

RESEARCH ARTICLE

Air quality measurements in the western Eagle Ford Shale

Geoffrey S. Roest^{*,†} and Gunnar W. Schade^{*}

The Eagle Ford Shale in southern Texas remains one of the most productive oil and gas regions in the US. Like the Permian Basin and Bakken Shale, ubiquitous natural gas flaring serves as an uncertain source of trace gas emissions within the Eagle Ford. A lack of ambient air quality data, especially in the western shale, impedes a thorough understanding of trace gas emissions within the shale and the subsequent local/regional air quality impacts. Meteorological and trace gas instrumentation was deployed to Shape Ranch in southwestern Dimmit County, near the US/Mexico border, from April to November of 2015. Mixing ratios of CO, NO_x, O₃, and VOCs did not exceed ambient air quality standards and were generally lower than mixing ratios measured in US cities, with the exception of alkanes. A non-negative matrix factorization demonstrated the dominance of oil and gas-sector emission sources in local trace gas variability, with combustion processes and transport of continental air also present. An analysis of NO_x/CO and NO_x/CO₂ ratios during periods of O₃ titration, identified by the ambient NO_x/O₃ ratio, suggested that combustion and biospheric sources contributed to emissions of NO_x, CO, and CO₂. In-plume NO_x/CO₂ ratios indicated relatively low-temperature combustion sources, with median NO_x/CO₂ ratios close to that expected for natural gas flaring (0.54 ppb/ppm), and much lower than emission ratios for internal combustion engines (>10 ppb/ppm). However, the NO_x/CO₂ ratio within these plumes exhibited a large variability, spanning more than an order of magnitude. Future research should focus on improving flaring emission factors and flaring volume estimates such that their air quality impacts can be better understood.

Keywords: Petroleum; Air quality; Oil and Gas; Emissions; Flaring; Eagle Ford Shale

1. Introduction

The Eagle Ford Shale (EFS) in south-central Texas has become one of the most productive shale plays in the United States – second in oil production and fifth in natural gas production among US shale plays in 2017 (U.S. Energy Information Administration, 2018a, 2018b). Like other shale oil and gas plays and conventional fossil fuel sources, production of oil and gas in the EFS has potential impacts on both the climate and local/regional air quality (Pacsi et al., 2015; Parrish et al., 2017; Peischl et al., 2018; Roest and Schade, 2017; Schade and Roest, 2015, 2016, 2018). Emissions of criteria air pollutants (CAPs), hazardous air pollutants (HAPs), and other pollutant precursors have been studied in active oil and gas producing regions across the US. Elevated mixing ratios of volatile organic compounds (VOCs) – important ozone (O₃) and particulate matter (PM) precursors – have been found in several active oil and gas producing regions, including the

Colorado Front Range (Gilman et al., 2013), the Marcellus Shale (Swarthout et al., 2015), the Permian Basin (Koss et al., 2017), and the EFS (Schade and Roest, 2018, 2016), among others. Nitrogen oxides (NO_x) – also important O₃ and PM precursors and respiratory irritants (Chauhan et al., 1998) – have also been observed in high abundance in several regions, such as the Bakken Shale in North Dakota (Prenni et al., 2016).

The proximity of the EFS to the city of San Antonio has raised air quality concerns as the city's O₃ design values – the three-year average of the fourth-highest annual daily maximum 8 hour average mixing ratio (U.S. Code of Federal Regulations, 2015) – have been above the EPA's 8-hour O₃ National Ambient Air Quality Standard (NAAQS) for several years. However, due to the large uncertainties in emissions from oil and gas and the sparsity of ambient air quality measurements, knowledge of the local and regional air quality impacts of the EFS remains limited. Furthermore, as the EFS rapidly developed in rural Texas, a limited pipeline network led to widespread flaring to destroy gas that would otherwise be emitted into the atmosphere. Flaring is used principally to destroy excess natural gas produced at oil wells, called “associated gas”.

* Department of Atmospheric Sciences, Texas A&M University, US

† School of Informatics, Computing, and Cyber Systems, Northern Arizona University, US

Corresponding author: Geoffrey S. Roest (geoffrey.roest@nau.edu)

The EFS, along with the Bakken Shale and the Permian Basin, has been identified in satellite imagery as a region with a high density of observed flares (Elvidge et al., 2015, Fig. S1) and coincident increases in NO_2 abundances have been detected using satellite data (Majid et al., 2017). This increased use of flaring introduces a particularly indeterminate emission source since flaring volumes are under-reported in Texas (Willyard and Schade, 2019) and flaring emission factors are highly uncertain (Stroscher, 2000; Torres et al., 2012; U.S. Environmental Protection Agency, 2016). Because flaring may have significant air quality impacts on a local- to regional-scale (Olague, 2012), flaring emissions need to be properly quantified before overall air quality impacts can be accurately assessed.

To better understand the air quality impacts of unconventional oil and gas development, increased monitoring and data analyses are needed to identify and characterize emission sources. Here, we present an assessment of ambient air quality measurements on Shape Ranch – a working bison ranch in southwestern Dimmit County in the western EFS. This site, which is far removed from urban emissions, may serve to represent ambient air quality in the more rural regions of the EFS where no other ambient air quality data exist. This assessment provides information about the exposure to pollution of rural communities in the EFS, and the air mass composition that may be advected into major urban areas, such as San Antonio. Emission source types were identified using a non-negative matrix factorization, and NO_x/CO_2 ratios were used to identify and characterize low temperature combustion plumes.

2. Measurements

2.1. Site location and environment

An air-conditioned trailer was placed on Shape Ranch – a large, working bison ranch near the southwestern corner of Dimmit County in southern Texas. The trailer was parked south of an olive tree grove adjacent to a house construction site toward the northeast. This location provided access to reliable line power and would provide the homeowners with information about their exposure to pollution near their new home. Some croplands are located within the region, especially to the northeast of the ranch near the cities of Carrizo Springs and Crystal City. These are also the nearest sizable urban centers (population less than 8,000 each), while other urban areas exist along the Rio Grande to the west and northwest of the site, especially near Eagle Pass, Texas. The city of Laredo (population ~250,000) is located 100 km south-southeast along the Rio Grande. The local and regional land cover consists dominantly of tropical or subtropical grass- and shrublands.

Mesquite trees dominated local vegetation, with black brush, live oak, date palms, twisted acacia, brasil bush, guajillo bush, and cacti also present. Among these species, only live oak and date palms were known isoprene emitters (Benjamin et al., 1996). More than seventy small olive trees, which are known to emit monoterpenes (Benjamin et al., 1996) were growing immediately north of the trailer. The nearest major road – Highway 186, known as Faith Ranch Road – was approximately 7 km east of the trailer,

running generally north to south with some east-west segments. This road serves as a main thoroughfare for local ranchers as well as commercial oil and gas industry vehicles. The nearest active drill site – 6.3 km east-southeast of the trailer – was located along a private dirt road which separates Shape Ranch on the north side from San Pedro Ranch to the south. A network of private dirt roads lay between Highway 186 and the trailer location. Several legacy oil and gas well pads exist around the trailer within a 2 km radius.

2.2. Instrumentation and data

The trailer was located within 100 m of the building site of the ranch owners' new home. An air conditioner on the trailer's roof cooled the inside, which housed all trace gas instrumentation, a Campbell Scientific CR1000 data logger, and a Dell Precision T3400 PC. Trace gas instrumentation included a NO/NO_x analyzer, a Gas-Filter-Correlation (GFC) CO analyzer, an NDIR $\text{CO}_2/\text{H}_2\text{O}$ analyzer, a UV-absorption O_3 analyzer, and a dual-channel gas chromatograph (GC) with FIDs (SRI model 8610). Specifications for these instruments are shown in Tables S1 and S2. Meteorological instrumentation from Campbell Scientific included a temperature and relative humidity sensor, a tri-cup anemometer, a barometer, a pyranometer, and a tipping-bucket rain gauge, all installed on a beam 3 m agl extending from the rear of the trailer toward the east. Air sampling was performed at the end of the beam through a single $\frac{1}{4}$ " ID PFA Teflon line with PTFE inlet filter using the air quality monitors' internal pumps and splitting gas paths inside the trailer while reducing PFA tubing diameters. Meteorology and trace gas mixing ratios, except for the hydrocarbon measurements, were recorded by the CR1000 data logger as 10 s averages.

The NO_x analyzer was set up to output mixing ratios expressed as a 20 s moving average. Zero and span calibrations (104 ppb NO in N_2 , prepared in July 2010 by Scott-Marrin, CA) were performed during each site visit – approximately once every ten days. Interference from nitric acid (HNO_3) was avoided via using a nylon membrane as the instrument's inlet filter. For data analyses using the 10 s averages, the NO_x data were shifted forward by 50 s to account for the instrument's response time relative to the O_3 and CO_2 measurements. The CO analyzer was also set up to record 20 s moving averages and it as well was zeroed and calibrated during each site visit (547 ppb CO in zero air, prepared in July 2011 by Scott-Marrin, CA). However, the measurements from this instrument are more temperature sensitive, and thus, automated internal zeroing was performed once every 6 h to counteract electronic drift. The O_3 analyzer was operated with a recent factory calibration and was regularly zero checked using an O_3 scrubber. CO_2 was measured using a single channel NDIR instrument (LI-840, Licor Inc.) calibrated using a 454 ppm CO_2 standard (prepared in August 2014 by Scott-Marrin, CA), while water vapor measurements relied on a factory calibration.

The gas chromatograph was set up to perform half-hourly hydrocarbon measurements to maximize temporal resolution, at the expense of measuring more compounds. Samples were collected at 100 mL min^{-1} onto

two-stage activated carbon adsorption micro-traps, one containing Carboxen B with Carbotrap X backup for C_5 and higher hydrocarbons, and one containing Carbotrap X and Carboxen 1000 to trap short-chain hydrocarbons. The former sample, injected into a 60 m, wide-bore MTX-624 column, was collected directly from ambient air before injection. The latter sample, injected into a 30 m, wide-bore PLOT-Alumina column, was first dried through a Nafion dryer (except for a period from 7–28 August when a Drierite desiccant tower was used instead); the flow then passed through an Ascarite II CO_2 scrubber with additional potassium perchlorate to remove any leftover moisture. Both samples were thermally desorbed directly into two wide-bore 5 m retention gaps ahead of the chromatographic columns, similar to Park et al. (2010, Table S2). Sampling time was twenty minutes for both samples, from 8 to 28 min into each 30 min period. Thermal desorption at 200°C occurred for 2 min at the beginning of each run, after which the traps cooled down with a fan to approximately room temperature before the next sample period. A ppm-level internal standard of 2,2-dimethylbutane (prepared in March 2015 by Scott-Marrin, CA) was diluted into the main sampling line to quantify observed hydrocarbon peaks in the chromatograms. Other standard gases, including mixes of normal alkanes, branched alkanes, alkenes, methyl ethyl ketone (MEK), benzene, toluene, ethylbenzene, and xylenes (collectively referred to as BTEX) were injected into the main sampling line inlet using a syringe during selected site visits. An automated zero-air run was performed once every 15 hours to monitor for contamination. A mixture of hydrogen gas and zero air produced onsite using hydrogen and zero air generators was used for the FIDs, as well as carrier gas (H_2). The chromatography data were re-analyzed offline using SRI PeakSimple Software (SRI Instruments, Version 3.88, 2010). Additional information regarding the quantification of VOCs is available in the supplement (Text S1).

The non-hydrocarbon trace gas and meteorological data were reduced to 30 min resolution to align with the hydrocarbon measurements. This was performed by finding the median measurement for each variable in 15 min periods and then averaging the medians for the two 15 min periods to create each 30 min data point. The exceptions were rainfall, which was totaled for each 30 min period, and wind speed and direction, for which a resultant mean (vector averaging) was used. Medians were chosen to remove artificial instantaneous spikes in the measurements. Once 30 min data were obtained, regular invalid measurements (e.g. automated zeroes) were flagged for removal and other invalid measurements (e.g. contamination, bad readings related to power issues, irregular zeroes and calibrations) were flagged and removed manually. For certain analyses discussed below, the original 10 s resolution data were used. In these instances, quality control was performed for the relevant subset of data.

3. Analysis methods

3.1. Source identification using non-negative matrix factorization

A number of emission source types were expected to contribute to observed trace gas variability on Shape Ranch.

In addition to emissions from local and regional oil and gas activities, possible emission sources included vegetation, soils, livestock and agriculture, distant urban sources, biomass burning, and local and regional automotive traffic. However, it should be noted that most of regional vehicle emissions came from traffic associated with the oil and gas industry, as traffic counts on the nearby FM 186 increased by more than a factor of 10 as the shale was developed between 2010 and 2015 (Texas Department of Transportation, 2018). Some emissions sources, including, e.g., biogenic emissions and evaporative hydrocarbon emissions, are strongly dependent on temperature (Guenther et al., 2012; Rubin et al., 2006), with biogenic and soil emissions also depending on soil moisture availability (Guenther et al., 2012; Vinken et al., 2014). Varying wind directions brought different air masses towards Shape Ranch: air over the Gulf of Mexico to the southeast is relatively clean (Gilman et al., 2009), while continental air masses, with emissions from distant urban centers such as San Antonio, are located to the east and to the north; westerly winds were very rare during the field campaign, as discussed in section 4.1.1. Emissions from biomass burning were assumed to have a negligible impact on trace gas variability due to the lack of detected fires during the field campaign (Andela et al., 2019).

Several methods are available to identify source types contributing to observed trace gas variability. Traditional matrix factorization techniques have been employed, such as singular value decomposition (SVD) and positive matrix factorization (PMF). The latter, which was originally developed by Paatero and Tapper (1994), is a preferred source apportionment method often used by the US EPA (Reff et al., 2007). Several new PMF variations have been developed and are referred to as non-negative matrix factorization (NMF), a name that was popularized by Lee and Seung (1999). NMF models have been developed for a variety of applications, including image decomposition (Lee and Seung, 1999), metagenomic analysis (Brunet et al., 2004), and atmospheric pollutant source attribution (e.g. Thiem et al., 2012). The NMF models work as follows:

1. A number of factors is prescribed by the user. This choice must be made carefully to allow for a sufficiently informative number of factors, though too many factors will cause results that cannot be physically explained, explain a negligible fraction of the total variability in the dataset, or are unstable.
2. Two matrices are initialized: the first is the contribution of individual variables (trace gases in the case of Shape Ranch) within each factor (\mathbf{W}), and the second is the expression of each factor during every time step in the database (\mathbf{H}), such that the matrix multiplication of \mathbf{W} and \mathbf{H} are an approximation of the observed dataset $\mathbf{V} \approx \mathbf{WH} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon}$ is the residual error. The initialization of \mathbf{W} and \mathbf{H} may be based on other matrix factorization methods, such as non-negative SVD; it may also be randomly generated based on the observed distributions of variables; or it may be prescribed by the user. For random initializations, the NMF model should be run several times to identify the best solution based

- on the large number of possible starting points.
- The initial guess is improved upon to reduce the error between the observed dataset and the modeled dataset – i.e., the matrix multiplication of the variables within each factor by the expression of each factor over time. There are a number of algorithms to reduce this error.

The least-squares NMF (LS-NMF) algorithm (Wang et al., 2006) was chosen because this method allows for the resulting factors to be weighted by uncertainties in the observed dataset, such that highly-uncertain variables have less of an influence on the solution than highly certain variables.

The NMF package (Gaujoux and Seoighe, 2010) in the R programming language (R Core Team, 2018) was used to perform a source apportionment of the Shape Ranch dataset. This package includes several initialization methods and error-reducing algorithms. Both the SVD and random-seed initializations were tested with the LS-NMF algorithm. The SVD has the advantage of performing a high-quality initial guess, resulting in relatively small errors between the observed and modeled datasets and a low computational cost. Meanwhile, the random-seed initialization, while relatively slow due to the large number of runs needed, may result in a lower residual error if one of the random initializations converges towards a true solution.

Data were prepared in a manner following Guha et al. (2015). Observations that were missing hydrocarbon measurements – e.g., during zero runs – were removed. Ethylbenzenes and xylenes were summed for the NMF due to periodic misidentifications associated with the near-coelution of ethylbenzene and m/p-xylene. The limit of detection (LOD) for each trace gas species was assumed to be the minimum non-missing measurement in the timeline. Missing CO and CO₂ values were linearly interpolated. There were no missing O₃ data during the NMF analysis period. VOC and NO_x mixing ratios below

the LOD were replaced with random values between zero and 0.5 × LOD. The top 1% of mixing ratios for each species were considered to be extreme. These values were scaled downward to equal the 99th percentile of the distribution of mixing ratios in an NMF run to limit the influence of episodic plumes on the overall source attribution. The background mixing ratio for each species, assumed to be the minimum mixing ratio in the timeline, was subtracted to remove the background but a value of e⁻⁵ was added back in to avoid mixing ratios of zero in the NMF model. Uncertainties were quantified based on instrument precision and accuracy and mixing ratio, as shown in **Table 1**, with absolute uncertainties converted to relative uncertainties using the mixing ratio time series for the associated species. Lastly, the mixing ratios for each species were normalized such that the maximum mixing ratio was equal to one for each species.

The US EPA has developed a VOC emission factor database for anthropogenic sources known as SPECIATE (U.S. Environmental Protection Agency, 2015a). The source factors from the NMF model were compared to the SPECIATE sources to identify source categories that may be associated with the NMF factors. The comparison was performed by calculating the dot product of normalized emission factors, following eq. 5 from Kota et al. (2014):

$$\theta = \frac{\sum_{i=1}^n f_i s_i}{\sqrt{\sum_{i=1}^n f_i^2 \sum_{i=1}^n s_i^2}}$$

where f_i and s_i are the NMF and SPECIATE scores for VOC i for an arbitrary pairing of emission factors from the NMF model output and SPECIATE database. A value of $\theta = 1$ implies a perfect match between the NMF and SPECIATE factors whereas $\theta = 0$ would result from factors that have no overlapping species.

Table 1: Assumed uncertainties for the LS-NMF dataset. DOI: <https://doi.org/10.1525/elementa.414.t1>

Species	Precision	Accuracy	Measurement uncertainty ^(a)
CO	±20 ppb ^(b)	±2% ^(c)	[CO] × 2% + 20 ppb
CO ₂	±1 ppm ^(b)	±1.5% ^(b)	[CO ₂] × 1.5% + 1 ppm
O ₃	±1.5 ppb ^(b)	±1.5 ppb ^(b)	3 ppb
NO _x ^(d)	±0.2 ppb ^(b)	±5% ^(c)	[NO _x] × 5% + 0.2 ppb
VOCs	±1.8% ^(e)	±8% ^(f)	2 × LOD for mixing ratios < LOD ^(g) $[(8\% \times [\text{VOC}])^2 + \text{LOD}^2]^{0.5(f,g)}$

^(a) Total uncertainty is assumed to be the sum of precision and accuracy, except for VOCs.

^(b) From instrument operation manual.

^(c) Accuracy of standard gas used for calibration.

^(d) NO_x was not included in the NMF analysis due to an instrument failure during a period of complete VOC data. Nonetheless, the uncertainty in the NO_x data is quantified in this table.

^(e) Variability (2σ) of the internal standard measurements.

^(f) Conservative estimate of 5% accuracy for each of two flow meters and 2% accuracy for standard gas concentration, added in quadrature.

^(g) From Guha et al. (2015).

3.2. Plume identification using trace gas ratios

3.2.1. General plume identification

Plumes from combustion sources were identified within the 30 min dataset using a combination of NO_x and O_3 mixing ratios. The background levels for NO_x , CO , and CO_2 were assumed to be moving 10th percentiles over a moving 12 h window. The ratio of NO_x/O_3 serves as a sensitive indicator of combustion plumes, particularly at night, as O_3 is rapidly titrated by NO emissions from combustion sources to form NO_2 . During the nighttime, NO_2 is not photolyzed and the NO_x/O_3 ratio remains elevated as the plume is transported downwind. However, during the daytime, NO_2 is photolyzed and O_3 increases again further downwind from the emission source. Therefore, this method does not capture plumes from distant combustion sources during the daytime. It may also fail to capture plumes that occur in a high background O_3 environment, as the NO_x/O_3 ratio may remain low even within a plume. NO_x/O_3 thresholds of 0.25 and 1 were both used to test the sensitivity of plume detections to the selected NO_x/O_3 ratio. Mixing ratios of selected VOCs within plumes were assessed for their relationships with one another and with CO_2 and NO_x . Alkanes and benzene were chosen due to their presence in emissions from combustion sources and their persistent signals above the LOD. Trace gas ratios within plumes were measured as the slope of a standard major axis (SMA) regression from – also known as a reduced major axis regression – which has been identified as a useful regression method for air quality data as the uncertainty in both variables is accounted for (Ayers, 2001).

3.2.2. In-plume ratios of carbon dioxide and nitrogen oxides

The high temporal resolution of the original 10 s data provided more insight into the plumes identified using the NO_x/O_3 ratio. However, due to the poor precision of the 10 s CO data, the NO_x/CO ratio was excessively noisy. Therefore, only the NO_x/CO_2 ratio was analyzed in more detail. A NO_x/O_3 ratio of 0.25 was used to filter 30 min data. Consecutive 30 min periods with NO_x/O_3 greater than 0.25 were assumed to be one continuous combustion plume event. For each of these events, 30 min was added before and after the time period to include pre- and post-plume observations. This expanded time period was then used to analyze the corresponding 10 s data. The moving 10th percentile from the 30 min dataset was linearly interpolated to a 10 s resolution to represent the background mixing ratios, which were then subtracted from the 10 s time series to obtain trace gas enhancements. For each plume event, the 10 s NO_x/CO_2 enhancement ratios were assessed using the slope of an SMA regression when ambient NO_x/O_3 exceeded 0.25.

4. Results

A brief discussion of air quality and meteorological observations is presented, followed by the results of the source attribution using non-negative matrix factorization. Next, combustion plumes are identified and characterized using trace gas correlations.

4.1. Ambient air quality at Shape Ranch

4.1.1. Meteorology

The observed meteorology at Shape Ranch in 2015 is compared to climatological means (1981–2010) at Carrizo Springs (Arguez et al., 2010) in Table S3. Temperatures and precipitation were near average from April through June. July and August, however, were hot and dry, with average daily maximum temperatures over 38°C and less than 1 mm of rain in July, followed by no rain in August. Below-average rain was observed in September and near-normal rain amounts in October and November. These months continued to be warmer than average, with daily mean temperatures and daily minimum temperatures exceeding the respective average temperatures by more than 3°C in both October and November. Winds were predominantly southeasterly – especially in the summer, while occasional cold fronts in the spring and fall brought northwesterly winds (Fig. S2). Wind speeds reached a daily minimum prior to sunrise before increasing during the day and peaking in the late evening and early overnight hours (Fig. S3). However, the late evening and overnight hours also had a relatively high frequency of calm observations, suggesting that the late-evening peak in wind speeds is not always present. While southeasterly winds remained prevalent throughout the diurnal cycle, easterly winds were frequently observed during the afternoon and evening hours.

4.1.2. Carbon dioxide

Figure S4 shows the diurnal cycle of CO_2 mixing ratios at Shape Ranch, with low CO_2 during the daytime and elevated CO_2 during the night. Several measurements during the early morning hours exceeded 500 ppm, with peak CO_2 mixing ratios reaching 567 ppm. Most of these elevated CO_2 mixing ratios measured at Shape Ranch occurred within two weeks of a precipitation pulse (3+ mm per hour, Hao et al., 2010), suggesting that high CO_2 events were generally associated with nighttime emissions due to enhanced ecosystem respiration as a result of increased soil moisture from recent rainfall, in addition to other local CO_2 sources (e.g. combustion, Fig. S5).

4.1.3. Carbon monoxide

Carbon monoxide levels at Shape Ranch were low, with a median mixing ratio of 97 ppb. This mixing ratio is lower than past observations in urban (Fig. S6, Baker et al., 2008) and at remote continental locations (Chin et al., 1994). Low CO mixing ratios compared to other remote areas are sensible considering the observed decreasing trend in CO in recent decades (Worden et al., 2013), since the measurements in Chin et al. (1994) were made in the late 1980's and early 1990's. The diurnal cycle of CO showed higher medians during the daytime as compared to overnight. However, the highest observations (above 200 ppb) often occurred during the late evening and overnight hours. These observations were likely associated with combustion plumes in a stable nighttime boundary layer.

4.1.4. Nitrogen oxides

Similar to CO, NO_x levels were also low at Shape Ranch, with median mixing ratios of 0.1 ppb for NO and 1.7 ppb for NO_x, compared to a median NO_x mixing ratio of 2.0 ppb in Karnes City (AQS code 482551070, Texas Commission on Environmental Quality, 2020) in the center of the EFS (Fig. S6). During the nighttime, total NO_x was elevated while NO remained low as any emissions will react with O₃ to produce NO₂ (Fig. S7). Increased NO and NO_x mixing ratios during the early morning hours may represent emissions from local traffic. While background NO remained low through the evening, several outliers above 4 ppb with concurrently elevated NO/NO_x ratios serve as evidence of NO plumes from nearby combustion sources.

Additionally, several high nocturnal NO_x observations, including the highest NO observations, were made shortly after minor rainfall events (3 mm), suggesting that some high-NO_x events were associated with higher local NO emissions from wetted soil (Williams et al., 1992).

4.1.5. Volatile organic compounds

Several sources impact ambient VOC mixing ratios at Shape Ranch – notably evaporative emissions of hydrocarbons and emissions from combustion sources. **Figure 1** shows the log-normal distributions of several hydrocarbons compared to past median observations in 28 US cities (Baker et al., 2008). The median mixing ratios for propane, butanes, pentanes, n-hexane, n-heptane, and

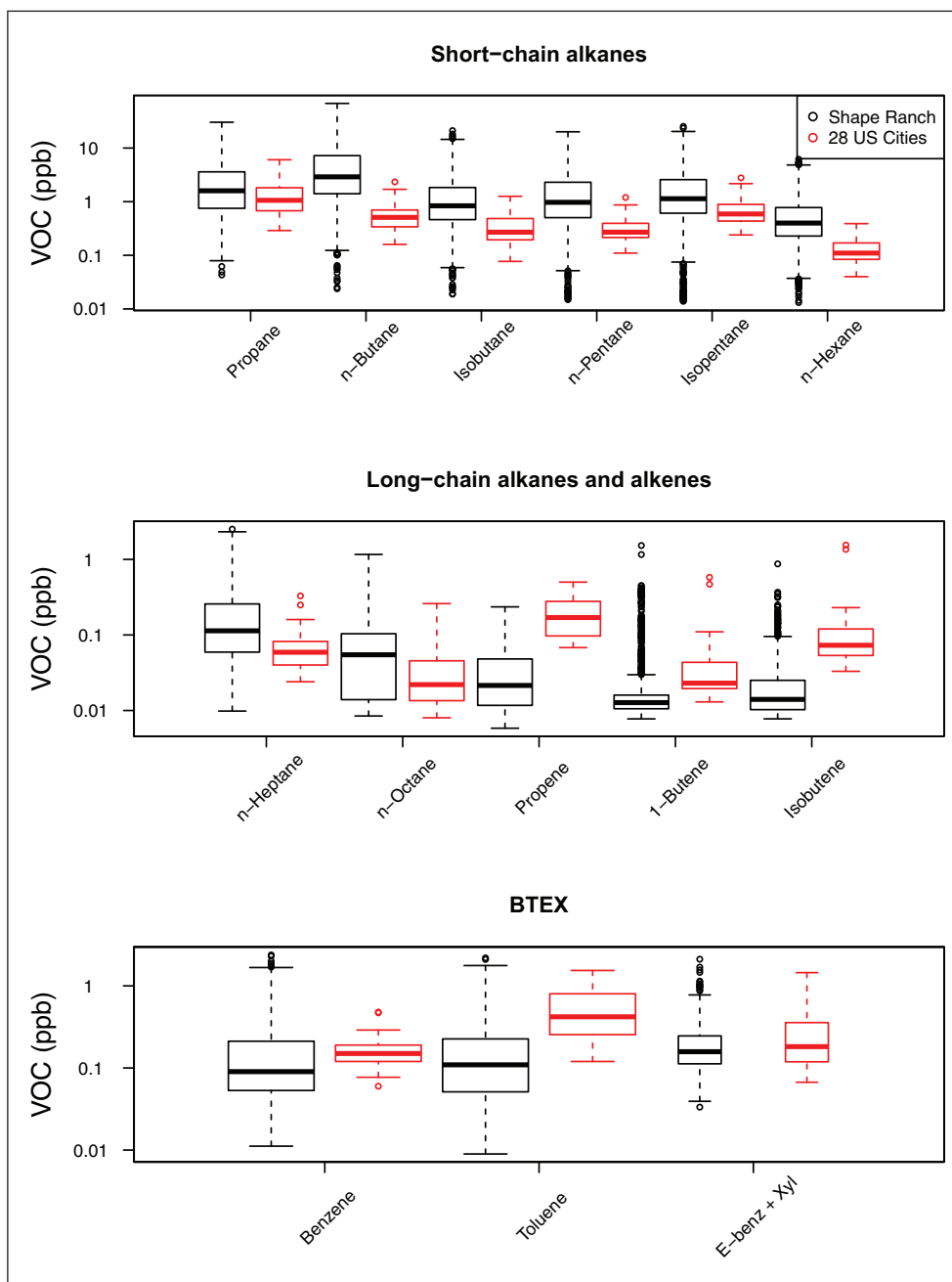


Figure 1: Comparison of VOCs at Shape Ranch and US cities. A boxplot comparison of VOC mixing ratios at Shape Ranch vs 28 US cities (Baker et al., 2008). Note that the y-axis is each panel is log-normal. Alkanes trended higher at Shape Ranch than in the 28 cities study while alkenes and BTEX trended lower. DOI: <https://doi.org/10.1525/elementa.414.f1>

n-octane were all higher than the median of the 28 US cities study. The ratio of isopentane to n-pentane of 1.14 (95% CI of 1.13 – 1.15, $R = 0.98$) was close to that observed for the pentane enhancement over the central EFS (Roest and Schade, 2017) and is much lower than in most urban areas. The high alkane levels and the isopentane to pentane ratio suggest that evaporative emissions from oil and gas industry sources dominate the observed alkane mixing ratios (Gilman et al., 2013). Mixing ratios of alkenes and BTEX species at Shape Ranch were generally lower than in the 28 cities (Figure 1), which was expected since internal combustion sources contribute to higher mixing ratios of these species in urban areas. The benzene to toluene ratio in the 28 cities study was 0.22 (0.17 – 0.29, $R = 0.72$, Figure 2) which is indicative of fresh automotive emissions combined with evaporative emissions of toluene (Baker et al., 2008). The ratio of 1.17 (1.15 – 1.19, $R = 0.93$) at Shape Ranch serves as evidence of a higher benzene to toluene ratio from local emission sources. This is confirmed by the retention of this ratio during high VOC episodes, which often result from an accumulation of local emissions in the boundary layer. However, there are observations of benzene, toluene, and o-xylene at Shape Ranch that are higher than all the median observations in the 28 cities study. Figure S8 shows that the alkanes and BTEX species were elevated at night, so high mixing ratios were likely associated with accumulation of local and regional emissions in a shallower nocturnal boundary layer. However, propene and isoprene were both elevated during the day. Isoprene is known to be emitted from certain green plants during the day, and the higher propene

during the day may be indicative of similar, local emissions during daytime (Goldstein et al., 1996).

4.1.6. Ozone

During the entire field campaign, the maximum daily 8-hour average O_3 mixing ratio exceeded the current EPA standard (U.S. Environmental Protection Agency, 2015b) on only one day (1 August 2015). The fourth-highest maximum daily 8-h average O_3 mixing ratio – the three-year average of which is used to compare against the EPA's standard (U.S. Code of Federal Regulations, 2015) – was 65 ppb, 5 ppb below the standard. The maximum daily 8-h average was generally higher than the nearby city of Laredo (Fig. S6, AQS code 484790016, Texas Commission on Environmental Quality, 2020), possibly due to O_3 titration in the urban air mass. The O_3 mixing ratios at Shape Ranch followed a typical diurnal pattern with high mixing ratios during the daytime, followed by decreasing mixing ratios throughout the night, reaching a minimum in the early morning hours as O_3 is titrated by local NO emissions.

4.2. Emission sources impacting Shape Ranch

4.2.1. Non-negative matrix factorization

A source attribution was performed for the Shape Ranch dataset for a period from 13:00 LST 4 September to 9:30 LST 5 October 2015, when most trace gas instruments functioned without interruption except for calibrations. However, due to an instrument failure, NO_x data are missing after 15 September and have therefore not been included in the NMF source attribution. During this time, temperatures were characteristically warm with only one

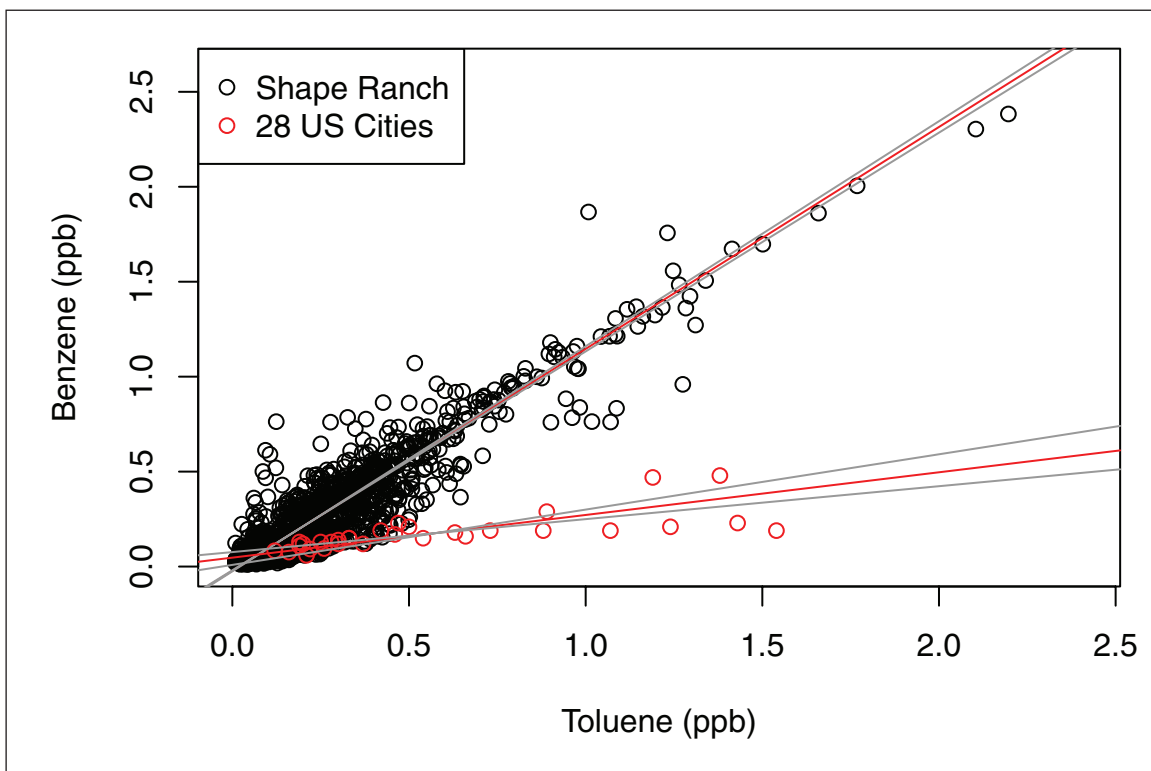


Figure 2: Benzene and toluene ratios at Shape Ranch and US cities. Benzene vs toluene at Shape Ranch and in 28 US cities study (Baker et al., 2008). The ratio at Shape Ranch was 1.17 (1.15 – 1.19, $R = 0.93$) while the ratio in the 28 cities was 0.22 (0.17 – 0.29, $R = 0.72$). DOI: <https://doi.org/10.1525/elementa.414.f2>

day of observed precipitation, and winds were generally from the SE to ESE.

Several LS-NMF runs were performed with varying numbers of factors and seeding methods. The performances of LS-NMF runs with two to seven factors were first evaluated against LS-NMF runs with randomized data. Runs with three, four, and five factors performed better as compared to two, six, or seven factor runs. The three-factor run was selected over the four- and five-factor runs based on the physical interpretation of the factors and the stability of the three-factor run, as measured by a bootstrap with replacement analysis (Norris et al., 2008). Each factor of the three-factor run was replicated with at least 96% accuracy after resampling the data 1,000 times, while the four- and five-factor runs had factors that were replicated

with as little as 11% and 8% accuracy, respectively. The random seed and non-negative SVD seed produced similar factors but with a slightly larger residual error for the non-negative SVD seed than for the random seed (Fig. S9). The SVD seed was chosen for presentation here due to its much lower computational cost.

Figures 3–5 show the results from the non-negative SVD-seed three-factor LS-NMF. The dominant factor (top panels of Figures 3–5) features most of the measured trace gases, except for O₃, propene, and isoprene (Figure 3). This factor appears when winds were from the easterly direction, with many outliers indicating episodic periods of high hydrocarbon abundances (Figure 4). This factor followed a diurnal pattern that is largely dictated by boundary layer dynamics, with surface emissions

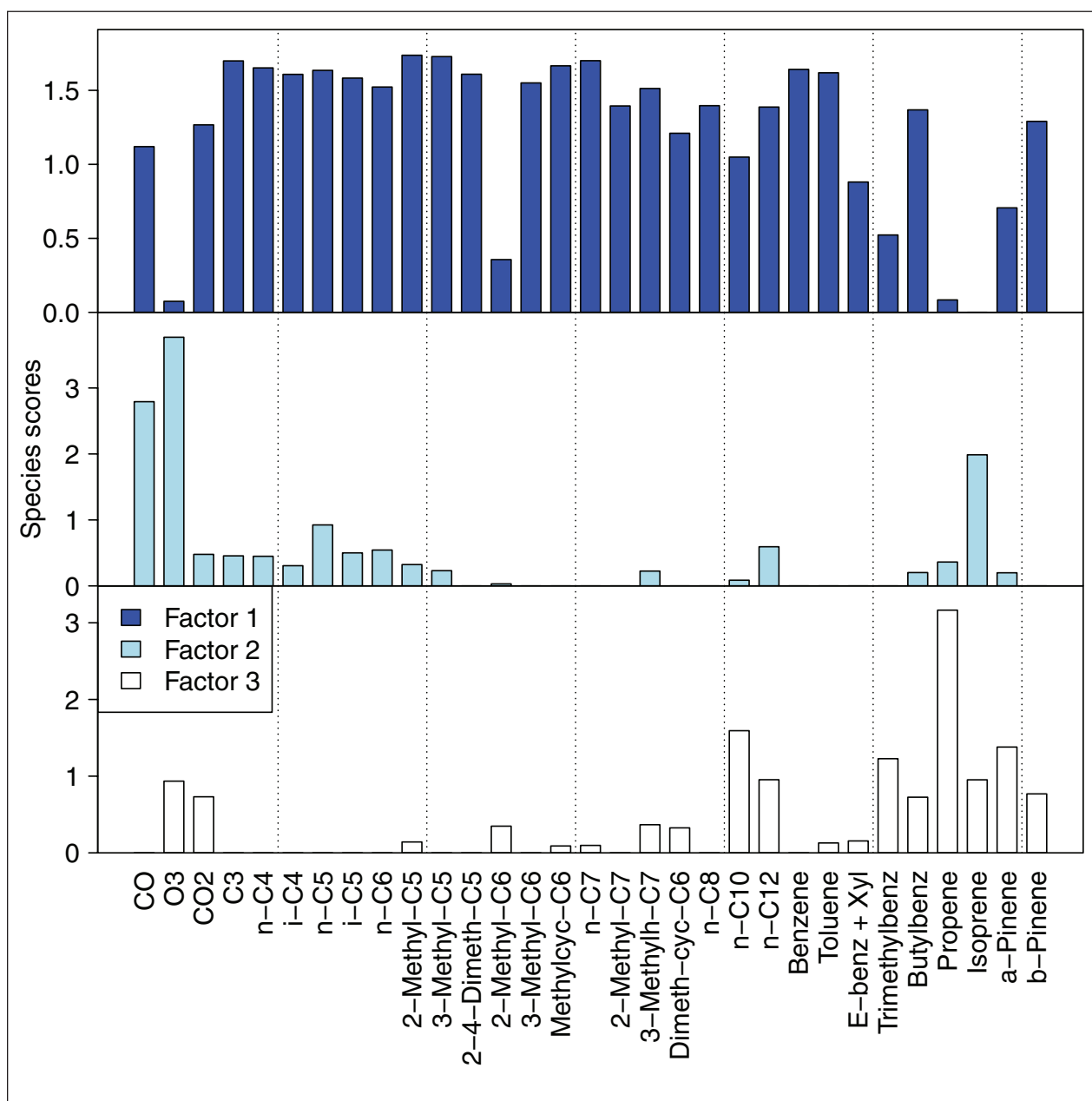


Figure 3: Species scores in factors from the LS-NMF model. The LS-NMF factor composition using all available species measured for the first factor (top panel), second factor (middle panel), and third factor (bottom panel). DOI: <https://doi.org/10.1525/elementa.414.f3>

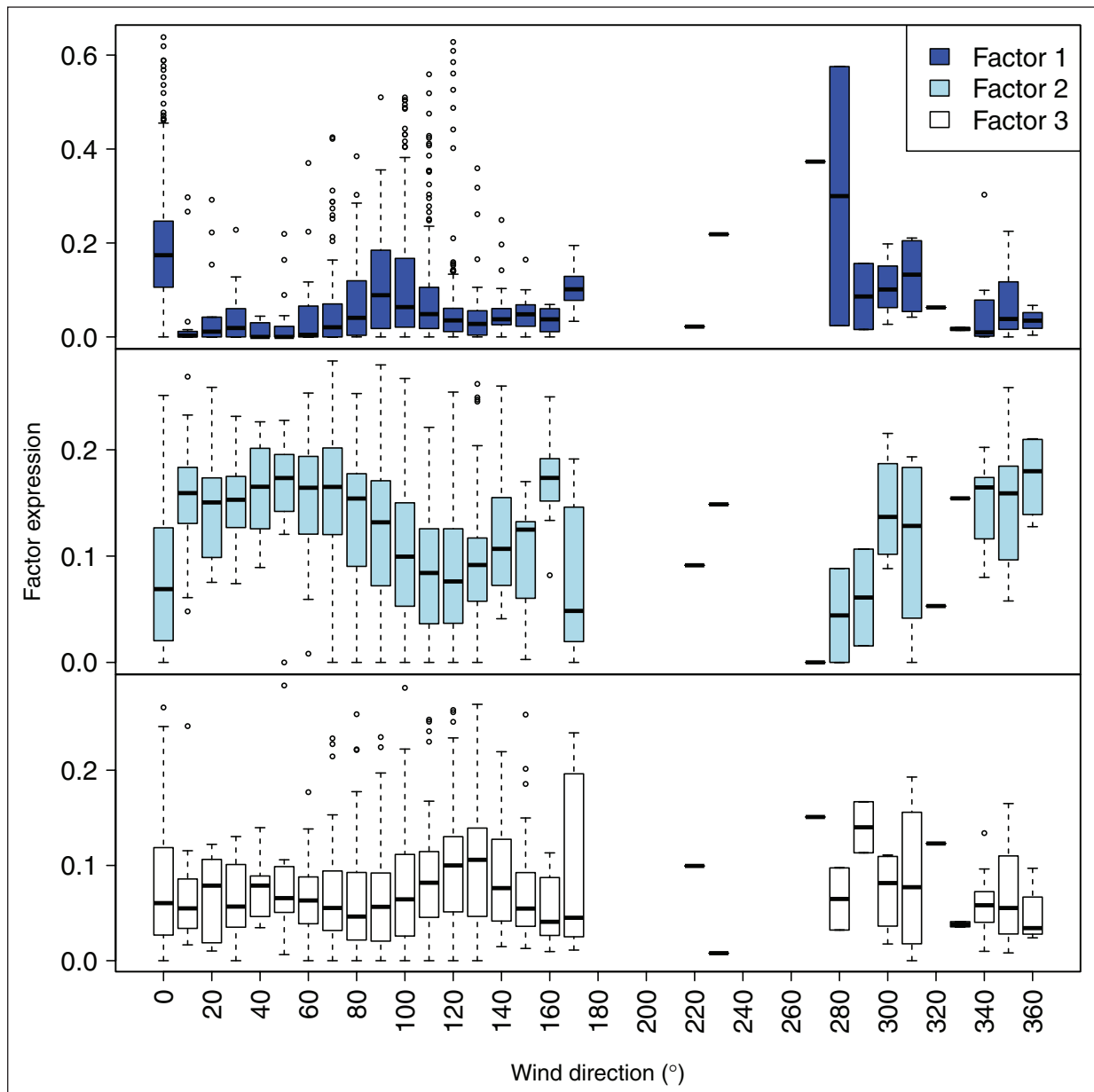


Figure 4: LS-NMF factor expression vs. wind direction. Boxplots of the wind direction (in 10-deg averages, zero degrees indicating calm conditions) dependence of the three factors shown in Fig. 3. Boxplots show medians (thick horizontal bar), interquartile ranges (lower and upper ends of rectangle), estimated 95% confidence limits (lower and upper whiskers), and associated outliers (open circles). DOI: <https://doi.org/10.1525/elementa.414.f4>

accumulating in a shallow nocturnal boundary layer before being mixed into a deeper layer during daytime (Figure 5). This factor most closely matched gasoline exhaust, gasoline headspace vapor, and oil field sources, including natural gas wells and condensate tanks, as listed in the EPA SPECIATE database. Hence, this factor was attributed to general oil field emissions. It explained 53% of the variability in the data set.

The second factor (middle panels of Figures 3–5) was dominated by CO and O₃ (Figure 3). The dominant hydrocarbon in this factor was isoprene, though several relatively long-lived alkanes had high scores as well. This factor had the highest scores when winds were from the northeast sector, indicating a continental air mass origin (Figure 4). The factor features a diurnal cycle that peaked

in the afternoon and slowly declined overnight, with a minimum near sunrise (Figure 5). There were no reasonable matches in the SPECIATE database because the dominant hydrocarbon – isoprene – is generally biogenic while the SPECIATE database consists only of anthropogenic sources. Due to the diurnal cycle and continental origin, we attributed this “source” to transport of regional O₃ and isoprene during the daytime, with distant urban sources explaining elevated CO levels due to direct emissions and hydrocarbon oxidation. This factor explained 28% of the data set’s variability.

The third factor (bottom panels of Figures 3–5) features propene as the dominant component, with CO₂, O₃, and some long-chain hydrocarbons also present (Figure 3). This factor was highest when winds were from

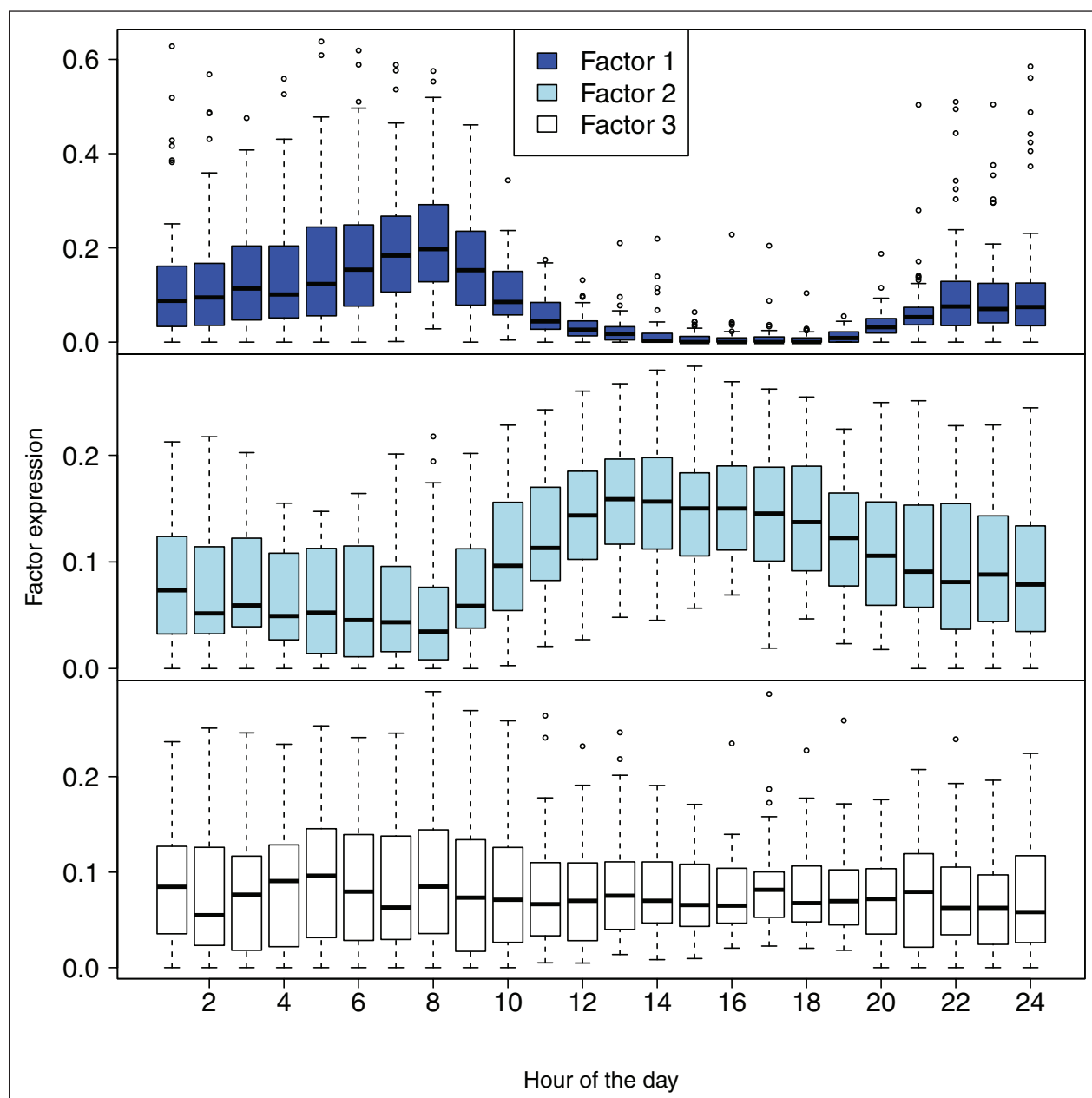


Figure 5: LS-NMF factor expression vs. hour of day. Boxplots of the derived diurnal cycle of the factors. Boxplots show medians (thick horizontal bar), interquartile ranges (lower and upper ends of rectangle), estimated 95% confidence limits (lower and upper whiskers), and associated outliers (open circles). DOI: <https://doi.org/10.1525/elementa.414.f5>

the southeast sector (**Figure 4**). There was no apparent diurnal cycle in the median scores for this factor, though the upper quartile showed elevated scores during the overnight and morning hours and slightly lower scores during the afternoon hours (**Figure 5**). This factor matched gasoline exhaust in the EPA SPECIATE database, and we attribute it to combustion sources including engine exhaust and flaring. It explained 18% of the data set's variability.

4.2.2. General plume identification using trace gas enhancement ratios

Initially, a NO_x/O_3 threshold of 0.25 was used to identify times of substantial O_3 reductions by NO based on ambient mixing ratios. There were 749 thirty-min periods when NO_x/O_3 exceeded 0.25. These periods

showed a strong diurnal dependence with increasing frequency after sunset, peaking during the early morning near 06:00 LST. Of the 749 periods, only one occurred between 11:00-17:00 LST. This is likely due to higher ambient O_3 mixing ratios during daylight hours, and efficient mixing in a turbulent daytime boundary layer reducing the number of defined plumes that reach the measurement site. The correlation between benzene and toluene when $\text{NO}_x/\text{O}_3 > 0.25$ was strong ($R = 0.86$) with a benzene/toluene slope of 1.23 (1.13 – 1.33). Most of these observations exceeded the median benzene and toluene observations during the field campaign, suggesting that this method does generally capture air masses with elevated mixing ratios of trace gases associated with combustion.

A weak correlation between CO_2 and NO_x was observed ($R = 0.11$, with a NO_x/CO_2 slope of 0.15 ppb/ppm (0.14 – 0.16). However, there were distinct clusters of data points with high CO_2 and low NO_x , low CO_2 and high NO_x , and high CO_2 and high NO_x . While high NO_x episodes were likely associated with combustion plumes and possibly episodic high soil NO emissions, high CO_2 events can also be associated with respiration, especially during wet periods (Fig. S5, Meyers, 2001). There was also a weak relationship between NO_x and CO when $\text{NO}_x/\text{O}_3 > 0.25$ ($R = 0.16$). The slope of NO_x/CO in that case was 0.17 (0.15–0.18). A higher NO_x/O_3 threshold of 1 was subsequently used to isolate plumes with increased ozone titration. This threshold identified 103 30 min observations, with a similar diurnal pattern as the observations when NO_x/O_3 exceeded 0.25. Benzene and toluene showed a slightly weaker but still highly significant relationship ($R = 0.81$). In contrast, the relationship between CO_2 and NO_x remained very weak ($R = 0.06$) and statistically insignificant, with a wide range of NO_x/CO_2 ratios. The CO to NO_x relationship was stronger ($R = 0.18$) with a slope of 0.18 (0.15 – 0.22), though the observations still exhibited a wide range of NO_x/CO ratios. The weak correlations between NO_x and both CO and CO_2 suggest that a general threshold NO_x/O_3 ratio for 30 min data and assumed background mixing ratios equal to a moving 10th percentile provides limited insight into the nature of specific combustion sources types affecting the air quality at Shape Ranch. Instead, it is likely that a variety of sources, including both high- and low-temperature combustion sources and biogenic emissions of NO and CO_2 , contributed to the observed enhancements above background.

4.2.3. VOCs within combustion plumes

Enhancements of several VOCs were also tested. **Figure 6** shows correlations of n-butane with n-hexane

and benzene within plumes ($\text{NO}_x/\text{O}_3 > 0.25$), one hour before and after plumes (referred to as plume shoulders), and other non-plume observations. Correlations of n-hexane and n-butane did not show any statistically significant differences between in-plume and out-of-plume observations. The n-butane to benzene ratio was significantly lower within plumes than out of plumes, possibly indicating benzene emissions from combustion, though the range of ratios was still large. This difference in ratios was not statistically significant when a plume definition of $\text{NO}_x/\text{O}_3 > 1$ was used. The sum of C_3 – C_7 alkanes were also compared to CO_2 and NO_x within plumes and out of plumes ($\text{NO}_x/\text{O}_3 > 0.25$, Fig. S10). While non-plume observations showed a positive correlation between alkanes with both NO_x and CO_2 , alkanes were not enhanced within plumes while both NO_x and CO_2 increased. This indicates that combustion plumes were not distinct sources of alkane emission on oil and gas well pads.

An episode with frequent plume detections was examined for evidence of VOC emissions from flaring. There were 62 individual 30 min observations when NO_x/O_3 exceeded 0.25 from 10 September through 13 September 2015 (>30% of the period), separated into nine continuous events. Supporting Figure S11 shows the same correlations as **Figure 6** for the four-day period, with in-plume, plume shoulder, and non-plume observations assessed separately. Like the general plume assessment shown in **Figure 6**, the ratios of n-butane and n-hexane were similar for all three scenarios and the correlations were poorly constrained. However, the n-butane/benzene ratio was reduced within combustion plumes compared to other observations, with the highest benzene mixing ratios occurring during plume passages. Likewise, supporting Figure S12 shows the same correlations as supporting Figure S10 for the four-day period, with lower

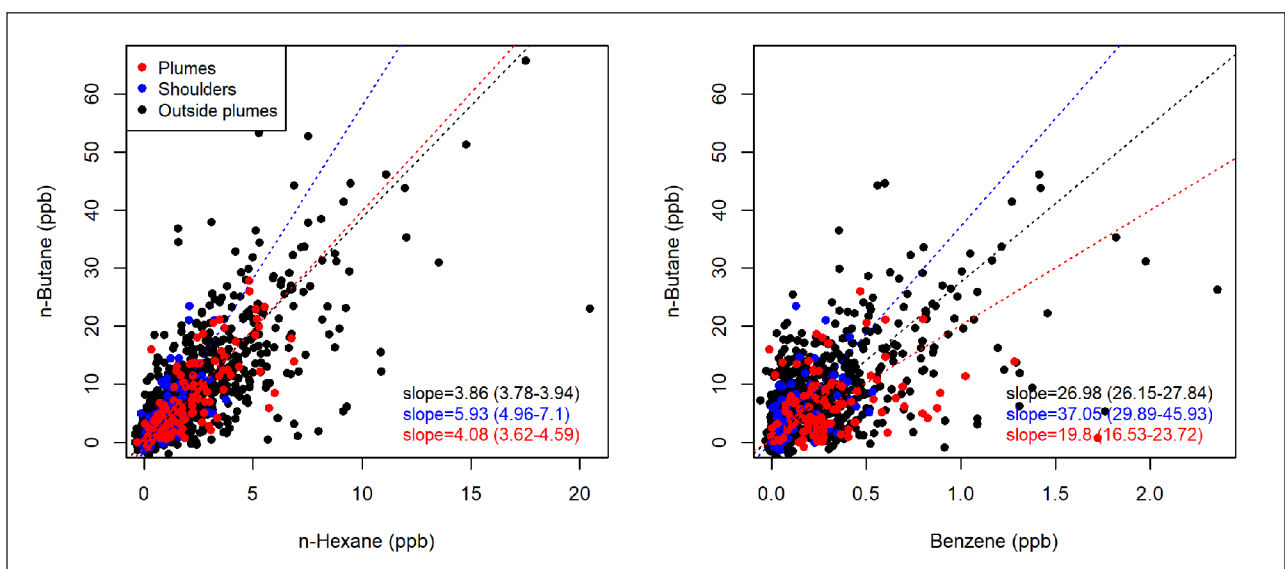


Figure 6: VOC ratios in and out of combustion plumes. Regressions for n-butane vs n-hexane and benzene were compared under three conditions: within plumes (defined as $\text{NO}_x/\text{O}_3 > 0.25$), during plume “shoulders” (1 hour before and after a plume), and all other observations. While uncertainties in the VOC ratios were generally large, a statistically significant difference in n-butane vs. benzene was observed for in-plume and out-of-plume observations, with relatively high benzene mixing ratios in plumes. DOI: <https://doi.org/10.1525/elementa.414.f6>

alkane/ NO_x and alkane/ CO_2 ratios in combustion plumes. While the correlations were weak, the highest NO_x , CO_2 , and alkane mixing ratios occurred in plumes. This four-day period with frequent combustion plume detections demonstrates that while there is evidence of elevated VOC mixing ratios in combustion plumes, the co-location of other VOC emission sources on well pads and the poorly constrained VOC ratios with other trace gases means that VOC emission factors cannot be assessed with confidence for the observed plumes.

4.2.4. Carbon dioxide and nitrogen oxide ratios within plumes

While the NO_x/O_3 ratios within the 30 min dataset provided limited insight into the general NO_x/CO and NO_x/CO_2 emission ratios within combustion plumes, it did succeed in identifying apparent plumes. A total of 176 plume events, defined as consecutive 30 min NO_x/O_3 ratios exceeding 0.25, were identified. These plumes were further analyzed using 10 s data, which allowed for a more robust statistical analysis of trace gas ratios within plumes. Many of the identified events had statistically insignificant relationships between NO_x and CO_2 enhancements (large uncertainty in the slope) or a low correlation coefficient. Furthermore, events with significant rainfall (>3 mm, Hao et al., 2010) observed in the previous three days were discarded to minimize the influence of nitrification and ecosystem respiration on elevated NO_x and CO_2 mixing ratios, respectively. Lastly, events prior to the end of May were discarded as the recording precision of the NO_x instrument data was too low during the early portion of the field campaign.

The correlation coefficient and uncertainty of the NO_x/CO_2 SMA slope, which we interpreted as the average NO_x/CO_2 emission ratio for the detected plumes, were used to analyze subsets of plumes. Three thresholds for the correlation coefficients were used (0.25, 0.50, and 0.75) and two thresholds for the certainty of the slope were used ($\pm 5\%$ and $\pm 10\%$). A summary of the identified plumes is shown in **Table 2**. Out of the 176 identified plume events, 34 met the least-restrictive criteria of $R > 0.25$ and a slope uncertainty of less than $\pm 10\%$, while only three plumes met the most restrictive criteria of $R > 0.75$ and a slope uncertainty of less than $\pm 5\%$. **Figure 7** shows the slope (ppb NO_x /ppm CO_2) and the R value for each of these events, with black points showing plumes with

slope uncertainties less than $\pm 5\%$ and red points showing slope uncertainties between $\pm 5 - 10\%$. Emission ratios of NO_x/CO_2 were generally lower for plumes with slope uncertainties less than $\pm 5\%$. Only one plume with a slope uncertainty less than $\pm 5\%$ had a NO_x/CO_2 ratio greater than the flaring emission ratio of 0.54 ppb/ppm based on AP-42 emission factors. Meanwhile, several plumes with higher slope uncertainties had higher NO_x/CO_2 ratios, with one plume exceeding a NO_x/CO_2 ratio of 4. Still, this ratio is more than a factor of two lower than the expected emission ratios for controlled diesel engines ($\text{NO}_x/\text{CO}_2 = 11.0$ ppm/ppb) and more than a factor of five lower than large, uncontrolled diesel engines ($\text{NO}_x/\text{CO}_2 = 25.7$, U.S. Energy Information Administration, 2016; U.S. Environmental Protection Agency, 2016).

An example of a plume event with a high correlation coefficient is presented in **Figure 8** and is circled in grey in **Figure 7**. For this event on the evening of 24 July 2015, the observed NO_x/CO_2 enhancement ratio of 1.95 was within a factor of four of the ratio of AP-42 emission factors for flares. The high correlation coefficient of the NO_x/CO_2 relationship ($R = 0.91$) suggests that this event was likely a plume from a low-temperature combustion source, such as a flare, although with an apparently higher temperature than the EPA and EIA emission factors for flares suggest, as higher temperature combustion produces more NO_x (Flagan and Seinfeld, 1988; Zeldovich, 1992).

An example for a combustion event that met the least restrictive criteria ($R > 0.25$, slope uncertainty $< \pm 10\%$) is shown in Fig. S13. On 17 July 2015, a prolonged episode of NO_x/CO_2 exceeding 0.25 occurred from approximately 05:00 to 07:00 LST. Most of this time period featured $\text{NO}_x/\text{O}_3 < 0.5$, except for two distinct NO_x plumes when NO_x/O_3 approached 3. During these plumes, NO_x increased by several ppb while CO_2 varied by less than 1 ppm. Though the overall NO_x/CO_2 ratio showed a relationship ($R = 0.47$, slope = 0.94), the two distinct NO_x plumes appeared as a cluster of data points with a much higher NO_x/CO_2 ratio. These two narrow plumes may have been associated with a high-temperature combustion source, such as vehicle exhaust. This demonstrates that the method of screening for plumes using the NO_x/O_3 ratio during episodes of O_3 titration is not perfect, and there exists a potential to mischaracterize the apparent NO_x/CO_2 emission ratio

Table 2: A summary of plumes identified using varying thresholds for the correlation coefficient and the uncertainty of the NO_x/CO_2 slope. DOI: <https://doi.org/10.1525/elementa.414.t2>

Correlation coefficient	Slope uncertainty	Number of identified plumes	NO_x/CO_2 ratios			
			Mean	Median	Min	Max
$R > 0.25$	$< \pm 10\%$	34	0.81	0.51	0.10	4.77
$R > 0.50$		18	0.85	0.45		
$R > 0.75$		7	1.47	0.92		
$R > 0.25$	$< \pm 5\%$	14	0.35	0.36	0.10	0.90
$R > 0.50$		11	0.33	0.27		
$R > 0.75$		3	0.48	0.46		

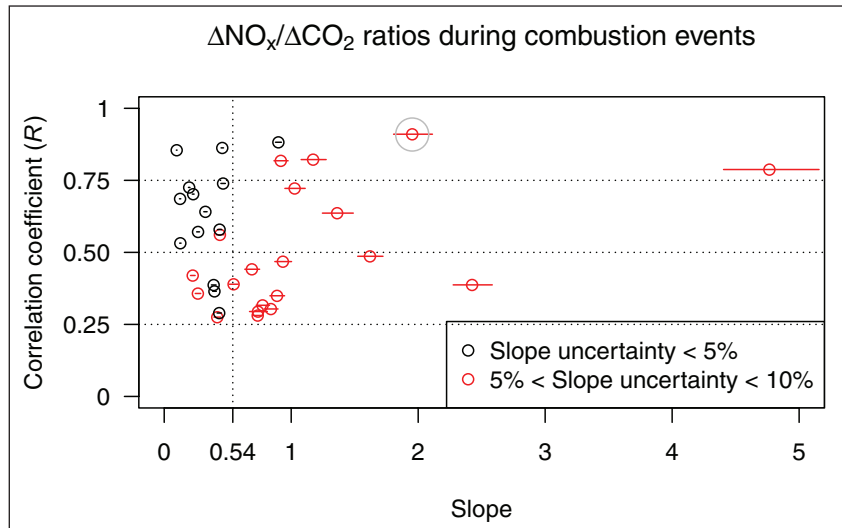


Figure 7: NO_x to CO₂ enhancement ratios within combustion plumes. The slope and *R* values for $\Delta\text{NO}_x/\Delta\text{CO}_2$ for each of the 34 combustion events with meaningful relationships ($R > 0.25$, $R > 0.50$, or $R > 0.75$), low uncertainty in the slope ($< \pm 5\%$ or $< \pm 10\%$), and without significant rainfall (> 3 mm, Hao et al., 2010) during the previous three days. Plumes with a lower slope uncertainty tended to have lower NO_x/CO₂ ratios, though plumes with strong correlation coefficients ($R > 0.75$) had NO_x/CO₂ ratios that spanned more than an order of magnitude. One data point identified with a grey circle is used as a case study in Fig. 8. The NO_x/CO₂ emission factor ratio for flares (0.54) is indicated with a vertical dotted line (U.S. Environmental Protection Agency, 2016). DOI: <https://doi.org/10.1525/elementa.414.f7>

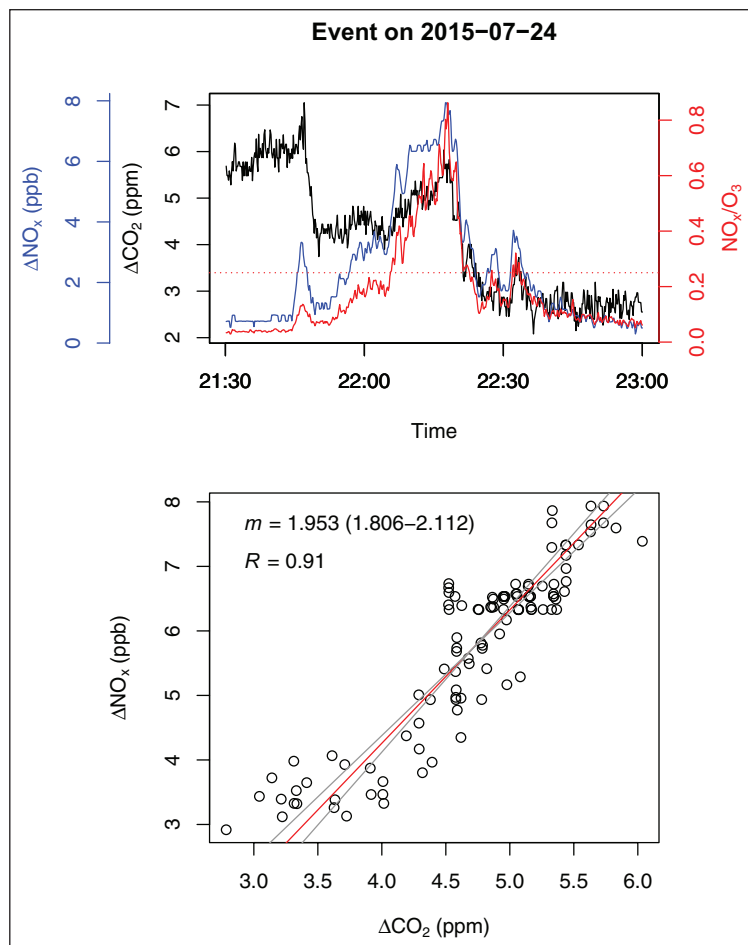


Figure 8: NO_x and CO₂ enhancements for a plume on 24 July 2015. (Top panel) Time series of CO₂ and NO_x measurements above background (black and blue lines, left axis), and NO_x/O₃ ratio (red, right axis) used to identify combustion plumes. The data points in the bottom panel correspond to the NO_x and CO₂ enhancements above background when NO_x/O₃ exceeded 0.25. DOI: <https://doi.org/10.1525/elementa.414.f8>

when a low threshold for the correlation is used. A second example of an apparent high-temperature combustion plume appeared on the morning of 10 September 2015, when several very narrow, distinct plumes occurred between 07:00 and 07:30 LST (Fig. S14). A marginally elevated NO_x/O_3 ratio during the preceding hours resulted in a negative correlation, with the distinct plumes appearing as outliers in the bottom panel of Fig. S14. Due to the very brief duration of these plumes, and the high NO_x/CO_2 ratio of the associated data points, it is likely that a nearby high-temperature combustion source, such as truck or generator exhaust, was responsible for those plumes.

Overall, the combustion events that met the set correlation and slope uncertainty criteria imply that low-temperature combustion sources with low NO_x/CO_2 emission ratios dominated combustion plumes observed at Shape Ranch. The highest observed ratio of 4.77 was more than a factor of two lower than the expected emission ratios for diesel engines. Most of the other plumes had NO_x/CO_2 ratios lower than 2, implying that low-temperature combustion sources were mostly responsible for the observed plumes at Shape Ranch. Regardless of the chosen correlation coefficient threshold, the median NO_x/CO_2 ratios were between 0.27 and 0.92 ppb/ppm – all within a factor of two of the ratio of 0.54 derived from the emission factors for flaring (U.S. Environmental Protection Agency, 2016). A more conservative correlation threshold of $R > 0.75$ resulted in generally higher NO_x/CO_2 ratios; though with fewer data points, the statistical measures of these slopes become less defensible.

Overall, these results imply that (i) low-temperature combustion sources were generally responsible for the observed plume events, which are assumed to be dominated by flares, but they have widely varying emission factors, and (ii) a combination of other sources of NO_x and CO_2 , including biogenic sources and high temperature combustion, as well as atmospheric transport and chemistry, may interfere with the detection of distinct plumes from distant combustion sources.

5. Conclusions

Uncertainties in emission estimates of both VOCs and NO_x from the shale industry prevent an accurate estimate of the regional air quality impacts associated with oil and gas production in the area. This is especially true in areas with widespread flaring, which has poorly quantified emissions due to uncertain emission factors. Ambient air quality was studied on Shape Ranch in the sparsely populated western portion of the EFS. Transport of continental air was associated with CO , O_3 , isoprene, and short-chain, longer lived alkanes. High alkane abundances indicated regional oil and natural gas production as a dominant emission source, which was confirmed by a non-negative matrix factorization analysis. Combustion was also identified as a local emission source, though mixing ratios of associated aromatics, alkenes, CO , and NO_x were generally low.

Flaring was identified as a significant source of NO_x and benzene emissions, similar to our previous work (Schade and Roest, 2018). Analyzed NO_x/CO_2 ratios during episodes of high NO_x/O_3 suggest that low-temperature combustion is very often responsible for plumes identified at

Shape Ranch. While flaring is the most likely source for many of these events, the NO_x/CO_2 ratio exhibited variability spanning more than an order of magnitude, suggesting that these low-temperature combustion sources have widely varying emission factors and/or other sources of NO_x and CO_2 interfere to introduce larger variability than would otherwise be expected. Regardless of the NO_x/CO_2 correlation and slope uncertainty thresholds used to identify combustion plumes, the observed median NO_x/CO_2 ratio in these plumes was within a factor of two of the expected ratio calculated from the EPA AP-42 flaring NO_x emissions factor.

Determining the relevant emissions from flares can be improved by studying emission factors associated with these relatively small diffusion flares in the field. Closely monitoring and reporting flaring activity data could improve inventories, which in turn would more accurately reflect real-world flaring and its impacts in liquids-rich shale plays such as the EFS, the Permian, and the Bakken shale.

Data accessibility statement

Data tables (.csv) and code (R scripts) are available through the Texas Data Repository (<https://doi.org/10.18738/T8/ISKZOC>). Ancillary data sources have been included in the bibliography.

Supplemental file

The Supplemental file for this article can be found as follows:

- **Supplemental material.** Figures S1–14. Tables S1–3. Text S1. References. DOI: <https://doi.org/10.1525/elementa.414.s1>

Acknowledgements

We thank Hugh Asa Fitzsimons III for providing access to his ranch land, and Freddy Longoria for his assistance with setting up the field site. We also thank Peter Bella for assistance in traveling and coordinating the field work on the ranch. Lastly, we thank Sarah Brooks, Jake Zenker, and Kristen Collier for assistance during the field campaign.

Funding information

This work was funded by a private grant from the owner of Shape Ranch, Hugh Asa Fitzsimons III, and was indirectly supported by internal funds from Texas A&M University.

Competing interests

The authors have no competing interests to declare.

Author contributions

- Contributed to conception and design: GWS designed the field study.
- Contributed to acquisition of data: GSR and GWS contributed to data acquisition and quality control.
- Contributed to analysis and interpretation of data: GSR and GWS contributed to the general air quality analysis and non-negative matrix factorization. GSR performed the analysis of trace gas ratios within combustion plumes.

- Drafted and/or revised the article: GSR drafted the article and GWS provided revisions and guidance.
- Approved the submitted version for publication: GSR and GWS have approved the submitted version for publication.

References

- Andela, N, Morton, DC, Giglio, L, Paugam, R, Chen, Y, Hantson, S, van der Werf, GR and Randerson, JT.** 2019. The Global Fire Atlas of individual fire size, duration, speed and direction. *Earth Syst. Sci. Data* **11**: 529–552. DOI: <https://doi.org/10.5194/essd-11-529-2019>
- Arguez, A, Durre, I, Applequist, S, Squires, M, Vose, R, Yin, X and Bilotta, R.** 2010. U.S. Climate Normals Product Suite (1981–2010). DOI: <https://doi.org/10.7289/V5PN93JP>
- Ayers, GP.** 2001. Comment on regression analysis of air quality data. *Atmos. Environ* **35**: 2423–2425. DOI: [https://doi.org/10.1016/S1352-2310\(00\)00527-6](https://doi.org/10.1016/S1352-2310(00)00527-6)
- Baker, AK, Beyersdorf, AJ, Doezema, LA, Katzenstein, A, Meinardi, S, Simpson, IJ, Blake, DR and Sherwood Rowland, F.** 2008. Measurements of nonmethane hydrocarbons in 28 United States cities. *Atmos. Environ* **42**: 170–182. DOI: <https://doi.org/10.1016/j.atmosenv.2007.09.007>
- Benjamin, MT, Sudol, M, Bloch, L and Winer, AM.** 1996. Low-emitting urban forests: A taxonomic methodology for assigning isoprene and monoterpene emission rates. *Atmos. Environ* **30**: 1437–1452. DOI: [https://doi.org/10.1016/1352-2310\(95\)00439-4](https://doi.org/10.1016/1352-2310(95)00439-4)
- Brunet, J-P, Tamayo, P, Golub, TR and Mesirov, JP.** 2004. Metagenes and molecular pattern discovery using matrix factorization. *Proc. Natl. Acad. Sci* **101**: 4164–4169. DOI: <https://doi.org/10.1073/pnas.0308531101>
- Chauhan, AJ, Krishna, MT, Frew, AJ and Holgate, ST.** 1998. Exposure to nitrogen dioxide (NO₂) and respiratory disease risk. *Rev. Environ. Health* **13**: 73–90.
- Chin, M, Jacob, DJ, Munger, JW, Parrish, DD and Doddridge, BG.** 1994. Relationship of ozone and carbon monoxide over North America. *J. Geophys. Res. Atmospheres* **99**: 14565–14573. DOI: <https://doi.org/10.1029/94JD00907>
- Elvidge, C, Zhizhin, M, Baugh, K, Hsu, F-C and Ghosh, T.** 2015. Methods for Global Survey of Natural Gas Flaring from Visible Infrared Imaging Radiometer Suite Data. *Energies* **9**: 14. DOI: <https://doi.org/10.3390/en9010014>
- Flagan, RC and Seinfeld, JH.** 1988. Fundamentals of air pollution engineering. Englewood Cliffs, N.J.: Prentice Hall.
- Gaujoux, R and Seoighe, C.** 2010. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* **11**: 367. DOI: <https://doi.org/10.1186/1471-2105-11-367>
- Gilman, JB, Kuster, WC, Goldan, PD, Herndon, SC, Zahniser, MS, Tucker, SC, Brewer, WA, Lerner, BM, Williams, EJ, Harley, RA, Fehsenfeld, FC, Warneke, C and de Gouw, JA.** 2009. Measurements of volatile organic compounds during the 2006 TexAQSGoMACCS campaign: Industrial influences, regional characteristics, and diurnal dependencies of the OH reactivity. *J. Geophys. Res. Atmospheres* **114**. DOI: <https://doi.org/10.1029/2008JD011525>
- Gilman, JB, Lerner, BM, Kuster, WC and de Gouw, JA.** 2013. Source Signature of Volatile Organic Compounds from Oil and Natural Gas Operations in Northeastern Colorado. *Environ. Sci. Technol* **3**: 1297–1305. DOI: <https://doi.org/10.1021/es304119a>
- Goldstein, AH, Fan, SM, Goulden, ML, Munger, JW and Wofsy, SC.** 1996. Emissions of ethene, propene, and 1-butene by a midlatitude forest. *J. Geophys. Res. Atmospheres* **101**: 9149–9157. DOI: <https://doi.org/10.1029/96JD00334>
- Gunther, AB, Jiang, X, Heald, CL, Sakulyanontvittaya, T, Duhl, T, Emmons, LK and Wang, X.** 2012. The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions. *Geosci Model Dev* **5**: 1471–1492. DOI: <https://doi.org/10.5194/gmd-5-1471-2012>
- Guha, A, Gentner, DR, Weber, RJ, Provencal, R and Goldstein, AH.** 2015. Source apportionment of methane and nitrous oxide in California's San Joaquin Valley at CalNex 2010 via positive matrix factorization. *Atmos Chem Phys* **15**: 12043–12063. DOI: <https://doi.org/10.5194/acp-15-12043-2015>
- Hao, Y, Wang, Y, Mei, X and Cui, X.** 2010. The response of ecosystem CO₂ exchange to small precipitation pulses over a temperate steppe. *Plant Ecol.* **209**: 335–347. DOI: <https://doi.org/10.1007/s11258-010-9766-1>
- Koss, A, Yuan, B, Warneke, C, Gilman, JB, Lerner, BM, Veres, PR, Peischl, J, Eilerman, S, Wild, R, Brown, SS, Thompson, CR, Ryerson, T, Hanisco, T, Wolfe, GM, Clair, JMSt, Thayer, M, Keutsch, FN, Murphy, S and de Gouw, J.** 2017. Observations of VOC emissions and photochemical products over US oil- and gas-producing regions using high-resolution H₃O⁺ CIMS (PTR-ToF-MS). *Atmospheric Meas. Tech* **10**: 2941–2968. DOI: <https://doi.org/10.5194/amt-10-2941-2017>
- Kota, SH, Park, C, Hale, MC, Werner, ND, Schade, GW and Ying, Q.** 2014. Estimation of VOC emission factors from flux measurements using a receptor model and footprint analysis. *Atmos. Environ* **82**: 24–35. DOI: <https://doi.org/10.1016/j.atmosenv.2013.09.052>
- Lee, DD and Seung, HS.** 1999. Learning the parts of objects by non-negative matrix factorization. *Nature* **401**: 789. DOI: <https://doi.org/10.1038/44565>
- Majid, A, Val Martin, M, Lamsal, LN and Duncan, BN.** 2017. A decade of changes in nitrogen oxides over regions of oil and natural gas activity in the United States. *Elem Sci Anth* **5**: 76. DOI: <https://doi.org/10.1525/elementa.259>
- Meyers, TP.** 2001. A comparison of summertime water and CO₂ fluxes over rangeland for well watered and drought conditions. *Agric. For. Meteorol*

- 106:** 205–214. DOI: [https://doi.org/10.1016/S0168-1923\(00\)00213-6](https://doi.org/10.1016/S0168-1923(00)00213-6)
- Norris, GA, Vedantham, R, Wade, K, Brown, S, Prouty, J and Foley, C.** 2008. EPA Positive Matrix Factorization (PMF) 3.0 Fundamentals & User Guide. Washington, DC: Environmental Protection Agency.
- Olaguer, EP.** 2012. The potential near-source ozone impacts of upstream oil and gas industry emissions. *J. Air Waste Manag. Assoc* **62**: 966–977. DOI: <https://doi.org/10.1080/10962247.2012.688923>
- Paatero, P and Tapper, U.** 1994. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* **5**: 111–126. DOI: <https://doi.org/10.1002/env.3170050203>
- Pacsi, AP, Kimura, Y, McGaughey, G, McDonald-Buller, EC and Allen, DT.** 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**: 3966–3973. DOI: <https://doi.org/10.1021/es5055012>
- Park, C, Schade, GW and Boedeker, I.** 2010. Flux measurements of volatile organic compounds by the relaxed eddy accumulation method combined with a GC-FID system in urban Houston, Texas. *Atmos. Environ* **44**: 2605–2614. DOI: <https://doi.org/10.1016/j.atmosenv.2010.04.016>
- Parrish, D, Kembal-Cook, S, Grant, J and Yarwood, G.** 2017. Science Synthesis Report: Atmospheric Impacts of Oil and Gas Development in Texas (Final Report No. 582-16-63215–16). Austin, Texas: Texas Commission on Environmental Quality.
- Peischl, J, Eilerman, SJ, Neuman, JA, Aikin, KC, de Gouw, J, Gilman, JB, Herndon, SC, Nadkarni, R, Trainer, M, Warneke, C and Ryerson, TB.** 2018. Quantifying Methane and Ethane Emissions to the Atmosphere From Central and Western U.S. Oil and Natural Gas Production Regions. *J. Geophys. Res. Atmospheres* **123**: 7725–7740. DOI: <https://doi.org/10.1029/2018JD028622>
- Prenni, AJ, Day, DE, Evanoski-Cole, AR, Sive, BC, Hecobian, A, Zhou, Y, Gebhart, KA, Hand, JL, Sullivan, AP, Li, Y, Schurman, MI, Desyaterik, Y, Malm, WC, Collett, JL, Jr. and Schichtel, BA.** 2016. Oil and gas impacts on air quality in federal lands in the Bakken region: an overview of the Bakken Air Quality Study and first results. *Atmospheric Chem. Phys* **16**: 1401–1416. DOI: <https://doi.org/10.5194/acp-16-1401-2016>
- R Core Team.** 2018. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
- Reff, A, Eberly, SI and Bhave, PV.** 2007. Receptor Modeling of Ambient Particulate Matter Data Using Positive Matrix Factorization: Review of Existing Methods. *J. Air Waste Manag. Assoc.* **57**: 146–154. DOI: <https://doi.org/10.1080/10473289.2007.10465319>
- Roest, G and Schade, G.** 2017. Quantifying alkane emissions in the Eagle Ford Shale using boundary layer enhancement. *Atmospheric Chem. Phys* **17**: 11163–11176. DOI: <https://doi.org/10.5194/acp-17-11163-2017>
- Rubin, JI, Kean, AJ, Harley, RA, Millet, DB and Goldstein, AH.** 2006. Temperature dependence of volatile organic compound evaporative emissions from motor vehicles. *J. Geophys. Res. Atmospheres* **111**. DOI: <https://doi.org/10.1029/2005JD006458>
- Schade, GW and Roest, GS.** 2015. Is the Shale Boom Reversing Progress in Curbing Ozone Pollution? *Eos* **96**. DOI: <https://doi.org/10.1029/2015EO028279>
- Schade, GW and Roest, G.** 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. Anthr* **4**: 000096. DOI: <https://doi.org/10.12952/journal.elementa.000096>
- Schade, GW and Roest, G.** 2018. Source apportionment of non-methane hydrocarbons, NO_x and H₂S data from a central monitoring station in the Eagle Ford shale, Texas. *Elem Sci Anth* **6**. DOI: <https://doi.org/10.1525/elementa.289>
- Stroscher, MT.** 2000. Characterization of Emissions from Diffusion Flare Systems. *J. Air Waste Manag. Assoc* **50**: 1723–1733. DOI: <https://doi.org/10.1080/10473289.2000.10464218>
- Swarthout, RF, Russo, RS, Zhou, Y, Miller, BM, Mitchell, B, Horsman, E, Lipsky, E, McCabe, DC, Baum, E and Sive, BC.** 2015. Impact of Marcellus Shale Natural Gas Development in Southwest Pennsylvania on Volatile Organic Compound Emissions and Regional Air Quality. *Environ. Sci. Technol* **49**: 3175–3184. DOI: <https://doi.org/10.1021/es504315f>
- Texas Commission on Environmental Quality.** 2020. Data by Year by Site by Parameter [WWW Document]. URL https://www.tceq.texas.gov/cgi-bin/compliance/monops/yearly_summary.pl?cams=1070 (accessed 1.30.20).
- Texas Department of Transportation.** 2018. AADT and AADT Trucks by Year for 1/1/2008 – 12/31/2017: Criteria: Pt = FM0186-KG (Transportation Planning and Programming Division's Statewide Traffic Analysis and Reporting System II).
- Thiem, A, Schlink, U, Pan, X-C, Hu, M, Peters, A, Wiedensohler, A, Breitner, S, Cyrus, J, Wehner, B, Rösch, C and Franck, U.** 2012. Using non-negative matrix factorization for the identification of daily patterns of particulate air pollution in Beijing during 2004–2008. *Atmospheric Chem. Phys. Discuss* **12**: 13015–13052. DOI: <https://doi.org/10.5194/acpd-12-13015-2012>
- Torres, VM, Herndon, S, Wood, E, Al-Fadhli, FM and Allen, DT.** 2012. Emissions of Nitrogen Oxides from Flares Operating at Low Flow Conditions. *Ind. Eng. Chem. Res* **51**: 12600–12605. DOI: <https://doi.org/10.1021/ie300179x>
- U.S. Code of Federal Regulations.** 2015. 40 CFR Chapter I Subchapter C Part 50§50.19, Code of Federal Regulations.

- U.S. Energy Information Administration.** 2016. Carbon Dioxide Emissions Coefficients [WWW Document]. URL http://www.eia.gov/environment/emissions/co2_vol_mass.cfm (accessed 5.23.16).
- U.S. Energy Information Administration.** 2018a. How much shale (tight) oil is produced in the United States? [WWW Document]. URL <https://www.eia.gov/tools/faqs/faq.php?id=847&t=6> (accessed 5.16.18).
- U.S. Energy Information Administration.** 2018b. How much shale gas is produced in the United States? [WWW Document]. URL <https://www.eia.gov/tools/faqs/faq.php?id=907&t=8> (accessed 5.16.18).
- U.S. Environmental Protection Agency.** 2015a. SPECIATE Version 4.5 through 4.0 [WWW Document]. *US EPA*. URL <https://www.epa.gov/air-emissions-modeling/speciate-version-45-through-40> (accessed 6.8.18).
- U.S. Environmental Protection Agency.** 2015b. National Ambient Air Quality Standards for Ozone; Final Rule.
- U.S. Environmental Protection Agency.** 2016. AP-42: Compilation of Air Emissions Factors [WWW Document]. *US EPA*. URL <https://www.epa.gov/air-emissions-factors-and-quantification/ap-42-compilation-air-emissions-factors> (accessed 5.17.18).
- Vinken, GCM, Boersma, KF, Maasakkers, JD, Adon, M and Martin, RV.** 2014. Worldwide biogenic soil NO_x emissions inferred from OMI NO₂ observations. *Atmospheric Chem. Phys* **14**: 10363–10381. DOI: <https://doi.org/10.5194/acp-14-10363-2014>
- Wang, G, Kossenkov, AV and Ochs, MF.** 2006. LS-NMF: A modified non-negative matrix factorization algorithm utilizing uncertainty estimates. *BMC Bioinformatics* **7**: 175. DOI: <https://doi.org/10.1186/1471-2105-7-175>
- Williams, EJ, Hutchinson, GL and Fehsenfeld, FC.** 1992. NO_x And N₂O Emissions From Soil. *Glob. Biogeochem. Cycles* **6**: 351–388. DOI: <https://doi.org/10.1029/92GB02124>
- Willyard, KA and Schade, GW.** 2019. Flaring in two Texas shale areas: Comparison of bottom-up with top-down volume estimates for 2012 to 2015. *Sci. Total Environ* **691**: 243–251. DOI: <https://doi.org/10.1016/j.scitotenv.2019.06.465>
- Worden, HM, Deeter, MN, Frankenberg, C, George, M, Nichitiu, F, Worden, J, Aben, I, Bowman, KW, Clerboux, C, Coheur, PF, de Laat, ATJ, Detweiler, R, Drummond, JR, Edwards, DP, Gille, JC, Hurtmans, D, Luo, M, Martínez-Alonso, S, Massie, S, Pfister, G and Warner, JX.** 2013. Decadal record of satellite carbon monoxide observations. *Atmospheric Chem. Phys* **13**: 837–850. DOI: <https://doi.org/10.5194/acp-13-837-2013>
- Zeldovich, YB.** 1992. Selected Works of Yakov Borisovich Zeldovich, Princeton Legacy Library. Princeton, New Jersey: Princeton University Press. DOI: <https://doi.org/10.1515/9781400862979>

How to cite this article: Roest, GS and Schade, GW. 2020. Air quality measurements in the western Eagle Ford Shale. *Elem Sci Anth*, 8: 18. DOI: <https://doi.org/10.1525/elementa.414>

Domain Editor-in-Chief: Detlev Helmig, Institute of Alpine and Arctic Research, University of Colorado Boulder, US

Guest Editor: Stefan Schwietzke, Environmental Defense Fund, US

Knowledge Domain: Atmospheric Science

Part of an *Elementa* Forum: Oil and Natural Gas Development: Air Quality, Climate Science, and Policy

Submitted: 11 August 2019

Accepted: 10 April 2020

Published: 08 May 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Elem Sci Anth is a peer-reviewed open access journal published by University of California Press.

OPEN ACCESS