Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-861-AC1, 2017

Atmospheric Chemistry and Physics Discussions
Open Access

Interactive comment on "Quantifying alkane emissions in the Eagle Ford Shale using boundary layer enhancement" *by* G. Roest and G. Schade

G. Roest and G. Schade

gsroest@tamu.edu

Received and published: 23 March 2017

The authors' responses to Anonymous Referee #1 are below. Each response is provided below the Referee's original comments.

Specific comments:

"Section 2.1: I had a hard time figuring out how many sites were used in the analysis based on the description in Section 2.1 and Figure 1. Table S1 was helpful for me to understand - consider moving it to the main paper. Indicate that there are five total sites. Section 2.1 says there are "several sites in Corpus Christi, including Hillcrest and Oak Park" but only those two are indicated in Table S1. I would also use a different color and/or more prominent symbol for the TCEQ sites in Figure 1 - as it is the symbols are hard to find, especially ones inside the circles indicating large cities."

Printer-friendly version

Discussion paper

Response: Table S1 has been moved to the main text as Table 1 and the selection of the two sites in Corpus Christi has been justified. Figure 1 has also been updated and is attached.

"Section 3.1: Mention Floresville. Even though the measurements didn't start until 2013, the signal is prominent in the figure."

Response: The elevated ethane mixing ratios from 2013-2015 have been added to the discussion.

"page 11, line 12: Specify that the emission rate displayed no trend *over the period 2013-2015*. (Otherwise seems inconsistent with results from Figure 2.)"

Response: This comment has been accepted and addressed in the text.

"There are many acronyms in this paper. Readability might be improved if the authors wrote some of them out instead. RNG (raw natural gas) and TG (tank gas), for example."

Response: Except for subscripts in equations, RNG and TG have been replaced with "raw natural gas" and "tank gas".

Technical comments:

"page 2, line 7: "associated with gas produced" instead of "associated gas produced""

Response: "Associated gas" is often used to refer to natural gas that is coproduced at oil wells. This change was not accepted. Instead, "associated gas" was italicized to indicate that it is a phrase.

"page 10, line 1: "significantly more constrained" instead of "significantly more constraint""

Response: This comment has been accepted and addressed in the text.

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-861, 2016.

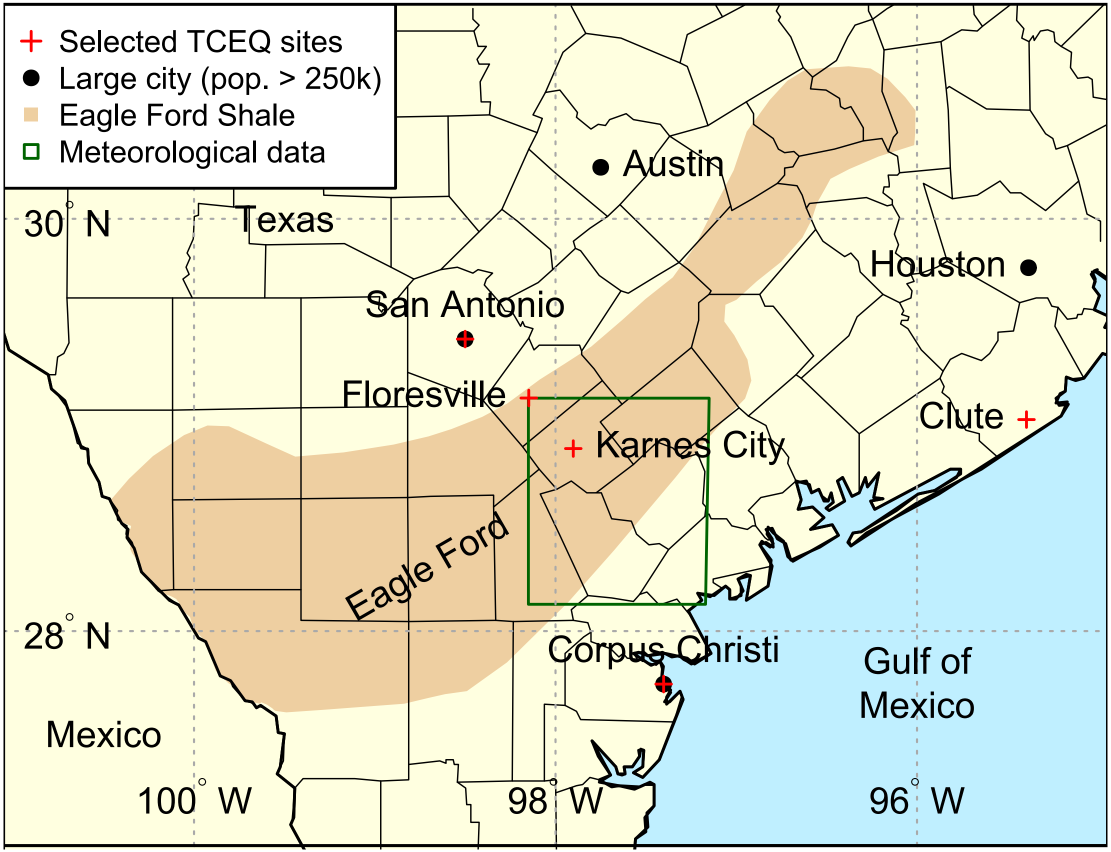

Fig. 1. Figure 1 in the manuscript has been updated to address the Referee's first comment.

Interactive comment

Printer-friendly version

Discussion paper

Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-861-AC2, 2017

Atmospheric Chemistry and Physics Discussions
Open Access

Interactive comment on "Quantifying alkane emissions in the Eagle Ford Shale using boundary layer enhancement" *by* G. Roest and G. Schade

G. Roest and G. Schade

gsroest@tamu.edu

Received and published: 23 March 2017

The authors' responses to Anonymous Referee #2 are below. Each response is provided below the Referee's original comments.

General comments

"This paper provides an emission estimate from a region with extensive oil and gas production whose emissions are not well known, and would therefore be an important addition to the body of knowledge concerning methane and alkane emissions from oil and gas production regions in the U.S. However, I have some concerns regarding the analysis that I think must be addressed before this paper is ready for publication. My main concerns are discussed in the next two paragraphs. Some lesser concerns are brought up in the Other Comments section."

"I am concerned with using Barnett tank alkane ratios to represent tank emissions from the Eagle Ford. First, the alkane ratios could be significantly different from the two regions. My understanding of the Eagle Ford shale is that produces a very wet mixture of hydrocarbons. Do the authors have data from any other oil producing regions in the U.S., like the Bakken or a traditional oil producing region? If they used those ratios, how would that affect the analysis results? I think more work will need to be done to show the effects of this assumption, especially since it plays such a large role in the results."

Response: We agree that using alkane ratios from liquid storage tank emission samples from the Eagle Ford Shale would be the most appropriate data to use in this study. However, we are not aware of such data publicly available for the Eagle Ford. While we have inquired about non-public data with two possible sources, our requests were unsuccessful. The available data, i.e. the sampled emissions from liquid storage tanks in the Barnett Shale are variable in composition and this is incorporated into our Monte-Carlo error analysis. The composition of emissions from oil and condensate storage tanks in other areas of Texas (Hendler et al., 2009) are also largely variable. We assume that the average composition of liquid storage tank emissions in the Eagle Ford falls within that variability of the Barnett Shale samples, although this assumption does introduce an uncertainty in our analysis. The text has been updated to emphasize this assumption and the associated uncertainty.

"Another concern I have is with the use of ground-based sampling to represent the entire vertical extent of the planetary boundary layer. I think some discussion of the location of possible sources of methane relative to the sampling sites is necessary. This is especially true for the Floresville site, which may be influenced by emissions that have not mixed completely through the planetary boundary layer on more days than just 18 March 2015."

Response: We agree that vertical measurements of alkane mixing ratios would validate the assumption that upwind emissions have been mixed through the planetary boundary layer (PBL). An aircraft campaign in the Barnett Shale has shown that methane emissions were thoroughly mixed downwind [Karion et al., 2015]. While we do not have alkane measurements in the vertical, the assumption of well-mixed upwind emissions is defensible. Our analysis shows that the long-term increasing trend in ethane enhancement between Corpus Christi and San Antonio parallels the development of the Eagle Ford Shale, suggesting that the Eagle Ford is responsible for the emissions that have led to said increase. Furthermore, the Eagle Ford Shale is sufficiently upwind of Floresville such that a discrete plume from a nearby source should be thoroughly dispersed in the afternoon PBL before it reaches Floresville. This is based on the finding that the nearest potential source is approximately 12 km upwind (a well pad south of the site), while all other sources are at least 15 or more kilometers upwind. Under typically conditions used in this study with boundary layers of 1.5-2 km depth and convective velocity scale values of 1-2 m s-1, vertical dispersion occurs within approximately 30 minutes, while horizontal transport at typical wind speeds of 5 m s-1 will require more than 30 min to reach the Floresville site.

Other comments

"p. 1, Line 25, is carbon monoxide a HAP? I don't see it here: https://www.epa.gov/haps/initial-list-hazardous-air-pollutants-modifications"

Response: This has been addressed in the text.

"p. 2, Line 10, please add Olaguer et al. to the References"

Response: Olaguer (2012) is already included in the references.

"p. 2, Line 14, How does the 5750 Gg of methane compare to the EPA GHG inventory? If they are the same, I'd cite the GHG inventory. If they are different, I'd still use the GHG inventory, but note the difference. Also, does this number include emissions from petroleum production as well as natural gas? Since you included associated gas in the leak rate calculation for the Eagle Ford on p. 11, line 27, the national leak rate should also include emissions from petroleum production."

Response: The methane emissions estimate from the EPA NEI Oil and Gas Emissions tool is lower than that of the GHG inventory because the GHG inventory includes emissions from federal offshore waters (the Outer Continental Shelf) while the EPA NEI Oil and Gas Emissions tool does not. Since the Energy Information Administration includes natural gas from federal waters in their U.S. natural gas production estimate, the emissions estimate of 5,750 Gg from the EPA NEI Oil and Gas Emissions tool has been replaced by the GHG inventory estimate of 6,616 Gg of methane emissions in 2011. This number includes emissions from both oil and gas wells.

"p. 3, Line 16, This emission rate is somewhat misleading. Schneising et al. reported an energy content leak rate, which is not the same as a natural gas leak rate. The energy content leak rate takes into account the oil produced as well as the natural gas. See Howarth [Energy and Emission Control Technologies, 2015, p. 48] or Peischl et al. [JGR-Atmospheres, 2016, p. 2] for a discussion on this issue."

Response: We agree that the context of an emission rate must be explicit. We chose to express the emission rate of natural gas as a fraction of produced natural gas to be consistent with most studies in other shale areas, including dry shale basins such as the Marcellus, as well as bottom-up greenhouse gas inventories (e.g. EPA GHG). We have revised the manuscript to clarify the context of the emission rate. If we were to compare the methane emissions to produced energy following Scheising et al., our emission rate would be lower as the produced energy in the Eagle Ford is largely in the form of oil. However, note that the combined emissions of ethane, propane, and butanes exceed the mass of methane emissions in our study. This suggests that a comparison of methane emissions alone to produced energy content may not be appropriate because it excludes the energy emitted in the form of non-methane VOCs. We consider this to be beyond the scope of this manuscript.

"p. 4, Line 21, Why did you use the EDAS 40 km dataset for meteorology over others?

Perhaps include a sentence explaining your choice."

Response: This dataset was chosen for computational efficiency while running the HYSPLIT trajectory model over a set of several years. Despite its relatively coarse grid, it will capture the general synoptic-scale flow, so it is sufficient to use to identify days with southeasterly flow. A sentence was added to the text. Note that a more robust meteorological dataset (the NARR) was used to calculate emissions.

"p. 5, Line 1, Please show a time series of the background upwind mixing ratios and the enhanced mixing ratios at Floresville. This will give the reader a sense of how well the background sites represent the background air impacting the Floresville site."

Response: A timeline of the ethane mixing ratios at the upwind site in Corpus Christi and the downwind site in Floresville has been added to the supplemental information document. This figure also shows the seasonality of ethane which was questioned in another comment.

"p. 7, Line 7, Please provide some discussion of the PBL height and what effect the uncertainty of the modeled PBL height has on the analysis. Has the modeled PBL height been verified using LIDAR or aircraft measurements?"

Response: The planetary boundary layer height does serve as a source of uncertainty. The uncertainty was estimated using the spatial variability of the PBL height in a subset of grid cells in between the upwind and downwind sites. Supporting Table S3 shows the height of the PBL for each afternoon and the standard deviation of the aforementioned cells. This standard deviation was used to introduce uncertainty into the Monte Carlo simulation. On average, the PBL height was 1789 m with a standard deviation of 164 m. This is a relative standard deviation of 9.2%. That is small compared to the relative uncertainty of the ethane enhancement and the alkane composition in the raw natural gas and tank gas samples. A few sentences have been added to the manuscript to emphasize the uncertainty in the meteorological variables. A reference has also been added for a study in which the NARR PBL heights were shown to have no strong bias compared to objectively determined PBL heights from sounding data, though the correlations were moderate and seasonally dependent.

"p. 9, Line 1, Did you not see a seasonal change in background ethane due to greater chemical loss during the summer?"

Response: The timeline of the ethane mixing ratios at the upwind and downwind sites shows the seasonality of background ethane mixing ratios. A brief discussion has been added to the text.

"p. 11, Line 33, A comparison with the EPA inventory estimate from petroleum production would be a fairer one, considering how much oil is produced in the Eagle Ford shale."

Response: Please see the response for the comment pertaining to p. 3, Line 16.

"Conclusions section, Please include a time frame for your emissions estimates. Are they for the entire study period? If so, please state this explicitly in the Conclusions."

Response: The emissions were estimated for a set of 68 days from August 2013 through August 2015. This was added to the text.

Grammar suggestions:

"p. 3, Lines 7-10, This is a long sentence. Consider splitting it up into two."

Response: Done

"p. 10, Line 1, I'm not sure “constraint” is a verb, unless it is an old-timey past tense."

Response: This was fixed in the text.

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-861, 2016.

The Eagle Ford Shale in southern Texas is home to a booming unconventional oil and gas industry, the climate and air quality impacts of which remain poorly quantified due to uncertain emissions estimates. We used the atmospheric enhancement of alkanes from Texas Commission on Environmental Quality volatile organic compound monitors across the shale, in combination with back trajectory and dispersion modeling, to quantify C2-C4 alkane emissions for a region in southern Texas, including the core of the Eagle Ford, for a set of 68 days from July 2013 to December 2015. Emissions were partitioned into raw natural gas and liquid storage tank sources using gas and headspace composition data, respectively, and observed enhancement ratios. We also estimate methane emissions based on typical ethane-to-methane ratios in gaseous emissions. The median emission rate from raw natural gas sources in the shale, calculated as a percentage of the total produced natural gas in the upwind region, was 0.8% with an interquartile range (IQR) of 0.5% -1.4%, close to the U.S. Environmental Protection Agency's (EPA) current estimates. However, storage tanks contributed 24% of methane emissions, 54% of ethane, 82% percent of propane, 90% of n-Butane, and 83% of isobutane emissions. The inclusion of liquid storage tank emissions results in an emission rate of 2.2% (IQR of 0.9-4.9%) relative to produced natural gas, ~~exceeding~~ overlapping the EPA~~'s~~ estimate ~~by a factor of two~~of roughly 1.6%. We conclude that ~~leaks~~ emissions from liquid storage tanks are likely a major source for the observed non-methane hydrocarbon enhancements in the northern hemisphere.

Quantifying alkane emissions in the Eagle Ford Shale using boundary layer enhancement

Geoffrey Roest[1] and Gunnar Schade[1]

[1]Department of Atmospheric Sciences, Texas A&M University, 3150 TAMU, College Station, Texas 77843-3150

Correspondence to: Geoffrey Roest (gsroest@tamu.edu)

Abstract. The Eagle Ford Shale in southern Texas is home to a booming unconventional oil and gas industry, the climate and air quality impacts of which remain poorly quantified due to uncertain emissions estimates. We used the atmospheric enhancement of alkanes from Texas Commission on Environmental Quality volatile organic compound monitors across the shale, in combination with back trajectory and dispersion modeling, to quantify C_2-C_4 alkane emissions for a region in southern Texas, including the core of the Eagle Ford, for a set of 68 days from July 2013 to December 2015. Emissions were partitioned into raw natural gas and liquid storage tank sources using gas and headspace composition data, respectively, and observed enhancement ratios. We also estimate methane emissions based on typical ethane-to-methane ratios in gaseous emissions. The median emission rate from raw natural gas sources in the shale, calculated as a percentage of the total produced natural gas in the upwind region, was 0.8% with an interquartile range (IQR) of 0.5% -1.4%, close to the U.S. Environmental Protection Agency's (EPA) current estimates. However, storage tanks contributed 24% of methane emissions, 54% of ethane, 82% percent of propane, 90% of *n*-Butane, and 83% of isobutane emissions. The inclusion of liquid storage tank emissions results in an emission rate of 2.2% (IQR of 0.9-4.9%) relative to produced natural gas, ~~exceeding the EPAestimate by a factor of two~~overlapping the EPA's estimate of roughly 1.6%. We conclude that ~~leaks~~ emissions from liquid storage tanks are likely a major source for the observed non-methane hydrocarbon enhancements in the northern hemisphere.

1 Introduction

The recent boom in onshore oil and gas production in the U.S. has heightened concerns over the environmental impacts of petroleum as an energy source. Technological advances in petroleum recovery methods, namely hydraulic fracturing and horizontal drilling (hereafter referred to as unconventional oil and gas or UOG development), have allowed for the production of previously inaccessible hydrocarbons in shale plays and tight sand formations. Emissions from the rapidly changing infrastructure in U.S. oil and gas fields have introduced new uncertainties in the climatological and ~~ecological~~ air quality impacts associated with petroleum production. However, independent scientific studies of these impacts are sparse in some shale plays that have developed rapidly, including the Eagle Ford Shale (EFS) in southern Texas.

Emissions from oil and gas production include methane, a potent greenhouse gas (GHG); non-methane volatile organic compounds (NMVOCs) and oxides of nitrogen (NO_x), important precursors for ozone formation; carbon monoxide; and hazardous air pollutants (HAPs) such as ~~carbon monoxide and~~ benzene, a known carcinogen (*U.S. Centers for Disease Control*, 2014). There are a variety of emission sources for each of these trace gases during oil and gas exploration, production, and distribution, resulting in co-emissions of GHGs, NMVOCs, NO_x, and HAPs. Direct emissions from oil and gas operations include hydrocarbons from flowback events and well completions, on-site equipment during the routine operation of the well, such as compressors and storage tanks, and unintentional emissions such as leaks from valves and pipelines. Indirect emissions of NMVOCs, HAPs, and NO_x come largely from combustion sources such as diesel engines in generators, drilling rigs, compressors, and trucks used to transport equipment, water, and petroleum (*Field et al.*, 2014). In addition, flaring is widely used in certain shale plays, including the EFS, to dispose of excess natural gas – mostly ~~associated gas~~ *associated gas* produced at oil wells (*Tedesco and Hiller*, 2014). Flaring is intended to efficiently dispose of hydrocarbons, resulting in emissions of carbon dioxide as opposed to methane and higher hydrocarbons, including HAPs. However, ~~inefficiencies in flaring have~~ the high temperatures in flares result in NO_x emissions, and inefficiencient flaring has been shown to release ~~NO_x, methane, unburned hydrocarbons ,~~ unburned hydrocarbons and NMVOCs produced via pyrolysis (*Strosher*, 2000; *Olaguer*, 2012; *Pikelnaya et al.*, 2013).

The ~~climatological~~ regional and global impacts of methane emissions from UOG development in the U.S. are poorly quantified due to widely varying and largely uncertain emissions estimates. According to the U.S. Environmental Protection Agency's (EPA) ~~Oil & Gas Emission Estimation Tool, available from the 2011 National Emissions Inventory documentation (?), approximately 5,750~~ 2016 greenhouse gas inventory (*U.S. Environmental Protection Agency*, 2016), approximately 6,570 Gg of methane was emitted from all ~~upstream~~ oil and gas ~~operations~~ field production in the U.S. ~~in~~ and in offshore federal waters in 2011 ~~, corresponding to 300~~ when voluntary emissions reductions are included. This corresponds to 428 billion cubic feet (bcf) of natural gas at 1 atm and 15°C, assuming an average methane content of 80% by volume in raw U.S. natural gas (*Pétron et al.*, 2012). During the same year, nationwide gross natural gas production ~~totaled 28,479~~ from oil and gas wells totaled 26,700 bcf (*U.S. Energy Information Administration*, 2017). Therefore, the ~~natural gas emitted from upstream~~ EPA's emission rate of natural gas from oil and gas ~~operations~~ field production in the U.S. in 2011 ~~corresponded to an~~ was 1.6% relative to the volume of produced natural gas. The EPA's 2016 greenhouse gas inventory estimated methane emissions of 6,985 Gg in 2013, which, when compared to the 2013 U.S. natual gas production of 30,139 bcf, yields a relative emission rate of ~~1.1%by volume~~ 1.5%, a slight reduction from 2011. However, the EPA's methane emission estimates from U.S. oil and gas field production in the 2016 greenhouse gas inventory increased from the 2015 EPA greenhouse gas inventory (*U.S. Environmental Protection Agency*, 2015a). In the 2015 greenhouse gas inventory, the estimated methane emissions from U.S. oil and gas field production for 2013 totaled 2,722 Gg or 0.6% of produced natural gas ~~.~~ by volume. The increase between the 2015 and 2016 greenhouse gas inventories is due, in part, to updated emission factors for methane emissions.

~~In contrast, several~~ Several recent top-down estimates of emissions from individual shale plays ~~exceed~~ exceeded bottom-up ~~estimates~~inventories, at times by an order of magnitude or more (*Karion et al.*, 2013; *Pétron et al.*, 2012, 2014; *Miller et al.*, 2013). Several aircraft measurements performed upwind and downwind of major natural gas-producing shale areas (*Peischl et al.*, 2015) suggest that higher-than-inventory emissions in one shale area might be compensated by lower-than-inventory emissions in another shale area, suggesting that national emissions might be represented correctly in the inventory. More recently, a series of studies in the Barnett Shale, which mostly produces natural gas, have attempted to reconcile the differences between bottom-up and top-down estimates (e.g. *Karion et al.*, 2015; *Lyon et al.*, 2015; *Smith et al.*, 2015). While the emissions estimates from the Barnett Shale studies ~~are~~ were lower than other top down estimates, the studies concluded (*Zavala-Araiza et al.*, 2015) that current emissions inventories, such as the EPA's Greenhouse Gas Reporting Program (GHGRP) (*U.S. Environmental Protection Agency*) and the Emission Database for Global Atmospheric Research (EDGAR) (*European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL)*, 2009), underestimate emissions from oil and gas production in the Barnett Shale, similar to a prior conclusion reached by *Zavala-Araiza et al.* (2014) based on hydrocarbon measurements collected throughout the shale area by the Texas Commission on Environmental Quality (TCEQ). Furthermore, a recent study using satellite measurements of tropospheric methane (*Turner et al.*, 2016) concluded that North American methane emissions have increased by approximately 30% between 2002 and 2014, most likely as a result of UOG development. In combination, these studies strongly ~~suggests~~ suggest that bottom-up inventories underestimate actual emissions from oil and gas systems.

The uncertainty and variability in methane emission estimates suggest that NMVOC emissions are also poorly quantified, as methane and NMVOCs are often co-emitted. The EFS may have particularly poorly quantified emission estimates of both methane and NMVOCs due to the widespread use of flaring (*Tedesco and Hiller*, 2014). The region has experienced a boom in UOG production in the region since 2008 (*Gebrekidan*, 2011)], and the EFS is currently one of the most productive shale plays in the United States (*U.S. Energy Information Administration*, 2016). In nearby San Antonio, Bexar County, ozone mixing ratios increased in tandem with oil and gas production rates during the initial growth years of the EFS (*Schade and Roest*, 2015). The increasing ozone levels have raised concerns over public health in the city, which is in danger of being designated as a nonattainment area by the EPA~~as~~. Bexar County ozone design values for recent years exceeded both the current and previous ozone standards of 70 and 75 ppb (*Alamo Area Council of Governments (AACOG)*, 2014; *U.S. Environmental Protection Agency*, 2015b), while ozone levels in Houston, a non-attainment area, have dropped below levels in San Antonio. Modelling studies for the San Antonio region (*Alamo Area Council of Governments (AACOG)*, 2013a; *Pacsi et al.*, 2015) have found that emissions from the EFS contribute significantly to regional ozone formation. However, these studies have utilized emissions inventories that may underestimate the magnitude of emissions from oil and gas operations in the shale area. *Schneising et al.* (2014) estimated methane emissions in the western EFS using trends in satellite based measurements of total column methane from 2006 to 2008 compared with 2009 to 2011. Their study yielded an emission rate ~~as a fraction of changes in production rates of~~ of 9.1% (2.9% – 15.3%), ~~nearly ten times higher than bottom-up inventories for nationwide relative emissions would suggest.~~ expressed as a fraction of increased methane emissions vs. increases in produced energy in the region. While their study focused on methane emissions, the co-emission of methane and NMVOCs suggests that existing VOC inventories may also be biased low, and the emissions of VOCs would result in the additional loss of produced energy. In another satellite data study, *Duncan et al.* (2016) showed that tropospheric NO_2 concentrations have increased in shale areas that practice flaring, including the EFS, while concentrations have decreased in US urban areas. Since most of the EFS is part of rural south Texas where NO_x concentrations are comparatively low, increases in NO_x can contribute strongly to ozone formation. Therefore, the increase in ozone associated with ozone precursor emissions from the EFS may be ~~much~~ higher than what the existing ozone modeling studies suggest.

This study uses the atmospheric enhancement of short-chain alkanes – ethane (C_2), propane (C_3), isobutane (iC_4), n-butane (nC_4), isopentane (iC_5), and n-pentane (nC_5) – between upwind and downwind measurement locations to estimate alkane emissions from a region in southeast Texas including the core of the EFS. Alkanes dominate atmospheric OH radical reactivity at a TCEQ monitoring site north of the EFS (*Schade and Roest*, 2016) and the emission estimates for these short-chain alkanes are needed to assess the potential air quality impacts from the EFS. We focus on ethane as a tracer for oil and gas emissions as it is the second most abundant compound in natural gas (*Xiao et al.*, 2008) and, unlike methane, it is not emitted by microbial sources in significant quantities (*Simpson et al.*, 2012). Recent increases in ethane abundance in the northern hemisphere have been linked to UOG production in the U.S. (*Franco et al.*, 2016; *Helmig et al.*, 2016; *Kort et al.*, 2016). Ethane has thus been used in previous oil and gas emissions estimates (*Schwietzke et al.*, 2014; *Smith et al.*, 2015), and statistically significant increases in ethane mixing ratios have been observed downwind of the EFS during its development (*Schade and Roest*, 2015). The C_3 and C_4 alkane-to-ethane enhancement ratios are used to estimate the relative contributions of raw natural gas emissions and vented gases from liquid storage tanks, two major sources of gaseous emissions from upstream UOG that have varying compositions (*Brantley et al.*, 2014; *Field et al.*, 2014; *Lyon et al.*, 2015; *Kort et al.*, 2016). A mass balance approach and a Monte Carlo simulation are then used to estimate the emissions of C_2-C_4 alkane from raw natural gas emissions and liquid storage tank venting and the associated uncertainties. Methane emissions are also estimated using methane-to-ethane ratios in raw natural gas and vented storage tank gas. Lastly, the methane emission rate is expressed as a fraction of the produced natural gas to compare our emission estimate with other top-down studies.

2 Methods

2.1 TCEQ Data

The TCEQ operates a network of air quality monitoring sites across the state of Texas, some of which measure NMVOCs including alkanes, alkenes, cycloalkanes, and aromatics. The TCEQ collects NMVOC data to the east and southeast of the EFS in Clute and at several sites in Corpus Christi, including "Hillcrest" and "Oak Park" (Fig. 1)~~.~~, which were selected for use in this study due to data availability and their location. Other sites in Corpus Christi are immediately downwind of major local point sources when winds are blowing from the Gulf of Mexico. To the northwest of the EFS, NMVOC data have been collected since summer 2013 in Floresville, a small city immediately north of the shale area, and in northwest San Antonio ("Old Highway 90"). ~~A description of the sites is compiled in Table S1 in the supporting information~~Descriptions of the five sites that are used in this study are presented in Table 1. Data from these sites have been previously used to demonstrate trends in ethane mixing ratios near the EFS (*Schade and Roest*, 2015). The emissions estimates in this study were performed using hourly ethane data from the automated ozone precursor monitoring sites in Floresville and Corpus Christi – Oak Park (hereafter referred to as "Oak Park"). While the Oak Park site was installed before the oil and gas boom in the EFS, data for Floresville are only available since 19 July 2013. Therefore, direct data comparisons were performed only for a thirty-month period from July 2013 through December 2015.

Alkane mixing ratios at Floresville and Oak Park were compared under south to southeasterly flow regimes, when Corpus Christi is upwind of the EFS and Floresville is downwind. South to southeasterly flow regimes were identified using 48 hour back-trajectories originating at Floresville from the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model (*Stein et al.*, 2015). The trajectories were run four times per day at an interval of six hours beginning at 06:00 UTC (00:00 LST). The EDAS 40 km dataset (*National Centers for Environmental Prediction (NCEP)*) was used for meteorology in the HYSPLIT model. This dataset, which adequately captures synoptic scale flow, was chosen for computational efficiency. A series of polygons, as discussed in supporting Text S1 and Table ~~S2~~ S1 and shown in supporting Fig. S1, were used to identify air mass origins. Trajectories that had continental origins during the previous 48 hours were ~~discarded~~ flagged in order to isolate those with marine origins. Only days with at least three out of four southeasterly trajectories and at least 75% completeness (i.e., at least eighteen hours of NMVOC data) at both Floresville and Oak Park were used. ~~Both of these sites~~

The TCEQ sites in Floresville and Oak Park measure hydrocarbons using nearly identical automated GC-FID systems which record continuous hourly data from 40 min, 600 mL air samples. Standard operating procedures for these instruments are available from the Field Operations Division of the TCEQ (*Texas Commission on Environmental Quality*, 2005). The method detection limit is 0.4 ppbC, and instrument precision is measured using weekly injections of propane and benzene standard gases. Data are quality assured by the TCEQ if the relative difference between standard gas measurements remains less than 20%. We have assumed that this value is representative of the two-standard-deviation uncertainty of an individual measurement. The mean afternoon alkane mixing ratios for each day were calculated at both sites by averaging hourly mixing ratios during the afternoon hours (15:00 to 18:00 LST), when daytime convection allows for mixing throughout the planetary boundary layer (*Stull*, 2009). The alkane enhancement for each day was determined by subtracting the mean of three afternoon alkane mixing ratios at the upwind site of Oak Park from the mean afternoon mixing ratio at the downwind site in Floresville. The relative standard error of the three hourly measurements at each site is 5.8%, and the uncertainty in the enhancement is equal to the sum of the absolute errors of the afternoon mixing ratios at each site.

2.2 Alkane Sources

In this study, we assumed that regional alkane emissions are dominated by UOG operations in the EFS. Other sources of ethane emissions were assumed to be negligible as no regional biomass burning was reported during the study period (*Randerson et al.*, 2015). However, the mixing ratios of longer chain alkanes (notably C_5 and higher) may be influenced by evaporative and tailpipe emissions from nearby automotive traffic (*Tsai et al.*, 2006; *Ho et al.*, 2009; *Simpson et al.*, 2012). Emissions from UOG come from multiple sources and can include emissions of raw natural gas from compressors, flowback events, and unintentional leaks. The composition of these gases are dominated by the most volatile hydrocarbons, i.e. methane and ethane. In comparison, emissions from storage tanks, used to store liquids from wells prior to transportation and further processing, have been shown to contribute largely to hydrocarbon emissions in UOG shale plays (*Lyon et al.*, 2015, 2016). Since gas produced at the well is separated from liquids prior to storage, the headspace in storage tanks is primarily composed of hydrocarbons heavier than ethane, notably short-chain alkanes such as propane, butanes, and pentanes, although methane and ethane ~~are still~~ may still be present. We assume that regional short-chain alkane emissions are dominated by gases produced at the wellhead (referred to as *raw natural gas*~~or RNG~~) and emissions from liquid storage tanks (referred to as *tank gas*~~or TG~~). Table 2 shows the available compositions of ~~RNG~~ raw natural gas samples from the EFS and ~~TG~~ tank gas samples from the Barnett Shale. ~~Due to a lack of TG samples from the EFS, we assumed that the composition of TG from~~ To our knowledge, no tank gas composition data are publicly available for the EFS. The sampled emissions from liquid storage tanks in the Barnett Shale ~~(*ENVIRON International Corporation*, 2010) is sufficiently similar to estimated TG emissions from the EFS. Since TG compositions are highly variable depending on liquids compositions, and may not be comparable between shale areas either, the TG values used here represent~~ are variable in composition and this is incorporated into our error analysis. The composition of emissions from oil and condensate storage tanks in conventional production areas in Texas (*Hendler et al.*, 2009) are also largely variable. We assume that the average composition of liquid storage tank emissions in the Eagle Ford falls within that variability of the Barnett Shale samples, although this assumption introduces a major source of uncertainty ~~to~~ in our analysis.

Observed alkane enhancement ratios can be partitioned into emissions from multiple sources, including ~~RNG and TG~~ raw natural gas and tank gas emissions. Equation 1 shows the partitioning of (e.g.) observed propane-to-ethane ratios into ~~RNG, TG~~raw natural gas, tank gas, and all other sources.

$$\left(\frac{C_3}{C_2}\right)_{observed} = f_{RNG}\left(\frac{C_3}{C_2}\right)_{RNG} + f_{TG}\left(\frac{C_3}{C_2}\right)_{TG} + f_{other}\left(\frac{C_3}{C_2}\right)_{other} \quad (1)$$

where $\left(\frac{C_3}{C_2}\right)_{RNG}$, $\left(\frac{C_3}{C_2}\right)_{TG}$, and $\left(\frac{C_3}{C_2}\right)_{other}$ represent the C_3/C_2 ratios in emissions from ~~RNG, TG~~raw natural gas, tank gas, and other sources, respectivey, and the relative contributions to the ~~observe~~ observed ratio from each source (f_{RNG}, f_{TG}, and f_{other}) add up to 1. If ~~RNG and TG~~ raw natural gas and tank gas sources dominate regional alkane emissions and other sources can be considered negligible, then $f_{RNG} + f_{TG} \approx 1$ and

$$\left(\frac{C_3}{C_2}\right)_{observed} = f_{RNG}\left(\frac{C_3}{C_2}\right)_{RNG} + (1 - f_{RNG})\left(\frac{C_3}{C_2}\right)_{TG} \quad (2)$$

$$f_{RNG} = \frac{\left(\frac{C_3}{C_2}\right)_{observed} - \left(\frac{C_3}{C_2}\right)_{TG}}{\left(\frac{C_3}{C_2}\right)_{RNG} - \left(\frac{C_3}{C_2}\right)_{TG}} \quad (3)$$

Here, f_{RNG} is found using C_3/C_2 and verified using iC_4/C_2 or nC_4/C_2. ratios. This number represents the fraction of ethane attributed to emissions from ~~RNG~~ raw natural gas sources, such that $C_{2,RNG} = f_{RNG} \cdot C_{2,observed}$ and $C_{2,TG} = (1 - f_{RNG}) \cdot C_{2,observed}$. The expected methane enhancement can be estimated using eq. 4.

$$C_1 = C_{2,observed}\left(f_{RNG}\left(\frac{C_1}{C_2}\right)_{RNG} + (1 - f_{RNG})\left(\frac{C_1}{C_2}\right)_{TG}\right) \quad (4)$$

Similarly, the methane enhancement estimate, along with other alkanes, can be attributed to ~~RNG and TG~~ raw natural gas and tank gas sources as follows.

$$C_{1,RNG} = C_{2,RNG} \cdot \left(\frac{C_1}{C_2}\right)_{RNG} \quad (5)$$

$$C_{1,TG} = C_{2,TG} \cdot \left(\frac{C_1}{C_2}\right)_{TG} \quad (6)$$

2.3 Mass Balance Approach

Short-chain alkane emissions from a region encompassing the central section of the EFS were quantified using a mass-balance approach that has been adapted to an area source (eq. 7), in which emissions are considered to be spatially and temporally homogenous.

$$F = \left(\bar{U} \cdot \cos\alpha\right) \cdot \bar{n} \cdot \int\limits_{Z_0}^{Z_{PBL}} \rho\,(z)\,dz \cdot \Delta x \tag{7}$$

A different form of this mass-balance method has been used in previous emissions estimates for emission plumes from oil and gas systems (*Karion et al.*, 2013, 2015; *Caulton et al.*, 2014; *Pétron et al.*, 2014; *Smith et al.*, 2015). This estimate can be biased low as it does not account for the entrainment of air from the free troposphere into the planetary boundary layer (PBL) (*Karion et al.*, 2015), but can also be biased high if nearby emissions produced unmixed plumes. We consider the Floresville site to be sufficiently downwind of ethane sources such that it is not impacted by discrete plumes if the PBL is well mixed. In our approach, we assume that the component of the wind that is parallel to the transect between upwind and downwind measurement sites ($\bar{U} \cdot \cos\alpha$, where α represents the angular deviation in wind from the direction of the transect), is representative of the general trajectories of air masses in the PBL being advected from the Gulf of Mexico. While actual trajectories that do not follow straight paths may stay over emissions sources for long periods of time, large spatial deviations in wind direction will result in a reduction of the magnitude of $\bar{U} \cdot \cos\alpha$. Therefore, the time an air mass spends over an emissions source will be reflected in the magnitude of the resultant wind. In a well-mixed PBL, the alkane mixing ratios are assumed to be near constant with height, and the mixing ratio enhancement ($\bar{n}$) multiplied by the integrated molar density ($\int_{Z_0}^{Z_{PBL}} \rho\,(z)\,dz$) from the surface ($Z_0 = 122$ m at Floresville) to the top of the PBL (Z_{PBL}) provides an estimate of the molar flux between the upwind and downwind locations. It is assumed that $\rho(z) = \rho_0 \cdot \exp\left(-\frac{z}{H}\right)$, where scale height $H = \frac{R_{air}T}{g}$, $R_{air} = 287$ J kg^{-1} K^{-1}, $g = 9.81$ m s^{-2}, and the molar density of air at sea level $\rho_0 = 42.29$ mol m^{-3} (*United States Committee on Extension to the Standard Atmosphere*, 1976). Lastly, a horizontal dimension (Δx) is necessary to produce an alkane flux for the region that is affecting the downwind receptor location. This was estimated as outlined in the Sect. 2.4.

Meteorological data used in the mass balance approach were obtained from NOAA's North American Regional Reanalysis (NARR) (*Mesinger et al.*, 2005), a combined model and assimilated dataset with a horizontal resolution of approximately 32 km. Temperature and ~~boundary layer~~ PBL height data for each date were obtained for the 3 hour period from 15:00 to 18:00 LST, representing general afternoon hours when the PBL is well mixed. Wind data were obtained for the previous 3 hour period of 12:00 to 15:00 LST when parcels were being advected over the EFS. Temperature and wind components at 950 mb were assumed to be representative of boundary layer conditions. Days with complicated meteorological conditions (e.g., precipitation, fronts, dry lines, or strong wind shear in ~~boundary layer~~the PBL) were discarded. The boundary layer heights from the NARR have been shown to have no strong bias compared to objectively determined PBL heights from sounding data at a site in Oklahoma (*Schmid and Niyogi*, 2012), although the correlation is moderate (as high as $R^2 = 0.58$ in the winter and as low as $R^2 = 0.39$ in the spring). While the use of the NARR introduces some uncertainty in the meteorological variables, we consider this to be the best available information for this data-sparse region where only surface observations are available.

2.4 Horizontal dimension and production reference areas

The horizontal dimension in previous mass balance applications using aircraft data has typically come from an observation of background mixing ratios at the "edge" of the emissions plume (e.g. *Karion et al.*, 2015), where upwind and downwind mixing ratios become indistinguishable. Since the EFS can be considered a line source, but only one downwind measurement site is available, we defined the "edge" of the emissions plume using HYSPLIT's backward dispersion modeling tool in STILT mode (*Hu et al.*, 2015). Model resolution was set to a 0.05° latitude × 0.05° longitude output grid (approximately 5 km resolution at these latitudes) using 12 km North American Mesoscale Model (NAM) meteorology input data. The model was set up to release 5,000 particles equally distributed in the PBL above the Floresville monitor site at 16:00 LST on each selected day using the estimated boundary layer depth from the NARR data. Particles were followed backwards for 20 hours and an integrated emissions impact map was created from particles entering the lowest layer (50 m agl). In almost all cases, the map was no longer changing after 8-14 hours of backward integration because all boundary layer particles had moved off-shore. Particle plots were used to further exclude days with significant wind shear in the boundary layer, as they do not fulfill the requirements for the mass balance technique.

The emissions impact map was assessed in two ways: (1) The near-field plume width was measured at the southern edge of the EFS as the representative horizontal measure necessary for the mass balance equation (Eq. 7) by assessing grid cell distances between the eastern and western edges of the plume. This choice was based on the assumption that alkane emissions are dominated by emissions in the EFS, that this width corresponds to the spread of back trajectory ensembles from the receptor location in Floresville, and that this dimension corresponds to the width of a plume under the uniform advection conditions necessary for mass balance had a continuous downwind measurement taken place for a source centered on the EFS. (2) The gridded map was overlaid with a map of natural gas and associated gas production for the thirty month period from July 2013 to December 2015, developed from county production data (*Railroad Commission of Texas*) equally distributed into the grid based on our assumption of a homogenously distributed source. All gas production in non-zero grid cells was accumulated to provide a reference number of upwind production potentially contributing to the measured downwind mixing ratios at the Floresville receptor. These numbers thus varied on a daily basis with wind direction and turbulence affecting the integrated impact map. Note that this estimate is based on the single receptor location, assuming it to be equivalent to an actual boundary layer "curtain" measurement undertaken via flying aircraft. Simulating the aircraft's "curtain" measurement via a particle release from numerous upwind locations would not substantially alter the result because of counter-acting consequences: A multi-point release throughout the downwind boundary layer would increase the width of the plume (impact map cross-section at southern EFS edge), increasing the total emissions estimate according to Eq. 7, but at the same time would also increase the production reference area where potential emissions occur. Thus our results would only change significantly if either upwind emissions or production were strongly non-homogenously distributed.

2.5 Monte Carlo Simulation

The errors arising from the variability and uncertainties of the alkane enhancement and the parameters derived from regional meteorology inputs were assessed using a Monte Carlo simulation, in which the emissions for each day were calculated one million times by randomly sampling the input parameters from either empirical or assumed probability distributions. The simulation was performed using the 'mc2d' package in R (*Pouillot and Delignette-Muller*, 2010). The absolute uncertainty in the afternoon alkane mixing ratios at each site associated with the precision of the instrument was represented by normal distributions about the afternoon alkane mixing ratios with relative standard deviations of 5.8%, as discussed in Sect. 2.1. The u and v components of the wind, the planetary boundary layer depth, and the temperature were also assigned normal distributions based on the spatial variability in the NARR composite data over a 1° latitude by 1° longitude box situated in the central Texas coastal plain, with Floresville located at the northwest corner (Fig. 1) . It is assumed that the meteorology in this region represents the general conditions to which air masses were subjected as they traveled inland from the Texas Coast towards Floresville. The compositions of four ~~RNG and four TG~~ raw natural gas and four tank gas samples shown in Table ~~1~~ 2 were used to produce normal distributions of the alkane ratios in ~~RNG and TG~~raw natural gas and tank gas. A Monte Carlo simulation was performed for each day, allowing for the temporal variability and the dependence of the emissions on input variables to be assessed.

3 Results and Discussion

3.1 Ethane trends

Schade and Roest (2015) briefly discussed long-term trends in ethane mixing ratios at TCEQ sites around the EFS, and an update is provided in Fig. 2. Here, we present the results from Kruskal-Wallis rank-sum and Dunns tests (*Kruskal and Wallis*, 1952; *Dunn*, 1964) performed on ethane mixing ratios vs. year. At the Corpus Christi – Hillcrest site, no set of years exhibited statistically significant ($p < 0.05$) differences in ethane mixing ratios under southeasterly wind regimes. At Clute, ethane was statistically significantly higher ($p < 0.05$) in 2015 than it was in 2007-2011, though no other years showed significant differences. We attribute recent increases in ethane mixing ratios in Clute to changes in emissions from local point sources, as the neighboring city of Freeport is a hub of petroleum processing and transportation (*Bonney*, 2014; *Ryan*, 2014). The data suggest that background ethane levels over the Gulf of Mexico did not significantly change during the development of the EFS. However, ethane mixing ratios at the San Antonio site are statistically significantly higher ($p < 0.05$) in later years than in earlier years. Ethane was higher in 2011 than 2007-2009, higher in 2012 than 2007-2011, and higher in 2013-2015 than in 2007-2010. The Floresville site was not installed until 2013 so the long term trend in ethane mixing ratios at that location is unknown. However, Floresville observed the highest ethane mixing ratios in the region from 2013-2015. While there is no evidence that ethane mixing ratios along the coast increased over time during southeasterly flow, ethane did increase downwind of the EFS during its development.

3.2 Alkane Enhancement

During the thirty-month period from July 2013 through December 2015, a total of 69 days were found to have 3 out of 4 trajectories identified as southeasterly, appropriate meteorological conditions, and 75% completeness at both Floresville and Oak Park. One of these days (18 March 2015) had alkane enhancement values that were outliers. Since we cannot exclude that the downwind measurement site of Floresville was influenced by a plume on this date, it was not ~~further~~ considered for analysis. The majority of the remaining 68 dates, which occured between August 2013 through August 2015, fell into the summer and fall months, when southerly and southeasterly flow are commonplace in south-central and coastal Texas (*Texas Commission on Environmental Quality*, 2015). Supporting Table ~~S3~~ S2 shows the observed alkane enhancements for each day, along with the alkane emission estimates (Sect. 3.3) and meteorology. ~~Briefly, the~~ The median alkane enhancements for the set of 68 days were as follows: ethane – 2.4 ppb with an interquartile range (IQR) of 2.0-3.1 ppb; propane – 1.9 (IQR of 1.4-2.5) ppb; *n*-butane – 0.8 (0.6-1.1) ppb; and isobutane – 0.4 (0.3-0.5) ppb. All observed alkane enhancements were positive.

Supporting Fig. S2 shows a timeline of the afternoon ethane mixing ratios at both Oak Park and Floresville for the set of 68 days with southeasterly flow. Ethane mixing ratios during the warm season (summer and fall) were generally low at both Oak Park and Floresville and higher at the two sites during the cool season (winter and spring). This seasonal variability conforms with current understanding of temporal hydrocarbon variability (*Helmig et al.*, 2016). The enhanced photochemical oxidation of ethane during the summer months explains the low background ethane observed in the onshore flow in Oak Park. (*Haman et al.*, 2012).

3.3 Partitioning of alkane sources

The enhancements of propane, butanes, and pentanes were highly correlated with ethane enhancements between Oak Park and Floresville, suggesting a co-emission from sources of natural gas. The strongest correlation was observed between ethane and propane (Fig. 3). Pentanes (and to a lesser extent, butanes) may be impacted by emissions from automotive traffic (*Tsai et al.*, 2006; *Ho et al.*, 2009; *Simpson et al.*, 2012) and ~~air~~ chemistry, as pentanes have atmospheric lifetimes of 1.5 days (298 K, $[OH] = 2.0 \times 10^6$) (*Atkinson and Arey*, 2003). Therefore, pentanes were not used to partition emissions from ~~RNG and TG~~ raw natural gas and tank gas sources and an emission rate was not calculated. Nonetheless, the isopentane to n-pentane enhancement ratio between Oak Park and Floresville (~~Fig. S2~~supporting Fig. S3), which will remain close to constant despite chemistry, was 1.17 ($p < 0.001$), indicating that alkane emissions are largely influenced by oil and gas (*Gilman et al.*, 2013; *Swarthout et al.*, 2016). By comparison, the same ratio at the Old Highway 90 site in San Antonio was 2.25, which falls within the bounds of the traffic-emission driven urban pentane ratio. For these reasons, the enhancements of short-chain alkanes during advection over the EFS are most likely dominated by emissions from oil and gas production. Note that the alkane enhancement distributions (Table ~~S3~~S2) are skewed, with the upper bounds likely representing the influence of individual plumes from UOG exploration activities.

Propane-to-ethane ratios were used to partition emissions from ~~RNG and TG~~ raw natural gas and tank gas sources. Within each Monte Carlo simulation, the fraction of the ethane enhancement from ~~RNG~~ raw natural gas sources varied largely, with the 95% confidence interval often bounded by physically unreasonable numbers. This was due to the large uncertainty of the propane-to-ethane ratios in ~~RNG and TG~~ raw natural gas and tank gas samples. However, the median fraction for each day showed less variability, with a median of 46% (IQR of 34-52%) of ethane attributed to ~~RNG~~ raw natural gas sources. The ~~RNG~~ raw natural gas source estimate was significantly more ~~constraint~~ constrained due to the availability of raw gas composition data, while no tank gas composition data were available for the EFS. The median methane enhancement estimate for all 68 days was 8.9 (IQR of 6.8-12.2) ppb, of which 76% (IQR of 66-80%) was attributed to ~~RNG~~ raw natural gas sources. Higher alkanes were also partitioned, and their relative contributions from ~~RNG~~ raw natural gas sources over the 68 days were: propane – 18% (IQR of 11-22%); *n*-butane – 10% (7-13%); and isobutane – 17% (11-21%). While the majority of methane emissions were due to emissions of raw natural gas, liquid storage tanks dominated the emissions of these higher alkanes

The ~~RNG~~ raw natural gas source fractions based on *n*-butane-to-ethane and isobutane-to-ethane ratios often did not match that of propane-to-ethane for individual days, although their probability distribution functions over all days overlapped (supporting Fig. ~~S3~~S4), suggesting that the partitioning is consistent and reasonable when integrated over multiple sources in a larger region. It should be noted, however, that Floresville is located closer to the oil producing window of the EFS than the gas producing window and the influence of raw natural gas production may be somewhat diminished due to this distance. Nonetheless, the contribution of ~~TG~~ tank gas sources to observed alkane enhancement ratios suggests that liquid storage tanks are a very important source of alkane emissions between Oak Park and Floresville, a result that agrees with hydrocarbon emission studies in many other liquids-rich U.S. shale plays.

3.4 Mass balance results

3.4.1 Dispersion plumes and upwind production areas

Figure 4 shows a representative backward dispersion plume output overlaid on a gridded production map to illustrate the relative emissions estimate. Emission plumes for other days varied in width, but generally overlapped the same counties. Since the EFS acts as a line source, all overlapping grid cells with non-zero production numbers were weighed equally. This area includes portions of several counties in the EFS – notably Atascosa, Bee, Karnes, Live Oak, and Wilson counties – which are members of AACOG (*Alamo Area Council of Governments (AACOG)*, 2015). *Pacsi et al.* (2015) found that, among all counties in the EFS, NO_x emissions from these counties had the greatest impact on ozone enhancement in Bexar County, home to the city of San Antonio. For these reasons, quantifying emissions from these counties in the core of the EFS is a particularly important step in assessing the air quality impacts of oil and gas operations in the San Antonio area. While production was considered over all grid cells overlapped by the backward dispersion plume, counties in the EFS dominate regional production and are likely responsible for the observed alkane enhancements.

3.4.2 Emission estimates

The alkane emissions estimates from the upwind production regions were estimated using a Monte Carlo simulation for each day, and the distributions of the median emission rates were found as follows: methane – 200 (IQR of 141-380) $\times 10^3$ kg day^{-1}; ethane – 98 (70-169) × 10^3 kg day^{-1}; propane – 110 (77-187) × 10^3 kg day^{-1}; *n*-butane – 65 (46-108) × 10^3 kg day^{-1}; and isobutane – 28 (19-53) × 10^3 kg day^{-1}. In comparison, VOC emissions estimated by AACOG (not including ethane) for the set of central EFS counties from southeast to southwest of Floresville were only 88 × 10^3 kg day^{-1} for calendar year 2012 (*Alamo Area Council of Governments (AACOG)*, 2013b), while total EFS VOC emissions were estimated at ~~206~~ 203 × 10^3 kg day^{-1}. This suggest that EFS VOC emissions used in ozone modeling may be underestimated by at least a factor of two, most likely more. To estimate relative methane losses, the ~~RNG-only~~ raw natural gas-only and the total mass emission of methane were converted into a volume of natural gas using the ideal gas law at standard temperature and pressure for natural gas volume reporting (*Texas Statutes*, 1977) and a natural gas methane content based on available ~~RNG~~ raw natural gas data (Table 2). The volume of the emitted natural gas was then compared to the produced natural gas at gas wells and associated gas at oil wells in the production reference area outlined in Sect. 2.4 and 3.4.1. Emission rates for individual days were highly uncertain, with the bounds of the 95% confidence interval often spanning an order of magnitude (supporting Fig. ~~S4~~S5). However, median emission rates over all days were less variable, with a median total emission rate of 2.2% (IQR of 0.9-4.9%) and ~~an RNG-only~~ a raw natural gas-only emission rate of 0.8% (0.5-1.4%). While the EPA's estimated emission rate of ~~1.1~~1.6% falls within the bounds of our total emission estimate, our median emission rate exceeds that of the EPA~~by a factor of two. The EPA's emission inventory is reasonable when only emissions from RNG sources are considered. The emission~~ . The emission rate displayed no significant trend over ~~time~~the 2013-2015 time period, and its correlations with independent meteorological variables and ethane enhancement over time were weak (supporting Fig. ~~S5 and~~ S6 and S7), which is to be expected for a continuous anthropogenic emission source. However, the uncertainty in the emission rate within each Monte Carlo simulation showed a strong dependence on the uncertainty of the ethane content in ~~RNG~~ raw natural gas samples, and, to a lesser extent, meteorology (supporting Fig. ~~S7~~S8). We find that the lack of data regarding the composition of both raw natural gas and vented gases from liquid storage tanks impedes a higher precision top-down emission rate estimate. Nonetheless, repeated emission estimates over a large set of days show consistent and reasonable emission rates that can be attributed to UOG operations.

4 Conclusions

Our study used ethane as a tracer for alkane emissions from UOG emissions for a region in southern Texas, where oil and gas production is dominated by the core of the EFS. Data from the TCEQ show that alkane mixing ratios downwind from the shale have increased in tandem with oil and gas production rates in the EFS. This trend, along with the strong correlation of ethane enhancements with both propane and butane enhancements, show that emissions from UOG production in the EFS are responsible for the observed alkane enhancements across the shale. Using a mass balance approach and a Monte Carlo error estimation, we calculate ethane emissions of 98 (IQR of 70-169) × 10^3 kg day^{-1} from areas between the Texas Coast and the downwind receptor site in Floresville ~~.~~ for a set of 68 days from August 2013 through August 2015. Using typical ethane-to-methane ratios, we estimate methane emissions of 200 (141-380) × 10^3 kg day^{-1}. These emissions represent 2.2% (0.9-4.9%) of the produced natural gas – including associated gas at oil wells – in the region upwind of Floresville. We show through the partitioning of ~~RNG and TG emissions that RNG~~ raw natural gas and tank gas emissions that raw natural gas sources account for three quarters of all methane emissions, with ~~an RNG only~~ a raw natural gas-only relative emission rate of 0.8% (0.5-1.4%). ~~TG sources~~ Note that these emission rates are expressed as a fraction of produced natural gas as opposed to produced energy. In liquids-rich shale plays such as the EFS, expressing emission rates as a fraction of produced energy may be a more appropriate measure of emissions, especially when comparing energy losses to other sources (e.g. coal). However, our findings suggest that energy losses in the form of VOC emissions may be an important consideration when estimating energy losses from liquids-rich shale plays.

We find that tank gas sources account for more than half of higher alkane emissions – notably 90% of *n*-butane emissions. Since the petroleum production in this region is dominated by counties within the EFS, we conclude that UOG activities in the shale area are largely responsible for these emissions. Our emission rate estimate falls within the bounds of many other top-down studies ~~, and while our estimated RNG only emission overlaps with~~ and overlaps the EPA's ~~1.1% methane inventory estimate~~most recent methane emissions estimates from its 2016 greenhouse gas inventory (*U.S. Environmental Protection Agency*, 2016). However, our median total emission rate including ~~TG sources exceeds it by a factor of two~~tank gas sources exceeds the EPA's emission rate. This suggests that methane emissions in the EFS ~~are~~ may be higher than current average nationwide emission rates in bottom-up inventories. Similarly, NMVOCs co-emitted with methane are likely underreported and underestimated in inventories used for air quality modeling studies in southern Texas. The partitioning of emissions from raw natural gas sources and liquid storage tanks confirms that tank gas is an important source of short-chain alkane emissions in the EFS, with the enhancement of propane during air mass transport over the EFS nearly as large as that of ethane. Furthermore, a recent assessment of hemispheric short-chain hydrocarbon emission trends highlighted an unknown source with a methane/ethane ratio that is lower than that of RNG (*Helmig et al.*, 2016). Existing data for ~~RNG and TG~~ raw natural gas and tank gas compositions show that the typical methane/ethane ratio in ~~TG~~ tank gas emissions are relatively low compared to that of ~~RNG~~ raw natural gas emissions. Our results show that emissions of alkanes from liquid storage tanks account for 24% of methane, 54% of ethane, 82% of propane, 90% of *n*-butane, and 83% of isobutane emissions from the EFS. These emissions are likely to contribute to the unknown NMHC source identified by *Helmig et al.* (2016).

While alkanes have been shown to dominate the OH reactivity at Floresville, the scarcity of trace gas measurements within the EFS prevents more thorough NMVOC emissions estimates. Our calculations indicate that propane and butanes emissions alone exceed current inventory numbers for the EFS by approximately a factor of two, which can have significant impacts on ozone modeling, particularly if NO_x emissions are underestimated as well. Hence, we stress the need for increased spatial coverage of VOC, NOx, and greenhouse gas monitoring in and around the shale area to improve upon existing emission inventories. Such improvements are needed before the air quality impacts of the EFS can be accurately quantified. As the unconventional oil and gas industry in the EFS continues to grow, the climate and air quality impacts associated with emissions from the shale need to be addressed. This is especially true as the San Antonio metropolitan area ~~is likely to~~ may be designated as a nonattainment area by the EPA. However, existing emissions estimates are uncertain and variable, and need to be improved before the impacts on air quality can be quantified.

Author contributions. G. Roest gathered and analyzed data, performed the emissions estimate, and prepared the manuscript. G. Schade provided input for the emissions estimate, performed the plume dispersion analyses and plume width estimate, and assisted in manuscript preparation.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We would like to thank TCEQ for providing automated GC data in the summer and fall of 2014, which are now available online at http://www17.tceq.texas.gov/tamis/. Funding was provided internally within the Department of Atmospheric Sciences at Texas A&M University.

References

Alamo Area Council of Governments (AACOG) (2013a), Development of the Extended June 2006 Photochemical Modeling Episode, *Tech. Rep.*, San Antonio, Texas.

Alamo Area Council of Governments (AACOG) (2013b), Oil and Gas Emission Inventory, Eagle Ford Shale, *Tech. Rep.*, San Antonio, Texas.

Alamo Area Council of Governments (AACOG) (2014), 2014 Ozone Watch, *Tech. Rep.*, San Antonio, Texas.

Alamo Area Council of Governments (AACOG) (2015), Cities and Counties, San Antonio, Texas.

Atkinson, R., and J. Arey (2003), Atmospheric Degradation of Volatile Organic Compounds, *Chem. Rev.*, *103*(12), 4605–4638, doi:10.1021/cr0206420.

Bonney, J. (2014), Port Freeport's Chemical Reaction, *J. of Commerce (15307557)*, *15*(12), 38.

Brantley, H. L., E. D. Thoma, W. C. Squier, B. B. Guven, and D. Lyon (2014), Assessment of Methane Emissions from Oil and Gas Production Pads using Mobile Measurements, *Environ. Sci. Technol.*, *48*(24), 14,508–14,515, doi:10.1021/es503070q.

Caulton, D. R., P. B. Shepson, R. L. Santoro, J. P. Sparks, R. W. Howarth, A. R. Ingraffea, M. O. L. Cambaliza, C. Sweeney, A. Karion, K. J. Davis, B. H. Stirm, S. A. Montzka, and B. R. Miller (2014), Toward a better understanding and quantification of methane emissions from shale gas development, *PNAS*, *111*(17), 6237–6242, doi:10.1073/pnas.1316546111.

Duncan, B. N., L. N. Lamsal, A. M. Thompson, Y. Yoshida, Z. Lu, D. G. Streets, M. M. Hurwitz, and K. E. Pickering (2016), A space-based, high-resolution view of notable changes in urban NOx pollution around the world (2005–2014), *J. Geophys. Res. Atmos.*, p. 2015JD024121, doi:10.1002/2015JD024121.

Dunn, O. J. (1964), Multiple Comparisons Using Rank Sums, *Technometrics*, *6*(3), 241–252, doi:10.1080/00401706.1964.10490181.

ENVIRON International Corporation (2010), Upstream Oil and Gas Tank Emission Measurements: TCEQ Project 2010 - 39, *Tech. Rep. 06-17477X*, Novato, CA.

European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL) (2009), Emission Database for Global Atmospheric Research (EDGAR), release version 4.2, *Tech. Rep.*

Field, R. A., J. Soltis, and S. Murphy (2014), Air quality concerns of unconventional oil and natural gas production, *Environ. Sci. Processes Impacts*, *16*(5), 954, doi:10.1039/c4em00081a.

Franco, B., E. Mahieu, L. K. Emmons, Z. A. Tzompa-Sosa, E. V. Fischer, K. Sudo, B. Bovy, S. Conway, D. Griffin, J. W. Hannigan, K. Strong, and K. A. Walker (2016), Evaluating ethane and methane emissions associated with the development of oil and natural gas extraction in North America, *Environ. Res. Lett.*, *11*(4), 044,010, doi:10.1088/1748-9326/11/4/044010.

Gebrekidan, S. (2011), Analysis: 100 years after boom, shale makes Texas oil hot again, *Reuters*.

Gilman, J. B., B. M. Lerner, W. C. Kuster, and J. A. de Gouw (2013), Source Signature of Volatile Organic Compounds from Oil and Natural Gas Operations in Northeastern Colorado, *Environ. Sci. Technol.*, *3*(47), 1297–1305, doi:10.1021/es304119a.

Haman, C. L., Lefer, B., and Morris, G. A.: Seasonal Variability in the Diurnal Evolution of the Boundary Layer in a Near-Coastal Urban Environment, Journal of Atmospheric and Oceanic Technology, 29, 697–710, doi:10.1175/JTECH-D-11-00114.1, http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-11-00114.1, 2012.

Helmig, D., S. Rossabi, J. Hueber, P. Tans, S. A. Montzka, K. Masarie, K. Thoning, C. Plass-Duelmer, A. Claude, L. J. Carpenter, A. C. Lewis, S. Punjabi, S. Reimann, M. K. Vollmer, R. Steinbrecher, J. W. Hannigan, L. K. Emmons, E. Mahieu, B. Franco, D. Smale, and

A. Pozzer (2016), Reversal of global atmospheric ethane and propane trends largely due to US oil and natural gas production, *Nat. Geosci.*, *advance online publication*, doi:10.1038/ngeo2721.

Hendler, A., Nunn, J., Lundeen, J., and McKaskle, R.: VOC emissions from oil and condensate storage tanks, Tech. rep., prepared for Texas Environmental Research Consortium, The Woodlands, TX, http://files.harc.edu/projects/airquality/projects/H051C/H051Cfinalreport.pdf, 2009.

Ho, K. F., S. C. Lee, W. K. Ho, D. R. Blake, Y. Cheng, Y. S. Li, S. S. H. Ho, K. Fung, P. K. K. Louie, and D. Park (2009), Vehicular emission of volatile organic compounds (VOCs) from a tunnel study in Hong Kong, *Atmos. Chem. Phys.*, *9*(19), 7491–7504, doi:10.5194/acp-9-7491-2009.

Hu, L., S. A. Montzka, J. B. Miller, A. E. Andrews, S. J. Lehman, B. R. Miller, K. Thoning, C. Sweeney, H. Chen, D. S. Godwin, K. Masarie, L. Bruhwiler, M. L. Fischer, S. C. Biraud, M. S. Torn, M. Mountain, T. Nehrkorn, J. Eluszkiewicz, S. Miller, R. R. Draxler, A. F. Stein, B. D. Hall, J. W. Elkins, and P. P. Tans (2015), U.S. emissions of HFC-134a derived for 2008–2012 from an extensive flask-air sampling network, *J. Geophys. Res. Atmos.*, *120*(2), 2014JD022,617, doi:10.1002/2014JD022617.

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

Karion, A., C. Sweeney, E. A. Kort, P. B. Shepson, A. Brewer, M. Cambaliza, S. A. Conley, K. Davis, A. Deng, M. Hardesty, S. C. Herndon, T. Lauvaux, T. Lavoie, D. Lyon, T. Newberger, G. Pétron, C. Rella, M. Smith, S. Wolter, T. I. Yacovitch, and P. Tans (2015), Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region, *Environ. Sci. Technol.*, *49*(13), 8124–8131, doi:10.1021/acs.est.5b00217.

Kort, E. A., M. L. Smith, L. T. Murray, A. Gvakharia, A. R. Brandt, J. Peischl, T. B. Ryerson, C. Sweeney, and K. Travis (2016), Fugitive emissions from the Bakken shale illustrate role of shale production in global ethane shift: Ethane emissions from the Bakken shale, *Geophys. Res. Lett.*, doi:10.1002/2016GL068703.

Kruskal, W. H., and W. A. Wallis (1952), Use of Ranks in One-Criterion Variance Analysis, *J. Am. Stat. Assoc.*, *47*(260), 583–621, doi:10.1080/01621459.1952.10483441.

Lyon, D. R., D. Zavala-Araiza, R. A. Alvarez, R. Harriss, V. Palacios, X. Lan, R. Talbot, T. Lavoie, P. Shepson, T. I. Yacovitch, S. C. Herndon, A. J. Marchese, D. Zimmerle, A. L. Robinson, and S. P. Hamburg (2015), Constructing a Spatially Resolved Methane Emission Inventory for the Barnett Shale Region, *Environ. Sci. Technol.*, *49*(13), 8147–8157, doi:10.1021/es506359c.

Lyon, D. R., R. A. Alvarez, D. Zavala-Araiza, A. R. Brandt, R. B. Jackson, and S. P. Hamburg (2016), Aerial Surveys of Elevated Hydrocarbon Emissions from Oil and Gas Production Sites, *Environ. Sci. Technol.*, *50*(9), 4877–4886, doi:10.1021/acs.est.6b00705.

Mesinger, F., Geoff DiMego, Eugenia Kalnay, Kenneth Mitchell, Perry C. Shafran, Wesley Ebisuzaki, Dusan Jovic, Jack Woollen, Eric Rogers, Ernesto H. Berbery, Michael B. Ek, Yun Fan, Robert Grumbine, Wayne Higgins, Hong Li, Ying Lin, Geoff Manikin, David Parrish, and Wei Shi (2005), NORTH AMERICAN REGIONAL REANALYSIS: A long-term, consistent, high-resolution climate dataset for the North American domain, as a major improvement upon the earlier global reanalysis datasets in both resolution and accuracy, *Bull. Am. Meteorol. Soc.*.

Miller, S. M., S. C. Wofsy, A. M. Michalak, E. A. Kort, A. E. Andrews, S. C. Biraud, E. J. Dlugokencky, J. Eluszkiewicz, M. L. Fischer, G. Janssens-Maenhout, B. R. Miller, J. B. Miller, S. A. Montzka, T. Nehrkorn, and C. Sweeney (2013), Anthropogenic emissions of methane in the United States, *PNAS*, *110*(50), 20,018–20,022, doi:10.1073/pnas.1314392110.

National Centers for Environmental Prediction (NCEP), Eta Data Assimilation System (EDAS40).

~~Office of Air Quality Planning and Standards, U. E. (2012), 2011 NationalEmissionsInventoryData& Documentation.~~

Olaguer, E. P. (2012), The potential near-source ozone impacts of upstream oil and gas industry emissions, *J. Air Waste Manage. Assoc.*, *62*(8), 966–977, doi:10.1080/10962247.2012.688923.

Pacsi, A. P., Y. Kimura, G. McGaughey, E. C. McDonald-Buller, and D. T. Allen (2015), Regional Ozone Impacts of Increased Natural Gas Use in the Texas Power Sector and Development in the Eagle Ford Shale, *Environ. Sci. Technol.*, *49*(6), 3966–3973, doi:10.1021/es5055012.

Peischl, J., T. B. Ryerson, K. C. Aikin, J. A. de Gouw, J. B. Gilman, J. S. Holloway, B. M. Lerner, R. Nadkarni, J. A. Neuman, J. B. Nowak, M. Trainer, C. Warneke, and D. D. Parrish (2015), Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions: CH4 emissions from shale gas production, *J. Geophys. Res. Atmos.*, *120*(5), 2119–2139, doi:10.1002/2014JD022697.

Pikelnaya, O., J. H. Flynn, C. Tsai, and J. Stutz (2013), Imaging DOAS detection of primary formaldehyde and sulfur dioxide emissions from petrochemical flares, *J. Geophys. Res. Atmos.*, *118*(15), 8716–8728, doi:10.1002/jgrd.50643.

Pouillot, R., and M.-L. Delignette-Muller (2010), Evaluating variability and uncertainty in microbial quantitative risk assessment using two R packages, *Int. J. Food Microbiol.*, *142*(3), 330–40.

Pring, M. (2012), Condensate Tank Oil and Gas Activities, *Final Report ERG. No. 0292.01.011.001*, Eastern Research Group, Inc., Morrisville, North Carolina.

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

Pétron, G., A. Karion, C. Sweeney, B. R. Miller, S. A. Montzka, G. J. Frost, M. Trainer, P. Tans, A. Andrews, J. Kofler, D. Helmig, D. Guenther, E. Dlugokencky, P. Lang, T. Newberger, S. Wolter, B. Hall, P. Novelli, A. Brewer, S. Conley, M. Hardesty, R. Banta, A. White, D. Noone, D. Wolfe, and R. Schnell (2014), A new look at methane and nonmethane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin: Hydrocarbon emissions in oil & gas basin, *J. Geophys. Res. Atmos.*, *119*(11), 6836–6852, doi:10.1002/2013JD021272.

Pétron, G., G. J. Frost, C. Sweeney, A. Karion, B. R. Miller, S. A. Montzka, E. Dlugokencky, P. Lang, A. Hirsch, S. Conley, J. Kofler, A. Andrews, D. Guenther, P. Novelli, T. Conway, K. Masarie, S. Wolter, T. Newberger, J. Higgs, B. Hall, R. Schnell, P. Tans (2012), Estimation of emissions from oil and natural gas operations in northeastern Colorado, Tampa, Florida.

Railroad Commission of Texas, Oil and Gas Production Data, *Dataset*, Austin, Texas.

Randerson, J. T., Van Der Werf, G.R., Giglio, L., Collatz, G.J., and Kasibhatla, P.S. (2015), Global Fire Emissions Database, Version 4, (GFEDv4), doi:10.3334/ORNLDAAC/1293.

Ryan, M. (2014), Boom times on the bay: Dow, other firms fuel development south of Houston, *Houston Business Journal*.

Schade, G. W., and G. Roest (2016), Analysis of non-methane hydrocarbon data from a monitoring station affected by oil and gas development in the Eagle Ford shale, Texas, *Elem. Sci. Anth.*, *4*, 000,096, doi:10.12952/journal.elementa.000096.

Schade, G. W., and G. S. Roest (2015), Is the Shale Boom Reversing Progress in Curbing Ozone Pollution?, *Eos*, *96*, doi:10.1029/2015EO028279.

Schmid, P. and Niyogi, D.: A Method for Estimating Planetary Boundary Layer Heights and Its Application over the ARM Southern Great Plains Site, Journal of Atmospheric and Oceanic Technology, 29, 316–322, doi:10.1175/JTECH-D-11-00118.1, http://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-11-00118.1, 2012.

Schneising, O., J. P. Burrows, R. R. Dickerson, M. Buchwitz, M. Reuter, and H. Bovensmann (2014), Remote sensing of fugitive methane emissions from oil and gas production in North American tight geologic formations: Remote sensing of fugitive methane emissions from oil and gas production, *Earth's Future*, *2*(10), 548–558, doi:10.1002/2014EF000265.

Schwietzke, S., W. M. Griffin, H. S. Matthews, and L. M. P. Bruhwiler (2014), Natural Gas Fugitive Emissions Rates Constrained by Global Atmospheric Methane and Ethane, *Environ. Sci. Technol.*, *48*(14), 7714–7722, doi:10.1021/es501204c.

Simpson, I. J., M. P. Sulbaek Andersen, S. Meinardi, L. Bruhwiler, N. J. Blake, D. Helmig, F. S. Rowland, and D. R. Blake (2012), Long-term decline of global atmospheric ethane concentrations and implications for methane, *Nature*, *488*(7412), 490–494, doi:10.1038/nature11342.

Smith, M. L., E. A. Kort, A. Karion, C. Sweeney, S. C. Herndon, and T. I. Yacovitch (2015), Airborne Ethane Observations in the Barnett Shale: Quantification of Ethane Flux and Attribution of Methane Emissions, *Environ. Sci. Technol.*, *49*(13), 8158–8166, doi:10.1021/acs.est.5b00219.

Stein, A. F., R. R. Draxler, G. D. Rolph, B. J. B. Stunder, M. D. Cohen, and F. Ngan (2015), NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System, *Bull. Amer. Meteor. Soc.*, *96*(12), 2059–2077, doi:10.1175/BAMS-D-14-00110.1.

Strosher, M. T. (2000), Characterization of Emissions from Diffusion Flare Systems, *J. Air Waste Manage. Assoc.*, *50*(10), 1723–1733, doi:10.1080/10473289.2000.10464218.

Stull, R. B. (2009), *An introduction to boundary layer meteorology*, no. 13 in Atmospheric and oceanographic sciences library, 1st ed ed., Springer, New York.

Swarthout, R. F., R. S. Russo, Y. Zhou, B. M. Miller, B. Mitchell, E. Horsman, E. Lipsky, D. C. McCabe, E. Baum, and B. C. Sive (2015), Impact of Marcellus Shale Natural Gas Development in Southwest Pennsylvania on Volatile Organic Compound Emissions and Regional Air Quality, *Environ. Sci. Technol.*, *49*(5), 3175–3184, doi:10.1021/es504315f.

Tedesco, J., and J. Hiller (2014), Up in Flames, *San Antonio Express News*.

Texas Commission on Environmental Quality (2005), Standard Operating Procedure for the Perkin-Elmer Auto Gas Chromatograph for VOC Ozone Precursors Analysis (FOSTAT-026).

Texas Commission on Environmental Quality (2015), Wind Roses.

Texas Statutes (1977), Natural Resources Code, Title 3, Subtitle B, Chapter 91, Section 91.052.

Todd, M. (2011), Proposed Rulemaking - Oil and Gas Sector Regulations Standards of Performance for New Stationary Sources: Oil and Natural Gas Production and Natural Gas Transmission and Distribution, *Memorandum Docket ID No. EPA-HQ-OAR-2010-0505*, American Petroleum Institute, Washington, DC.

Tsai, W. Y., L. Y. Chan, D. R. Blake, and K. W. Chu (2006), Vehicular fuel composition and atmospheric emissions in South China: Hong Kong, Macau, Guangzhou, and Zhuhai, *Atmos. Chem. Phys.*, *6*(11), 3281–3288, doi:10.5194/acp-6-3281-2006.

Turner, A. J., D. J. Jacob, J. Benmergui, S. C. Wofsy, J. D. Maasakkers, A. Butz, O. Hasekamp, and S. C. Biraud (2016), A large increase in U.S. methane emissions over the past decade inferred from satellite data and surface observations, *Geophys. Res. Lett.*, *43*(5), 2016GL067,987, doi:10.1002/2016GL067987.

United States Committee on Extension to the Standard Atmosphere (1976), *U.S. standard atmosphere, 1976*, National Oceanic and Atmospheric Administration: for sale by the Supt. of Docs., U.S. Govt. Print. Off.

U.S. Centers for Disease Control (2014), Benzene - NIOSH Publications and Products.

U.S. Energy Information Administration (2017), Natural Gas Summary, ~~International EnergyStatistics,~~ *Tech. Rep.* https://www.eia.gov/dnav/ng/ng_sum_lsum_dcu_nus_a.htm.

U.S. Energy Information Administration (2016), Drilling productivity report for key tight oil and shale gas regions, *Tech. Rep.*

U.S. Environmental Protection Agency, Greenhouse Gas Reporting Program, *Tech. Rep.*

U.S. Environmental Protection Agency~~(~~: Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2013, Tech. Rep. EPA 430-R-15-004, Washington, DC, https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2013, 2015~~),~~ a.

U.S. Environmental Protection Agency: National Ambient Air Quality Standards for Ozone; Final Rule, http://www.gpo.gov/fdsys/pkg/FR-2015-10-26/pdf/2015-26594.pdf.

U.S. Environmental Protection Agency: Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2014 (2016), *Tech. Rep.*, EPA 430-R-16-002, Washington, DC, https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2014.

Xiao, Y., J. A. Logan, D. J. Jacob, R. C. Hudman, R. Yantosca, and D. R. Blake (2008), Global budget of ethane and regional constraints on U.S. sources, *J. Geophys. Res.*, *113*(D21), doi:10.1029/2007JD009415.

Zavala-Araiza, D., D. W. Sullivan, and D. T. Allen (2014), Atmospheric Hydrocarbon Emissions and Concentrations in the Barnett Shale Natural Gas Production Region, *Environ. Sci. Technol.*, *48*(9), 5314–5321, doi:10.1021/es405770h.

Zavala-Araiza, D., D. R. Lyon, R. A. Alvarez, K. J. Davis, R. Harriss, S. C. Herndon, A. Karion, E. A. Kort, B. K. Lamb, X. Lan, A. J. Marchese, S. W. Pacala, A. L. Robinson, P. B. Shepson, C. Sweeney, R. Talbot, A. Townsend-Small, T. I. Yacovitch, D. J. Zimmerle, and S. P. Hamburg (2015), Reconciling divergent estimates of oil and gas methane emissions, *PNAS*, *112*(51), 15,597–15,602, doi:10.1073/pnas.1522126112.

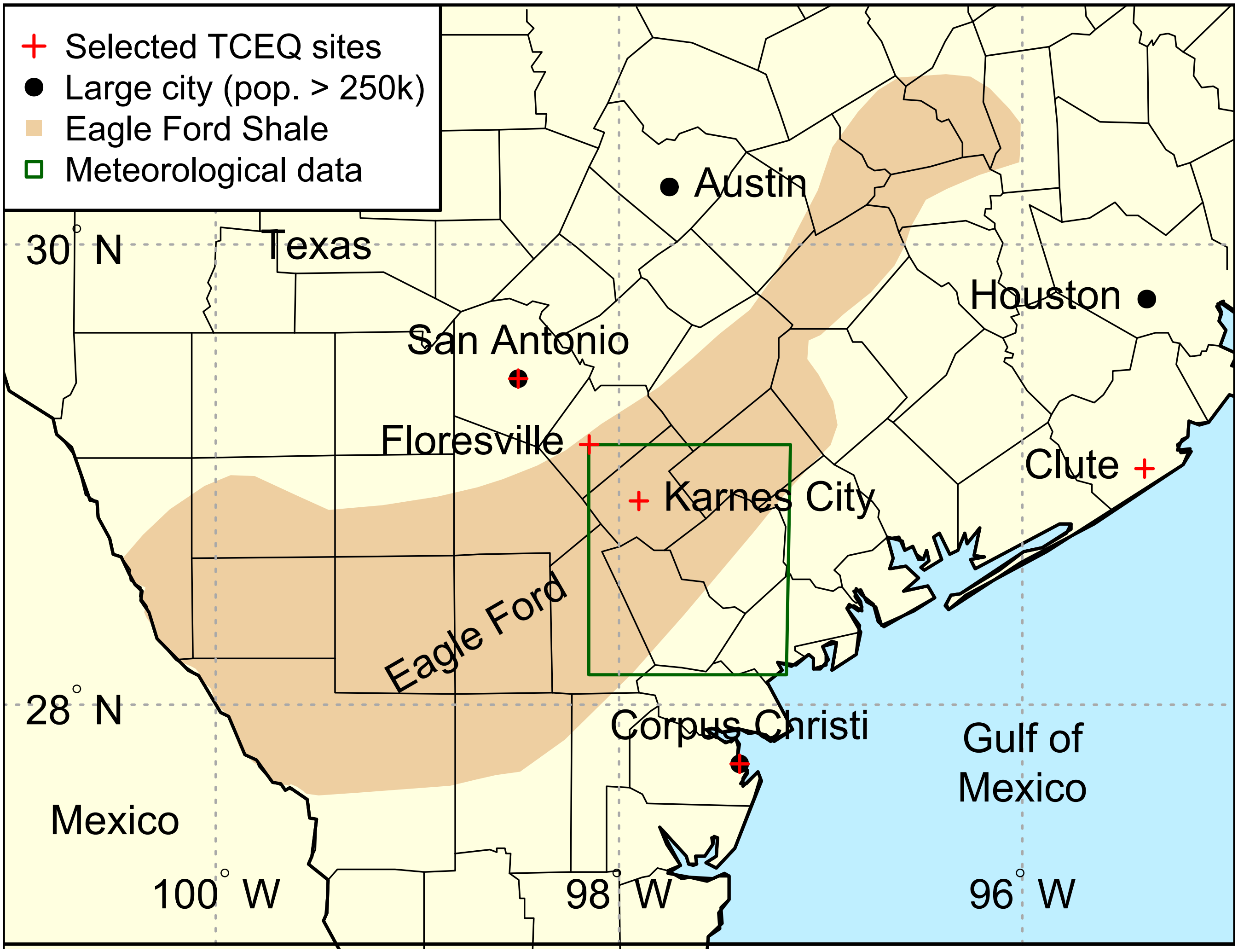

Figure 1. Selected TCEQ NMVOC monitoring sites and large cities near the Eagle Ford. The ~~red~~ green box shows the 1° latitude by 1° longitude box in which the meteorology was assessed.

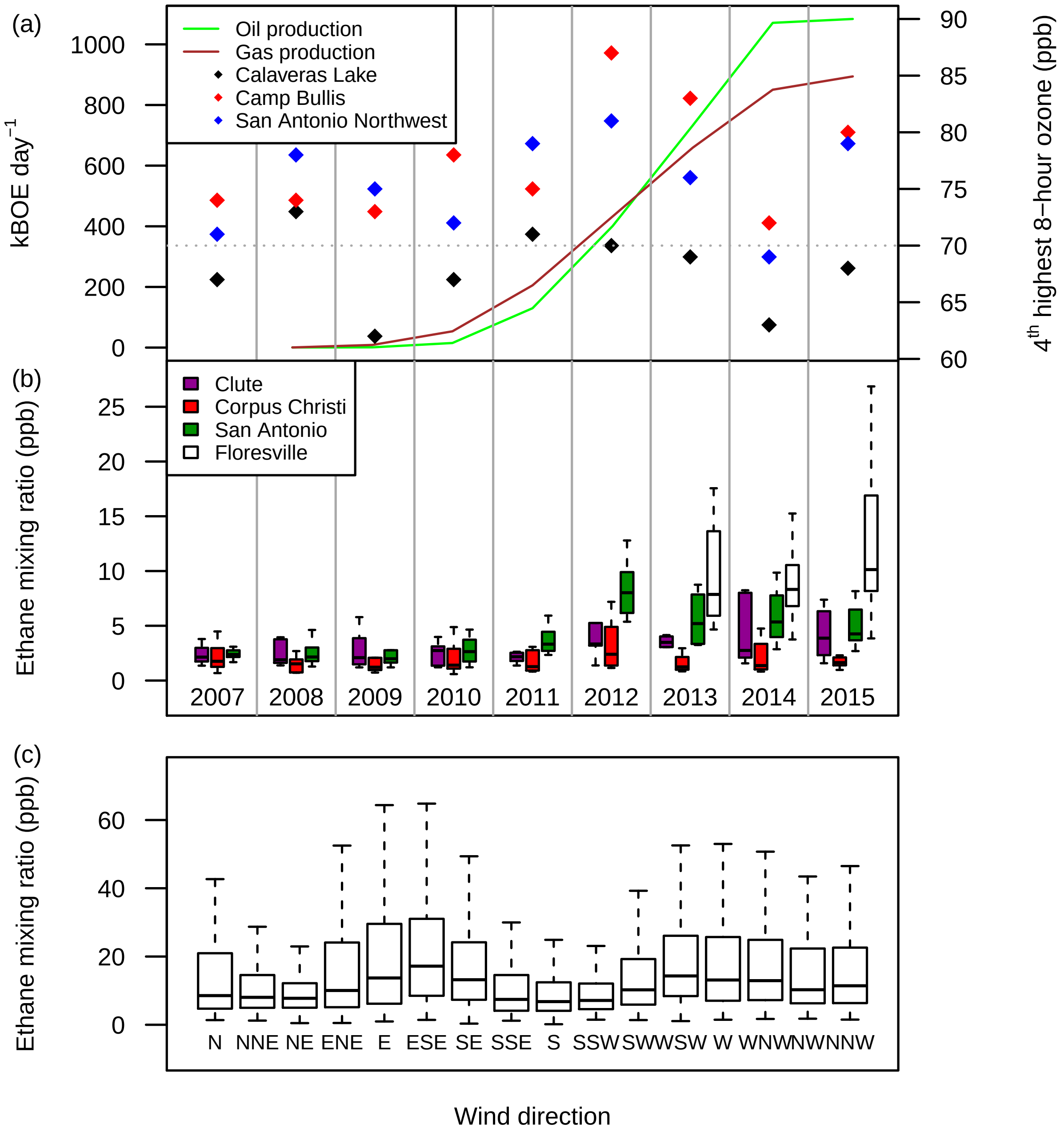

Figure 2. Adapted version of Fig. 2 in *Schade and Roest* (2015), updated to include data through 2015. (a) Oil and gas production rates in the Eagle Ford and 4th highest maximum 8-hour ozone values at 3 sites in San Antonio. (b) Timeline of 24-hour ethane mixing ratios at 4 sites near the Eagle Ford Shale. Days were used only if 3 out of 4 back-trajectories originating from San Antonio – Old Highway 90 were binned as southeasterly. Data at Floresville begins in July of 2013. (c) Ethane mixing ratios vs. wind direction at Floresville, with elevated mixing ratios under E to SE or SW to W winds, when Floresville is downwind of the Eagle Ford. Ethane is also elevated under NW winds, likely due to higher ethane in continental air masses and local emissions from the San Antonio metropolitan area.

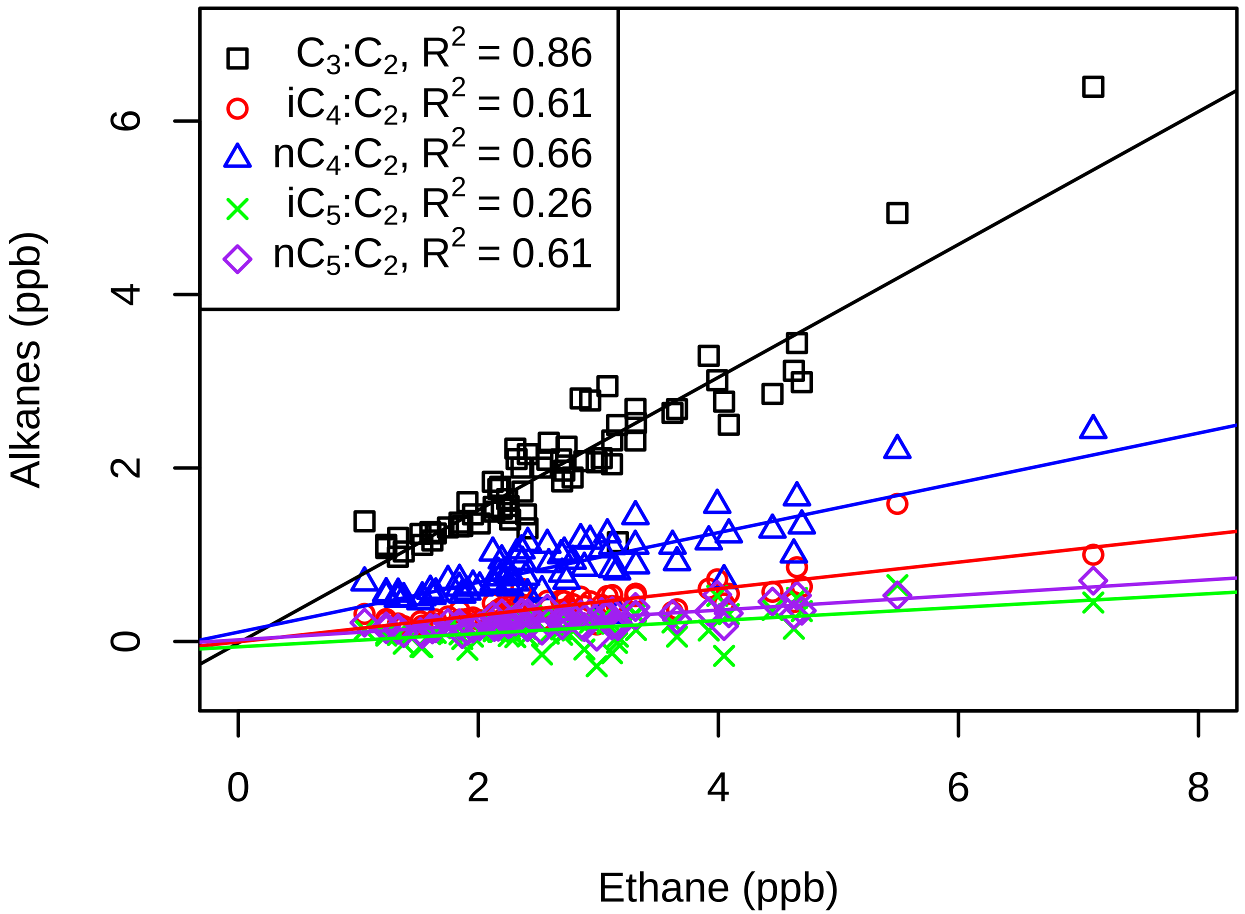

Figure 3. Correlations of propane, butane, and pentane enhancements with ethane enhancement. All correlations were highly statistically significant ($p < 0.001$). While ethane and propane showed the strongest correlation, all alkanes showed a positive correlation with ethane.

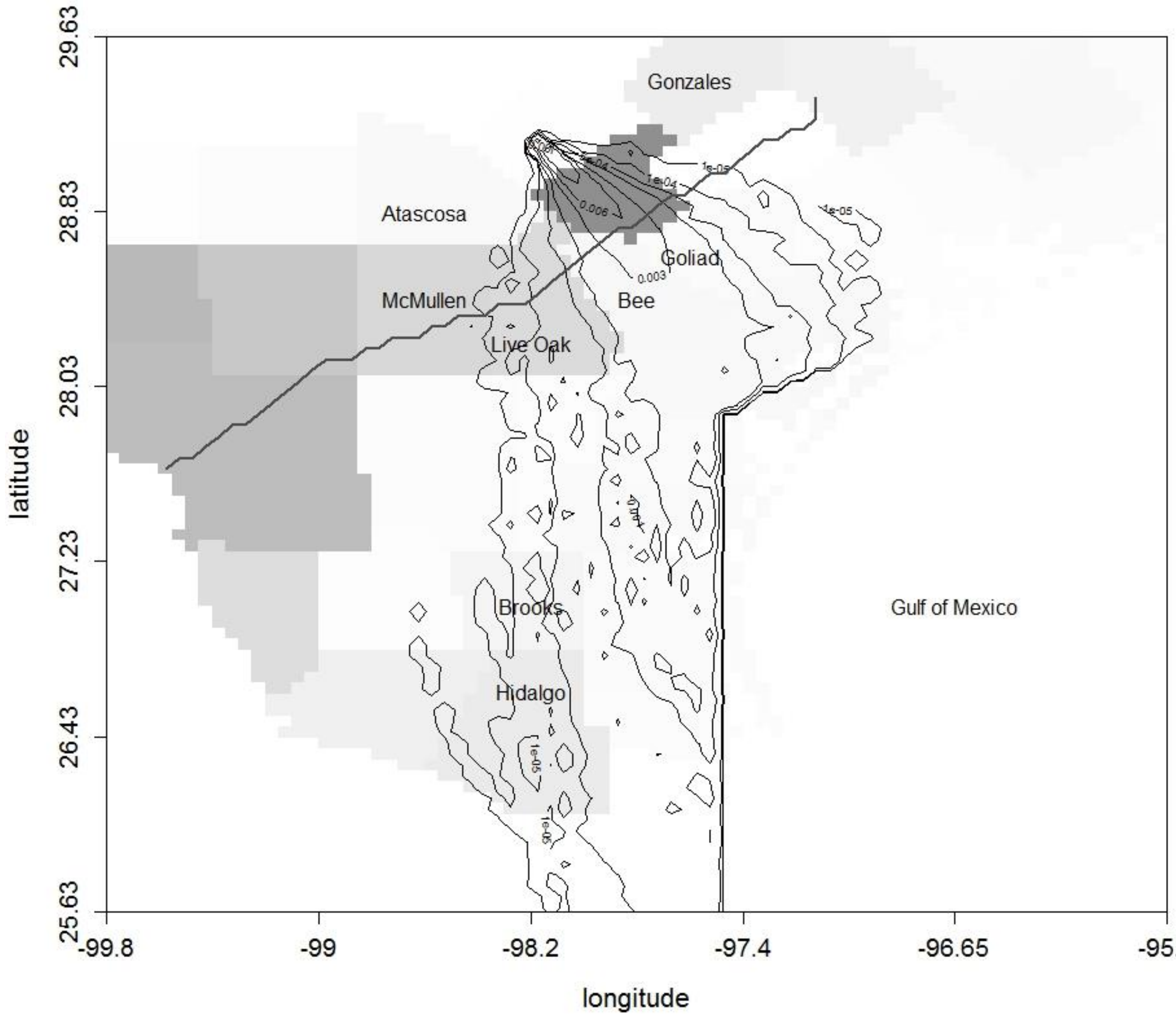

Figure 4. An example of an integral backward dispersion plume map created from 5000 particles released above the Floresville receptor location and followed backwards in time (20 hours) into the model lowest vertical level (<50 m agl) to assess surface emitter impacts. The raw map was normalized to its total after removing all surface impacts over the Gulf of Mexico. The grey shading underlying the plume map identifies counties with the sum of natural gas and associated gas production, darker shading indicating higher production rates. The darkest shading indicates Karnes County, not labeled for clarity. The dark grey, jagged line extending from the west-southwest near the Mexican border to the east-northeast south of Gonzales County marks the southern "edge" of the EFS.

Table 1. Description of TCEQ sites used in this study.

Site Name	AQS Code	Lat (°N)	Lon (°W)	Sample Duration	Samples Collected	Use in Study
Clute	480391003	29.01	95.40	24 hour	Once every 6 days	Long-term trends
Corpus Christi – Hillcrest	483550029	27.81	97.42	24 hour	Once every 6 days	Long-term trends
Corpus Christi – Oak Park	483550035	27.80	97.43	40 min	Hourly, automated	Emission estimate
Floresville Hospital Boulevard	484931038	29.13	98.15	40 min	Hourly, automated	Emission estimate
San Antonio – Old Highway 90	480290677	29.42	98.58	24 hour	Once every 6 days	Long-term trends

Table 2. Ethane content in raw natural gas and tank gas samples by mol percent and associated ethane/alkane ratios

Ratio	C_2 (mol %)	C_2/C_1	C_2/C_3	C_2/iC_4	C_2/nC_4	$C_2/(iC_5+nC_5)$
Raw Natural Gas	4.51[a]	0.05	2.20	9.40	8.84	11.00
	9.15[b]	0.11	2.97	8.55	9.24	9.63
	13.20[b]	0.17	2.63	12.34	10.08	18.86
	15.88[b]	0.22	2.55	36.93	12.70	31.76
Mean	10.69	0.14	2.59	16.80	10.22	17.79
Tank Gas	13.07[c]	0.84	0.75	2.52	1.08	0.47
	16.83[c]	0.72	1.13	3.21	1.59	0.78
	14.04[c]	0.61	0.89	3.32	1.45	0.65
	13.58[c]	0.47	0.96	4.02	1.54	0.64
Mean	13.48	0.66	0.93	3.27	1.42	0.64

[a] (*Pring*, 2012), [b] (*Todd*, 2011), [c] (*ENVIRON International Corporation*, 2010)

Introduction

This document contains a description of the methodology used to bin HYSPLIT trajectories based on air mass origins in Text S1, with Table S2 and Figure S1 outlining the polygons used to test the paths of trajectories. Additionally, TCEQ sites where VOC data are collected are described in Table S1 and input variables for the Monte Carlo simulation are outlined in Table S3. Figures S2 through S7 support the results in the main text.

Text S1. In order to identify days with mostly southeasterly flow, 48 hour back trajectories were obtained from the HYSPLIT model. The start times for these trajectories were 06:00, 12:00, and 18:00 UTC each calendar day, and 00:00 UTC the following calendar day. The origins of the trajectories were binned by their passage through a series of polygons, as shown in Figure S1. The vertices of the polygons are provided in Table S2.

Back trajectories ending at the San Antonio – Old Highway 90 site were used to identify southeasterly flow when assessing long term alkane trends at San Antonio – Old Highway 90, Clute, and Corpus Christi – Hillcrest. All trajectories that passed through Polygon 1 were assumed to have continental origins and were removed. Polygon 2 was selected to represent the central Texas Coast region, roughly extending from the coastal waters southeast of Corpus Christi to the waters south of Clute. Trajectories that did not pass through this polygon were also removed, leaving generally southeasterly trajectories of maritime origin remaining. Days with 3 out of 4 southeasterly trajectories were used to compare long term alkane trends at these sites.

For the alkane emission calculation using data from Corpus Christi – Oak Park and Floresville, back trajectories ending at Floresville were removed if they passed through Polygon 1. Again, this was to remove air masses which were influenced by continental emissions prior to moving ashore. Trajectories were also removed if they did not pass through Polygon 3, which encompasses Corpus Christi and the surrounding region of the Texas Coast. Again, days with 3 out of 4 trajectories were used to quantify the afternoon alkane enhancement between Corpus Christi – Oak Park and Floresville.

Table S1. Vertices of polygons used to bin HYSPLIT trajectories as described in Text S1.

Polygon	Vertex 1	Vertex 2	Vertex 3	Vertex 4	Vertex 5	Vertex 6
1	29.16° N, 96.12° W	29.16° N, 83.00° W	31.00° N, 83.00° W	31.00° N, 103.00° W	29.50° N, 103.00° W	29.50° N, 96.12° W
2	27.04° N, 96.75° W	27.43° N, 97.07° W	28.67° N, 95.50° W	28.28° N, 95.19° W	-	-
3	27.27° N, 97.48° W	27.81° N, 98.00° W	28.16° N, 97.64° W	27.62° N, 97.12° W	-	-

Table S2. Separate Excel file showing average alkane mixing ratios (ppb) during afternoon hours at Floresville and Corpus Christi – Hillcrest during 68 days with appropriate meteorological conditions. Also shown are input meteorology from the North American Regional Reanalysis (NARR) that were used as input for the mass balance approach and Monte Carlo simulation. Lastly, alkane emission rates and methane emissions relative to production are provided for each day.

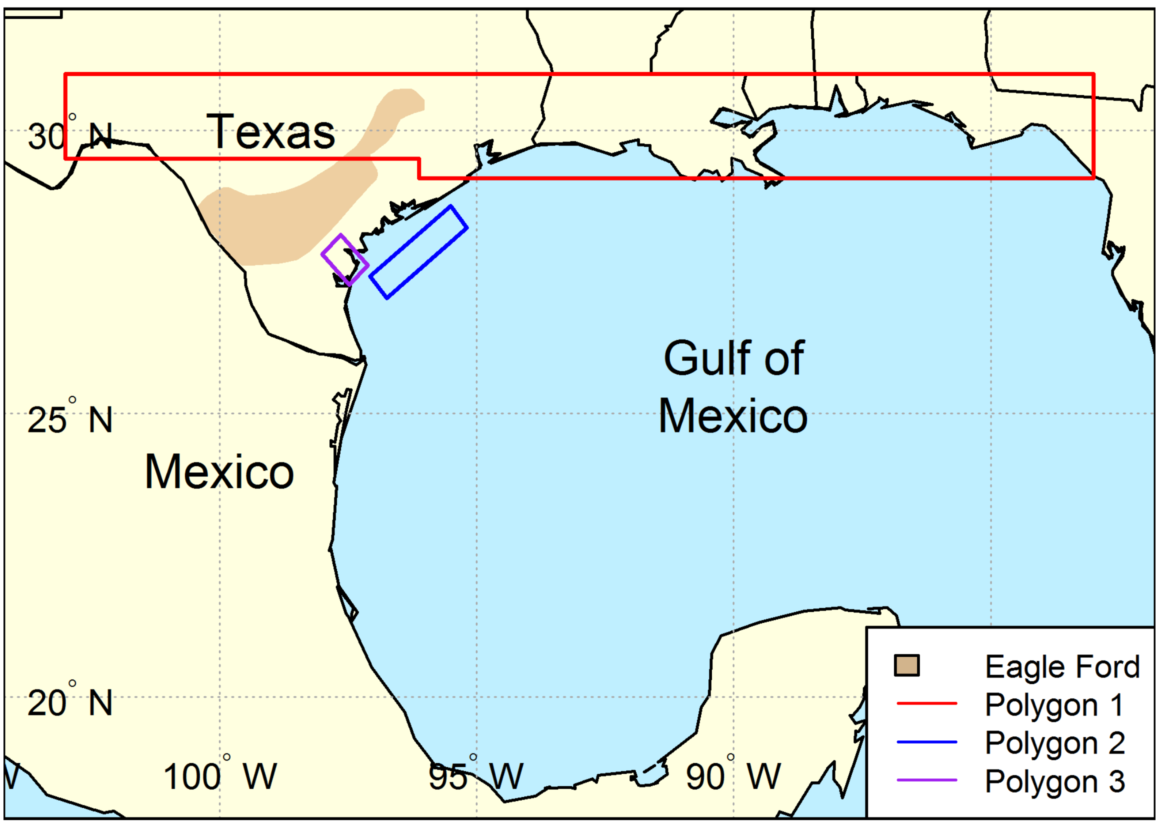

Figure S1. Polygons used to identify trajectories as southeasterly with maritime origins. See Text S1 for a description of the methods and Table S2 for corners of the vertices.

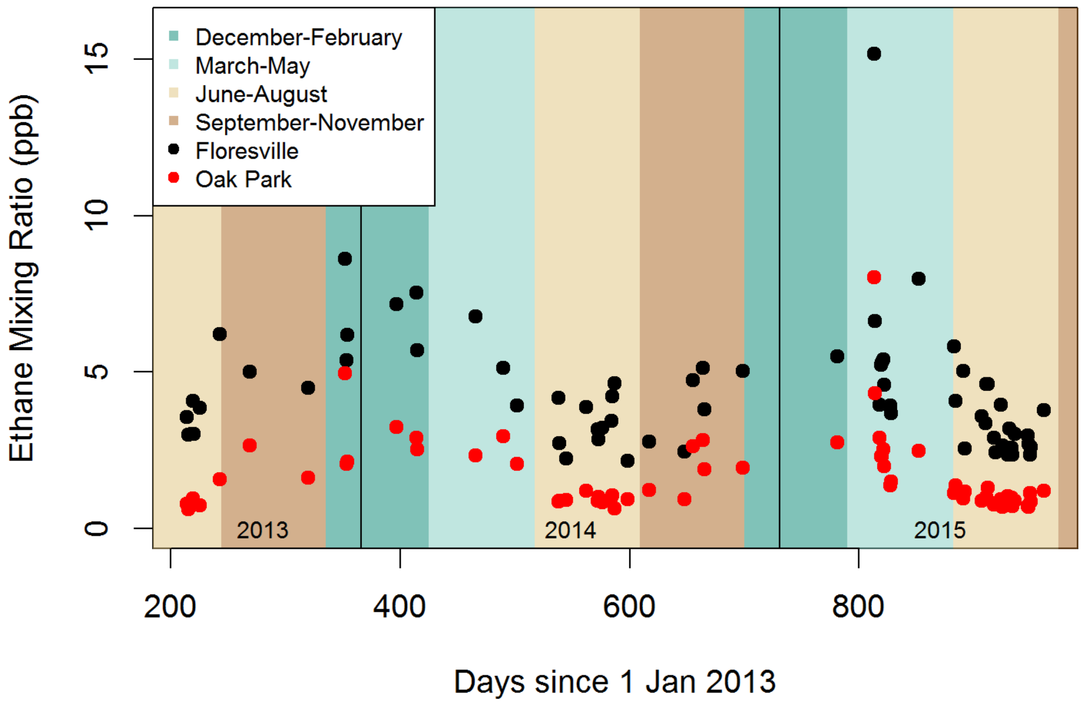

Figure S2. Scatterplot of afternoon ethane mixing ratios at the upwind site (Oak Park), the downwind site (Floresville) over 68 days with southeasterly flow. Note that each mixing ratio has a relative uncertainty of ±5.8% and the uncertainty of the enhancement is equal to the sum of the uncertainties of the upwind and downwind sites. The background colors show the warm and cool seasons. Ethane mixing ratios are generally lower at both the upwind and downwind sites during the summer and fall and higher in the winter and spring.

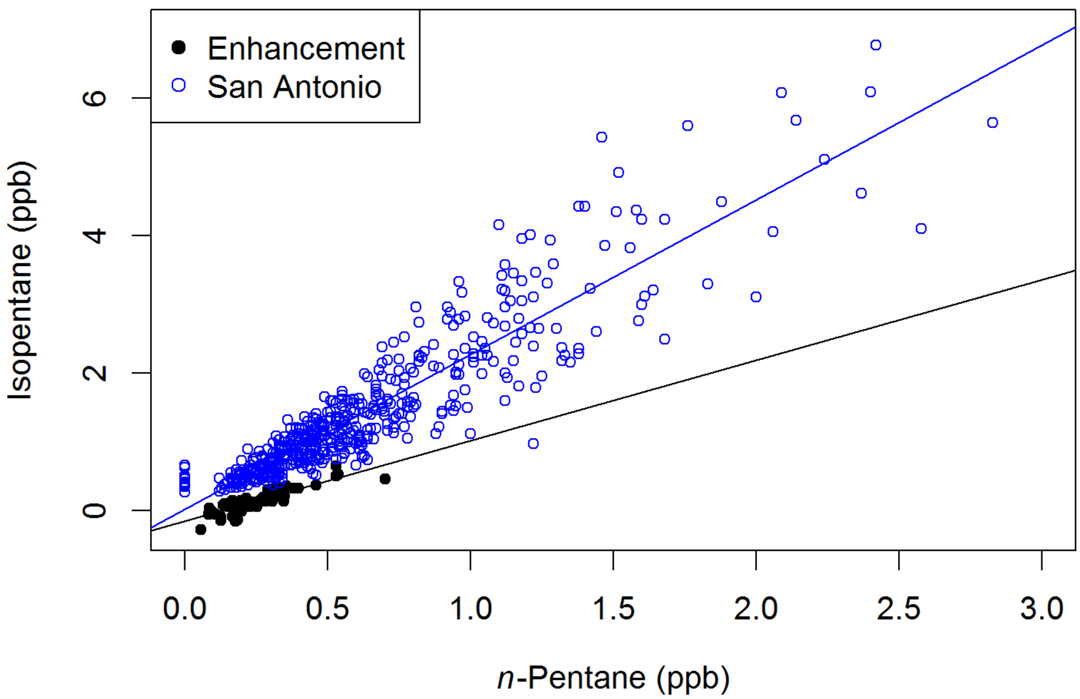

Figure S3. Scatterplot of afternoon isopentane and *n*-pentane enhancements observed between Oak Park and Floresville over 68 days with southeasterly flow. Also shown are isopentane and *n*-pentane mixing ratios in San Antonio for 24 hour canister samples from 2007 to 2015. The slopes of both linear regressions are highly statistically significant ($p < 0.001$). The isopentane-to-n-pentane ratio in San Antonio is 2.25, which is indicative of urban emissions. Meanwhile, the slope of the ratio for the enhancements is 1.17, indicating emissions from petroleum production.

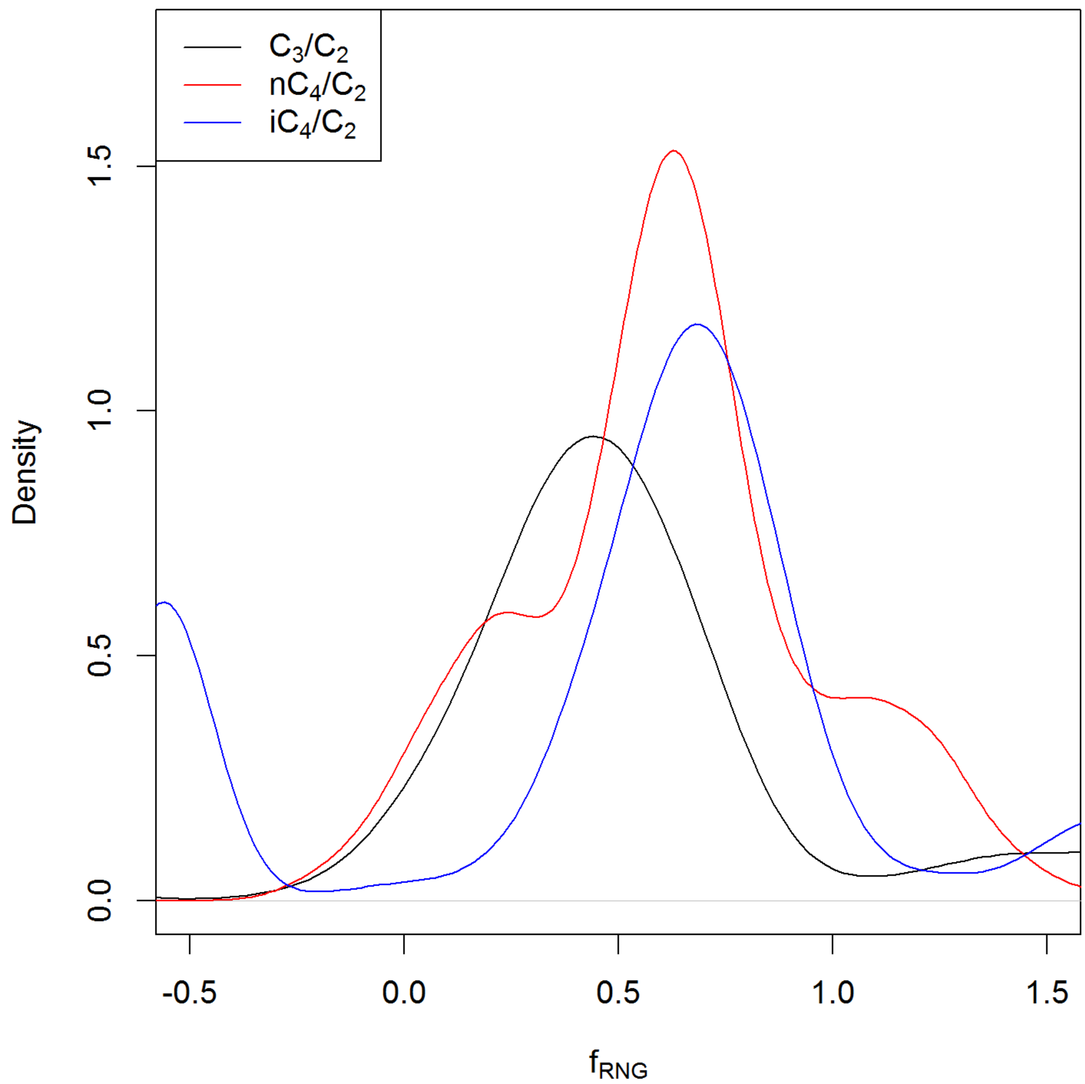

Figure S4. Probability distribution functions for the fraction of observed alkane ratios that can be explained by emissions of raw natural gas (f_{RNG}). Note that the x-axis extends to numbers that are physically unreasonable. This is due to the assumption that alkane ratios can be explained by emissions from raw natural gas and vented tank gas alone (the compositions of which are highly uncertain), and other sources and sinks are neglected. Nonetheless, there is general agreement between the fractions derived from the C_3/C_2 ratio, the nC_4/C_2 ratio, and the iC_4/C_2 ratio, suggesting that the above assumption creates results that are reproducible using these three alkane ratios.

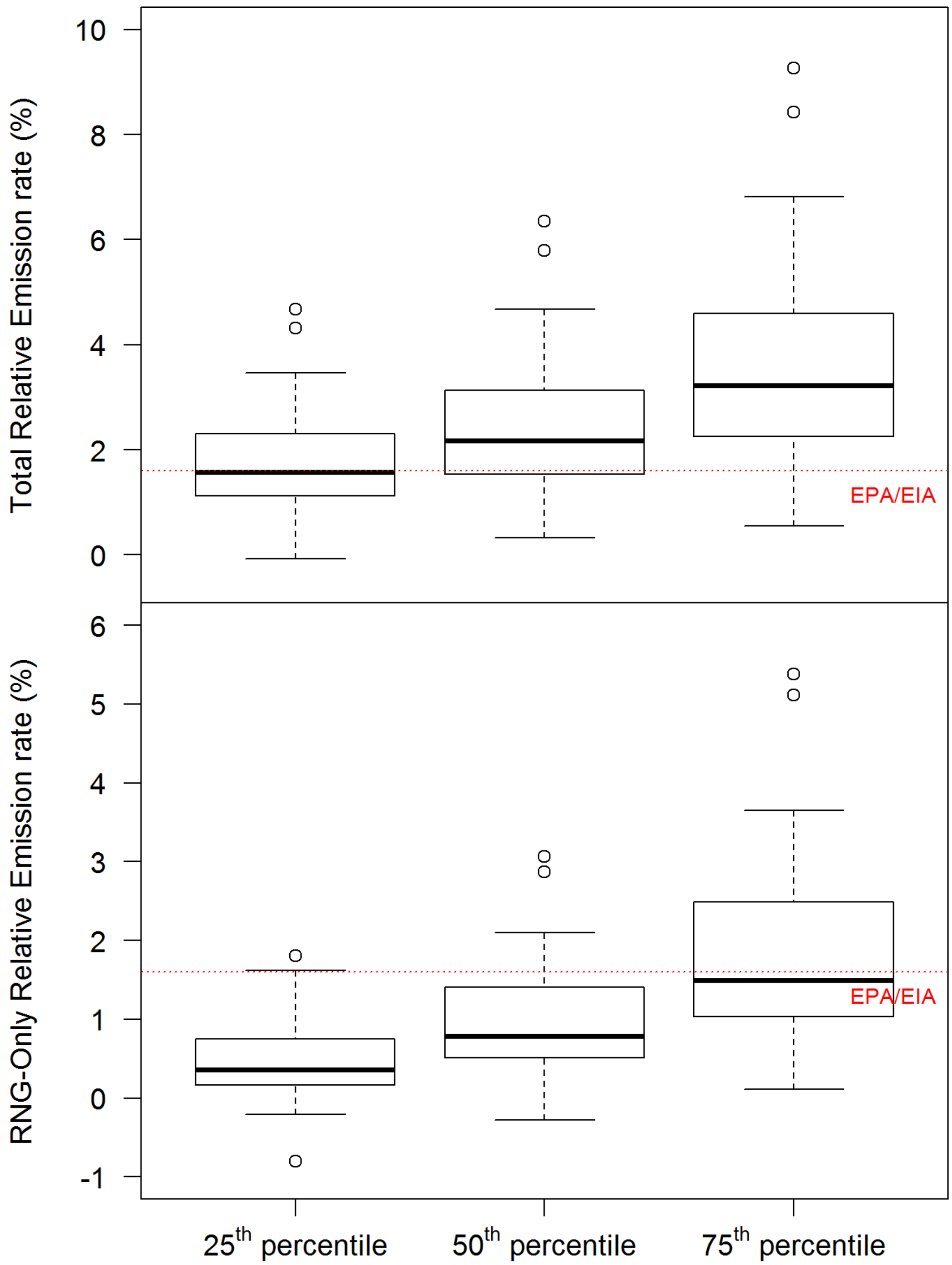

Figure S5. Distributions of the 25th, 50th, and 75th percentile of relative emissions over all 68 days. Methane emissions were converted to a volume of natural gas using methane to natural gas ratios (Table 1 in main text) and compared to natural gas production. The top panel shows total methane emissions while the bottom panel shows only emissions from raw natural gas (RNG) sources and excludes emissions from liquid storage tanks. Some outliers of the lower bound were less than zero due to large uncertainties in the methane enhancement which occasionally produced a negative methane flux within the Monte Carlo simulations. The ~~lower bound of the~~ total emission rate is often close to that of the EPA/EIA emission estimate for nationwide natural gas and associated gas production in 2011~~ while the median total emission rate estimate was nearly twice as large as the EPA/EIA~~. The emission rate for raw natural gas emissions alone was generally ~~close to~~lower than the EPA/EIA estimate. Note that the emission rate estimate has a slightly skewed distribution, with the upper bound possibly influenced by plumes from large emitters.

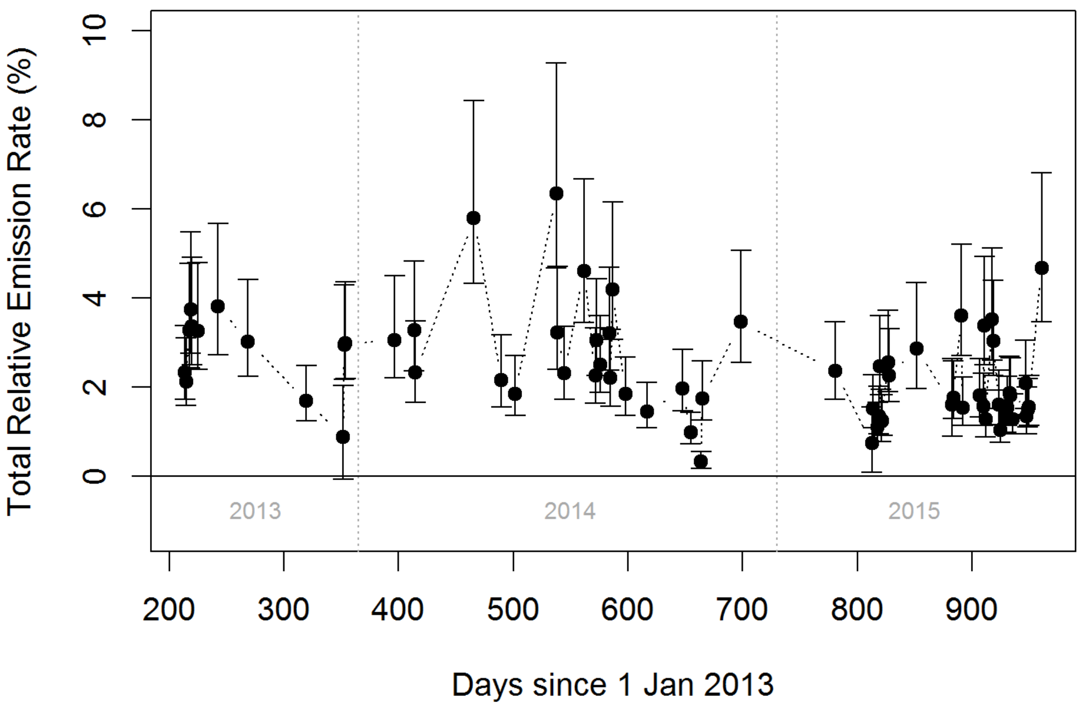

Figure S6. Timeline of the median emission rate for each day with the interquartile range represented by whiskers. The emission rate showed neither apparent seasonality nor trend over time, which is to be expected of a continuous emission source that does not depend on meteorological variables.

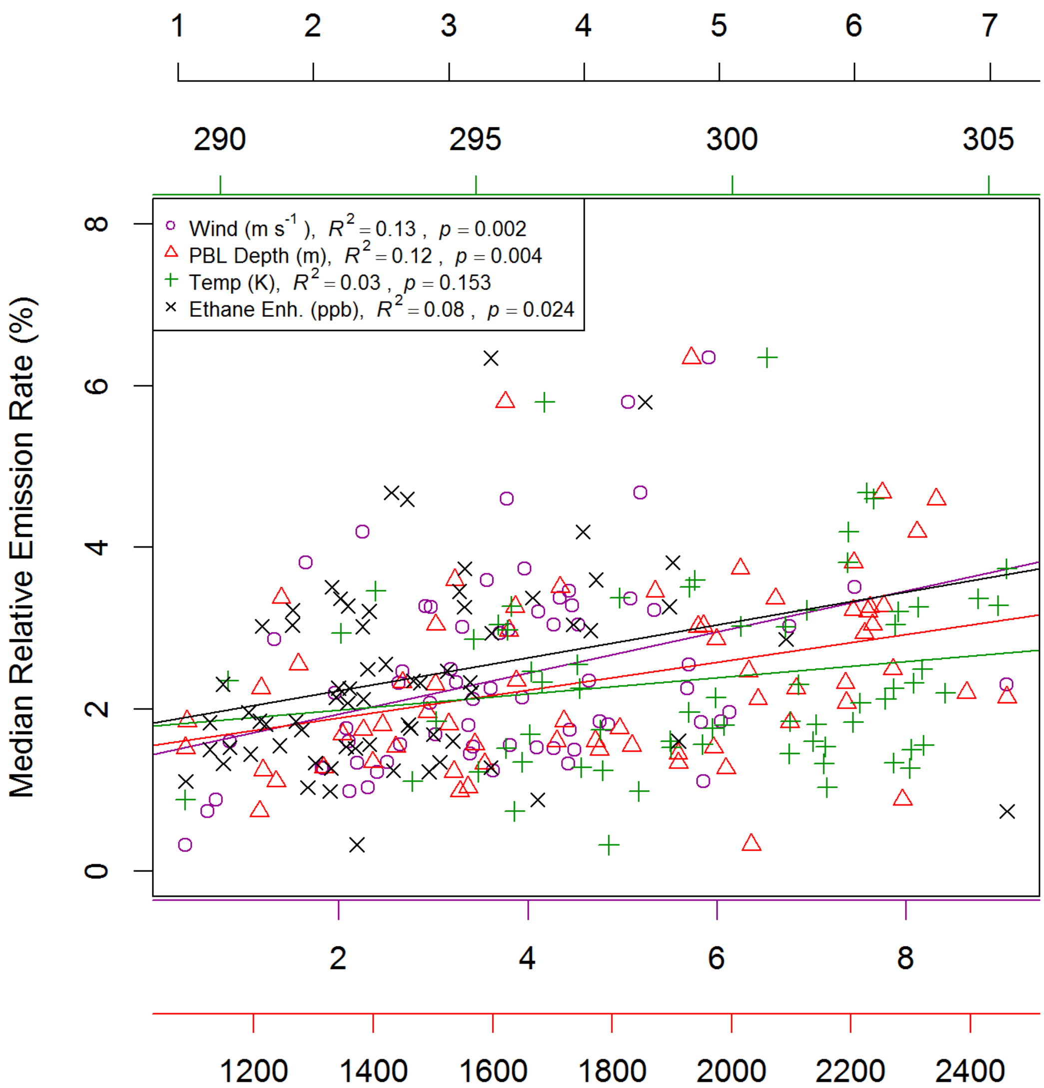

Figure S7. The median emission rate for each of the 68 days plotted against wind speed (component parallel to the transect between Corpus Christi and Floresville), PBL depth, temperature, and ethane enhancement. While there were statistically significant correlations with wind speed, temperature, and ethane enhancement, these correlations were week. While wind speed shared the strongest correlation with the emission rate ($R^2 = 0.13$), this suggests that the emission rate was not strongly driven by meteorological variability.

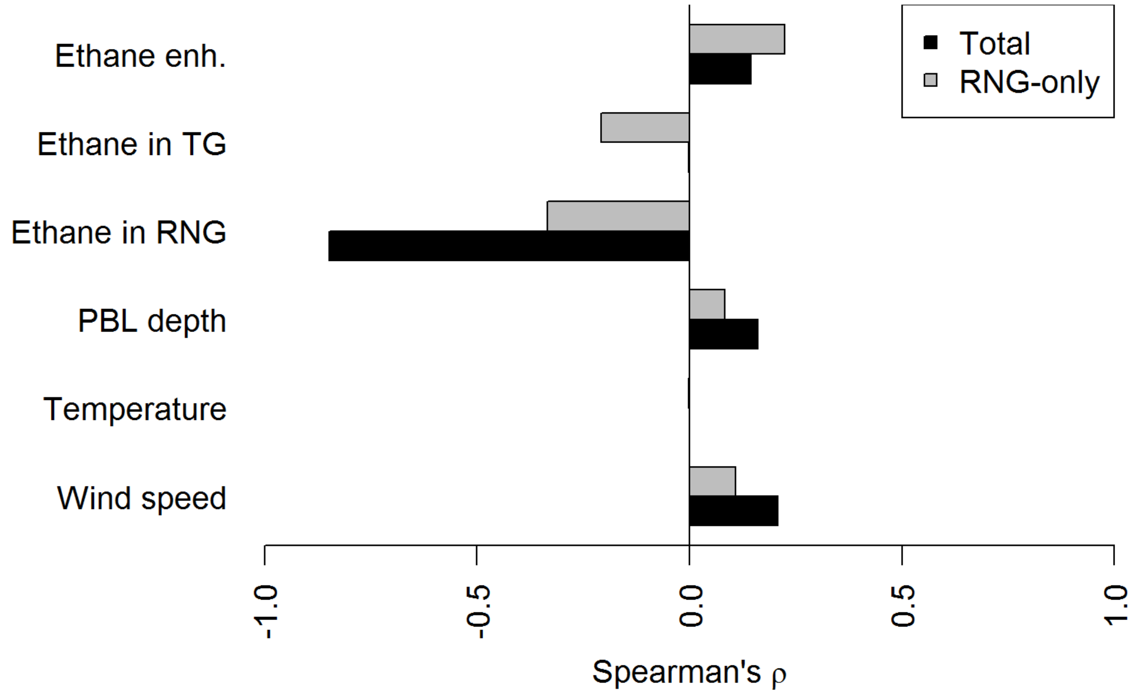

Figure S8. Example of a tornado plot from a Monte Carlo simulation for 2 August 2013. The total emission rate depended largely on the composition of raw natural gas (RNG), with lower ethane content in natural gas resulting in a higher emission rate. The RNG-only emission rate shows a negative correlation with the ethane content in both RNG and vented gas from liquid storage tanks (TG), as these numbers were used to partition emissions between RNG and TG sources. Both the total and RNG-only emission rates showed positive correlations with wind speed and planetary boundary layer (PBL) depth, but not temperature. The rates were also positively correlated with the ethane enhancement.