Response to Referee # 1 comments

General comments by Referee # 1

The authors use positive matrix factorization (PMF; a factor analysis method) and a suite of trace gas measurements from the CALNEX campaign during May and June, 2010 to resolve source categories of CH4 and N2O affecting a measurement site near Bakersfield, CA, in the San Joaquin Valley. PMF, combined with the large suite of measurements at half-hourly time resolution, seems to be an appropriate tool for this problem. The main conclusions are that none of the observed enhancements in CH4 come from the oil and gas sector, despite significant activity nearby, and that none of the N2O comes from vehicle emissions, in sharp contrast to California Air Resources Board inventories. Both results are somewhat surprising, but the analysis, as presented, is convincing. I recommend this paper for publication after the authors address the comments below to the satisfaction of the editor.

1) The paper needs editing to reduce excess words and improve clarity of the discussion.

   Author's changes: We have substantially edited the manuscript to reduce words and improve clarity of the discussion as requested. Changes suggested in the final PhD dissertation chapter associated with this manuscript have also been included to improve the document.

2) Given the strong latitudinal and vertical gradients in CH4, comparison of measurements at Bakersfield and Mauna Loa Observatory are not appropriate. A suitable site at mid-latitudes would be more appropriate. (also related is reviewer's specific comment on p6085, l10-11)

   Author response:

   Along with comparison of our measurements to the observed backgrounds at Mauna Loa (~19.48 N latitude), we have also compared them to observations from the NOAA station at Trinidad Head (40.97 N latitude; only CH4 data exists at THD) in Northern California simultaneously. The referee is correct that the CH4 measurements at THD point to reasonably higher concentrations than at MLO, although the backgrounds at Bakersfield are even higher than that at THD. But there are two issues with using THD measurements for the comparison: firstly, the measured samples at THD are discrete flask samples collected very few days at a random time of day (not necessarily a daily low signal) while MLO measurements do contain information about daily minimum concentrations during the same measurement period as Bakersfield. Secondly, since we did not have N2O data at THD, we prefer comparing the backgrounds at Bakersfield (35.36 N latitude) to MLO rather than THD just to be consistent using one station (MLO) for comparison of both CH4 and N2O backgrounds.

Author's changes: We have added the CH4 background information from THD to Section 4.1 in the text for reference as well and also clarified why we use MLO data. I do not think the abstract can or should be altered to maintain brevity

3) Many experimental details, for example choice of standards, are glossed over. Was the water vapor correction checked experimentally? If so, how? If not, how can you trust it? (also related is reviewer's specific comment on p6085, l10-11)

Author's Response:

Los Gatos Research Inc. is the manufacturer of the instruments that were used in this experiment and also our partners in this campaign. LGR supplied us with analyzers that were calibrated using the primary WMO standard from the Global Monitoring Division (GMD) at the NOAA Earth System Research Lab. The **scuba tanks were secondary references and were calibrated before and after the experiments using the primary standards. The LGR co-author in this manuscript has responded that the calibration tests confirmed that there is no issue in short term stability of these species.**

The water vapor correction is a correction for the dilution effect of water. As water changes concentration, the effect is to dilute the target molecule. This is a mathematical formula applied to the wet values" {X}dry={Xwet}/(1-([H2O]/1e6)) where [h2o] is the measured water concentration in ppm. We are not aware of published results confirming the LGR vapor correction and did not separately perform the test ourselves. I know if has been confirmed by researchers at NOAA and EMPA, but I do not think the results were published. Hence we assume the dry mole concentrations of the tracers based on the LGR internal correction to be sufficiently accurate.

Author's changes: We have clarified these points in the main text in Section 3.1 (trace gas measurements and instrumentation).

4) Lack of benzene measurements seems odd, given the spectrum of VOCs reported.

Author's Response:

There were instrumental and gas column related challenges that prevented us from being able to capture benzene is a quantitative manner.

Author's changes: We have specifically added in Section 3.1 (trace gas measurements and instrumentation) as to why benzene was not an included tracer in the analysis. Also a reference is added to a different manuscript that addresses this topic.

5) Uncertainties are not used consistently nor treated clearly. It seems odd to state the fraction of total emissions of a gas from a particular source as a range, then give an uncertainty.

Author's Response:

I agree that the presentation of the averaged diurnal range with bootstrapping uncertainties is confusing and misleading as these two physical quantities are not related in the context of the how they are derived.

Uncertainties in mass fractions derived from PMF analysis are calculated and reported in rather unconventional ways in past literature and often not mentioned at all. A detailed description of how the uncertainty in the relative apportionment of a particular gas, say CH4, in each source type is documented in section S3 in the Supplement. The uncertainty of a tracer in each source factor is derived from the 1-sigma deviation of the averaged mass fraction of that tracer in that source from 100 bootstrapping runs. This is the quantity I report and call 'uncertainty' in the abstract and the main manuscript.

The percentage range for CH4 and N2O that I report for the major source types has nothing to do with the bootstrapping run. That range is derived from the relative apportionment of CH4 and N2O to the specific source from the 653 hourly samples collected during the experiment period in the PMF analysis. This range is reported to demonstrate the diurnal variation typically observed and also seen in Figure 7.

Author's changes: I have used clarity and a further explanation when referring to the range and its related uncertainty throughout the text as a direct response to the comments from both reviewers. I have edited the abstract such that the bootstrapping uncertainty is not mentioned when reporting the diurnal range. In the main text, I provide a deeper explanation of the uncertainty estimates associated with tracer contributions and how are they derived.

6) Information in the introduction should be updated to the most recent IPCC report and original literature should be cited where possible.

Author's changes: I have incorporated the suggestion and the requested change in the revised manuscript.

7) How is the footprint of the observations affected by differing night and day meteorology? Is it reasonable to lump measurements from both periods? Aren't most of the enhancements coming from nightime build-up of species in the shallow boundary layer when the site's footprint would be much smaller? (also related Specific Comments p6088, l22)

Author comments:

It is true that the footprints differ between day and night and also based on wind speed and wind direction. But that is the advantage of this statistical technique that even in a nighttime buildup due to shallow boundary layer, difference in wind speeds and wind directions over the experiment period cause significant variations in the amplitude and timing of peak enhancements in various trace gas species. Similarly during day time, while the absolute enhancements are tinier as compared to nighttime for primary gases, the timing differences in day time lows, and high correlation with other species behaving similarly are important features that contribute to the overall statistical analysis. PMF is able to detect the differences in covariances of several groups of tracers varying together to create multiple sets of solutions with varying Q/Qexp ratio and allows the user to determine the most likely and plausible combination. If we split the daytime and nighttime data and attempt to run PMF on these separate periods (which we have done experimentally to verify), the PMF analysis is not conclusive because most importantly, a distinct diurnal profile of the enhancements being apportioned is missing which limits quality of the analysis and renders source factors that are not quite distinguishable and interpretable. You also cut down dramatically on the total number of samples (by half) by treating daytime and nighttime data separately so the degrees of freedom in the analysis suffer and impact the apportionment.

In the overall sense, since multiple biological and fugitive sources of GHGs are randomly distributed all across and also in close proximity to the experiment site at Bakersfield, we believe that the varying size of footprints does not cause certain GHG sources to go completely out of range except perhaps the larger oilfields to the east. But there are several operational oil pump jacks within and around the main urban core as well, and those are definitely within the footprints typically observed at the experiment site irrespective of day or night and wind direction.

Author's changes: None. I hope the reviewer/editor is satisfied with the explanation mentioned above

8) Are comparisons appropriate of these results for the southern portion of the SJV with CARB inventories for the entire state?

Author response:

The comparisons are made to point to some discrepancies in the bottom-up and top-down measurements and to provide a frame of reference to a reader who may not be aware of California's GHG inventory. We elucidate that emissions in the inventory are a product of combination of activity data and emission factors. Specific comparison of emission factors of the evaporative and fugitive source factor (Table 2) and the vehicle emission factor (Table 3) to those derived from southern SJV and Bakersfield-specific studies are used first to establish the origin of these PMF factors. These emission factors are then compared to the state-specific factors. Why? Because, in theory, the measured emission factors for specific source categories should not be significantly different from the state inventory emission factors although activity data (fuel sale, # of vehicles etc.) does vary greatly from region to region which we recognize. In our analysis, we point out to the mismatch of some of these emission factors when comparing local source factor to CARB sources. We are using this argument to suggest that CARB emission factors for certain sources need further validation through top-down data and by that extension question the accuracy of the inventory.

Author's changes: To make this point clear to other readers as well, I have added a sentence in the Implications section in paragraph 4 - "It should be noted that there is a significant difference in the regional vehicle activity data in the Bakersfield region as compared to the more urban Los Angeles and San Francisco regions of California and this can also be an important reason for difference in varying N_2O apportionment seen in regional vs state-specific N_2O inventory distributions. However, the discord of theoretical emission factors (EMFAC) from the measured vehicle emission factors in this study also points to the likelihood of erroneous estimates in any inventory when the theoretical but inaccurate emission factors are combined with activity data to calculate emissions."

All editing comments and specific comments by Referee # 1 have been attended to and addressed in the revised manuscript.

Response to Referee # 2 comments

General comments by Referee # 2

This study adds useful information on source apportionment of CH4 and N2O emissions from an area with important air quality control problems. The research was conducted carefully, the manuscript is generally well-written, and the results are interesting. This work could reasonably be published with minor revisions. My only significant criticism is that despite commenting on agriculture and energy resource related emissions from the southern valley (e.g., page 6083) the authors compare the relative source strengths derived from this 6 week (May-June) study with annual-average state-wide CH4 and N2O emissions (that contain significant contributions from coastal urban areas. The authors might consider revising the abstract and discussion to be specific that their results likely differ from state-wide annual average emissions, or better yet, also attempt to compare with an inventory-based emission estimates specific to the summer-time central valley.

Author response:

Please read response to Referee # 1 on a similar line of questioning (answer # 8 in previous pages).

Author's changes: Additionally, I have taken Referee # 2's suggestion and made it amply clear in the abstract and discussion that we expect differences in the relative proportion of emissions from major sources in local and state emissions distributions and focus on the differences arising in the emission factors of specific sources that should remain consistent in all comparisons.

Specific comments:
page 6079, line 1: Would it be correct to state that given the overwhelming signal from livestock that the PMF analysis is consistent with the current CA inventory estimate that only _ 5% of regional CH4 emissions are derived from oil and gas operations ?

Author's changes: I understand what you are suggesting and I have incorporated another sentence into the abstract:

"The evaporative/fugitive source profile resembles a mix of petroleum operation and non-tailpipe evaporative gasoline sources, but was not responsible for any observed PMF resolved-CH4 enhancements. **The uncertainty in the CH_4 estimates from the oil and gas sector in the bootstrapping analysis is consistent with the ~ 3 % contribution of fugitive emissions to the statewide CH4 inventory."**

page 6089, line 15-25: Why assign uncertainty to GHG and CO measurements in proportion to the square root of hourly GHG enhancement rather than measurement uncertainty? Do the PMF results change significantly if the uncertainty for each time point is estimated in proportion to the standard deviation of the sub-hourly measurements used to construct each hourly average ?

Author's response:

Yes, if the uncertainties ascribed to the individual data points for different species differ significantly as they certainly do in this study in case of gas chromatography-measured VOCs versus cavity ringdown spectroscopy-measured GHGs and CO, the PMF analysis is impacted. Since PMF is attempting to reproduce the time series of each tracer within its ascribed uncertainties, applying a very narrow range of uncertainties to GHGs and CO results in PMF not apportioning these species into multiple sources and instead lumping these species almost completely and exclusively into one source factor profile. This is done to keep the Q/Qexp ratio to a minimum which is what the statistical technique is trying to achieve but this defeats the purpose of using this method to apportion the tracer time series into multiple factor contributions. Hence, a different technique that allots higher uncertainties to GHGs and CO, and at the same time makes these relative uncertainties inversely vary in proportion to the magnitude of enhancements is adopted. This brings the uncertainties in line with those ascribed to VOCs.

Author's changes: None.

Summary of final changes to the ACPD published manuscript

1) Removed space between number and % or per mil symbol everywhere (reviewer 1 comment)

2) Removed excess and unnecessary words throughout text (reviewer 1 comment)

3) Page 2, column 2, line 14 – updated GWP to AR5 values with appropriate change in citation (reviewer 1 comment)

4) Page 2, column 2, line 27, added new sentence after '*of* N_2O*.*'. The new sentence is '*The* N_2O *attribution to the agriculture and soil management factor had a relatively high uncertainty most likely due to an asynchronous pattern of soil-mediated* N_2O *emissions from fertilizer usage and collocated biogenic emissions from crops from the surrounding agricultural operations that is difficult for PMF to detect and apportion.*'

5) Starting at Page 2, column 2, line 27 and ending at Page 3, column 1, line 4, replaced two entire sentences '*The evaporative/fugitive.....and livestock factor.*' with '*The evaporative / fugitive source profile which resembled a mix of petroleum operation and non-tailpipe evaporative gasoline sources did not include a PMF resolved-*CH_4 *contribution that was significant (< 2 %) compared to the uncertainty in the livestock-associated* CH_4 *emissions. The vehicle emission source factor broadly matched VOC profiles of on-road exhaust sources. This source factor had no statistically significant detected contribution to the* N_2O *signals (confidence interval of 3 % of livestock* N_2O *enhancements) and negligible* CH_4 *(confidence interval of 4 % of livestock* CH_4 *enhancements) in the presence of a dominant dairy and livestock factor.*'

6) Page 3, column 1, line 18, added reference '*McMillan et al., 2007*' to existing '*Owen and Silver 2014*'. Added citation to bibliography too

McMillan, A. M. S., Goulden, M. L. and Tyler, S. C.: Stoichiometry of CH_4 and CO_2 flux in a California rice paddy, J. Geophys. Res., 112(G1), G01008, doi:10.1029/2006JG000198, 2007.

7) Page 6, column 2, line 22, added '*accuracy of the*' so that the sentence reads '*assess the accuracy of the inventory*'.

8) Page 6, column 2, line 22, at the end of the column added the following sentences – '*We hypothesize that the PMF analysis will be able to parse the atmospheric observations into unique statistical source combinations that, as an analyst, I will be able to distinguish from each other on the basis of unique VOC source markers and appropriately attribute the* CH_4 *and* N_2O *apportioned to each of these factor profiles to a major source category. We then proceed to answer the scientific question if our top down assessment of the* CH_4 *and* N_2O *inventory can improve our understanding of the bottom-up CARB inventory in the region.*'

9) Page 8, column 2, line 26: As a response to reviewer 1's comments on precision, we added another sentence to make the point clear: "*Prior to the campaign, the precision of measurements of*

each instrument used in this study were determined as the 1-sigma standard deviation of a data set over a given length of time measuring a fixed standard (scuba tank)."

10) Page 8, column 2, line 10: Clarified reviewer 1's comment on "scuba tanks" by adding another sentence: "*The scuba tanks were secondary references and were calibrated before and after the experiments using primary WMO standards obtained from the Global Monitoring Division (GMD) at the NOAA Earth System Research Lab. The calibration tests confirmed that there was no issue in short term stability of these species.*"

11) Page 9, column 3, line 25-26, order of the two references was reversed so that the most recent publication is the last one.

12) Page 9, column 3, line 28, order of the two references was reversed so that the most recent publication is the last one. This should be applied to all cases on multiple references.

13) Page 12, column 1, line 6- addressed Reviewer 1's comments on lack of benzene measurements

14) Page 13, column 2, line 5, replaced (Williams et al., 2010) with Williams et al. (2010)

15) Page 14, end of Section 3.5, new sentences have been added to clearly state what the percentage range and the uncertainty reported throughout the text means. This has been done to address one of Reviewer 1's main comments/critique.

16) Page 15, column 1, line 9 - addressed Reviewer 1 's comments on effect of dilution and concentration of boundary layer

17) Page 16, column 1, line 6, added '*higher*' in blank space '*in the* _____ *PMF-derived uncertainties*'.

18) Page 16, column 2, line 15-20, added more information on mid-latitude NOAA site to address reviewer 1's comments

19) Page 19, column 3, line 24. After end of last sentence '*…(EMFAC, 2011)*', added a new sentence – '*While it is certainly a possibility that current in-use CH_4 emission factor in the inventory may be an underestimation, it seems more logical that the relatively high proportion of CH_4 signal in the vehicle source factor profile is due to contributions from coincident urban sources (e.g. natural gas leaks) mixed into the vehicle gasoline exhausts resulting in a 'mixing' phenomena as discussed in the supplement.*'

20) Page 21, column 1, line 13, replaced '*is*' with '*are*'

21) Page 22, column 1, line 8, at the end, added one more sentence – '*We do find the proportion of regional N_2O enhancements attributed to this sector to be a much larger proportion of total emissions as compared to the state inventory.*'

22) Page 23, column 2, line 7: Addressed Reviewer 1's comments about why CO2 is not included in PMF analysis by adding a sentence.

23) Page 23, column 2, line 19, replaced the entire first sentence 'Factor 4…70%' with the following modified version of that sentence – '*Factor 4 is a significant source of GHGs contributing about 20 - 25 % of the total* N_2O *enhancements in the diurnal cycle (Fig. 7c) with a relatively large 1σ confidence interval of 70 % of the PMF-derived enhancements.*'

24) Page 24, column 1, line 14, replaced '*Goldstein et al., 2014*' with '*Goldstein et al., in prep*'

25) Page 27, column 2, line 16-17, order of the references should be chronological such that the most recent publication is the last one. This should be applied to all cases of multiple references. I have also added one more reference in this particular instance so the citation call in the text should look like '(Pétron et al., 2012; Karion et al., 2013; Miller et al., 2013; Kort et al., 2014)'

Kort, E. A., C. Frankenberg, K. R. Costigan, R. Lindenmaier, M. K. Dubey, and D. Wunch (2014), Four corners: The largest US methane anomaly viewed from space, Geophys. Res. Lett., 41, 6898–6903, doi:10.1002/2014GL061503.

26) TS10 – the doi is 10.1073/pnas.1212272109 so please correct accordingly.

27) TS17 – reference presenetd as

"*Hopke, P.: A guide to positive matrix factorization, Department of Chemistry, Calrkson University, Potsdam, NY pp 1–16, http://www.epa.gov/ttnamti1/files/ambient/pm25/workshop/laymen.pdf.*"

28) TS20 – reference added to bibliography

"*Jordan, C., Fitz, E., Hagan, T., Sive, B., Frinak, E., Haase, K., Cottrell, L., Buckley, S., and Talbot, R.: Long-term study of VOCs measured with PTR-MS at a rural site in New Hampshire with urban influences, Atmos. Chem. Phys., 9, 4677-4697, doi:10.5194/acp-9-4677-2009, 2009.*"

29) TS25 and TS26 – Change this reference to following instead of Pacala et al., 2010. Aslo made the appropriate change in the text in Page 3, Column 2, Line 20

"*NRC (2010). National Research Council Report - Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements, The National Academies Press, Washington, DC, 124 pp., 2010.*"

30) I have revised four tables namely Table 1, Table 2, Table 3 and Table 4 I have reattached the tables in a word document that I will upload along with this summary document:

a) Table 1, page 37 Replace all instances of 'pptv' with 'ppt','ppbv' with 'ppb' and 'ppmv' with 'ppm', respectively

Table 1. PMF dataset with total samples (N) and mixing ratio range (in ppt).

Class	Compound	N	1st percentile	99th percentile	Background
GHG	CH_4[a,c]	619	1855.0	3400.8	1813.6
	CO_2[b,c]	619	390.8	468.3	390.0
	N_2O[a,d]	490	323.3	339.5	323.2
combustion tracer	CO[a,d]	653	118.9	330.6	102.1
straight chain alkanes	propane	592	580.8	30839.0	455.5
	n-butane	587	96.4	12649.0	73.6
	n-pentane	647	93.2	3805.4	64.4
	n-hexane	647	23.1	960.5	17.2
	dodecane	643	1.56	54.3	0
branched alkanes	isopentane	646	165.4	7490.5	100.4
	2,3-dimethylbutane	650	52.5	1747.7	41.1
	2,5-dimethylhexane	651	2.37	145.8	0
	isooctane	647	16.6	476.9	12.3
	4-ethylheptane	651	1.45	52.6	0
	dimethyl undecane	643	0.46	24.9	0
cyclo alkanes	methylcyclopentane	647	23.3	1329.6	20.3
	methylcyclohexane	649	8.10	813.9	0
	ethylcyclohexane	651	1.78	169.1	0
alkenes	propene	592	34.7	3299.9	28.6
	isobutene	595	16.7	422.1	10.7
aromatics	toluene	647	48.8	1749.5	33.1
	ethylbenzene	647	5.83	282.0	0
	m,p-xylene	647	21.8	1127.1	21.8
	o-xylene	647	4.31	405.0	0
	cumene	640	0.55	22.8	0
	1-ethyl-3,4-methylbenzene	651	2.22	358.6	0
	p-cymene	649	0.84	93.9	0
	indane	647	0.45	27.9	0
	1,3-dimethyl-4-ethylbenzene	635	0.46	23.9	0
	naphthalene	654	0.44	19.9	0
unsaturated aldehyde	methacrolein	573	14.2	337.0	0
alchohol	methanol	429	2636.81	88691.8	1085.2
	ethanol	598	1021.93	65759.8	1021.9
	isopropyl alcohol	583	25.7	2001.0	25.7
ketone	acetone	663	142.9	3505.8	142.9
	methyl ethyl ketone	605	8.55	1111.2	0
	methyl isobutyl ketone	629	2.03	71.9	0
aldehyde	propanal	636	3.68	140.8	0
	butanal	589	1.72	35.1	0
biogenics	isoprene	651	9.70	310.0	0
	alpha-pinene	740	1.67	525.8	0
	d-limonene	641	1.10	357.1	0
	nopinone	614	0.78	89.5	0
	alpha-thujene	591	0.52	23.8	0
	camphene	645	0.72	100.3	0
chloroalkanes	chloroform	647	34.1	209.3	31.6
	tetrachloroethylene	641	3.41	120.9	0
	1,2-dichloroethane	640	20.6	103.8	20.6
	1,2-dichloropropane	627	2.40	28.4	0
sulfides	carbon disulfide	610	7.84	133.7	0
thiol	ethanethiol	491	4.54	685.8	0

[a] parts per billion volume (ppb)
[b] parts per million (ppm)
[c] measured using LGR Fast Green House Gas Analyzer
[d] measured using LGR N_2O/CO analyzer

b) Table 2, Page 38
- Replace '*Gentner et al. (2014)*' with '*Gentner et al. (2014a)*' in row 2 of table
- Replace '*Gentner et al. (2014)*' with '*Gentner et al. (2014a)*' in footnote b
- Replace '*SD*' with '*standard deviation*' in footnote a, d, and f

Table 2. Comparison of light alkane ratios to propane (gC $(gC)^{-1}$) from PMF fugitive and evaporative factor with those from other PMF studies and oil and gas operations.

Study	Source	propane	n-butane	n-pentane	n-hexane	isopentane
Bakersfield PMF evaporative and fugitive factor[a]	This study	1	0.52 ± 0.02	0.18 ± 0.01	0.06 ± 0.003	0.33 ± 0.02
Bakersfield petroleum operations source profile[b]	Gentner et al. (2014a)	1	0.53 ± 0.1	0.09 ± 0.02	0.04 ± 0.01	0.08 ± 0.02
Mexico city PMF LPG factor[c]	Bon et al. (2011)	1	0.5 (0.4 - 0.7)	0.05 (0.04 - 0.07)	0.02 (0.02 - 0.03)	0.07 (0.06 - 0.1)
Wattenberg field BAO, Colorado[d]	Gilman et al. (2013)	1	0.75 ± 1.37	0.32 ± 0.6	0.08 ± 0.13	0.28 ± 0.52
Wattenberg field BAO, Colorado[e]	Petron et al. (2012)	1	0.58 - 0.65	0.22 - 0.31	NA	0.22 - 0.31
PMF natural gas and evaporation factor, Houston Ship Channel[g]	Leuchner and Rappengluck (2010)	1	0.33	0.27	0.12	0.37
PMF natural gas factor, Houston Ship Channel[h]	Buzcu and Fraser (2006)	1	0.67 ± 0.16	0.07 ± 0.18	NA	NA

[a] Uncertainties calculated from propagation of errors (standard deviations) over FPEAK range of -1.6 to 0.4.
[b] Ratios calculated from Table 4, Gentner et al., 2014a; uncertainties defined as ±20% to account for variability in oil well data.
[c] Uncertainties calculated from propagation of uncertainties over FPEAK range of -3 to 3.
[d] Emission ratios derived from multivariate regression analysis; error bars derived from propagation of uncertainty using mean and standard deviation of samples.
[e] Range over 5 regressions conducted over data collected in different seasons and from mobile lab samples.
[f] Ratios derived from mean and standard deviations, with propagation of uncertainty.
[g] Estimated from Figure 2, Leuchner and Rappengluck, 2010.
[h] Estimated from Figure 2, Buzcu and Frazer, 2006.

c) Table 3, Page 39
- Replace '*Gentner et al. (2014)*' with '*Gentner et al. (2012)*' in row 2 of the table
- In row 8 of table named n-dodecane, replace '*(0.04 +/- 0.004)*[d]' with '*(0.02 +/- 0.007)*[d]'
- In footnote d, please add '*and gasoline and diesel fuel sale data in Kern County (Table S1, Gentner et al., 2012).*' at the end of the existing sentence
- Replace '*SD*' with '*standard deviation*' in footnote a

Table 3. Comparison of hydrocarbon ratios to toluene (gC (gC)$^{-1}$) from PMF vehicle emission factor with similar ratios from other California specific studies.

Study	Bakersfield PMF vehicle emissions factor[a]	Bakersfield gasoline source profile[b,c]	Riverside liquid gasoline profile[e]	CalNex Los Angeles ambient emission ratios[g]
Source	This study	Gentner et al. (2012)	Gentner et al. (2009)	Borbon et al. (2013)
CH$_4$	8.1 ± 2.1	NA	NA	NA
CO	14.0 ± 0.4	NA	NA	45
toluene	1	1	1	1
isopentane	0.69 ± 0.01	0.77 ± 0.04	0.64-0.84	1.95
isooctane	0.29 ± 0.03	0.34 ± 0.02	0.64-0.80	NA
n-dodecane	0.03 ± 0.001	(0.02 ± 0.007)[d]	NA	NA
methylcyclopentane	0.24 ± 0.01	0.32 ± 0.02	NA	NA
ethyl benzene	0.17 ± 0.01	0.14 ± 0.01	NA	0.2
m/p - xylene	0.65 ± 0.01	0.65 ± 0.03	(0.45-0.52)[f]	0.64
o - xylene	0.22 ± 0.01	0.23 ± 0.01	NA	0.24

[a] errors are standard deviation of 12 unique PMF solutions between FPEAK =-1.6 to +0.4; see section S2.
[b] derived from liquid gasoline fuel speciation profile (Table S9; Gentner et al., 2012).
[c] errors bars derived from propagation of uncertainties.
[d] derived by combining diesel fuel and gasoline speciation profile (Table S9 and S10; Gentner et al., 2012) and gasoline and diesel fuel sale data in Kern County (Table S1, Gentner et al., 2012).
[e] summer data.
[f] only m-xylene.
[g] derived from Linear Regression Fit slope of scatterplot from CalNex Pasadena supersite samples.

d) Table 4, Page 40
- Replace '*Gentner et al. (2014)*' with '*Gentner et al. (2014a)*' in penultimate row of table
- Replace '*Guha et al. (2014)*' with '*Guha et al. (in prep)*' in last row of table
- In footnote b, please make '*4*' in '*CH4*' subscripted like it is in footnote a so it reads 'CH_4'

Table 4. Comparison of PMF dairy and livestock emission rates (mmol mol^{-1}) with previous studies.

Study	Source	Cow/manure type (if applicable)	methanol / methane EF avg. (range)	ethanol / methane EF avg. (range)
PMF analysis of regional measurements	This study		15 - 47	9 - 32.2
Environmental chamber with cows and/or manure	Shaw et al. (2008)	Dry	3.2 (0.6 - 7.4)	NA
		Lactating	1.9 (0.8 - 3.6)	NA
Environmental chamber with cows and/or manure	Sun et al. (2008)	Dry	13.4 (4 - 25)	14.4 (11 - 19)
		Lactating	19.2 (15 - 25)	24.2 (18 - 32)
Cowshed with regular dairy operations (winter)	Ngwabie et al. (2008)		2.0 (1.6 - 2.4)	9.3 (4 - 16)
Cow stall area with regular dairy operations (summer)	Filipy et al. (2006)		NA	(42 - 127)[a]
Manure from cattle feedlot	Miller and Varel (2001)	Fresh (< 24 hr)	NA	14[b]
		Aged (> 24 hr)		118[b]
Measured slope of regression (CalNex 2010)	Gentner et al. (2014a)		7.4 (7 - 16)[c]	18[d]
Sampling of dairy plumes from aircraft (CABERNET 2011)	Guha et al. (*in prep*)		9.6 (9 - 30)[c]	NA

---

[a] calculated based on CH_4 emission rate of 4160 µg cow^{-1} s^{-1} for mid-lactating cows (Shaw et al., 2007).

[b] calculated based on CH_4 emission rate of 4160 µg cow-1 s-1 for mid-lactating cows (Shaw et al., 2007) and ethanol emission rate for fresh and aged manure of 175 and 1223 µg cow-1 s-1, respectively, derived by Filipy et al. (2006).

[c] slope of regression with range of measured slopes (in parentheses) from sampling of dairy plumes by aircraft.

[d] ground site data; lower limit of slope of non-vehicular ethanol versus methane

31) Figure 3 label: added the following underlined words in the caption "nighttime (17:00–06:00 hour) during the experiment period in summer 2010. The concentric…"

32) Figure 4 and 5 label - x and y axis label was reversed .

Source apportionment of methane and nitrous oxide in California's San Joaquin Valley at CalNex 2010 via positive matrix factorization

A. Guha[1*], D. R. Gentner[2†~~*~~], R. J. Weber[1], R. Provencal[3], and A. H. Goldstein[1,2]

[1] Department of Environmental Science, Policy and Management, University of California, Berkeley, USA

[2] Department of Civil and Environmental Engineering, University of California, Berkeley, USA

[3] Los Gatos Research Inc., Mountain View, California, USA

*now at: Planning and Climate Protection Division, Bay Area Air Quality Management District, San Francisco, CA, USA

†now at: Department of Chemical & Environmental Engineering, Yale University, New Haven, CT, USA

Correspondence to: A. Guha (aguha@baaqmd.gov~~abhinavguha@berkeley.edu~~)

Abstract

Sources of methane (CH_4) and nitrous oxide (N_2O) were investigated using measurements from a site in southeast Bakersfield as part of the CalNex (California at the Nexus of Air Quality and Climate Change) experiment from ~~May 15th to June 30th, 2010~~mid-May to June-end, 2010. Typical daily minimum mixing ratios of CH_4 and N_2O were higher than daily ~~averages~~ minima that were simultaneously observed at a ~~similar latitude~~mid-oceanic background station (NOAA, Mauna Loa) by approximately 70 ppb and 0.5 ppb, respectively. Substantial enhancements of CH_4 and N_2O (hourly averages > 500 ppb and > 7 ppb, respectively) were routinely observed suggesting the presence of large regional sources. Collocated measurements of carbon monoxide (CO) and a range of volatile organic compounds (VOCs) (e.g. straight-chain and branched alkanes, cycloalkanes, chlorinated alkanes, aromatics, alcohols, isoprene, terpenes and ketones) were used with a Positive Matrix Factorization (PMF) source apportionment method to estimate the contribution of regional sources to observed enhancements of CH_4 and N_2O.

The PMF technique provided a "top-down" deconstruction of ambient gas-phase observations into broad source categories, yielding a 7-factor solution. We identified these source factors as emissions from evaporative and fugitive; motor vehicles; livestock and dairy; agricultural and soil management; daytime light and temperature driven; non-vehicular urban; and nighttime terpene

biogenics and anthropogenics. The dairy and livestock factor accounted for a majority of the CH_4 (70 - 90%) enhancements during the duration of experiments. ~~Propagation of uncertainties in the PMF-derived factor profiles and time series from bootstrapping analysis resulted in a 29 % error in the CH_4 apportionment to this factor.~~ The dairy and livestock factor was also a principal contributor to the daily enhancements of N_2O (60 – 70%)~~ with an uncertainty of 33 %~~. Agriculture and soil management accounted for ~ 20 - 25% of N_2O enhancements over a 24-hour cycle~~the course of a day (70 % uncertainty)~~, not surprisingly given that organic and synthetic fertilizers are known to be a major source of N_2O. The N_2O attribution to the agriculture and soil management factor had a high uncertainty in the conducted bootstrapping analysis. This is most likely due to an asynchronous pattern of soil-mediated N_2O emissions from fertilizer usage and collocated biogenic emissions from crops from the surrounding agricultural operations that is difficult to apportion statistically when using PMF. The evaporative / fugitive source profile, which resembled a mix of petroleum operation and non-tailpipe evaporative gasoline sources, did not include a PMF resolved-CH_4 contribution that was significant (< 2%) compared to the uncertainty in the livestock-associated CH_4 emissions. The uncertainty of the CH_4 estimates in this source factor, derived from the bootstrapping analysis, is consistent with the ~ 3% contribution of fugitive oil and gas emissions to the statewide CH_4 inventory. The vehicle emission source factor broadly matched VOC profiles of on-road exhaust sources. This source factor had no statistically significant detected contribution to the N_2O signals (confidence interval of 3% of livestock N_2O enhancements) and negligible CH_4 (confidence interval of 4% of livestock CH_4 enhancements) in the presence of a dominant dairy and livestock factor.~~The evaporative/fugitive source profile resembles a mix of petroleum operation and non-tailpipe evaporative gasoline sources, but was not responsible for any observed PMF resolved-CH_4 enhancements. The vehicle emission source factor broadly matches VOC profiles of on-road exhaust sources and had no detected contribution to the N_2O signals and negligible CH_4 in the presence of a dominant dairy and livestock factor. This~~ The CalNex PMF study provides a measurement-based assessment of the state CH_4 and N_2O inventories for the southern San Joaquin valley. The state inventory attributes ~ 18% of ~~the~~ total N_2O emissions to the transportation sector. Our PMF analysis directly contradicts the state inventory and demonstrates there were no discernible N_2O emissions from the transportation sector in the southern SJV region.

1. Introduction

Methane (CH_4) and nitrous oxide (N_2O) are the two most significant non-CO_2 greenhouse gases (GHGs) contributing about 50% and 17~~6~~% of the total direct non-CO_2 GHG radiative forcing (~ 1 W m^{-2}), respectively (Figure SPM.5; IPCC, 2013~~Forster et al., 2007~~). CH_4, with a lifetime of ~ 10 years and Global Warming Potential (GWP) of 34~~25~~ on a 100-year basis, accounting for climate-carbon feedbacks (Table 8.7, Myhre et al., 2013;~~Forster et al., 2007;~~ Montzka et al., 2011), is emitted by both anthropogenic and natural sources (e.g. wetlands, oceans, termites etc.). Anthropogenic global CH_4 emissions are due to agricultural activities (enteric fermentation in livestock, manure management and rice cultivation)(McMillan et al., 2007; Owen and Silver, 2014), energy sector (oil and gas operations and coal mining), waste management (landfills and waste water treatment), and biomass burning (some of which is natural) (Smith et al., 2007; NRC, 2010). N_2O has a higher persistence in the atmosphere (lifetime of ~ 120 years) and stronger infrared radiation absorption characteristics than CH_4 giving it a GWP of 298 (Table 8.7, Myhre et al., 2013~~Forster et al., 2007~~; Montzka et al., 2011). Agriculture is the biggest source of anthropogenic N_2O emissions since the use of synthetic fertilizers and manure leads to microbial N_2O emissions from soil (Crutzen et al., 2007; Galloway et al., 2008). Management of livestock and animal waste is another important ~~agricultural~~ biological source of N_2O, while industrial processes including fossil fuel combustion have been estimated to account for ~10~~5~~% of total global anthropogenic N_2O emissions (Denman et al., 2007).

In 2006, the state of California adopted Assembly Bill 32 (AB32) into a law known as the Global Warming Solutions Act, which committed the state to cap and reduce anthropogenic GHG emissions to 1990 levels by 2020. A statewide GHG emission inventory (CARB, 2013) maintained by the Air Resources Board of California (CARB) is used to report, verify and regulate emissions from GHG sources. In 2011, CH_4 accounted for 32.5 million metric tonnes (MMT) CO_2-eq representing 6.2% of the statewide GHG emissions, while N_2O emissions totaled 13.4~~6~~ MMT CO_2-eq representing about 3% of the GHG emissions inventory (Figure 1). CARB's accurate knowledge of GHG sources and statewide emissions is a key component to the success of any climate change mitigation strategy under AB32. CARB's GHG inventory is a ''bottom–up'' summation of emissions derived from emission factors and activity data. The bottom-up approach is reasonably accurate for estimation and verification of emissions from mobile and point sources (vehicle tailpipes, power plant stacks etc.) where the input variables are well-understood and well-quantified. The main anthropogenic sources of CH_4 in the CARB inventory ~~include~~ are ruminant livestock and manure management, landfills, wastewater treatment, fugitive and process losses from oil and gas production and transmission, and rice cultivation while the major N_2O sources are agricultural soil management, livestock manure management and vehicle fuel combustion (CARB, 2013). The emission factors for many of these sources have large uncertainties as they are biological in nature and their production and release mechanisms are inadequately understood thus making these sources unsuitable for direct measurements (e.g. emissions of N_2O from farmlands). Many of these sources (e.g. CH_4 from landfills) are susceptible to spatial heterogeneity and seasonal variability. Unfortunately, a more detailed understanding of source characteristics is made difficult because CH_4 and N_2O are often emitted from a mix of point and area sources within the same source facility (e.g. dairies in the agricultural sector) making bottom-up estimation uncertain. There is a lack of direct measurement data or "top-down" measurement-based approaches to independently validate seasonal trends and inventory estimates of CH_4 and N_2O in California's Central Valley, which has a mix of several agricultural sources and oil and gas operations, both of which are known major sources of GHGs.

In the recent past, regional emission estimates derived from measurements from a tall tower at Walnut Grove in Central California coupled with inverse dispersion techniques (Fischer et al., 2009) reported underestimation of CH_4 and N_2O emissions especially in the Central Valley. Comparison of regional surface footprints determined from WRF-STILT algorithm between Oct-Dec 2007 indicate posterior CH_4 emissions to be higher than California-specific inventory estimates by 37 ± 21% (Zhao et al., 2009). Predicted livestock CH_4 emissions are 63 ± 22% higher than *a priori* estimates. A study over a longer period (Dec 2007 - Nov 2008) at the same tower (Jeong et al., 2012a) generated posterior CH_4 estimates that were 55 - 84% larger than California-specific prior emissions for a region within 150 km from the tower. For N_2O, inverse estimates for the same sub-regions (using either EDGAR32 and EDGAR42 *a priori* maps) were about twice as much as *a priori* EDGAR inventories (Jeong et al., 2012b). Recent studies have incorporated WRF-STILT inverse analysis on airborne observations across California (Santoni et al., 2012). The authors conclude that CARB CH_4 budget is being underestimated by a factor of 1.64 with aircraft-derived emissions from cattle and manure management, landfills, rice, and natural gas infrastructure being around 75%, 22%, 460%, and 430% more than CARB's current estimates for these categories, respectively. Statistical source footprints of CH_4 emissions generated using FLEXPART-WRF modeling and CalNex-Bakersfield CH_4 concentration data are consistent with locations of dairies in the region (Gentner et al., 2014a). The authors conclude that the majority of CH_4 emissions in the region originate from dairy operations. Scaled-up CH_4 rice cultivation estimates derived from aircraft CH_4/CO_2 flux ratio observations over rice paddies in the Sacramento valley during the growing season when emissions are at their strongest (Peischl et al., 2012) are around three times larger than inventory estimates. CH_4 budgets derived for the Los Angeles (LA) basin from aircraft observations (Peischl et al., 2013) and studies involving comparison with CO enhancements and inventory at Mt. Wilson (Hsu et al., 2010; Wunch et al., 2009) indicate higher atmospheric CH_4 emissions in the LA basin than expected from bottom-up accounting.

Recent literature seems to suggest that the CARB bottom-up inventory is underestimating CH_4 and N_2O sources, especially from the livestock sector and perhaps from the oil and gas industry as well. Source apportionment studies of non-CO_2 GHGs over the Central Valley can provide critical information about under-inventoried or unknown sources that seek to bridge the gap between 'bottom-up' and 'top-down' methods. GHG emission inventories can potentially be constrained through simultaneous measurements of GHGs and multiple gas species (VOCs) that are tracers of various source categories. This study provides CH_4 and N_2O source attribution during a six-week study involving a complete suite of continuous GHG and VOC tracer measurements during the CalNex 2010 campaign in Bakersfield which is~~,~~ located in the southern part of the Central Valley (May - June 2010). The objective of this study is to partition the measured CH_4, N_2O and VOC enhancements into statistically unique combinations using Positive Matrix Factorization (PMF) apportionment technique. We classify these combinations as plausible source factors based on our prior knowledge of the chemical origin of mutually co-varying groups of VOC tracers found in each statistical combination. We examine the source categorization using observations from source-specific, ground site and airborne measurements and results from other source apportionment studies. We also compare the relative abundance of CH_4 and N_2O enhancements in each source factor with the CARB inventory estimates in order to assess the accuracy of the inventory. We hypothesize that the PMF analysis will be able to parse the atmospheric observations into unique statistical source combinations that, as analysts, we would be able to distinguish from each other on the basis of unique VOC source markers. We should, thus, be able to appropriately attribute the CH_4 and N_2O apportioned to each of these factor profiles to a major source category. We then proceed to answer the scientific question if our top down assessment of the CH_4 and N_2O inventory can improve our understanding of the bottom-up CARB inventory in the region.

2. Experimental Setup

2.1. Field Site and Meteorology

Measurements were conducted from 19 May to 25 June 2010 at the Bakersfield CalNex supersite (35.3463°N, 118.9654°W) (Figure 2) in the southern San Joaquin Valley (SJV) (Ryerson et al., 2013). The SJV represents the southern half of California's Central Valley. It is 60 to 100 km wide, surrounded on three sides by mountains, with the Coastal Ranges to the west, the Sierra Nevada Mountains to the east, and the Tehachapi Mountains to the southeast.

The measurement site was located to the southeast of the Bakersfield urban core in Kern County (Figure 2). The east-west Highway 58 is located ~~about~~ ~ 0.8 km ~~to the~~ north; the north-south Highway 99 is ~~about~~ ~ 7 km ~~to the~~ west. The city's main waste water treatment plant (WWTP) and its settling ponds are located ~~to the~~ east and south ~~of the site~~ (< 2.5 km), respectively. Numerous dairy and livestock operations are located ~~to the~~ south-southwest ~~of the site~~ at 10 km distance ~~or farther~~and onwards. The metropolitan region has three major oil refineries located within 10 km ~~from the site~~ (two to the northwest; one to the southeast). A majority of Kern County's high-production active oil fields (> 10000 barrels (bbl) per day) (CDC, 2013) are located to the west/northwest and are distant (~ 40 - 100 km). Kern River oilfield (~ 60000 bbl day^{-1}), one of the largest in the country, and Kern Front (~ 11000 bbl day^{-1}) are located about 10 - 15 km to the north. There are several other oil fields dotted within the urban core (5 - 20 km) which are less productive (< 2000 bbl day^{-1}) or not active (< 100 bbl day^{-1}). The whole region is covered with agricultural farmlands with almonds, grapes, citrus, carrots and pistachios amongst the top commodities by value and acreage (KernAg, 2010).

The meteorology and transport of air masses in the southern SJV is complex and has been addressed previously (Bao et al., 2007; Beaver and Palazoglu, 2009). The wind rose plots (Figure 3) shown here present a simplified distribution of microscale wind speed and direction at the site for the campaign duration, the latter often being non-linear over larger spatial scales. The plots depict broad differences in meteorology during daytime and nighttime. A mesoscale representation of the site meteorology during this study period was evaluated through back-trajectory footprints generated from each hourly sample using FLEXPART Lagrangian transport model with WRF meteorological modeling (Gentner et al., 2014a). The 6-h and 12-h back trajectory footprints are generated on a 4 × 4 km resolution with simulations originating from top of the 18-m tall tower. The site experiences persistent up-valley flows from the north and northwest during afternoons and evenings, usually at high wind speeds. The direction and speed of the flow during ~~the~~ nights is quite variable (Figure 3). On some nights, the up-valley flows diminish as night-time inversion forms a stable layer near the ground, and eventually downslope flows off the nearby mountain ranges bring winds from the east and south during late night and early morning periods. On other nights, fast moving northwesterly flows extend in to middle of the night leading to unstable conditions through the night. The daytime flows bring plumes from the upwind metropolitan region (Figure 3), and regional emissions from sources like dairies and farmlands located further upwind. The slow nighttime flows and stagnant conditions cause local source contributions to be more significant than during daytime, including those from nearby petroleum operations and dairies (Gentner et al., 2014a), and agriculture (Gentner et al., 2014b).

3. Methods

3.1. Trace gas measurements and instrumentation

Ambient air was sampled from the top of a tower (18.7 m a.g.l~~AGL~~) through Teflon inlet sampling lines with Teflon filters to remove particulate matter from the gas stream. CH_4, CO_2 and H_2O were measured using a Los Gatos Research (LGR Inc., Mountain View, CA) Fast Greenhouse Gas Analyzer (FGGA, Model 907-0010). N_2O and CO were measured by another LGR analyzer (Model 907-0015) with time response of ~ 0.1 to 0.2 Hz. These instruments use off-axis Integrated Cavity Output Spectroscopy (ICOS) (O'Keefe, 1998; Paul et al., 2002; Hendriks et al., 2008; Parameswaran et al., 2009). The FGGA instrument internally calculates and automatically applies a water vapor correction to counter the dilution effect of water on a target molecule ~~corrects for water vapor dilution~~ and ~~reports~~ calculates CH_4 and CO_2 on a dry (and wet) mole fraction basis. We report dry mole fraction mixing ratios in this study. The FGGA instrument had a 1σ-precision of 1 ppb (for CH_4) and 0.15 ppm (for CO_2) while the N_2O/CO instrument had a 1σ-precision of 0.3 ppb, respectively over short time periods (< 10 s). Prior to the campaign, the precision of measurements of each instrument used in this study were determined as the 1-sigma standard deviation of a data set over a given length of time measuring a fixed standard (scuba tank) and found to conform to the manufacturer specifications. The instruments were housed at ground level in a thermally insulated temperature controlled 7-foot wide cargo wagon trailer developed by the ~~GHG~~ instrument manufacturers (Los Gatos Research Inc.). CO was coincidentally measured using another instrument (Teledyne API, USA, Model # M300EU2) with a precision of 0.5-% of reading and output as 1-minute averages~~s~~. The mixing ratios from the two collocated CO instruments correlated well (r ~ 0.99) and provided a good stability check for the LGR instrumentation. Scaled Teledyne CO data was used to gap-fill the LGR CO data. The coincident gas-phase VOC measurements were made using a gas chromatograph (GC) with a quadrapole mass selective detector and a flame ionization detector (Gentner et al., 2012).

Hourly calibration checks of the three GHGs and CO were performed using near-ambient level scuba tank standards through the entire campaign. The scuba tanks were secondary references and were calibrated before and after the experiments using primary standards conforming to the WMO mole fraction scale obtained from the Global Monitoring Division (GMD) at the NOAA Earth System Research Lab. The calibration tests confirmed that there was no issue in short term stability of these species. During data processing, final concentrations were generated from the raw data values using scaling factors obtained from comparison of measured and target concentrations during secondary calibration checks. Diurnal plots of measured species are generated from 1-min averages. PMF analyses in the following sections are based on 30-minute averages to match the time resolution of VOC measurements. The meteorological data measured at the top of the tower included relative humidity (RH), temperature (T), and wind speed (WS) and direction (WD).

3.2. Positive Matrix Factorization (PMF)

Source apportionment techniques like PMF have been used in the past to apportion ambient concentration datasets into mutually co-varying groups of species. PMF is especially suitable for studies where *a priori* knowledge of contributing number of sources impacting the measurements, chemical nature of source profiles and relative contribution of each source to the concentration time series of a measured compound are unknown or cannot be assumed. PMF has been applied to ambient particulate matter studies (Lee et al., 1999; Kim et al., 2004~~; Lee et al., 1999~~); in determining sources of atmospheric organic aerosols (OA) (Ulbrich et al., 2009; Slowik et al., 2010; Williams et al., 2010); and in gas phase measurements of VOCs in major metropolitan cities (Brown et al., 2007; Bon et al., 2011~~; Brown et al., 2007~~). PMF is a receptor-only unmixing model which breaks down a measured data set containing time series of a number of compounds into a mass balance of an arbitrary number of constant source factor profiles (FP) with varying concentrations over the time of the data set (time series *or* TS) (Ulbrich et al., 2009).

In real world ambient scenarios, ~~sources of~~ emissions sources ~~-~~are often not known or well-understood. PMF technique requires no a priori information about the number or composition of factor profiles or time trends of those profiles. The constraint of non-negativity in PMF ensures that all values in the derived factor profiles and their contributions are constrained to be positive leading to physically meaningful solutions. PMF requires the user to attribute~~s~~ a measure of experimental uncertainty (or weight) to each input measurement. Data point weights allow the level of influence to be related to the level of confidence the analyst has in the measured data (Hopke, 2000). In this way, problematic data such as outliers, below-detection-limit (BDL), or altogether missing data can still be substituted into the model with appropriated weight adjustment (Comero et al., 2009) allowing for a larger input data set, and hence a more robust analysis. PMF results are quantitative; it is possible to obtain chemical composition of sources determined by the model (Comero et al., 2009). PMF is not data-sensitive and can be applied to data sets that are not homogenous and/or require normalization without introducing artifacts.

3.3. Mathematical Framework of PMF

The PMF model is described in greater detail elsewhere (Paatero and Tapper, 1994; Paatero 1997; Comero et al., 2009; Ulbrich et al., 2009) and we will briefly mention some concepts relevant to the understanding of the analysis carried out in this study. The PMF input parameters involve a $m \times n$ data matrix **X** with i rows containing mixing ratios at sampling time t_i and j columns containing time series of each tracer_j. A corresponding uncertainty matrix **S** reports measurement precision (uncertainty) of the signal of each tracer_j at every t_i (s_{ij}). The PMF model can then be resolved as:

$$\boldsymbol{X}_{ij} = \sum_{p} g_{ip} f_{pj} + e_{ij} \tag{1}$$

where p refers to the number of contributing factors in the solution as determined by the analyst (discussed below), g_{ij} (mass concentration) are elements of a $m \times p$ matrix **G** whose columns represent the factor time series while f_{ij} (mass fraction) are elements of a $p \times n$ matrix **F** whose rows represent the factor chemical profiles. e_{ij} are the elements of a $m \times n$ matrix **E** containing residuals not fit by the model matrix at each data point.

The PMF algorithm uses a least-squares algorithm to iteratively fit the values of **G** and **F** by minimizing a "~~a~~ quality of fit" parameter Q (Bon et al., 2011), defined as:

$$Q = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(e_{ij}/s_{ij}\right)^2 \tag{2}$$

In this way, PMF minimizes the sum of squares of error-weighted model-measurement deviations. The theoretical value of Q, denoted by Q-expected (Q_{exp}) can be estimated as:

$$Q_{exp} \equiv (m \times n) - p \times (m + n) \tag{3}$$

If all the errors have been estimated within the uncertainty of the data points (i.e. $e_{ij}\ s_{ij}^{-1} \sim 1$) and the model fits the data perfectly, then Q should be approximately equal to Q_{exp}.

3.4. Data preparation for PMF analysis

For this study, measurements from the FGGA, LGR N_2O/CO analyzer and the GC were combined into a unified data set to create matrices **X** and **S**. Only VOCs that are a part of broad chemical composition of nearby sources (like dairies and vehicle emissions) or could potentially serve as source specific tracers (e.g. iso-octane as a tailpipe emissions tracer; isoprene as a biogenic tracer) were included. Isomers were limited (e.g. 2,3-Dimethylbutane over 2,2-Dimethylbutane) and VOCs with large number of missing values were not included. The input data set represented major chemical families like straight-chain and branched alkanes, cycloalkanes, alkenes, aromatics, alcohols, aldehydes, ketones and chlorinated as well as organosulfur compounds. In spite of best efforts by the authors, it was not possible to quantify the magnitude of observed concentrations of benzene relative to the positive artifacts coming from the Tenax TA adsorbent (previously documented elsewhere). Hence, benzene was not included in the PMF analysis. In all, there were a total of 653 half-hour samples of data collected ~~covering a period~~ from 22 May to 25 June. In the days prior to and after this period, there was no N_2O and/or VOC data collected and hence the PMF analysis is limited to this period. Table 1 lists all the compounds included in the PMF analysis along with a spectrum of observed and background concentrations.

PMF analysis resolves the covariance of mixing ratio enhancements and thus characterizes the chemical composition of emissions from various sources. Hence, for this analysis, only enhancements representing local emissions were included in the data set after subtracting local background concentrations from the original signals. Background concentrations were derived as the minima in the time series (0th percentile) for each of the 50 tracers included in the PMF analysis (CH_4, N_2O, CO and 46 VOCs). For VOCs, tracers with a minimum value less than two times the limit of detection (LOD, in ppt) and a maximum value larger than hundred times the LOD were assumed to have a negligible background (0 ppt) (Table 1). The 99th percentile for each tracer was treated as the effective-maximum mixing ratio and the upper limit of the range for the "normalization" of time series. Enhancements above the 99th percentile are often extreme values~~Data points representing enhancement values above the 99th percentile are often extreme data points~~. Such outliers, even if true enhancements, represent isolated and short-duration footprints of high-emission events that are difficult for PMF to reconstruct. In order to maintain the robustness of PMF analysis, outliers were selectively down-weighted by increasing their uncertainty in proportion to the uncertainty of other data points (described below). Finally, the enhancements in each time series were "normalized" by dividing every sample by the difference in the 99th percentile and background (the range) as seen in Equation 4. This process scaled the enhancements in each time series (final data points in **X**) within a range of 0 to 1. This allowed for a consistent scheme to represent tracers with vastly different concentrations (e.g. ppm level of CH_4 vs ppt level of propene) and improve the visual attributes of PMF output plots to follow. Data points denoting zero enhancement (lower limit) were replaced by a very small positive number (i.e. exp(-5)) to avoid 'zeros' in the data matrix **X**.

$$x_{ij} = \left(Mixing\ ratio_{ij} - Background_j\right) \Big/ \left(Maximum\ mixing\ ratio_j - Background_j\right) \tag{4}$$

For the VOCs, guidelines set forth by Williams et al., (2010)~~(Williams et al., 2010)~~ were adopted to calculate the uncertainty estimates. An analytical uncertainty (AU) of 10~~-~~% was used; a limit of detection (LOD) of 1 ppt and a limit of quantification (LOQ) of 2 ppt (Gentner et al., 2012) was used to calculate the total uncertainty for each x_{ij}:

$$s_{ij} \equiv 2 \times LOD, \quad \text{if } x_{ij} \leq \text{LOD}, \tag{5a}$$

$$s_{ij} \equiv LOQ, \quad \text{if LOD} < x_{ij} \leq \text{LOQ}, \tag{5b}$$

$$s_{ij} \equiv \left(\left(AU \times x_{ij}\right)^2 + (LOD)^2\right)^{0.5}, \quad \text{if } x_{ij} > \text{LOQ} \tag{5c}$$

Using this approach, ~~the~~ detection limit dictates the errors for low enhancements (near LOD) while~~ the~~ errors for larger enhancements of VOCs are tied more to the magnitude of the data value (x_{ij}) itself.

The GHG and CO measurements have high precision and significantly lower detection limits than ambient levels. The relatively low values of GHGs in the uncertainty matrix, compared to VOCs, is substituted with those calculated using a custom approach. The GHG and CO uncertainties are assumed to be proportional to the square root of the data value and an arbitrary scaling factor determined through trial and error in order to produce lower values of $Q\ Q_{exp}^{-1}$:

$$s_{ij} \equiv A \times \left(x_{ij}\right)^{0.5}, \text{ where A = 1 (for } CH_4\text{), 0.25 (for } CO_2\text{), 0.5 (for CO), 0.1 (for } N_2O\text{)} \tag{6}$$

This method attributes larger percentage uncertainties to smaller enhancements and hence lesser weight in the final solution and vice versa. This approach leads to an uncertainty matrix that attributes relatively similar percentage errors to both GHGs and VOCs, which should lead to a better fitting of the data through PMF.

Missing values are replaced by geometric mean of the tracer time series and their accompanying uncertainties are set at four times this geometric mean (Polissar et al., 1998) to decrease their weight in the solution. Based on the *a priori* treatment of the entire input data (scaling) and the corresponding outputs of the PMF analysis, a weighting -approach (for measurements from different instruments) as used in (Slowik et al., 2010) is not found to be necessary.

3.5. PMF source analysis

We use the customized software tool (PMF Evaluation Tool v2.04, PET) developed by Ulbrich et al. (2009) in Igor Pro (Wavemetrics Inc., Portland, Oregon) to run PMF, evaluate the outputs and generate statistics. The PET calls the PMF2 algorithm (described in detail in Ulbrich et al., 2009) to solve the bilinear model for a given set of matrices **X** and **S** for different numbers of factors *p* and for different values of FPEAK or SEED (defined and described later). The tool also stores the results for each of these combinations in a user friendly interface that allows simultaneous display of ~~the~~ factor profiles (FP) and time series (TS) of a chosen solution along with residual plots for individual tracers. A detailed explanation of PMF analysis performed in this study is provided in the Supplement (see Section S). The supplement describes the PMF methodology of how the final number of user-defined factors was chosen (Section S1), the outcomes of linear transformations (rotations) of various PMF solutions (Section S2) and how uncertainties in the chosen solution were derived (Section S3). The standard deviations in the mass fractions of individual tracers in each factor profile and time series of each factor mass is evaluated using a bootstrapping analysis (Norris et al., 2008; Ulbrich et al., 2009) and described in Section S3. The uncertainty of a tracer contribution to a source factor is derived from the 1-sigma deviation of the averaged mass fraction of that tracer to that factor from 100 bootstrapping runs. This is the quantity we report and refer to as 'uncertainty' throughout the text in Section 4. The percentage ranges reported in the abstract and in Section 4 are derived from the relative apportionment of CH_4 and N_2O to different source factors over the 653 half-hourly samples collected during the experiment period. This range represents the mean diurnal range observed and as seen in Figure 7. ~~These error estimates are combined and propagated to derive PMF-based uncertainties for each factor's contribution to source-apportioned diurnal enhancements for a specific compound (Section 4).~~ This diurnal range combined with bootstrapping based uncertainty estimates is used to understand better the contribution of each source factor to the observed enhancements of a target GHG and the analyst's confidence in those estimates.

4. Results and Discussion

In Bakersfield, there are a multitude of pollutant sources, ranging from local to regional, from biogenic to anthropogenic, and from primary to secondary. We recognize that PMF analysis is not capable of precise separation of all sources. In PMF analysis, the analyst chooses the number of factor profiles to include in the solution and assigns a source category interpretation for each identified factor. The PMF factors are not unique sources but really statistical combinations of coincident sources. The chemical profile of each factor may contain some contributions from multiple sources that are collocated~~co-located~~, or have a similar diurnal pattern of emissions. The cycle of daytime dilution of the boundary layer and nighttime inversion can also result in a covariance that can lead to emissions from unrelated sources being apportioned to a single source factor. Such limitations have been observed previously by Williams et al. (2010) while applying PMF in an urban-industrial setting like Riverside, California. The user must infer the dominant source contributions to these individual factors. Our factor profile (FP) nomenclature is based on the closest explanation of the nature and distribution of emission sources in the region. The source factor names should be treated with caution bearing in mind the physical constraints of the solution and not used to over-explain our interpretation of the region's CH_4 and N_2O inventories.

A seven factor solution has been chosen to optimally explain the variability of the included trace gases. The factors have been named based on our interpretation of the emission "source" categories they represent, with corresponding colors which remain consistent in the discussion across the rest of the paper: evaporative and fugitive (black), dairy and livestock (orange), motor vehicles (red), agricultural + soil management (purple), daytime biogenics + secondary organics (light blue), non-vehicular urban (green) and nighttime anthropogenic + terpene biogenics (navy blue). Figure 4 presents the Factor Profile (FP) plots of each factor. The sum of the normalized contributions of the 50 species in each "source" is equal to 1 in the FP plots. Figures 5a through 5g present the diurnal profiles based on mean hourly concentrations (in normalized units) of each PMF factor with standard deviations explaining the variability. The interpretation of the individual FPs is discussed below (in Section 4.2-4.8). Molar emission factor (EF) of tracers with respect to (w.r.t) one another can be derived for each FP. These EFs can then be compared to those from previous source-specific and apportionment studies (Table 2 through 5). The ratio of PMF-derived total CH_4 enhancement to the input measured CH_4 enhancement ranges from 0.90 to 0.95 (mol mol^{-1}) through the whole time series except outliers with really high values (> 500 ppb). For N_2O, the ratio is somewhat lower (0.82-0.92 mol mol^{-1}) and this is reflected in the higher PMF-derived uncertainties. The apportionment of some N_2O mass into a statistically weak and time-varying factor is discussed in Section 4.5. The general assessment is that PMF analysis is able to reconstruct a majority of the measured enhancements for both CH_4 and N_2O.

4.1. Time trends of measured CH_4, CO_2, CO, and N_2O

The time series of CH_4, CO_2, CO, and N_2O mixing ratios have been plotted in Figures 6a through 6d while the diurnal variations have been plotted in Figures 6e-6h, respectively. The color markers in each plot indicate the median wind direction. The daily minima for the three GHGs and CO occur during the late afternoon period when daytime heating, mixing and subsequent dilution occurs rapidly. The daily minimum values of CH_4 and N_2O were larger than that observed at National Oceanic and Atmospheric Administration's (NOAA) Mauna Loa station at 19.48°N latitude in Hawaii (Dlugokencky et al., 2014) by at least 70 ppb and 0.5 ppb, respectively, for this period. We also compare Bakersfield (at 35.36°N latitude) observations to that from NOAA's Trinidad Head station which is located on the coast in Northern California and is more representative of mid-latitudes at 40.97°N latitude. Although there was no N_2O data collected at Trinidad Head, the CH_4 concentrations observed in discrete flask samples collected every few days during summer of 2010 (not necessarily a daily low background) were consistently lower than the daily minimum CH_4 concentration curve at Bakersfield by 10 - 15 ppb. This indicates that there are significant GHG ~~large~~ emissions from regional sources ~~of these two GHGs~~ around Bakersfield that get added to the already higher local background concentrations, thus keep the local mixing ratio levels quite high. Winds during the highest temperature period between noon and evening (12:00 - 20:00 hour local time) almost always arrive through the urban core in the northwest. Any PMF factor whose dominant source direction is northwest is likely to contain contributions from VOCs emitted from urban sources, regional sources further upwind or contain contributions from secondary tracers generated from photochemical processing during the day. The three GHGs show a sharp increase during ~~the~~ nighttime when the inversion layer builds up and traps primary emissions close to the ground. For CO, measured concentrations show two distinct peaks in the diurnal plot (Figure 6g). The observed early morning peak in the concentration is a combination of decreased dilution and fresh emissions from the morning motor vehicle traffic. The late evening peak in CO concentrations is not coincident with rush hour and is a result of build-up of evening emissions in the boundary layer that is getting shallower as the night progresses. Figure 6a indicates CH_4 enhancements of 500 ppb or more on almost every night with peak mixing ratios exceeding 3000 ppb on several occasions indicating an active methane source(s) in the region. Figure 6d shows that peak N_2O mixing ratios rise above 330 ppb on almost every night suggesting large sources in the region. Huge enhancements of CH_4, CO_2 and N_2O (on DOY 157,164, and 165) (in Figures 6a, 6b and 6d, respectively) may appear well-correlated to each other due to regional sources emitting into the inversion layer. However, the shapes of the diurnal cycles differ indicating different emission distributions, with the early morning maximum in CH_4 occurring before the maxima for CO_2 and N_2O, and the morning maximum for CO occurring slightly later. These differences in timing allow PMF analysis to differentiate their contributions into separate factors.

4.2. Factor 1: Evaporative and fugitive emissions

Factor 1 has a chemical signature indicative of evaporative and fugitive losses of VOCs. The FP of this source is dominated by C_3 to C_6 straight-chain and branched alkanes and some cycloalkanes (Figure 4). The average diurnal cycle of Factor 1 (Figure 5a) shows a broad peak during late night and early morning hours after which the concentrations begin to decrease as the day proceeds reaching a minimum at sunset before beginning to rise again. This is strong indication of a source containing primary emissions that build up in the shallow pronounced nighttime inversions of southern SJV. The subsequent dilution of primary emissions as the mixed layer expands leads to low concentrations during the daytime.

Most of the propane, n-butane and pentanes signal is apportioned to this factor, but not the typical vehicle emission tracers like isooctane or CO or any of the alkenes or aromatics. Absence of these tracers in the FP suggests this factor is not related to vehicular exhaust and is a combination of non-tailpipe emissions and fugitive losses from petroleum operations. None of the CH_4 signal at the SJV site is apportioned to this factor, but almost all of the small straight-chain alkanes, exclusively apportion to this factor. This is in agreement with Gentner et al. (2014a) where the authors show ~~which concluded~~ that VOC emissions from petroleum operations are due to fugitive losses of associated gas from condensate tanks following separation from CH_4. Table 2 compares EFs derived from this PMF study for the non-tailpipe (evaporative) and fugitive petroleum operation source factor with those from the Gentner et al. (2014a) study done on the same CalNex dataset using an independent source receptor model with chemical mass balancing and effective variance weighting method, ~~-~~and also to, reports of fugitive emissions from the oil and natural gas sources (Pétron et al., 2012; Gilman et al., 2013) and similar factors produced by other PMF studies (Buzcu and Fraser, 2006; Leuchner and Rappenglück, 2010; Bon et al., 2011). Good agreement of Factor 1 VOC EFs with those from the mentioned studies confirms petroleum operations in Kern County as the major source contributing to this factor. The PMF apportionment indicates that this source factor does not contribute to CH_4 enhancements observed at the SJV site (Figure 7a) and thus most of the ‘associated’ CH_4 is likely separated from the condensate prior to emission. As mentioned before, a tiny fraction (~ 5~~-~~%; Section 4) of the total input CH_4 enhancement is not resolved into source-apportioned contributions. There could be a minor contribution to CH_4 signal from this source, which is unresolved within the framework of uncertainties in the PMF analysis.

4.3. Factor 2: Motor vehicle emissions

Factor 2 has a chemical signature consistent with the tailpipe exhausts of gasoline and diesel motor vehicles. This source factor includes the combustion tracer CO, and other vehicular emissions tracers, such as isooctane (Figure 4). Alkenes are a product of incomplete fuel combustion in motor vehicles, and almost all of the propene and a significant portion of the isobutene signal are attributed to this source factor. The diurnal variation of Factor 2 shows two distinctive peaks (Figure 5b). The first peak occurs in the morning between 06:00 and 07:00 local time and is influenced by morning rush hour traffic, with suppressed mixing allowing vehicle emissions to build up. As the day proceeds, accelerated mixing and dilution (and perhaps chemical processing of reactive VOCs) reduce the enhancements to a minimum by late afternoon. The evening peak mainly occurs as the dilution process slows down after sunset and emissions build up. The increased motor vehicle traffic in the evening adds more emissions to the shrinking boundary layer. This build-up reaches a peak around 22:00~~ in the night~~. The occasional high wind events from the northwest (unstable conditions) and ~~the reasonably lesser number of~~fewer vehicles operating on the roads during ~~the ~~late nighttime hours contribute to the relatively lower levels of enhancements as compared to the peaks on either side of this nighttime period.

Table 3 compares selective PMF derived EFs from vehicle emissions ~~factor~~ with the measured gasoline composition collected during CalNex in Bakersfield (Gentner et al., 2012), analysis of gasoline samples from Riverside in Los Angeles basin (Gentner et al., 2009) and ambient VOC emission ratios measured during CalNex at the Pasadena supersite (Borbon et al., 2013). Although, the two Bakersfield studies employ different source apportionment techniques (and so do the studies conducted in the Los Angeles basin), we observe a broad agreement of relative emission rates of vehicular emission tracers. This agreement validates our assertion that Factor 2 represents a broad suite of vehicular tailpipe emissions.

The PMF derived CH$_4$/CO EF in Factor 2 is 0.58 (mol mol^{-1}) and is significantly higher than the range of 0.03 - 0.08 (mol mol^{-1}) calculated from results of a vehicle dynamometer study of 30 different cars and trucks (Nam et al., 2004) and an EF of 0.014 (mol mol^{-1}) calculated for SJV district during summer of 2010 using EMFAC, which is CARB's model for estimating emissions from on-road vehicles operating in California (EMFAC, 2011). While it is certainly a possibility that current in-use CH$_4$ emission factor in the inventory may be an underestimation, it seems more logical that the relatively high proportion of CH$_4$ signal in the vehicle source factor profile is due to contributions from coincident urban sources (e.g. natural gas leaks) mixed into the vehicle gasoline exhausts resulting in a 'mixing' phenomena as discussed in the supplement. In spite of the non-negligible proportion of CH$_4$ in the

Factor 2 source profile, the contribution of the factor to CH_4 enhancements (Figure 7a) at Bakersfield is negligible relative to the dairy and livestock factor.

The state GHG inventory attributes about 18% of the 2010 statewide N_2O emissions to the on-road transportation sector (CARB 2012). Our PMF analysis shows essentially a negligible enhancement of N_2O associated with the vehicle emission Factor 2 with a PMF derived N_2O/CO EF of 0.00015 (mol mol^{-1}). The EMFAC generated N_2O/CO EF in SJV during summer of 2010 is more than 20 times higher at 0.0034 (mol mol^{-1}). The PMF derived 'vehicle emissions' contribution to N_2O is in stark contrast to the inventory and is an important outcome suggesting a significant error in EFs used to derive the statewide inventory for N_2O.

4.4. Factor 3: Dairy and livestock emissions

Factor 3 has a chemical signature indicative of emissions from dairy operations. This source factor is the largest contributor to CH_4 enhancements (Figure 7a) and a significant portion of the N_2O signal (Figure 7c). The FP also has major contributions from methanol (MeOH) and ethanol (EtOH), with minor contributions from aldehydes and ketones (Figure 4). A separate PMF analysis with a broader set of VOC measurements at the same site showed that most of the acetic acid (CH_3COOH) and some formaldehyde (HCHO) signal attributed to this factor as well (Allen Goldstein, personal communication, 2014). All the above-mentioned VOCs are emitted in significant quantities from dairy operations and cattle feedlots (Filipy et al., 2006; Shaw et al., 2007; Ngwabie et al., 2008; Chung et al., 2010). About 70-90% of the diurnal CH_4 signal is attributed to this factor (Figure 7a) depending on the time of day. The CH_4 dairy and livestock mass fraction which is used to calculate this diurnal range has ~~From propagation of errors,~~ an uncertainty of 29% derived using the bootstrapping method~~is determined in the diurnal CH_4 enhancements in Factor 3~~. This source factor contributes about 60 - 70% of the total N_2O daily enhancements as seen in Figure 7c. The bootstrapping ~~with an~~ uncertainty ~~of~~ in the N_2O dairy and livestock mass fraction is 33%.

Comparing the Factor 3 profile to dairy source profiles from various studies is challenging. A dairy is, in essence, a collection of area sources with distinct emission pathways and chemical characteristics. Hence, a lot of dairy studies do not look at facility-wide emissions instead focusing on specific area sources within the facility. In contrast, PMF captures the covariance of CH_4, N_2O, and VOCs emitted from the ensemble source as downwind plumes from dairies arrive at the site. Table 4 compares the PMF derived EFs of CH_4 w.r.t MeOH and EtOH with those from other studies. Previously, cow chamber experiments (Shaw et al., 2007; Sun et al., 2008) have measured emissions from ruminants and their fresh manure; emissions have also been studied in a German cowshed (Ngwabie et al., 2008) and EFs have been derived from SJV dairy plumes sampled from aircraft~~s~~ (Gentner et al., 2014a; Guha et al., in prep). Since enteric fermentation and waste manure is the predominant CH_4 source in dairies, CH_4 emission rates calculated by Shaw et al. (2007) are representative of a whole facility. However, their MeOH/CH_4 ratios are lower than those ~~measured~~ determined by PMF and aircraft studies. Animal feed and silage are the dominant sources of many VOCs including MeOH and EtOH (Alanis et al., 2010; Howard et al., 2010) and the ratios in (Shaw et al., 2007) do not reflect these emissions. In ~~(~~Ngwabie et al. (~~,~~2008), experiments were performed in cold winter conditions (-2 to 8°C) when temperature dependent VOC emissions from silage and feed are at a minimum. The authors comment that MeOH emissions from California dairies ~~is~~are likely higher, as the alfalfa-based feed is a big source of MeOH owing to its high pectin content (Galbally and Kirstine, 2002). These observations explain why MeOH/CH_4 ratios in these studies are lower than PMF derived ratios. The PMF range for EtOH /CH_4 EF for Factor 3 agrees with the slope derived from ground-site data (Gentner et al., 2014a) and is similar, but somewhat larger than the German dairy study (Ngwabie et al., 2008) . Miller and Varel (2001) and Filipy et al. (2006) did not measure CH_4 emission rates so a direct derivation of EF w.r.t CH_4 is not possible. These studies, however, reported EtOH emission rates (from dairies and feedlots in United States) which are used to derive EFs w.r.t to CH_4 using an averaged CH_4 emission rate from (Shaw et al., 2007). Using this method, we get EFs that are comparable to PMF derived EF of CH_4/EtOH (Table 4). Hence, we demonstrate within reasonable terms that the relative fractions of masses in Factor 3 are consistent with CH_4 and VOC emissions from dairies.

Enteric fermentation is a part of the normal digestive process of livestock such as cows, and is a large source of CH_4 while the storage and management of animal manure in lagoons or holding tanks is also a major source of CH_4. According to the state GHG inventory (CARB, 2013), ~ 58% of the statewide CH_4 emissions results from a combination of these two processes. N_2O is also emitted during the breakdown of nitrogen in livestock manure and urine and accounts for about 10% of the statewide N_2O emission inventory. Kern County has a big dairy industry with about 160,000 milk cows representing 10 % of the dairy livestock of the state in 2012 and another 330,000 heads of cattle for beef (KernAg, 2011; CASR, 2013). The dominant contributions to CH_4 and N_2O signal and the general agreement of dairy EFs with PMF EFs from Factor 3 indicate that the extensive cattle operations in the county are a big source of these emissions. We do observe that the proportion of regional N_2O enhancements attributed to this sector is a significantly larger proportion of the total N_2O emissions as compared to the state inventory.

4.5. Factor 4: Agricultural and soil management emissions

The chemical profile of Factor 4 is a mix of emissions from agricultural activities around the site. Factor 4 includes a major portion of the N_2O signal along with a number of VOCs that have crop/plant signatures like methacrolein, methyl ethyl ketone (Jordan et al., 2009; McKinney et al., 2011), methanol and acetone (Goldstein and Schade, 2000; Hu et al., 2013; Gentner et al., 2014b) (Figure 4). While many of these oxygenated VOCs have several prominent sources, studies have reported substantial simultaneous emissions from natural vegetation and agricultural crops. At a rural site in the Northeast, Jordan et al. (2009) reported high concentrations of oxygenated VOCs and correlations between the diurnal concentrations of acetone, methanol, and methyl ethyl ketone. Kern County is one of the most prolific agricultural counties in California. The four main crops grown (by value as well as acreage) in 2010 were almonds, grapes, citrus and pistachios (KernAg, 2011). Table 5 compares the PMF derived EFs for acetone/MeOH from Factor 4 with ratios of basal emission factors (BEFs) from crop-specific greenhouse and field measurements (Fares et al., 2011, 2012; Gentner et al., 2014b). The good agreement of the ratios confirms that the FP of Factor 4 is an aggregate of biogenic VOC emissions from the agricultural sector. Nitrous oxide is emitted when nitrogen is added to soil through use of synthetic fertilizers and animal manure, while crops and plants are responsible for the VOC emissions. Hence this source factor is a combination of collocated sources (soils and crops). The PMF solution to this factor has uncertainties greater than those for other factors (Figure S4). This is potentially because not all crops emit the same combination of VOCs nor are all agricultural fields fertilized at the same time. The existence of this statistically weak factor is confirmed by bootstrapping runs (Section S3) and numerous PMF trials all of which produce a distinct factor with N_2O as a dominant contributor along with certain biogenic VOCs, though often in varying proportions. CO_2 is not included in the PMF analysis reported in the paper, most importantly because negative CO_2 fluxes during daytime can introduce artifacts in PMF analysis and result in erroneous apportionment.~~,~~ B~~b~~ut ~~-~~PMF runs involving CO_2 indicate that most of the CO_2 is apportioned to this factor. Plant and soil respiration (especially during the night) is a major source of CO_2 and the apportionment of CO_2 to Factor 4 confirms the nature of this source. The temporal correlation between CO_2 and N_2O is also evident in their average diurnal cycles (Figures 6f and 6h), which have a coincident early morning peak. The absence of monoterpenes from the FP of this factor can be explained by their shorter atmospheric lifetimes compared to VOCs like acetone and MeOH and the rapid daytime mixing which dilutes the terpenoid emissions arriving at the site during the day. At night~~s~~, when ~~the~~ atmospheric dilution ~~has been reduced to a~~is low, monoterpenes emissions from agriculture are more likely to get apportioned into a separate source factor dominant during nighttime, when temperature-sensitive biogenic emissions of MeOH and acetone can be expected to be a minor constituent in the FP (see Section 4.8).

Factor 4 is a significant source of GHGs contributing about 20 - 25% of the total N_2O enhancements in the diurnal cycle (Fig. 7c) but with a relatively large 1σ confidence interval of 70% in the agriculture and soil N_2O mass fraction.~~Factor 4 is a significant source of GHGs contributing about 20-25 % of the total N_2O enhancements in the diurnal cycle (Figure 7c) with a relatively larger uncertainty of 70 %.~~ Kern County is one of the premier agricultural counties of California accounting for \$4.2 billion (about 18%) of the total agricultural revenue from fruits and nuts, vegetables and field crops (KernAg, 2011; CASR, 2013) and is also the biggest consumer of synthetic fertilizers. Agricultural soil management accounts for about 60% of the statewide N_2O emission inventory (CARB, 2013). Our assessment of diurnal source distribution of N_2O emissions from the agriculture source factor (Figure 7c) in presence of another dominating source (dairy and livestock) is consistent with the inventory estimates from agricultural and soil management sector.

4.6. Factor 5: Daytime biogenics and secondary organics

The chemical composition and diurnal profile of Factor 5 points to a source whose emissions are either primary biogenic VOCs with temperature-dependent emissions (e.g. isoprene), or products of photochemical oxidation of primary VOCs (e.g. acetone) (Figure 4). Isoprene is a dominant component of the source FP and is mostly apportioned to Factor 5. Figure 5e~~c~~ shows a steady increase in the PMF factor mass concentration during the daytime hours that hits a peak during afternoons indicating that this source is dependent on sunlight and temperature. Potential source contributions come from oak forests on the foothills of the western edge of the SJV or scattered isoprene producing plants in the SJV (note that most crops do not emit significant amounts of isoprene). Factor 5 includes contribution from VOCs that have primary light and temperature driven (crops), as well as secondary sources in the Central Valley e.g. acetone (Goldstein and Schade, 2000), methanol (Gentner et al., 2014b) and aldehydes. A similar PMF analysis with a different objective (Goldstein et al., in prep) shows that secondary organics like glyoxal, formaldehyde and formic acid mostly apportion to Factor 5. The CO apportioned to this factor could potentially be a product of mobile and/or stationary combustion co-located or up/downwind of the biogenic VOC source. CO can also come from coincident isoprene oxidation (Hudman et al., 2008). This daytime source is not responsible for any of the observed CH_4 and N_2O enhancements.

4.7. Factor 6: Non-vehicular/miscellaneous urban emissions

The chemical signature of Factor 6 is composed of VOCs associated with an array of applications and processes, including solvents, fumigants, industrial-byproducts, etc. The diurnal profile of Factor 6 (Figure 5f~~e~~) is somewhat different from that of evaporative and fugitive source (Figure 5a) and dairies (Figure 5c) in that even during the middle of the day when vertical mixing is at its strongest, the enhancements contributing to the factor are substantial. This suggests that the source(s) is in close proximity to the site and hence most likely located within the urban core. The FP has CO as an important component but relative absence of fugitive source markers (e.g. light alkanes) and vehicle emissions tracers (e.g. isooctane, cycloalkanes etc.) indicate that the origin of this source factor is potentially non-mobile combustion. Also present in a major proportion is carbon disulfide (CS_2), chlorinated alkanes like 1,2-dichloroethane and 1,2-dichlorpropane, isobutene (product of incomplete combustion), and minor contributions from aromatics and aldehydes (Figure 4). There are~~is~~ a myriad of potential sources that could be contributing to this factor, and we don't have specific tracers or other information to ascribe it to a single source or group of sources. Hence we call Factor 6 an ‘urban emissions source’. There is a very minor CH_4 contribution from this factor which results in a tiny and negligible contribution to the PMF source apportionment of CH_4 (Figure 7a). The source factor does not contribute to the N_2O enhancements.

4.8. Factor 7: Nighttime anthropogenic and terpene biogenic emissions

Factor 7 is primarily composed of biogenic compounds belonging to the terpene family and p-cymene (Figure 4). Factor 7 mostly influences the site during late night and early morning hours (Figure 5g) when nighttime downslope flows usually dominate bringing winds from the east and south to the site. The entire flow path from the base of the foothills to the site is covered with agricultural crops emitting into a shallow nighttime boundary layer. These crops include grapes, almonds, citrus and pistachios, which are the top four agricultural commodities grown in the county (KernAg, 2011; CASR, 2013), and these produce considerable monoterpenoid emissions (Fares et al., 2012; Gentner et al., 2014b). The spatial distribution of terpenoid compounds from statistical source footprint derived from FLEXPART back-trajectories is consistent with the location of croplands in southern SJV (Gentner et al., 2014b). Biogenic VOCs emitted from forests in the foothills are likely minor contributors to the downslope flows arriving at the site owing to their lifetime and distance (> 50 km) (Tanner and Zielinska, 1994).

Following the rapid rise in enhancements in the early morning hours, contributions of Factor 7 to total signal decrease rapidly when the flow moves to more typical daytime wind directions (Figure 5g). A nearby source (e.g. the WWTP), that is ~~in the~~ upwind ~~direction~~ of the site for only a certain part of the diurnal cycle, is expected to be more directionally constrained and emissions profile from such a source will look similar to the diurnal profile of Factor 7. Among source factors which contain non-negligible fractional contribution of both CH_4 and N_2O (i.e. dairies, agriculture and soil management, and Factor 7), the PMF derived CH_4/N_2O EF of 42 ± 20 (gC gN^{-1})~~(gC gC^{-1})~~ from Factor 7 is most similar to the bottom-up inventory EF of 56 (gC gN^{-1}) ~~(gC gC^{-1})~~ for waste water treatment in Kern County (KernGHG, 2012). Given the proximity of the WWTP and previous observations of GHGs from them, it is possible that there is a minor but noticeable contribution (~ 5%) to CH_4 and N_2O enhancements from this nighttime source (Figures 7a and 7c).

5. Implications

This study demonstrates the potential of ~~the~~ PMF technique to apportion atmospheric gas-phase observations of CH_4 and N_2O into source categories using a broad array of tracers. PMF is not commonly employed to perform for source attribution of these GHGs because studies generally lack simultaneous measurements of specific source-markers. Applying this statistical technique on a GHG-VOC unified data set, well-represented by a broad suite of VOC classes, allows a set of compounds acting as source markers to be partitioned into separate profiles leading to easier identification of their sources.

We provide clear analysis that dairy and livestock operations are the largest sources of emissions in the Bakersfield region accounting for a majority of the CH_4 (70 - 90%) and N_2O (~~50-60~~ 60 - 70%) emissions. As per the CARB inventory (Figure~~s~~ 1)~~a and 1b)~~, dairy operations are the dominant source of non-CO_2 GHGs in the state and our analysis agrees with that broad trend. However, in the recent past, a number of top-down CH_4 and N_2O emission studies in the Central Valley have reported underestimation of the non-CO_2 GHG inventory (Zhao et al., 2009; Santoni et al., 2012; Jeong et al., 2012a, 2012b; Miller et al., 2013). These studies attribute a majority of this underestimation to the dairy sector. Our results emphasize the significance of this sector in the SJV although we do not derive total emission estimates to compare directly with the inventory.

The contribution of fugitive emissions from the oil and gas industry in Bakersfield to CH_4 emissions is found to be negligible especially in the presence of the much larger dairy source. The PMF analysis, though, clearly establishes an evaporative and fugitive source that contributes to emissions of lighter hydrocarbons. This supports the conclusion that the majority of the CH_4 is being separated at the point of extraction from the 'associated gas' and is not released with fugitive emissions (Gentner et al., 2014a). Kern County produces 75% of all the oil produced in California (~ 6% of US production) and has 81% of the state's 60000+ active oil wells (CDC, 2013). There is, however, a surprising scarcity of measured data to quantify the estimates of fugitive CH_4 from the prolific oil fields in the county and validate the bottom-up, activity data-based inventory. Currently, fugitive emissions from fossil fuel extraction and distribution contribute ~ 5% to the county's CH_4 emissions inventory (KernGHG, 2012). Nationwide, a number of recent studies have reported significantly higher emissions of fugitive CH_4 from oil and gas production operations in other regions (Pétron et al., 2012; Karion et al., 2013; Miller et al., 2013; Kort et al., 2014) ~~(Karion et al., 2013; Miller et al., 2013; Pétron et al., 2012)~~. The PMF apportionment in this study (~ 2%) is consistent with the fraction of fugitive CH_4 emissions in the regional and state inventories~~y~~ ~~(< 5 %)~~ but the PMF method, by itself, cannot accurately constrain emissions from minor sources owing to the uncertainties in the dominant sources~~in itself is limited in accurate partitioning of minor sources~~.

We find that the vehicle emissions source factor identified in this study makes no detectable contribution to observed N_2O enhancements. Our findings do not agree with the significant contribution (~ 18%) of the transportation sector to the state's N_2O emission inventory (CARB, 2013). Vehicle dynamometer studies have indicated rapidly declining N_2O EFs with advancement in catalyst technologies, declining sulfur content in fuel and newer technology vehicles (Huai et al., 2004). N_2O emissions from California vehicles, required to meet progressively stringent emission standards, are expected to decline and should have a minimal contribution to the CARB inventory in this decade. However, it seems the updates to the mobile N_2O emissions inventory is not keeping in pace with the improvements in vehicle catalyst technologies and corresponding decline in tailpipe N_2O emissions. Bakersfield is a fairly large population urban region (~ 500,000) and the essentially non-existent contribution of the PMF vehicle emissions source to the N_2O apportionment and large divergence of the PMF derived N_2O/CO EF from the state inventory EF for motor vehicles is a significant outcome pointing to overestimation of N_2O from motor vehicles in the inventory.

Acknowledgements

This work was supported by a contract from the California Air Resources Board (CARB) # 08-316 and in kind support from Los Gatos Research (LGR) Inc. We thank Dr. John Karlik, the staff at UC

Cooperative Extension at Bakersfield, Aaron Gardner and Dr. Doug Baer at LGR Inc., and Dr. Ying Hsu at CARB for their logistical support. We also appreciate the subject matter guidance from Dr. Marc Fischer at Lawrence Berkeley National Laboratory.

References

Alanis, P., Ashkan, S., Krauter, C., Campbell, S. and Hasson, A. S.: Emissions of volatile fatty acids from feed at dairy facilities, Atmos. Environ., 44(39), 5084–5092, doi:10.1016/j.atmosenv.2010.09.017, 2010.

Bao, J.-W., Michelson, S. a., Persson, P. O. G., Djalalova, I. V. and Wilczak, J. M.: Observed and WRF-Simulated Low-Level Winds in a High-Ozone Episode during the Central California Ozone Study, J. Appl. Meteorol. Climatol., 47(9), 2372–2394, doi:10.1175/2008JAMC1822.1, 2007.

Beaver, S. and Palazoglu, A.: Influence of synoptic and mesoscale meteorology on ozone pollution potential for San Joaquin Valley of California, Atmos. Environ., 43(10), 1779–1788, doi:10.1016/j.atmosenv.2008.12.034, 2009.

Bon, D. M., Ulbrich, I. M., de Gouw, J. a., Warneke, C., Kuster, W. C., Alexander, M. L., Baker, a., Beyersdorf, a. J., Blake, D., Fall, R., Jimenez, J. L., Herndon, S. C., Huey, L. G., Knighton, W. B., Ortega, J., Springston, S. and Vargas, O.: Measurements of volatile organic compounds at a suburban ground site (T1) in Mexico City during the MILAGRO 2006 campaign: measurement comparison, emission ratios, and source attribution, Atmos. Chem. Phys., 11(6), 2399–2421, doi:10.5194/acp-11-2399-2011, 2011.

Borbon, A., Gilman, J. B., Kuster, W. C., Grand, N., Chevaillier, S., Colomb, a., Dolgorouky, C., Gros, V., Lopez, M., Sarda-Esteve, R., Holloway, J., Stutz, J., Petetin, H., McKeen, S., Beekmann, M., Warneke, C., Parrish, D. D. and de Gouw, J. a.: Emission ratios of anthropogenic volatile organic compounds in northern mid-latitude megacities: Observations versus emission inventories in Los Angeles and Paris, J. Geophys. Res. Atmos., 118(4), 2041–2057, doi:10.1002/jgrd.50059, 2013.

Brown, S. G., Frankel, A. and Hafner, H. R.: Source apportionment of VOCs in the Los Angeles area using positive matrix factorization, Atmos. Environ., 41(2), 227–237, doi:10.1016/j.atmosenv.2006.08.021, 2007.

Buzcu, B., and M. P. Fraser (2006), Source identification and apportionment of volatile organic compounds in Houston, TX, Atmos. Environ., 40(13), 2385–2400, doi:10.1016/j.atmosenv.2005.12.020.

CARB (2013), California Greenhouse Gas Inventory for 2000-2012 - by IPCC Category. http://www.arb.ca.gov/cc/inventory/data/data.htm. accessesd on September 23, 2014.

CASR (2011), California Agricultural Statistics Review, 2013-14. California Department of Food and Agriculture, Sacramento, CA.

CDC (2013), 2012 Preliminary Report Of California Oil And Gas Production Statistics. California Department of Conservation, Division of Oil, Gas and Geothermal Resources. ftp://ftp.consrv.ca.gov/pub/oil/annual_reports/2012/PR03_PreAnnual_2012.pdf. accessed on September 23, 2014.

Chung, M. Y., M. Beene, S. Ashkan, C. Krauter, and A. S. Hasson (2010), Evaluation of non-enteric sources of non-methane volatile organic compound (NMVOC) emissions from dairies, Atmos. Environ., 44(6), 786–794, doi:10.1016/j.atmosenv.2009.11.033.

Comero, S., Capitani, L., and Gawlik, B. M.: Positive Matrix Factorization - An introduction to the chemometric evaluation of environmental monitoring data using PMF, JRC Scientific and Technical Reports, EUR 23946 EN-2009.

Crutzen, P. J., Mosier, a. R., Smith, K. a. and Winiwarter, W.: N_2O release from agro-biofuel production negates global warming reduction by replacing fossil fuels, Atmos. Chem. Phys. Discuss., 7(4), 11191–11205, doi:10.5194/acpd-7-11191-2007, 2007.

Denman, K.L., G. Brasseur, A. Chidthaisong, P. Ciais, P.M. Cox, R.E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. Jacob, U. Lohmann, S Ramachandran, P.L. da Silva Dias, S.C. Wofsy and X. Zhang, 2007: Couplings Between Changes in the Climate System and Biogeochemistry. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Dlugokencky, E.J., A.M. Crotwell, P.M. Lang, K.A. Masarie (2014), Atmospheric Methane Dry Air Mole Fractions from quasi-continuous measurements at Barrow, Alaska and Mauna Loa, Hawaii, 1986-2013, Version: 2014-08-12, Path: ftp://ftp.cmdl.noaa.gov/ccg/ch4/in-situ/.

EMFAC (2011). 2011 Mobile Source Emission Inventory - Current methods and data. California Air Resources Board, Sacramento, CA. http://www.arb.ca.gov/msei/modeling.htm. accessed on September 20, 2014.

Fares, S., Gentner, D. R., Park, J.-H., Ormeno, E., Karlik, J. and Goldstein, A. H.: Biogenic emissions from Citrus species in California, Atmos. Environ., 45(27), 4557–4568, doi:10.1016/j.atmosenv.2011.05.066, 2011.

Fares, S., Park, J.-H., Gentner, D. R., Weber, R., Ormeño, E., Karlik, J. and Goldstein, a. H.: Seasonal cycles of biogenic volatile organic compound fluxes and concentrations in a California citrus orchard, Atmos. Chem. Phys., 12(20), 9865–9880, doi:10.5194/acp-12-9865-2012, 2012.

Filipy, J., B. Rumburg, G. Mount, H. Westberg, and B. Lamb (2006), Identification and quantification of volatile organic compounds from a dairy, Atmos. Environ., 40(8), 1480–1494, doi:10.1016/j.atmosenv.2005.10.048.

Fischer, M.L., C. Zhao., W.J. Riley, and A.C. Andrews. 2009. Observation of CH4 and other Non-CO2 Green House Gas Emissions from California. California Energy Commission, PIER Energy-Related Environmental Research. (Report # 500-2009-096).

~~Forster, P. ., Ramaswamy, V. ., Artaxo, P. ., Berntsen, T. ., Betts, R. ., Fahey, D. W. ., Haywood, J. ., Lean, J. ., Lowe, D. C. ., Myhre, G. ., Nganga, J. ., Prinn, R. ., Raga, G. ., Schulz, M. . and Dorland, R. V.: 2007: Changes in Atmospheric Constituents and in Radiative Forcing, in Climate Change 2007: The Physical~~

~~Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M.Tignor, and H. L. Miller, p. 996, Cambridge University Press. [online] Available from: http://www.ipcc.ch/publications_and_data/ar4/wg1/en/ch2.html, 2007.~~

Galbally, I. E., and Kirstine, W.: The Production of Methanol by Flowering Plants and the Global Cycle of Methanol, Journal of Atmospheric Chemistry, 43, 195-229, 10.1023/A:1020684815474, 2002.Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. a, Seitzinger, S. P. and Sutton, M. a: Transformation of the nitrogen cycle: recent trends, questions, and potential solutions., Science, 320(5878), 889–92, doi:10.1126/science.1136674, 2008.

Galloway, J. N., A. R. Townsend, J. W. Erisman, M. Bekunda, Z. Cai, J. R. Freney, L. a Martinelli, S. P. Seitzinger, and M. a Sutton (2008), Transformation of the nitrogen cycle: recent trends, questions, and potential solutions., Science, 320(5878), 889–92, doi:10.1126/science.1136674.

Gentner, D. R., Ford, T. B., Guha, a., Boulanger, K., Brioude, J., Angevine, W. M., de Gouw, J. a., Warneke, C., Gilman, J. B., Ryerson, T. B., Peischl, J., Meinardi, S., Blake, D. R., Atlas, E., Lonneman, W. a., Kleindienst, T. E., Beaver, M. R., Clair, J. M. St., Wennberg, P. O., VandenBoer, T. C., Markovic, M. Z., Murphy, J. G., Harley, R. a. and Goldstein, a. H.: Emissions of organic carbon and methane from petroleum and dairy operations in California's San Joaquin Valley, Atmos. Chem. Phys., 14(10), 4955–4978, doi:10.5194/acp-14-4955-2014, 2014a.

Gentner, D. R., Harley, R. a, Miller, A. M. and Goldstein, A. H.: Diurnal and seasonal variability of gasoline-related volatile organic compound emissions in Riverside, California., Environ. Sci. Technol., 43(12), 4247–52 [online] Available from: http://www.ncbi.nlm.nih.gov/pubmed/19603630, 2009.

Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W. H., Dallmann, T. R., Davis, L., Liu, S., Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A. and Harley, R. A.: Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions, , doi:10.1073/pnas.1212272109/~~-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1212272109~~, 2012.

Gentner, D. R., Ormeño, E., Fares, S., Ford, T. B., Weber, R., Park, J.-H., Brioude, J., Angevine, W. M., Karlik, J. F. and Goldstein, a. H.: Emissions of terpenoids, benzenoids, and other biogenic gas-phase organic compounds from agricultural crops and their potential implications for air quality, Atmos. Chem. Phys., 14(11), 5393–5413, doi:10.5194/acp-14-5393-2014, 2014b.

Gilman, J. B., B. M. Lerner, W. C. Kuster, and J. a de Gouw (2013), Source signature of volatile organic compounds from oil and natural gas operations in northeastern colorado., Environ. Sci. Technol., 47(3), 1297–305, doi:10.1021/es304119a.

Goldstein, A. H. and Schade, G. W.: Quantifying biogenic and anthropogenic contributions to acetone mixing ratios in a rural environment, Atmos. Environ., 34(29-30), 4997–5006, doi:10.1016/S1352-2310(00)00321-6, 2000.

Goldstein, A. H. et al.: Source apportionment of OH reactivity at Bakersfield Positive Matrix Factorization (PMF) during CalNex 2010, *in preparation*.

Guha, A., Misztal, P. K., Karl, T., Peischl, J., Ryerson, T. B., , Jonsson, H.H., and Goldstein, A.H.: Identifying and mapping methane (CH_4) sources over California from airborne mixing ratio, eddy covariance flux and VOC source tracer measurements, in preparation, 2015.

~~Guha et al.: Identifying and mapping methane (CH_4) sources over California from mixing ratio, airborne flux and VOC source tracer measurements, *in preparation*.~~

Hendriks, D. M. D., a. J. Dolman, M. K. van der Molen, and J. van Huissteden (2008), A compact and stable eddy covariance set-up for methane measurements using off-axis integrated cavity output spectroscopy, Atmos. Chem. Phys., 8(2), 431–443, doi:10.5194/acp-8-431-2008.

Hopke, P.: A guide to positive matrix factorization, Department of Chemistry, Calrkson University, Potsdam, NY pp 1–16, http://www.epa.gov/ttnamti1/files/ambient/pm25/workshop/laymen.pdf., 2000

~~Hopke, P.: A guide to positive matrix factorization, Work. UNMIX PMF as Appl. to PM2, 1–16 [online] Available from: ftp://128.153.5.141/users/h/o/hopkepk/IAEA/PMF-Guidance.pdf (Accessed 13 March 2013), 2000.~~

Howard, C. J., Kumar, A., Malkina, I., Mitloehner, F., Green, P. G., Flocchini, R. G. and Kleeman, M. J.: Reactive organic gas emissions from livestock feed contribute significantly to ozone production in central California., Environ. Sci. Technol., 44(7), 2309–14, doi:10.1021/es902864u, 2010.

Hsu, Y.-K., VanCuren, T., Park, S., Jakober, C., Herner, J., FitzGibbon, M., Blake, D. R. and Parrish, D. D.: Methane emissions inventory verification in southern California, Atmos. Environ., 44(1), 1–7, doi:10.1016/j.atmosenv.2009.10.002, 2010.

Hu, L., D. B. Millet, S. Y. Kim, K. C. Wells, T. J. Griffis, E. V. Fischer, D. Helmig, J. Hueber, and a. J. Curtis (2013), North American acetone sources determined from tall tower measurements and inverse modeling, Atmos. Chem. Phys., 13(6), 3379–3392, doi:10.5194/acp-13-3379-2013.

Huai, T., Durbin, T. D., Wayne Miller, J. and Norbeck, J. M.: Estimates of the emission rates of nitrous oxide from light-duty vehicles using different chassis dynamometer test cycles, Atmos. Environ., 38(38), 6621–6629, doi:10.1016/j.atmosenv.2004.07.007, 2004.

Hudman, R. C., Murray, L. T., Jacob, D. J., Millet, D. B., Turquety, S., Wu, S., Blake, D. R., Goldstein, a. H., Holloway, J. and Sachse, G. W.: Biogenic versus anthropogenic sources of CO in the United States, Geophys. Res. Lett., 35(4), L04801, doi:10.1029/2007GL032393, 2008.

IPCC, 2013: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jeong, S., C. Zhao, A. E. Andrews, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012a), Seasonal variation of CH 4 emissions from central California, J. Geophys. Res., 117(D11), D11306, doi:10.1029/2011JD016896.

Jeong, S., C. Zhao, A. E. Andrews, E. J. Dlugokencky, C. Sweeney, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012b), Seasonal variations in N 2 O emissions from central California, Geophys. Res. Lett., 39(16), n/a–n/a, doi:10.1029/2012GL052307.

Jordan, C., Fitz, E., Hagan, T., Sive, B., Frinak, E., Haase, K., Cottrell, L., Buckley, S., and Talbot, R.: Long-term study of VOCs measured with PTR-MS at a rural site in New Hampshire with urban influences, Atmos. Chem. Phys., 9, 4677-4697, doi:10.5194/acp-9-4677-2009, 2009.

~~Jordan, C., Fitz, E., Hagan, T., Sive, B., Frinak, E., Haase, K., Cottrell, L., Buckley, S. and Talbot, R.: and Physics Long-term study of VOCs measured with PTR-MS at a rural site in New Hampshire with urban influences, , (Table 1), 4677–4697, 2009.~~

KernGHG (2012), Kern County Communitywide Greenhouse Gas Emission Inventory, Methodology Documents Volume 2; Final Report - May 2012. Prepared by San Joaquin Air Pollution Control District for Kern County Planning and Community Development. http://www.kerncog.org/climate-change.

KernAg (2010). 2010 Kern County Agricultural Crop Report. Department of Agriculture and Measurement Standards. http://www.kernag.com/caap/crop-reports/crop10_19/crop2010.pdf. accessed on September 22 2014.

Karion, A., Sweeney, C., Pétron, G., Frost, G., Michael Hardesty, R., Kofler, J., Miller, B. R., Newberger, T., Wolter, S., Banta, R., Brewer, A., Dlugokencky, E., Lang, P., Montzka, S. a., Schnell, R., Tans, P., Trainer, M., Zamora, R. and Conley, S.: Methane emissions estimate from airborne measurements over a western United States natural gas field, Geophys. Res. Lett., 40(16), 4393–4397, doi:10.1002/grl.50811, 2013.

Kim, E., Hopke, P. K. and Edgerton, E. S.: Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization, Atmos. Environ., 38(20), 3349–3362, doi:10.1016/j.atmosenv.2004.03.012, 2004.

Kort, E. A., C. Frankenberg, K. R. Costigan, R. Lindenmaier, M. K. Dubey, and D. Wunch (2014), Four corners: The largest US methane anomaly viewed from space, Geophys. Res. Lett., 41, 6898–6903, doi:10.1002/2014GL061503.

Lee, E., Chan, C. K. and Paatero, P.: Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong, Atmos. Environ., 33(19), 3201–3212, doi:10.1016/S1352-2310(99)00113-2, 1999.

Leuchner, M., and B. Rappenglück (2010), VOC source–receptor relationships in Houston during TexAQS-II, Atmos. Environ., 44(33), 4056–4067, doi:10.1016/j.atmosenv.2009.02.029.

McKinney, K. a., Lee, B. H., Vasta, a., Pho, T. V. and Munger, J. W.: Emissions of isoprenoids and oxygenated biogenic volatile organic compounds from a New England mixed forest, Atmos. Chem. Phys., 11(10), 4807–4831, doi:10.5194/acp-11-4807-2011, 2011.

McMillan, A. M. S., Goulden, M. L. and Tyler, S. C.: Stoichiometry of CH$_4$ and CO$_2$ flux in a California rice paddy, J. Geophys. Res., 112(G1), G01008, doi:10.1029/2006JG000198, 2007.

Miller, D. N., and Varel, V. H.: In vitro study of the biochemical origin and production limits of odorous compounds in cattle feedlots, Journal of Animal Science, 79, 2949-2956, 2001.

Miller, S. M., Wofsy, S. C., Michalak, A. M., Kort, E. a, Andrews, A. E., Biraud, S. C., Dlugokencky, E. J., Eluszkiewicz, J., Fischer, M. L., Janssens-Maenhout, G., Miller, B. R., Miller, J. B., Montzka, S. a, Nehrkorn, T. and Sweeney, C.: Anthropogenic emissions of methane in the United States., Proc. Natl. Acad. Sci. U. S. A., 110(50), 20018–22, doi:10.1073/pnas.1314392110, 2013.

Montzka, S. a, Dlugokencky, E. J. and Butler, J. H.: Non-CO2 greenhouse gases and climate change., Nature, 476(7358), 43–50, doi:10.1038/nature10322, 2011.

Myhre, G., D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura and H. Zhang, 2013: Anthropogenic and Natural Radiative Forcing. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York,NY, USA, pp. 659–740, doi:10.1017/ CBO9781107415324.018.

Nam, E., Jensen, T. and Wallington, T.: Methane emissions from vehicles, Environ. Sci. …, 38(7), 2005–2010 [online] Available from: http://pubs.acs.org/doi/abs/10.1021/es034837g (Accessed 16 April 2013), 2004.

NRC: National Research Council Report - Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements, The National Academies Press, Washington, DC, 124 pp., 2010.

Norris, G., Vedantham, R., Wade, K., Brown, S., Prouty, J., and Foley, C.: EPA Positive Matrix Factorization (PMF) 3.0 Fundamentals and User Guide, Washington DC, USA, 2008.

Ngwabie, N. M., Schade, G. W., Custer, T. G., Linke, S. and Hinz, T.: Abundances and Flux Estimates of Volatile Organic Compounds from a Dairy Cowshed in Germany, J. Environ. Qual., 37(2), 565, doi:10.2134/jeq2006.0417, 2008.

O'Keefe, A.: 1998: integrated cavity output analysis of ultra-weak absorption, Chem. Phys. Lett.,293, 331–336, doi:10.1016/S0009-2614(98)00785-4, 1998.

Owen, J. J. and Silver, W. L.: Greenhouse gas emissions from dairy manure management: a review of field-based studies., Glob. Chang. Biol., doi:10.1111/gcb.12687, 2014.

Paatero, P. (1997), Least squares formulation of robust non-negative factor analysis, Chemom. Intell. Lab. Syst., 37(1), 23–35, doi:10.1016/S0169-7439(96)00044-5.

Paatero, P., and U. Tapper (1994), Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, Environmetrics, 5(April 1993), 111–126.

Parameswaran, K. R., Rosen, D. I., Allen, M. G., Ganz, A. M., and Risby, T. H.: Off-axis integrated cavity output spectroscopy with a mid-infrared interband cascade laser for real-time breath ethane measurements., Appl. Optics, 48, B73–79, doi: 10.1364/AO.48.000B73, 2009.

~~Parameswaran, K. R., D. I. Rosen, M. G. Allen, A. M. Ganz, and T. H. Risby (2009), Off-axis integrated cavity output spectroscopy with a mid-infrared interband cascade laser for real-time breath ethane measurements., Appl. Opt., 48(4), B73–9.~~

Paul, J. B., Scherer, J. J., O'Keefe, A., Lapson, L., Anderson, J. R., Gmachl, C. F., Capasso, F., and Cho, A. Y.: Infrared cavity ringdown and integrated cavity output spectroscopy for trace species monitoring, 2002, 1-11, Proc. SPIE 4577, Vibrational Spectroscopy-based Sensor Systems, 1 (February 13, 2002); doi:10.1117/12.455722

Peischl, J., Ryerson, T. B., Brioude, J., Aikin, K. C., Andrews, a. E., Atlas, E., Blake, D., Daube, B. C., de Gouw, J. a., Dlugokencky, E., Frost, G. J., Gentner, D. R., Gilman, J. B., Goldstein, a. H., Harley, R. a., Holloway, J. S., Kofler, J., Kuster, W. C., Lang, P. M., Novelli, P. C., Santoni, G. W., Trainer, M., Wofsy, S. C. and Parrish, D. D.: Quantifying sources of methane using light alkanes in the Los Angeles basin, California, J. Geophys. Res. Atmos., 118(10), 4974–4990, doi:10.1002/jgrd.50413, 2013.

Peischl, J., Ryerson, T. B., Holloway, J. S., Trainer, M., Andrews, a. E., Atlas, E. L., Blake, D. R., Daube, B. C., Dlugokencky, E. J., Fischer, M. L., Goldstein, a. H., Guha, a., Karl, T., Kofler, J., Kosciuch, E., Misztal, P. K., Perring, a. E., Pollack, I. B., Santoni, G. W., Schwarz, J. P., Spackman, J. R., Wofsy, S. C. and Parrish, D. D.: Airborne observations of methane emissions from rice cultivation in the Sacramento Valley of California, J. Geophys. Res., 117, D00V25, doi:10.1029/2012JD017994, 2012.

Pétron, G., Frost, G., Miller, B. R., Hirsch, A. I., Montzka, S. a., Karion, A., Trainer, M., Sweeney, C., Andrews, A. E., Miller, L., Kofler, J., Bar-Ilan, A., Dlugokencky, E. J., Patrick, L., Moore, C. T., Ryerson, T. B., Siso, C., Kolodzey, W., Lang, P. M., Conway, T., Novelli, P., Masarie, K., Hall, B., Guenther, D., Kitzis, D., Miller, J., Welsh, D., Wolfe, D., Neff, W. and Tans, P.: Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study, J. Geophys. Res., 117(D4), D04304, doi:10.1029/2011JD016360, 2012.

Polissar, A. V.*,* P. K. Hopke*,* P. Paatero*,* W. C. Malm*, and* J. F. Sisler *(*1998*),* Atmospheric aerosol over Alaska: 2. Elemental composition and sources*,* J. Geophys. Res.*,* 103*(*D15*),* 19045–19057*, doi:*10.1029/98JD01212*.*

Ryerson, T. B., Andrews, a. E., Angevine, W. M., Bates, T. S., Brock, C. a., Cairns, B., Cohen, R. C., Cooper, O. R., de Gouw, J. a., Fehsenfeld, F. C., Ferrare, R. a., Fischer, M. L., Flagan, R. C., Goldstein, a. H., Hair, J. W., Hardesty, R. M., Hostetler, C. a., Jimenez, J. L., Langford, a. O., McCauley, E., McKeen, S. a., Molina, L. T., Nenes, a., Oltmans, S. J., Parrish, D. D., Pederson, J. R., Pierce, R. B., Prather, K., Quinn, P. K., Seinfeld, J. H., Senff, C. J., Sorooshian, a., Stutz, J., Surratt, J. D., Trainer, M., Volkamer, R., Williams, E. J. and Wofsy, S. C.: The 2010 California Research at the Nexus of Air Quality and Climate Change (CalNex) field study, J. Geophys. Res. Atmos., 118(11), 5830–5866, doi:10.1002/jgrd.50331, 2013.

Santoni, G. W.; Xiang, B.; Kort, E. A.; Daube, B.; Andrews, A. E.; Sweeney, C.; Wecht, K.; Peischl, J.; Ryerson, T. B.; Angevine, W. M.; Trainer, M.; Nehrkorn, T.; Eluszkiewicz, J.; Wofsy, S. C. (2012) California's Methane Budget derived from CalNex P-3 Aircraft Observations and the WRF-STILT Lagrangian Transport Model. AGU Fall Meeting 2012, San Francisco, CA.

Shaw, S. L., Mitloehner, F. M., Jackson, W., Depeters, E. J., Fadel, J. G., Robinson, P. H., Holzinger, R., and Goldstein, A. H.: Volatile organic compound emissions from dairy cows 5 and their waste as measured by proton-transfer-reaction mass spectrometry., Environ. Sci.Technol., 41, 1310–1316, doi: 10.1021/es061475e, 2007.

~~Shaw, S. L., Mitloehner, F. M., Jackson, W., Depeters, E. J., Fadel, J. G., Robinson, P. H., Holzinger, R. and Goldstein, A. H.: Volatile organic compound emissions from dairy cows and their waste as measured by proton-transfer-reaction mass spectrometry., Environ. Sci. Technol., 41(4), 1310–6 [online] Available from: http://www.ncbi.nlm.nih.gov/pubmed/17593735, 2007.~~

Slowik, J. G., Vlasenko, a., McGuire, M., Evans, G. J. and Abbatt, J. P. D.: Simultaneous factor analysis of organic particle and gas mass spectra: AMS and PTR-MS measurements at an urban site, Atmos. Chem. Phys., 10(4), 1969–1988, doi:10.5194/acp-10-1969-2010, 2010.

Smith, P., D. Martino, Z. Cai, D. Gwary, H. Janzen, P. Kumar, B. McCarl, S. Ogle, F. O'Mara, C. Rice, B. Scholes, O. Sirotenko, 2007: Agriculture. In Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Sun, H., Trabue, S. L., Scoggin, K., Jackson, W. a., Pan, Y., Zhao, Y., Malkina, I. L., Koziel, J. a. and Mitloehner, F. M.: Alcohol, Volatile Fatty Acid, Phenol, and Methane Emissions from Dairy Cows and Fresh Manure, J. Environ. Qual., 37(2), 615, doi:10.2134/jeq2007.0357, 2008.

Tanner, R. L. and Zielinska, B.: Determination of the biogenic emission rates of species contributing to VOC in the San Joaquin Valley OF California, Atmos. Environ., 28(6), 1113–1120, doi:10.1016/1352-2310(94)90288-7, 1994.

Ulbrich, I. M. et al. (2009), Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9(9), 2891–2918, doi:10.5194/acp-9-2891-2009.

Williams, B. J., Goldstein, a. H., Kreisberg, N. M., Hering, S. V., Worsnop, D. R., Ulbrich, I. M., Docherty, K. S. and Jimenez, J. L.: Major components of atmospheric organic aerosol in southern California as determined by hourly measurements of source marker compounds, Atmos. Chem. Phys., 10(23), 11577–11603, doi:10.5194/acp-10-11577-2010, 2010.

Wunch, D., Wennberg, P. O., Toon, G. C., Keppel-Aleks, G. and Yavin, Y. G.: Emissions of greenhouse gases from a North American megacity, Geophys. Res. Lett., 36(15), L15810, doi:10.1029/2009GL039825, 2009.

Zhao, C., Andrews, A. E., Bianco, L., Eluszkiewicz, J., Hirsch, A., MacDonald, C., Nehrkorn, T. and Fischer, M. L.: Atmospheric inverse estimates of methane emissions from Central California, J. Geophys. Res., 114(D16), D16302, doi:10.1029/2008JD011671, 2009.

Table 1. PMF dataset with total samples (N) and mixing ratio range (in ppt~~v~~).

Class	Compound	N	1st percentile	99th percentile	Background
GHG	$CH_4^{a,c}$	619	1855.0	3400.8	1813.6
	$CO_2^{b,c}$	619	390.8	468.3	390.0
	$N_2O^{a,d}$	490	323.3	339.5	323.2
combustion tracer	$CO^{a,d}$	653	118.9	330.6	102.1
straight chain alkanes	propane	592	580.8	30839.0	455.5
	n-butane	587	96.4	12649.0	73.6
	n-pentane	647	93.2	3805.4	64.4
	n-hexane	647	23.1	960.5	17.2
	dodecane	643	1.56	54.3	0
branched alkanes	isopentane	646	165.4	7490.5	100.4
	2,3-dimethylbutane	650	52.5	1747.7	41.1
	2,5-dimethylhexane	651	2.37	145.8	0
	isooctane	647	16.6	476.9	12.3
	4-ethylheptane	651	1.45	52.6	0
	dimethyl undecane	643	0.46	24.9	0
cyclo alkanes	methylcyclopentane	647	23.3	1329.6	20.3
	methylcyclohexane	649	8.10	813.9	0
	ethylcyclohexane	651	1.78	169.1	0
alkenes	propene	592	34.7	3299.9	28.6
	isobutene	595	16.7	422.1	10.7
aromatics	toluene	647	48.8	1749.5	33.1
	ethylbenzene	647	5.83	282.0	0
	m,p-xylene	647	21.8	1127.1	21.8
	o-xylene	647	4.31	405.0	0
	cumene	640	0.55	22.8	0
	1-ethyl-3,4-methylbenzene	651	2.22	358.6	0
	p-cymene	649	0.84	93.9	0
	indane	647	0.45	27.9	0
	1,3-dimethyl-4-ethylbenzene	635	0.46	23.9	0
	naphthalene	654	0.44	19.9	0
unsaturated aldehyde	methacrolein	573	14.2	337.0	0
alchohol	methanol	429	2636.81	88691.8	1085.2
	ethanol	598	1021.93	65759.8	1021.9
	isopropyl alcohol	583	25.7	2001.0	25.7
ketone	acetone	663	142.9	3505.8	142.9
	methyl ethyl ketone	605	8.55	1111.2	0
	methyl isobutyl ketone	629	2.03	71.9	0
aldehyde	propanal	636	3.68	140.8	0
	butanal	589	1.72	35.1	0
biogenics	isoprene	651	9.70	310.0	0
	alpha-pinene	740	1.67	525.8	0
	d-limonene	641	1.10	357.1	0
	nopinone	614	0.78	89.5	0
	alpha-thujene	591	0.52	23.8	0
	camphene	645	0.72	100.3	0
chloroalkanes	chloroform	647	34.1	209.3	31.6

	tetrachloroethylene	641	3.41	120.9	0
	1,2-dichloroethane	640	20.6	103.8	20.6
	1,2-dichloropropane	627	2.40	28.4	0
sulfides	carbon disulfide	610	7.84	133.7	0
thiol	ethanethiol	491	4.54	685.8	0

[a] parts per billion volume (ppb~~v~~)
[b] parts per million (ppm~~v~~)
[c] measured using LGR Fast Green House Gas Analyzer
[d] measured using LGR N_2O/CO analyzer

Table 2. Comparison of light alkane ratios to propane ($gC\ gC^{-1}$) from PMF fugitive and evaporative factor with those from other PMF studies and oil and gas operations.

~~Study~~	~~Source~~	~~propane~~	~~n-butane~~	~~n-pentane~~	~~n-hexane~~	~~isopentane~~
~~Bakersfield PMF evaporative and fugitive factor[a]~~	~~This study~~	~~1~~	~~0.52 ± 0.02~~	~~0.18 ± 0.01~~	~~0.06 ± 0.01~~	~~0.33 ± 0.02~~
~~Bakersfield petroleum operations source profile[b]~~	~~Gentner et al. (2014)~~	~~1~~	~~0.53 ± 0.1~~	~~0.09 ± 0.02~~	~~0.04 ± 0.01~~	~~0.08 ± 0.02~~
~~Mexico city PMF LPG factor[c]~~	~~Bon et al. (2011)~~	~~1~~	~~0.5 (0.4 - 0.7)~~	~~0.05 (0.04 - 0.07)~~	~~0.02 (0.02 - 0.03)~~	~~0.07 (0.06 - 0.1)~~
~~Wattenberg field BAO, Colorado[d]~~	~~Gilman et al. (2013)~~	~~1~~	~~0.75 ± 1.37~~	~~0.32 ± 0.6~~	~~0.08 ± 0.13~~	~~0.28 ± 0.52~~
~~Wattenberg field BAO, Colorado[e]~~	~~Petron et al. (2012)~~	~~1~~	~~0.58 - 0.65~~	~~0.22 - 0.31~~	~~NA~~	~~0.22 - 0.31~~
~~PMF natural gas and evaporation factor, Houston Ship Channel[g]~~	~~Leuchner and Rappengluck (2010)~~	~~1~~	~~0.33~~	~~0.27~~	~~0.12~~	~~0.37~~
~~PMF natural gas factor, Houston Ship Channel[h]~~	~~Buzcu and Fraser (2006)~~	~~1~~	~~0.67 ± 0.16~~	~~0.07 ± 0.18~~	~~NA~~	~~NA~~

~~[a] Uncertainties calculated from propagation of errors (standard deviations) over FPEAK range of -1.6 to 0.4.~~
~~[b] Ratios calculated from Table 4, Gentner et al., 2014; uncertainties defined as ±20% to account for variability in oil well data.~~
~~[c] Uncertainties calculated from propagation of uncertainties over FPEAK range of -3 to 3.~~
~~[d] Emission ratios derived from multivariate regression analysis; error bars derived from propogation of uncertainty using mean and standard deviation of samples.~~
~~[e] Range over 5 regressions conducted over data collected in different seasons and from mobile lab samples.~~
~~[f] Ratios derived from mean and standard deviations, with propagation of uncertainty.~~
~~[g] Estimated from Figure 2, Leuchner and Rappengluck, 2010.~~
~~[h] Estimated from Figure 2, Buzcu and Frazer, 2006.~~

Study	Source	propane	n-butane	n-pentane	n-hexane	isopentane
Bakersfield PMF evaporative and fugitive factor[a]	This study	1	0.52 ± 0.02	0.18 ± 0.01	0.06 ± 0.003	0.33 ± 0.02
Bakersfield petroleum operations source profile[b]	Gentner et al. (2014a)	1	0.53 ± 0.1	0.09 ± 0.02	0.04 ± 0.01	0.08 ± 0.02
Mexico city PMF LPG factor[c]	Bon et al. (2011)	1	0.5 (0.4 - 0.7)	0.05 (0.04 - 0.07)	0.02 (0.02 - 0.03)	0.07 (0.06 - 0.1)
Wattenberg field BAO, Colorado[d]	Gilman et al. (2013)	1	0.75 ± 1.37	0.32 ± 0.6	0.08 ± 0.13	0.28 ± 0.52
Wattenberg field BAO, Colorado[e]	Petron et al. (2012)	1	0.58 - 0.65	0.22 - 0.31	NA	0.22 - 0.31
PMF natural gas and evaporation factor, Houston Ship Channel[g]	Leuchner and Rappengluck (2010)	1	0.33	0.27	0.12	0.37
PMF natural gas factor, Houston Ship Channel[h]	Buzcu and Fraser (2006)	1	0.67 ± 0.16	0.07 ± 0.18	NA	NA

---

[a] Uncertainties calculated from propagation of errors (standard deviations) over FPEAK range of -1.6 to 0.4.
[b] Ratios calculated from Table 4, Gentner et al., 2014a; uncertainties defined as ±20% to account for variability in oil well data.
[c] Uncertainties calculated from propagation of uncertainties over FPEAK range of -3 to 3.
[d] Emission ratios derived from multivariate regression analysis; error bars derived from propagation of uncertainty using mean and standard deviation of samples.
[e] Range over 5 regressions conducted over data collected in different seasons and from mobile lab samples.
[f] Ratios derived from mean and standard deviations, with propagation of uncertainty.
[g] Estimated from Figure 2, Leuchner and Rappengluck, 2010.
[h] Estimated from Figure 2, Buzcu and Frazer, 2006.

Table 3. Comparison of hydrocarbon ratios to toluene (gC gC^{-1}) from PMF vehicle emission factor with similar ratios from other California specific studies.

~~Study~~	~~Bakersfield PMF vehicle emissions factor[a]~~	~~Bakersfield gasoline source profile[b,c]~~	~~Riverside liquid gasoline profile[e]~~	~~CalNex Los Angeles ambient emission ratios[g]~~
~~Source~~	~~This study~~	~~Gentner et al. (2014)~~	~~Gentner et al. (2009)~~	~~Borbon et al. (2013)~~
~~CH_4~~	~~8.1 ± 2.1~~	~~NA~~	~~NA~~	~~NA~~
~~CO~~	~~14.0 ± 0.4~~	~~NA~~	~~NA~~	~~22.26~~
~~toluene~~	~~1~~	~~1~~	~~1~~	~~1~~
~~isopentane~~	~~0.69 ± 0.01~~	~~0.77 ± 0.04~~	~~0.64-0.84~~	~~1.95~~
~~isooctane~~	~~0.29 ± 0.03~~	~~0.34 ± 0.02~~	~~0.64-0.80~~	~~NA~~
~~n-dodecane~~	~~0.03 ± 0.001~~	~~(0.04 ± 0.004)[d]~~	~~NA~~	~~NA~~
~~methylcyclopentane~~	~~0.24 ± 0.01~~	~~0.32 ± 0.02~~	~~NA~~	~~NA~~
~~ethyl benzene~~	~~0.17 ± 0.01~~	~~0.14 ± 0.01~~	~~NA~~	~~0.2~~
~~m/p – xylene~~	~~0.65 ± 0.01~~	~~0.65 ± 0.03~~	~~(0.45-0.52)[f]~~	~~0.64~~
~~o – xylene~~	~~0.22 ± 0.01~~	~~0.23 ± 0.01~~	~~NA~~	~~0.24~~

~~[a] errors are standard deviation of 12 unique PMF solutions between FPEAK = -1.6 to +0.4; see section S2.~~
~~[b] derived from liquid gasoline fuel speciation profile (Table S9; Gentner et al., 2012).~~
~~[c] errors bars derived from propagation of uncertainties.~~
~~[d] derived by combining diesel fuel and gasoline speciation profile (Table S9 and S10; Gentner et al., 2012).~~
~~[e] summer data.~~
~~[f] only m-xylene.~~
~~[g] derived from Linear Regression Fit slope of scatterplot from CalNex Pasadena supersite samples.~~

Study	Bakersfield PMF vehicle emissions factor[a]	Bakersfield gasoline source profile[b,c]	Riverside liquid gasoline profile[e]	CalNex Los Angeles ambient emission ratios[g]

Source	This study	Gentner et al. (2012)	Gentner et al. (2009)	Borbon et al. (2013)
CH_4	8.1 ± 2.1	NA	NA	NA
CO	14.0 ± 0.4	NA	NA	45
toluene	1	1	1	1
isopentane	0.69 ± 0.01	0.77 ± 0.04	0.64-0.84	1.95
isooctane	0.29 ± 0.03	0.34 ± 0.02	0.64-0.80	NA
n-dodecane	0.03 ± 0.001	(0.02 ± 0.007)[d]	NA	NA
methylcyclopentane	0.24 ± 0.01	0.32 ± 0.02	NA	NA
ethyl benzene	0.17 ± 0.01	0.14 ± 0.01	NA	0.2
m/p - xylene	0.65 ± 0.01	0.65 ± 0.03	(0.45-0.52)[f]	0.64
o - xylene	0.22 ± 0.01	0.23 ± 0.01	NA	0.24

[a] errors are standard deviation of 12 unique PMF solutions between FPEAK =-1.6 to +0.4; see section S2.
[b] derived from liquid gasoline fuel speciation profile (Table S9; Gentner et al., 2012).
[c] errors bars derived from propagation of uncertainties.
[d] derived by combining diesel fuel and gasoline speciation profile (Table S9 and S10; Gentner et al., 2012) and gasoline and diesel fuel sale data in Kern County (Table S1, Gentner et al., 2012).
[e] summer data.
[f] only m-xylene.
[g] derived from Linear Regression Fit slope of scatterplot from CalNex Pasadena supersite samples.

Table 4. Comparison of PMF dairy and livestock emission factors (mmol mol^{-1}) with previous studies.

~~Study~~	~~Source~~	~~Cow/manure type (if applicable)~~	~~methanol / methane EF avg. (range)~~	~~ethanol / methane EF avg. (range)~~
~~PMF analysis of regional measurements~~	~~This study~~	-	~~15 – 47~~	~~9 – 32.2~~
~~Environmental chamber with cows and/or manure~~	~~Shaw et al. (2008)~~	~~Dry~~	~~3.2 (0.6 – 7.4)~~	~~NA~~
		~~Lactating~~	~~1.9 (0.8 – 3.6)~~	~~NA~~
~~Environmental chamber with cows and/or manure~~	~~Sun et al. (2008)~~	~~Dry~~	~~13.4 (4 – 25)~~	~~14.4 (11 – 19)~~
		~~Lactating~~	~~19.2 (15 – 25)~~	~~24.2 (18 – 32)~~
~~Cowshed with regular dairy operations (winter)~~	~~Ngwabie et al. (2008)~~	-	~~2.0 (1.6 – 2.4)~~	~~9.3 (4 – 16)~~
~~Cow stall area with regular dairy operations (summer)~~	~~Filipy et al. (2006)~~	-	~~NA~~	~~(42 – 127)~~[a]
~~Manure from cattle feedlot~~	~~Miller and Varel (2001)~~	~~Fresh (< 24 hr)~~	~~NA~~	~~14~~[b]
		~~Aged (> 24 hr)~~		~~118~~[b]
~~Measured slope of regression (CalNex 2010)~~	~~Gentner et al. (2014)~~	-	~~7.4 (7 – 16)~~[c]	~~18~~[d]
~~Sampling of dairy plumes from aircraft (CABERNET 2011)~~	~~Guha et al. (*in prep*)~~	-	~~9.6 (9 – 30)~~[c]	~~NA~~

---

[a] ~~calculated based on CH$_4$ emission rate of 4160 µg cow^{-1} s^{-1} for mid-lactating cows (Shaw et al., 2007).~~

[b] ~~calculated based on CH4 emission rate of 4160 µg cow-1 s-1 for mid-lactating cows (Shaw et al., 2007) and ethanol emission rate for fresh and aged manure of 175 and 1223 µg cow-1 s-1, respectively, derived by Filipy et al. (2006).~~

[c] ~~slope of regression with range of measured slopes (in parentheses) from sampling of dairy plumes by aircraft.~~

[d] ~~ground site data; lower limit of slope of non-vehicular ethanol versus methane~~

Study	Source	Cow/manure type (if applicable)	methanol / methane EF avg. (range)	ethanol / methane EF avg. (range)
PMF analysis of regional measurements	This study	-	15 - 47	9 - 32.2
Environmental chamber with cows and/or manure	Shaw et al. (2008)	Dry	3.2 (0.6 - 7.4)	NA
		Lactating	1.9 (0.8 - 3.6)	NA
Environmental chamber with cows and/or manure	Sun et al. (2008)	Dry	13.4 (4 - 25)	14.4 (11 - 19)
		Lactating	19.2 (15 - 25)	24.2 (18 - 32)
Cowshed with regular dairy operations (winter)	Ngwabie et al. (2008)	-	2.0 (1.6 - 2.4)	9.3 (4 - 16)
Cow stall area with regular dairy operations (summer)	Filipy et al. (2006)	-	NA	(42 - 127)[a]
Manure from cattle feedlot	Miller and Varel (2001)	Fresh (< 24 hr)	NA	14[b]
		Aged (> 24 hr)		118[b]
Measured slope of regression (CalNex 2010)	Gentner et al. (2014a)	-	7.4 (7 - 16)[c]	18[d]
Sampling of dairy plumes from aircraft (CABERNET 2011)	Guha et al. (*in prep*)	-	9.6 (9 - 30)[c]	NA

[a] calculated based on CH_4 emission rate of 4160 µg cow^{-1} s^{-1} for mid-lactating cows (Shaw et al., 2007).

[b] calculated based on CH_4 emission rate of 4160 µg cow-1 s-1 for mid-lactating cows (Shaw et al., 2007) and ethanol emission rate for fresh and aged manure of 175 and 1223 µg cow-1 s-1, respectively, derived by Filipy et al. (2006).

[c] slope of regression with range of measured slopes (in parentheses) from sampling of dairy plumes by aircraft.

[d] ground site data; lower limit of slope of non-vehicular ethanol versus methane

Table 5. Comparison of PMF agricultural and soil management emission factor for acetone versus methanol ($gC\ gC^{-1}$) with ratios of basal emission factors generated for major crops grown in the Kern County. Errors denote standard deviations computed by propagation of uncertainty.

Bakersfield PMF agricultural and soil management factor	Almond greenhouse summer 2008	Table grape greenhouse summer 2008	Pistachio greenhouse summer 2008	Navel oranges greenhouse summer 2008[a]	Valencia oranges greenhouse summer 2008
This study	Gentner et al. (2014b)	Gentner et al. (2014b)	Gentner et al. (2014b)	Fares et al. (2011)	Fares et al. (2012)
0.58 ± 0.37	0.14 ± 0.2	0.04 ± 0.02	0.5 ± 0.6	0.57 ± 0.1	0.5 ± 0.3

[a] branch with flowers not removed.

Figures

Figure 1

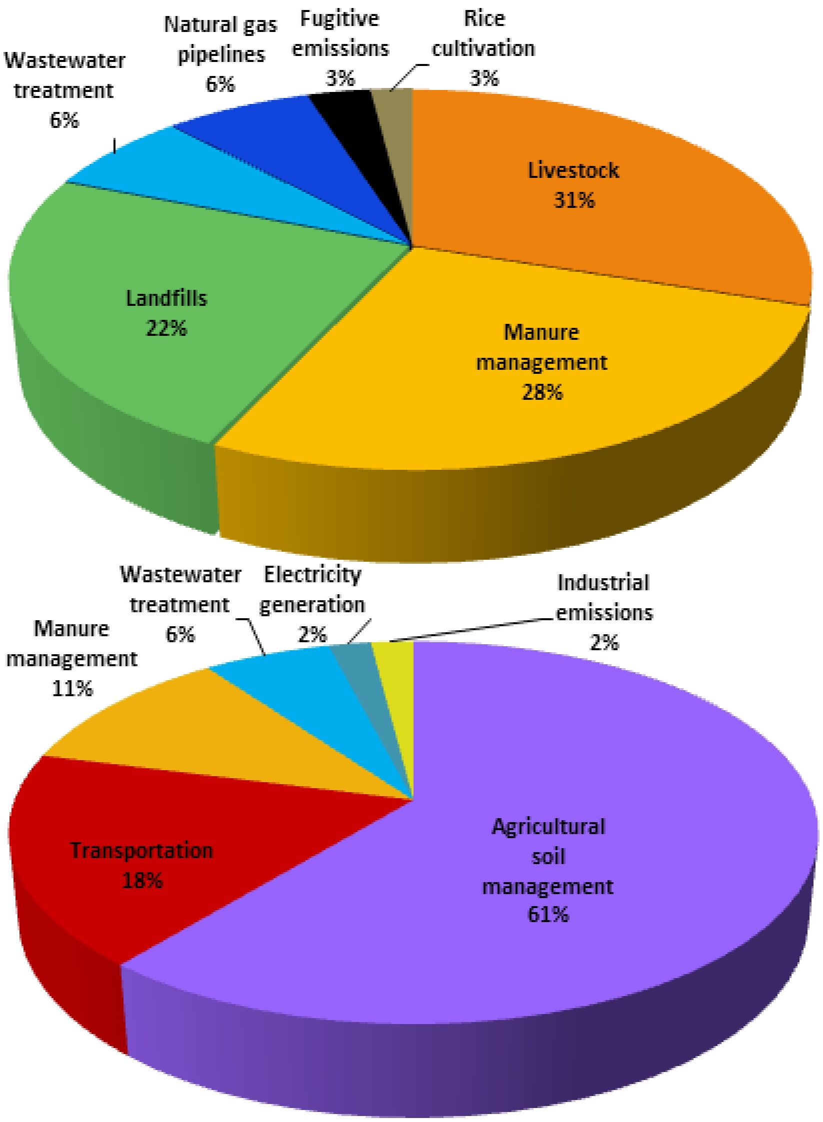

Figure 1. 2011 California emission inventory for **(a)** methane (CH_4) - 32.5 million ton CO_2eq at GWP = 25; and **(b)** nitrous oxide (N_2O) - 13.4 million ton CO_2eq at GWP = 298. (Source: CARB GHG Inventory Tool, Aug 2013)

Figure 2

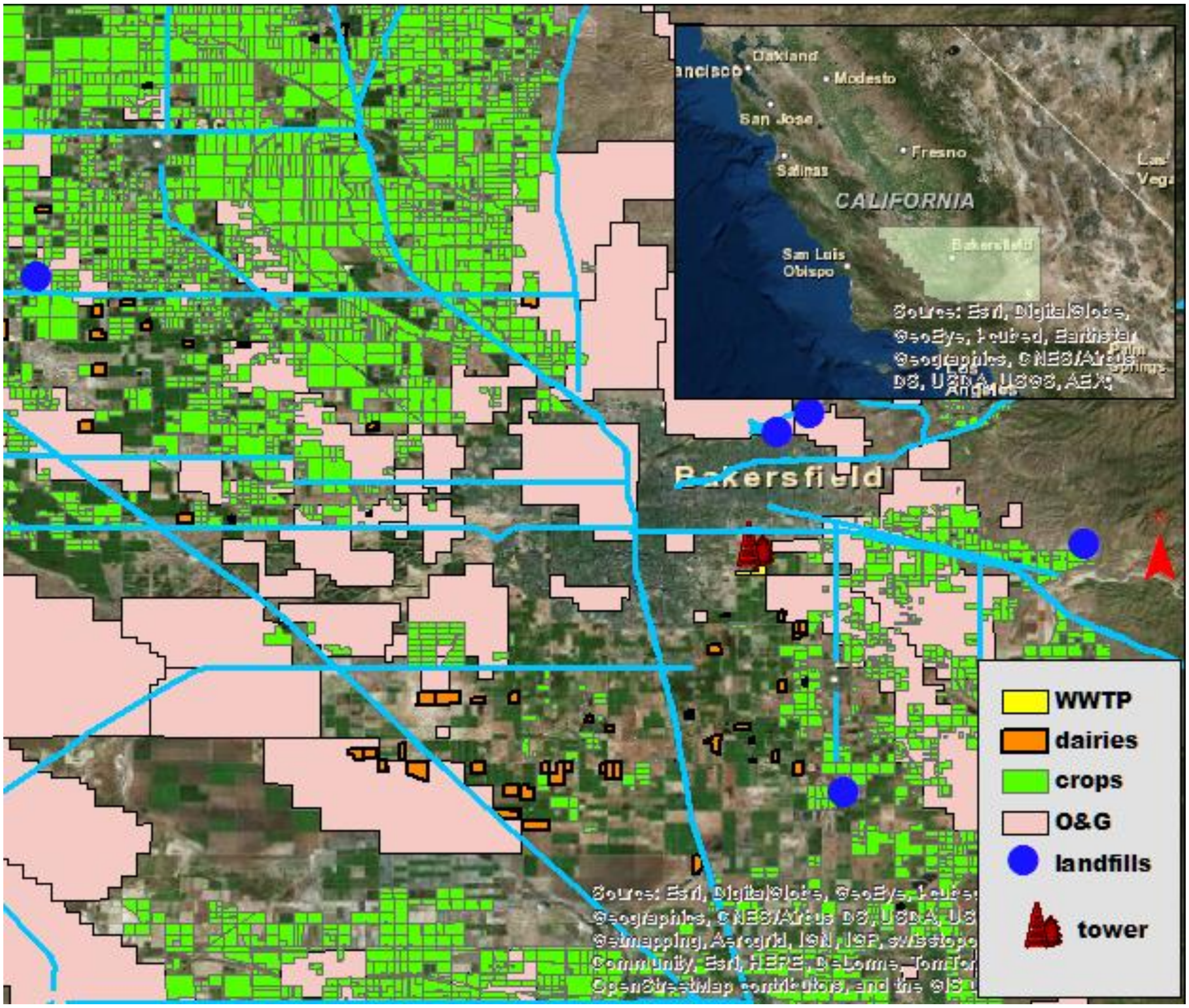

Figure 2. Map of potential sources of methane and nitrous oxide in/around the city of Bakersfield and the surrounding parts of the valley. The inset map is a zoomed out image of the southern part of San Joaquin Valley (SJV) with location of Kern County superimposed. The light blue lines mark the highways, WWTP stands for waste water treatment plant, and O&G stands for oil and gas fields. The location of the CalNex experiment site is marked by the 'tower' symbol.

Figure 3

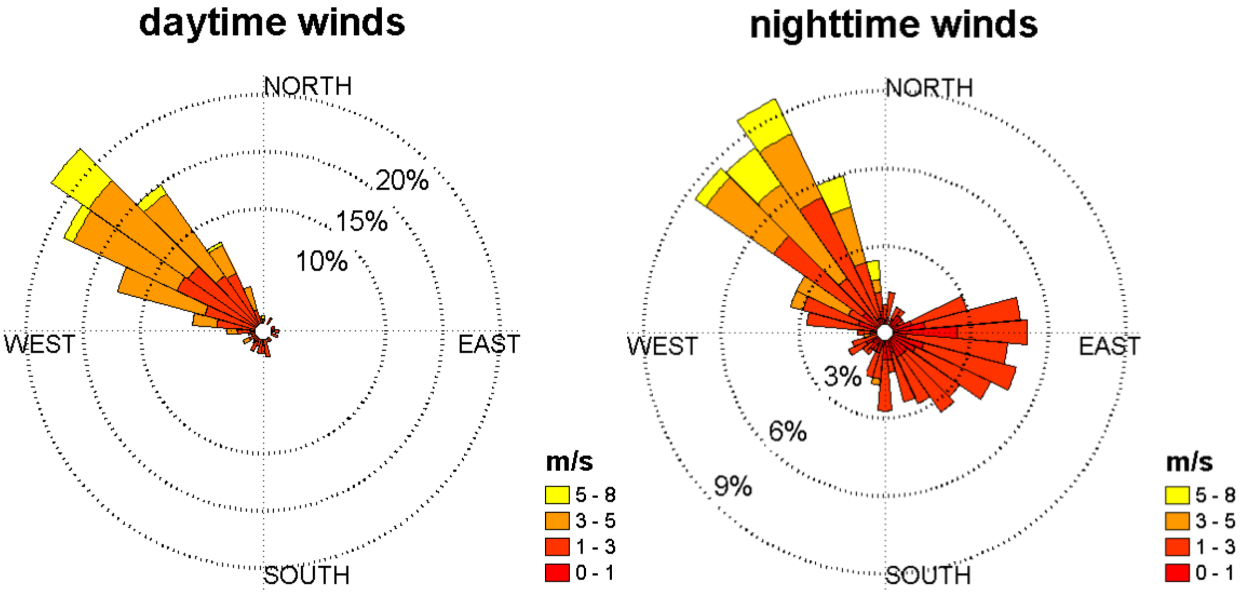

Figure 3. Wind rose plots showing mean wind direction measured at the site during **(left)** day time (07:00-16:00 hour), and **(right)** nighttime (17:00-06:00 hour) during the experiment period in summer 2010. The concentric circles represent the percentage of total observations; each colored pie represents a range of 10° while the colors denote different wind speed ranges.

Figure 4

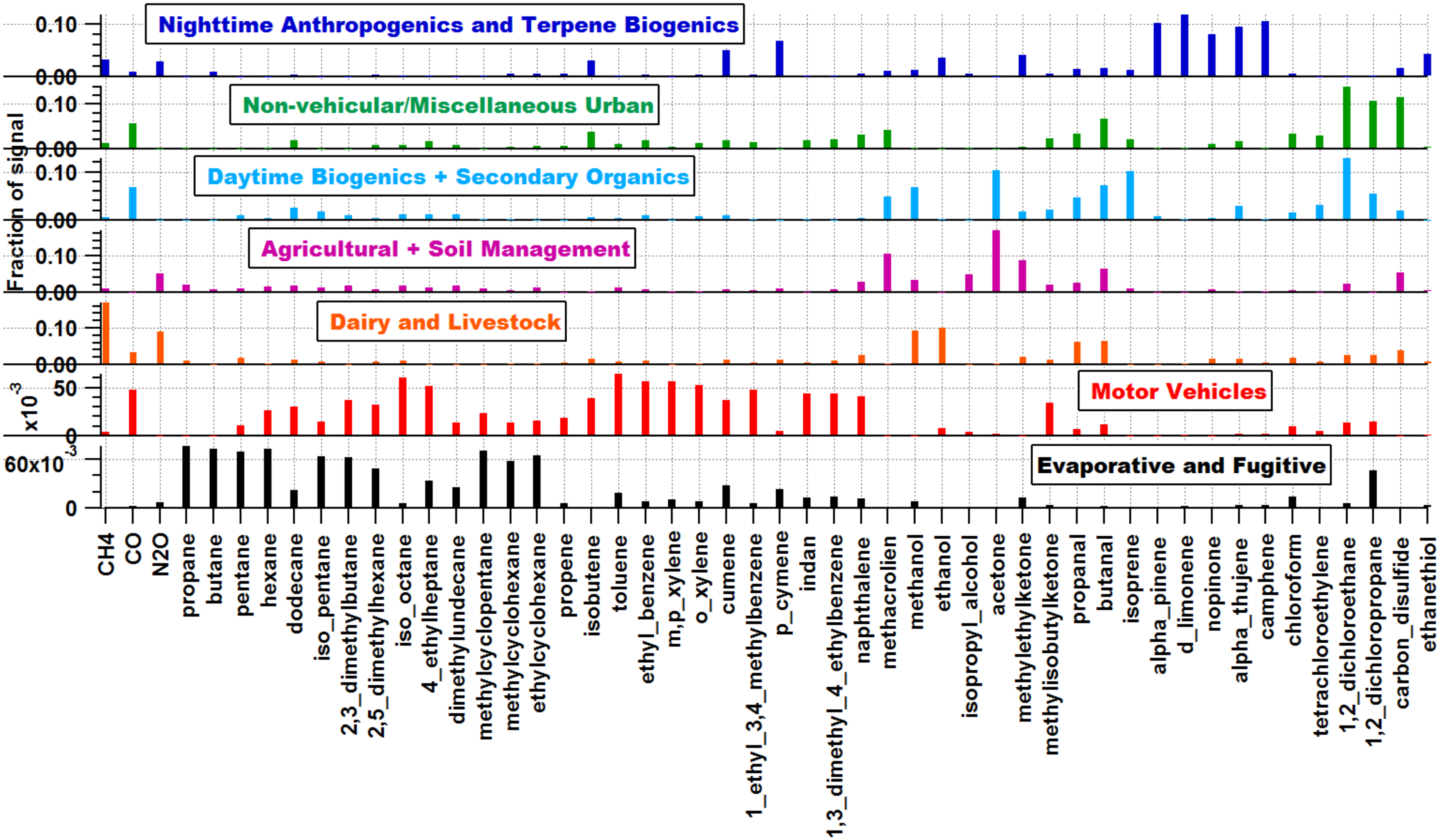

Figure 4. Source profile of the seven factors derived using PMF. The source factors are evaporative and fugitive, motor vehicles, dairy and livestock, agricultural + soil management, daytime biogenics + secondary organics, urban, and nighttime anthropogenics + terpene biogenics. The

y~~x~~-axis represents the normalized fraction of mass in each source factor, while the x~~y~~-axis lists all the chemical species included in the PMF analysis.

Figure 5

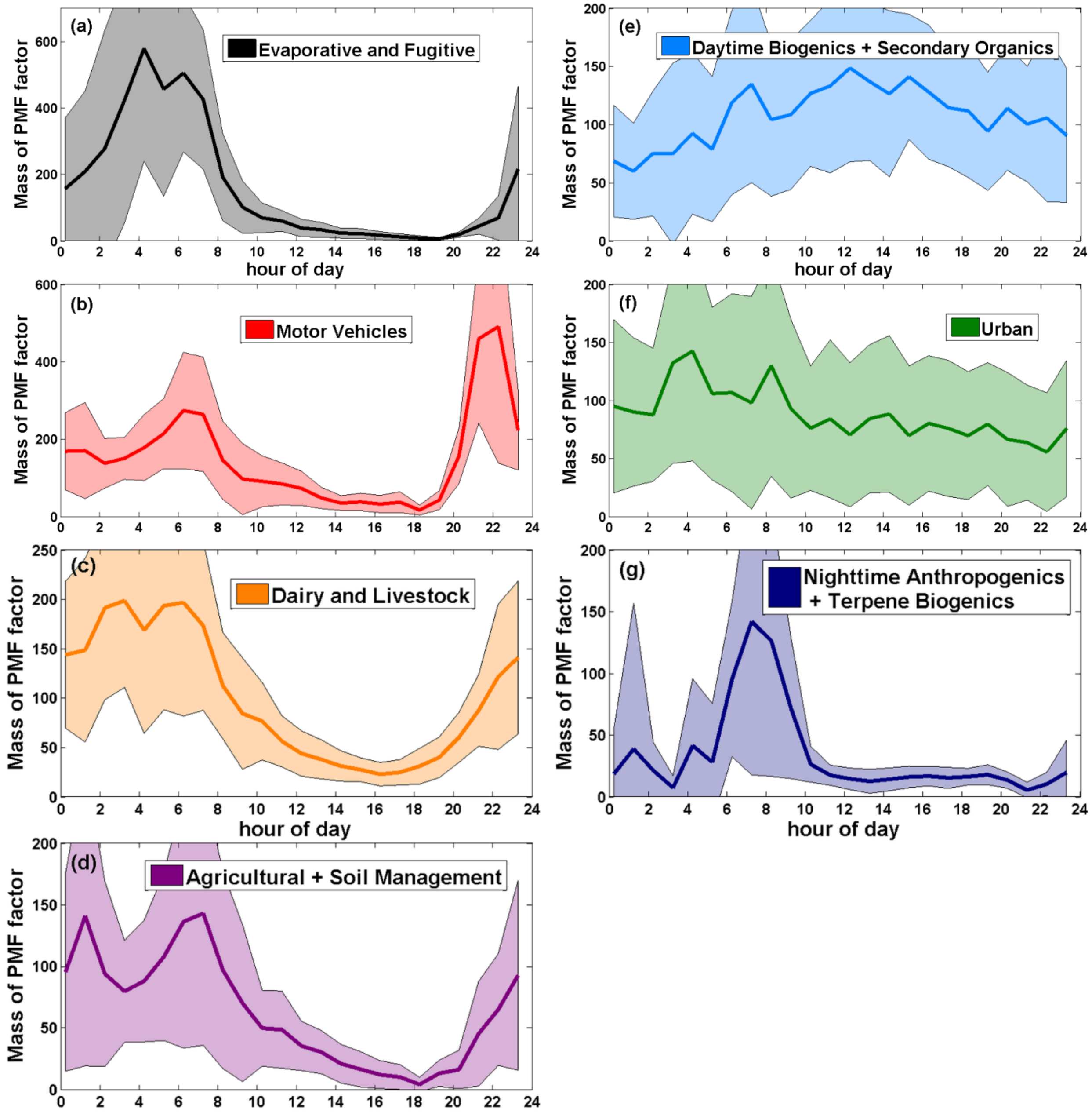

Figure 5. Mean hourly diurnal plots of PMF source factor concentration enhancements for **(a)** evaporative and fugitive, **(b)** motor vehicles, **(c)** dairy and livestock, **(d)** agricultural + soil management, **(e)** daytime biogenics and secondary organics, **(f)** non-vehicular/miscellaneous urban and **(g)** nighttime anthropogenics + terpene biogenics. The y~~x~~-axis represents sum of normalized mass concentrations from all tracers contributing to the factor. The x~~y~~-axis is hour of day (local time). The solid lines represent the mean and the shaded area represents the standard deviation (variability) at each hour.

Figure 6

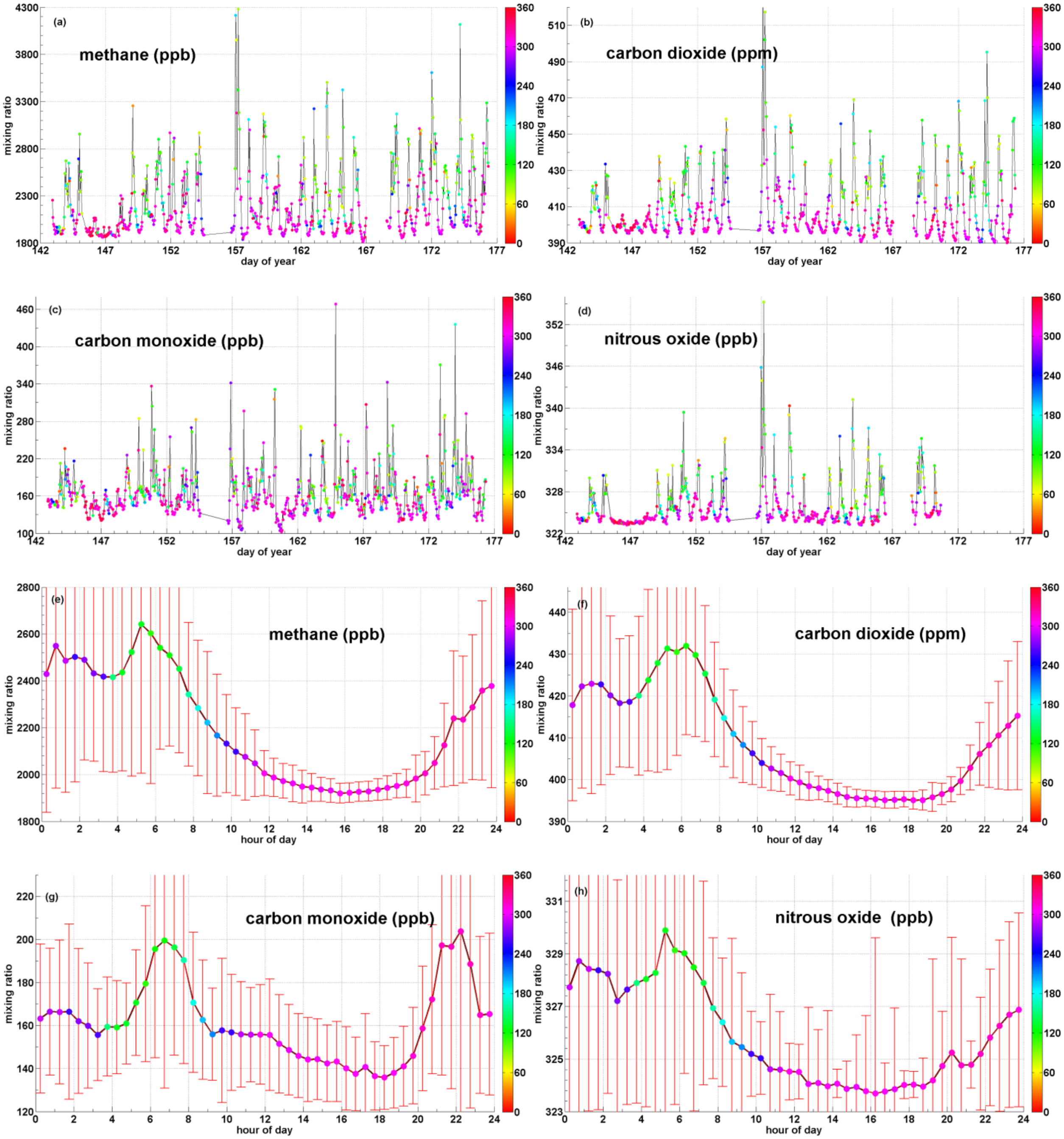

Figure 6. Time series of **(a)** CH_4, **(b)** CO_2, **(c)** CO, and **(d)** N_2O obtained from 30-min averages ~~over the entire sampling period~~from May 15- June 30, 2010. The color bar indicates the average wind direction during each 30-min period. Mixing ratios plotted as average diurnal cycles for **(e)** CH_4, **(f)** CO_2, **(g)** CO and **(h)** N_2O along with wind direction. The curve and the red whiskers represent the mean and the standard deviations about the mean, respectively.

Figure 7

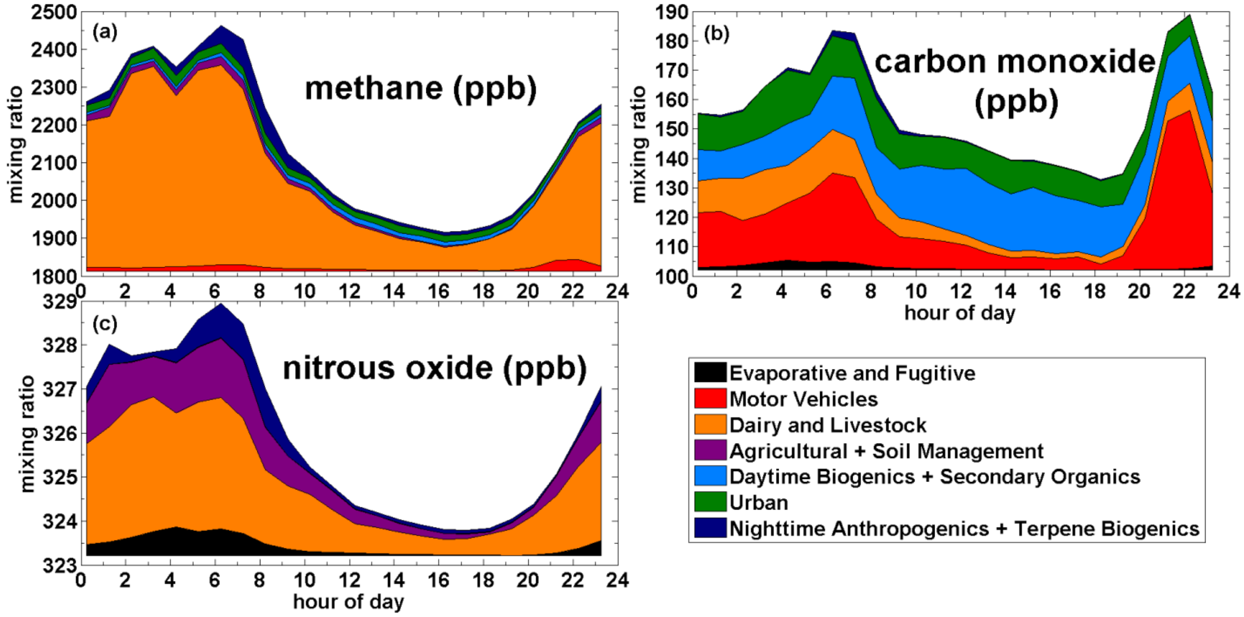

Figure 7. Diurnal plot of PMF derived **(a)** CH_4, **(b)** CO, and **(c)** N_2O concentrations sorted by PMF source category. The legend on the bottom right shows the names of the PMF source factor which each color represents. The PMF derived enhancements from each source have been added to the background concentrations.