Analysis of Oil and Gas Ethane and Methane Emissions in the Southcentral and Eastern United States Using Four Seasons of Continuous Aircraft Ethane Measurements

Z. R. Barkley[1], K. J. Davis[1], S. Feng[1], Y. Y. Cui[1], A. Fried[2], P. Weibring[2], D. Richter[2], J. G. Walega[2], S. M. Miller[3], M. Eckl[5], A. Roiger[5], A. Fiehn[5], J. Kostinek[5]

[1]The Pennsylvania State University, University Park, PA, USA
[2]Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, USA
[3]Department of Environmental Health and Engineering, Johns Hopkins University
[4]National Aeronautics and Space Administration, Langley Research Center, Hampton, VA, USA
[5]Deutsches Zentrum für Luft- und Raumfahrt e.V., Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

Key Points:

- This study uses ethane observations to quantify both ethane and methane emissions from the United States oil and gas sector.
- Ethane emissions in the central and eastern United States are larger than existing inventories by more than a factor of 2.
- Ethane-methane ratios indicate that the US EPA methane inventory is underestimating leak rates from the oil and gas sector by at least 50%.

Corresponding author: Zachary R. Barkley, zrb5027@psu.edu

Abstract

In the last decade, much work has been done to better understand methane (CH_4) emissions from the oil and gas (O&G) industry in the United States. Ethane (C_2H_6), a gas that is co-emitted with thermogenic sources of CH_4, is emitted in the US almost entirely by the O&G sector. In this study, we perform an inverse analysis on 300 hours of atmospheric boundary layer C_2H_6 measurements to estimate C_2H_6 emissions from the US O&G sector. Measurements were collected from 2017-2019 as part of the Atmospheric Carbon and Transport (ACT) America aircraft campaign and encompass much of the central and eastern United States. We find that for the fall, winter, and spring campaigns, C_2H_6 data consistently exceeds values that would be expected based on EPA O&G leak rate estimates. C_2H_6 observations from the summer 2019 dataset show significantly lower C_2H_6 enhancements in the southcentral region that cannot be reconciled with data from the other three seasons, either due to complex meteorological conditions or a temporal shift in the emissions. Converting the fall, winter, and spring season posterior C_2H_6 emissions estimate to an inventory of O&G CH_4 emissions, we estimate that O&G CH_4 emissions are larger than EPA inventory values by more than 50%. Uncertainties in the gas composition data limit the effectiveness of using C_2H_6 as a proxy for O&G CH_4 emissions. These limits could be resolved retroactively by increasing the availability of industry-collected gas composition data.

Plain Language Summary

Methane is a potent greenhouse gas responsible for a quarter of the warming the climate has experienced thus far. The oil and gas sector is a significant source of methane through leaks in its infrastructure. Recent studies of individual basins have found emissions from oil and gas in the US to be greater than inventory estimates, but difficulties arise with source attribution in broader scale studies due to the numerous potential sources of methane. This study quantifies methane emissions from oil and gas by looking at atmospheric ethane, a gas whose emissions stem almost entirely from oil and gas in the US. Hundreds of hours of ethane observations were collected via aircraft over the course of 4 seasons between 2017-2019. These observations are compared with model-projected ethane values based on our current knowledge of ethane emissions, and those emissions are adjusted to best match the observed data. We find ethane emissions are grossly underestimated in the US. Because ethane is co-emitted with oil and gas methane sources, this underestimation of ethane reflects similar underestimations in oil and gas methane emissions. We conclude that US inventories are underestimating methane emissions from oil and gas by more than 50%.

1 Introduction

Methane (CH_4) is an important greenhouse gas with 28-35 times the warming potential of carbon dioxide over a 100 year period (Myhre et al., 2013). Global CH_4 concentrations in the atmosphere have nearly tripled since pre-industrial times, mainly driven by anthropogenic activity and are responsible for a fourth of the increased radiative forcing on the planet (Myhre et al., 2013). Although CH_4 concentrations stabilized for a brief period in the early 2000s, global concentrations began increasing again by 2007, with an increasing growth rate continuing through present date (Nisbet et al., 2019).

Efforts to understand causes for increasing global CH_4 trends are hampered by difficulties related to source attribution (Saunois et al., 2020). CH_4 has numerous anthropogenic sources, including animal agriculture, fossil fuel extraction, and waste management. In addition to anthropogenic emitters, there are natural sources of CH_4 emissions that play a large role in the global CH_4 budget. Of particular importance to the global CH_4 budget are emissions from anaerobic respiration in wetlands, which create unique challenges to CH_4 source attribution. Unlike most anthropogenic sources, whose emis-

sions can be spatially mapped out using inventory techniques and are relatively consistent in magnitude on an annual timeframe, wetland emissions are more difficult to describe, with an uncertain spatial pattern and large seasonal variability in emissions based on soil temperature and moisture (Yvon-Durocher et al., 2014). Many prior studies have examined seasonal emissions and magnitudes of local wetland sources using flux tower measurements (Grant & Roulet, 2002; Matthes et al., 2014), but extrapolating these results to a continental or global scale is challenging.

Difficulties with source attribution extend into the United States, where both anthropogenic and natural sources play a large role in the country's CH_4 budget. In the US, greenhouse gas emissions from anthropogenic sources are quantified through a bottom-up inventory created by the United States Environmental Protection Agency (EPA). The largest sources of anthropogenic CH_4 emissions projected by this inventory are from leaks in oil and gas (O&G) infrastructure, enteric fermentation and manure management related to livestock, and anerobic respiration occurring in landfills (US Environmental Protection Agency, 2020). These sources were responsible for 83% of US anthropogenic CH_4 emissions in the 2018 inventory estimate. Natural sources in the US are dominated by wetlands and are not tracked by the EPA's inventory. WetCHARTs, a global wetland emissions ensemble, is generally used as the prior for CH_4 emissions from wetlands in the US (Bloom et al., 2017; Sheng et al., 2017; Maasakkers et al., 2016). WetCHARTs ensemble members estimates the country's wetland emissions to vary from values that are insignificant to totals rivaling those from anthropogenic emissions, with seasonal dependence and spatial variability between ensemble members, resulting in large uncertainties in the overall US CH_4 budget.

Contrary to recent global increases in the atmospheric growth rate of CH_4, the EPA's inventory estimates of anthropogenic CH_4 emission in the US show a 10% decrease in the last decade, from 28 Tg in 2008 to 25 Tg in 2018. This decrease is driven primarily by projected decreases in emissions from the energy sector, despite a 50% increase in gas production and a >200% increase in oil production during the 10 year period (US Energy Information Administration, 2020b). Over the last several years, various atmospheric studies monitoring emissions from O&G from individual wellpads (Rella et al., 2015; Robertson et al., 2017; Caulton et al., 2019), basins (Karion et al., 2015; Barkley et al., 2017; Peischl et al., 2018), and entire regions (Barkley et al., 2019b) have consistently found emission rates larger than the EPA inventory, raising concerns of a broad underestimation of leaks from the O&G sector (Alvarez et al., 2018). However, large-scale CH_4 inversion studies involving the US have not been as conclusive, with differing opinions, as to the accuracy of inventory emissions from O&G, animal agriculture, and wetlands (Sheng et al., 2018; Yu et al., 2020; Maasakkers et al., 2019). The enormous spread of uncertainty regarding the magnitude, spatial distribution, and seasonality of CH_4 emissions from wetlands, as well as the numerous other potential sources of CH_4, poses a challenge to large scale studies attempting source attribution of detected CH_4 signals.

One common method to disaggregate certain sources is to measure both CH_4 and ethane (C_2H_6) concentrations (McKain et al., 2015; Barkley et al., 2019a). Sources of C_2H_6 coincide with thermogenic CH_4 emitters (O&G extraction, biomass burning), but not biogenic sources (wetlands, animal agriculture, landfills). Thus, a correlation (or lack of) between C_2H_6 and CH_4 can provide information on the sector responsible for observed CH_4 enhancements. While global C_2H_6 emissions are split between the energy sector and biomass burning (Xiao et al., 2008), the National Emissions Inventory (NEI) 2011 C_2H_6 inventory attributes 95% of C_2H_6 emissions in the US to the energy sector, making it a simple tracer for identifying O&G emissions in the region ((US Environmental Protection Agency, 2014)). This technique has been used in various smaller-scale aircraft mass balance studies across individual O&G basins to verify that the measured enhancements are associated with O&G activity (Peischl et al., 2018; Roscioli et al., 2015). Sim-

ilarly, C_2H_6 concentrations from a network of flask measurements have been used to infer trends in emissions from O&G emissions on broader scales (Xiao et al., 2008; Tzompa-Sosa et al., 2017).

Advancements in technology have resulted in the increased availability of instrumentation capable of measuring C_2H_6 concentrations precisely at high temporal resolution (Weibring et al., 2020; Kostinek et al., 2019; Yacovitch et al., 2014), expanding our capabilities of solving for C_2H_6 emissions. The Atmospheric Carbon and Transport - America (ACT-America) mission was a 5-season aircraft campaign across the central and eastern US designed to examine various trace gases and their transport in the atmosphere. During four of the seasonal campaigns, continuous C_2H_6 measurements were collected, producing more than 500 hours of C_2H_6 data capable of capturing various plume structures related to O&G activity on a regional scale. In this study, we examine the characteristics of this unprecedented dataset and use it to quantify C_2H_6 emissions compared to bottom-up inventory estimates. These C_2H_6 emissions are then combined with US gas composition data to infer CH_4 leak rates from major US O&G basins.

2 Methods

2.1 ACT-America Aircraft Campaign and Observational Dataset

The ACT-America campaign was a NASA Earth Venture suborbital aircraft mission with flight activities spanning summer 2016 to summer 2019. During this time, 5 individual campaigns covering all 4 seasons (summer twice) were conducted using two aircraft collecting meteorological, greenhouse gas, and trace gas data within fair weather and frontal weather patterns. For all 5 campaigns, continuous CH_4 data were collected on both aircraft using a commercial PICARRO G2401-m instrument adapted with a custom inlet system for drying and conditioning the sample air (DiGangi et al., 2018). During the Winter 2017, Fall 2017, Spring 2018, and Summer 2019 campaigns, continuous C_2H_6 data were collected on the B200 aircraft using the CAMS-2 (Compact Airborne Multi-Species Spectrometer) instrument (Weibring et al., 2020). The CAMS-2 C_2H_6 measurements when averaged over time and linearly regressed versus NOAA portable flask packages collected during the flights yielded slopes in the 0.990 to 1.031 range across seasonal campaigns, with r^2 values between the two measurements of 0.996 (Baier et al., 2020; Weibring et al., 2020), providing high confidence in the accuracy of its measurements. Additionally, during the Fall 2017 and Summer 2019 campaigns, the C130 aircraft was equipped with a quantum and interband cascade laser spectrometer (QCLS) capable of continuous in situ C_2H_6 measurements (Kostinek et al., 2019). The QCLS instrument performed in-flight two-point calibrations every three to ten minutes, ensuring accurate measurements throughout the flights. Altogether, the entire C_2H_6 dataset contains more than 500 hours of continuous airborne C_2H_6 observations, of which 300 hours were within the atmospheric boundary layer (ABL), making it an ideal dataset to study atmospheric C_2H_6 plumes and structures throughout the central and eastern US.

From the observational dataset, large C_2H_6 plumes were consistently observed downwind of Texas/Oklahoma/Louisianna in the southcentral US and along the western Appalachians in the northeast, corresponding with O&G activity in each of these regions (Figure 1). C_2H_6 plumes associated with southcentral O&G sources could be observed in the midwestern flights as far north as Minnesota when consistent southerly winds were present. C_2H_6 mixing ratios within the plumes were largest during the fall and winter campaign, likely related to the low boundary layer mixing depth during those seasons. During the Summer 2019 campaign, large C_2H_6 plumes were detected in the western Appalachian similar to other seasons, but signals in the southcentral were substantially smaller in a way that boundary layer depth alone cannot explain. Additionally, boundary layer

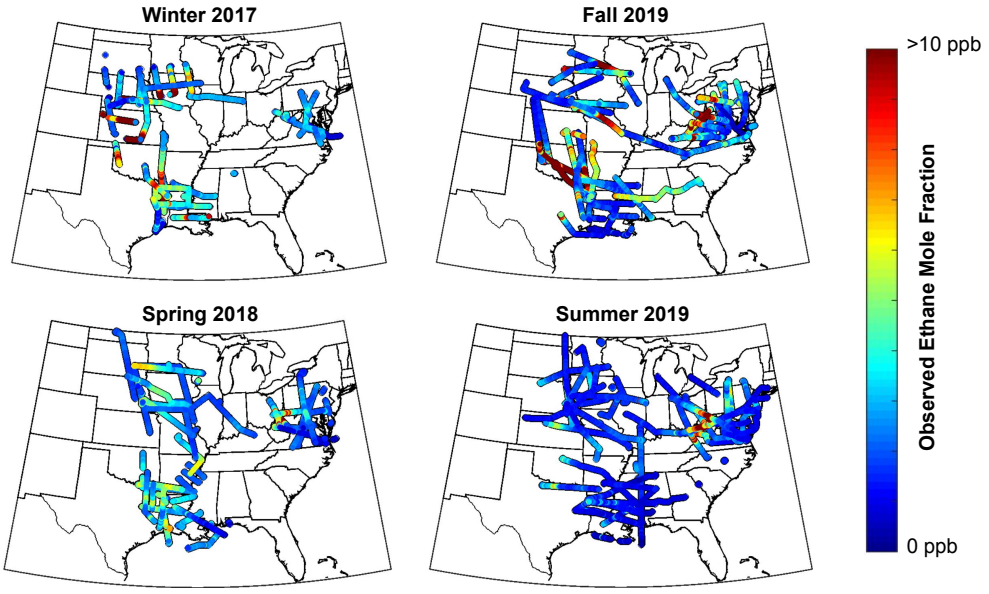

Figure 1. Continuous boundary layer C_2H_6 measurements collected by the B200 and C130 aircraft during each seasonal campaign. Only data within the atmospheric boundary layer (¡1000 m AGL) are shown.

C_2H_6 concentrations in the summer were often similar in value to their free tropospheric counterparts, a trait not observed during other seasons (Figure S2).

Influence functions for observations were created using the Lagrangian particle dispersion model FLEXPART-WRF (Brioude et al., 2013) to provide information on the area captured by the ACT-America C_2H_6 dataset. The model was run at 27 km resolution over the North American domain and was meteorologically driven using WRF-Chem simulations developed as part of the ACT-America campaign (Feng et al., 2019). Every 30 seconds of flight time, 5,000 particles were released from the aircraft location and traced back 10 days in time. Further details on the development of the influence functions can be found in Cui et al. (2015).

Averaging the influence functions provides information on which O&G basins have substantial coverage by the ACT flight campaign (Figure 2). Much of the midwestern and southcentral US is captured across all four seasons, encompassing many of the major O&G basins including Haynesville, Fayetteville, Anadarko, Eagle Ford, Woodford, and the Barnett. A sharp decrease in the influence map can be observed west of central Texas. For this reason, we do not include the Permian basin in western Texas/New Mexico in our analysis of the southcentral US and consider it and other western basins (Bakken, Denver-Julesburg, four-corners region, California) to be outside the scope of this study. In the northeast, many flight tracks were designed to fly downwind of the Marcellus and Utica gas basins in the western Appalachian regions of West Virginia and western Pennsylvania, and thus this area has significant coverage in the influence maps. An exception to this occurs during the winter campaign, where many of the downwind measurements were performed by the C130 aircraft which did not have an C_2H_6 instrument at the time, resulting in only 3 flights with direct C_2H_6 measurements downwind of the region (see Figure S3 for a seasonal breakdown of the influence functions).

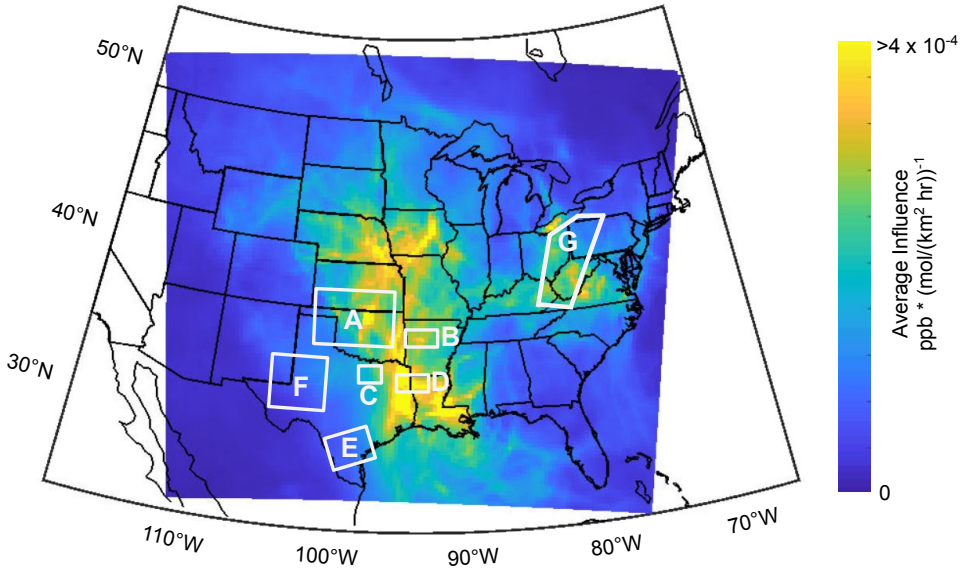

Figure 2. Averaged influence function from the boundary layer observations used in the C_2H_6 inversion. Brighter colors indicate areas whose surface interactions were captured more frequently by the boundary layer observations. O&G basins pertinent to this study are highlighted in boxes and are labeled as follows: (A.) Anadarko. (B.) Fayetteville. (C.) Barnett. (D.) Haynesville. (E.) Eagle Ford. (F.) Permian. (G.) Appalachia.

2.2 Inversion Framework

Observational C_2H_6 data collected from the ACT-America campaign is used in this study to perform an inversion and learn about C_2H_6 emissions from O&G sources in the flight domain. The basic inversion framework used is similar to previous work (Lauvaux et al., 2012; Sheng et al., 2018) and is expressed by cost function,

$$J(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{y} - \mathbf{H}\boldsymbol{x})^T\mathbf{R}^{-1}(\boldsymbol{y} - \mathbf{H}\boldsymbol{x}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x_0})^T\mathbf{B}^{-1}(\boldsymbol{x} - \boldsymbol{x_0}) \quad (1)$$

In this equation, we solve for a posterior emissions grid $\boldsymbol{x}$ that minimizes the cost function J using influence functions ($\mathbf{H}$) that translate the flux field to a modelled enhancement ($\mathbf{H}\boldsymbol{x}$). In the cost function, two terms control the solution. The first term is a cost related to the mismatch between the posterior modelled enhancements versus observed concentrations ($\boldsymbol{y}$), with greater discrepancies resulting in a larger cost term. Here, $\mathbf{R}$ is the observation error covariance matrix, and weights the first term based on the confidence in the observations and model transport. The second term in the cost function equation is a cost related to the change between the posterior flux ($\boldsymbol{x}$) and the prior flux ($\boldsymbol{x_0}$), with larger changes resulting in a greater cost. Here, $\mathbf{B}$ is the flux error covariance matrix, and weights the second term based on the confidence in the prior flux field. Minimizing the cost function with respect to $\boldsymbol{x}$ yields

$$\boldsymbol{x} = \boldsymbol{x_0} + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\boldsymbol{y} - \mathbf{H}\boldsymbol{x_0}) \quad (2)$$

and solving for $\boldsymbol{x}$ yields the posterior flux field. For this study, observations from each season are grouped together and an posterior flux map is solved for each of the seasonal flight campaigns.

A best guess C_2H_6 emissions map was created to serve as a prior for the inversion (Figure 4). To create this prior, CH_4 emissions from the O&G sector of the EPA Gridded 2012 CH_4 Emissions Inventory (Maasakkers et al., 2016) were multiplied by expected molar C_2H_6/CH_4 ratios of each basin (Table S2), resulting in an C_2H_6 emissions map. For Eagle Ford, Haynesville, Fayetteville, Barnett, Permian, Denver-Julesburg, and the Bakken O&G basins, atmospheric measurements from NOAA aircraft studies were available to derive observed basin-wide C_2H_6/CH_4 ratios (Peischl et al., 2015, 2018; Tzompa-Sosa et al., 2017). For the Anadarko and Appalachian region where representative atmospheric measurements were not available, data from the United States Geological Survey were used to create a spatial map of C_2H_6/CH_4 ratios to apply to these regions (US Geological Survey, 2018; Kitanidis, 1997)(See supplemental section S1 for additional info). For all emissions related to transmission, storage, and distribution, an C_2H_6/CH_4 ratio of 0.027 was applied (Plant et al., 2019). For prior C_2H_6 emissions in Canada and Mexico, we use CH_4 emissions from O&G facilities provided in Sheng et al. (2017) and apply the mean C_2H_6/CH_4 ratio from the USGS datbase of 0.085 to convert it to C_2H_6 emissions. These sources outside the US have minimal representation in our influence functions and have no impact on the overall solution. Although biomass burning and biofuels are also significant sources of C_2H_6 on a global scale, these sources are estimated to be only a few percent of the total C_2H_6 emissions in the US (Tzompa-Sosa et al., 2017). Overall, this C_2H_6 emissions map represents our best guess as to representing C_2H_6 emissions based on the EPA's bottom-up O&G CH_4 emissions inventory, and will be referred to as the "Default" map henceforth.

Two additional C_2H_6 emission maps were created to test the sensitivity of the inversion to different priors. The first alternative map was created by taking the EPA CH_4 emissions map used in the creation of the Default inventory and applying a flat C_2H_6/CH_4 ratio of 0.085. This ratio preserves the total C_2H_6 emissions from the "Default" map but redistributes them in a way that removes knowledge of the unique gas compositions of different basins, and is referred to henceforth as the "Flat Rate" inventory. The second alternative map is based on ethane emissions provided by the US National Emissions Inventory 2011 inventory (NEI2011) C_2H_6, an C_2H_6 map available in GEOS-Chem and used in Tzompa-Sosa et al. (2017). In addition to having a different spatial distribution compared to the Default inventory, total emissions in the NEI2011 inventory are roughly half the Default total. C_2H_6 emissions from the Emissions Database for Global Atmospheric Research (EDGAR) v4.3.2 were originally considered as well for this analysis, but total emissions were 8 times lower than our Default inventory and were decided to be too inaccurate to serve as a useful prior (Huang et al., 2017).

Continuous boundary layer C_2H_6 measurements from the ACT-America campaign were used as observational input for this study, with their associated FLEXPART-derived influence functions used for $\mathbf{H}$ in the inversion (section 2.1). Observations greater than 1000 meters above ground level were not included in the analysis to remove any data above the ABL. For this study, we restrict the domain of our inversion to the area within the 4 corners (23.7°N, 110.72°W), (23.0°N, 77.5°W), (49.9°N ,-67.3°W), (51.1°N, -119.0°W, domain shown as the colored region in Figure 2). Because the influence functions only provide a local enhancement inside the study domain, for each flight date a background value is determined to represent the C_2H_6 mixing ratios entering the domain. This value is chosen by taking the 5th percentile of the observed boundary layer C_2H_6 concentrations on a given flight and subtracting it from the observations, producing an observed C_2H_6 enhancement. The 5th percentile of model-projected enhancements along the flight track is then added onto the observed enhancements in order to align the modelled and observed background values. This final step is necessary in rare scenarios where mod-

elled O&G enhancements are influencing the entire aircraft transect, thus impacting the observed background concentrations (see Barkley et al. (2019b) for more details). In calculating the modelled C_2H_6 enhancements for this study, we treat C_2H_6 as an inert gas rather than a reactive one due to its long average lifetime (weeks to months) relative to the length of time the local plumes travel from the source to the aircraft (hours to days). For more information on the possible influence of C_2H_6 loss rates, see supplemental section S1.

To run an inversion, values must be assigned to the **R** and **B** matrices related to the uncertainty in the observation/transport and prior flux fields respectively. For the observational/transport uncertainty matrix **R**, we use a method similar to the residual error method discussed in Sheng et al. (2017). For each flight, modelled enhancements are first scaled by a constant to have the same overall enhancement as the observed enhancements. This step is performed to remove any existing overall bias that may exist in the prior inventory for the calculation of **R**. After removing this bias, the residual errors are calculated between the model and observation and the standard deviation of this error is used to represent the **R** value along its diagonal for a given flight, with no value assigned for the off-diagonals. Values for the diagonal elements of **R** varied across flights, but seasonal averages for the standard deviation of the error ranged from 0.7 ppb in the spring to 1.8 ppb in the fall. This process results in flights with large observational and transport uncertainty on days with large enhancements and poor correlations between the observed and modelled values, thus giving these flights less weight in the inversion solution. Similarly, flights where observed and modelled plume structures align have a smaller R value assigned and thus are given greater weight in the overall solution. This method for classifying transport uncertainty is particularly effective for an C_2H_6 inversion study, since the locations of the sources (i.e. O&G infrastructure) are known with high confidence such that misaligned plumes would most likely be caused by errors in the transport rather than problems with the spatial mapping of the flux.

For the flux uncertainty matrix **B**, there lacks a clear answer on what the uncertainty of the prior fluxes should be. Since the primary source of C_2H_6 in the US is from O&G production and processing, the location of C_2H_6 emitters in the US should be accurate. However, leak rates from O&G activity have been shown in various studies to be beyond the uncertainty bounds of the EPA 2012 Gridded CH_4 Inventory (Alvarez et al., 2018; Maasakkers et al., 2016). For this reason, a value for the flux uncertainty **B** is selected based on mathematical constraints of the inversion rather than an understanding of the prior flux map. Specifically, all flux grids are assigned an error as a percentage of their prior, where the percent uncertainty is selected using a chi-squared metric, defined as

$$\chi^2 = \frac{1}{m}(\boldsymbol{y} - \mathbf{H}\boldsymbol{x_0})^T(\mathbf{HBH'} + \mathbf{R})^{-1}(\boldsymbol{y} - \mathbf{H}\boldsymbol{x_0}) \tag{3}$$

where m is the number of observations and χ^2 is the chi-squared metric used to assess whether the inversion errors satisfy a Gaussian distribution. Here, we select a percent error for the flux uncertainty in **B** that brings χ^2 close to 1. For the Default, Flat Rate, and NEI2011 inventory, the assigned flux errors along the diagonal of **B** were 50, 50, and 80 percent of the total grid emissions respectively. Additionally, a correlation length with an e-folding decay length of 50 km is assigned to the off-diagonal elements in **B**. This added correlation provides consistent shifts in the emissions within a basin, while allowing changes across basins to behave uniquely. Implications of the **R** and **B** matrices selected for this study are explored further in the sensitivity analysis (Table S4).

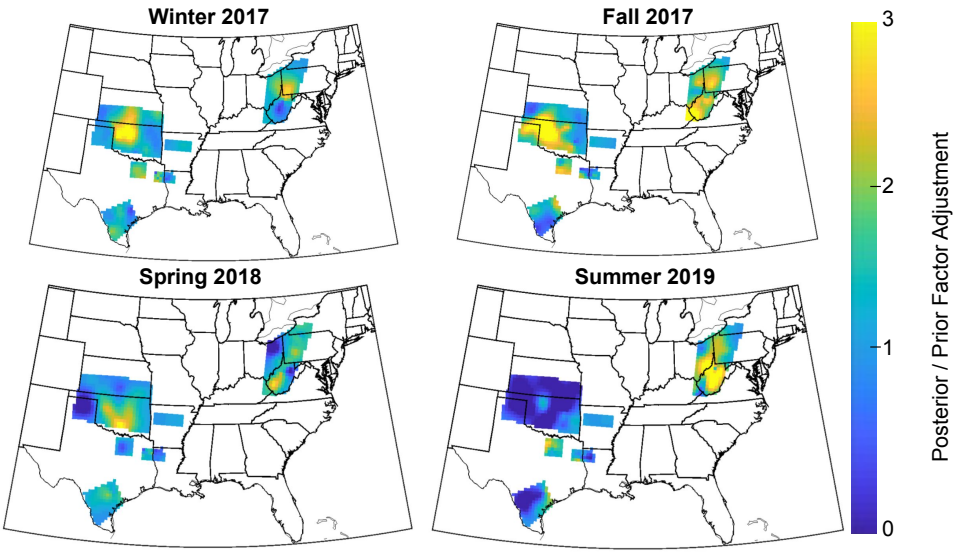

Figure 3. Fractional changes between the Default posterior and prior C_2H_6 flux map created by an inversion conducted for each individual season. Only basins within the scope of this study are shown on the map.

3 Results and Discussion

3.1 C_2H_6 Inversion Results

For each season, posterior C_2H_6 flux maps were successful in reducing both the error and bias and increasing the correlation between the observed and modelled signal for that season (Table 1, Figures S4-S9). Notably, across all 4 seasons there was an overall low bias between the modelled prior and observed C_2H_6 enhancements, resulting in seasonal posterior maps that generally increased C_2H_6 emissions in order to compensate. Regionally, observed C_2H_6 plumes were underestimated the largest in the western Appalachian region across all 4 seasons (Figure 3). In this region, posterior C_2H_6 emissions were increased by 50 to 150% more than the prior to correct for the underestimation. In the southcentral US, a similar low bias was observed in the winter, fall, and spring campaign. This led to posterior solutions for these three seasons that show a systematic increase to the total posterior flux of 50%. This increase is mostly uniform across the individual O&G basins, with a notable exception for the Haynesville basin whose posterior is consistently within 11% of the Default prior across all seasons.

Observations from the southcentral and midwestern US during the Summer 19 campaign behave like outliers compared to the rest of the dataset. Of the 87 flights used in the inversion analysis, only 12 had an overall negative bias when comparing the observed C_2H_6 to the model prior (observed enhancement less than modelled enhancement), 8 of which occurred in the southcentral and midwestern portion of the Summer 19 campaign. These low observations have a profound effect on the posterior emissions in the southcentral US for the Summer 19 campaign, with total emissions from this region that are less than a third of posterior emission maps from the other three seasonal campaigns. To demonstrate how unrepresentative the Summer 19 results are compared to the rest of the dataset, we take the posterior inventory derived for each season and apply it to the entire 4 season dataset (Table 2). In doing this, we find that the winter, fall, and spring posteriors all produce similar statistical improvements to the overall dataset compared

	Default Inventory Winter 2017 Performance		Default Inventory Fall 2017 Performance		Default Inventory Spring 2018 Performance		Default Inventory Summer 2019 Performance	
	Prior	**Posterior**	**Prior**	**Posterior**	**Prior**	**Posterior**	**Prior**	**Posterior**
Southcentral Total C_2H_6 (mol s^{-1})	488	709	488	777	488	601	488	199
Western Appalachia Total C_2H_6 (mol s^{-1})	125	184	125	294	125	185	125	269
Mean Absolute Error (ppb)	1.39	0.99	1.65	1.16	0.68	0.54	0.85	0.59
Mean Bias (ppb)	1.20	0.61	1.50	0.76	0.55	0.37	0.25	0.20
y,Hx Correlation	0.85	0.90	0.75	0.83	0.61	0.72	0.44	0.74

Table 1. Table describing the performance of each of the seasonal posteriors relative to the observations from that season.

to the prior. However, when applying the summer C_2H_6 posterior to the 4 season dataset, not only does it perform substantially worse than each of the other seasonal posteriors, it also performs worse than the prior, with increases to the absolute error, bias, and a substantial decrease in the model-obs correlation coefficient from 0.71 to 0.50.

One possible explanation for the discrepancy in summer could be related to errors in the model transport simulation. The mean absolute error in the modelled boundary layer wind direction compared to flight observations were on average 30 degrees per flight in the summer campaign versus 15 in each of the other three seasons, likely related to the slower wind speeds observed during summer (Table S3). These directional transport errors can create misalignments in observed versus modelled plumes, leading to an overall reduction in the posterior emissions relative to the truth. Compounding on these wind issues is the possibility of increased convective activity in summer. While boundary layer statistics are tracked in the model and compared to observations, it is more difficult to assess how much of the signal is lost out the top of the boundary layer due to summertime convective processes. Another, non-transport related possibility is that there could be a non-trivial loss of C_2H_6 due to an increase in the OH chemical sink during the summer months. In this study, the impacts of C_2H_6 loss were not considered, as C_2H_6 has an average lifetime of two months and most plumes were captured within 48 hours of release from the source (Burkholder et al., 2015). However, in conditions with excessive heat and large OH concentrations that can occur in the summer, the lifetime of C_2H_6 can be reduced to as little as four days in the most extreme conditions (see supplemental section S2 for more details). Even so, a lifetime of four days would only have minor impacts on local plumes, and the extreme conditions leading to high C_2H_6 loss would only exist for a short period in the afternoon hours. Furthermore, summer C_2H_6 observations from the northeast were elevated to levels similar to those observed during the other seasons, seemingly unaffected by summertime conditions. For these reasons, it is unlikely that a chemical loss could be responsible for the low values observed in the southcentral US. One final possibility is that the low C_2H_6 enhancements observed in the Summer 19 campaign are due to a real and significant temporal change in C_2H_6 emissions in the southcentral US during this period, particularly in the Anadarko basin. However, the near-zero emission rate solution provided by the inversion posterior in the Anadarko lacks any sort of real-world explanation for such a large shift compared to previous sea-

	Default Inventory Prior	Default Inventory Winter 2017 Posterior	Default Inventory Fall 2017 Posterior	Default Inventory Spring 2018 Posterior	Default Inventory Summer 2019 Posterior
Southcentral Total C_2H_6 (mol s^{-1})	488	709	777	601	199
Western Appalachia Total C_2H_6 (mol s^{-1})	125	184	294	185	269
Four Season Mean Absolute Error (ppb)	1.15	1.02	0.96	1.04	**1.18**
Four Season Mean Bias (ppb)	0.86	0.21	0.23	0.58	**0.94**
Four Season y,Hx Correlation	0.71	0.73	0.78	0.71	**0.50**

Table 2. Table describing the performance of each of the individual seasonal posterior fluxes when each is applied relative to all four seasons of observations. Green areas highlight statistical improvement compared to the prior, whereas red boxes show degradation.

sons. Production data shows the Anadarko basin was at its peak O&G production rates during the period (US Energy Information Administration, 2020a). Regardless of the reason, the low C_2H_6 observations measured in the southcentral during the Summer 19 months are not representative of data from the winter, fall, and spring campaigns, which show strong consistency in the location and expected magnitudes of regional C_2H_6 enhancements. For this work, we choose to discard summer data in evaluating our best estimate of US C_2H_6 emissions, but consider it a source of uncertainty and intrigue for future research.

To create our best guess regarding C_2H_6 emissions across the southcentral and eastern US, we weight the Default posterior flux maps from winter, fall, and spring with equal confidence and average them together (Figure 4). This averaged posterior solution provides consistent improvement compared to the prior, reducing the absolute error between the model and observations in 84% of flights, providing confidence that this averaged solution serves as an improvement to a large majority of the observational data for the three seasons it represents. C_2H_6 emissions from the 3 season posterior are almost universally increased compared to the prior, with a 43% increase overall in the southcentral US and a 76% increase in the Appalachia. With these large increases in the emissions the mean obs-model bias is reduced, decreasing from 1.14 ppb to 0.70 ppb. The inability to eliminate this bias is a natural result of errors in the transport preventing the inversion from solving for misaligned observed plumes, and thus the emission increases in the 3 season posterior are likely still an underestimation of the true emissions.

To better understand the sensitivity of our 3 season posterior to the prior, the inversion was rerun using the two alternative prior maps discussed in section 2.2 and a 3 season posterior was created from each set of priors (Figures 4 and 5) For every basin, posteriors from the three inventories converged towards a similar solution compared to their priors. An example of this is the Haynesville basin, where the two alternative prior inventories have a factor of 4 spread between their emissions, but their posteriors con-

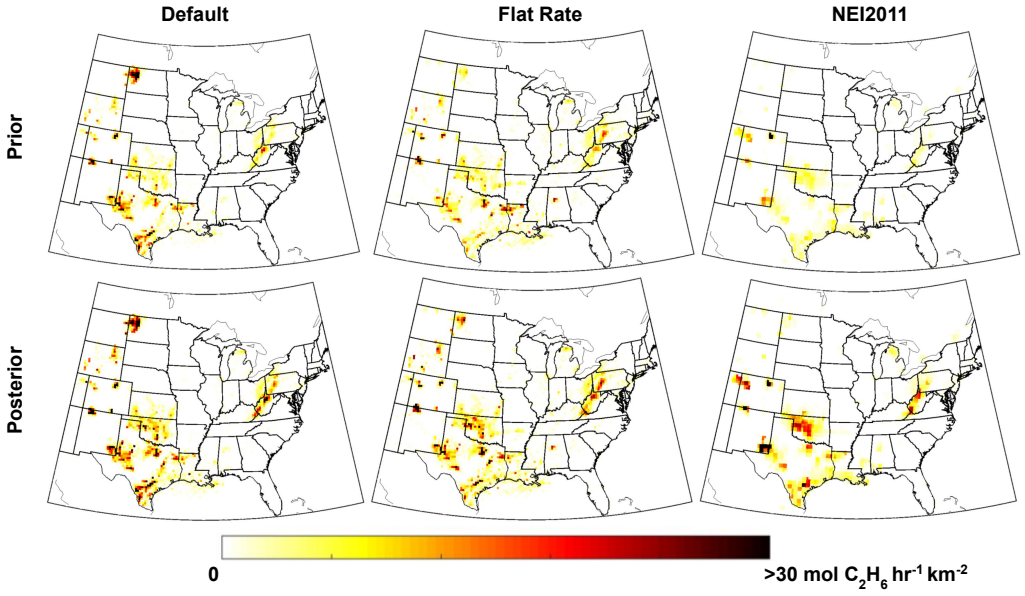

Figure 4. Prior C_2H_6 inventories used in this study and their respective 3 season mean posteriors. "Default" represents the best guess prior from this study based on multiplying the O&G sector of the Gridded EPA 2012 CH_4 Emissions Inventory by projected C_2H_6/CH_4 ratios of individual basins. "Flat Rate" multiplies the same CH_4 inventory by a flat rate C_2H_6/CH_4 ratio of 0.085, producing a similar total as "Default" with a different spatial representation. "NEI2011" comes from the NEI 2011 C_2H_6 inventory.

verge to within 20% of the Default posterior solution. Ultimately, the 2 alternative posteriors produced solutions for the entire southcentral US that were within 1 mol/s of each other, despite starting 183 mol/s apart. Their similar solutions for the southcentral region are 12% lower than the posterior calculated using the Default inventory. This difference is driven primarily by the Eagle Ford basin, which had the largest absolute spreads between prior inventories and whose western portion lies on the edge of the region of influence captured by the ACT-America flights (Figure 2), making it more difficult for the inversion to constrain using the available dataset. In the northeast, there was a similar consensus among the three inventories of a substantial increase in emissions required to recreate the large plumes observed there. While the emissions in the NEI2011 posterior do not increase to the same levels as the other two priors (171 vs 221 and 239 mol/s), the NEI2011 prior in the northeast was missing nearly all of the emissions from conventional gas activity in western Pennsylvania, resulting in a posterior map that still appears to be underestimating the sources in that region even after substantial increases given the large positive bias between the observations and modelled posterior. Overall the NEI2011 prior, whose total C_2H_6 emissions are 53% less than the Default prior, leads to a total posterior C_2H_6 emissions estimate that is only 14% lower than the Default posterior.

In addition to testing the sensitivity of the inversion posterior to different priors, we also test its sensitivity to numerous other conditions, including adjustments to the magnitude of the prior, changes to the selection of the background C_2H_6 term, elimination of the length correlations applied to the prior flux uncertainty, adjustment of the observation matrix error term to a constant (giving equal uncertainty to all observations), and solving for a posterior solution using all three seasons of data simultaneously rather

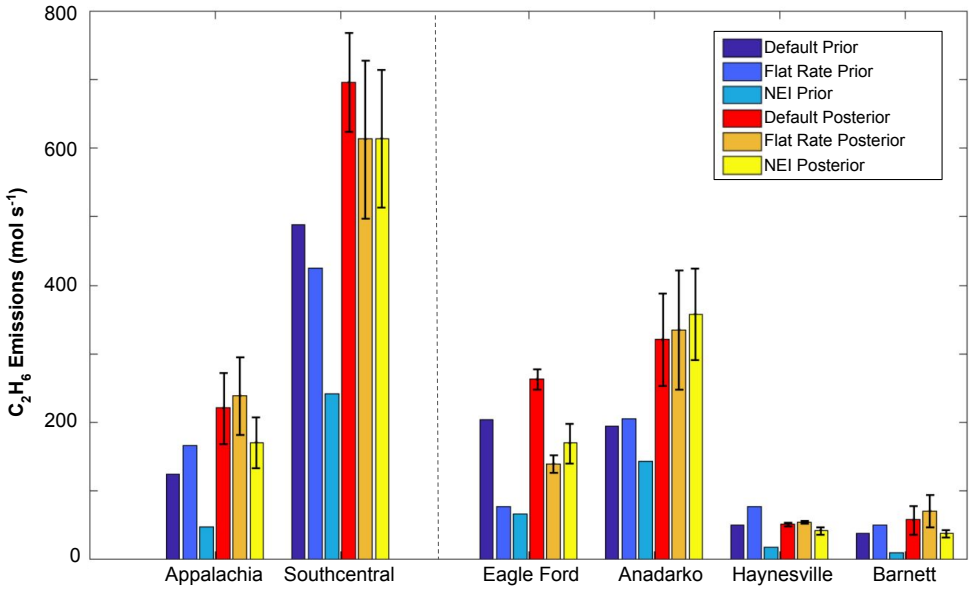

Figure 5. Priors and their respective 3 season posterior C_2H_6 emissions for major basins in this study. "Southcentral" is the sum of Eagle Ford, Anadarko, Haynesville, and Barnett. Error bars show the standard deviation of the individual season solutions for each basin.

than averaging 3 unique posterior maps (Table S4). All resulting posteriors produce similar results compared to the Default posterior, providing further confidence in the solution. In particular, using our Default prior multiplied by a factor of 3 converged to a similar solution as our Default posterior, reducing concerns that using a prior with a low bias could be resulting in an large underestimation of the solution. One source of uncertainty that is difficult to address in our sensitivity tests are errors in the transport used to create the influence functions. Average flight wind speed biases were under 1.3 m/s and wind direction absolute errors were close to 15 degrees for the winter, fall, and spring campaigns when compared to ACT-America boundary layer aircraft data (Table S3). However, modelled boundary layer heights were 15 to 22 percent lower on average seasonally compared to observations. A negative bias in the modelled boundary layer would result in an equally proportionally higher value in the influence functions and model-projected enhancements. Applying a correction factor to the influence functions relative to the mean bias for each season and rerunning the inversion produces an overall posterior that is 22% higher than the Default posterior. Consistent with our past work that has applied corrections to simulated errors in ABL depth and winds (Barkley et al., 2017, 2019a, 2019b), we consider the posterior created using the ABL depth correction to be an equally plausible solution in our best estimate of the C_2H_6 emissions, as the boundary layer bias is a source of error with a known and somewhat correctable bias on the solution. The boundary layer bias, along with a negative enhancement bias still present in the model versus observational comparison of the posterior C_2H_6 enhancements, are both potential reasons to suspect the Default posterior on its own may still be underestimating overall C_2H_6 emissions despite the large increase in the emissions relative to the prior.

3.2 Interpretation of CH_4 Emissions from the O&G Sector

The Default prior inventory developed in this study is created by multiplying the EPA 2012 Gridded CH_4 Emissions Inventory for O&G sources by the suspected C_2H_6/CH_4

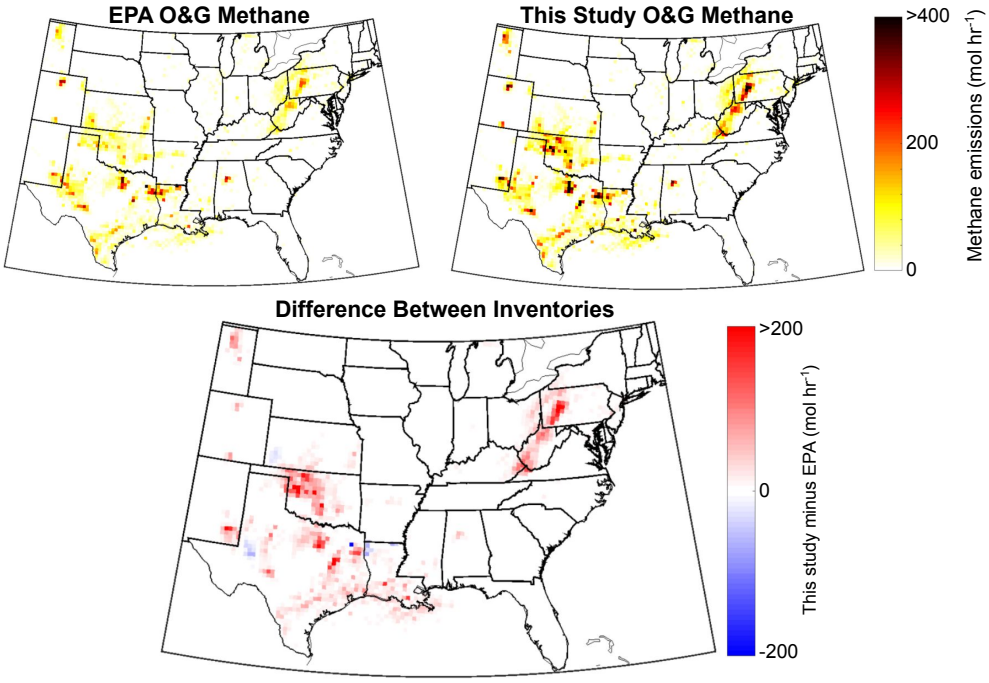

Figure 6. Top-left: Oil and gas CH_4 emissions from the EPA 2012 Gridded CH_4 Inventory. Top-right: Oil and gas CH_4 emissions estimated from the C_2H_6 posterior in this study. Bottom: Difference between the two inventories.

ratio of each grid's emissions. If the assumed ratios are correct, and emissions primarily occur in situations where the gas content is unaltered (i.e. processes unrelated to gas separation), then changes between the posterior and prior Default C_2H_6 inventory should proportionally reflect changes in the EPA 2012 CH_4 inventory for O&G. Thus, we can use our Default posterior C_2H_6 inventory to create our best interpretation of O&G CH_4 emissions in the southcentral and eastern US.

Using the converted C_2H_6 posterior as a proxy for O&G CH_4 emissions, our inventory projects that O&G CH_4 emissions are almost universally larger than the 2012 EPA Inventory estimates (Figure 6). In the southcentral US, we estimate emissions to be 48% higher than inventory estimates (77% using the ABL-adjusted posterior). Of this increase, two-thirds of it is driven by increases in the Anadarko basin, which had the largest proportional change in the C_2H_6 posterior and is the largest source of CH_4 emissions in the region. Of the four remaining southcentral basins captured in this study, the Haynesville basin is the only basin that did not see a significant increase in its emissions relative to the prior. As noted in section 2.1, our solution does not include potential changes from the Permian basin which is outside of the region of influence captured by ACT-America campaign. O&G production in the Permian has increased by more than a factor of 3 since 2012 and some of the largest CH_4 signals from the basin have been measured using satellite observations in the area (Zhang et al., 2020). As such, it is likely that emissions in the Permian follow a similar pattern of underestimation observed for other southcentral basins in our study.

In the Appalachians, we estimate CH_4 emissions to be 77% higher than EPA 2012 inventory estimates (105% using the ABL-adjusted posterior), the largest difference of any basin in this study. Part of the discrepancy between inventory results and the posterior may be related to the increased presence of unconventional natural gas activity

in the Marcellus shale. Between 2012 and 2018 Pennsylvania and West Virginia underwent some of the largest gas production growth in the US, with annual production tripling during the period (US Energy Information Administration, 2020a), a change that would not be captured in the EPA 2012 inventory. However, Pennsylvania state inventories, which provide annual inventory estimates of unconventional natural gas activity in the state using methodologies similar to the EPA, show CH_4 emissions from unconventional activity only increased by 20% during this period, and that these unconventional wells only represent a portion of O&G emissions in the region (Omara et al., 2016), with much of the emissions coming from pre-existing conventional activity. Thus, it is unlikely that changes in unconventional activity between 2012 and the time of this study would be responsible for the 77% increase in regional emission rates found from the inversion results relative to the EPA 2012 Gridded Inventory, and that the discrepancy would still be present in an updated EPA inventory.

The CH_4 inventory estimates for individual basins from this study generally align with mass balance studies of corresponding basins. In the Haynvesville basin we calculate an O&G emission rate of 43 Mg/hr, compared to 42 Mg/hr from Peischl et al. (2018) and 76 Mg/hr from Cui et al. (2017), which includes non-O&G values in its total as well. In the Barnett, we calculate emissions to be 57 Mg/hr, larger than 46 Mg/hr found in Peischl et al. (2018) but close to the 60 Mg/hr found in Karion et al. (2015). In Eagle Ford, both the western and eastern basin in this study had a combined emission rate of 68 Mg/hr versus 83 Mg/hr in Peischl et al. (2018). Additionally, the large values observed in the Appalachian match findings that show emissions from unconventional O&G infrastructure in the Marcellus are greatly underestimated by EPA inventory values (Barkley et al., 2019a; Caulton et al., 2019).

The interpretation of O&G CH_4 emissions using C_2H_6 observations has a unique advantage compared to more traditional methodologies that rely on CH_4 measurements due to the simplicity of C_2H_6 sources. In the US where CH_4 emissions have near equal contributions from fossil fuels, agriculture, and wetlands, each of which have their own uncertainties, C_2H_6 emissions are dominated almost entirely by the O&G sector. Furthermore, there is high confidence in the spatial mapping of O&G sources in the US due to extensive documentation of the various components associated with O&G extraction, simplifying interpretation of atmospheric C_2H_6 data. As an example, in the ACT-America campaign, the majority of aircraft flight tracks were hundreds to thousands of kilometers away from O&G basins. Despite this, the model prior was able to consistently track C_2H_6 enhancements from these sources in the winter, fall, and spring, with correlations between the model vs observed boundary layer C_2H_6 enhancements of 0.85, 0.75, and 0.61 respectively. The high skill in tracking enhancements from a single sector with well-defined locations creates a scenario where a stable posterior solution can be generated through various inverse methodologies (Table S4).

Despite high confidence in the C_2H_6 posterior, the conversion of this posterior to O&G CH_4 emissions is entirely dependent on the quality and availability of information related to the C_2H_6/CH_4 ratio for each basin. During the time observations from this study were collected, there were numerous recent flights available from a separate study that captured the C_2H_6/CH_4 ratios of various basins (Peischl et al., 2018), providing confidence in our C_2H_6 to CH_4 conversions for those locations. However, basins that are geographically broad, such as the Anadarko and Marcellus, are more difficult to characterize a ratio for using atmospheric data. Furthermore, the average gas composition of a basin can change over time, making ratios found in older studies less applicable (Lan et al., 2019). Publicizing upstream gas composition data collected by major O&G companies would be one immediate solution and effectively eliminate any uncertainty introduced in an C_2H_6 to CH_4 inventory conversion.

One other source of uncertainty related to the C_2H_6 to CH_4 inventory conversion presented in this study is the presence of coal mines in the Appalachia that overlap with

regional gas production. These mines are a dominant source of CH_4 emissions in the region, but little information is available regarding C_2H_6 emissions associated with them. For this study, we treat C_2H_6 emissions from these sources to be negligible based measured values from a 1973 geological survey that found little to no C_2H_6 in many of the mines relevant to this study (Kim, 1973). However, ratios as high as 0.08 have been observed in western Kentucky (Strapoć et al., 2007), raising the possibility that some portion of the C_2H_6 observed in this study in the Appalachia could be attributable to underground mines. Air from major coal ventilation shafts is sampled 4 times each year to quantify the amount of CH_4 present for safety precautions (US Environmental Protection Agency, 2017). Measuring C_2H_6 mixing ratios in these samples would provide additional information to expand our understanding of C_2H_6 emissions in the northeastern US.

4 Conclusion

Using the largest collection of airborne boundary layer C_2H_6 data to date, an inversion was performed to estimate C_2H_6 and CH_4 emissions from various O&G basins across the eastern US. From this dataset, we estimate that a large portion of C_2H_6 emissions cannot be explained using O&G emission data from the EPA 2012 Gridded CH_4 Inventory and existing C_2H_6 to CH_4 emissions ratio data. We conclude that the EPA CH_4 emissions inventory significantly underestimates CH_4 emissions from O&G sources. From this dataset, our results yield CH_4 emissions estimates similar in magnitude to national estimates in Alvarez et al. (2018). In particular, this study finds the largest discrepancies occurring in the Anadarko and western Appalachian O&G plays, though all basins other than Haynesville showed an increase in expected C_2H_6 emissions relative to the prior. We also find C_2H_6 emissions from O&G that are much higher than existing C_2H_6 inventories, a fact that should be taken into account in any future efforts to investigate O&G CH_4 emissions using atmospheric C_2H_6 observations, as well as studies relying on existing C_2H_6 inventories to account for certain chemical reactions.

This study reveals the potential to use broad-scale continuous C_2H_6 data to constrain CH_4 emissions from the O&G sector on a nationwide scale. Our confidence in the spatial distribution of emissions associated with the O&G sector and the dominant role of O&G in US C_2H_6 emissions allows for modelling and interpretation of observed signals without large concerns for source attribution. Furthermore, misalignment in observed versus modelled plumes may serve as a useful diagnostic of model transport errors due to the well-documented spatial knowledge of C_2H_6 emitters in the US. One current weakness with relying on C_2H_6 observations to understand CH_4 emissions from O&G is the requirement of knowing the C_2H_6 to CH_4 ratio of gas composition and basin emissions in order to convert C_2H_6 emissions into CH_4 emissions. Gas composition data is collected by individual companies but not shared publicly. Should some form of this information be made available for the public, it would greatly enhance the confidence of CH_4 emission estimates of the O&G industry using C_2H_6 and other trace gases.

Acknowledgments

The Atmospheric Carbon and Transport (ACT) - America project is a NASA Earth Venture Suborbital 2 project funded by NASA's Earth Science Division (Grant NNX15AG76G to Penn State). ACT-America aircraft data used in this study can be found at https://daac.ornl.gov/cgi-bin/dataset_lister.pl?p=37. We thank C. Sweeney and B. Baier for flask sampling collection during the campaign, and P. Lang, E. Moglia, B. Miller, and M. Crotwell for analysis of the flask samples. We also thank everyone at NASA Langley involved with the data collection and maintenance of the dataset. The authors declare no conflicts of interest with the submission of this manuscript.

References

Alvarez, R. A., Zavala-Araiza, D., Lyon, D. R., Allen, D. T., Barkley, Z. R., Brandt, A. R., ... Hamburg, S. P. (2018). Assessment of methane emissions from the U.S. oil and gas supply chain. *Science*. Retrieved from `http://science.sciencemag.org/content/early/2018/06/20/science.aar7204` doi: 10.1126/science.aar7204

Baier, B. C., Sweeney, C., Choi, Y., Davis, K. J., DiGangi, J. P., Feng, S., ... Weibring, P. (2020). Multispecies assessment of factors influencing regional CO_2 and CH_4 enhancements during the Winter 2017 ACT-America campaign. *Journal of Geophysical Research: Atmospheres*, *125*(2), e2019JD031339. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD031339` (e2019JD031339 10.1029/2019JD031339) doi: 10.1029/2019JD031339

Barkley, Z. R., Davis, K. J., Feng, S., Balashov, N., Fried, A., DiGangi, J., ... Halliday, H. S. (2019b). Forward modeling and optimization of methane emissions in the south central United States using aircraft transects across frontal boundaries. *Geophysical Research Letters*, *46*(22), 13564-13573. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL084495` doi: 10.1029/2019GL084495

Barkley, Z. R., Lauvaux, T., Davis, K. J., Deng, A., Fried, A., Weibring, P., ... Dickerson, R. R. (2019a). Estimating methane emissions from underground coal and natural gas production in southwestern Pennsylvania. *Geophysical Research Letters*, *46*(8), 4531-4540. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL082131` doi: 10.1029/2019GL082131

Barkley, Z. R., Lauvaux, T., Davis, K. J., Deng, A., Miles, N. L., Richardson, S. J., ... Maasakkers, J. D. (2017). Quantifying methane emissions from natural gas production in north-eastern Pennsylvania. *Atmospheric Chemistry and Physics*, *17*(22), 13941–13966. Retrieved from `https://www.atmos-chem-phys.net/17/13941/2017/` doi: 10.5194/acp-17-13941-2017

Bloom, A. A., Bowman, K. W., Lee, M., Turner, A. J., Schroeder, R., Worden, J. R., ... Jacob, D. J. (2017). A global wetland methane emissions and uncertainty dataset for atmospheric chemical transport models (WetCHARTs version 1.0). *Geoscientific Model Development*, *10*(6), 2141–2156. Retrieved from `https://www.geosci-model-dev.net/10/2141/2017/` doi: 10.5194/gmd-10-2141-2017

Brioude, J., Arnold, D., Stohl, A., Cassiani, M., Morton, D., Seibert, P., ... Wotawa, G. (2013, 11). The Lagrangian particle dispersion model FLEXPART-WRF version 3.1. *Geoscientific Model Development*, *6*, 1889-1904. doi: 10.5194/gmd-6-1889-2013

Burkholder, J. B., Sander, S. P., Abbatt, J., Barker, J. R., Huie, R. E., Kolb, C. E., ... Wine, P. H. (2015). *Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies: Evaluation No. 18.* JPL Publication 15-10. Jet Propulsion Laboratory.

Caulton, D., Lu, J., Lane, H., Buchholz, B., Fitts, J., Golston, L., ... Zondlo, M. (2019, 03). Importance of super-emitter natural gas well pads in the Marcellus shale. *Environmental Science & Technology*, *53*. doi: 10.1021/acs.est.8b06965

Cui, Y. Y., Brioude, J., Angevine, W. M., Peischl, J., McKeen, S. A., Kim, S.-W., ... Trainer, M. (2017). Top-down estimate of methane emissions in California using a mesoscale inverse modeling technique: The San Joaquin Valley. *Journal of Geophysical Research: Atmospheres*, *122*(6), 3686-3699. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD026398` doi: 10.1002/2016JD026398

Cui, Y. Y., Brioude, J., McKeen, S. A., Angevine, W. M., Kim, S.-W., Frost, G. J., ... Trainer, M. (2015). Top-down estimate of methane emissions in Cali-

fornia using a mesoscale inverse modeling technique: The South Coast Air Basin. *Journal of Geophysical Research: Atmospheres*, *120*(13), 6698-6711. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD023002 doi: 10.1002/2014JD023002

DiGangi, J. P., Choi, Y., Nowak, J. B., Halliday, H., & Yang, M. M. (2018). *ACT-America: L2 in situ atmospheric CO_2, CO, CH_4, and O_3 concentrations, Eastern USA.* (ORNL DAAC) doi: 10.3334/ORNLDAAC/1556

Feng, S., Lauvaux, T., Davis, K. J., Keller, K., Zhou, Y., Williams, C., ... Baker, I. (2019). Seasonal characteristics of model uncertainties from biogenic fluxes, transport, and large-scale boundary inflow in atmospheric CO_2 simulations over North America. *Journal of Geophysical Research: Atmospheres*, *124*(24), 14325-14346. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JD031165 doi: 10.1029/2019JD031165

Grant, R. F., & Roulet, N. T. (2002). Methane efflux from boreal wetlands: Theory and testing of the ecosystem model ecosys with chamber and tower flux measurements. *Global Biogeochemical Cycles*, *16*(4), 2-1-2-16. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001GB001702 doi: 10.1029/2001GB001702

Huang, G., Brook, R., Crippa, M., Janssens-Maenhout, G., Schieberle, C., Dore, C., ... Friedrich, R. (2017). Speciation of anthropogenic emissions of non-methane volatile organic compounds: a global gridded data set for 1970–2012. *Atmospheric Chemistry and Physics*, *17*(12), 7683–7701. Retrieved from https://www.atmos-chem-phys.net/17/7683/2017/ doi: 10.5194/acp-17-7683-2017

Karion, A., Sweeney, C., Kort, E. A., Shepson, P. B., Brewer, A., Cambaliza, M., ... Tans, P. (2015). Aircraft-based estimate of total methane emissions from the Barnett shale region. *Environmental Science & Technology*, *49*(13), 8124-8131. Retrieved from https://doi.org/10.1021/acs.est.5b00217 (PMID: 26148550) doi: 10.1021/acs.est.5b00217

Kim, A. G. (1973). *The composition of coalbed gas.* (Pittsburgh Mining and Safety Research Center. Available at https://www.cdc.gov/niosh/mining/works/coversheet1285.html. Accessed October 2018)

Kitanidis, P. (1997). *Introduction to geostatistics: Applications in hydrogeology.* Cambridge University Press. Retrieved from https://books.google.com/books?id=ZvoibTTS9QwC

Kostinek, J., Roiger, A., Davis, K. J., Sweeney, C., DiGangi, J. P., Choi, Y., ... Butz, A. (2019). Adaptation and performance assessment of a quantum and interband cascade laser spectrometer for simultaneous airborne in situ observation of CH_4, C_2H_6, CO_2, CO and N_2O. *Atmospheric Measurement Techniques*, *12*(3), 1767–1783. Retrieved from https://www.atmos-meas-tech.net/12/1767/2019/ doi: 10.5194/amt-12-1767-2019

Lan, X., Tans, P., Sweeney, C., Andrews, A., Dlugokencky, E., Schwietzke, S., ... Biraud, S. C. (2019). Long-term measurements show little evidence for large increases in total U.S. methane emissions over the past decade. *Geophysical Research Letters*, *46*(9), 4991-4999. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL081731 doi: 10.1029/2018GL081731

Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., ... Davis, K. J. (2012). Constraining the CO_2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system. *Atmospheric Chemistry and Physics*, *12*(1), 337-354. Retrieved from https://www.atmos-chem-phys.net/12/337/2012/ doi: 10.5194/acp-12-337-2012

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J.-X., ... Parker, R. J. (2019). Global distribution of methane emissions,

emission trends, and OH concentrations and trends inferred from an inversion of GOSAT satellite data for 2010–2015. *Atmospheric Chemistry and Physics*, *19*(11), 7859–7881. Retrieved from https://acp.copernicus.org/articles/19/7859/2019/ doi: 10.5194/acp-19-7859-2019

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., ... Fischer, M. L. (2016). Gridded national inventory of U.S. methane emissions. *Environmental Science & Technology*, *50*(23), 13123-13133. (PMID: 27934278) doi: 10.1021/acs.est.6b02878

Matthes, J. H., Sturtevant, C., Verfaillie, J., Knox, S., & Baldocchi, D. (2014). Parsing the variability in CH_4 flux at a spatially heterogeneous wetland: Integrating multiple eddy covariance towers with high-resolution flux footprint analysis. *Journal of Geophysical Research: Biogeosciences*, *119*(7), 1322-1339. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JG002642 doi: 10.1002/2014JG002642

McKain, K., Down, A., Raciti, S. M., Budney, J., Hutyra, L. R., Floerchinger, C., ... Wofsy, S. C. (2015). Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. *Proceedings of the National Academy of Sciences*, *112*(7), 1941–1946. Retrieved from http://www.pnas.org/content/112/7/1941 doi: 10.1073/pnas.1416261112

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., ... Zhang, H. (2013). Anthropogenic and natural radiative forcing [Book Section]. In T. Stocker et al. (Eds.), *Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change* (p. 659–740). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Retrieved from www.climatechange2013.org doi: 10.1017/CBO9781107415324.018

Nisbet, E. G., Manning, M. R., Dlugokencky, E. J., Fisher, R. E., Lowry, D., Michel, S. E., ... White, J. W. C. (2019). Very strong atmospheric methane growth in the 4 years 2014–2017: Implications for the Paris agreement. *Global Biogeochemical Cycles*, *33*(3), 318-342. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GB006009 doi: 10.1029/2018GB006009

Omara, M., Sullivan, M. R., Li, X., Subramanian, R., Robinson, A. L., & Presto, A. A. (2016). Methane emissions from conventional and unconventional natural gas production sites in the Marcellus shale basin. *Environmental Science & Technology*, *50*(4), 2099-2107. Retrieved from https://doi.org/10.1021/acs.est.5b05503 (PMID: 26824407) doi: 10.1021/acs.est.5b05503

Peischl, J., Eilerman, S. J., Neuman, J. A., Aikin, K. C., de Gouw, J., Gilman, J. B., ... Ryerson, T. B. (2018). Quantifying methane and ethane emissions to the atmosphere from central and western U.S. oil and natural gas production regions. *Journal of Geophysical Research: Atmospheres*, *123*(14), 7725-7740. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD028622 doi: 10.1029/2018JD028622

Peischl, J., Ryerson, T. B., Aikin, K. C., de Gouw, J. A., Gilman, J. B., Holloway, J. S., ... Parrish, D. D. (2015). Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions. *Journal of Geophysical Research: Atmospheres*, *120*(5), 2119-2139. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD022697 doi: 10.1002/2014JD022697

Plant, G., Kort, E. A., Floerchinger, C., Gvakharia, A., Vimont, I., & Sweeney, C. (2019). Large fugitive methane emissions from urban centers along the U.S. east coast. *Geophysical Research Letters*, *46*(14), 8500-8507. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL082635 doi: 10.1029/2019GL082635

Rella, C. W., Tsai, T. R., Botkin, C. G., Crosson, E. R., & Steele, D. (2015). Mea-

suring emissions from oil and natural gas well pads using the mobile flux plane technique. *Environmental Science & Technology*, *49*(7), 4742-4748. (PMID: 25806837) doi: 10.1021/acs.est.5b00099

Robertson, A. M., Edie, R., Snare, D., Soltis, J., Field, R. A., Burkhart, M. D., ... Murphy, S. M. (2017). Variation in methane emission rates from well pads in four oil and gas basins with contrasting production volumes and compositions. *Environmental Science & Technology*, *51*(15), 8832-8840. Retrieved from https://doi.org/10.1021/acs.est.7b00571 (PMID: 28628305) doi: 10.1021/acs.est.7b00571

Roscioli, J. R., Yacovitch, T. I., Floerchinger, C., Mitchell, A. L., Tkacik, D. S., Subramanian, R., ... Marchese, A. J. (2015). Measurements of methane emissions from natural gas gathering facilities and processing plants: measurement methods. *Atmospheric Measurement Techniques*, *8*(5), 2017–2035. Retrieved from https://amt.copernicus.org/articles/8/2017/2015/ doi: 10.5194/amt-8-2017-2015

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., ... Zhuang, Q. (2020). The global methane budget 2000–2017. *Earth System Science Data*, *12*(3), 1561–1623. Retrieved from https://essd.copernicus.org/articles/12/1561/2020/ doi: 10.5194/essd-12-1561-2020

Sheng, J.-X., Jacob, D. J., Maasakkers, J. D., Sulprizio, M. P., Zavala-Araiza, D., & Hamburg, S. P. (2017). A high-resolution (0.1°x 0.1°) inventory of methane emissions from Canadian and Mexican oil and gas systems. *Atmospheric Environment*, *158*, 211 - 215. Retrieved from http://www.sciencedirect.com/science/article/pii/S1352231017301164 doi: https://doi.org/10.1016/j.atmosenv.2017.02.036

Sheng, J.-X., Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sulprizio, M. P., Bloom, A. A., ... Wunch, D. (2018). High-resolution inversion of methane emissions in the southeast us using SEAC^{4}RS aircraft observations of atmospheric methane: anthropogenic and wetland sources. *Atmospheric Chemistry and Physics*, *18*(9), 6483–6491. Retrieved from https://www.atmos-chem-phys.net/18/6483/2018/ doi: 10.5194/acp-18-6483-2018

Strapoć, D., Mastalerz, M., Eble, C., & Schimmelmann, A. (2007). Characterization of the origin of coalbed gases in southeastern illinois basin by compound-specific carbon and hydrogen stable isotope ratios. *Organic Geochemistry*, *38*(2), 267 - 287. Retrieved from http://www.sciencedirect.com/science/article/pii/S0146638006002385 doi: https://doi.org/10.1016/j.orggeochem.2006.09.005

Tzompa-Sosa, Z. A., Mahieu, E., Franco, B., Keller, C. A., Turner, A. J., Helmig, D., ... Fischer, E. V. (2017). Revisiting global fossil fuel and biofuel emissions of ethane. *Journal of Geophysical Research: Atmospheres*, *122*(4), 2493-2512. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD025767 doi: 10.1002/2016JD025767

US Energy Information Administration. (2020a). *Natural gas gross withdrawls and production.* (Available at https://www.eia.gov/dnav/ng/ng_prod_sum_a_EPG0_VGM_mmcf_a.htm. Accessed Jul 6 2020)

US Energy Information Administration. (2020b). *United States energy information administration homepage.* (Available at https://www.eia.gov/. Accessed August 2020)

US Environmental Protection Agency. (2014). *2011 national emissions inventory data.* (Available at https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data. Accessed October 2020)

US Environmental Protection Agency. (2017). *U.S. underground coal mine ventilation air methane exhaust characterization 2011-2015.*

(https://www.epa.gov/sites/production/files/2016-03/documents/vam-exhaust-characterization-july2010.pdf. Accessed October 2018)

US Environmental Protection Agency. (2020). *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2018.* (Available at https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks. Accessed August 2020)

US Geological Survey. (2018). *Geochemistry laboratory database.* (Available at https://certmapper.cr.usgs.gov/data/apps/geochem-db/ Accessed July 2018)

Weibring, P., Richter, D., Walega, J. G., Fried, A., DiGangi, J., Halliday, H., ... Obland, M. D. (2020). Autonomous airborne mid-ir spectrometer for high precision measurements of ethane during the NASA ACT-America studies. *Atmospheric Measurement Techniques Discussions*, *2020*, 1–42. Retrieved from `https://amt.copernicus.org/preprints/amt-2020-210/` doi: 10.5194/amt-2020-210

Xiao, Y., Logan, J. A., Jacob, D. J., Hudman, R. C., Yantosca, R., & Blake, D. R. (2008). Global budget of ethane and regional constraints on U.S. sources. *Journal of Geophysical Research: Atmospheres*, *113*(D21). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JD009415` doi: 10.1029/2007JD009415

Yacovitch, T., Herndon, S., Roscioli, J., Floerchinger, C., Mcgovern, R., Agnese, M., ... Kolb, C. (2014, 06). Demonstration of an ethane spectrometer for methane source identification. *Environmental science & technology*, *48*. doi: 10.1021/es501475q

Yu, X., Millet, D. B., Wells, K. C., Henze, D. K., Cao, H., Griffis, T. J., ... Bloom, A. A. (2020). Aircraft-based inversions quantify the importance of wetlands and livestock for upper midwest methane emissions. *Atmospheric Chemistry and Physics Discussions*, *2020*, 1–34. Retrieved from `https://acp.copernicus.org/preprints/acp-2020-826/` doi: 10.5194/acp-2020-826

Yvon-Durocher, G., Allen, A. P., Bastviken, D., Conrad, R., Gudasz, C., St-Pierre, A., ... del Giorgio, P. A. (2014, 03). Methane fluxes show consistent temperature dependence across microbial to ecosystem scales. *Nature*, *507*, 488-491. doi: 10.1038/nature13164

Zhang, Y., Gautam, R., Pandey, S., Omara, M., Maasakkers, J. D., Sadavarte, P., ... Jacob, D. J. (2020). Quantifying methane emissions from the largest oil-producing basin in the United States from space. *Science Advances*, *6*(17). Retrieved from `https://advances.sciencemag.org/content/6/17/eaaz5120` doi: 10.1126/sciadv.aaz5120