Optimizing the isoprene emission model MEGAN with satellite and ground-based observational constraints

Christian A. DiMaria[1], Dylan B. A. Jones[1], Helen M Worden[2], A. Anthony Bloom[3], Kevin W. Bowman[4], Trissevgeni Stavrakou[5], Kazuyuki Miyazaki[6], John R. Worden[7], Alex B. Guenther[8], Chinmoy Sarkar[9], Roger Seco[10], Jeonghoo Park[11], Julio Tota[12], Eliane Gomes Alvez[13], and Valerio Ferracci[14]

[1]University of Toronto
[2]National Center for Atmospheric Research (UCAR)
[3]Jet Propulsion Laboratory, California Institute of Technology
[4]Jet Propulsion Lab (NASA)
[5]Royal Belgian Institute for Space Aeronomy
[6]Jet Propulsion Laboratory
[7]JPL / Caltech
[8]University of California, Irvine
[9]University of California Irvine
[10]Institute of Environmental Assessment and Water Research (IDAEA-CSIC)
[11]National Institute of Environmental Research
[12]Universidade Federal do Oeste do Para
[13]Max Planck Institute for Biogeochemistry
[14]Cranfield University

November 22, 2022

Abstract

Isoprene is a hydrocarbon emitted in large quantities by terrestrial vegetation. It is a precursor to several air quality and climate pollutants including ozone. Emission rates vary with plant species and environmental conditions. This variability can be modelled using the Model of Emissions of Gases and Aerosols from Nature (MEGAN). MEGAN parameterizes isoprene emission rates as a vegetation-specific standard rate which is modulated by scaling factors that depend on meteorological and environmental driving variables. Recent experiments have identified large uncertainties in the MEGAN temperature response parameterization, while the emission rates under standard conditions are poorly constrained in some regions due to a lack of representative measurements and uncertainties in landcover. In this study, we use Bayesian model-data fusion to optimize the MEGAN temperature response and standard emission rates using satellite- and ground-based observational constraints. Optimization of the standard emission rate with satellite constraints reduced model biases but was highly sensitive to model input errors and drought stress and was found to be inconsistent with ground-based constraints at an Amazonian field site, reflecting large uncertainties in the satellite-based emissions. Optimization of the temperature response with ground-based constraints increased the temperature sensitivity of the model by a factor of five at an Amazonian field site but had no impact at a UK field site, demonstrating significant ecosystem-dependent variability of the isoprene emission temperature sensitivity. Ground-based measurements of isoprene across a wide range of ecosystems will be key for obtaining an accurate representation of isoprene emission temperature sensitivity in global biogeochemical models.

Optimizing the isoprene emission model MEGAN with satellite and ground-based observational constraints

Christian A. DiMaria[1], Dylan B. A. Jones[1], Helen Worden[2], A. Anthony Bloom[3], Kevin Bowman[3], Trissevgeni Stavrakou[4], Kazuyuki Miyazaki[3], John Worden[3], Alex Guenther[5], Chinmoy Sarkar[6†], Roger Seco[7], Jeong-Hoo Park[8], Julio Tota[9], Eliane Gomes Alvez[10], and Valerio Ferracci[11]

[1]Department of Physics, University of Toronto, Toronto, Ontario, Canada, [2]Atmospheric Chemistry Observations & Modeling Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [4]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium, [5]Department of Earth System Science, University of California, Irvine, CA, USA, [6]Air Quality Research Center, University of California, Davis, CA, USA, [7]Institute of Environmental Assessment and Water Research (IDAEA-CSIC), Barcelona, Catalonia, Spain, [8]Air Quality Reserach Division, National Institute of Environmetal Research, 22689, Republic of Korea, [9]Julio Tota, Instituto de engenharia e Geociências, Universidade Federal do Oeste do Pará, UFOPA, Santarém, Pará, Brazil, [10]Department of Biogeochemical Processes, Max Planck Institute for Biogeochemistry, Jena, Germany, [11]School of Water, Energy and Environment, Cranfield University, Cranfield, UK

Corresponding author: Christian DiMaria (christian.dimaria@mail.utoronto.ca)

† Now at Thermo Fisher Scientific, San Francisco, CA, USA

Key Points:

- Satellite and ground based observations were used to optimize an isoprene emission model in a Bayesian model-data fusion framework
- Optimization with satellite observations was highly uncertain due to observation biases and a high sensitivity to model input errors
- Ground-based observations showed that Amazonian isoprene emissions were 5x more sensitive to temperature than UK isoprene emissions

Abstract

Isoprene is a hydrocarbon emitted in large quantities by terrestrial vegetation. It is a precursor to several air quality and climate pollutants including ozone. Emission rates vary with plant species and environmental conditions. This variability can be modelled using the Model of Emissions of Gases and Aerosols from Nature (MEGAN). MEGAN parameterizes isoprene emission rates as a vegetation-specific standard rate which is modulated by scaling factors that depend on meteorological and environmental driving variables. Recent experiments have identified large uncertainties in the MEGAN temperature response parameterization, while the emission rates under standard conditions are poorly constrained in some regions due to a lack of representative measurements and uncertainties in landcover. In this study, we use Bayesian model-data fusion to optimize the MEGAN temperature response and standard emission rates using satellite- and ground-based observational constraints. Optimization of the standard emission rate with satellite constraints reduced model biases but was highly sensitive to model input errors and drought stress and was found to be inconsistent with ground-based constraints at an Amazonian field site, reflecting large uncertainties in the satellite-based emissions. Optimization of the temperature response with ground-based constraints increased the temperature sensitivity of the model by a factor of five at an Amazonian field site but had no impact at a UK field site, demonstrating significant ecosystem-dependent variability of the isoprene emission temperature sensitivity. Ground-based measurements of isoprene across a wide range of ecosystems will be key for obtaining an accurate representation of isoprene emission temperature sensitivity in global biogeochemical models.

Plain Language Summary

Isoprene is a reactive hydrocarbon emitted by plants into the atmosphere, where it impacts air quality and regional climate. The emission rate depends on plant species and environmental conditions. Isoprene emission rates can be estimated with computer models, but these models are uncertain. For example, the emission capacity of many ecosystems ("standard emission rate") is not well known due to a lack of measurements, and recent experiments have revealed variability in the temperature dependance of isoprene emissions that is not captured by current models. Our goal in this study was to improve a widely-used computer model of isoprene emissions by updating the standard emission rate and temperature response with observations. We used

satellite observations to update the standard emission rate, but these results were uncertain due to observation uncertainties and sensitivity to model input errors. Using ground-based observations, we found that isoprene emissions were five times more sensitive to temperature at an Amazonian field site than they were at a UK field site, suggesting the need for an ecosystem-dependent temperature sensitivity parameterization in the model. Such a parameterization has the potential to improve models of isoprene emissions during extreme heat events and in a warming climate.

1. Introduction

Isoprene is a reactive volatile organic compound (VOC) emitted by terrestrial vegetation, possibly contributing to abiotic stress tolerance (Monson et al., 2021). Annual isoprene emissions exceed all other non-methane VOCs, with an estimated global emission rate of 360 - 800 Tg yr^{-1} (Guenther et al., 2012). Emissions vary widely among plant species and are concentrated in tropical terrestrial ecosystems (Guenther et al., 1995, 2006), but can also be significant in temperate regions during the growing season (Wiedinmyer et al., 2005; Potosnak et al., 2014; Seco et al., 2015). Emission rates vary in response to temperature and sunlight (Guenther et al., 1993), and are modulated on seasonal and interannual timescales by landcover and canopy environment changes (Alves et al., 2016; Chen et al., 2018) as well as environmental stressors such as drought (X. Jiang et al., 2018; Potosnak et al., 2014; Seco et al., 2015) and CO_2 inhibition (Heald et al., 2009).

Isoprene has a short atmospheric lifetime on the order of hours due to rapid oxidation by the hydroxyl (OH) radical (Sprengnether et al., 2002). This oxidation contributes to the formation of air quality and climate pollutants, including ozone (Trainer et al., 1987), secondary organic aerosol (Claeys at al., 2004), and carbon monoxide (CO) (Z. Jiang et al., 2017). The biogenic CO source from isoprene oxidation exceeds anthropogenic and biomass burning emissions in many regions (Hudman et al., 2008; Worden et al., 2019), and the relative contribution of biogenic CO to the total atmospheric CO burden may be increasing due to declining anthropogenic emissions (Buchholz et al., 2021; Worden et al., 2013). Isoprene is a major sink for OH, which influences the lifetime of methane and other trace gases (Karl et al., 2007, 2013). Quantifying these atmospheric chemistry impacts requires accurate isoprene emission estimates.

The Model of Emissions of Gases and Aerosols from Nature (MEGAN) is widely used to estimate isoprene emissions (Guenther et al., 2006, 2012). In MEGAN, isoprene emissions are the

product of a vegetation-specific standard emission rate and dimensionless scaling functions which depend on temperature, sunlight, leaf area index, leaf age, and soil moisture. MEGAN isoprene emissions are highly sensitive to meteorology and landcover (Arneth et al., 2011; Misztal et al., 2016) and are thus dependent on the accuracy and resolution of these model inputs (Ashworth et al., 2010; Pugh et al., 2013). Model performance is generally good when accurate driving variables are used (Situ et al., 2014; see also Filella et al., 2018; Sarkar et al., 2020), but significant sources of uncertainty remain. These include the empirical parameterization of the standard emission rates and the temperature response algorithm (Emmerson et al., 2020; Guenther et al., 2006).

The standard emission rates are based on leaf- or canopy-scale isoprene flux measurements which can be extrapolated globally over regions with similar landcover (Guenther et al., 1995). This methodology introduces biases due to the large variability of isoprene emission rates among plant species (Guenther et al., 1993) and within ecosystems (Batista et al., 2019; Li et al., 2021). Emission rate biases are exacerbated by landcover uncertainties (Opacka et al., 2021), particularly in tropical regions (Fang et al., 2013; Gu et al., 2017; Mougin et al., 2019). An additional source of uncertainty is the relative lack of isoprene flux measurements to constrain emission rates in the tropics (Guenther et al., 2006; Marais et al., 2014). Constraining the standard emission rates with satellite observations has been shown to reduce emission biases in MEGAN (Bauwens et al., 2016; Marais et al., 2014), allowing the model to be improved in remote regions.

MEGAN's temperature response algorithm, which simulates the exponential increase in emissions with temperature up to an optimum value (Guenther et al., 1993, 1995), is another source of uncertainty. Recent experiments with Arctic vegetation (Angot et al., 2020; Kramshøj et al., 2016; Seco et al., 2020, 2022) and Australian *eucalypt* trees (Emmerson et al., 2020) have found that MEGAN significantly underestimates isoprene emissions from these species at high temperatures. Updating the parameterization of the temperature response with species-specific measurements in Australia has been shown to improve MEGAN isoprene emissions estimates, with significant consequences for predictions of future ozone pollution in a warming climate (Emmerson et al., 2020). Because temperature and sunlight are the primary drivers of short-term isoprene emission variability (Guenther et al., 1993), uncertainties in the MEGAN temperature response have significant implications for isoprene emission modeling.

In this study we use Bayesian model-data fusion, a form of data assimilation which combines information from models and observations in a statistically rigorous way, to optimize

the parameterization of the MEGAN standard emission rate and temperature response. Using top-down isoprene emissions based on satellite formaldehyde retrievals from the Ozone Monitoring Instrument (OMI; Bauwens et al., 2016) and the TROPOspheric Monitoring Instrument (TROPOMI), we constrain the standard emission rate in several regions to reduce model biases. Eddy covariance isoprene measurements at the BR-Sa1 AmeriFlux site in the Tapajós National Forest in the Brazilian Amazon and isoprene atmospheric mixing ratio measurements at Wytham Woods near Oxford, United Kingdom, are used to constrain the temperature response parameterization. The sensitivity of our optimization to model input errors and drought stress, discrepancies between satellite- and ground-based constraints, and the variability of the temperature response between the two field sites are discussed.

2. Methods and Data

In Section 2.1 we describe the Model of Emissions of Gases and Aerosols from Nature (MEGAN), with a particular focus on the model sensitivity to temperature. Section 2.2 introduces the ground-based isoprene measurements (2.2.1) and satellite-based top-down isoprene flux estimates (2.2.2). Finally, Section 2.3 describes the Bayesian model-data fusion and validates the method with a simulated observation experiment.

2.1 MEGAN and the temperature activity factor $\boldsymbol{\gamma_T}$

MEGAN is an empirical model in which isoprene emissions are the product of a vegetation-specific standard rate and activity factors (i.e., dimensionless scaling functions) (Guenther et al., 2006) which depend on meteorology and leaf phenology (Guenther et al., 1991, 1993). This is shown in the following equation:

$$E_{ISOP} = E_0 \times C_{CE} \times \gamma_T \times \gamma_{PAR} \times \gamma_{LAI} \times \gamma_{AGE} \times \gamma_{CO2} \times \gamma_{SM}, \quad (1)$$

where E_{ISOP} is the isoprene emission rate, E_0 is the standard emission rate, and $\boldsymbol{\gamma}_{\mathrm{T}}$, $\boldsymbol{\gamma}_{\mathrm{PAR}}$, $\boldsymbol{\gamma}_{\mathrm{LAI}}$, $\boldsymbol{\gamma}_{\mathrm{AGE}}$, $\boldsymbol{\gamma}_{\mathrm{CO2}}$, and $\boldsymbol{\gamma}_{\mathrm{SM}}$ are activity factors that represent the emission sensitivity to temperature (T), photosynthetically active radiation (PAR), leaf area index (LAI), leaf age, CO_2 partial pressure, and soil moisture (SM) stress respectively (Guenther et al., 2006, 2012). Diurnal variability of modelled emissions is controlled by temperature and sunlight, whereas longer term variability is influenced by changes in leaf phenology, landcover, drought stress, and CO_2 inhibition (Alves et al., 2016; Chen et al., 2018; Guenther et al., 1993; Heald et al., 2009; Opacka et al., 2021; X. Jiang

et al., 2018). C_{CE} is a normalization constant which ensures that $E_{ISOP} = E_0$ under standard conditions (see Supplementary Text S1).

We use the Parameterized Canopy Environment Emission Activity (PCEEA) algorithm for the activity factors $\boldsymbol{\gamma}_{\mathrm{T}}$, $\boldsymbol{\gamma}_{\mathrm{PAR}}$, and $\boldsymbol{\gamma}_{\mathrm{LAI}}$ and the empirical parameter values contained therein (Guenther et al., 2006). In particular, the temperature response function $\boldsymbol{\gamma}_{\mathrm{T}}$ is given by

$$\gamma_T = E_{Opt}\left[\frac{C_{T2}\times\exp(C_{T1}\times x)}{\left(C_{T2}-C_{T1}\times(1-\exp(C_{T2}\times x))\right)}\right], \tag{2}$$

where C_{T1} (80 kJ mol^{-1}) and C_{T2} (200 kJ mol^{-1}) are fitting parameters, x is a temperature dependent variable given by

$$x = \left[\frac{\frac{1}{T_{Opt}}-\frac{1}{T}}{R}\right], \tag{3}$$

where T is the air temperature (K) (assumed to be equal to the leaf temperature in the PCEEA algorithm) and R is the ideal gas constant (0.00831 kJ K^{-1} mol^{-1}). Equation (2) is based on a photosynthetic electron transport model in which activity increases with temperature up to an optimum value (Guenther et al., 1991). E_{Opt} is an empirical parameter given by the equation

$$E_{Opt} = C_{EO}\times\exp\left(K_2\left(T_{daily}-297\right)\right), \tag{4}$$

where C_{E0} (1.75) and K_2 (0.08) are empirical coefficients and T_{daily} is the average air temperature of the past 24 hours (Guenther et al., 2006). T_{Opt} is an empirical parameter defined as follows:

$$T_{Opt} = T_{Max} + \left(K_1\left(T_{daily}-297\right)\right) \tag{5}$$

where T_{Max} (313 K) is the peak emission temperature under standard conditions and K_1 (0.6) is an empirical coefficient. Equations (4) and (5) describe how the peak isoprene emission rate, as well as the temperature at which it occurs, changes as a function of T_{daily} (see Supplementary Text S2). This hysteresis, or long-term temperature response, is based on a small number of experiments and is highly uncertain (Geron et al., 2000; Hanson & Sharkey, 2001; Monson et al., 1994; Pétron et al., 2001).

We quantify the sensitivity of $\boldsymbol{\gamma}_{\mathrm{T}}$ to temperature using the Q_{10} metric, which is defined as the multiplicative factor by which $\boldsymbol{\gamma}_{\mathrm{T}}$ changes when going from 303 – 313 K (Seco et al., 2020). All Q_{10} values presented in this study are calculated using Equations (2) – (5) with an assumed

179 T_{daily} of 297 K (standard conditions) to ensure a fair comparison between different field sites with
180 different T_{daily}.

181 Unless otherwise stated, we drive MEGAN2.1 (Guenther et al., 2012) with hourly
182 meteorology fields from the Modern-Era Retrospective analysis for Research and Applications,
183 Version 2 (MERRA-2; Gelaro et al; 2017) at 0.5° × 0.625° spatial resolution and 8-day average
184 LAI at 2° × 2.5° resolution from the Moderate Resolution Imaging Spectroradiometer (MODIS)
185 aboard NASA's Terra and Aqua satellites (Yuan et al., 2011). Local temperature measurements
186 are used to drive the model at BR-Sa1 (Sarkar et al., 2020) and Wytham Woods (Ferracci et al.,
187 2020) instead of MERRA-2 temperature data. We used hourly Goddard Earth Observing System
188 - Forward Processing (GEOS-FP; provided by the Global Modeling and Assimilation Office
189 (GMAO) at NASA Goddard Space Flight Center) PAR and windspeed data to filter the Wytham
190 Woods observations due to their higher native resolution (0.25° × 0.3125°) relative to MERRA-2.
191 We used gridded $\boldsymbol{\gamma}_{SM}$ values at 0.5° × 0.5° from 2005 – 2015 as obtained with Community Land
192 Model (CLM) 4.5 and the MEGAN3 soil moisture algorithm (X. Jiang et al, 2018). We fix $\boldsymbol{\gamma}_{CO2}$ =
193 1 due to the insignificance of CO_2 inhibition on short timescales (Heald et al., 2009). A priori E_0
194 at 0.25° × 0.3125° spatial resolution are based on the CLM plant functional type distribution
195 (Lawrence et al., 2011).
196

197 ## 2.2 Data

198 ### 2.2.1 Eddy covariance and atmospheric mixing ratio measurements

199 Isoprene emissions can be estimated from tower- or aircraft- based eddy covariance
200 measurements (Guenther & Hills., 1998; Karl et al., 2009). This approach is useful for constraining
201 the diurnal variability of isoprene emissions, making it well-suited for characterizing the
202 temperature response (Misztal et al., 2014; Seco et al., 2015, 2022; Yu et al., 2017). In this study,
203 we use tower-based eddy covariance isoprene flux measurements from the AmeriFlux site BR-Sa1
204 in Brazil's Tapajós National Forest (2.86°S, 54.96°W) (Sarkar et al, 2020), as well as isoprene
205 mixing ratio measurements from the 2018 Wytham Isoprene iDirac Oak Tree Measurements
206 (WIsDOM) campaign in Wytham Woods near Oxford, UK (51.46°N, 1.20°W) (Ferracci et al,
207 2020). The locations of these field sites are shown in Figure 1. The field sites were selected to
208 represent contrasting ecosystems based on the availability of concurrent isoprene and temperature
209 measurements. Tapajós National Forest is a protected Amazonian old-growth closed-canopy

evergreen tropical forest (Sarkar et al., 2020), while the Wytham Woods site is a mixed temperate woodland where *Quercus robur* (European oak) is the dominant isoprene emitter (Butt et al., 2009).

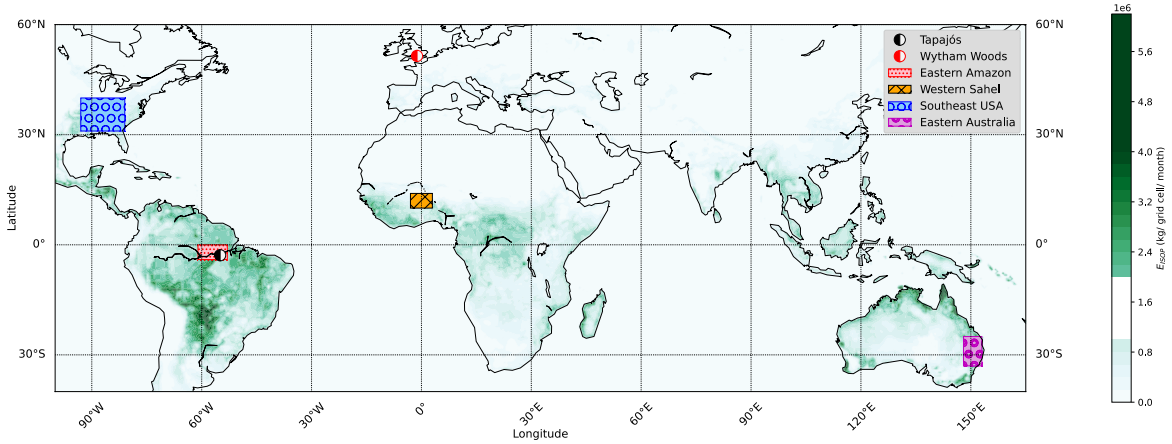

Figure 1: Locations of the 4 study regions used for the optimization of E_0 with top-down constraints (rectangles) and the two field sites used for the optimization of $\boldsymbol{\gamma}_T$ with eddy covariance and mixing ratio constraints (circles). The colorbar represents the mean isoprene emission rate (kg / 0.5°×0.5° grid cell / month) from the OMI-based GlobEmission top-down isoprene emission dataset from 2005 - 2014.

The BR-Sa1 dataset consists of hourly isoprene emission measurements and surface air temperature measurements from 1–16 June 2014 (Sarkar et al., 2020). The isoprene fluxes were calculated from proton transfer reaction time-of-flight mass spectrometry (PTR-TOF-MS) isoprene mixing ratio measurements using eddy covariance techniques. The spatiotemporal overlap of the BR-Sa1 data with the OMI-based GlobEmission data allows for a direct comparison of top-down and eddy covariance constraints. Uncertainties were estimated to be 15%, based on measurement errors and the variability of the observations (Sarkar et al., 2020). MEGAN was driven with locally measured temperature data at this site. The soil moisture response $\boldsymbol{\gamma}_{SM}$ was equal to 1 at BR-Sa1 throughout the study period (X. Jiang et al., 2018). Further details about the BR-Sa1 measurements, as well as a more detailed site description, are available in Sarkar et al (2020).

The WIsDOM dataset used in this study consists of isoprene mixing ratio and temperature measurements taken at a height of 15.55m above ground level in a UK mixed forest canopy from 25 May to 6 November 2018 (Ferracci et al., 2020). Isoprene mixing ratios were measured using

the iDirac (Bolas et al., 2020) portable gas chromatograph with photo-ionization detection (GC-PID) with a measurement precision of 10% and a time resolution of 20 minutes. Temperature was measured with an EasyLog probe (EL-USB-2, Lascard Ltd.) placed next to the iDirac inlet; this temperature dataset was used to drive MEGAN at this site. A long and uninterrupted heatwave occurred at Wytham Woods from 22 June to 8 August 2018 and the associated drought had a significant impact on local isoprene emission rates (Ferracci et al., 2020; Otu-Larbi et al., 2020). To avoid the impact of drought stress we use only the data from 25 May to 21 June, hereafter referred to as the preheatwave period. Local PAR measurements were not available at Wytham Woods during the preheatwave period, so we used hourly GEOS-FP PAR at 0.25° × 0.325° resolution to drive MEGAN. Following Ferracci et al (2020), we used a linear mapping to account for any systematic offset between the GEOS-FP PAR and the locally measured PAR which was available at Wytham Woods in September 2018 (see Supplementary Text S4). Additionally, hourly GEOS-FP 10-m wind speed data was used for data filtering (see Section 3.3.2). Wytham Woods data are available for download at https://archive.ceda.ac.uk. Further details about the WIsDOM campaign are available in Ferracci et al (2020).

2.2.2 Top-down emissions

Isoprene emissions can be estimated from satellite retrievals of its high-yield oxidation product formaldehyde (CH_2O) (Palmer et al., 2003). Regional and global estimates have been obtained from several sensors including the Global Ozone Monitoring Experiment (GOME) (Palmer et al., 2003), the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) (Barkley et al., 2013; Stavrakou et al., 2009a), GOME-2 (Stavrakou et al., 2015), the Ozone Monitoring Instrument (OMI) (Bauwens et al., 2016; Kaiser et al., 2018; Marais et al., 2012), and the TROPOspheric Monitoring Instrument (TROPOMI). Here we use top-down isoprene estimates from (i) the GlobEmission global dataset (2005 - 2014) constrained by OMI CH_2O columns (Bauwens et al., 2016), and (ii) the SOLFEO regional dataset over South America constrained by TROPOMI CH_2O columns in 2018.

The 10-year (2005–2014) GlobEmission inventory (Bauwens et al., 2016) was created by assimilating OMI CH_2O column retrievals into the IMAGESv2 chemistry-transport model (Müller & Stavrakou, 2005; Stavrakou et al., 2009b, c) using an adjoint-based variational data assimilation scheme. The inversion is performed at a spatial resolution of 2°×2.5° on a monthly basis. The top-

down emissions are further downscaled to a daily inventory resolution at 0.5°×0.5°, based on the spatiotemporal variability of the a priori MEGAN inventory used in the inversions. We use the GlobEmission data in the Eastern Amazon basin (0° – 4°S, 53°W – 61°W), the Western Sahel (10°N – 14°N, 3°W – 3°E), the Southeastern United States (31°N – 40°N, 81°W – 93°W), and Eastern Australia (25°S – 33°S, 148°E – 153°E). Data for the year 2005 was omitted from our analysis due to missing MEGAN driving variables. Figure 1 shows the mean GlobEmission isoprene emissions, with the locations of the study regions indicated. Discrepancies between MEGAN and top-down flux estimates have been previously documented in each of these regions (Bauwens et al., 2016; Marais et al., 2014; Worden et al., 2019), with the Amazon and Southeast USA accounting for a significant portion of the total global annual isoprene emissions (Guenther et al., 2006). The SOLFEO dataset (https://emissions.aeronomie.be/index.php/tropomi-based/isoprene-sa) provided an independent set of top-down constraints for the E_0 optimization in the Eastern Amazon region. This dataset provides monthly average isoprene emission rate estimates for 2018 at 0.5° × 0.5° resolution in South America, based on an assimilation of TROPOMI CH_2O observations into the regional chemistry-transport model MAGRITTEv1.1 (Müller et al., 2019; Stavrakou et al., 2016). Both top-down emission datasets are publicly available from the Emission portal at https://emissions.aeronomie.be/.

A validation study using a network of ground-based remotely-sensed CH_2O columns revealed a substantial bias in the TROPOMI CH_2O columns (Vigouroux et al., 2020, De Smedt et al., 2021). The estimated bias is found to be low for high columns (-30% for values higher than 8×10^{15} molec. cm^{-2}) and high for low columns (+26% for values lower than 2.5×10^{15} molec. cm^{-2}). Based on those comparisons, bias-corrected TROPOMI columns are given by the linear regression relationship: $1.56\times C - 1.72\times10^{15}$ molec. cm^{-2}, where C is the TROPOMI CH_2O column (in molec. cm^{-2}). The adjustment increases the columns by 20-50% for TROPOMI columns within (5-40) $\times10^{15}$ molec. cm^{-2}. Those higher columns would entail substantially higher top-down isoprene fluxes than those derived based on the standard TROPOMI product. Since the OMI and TROPOMI formaldehyde products were retrieved using similar algorithms (De Smedt et al., 2018), the top-down estimates of the GlobEmission dataset are also likely similarly underestimated.

The optimization of E_0 was found to be largely insensitive to the assumed uncertainty of the top-down emissions (see Supplementary Text S3). A relatively large uncertainty of 30% was used to avoid overfitting the model to the top-down emissions, but this is likely an underestimate

of the true uncertainty in some regions due to the biases in the satellite CH_2O retrievals (Vigouroux et al., 2020, De Smedt et al., 2021).

2.3 Parameter estimation using Bayesian model-data fusion

We use Bayesian model-data fusion, a form of data assimilation used to combine information from models and observations in a statistically rigorous way, to constrain the MEGAN parameters with observations. Given a parameter vector $\boldsymbol{x}$ and observations $\boldsymbol{O}$, we can derive the posterior probability density function of $\boldsymbol{x}$ as

$$P(\boldsymbol{x}|\boldsymbol{O}) \propto P(\boldsymbol{x})P(\boldsymbol{O}|\boldsymbol{x}) \quad (6)$$

where $P(\boldsymbol{x})$ is the a priori probability distribution of $\boldsymbol{x}$ and $P(\boldsymbol{O}|\boldsymbol{x})$ is the observation probability given $\boldsymbol{x}$, also called the model likelihood function. In this study we assume a non-informative uniform $P(\boldsymbol{x})$ for all parameters such that $P(x) = 1$ for all $\boldsymbol{x} \in [\boldsymbol{x}_{min}, \boldsymbol{x}_{max}]$ and $P(x) = 0$ outside of this range. The limits $\boldsymbol{x}_{min}$ and $\boldsymbol{x}_{max}$ were respectively set to 1/5 and 5 times the a priori parameter value for all parameters except for T_{Max}, which was constrained to $T_{Max} \in$ [20°C, 60°C] to avoid unphysical parameterizations. For all $\boldsymbol{x} \in [\boldsymbol{x}_{min}, \boldsymbol{x}_{max}]$ the posterior probability $P(\boldsymbol{x}|\boldsymbol{O})$ is then proportional to the observation probability $P(\boldsymbol{O}|\boldsymbol{x})$, which we define for N observations as

$$P(\boldsymbol{O}|\boldsymbol{x}) = \exp\left(-0.5\sum_{n=1}^{N}\frac{(M_n - O_n)^2}{\sigma_n^2}\right) \quad (7)$$

where O_n is the nth observation (top-down or eddy covariance isoprene measurements), M_n is the corresponding model state (in our case, MEGAN isoprene emissions), and σ_n^2 is the observation error variance. This definition of $P(\boldsymbol{O}|\boldsymbol{x})$ assumes Gaussian error statistics and no covariance between observation errors.

We sample the probability distribution $P(\boldsymbol{x}|\boldsymbol{O})$ using an adaptive Metropolis-Hastings Markov Chain Monte Carlo (MHMCMC) algorithm (Haario et al., 2001). The MHMCMC algorithm has been previously applied to parameter estimation problems in the context of ecosystem modeling (Bloom et al., 2015; Bloom & Williams, 2015; Xu et al., 2006; Ziehn et al., 2012), and consists of the following 4 basic steps:

1. Choose an initial parameter state vector: $\boldsymbol{x_i}$
2. Perturb the parameters: $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \Delta\boldsymbol{x}$
3. Run model with both sets of parameters: $\boldsymbol{M}(\boldsymbol{x}_i)$ and $\boldsymbol{M}(\boldsymbol{x}_{i+1})$

4. Accept new parameters if $\frac{P(\boldsymbol{x}_{i+1}|\boldsymbol{O})}{P(\boldsymbol{x}_i|\boldsymbol{O})} > U \in [0,1]$; else $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i$

where $\boldsymbol{x}_i$ is the i^{th} iteration of the parameter state vector, $\Delta\boldsymbol{x}$ is the parameter perturbation step-size, U is a uniform distribution (Ziehn et al., 2012; see also Bloom & Williams, 2015), and $P(\boldsymbol{x}|\boldsymbol{O})$ is the observation probability, given by Equation (7) for all $\boldsymbol{x} \in [\boldsymbol{x}_{min}, \boldsymbol{x}_{max}]$ and equal to 0 otherwise. Our implementation of the MHMCMC algorithm follows Bloom et al. (2020) and uses a Matlab code developed by Yang et al (2021, 2022).

The use of an adaptive step size $\Delta\boldsymbol{x}$ reduces the number of iterations required to explore the parameter space, with smaller step sizes being used in regions of parameter space where the probability function in (7) is steeper. Step 4 of the MHMCMC algorithm ensures that parameter values which maximize the probability in (7) are more likely to be accepted. Because only the ratio of probabilities is used, $P(\boldsymbol{x}|\boldsymbol{O})$ does not need to be normalized. In our experiments, Steps 1 - 4 were repeated for 4×10^4 iterations. The first half of the samples were discarded to allow for the statistical properties of the parameter distribution to stabilize, and the remaining half were subsampled by a factor of 20 to reduce correlations between samples (Ziehn et al., 2012), giving a final distribution of 1000 parameter samples for each experiment.

We validated our MHMCMC approach using a series of simulated observation experiments. The purpose of these experiments was to ensure that the MHMCMC scheme was capable of estimating MEGAN parameters when provided with suitable observational constraints. In the initial experiment, the posterior $\boldsymbol{\gamma}_T$ parameterization of Emmerson et al (2020) (based on *eucalypt* measurements) was used to calculate a "true" temperature response with Equation (2). We sampled this "true" temperature response from 290 – 330 K at 2 K intervals to produce pseudo-observations, which were perturbed with Gaussian noise ($\boldsymbol{\sigma}$ = 0.1 – 2.0, approximately 3% – 60% of the mean "true" $\boldsymbol{\gamma}_T$) to simulate measurement error. The PCEEA parameterization of $\boldsymbol{\gamma}_T$ was used as the a priori, as in all other experiments in this study. Figure 2 (a) shows the results of a simulated observation temperature optimization experiment, with $\boldsymbol{\sigma}$ = 0.5. The optimized posterior $\boldsymbol{\gamma}_T$ is in good agreement with the truth and is much more sensitive to temperature (Q_{10} = 3.01) than the a priori (Q_{10} = 1.73). The MEGAN $\boldsymbol{\gamma}_T$ parameters T_{Max}, C_{EO}, C_{T1}, and C_{T2} could be constrained with pseudo-observations provided that the observation errors were sufficiently small. Figure 2 (b) – (e) show that the optimized parameters are consistent with those derived by Emmerson et al (2020). These results demonstrate that MHMCMC is a suitable method for constraining MEGAN parameters with observations.

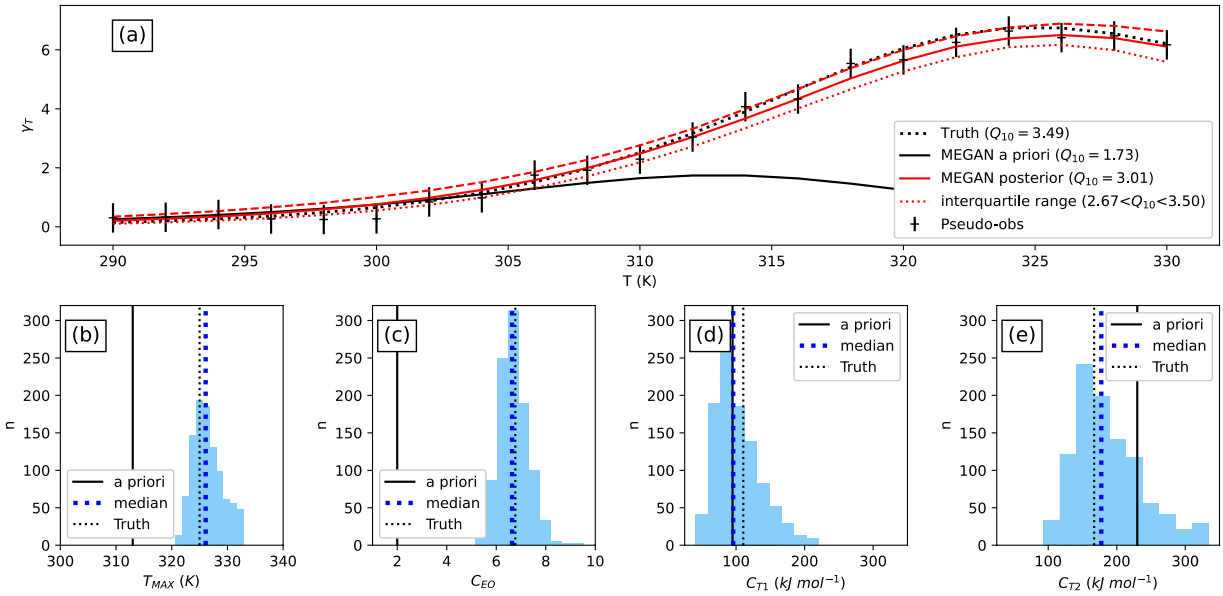

Figure 2: (a) A priori MEGAN $\boldsymbol{\gamma}_T$ (black), median posterior $\boldsymbol{\gamma}_T$ (solid red) and interquartile range (dashed red) as a function of temperature compared with pseudo-observations. The dotted black line represents the "true" $\boldsymbol{\gamma}_T$ as calculated using the posterior parameterization of Emmerson et al (2020). All $\boldsymbol{\gamma}_T$ curves in (a) were calculated using Equation (2). The posterior parameter distributions for T_{Max}, C_{EO}, C_{T1}, and C_{T2} (from Equations (2), (4), and (5)) are shown in (b) – (e) respectively, with the median values (dotted blue line) agreeing closely with the "true" values (thin dotted black line). The a priori parameter values are indicated by the solid black lines. The a priori probability distribution for each parameter (not shown) was uniform and spanned from 1/5 – 5 times the a priori values, except for T_{Max} which was constrained between 293 – 333K.

Additional simulated observation experiments were used to assess the impact of observation errors and temporal resolution on the optimization using time series pseudo-observations. We found that while the standard emission rate E_0 could be easily constrained with relatively imprecise (< 50% error) or infrequent (e.g., monthly) observations (see Supplementary Text S3), the $\boldsymbol{\gamma}_T$ parameters could only be properly constrained with more precise and higher frequency observations (see Supplementary Text S5). A sensitivity analysis revealed that MEGAN is highly sensitive to the E_0 scaling factor but less sensitive to the $\boldsymbol{\gamma}_T$ parameters at typical ambient temperatures (see Supplementary Text S6). We therefore limit our $\boldsymbol{\gamma}_T$ parameter optimization to the BR-Sa1 and Wytham Woods sites, where sufficiently precise and frequent isoprene measurements are available to constrain the temperature response.

3. Results

Section 3.1 presents the optimization of the standard emission rate E_0 using satellite-based top-down constraints in our four study regions (3.1.1) and using eddy covariance measurements at the Amazonian BR-Sa1 site (3.1.2). Section 3.2 presents the optimization of the $\boldsymbol{\gamma}_{\mathrm{T}}$ parameters using eddy covariance observations at BR-Sa1 (3.2.1) and using isoprene mixing ratio measurements at the UK Wytham Woods site (3.2.2).

3.1. Optimization of standard emission rate E_0

3.1.1 Top-down constraints

Monthly OMI-based GlobEmission data were used to constrain the standard emission rate E_0 in four study regions (Figure 1) for the period of 2006 – 2014. The study regions were selected to represent a diversity of ecosystems with substantial isoprene emissions. The MEGAN a priori emissions were calculated as described in Section 2.1. We performed the optimization both with and without the MEGAN3 soil moisture response $\boldsymbol{\gamma}_{\mathrm{SM}}$ enabled to assess the impact of this uncertain model component. All MEGAN $\boldsymbol{\gamma}$ factor parameters were fixed to their a priori (PCEEA) values, as defined in Section 2.1.

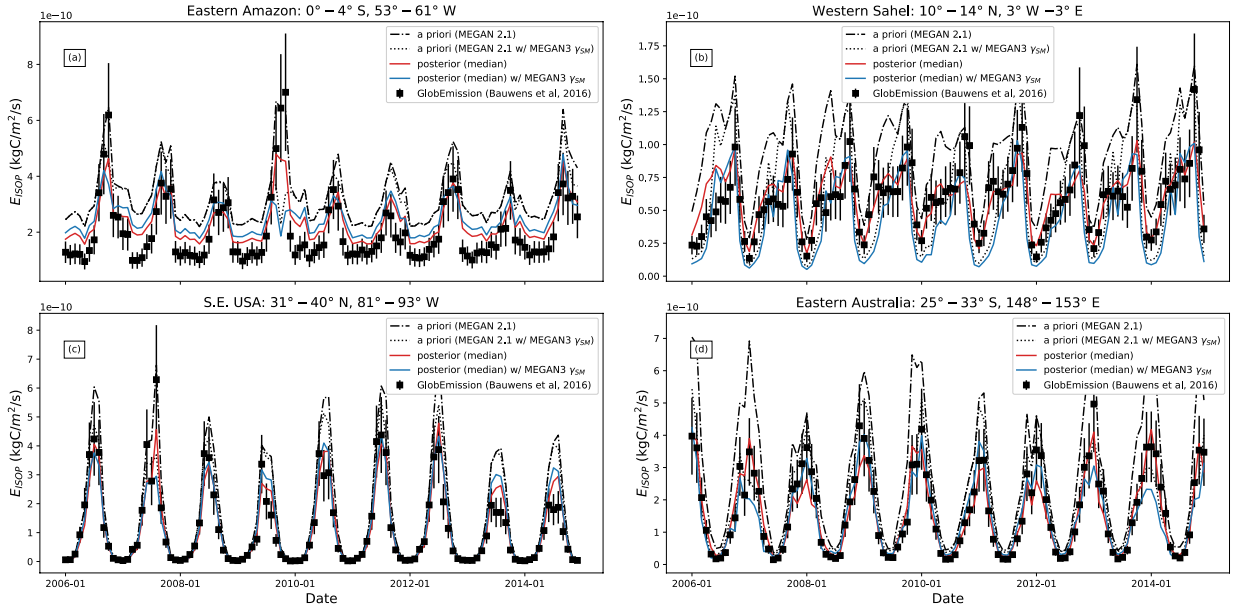

Figure 3: Comparison of a priori and a posterioi fluxes with top-down GlobEmission fluxes for (a) the eastern Amazon, (b) the western Sahel, (c) the southeastern US, and (d) eastern Australia. Shown are the a priori MEGAN2.1 fluxes (dashed-dotted line), the a priori fluxes with the inclusion of the MEGAN3 drought stress response $\boldsymbol{\gamma}_{\mathrm{SM}}$ (dotted line),

the median a posteriori MEGAN2.1 fluxes (solid red line), and the median a posteriori MEGAN2.1 fluxes with the inclusion of the MEGAN3 drought stress response $\boldsymbol{\gamma}_{SM}$ (solid blue line).

Optimization of the standard emission rate E_0 reduced MEGAN isoprene emissions in all four study regions, leading to better agreement between MEGAN and the OMI-based GlobEmission flux estimates (Figure 3). The optimized E_0 values were well-constrained and significantly reduced relative to their a priori values (Figure 4). Constraining E_0 with the TROPOMI-based SOLFEO fluxes produced a consistent result in the Eastern Amazon (Figure 4), indicating that the biases observed in this region are not unique to the GlobEmission dataset. Our results are consistent with previous studies which have used top-down constraints to reduce biases in modelled isoprene emissions (Kaiser et al., 2018) and in E_0 (Marais et al., 2014).

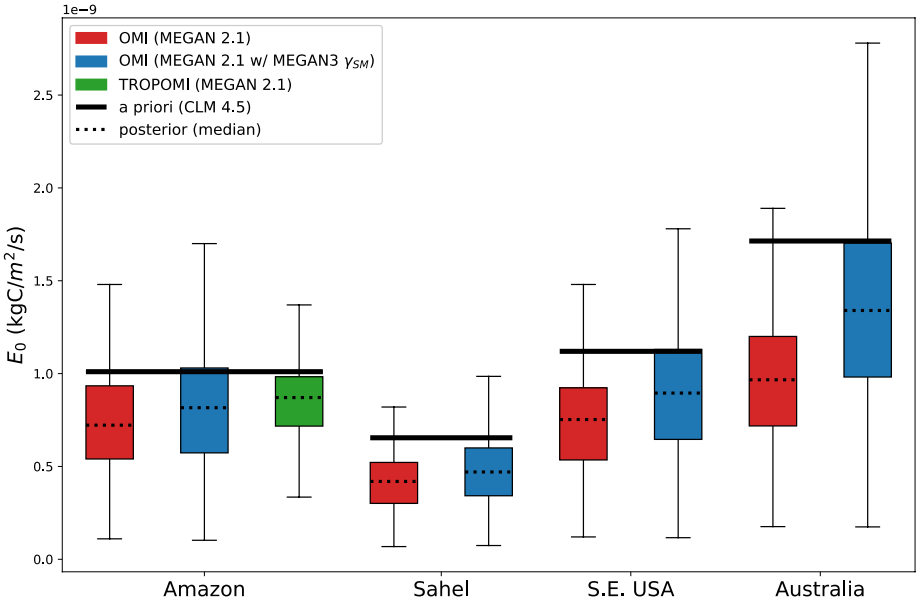

Figure 4: Comparison of the a posteriori E_0 inferred from the top-down constraints in all four study regions. The red boxes indicate the E_0 inferred from the OMI-based constraints and MEGAN2.1,whereas the blue boxes denote the E_0 estimated from the OMI-based constraints and MEGAN2.1 with the inclusion of the MEGAN3 drought stress response $\boldsymbol{\gamma}_{SM}$. The green bar for the Amazon shows the a posteriori E_0 obtained from the TROPOMI-based SOLFEO fluxes.

Enabling $\boldsymbol{\gamma}_{SM}$ reduces the a priori emissions in each region (Figure 3). The optimized E_0 values are consequently larger than in the no-$\boldsymbol{\gamma}_{SM}$ case (Figure 4) because some of the bias between MEGAN and GlobEmission has already been accounted for by $\boldsymbol{\gamma}_{SM}$. This change was most significant in Eastern Australia, where the difference between the two optimization results is over

40% due to the large impact of γ_{SM} on the a priori emissions. The significance of drought stress for Eastern Australian isoprene emissions has been previously reported (Emmerson et al., 2019) and our results are consistent with this. Although the impact on the a priori emissions and posterior E_0 was large, inclusion of γ_{SM} had little impact on the posterior emissions in all four regions. This is an example of equifinality, in which different combinations of model inputs and parameters (in this case E_0 and γ_{SM}) produce the same output. The sensitivity of the optimization results to γ_{SM} is a source of uncertainty which is discussed further in Section 4.1.

3.1.2 Eddy covariance constraints

The standard emission rate E_0 was constrained using hourly eddy covariance measurements at the BR-Sa1 site. For this experiment, MEGAN was driven with locally measured air temperature data. The a priori MEGAN emissions are biased low relative to the BR-Sa1 observations, but this bias is largely corrected for in the optimization (Figure 5 (a)). The posterior E_0 is well-constrained and approximately 36% larger than the a priori value (Figure 5 (b)).

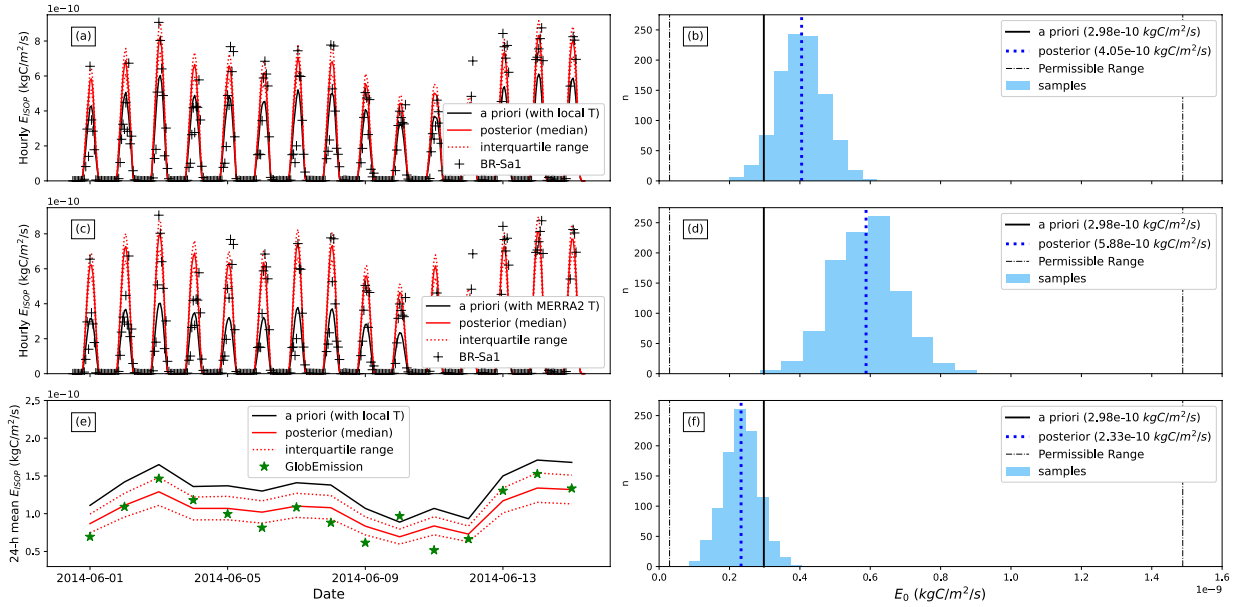

Figure 5: A priori and a posteriori isoprene fluxes and E_0 estimated with BR-Sa1 eddy covariance data (a–d) and GlobEmission fluxes (e–f). (a) Comparison of BR-Sa1 eddy covariance fluxes (+ symbols) with a priori (black line) and a posteriori (red line) MEGAN fluxes, estimated with local temperature data. (b) A priori E_0 and a posteriori distribution of E_0 calculated with local temperature data. (c) Comparison of BR-Sa1 eddy covariance fluxes (+ symbols) with a priori (black line) and a posteriori (red line) MEGAN fluxes, calculated with MERRA-2 temperature

data. (d) A priori E_0 and a posteriori distribution of E_0 inferred with MERRA-2 temperature data. (e) Comparison of GlobEmission isoprene fluxes (green stars) with 24-h mean a priori (black line) and a posteriori (red line) MEGAN fluxes, estimated with local BR-Sa1 temperature data. (f) A priori E_0 and a posteriori distribution of E_0 inferred with local temperature data. The dotted red lines in panels (a), (c), and (e) denote the interquartile range. The dashed black vertical lines in (b), (d), and (f) denote the uniform a priori E_0 distribution, indicating the permissible range of E_0 values.

The optimization was highly sensitive to the temperature input data. Figure 5 (c) shows that driving MEGAN with MERRA-2 temperatures leads to a significant reduction in the a priori emissions. This is due to the negative bias of the MERRA-2 temperature relative to the measured temperature at BR-Sa1 (approximately 1 – 2 degrees K), which was also reported by Sarkar et al (2020). The enhanced negative emission bias is compensated for by increasing the posterior E_0 by 45% relative to the local temperature case (Figure 5 (d)). While the posterior emissions are similar in both cases (Figures 5 (a) and 5 (c)), the E_0 values which produce those emissions are very different. This is another example of equifinality, as described in Section 3.1.1. Sensitivity to model input data is a significant source of uncertainty and is discussed further in Section 4.2.

Optimizing MEGAN with daily top-down emissions from GlobEmission (Figure 5 (e)) resulted in a 22% reduction in E_0 relative to the a priori (Figure 5 (f)). The large discrepancy between the top-down and eddy covariance constraints at BR-Sa1, discussed further in Section 4.3, has been previously reported in the Amazon (Gu et al., 2017). It could be attributed to the low bias in the OMI CH_2O columns, which translates to low-biased GlobEmission fluxes (see Section 2.2.2).

3.2. Optimization of temperature response $\boldsymbol{\gamma_T}$

3.2.1 Tapajós National Forest (AmeriFlux Site BR-Sa1, Brazil)

Figure 6 (a) shows the BR-Sa1 isoprene flux time series from 1-16 June 2014, along with the MEGAN a priori isoprene emissions calculated using locally measured temperature data (Sarkar et al., 2020). The time series of MEGAN $\boldsymbol{\gamma}$ factors plotted in Figure 6 (b) shows that the variability in MEGAN is dominated by $\boldsymbol{\gamma}_{\mathrm{T}}$ and $\boldsymbol{\gamma}_{\mathrm{PAR}}$, while $\boldsymbol{\gamma}_{\mathrm{LAI}}$, $\boldsymbol{\gamma}_{\mathrm{AGE}}$, and $\boldsymbol{\gamma}_{\mathrm{SM}}$ have constant values at or very close to 1 throughout the time series. Figure 6 (c) shows that both the modelled and observed emission variability is directly proportional to the temperature and sunlight response ($\boldsymbol{\gamma}_{\mathrm{T}} \times \boldsymbol{\gamma}_{\mathrm{PAR}}$), with a correlation of $r^2 = 0.799$ for the observed emissions. The steeper slope of the

observed linear fit relative to MEGAN in Figure 6 (c) is due to the large afternoon emission peaks observed in Figure 6 (a) that are not apparent in the a priori emissions. The linear fit allows us to derive an observed temperature response γ'_T from the BR-Sa1 flux measurements.

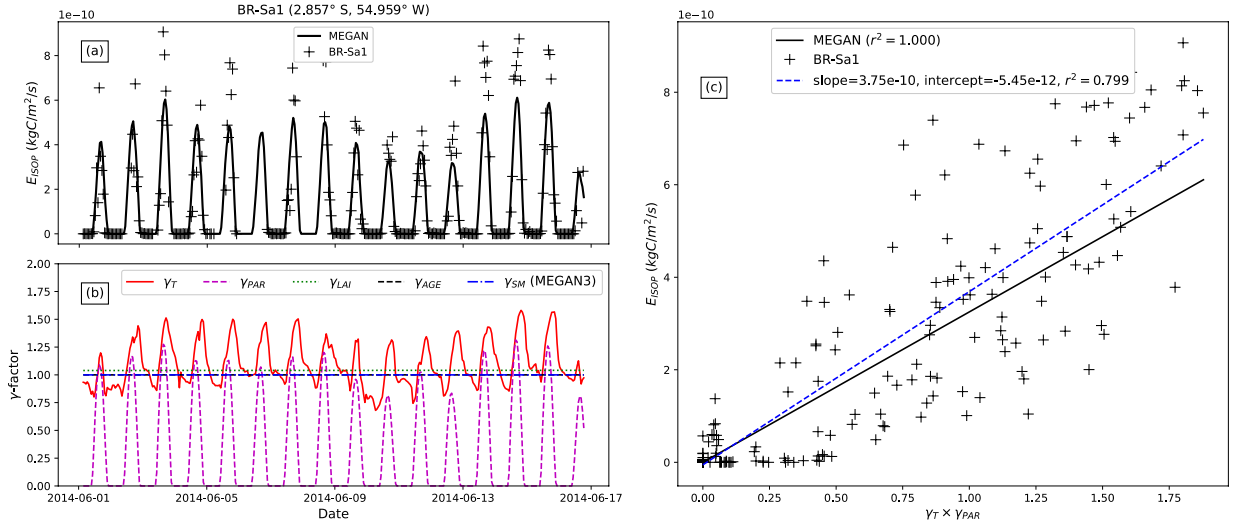

Figure 6: (a) Time series of BR-Sa1 (+ symbols) and MEGAN (solid black line) fluxes between 1–16 June 2014. (b) Time series of $\boldsymbol{\gamma}_{\mathrm{T}}$, $\boldsymbol{\gamma}_{\mathrm{PAR}}$, $\boldsymbol{\gamma}_{\mathrm{LAI}}$, $\boldsymbol{\gamma}_{\mathrm{AGE}}$, and $\boldsymbol{\gamma}_{\mathrm{SM}}$ between 1–17 June 2014. (c) Correlation between the observed BR-Sa1 fluxes (+ symbols) and the product $\boldsymbol{\gamma}_{\mathrm{T}} \times \boldsymbol{\gamma}_{\mathrm{PAR}}$ in MEGAN. The blue dashed line in (c) indicates the linear fit between the BR-Sa1 fluxes and $\boldsymbol{\gamma}_{\mathrm{T}} \times \boldsymbol{\gamma}_{\mathrm{PAR}}$, whereas the solid black line denotes the fit between the MEGAN fluxes and $\boldsymbol{\gamma}_{\mathrm{T}} \times \boldsymbol{\gamma}_{\mathrm{PAR}}$.

The linear fit in Figure 6 (c) gives

$$\gamma'_T = \frac{E_{ISOP} - b}{m\gamma_{PAR}}, \tag{8}$$

where E_{ISOP} is the observed isoprene emission rate, $\boldsymbol{\gamma'_T}$ is the observed temperature response, γ_{PAR} is the MEGAN sunlight response, and m and b are the slope and intercept of the linear fit, respectively. This is equivalent to solving the Equation (1) for $\boldsymbol{\gamma}_{\mathrm{T}}$, with a slope of $m = E_0 \times C_{CE} \times \gamma_{LAI} \times \gamma_{AGE} \times \gamma_{CO2} \times \gamma_{SM}$ and an intercept of b = 0. The strong correlation in Figure 6 (c) indicate that the BR-Sa1 flux observations are broadly consistent with the diurnal variability in MEGAN which has been previously reported at this site (Sarkar et al., 2020).

This derivation of $\boldsymbol{\gamma'_T}$ is sensitive to errors in model inputs and the other MEGAN $\boldsymbol{\gamma}$ factors. Uncertainties in E_0, $\boldsymbol{\gamma}_{\mathrm{LAI}}$, $\boldsymbol{\gamma}_{\mathrm{AGE}}$, and $\boldsymbol{\gamma}_{\mathrm{SM}}$, which are implicit in the slope m, are negligible as all four terms are constant for the BR-Sa1 time series. Constant errors in these terms would not impact

the agreement between the MEGAN γ_T and observed γ'_T because both quantities would be scaled by the same constant slope *m* in Equation (8). Uncertainties in γ_{PAR} and in the PAR input data are important due to the large influence of γ_{PAR} on emission variability on short time scales, particularly given the lack of available local PAR measurements at BR-Sa1. We filtered the observations for MERRA-2 PPFD > 650 μmol/m^2/s (equivalent to PAR > 136 W/m^2) to limit the impact of γ_{PAR} uncertainties on our calculation of γ'_T. This restricts our analysis to intervals centered around midday (approximately 10 am – 4 pm local time), limiting the variability of γ_{PAR} and preventing uncertainties in γ_{PAR} from being amplified under low-light conditions when the denominator in (8) becomes small. Further increasing the PPFD threshold beyond 650 μmol/m^2/s did not significantly reduce scatter but led to excessive data loss. To avoid errors due to model inputs, only the locally measured temperature was used to drive MEGAN in the γ_T optimization.

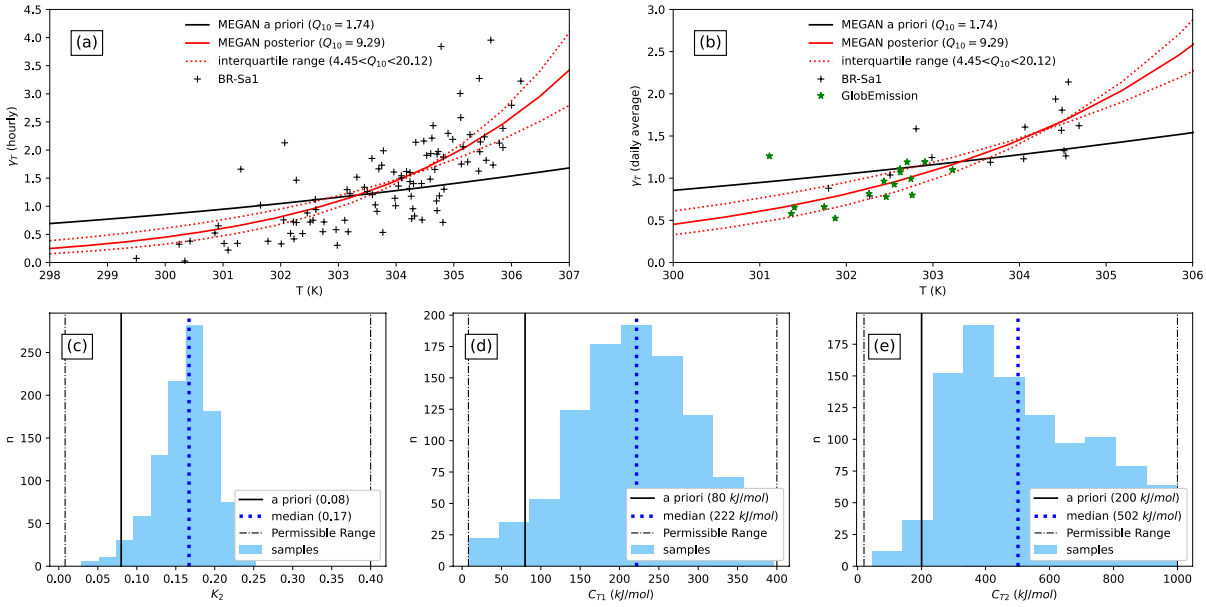

Figure 7: (a) A priori MEGAN γ_T (black), median posterior γ_T (solid red) and interquartile range (dotted red) as a function of temperature (calculated using Equation (2)) compared with the observed γ'_T (+ symbols) at BR-Sa1 (calculated using Equation (8)). (b) As in (a), but resampled from hourly to daily averages for comparison with OMI-based GlobEmission-derived temperature response (green stars). The Q_{10} values in (a) and (b) indicate the fractional change in γ_T between 303 – 313K. Posterior parameter distributions for K_2, C_{T1}, and C_{T2} are shown in panels (c) – (e) (light blue). The median posterior values are indicated by the dashed blue lines in (c) – (e), while the a priori values

509 are indicated by the solid black lines. The uniform a priori parameter distribution is indicated by the dashed black
510 vertical line in (c) – (e).

511 The filtered $\boldsymbol{\gamma}'_{\boldsymbol{T}}$ observations are shown in Figure 7 (a), along with the MEGAN a priori $\boldsymbol{\gamma}_{\mathrm{T}}$.
512 The observed $\boldsymbol{\gamma}'_{\boldsymbol{T}}$ is more sensitive to temperature compared to the a priori, which is consistent
513 with the sharp afternoon emission peaks in the BR-Sa1 time series data that are not captured by
514 MEGAN in Figure 6 (a).

515 MEGAN has different sensitivity to each of the $\boldsymbol{\gamma}_{\mathrm{T}}$ parameters T_{Max}, K_1, K_2, C_{T1}, and C_{T2}
516 (see Supplementary Text S6). We therefore ran optimization experiments using all 31 unique
517 combinations of these 5 parameters to identify the largest combination that could be reliably
518 constrained by the observations while significantly improving model-observation agreement. The
519 subset of K_2, C_{T1}, and C_{T2} yielded the best results, with the posterior $\boldsymbol{\gamma}_{\mathrm{T}}$ agreeing closely with the
520 observed $\boldsymbol{\gamma}'_{\boldsymbol{T}}$ (Figure 7 (a)). The posterior $\boldsymbol{\gamma}_{\mathrm{T}}$ is more than 5 times more sensitive to temperature
521 (Q_{10} = 9.29) than the a priori (Q_{10} = 1.74). This heightened temperature sensitivity is consistent
522 with previous studies, which have found that the a priori MEGAN $\boldsymbol{\gamma}_{\mathrm{T}}$ parameterization
523 underestimates emissions at high temperatures for Australian *eucalypt* species (Emmerson et al.,
524 2020) and Arctic vegetation (Angot et al., 2020; Kramshøj et al., 2016; Seco et al., 2020, 2022).
525 Figure 7 (b) shows that the median posterior $\boldsymbol{\gamma}_{\mathrm{T}}$ is also consistent with the observed $\boldsymbol{\gamma}'_{\boldsymbol{T}}$ as
526 calculated with daily top-down GlobEmission fluxes, indicating that the top-down and eddy
527 covariance constraints are consistent with respect to the temperature dependence of emissions. The
528 K_2, C_{T1}, and C_{T2} distributions from the MHMCMC optimization are shown in Figure 7 (c) – (e).
529 All three parameters are well constrained and significantly different from the MEGAN a priori
530 values (K_2: posterior = 0.17, prior = 0.08; C_{T1}: posterior = 222 kJ/mol, prior = 80 kJ/mol; C_{T2}:
531 posterior = 502 kJ/mol, prior = 200 kJ/mol). Optimization of an alternative parameter combination
532 is shown in Supplementary Text S7.

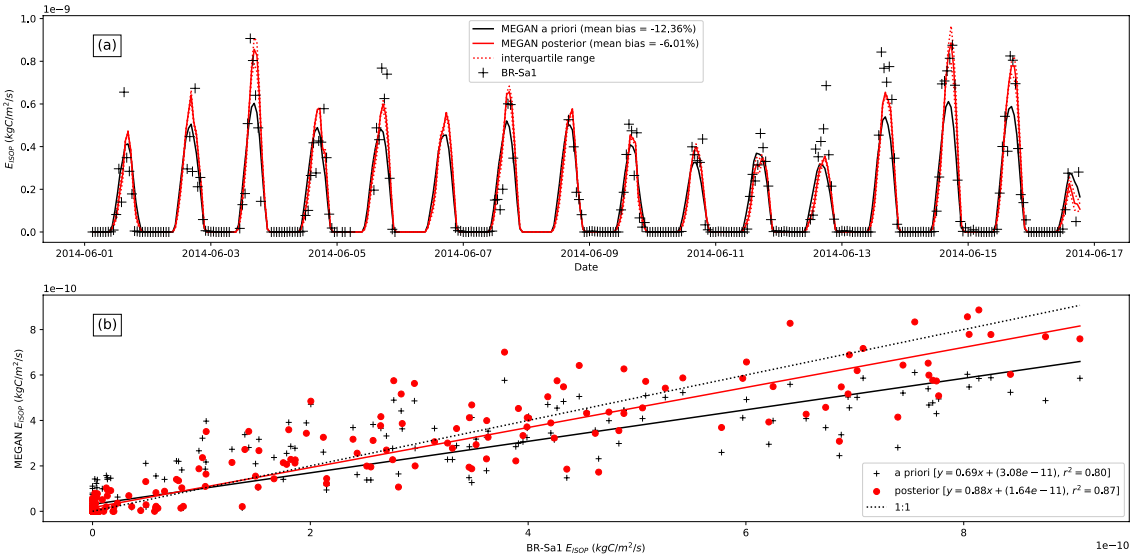

Figure 8: (a) A priori and posterior MEGAN isoprene flux estimates at BR-Sa1 from 1–16 June 2014. The posterior emissions were calculated using the optimized $\boldsymbol{\gamma}_T$ based on eddy covariance observations (+ symbols). The dotted red line denotes the interquartile range on the posterior emission estimate. The mean bias of the a priori and a posteriori MEGAN emissions relative to the BR-Sa1 measurements are indicated. (b) Correlation between observed and modelled hourly mean isoprene emission rates at BR-Sa1. The solid red line is a linear fit to the posterior modelled emissions (red circles), whereas the solid black curve is a linear fit to the a priori emissions (+ symbols).

Figure 8 (a) shows that the posterior MEGAN isoprene emissions are in good agreement with the BR-Sa1 measurements, with a mean bias reduction of ~6% relative to the a priori. In particular, the posterior estimate is better able to capture the observed day-to-day variability in the afternoon peak emissions. This contrasts with the E_0 optimization at BR-Sa1 (Figure 5 (a)). In that case, the temporal variability of the model was fixed in the optimization because E_0 is a simple scaling factor. Figure 8 (b) shows the improved correlation between the posterior and observed emissions (r^2=0.87 for the posterior, versus r^2=0.80 for the a priori).

3.2.2 Wytham Woods (WIsDOM campaign, Oxford, UK)

Figure 9 (a) shows the WIsDOM isoprene mixing ratio measurements from 25 May to 21 June 2018 at Wytham Woods (Ferracci et al., 2020), resampled to hourly averages. Because the WIsDOM measurements are isoprene mixing ratios, a direct comparison with MEGAN is not possible without the use of an atmospheric model. However, an observed temperature response $\boldsymbol{\gamma}'_T$ can be obtained using the same methodology as at BR-Sa1 if the variability in the observed

isoprene mixing ratios is primarily due to the temperature and sunlight emission response ($\gamma_T \times \gamma_{PAR}$). This requires that we filter the WIsDOM observations to minimize variability in photochemical loss rates and dispersion, the two main isoprene removal pathways (Ferracci et al., 2020). Of these two pathways, dispersion is the primary loss mechanism due to the short residence time of isoprene within the forest canopy (seconds to minutes; Fuentes et al., 2007; Gerken et al., 2017) compared to its photochemical oxidation lifetime (30min - 1 hour) (Ferracci et al., 2020).

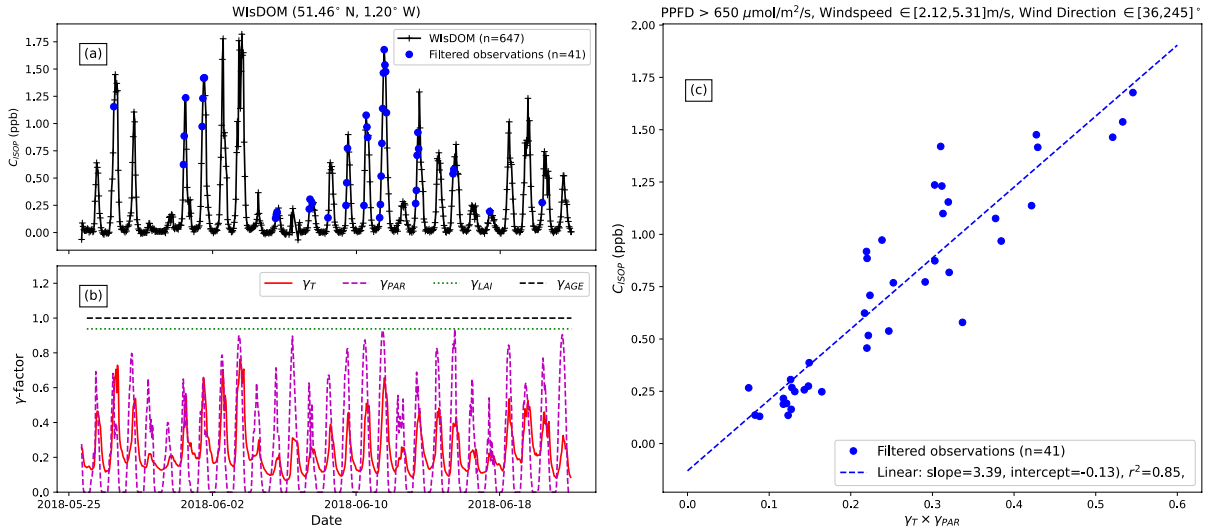

Figure 9: (a) Isoprene mixing ratios at the top of the canopy at Wytham Woods from 26 May – 21 June 2018 resampled to hourly averages. The blue circles indicate the observations which remain after filtering for PPFD > 650 μmol/m^2/s, wind speed between 1.88 m/s and 4.73 m/s, and wind direction between 36° and 245°. (b) Time series of a priori MEGAN γ factors at Wytham Woods. (c) Correlation between the filtered WIsDOM isoprene mixing ratio measurements (blue circles) and the product $\gamma_T \times \gamma_{PAR}$ in MEGAN. The blue dashed line in (c) indicates the linear fit between the WIsDOM measurements and $\gamma_T \times \gamma_{PAR}$.

We filter the preheatwave isoprene measurements for PPFD > 650 μmol/m^2/s to reduce variability in photochemical isoprene loss rates and in the sunlight-driven emission response γ_{PAR} following the methodology of Ferracci et al (2020). We then keep only those measurements with a concurrent GEOS-FP 10-metre windspeed within one standard deviation of the mean preheatwave windspeed (2.12 – 5.31 m s^{-1}) to minimize variability in isoprene dispersion rates. As a final step, only data with a GEOS-FP 10-metre wind direction within one standard deviation of the mean (between 36° and 245°) were kept to minimize variability in isoprene source strength.

The filtered time series is shown in Figure 9 (a). Additional information about this data filtering methodology is available in Ferracci et al (2020).

The MEGAN $\boldsymbol{\gamma}$ factors for the preheatwave period at Wytham Woods are shown in Figure 9 (b). The MEGAN3 $\boldsymbol{\gamma}_{\mathrm{SM}}$ was not available for 2018, but the insignificance of drought stress at Wytham Woods during the preheatwave period has been previously reported (Ferracci et al., 2020; Otu-Larbi et al., 2020). The leaf phenology activity factors $\boldsymbol{\gamma}_{\mathrm{LAI}}$ and $\boldsymbol{\gamma}_{\mathrm{AGE}}$ exhibited minimal variability (Ferracci et al., 2020). The variability of observed isoprene mixing ratios (Figure 9 (a)) closely follows the variability of $\boldsymbol{\gamma}_{\mathrm{T}}$ and $\boldsymbol{\gamma}_{\mathrm{PAR}}$ (Figure 9 (b)). This is confirmed in Figure 9 (c), which shows the strong linear correlation between the filtered Wytham Woods isoprene observations and $\boldsymbol{\gamma}_{\mathrm{T}} \times \boldsymbol{\gamma}_{\mathrm{PAR}}$. Following the same methodology as at BR-Sa1, we define an observed temperature response $\boldsymbol{\gamma}'_{\boldsymbol{T}}$ as

$$\gamma'_T = \frac{C_{ISOP} - b}{m\gamma_{PAR}}, \tag{9}$$

where C_{ISOP} is the filtered isoprene mixing ratio, $\boldsymbol{\gamma}_{\mathrm{PAR}}$ is the MEGAN sunlight response, and m and b are the slope and intercept of the linear fit in Figure 9 (c). Figure 10 (a) shows that there is good agreement between the observed $\boldsymbol{\gamma}'_{\boldsymbol{T}}$ and the a priori MEGAN $\boldsymbol{\gamma}_{\mathrm{T}}$, which is supported by the modeling work by Otu-Larbi et al. (2020) for the preheatwave period. This is in stark contrast to the behavior observed in Figure 7 (a) at BR-Sa1.

The observed $\boldsymbol{\gamma}'_{\boldsymbol{T}}$ was used to constrain the MEGAN $\boldsymbol{\gamma}_{\mathrm{T}}$ parameters at Wytham Woods. We tested all 31 parameter combinations, but for consistency with BR-Sa1 we present only the optimization results for the subset K_2, C_{T1}, and C_{T2} (Figure 10). The posterior $\boldsymbol{\gamma}_{\mathrm{T}}$ (Q_{10} = 2.29) is within error of the MEGAN a priori (Q_{10} = 1.74), and both quantities are in good agreement with the observed $\boldsymbol{\gamma}'_{\boldsymbol{T}}$ (Figure 10 (a)). Figures 10 (b) – (d) show that the $\boldsymbol{\gamma}_{\mathrm{T}}$ parameters are not as precisely constrained as at BR-Sa1, which is due to the lower sensitivity of MEGAN to K_2 and C_{T1} at the lower ambient temperatures of Wytham Woods (see Supplementary Text S6). Unlike at BR-Sa1 (Figures 8 (c) – (e)), the median posterior parameters are all relatively close to their a priori values. This indicates that the temperature sensitivity of isoprene emissions at Wytham Woods during the preheatwave period is accurately represented by the a priori parameterization of $\boldsymbol{\gamma}_{\mathrm{T}}$, in contrast to what was observed at BR-Sa1. This is discussed further in Section 4.4.

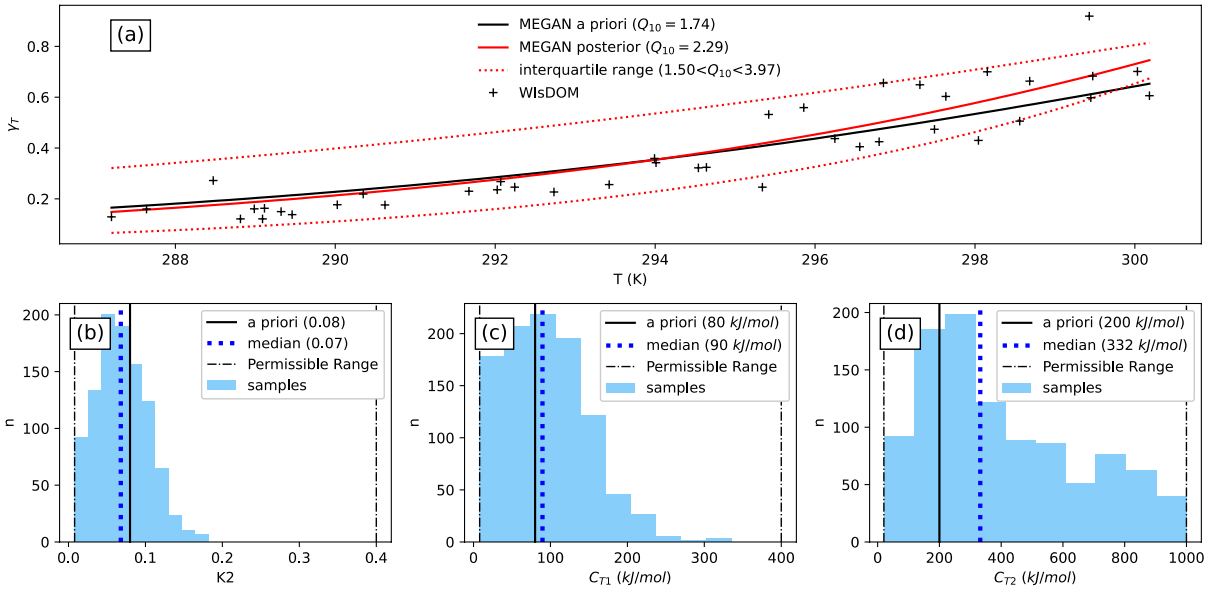

Figure 10: (a) A priori MEGAN $\boldsymbol{\gamma}_T$ (black), median posterior $\boldsymbol{\gamma}_T$ (solid red) and interquartile range (dotted red) as a function of temperature (calculated using Equation (2)) compared with the observed γ'_T (+ symbols) (calculated using Equation (9) from filtered mixing ratio observations) at Wytham Woods. Posterior parameter distributions for K_2, C_{T1}, and C_{T2} are shown in panels (b) – (d) (light blue). The median posterior values are indicated by the dashed blue lines in (b) – (d), while the a priori values are indicated by the solid black lines. The uniform a priori parameter distribution is indicated by the dashed black vertical line in (b) – (d).

4. Discussion

4.1 Sensitivity to drought stress

The sensitivity of MEGAN to drought stress is a source of error in our parameter optimization experiments. Our optimization of E_0 in Eastern Australia was highly sensitive to the MEGAN3 drought stress activity factor $\boldsymbol{\gamma}_{SM}$ (Figure 3), consistent with previous studies (Emmerson et al., 2019). Enabling $\boldsymbol{\gamma}_{SM}$ increased the posterior E_0 by 40% relative to the no-$\boldsymbol{\gamma}_{SM}$ optimization because much of the bias between MEGAN and GlobEmission was already accounted for by the drought stress factor. We can mitigate this source of uncertainty by focusing our optimization efforts on regions and time periods which are not subject to drought stress, such as the Eastern Amazon during the BR-Sa1 measurement period and Wytham Woods during the 2018 preheatwave period.

While in principle we could enable the MEGAN3 γ_{SM} in all experiments to account for drought stress, there are uncertainties in γ_{SM} itself which limit the usefulness of such an approach. In particular, γ_{SM} does not account for the observed increase in isoprene emission rates that accompany moderate drought stress (Ferracci et al., 2020; Otu-Larbi et al., 2020; Potosnak et al., 2014; Saunier et al., 2020). Including γ_{SM} in our optimization experiments therefore risks misattributing uncertainties in γ_{SM} to other model parameters including E_0 and γ_T. A recent study has found that the parameterization of the MEGAN γ_{SM} can be improved using a combination of ecosystem-scale isoprene flux measurements and satellite-derived soil moisture (Opacka et al., 2022). Future work in this direction, along with updates to the γ_{SM} algorithm to account for the impact of moderate drought stress on emissions, should improve the utility of the MEGAN drought stress response and allow parameter optimization experiments to be reliably performed under drought conditions.

4.2 Sensitivity to errors in model input data: temperature and LAI

The sensitivity of MEGAN to meteorology and landcover inputs is another source of uncertainty in our parameter optimization. Biases in temperature and LAI inputs are particularly important, as these are among the primary drivers of short- and long-term variability in MEGAN, respectively (Alves et al., 2016, 2018; Chen et al., 2018; Guenther et al., 1993; Opacka et al., 2021). The impact of input biases on our optimization is apparent in Figures 5 (a) – (d), where there is a 45% difference between the posterior E_0 at BR-Sa1 depending on whether local or MERRA-2 temperature data was used due to discrepancies between the two temperature datasets. These discrepancies could reflect region-specific biases in the MERRA-2 temperature data (e.g., Draper et al., 2018; Gupta et al., 2020) or could simply be due to the coarse spatial resolution of the MERRA-2 data.

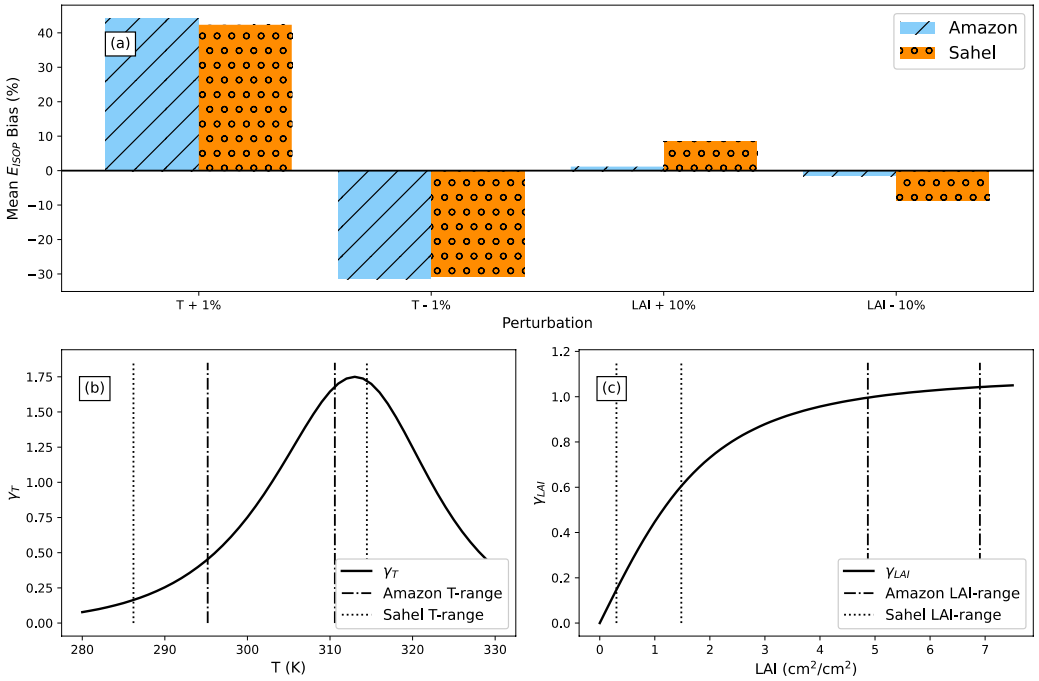

Figure 11: (a) Mean annual (2014) isoprene emission bias as calculated with MEGANv2.1 in the Amazon (light blue) and the Sahel (orange) when running the model with biased temperature (T ± 1 %) or leaf area index (LAI ± 10 %) input data. (b) A priori MEGAN $\boldsymbol{\gamma}_{\mathrm{T}}$ (calculated using Equation (2)) superimposed with range of temperatures in the un-perturbed MERRA2 input data for both the Amazon (dash-dotted line) and the Sahel (dotted line) in 2014. (c) A priori MEGAN $\boldsymbol{\gamma}_{\mathrm{LAI}}$ (calculated using Equation S5) superimposed with range of LAI values in the unperturbed 8-day average MODIS LAI input data for both the Amazon and the Sahel in 2014. The heightened sensitivity of MEGAN to LAI in the Sahel is indicated by the steeper slope of $\boldsymbol{\gamma}_{\mathrm{LAI}}$ in the Sahel LAI range compared to the Amazon.

The sensitivity of MEGAN isoprene emissions to temperature and LAI inputs is illustrated in Figure 11, which shows the impact of temperature and LAI biases on annual mean (2014) MEGAN isoprene emission rates in the Eastern Amazon and Western Sahel regions. A 1% temperature bias (~2.5–3.5K, consistent with the difference between the local BR-Sa1 and MERRA-2 temperatures) leads to emission biases exceeding 30% (Figure 11 (a)) in both regions. Note that this sensitivity test used the a priori parameterization of $\boldsymbol{\gamma}_{\mathrm{T}}$ (Q_{10} = 1.74). The impact of temperature input errors would be even more significant if we instead used the posterior parameterization from BR-Sa1 due to the enhanced temperature sensitivity of isoprene emissions at that site (Q_{10} = 9.29). The sensitivity to LAI is much lower and varies significantly between the two regions, with a 10% LAI bias (based on the estimated uncertainty of the MODIS LAI product (Fang et al., 2013)) leading to a 10% emission bias in the Sahel but only a 2% emission bias in the

Amazon (Figure 11 (a)). This variable sensitivity is due to the saturation of γ_{LAI} in high-LAI environments such as the Amazon (Figure 11 (c)), which represents the shading of foliage by the upper layers of the forest canopy (Guenther et al., 2006) (see Equation S5 in Supplementary Text S8).

The high sensitivity of MEGAN emissions to temperature biases shows the importance of using reliable locally constrained model inputs to drive MEGAN, such as the locally measured temperature at BR-Sa1 and Wytham Woods. Biases in LAI are unlikely to have any significant impact on our optimization in high-LAI environments such as the Amazon but may introduce biases of 10% or more in low-LAI environments like the Sahel. In such cases, locally constrained LAI data may be necessary to mitigate this source of error and obtain reliable posterior parameters. The spatial variability of the MEGAN temperature and LAI sensitivity on a global scale is presented in Supplementary Text S8.

It is possible that the sensitivity to LAI presented here is an underestimate of the true sensitivity because it does not account for leaf age fractionation (i.e., what fraction of the total LAI is made up of immature, mature, or senescing leaves). Isoprene emissions are known to vary as a function of leaf age (Guenther et al., 2006), and changes in leaf age fractionation have been shown to significantly impact photosynthesis rates and isoprene production in the Amazon basin (Alves et al., 2018; Wu et al., 2016). While this effect is modeled by the leaf age activity factor γ_{AGE} in MEGAN, the leaf age fractionations used to calculate γ_{AGE} are derived from changes in the MODIS 8-Day LAI product (Guenther et al., 2006). This method of calculating leaf age fractionations inherently reduces the variability of γ_{AGE} in environments with relatively constant total LAI such as the Amazon. The use of directly measured leaf age fractionations has been shown to improve modelled isoprene emissions in the Amazon on seasonal time scales (Alves et al., 2018). Given the short duration of the BR-Sa1 and preheatwave WIsDOM time series it is unlikely that γ_{AGE} would have a significant impact on our temperature optimization experiments, but it may lead to significant seasonally-dependent biases to our 9-year E_0 optimization experiments.

4.3 Discrepancies between top-down and eddy covariance constraints

The a priori MEGAN isoprene emissions at BR-Sa1 were biased high relative to the top-down constraints from GlobEmission (OMI-based) and SOLFEO (TROPOMI-based) but biased low relative to the local eddy covariance measurements, leading to inconsistent posterior E_0 values.

The disagreement between the top-down and eddy covariance constraints is likely due to the low bias of the OMI and TROPOMI CH_2O columns, discussed in Section 2.2.1. Note that similar discrepancies between top-down and eddy covariance isoprene fluxes have previously been reported in the Amazon (Gu et al., 2017), which may also have been due to underestimated spaceborne columns. Correcting the low OMI and TROPOMI CH_2O column biases with ground-based validation studies in the Amazon would result in higher top-down emissions than in GlobEmission and SOLFEO, and consequently smaller discrepancies between the top-down and eddy covariance constraints. Other uncertainties in top-down emissions may also contribute to these discrepancies. Top-down estimates are sensitive to CH_2O retrieval errors (Millet et al., 2006), chemistry-transport model errors (Barkley et al., 2013; Stavrakou et al., 2015), spatial smearing errors due to non-local CH_2O sources (Palmer et al., 2003; Turner et al., 2012), and non-biogenic CH_2O background sources such as biomass burning and methane oxidation (Marais et al., 2012; Wolfe et al., 2016).

Chemistry-transport model errors have been identified as a source of particular concern in top-down isoprene emissions in tropical regions due to their impacts on the CH_2O yield from isoprene oxidation as well as the diurnal variability of isoprene and CH_2O concentrations (Kefauver et al., 2014). This is particularly relevant in the Amazon due to the large uncertainties in regional NO_2 and OH, both of which impact isoprene oxidation chemistry (Wells et al., 2020; Liu et al., 2016; Liu et al., 2018; Jeong et al., 2022; Wolfe et al., 2016). Stronger constraints on this chemistry, such as direct space-based isoprene retrievals from the Cross-track Infrared Sounder (CrIS) (Fu et al., 2019; Wells et al., 2020), may mitigate some of these uncertainties. For example, CrIS isoprene column retrievals have recently been validated using ground-based remotely sensed isoprene columns in the Amazon and combined with OMI CH_2O and NO_2 retrievals to reveal large day-to-day variability in isoprene oxidation lifetime (Wells et al., 2022). Taking full advantage of top-down emissions estimates to optimize MEGAN will require a better understanding of the sensitivity of these estimates to chemistry-transport model errors. This sensitivity could be quantified by obtaining top-down emission estimates using multiple chemistry-transport models (Miyazaki et al., 2020). Simulated inversion experiments could also be performed to directly probe the impact of chemistry model errors on top-down emissions based on satellite pseudo-observations.

728 An additional important issue is the representativeness of the eddy covariance observations
729 at BR-Sa1. The top-down fluxes are provided at 0.5° × 0.5° spatial resolution, corresponding to a
730 pixel size of roughly 3000 km^2 near the equator (Bauwens et al., 2016). Extrapolating the BR-Sa1
731 flux measurements to this large pixel size can incur biases due to the spatial heterogeneity of
732 isoprene emissions (Batista et al., 2019; Li et al., 2021). However, independent aircraft-based eddy
733 covariance isoprene flux measurements from the Green Ocean Amazon campaign (Gu et al., 2017)
734 show similar discrepancies with OMI-based emissions, which suggests that this problem cannot
735 be entirely due to the representativeness of the BR-Sa1 measurements.
736

737 ### 4.4 Variability of γ_T between ecosystems

738 The optimized γ_T at the Amazonian BR-Sa1 site was roughly five times more sensitive to
739 temperature (Q_{10} = 9.29) than the MEGAN a priori (Q_{10} = 1.74). Due to the increased temperature
740 sensitivity of the posterior γ_T, applying this parameterization to the broader Amazon region would
741 have potentially significant impacts on regional isoprene emissions. However, it is difficult to
742 extrapolate results from a single measurement site to a broader geographic area due to landcover
743 and species distribution heterogeneity. Measurements from other Amazonian sites could be used
744 to determine whether this posterior parameterization is representative of the region.
745 Unlike at BR-Sa1, the optimized γ_T at the UK Wytham Woods site was not significantly
746 different from the a priori. This demonstrates that the performance of γ_T varies across ecosystems,
747 and that any updates to the parameterization of γ_T should be applied on an ecosystem-specific
748 scale. This has been done previously for Australian *eucalypt* species (Emmerson et al., 2020), and
749 other studies have demonstrated that this may also be necessary for Arctic vegetation (Angot et
750 al., 2020; Kramshøj et al., 2016; Seco et al., 2020, 2022). Accurately modeling the sensitivity of
751 isoprene emissions to temperature will depend on the development of an ecosystem-specific
752 parameterization for γ_T, which could be derived using our methodology wherever suitable
753 observations are available. This has the potential to significantly improve models of atmospheric
754 chemistry in a warming climate or during severe heat wave events (Emmerson et al., 2020) in
755 addition to improving day-to-day emission variability.
756 An ecosystem-specific parameterization would be particularly useful when relying on
757 simplified canopy models such as the PCEEA algorithm, which is primarily weighted towards
758 warm broadleaf forests yet is applied globally to all ecosystem types (Guenther et al., 2006). The

PCEEA algorithm has been shown to introduce local emission biases of up to 25% relative to more sophisticated multi-level canopy model versions of MEGAN even though global mean emissions are consistent to within 5% (Guenther et al., 2006). However, more sophisticated canopy physics models can still introduce ecosystem-specific temperature and sunlight biases, leading to vegetation-dependent emission errors (Silva et al., 2020).

An unresolved source of uncertainty is the impact of drought stress on the emission temperature response. While the MEGAN a priori $\boldsymbol{\gamma}_T$ adequately describes the observed temperature response at Wytham Woods during the preheatwave period, increased temperature sensitivity was observed after the onset of a severe drought and heatwave in June 2018 (Ferracci et al., 2020; Otu-Larbi et al., 2020). A similar drought response was observed in an oak-dominated forest in the Missouri Ozarks (Seco et al., 2015). These observations are consistent with current conceptual models of the drought stress response (Potosnak et al., 2014), and show that the temperature sensitivity of emissions is a function not only of ecosystem type but also of current environmental conditions including drought stress.

5. Summary and conclusions

We have used Bayesian model-data fusion to optimize the standard emission rate and the temperature activity factor in MEGAN, using top-down isoprene fluxes derived from satellite observations in four regions (Amazon, Southeast USA, Western Sahel, and Eastern Australia), eddy covariance isoprene flux measurements in the Amazon, and isoprene atmospheric mixing ratio measurements in the United Kingdom.

Optimization of the standard emission rate E_0 with satellite constraints reduced model biases in the Amazon, the Southeast USA, the Western Sahel, and Eastern Australia. The optimized E_0 values were highly sensitive to model input errors. Sensitivity to temperature errors was extremely high in all regions, while sensitivity to total LAI errors was only significant in low-LAI environments such as the Western Sahel. The impact of the MEGAN3 drought stress response on the optimization was spatially and temporally variable, with the largest impact being seen in Eastern Australia. Uncertainties in the drought stress response are a major obstacle for reliable parameter optimization under drought conditions.

We optimized E_0 at the Amazonian BR-Sa1 field site using both satellite and eddy covariance constraints and found that the two results were inconsistent with one another. This

mismatch may be largely due to the low biases identified in the satellite CH_2O retrievals used as constraints to derive the top-down emissions, leading to an underprediction of the top-down emission fluxes. Chemistry-transport model errors may add more uncertainty to the top-down emissions. Future optimization work using top-down emissions as constraints will be dependent on understanding and reducing the uncertainties in top-down emissions, which could be done through a combination of satellite validation studies with ground-based measurements and modeling studies to assess the impact of chemistry-transport model errors on top-down emissions. Stronger constraints could be placed on isoprene oxidation chemistry in models by combining satellite retrievals of CH_2O with retrievals of isoprene (e.g., Wells et al., 2020), while simultaneous optimization of isoprene and NO_X emissions may also improve top-down constraints due to the strong dependence of isoprene oxidation chemistry on ambient NO_X concentrations (Miyazaki et al., 2020).

Optimization of the temperature response $\boldsymbol{\gamma}_T$ with eddy covariance isoprene emission measurements increased the temperature sensitivity of the model by a factor of 5 (posterior Q_{10} = 9.29 compared to priori Q_{10} = 1.74) at BR-Sa1 and reduced model biases by 6%. By contrast, optimizing $\boldsymbol{\gamma}_T$ with isoprene mixing ratio measurements at the UK-based Wytham Woods site had no significant impact on the model parameters due to the good agreement between the MEGAN a priori $\boldsymbol{\gamma}_T$ and the observations. Enhanced sensitivity of isoprene emissions to temperature, and more specifically the underestimation of emissions at high temperatures, has now been observed in several ecosystem types, including an Amazonian old growth forest (this study), Australian *eucalpyt* trees (Emmerson et al., 2020), and various species of Arctic vegetation (Angot et al., 2020; Kramshøj et al., 2016; Seco et al., 2020, 2022). Drought stress has also been shown to increase temperature sensitivity at multiple temperate sites (Ferracci et al., 2020; Otu-Larbi et al., 2020; Seco et al., 2015). Accurate modeling of isoprene emissions, as well as their impacts on air quality and climate, will require an ecosystem-specific parameterization of the temperature emission response as well as an improved understanding of the drought stress emission response. Such a parameterization could be derived from ground-based isoprene and temperature measurements in a wide range of ecosystems. Existing measurements should be used to evaluate and reparametrize the temperature response where possible, while future measurement campaigns should target a wide range of ecosystem types. The use of longer measurement time series than presented in this paper would allow seasonal impacts such as drought to be investigated as well.

The significance of isoprene emissions at high temperatures or under drought conditions is expected to increase in a warming climate (Emmerson et al., 2020; Saunier et al., 2020), further highlighting the need for a reliable parameterization of the emission temperature response.

Acknowledgments

C.A. DiMaria acknowledges a Canada Graduate Scholarhip - Doctoral (CGS D) grant funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) (application no. PGSD3 - 546721 - 2020). R. Seco acknowledges a Ramón y Cajal grant (RYC2020-029216-I) funded by MCIN/AEI/ 10.13039/501100011033 and by "ESF Investing in your future". IDAEA-CSIC is a Severo Ochoa Centre of Research Excellence (MCIN/AEI, Project CEX2018-000794-S). The BR-Sa1 field measurements were supported by Núcleo de Apoio à Pesquisa no Pará (NAPPA) em Santarém-Pa/Instituto Nacional de Pesquisas da Amazônia (INPA), Programa de Grande Escala Biosfera Atmosfera na Amazônia (LBA) and Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) em Santarém-Pa. V. Ferracci acknowledges funding from the Natural Environmental Research Council (NERC) project "Biodiversity and land-use impacts (BALI) on tropical ecosystems" (NE/K016377/1) in support of the Wytham Woods measurements. Part of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA).

Open Research

Data Availability Statement

The MEGAN 2.1 source code (Guenther et al., 2012) and the MEGAN 3 drought stress activity factors (X. Jiang et al., 2018) can be obtained from https://bai.ess.uci.edu/megan/data-and-code. The MHMCMC Matlab code (Yang et al., 2021) is available at https://doi.org/10.5281/zenodo.4904195. The global OMI-based top-down isoprene flux estimates (Bauwens et al., 2016) are available for download at https://emissions.aeronomie.be/index.php/omi-based/biogenic. The TROPOMI-based top-down isoprene flux estimates for South America are available for download at https://emissions.aeronomie.be/index.php/tropomi-based/isoprene-sa. The data from the WIsDOM campaign (Ferracci et al., 2020) are available from the Natural Environment Research

Council (NERC) Centre for Environmental Data Analysis (CEDA) archive at https://catalogue.ceda.ac.uk/uuid/0c39809848ce47bb850d8ca2045e40f2. The BR-Sa1 isoprene flux and temperature measurements (Sarkar et al., 2020) can be obtained by sending an email to chinmoysarkar8@gmail.com or alex.guenther@uci.edu. The MERRA-2 and GEOS-FP data used in this study have been provided by the Global Modeling and Assimilation Office (GMAO) at NASA Goddard Space Flight Center. The particular files used to drive MEGAN in this study were obtained from the GEOS-Chem (Bey et al., 2001) met-field archive at http://geoschemdata.wustl.edu. The MODIS 8-day LAI product (Yuan et al., 2011, 2020) is available at http://globalchange.bnu.edu.cn/research/laiv6.

References

Alves, E. G., Jardine, K., Tota, J., Jardine, A., Yãnez-Serrano, A. M., Karl, T., Tavares, J., Nelson, B., Gu, D., Stavrakou, T., Martin, S., Artaxo, P., Manzi, A., & Guenther, A. (2016). Seasonality of isoprenoid emissions from a primary rainforest in central Amazonia. *Atmospheric Chemistry and Physics*, *16*(6), 3903–3925. https://doi.org/10.5194/acp-16-3903-2016

Alves, E. G., Tóta, J., Turnipseed, A., Guenther, A. B., Vega Bustillos, J. O., Santana, R. A., Cirino, G. G., Tavares, J. V., Lopes, A. P., Nelson, B. W., de Souza, R. A., Gu, D., Stavrakou, T., Adams, D. K., Wu, J., Saleska, S., & Manzi, A. O. (2018). Leaf phenology as one important driver of seasonal changes in isoprene emissions in Central Amazonia. *Biogeosciences*, *15*(13), 4019–4032. https://doi.org/10.5194/bg-15-4019-2018

Angot, H., McErlean, K., Hu, L., Millet, D. B., Hueber, J., Cui, K., Moss, J., Wielgasz, C., Milligan, T., Ketcherside, D., Bret-Harte, M. S., & Helmig, D. (2020). Biogenic volatile organic compound ambient mixing ratios and emission rates in the Alaskan Arctic tundra. *Biogeosciences*, *17*(23), 6219–6236. https://doi.org/10.5194/bg-17-6219-2020

Arneth, A., Schurgers, G., Lathiere, J., Duhl, T., Beerling, D. J., Hewitt, C. N., Martin, M., & Guenther, A. (2011). Global terrestrial isoprene emission models: Sensitivity to variability in climate and vegetation. *Atmospheric Chemistry and Physics*, *11*(15), 8037–8052. https://doi.org/10.5194/acp-11-8037-2011

Ashworth, K., Wild, O., & Hewitt, C. N. (2010). Sensitivity of isoprene emissions estimated using MEGAN to the time resolution of input climate data. *Atmospheric Chemistry and Physics*, *10*(3), 1193–1201. https://doi.org/10.5194/acp-10-1193-2010

Barkley, M. P., Smedt, I. D., Van Roozendael, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D., Guenther, A., Paulot, F., Marais, E., & Mao, J. (2013). Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns. *Journal of Geophysical Research: Atmospheres*, *118*(12), 6849–6868. https://doi.org/10.1002/jgrd.50552

Batista, C. E., Ye, J., Ribeiro, I. O., Guimarães, P. C., Medeiros, A. S., Barbosa, R. G., Oliveira, R. L., Duvoisin, S., Jardine, K. J., Gu, D., Guenther, A. B., McKinney, K. A., Martins, L. D., Souza, R. A., & Martin, S. T. (2019). Intermediate-scale horizontal isoprene concentrations in the near-canopy forest atmosphere and implications for emission heterogeneity. *Proceedings of the National Academy of Sciences*, *116*(39), 19318–19323. https://doi.org/10.1073/pnas.1904154116

Bauwens, M., Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Wiedinmyer, C., Kaiser, J. W., Sindelarova, K., & Guenther, A. (2016). Nine years of global hydrocarbon emissions based on source inversion of OMI formaldehyde observations. *Atmospheric Chemistry and Physics*, *16*(15), 10133–10158. https://doi.org/10.5194/acp-16-10133-2016

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., & Schultz, M. G. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model Description and Evaluation. *Journal of Geophysical Research: Atmospheres*, *106*(D19), 23073–23095. https://doi.org/10.1029/2001jd000807

Bloom, A. A., Bowman, K. W., Liu, J., Konings, A. G., Worden, J. R., Parazoo, N. C., Meyer, V., Reager, J. T., Worden, H. M., Jiang, Z., Quetin, G. R., Smallman, T. L., Exbrayat, J.-F., Yin, Y., Saatchi, S. S., Williams, M., & Schimel, D. S. (2020). Lagged effects regulate the inter-annual variability of the Tropical Carbon Balance. *Biogeosciences*, *17*(24), 6393–6422. https://doi.org/10.5194/bg-17-6393-2020

Bloom, A. A., & Williams, M. (2015). Constraining ecosystem carbon dynamics in a data-limited world: integrating ecological "common sense" in a model–data fusion framework. *Biogeosciences*, *12*(5), 1299–1315. https://doi.org/10.5194/bg-12-1299-2015

Bloom, A. A., Worden, J., Jiang, Z., Worden, H., Kurosu, T., Frankenberg, C., & Schimel, D. (2015). Remote-sensing constraints on South America fire traits by Bayesian fusion of atmospheric and surface data. *Geophysical Research Letters*, *42*(4), 1268–1274. https://doi.org/10.1002/2014gl062584

Buchholz, R. R., Worden, H. M., Park, M., Francis, G., Deeter, M. N., Edwards, D. P., Emmons, L. K., Gaubert, B., Gille, J., Martínez-Alonso, S., Tang, W., Kumar, R., Drummond, J. R., Clerbaux, C., George, M., Coheur, P.-F., Hurtmans, D., Bowman, K. W., Luo, M., … Kulawik, S. S. (2021). Air Pollution Trends measured from Terra: Co and AOD over industrial, fire-prone, and background regions. *Remote Sensing of Environment*, *256*, 112275. https://doi.org/10.1016/j.rse.2020.112275

Butt, N., Campbell, G., Malhi, Y., Morecroft, M., Fenn, K., & Thomas, M. (2009). *Initial results from establishment of a long-term broadleaf monitoring plot at Wytham Woods*. Oxford: UK. University of Oxford.

Chen, W. H., Guenther, A. B., Wang, X. M., Chen, Y. H., Gu, D. S., Chang, M., Zhou, S. Z., Wu, L. L., & Zhang, Y. Q. (2018). Regional to global biogenic isoprene emission responses to changes in vegetation from 2000 to 2015. *Journal of Geophysical Research: Atmospheres*, *123*(7), 3757–3771. https://doi.org/10.1002/2017jd027934

Claeys, M., Graham, B., Vas, G., Wang, W., Vermeylen, R., Pashynska, V., Cafmeyer, J., Guyon, P., Andreae, M. O., Artaxo, P., & Maenhaut, W. (2004). Formation of secondary organic aerosols through photooxidation of isoprene. *Science*, *303*(5661), 1173–1176. https://doi.org/10.1126/science.1092805

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H., and Van Roozendael, M.: Comparative assessment of

TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements, Atmos. Chem. Phys., 21, 12561–12593, https://doi.org/10.5194/acp-21-12561-2021, 2021.

De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compernolle, S., Van Roozendael, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma, K. F., and Veefkind, P.: Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, Atmos. Meas. Tech., 11, 2395–2426, https://doi.org/10.5194/amt-11-2395-2018, 2018.

Draper, C. S., Reichle, R. H., & Koster, R. D. (2018). Assessment of merra-2 land surface energy flux estimates. *Journal of Climate*, *31*(2), 671–691. https://doi.org/10.1175/jcli-d-17-0121.1

Emmerson, K. M., Palmer, P. I., Thatcher, M., Haverd, V., & Guenther, A. B. (2019). Sensitivity of isoprene emissions to drought over south-eastern Australia: Integrating models and satellite observations of soil moisture. *Atmospheric Environment*, *209*, 112–124. https://doi.org/10.1016/j.atmosenv.2019.04.038

Emmerson, K. M., Possell, M., Aspinwall, M. J., Pfautsch, S., & Tjoelker, M. G. (2020). Temperature response measurements from eucalypts give insight into the impact of Australian isoprene emissions on air quality in 2050. *Atmospheric Chemistry and Physics*, *20*(10), 6193–6206. https://doi.org/10.5194/acp-20-6193-2020

Fang, H., Jiang, C., Li, W., Wei, S., Baret, F., Chen, J. M., Garcia-Haro, J., Liang, S., Liu, R., Myneni, R. B., Pinty, B., Xiao, Z., & Zhu, Z. (2013). Characterization and Intercomparison of Global Moderate Resolution Leaf Area Index (LAI) products: Analysis of climatologies and theoretical uncertainties. *Journal of Geophysical Research: Biogeosciences*, *118*(2), 529–548. https://doi.org/10.1002/jgrg.20051

Ferracci, V., Bolas, C. G., Freshwater, R. A., Staniaszek, Z., King, T., Jaars, K., Otu-Larbi, F., Beale, J., Malhi, Y., Waine, T. W., Jones, R. L., Ashworth, K., & Harris, N. R. (2020). Continuous isoprene measurements in a UK temperate forest for a whole growing season:

Effects of drought stress during the 2018 heatwave. *Geophysical Research Letters*, *47*(15). https://doi.org/10.1029/2020gl088885

Ferracci, V.; Harris, N.; Bolas, C.; Jones, R.; Staniaszek, Z. (2020): Biodiversity and Land Use Impacts on Tropical Forest Ecosystem Function (BALI): Isoprene concentration measurements at Wytham Woods (UK) during the summer of 2018 [Dataset]. Centre for Environmental Data Analysis. https://catalogue.ceda.ac.uk/uuid/0c39809848ce47bb850d8ca2045e40f2

Filella, I., Zhang, C., Seco, R., Potosnak, M., Guenther, A., Karl, T., Gamon, J., Pallardy, S., Gu, L., Kim, S., Balzarolo, M., Fernandez-Martinez, M., & Penuelas, J. (2018). A MODIS photochemical reflectance index (PRI) as an estimator of isoprene emissions in a temperate deciduous forest. *Remote Sensing*, *10*(4), 557. https://doi.org/10.3390/rs10040557

Fu, D., Millet, D. B., Wells, K. C., Payne, V. H., Yu, S., Guenther, A., & Eldering, A. (2019). Direct retrieval of isoprene from satellite-based infrared measurements. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-11835-0

Fuentes, J. D., Wang, D., Bowling, D. R., Potosnak, M., Monson, R. K., Goliff, W. S., & Stockwell, W. R. (2007). Biogenic hydrocarbon chemistry within and above a mixed deciduous forest. *Journal of Atmospheric Chemistry*, *56*(2), 165–185. https://doi.org/10.1007/s10874-006-9048-4

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., … Zhao, B. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/jcli-d-16-0758.1

Gerken, T., Chamecki, M., & Fuentes, J. D. (2017). Air-parcel residence times within forest canopies. *Boundary-Layer Meteorology*, *165*(1), 29–54. https://doi.org/10.1007/s10546-017-0269-7

995 Geron, C., Guenther, A., Sharkey, T., & Arnts, R. R. (2000). Temporal variability in basal
996 isoprene emission factor. *Tree Physiology*, *20*(12), 799–805.
997 https://doi.org/10.1093/treephys/20.12.799

998 Gu, D., Guenther, A. B., Shilling, J. E., Yu, H., Huang, M., Zhao, C., Yang, Q., Martin, S. T.,
999 Artaxo, P., Kim, S., Seco, R., Stavrakou, T., Longo, K. M., Tóta, J., de Souza, R. A.,
1000 Vega, O., Liu, Y., Shrivastava, M., Alves, E. G., … Hu, Z. (2017). Airborne observations
1001 reveal elevational gradient in tropical forest isoprene emissions. *Nature Communications*,
1002 *8*(1). https://doi.org/10.1038/ncomms15541

1003 Guenther, A. B., & Hills, A. J. (1998). Eddy covariance measurement of isoprene fluxes. *Journal*
1004 *of Geophysical Research: Atmospheres*, *103*(D11), 13145–13152.
1005 https://doi.org/10.1029/97jd03283

1006 Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., &
1007 Wang, X. (2012). The model of emissions of gases and aerosols from nature version 2.1
1008 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions.
1009 *Geoscientific Model Development*, *5*(6), 1471–1492. https://doi.org/10.5194/gmd-5-1471-
1010 2012

1011 Guenther, A. B., Monson, R. K., & Fall, R. (1991). Isoprene and monoterpene emission rate
1012 variability: Observations with eucalyptus and emission rate algorithm development.
1013 *Journal of Geophysical Research*, *96*(D6), 10799. https://doi.org/10.1029/91jd00960

1014 Guenther, A. B., Zimmerman, P. R., Harley, P. C., Monson, R. K., & Fall, R. (1993). Isoprene
1015 and monoterpene emission rate variability: Model evaluations and sensitivity analyses.
1016 *Journal of Geophysical Research*, *98*(D7), 12609–12617.
1017 https://doi.org/10.1029/93jd00527

1018 Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., Harley, P., Klinger,
1019 L., Lerdau, M., Mckay, W. A., Pierce, T., Scholes, B., Steinbrecher, R., Tallamraju, R.,
1020 Taylor, J., & Zimmerman, P. (1995). A global model of natural volatile organic
1021 compound emissions. *Journal of Geophysical Research*, *100*(D5), 8873–8892.
1022 https://doi.org/10.1029/94jd02950

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., & Geron, C. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (model of emissions of gases and aerosols from nature). *Atmospheric Chemistry and Physics*, *6*(11), 3181–3210. https://doi.org/10.5194/acp-6-3181-2006

Gupta, P., Verma, S., Bhatla, R., Chandel, A. S., Singh, J., & Payra, S. (2020). Validation of surface temperature derived from Merra-2 reanalysis against IMD gridded data set over India. *Earth and Space Science*, *7*(1). https://doi.org/10.1029/2019ea000910

Haario, H., Saksman, E., & Tamminen, J. (2001). An adaptive Metropolis algorithm. Bernoulli, 7(2), 223. https://doi.org/10.2307/3318737

Hanson, D. T., & Sharkey, T. D. (2001). Rate of acclimation of the capacity for isoprene emission in response to light and temperature. *Plant, Cell & Environment*, *24*(9), 937–946. https://doi.org/10.1046/j.1365-3040.2001.00745.x

Heald, C., Wilkinson, M., Monson, R., Alo, C., Wang, G., & Guenther, A. (2009). Response of isoprene emission to ambient CO2 changes and implications for global budgets. *Global Change Biology*, *15*(5), 1127–1140. https://doi.org/10.1111/j.1365-2486.2008.01802.x

Hudman, R. C., Murray, L. T., Jacob, D. J., Millet, D. B., Turquety, S., Wu, S., Blake, D. R., Goldstein, A. H., Holloway, J., & Sachse, G. W. (2008). Biogenic versus anthropogenic sources of CO in the United States. *Geophysical Research Letters*, *35*(4). https://doi.org/10.1029/2007gl032393

Jeong, D., Seco, R., Emmons, L., Schwantes, R., Liu, Y., McKinney, K. A., Martin, S. T., Keutsch, F. N., Gu, D., Guenther, A. B., Vega, O., Tota, J., Souza, R. A., Springston, S. R., Watson, T. B., & Kim, S. (2022). Reconciling observed and predicted Tropical Rainforest oh concentrations. *Journal of Geophysical Research: Atmospheres*, *127*(1). https://doi.org/10.1029/2020jd032901

Jiang, X., Guenther, A., Potosnak, M., Geron, C., Seco, R., Karl, T., Kim, S., Gu, L., & Pallardy, S. (2018). Isoprene emission response to drought and the impact on global atmospheric chemistry. *Atmospheric Environment*, *183*, 69–83. https://doi.org/10.1016/j.atmosenv.2018.01.026

Jiang, Z., Worden, J. R., Worden, H., Deeter, M., Jones, D. B., Arellano, A. F., & Henze, D. K. (2017). A 15-year record of CO emissions constrained by MOPITT CO observations. *Atmospheric Chemistry and Physics*, *17*(7), 4565–4583. https://doi.org/10.5194/acp-17-4565-2017

Kaiser, J., Jacob, D. J., Zhu, L., Travis, K. R., Fisher, J. A., González Abad, G., Zhang, L., Zhang, X., Fried, A., Crounse, J. D., St. Clair, J. M., & Wisthaler, A. (2018). High-resolution inversion of OMI formaldehyde columns to quantify isoprene emission on ecosystem-relevant scales: Application to the southeast US. *Atmospheric Chemistry and Physics*, *18*(8), 5483–5497. https://doi.org/10.5194/acp-18-5483-2018

Karl, T., Apel, E., Hodzic, A., Riemer, D. D., Blake, D. R., & Wiedinmyer, C. (2009). Emissions of volatile organic compounds inferred from airborne flux measurements over a megacity. *Atmospheric Chemistry and Physics*, *9*(1), 271–285. https://doi.org/10.5194/acp-9-271-2009

Karl, T., Guenther, A., Yokelson, R. J., Greenberg, J., Potosnak, M., Blake, D. R., & Artaxo, P. (2007). The Tropical Forest and Fire Emissions Experiment: Emission, chemistry, and transport of biogenic volatile organic compounds in the lower atmosphere over Amazonia. *Journal of Geophysical Research*, *112*(D18). https://doi.org/10.1029/2007jd008539

Karl, T., Misztal, P. K., Jonsson, H. H., Shertz, S., Goldstein, A. H., & Guenther, A. B. (2013). Airborne flux measurements of BVOCs above Californian oak forests: Experimental investigation of surface and entrainment fluxes, oh densities, and Damköhler numbers. *Journal of the Atmospheric Sciences*, *70*(10), 3277–3287. https://doi.org/10.1175/jas-d-13-054.1

Kefauver, S. C., Filella, I., & Peñuelas, J. (2014). Remote sensing of atmospheric biogenic volatile organic compounds (BVOCs) via satellite-based formaldehyde vertical column assessments. *International Journal of Remote Sensing*, *35*(21), 7519–7542. https://doi.org/10.1080/01431161.2014.968690

Kramshøj, M., Vedel-Petersen, I., Schollert, M., Rinnan, Å., Nymand, J., Ro-Poulsen, H., & Rinnan, R. (2016). Large increases in arctic biogenic volatile emissions are a direct effect of warming. *Nature Geoscience*, *9*(5), 349–352. https://doi.org/10.1038/ngeo2692

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., & Slater, A. G. (2011). Parameterization improvements and functional and structural advances in version 4 of the Community Land Model. *Journal of Advances in Modeling Earth Systems*, *3*(3). https://doi.org/10.1029/2011ms000045

Li, Y., Liu, B., Ye, J., Jia, T., Khuzestani, R. B., Sun, J. Y., Cheng, X., Zheng, Y., Li, X., Wu, C., Xin, J., Wu, Z., Tomoto, M. A., McKinney, K. A., Martin, S. T., Li, Y. J., & Chen, Q. (2021). Unmanned aerial vehicle measurements of volatile organic compounds over a subtropical forest in China and implications for emission heterogeneity. *ACS Earth and Space Chemistry*, *5*(2), 247–256. https://doi.org/10.1021/acsearthspacechem.0c00271

Liu, Y., Brito, J., Dorris, M. R., Rivera-Rios, J. C., Seco, R., Bates, K. H., Artaxo, P., Duvoisin, S., Keutsch, F. N., Kim, S., Goldstein, A. H., Guenther, A. B., Manzi, A. O., Souza, R. A., Springston, S. R., Watson, T. B., McKinney, K. A., & Martin, S. T. (2016). Isoprene photochemistry over the amazon rainforest. *Proceedings of the National Academy of Sciences*, *113*(22), 6125–6130. https://doi.org/10.1073/pnas.1524136113

Liu, Y., Seco, R., Kim, S., Guenther, A. B., Goldstein, A. H., Keutsch, F. N., Springston, S. R., Watson, T. B., Artaxo, P., Souza, R. A., McKinney, K. A., & Martin, S. T. (2018). Isoprene photo-oxidation products quantify the effect of pollution on hydroxyl radicals over Amazonia. *Science Advances*, *4*(4). https://doi.org/10.1126/sciadv.aar2547

Marais, E. A., Jacob, D. J., Guenther, A., Chance, K., Kurosu, T. P., Murphy, J. G., Reeves, C. E., & Pye, H. O. (2014). Improved model of isoprene emissions in Africa using Ozone Monitoring Instrument (OMI) satellite observations of formaldehyde: Implications for oxidants and particulate matter. *Atmospheric Chemistry and Physics*, *14*(15), 7693–7703. https://doi.org/10.5194/acp-14-7693-2014

Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., Mills, G., Casadio, S., Millet, D. B., Barkley, M. P., Paulot, F., & Mao, J. (2012). Isoprene

emissions in Africa inferred from OMI observations of formaldehyde columns. *Atmospheric Chemistry and Physics*, *12*(14), 6219–6235. https://doi.org/10.5194/acp-12-6219-2012

Millet, D. B., Jacob, D. J., Turquety, S., Hudman, R. C., Wu, S., Fried, A., Walega, J., Heikes, B. G., Blake, D. R., Singh, H. B., Anderson, B. E., & Clarke, A. D. (2006). Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission. *Journal of Geophysical Research*, *111*(D24). https://doi.org/10.1029/2005jd006853

Misztal, P. K., Avise, J. C., Karl, T., Scott, K., Jonsson, H. H., Guenther, A. B., & Goldstein, A. H. (2016). Evaluation of regional isoprene emission factors and modeled fluxes in California. *Atmospheric Chemistry and Physics*, *16*(15), 9611–9628. https://doi.org/10.5194/acp-16-9611-2016

Misztal, P. K., Karl, T., Weber, R., Jonsson, H. H., Guenther, A. B., & Goldstein, A. H. (2014). Airborne flux measurements of biogenic isoprene over California. *Atmospheric Chemistry and Physics*, *14*(19), 10631–10647. https://doi.org/10.5194/acp-14-10631-2014

Miyazaki, K., Bowman, K. W., Yumimoto, K., Walker, T., and Sudo, K.: Evaluation of a multi-model, multi-constituent assimilation framework for tropospheric chemical reanalysis, Atmos. Chem. Phys., 20, 931–967, https://doi.org/10.5194/acp-20-931-2020, 2020.) -- Kazu Miyazaki

Monson, R. K., Harley, P. C., Litvak, M. E., Wildermuth, M., Guenther, A. B., Zimmerman, P. R., & Fall, R. (1994). Environmental and developmental controls over the seasonal pattern of isoprene emission from Aspen leaves. *Oecologia*, *99*(3-4), 260–270. https://doi.org/10.1007/bf00627738

Monson, R. K., Weraduwage, S. M., Rosenkranz, M., Schnitzler, J.-P., & Sharkey, T. D. (2021). Leaf isoprene emission as a trait that mediates the growth-defense tradeoff in the face of climate stress. *Oecologia*, *197*(4), 885–902. https://doi.org/10.1007/s00442-020-04813-7

Mougin, E., Diawara, M. O., Soumaguel, N., Maïga, A. A., Demarez, V., Hiernaux, P., Grippa, M., Chaffard, V., & Ba, A. (2019). A leaf area index data set acquired in Sahelian

Rangelands of gourma in Mali over the 2005–2017 Period. *Earth System Science Data*, *11*(2), 675–686. https://doi.org/10.5194/essd-11-675-2019

Müller, J.-F., & Stavrakou, T. (2005). Inversion of CO and NOx emissions using the adjoint of the IMAGES model. *Atmospheric Chemistry and Physics*, *5*(5), 1157–1186. https://doi.org/10.5194/acp-5-1157-2005

Müller, J.-F., Stavrakou, T., & Peeters, J. (2019). Chemistry and deposition in the Model of Atmospheric composition at Global and Regional scales using Inversion Techniques for Trace gas Emissions (MAGRITTE v1.1) – part 1: Chemical mechanism. *Geoscientific Model Development*, *12*(6), 2307–2356. https://doi.org/10.5194/gmd-12-2307-2019

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V., Underwood, E. C., D'amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., & Kassem, K. R. (2001). Terrestrial ecoregions of the world: A new map of life on Earth. *BioScience*, *51*(11), 933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:teotwa]2.0.co;2

Opacka, B., Müller, J.-F., Stavrakou, T., Bauwens, M., Sindelarova, K., Markova, J., & Guenther, A. B. (2021). Global and regional impacts of land cover changes on isoprene emissions derived from spaceborne data and the MEGAN model. *Atmospheric Chemistry and Physics*, *21*(11), 8413–8436. https://doi.org/10.5194/acp-21-8413-2021

Opacka, B., Müller, J.-F., Stavrakou, T., Miralles, D. G., Koppa, A., Pagán, B. R., Potosnak, M. J., Seco, R., De Smedt, I., & Guenther, A. B. (2022). Impact of drought on isoprene fluxes assessed using field data, satellite-based gleam soil moisture and HCHO observations from omi. *Remote Sensing*, *14*(9), 2021. https://doi.org/10.3390/rs14092021

Otu-Larbi, F., Bolas, C. G., Ferracci, V., Staniaszek, Z., Jones, R. L., Malhi, Y., Harris, N. R., Wild, O., & Ashworth, K. (2020). Modelling the effect of the 2018 summer heatwave and drought on isoprene emissions in a UK woodland. *Global Change Biology*, *26*(4), 2320–2335. https://doi.org/10.1111/gcb.14963

Palmer, P. I., Jacob, D. J., Fiore, A. M., Martin, R. V., Chance, K., & Kurosu, T. P. (2003). Mapping isoprene emissions over North America using formaldehyde column

observations from space. *Journal of Geophysical Research: Atmospheres*, *108*(D6). https://doi.org/10.1029/2002jd002153

Pétron, G., Harley, P., Greenberg, J., & Guenther, A. (2001). Seasonal temperature variations influence isoprene emission. *Geophysical Research Letters*, *28*(9), 1707–1710. https://doi.org/10.1029/2000gl011583

Potosnak, M. J., LeStourgeon, L., Pallardy, S. G., Hosman, K. P., Gu, L., Karl, T., Geron, C., & Guenther, A. B. (2014). Observed and modeled ecosystem isoprene fluxes from an oak-dominated temperate forest and the influence of drought stress. *Atmospheric Environment*, *84*, 314–322. https://doi.org/10.1016/j.atmosenv.2013.11.055

Pugh, T. A. M., Ashworth, K., Wild, O., & Hewitt, C. N. (2013). Effects of the spatial resolution of climate data on estimates of biogenic isoprene emissions. *Atmospheric Environment*, *70*, 1–6. https://doi.org/10.1016/j.atmosenv.2013.01.001

Sarkar, C., Guenther, A. B., Park, J.-H., Seco, R., Alves, E., Batalha, S., Santana, R., Kim, S., Smith, J., Tóta, J., & Vega, O. (2020). PTR-TOF-MS eddy covariance measurements of isoprene and monoterpene fluxes from an eastern Amazonian rainforest. *Atmospheric Chemistry and Physics*, *20*(12), 7179–7191. https://doi.org/10.5194/acp-20-7179-2020

Saunier, A., Ormeño, E., Piga, D., Armengaud, A., Boissard, C., Lathière, J., Szopa, S., Genard-Zielinski, A.-C., & Fernandez, C. (2020). Isoprene contribution to ozone production under climate change conditions in the French Mediterranean Area. *Regional Environmental Change*, *20*(4). https://doi.org/10.1007/s10113-020-01697-4

Seco R, Holst T, Davie-Martin CL, Simin T, Guenther A, Pirk N, Rinne J, Rinnan R (2022) Strong isoprene emission response to temperature in tundra vegetation. Proceedings of the National Academy of Sciences of the USA, 119: e2118014119. DOI: 10.1073/pnas.2118014119

Seco, R., Holst, T., Matzen, M. S., Westergaard-Nielsen, A., Li, T., Simin, T., Jansen, J., Crill, P., Friborg, T., Rinne, J., & Rinnan, R. (2020). Volatile organic compound fluxes in a subarctic peatland and lake. *Atmospheric Chemistry and Physics*., 20, 13399–13416, https://doi.org/10.5194/acp-20-13399-2020

Seco, R., Karl, T., Guenther, A., Hosman, K. P., Pallardy, S. G., Gu, L., Geron, C., Harley, P., & Kim, S. (2015). Ecosystem-scale volatile organic compound fluxes during an extreme drought in a broadleaf temperate forest of the Missouri Ozarks (central USA). *Global Change Biology*, *21*(10), 3657–3674. https://doi.org/10.1111/gcb.12980

Silva, S. J., Heald, C. L., & Guenther, A. B. (2020). Development of a reduced-complexity plant canopy physics surrogate model for use in chemical transport models: A case study with GEOS-Chem V12.3.0. *Geoscientific Model Development*, *13*(6), 2569–2585. https://doi.org/10.5194/gmd-13-2569-2020

Situ, S., Wang, X., Guenther, A., Zhang, Y., Wang, X., Huang, M., Fan, Q., & Xiong, Z. (2014). Uncertainties of isoprene emissions in the MEGAN model estimated for a coniferous and broad-leaved mixed forest in southern China. *Atmospheric Environment*, *98*, 105–110. https://doi.org/10.1016/j.atmosenv.2014.08.023

Sprengnether, M., Demerjian, K. L., Donahue, N. M., & Anderson, J. G. (2002). Product analysis of the OH oxidation of isoprene and 1,3-butadiene in the presence of NO. *Journal of Geophysical Research*, *107*(D15). https://doi.org/10.1029/2001jd000716

Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Lerot, C., Van Roozendael, M., Coheur, P.-F., Clerbaux, C., Boersma, K. F., van der A, R., & Song, Y. (2016). Substantial underestimation of post-harvest burning emissions in the North China Plain revealed by multi-species space observations. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep32307

Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Van Roozendael, M., De Mazière, M., Vigouroux, C., Hendrick, F., George, M., Clerbaux, C., Coheur, P.-F., & Guenther, A. (2015). How consistent are top-down hydrocarbon emissions based on formaldehyde observations from GOME-2 and OMI? *Atmospheric Chemistry and Physics*, *15*(20), 11861–11884. https://doi.org/10.5194/acp-15-11861-2015

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., Kanakidou, M., Vrekoussis, M., Wittrock, F., Richter, A., & Burrows, J. P. (2009b). The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling.

Atmospheric Chemistry and Physics, *9*(21), 8431–8446. https://doi.org/10.5194/acp-9-8431-2009

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Giglio, L., & Guenther, A. (2009c). Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns. *Atmospheric Chemistry and Physics*, *9*(3), 1037–1060. https://doi.org/10.5194/acp-9-1037-2009

Stavrakou, T., Müller, J.-F., De Smedt, I., Van Roozendael, M., van der Werf, G. R., Giglio, L., & Guenther, A. (2009a). Global emissions of non-methane hydrocarbons deduced from SCIAMACHY formaldehyde columns through 2003–2006. *Atmospheric Chemistry and Physics*, *9*(11), 3663–3679. https://doi.org/10.5194/acp-9-3663-2009

Trainer, M., Williams, E. J., Parrish, D. D., Buhr, M. P., Allwine, E. J., Westberg, H. H., Fehsenfeld, F. C., & Liu, S. C. (1987). Models and observations of the impact of natural hydrocarbons on rural ozone. *Nature*, *329*(6141), 705–707. https://doi.org/10.1038/329705a0

Turner, A. J., Henze, D. K., Martin, R. V., & Hakami, A. (2012). The spatial extent of source influences on modeled column concentrations of short-lived species. *Geophysical Research Letters*, *39*(12). https://doi.org/10.1029/2012gl051832

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., De Smedt, I., Grutter, M., Hannigan, J. W., Jones, N., Kivi, R., Loyola, D., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J.-M., Morino, I., Murata, I., Nagahama, T., … Winkler, H. (2020). Tropomi–sentinel-5 precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations. *Atmospheric Measurement Techniques*, *13*(7), 3751–3767. https://doi.org/10.5194/amt-13-3751-2020

Wells, K. C., Millet, D. B., Payne, V. H., Deventer, M. J., Bates, K. H., de Gouw, J. A., Graus, M., Warneke, C., Wisthaler, A., & Fuentes, J. D. (2020). Satellite isoprene retrievals constrain emissions and atmospheric oxidation. *Nature*, *585*(7824), 225–233. https://doi.org/10.1038/s41586-020-2664-3

Wells, K. C., Millet, D. B., Payne, V. H., Vigouroux, C., Aquino, C. A., De Mazière, M., de Gouw, J. A., Graus, M., Kurosu, T., Warneke, C., & Wisthaler, A. (2022). Next-generation

isoprene measurements from space: Detecting daily variability at high resolution. *Journal of Geophysical Research: Atmospheres*, *127*(5). https://doi.org/10.1029/2021jd036181

Wiedinmyer, C., Greenberg, J., Guenther, A., Hopkins, B., Baker, K., Geron, C., Palmer, P. I., Long, B. P., Turner, J. R., Pétron, G., Harley, P., Pierce, T. E., Lamb, B., Westberg, H., Baugh, W., Koerber, M., & Janssen, M. (2005). Ozarks Isoprene Experiment (OZIE): Measurements and modeling of the "Isoprene Volcano." *Journal of Geophysical Research*, *110*(D18). https://doi.org/10.1029/2005jd005800

Wolfe, G. M., Kaiser, J., Hanisco, T. F., Keutsch, F. N., de Gouw, J. A., Gilman, J. B., Graus, M., Hatch, C. D., Holloway, J., Horowitz, L. W., Lee, B. H., Lerner, B. M., Lopez-Hilifiker, F., Mao, J., Marvin, M. R., Peischl, J., Pollack, I. B., Roberts, J. M., Ryerson, T. B., … Warneke, C. (2016). Formaldehyde production from isoprene oxidation across NOx regimes. *Atmospheric Chemistry and Physics*, *16*(4), 2597–2610. https://doi.org/10.5194/acp-16-2597-2016

Worden, H. M., Bloom, A. A., Worden, J. R., Jiang, Z., Marais, E. A., Stavrakou, T., Gaubert, B., & Lacey, F. (2019). New constraints on biogenic emissions using satellite-based estimates of carbon monoxide fluxes. *Atmospheric Chemistry and Physics*, *19*(21), 13569–13579. https://doi.org/10.5194/acp-19-13569-2019

Worden, H. M., Deeter, M. N., Frankenberg, C., George, M., Nichitiu, F., Worden, J., Aben, I., Bowman, K. W., Clerbaux, C., Coheur, P. F., de Laat, A. T., Detweiler, R., Drummond, J. R., Edwards, D. P., Gille, J. C., Hurtmans, D., Luo, M., Martínez-Alonso, S., Massie, S., … Warner, J. X. (2013). Decadal record of Satellite Carbon Monoxide Observations. *Atmospheric Chemistry and Physics*, *13*(2), 837–850. https://doi.org/10.5194/acp-13-837-2013

Wu, J., Albert, L. P., Lopes, A. P., Restrepo-Coupe, N., Hayek, M., Wiedemann, K. T., Guan, K., Stark, S. C., Christoffersen, B., Prohaska, N., Tavares, J. V., Marostica, S., Kobayashi, H., Ferreira, M. L., Campos, K. S., da Silva, R., Brando, P. M., Dye, D. G., Huxman, T. E., … Saleska, S. R. (2016). Leaf development and demography explain photosynthetic seasonality in Amazon evergreen forests. *Science*, *351*(6276), 972–976. https://doi.org/10.1126/science.aad5068

Xu, T., White, L., Hui, D., & Luo, Y. (2006). Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction. *Global Biogeochemical Cycles*, *20*(2). https://doi.org/10.1029/2005gb002468

Yang, Y., Bloom, A. A., Ma, S. Levine, P., Norton, A., Parazoo, N. C., Reager, J. T., Worden, J., Quetin, G. R., Smallman, T. L., Williams, M., Xu, L., & Saatchi, S.. (2021). CARDAMOM-FluxVal Version 1.0 [Software]. https://doi.org/10.5281/zenodo.4904195

Yang, Y., Bloom, A. A., Ma, S., Levine, P., Norton, A., Parazoo, N. C., Reager, J. T., Worden, J., Quetin, G. R., Smallman, T. L., Williams, M., Xu, L., & Saatchi, S. (2022). CARDAMOM-FluxVal version 1.0: A FLUXNET-based validation system for Cardamom Carbon and water flux estimates. *Geoscientific Model Development*, *15*(4), 1789–1802. https://doi.org/10.5194/gmd-15-1789-2022

Yu, H., Guenther, A., Gu, D., Warneke, C., Geron, C., Goldstein, A., Graus, M., Karl, T., Kaser, L., Misztal, P., & Yuan, B. (2017). Airborne measurements of isoprene and monoterpene emissions from Southeastern U.S. forests. *Science of The Total Environment*, *595*, 149–158. https://doi.org/10.1016/j.scitotenv.2017.03.262

Yuan, H., Dai, Y., Li, S., 2020. Reprocessed MODIS Version 6 Leaf Area Index data sets for land surface and climate modelling. Sun Yat-sun University. http://globalchange.bnu.edu.cn/research/laiv6#cite

Yuan, H., Dai, Y., Xiao, Z., Ji, D., & Shangguan, W. (2011). Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. *Remote Sensing of Environment*, *115*(5), 1171–1187. https://doi.org/10.1016/j.rse.2011.01.001

Ziehn, T., Scholze, M., & Knorr, W. (2012). On the capability of Monte Carlo and adjoint inversion techniques to derive posterior parameter uncertainties in terrestrial ecosystem models. *Global Biogeochemical Cycles*, *26*(3). https://doi.org/10.1029/2011gb004185

JGR Atmospheres

Supporting Information for

Optimizing the isoprene emission model MEGAN with satellite and ground-based observational constraints

Christian A. DiMaria[1*], Dylan B. A. Jones[1], Helen Worden[2], A. Anthony Bloom[3], Kevin Bowman[3], Trissevgeni Stavrakou[4], Kazuyuki Miyazaki[3], John Worden[3], Alex Guenther[5], Chinmoy Sarkar[6†], Roger Seco[7], Jeong-Hoo Park[8], Julio Tota[9], Eliane Gomes Alvez[10], and Valerio Ferracci[11]

[1]Department of Physics, University of Toronto, Toronto, Ontario, Canada, [2]Atmospheric Chemistry Observations & Modeling Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [4]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium, [5]Department of Earth System Science, University of California, Irvine, CA, USA, [6]Air Quality Research Center, University of California, Davis, CA, USA, [7]Institute of Environmental Assessment and Water Research (IDAEA-CSIC), Barcelona, Catalonia, Spain, [8]Air Quality Reserach Division, National Institute of Environmetal Research, 22689, Republic of Korea, [9]Julio Tota, Instituto de engenharia e Geociências, Universidade Federal do Oeste do Pará, UFOPA, Santarém, Pará, Brazil, [10]Department of Biogeochemical Processes, Max Planck Institute for Biogeochemistry, Jena, Germany, [11]School of Water, Energy and Environment, Cranfield University, Cranfield, UK

Corresponding author: Christian DiMaria (christian.dimaria@mail.utoronto.ca)

† Now at Thermo Fisher Scientific, San Francisco, CA, USA

Contents of this file

Text S1. MEGAN normalization with canopy environment coefficient

The canopy environment coefficient C_{CE} is a normalization constant which ensures that $E_{ISOP} = E_0$ under standard conditions. These conditions are listed in Table S1. The normalization constant is defined as

$$C_{CE} = \frac{1}{(\gamma_T \times \gamma_{PAR} \times \gamma_{LAI} \times \gamma_{AGE} \times \gamma_{CO2} \times \gamma_{SM})_{Std}} \quad \text{(S1)}$$

where the subscript *Std* indicates that each $\boldsymbol{\gamma}$-factor is calculated at standard conditions. The definition of C_{CE} makes it sensitive to certain $\boldsymbol{\gamma}_T$ parameters, including T_{Max}, C_{T1}, and C_{T2}, while it is insensitive to other parameters including K_1 and K_2. The $\boldsymbol{\gamma}_T$ parameter C_{EO} has no impact on MEGAN isoprene emissions because it appears as a multiplicative factor in $\boldsymbol{\gamma}_T$ and in $\boldsymbol{\gamma}_{T,Std}$ (i.e., in both the numerator and denominator of the MEGAN equation). Because C_{CE} is sensitive to the values of the $\boldsymbol{\gamma}_T$ parameters, we multiply our posterior $\boldsymbol{\gamma}_T$ by the updated normalization constant when calculating posterior MEGAN

emissions. This is necessary to ensure $E_{ISOP} = E_0$ under standard conditions when using the posterior parameterization. All plots of γ_T in the main text are thus multiplied by their corresponding C_{CE} values (i.e., either the a priori or posterior C_{CE}) to ensure they are all normalized to the same standard conditions.

Text S2. Long term temperature response in MEGAN γ_T

The MEGAN temperature response γ_T exhibits hysteresis with respect to past ambient temperatures. Emissions increase up with temperature up to an optimal value E_{Opt}, given by

$$E_{Opt} = \left[C_{EO} \times \exp\left(K_2\left(T_{daily} - 297\right)\right)\right], \quad \text{(S2)}$$

which occurs at temperature T_{Opt}, given by

$$T_{Opt} = T_{Max} + \left(K_1\left(T_{daily} - 297\right)\right) \quad \text{(S3)}$$

where C_{EO} is a scaling factor, T_{Daily} is the ambient temperature (in K) of the past 24 hours, T_{Max} (313 K) is the standard optimal temperature, and K_1 (0.6) and K_2 (0.08) are coefficients which determine the sensitivity of E_{Opt} and T_{Opt} to variations in T_{Daily} (Guenther et al., 2006). Values of T_{Daily} greater than 297 K will increase the optimum value of γ_T (E_{Opt}) as well as cause this optimum to occur at a higher temperature (T_{Opt}). The opposite occurs when T_{Daily} falls below 297 K. The magnitude of these changes in E_{Opt} and T_{Opt} is set by K_1 and K_2, with larger values of these parameters giving a higher sensitivity to T_{Daily} and vice versa. These impacts are illustrated in Figure S1.

This algorithm is based on a limited number of experiments (Geron et al., 2000; Hanson & Sharkey, 2001; Monson et al., 1994; Pétron et al., 2001), leading to relatively large uncertainties in the empirical parameters C_{EO}, T_{Max}, K_1, and K_2. We therefore treat these as free parameters in our optimization experiments. However, the parameter C_{EO} has no impact on modelled emissions when γ_T is used within the full MEGAN equation (Equation (1) in the main text) due to the normalization of emissions to E_0 under standard environmental conditions (see Supplementary Text S1).

Text S3. Sensitivity of E_0 optimization to observation errors

In our top-down optimization of E_0 we assumed a constant relative error in the observations of 30%. This is consistent with early OMI-based top-down emission inventories (Millet et al., 2006). It is possible that we are overestimating the uncertainty in some regions, but we are likely underestimating it in others, due to spatially and temporally varying sources of error in the satellite-derived products (including CH_2O retrieval biases (Vigouroux et al., 2020; De Smedt et al., 2021), chemistry-transport model errors (Barkley et al., 2013; Stavrakou et al., 2015), spatial smearing errors (Palmer et al., 2003; Turner et al., 2012), and background CH_2O sources (Marais et al., 2012; Wolfe et al., 2016)). We therefore had to account for the potential impact of our uncertainty assumption on the optimization of E_0.

The optimization of model parameters is dependent on the accuracy and precision of the observational constraints. While we assume observations are unbiased during our experiments, we account for the precision (i.e., measurement noise and other random errors) of the observations explicitly in our MHMCMC scheme through the model likelihood function

$$P(\boldsymbol{x}_i) = \exp\left(-0.5\sum\left(\frac{(\boldsymbol{M}(\boldsymbol{x}_i) - \boldsymbol{O})^2}{\sigma_O^2}\right)\right), \quad \text{(S4)}$$

where $\boldsymbol{M}(\boldsymbol{x}_i)$ is the model output vector (in our case MEGAN isoprene emissions), $\boldsymbol{O}$ is the observation vector, and $\boldsymbol{\sigma}_O$ is the observation error. The presence of $\boldsymbol{\sigma}_O$ in the denominator of Equation (S4) ensures that more precise observations will give a more strongly peaked probability distribution and therefore stronger constraints on the model parameters.

We performed a series of simulated observation experiments to test the sensitivity of the E_0 optimization to the assumed observation error $\boldsymbol{\sigma}_O$. MEGAN was used to generate a "true" isoprene emission time series which was then sampled at user-specified intervals (hourly, daily, and monthly) and perturbed with Gaussian noise ($\boldsymbol{\sigma}_O$ = 1 % – 50%) to produce pseudo-observations. The pseudo-observations were then used to constrain E_0 in the MHMCMC scheme. Figure S2 shows a sample of how the posterior E_0 distribution changes as a function of observation frequency and error. Figure S2 (a) and (b) show the posterior E_0 as constrained with monthly observations with 5% and 25% error, respectively, while Figure S2 (c) and (d) show the results for hourly observations with the same errors as (a) and (b). The observation frequency had no significant impact on the optimization of E_0, whereas larger observation errors resulted in a broader posterior E_0 distribution. Despite the weaker constraints on E_0 with larger observation errors, the median posterior value is still well-within one standard deviation of the true E_0. The optimization was consistently able to recover E_0 up to an observation error of nearly 50%, at which point signal-to-noise was insufficient to constrain the emissions.

These simulated observation experiments suggest that top-down emission estimates are sufficiently precise to constrain E_0, and that our assumption of a 30% error has little impact on the posterior E_0 values other than potentially over or underestimating the standard deviation of the posterior E_0 distribution. This is a relatively minor issue when compared to other sources of error in the E_0 optimization, including the impact of drought stress, model input uncertainties, and discrepancies between different observation datasets (see main text Section 4).

Text S4. Mapping locally measured PAR to GEOS-FP PAR

Local PAR measurements are available at Wytham Woods in September 2018 but they are not available during the pre-heatwave period (25 May – 21 June). We consequently used hourly GEOS-FP PAR at 0.25° × 0.325° spatial resolution to drive MEGAN in our optimization experiments. To avoid introducing biases due to discrepancies between the local PAR measurements and the GEOS-FP PAR, we generated a linear mapping between the two datasets for September 2018 following the methodology of Ferracci et al (2020). This mapping is shown in Figure S3, in units of photosynthetic photon flux density (PPFD). There is a strong correlation between the local WIsDOM measurements and the GEOS-FP values (r^2 = 0.83). However, there is also a ~30% offset between the two datasets as indicated by the slope of the linear map in Figure S3, with WIsDOM PPFD measurements being consistently lower than the concurrent GEOS-FP values. We applied the linear map in Figure S3 to the pre-heatwave (25

May– 21 June) GEOS-FP PPFD data to account for this offset, effectively transforming the GEOS-FP PPFD into the local measurement space.

The linear map in Figure S3 is a source of error in our optimization, partially due to the large amount of scatter in the measurements and partially due to our assumption that a map based on September 2018 observations is applicable during the pre-heatwave period (25 May – 21 June). The impact of this source of error is limited by filtering the WIsDOM measurements for PPFD > 650 µmol/m^2/s (equivalent to PAR > 136 W/m^2). This effectively limits our observations to mid-day scenes (roughly 10am – 4pm local time, depending on cloud cover) and greatly reduces the sunlight-driven variability of the isoprene measurements. This is described in more detail in Section 3.2 of the main text, as well as in Ferracci et al (2020).

Text S5. Sensitivity of $\boldsymbol{\gamma}_T$ optimization to observation errors

We performed a series of simulated observation experiments to test the sensitivity of the $\boldsymbol{\gamma}_T$ optimization to the assumed observation error $\boldsymbol{\sigma}_O$. In these simulated observation experiments, MEGAN was used to generate a "true" isoprene emission time series which was then sampled at user-specified intervals (hourly, daily, and monthly) and perturbed with Gaussian noise ($\boldsymbol{\sigma}_O$ = 1 % – 50%) to produce pseudo-observations. The pseudo-observations were then used to constrain the $\boldsymbol{\gamma}_T$ parameters in the MHMCMC scheme. All 31 unique combinations of the $\boldsymbol{\gamma}_T$ parameters T_{Max}, K_1, K_2, C_{T1}, and C_{T2} were tested. For brevity, we present here only the results for the optimization of T_{Max} with all other parameters fixed to their a priori values.

Figure S4 shows how the posterior T_{Max} distribution changes as a function of observation error and observation frequency. The MHCMC can constrain T_{Max} given hourly observations with 1% relative error, resulting in a very narrow probability distribution (Figure S4 (a)). Increasing the observation error to 5% results in a much broader probability distribution and an incorrect median posterior value (Figure S4 (b)), though the mode of the distribution still coincides with the true parameter value. With an observation error of 25% the MHMCMC is unable to constrain T_{Max} even with hourly observations (Figure S4 (c)). Unlike with the E_0 optimization (Figure S2), the optimization of the $\boldsymbol{\gamma}_T$ parameters was sensitive to observation frequency. Figure S4 (d) shows that monthly observations with 1% error can constrain T_{Max}, but the probability distribution is broader than with hourly observations (Figure S4 (a)). T_{Max} could not be constrained with monthly observations with 5% error or greater (Figure S4 (e) and (f)).

None of the 31 combinations of $\boldsymbol{\gamma}_T$ parameters could be reliably constrained with observation errors exceeding 20% in our simulated observation experiments. This effectively ruled out the use of top-down isoprene emission estimates as constraints on $\boldsymbol{\gamma}_T$. Because of this, we only attempted to optimize $\boldsymbol{\gamma}_T$ using the BR-Sa1 eddy covariance flux measurements (observation error of ~15%; Sarkar et al., 2020) and WIsDOM isoprene mixing ratio measurements (observation error of ~10%; Ferracci et al., 2020).

Text S6. Sensitivity of MEGAN to E_0 and $\boldsymbol{\gamma}_T$ parameters

MEGAN has a different sensitivity to each of its empirical parameters, defined here as the change in model output for a given perturbation to an empirical parameter. This has important

implications for optimizing MEGAN parameters with observations using MHMCMC, because higher sensitivity parameters can be more readily constrained due to their larger impact on model output (i.e., perturbations to these parameters by the MHMCMC algorithm are more likely to produce a signal in the model output which rises above the noise in the observational constraints). Conversely, lower sensitivity parameters can only be constrained with more precise observations.

The sensitivity of MEGAN to E_0 and each of the $\boldsymbol{\gamma}_T$ parameters (T_{Max}, K_1, K_2, C_{T1}, and C_{T2}) is illustrated in Figure S5 (note that the $\boldsymbol{\gamma}_T$ parameter C_{EO} was omitted due to the normalization of MEGAN described in Supplementary Text S1). Each panel in Figure S5 shows the mean isoprene emission rate at BR-Sa1 from 1 – 16 June 2014 as calculated with MEGAN 2.1. The red circle in each panel represents the unperturbed emissions, using the a priori (PCEEA) value parameter values and the measured mean temperature at BR-Sa1 (this measured temperature is indicated by the vertical dotted black line). Deviations from this point on the y-axis represent perturbations to the MEGAN parameters, while deviations along the x-axis represent perturbations to the ambient daily temperature. The solid black contours represent the measured mean isoprene emission rate at BR-Sa1 ± 5 %.

By following the dotted black line from the red circle to the solid black contours in each panel, we can see that MEGAN could be brought into agreement with the BR-Sa1 observations (at least in terms of the mean emission rate for the whole time series) by: (a) increasing E_0 from 1 to 1.2 (in units of 10^9 kgC/m^2/s); (b) reducing T_{Max} from 313 K to 303 K; (c) reducing K_1 from 0.6 – 0.3; (d) increasing K_2 from 0.08 – 0.11; (e) reducing C_{T1} from 80 – 40 kJ/mol; or (f) reducing C_{T2} from 200 – 100 kJ/mol. The gradient along each path is a measure of the sensitivity of MEGAN to each parameter, with steeper gradients (e.g., E_0, K_2) indicating a higher sensitivity and shallower gradients (e.g., T_{Max}, K_1, C_{T1}, C_{T2}) indicating a lower sensitivity. This is illustrated in Figure S6, where we have used MEGAN 2.1 to calculate weekly mean isoprene emissions for 2014 in the Western Sahel regions with large perturbations to each of the MEGAN parameters. Large perturbations in E_0 (called "AEF" = "annual emission factor" in Figure S6) and K_2 lead to significant changes to the modelled emissions, while perturbations to K_1, T_{Max}, C_{T1}, and C_{T2} have a more minor impact. These sensitivities vary in time due to changes in ambient temperature; this is the same effect as is visible in Figure S5, where the gradient along the y-axis (i.e., parameter perturbation axis) changes as a function of position along the x-axis (i.e., the temperature axis).

These sensitivities are a proxy for our ability to constrain each parameter with observations, indicating that most of the $\boldsymbol{\gamma}_T$ parameters (except for K_2) are more difficult to constrain than E_0. However, it must be noted that the joint sensitivities which emerge when optimizing multiple MEGAN parameters simultaneously are not necessarily the same as those shown in Figure S5 or S6 where parameters were perturbed in isolation. It was therefore necessary to test all possible combinations of parameters in our optimization experiments to determine which subsets could be reasonably well constrained (see Supplementary Text S5).

Text S7. Optimization of T_{Max} and K_1 at BR-Sa1 using eddy covariance fluxes

There are 31 unique combinations of the 5 $\boldsymbol{\gamma}_T$ parameters T_{Max}, K_1, K_2, C_{T1}, and C_{T2}. Due to the differing sensitivity of MEGAN to each of these parameters, particularly when multiple

parameters are perturbed simultaneously, all 31 combinations were tested in our MHMCMC optimization to identify the best possible subset. Here "best" refers to the largest subset in which each parameter is reliably constrained while also significantly improving model-observation agreement. The γ_T parameter subset K_2, C_{T1}, and C_{T2} was identified as the best subset at BR-Sa1 using eddy covariance flux constraints (see Main Text Figure 7 and Figure 8). Many subsets contained poorly constrained parameters, or alternatively they contained well-constrained parameters but failed to significantly improve model-observation agreement.

Figure S7 shows the optimization results for the subset of T_{Max} and K_1 (the two empirical parameters in Equation S2). The posterior γ_T is extremely similar to the a priori over the range of observed temperatures and is not in significantly better agreement with the observations. Furthermore, the interquartile range of the posterior γ_T shows that T_{Max} and K_1 have a negligible impact on γ_T at lower temperatures but a very large impact at high temperatures. This is expected from Equation S2, as both parameters simply modulate the location of the high-temperature emission peak. The histograms in Figure S7 show that neither parameter is well-constrained by the optimization; the MHMCMC can effectively rule out a reduction in T_{Max} but is unable to distinguish between higher values of T_{Max} due to the low and asymmetric sensitivity of MEGAN to this parameter (see Figure S5 (b)). The K_1 parameter is dragged towards zero by the optimization, which reduces the model's sensitivity to T_{Daily} and keeps the emission peak closer to T_{Max}. This has the net effect of slightly increasing modelled emissions by moving the emission peak T_{Opt} closer to ambient temperatures, thereby slightly reducing model biases relative to the observations.

While the median posterior parameter values indicated in Figure S7 do reduce model biases (Figure S8 (a)), they do not improve the temporal variability of the estimate (i.e., the correlation between the model and the observations as indicated by the r^2 values in Figure S8 (b)). This contrasts with the subset of K_2, C_{T1}, and C_{T2} shown in the main text (Figure 8).

Text S8. Variability of MEGAN sensitivity T and LAI input errors

MEGAN isoprene emissions are highly sensitive to temperature due to the exponential form of γ_T over typical terrestrial temperature ranges (see Figure 11 in the main text). As a result, small errors in the temperature data used to drive MEGAN can lead to large errors in the modelled emissions. This can be seen in Figure S9 (a), which shows the relative change in seasonally averaged (December-January-February, 2016–2017) MEGAN isoprene emissions when positively biased temperature data are used to drive the model. A 1% temperature bias leads to emission biases in excess of 40% in most regions, consistent with our findings at BR-Sa1 when comparing the MERRA2 and locally measured temperature data (see Figure 5 (c) and (d) in the main text). From the viewpoint of model parameter optimization, it is clear that accurate temperature data must be used to drive MEGAN to obtain reliable emissions estimates and posterior parameters in all regions.

The sensitivity of MEGAN to LAI is much more variable, being high in low-LAI environments like the Sahel and negligible in high-LAI environments like the Amazon (see Figure 11 in the main text). This spatial variability is a direct consequence of the functional form of γ_{LAI} (Guenther et al., 2006),

$$\gamma_{LAI} = \frac{0.49LAI}{\sqrt{1 + 0.2LAI^2}} \quad \text{(S5)}$$

and is illustrated in Figure S9 (b),which shows the relative change in seasonally averaged (Dec-Jan-Feb) MEGAN isoprene emissions when positively biased LAI data are used to drive the model. The season Dec-Jan-Feb was chosen because this coincides with the minimum of the MODIS 8-Day LAI product in the Sahel of North Africa, allowing the spatial variability of LAI sensitivity in North African to be more easily seen. A 5% LAI bias leads to a 5% emission bias in the savannas of Northern and Southern Africa as well as Western Australia but has barely any impact (<1%) on emissions in tropical rainforest environments such as the Amazon or Congo River basins. Note that the large relative impact at high Northern latitudes in Figure S9 (b) is small in absolute terms, as total isoprene emissions from these regions are very low during boreal winter. The low sensitivity of MEGAN to LAI in forested regions implies that errors in LAI inputs are a likely only a very minor source of error (a few percent at most) in our model optimization in the Eastern Amazon, the Southeast US, and the UK temperate forest at Wytham Woods. On the other hand, LAI input errors may be on the order of 10% in the Western Sahel and Eastern Australia given the estimated uncertainty in the MODIS 8-day LAI product (Fang et al., 2013) and the higher sensitivity of MEGAN to LAI in low-LAI environments.

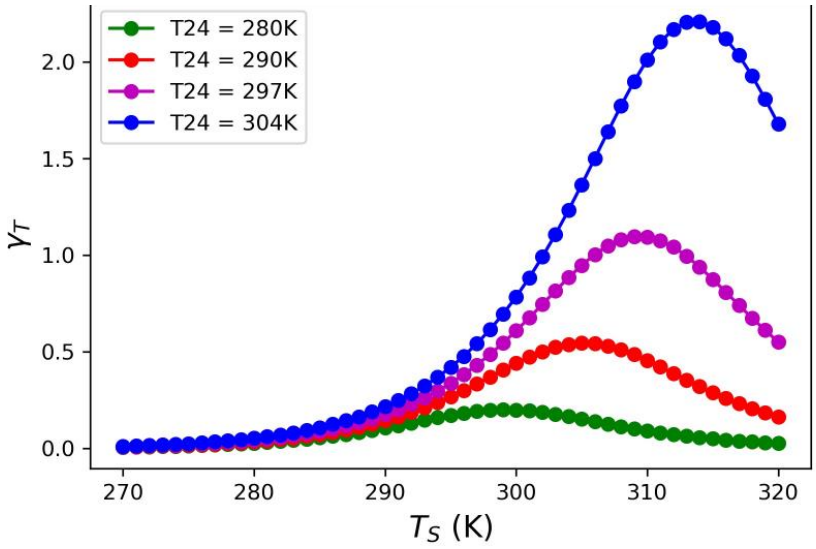

Figure S1. MEGAN temperature response γ_T plotted as a function of temperature for 4 different 24-hour average temperatures (T24, equivalent to T_{Daily} in Equation (S1) and (S2)). The standard value of T24 is 297 K, such that γ_T = 1 at 303 K (standard instantaneous temperature) and the emission peak is located at T_{Max} = 313 K. Changing the past ambient temperature changes both the location and height of the γ_T peak. The magnitude of these changes is controlled by K_1 and K_2 in Equation (S1) and (S2).

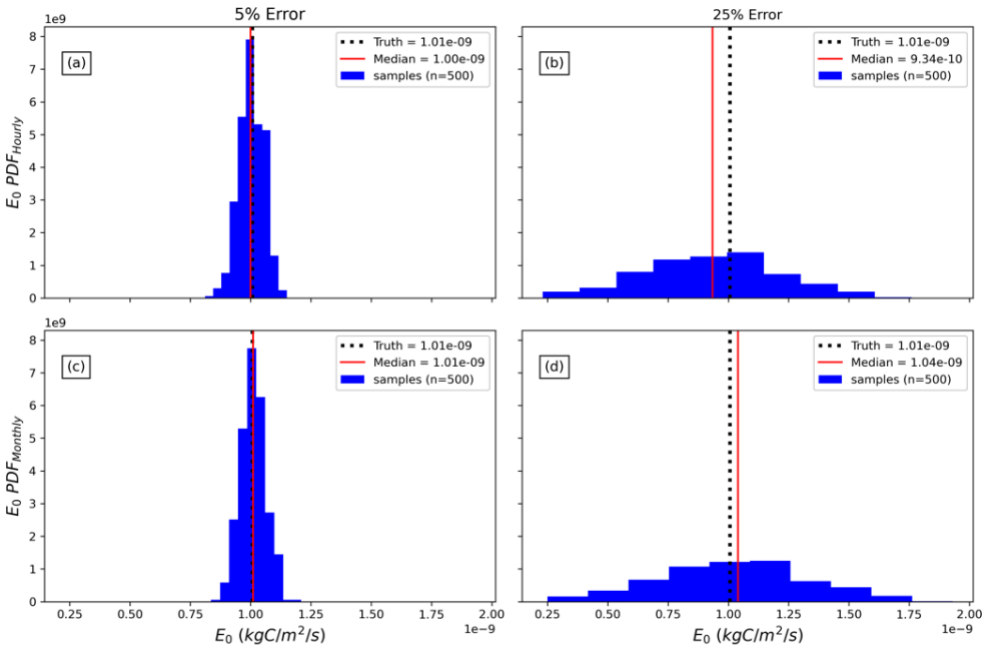

Figure S2. Posterior E_0 distribution for a selection of simulated observation experiments using 2014 MEGAN isoprene emissions in the Eastern Amazon region as the truth. The observation characteristics were (a) hourly with 5% error, (b) hourly with 25% error, (c) monthly with 5% error, and (d) monthly with 25% error. The true E_0 value is represented by the solid black line and the median posterior E_0 for each experiment is the solid red line. The blue histograms show a subsample of 500 E_0 values from the MHMCMC output (out of 5000 total accepted samples), which represents the posterior E_0 probability distribution.

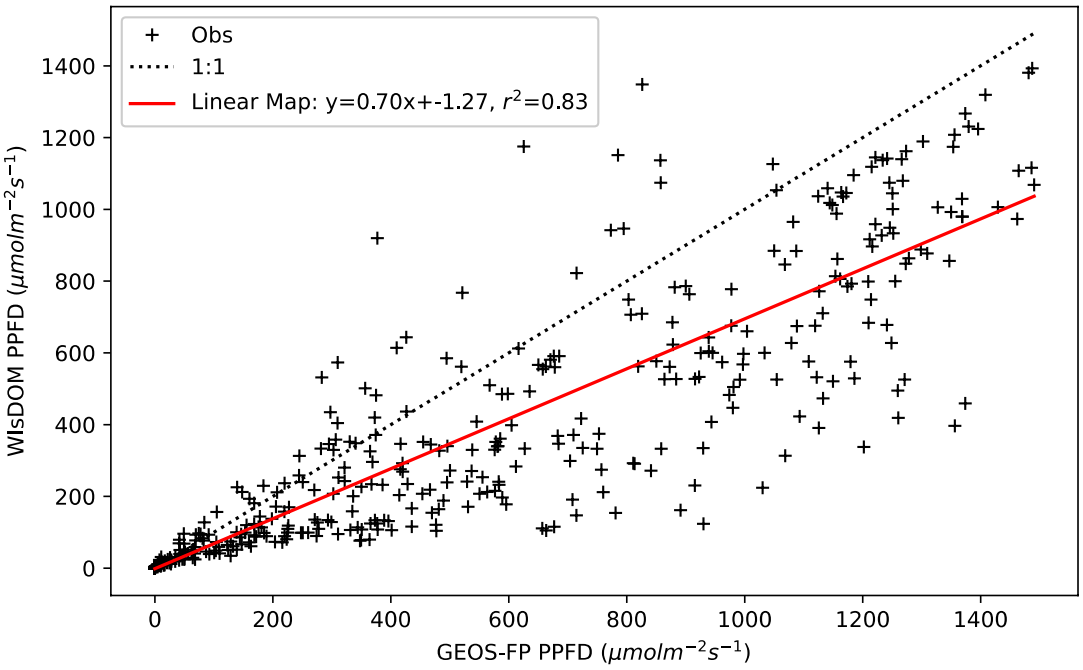

Figure S3. Scatter plot of local PAR measurements with GEOS-FP PAR for September 2018. The linear mapping (red line) was used to account for the systematic bias between the GEOS-FP PAR and the locally measured PAR in September 2018, and then applied to the GEOS-FP PAR during the pre-heatwave period.

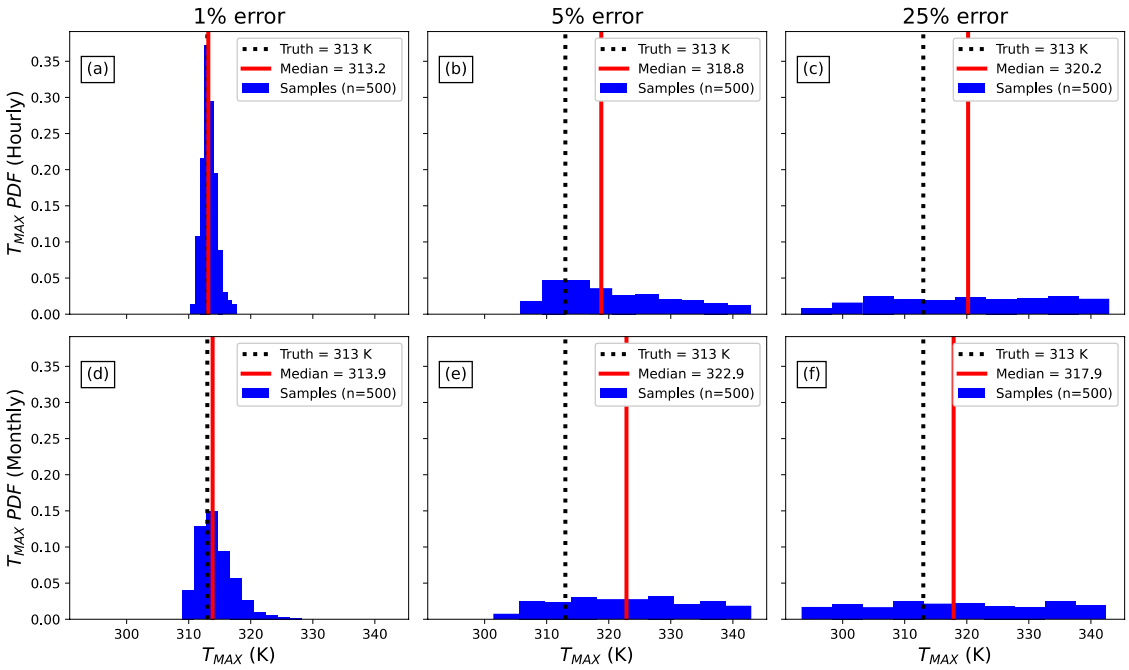

Figure S4. Posterior T_{Max} distribution for a selection of simulated observation experiments using 2014 MEGAN isoprene emissions in the Eastern Amazon region as the truth. The observation characteristics were (a) hourly with 1% error, (b) hourly with 5% error, (c) hourly with 25% error, (d) monthly with 1% error, (e) monthly with 5% error, and (f) monthly with 25% error. The true T_{Max} value is represented by the solid black line and the median posterior T_{Max} for each experiment is the solid red line. The blue histograms show a subsample of 500 T_{Max} values from the MHMCMC output (out of 5000 total accepted samples), which represents the posterior T_{Max} probability distribution.

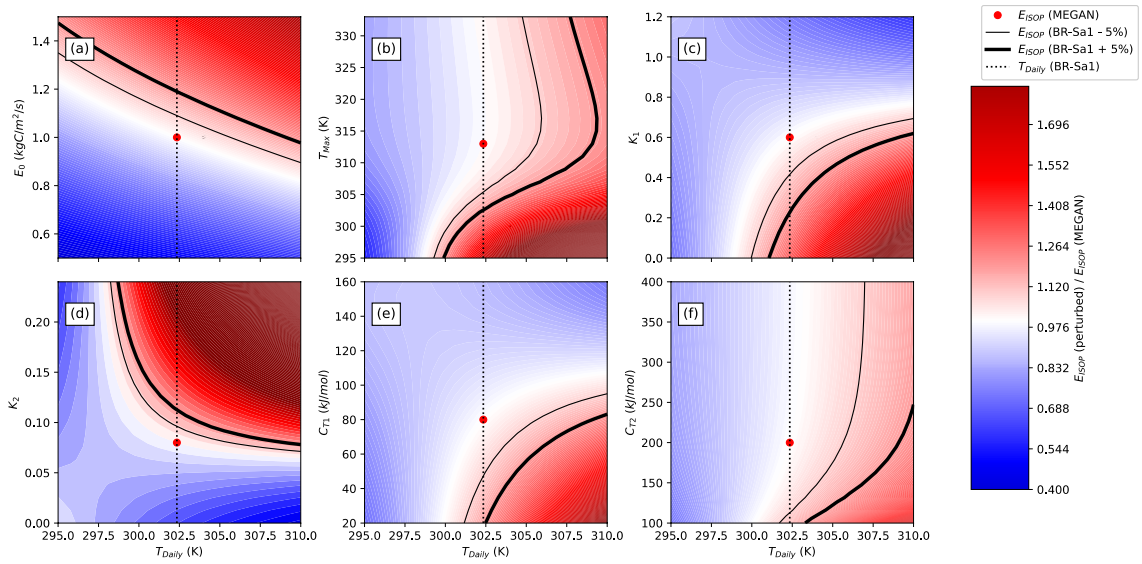

Figure S5. Sensitivity of mean MEGAN isoprene emissions at BR-Sa1 from 1 - 16 June 2014 to (a) E_0, (b) T_{Max}, (c) K_1, (d) K_2, (e) C_{T1}, and (f) C_{T2}. The red circle indicates the unperturbed MEGAN emissions, and the dashed vertical black line indicates the measured daily average temperature T_{Daily}. The y-axis of each panel shows parameter values, while the x-axis shows perturbations to the daily average temperature to illustrate the temperature dependence of the parameter sensitivity. The sensitivity is indicated by the gradient of each contour plot. The solid black lines indicate the average measured isoprene emission rate at BR-Sa1 ± 5 %.

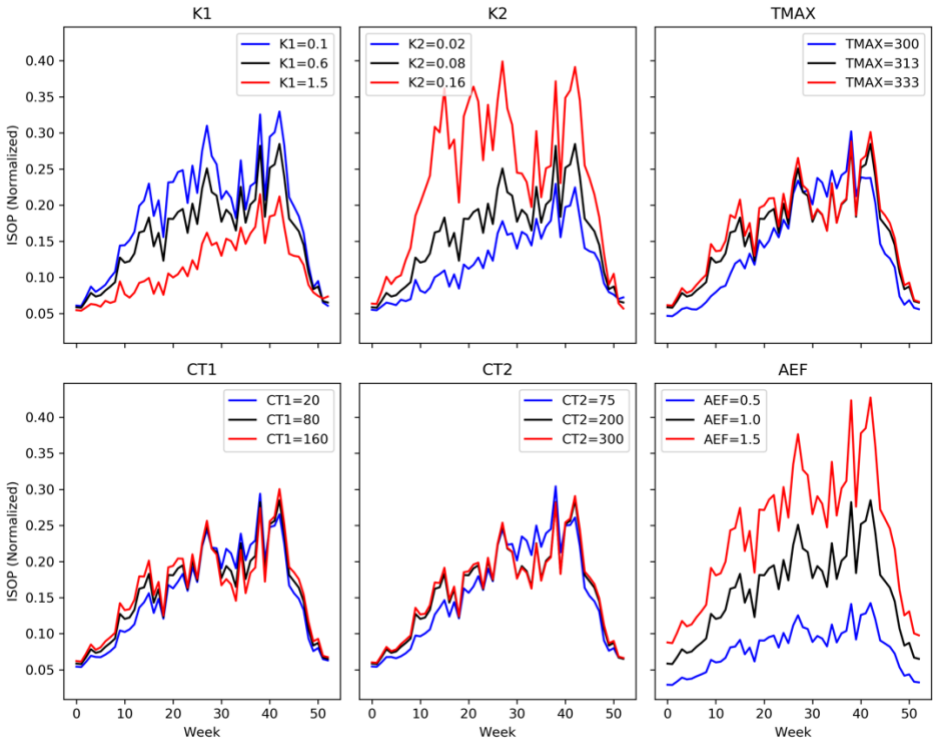

Figure S6. Weekly average isoprene emission rates (2014) in the Western Sahel region as calculated with MEGAN 2.1 using default (PCEEA) parameter values, normalized by AEF (note that AEF is equivalent to E_0). Positive (red) and negative (blue) perturbations to each parameter have an impact on the modelled emissions, with higher sensitivity parameters (AEF, K_2) impacting the emissions more than lower sensitivity parameters (K_1, T_{Max}, C_{T1}, C_{T2}).

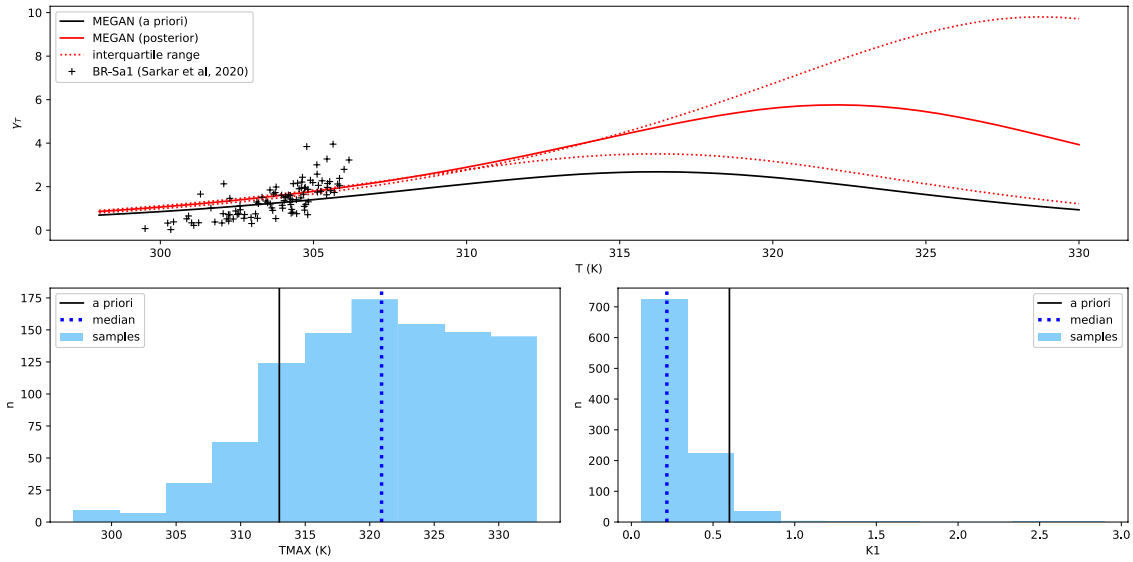

Figure S7. (top) A priori MEGAN $\boldsymbol{\gamma}_T$ (black), median posterior $\boldsymbol{\gamma}_T$ (solid red) and interquartile range (dotted red) as a function of temperature compared with the observed $\boldsymbol{\gamma}'_T$ (+ symbols) at BR-Sa1. Posterior parameter distributions for T_{Max} and K_1 are shown in the lower left and lower right panels, respectively (light blue). The median posterior values are indicated by the dashed blue lines in the lower panels, while the a priori values are indicated by the solid black lines.

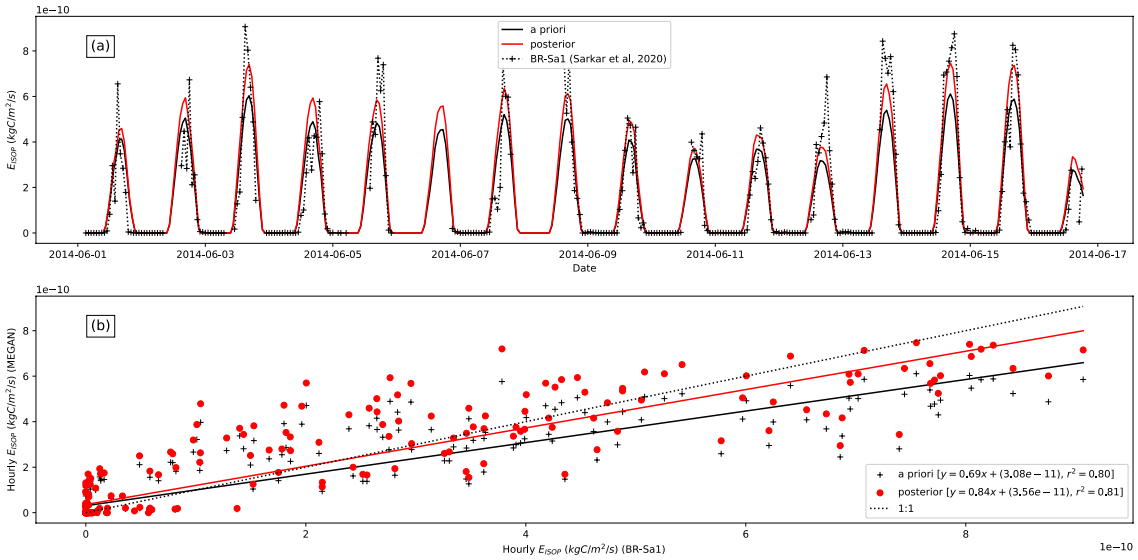

Figure S8. (a) A priori and posterior MEGAN isoprene flux estimates at BR-Sa1 from 1 – 16 June 2014, using the parameter subset of T_{Max} and K_1. The posterior emissions were calculated using the optimized $\boldsymbol{\gamma}_T$ based on eddy covariance observations (+ symbols). The dotted red line denotes the interquartile range on the posterior emission estimate. (b) Correlation between observed and modelled hourly mean isoprene emission rates at BR-Sa1. The solid red line is a

linear fit to the posterior modelled emissions (red circles), whereas the solid black curve is a linear fit to the a priori emissions (+ symbols).

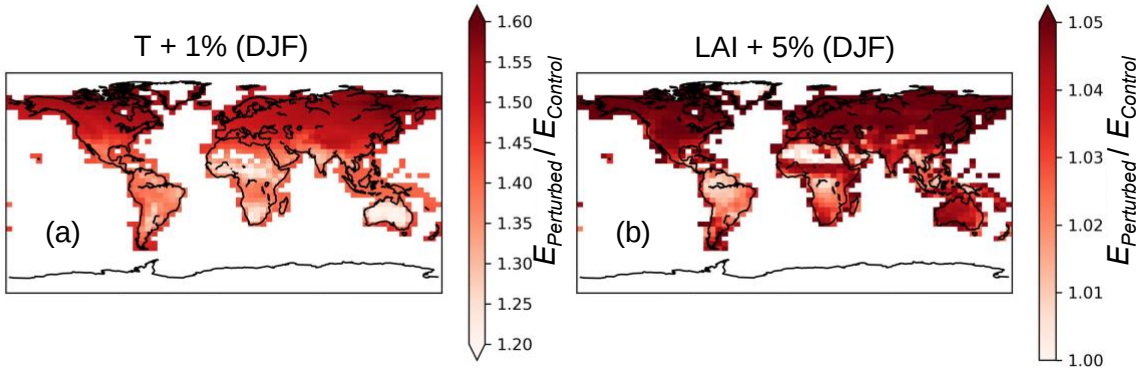

Figure S9. (a) Ratio of the biased ("Perturbed") and unbiased ("Control") seasonally averaged MEGAN isoprene emissions in December-January-February 2016–2017 when running the model with biased (+1%) temperature input data. (b) As in (a) but using biased (+5%) LAI input data. Note the scale of the colour bar in (b) is greatly reduced compared to (a), reflected the lower sensitivity of MEGAN to LAI.

Variable	Description	Standard Value
LAI	Leaf area index	5 cm^2/cm^2
F_{MAT}	Fraction of mature leaves	80%
F_{GRO}	Fraction of growing leaves	10%
F_{OLD}	Fraction of old leaves	10%
φ	Solar elevation angle	60°
τ_{PPFD}	PPFD transmission from top of atmosphere to canopy	0.6
T	Air temperature	303 K
T24	Average air temperature of past 24 hours	297 K
$PPFD_{SUN}$	Photosynthetic photon flux density (Sun leaves)	200 $\mu mol/m^2/s$
$PPFD_{SHADE}$	Photosynthetic photon flux density (shade leaves)	50 $\mu mol/m^2/s$

Table S1. Standard conditions for MEGAN γ-factors as defined by Guenther et al (2006). Under these conditions, isoprene emissions are equal to the standard emission rate ($E_{ISOP} = E_0$). We list here only the standard conditions which appear in the PCEEA implementation of MEGAN. Other standard driving variables, including wind speed and humidity, are only required when using the full MEGAN canopy environment model.

References From the Supporting Information

Barkley, M. P., Smedt, I. D., Van Roozendael, M., Kurosu, T. P., Chance, K., Arneth, A., Hagberg, D., Guenther, A., Paulot, F., Marais, E., & Mao, J. (2013). Top-down isoprene emissions over tropical South America inferred from SCIAMACHY and OMI formaldehyde columns. *Journal of Geophysical Research: Atmospheres*, *118*(12), 6849–6868. https://doi.org/10.1002/jgrd.50552

De Smedt, I., Pinardi, G., Vigouroux, C., Compernolle, S., Bais, A., Benavent, N., Boersma, F., Chan, K.-L., Donner, S., Eichmann, K.-U., Hedelt, P., Hendrick, F., Irie, H., Kumar, V., Lambert, J.-C., Langerock, B., Lerot, C., Liu, C., Loyola, D., Piters, A., Richter, A., Rivera Cárdenas, C., Romahn, F., Ryan, R. G., Sinha, V., Theys, N., Vlietinck, J., Wagner, T., Wang, T., Yu, H., and Van Roozendael, M.: Comparative assessment of TROPOMI and OMI formaldehyde observations and validation against MAX-DOAS network column measurements, Atmos. Chem. Phys., 21, 12561–12593, https://doi.org/10.5194/acp-21-12561-2021, 2021.

Fang, H., Jiang, C., Li, W., Wei, S., Baret, F., Chen, J. M., Garcia-Haro, J., Liang, S., Liu, R., Myneni, R. B., Pinty, B., Xiao, Z., & Zhu, Z. (2013). Characterization and Intercomparison of Global Moderate Resolution Leaf Area Index (LAI) products: Analysis of climatologies and theoretical uncertainties. *Journal of Geophysical Research: Biogeosciences*, *118*(2), 529–548. https://doi.org/10.1002/jgrg.20051

Ferracci, V., Bolas, C. G., Freshwater, R. A., Staniaszek, Z., King, T., Jaars, K., Otu-Larbi, F., Beale, J., Malhi, Y., Waine, T. W., Jones, R. L., Ashworth, K., & Harris, N. R. (2020). Continuous isoprene measurements in a UK temperate forest for a whole growing season: Effects of drought stress during the 2018 heatwave. *Geophysical Research Letters*, *47*(15). https://doi.org/10.1029/2020gl088885

Geron, C., Guenther, A., Sharkey, T., & Arnts, R. R. (2000). Temporal variability in basal isoprene emission factor. *Tree Physiology*, *20*(12), 799–805. https://doi.org/10.1093/treephys/20.12.799

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., & Geron, C. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (model of emissions of gases and aerosols from nature). *Atmospheric Chemistry and Physics*, *6*(11), 3181–3210. https://doi.org/10.5194/acp-6-3181-2006

Hanson, D. T., & Sharkey, T. D. (2001). Rate of acclimation of the capacity for isoprene emission in response to light and temperature. *Plant, Cell & Environment*, *24*(9), 937–946. https://doi.org/10.1046/j.1365-3040.2001.00745.x

Marais, E. A., Jacob, D. J., Kurosu, T. P., Chance, K., Murphy, J. G., Reeves, C., Mills, G., Casadio, S., Millet, D. B., Barkley, M. P., Paulot, F., & Mao, J. (2012). Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns. *Atmospheric Chemistry and Physics*, *12*(14), 6219–6235. https://doi.org/10.5194/acp-12-6219-2012

Millet, D. B., Jacob, D. J., Turquety, S., Hudman, R. C., Wu, S., Fried, A., Walega, J., Heikes, B. G., Blake, D. R., Singh, H. B., Anderson, B. E., & Clarke, A. D. (2006). Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission. *Journal of Geophysical Research*, *111*(D24). https://doi.org/10.1029/2005jd006853

Monson, R. K., Harley, P. C., Litvak, M. E., Wildermuth, M., Guenther, A. B., Zimmerman, P. R., & Fall, R. (1994). Environmental and developmental controls over the seasonal pattern of isoprene emission from Aspen leaves. *Oecologia*, *99*(3-4), 260–270. https://doi.org/10.1007/bf00627738

Palmer, P. I., Jacob, D. J., Fiore, A. M., Martin, R. V., Chance, K., & Kurosu, T. P. (2003). Mapping isoprene emissions over North America using formaldehyde column observations from space. *Journal of Geophysical Research: Atmospheres*, *108*(D6). https://doi.org/10.1029/2002jd002153

Pétron, G., Harley, P., Greenberg, J., & Guenther, A. (2001). Seasonal temperature variations influence isoprene emission. *Geophysical Research Letters*, *28*(9), 1707–1710. https://doi.org/10.1029/2000gl011583

Sarkar, C., Guenther, A. B., Park, J.-H., Seco, R., Alves, E., Batalha, S., Santana, R., Kim, S., Smith, J., Tóta, J., & Vega, O. (2020). PTR-TOF-MS eddy covariance measurements of isoprene and monoterpene fluxes from an eastern Amazonian rainforest. *Atmospheric Chemistry and Physics*, *20*(12), 7179–7191. https://doi.org/10.5194/acp-20-7179-2020

Stavrakou, T., Müller, J.-F., Bauwens, M., De Smedt, I., Van Roozendael, M., De Mazière, M., Vigouroux, C., Hendrick, F., George, M., Clerbaux, C., Coheur, P.-F., & Guenther, A. (2015). How consistent are top-down hydrocarbon emissions based on formaldehyde observations from GOME-2 and OMI? *Atmospheric Chemistry and Physics*, *15*(20), 11861–11884. https://doi.org/10.5194/acp-15-11861-2015

Turner, A. J., Henze, D. K., Martin, R. V., & Hakami, A. (2012). The spatial extent of source influences on modeled column concentrations of short-lived species. *Geophysical Research Letters*, *39*(12). https://doi.org/10.1029/2012gl051832

Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., De Smedt, I., Grutter, M., Hannigan, J. W., Jones, N., Kivi, R., Loyola, D., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J.-M., Morino, I., Murata, I., Nagahama, T., … Winkler, H. (2020). Tropomi–sentinel-5 precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations. *Atmospheric Measurement Techniques*, *13*(7), 3751–3767. https://doi.org/10.5194/amt-13-3751-2020

Wolfe, G. M., Kaiser, J., Hanisco, T. F., Keutsch, F. N., de Gouw, J. A., Gilman, J. B., Graus, M., Hatch, C. D., Holloway, J., Horowitz, L. W., Lee, B. H., Lerner, B. M., Lopez-Hilifiker, F., Mao, J., Marvin, M. R., Peischl, J., Pollack, I. B., Roberts, J. M., Ryerson, T. B., … Warneke, C. (2016). Formaldehyde production from isoprene oxidation across NOx regimes. *Atmospheric Chemistry and Physics*, *16*(4), 2597–2610. https://doi.org/10.5194/acp-16-2597-2016