Satellite isoprene retrievals constrain emissions and atmospheric oxidation

Isoprene is the dominant non-methane organic compound emitted to the atmosphere\textsuperscript{1-3}. It drives ozone and aerosol production, modulates atmospheric oxidation and interacts with the global nitrogen cycle\textsuperscript{4-8}. Isoprene emissions are highly uncertain\textsuperscript{1,9}, as is the nonlinear chemistry coupling isoprene and the hydroxyl radical, OH—its primary sink\textsuperscript{10-13}. Here we present global isoprene measurements taken from space using the Cross-track Infrared Sounder. Together with observations of formaldehyde, an isoprene oxidation product, these measurements provide constraints on isoprene emissions and atmospheric oxidation. We find that the isoprene–formaldehyde relationships measured from space are broadly consistent with the current understanding of isoprene–OH chemistry, with no indication of missing OH recycling at low nitrogen oxide concentrations. We analyse these datasets over four global isoprene hotspots in relation to model predictions, and present a quantification of isoprene emissions based directly on satellite measurements of isoprene itself. A major discrepancy emerges over Amazonia, where current underestimates of natural nitrogen oxide emissions bias modelled OH and hence isoprene. Over southern Africa, we find that a prominent isoprene hotspot is missing from bottom-up predictions. A multi-year analysis sheds light on interannual isoprene variability, and suggests the influence of the El Niño/Southern Oscillation.

Isoprene (2-methyl-1,3-butadiene), produced during photosynthetic metabolism and emitted mainly from the leaves of woody plants, has global emissions comparable to those of methane and considerably greater than the sum of anthropogenic volatile organic compounds (VOCs)\textsuperscript{1-3}. Isoprene is highly reactive (lifetime <1 h at [OH] = 5 × 10\textsuperscript{6} molecules per cm\textsuperscript{3}) and plays a pivotal role in atmospheric oxidation, ozone and aerosol formation\textsuperscript{4-8}. Air quality and chemistry–climate models thus require accurate isoprene emission inputs; however, current estimates span a wide range (~210–990 Tg C yr\textsuperscript{-1} globally\textsuperscript{1,9}). The degree to which isoprene oxidation occurs at low nitrogen oxide (NO\textsubscript{x}) levels depletes versus sustains the abundance of hydroxyl radicals (OH)—the principal atmospheric oxidant\textsuperscript{10-13}.—is also uncertain. Space-borne measurements of formaldehyde (HCHO, an isoprene oxidation product) can provide top-down constraints\textsuperscript{4,5}, but alone its use as an isoprene proxy is hampered by uncertainties in the NO\textsubscript{x}-dependent chemistry governing the formaldehyde production yield and timescale\textsuperscript{6,9}, and by competing non-isoprene formaldehyde sources\textsuperscript{14-18}. Fu et al.\textsuperscript{19} recently demonstrated the viability of direct space-borne isoprene retrievals using infrared (IR) radiance measurements from the Cross-track Infrared Sounder (CrIS). That study employed optimal estimation to retrieve isoprene column abundances (\(\Omega\text{_{isoprene}}\); see Supplementary Note I and Supplementary Fig. I) over Amazonia, with results validated using aircraft measurements. Here, we build on that work to develop an artificial neural network (ANN)-based algorithm for deriving global isoprene columns from the CrIS measurements. The computational efficiency of the ANN allows fuller exploitation of the dense CrIS sampling (~9 × 10\textsuperscript{6} spectra per day) for understanding spatial and temporal drivers of atmospheric isoprene. We thus derive global observations of atmospheric isoprene from space, and use this dataset to evaluate current understanding of its emissions and atmospheric oxidation.

Isoprene spectral index

As described in Methods, we use the CrIS-measured brightness temperature difference (\(\Delta T\nu\)) between the peak of the \(\nu\text{\textsubscript{isoprene}}\) band\textsuperscript{20} and nearby off-peak channels (see Extended Data Fig. 1a) as a spectral index for deriving isoprene column abundances from the satellite data. Analogous methodologies have been used successfully for a variety of other atmospheric species\textsuperscript{21-24}. Extended Data Fig. 1b shows the \(\Delta T\nu\)—isoprene relationship as simulated by a forward radiative transfer model for diverse conditions spanning the global atmosphere over land (Methods). The relationship is approximately linear with slope varying as a function of thermal contrast (atmosphere–surface temperature difference; see Methods and Extended Data Fig. 1c). Interfering species likewise play a role and need to be accounted for, as discussed later.

**Feferal Note 1**

E-mail: dbm@umn.edu

---

*Department of Soil, Water, and Climate, University of Minnesota, St Paul, MN, USA. 1Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. 2School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. 3Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA. 4Department of Chemistry, University of Colorado, Boulder, CO, USA. 5Department of Atmospheric and Cryogenic Sciences, University of Innsbruck, Innsbruck, Austria. 6Institute for Ion Physics and Applied Physics, University of Innsbruck, Innsbruck, Austria. 7NOAA Chemical Sciences Laboratory, Boulder, CO, USA. 8Institute for Ion Physics and Applied Physics, University of Innsbruck, Innsbruck, Austria. 9Department of Chemistry, University of Oslo, Oslo, Norway. 10Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, PA, USA. 11Present address: Bioclimatology, University of Göttingen, Göttingen, Germany. 12e-mail: dbm@umn.edu**
Figure 1 maps the global and seasonal $\Delta T_b$ distribution measured by CrIS. Clear enhancements are seen over many predicted isoprene source regions: Amazonia, northern Australia (January), central Africa (April) and the southeast United States (July). However, $\Delta T_b$ enhancements also manifest over regions not predicted by the GEOS-Chem chemical transport model (CTM; Methods) to have large isoprene sources (for example, equatorial eastern Africa and the Arabian Peninsula, Pakistan and the southwest United States in July, Angola/Zambia in January and April). Elevated $\Delta T_b$ values also occur across the tropics, with a spatial distribution resembling that of water vapour. As will be seen, $\Delta T_b$ enhancements not associated with high modelled isoprene can reveal locations where emissions are much higher than presently thought—many parts of the world lack flux measurements for regionally important plant species. However, we show later that the rest of these anomalous features disappear once thermal contrast, water vapour and related factors are properly accounted for via the ANN.

ANN-based isoprene measurements

We use a supervised feed-forward ANN to derive isoprene columns from the CrIS $\Delta T_b$ data and contemporaneous observations of relevant surface and atmospheric properties (Methods). The ANN used (representing the mean of 10 networks) reproduces 93% of the isoprene column variance across the full training dataset. Prediction uncertainty is typically <30% for elevated isoprene columns ($>1 \times 10^{16}$ molecules per cm$^2$), increasing to 50% or more for low isoprene amounts/low thermal contrast.

We apply the trained ANN to the space-borne CrIS $\Delta T_b$ measurements to derive global isoprene distributions for January, April, July and October 2013 (Methods). Because the statistical performance of the ANN summarized above does not necessarily represent the full observational uncertainty, we further evaluate our results with the previously validated optimal estimation retrievals and with independent aircraft measurements from two campaigns over the southeast United States.

Figure 2a compares the ANN and optimal estimation isoprene measurements over Amazonia for September 2014, revealing strong agreement between the two (correlation $r = 0.9$, slope $m = 0.8$). Furthermore, Fig. 2b–d shows that the aircraft–model comparisons (see Methods) yield slopes ($m = 1.2–1.3$) and correlations ($r = 0.5–0.7$) that statistically match the CrIS–model comparison ($m = 1.3$, $r = 0.6$), thus providing indirect validation of the CrIS data. The aircraft measurements also reveal key spatial features that are consistent with CrIS but not captured by
Fig. 2 | Comparison of the CrIS ANN isoprene columns with other datasets. 

a, Comparison of ANN- and optimal estimation (OE)-derived isoprene estimates. Both are derived from cloud-screened CrIS radiance data for September 2014; ANN results employ GEOS-Chem HNO₃ as CrIS HNO₃ data were unavailable for this timeframe. The maps (left) display columns normalized to their domain means, and the scatterplot (right) compares the absolute columns (absolute columns are mapped in Extended Data Fig. 8).

b–d, Evaluation of CrIS ANN isoprene measurements using aircraft observations and GEOS-Chem model output. b, Monthly mean July 2013 isoprene columns as measured by CrIS (~13:30 LT) and simulated by GEOS-Chem (12:00–15:00 LT mean). c, d, Ambient isoprene concentrations as measured during the SENEX (June – July 2013; c) and SEAC4RS (August – September 2013; d) aircraft campaigns and simulated by GEOS-Chem along the flight tracks. Data are plotted as campaign-average density-weighted boundary layer number densities (P > 800 hPa). The error bars in the scatter plots for a and b indicate the standard deviation across the 10 ANN-based columns (see Methods; in some cases error bars are smaller than the data points), the red dashed lines indicate the range in slopes across ANNs and black dashed lines indicate the 1:1 relation. Stated slope uncertainties and grey shaded regions represent the bootstrapped standard error of regression.
emissions, with the slope corresponding to the isoprene lifetime. At lower NOx, the isoprene columns increase superlinearly with emissions. In this regime (occurring in the model most notably over Amazonia), elevated isoprene suppresses OH and therefore its own sink, leading to runaway concentrations.

As an isoprene oxidation product, formaldehyde is more buffered with respect to OH variability: (1) photolysis ensures that HCHO removal continues even at low OH levels, and (2) its production is proportional to isoprene + OH, which is more stable than either quantity alone when elevated isoprene suppresses OH. Because of these differing sensitivities, the isoprene:HCHO column ratio is a proxy for the atmosphere’s oxidizing capacity over isoprene source regions. Figure 3b illustrates this relationship: on a global basis, across all locations and seasons, the monthly mean 13:30 LT Ω\text{isoprene}/\Omega\text{HCHO} ratios simulated by GEOS-Chem scale tightly with 1/[OH] (r = 0.94; Supplementary Note 2 discusses the factors driving this relationship). A sensitivity analysis using an alternate isoprene oxidation mechanism (Mini-CIM; see Methods) yields a similarly strong correlation (Supplementary Fig. 2), with details presented in Supplementary Note 3.

The strong correlation in Fig. 3b encompasses the full global range of chemical regimes for isoprene oxidation: from unpolluted situations where isoprene-derived peroxy radicals (RO2) are long lived and react mainly with hydroperoxyl radicals (HO2) and other RO2, or isomerize, to polluted areas where isoprene-derived RO2 react quickly with NO (refs. 10,12,32). This globally aggregated Ω\text{isoprene}/\Omega\text{HCHO} versus 1/[OH] slope is weighted to isoprene-rich, OH-poor conditions: Supplementary Note 3 shows that the modelled slope varies across our analysis regions from 0.18 to 0.49. A sensitivity study with the independent Mini-CIM mechanism further shows systematic adjustments of 28–56% depending on location (Supplementary Fig. 4); factors such as non-isoprene biogenic VOC emissions and model mixing assumptions (which influence the column-integrated OH—isoprene reaction rate) also influence the slope (Supplementary Note 3). Overall, however, results here clearly demonstrate that the Ω\text{isoprene}/\Omega\text{HCHO} ratio provides a strong proxy of atmospheric oxidation that is observable from space.

We can therefore derive new constraints on isoprene–OH chemistry globally by combining the CrIS isoprene measurements derived here with space-based HCHO columns from OMI (Ozone Monitoring Instrument; Methods). Specifically, we employ the measured isoprene:HCHO ratios from CrIS and OMI as a direct proxy of 1/[OH] (and hence the isoprene lifetime) that can be used to test chemical models. To that end, Fig. 4 plots the Ω\text{isoprene}/\Omega\text{HCHO} ratios measured by CrIS + OMI and simulated by GEOS-Chem. Data are shown as a function of isoprene and NO2 (from ref. 33) for months spanning all four seasons (January, April, July, and October), and confined to locations with elevated surface temperatures (>293K) to limit noise due to low isoprene/thermal contrast. In both satellite-based and modelled relationships, we see a low-OH (and long isoprene lifetime) regime when isoprene is elevated and NO2 is low, and an opposing higher-OH (short lifetime) regime when the reverse is true. These oxidative regimes, and the chemical transitions between them, are generally consistent between model and observations, with the corresponding Ω\text{isoprene}/\Omega\text{HCHO} ratios and thus OH agreeing to within 10–40% at low to moderate NO2 (≤1015 molecules per cm2). One clear discrepancy is that the model population of extremely high isoprene at extremely low NO2 is not seen in the data; as we will see, this primarily reflects model NO2 errors over Amazonia. Some disparities also emerge at elevated NO2; however, the observed values in this range are subject to greater error due to limited measurements and lower isoprene columns with more uncertainty (Extended Data Fig. 3).

The above comparison supports the current model treatment of OH chemistry in the presence of isoprene. In particular, it argues against any substantial missing OH recycling at low NO2 (refs. 10,12,32)—instead, the modelled OH levels are modestly higher than implied by the satellite data. A sensitivity analysis using the Mini-CIM isoprene oxidation
The CrIS isoprene columns over Amazonia reveal strong seasonal variability in both the magnitude and location of the isoprene maxima. For the months examined, observed columns in west Amazonia (Fig. 5b, Extended Data Fig. 5) are highest in October and April and lowest in July. This is consistent with local ground-based measurements during GoAmazon33, which exhibit a June minimum and increase nearly twofold from then to October (Fig. 5b). Wei et al.33 attribute this seasonal minimum to leaf-flushing between wet and dry seasons; other studies34,35 also infer low isoprene emissions during new leaf growth in June–July. This seasonality is not well represented in GEOS-Chem, which instead peaks in April and exhibits only a 5% July–October column increase.

Also apparent from Fig. 5a is that the regions with long isoprene lifetimes and low OH concentrations based on the $\frac{\Delta_{\text{isoprene}}}{\Delta_{\text{HCHO}}}$ ratio, with measurement–model differences used to inform present understanding of emissions and atmospheric NOx.

Figure 4 shows the global CrIS isoprene columns and corresponding GEOS-Chem predictions for January, April, July and October 2013. The CrIS data reveal a number of isoprene hotspots that are consistent with the known isoprene sources discussed earlier—in particular, Amazonia, Central Africa, Australia and the Ozarks of the southeast United States. These regions stand out because they combine strong emissions with a chemical regime where isoprene is sufficiently long-lived to be detectable from space (unlike, for example, China in July, with elevated emissions but shorter isoprene lifetimes; Extended Data Fig. 2). For the months shown, the Central Africa and southeast United States enhancements peak in April and July, respectively, consistent with model predictions.

These dominant isoprene features are robust across the suite of ANN predictions: the column standard deviation across networks is typically <25% in these regions (Methods; Extended Data Fig. 4). The anomalous $\Delta_{T_o}$ enhancements discussed earlier in the context of spectral interferences do not emerge as enhancements in the CrIS isoprene maps, showing that the ANN is effectively accounting for non-isoprene factors influencing $\Delta_{T_o}$. A notable feature not predicted by GEOS-Chem is the strong observed isoprene enhancement over southern Africa in January and, to a lesser degree, in April; this is explored later in the text.

The following sections examine each of the above hotspots in terms of their implications for present understanding of atmospheric isoprene. For each region, we apply the corresponding $\frac{\Delta_{\text{isoprene}}}{\Delta_{\text{HCHO}}}$ versus $1/\text{[OH]}$ relationship in Supplementary Fig. 3 as a transfer function to quantify OH and the isoprene lifetime from the measured isoprene:HCHO ratios. The same transfer function is likewise applied to the model ratios (in this way, all relative model–measurement lifetime discrepancies arise solely from the underlying isoprene and HCHO column data, and are unaffected by any transfer function uncertainty). We also use the satellite measurements to provide an initial quantification of isoprene (and NO2) emissions over the same global hotspots, as detailed in Supplementary Notes 5, 6. From this analysis, we identify and discuss emergent gaps in current bottom-up understanding of isoprene emissions. Results are summarized in Figs. 5, 6 and in Supplementary Figs. 6, 7, 12–17.
surface observations during GoAmazon38 (Supplementary Note 5). From in situ measurements, Liu et al.39 likewise infer a large near-surface NO\textsubscript{x} bias in GEOS-Chem predictions for this region, which they attribute to underestimated soil emissions. Our satellite-based optimization described in Supplementary Note 5 (Supplementary Figs. 9–11, Supplementary Table 1) leads to substantial Amazonian NO\textsubscript{x} emission increases that agree well with the findings of Liu et al.39. Supplementary Figure 11 further shows that our NO\textsubscript{x} optimization successfully reduces the large isoprene lifetime biases over Amazonia in the prior model—providing independent confirmation of the results and supporting this isoprene emission quantification using CrIS. We thus derive monthly Amazonian isoprene emissions that point to substantial and coherent spatial errors in the bottom-up inventory (details in Supplementary Note 5). Overall, these results highlight the critical need to better understand NO\textsubscript{x} sources for this part of the world, and to elucidate the mechanisms driving isoprene emission variability in the tropics.

**Africa.** Two African isoprene hotspots are observed by CrIS: one in central Africa in April and one in the Miombo and transitional woodlands of Angola peaking in January (Fig. 5c–d). Although GEOS-Chem captures the timing of the central African enhancement, the CrIS data show the predicted isoprene peak to be too strong and too far north—as found previously based on OMI HCHO41 (model predictions using Mini-CIM are similar; Supplementary Fig. 6).

The Miombo/Angola peak has not been previously identified to this extent, though elevated leaf-level isoprene fluxes have been observed...
in woody savannahs here\textsuperscript{42}. Furthermore, while the CrIS-observed hotspot is largely missing from MEGANv2.1 (ref.\textsuperscript{1}), it matches the location and season of highest emissions according to a regional inventory from Otter et al.\textsuperscript{43} that incorporates detailed local land-cover information. The enhancement occurs in a low-NO\textsubscript{x} (and therefore low OH) area, leading to large isoprene enhancements relative to the corresponding emissions and HCHO (Figs. 3a, 3c) and explaining why a correspondingly strong HCHO peak is not seen. The CrIS seasonality over southern Africa also compares well with Otter et al.\textsuperscript{44} (Fig. 5d, Extended Data Fig. 5), with a January maximum and July minimum. GEOS-Chem, conversely, peaks in April with isoprene columns 2–4 times lower than CrIS.

The total isoprene emissions inferred from CrIS over southern Africa are higher than the prior estimate during January and April (Supplementary Fig. 13), and imply an emission overestimate north of the Equator and underestimate to the south (particularly over Angola/Namibia). These emission adjustments broadly support previous HCHO-based findings\textsuperscript{44,41,44,45}. As described in Supplementary Note 5, our CrIS-derived isoprene emissions for all of sub-equatorial Africa are highly consistent with the Otter et al.\textsuperscript{45} estimates, but substantially higher than MEGANv2.1. Such large discrepancies reveal a need for further investigation of isoprene sources in this understudied region.

**Southeast United States.** CrIS isoprene columns over the southeast United States peak in July over the ‘isoprene volcano’ in Missouri/Arkansas, where surface mixing ratios up to 36 parts per billion (ppb) have been observed\textsuperscript{46}. The aircraft data shown in Fig. 2b–d corroborate the CrIS isoprene distribution over this region, and OMI HCHO columns (Extended Data Fig. 6) likewise peak over the same part of the Ozarks during this time.

The GEOS-Chem isoprene maximum is shifted southwards with lower column amounts than CrIS (Fig. 6a, Supplementary Fig. 7). Kaiser et al.\textsuperscript{47} emphasize the importance of correcting NO\textsubscript{x} biases when inferring isoprene emissions, and modelled NO\textsubscript{2} columns exhibit substantial, spatially varying biases over this region (Fig. 6a and Supplementary Fig. 8). The isoprene lifetime predicted by the standard model is ~2 times the satellite-inferred value over the southern portion of the domain (where model isoprene is biased high) and 30–50% too low over Missouri (where the model is too low). However, the model does capture the observed regional isoprene seasonality\textsuperscript{48} (Fig. 6b).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig6.png}
\caption{Seasonality of space-based isoprene over the southeast United States and Australia. a, c, CrIS and GEOS-Chem isoprene columns, OMI and GEOS-Chem tropospheric NO\textsubscript{2} columns and space-based and GEOS-Chem isoprene lifetimes for the southeast United States (a) and southeast Australia (e) during January, April, July and October 2013 (top to bottom). The CrIS isoprene and space-based isoprene lifetimes are shown for snow-free, above detection limit locations ($\Omega$\textsubscript{isoprene} and $\Omega$\textsubscript{HCHO} values $>2 \times 10^{16}$ molecules per cm$^2$). b, d, Regional mean CrIS (black; error bars indicate the range across ANN predictions) and GEOS-Chem (red) isoprene columns for the southeast United States (b; regions defined in Extended Data Fig. 5) and southeast Australia (d).}
\end{figure}
correcting the NO2 biases above, we derive from the CrIS data moderate downward isoprene emission adjustments over Louisiana, Mississippi and Alabama, offset by increases over Missouri, Illinois and eastern Texas (Supplementary Fig. 14).

**Australia.** CrIS isoprene columns over Australia are highest in the north during January and April, with smaller enhancements along portions of the eastern and southern coasts (Fig. 6c). The northern Australia hotspot matches the location and timing of peak OMI HCHO (Extended Data Fig. 6). GEOS-Chem does not capture the observed spatial distribution, instead predicting peak enhancements over eastern Australia in January and weaker enhancements to the north and south (Fig. 6c and Supplementary Fig. 7). As over the southeast United States, spatially varying NO2 biases are apparent and play a role in the above isoprene discrepancies.

Over southeastern Australia, the CrIS isoprene columns peak in January, with an ~25% decrease from January to April and a July minimum. GEOS-Chem predicts a much larger (~90%) January–April drop, with mean columns 40–95% lower than observed. In situ measurements from the Sydney Particle Study49 support the weaker seasonality seen by CrIS (Fig. 6d). The CrIS-based source optimization shows that this modest seasonality also manifests in the underlying isoprene emissions (Supplementary Fig. 15).

**Future outlook**

We presented a global picture of isoprene from space, derived from CrIS radiances using an ANN. The reliability of the CrIS measurements is supported by comparisons to aircraft data and to (previously validated) optimal estimation measurements. However, more extensive validation data are needed to better quantify uncertainties and refine the measurement approach presented here.

Combining the CrIS measurements with contemporaneous HCHO observations provides a new space-based constraint on isoprene lifetimes, OH levels and emissions. The satellite-derived isoprene/HCHO column ratios support current understanding of isoprene–OH chemistry as represented in GEOS-Chem. In particular, the satellite data provide no indication of substantial missing OH recycling under high-isoprene, low-NOx conditions. A comparison between measured and predicted isoprene columns over key hotspot regions elucidates spatial and temporal biases in modelled isoprene emissions and NOx, which highlight the need for better mechanistic understanding of the drivers of tropical isoprene and NOx sources.

Finally, this work lays a foundation for multi-year studies examining seasonal-to-interannual isoprene changes and their impacts on atmospheric chemistry. Supplementary Note 7 illustrates this potential by applying the CrIS ANN retrieval from 2012–2018 over Amazonia and southern Africa (Supplementary Fig. 18). Results show that the strong seasonal patterns discussed earlier persist from year to year, but also reveal interannual differences tied to temperature shifts and climate features such as El Niño. Future analyses of the full global CrIS isoprene record can therefore elucidate key drivers of interannual ecosystem variability, including drought and other disturbance, and the couplings between climate, ecosystems and atmospheric chemistry.

**Online content**

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41556-020-2664-3.

---

1. Guenther, A. B. et al. The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions. Geosci. Model Dev. 5, 1471–1492 (2012).

2. Saunois, M. et al. The global methane budget 2000–2012. Earth Syst. Sci. Data 8, 697–751 (2016).

3. Huang, G. L. et al. Speciation of anthropogenic emissions of non-methane volatile organic compounds: a global gridded data set for 1970–2012. Atmos. Chem. Phys. 17, 7683–7701 (2017).

4. Trainer, M. et al. Models and observations of the impact of natural hydrocarbons on rural ozone. Nature 329, 705–707 (1987).

5. Hewitt, C. N. et al. Ground-level ozone influence by circadian control of isoprene emissions. Nat. Geosci. 4, 671–674 (2011).

6. Mao, J. Q. et al. Ozone and organic nitrates over the eastern United States: sensitivity to isoprene chemistry. J. Geophys. Res. Atmos. 118, 11256–11268 (2013).

7. Lin, Y. H. et al. Epoxide as a precursor to secondary organic aerosol formation from isoprene photooxidation in the presence of nitrogen oxides. Proc. Natl Acad. Sci. USA 110, 6718–6723 (2013).

8. Bates, K. H. & Jacob, D. J. A new model mechanism for atmospheric oxidation of isoprene: global effects on oxidants, nitrogen oxides, organic products, and secondary organic aerosol. Atmos. Chem. Phys. 16, 9633–9640 (2016).

9. Arneth, A. et al. Global terrestrial isoprene emission models: sensitivity to variability in climate and vegetation. Atmos. Chem. Phys. 11, 8037–8052 (2011).

10. Lelieveld, J. et al. Atmospheric oxidation capacity sustained by a tropical forest. Nature 452, 737–740 (2008).

11. Fuchs, H. et al. Experimental evidence for efficient hydroxyl radical regeneration in isoprene oxidation. Nat. Geosci. 6, 1023–1026 (2013).

12. Feiner, P. A. et al. Testing atmospheric oxidation in an Alabama forest. J. Geophys. Res. Atmos. 121, 478–493 (2016).

13. Boeke, N. L. et al. Formaldehyde columns from the Ozone Monitoring Instrument: urban versus background levels and evaluation using aircraft data and a global model. J. Geophys. Res. Atmos. 116, D25S03 (2011).

14. Fu, D. et al. Direct retrieval of isoprene from satellite-based infrared measurements. Nat. Commun. 10, 3811 (2019).

15. Brauer, C. S. et al. Quantitative infrared absorption cross sections of isoprene for atmospheric measurements. Atmos. Meas. Tech. 7, 3839–3847 (2014).

16. Razavi, A. et al. Global distributions of methanol and formic acid retrieved for the first time from the IASI/MetOp thermal infrared sounder. Atmos. Chem. Phys. 11, 857–872 (2011).

17. Clarisse, L., Clerbaux, C., Dentener, F., Hurtmans, D. & Coheur, P. F. Global ammonia distribution derived from infrared satellite observations. Nat. Geosci. 2, 479–483 (2009).

18. Franco, B. et al. A general framework for global retrievals of trace gases from IASI: application to methanol, formic acid, and PAN. J. Geophys. Res. Atmos. 123, 13983–13984 (2018).

19. Whitburn, S. et al. A flexible and robust neural network IASI-NH, retrieval algorithm. J. Geophys. Res. Atmos. 121, 6581–6589 (2016).

20. Warnke, C. et al. Instrumentation and measurement strategy for the NOAA SEVENET aircraft campaign as part of the Southeast Atmosphere Study 2013. Atmos. Meas. Tech. 9, 3063–3069 (2016).

21. Toon, O. B. et al. Planning, implementation, and scientific goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) field mission. J. Geophys. Res. Atmos. 121, 4967–5009 (2016).

22. Xie, Y. et al. Understanding the impact of recent advances in isoprene photooxidation on simulations of regional air quality. Atmos. Chem. Phys. 13, 8439–8455 (2013).

23. Teng, A. P., Crouse, J. D. & Wennberg, P. O. Isoprene peroxy radical dynamics. J. Am. Chem. Soc. 139, 5367–5377 (2017).

24. Kim, S.-W., Barth, M. C. & Trainer, M. Impact of turbulent mixing on isoprene chemistry. Geophys. Res. Lett. 43, 7701–7708 (2016).

25. de Smedt, I. et al. Algorithm theoretical baseline for formaldehyde retrievals from SSP TROPOMI and from the QA4ECV project. Atmos. Meas. Techn. 11, 2395–2426 (2018).

26. Boersma, B. et al. Improving algorithms and uncertainty estimates for satellite NO2 retrievals: results from the quality assurance for the essential climate variables (QA4ECV) project. Atmos. Meas. Tech. 11, 6651–6678 (2018).

27. de Gouw, J. A. et al. Hydrocarbon removal in power plant plumes shows nitrogen oxide dependence of hydroxyl radicals. Geophys. Res. Lett. 46, 7752–7760 (2019).

28. Wei, D. D. et al. Environmental and biological controls on seasonal patterns of isoprene above a rain forest in central Amazonia. Agric. For. Meteorol. 256–257, 391–406 (2018).

29. Barley, M. P. et al. Regulated large-scale annual shutdown of Amazonian isoprene emissions? Geophys. Res. Lett. 36, L04803 (2009).

30. Silsers, B. et al. Leaf phenology as one important driver of seasonal changes in isoprene emissions in central Amazonia. Biogeosciences 15, 4019–4032 (2018).

31. Belmonte Rivas, M. et al. OM1 tropospheric NOx profiles from cloud-slicing: constraints on surface emissions, convective transport and lightning NOx. Atmos. Chem. Phys. 15, 13519–13553 (2015).
38. Martin, S. T. et al. Introduction: observations and modeling of the Green Ocean Amazon (GoAmazon2014/5). Atmos. Chem. Phys. 16, 4785–4797 (2016).
39. Liu, Y. et al. Isoprene photochemistry over the Amazon rainforest. Proc. Natl Acad. Sci. USA 113, 6125–6130 (2016).
40. Guenther, A. et al. Isoprene emission estimates and uncertainties for the Central African EXPRESSO study domain. J. Geophys. Res. Atmos. 104, 30625–30639 (1999).
41. Marais, E. A. et al. Isoprene emissions in Africa inferred from OMI observations of formaldehyde columns. Atmos. Chem. Phys. 12, 6219–6235 (2012).
42. Otter, L. B., Guenther, A. & Greenberg, J. Seasonal and spatial variations in biogenic hydrocarbon emissions from southern African savannas and woodlands. Atmos. Environ. 36, 4265–4275 (2002).
43. Otter, L. et al. Spatial and temporal variations in biogenic volatile organic compound emissions for Africa south of the equator. J. Geophys. Res. Atmos. 108, 8505 (2003).
44. Stavrakou, T. et al. Global emissions of non-methane hydrocarbons deduced from SCIAMACHY formaldehyde columns through 2003–2006. Atmos. Chem. Phys. 9, 3663–3679 (2009).
45. Marais, E. A. et al. Improved model of isoprene emissions in Africa using Ozone Monitoring Instrument (OMI) satellite observations of formaldehyde: implications for oxidants and particulate matter. Atmos. Chem. Phys. 14, 7693–7703 (2014).
46. Wiedinmyer, C. et al. Ozarks Isoprene Experiment (OZIE): measurements and modeling of the “isoprene volcano”. J. Geophys. Res. Atmos. 110, D18307 (2005).
47. Kaiser, J. et al. High-resolution inversion of OMI formaldehyde columns to quantify isoprene emission on ecosystem-relevant scales: application to the southeast US. Atmos. Chem. Phys. 18, 5483–5497 (2018).
48. Hansen, D. A. et al. The southeastern aerosol research and characterization study: Part 1-overview. J. Air Waste Manag. Assoc. 53, 1460–1471 (2003).
49. Emmerson, K. M. et al. Current estimates of biogenic emissions from eucalypts uncertain for southeast Australia. Atmos. Chem. Phys. 16, 6997–7011 (2016).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
© The Author(s), under exclusive licence to Springer Nature Limited 2020
Isoprene signal and brightness temperature difference

Isoprene has two IR absorption features ($\nu_2$ and $\nu_3$) in the vicinity of 900 cm$^{-1}$ that are associated with the wagging vibrational mode for each of the molecule’s $sp^2$-CH groups. Extended Data Fig. 1a illustrates the radiance signal arising from those absorption features, plotted as the simulated difference in brightness temperature between an atmosphere with and without isoprene, assuming an isoprene profile with 5 ppb in the boundary layer and the US Standard Atmosphere for interfering species. Fu et al. demonstrated previously that the $\nu_2$ and $\nu_3$ features shown in Extended Data Fig. 1a are detectable from individual CrIS spectra over high-isoprene regions.

We start here from single-footprint Level 2 CrIS radiances that have been subsetted (1 of each 3 × 3 pixel array; FOV 6), cloud screened and gridded to 0.5° latitude × 0.625° longitude. The $\Delta T_b$ values are then calculated as the difference between off-peak (mean of the spectral points at 894.375 and 895 cm$^{-1}$) and on-peak (mean of the spectral points at 893.125 and 893.75 cm$^{-1}$) $T_b$ values at the $\nu_3$ feature.

Cloud screening is based on the observed difference between the 900 cm$^{-1}$ brightness temperature and the surface skin temperature. We simulate this difference for clear-sky conditions as a function of water vapour column density (solid black line in Extended Data Fig. 7a) using the Line-by-Line Radiative Transfer Model and employ a conservative linear approximation (solid red line in Extended Data Fig. 7a) to screen the observations. Temperature and water vapour information is from MERRA-2 reanalysis and interpolated to the time of CrIS overpass. We find good spatial correspondence between the location of our cloud-screened pixels and cloud flags derived from other space-borne sensors such as VIIRS and MODIS.

Given the demonstrated importance of careful cloud screening for optimal estimation isoprene retrievals from CrIS, we test the sensitivity of our results to cloud effects by employing a less stringent (by 2 K) brightness temperature threshold (dashed red line in Extended Data Fig. 7a). The results of this test are summarized in Extended Data Fig. 7b, c, and show that the resulting $\Delta T_b$ and isoprene changes are generally less than 15%, and less than 5% for enhanced isoprene levels. This suggests that the uncertainty in results presented here is not dominated by cloud effects.

Extended Data Figure 1a shows that other atmospheric species (specifically water vapour, nitric acid, ammonia and CFC-12) also have absorption features in the vicinity of the $\nu_2$ and $\nu_3$ isoprene peaks. We specifically chose to use $\nu_3$ in computing $\Delta T_b$ as it is the stronger of the two bands and less subject to such interferences. Nevertheless, variability in these other atmospheric species (and in factors such as surface–atmosphere thermal contrast, surface elevation and satellite viewing angle) can still affect the $\Delta T_b$–isoprene relationship, and are therefore accounted for in the estimation process described in the following section.

While other biogenically derived VOCs with terminal $=\text{CH}_2$ groups may also absorb in the vicinity of the isoprene peaks, Fu et al. showed that the relevant primary biogenic species (including monoterpenes) with published absorption cross-sections have much weaker absorption signals (<0.01 K) than isoprene at $\nu_3$. Since we focus here on isoprene hotspots, we assume such effects to be minor for our analysis. Relevant absorption cross-sections for key non-HCHO isoprene oxidation products (methyl vinyl ketone, methacrolein, isoprene hydroxyhydroperoxides) have not been reported, but available analogues indicate that their
overpass with daily cloud screening. Correlations span $r = 0.43–0.72$.

clustering to ensure representative sampling across the full range of

tangent (sigmoid) and linear transfer functions, respectively. The train-

hidden layers and one (single-neuron) output layer using hyperbolic

performance for a three-layer model containing two (six- and three-neuron)

als, which incorporate varying amounts of prior information depending

spectroscopy. This is a key distinction from optimal estimation retriev-

associated factors that is mechanistically defined by the underlying

$\Delta T_b$ values thus spatially correlate with isoprene

predictions over known source regions to a degree commonly found

for model–measurement comparisons of isoprene itself.

**ANN training and forward prediction**

We describe here a supervised feed-forward (that is, non-cyclic) ANN to derive isoprene columns from the CrIS $\Delta T_b$ observations. The approach employs a multilayer perceptron with training via Levenberg–Marquardt backpropagation to account for the interfering effects mentioned above based on contemporaneous observations of other relevant surface and atmospheric properties.

Given a set of input variables $x$ (in our case, $\Delta T_b$ and related parameters summarized in Extended Data Table 2), an ANN can be used to approximate an output $f(x)$ (in our case, $\Omega_{\text{isoprene}}$) that depends on $x$ in an unknown and possibly nonlinear way. This approximation occurs via a transfer function, $Y(W, x)$, where $W$ represents the weights of the function $Y$.

The weights are determined here with a synthetic dataset, constructed based on a full year of simulated radiances (from the Earth Limb and Nadir Operational Retrieval (ELANOR) model), which also serves as the operational forward model for the Tropospheric Emission Spectrometer (TES). ELANOR model inputs include temperature and water vapour profiles (using assimilated meteorological data from NASA GMAO) and climatological non-isoprene trace gas profiles (from the MOZART CTM). Isoprene profiles are taken from daily mid-afternoon (12:00–15:00 LT) GEOS-Chem predictions with 100% (1σ) Gaussian noise applied. We then apply global sampling (afternoon overlap, following the along-track separation of measurements from the global sampling strategy of TES, land locations only) to arrive at a representa-
tive input dataset of appropriate size for ANN training. Finally, the resulting radiances are simulated (using temperature-dependent isoprene absorption look-up tables) for three satellite viewing angles (selected randomly for each location). The full synthetic dataset comprises ~165,000 simulated spectra, from which we compute $\Delta T_b$ as above.

We then train the ANN to predict isoprene column densities based on six predictors (each taken as a firm constraint): $\Delta T_b$, water vapour column density ($\Omega_{\text{H}_2\text{O}}$), column nitric acid density ($\Omega_{\text{HNO}_3}$), thermal contrast (taken as the difference between the surface skin and 2 m air temperatures), surface pressure and satellite viewing angle. Alternate ANNs accounting for other potential interferers (such as CFCs and ammonia) were tested but ultimately discarded as they contributed little additional power to the isoprene predictions. No location-specific information is included in the training; the network thus describes the general global relationship between $\Delta T_b$, isoprene columns and associated factors that is mechanistically defined by the underlying spectroscopy. This is a key distinction from optimal estimation retrieval-

s, which incorporate varying amounts of prior information depending on the location-specific sensitivity.

We assessed multiple network architectures and found the best performance for a three-layer model containing two (six- and three-neuron) hidden layers and one (single-neuron) output layer using hyperbolic tangent (sigmoid) and linear transfer functions, respectively. The training occurs on ten random extractions of the synthetic dataset (after clustering to ensure representative sampling across the full range of isoprene column densities), with each extraction subsetted for training (50%), validation (30%) and testing (20%). The validation subset is used to determine when training can cease, and the testing subset is used subsequently to independently confirm network performance. Output from the resulting ten networks are then averaged to provide the final ANN prediction.

Finally, we apply the trained ANN to the space-borne CrIS $\Delta T_b$ measurements to derive global isoprene distributions for January, April, July and October 2013. Temperature and water vapour data are taken from the MERRA-2 reanalysis and interpolated to the CrIS overpass time, whereas nitric acid column observations are from the CrIS CLIMCAPS product. All input variables are cloud-screened as described above before calculation of the gridded ($2^\circ 	imes 2.5^\circ$) 13:30 LT monthly mean.

Fewer than 1% of the employed input variables fall outside the range used for ANN training (none of which occur over isoprene source regions), confirming that our training set is well-generalized.

Unlike a conventional optimal estimation retrieval, the ANN-based approach does not provide an estimate of the measurement vertical sensitivity (that is, averaging kernel) and associated uncertainty for every individual location. However, the ANN training statistics provide a quantification of the overall network performance, and therefore of the expected uncertainties for isoprene column abundances inferred from CrIS data. We find here that the six-predictor ANN can reproduce 93% of the variance in the isoprene total columns across the full synthetic dataset (Extended Data Fig. 7d). The performance of each of the ten networks relative to the independent testing set is similar ($r^2 = 0.92–0.93$, slopes ~1.0). This explanatory skill is lost when $\Delta T_b$ is withheld from training ($r^2 = 0.28$; Extended Data Fig. 7e) — confirming that the predictive power of the ANN is driven by the isoprene spectral signal rather than by the ancillary variables.

The relative uncertainty of the ANN predictions varies as a function of both isoprene amount and thermal contrast (Extended Data Fig. 7f). For enhanced isoprene columns ($>1 \times 10^{17}$ molecules per cm$^2$) the prediction uncertainty is typically less than 30%, even with very low thermal contrast. Uncertainty increases for lower isoprene amounts, exceeding 50% for columns less than $2 \times 10^{16}$ molecules per cm$^2$, and for columns less than $5 \times 10^{15}$ molecules per cm$^2$ at low thermal contrast (0–5 K; Extended Data Fig. 1c shows thermal contrast maps for January, April, July and October). These can be considered limits of detection for the 13:30 LT monthly mean isoprene columns derived from CrIS.

The statistical performance of the ANN as summarized above does not necessarily represent the full uncertainty of the CrIS isoprene measurements, as other factors (for example, cross-section or radiative transfer errors, uncertainties in ancillary datasets used for water vapour, temperature and HNO$_3$, uncertainties in the vertical profiles of isoprene used to train the ANN, residual column impacts) may also contribute. We therefore evaluate the CrIS isoprene columns using (1) the previously published and validated optimal estimation retrievals and (2) independent atmospheric measurements, as described below and in the main text.

**CrIS evaluation via aircraft-model intercomparison**

Direct evaluation of the CrIS isoprene measurements is difficult due to lack of either ground-based isoprene column observations in isoprene hotspot regions, or a statistically sufficient ensemble of full airborne profiles over isoprene source regions at the satellite overpass time. Instead, we perform an indirect validation (Fig. 2b–d) using measurements from two aircraft campaigns over the southeast United States: SENEX (Southeast Nexus; 27 May–10 July 2013; ref. 25) and SEAC4RS (Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys; 1 August–23 September 2013; ref. 26). In each case, we employ the GEOS-Chem model as an intercomparison platform to quantify the level of consistency between CrIS and the in situ aircraft data. Since any model isoprene bias should manifest in a consistent way relative to independent observational datasets for
the same region and time period, the consistency between the CrIs/GEOS-Chem regression and the aircraft/GEOS-Chem regression reflects the agreement between the CrIs and in situ isoprene datasets.

To perform this intercomparison, we sample the model at the time and location of the aircraft measurements (which are restricted to ±2 h from the CrIs overpass time). Results discussed in the main text are aggregated to the model resolution and averaged vertically for each campaign by calculating a density-weighted mean boundary layer (pressure $P > 800$ hPa) number density for each latitude × longitude grid cell.

OMI HCHO and NO$_2$ data

We use here the Quality Assurance for Essential Climate Variables (QA4ECV) version 1.0 Level 2 HCHO product from the OMI satellite sensor. OMI is a near-ultraviolet–visible spectrometer onboard NASA's EOS Aura satellite, which has an Equator overpass time (13:40 LT) close to that of Suomi-NPP. The HCHO slant column density is determined via fitting of OMI radiances and subsequently converted to vertical column densities using a modelled shape factor. The Q4AECV retrieval uses a single, extended fitting interval (328.5–359.0 nm), whereas the precursor BIRA HCHO retrieval employed a smaller window with prebits for O$_2$–O$_2$ and BrO slant columns. Although the Q4AECV data have yet to be fully validated, recent work has demonstrated their improved performance over the earlier BIRA retrieval. Zhu et al. previously found the BIRA v14 HCHO retrieval to exhibit a 12% low bias (with use of an accurate shape factor) relative to aircraft measurements, and subsequent analysis has supported these findings. We find here that a global Q4AECV versus BIRA v14 comparison for the timeframe of our analysis yields a slope of 1.1–1.4 (0.9–1.8 over our targeted subregions), and we therefore do not apply any bias correction to the Q4AECV HCHO data. Repeating our analysis using instead the bias-corrected BIRA v14 data (Supplementary Figs. 20–22) leads to no substantive differences in our core results.

Standard data processing and screening procedures are followed. We restrict the data to solar zenith angle <70° and cloud fraction <0.4. The OMI data are then gridded to the 2° × 2.5° GEOS-Chem resolution. For all comparisons the model is sampled according to the OMI HCHO observation operator at the time and location of the satellite overpass. Tropospheric NO$_2$ column data are from the OMI Q4AECV v1.1 monthly NO$_2$ product. The Q4AECV retrieval employs updated NO$_2$ spectral fitting that accounts for liquid water absorption and includes an intensity offset correction. This improves the quality of the product, particularly over clear-sky ocean locations. OMI Q4AECV tropospheric NO$_2$ columns exhibited good agreement (bias = -2% and root-mean-square difference = 16%) with ground-based column measurements in China. Comparisons in this work are performed with respect to monthly mean GEOS-Chem tropospheric NO$_2$ columns sampled at the time of the satellite overpass, with no observation operator applied.

Data availability

The CrIs Level 1B data used in this work are publicly available at https://snpp.sounder.gesdisc.eosdis.nasa.gov/data/SNPP_Sounder_Level1/SNPPCRISLBNSR.1/. The isoprene column data employed in this work are available at https://doi.org/10.13020/v959-dr15. The airborne data are publicly available for SENEX at http://esrl.noaa.gov/csd/projects/senex/ and for SEAC4RS at http://www-air.larc.nasa.gov/missions/seac4rs/index.html. OMI Q4AECV HCHO and NO$_2$ data are publicly available at http://www.qa4ecv.eu/ecvs.

Code availability

GEOS-Chem model code is publicly available at http://www.geos-chem.org. The LBLRTM$^{70,82}$, which is used to calculate the molecular absorption look-up tables employed in ELANOR$^{45}$, is publicly available at http://rtweb.aer.com/lblrtm.html.
86. Emmons, L. K. et al. Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4). Geosci. Model Dev. 3, 43–67 (2010).
87. Beer, R. TES on the Aura mission: scientific objectives, measurements, and analysis overview. IEEE Trans. Geosci. Remote Sens. 44, 1102–1105 (2006).
88. Smith, N. & Barnet, C. D. Uncertainty characterization and propagation in the Community Long-Term Infrared Microwave Combined Atmospheric Product System (CLIMCAPS). Remote Sens. 11, 1227 (2019).
89. Wells, K. C. et al. Tropospheric methanol observations from space: retrieval evaluation and constraints on the seasonality of biogenic emissions. Atmos. Chem. Phys. 12, 5897–5912 (2012).
90. Chaliyakunnel, S., Millet, D. B., Wells, K. C., Cady-Pereira, K. E. & Shephard, M. W. A large underestimate of formic acid from tropical fires: constraints from space-borne measurements. Environ. Sci. Technol. 50, 5631–5640 (2016).
91. De Smedt, I. et al. QA4ECV HCHO Tropospheric Column Data from OMI (Version 1.1) (KNMI, 2017); https://doi.org/10.18758/71021031.
92. Zara, M. et al. Improved slant column density retrieval of nitrogen dioxide and formaldehyde for OMI and GOME-2A from QA4ECV: intercomparison, uncertainty characterisation, and trends. Atmos. Meas. Technol. 11, 4033–4058 (2018).
93. Zhu, L. et al. Observing atmospheric formaldehyde (HCHO) from space: validation and intercomparison of six retrievals from four satellites (OMI, GOME-2A, GOME-2B, OMPS) with SEAC4RS aircraft observations over the southeast US. Atmos. Chem. Phys. 16, 13477–13490 (2016).
94. Shen, L. et al. The 2005–2016 trends of formaldehyde columns over China observed by satellites: increasing anthropogenic emissions of volatile organic compounds and decreasing agricultural fire emissions. Geophys. Res. Lett. 46, 4468–4475 (2019).
95. Boersma, K. F. et al. QA4ECV NO2 Tropospheric and Stratospheric Vertical Column Data from GOME-2A (Version 1.1) (KNMI, 2017); https://doi.org/10.21944/qa4ecv-no2-gome2a-v1.1.

Acknowledgements This work was supported by the NASA Atmospheric Composition Modeling and Analysis Program (Grant Number NNX17AF61G) and by the Minnesota Supercomputing Institute. We thank D. Fu for providing optimal estimation isoprene retrievals over Amazonia and input on this manuscript, C. Barnet, E. Manning and R. Monarrez for providing CLIMCAPS HNO3 retrievals, M. Alvarado, K. Cady-Pereira, D. Gombos, J. Hegarty and I. Strickland for generating and testing isoprene absorption look-up tables employed here, and E. Edgerton for providing isoprene data from the SouthEastern Aerosol Research and Characterization (SEARCH) network. The SEARCH network was sponsored by the Southern Company and the Electric Power Research Institute. Isoprene measurements aboard the NASA DC-8 during SEAC4RS were supported by the Austrian Federal Ministry for Transport, Innovation and Technology (bmvi) through the Austrian Space Applications Programme (ASAP) of the Austrian Research Promotion Agency (FFG). T. Mikoviny is acknowledged for his support during SEAC4RS. We thank S. Springston for GoAmazon T3 data, which were supported by the ARM Climate Research Facility, the Central Office of the Large-Scale Biosphere Atmosphere Experiment in Amazonia (LBA), the Instituto Nacional de Pesquisas da Amazonia (INPA) and the Universidade do Estado do Amazonia (UEA). Part of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract to NASA.

Author contributions D.B.M. planned the project and oversaw the scientific interpretation. K.C.W. performed the ΔTb calculations, ANN training, CrIS isoprene retrievals and evaluation, lifetime calculations, emission optimization, and synthesis of results for major source regions. V.H.P. performed radiative transfer model simulations and provided guidance with CrIS data analysis. M.J.D. assisted with ANN training and application. D.B.M. and K.C.W. conducted the GEOS-Chem model simulations. K.H.B. worked on the development of RCIM and Mini-CIM and incorporated them into GEOS-Chem. J.A. d’G., M.G., C.W. and A.W. carried out the aircraft measurements used for evaluation. J.D.F. provided ground-based isoprene measurements from Amazonia. K.C.W. and D.B.M. wrote the manuscript. All authors reviewed and commented on the paper.

Competing interests The authors declare no competing interests.

Additional information Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2664-3.
Correspondence and requests for materials should be addressed to D.B.M.

Reprints and permissions information is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | Simulated spectral signals near 900 cm⁻¹ for the CrIS sensor. a, Brightness temperature ($T_b$) difference for simulated spectra with and without isoprene (black), nitric acid (red), ammonia (violet) and CFC-12 (yellow) and a 10% perturbation in water vapour (cyan). Red and blue arrows indicate the $\nu_{28}$ on-peak and off-peak spectral points used to calculate $\Delta T_b$. Simulations were performed with LBLRTM for an isoprene profile with 5 ppb in the boundary layer ($P > 800$ hPa) that decays exponentially aloft, and US standard atmosphere profiles of temperature, water vapour and nitric acid. b, Relationship between $\Delta T_b$ and isoprene column density, shaded by thermal contrast, for the full synthetic dataset used in this work. c, Global distribution of surface-atmosphere thermal contrast at the time of the CrIS overpass. Maps are derived from time-interpolated GMAO temperatures for January, April, July and October.
Extended Data Fig. 2 | Global distribution of isoprene columns, emissions and lifetime as predicted by GEOS-Chem. Predicted columns (left), emissions (middle) and lifetime (z < 500 m; right) are shown at 13:30 LT for January, April, July and October 2013 (top to bottom).
Extended Data Fig. 3 | Statistical uncertainty in the global distribution of monthly mean isoprene:HCHO ratios as a function of isoprene and NO$_x$ regime. a, Relative 95% confidence interval in the mean ratio for each isoprene and tropospheric NO$_x$ bin. b, Number of observations in each bin.
Extended Data Fig. 4 | Global distribution of isoprene column densities derived from CrIS. Plotted are the mean (left) and relative standard deviation (right) across the 10 ANNs for January, April, July and October 2013 (top to bottom).
Extended Data Fig. 5 | Boundaries of the four regions examined in the seasonal bar plots shown in Figs. 5, 6.
Extended Data Fig. 6 | Measured and simulated HCHO columns. Plotted are the HCHO columns measured by OMI (left) and simulated by GEOS-Chem (right) at ~13:30 LT for January, April, July and October 2013 (top to bottom).
Extended Data Fig. 7 | CrIS cloud screening and ANN performance.

**a**, Function used for cloud screening CrIS L1B data before $\Delta T_b$ calculation. The black line shows the modelled clear-sky difference between the 900 cm$^{-1}$ brightness temperature and surface skin temperature, as a function of water vapour column density (calculated using LBLRTM$^{80,81}$). The solid red line is the linear approximation used here, and the dashed red line represents a less stringent threshold used to test the sensitivity of the results to our cloud screening approach.

**b**, Sensitivity of the CrIS brightness temperature differences ($\Delta T_b$) and isoprene columns ($c$) to cloud screening. Data shown represent the median relative differences between the base-case results (derived using the solid red line in **a**) and those derived using the less stringent cloud screening threshold (dashed red line in **a**). **d**, **e**, Scatterplots of the predicted versus true isoprene columns for the six-predictor ANN (d) and an ANN in which $\Delta T_b$ is withheld as a predictor variable (e). Red dots show the mean of the ten ANN predictions, and blue error bars show the standard deviation across the predictions. **f**, The relative uncertainty (based on the difference between the mean ANN predicted value and the true value) for the six-predictor ANN, binned as a function of thermal contrast and isoprene column density.
Extended Data Fig. 8 | CrIS isoprene measurements over Amazonia. The maps were derived using ANN- (left) and optimal estimation- (right) based approaches. Data are shown for September 2014 and displayed as absolute columns.
Extended Data Table 1 | Spatial correlation between monthly mean CrIS $\Delta T_b$ and monthly mean 13:30 LT isoprene columns predicted by GEOS-Chem at $2^{\circ} \times 2.5^{\circ}$ resolution for select regions

| Region               | Month  | $\Delta T_b$-GEOS-Chem isoprene correlation, $r$ | # data points |
|----------------------|--------|--------------------------------------------------|---------------|
| Australia            | January| 0.54                                             | 323           |
| Central Africa       | April  | 0.43                                             | 357           |
| Southeast United States | July  | 0.72                                             | 90            |
| Amazonia             | October| 0.57                                             | 340           |
## Extended Data Table 2 | Data sources for the six input parameters used for ANN training and retrievals

| Input parameter     | Source for training set                          | Source for ANN-based retrieval               |
|---------------------|--------------------------------------------------|---------------------------------------------|
| $\Delta T_d$        | ELANOR simulation                                | CrIS L1B radiances                         |
| H$_2$O vapor column | Assimilated meteorology (GMAO; TES-like sampling)| Assimilated meteorology (GMAO; CrIS collocation) |
| HNO$_3$ column      | MOZART CTM                                       | CrIS CLIMCAPS                               |
| Thermal contrast    | Assimilated meteorology (GMAO; TES-like sampling)| Assimilated meteorology (GMAO; CrIS collocation) |
| Pressure            | Assimilated meteorology (GMAO; TES-like sampling)| Assimilated meteorology (GMAO; CrIS collocation) |
| Satellite view angle| Randomly defined                                 | CrIS satellite pointing angle               |