# Next-generation Isoprene Measurements from Space: Quantifying Daily Variability at High Resolution

**K. C. Wells[1], D. B. Millet[1], V. H. Payne[2], C. Vigouroux[3], C. A. B. Aquino[4], M. De Mazière[3], J. A. de Gouw[5,6], M. Graus[7], T. Kurosu[2], C. Warneke[5,8], and A. Wisthaler[9,10]**

[1]Department of Soil, Water, and Climate, University of Minnesota, Saint Paul, MN, USA.

[2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.

[3]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium.

[4]Instituto Federal de Educaçao, Ciência e Tecnologia de Rondônia (IFRO), Porto Velho, Brazil.

[5]Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, CO, USA.

[6]Department of Chemistry, University of Colorado, Boulder, CO, USA.

[7]Department of Atmospheric and Cryogenic Sciences, University of Innsbruck, Innsbruck, Austria

[8]NOAA Earth System Research Laboratory, Boulder, CO, USA.

[9]Institute for Ion Physics and Applied Physics, University of Innsbruck, Innsbruck, Austria.

[10]Department of Chemistry, University of Oslo, Oslo, Norway.

Corresponding author: Dylan Millet (dbm@umn.edu)

**Key Points:**

- We present a next-generation spaced-based isoprene retrieval with higher sensitivity and resolution than previous approaches
- Global, daily isoprene distributions are derived from 2012-2020 that compare well with the first ground-based isoprene column observations
- High-resolution results over isoprene hotspots highlight processes controlling isoprene abundance and its daily-to-interannual variability

## Abstract

Isoprene is the dominant non-methane organic compound emitted to the atmosphere, where it drives ozone and aerosol production, modulates atmospheric oxidation, and interacts with the global nitrogen cycle. Isoprene emissions are highly variable and uncertain, as is the non-linear chemistry coupling isoprene and its primary sink, the hydroxyl radical (OH). Space-based isoprene measurements can help close the gap on these uncertainties, and when combined with concurrent formaldehyde data provide a new constraint on atmospheric oxidation regimes. Here we present a next-generation machine-learning isoprene retrieval for the Cross-track Infrared Sounder (CrIS) that provides improved sensitivity, lower noise, and thus higher space-time resolution than earlier approaches. The Retrieval of Organics with CrIS Radiances (ROCR) isoprene measurements compare well with previous space-based retrievals as well as with the first-ever ground-based isoprene column measurements, with 20-50% discrepancies that reflect differing sources of systematic uncertainty. An ensemble of sensitivity tests points to the spectral background and isoprene profile specification as the most relevant uncertainty sources in the ROCR framework. We apply the ROCR isoprene algorithm to the full CrIS record from 2012-2020, showing that it can resolve fine-scale spatial gradients at daily resolution over the world's isoprene hotspots. Results over North America and Amazonia highlight emergent connections between isoprene abundance and daily-to-interannual variations in temperature, nitrogen oxides, and drought stress.



## Plain Language Summary

Isoprene is a naturally occurring trace gas emitted primarily from the leaves of woody plants. Isoprene has important impacts on both air quality and climate; however, these impacts are difficult to assess and predict given large uncertainties in its sources and atmospheric chemistry. Space-based measurements can help to address these uncertainties. Here we present new satellite measurements of isoprene from the Cross-track Infrared Sounder (CrIS), using a computationally efficient machine-learning framework (Retrieval of Organics from CrIS Radiances; ROCR). ROCR measurements provide improved sensitivity and richer spatiotemporal information on atmospheric isoprene than was previously available. Results compare well to previous satellite-based approaches and to new ground-based observations. We apply the ROCR framework to measure daily, global isoprene distributions from 2012-2020. Results over North America and

Amazonia highlight the processes controlling isoprene abundances and their variability over time.

## 1 Introduction

Isoprene is the dominant non-methane volatile organic compound (VOC) emitted to the atmosphere (Guenther et al., 2012). Produced mainly in the leaves of woody plants, isoprene is highly reactive and drives ozone and aerosol production (Lin et al., 2013; Paulot et al., 2012), modulates atmospheric oxidation (Bates and Jacob, 2019), and affects the global nitrogen cycle (Mao et al., 2013; Paulot et al., 2013). Accurate flux estimates are critical for assessing and predicting these impacts; however, bottom-up isoprene inventories are highly uncertain as they i) rely on emission factors extrapolated from limited point measurements and ii) are sensitive to model assumptions for land cover, meteorology, and plant canopy structure (Arneth et al., 2011; Messina et al., 2016; Ganzeveld et al., 2002). Particular uncertainties have been identified in the world's isoprene hotspots such as Amazonia, where studies show that isoprene exhibits much stronger seasonal and spatial variability than can be explained by current models (Barkley et al., 2009; Alves et al., 2018; Wei et al., 2018; Gu et al., 2017; Batista et al., 2019). Isoprene's impact on atmospheric oxidation has been a further subject of debate, and specifically the degree to which it acts to sustain or deplete hydroxyl radical (OH) concentrations under low-$NO_x$ conditions (Lelieveld et al., 2008; Fuchs et al., 2013; Feiner et al., 2016).

Satellite-based measurements of isoprene are beginning to provide powerful new information for addressing these gaps (Fu et al., 2019). Our team recently developed the first global isoprene measurements from space by applying an efficient machine-learning algorithm to thermal infrared radiances from the Cross-track Infrared Sounder (CrIS; Wells et al., 2020). That approach derived isoprene column abundances from the CrIS-measured on-peak/off-peak brightness temperature difference ($\Delta T_b$) at the $\nu_{28}$ absorption feature, and showed that combining these data with concurrent formaldehyde observations affords joint constraints on isoprene emissions and chemistry over source regions. A major finding of that work was that the isoprene:formaldehyde relationship observed from space supports current understanding of isoprene-OH chemistry, with no indication of missing OH recycling at low $NO_x$. The new measurements also pinpointed regions where emission errors for both isoprene and $NO_x$ cause major prediction biases in current models.

In this paper, we present the Retrieval of Organics from CrIS Radiances (ROCR) next-generation isoprene retrieval, which improves on our original machine learning algorithm by employing a hyperspectral range index (HRI) to quantify column abundances. The HRI approach has been used previously with the Infrared Atmospheric Sounding Interferometer (IASI) to retrieve global distributions of ammonia (Whitburn et al., 2016), methanol, formic acid, peroxyacetyl nitrate (Franco et al., 2018), acetone (Franco et al., 2019), and acetic acid (Franco et al., 2020). By leveraging a broader spectral range than the brightness temperature difference, the HRI increases near-surface sensitivity while reducing impacts from interferences; our updated algorithm thus enables isoprene detection at unprecedented resolution while maintaining high computational efficiency. We apply the ROCR algorithm here to obtain global isoprene distributions on a daily basis from 2012 through 2020. We evaluate the results against other observations, including the first ground-based column retrievals of atmospheric isoprene, and characterize pertinent sources of measurement uncertainty. Finally, we explore this long-term global dataset, and highlight the new, high-resolution information provided over two key isoprene hotspots.

## 2 Materials and Methods

CrIS is a Fourier transform spectrometer flying in a sun-synchronous orbit onboard Suomi-NPP (SNPP, launched 10/2011) and JPSS-1/NOAA-20 (launched 11/2017), with a third instrument planned for inclusion on JPSS-2 (launch expected 09/2022). CrIS has 0.625 $cm^{-1}$ spectral resolution in the longwave IR (LWIR; 650-1095 $cm^{-1}$), and an angular field of regard consisting of a 3×3 pixel array with a 14-km diameter footprint at nadir and a 2200 km cross-track scan width that provides near-global coverage twice daily. This sampling strategy affords the opportunity for high-resolution quantification of daily isoprene distributions. The early-afternoon daytime overpass (~1330 LT for SNPP and ~1240 LT for JPSS-1/NOAA-20) typically corresponds with peak isoprene emissions, enhanced vertical mixing, and strong land-atmosphere thermal contrast—all of which increase sensitivity to near-surface absorbers. CrIS also features significantly lower noise (e.g., ~0.04 K at 900 $cm^{-1}$ and 280 K) than other atmospheric sounders (Zavyalov et al., 2013).

### 2.1 CrIS HRI derivation

The ROCR algorithm derives isoprene column abundances based on the HRI (Eq. 1), which is a dimensionless quantity measuring the spectral signature of a target atmospheric species (Walker

et al., 2011). The retrieval begins with single-footprint Level 1B CrIS spectra (available from 02/2012 onward from SNPP) over land. We cloud-screen the spectra based on the difference between the MERRA-2 (Gelaro et al., 2017) surface temperature and the CrIS-measured 900 $cm^{-1}$ brightness temperature (Fig. S1), following Wells et al. (2020). The HRI is then computed for each spectrum via (Franco et al., 2018):

$$\mathrm{HRI} = \frac{1}{N}\frac{\mathbf{K}^T\mathbf{S}_{\bar{y}}^{-1}(\boldsymbol{y}-\bar{\boldsymbol{y}})}{\sqrt{\mathbf{K}^T\mathbf{S}_{\bar{y}}^{-1}\mathbf{K}}} \quad (1)$$

Here, $\boldsymbol{y}$ is the measured spectrum, while $\bar{\boldsymbol{y}}$ and $\mathbf{S}_y$ are respectively the mean background spectrum and background covariance matrix, both calculated from scenes in which isoprene is undetectable. $\mathbf{S}_y$ characterizes the expected correlations between spectral channels due to factors other than isoprene, such as interfering trace gases (Whitburn et al., 2016). $\mathbf{K}$ is the spectral Jacobian for a change in the target species, which we calculate at the midpoint of 12 equally-spaced viewing angle bins (0-5° to 55-60°) using radiances generated by the Earth Limb and Nadir Operational Retrieval (ELANOR) radiative transfer model (Clough et al., 2006) and simulated isoprene output from GEOS-Chem (v11-02e, www.geos-chem.org). The employed spectral range (890-910 $cm^{-1}$) encompasses the two $=CH_2$ wag absorption peaks exhibited by the isoprene molecule (Brauer et al., 2014).

The background quantities $\bar{\boldsymbol{y}}$ and $\mathbf{S}_y$ are calculated monthly for each angle bin from CrIS spectra that are selected iteratively following the approach used for IASI VOC retrievals (Franco et al., 2018). We start by calculating the HRI values for all spectra in a given month and angle bin. After removing spectra with an HRI exceeding a specified threshold, we rederive $\bar{\boldsymbol{y}}$, $\mathbf{S}_y$, and the corresponding HRIs, and iterate until convergence upon a spectral ensemble with below-threshold values. Each iteration step involves normalization (by $N$) to maintain an HRI mean of 0.0 and standard deviation of 1.0 for background conditions. $N$ is calculated as the HRI standard deviation over a region where target species detection is not expected; we employ spectra over central Australia (20°-30° S, 122°-137° E) for this purpose and test the use of an alternate normalization region in Section 3.2. We achieve enhanced sensitivity for isoprene with a background HRI threshold of 1.0, consistent with prior VOC findings for IASI (Franco et al., 2018); we explore the retrieval sensitivity to this threshold in Section 3.2. Once iteration is complete, we generate daily gridded HRI maps at 0.5° × 0.625° resolution.

By encompassing a broader spectral range over which the target species is optically active, the HRI generated as above delivers improved sensitivity over $\Delta T_b$-based and other approaches. It also lessens the impacts of interferents by using measured radiances to account for spectral correlations under background conditions. Figure S2 demonstrates this improvement by comparing the isoprene HRI and $\Delta T_b$ (at $\nu_{28}$) for synthetically-generated radiances under varying environmental conditions and a given viewing geometry. The HRI exhibits a tighter linear correlation with isoprene (r = 0.78 versus 0.51 in this example) with significantly less scatter for low and moderate column densities ($\Omega_{\text{isoprene}} < 1 \times 16$ molec cm$^{-2}$). Thermal contrast, water vapor and other factors still drive some variability in the HRI-isoprene relationship, and we account for these residual effects using a neural network as described next.

## 2.2 Machine-learning retrieval

Isoprene abundances are derived from the CrIS HRIs using a feed-forward neural network (NN) as employed in our previous work (Wells et al., 2020). The NN is trained using a synthetic HRI dataset generated from a full year of global overland ELANOR radiances simulated at the midpoint of the 12 angle bins described earlier. ELANOR inputs include temperature and water vapor profiles from the NASA Goddard Modeling and Assimilation Office (GMAO) and isoprene profiles from GEOS-Chem (Wells et al., 2020), with the latter subjected to 100% 1σ Gaussian noise to ensure that NN predictions reflect the spectroscopic effects of the input variables rather than any prior correlations between them. We replicate the ELANOR spectral output 25 times per scene, with CrIS-like noise applied each time, and compute the resulting simulated HRIs as:

$$\text{HRI}_{sim} = \frac{\mathbf{K}^T \mathbf{S}_y^{-1} (\boldsymbol{y} - \boldsymbol{y_o})}{\sqrt{\mathbf{K}^T \mathbf{S}_y^{-1} \mathbf{K}}} \qquad (2)$$

where $\boldsymbol{y}$ and $\boldsymbol{y_o}$ are simulated spectra with and without the target species, and $\mathbf{S}_y$ is the CrIS-observed background spectral covariance matrix for the same month-of-year. Using spectral pairs with and without the target species for the simulated HRI reduces retrieval sensitivity to forward model errors, while preserving a mean of zero in the absence of the target gas (Franco et al., 2018). The final $\text{HRI}_{sim}$ for each scene is obtained as the mean across the 25 replications.

We then train the NN to convert the HRI values to isoprene columns while accounting for the additional factors that affect that relationship. For isoprene we find that these comprise water

vapor column, thermal contrast, surface pressure, and viewing angle. Our previous $\Delta T_b$-based retrieval also included nitric acid as a predictor (Wells et al., 2020), but we find here that its inclusion does not improve predictive power, implying that its effects are already captured in the HRI computation. Other factors found here (surface emissivity) or previously (ammonia, CFCs; Wells et al., 2020) to confer no predictive benefit are similarly omitted.

After testing multiple architectures, we find optimum performance for a two-layer network with 10 and 5 nodes. The training proceeds on 10 random extractions of the data, with 50% of each extraction used for training, 30% for validation, and 20% for testing. Finally, we apply the resulting 10 NNs to the CrIS-measured HRIs with temperature and water information obtained from MERRA-2 reanalysis (Gelaro et al., 2017), and derive the final retrieved isoprene column as the mean output across the 10 NNs.

We have applied the above ROCR algorithm to obtain global, daily distributions of isoprene at $0.5° \times 0.625°$ resolution spanning most of the CrIS SNPP record (2012-2020). As we demonstrate below, the retrieval exhibits significantly less noise than our previous version outlined in Wells et al. (2020). We do obtain some artifacts at high latitudes that we attribute to surface effects in the presence of ice and snow. We therefore postfilter scenes with surface temperatures below 273 K as we do not expect to detect isoprene in such conditions; a similar approach is used for IASI VOC retrievals (Franco et al., 2018; Franco et al., 2019).

Next, in Section 3, we present a series of intercomparisons and sensitivity tests to characterize relevant uncertainties in the ROCR isoprene retrievals. Subsequently, in Section 4 we explore the spatial and temporal information provided by these new satellite-based isoprene measurements over key global source regions.

## 3 Validation and error characterization

### 3.1 Comparison to OE and brightness-temperature difference approach

We begin with a spatial evaluation of the next-generation ROCR isoprene retrievals against previously published results generated using optimal estimation (OE; Fu et al., 2019) and brightness-temperature difference (Wells et al., 2020) methods. Fig. 1 compares these three datasets over Amazonia during September 2014. The high degree of spatial consistency (panels a-c) and strong correlation (r = 0.92-0.93, panels d-e) among them provides confidence in the

isoprene distribution provided by these separate approaches. Furthermore, we see reduced scatter at low-to-moderate isoprene for the HRI-OE comparison (Fig. 1e) compared to the $\Delta T_b$-OE case (Fig. S3)—demonstrating the HRI signal-to-noise improvement over the earlier $\Delta T_b$ approach.

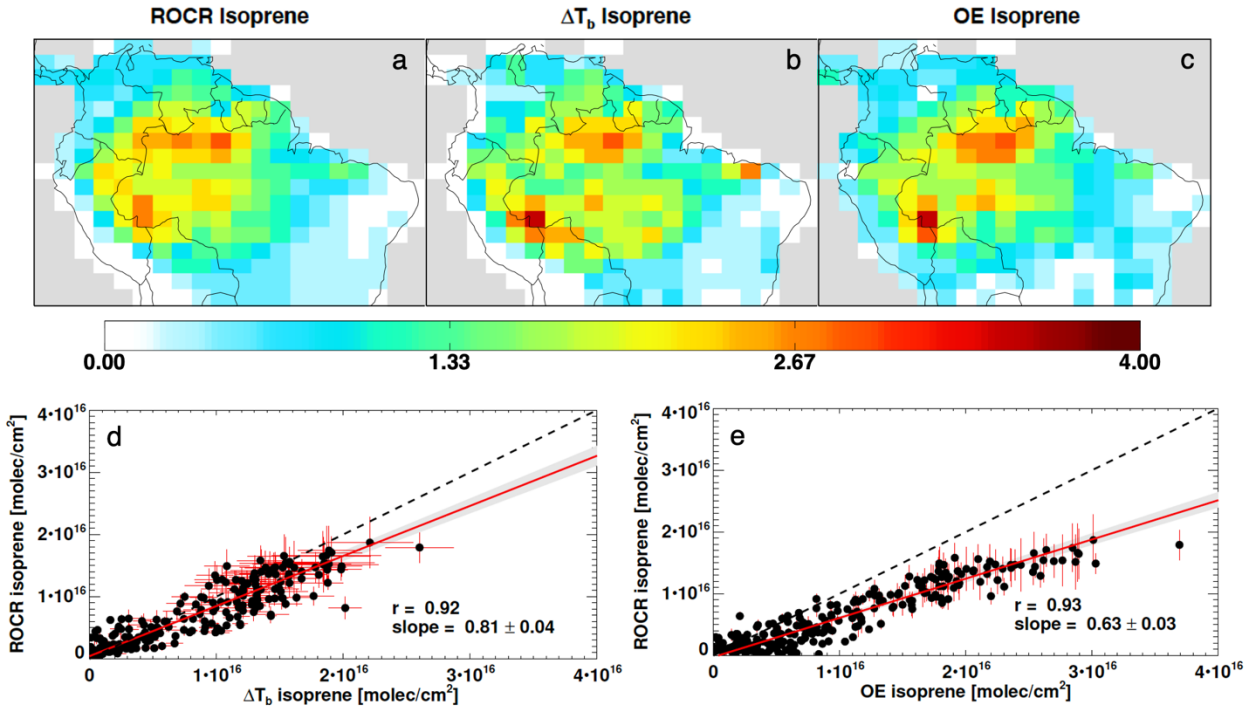


**Figure 1. Isoprene retrieval comparison over Amazonia during September 2014. Maps display monthly-mean columns averaged on a 2° × 2.5° grid and normalized to their domain means for (a) the ROCR HRI-based retrieval, (b) the $\Delta T_b$- based retrieval presented in Wells et al. (2020), and (c) the OE-based retrieval presented in Fu et al. (2019). Scatter plots compare the corresponding absolute column densities. Error bars in the scatter plots show the standard deviation across the 10 NNs; the red line and gray shaded area indicate the reduced major axis regression and bootstrapped 95% confidence interval, respectively. The dashed line shows the 1:1 relationship.**

The major-axis intercomparison slope for the ROCR HRI-based isoprene columns is 0.81 vs. the $\Delta T_b$-based columns and 0.63 vs. the OE-based columns (Fig. 1d-e). As discussed by Franco et al. (2018), each of these retrieval approaches is subject to its own uncertainties, some of which may be systematic. In the OE case, these include forward model assumptions, fitting errors for interfering species, and the dependence on a priori target species information. In particular, the OE columns in Fig. 1 (Fu et al., 2019) employ isoprene columns from GEOS-Chem as prior, which have been shown to have a major high bias over Amazonia (Wells et al., 2020); this may partly explain the difference seen in Fig. 1e. Uncertainties for the HRI-based approach include assumptions in the forward model set-up and in the HRI derivation. We examine the impact of these in Section 3.2, and present a comparison of the ROCR HRI-based isoprene retrievals to independent ground-based column observations in Section 3.3.

## 3.2 Uncertainty analysis

We assess the uncertainties of the HRI-based results through a series of sensitivity analyses targeting specific aspects of the retrieval. We focus in particular on the HRI background computation, cloud screening, normalization approach, vertical mixing, and potential interferences from monoterpenes.

### *3.2.1 Background definition*

The background spectrum and covariance matrix ($\bar{\boldsymbol{y}}$, $\mathbf{S}_y$) were derived in Section 2.1 using an HRI < 1 threshold following Franco et al. (2018). We tested the impact of this selection by reanalyzing a full year of CrIS spectra with an alternate background definition (HRI < 2). The resulting $\mathbf{S}_y$ matrix was used to rederive the simulated HRI values, with the NN then retrained and applied to the updated CrIS HRI fields.

Figure 2a-b compares the resulting monthly-mean isoprene columns to those obtained with our standard approach. The comparison slope across the entire dataset is 0.774 ± 0.002 (bootstrapped 95% confidence interval); however, differences occur primarily for points with $\Omega_{\text{isoprene}} < 1 \times 10^{16}$ molec cm$^{-2}$. Use of a higher threshold leads to retention of more low-to-moderate isoprene scenes in the background calculation, yielding a negative bias at low column amounts. This impact is strongly reduced when isoprene is elevated, with mean differences decreasing from 20% at $1 \times 10^{16}$ molec cm$^{-2}$ to near-zero at higher column amounts.

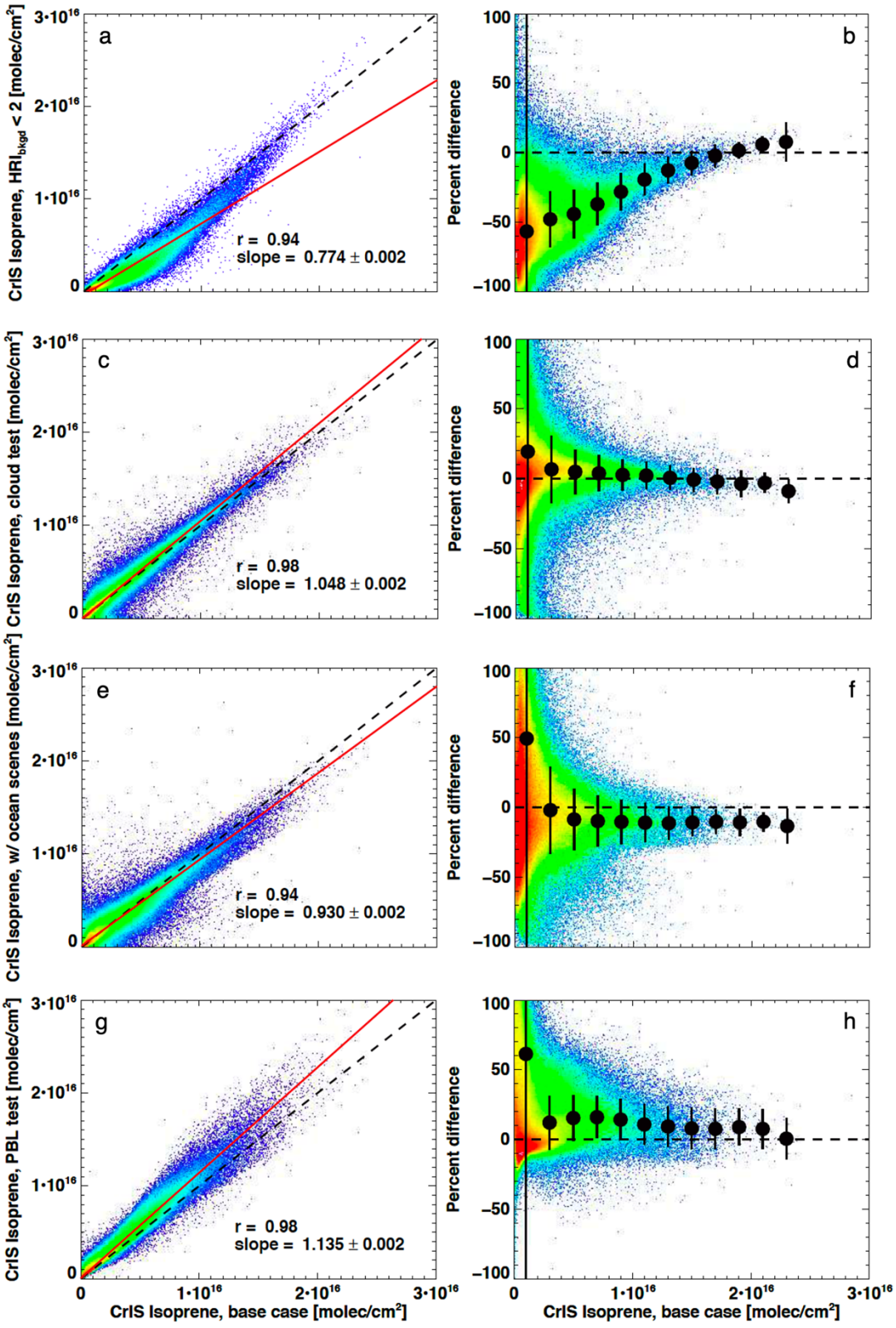


**Figure 2. ROCR isoprene sensitivity to different aspects of the retrieval: (a) Comparison of monthly-mean isoprene columns derived with a background definition of HRI<2 against the baseline approach using HRI<1. (c) Comparison of monthly-mean isoprene columns derived with a more stringent cloud-screening threshold**

**against the baseline approach (see text). (e) Comparison of monthly-mean isoprene columns derived when including ocean scenes against the baseline land-only approach. (g) Comparison of monthly-mean isoprene columns derived when using an alternative PBL mixing scheme during NN training. The red and dashed lines show the reduced major axis regression and 1:1 relationship, respectively. (b, d, f, and h) Percent difference ($100\% \times [\Omega_{test} - \Omega_{base}]/\Omega_{base}$) for each test as a function of the baseline column amount, with black dots and error bars denoting the mean and standard deviation for each bin. Coloring in each panel indicates the data density.**

### *3.2.2 Cloud-screening*

Clouds are an important consideration in thermal-IR retrievals, generally acting to reduce signal by obscuring the below-cloud portion of the trace gas column (Whitburn et al., 2016). For that reason we cloud-screen the CrIS spectra as described in Section 2.1; if such screening were inadequate we would expect a low retrieval bias due to retained cloudy scenes. To characterize the uncertainties associated with cloud-screening, we performed a sensitivity test using a more conservative cloud threshold (based on the clear-sky difference between the 900 $cm^{-1}$ brightness temperature and surface skin temperature, as shown in Fig. S1). As above, we then reanalyzed a year of CrIS spectra, regenerated the training set using the updated $\mathbf{S}_y$, retrained the NN, and applied it to the observed HRI fields. Results plotted in Fig. 2c-d show that for all but the lowest isoprene columns the impact of this alternate cloud treatment is minimal, with mean differences of < 5% relative to the base-case. We conclude that clouds are not a predominant source of uncertainty in these retrievals.

### *3.2.3 Inclusion of ocean scenes and HRI normalization*

Given isoprene's terrestrial sources and short atmospheric lifetime, we restrict the base-case retrieval to land scenes for computational efficiency. The HRI normalization step, which requires scenes lacking the target species, then employs spectra over central Australia as outlined in Section 2.1. We tested the impact of this treatment by reanalyzing a year of CrIS spectra for both ocean and land scenes, with HRI normalization based on spectra over the remote Pacific (10° to 30° S, 180° to 130° W) rather than Australia. Results are shown in Fig. 2e-f. Isoprene columns derived when including vs. excluding ocean scenes generally agree to within 5-10% on average for retrievals above $0.5 \times 10^{16}$ molec $cm^{-2}$, showing that this data selection criterion does not impart significant uncertainty to the results.

### *3.2.4 Vertical mixing*

Satellite-based measurement sensitivity in the thermal IR depends on the vertical distribution of the absorber and hence on atmospheric mixing. To test the sensitivity of the retrievals to model vertical mixing assumptions, we generated a new HRI training set from ELANOR simulations driven by GEOS-Chem isoprene profiles produced using an alternate planetary boundary layer (PBL) scheme (Wu et al., 2007). Here, all surface emissions are mixed instantaneously through the vertical extent of the boundary layer, resulting in higher lofting of isoprene than in the standard case, which employs the GEOS-Chem default non-local PBL mixing scheme (Lin and Mcelroy, 2010). Fig. S4 compares results for these two mixing treatments over Amazonia. We then trained a NN with this modified output and applied it to a year of CrIS-measured HRI values.

Figure 2g-h shows that the resulting columns have mean differences of up to 20% between these two cases, with higher columns generally obtained with the full-mixing scheme. By itself we would expect the enhanced isoprene lofting to yield a higher HRI for a given isoprene abundance—therefore leading to lower column predictions from the CrIS HRI fields. This is indeed what occurs under dry conditions (Fig. S5). However, isoprene is predominantly emitted in humid tropical climates, and under wet conditions enhanced water-driven IR absorption below the lofted isoprene yields the opposite effect (the water distribution is prescribed by the assimilated meteorological fields driving the simulation, and is not altered by the change in mixing scheme). Overall, however, we can consider the <20% mean differences arising from this sensitivity test (as shown in Fig. 2d) to reflect the envelope of uncertainty arising from vertical mixing as currently represented in chemical transport models.

### *3.2.5 Potential interference from monoterpenes*

In our previous work (Fu et al., 2019; Wells et al., 2020) we assessed the potential for other molecules with terminal $=CH_2$ groups to interfere with the isoprene signal, and concluded that monoterpenes are likely the most important such species to consider. To estimate the extent of their potential impact, we compute and compare optical depths for isoprene and total monoterpenes at the peak of the $\nu_{28}$ isoprene feature. For purposes of this assessment, optical depths are approximated as the product of the GEOS-Chem column densities and the corresponding absorption cross sections (Gordon et al., 2017) at $\nu_{28}$ and 298 K. Monoterpenes are simulated in GEOS-Chem as MTPA ($\alpha$-pinene + $\beta$-pinene + sabinene + carene), MTPO

(terpinene + terpinolene + myrcene + ocimene + other monoterpenes), and limonene. For MTPA we employ the $\beta$-pinene cross-section (which is the largest reported among constituent species) whereas for MTPO we employ the myrcene cross-section (the only one reported among constituent species).

While these biogenic species share a number of emission hotspots, Fig. 3 shows that the total estimated monoterpene optical depth is typically <5% that of isoprene for major isoprene source regions. The monoterpene signal only becomes relevant for isoprene columns < $5 \times 10^{14}$ molec $cm^{-2}$, well below the CrIS limit of detection (Wells et al., 2020). Furthermore, this influence is likely to be an upper limit, for two reasons. The first is our assumption that all MTPO species absorb as efficiently as $\beta$-pinene. The second is that all the published monoterpene cross sections in 890-910 $cm^{-1}$ (Gordon et al., 2017) are spectrally distinct from that of isoprene, diminishing any potential impact on the HRI. We therefore consider monoterpene interferences to be minimal for the isoprene retrievals shown here.

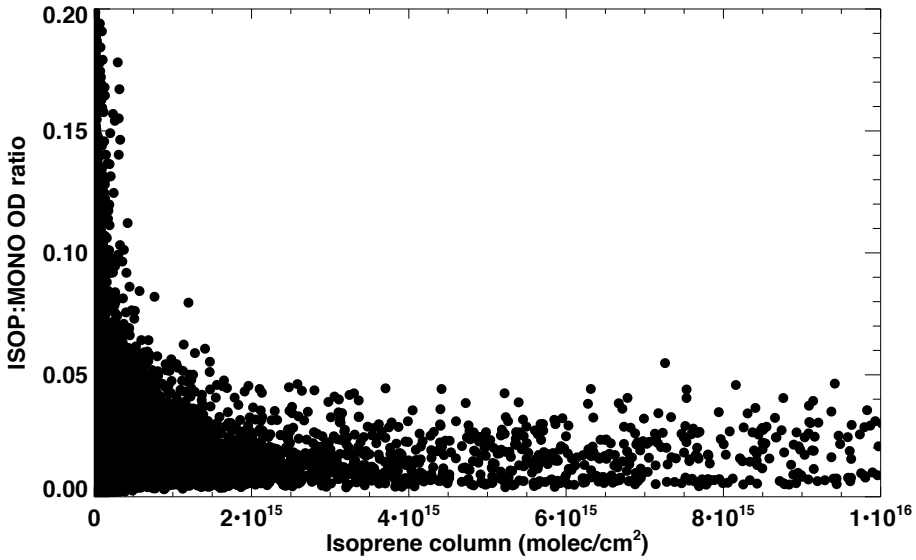


**Figure 3. Potential impact of monoterpenes on the CrIS isoprene retrievals. Shown is the estimated isoprene:monoterpene optical depth ratio as a function of isoprene column. Optical depths are derived from GEOS-Chem model predictions as described in-text.**

### 3.3 Comparison to ground-based isoprene retrievals

We showed in the last section that different retrieval assumptions, particularly the HRI background definition and vertical mixing treatment in the forward model, can each lead to systematic uncertainties on the order of 10-20% for isoprene columns > $1 \times 10^{16}$ molec $cm^{-2}$ (and higher at lower column amounts). For that reason, independent evaluation is key. Here, we

compare the ROCR isoprene retrieval from CrIS against ground-based Fourier Transform InfraRed (FTIR) isoprene column measurements at Porto Velho (8.77° S, 63.87° W), on the border between the Brazilian states of Rondônia and Amazonas.

The ground-based instrument is a Bruker 125M high-resolution (up to 0.006 $cm^{-1}$) spectrometer previously deployed at Saint-Denis, Réunion Island (Vigouroux et al., 2009, 2012). Since 2016, the instrument has been used at Porto Velho for satellite validation of formaldehyde, methane, and carbon monoxide (Vigouroux et al., 2020; Sha et al., 2021). Measurements in the spectral range needed for isoprene detection started in June 2019 and ended October 2019.

These measurements represent the first retrievals of isoprene from ground-based FTIR spectra. Information on retrieval settings is provided in Table 1 and we refer to Vigouroux et al. (2012) for further details. Retrievals employ the SFIT4 algorithm commonly used in the InfraRed Working Group of the Network for the Detection of Atmospheric Composition Change (NDACC), with the spectral range extended to 917 $cm^{-1}$ to encompass the $CCl_2F_2$ feature maximum. Figure S6 shows an example of the retrieved spectral signatures. As Porto Velho is located in a tropical environment, the $H_2O$ lines are very strong (Fig. S6b) and we therefore apply a de-weighting (Signal-to-Noise Ratio of 0 instead of 200) to some of these lines that are not well-fitted (see residuals in Fig. S6c). The theoretical (OE) mean total random uncertainty is approximately 7% ($3.4 \times 10^{14}$ molec $cm^{-2}$) and is due primarily to temperature uncertainties (~6%; Vigouroux et al., 2018) and measurement noise (~3%). The mean theoretical total systematic uncertainty is approximately 21%, dominated by spectroscopic (~20%) and temperature (~4%) components.

Figure 4a shows that the monthly-mean ground-based retrievals derived as above agree well with the CrIS observations, with both datasets exhibiting consistent seasonal increases between June and September. We also see significant CrIS-FTIR correlation on a daily timescale ($r = 0.47$ when employing FTIR spectra within ± 2 h of the CrIS overpass), showing that both datasets are capturing short-term ambient fluctuations. The daily isoprene columns from CrIS are somewhat higher than the ground-based results (slope: 1.5; 95% confidence interval: 1.1-2.0, Fig. 4b), whereas they were somewhat lower than the previously published OE and $\Delta T_b$ retrievals shown in Section 3.1.

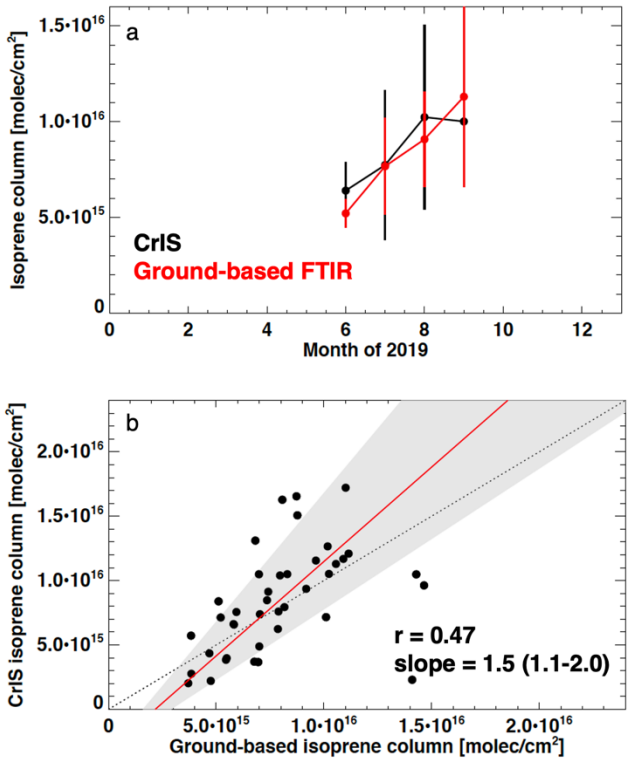


**Figure 4. Comparison of CrIS ROCR and ground-based FTIR isoprene retrievals at Porto Velho, Brazil for June-September 2019. (a) Monthly mean (± standard deviation) isoprene columns based on the CrIS (black) and ground-based (red) retrievals. (b) Daily mean CrIS vs. ground-based (±2 hours of the CrIS overpass) isoprene columns; the reduced major axis regression and bootstrapped 95% confidence interval are indicated by the red line and gray shaded area, respectively. The dotted line shows the 1:1 relationship.**

The Porto Velho measurements are slated to resume in the future, and with time should provide an increasingly robust resource for quantitatively testing the space-based retrievals. In theory, such measurements can also be performed across the global network of ground-based solar FTIR stations (Vigouroux et al., 2020). However, this site in Amazonia is uniquely situated for the validation of isoprene, which is measurable from space primarily over hotspot regions (Wells et al., 2020).

## 4 Global distribution and key hotspot results

Figure 5 shows the global distribution of monthly-mean ROCR isoprene retrievals as a multi-year mean over the 2012-2020 SNPP CrIS record. These data represent the first space-based quantification showing the full seasonal cycle of atmospheric isoprene. We see the highest columns over Amazonia, with a widespread regional maximum in September and a secondary more localized maximum over northwestern Brazil peaking in April. Within the northern

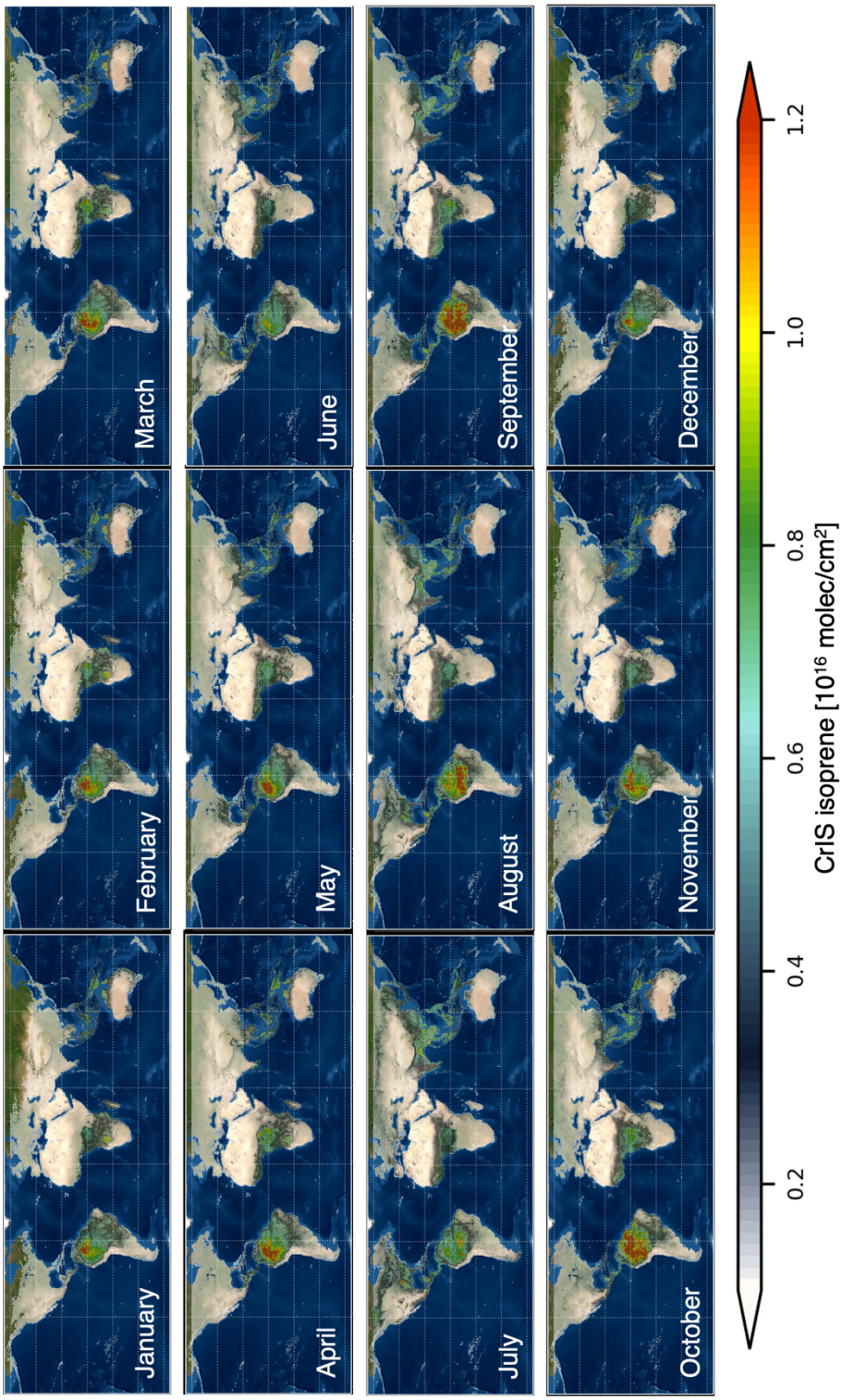


**Figure 5. Global distribution of gridded (0.5° × 0.625°) isoprene column densities as measured from SNPP CrIS, averaged from 2012-2020 for each month of the year (2013-2020 for January). Data are plotted with partial transparency on the underlying land cover; missing data (e.g., in the wintertime high northern latitudes) have full transparency.**

midlatitudes, the highest isoprene columns occur during summer over the Ozarks region of the US, with other summertime enhancements detected over the forests of Canada and Russia.

Over Africa, CrIS reveals a distinct north-south seasonal shift, with isoprene enhancements in the southeast (Angola/Zambia) peaking in January-February, throughout the Sahel during May-October, and in central Africa during March-April and September-October. A similar seasonal pattern manifests over Oceania and East Asia: during February-April the highest columns occur over northern Australia, with this peak then shifting northward to Southeast Asia, eastern China, and the eastern Indian subcontinent by July-September. Detection for many of the above patterns is made feasible by the improved sensitivity of the HRI retrieval: over East Asia, for example, the short isoprene lifetimes and lower column amounts hindered detection with the previous $\Delta T_b$ approach (Wells et al., 2020).

Below, we explore these results in more detail over two key isoprene source regions—the United States+Mexico and Amazonia—demonstrating in particular the capabilities of this new dataset for resolving patterns of isoprene variability at high spatial resolution and on daily timescales.

### 4.1 United States and Mexico

Figure 6 shows the 2012-2020 ROCR isoprene retrievals over the US and Mexico averaged by month for May-October. The data exhibit strong spatial heterogeneity at unprecedented resolution, with major hotspots over the "isoprene volcano" in the southern Missouri/northern Arkansas Ozarks (Wiedinmyer et al., 2005) and over the Yucatán Peninsula in Mexico, an area with some of the highest predicted emission rates globally (Opacka et al., 2021). Substantial isoprene enhancements are also observed over the South-Central Plains in eastern Texas and western Louisiana, where an emission underestimate was previously inferred from satellite-based measurements of formaldehyde (Kaiser et al., 2018). We also see elevated isoprene columns over source regions along the Piedmont Plateau in Appalachia, in coastal Mexico, in the national forests of southwestern Colorado, in northern California, and in the Sierra Nevada foothills.

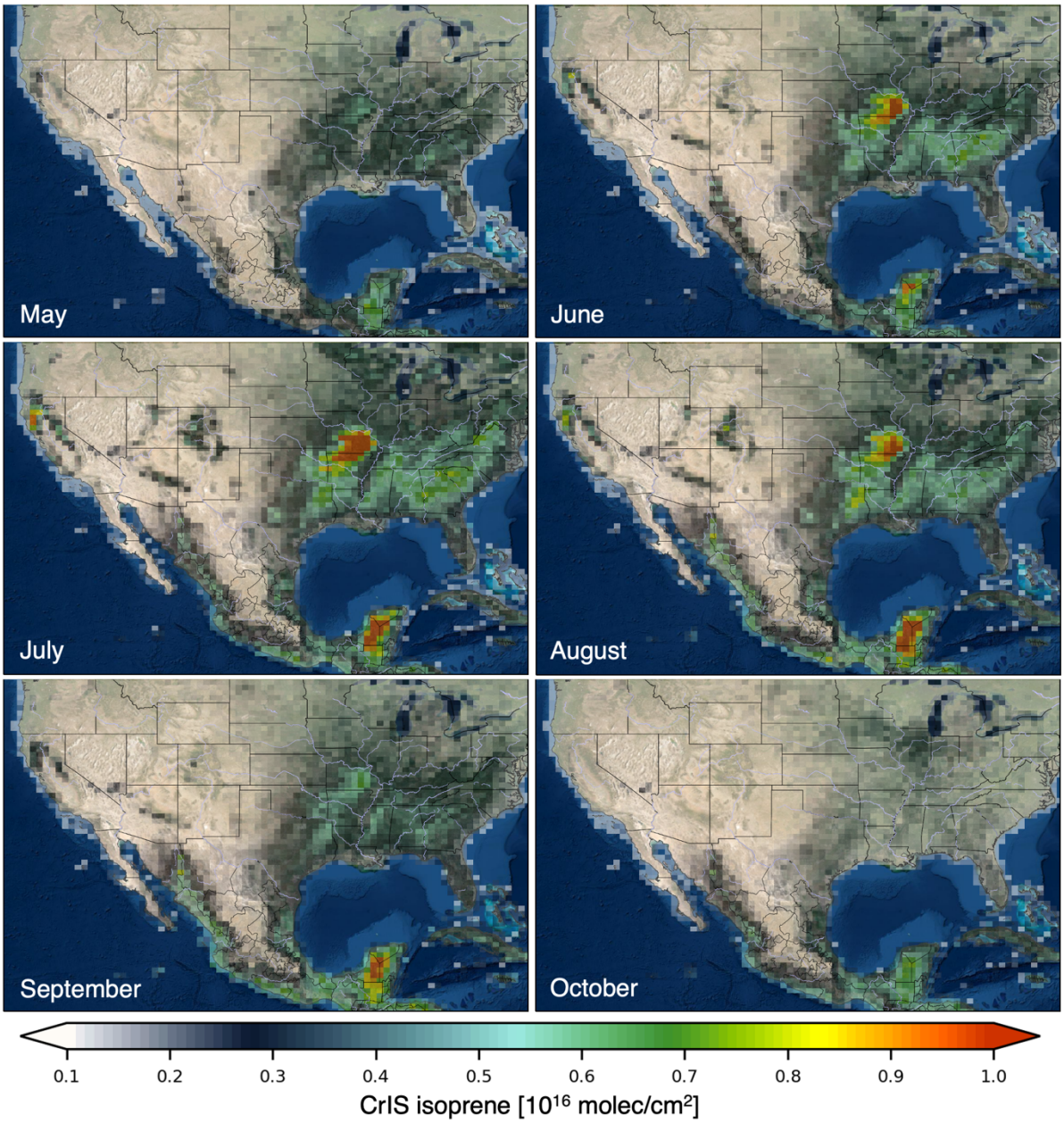


**Figure 6. Distribution of gridded (0.5° × 0.625°) isoprene column densities over the US and Mexico as measured from SNPP CrIS, averaged from 2012-2020 by month for May-October.**

The fine-scale patterns of isoprene variability seen by CrIS over North America are also supported by independent data. For example, Fig. 7 shows that aircraft measurements during the SENEX (Warneke et al., 2016) and SEAC$^4$RS (Toon et al., 2016) campaigns in the southeast US exhibit spatial structure that agrees well with CrIS, with all three datasets identifying the Ozarks as the primary isoprene hotspot in this region. Over this hotspot, the CrIS measurements are further able to resolve daily isoprene variability correlating strongly with surface temperature

during summer (Fig. 8), reflecting the underlying emission dependence (Guenther et al., 1993). Quantifying this isoprene-temperature relationship on a daily basis represents a major advance over our previous retrievals, which required monthly averaging to enhance signal-to-noise (Wells et al., 2020). We see in Fig. 8j that the isoprene-temperature dependence seen from space by CrIS corroborates the bottom-up predictions, with an absolute offset consistent with the isoprene lifetime underestimate inferred previously for GEOS-Chem over this region (Wells et al., 2020).

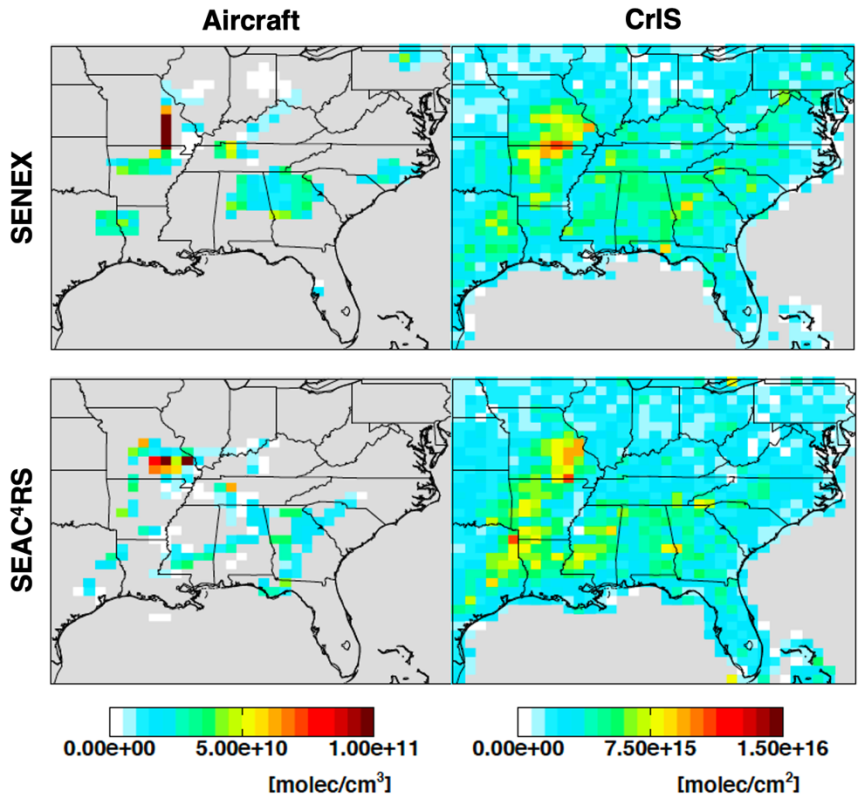


**Figure 7. Spatial distribution of isoprene over the Southeast US as measured from aircraft (left) and CrIS (right) during the SENEX (top, June-July 2013) and SEAC$^4$RS (bottom, August-September 2013) campaigns. Aircraft data are displayed as campaign mean (daytime flights only) density weighted boundary layer (P > 800 hPa) concentrations (molec cm$^{-3}$); CrIS data are plotted as column densities (molec cm$^{-2}$) averaged over the flight days of each campaign. Areas in gray have no available data.**

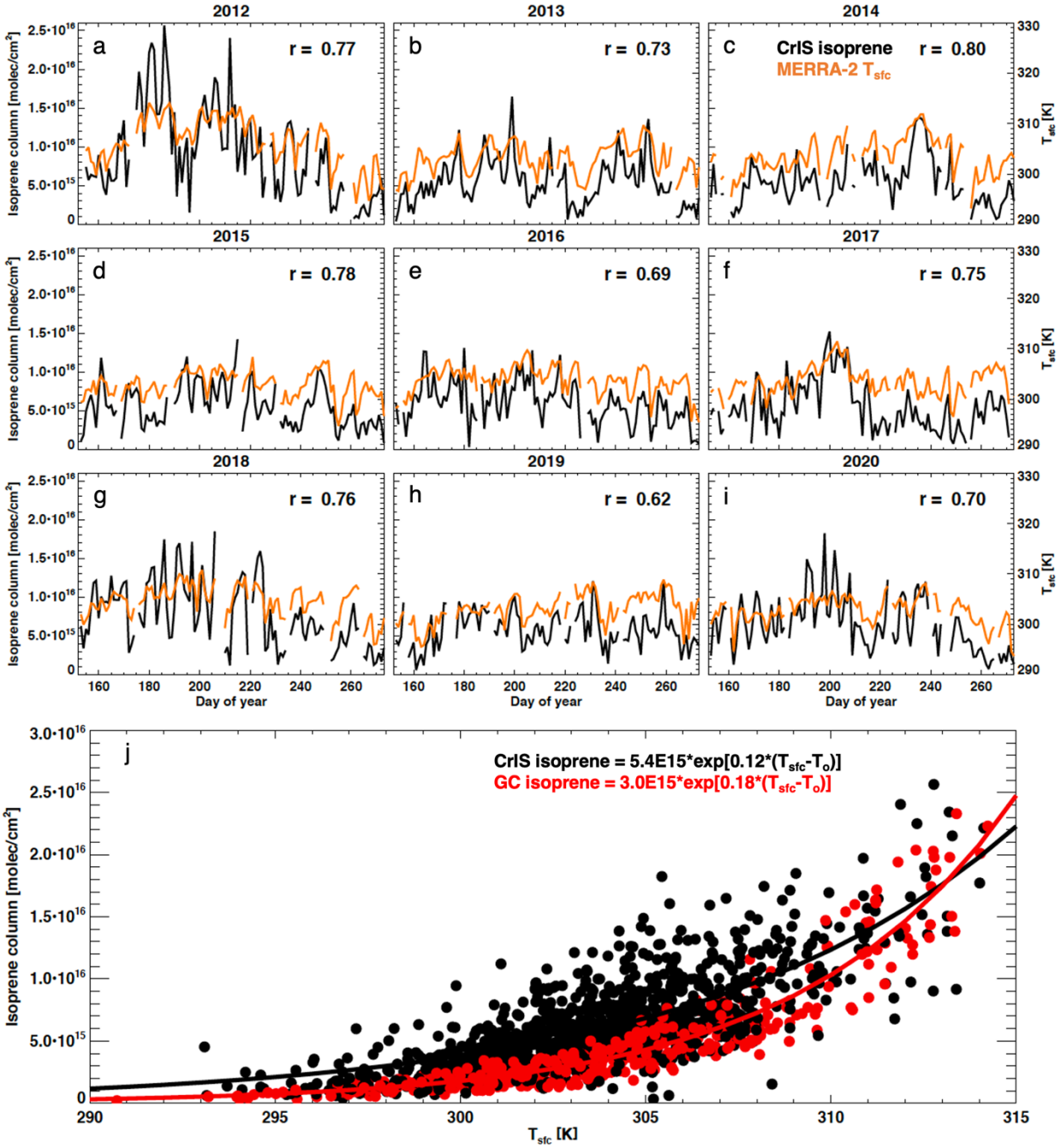


**Figure 8. Daily evolution of summertime (June-September) isoprene over the US Ozarks. (a-i) CrIS isoprene (black) and MERRA-2 surface air temperature ($T_{sfc}$, orange) timeseries averaged over the Ozarks region for each year from 2012-2020. (j) Daily CrIS (black) and GEOS-Chem (red) isoprene columns for the same region and timeframes plotted as a function of surface air temperature. An exponential fit to each data set is shown with $T_o$ = 330 K in each case.**

The strong temporal isoprene variability over the Ozarks that is revealed by CrIS in Fig. 8 is corroborated not only by the above temperature correlation but also by separate aircraft and ground-based observations. For example, Fig. S7 shows that day-to-day differences of

comparable magnitude are detected over this region from both space-based and airborne platforms. Furthermore, the highest CrIS isoprene columns across the entire Ozark data record occur in early July of 2012 (Fig. 8a), when anomalously large isoprene fluxes were measured in situ (Seco et al., 2015) during the early phase of an extreme drought.

### 4.2 Amazonia

Figure 9 shows CrIS isoprene columns over Amazonia averaged across the SNPP record for January, March, May, July, September, and November. Significant spatial variability is revealed by the high-resolution CrIS isoprene data over this region. The highest columns are observed over the western Brazilian state of Acre (on the border with Peru and Bolivia), followed by a widespread enhancement over the northern Amazon Basin (Brazilian states of Amazonas, Pará, and Roraima) and a more localized hotspot over Maranhão in eastern Brazil. Persistently low isoprene columns are detected along the Amazon mainstem and over other major regional floodplains (Fig. S8; Hamilton et al., 2002), consistent with the elevational and phenological gradients in isoprene emissions discussed by Gu et al. (2017).

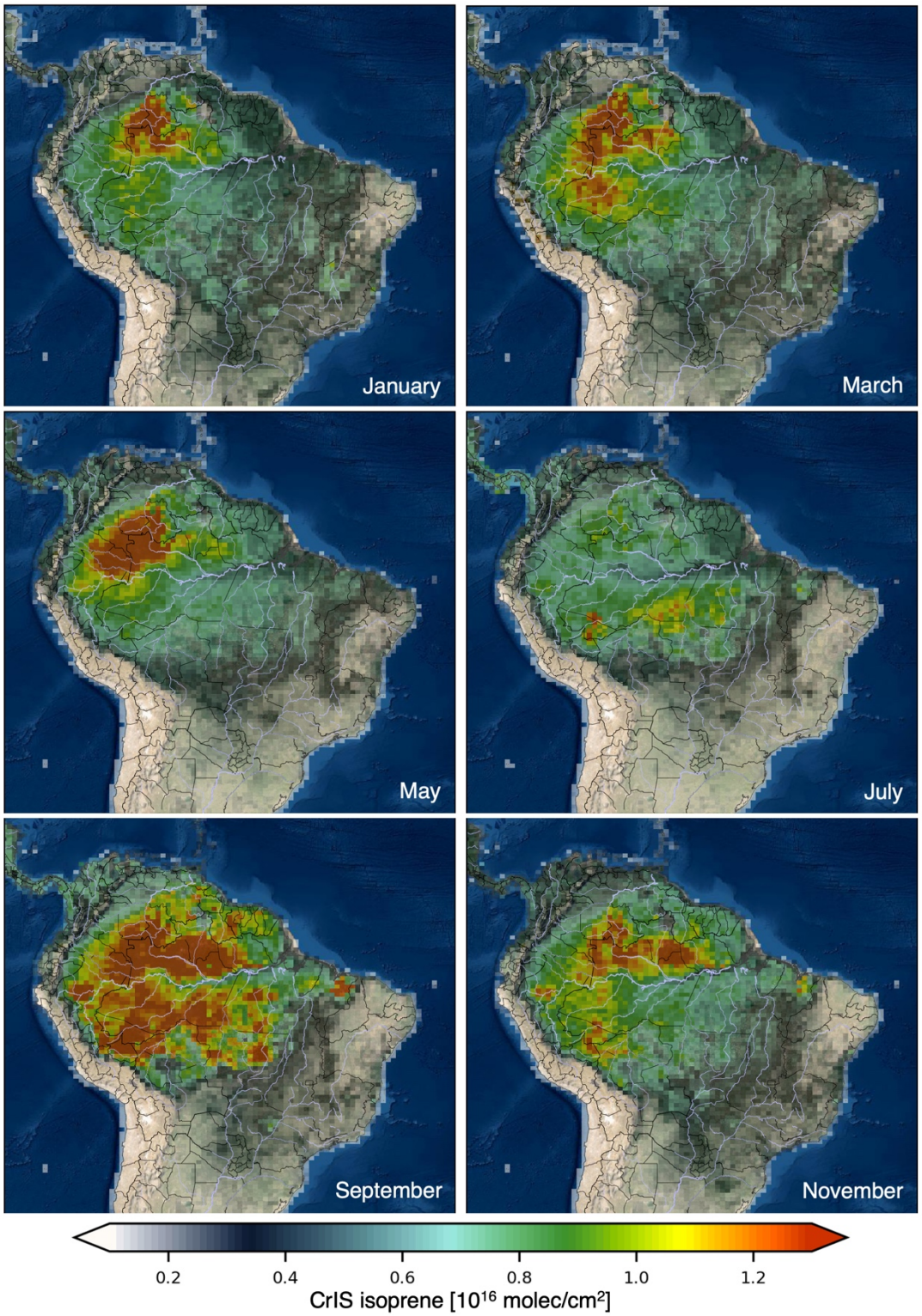


**Figure 9. Distribution of gridded (0.5° × 0.625°) isoprene column densities over Amazonia as measured from SNPP CrIS, averaged from 2012-2020 by month for January, March, May, July, September, and November.**

Strong seasonal shifts are seen in the location and magnitude of the above isoprene enhancements. In particular, a widespread maximum is detected during the dry season

458 (September-October) when temperatures and leaf area are highest (Wei et al., 2018). A more
459 localized maxima then emerges over the northwest Amazon basin during the wet season
460 (January-May), with a minimum during the transition between these two periods (June-July)
461 when leaf flushing is believed to cause a large-scale shutdown of isoprene emissions (Barkley et
462 al., 2009).

463 As was the case over the southeastern US, we also observe significant day-to-day isoprene
464 variability over Amazonia. Figure 10a shows an example isoprene timeline for year-2013 over
465 the state of Acre in western Brazil, where the highest columns in the entire basin are often
466 detected. Isoprene enhancements of similar magnitude occur during both the wet and dry
467 seasons—despite the much higher temperatures, and thus presumably higher emissions, in the
468 dry season (Fig. 10c). Further, the daily isoprene columns have a robust temperature correlation
469 during the dry season ($r = 0.58$, Fig. 10c) but not during the wet season ($r = 0.11$, Fig. 10b) when
470 the temperature range is small.

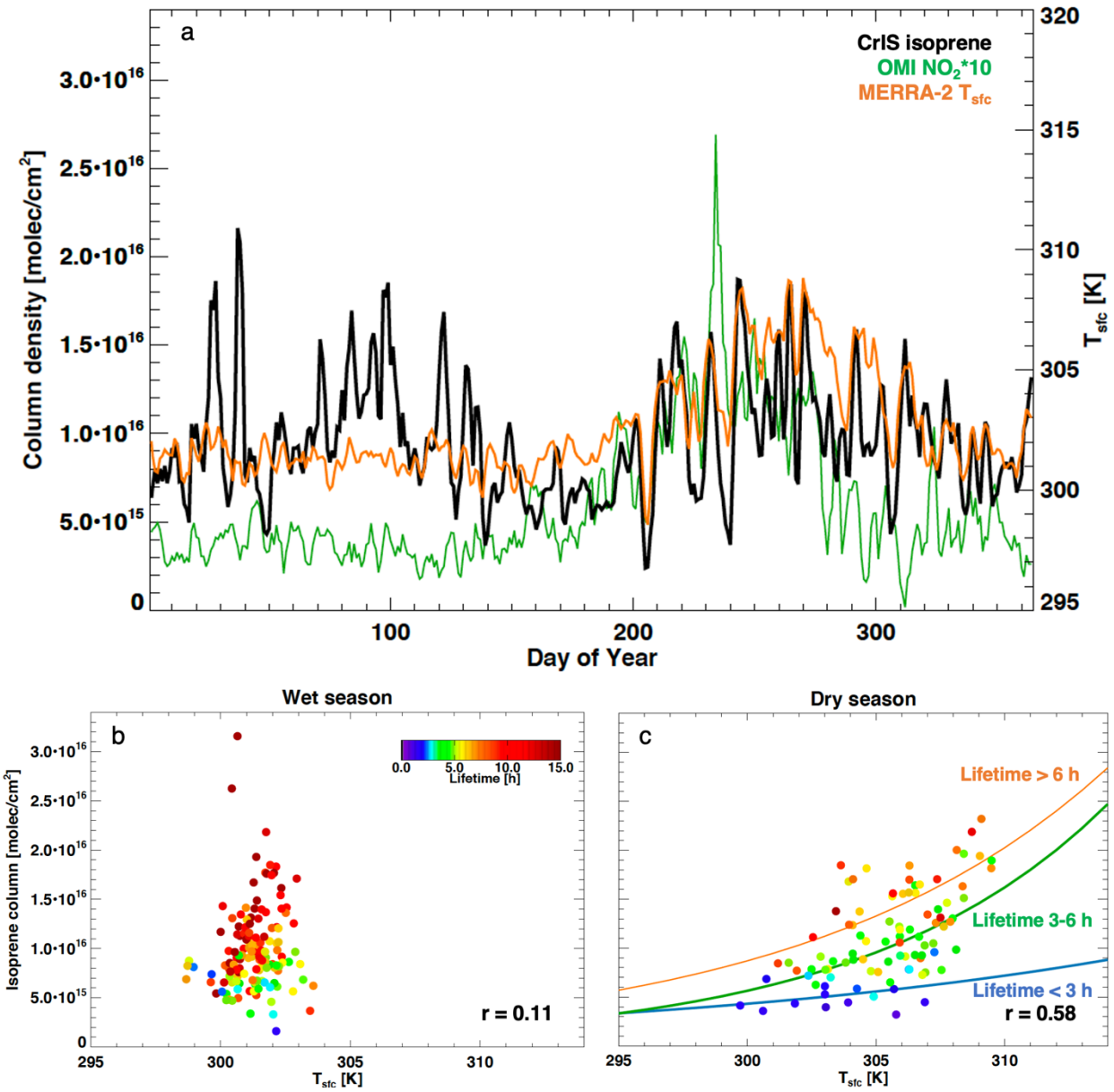


**Figure 10. Seasonal variability in atmospheric isoprene over the state of Acre in western Brazil as seen from CrIS. (a) Temporal evolution of CrIS isoprene (black), OMI tropospheric $NO_2$ (×10, green; Boersma et al., 2017), and MERRA-2 surface temperature (orange) averaged over the region during year-2013. (b) CrIS isoprene columns plotted as a function of MERRA-2 surface temperature and shaded by lifetime for the Amazonian wet season (January-May) and (c) dry season (August-October). Isoprene lifetimes are derived from the isoprene:HCHO column ratios as described by Wells et al. (2020) using HCHO data from OMI (De Smedt et al., 2017). Exponential fits to the dry season data are shown for isoprene lifetimes < 3 h (blue), 3-6 h (green), and > 6 h (orange). The timelines are displayed as a three-day running mean; scatter plots show daily data.**

The above patterns arise because of seasonal $NO_x$-driven differences in the isoprene lifetime. Wells et al. (2020) showed previously that the isoprene lifetime can be directly estimated from the satellite-measured isoprene:formaldehyde column ratio. Applying the same approach here,

484 we see that in the wet season, when $NO_x$ is low (< $5 \times 10^{14}$ molec $cm^{-2}$, Fig. 10a), isoprene
485 lifetimes over Acre are far longer (often 10-20+ h, Fig. 10b) than they are in the dry season
486 (typically < 6 h, Fig. 10c) when $NO_x$ is elevated (> $1 \times 10^{15}$ molec $cm^{-2}$, Fig. 10a). This
487 suppressed OH then leads to runaway wet season concentrations in spite of the relatively low
488 temperatures. Furthermore, we find during the dry season that the isoprene column:temperature
489 relationship varies in a coherent way with the isoprene lifetime: when lifetimes are longer the
490 temperature dependence is steeper, and vice versa (Fig. 10c). This coherence further
491 demonstrates the fidelity of the daily CrIS isoprene measurements.

492 **5 Conclusions**

493 We have described the next-generation ROCR isoprene retrieval, and applied it to obtain daily
494 global isoprene distributions from the CrIS satellite sensor for 2012-2020. The retrieval employs
495 the HRI as isoprene spectral index in a machine learning framework, building on previous work
496 for other VOCs with the IASI sensor (Franco et al., 2018; Franco et al., 2019; Franco et al.,
497 2020). We show that the ROCR isoprene retrieval provides enhanced sensitivity over our
498 previous method while maintaining computational efficiency to fully exploit the dense global
499 sampling of the CrIS instruments. Over Amazonia, the spatial distribution of the new isoprene
500 retrievals is highly consistent with prior OE and $\Delta T_b$-based results, with daily variability that
501 compares well with ground-based column observations. The derived CrIS isoprene columns are
502 lower in magnitude (20-40%) than earlier retrieval versions but higher (up to 50% for daily data)
503 than the ground-based results, likely reflecting different sources of systematic uncertainty in the
504 various approaches. A series of sensitivity tests identifies the background definition and isoprene
505 vertical profile as the most relevant uncertainty sources in the ROCR retrieval. Independent
506 validation continues to be a critical need for robust interpretation of the isoprene column
507 abundances in terms of controlling process.

508 The long-term global CrIS dataset presented here has unprecedented resolution, revealing strong
509 seasonal and spatial variability in atmospheric isoprene. Clear seasonal cycles are seen over the
510 world's isoprene hotspots, including East Asia and India where sensitivity was limited with our
511 previous approach. Over the US and Mexico, isoprene enhancements are observed across a range
512 of ecosystems and in particular over the "isoprene volcano" in the Ozarks. There, the patterns of
513 variability detected by CrIS are supported by independent data, reflecting emission drivers on

daily (e.g., temperature) to interannual (e.g., drought stress) timescales. Over Amazonia, strong spatial gradients are observed that vary temporally according to seasonal phenology of the underlying landscape features. Combining the CrIS data with space-based formaldehyde and $NO_2$ measurements reveals daily lifetime variations that elucidate the dual controls of emissions and chemistry on the isoprene abundance.

Overall, results presented here demonstrate the ability of these high-resolution CrIS retrievals to resolve daily isoprene variability from space, thus enabling new investigations into emission processes across a range of plant functional types. Our initial analyses highlighted some of the emergent connections between isoprene abundance and daily-to-interannual variations in temperature, $NO_x$, and drought stress. Because the CrIS record is planned to extend through at least 2030, the ROCR isoprene retrievals will provide valuable long-term information for diagnosing ecosystem variability and the links between surface-atmosphere exchange, climate, and chemistry in biogenic source regions.

## Acknowledgments

This work was supported by the NASA Aura Science Team (Grant Number 80NSSC20K0927) and the Minnesota Supercomputing Institute. M. Alvarado, K. Cady-Pereira, D. Gombos, J. Hegarty, and I. Strickland are acknowledged for generating and testing isoprene absorption look-up tables employed here. G. Toon is acknowledged for providing the pseudo-lines used in the ground-based FTIR isoprene retrievals at Porto Velho. D. Fu is acknowledged for providing OE isoprene retrievals over Amazonia. Isoprene measurements aboard the NASA DC-8 during SEAC$^4$RS were supported by the Austrian Federal Ministry for Transport, Innovation and Technology (bmvit) through the Austrian Space Applications Programme (ASAP) of the Austrian Research Promotion Agency (FFG). T. Mikoviny is acknowledged for his support during SEAC$^4$RS. Part of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract to NASA.

## Open Research

The CrIS Level 1B data used in this work are publicly available at https://snpp-sounder.gesdisc.eosdis.nasa.gov/data/SNPP_Sounder_Level1/SNPPCrISL1BNSR.2/. The ROCR isoprene column data analyzed here are available at http://z.umn.edu/ROCR_isoprene. Employed

airborne data is publicly available at http://esrl.noaa.gov/csd/projects/senex/ (SENEX) and at http://www-air.larc.nasa.gov/missions/seac4rs/index.html (SEAC$^4$RS). GEOS-Chem model code is available at www.geos-chem.org.

## References

Alves, E. G., Tota, J., Turnipseed, A., Guenther, A. B., Bustillos, J., Santana, R. A., et al. (2018). Leaf phenology as one important driver of seasonal changes in isoprene emissions in central Amazonia. *Biogeosciences*, 15, 4019-4032, https://doi.org/10.5194/bg-15-4019-2018.

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

Barkley, M. P., Palmer, P. I., De Smedt, I., Karl, T., Guenther, A., & Van Roozendael, M. (2009). Regulated large-scale annual shutdown of Amazonian isoprene emissions? Geophysical Research Letters, 36, https://doi.org/10.1029/2008gl036843.

Bates, K. H. & Jacob, D. J. (2019). A new model mechanism for atmospheric oxidation of isoprene: global effects on oxidants, nitrogen oxides, organic products, and secondary organic aerosol. *Atmospheric Chemistry and Physics*, 19, 9613-9640, https://doi.org/10.5194/acp-2019-328.

Batista, C. E., Ye, J. H., Ribeiro, I. O., Guimaraes, P. C., Medeiros, A. S. S., Barbosa, R. G., et al. (2019). Intermediate-scale horizontal isoprene concentrations in the near-canopy forest atmosphere and implications for emission heterogeneity. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 19318-19323, https://doi.org/10.1073/pnas.1904154116, 2019.

Boersma, K. F., Eskes, H., Richter, A., De Smedt, I., Lorente, A., Beirle, S., et al. (2017) QA4ECV $NO_2$ tropospheric and stratospheric vertical column data from OMI (Version 1.1). Royal Netherlands Meteorological Institute (KNMI). https://doi.org/10.21944/qa4ecv-no2-omi-v1.1.

Brauer, C. S., Blake, T. A., Guenther, A. B., Sharpe, S. W., Sams, R. L., & Johnson, T. J. (2014). Quantitative infrared absorption cross sections of isoprene for atmospheric measurements. *Atmospheric Measurement Techniques*, 7, 3839-3847, https://doi.org/10.5194/amt-7-3839-2014.

Clough, S. A., Shephard, M. W., Worden, J., Brown, P. D., Worden, H. M., Luo, M., et al. (2006). Forward model and Jacobians for Tropospheric Emission Spectrometer retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 1308-1323, https://doi.org/10.1109/tgrs.2005.860986.

De Smedt, I., Yu, H., Richter, A., Beirle, S., Eskes, H., Boersma, K. F., et al. (2017). QA4ECV HCHO tropospheric column data from OMI (Version 1.1). https://doi.org/10.18758/71021031.

Feiner, P. A., Brune, W. H., Miller, D. O., Zhang, L., Cohen, R. C., Romer, P. S., et al. (2016). Testing atmospheric oxidation in an Alabama forest. *Journal of the Atmospheric Sciences*, 73, 4699-4710, https://doi.org/10.1175/jas-d-16-0044.1.

Franco, B., Clarisse, L., Stavrakou, T., Muller, J. F., Pozzer, A., Hadji-Lazaro, J., et al. (2019). Acetone atmospheric distribution retrieved from space. *Geophysical Research Letters*, 46, 2884-2893, https://doi.org/10.1029/2019gl082052.

Franco, B., Clarisse, L., Stavrakou, T., Muller, J. F., Van Damme, M., Whitburn, S., et al. (2018). A general framework for global retrievals of trace gases From IASI: Application to methanol, formic acid, and PAN. *Journal of Geophysical Research-Atmospheres*, 123, 13963-13984, https://doi.org/10.1029/2018jd029633.

Franco, B., Clarisse, L., Stavrakou, T., Muller, J. F., Taraborrelli, D., Hadii-Lazaro, et al. (2020). Spaceborne measurements of formic and acetic acids: A global view of the regional sources. *Geophysical Research Letters*, 47, https://doi.org/10.1029/2019gl086239, 2020.

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

Fuchs, H., Hofzumahaus, A., Rohrer, F., Bohn, B., Brauers, T., Dorn, H. P., et al. (2013). Experimental evidence for efficient hydroxyl radical regeneration in isoprene oxidation. *Nature Geoscience*, 6, 1023-1026, https://doi.org/10.1038/ngeo1964.

Ganzeveld, L. N., Lelieveld, J., Dentener, F. J., Krol, M. C., & Roelofs, G. J. (2002). Atmosphere-biosphere trace gas exchanges simulated with a single-column model. *Journal of Geophysical Research-Atmospheres*, 107, https://doi.org/10.1029/2001jd000684.

Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, 30, 5419-5454, https://doi.org/10.1175/jcli-d-16-0758.1.

Gordon, I. E., Rothman, L. S., Hill, C., Kochanov, R. V., Tan, Y., Bernath, P. F., et al. (2017). The HITRAN2016 molecular spectroscopic database. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 203, 3-69, https://doi.org/10.1016/j.jqsrt.2017.06.038.

Gu, D., Guenther, A. B., Shilling, J. E., Yu, H. F., Huang, M. Y., Zhao, C., et al. (2017). Airborne observations reveal elevational gradient in tropical forest isoprene emissions. *Nature Communications*, 8, https://doi.org/10.1038/ncomms15541.

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

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., & Wang, X. (2012). The Model of Emissions of Gases and Aerosols from Nature version 2.1

(MEGAN2.1): an extended and updated framework for modeling biogenic emissions. *Geoscientific Model Development*, 5, 1471-1492, https://doi.org/10.5194/gmd-5-1471-2012.

Hamilton, S. K., Sippel, S. J., & Melack, J. M. (2002). Comparison of inundation patterns among major South American floodplains. *Journal of Geophysical Research-Atmospheres*, 107, https://doi.org/10.1029/2000jd000306.

Hase, F., Wallace, L. McLeod, S. D., Harrison, J. J., & Bernath, P. F. (2010). The ACE-FTS atlas of the infrared solar spectrum. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111, 521-528, https://doi.org/10.1016/j.jqrst.2009.10.020.

Kaiser, J., Jacob, D. J., Zhu, L., Travis, K. R., Fisher, J. A., Abad, G. G., et al. (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, 5483-5497, https://doi.org/10.5194/acp-18-5483-2018.

Lelieveld, J., Butler, T. M., Crowley, J. N., Dillon, T. J., Fischer, H., Ganzeveld, L., et al. (2008). Atmospheric oxidation capacity sustained by a tropical forest. *Nature*, 452, 737-740, https://doi.org/10.1038/nature06870.

Lin, J. T. & McElroy, M. B. (2010). Impacts of boundary layer mixing on pollutant vertical profiles in the lower troposphere: Implications to satellite remote sensing. *Atmospheric Environment*, 44, 1726-1739, https://doi.org/10.1016/j.atmosenv.2010.02.009.

Lin, Y. H., Zhang, H. F., Pye, H. O. T., Zhang, Z. F., Marth, W. J., Park, S., et al. (2013). Epoxide as a precursor to secondary organic aerosol formation from isoprene photooxidation in the presence of nitrogen oxides. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 6718-6723, https://doi.org/10.1073/pnas.1221150110.

Mao, J. Q., Paulot, F., Jacob, D. J., Cohen, R. C., Crounse, J. D., Wennberg, P. O., et al. (2013). Ozone and organic nitrates over the eastern United States: Sensitivity to isoprene chemistry. *Journal of Geophysical Research-Atmospheres*, 118, 11256-11268, https://doi.org/10.1002/jgrd.50817.

Messina, P., Lathiere, J., Sindelarova, K., Vuichard, N., Granier, C., Ghattas, J., et al. (2016). Global biogenic volatile organic compound emissions in the ORCHIDEE and MEGAN models and sensitivity to key parameters. *Atmospheric Chemistry and Physics*, 16, 14169-14202, https://doi.org/10.5194/acp-16-14169-2016.

Opacka, B., Muller, 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, 8413-8436, https://doi.org/10.5194/acp-21-8413-2021.

Paulot, F., Henze, D. K., & Wennberg, P. O. (2012). Impact of the isoprene photochemical cascade on tropical ozone. *Atmospheric Chemistry and Physics*, 12, 1307-1325, https://doi.org/10.5194/acp-12-1307-2012.

Paulot, F., Jacob, D. J., & Henze, D. K. (2013). Sources and processes contributing to nitrogen deposition: An adjoint model analysis applied to biodiversity hotspots worldwide. *Environmental Science & Technology*, 47, 3226-3233, https://doi.org/10.1021/es3027727.

Seco, R., Karl, T., Guenther, A., Hosman, K. P., Pallardy, S. G., Gu, L. H., et al. (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, 3657-3674, https://doi.org/10.1111/gcb.12980.

Sha, M. K., Langerock, B., Blavier, J.-F. L., Blumenstock, T., Borsdoff, T., Buschmann, M., et al. (2021). Validation of methane and carbon monoxide from Sentinal-5 Precursor using TCCON and NDACC-IRWG stations. *Atmospheric Measurement Techniques*, 14, 6249-6304, https://doi.org/10.5194/amt-14-6249-2021.

Tikhonov, A. N. (1963). On the solution of incorrectly posed problem and the method of regularization. *Soviet Mathematics*, 4, 1035-1038.

Toon, O. B., Maring, H., Dibb, J., Ferrare, R., Jacob, D. J., Jensen, E. J., et al. (2016). Planning, implementation, and scientific goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC$^4$RS) field mission. *Journal of Geophysical Research-Atmospheres*, 121, 4967-5009, https://doi.org/10.1002/2015jd024297.

Vigouroux, C., Hendrick, F., Stavrakou, T., Dils, B., De Smedt, I., Hermans, C., et al. (2009). Ground-based FTIR and MAX-DOAS observations of formaldehyde at Reunion Island and comparisons with satellite and model data. *Atmospheric Chemistry and Physics*, 9, 9523-9544, https://doi.org/10.5194/acp-9-9523-2009.

Vigouroux, C., Stavrakou, T., Whaley, C., Dils, B., Duflot, V., Hermans, C., et al. (2012). FTIR time-series of biomass burning products (HCN, $C_2H_6$, $C_2H_2$, $CH_3OH$, and HCOOH) at Reunion Island (21° S, 55° E) and comparisons with model data. *Atmospheric Chemistry and Physics*, 12, 10367-10385, https://doi.org/10.5194/acp-12-10367-2012.

Vigouroux, C., Aquino, C. A. B., Bauwens, M., Becker, C., Blumenstock, T., De Mazière, M., et al. (2018). NDAAC harmonized formaldehyde time series from 21 FTIR stations covering a wide range of column abundances. *Atmospheric Measurement Techniques*, 11, 5049-5073, https://doi.org/10.5194/amt-11-5049-2018.

Vigouroux, C., Langerock, B., Aquino, C. A. B., Blumenstock, T., Cheng, Z. B., De Mazière, M., et al. (2020). TROPOMI-Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations. *Atmospheric Measurement Techniques*, 13, 3751-3767, https://doi.org/10.5194/amt-13-3751-2020.

Walker, J. C., Dudhia, A., & Carboni, E. (2011). An effective method for the detection of trace species demonstrated using the MetOp Infrared Atmospheric Sounding Interferometer. *Atmospheric Measurement Techniques*, 4, 1567-1580, https://doi.org/10.5194/amt-4-1567-2011.

Warneke, C., Trainer, M., de Gouw, J. A., Parrish, D. D., Fahey, D. W., Ravishankara, A. R., et al. (2016). Instrumentation and measurement strategy for the NOAA SENEX aircraft campaign

as part of the Southeast Atmosphere Study 2013. *Atmospheric Measurement Techniques*, 9, 3063-3093, https://doi.org/10.5194/amt-9-3063-2016.

Wei, D. D., Fuentes, J. D., Gerken, T., Chamecki, M., Trowbridg, A. M., Stoy, P. C., et al. (2018). Environmental and biological controls on seasonal patterns of isoprene above a rain forest in central Amazonia. *Agricultural and Forest Meteorology*, 256, 391-406, https://doi.org/10.1016/j.agrformet.2018.03.024.

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

Whitburn, S., Van Damme, M., Clarisse, L., Bauduin, S., Heald, C. L., Hadji-Lazaro, J., et al. (2016). A flexible and robust neural network IASI-$NH_3$ retrieval algorithm. *Journal of Geophysical Research-Atmospheres,* 121, 6581-6599, https://doi.org/10.1002/2016jd024828.

Wiedinmyer, C., Greenberg, J., Guenther, A., Hopkins, B., Baker, K., Geron, C., et al. (2005). Ozarks Isoprene Experiment (OZIE): Measurements and modeling of the "isoprene volcano". *Journal of Geophysical Research-Atmospheres*, 110, https://doi.org/10.1029/2005jd005800.

Wu, S. L., Mickley, L. J., Jacob, D. J., Logan, J. A., Yantosca, R. M., & Rind, D. (2007). Why are there large differences between models in global budgets of tropospheric ozone? *Journal of Geophysical Research-Atmospheres*, 112, https://doi.org/10.1029/2006jd007801.

Zavyalov, V., Esplin, M., Scott, D., Esplin, B., Bingham, G., Hoffman, E., et al. (2013). Noise performance of the CrIS instrument. *Journal of Geophysical Research-Atmospheres*, 118, 13108-13120, https://doi.org/10.1002/2013jd020457.

713 **Table 1: Settings for the ground-based FTIR isoprene retrievals at Porto Velho.**

| | |
|---|---|
| Retrieval code | SFIT4v09.4.4; available at https://www2.acom.ucar.edu/irwg/ |
| Spectral window | 891.6-917.0 $cm^{-1}$ |
| De-weighted signatures ($H_2O$) | 896.4-896.6; 897.6-897.8; 902-8-903.2; 906.5-907.0; 907.7-908.9; 910.1-910.4; 910.6-910.8; 911.2-911.3; 914.5-915.0 $cm^{-1}$ |
| Retrieved species (**target** + interfering) | Profile retrieval: **$C_5H_8$**, $H_2O$, $NH_3$<br>Scaling of the a priori profile: $CO_2$, $HNO_3$, $H_2{}^{18}O$, $C_2H_4$, $CCl_2F_2$, F142b, solar CO |
| Spectroscopic parameters | $C_5H_8$: pseudo line list by G. Toon (JPL): https://mark4sun.jpl.nasa.gov/pseudo.html, constructed from the cross-sections of Brauer et al. (2014).<br>$H_2O$, $H_2{}^{18}O$: from Toth (2003), available at http://mark4sun.jpl.nasa.gov/data/spec/H2O/RAToth_H2O.tar<br>$NH_3$, $HNO_3$, $C_2H_4$: HITRAN 2012<br>$CO_2$: HITRAN 2008<br>$CCl_2F_2$, F142b: pseudo line list by G. Toon (JPL)<br>Solar CO: updated from Hase et al. (2010), see SFIT4v09.4.4 package. |
| A priori profiles of atmospheric species | $H_2O$, $H_2{}^{18}O$: 6-hourly profiles from NCEP<br>Other species: climatology from WACCM v4 model |
| P, T profiles | 6-hourly profiles from NCEP |
| Regularization | Tikhonov L1 regularization (Tikhonov, 1963; see also Vigouroux et al., 2012 for application to FTIR); $\alpha$-value = 50, 5, and 0.1 for $C_5H_8$, $H_2O$ and $NH_3$, respectively. |
| Uncertainties | Random: 7% ($3.4 \times 10^{14}$ molec/$cm^2$); Systematic: 21% |

714