Hydrocarbon Tracers Suggest Methane Emissions from Fossil Sources Occur Predominately Before Gas Processing and That Petroleum Plays Are a Significant Source


Cite This: https://doi.org/10.1021/acs.est.2c00927

ABSTRACT: We use global airborne observations of propane (C3H8) and ethane (C2H6) from the Atmospheric Tomography (ATom) and HIAPER Pole-to-Pole Observations (HIPPO), as well as U.S.-based aircraft and tower observations by NOAA and from the NCAR FRAPPE campaign as tracers for emissions from oil and gas operations. To simulate global mole fraction fields for these gases, we update the default emissions’ configuration of C3H8 used by the global chemical transport model, GEOS-Chem v13.0.0, using a scaled C2H6 spatial proxy. With the updated emissions, simulations of both C3H8 and C2H6 using GEOS-Chem are in reasonable agreement with ATom and HIPPO observations, though the updated emission fields underestimate C3H8 accumulation in the arctic wintertime, pointing to additional sources of this gas in the high latitudes (e.g., Europe). Using a Bayesian hierarchical model, we estimate global emissions of C2H6 and C3H8 from fossil fuel production in 2016−2018 to be 13.3 ± 0.7 (95% CI) and 14.7 ± 0.8 (95% CI) Tg/year, respectively. We calculate bottom-up hydrocarbon emission ratios using basin composition measurements weighted by gas production and find that their magnitude is higher than expected and is similar to ratios informed by our revised alkane emissions. This suggests that emissions are dominated by pre-processing activities in oil-producing basins.

KEYWORDS: ethane, propane, natural gas, methane, energy

INTRODUCTION

Many studies have diagnosed recent methane (CH4) trends (both global and regional) using ethane (C2H6) atmospheric ratio signatures. However, the rejection of C2H6 by oil and gas producers (in the U.S., and presumably in countries following similar economic trends, Figure S1) results in an increase in the mole fraction of C3H8 in the natural gas pipelines. Thus, to the extent that losses occur in the pipelines and at the end users of natural gas, emissions of C3H8 may not necessarily directly reflect CH4 emissions, adding additional uncertainty in CH4 emission estimates from natural gas operations. In addition, a global uptick in hydraulic fracturing has shifted production from dry to wet fields, resulting in an increase in the ratios of both C3H8 and C2H6 to CH4, further complicating the use of the alkanes to diagnose the underlying CH4 emission sources.7−4

Given the uncertainty in using C2H6 alone as a tracer for CH4 emissions, we use both C3H8 and propane (C3H8) to diagnose whether significant CH4 emissions from natural gas and petroleum occur before gas processing. Unlike C2H6, C3H8 has a much higher market value and therefore does not undergo “rejection.”5 Provided downstream losses are minimal and the raw gas ratio of C3H8 to CH4 is known, C3H8 can provide a constraint for CH4 emissions from raw, unprocessed natural gas.

In this study, we employ global observations from aircraft, including the 2009−2011 High-Performance Instrumented Airborne Platform for Environmental Research (HIAPER) Pole-to-Pole Observations (HIPPO)6 and the 2016−2018 Atmospheric Tomography (ATom)7 missions, which provide vertical profiles of a variety of constituents, including C2H6 and C3H8, around the remote atmospheres of the globe. Together with the large-scale chemical transport model GEOS-Chem, we estimate global fossil emissions of C2H6 and C3H8.

Received: February 11, 2022
Revised: May 18, 2022
Accepted: May 18, 2022
Observations from the National Oceanic and Atmospheric Administration (NOAA) Global Monitoring Laboratory (GML). Measurements of CH₄⁸⁻⁻⁻⁻¹⁰, C₂H₆, and C₃H₈¹¹ from flask air collected by the NOAA GML tower¹² and aircraft¹³ near oil/natural gas basins were used as unprocessed gas references. We refer to site locations using state abbreviations. More information on data processing/spatial coverage is included in section S2, Table S1, and Figure S5. To better quantify geophysical variability and generate a confidence interval in the correlation between C₂H₆ and C₃H₈ mole fractions, we implement a pairs bootstrap to generate replicates of C₂H₆ and C₃H₈ observations. The CIs calculated from the bootstrapped samples are much broader than those calculated assuming the noise in the measurements is dominated by analytical errors. This suggests that geophysical noise induced by differences in transport and chemistry dominates the statistics. See section S2.1 for more details.

FRAPPE Observations. FRAPPE C130 flight data were taken within the Colorado Front Range between July 26 and August 19, 2014. We accessed the data on October 6, 2021 from www.air.larc.nasa.gov from the WAS C130 merge file. Our data processing for FRAPPE is similar to our methods for the NOAA in situ samples. A spatial illustration of FRAPPE observations is shown in Figure S6.

HIAPER Pole-to-Pole Observations and ATmospheric Tomography Data. The HIPOPO campaign was a sequence of five global measurement campaigns which sampled from near the North Pole to the coastal waters of Antarctica, covering different seasons between 2009 and 2011. Similarly, ATom took place from 2016 and 2018. Flight paths of HIPOPO and ATom campaigns are illustrated in Figure S22, and specific details about the data sources are included in section 3 of the Supporting Information (SI).

Only data observed at >20° north were used since the majority of emissions of these short-lived gases of interest lie in the northern hemisphere. The lifetime of C₂H₆ and C₃H₈ are on the order of a few months or shorter during the summer, and the time it takes for mixing between the northern to southern hemispheres is on the order of a year,¹⁴ so the mole fraction of C₂H₆ is most sensitive to their total emissions of C₂H₆ and C₃H₈. To exclude subtropical air, we only analyze measurements with tropopause pressure above 100 hPa (about 5% of the data were excluded under this constraint) for both ATom and HIPOPO, which was sufficient to reduce the influence of tropical intrusions. See Table S2 for a comprehensive outline of the filters we use.

As in other studies,¹⁶⁻¹⁷ we use potential temperature (θ, in units of Kelvin) in our analysis as a zonal coordinate. Potential temperature is conserved following adiabatic flow, and in the extratropics, variability within large-scale circulation can be well captured using this coordinate system. As a result, trace gases that have long lifetimes compared to synoptic-scale meteorology, which has a horizontal length scale on an order of 1000 km or more and a time scale of about 10 days,¹⁸⁻¹¹ will be well correlated with θ. Using θ as a dynamical coordinate allows us to more accurately compare low spatial resolution GEOS-Chem simulations with the aircraft in situ measurements (compared with simply using altitude and latitude coordinates, Figure S27). Potential temperature is not well-correlated with trace gases in the tropics or boundary layer, where moist convection and surface-drag-driven turbulence can result in nonunique pairs or when the photochemical lifetimes are short (summer).

GEOS-Chem Simulations. We simulated HIPOPO and ATom measurements using the GEOS-Chem “classic” 3-D chemical transport model with default settings (details about the simulations and emissions are provided in the SI, section 4). We use the same constraints as the aircraft observations, except we use a boundary layer height parameter. As described below, we use a Bayesian model to provide a best estimate for global emissions of C₂H₆ and C₃H₈ and their uncertainty. One contribution to the error estimate is transport errors in GEOS-Chem. To capture some of the uncertainty in the transport field, we sample the GEOS-Chem model several days before and after the in situ sampling time along the aircraft flight path, which we refer to as “synoptic replicates”. Finally, all GEOS-Chem simulations of C₂H₆ and C₃H₈ were interpolated on the vertical coordinate using θ to the aircraft measurements. As expected, GEOS-Chem synoptic replicates show less consistency in altitude (Figure S27), providing support for using θ as an analysis coordinate.

Bayesian Inference. We wish to quantify the global emissions of C₂H₆ and C₃H₈ using the observed mole fraction of these alkane during ATom and HIPOPO. The ambient mole fraction of C₂H₆ and C₃H₈ is most sensitive to their total northern hemisphere emissions during the winter/fall/spring when there is decreased sunlight/oxidation. As such, we assume differences between the GEOS-Chem simulations and the aircraft observations can be largely attributed to the underlying emissions grid, such that

\[ a = gcs \cdot \alpha \]  

where \( a \) is the aircraft C₂H₆ or C₃H₈ mole fraction, gcs is the GEOS-Chem simulation of C₂H₆ or C₃H₈ mole fraction, and \( \alpha \) is a scalar that represents the missing emissions of C₂H₆ and C₃H₈ from default emissions. We developed a Bayesian hierarchical model to estimate the missing emissions, where eq 1 forms the basis of our model. Our model only uses the GEOS-Chem simulated alkane mole fraction data (synoptic replicates), the alkane mole fractions observed by the aircraft, tropopause height, and UTC time. Our complete statistical model and its development, priors (Figures S37–S41), as well as the software used, are included in the SI, section 5.
Hydrocarbon Percent Composition Literature Compilation and Bootstrapping. We gather literature measurements of hydrocarbon composition from unprocessed gas from oil- and gas-producing basins in the U.S. and around the globe to calculate emission ratios. Summary statistics of the percent composition by region and the corresponding literature source is included in Tables S4 and S5 (SI). Gas composition varies significantly across basins, so we perform bootstrap calculations for data samples within each basin separately. For each basin, we draw random pairs of hydrocarbon composition measurements (C₂H₆, C₃H₈, and C₄H₁₀) for the size of the data set and then take the mean and repeat this 10,000 times. We use these bootstrap samples in subsequent calculations (eq 2) to arrive at emission ratios (Figure 6).

RESULTS AND DISCUSSION

CH₄ Leaks from U.S. Energy Activities Are Dominated by Emissions from Unprocessed Natural Gas. C₃H₈ and C₂H₆ are highly correlated at the NOAA sites (Figure 1). The data shown were obtained across 2–11 years and include tower and aircraft data (Table S1). The cross plot of C₂H₆ and C₃H₈ illustrates two distinct chemical regimes, similar to those described by Parrish et al.19 Above 1 ppb C₃H₈, the distribution is nearly linear, consistent with the mixing of fresh non-photochemically aged emissions into the background atmosphere. At mole fractions below 1 ppb, a second regime is defined by mixing of the aged emissions (the lifetime of C₃H₈ is much less than that of C₂H₆). To explore the characteristics of unprocessed natural gas emissions, we study the ratio of these gases within the 50th highest percentile (everything above 1 ppb C₃H₈) that would indicate fresh emissions. After this filtering, two sites, Northwestern CO and Western UT (site codes NWR and UTA), did not have any data in the fresh emission regime and are not included in further analysis (more detail in Figure S8).

Figure 1. Measurements of C₂H₆ and C₃H₈ from ongoing NOAA GML tower and aircraft sites (Table S1) from 2005 to 2018. The data follow the photochemical aging distribution described in Parrish et al., where the data below 1 ppb C₃H₈ are affected by photochemically aged emissions and mixing processes. As such, we only study the ratio of these gases in the 50th highest percentile (everything above 1 ppb C₃H₈) that would indicate fresh emissions. After this filtering, two sites, Northwestern CO and Western UT (site codes NWR and UTA), did not have any data in the fresh emission regime and are not included in further analysis (more detail in Figure S8).

Default GEOS-Chem Simulations Underestimate C₃H₈ Compared to Aircraft Observations. We compare the cross plot of C₃H₈ to C₂H₆ from the HIPPO and ATom aircraft measurements and GEOS-Chem simulations to the NOAA measurements (Figure 3). As expected, both the aircraft observations and GEOS-Chem simulations fall under ([0.62, 0.67] and [0.63, 0.71], respectively, ppb/ppb 95% CI, Figure S9) and over the entire time series (Figure S19). An unchanging C₃H₈/C₂H₆ ratio over time across the U.S. despite large changes in the C₃H₈/C₂H₆ ratio in processed gas during the same years (Figure S1) suggests that the majority of the alkane (and likely CH₄) emissions occur before the gas processing stage. (During gas processing, most of the C₂H₆ and sometimes much of the C₃H₈ are separated from the raw gas.) Conversely, post-processing (pipeline) composition of C₂H₆ at Playa del Rey in Southern California follows rejection trends, where the C₂H₆/C₃H₈ ratio has decreased by 8% from 2008 to 2018 and in recent years is about 18% lower in magnitude compared to the NOAA ratio (Figure S4).

Processed gas in California is a good representation of typical gas composition of domestic and globally imported consumer-grade gas.21 Our results are in agreement with Rutherford et al.’s U.S.-based model for CH₄ emissions, which finds that unintentional emissions from the production segment (namely, liquid storage tanks and other equipment leaks) are the largest contributors to divergence with the EPA’s GHGI.21

We use the same chemical aging approach to construct a background for NOAA and FRAPPE CH₄ observations (Figures S7 and S11). Since we only focus on the linear part of the chemical aging distribution, our analysis is not terribly sensitive to how the CH₄ anomaly is determined (it simply produces varying intercepts, Figure S10). Consistent with the analysis of Lan et al.2 (see Figures S20 and S21), the ratio of C₂H₆ to CH₄ has increased with time, reflecting a growing importance of oil exploration on CH₄ emissions in the U.S.

C₃H₈ is demonstrated to be a useful tracer that constrains oil- and gas-related CH₄ emissions. Given that the fit of C₂H₆ versus CH₄ (and C₃H₈ versus CH₄) have similar precision over the whole record, using C₃H₈ as a tracer likely separated nearby competing non-oil and gas CH₄ emissions. However, if there were to be non-oil and gas CH₄ emissions that were spatially coherent to the NOAA observation sites, our C₂H₆ versus CH₄ and C₃H₈ versus CH₄ emission ratios would be impacted. Investigating and potentially separating spatially coherent emissions of nonfossil origins would be the topic of future studies.

NOAA observations at Oklahoma ARM site are especially impacted by nearby unprocessed gas emissions, as C₂H₆ and C₃H₈ correlations with CH₄ have less noise compared to other sites (Figures S12 and S13). While the ratios of C₂H₆ and C₃H₈ versus CH₄ have increased by 50% at NOAA Oklahoma site since 2010, we find that both C₂H₆ and C₃H₈ versus CH₄ ratios are fractionally increasing at the same rate (Figure 2). That the atmospheric C₂H₆ and C₃H₈ increase fractionally the same suggests that the ratio of the alkanes in the reservoirs producing these emissions do not change significantly over the time of this record. Below, we use the 2017 average C₂H₆/CH₄ and C₃H₈/CH₄ from NOAA Oklahoma site ([0.060, 0.061] and [0.086, 0.088] ppb/ppb, respectively, Figure S21) to compare to ratios between our emission estimates for C₂H₆ and C₃H₈ to published estimates of CH₄ emissions from oil and gas exploration.

Environmental Science & Technology
the photochemically aged emissions part of the NOAA distribution. While the aircraft data overlay the NOAA measurements almost perfectly (especially in the winter when the lifetimes of both gases are the longest), GEOS-Chem underestimates C3H8, particularly over the Atlantic curtain (Figures 3 and 4). The same conclusion is drawn for HIPPO time periods (Figures S31, S32, and S34). Because the atmospheric lifetimes of C2H6 and C3H8 are different and vary seasonally due to the much higher concentrations of OH in the summer, our estimate of global C2H6 and C3H8 emissions from

![Figure 2](https://doi.org/10.1021/acs.est.2c00927)

**Figure 2.** Yearly correlation between NOAA hydrocarbon versus CH4 anomaly in Oklahoma. We show the percent change of anomalies/year with respect to the mean hydrocarbon and methane anomalies. The trend for C2H6/CH4 is 7.13 ± 1.44% with an R2 of 0.71. The trend for C3H8/CH4 is 5.87 ± 1.26% with an R2 of 0.69. The variability in the trend comes from the standard error of a linear regression. The variability in the individual points comes from the 95% confidence interval of a pairs bootstrap of the alkanes and CH4 anomalies. (We ran a pairs bootstrap for co-measurements of C2H6 and ΔCH4 and compute the slope of the correlation for each bootstrap sample and repeated this for every year in the data; please see the Materials and Methods section.) This trend in units of ppt/ppb/year is shown in Figure S21.

![Figure 3](https://doi.org/10.1021/acs.est.2c00927)

**Figure 3.** Comparison of C2H6 versus C3H8 for NOAA, ATom aircraft, and GEOS-Chem simulations during fall/winter seasons. NOAA photochemically aged measurements (all sites, 2005–2018), as explained in the text, are shown on the heat map (colored by the number density of data). The spring/summer seasons are included in Figures S29, S30, and S33. HIPPO is shown in Figure S34.

![Figure 4](https://doi.org/10.1021/acs.est.2c00927)

**Figure 4.** Impact of revised C3H8 emissions on GEOS-Chem simulation. Combined Pacific and Atlantic transects for ATom 4 aircraft campaign, which took place during Spring 2018, are shown in gold. The GEOS-Chem simulation using default C3H8 emissions are shown in blue and orange, referring to the Pacific and Atlantic transects, respectively. The GEOS-Chem simulation after implementing the revised C3H8 emissions is shown in green. The rest of the ATom and HIPPO campaigns are shown in Figures S35 and S36.

GEOS-Chem comparisons are sensitive to the a priori spatial distribution of these emissions.

Relative to C2H6, the default GEOS-Chem v13.0.0 C3H8 emissions result in a much larger underestimate of C3H8 mole fractions over the Atlantic transect compared to the Pacific (Figures 3 and 4), implying an underestimate over North America. This pattern is most clearly visible in the summer when the C3H8 lifetime is short. In contrast, the default C2H6 emissions produce a good simulation of the ATom data (within 5%) over both ocean basins (Figures 3 and 4). As such, we use the mean ratio observed in the linear regime of the NOAA data (0.67 C3H8/C2H6 ppb/ppb Figure S9) as the default global ratio of their emissions to update GEOS-Chem. In mass units (as used in the GEOS-Chem emissions), this is ≈0.99 kg C3H8/kg C2H6. Given the remarkable coherence in both the large-scale fields from the aircraft over both the Atlantic and Pacific transects and in the NOAA data, the spatial distribution of the emissions ratios for both gases must be very similar upwind of the Pacific (e.g., Asia) and the
Atlantic (North America). As such, in our revised emission fields for C\textsubscript{2}H\textsubscript{6}, we simply used the default C\textsubscript{2}H\textsubscript{6} emissions configuration used by GEOS-Chem v13.0.0. This scaling substantially altered the spatial distribution of C\textsubscript{2}H\textsubscript{6} emissions (Figure S28). The effect on the C\textsubscript{2}H\textsubscript{6} simulation is shown in Figure 4 using ATom 4 as an example, where updating the emissions resulted in a much better agreement between GEOS-Chem C\textsubscript{2}H\textsubscript{6} and aircraft measurements (other campaigns are included in Figures S33 and S36). Although simulations are greatly improved using the revised emissions, it appears there is a missing high latitude source of C\textsubscript{2}H\textsubscript{6} and C\textsubscript{3}H\textsubscript{8} (Figures S47, S48, S54, and S55).

**Bayesian Model Suggests Decadal Increase in Global C\textsubscript{2}H\textsubscript{6} and C\textsubscript{3}H\textsubscript{8} Anthropogenic Fossil Emissions.** The results of our Bayesian inference were satisfactory. We had good sampling of our posterior, and the sampling diagnostics were excellent (Figures S43, S44, S50, and S51). We performed tests that verified our inference procedure could capture the ground truth (using simulated data for which the ground truth is known), and that our posterior was more concentrated around the ground truth than the prior (Figure S42). Furthermore, our model could generate the measured data reasonably well; the majority of the measured aircraft data fell into the 30th and 50th percentile of the simulated Bayesian model data (Figures S47, S48, S54, and S55). The exception to this was the summer season, where the Bayesian model does not capture the measured aircraft data well. This is expected, since during the summer we do not observe a robust relationship between potential temperature and C\textsubscript{2}H\textsubscript{6} or C\textsubscript{3}H\textsubscript{8}. The model remained robust even when varying the sensitivity to low mole fractions of alkanes with tropical origin (Figures S58–S60).

During ATom 2 (winter) Atlantic curtain, the GEOS-Chem simulations poorly capture the observed C\textsubscript{2}H\textsubscript{6} and C\textsubscript{3}H\textsubscript{8} at low potential temperature compared to the aircraft. These measurements are samples obtained at low altitude, high latitude, and cold temperatures (Figure S49). During the winter, these arctic airmasses are often characterized by stagnant conditions, with less mixing with the midlatitudes.\textsuperscript{25} As a result, emissions that occur at high latitudes during the winter can be trapped there unless the zonal flow is disrupted. Additionally, emissions of C\textsubscript{2}H\textsubscript{6} and C\textsubscript{3}H\textsubscript{8} near the arctic during the winter will oxidize more slowly relative to midlatitudes due to the cold temperature and minimal sunlight. These conditions result in high C\textsubscript{2}H\textsubscript{6} and C\textsubscript{3}H\textsubscript{8} mole fraction over the arctic relative to remote midlatitude chemical regimes. The measurements subject to arctic conditions during ATom were too few to make a substantial impact on the overall Bayesian emission estimates, but during HIPPO winter flights, few samples were obtained, resulting in a large bias toward arctic data. We place much more weight on the ATom winter C\textsubscript{2}H\textsubscript{6} analysis estimate of global C\textsubscript{2}H\textsubscript{6} emissions since the arctic represents only a small fraction of the atmosphere.

It is likely that GEOS-Chem v13.0.0 is missing a high latitude emissions source (Figures S47–S49, S54, and S55). Underestimation of C\textsubscript{2}H\textsubscript{6} and C\textsubscript{3}H\textsubscript{8} at high latitudes is consistent with other studies, which found that fossil fuel emissions from Eurasia accounted for the largest underestimation.\textsuperscript{23} It is possible that emissions from northern Europe may account for this discrepancy, as fossil emissions were found to be underestimated,\textsuperscript{27} and our revised C\textsubscript{2}H\textsubscript{6} emissions decreased in that region after implementing the emission C\textsubscript{2}H\textsubscript{6} proxy (Figure S28). This, combined with a relatively lower number of HIPPO aircraft observations at lower latitudes, results in a substantial positive bias on the overall Bayesian emissions estimate for C\textsubscript{2}H\textsubscript{6} during winter 2009 (Figure S62).

We report our Bayesian estimates for each seasonal campaign and ocean transect during 2009–2011 and 2016–2018 and what the GEOS-Chem v13.0.0 default emissions grids should be scaled by, according to our analysis (Figures S45, S46, S52, and S53 and SI section 5.8). We estimate global emissions of C\textsubscript{2}H\textsubscript{6} and C\textsubscript{3}H\textsubscript{8} from fossil fuel production from 2016 to 2018 to be 13.3 \pm 0.7 (95% CI) and 14.7 \pm 0.8 (95% CI) Tg/year, respectively. Our results compare well to other studies (Figure 5 and Figure S62). Our estimates suggest emissions of C\textsubscript{2}H\textsubscript{6} have increased by about 15% from 2010 to 2017 when comparing the mean revised C\textsubscript{2}H\textsubscript{6} emissions during those time periods (Figure S61). Emissions of C\textsubscript{3}H\textsubscript{8} are calculated to have increased more (65%, Figure S62), but this estimate is highly uncertain due to the few samples obtained during HIPPO and the impact of the Arctic winter pooling in both campaigns. Nevertheless, these increases are consistent with greater oil production emissions contribution. Similarly, Helmig et al., which used data from a global surface network and atmospheric column observations, found about a 22% increase in C\textsubscript{2}H\textsubscript{6} emissions between 2009 and 2014.\textsuperscript{24}

**Oil Exploration Plays a More Significant Role in Global CH\textsubscript{4} Compared to Dry Gas.** We calculate an emission ratio of C\textsubscript{3}H\textsubscript{8}/CH\textsubscript{4} for n basins or countries using the following equation:

\[
E\left\{\frac{C_k}{C_i}\right\} = \sum_{i=1}^{n} \left( P_{DNG} \times \frac{C_{i,\text{tot}}}{C_i} + P_{DNG} \times \frac{C_{k,\text{tot}}}{C_k} \right)
\]

where \(k\) represents anthropogenic fossil emissions, except for the work by Dalsøren et al., which also includes biofuel, agriculture, and waste. We obtained the CEDS CMIP6 estimate from Dalsøren et al. Our emissions estimates do not include biomass burning or biofuels. Propane emissions are included in Figure S62.
where \( P_{\text{NGC}} \) is the dry natural gas production (in million tonnes), \( C_1 \) and \( C_3 \) are the bootstrapped samples of measured hydrocarbon fractions in raw natural gas samples (in mass%, details on the bootstrapping in the Materials and Methods section), and \( \text{tot} \) is the sum of the bootstrapped samples of measured hydrocarbon fractions for \( \text{CH}_4 \), \( \text{C}_2\text{H}_6 \), and \( \text{C}_3\text{H}_8 \).

(Note that our \( E\left(\frac{C_1}{C_3}\right) \) emission ratios are inherently weighted by natural gas production by basin, \( P_{\text{NGC}} \), \( E\left(\frac{C_1}{C_3}\right) \) is calculated similarly. We refer to the emission ratio in eq 2 as the “literature” emission ratio, since we combine a variety of natural gas composition measurements for wet and dry basins.

We calculate a global literature emission ratio using hydrocarbon and dry natural gas production data from the top 5 producing natural gas basins around the world that made up 50% of the total natural gas production in 2019.\(^{26,29}\) We also calculate a U.S. literature emission ratio using the top 7 natural-gas-producing basins that account for 86% of total U.S. natural gas production.\(^{30,31}\) We show the relative production of the top global and U.S. basins used in our analysis in Figure S3.

Additional summary statistics and sources for the composition measurements are included in Tables S4 and S5.

Separately, we calculate an “observationally informed emission ratio (OIER)” by taking the ratio between our revised \( \text{C}_3\text{H}_8 \) emissions with literature estimates of \( \text{CH}_4 \) emissions from oil and natural gas processes. The OIER for \( \text{C}_3\text{H}_8/\text{CH}_4 \) is calculated similarly. Previous studies have constrained global \( \text{CH}_4 \) emissions from natural gas/petroleum systems to range from 63 to 91 Tg/year.\(^{25,32−37}\) Given our estimate for revised \( \text{C}_3\text{H}_8 \) and \( \text{CH}_4 \) emissions, this implies a mean alkane ratio of 100: [8.0, 10.0]: [6.5, 7.3] molar % (\( \text{CH}_4/\text{C}_3\text{H}_8/\text{C}_2\text{H}_6 \) in 2016−2018. We compare our literature ratio with several OIER using global estimates in Figure 6 and with U.S. estimates in Figure S6. The small abundance of spatial and temporal literature measurements of raw gas composition throughout basins most affect the final uncertainty in the emission ratio comparison.

Unprocessed dry gas has a smaller \( \text{C}_3\text{H}_8/\text{CH}_4 \) and \( \text{C}_2\text{H}_6/\text{CH}_4 \) ratio compared to unprocessed associated gas from oil-producing basins (“wet” gas). A greater contribution of emissions from dry basins would decrease the magnitude of the overall literature ratio, assuming minimal \( \text{C}_3\text{H}_8 \) leakage after separation from raw gas, given the high market value of \( \text{C}_3\text{H}_8 \). In the U.S., the basin with the highest gas production is the Appalachian (East Coast) (Figure S3). If \( \text{CH}_4 \) leaks were proportional to production, we would expect a “dry” (small) emission ratio that resembles the composition of the Appalachian region (6% mass/mass of \( \text{C}_3\text{H}_8/\text{CH}_4 \) calculated from Table S5). However, we find the production-weighted literature ratio to be much larger than expected (15% mass/mass for \( \text{C}_3\text{H}_8/\text{CH}_4 \), Figure S6). The second-largest gas-producing region, the Permian (Southwestern U.S.), is also the largest oil producer in the U.S. and vastly powers the Appalachian in terms of oil production (Figure S3). The magnitude of our production-weighted raw gas “literature” emission ratios suggest a significant contribution from wet gas and emissions that are biased toward oil-producing basins. We find similar results for global emission ratios (Figure 6).

The Global Carbon Project \( \text{CH}_4 \) emissions estimate implies an OIER that is dry relative to the production-weighted raw gas ratio (literature ratio) (Figure 6). Instead, the IEA (76 Tg/year) and Scarpelli et al. (66 Tg/year) \( \text{CH}_4 \) emission estimates yield an OIER that is within a few percent of the production-weighted raw gas literature ratio, given our revised \( \text{C}_3\text{H}_8 \) and \( \text{C}_2\text{H}_6 \) emissions. Both of those studies estimate oil production emissions to have a relatively higher contribution to the global footprint compared to dry gas production.

The FRAPPE and NOAA Oklahoma observed emission ratios compare well to the global OIER (Figure 6) and U.S. OIER (Figure S6), suggesting high emissions from oil production. Indeed, there are substantial oil production activities (Figures S15−S18) surrounding the NOAA Oklahoma and FRAPPE observation sites. Increasing trends in the Oklahoma emission ratios are consistent with production trends: oil production in Oklahoma tripled from 2010 to 2017\(^{38,39}\) (compared to doubling of gas production), and in 2020, Oklahoma was the fourth-largest oil producer in the U.S.\(^{40}\) (Note that Oklahoma was not included in calculations for the literature ratio, since Oklahoma natural gas production does not rank in the top 7.) Several factors may reduce incentive or the ability for oil producers to capture associated

Since our findings suggest that CH$_4$ losses are likely greater and biased toward oil-producing sites, a significant fraction of bottom-up estimates of CH$_4$ emissions may be misallocated to dry CH$_4$ production when they should instead be included with the oil production sector. Correctly attributing CH$_4$ emissions to oil production would increase the greenhouse gas footprint of petroleum-based transportation, while decreasing the greenhouse gas emissions ascribed to natural-gas-powered power plants. At a minimum, the CO$_2$ equivalent footprint of the global transportation sector would increase by roughly 5%, using IEA’s estimate of 76 Tg/year CH$_4$ emissions from oil and natural gas and recent transportation CO$_2$ emission estimates (section 6.2, SI).\footnote{2} This estimate will only increase when accounting for vented and flared losses of associated natural gas that is not accounted for in marketed associated gas (which we use to calculate these numbers).

### ASSOCIATED CONTENT

**Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.2c00927.

U.S. and global economic and production trends of oil and natural gas; statistical methods and chemical aging approach to NOAA and FRAPPE observations; methane anomaly determination for NOAA and FRAPPE observations; yearly methane anomaly ratios for NOAA observations; ATom and HIPPO aircraft campaign filtering methods; GEOS-Chem simulation results after revising propane emissions; Bayesian inference results; emission ratio calculation details; analysis of reallocating methane emissions to the transportation sector (PDF)

### AUTHOR INFORMATION

**Corresponding Author**

Paul O. Wennberg -- Division of Engineering and Applied Science and Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, United States; orcid.org/0000-0002-6126-3854; Email: wennberg@caltech.edu

### References

5. U.S. Energy Information Administration. Hydrocarbon gas liquids explained; https://www.eia.gov/energyexplained/hydrocarbon-gas-


