Supporting Information for "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], Alina 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

Contents

S1. USGS C_2H_6/CH_4 Ratio Map

To fill in C_2H_6/CH_4 ratios for areas not measured in recent flight campaigns, ratio data was used based on chemical composition samples taken from the United States Geological Survey (USGS) Geochemical Laboratory Database (US Geological Survey, 2018). This database contains 13,000 representative samples of natural gas composition from wells across the US. USGS provides geographic coordinates for each data point but not other identifying information to ensure that the data origin remains anonymous. We grid these observations to a 0.25 degree latitude by 0.25 degree longitude resolution using a statistical interpolation approach known as ordinary kriging (Kitanidis, 1997) (Figure S1).

S2. C_2H_6 Chemical Sink

One can examine the seasonal dependence of the C_2H_6 lifetime (in days) due to reactions with OH, its primary sink in the ABL. The C_2H_6 lifetime is calculated from Eq. S1 below:

$$t = 1/(k_{OH} * [OH]) \tag{1}$$

where t is the lifetime of C_2H_6 in seconds. The temperature-dependent reaction rate constant (k_{OH}) is obtained from the JPL Kinetics database (Burkholder et al., 2015) is 7.66 x 10^{-12} $e^{-(1020/T)}$ cm^3 $molecules^{-1}$ sec^{-1} (where T is the temperature in K). Assuming a wintertime temperature of 20°F (266 K) and a typical OH concentration of 10^6 molecules cm^{-3} during unpolluted conditions, one calculates an C_2H_6 lifetime of 70 days. During the summer, assuming a temperature of 104°F (313 K) and a very high OH concentration of 10^7 molecules cm^{-3}, this lifetime reduces to 4 days (Table S1).

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

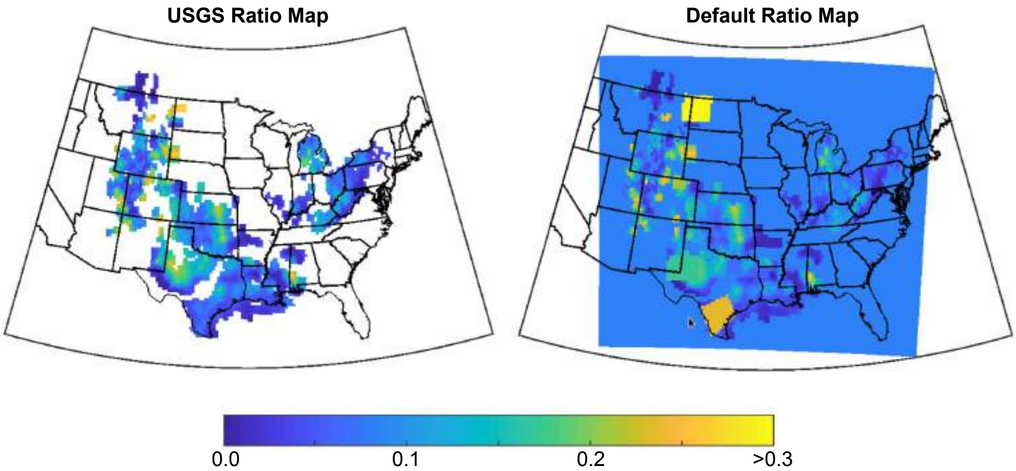

Figure 1. (left) C_2H_6/CH_4 ratios derived from the USGS Geological Database. Grid cells that are outside the domain or did not contain at least one data point from the USGS database are left blank. (right) C_2H_6/CH_4 ratios used for the production sector in the Default prior from this study, integrating data from the USGS database with basinwide ratios observed from aircraft measurements. Areas with no data were assigned a value of 0.085.

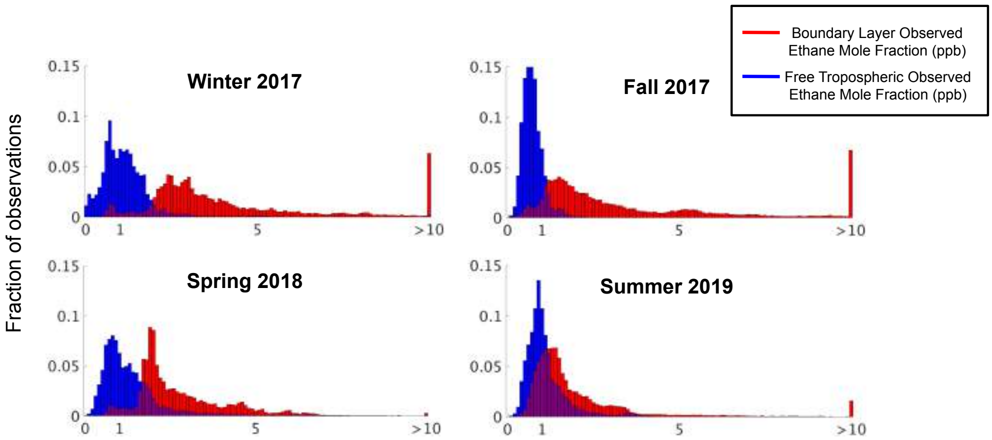

Figure 2. A histogram of all C_2H_6 observations observed during each seasonal flight campaign, separated into boundary layer (<1000 m AGL) and free troposphere (>2000 m AGL)

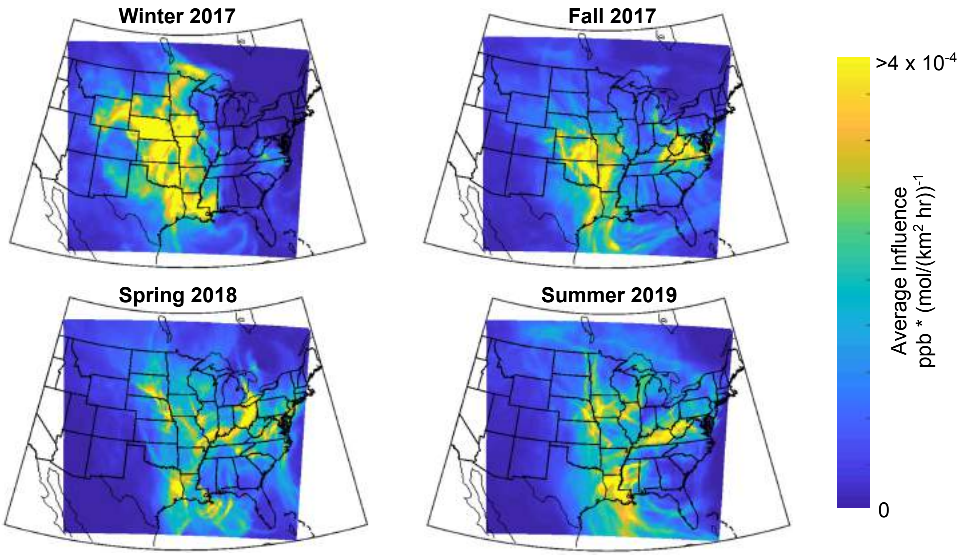

Figure 3. Averaged influence functions from the boundary layer observations used in the C_2H_6 inversion for each season. Brighter colors indicate areas whose surface interactions were captured more frequently by the boundary layer observations.

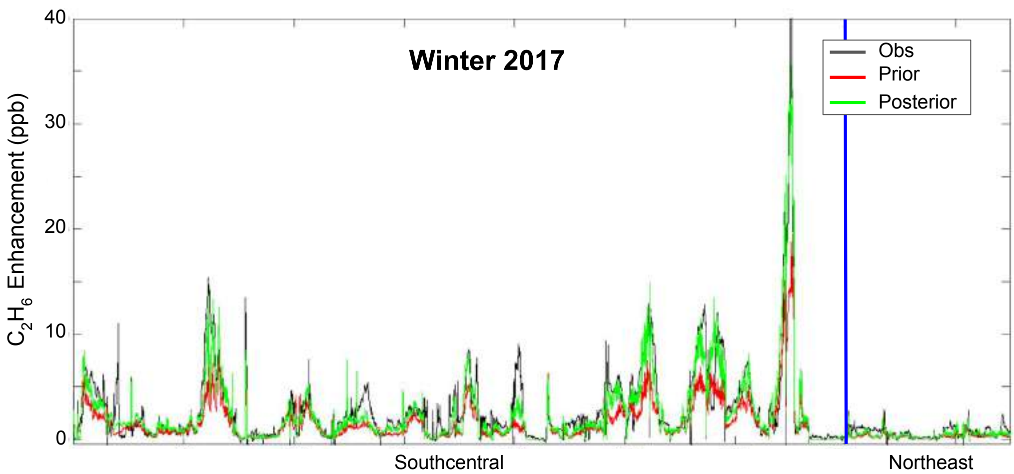

Figure 4. Observed, modelled Default prior, and modelled Default posterior C_2H_6 enhancements from all flights in the Winter 2017 campaign. The blue line denotes a shift in the region the observations were collected.

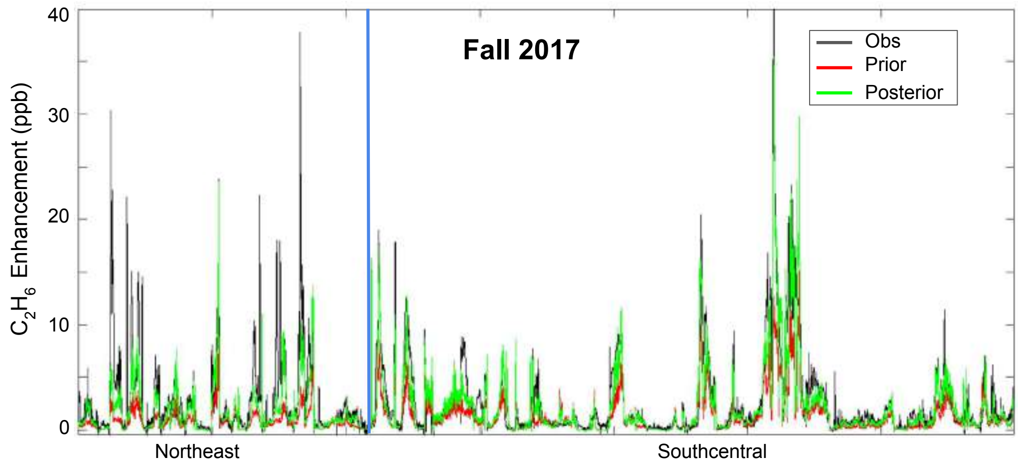

Figure 5. Observed, modelled Default prior, and modelled Default posterior C_2H_6 enhancements from all flights in the Fall 2017 campaign. The blue line denotes a shift in the region the observations were collected.

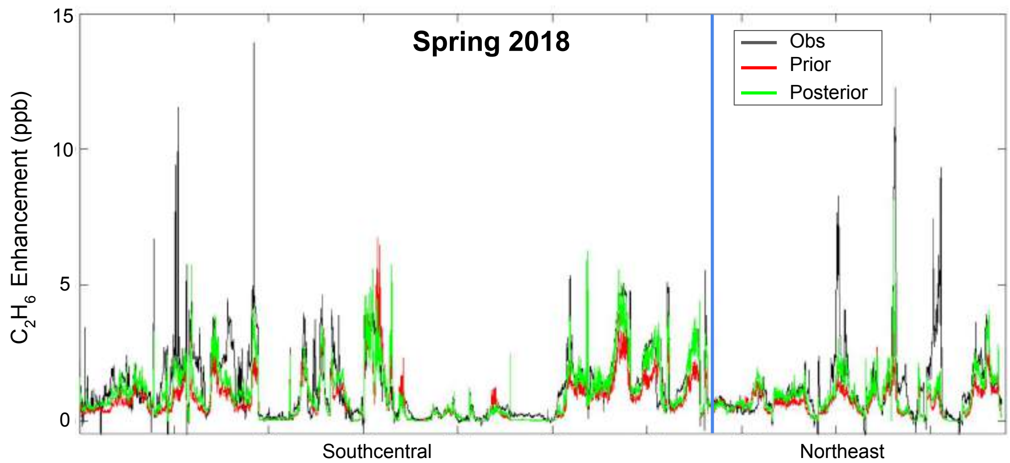

Figure 6. Observed, modelled Default prior, and modelled Default posterior C_2H_6 enhancements from all flights in the Spring 2018 campaign. The blue line denotes a shift in the region the observations were collected.

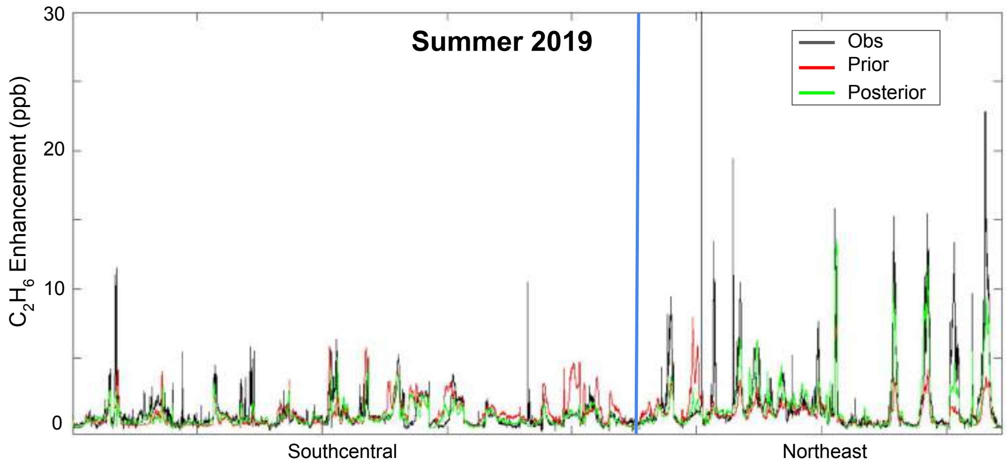

Figure 7. Observed, modelled Default prior, and modelled Default posterior C_2H_6 enhancements from all flights in the Summer 2019 campaign. The blue line denotes a shift in the region the observations were collected.

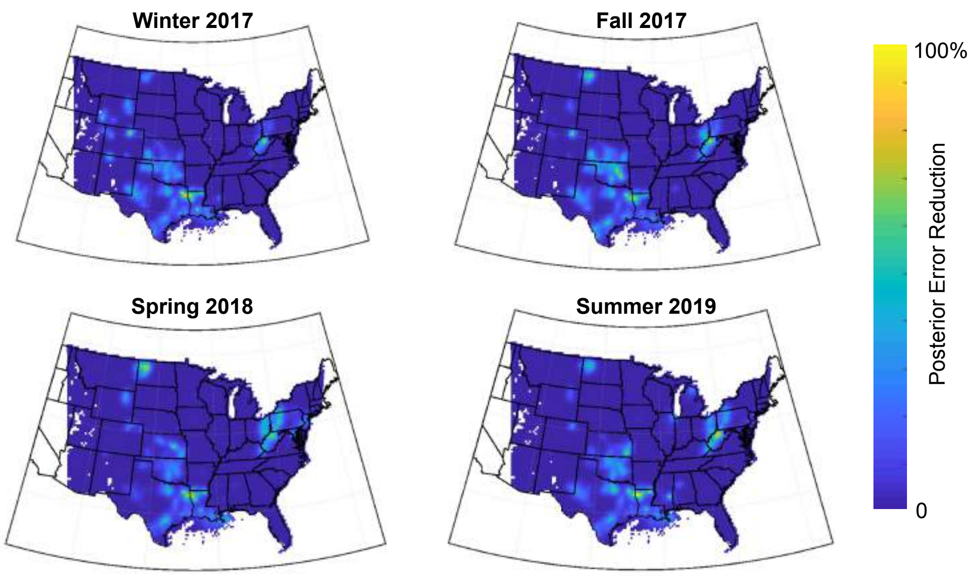

Figure 8. Fractional change in the posterior/prior B matrix showing areas with the largest improvement in the flux errors. Results shown using the Default prior and posterior.

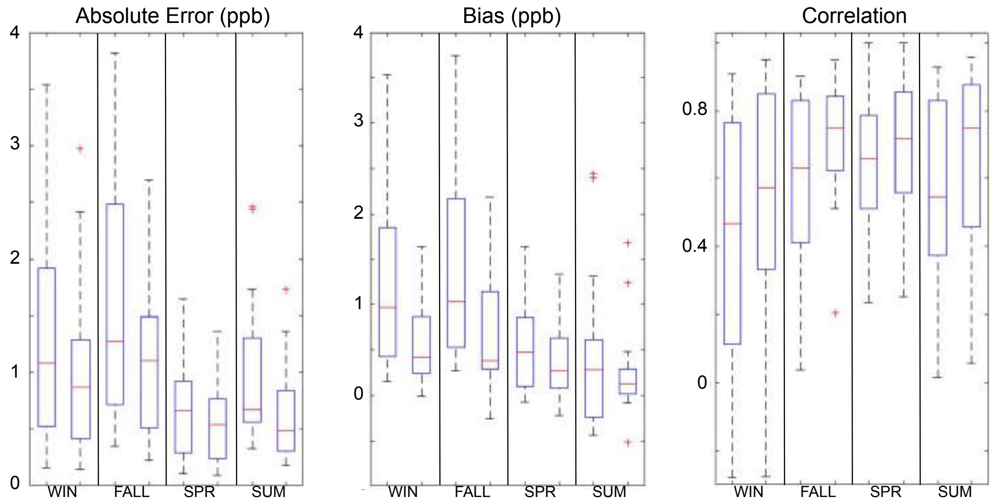

Figure 9. Box and whisker plots showing flight-by-flight statistical performances of the Default prior (left box in each column) and individual season posteriors (right box in each column).

Temperature (℃)	[OH] (molecule/cm^3)	Lifetime (days)	C_2H_6 loss after 24 hours
15	1 x 10^6	54.5	1.8%
25	1 x 10^6	48.2	2.1%
40	1 x 10^6	40.5	2.4%
40	5 x 10^6	8.1	11.6%
40	1 x 10^7	4.1	21.9%

Table 1. Table showing loss rate of C_2H_6 at different temperatures and OH concentrations

Basin	C_2H_6 / CH_4 Ratio	Reference
Anadarko	Varies	USGS Database
Appalachia	Varies	USGS Database
Bakken	0.514	Peischl et. al., 2018
Barnett	0.068	Peischl et. al., 2018
Eagle Ford	0.238	Peischl et. al., 2018
Fayetteville	0.006	Peischl et. al., 2015
Haynesville	0.057	Peischl et. al., 2018
Permian	0.170	Tzompa-Sosa et. al., 2017

Table 2. Table showing C_2H_6/CH_4 ratios applied to production sites of different O&G basins to convert between C_2H_6 and CH_4 emissions.

	Observed Mean Wind Speed (m/s)	Model Mean Wind Speed Bias (m/s)	Observed Mean Boundary Layer Height (m)	Model Mean Boundary Layer Bias (m)	H Correction Factor	Model Wind Direction Mean Absolute Error (degrees)
Winter 2017	8.9	1.3	1204	-260	0.90	14.6°
Fall 2017	7.8	0.2	1044	-153	0.88	16.1°
Spring 2018	8.4	-0.2	1544	-323	0.77	15.8°
Summer 2019	5.8	-0.4	1314	-35	0.91	30.5°

Table 3. Table showing performance of the WRF 27 km model run used to drive the meteorology in FLEXPART-WRF. H correction factor is derived from Barkley et al. (2017) based on seasonal biases in the model wind speed and boundary layer and used in the H Correction column of the sensitivity test in Table S4.

	Default Prior	Default Posterior	Default Posterior H Correction	Default Posterior Prior x 3	Default Posterior BG = 10%	Default Posterior No Flux Correlation	Default Posterior Equal Weighting (Constant R)	Default Posterior 3 Season Combined	Flat Rate Posterior	NEI2011 Posterior
Southcentral Total C_2H_6 (mol s^{-1})	488	696	864	706	605	658	730	644	613	614
Western Appalachia Total C_2H_6 (mol s^{-1})	125	221	255	253	217	170	228	203	238	171
Three Season Mean Absolute Error (ppb)	1.28	1.02	0.96	0.81	1.04	1.11	0.92	0.95	1.02	1.09
Three Season Mean Bias (ppb)	1.14	0.70	0.43	0.50	0.70	0.85	0.64	0.60	0.65	0.82
Three Season y,Hx Correlation	0.78	0.82	0.81	0.81	0.81	0.80	0.84	0.83	0.81	0.80

Table 4. Emission totals and performance statistics of different posterior maps created through a sensitivity analysis.
Default Prior: The main prior used in this study.
Default Posterior: The main 3 season posterior used in this study.
Default Posterior H Correction: Similar to the Default posterior method but with an adjusted H matrix to account for ABL and wind speed biases in each season (Table S3).
Default Posterior Prior x 3: Posterior solution created from the default prior multiplied by a factor of 3.
Default Posterior BG = 10%: Similar to the Default posterior method, but with a background value calculated using the 10th percentile of boundary layer C_2H_6 values rather than 5th percentile.
Default Posterior No Flux Correlation: Similar to the Default posterior method, but with the flux correlation length changed from 54 km to 0 (off-diagonal elements removed from $\mathbf{B}$).
Default Posterior Equal Weighting: Similar to the Default posterior method, but with the observation-transport error matrix $\mathbf{R}$ set to a constant value for each season, resulting in all observations receiving equal uncertainty.
Default Posterior 3 Season Combined. A posterior solution created by performing an inversion on all 3 season at once rather than averaging 3 individual posteriors associated with each season.
Flat Rate Posterior: A posterior solution created by using a prior that assigned a flat C_2H_6/CH_4 ratio of 0.085 to all O&G CH_4 emissions in the US.
NEI2011 Posterior: A posterior solution created by using the NEI2011 C_2H_6 inventory as a prior.