**Atmospheric
Chemistry
and Physics**

Emissions of organic carbon and methane from petroleum and dairy operations in California's San Joaquin Valley

D. R. Gentner[1,*], T. B. Ford[2], A. Guha[3], K. Boulanger[1,], J. Brioude[4,5], W. M. Angevine[4,5], J. A. de Gouw[4,5], C. Warneke[4,5], J. B. Gilman[4,5], T. B. Ryerson[4,5], J. Peischl[4,5], S. Meinardi[6], D. R. Blake[6], E. Atlas[7], W. A. Lonneman[8], T. E. Kleindienst[8], M. R. Beaver[9,***], J. M. St. Clair[9], P. O. Wennberg[9], T. C. VandenBoer[10,****], M. Z. Markovic[10,*****,******], J. G. Murphy[10], R. A. Harley[1], and A. H. Goldstein[1,3]**

[1]Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720, USA
[2]Department of Chemistry, University of California, Berkeley, CA, 94720, USA
[3]Department of Environmental Science, Policy and Management, University of California, Berkeley, CA, 94720, USA
[4]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA
[5]Chemical Sciences Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305, USA
[6]Department of Chemistry, University of California, Irvine, CA, 92697, USA
[7]Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL, 33149, USA
[8]National Exposure Research Laboratory, Environmental Protection Agency, Research Triangle Park, NC, USA
[9]California Institute of Technology, Pasadena, CA, USA
[10]Department of Chemistry, University of Toronto, Toronto, ON, CA
[*]now at: Department of Chemical and Environmental Engineering, Yale University, New Haven, CT 06511, USA
[**]now at: Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[***]now at: National Exposure Research Laboratory, Environmental Protection Agency, Research Triangle Park, NC, USA
[****]now at: Department of Chemistry, Memorial University of Newfoundland, NL, CA
[*****]now at: Chemical Sciences Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305, USA
[******]now at: Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA

Correspondence to: D. R. Gentner (drew.gentner@yale.edu)

Abstract. Petroleum and dairy operations are prominent sources of gas-phase organic compounds in California's San Joaquin Valley. It is essential to understand the emissions and air quality impacts of these relatively understudied sources, especially for oil/gas operations in light of increasing US production. Ground site measurements in Bakersfield and regional aircraft measurements of reactive gas-phase organic compounds and methane were part of the CalNex (California Research at the Nexus of Air Quality and Climate Change) project to determine the sources contributing to regional gas-phase organic carbon emissions. Using a combination of near-source and downwind data, we assess the composition and magnitude of emissions, and provide average source profiles. To examine the spatial distribution of emissions in the San Joaquin Valley, we developed a statistical modeling method using ground-based data and the FLEXPART-WRF transport and meteorological model. We present evidence for large sources of paraffinic hydrocarbons from petroleum operations and oxygenated compounds from dairy (and other cattle) operations. In addition to the small straight-chain alkanes typically associated with petroleum operations, we observed a wide range of branched and cyclic alkanes,

most of which have limited previous in situ measurements or characterization in petroleum operation emissions. Observed dairy emissions were dominated by ethanol, methanol, acetic acid, and methane. Dairy operations were responsible for the vast majority of methane emissions in the San Joaquin Valley; observations of methane were well correlated with non-vehicular ethanol, and multiple assessments of the spatial distribution of emissions in the San Joaquin Valley highlight the dominance of dairy operations for methane emissions. The petroleum operations source profile was developed using the composition of non-methane hydrocarbons in unrefined natural gas associated with crude oil. The observed source profile is consistent with fugitive emissions of condensate during storage or processing of associated gas following extraction and methane separation. Aircraft observations of concentration hotspots near oil wells and dairies are consistent with the statistical source footprint determined via our FLEXPART-WRF-based modeling method and ground-based data. We quantitatively compared our observations at Bakersfield to the California Air Resources Board emission inventory and find consistency for relative emission rates of reactive organic gases between the aforementioned sources and motor vehicles in the region. We estimate that petroleum and dairy operations each comprised 22 % of anthropogenic non-methane organic carbon at Bakersfield and were each responsible for 8–13 % of potential precursors to ozone. Yet, their direct impacts as potential secondary organic aerosol (SOA) precursors were estimated to be minor for the source profiles observed in the San Joaquin Valley.

1 Introduction

California's San Joaquin Valley contains a large density of dairy farms and is an important region for oil and natural gas production in the United States. Both sources are prominent in the California Air Resources Board (CARB) emission inventory of reactive organic gases (ROG) in the San Joaquin Valley (California Air Resources Board, 2010). Recent work has described large emissions and impacts from new oil/gas operations with increased US production (Petron et al., 2012; Gilman et al., 2013; Carter and Seinfeld, 2012; Schnell et al., 2009; Kemball-Cook et al., 2010; Pacsi et al., 2013). Petroleum operations include extraction, storage, transport, and processing; all of which can have varying degrees of fugitive emissions of methane and other gas-phase organic carbon, such as volatile organic compounds (VOCs) (Leuchner and Rappengluck, 2010; Buzcu and Fraser, 2006, Katzenstein et al., 2003; Petron et al., 2012; Gilman et al., 2013). Crude oil and unrefined natural gas are composed of a suite of organic compounds that span a range of vapor pressures, and are either produced by thermogenic or biogenic processes in the reservoirs (Lillis et al., 2007; Ryerson et al., 2011). Thermogenic gas is geochemically produced via

the cracking of larger compounds in oil and can either be termed associated or non-associated depending on the presence of oil (Lillis et al., 2007). The vast majority of wells in the San Joaquin Valley are oil wells and most have associated gas, also known as wet thermogenic gas (Lillis et al., 2007). Thermogenic wet gas is predominately found in oil wells and contains substantial amounts of non-methane hydrocarbons ranging from 3 to 40 % C_2 and greater content (e.g., Table 1) (Lillis et al., 2007). The San Joaquin Valley has historically been an active region for oil/gas production. In 2010, crude oil production in Kern County, located at the southern end of the San Joaquin Valley, was 450 000 barrels day^{-1}, which represents 69 % of production within California and 8 % of national production (US EIA, 2010; Sheridan, 2006).

There have been several studies on fugitive emissions from oil and gas operations, including emissions from isolated facilities at oil or gas fields, extraction facilities using advanced recovery methods (i.e., hydraulic fracturing), and urban areas with industrial storage and processing facilities (Leuchner and Rappengluck, 2010; Buzcu and Fraser, 2006; Katzenstein et al., 2003; Petron et al., 2012; Gilman et al., 2013). These studies all provide important advances in the characterization of emissions from petroleum operations, but there is considerable variability between regions due to differences in reservoirs and production methods. The specific equipment/processes, state/county regulations, and regional composition of crude oil and natural gas are critical for determining the potential emission pathways and composition of fugitive emissions. So, regional studies remain important to effectively characterize petroleum operation sources.

Previous research on dairy farms and livestock operations has reported emissions of methane, alcohols, carbonyls, esters, acids, and other organic hydrocarbons. Among these, emissions are dominated by methane, methanol, ethanol, and acetic acid (Alanis et al., 2010; Chung et al., 2010; Hafner et al., 2013; Howard et al., 2010a, b; Malkina et al., 2011; Sun et al., 2008; Shaw et al., 2007). Howard et al. (2010b) recently concluded that emissions from dairy operations are dominant contributors to ozone production in California's Central Valley (comprised of the San Joaquin Valley and the Sacramento Valley to the north), but modeling studies suggest a larger role for VOC emissions from motor vehicles (Hu et al., 2012). Methane and oxygenated organic compounds are emitted via several pathways and sources, all co-located at dairies (and their farms). Silage processing/fermentation, bovine enteric fermentation, and animal waste are among the most dominant sources (Alanis et al., 2010; Chung et al., 2010; Hafner et al., 2013; Malkina et al., 2011; Sun et al., 2008; Shaw et al., 2007). The composition of emissions from each of these sources is different and varies widely depending on factors such as feed composition. The animal feed, also known as total mixed rations, is typically comprised of corn and other grains (i.e., silage), with corn being the abundant type in the US (Hafner et al., 2013). The silage is fermented on-site in large piles and mixed with various adjuncts

Table 1. Unrefined natural gas composition for thermogenic wet wells in the San Joaquin Valley from USGS samples ($N = 49$ wells) (Lillis et al., 2007).

	wtC %	Std. Dev.	k_{OH}	MIR
methane	82.3	9.2	0.0064	0.014
ethane	5.33	3.46	0.248	0.28
propane	4.42	3.50	1.09	0.49
isobutane	0.920	0.837	2.12	1.23
n-butane	1.55	2.17	2.36	1.15
isopentane	0.223	0.401	3.6	1.45
n-pentane	0.273	0.405	3.80	1.31
neopentane	0.061	0.182	0.825	0.67
n-hexane	0.105	0.108	5.20	1.24
n-heptane	0.049	0.041	6.76	1.07

Notes:
– k_{OH} is in cm^3 s^{-1} molecules^{-1} $\times 10^{12}$ from Atkinson and Arey, (2003).
– MIR is in gO$_3$ g^{-1} from Carter 2007.
– The observed source profile for petroleum gas emissions at the Bakersfield site is well represented by the composition of non-methane organic carbon shown here.

(e.g., almond shells, fruit, fat). The site-by-site heterogeneity in feed composition and the processing of both animal feed and waste leads to variability in the source profile and emission ratios of organic compounds from dairy operations. This work aims to reduce this uncertainty by estimating the average source profile for dairy operation emissions in the San Joaquin Valley.

The objectives of this work are to examine the magnitude, chemical composition, and spatial distribution of organic carbon emissions from petroleum and dairy operations in the San Joaquin Valley. This is accomplished using multiple gas-phase organic carbon data sets from stationary ground sites and aircraft platforms. Our approach includes the development of a method to assess the spatial distribution of sources (i.e., a statistical source footprint) via ground site measurements and meteorological modeling. We examine the relative abundance of emissions from petroleum and dairy operations against other prominent anthropogenic sources in the San Joaquin Valley, and evaluate their potential to impact air quality. We also provide a quantitative assessment of petroleum and dairy operations emissions relative to motor vehicle emissions in the CARB emission inventory.

2 Materials and methods

2.1 Measurement sites and instrumentation

Gas-phase organics and other gases were measured 18 May–30 June 2010 in Bakersfield, CA during the CalNex (California Research at the Nexus of Air Quality and Climate Change) project. The ground supersite (35.3463° N, 118.9654° W) was located in southeast Bakersfield, a city in the southern San Joaquin Valley. With the exception of

gas-sampling canisters and ion chromatography to measure acids, measurements were made from the top of an 18 m tower. Measurements of a few light VOCs are included from canister measurements at ground level to further characterize the observed sources. Canisters were taken as 3 h averages in the morning (05:00–08:00 PST) and analyzed via US Environmental Protection Agency (EPA) methods for an array of organic compounds (Klouda et al., 2002). Supporting methane measurements were made using integrated cavity output spectroscopy (Los Gatos Research, Fast Greenhouse Gas Analyzer) with 1 min time resolution. Acetic acid and other acids were measured using both chemical ionization mass spectrometry (CIMS) and ambient ion monitor–ion chromatography (AIM-IC). These two instruments were located at different heights on the sampling tower in Bakersfield and had different measurement frequencies. With both sets of data averaged to hourly resolution, the acetic acid data were well correlated to each other ($r = 0.84$) with a slope near unity. Details on their sampling and measurement methods have been published previously (Crounse et al., 2006; Markovic et al., 2012).

As part of the CalNex project, measurements were also made from the National Oceanic and Atmospheric Administration (NOAA) WD-P3 research aircraft. VOC canister samples were collected on the aircraft and analyzed offline (Barletta et al., 2013). High time resolution data on selected organic compounds and methane were collected on the aircraft using proton transfer reaction mass spectrometry (PTR-MS) and a Picarro flight-ready greenhouse gas analyzer (model 1301 m), respectively (de Gouw and Warneke, 2007; Peischl et al., 2012). High-resolution data were averaged to 1 min intervals and select flights in the Central Valley were used to evaluate the spatial distribution of methane concentrations (flight dates: 5/7, 5/11, 5/12, 6/14, 6/16, 6/18, 2010).

2.2 Source apportionment methods

2.2.1 Petroleum operations

Using six weeks of in situ VOC data from the Bakersfield ground site, we assessed emissions from petroleum operations during spring and summer 2010. Contributions to observed VOC concentrations at the site from petroleum operations were determined (along with other motor vehicle-related sources) using a source receptor model with chemical mass balancing and effective variance weighting focused on hydrocarbon emissions from petroleum-related sources (Gentner et al., 2012). The model used 10 compounds emitted from the sources of interest (petroleum operations, non-tailpipe gasoline emissions, gasoline exhaust, and diesel exhaust) along with reliable information on the fractional composition of the 10 compounds from each of the sources (i.e., source profiles). The 10 compounds used were dependent species, but the model also calculated the predicted concentrations of all the independent compounds not included in

the model, but emitted by the petroleum-related sources and measured at the site.

The compounds used in the over-constrained (i.e., more tracer compounds than sources) source receptor model were propane, n-butane, n-pentane, isopentane, m/p-xylene, o-xylene, isooctane, n-nonane, n-undecane, and n-dodecane to model motor vehicle and petroleum operation sources. Due to high background concentrations, measurements of propane and n-butane were corrected by local background values of 500 and 100 pptv, respectively. The 10 tracer compounds were carefully selected because together they captured the dynamics of all four petroleum-related sources. The atmospheric lifetimes of the more reactive species did not bias the model since the vast majority of contributions (i.e., emissions) were within short transport times to the site. The petroleum operations source had the longest transport times (up to 6 h) from source to field site, which did not present a problem because that source was represented and modeled by the least reactive species that had negligible degradation during transport. Extensive details on these methods and model validation are described in detail in Gentner et al. (2012).

A priori source profile information for the model was constructed using US Geological Survey data on associated thermogenic natural gas composition from wells in the San Joaquin Valley (Table 1) (Lillis et al., 2007) and regional gasoline/diesel fuel composition data (Gentner et al., 2012). There was substantial variability between wells and sampling methods in the data compiled by the USGS, so standard deviations for the petroleum operations source profile were ±80–300 %. Due to this large uncertainty, we represented the uncertainty for all the source profiles in the model by standard errors (similar to the US EPA CMB 8.2 model), defined as the standard deviation divided by the square root of the sample size ($N = 49$).

The source receptor model effectively modeled the compounds included in the initial petroleum operations source profile (Table 1), but there were an array of hydrocarbons (not among the compounds used in the model) that episodically exceed predicted concentrations based on emissions from motor vehicles. Many of the excess hydrocarbon concentrations were well correlated with each other and the petroleum operations source factor, likely indicating emissions from the petroleum operations source. Emissions of additional compounds from petroleum operations (not present in the initial limited petroleum gas profile) are derived from the residual mass that is well correlated with the petroleum operations source. The residuals, or excess concentrations beyond contributions from motor vehicles, were filtered for values that exceeded the uncertainties of model calculations, which are determined in part by the 10–20 % variability in gasoline and diesel fuel.

Similarly, we calculated the expected ethanol emissions from gasoline vehicles for hourly data. Taking the difference between these predicted concentrations and total observed ethanol results in non-vehicular ethanol concentrations that must be attributed to other ethanol sources, but were not correlated with the petroleum operations source.

2.2.2 Dairy operations

A reliable source profile for dairy operations in the San Joaquin Valley was not available in the literature for all the compounds of interest in this study, so the source profile was established using a mix of aircraft and ground measurements. The emission ratios of organic compounds to methane were calculated using flight and ground data for compounds that had evident, quantifiable emissions from dairy operations to construct the source profile. The ratio of methanol to methane in dairy operation emissions was determined using 1 min aircraft data points sampled in the plumes from farms and facilities in the San Joaquin Valley. Acetic acid and ethanol ratios could not be determined using the flight data due to a lack of measurements and spatial incongruence of canister to methane data, respectively. Ratios of these two compounds to methane were determined using ground site data from Bakersfield. Dairies have been shown in previous studies to be major sources of methane, methanol, ethanol, acetic acid, and other oxygenated species; and there is a large concentration of dairies in the San Joaquin Valley (Alanis et al., 2010; Chung et al., 2010; Howard et al., 2010a, b; Malkina et al., 2011; Sun et al., 2008; Shaw et al., 2007). So each compound is compared to methane via regression with close attention to enhancements from other sources that may skew the observed dairy operations emission ratio.

Predicted concentrations in Bakersfield of methanol, ethanol, and acetic acid from dairy operations were estimated using the determined emission ratios to methane and measurements of methane at the Bakersfield ground site. A local background methane concentration of 1.87 ppmv was subtracted prior to multiplication by the emission ratio. These predicted concentrations were compared with observed concentrations to determine the fraction of each compound emitted from dairy operations.

OH reactivities and ozone formation potentials reported in this paper are from literature on OH reaction constants and maximum incremental reactivities (MIRs), respectively (Carter, 2007; Atkinson and Arey, 2003).

2.3 Methods to determine spatial distribution of emissions

Several methods are used in this work to assess the spatial distribution of organic carbon sources. In addition to the use of aircraft data collected from the NOAA WD-P3 mobile platform during the CalNex campaign, we developed a method that uses a Lagrangian transport and meteorological model (FLEXPART-WRF) to calculate the distribution of air parcels (i.e., back-trajectory footprints) for each hourly sample prior to measurement at a ground site. We combine these footprints with ambient compound data from the CalNex site

to assess the spatial distribution of emissions for a given compound in a region. Our method builds upon previous techniques (i.e., TrajStat) to estimate source location(s) using ground site data and the Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPLIT) (Wang et al., 2009).

We generated 6 and 12 h back-trajectory footprints with 4×4 km resolution for each hourly sample using the FLEXPART Lagrangian transport model with WRF meteorological modeling (Fig. 1). Simulations were initiated from the top of the 18 m tower using WRF runs EM4N in Angevine et al. (2012); further details on FLEXPART and WRF modeling can also be found in Brioude et al. (2012) and Metcalf et al. (2012). Here, we integrate this transport/meteorological model with statistical back-trajectory analysis to explore the distribution and relative magnitude of gas-phase organic carbon sources at ground level.

The back-trajectory footprint produced by FLEXPART-WRF represents the area where the air parcel(s) of interest (i.e., a 30 min VOC sample) contacts the surface layer. The statistical source footprint (the final output) represents the calculated distribution of ground-level emissions. Utilizing this concentration-weighted trajectory analysis allows us to find the emissions potential of every point in a region, which is represented by the average concentration of a compound in each cell ($\bar{C}_{ij}$) on a gridded map with i and j representing the axes:

$$\bar{C}_{ij} = \frac{1}{\sum_0^t (\tau_{ijt})} \sum_0^t (c_t \cdot \tau_{ijt}), \qquad (1)$$

where τ_{ijt} is the time each back-trajectory footprint spends at ground level (< 100 m) in the 2-dimensional cell ij for the VOC sample at time t, and c_t is the measured concentration of a compound at the ground site. Each cell has a corresponding n_{ij} value, representing the number of individual footprints included in each cell, which was determined as the number of samples contributing to a cell's average ($\bar{C}_{ij}$) (Seibert et al., 1994). To reduce bias from cells with few samples (i.e., low n_{ij} values), a weighting function multiplies the ($\bar{C}_{ij}$) result by a factor of 1, 0.7, 0.4, or 0.05 for cells with n_{ij} values above the Q_{90}, Q_{75}, Q_{50} or below the Q_{50} percentiles, respectively (Polissar et al., 2001). Contour maps were then plotted using these final values and shown with a 1 arcsec elevation map obtained from the USGS National Map Seamless Server (http://viewer.nationalmap.gov/viewer/). It is insufficient to only consider the distribution of wind directions against compound concentrations when complex meteorology affects the transport of air masses. This is the case in California's Central Valley. Similarly, basic single HYSPLIT back-trajectory analysis can oversimplify the footprint of measurements into one single path and not accurately represent the distribution of ground-level residence times for an air parcel (Fig. 2).

3 Results and discussion

Figure 3 shows measurements of a selection of compounds plotted against carbon monoxide, a common technique to assess contributions from anthropogenic emissions (after filtering for biomass burning events). Some compounds have ratios to CO consistent with measurements from the Los Angeles air basin during the same time period (Borbon et al., 2013). However, there are several compounds with frequent enhancements above the Los Angeles slope, indicating additional sources of these compounds that are not abundant in LA. Most of the compounds shown in Fig. 3 have been previously linked to petroleum and dairy operations (e.g., Gilman et al., 2013 and Shaw et al., 2007), and their enhancements here are evidence for substantial emissions in the San Joaquin Valley.

### 3.1 Emissions from petroleum operations

Petroleum operations emit a significant mass of numerous hydrocarbons, which have a distribution of molecular weights smaller than emissions from gasoline sources. The 25th percentiles for propane and n-butane are similar to other urban ground sites during the summer, but higher concentrations were observed for the 50th and 75th percentiles, by up to a factor of 2 compared to Pittsburgh, PA (2002) (Millet et al., 2005). The 75th percentiles in the San Joaquin Valley are also higher by 25–50 % compared to measurements from 2005 in Riverside, CA, a much more populated region (Gentner et al., 2009). Between the CalNex field sites at Bakersfield and Pasadena, median and smaller values (10th and 25th percentiles) were similar and lower at Bakersfield, respectively. Yet, 75th percentile concentrations were greater at Bakersfield by 53 % for propane (5.6 vs. 3.7 ppbv) and 16 % for n-butane (5.6 vs. 3.7 ppbv). Previous work in the South Coast air basin has also reported emissions of light alkanes from oil/gas operations, but there is a lesser prevalence of oil/gas fields in that air basin compared to the San Joaquin Valley (Peischl et al., 2013).

The source receptor model with chemical mass balancing used in Gentner et al. (2012) effectively modeled emissions of most compounds in a motor vehicle emissions study at the Caldecott tunnel and many of the compounds that are most prevalent in gasoline and diesel emissions at Bakersfield. We used the non-methane composition of thermogenic wet gas reported by the USGS (Table 1) to construct the initial petroleum operations source profile in our source receptor model. The composition of unrefined natural gas has substantial variability among all the wells sampled, but the average composition was very effective for modeling the in situ data from Bakersfield. In many cases, ratios in ambient data can be impacted by differences in the rates of chemical reaction in the atmosphere; as is the case in Los Angeles (Borbon et al., 2013). At Bakersfield, the timescales for transport from source to measurement site are much shorter than the
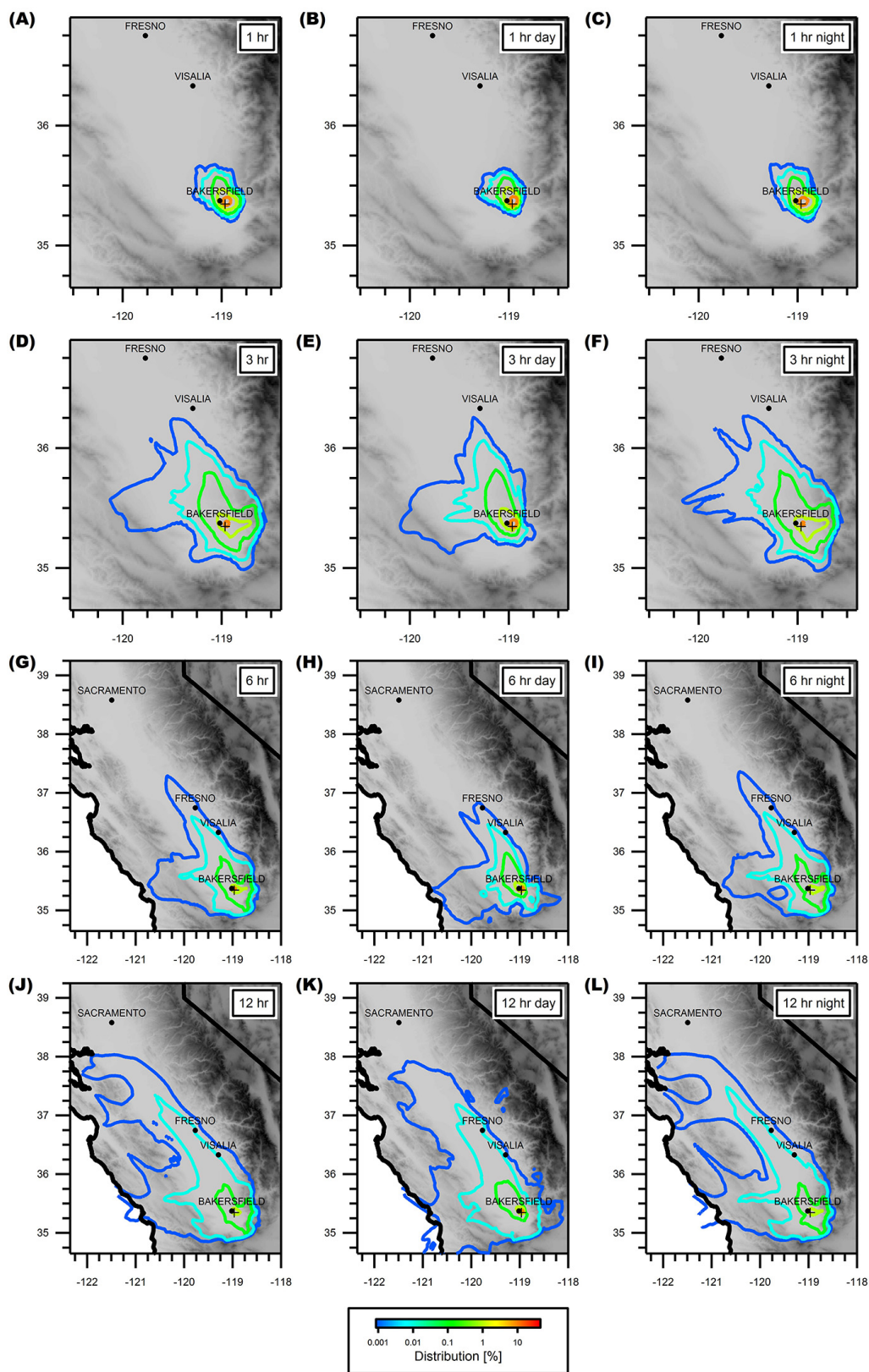
Figure 1. 1, 3, 6 and 12 h statistical footprints for the Bakersfield ground site (marked by +) averaged across the entire CalNex campaign (*y* and *x* axes represent latitude and longitude, respectively). Daytime (**B, E, H, K**) and nighttime (**C, F, I, L**) averages are filtered for 08:00–20:00 PST and 21:00–06:00 PST, respectively, and are shown with overall averages (**A, D, G, J**).
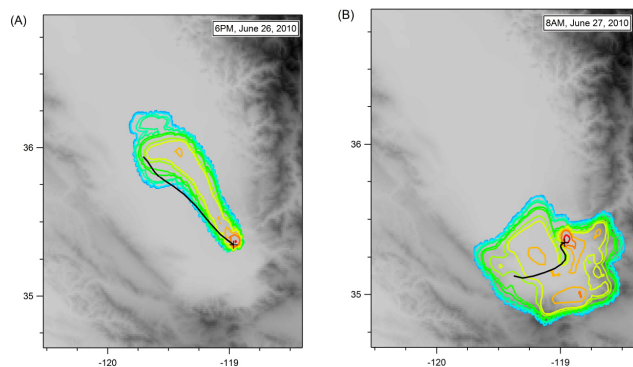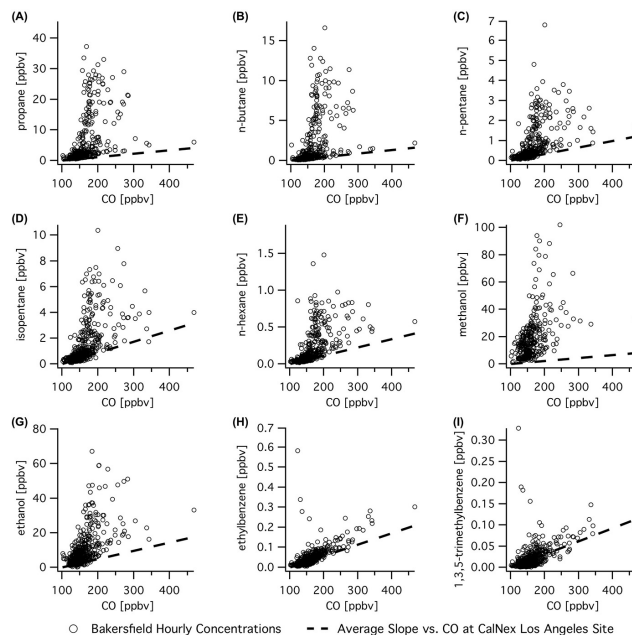
Figure 2. Examples of individual probability distribution back-trajectory footprints produced using FLEXPART-WRF (contours with log color scale – red: max, blue: min) for the southern San Joaquin Valley with air parcels arriving at the CalNex-Bakersfield ground site. Two examples show results for the previous 6 h with air parcels coming **(A)** along a concentrated northwest flowpath and **(B)** a more dispersed footprint from the southern tip of the valley. Dates and arrival times are superimposed on the panels. Also shown are comparisons of single-path HYSPLIT back-trajectories (black lines) and FLEXPART-WRF footprints.

Figure 3. Concentrations of several compounds from Bakersfield, CA shown against carbon monoxide with the average slope of compounds vs. CO during the same time period at the CalNex-LA site in Pasadena, CA (Bourbon et al., 2013). Concentration enhancements above VOC/CO line are due to emissions from **(A–E)** petroleum operations and **(F–G)** dairy operations, neither of which emit CO. **(H–I)** are shown as examples of compounds that largely agree between Bakersfield and Los Angeles.

timescales of reaction for the species considered here. So, variability due to chemical processing is negligible for all but the most reactive primary emitted compounds in our Bakersfield data (Gentner et al., 2012).

In addition to the compounds known to be in thermogenic wet gas (Table 1), the model under-predicted the observed concentrations of numerous alkanes. These compounds are summarized in Table 2 and Fig. 4, which show their average unexplained concentrations and the percent of total mass that is unexplained (determined by the residuals in the chemical mass balance source receptor model). Most of the unexplained concentrations of these alkanes were well correlated ($r \geq 0.75$) with the petroleum operations source contribution from the model and are attributed to this source. The presence of the branched and cyclic alkanes in unrefined petroleum gas is not surprising as there are significant amounts of C_{5-7} straight chain alkanes in the reported composition (Table 1) and a select few have been measured in other studies (Gilman et al., 2013; Ryerson et al., 2011). Yet, there are limited previous in situ measurements for many of the compounds reported here, especially many of the cyclic alkanes. Concentrations of aromatics observed at Bakersfield matched predicted concentrations from motor vehicle sources in our model, but other studies have observed aromatic emissions from petroleum operations (e.g., Gilman et al., 2013).

The additional compounds attributed here to the petroleum operations source profile increase the mass of emissions by 10.6 % as shown by the regression of the correlated "unexplained" compounds with the petroleum gas source ($r = 0.95$) (Fig. 5). The weight fraction of each correlated compound in the "unexplained" mass is given in Table 2 with similar fractions in the overall source profile as the known C_{5-7} compounds in thermogenic wet gas (Table 1). In all, the interquartile range of the unrefined petroleum gas source contribution was 7.6–89 ppbC, with a diurnal pattern that was strongly dependent on meteorological dilution (Supplement Fig. S3). This source represented a substantial fraction of anthropogenic emissions. For comparison, the mass concentration of compounds emitted by the observed petroleum operations source ranged from 30–40 % to 100–150 % of the sum of compounds from motor vehicles during the afternoon and nighttime, respectively (Supplement Fig. S4).

The remaining branched and cyclic compounds that were not highly correlated with the petroleum gas source represent a relatively small amount of mass and we could not confidently infer a specific source for these compounds. The excess C_{13-16} branched alkanes were well correlated ($r \geq 0.80$) with each other, but not with any other compounds. The excess concentrations of C_{10-11} branched alkanes were correlated with each other, and one of the compounds, 2,6-dimethyloctane, was well correlated ($r \geq 0.80$) with the three C_9 cycloalkanes that do not correlate well with the petroleum operations source. These remaining compounds have ozone formation potentials similar to other observed compounds,

Table 2. Interquartile ranges and MIRs for alkanes discussed in this work.

Compound name	# in Fig. 4	Interquartile range [pptv]	WtC % of unexplained mass	MIR [gO$_3$ g^{-1}]
propane	–	1133–5602		0.49
n-butane	–	230–6397		1.15
n-pentane	–	221–2127		1.31
2-2-dimethylbutane	1	28.0–76.6		1.17
2-methylpentane & 2,3-dimethylbutane	2	121.6–501.0	9.92	1.2
3-methylpentane	3	50.1–253.9	7.67	1.80
2,4- & 2,2-dimethylpentane	4	13.7–54.7		1.3
3,3-dimethylpentane	5	4.0–16.6		1.20
2,3-dimethylpentane	6	19.7–93.0		1.34
2-methylhexane	7	23.2–90.3	2.73	1.19
3-methylhexane	8	28.0–124.6	3.48	1.61
2,2-dimethylhexane	9	1.0–4.0		1.02
2,5-dimethylhexane	10	6.2–35.8	1.44	1.46
2,4-dimethylhexane	11	7.4–32.0	0.84	1.73
2,2,3-trimethylpentane	12	2.7–12.1		1.22
isooctane	13	39.1–115.3		1.26
2,3,4-trimethylpentane & ctc-1,2,3-trimethylcyclopentane	14	31.6–160.2	7.38	1.3
2,3,3-trimethylpentane & 2,3-dimethylhexane	15	11.3–32.8		1.1
2-methylheptane	16	10.2–48.8	1.29	1.07
4-methylheptane	17	4.3–20.7		1.25
3-methylheptane	18	9.3–43.6	1.79	1.24
2,2,5-trimethylhexane	19	5.4–16.3		1.13
2,6-dimethylheptane	20	5.4–30.7	1.86	1.04
3,5-dimetylheptane	21	2.2–10.3		1.56
2,3-dimethylheptane	22	0.9–4.7		1.09
2- & 4-methyloctane	23	2.9–12.7		0.9
3-methyloctane & 4-ethylheptane	24	3.1–12.9		1.1
2,2,5-trimethylheptane	25	0.7–1.7		1.26
2,2,4-trimethylheptane	26	0.8–2.4		1.16
C10 branched alkanes (5 unknown isomers)	27	3.0–11.5		0.94
2,6-dimethyloctane	28	0.7–3.2		1.08
2- & 3- & 4-methylnonane & 3- & 4-ethyloctane & 2,3-dimetyloctane	29	6.9–24.6		0.94
C11 branched alkanes (3 unknown isomers)	30	0.7–2.6		0.73
C11 branched alkanes (10 unknown isomers)	31	5.4–17.5		0.73
dimethylundecane isomer #1	32	0.8–3.3		0.6
dimethylundecane isomer #2	33	0.8–2.6		0.6
C13 branched alkanes (2 unknown isomers)	34	2.3–5.8		0.6
C14 branched alkanes (6 unknown isomers)	35	4.4–11.3		0.55
C16 branched alkane (unknown)	36	1.3–3.1		0.47
cyclopentane	37	36.7–164.5	4.14	2.39
methylcyclopentane	38	57.4–315.3	9.24	2.19
cis-1,3-dimethylcyclopentane	39	14.8–100.1	5.09	1.94
trans-1,3-dimethylcyclopentane	40	16.4–177.7	7.70	1.94
ethylcyclopentane	41	7.9–44.4	1.89	2.01
ctc-1,2,4-trimethylcyclopentane	42	5.4–52.2	4.09	1.53
ctt-1,2,4-trimethylcyclopentane	43	1.7–15.5	1.29	1.53
Unknown methylethylcyclopentane	44	0.7–4.3		1.6
iso-propylcyclopentane	45	1.1–5.9	0.35	1.69
n-propylcyclopentane	46	2.1–10.0	0.56	1.69
cyclohexane	47	27.5–154.0	6.10	1.25
methylcyclohexane	48	20.4–147.0	7.17	1.70
cis-1,3- & 1,1-dimethylcyclohexane	49	4.6–38.4	2.91	1.4
trans-1,2-dimethylcyclohexane	50	4.6–42.4	3.27	1.41
trans-1,3-dimethylcyclohexane	51	2.9–17.8	0.91	1.52
cis-1,2-dimethylcyclohexane	52	1.9–9.8	0.51	1.41
ethylcyclohexane	53	4.8–31.9	2.31	1.47
ccc-1,3,5-trimethylcyclohexane	54	1.0–6.6		1.15
1,1,3-trimethylcyclohexane	55	2.0–20.4	2.26	1.19
1,1,4-trimethylcyclohexane	56	1.1–8.8		1.2
ctt-1,2,4- & cct-1,3,5-trimethylcyclohexane	57	0.7–3.9		1.2
ctc-1,2,4-trimethylcyclohexane	58	1.2–9.6		1.2
1,1,2-trimethylcyclohexane and isobutylcyclopentane	59	0.7–2.0		1.3
methylethylcyclohexane isomer #1	60	0.8–4.5	0.32	1.4
methylethylcyclohexane isomer #2	61	0.7–3.7	0.27	1.4
iso-propylcyclohexane	62	0.9–5.2		1.3
n-propylcyclohexane	63	2.9–15.5		1.29
unidentified C10 cyclohexane	64	2.5–7.8		1.07
unidentified C10 cyclohexanes	65	0.7–2.7		1.07
unidentified C9 cycloalkane	66	1.2–11.0	1.23	1.36
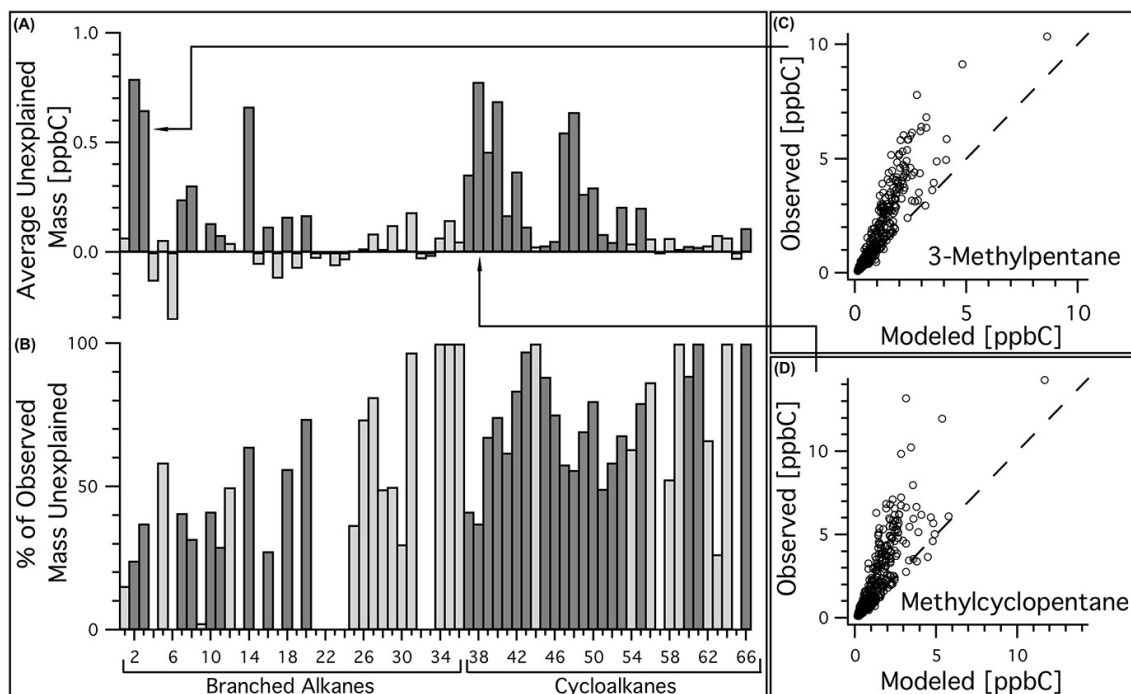
Figure 4. Many branched and cyclic alkanes exceeded predicted concentrations based on source profiles for motor vehicles. **(A–B)** the average unexplained concentration of each compound and the percentage of unexplained mass out of total observed mass. Compounds that are well correlated ($r \geq 0.75$) with the petroleum gas source are shown with shaded bars. A few compounds have negative residuals. **(C–D)** Examples of exceedances of observed over-predicted values are shown with a 1 : 1 line.
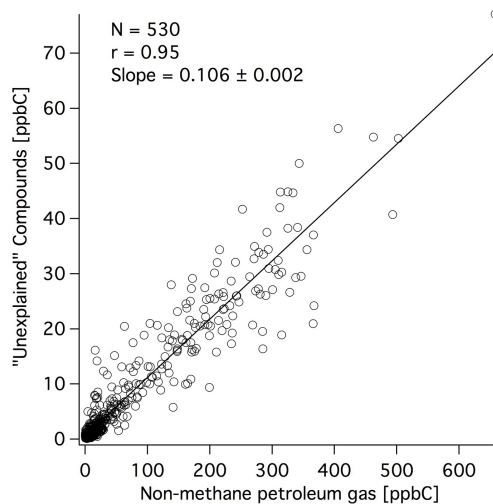
Figure 5. The sum of unexplained compounds was very well correlated with gas-phase emissions from the modeled petroleum operations source with a slope of 0.106. This increases emissions by 10.6 % from the original profile.

ranging from 0.6 to 1.6 $\mathrm{gO_3\,g^{-1}}$, but their excess concentrations after modeling were minimal – average values from 0 to 0.15 ppbC each (Fig. 4). Work by Liu et al. (2012) and Chan et al. (2013) at CalNex-Bakersfield inferred a source

of higher molecular weight organic carbon, potentially from petroleum operations, but we did not observe any significant correlation with their data.

Unrefined thermogenic wet gas is largely comprised of methane when extracted at the wells. Yet, at the Bakersfield ground site observations of methane and contributions from the petroleum operations source were not well correlated (Supplement Fig. S5). Additionally, the potential methane emissions expected based on the thermogenic wet gas source profile (Table 1) would exceed all of the observed methane enhancements above background concentrations by over 30 %. Despite the absent methane emissions, the large source of hydrocarbons is well modeled by the source profile from unrefined thermogenic wet gas in the San Joaquin Valley when using propane and larger compounds.

We compared the relative ratios of hydrocarbons in the thermogenic wet gas profile data to regression slopes of in situ data and canister data to further explore emissions from petroleum operations using ethane and isobutane, which were not available in our in situ data. The light alkanes discussed here were very well correlated in measurements from Bakersfield. Regressions with C_5 and larger compounds have more scatter due to emissions from gasoline-related sources, so they are excluded here and addressed using the source receptor model (example in Supplement Fig. S2). For the light alkanes, which have relatively minimal contributions from

Table 3. Observed light alkane ratios (gC gC^{-1}) from this and other studies.

	Data Source	ethane/ propane	propane/ n-butane	n-butane/ isobutane	ethane/ n-butane	Sample size (N)
Unrefined SJV thermogenic wet gas [±std. err. (±std. dev.)]	This study/USGS	1.2 ± 0.2 (±1.2)	2.9 ± 0.7 (±4.6)	1.7 ± 0.4 (±2.8)	3.4 ± 0.8 (±5.3)	49
Bakersfield ambient canister measurements [±std. dev. (r)]	This study	0.6 ± 0.03 (r = 0.93)	1.8 ± 0.1 (r = 0.98)	1.7 ± 0.04 (r = 0.99)	1.1 ± 0.1 (r = 0.90)	46
Bakersfield ambient in situ measurements [±std. dev. (r)][a]	This study	–	1.9 ± 0.01 (r = 0.98)	–	–	693
Colorado Front Range ambient in situ measurements [±std. err. (±std. dev.)][c,d]	Gilman et al. (2013)	0.86 ± 0.06 (±1.41)	1.5 ± 0.1 (±2.6)	2.3 ± 0.2 (±4.6)	1.3 ± 0.1 (±2.2)	554
Colorado Front Range ambient canister measurements [range (±std. dev.) (r)][b,d]	Petron et al. (2012)	–	1.5-1.7 (±0.01) (r ~ 1)	–	–	25+[b]
SW US (fall) ambient canister measurements [±std. err.][c,d]	Katzenstein et al. (2003)	1.1 ± 0.2	1.7 ± 0.4	2.2 ± 0.5	1.9 ± 0.4	85
SW U.S. (spring) ambient canister measurements [±std. err.][c,d]	Katzenstein et al. (2003)	1.4 ± 0.1	2.0 ± 0.3	2.0 ± 0.3	2.9 ± 0.4	261
East Texas Condensate Tanks [±std. err. (±std. dev.)]	Hendler et al. (2009)	0.64 ± 0.04 (±0.20)	1.3 ± 0.1 (±0.4)	1.9 ± 0.2 (±0.8)	0.78 ± 0.07 (±0.33)	24

Notes:
– Comparison done using C$_4$ alkanes and smaller as there are large contributions/interference from motor vehicle sources for C$_5$ and larger compounds at Bakersfield.
– Standard error (aka standard deviation of the mean) is reported as the primary uncertainty for the unrefined natural gas profile and others where appropriate, and represents the variability of the average within large highly variable data sets. Further information on statistical definitions/differences can be found in Altman and Bland (2005). Both the standard error and deviation are provided so the reader can judge the uncertainty and variability.
– Results of positive matrix factorization (PMF), and similar studies are excluded from this comparison (Peischl et al., 2013; Buzcu and Fraser, 2006; Leuchner and Rappengluck, 2010).
[a] Measurements of ethane and isobutane were unavailable from Bakersfield in situ data.
[b] Range of 5 data regressions, each with 25 or more samples and very small uncertainty. Other regressions were not reported in Petron et al. (2012).
[c] Ratios calculated from mean mixing ratios and their standard deviations, with propagation of uncertainty.
[d] Studies focused on regions with large oil and gas operations.

motor vehicles at the site, we compare ratios between atmospheric data and the source profile expected for petroleum operations (Table 1) with the results summarized in Table 3. Ratios of n-butane to isobutane strongly support the conclusion of a petroleum operations source, as they are identical with 1.7 ± 0.4 and 1.7 ± 0.04 (r = 0.99) in the oil well data and in canister measurements from Bakersfield, respectively. The process of methane separation from the associated petroleum gas can remove a fraction of very light alkanes (i.e., C$_{2-3}$) and affect their relative composition to other hydrocarbons in the condensate (Armendariz, 2009; Hendler et al., 2009). This is consistent with our observations of ratios involving C$_{2-3}$ alkanes. The ethane to propane ratio (gC gC^{-1}) observed via canister measurements at the Bakersfield site (0.6 ± 0.06, r = 0.93) (Supplement Fig. S1) is significantly lower than expected based on the thermogenic wet well composition in the San Joaquin Valley (1.2 ± 0.2). Similarly, the ethane to n-butane ratio is significantly lower in

the canister data (1.1 ± 0.1) relative to the unrefined gas data (3.4 ± 0.8). The propane to n-butane ratios in the in situ and canister data (1.9 ± 0.01 (r = 0.98) & 1.8 ± 0.1 (r = 0.98)) were slightly lower than in the oil well data (2.9 ± 0.7). The selective removal of ethane and propane along with methane changes the overall petroleum operations source profile observed at Bakersfield, primarily for ethane, which was not used in our source receptor model. This also results in a 33 % decrease in the propane weight fraction of the source profile. A revised source profile is shown in Table 4 with the addition of the previously "unexplained" compounds. We modified the propane content of the source profile to reflect this slight change in the propane composition relative to n-butane, and it resulted in very minor changes to the source receptor model outputs and maintained the same robust model diagnostics. The results reported throughout the paper reflect this minor change. The new source profile (Table 4) does affect the overall ozone formation potential. Including

Table 4. Observed petroleum operations source profile at Bakersfield.

Compound	wtC %
ethane	19.72
propane	34.02
n-butane	17.87
n-pentane	3.15
n-hexane	1.21
n-heptane	0.57
isobutane	10.61
isopentane	2.57
neopentane	0.70
2-methylpentane & 2,3-dimethylbutane	0.95
3-methylpentane	0.73
2-methylhexane	0.26
3-methylhexane	0.33
2,5-dimethylhexane	0.14
2,4-dimethylhexane	0.08
2,3,4-trimethylpentane & ctc-1,2,3-trimethylcyclopentane	0.71
2-methylheptane	0.12
3-methylheptane	0.17
2,6-dimethylheptane	0.18
cyclopentane	0.40
methylcyclopentane	0.89
cis-1,3-dimethylcyclopentane	0.49
trans-1,3-dimethylcyclopentane	0.74
ethylcyclopentane	0.18
ctc-1,2,4-trimethylcyclopentane	0.39
ctt-1,2,4-trimethylcyclopentane	0.12
iso-propylcyclopentane	0.03
n-propylcyclopentane	0.05
cyclohexane	0.58
methylcyclohexane	0.69
cis-1,3- & 1,1-dimethylcyclohexane	0.28
trans-1,2-dimethylcyclohexane	0.31
trans-1,3-dimethylcyclohexane	0.09
cis-1,2-dimethylcyclohexane	0.05
ethylcyclohexane	0.22
1,1,3-trimethylcyclohexane	0.22
methylethylcyclohexane isomer #1	0.03
methylethylcyclohexane isomer #2	0.03
unidentified C9 cycloalkane	0.12

Notes:
– Source profile carbon fraction is 0.82.
– Uncertainties are defined as standard errors and conservatively ($\pm 20\%$) mainly due to the variability in the oil well data.

these "new" compounds increases the ozone forming potential of the reported petroleum operations source profile to $0.82\,\mathrm{gO_3\,g^{-1}}$, due to the addition of more reactive cycloalkanes and branched alkanes to the initial source profile (Table 1).

The successful modeling of these emissions using the source profile constructed from well data and the consistency of hydrocarbon ratios between wells and Bakersfield measurements (canisters and in situ data) contributes to the strong evidence of emissions from petroleum operations. Overall, our results infer that the VOC source characterized and classified as petroleum operations in this analysis

is not a major source of methane in this region. In many cases, methane emissions are coincident with emissions of non-methane hydrocarbons at petroleum extraction or processing sites due to either co-emission from the same equipment/reservoir or co-located emission pathways at the same facility (Katzenstein et al., 2003; Petron et al., 2012; Gilman et al., 2013). For comparison, we include light hydrocarbon ratios from other relevant studies in Table 3. Given regional variability in oil/gas deposit composition, direct extrapolation between regions should only be done with careful attention to compositional differences in wells and other fuels, especially in urban areas where there are numerous sources of light hydrocarbons. Despite this expected heterogeneity, ratios are similar between most of the studies within the calculated uncertainties. The consistency between ratios of ethane to propane and n-butane between our ambient measurements and condensate tank samples in Hendler et al. (2009) supports the case for emissions from condensate storage tanks or associated equipment. Our observation of a major petroleum operations source with minimal coincident methane is consistent with composition measurements of condensate storage tank emissions in two Texas-based studies. The tanks contain the separated non-methane liquids and emissions were dominated by non-methane hydrocarbons (Armendariz, 2009; Hendler et al., 2006). The studies demonstrated that condensate tanks emit 4–6 times more VOCs than methane, whereas all other emission pathways emit 3–15 times more methane than VOCs, and methane was on average only 15 ± 11 wt % of 20 vent gas samples from condensate tanks (Armendariz, 2009; Hendler et al., 2009).

Similar results can also be found in positive matrix factorization (PMF) studies in the urban area of Houston, a prominent region for petroleum imports and refining. They reported considerable emissions attributed to oil/gas operations and petrochemical production of other chemicals (Leuchner and Rappengluck, 2010; Buzcu and Fraser, 2006). One evident source, termed oil/natural gas evaporation from refineries, was comprised of C_{2-7} straight and branched alkanes, as well as cyclopentane, cyclohexane, and methylcyclopentane. In Leuchner and Rappengluck (2010), a similar source accounted for 27 % of observed VOC mass at the urban site outside of the Houston shipping channel, and resulted in atmospheric concentrations ranging from 10–40 ppbC diurnally from that source.

The good agreement of the observed non-methane hydrocarbon source profile with the measured composition of associated gas in oil wells (accounting for the selective reduction of C_{2-3} alkanes) suggests that emissions occurs via a pathway involving volatile non-methane components separated from thermogenic wet gas. This is very likely a fugitive emission pathway(s), occurring predominantly after methane separation, during the extraction, storage, or processing of crude oil, associated gas, or condensate. In 2012 and 2013, California issued targeted standards to reduce emissions of VOCs and methane from oil and natural gas operations.

Table 5. Interquartile range ($Q_{25} - Q_{75}$) at Bakersfield shown with the source profile of dairy operations (determined using ground site Bakersfield data and aircraft measurements in the San Joaquin Valley), and the ozone formation potential (MIR) of individual components.

Compound	IQR [ppbv]	wt %	MIR [gO$_3$ g^{-1}]	% of Observed concentrations from dairy operations during CalNex-Bakersfield [avg. (range)]
Methane	1950–2380	93.3	0.014	–
Methanol	9.5–25.5	1.4	0.67	27 % (22–37 %)
Ethanol	3.9–14.3	4.9	1.57	45 % (18–67 %)
Acetic acid	0.79–2.5	0.45	0.68	28 % (11–44 %)

Note:
There are potential contributions from other organic compounds (e.g., carbonyls, larger alcohols, acids, alkenes).
Based on our data, they are either minor or much more reactive than measured species, as they could not be
apportioned with significance in ambient measurements. Nevertheless, there are potentially other compounds emitted
from dairy operations that have high ozone formation potential.

These efforts to control VOCs are primarily directed at storage tanks and other relevant equipment, with a focus on emissions during production and transmission from equipment that stores and handles crude oil or condensate, and effective control technologies (California Air Resources Board, 2012, 2013). Spatial mapping of emissions in Sect. 3.3 suggests an area source with a similar distribution to oil wells in the San Joaquin Valley.

The results of this section along with the following sections form and augment the conclusion that the vast majority of methane enhancements observed in the San Joaquin Valley are due to emissions from dairy operations. In particular, Sect. 3.3 shows the statistical source footprint of emissions from petroleum operations in stark contrast to both the statistical source footprint of methane emissions and the spatial distribution of methane concentrations measured via aircraft in California's Central Valley with large spikes over areas with high concentrations of dairies. It is very possible that there are emissions of methane in the San Joaquin Valley from other petroleum operations that are downstream from our observed source, perhaps related to natural gas marketing. The results of this study infer that these emissions are minor compared to dairy operations, and are predominantly not co-located with our characterized petroleum operations source.

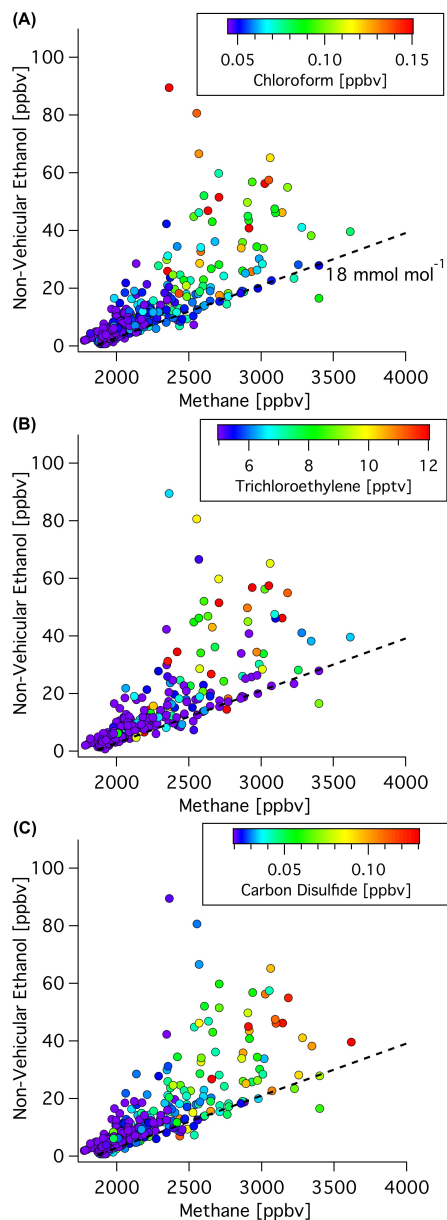## 3.2 Emissions from dairy operations

We observed evidence for substantial emissions from dairy operations in the San Joaquin Valley. These emissions, unlike the petroleum operations source, were dominated by small alcohols, acetic acid, and methane. Concentrations of the major non-methane organic compounds – methanol, ethanol, and acetic acid (average and interquartile range concentrations in Table 5) – are higher than previous measurements at other locations. Compared to another urban ground site in Pittsburgh during summer 2002 (Millet et al., 2005), the ethanol and methanol interquartile ranges and geometric
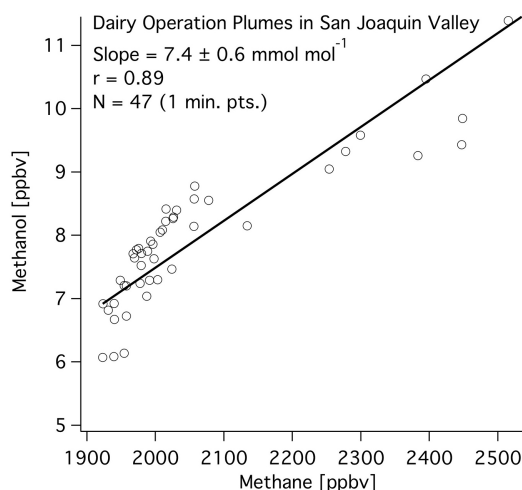
Figure 6. Methanol and methane concentrations are well correlated in dairy operation plumes sampled via aircraft (flight dates: 5/7, 6/14, 2010). Ratios of methanol to methane average 7.4 ± 0.6 mmol mol^{-1} and range up to 16 mol mol^{-1} due to the heterogeneity in emission pathways at dairy operations. Note: the data shown here represent a subset of dairies in the valley.
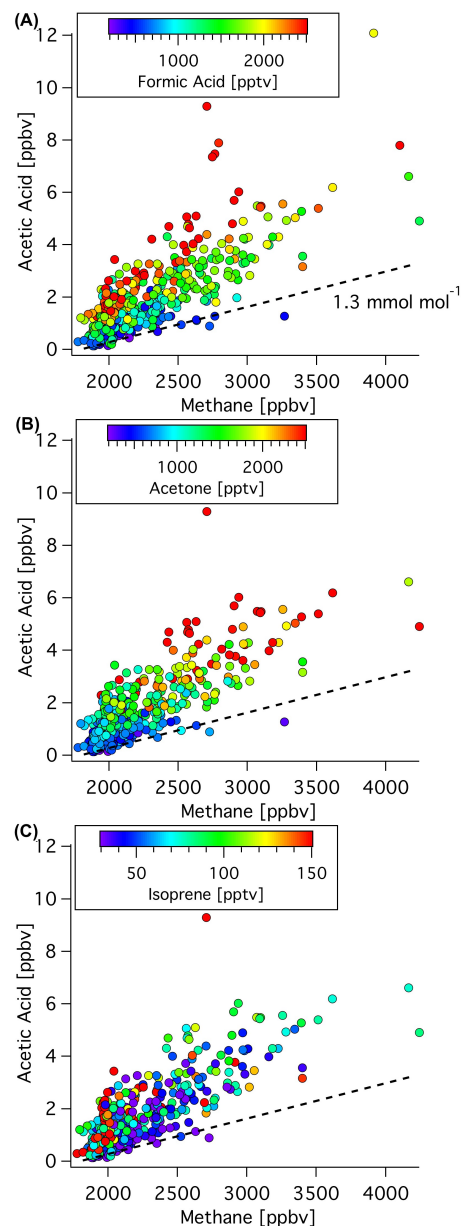
means were greater in Bakersfield, by approximately 300 % and 50 %, respectively. Despite the larger human population of the South Coast air basin, nighttime geometric means were 70 % and 240 % greater in Bakersfield compared to coincident measurements at Pasadena, CA (CalNex) for ethanol and methanol, respectively. The mean and median ethanol concentrations at the urban Bakersfield site were 12.8 and 7.6 ppbv, respectively. These values are several times greater than observations of urban and continental ethanol mixing ratios globally, as reported by Kirstine et al. (2012). However, a comparison of methanol concentrations is within the typical range of observed values globally (Heikes et al., 2002).

The methanol to methane emission ratio in dairy operation plumes measured on the aircraft was 7.4 ± 0.6 mmol mol^{-1}

Figure 7. Observations of non-vehicular ethanol vs. methane are correlated and shown with the inferred emission ratio from dairy operations. Enhancements of ethanol from a source other than the dominant source of methane and ethanol are shown by enhancements in **(A)** chloroform, **(B)** trichloroethylene, and **(C)** carbon disulfide. No major enhancements of methane are observed beyond the inferred slope with non-vehicular ethanol.

Figure 8. Acetic acid vs. methane shown with the inferred acetic acid : methane emission ratio from dairy operations. Acetic acid exceedances above the emission ratio are due to other sources of acetic acid coincident with emissions of **(A)** formic acid, **(B)** acetone, and **(C)** isoprene.

(aka ppbv ppmv^{-1}); this slope of the regression ($r = 0.89$) is nearer to the lower limit of the 7–16 mmol mol^{-1} range in the plumes (Fig. 6). This ratio was constructed from multiple transects and shows a range of ratios indicating some near-source variability in emissions from the different pathways of emissions. This ratio could be improved by collect-

ing a larger sample size of data from more locations in future source characterization studies.

Ground site ethanol and acetic acid data were compared to methane to determine their emission ratios with close attention to enhancements from other sources. For ethanol and somewhat for acetic acid, there is a clear slope that emerges (Figs. 7–8) against methane with occasional enhancements in ethanol or acetic acid that are coincident with high concentrations of tracers for other sources. In contrast, there were

no enhancements in methane concentrations past these baseline slopes in the data. This is indicative of a singular major source of methane that is clearly related to ethanol and acetic acid. This result, along with the results of Sect. 3.3 showing the agreement of dairy locations with the spatial distribution of concentrations (measured via aircraft) and the statistical source footprint of both methane and ethanol, supports the conclusion that dairies are the predominant source of methane in the San Joaquin Valley and emissions from petroleum are minor in comparison. To calculate emission ratios, data points with enhancements due to other sources (determined and shown by correlation with other tracer compounds) were not considered in the emission ratio assessment. Thus isolating the ethanol and acetic acid associated with dairy operations. With dairy (and other cattle) operations responsible for the vast majority of methane emissions observed at the Bakersfield site, the emission ratios of ethanol and acetic acid to methane are effectively the lower limit of slopes versus methane when enhancements from other sources of ethanol or acetic acid are at their minimum.

At the Bakersfield ground site, concentrations of non-vehicular ethanol (calculated via the source receptor model) were well correlated with methane, except for outliers with enhancements in ethanol that were coincident with large enhancements in tracers of other ethanol sources (Fig. 7). Other potential sources of alcohols and oxygenated gas-phase organic carbon are wastewater treatment, vegetation, soil processes, motor vehicles, and landfill/composting facilities. At low concentrations of these tracers, non-vehicular ethanol and methane are very well correlated with a slope of $18\,\mathrm{mmol\,mol^{-1}}$. Chloroform, trichloroethylene, and carbon disulfide correlate with different points that deviate from the emission ratio, suggesting multiple other minor sources of ethanol.

The results of the acetic acid versus methane assessment (Fig. 8) at the Bakersfield ground site produced similar results to that of non-vehicular ethanol versus methane. The enhancements of acetic acid above the emission ratio slope coincided with tracers of other primary and secondary sources. We calculated an emission ratio for acetic acid to methane of $1.3\,\mathrm{mmol\,mol^{-1}}$. This value represents a lower limit of acetic acid emissions associated with dairy operations. There is remaining uncertainty in this emission ratio and, based on the data shown in Fig. 8, the ratio of acetic acid to methane could be up to 50 % greater. The diurnal profile of acetic acid also suggests emissions from local/regional sources since concentrations are at their maxima during the night when emissions accumulate in the nocturnal boundary layer with minimal horizontal or vertical dilution. The results of our study show that there are high concentrations of acetic acid that are associated with methane, formic acid, acetone, or isoprene. This indicates that there are multiple major biogenic and anthropogenic sources of acetic acid in the San Joaquin Valley.

Rice cultivation could also be an important source of light alcohols and methane (Peischl et al., 2012), but there is little rice cultivation in the San Joaquin Valley. The bulk of Californian rice cultivation is located in the Sacramento Valley – the northern portion of California's Central Valley. In the San Joaquin Valley, emissions from dairy operations should far outweigh those from rice cultivation. This work is focused on sources in the San Joaquin Valley, but data from aircraft canister measurements suggest that dairy operations and rice cultivation have different emission ratios of ethanol to methanol (Supplement Fig. S6). In general, observations between the two valleys are heavily influenced by the major source that dominates in each air basin (Figs. 13, S11).

Constructing an overall source profile for dairy operations is difficult since methane, light alcohols, and acetic acid all have different emission rates from specific source pathways at dairies. Previous studies report that methane emissions are minimal from animal waste and greatest from enteric fermentation in cows, whereas emissions of non-methane gas-phase organic carbon come predominately from animal feed, followed by waste, with minor contributions from the animals themselves (Chung et al., 2010; Howard et al., 2010b; Shaw et al., 2007; Sun et al., 2008). Further variability is introduced by factors such as feed composition, temperature, and specifics of feed and waste handling. Table 5 summarizes the average regional source profile for dairy operations, determined via downwind sampling of a large collection of individual farms/feedlots in the San Joaquin Valley. Comparison against other studies is limited by the lack of a similar set of compounds. Previous studies report high emission rates for a selection of the primary compounds in Table 5, but there is no full set for comparison, and other work is focused on singular emission pathways rather than the overall source profile. Extrapolation to other regions must be done with caution, as the emission ratios reported here are region specific. So here we compare our results to other studies to the extent that it is possible.

In this and other studies, emissions of ethanol are typically greater than methanol, ranging $1.3\text{–}2.4\,\mathrm{mol\,mol^{-1}}$. Based on the literature and our results, it is apparent that the ratios of the two main alcohols to methane can vary depending on the relative amount of animals versus feed and waste, and the specifics of feed/waste storage and processing. Our reported ratios represent the average for the region; the ratio of ethanol to methane reported by Sun et al. (2008) for lactating cows and waste ($24\,\mathrm{mmol\,mol^{-1}}$) is slightly higher than our value ($18\,\mathrm{mmol\,mol^{-1}}$). Their ratio of methanol to methane ($19\,\mathrm{mmol\,mol^{-1}}$) was greater by 150 %, but is within the range observed in our analysis of aircraft data. The differences between results can potentially be attributed to the absence of feed, which will increase alcohol emissions. Measurements of acetic acid are less common so there are few studies to compare emission ratios. Shaw et al. (2007) reported ratios of acetic acid to methanol ranging from 0.05 to $0.94\,\mathrm{mol\,mol^{-1}}$ for cows and their waste. In this work, we observed a ratio of $0.18\,\mathrm{mol\,mol^{-1}}$.

Emissions of other compounds have been reported from dairy and other livestock operations, most in relatively minor quantities compared to the dominant compounds presented in this work. There are likely small emissions of low molecular weight aldehydes (e.g., propanal, butanal), ketones (e.g., acetone), other alcohols (e.g., propanol, phenols), alkenes, and esters (e.g., propyl acetate, propyl propionate) from dairy operations (Chung et al., 2010; Howard et al., 2010b; Malkina et al., 2011). In general, a major source of many oxygenated species is secondary production from the chemical oxidation of other compounds. The measurements used in this study similarly suggest substantial contributions from secondary production for many of the measured carbonyls and acids. At the ground site and from the aircraft, emissions of many of these carbonyls from dairy operations could not be detected due to the magnitude of other sources, and there were no measurements of esters or larger alcohols. In this study, dairy operation emissions of these minor compounds (acetone, methyl ethyl ketone, propanal, butanal, and other oxygenated VOCs measured at the Bakersfield site) make only minimal contributions to total emissions of these compounds on a valley-wide basis. One potential exception is acetaldehyde; previous work reported emissions equivalent to 20–110 % of ethanol emissions from feed and relatively minor emissions from cows and their manure (Makina et al., 2007; Shaw et al., 2007). In this study, no significant correlation was observed between acetaldehyde and methane in the dairy plumes measured by aircraft, and insufficient data exist from the ground site to check for emissions of acetaldehyde. Also, neither methyl ethyl ketone nor acetone were well correlated ($r = 0.55$–0.65) with methane in the dairy plumes measured by the aircraft. Other studies on volatile organic acids have also reported emissions of propanoic acid and butanoic acid with relative emission rates ranging from an order of magnitude below acetic acid to the same order of magnitude as acetic acid (Alanis et al., 2010; Shaw et al., 2007; Sun et al., 2008). We did not measure propanoic or butanoic acid, but we did not observe any correlation between measured concentrations of either formic or oxalic acid and the prominent compounds emitted from dairies at the Bakersfield ground site. Based on our work and the literature, acetic acid appears to be the most prominent acid emitted by dairy operations.

One of the objectives of this study was to provide a basic source profile, averaged over the bulk of dairy operations in the San Joaquin Valley with the understanding that the profile can potentially vary between individual operations. Methanol, ethanol, and acetic acid were the predominant non-methane compounds emitted from dairy operations. Figure 9 shows comparisons of the concentrations of these compounds attributed to dairy operations versus the total observed concentrations for each hourly sample in Bakersfield. The percentage of each compound from dairies ranged widely with some significant diurnal patterns (Supplement Fig. S8). On average, 27 % of observed methanol was from dairies with hourly averages ranging diurnally 22–

37 %. 28 % of observed acetic acid was from dairies with a diurnal range of 11–44 %. As mentioned previously, the emission ratios for methanol and acetic acid are conservative estimates that may tend towards lower limits. In this case, the fraction of methanol and acetic acid from dairy operations will increase slightly, but since ethanol makes up a dominant fraction of the non-methane source profile (Table 5) these changes will have a negligible impact on the overall source profile and implications of dairy operations on air quality in the valley (Sect. 3.4). Due to the increased use of gasoline, 9.6 ± 5.8 % of ethanol was emitted by gasoline-related sources. Of the remainder, 48 % was from dairy operations on average with a diurnal range of 30–71 %.

The diurnal average of the percent contribution from dairy sources (Supplement Fig. S8) shows minima during the daytime for acetic acid and non-vehicular ethanol. These ratios vary widely with time of day and meteorology. This daytime minimum can be attributed in part to biogenic emissions of ethanol when emissions from natural vegetation and agriculture are likely highest. For acetic acid, the minimum is likely due to secondary production from the oxidation of isoprene and other reactive precursors. Methanol did not have a strong diurnal pattern, since other major day and nighttime sources have similar emission patterns (e.g., vegetation). The remaining methanol observed at the Bakersfield site can be attributed to a mix of emissions from anthropogenic urban sources, natural vegetation, and biogenic emissions from agriculture. A recent study by Hu et al. (2011) found that 90 % of methanol was biogenic during the summer in the Midwestern US, with the remainder being anthropogenic. Heikes et al. (2002) reports a similar value with primary biogenic emissions responsible for 81 % of non-oceanic emissions. Dairies are an important source of methanol in the San Joaquin Valley along with emissions from agriculture and natural vegetation. The methods used in these studies to allocate emissions will determine whether dairy (and other cattle) operations are categorized as biogenic or anthropogenic sources. In this work we consider emissions from dairy operations to be anthropogenic, similar to the CARB inventory.

Pusede et al. (2014) found that daytime average concentrations of light alcohols, aldehydes, and acids at the Bakersfield site increased with daily maximum temperature. It is possible that increases in ambient temperature could lead to increases in silage emissions due to enhanced volatilization of some compounds (e.g., alcohols), which would change the reported source profiles slightly. Yet, ethanol was the most prominent non-methane compound in our source profile and results from Pusede et al. (2014, Table A2) show that daytime averages of ethanol did not increase between moderate and high temperatures. So, we do not expect major changes with temperature for the dairy source profile reported in this work and recommend further research to identify other high-temperature sources of oxygenated compounds.

Table 6. Quartiles [ppbC] for ambient concentrations from major petroleum-based sources measured at the Bakersfield site (does not include methane emissions) shown with average maximum incremental reactivity (MIR) values and secondary organic aerosol (SOA) yields for each source.

	Q_{25}	Q_{50}	Q_{75}	MIR [$gO_3\,g^{-1}$]	SOA yield [$gSOA\,g^{-1}$]
Gasoline exhaust	12	21	35	4.5	0.023 ± 0.007
Diesel exhaust	15	28	54	2.5	0.15 ± 0.05
Non-tailpipe gasoline	4.1	8.1	18	2.0	0.0024 ± 0.0001
Petroleum gas source	7.6	20	89	0.82	~ 0
Dairy operations	5.7	11	26	1.3	~ 0

Note:
– Gasoline and diesel exhaust include both emissions of unburned fuel and products of incomplete combustion. MIR and SOA yield values for motor vehicle sources shown for comparison from Gentner et al. (2013) and Gentner et al. (2012) for comparison.
– Dairy operations include other cattle farming in the San Joaquin Valley, and the MIR value is for the NMOC fraction of the source profile.
– The average ozone formation potential (MIR) value is potentially an underestimate due to other organic compounds emitted, which may also impact the SOA formation potential (see Table 5 note).

3.3 Spatial distribution of sources

Using FLEXPART-WRF meteorological data and methods, distributions of back-trajectories were calculated for 6 and 12 h prior to arrival and measurement at the Bakersfield site. Overall averages, as well as day and nighttime averages, are shown for the entire campaign in Fig. 1. The influence of local emissions near the site is important at all times. Daytime measurements are largely impacted by transport from the north-northwest due to consistent up-valley flows during the day. In contrast, at night the wind speeds and direction are more variable and irregular with flows that arrive from all directions, but originate as up-valley flows from the north-northwest. Extensive reviews of meteorology and flow patterns in the San Joaquin Valley found elsewhere are consistent with the results presented in this work (Bao et al., 2007; Beaver and Palazoglu, 2009). The footprint analysis used in this study provides a good representation of the distribution of surface-level areas that influence parcels' contact with the surface layer and associated sources, but potentially has some uncertainty given the complexities of Bakersfield meteorology (Angevine et al., 2013).

Statistical meteorological modeling using ground site data resulted in a spatial distribution of petroleum gas emissions similar to that of oil wells in the southern San Joaquin Valley (Fig. 10). Additionally, canister samples taken via aircraft in the region show higher propane (a major component of the source profile) concentrations for some points in the southern part of the valley (Fig. 10c). Given the co-location of oil wells in the region and the spatial distribution of elevated concentrations of petroleum gas compounds, it is probable that the observed emissions occur at or near the wells during extraction, storage, and initial processing.

The statistical distribution of emissions of non-vehicular ethanol and methane were similar for both 6 and 12 h back-trajectories. The map of emissions is consistent with the distribution of dairies in the San Joaquin Valley (Figs. 11, 12)

and aircraft measurements of ethanol and methane (Figs. 13, 14). While there are dairy operations within the 12 h footprint and the emitted methane and light alcohols have long atmospheric lifetimes, the dairies within the 6 h footprint are much more influential on elevated concentrations, especially at night. The spatial distributions of petroleum and dairy operation emissions clearly show that they are coming from different parts of the valley. The maps in this section provide strong supporting evidence that the vast majority of methane is emitted from dairy (and other cattle) operations.

The statistical emissions mapping method developed in this paper is a useful integration of concentration-weighting trajectory methods with the FLEXPART-WRF modeling platform. This emissions mapping tool is effective at locating point and area sources, especially for prominent sources in the San Joaquin Valley. The analyses of the spatial distribution of emissions from petroleum and dairy operations shown in this work are two applications of this technique. For these purposes, either concentration data or modeling outputs (e.g., source receptor models) can be used, both of which appear in this work. Further development of this approach will continue to improve its utility and quantitative outputs, but caution must be given to the transport timescales and tracer lifetime. There is one limitation to the current version of the statistical source footprint analysis. The area of analysis is limited to the distribution of sample footprints across all runs, and there is likely insufficient data to assess areas outside that total footprint. Nevertheless, the current method is excellent for looking at the most important sources that impact an area, such as Bakersfield in this study. Coverage could be improved in other studies by using data from multiple sites in a region, but care must be exercised to ensure the data is properly weighted. Overall, this work demonstrates the efficacy and usefulness of this tool, warrants further development, and future work should apply it on regional and continental scales, as appropriate, to locate primary sources of pollution.
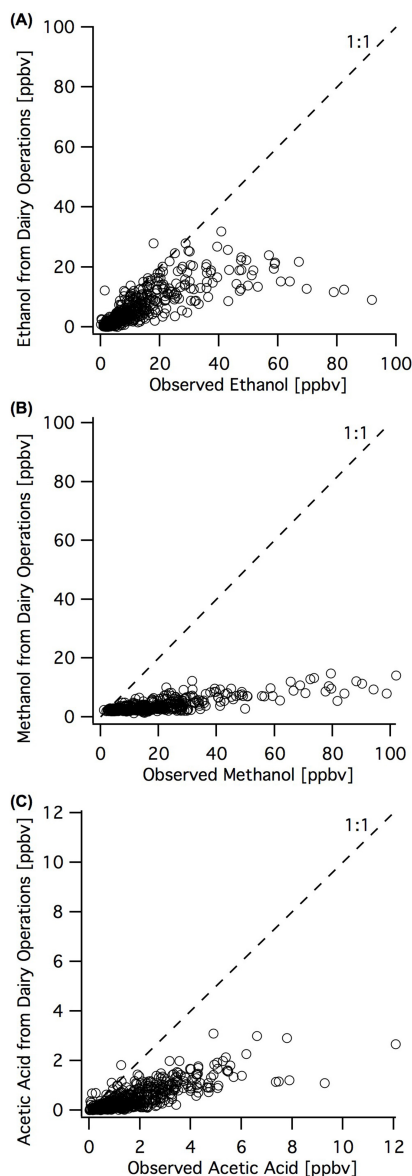
Figure 9. Estimated concentrations of non-methane organic compounds emitted by dairy operations shown against ambient observations at the Bakersfield ground site. Emissions are apportioned to dairy operations using the emission ratios of methane determined using aircraft and ground site measurements. On average, 45 % of observed **(A)** ethanol is from dairies, whereas, smaller fractions of **(B)** methanol (27 %) and **(C)** acetic acid (28 %) are from dairy operations. These fractions vary with time of day and source strength. Diurnal patterns of percent contributions from dairy operations are shown in Supplement Fig. S8.

3.4 Implications for air quality and emissions inventories

Both petroleum and dairy (and other cattle) operations are important sources of reactive organic carbon in the San Joaquin Valley. On a mass basis, observed VOC concentrations from petroleum extraction/processing were on the same order as emissions from motor vehicles. Yet, they represent a relatively minor contribution to potential ozone formation, as the average MIR value for the source ($0.82\,\mathrm{gO_3\,g^{-1}}$) is $\sim$ 3–6 times less than that of motor vehicle sources. Direct contributions to secondary organic aerosol (SOA) from the petroleum operations source profile in this study are likely to be minimal, given that the yields for all of the alkanes with eight or less carbon atoms will be $0.002\,\mathrm{gSOA\,g^{-1}}$ at most, with an organic particle loading of $10\,\mathrm{\mu g\,m^{-3}}$ (Gentner et al., 2012). The potential ozone and SOA implications of petroleum operation emissions will depend greatly upon composition, which varies between regions. We did not observe any aromatic content, but other studies have observed aromatic and other larger compound fractions (Carter and Seinfeld, 2012; Gilman et al., 2013). Aromatics have been shown to be very effective precursors to SOA and ozone (Gentner et al., 2012; Carter, 2007). So, their presence in oil/gas emissions will have further implications for air quality.

Dairy operations in the San Joaquin Valley are largely responsible for the higher than typical ethanol concentrations in the San Joaquin Valley. Based on the primary compounds observed from dairy operations (ethanol, methanol, acetic acid), we infer that emissions have minor impacts on SOA formation, but have a greater potential to impact ozone formation with an MIR of $1.3\,\mathrm{gO_3\,g^{-1}}$. The inclusion of other oxygenated compounds previously observed from dairy operations (e.g., Hafner et al., 2013) to the basic source profile in Table 5 may increase the ozone and SOA formation potential. Yet, in this study they were minor and not significantly correlated with other dairy emissions (see Sect. 3.2).

In Bakersfield during spring/summer, dairy operations were responsible for 22 % of anthropogenic non-methane organic carbon emissions and 13 % of potential anthropogenic ozone formation. Similarly, petroleum operations were responsible for 22 % of anthropogenic emissions and 8 % of potential ozone. Motor vehicles were responsible for the remaining 56 % of anthropogenic emissions, 79 % of anthropogenic potential ozone formation, and essentially all of the potential anthropogenic SOA formation. It is important to note that emissions from petroleum and dairy operations have substantial potential to impact the atmospheric chemistry leading to secondary pollution, but they themselves are not a major source of SOA precursors (note: does not consider aqueous chemical processing). These results apply to the emissions of VOCs from petroleum operations observed and characterized in this work; other recent work on petroleum operations has reported emissions of larger hydrocarbons that have higher SOA yields (Chan et al., 2013; Gilman et al., 2013). These five main sources are summarized in Fig. 15 and are very important sources for the San Joaquin Valley. There are other anthropogenic sources that likely contribute emissions on smaller urban scales that are
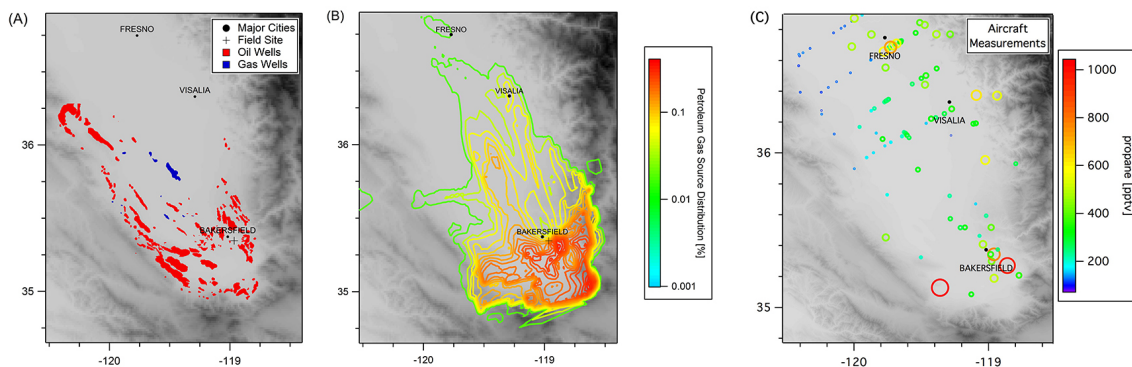
Figure 10. Maps of the southern part of the San Joaquin Valley with **(A)** the location of oil and gas wells, **(B)** the spatial distribution of petroleum gas emissions determined using statistical footprint analysis (6 h), and **(C)** aircraft canister measurements of propane, sized and colored by concentration. Together the maps show a similar distribution of wells and emissions in the region. Note: meteorological conditions, altitude, and local dilution varies between canister measurements.

Table 7. Comparison of the relative abundance of VOC emissions from each source observed in Bakersfield to CARB inventory for the San Joaquin Valley (SJV) and the portion of Kern County in the San Joaquin Valley including Bakersfield (SJV-Kern).

|	Relative mass abundance in Bakersfield (this study)	Fraction of emissions [%] in inventory (absolute emission rate [tons day^{-1}])	
		SJV inventory	SJV-Kern inventory
Petroleum operations	22 %	15 % (28)	53 % (26)
Dairy operations	22 %	30 % (57)	9 % (4.5)
On- & off-road motor vehicles	56 %	55 % (104)	38 % (19)

Notes:
– Motor vehicle emissions are sum of on- and off-road since ambient source apportionment cannot discern between them; includes gasoline and diesel exhaust, and non-tailpipe gasoline emissions.
– Comparison is limited to discussed sources, biogenic emissions and other potentially important sources are excluded (for biogenic emissions from agriculture see Gentner et al., 2014).

not enumerated in this work. Such as the contributions of biogenic sources, which are another major factor for air quality in California's Central Valley.

In the comparison of the sources discussed in this work, the percent contribution of vehicular sources is larger in Bakersfield than it would be most places in the region. In nonurban areas of the San Joaquin Valley, motor vehicle emissions will still be important, but emissions from petroleum and dairy operations will make up a greater fraction of nonmethane organic carbon in the atmosphere and will be responsible for a greater fraction of potential ozone formation. The results from Bakersfield in this study confirm the transport and importance of emissions from dairy operations throughout the San Joaquin Valley. Our results for potential ozone give a 3.5 : 1 ratio of potential ozone from gasoline vehicles to dairy operations in Bakersfield. When considering that there is a greater prevalence of motor vehicles around our measurement site and most dairy emissions are outside the

county (Table 7), the ratio is similar to the valley-wide ratio of 3 : 2 for light-duty vehicles to livestock feed modeled by Hu et al. (2012). Overall, this, and other recent work (Howard et al., 2010a; Hu et al., 2012), demonstrates that motor vehicles and multiple source pathways at dairy operations are major emitters of reactive ozone precursors throughout the San Joaquin Valley. Elevated concentrations of non-vehicular ethanol that are largely linked to dairy operations warrants further evaluation of emission processes involving livestock silage, as ethanol has been demonstrated as a primary component of those emissions (Hafner et al., 2013; Howard et al., 2010a; Malkina et al., 2011).

Our results on the relative contributions from each source indicate a mix of influential sources. Given our location in an urban area in the southern San Joaquin Valley, where oil wells are concentrated, emissions from motor vehicles and petroleum operations are likely higher than other parts of the valley. The San Joaquin Valley has an abundance of
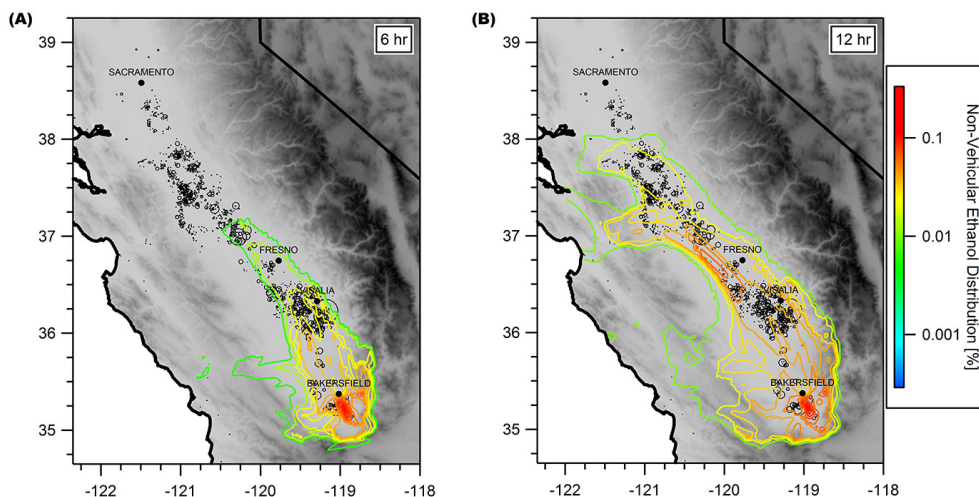
Figure 11. Statistical distribution of emissions of non-vehicular ethanol in the San Joaquin Valley shown as colored contours for 6 and 12 h footprints. Modeling results shown with the location of dairies as markers (o) scaled by the size of each dairy.
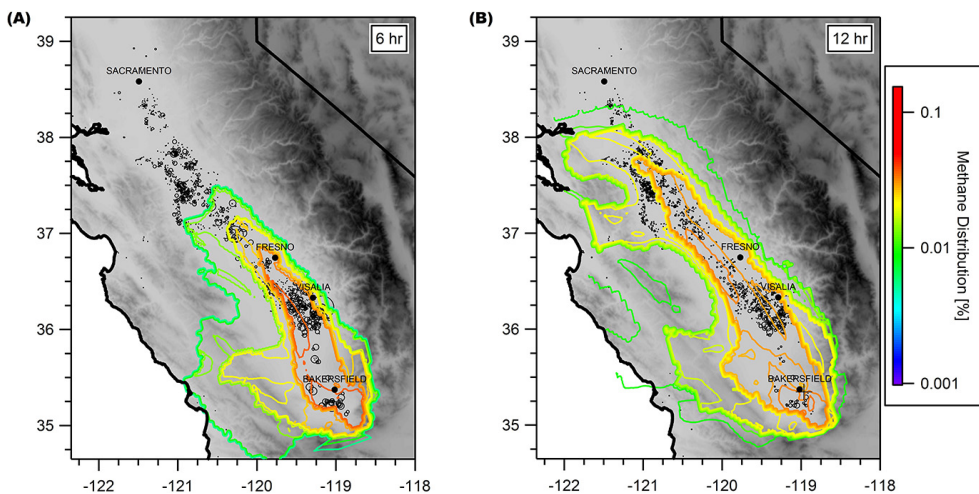

Figure 12. Statistical distribution of emissions of methane in the San Joaquin Valley shown as colored contours for 6 and 12 h footprints. Modeling results shown with the location of dairies as markers (o) scaled by the size of each dairy.

agriculture and is surrounded by natural vegetation that represents a large potential source of emissions following transport to other parts of the valley. Comprehensive modeling assessments need to evaluate the sources discussed here along with biogenic emissions of reactive organic gases from both agriculture and natural vegetation.

Comparing different assessments for emissions from multiple sources presents challenges relating to the definition of sources and spatial boundaries. Here, we provide a comparison of our relative emission magnitudes at the Bakersfield site to the CARB emission inventory for the San Joaquin Valley (Table 7). To promote consistency with our observed sources, we compare our petroleum operations source to emissions from oil/gas production and refining, and exclude petroleum marketing (and combustion from petroleum oper-

ations) since our observed source is clearly related to unrefined petroleum. While there are likely some differences in emissions, it is difficult to separate dairy cattle from other cattle, so we have assumed that we are observing all cattle in this study and include them with dairy operations. Although in the CARB inventory, dairy cattle represent almost 80 % of cattle-related emissions in the San Joaquin Valley. Similarly, we compare these sources to on- and off-road mobile sources as that is the best representation of the observed motor vehicle sources in our source apportionment.

There are potential seasonal effects among the 5 sources shown in Tables 6 and 7 and Fig. 15. The composition of gasoline fuel changes seasonally to reduce volatility by varying formulation, which affects the composition and magnitude of emissions. In the CARB almanac, VOC emissions
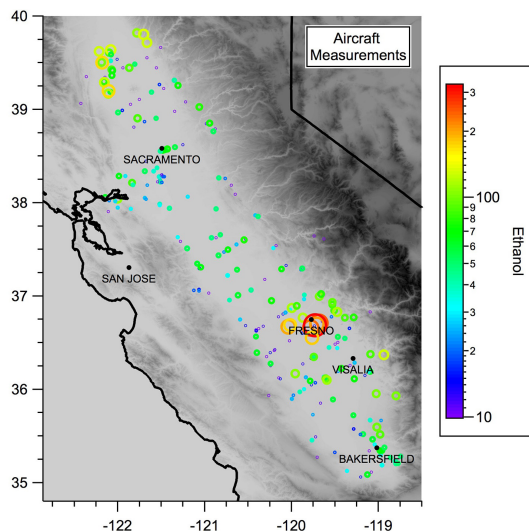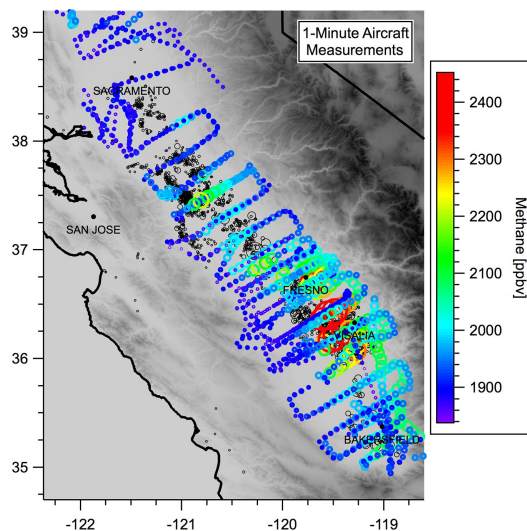
Figure 13. Aircraft canister measurements of ethanol in California's Central Valley shown as individual circles, sized and colored by ethanol concentration. Data were taken at varying altitudes above and below the boundary layer with a general filter for below 1000 m. Vertical gradients are responsible for some variability, but aircraft data support conclusions of other analyses showing large ethanol sources in the Central Valley: dairy operations in the San Joaquin Valley and rice cultivation in the Sacramento Valley. Note: meteorological conditions and local dilution varies between canister measurements. Also, alcohol measurements made using the canisters were prone to significant losses, so their use is only relative.
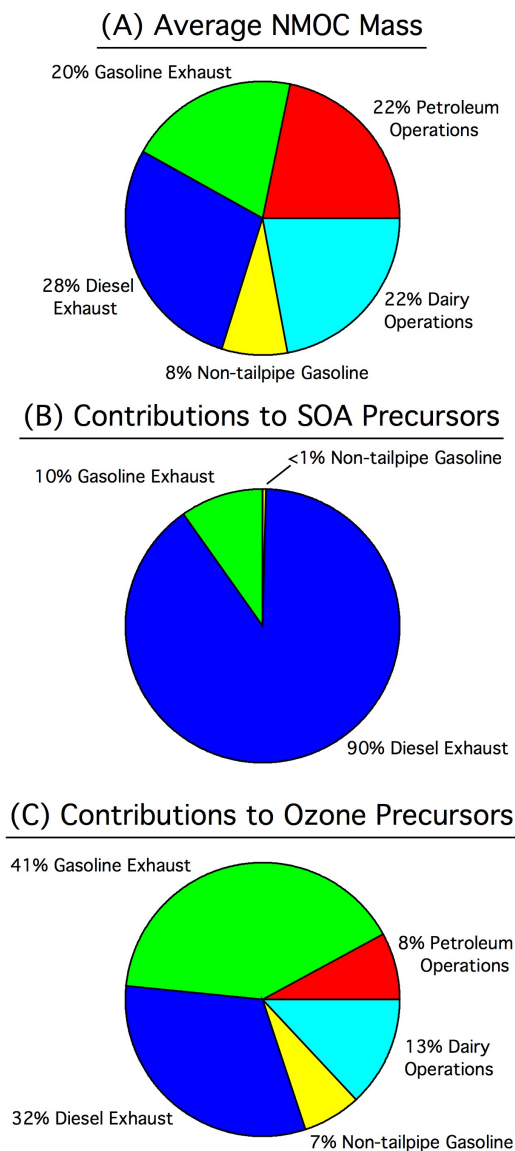

Figure 14. Map of observed methane concentrations over 7 flights in California's Central Valley shown as individual circles, sized and colored by methane concentration. Data were taken at varying altitudes above and below the boundary layer with general filter for below 1000 m. Vertical gradients and multiple flights are responsible for some variability, but methane enhancements in aircraft data show good correlation with the location of dairy operations (open black circles sized by bovine population). A map including the entire Sacramento Valley can be found in the Supplement (Fig. S11).

from dairy operations and petroleum production and refining have no seasonal change between summer and winter. The emissions we observe from both sources could be hypothesized to volatilize more in warmer weather, but we have insufficient data to assess seasonal changes and effects other than temperature may potentially play a role.

The CARB emissions inventory for the San Joaquin Valley reports an average of 28 tons ROG per day from petroleum operations (production and refining), which is equal to 27 % of on-and off-road mobile source emissions (72+32 tons per day) in the air basin (California Air Resources Board, 2010). This value is consistent with daytime ratios (18–51 %) observed at the Bakersfield site (Supplement Fig. S4) when vehicular emissions are greatest, but is smaller than nighttime ratios (62–120 %) and the overall ratio (39 %). Bakersfield is in much closer proximity to potential petroleum operations sources compared to other parts of the air basin, so nighttime ratios are significantly higher with relatively less vehicular traffic and local emissions play a larger role when there is less atmospheric dilution. A comparison on a smaller scale for the portion of Kern County in the San Joaquin Valley demonstrates the local importance of petroleum operations, as much of the San Joaquin Valley's petroleum operation emissions are in this county. For this area, petroleum production/refining emissions in the CARB inventory are equivalent

to 139 % of on- and off-road mobile sources (California Air Resources Board, 2010). These observations are consistent with the statistical footprints shown in this work as daytime footprints encompass a larger area that stretches into other counties, while nighttime footprints are more heavily influenced by local emissions.

According to the CARB emission inventory, dairy and other cattle operations in the San Joaquin Valley emit 57 tons ROG per day, which is 80 % of non-vegetation farming-related emissions (California Air Resources Board, 2010). These emissions from dairy and cattle operations are equivalent to 55 % of on- and off-road motor vehicle emissions in the inventory, which is higher than the average non-methane organic carbon (NMOC) mass comparison at the Bakersfield measurement site (40 %). The CARB inventory for the San Joaquin Valley states that emissions from dairy operations are twice those from petroleum operations (dairy & other cattle operations ROG emissions = 2.0 × oil/gas production and refining ROG emissions). The average measured contributions from petroleum and dairy sources were equivalent at the Bakersfield site (Fig. 15). This is largely dependent on the distribution of petroleum operations relative to dairy operations, which is greatest in the southern part of the San Joaquin Valley (e.g., Bakersfield) where the oil wells and related operations are concentrated. Thus, the ratio of petroleum to dairy operation contributions goes up by several factors with decreased dilution and a greater influence of local sources

(A) Average NMOC Mass

20% Gasoline Exhaust

22% Petroleum Operations

28% Diesel Exhaust

22% Dairy Operations

8% Non-tailpipe Gasoline

(B) Contributions to SOA Precursors

10% Gasoline Exhaust

<1% Non-tailpipe Gasoline

90% Diesel Exhaust

(C) Contributions to Ozone Precursors

41% Gasoline Exhaust

8% Petroleum Operations

13% Dairy Operations

32% Diesel Exhaust

7% Non-tailpipe Gasoline

Figure 15. Breakdown of the contributions of prominent anthropogenic sources in Bakersfield for **(A)** total non-methane organic carbon (NMOC) mass (g), **(B)** precursors to secondary organic aerosol (SOA), and **(C)** precursors to ozone. Other sources/compounds may impact SOA formation indirectly via changes in photochemistry. The exhaust values here include unburned fuel emissions and products of incomplete combustion, and dairy operations include other cattle farming. Biogenic emissions from natural vegetation are excluded, but are likely to have important contributions to emissions and air quality in the San Joaquin Valley, but less so in the urban core of Bakersfield, CA. Note: the NMOC mass comparison in panel A is in terms of mass (similar to inventories), so ratios of sources will be slightly different from Table 5 where they are in mol carbon.

(Table 6). This is likely also the reason for the greater contribution from motor vehicles relative to dairy operations at the Bakersfield site versus the inventory. The greater prevalence of motor vehicles near the site increases its impact relative to the whole valley.

A comparison of the dairy operations source profile (Table 5) with the CARB emission inventory reveals that the ratio of methane to NMOC is consistent between our results and the inventory, 93 % vs. 92 % methane. Additionally, the existing CARB inventory for the San Joaquin Valley reflects the difference in the magnitude of methane emissions between the two sources, with total methane emissions from dairy (and other cattle) operations being an order of magnitude greater than petroleum production operations, and responsible for at least 87 % of methane emissions. Furthermore, for petroleum operations, the majority (81 %) of fugitive methane (and ethane by inventory definition) emissions are from oil/gas marketing rather than production/refining (California Air Resources Board, 2010). Overall, these intercomparisons, while rough, provide validation of the CARB emission inventory for relative emission rates of dairy and petroleum operations in the San Joaquin Valley.

The San Joaquin Valley, and the Central Valley as a whole, contains a complex mixture of both anthropogenic and biogenic sources of reactive gas-phase organic carbon on both regional and urban scales. Our focus in this paper has been quantifying regional emissions from petroleum and dairy operations, comparing their emission rates to other anthropogenic sources, and evaluating their importance for air quality in the urban area of Bakersfield and the San Joaquin Valley. The dairy and petroleum sources are clearly relevant to air quality on both local and regional scales for ozone formation, but are likely not as important as sources of precursors to SOA. This study provides important new information, expanding knowledge on the suite of compounds emitted from these sources and providing new useful information on their sources profiles.

The Supplement related to this article is available online at doi:10.5194/acp-14-4955-2014-supplement.

not constitute endorsement or recommendation for use.

Edited by: A. Carlton

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, 5084–5092, 2010.

Altman, D. G. and Bland, J. M.: Standard deviations and standard errors, BMJ, 331, 903, 2005.

Angevine, W. M., Eddington, E., Durkee, K., Fairall, C., Bianco, L., Brioude, J.: Meteorological Model Evaluation for CalNex 2010, Mon. Weather Rev., 140, 3885–3906, 2012.

Angevine, W. M., Brioude, J., McKeen, S., Holloway, J. S., Lerner, B. M., Goldstein, A. H., Guha, A., Andrews, A., Nowak, J. B., Evan, S., Fischer, M. L., Gilman, J. B., and Bon, D.: Pollutant transport among California regions, J. Geophys. Res. Atmos., 118, 6750–6763, doi:10.1002/jgrd.50490, 2013.

Armendariz, A.: Emissions from Natural Gas Production in the Barnett Shale Area and Opportunities for Cost-Effective Improvements, available at: http://www.edf.org/sites/default/files/9235_Barnett_Shale_Report.pdf (last access: October 2012), 2009.

Atkinson, R. and Arey, J.: Atmospheric degradation of volatile organic compounds, Chem. Rev., 103, 4605–4638, 2003.

Bao J. W., Michelson, S. A., Persson, P. O. G., Djalalova, I. V., and Wilczak, J. M.: Observed and simulated low-level winds in a high ozone episode during the central California ozone study, J. Appl. Meteor. Climatol., 47, 2372–2394, 2007.

Barletta, B., Carreras-Sospedra, M., Cohan, A., Nissenson, P., Dabdub, D., Meinard, S., Atlas, E., Lueb, R., Holloway, J. S., Ryerson, T. B., Pederson, J., VanCuren, R. A., and Blake, D. R.: Emission estimates of HCFCs and HFCs in California from the 2010 CalNex study, J. Geophys. Res.-Atmos., 118, 2019–2030, 2013.

Beaver S. and Palazoglu A.: Influence of synoptic and mesoscale meteorology on ozone pollution potential for San Joaquin Valley of California, Atmos. Environ., 43, 1779–1788, 2009.

Brioude, J., Angevine, W. M., McKeen, S. A., and Hsie, E.-Y.: Numerical uncertainty at mesoscale in a Lagrangian model in complex terrain, Geosci. Model Dev., 5, 1127–1136, doi:10.5194/gmd-5-1127-2012, 2012.

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. S., Stutz, J., McKeen, S., Petetin, H., Beekmann, M., Warneke, C., Parrish, D. D., and de Gouw, J. A.: Emissions of anthropogenic VOCs in northern mid-latitude megacities: observations vs. emission inventories in Los Angeles and Paris, J. Geophys. Res., 118, 2041–2057, 2013.

Buzcu B. and Fraser M. P.: Source identification and apportionment of volatile organic compounds in Houston, TX, Atmos. Envron., 40, 2385–2400, 2006.

California Air Resources Board: The California Almanac of Air Quality & Emissions – 2009 Edition, available at: http://www.arb.ca.gov/app/emsinv/fcemssumcat2009.php (last access: May 2012), 2010.

California Air Resources Board: Overview of final amendments to air regulations for the oil and natural gas industry, available at: http://www.epa.gov/airquality/oilandgas/pdfs/20120417fs.pdf (last access: accessed March 2014), 2012.

California Air Resources Board: Final updates to requirements for storage tanks used in oil and natural gas production and transmission, available at: http://www.epa.gov/airquality/oilandgas/pdfs/20130805fs.pdf (last access: accessed March 2014), 2013.

Carter, W. P. L.: SAPRC Atmospheric Chemical Mechanisms and VOC Reactivity Scales, availableat:http://www.engr.ucr.edu/~carter/SAPRC/ (last access: August 2012), 2007.

Carter, W. P. L. and Seinfeld, J. H.: Winter ozone formation and VOC incremental reactivities in the Upper Green River Basin of Wyoming, Atmos. Environ., 50, 255–266, doi:10.1016/j.atmosenv.2011.12.025, 2012.

Chan, A. W. H., Isaacman, G., Wilson, K. R., Worton, D. R., Ruehl, C. R., Nah, T., Gentner, D. R., Dallmann, T. R., Kirchstetter, T. W., Harley, R. A., Gilman, J. B., Kuster, W. C., deGouw, J. A., Offenberg, J. H., Kleindienst, T. E., Lin, Y. H., Rubitschun, C. L., Surratt, J. D., Hayes, P. L., Jimenez, J. L., and Goldstein, A. H.: Detailed Chemical Characterization of Unresolved Complex Mixtures in Atmospheric Organics: Insights into Emission Sources, Atmospheric Processing and Secondary Organic Aerosol Formation. J. Geophys. Res., 118, 6783–6796, doi:10.1002/jgrd.50533, 2013.

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

Crounse, J. D., McKinney, K. A., Kwan, A. J., and Wennberg, P. O.: Measurement of gas-phase hydroperoxides by chemical ionization mass spectrometry, Anal. Chem., 78, 6726–6732, 2006.

de Gouw, J. A. and Warneke, C.: Measurements of volatile organic compounds in the Earth's atmosphere using proton-transfer-reaction mass spectrometry, Mass Spectrom. Rev., 26, 223–257, 2007.

US Environmental Protection Agency, EPA-CMB8.2, available at: http://www.epa.gov/scram001/receptor_cmb.htm, last access: January 2014.

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, 4247–4252, 2009.

Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W., Dallmann, T. R., Davis, L., Liu, S., Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A., Harley, R. A., and Goldstein, A. H.: Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed characterization of organic carbon emissions, P. Natl. Acad. Sci. USA, 109, 18318–18323, 2012.

Gentner, D. R., Worton, D. R., Isaacman, G., Davis, L., Dallmann, T. R. Wood, E. C. Herndon, S. C, Goldstein, A. H., and Harley, R. A.: Chemical speciation of gas-phase organic carbon emissions from motor vehicles and implications for ozone production potential. Environ. Sci. Technol., 47, 11837–11848, 2013.

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., in press, 2014.

Gilman, J. B., Lerner, B. M., Kuster, W. C., and de Gouw, J. A.: Source Signature of Volatile Organic Compounds from Oil and

Natural Gas Operations in Northeastern Colorado, Environ. Sci. Technol., 47, 1297–1305, 2013.

Hafner, S. D., Howard, C., Muck, R. E., Franco, R. B., Montes, F., Green, P. G., Mitloehner, F., Trabue, S. L., and Rotz, C. A.: Emission of volatile organic compounds from silage: Compounds, sources, and implications, Atmos. Environ., 77, 827–839, 2013.

Hendler, A., Nunn, J., Lundeen, J., and McKaskle, R.: VOC Emissions from Oil and Condensate Storage Tanks, report prepared for the Houston Advanced Research Center, available at: http://files.harc.edu/Projects/AirQuality/Projects/H051C/H051CFinalReport.pdf (last access: October 2012), 2009.

Heikes, B. G., Chang, W. N., Pilson, M. E. Q., Swift, E., Singh, H. B., Guenther, A., Jacob, D. J., Field, B. D., Fall, R., Riemer, D., and Brand, L.: Atmospheric methanol budget and ocean implication, Global Biogeochem. Cy., 16, 1133, doi:10.1029/2002GB001895, 2002.

Howard, C. J., Kumar, A., Mitloehner, F., Stackhouse, K., Green, P. G., Flocchini, R. G., and Kleeman, M. J.: Direct Measurements of the Ozone Formation Potential from Livestock and Poultry Waste Emissions, Environ. Sci. Technol., 44, 2292–2298, 2010a.

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, 2309–2314, 2010b.

Hu, J., Howard, C. J., Mitloehner, F., Green, P. G., and Kleeman, M. J.: Mobile source and livestock feed contributions to regional ozone formation in Central California, Environ. Sci. Technol., 46, 2781–2789, 2012.

Hu, L., Millet, D. B., Mohr, M. J., Wells, K. C., Griffis, T. J., and Helmig, D.: Sources and seasonality of atmospheric methanol based on tall tower measurements in the US Upper Midwest, Atmos. Chem. Phys., 11, 11145–11156, doi:10.5194/acp-11-11145-2011, 2011.

Katzenstein, A. S., Doezema, L. A., Simpson, I. J., Blake, D. R., and Rowland, F. S.: Extensive regional atmospheric hydrocarbon pollution in the southwestern United States, P. Natl. Acad. Sci. USA, 100, 11975–11979, 2003.

Kemball-Cook, S., Bar-Ilan, A., Grant, J., Parker, L., Jung, J., Santamaria, W., Mathews, J., and Yarwood, G.: Ozone impacts of natural gas development in the Haynesville Shale, Environ. Sci. Technol., 44, 9357–9363, doi:10.1021/es1021137, 2010.

Kirstine, W. V. and Galbally, I. E.: The global atmospheric budget of ethanol revisited, Atmos. Chem. Phys., 12, 545–555, doi:10.5194/acp-12-545-2012, 2012.

Klouda, G. A., Lewis, C. W., Stiles, D. C., Marolf, J. L., Ellenson, W. D., and Lonneman, W. A.: Biogenic contributions to atmospheric volatile organic compounds in Azusa, CA, J. Geophys. Res., 107, doi:10.1029/2001000758, 2002.

Leuchner, M. and Rappengluck, B.: VOC source-receptor relationships in Houston during TexAQS-II, Atmos. Environ., 44, 4056–4067, 2010.

Lillis, P. G., Warden, A., Claypool, G. E., and Magoon, L. B.: Petroleum Systems of the San Joaquin Basin Province – Geochemical Characteristics of Gas Types, Ch. 10, in: Petroleum Systems and Geologic Assessment of Oil and Gas in the San Joaquin Basin Province, California, edited by: Scheirer, A. H.,

U.S. Geological Survey, available at: http://pubs.usgs.gov/pp/pp1713/, 2007.

Liu, S., Ahlm, L., Day, D., Russell, L., Zhao, Y., Gentner, D., Weber, R., Goldstein, A., Jaoui, M., Offenberg, J., Kleindienst, T., Rubitschun, C., Surratt, J., Sheesley, R., and Scheller, S.: Secondary organic aerosol formation from fossil fuel sources contribute majority of summertime organic mass at bakersfield, J. Geophys. Res., 117, D00V26, doi:10.1029/2012JD018170, 2012.

Malkina, I., Kumar, A., Green, P. G., and Mitloehner, F.: Identification and Quantification of Volatile Organic Compounds Emitted from Dairy Silages and Other Feedstuffs, J. Environ. Qual., 40, 28–36, 2011.

Markovic, M. Z., VandenBoer, T. C., and Murphy, J. G.: Characterization and optimization of an online system for the simultaneous measurement of atmospheric water-soluble constituents in the gas and particle phases, J. Environ. Monitor., 14, 1872–1884, 2012.

Metcalf, A. R., Craven, J. S., Ensberg, J. J., Brioude, J., Angevine, W. M. M., Sorooshian, A., Duong, H. T., Jonsson, H. H., Flagan, R. C., and Seinfeld, J. H.: Black carbon aerosol over the Los Angeles Basin during CalNex, J. Geophys. Res., 117, D00V13, doi:10.1029/2011JD017255, 2012.

Millet, D. B., Donahue, N. M., Pandis, S. N., Polidori, A., Stanier, C. O., Turpin, B. J., and Goldstein, A. H.: Atmospheric volatile organic compound measurements during the Pittsburgh Air Quality Study: Results, interpretations, and quantification of primary and secondary contributions, J. Geophys. Res., 110, D07S07, doi:10.1029/2004JD004601, 2005.

Pacsi, A. P., Alhajeri, N. S., Zavala-Araiza, D., Webster, M. D., and Allen, D. T.: Regional air quality impacts of increased natural gas production and use in Texas, Environ. Sci. Technol., 47, 3521–3527, doi:10.1021/es3044714, 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.

Peischl, J., Ryerson, T. B., Aikin, K. C., Andrews, A. E., Atlas, E., Blake, D., Brioude, J., 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., 118, 4974–4990, 2013.

Petron, 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., Jr., 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, D04304, doi:10.1029/2011JD016360, 2012.

Polissar, A. V., Hopke, P. K., and Harris, J. M.: Source regions for atmospheric aerosol measured at Barrow, Alaska. Environ. Sci. Technol., 35, 4214–4226, 2001.

Pusede, S. E., Gentner, D. R., Wooldridge, P. J., Browne, E. C., Rollins, A. W., Min, K.-E., Russell, A. R., Thomas, J., Zhang, L., Brune, W. H., Henry, S. B., DiGangi, J. P., Keutsch, F. N., Harrold, S. A., Thornton, J. A., Beaver, M. R., St. Clair, J. M., Wennberg, P. O., Sanders, J., Ren, X., VandenBoer, T. C., Markovic, M. Z., Guha, A., Weber, R., Goldstein, A. H., and Cohen, R. C.: On the temperature dependence of organic reactivity, nitrogen oxides, ozone production, and the impact of emission controls in San Joaquin Valley, California, Atmos. Chem. Phys., 14, 3373–3395, doi:10.5194/acp-14-3373-2014, 2014.

Ryerson, T. B., Aikin, K. C., Angevine, W. M., Atlas, E. L., Blake, D. R., Brock, C. A., Fehsenfeld, F. C., Gao, R.-S., de Gouw, J. A., Fahey, D. W., Holloway, J. S., Lack, D. A., Lueb, R. A., Meinardi, S., Middlebrook, A. M., Murphy, D. M., Neuman, J. A., Nowak, J. B., Parrish, D. D., Peischl, J., Perring, A. E., Pollack, I. B., Ravishankara, A. R., Roberts, J. M., Schwarz, J. P., Spackman, J. R., Stark, H., Warneke, C., and Watts, L. A.: Atmospheric emissions from the Deepwater Horizon spill constrain air-water partitioning, hydrocarbon fate, and leak rate, Geophys. Res. Lett., 3, L07803, doi:10.1029/2011GL046726, 2011.

Schnell, R. C., Oltmans, S. J., Neely, R. R., Endres, M. S., Molenar, J. V. and White, A. B.: Rapid photochemical production of ozone at high concentrations in a rural site during winter, Nat. Geosci., 2, 120–122, doi:10.1038/ngeo415, 2009.

Seibert, P., Kromp-Kolb, H., Baltensperger, U., Jost, D. T., Schwikowski, M., Kasper, A., and Puxbaum, H.: Trajectory analysis of aerosol measurements at high alpine sites, in: Transport and Transformation of Pollutants in the Troposphere, edited by: Borrell, P., Borrell, P. M., Cvitas, T., Kelly, K., Midgley, P., and Seiler, W., Academic Publishing, Den Haag, 689–693, 1994.

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., 31, 1310–1316, 2007.

Sheridan, M.: California Crude Oil Production and Imports, California Energy Commission, available at: http://www.energy.ca.gov/2006publications/CEC-600-2006-006/CEC-600-2006-006.PDF, 2006.

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, 615–622, 2008.

US Energy Information Administration, U.S. Crude Oil Production by State, available at: http://www.eia.gov/dnav/pet/PET_CRD_CRPDN_ADC_MBBLPD_A.htm (last access: August 2012), 2010.

Wang, Y. Q., Zhang, X. Y., and Draxler, R. R.: TrajStat: GIS-based software that uses various trajectory statistical analysis methods to identify potential sources from long-term air pollution measurement data, Environ. Modell. Softw., 24, 938–939, 2009.