Apportionment and Inventory Optimization of Agriculture and Energy Sector Methane Emissions using Multi-month Trace Gas Measurements in Northern Colorado

Griffin J Mead[1], Daniel I Herman[1], Fabrizio R Giorgetta[1], Nathan A Malarich[1], Esther Baumann[1], Brian R Washburn[1], Nathan R Newbury[1], Ian Coddington[1], and Kevin C Cossel[1]

[1]National Institute of Standards and Technology, Spectrum Technology and Research Division

May 31, 2023

Abstract

Quantifying sector-resolved methane fluxes in complex emissions environments is challenging yet necessary for inventory validations. We separate energy and agriculture sector methane using a dynamic linear model of methane, ethane, and ammonia mixing ratios measured at a Northern Colorado site from November 2021 to January 2022. Combining observations with spatially resolved inventories and inverse methods, energy and agriculture methane fluxes are constrained across a ˜850 km^2 area. Optimized energy sector fluxes were 22% lower than the inventory despite a ˜360% increase in regional energy production since the inventory was constructed, suggesting a regional decline in emissions factors. In contrast, optimized agriculture fluxes were 3× larger than the inventory; we demonstrate this discrepancy is consistent with the spatial distribution of agricultural sources. These results highlight the utility of sector-apportioned methane observations for multi-sector inventory optimization in complex environments, which may prove valuable for national and global quantification of sector-resolved methane fluxes.

Apportionment and Inventory Optimization of Agriculture and Energy Sector Methane Emissions using Multi-month Trace Gas Measurements in Northern Colorado

Griffin J. Mead,[1] Daniel I. Herman,[1,2] Fabrizio R. Giorgetta,[1,2] Nathan A. Malarich,[1] Esther Baumann,[1,2] Brian R. Washburn,[1] Nathan R. Newbury,[1] Ian Coddington,[1] Kevin C. Cossel[1,*]

[1] National Institute of Standards and Technology, Spectrum Technology and Research Division, Boulder, CO 80305

[2] University of Colorado, Boulder, Department of Physics, Boulder, CO 80309

Corresponding author: Kevin Cossel (kevin.cossel@nist.gov)

Key Points:

- A dynamic linear model (DLM) extracts sector-apportioned methane from multi-month trace gas measurements in Northern Colorado.
- Bayesian optimization constrained by the DLM analysis indicates a regional decrease in oil and natural gas emissions factors since 2012.
- Optimized methane emissions from agriculture are higher than inventory predictions, in part due to spatial misallocation in the inventory.

Abstract

Quantifying sector-resolved methane fluxes in complex emissions environments is challenging yet necessary to improve emissions inventories and guide policy decisions. We separate energy and agriculture sector methane using a dynamic linear model analysis of methane, ethane, and ammonia measurements at a Northern Colorado site from November 2021 to January 2022. By combining observations with spatially resolved inventories and Bayesian inverse methods, energy and agriculture methane fluxes are optimized across a ~850 km^2 area. Posterior energy sector fluxes were 22% lower than the inventory despite a ~360% increase in regional energy production since the inventory was constructed, suggesting a regional decline in emissions factors. In contrast, optimized agriculture fluxes were 3× larger than inventory estimates; we demonstrate this discrepancy is consistent with differences in the actual vs. modeled spatial distribution of agricultural sources. These results highlight how sector-apportioned methane observations can yield multi-sector inventory optimizations even in complex environments.

Plain Language Summary

Knowledge of the locations, fluxes, and kinds of methane sources is important for implementing effective emissions mitigation technologies and regulations. Methane emissions are often challenging to categorize because a wide variety of sources can emit methane, and these disparate sources are often intermingled at the spatial resolution of gridded inventories. We demonstrate how a dynamic linear model can use multi-month time series of two tracer gases, ethane and ammonia, to effectively separate methane into contributions from the energy and

agriculture sectors. We further demonstrate how the sector-apportioned methane can be incoroporated into a Bayesian inversion analysis which refines the magnitude and distribution of the inventory's methane fluxes. Our results suggest that emissions factors for energy infrastructure have decreased three-fold since 2012 in the study area, and that our methodology is sensitive to the spatial distribution of methane sources in the region.

1 Introduction

While short-lived in the atmosphere, methane has ~30× greater global warming potential than carbon dioxide over a 100-year timescale. United States methane inventories estimate that the energy and agriculture sectors each contribute about a third of total annual U.S. anthropogenic emissions (Maasakkers et al., 2016). Refining energy and agriculture inventories is an important first step towards identifying emissions reduction strategies. This, however, is itself a difficult task: energy and agriculture infrastructures are often spatially overlapped at typical inventory resolutions, and there are challenges with attributing methane to one or the other sector. These hurdles must be overcome in observational studies seeking to optimize and contrain methane emissions from these two important sectors.

Here, we demonstrate how tracer gas measurememnts can help optimize regional energy and agriculture methane inventories despite substantial spatial overlap between the two sectors. Our study area is the Northern Colorado Front Range Urban Corridor (FRUC), where oil and natural gas infrastructure accessing the Wattenberg gas field are intermingled with large livestock developments (Figs 1a,b). We measured methane, ethane (a tracer gas for energy emissions), and ammonia (a tracer for agriculture) across a multi-month period using an open-path, mid-infrared dual-comb spectrometer (MIR-DCS) (Coddington et al., 2016; Giorgetta et al., 2021; Ycas et al., 2018) and a cavity ring-down spectrometer (CRDS). Unlike previously described, short duration tracer gas studies (Kille et al., 2019; Pollack et al., 2022; Yacovitch et al., 2014, 2015), our extended time series required a dynamic linear model (DLM) to capture variations in the tracer gas coefficients over time (West & Harrison, 1997). A Bayesian inversion then optimized energy and agriculture methane fluxes using the DLM-derived energy and agriculture-sector methane observations and an atmospheric transport model.

Optimized energy fluxes in the region around the measurement site were similar to 2012 inventory estimates despite a ~360% increase in energy production; this supports other findings which suggest the percent of methane emitted as a function of production has changed since the inventory was constructed (Lu et al., 2023; Peischl et al., 2018). In contrast, inferred agricultural methane fluxes were 3.0× greater than inventory estimates. We demonstrate that this discrepancy arises partially from the spatial distribution of livestock which is not captured in the inventory model. Our work highlights that tracer gas measurements can guide inventory optimizations even in complex emissions environemnts.

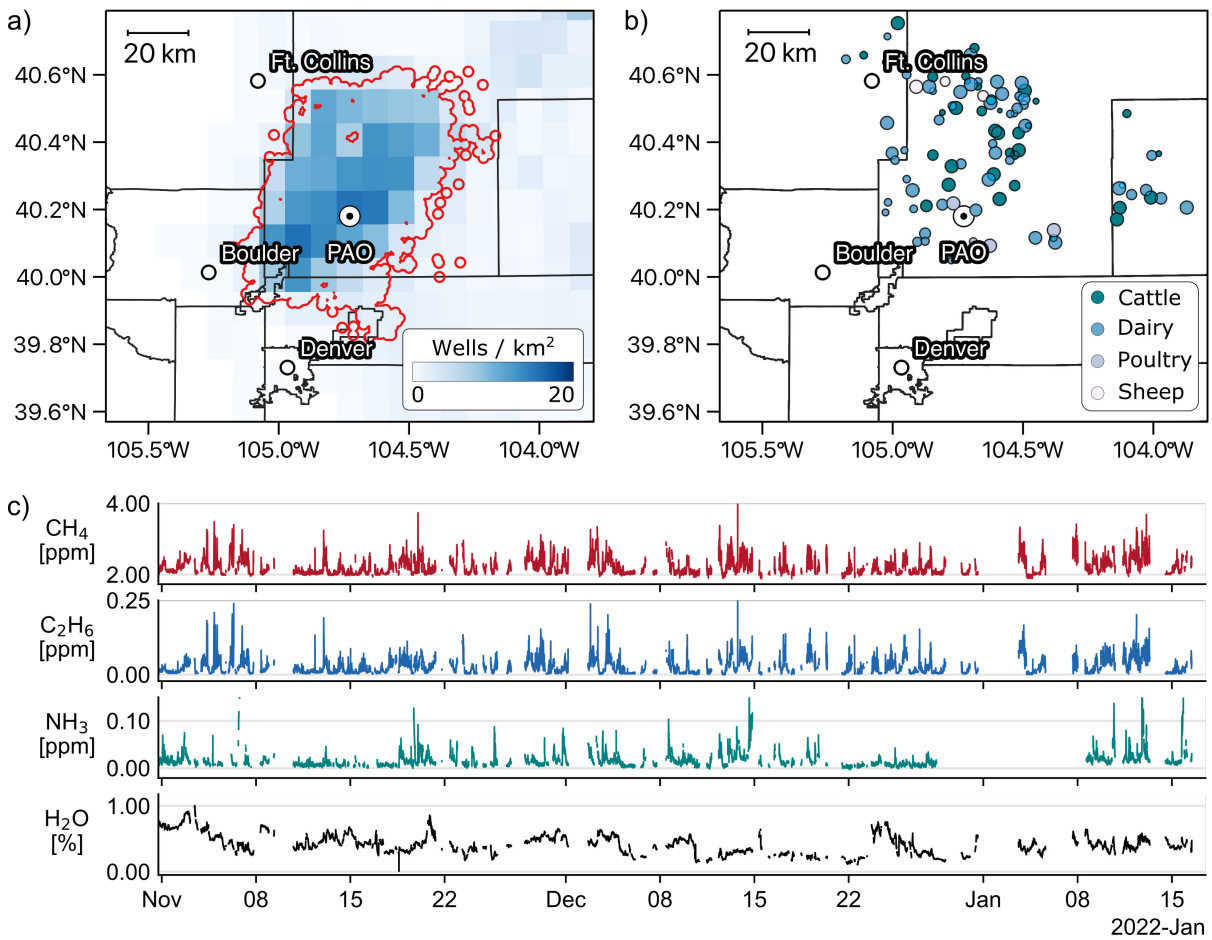


Figure 1 Energy and agriculture sources of methane are intermingled around the Platteville Atmospheric Observatory (PAO) measurement site. a) Thousands of wellheads (shown as a density map) extract oil and gas from the Wattenberg field (red outline). Locations of other down-stream components of the extraction process are not shown. Counties are outlined in black. b) Agricultural developments, in particular concentrated animal feeding operations (CAFOs, color coded by livestock and scaled to relative expected emissions magnitude), are widely distributed and spatially overlapped with energy infrastructure. c) The full multi-month methane, ethane, ammonia, (expressed as dry mixing ratios) and water time series recorded at PAO.

2 Materials and Methods

First we discuss the collection of time series methane and tracer gas data, and subsequent sector apportionment using a dynamic linear model. Next, we give a brief description of the atmospheric transport model and sector-resolved emissions inventory used in this work. Finally, we describe the Bayesian inversion approach which generates the optimized posterior emissions inventories.

2.1 Observational data collection

Methane (CH_4), ethane (C_2H_6), and water (H_2O) concentrations were measured at the Platteville Atmospheric Observatory (PAO) from 1 November 2021 to 17 January 2022 with an

open-path MIR-DCS instrument; ammonia (NH_3) was measured with a commercial CRDS. Ammonia data were interpolated onto the 2-minute MIR-DCS time base, which is set by the MIR-DCS spectral averaging time. Figure 1c shows the interpolated dry air mole fractions CH_4, C_2H_6, and NH_3 time series, reported in ppm [umol/mol] and/or ppb [nmol/mol]. Subsequent analysis relies on periods where all three species were measured. A map of the measurement site is provided in Figure S1.

The MIR-DCS system is similar to previously reported designs (Giorgetta et al., 2021; Ycas et al., 2019, 2020) and is described in more detail in another publication (D. Herman et al., 2023). Briefly, the instrument measures an optical bandwidth spanning 2400 cm^{-1} to 3100 cm^{-1} with 0.006 cm^{-1} spectral resolution (Figure S2b). A telescope transmits MIR light along a 380 m open-air path to a retroreflector ~5 m above ground level. The reflected light is collected by the transmit/receive telescope, detected by a thermoelectrically cooled mercury cadmium telluride detector, digitized at 200 MHz, and coherently averaged (Roy et al., 2012; Ycas et al., 2018). Path-averaged CH_4, C_2H_6, and H_2O concentrations were retrieved from the spectra using the HITRAN2020 database (Gordon et al., 2022).

2.2 Dynamic linear model tracer gas analysis

Energy and agriculture contributions in a methane time series can be extracted using correlations with ethane and ammonia (Kille et al., 2019). Generally this is achieved by fitting the methane data to a linear regression model comprised of energy sector methane ($y_{Energy} = \beta_1\,[C_2H_6]$), agricultural sector methane ($y_{Agri} = \beta_2\,[NH_3]$), a background term ($\beta_0$), and a Gaussian noise term (ϵ):

$$[CH_4] = \beta_0 + \beta_1[C_2H_6] + \beta_2\,[NH_3] + \epsilon$$

This model is appropriate for the FRUC region since the majority of methane emissions are from energy and agriculture. While landfills can emit substantial volumes of methane, landfill emissions are not included in the analysis because all major sites were outside this work's area of sensitivity.

Fluctuations in the β_0, β_1, and β_2 tracer gas coefficients are expected duriung the multi-month study at PAO; the background methane concentration β_0 varies diurnally, and the two tracer gas coefficients, β_1 and β_2, change as different methane sources are transported to PAO. Since a static linear regression analysis cannot capture these variations, and to avoid dividing the time series into arbitrarily smaller segments, we perform the tracer gas analysis using a dynamic linear model (West & Harrison, 1997). Methane data are modelled with the observation equation,

$$[CH_4]_t = F_t'\theta_t + \nu_t, \qquad \nu_t \sim N[0, V_t],$$

and the system equation,

$$\theta_t = \theta_{t-1} + \omega_t, \qquad \omega_t \sim N[0, W_t],$$

where *t* is an index representing data time steps. Tracer gas observations, along with a constant unity term which models the intercept, are represented by the regression vector $F_t =$

$(1, [C_2H_6]_t, [NH_3]_t)$. Observations are assumed subject to Gaussian noise ν_t with a mean of zero and a variance V_t (defined here as the variance of the point-wise difference of the methane time series). The state vector $\theta_t = (\beta_{0,t}, \beta_{1,t}, \beta_{2,t})$ evolves over time as a function of the θ_{t-1} state vector and the evolution variance vector W_t. Because the variance is difficult to directly estimate and may not be time-invariant, DLMs are often solved using a discount factor δ instead as a proxy for the "memory" of the system over time (West & Harrison, 1997). The discount factor is defined as $\delta = P_t/(W_t + P_t)$, where P_t is the prior variance corresponding to a state vector with zero stochastic change ($W_t = 0$). In that limiting case, $\delta = 1$ (irrespective of the actual value of P_t) and the DLM is identical to a static linear regression model. An optimal discount factor can be determined through minimizing the model's mean standard error, but in practice this minimization becomes expensive for large data sets. For this analysis, 100 DLM fits were performed over the full times series data; discount factors were sampled from a random uniform distribution spanning [0.98,0.999]; the mean values from the 100 DLM fits are used throughout. (Discount values below 0.98 lead to numerical instability; data where the fractional variance of either β_1 or β_2 was greater than 100% of the fit value are excluded in subsequent analysis.)

2.3 Atmospheric transport modelling

We use the STILT-R atmospheric transport model and 3-km High Resolution Rapid Refresh (HRRR) meteorological data to calculate influence footprints in a 3°×3° domain centered on PAO (Benjamin et al., 2016; Fasoli et al., 2018; Lin, 2003). Each influence footprint $\boldsymbol{H}(z_r, T_r | z_i, T_i)$ (units of [ppm m^2 s / μmol CH_4]) connects sector-specific emissions throughout the spatial domain, at location z_i and time T_i, to observed sector-apportioned methane mixing ratios at PAO (z_r) at time T_r. Footprints were calculated for each hour in an 8-week period of observations from November and December 2021. Each footprint is the sum of a 2-day duration back trajectory of 100 particles originating from PAO, calculated at 0.1° resolution and hourly step size with hyper near field effects enabled.

2.4 Emissions inventories

Energy and agriculture emissions are estimated using 0.1°×0.1° sector-resolved methane flux maps derived from the 2012 EPA national methane inventory (Maasakkers et al., 2016). The energy inventory, x_{Energy}, is the sum of IPCC categories 1B2b (Natural Gas Production + Processing + Transmission + Distribution) and 1B2a (Petroleum); coal methane emissions are not considered (IPCC, 1996). The agriculture inventory, x_{Agri}, is the sum of IPCC categories 4A (Enteric Fermentation) and 4B (Manure Management).

2.5 Bayesian inversion

Each sector-resolved methane time series (y_{Energy} and y_{Agri}) can be modelled as the sum of the product of a time-independent methane inventory (x_{Energy} and x_{Agri}) and the time-varying series of influence footprints $\boldsymbol{H}$ at each grid cell, plus an error term ϵ,

$$y_{Energy} = \boldsymbol{H}\, x_{Energy} + \epsilon$$
$$y_{Agri} = \boldsymbol{H}\, x_{Agri} + \epsilon$$

Bayesian inverse modelling uses observational constraints ($y^{Obs}_{Energy/Agri}$) to generate maximum a posteriori (MAP) inventory estimates, $x^{Posterior}_{Energy/Agri}$, using the prior information provided by the

inventories, $x^{Prior}_{Energy/Agri}$ (D. H. Cusworth et al., 2020). The observation vector $y^{Obs}_{Energy/Agri}$ are the hourly mean mixing ratios of energy and agriculture methane averaged from the 2-minute time series. Following other studies, data are restricted to between the hours 11-16 local time when the boundary layer is well mixed and better captured by the meteorological models, for a total of 238 data points for each observation vector (Fasoli et al., 2018; Kunik et al., 2019; McKain et al., 2015; Sargent et al., 2018). The $\boldsymbol{H}$ matrix contains the corresponding STILT footprint for each hour, where each footprint is restricted to an area +/- 2.9° latitude and longitude centered on PAO at 0.1° resolution for a total of 3422 state vector elements; footprints were flattened and stacked to yield the final $\boldsymbol{H}$ matrix with shape (238×3422). The priors $x^{Prior}_{Energy/Agri}$ were interpolated onto the grid of the STILT footprints. Optimization of the prior and observational error covariance matrices required for the MAP estimation is discussed in the SI (Michalak, 2004; Michalak et al., 2005). The averaging kernel sensitivity matrix (Figure S4) indicates the posterior is constrained by observations in an 850 km^2 area centerd around PAO. This region is highlighted with a dashed rectangular outline in Figs 3 and 4.

3 Time-resolved sector apportioned methane

We first examine the dynamic linear model tracer gas results which provide the key observational constraints for the Bayesian inversion. Three illustrative examples are shown in Figure 2. The DLM analysis captures not only how each tracer gas coefficient varies as different sources are transported to PAO but also how uncertainty in the coefficients evolve. During periods with a low tracer gas concentration or little variation in the tracer gas, uncertainty in the respective coefficient increases. Alternatively, a sharp increase in one tracer gas concentration rapidly shrinks the uncertainty in the respective DLM coefficient.

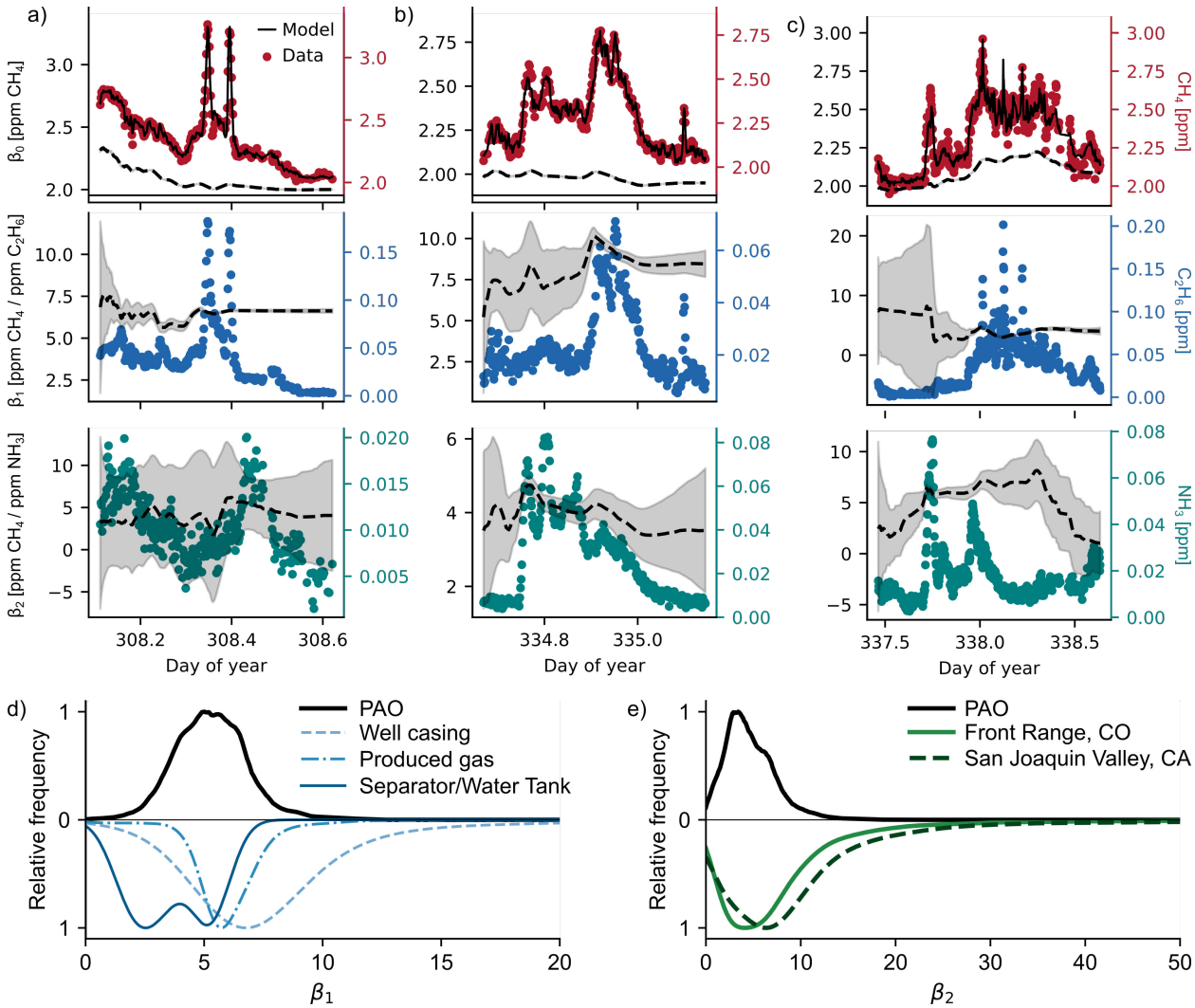


Figure 2 Three methane plumes (a,b,c) illustrate how the DLM apportions methane into contributions from the energy and agriculture sectors. The tracer gas coefficients (dashed lines, left axis) are shown with uncertainties in gray shading. In addition, the top panel shows both the full modeled methane concentration (solid line, right axis) and the measured methane concentration (red circles, right axis). The second and third rows show the ethane and ammonica measurements (colored dots, right axis). Panels d) and e) compare the β_1 and β_2 coefficients from the full time series to other literature. The range of β_1 coefficients observed at PAO are consistent with coefficients calculated from COGCC sampling data and may reflect contributions from different sources around PAO. e) β_2 coefficients at PAO are consistent with other studies performed in Colorado (Eilerman et al., 2016) and California (Miller et al., 2015).

DLM analysis produces tracer gas coefficient time series which can provide insight into emission source characteristics. Figure 2d,e show kernel density estimates of the energy (β_1) and agriculture (β_2) tracer gas coefficients over the multi-month observation period. In the case of β_1, this ratio has been observed to vary as natural gas is extracted, processed, and transported (Cardoso-Saldaña et al., 2019; Peischl et al., 2013). Ethane and methane mole fractions for natural gas samples collected after 2010 in the Front Range Urban Corridor by the Colorado Oil

and natural gas Conservation Commission (COGCC) provide a direct comparison to our estimates for β_1 (Figure 2d) (Colorado Oil and Gas Conservation Commission, 2022). COGCC recorded data for a range of sample locations, including well casings (bradenheads, well tubing, and surface, intermediate, and production casings), produced gas, and separators and water tanks. The β_1 values determined from the PAO data span the lower end of values for well casing and the higher range of values for separator and water tank emissions, but are most consistent with produced gas emissions.

Similarly, β_2 is expected to vary as emissions from different livestock species can have substantially different ratios of methane and ammonia concentrations (Golston et al., 2020). Other sources of variation could include atmospheric chemical effects such as deposition and reactivity (primarily for NH_3). We compare our β_2 results with two mobile measurement studies in Figure 2e. While extensive sampling of ammonia/methane ratios throughout the state are not available, studies in both the San Joaquin Valley of California and the FRUC overlap well with β_2 results obtained at PAO, indicating a consistent, if broad, distribution of β_2 values for agriculture across the western United States (Eilerman et al., 2016; Miller et al., 2015).

Significant day-to-day variations in tracer gas coefficients observed in this analysis emphasizes the difficulty determining a unique set of energy and agriculture coefficients, even for measurements conducted in a single location (Lan et al., 2019). Despite these complexities, the DLM approach successfully generates energy and agriculture sector-apportioned methane time series which will serve as the observational constraints for inventory optimization.

4 Methane inventory optimization

Changes in emissions between the prior and posterior inventories are reported in several ways. First, we report the mean absolute differences between the observed time series y^{Obs}, and the prior and posterior predicted time series, y^{Prior} and $y^{Posterior}$. Second, we report Welch's two-sided t-test results which compare mean mixing ratios of y^{Obs} to y^{Prior} and $y^{Posterior}$. Third, we compare mean fluxes from x^{Prior} and $x^{Posterior}$ within the 850 km^2 region identified by the averaging kernel sensitivity matrix.

4.1 Energy sector

A mean absolute difference of 15.36 ± 55.98 ppb CH_4 between y^{Obs}_{Energy} and y^{Prior}_{Energy} was reduced in $y^{Posterior}_{Energy}$ to 11.87 ± 25.76 ppb CH_4. The two-sided Welch's t-test found that observations were indistinguishable from both the predicted prior (t=-1.15, p-value=0.27) and posterior (t=0.18, p-value = 0.85) mixing ratios, consistent with the minor changes in the mean difference between the time series. Spatially, mean energy fluxes within the region of maximum sensitivity (dashed rectangle in Figure 3) were 22% lower in the posterior solution (78.4 ± 3.5 nmol CH_4 m^{-2} s^{-1}) compared to the prior (100.0 ± 53.0 nmol CH_4 m^{-2} s^{-1}), with the posterior state vector emissions slightly reduced towards the north-east of PAO (Figure 3c).

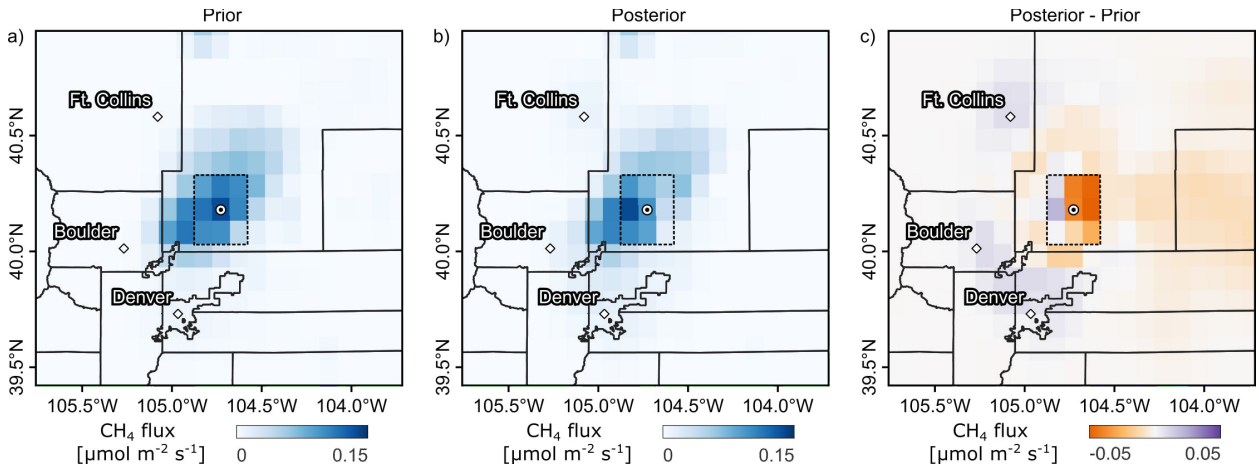


Figure 3 The spatial distribution of methane emissions from the energy sector are optimized with a Bayesian inversion using the energy sector methane time series observed at PAO. a-b) Prior (x_{Energy}^{Prior}) and posterior ($x_{Energy}^{Posterior}$) surface flux maps for energy sector methane emissions remain largely similar in both distribution and magnitude of emissions. c) Difference between prior and posterior emissions indicate a slight reduction in emissions north-east of PAO.

It is noteworthy that the means of y_{Energy}^{Obs} (measured in 2021) and y_{Energy}^{Prior} (calculated from the 2012 EPA inventory) are within ~20% of each other. Between 2012 to 2021, the Wattenberg field's oil and natural gas production volumes increased by 370% and 360% respectively. Several aircraft mass-balance studies of the Wattenberg field from 2008 to 2021 show relatively constant emissions over this time period that are also consistent with the EPA inventory (D. Cusworth et al., 2022; Peischl et al., 2018; Pétron et al., 2012, 2014). Thus, the agreement between our observations and the 2012 inventory suggests either that 1) emissions factors have declined since 2012, or 2) production within the sensitivity region around PAO remained unchanged relative to 2012. An spatially gridded dataset reporting annual oil and gas production volumes and new well installations (Skinner et al., 2022) demonstrates that, although the distribution of production across the region did become more localized and heterogeneous, production immediately around PAO increased at roughly the same rate as the Wattenberg field overall. This indicates that emissions factors have likely declined since 2012.

While similar trends have been seen across the US (Lu et al., 2023), the exact causes remain undetermined. Data from Skinner et al., 2022 indicate a significant change in regional well infrastructure and production volumes over the past decade. Following trends in the Permian and other major basins, horizontally drilled well installations became ubiquitous in the Wattenberg field between 2010-2012; large increases in oil and natural gas production followed shortly after. Given the correlated change in well infrastructure and extraction efficiency, we speculate that horizontal well emissions factors differ from those used to construct the inventory model. Reasons for this could include the consolidated infrastructure and multiple well heads at horizontally drilled sites, which may lead to higher rates of successful leak detection and repair than traditional dispersed, vertically drilled single well installations (Robertson et al., 2017). Colorado State's 2014 adoption of stricter air quality standards may have further mitigated energy sector emissions. Indeed, the observed ~3.6× decrease in regional emissions factors since

2012 is significantly larger than the US-average decrease of 1.6× (Lu et al., 2023), further hinting that state regulations had a significant impact.

4.2 Agriculture sector

In contrast to the energy sector, y_{Agri}^{Obs} was 3.1× greater than y_{Agri}^{Prior}, with a mean hourly difference of 32.21 ± 25.72 ppb CH_4. This difference reduced to 8.40 ± 17.33 ppb CH_4 in $y_{Agri}^{Posterior}$ as methane fluxes around PAO increased from 13.7 ± 16.4 nmol CH_4 m^{-2} s^{-1} to 41.4 ± 1.1 nmol CH_4 m^{-2} s^{-1} (Figure 4a,b). Posterior mean mixing ratios were indistinguishable from observations (t=1.35, p-value=0.20) in contrast to the distinctly different prior mean mixing ratio (t=10.25 , p-value=6.7e-6). The discrepancy between y_{Agri}^{Obs} and y_{Agri}^{Prior} is surprising given that the total permitted livestock population around PAO has remained roughly constant since 2012 (National Agricultural Statistics Service, n.d.). A threefold error in livestock emissions factors or increase since 2012 are both improbable; instead, we propose a spatial misallocation of regional emissions is likely to blame. Comparing the prior (Figure 4a) to locations of registered concentrated animal feeding operations (CAFOs, Figure 1b) clearly demonstrates that fluxes are not localized around CAFOs. This is a result of methodology: the agriculture inventory was generated by probabilistically distributing county-level livestock headcounts throughout each county using multiple livestock occurance probability maps (Maasakkers et al., 2016). For some livestock, such as beef cattle which graze in pastures for parts of the year, this is a logical approach; however, poultry and dairy cattle are often on CAFOs throughout the animal lifespan.

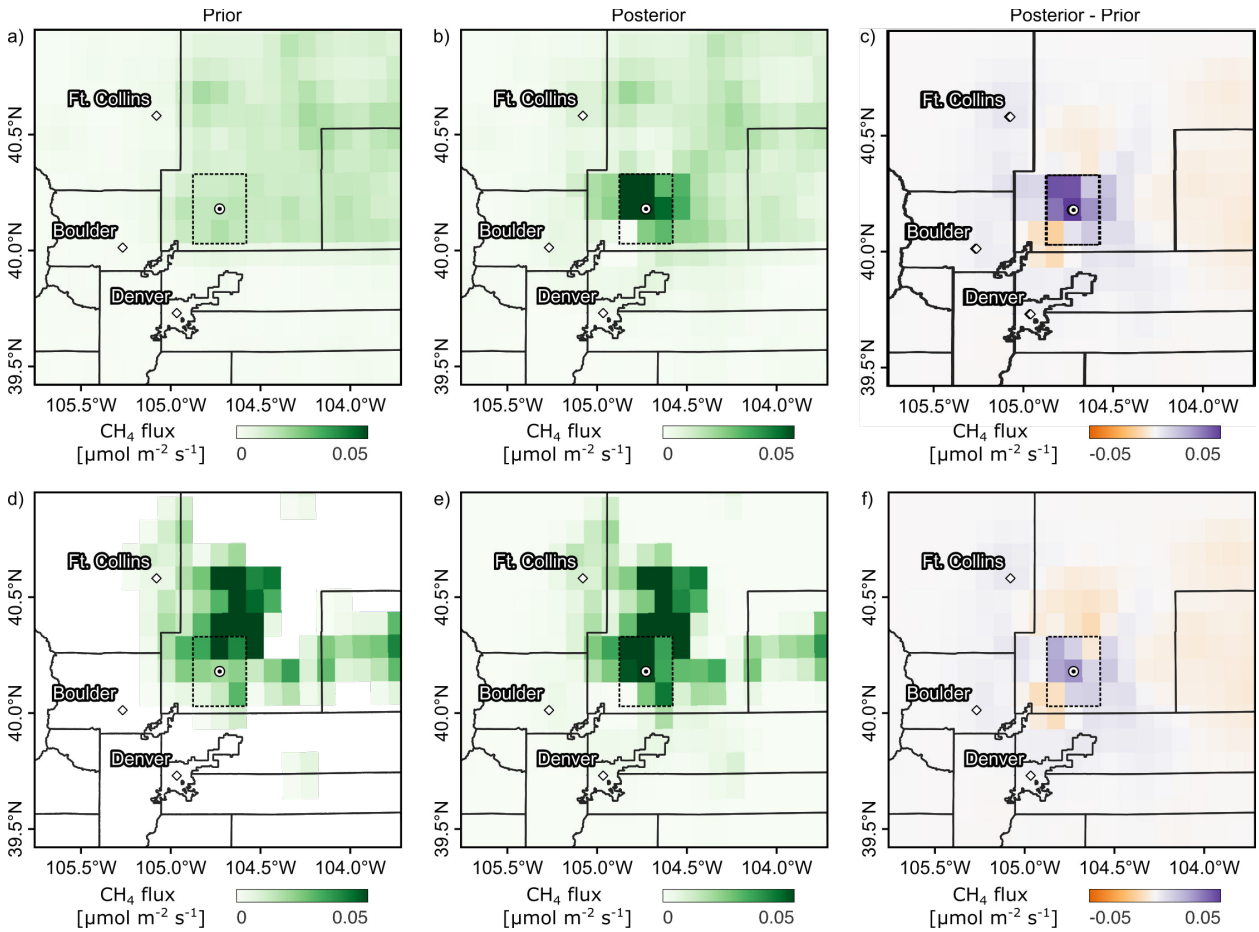


Figure 4 Comparison of agriculture-sector methane observed at PAO to the original inventory and posterior predictions (top row), and to the re-distributed inventory and posterior predictions (bottom row). a-b) Posterior ($x_{Agri}^{Posterior}$) agriculture methane are more localized around PAO than in the prior (x_{Agri}^{Prior}). c) Difference between prior and posterior emissions are significant, with a several-fold increase in emissions to the north-west. d-e) Redistributed posterior agriculture methane is distributed similarly to the redistributed prior. f) Differences between the re-distributed prior and posterior are slight, suggesting that the re-distribution better captures the distribution of emissions around PAO.

To determine if localizing emissions to CAFOs improves agreement with observations, county-level inventory emissions were calculated and redistributed to CAFO locations within each county proportionate to the total animal equivalent emissions units at each CAFO (Golston et al., 2020). Total county level emissions were unchanged, reflecting our assumption that agricultural emissions factors have remained constant. Redistributed emissions uncertainties were calculated using equations from Maasakkers. Differences between the redistributed prior (Figure 4d) and posterior (Figure 4e) were substantially smaller (Figure 4f) than those observed with the original inventory (Figure 4c). The $y_{Redist\ Agri}^{Prior}$ and y_{Agri}^{Obs} time series had a decreased mean absolute difference of 20.95 ± 40.99 ppb CH_4 although the two time series remained distinct (t=5.86, p-value =1.6e-3). Mean absolute difference was reduced by $y_{Redist\ Agri}^{Posterior}$ to 7.93 ± 17.38 ppb CH_4

and was found to be indistinguishable from y_{Agri}^{Obs} (t=1.17, p-value =0.27). Mean fluxes in $x_{Redist\ Agri}^{Posterior}$ were increased to 42.8 ± 1.9 nmol CH_4 m^{-2} s^{-1}, consistent with the $x_{Agri}^{Posterior}$ results.

5 Conclusions

We constrain energy and agriculture methane emissions in a ~850 km^2 region in the Front Range Urban Corridor by analyzing long-term measurements of methane, ethane, and ammonia with a dynamic linear model and Bayesian inversion. (While two instruments were used in this work, in the future all three gases could be measured using a single DCS instrument with adequate spectral coverage (D. I. Herman et al., 2021).) Comparison with the 2012 gridded EPA inventory showed a small decrease in energy sector methane emissions which is suggestive of a significant decrease in regional energy emissions factors from 2012 to 2021. Adoption of horizontal drilling and stricter state-level regulations around 2010-2014 may have contributed to these inferred changes in emissions factors. Furthermore, the significant increase and clustering of agricultural methane emissions in the posterior helped identify issues in the spatial composition of the regional agriculture inventory. Redistributing emissions to known CAFO locations reduced the spatial differences between the redistributed prior and posterior flux maps, although observations still suggest agriculture emissions are ~1.6× higher than even the redistributed inventory. Improvements in the spatial distribution of emissions in the inventory are critical for regional scale studies using aircraft or satellite observations where multiple tracer gas observations are not present (D. H. Cusworth et al., 2021; Peischl et al., 2018) . While conclusions from our single-sensor study can be improved with a distributed sensor network, it is noteworthy this approach can refine sector-resolved methane emission across areas comparable to the footprints of many methane observing satellites (D. H. Cusworth et al., 2021; Ware et al., 2019).

Acknowledgments

Funding. G.M. acknowledges support from the NIST NRC fellowship program. Funding for this work was provided by NIST and the NASA Instrument Incubator Program.

Acknowledgments. The authors thank NOAA (Eric Williams) for access to the Platteville Atmospheric Observatory, Andy Neuman for loan of the ammonia CRDS instrument, and Dan Zimmerle, Dan Bon, and Chad DeVolin for CAFO information. Finally, the authors thank Newton Nguyen, Israel Lopez-Coto, Zachary Grey, and Subhomoy Ghosh for helpful suggestions and comments.

Disclosures. The authors declare no conflicts of interest. Official contribution of the National Institute of Standards and Technology; not subject to copyright in the United States.

Data availability. Data underlying the results presented in this paper are available from the authors upon reasonable request.

Open Research

AGU requires an Availability Statement for the underlying data needed to understand, evaluate, and build upon the reported research at the time of peer review and publication. Additionally, authors should include an Availability Statement for the software that has a significant impact on the research. Details and templates are in the Availability Statement section of the Data & Software for Authors Guidance. For physical samples, use the IGSN persistent identifier, see the International Geo Sample Numbers section.

References

Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., et al. (2016). A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review*, *144*(4), 1669–1694. https://doi.org/10.1175/MWR-D-15-0242.1

Cardoso-Saldaña, F. J., Kimura, Y., Stanley, P., McGaughey, G., Herndon, S. C., Roscioli, J. R., et al. (2019). Use of Light Alkane Fingerprints in Attributing Emissions from Oil and Gas Production. *Environmental Science & Technology*, *53*(9), 5483–5492. https://doi.org/10.1021/acs.est.8b05828

Coddington, I., Newbury, N., & Swann, W. (2016). Dual-comb spectroscopy. *Optica*, *3*(4), 414–426. https://doi.org/10.1364/OPTICA.3.000414

Colorado Oil and Gas Conservation Commission. (2022, September 15). Oil and Gas Well Analytical Data. Retrieved from https://cogcc.state.co.us/documents/data/downloads/environmental/ProdWellDownLoad.html

Cusworth, D., Thorpe, A., Ayasse, A., Stepp, D., Heckler, J., Asner, G., et al. (2022). *Strong methane point sources contribute a disproportionate fraction of total emissions across multiple basins in the U.S.* (preprint). Earth Sciences. https://doi.org/10.31223/X53P88

Cusworth, D. H., Duren, R. M., Yadav, V., Thorpe, A. K., Verhulst, K., Sander, S., et al. (2020). Synthesis of Methane Observations Across Scales: Strategies for Deploying a Multitiered Observing Network. *Geophysical Research Letters*, *47*(7). https://doi.org/10.1029/2020GL087869

Cusworth, D. H., Bloom, A. A., Ma, S., Miller, C. E., Bowman, K., Yin, Y., et al. (2021). A Bayesian framework for deriving sector-based methane emissions from top-down fluxes. *Communications Earth & Environment*, *2*(1), 242. https://doi.org/10.1038/s43247-021-00312-6

Eilerman, S. J., Peischl, J., Neuman, J. A., Ryerson, T. B., Aikin, K. C., Holloway, M. W., et al. (2016). Characterization of Ammonia, Methane, and Nitrous Oxide Emissions from Concentrated Animal Feeding Operations in Northeastern Colorado. *Environmental Science & Technology*, *50*(20), 10885–10893. https://doi.org/10.1021/acs.est.6b02851

Fasoli, B., Lin, J. C., Bowling, D. R., Mitchell, L., & Mendoza, D. (2018). Simulating atmospheric tracer concentrations for spatially distributed receptors: updates to the Stochastic Time-Inverted Lagrangian Transport model's R interface (STILT-R version 2). *Geoscientific Model Development*, *11*(7), 2813–2824. https://doi.org/10.5194/gmd-11-2813-2018

Giorgetta, F. R., Peischl, J., Herman, D. I., Ycas, G., Coddington, I., Newbury, N. R., & Cossel, K. C. (2021). Open-Path Dual-Comb Spectroscopy for Multispecies Trace Gas Detection

in the 4.5–5 µm Spectral Region. *Laser & Photonics Reviews*, *15*(9), 2000583. https://doi.org/10.1002/lpor.202000583

Golston, L. M., Pan, D., Sun, K., Tao, L., Zondlo, M. A., Eilerman, S. J., et al. (2020). Variability of Ammonia and Methane Emissions from Animal Feeding Operations in Northeastern Colorado. *Environmental Science & Technology*, *54*(18), 11015–11024. https://doi.org/10.1021/acs.est.0c00301

Gordon, I. E., Rothman, L. S., Hargreaves, R. J., Hashemi, R., Karlovets, E. V., Skinner, F. M., et al. (2022). The HITRAN2020 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *277*, 107949. https://doi.org/10.1016/j.jqsrt.2021.107949

Herman, D., Mead, G., Giorgetta, F. R., Baumann, E., Malarich, N., Washburn, B. R., et al. (2023). Open-path measurement of stable water isotopologues using mid-infrared dual-comb spectroscopy.

Herman, D. I., Weerasekara, C., Hutcherson, L. C., Giorgetta, F. R., Cossel, K. C., Waxman, E. M., et al. (2021). Precise multispecies agricultural gas flux determined using broadband open-path dual-comb spectroscopy. *Science Advances*, *7*(14), eabe9765. https://doi.org/10.1126/sciadv.abe9765

IPCC. (1996). Emission Factor Database. Retrieved from https://www.ipcc-nggip.iges.or.jp/EFDB/find_ef.php

Kille, N., Chiu, R., Frey, M., Hase, F., Sha, M. K., Blumenstock, T., et al. (2019). Separation of Methane Emissions From Agricultural and Natural Gas Sources in the Colorado Front Range. *Geophysical Research Letters*, *46*(7), 3990–3998. https://doi.org/10.1029/2019GL082132

Kunik, L., Mallia, D. V., Gurney, K. R., Mendoza, D. L., Oda, T., & Lin, J. C. (2019). Bayesian inverse estimation of urban CO2 emissions: Results from a synthetic data simulation over Salt Lake City, UT. *Elementa: Science of the Anthropocene*, *7*, 36. https://doi.org/10.1525/elementa.375

Lan, X., Tans, P., Sweeney, C., Andrews, A., Dlugokencky, E., Schwietzke, S., et al. (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. https://doi.org/10.1029/2018GL081731

Lin, J. C. (2003). A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *Journal of Geophysical Research*, *108*(D16), ACH 2-1-ACH 2-17. https://doi.org/10.1029/2002JD003161

Lu, X., Jacob, D. J., Zhang, Y., Shen, L., Sulprizio, M. P., Maasakkers, J. D., et al. (2023). Observation-derived 2010-2019 trends in methane emissions and intensities from US oil and gas fields tied to activity metrics. *Proceedings of the National Academy of Sciences*, *120*(17), e2217900120. https://doi.org/10.1073/pnas.2217900120

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., et al. (2016). Gridded National Inventory of U.S. Methane Emissions. *Environmental Science & Technology*, *50*(23), 13123–13133. https://doi.org/10.1021/acs.est.6b02878

McKain, K., Down, A., Raciti, S. M., Budney, J., Hutyra, L. R., Floerchinger, C., et al. (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. https://doi.org/10.1073/pnas.1416261112

Michalak, A. M. (2004). A geostatistical approach to surface flux estimation of atmospheric trace gases. *Journal of Geophysical Research*, *109*(D14), D14109. https://doi.org/10.1029/2003JD004422

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., & Tans, P. P. (2005). Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions. *Journal of Geophysical Research*, *110*(D24), D24107. https://doi.org/10.1029/2005JD005970

Miller, D. J., Sun, K., Tao, L., Pan, D., Zondlo, M. A., Nowak, J. B., et al. (2015). Ammonia and methane dairy emission plumes in the San Joaquin Valley of California from individual feedlot to regional scales. *Journal of Geophysical Research: Atmospheres*, *120*(18), 9718–9738. https://doi.org/10.1002/2015JD023241

National Agricultural Statistics Service. (n.d.). Colorado Annual Agricultural Statistics Bulletin. United States Department of Agriculture. Retrieved from https://www.nass.usda.gov/Statistics_by_State/Colorado/Publications/Annual_Statistical_Bulletin/index.php

Peischl, J., Ryerson, T. B., Brioude, J., Aikin, K. C., Andrews, A. E., Atlas, E., et al. (2013). Quantifying sources of methane using light alkanes in the Los Angeles basin, California: SOURCES OF METHANE IN L.A. *Journal of Geophysical Research: Atmospheres*, *118*(10), 4974–4990. https://doi.org/10.1002/jgrd.50413

Peischl, J., Eilerman, S. J., Neuman, J. A., Aikin, K. C., de Gouw, J., Gilman, J. B., et al. (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*. https://doi.org/10.1029/2018JD028622

Pétron, G., Frost, G., Miller, B. R., Hirsch, A. I., Montzka, S. A., Karion, A., et al. (2012). Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study. *Journal of Geophysical Research: Atmospheres*, *117*(D4), D04304. https://doi.org/10.1029/2011JD016360

Pétron, G., Karion, A., Sweeney, C., Miller, B. R., Montzka, S. A., Frost, G. J., et al. (2014). A new look at methane and nonmethane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin. *Journal of Geophysical Research: Atmospheres*, *119*(11), 6836–6852. https://doi.org/10.1002/2013JD021272

Pollack, I. B., McCabe, M. E., Caulton, D. R., & Fischer, E. V. (2022). Enhancements in Ammonia and Methane from Agricultural Sources in the Northeastern Colorado Front Range Using Observations from a Small Research Aircraft. *Environmental Science & Technology*, *56*(4), 2236–2247. https://doi.org/10.1021/acs.est.1c07382

Robertson, A. M., Edie, R., Snare, D., Soltis, J., Field, R. A., Burkhart, M. D., et al. (2017). Variation in Methane Emission Rates from Well Pads in Four Oil and Gas Basins with Contrasting Production Volumes and Compositions. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.7b00571

Roy, J., Deschênes, J.-D., Potvin, S., & Genest, J. (2012). Continuous real-time correction and averaging for frequency comb interferometry. *Optics Express*, *20*(20), 21932–21939. https://doi.org/10.1364/OE.20.021932

Sargent, M., Barrera, Y., Nehrkorn, T., Hutyra, L. R., Gately, C. K., Jones, T., et al. (2018). Anthropogenic and biogenic CO_2 fluxes in the Boston urban region. *Proceedings of the National Academy of Sciences*, *115*(29), 7491–7496. https://doi.org/10.1073/pnas.1803715115

Skinner, C. C., Miller, R. F., Kinney, S. A., Gianoutsos, N. J., Gunther, G., & Shorten, C. M. (2022). Aggregated Oil and Natural Gas Drilling and Production History of the United States [Data set]. U.S. Geological Survey. https://doi.org/10.5066/P9UIR5HE

Ware, J., Kort, E. A., Duren, R., Mueller, K. L., Verhulst, K., & Yadav, V. (2019). Detecting Urban Emissions Changes and Events With a Near-Real-Time-Capable Inversion System. *Journal of Geophysical Research: Atmospheres*, *124*(9), 5117–5130.

West, M., & Harrison, J. (1997). *Bayesian forecasting and dynamic models* (2nd ed). New York: Springer.

Yacovitch, T. I., Herndon, S. C., Roscioli, J. R., Floerchinger, C., McGovern, R. M., Agnese, M., et al. (2014). Demonstration of an Ethane Spectrometer for Methane Source Identification. *Environmental Science & Technology*, *48*(14), 8028–8034. https://doi.org/10.1021/es501475q

Yacovitch, T. I., Herndon, S. C., Pétron, G., Kofler, J., Lyon, D., Zahniser, M. S., & Kolb, C. E. (2015). Mobile Laboratory Observations of Methane Emissions in the Barnett Shale Region. *Environmental Science & Technology*, *49*(13), 7889–7895. https://doi.org/10.1021/es506352j

Ycas, G., Giorgetta, F. R., Baumann, E., Coddington, I., Herman, D., Diddams, S. A., & Newbury, N. R. (2018). High-coherence mid-infrared dual-comb spectroscopy spanning 2.6 to 5.2 μm. *Nature Photonics*, *12*(4), 202–208. https://doi.org/10.1038/s41566-018-0114-7

Ycas, G., Giorgetta, F. R., Cossel, K. C., Waxman, E. M., Baumann, E., Newbury, N. R., & Coddington, I. (2019). Mid-infrared dual-comb spectroscopy of volatile organic

compounds across long open-air paths. *Optica*, *6*(2), 165–168. https://doi.org/10.1364/OPTICA.6.000165

Ycas, G., Giorgetta, F. R., Friedlein, J. T., Herman, D., Cossel, K. C., Baumann, E., et al. (2020). Compact mid-infrared dual-comb spectrometer for outdoor spectroscopy. *Optics Express*, *28*(10), 14740–14752. https://doi.org/10.1364/OE.385860

References From the Supporting Information

Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., et al. (2016). A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review*, *144*(4), 1669–1694. https://doi.org/10.1175/MWR-D-15-0242.1

Boulder County Public Health, Boulder A.I.R., LLC, & Colorado Department of Public Health and Environment. (n.d.). Boulder Reservoir Atmospheric Chemical Conditions. Retrieved from https://www.bouldair.com/boulder.htm#header

City of Longmont, & Boulder A.I.R., LLC. (n.d.). Longmont Union Reservoir Atmospheric Chemical Conditions. Retrieved from https://www.bouldair.com/longmont.htm

Cusworth, D. H., Duren, R. M., Yadav, V., Thorpe, A. K., Verhulst, K., Sander, S., et al. (2020). Synthesis of Methane Observations Across Scales: Strategies for Deploying a Multitiered Observing Network. *Geophysical Research Letters*, *47*(7). https://doi.org/10.1029/2020GL087869

Gordon, I. E., Rothman, L. S., Hargreaves, R. J., Hashemi, R., Karlovets, E. V., Skinner, F. M., et al. (2021). The HITRAN2020 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 107949. https://doi.org/10.1016/j.jqsrt.2021.107949

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., et al. (2016). Gridded National Inventory of U.S. Methane Emissions. *Environmental Science & Technology*, *50*(23), 13123–13133. https://doi.org/10.1021/acs.est.6b02878

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R., Peters, W., & Tans, P. P. (2005). Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace

gas surface flux inversions. *Journal of Geophysical Research*, *110*(D24), D24107. https://doi.org/10.1029/2005JD005970

Newville, Matthew, Stensitzki, Till, Allen, Daniel B., & Ingargiola, Antonino. (2014, September 21). LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python. Zenodo. https://doi.org/10.5281/zenodo.11813

Geophysical Research Letters

Supporting Information for

Apportionment and Inventory Optimization of Agriculture and Energy Sector Methane Emissions using Multi-month Trace Gas Measurements in Northern Colorado

Griffin J. Mead,[1] Daniel I. Herman,[1,2] Fabrizio R. Giorgetta,[1,2] Nathan A. Malarich,[1] Esther Baumann,[1,2] Brian R. Washburn,[1] Nathan R. Newbury,[1] Ian Coddington,[1] Kevin C. Cossel[1,*]

[1] National Institute of Standards and Technology, Spectrum Technology and Research Division, Boulder, CO 80305

[2] University of Colorado, Boulder, Department of Physics, Boulder, CO 80309

Contents of this file

Text S0: Data sources and analytical tools

This work relies upon observational time series data, meteorological model data, and surface flux data. Time series data were produced by calculating the dry mixing ratios of methane and ethane for each 2-minute averaged spectra recorded at PAO. Molecular linelists were provided by HITRAN2020 (Gordon et al., 2021); spectral fitting was performed using the LMFIT Python package (Newville, Matthew et al., 2014). High Resolution Rapid Refresh (HRRR) meteorological model (Benjamin et al., 2016) data was retrieved from the National Oceanic and Atmospheric Administration's Air Resources Laboratory FTP server (ftp://ftp.arl.noaa.gov/nams). Transport modeling using the HRRR data was performed with the STILT-R package (https://github.com/uataq/stilt/). Surface flux data was published in Maasakkers et al, 2016. Much of the geospatial processing and plotting was performed in QGIS. The NOAA meteorological station data can be accessed here: https://psl.noaa.gov/data/obs/sites/view_site_details.php?siteID=pvl.

Text S1: Experimental setup at the Platteville Atmospheric Observatory

An overview of instrument locations at Platteville Atmospheric Observatory is shown in Figure S1. Mid-infrared (MIR) light from the dual-comb spectrometer (DCS) was launched from a gimbal-mounted 10-cm aperture transmit/receive telescope to a 12.5-cm diameter gold hollow corner-cube retroreflector mounted on a telephone pole. The ammonia cavity ring-down spectrometer (CRDS) analyzer was housed in a nearby trailer which was offset ~50 meters perpendicular to the MIR beam path. A ~4-meter ¼" PTFE tubing clad in copper tubing and wrapped in heater tape and foil tape extended from the CRDS analyzer to sample air above the trailer at ~4 meters AGL. The heater tape was regulated to maintain a temperature of 47 ˚C. Data were collected at 1 Hz and interpolated onto the 2-minute DCS time base.

Zero-air measurements were performed on the CRDS instrument every 1-2 weeks at PAO by overflowing the inlet with zero grade dry air. Over all measurements (N=10), the mean reported

ammonia concentration was 1.8 ppb +/- 1.0 ppb. This concentration offset was comparable to manufacturer specifications and subtracted from the NH_3 data. Equivalent zero-air reference measurements were not possible with the open-path DCS instrument; as a result, the reported accuracy of the methane and ethane data are estimated. For ethane, a minimum detected concentration of 4 ppb +/- 1 ppb was observed on the night of 31 October 2021. Since this measurement includes contributions from regional oil and natural gas emissions, ethane's accuracy is estimated at <=4 ppb with a precision of 1 ppb. Methane precision was determined to be +/- 1 ppb over the same 31 October 2021 nighttime period.

A key assumption of this work is that the mixing ratios of methane, ethane, and ammonia are uniform across the DCS beam path and the CRDS inlet. Any variation which does occur, for example due to venting from a nearby tank battery, is further assumed to be detected by both systems with only a small offset in time as the plume travels from one system to the other. Depending upon the wind direction, a local plume would need to be transported 50 m-300 m between the CRDS and DCS beam path. With a typical 2 m/s wind speed, this would result in an arrival time offset of 25-150 s, which is less than or equal to the 2-minute time base of a single data point. Therefore, the locations of the two systems are not expected to introduce any substantial error into the analysis. Data from a CH_4 CRDS instrument collocated with the NH_3 CRDS instrument closely matched the DCS observations ($CH_4^{CRDS} = 0.994 \times CH_4^{DCS}$, R^2 = 0.998).

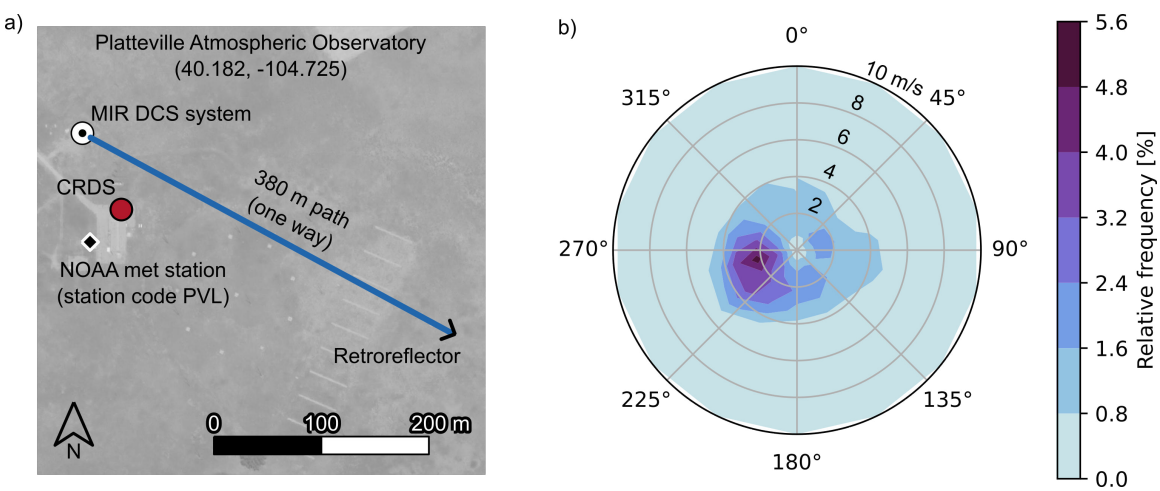


Figure S1. a) Map of instrument locations at PAO. The MIR DCS beam path traverses 380 m from a transmit/receive telescope to a retroreflector. The cavity ringdown spectrometer (CRDS) is co-located to retrieve ammonia. Satellite background image ©2023 Google and Maxar Technologies. b) Polar histogram of wind direction and speed recorded by the NOAA PVL met station.

Text S2: Dual-comb spectroscopy

A basic overview of the dual-comb spectroscopy method is shown in Figure S2a. Two mid-infrared frequency combs travel across an open-air beam path where molecular absorption occurs. Collection and digitization of the two combs on a photodiode down-converts information at optical frequencies to radio frequencies. Using known comb parameters, the optical spectrum can be reproduced from the radio signal (Figure S2b). Methane, ethane, and water concentrations were fit in the 2918-2968 cm^{-1} spectral region for all data in this analysis. The spectral baseline was modelled with a fifth-order polynomial.

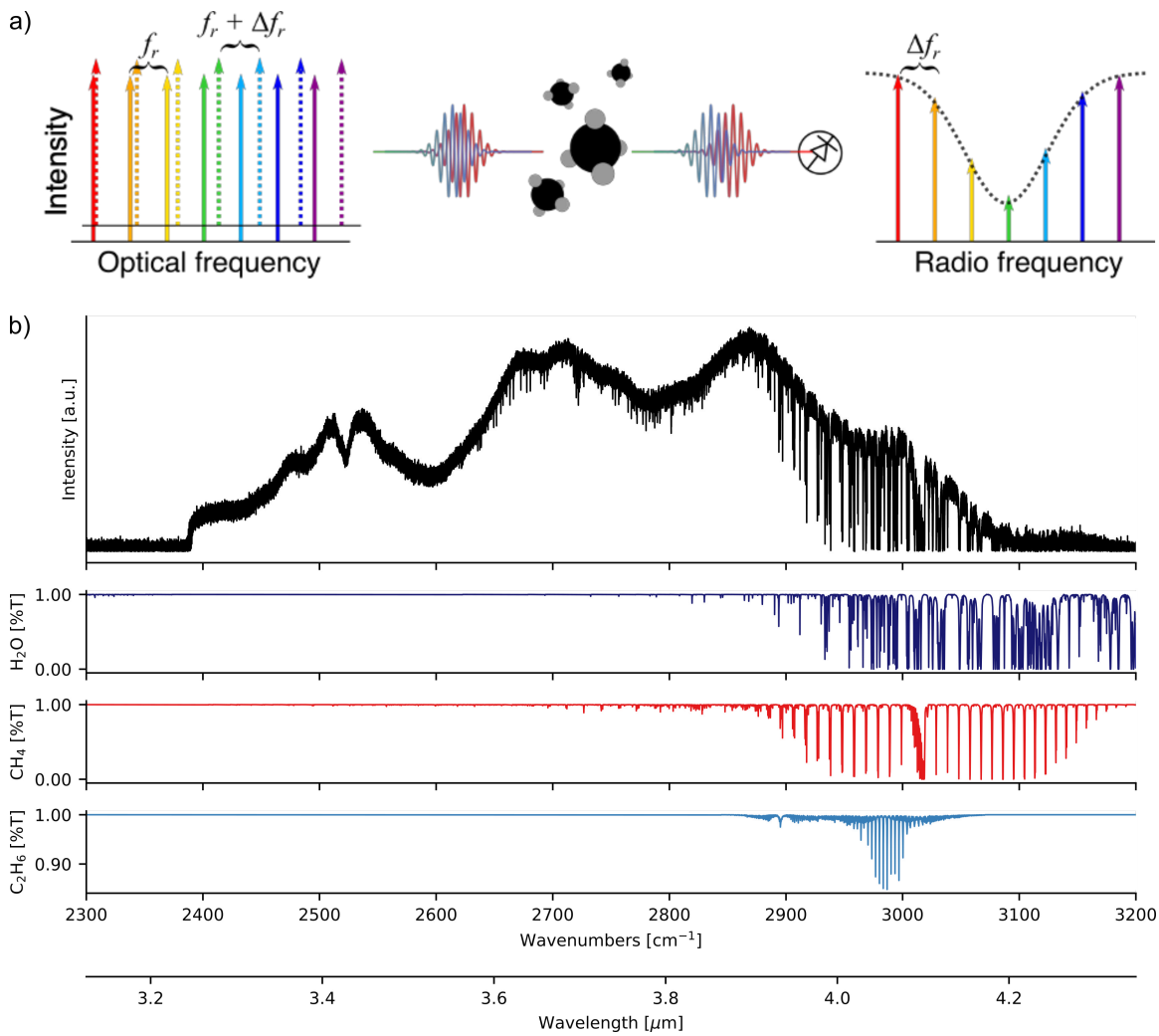


Figure S2. Mid-infrared dual-comb spectroscopy was used to measure methane, ethane, and water at PAO. a) A model of the dual-comb spectroscopy technique. Two frequency combs with repetition rates f_r and $f_r + \Delta f$ propagate along an open-air path. Molecules in the path absorb portions of the frequency comb light. The two combs then interfere on the detector, down-converting the optical comb and molecular absorption spectra to radio frequencies. b) A representative 2-minute averaged MIR spectrum for the PAO measurements (black trace). The overall shape is dominated by the comb spectra themselves, but the smaller sharp lines indicate transitions from multiple gas species. The lower graphs show the absorption spectra of water, methane, and ethane, scaled to the concentrations retrieved from the MIR spectrum on top.

Text S3: Dynamic linear model background estimation

A dynamic linear model tracer gas analysis is used to separate the observed methane time series into contributions from energy and agriculture emissions, and a background term:

$$[CH_4] = \beta_0 + \beta_1[C_2H_6] + \beta_2\,[NH_3] + \epsilon$$

In the Front Range Urban Corridor, the β_0 term is the 'background' methane mixing ratio which would be measured in the absence of energy and agricultural methane emissions. We compared data from PAO with methane and ethane time series data collected at two regional air quality monitoring sites shown in Figure S3a: Boulder Reservoir (BRZ) (Boulder County Public Health et

al., n.d.) and Longmont Union Reservoir (LUR) (City of Longmont & Boulder A.I.R., LLC, n.d.). (Ammonia data is not recorded at either BRZ or LUR.) For each site, background methane mixing ratios were defined as those time periods when the observed ethane mixing ratio was less than or equal to the 5th percentile of the ethane time series. Background methane mixing ratios calculated using this approach are shown in the downward-going kernel density estimates in Figure S3b. Background methane mixing ratios decreased from PAO to LUR to BRZ as the density of oil and natural gas infrastructure around each site decreased. The independently determined β_0 regression results (Figure S3b, upward-going kernel density estimate) overlapped with these three background estimates, indicating that the dynamic linear model tracer gas analysis provides a reasonable estimate of background conditions at PAO.

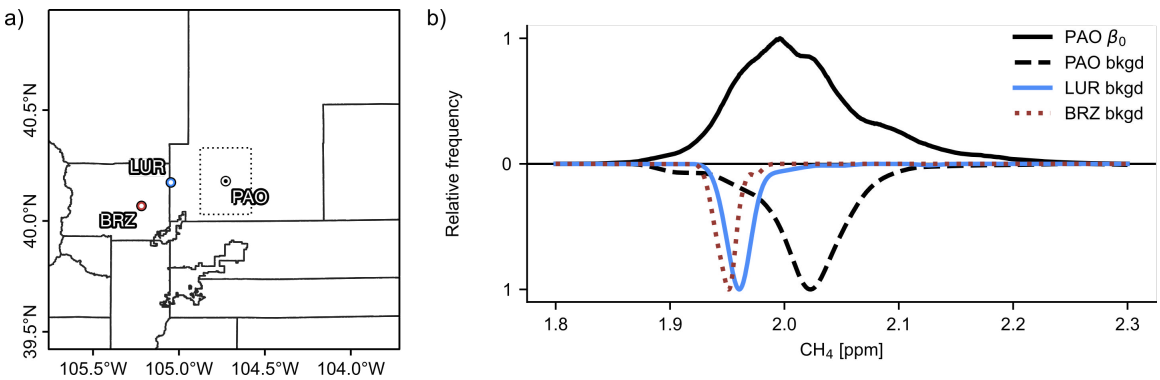


Figure S3. a) Additional methane and ethane mixing ratios measured at the Boulder Reservoir (BRZ) and Longmont Union Reservoir (LUR) air quality monitoring sites was used to independently estimate regional background methane mixing ratios. b) Ethane-only estimates from BRZ, LUR, and PAO ('PAO/LUR/BRZ bkgd') are consistent with the DLM analysis of the PAO data ('PAO β_0').

Text S4: Bayesian inversion analysis details

Using the observed sector-apportioned methane time series, the 8-week hourly resolution primary back trajectory simulations, and emission inventories, we generated optimized posterior inventories by calculating the maximum a posteriori (MAP) estimate $x^{Posterior}$ and posterior error covariance matrix $\widehat{\boldsymbol{S}}$ (Cusworth et al., 2020),

$$x^{Posterior} = x^{Prior} + \boldsymbol{S}\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T + \boldsymbol{R})^{-1}(y^{Obs} - \boldsymbol{H}x^{Prior})$$

$$\widehat{\boldsymbol{S}} = (\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H} + \boldsymbol{S}^{-1})^{-1}$$

The solutions $x^{Posterior}$ and $\widehat{\boldsymbol{S}}$ require estimates for the observational error covariance matrix $\boldsymbol{R}$ and the prior error covariance matrix $\boldsymbol{S}$, which are determined by finding solutions $\theta = \{\sigma_R, \sigma_S\}$ that minimize the cost function (Cusworth et al., 2020; Michalak et al., 2005):

$$L_\theta = |\boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T + \boldsymbol{R}| + (\boldsymbol{H}x^{Prior} - y^{Obs})^T(\boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T + \boldsymbol{R})^{-1}(\boldsymbol{H}x^{Prior} - y^{Obs})$$

The two error covariance matrices are then constructed as identity matrices multiplied by the respective scalar variances:

$$\boldsymbol{S} = \sigma_S\,\boldsymbol{I}, \qquad \boldsymbol{R} = \sigma_R\boldsymbol{I}$$

Different regions contribute to the posterior state vector to varying degrees, which can be determined from the averaging kernel matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = \boldsymbol{I} - \widehat{\boldsymbol{S}}\boldsymbol{S}^{-1}$$

The ideal averaging kernel matrix is $\boldsymbol{I}$; non-ideal deviations due to measurement design, model errors, etc, produce off-diagonal matrix elements which undesirably spread information across multiple grid cells.

Posterior estimates will only be optimized in regions where the averaging kernel is appreciably greater than 0. The diagonal elements of this study's averaging kernel matrix is shown in Figure S4, which indicates that our observations will only meaningfully optimize the emissions inventory in an area of approximately 850 km^2 centered around PAO. This sensitivity region is shown as a rectangular outline centered on PAO in the main text's Figures 3 and 4.

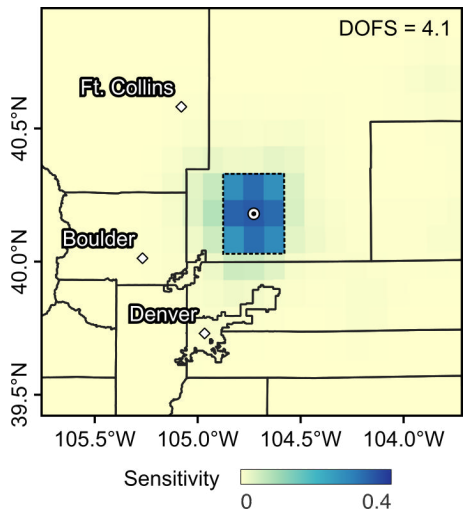


Figure S4. The diagonal elements of the averaging kernel sensitivity matrix $\boldsymbol{A}$ for observations at PAO. Posterior emissions are most strongly constrained by observations within a 850 km^2 (0.3° × 0.3°) area centered around PAO (outlined in black dashed rectangle). The degrees of freedom of signal (DOFS) provided by the observations is 4.1.

Text S5: Redistributing agriculture methane emissions to CAFO locations

EPA agriculture emissions were redistributed to known CAFO locations using QGIS tools. First, the total emissions (4A+4B) for each county were calculated using the Zonal Statistics tool. This step reverses the probabilistic distribution of emissions throughout each county which was used to produce the EPA inventory. After exporting these county-level data to a comma separated variable format, the total emissions for each county were distributed to every CAFO within that county proportionate to the fractional animal equivalent units of livestock permitted at each CAFO. This results in the same total agriculture emissions at the county level but spatially re-distributed to known CAFO locations according to relative CAFO size and primary livestock type. CAFO distributed emissions data were then re-imported to QGIS; after generating a 0.1° grid with the same spatial extent as the EPA inventory, the total emissions per grid element was calculated using the Points in Polygon query. Finally, the emissions were converted to a raster and exported as a netCDF file, producing an updated AG emissions inventory with the same extent and resolution as the original EPA inventory. (This redistributed inventory only defines emissions within the state of Colorado.)

Text S6: STILT-R simulation parameters

A range of atmospheric dispersion simulations were run to determine the sensitivity of the influence footprint to input parameters. Footprint variability was estimated by performing simulations over a range of spatial resolutions (0.03° and 0.1°), back-trajectory durations (24 and 48 hours), numbers of particles (50, 100, and 200 particles), and with and without hyper-near field effects for a two-week period in November. Variations in $y^{Prior}_{Energy/Agri}$ due to simulation inputs was found to be quite small compared to the estimated spatial inventory uncertainties specified in Maasakkers (Maasakkers et al., 2016).

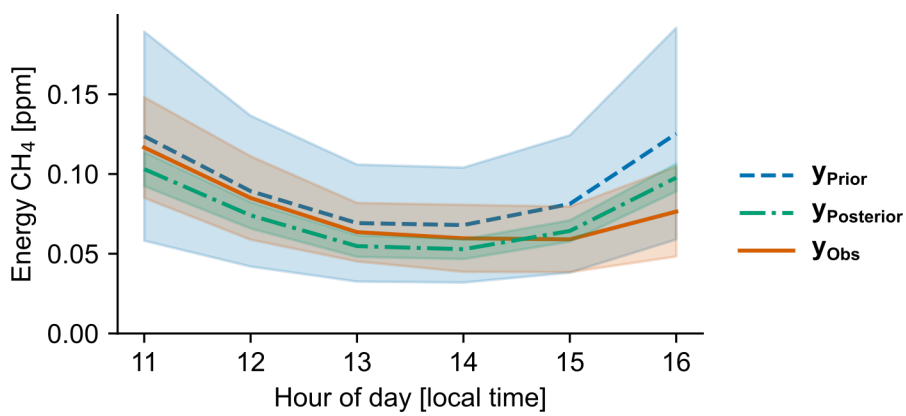
Text S7: Uncertainty analysis of Bayesian inversion results

Uncertainties in $y^{Obs}_{Energy/Agri}$ were estimated with a bootstrap method by randomly sampling 20% of the sector apportioned methane time series 500 times, calculating the hourly mean for each sample, and calculating the variance of the spread of the hourly means. The uncertainty of y^{Prior} was based on the sector-dependent uncertainties described in Maasakkers. Uncertainty in $y^{Posterior}$ was calculated using the posterior error covariance matrix (see Text S4).

In the following tables, the mean hourly mixing ratios for y^{Obs}, y^{Prior}, and $y^{Posterior}$ are listed along with uncertainties. Mean absolute differences were calculated as $|y^{Obs} - y^{Prior/Posterior}|$, and uncertainties were propagated as $\sqrt{\delta y^2_{Obs} + \delta y^2_{Prior/Posterior}}$.

S7.1 Energy sector

S7.1.1 Mean diurnal mixing ratios



S7.1.2 Prior vs. Observations

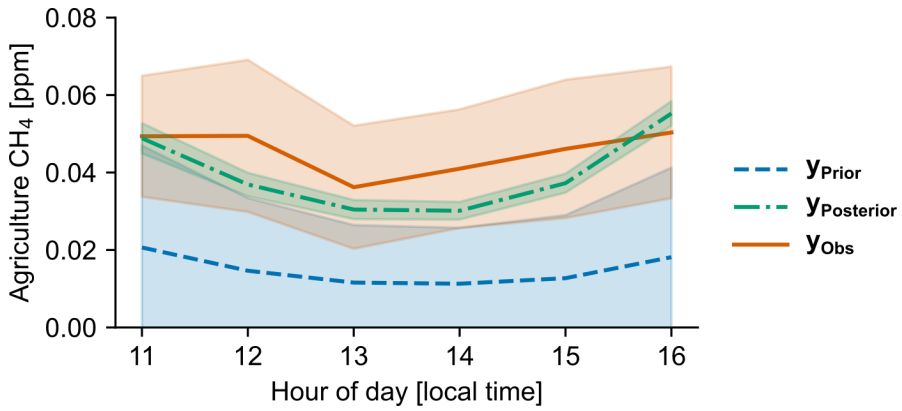
Hour	y^{Obs} [ppm]	δy^{Obs} [ppm]	y^{Prior} [ppm]	δy^{Prior} [ppm]	$\lvert y^{Obs} - y^{Prior}\rvert$ [ppm]	$\sqrt{\delta y^2_{Obs} + \delta y^2_{Prior}}$ [ppm]
11	0.11947	0.03048	0.12370	0.06556	0.00423	0.07230
12	0.08577	0.02663	0.08925	0.04730	0.00348	0.05429
13	0.06337	0.01786	0.06923	0.03669	0.00586	0.04080
14	0.06068	0.02004	0.06799	0.03604	0.00731	0.04123
15	0.06053	0.02204	0.08123	0.04305	0.0207	0.04837
16	0.07475	0.03005	0.12530	0.06641	0.05055	0.07289
				CH_4 [ppm]	0.01536	0.05498
				CH_4 [ppb]	15.36	54.98

S7.1.3 Posterior vs. Observations

Hour	y^{Obs} [ppm]	δy^{Obs} [ppm]	$y^{Posterior}$ [ppm]	$\delta y^{Posterior}$ [ppm]	$\lvert y^{Obs} - y^{Posterior} \rvert$ [ppm]	$\sqrt{\delta y_{Obs}^2 + \delta y_{Posterior}^2}$ [ppm]
11	0.11947	0.03048	0.10307	0.01067	0.0164	0.03229
12	0.08577	0.02663	0.07407	0.00827	0.0117	0.02788
13	0.06337	0.01786	0.05478	0.00674	0.00859	0.01909
14	0.06068	0.02004	0.05285	0.00622	0.00783	0.02098
15	0.06053	0.02204	0.06429	0.00669	0.00376	0.02304
16	0.07475	0.03005	0.09772	0.00867	0.02297	0.03127
				CH_4 [ppm]	0.01187	0.02576
				CH_4 [ppb]	11.87	25.76

S7.2 Agriculture sector

S7.2.1 Mean diurnal mixing ratios



S7.2.2 Prior vs. Observations

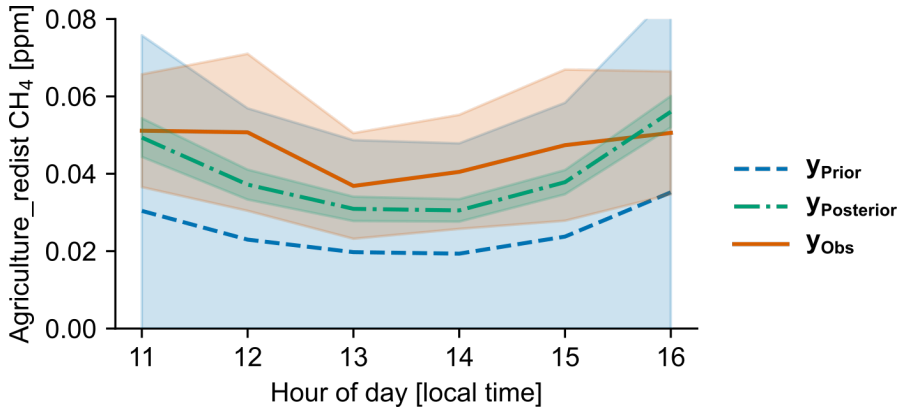
Hour	y^{Obs} [ppm]	δy^{Obs} [ppm]	y^{Prior} [ppm]	δy^{Prior} [ppm]	$\lvert y^{Obs} - y^{Prior} \rvert$ [ppm]	$\sqrt{\delta y_{Obs}^2 + \delta y_{Prior}^2}$ [ppm]
11	0.05178	0.01524	0.02064	0.02634	0.03114	0.03044
12	0.05236	0.02016	0.01465	0.01877	0.03771	0.02755
13	0.03796	0.01452	0.01159	0.01485	0.02637	0.02077
14	0.04135	0.01526	0.01129	0.01442	0.03006	0.02100
15	0.04709	0.01981	0.01272	0.01627	0.03437	0.02564
16	0.05177	0.01742	0.01815	0.02311	0.03362	0.02894
				CH_4 [ppm]	0.03221	0.02572
				CH_4 [ppb]	32.21	25.72

S7.2.3 Posterior vs. Observations

Hour	y^{Obs} [ppm]	δy^{Obs} [ppm]	$y^{Posterior}$ [ppm]	$\delta y^{Posterior}$ [ppm]	$\lvert y^{Obs} - y^{Posterior} \rvert$ [ppm]	$\sqrt{\delta y_{Obs}^2 + \delta y_{Posterior}^2}$ [ppm]
11	0.05178	0.01524	0.04885	0.00391	0.00293	0.01574
12	0.05236	0.02016	0.03692	0.00303	0.01544	0.02039
13	0.03796	0.01452	0.03044	0.00247	0.00752	0.01473
14	0.04135	0.01526	0.03013	0.00228	0.01122	0.01543
15	0.04709	0.01981	0.03725	0.00245	0.00984	0.01996
16	0.05177	0.01742	0.05524	0.00318	0.00347	0.01771
				CH_4 [ppm]	0.00840	0.01733
				CH_4 [ppb]	8.40	17.33

S7.3 Redistributed agriculture sector

S7.3.1 Mean diurnal mixing ratios



S7.3.2 Prior vs. Observations

Hour	y^{Obs} [ppm]	δy^{Obs} [ppm]	y^{Prior} [ppm]	δy^{Prior} [ppm]	$\lvert y^{Obs} - y^{Prior} \rvert$ [ppm]	$\sqrt{\delta y_{Obs}^2 + \delta y_{Prior}^2}$ [ppm]
11	0.05108	0.01491	0.03041	0.04536	0.02067	0.04774
12	0.05238	0.02022	0.02298	0.03393	0.02940	0.03950
13	0.03744	0.01459	0.01974	0.02895	0.01770	0.03242
14	0.04146	0.01601	0.01936	0.02851	0.02210	0.03270
15	0.04490	0.01940	0.02376	0.03458	0.02114	0.03965
16	0.04988	0.01664	0.03522	0.05132	0.01466	0.05395
				CH_4 [ppm]	0.02095	0.04099
				CH_4 [ppb]	20.95	40.99

S7.3.3 Posterior vs. Observations

Hour	y^{Obs} [ppm]	δy^{Obs} [ppm]	$y^{Posterior}$ [ppm]	$\delta y^{Posterior}$ [ppm]	$\lvert y^{Obs} - y^{Posterior} \rvert$ [ppm]	$\sqrt{\delta y_{Obs}^2 + \delta y_{Posterior}^2}$ [ppm]
11	0.05108	0.01491	0.04936	0.00500	0.00172	0.01573
12	0.05238	0.02022	0.03720	0.00387	0.01518	0.02059
13	0.03744	0.01459	0.03093	0.00316	0.00651	0.01492
14	0.04146	0.01601	0.03053	0.00291	0.01093	0.01627
15	0.04490	0.01940	0.03783	0.00313	0.00707	0.01965
16	0.04988	0.01664	0.05604	0.00405	0.00616	0.01713
				CH_4 [ppm]	0.00793	0.01738
				CH_4 [ppb]	7.93	17.38