

Space-based observations of tropospheric ethane map emissions from fossil fuel extraction

Received: 18 December 2023

Accepted: 23 August 2024

Published online: 07 September 2024

 Check for updates

Jared F. Brewer¹, Dylan B. Millet¹✉, Kelley C. Wells¹, Vivienne H. Payne², Susan Kulawik³, Corinne Vigouroux⁴, Karen E. Cady-Pereira⁵, Rick Pernak⁵ & Minqiang Zhou⁶

Ethane is the most abundant non-methane hydrocarbon in the troposphere, where it impacts ozone and reactive nitrogen and is a key tracer used for partitioning emitted methane between anthropogenic and natural sources. However, quantification has been challenged by sparse observations. Here, we present a satellite-based measurement of tropospheric ethane and demonstrate its utility for fossil-fuel source quantification. An ethane spectral signal is detectable from space in Cross-track Infrared Sounder (CrIS) radiances, revealing ethane signatures associated with fires and fossil fuel production. We use machine-learning to convert these signals to ethane abundances and validate the results against surface observations ($R^2 = 0.66$, mean CrIS/surface ratio: 0.65). The CrIS data show that the Permian Basin in Texas and New Mexico exhibits the largest persistent ethane enhancements on the planet, with regional emissions underestimated by seven-fold. Correcting this underestimate reveals Permian ethane emissions that represent at least 4–7% of the global fossil-fuel ethane source.

Ethane (C_2H_6) is one of the most abundant volatile organic compounds (VOCs) present in the global atmosphere^{1–3}. It is a precursor of tropospheric ozone, affects the transport and fate of reactive nitrogen through the formation of peroxyacetyl nitrate (PAN)⁴, and influences the atmosphere's oxidizing capacity^{3,5–7}. Ethane is also co-emitted with methane from fossil fuel sources and has become an important tracer for fingerprinting and quantifying anthropogenic methane emissions on regional^{8,9} and national¹⁰ scales. Attempts have been made to replicate this technique at the global^{11–13} scale, but success has proven difficult due in part to insufficient observational coverage and uncertainty in the methane:ethane emission ratios^{11–13}.

The sources of ethane are primarily anthropogenic: recent estimates of fossil fuel-related ethane emissions range from 7–13 teragrams per year ($Tg\ yr^{-1}$), with biofuel and biomass burning contributing an additional 2–8 $Tg\ yr^{-1}$ ^{13,14,15}. Geologic emissions of ethane are uncertain,

and not always included in budget assessments, though some studies estimate them at 2–6 $Tg\ yr^{-1}$ ^{11,3,16–18}. Natural ethane sources, including oceanic and biogenic emissions, are believed to be globally negligible^{5,7,17,19}. Atmospheric ethane is primarily destroyed by reaction with hydroxyl radicals (OH) on a global-mean timescale of ~2 months²⁰, sufficiently long that it provides a valuable air mass and source tracer over long distances^{21–23}.

Uncertainty in the fossil fuel ethane source is large, with top-down studies often indicating substantially higher emissions than inventories predict (e.g., a factor of 1.4 to 3 more^{3,14,24,25}). Fossil ethane emission ratios relative to methane also vary temporally and spatially in a manner that is not well constrained^{10,14}. These two issues challenge ethane's use as a source tracer along with efforts to quantify its atmospheric impacts.

Satellite-based measurements of tropospheric ethane would offer a major advance by providing globally continuous and long-term

¹University of Minnesota, Department of Soil, Water, and Climate, Saint Paul, MN, USA. ²NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. ³BAER Institute, Petaluma, CA, USA. ⁴Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium. ⁵Atmospheric and Environmental Research, Lexington, MA, USA. ⁶Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China. ✉e-mail: dbm@umn.edu

observations to address the above challenges^{26–28}. Ethane is a relatively weak absorber but features distinctive bands in the shortwave infrared (IR; v7 band near 3000 reciprocal centimeters [cm^{-1}])²⁹ and thermal IR (v9 band near 815 cm^{-1})³⁰ that are amenable to remote sensing. The v7 feature has been used for ground-based solar Fourier Transform Infrared (FTIR) measurements such as those produced by the Network for the Detection of Atmospheric Composition Change (NDACC)^{31–33}. Satellite-based ethane measurements have so far been performed by solar occultation (using v7)³⁴ and by limb sounding (using v9)³⁵. However, these ethane retrievals are limited to the upper troposphere and stratosphere and are unable to capture spatial and temporal variability linked to surface emissions.

In this work, we present a retrieval of tropospheric ethane using the Cross-Track Infrared Sounder (CrIS; see Methods) onboard the Suomi-National Polar-orbiting Partnership (Suomi-NPP) satellite, an instrument well-suited for the task. CrIS has low noise in the longwave IR (LWIR) relative to similar sensors³⁶, an afternoon overpass that maximizes sensitivity in the troposphere²⁷, and high resolution in both space and time. Compared to observations at shorter wavelengths, the LWIR radiances detected by CrIS are much less impacted by atmospheric and surface scattering but have lower near-surface sensitivity due to atmospheric absorption. This issue is partly counteracted by the CrIS nadir view, which provides far more tropospheric sensitivity than the limb-based observations previously used for space-based ethane detection^{26,28}. Below, we first demonstrate the spectral detection of ethane in the CrIS radiances and highlight the coherence of this signal with known fossil fuel and biomass burning sources. We next employ a neural network to derive ethane column concentrations from these spectral signals and evaluate the results against established ground-based observations. Finally, we combine this dataset with chemical transport model (CTM) simulations to investigate the largest persistent ethane enhancements visible in the global CrIS data record, which occur over the Permian basin in the southwestern United States (US).

Results

Spectral detection of ethane from space

We use the hyperspectral range index (HRI) to characterize the ethane spectral signal detected by CrIS. The HRI is a dimensionless measure

of the spectral signature of a target atmospheric gas^{37,38} that is computed via²⁸:

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

where \mathbf{y} is a measured spectrum from CrIS and $\bar{\mathbf{y}}$ is the mean background spectrum calculated iteratively from background scenes in which ethane is undetectable³⁹. The same background scenes are used to calculate the spectral covariance matrix \mathbf{S}_y , which accounts for correlations between spectral channels due to factors other than ethane⁴⁰. \mathbf{K} is the spectral Jacobian reflecting the change in absorption per change in ethane³⁹ (Supplementary Fig. 1). A normalization factor N is applied based on the HRI standard deviation over a region where no ethane enhancements are expected. HRI values computed in this way specifically quantify the above-background signal strength, and by definition have a mean of 0 and a standard deviation of 1 under background conditions. Higher values imply a strong signal from the target gas. Initially developed and applied to quantify ammonia and sulfur dioxide from the Infrared Atmospheric Sounding Interferometer (IASI)^{38,40}, the HRI has subsequently been employed to measure a range of VOCs from IASI²⁸ and to quantify isoprene from CrIS³⁹. Here, we use an updated version of the Retrieval of Organics from CrIS Radiances³⁹ (ROCRv2) to iteratively derive an ethane HRI between 800 and 850 cm^{-1} , a range that encompasses the v9 absorption feature (Supplementary Fig. 1). The subsequent conversion of these spectral signals to ethane column abundances is described below.

Figure 1 illustrates the spectral detection of ethane from CrIS using the HRI approach. Plotted are data from January 2nd, 2020, over the South Pacific downwind of Australian megafires that occurred during that period. Plumes from these fires provide a useful test-case scenario for ethane detection, given (1) strong expected ethane emissions from the fires, (2) lofting of those emissions high into the troposphere and lower stratosphere⁴¹, and (3) over-ocean transport, which minimizes any spectral influences from the surface (as ocean emissivity is less variable than over land). Three independent characterizations of the transported biomass burning plume are shown: (a)

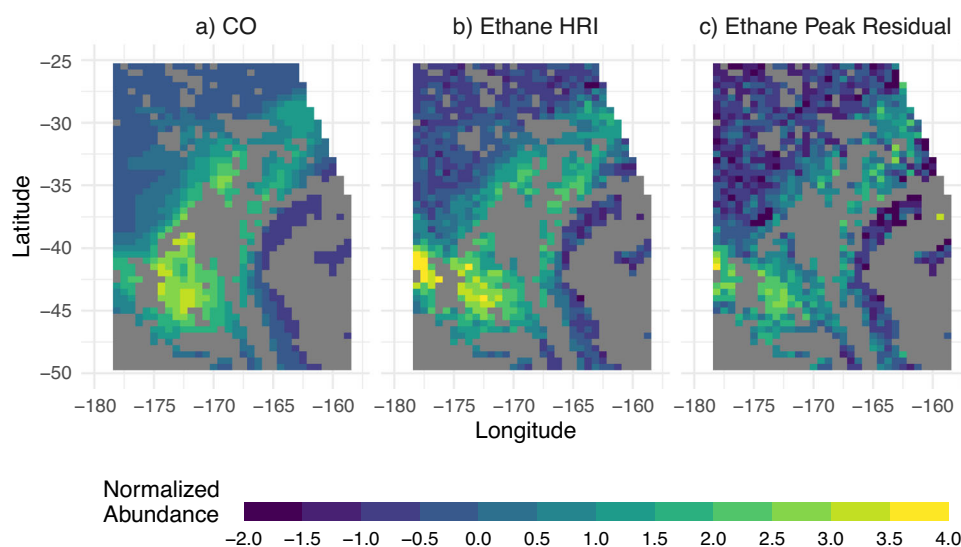


Fig. 1 | Spectral detection of ethane in radiances measured by the Cross-track Infrared Sounder (CrIS). Plotted are (a) CrIS carbon monoxide (CO) columns⁴², (b) ethane hyperspectral range indices (HRIs), and (c) ethane brightness temperature differences for a fire plume over the South Pacific on January 2, 2020. All quantities

are normalized and screened for clouds using the 900 cm^{-1} brightness/surface-skin temperature difference²⁷. The CrIS data shown is primarily from granule 13, with additional data from granules 12, 14, 29, 232, 233, and 234.

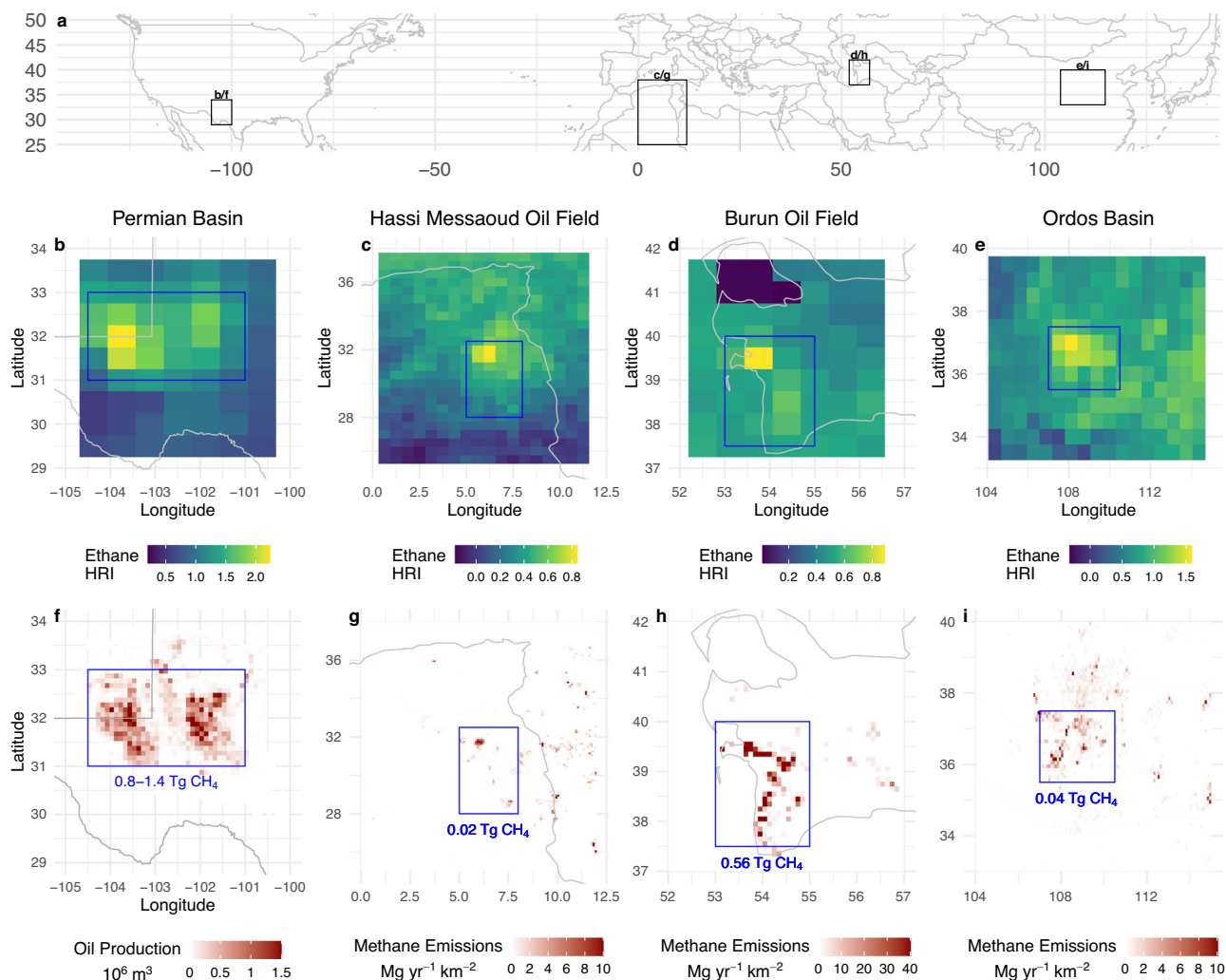


Fig. 2 | Ethane hyperspectral range indices (HRIs) derived from CrIS radiances over fossil fuel production regions. Panel (a) shows the location and spatial extent of the other panels, which contain data for the Permian basin in Texas and New Mexico, US⁴⁶ (b, f), the Hassi Messaoud oil field in Algeria⁴⁷ (c, g), the Burun oil field in Turkmenistan⁴⁷ (d, h), and the Ordos basin in northern China⁴⁸ (e, i). Panels (b–e) show the 2013–2021 mean ethane HRI values as detected by CrIS over these

areas. Panels (f–i) show bottom-up fossil fuel data for the same locations based on oil production by well⁴⁹ (f; 2018 data) or based on the estimated methane emissions from oil extraction and production⁵⁰ (g–i). Blue rectangles highlight high-HRI regions on each set of paired figures, with annual oil-related methane emissions from within those squares listed beneath⁵⁰. Color scale maxima vary between regions.

the CrIS carbon monoxide (CO) column abundance obtained via optimal estimation⁴²; (b) the CrIS ethane HRI computed as above; and (c) the mean on-peak/off-peak brightness temperature difference at the wavenumbers with highest ethane absorption (Supplementary Fig. 1). For (c), we use CrIS spectral residuals processed with the MULTiSpEctra, MULTiSpEcies, Multi-SEnsors (MUSES) algorithm from the Tropospheric Ozone and its Precursors from Earth System Sounding (TROPESS) project^{43,44}. Several modifications were made to the MUSES/TROPESS processing strategy for this work, and these are described in Methods. Once the spectral residuals are derived, values at the eight largest ethane peaks in the v9 band are then each background-corrected by subtracting the mean adjacent off-peak value as indicated in Supplementary Fig. 1 and Supplementary Table 1. These background-corrected ethane residuals are averaged to obtain the single spectral index shown in Fig. 1c. All quantities (a–c) are cloud-screened using the difference between the 900 cm⁻¹ brightness temperature and surface skin temperature²⁷.

Figure 1 shows strong consistency between the above measurements, with all three mapping the same large fire plume stretching from (50° S, 180° E) to (30° S, 165° E). This coherence reflects the physical structure of the fire plume and the distribution of co-emitted

absorbers within that plume (in this case, CO and ethane). The brightness-temperature difference ethane metric is noisier than the corresponding HRI, reflecting the sensitivity and noise advantages of the latter. While factors such as water vapor and ash can interfere with thermal IR retrievals in wildfire smoke⁴⁵, the spatial consistency between the two entirely different indices shown in Fig. 1 demonstrates that both are capturing the underlying signal due to ethane rather than that of potential interferences.

Spatial patterns in the ethane signal

Along with fires, the strongest ethane HRI enhancements that emerge in the global CrIS record are associated with major fossil fuel production regions. This is illustrated in Fig. 2, which shows CrIS ethane HRI values over four oil and gas basins: the Permian in Texas and New Mexico, US⁴⁶, the Hassi Messaoud oil field in Algeria⁴⁷, the Burun oil field in Turkmenistan⁴⁷, and the Ordos basin in northern China⁴⁸. Of these four areas, we see the highest ethane HRI values over the Permian. In fact, this finding extends across the entire CrIS data record—the Permian basin nearly always exhibits the highest ethane HRI at any location on the globe. The Permian also shows a clear spatial correlation between fossil fuel production and ethane signal: Fig. 2b, f both

show a clear two-lobed distribution corresponding to the Delaware (western) and Midland (eastern) component basins. Of these, the Delaware basin has both a higher mean ethane HRI and a higher total fossil fuel extraction rate, producing 54% and 61% of Permian oil and natural gas, respectively^{46,49}. The corresponding methane emissions were estimated at 1.7 Tg CH₄ yr⁻¹ (Delaware) and 1.0 Tg CH₄ yr⁻¹ (Midland) in 2018–2019 using TROPospheric Monitoring Instrument (TROPOMI) satellite-based observations⁴⁶.

The other basins highlighted in Fig. 2 show lower ethane HRI values, but nevertheless exhibit a tight spatial coherence between the ethane spectral signal and underlying fossil fuel extraction activities. We highlight this coherence in Fig. 2 with boxes delineating the HRI enhancements for each basin (panels b–e); these boxes are then replicated for reference over the corresponding emission/production maps (panels f–i). This coherence is particularly strong over the Hassi Messaoud oil field. Ethane HRI enhancements over the Burun oil field in Turkmenistan have a distribution that matches that of oil-extraction methane emissions, but with the highest ethane signals at the base of the Cheleken peninsula. The Ordos basin in China features the second-highest ethane HRI values across the regions in Fig. 2 but with much lower methane emissions (50 Gg CH₄ yr⁻¹) than those in Burun (690 Gg CH₄ yr⁻¹)⁵⁰.

These observations represent an opportunity to characterize atmospheric ethane sources through time and space, and for understanding their connections to fossil fuel production and to the co-emission of methane. However, the HRI is a unitless measure that depends not only on ethane abundance but also on spectral and environmental factors as described previously^{39,45}. We therefore employ a retrieval (described below) to convert the CrIS HRI values into ethane column amounts.

Retrieval of Organics from CrIS Radiances (ROCR) algorithm updates

We convert the satellite-derived HRI fields to ethane columns using an artificial neural network (ANN). The retrieval involves two significant updates to version 1 of the Retrieval of Organics from CrIS Radiances (ROCR) algorithm described by Wells et al.³⁹. In ROCRv2, we use the Line-by-Line Radiative Transfer Model^{51,52} (LBLRTM) rather than the Earth Limb and Nadir Operational Retrieval⁵³ (ELANOR) model to calculate spectral Jacobians and synthetic radiances. LBLRTM computes molecular optical depths line-by-line rather than via pre-computed lookup tables, thus increasing accuracy and species flexibility. The second update involves an explicit and customizable means of accounting for vertical retrieval sensitivity and is described later in this section.

Reactive trace gas fields (VOCs, ozone, ammonia) for ANN training use 2019 output from the GEOS-Chem CTM (Version 13.3⁵⁴; see Methods) for 12:00–15:00 local time (LT), reflecting the -13:30 LT CrIS observations. The model is sampled between 60° S and 75° N every 16 days throughout the year (total: 23 days), with random positive scaling ($\sigma = 1$, $\mu = 1$) applied to the VOC profiles to minimize any effects from correlations with atmospheric conditions³⁹. Profiles for other spectrally relevant species (e.g., chlorofluorocarbons, carbon dioxide, methane, nitrous oxide, carbon tetrachloride) are based on a climatology from the Model for Ozone and Related chemical Tracers (MOZART)⁵⁵. Meteorological fields (temperature, water vapor) use the Modern-Era Retrospective analysis for Research and Applications (MERRA-2).

The above input data are used in LBLRTM to generate synthetic HRIs as would be detected by CrIS. For each model scene, we simulate top-of-atmosphere radiances with and without atmospheric ethane for three view angles selected randomly from each of three bins (0–16°, 17–32°, 33–48°). We then add CrIS-like random noise, with zero mean and a standard deviation derived from values reported in the CrIS LIB product. These CrIS noise values are (1) lowest in the middle of the

LWIR band (700–1000 cm⁻¹) and (2) lower at higher brightness temperatures³⁶. We next calculate our set of training HRIs by combining the LBLRTM-derived synthetic radiances with background spectral covariances obtained from the CrIS measurements themselves, as detailed by Franco et al. for IASI²⁸. Calculation of the synthetic HRIs against a background of zero ethane reduces sensitivity to model errors, but (as will be discussed) if the true ethane background is above the CrIS detection threshold it would then need to be added post-hoc to the ANN output prior to data analysis²⁸. Through the above procedure we obtain a training HRI dataset for which the underlying ethane columns are known. The range of synthetic HRIs encompasses that of the CrIS HRIs, so that the ANN predictions described next are in-sample.

A feed-forward ANN is trained on this dataset to predict ethane column amounts from the corresponding HRI values. The employed ANN includes two hidden layers (with 20 and 10 nodes, respectively) and a single output node, and is trained on 10 random extractions of the input dataset. The ANN includes as predictors relevant ancillary variables that affect the spectroscopy. For ethane, these include water vapor (column amount), surface skin temperature, atmospheric temperatures (surface air temperature, -1 kilometer, -4 km, -10 km), surface pressure, and view angle. As with ROCRv1³⁹, latitude and longitude are not included as predictors. As an improvement over ROCRv1³⁹ we introduce here an additional predictor to account for the sensitivity of thermal IR satellite measurements to the vertical profile of the absorber: P₉₀, the atmospheric pressure below which 90% of the ethane column resides. For training purposes these vertical profiles are obtained from the randomized GEOS-Chem simulations referenced above; inclusion of P₉₀ in the ANN then allows the vertical sensitivity of the measurement to be explicitly represented via a single variable. As will be outlined below, it also allows the retrieved ethane column to be adjusted *a posteriori* to any assumed vertical profile shape—e.g., for an internally consistent comparison with model fields, or where an observational constraint on the profile shape is available.

Training performance and uncertainty

Figure 3 summarizes the ANN training performance. The network mean can reproduce 88% of the ethane column variance in the training set with a root-mean-square error of 4.1×10^{15} molecules cm⁻². Figure 3b and Supplementary Fig. 2 compare the relative impacts of the input variables on the column predictions, demonstrating that the HRI is most important—particularly when ethane is enhanced. Next in importance is P₉₀, and its influence is likewise most pronounced for higher ethane columns. The leading importance of the HRI and P₉₀ variables show that the retrieval is predominantly determined by the ethane column amount and its vertical distribution, as expected for a thermal IR measurement. The five temperature variables individually have only modest impacts on the predictions (Supplementary Fig. 2) but are collectively important given that thermal IR retrieval sensitivity depends on the surface-atmosphere thermal contrast.

The relative error in the predicted value increases as column values decrease, from -10% at a column abundance of 5×10^{16} molecules cm⁻² to -30% at 1×10^{16} molecules cm⁻². Supplementary Fig. 3 plots the prediction error (the absolute value of the difference from the true value) and bias (the simple difference from the true value) as a function of both thermal contrast and P₉₀. The mean error exceeds 100% for ethane columns below 4×10^{15} molecules cm⁻², defining an approximate detection threshold for the CrIS observations (this threshold does not vary significantly with thermal contrast or P₉₀). The lack of a strong relationship between error and thermal contrast seen in Supplementary Fig. 3 is notably different from the isoprene case³⁹. We attribute this difference to the longer atmospheric lifetime of ethane: isoprene is highly concentrated in the lowermost atmosphere and thus more sensitive to near-surface thermal contrast. Ethane is more dispersed through the vertical column and this dependence is thus

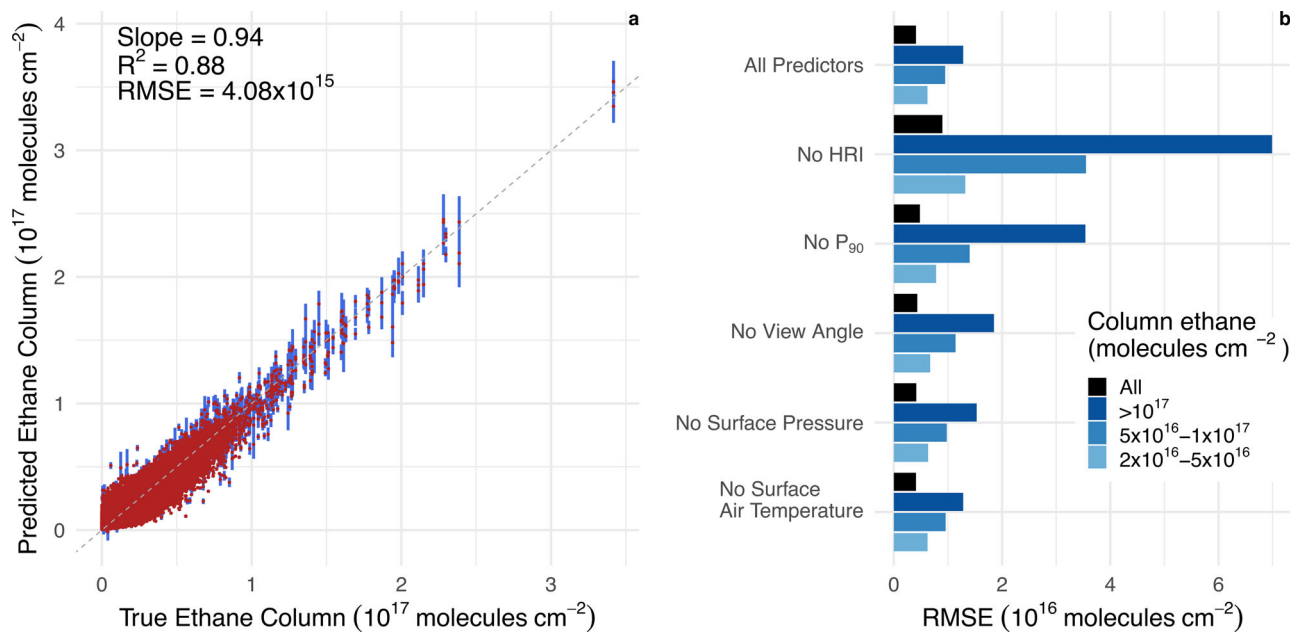


Fig. 3 | Artificial Neural Net (ANN) training set and predictor importance. Panel (a) displays predicted vs true ethane columns for the full training set, shown as the mean (red dots) and standard deviation (blue error bars) across the 10 individual ANNs. Panel (b) displays the relative importance of selected individual predictors

via their impact on the overall 10-ANN root-mean-square error (RMSE). ‘All Predictors’ shows the RMSE when including all predictors; other results show the RMSE when the designated predictor is omitted. A version of this figure including all predictors is included in the supplement (Supplementary Fig. 2).

distributed amongst the suite of atmospheric temperature inputs. The bias plots in Supplementary Fig. 3 further reveals a modest tendency for the retrieval to overpredict the training values when ethane column amounts are low and to underpredict them at high column amounts and when ethane is lofted vertically (by a mean of 12% for columns $>2 \times 10^{16}$ molecules cm^{-2} and $P_{90} < 300$ hectopascals [hPa]). This moderate underestimate at high column amounts, revealed in Supplementary Fig. 3, also manifests in the predicted vs. true regression slope (0.94; Fig. 3).

Ethane column derivation and comparison with observations

We use the 10-ANN ensemble described in the prior section to derive ethane column densities from the CrIS-measured HRI values. Meteorological inputs for the retrieval use MERRA-2 reanalysis. We perform the ANN retrieval on a $0.5^\circ \times 0.625^\circ$ horizontal grid and for a range of P_{90} values (150, 250, 350, 450, 550, and 650 hPa; chosen based on global P_{90} statistics from a GEOS-Chem simulation). The CrIS retrieval output for a given location can hence be adapted to any assumed vertical profile by interpolating between the ethane columns derived for the two proximal P_{90} values.

We evaluate the resulting CrIS ethane measurements against surface-based solar FTIR observations collected by the Network for the Detection of Atmospheric Composition Change (NDACC; see Methods). A map of NDACC stations used in this analysis is shown in Supplementary Fig. 4. NDACC ethane retrievals are well-characterized, with $<6\%$ systematic error and $<3\%$ random error³³, exhibit good agreement with in-situ observations⁵⁶, and have a long history of use^{3,21,32,33,57–63}. In our baseline comparison approach we use the NDACC *a posteriori* profiles to specify the ethane P_{90} value used in the CrIS retrieval for each time and location. This provides a CrIS-independent constraint and ensures internally consistent profile assumptions between the two datasets.

Figure 4 displays the CrIS-NDACC comparison results obtained in this way. The two datasets are significantly correlated ($R^2 = 0.66$), showing that CrIS can capture variability in atmospheric ethane across the NDACC locations and sampling times. The CrIS data are lower than

the surface column observations, with a mean CrIS/NDACC column ratio of 0.65 and a major axis slope of 0.60 (bootstrapped 95% confidence interval: 0.59–0.61). The intercept of $(-1.1 [-1.2 \text{ to } -0.9]) \times 10^{15}$ molecules cm^{-2} is near-zero and below the CrIS limit of detection, obviating any background addition that would otherwise be required from the HRI formulation. Restricting the analysis to the single CrIS grid cell containing the NDACC station rather than the 3×3 matrix mean increases the slope to 0.75 (0.72–0.77), increases the scatter ($R^2 = 0.48$), and shifts the intercept slightly $(-3.2 [-3.4 \text{ to } -2.9]) \times 10^{15}$. Despite the greater noise, this latter approach likely provides a more robust comparison near source regions (e.g., for the St. Petersburg, Bremen, Xianghe, and Toronto sites). Relaxing the ± 2 h temporal overlap window to ± 8 hours, or excluding the altitude-dependent scaling for high-elevation sites (see Methods), does not appreciably change these results. Averaging the results in Fig. 4 by station (see Supplementary Fig. 5) shows that the CrIS:NDACC comparisons adhere to a single statistical relationship ($R^2 = 0.92$; CrIS/NDACC column ratio = 0.65; slope = 0.45) across the global distribution of observing stations—arguing against any large regional differences in retrieval performance.

The NDACC degrees of freedom for signal (averaging -1.5) indicate that their *a posteriori* profile shapes and associated P_{90} values are largely determined by the NDACC prior rather than by the observations. For that reason, we performed an additional CrIS-NDACC comparison instead using P_{90} values predicted by GEOS-Chem for each time and location. Results (slope = 0.66; $R^2 = 0.55$) are consistent with the baseline comparison, supporting the robustness of the conclusions above.

Possible reasons for the non-unity CrIS:NDACC slope shown in Fig. 4 could include inconsistency between the employed thermal and shortwave infrared cross-sections, a misspecification of the CrIS noise in the vicinity of the ν_9 ethane feature (as used for deriving the synthetic HRIs), or radiative transfer model uncertainties (different forward models are used for the two datasets). Future refinements of the CrIS ethane product presented here will explore such factors for potential improvement.

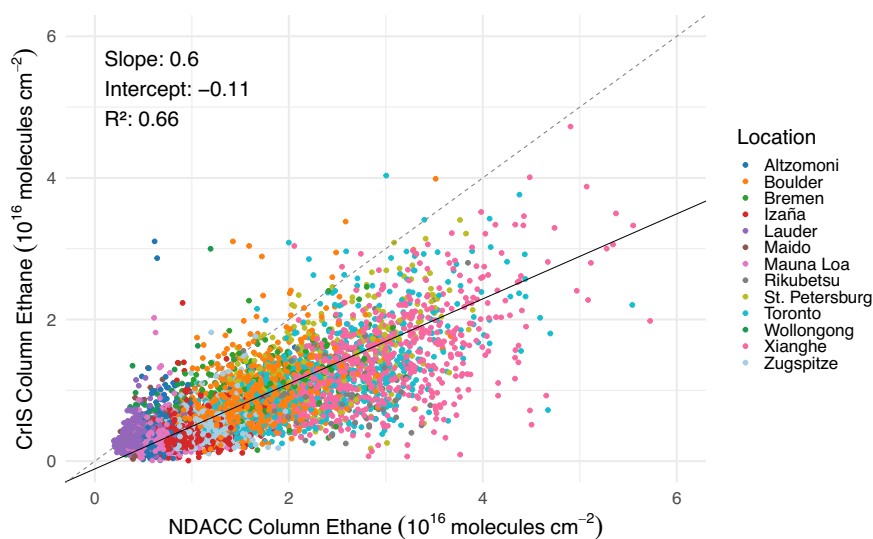


Fig. 4 | Evaluation of Suomi-NPP Cross-track Infrared Sounder (CrIS) ethane retrievals against ground-based observations from the Network for the Detection of Atmospheric Composition Change (NDACC). NDACC observations are averaged over ± 2 h of the local satellite overpass time. CrIS observations reflect a 3×3 pixel mean ($0.5^\circ \times 0.625^\circ$ resolution) centered on the NDACC site. The black

line shows a standard major axis fit, with parameters given inset (intercept units: 10^{16} molecules cm^{-2}). Vertical scaling corrections (Methods) have been applied to the CrIS columns for the six high-altitude sites (Altzomoni, Mado, Mauna Loa, Izaña, Rikubetsu, and Zugspitze). The 1:1 line is also shown (dashed gray line).

Ethane in the Permian Basin

The ~ 400 km² Permian basin in New Mexico and west Texas consistently exhibits the largest ethane enhancements over the 2013–2021 CrIS record (Fig. 2). Over the last decade, with advances in drilling technology this region has become the most productive oil-producing basin in the US and one of the largest in the world⁴⁶. As a result, the Permian has been a focus of methane-related research^{46,64,65}, with a recent top-down analysis concluding that leakage rates 60% higher than the national average led to a methane flux of 2.7 ± 0.5 Tg yr⁻¹ in 2018–2019—the largest reported amount from any US oil/gas production region⁴⁶.

The accompanying ethane emissions are less well known. Bottom-up inventories used by GEOS-Chem⁶ suggest an ethane flux of 0.074 Tg yr⁻¹ from the Permian: however, this leads to a dramatic underestimate of the resulting column amounts compared to those seen by CrIS, indicating that the true flux is much higher. Barkley et al.⁶⁶ obtained a regional ethane flux estimate of 0.28 Tg yr⁻¹ based on airborne sampling downwind of the Permian that mainly took place in 2017, though the authors report low confidence in this finding due to limited regional sampling coverage. A recent bottom-up estimate for 2015 from NOAA's Fuel-based Oil and Gas (FOG) inventory⁶⁷ suggests considerably higher Permian ethane emissions for 2015 (0.56 Tg) than those inferred by Barkley et al.⁶⁶ for 2017 (0.28 Tg) or predicted by GEOS-Chem.

We performed a series of sensitivity simulations with the GEOS-Chem CTM (configured as described in Methods) to determine the Permian ethane flux required to match the CrIS observations. Simulations were first performed for 2019 over a nested $0.5^\circ \times 0.625^\circ$ domain containing the Permian basin (29° – 34° N; 100° – 106° W). In addition to the reference run with standard model emissions, simulations were performed with ethane emissions from the subregion encompassing the Delaware and Midland basins (31° – 34° N, 101° – 105° W) scaled by factors of 2, 5, 6, 7, 8, 10 and 11. The model was sampled from 12:00 to 15:00 LT daily to match the CrIS overpass, with the simulated profile shapes then used to determine the grid-cell and time-specific P_{90} values for the corresponding CrIS retrievals.

The top panel of Fig. 5 plots the mean CrIS/GEOS-Chem column difference over the full Permian domain as a function of the emission scale factor over employed in the model. The intercept of 7.4 (95%

confidence interval: 7.3–7.6) reflects the factor by which the bottom-up Permian emissions need to be scaled to be consistent with the space-borne constraints. The above confidence interval was generated by bootstrapping the statistical fit shown in Fig. 5 and thus does not account for any systematic satellite or model errors. The result implies an ethane flux of 0.53 Tg in 2019, with the Permian basin alone then responsible for 4–7% of the total estimated fossil-fuel ethane emissions worldwide^{3,14}.

Figure 5 also spatially compares the CrIS ethane columns over the Permian with those derived by the base-case GEOS-Chem simulation and by an optimized simulation with regional emissions scaled by the derived factor of 7.4. The updated simulation captures the mean ethane column magnitude observed by CrIS over the Permian but not the underlying spatial features, because the bottom-up inventory employed in the model (which is scaled uniformly over the indicated subdomain) inaccurately maps the distribution of emissions across the Delaware and Midland basins. Advancing ethane emission inventories, and improving on the Permian flux estimate presented here, will thus require improved and up-to-date spatial mapping of the underlying activities (e.g., as provided via GFEI and FOG^{50,67}).

We extended the above inversion technique through the 2014–2019 period to characterize the associated ethane emission trend as seen by CrIS. Figure 6 shows the results and compares them to US Energy Information Agency records of regional oil production^{46,47,66–68} (natural gas follows a similar trend) and to other bottom-up and top-down constraints. The Permian fossil fuel production rates rise strongly through this period, with slightly lower rates of increase in 2016–2017. The CrIS-derived ethane emissions likewise show an overall increase, with a 2016 dip followed by a through the remainder of the period. The CrIS ethane emission trend arises entirely from changes in the measured ethane columns, as the prior fluxes in GEOS-Chem are static over this time³.

Figure 6 also shows that the CrIS-derived Permian emissions agree match well with the Barkley et al.⁶⁶ estimate for 2017 (0.31 Tg vs 0.28 Tg, respectively) but are $\sim 2 \times$ lower than the FOG bottom-up estimate for 2015, and are also lower than those implied by applying a regional 17–18% mol/mol ethane:methane emission ratio (32–34% on a mass basis)^{3,66,69} to recent satellite-inferred methane fluxes^{46,65}. Applying a

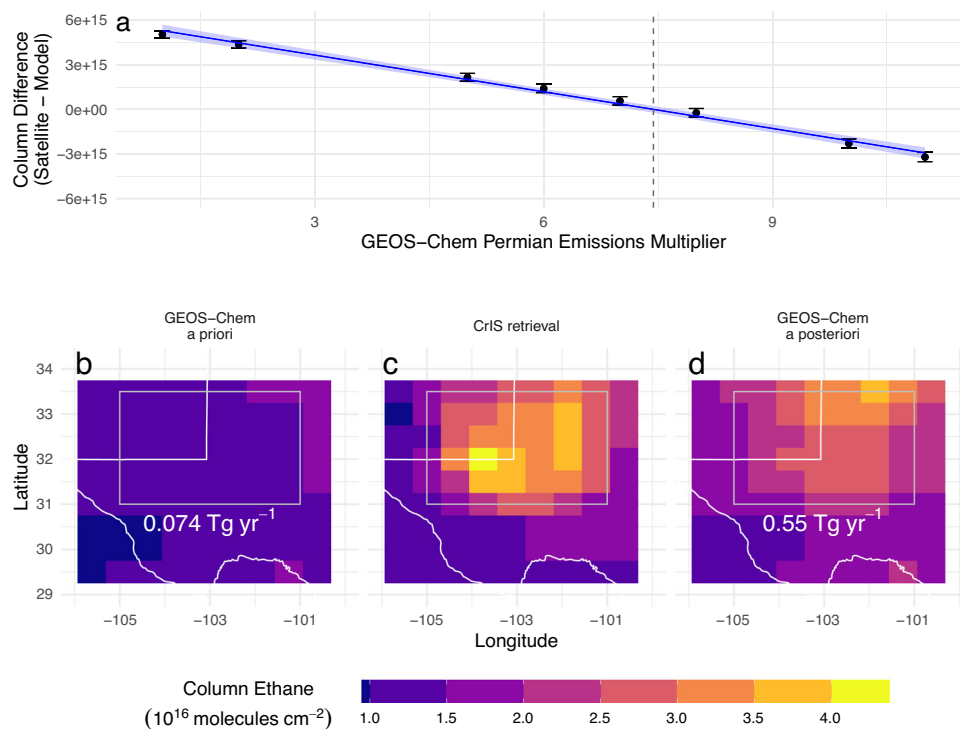


Fig. 5 | Derivation of 2019 mean ethane columns over the Permian basin. Panel (a) displays the difference in the 2019-mean ethane columns over the Permian basin between CrIS and GEOS-Chem as a function of the regional emissions multiplier used in the model. Error bars reflect 95% confidence intervals for each mean. The blue line and shaded region show a linear least-squares fit and associated 95% confidence interval, with the intercept indicating a true Permian emission scale factor of 7.4 ± 0.1 (vertical dashed line). The bottom panel shows 2019 mean ethane columns over the Permian basin as (b) simulated by the GEOS-Chem base-case (GC

a priori), (c) observed from space by CrIS, and (d) simulated by the updated GEOS-Chem run based on a Permian emission scale factor of 7.4 (GC a posteriori). Vertical profile shape information (P_{90} values, defined as the pressure level above 90% of the ethane column) for the CrIS retrievals are obtained from the corresponding GC a posteriori output. The gray square indicates the subdomain over which the emission scale factors were applied in the model. Total annual ethane emissions from the full Permian domain are indicated in white on panels (b, d).

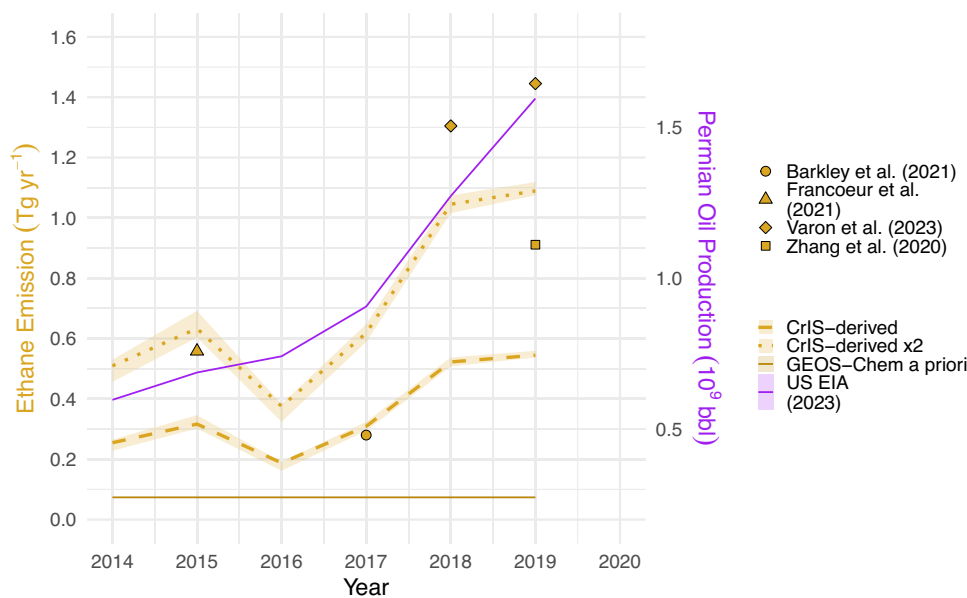


Fig. 6 | Trends in oil production and emissions over the Permian Basin. Annual mean ethane emission estimates are plotted in gold on the left axis, while annual oil production rates for the same region⁶⁸ are plotted in purple on the right axis. The Barkley et al.⁶⁶ flux estimate was derived from airborne ethane measurements; the Francoeur et al.⁶⁷ estimate is a bottom-up estimate of ethane emissions; the GEOS-

Chem ethane emissions are from Tzompa-Sosa et al.³; and the Varon et al.⁶⁵ and Zhang et al.⁴⁶ values were obtained by applying a basin-average ethane:methane emission ratio of 18%^{3,66,69} to their satellite-derived methane flux estimates. The Cross-track Infrared Sounder (CrIS) ethane emissions in this plot were derived using the same methodology as shown in Fig. 5.

factor of two correction to the CrIS values, as would be implied by the NDACC comparison in Fig. 4, yields 2014–2019 emissions that more closely match FOG and fall between the two methane-based estimates –while degrading agreement with Barkley et al.⁶⁶.

Discussion

We present a space-based retrieval of tropospheric ethane based on thermal infrared radiance observations from the CrIS satellite sensor onboard Suomi-NPP. Spectral detection of ethane is first demonstrated using plume observations downwind of the January 2020 Australian wildfires. Other ethane enhancements seen from CrIS correspond with known fossil fuel extraction activities, and we show examples from four major production regions: the Permian basin (in Texas and New Mexico); the Hassi Messaoud oil field (in Algeria); the Burun oil field (in Turkmenistan); and the Ordos basin (in Northern China). The CrIS ethane HRIs are sensitive enough to reveal sub-basin variability that reflects the distribution of oil/gas extraction and processing facilities.

We then employ an artificial neural network to convert the CrIS ethane enhancements to atmospheric ethane column amounts, while accounting for the ancillary variables that affect the spectroscopy. Our ethane retrieval can reproduce 88% of the variance in the training set with low bias and an RMSE of 4.1×10^{15} molecules cm^{-2} , while comparison with ground-based column observations from the NDACC network reveals a significant correlation ($R^2 = 0.48\text{--}0.66$) and a mean CrIS/NDACC ratio of 0.61–0.65. We focus subsequent analysis on what the CrIS HRI observations reveal to be the most pronounced ethane hotspot on the globe: the Permian basin in Western Texas and Southern New Mexico. By combining our ROCRv2 ethane retrieval with GEOS-Chem using a mass-balance inversion technique, we obtain ethane emission estimates that increase with Permian fossil fuel production rates. Our analysis suggests that the Permian alone represents at least 4–7% of the total global fossil fuel-related ethane source; considering the CrIS:NDACC comparison results, this estimate may be conservative^{3,14}. The analysis in this paper represents a first step in using the CrIS observations to better understand global ethane emissions from fossil and non-fossil sources.

This Suomi-NPP CrIS dataset spans 12 years (2012–2023) of daily global ethane column observations, opening opportunities for understanding the underlying emission magnitudes, variability, and trends around the world. Suomi-NPP CrIS ceased LWIR operation in August 2023, but successor CrIS instruments are currently onboard the Joint Polar Satellite System JPSS-1/NOAA-20 and JPSS-2/NOAA-21 satellites and slated for launch onboard JPSS-3 and JPSS-4 –offering data continuity well into the 2030 s. This long-term record is a key community resource for advancing understanding of ozone, reactive nitrogen budgets, and fossil fuel emissions worldwide.

Methods

Instrumentation

CrIS is a Fourier transform spectrometer onboard the Suomi-NPP (launched 2011), JPSS-1/NOAA-20 (2017), and JPSS-2/NOAA-21 (2022) satellites, which are all in sun-synchronous low-Earth orbits with $\sim 01:30/13:30$ equator crossing times. Additional CrIS instruments are planned for the JPSS-3 and -4 missions. CrIS measures in three spectral bands covering the LWIR ($650\text{--}1095\text{ cm}^{-1}$), midwave IR ($1210\text{--}1750\text{ cm}^{-1}$), and shortwave IR ($2155\text{--}2550\text{ cm}^{-1}$). We employ here daytime-overpass LWIR data from Suomi-NPP, which is available from 02/2012 to 08/2023 at 0.625 cm^{-1} spectral resolution. The CrIS field of regard consists of a 3×3 pixel array each with 14-km nadir footprint diameter; the 2200 km cross-track scan width provides twice-daily global coverage. Single-pixel results from CrIS are averaged here to the MERRA-2 grid⁷⁰ prior to analysis using a drop-in-the-box method with no temporal or spatial co-adding.

CrIS spectral residual processing

We use CrIS spectral residuals processed with the Multi-SpEctra, MultiSpEctra, Multi-SEnsors (MUSES) algorithm from the Tropospheric Ozone and its Precursors from Earth System Sounding (TROPESS) project. The MUSES algorithm leverages over 20 years of heritage from the Aura TES optimal estimation algorithm. The strategy for the residuals differs from the standard TROPESS CrIS strategy as the latter aims to estimate the best values for all relevant absorbers, whereas our goal is to remove non-target absorber signals. Following the standard initial guess refinement for clouds and surface temperature⁷¹, the residual strategy used here therefore diverges from the standard TROPESS processing. A joint retrieval for H_2O , CO_2 , O_3 , NH_3 , peroxyacetyl nitrate (PAN), HDO, N_2O , CH_4 , cloud properties, surface temperature, and atmospheric temperature is performed. The radiative transfer also accounts for fixed climatological values for SO_2 , HNO_3 , OCS, HCN, SF_6 , HCOOH, C_2H_4 , CH_3OH , CFC-11, CFC-12, and CCl_4 . All spectroscopic parameters come from the AER line file (aer_v_3.8.1: https://github.com/AER-RC/AER_Line_File). These parameters were either obtained from the HITRAN database^{30,72,73} (this is the case for ethane) or were custom-derived at AER (for certain spectral ranges affected by H_2O , CO_2 , O_3 and O_2). The windows span 650 to 1750 cm^{-1} with gaps imposed where radiances do not fit well (e.g. due to spectroscopy). Following the joint step, cloud properties, surface properties, and atmospheric temperature are fixed, and PAN is retrieved between 780 and 790 cm^{-1} (with a gap between 783.125 and 786.250 cm^{-1} to window around water interference). Next, ozone and ammonia are updated with windows between 923.125 and 1317.5 cm^{-1} . Finally, the spectral residual is calculated from 650 to 1095 cm^{-1} and from 1215 to 1750 cm^{-1} , accounting for H_2O , HDO, CO_2 , O_3 , N_2O , CH_4 , SO_2 , NH_3 , HNO_3 , OCS, HCN, SF_6 , HCOOH, C_2H_4 , CH_3OH , PAN, CFC-11, CFC-12, and CCl_4 in the first window and H_2O , HDO, CO_2 , O_3 , N_2O , CH_4 , SO_2 , NH_3 , HNO_3 , HCN, HCOOH, C_2H_4 , CH_3OH , and PAN in the second.

Modeling the Permian Basin with GEOS-Chem

We use version 14.1 Classic (<https://doi.org/10.5281/zenodo.7600579>) of the GEOS-Chem CTM for our simulations of the Permian basin. Simulations are for 2019 and employ MERRA-2 meteorology, Community Emissions Data System global emissions⁷⁴, and US ethane emissions from the Tzompa-Sosa et al. inventory³, which is scaled from the 2011 US National Emissions Inventory⁷⁵ based on observational comparisons. Model runs apply ethane emission scale factors (1, 2, 5, 6, 7, 8, 10, 11) over the Delaware and Midland basins ($31^\circ\text{--}34^\circ\text{ N}$, $101^\circ\text{--}105^\circ\text{ W}$). A 6-month spinup is used to initialize year-long, scenario-specific global runs at $2 \times 2.5^\circ$ resolution that then provide boundary conditions for matching nested simulations over the Permian ($0.5^\circ \times 0.625^\circ$; $29^\circ\text{--}34^\circ\text{ N}$, $100^\circ\text{--}106^\circ\text{ W}$). Model columns are sampled daily from 12:00 to 15:00 LT to match the CrIS overpass. A scenario-specific P_{90} value for each column amount is then calculated from the model output and used for the CrIS-model comparisons.

Surface-based ethane column observations

We use 2013–2021 data from 13 sites (12 NDACC stations plus one candidate NDACC station at Xianghe, China) between 50°S - 60°N (Supplementary Fig. 4) and restrict the comparison to observations within $\pm 2\text{ h}$ of the local satellite overpass time. Observation pairs with skin temperatures below 270 K are omitted to avoid retrieval artifacts associated with snow or ice cover. We compare the NDACC columns to (1) the mean CrIS value for a 3×3 matrix of $0.5^\circ \times 0.625^\circ$ grid cells centered on the site in question (to reduce noise), and (2) to the CrIS value for the individual underlying grid cell (to minimize any signal suppression for near-source stations). Six of the sites used here (Altzomoni, Maidu, Mauna Loa, Rikubetsu, Izaña, and Zugspitze) are located at altitudes considerably higher than the average altitude over the corresponding CrIS pixels. We therefore apply a scaling correction to the CrIS values for these sites to account for this altitude difference

based on the shape of the modeled GEOS-Chem profile for each date and location.

Supplementary Fig. 6 plots the mean column averaging kernels for the employed NDACC datasets, showing sensitivity through the tropospheric column that increases with altitude. Also plotted in Supplementary Fig. 6 is the HRI/column ratio for the corresponding CrIS observations plotted as a function of the P_{90} value employed in the retrieval, reflecting the measurement sensitivity to the vertical location of ethane in the atmosphere. Simulations from GEOS-Chem suggest that the range of ambient ethane profiles encountered at these sites correspond to atmospheric P_{90} values of ~200–400 hPa (shaded regions in Supplementary Fig. 6B). Supplementary Fig. 7 shows the number of NDACC observations by site and year used in Fig. 4 and in Supplementary Figs. 5 and 6.

Data availability

The data required to replicate the figures in this work have been made publicly available as a capsule via the CodeOcean platform⁷⁶ at <https://doi.org/10.24433/CO.4014696.v1>. CrIS ethane columns, P_{90} values, HRI values, and processed NDACC and US EIA data used here are posted on CodeOcean. NDACC data is also publicly available at <https://ndacc.larc.nasa.gov/>. US EIA oil and gas production data is available publicly at <https://www.eia.gov/>. Additional data was taken from assorted other published papers^{3,46,47,50,65,66} and is also included in the CodeOcean repository linked above.

Code availability

The R code and libraries required to replicate the figures in this work have been made publicly available as a capsule via the CodeOcean platform⁷⁶ at <https://doi.org/10.24433/CO.4014696.v1>. The code can be examined and run interactively at that platform. GEOS-Chem model code is publicly available via <https://geoschem.github.io/>.

References

1. Simpson, I. J. et al. Long-term decline of global atmospheric ethane concentrations and implications for methane. *Nature* **488**, 490–494 (2012).
2. Wofsy, S. C. & ATom Science Team. ATom: Aircraft flight track and navigational data. 200.237798 MB <https://doi.org/10.3334/ORNLDAAC/1613> (2018).
3. Tzompa-Sosa, Z. A. et al. Revisiting global fossil fuel and biofuel emissions of ethane. *J. Geophys. Res. Atmos.* **122**, 2493–2512 (2017).
4. Fischer, E. V. et al. Atmospheric peroxyacetyl nitrate (PAN): A global budget and source attribution. *Atmos. Chem. Phys.* **14**, 2679–2698 (2014).
5. Rudolph, J. The tropospheric distribution and budget of ethane. *J. Geophys. Res. Atmos.* **100**, 11369–11381 (1995).
6. Tzompa-Sosa, Z. A. & Fischer, E. V. Impacts of emissions of C₂–C₅ alkanes from the U.S. oil and gas sector on ozone and other secondary species. *J. Geophys. Res. Atmos.* **126**, e2019JD031935 (2021).
7. Xiao, Y. et al. Global budget of ethane and regional constraints on U.S. sources. *J. Geophys. Res. Atmos.* **113**, D21306 (2008).
8. Pollack, I. B., McCabe, M. E., Caulton, D. R. & Fischer, E. V. Enhancements in ammonia and methane from agricultural sources in the northeastern Colorado Front Range using observations from a small research aircraft. *Environ. Sci. Technol.* **56**, 2236–2247 (2022).
9. Wunch, D. et al. Quantifying the loss of processed natural gas within California's South Coast Air Basin using long-term measurements of ethane and methane. *Atmos. Chem. Phys.* **16**, 14091–14105 (2016).
10. Ramsden, A. E. et al. Quantifying fossil fuel methane emissions using observations of atmospheric ethane and an uncertain emission ratio. *Atmos. Chem. Phys.* **22**, 3911–3929 (2022).
11. Helmig, D. et al. Reversal of global atmospheric ethane and propane trends largely due to US oil and natural gas production. *Nat. Geosci.* **9**, 490–495 (2016).
12. Turner, A. J., Frankenberg, C. & Kort, E. A. Interpreting contemporary trends in atmospheric methane. *Proc. Natl. Acad. Sci. USA* **116**, 2805–2813 (2019).
13. Hausmann, P., Sussmann, R. & Smale, D. Contribution of oil and natural gas production to renewed increase in atmospheric methane (2007–2014): top-down estimate from ethane and methane column observations. *Atmos. Chem. Phys.* **16**, 3227–3244 (2016).
14. Tribby, A. L. et al. Hydrocarbon tracers suggest methane emissions from fossil sources occur predominately before gas processing and that petroleum plays are a significant source. *Environ. Sci. Technol.* **56**, 9623–9631 (2022).
15. Wiedinmyer, C. et al. The Fire Inventory from NCAR version 2.5: An updated global fire emissions model for climate and chemistry applications. *Geosci. Model Dev.* **16**, 3873–3891 (2023).
16. Nicewonger, M. R., Aydin, M., Prather, M. J. & Saltzman, E. S. Large changes in biomass burning over the last millennium inferred from paleoatmospheric ethane in polar ice cores. *Proc. Natl. Acad. Sci. USA* **115**, 12413–12418 (2018).
17. Etiope, G. & Cicciooli, P. Earth's degassing: A missing ethane and propane source. *Science* **323**, 478–478 (2009).
18. Mazzini, A. et al. Relevant methane emission to the atmosphere from a geological gas manifestation. *Sci. Rep.* **11**, 4138 (2021).
19. Stein, O. & Rudolph, J. Modeling and interpretation of stable carbon isotope ratios of ethane in global chemical transport models. *J. Geophys. Res. Atmos.* **112**, D14308 (2007).
20. Rudolph, J. & Ehhalt, D. H. Measurements of C₂–C₅ hydrocarbons over the North Atlantic. *J. Geophys. Res. Oceans* **86**, 11959–11964 (1981).
21. Franco, B. et al. Evaluating ethane and methane emissions associated with the development of oil and natural gas extraction in North America. *Environ. Res. Lett.* **11**, 044010 (2016).
22. McKeen, S. A. et al. Hydrocarbon ratios during PEM-WEST A: A model perspective. *J. Geophys. Res. Atmos.* **101**, 2087–2109 (1996).
23. Wang, Y. & Zeng, T. On tracer correlations in the troposphere: The case of ethane and propane. *J. Geophys. Res. Atmos.* **109**, D24306 (2004).
24. Adedeji, A. R. et al. Measurement report: Assessment of Asian emissions of ethane and propane with a chemistry transport model based on observations from the island of Hateruma. *Atmos. Chem. Phys.* **23**, 9229–9244 (2023).
25. Dalsøren, S. B. et al. Discrepancy between simulated and observed ethane and propane levels explained by underestimated fossil emissions. *Nat. Geosci.* **11**, 178–184 (2018).
26. Payne, V. H. et al. Satellite measurements of peroxyacetyl nitrate from the Cross-track Infrared Sounder: comparison with ATom aircraft measurements. *Atmos. Meas. Tech.* **15**, 3497–3511 (2022).
27. Wells, K. C. et al. Satellite isoprene retrievals constrain emissions and atmospheric oxidation. *Nature* **585**, 225–233 (2020).
28. 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**, 13963–13984 (2018).
29. Harrison, J. J., Allen, N. D. C. & Bernath, P. F. Infrared absorption cross sections for ethane (C₂H₆) in the 3 μ m region. *J. Quant. Spectrosc. Radiat. Transf.* **111**, 357–363 (2010).
30. Gordon, I. E. et al. The HITRAN2016 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **203**, 3–69 (2017).
31. De Mazière, M. et al. The network for the detection of atmospheric composition change (NDACC): History, status and perspectives. *Atmos. Chem. Phys.* **18**, 4935–4964 (2018).
32. Ehhalt, D. H., Schmidt, U., Zander, R., Demoulin, P. & Rinsland, C. P. Seasonal cycle and secular trend of the total and tropospheric

- column abundance of ethane above the Jungfrauoch. *J. Geophys. Res. Atmos.* **96**, 4985–4994 (1991).
33. Franco, B. et al. Retrieval of ethane from ground-based FTIR solar spectra using improved spectroscopy: Recent burden increase above Jungfrauoch. *J. Quant. Spectrosc. Radiat. Transf.* **160**, 36–49 (2015).
34. González Abad, G. et al. Ethane, ethyne and carbon monoxide concentrations in the upper troposphere and lower stratosphere from ACE and GEOS-Chem: a comparison study. *Atmos. Chem. Phys.* **11**, 9927–9941 (2011).
35. von Clarmann, T. et al. MIPAS measurements of upper tropospheric C₂H₆ and O₃ during the southern hemispheric biomass burning season in 2003. *Atmos. Chem. Phys.* **7**, 5861–5872 (2007).
36. Zavalyov, V. et al. Noise performance of the CrIS instrument. *J. Geophys. Res. Atmos.* **118**, 13108–13120 (2013).
37. Franco, B. et al. Ethylene industrial emitters seen from space. *Nat. Commun.* **13**, 6452 (2022).
38. Walker, J. C., Dudhia, A. & Carboni, E. An effective method for the detection of trace species demonstrated using the MetOp Infrared Atmospheric Sounding Interferometer. *Atmos. Meas. Tech.* **4**, 1567–1580 (2011).
39. Wells, K. C. et al. Next-generation isoprene measurements from space: Detecting daily variability at high resolution. *J. Geophys. Res. Atmos.* **127**, e2021JD036181 (2022).
40. Whitburn, S. et al. A flexible and robust neural network IASI-NH3 retrieval algorithm. *J. Geophys. Res. Atmos.* **121**, 6581–6599 (2016).
41. Kloss, C. et al. Australian Fires 2019–2020: Tropospheric and stratospheric pollution throughout the whole fire season. *Front. Env. Sci.* **9**, 652024 (2021).
42. Worden, H. M. et al. TROPES/CrIS carbon monoxide profile validation with NOAA GML and ATom in situ aircraft observations. *Atmos. Meas. Tech.* **15**, 5383–5398 (2022).
43. Fu, D. et al. Retrievals of tropospheric ozone profiles from the synergism of AIRS and OMI: methodology and validation. *Atmos. Meas. Tech.* **11**, 5587–5605 (2018).
44. Worden, J. R. et al. Characterization and evaluation of AIRS-based estimates of the deuterium content of water vapor. *Atmos. Meas. Tech.* **12**, 2331–2339 (2019).
45. Sorenson, B. T. et al. Thermal infrared observations of a western United States biomass burning aerosol plume. *Atmos. Chem. Phys.* **24**, 1231–1248 (2024).
46. Zhang, Y. et al. Quantifying methane emissions from the largest oil-producing basin in the United States from space. *Sci. Adv.* **6**, eaa5120 (2020).
47. Varon, D. J. et al. High-frequency monitoring of anomalous methane point sources with multispectral Sentinel-2 satellite observations. *Atmos. Meas. Tech.* **14**, 2771–2785 (2021).
48. Chen, S. et al. Factors controlling the reservoir accumulation of Triassic Chang 6 Member in Jiyuan-Wuqi area, Ordos Basin, NW China. *Pet. Explor. Dev.* **46**, 253–264 (2019).
49. Enverus DrillingInfo. DI Desktop. (2019).
50. Scarpelli, T. R. et al. Updated Global Fuel Exploitation Inventory (GFEI) for methane emissions from the oil, gas, and coal sectors: evaluation with inversions of atmospheric methane observations. *Atmos. Chem. Phys.* **22**, 3235–3249 (2022).
51. Clough, S. A. et al. Atmospheric radiative transfer modeling: a summary of the AER codes. *J. Quant. Spectrosc. Radiat. Transf.* **91**, 233–244 (2005).
52. Clough, S. A., Iacono, M. J. & Moncet, J.-L. Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor. *J. Geophys. Res. Atmos.* **97**, 15761–15785 (1992).
53. Clough, S. A. et al. Forward model and Jacobians for Tropospheric Emission Spectrometer retrievals. *IEEE Trans. Geosci. Remote Sens.* **44**, 1308–1323 (2006).
54. Bey, I. et al. Global Modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *J. Geophys. Res.* **106**, 23073–23095 (2001).
55. 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).
56. Rinsland, C. P. et al. Infrared spectroscopic measurements of the ethane (C₂H₆) total column abundance above Mauna Loa, Hawaii—seasonal variations. *J. Quant. Spectrosc. Radiat. Transf.* **52**, 273–279 (1994).
57. Maddanu, F. & Proietti, T. Trends in atmospheric ethane. *Clim. Change* **176**, 53 (2023).
58. Zhou, M. et al. Understanding the variations and sources of CO, C₂H₂, C₂H₆, H₂CO, and HCN columns based on 3 years of new ground-based Fourier transform infrared measurements at Xianghe, China. *Atmos. Meas. Tech.* **16**, 273–293 (2023).
59. Zeng, G. et al. Trends and variations in CO, C₂H₆, and HCN in the Southern Hemisphere point to the declining anthropogenic emissions of CO and C₂H₆. *Atmos. Chem. Phys.* **12**, 7543–7555 (2012).
60. Vigouroux, C. et al. FTIR time-series of biomass burning products (HCN, C₂H₆, C₂H₂, CH₃OH, and HCOOH) at Reunion Island (21 S, 55 E) and comparisons with model data. *Atmos. Chem. Phys.* **12**, 10367–10385 (2012).
61. Lutsch, E., Dammers, E., Conway, S. & Strong, K. Long-range transport of NH₃, CO, HCN, and C₂H₆ from the 2014 Canadian wildfires. *Geophys. Res. Lett.* **43**, 8286–8297 (2016).
62. Yamanouchi, S. et al. Network for the Detection of Atmospheric Composition Change (NDACC) Fourier transform infrared (FTIR) trace gas measurements at the University of Toronto Atmospheric Observatory from 2002 to 2020. *Earth Sys. Sci. Data* **15**, 3387–3418 (2023).
63. Makarova, M. V. et al. Evaluation of emissions of atmospheric trace gases from remote sensing and local observations at the St. Petersburg site: anthropogenic emissions and wildfires. In *24th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics* (eds. Romanovskii, O. A. & Matvienko, G. G.) 39 (SPIE, Tomsk, Russian Federation, 2018). <https://doi.org/10.1117/12.2503106>.
64. Irakulis-Loitxate, I. et al. Satellite-based survey of extreme methane emissions in the Permian basin. *Sci. Adv.* **7**, eabf4507 (2021).
65. Varon, D. J. et al. Continuous weekly monitoring of methane emissions from the Permian Basin by inversion of TROPOMI satellite observations. *Atmos. Chem. Phys.* **23**, 7503–7520 (2023).
66. Barkley, Z. R. et al. Analysis of oil and gas ethane and methane emissions in the Southcentral and Eastern United States using four seasons of continuous aircraft ethane measurements. *J. Geophys. Res. Atmos.* **126**, e2020JD034194 (2021).
67. Francoeur, C. B. et al. Quantifying methane and ozone precursor emissions from oil and gas production regions across the contiguous US. *Environ. Sci. Technol.* **55**, 9129–9139 (2021).
68. US EIA. *Drilling Productivity Report*. <https://www.eia.gov/petroleum/drilling/index.php> (2023).
69. Caulton, D. R. et al. Abnormal tank emissions in the Permian Basin identified using ethane to methane ratios. *Elem. Sci. Anthro.* **11**, 00121 (2023).
70. Gelaro, R. et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **30**, 5419–5454 (2017).
71. Bowman, K. W. et al. *TROPES Level 2 Algorithm Theoretical Basis Document (ATBD) V1*. https://docserv.gesdisc.eosdis.nasa.gov/public/project/TROPES/TROPES_ATBDv1.1.pdf (2021).
72. Malathy Devi, V. et al. Multispectrum measurements of spectral line parameters including temperature dependences of N₂- and self-broadened half-width coefficients in the region of the v9 band of ¹²C₂H₆. *J. Quant. Spectrosc. Radiat. Transf.* **111**, 2481–2504 (2010).

73. Malathy Devi, V., Rinsland, C. P., Chris Benner, D., Sams, R. L. & Blake, T. A. Multispectrum analysis of the v9 band of $^{12}\text{C}_2\text{H}_6$: Positions, intensities, self- and N_2 -broadened half-width coefficients. *J. Quant. Spectrosc. Radiat. Transf.* **111**, 1234–1251 (2010).
74. McDuffie, E. E. et al. A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): an application of the Community Emissions Data System (CEDS). *Earth Sys. Sci. Data* **12**, 3413–3442 (2020).
75. EPA. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2012*. (2013).
76. Clyburne-Sherin, A., Fei, X. & Green, S. A. Computational reproducibility via containers in psychology. *Meta-Psychology* **3**, MP.2018.892 (2019).

Acknowledgements

This research was supported by the National Aeronautics and Space Administration (NASA; Grant #80NSSC21K1975, DBM, KCW, VHP, KCP; Grant #80NSSC21K1963, KCP, RP), the National Oceanic and Atmospheric Administration (NOAA; Grant #NA22OAR4310200, DBM, KCW), and the Minnesota Supercomputing Institute (MSI). Part of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract to NASA (#80NM0018D0004, VHP). We also acknowledge David Porter for help in architecting our training program. Ground-based data used in this publication were obtained from James Hannigan, Ivan Ortega, Nicholas Jones, Martine De Mazière, Justus Notholt, Dan Smale, Michel Grutter de la Mora, Wolfgang Stremme, Alejandro Bezanilla, Tomoo Nagahama, Isamu Morino, Kimberly Strong, Ralf Sussmann, Matthias Schneider, Omaira E. Garcia, and Maria Makarova as part of the Network for the Detection of Atmospheric Composition Change (NDACC) and are available through the NDACC website (<https://ndacc.org>). The NDACC FTIR observation programs at Boulder, CO and Mauna Loa, HI are operated by the National Center for Atmospheric Research (NCAR, sponsored by the National Science Foundation) and supported under contract by NASA. We also thank NOAA for support at the MLO station. The Toronto NDACC measurements were made at the University of Toronto Atmospheric Observatory, supported by the Natural Sciences and Engineering Research Council (NSERC) and by Environment and Climate Change Canada (ECCC). We acknowledge the support provided by the members of the Infrared Team of the Royal Belgian Institute for Space Aeronomy as well as the support provided by J.-M. Metzger and colleagues of the Observatoire de Physique de l'Atmosphère de la Réunion (OPAR). We also acknowledge the financial support that NDACC receives from the Belgian Science Policy Office through the ACTRIS-BE Grant (CV). The Xianghe FTIR ethane measurements are supported by the National Key Research and Development Program (#2023YFC3705202, MZ).

Author contributions

DBM, KCW, VHP, CV and KCP were responsible for funding acquisition. JFB, DBM, KCW, VHP and KCP designed the research. SK, MZ and KCP prepared raw data used in this work. JFB, KCW and RP designed the code used in this work. JFB, KCW and DBM performed the data analysis, and JFB and DBM wrote the paper. JFB, DBM, KCW, VHP, SK, CV, MZ and KCP edited the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-52247-z>.

Correspondence and requests for materials should be addressed to Dylan B. Millet.

Peer review information *Nature Communications* thanks Ira Leifer, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024