Supporting Information for "Quantification of Oil and Gas Methane Emissions in the Delaware and Marcellus Basins Using a Network of Continuous Tower-Based Measurements"

Z. R. Barkley[1], K. J. Davis[1], N. Miles[1], S. Richardson[1], A. Deng[2], B. Hmiel[3], D. Lyon[3], T. Lauvaux[4]

[1]The Pennsylvania State University, University Park, PA, USA
[2]Utopus Insights, Inc, Valhalla, NY, USA
[3]Environmental Defense Fund. 301 Congress Ave., Suite 1300, Austin, TX, USA
[4]GSMA, University of Reims-Champagne Ardenne, UMR CNRS 7331, Reims, France

1 S1. Inversion Sensitivity Analysis

To estimate emissions using the Bayesian inversion described in Section 2.5, numerous decisions are made that can affect the the resulting posterior solution. In this sensitivity analysis we explore adjusting different parameters and the consequences these choices have on the resulting posterior emissions map. These changes include adjustments to the prior flux magnitude, adjustments to the uncertainty of the prior flux, adjustments to the error correlation length scale, changes to the hours which define the afternoon period, and changes to how the background tower is defined. These changes are described in further detail below, and results can be found in Tables S2-S5 as well as in Figure 3 of the main manuscript.

Adjusting the magnitude to the prior can help to determine how much the total of the posterior solution is being pre-determined by its prior. In an ideal scenario, three priors with similar spatial structures but different magnitudes would converge to a similar posterior solution. However, transport uncertainties or insufficient influence function coverage can result in a posterior emission map that still contains characteristics of the prior's flux magnitude. In this test, an inversion is run for each prior, adjusting its magnitude by 50% and 150% to see the level of convergence between the posterior magnitudes. For this experiment, the error covariance matrix is doubled from its original value in both cases to give the inversion the flexibility to shift heavily from the prior, as the prior itself no longer represents a best estimate of the truth. In the Delaware basin, using the EI_{ME} prior at 50% and 150% (88 and 264 Mg/hr) produces mean posterior solutions of 146 (+58) and 210 (-54) Mg/hr respectively, showing some but not total convergence to the value of 174 Mg/hr using the original prior. In the Marcellus using the Production-based prior at 50% and 150% (12 and 37 Mg/hr) produces mean posterior solutions of 19 (+7) and 26 (-11) Mg/hr respectively. Interestingly, using the PADEP inventory as a prior (7 Mg/hr), both the 50% and 150% solutions (16 and 19 Mg/hr) converge to values above the 150% prior, strongly indicating that the PADEP prior is far too low and as a consequence may be biasing the posterior solutions with it.

Adjustments to the correlation length scale of the prior flux and adjustments to the uncertainty of the prior flux are both ways of changing the structure of the error covariance matrix used in the inversion to solve for a posterior flux. Here, we experiment with adjusting the correlation length, running an inversion using a length of 10 km. By increasing the correlation length, it forces the inversion to move large sections of the posterior solution towards the same relative change in the flux. resulting in an inversion more inclined to solve for the mean change across the basin and less capable of attempting to solve for small scale changes. Additionally, increasing the correlation length allows the

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

inversion to adjust grids on the outskirts of the domain where footprint coverage may be less prevalent. In both the Delaware and Marcellus inversions, solutions using the 10 km correlation length were similar in magnitude and performance to their 5 km counterparts, indicating that the selection of reasonable correlation length has little impact on our overall posterior solutions.

In addition to adjusting the correlation length, we also run an inversion scenario where the flux error matrix is doubled. This gives the posterior solution the ability to stray further from the prior and achieve a posterior solution that matches more closely with the observations, but can result in an underconstrained solution and create unrealistically high or negative flux values to achieve the optimal match with observations, underweighting noise related to transport errors. In both the Delaware and Marcellus inversions, doubling the flux error matrix produces solutions with minor improvements to the obs-model statistical comparisons relative to the default posterior solution, but with similar total fluxes.

Throughout the study, afternoon hours are defined as 20-23 UTC in the Delaware and 18-21 UTC in the Marcellus. These hours are selected based on the time within the transport model when boundary layer heights are at their peak and have stabilized (Figure S8), providing more confidence in the overall solution as mismatches in timing of the boundary layer development in the model compared to reality would have minimal impacts on the projected size of the enhancements. However, theoretically the tower observations are measuring enhancements from the study domain mixed within the boundary layer at earlier times as well, even if the boundary layer is still developing. In the Early scenario, we expand the hours included in the afternoon definition to 16-23 UTC in the Permian and 14-21 UTC in the Marcellus. In both the Delaware and Marcellus inversions, using the early hour data produces posterior solutions that statistically perform slightly worse than the default posterior, but have similar total emissions.

In the default inversion, the observed enhancement is calculated by subtracting off a background value based on the tower(s) that have the lowest mole fraction, or have the smallest model enhancement. The main advantage of this method is that the tower(s) selected are most likely to be clean of contamination from local sources that could create errors in estimating the background of the air mass entering the model domain. However, the tower(s) selected may not lie directly upwind of the towers downwind of the O&G sources, and could at times be representative of an air mass different from those downwind towers. To account for this, an alternative background tower selection method is performed, selecting the tower directly upwind of the O&G sources based on the mean afternoon wind direction. In both the Delaware and the Marcellus, this alternative background selection produces a posterior solution that statistically performs substantially worse than the default selection in terms of the mean absolute error and correlation between model and observations and overall has a marginally lower emissions total. Despite the worse statistical performance of the alternative background posteriors, the overall bias between the model and observed enhancements is closer to 0 in all alternative background posterior solutions. This result is contrary to the default inversion posteriors which all maintain a negative bias (model enhancements larger than observations) and may indicate that the default background methodology results in a background that is too low, artificially creating enhancements at downwind tower sites that the inversion is not able to reconcile. Regardless of this discrepancy, both background selections produce posterior emission rates that are within 15% of each other for all prior inventories.

2 S2. Detailed Background Sensitivity Analysis

In this section we examine the effects various background selection methods have on the posterior flux and the skill of the model dataset relative to the observations. A total of 6 different selection methods are tested, described below.

-Minimum Tower: The background tower is selected based on the tower with the lowest observed afternoon methane mole fraction.

-Minimum Tower (no model subtraction): The background tower is selected based on the tower with the lowest observed afternoon methane mole fraction. However, in this case, the modeled methane enhancements are not subtracted off of the background. This method assumes that the tower with the lowest observed methane values should not be heavily influenced by sources within the model domain, and that zeroing any model enhancements at the tower site eliminates the possibility of further reducing the background value due to errant plumes within the model.

-Minimum Model: The background tower is selected based on the tower with the lowest **modelled** afternoon methane enhancement.

-Hybrid: The background value is selected by averaging the results of the minimum tower methods and the minimum model method. This is the method used in the "default" inversion.

-Hybrid Filter: The background value is selected by averaging the results of the minimum tower methods and the minimum model method. If the resulting background value using each of the two methods is not within 10 ppb, the day is not used in the inversion.

-Upwind: The background value is selected based on the tower or towers that are most directly upwind of the O&G basin using the afternoon mean wind direction.

Tables S7-S8 show summaries of the results using each one of these background methods for each basin. In both basins, the Hybrid and Hybrid Filter approaches produce the best overall posterior results, with low mean absolute errors and high correlations between model and observations, and more consistent background values in a 15 day moving std test. Though the Hybrid Filter approach performs better than the Hybrid approach used as the default method in the inversion, this method comes at the cost of eliminating data on days with complex background conditions (30% in the Delaware and 41% in the Marcellus). Of the remaining methods, none perform consistently well across both basins.

Figure S21 shows how each background methodology affects the temporal timeseries of posterior emission rates. Generally, the background methodology selected will change the overall magnitude of the emission rate but not affect the trendline. It is important to note that each of these solutions should not be considered equally plausible. For example, the Minimum Tower method would be expected to produce a background that is biased low (and thus an emission rate that is biased high) and the Minimum Model method would produce an background that is biased high (and thus an eission rate that is biased low) as discussed in the main text.

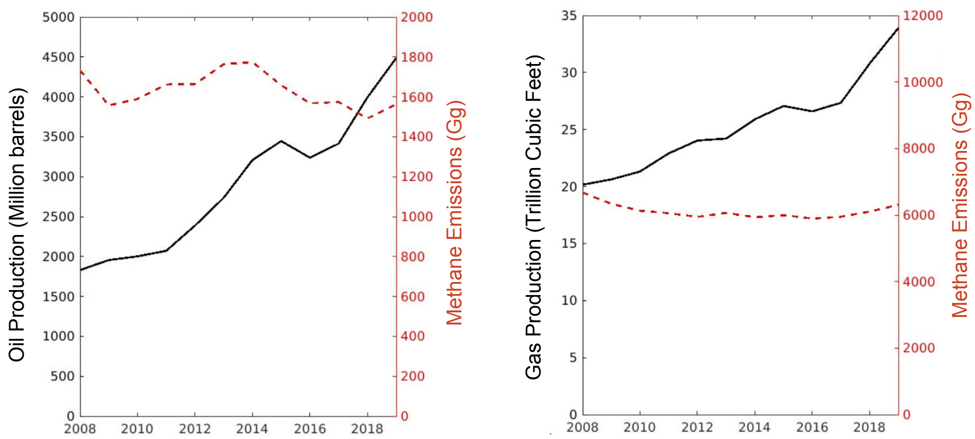

Figure S1. (left) U.S. annual oil production in black, with EPA bottom-up inventory emission estimates of methane emissions from the petroleum sector in red. (right) U.S. annual natural gas production in black, with EPA bottom-up inventory estimates of methane emissions from the natural gas sector in red.

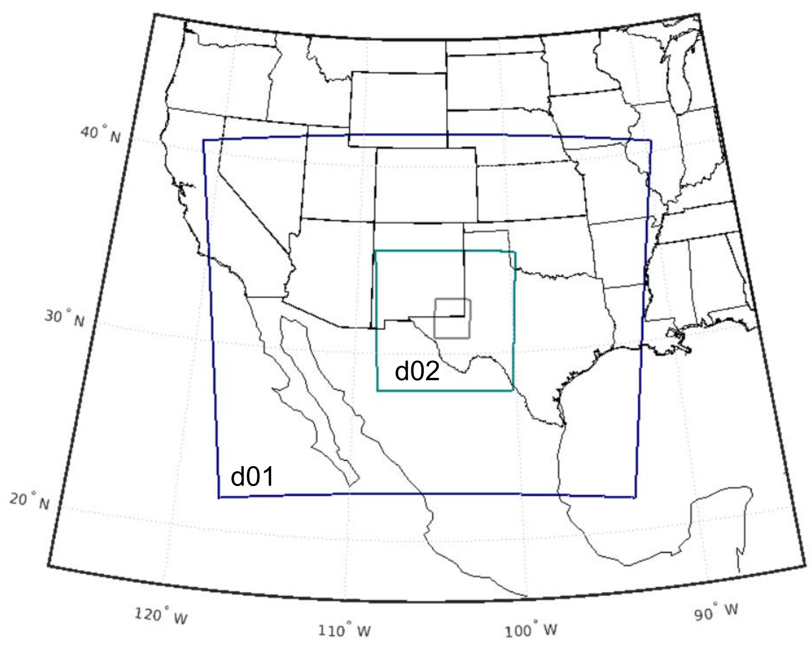

Figure S2. A map of the 9 km and 3 km model domain used to generate meteorology for influence functions in the Delaware tower analysis, with the study domain illustrated within the 3 km model domain.

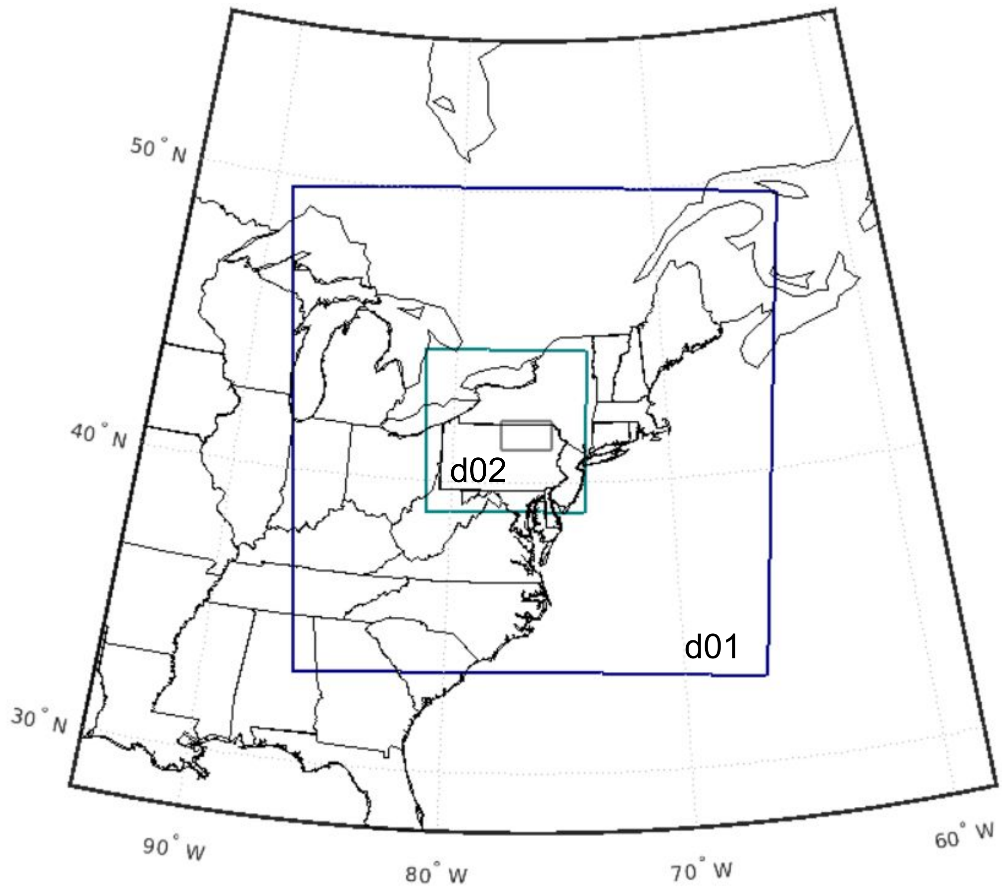

Figure S3. A map of the 9 km and 3 km model domain used to generate meteorology for influence functions in the northeastern Marcellus tower analysis, with the study domain illustrated within the 3 km model domain.

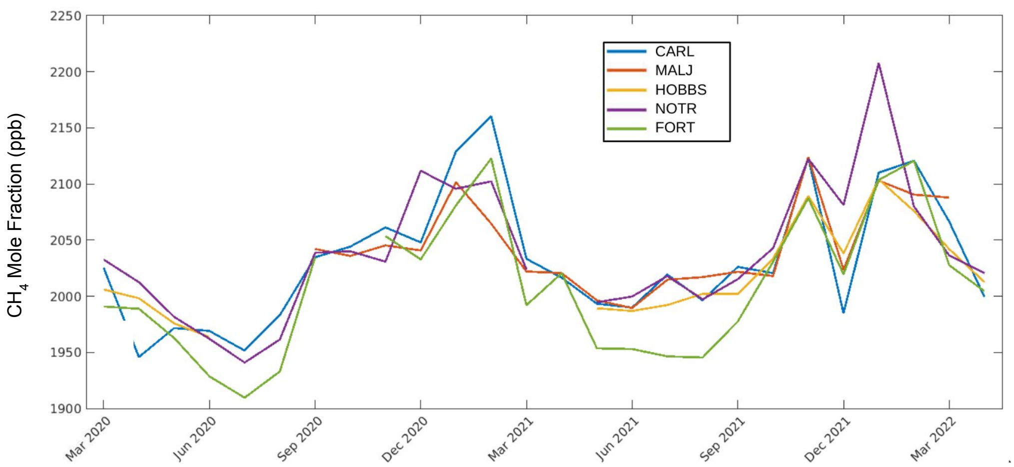

Figure S4. Monthly means of observed afternoon mole fractions (20-23 UTC) from the Delaware basin tower network from March 2020 through April 2022.

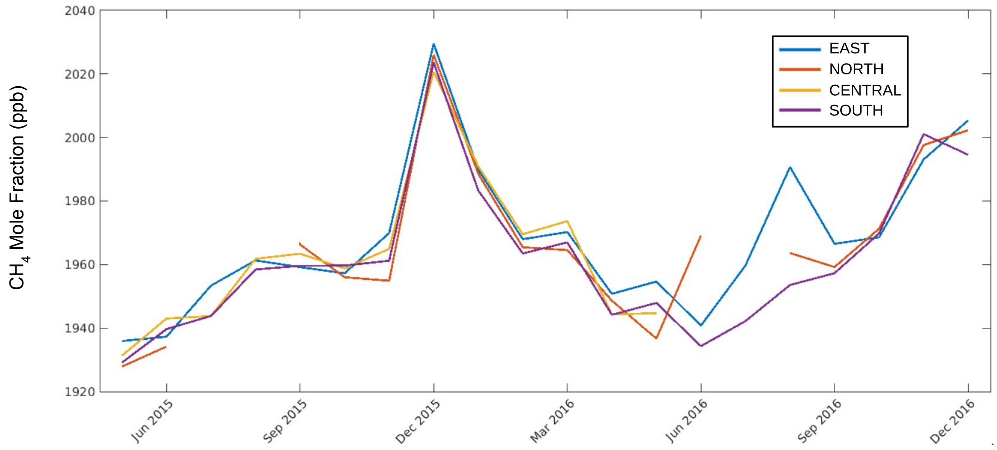

Figure S5. Monthly means of observed afternoon mole fractions (18-21 UTC) from the Marcellus basin tower network from May 2015 through December 2016.

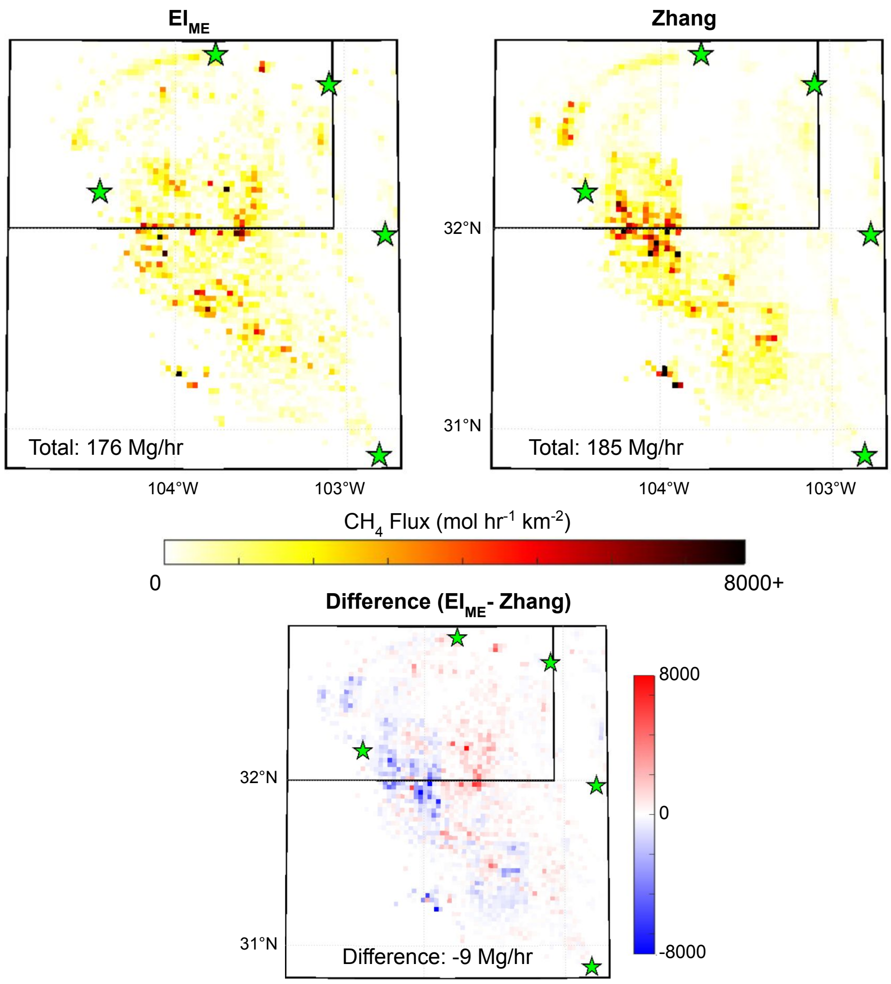

Figure S6. A comparison between O&G emissions from the two priors used in this study for the Delaware basin. (left) The EI_{ME} emission map constructed from site-level data and used in Zhang et al. (2020). (right) The posterior emission map from Zhang et al. (2020), used as an alternative prior for this study. (bottom) The difference between the two priors.

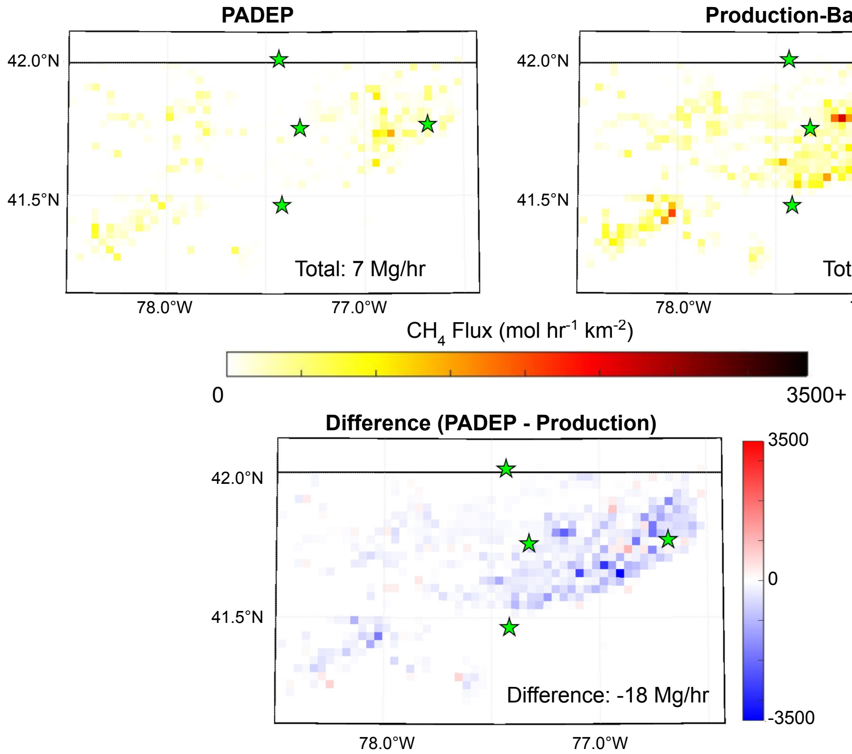

Figure S7. A comparison between O&G emissions from the two priors used in this study for the Delaware basin. (left) An emission map of unconventional natural gas activity constructed by the Pennsylvania Department of Environmental Protection. (right) An alternative emissions map created by taking the annual production of unconventional gas wells during the 2015-2016 time period and assuming a mean emission rate of 0.4% of production. (bottom) The difference between the two priors.

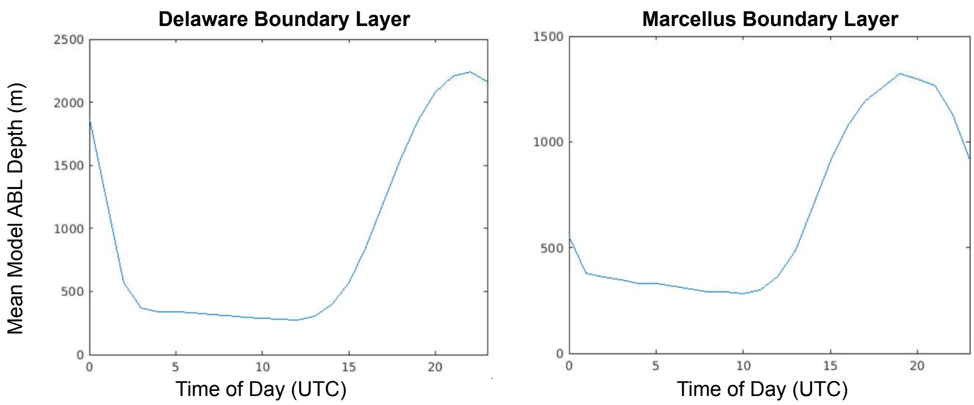

Figure S8. Average model boundary layer depths in the Delaware and Marcellus basins based on time of day.

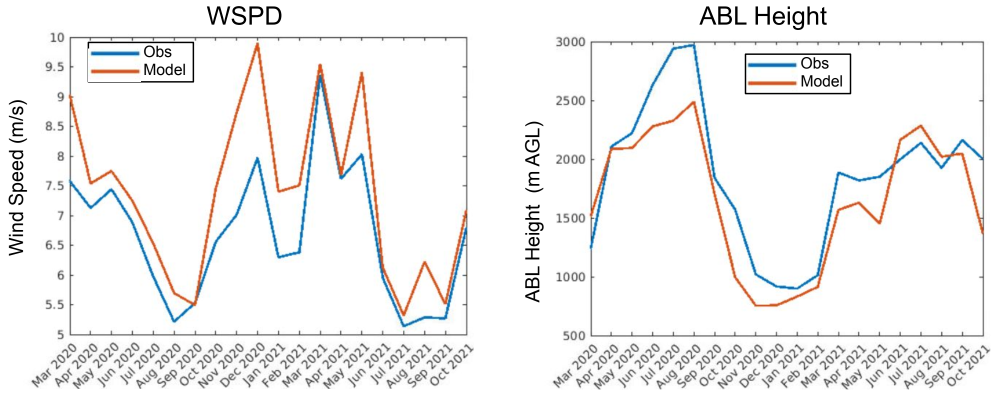

Figure S9. Observed vs modelled monthly mean boundary layer wind speeds and boundary layer heights at the location of the radiosonde in Midland, Texas (location: 31.95°N, 102.18°W).

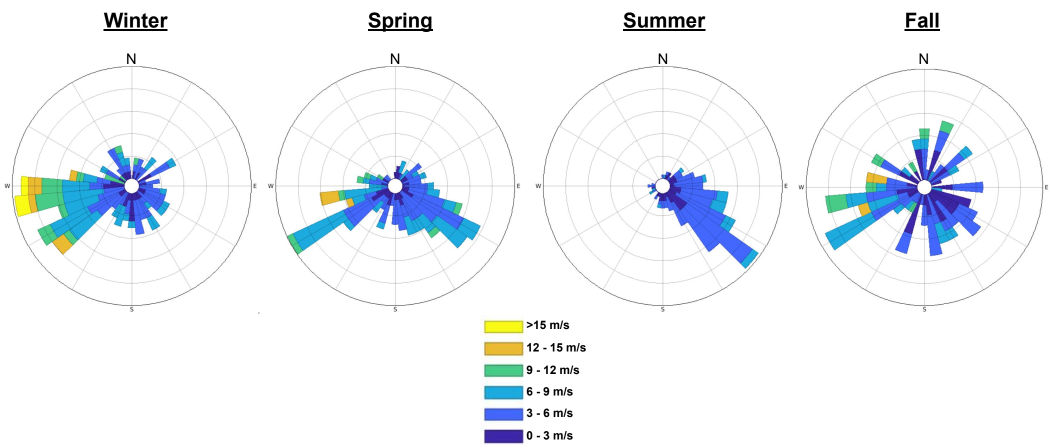

Figure S10. A wind rose for the Delaware basin showing the speed and frequency of afternoon winds by season. Wind data is 100 m AGL from the WRF-Chem model simulation averaged over the study domain.

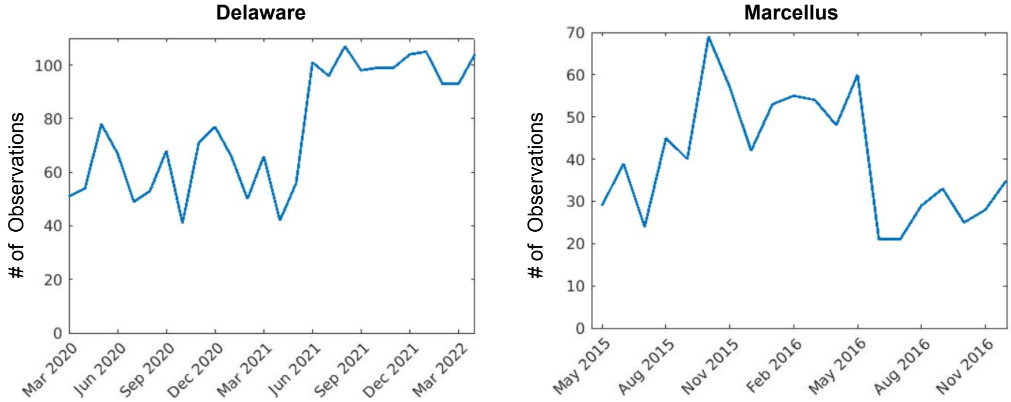

Figure S11. Number of afternoon downwind tower observations used in the Delaware and Marcellus inversions for each month

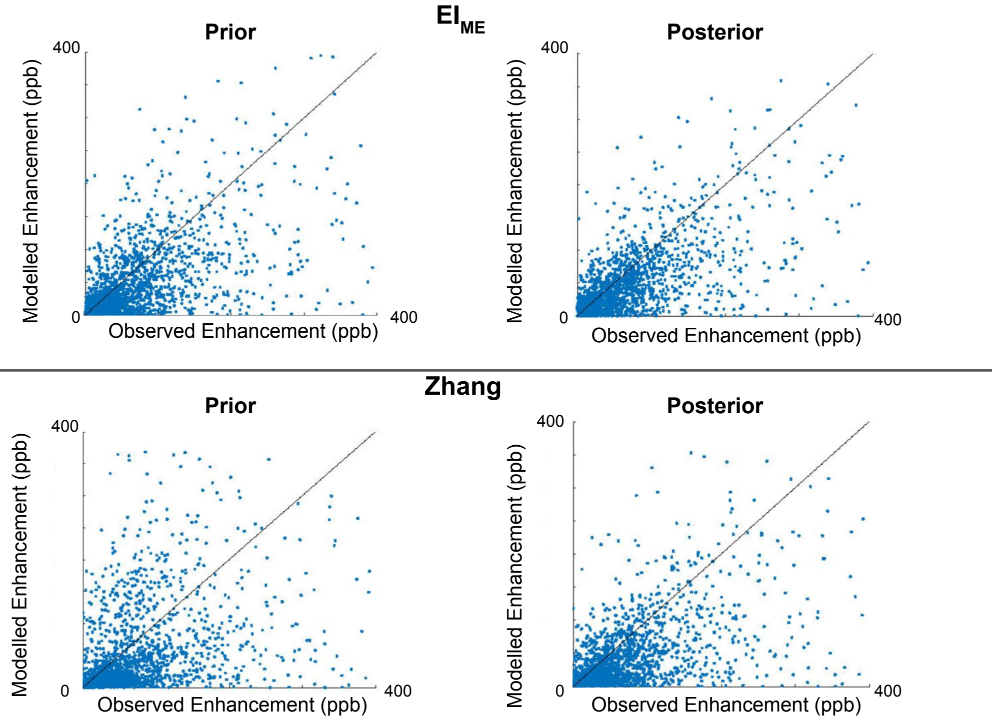

Figure S12. (top) A comparison of modelled vs observed O&G methane enhancement for based on the EI_{ME} prior and monthly posterior emission maps. (bottom) Similar to top, but using the Zhang prior and monthly posterior emission maps. The black line on all plots is the identity line.

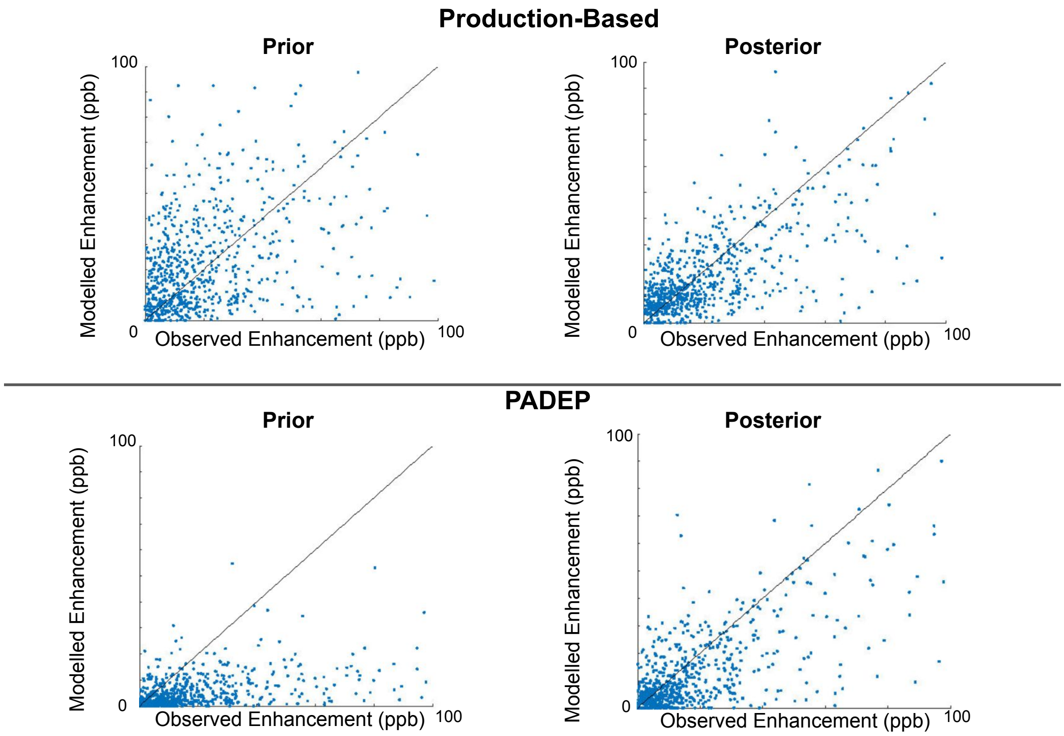

Figure S13. (top) A comparison of modelled vs observed O&G methane enhancement for based on the Production-based prior and monthly posterior emission maps. (bottom) Similar to top, but using the PADEP prior and monthly posterior emission maps. The black line on all plots is the identity line.

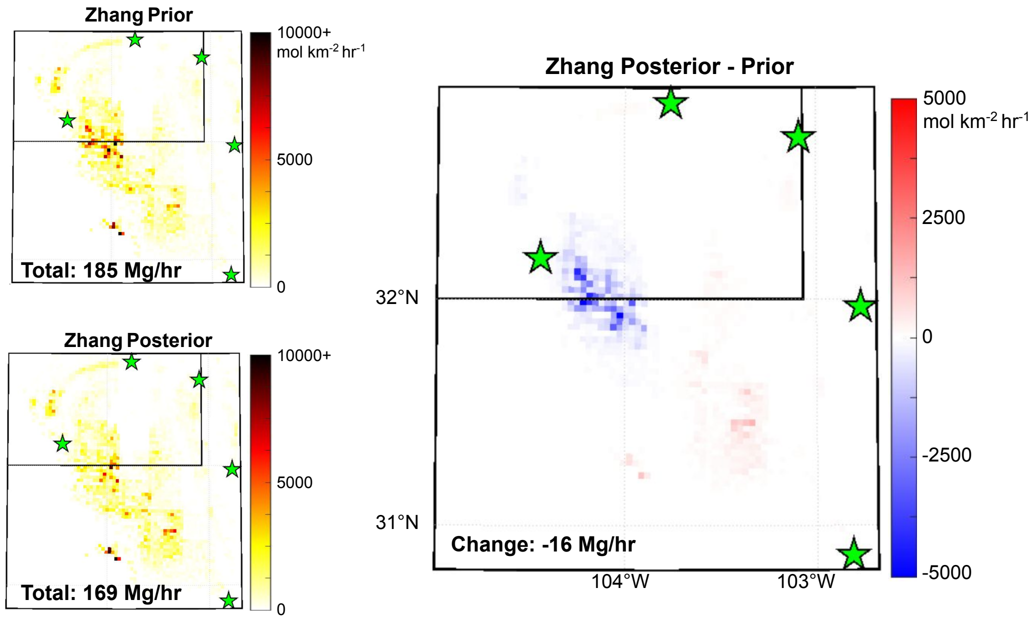

Figure S14. Prior, posterior, and difference between the Zhang prior and posterior maps for the Delaware domain, averaged across all months.

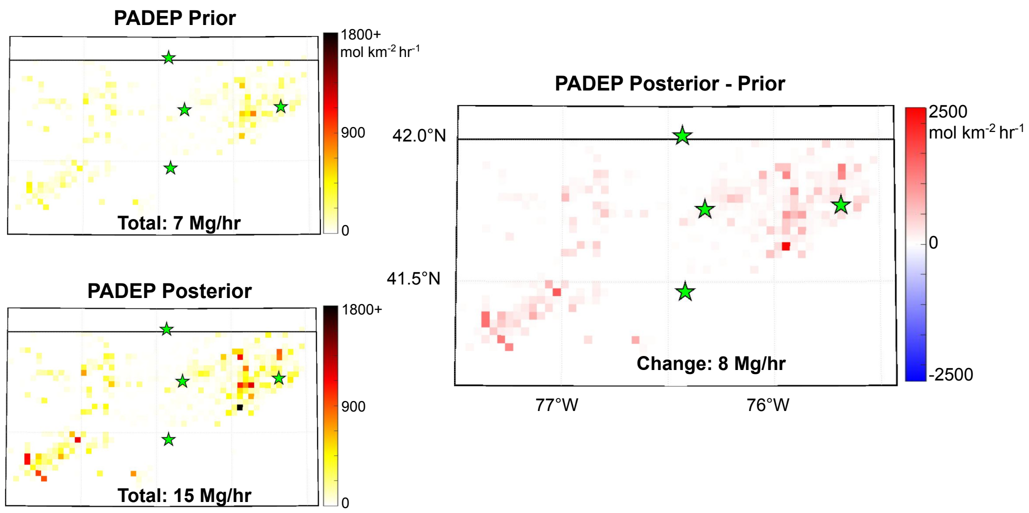

Figure S15. Prior, posterior, and difference between the Production-based prior and posterior maps for the Marcellus domain, averaged across all months.

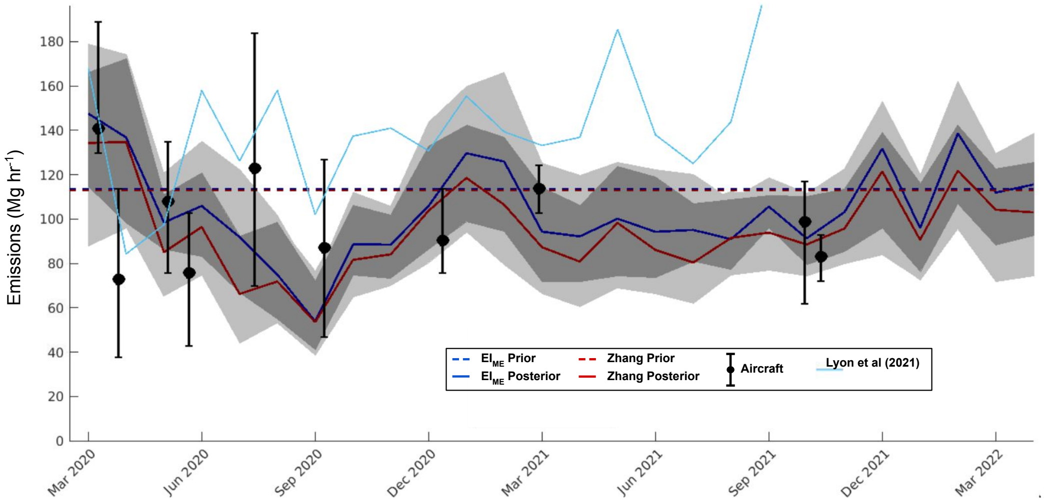

Figure S16. Prior and posterior monthly O&G methane emission totals for the 100 x 100 km Delaware domain used in Lyon et al. (2021) based on the EI$_{ME}$ prior (blue) and Zhang prior from this study (red). The shaded area represents the minimum and maximum emission rate for each month based on the range of results by adjusting the inversion as described in the sensitivity analysis in supplemental section S1. Emission results from the aircraft campaign performed in the same domain from Lyon et al. (2021) are plotted overtop, as are the monthly mean emission estimates using techniques from the tower analysis in Lyon et al. (2021)

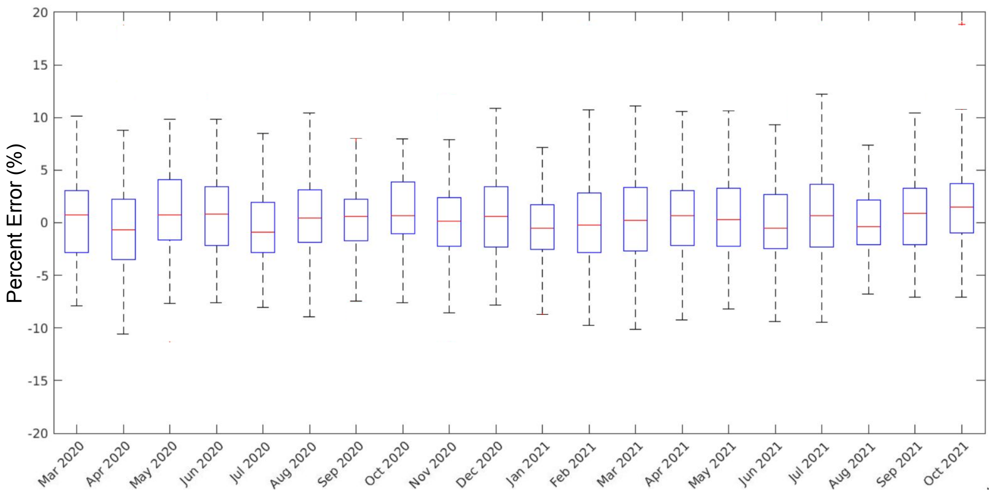

Figure S17. Box plot from the OSSE experiment in the Delaware basin showing the range of errors of the monthly posterior emission rates compared to the true mean monthly emission rate over the course of 100 simulations.

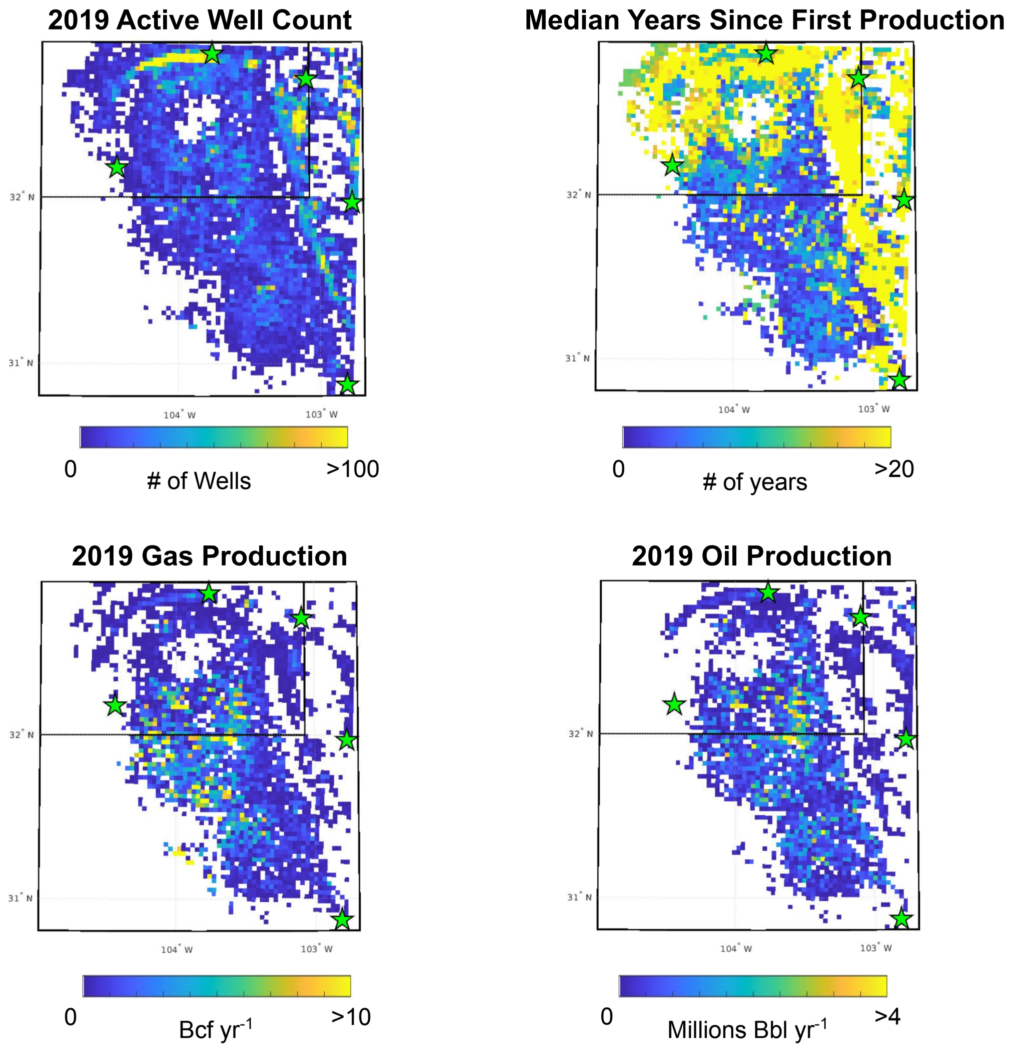

Figure S18. Well counts, median years since production, annual natural gas production, and annual oil production inside the Delaware study domain for the year 2019. Values are aggregated at 3 x 3 km resolution matching the model grid information from this study.

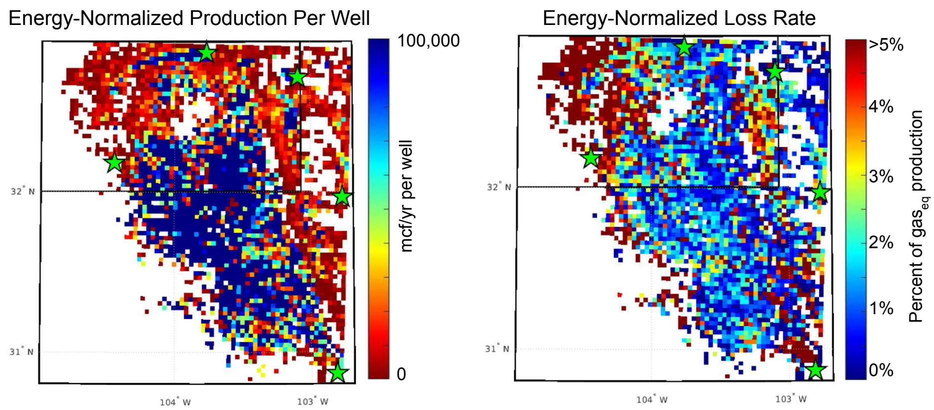

Figure S19. (left) Energy-normalized average production per well in each 3 x 3 km grid of the study domain. 1 barrel of oil = 7 mcf of gas produced assuming a methane composition of 80% in the Delaware basin. (right) Average energy-normalized loss rates in the study domain based on the EI$_{ME}$ posterior.

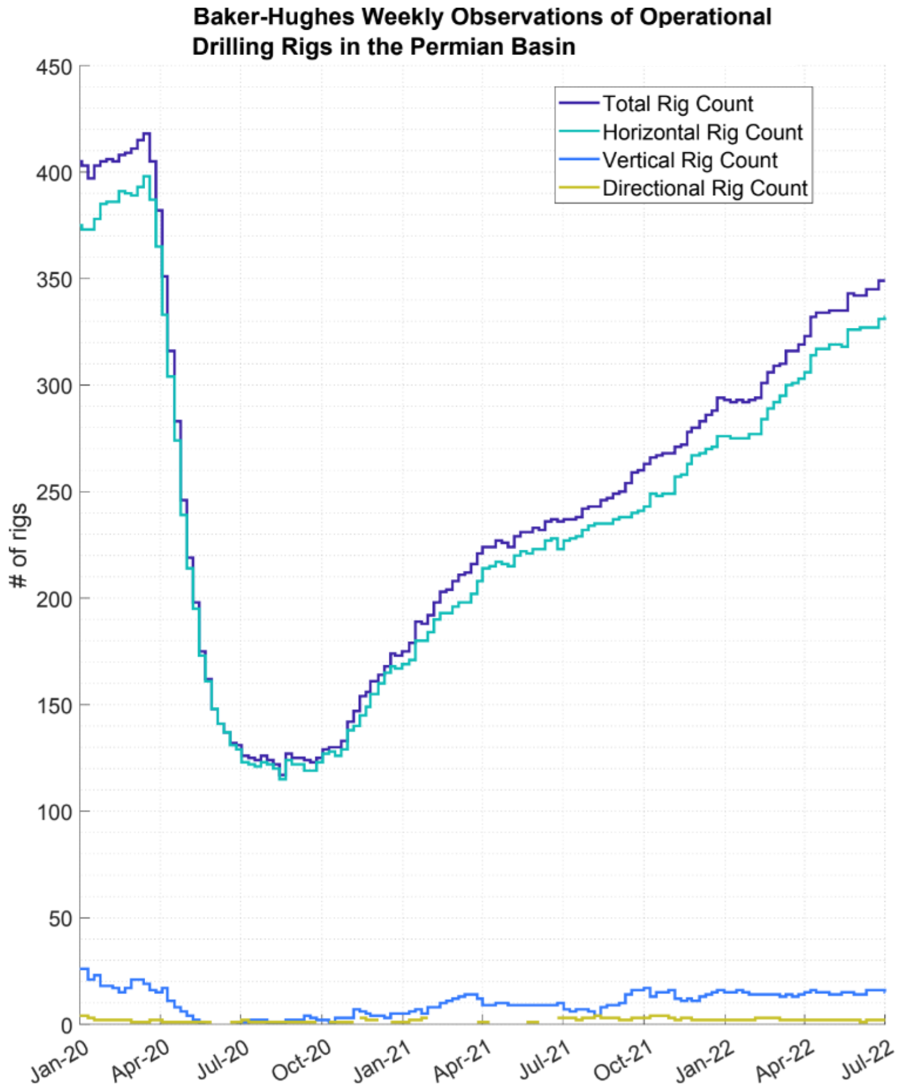

Figure S20. Rigs count in the Permian basin based on data from Baker Hughes (https://rigcount.bakerhughes.com/na-rig-count).

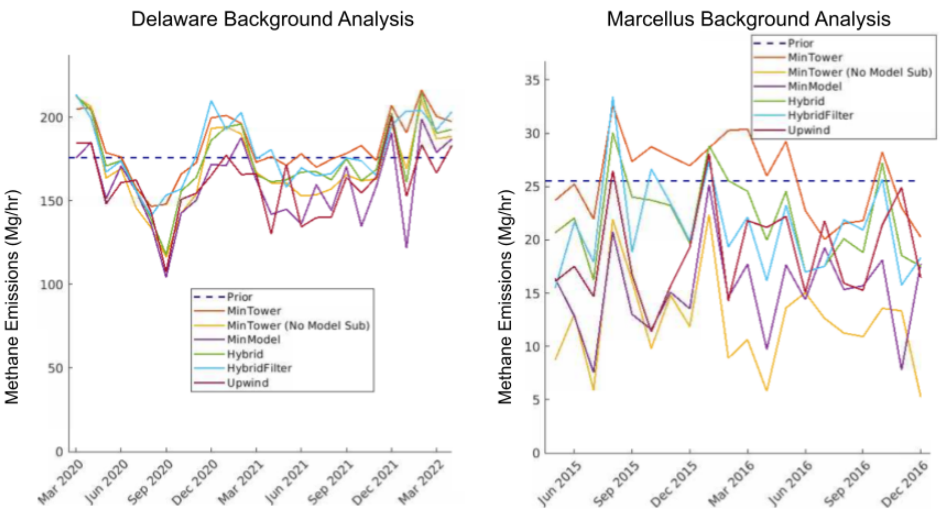

Figure S21. Posterior timeseries of monthly emission rates for the Delaware and Marcellus basins using the various background selection methodologies outlined in supplemental section 2. For the Delaware, the prior used was the EI_{ME}. For the Marcellus, the prior used was the Production-based prior.

DELAWARE

	EI_{ME} Prior	Zhang Prior	EI_{ME} Posterior (Default)	Zhang Posterior (Default)
Mean Total O&G Emissions (Mg/hr)	176	185	174	172
Mean Absolute Error (ppb)	46	55	41	46
Model Bias (Model - Obs) (ppb)	-21	-20	-20	-28
Correlation of Model, Obs Enhancements	0.49	0.46	0.60	0.58

MARCELLUS

	PADEP Prior	Production-Scaled Prior	PADEP Posterior (Default)	Production-Scaled Posterior (Default)
Mean Total O&G Emissions (Mg/hr)	7	25	14	22
Mean Absolute Error (ppb)	15	18	11	11
Model Bias (Model - Obs) (ppb)	-11	3	-4	-3
Correlation of Model, Obs Enhancements	0.37	0.44	0.66	0.74

Table S1. (left) Table describing the performance of the monthly posterior emission maps using the default settings relative to the priors for the Delaware basin. (right) Same as left, but for the northeast Marcellus basin.

	EI_{ME} Prior	EI_{ME} Posterior (Default)	EI_{ME} Posterior (Prior x 1.5)	EI_{ME} Posterior (Prior x 0.5)	EI_{ME} Posterior (Early Hours)	EI_{ME} Posterior (Flux Error x2)	EI_{ME} Posterior (10 km correlation)	EI_{ME} Posterior (Alternative Background)
Mean O&G Emissions (Mg/hr)	176	174	210	146	160	178	178	160
Mean Absolute Error (ppb)	46	41	37	38	48	37	40	47
Model Bias (Model - Obs) (ppb)	-21	-20	-12	-19	-23	-16	-18	-7
Correlation of Model, Obs Enhancements	0.49	0.60	0.68	0.68	0.60	0.68	0.63	0.51

Table S2. Table describing the performance of the monthly posterior emission maps for the Delaware basin created using the various methods described in Section S1 using the EI_{ME} prior.

	Zhang Prior	Zhang Posterior (Default)	Zhang Posterior (Prior x 1.5)	Zhang Posterior (Prior x 0.5)	Zhang Posterior (Early Hours)	Zhang Posterior (Flux Error x2)	Zhang Posterior (10 km correlation)	Zhang Posterior (Alternative Background)
Mean O&G Emissions (Mg/hr)	185	172	219	146	156	183	177	156
Mean Absolute Error (ppb)	55	46	41	43	54	41	44	52
Model Bias (Model - Obs) (ppb)	-20	-28	-19	-29	-31	-24	-27	-15
Correlation of Model, Obs Enhancements	0.46	0.58	0.65	0.65	0.57	0.65	0.61	0.48

Table S3. Table describing the performance of the monthly posterior emission maps for the Delaware basin created using the various methods described in Section S1 using the Zhang prior.

	Production Prior	Production Posterior (Default)	Production Posterior (Prior x 1.5)	Production Posterior (Prior x 0.5)	Production Posterior (Early Hours)	Production Posterior (Flux Error x2)	Production Posterior (10 km correlation)	Production Posterior (Alternative Background)
Mean O&G Emissions (Mg/hr)	25	22	28	19	21	23	22	20
Mean Absolute Error (ppb)	18	11	9	9	14	9	11	13
Model Bias (Model - Obs) (ppb)	3	-3	-3	-2	-4	-2	-3	-1
Correlation of Model, Obs Enhancements	0.44	0.74	0.80	0.80	0.71	0.81	0.74	0.69

Table S4. Table describing the performance of the monthly posterior emission maps for the Marcellus basin created using the various methods described in Section S1 using the Production-based prior.

	PADEP Prior	PADEP Posterior (Default)	PADEP Posterior (Prior x 1.5)	PADEP Posterior (Prior x 0.5)	PADEP Posterior (Early Hours)	PADEP Posterior (Flux Error x2)	PADEP Posterior (10 km correlation)	PADEP Posterior (Alternative Background)
Mean O&G Emissions (Mg/hr)	7	14	19	16	13	18	17	12
Mean Absolute Error (ppb)	15	11	9	9	13	9	11	13
Model Bias (Model - Obs) (ppb)	-11	-4	-2	-3	-5	-3	-3	-2
Correlation of Model, Obs Enhancements	0.37	0.66	0.75	0.75	0.66	0.75	0.67	0.64

Table S5. Table describing the performance of the monthly posterior emission maps for the Marcellus basin created using the various methods described in Section S1 using the PADEP prior.

	EI_{ME} Prior	**EI_{ME} Posterior (Default)**	**OSSE Posterior (min max)**
Mean Total O&G Emissions (Mg/hr)	176	174	176-176
Mean Absolute Error (ppb)	46	41	4-7
Model Bias (Model - Obs) (ppb)	-21	-20	(-1)-1
Correlation of Model, Obs Enhancements	0.49	0.60	0.85-0.98

Table S6. Table describing the performance of the monthly posterior emission maps for the 100 OSSE-based inversions investigating the effects of intermittent emitters on error in the inverse solution. The range of values show the minimum and maximum values from the 100 simulations. The low mean absolute error, low bias, and high correlation in all cases relative to the real-world inversion show that intermittent sources do not explain the errors and biases we see using real observations.

Delaware Background Analysis

	MinTower	MinTower (no model subtraction)	MinModel	Hybrid	Hybrid Filter	Upwind
Available Days	782	782	782	782	554	716
Mean Background Mole Fraction (ppb)	1955	1973	1982	1967	1965	1980
Background 15 Day STD (ppb)	38	41	62	41	37	56
Posterior Flux (Mg/hr)	182	170	159	174	180	160
Correlation	0.55	0.58	0.51	0.60	0.65	0.51
MAE (ppb)	47	38	47	41	37	47
Bias (ppb)	-33	-19	-5	-20	-21	-7

Table S7. Inversion posterior statistics for the Delaware basin (EI_{ME} prior) using different background methods described in section S1.2. Available Days is the number of days with a calculable background using the specified methodology. Mean Background Mole Fraction describes the mean calculated background value across all available days for each method. Background 15 Day STD is the mean of a 15 day rolling standard deviation. Remaining rows are similar to those reported in previous tables, providing information on the performance of the posterior model enhancements relative to observations.

Marcellus Background Analysis

	MinTower	MinTower (no model subtraction)	MinModel	Hybrid	Hybrid Filter	Upwind
Available Days	531	531	531	531	311	385
Mean Background Mole Fraction (ppb)	1928	1953	1943	1938	1939	1945
Background 15 Day STD (ppb)	27	26	28	25	25	29
Posterior Flux (Mg/hr)	26	12	15	22	21	19
Correlation	0.56	0.65	0.71	0.74	0.73	0.69
MAE (ppb)	18	12	13	11	10	13
Bias (ppb)	-10	2	1	3	2	1

Table S8. Same as Table S7 but for the Marcellus basin (using the Production-based prior)