Solar-induced fluorescence products show variable skill in constraining global patterns in biospheric CO2 fluxes

Mingyang Zhang[1], Joe Berry[2], Yoichi Paolo Shiga[3], Russell Doughty[4], Nima Madani[5], Xing Li[6], Jingfeng Xiao[7], Jiaming Wen[8], Ying Sun[8], Ruixue Lei[1], and Scot Miller[1]

[1]Johns Hopkins University
[2]Carnegie Institution for Science
[3]Independent Researcher
[4]California Institute of Technology
[5]Jet Propulsion Laboratory
[6]Seoul National University
[7]University of New Hampshire
[8]Cornell University

December 7, 2022

Abstract

Solar-induced fluorescence (SIF) shows enormous promise as a proxy for photosynthesis and as a tool for modeling variability in gross primary productivity (GPP) and net biosphere exchange (NBE). In this study, we explore the skill of SIF and other vegetation indicators in predicting variability in global atmospheric CO2 observations, and thus global variability in NBE. We do so using a four-year record of global CO2 observations from NASA's Orbiting Carbon Observatory 2 (OCO-2) satellite and using a geostatistical inverse model. We find that existing SIF products closely correlate with space-time variability in atmospheric CO2 observations in the extra-tropics but show weaker explanatory power across the tropics. In the extra-tropics, all SIF products exhibit greater skill in explaining variability in atmospheric CO2 observations compared to an ensemble of process-based CO2 flux models and other vegetation indicators. Furthermore, we find that using SIF as a predictor variable in the geosatistical inverse model shifts the seasonal cycle of estimated NBE and yields an earlier end to the growing season relative to other vegetation indicators. In tropical biomes, by contrast, the seasonal cycles of SIF products and estimated NBE are out of phase, and existing respiration and biomass burning estimates do not reconcile this discrepancy. Overall, our results highlight several advantages and challenges of using SIF products to help predict global variability in GPP and NBE.

Solar-induced fluorescence products show variable skill in constraining global patterns in biospheric CO_2 fluxes

Mingyang Zhang[1], Joseph A. Berry[2], Yoichi P. Shiga[3], Russell B. Doughty[4], Nima Madani[5,6], Xing Li[7], Jingfeng Xiao[8], Jiaming Wen[9], Ying Sun[9], Ruixue Lei[1], Scot M. Miller[1]

[1]Department of Environmental Health and Engineering, Johns Hopkins University, Baltimore, MD, USA
[2]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA
[3]Independent Researcher, San Francisco, CA, USA
[4]College of Atmospheric & Geographic Sciences, University of Oklahoma, Norman, OK, USA
[5]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA
[6]UCLA Joint Institute for Regional Earth System Science and Engineering. Los Angeles, CA
[7]Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, South Korea
[8]Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH, USA
[9]School of Integrative Plant Science, Soil and Crop Sciences Section, Cornell University, Ithaca, NY, USA

Key Points:

- SIF products adeptly explain variability in atmospheric CO2 observations, and thus in CO2 fluxes, in the extra-tropics but not the tropics.
- Inverse model estimates of net biospheric exchange (NBE) that are informed by SIF exhibit a different seasonal cycle in the extra-tropics
- The seasonal cycle of SIF products in tropical biomes is out of phase with inverse estimates of NBE.

Corresponding author: Mingyang Zhang, `mzhang78@jhu.edu`

Abstract

Solar-induced fluorescence (SIF) shows enormous promise as a proxy for photosynthesis and as a tool for modeling variability in gross primary productivity (GPP) and net biosphere exchange (NBE). In this study, we explore the skill of SIF and other vegetation indicators in predicting variability in global atmospheric CO_2 observations, and thus global variability in NBE. We do so using a four-year record of global CO_2 observations from NASA's Orbiting Carbon Observatory 2 (OCO-2) satellite and using a geostatistical inverse model. We find that existing SIF products closely correlate with space-time variability in atmospheric CO_2 observations in the extra-tropics but show weaker explanatory power across the tropics. In the extra-tropics, all SIF products exhibit greater skill in explaining variability in atmospheric CO_2 observations compared to an ensemble of process-based CO_2 flux models and other vegetation indicators. Furthermore, we find that using SIF as a predictor variable in the geosatistical inverse model shifts the seasonal cycle of estimated NBE and yields an earlier end to the growing season relative to other vegetation indicators. In tropical biomes, by contrast, the seasonal cycles of SIF products and estimated NBE are out of phase, and existing respiration and biomass burning estimates do not reconcile this discrepancy. Overall, our results highlight several advantages and challenges of using SIF products to help predict global variability in GPP and NBE.

1 Introduction

CO_2 uptake by photosynthesis, also known as gross primary productivity (GPP), is a key driver of the carbon cycle (e.g., Beer et al., 2010; Field et al., 1995). However, global-scale patterns in GPP are difficult to estimate. For example, terrestrial biospheric flux models (TBMs) give widely different estimates of GPP; models do not show consensus on the global magnitude of GPP, seasonal amplitude, or inter-annual variability – often due to divergent model responses to environmental conditions (e.g., Anav et al., 2015; Huntzinger et al., 2012, 2017). Huntzinger et al. (2012) further argue that uncertainties in estimated GPP dominate uncertainties in modeled net biospheric exchange (NBE), at least in an analysis of North America. In addition to models, multiple data-driven GPP estimates are available, like those generated from eddy flux towers. However, eddy flux sites are unevenly distributed across the globe, and the data are often up-scaled using machine learning algorithms to obtain a global GPP estimate (Jung et al., 2019). These estimates also show numerous differences relative to TBMs (Jung et al., 2020).

These uncertainties have motivated a longstanding interest in generating remote sensing products that can help predict space-time patterns in GPP. Numerous studies have argued that solar-induced fluorescence (SIF) holds particular promise in this regard (e.g., Damm et al., 2015; Frankenberg et al., 2011; Guan et al., 2015; Guanter et al., 2014; Köhler et al., 2018; X. Li, Xiao, & He, 2018; X. Li, Xiao, He, Arain, et al., 2018; Luus et al., 2017; MacBean et al., 2018; Shiga et al., 2018a; Y. Sun et al., 2018; Verma et al., 2017; Wood et al., 2017). SIF is radiation emitted in the red and near-infrared by chlorophyll. Hence, it can serve as an indicator of sunlight absorption by chlorophyll and therefore has the potential as a predictor of photosynthesis in plants. In the past decade, a growing number of space-based sensors provide information on SIF, opening a new window into studying photosynthesis and GPP at local to global spatial scales.

Over the past decade, there has been a profusion of research on SIF and the carbon cycle. Several studies quantify the relationships between SIF and GPP, explore the linearity or non-linearity of those relationships, and how those relationships vary across different vegetation types (e.g., A. Chen et al., 2021; Gu et al., 2019; Helm et al., 2020; Kim et al., 2021; Z. Li et al., 2020; X. Li & Xiao, 2022; Magney et al., 2017; Magney, Frankenberg, et al., 2019; Marrs et al., 2020; Y. Sun et al., 2018; Verma et al., 2017; Wood et al., 2017). Additional studies explore how SIF varies during climate anomalies like heatwaves or drought (e.g., Guan et al., 2016; He et al., 2019; Helm et al., 2020; Jiao et al., 2019; Shekhar et al., 2020; L. Zhang et al., 2019).

In general, most studies show a close correlation between SIF and GPP at space-time scales that are observable by current satellite observations of SIF (e.g., Frankenberg et al., 2011; Guanter et al., 2012, 2014; X. Li, Xiao, He, Arain, et al., 2018; Y. Sun et al., 2018; Verma et al., 2017; Wood et al., 2017). By contrast, the relationship between SIF and GPP can be non-linear and/or weak at the scale of individual leaves or plants, partly due to variability in photosynthetic efficiency at these scales (e.g., Magney et al., 2020). With that said, the non-linearities found at sub-canopy scales often average out at kilometer spatial scales (Magney et al., 2020) and when SIF is upscaled from instantaneous to daily or monthly scales (e.g., Hu et al., 2018; Pierrat et al., 2022). Specifically, SIF-GPP relationships are likely strongest across coarser space-time scales where GPP closely correlates with absorbed photosynthetically active radiation (APAR) (e.g., Magney et al., 2020; Marrs et al., 2020).

Numerous works also compare and contrast SIF against other vegetation indicators that are commonly used to model the global carbon cycle, including the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) (e.g., Chang et al., 2019; Doughty et al., 2021; Jeong et al., 2017; Magney, Bowling, et al., 2019; Shiga et al., 2018a; Yang et al., 2015; J. Zhang et al., 2022; Zuromski et al., 2018). In general, SIF appears to reflect changes in GPP induced by seasonal or climate-related variability more quickly than either EVI or NDVI (e.g., Luus et al., 2017; Jeong et al., 2017; Magney, Bowling, et al., 2019; Meroni et al., 2009; Shekhar et al., 2020; F. Wang et al., 2020).

In addition to SIF, another vegetation indicator, known as the near-infrared reflectance of vegetation (NIRv), has recently gained attention as a potential proxy for GPP. NIRv is the estimated portion of near-infrared (NIR) reflectance that is due to vegetation (Badgley et al., 2017, 2019; Dechant et al., 2020, 2022). One motive for the creation of NIRv is to decrease contamination due to non-vegetation (branches, litter, etc.) that can be present in other vegetation indicators like EVI and NDVI. Existing studies show that NIRv correlates with GPP (Badgley et al., 2017), largely because NIRv indicates variation in canopy structure, which was shown to correlate with light use efficiency and GPP at several crop sites (Dechant et al., 2020).

There has also been substantial interest in incorporating SIF within TBMs to improve regional to global estimates of GPP and NBE (e.g., Bacour et al., 2019; Luus et al., 2017; MacBean et al., 2018; Parazoo et al., 2020; Thum et al., 2017). However, it is challenging to evaluate the relationships between SIF and GPP or NBE across large regions and across the entire globe. For example, existing studies often evaluate these relationships at eddy flux sites, which have a very localized footprint (e.g., Dechant et al., 2022; S. Wang et al., 2021; Wood et al., 2017) or evaluate these relationships across larger regions using model estimates of GPP (e.g., Byrne et al., 2018; Frankenberg et al., 2011; Verma et al., 2017).

Atmospheric CO_2 observations, by contrast, provide an opportunity to evaluate the skill of vegetation indicators in describing space-time variability in GPP and NBE across larger regions. Satellites like OCO-2 provide global coverage of CO_2 observations, including regions with sparse ground-based atmospheric or eddy flux CO_2 data. The task of evaluating vegetation indicators using atmospheric CO_2 observations entails several challenges. First, atmospheric observations are influenced by all types of CO_2 fluxes, not just GPP. Second, atmospheric CO_2 observations do not provide a direct measure of surface CO_2 fluxes and typically necessitate an atmospheric model and/or inverse model to relate observations to surface fluxes.

Several existing studies provide a possible road map for how to connect atmospheric CO_2 observations with vegetation indicators. For example, a handful of studies directly compare the seasonal cycles of SIF against satellite-based CO_2 observations across the Amazon (e.g., Parazoo et al., 2013; Albright et al., 2022); these studies report that the seasonality of SIF and atmospheric CO_2 are modestly anti-correlated, implying that GPP is driving the seasonality of NBE in this region (Parazoo et al., 2013). Additional studies use SIF to help interpret space-time patterns in NBE estimated using atmospheric CO_2 observations and inverse modeling (e.g., Liu et al., 2017, 2020; Byrne et al., 2021).

Shiga et al. (2018a) use a different approach to evaluate vegetation indicators using atmospheric CO_2 observations across North America. The authors estimate NBE using a geostatistical inverse model (GIM) that is paired with an atmospheric transport model. They test out different vegetation indicators as predictor variables of NBE within the inverse model and evaluate how well each helps the inverse model match atmospheric CO_2 observations. Using this framework, the authors find a stronger correlation between atmospheric CO_2 observations and SIF relative to other vegetation indicators. The authors posit that SIF better captures peak CO_2 uptake in croplands and better describes seasonal transitions in boreal evergreeen forests.

In the present study, we use atmospheric CO_2 observations and a GIM to evaluate the ability of SIF products, NIRv, and other vegetation indicators to help constrain global patterns in NBE. We are particularly interested in why SIF products are (or are not) better predictors of space-time variability in atmospheric CO_2 observations relative to other vegetation indicators. We then compare the ability of SIF products to help describe variability in atmospheric CO_2 observations against state-of-the-art TBMs. We are also interested in how estimated NBE changes globally when we use different vegetation indicators as predictors.

2 Methods

The overall approach in this study is to incorporate different vegetation indicators as predictor variables of NBE in a geostatistical inverse model (GIM). The GIM is fitted to global CO_2 observations from OCO-2 or in situ CO_2 observations. The approach used here follows that of Shiga et al. (2018a), who evaluate the relationships between SIF and NBE across North America using in situ CO_2 observations. It also builds upon the methodology developed in previous GIM studies using both in situ and satellite CO_2 observations (e.g., Gourdji et al., 2008; Shiga et al., 2018b; S. M. Miller et al., 2020a; Z. Chen et al., 2021a, 2021b).

2.1 The atmospheric inverse model

A GIM will produce an estimate of CO_2 fluxes (in this case, the sum of NBE, anthropogenic emissions, and ocean fluxes) using atmospheric CO_2 observations, an atmospheric transport model, and predictor variables of CO_2 fluxes. Specifically, a GIM models CO_2 fluxes as the sum of two different components (e.g., Kitanidis & Vomvoris, 1983; Michalak et al., 2004; Fang et al., 2014; Z. Chen et al., 2021a):

$$\boldsymbol{s} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\zeta} \tag{1}$$

The first component ($\mathbf{X}\boldsymbol{\beta}$) is a linear model of different predictor variables that may help describe variability in NBE. In the above equation, $\boldsymbol{s}$ are total CO_2 fluxes, the sum of NBE, ocean fluxes, and anthropogenic emissions. The dimensions of $\boldsymbol{s}$ are $m \times 1$, where m is the number of model grid cells at all different times and locations. The variable $\mathbf{X}$ is a matrix of different predictor variables, and each column of $\mathbf{X}$ is a different predictor. If there are p predictor variables, then $\mathbf{X}$ has dimensions $m \times p$. In previous studies, these predictor variables have included vegetation indicators, environmental data (e.g., estimates of soil moisture or PAR), estimates of biomass burning CO_2 fluxes, estimates of ocean CO_2 fluxes, and estimates of anthropogenic CO_2 emissions (e.g., Gourdji et al., 2008, 2012; Fang et al., 2014; Fang & Michalak, 2015; Shiga et al., 2018a, 2018b; Z. Chen et al., 2021a, 2021b). The coefficients ($\boldsymbol{\beta}$, dimensions $p\times1$) scale the overall magnitude of each predictor variable. These coefficients are estimated as part of the GIM to optimize the model fit against CO_2 observations, a topic discussed later in this section.

It is unlikely that any combination of predictor variables will be able to perfectly match actual CO_2 fluxes (e.g., Gourdji et al., 2008, 2012; Z. Chen et al., 2021a). There may be errors in these predictor variables, and/or there may be complex processes governing NBE that cannot be explained by a linear combination of available predictor variables. In the second component of Eq. 1, denoted $\boldsymbol{\zeta}$ (dimensions $m\times1$), the GIM will quantify additional patterns in CO_2 fluxes such that the CO_2 flux estimate better matches atmospheric CO_2 observations (e.g., Michalak et al., 2004). This component can vary in each model grid box and at each time step (in this case, each day).

Note that we subtract anthropogenic emissions from our estimate of $\boldsymbol{s}$ to obtain an estimate of NBE. In this study, we use an anthropogenic emissions estimate from the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC, Oda et al., 2018) as a predictor variable in $\mathbf{X}$, and we subsequently subtract ODIAC from our estimate of $\boldsymbol{s}$ to obtain an estimate of NBE. Note that it is standard practice in existing inverse modeling studies of CO_2 to subtract the influence of anthropogenic emissions in order to obtain an estimate of NBE (e.g., Gourdji et al., 2012; Z. Chen et al., 2021a, 2021b; Peiro et al., 2022). We do not attempt to further partition the NBE estimate into GPP, respiration, or biomass burning fluxes. We argue that the atmospheric CO_2 observations used in this study do not provide sufficient information to confidently partition the NBE estimate into these categories. Furthermore, no existing GIM study has attempted this type of partitioning (e.g., Michalak et al., 2004; Gourdji et al., 2008, 2012; Fang et al., 2014; Fang & Michalak, 2015; Shiga et al., 2018a, 2018b; Z. Chen et al., 2021a, 2021b).

The CO_2 fluxes from the inverse model ($\boldsymbol{s}$), when passed through an atmospheric model ($h()$), should match atmospheric CO_2 observations ($\boldsymbol{z}$) (e.g., Fang et al., 2014; Z. Chen et al., 2021a):

$$\boldsymbol{z} = h(\boldsymbol{s}) + \boldsymbol{\epsilon} \tag{2}$$

Specifically, the CO_2 fluxes ($\boldsymbol{s}$) should match the atmospheric observations ($\boldsymbol{z}$) within a margin of error specified by the inverse modeler ($\boldsymbol{\epsilon}$). The variance of $\boldsymbol{\epsilon}$ needs to be specified by the user before running the inverse model, and this point is discussed in greater detail in the SI. $\boldsymbol{z}$ and $\boldsymbol{\epsilon}$ have dimensions $n\times1$, where n are the number of observations. In this study, we use 10-second averages of version 10r CO_2 observations from OCO-2 (land nadir, land glint, and target observations only) (e.g. Peiro et al., 2022) and in situ observations from the NOAA CO_2 Obspack v3.2 developed for the OCO-2 model intercomparison project (MIP) (NOAA Global Monitoring Laboratory, 2021). In addition, we use the GEOS-Chem forward and adjoint models (version 9-02) for the atmospheric transport ($h()$). The global simulations here are driven by winds from the Modern-Era Retrospective Analysis for Research and Applications 2 (MERRA-2) (Gelaro et al., 2017) and have a spatial resolution of 4° latitude by 5° longitude. Furthermore, the simulations in this study cover September 2014 through December 2018. Note that we initialize model simulations for September 2014 using estimated atmospheric CO_2 fields from NOAA's CarbonTracker CT2019 product (Jacobson et al., 2020). We discard 2014 as a model spin-up period, following the procedure used in Z. Chen et al. (2021a) and Z. Chen et al. (2021b).

Both the CO_2 fluxes ($\boldsymbol{s}$) and coefficients ($\boldsymbol{\beta}$) are unknown and must be estimated as part of the GIM. The coefficients are calculated by solving a linear equation (e.g., Fang et al., 2014; Z. Chen et al., 2021a):

$$\hat{\boldsymbol{\beta}} \ = \ (h(\mathbf{X})^T \boldsymbol{\Psi}^{-1} h(\mathbf{X}))^{-1} h(\mathbf{X})^T \boldsymbol{\Psi}^{-1} \boldsymbol{z} \tag{3}$$

In this equation, $\mathbf{\Psi}$ (dimensions $n \times n$) is a covariance matrix that describes the uncertainties in the model-data system (i.e., describes the residuals $\boldsymbol{z} - h(\mathbf{X}\boldsymbol{\beta})$). This matrix is defined by the modeler and is described in more detail in the SI. Furthermore, this equation requires inputting each vegetation indicator and predictor variable into the atmospheric transport model in place of a CO_2 flux estimate. In other words, we input each vegetation indicator in GEOS-Chem as if it were a CO_2 flux to calculate $h(\mathbf{X})$. GEOS-Chem will then translate patterns in surface-level variables like SIF into an atmospheric tracer with patterns defined by SIF. Note that the units or absolute magnitude of the vegetation indicators are not important in this framework (e.g., Shiga et al., 2018a, 2018b). Rather, the coefficients ($\boldsymbol{\beta}$) estimated using the GIM will scale the magnitude of each vegetation indicator or predictor variable to create a model of CO_2 fluxes that optimally matches atmospheric CO_2 observations ($\boldsymbol{z}$).

Note that we run the GIM for the entire globe, but we estimate different coefficients ($\boldsymbol{\beta}$) for different biomes and for different years, the same approach used in S. M. Miller et al. (2018), S. M. Miller and Michalak (2020b), Z. Chen et al. (2021a), and Z. Chen et al. (2021b). These biomes are shown in Fig. S1 and are the same biomes used in the aforementioned studies. This setup accounts for the fact that the relationships between NBE and vegetation indicators like SIF may be different in boreal forests versus deserts or tropical grasslands (e.g., A. Chen et al., 2021). We also analyze the results at this biome level. The use of different coefficients in different years also means that the model is being fitted to spatial and seasonal variability within each year; spurious multi-year trends in the predictor variables will not adversely impact the model-data fit. With that said, we present the estimated coefficients ($\hat{\boldsymbol{\beta}}$) and NBE estimate averaged across the four-year study period (2015–2018). The SI, by contrast, presents year-to-year differences in the estimated coefficients.

The choice of biomes here is specifically informed by several previous studies of OCO-2 observations (S. M. Miller et al., 2018; S. M. Miller & Michalak, 2020b). These studies find that recent versions of the observations can be used to constrain NBE across large biome-based regions in most seasons. If we attempt to estimate different coefficients ($\boldsymbol{\beta}$) for smaller regions, we generally obtain unrealistic and unphysical estimates because OCO-2 observations do not provide sufficient information to constrain the coefficients across smaller regions. Hence, the biomes used here balance our desire to obtain detailed information about NBE with the limitations of currently-available CO_2 observations.

In contrast to the coefficients ($\boldsymbol{\beta}$), estimating the CO_2 fluxes ($\boldsymbol{s}$) is more involved. This step requires minimizing a cost function (e.g., Kitanidis & Vomvoris, 1983; Michalak et al., 2004):

$$L = \frac{1}{2}(\boldsymbol{z} - h(\boldsymbol{s}))^T \mathbf{R}^{-1}(\boldsymbol{z} - h(\boldsymbol{s})) + \frac{1}{2}(\boldsymbol{s} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{Q}^{-1}(\boldsymbol{s} - \mathbf{X}\boldsymbol{\beta}) \quad (4)$$

The first component of the cost function quantifies how well the estimated CO_2 fluxes , when passed through GEOS-Chem, match the atmospheric observations ($\boldsymbol{z}-h(\boldsymbol{s})$, aka $\boldsymbol{\epsilon}$). They should match within a margin specified by $\mathbf{R}$ (dimensions $n \times n$), a covariance matrix defined by the modeler. The second term governs the properties of $\boldsymbol{\zeta}$ or, equivalently, $\boldsymbol{s}-\mathbf{X}\boldsymbol{\beta}$ (Eq. 1). $\boldsymbol{\zeta}$ should have spatial and temporal properties that match the covariance matrix $\mathbf{Q}$ (dimensions $m \times m$), which is also defined by the modeler. These covariance matrices are described in greater detail in the SI. We minimize the cost function with respect to $\boldsymbol{s}$ using an iterative solver, described in S. M. Miller et al. (2020a), Z. Chen et al. (2021a), and Z. Chen et al. (2021b).

2.2 Remote sensing vegetation indicators and predictor variables

We incorporate several vegetation indicators as predictor variables of NBE in the GIM. We specifically use four different SIF products based on SIF retrievals from OCO-2 – $\mathrm{SIF}_{\mathrm{OCO2\_005}}$ (Yu et al., 2019a), CSIF (Y. Zhang et al., 2018), GOSIF (X. Li & Xiao, 2019b), and a SIF product from scientists at the Jet Propulsion Laboratory (referred to as JPL SIF) (Madani et al., 2022). The first three products are created by interpolating SIF retrievals from OCO-2 onto a global grid using machine learning algorithms. The last product (JPL SIF) is created by binning high-quality OCO-2 SIF retrievals into monthly, 4° latitude by 4° longitude grid boxes and taking a simple mean of all retrievals in each grid box. Note that other satellite sensors also provide SIF retrievals (e.g., the Greenhouse Gases Observing Satellite (GOSAT) and the TROPOspheric Monitoring Instrument (TROPOMI)), but we specifically use SIF products from OCO-2 retreivals because these products are available for the same time period as the modeling simulations in this study (2015–2018) and are generated from the same satellite (OCO-2) as the CO_2 observations used in this study.

Although all four SIF products are based on OCO-2 SIF observations, these products show several notable differences. First, the JPL SIF product is based on a much simpler method than any of the machine-learning (ML) based products. Second, different OCO-2 SIF retrievals are used in these products. GOSIF and CSIF use the 757 nm wavelenth, while SIF_{OCO2_005} uses the average of 757nm and 771 nm wavelengths, and JPL SIF uses the 740 nm wavelength. Third, although three of the four SIF products are ML-based, they use either a Cubist regression tree-based method (GOSIF) or feed-forward neural networks (CSIF and SIF_{OCO2_005}). However, Wen et al. (2020) find a similar prediction performance of these two types of ML methods, and therefore the difference in the ML method alone may not yield notable differences in the final SIF products. Fourth, different sets of predictor variables are used in each of the three ML-based products. The nadir bidirectional reflectance distribution adjusted reflectance (NBAR) from MODIS is the only predictor in the model for CSIF (MCD43C4) and SIF_{OCO2_005} (MCD43A4 and MCD43C4) products, and the MCD43 surface reflectance may contain some missing values in tropical forests due to clouds (Yu et al., 2019a; Y. Zhang et al., 2018). By contrast, the GOSIF study does not include the MODIS NBAR as a predictor. Instead, they use environmental data such as EVI from MODIS (MCD12), PAR, vapor pressure deficit, and air temperature from MERRA-2 (X. Li & Xiao, 2019b). Lastly, these studies use different strategies to fit the ML model. Specifically, GOSIF and CSIF are fitted globally, while SIF_{OCO2_005} is fitted separately for each individual biome.

Note that for the setup here, we aggregate the GOSIF, CSIF, and SIF_{OCO2_005} products and other predictor variables (described below) to a 16-day time resolution before inputting them into the GIM, ensuring a fairer comparison among different GIM simulations using different vegetation indicators (e.g., Shiga et al., 2018a; Z. Chen et al., 2021a). Note that JPL SIF is available at a monthly time resolution; OCO-2 SIF observations are spatially sparse, and the monthly time resolution ensures that there are enough observations in each grid box to obtain a reliable mean. Refer to SI Sect. S3 for more discussion of this point.

In addition to SIF, we also include several additional vegetation indicators as predictor variables in the inverse model – NDVI, EVI, and NIRv. NDVI is the difference between NIR reflectance and visible red reflectance, divided by the sum of these two quantities (e.g., NASA, 2000). Green vegetation reflects in the NIR but not in the visible red, and NDVI therefore provides a measure of vegetation greenness. EVI additionally includes a correction for atmospheric effects and background noise (e.g., USGS, 2022). We use 16-day EVI and NDVI from MODIS Terra (product MOD13C1 with best, good, and mixed quality assurance flags; QA =0, 1 and 2, respectively). NIRv, by contrast, is the product of vegetation indicator NIR reflectance and NDVI, and we construct NIRv using red (620–670 nm) and NIR (841–876 nm) reflectance data from MODIS Terra (product MCD43, reflectance bands 1 and 2). We further re-grid each product from the original 0.05° by 0.05° degree resolution provided by NASA to the 4° by 5° resolution of GEOS-Chem.

In addition to these vegetation indicators, we include environmental driver data (e.g., estimated meteorological variables) as predictor variables of NBE. These additional predictor variables may help account for space-time variability in respiration, and may describe additional patterns in GPP not described by the vegetation indicators. The use of environmental driver data in the GIM follows numerous existing studies (Gourdji et al., 2008, 2012; Fang et al., 2014; Fang & Michalak, 2015; Shiga et al., 2018a, 2018b; Z. Chen et al., 2021a, 2021b). We specifically consider driver data from MERRA-2 – 2 m air temperature, precipitation, PAR, surface downwelling shortwave radiation, soil temperature at 10 cm depth, soil moisture at 10 cm depth, specific humidity, and relative humidity. We also include a non-linear function of air temperature from Mahadevan et al. (2008). In a recent GIM study using CO_2 observations from OCO-2, Z. Chen et al. (2021a) find that air temperature is a poor predictor variable and does little to help the inverse model describe patterns in CO_2 observations. Rather, they find that a non-linear function of air temperature has much better explanatory power in the GIM. Refer to SI Sect. S3 for a discussion of uncertainties in these environmental driver data.

We do not assimilate all of these environmental driver data from MERRA-2 as predictor variables in the GIM; several of these variables are highly correlated or colinear and including all of these variables would likely overfit the CO_2 observations from OCO-2 and in situ sites. Instead, we use model selection based on the Bayesian Information Criterion (BIC) to determine which combination of variables in different biomes can best complement the vegetation indicators and optimize model-data fit against CO_2 observations from OCO-2. Numerous GIM studies to date employ the BIC to decide on a set of environmental predictor variables for the GIM (e.g., Gourdji et al., 2012; Fang et al., 2014; Fang & Michalak, 2015; S. Miller et al., 2014a; S. M. Miller et al., 2016a; Shiga et al., 2018a, 2018b; Z. Chen et al., 2021a, 2021b). Furthermore, the approach used here mirrors that used in Z. Chen et al. (2021a) and Z. Chen et al. (2021b) and is described in greater detail in the Supplement.

In some GIM simulations (e.g., Sect. 3.3), we also consider respiration estimates from an ensemble of TBMs for use as predictor variables in the GIM. We specifically incorporate respiration estimates from 15 flux models that are part of the Global Carbon Projects' Trends in Net Land Atmosphere Carbon Exchanges (TRENDY) model intercomparison (version 8, Friedlingstein et al., 2019; Sitch et al., 2015). Note that we use TRENDY scenario three, which includes all forcings (e.g., climate and land use forcings). These respiration estimates are available for the time period of this study (2014–2018) and are reported at variable spatial resolution and monthly temporal resolution, described in the Supplement.

All model simulations in this study further include estimates for other CO_2 source types: fossil fuel fluxes from ODIAC (Oda et al., 2018), ocean fluxes from the Circulation and Climate of the Ocean consortium (ECCO-Darwin, Carroll et al., 2020), and Global Fire Emissions Database (GFED) version 4.1 (Giglio et al., 2013). These flux estimates are incorporated as predictor variables in the GIM (in $\mathbf{X}$, as in Z. Chen et al., 2021a, 2021b).

2.3 Analysis using the inverse model and predictor variables

We use the inverse model or GIM to conduct two analyses. In the first analysis, we evaluate how well we are able to reproduce patterns in atmospheric CO_2 observations using a linear model of vegetation indicators and other predictor variables. For this analysis, we use the first component of the CO_2 flux estimate from the GIM ($\mathbf{X}\boldsymbol{\beta}$ in Eq. 1). We input this component of the flux estimate into GEOS-Chem and compare the results against CO_2 observations. This linear model provides a convenient way to evaluate vegetation indicators like SIF and NIRv using atmospheric CO_2 observations.

For the second analysis, we investigate how NBE estimated by the inverse model changes when we incorporate different vegetation indicators as predictors. For this analysis, we evaluate the full NBE estimate from different GIM simulations that use different vegetation indicators. This analysis illustrates the potential of SIF products to inform inverse estimates of NBE; it highlights the additional information on NBE provided by SIF products compared to inverse estimates of NBE using other vegetation indicators (e.g., Shiga et al., 2018a). Note that estimating the linear model ($\mathbf{X}\boldsymbol{\beta}$) using Eq. 3 requires relatively little computing time while estimating NBE using Eq. 4 requires several weeks on a supercomputer cluster. Hence, we only discuss a few representative examples for this second set of analyses using estimated NBE.

The vegetation indicators evaluated here are often used as proxies for GPP, and we acknowledge that GPP is not the only component of NBE. With that said, we argue that atmospheric CO_2 observations and the GIM can help inform the use of these vegetation indicators. First, in many biomes, GPP likely dominates large-scale space-time patterns in NBE (e.g., Parazoo et al., 2013; Shiga et al., 2018a; W. Sun et al., 2021). Specifically, GPP is a large component of NBE, and other flux processes like autotrophic respiration likely exhibit similar seasonal patterns as GPP (albeit with opposite sign) (e.g., Huntzinger et al., 2012). Furthermore, existing regional-scale atmospheric studies show a correlation between GPP and NBE in both tropical and extra-tropical regions (e.g., Parazoo et al., 2013; Shiga et al., 2018a; W. Sun et al., 2021).

Second, our primary goal is not to evaluate the absolute performance of the GIM relative to atmospheric CO_2 observations. Rather, we are interested in the relative performance of simulations that use vegetation indicators as one of several predictor variables of NBE. Uncertainties in GFED and the environmental driver data, among other uncertainties, could lower overall model performance relative to the CO_2 observations, but these uncertainties are unlikely to erroneously make one vegetation indicator appear more skilled than another.

Third, in instances where SIF products do not show favorable results in the GIM compared to other vegetation indicators, we also examine the possible role of uncertainties in respiration and biomass burning to explain the discrepancies.

3 Results and discussion

3.1 Summary of global results

A linear model of SIF products is able to describe substantial variability in CO_2 observations from OCO-2. Specifically, a model that consists of a linear combination of SIF products can describe between 40–85% of all variability in the OCO-2 observations, depending upon the biome (Fig. 1a). This result reaffirms the skill of SIF to describe regional spatial and seasonal patterns in NBE. Existing studies using atmospheric CO_2 observations indicate a strong correlation between SIF and NBE in a handful of regions – across southern Amazonia (Parazoo et al., 2013) and across North America (Shiga et al., 2018a), and this study suggests strong correlations across much broader global biomes.

SIF products are also more skilled at predicting variability in CO_2 observations compared to other vegetation indicators, at least in the extra-tropics (Fig. 1a). In these biomes, model-data comparisons using the SIF-based linear model show higher R^2 values compared to models based on EVI, NDVI, and NIRv (Fig. 1a). Indeed, this result complements several studies of eddy flux data that show a stronger relationship between SIF and GPP than NDVI or EVI and GPP (e.g., Guan et al., 2016; Magney, Bowling, et al., 2019; Shiga et al., 2018a; Yang et al., 2015; J. Zhang et al., 2022; Zuromski et al., 2018). It also mirrors regional studies that find large-scale, seasonal decoupling between GPP and measures of greenness like EVI and NDVI in numerous extra-tropical biomes (e.g., Walther et al., 2016; Jeong et al., 2017; Luus et al., 2017; Pierrat et al., 2021). In Sect.

3.2, we discuss in detail why a linear model of SIF is a better fit against OCO-2 observations in the extra-tropics, and we explore how inverse estimates of NBE that are informed by SIF differ from those informed by other vegetation indicators.

By contrast, SIF products do not show the same advantage relative to other vegetation indicators in the tropics (Fig. 1a). Linear models using SIF products, EVI, and NIRv exhibit similar R^2 values relative to OCO-2 observations in tropical biomes. This topic is the focus of Sect. 3.3.

We also note that all model simulations (i.e., using EVI, NDVI, SIF, and NIRv) exhibit a lower R^2 value relative to OCO-2 observations in the tropics (Fig. 1a). This lower model skill could be explained by the fact that NBE often has a larger seasonal cycle in the extra-tropics than in the tropics, and seasonal patterns in CO_2 observations are therefore likely easier to fit in the former biomes than in the latter. In addition, the atmospheric transport model (GEOS-Chem coupled with winds from MERRA-2) may be subject to larger errors in the tropics than in the extra-tropics; OCO-2 has a sun-synchronous orbit and passes over every location at approximately 1pm local time. At this time of day, there is often heterogeneous convection in biomes like tropical forests, and these features may be challenging to model using a global chemical transport model like GEOS-Chem (e.g., Jiang et al., 2013). Relatedly, biomass burning events can also create convection that is difficult to capture in atmospheric transport models (e.g., S. M. Miller et al., 2008), particularly in the tropics where biomass burning emissions are highly variable (e.g., Giglio et al., 2013). These factors may help explain why the overall model skill is lower in the tropics than in the extra-tropics. However, it does not explain why the model simulations using SIF products do not outperform the other vegetation indicators like EVI or NDVI in tropical biomes, as is the case in extra-tropical biomes.

We further find that linear model simulations using different SIF products yield different model-data fit against CO_2 observations (Fig. 1a). Globally, we find that a linear model using either GOSIF or JPL SIF yields a slightly higher R^2 in Fig. 1 relative to other SIF products. This result is perhaps surprising because JPL SIF is a much simpler product than the other SIF products (see Sect. 2.2). With that said, JPL SIF only relies on OCO-2 SIF data, while SIF products that yield lower R^2 values rely on MODIS reflectance to interpolate the OCO-2 SIF data. The use of MODIS data makes it possible to produce SIF estimates at a much finer spatial and temporal grid than available from JPL SIF (see Table S2), but these products may partly mirror patterns in that MODIS data instead of SIF. We focus on GOSIF in subsequent analyses – because it is one of two SIF products that yield the highest R^2 values.

In a subsequent analysis, we add additional predictor variables to the linear model – environmental data, including precipitation and a function of air temperature (Fig. 1b; see also Sect. 2.2 and Table S4). Vegetation indicators like SIF and EVI are often used as predictors of GPP. However, atmospheric CO_2 observations, like those used in the model-data comparisons here, are influenced by many different types of CO_2 fluxes, including GPP and respiration. The inclusions of environmental data may help the model better describe variability in CO_2 observations caused by respiration and may help describe additional variability in GPP that is not described by the vegetation indicators.

We find that the inclusion of additional predictor variables does little to improve the model-data fit (R^2, Fig. 1b). The result may appear surprising, yet it parallels existing studies of in situ atmospheric CO_2 observations focused on North America (Shiga et al., 2018a; W. Sun et al., 2021). Shiga et al. (2018a) construct a similar linear model of SIF and other vegetation indicators. They find that the inclusion of additional predictor variables yields a better model-data fit in croplands but does little to change model-data fit in other North American biomes. W. Sun et al. (2021) also find that a SIF-based model is as adept at describing variability in atmospheric CO_2 observations as NBE estimates from many of the TRENDY models and from FLUXCOM. In addition, GPP estimates from TRANSCOM, TRENDY, and MsTMIP are often able to match patterns in atmospheric CO_2 observations as well as NBE estimates from these products.

Several factors may help explain why the inclusion of additional predictor variables does little to improve model-data fit. First, Shiga et al. (2018a) note that GPP and NBE are highly correlated in many regions, and this fact may explain why predictors of GPP like SIF are able to explain a large percentage of variability in atmospheric CO_2 observations. Second, this result may also reflect the limits of using OCO-2 observations to constrain NBE. CO_2 observations from OCO-2 are spatially sparse and represent columns averages, and these observations are not as sensitive as many eddy flux or ground-based observations to fine-scale variability in NBE or the individual components of NBE. Hence, this result speaks to the positive ability of SIF to help predict global-scale patterns in NBE, but this result likely also speaks to the limitations of using OCO-2 observations to constrain detailed space-time patterns in NBE.

In the above analyses, we fit the linear model of predictor variables to CO_2 observations from OCO-2. We also conduct a parallel analysis using CO_2 observations from in situ monitoring sites and find similar results (Fig. 2). This similarity indicates that the results are robust to the specific type of the CO_2 observations used in the analysis, and that the results in Fig. 1 are unlikely to be aliased or unduly contaminated by observational errors. Note that much of the analysis in the remainder of the manuscript focuses on results using CO_2 observations from OCO-2 because it provides better data coverage (Fig. S2) across the tropics relative to the in situ observing network (Fig. S3).

We also explore year-to-year differences in model-data fit relative to atmospheric CO_2 observations. The first half of the study period (2015-2016) corresponds to a large El Niño event, whereas the second half of the study period (2017-2018) does not. We find that model-data fit is similar in El Niño versus non El Niño years – within an R^2 of 0.05 for all of the simulations using different vegetation indicators. In other words, no vegetation indicator shows a discernible advantage in describing CO_2 observations during El Niño conditions.

We further find that the linear model using SIF products is a better model-data fit against OCO-2 observations in the extra-tropics than NBE estimates from the TBMs in TRENDY, which do not incorporate SIF (Fig. 3). We specifically compare the correlation (R^2) with OCO-2 observations when we use NBE estimates from 15 different TRENDY models in place of the SIF-based linear model. Note that the vegetation indicators used here have a 16-day temporal resolution while the NBE estimates from TRENDY are available at a monthly resolution, though this fact is unlikely to place the TBMs a noticeable disadvantage (refer to the sensitivity study in SI Sect. S3). Furthermore, the TRENDY models have not been calibrated to CO_2 observations from OCO-2. These facts not withstanding, the results suggest that SIF could be beneficial for improving TBM flux estimates, at least in the extra-tropics, while the SIF products discussed here are unlikely to yield a similar benefit in the tropics.

Note that the inverse modeling analysis described in this section involves several uncertainties that may impact our ability to model variability in atmospheric CO_2 observations. These include uncertain biomass burning and anthropogenic CO_2 emissions, NBE variability due to land use change, possible atmospheric transport errors, and issues related to the atmospheric CO_2 observations (e.g., observational errors or variability in data coverage). Indeed, these challenges are common to inverse modeling studies using atmospheric CO_2 observations. We specifically incorporate biomass burning emissions from GFED and anthropogenic emissions from ODIAC as predictor variables in the inverse model, and these sources are accounted for in all modeling simulations conducted in this study (Sect. 2.2). However, errors in either emissions estimate could lower the overall model-data fit relative to atmospheric CO_2 observations. With that said, we are primarily interested in the relative model-data fit, not absolute model-data fit, of model simulations informed by SIF relative to those informed by other vegetation products. In addition, variability in NBE due to land use change may also be a source of uncertainty. Land use changes that alter GPP should also impact SIF (e.g., del Rosario Uribe & Dukes, 2021; Ding et al., 2021) and should therefore be accounted for in model simulations. With that said, existing studies show that CO_2 observations from OCO-2 can only be used to constrain very broad, seasonal, biome-level variability in NBE (e.g., S. M. Miller et al., 2018; S. M. Miller & Michalak, 2020b). Land use changes that unfold over decades will undoubtedly change biome-level patterns in NBE but may not be detectable across the relatively short, four-year duration of this study. Lastly, atmospheric transport errors and issues related to the CO_2 observations undoubtedly lead to uncertainties in the inverse model. These issues are common to inverse modeling studies and reiterate the importance of setting accurate uncertainties (i.e., covariance matrix parameters) within the inverse model. We discuss the covariance matrix parameters in Sect. S1.

3.2 The extra-tropics

In this section, we estimate NBE using SIF products and other vegetation indicators as predictor variables in the inverse model. We then evaluate how the resulting NBE estimates differ among these different inverse modeling simulations.

We find that incorporating SIF as a predictor variable in the inverse model leads to a different seasonal variability in NBE across the extra-tropics relative to an inverse model that incorporates EVI (Fig. 4). Specifically, NBE estimated using GOSIF show less CO_2 uptake in the fall than NBE estimated using EVI. Furthermore, we see this result in all extra-tropical biomes.

This result, using a global-scale inverse model, parallels studies that compare vegetation indicators against satellite-based GPP products and eddy flux observations. Several satellite-based studies report that the seasonality of greenness is decoupled from the seasonality of GPP in multiple extra-tropical biomes, including both evergreen and deciduous forests (e.g., Walther et al., 2016; Jeong et al., 2017; Luus et al., 2017; Y. Zhang et al., 2020). For example, Jeong et al. (2017) compare the seasonal cycle of SIF, NDVI, and satellite-based estimates of GPP for extratropical forests between $40° - 55°$N latitude globally. They find that the growing season determined by NDVI is $46{\pm}11$ days longer than estimated using SIF. Studies that leverage eddy flux observations reach similar conclusions (e.g., Churkina et al., 2005; Gonsamo et al., 2012).

This seasonal discrepancy is likely because leaves reduce their photosynthetic output during late summer and early autumn. However, their optical properties do not change as quickly, and are unlikely to be detected by greenness indicators like EVI or NDVI (e.g., Jeong et al., 2017). This change is probably caused by a seasonal reduction in incoming solar radiation (e.g., Bauerle et al., 2012). Jeong et al. (2017) specifically find that seasonal changes and SIF and GPP products during fall correlate with changes in incoming shortwave radiation, whereas seasonal changes in greenness indicators like NDVI correlate with changes in temperature (e.g., F. Wang et al., 2020). Furthermore, Jeong et al. (2017) find that temperature and NDVI changes in fall are not linked to GPP and more likely reflect the timing of chlorophyll reduction and leaf drop (Jeong & Medvigy, 2014). Hence, greenness indicators are less effective than SIF at predicting end-of-season changes in GPP. In addition, this decrease in photosynthetic activity could reflect drought stress as soil moisture is depleted through the summer (e.g., P. A. Schwarz et al., 2004). These seasonal changes may not be reflected in greenness indicators (e.g., Goerner et al., 2009). Also, data contamination cannot be ruled out; leaf litter and plant material that has not yet fallen from the plant can increase greenness indicators in the fall, yielding erroneous estimates for the end of growing season (e.g., Gonsamo et al., 2012; Walther et al., 2016).

Despite the differences between SIF and other vegetation indicators, Dechant et al. (2022) propose multiplying greenness indicators by PAR, and they argue that the result may serve as an effective structural proxy for SIF and for photosynthesis. The authors of that study focus on the product of NIRv and PAR (NIRvP) but show that other greenness indicators, when multiplied by PAR, also correlate with SIF at multiple scales – when compared to both tower and satellite observations. The product of greenness indicators and PAR may help overcome the seasonal decoupling between greenness and SIF, as discussed in the previous paragraph. Furthermore, the development of a SIF proxy could hold practical applications. Such a proxy could be used in place of SIF in time periods or locations when SIF is not available. For example, satellites like OCO-2 provide spatially sparse SIF observations, and existing studies assimilate SIF with other vegetation indicators in a machine learning algorithm to create interpolated SIF maps. Other products, like those proposed by Dechant et al. (2022), may serve as a better proxy.

Dechant et al. (2022) also provide a theoretical underpinning for the multiplication of greenness indicators and PAR. They start with equations for GPP and SIF:

$$GPP = APAR \times LUE \tag{5}$$

$$SIF = APAR \times f_{esc} \times \Phi_F \tag{6}$$

where LUE is light use efficiency, f_{esc} is the canopy escape fraction, and Φ_F is the fluorescence emission yield. Dechant et al. (2020) argue that f_{esc} is correlated with LUE at seasonal time scales, at least at the agricultural sites examined, and that f_{esc} therefore plays a key role in the seasonal relationship between SIF and GPP. By contrast, they argue that Φ_F shows poor correlation with LUE. f_{esc} can be approximated by a greenness or vegetation indicator (VI) and fPAR (Zeng et al., 2019):

$$f_{esc} \approx VI/fPAR \tag{7}$$

where fPAR is the fraction of absorbed PAR. Given that $APAR = PAR \times fPAR$ the following relationship should hold:

$$APAR \times f_{esc} \approx VI \times PAR \tag{8}$$

Following this logic a greenness indicator like NIRv, EVI, and/or NDIV, when multiplied by PAR, may be a reasonable structural proxy for SIF.

Indeed, we find that a linear model of NIRv × PAR (NIRvP), EVIP, and NDVIP are just as skilled at matching variability in CO_2 observations compared to a linear model using GOSIF (Figs. 5a-b and S6). In addition, NBE estimated using EVIP exhibits a similar seasonality in the fall relative to results using GOSIF (Fig. 5c). For the large biome-based regions examined here, the product of greenness indicators and PAR may, in fact, be an effective structural proxy for SIF and overcome the seasonal decoupling described earlier in this section. By contrast to the extra-tropics, we find that multiplying vegetation indicators by PAR does relatively little to improve model data fit against OCO-2 observations in the tropics (Figs. 5a-b and S6). Furthermore, these predictors actually worsens model-data fit relative to OCO-2 observations in tropical grasslands (Figs. 5a-b and S6). High levels of PAR can indicate decreased photosynthesis associated with seasonal drought and low levels of PAR can indicate increased photosynthesis associated with seasonal rainfall, a possible reason why PAR worsens model-data fit in tropical grasslands (e.g., Ma et al., 2014). The next section describes inverse modeling results for tropical biomes in depth.

3.3 The tropics

We find that the seasonal cycle of SIF products in tropical biomes is shifted compared to that of NBE estimated by the inverse model. Furthermore, we are unable to reconcile these different seasonal cycles using existing estimates of respiration and biomass burning. This seasonal mismatch may help explain why a linear model of SIF products is not able to explain any more variability in OCO-2 observations relative to other vegetation indicators, a result discussed previously in Sect. 3.1.

This mismatch is apparent in Fig. 6, which compares the seasonal cycle of GOSIF against two NBE estimates in different tropical biomes. In the Southern Hemisphere and in Northern Hemisphere tropical forests, GOSIF indicates an onset of seasonal CO_2 uptake before the two NBE estimates. By contrast, GOSIF predicts peak CO_2 uptake in roughly the same months as the two NBE estimates. Note that the seasonal cycle of each estimate in these panels has been normalized to have a mean of zero and a standard deviation of one in order to make GOSIF directly comparable with the NBE estimates. The blue line in Fig. 6 displays NBE estimated by the inverse model. This inverse model incorporates GOSIF as a predictor variable, yet the the seasonal cycle of NBE looks very different from that of GOSIF. In addition, the red line shows the NBE estimate from ORCHIDEE; of all TBMs in TRENDY, ORCHIDEE exhibits the best model-data fit (R^2) compared to OCO-2 observations in the tropics (see Fig. S9). It is possible, although unlikely, that CO_2 observations from OCO-2 do not provide a unique constraint on the seasonal cycle of NBE in tropical biomes. However, the consistency between the seasonal cycle of the inverse model and of ORCHIDEE suggest otherwise.

We further analyze the seasonal cycle of SIF data from the TROPOMI instrument to see whether the seasonal cycle of SIF from this instrument is in any better agreement with NBE estimates (Fig. 6). Note that the TROPOMI-SIF data do not cover the full range of our study time period (Jan. 2015 to Dec. 2018), and we use a multi-year monthly average (May 2018 to Dec. 2021) in this analysis for each tropical biome. We find that the GOSIF and TROPOMI-SIF have similar seasonal cycles in tropical grasslands and forests, which indicates that possible biases in the seasonal cycle of OCO-2 SIF products are unlikely to be the cause of the seasonal discrepancies in Fig. 6.

We also find that seasonal patterns in respiration and biomass burning cannot reconcile the seasonal differences between SIF products and NBE. We include GOSIF and environmental driver variables in the inverse model to help predict NBE (as in Shiga et al., 2018a, ; Table S4). It is possible that this combination of predictors is skilled at capturing space-time patterns in NBE due to GPP but not due to respiration. To explore this possibility, we re-run the linear model 15 times and each time use a different respiration estimate from TRENDY in place of using environmental driver variables (Fig. S10a). We find that these respiration estimates do little to improve the model-data fit for the SIF-based linear model; the ORCHIDEE model reproduces variability in OCO-2 observations better than any of the linear models using GOSIF and TRENDY respiration estimates (i.e., has a higher R^2), at least in tropical biomes (Fig. S10b).

Similarly, biomass burning cannot reconcile the differing seasonal cycles between SIF products and the NBE estimates. Fig. 6 displays the seasonal cycle of biomass burning emissions from GFED. In this figure, we normalize GFED such that the seasonal cycle is easier to compare against GOSIF and the NBE estimates. In most biomes, the peak in biomass burning has similar timing to the minimum CO_2 uptake predicted by GOSIF. However, this peak is earlier in most tropical biomes than the peak in NBE (i.e., maximum seasonal CO_2 release). Note that all modeling simulations in this study include GFED as a predictor variable.

It is unclear why SIF products are not more skilled predictors of NBE in the tropics relative to other vegetation indicators; there are several possible reasons. First, TBMs disagree on the seasonality of respiration in the tropics, and it is possible that none of the TBMs used here provides a skilled respiration estimate. Indeed, the seasonal mismatch between SIF products and NBE in Fig. 6 is most pronounced around the peak in NBE (i.e., maximum CO_2 release), indicating that deficiencies in the respiration estimates may be at play.

Second, the SIF products used in this study may not be an accurate representation of SIF. SIF retrievals from OCO-2 must be interpolated to fill data gaps and create a continuous, gridded SIF map, and uncertainties in the gap-filling process can impact the accuracy of the resulting SIF products, particularly in the tropics. For example, Yu et al. (2019b) point out that SIF retrievals for some tropical regions (e.g., grasslands and shrublands) exhibit a lower signal-to-noise ratio due to lower overall SIF values in those biomes, and they explain that those biomes could experience rapid changes in photosynthesis that may not be captured by OCO-2 with a 16-day revisit time. Substantial differences among the OCO-2 based SIF products further highlights the uncertainty associated with interpolation and gap-filling. In addition, SIF observations from OCO-2 are sensitive to atmospheric cloud/aerosol contamination or sun–sensor geometry which can confound the real seasonality photosynthesis, particularly in tropical forests (e.g., X. Li, Xiao, He, Arain, et al., 2018; Yu et al., 2019b).

Third, canopy-level and/or remotely sensed SIF may not be a skilled proxy for GPP. This explanation, however, seems less likely given that several site-level studies find good correlation between SIF and GPP in a variety of tropical biomes (e.g., C. Wang et al., 2019; Mengistu et al., 2021). Satellite measurements may not detect all photosynthetic activity in tropical forests, including in the understory, mid-canopy, and the dense canopy. Differences in photosynthesis among these different levels can be key to estimating GPP and NBE of the entire forest; the canopy and understory can have very different seasonal dynamics in tropical forests, dynamics that may not be captured by satellite-based SIF. For example Tang and Dubayah (2017) find that leaf area index in the canopy and understory are anti-correlated in tropical forests, and that dry season leaf loss from the canopy is associated with opportunistic leaf growth in the understory.

4 Conclusion

Remote sensing products like SIF and NIRv have shown enormous promise as predictors of the global carbon cycle. Indeed, we find that existing SIF products are skilled at predicting variability in atmospheric CO_2 observations, and thus in predicting variability in NBE, across the extra-tropics, particularly when compared to other vegetation indicators and to state-of-the-art TBMs that do not assimilate SIF. Specifically, inverse estimates of NBE that assimilate SIF products exhibit a different seasonal cycle, particularly during the fall months when CO_2 uptake by plants may decline more quickly than changes in vegetation greenness. However, we find that other vegetation indicators like NIRv, EVI, and NDVI are just as skilled at predicting patterns in CO_2 observations across large global biomes when we multiply these indicators by PAR, suggesting that NIRvP, EVIP, and NDVIP may, indeed, be reasonable structural proxies for SIF at global scales. By contrast, existing SIF products do not show the same advantage relative to other vegetation indicators in the tropics. Notably, the seasonal cycle of SIF products does not match inverse estimates of NBE nor does it match the seasonal cycle of TBMs that are skilled at predicting patterns in CO_2 observations from OCO-2. We are not able to reconcile this discrepancy using respiration estimates from the 15 TBMs analyzed in this study or using a biomass burning emissions estimate.

Overall, the results suggest that interpolated SIF products can be a powerful tool to improve bottom-up NBE estimates across the global extra-tropics. Specifically, the direct use of SIF within diagonistic TBMs (i.e., those that use forcing data or vegetation characteristics from an external source) could improve the characterization of seasonal variability in GPP and NBE across the extra-tropics, while SIF could serve as an effective tool for evaluating or tuning seasonal variability of GPP in prognostic TBMs (i.e., those that calculate forcing data and vegetation characteristics internally). Indeed, several prognostic TBMs can be used to predict SIF, but these TBMs show wide disagreement on both SIF and GPP, at least at the site level Parazoo et al. (2020). However, this study suggests that there is more work to be done to understand the relationships between SIF, GPP, and NBE in the tropics. We argue that there is a need for more atmospheric CO_2 observations in the tropics that can be used to evaluate relationships between SIF, GPP, and NBE at intermediate regional scales. These include observations from aircraft or tall towers, as in Alden et al. (2016), and/or geostationary satellites like the Geostationary Carbon Cycle Observatory (GeoCarb). Such observations could help bridge the gap between site-level evaluation (e.g., Irteza et al., 2021; C. Wang et al., 2019; Doughty et al., 2019) and global-scale efforts like the present study.

5 Open Research

The atmospheric CO_2 observations from OCO-2 (b10, 10-second averages) and from the NOAA Obspack are publicly available from Baker (2021) and NOAA Global Monitoring Laboratory (2021), respectively. The SIF products are available from Yu et al. (2019a), Y. Zhang et al. (2018), and X. Li and Xiao (2019b); EVI and NDVI are available from NASA MODIS at (Didan, 2022), and the inputs required to calculate NIRv are also available from NASA MODIS (NASA, 2022). Furthermore, the meteorological variables used in this study, including PAR, are available from NASA at NASA Global Modeling and Assimilation Office (2019).

In addition, the inverse modeling simulations in this study use the code published in S. M. Miller and Saibaba (2019).

Acknowledgments
We thank David Baker and Andrew Jacobson for providing the 10 s averaged OCO-2 and Obspack data files, respectively. We also thank Grayson Badgley for his help constructing NIRv and with the research more broadly. We further thank the research teams that generated the OCO-2 SIF products used in this study – led by Longlong Yu, and Yao Zhang. We also thank the Global Carbon Project TRENDY model team, including Stephen Sitch, Pierre Friedlingstein, Emilie Joetzjer, Vladislav Bastrikov, Daniel S. Goll, Vanessa Haverd, Atul K. Jain, Etsushi Kato, Sebastian Lienert, Danica L. Lombardozzi, Patrick C. McGuire, Joe R. Melton, Julia E. M. S. Nabel, Benjamin Poulter, Hanqin Tian, Andrew J. Wiltshire, and Sönke Zaehle. This work is funded by NASA grants 80NSSC18K0976 and 80NSSC21K1073. Jingfeng Xiao is supported by NSF (DEB-2017870).

References

Albright, R., Corbett, A., Jiang, X., Creecy, E., Newman, S., Li, K.-F., . . . Yung, Y. L. (2022). Seasonal variations of solar-induced fluorescence, precipitation, and carbon dioxide over the amazon. *Earth and Space Science*, *9*(1), e2021EA002078. doi: 10.1029/2021EA002078

Alden, C. B., Miller, J. B., Gatti, L. V., Gloor, M. M., Guan, K., Michalak, A. M., . . . Diffenbaugh, N. S. (2016). Regional atmospheric co2 inversion reveals seasonal and geographic differences in amazon net biome exchange. *Global Change Biology*, *22*(10), 3427-3443. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.13305 doi: 10.1111/gcb.13305

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., . . . Zhao, M. (2015). Spatiotemporal patterns of terrestrial gross primary production: A review. *Reviews of Geophysics*, *53*(3), 785-818. doi: 10.1002/2015RG000483

Bacour, C., Maignan, F., MacBean, N., Porcar-Castell, A., Flexas, J., Frankenberg, C., . . . Bastrikov, V. (2019). Improving estimates of gross primary productivity by assimilating solar-induced fluorescence satellite retrievals in a terrestrial biosphere model using a process-based SIF model. *Journal of Geophysical Research: Biogeosciences*, *124*(11), 3281-3306. doi:

https://doi.org/10.1029/2019JG005040
Badgley, G., Anderegg, L. D. L., Berry, J. A., & Field, C. B. (2019). Terrestrial gross primary production: Using NIRv to scale from site to globe. *Global Change Biology*, *25*(11), 3731-3740. doi: https://doi.org/10.1111/gcb.14729
Badgley, G., Field, C. B., & Berry, J. A. (2017). Canopy near-infrared reflectance and terrestrial photosynthesis. *Science Advances*, *3*(3). doi: https://doi.org/10.1126/sciadv.1602244
Baker, D. F. (2021). *OCO-2 b10c 10sec* [dataset]. Retrieved 8 Feb. 2022, from `ftp.cira.colostate.edu:/ftp/BAKER/OCO2_b10c_10sec_GOOD_r5.nc4` doi: 0.1002/essoar.10505688.1
Bauerle, W. L., Oren, R., Way, D. A., Qian, S. S., Stoy, P. C., Thornton, P. E., ... Reynolds, R. F. (2012). Photoperiodic regulation of the seasonal pattern of photosynthetic capacity and the implications for carbon cycling. *Proceedings of the National Academy of Sciences*, *109*(22), 8612-8617. Retrieved from `https://www.pnas.org/doi/abs/10.1073/pnas.1119131109` doi: 10.1073/pnas.1119131109
Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., ... others (2010). Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science*, *329*(5993), 834–838. doi: https://doi.org/10.1126/science.1184984
Byrne, B., Liu, J., Lee, M., Yin, Y., Bowman, K. W., Miyazaki, K., ... Paton-Walsh, C. (2021). The carbon cycle of southeast Australia during 2019–2020: Drought, fires, and subsequent recovery. *AGU Advances*, *2*(4), e2021AV000469. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021AV000469` (e2021AV000469 2021AV000469) doi: 10.1029/2021AV000469
Byrne, B., Wunch, D., Jones, D. B. A., Strong, K., Deng, F., Baker, I., ... Roehl, C. M. (2018). Evaluating GPP and respiration estimates over northern mid-latitude ecosystems using solar-induced fluorescence and atmospheric CO_2 measurements. *Journal of Geophysical Research: Biogeosciences*, *123*(9), 2976-2997. doi: 10.1029/2018JG004472
Carroll, D., Menemenlis, D., Adkins, J. F., Bowman, K. W., Brix, H., Dutkiewicz, S., ... Zhang, H. (2020). The ECCO-Darwin data-assimilative global ocean biogeochemistry model: Estimates of seasonal to multidecadal surface ocean pCO_2 and air-sea CO_2 flux. *Journal of Advances in Modeling Earth Systems*, *12*(10), e2019MS001888. doi: https://doi.org/10.1029/2019MS001888
Chang, Q., Xiao, X., Jiao, W., Wu, X., Doughty, R., Wang, J., ... Qin, Y. (2019). Assessing consistency of spring phenology of snow-covered forests as estimated by vegetation indices, gross primary production, and solar-induced chlorophyll fluorescence. *Agricultural and Forest Meteorology*, *275*, 305-316. doi: https://doi.org/10.1016/j.agrformet.2019.06.002
Chen, A., Mao, J., Ricciuto, D., Xiao, J., Frankenberg, C., Li, X., ... Knapp, A. K. (2021). Moisture availability mediates the relationship between terrestrial gross primary production and solar-induced chlorophyll fluorescence: Insights from global-scale variations. *Global Change Biology*, *27*(6), 1144-1156. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.15373` doi: 10.1111/gcb.15373
Chen, Z., Huntzinger, D. N., Liu, J., Piao, S., Wang, X., Sitch, S., ... Miller, S. M. (2021b). Five years of variability in the global carbon cycle: comparing an estimate from the Orbiting Carbon Observatory-2 and process-based models. *Environmental Research Letters*, *16*(5), 054041. doi: 10.1088/1748-9326/abfac1
Chen, Z., Liu, J., Henze, D. K., Huntzinger, D. N., Wells, K. C., Sitch, S., ... Miller, S. M. (2021a). Linking global terrestrial CO_2 fluxes and environmental drivers: inferences from the Orbiting Carbon Observatory 2 satellite and terrestrial biospheric models. *Atmospheric Chemistry and Physics*, *21*(9), 6663–6680. doi:

10.5194/acp-21-6663-2021
Churkina, G., Schimel, D., Braswell, B. H., & Xiao, X. (2005). Spatial analysis of growing season length control over net ecosystem exchange. *Global Change Biology*, *11*(10), 1777-1787. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2486.2005.001012.x doi: 10.1111/j.1365-2486.2005.001012.x
Damm, A., Guanter, L., Paul-Limoges, E., van der Tol, C., Hueni, A., Buchmann, N., … Schaepman, M. (2015). Far-red sun-induced chlorophyll fluorescence shows ecosystem-specific relationships to gross primary production: An assessment based on observational and modeling approaches. *Remote Sensing of Environment*, *166*, 91-105. doi: https://doi.org/10.1016/j.rse.2015.06.004
Dechant, B., Ryu, Y., Badgley, G., Köhler, P., Rascher, U., Migliavacca, M., … Berry, J. A. (2022). Nirvp: A robust structural proxy for sun-induced chlorophyll fluorescence and photosynthesis across scales. *Remote Sensing of Environment*, *268*, 112763. doi: https://doi.org/10.1016/j.rse.2021.112763
Dechant, B., Ryu, Y., Badgley, G., Zeng, Y., Berry, J. A., Zhang, Y., … Moya, I. (2020). Canopy structure explains the relationship between photosynthesis and sun-induced chlorophyll fluorescence in crops. *Remote Sensing of Environment*, *241*, 111733. Retrieved from https://www.sciencedirect.com/science/article/pii/S0034425720301036 doi: 10.1016/j.rse.2020.111733
del Rosario Uribe, M., & Dukes, J. S. (2021, apr). Land cover change alters seasonal photosynthetic activity and transpiration of amazon forest and cerrado. *Environmental Research Letters*, *16*(5), 054013. Retrieved from https://doi.org/10.1088/1748-9326/abf60d doi: 10.1088/1748-9326/abf60d
Didan, K. (2022). *MODIS Vegetation Index Products (NDVI and EVI)* [dataset]. Retrieved 8 Feb. 2022, from https://modis.gsfc.nasa.gov/data/dataprod/mod13.php doi: 10.5067/MODIS/MOD13C1.006
Ding, Y., Wang, F., Mu, Q., Sun, Y., Cai, H., Zhou, Z., … Shi, H. (2021). Estimating land use/land cover change impacts on vegetation response to drought under 'grain for green' in the loess plateau. *Land Degradation & Development*, *32*(17), 5083-5098. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/ldr.4093 doi: 10.1002/ldr.4093
Doughty, R., Köhler, P., Frankenberg, C., Magney, T. S., Xiao, X., Qin, Y., … Moore, B. (2019). Tropomi reveals dry-season increase of solar-induced chlorophyll fluorescence in the amazon forest. *Proceedings of the National Academy of Sciences*, *116*(44), 22393-22398. Retrieved from https://www.pnas.org/doi/abs/10.1073/pnas.1908157116 doi: 10.1073/pnas.1908157116
Doughty, R., Xiao, X., Köhler, P., Frankenberg, C., Qin, Y., Wu, X., … Moore III, B. (2021). Global-scale consistency of spaceborne vegetation indices, chlorophyll fluorescence, and photosynthesis. *Journal of Geophysical Research: Biogeosciences*, *126*(6), e2020JG006136. doi: https://doi.org/10.1029/2020JG006136
Fang, Y., & Michalak, A. M. (2015). Atmospheric observations inform CO_2 flux responses to enviroclimatic drivers. *Global Biogeochemical Cycles*, *29*(5), 555–566. (2014GB005034) doi: https://doi.org/10.1002/2014GB005034
Fang, Y., Michalak, A. M., Shiga, Y. P., & Yadav, V. (2014). Using atmospheric observations to evaluate the spatiotemporal variability of CO_2 fluxes simulated by terrestrial biospheric models. *Biogeosciences*, *11*(23), 6985–6997. Retrieved from https://bg.copernicus.org/articles/11/6985/2014/ doi: 10.5194/bg-11-6985-2014
Field, C. B., Randerson, J. T., & Malmström, C. M. (1995). Global net primary production: Combining ecology and remote sensing. *Remote Sensing of Environment*, *51*(1), 74-88. (Remote Sensing of Land Surface for Studies of Global Chage) doi: https://doi.org/10.1016/0034-4257(94)00066-V
Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E.,

... others (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, *38*(17). doi: https://doi.org/10.1029/2011GL048738

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., ... Zaehle, S. (2019). Global carbon budget 2019. *Earth System Science Data*, *11*(4), 1783–1838. doi: https://doi.org/10.5194/essd-11-1783-2019

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., ... Zhao, B. (2017). The modern-era retrospective analysis for research and applications, version 2 (merra-2). *Journal of Climate*, *30*(14), 5419 - 5454. Retrieved from `https://journals.ametsoc.org/view/journals/clim/30/14/jcli-d-16-0758.1.xml` doi: 10.1175/JCLI-D-16-0758.1

Giglio, L., Randerson, J. T., & van der Werf, G. R. (2013). Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *Journal of Geophysical Research: Biogeosciences*, *118*(1), 317–328. doi: https://doi.org/10.1002/jgrg.20042

Goerner, A., Reichstein, M., & Rambal, S. (2009). Tracking seasonal drought effects on ecosystem light use efficiency with satellite-based pri in a mediterranean forest. *Remote Sensing of Environment*, *113*(5), 1101-1111. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0034425709000285` doi: 10.1016/j.rse.2009.02.001

Gonsamo, A., Chen, J. M., Price, D. T., Kurz, W. A., & Wu, C. (2012). Land surface phenology from optical satellite measurement and CO_2 eddy covariance technique. *Journal of Geophysical Research: Biogeosciences*, *117*(G3). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JG002070` doi: 10.1029/2012JG002070

Gourdji, S. M., Mueller, K. L., Schaefer, K., & Michalak, A. M. (2008). Global monthly averaged CO_2 fluxes recovered using a geostatistical inverse modeling approach: 2. results including auxiliary environmental data. *Journal of Geophysical Research: Atmospheres*, *113*(D21). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD009733` doi: 10.1029/2007JD009733

Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., ... Michalak, A. M. (2012). North American CO_2 exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion. *Biogeosciences*, *9*(1), 457–475. doi: https://doi.org/10.5194/bg-9-457-2012

Gu, L., Han, J., Wood, J. D., Chang, C. Y.-Y., & Sun, Y. (2019). Sun-induced chl fluorescence and its importance for biophysical modeling of photosynthesis based on light reactions. *New Phytologist*, *223*(3), 1179–1191. doi: https://doi.org/10.1111/nph.15796

Guan, K., Berry, J. A., Zhang, Y., Joiner, J., Guanter, L., Badgley, G., & Lobell, D. B. (2016). Improving the monitoring of crop productivity using spaceborne solar-induced fluorescence. *Global Change Biology*, *22*(2), 716-726. doi: https://doi.org/10.1111/gcb.13136

Guan, K., Pan, M., Li, H., Wolf, A., Wu, J., Medvigy, D., ... Lyapustin, A. I. (2015). Photosynthetic seasonality of global tropical forests constrained by hydroclimate. *Nature Geoscience*, *8*(4), 284–289. doi: https://doi.org/10.1038/ngeo2382

Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P. E., Gómez-Dans, J., Kuze, A., ... Grainger, R. G. (2012). Retrieval and global assessment of terrestrial chlorophyll fluorescence from gosat space measurements. *Remote Sensing of Environment*, *121*, 236-251. doi: https://doi.org/10.1016/j.rse.2012.02.006

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., ... others (2014). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proceedings of the National Academy of Sciences*, *111*(14), E1327–E1333. doi: https://doi.org/10.1073/pnas.1320008111

He, M., Kimball, J. S., Yi, Y., Running, S., Guan, K., Jensco, K., … Maneta, M. (2019, Jul). Impacts of the 2017 flash drought in the US northern plains informed by satellite-based evapotranspiration and solar-induced fluorescence. *Environmental Research Letters*, *14*(7), 074019. doi: https://doi.org/10.1088/1748-9326/ab22c3

Helm, L. T., Shi, H., Lerdau, M. T., & Yang, X. (2020). Solar-induced chlorophyll fluorescence and short-term photosynthetic response to drought. *Ecological Applications*, *30*(5), e02101. doi: 10.1002/eap.2101

Hu, J., Liu, L., Guo, J., Du, S., & Liu, X. (2018). Upscaling solar-induced chlorophyll fluorescence from an instantaneous to daily scale gives an improved estimation of the gross primary productivity. *Remote Sensing*, *10*(10), 1663. doi: https://doi.org/10.3390/rs10101663

Huntzinger, D., Michalak, A., Schwalm, C., Ciais, P., King, A., Fang, Y., … others (2017). Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Scientific reports*, *7*(1), 1–8. doi: https://doi.org/10.1038/s41598-017-03818-2

Huntzinger, D., Post, W., Wei, Y., Michalak, A., West, T., Jacobson, A., … Cook, R. (2012). North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison. *Ecological Modelling*, *232*, 144-157. doi: 10.1016/j.ecolmodel.2012.02.004

Irteza, S. M., Nichol, J. E., Shi, W., & Abbas, S. (2021). Ndvi and fluorescence indicators of seasonal and structural changes in a tropical forest succession. *Earth Systems and Environment*, *5*(1), 127–133. doi: 10.1007/s41748-020-00175-5

Jacobson, A., Schuldt, K., Miller, J., Oda, T., Tans, P., Andrews, A., … Zimnoch, M. (2020, May). *CarbonTracker documentation CT2019 release* [dataset]. Retrieved Dec. 29, 2021, from `https://gml.noaa.gov/ccgg/carbontracker/CT2019/CT2019_doc.php` doi: https://doi.org/10.25925/39m3-6069

Jeong, S.-J., & Medvigy, D. (2014). Macroscale prediction of autumn leaf coloration throughout the continental United States. *Global Ecology and Biogeography*, *23*(11), 1245-1254. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/geb.12206` doi: 10.1111/geb.12206

Jeong, S.-J., Schimel, D., Frankenberg, C., Drewry, D. T., Fisher, J. B., Verma, M., … Joiner, J. (2017). Application of satellite solar-induced chlorophyll fluorescence to understanding large-scale variations in vegetation phenology and function over northern high latitude forests. *Remote Sensing of Environment*, *190*, 178-187. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0034425716304680` doi: 10.1016/j.rse.2016.11.021

Jiang, Z., Jones, D. B. A., Worden, H. M., Deeter, M. N., Henze, D. K., Worden, J., … Schuck, T. J. (2013). Impact of model errors in convective transport on co source estimates inferred from MOPITT CO retrievals. *Journal of Geophysical Research: Atmospheres*, *118*(4), 2073-2083. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50216` doi: 10.1002/jgrd.50216

Jiao, W., Chang, Q., & Wang, L. (2019). The sensitivity of satellite solar-induced chlorophyll fluorescence to meteorological drought. *Earth's Future*, *7*(5), 558-573. doi: https://doi.org/10.1029/2018EF001087

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., … Reichstein, M. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific data*, *6*(1), 1–14. doi: 10.1038/s41597-019-0076-8

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., … Reichstein, M. (2020). Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. *Biogeosciences*, *17*(5), 1343–1365. Retrieved from `https://bg.copernicus.org/articles/`

17/1343/2020/ doi: 10.5194/bg-17-1343-2020

Kim, J., Ryu, Y., Dechant, B., Lee, H., Kim, H. S., Kornfeld, A., & Berry, J. A. (2021). Solar-induced chlorophyll fluorescence is non-linearly related to canopy photosynthesis in a temperate evergreen needleleaf forest during the fall transition. *Remote Sensing of Environment*, *258*, 112362. doi: https://doi.org/10.1016/j.rse.2021.112362

Kitanidis, P. K., & Vomvoris, E. G. (1983). A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resour. Res.*, *19*(3), 677-690. doi: https://doi.org/10.1029/WR019i003p00677

Köhler, P., Frankenberg, C., Magney, T. S., Guanter, L., Joiner, J., & Landgraf, J. (2018). Global retrievals of solar-induced chlorophyll fluorescence with TROPOMI: First results and intersensor comparison to OCO-2. *Geophysical Research Letters*, *45*(19), 10–456. doi: https://doi.org/10.1029/2018GL079031

Li, X., & Xiao, J. (2019a). A global, 0.05-degree product of solar-induced chlorophyll fluorescence derived from OCO-2, MODIS, and reanalysis data. *Remote Sensing*, *11*(5). doi: 10.3390/rs11050517

Li, X., & Xiao, J. (2019b). *GOSIF - global, OCO-2 based SIF product, v2* [dataset]. Global Ecology Data Repository. Retrieved 9 Sep. 2021, from https://globalecology.unh.edu/data/GOSIF.html doi: 10.3390/rs11050517

Li, X., & Xiao, J. (2022). TROPOMI observations allow for robust exploration of the relationship between solar-induced chlorophyll fluorescence and terrestrial gross primary production. *Remote Sensing of Environment*, *268*, 112748. doi: https://doi.org/10.1016/j.rse.2021.112748

Li, X., Xiao, J., & He, B. (2018). Chlorophyll fluorescence observed by OCO-2 is strongly related to gross primary productivity estimated from flux towers in temperate forests. *Remote Sensing of Environment*, *204*, 659-671. doi: https://doi.org/10.1016/j.rse.2017.09.034

Li, X., Xiao, J., He, B., Arain, M. A., Beringer, J., Desai, A. R., ... others (2018). Solar-induced chlorophyll fluorescence is strongly correlated with terrestrial photosynthesis for a wide variety of biomes: First global analysis based on OCO-2 and flux tower observations. *Global change biology*, *24*(9), 3990–4008. doi: https://doi.org/10.1111/gcb.14297

Li, Z., Zhang, Q., Li, J., Yang, X., Wu, Y., Zhang, Z., ... Zhang, Y. (2020). Solar-induced chlorophyll fluorescence and its link to canopy photosynthesis in maize from continuous ground measurements. *Remote Sensing of Environment*, *236*, 111420. doi: https://doi.org/10.1016/j.rse.2019.111420

Liu, J., Bowman, K. W., Schimel, D. S., Parazoo, N. C., Jiang, Z., Lee, M., ... Eldering, A. (2017). Contrasting carbon cycle responses of the tropical continents to the 2015–2016 El Niño. *Science*, *358*(6360), eaam5690. Retrieved from https://www.science.org/doi/abs/10.1126/science.aam5690 doi: 10.1126/science.aam5690

Liu, J., Wennberg, P. O., Parazoo, N. C., Yin, Y., & Frankenberg, C. (2020). Observational constraints on the response of high-latitude northern forests to warming. *AGU Advances*, *1*(4), e2020AV000228. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020AV000228 (e2020AV000228 2020AV000228) doi: 10.1029/2020AV000228

Luus, K. A., Commane, R., Parazoo, N. C., Benmergui, J., Euskirchen, E. S., Frankenberg, C., ... Lin, J. C. (2017). Tundra photosynthesis captured by satellite-observed solar-induced chlorophyll fluorescence. *Geophysical Research Letters*, *44*(3), 1564-1573. doi: 10.1002/2016GL070842

Ma, X., Huete, A., Yu, Q., Restrepo-Coupe, N., Beringer, J., Hutley, L. B., ... Eamus, D. (2014). Parameterization of an ecosystem light-use-efficiency model for predicting savanna GPP using MODIS EVI. *Remote Sensing of Environment*, *154*, 253-271. Retrieved from https://www.sciencedirect.com/science/article/pii/S0034425714003228 doi: 10.1016/j.rse.2014.08.025

MacBean, N., Maignan, F., Bacour, C., Lewis, P., Peylin, P., Guanter, L., . . . Disney, M. (2018). Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data. *Scientific reports*, *8*(1), 1–12. doi: https://doi.org/10.1038/s41598-018-20024-w

Madani, N., Parazoo, N. C., & Miller, C. E. (2022). Climate change is enforcing physiological changes in arctic ecosystems. *Nature Communications Earth & Environment*, *0*(0), in review.

Magney, T. S., Barnes, M. L., & Yang, X. (2020). On the covariation of chlorophyll fluorescence and photosynthesis across scales. *Geophysical Research Letters*, *47*(23), e2020GL091098. doi: 10.1029/2020GL091098

Magney, T. S., Bowling, D. R., Logan, B. A., Grossmann, K., Stutz, J., Blanken, P. D., . . . Frankenberg, C. (2019). Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence. *Proceedings of the National Academy of Sciences*, *116*(24), 11640-11645. doi: 10.1073/pnas.1900278116

Magney, T. S., Frankenberg, C., Fisher, J. B., Sun, Y., North, G. B., Davis, T. S., . . . Siebke, K. (2017). Connecting active to passive fluorescence with photosynthesis: A method for evaluating remote sensing measurements of chl fluorescence. *New phytologist*, *215*(4), 1594–1608. doi: https://doi.org/10.1111/nph.14662

Magney, T. S., Frankenberg, C., Köhler, P., North, G., Davis, T. S., Dold, C., . . . others (2019). Disentangling changes in the spectral shape of chlorophyll fluorescence: Implications for remote sensing of photosynthesis. *Journal of Geophysical Research: Biogeosciences*, *124*(6), 1491–1507. doi: 10.1029/2019JG005029

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., . . . Gottlieb, E. W. (2008). A satellite-based biosphere parameterization for net ecosystem CO_2 exchange: Vegetation photosynthesis and respiration model (VPRM). *Global Biogeochemical Cycles*, *22*(2), GB2005. doi: https://doi.org/10.1029/2006GB002735

Marrs, J. K., Reblin, J. S., Logan, B. A., Allen, D. W., Reinmann, A. B., Bombard, D. M., . . . Hutyra, L. R. (2020). Solar-induced fluorescence does not track photosynthetic carbon assimilation following induced stomatal closure. *Geophysical Research Letters*, *47*(15), e2020GL087956. (e2020GL087956 2020GL087956) doi: https://doi.org/10.1029/2020GL087956

Mengistu, A. G., Mengistu Tsidu, G., Koren, G., Kooreman, M. L., Boersma, K. F., Tagesson, T., . . . Peters, W. (2021). Sun-induced fluorescence and near-infrared reflectance of vegetation track the seasonal dynamics of gross primary production over africa. *Biogeosciences*, *18*(9), 2843–2857. Retrieved from https://bg.copernicus.org/articles/18/2843/2021/ doi: 10.5194/bg-18-2843-2021

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

Michalak, A. M., Bruhwiler, L., & Tans, P. P. (2004). A geostatistical approach to surface flux estimation of atmospheric trace gases. *Journal of Geophysical Research: Atmospheres*, *109*(D14). doi: https://doi.org/10.1029/2003JD004422

Miller, S., Worthy, D. J., Michalak, A., Wofsy, S., Kort, E., Havice, T., . . . Zhang, B. (2014a). Observational constraints on the distribution, seasonality, and environmental predictors of north american boreal methane emissions. *Global Biogeochemical Cycles*, *28*(2), 146-160. doi: 10.1002/2013GB004580

Miller, S. M., Matross, D. M., Andrews, A. E., Millet, D. B., Longo, M., Gottlieb, E. W., ... Wofsy, S. C. (2008). Sources of carbon monoxide and formaldehyde in North America determined from high-resolution atmospheric data. *Atmospheric Chemistry and Physics*, *8*(24), 7673–7696. Retrieved from https://acp.copernicus.org/articles/8/7673/2008/ doi: 10.5194/acp-8-7673-2008

Miller, S. M., & Michalak, A. M. (2020b). The impact of improved satellite retrievals on estimates of biospheric carbon balance. *Atmospheric Chemistry and Physics*, *20*(1), 323–331. doi: 10.5194/acp-20-323-2020

Miller, S. M., Michalak, A. M., Yadav, V., & Tadić, J. M. (2018). Characterizing biospheric carbon balance using CO_2 observations from the OCO-2 satellite. *Atmospheric Chemistry and Physics*, *18*(9), 6785–6799. doi: 10.5194/acp-18-6785-2018

Miller, S. M., Miller, C. E., Commane, R., Chang, R. Y., Dinardo, S. J., Henderson, J. M., ... Michalak, A. M. (2016a). A multiyear estimate of methane fluxes in alaska from carve atmospheric observations. *Global Biogeochemical Cycles*, *30*(10), 1441-1453. doi: 10.1002/2016GB005419

Miller, S. M., & Saibaba, A. (2019). *Geostatistical inverse modeling with large atmospheric datasets.* Zenodo. doi: 10.5281/zenodo.3595574

Miller, S. M., Saibaba, A. K., Trudeau, M. E., Mountain, M. E., & Andrews, A. E. (2020a). Geostatistical inverse modeling with very large datasets: an example from the Orbiting Carbon Observatory 2 (OCO-2) satellite. *Geoscientific Model Development*, *13*(3), 1771–1785. doi: 10.5194/gmd-13-1771-2020

NASA. (2000). *Normalized difference vegetation index (ndvi).* Retrieved 15 July 2022, from https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php

NASA. (2022). *MCD43C3 - MODIS/Terra+Aqua albedo 16-day L3 global 0.05deg CMG* [dataset]. Retrieved 8 Feb. 2022, from https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MCD43C3/ doi: 10.5067/MODIS/MCD43C3.006

NASA Global Modeling and Assimilation Office. (2019). *Modern-Era Retrospective analysis for Research and Applications, version 2* [dataset]. Retrieved 8 Feb. 2022, from https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/ doi: 10.5067/L0T5GEG1NYFA

NOAA Global Monitoring Laboratory. (2021). *Observation package (ObsPack) data products, v3.2* [dataset]. Retrieved Dec. 29, 2021, from https://gml.noaa.gov/ccgg/obspack/data.php?id=obspack_co2_1_OCO2MIP_v3.2_2021-05-20 doi: https://doi.org/10.25925/20210519

Oda, T., Maksyutov, S., & Andres, R. J. (2018). The Open-source Data Inventory for Anthropogenic CO_2, version 2016 (ODIAC2016): a global monthly fossil fuel CO_2 gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth System Science Data*, *10*(1), 87–107. doi: https://doi.org/10.5194/essd-10-87-2018

Parazoo, N. C., Bowman, K., Frankenberg, C., Lee, J.-E., Fisher, J. B., Worden, J., ... Gerbig, C. (2013). Interpreting seasonal changes in the carbon balance of southern amazonia using measurements of XCO_2 and chlorophyll fluorescence from gosat. *Geophysical Research Letters*, *40*(11), 2829-2833. doi: 10.1002/grl.50452

Parazoo, N. C., Magney, T., Norton, A., Raczka, B., Bacour, C., Maignan, F., ... Frankenberg, C. (2020). Wide discrepancies in the magnitude and direction of modeled solar-induced chlorophyll fluorescence in response to light conditions. *Biogeosciences*, *17*(13), 3733–3755. Retrieved from https://bg.copernicus.org/articles/17/3733/2020/ doi: 10.5194/bg-17-3733-2020

Peiro, H., Crowell, S., Schuh, A., Baker, D. F., O'Dell, C., Jacobson, A. R., ... Baker, I. (2022). Four years of global carbon cycle observed from the orbiting carbon observatory 2 (OCO-2) version 9 and in situ data and comparison to OCO-2 version 7. *Atmospheric Chemistry and Physics*, *22*(2), 1097–1130. Retrieved from https://acp.copernicus.org/articles/22/1097/2022/ doi: 10.5194/acp-22-1097-2022

Pierrat, Z., Magney, T., Parazoo, N. C., Grossmann, K., Bowling, D. R., Seibt, U., … Stutz, J. (2022). Diurnal and seasonal dynamics of solar-induced chlorophyll fluorescence, vegetation indices, and gross primary productivity in the boreal forest. *Journal of Geophysical Research: Biogeosciences*, *127*(2), e2021JG006588. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JG006588 (e2021JG006588 2021JG006588) doi: 10.1029/2021JG006588

Pierrat, Z., Nehemy, M. F., Roy, A., Magney, T., Parazoo, N. C., Laroque, C., … Stutz, J. (2021). Tower-based remote sensing reveals mechanisms behind a two-phased spring transition in a mixed-species boreal forest. *Journal of Geophysical Research: Biogeosciences*, *126*(5), e2020JG006191. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JG006191 (e2020JG006191 2020JG006191) doi: 10.1029/2020JG006191

Saibaba, A. K., & Kitanidis, P. K. (2015). Fast computation of uncertainty quantification measures in the geostatistical approach to solve inverse problems. *Advances in Water Resources*, *82*(0), 124-138. doi: https://doi.org/10.1016/j.advwatres.2015.04.012

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461 – 464. doi: https://doi.org/10.1214/aos/1176344136

Schwarz, P. A., Law, B. E., Williams, M., Irvine, J., Kurpius, M., & Moore, D. (2004). Climatic versus biotic constraints on carbon and water fluxes in seasonally drought-affected ponderosa pine ecosystems. *Global Biogeochemical Cycles*, *18*(4). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004GB002234 doi: 10.1029/2004GB002234

Shekhar, A., Chen, J., Bhattacharjee, S., Buras, A., Castro, A. O., Zang, C. S., & Rammig, A. (2020). Capturing the impact of the 2018 european drought and heat across different vegetation types using OCO-2 solar-induced fluorescence. *Remote Sensing*, *12*(19). doi: https://doi.org/10.3390/rs12193249

Shiga, Y. P., Michalak, A. M., Fang, Y., Schaefer, K., Andrews, A. E., Huntzinger, D. H., … Wei, Y. (2018a). Forests dominate the interannual variability of the North American carbon sink. *Environ. Res. Lett.*, *13*(8), 084015. doi: https://doi.org/10.1088/1748-9326/aad505

Shiga, Y. P., Michalak, A. M., Fang, Y., Schaefer, K., Andrews, A. E., Huntzinger, D. H., … Wei, Y. (2018b, aug). Forests dominate the interannual variability of the North American carbon sink. *Environmental Research Letters*, *13*(8), 084015. doi: 10.1088/1748-9326/aad505

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlstrom, A., … Myneni, R. (2015). Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, *12*(3), 653–679. doi: https://doi.org/10.5194/bg-12-653-2015

Sun, W., Fang, Y., Luo, X., Shiga, Y. P., Zhang, Y., Andrews, A. E., … Michalak, A. M. (2021). Midwest US croplands determine model divergence in North American carbon fluxes. *AGU Advances*, *2*(2), e2020AV000310. doi: 10.1029/2020AV000310

Sun, Y., Frankenberg, C., Jung, M., Joiner, J., Guanter, L., Köhler, P., & Magney, T. (2018). Overview of solar-induced chlorophyll fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP. *Remote Sensing of Environment*, *209*, 808-823. doi: https://doi.org/10.1016/j.rse.2018.02.016

Tang, H., & Dubayah, R. (2017). Light-driven growth in amazon evergreen forests explained by seasonal variations of vertical canopy structure. *Pro-*

ceedings of the National Academy of Sciences, *114*(10), 2640-2644. doi: https://doi.org/10.1073/pnas.1616943114

Thum, T., Zaehle, S., Köhler, P., Aalto, T., Aurela, M., Guanter, L., … Markkanen, T. (2017). Modelling sun-induced fluorescence and photosynthesis with a land surface model at local and regional scales in northern europe. *Biogeosciences*, *14*(7), 1969–1987. doi: https://doi.org/10.5194/bg-14-1969-2017

USGS. (2022). *Landsat enhanced vegetation index.* Retrieved 15 July 2022, from `https://www.usgs.gov/landsat-missions/landsat-enhanced-vegetation-index#:~:text=Landsat%20Enhanced%20Vegetation%20Index%20(EVI,in%20areas%20with%20dense%20vegetation.`

Verma, M., Schimel, D., Evans, B., Frankenberg, C., Beringer, J., Drewry, D., … Eldering, A. (2017). Effect of environmental conditions on the relationship between solar-induced fluorescence and gross primary productivity at an ozflux grassland site. *Journal of Geophysical Research: Biogeosciences*, *122*. doi: https://doi.org/10.1002/2016JG003580

Walther, S., Voigt, M., Thum, T., Gonsamo, A., Zhang, Y., Köhler, P., … Guanter, L. (2016). Satellite chlorophyll fluorescence measurements reveal large-scale decoupling of photosynthesis and greenness dynamics in boreal evergreen forests. *Global Change Biology*, *22*(9), 2979-2996. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.13200` doi: 10.1111/gcb.13200

Wang, C., Beringer, J., Hutley, L. B., Cleverly, J., Li, J., Liu, Q., & Sun, Y. (2019). Phenology dynamics of dryland ecosystems along the north Australian tropical transect revealed by satellite solar-induced chlorophyll fluorescence. *Geophysical Research Letters*, *46*(10), 5294-5302. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL082716` doi: 10.1029/2019GL082716

Wang, F., Chen, B., Lin, X., & Zhang, H. (2020). Solar-induced chlorophyll fluorescence as an indicator for determining the end date of the vegetation growing season. *Ecological Indicators*, *109*, 105755. doi: https://doi.org/10.1016/j.ecolind.2019.105755

Wang, S., Zhang, Y., Ju, W., Qiu, B., & Zhang, Z. (2021). Tracking the seasonal and inter-annual variations of global gross primary production during last four decades using satellite near-infrared reflectance data. *Science of The Total Environment*, *755*, 142569. doi: https://doi.org/10.1016/j.scitotenv.2020.142569

Wen, J., Köhler, P., Duveiller, G., Parazoo, N., Magney, T., Hooker, G., … Sun, Y. (2020). A framework for harmonizing multiple satellite instruments to generate a long-term global high spatial-resolution solar-induced chlorophyll fluorescence (SIF). *Remote Sensing of Environment*, *239*, 111644. doi: https://doi.org/10.1016/j.rse.2020.111644

Wood, J. D., Griffis, T. J., Baker, J. M., Frankenberg, C., Verma, M., & Yuen, K. (2017). Multiscale analyses of solar-induced florescence and gross primary production. *Geophysical Research Letters*, *44*(1), 533–541. doi: https://doi.org/10.1002/2016GL070775

Yang, X., Tang, J., Mustard, J. F., Lee, J.-E., Rossini, M., Joiner, J., … Richardson, A. D. (2015). Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. *Geophysical Research Letters*, *42*(8), 2977-2987. doi: https://doi.org/10.1002/2015GL063201

Yu, L., Wen, J., Chang, C. Y., Frankenberg, C., & Sun, Y. (2019a). *High resolution global contiguous SIF estimates derived from OCO-2 SIF and MODIS* [dataset]. ORNL Distributed Active Archive Center. Retrieved 9 Sep. 2021, from `https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1696` doi: 10.3334/ORNLDAAC/1696

Yu, L., Wen, J., Chang, C. Y., Frankenberg, C., & Sun, Y. (2019b). High-resolution global contiguous SIF of OCO-2. *Geophysical Research Letters*, *46* (3), 1449-1458. doi: 10.1029/2018GL081109

Zeng, Y., Badgley, G., Dechant, B., Ryu, Y., Chen, M., & Berry, J. (2019). A practical approach for estimating the escape ratio of near-infrared solar-induced chlorophyll fluorescence. *Remote Sensing of Environment*, *232*, 111209. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0034425719302226` doi: 10.1016/j.rse.2019.05.028

Zhang, J., Xiao, J., Tong, X., Zhang, J., Meng, P., Li, J., ... Yu, P. (2022). NIRv and SIF better estimate phenology than NDVI and EVI: Effects of spring and autumn phenology on ecosystem production of planted forests. *Agricultural and Forest Meteorology*, *315*, 108819. doi: 10.1016/j.agrformet.2022.108819

Zhang, L., Qiao, N., Huang, C., & Wang, S. (2019). Monitoring drought effects on vegetation productivity using satellite solar-induced chlorophyll fluorescence. *Remote Sensing*, *11* (4). doi: https://doi.org/10.3390/rs11040378

Zhang, Y., Commane, R., Zhou, S., Williams, A. P., & Gentine, P. (2020). Light limitation regulates the response of autumn terrestrial carbon uptake to warming. *Nature Climate Change*, *10* (8), 739–743. doi: 10.1038/s41558-020-0806-0

Zhang, Y., Joiner, J., Alemohammad, S. H., Zhou, S., & Gentine, P. (2018). A global spatially contiguous solar-induced fluorescence (CSIF) dataset using neural networks. *Biogeosciences*, *15* (19), 5779–5800. doi: 10.5194/bg-15-5779-2018

Zuromski, L. M., Bowling, D. R., Köhler, P., Frankenberg, C., Goulden, M. L., Blanken, P. D., & Lin, J. C. (2018). Solar-induced fluorescence detects interannual variation in gross primary production of coniferous forests in the western united states. *Geophysical Research Letters*, *45* (14), 7184-7193. doi: https://doi.org/10.1029/2018GL077906

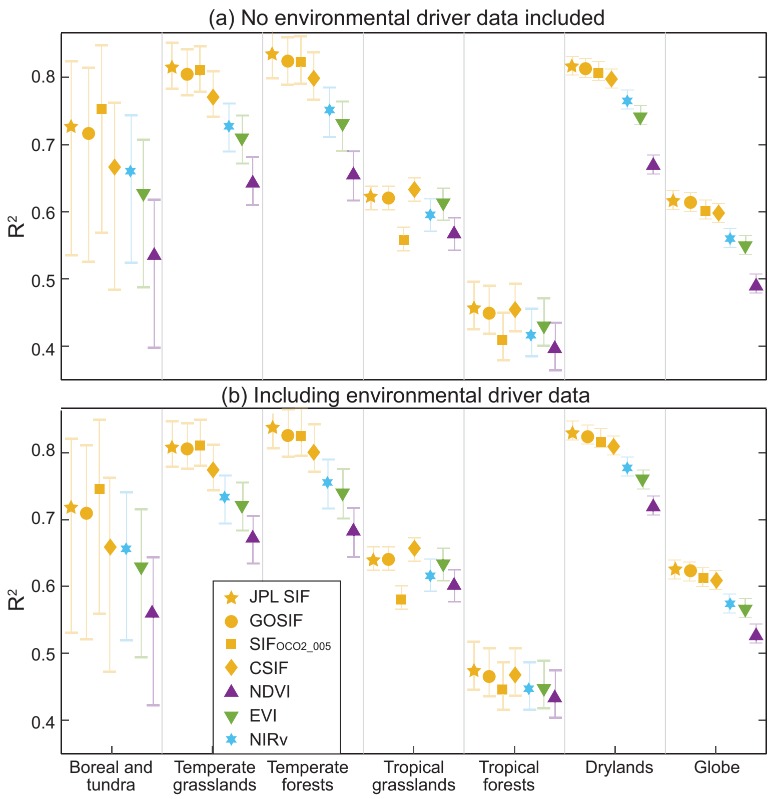

Figure 1. Results from the linear model using vegetation indicators as predictor variables (a) and using vegetation indicators plus environmental driver variables (b). This figure specifically shows the linear model fit (R^2) when compared against CO_2 observations from OCO-2. Overall, we find that SIF products yield a better model-data fit (R^2) compared to other vegetation indicators across the extra-tropics, but SIF products do not exhibit the same advantage in the tropics. We also find that the inclusion of additional predictor variables to help better describe variability in NBE (panel b) does not substantially improve or otherwise change the model-data fit. Note that we combine boreal and tundra biomes for OCO-2 simulations, due to the paucity of OCO-2 observations over the tundra.

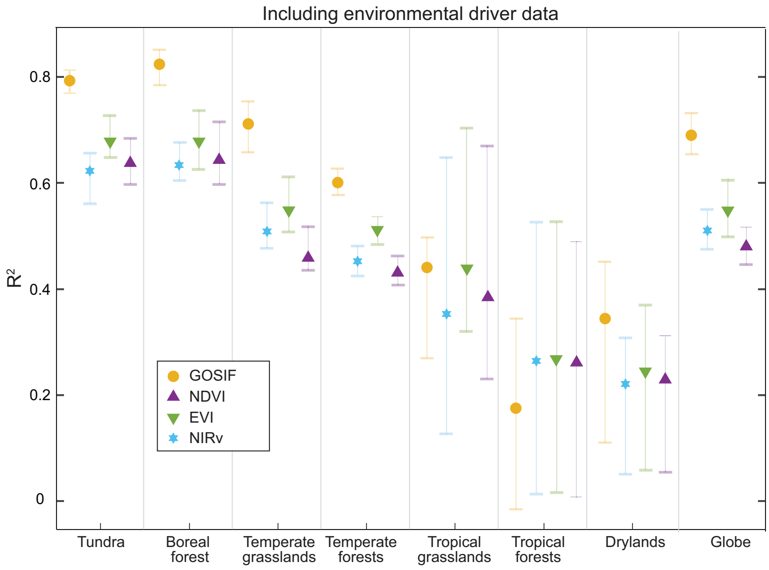

Figure 2. Results from the linear model using in situ CO_2 observations instead of CO_2 observations from OCO-2. The linear model results using in situ CO_2 observations broadly parallel results using OCO-2 observations. Notably, a linear model using GOSIF yields a better fit to in situ CO_2 observations than other vegetation indicators, at least in the extra-tropics.

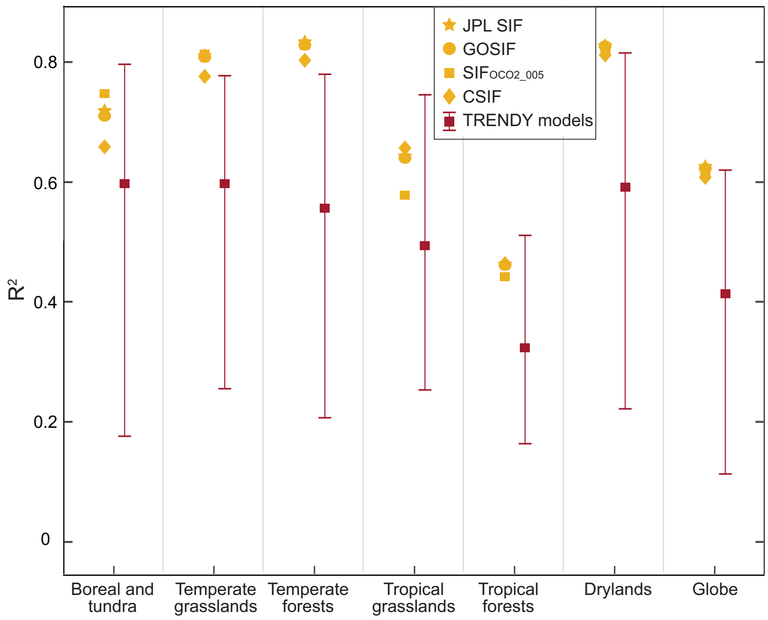

Figure 3. A comparison between the SIF-based linear model and NBE estimates from 15 bottom-up models in TRENDY. This figure displays the model-data fit (R^2) of the linear model and TRENDY models against CO_2 observations from OCO-2. In several biomes (temperate grasslands, temperate forests, and drylands), the SIF-based results are a better fit than the TRENDY models, which do not assimilate SIF. By contrast, in other biomes (e.g., tropical biomes), several TRENDY models are a better fit than the SIF-based linear model. Note that the red box indicates the mean R^2 value of the 15 TRENDY models and the vertical bar is the range of R^2 values from the 15 models.

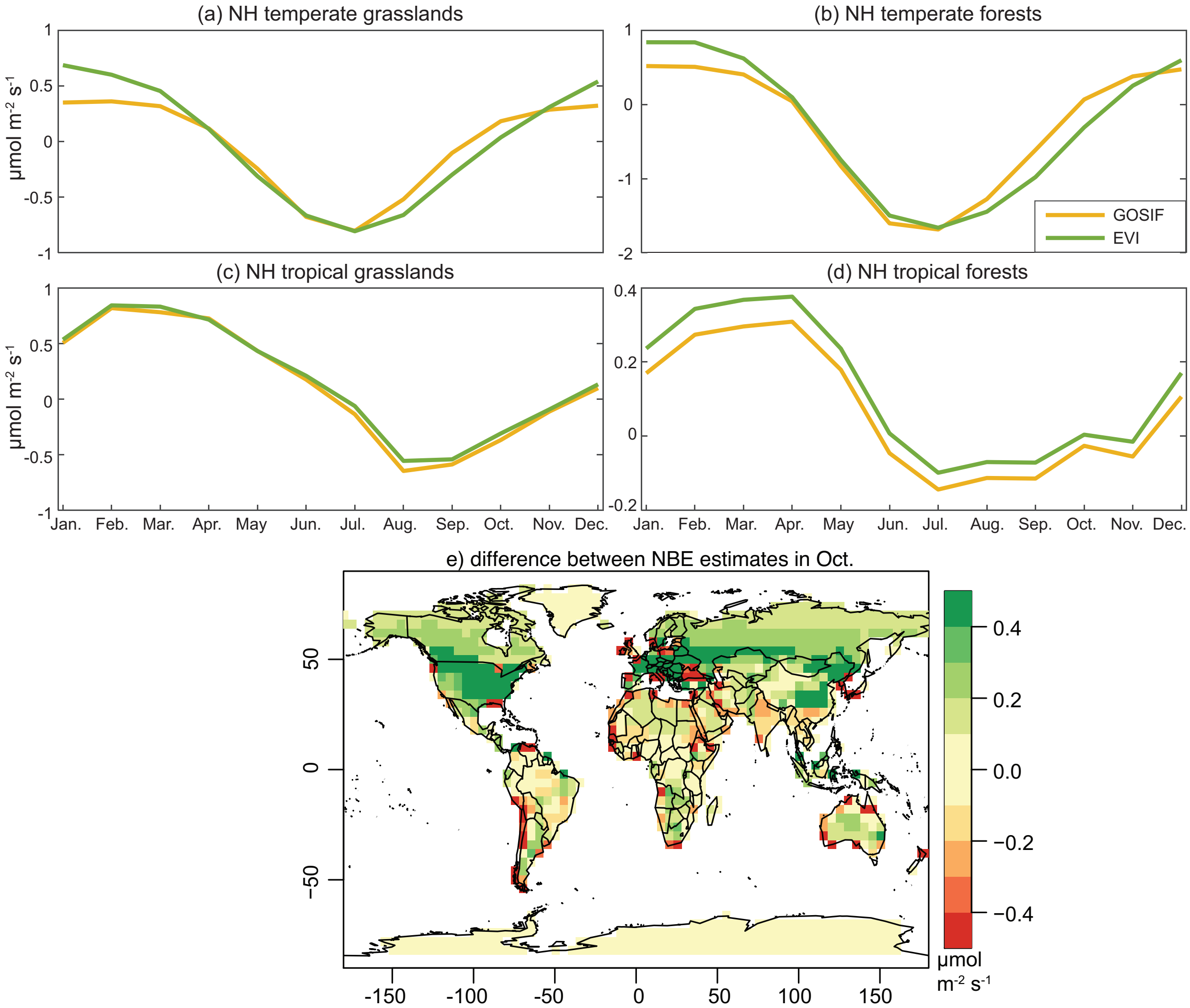

Figure 4. Estimated NBE from inverse modeling simulations that use GOSIF (yellow) and EVI (green) as predictor variables. Panels a-d compare the seasonal cycle. Inverse modeling simulations that incorporate GOSIF yield an different seasonal cycle in the extra-tropics relative to simulations using EVI (panels a and b). Specifically, CO_2 uptake during northern hemisphere fall declines more quickly in the SIF simulations. By contrast, results for tropical biomes (panels c and d) show little difference between the two inverse modeling simulations. In addition, panel e compares spatial patterns in estimated NBE during October (i.e., GOSIF simulations minus EVI simulations), a month when the estimates yield different seasonal patterns across the northern extra-tropics. Green colors indicate greater CO_2 uptake (or less CO_2 release to the atmosphere) in simulations using EVI compared to those using GOSIF. Overall, panel e indicates broad differences between the NBE estimates across the northern extra-tropics.

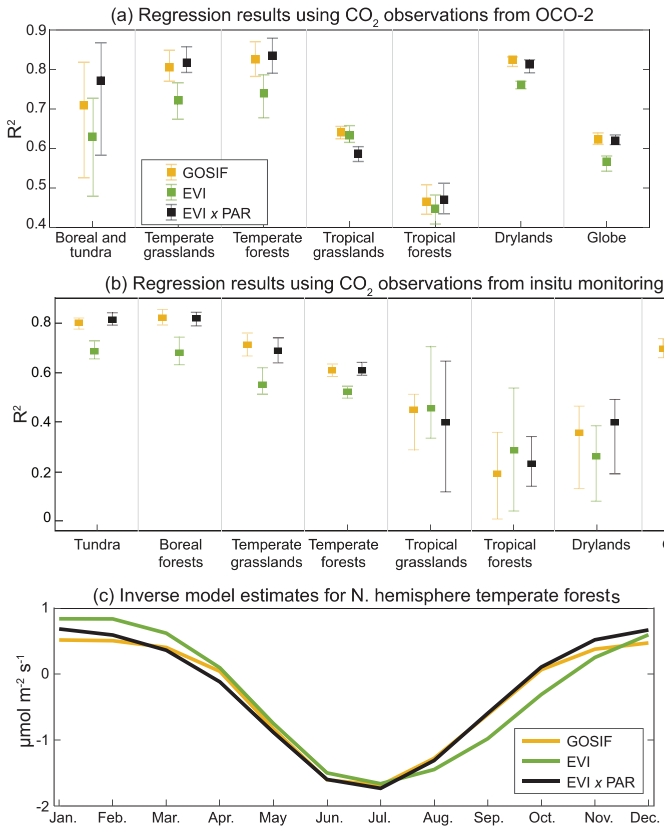

Figure 5. The fit (R^2) against OCO-2 observations using a linear model of GOSIF, EVI, and EVI $\times$ PAR as predictor variables. Panel (c) displays NBE estimated by the inverse model when using GOSIF, EVI and EVIP as predictor variables. Note that panels (a) and (c) assimilate CO_2 observations from OCO-2 while panel (b) shows the results of analysis using CO_2 observations from in situ CO_2 monitoring sites. Across all simulations, we find that EVI, when multiplied by PAR, is as skillful a predictor of NBE in the extra-tropics as SIF. Furthermore, NBE estimated in the inverse model using EVI $\times$ PAR as a predictor variable exhibits a similar seasonal cycle as NBE estimated using GOSIF as a predictor.

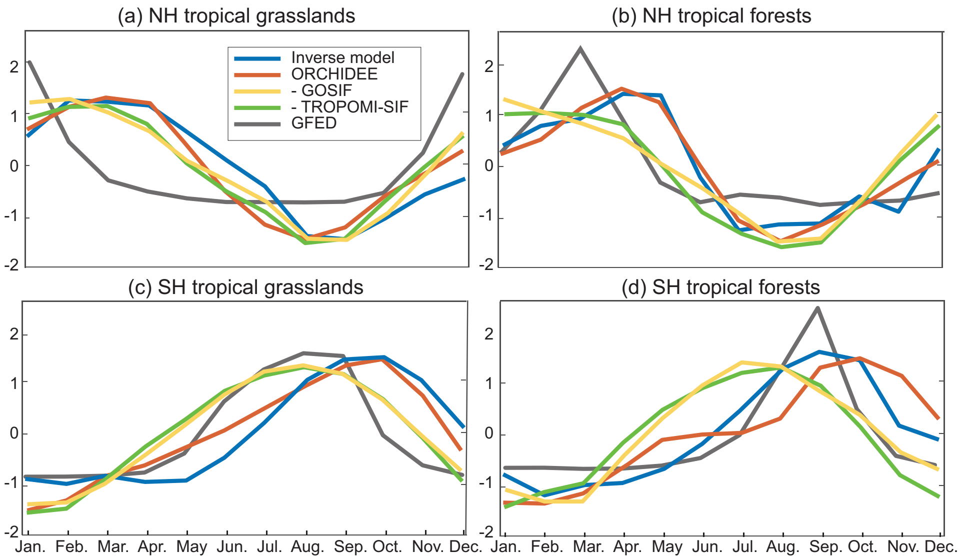

Figure 6. The seasonal cycle of NBE from the inverse model (blue), the ORCHIDEE model (red), GOSIF (yellow), TROPOMI-SIF (green), and GFED (charcoal) in different tropical biomes. Each product has been normalized to have a mean of zero and standard deviation of one for easier comparison. SIF in the tropics is out of phase with the seasonal cycle of the inverse modeling estimate and with ORCHIDEE, a bottom-up model that is more skillful at predicting CO_2 observations from OCO-2 relative to other TRENDY models.