Todd-Brown et al., 2014). A common method for model evaluation is to directly compare spatial or temporal variations in model properties (e.g., leaf area index) or processes (e.g., gross primary productivity) against observations (Randerson, 2009; Turner et al., 2006). Such comparisons assess model fidelity under present day climate, but may not ensure future predictivity of the model. The use of functional response metrics, which evaluate the relationship between a model process and an underlying driver, may ensure that the model captures the sensitivities required to predict future

evolution (Collier et al., 2018, Keppel-Aleks et al., 2018). A third benchmarking approach is to use hypothesis-driven approaches or experimental manipulations to evaluate processes (Medlyn et al., 2015). It is likely that these methods will have maximum utility when combined within a benchmarking framework (e.g., Collier et al., 2018; Hoffman et al., 2017) since they evaluate different aspects of model predictive capability.

Although a lack of direct respiration observations remains a gap for model evaluation, indirect proxies for respiration may be obtained from atmospheric $CO_2$, which reflects the balance of all carbon exchange processes between the atmosphere and biosphere. Previous work has shown that atmospheric $CO_2$ observations are inherently sensitive to HR across a range of timescales. For example, at seasonal timescales, improving the parameterization for litterfall in the CASA model improved phasing – i.e., the timing of seasonal maxima, minima, and inflection points – for the simulated atmospheric $CO_2$ annual cycle (Randerson et al., 1996). At interannual timescales, variations in the Northern Hemisphere $CO_2$ seasonal minimum are hypothesized to arise from variations in respiration (Wunch et al., 2013) and variations in the growth rate have been linked to tropical respiration and its temperature sensitivity (Anderegg et al., 2015). Here, we hypothesize that atmospheric $CO_2$ data can be used to evaluate simulations of soil heterotrophic respiration and differentiate between the chemical and microbial parameterizations used in state-of-the-art models. In this analysis, we simulate atmospheric $CO_2$ distributions using three different soil model representations that are part of a soil biogeochemical testbed (Wieder et al., 2018). The three sets of HR fluxes, shown in Wieder et al., (2018) to have distinct patterns at seasonal timescales, are used as boundary conditions for a 3-dimensional atmospheric transport model. We evaluate temporal variability in the resulting $CO_2$ simulations against observations, quantify the functional relationships between $CO_2$ variability and temperature variability, and quantify the regional influences of land carbon fluxes on global $CO_2$ variability. The methods and results are presented in section 2 and 3, and discussion of the implications for benchmarking and our understanding of drivers of atmospheric $CO_2$ variability are presented in section 4.

## 2. Data and Methods

We used a combined biosphere-atmosphere modeling approach to diagnose the signatures of land fluxes on atmospheric $CO_2$ (Fig. 1). At the heart of this approach is comparison of simulated

atmospheric $CO_2$ owing to individual processes and regions to atmospheric $CO_2$ observations. The observations and models used are described below.

### 2.1 Observations and timeseries analysis

For this analysis we use reference $CO_2$ measurements reported in parts per million (ppm) from 34 marine boundary layer sites (MBL, Table S1) within the NOAA Earth System Research Laboratory sampling network (ESRL, Fig. 2; Dlugokencky et al., 2016). These sites were chosen to minimize the influence of local anthropogenic emissions and had at least 50% data coverage over the 29-year period between 1982 and 2010. Following the approach in Keppel-Aleks et al., 2018, we aggregate site specific $CO_2$ by averaging measurement timeseries across six latitude zones (Fig. 2, solid lines): Northern Hemisphere high latitudes (NHL: 61 to 90°N), midlatitudes (NML: 24 to 60°N), tropics (NT: 1 to 23°N), Southern Hemisphere tropics (ST: 0 to 23°S), and two Southern extratropics bands: the southern midlatitudes (SML, 24-60 S), and the southern high latitudes (SHL, 61-90 S). The global mean $CO_2$ timeseries is constructed as an area-weighted average of these six atmospheric zones.

We detrend all timeseries data using a third-order polynomial fit to remove the impact of annually increasing atmospheric concentration in our seasonal and interannual calculations (SFig. 1). Using the detrended $CO_2$ data, we calculate a period median annual cycle by averaging all observations for a given calendar month. To calculate $CO_2$ interannual variability ($CO_2$ IAV), the median annual cycle is subtracted from the detrended timeseries (SFig. 1). The magnitude of $CO_2$ IAV is calculated as one standard deviation on the detrended, deseasonalized timeseries, unless otherwise noted. Model simulated $CO_2$ seasonality and interannual variability is calculated using the same methods.

### 2.2 Soil testbed representations of heterotrophic respiration

We used a soil biogeochemical testbed (Fig. 1; Wieder et al., 2018), which generates daily estimates of soil carbon stocks and fluxes at global scale without the computational burden of running a full land model. All testbed fluxes are output in grams of carbon per meter square at a daily temporal resolution and then converted to petagrams (Pg C, over a region). The testbed is a chain of model simulations where soil models with different structures can be run under the same forcing data, including the same gross primary productivity (GPP) fluxes, soil temperature, and

soil moisture. The testbed produces its own estimates of net primary production (NPP), the difference between GPP and autotrophic respiration (AR; Eqn. 1). Each testbed soil model in this analysis produces unique gridded heterotrophic respiration (HR) values based on its own underlying mechanism and soil C stocks. Currently, the testbed is run with a carbon-only configuration.

For the simulations described in this paper, the chain starts with the Community Land Model 4.5 (CLM4.5; Oleson et al., 2013), run with satellite phenology with CRU-NCEP climate reanalysis as forcing data (Jones et al., 2013; Kalnay et al., 1996; Le Quéré et al., 2018). In this simplified formulation of CLM, a single plant functional type is assumed in each 2° by 2° gridcell. Daily values for gross primary productivity (GPP), soil moisture, soil temperature, and air temperature from CLM4.5 are passed to the Carnegie-Ames Stanford Approach terrestrial model (CASA-CNP; Potter et al., 1993; Randerson et al., 1996; Randerson et al., 1997; Wang et al., 2010). The CASA-CNP plant model uses the data from CLM4.5 to calculate NPP and carbon allocation to roots, wood, and leaves. This module also determines the timing of litterfall. Finally, metabolic litter, structural litter, and decomposing coarse woody debris (CWD) are then passed to the soil biogeochemical models to simulate HR.

From the testbed output we calculate the net ecosystem productivity (NEP; Eqn. 3). In the analysis presented here, CASA NPP was used across the testbed ensemble in the NEP calculation, thus highlighting differences in the timing and magnitude of HR fluxes from the individual soil models. From a land perspective (positive NEP fluxes into land), NEP is calculated as NPP – HR, where respiration release of $CO_2$ decreases net carbon gains through photosynthesis. Here, we use an atmospheric perspective for NEP (positive NEP fluxes into the atmosphere) by reversing the sign on the NPP flux and taking HR as positive (Eqn. 3).

$$NPP = GPP - AR \quad (1)$$

$$NEP = HR + (-NPP) \quad (2)$$

The three soil models make distinct assumptions about microbial processes. More details regarding these formulations and their implementation in the testbed are found in Wieder et al. (2018), but we provide brief descriptions here. The CASA-CNP soil model computes first-order, linear decay rates modified by soil temperature and moisture, implicitly representing microbial

activity and soil carbon turnover through a cascade of organic matter pools (CASA: Randerson et al., 1997; CASA-CNP: CASA carbon cycling with additional nitrogen, and phosphorus cycling, Wang et al., 2010). These include metabolic and structural litter, as well as a fast, slow, and passive soil carbon pools. The Microbial-Mineralization Carbon Stabilization model (MIMICS; Wieder et al., 2014; Wieder et al., 2015) explicitly represents microbial activity with a temperature-sensitive reverse Michaelis-Menten kinetics (Buchkowski et al., 2017; Moorhead and Weintraub, 2018) but has no soil moisture controls. The decomposition pathway is set up with two litter pools (identical to those simulated by CASA-CNP), three soil organic matter pools (available, chemically and physically protected), and two microbial biomass pools for copiotrophic (fast) and oligotrophic (slow) microbial functional groups. The Carbon, Organisms, Rhizosphere, and Protection in the Soil Environment model (CORPSE) is also microbially explicit and uses reverse Michaelis-Menten kinetics, but it assumes different microbial and soil carbon pools. Surface litter and soil C pools are considered separately, but only soil C has a parallel set of physically protected pools that are isolated from microbial decomposition. CORPSE includes a temperature dependent Maximum Reaction Velocity ($V_{max}$) parameter, but also includes a term for the soil moisture controls on decomposition rates that uses volumetric liquid soil water content. For all three models, soil texture inputs were also derived from the CLM surface data set (Oleson et al., 2013). We acknowledge that one potential limitation of the approach is a lack of vertical resolution in terms of temperature or frozen fraction of soil moisture (Koven et al., 2013). Overall, while the testbed approach contains necessary simplifications, it provides the ability to query the role of model structure, including assumptions about the number of soil carbon pools, the role of microorganisms, and the sensitivity to environmental factors, in driving HR flux differences when NPP and environmental controls are held in common.

The testbed fluxes are used in two ways: first, we analyze monthly-averaged, regional fluxes for net primary production (NPP) from CASA-CNP and HR simulated by CASA-CNP, CORPSE and MIMICS. Second, we use the raw daily fluxes as boundary conditions for global GEOS-Chem runs to simulate the influence of these fluxes on atmospheric $CO_2$, as described in the following section.

### 2.3 GEOS-Chem atmospheric transport modeling of $CO_2$

We simulate the imprint of the testbed fluxes on atmospheric $CO_2$ using GEOS-Chem, a 3-D atmospheric transport model. We run the GEOS-Chem v12.0.0 $CO_2$ simulation between 1980 and 2010 at a resolution of 2.0° in latitude by 2.5° in longitude with 47 vertical levels. The model is driven by hourly meteorological data from the Modern-Era Retrospective analysis for Research and Application version 2 (MERRA2) reanalysis data (Gelaro et al., 2017; https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/), with the dynamic timestep set to be 600 seconds. The model is initialized with globally-uniform atmospheric $CO_2$ mole fraction equal to 350 ppm. The testbed fluxes from 1980 to 2010 are used for land emissions to simulate the imprint of these different soil model configurations on atmospheric $CO_2$ (Fig. 1). In our simulations, HR and NPP fluxes were separated into the five regions listed above (NHL, NML, NT, ST, SE) so that the influence of carbon fluxes originating from these individual regions on global atmospheric $CO_2$ mole fraction could be quantified. We initialized separate species of $CO_2$ in the atmospheric model, one for each flux (HR or NPP) and region (NHL, NML, etc.). Since we considered four fluxes (CASA NPP and three types of HR) originating in five regions, we simulated a total of 20 species. These species were tracked throughout the simulation as their spatiotemporal distribution changed due to the combined influence of $CO_2$ fluxes at the surface and atmospheric weather. Although these species are simulated individually, we can simply sum the regional atmospheric species for a given flux (e.g., CASA HR) to determine the atmospheric $CO_2$ arising from all fluxes over the globe. We also simulated the fossil and ocean imprint on atmospheric $CO_2$ using boundary conditions from $CO_2$ CAMS inversion 17r1 (https://atmosphere.copernicus.eu/sites/default/files/2018-10/CAMS73_2015SC3_D73.1.4.2-1979-2017-v1_201807_v1-1.pdf). However, at the temporal scales of this analysis, ocean and fossil fuel fluxes had a much smaller influence on regional patterns of atmospheric $CO_2$ than did land fluxes. Across the six latitude bands, the detrended $CO_2^{NEP}$ annual amplitude ranges from a factor of 1.5 (in the tropics) to an order of magnitude larger (at high latitudes) than $CO_2$ from ocean fluxes and fossil fuel emissions. Likewise, the IAV from fossil and ocean-derived $CO_2$ was at most 25% that of NEP-derived $CO_2$ at most latitude bands. These results are consistent with previous studies that have demonstrated that NEP drives most of the atmospheric $CO_2$ seasonality (> 90%; Nevison et al., 2008; Randerson et al., 1997) and interannual variability (e.g., Rayner et al., 2008; Battle et al., 2000). Given that patterns of IAV in ocean and fossil $CO_2$ partially cancel each other and the large uncertainty in ocean fluxes, we choose to omit these $CO_2$ species from our analysis.

We discard the first two years of the atmospheric simulations for model spin-up, and we analyze the monthly average model outputs for the period 1982–2010. We sample the gridded atmospheric simulation output at the 34 marine boundary layer (MBL) sites identified in section 2.1, using the 3rd vertical level to minimize influence of land-atmosphere boundary layer dynamics. We then calculate the latitude zone average, median annual cycle and interannual variability using the methods described for $CO_2$ observations (see section 2.1). Averaging $CO_2$ from all sites within a latitude band is consistent with our hypothesis that atmospheric $CO_2$ may provide constraints on large-scale patterns of heterotrophic respiration, but individual sites may be too heavily influenced by local characteristics not accounted for by the testbed fluxes. As such, averaging simulated and observed $CO_2$ across latitude zones smooths local information while retaining information about regional scale fluxes.

Throughout the manuscript, we refer to $CO_2$ originating from these NPP and HR component fluxes as $CO_2^{NPP}$ and $CO_2^{HR}$, respectively. We use a sign convention for the fluxes whereby a positive value indicates a source of carbon to the atmosphere, which means we can combine the $CO_2$ tracers from NPP and HR to calculate the expected atmospheric variation owing to NEP using (Eqn. 3):

$$CO_2^{NEP} = CO_2^{HR} + CO_2^{NPP} \qquad (3)$$

We note that the net $CO_2$ response from the model (i.e., $CO_2^{NEP}$) is approximately equivalent to observations in terms of seasonal and interannual variations, although we neglect ocean fluxes and emissions from fossil fuels, land use and land cover change, and disturbance. In the results below, the subscript notation will be used to denote the testbed ensemble sources. For example, $CO_2^{HR}$ simulated from CORPSE fluxes is defined as $CO_2^{\text{CORPSE HR}}$, similarly for $CO_2^{\text{CORPSE NEP}}$.

### 2.4 Global temperature sensitivity and separation of regional influences

For insight on a functional climate response, we investigate the global temperature sensitivity of the atmospheric $CO_2$ growth rate and the testbed ensemble fluxes. Rates of change were derived from monthly and annual timeseries to calculate the temperature sensitivity of the testbed fluxes, the modeled $CO_2$, and the observed $CO_2$ values. The $CO_2$ growth rate anomaly was calculated as the difference between timestep n and n-1 in both the monthly and annual $CO_2$ IAV

timeseries. As a result of this technique, the monthly $CO_2$ growth rate anomalies were centered on the first day of the corresponding months. To compare flux information with $CO_2$ growth rate anomalies, daily testbed flux timeseries were averaged to monthly resolution and then interpolated by averaging between months to center values on the first day of each month.

Following Arora et al. (2013), we calculate temperature sensitivity ($\gamma$) using an ordinary linear regression (OLR). We calculate OLR for the interannual variability timeseries of CASA-CNP soil temperature (T IAV) against 1) atmospheric $CO_2$ growth rate anomalies, and 2) land flux IAV (see section 2.2). For atmospheric $CO_2$ growth rate anomalies, each timeseries was converted from ppm $y^{-1}$ to PgC $y^{-1}$ based on the global mass of atmospheric dry air. Thus, all global temperature sensitivity values are reported in units of PgC $y^{-1}$ $K^{-1}$. The global temperature sensitivity value for the observed $CO_2$ growth rate anomaly was calculated for 1982 to 2010 using ESRL $CO_2$ observations and the Climatic Research Unit's gridded temperature product (CRU TS4; Jones et al., 2012), which is derived from interpolated ground station measurements.

We also assess the influence of individual regions on the global mean signal for both component land fluxes (NPP, HR) and simulated atmospheric $CO_2$ ($CO_2^{NPP}$, $CO_2^{HR}$, $CO_2^{NEP}$). We first quantify the magnitude of variability in each region relative to the magnitude of global variability ($\sigma_{REL}$) as the ratio of regional IAV standard deviation to global IAV standard deviation. This ratio is calculated for monthly flux IAV from each of the five flux regions and for the global-mean $CO_2$ timeseries that arises from fluxes in each of the five flux regions (e.g., the global $CO_2$ response to NHL fluxes, or the global $CO_2$ response to NML fluxes, etc.). The value of $\sigma_{REL}$ has a lower bound of 0, which would indicate that a region contributes no IAV, but has no upper bound, since a value greater than 1 simply indicates that the fluxes in a given region are more variable than global fluxes.

We note that the timing of IAV in a given region may be independent of IAV in other regions, and thus may or may not be temporally in-phase with global IAV. We therefore also calculate correlation coefficients ($r$) for the timeseries of regional flux IAV and $CO_2$ IAV with the global signal. Thus, if an individual region were responsible for all observed global flux or $CO_2$ variability, it would have both $\sigma_{REL}$ and $r$ values equal to 1 in this comparison. The value for $r$ will be small if a regional signal is not temporally coherent with the global signal, even if the magnitude of variability is high.

## 3. Results

### 3.1 Seasonal imprint of heterotrophic respiration

Our evaluation of $CO_2$ simulated using testbed fluxes revealed that all testbed models overestimated the mean annual cycle amplitude of atmospheric $CO_2$ observations. In the Northern Hemisphere, the bias was largest for MIMICS, as the $CO_2^{\text{MIMICS NEP}}$ amplitude was overestimated by up to 100% (Fig. 3). The mismatch was smallest in $CO_2^{\text{CORPSE NEP}}$, which was within 70% of the observed annual cycle amplitude where CORPSE simulates the largest seasonal HR fluxes (Fig. 3a-c, Table 1). Within the modeled carbon dioxide concentrations resulting from land fluxes, $CO_2^{\text{NPP}}$ and $CO_2^{\text{HR}}$, show largest seasonality in the NHL, with seasonal amplitudes decaying toward the tropics and Southern Hemisphere. In the NHL, the peak-to-trough amplitude of $CO_2^{\text{NPP}}$ is 39±2 ppm, with a seasonal maximum in April and a seasonal minimum in August (Fig. 4a; note, this $CO_2^{\text{NPP}}$ peak reflects the sign reversal in the driving NPP flux (section 2.3). The seasonal cycles for $CO_2^{\text{HR}}$ simulated from all testbed models are out of phase with that of $CO_2^{\text{NPP}}$, and there are large amplitude differences in $CO_2^{\text{HR}}$ among the model ensemble members. Specifically, the NHL amplitude of $CO_2^{\text{CORPSE HR}}$ is 28±3 ppm, while the amplitudes for $CO_2^{\text{MIMICS HR}}$ and $CO_2^{\text{CASA-CNP HR}}$ are only 17±1 ppm, accounting for about 40-70% of the amplitude from $CO_2^{\text{NPP}}$ (Table 1). However, in all latitude bands, the largest $CO_2^{\text{HR}}$ amplitude comes from the microbially explicit model – CORPSE for the Northern Hemisphere. In the Southern Hemisphere extratropics, the amplitudes for all components were less 3 ppm (Table 1).

The three soil carbon models in the testbed impart different fingerprints on atmospheric $CO_2$ variability. Specifically, the phasing of $CO_2^{\text{HR}}$ is an important driver of the overall comparison between $CO_2^{\text{NEP}}$ and observed $CO_2$ seasonality (Fig. 3). When the contributions of NPP and HR seasonality are considered together (i.e., $CO_2^{\text{HR}}$ + $CO_2^{\text{NPP}}$), the simulated amplitude of $CO_2^{\text{NEP}}$ is larger than the observed $CO_2$ across all latitude bands (Fig. 3). The largest mismatch is in the NHL zone, where the observed mean annual cycle is 15±0.9 ppm, while the peak-to-trough $CO_2^{\text{NEP}}$ ranges from 23±1.3 ppm for CORPSE to 33±1.4 ppm for MIMICS (Fig. 3a). The smaller $CO_2^{\text{NEP}}$ amplitude simulated by CORPSE is due to the large $CO_2^{\text{HR}}$ seasonality that counteracts the seasonality in NPP (Fig. 4a-b). Furthermore, $CO_2^{\text{MIMICS HR}}$ and $CO_2^{\text{CASA-CNP HR}}$ have similar amplitudes in the NHL (Fig. 4a; Table 1), but the $CO_2^{\text{NEP}}$ amplitude from these two models differs (33±1.2 ppm versus 26±1 ppm, respectively; Fig. 3a; Table 1). This occurs because $CO_2^{\text{MIMICS HR}}$ peaks one-month later than $CO_2^{\text{CASA-CNP HR}}$, and has a zero-crossing that is more closely aligned

with the trough of $CO_2^{NPP}$ (Fig. 4a), leading to the larger amplitude in $CO_2^{MIMICS\ NEP}$ (Fig. 3a; Table 1). Although the amplitude mismatch decreases towards the south (Fig. 3b-f), the overall bias in the Northern Hemisphere suggests that either the seasonality of NPP is too large, or that all testbed models underestimate the seasonality of HR. Within the ST region, ensemble $CO_2^{HR}$ minima are opposite to that in $CO_2^{NPP}$, leading to a small annual cycle in simulations, consistent in magnitude with that of the observations (Figs. 3d, 4d).

### 3.2 Interannual imprint of heterotrophic respiration

The testbed ensemble reasonably simulates the magnitude and timing of interannual variability (IAV) compared with $CO_2$ observations (Fig. 5). Across the six latitude bands analyzed, simulated $CO_2^{NEP}$ IAV generally falls within one standard deviation of the median variation from observations for most of the study period (Fig. 5). Taking a closer look at the $CO_2$ from the component fluxes (NPP and HR), across all six latitude bands, the $CO_2^{NPP}$ IAV standard deviation is between 0.9 and 1.1 ppm for component fluxes (Fig. 6b). $CO_2^{CASA\text{-}CNP\ HR}$ IAV shows similar standard deviation as $CO_2^{NPP}$ IAV, whereas the standard deviations of $CO_2^{CORPSE\ HR}$ and $CO_2^{MIMICS\ HR}$ range from 0.7-1.4 ppm and 0.5-1.1 ppm, respectively (Fig. 6b).

Combining the $CO_2$ responses from component fluxes to $CO_2^{NEP}$ reveals a latitudinal gradient in IAV standard deviation similar to that of ESRL observations, with largest standard deviation found in the northern extratropics (Fig. 6a). Among the three testbed models, the standard deviation of $CO_2^{CASA\ NEP}$ agrees best with observations across all latitude bands ($CO_2^{CASA\ NEP}$: 0.5-0.9 ppm; ESRL: 0.6-1.0 ppm; Fig. 6a). $CO_2^{CORPSE\ NEP}$ overestimates IAV by up to 30% in NHL and NML, but agrees better with observations in the tropics and Southern Hemisphere. $CO_2^{MIMICS\ NEP}$ overestimates IAV standard deviations across all latitude bands (Fig. 6a). Interestingly, in the NHL, the overestimation is 20% even though $CO_2^{MIMICS\ HR}$ shows similar IAV as $CO_2^{NPP}$ (both 1.1 ppm; Fig. 6b). This suggests that the atmospheric $CO_2$ diagnostic for IAV, like that for amplitude, is critically sensitive to the phasing of IAV in heterotrophic respiration relative to the IAV of NPP.

Both global NPP and HR fluxes are sensitive to temperature variations at interannual timescales, with increased build-up of $CO_2$ in the atmosphere at higher temperatures, in part because the rate of HR increases at higher temperature and in part because most latitude bands show a reduction in NPP at above-average temperatures. For CASA-CNP, the temperature

sensitivity (γ) for globally integrated NPP and HR fluxes is 2.5 PgC $yr^{-1}$ $K^{-1}$ and 1.7 PgC $yr^{-1}$ $K^{-1}$; respectively (Fig. 7b). The temperature sensitivity of HR was higher for the microbially explicit models: 2.1 PgC $yr^{-1}$ $K^{-1}$ for CORPSE and 4.2 PgC $yr^{-1}$ $K^{-1}$ for MIMICS (Fig. 7b). For any given testbed flux (NPP, HR, or NEP), the temperature sensitivity of the resulting global mean $CO_2$ growth rate anomaly is higher than that of the underlying flux IAV. For example, the temperature sensitivity of the globally integrated NPP flux IAV (γNPP) is 2.5 PgC $yr^{-1}$ $K^{-1}$ whereas $\gamma CO_2^{NPP}$ is 3.2 PgC $yr^{-1}$ $K^{-1}$. The apparent amplification of the temperature sensitivity was even larger for HR. For example, the temperature sensitivity of MIMICS HR IAV ($\gamma HR^{MIMICS}$) was 4.2 PgC $yr^{-1}$ $K^{-1}$, whereas $\gamma CO_2^{MIMICS\ HR}$ was 7.7 PgC $yr^{-1}$ $K^{-1}$ (Fig. 7b). The simulated $\gamma CO_2^{NEP}$ simulated by the testbed models all overestimate the temperature sensitivity of the observed atmospheric $CO_2$ growth rate anomaly (6.1±2.5 PgC $yr^{-1}$ $K^{-1}$; Fig. 7a). CASA-CNP and CORPSE have temperature sensitivities within the range of the observed sensitivity (5.16±0.9 PgC $yr^{-1}$ $K^{-1}$, Cox et al., 2013; 6.5±1.8 PgC $yr^{-1}$ $K^{-1}$; Keppel-Aleks et al., 2018), but $\gamma CO_2^{MIMICS\ NEP}$ is 80% larger than observed value (10.9 PgC $yr^{-1}$ $K^{-1}$; Fig. 7a). We note that the γHR and $\gamma CO_2^{HR}$ are an emergent property that reflects both direct and indirect temperature influences, including the impact of temperature variability on NPP and litterfall (Table S3). Nevertheless, these results suggest that the direct temperature sensitivity of MIMICS HR is too high relative to observational constraints.

### 3.3 Geographic origins of $CO_2$ IAV

The interannual variability (IAV) in global NPP and HR originate from different geographic regions. The IAV in global NPP fluxes are dominated by variations within the tropics (both NT and ST regions), with relative standard deviation $\sigma_{REL}$ ~ 0.5 and correlation coefficient $r$ ~ 0.6 (Fig. 8a-b). The NML region also has a similar contribution to the NT in magnitude, but with a lower timing coherence ($r$ = 0.44; Fig. 8a-b). In contrast to the dominance of the tropics in contributing to the interannual variability of global NPP, the NML region contributes most to IAV in global HR, with $\sigma_{REL} \geq 0.6$ and $r$ ~ 0.8 for all three testbed models (Fig. 8c-d). The NHL region is also important in driving global HR flux variability based on CORPSE model results ($\sigma_{REL}$ = 0.59 and $r$ = 0.82; Fig. 8c-d). Despite high NPP variability in the tropics, the magnitude of tropical HR variability is only about 10-30% of global HR variability, and the timing coherence with the global signal is generally low ($r$ < 0.45; Fig. 8a-b). MIMICS HR IAV is the exception for the ST measuring close to 40% of global HR IAV magnitude and relatively high correlation ($r$ = 0.58;

Fig. 8c-d). Together, the tropics and NML contribute roughly equally to the magnitude of global NEP variability ($\sigma_{REL}$ between 0.44-0.55; Fig. 8e). Although the NML and NT show relatively high timing coherence (0.41-0.55), the ST shows the strongest timing coherence with global NEP IAV ($r > 0.7$; Fig. 8f).

Atmospheric transport modifies patterns of IAV in fluxes, emphasizing tropical flux patterns and de-emphasizing northern hemisphere flux patterns. For example, the role of ST in driving global $CO_2^{NPP}$ variability is amplified compared to the underlying fluxes, as the timing coherence with the global signal increases from $r = 0.64$ for flux IAV to $r = 0.88$ for $CO_2^{NPP}$ IAV for this region (Fig. 8b). Conversely, the role of NML is dampened, with timing coherence decreasing to $r = 0.33$ for $CO_2^{NPP}$ IAV versus $r = 0.44$ for NPP IAV (Fig. 8b). Similarly, timing coherence for tropical $CO_2^{HR}$ IAV is substantially higher than that for HR fluxes in ST and NT (>0.7), although the atmospheric transport impact differs across the three testbed models (Fig. 8d). In contrast to closely aligned NML correlation values for $CO_2^{HR}$ and HR ($r \sim 0.8$-$0.9$), NML $CO_2^{HR}$ IAV shows $\sigma_{REL}$ between 0.45 and 0.58, a decrease from the HR IAV contribution (NML HR IAV $\sigma_{REL}$ range: 0.57 to 0.74; Fig. 8c). For $CO_2^{NEP}$ IAV, the regional contribution is more consistent with similar $\sigma_{REL}$ and $r$ to flux IAV (Fig. 8e-f). These effects of numerical techniques should be considered when evaluating the underlying processes controlling modeled land fluxes.

## 4. Discussion

Modeled differences in heterotrophic respiration impart discernible signatures on atmospheric $CO_2$, suggesting that atmospheric $CO_2$ observations may be able to help evaluate broad differences in the timing and magnitude fluxes simulated by different vegetation and soil biogeochemical models. We used a 3-D atmospheric transport model to analyze the imprint of the atmospheric $CO_2$ resulting from soil heterotrophic respiration and net ecosystem exchange fluxes from soil testbed ensemble with three representations of soil biogeochemistry (CASA-CNP, CORPSE, MIMICS). Results show that the phasing of heterotrophic respiration fluxes relative to NPP fluxes is an important source of bias in evaluating simulated $CO_2$ against atmospheric observations at both seasonal and interannual timescales. Regional patterns of heterotrophic respiration variability provide non-negligible contributions to global $CO_2$ variability. Here we discuss these findings in

more detail as well as implications for the use of $CO_2$ observations for flux evaluation and model benchmarking.

### 4.1 Impacts of heterotrophic respiration on seasonality

Our evaluation of $CO_2$ simulated using testbed fluxes revealed that all testbed models overestimated the mean annual cycle amplitude of atmospheric $CO_2$ observations. In the Northern Hemisphere, the bias was largest for MIMICS, which had a $CO_2$ amplitude from net ecosystem production that was overestimated by up to 100% (Fig. 3). The mismatch in the amplitude of the Northern Hemisphere NEP fluxes was smallest from CORPSE, despite CORPSE also simulating the largest seasonal amplitude in HR fluxes (Fig. 3a-c, Table 1). By contrast, in the Southern Hemisphere the simulated $CO_2$ annual cycle amplitudes were similar across all three models, with small absolute mismatches (about 1 ppm) compared to observations (Fig. 3). We note that the differences in the amplitude of NEP fluxes across all three testbed formulations could be due to biases in the timing and magnitude of NPP and HR fluxes simulated by models in the testbed. However, an advantage of the testbed approach is that, because all of the models are driven by the same GPP and climate variables, the differences in the timing and magnitude of NEP fluxes are all related to differences in HR fluxes that are simulated by different soil models in the testbed. With future work we would like to consider forcing uncertainty that could be generated by using different inputs of productivity, temperature, and moisture from land model ensembles (e.g. TRENDY simulations, CMIP6 models, etc). From these results, however, it appears that the seasonal amplitude of atmospheric $CO_2$ fluxes from net ecosystem production that are simulated in the Northern high- and mid-latitudes are higher than atmospheric observations for all of the models tested here, but especially MIMICS.

One challenge in using atmospheric $CO_2$ to evaluate HR representation in soil models is the influence of productivity (NPP) on both HR fluxes and atmospheric $CO_2$ variations. The seasonal diagnostics we present are very sensitive to the phasing of HR fluxes relative to NPP. For example, in NHL a one-month lag in the seasonal maximum of $CO_2^{HR}$ between MIMICS and CASA-CNP (Fig. 4a) leads to a 7 ppm difference in the overall amplitude of $CO_2^{NEP}$— this despite identical amplitudes of $CO_2^{HR}$ for the two models (Fig. 3a). Although the substantial impacts of subtle phase differences complicate benchmarking, the sensitivity reveals interesting and important differences related to model structural choices (i.e. first order versus microbially

explicit). Wieder et al., (2018) noted that the microbially explicit models in the testbed had seasonal HR fluxes that peaked in the fall, about a month later than the HR fluxes simulated by CASA-CNP. Annual phasing of HR is altered with the addition of microbial processes but also reflects NPP seasonality. The timing of CASA-CNP fluxes largely depend on soil temperature (highest HR flux when temperature is highest), whereas MIMICS and CORPSE have maximum HR fluxes set by trade-offs between the timing of maximal temperature and maximal microbial biomass, which is more tightly linked with litterfall (Fig. 7 from Wieder et al., 2018). Thus, phasing of HR is a sensitive diagnostic for benchmarking, especially if additional constraints on the magnitude and phasing of NPP are available.

In this study, determining the unique contribution from HR was possible since NPP was common among the three soil models used in the testbed, but the contribution of NPP will need to be resolved for model evaluation in other contexts. For example, long-term records of vegetation productivity at regional and global scales have been observed via satellite vegetation indices (Hicke et al., 2002; Meroni et al., 2009; Running et al., 2004), and more recently chlorophyll fluorescence (Frankenberg et al., 2011; Guan et al., 2016; Köhler et al., 2018; Li et al., 2018). Our study underscores the importance of developing methods to use these datasets together with atmospheric $CO_2$ to inform the dynamics of carbon cycling and its component fluxes. Current benchmarks used to evaluate carbon cycle metrics in land models include globally gridded estimates of fluxes (GPP, NEE, ecosystem respiration) and C stocks (leaf area index, vegetation biomass, and soil C; Collier et al., 2018). This is an excellent starting point, but provides a rather coarse estimate for the component fluxes we are trying to evaluate with this analysis. Notably, current benchmarks but do not yet consider the other metrics like NPP, litterfall, or root turnover and exudation that are important drivers of ecosystem, soil, and heterotrophic respiration. Globally gridded estimates of annual soil respiration have been upscaled using machine learning techniques (Zhao et al., 2017), and we recognize the value in using this and similar data products to provide an independent benchmark to evaluate C fluxes that are simulated by models in the testbed or other model ensembles. These annual estimates are useful for looking at the spatial distribution of fluxes and infering information about simulated trends, but they will not help resolve differences in the timing of heterotrophic respiration fluxes (Fig. 4) that are driving differences in net ecosystem production in the testbed models (Fig. 3). Instead, additional work with databases of soil and heterotrophic respiration (e.g. Bond-Lamberty and Thomson, 2010; Schädel et al. 2019) will be

critical to evaluating the seasonal dynamics and environmental sensitivities of soil and heterotopic respiration fluxes.

### 4.2 Impacts of heterotrophic respiration on interannual variability

Capturing appropriate interannual variability is important to generating credible land C-cycle representations (Cox et al., 2013; Piao et al., 2019). To a first approximation, all models in the testbed generated interannual variability in NEP fluxes that matched latitudinal distributions from atmospheric observations (Fig. 5). Similar to the analyses on seasonal cycles, the testbed ensemble simulations showed a higher interannual variability of $CO_2$ fluxes associated with explicit microbial representation – especially for heterotrophic respiration fluxes with CORPSE in the Northern high-latitudes (Fig. 5a, 6).

Interestingly, in the tropics and Southern extra-tropics, the interannual variability of heterotrophic respiration fluxes simulated by MIMICS is only slightly higher than CASA-CNP or CORPSE (Fig. 6b), but the interannual variability of NEP fluxes simulated by MIMICS was 20-30% higher than that of other models (Fig. 6a). Further, in these regions the interannual variability of heterotrophic respiration fluxes simulated by MIMICS also shows an inverse, but highly correlated relationship with the interannual variability of NPP ($R^2 > 0.60$, Table S3). This suggests that the large interannual variability of NEP fluxes simulated by MIMICS may result from differences in phasing between NPP and MIMICS HR fluxes, similar to phasing between MIMICS NPP and HR affecting the shape of the $CO_2^{NEP}$ annual cycle in Northern high-latitudes. In the Northern high-latitudes, all testbed models show interannual variability of heterotrophic respiration is correlated with the interannual variability of both NPP and temperature ($R^2$ of 0.32 to 0.77; Table S3). Additionally, the interannual variability NPP is sensitive to temperature variability ($\gamma = 0.15$, $R^2 = 0.43$; Table S3). As in section 4.1, better diagnostics to partition the interannual variability of atmospheric $CO_2$ measurements into environmental sensitivities of heterotrophic respiration and productivity are required, especially at high latitudes, but our results suggest that the carbon cycle simulated by the MIMICS model shows interannual variability of $CO_2$ fluxes that are higher than atmospheric observations.

This high interannual variability of NEP simulated by MIMICS is consistent with this model having the highest global temperature sensitivity, overestimating observed value by 80% (Fig. 7a). CORPSE, the other microbially explicit model, had a 30% higher temperature sensitivity

in $CO_2^{NEP}$ than observed globally (Fig. 7a). This large bias in temperature sensitivity demonstrates uncertainties in the model structure and parameterization that is associated with soil biogeochemical models (Sulman et al., 2018). And although the temperature sensitivity of microbial kinetics simulated in MIMICS was parameterized with observations from enzyme assays from laboratory experiments (German et al., 2012; Wieder et al., 2014, 2015). Additional factors, including substrate availability, exert important proximal controls over the ultimate temperature sensitivity of soil C decomposition (Conant et al., 2011; Dungait et al., 2012). Recently, Zhang and co-authors (2020) uses observations from >200 sites in Europe and China to calibrate parameters for MIMICS, but these parameters have not yet been tested globally. Future work should similarly leverage local observations for model calibration to develop parameters that can be applied in subsequent global-scale simulations. The work presented here establishes a framework that uses a top-down constraint of atmospheric $CO_2$ observations to then evaluate, or benchmark, the $CO_2$ fluxes that are simulated by the revised model(s). As with larger land models (Collier et al., 2018), we see this interplay of model parameterization, testing, and evaluation as critical to refining and improving confidence in projections from soil biogeochemical models (Bradford et al., 2016).

### 4.3 Implications for model benchmarking using atmospheric $CO_2$

Our results provide useful insights for model benchmarking using atmospheric $CO_2$. On a global scale, interannual variability of simulated atmospheric $CO_2$ was shown to be affected by the variability in component fluxes (NPP, HR) from different land regions (Figs. 6-8). The tropics dominate the interannual variability in global NPP, while northern extratropics dominate the interannual variability in global heteroptophic respiration (Fig. 8a-d). Taken together, NEP variability reflects roughly equal contributions from northern hemisphere temperate ecosystems (NML) and tropical ecosystems (NT and ST; Fig. 8e-f). These results suggest that the interannual variability of atmospheric $CO_2$ results from two different processes (respiration and productivity) across multiple ecoclimatological regions, whereas previous studies have mostly identified tropical (e.g. Cox et al., 2013; Wang et al., 2013) or subtropical, semi-arid regions (e.g. Ahlström et al., 2015; Poulter et al., 2014) as dominant controls on the global interannual variability of atmospheric $CO_2$ observations. Additional analyses are needed to test the robustness of this finding

with different forcings and soil models, but these results emphasize the importance different processes and regions as sources of variability in the terrestrial carbon cycle.

Our analysis underscores that patterns of variability in atmospheric $CO_2$ are tied not only to variabilities in the underlying fluxes, but also to atmospheric transport. For example, we showed that the temperature sensitivity of $CO_2$ growth rate anomalies was larger than the sensitivity estimated from the fluxes themselves (Fig. 7). The enhanced temperature sensitivity for $CO_2^{HR}$ was larger than for that of $CO_2^{NPP}$, which suggests that the geographic origin of the fluxes relative to dominant patterns of transport affects the result (Fig. 7b). This transport enhancement of the apparent temperature sensitivity of $CO_2$ growth rate anomalies is consistent with results from Keppel-Aleks et al. (2018). While these results may be tied to the choice of GEOS-Chem to simulate atmospheric transport, they do underscore that (1) atmospheric $CO_2$ must be simulated from land fluxes to be use as a benchmark and (2) atmospheric observations should not be assumed to be a direct proxy for fluxes themselves.

We employed several benchmarking approaches, including timeseries comparison and functional response to temperature, to evaluate if $CO_2$ patterns reflect underlying representations of soil heterotrophic respiration. We found that soil heterotrophic respiration leaves non-negligible imprints on atmospheric $CO_2$, leaving open the possibility of more explicitly accounting for respiration variability using atmospheric $CO_2$ observations. Given that HR links to NPP, soil C pools, and temperature, we recommend synergistically using datasets that reflect these variables (instead of identifying metrics in isolation). This could provide better model process evaluation if implemented in a larger benchmarking framework, such as the International Land Model Benchmarking Project (ILAMB; Collier et al., 2018; Hoffman et al., 2017). Model development will be crucial in the next decade of carbon cycle research, but so will tools to test mechanistic understanding and elucidate a coherent picture of the land-atmosphere carbon response to a changing climate.

## Code and Data Availability

NOAA Earth System Research Laboratory $CO_2$ measurements (Dlugokencky et al., 2016; ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/) and the Climatic Research Unit's

gridded temperature product (Jones et al., 2012; http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__ACTIVITY_0c08abfc-f2d5-11e2-a948-00163e251233) are publicly available online. CASA testbed information and fluxes have been previously published in Wieder et al., 2018. GEOSChem $CO_2$ response data is available at the University of Michigan Library Deep Blue online repository (Basile et al., 2019; https://deepblue.lib.umich.edu/data/concern/data_sets/gt54kn02m).

## Author Contributions

Samantha J. Basile and Gretchen Keppel-Aleks designed the research. William R. Wieder, Melannie D. Hartman, and Xin Lin contributed model components. Samantha J. Basile conducted the analysis. All authors contributed to discussions. Samantha Basile, Gretchen Keppel-Aleks, and William R. Wieder wrote the manuscript.

## Competing Interests

The authors declare that they have no conflict of interest.

## Acknowledgements

Funding for this work was provided through the NASA ROSES Interdisciplinary Science Grant NNX17AK19G and through the RUBISCO Science Focus Area sponsored by the DoE Regional and Global Model Analysis program. WRW and MDH were also supported by grants from the US Department of Agriculture, National Institute of Food and Agriculture Award 2015-67003-23485 and the US Department of Energy, Biological and Environmental Research Awards: TES DE-SC0014374 & BSS DE-SC0016364.

We thank NOAA ESRL for providing observations of atmospheric $CO_2$. We thank the Climate Research Unit for their historically gridded temperature product.

## References

Ahlström, A., Raupach, M., Schurgers, G., Smith, B., Arneth, A., Jung, M., Reichstein, M., Canadell, J., Friedlingstein, P., Jain, A., Kato, E., Poulter, B., Sitch, S., Stocker, B., Viovy, N., Wang, Y. P., Wiltshire, A., Zaehle, S. and Zeng, N.: The dominant role of semi-arid ecosystems in the trend and variability of the land $CO_2$ sink, Science (80-. )., 348(6237), 895–899,

doi:10.1002/2015JA021022, 2015.

Anderegg, W.R., Ballantyne, A.P., Smith, W.K., Majkut, J., Rabin, S., Beaulieu, C., Birdsey, R., Dunne, J.P., Houghton, R.A., Myneni, R.B. and Pan, Y.: Tropical nighttime warming as a dominant driver of variability in the terrestrial carbon sink, Proceedings of the National Academy of Sciences, 112(51), 15591-15596, doi:10.1073/pnas.1521479112, 2015.

Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F. and Wu, T.: Carbon-concentration and carbon-climate feedbacks in CMIP5 earth system models, J. Clim., 26(15), 5289–5314, doi:10.1175/JCLI-D-12-00494.1, 2013.

Baldocchi, D.: TURNER REVIEW No. 15. "Breathing" of the terrestrial biosphere: Lessons learned from a global network of carbon dioxide flux measurement systems, Aust. J. Bot., 56(1), 1–26, doi:10.1071/BT07151, 2008.

Barba, J., Cueva, A., Bahn, M., Barron-Gafford, G. A., Bond-Lamberty, B., Hanson, P. J., Jaimes, A., Kulmala, L., Pumpanen, J., Scott, R. L., Wohlfahrt, G. and Vargas, R.: Comparing ecosystem and soil respiration: Review and key challenges of tower-based and soil measurements, Agric. For. Meteorol., 249(August 2017), 434–443, doi:10.1016/j.agrformet.2017.10.028, 2018.

Basile, S., Lin, X., Keppel-Aleks, G.: Simulated $CO_2$ dataset using the atmospheric transport model GEOSChem v12.0.0: Response to regional land carbon fluxes, doi:10.7302/xjzc-xy05, 2019.

Battle, M., Bender, M.L., Tans, P.P., White, J.W.C., Ellis, J.T., Conway, T. and Francey, R.J.: Global carbon sinks and their variability inferred from atmospheric O2 and δ13C, Science, 287(5462), 2467-2470, doi:10.1126/science.287.5462.2467, 2000.

Bond-Lamberty, B. and Thomson, A.: A global database of soil respiration data, Biogeosciences, 7(6), 1915–1926, doi:10.5194/bg-7-1915-2010, doi:10.1038/s41586-018-0358-x, 2010.

Bond-Lamberty, B., Wang, C. and Gower, S.T.: A global relationship between the heterotrophic and autotrophic components of soil respiration?, Glob. Change Biol., 10, 1756-1766, doi:10.1111/j.1365-2486.2004.00816.x, 2004.

Bond-Lamberty, B., Bronson, D., Bladyka, E., & Gower, S. T.: A comparison of trenched plot techniques for partitioning soil respiration. Soil Biol. Biochem., 43(10), 2108-2114, doi:10.1016/j.soilbio.2011.06.011, 2011.

Bond-Lamberty, B., Bailey, V. L., Chen, M., Gough, C. M. and Vargas, R.: Globally rising soil heterotrophic respiration over recent decades, Nature, 560(7716), 80–83, doi:10.1038/s41586-018-0358-x, 2018.

Bradford, M. A., Wieder, W. R., Bonan, G. B., Fierer, N., Raymond, P. A. and Crowther, T. W.: Managing uncertainty in soil carbon feedbacks to climate change, Nat. Clim. Chang., 6(8), 751–758, doi:10.1038/nclimate3071, 2016.

Bruhwiler, L., Michalak, A. M., Birdsey, R., Huntzinger, D. N., Fisher, J. B., Miller, J. and Houghton, R. A.: Overview of the Global Carbon Cycle, Second State Carbon Cycle Rep., 1–33, doi:10.7930/SOCCR2.2018.Ch1, 2018.

Buchkowski, R. W., Bradford, M. A., Grandy, A. S., Schmitz, O. J. and Wieder, W. R.: Applying population and community ecology theory to advance understanding of belowground biogeochemistry, Ecol. Lett., 20(2), 231–245, doi:10.1111/ele.12712, 2017.

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M. and Randerson, J. T.: The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation, J. Adv. Model. Earth Syst., 10(11), 2731–2754, doi:10.1029/2018MS001354, 2018.

Conant, R.T., Ryan, M.G., Ågren, G.I., Birge, H.E., Davidson, E.A., Eliasson, P.E., Evans, S.E., Frey, S.D., Giardina, C.P., Hopkins, F.M. and Hyvönen, R.: Temperature and soil organic matter decomposition rates – synthesis of current knowledge and a way forward, Glob. Chang. Biol., 17(11), 3392-3404, doi:10.1111/j.1365-2486.2011.02496.x, 2011.

Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D. and Luke, C. M.: Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability., Nature, 494(7437), 341–4, doi:10.1038/nature11882, 2013.

Crisp, D., Pollock, H., Rosenberg, R., Chapsky, L., Lee, R., Oyafuso, F., Frankenberg, C., Dell, C., Bruegge, C., Doran, G., Eldering, A., Fisher, B., Fu, D., Gunson, M., Mandrake, L., Osterman, G., Schwandner, F., Sun, K., Taylor, T., Wennberg, P. and Wunch, D.: The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products, Atmos. Meas. Tech., 10(1), 59–81, doi:10.5194/amt-10-59-2017, 2017.

Davidson, E. A., Savage, K., Verchot, L. V. and Navarro, R.: Minimizing artifacts and biases in chamber-based measurements of soil respiration, Agric. For. Meteorol., 113(1–4), 21–37, doi:10.1016/S0168-1923(02)00100-4, 2002.

Dlugokencky, E. J., Lang P. M., Mund J. W., Crotwell A. M., Crotwell M. J., and Thoning K. W.: Atmospheric carbon dioxide dry air mole fractions from the NOAA ESRL carbon cycle cooperative global air sampling network, 1968–2015, version 2016-08-30. NOAA, accessed 4 January 2017, ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/, 2016.

Dungait, J. A. J., Hopkins, D. W., Gregory, A. S., & Whitmore, A. P.: Soil organic matter turnover is governed by accessibility not recalcitrance, Glob. Chang. Biol., 18(6), 1781-1796, doi:10.1111/j.1365-2486.2012.02665.x, 2012.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J. E., Toon, G. C., Butz, A., Jung, M., Kuze, A. and Yokota, T.: New global observations of the terrestrial carbon cycle from GOSAT : Patterns of plant fluorescence with gross primary productivity, , 38, 1–6, doi:10.1029/2011GL048738, 2011.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M. and Zhao, B.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), J. Clim., 30(14), 5419–5454, doi:10.1175/JCLI-D-16-0758.1, 2017.

German, D. P., Marcelo, K. R. B., Stone, M. M., & Allison, S. D.: The Michaelis-Menten kinetics of soil extracellular enzymes in response to temperature: a cross-latitudinal study., Glob. Chang. Biol., 18(4), 1468-1479. doi:10.1111/j.1365-2486.2011.02615.x, 2012.

Guan, K., Berry, J. A., Zhang, Y., Joiner, J., Guanter, L., Badgley, G. and Lobell, D. B.: Improving the monitoring of crop productivity using spaceborne solar-induced fluorescence, Glob. Chang. Biol., 22(2), 716–726, doi:10.1111/gcb.13136, 2016.

Hicke, J. A., Asner, G. P., Randerson, J. T., Tucker, C., Los, S., Birdsey, R., Jenkins, J. C. and Field, C.: Trends in North American net primary productivity derived from satellite observations, 1982-1998, Global Biogeochem. Cycles, 16(2), 2-1-2–14, doi:10.1029/2001gb001550, 2002.

Hoffman, F. M., Koven, C. D., Keppel-Aleks, G., Lawrence, D. M., Riley, W. J., Randerson, J. T., Ahlström, A., Abramowitz, G., Baldocchi, D. D., Best, M. J., Bond-Lamberty, B., De Kauwe, M. G., Denning, A. S., Desai, A. R., Eyring, V., Fisher, J. B., Fisher, R. A., Gleckler, P. J., Huang, M., Hugelius, G., Jain, A. K., Kiang, N. Y., Kim, H., Koster, R. D., Kumar, S. V., Li, H., Luo, Y., Mao, J., McDowell, N. G., Mishra, U., Moorcroft, P. R., Pau, G. S. H., Ricciuto, D. M., Schaefer, K., Schwalm, C. R., Serbin, S. P., Shevliakova, E., Slater, A. G., Tang, J., Williams, M., Xia, J., Xu, C., Joseph, R. and Koch, D.: 2016 International Land Model Benchmarking (ILAMB) Workshop Report, , doi:10.2172/1330803, 2017.

Humphrey, V., Zscheischler, J., Ciais, P., Gudmundsson, L., Sitch, S. and Seneviratne, S. I.: Sensitivity of atmospheric $CO_2$ growth rate to observed changes in terrestrial water storage, Nature, 560(7720), 628–631, doi:10.1038/s41586-018-0424-4, 2018.

Jenkinson, A. D. S., Andrew, S. P. S., Lynch, J. M., Goss, M. J., Tinker, P. B. and Jenkinson, D. S.: The turnover of organic carbon and nitrogen in soil, Philos. Trans. R. Soc. London. Ser. B Biol. Sci., 329(1255), 361–368, doi:10.1098/rstb.1990.0177, 1990.

Jones, P. D., Lister, D. H., Osborn, T. J., Harpham, C., Salmon, M. and Morice, C. P.: Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010, J. Geophys. Res. Atmos., 117(5), doi:10.1029/2011JD017139, 2012.

Jones, P., and Harris, I.: CRU TS3. 21: Climatic Research Unit (CRU) Time-Series (TS) version 3.21 of high resolution gridded data of month-by-month variation in climate (Jan. 1901–Dec. 2012). NCAS British Atmospheric Data Centre, http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__ACTIVITY_0c08abfc-f2d5-11e2-a948-

00163e251233, doi:10.5285/D0E1585D-3417-485F-87AE-4FCECF10A992, 2013.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J. and Zhu, Y.: The NCEP/NCAR 40-year reanalysis project, B. Am. Met. Soc., 77(3), 437-472, doi:10.1175/1520-0477(1996)077%3C0437:TNYRP%3E2.0.CO;2, 1996.

Keeling, C. D., Piper, S. C., Whorf, T. P. and Keeling, R. F.: Evolution of natural and anthropogenic fluxes of atmospheric CO 2 from 1957 to 2003, Tellus, Ser. B Chem. Phys. Meteorol., 63(1), 1–22, doi:10.1111/j.1600-0889.2010.00507.x, 2011.

Keppel-Aleks, G., Randerson, J. T., Lindsay, K., Stephens, B. B., Keith Moore, J., Doney, S. C., Thornton, P. E., Mahowald, N. M., Hoffman, F. M., Sweeney, C., Tans, P. P., Wennberg, P. O. and Wofsy, S. C.: Atmospheric carbon dioxide variability in the community earth system model: Evaluation and transient dynamics during the twentieth and twenty-first centuries, J. Clim., 26(13), 4447–4475, doi:10.1175/JCLI-D-12-00589.1, 2013.

Keppel-Aleks, G., Wolf, A. S., Mu, M., Doney, S. C., Morton, D. C., Kasibhatla, P. S., Miller, J. B., Dlugokencky, E. J. and Randerson, J. T.: Small phytoplankton drive high summertime carbon and nutrient export in the Gulf of California and Eastern Tropical North Pacifi, Global Biogeochem. Cy., 29, 1295–1310, doi:10.1002/2014GB004890.Received, 2014.

Keppel-Aleks, G., Basile, S. J. and Hoffman, F. M.: A functional response metric for the temperature sensitivity of tropical ecosystems, Earth Interact., 22(7), doi:10.1175/EI-D-17-0017.1, 2018.

Köhler, P., Frankenberg, C., Magney, T.S., Guanter, L., Joiner, J. and Landgraf, J.: Global retrievals of solar ‐ induced chlorophyll fluorescence with TROPOMI: First results and intersensor comparison to OCO ‐ 2, Geophys. Res. Lett., 45(19), 10-456, doi:10.1029/2018GL079031, 2018.

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., Bonan, G. B., Lawrence, D. M. and Swenson, S. C.: The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4, Biogeosciences, 10(11), 7109–7131, doi:10.5194/bg-10-7109-2013, 2013.

Lavigne, M. B., Ryan, M. G., Anderson, D. E., Baldocchi, D. D., Crill, P. M., Fitzjarrald, D. R., Goulden, M. L., Gower, S. T., Massheder, J. M., McCaughey, J. H., Rayment, M. and Striegl, R. G.: Comparing nocturnal eddy covariance measurements to estimates of ecosystem respiration made by scaling chamber measurements at six coniferous boreal sites, J. Geophys. Res. Atmos., 102(D24), 28977–28985, doi:10.1029/97jd01173, 1997.

Lehmann, J. and Kleber, M.: The contentious nature of soil organic matter, Nature, 528(7580), 60–68, doi:10.1038/nature16069, 2015.

Medlyn, B. E., Zaehle, S., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hanson, P. J., Hickler, T., Jain, A. K., Luo, Y., Parton, W., Prentice, I. C., Thornton, P. E., Wang, S., Wang, Y. P., Weng, E., Iversen, C. M., Mccarthy, H. R., Warren, J. M., Oren, R. and Norby, R. J.: Using ecosystem experiments to improve vegetation models, Nat. Clim. Chang., 5(6), 528–534,

doi:10.1038/nclimate2621, 2015.

Meroni, M., Rossini, M., Guanter, L., Alonso, L., Rascher, U., Colombo, R. and Moreno, J.: Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications, Remote Sens. Environ., 113(10), 2037–2051, doi:10.1016/j.rse.2009.05.003, 2009.

Moorhead, D. L. and Weintraub, M. N.: The evolution and application of the reverse Michaelis-Menten equation, Soil Biol. Biochem., 125(May), 261–262, doi:10.1016/j.soilbio.2018.07.021, 2018.

Nevison, C. D., Mahowald, N. M., Doney, S. C., Lima, I. D., van der Werf, G. R., Randerson, J. T., Baker, D. F., Kasibhatla, P. and McKinley, G. A.: Contribution of ocean, fossil fuel, land biosphere, and biomass burning carbon fluxes to seasonal and interannual variability in atmospheric $CO_2$, J. Geophys. Res. Biogeosciences, 113(1), 1–21, doi:10.1029/2007JG000408, 2008.

Oleson, K. W., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Charles, D., Levis, S., Li, F., Riley, W. J., Zachary, M., Swenson, S. C., Thornton, P. E., Bozbiyik, A., Fisher, R., Heald, C. L., Kluzek, E., Lamarque, F., Lawrence, P. J., Leung, L. R., Muszala, S., Ricciuto, D. M. and Sacks, W.: Technical description of version 4.5 of the Community Land Model (CLM), NCAR Technical Note NCAR/TN-503+STR, Natl. Cent. Atmos. Res. Boulder, CO, (July), 420pp, doi:10.5065/D6RR1W7M, 2013.

Parton, W. J.: The CENTURY Model, in: Evaluation of Soil Organic Matter Models, edited by: Powlson, D. S., Smith, P. and Smith, J. U., Springer-Verlag, Berlin, Heidelberg, Germany, 283–291, 1996.

Piao, S., Wang, X., Wang, K., Li, X., Bastos, A., Canadell, J.G., Ciais, P., Friedlingstein, P. and Sitch, S.: Interannual variation of terrestrial carbon cycle: Issues and perspectives., Glob. Chang. Biol., 26(1), 300-318, doi:10.1111/gcb.14884, 2019.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., and Klooster, S. A.: Terrestrial ecosystem production: A process model based on global satellite and surface data, Global Biogeochem. Cy., 7(4), 811–841, doi:10.1029/93GB02725, 1993.

Poulter, B., Frank, D., Ciais, P., Myneni, R.B., Andela, N., Bi, J., Broquet, G., Canadell, J.G., Chevallier, F., Liu, Y.Y., Running, S.W., Stich, S. and van der Werf, G.R.: Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle, Nature, 509(7502), 600–603, doi:10.1038/nature13376, 2014.

Pumpanen, J., Kolari, P., Ilvesniemi, H., Minkkinen, K., Vesala, T., Niinistö, S., Lohila, A., Larmola, T., Morero, M., Pihlatie, M., Janssens, I., Yuste, J. C., Grünzweig, J. M., Reth, S., Subke, J. A., Savage, K., Kutsch, W., Østreng, G., Ziegler, W., Anthoni, P., Lindroth, A. and Hari, P.: Comparison of different chamber techniques for measuring soil $CO_2$ efflux, Agric. For. Meteorol., 123(3–4), 159–176, doi:10.1016/j.agrformet.2003.12.001, 2004.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., Bakker, D. C. E., Barbero, L., Becker, M., Betts, R. A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Cosca, C. E., Cross, J., Currie, K., Gasser, T., Harris, I., Hauck, J., Haverd, V., Houghton, R. A., Hunt, C. W., Hurtt, G., Ilyina, T., Jain, A. K., Kato, E., Kautz, M., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lima, I., Lombardozzi, D., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S., Nojiri, Y., Padín, X. A., Peregon, A., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Reimer, J., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Tian, H., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Viovy, N., Vuichard, N., Walker, A. P., Watson, A. J., Wiltshire, A. J., Zaehle, S. and Zhu, D.: Global Carbon Budget 2017, Earth Syst. Sci. Data Discuss., 1–79, doi:10.5194/essd-2017-123, 2018.

Randerson, J. T., Thompson, M. V., Malmstrom, C. M.: Substrate Limitations for Heterotrophs: Implications for models that estimate the seasonal cycle of atmospheric $CO_2$, Global Biogeochem. Cy., 10(4), 585–602, doi:10.1029/96GB01981, 1996.

Randerson, J. T., Thompson, M. V., Conway, T. J., Fung, I. Y. and Field, C. B.: The contribution of sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, Global Biogeochem. Cy., 11(4), 535–560, doi:10.1029/97GB02268, 1997.

Randerson, J. T., Hoffman, F. M., Thornton, P. E., Mahowald, N. M., Lindsay, K., Lee, Y. H., Nevison, C. D., Doney, S. C., Bonan, G., Stöckli, R., Covey, C., Running, S. W. and Fung, I. Y.: Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, Glob. Chang. Biol., 15(10), 2462–2484, doi:10.1111/j.1365-2486.2009.01912.x, 2009.

Rasmussen, C., Heckman, K., Wieder, W. R., Keiluweit, M., Lawrence, C. R., Berhe, A. A., Blankinship, J. C., Crow, S. E., Druhan, J. L., Hicks Pries, C. E., Marin-Spiotta, E., Plante, A. F., Schädel, C., Schimel, J. P., Sierra, C. A., Thompson, A. and Wagai, R.: Beyond clay: towards an improved set of variables for predicting soil organic matter content, Biogeochemistry, 137(3), 297–306, doi:10.1007/s10533-018-0424-3, 2018.

Rayner, P. J., Law, R. M., Allison, C. E., Francey, R. J., Trudinger, C. M. and Pickett-Heaps, C.: Interannual variability of the global carbon cycle (1992-2005) inferred by inversion of atmospheric $CO_2$ and $\delta$ 13 $CO_2$ measurements, Global Biogeochem. Cy., 22(3), 1–12, doi:10.1029/2007GB003068, 2008.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M. and Hashimoto, H.: A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production, Bioscience, 54(6), 547, doi:10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2, 2004.

Ryan, M. G. and Law, B. E.: Interpreting, measuring, and modeling soil respiration, Biogeochemistry, 73(1), 3–27, doi:10.1007/s10533-004-5167-7, 2005.

Schädel, C., Beem-Miller, J., Aziz Rad, M., Crow, S. E., Hicks Pries, C., Ernakovich, J., Hoyt, A. M., Plante, A., Stoner, S., Treat, C. C., and Sierra, C. A.: Decomposability of soil organic matter

over time: The Soil Incubation Database (SIDb, version 1.0) and guidance for incubation procedures, Earth Syst. Sci. Data Discuss., doi:10.5194/essd-2019-184, in review, 2019.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D. A. C., Nannipieri, P., Rasse, D. P., Weiner, S. and Trumbore, S. E.: Persistence of soil organic matter as an ecosystem property., Nature, 478(7367), 49–56, doi:10.1038/nature10386, 2011.

Schuur, E.A. and Mack, M.C.. Ecological response to permafrost thaw and consequences for local and global ecosystem services, Annu. Rev. Ecol. Evol. S., 49, 279-301, doi:10.1146/annurev-ecolsys-121415-032349, 2018.

Sulman, B. N., Phillips, R. P., Oishi, A. C., Shevliakova, E. and Pacala, S. W.: Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO 2, Nat. Clim. Chang., 4(12), 1099–1102, doi:10.1038/nclimate2436, 2014.

Sulman, B. N., Moore, J. A. M., Abramoff, R., Averill, C., Kivlin, S., Georgiou, K., Sridhar, B., Hartman, M. D., Wang, G., Wieder, W. R., Bradford, M. A., Luo, Y., Mayes, M. A., Morrison, E., Riley, W. J., Salazar, A., Schimel, J. P., Tang, J. and Classen, A. T.: Multiple models and experiments underscore large uncertainty in soil carbon dynamics, Biogeochemistry, 141(2), 109–123, doi:10.1007/s10533-018-0509-z, 2018.

Todd-Brown, K. E. O., Randerson, J. T., Hopkins, F., Arora, V., Hajima, T., Jones, C., Shevliakova, E., Tjiputra, J., Volodin, E., Wu, T., Zhang, Q. and Allison, S. D.: Changes in soil organic carbon storage predicted by Earth system models during the 21st century, Biogeosciences, 11(8), 2341–2356, doi:10.5194/bg-11-2341-2014, 2014.

Turner, D. P., Ritts, W. D., Cohen, W. B., Gower, S. T., Running, S. W., Zhao, M., Costa, M. H., Kirschbaum, A. A., Ham, J. M., Saleska, S. R. and Ahl, D. E.: Evaluation of MODIS NPP and GPP products across multiple biomes, Remote Sens. Environ., 102(3–4), 282–292, doi:10.1016/j.rse.2006.02.017, 2006.

Wang, Y. P., Law, R. M. and Pak, B.: A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere, Biogeosciences, 7(7), 2261–2282, doi:10.5194/bg-7-2261-2010, 2010.

Wieder, W. R., Bonan, G. B. and Allison, S. D.: Global soil carbon projections are improved by modelling microbial processes, Nat. Clim. Chang., 3(10), 909–912, doi:10.1038/nclimate1951, 2013.

Wieder, W. R., Grandy, A. S., Kallenbach, C. M. and Bonan, G. B.: Integrating microbial physiology and physio-chemical principles in soils with the MIcrobial-MIneral Carbon Stabilization (MIMICS) model, Biogeosciences, 11(14), 3899–3917, doi:10.5194/bg-11-3899-2014, 2014.

Wieder, W. R., Grandy, A. S., Kallenbach, C. M., Taylor, P. G. and Bonan, G. B.: Representing

life in the Earth system with soil microbial functional traits in the MIMICS model, Geosci. Model Dev., 8(6), 1789–1808, doi:10.5194/gmd-8-1789-2015, 2015.

Wieder, W. R., Hartman, M. D., Sulman, B. N., Wang, Y. P., Koven, C. D. and Bonan, G. B.: Carbon cycle confidence and uncertainty: Exploring variation among soil biogeochemical models, Glob. Chang. Biol., 24(4), 1563–1579, doi:10.1111/gcb.13979, 2018.

Wunch, D., Wennberg, P. O., Messerschmidt, J., Parazoo, N. C., Toon, G. C., Deutscher, N. M., Keppel-Aleks, G., Roehl, C. M., Randerson, J. T., Warneke, T. and Notholt, J.: The covariation of Northern Hemisphere summertime $CO_2$ with surface temperature in boreal regions, Atmos. Chem. Phys., 13(18), 9447–9459, doi:10.5194/acp-13-9447-2013, 2013.

Yang, Z., Washenfelder, R.A., Keppel-Aleks, G., Krakauer, N.Y., Randerson, J.T., Tans, P.P., Sweeney, C. and Wennberg, P.O.: New constraints on Northern Hemisphere growing season net flux, Geophys. Res. Lett., 34(12), doi:10.1029/2007GL029742, 2007.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H. and Maksyutov, S.: Global Concentrations of $CO_2$ and $CH_4$ Retrieved from GOSAT: First Preliminary Results, Sola, 5, 160–163, doi:10.2151/sola.2009-041, 2009.

Zhao, Z., Peng, C., Yang, Q., Meng, F. R., Song, X., Chen, S., Epule, T. E., Li, P. and Zhu, Q.: Model prediction of biome-specific global soil respiration from 1960 to 2012, Earth's Futur., 5(7), 715–729, doi:10.1002/2016EF000480, 2017.

Zhang, H., Goll, D. S., Wang, Y., Ciais, P., Wieder, W. R., Abramoff, R., Huang, Y., Guenet, B., Prescher, A., Viscarra Rossel, R., A., Barré, P., Chenu, C., Zhou, G., Tang, X.: Microbial dynamics and soil physicochemical properties explain large scale variations in soil organic carbon, Glob Chang Biol, in press., 2020.

Table 1 Atmospheric $CO_2$ mean annual cycle amplitude (in ppm) simulated from heterotrophic respiration (HR), net primary productivity (NPP), and net ecosystem productivity (NEP). The median annual cycle amplitudes for observed $CO_2$ ($CO_2^{OBS}$) averaged over latitude bands are also reported.

| | 61°-90°N | 24°-60°N | 0°-23°N | 1°-23°S | 24°-60°S | 61°-90°S |
|---|---|---|---|---|---|---|
| $CO_2^{CASA\text{-}CNP\ HR}$ | 17.6 | 11.4 | 4.3 | 4.3 | 1.1 | 1.9 |
| $CO_2^{CORPSE\ HR}$ | 28.2 | 16.6 | 6.4 | 4.9 | 1.4 | 2.2 |
| $CO_2^{MIMICS\ HR}$ | 17.2 | 11.8 | 5.1 | 4.4 | 1.9 | 2.5 |
| $CO_2^{CASA\text{-}CNP\ NPP}$ | 39.3 | 24.6 | 11.9 | 6.0 | 3.1 | 3.1 |
| $CO_2^{CASA\text{-}CNP\ NEP}$ | 26.2 | 16.3 | 9.3 | 1.6 | 2.2 | 2.2 |
| $CO_2^{CORPSE\ NEP}$ | 23.4 | 14.8 | 8.7 | 1.3 | 2.2 | 2.4 |
| $CO_2^{MIMICS\ NEP}$ | 32.8 | 19.0 | 10.4 | 1.7 | 1.9 | 2.1 |
| $CO_2^{OBS}$ | 15.3 | 10.6 | 6.1 | 0.9 | 0.8 | 1.4 |

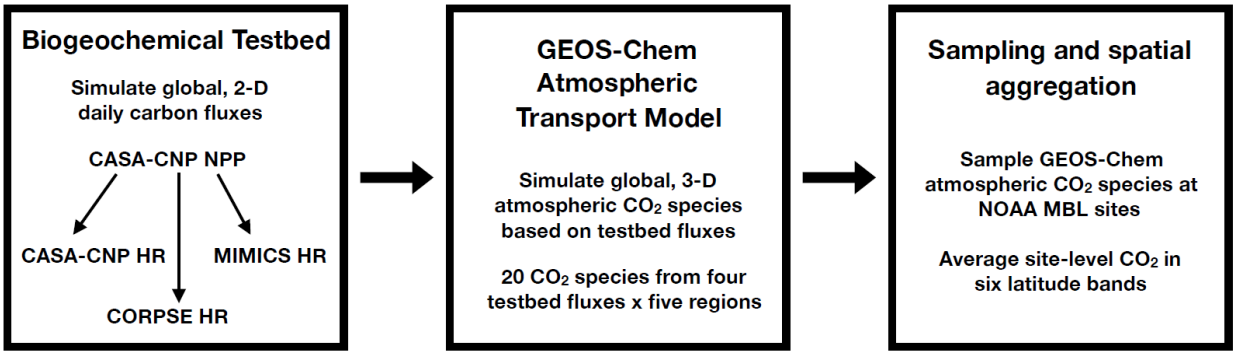


Figure 1: Flow chart depiction of the analysis process from soil model fluxes to simulated $CO_2$ concentration and comparison with NOAA observations.

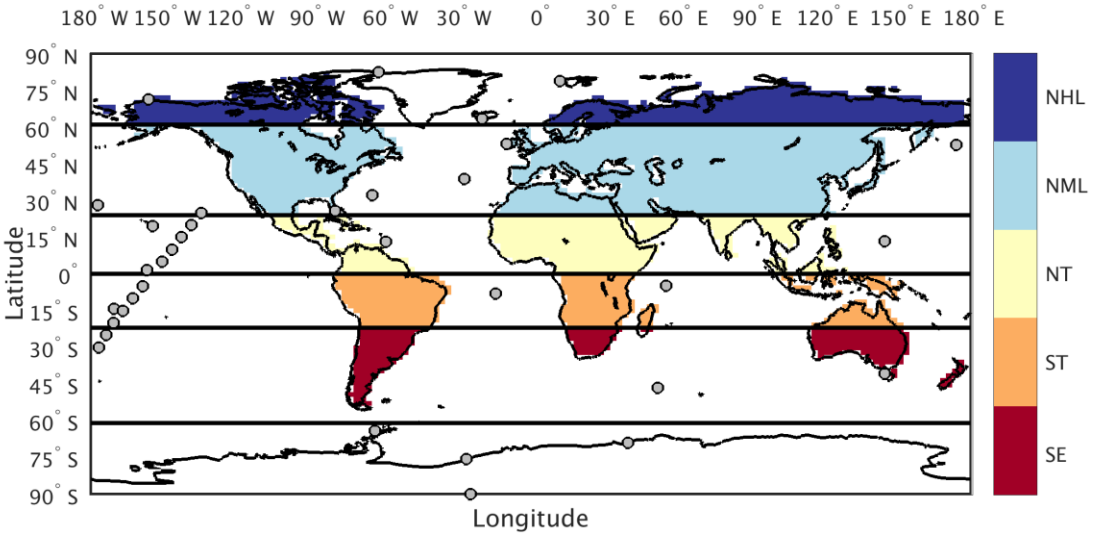


Figure 2: Tagged flux regions and marine boundary layer $CO_2$ observing sites used in our analysis. The 5 tagged flux regions are shown in color fill: Northern High Latitude (NHL), Northern Mid-Latitude (NML), Northern Tropics (NT), Southern Tropics (ST) and Southern Extratropics (SE). For sampling simulated $CO_2$ consistent with the tagged flux regions, we aggregate marine boundary layer sites (filled circles) into 6 latitude bands defined by the black lines.

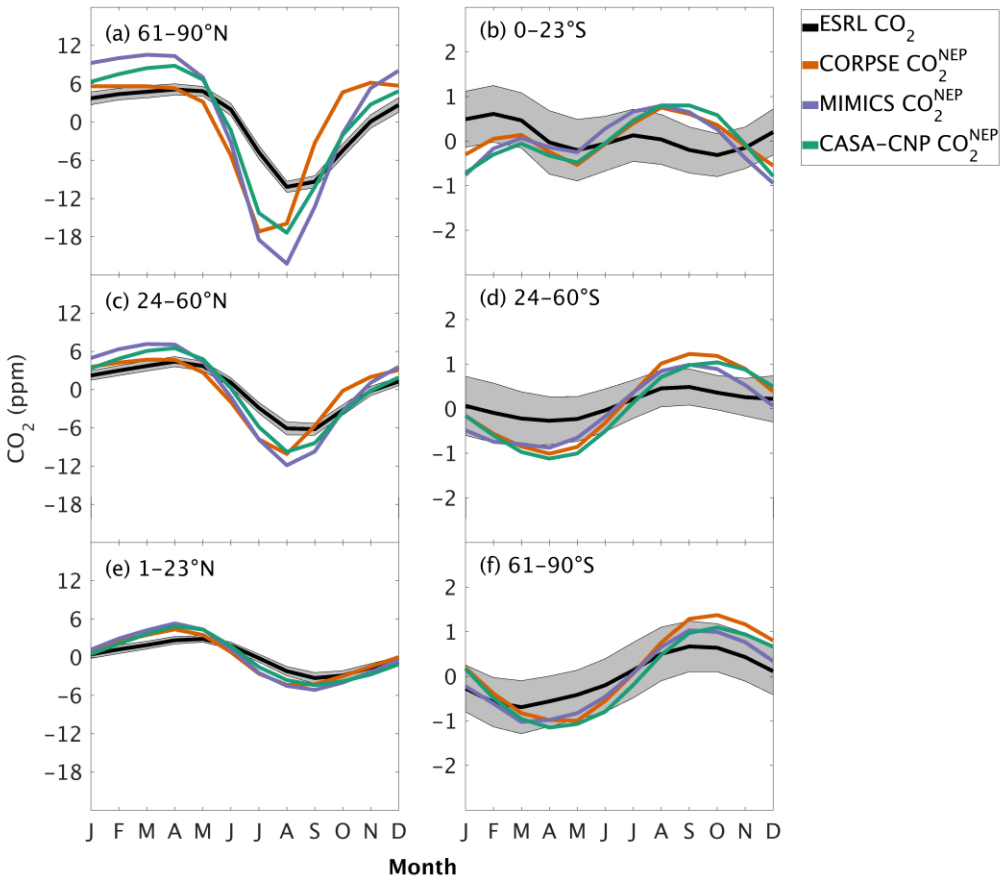


Figure 3: Climatological annual cycle (median) of $CO_2$ for observations (black) and global net ecosystem productivity flux ($CO_2^{NEP}$, colors) between 1982 and 2010. Monthly climatology values were created after detrending the $CO_2$ timeseries for atmospheric sampling bands in the (a-c) Northern Hemisphere (d-f) and Southern Hemisphere. Note the change in y-axis scale between the two hemispheres and the sign of $CO_2^{NEP}$ reflects the combination of $CO_2^{NPP}$ and $CO_2^{HR}$ (Eqn. 3). Shading on the observed line represents one standard deviation due to interannual variability in the seasonal cycle.

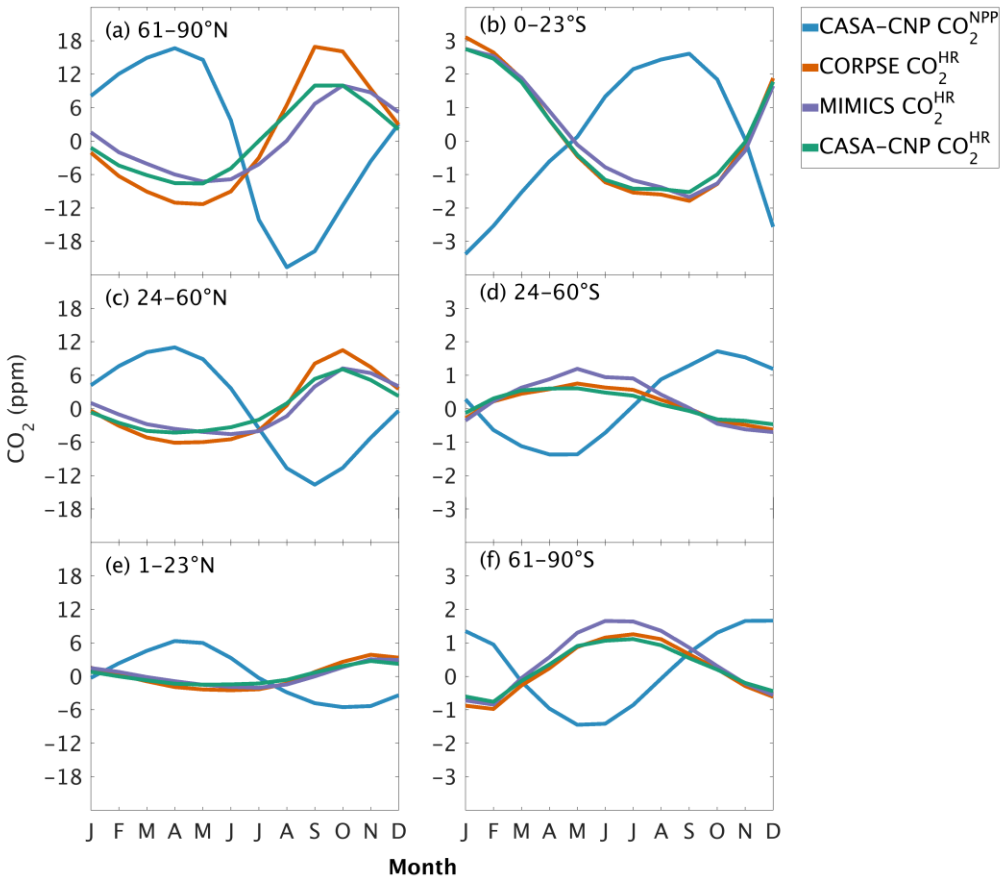


Figure 4: Climatological annual cycle (median) of atmospheric $CO_2$ simulated from land fluxes ($CO_2^{NPP}$, $CO_2^{HR}$) between 1982 and 2010. Monthly climatology values were created after detrending the $CO_2$ timeseries for atmospheric sampling bands in the (a-c) Northern Hemisphere (d-f) and Southern Hemisphere. Note the change in y-axis scale between the two hemispheres and the sign of $CO_2^{NPP}$ reflects the sign reversal of the underlying NPP (positive flux to the atmosphere; Eqn. 2).

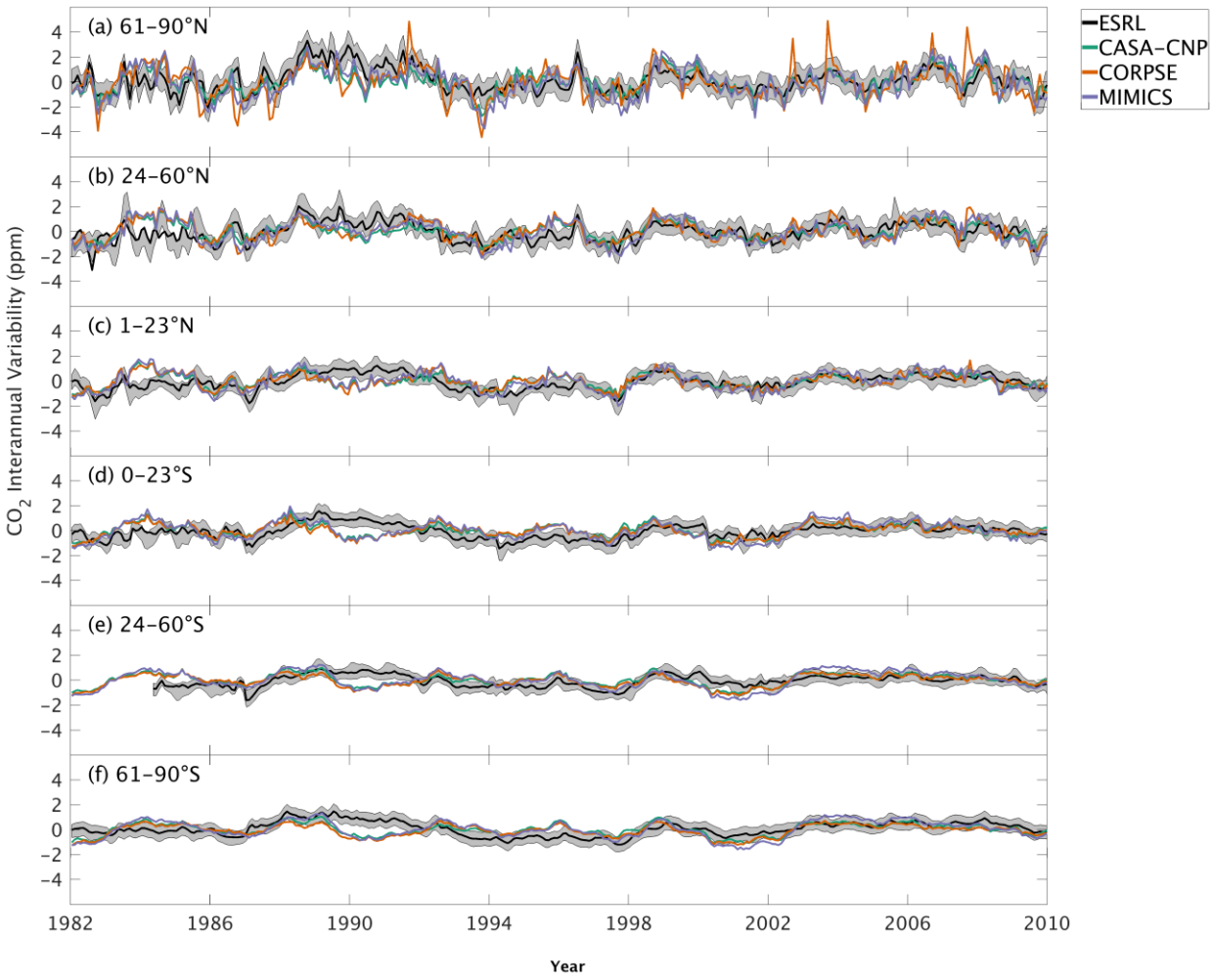


Figure 5: Interannual variability of $CO_2$ from global net ecosystem productivity ($CO_2^{NEP}$ IAV) for testbed models (colors) and marine boundary layer observations from the NOAA ESRL network (black). Gray shading outlines one standard deviation of observed $CO_2$ interannual variability. High-latitude, mid-latitude and tropical land belts are shown for the Northern Hemisphere (a-c) and Southern Hemisphere (d-f).

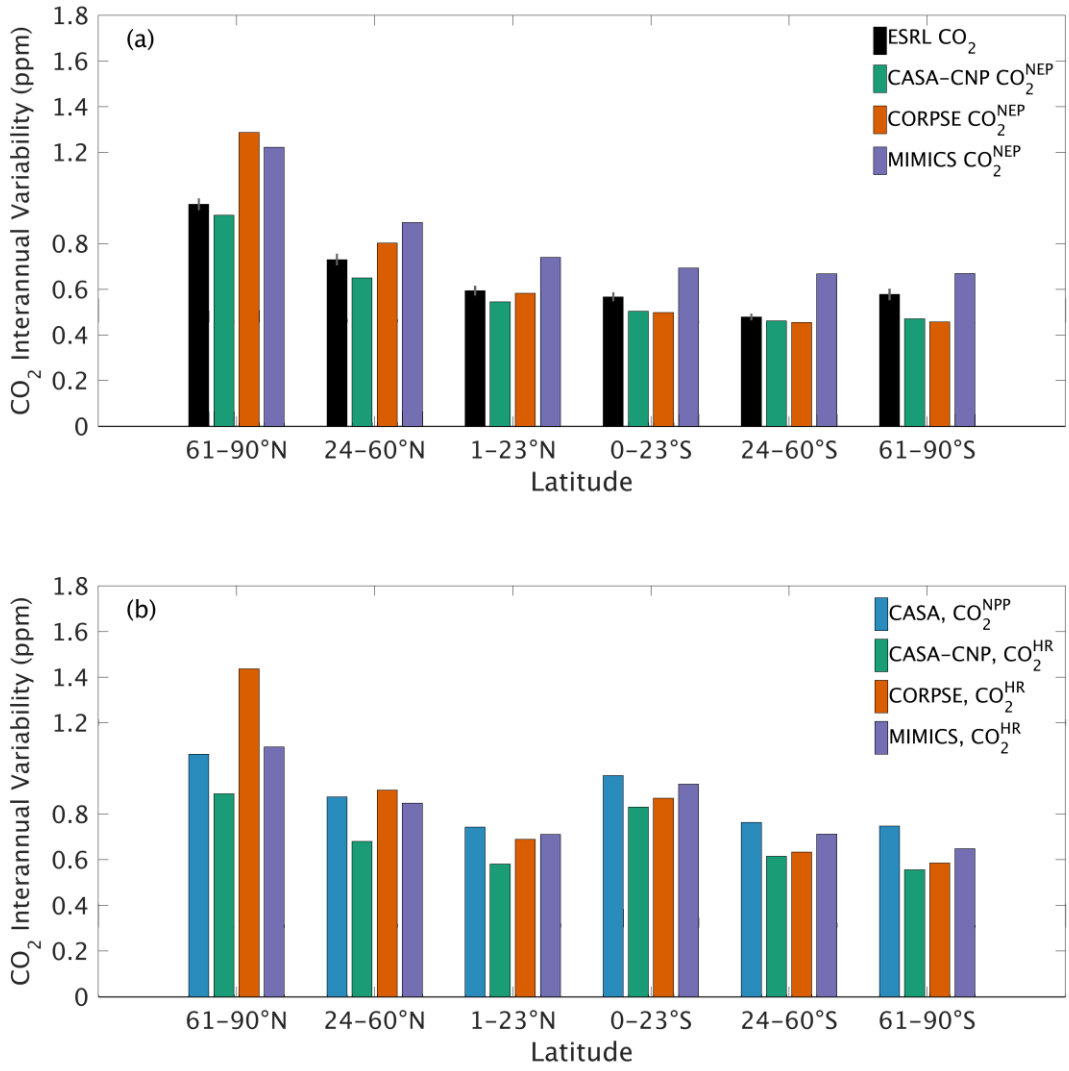


Figure 6: Magnitude of $CO_2$ interannual variability resulting from (a) individual flux components ($CO_2^{NPP}$ IAV, $CO_2^{HR}$ IAV) and (b) global net ecosystem productivity ($CO_2^{NEP}$ IAV). Observed $CO_2$ IAV from NOAA ESRL network are shown with black bars whereas colors represent simulated data. Errorbars shown on the observed IAV represent two standard deviations, calculated as the median magnitude after removing a 12 month sliding window from the IAV timeseries.

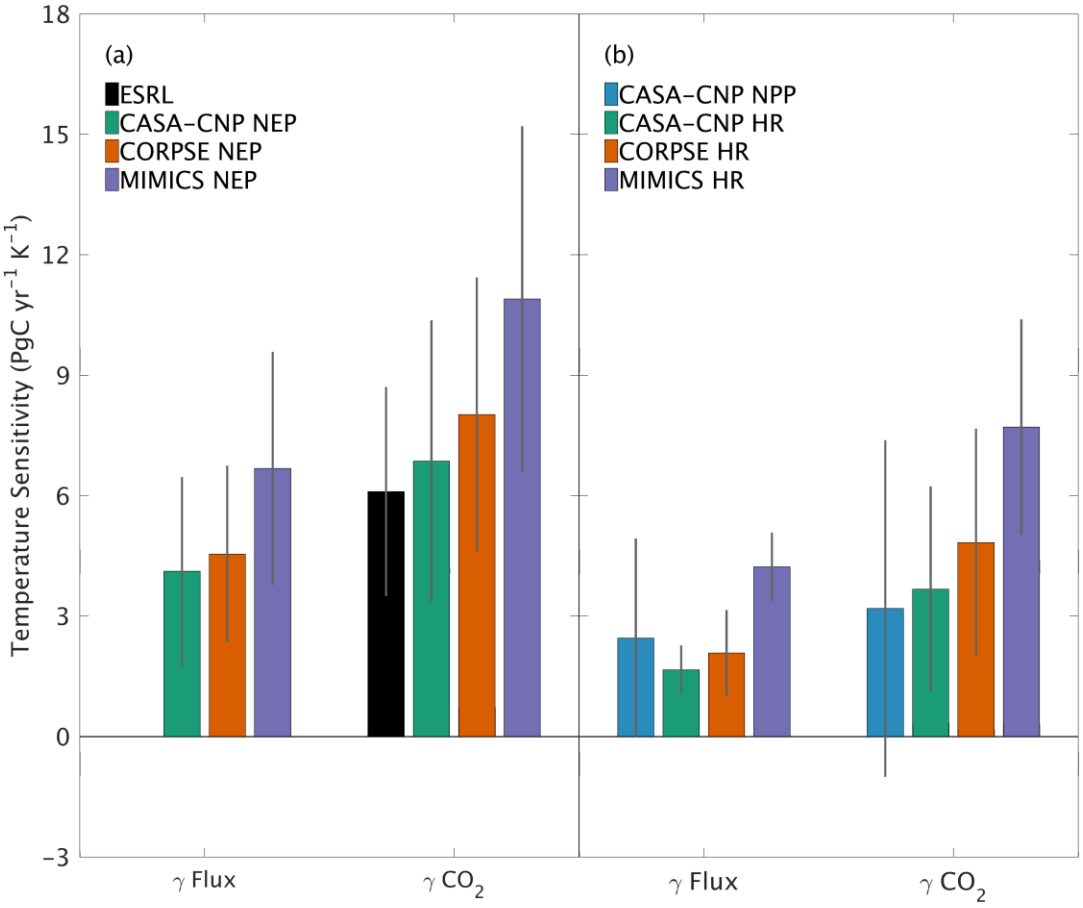


Figure 7: Temperature sensitivity (γ) calculated for interannual variability (IAV) of CASA-CNP air temperature and (a) flux IAV and corresponding $CO_2$ growth rate anomalies, (b) NEP IAV and $CO_2^{NEP}$ growth rate anomalies. Reference sensitivity value (black) was calculated using NOAA ESRL $CO_2$ and CRU TS4 air temperature. Sensitivity values were calculated as the ordinary linear regression coefficient between IAV timeseries for 1982 to 2010. Errorbars represent the 95% confidence interval for coefficient values.

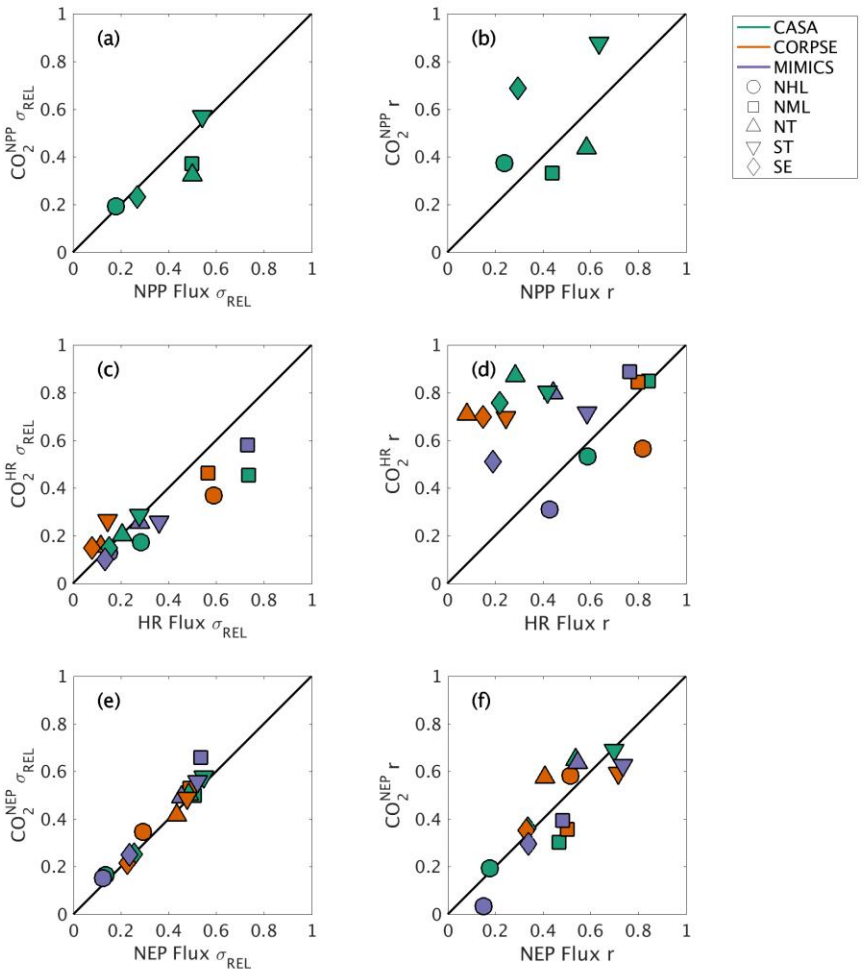


Figure 8: Comparison of regional and global interannual variability (IAV) from land fluxes and resulting atmospheric $CO_2$ between 1982 and 2010. (a, c, e) Normalized ratio taken between regional IAV and global IAV magnitude. (b, d, f) Linear correlation between regional IAV and global IAV. The scatterplot shows a direct comparison of ratio and correlation values for land flux values (x-axes) and corresponding $CO_2$ (y-axes). Shapes denote the source regions for both land fluxes and $CO_2$ response.