DISSERTATION

ATMOSPHERIC AND AIR QUALITY IMPLICATIONS OF C_2-C_5 ALKANE EMISSIONS FROM THE OIL AND GAS SECTOR

Submitted by

Zitely Asafay Tzompa Sosa

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2018

Doctoral Committee:

Advisor: Emily Fischer

Sonia M. Kreidenweis
Jeffrey Pierce
Shantanu Jathar

ABSTRACT

ATMOSPHERIC AND AIR QUALITY IMPLICATIONS OF C_2-C_5 ALKANE EMISSIONS FROM THE OIL AND GAS SECTOR

Emissions of C_2-C_5 alkanes from the U.S. oil and gas sector have changed rapidly over the last decade. This dissertation quantifies the role of the oil and gas sector on light alkane emissions and abundances at local, regional, and global scales. First, we present an updated global ethane (C_2H_6) emission inventory based on 2010 satellite-derived CH_4 fluxes with adjusted C_2H_6 emissions over the U.S. from the National Emission Inventory (NEI 2011). We contrast our global 2010 C_2H_6 emission inventory with one developed for 2001. The C_2H_6 difference between global anthropogenic emissions is subtle (7.9 versus 7.2 Tg yr^{-1}), but the spatial distribution of the emissions is distinct. In the 2010 C_2H_6 inventory, fossil fuel sources in the Northern Hemisphere represent half of global C_2H_6 emissions and 95% of global fossil fuel emissions. Over the U.S., un-adjusted NEI 2011 C_2H_6 emissions produce mixing ratios that are 14-50 % of those observed by aircraft observations (2008-2014). When the NEI 2011 C_2H_6 emission totals are scaled by a factor of 1.4, the GEOS-Chem model largely reproduces a regional suite of observations, with the exception of the central U.S., where it continues to under-predict observed mixing ratios in the lower troposphere.

Second, we use a nested GEOS-Chem simulation driven by updated 2011NEI emissions with aircraft, surface and column observations to 1) document spatial patterns in the emissions and observed atmospheric abundances of C_2-C_5 alkanes over the U.S., and 2) estimate the contribution of emissions from the U.S. oil and gas industry to these patterns. The oil and gas

sector in the updated 2011NEI contributes >80% of the total U.S. emissions of C_2H_6 and propane (C_3H_8), and emissions of these species are largest in the central U.S. Observed mixing ratios of C_2-C_5 alkanes show enhancements over the central U.S. below 2 km. A nested GEOS-Chem simulation underpredicts observed C_3H_8 mixing ratios in the boundary layer over several U.S. regions and the relative underprediction is not consistent, suggesting C_3H_8 emissions should receive more attention moving forward. Our decision to consider only C_4-C_5 alkane emissions as a single lumped species produces a geographic distribution similar to observations. Due to the increasing importance of oil and gas emissions in the U.S., we recommend continued support of existing long-term measurements of C_2-C_5 alkanes. We suggest additional monitoring of C_2-C_5 alkanes downwind of northeastern Colorado, Wyoming and western North Dakota to capture changes in these regions. The atmospheric chemistry modeling community should also evaluate whether chemical mechanisms that lump $\leq C_6$ alkanes are sufficient to understand air quality issues in regions with large emissions of these species.

Finally, we investigate the contribution of C_2-C_5 alkane emissions from the U.S. oil and gas industry to O_3 abundances at regional and global scales. Emissions of C_2-C_5 alkanes from the oil and gas sector make the largest contribution to ozone (O_3) production over the central U.S. compared to other regions. The Colorado Front Range is the 8-hour O_3 non-attainment area with the highest summertime daytime average O_3 enhancement attributed to the U.S. oil and gas sector. The global tropospheric contribution of C_2-C_5 alkane emissions from the U.S. oil and gas sector to the O_3 burden is 0.5 Tg for the year 2011, which represents 0.27% of the Northern Hemisphere tropospheric O_3 burden.

ACKNOWLEDGEMENTS

Funding for Zitely A. Tzompa-Sosa was provided by Consejo Nacional de Ciencia y Tecnología (CONACYT) under fellowship No. 216028, Mario Molina para Ciencias Ambientales fund, the Colorado State University Department of Atmospheric Science Assisting Students, Cultivating Excellence, Nurturing Talent (ASCENT) fund, and NOAA under award number NA14OAR4310148. I acknowledge NDACC for FTIR solar data provision. The FTIR data used in this publication are publicly available at http://www.ndacc.org. The Toronto measurements were made at the University of Toronto Atmospheric Observatory (TAO), which has been supported by CFCAS, ABB Bomem, CFI, CSA, ECCC, NSERC, ORDCF, PREA, and the University of Toronto. Analysis of the Toronto NDACC data was supported by the CAFTON project, funded by the Canadian Space Agency's FAST Program. Thanks to the International Foundation High Altitude Research Stations Jungfraujoch and Gornergrat (HFSJG, Bern) for supporting the facilities needed to perform the Jungfraujoch observations. The global VOC flask analyses are a component of NOAA's Cooperative USA and global-scale Greenhouse Gas Reference flask sampling network, which is supported in part by NOAA Climate Program Office's AC4 program. The National Center for Atmospheric Research is sponsored by the National Science Foundation. The NCAR FTIR program is supported under contract by the National Aeronautics and Space Administration (NASA). Portions of the research described in this dissertation have been reviewed by the U.S. Environmental Protection Agency and approved for publication. Approval does not signify that the contents necessarily reflect the views and the policies of the Agency nor does mention of trade names or commercial products constitute endorsement or recommendation for use. The 2010 global C_2H_6 emission inventory and the

updated 2011NEI emission inventory can be accessed via the Colorado State University Digital Repository at http://hdl.handle.net/10217/178758 and https://hdl.handle.net/10217/187477, respectively.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION[1]

The rise in oil prices combined with the expansion of unconventional techniques of extraction (horizontal drilling and hydraulic fracturing) caused domestic production of oil and gas to experience a rapid growth in the U.S. since 2005 (U.S. EIA, 2017), increasing emission rates of many trace gases over oil and gas-producing basins (de Gouw et al., 2014; Kort et al., 2016). Between 2005 and 2017, U.S. natural gas production increased 42% (U.S. EIA, 2017). Emission sources associated with oil and gas production leak a variety of volatile organic compounds (VOCs) to the atmosphere (Collett et al., 2016; Gilman et al., 2013; Lee et al., 2006; Roest and Schade, 2017; Swarthout et al., 2013; Swarthout et al., 2015). VOC emissions from the oil and gas sector occur during well development and production phases (Collett et al., 2016; Pacsi et al., 2015), and emissions to the atmosphere also continue when wells are abandoned (Kang et al., 2014). These emissions can impact climate (Brandt et al., 2014; Brantley et al., 2014; Franco et al., 2016; Mitchell et al., 2015; Roscioli et al., 2015), the formation of ozone (O_3) and aerosols (Field et al., 2015; Guo, 2012; Koss et al., 2015; Pacsi et al., 2015; Phillips-Smith et al., 2017; Pusede and Cohen, 2012; Rappenglück et al., 2014), and human exposure to air toxics (Brantley et al., 2015; Halliday et al., 2016; Zielinska et al., 2014). Observations suggest that depending on the lifetime and emission rate of each species, the impact on atmospheric abundances of VOCs emitted by oil and gas sources can be substantial at local, regional, and global scales. For example, inside the Denver-Julesburg Basin Gilman et al. (2013) estimated that oil and gas sources are the dominant source (72-96 %) of regional C_2 to C_7 alkane emissions. Similarly, in the Uintah Basin, oil and gas leakage contributes 43-82% of observed

[1] This chapter contains published work from: Tzompa-Sosa, Z. A., et al. (2017), Revisiting global fossil fuel and biofuel emissions of ethane, Journal of Geophysical Research: Atmospheres, 122, doi:10.1002/2016JD025767.

abundances of C_2-C_5 alkanes (Helmig et al., 2014b; Swarthout et al., 2015). In the Marcellus shale region, multiple studies show that unconventional oil and gas production is responsible for recent positive trends in the observed abundances of methane (CH_4) and ethane (C_2H_6) (Goetz et al., 2017; Peischl et al., 2015b; Vinciguerra et al., 2015). In the Northern Hemisphere annual growth rates of C_2H_6 abundances of 3–5% yr^{-1} between 2009–2014 have been attributed to the recent increase of oil and gas extraction in North America (Franco et al., 2016; Helmig et al., 2016).

In the context of rapidly changing industrial activities and the fact that production is often driven by transitory economics, updating emission inventories for the U.S. oil and gas sector is a challenge. In addition to the rapid growth of the oil and gas industry, there are a number of factors that make constraining VOC emissions from this industry difficult: 1) Natural gas composition varies with the type of reservoir (e.g., tight gas vs. shale gas) (Kort et al., 2016; Tzompa-Sosa et al., 2017; Warneke et al., 2014); 2) Emissions depend on the stage (e.g., development, production or abandoned) of a well. Most of the VOC emissions occur during production (Pacsi et al., 2015), but emissions can continue for decades even after the well has been abandoned (Kang et al., 2014); 3) Emission inventories rely on activity factors and emission factors that represent typical emission rates for oil and gas wells. However, Brandt et al. (2016) found that in the U.S. 5% of the wells contribute over 50% of the total leakage volume of CH_4. These emission outliers (so-called "super-emitters") are poorly understood and not represented in emission inventories; 4) National and state regulations vary with respect to in situ emission control technologies (U.S. EPA, 2016a).

Most of the recent studies have focused on C_2H_6 because it has been proposed as a tracer for fugitive emissions from natural gas production (Schwietzke et al., 2014; Swarthout et al., 2013;

Vinciguerra et al., 2015) and there are multiple lines of evidence indicating that its abundance has increased over the Northern Hemisphere since 2009 (Franco et al., 2015; Helmig et al., 2016). In locations with multiple CH_4 sources (*e.g.* cows, oil and gas, rice production, wetlands), C_2H_6 can be used as a tracer for fossil fuel CH_4 emissions (McKain et al., 2015; Roscioli et al., 2015). Natural gas leakage contributes about ~60% of C_2H_6 emissions globally (Xiao et al., 2008), and up to 70% in regions with active oil and gas development (Gilman et al., 2013). Other important sources of C_2H_6 are biomass burning and biofuel consumption (domestic woodfuels), and each of these sources is estimated to individually account for ~20% of global emissions (Rudolph and Ehhalt, 1981; Singh and Zimmerman, 1992; Xiao et al., 2008; Zimmerman et al., 1988). Biogenic and oceanic emissions of C_2H_6 are considered negligible on a global scale (Plass-Dülmer et al., 1995; Rudolph, 1995; Zimmerman et al., 1988).

Ethane is one of the most abundant volatile organic compounds (VOC) in the atmosphere after CH_4. Observed C_2H_6 mixing ratios near the surface range from ~0.2 ppbv over remote regions of the Southern Hemisphere (Wofsy et al., 2012) and up to 1500 ppbv over oil and natural gas basins (Gilman et al., 2013; Helmig et al., 2014b; Thompson et al., 2014). The primary tropospheric sink of C_2H_6 is oxidation via reaction with hydroxyl radicals (OH). This loss pathway gives atmospheric C_2H_6 a strong seasonality and a seasonally dependent lifetime with a global annual average of ~2 months (Rudolph and Ehhalt, 1981). Based upon an approximate CH_4/C_2H_6 ratio of 2000 ppbv/2 ppbv and their relative reaction rates with OH, C_2H_6 can make an instantaneous contribution of 4-7% of the total OH loss for these two species combined (depending upon temperature and the specific enhancements encountered). Strong increases in C_2H_6 relative to CH_4 have been found in shale gas-producing areas such as the Bakken (Kort et al., 2016), and thus the contribution of C_2H_6 to OH reactivity may become more

important in the future. Other smaller tropospheric sinks of C_2H_6 include reaction with chlorine (Cl) radicals (Aikin et al., 1982; Sherwen et al., 2016), and loss via transport into the stratosphere (Rudolph, 1995). The relatively long lifetime of C_2H_6 allows it to be subject to long-range transport and to be relatively well mixed in the troposphere within each hemisphere. Since most of the anthropogenic C_2H_6 sources are concentrated in the Northern Hemisphere, and its lifetime is shorter than the inter-hemispheric exchange rate, there is a strong hemispheric gradient in C_2H_6 (Aydin et al., 2011; Helmig et al., 2016; Pozzer et al., 2010; Rudolph, 1995; Simpson et al., 2012).

Global C_2H_6 emissions have significantly changed over the last century. The recent literature is summarized in Figure 1.1. Briefly, measurements in firn air from Greenland and Antarctica show rising concentrations of C_2H_6 starting in the 1900s and peaking in the 1970s, followed by a decrease that lasted until the late 2000s (Aydin et al., 2011; Helmig et al., 2014a). The decrease in C_2H_6 between 1970 and 2006 observed by Aydin et al. (2011) was attributed to a reduction in fugitive emissions from the fossil fuel sector. Simpson et al. (2012) observed the same decreasing trend from surface flask measurements and found a strong correlation between global average C_2H_6 mixing ratios and CH_4 growth rates from 1985-2010, suggesting that these light alkanes have a common source.

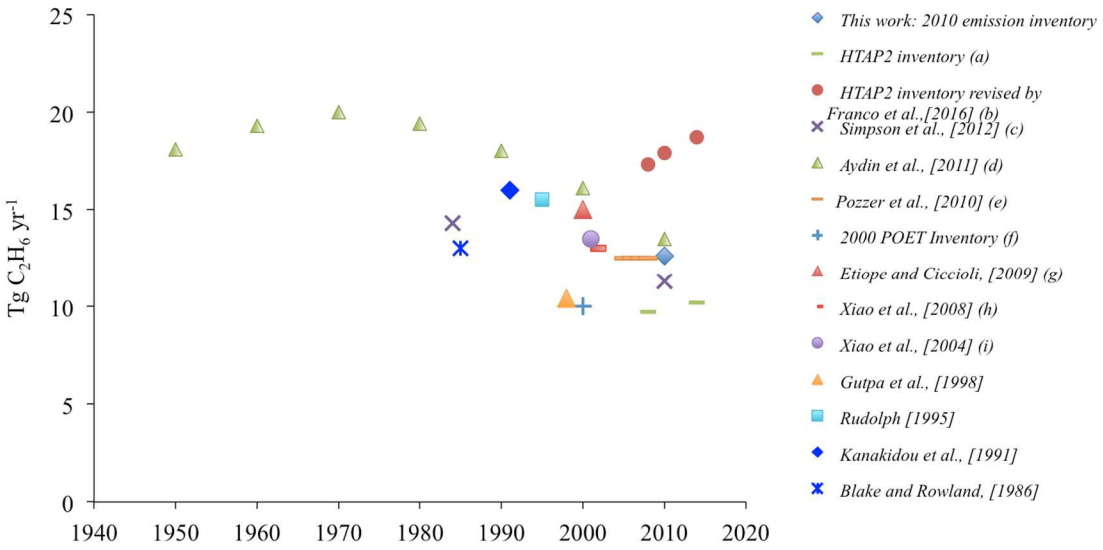


Figure 1.1: Estimated global annual emissions of C_2H_6 in Tg yr^{-1}. Total emissions correspond to the base year of each global estimate, if any, otherwise they correspond to the year they were published. (a) HTAP2 inventory for 2008 and 2014 as reported by Franco et al. (2016). Total emissions from anthropogenic sources (7.5 Tg yr^{-1}), biomass burning (1.8-2.3 Tg yr^{-1}), and biogenic (0.4 Tg yr^{-1}). (b) HTAP2 global anthropogenic emissions for 2008 were doubled for all years prior to 2009, with increasing North American emissions after 2009. (c) Total emissions from fossil fuels (8.0–9.2 Tg yr^{-1}), biofuels (2.6 Tg yr^{-1}), and biomass burning (2.4– 2.8 Tg yr^{-1}). (d) Emissions histories of total emissions from fossil fuels, biofuels and biomass burning. (e) Total emissions from anthropogenic (9.2 Tg yr^{-1}), biomass burning (2.8 Tg yr^{-1}), and oceanic (0.5 Tg yr^{-1}) sources. (f) POET inventory for 2000 as reported by Etiope and Ciccioli (2009). Total emissions from anthropogenic (5.7 Tg yr^{-1}), forest-savanna burning (2.6 Tg yr^{-1}), biogenic (0.8 Tg yr^{-1}), and ocean (0.8 Tg yr^{-1}). (g) Total emissions from POET Inventory base year 2000, with additional geologic emissions (2-4 Tg yr^{-1}). (h) Total emissions from fossil fuels (8.0 Tg yr^{-1}), biofuels (2.6 Tg yr^{-1}), and biomass burning (2.4 Tg yr^{-1}). (i) Total emissions reported by Xiao et al. (2008). (j) As reported by Gupta et al. (1998). (k) As reported by Rudolph (1995). (l) As reported by Kanakidou et al. (1991). (m) As reported by *Blake and Rowland* [1986].

There is evidence that the long-term decline in C_2H_6 in the Northern Hemisphere, recently reversed (Franco et al., 2015; Helmig et al., 2016). The change is postulated to be due to increased emissions tied to the recent growth of shale gas exploration and development in the U.S. (Franco et al., 2016; Helmig et al., 2016). Helmig et al. (2016) estimate a mean C_2H_6 annual emission increase of 0.42 ± 0.19 Tg yr^{-1} between 2009 and 2014 in the Northern Hemisphere, corresponding to an overall 2.1 ± 1.0 Tg yr^{-1} increase of C_2H_6 emissions for the same period. Franco et al. (2015) report a sharp increase (4.90 ± 0.91 % yr^{-1}) in measurements of C_2H_6

columns (molecules cm^{-2}) over the Jungfraujoch site in the Swiss Alps between 2009 and 2014. Vinciguerra et al. (2015) also showed a ~25% increase (1.1 ppbv) in hourly mean C_2H_6 surface mixing ratios from 2004 to 2013 at different sites downwind of the Marcellus shale play, one of the largest natural gas producing regions in the U.S. Several recent field measurement campaigns over U.S. natural gas basins have reported very high average mixing ratios of C_2H_6 (up to 300 ± 169 ppbv (1σ) (Koss et al., 2015)), along with other VOCs (Gilman et al., 2013; Helmig et al., 2014b; Katzenstein et al., 2003; Pekney et al., 2014; Pétron et al., 2012; Swarthout et al., 2013; Thompson et al., 2014), and several studies have found that C_2H_6 is the quantitatively largest non-methane VOC emitted during oil and natural gas exploitation (Field et al., 2015; Kort et al., 2016; Vinciguerra et al., 2015; Warneke et al., 2014).

Xiao et al. (2008) presented a 2001 global budget for C_2H_6 based on CH_4 emission estimates. They considered the geographical distributions of natural gas production based on production statistics and locations of major oil and gas wells compiled by Fung et al. (1991) and compared their results to a suite of observations collected prior to 2004. Therefore, this inventory is expected to be outdated, at least for North America, where the majority of the oil and gas development has occurred since 2004. Though we do not focus on it here, the Hemispheric Transport of Air Pollutants, Phase II (HTAP2) is also likely outdated as it requires an annual additional 1.2 Tg C_2H_6 emissions from North American sources in 2014 over 2008 emission rates to match C_2H_6 column observations (Franco et al., 2016). Note that Franco et al. (2016) applied that scaling uniformly without focusing on particular geographic regions.

Although at a national level anthropogenic VOC emissions decreased 11% from 2002 to 2011, VOC emissions from the U.S. oil and gas sector increased by 400% over the same period (Allen, 2016). Changes in isomeric ratios of C_4-C_5 alkanes observed over the continental U.S.

between 2001 and 2015 suggest an increasing influence of emissions from the oil and gas sector in this region (Rossabi and Helmig, 2018). However, despite the high emission rates and the known atmospheric impacts of C_2-C_5 alkanes inside U.S. oil and gas basins, relatively few studies have examined the contribution of these species to ozone (O_3) production. The total VOC contribution to O_3 mixing ratios highly depends on meteorological, seasonal, and geographical conditions, as well as regional transport and local VOC and NO_x sources, thus the estimates cannot be generalized or extrapolated to every oil and gas basin.

More than one third of the U.S. population currently live in counties that do not meet the 8-hour National Ambient Air Quality Standard (NAAQS) for O_3 (U.S. EPA, 2010). Although U.S. O_3 levels nationwide have been decreasing since 1980, from 2010 to 2017 the trend flattened out (U.S. EPA, 2018). Cheadle et al. (2017) estimated that oil and gas O_3 precursors contribute up to 30 ppb to summertime O_3 mixing ratios on individual days in the Northern Front Range of Colorado, where oil and gas development is the primary VOC source by mass (Eisele et al., 2009; Gilman et al., 2013). In this region, oil and gas alkane emissions contribute on average 20% to regional photochemical O_3 production and ~50% to the regional VOC OH reactivity during summer months (McDuffie et al., 2016). On high O_3 days, 30-40% of the ozone production can be attributed to the oil and gas sector (Pfister et al., 2017). Helmig et al. (2016) estimated that the current increase in C_2-C_5 alkanes inside and downwind U.S. oil and gas-producing regions can enhance average surface O_3 by 0.5 ppb yr^{-1} during summertime.

Overview of Chapters in this dissertation

Chapter 2 of this dissertation is motivated by the apparent dynamic nature of C_2H_6 mixing rations and the plausible use of C_2H_6 as a tracer for CH_4 leakage from the fossil fuel industry. In this chapter, we present a 2010 C_2H_6 emission inventory (beginning of the increasing trend in the

abundance of C_2H_6 reported by Franco et al. (2015)) and evaluate the differences between this and a previous 2001 C_2H_6 global emission inventory. To estimate C_2H_6 emissions for the year 2010 outside the U.S., we use a similar approach to Xiao et al. (2008), but based on CH_4 emissions derived from 2010 space-borne CH_4 observations from the Greenhouse Gases Observing SATellite (GOSAT), and we combined this with adjusted C_2H_6 emissions from the U.S. National Emission Inventory version 1 (2011NEIv1). We implement the emission inventories into the GEOS-Chem chemical transport model and compare the C_2H_6 simulation to a global suite of surface air observations, column measurements, and aircraft profiles.

Chapter 3 examines C_2-C_5 alkane emissions from the most recently updated 2011 National Emission Inventory (NEI), which includes updates over important oil-and-gas-producing basins and revised speciation profiles. We use those emissions to estimate the contribution to atmospheric abundances of C_2-C_5 alkanes over the U.S. from this industry. Also, we compare abundances of C_2-C_5 to a suite of surface observations, column measurements, and aircraft profiles. There have been several modeling studies that have begun to explore this issue (Kort et al., 2016; Thompson et al., 2017). The present work is an important addition to the existing literature because 1) we examine multiple species in tandem, 2) we performed higher resolution simulations over larger periods of time, and 3) we take a national scale perspective.

Chapter 4 investigates the contribution of C_2-C_5 alkane emissions from the U.S. oil and gas industry to O_3 abundances at regional and global scales. In this chapter, we use a similar set of simulations as the ones developed in Chapter 3. Again, there have been several modeling studies that have begun to explore the impact of C_2-C_5 alkanes on O_3 production on a variety of spatial and temporal scales (Helmig et al., 2016; Kort et al., 2016; Thompson et al., 2017). However, this work improves our understanding by 1) using an updated national emission

inventory for the U.S., and 2) improved global emission estimates of C_2H_6 based on the work presented in Chapter 2. The last Chapter summarizes our findings from the 3 previous Chapters and points out remaining challenges associated with quantifying the impact of emissions of C_2-C_5 from the U.S. oil and gas sector on atmospheric composition.

CHAPTER 2. REVISITING GLOBAL FOSSIL FUEL AND BIOFUEL EMISSIONS OF ETHANE[2]

2.1 GEOS-Chem Model description and configuration

We use the 3D chemical transport model (CTM) GEOS-Chem version 10-01 with tropospheric chemistry driven by GEOS-5 assimilated meteorological fields, from the Goddard Earth Observing System (GEOS) of the NASA Global Modeling and Assimilation Office (GMAO) (Bey et al., 2001). This model version includes the Harvard-NASA Emissions Component (HEMCO) version 1.1.005. HEMCO is a stand-alone software component for computing emissions from different sources, regions, and species on a user-defined grid that gives the user the opportunity to combine, overlay, and update a set of data inventories and scale factors (Keller et al., 2014). Our analysis is based on a 2°x2.5° resolution simulation for 2010. We found that concentrations at the surface were highly sensible to the spin-up time. A smaller spin-up time of 12 months produced mixing rations for the first month on average 0.8 ppbv lower compared to a 18 month spin up over the northern hemisphere. Therefore, we performed an 18 month spin-up for all simulations in this dissertation. The GEOS-Chem NO_x-O_x-HC-Aer-Br chemistry mechanism includes tropospheric C_2H_6 loss via reaction with OH, Br, and NO_3, with rate constants of $7.66 \times 10^{-12} \exp(-1020/T)$ cm^3 $molecule^{-1}$ s^{-1} (Sander et al., 2011), $2.36 \times 10^{-10} \exp(-6411/T)$ cm^3 $molecule^{-1}$ s^{-1} (Parrella et al., 2012) and 1.4×10^{-18} cm^3 $molecule^{-1}$ s^{-1}, respectively. The reaction rate with NO_3 is slow and is considered unimportant for the lifetime of C_2H_6 (Atkinson, 1991; Atkinson et al., 2006; Calvert et al., 2008). Stratospheric removal of C_2H_6 by Cl is not considered in our simulation, since past studies have estimated it to account for only

[2] This chapter contains published work from: Tzompa-Sosa, Z. A., et al. (2017), Revisiting global fossil fuel and biofuel emissions of ethane, Journal of Geophysical Research: Atmospheres, 122, doi:10.1002/2016JD025767.

~2% of total global loss (Gupta et al., 1998). The annual mass-weighted mean OH concentration of $8.5x10^5$ molecules cm^{-3} in our GEOS-Chem simulation, yields a global tropospheric (> 100 hPa) annual mean lifetime for C_2H_6 of 93 days. In the boundary layer (> 868 hPa), we estimated averaged lifetimes of 67 days globally, 41 days over the tropics (23°N-23°S) and 105 days in the mid- to high latitudes (23°-66°N, 23°-66°S). Based on the analysis in Naik et al. (2013) for other models, our global mean OH abundance of $8.5x10^5$ molecules cm^{-3} would approximately produce CH_4 and methyl chloroform (CH_3CCl_3) lifetimes of ~11.6 and ~6.7 years respectively. Both lifetime values are consistent with observation-derived lifetime estimates from Prinn et al. (2005) and Prather et al. (2012) which range from 10 .2 – 11.2 years for CH_4, and 6.0 – 6.3 years for CH_3CCl_3.

The public release version of GEOS-Chem v10-01 (used here) does not include tropospheric halogens other than Br, and this is a source of uncertainty in the following analysis. Though prior studies have shown Cl to be a minor sink for C_2H_6 (Gupta et al., 1998), in a very recent paper Sherwen et al. (2016) concludes that Cl may be an important C_2H_6 sink that can decrease the simulated global burden of C_2H_6 by about ~20%. The lifetime of C_2H_6 is very sensitive to simulated OH, and thus the interpretation of model-measurement comparisons is always limited by our ability to adequately represent the emissions of other trace gases that compete for reaction with OH.

We use Global Fire Emissions Database Version 3 (GFED3) biomass-burning emissions of C_2H_6 (van der Werf et al., 2010) in the simulations presented in this Chapter. The GFED3 emission inventory is based on global satellite-derived burned area information from the MODerate resolution Imaging Spectroradiometer (MODIS) sensor. At a global scale, the estimated uncertainty for biomass burning carbon emissions is around 20% (van der Werf et al.,

2010). GFED3 does not account for many small fires; this may be particularly relevant in the southeastern U.S. during time periods or locations with significant agricultural/prescribed burning (Randerson et al., 2012). There is interannual variability in the emissions of C_2H_6 from fires globally and over the U.S. (23-50°N, -130 -60°W). We compared emissions during 11 years (2001-2010), and found that averaged biomass burning C_2H_6 emissions from GFED3 are 2.1±0.35 (1σ) Tg/yr and 0.011±0.0049 (1σ) Tg/yr globally and over the U.S., respectively. During 2001-2011, global C_2H_6 emissions from biomass burning were highest in 2010; however, over the U.S., C_2H_6 emissions were equal to the average emissions for this period.

A detailed description of fossil fuel and biofuel C_2H_6 emissions in our simulations is discussed in section 2.3.3. For emissions of other species such as CO, NO, SO_X, and other VOCs, we use global emission inventories (HTAP v2, Emissions Database for Global Atmospheric Research inventory version 4.2 - EDGAR v4.2) overwritten by available regional emission inventories for Asia, Canada, Europe, Mexico, and the U.S. The composite of emission inventories corresponds to the public release version of GEOS-Chem v10-01.

We present updated anthropogenic (fossil fuel and biofuel) emissions of C_2H_6 for the year 2010 and compare them to a previous C_2H_6 emission inventory for the year 2001. We also compare the C_2H_6 model simulations based on both emission inventories to a global suite of observations. Our goal is to showcase the differences in anthropogenic emission totals and geographical distributions that are borne out by using different inventories at different points in time. Lastly, we document the impact of C_2H_6 on 2010 simulated atmospheric abundances of O_3 and PAN.

2.2 Global observations

We compare model simulations to an exhaustive database of recent C_2H_6 observations at the surface (2010-2011) and airborne campaigns (2008-2014). All observations are summarized in Table 2.1 and the regions of interest are depicted in Figure 2.1. We include surface flask measurements made at the Institute of Artic and Alpine Research (INSTAAR) Global Monitoring Program from samples collected by the National Oceanic and Atmospheric Administration (NOAA) Global Greenhouse Gas Reference Network (GGGRN) (http://instaar.colorado.edu/arl/Global_VOC.html), C_2H_6 column measurements derived from ground-based Fourier transform infrared (FTIR) solar observations from the Network for the Detection of Atmospheric Composition Change (NDACC, http://www.ndsc.ncep.noaa.gov/), and data from recent aircraft campaigns including the Arctic Research of the Composition of the Troposphere from Aircraft and Satellites (ARCTAS) (Simpson et al., 2011; Simpson et al., 2010), the Hiaper Pole-to-Pole (HIPPO) campaign (Wofsy et al., 2012), the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) (Blake et al., 2014; Schauffler, 2014), the 2014 Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) (Yacovitch and Herndon, 2014) campaign, and the Front Range Air Pollution and Photochemistry Éxperiment (FRAPPÉ) (Richter et al., 2015). We also include reported surface measurements from the Boulder Atmospheric Observatory (BAO) (Gilman et al., 2013; Swarthout et al., 2013), and data from 43 Chinese cities (Barletta et al., 2005).

Table 2.1: C_2H_6 observations from surface sites and airborne campaigns used to evaluate the model.

Aircraft measurements				
Figure 2.1 Region #	Mission	Location	Period	Reference
1	ARCTAS	40°-180°W, 32°-90° N	Apr, Jun-Jul, 2008	Simpson et al. (2010) Simpson et al. (2011)
2	HIPPO	150° E-84° W, 80° N-67° S	Jan, Oct-Nov, 2009, Mar-Apr, 2010, Jun-Sep, 2011	Wofsy et al. (2012)
3	SEACR4S	80°-126° W, 19°-50° N	Aug-Sep, 2013	Blake et al. (2014) Schauffler (2014)
4	DISCOVER-AQ	103°-105° W, 38°-42° N	Jul-Aug, 2014	Yacovitch and Herndon (2014)
5	FRAPPÉ	101°-109° W, 38°-42° N	Jul-Aug, 2014	Richter et al. (2015)

2010 Column measurements from the NDACC Network				
Code	Site	Location	Altitude (masl)	Reference
North America				
TAO	Toronto, Canada	112° W, 32° N	2,158	Wiacek et al. (2007)
BLD	Boulder, Colorado, United States	69° W, 77° N	30	Hannigan et al. (2009)
Europe				
JFJ	Jungfraujoch, Switzerland	8° W, 47° N	3,580	Franco et al. (2015)
KRN	Kiruna, Sweden	20° E, 68° N	419	Blumenstock et al. (2009) Kohlhepp et al. (2011)
North Africa				
IZO	Izaña, Tenerife, Spain	16° W, 28° N	2,367	García et al. (2012) Schneider et al. (2010)

2010 Surface flask measurements from the NOAA/INSTAAR Global VOC Monitoring Program			
Code	Site	Location	Altitude (masl)
North America			
ALT	Alert, Nunavut, Canada	62.51° W, 82.45° N	205
BMW	Tudor Hill, Bermuda, United Kingdom	64.88° W, 32.26° N	60
BRW	Barrow, Alaska, United States	156.61° W, 71.32° N	16
CBA	Cold Bay, Alaska, United States	162.72° W, 55.21° N	57
KEY	Key Biscayne, Florida, United States	80.16° W, 25.67° N	6
LEF	Park Falls, Wisconsin, United States	90.27° W, 45.95° N	868
MID	Sand Island, Midway, United States	177.38° W, 28.21° N	15
SUM	Summit, Greenland	38.42° W, 72.6° N	3,215
THD	Trinidad Head, California, United States	124.15° W, 41.05° N	112
UTA	Wendover, Utah, United States	113.72° W, 39.9° N	1,332
Europe			
HPB	Hohenpeissenberg, Germany	11.02° E, 47.8° N	941
ICE	Storhofdi, Vestmannaeyjar, Iceland	20.29° W, 63.4° N	127
MHD	Mace Head, County Galway, Ireland	9.9° W, 53.33° N	26
OXK	Ochsenkopf, Germany	11.81° E, 50.03° N	1,172
PAL	Pallas-Sammaltunturi, GAW Station, Finland	24.12° E, 67.97° N	565
ZEP	Ny-Alesund, Svalbard, Norway and Sweden	11.89° E, 78.91° N	479
East Asia			
SHM	Shemya Island, Alaska, United States	174.13° E, 52.71° N	28

2010 Surface flask measurements from the NOAA/INSTAAR Global VOC Monitoring Program			
Code	Site	Location	Altitude (masl)
TAP	Tae-ahn Peninsula, Republic of Korea	126.13° E, 36.74° N	21
Central America			
KUM	Cape Kumukahi, Hawaii, United States	154.82° W, 19.52° N	8
MEX	High Altitude Global Climate Observation Center, Mexico	97.31° W, 18.98° N	4,469
MLO	Mauna Loa, Hawaii, United States	155.58° W, 19.54° N	3,402
North Africa			
IZO	Izana, Tenerife, Canary Islands, Spain	16.5° W, 28.31° N	2,378
ASK	Assekrem, Algeria	5.63° E, 23.26° N	2,715
South Asia			
GMI	Mariana Islands, Guam	144.66° E, 13.39° N	5
Australia			
BKT	Bukit Kototabang, Indonesia	100.32° E, 0.2° S	850
CGO	Cape Grim, Tasmania, Australia	144.69° E, 40.68° S	164
South Africa			
ASC	Ascension Island, United Kingdom	14.4° W, 7.97° S	90
CRZ	Crozet Island, France	51.85° E, 46.43° S	202
HBA	Halley Station, Antarctica, United Kingdom	26.21° W, 75.61° S	35
MKN	Mt. Kenya, Kenya	37.3° E, 0.06° S	3,649
SEY	Mahe Island, Seychelles	55.53° E, 4.68° S	6
SYO	Syowa Station, Antarctica, Japan	39.58° E, 69° S	3
SPO	South Pole, Antarctica, United States	24.8° W, 89.98° S	2,815
South America			
EIC	Easter Island, Chile	109.43° W, 27.16° S	69
PSA	Palmer Station, Antarctica, United States	64° W, 64.92° S	15
SMO	Tutuila, American Samoa	170.56° W, 14.25° S	60
TDF	Tierra Del Fuego, Ushuaia, Argentina	68.31° W, 54.85° S	32

Surface observations				
Code / Figure 2.1 Region #	Site	Location	Period	Reference
BAO	Boulder Atmospheric Observatory	105.01° W, 40.05° N	Feb-Mar, 2011	Gilman et al. (2013)
BAO	Boulder Atmospheric Observatory	105.01° W, 40.05° N	Feb-Mar, 2011	Swarthout et al. (2013)
6	43 Chinese cities averaged horizontally every 20°x10° (longitude, latitude)	100°-130°E, 20°-45° N	Jan – Feb 2001	Barletta et al. (2005)

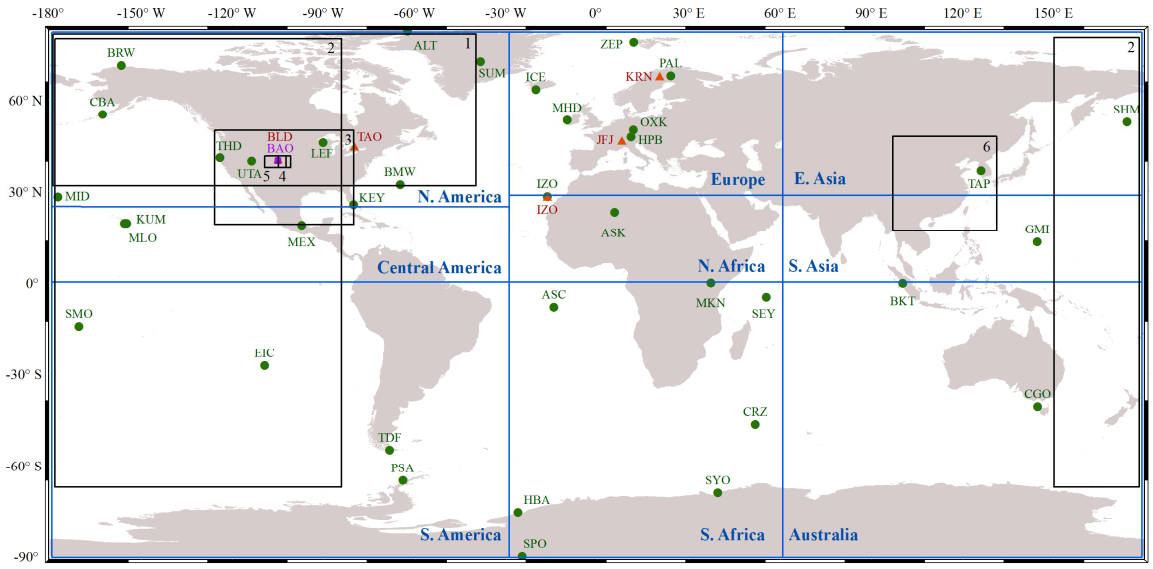


Figure 2.1: Regions for C_2H_6 emissions analysis and locations of C_2H_6 observations. Black boxes cover regions of aircraft measurements, green circles represent surface flask measurements, orange triangles locate C_2H_6 column measurements and the purple square shows BAO surface measurements. Regions delimited to calculate C_2H_6 emissions presented on Table 2.1 are encompassed by blue boxes.

2.3 Global C_2H_6 Emission Inventories

2.3.1 2001 C_2H_6 emission inventory

Prior to this work, the most recent global C_2H_6 emission inventory implemented in GEOS-Chem model version 10-01 was based on the year 2001 (Xiao et al., 2008). The model sets this C_2H_6 inventory as default for any simulation. Briefly, this inventory is derived from a previous C_2H_6 emission inventory by Xiao et al. (2004), which scales C_2H_6 emissions to CH_4 fossil fuel sources using fixed regional ratios, and bases the geographical distribution for the emissions on data from 1978-1986 (Fung et al., 1991; Wang et al., 2004). Major changes to the distribution of fossil fuel sources may have occurred globally during the period from which they draw data for the model evaluation. Xiao et al. (2008) estimate global C_2H_6 emissions for three different source types: 1) fossil fuel, 2) biofuel (domestic wood fuels), and 3) biomass burning. However, the only global C_2H_6 emission inventory from Xiao et al. (2008) implemented in GEOS-Chem

version 10-01 is from fossil fuel sources. In the model, the C_2H_6 emission fluxes from the fossil fuel inventory from Xiao et al. (2008) have no seasonality, and no scaling factors are available to scale them to other years. To simulate global biofuel sources, we use the biofuel C_2H_6 emission inventory derived by Yevich and Logan (2003) and the GEFD3 emission inventory (van der Werf et al., 2010) for biomass burning C_2H_6 emissions.

2.3.2 2010 C_2H_6 emission inventory

Global C_2H_6 emissions

We develop an updated global C_2H_6 emission inventory for 2010, by scaling C_2H_6 to CH_4 emissions following a similar approach to previous studies (Blake and Rowland, 1986; Etiope and Ciccioli, 2009; Franco et al., 2016; Rudolph, 1995; Xiao et al., 2004; Xiao et al., 2008). There are many approaches that can be used to estimate CH_4 emissions (*i.e.,* top-down studies, bottom-up models, inventories, and data-driven approaches), and differing approaches can yield different emission totals, attribution, or geographical distributions (Saunois et al., 2016). In this study, the CH_4 fluxes were derived from the Greenhouse Gases Observing SATellite (GOSAT) by Turner et al. (2015) for the year 2010. To derive anthropogenic CH_4 emissions, Turner et al. (2015) used a priori emissions from EDGARv4.2 (http://edgar.jrc.ec.europa.eu/). The EDGAR emission inventory combines Tier 1 and region-specific Tier 2 emission factors, which have multiple uncertainties associated with them. A detailed description of these uncertainties is presented by Olivier (2002). The estimated uncertainty of satellite CH_4 single-retrievals is 0.8% (Parker et al., 2011). A description of the error characterization and the uncertainties associated with the North American CH_4 inversions can be found in Turner and Jacob (2015). Turner et al. (2015) infer a 2009-2011 U.S. anthropogenic emission source of 40.2 – 42.7 Tg a^{-1}, and attribute

22-31% to oil and gas activities. Other inverse studies have inferred a larger range of anthropogenic emissions (30.0 – 44.5 Tg a^{-1}) [see Turner et al. (2015) and references within]. It is important to note that over regions with CH_4 emissions from oil and natural gas activities and livestock, the source attribution is very sensitive to assumptions made in the prior distribution. Uncertainties associated with the CH_4 emissions, or their attribution, is only one of several sources of uncertainty in using CH_4 fluxes to estimate C_2H_6 fluxes. As we discuss later in this section, a second major issue is the choice of C_2H_6/CH_4 emission ratio.

We implement two grid resolutions for the Turner et al. (2015) CH_4 fluxes for the year 2010. For North America, we use CH_4 emission fluxes at 1/2°x2/3° resolution, and at 4°x5° resolution for the rest of the world. Considering the uncertainties in the attribution of fluxes, we expect a better agreement of CH_4 anthropogenic sources at a coarse resolution compared to the finer resolution. We can have the most confidence in the total fluxes, rather than fine sectorial attribution. From the 12 anthropogenic CH_4 source categories derived in Turner et al. (2015), three are relevant to C_2H_6: natural gas activity, biofuel usage, and biomass burning.

We consider natural gas activity and biofuel source categories and retained the GFED3 emission inventory for emissions of C_2H_6 from biomass burning during 2010. We treated biofuel consumption (both from home cooking and heating) as residential biomass burning, and thus applied a temperate forest fuel ratio of 15.2 (% mol C_2H_6/ mol CH_4) as estimated by Akagi et al. (2011). To derive C_2H_6 emissions from CH_4 fluxes associated with natural gas activity, we used a ratio of 4.3 (% mol C_2H_6/ mol CH_4) based on mixing ratio enhancements estimated from the South Central U.S. by Katzenstein et al. (2003). Warneke et al. (2014) observed similar emission ratios during wintertime 2012 over the Uintah Basin. In this study, we assume observed enhancement ratios (slopes of the linear fits) are approximately equivalent to emission ratios

since C_2H_6 is a relatively long-lived species and in situ measurements are taken close to the sources.

Constraints on C_2H_6 Emissions over Mexico and Asia

An analysis of the resulting global C_2H_6 emissions immediately points to likely problems with the underlying CH_4 fluxes or the 4.3 (% mol C_2H_6/ mol CH_4) ratio over Mexico and Asia. Also, estimated fossil fuel C_2H_6 emission totals derived from CH_4 fluxes over Mexico are two times higher (0.23 Tg) than the 2001 C_2H_6 emission inventory. Similar differences of 0.36 Tg and 0.13 Tg occur when comparing to RETRO (REanalysis of the TROposhperic chemical composition 2000, http://gcmd.gsfc.nasa.gov/records/GCMD_GEIA_RETRO.html) and GEIA (Global Emissions InitiAtive 1985, http://www.geiacenter.org) emission inventories, respectively. Additionally, when analyzing the spatial distribution of fossil fuel C_2H_6 emissions over Mexico derived from CH_4 fluxes, we find that the C_2H_6 emission sources are located away from oil and natural gas production areas. Second, total fossil fuel satellite-derived emissions of C_2H_6 over Asia are half (~1.2 Tg) of the 2001 C_2H_6 emission inventory and RETRO, respectively. A simulation with these emissions produces C_2H_6 mixing ratios that are 1/6 of observed mixing ratios during wintertime in 2001 by Barletta et al. (2005) (note also the time difference between these in situ observations (2001) and the inversion (2010)). Finally, a comparison between the spatial distribution of fossil fuel C_2H_6 emissions over China from the 2010 C_2H_6 emission inventory and the emissions derived from CH_4 fluxes shows that C_2H_6 emissions from CH_4 fluxes are clustered in south central China, while the Xiao et al. (2008) C_2H_6 emissions distribution covers urban and known oil and natural gas-producing regions in China. In summary, there is evidence that scaling C_2H_6 emissions derived from CH_4 fluxes does not produce realistic C_2H_6 emission over Mexico and China. In order to address the two regional

discrepancies above, we substitute the C_2H_6 emissions derived from CH_4 fluxes with the Xiao et al. (2008) C_2H_6 emission inventory over Mexico and Asia (including: China, India, Indonesia, Japan, Mongolia, North and South Korea).

Constraints on C_2H_6 Emissions over the U.S.

The 2011NEIv1 emissions data are provided by state and local agencies based on industrial, commercial, and area sources. We incorporate 2011NEIv1 version 2 C_2H_6 emissions on a 0.1^o x 0.1^o grid for biofuel and six anthropogenic source categories, including oil and gas activities (U.S. EPA, 2013). GEOS-Chem version 10-01 uses a scaling factor of 1.016 to apply 2011NEIv1 C_2H_6 emissions to the year 2010. For other species such as CO, NO, and other VOCs, scaling factors are assigned based on government statistics and documents. For industrial emissions, the scaling factors are based on reported trends from the Environmental Protection Agency Acid Rain Program (https://www.epa.gov/airmarkets/acid-rain-program). For other emissions the scaling factors come from the National Emissions Inventory Air Pollutant Emissions Trends Data (https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data).

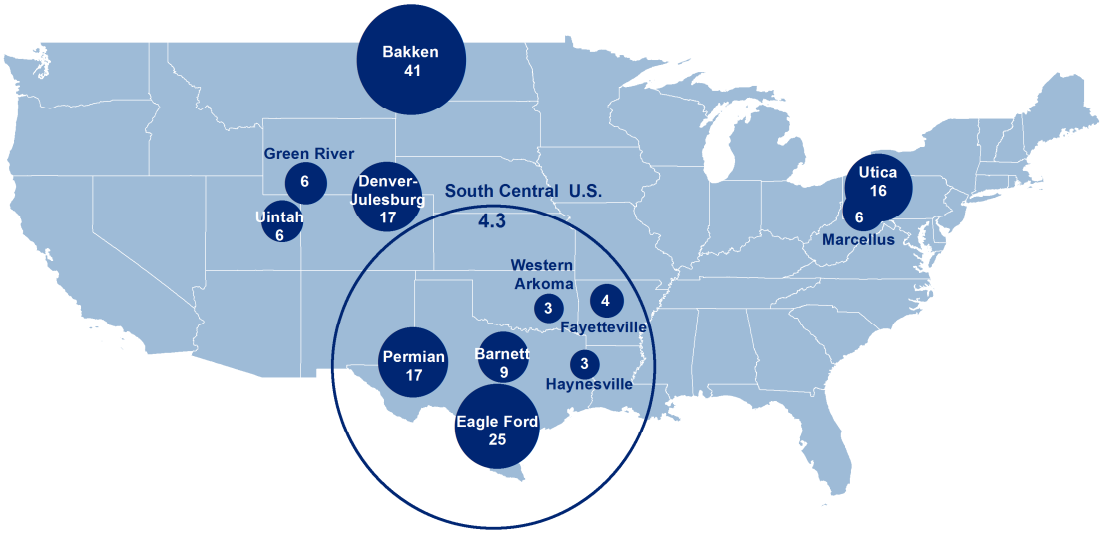
The 2011NEIv1 C_2H_6 emission sources appear to align with the distribution of active oil and natural wells over the U.S. (see Figure 2.4); however, when the GEOS-Chem simulated C_2H_6 is compared to aircraft measurements over the U.S. from five recent field campaigns (2008-2014) and 2010 surface flask observations from the NOAA GGGRN, the use of the 2011NEIv1 emissions produce mixing ratios at the surface and throughout the column that are 14-50 % of those observed. Consequently, we tested uniformly scaling C_2H_6 emissions from all the categories in the 2011NEIv1 by factors between 1.2 and 2, and we compared the results to observations. For all factors, a linear regression of 2010 monthly mean surface flask observations

over the U.S. versus model output yields coefficient of determination (R^2) values between 0.59 and 0.64. The slopes range from 0.8 to 1.0. Of the scaling factors tested, 1.4 produces the best agreement between the GEOS-Chem simulation and observations in regions *without* major oil and gas operations. Therefore, we multiplied 2011NEIv1 C_2H_6 emissions by 1.4, which represents an addition of 0.5 Tg of C_2H_6 compared to the base 2011NEIv1. Scaling beyond 1.4 results in an overestimate of observations in these regions. Following the adjustment of 2011NEIv1 C_2H_6 emissions, we refer to the resulting global C_2H_6 emission inventory as *2010 C_2H_6 emission inventory*. Thus, the 2010 C_2H_6 emission inventory combines a global C_2H_6 emission inventory derived from satellite CH_4 observations, except for Mexico and Asia where we apply previous emission estimates, and a regional C_2H_6 emission inventory derived by adjusting 2011NEIv1 C_2H_6 emissions.

Uncertainties

Though the approach of deriving C_2H_6 from CH_4 emissions is consistent with past global budget studies, large uncertainties are associated with the use of few C_2H_6/CH_4 emission ratios, especially for the natural gas industry, which in the last decade has been subject to multiple emission controls in many countries. Emission ratios depend on the type of oil and natural gas reservoir (*e.g.,* tight gas vs. shale gas), the VOC composition of the natural gas (Warneke et al., 2014), the production stage of a producing well (Kang et al., 2014; Pacsi et al., 2015), among other characteristics. There has been significant attention devoted to documenting C_2H_6 to CH_4 enhancement ratios. Given the lifetime of each species, enhancement ratios observed near defined sources are often a reasonable surrogate for emission ratios. Figure 2.2 presents a summary of averaged percentage molar C_2H_6/CH_4 ratios observed in different oil and natural gas basins over the contiguous U.S. Reported ratios (% mol C_2H_6/ mol CH_4) have a large range, for

example: Kort et al. (2016) report 40.5 for the Bakken, more than an order of magnitude larger than the ratio reported for some oil and gas basins in the central U.S. (Peischl et al., 2015b). There are a number of problems associated with basing C_2H_6 emissions on CH_4 emissions, and dynamic C_2H_6/CH_4 emission ratios. As we will show later, using a constant C_2H_6/CH_4 emission ratio over regions with high emission gradients (*e.g.,* U.S.) does not represent the geographical distributions of the emissions and the resulting atmospheric abundances of C_2H_6. Section 2.5.1 presents the sensitivity of our findings to the choice of C_2H_6 to CH_4 molar ratios through simulations with a fixed ratio applied broadly across the U.S. using the low and high ratios available from the recently published literature (Figure 2.2).



South Central U.S.:	Calculated using annual emissions of C_2H_6 and CH_4 reported by Katzenstein et al. (2003).
Bakken:	Brandt et al. (2015) as reported by Kort et al. (2016).
Barnett:	Speight (2013) as reported by Kort et al. (2016).
Denver-Julesburg:	Peischl et al. (2015a).
Eagle Ford:	Conder and Lawlor (2014) and Ghandi et al. (2015) as reported by Kort et al. (2016).
Fayetteville:	Average from Peischl et al. (2015b) and Speight (2013) as reported by Kort et al. (2016).
Green River:	Peischl et al. (2015a).
Haynesville:	Average from Peischl et al. (2015b) and Speight (2013) as reported by Kort et al. (2016).
Marcellus:	Average from Peischl et al. (2015b), 2009 U.S. Geological Survey (USGS) database as reported by Peischl et al. (2015b), and Conder and Lawlor (2014) as reported by Kort et al. (2016).
Permian:	Peischl et al. (2015a).
Western Arkoma:	Average from Peischl et al. (2015b), 2009 U.S. Geological Survey (USGS) database as reported by Peischl et al. (2015b).
Uintah:	Average from Helmig et al. (2014b) and Warneke et al. (2014).
Utica:	Conder and Lawlor (2014) and Ghandi et al. (2015) as reported by Kort et al. (2016).

Figure 2.2: Spatial distribution of averaged percentage molar C_2H_6/CH_4 ratios in oil and natural gas basins over the contiguous U.S. The values and sizes of the circles represent the magnitude of the ratios in each basin. South Central U.S.: calculated using annual emissions of C_2H_6 and CH_4 reported by Katzenstein et al. (2003). Bakken: Brandt et al. (2015) as reported by Kort et al. (2016). Barnett: Speight (2013) as reported by Kort et al. (2016). Denver-Julesburg: Peischl et al. (2015a). Eagle Ford: Conder and Lawlor (2014) and Ghandi et al. (2015) as reported by Kort et al. (2016). Fayetteville: average from Peischl et al. (2015b) and Speight (2013) as reported by Kort et al. (2016). Green River: Peischl et al. (2015a). Haynesville: average from Peischl et al. (2015b) and Speight (2013) as reported by Kort et al. (2016). Marcellus: average from Peischl et al. (2015b), 2009 U.S. Geological Survey (USGS) database as reported by Peischl et al. (2015b), and Conder and Lawlor (2014) as reported by Kort et al. (2016). Permian: Peischl et al. (2015a). Western Arkoma: average from Peischl et al. (2015b), 2009 U.S. Geological Survey (USGS) database as reported by Peischl et al. (2015b). Uintah: average from Helmig et al. (2014b) and Warneke et al. (2014). Utica: Conder and Lawlor (2014) and Ghandi et al. (2015) as reported by Kort et al. (2016).

2.3.3 Comparison between the 2001 and 2010 C_2H_6 emission inventories

Table 2.2 shows global and regional C_2H_6 emission estimates for both emission inventories. For the 2010 C_2H_6 emission inventory Northern Hemisphere fossil fuel sources represent half of global C_2H_6 emissions and 95% of global fossil fuel emissions.

The C_2H_6 emission totals are only subtly different between both global inventories; however, the spatial distributions of the emissions are quite distinct. In our recommended 2010 inventory, C_2H_6 emissions increase over intense oil and gas producing regions, including the central and northeastern U.S., Venezuela, eastern Russia, and the northern part of the Middle East (Figure 2.3). We point this out because it may indicate that emissions from the oil and natural gas industry in these regions could be important, but may not be accounted for in commonly used inventories. Over Europe, Xiao et al. (2008) concluded that their inventory overestimated observed C_2H_6 mixing ratios by 20-30%, and they attributed this in part to an overestimation of European sources. Our 2010 C_2H_6 emission inventory shows a similar reduction of C_2H_6 European sources (Table 2.2). Over the contiguous U.S., we find important differences in the geographical distribution and magnitude when comparing the fossil fuel C_2H_6 emission fluxes from the 2010 C_2H_6 emission inventory to the Xiao et al. (2008) 2001 emission inventory (Figure 2.4). Fossil fuel C_2H_6 emission fluxes are smaller over the northeastern part and larger over the central and south central part of the U.S.

Table 2.2: C_2H_6 emissions in Tg yr^{-1} by region for the 2001 and 2010 C_2H_6 emission inventories.

Region	Fossil Fuel (Tg yr^{-1})	Biofuel (Tg yr^{-1})	Biomass Burning (Tg yr^{-1})
2001 C_2H_6 emission Inventory			
Global	**7.9**	**2.5**	**2.7**
Northern Hemisphere	7.2	2.1	1.1
North America	1.9	<0.05	0.1
Europe	2.1	0.3	<0.05
East Asia	1.6	0.4	0.1
Central America	0.2	0.1	<0.05
North Africa	0.6	0.3	0.4
South Asia	0.8	1.0	0.4
Southern Hemisphere	0.7	0.4	1.7
Australia	0.3	0.1	<0.05
South Africa	0.2	0.2	0.7
South America	0.1	0.2	1.0
2010 C_2H_6 emission Inventory			
Global	**7.1**	**2.8**	**2.7**
Northern Hemisphere	6.7	2.4	1.1
North America	1.7	<0.05	0.1
Europe	1.6	0.4	<0.05
East Asia	1.9	0.4	0.1
Central America	0.4	0.1	<0.05
North Africa	0.4	0.4	0.4
South Asia	0.8	1.0	0.4
Southern Hemisphere	0.4	0.4	1.7
Australia	0.1	0.1	<0.05
South Africa	0.1	0.3	0.7
South America	0.2	0.1	1.0

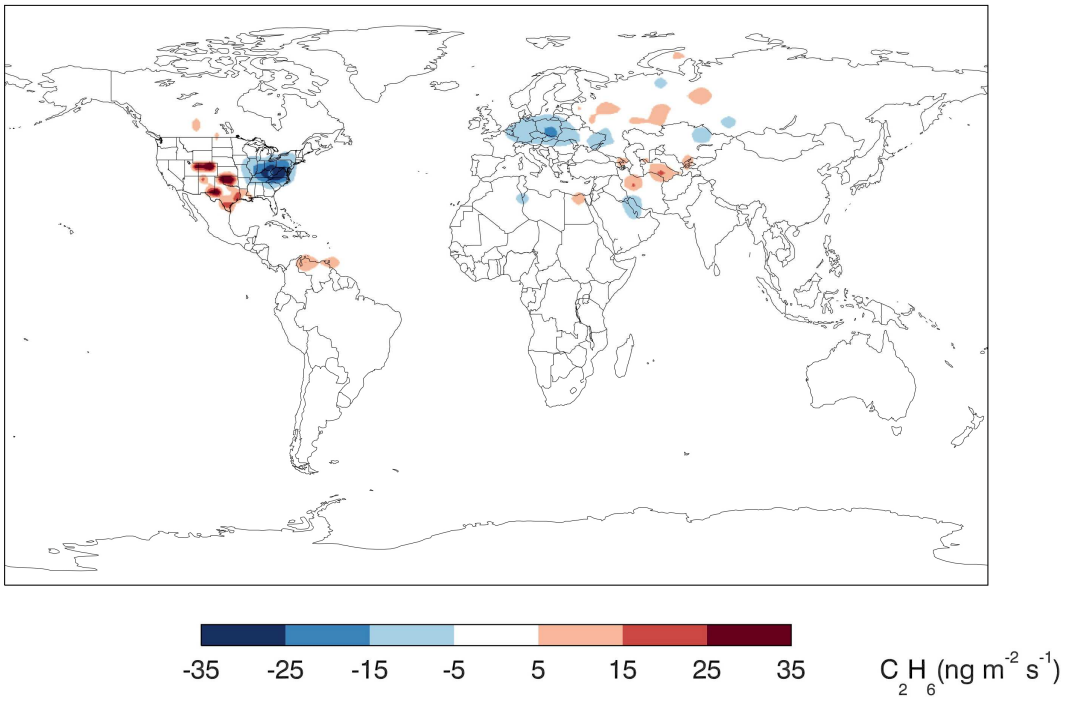


Figure 2.3: Global comparison between modeled distributions of fossil fuel C_2H_6 emissions for 2001 and 2010 C_2H_6 emission inventories (2010-2001). Positive values (warmer colors) represent increases in modeled annual mean C_2H_6 emission fluxes.

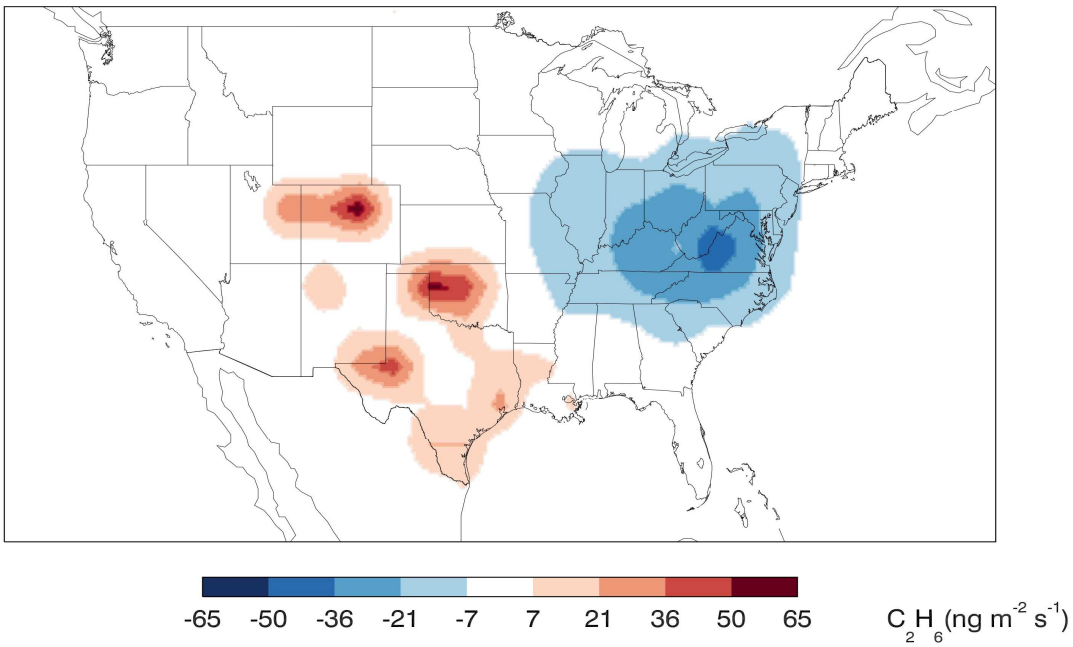


Figure 2.4: Comparison between modeled distributions of fossil fuel C_2H_6 emissions for 2001 and 2010 C_2H_6 emission inventories (2010-2001) over the U.S. Positive values (warmer colors) represent increases in modeled annual mean C_2H_6 emission fluxes.

The 2010 C_2H_6 emission inventory shows increased emission regions encompassing major U.S. natural gas production basins (Figure 2.5). The simulated surface C_2H_6 abundances produced by the 2010 C_2H_6 emission inventory closely align with oil and gas activity over the U.S. Although, 2010 C_2H_6 emissions show significant increases in fossil fuel C_2H_6 emissions over these regions, they continue to underestimate the most recent vertical and surface observations of C_2H_6 mixing ratios over the central U.S. as described in section 4. Despite the underestimation of C_2H_6 abundances over the central U.S., the 2010 C_2H_6 emission inventory produces a better geographical distribution of fossil fuel C_2H_6 sources over North American regions and elsewhere compared to the 2001 C_2H_6 emission inventory.

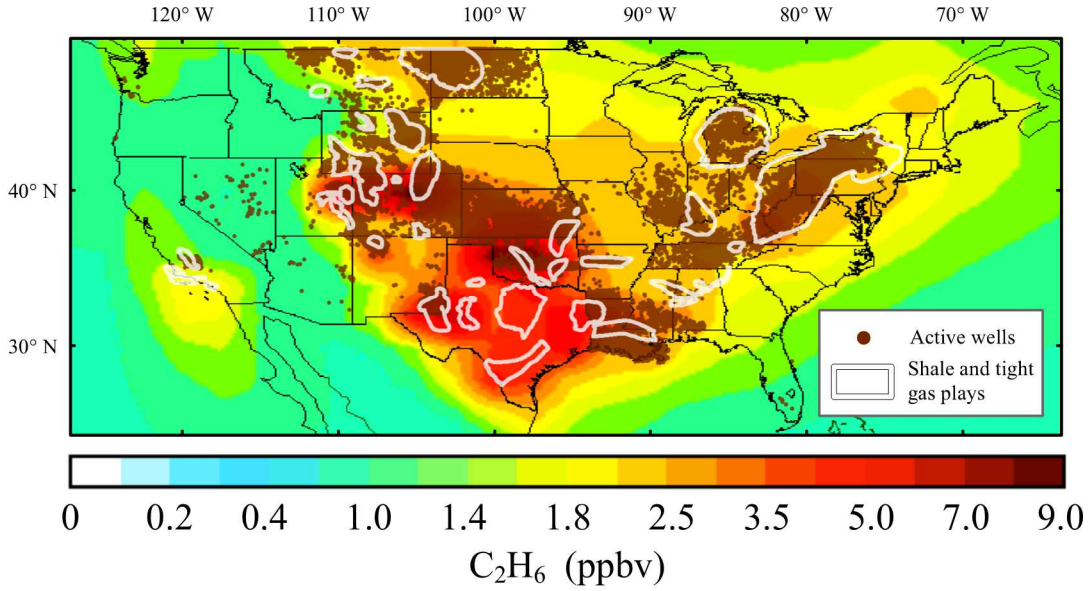


Figure 2.5: Modeled annual mean surface mixing ratios of the 2010 C_2H_6 emission inventory and spatial distribution of active wells (FracTracker, accessed Nov. 2015, www.fractracker.org; data for Maryland, North Carolina, and Texas are missing). Shale and tight gas plays (Energy Information Administration, accessed Dec. 2014, www.eia.gov/dnav/ng/ng_sum_lsum_a_EPG0_xdg_count_a.htm) are shown to provide a sense for well distribution over states where well location data is missing.

2.4 Model evaluation

2.4.1 Ground-based C_2H_6 column observations

Comparisons between each modeled emission inventory and monthly mean C_2H_6 total columns at selected NDACC stations over the Northern Hemisphere for both C_2H_6 emission inventories are shown in Figure 2.6. We note that observations over Northern Hemisphere continental regions such as Asia and the Middle East are needed to evaluate model outputs in other oil and natural gas producing regions.

Ethane columns are derived from ground-based FTIR solar observations following the methodology presented by Franco et al. (2015). The information content as well as the vertical sensitivity for the FTIR retrievals from all the sites we analyze in this paper is similar to the one presented in Franco et al. (2015). At altitudes below 13 km, 99 % of the information content is independent from the a priori profile; indicating a very good sensitivity to the true state of the atmosphere (Franco et al., 2015).

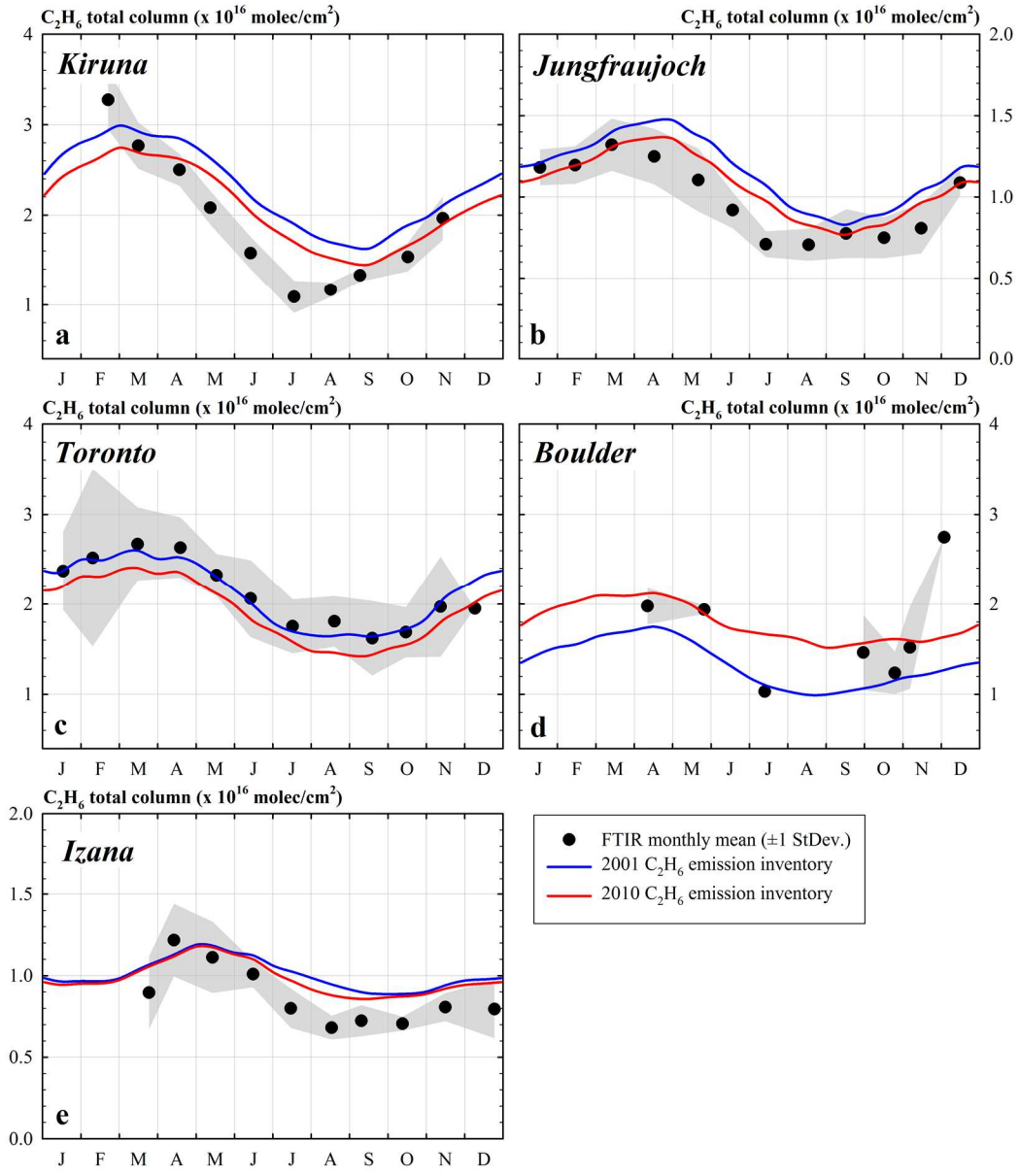


Figure 2.6: Comparison of 2010 C_2H_6 total columns to modeled 2001 and 2010 C_2H_6 emission inventories. Black dots represent monthly mean C_2H_6 total columns and grey areas denote their associated 1σ standard deviation. Lines represent modeled total columns for different emission inventories.

The FTIR total column observations largely reflect the tropospheric background, and not solely the surface C_2H_6 abundances. Thus, similarities between our two simulations can be expected, particularly since most FTIR stations involved here are not located in source regions

and C_2H_6 has a relative long lifetime that allows emissions to impact abundances at a hemispheric scale. Our model simulations largely reproduce C_2H_6 column observations at the selected stations, suggesting that the OH losses and emissions of other OH-reactive species are being well represented. Additionally, the spatial variability indicates that the model reproduces the major features of C_2H_6 emissions and the most important transport processes. The agreement with observations is particularly good at Toronto and Boulder (see Figure 2.9, explained on section 2.5.2). However, during summertime, GEOS-Chem overestimates the C_2H_6 column at three sites (Izana, Kiruna, and Jungfraujoch, Figure 2.6 a,b,e). A similar bias was also reported by Franco et al. (2016) over remote sites in the Arctic and the Tropics. The same version of the GEOS-Chem model we use here was also used in that study, and the summertime bias may be driven by ~10-15% lower global mean OH in GEOS-Chem version 10-01 compared to previous versions. Xiao et al. (2008) used GEOS-Chem version 6.01.03, and their results do not show a summertime bias compared to surface data. GEOS-Chem version 6.01.03 had a 15% greater C_2H_6-OH rate constant ($8.7 \times 10^{-12} \exp(-1020/T)$ cm^3 $molecule^{-1}$ s^{-1}, Sander et al. (2003)), and ~3% lower annual mean OH concentration (10.2×10^5 molecules cm^{-3}) compared to GEOS-Chem version 10-01 used in this study.

2.4.2 Surface observations

Figure 2.7 presents a comparison between 2010 Northern Hemisphere C_2H_6 surface mixing ratios and the 2001 and 2010 C_2H_6 emission inventories. The Southern Hemisphere comparison is presented in the supplemental material (Figure A1). In both figures, sampling stations are ordered from higher to lower latitudes. The NOAA sampling stations are mostly located at remote locations around the globe, largely avoiding the impact of local anthropogenic emissions.

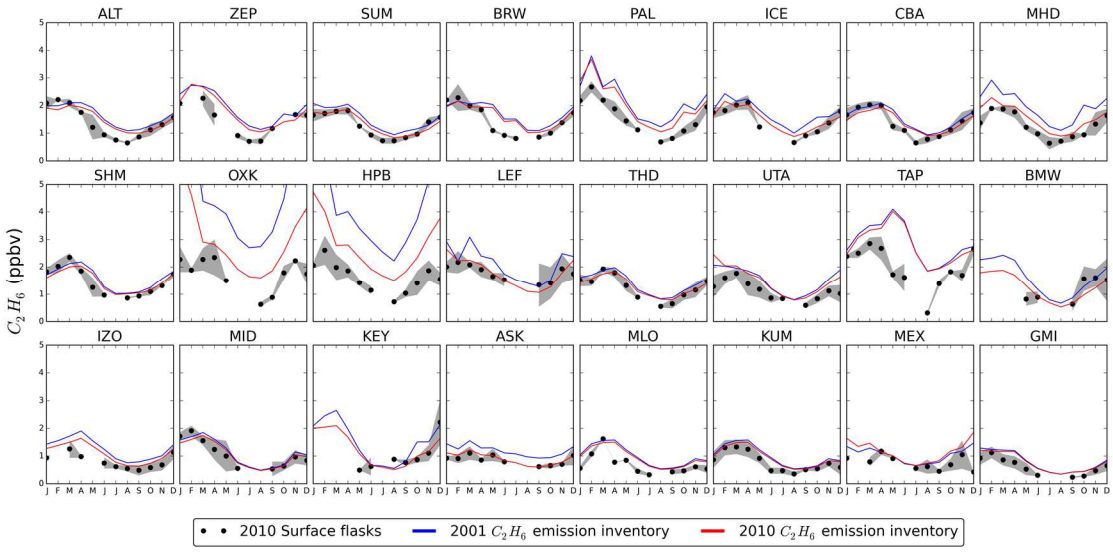


Figure 2.7: Comparison of Northern Hemisphere 2010 C_2H_6 surface mixing ratios to modeled 2001 and 2010 C_2H_6 emission inventories. Black dots represent C_2H_6 observations from NOAA GGGRN global surface flask network and grey areas denote their associated 1σ standard deviation. Lines represent model mixing ratios at the surface from both C_2H_6 emission inventories. Stations are ordered from higher to lower latitudes.

Mixing ratios at the surface reflect the C_2H_6 latitudinal gradient with values decreasing with decreasing latitude. Our model simulations largely capture the C_2H_6 seasonal cycles. However, the model overestimates surface mixing ratios over Europe (OXK and HPB stations) by ~160% and ~80% on average for the 2001 and 2010 C_2H_6 emission inventories, respectively. The difference between model outputs and observations might be due to an overestimation of European sources. The overestimation we find when simulating 2010 C_2H_6 emissions may be related to incorrect source attribution of European C_2H_6 anthropogenic sources derived from the EDGARv4.2 inventory used as a prior for the satellite-derived CH_4 fluxes. In Asia, the model over-predicts the C_2H_6 abundance at the TAP station located in the southern part of the Republic of Korea by ~1 ppbv throughout the year. TAP is likely heavily influenced by emissions from both Russia and China. However, a comparison with surface measurements in 43 Chinese cities (Barletta et al., 2005) shows that the model under predicts surface mixing ratios for the 2001

wintertime by up to a factor of 3. Given the limited and outdated observations in this region we cannot determine whether this is a result of incorrect emissions or an incorrect distribution of emissions. Compared to the rest of the observations we show, there is also a large temporal mismatch between the simulation year (2010) and the Chinese observations (2001). Although our analysis focuses on the Northern Hemisphere, we note that over the Southern Hemisphere the station CGO, located in Tasmania, Australia, is the only station (Figure A1) with large differences between simulations and observations. Specifically, 2001 C_2H_6 emissions produce mixing ratios at the surface almost twice as large as those observed. This difference can be attributed to lower C_2H_6 emissions derived from CH_4 satellite observations compared to the 2001 C_2H_6 emission inventory.

2.4.3 Vertical distribution

A comparison between the observed global distribution of C_2H_6 and the GEOS-Chem model output for 2010 C_2H_6 emissions from the surface to 10 km is presented in Figure 2.8. The background solid contours are model output from 2010 and filled circles represent seasonal averages from observations compiled in Table 2.1. Measurements at the surface (lower panel) correspond to seasonal averages from the 2010 surface flasks sampling sites and two measurements taken at the Boulder Atmospheric Observatory (BAO) in Colorado, U.S., during February and March of 2011 (Gilman et al., 2013; Swarthout et al., 2013). The aircraft observations are plotted as the averages for each altitude range for individual 20° x 10° grid boxes. We plot the data this way, instead of averaged vertical profiles for individual regions/campaigns, because the observations show large gradients over relatively small areas. We compare model output from 2010 to aircraft observations collected over the period 2008-2014. Since the aircraft data represent a relatively short snapshot of time (15-59 flights over 3-20

weeks), this represents a source of uncertainty in the model evaluation as the model output represents seasonal averages.

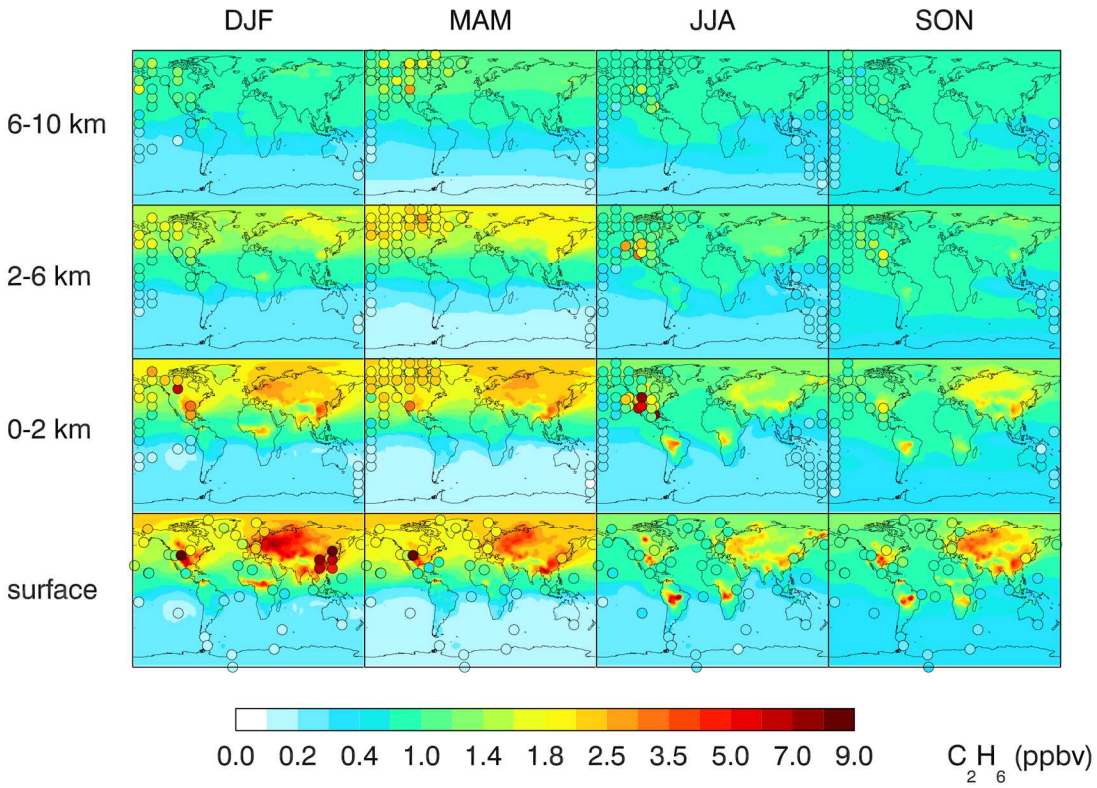


Figure 2.8: Global mean distribution of C_2H_6 for different seasons and altitude ranges compared to observations from Table 2.1. Background solid contours are model outputs for 2010 C_2H_6 emissions. Filled circles represent seasonal averages from observations (2008–2014). Aircraft measurements (panels 0-2, 2-6, and 6-10 km) are averaged vertically for each altitude range and horizontally every 20°x10° (longitude, latitude). Wintertime surface measurements over 43 Chinese cities are averaged horizontally every 20°x10° (longitude, latitude). Overlapping circles represent averaged results from various observations.

Surface measurements are generally well simulated by the model throughout the year. However, during the winter the model tends to over-predict the surface mixing ratios over Europe. The vertical distribution of C_2H_6 mixing ratios over the U.S. will be discussed in the next section. A comparison between Northern Hemisphere observations and modeled mixing ratios outside the U.S. (not shown) reveals that springtime C_2H_6 mixing ratios from HIPPO over latitudes < ~35 degrees are the most under predicted observations (up to 1.5 ppbv). Springtime

HIPPO measurements were taken during the same year of the model simulation and the 2010 C_2H_6 emission inventory. The missing C_2H_6 source might be due to an underestimation of biomass burning over the tropics and subtropics, or potentially point to remaining issues with East Asian emission inventories.

2.5 Model-data comparison over the contiguous U.S.

2.5.1 Model comparison to aircraft campaigns and surface observations

Multiple recent field campaigns, as well as surface observations over the contiguous U.S. allow us to deepen our analysis of anthropogenic C_2H_6 sources and their regional effects on atmospheric mixing ratios. Figure 2.9 presents a comparison between the observed distribution of C_2H_6 and the GEOS-Chem model output for 2010 C_2H_6 emissions over the U.S. At the surface (lower panel) we show seasonal averages from five surface flask-sampling stations and two sets of measurements collected at the Boulder Atmospheric Observatory during February and March of 2011 (Gilman et al., 2013; Swarthout et al., 2013). In this figure, aircraft measurements (0-10 km above the ground) were averaged vertically for each altitude range and horizontally every 5°x5°.

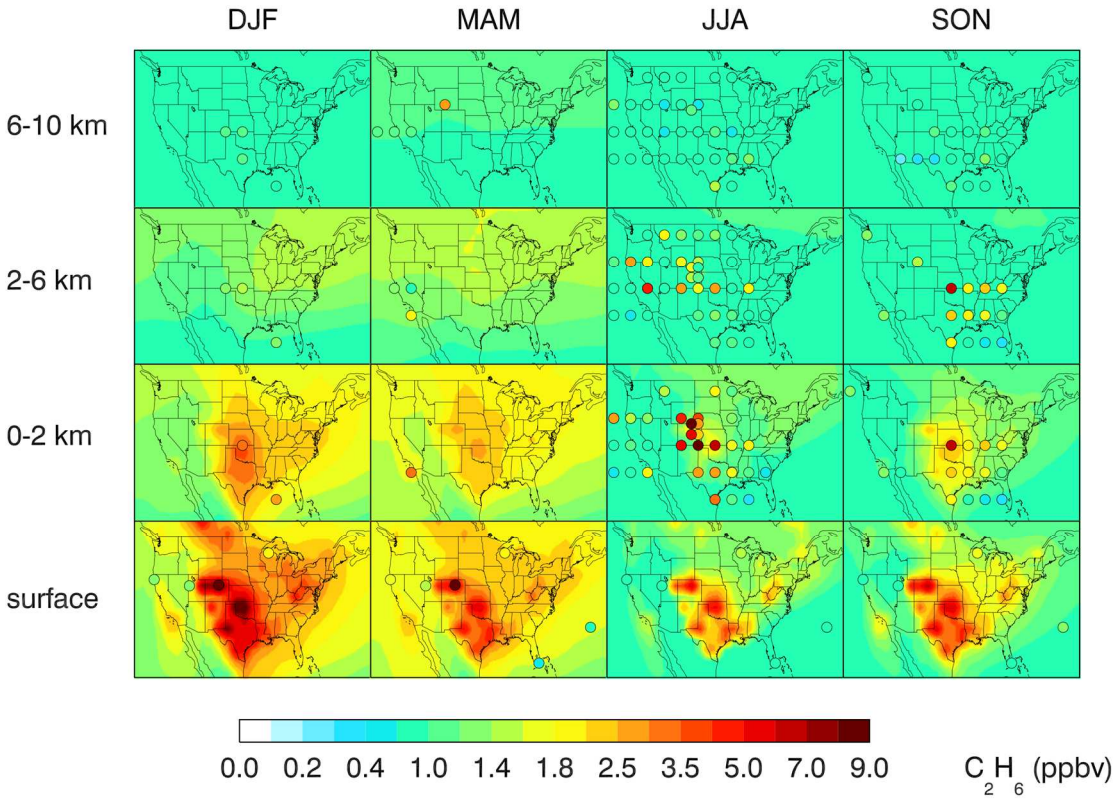


Figure 2.9: Mean mixing ratios of 2010 C_2H_6 emissions over the U.S. for different seasons and altitude ranges compared to observations from Table 2.1. Background solid contours are model outputs. Filled circles represent seasonal averages from observations (2008–2014). Aircraft measurements (panels 0-2, 2-6, and 6-10 km) were averaged vertically for each range of altitude and horizontally every 5°x5° (longitude, latitude). Overlapping circles represent averaged results from various observations.

In the upper troposphere (upper panel; 6-10 km) the model agrees with observations except for the central U.S. Here the model overestimates summertime observed mixing ratios by up to a factor of ~3 (0.5 ppbv). Below 6 km, the model is consistently lower than observations over the central and southeast U.S. Despite scaling up the 2011NEIv1 emissions, the model continues to underpredict by a factor of ~4 the observed abundances of C_2H_6 in the central states of Colorado, Wyoming, Nebraska, and Kansas where there is substantial natural gas and oil extraction (Figure 2.5). The model resolution of 2x2.5 (latitude x longitude degrees) can explain some of the discrepancies between observations and model output. Also, we note that the data presented in

Figure 2.9 are primarily from aircraft campaigns, and the uncertainties (as discussed in section 2.3.2) associated to the temporal mismatch between the aircraft data (2008-2014) and the model (2010) can potentially be greater due to the increase of oil and natural gas extraction over the U.S. during this period (U.S. EPA, 2015). There are greater discrepancies between the most recent field campaigns (SEACR4S (Blake et al., 2014; Schauffler, 2014), DISCOVER-AQ (Yacovitch and Herndon, 2014), and FRAPPÉ (Richter et al., 2015) and the 2010 model outputs. In order to explore this issue, we used the observed 5% yr^{-1} annual rate of change of C_2H_6 total column over the 2009–2014 time period at mid-latitudes reported by Franco et al. (2016) to scale the observations to the year 2010. Unfortunately, underpredictions of mixing ratios across the entire column persist, suggesting that the annual rates of change over these intense oil and natural gas regions are greater than 5% yr^{-1} (as reported at the Boulder FTIR station (Franco et al., 2016)). We cannot rule out the influence of smoke plumes on aircraft measurements during SEACR4S, since 6 out of 26 flights intercepted smoke plumes. However, the FRAPPÉ flights did not intercept major biomass burning plumes, suggesting that neither agricultural burning or wildfire smoke is responsible for the observed discrepancy, at least during summer 2014. The wide spatial coverage of the underprediction suggests a regional impact of oil and natural gas emissions to C_2H_6 mixing ratios.

To investigate the impact of considering the most extreme measured molar C_2H_6/CH_4 ratio observed over oil and natural gas regions, we did an additional simulation to derive global C_2H_6 emissions from CH_4 fluxes of fossil fuels, but using a value of 40.5 (percentage of mol C_2H_6/ mol CH_4) (Kort et al., 2016). This extreme choice results in C_2H_6 emissions that are ~7 times greater than the 2011NEIv1 oil and gas source category. Even though, the higher molar ratio (40.5 percentage of mol C_2H_6/ mol CH_4) is based on observations taken over the Bakken basin,

the use of this ratio produces atmospheric C_2H_6 distributions consistent with aircraft observations over the U.S. central region up to 6 km above the ground. The similar values are consistent with the strong underestimation of the underlying 2011NEIv1 C_2H_6 emissions over the Central U.S.. It is not appropriate to apply the Bakken molar ratio broadly, but the improvement in representing observed mixing ratios over intense oil and natural gas regions suggests that where sufficient information is available, the use of regional emission ratios instead of a fixed nationwide value for all anthropogenic sources may better reproduce C_2H_6 atmospheric abundances. Emission impacts from basins/regions with rapid increase oil and natural gas production like the Bakken shale (Kort et al., 2016) could be better represented by using local emission estimates, which account for the type of oil and natural gas reservoir and the local VOC composition of the natural gas. Our simulations do not show significant emissions from the Bakken shale since the rapid production increase of this basin began in the year 2010 (same year of our simulation, Peischl et al. (2016)).

2.5.2 Boulder C_2H_6 column observations

Over the U.S., we present results of 2010 C_2H_6 column measurements from the National Center for Atmospheric Research Boulder station in Figure 2.6. The Boulder station is located in a region with intensive oil and natural gas production (Gilman et al., 2013; Pétron et al., 2012; Pétron et al., 2014; Swarthout et al., 2013; Thompson et al., 2014). Unfortunately, there are limited data available for 2010; therefore, we present in Figure 2.10 a comparison between model output and C_2H_6 total column observations using de-trended and scaled data for the period 2010-2014 onto the year 2010 using the same methodology and according to the annual rates of change reported by Franco et al. (2016). This removes the effect of the observed C_2H_6 decrease prior to 2009 and its increase from 2009 onwards (Franco et al., 2016). For the Boulder site, the

C_2H_6 total column and GEOS-Chem model output comparison accounts for the altitude difference between the Boulder station and the coarse model grid cell, as explained in Franco et al. (2016). The individual mixing ratio profiles from GEOS-Chem were re-gridded onto the vertical FTIR layer schemes using a mass-conservative interpolation that preserves the total C_2H_6 mass above the station altitude and ignores the mass underneath.

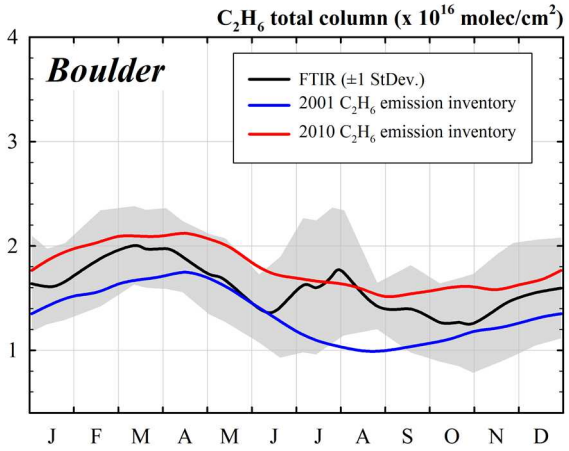


Figure 2.10: Comparison of C_2H_6 total column to 2010 model output at the Boulder site. The black line represent measurements of C_2H_6 total columns over the period (2010–2014) de-trended and scaled to the year 2010, and grey areas their associated 1σ standard deviation. Green, blue, and red lines represent modeled total columns for different emission scenarios. We note that 2012 was a high wildfire year for the Rocky Mountain region.

The model outputs for both C_2H_6 emission inventories encompass observed C_2H_6 total columns. Compared to the 2001 C_2H_6 emission inventory, the C_2H_6 total columns produced by the 2010 inventory stay almost all year round within a one standard deviation from observations. Also, the 2010 C_2H_6 emission inventory produces a seasonality with broader maximum and minimum features that are in line with observations.

CHAPTER 3. ATMOSPHERIC IMPLICATIONS OF LARGE C_2-C_5 ALKANE EMISSIONS FROM THE U.S. OIL AND GAS INDUSTRY

3.1 Methods

In order to investigate the oil and gas contribution to atmospheric abundances of C_2-C_5 alkanes over the U.S., we use emission fluxes from the model-ready version of the 2011v6.3 emissions modeling platform (more specifically the 2011ek modeling case) and incorporate them into the Goddard Earth Observing System global chemical transport model (GEOS-Chem). In this section, we explain the regridding and unit conversion process of the 2011v6.3 emissions modeling platform fluxes, the creation of year-round daily emission fluxes for the year 2011, and the implementation of year-round daily emission fluxes into GEOS-Chem.

3.1.1 Updated 2011NEI emission fluxes over the U.S.

In the U.S., the NEI is released every three years. It is based on activity data from state and local agencies. Here we use an updated version of the 2011NEI that is part of the EPA 2011v6.3 emissions modeling platform (U.S. EPA, 2016b) (https://www.epa.gov/air-emissions-modeling/2011-version-63-technical-support-document). Specifically, the modeling case used for the emissions is from the initial version of the 2011v6.3 platform and is also known as "2011ek". This platform uses the Carbon Bond Mechanism version 6 (CB06) to compute emissions for use as inputs to chemical transport models that require hourly and gridded emissions of chemical species. Relevant to this study, CB06 includes chemical reactions to treat explicit VOC species, such as C_3H_8, benzene, and acetone. In previous model versions, these explicit species were lumped in the paraffin (PAR) species. This dissertation uses specific emissions of propane, benzene and acetone for all model simulations (see Section 3.1.4). Also,

the handling of PAR species in our GEOS-Chem model simulations is explained in Section 3.1.4.

In the 2011v6.3 modeling platform, oil and gas emission sources are divided into point and non-point sources. Oil and gas point sources include extraction and distribution of oil and natural gas, pipeline transportation, and support activities for oil and gas operations. Non-point oil and gas sources include drill rigs, workover rigs, artificial lifts, hydraulic fracturing engines, pneumatic pumps and other devices, storage tanks, flares, truck loading, compressor engines, and dehydrators. The 2011v6.3 platform is expected to better represent the spatial distribution, amount, and type of species emitted from oil and gas sources due to the incorporation of updates over important oil-and-gas-producing basins and speciation profiles based on the Western Regional Air Partnership (WRAP, www.wrapair2.org). WRAP is a voluntary partnership of states, tribes, federal land managers, local air agencies, and the U.S. Environmental Protection Agency (EPA) within the contiguous U.S. West plus North and South Dakota. The WRAP region encompasses several major U.S. natural gas production basins. Incorporating WRAP data into the 2011v6.3 platform is part of multiple efforts by the EPA to revisit and understand the dynamic nature of oil and gas emissions.

3.1.2 Regridding and unit conversion process of emission fluxes

The air quality model-ready emissions in the platform dataset contain daily files with hourly primary emission fluxes in moles per second (mol/s) on a curvilinear grid at 12 km x 12 km horizontal resolution for all states inside the contiguous U.S. (U.S. EPA, 2017). These model-ready emission files were created using the Sparse Matrix Operator Kernel Emissions modeling system (SMOKE, http://www.smoke-model.org/) version 3.7. We converted the emission fluxes using mass conservative interpolation into kilograms per square meter per

second ($kg/m^2/s$) and regridded them onto a rectilinear grid at a 0.1° longitude by 0.1° latitude resolution (equivalent to approximately 8 km x 11 km over the U.S.). We used the Earth System Modeling Framework (ESMF) software to interpolate from the curvilinear SMOKE data grid to the rectilinear GEOS-Chem grid (Appendix B contains scripts used for regridding 2011v6.3 platform anthropogenic emission sources onto the GEOS-Chem grid). Table 3.1 shows a list of the 2011v6.3 platform anthropogenic emission sources considered in this study.

3.1.3 Creation of year-round daily emission fluxes

Each emission sector in the 2011v6.3 platform has daily emission flux files, presented in one out of four different temporal resolutions: daily, according to the day-of-week, weekly, and monthly. Sectors with daily temporal resolution have hourly emissions computed for every day of 2011. Sectors with a temporal approach according to the day-of-week have hourly emissions for four representative days per month: a Saturday, Sunday, Monday and weekday (representing Tuesday through Friday). For sectors with a weekly temporal approach, hourly emissions are computed for all seven days of one representative week in each month. Additionally, the day-of-week and weekly temporal resolutions include emission files for holidays and the consecutive day after each holiday. Table 3.1 summarizes the temporal resolution approach for each of the emission sectors considered. For the seven sectors without a daily temporal resolution, year-round daily emission files were created by reproducing the emission flux files according to the temporal resolution of each sector. For example, each emission sector with a monthly temporal resolution had twelve emission flux files; thus, each monthly file was reproduced according to the number of days of the month it represented.

The complete emissions dataset in 2011v6.3 platform contains additional emission sources and species than the ones considered in this work. Thus, hereafter we will refer to the implemented emissions from 2011v6.3 platform as updated 2011NEI emissions.

Table 3.1: Characteristics of emission sources from the 2011v6.3 platform emissions dataset included in this work.

Source category	Emission sector	Temporal resolution	Number of files
Point	Electric generating units	daily	365
	Point oil and gas	day-of-week	64
	Other point sources	day-of-week	64
Non-point	Agricultural ammonia	daily	365
	Commercial marine vessels	monthly	12
	Non-point oil and gas	weekly	100
	Other non-point sources	weekly	100
	Railroads	monthly	12
	Residential wood combustion	daily	365
On-road		daily	365
Non-road		day-of-week	64

Note: day-of-week and weekly temporal resolutions include emission files for holidays and the consecutive day after each holiday.

3.1.4 GEOS-Chem simulations

We conducted two nested simulations (0.5°x0.6°) over North America (40° to 140° W, 10° to 70° N) using the 3-D chemical transport model GEOS-Chem version 10-01 (Bey et al. (2001), http://www.geos-chem.org) for the year 2011. The GEOS-Chem model was driven by off-line GEOS-5 assimilated meteorological fields (https://gmao.gsfc.nasa.gov/GEOS/) with 47 vertical levels. Global simulations at 2°x2.5° resolution with a spin-up of 18 months were used as boundary conditions for the nested simulations. The emissions and injection timesteps were set to 20 minutes; the transport timestep was set to 10 minutes.

In our baseline simulation, we implemented the updated 2011NEI emission fluxes into GEOS-Chem using the stand-alone software component for computing emissions, Harvard-NASA

Emissions Component (HEMCO) version 1.1.005 (Keller et al., 2014). Over the U.S. (CONUS), all anthropogenic and biofuel emissions were derived from the updated 2011NEI. Outside the CONUS, we used anthropogenic emissions from the Emissions Database for Global Atmospheric Research (EDGAR v4.2) and VOC emissions from the Reanalysis of the Tropospheric chemical composition (RETRO) emission inventory, except for C_2H_6 and C_3H_8 for which we used the Tzompa-Sosa et al. (2017) and Xiao et al. (2008) emission inventories, respectively. We also include regional anthropogenic emission inventories for northern Mexico (Kuhns et al., 2003), Canada (https://www.canada.ca/en/environment-climate-change/services/national-pollutant-release-inventory/tools-resources-data.html), Europe (http://www.ceip.at/), and Asia (http://meicmodel.org/dataset-mix.html). Non-U.S. biofuel emissions were from the Yevich and Logan (2003) emission inventory, with two exceptions. 1) We used ammonia (NH_3) emissions from the Global Emissions InitiAtive (GEIA), and 2) C_2H_6 emissions from Tzompa-Sosa et al. (2017). Shipping, aviation and natural sources are expected to make minor contributions to the emissions of C_2-C_5 alkanes; the default global datasets incorporated into GEOS-Chem were used for these sectors.

Table 3.2 shows the summary of the emission inventories that we used in the baseline simulation.

For the second simulation (hereafter updated 2011NEI: OG off), we maintained the same configuration of emission inventories as the baseline simulation, but we turned off the updated 2011NEI emissions of ≥ C_2 alkanes from the oil and gas sector. We used this simulation to investigate the contribution of oil and gas related activities to the abundance of the ≥ C_2 alkanes over the CONUS.

The alkane speciation used in GEOS-Chem was originally based on the Lurmann et al. (1986) condensed gas phase chemical mechanism. The current mechanism in GEOS-Chem treats C_3H_8 and C_2H_6 as explicit species. ≥ C_4 alkanes are lumped into one tracer, originally named ALKA in Lurmann et al. (1986), and currently named ALK4 in GEOS-Chem. The rate constant in GEOS-Chem associated with the reaction of OH with ALK4 is based on the absolute rate coefficient of butane ($9.1 \times 10^{-12} e^{(-405/T)}$ cm^3 $molecule^{-1}$ s^{-1}, Atkinson et al. (2006)), but is used to represent the chemistry of all ≥ C_4 alkanes. Both of our simulations have a global annual mean tropospheric mass-weighted OH concentration of 1.3×10^6 molecule cm^{-3}, which is very close to the upper bound of previous model studies in the literature (Naik et al., 2013; Voulgarakis et al., 2013). In this study, we consider ALK4 specifically as n-butane, i-butane, n-pentane, and i-pentane (hereafter referred to as C_4-C_5 alkanes), rather than a more inclusive ≥ C_4 alkanes. Thus, one obvious challenge with the model-observation comparison that we present later in the paper is the use of a single reaction rate for C_4-C_5 alkanes. As explained in Section 3.1.1, 2011v6.3 platform emissions of PAR species include alkanes. We assigned a fraction of PAR to C_4-C_5 alkane species based on Simon et al. (2010). Simon et al. [2010] summarize the 50 VOCs with the largest emissions over the U.S.; for an example day C_4-C_5 alkanes correspond to 36% of

these emissions over the U.S. Thus, we set ALK4 emissions as 36% of the total PAR emitted species. The remaining reactive carbon in the PAR species is not considered here. Omitting such a large fraction of reactive carbon limits our ability to provide a full view of the impact of oil and gas operations and urban activities on atmospheric composition. This is a known limitation to our approach. However, we investigated the impact of attributing 100% of PAR to ALK4 versus 36% of PAR to ALK4 on ALK4 lifetime. Omitting such a large fraction of carbon changes the lifetime of ALK4 over the U.S. by < 5%. As discussed in the conclusions, we suggest the addition of a new GEOS-Chem tracer for C_6-C_8 alkanes, which based on the 50 VOCs with largest emissions over the U.S. (Simon et al., 2010), account for ~40% of PAR (see Table B1 in Appendix B). Thus, adding such fraction of the remaining reactive carbon from the PAR species could provide a better estimate of the full impact of the emissions from this sector.

The GEOS-Chem mechanism does not include other paraffin compounds, such as alkynes, and higher aromatic VOCs that have also been found in high abundances (compared to background values) over oil and gas basins (Abeleira et al., 2017; Gilman et al., 2013; Helmig et al., 2014b; Pétron et al., 2012; Pétron et al., 2014; Swarthout et al., 2013; Swarthout et al., 2015; Thompson et al., 2014; Zielinska et al., 2014). Additionally, our model simulations do not include tropospheric chlorine chemistry; thus reaction with OH is the only tropospheric sink of C_2-C_5 alkane species. Sherwen et al. (2016) used the same initial GEOS-Chem version, but they added tropospheric halogen chemistry (Cl, Br, I). In their study, adding a chlorine sink term led to decreases in the tropospheric global burdens of C_2H_6, C_3H_8, and ≥ C_4 alkanes of 19%, 14% and 12%, respectively. However global tropospheric burden changes are heterogeneous, and in general they are lower over land compared to oceans. At the surface over the U.S., the annual average changes are smaller (typically <10% for C_2H_6 and less for C_3H_8, and ≥ C_4 alkanes -

Sherwen et al. [2016]). The inclusion of halogen chemistry would decrease the O_3 burden, and thus the OH burden as well. This would increase the lifetimes of C_2H_6, C_3H_8, and ≥ C_4 alkanes against OH oxidation. Given that the inclusion of updated 2011NEI emissions in our model produces significant increases in C_2-C_5 alkane fluxes over the U.S. compared to the emission inventory used by Sherwen et al. (2016), a comprehensive understanding of these two model developments would require additional model simulations.

Table 3.2: Configuration of emission inventories in our baseline simulation.

Emission inventory	Region	Base year	Species in GEOS-Chem
Anthropogenic			
Updated 2011NEI	CONUS	2011	ACET, ALD2, ALK4, BCPI, BCPO, BENZ, C_2H_4, C_2H_6, C_3H_8, CO, EOH, FORM, HONO, MACR, MEK, MOH, NH_3, NO, NO_2, OCPI, OCPO, PRPE, RCHO, SO_2, SO_4, TOLU, XYLE
BRAVO	Northern Mexico	1999	CO, NO, SO_2, SO_4
CANADA*	Canada	2002	CO, NO, SO_2, SO_4
		2008	NH_3
EMEP	Europe	2011	CO, NH_3, NO, SO_2
		2000	ALD2, ALK4, MEK, PRPE
MIX*	Asia	2010	ALD2, ALK4, CH_2O, CO, NH_3, NO, SO_2, SO_4, MEK, PRPE
EDGAR*	global	2008	CO, NAP, NH_3, NO, SO_2, SO_4
Tzompa-Sosa et al. (2017)	global	2010	C_2H_6
Xiao et al. (2008)	global	1985	C_3H_8
RETRO	global	2000	ACET, ALD2, ALK4, BENZ, CH_2O, C_2H_2, C_2H_4, MEK, PRPE, TOLU, XYLE
Biofuel			
Updated 2011NEI	CONUS	2011	ACET, ALD2, ALK4, BCPI, BCPO, BENZ, C_2H_4, C_2H_6, C_3H_8, CO, EOH, FORM, HONO, MACR, MEK, MOH, NH_3, NO, NO_2, OCPI, OCPO, PRPE, RCHO, SO_2, SO_4, TOLU, XYLE
Yevich and Logan (2003)	global, except the CONUS	1985	ACET, ALD2, ALK4, BENZ, CH_2O, C_2H_2, C_2H_4, C_3H_8, CO, GLYC, GLYX, HAC, MEK, MGLY, NAP, NO, PRPE, SO_2, TOLY, XYLE
GEIA	global	1998	NH_3
Tzompa-Sosa et al. (2017)	global	2010	C_2H_6
Shipping			
EMEP	Europe	2011	CO, NO, SO_2
ARCTAS	global	2008	SO_2
ICOADS	global	2002	CO, NO
Aviation			
AEIC	global	2005	ACET, ALD2, ALK4, BC, CH_2O, C_2H_6, C_3H_8, CO, MACR, NO, SO_2, SO_4, OC, PRPE, RCHO

Emission inventory	Region	Base year	Species in GEOS-Chem
Natural sources			
GEOS-Chem default	global	2000	NO
		1985	DMS
		2009	SO_2
GEIA	global	1990	NH_3

Notes:
1. Over the CONUS, all anthropogenic and biofuel emissions in the baseline simulation come from the updated 2011NEI.
2. Unless otherwise noted, the simulation uses the same year as the base year for the 2011 simulation.
* Projected to 2010 using GEOS-Chem default annual scaling factors.

3.2 Results and discussion

3.2.1 Contribution of the oil and gas sector to emissions of C_2-C_5 alkanes

Ethane and Propane

The oil and gas sector emits C_2H_6 and C_3H_8 primarily due to leakage during the production, processing, and transportation of natural gas (Gilman et al., 2013; Kort et al., 2016; Pétron et al., 2012; Roest and Schade, 2017). Trace amounts of C_2H_6 and C_3H_8 can also be produced during hydrocarbon combustion processes (Basevich et al., 2012; Gomer and Kistiakowsky, 1951; Sangwan et al., 2015; Thynne, 1962). In the updated 2011NEI, total emissions of C_2H_6 and C_3H_8 are dominated by oil and gas sources (point sources - e.g. oil and gas extraction, distribution, pipelines - and non-point sources - e.g. flares, drill and workover rigs) with an estimated contribution to total anthropogenic emissions for the U.S. of 89% and 82%, respectively (Figure 3.1). The remaining percentage of C_2H_6 and C_3H_8 emissions is distributed among other sources such as vehicles and residential wood combustion. We note that from the two oil and gas sectors considered, non-point sources are the biggest contributors, accounting for 95% and 97% of the total oil and gas contribution estimate.

The total updated 2011NEI emissions of C_2H_6 from the present study are 15% lower than the Tzompa-Sosa et al. (2017) C_2H_6 emission inventory estimate, which was calculated by scaling C_2H_6 emissions of 2011NEI version 1 (2011NEIv1) by a factor of 1.4 based on a comparison to existing observations. It is important to note that the 2011NEIv1 did not contain updates over oil and gas basins based on the WRAP (the updated 2011NEI used here includes the WRAP data), causing oil and gas regions like the Uintah basin to have minimal C_2H_6 emission fluxes. In this region, where studies have found important C_2H_6 emission enhancements (Helmig et al., 2014b; Koss et al., 2015; Warneke et al., 2014), emission fluxes from oil and gas sources in 2011NEIv1 are close to zero, thus upward scaling by 1.4 still results in a small flux. Thus, if the scaling in Tzompa-Sosa et al. (2017) were applied in this study, the result would be a different spatial distribution and amount of emission fluxes compared to the updated 2011NEI used here. Lastly, we notice that between 2011NEIv1 and the updated 2011NEI used in this work, emissions of C_3H_8 have higher emission flux increases compared to C_2H_6. The highest emission fluxes occur over oil and gas regions in the central U.S., with increases of up to 300 ng m^{-2} s^{-1}.

C_4-C_5 alkanes

Over the last decade, leakage from oil and gas sources has become an important contributor to the emissions of C_4-C_5 alkanes (Gilman et al., 2013; Johansson et al., 2014; Roest and Schade, 2017; Swarthout et al., 2013; Swarthout et al., 2015), which historically were dominated by automobile combustion, and fugitive emissions from gasoline and diesel distribution (Lee et al., 2006; Schauer et al., 2002). Thus, urban areas are the locations where enhancements of C_4-C_5 alkanes commonly observed (Aceves and Grimalt, 1993; Bi et al., 2005; Lee et al., 2006). Rossabi and Helmig (2018) recently used data collected between 2001and 2015

over the U.S. to show a predominantly decreasing trend in C_4-C_5 alkanes surface mixing ratios, but they found a relative increase in the predominance of the n-isomers. They attributed this pattern to changes in isomeric ratios in gasoline sector emissions, and emissions from the oil and gas industry. The emergence of U.S. oil and gas development as a larger source of C_4-C_5 alkanes has increased their atmospheric abundances in areas with low population density (Gilman et al., 2013; Pétron et al., 2014; Warneke et al., 2014). Gilman et al. (2013) estimated that the mean oil and gas contribution to C_4-C_5 alkane emissions in northeastern Colorado is 93%-96%. Based on the updated 2011NEI emissions (Figure 3.1), we estimate that oil and gas sources (including both point and non-point sources) over the CONUS are the third most important emission source of C_4-C_5 alkanes with an annual contribution of 26% of the total emissions.

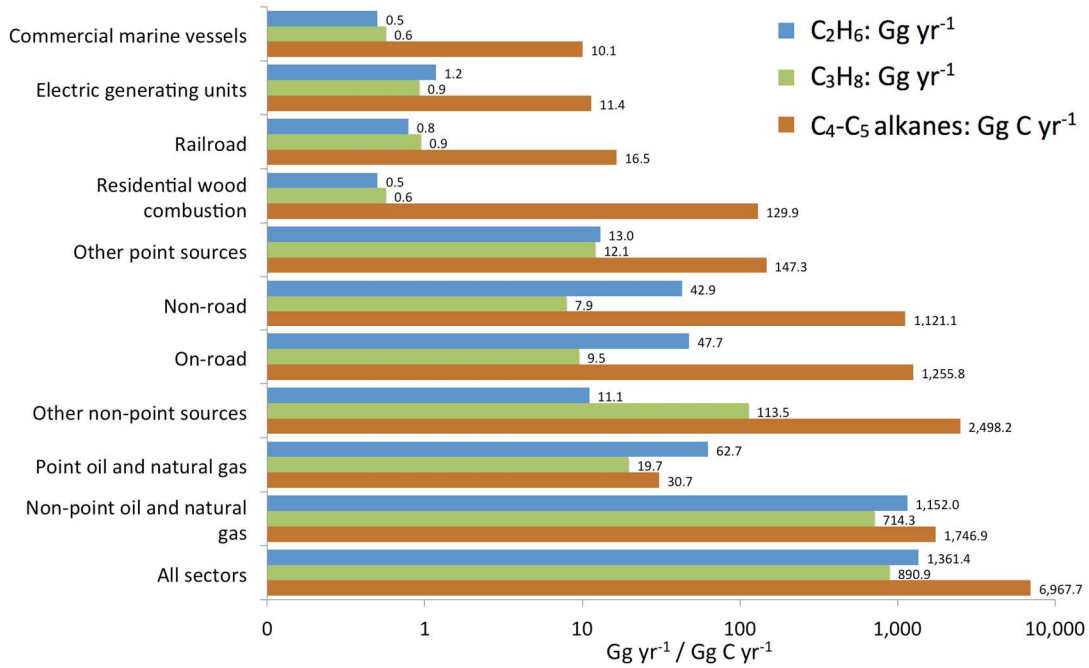


Figure 3.1: Updated 2011NEI emissions of C_2H_6, C_3H_8, and C_4-C_5 alkanes by sector. C_4-C_5 alkanes are presented as 36% of PAR emissions. Units for C_2H_6 and C_3H_8, are in Gg yr^{-1}; and units for C_4-C_5 alkanes are presented in Gg C yr^{-1}.

3.2.2 Geographical distribution of oil and gas C_2-C_5 alkane emissions and its contribution to U.S. total anthropogenic emissions

In the U.S., emissions of C_2H_6 and C_3H_8 are mainly clustered inside oil and gas basins, where the contribution of the oil and gas sector to total anthropogenic emissions is > 90% (Figure 3.2, panels a and b) For C_4-C_5 alkanes, the emissions not only occur inside oil and gas basins, but also in urban areas due to the importance of other fossil fuel sources. The contribution of urban sources to total emissions of C_4-C_5 alkanes over oil-and-gas-producing regions reduces the overall percentage contribution of these sources (Figure 3.2, panel c).

A comparison between regional emissions of C_2H_6, C_3H_8, and C_4-C_5 alkanes shows that the central region of the U.S. is the most important contributor to total CONUS C_2-C_5 alkane emissions in 2011, contributing ~70% of C_2H_6 and C_3H_8 total CONUS emissions, and ~40% of the emissions of C_4-C_5 alkanes (Figure 3.3). The central region fully encompasses four U.S. oil and gas basins: Eagle Ford, Permian, Niobrara, and Bakken. This estimate is likely to be higher for years later than 2011 for C_2H_6 due to the massive increase in oil and gas exploitation in the Bakken basin (Kort et al., 2016; Peischl et al., 2016).

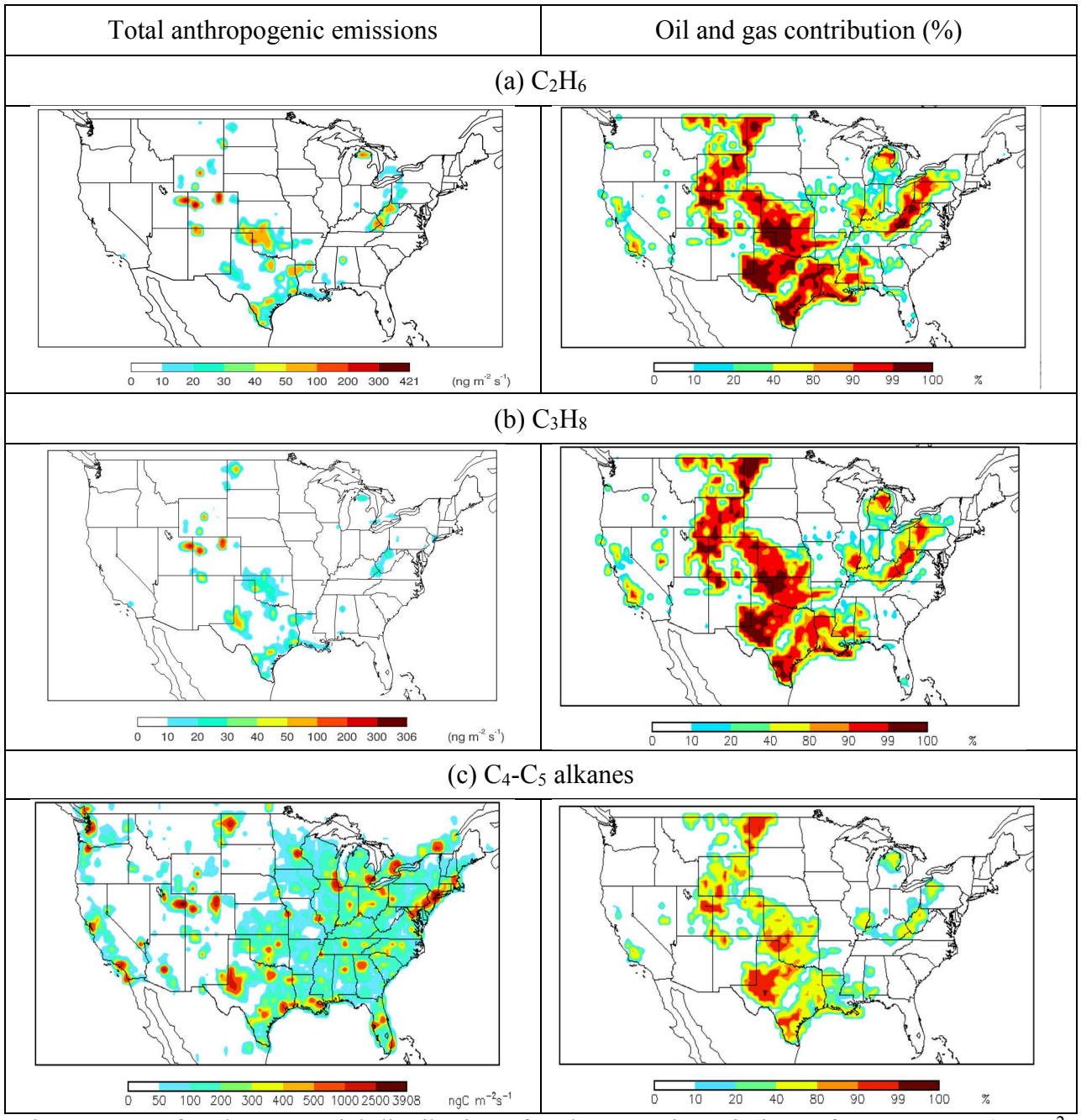


Figure 3.2: Left column: spatial distribution of anthropogenic emissions of C_2H_6, C_3H_8, (ng m^{-2} s^{-1}) and C_4-C_5 alkanes (ng C m^{-2} s^{-1}). Right column: spatial distribution of the percent contribution of oil and gas emissions to total anthropogenic emissions of C_2H_6, C_3H_8, and C_4-C_5 alkanes. C_2-C_5 alkane emissions data from the updated 2011NEI. C_4-C_5 alkanes are presented as 36% of PAR emissions.

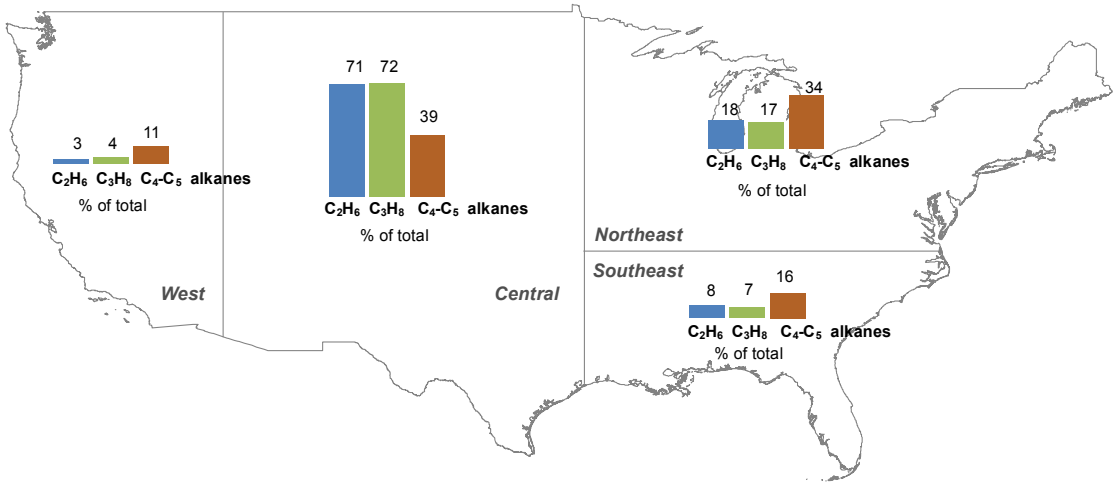


Figure 3.3: Regional contributions (as %) to U.S. total anthropogenic emissions of C_2H_6, C_3H_8, and C_4-C_5 alkanes. C_2-C_5 alkane emissions data from the updated 2011NEI. C_4-C_5 alkanes are presented as 36% of PAR emissions.

3.3 Model comparison to observations and oil and gas contribution to atmospheric abundances of C_2-C_5 alkanes

We compare 2011 abundances of C_2H_6, C_3H_8, and C_4-C_5 alkanes from a GEOS-Chem simulation to a suite of observations over North America (Table 3.3 and Figure 3.4). To the best of our knowledge, this constitutes the largest compendium of C_2-C_5 alkane observations compared to model output for this region. Also, we estimate the contribution of oil and gas to atmospheric abundances of C_2H_6, C_3H_8, and C_4-C_5 alkanes by turning off the emissions of these species from the oil and gas sector in a separate GEOS-Chem simulation (updated 2011NEI: OG off).

Table 3.3: Observations from surface sites and airborne campaigns, ordered by type and date.

2011 FTIR column measurements

Species	Site	Location	Period	Reference
C_2H_6	Toronto, Ontario, Canada	79.4° W, 43.6° N	Jan-Dec 2011	Wiacek et al. (2007)
C_2H_6	Boulder, Colorado, USA	105.3° W, 40.4° N	Jan-Dec 2011	Hannigan et al. (2009)

Aircraft campaigns

Species	Field campaign	Region	Period	Reference
Figures 3.7-3.9				
$C_{2\text{-}5}$	ARCTAS	110° to 126° W, 30° to 50° N	Apr, Jun-Jul 2008	Simpson et al. (2010) Simpson et al. (2011)
$C_{2\text{-}5}$	HIPPO	90° to 116° W, 25° to 50° N	Jun-Sep 2011	Wofsy et al. (2012)
$C_{2\text{-}5}$	SEACR4S	80° to 126° W, 25° to 50° N	Aug-Sep 2013	Blake et al. (2014) Schauffler (2014)
$C_{2\text{-}5}$	FRAPPÉ	101° to 109° W, 38° to 46° N	Jul-Aug 2014	Richter et al. (2015)

2011 Surface flask measurements from National Oceanic and Atmospheric Administration (NOAA) / Institute of Arctic and Alpine Research (INSTAAR) Global VOC Monitoring Program

Species	Site	Location	Period	Website
$C_{2\text{-}5}$	Key Biscayne, Florida (KEY), USA	80.16°W, 25.67°N	Jan-Dec 2011	http://instaar.colorado.edu/arl/Global_VOC.html
$C_{2\text{-}5}$	Park Falls, Wisconsin (LEF), USA	90.27° W, 45.95° N	Jan-Dec 2011	
$C_{2\text{-}5}$	Southern Great Plains, Oklahoma, (SGP), USA	97.5° W, 36.8° N	Jan-Dec 2011	
$C_{2\text{-}5}$	Trinidad Head, California (THD), USA	124.15° W, 41.05° N	Jan-Dec 2011	
$C_{2\text{-}5}$	Wendover, Utah (UTA), USA	113.72° W, 39.9° N	Jan-Dec 2011	

2011 Surface flask measurements from Photochemical Assessment Monitoring Stations (PAMS)

Species	Site	Location	Period	Website
$C_{2\text{-}5}$	Baltimore, Maryland (BAL), USA	76.6° W, 39.3° N	Jun-Aug 2011	https://www.airnowtech.org
$C_{2\text{-}5}$	Boston, Massachusetts (BOS), USA	71.1° W, 42.4° N	Jun-Aug 2012	
$C_{2\text{-}5}$	El Paso, Texas (ELP), USA	106.4° W, 31.8° N	Jan-Dec 2011	
$C_{2\text{-}5}$	Gary, Indiana (GAR), USA	87.3° W, 41.6° N	Jun-Dec 2011	
$C_{2\text{-}5}$	Houston, Texas (HOU), USA	95.4° W, 29.8° N	Jan-Dec 2011	
$C_{2\text{-}5}$	Los Angeles, California (LAX), USA	118.3° W, 34.1° N	Jan-Dec 2011	
$C_{2\text{-}5}$	Philadelphia, Pennsylvania (PHI), USA	75.2° W, 40° N	May-Oct 2011	
$C_{2\text{-}5}$	Atlanta, Georgia (SDK), USA	84.4° W, 33.8° N	Jun-Aug 2011	
$C_{2\text{-}5}$	Springfield, Massachusetts (SPR), USA	72.5° W, 42.1° N	Jun-Aug 2011	

Surface observations				
Species	Site	Location	Period	Reference
C_{4-5}	Houston Ship Channel, Texas (HSC), USA	95.03° W, 29.65° N	Sep 2006	Johansson et al. (2014)
C_{3-5}	San Francisco, California (STR), USA	122.45° W, 37.76° N	Jun-Aug 2007-2010	Pétron et al. (2012)[1]
C_{3-5}	Walnut Grove, California (WGC), USA	121.49° W, 38.26° N	Jun-Aug 2007-2010	Pétron et al. (2012)[1]
C_{3-5}	Moody, Texas (WKT), USA	97.33° W, 31.32° N	Jun-Aug 2007-2010	Pétron et al. (2012)[1]
C_{3-5}	Park Falls, Wisconsin (LEF), USA	90.27° W, 45.93° N	Jun-Aug 2007-2010	Pétron et al. (2012)[1]
C_5	Barnett Shale, Texas (BST), USA	97.42° W, 33.27° N	May 2010	Zielinska et al. (2014)
C_{3-5}	Boulder Atmospheric Observatory, Colorado (BAO),	105.01° W, 40.05° N	Aug 2007-Apr 2010	Pétron et al. (2012)
C_{2-5}	Boulder Atmospheric Observatory, Colorado (BAO),	105.01° W, 40.05° N	Feb-Mar 2011	Gilman et al. (2013)
C_{2-5}	Boulder Atmospheric Observatory, Colorado (BAO),	105.01° W, 40.05° N	Feb-Mar 2011	Swarthout et al. (2013)
C_{2-5}	Beaumont Downtown, Texas (BDT), USA	94.07° W, 30.04° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Cesar Chavez HS, Texas (CCH), USA	95.25° W, 29.68° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Channelview, Texas (CNV), USA	95.13° W, 29.8° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Clinton, Texas (CLT), USA	95.26° W, 29.73° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Corpus Christi Oak Park, Texas (CCO), USA	97.43° W, 27.8° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Corpus Christi Palm, Texas (CCP), USA	97.42° W, 27.8° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Corpus Christi Solar Estates, Texas (CCS), USA	97.54° W, 27.83° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Dallas Hinton, Texas (DHT), USA	96.86° W, 32.82° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Danciger, Texas (DNG), USA	95.76° W, 29.14° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Decatur Thompson, Texas (DTS), USA	97.58° W, 33.22° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Deer Park, Texas (DPK), USA	95.13° W, 29.67° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Dish Airfield, Texas (DAF), USA	97.3° W, 33.13° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	Eagle Mtn Lake, Texas (EML), USA	97.48° W, 32.99° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	El Paso Chamizal, Texas (EPC), USA	106.46° W, 31.77° N	Jan-Dec 2011	TCEQ (2012)[2]
C_{2-5}	El Paso Delta, Texas (EPD), USA	106.41° W, 31.76° N	Jan-Dec 2011	TCEQ (2012)[2]

$C_{2\text{-}5}$	Everman Johnson Park, Texas (WJP), USA	97.29° W, 32.62° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Flower Mound, Texas (FWM), USA	97.13° W, 33.05° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Fort Worth NW, Texas (FWN), USA	97.36° W, 32.81° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	HRM3, Texas (HRM), USA	95.18° W, 29.76° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Lake Jackson, Texas (LJK), USA	95.47° W, 29.04° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Lynchburg Ferry, Texas (LBF), USA	95.08° W, 29.76° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Milby Park, Texas (MPK), USA	95.26° W, 29.71° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Nederland HS, Texas (NDL), USA	94.01° W, 29.98° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Odessa Hays, Texas (OHY), USA	102.34° W, 31.84° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Texas City 34th St., Texas (TXC), USA	94.95° W, 29.41° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Wallisville Rd, Texas (WVR), USA	94.99° W, 29.82° N	Jan-Dec 2011	TCEQ (2012) [2]
$C_{2\text{-}5}$	Hickory, Pennsylvania (HKY), USA	80.30° W, 40.30° N	July 2012	Swarthout et al. (2015)
$C_{2\text{-}5}$	Racoon Creek State Park, Pennsylvania (RCS), USA	80.50° W, 40.50° N	July 2012	Swarthout et al. (2015)
$C_{2\text{-}5}$	Erie, Colorado (ERC), USA	105.05° W, 40.05° N	Mar-Jun 2013	Thompson et al. (2014)
$C_{2\text{-}5}$	Boulder Atmospheric Observatory, Colorado (BAO),	105.01° W, 40.05° N	Mar-May, Jul-Sep, 2015	Abeleira et al. (2017)

Notes:
1. C_4 observations only include n-C_4H_{10}.
2. C_4-C_5 observations only include n-C_4H_{10} and n-C_5H_{12}.

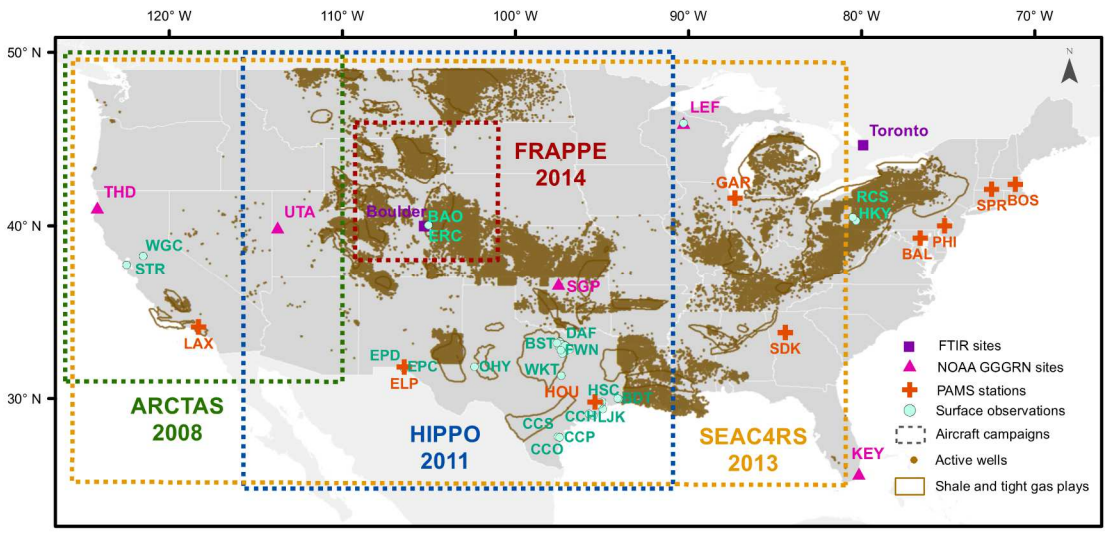


Figure 3.4: Summary of observations listed in Table 3.3. Labels of overlapping surface observations are not shown. Locations of active wells come from FracTracker (accessed Nov 2015, www.fractracker.org). In order to provide a sense for well spatial distribution over states with missing data, shale and tight gas plays (Energy Information Administration, accessed Dec 2014, www.eia.gov/dnav/ng/ ng_sum_lsum_a_EPG0_xdg_count_a.htm) are shown.

3.3.1 Comparison to ground-based FTIR C_2H_6 column observations

In this section, we compare 2011 C_2H_6 total columns derived from ground-based Fourier Transform Infrared (FTIR) solar observations at the Boulder and Toronto to GEOS-Chem simulated C_2H_6 total columns stations for our two emission scenarios (Figure 3.5). The C_2H_6 total columns were consistently determined at both sites following the methodology described in Franco et al. (2015). This latter paper further provides information on the typical systematic and random uncertainties affecting the column measurements. The first emission scenario considers all emissions and sectors from the updated 2011NEI. In the second emission scenario, C_2H_6 emissions from the oil and gas industry are turned off (updated 2011NEI: OG off). Finally, the oil and gas contribution to C_2H_6 total columns is calculated by subtracting the results of the second scenario from those of the first scenario (updated 2011NEI - updated 2011NEI: OG off).

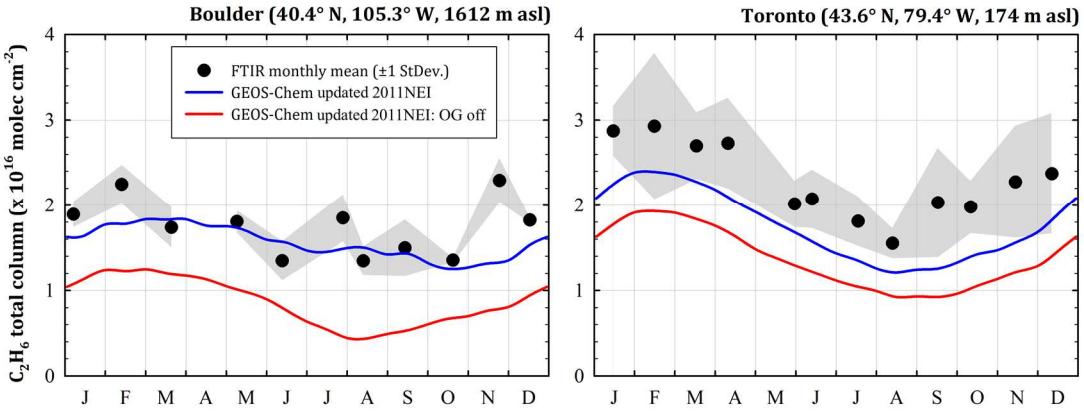


Figure 3.5: Comparison of 2011 FTIR C_2H_6 total columns to GEOS-Chem C_2H_6 columns using a simulation with and without oil and gas sources from the updated 2011NEI. Black dots represent FTIR monthly mean C_2H_6 total columns, and the grey shading denotes their associated 1σ standard deviation. Monthly means are displayed proportionally to the observations available in each month. The blue line represents modeled C_2H_6 total columns using all sectors from the updated 2011NEI. The red line represents modeled C_2H_6 total columns with C_2H_6 emissions from oil and gas sector turned off (updated 2011NEI: OG off). The blue and red lines are running mean fits to the daily-averaged model columns (with a 6-week wide integration time and a 15-day time step).

At the Boulder station, C_2H_6 emissions from the updated 2011NEI reproduce observed C_2H_6 total columns outside of winter months. A difference of ~0.2-1x10^{16} molecules cm^{-2} is observed during the winter season (including November). At the Toronto station, modeled updated 2011NEI C_2H_6 emissions underestimate (on average by ~0.5x10^{16} molecules cm^{-2}) the observed C_2H_6 total column throughout the year 2011. We note that as shown in Table 3.2, our simulation does not include recent updates to C_2H_6 emission fluxes made by Environment and Climate Change Canada. The difference in observed and modeled C_2H_6 total columns might be due to a combination of underestimated urban C_2H_6 leakage from natural gas delivery and end use, residential wood combustion, and the higher resolution (0.5°x0.6°) analysis made in this study using 2°x2.5° C_2H_6 emissions derived by Tzompa-Sosa et al. (2017). The coarser resolution of the C_2H_6 emissions over Toronto limits the ability of our higher resolution model

simulation to capture local enhancements. The total C_2H_6 columns observed and produced by both emission scenarios over Toronto are larger than over Boulder. Considering that the column measurements are sensitive to the whole troposphere and lower stratosphere, the column difference between Toronto and Boulder can be explained by the altitude difference between both stations (~1.5 km). Another possible explanation of the column difference is the latitudinal gradient in C_2H_6, with higher abundances towards the Artic (Helmig et al., 2016; Simpson et al., 2012).

There is a greater contribution to modeled total C_2H_6 columns from emissions from the oil and gas sector (~$0.7x10^{16}$ molecules cm^{-2}) at the Boulder station compared to the Toronto station (~$0.4x10^{16}$ molecules cm^{-2}). This finding is consistent with results presented in Franco et al. (2016) for the 2009 - 2014 period. Among six FTIR stations (including Toronto) located at different latitudes across the Northern Hemisphere, they showed that the Boulder station had the highest rate of change in the C_2H_6 total column over this time period, presumably associated with the oil and gas development in the central U.S. The high contribution of the oil and gas sector over Boulder is also shown in results from this study. From the four regions analyzed in Section 3.3.3 (Figure 3.10), the region where the Boulder station is located shows the highest percentage contribution from the oil and gas sector to total abundances of C_2-C_5 alkanes throughout the troposphere.

3.3.2 Comparison to surface flask observations

In this section, we compare 2011 simulated alkane mixing ratios with and without oil and gas sources (blue and red lines, respectively in Figure 3.6) to measured C_2-C_5 surface mixing ratios from samples collected at selected U.S. stations (Table 3.3) from the NOAA Global Greenhouse Gas Reference Network (GGGRN) (Figure 3.6). Stations are ordered from higher to

lower latitudes due to the observed strong latitudinal gradient of C_2-C_5 alkane abundances (Helmig et al., 2016; Simpson et al., 2012).

Differences between our simulations with and without emissions of C_2H_6 and C_3H_8 from the oil and gas sector (blue versus red lines in Figure 3.6), suggest that the SGP station was more impacted by emissions from this industry compared to the rest of the stations throughout 2011. The estimated annual oil and gas source contributions to surface C_2H_6 and C_3H_8 mixing ratios at the SGP station is 86% for both species. The higher oil and gas impact at the SGP station is expected because it is located inside an oil and gas region. Typical i-pentane/n-pentane ratio for regions dominated by emissions from the oil and gas sector range from 0.89 – 1.10 (Gilman et al., 2013). The calculated 2011 i-pentane/n-pentane ratio at SGP is 0.97, corroborating that air masses in this area are highly impacted by oil and gas sources. LEF and KEY are two other stations where the model predicts that oil and gas activities make a large contribution to atmospheric abundances of C_2-C_5 alkanes (~25% for C_2H_6 and C_3H_8). The relatively high oil and gas contributions are consistent with Helmig et al. (2016); their analysis shows higher rates of changes between 2009-2014 in C_2H_6 and C_3H_8 occurring in sites downwind the central and eastern U.S.

For C_4-C_5 alkanes, the model overestimates monthly mean mixing ratios especially at SGP, LEF, and KEY stations, where the monthly overestimations are as high as 22 ppbC (bottom row of Figure 3.6). This overestimation can stem from our choice of assigning C_4-C_5 alkanes as a continuous fraction of total PAR emitted species across the CONUS (see Section 2.4), meaning that 36% of PAR is a high fraction for C_4-C_5 alkanes in these areas. Another possible cause is the overestimation of total PAR emissions in these areas. At the THD and UTA stations, the overestimation is not nearly as dramatic as for the other stations; the average annual

overestimation for both sites is ~1.2 ppbC. Both stations, THD and UTA are located away from urban areas. Most importantly, we note that only a few of long-term stations are ideally located to capture changes related to the major oil and gas source regions that have the highest emissions in the updated 2011NEI. Long-term monitoring stations located in northeastern Colorado, Wyoming, and North Dakota should be considered in order to capture emission changes in the oil and gas sector.

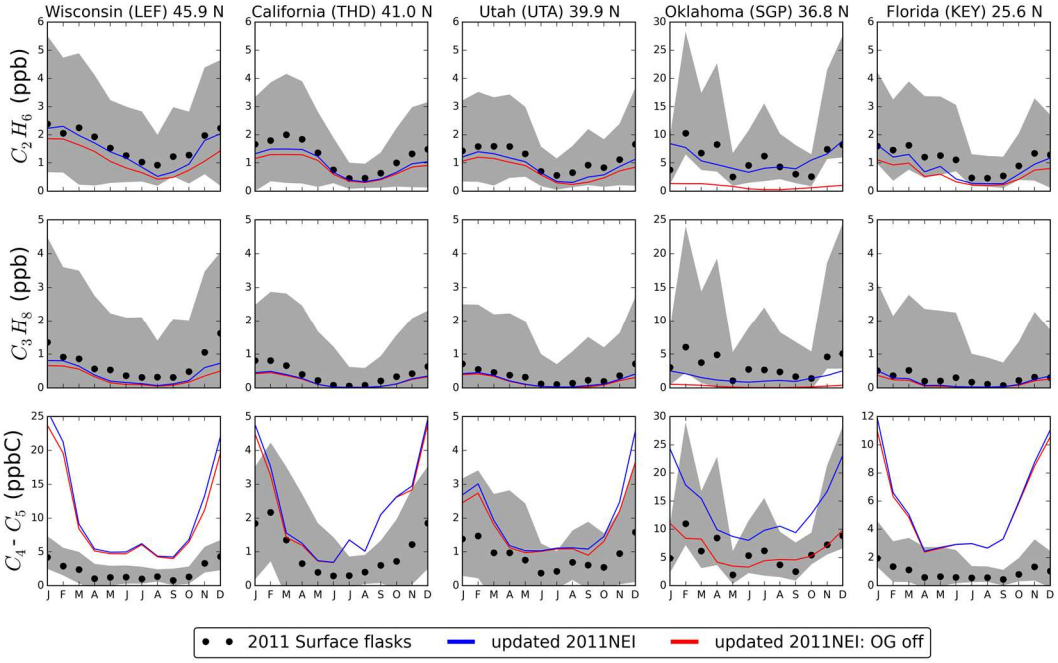


Figure 3.6. Comparison of 2011 surface mixing ratios for C_2H_6, C_3H_8 and C_4-C_5 alkanes (from top to bottom) to modeled 2011 emissions from the updated 2011NEI with and without oil and gas sources. Black dots represent monthly mean observations from NOAA GGGRN global surface flask network (Table 3.3), and the grey areas denote their associated 90th percentile. The blue line represents monthly mean simulated surface mixing ratios using emissions from all sectors of the updated 2011NEI. The red line represents mixing ratios from the updated 2011NEI: OG off simulation. The stations are ordered from higher (left) to lower (right) latitudes. Note the various vertical scales.

3.3.3 Seasonal comparison to averaged observational datasets

GEOS-Chem averaged seasonal model output for the year 2011 and observed abundances of C_2H_6, C_3H_8, and C_4-C_5 alkanes are shown in Figures 3.7, 3.8, and 3.9, respectively. Filled circles correspond to vertically averaged aircraft measurements (0 - 10 km above the ground) for each season, altitude range and horizontally every 1° × 1°. In the lower panel, filled circles represent seasonal averages of daily surface flask measurements and other surface observations as averages of their specific sampling period in the corresponding season when they occurred. Time periods and locations for each dataset are presented in Table 3.3. As can be noted in Table 3.3, there are very few aircraft observations for 2011. Figures 3.7-3.9 present aircraft observations from other years. Given high rates of change in C_2H_6 and C_3H_8 since 2009, particularly over the central U.S. [*Franco et al.,* 2016; *Helmig et al.,* 2016], directly comparing the model to these observations is a challenge. We use the observations and model to show seasonal, horizontal and vertical gradients in these species across the U.S. We provide comparisons where we can do so conservatively.

Across the U.S., there is a seasonal gradient in C_2-C_5 alkane mixing ratios due to the seasonal variations in OH concentrations; there are higher C_2-C_5 alkane mixing ratios during fall and winter compared to spring and summer (more on this in Section 3.3.2). Most of the aircraft observations (0-10 km) presented here were collected during summer months. In this season, the observations cover most of the CONUS. Although the aircraft campaigns occurred during different years (2008-2014), the almost full coverage of the CONUS provides an overview of the spatial distribution of C_2H_6, C_3H_8, and C_4-C_5 mixing ratios. C_2-C_5 alkane abundances are more homogenous above 2 km compared to the boundary layer. From 2 to 10 km, mixing ratios primarily reflect northern hemisphere background abundances, while in the boundary layer

enhancements mirror the spatial distribution of emissions. The largest boundary layer enhancements for these species occur over Colorado, Texas and Oklahoma. The model output shown in Figures 3.7-3.9 corresponds to seasonal averages of monthly means. The spatial distribution of tropospheric abundances is determined by the atmospheric lifetime of each C_2-C_5 alkane. Consequently in our model simulations, tropospheric abundances of C_2H_6 and C_3H_8 are more homogeneous across the CONUS compared to abundances of C_4-C_5, which show stronger local enhancements below 2 km.

In some regions (e.g., the Colorado Front Range), most of the observed abundances of C_4-C_5 alkanes have been attributed to oil and gas activities (Abeleira et al., 2017; Gilman et al., 2013). This is often diagnosed using the ratio of the isomers of butane and pentane (Pétron et al., 2014; Thompson et al., 2014). Enhanced abundances of the C_4-C_5 alkanes compared to background values are well documented over oil and gas regions (Abeleira et al., 2017; Johansson et al., 2014; Swarthout et al., 2013). For example in the Colorado Front Range, ratios of n-butane, i-pentane, and n-pentane to C_3H_8 in air masses impacted by oil and gas emissions have ranges of 0.43-0.56, 0.13-0.16, and 0.13-0.19, respectively (Gilman et al., 2013; Pétron et al., 2014; Swarthout et al., 2013). In addition to oil and gas sources, typical urban sources of C_4-C_5 alkanes are landfills (U.S. EPA, 2009) and traffic (Abeleira et al., 2017; Kirchstetter et al., 1996). A course comparison of observed vs. modeled regional vertical profiles and simulated oil and gas contribution is presented in Figure B1 in Appendix B.

A portion of the HIPPO flights did occur in 2011, but only two HIPPO flights (9 and 11 August 2011) cross our region of interest during this period. However, these two HIPPO flights allow us to make a direct comparison between aircraft observations and model output. We sampled the model at the coincident time and location of the observations from the HIPPO

flights on 9 and 11 August 2011. Model bias is largest for C_3H_8 and is only apparent below 700 hPa. The normalized mean model bias for C_3H_8 (NMB = sum(model – obs)/(sum(obs)) is -44% for the 9 August 2011 flight and -33% for the 11 August 2011 flight. This supports the conclusion drawn from the surface observation comparison for 2011 (Figure 3.8 lowest row) that the model underpredicts C_3H_8.

The SEAC^{4}RS observations cover a region of the U.S. where we would expect a large influence from emissions from the oil and gas sector. Though there are few long-term monitors in this region, it is likely that there were changes in the average abundance of C_2-C_5 in this region, even over the two-year period between 2011 (model output) and 2013 (SEAC^{4}RS observations). *Helmig et al.* [2016] summarize observed trends in C_2H_6 and C_3H_8 over the period 2009 – 2014 for long-term surface sites. The largest trend in C_3H_8 over the period 2009 – 2014 in this region in *Helmig et al.* [2016] is for Moody, Texas. The change is 286 pptv/year. The model, based on 2011 emissions and meteorology, underpredicts the observed 2013 SEACR^{4}S C_3H_8 mixing ratios over Texas below 2 km by ~ 1ppbv. Even a trend of 286 pptv/year applied to the SEAC^{4}RS data is unlikely to close the model-measurement gap. When we compare average model output to the SEAC^{4}RS observations, hypothetical de-trending would still result in an underprediction of the observations by >400 pptv. This rough calculation also supports the conclusion that C_3H_8 is underpredicted by the model.

There is a strong diurnal cycle in the mixing ratios of alkanes within the boundary layer (Abeleira et al., 2017; Vinciguerra et al., 2015). The model output in Figures 3.7-3.9 represents seasonal means, thus the model represents an average of the entire diurnal cycle over these seasons. In contrast, the majority of the aircraft observations were collected during the day when local emissions are mixed into a larger volume and reacting with OH. Despite this, the simulated

abundances of C_2H_6 and C_3H_8 at altitudes below 2 km are on average 5 and 3 ppb lower, respectively (both modeled and observed abundances are horizontally averaged every 1° × 1°, Figures 3.7 and 3.8). The discrepancy between the model and the observations is largest for the FRAPPÉ aircraft campaign, which is also the most recent field campaign presented in this study (2014) and encompasses the region with higher annual rates of change of C_2H_6 total column from 2009 to 2014 as estimated by Franco et al. (2016).

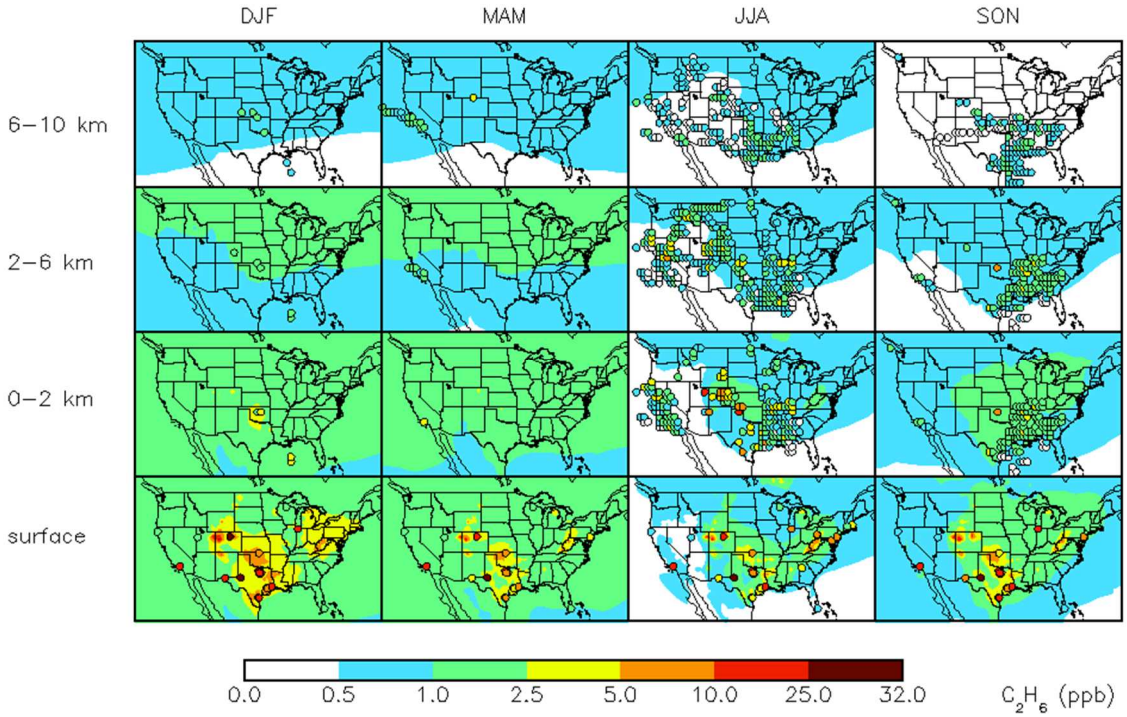


Figure 3.7: Mean distribution of C_2H_6 abundances for different seasons and altitude ranges compared to observations from aircraft campaigns and surface measurements (Table 3.3). The background contours are model outputs for 2011. The filled circles represent seasonally averaged observations. Aircraft measurements (0–2, 2–6, and 6–10 km) are averaged vertically for each altitude range and horizontally every 1° × 1°.

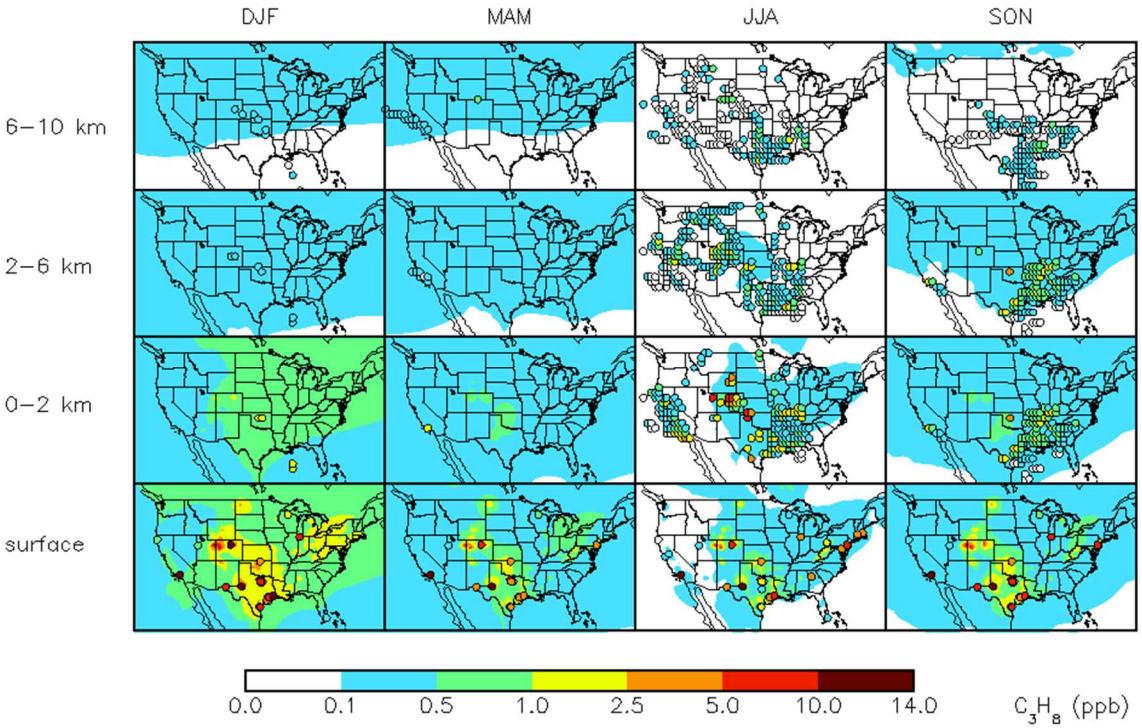


Figure 3.8: Mean distribution of C_3H_8 abundances for different seasons and altitude ranges compared to observations from aircraft campaigns and surface measurements (Table 3.3). The background contours are model outputs for 2011. The filled circles represent seasonally averaged observations. Aircraft measurements (0–2, 2–6, and 6–10 km) are averaged vertically for each altitude range and horizontally every 1° × 1°.

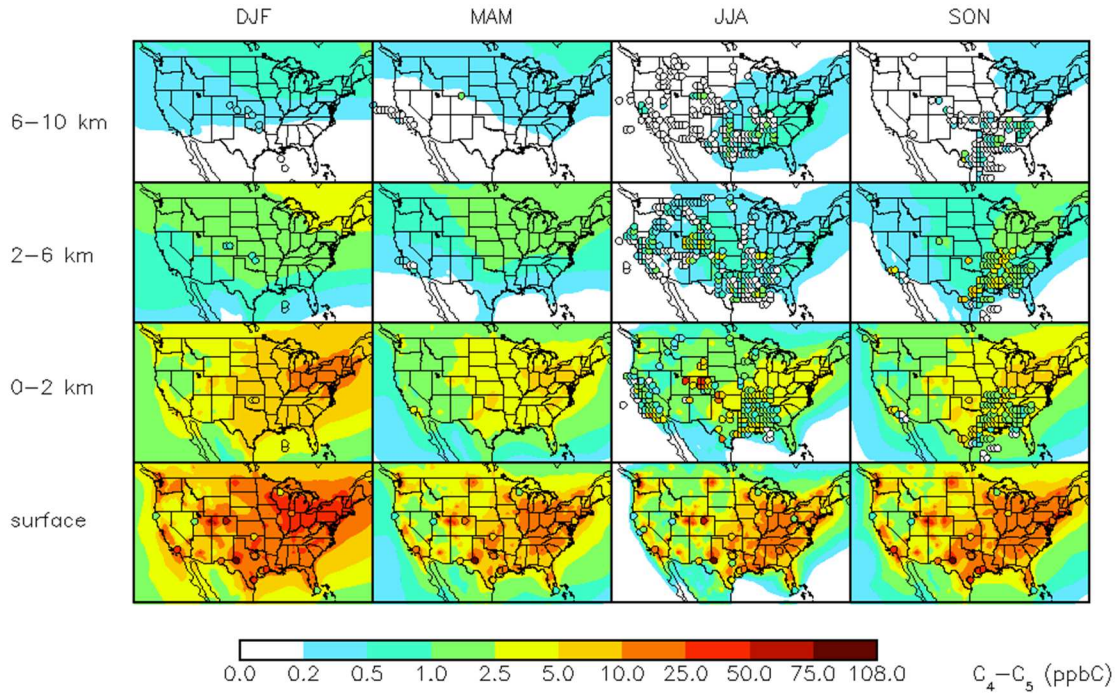


Figure 3.9: Mean distribution of C_4-C_5 alkane abundances for different seasons and altitude ranges compared to observations from aircraft campaigns and surface measurements (Table 3.3). The background contours are model outputs for 2011. The filled circles represent seasonally averaged observations. Aircraft measurements (0–2, 2–6, and 6–10 km) are averaged vertically for each altitude range and horizontally every 1° × 1°.

Figures 3.10-3.12 present the simulated percentage contribution from the oil and gas sector to total abundances of C_2-C_5 alkanes. We use the updated 2011NEI: OG off simulation to estimate the percentage contribution of emissions from this sector to total C_2-C_5 alkane mixing ratios.

Of the regions examined here, the lowest contribution of U.S. oil and gas activity to surface mixing ratios of C_2-C_5 alkanes is over California, which has relatively little local oil and gas development compared to the other regions of the U.S. Gentner et al. (2009) reported i-pentane/n-pentane ratios for California during summertime ranging from 2.9 for liquid gasoline to 3.8 for gasoline vapors. The June-July mean i-pentane/n-pentane ratios over this area for the 2008 ARCTAS aircraft campaign and 2011 surface flask observations at the LAX station were

2.0 and 2.2, respectively; suggesting that air masses are dominated by urban sources (Figure 3.11).

Over the central and southeastern U.S. the model attributes a higher percentage of the near-surface C_2H_6 and C_3H_8 to oil and gas related activities, compared to C_4-C_5 alkanes. Figure 3.11 shows that the model attributes most of the C_3H_8 at the surface to emissions from the oil and gas sector. The estimated oil and gas contribution to near surface C_2H_6 and C_3H_8 mixing ratios over Colorado is consistent with results from Gilman et al. (2013), who estimated mean percentage contributions of 72% and 90%, respectively. However, our estimated contribution from oil and gas sources to C_4-C_5 alkanes is only half of the Gilman et al. (2013) calculation of 93-96%. This difference suggests that our choice to assign C_4-C_5 alkanes as 36% of the total emitted PAR species over the U.S. for all emissions sources, should be revisited over regions where oil and gas activities abut urban areas, like Colorado. For this region, the percentage of C_4-C_5 alkanes from the oil and gas sector is likely much higher than 36%.

Figures 3.10, 3.11 and 3.12 suggest that emissions over oil and gas regions can impact atmospheric abundances over much of the U.S. lower-to-mid free troposphere. This does not imply that the atmosphere is well mixed over a given area from the surface to 10 km on a given day. Rather, it reflects typical characteristic time scales for vertical transport which are ~1 week for mixing in the lower free troposphere and ~1 month for mixing throughout the troposphere. The lifetimes of C_2H_6 and C_3H_8 are sufficiently long such that these species can be mixed vertically. Figures 3.10 and 3.11 reflect more vigorous mixing in summer months.

Figures 3.10, 3.11 and 3.12 suggest that emissions over oil and gas regions can impact atmospheric abundances over much of the U.S. lower-to-mid free troposphere. This does not imply that the atmosphere is well mixed over a given area from the surface to 10 km on a given

day. Rather, it reflects typical characteristic time scales for vertical transport which are ~1 week for mixing in the lower free troposphere and ~1 month for mixing throughout the troposphere. The lifetimes of C_2H_6 and C_3H_8 are sufficiently long such that these species can be mixed vertically. Figures 3.10 and 3.11 reflect more vigorous mixing in summer months.

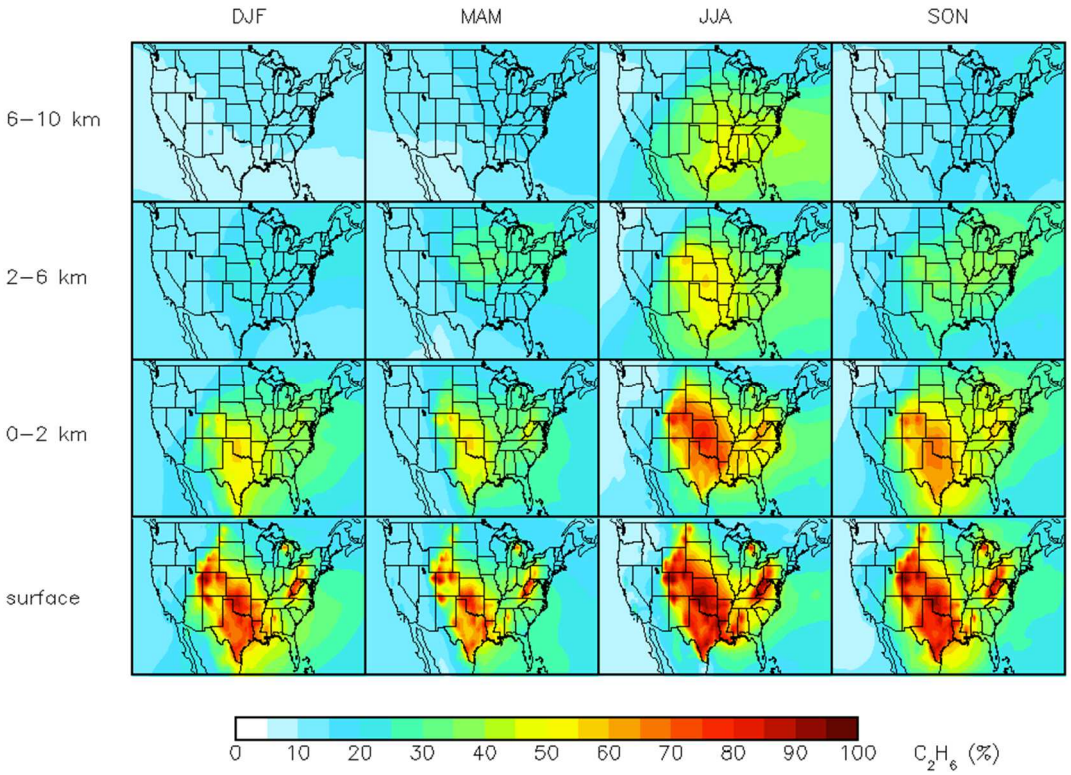


Figure 3.10: 2011 simulated percentage contribution from the oil and gas sector to total abundances of C_2H_6.

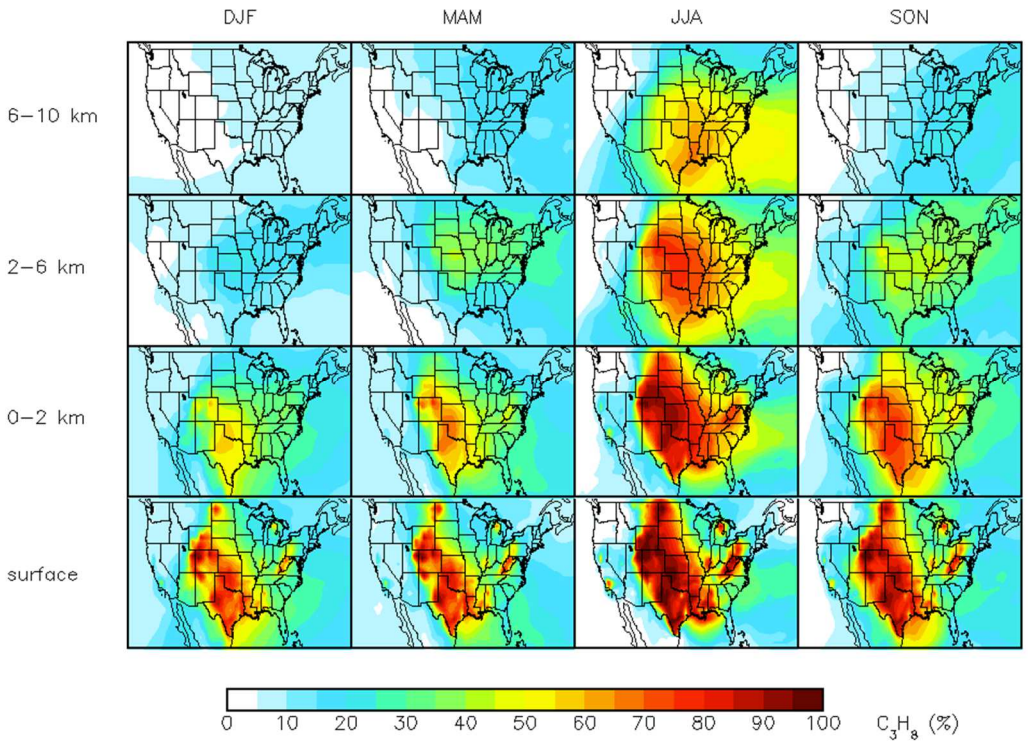


Figure 3.11: 2011 simulated percentage contribution from the oil and gas sector to total abundances of C_3H_8.

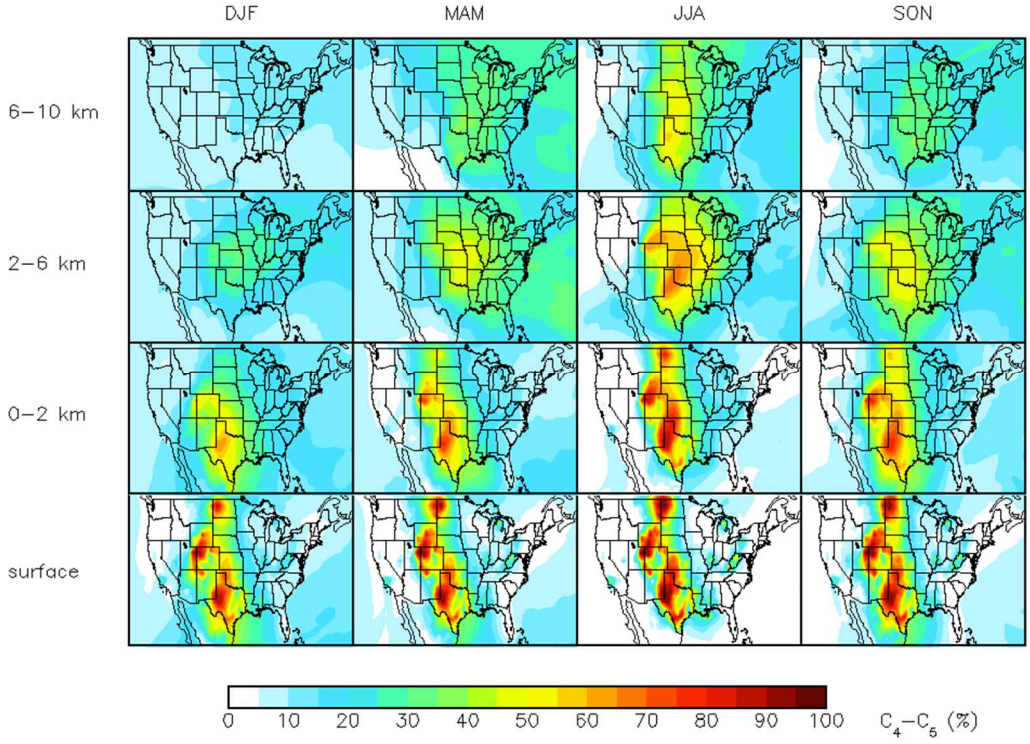


Figure 3.12: 2011 simulated percentage contribution from the oil and gas sector to total abundances of C_4-C_5 alkanes.

Figures 3.10, 3.11 and 3.12 suggest that emissions over oil and gas regions can impact atmospheric abundances over much of the U.S. lower-to-mid free troposphere. This does not imply that the atmosphere is well mixed over a given area from the surface to 10 km on a given day. Rather, it reflects typical characteristic time scales for vertical transport which are ~1 week for mixing in the lower free troposphere and ~1 month for mixing throughout the troposphere. The lifetimes of C_2H_6 and C_3H_8 are sufficiently long such that these species can be mixed vertically. Figures 3.10 and 3.11 reflect more vigorous mixing in summer months.

CHAPTER 4. IMPACTS TO U.S. AND GLOBAL SURFACE OZONE FROM OIL AND GAS ALKANE EMISSIONS[3]

4.1 Model Configuration

Here we again use GEOS-Chem model version 10-01 (Bey et al., 2001) driven by off-line GEOS-5 assimilated meteorological fields with 47 vertical levels. The emission inventory configuration used in this study is described in detail in Chapter 3. The U.S. anthropogenic emissions correspond to an updated version of the 2011NEI, which includes updated emission estimates from the oil and gas industry and adjustments to C_4-C_5 alkane emissions. In our model simulations, NO_X ($NO_X = NO + NO_2$) emission fluxes correspond to a U.S. NEI total of 3.6 Tg N (nitrogen) for the year 2011. This total is lower compared to modeled emissions of the 2011 NEIv1 of 4.3 Tg N, and similar to an unadjusted 2013 estimate of 3.5 Tg N by Travis et al. (2016). U.S. total emissions of C_2H_6, C_3H_8, and C_4-C_5 alkanes are 1.4 Tg, 0.9 Tg, and 7.0 TgC, respectively for the year 2011 in our baseline simulation. Our C_2H_6 total is ~50% higher compared to modeled emissions of the 2011 NEIv1 and ~10% higher compared to the estimates presented in Chapter 2. For C_3H_8, the U.S. total here is much higher compared to 2011 NEIv1 emissions (216% greater). Lastly, U.S. total emissions of C_4-C_5 alkanes are estimated as 36% of the total paraffin emissions. A detailed description of this calculation can be found in Chapter 3.

In this Chapter, we investigate the contribution of C_2-C_5 alkane emissions from the U.S. oil and gas industry to O_3 abundances at regional and global scales using two emission scenarios. The first emission scenario (baseline) uses a complete emission configuration, while in the second we turn off all C_2-C_5 alkane emissions from the U.S. oil and gas industry (OG-off). U.S.

[3] This chapter contains some published work from: Tzompa-Sosa, Z. A., et al. (2017), Revisiting global fossil fuel and biofuel emissions of ethane, Journal of Geophysical Research: Atmospheres, 122, doi:10.1002/2016JD025767.

impacts to O_3 abundances were estimated using two nested simulations (0.5°x0.6°) over North America (40° to 140° W and 10° to 70° N) for the year 2011. Each nested simulation uses boundary conditions from a 2°x2.5° simulation with the same emission inventory configuration as the nested runs. We use the 2°x2.5° simulations to calculate global impacts from the U.S. oil and gas industry to global and hemispheric O_3 abundances.

4.2 Results

4.2.1 Modeled U.S. Emissions of O_3 Precursors

Anthropogenic NO_X and biogenic VOCs

Our model simulation includes seven different anthropogenic and biogenic sources of NO_X. Anthropogenic sources (including aircraft, biofuels, and fertilizers) are responsible for 72% of the total U.S. emissions of NO_x. Emissions of NO_x are largest in urban areas. Consequently, NO_x emissions show a west-to-east gradient with higher emissions over the eastern U.S., where more urban areas are located (Figure 4.1a). However, some regions with high NO_X emissions also correspond to semi-rural areas where oil and gas extraction is abundant (e.g. Northern Front Range in Colorado, Texas Panhandle, Uintah Basin in Utah, and Upper River Basin in Wyoming). Even though NO_X emissions from the oil and gas sector only represent 4.7% of the total U.S. NO_X emissions, these emissions have increased by 94% between 2002 and 2011 compared to a 40% decrease of total anthropogenic emissions (Allen, 2016). Figure 4.1a shows monthly average NO emissions for the year 2011. NO emission fluxes have low seasonality in terms of their spatial distribution and magnitude.

Biogenic VOCs are important O_3 precursors in some U.S regions. The terrestrial biosphere is the largest source of VOCs in summer over the U.S.; the highest emission fluxes

occur over the southeastern U.S. and they peak during summer months (Geron et al., 1997; Guenther et al., 2006; Sindelarova et al., 2014; Ying et al., 2015). To illustrate the typical spatial distribution of biogenic emissions, Figure 4.1b shows summertime (JJA) isoprene emission fluxes.

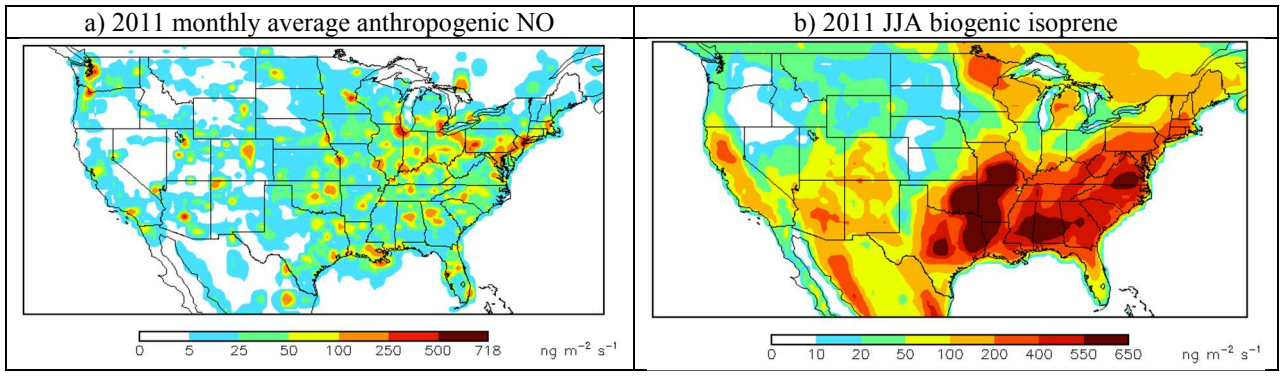


Figure 4.1: 2011 monthly average U.S. anthropogenic emission fluxes of NO (left panel) and summertime biogenic emissions of isoprene (right panel).

Anthropogenic C_2-C_5 alkanes and contribution from the oil and gas sector

Anthropogenic emission fluxes of C_2-C_5 alkanes have low seasonality and occur across the entire U.S. (upper left panel on Figure 4.2). The oil and gas sector is a major source of C_2-C_5 alkanes. Emissions of these species from the oil and gas have also low seasonality and are mainly clustered over the central U.S. (upper right panel on Figure 4.2). However, while oil and gas sources dominate emissions of C_2H_6 and C_3H_8, urban sources (e.g. automobile combustion and fugitive emissions from gasoline and diesel distribution) are the primarily emission sources of C_4-C_5 alkanes at a national level. Interestingly, inside oil-and-gas-producing basins C_2-C_5 alkane emissions from this sector dominate local anthropogenic (urban and rural) emission fluxes of C_2-C_5 alkanes with contributions >70% (lower panel, Figure 4.2). Our estimates of the oil and gas contribution to local total anthropogenic emissions of C_2-C_5 alkanes are similar to

conclusions drawn from other studies inside different U.S. oil and gas basins (Gilman et al., 2013; Peischl et al., 2016; Warneke et al., 2014).

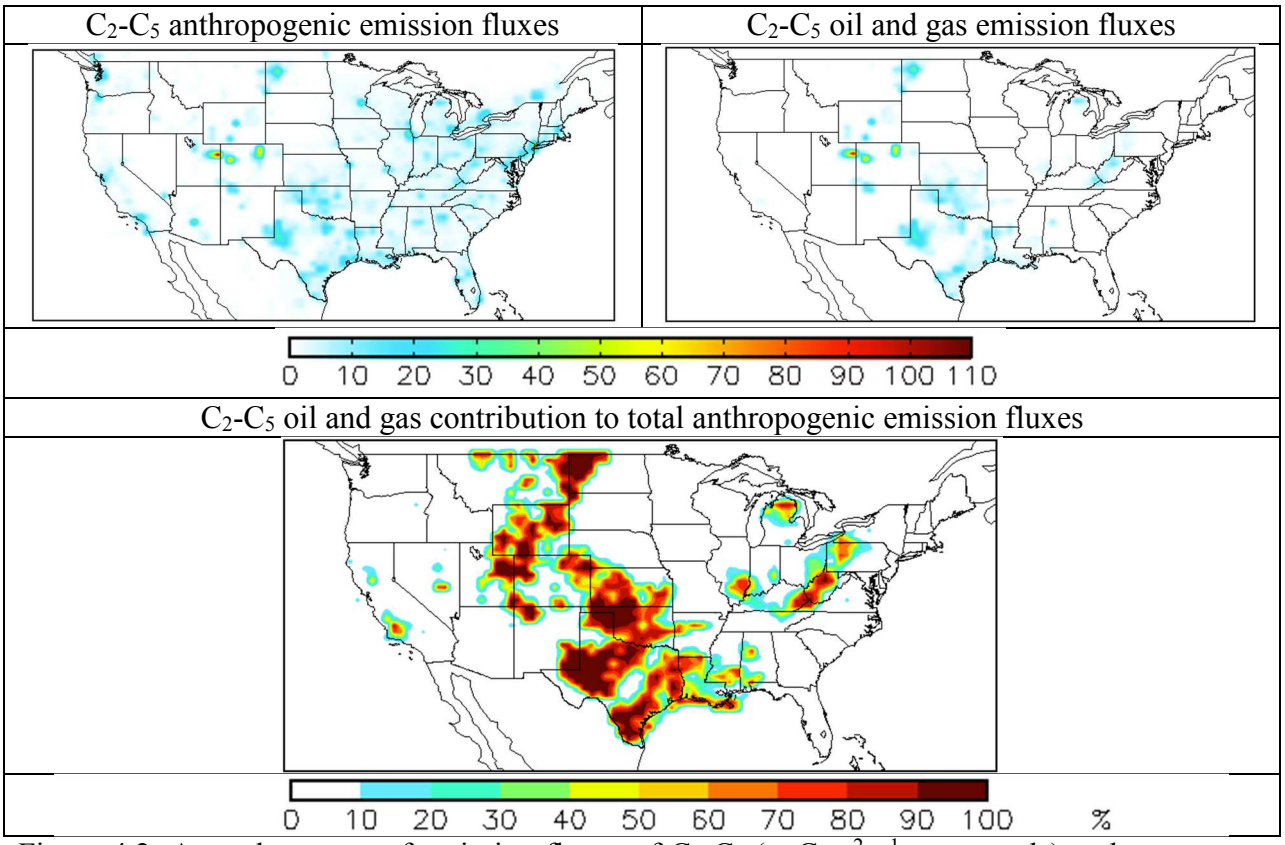


Figure 4.2: Annual average of emission fluxes of C_2-C_5 (ngC m^2 s^{-1}, top panels) and percentage contribution of oil and gas emission sources to total anthropogenic fluxes of C_2-C_5 (lower panel).

4.2.2 Modeled daytime O_3 mixing ratios at the surface

The simulation described on the previous section, results in a consistent nation-wide overestimation of daytime O_3 mixing ratios at the surface. We compared our simulated surface O_3 mixing ratios to Air Quality System (AQS, https://www.epa.gov/aqs) data, and the normalized mean bias (NMB) for each month ranges from 8.6% to 41.1%. The upper panel of Figure 4.3 shows a comparison between modeled and observed daytime surface O_3 mixing ratios for the month of August, which is the month with not only the highest averaged daytime surface

O_3 bias (17.2 ppb), but also with the highest tropospheric O_3 throughout the column over the CONUS. Travis et al. (2016) investigated O_3 mixing ratios in the southeastern U.S. Scaling the 2011NEIv1 to 2013 (multiplying NO_X emissions by 0.89), they found a 26-31% bias when comparing modeled versus observed O_3 vertical profiles. After reducing NO_X emissions of non-power-plant NEI emissions by 60%, their O_3 column bias was reduced to a NMB of 4.5%. They estimated that total U.S. anthropogenic fuel NO_X emissions were between 1.7 and 2.6 Tg N for the year 2013. We note that between the 2011NEIv1 (scaled to 2013 in their study) and the updated 2011NEI used in this study, there are important changes in U.S. anthropogenic O_3 precursors (see Table C1 in Appendix C). We tested a reduction of NO_X emissions considering the total U.S. NO emission range estimated by Travis et al. (2016). We reduced the non-power plant NO_X emissions (excluding commercial marine vessels and railroad emissions) in the updated 2011NEI and kept the rest of the emissions as in our baseline simulation. The resulting total U.S. anthropogenic NO_X emissions in this iteration was 2.1 Tg N. However in our baseline simulation, we obtain increased surface O_3 mixing ratios throughout the year; thus, a higher bias compared to observations. Our larger surface O_3 mixing ratios is the combined result of increased of NO and VOC emissions from the updated version of the 2011NEI used in this study compared to 2011NEIv1 used by Travis et al. (2016).

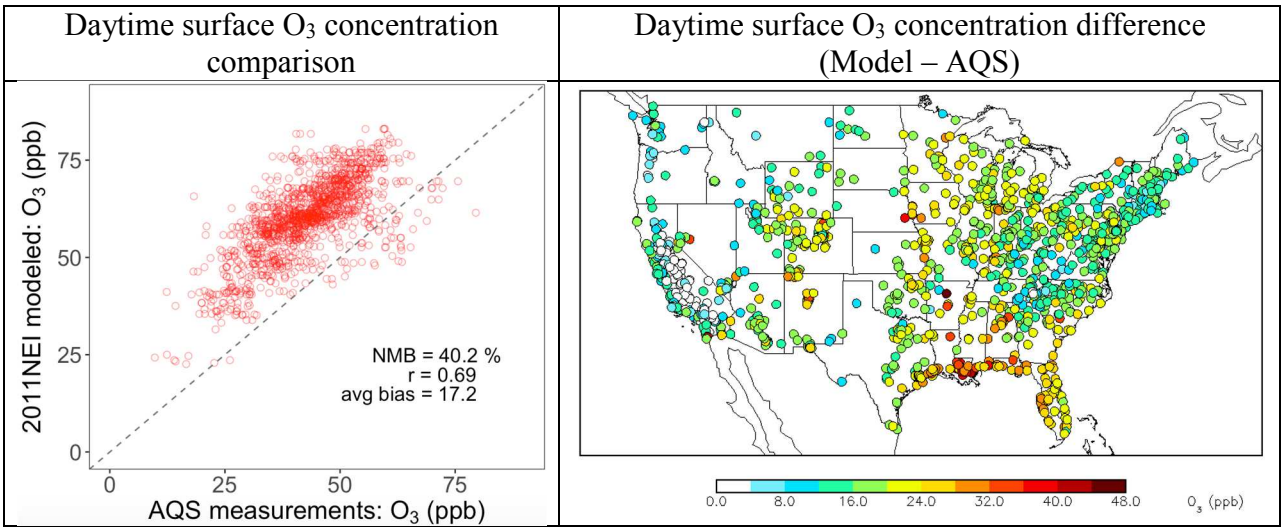


Figure 4.3: Comparison between observed and modeled averaged daytime 3-hour instantaneous output of surface O_3 mixing ratios for August 2011. In the left figure, the dash line represents 1:1 line, and the correlation coefficient (r) and normalized mean bias (NMB) are also shown. In the right panel, the filled circles represent the locations of AQS network sampling stations with 2011 O_3 data. The color of each filled circle represents the difference between modeled and observed (Model – AQS) surface O_3 mixing ratios.

4.2.3 Impact of C_2-C_5 alkane emissions from Oil and Gas on U.S. O_3 abundances

The difference between the baseline simulation and the simulation without emissions of C_2-C_5 alkanes from the oil and gas sector provides an estimate of the impact of these emissions on O_3 mixing ratios throughout the column. Figure 4.4 shows seasonal average O_3 enhancements in the boundary layer (defined here as the average of model levels below 2 km). The impact of these emissions on boundary layer O_3 mixing ratios varies by season and altitude. Higher O_3 enhancements occur during summertime over the central U.S., and these enhancements are largest below 1 km. Maximum enhancements occur over major oil and gas-producing areas in eastern Colorado, Kansas, and the Texas-Oklahoma Panhandle.

The highest surface O_3 enhancements due to oil and gas C_2-C_5 alkane emissions occur in August. Figure 4.5 shows average daytime surface O_3 enhancements from 3-hr instantaneous mean model output for August 2011. Blue contour areas correspond to 8-Hr O_3 non-attainment

areas (2008 standard; includes all classifications: Marginal, Moderate, Serious, Severe 15, Severe 17, and Extreme). The central part of the U.S. is the most impacted region with monthly daytime average O_3 enhancements between 2 and 3 ppb. Figure 4.6 compares August 2011 daytime O_3 enhancements over three O_3 non-attainment areas located inside major oil and gas-producing basins. The model produces daytime O_3 enhancements in the Northern Colorado Front Range > 4 ppb. Outside the central part of the U.S., the enhancements are lower. The model produces lower O_3 enhancements from oil and gas C_2-C_5 alkane emissions over both Dallas, TX and Pittsburgh, PA compared to the Colorado Front Range; the largest simulated 3-hour instantaneous O_3 enhancements in these two other regions are 1.5 ppb and 0.7 ppb, respectively.

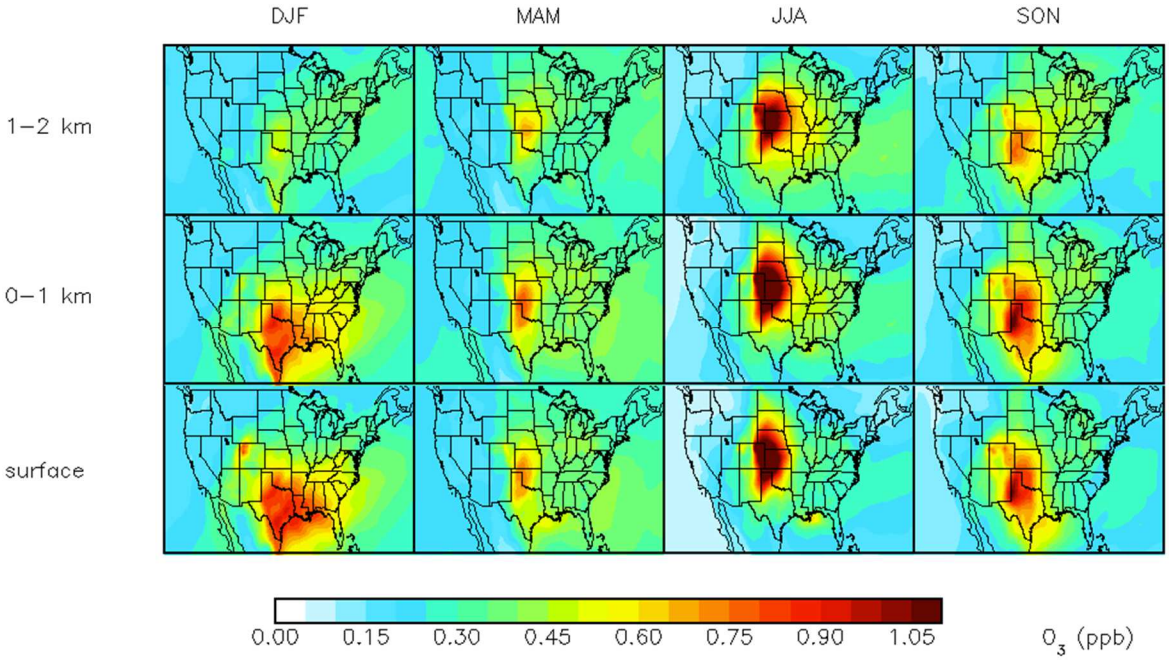


Figure 4.4: 2011 seasonal mean O_3 enhancements driven by emissions of C_2-C_5 alkanes from the U.S. oil and gas industry.

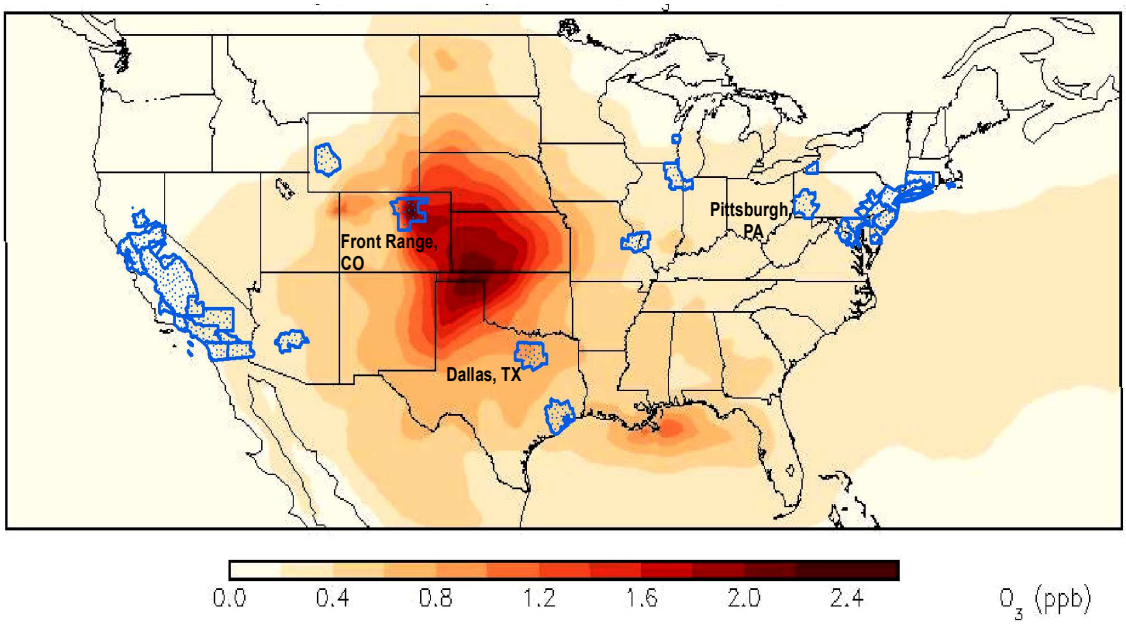


Figure 4.5: August 2011 daytime average O_3 enhancements (calculated from simulated 3-hr instantaneous means) due to C_2-C_5 alkanes emitted by the U.S. oil and gas sector. Blue contoured areas correspond to 8-Hr O_3 non-attainment areas (2008 standard; includes all classifications: Marginal, Moderate, Serious, Severe 15, Severe 17, and Extreme).

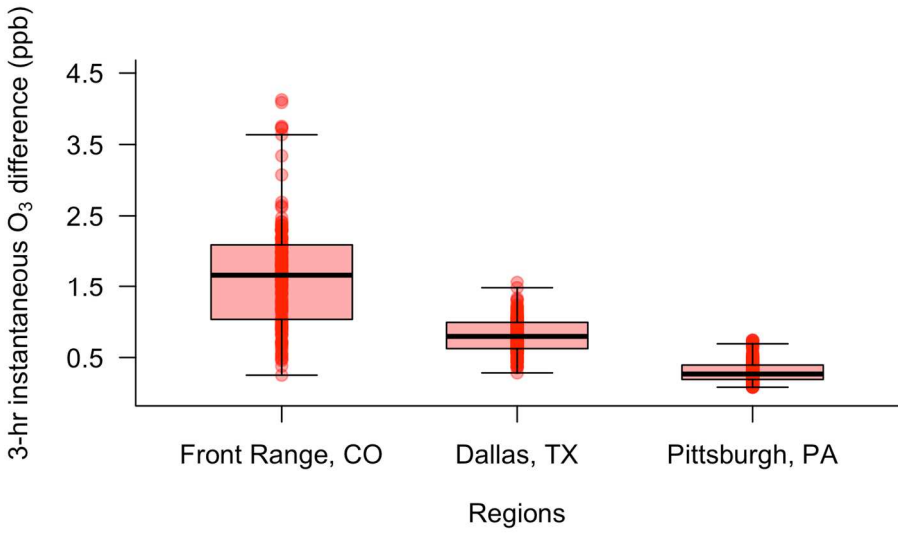


Figure 4.6: August 2011 3-hour daytime instantaneous O_3 enhancements over three O_3 non-attainment areas located inside important oil and gas-producing basins.

4.2.4 Contribution of U.S. C_2-C_5 alkane emissions from the oil and gas industry to the hemispheric O_3 burden

We estimate the contribution of C_2-C_5 alkane emissions from the oil and gas sector to the hemispheric O_3 burden using two 2°x2.5° global simulations based on the baseline emission scenario and the scenario without emissions of these species from the oil and gas sector (OG-off). The global tropospheric contribution of C_2-C_5 alkane emissions from the U.S. oil and gas sector to the O_3 burden is 0.5 Tg for the year 2011, which represents 0.17% of the global tropospheric O_3 burden. Due to the lifetime of C_2-C_5 alkanes, the highest O_3 enhancements are located over the Northern Hemisphere, where their contribution to the tropospheric O_3 burden is 0.27%. The highest contribution to O_3 surface mixing ratios over land areas outside the U.S. occurs during August over northern Africa, where the enhancement is up to 0.32 ppb (Figure 4.7). Although small, we note that these contributions to the O_3 burden result from VOCs emitted from oil and gas activities that are not considered to be efficient at producing O_3 (Russell et al., 1995). Future work should focus on the impact of more reactive compounds emitted alongside the C_2-C_5 alkanes.

This study does not include tropospheric halogen chemistry. Sherwen et al. (2016) study the tropospheric impacts of Cl, Br, and I chemistry and found that adding a chlorine sink decreases the tropospheric global burdens of C_2-C_5 alkanes by 12% to 19%, depending on the specie. Their global tropospheric burdens were more impacted over the oceans compared to land areas. In this study we use a configuration of emission inventories with higher U.S. emissions of C_2-C_5 alkanes, thus further examination of the impacts of halogen chemistry on C_2-C_5 alkanes global burdens and their impacts on the global tropospheric burden of O_3 is needed.

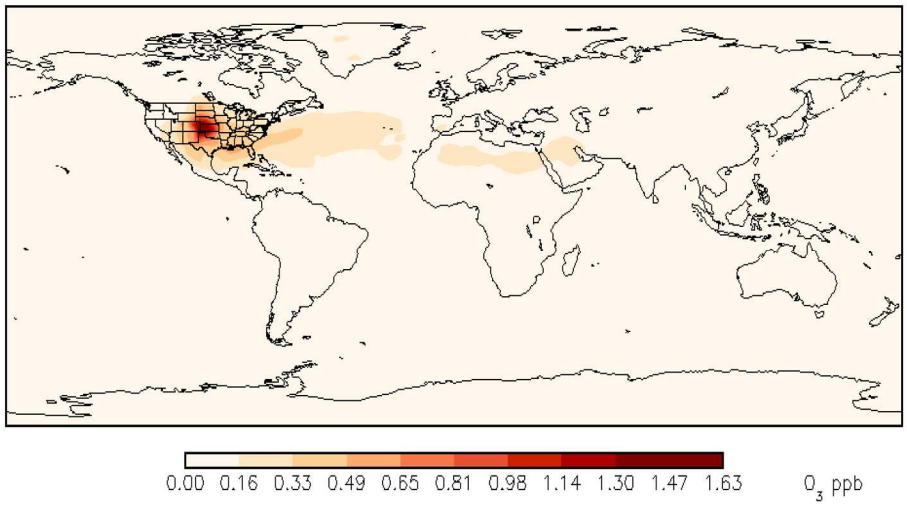


Figure 4.7: Averaged surface O_3 enhancements due to U.S. emissions of C_2-C_5 alkanes from the oil and gas sector for August 2011.

4.2.5 Global contribution of fossil fuel C_2H_6 emissions to O_3 and PAN mixing ratios

In section 4.2.4 we examine the impact of emissions of C_2-C_5 alkanes from the U.S. oil and gas sector to the global O_3 burden. In addition to the previous analysis we also examine the contribution of global C_2H_6 emissions on the global O_3 burden. These results were derived using the simulations developed in Chapter 2.

Ethane is a precursor of carbon monoxide (CO), O_3, and peroxyacetyl nitrate (PAN) in the troposphere (Aikin et al., 1982). C_2H_6 degradation can lead to the production of O_3 via two pathways: 1) C_2H_6 oxidation by OH radicals in the presence of nitrogen oxide radicals (NO_x = NO + NO_2), and 2) by serving as a precursor for PAN. PAN acts as a reservoir for NO_x (Aikin et al., 1982; Fischer et al., 2014), and its thermal decomposition over remote areas can efficiently produce O_3 (Fischer et al., 2011). C_2H_6 impacts the distribution of several atmospherically relevant species due to its main removal process via reaction with OH radicals (Blake and Rowland, 1986); however, this impact is smaller compared to other species such as CO, CH_4, and isoprene.

In this section, we estimate the contribution of C_2H_6 fossil fuel sources to the global burden and surface mixing ratios of O_3 and PAN rather than focus on specific oil and gas producing regions. This estimate is based on the comparison of a 2°x 2.5° degree simulation without fossil fuel sources of C_2H_6 to the results produced using the 2010 C_2H_6 emissions presented in Chapter 2. The global contribution of fossil fuel C_2H_6 emissions to O_3 and PAN surface mixing ratios has a strong inter-hemispheric gradient. Due to the C_2H_6 lifetime, which allows its transport to remote areas, the largest impacts on surface O_3 and PAN occur over regions with low emissions of highly reactive hydrocarbons (Figures 4.8 and 4.9). The highest impacts on O_3 and PAN surface concentrations occur over the Northern Hemisphere. For O_3, the highest contribution of fossil fuel C_2H_6 emissions to surface mixing ratios is 0.58 ppbv during spring and summertime. Over land areas of the mid-latitude Northern Hemisphere, fossil fuel C_2H_6 emissions increase annual average O_3 mixing ratios at the surface by 0.4 ppbv (~1%). The effect is slightly smaller (0.3 ppbv) from 50°-70°N. For PAN, the highest contributions to surface mixing ratios occur during spring (up to 30 pptv). Fossil fuel C_2H_6 emissions enhance mid-latitude Northern Hemisphere continental PAN mixing ratios up to 26 pptv, with an average contribution of ~8% at the surface. The impacts of C_2H_6 oxidation on atmospheric mixing ratios of O_3 and PAN in the free troposphere are more homogeneous across all longitudes, but similar in magnitude to the impacts near the surface. As C_2H_6 emissions increase due to fossil fuel sources like oil and natural gas activities, we anticipate greater contributions to both average mixing ratios of O_3 and PAN. Given the similar emission totals between both C_2H_6 emission inventories, our estimated contribution to global PAN annual burden is consistent with the Fischer et al. (2014) estimate of 6%, which was based on the Xiao et al. (2008) C_2H_6 emission inventory implemented in GEOS-Chem v.9.01.01.

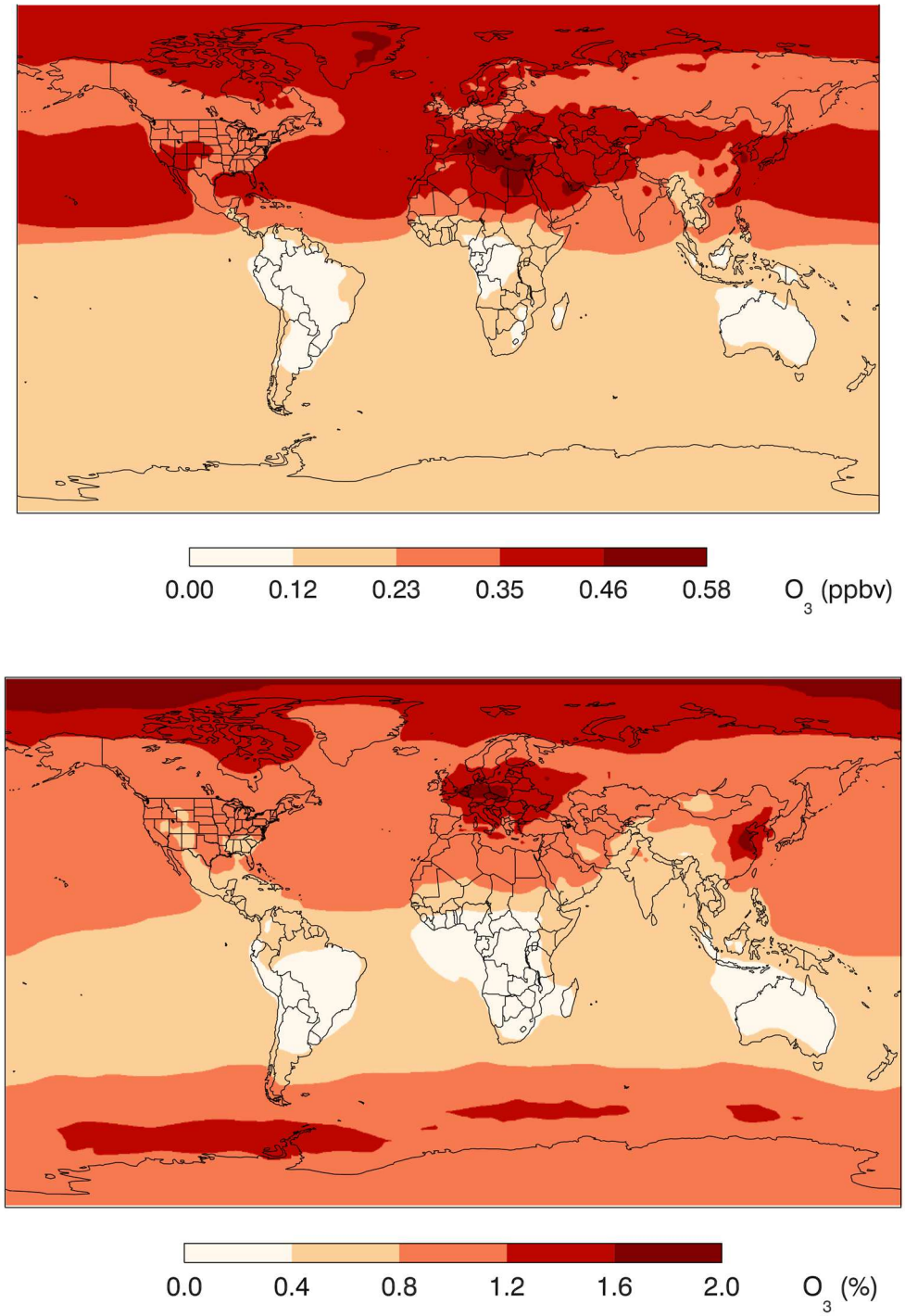


Figure 4.8: Absolute (top) and percent (bottom) averaged annual contribution of oxidation of C_2H_6 from fossil fuel sources to surface O_3 mixing ratios. Modeled C_2H_6 fossil fuel sources correspond to the 2010 C_2H_6 emission inventory.

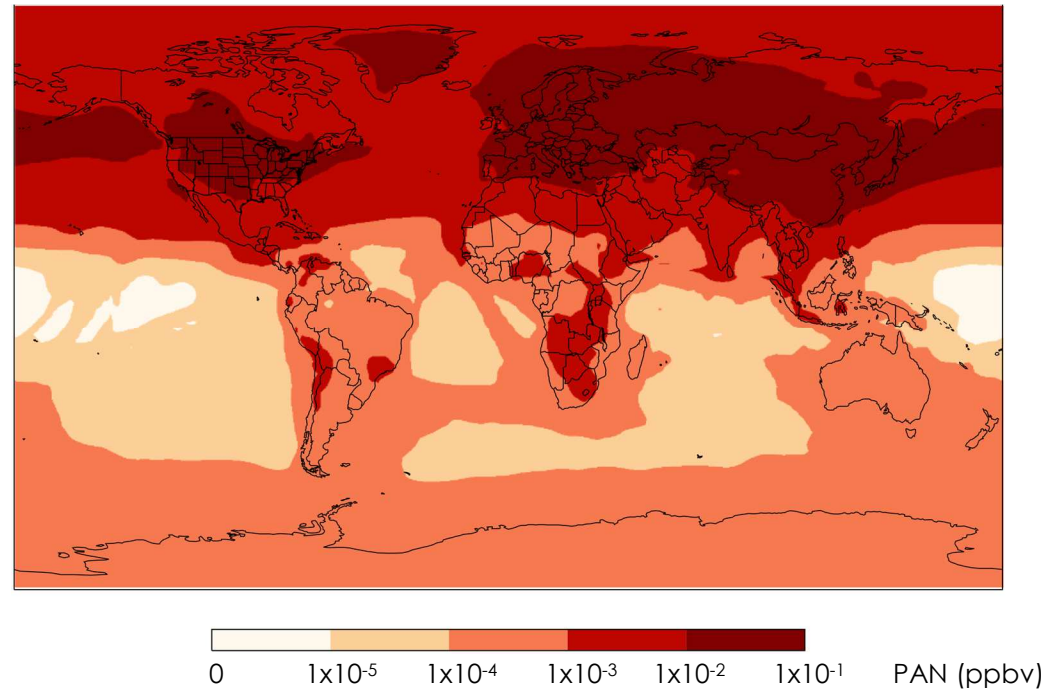


Figure 4.9: Absolute (top) and percent (bottom) averaged annual contribution of oxidation of C_2H_6 from fossil fuel sources to surface mixing ratios of PAN. Modeled C_2H_6 fossil fuel sources correspond to the 2010 C_2H_6 emission inventory.

CHAPTER 5. CONCLUSIONS AND FUTURE WORK[4]

In Chapter 2, we update a global simulation of C_2H_6 in the GEOS-Chem model by implementing a global C_2H_6 emission inventory estimated from CH_4 fluxes derived from satellite observations and a regional U.S. emission inventory derived by adjusting C_2H_6 emissions from the 2011NEIv1 upward. We contrast two global C_2H_6 emission inventories for the years 2001 and 2010. We show that these C_2H_6 emission inventories have similar emission totals, but very different spatial distributions. In particular, the distribution of emissions differs over the U.S., Europe, Russia, and the Middle East. Our 2010 C_2H_6 emission inventory, which includes C_2H_6 emissions from 2010 satellite-derived CH_4 fluxes and adjusted C_2H_6 emissions from 2011NEIv1, produces C_2H_6 emissions that are systematically larger over intense gas-producing regions and systematically lower over regions with low natural gas production compared to 2001 C_2H_6 emissions. Globally, the fossil fuel C_2H_6 emissions in 2010 decrease by 0.8 Tg compared to 2001. This difference is consistent with the long-term global decline over this period ending in 2009 (Franco et al., 2015; Helmig et al., 2014a).

When compared to a suite of global observations of C_2H_6, the model simulations capture the C_2H_6 seasonal cycle, the inter-hemispheric and vertical gradients, surface mixing ratios, and the C_2H_6 columns in most regions. However, over some intensive natural gas production regions over the U.S., aircraft measurements reveal greater C_2H_6 mixing ratios compared to the model, especially below 2 km. Given the reported strong increasing trend of the C_2H_6 atmospheric burden that started in 2009 (Franco et al., 2015; Helmig et al., 2016) and an estimated increase of anthropogenic emissions in North America of 75% from 2008 to 2014 (Franco et al., 2016), one

[4] This chapter contains published work from: Tzompa-Sosa, Z. A., et al. (2017), Revisiting global fossil fuel and biofuel emissions of ethane, Journal of Geophysical Research: Atmospheres, 122, doi:10.1002/2016JD025767.

plausible reason for this discrepancy could be the time difference between the measurements (2013-2014) and the 2010 C_2H_6 emissions implemented in the model. For these potentially fast-changing emission areas, large assumptions of continental or nationwide C_2H_6/CH_4 emission ratios or emission inventory scaling, are not likely to accurately represent the amount, distribution, and mixing ratio impacts of major local sources. Due to the limited observations and the scarcity of long-term in situ C_2H_6 measurements within or downwind of oil and natural gas producing regions, we recommend the use of different approaches to estimate C_2H_6 emissions for a particular region/basin depending on the type of data available (CH_4/C_2H_6 enhancement ratios, natural gas composition, etc.).

In Chapter 3, we use a GEOS-Chem nested simulation driven by updated 2011NEI emissions in combination with a collection of observations over the U.S. to 1) document the spatial patterns in observed atmospheric abundances of C_2-C_5 alkanes, and 2) estimate the contribution of the U.S. oil and gas industry to the observed patterns. The updated 2011NEI, which includes updates over U.S. oil-and-gas-producing basins, indicates that the oil and gas sector dominated U.S. emissions of C_2H_6 and C_3H_8 with a contribution to total emissions of 89% and 82%, respectively (U.S. EPA, 2017). Emissions of these two species are clustered inside U.S. oil and gas basins. As implemented in GEOS-Chem, oil and gas sources represent the third most important emission source for C_4-C_5 alkanes. Other fossil fuel sources contribute significantly to the emissions of these larger alkanes, thus their emissions are located not only inside oil-and-gas-producing basins, but also within urban and industrial areas.

Aircraft observations over the period 2008-2014 show that the highest mixing ratios of C_2-C_5 alkanes were encountered over the central U.S. boundary layer (mainly over Colorado, Texas and Oklahoma) during this period. Observations were much more homogenous above 2

km for all the species considered here. Both, the suite of observations and modeled C_2-C_5 alkane abundances, show that U.S. oil and gas emissions impact large regions of the lower troposphere especially over the central and eastern U.S. The surface and limited aircraft observation-model comparisons for C_3H_8 suggest that the emissions of C_3H_8 in the updated 2011NEI may continue to be too low.

Given that increases in C_2-C_5 alkane abundances driven by emissions from the U.S. oil and gas industry began in 2009, we do not recommend using the updated 2011NEI for prior years. There are many locations where oil and gas development is relatively recent. Similarly, the updated 2011NEI precedes much of the extraction of oil and gas in the Bakken. Thus if simple scaling factors were to be applied to this inventory for simulations after 2011, we would not expect that the resulting emissions would represent this area well. Furthermore, the reported increasing trends in atmospheric concentrations of oil and natural gas related emissions during 2010-2015 (Franco et al., 2016; Helmig et al., 2016; Vinciguerra et al., 2015), suggest that the C_2-C_5 alkane emission estimates in this paper are likely a low estimate for years following 2011.

Due to the increasing importance of oil and gas emissions in the U.S., long-term measurements of C_2-C_5 alkanes are needed in order to document how the emissions of these species are changing. We recommend continued support of existing long-term measurements of C_2-C_5 alkanes. We also suggest continuous consistent monitoring of surface mixing ratios in northeastern Colorado, Wyoming and North Dakota. Further, we suggest that the community evaluate whether chemical mechanisms that lump larger alkanes are sufficient to understand air quality issues in regions with large emissions of these species.

In Chapter 4 we show that C_2-C_5 alkane emissions from the oil and gas sector produce the highest O_3 enhancements during summertime over the central U.S. with maximum enhancements

over major oil and gas-producing areas in eastern Colorado, Kansas, and the Texas-Oklahoma Panhandle. These O_3 enhancements are largest below 1 km. Simulated August 2011 daytime average O_3 enhancements show that the Colorado Front Range is the most impacted 8-Hr O_3 non-attainment area in the U.S. with enhancements > 4 ppb.

Additionally, we find that global fossil fuel C_2H_6 emissions make the largest relative contributions to O_3 and PAN over remote areas without large emissions of highly reactive hydrocarbons. Over continental areas in the Northern Hemisphere, we estimate an average increase of ~1% and ~8% to mean annual O_3 and PAN surface mixing ratios, respectively, due to fossil fuel C_2H_6 emissions. On a global scale, these results appear to be largely insensitive to the distribution of C_2H_6 emissions over North America. These contributions from C_2H_6 oxidation to O_3 and PAN abundance are expected to be greater in years following 2010 due to increased emissions from oil and natural gas extraction over the U.S. over this period (U.S. EPA, 2015).

As the number of C_2-C_5 alkane observations increases, both the global C_2H_6 emission inventory presented in Chapter 2 and the emissions from the updated 2011NEI presented in Chapter 3, could be used as the prior emission vectors in inverse modeling studies to improve emission flux estimates. The two emission inventories developed in this dissertation produce mixing ratios that better represent observations. Therefore, the use of these inventories in inverse modeling studies is expected to produce lower characterization errors (Turner and Jacob, 2015).

Despite the enhanced emission fluxes of C_2-C_5 alkanes over oil and gas-producing regions, these alkanes represent just a small fraction of the total number of compounds emitted by the oil and gas industry. Furthermore, C_2-C_5 alkanes have low efficiency at producing O_3 compared to other VOCs emitted by this industry. Thus, future work should focus on investigating the atmospheric and air quality implications of other highly reactive compounds emitted inside oil

and gas basins. Additionally, investigating other VOCs can provide more data related to the use of C_2-C_5 alkanes as tracers for other gases. The scientific modeling community can benefit from the development of VOC ratios (e.g. CH_4/C_2H_6 ratio used in Chapter 2) for urban, rural, and oil and gas impacted air masses, especially in regions of the world where observational data is scarce. Finally, as pointed out in Chapters 2 and 3, emission inventories are challenged by emerging and rapidly changing sources like the oil and gas sector. The use of oil and gas VOC ratios can also help develop updated emission factors, which are widely used in emission inventories. Given the wide range of VOCs emitted by the oil and gas industry, both criteria and greenhouse gas emission inventories can benefit from the development of such emission factors. Efforts not only in the U.S., but around the world will benefit from having a better understanding of emission sources that have effects on air quality and climate change.

REFERENCES

Abeleira, A., Pollack, I.B., Sive, B., Zhou, Y., Fischer, E.V., and Farmer, D.K. (2017). Source characterization of volatile organic compounds in the Colorado Northern Front Range Metropolitan Area during spring and summer 2015. Journal of Geophysical Research: Atmospheres.

Aceves, M., and Grimalt, J.O. (1993). Seasonally dependent size distributions of aliphatic and polycyclic aromatic hydrocarbons in urban aerosols from densely populated areas. Environmental science & technology *27*, 2896-2908.

Aikin, A.C., Herman, J.R., Maier, E.J., and McQuillan, C.J. (1982). Atmospheric chemistry of ethane and ethylene. Journal of Geophysical Research *87*, 3105.

Akagi, S.K., Yokelson, R.J., Wiedinmyer, C., Alvarado, M.J., Reid, J.S., Karl, T., Crounse, J.D., and Wennberg, P.O. (2011). Emission factors for open and domestic biomass burning for use in atmospheric models. Atmospheric Chemistry and Physics *11*, 4039-4072.

Allen, D.T. (2016). Emissions from oil and gas operations in the United States and their air quality implications. J Air Waste Manag Assoc *66*, 549-575.

Atkinson, R. (1991). Kinetics and Mecanisms of the Gas-Phase Reactions ot the NO3 Radical with Organic Compounds. Journal of Physical Chemistry *20*, 459-507.

Atkinson, R., Baulch, D.L., Cox, R.A., Crowley, J.N., Hampson, R.F., Hynes, R.G., Jenkin, M.E., Rossi, M.J., and Troe, J. (2006). Evaluated kinetic and photochemical data for atmospheric chemistry: Volume II - gas phase reactions of organic species. Atmospheric Chemistry and Physics *6*, 3625-4055.

Aydin, M., Verhulst, K.R., Saltzman, E.S., Battle, M.O., Montzka, S.A., Blake, D.R., Tang, Q., and Prather, M.J. (2011). Recent decreases in fossil-fuel emissions of ethane and methane derived from firn air. Nature *476*, 198-201.

Barletta, B., Meinardi, S., Sherwood Rowland, F., Chan, C.-Y., Wang, X., Zou, S., Yin Chan, L., and Blake, D.R. (2005). Volatile organic compounds in 43 Chinese cities. Atmospheric Environment *39*, 5979-5990.

Basevich, V.Y., Belyaev, A.A., Medvedev, S.N., Posvyanskii, V.S., and Frolov, S.M. (2012). Oxidation and combustion mechanisms of paraffin hydrocarbons: Transfer from C1-C7 to C8H18, C9H20, and C10H22. Russian Journal of Physical Chemistry B *5*, 974-990.

Bey, I., Jacob, D.J., Yantosca, R.M., Logan, J.A., Field, B.D., Fiore, A.M., Li, Q., Liu, H.Y., Mickley, L.J., and Schultz, M.G. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. Journal of Geophysical Research *106*, 23073.

Bi, X., Sheng, G., Peng, P.a., Chen, Y., and Fu, J. (2005). Size distribution of n-alkanes and polycyclic aromatic hydrocarbons (PAHs) in urban and rural atmospheres of Guangzhou, China. Atmospheric Environment *39*, 477-487.

Blake, D.R., and Rowland, F.S. (1986). Global atmospheric concentrations and source strength of ethane. Nature *321*, 3.

Blake, N.J., Simpson, I.J., Barletta, B., Schroeder, J., Marrero, J., Hughes, S., Hartt, G., Meinardi, S., Blake, D.R., Emmons, L.K., et al. (2014). Spatial Distributions and Source Characterization of Trace Organic Gases during SEAC4RS and comparison to DC3. In American Geophysical Union (San Francisco).

Blumenstock, T., Hase, F., Kramer, I., Mikuteit, S., Fischer, H., Goutail, F., and Raffalski, U. (2009). Winter to winter variability of chlorine activation and ozone loss as observed by ground-based FTIR measurements at Kiruna since winter 1993/94. International Journal of Remote Sensing *30*, 4055-4064.

Brandt, A.R., Heath, G.A., and Cooley, D. (2016). Methane Leaks from Natural Gas Systems Follow Extreme Distributions. Environmental science & technology *50*, 12512-12520.

Brandt, A.R., Heath, G.A., Kort, E.A., O'Sullivan, F., Pétron, G., Jordaan, S.M., Tans, P., Wilcox, J., Gopstein, A.M., Arent, D., et al. (2014). Methane leaks from North American natural gas systems. Science *343*, 733-735.

Brandt, A.R., Yeskoo, T., McNally, S., Vafi, K., Cai, H., and Wang, M.Q. (2015). Energy Intensity and Greenhouse Gas Emissions from Crude Oil Production in the Bakken Formation: Input Data and Analysis Methods. (Energy Systems Division Argonne National Laboratory).

Brantley, H.L., Thoma, E.D., and Eisele, A.P. (2015). Assessment of volatile organic compound and hazardous air pollutant emissions from oil and natural gas well pads using mobile remote and on-site direct measurements. Journal of the Air & Waste Management Association *65*, 1072-1082.

Brantley, H.L., Thoma, E.D., Squier, W.C., Guven, B.B., and Lyon, D. (2014). Assessment of methane emissions from oil and gas production pads using mobile measurements. Environmental science & technology *48*, 14508-14515.

Calvert, J.G., Derwent, R.G., Orlando, J.J., Tyndall, G.S., and Wallington, T.J. (2008). Mechanisms of Atmospheric Oxidation of the Alkanes. (Oxford University Press, USA).

Cheadle, L.C., Oltmans, S.J., Petron, G., Schnell, R.C., Mattson, E.J., Herndon, S.C., Thompson, A.M., Blake, D.R., and McClure-Begley, A. (2017). Surface ozone in the Colorado northern Front Range and the influence of oil and gas development during FRAPPE/DISCOVER-AQ in summer 2014. Elementa Science of the Anthropocene *5*, 61.

Collett, J.L., Ham, J., and Hecobian, A. (2016). North Front Range Oil and Gas Air Pollutant Emission and Dispersion Study. (Fort Collins, Colorado, USA: Colorado State University).

Conder, M.W., and Lawlor, K.A. (2014). Production characteristics of liquids-rich resource plays challenge facility design.

de Gouw, J.A., Parrish, D.D., Frost, G.J., and Trainer, M. (2014). Reduced emissions of CO2, NOx, and SO2 from U.S. power plants owing to switch from coal to natural gas with combined cycle technology. Earth's Future *2*, 75-82.

Eisele, A.P., Hannigan, M., Milford, J., Helmig, D., and Milmoe, P. (2009). Understanding Air Toxics and Carbonyl Pollutant Sources in Boulder County, Colorado. (U.S. Environmental Protection Agency, Region 8).

Etiope, G., and Ciccioli, P. (2009). Earth's Degassing: A Missing Ethane and Propane Source. Science *323*, 1.

Field, R.A., Soltis, J., McCarthy, M.C., Murphy, S., and Montague, D.C. (2015). Influence of oil and gas field operations on spatial and temporal distributions of atmospheric non-methane hydrocarbons and their effect on ozone formation in winter. Atmos. Chem. Phys. *15*, 3527-3542.

Fischer, E.V., Jacob, D.J., Yantosca, R.M., Sulprizio, M.P., Millet, D.B., Mao, J., Paulot, F., Singh, H.B., Roiger, A., Ries, L., et al. (2014). Atmospheric peroxyacetyl nitrate (PAN): a global budget and source attribution. Atmospheric Chemistry and Physics *14*, 2679-2698.

Fischer, E.V., Jaffe, D.A., and Weatherhead, E.C. (2011). Free tropospheric peroxyacetyl nitrate (PAN) and ozone at Mount Bachelor: potential causes of variability and timescale for trend detection. Atmospheric Chemistry and Physics *11*, 5641-5654.

Franco, B., Bader, W., Toon, G.C., Bray, C., Perrin, A., Fischer, E.V., Sudo, K., Boone, C.D., Bovy, B., Lejeune, B., et al. (2015). Retrieval of ethane from ground-based FTIR solar spectra using improved spectroscopy: recent burden increase above Jungfraujoch. J. Quant. Spectrosc. Radiat. Transfer *160*.

Franco, B., Mahieu, E., Emmons, L.K., Tzompa-Sosa, Z.A., Fischer, E.V., Sudo, K., Bovy, B., Conway, S., Griffin, D., Hannigan, J.W., et al. (2016). Evaluating ethane and methane emissions associated with the development of oil and natural gas extraction in North America. Environmental Research Letters *11*, 044010.

Fung, I., John, J., Lerner, J., Matthews, E., Prather, M.J., Steele, L.P., and Fraser, P.J. (1991). Three-Dimensional Model Synthesisof the Global Methane Cycle. Journal of Geophysical Research *96*, 13033-13065.

García, O.E., Schneider, M., Redondas, A., González, Y., Hase, F., Blumenstock, T., and Sepúlveda, E. (2012). Investigating the long-term evolution of subtropical ozone profiles applying ground-based FTIR spectrometry. Atmospheric Measurement Techniques *5*, 2917-2931.

Gentner, D.R., Harley, R.A., Miller, A.M., and Goldstein, A.H. (2009). Diurnal and Seasonal Variability of Gasoline-Related Volatile Organic Compound Emissions in Riverside, California. Environmental science & technology *43*, 4247-4252.

Geron, C.D., Nie, D., Arnts, R.R., Sharkey, T.D., Singsaas, E.L., Vanderveer, P.J., Guenther, A., Sickles, J.E., and Kleindienst, T.E. (1997). Biogenic isoprene emission: Model evaluation in a southeastern United States bottomland deciduous forest. Journal of Geophysical Research: Atmospheres *102*, 18889-18901.

Ghandi, A., Yeh, S., Brandt, A.R., Vafi, K., Cai, H., Wang, M.Q., Scanlon, B.R., and Reedy, R.C. (2015). Energy intensity and greenhouse gas

emissions from crude oil production in the Eagle Ford Region: Input data and analysis methods. (Institute of Transportation Studies, University of California, Davis).

Gilman, J.B., Lerner, B.M., Kuster, W.C., and de Gouw, J.A. (2013). Source signature of volatile organic compounds from oil and natural gas operations in northeastern Colorado. Environmental science & technology *47*, 1297-1305.

Goetz, J.D., Avery, A., Werde, B., Floerchinger, C., Fortne, E.C., Wormhoudt, J., Massol, P., Herndon, S.C., Kolb, C.E., Knighton, W.B., et al. (2017). Analysis of local-scale background concentrations of methane and other gas-phase species in the Marcellus Shale.

Gomer, R., and Kistiakowsky, G.B. (1951). The Rate Constant of Ethane Formation from Methyl Radicals. The Journal of Chemical Physics *19*, 85-91.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P.I., and Geron, C.D. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). Atmospheric Chemistry and Physics *6*, 3181-3210.

Guo, H. (2012). Volatile Organic Compounds (VOCs) Emitted from Petroleum and their Influence on Photochemical Smog Formation in the Atmosphere. Journal of Petroleum & Environmental Biotechnology *03*.

Gupta, M.L., Cicerone, R.J., Blake, D.R., Rowland, F.S., and Isaksen, I.S.A. (1998). Global atmospheric distributions and source strengths of light hydrocarbons and tetrachloroethene. Journal of Geophysical Research: Atmospheres *103*, 28219-28235.

Halliday, H.S., Thompson, A.M., Wisthaler, A., Blake, D.R., Hornbrook, R.S., Mikoviny, T., Muller, M., Eichler, P., Apel, E.C., and Hills, A.J. (2016). Atmospheric benzene observations from oil and gas production in the Denver-Julesburg Basin in July and August 2014. Journal of Geophysical Research: Atmospheres *121*, 11055-11074.

Hannigan, J.W., Coffey, M.T., and Goldman, A. (2009). Semiautonomous FTS Observation System for Remote Sensing of Stratospheric and Tropospheric Gases. Journal of Atmospheric and Oceanic Technology *26*, 1814-1828.

Helmig, D., Petrenko, V., Martinerie, P., Witrant, E., Röckmann, T., Zuiderweg, A., Holzinger, R., Hueber, J., Thompson, C., White, J.W.C., et al. (2014a). Reconstruction of Northern Hemisphere 1950-2010 atmospheric non-methane hydrocarbons. Atmospheric Chemistry and Physics *14*, 1463-1483.

Helmig, D., Rossabi, S., Hueber, J., Tans, P., Montzka, S.A., Masarie, K., Thoning, K., Plass-Duelmer, C., Claude, A., Carpenter, L.J., et al. (2016). Reversal of global atmospheric ethane and propane trends largely due to US oil and natural gas production. Nature Geoscience.

Helmig, D., Thompson, C.R., Evans, J., Boylan, P., Hueber, J., and Park, J.H. (2014b). Highly elevated atmospheric levels of volatile organic compounds in the Uintah Basin, Utah. Environmental science & technology *48*, 4707-4715.

Johansson, J.K.E., Mellqvist, J., Samuelsson, J., Offerle, B., Lefer, B., Rappenglück, B., Flynn, J., and Yarwood, G. (2014). Emission measurements of alkenes, alkanes, SO2, and NO2from stationary sources in Southeast Texas over a 5 year period using SOF and mobile DOAS. Journal of Geophysical Research: Atmospheres *119*, 1973-1991.

Kanakidou, M., Singh, H.B., Valentin, K.M., and Crutzen, P.J. (1991). A two-dimensional study of ethane and propane oxidation in the troposphere. Journal of Geophysical Research *96*, 15,395-315,413.

Kang, M., Kanno, C.M., Reid, M.C., Zhang, X., Mauzerall, D.L., Celia, M.A., Chen, Y., and Onstott, T.C. (2014). Direct measurements of methane emissions from abandoned oil and gas wells in Pennsylvania. Proceedings of the National Academy of Sciences of the United States of America *111*, 18173-18177.

Katzenstein, A.S., Doezema, L.A., Simpson, I.J., Blake, D.R., and Rowland, F.S. (2003). Extensive regional atmospheric hydrocarbon pollution in the southwestern United States. Proceedings of the National Academy of Sciences of the United States of America *100*, 11975-11979.

Keller, C.A., Long, M.S., Yantosca, R.M., Da Silva, A.M., Pawson, S., and Jacob, D.J. (2014). HEMCO v1.0: a versatile, ESMF-compliant component for calculating emissions in atmospheric models. Geoscientific Model Development *7*, 1409-1417.

Kirchstetter, T.W., Singer, B.C., Harley, R.A., Kendall, G.R., and Chan, W. (1996). Impact of Oxygenated Gasoline Use on California Light-Duty Vehicle Emissions. Environmental science & technology *30*, 661-670.

Kohlhepp, R., Barthlott, S., Blumenstock, T., Hase, F., Kaiser, I., Raffalski, U., and Ruhnke, R. (2011). Trends of HCl, ClONO₂, and HF column abundances from ground-based FTIR measurements in Kiruna (Sweden) in comparison with KASIMA model calculations. Atmospheric Chemistry and Physics *11*, 4669-4677.

Kort, E.A., Smith, M.L., Murray, L.T., Gvakharia, A., Brandt, A.R., Peischl, J., Ryerson, T.B., Sweeney, C., and Travis, K. (2016). Fugitive emissions from the Bakken shale illustrate role of shale production in global ethane shift. Geophysical Research Letters *43*, 4617-4623.

Koss, A.R., de Gouw, J.A., Warneke, C., Gilman, J.B., Lerner, B.M., Graus, M., Yuan, B., Edwards, P., Brown, S.S., Wild, R., et al. (2015). Photochemical aging of volatile organic compounds associated with oil and natural gas extraction in the Uintah Basin, UT, during a wintertime ozone formation event. Atmospheric Chemistry and Physics *15*, 5727-5741.

Kuhns, H., Green, M., and Etyemezian, V. (2003). Big Bend Regional Aerosol and Visibility Observational (BRAVO) Study Emissions Inventory. (Las Vegas, Nevada, USA: Desert Research Institute).

Lee, B.H., Munger, J.W., Wofsy, S.C., and Goldstein, A.H. (2006). Anthropogenic emissions of nonmethane hydrocarbons in the northeastern United States: Measured seasonal variations from 1992–1996 and 1999–2001. Journal of Geophysical Research *111*.

Lurmann, F.W., Lloyd, A.C., and Atkinson, R. (1986). A Chemical Mechanism for Use in Long-Range Transport/Acid Deposition Computer Modeling. Journal of Geophysical Research *91*, 10905-10936.

McDuffie, E.E., Edwards, P.M., Gilman, J.B., Lerner, B.M., Dubé, W.P., Trainer, M., Wolfe, D.E., Angevine, W.M., deGouw, J., Williams, E.J., et al. (2016). Influence of oil and gas emissions on summertime ozone in the Colorado Northern Front Range. Journal of Geophysical Research: Atmospheres *121*, 8712-8729.

McKain, K., Down, A., Raciti, S.M., Budney, J., Hutyra, L.R., Floerchinger, C., Herndon, S.C., Nehrkorn, T., Zahniser, M.S., Jackson, R.B., et al. (2015). Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. Proceedings of the National Academy of Sciences of the United States of America *112*, 1941-1946.

Mitchell, A.L., Tkacik, D.S., Roscioli, J.R., Herndon, S.C., Yacovitch, T.I., Martinez, D.M., Vaughn, T.L., Williams, L.L., Sullivan, M.R., Floerchinger, C., et al. (2015). Measurements of methane emissions from natural gas gathering facilities and processing plants: measurement results. Environmental science & technology *49*, 3219-3227.

Naik, V., Voulgarakis, A., Fiore, A.M., Horowitz, L.W., Lamarque, J.F., Lin, M., Prather, M.J., Young, P.J., Bergmann, D., Cameron-Smith, P.J., et al. (2013). Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). Atmospheric Chemistry and Physics *13*, 5277-5298.

Olivier, J.G.J. (2002). On the Quality of Global Emission Inventories. Approaches, Methodologies, Input Data, and Uncertainties. (Netherlands: Utrecht University).

Pacsi, A.P., Kimura, Y., McGaughey, G., McDonald-Buller, E.C., and Allen, D.T. (2015). Regional ozone impacts of increased natural gas use in the Texas power sector and development in the Eagle Ford shale. Environmental science & technology *49*, 3966-3973.

Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P.I., Messerschmidt, J., Deutscher, N., Griffith, D.W.T., Notholt, J., et al. (2011). Methane observations from the Greenhouse Gases Observing SATellite: Comparison to ground-based TCCON data and model calculations. Geophysical Research Letters *38*, n/a-n/a.

Parrella, J.P., Jacob, D.J., Liang, Q., Zhang, Y., Mickley, L.J., Miller, B., Evans, M.J., Yang, X., Pyle, J.A., Theys, N., et al. (2012). Tropospheric bromine chemistry: implications for present and pre-industrial ozone and mercury. Atmospheric Chemistry and Physics *12*, 6723-6740.

Peischl, J., Aikin, K.C., Eilerman, S., Gilman, J.B., de Gouw, J.A., Herndon, S.C., Lerner, B.M., Neuman, A., Tokarek, T., Trainer, M., et al. (2015a). Quantification of methane emissions from oil and natural gas extraction regions in the Central/Western U.S. and comparison to previous studies. In American Geophysical Union Fall Meeting (San Francisco, CA, USA).

Peischl, J., Karion, A., Sweeney, C., Kort, E.A., Smith, M.L., Brandt, A.R., Yeskoo, T., Aikin, K.C., Conley, S.A., Gvakharia, A., et al. (2016). Quantifying atmospheric methane emissions from oil and natural gas production in the Bakken shale region of North Dakota. Journal of Geophysical Research: Atmospheres *121*, 6101-6111.

Peischl, J., Ryerson, T.B., Aikin, K.C., de Gouw, J.A., Gilman, J.B., Holloway, J.S., Lerner, B.M., Nadkarni, R., Neuman, J.A., Nowak, J.B., et al. (2015b). Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions. Journal of Geophysical Research: Atmospheres.

Pekney, N.J., Veloski, G., Reeder, M., Tamilia, J., Rupp, E., and Wetzel, A. (2014). Measurement of atmospheric pollutants associated with oil and natural gas exploration and production activity in Pennsylvania's Allegheny National Forest. Journal of the Air & Waste Management Association *64*, 1062-1072.

Pétron, G., Frost, G.J., Miller, B.R., Hirsch, A.I., Montzka, S.A., Karion, A., Trainer, M., Sweeney, C., Andrews, A.E., Miller, L., et al. (2012). Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study. Journal of Geophysical Research *117*.

Pétron, G., Karion, A., Sweeney, C., Miller, B.R., Montzka, S.A., Frost, G.J., Trainer, M., Tans, P., Andrews, A.E., Kofler, J., et al. (2014). A new look at methane and nonmethane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin. Journal of Geophysical Research: Atmospheres.

Pfister, G., Flocke, F., Hornbrook, R.S., Orlando, J.J., Lee, S., and Schroeder, J. (2017). Process-Based and Regional Source Impact Analysis for FRAPPÉ and DISCOVER-AQ 2014. (Boulder, CO, USA.: National Center for Atmospheric Research and Atmospheric Chemistry Observations and Modeling Laboratory).

Phillips-Smith, C., Jeong, C.-H., Healy, R.M., Dabek-Zlotorzynska, E., Celo, V., Brook, J.R., and Evans, G. (2017). Sources of Particulate Matter in the Athabasca Oil Sands Region: Investigation through a Comparison of Trace Element Measurement Methodologies. Atmospheric Chemistry and Physics Discussions, 1-34.

Plass-Dülmer, C., Koppmann, R., Ratte, M., and Rudolph, J. (1995). Light nonmethane hydrocargons in seawater. Global Biogeochemical Cycles *9*, 79-100.

Pozzer, A., Pollmann, J., Taraborrelli, D., Jöckel, P., Helmig, D., Tans, P., Hueber, J., and Lelieveld, J. (2010). Observed and simulated global distribution and budget of atmospheric C2-C5 alkanes. Atmospheric Chemistry and Physics *10*, 4403-4422.

Prather, M.J., Holmes, C.D., and Hsu, J. (2012). Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry. Geophysical Research Letters *39*, n/a-n/a.

Prinn, R.G., Huang, J., Weiss, R.F., Cunnold, D.M., Fraser, P.J., Simmonds, P.G., McCulloch, A., Harth, C., Reimann, S., Salameh, P., et al. (2005). Evidence for variability of atmospheric hydroxyl radicals over the past quarter century. Geophysical Research Letters *32*, n/a-n/a.

Pusede, S.E., and Cohen, R.C. (2012). On the observed response of ozone to NOx and VOC reactivity reductions in San Joaquin Valley California 1995–present. Atmospheric Chemistry and Physics *12*, 8323-8339.

Randerson, J.T., Chen, Y., van der Werf, G.R., Rogers, B.M., and Morton, D.C. (2012). Global burned area and biomass burning emissions from small fires. Journal of Geophysical Research: Biogeosciences *117*, n/a-n/a.

Rappenglück, B., Ackermann, L., Alvarez, S., Golovko, J., Buhr, M., Field, R.A., Soltis, J., Montague, D.C., Hauze, B., Adamson, S., et al. (2014). Strong wintertime ozone events in the Upper Green River basin, Wyoming. Atmospheric Chemistry and Physics *14*, 4909-4934.

Richter, D., Weibring, P., Walega, J.G., Fried, A., Spuler, S.M., and Taubman, M.S. (2015). Compact highly sensitive multi-species airborne mid-IR spectrometer. Applied Physics B *119*, 119-131.

Roest, G., and Schade, G. (2017). Quantifying alkane emissions in the Eagle Ford Shale using boundary layer enhancement. Atmospheric Chemistry and Physics *17*, 11163-11176.

Roscioli, J.R., Yacovitch, T.I., Floerchinger, C., Mitchell, A.L., Tkacik, D.S., Subramanian, R., Martinez, D.M., Vaughn, T.L., Williams, L., Zimmerle, D., et al. (2015). Measurements of methane emissions from natural gas gathering facilities and processing plants: measurement methods. Atmospheric Measurement Techniques *8*, 2017-2035.

Rossabi, S., and Helmig, D. (2018). Changes in Atmospheric Butanes and Pentanes and their Isomeric Ratios in the Continental United States. Journal of Geophysical Research: Atmospheres.

Rudolph, J. (1995). The tropospheric distribution and budget of ethane. Journal of Geophysical Research *100*, 11.

Rudolph, J., and Ehhalt, D.H. (1981). Measurements of C2–C5hydrocarbons over the North Atlantic. Journal of Geophysical Research *86*, 11959.

Russell, A., Milford, J., Bergin, M.S., McBride, S., McNair, L., Yang, Y., Stockwell, W.R., and Croes, B. (1995). Urban ozone control and atmospheric reactivity of organic gases. Science *269*, 5.

Sander, S.P., Friedl, R.R., Abbatt, J., Barker, J.R., Burkholder, J.B., Golden, D.M., Kolb, C.E., Kurylo, M.J., Moortgat, G.K., H., W.P., et al. (2011). Chemical Kinetics and Photochemical

Data for Use in Atmospheric Studies, Evaluation No. 17. J.P. Laboratory, ed. (Pasadena, CA: JPL Publication).

Sander, S.P., Friedl, R.R., Golden, D.M., Kurylo, M.J., Huie, R.E., Orkin, V.L., Moortgat, G.K., Ravishankara, A.R., Kolb, C.E., Molina, M.J., et al. (2003). Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies, Evaluation No. 14. J.P. Laboratory, ed. (Pasadena, CA, USA: JPL Publication).

Sangwan, M., Yan, C., Chesnokov, E.N., and Krasnoperov, L.N. (2015). Reaction CH3 + CH3 --> C2H6 Studied over the 292-714 K Temperature and 1-100 bar Pressure Ranges. The Journal of Physical Chemistry *119*, 7847-7857.

Saunois, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J.G., Dlugokencky, E.J., Etiope, G., Bastviken, D., Houweling, S., et al. (2016). The Global Methane Budget: 2000-2012. Earth System Science Data Discussions, 1-79.

Schauer, J.J., Kleeman, M.J., Cass, G.R., and Simoneit, B.R. (2002). Measurement of emissions from air pollution sources. 5. C-1−C-32 organic compounds from gasoline-powered motor vehicles. Environ. Sci. Technol. *36*, 12.

Schauffler, S.A., E. L.; Navarro, M. A.; Pan, L.; Blake, Donald R.; Blake, N. J.; Kinnison, D. E.; Meinardi, S.; Lueb, R.; Zhu, X.; Pope, L. (2014). Organic Halogen and Hydrocarbon Distributions During SEAC4RS Measured from the ER-2 and DC-8. In American Geophysical Union (San Francisco, CA).

Schneider, M., Romero, P.M., Hase, F., Blumenstock, T., Cuevas, E., and Ramos, R. (2010). Continuous quality assessment of atmospheric water vapour measurement techniques: FTIR, Cimel, MFRSR, GPS, and Vaisala RS92. Atmospheric Measurement Techniques *3*, 323–338.

Schwietzke, S., Griffin, W.M., Matthews, H.S., and Bruhwiler, L.M. (2014). Natural gas fugitive emissions rates constrained by global atmospheric methane and ethane. Environmental science & technology *48*, 7714-7722.

Sherwen, T., Schmidt, J.A., Evans, M.J., Carpenter, L.J., Großmann, K., Eastham, S.D., Jacob, D.J., Dix, B., Koenig, T.K., Sinreich, R., et al. (2016). Global impacts of tropospheric halogens (Cl, Br, I) on oxidants and composition in GEOS-Chem. Atmospheric Chemistry and Physics *16*, 12239-12271.

Simon, H., Beck, L., Bhave, P.V., Divita, F., Hsu, Y., Luecken, D., Mobley, J.D., Pouliot, G.A., Reff, A., Sarwar, G., et al. (2010). The development and uses of EPA's SPECIATE database. Atmospheric Pollution Research *1*, 196-206.

Simpson, I.J., Akagi, S.K., Barletta, B., Blake, N.J., Choi, Y., Diskin, G.S., Fried, A., Fuelberg, H.E., Meinardi, S., Rowland, F.S., et al. (2011). Boreal forest fire emissions in fresh Canadian smoke plumes: C1-C10 volatile organic compounds (VOCs), CO2, CO, NO2, NO, HCN and CH3CN. Atmospheric Chemistry and Physics *11*, 6445-6463.

Simpson, I.J., Blake, N.J., Barletta, B., Diskin, G.S., Fuelberg, H.E., Gorham, K., Huey, L.G., Meinardi, S., Rowland, F.S., Vay, S.A., et al. (2010). Characterization of trace gases measured over Alberta oil sands mining operations: 76 speciated C2–C10 volatile organic compounds (VOCs), CO2, CH4, CO, NO, NO2, NOy, O3 and SO2. Atmospheric Chemistry and Physics *10*, 11931-11954.

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

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.F., Kuhn, U., Stefani, P., and Knorr, W. (2014). Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. Atmospheric Chemistry and Physics *14*, 9317-9341.

Singh, H.B., and Zimmerman, P.B. (1992). Atmospheric distribution and sources of nonmethane hydrocarbons. In Gaseous Pollutants: Characterization and cycling. J. O.Nriagu, ed. (New York: Wiley), pp. 177-235.

Speight, J.G. (2013). Shale Gas Properties and Processing. (Boston: Gulf Professional Publishing).

Swarthout, R.F., Russo, R.S., Zhou, Y., Hart, A.H., and Sive, B.C. (2013). Volatile organic compound distributions during the NACHTT campaign at the Boulder Atmospheric Observatory: Influence of urban and natural gas sources. Journal of Geophysical Research *118*, 24.

Swarthout, R.F., Russo, R.S., Zhou, Y., Miller, B.M., Mitchell, B., Horsman, E., Lipsky, E., McCabe, D.C., Baum, E., and Sive, B.C. (2015). Impact of Marcellus Shale natural gas development in southwest Pennsylvania on volatile organic compound emissions and regional air quality. Environmental science & technology *49*, 3175-3184.

TCEQ (2012). Data collected by Automated Gas Chromatographs (Auto-GCs). In daily (Texas, USA: Texas Commission on Environmental Quality).

Thompson, C.R., Hueber, J., and Helmig, D. (2014). Influence of oil and gas emissions on ambient atmospheric non-methane hydrocarbons in residential areas of Northeastern Colorado. Elementa: Science of the Anthropocene *2*, 000035.

Thompson, T.M., Shepherd, D., Stacy, A., Barna, M.G., and Schichtel, B.A. (2017). Modeling to Evaluate Contribution of Oil and Gas Emissions to Air Pollution. Journal of the Air & Waste Management Association *67*, 445-461.

Thynne, J.C.J. (1962). Reactions of alkyl radicals. Part 1.—Methyl radical photosensitized decomposition of ethyl formate. Trans. Faraday Soc. *58*, 676-684.

Travis, K.R., Jacob, D.J., Fisher, J.A., Kim, P.S., Marais, E.A., Zhu, L., Yu, K., Miller, C.C., Yantosca, R.M., Sulprizio, M.P., et al. (2016). Why do models overestimate surface ozone in the Southeast United States? Atmospheric Chemistry and Physics *16*, 13561-13577.

Turner, A.J., and Jacob, D.J. (2015). Balancing aggregation and smoothing errors in inverse models. Atmospheric Chemistry and Physics *15*, 7039-7048.

Turner, A.J., Jacob, D.J., Wecht, K.J., Maasakkers, J.D., Lundgren, E., Andrews, A.E., Biraud, S.C., Boesch, H., Bowman, K.W., Deutscher, N.M., et al. (2015). Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data. Atmospheric Chemistry and Physics *15*, 7049-7069.

Tzompa-Sosa, Z.A., Mahieu, E., Franco, B., Keller, C.A., Turner, A.J., Helmig, D., Fried, A., Richter, D., Weibring, P., Walega, J.G., et al. (2017). Revisiting global fossil fuel and biofuel emissions of ethane. Journal of Geophysical Research: Atmospheres *122*.

U.S. EIA (2017). U.S. Natural gas gross withdrawals and production. Energy Information Administration, ed.

U.S. EPA (2009). Municipal Solid Waste Landfills. A.E.F.a. Quantification, ed.

U.S. EPA (2010). 8-hour Ozone (2008) Nonattainment Area Summary.

U.S. EPA (2013). National Emissions Inventory 2011, version 1. Environmental Protection Agency, ed.

U.S. EPA (2015). Drilling Productivity Report for key tight oil and shale gas regions. E.P. Agency, ed.

U.S. EPA (2016a). Control Techniques Guidelines for the Oil and Natural Gas Industry. O.o.A.a.R.a.O.o.A.Q.P.a. Standards, ed. (Research Triangle Park, North Carolina, USA: U.S. Environmental Protection Agency,).

U.S. EPA (2016b). Technical Support Document: Preparation of Emissions Inventories for the Version 6.3, 2011 Emissions Modeling Platform. Environmental Protection Agency, ed.

U.S. EPA (2017). National Emissions Inventory 2011, version 6.3ek. Environmental Protection Agency, ed.

U.S. EPA (2018). Ozone trends.

van der Werf, G.R., Randerson, J.T., Giglio, L., Collatz, G.J., Mu, M., Kasibhatla, P.S., Morton, D.C., DeFries, R.S., Jin, Y., and van Leeuwen, T.T. (2010). Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). Atmospheric Chemistry and Physics *10*, 11707-11735.

Vinciguerra, T., Yao, S., Dadzie, J., Chittams, A., Deskins, T., Ehrman, S., and Dickerson, R.R. (2015). Regional air quality impacts of hydraulic fracturing and shale natural gas activity: Evidence from ambient VOC observations. Atmospheric Environment *110*, 144-150.

Voulgarakis, A., Naik, V., Lamarque, J.F., Shindell, D.T., Young, P.J., Prather, M.J., Wild, O., Field, R.D., Bergmann, D., Cameron-Smith, P., et al. (2013). Analysis of present day and future OH and methane lifetime in the ACCMIP simulations. Atmospheric Chemistry and Physics *13*, 2563-2587.

Wang, J.S., Logan, J.A., McElroy, M.B., Duncan, B.N., Megretskaia, I.A., and Yantosca, R.M. (2004). A 3-D model analysis of the slowdown and interannual variability in the methane growth rate from 1988 to 1997. Global Biogeochemical Cycles *18*, n/a-n/a.

Warneke, C., Geiger, F., Edwards, P.M., Dube, W., Pétron, G., Kofler, J., Zahn, A., Brown, S.S., Graus, M., Gilman, J.B., et al. (2014). Volatile organic compound emissions from the oil and natural gas industry in the Uintah Basin, Utah: oil and gas well pad emissions compared to ambient air composition. Atmos. Chem. Phys. *14*.

Wiacek, A., Taylor, J.R., Strong, K., Saari, R., Kerzenmacher, T.E., Jones, N.B., and Griffith, D.W.T. (2007). Ground-Based Solar Absorption FTIR Spectroscopy: Characterization of Retrievals and First Results from a Novel Optical Design Instrument at a New NDACC Complementary Station. Journal of Atmospheric and Oceanic Technology *24*, 432-448.

Wofsy, S.C., Daube, B.C., Jimenez, R., Kort, E., Pittman, J.V., Park, S., Commane, R., Xiang, B., Santoni, G., Jacob, D., et al. (2012). HIPPO Combined Discrete Flask and GC Sample GHG, Halo-, Hydrocarbon Data (R_20121129). O.R.N.L. Carbon Dioxide Information Analysis Center, Oak Ridge, Tennessee, U.S.A. , ed.

Xiao, Y.P., Jacob, D.J., Wang, J.S., Logan, J.A., Palmer, P.I., Suntharalingam, P., Yantosca, R.M., Sachse, G.W., Blake, D.R., and Streets, D.G. (2004). Constraints on Asian and European sources of methane from CH4-C2H6-CO correlations in Asian outflow. Journal of Geophysical Research *109*.

Xiao, Y.P., Logan, J.A., Jacob, D.J., Hudman, R.C., Yantosca, R.M., and Blake, D.R. (2008). Global budget of ethane and regional constraints on US sources. Journal of Geophysical Research: Atmospheres *113*.

Yacovitch, T.I., and Herndon, S.C. (2014). TILDAS Ethane Quality Assurance Document, Discover AQ Denver 2014.

Yevich, R., and Logan, J.A. (2003). An assessment of biofuel use and burning of agricultural waste in the developing world. Global Biogeochemical Cycles *17*.

Ying, Q., Li, J., and Kota, S.H. (2015). Significant Contributions of Isoprene to Summertime Secondary Organic Aerosol in Eastern United States. Environmental science & technology *49*, 7834-7842.

Zielinska, B., Campbell, D., and Samburova, V. (2014). Impact of emissions from natural gas production facilities on ambient air quality in the Barnett Shale area: A pilot study. Journal of the Air & Waste Management Association *64*, 1369-1383.

Zimmerman, P.R., Greenberg, J.P., and Westberg, C.E. (1988). Measurements of Atmospheric Hydrocarbons and Biogenic Emission Fluxes in the Amazon Boundary Layer. Journal of Geophysical Research *93*, 10.

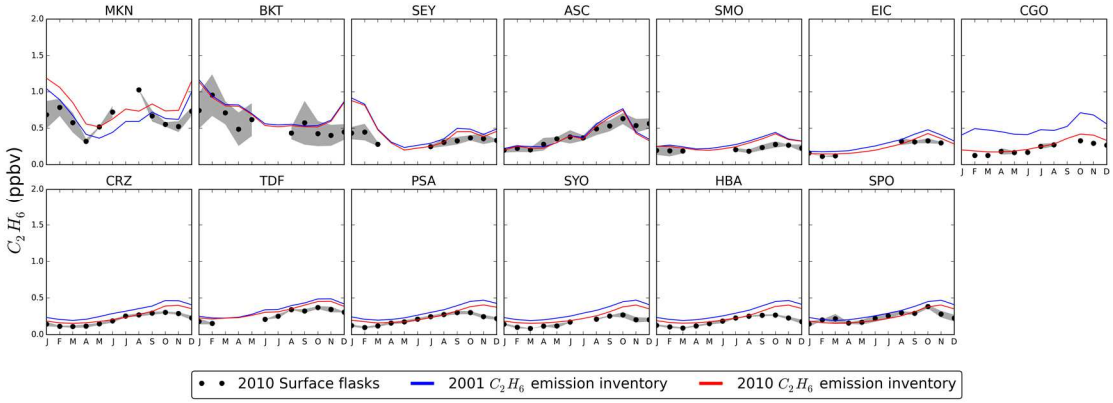


Figure A1: Comparison of Southern Hemisphere 2010 C_2H_6 surface mixing ratios to modeled 2001 and 2010 C_2H_6 emission inventories. Black dots represent C_2H_6 observations from NOAA GGGRN global surface flask network and grey areas denote their associated 1σ standard deviation. Lines represent model mixing ratios at the surface from both C_2H_6 emission inventories. Stations are ordered from higher to lower latitudes.

APPENDIX B

NCL script to regrid the 2011NEIv6.3ek point emission source data from Lambert conformal conic projection at a 12 km x 12 km grid to a rectilinear 0.1x0.1 degree grid that can be read by GEOS-Chem

```
load "$NCARG_ROOT/lib/ncarg/nclscripts/csm/contributed.ncl"
load "$NCARG_ROOT/lib/ncarg/nclscripts/esmf/ESMF_regridding.ncl"

begin

print("***** START REGRIDDING SOURCES *********")

split_number_3 = 3
;-----------------------------------------
; All non-point sources are: ag cmv nonpt nonroad np_oilgas onroad rail rwc
sourcelist = (/"ag"/)
i_sourcelist = dimsizes(sourcelist)-1

;----
; ag
;----
; The agricultural sector only has NH3 and NH3_FERT emissions
VARLIST_ag = (/"NH3","NH3_FERT"/)

;---------------------------------------------------------------
; Sources that have all species: cmv, nonpt, nonroad, rail, rwc
;---------------------------------------------------------------
VARLIST_all =
(/"CO","NO","NO2","HONO","FORM","PAR","IOLE","OLE","ETH","ETHA","ETOH","MEO
H",
"BENZ","TOL","XYL","SO2","SULF","PEC","POC","PSO4","NH3","ACROLEIN","ALD2",
"ALDX","ISOP","PRPA","ACET","KET","NH3_FERT","ALD2_PRIMARY","FORM_PRIMA
RY"/)

;-----------
; np_oilgas
;-----------
; The np_oilgas has a different split_number and does not have NH3, NH3_FERT
split_number_5 = 5
VARLIST_np_oilgas =
(/"CO","NO","NO2","HONO","FORM","PAR","IOLE","OLE","ETH","ETHA","ETOH","MEO
```

```
H", "BENZ","TOL","XYL","SO2","SULF","PEC","POC","PSO4","ACROLEIN","ALD2",
"ALDX","ISOP","PRPA","ACET","KET","ALD2_PRIMARY","FORM_PRIMARY"/)

;--------
; onroad
;--------
; The onroad source does not have SULF,NH3_FERT
VARLIST_onroad =
(/"CO","NO","NO2","HONO","FORM","PAR","IOLE","OLE","ETH","ETHA","ETOH","MEO
H", "BENZ","TOL","XYL","SO2","PEC","POC","PSO4","NH3","ACROLEIN","ALD2",
"ALDX","ISOP","PRPA","ACET","KET","ALD2_PRIMARY","FORM_PRIMARY"/)

; Counter for the number of files regridded
counter = 0

do ii=0,i_sourcelist
       source = sourcelist(ii)
       print("_____")
       print("  Starting to regrid with ncl code for "+source+" files")
       print("_____")
          ;*********************************************
          ;**  SELECT VARLIST OPTION AND SPLIT_NUMBER  **
          ;*********************************************
      if (source .eq. "ag") then
        split_number  = split_number_3
        VARLIST       = VARLIST_ag
      ; ":=" indicates that the variable will be overwritten
      else if (source .eq. "cmv") then
        split_number := split_number_3
        VARLIST      := VARLIST_all
      else if (source .eq. "nonpt") then
        split_number := split_number_3
        VARLIST      := VARLIST_all
      else if (source .eq. "nonroad") then
        split_number := split_number_3
        VARLIST      := VARLIST_all
      else if (source .eq. "np_oilgas") then
        split_number := split_number_5
        VARLIST      := VARLIST_np_oilgas
      else if (source .eq. "onroad") then
        split_number := split_number_3
        VARLIST      := VARLIST_onroad
      else if (source .eq. "rail") then
        split_number := split_number_3
        VARLIST      := VARLIST_all
      else if (source .eq. "rwc") then
```

```
 split_number := split_number_3
 VARLIST      := VARLIST_all
end if ; ag
end if ; cmv
end if ; nonpt
end if ; nonroad
end if ; np_oilgas
end if ; onroad
   end if ; rail
   end if ; rwc

   ;*************************************************************
   ; Path where scripts/regridding files are/will be located
   scripts_dir = "/home/ztzompa/NEI2011ek_scripts/area_sources/"
   ; weight file with the remapping weights
wgtFile= scripts_dir+"nei11_to_0.1x0.1_wgts.nc"
; input directory
   ;  indir    = "./"    ; input directory
   indir = "/fischer-scratch/ztzompa/NEI2011ek/"+source+"/input/"
   ; Open nc file; Read dimensions of grid. These are file variables.
   myFiles = systemfunc("ls " + indir + "*2011*.ncf")
;print(myFiles)
   ;*************************************************************
   ; Loop over files
   ;*************************************************************
   do f      = 0, dimsizes(myFiles)-1
     infile  =  myFiles(f)
     fd      = addfile(infile, "r")
  ;print(infile)
  print("---------------------------------------------------------")
  print("  Now regridding "+infile)
  print("---------------------------------------------------------")

     nlat    = fd@NROWS
     mlon    = fd@NCOLS
     nlev    = fd@NLAYS
     gdnam   = str_right_strip(fd@GDNAM)     ; removes ending spaces and TABs

     tt   = "VAR-LIST"
     VARS = fd@$tt$
     ;VARLIST = str_split(VARS," ")

   ;*************************************************************
   ; Create date string for this file
   ;*************************************************************
     tmpstr  = str_split(infile, ".")
```

```
  datestr = str_split(tmpstr(0), "_")
  tttime=str_split_by_length(datestr(split_number),(/4,2,2/))

  ;print (tmpstr)
  ;print (datestr)
  print (tttime)

  time          = fspan(0.,23.,24)
  time!0        = "time"
  time&time     = time
  time@long_name = "Time"
  time@units    = "hours since "+tttime(0)+"-"+tttime(1)+"-"+tttime(2)+" 00:00:00"
  time@calendar  = "standard"

;*****************************************************************
; Check / create weight files for remapping
;*****************************************************************
  exists=isfilepresent(wgtFile)
  if( .not. exists ) then
    ; source grid
    srcGridName = scripts_dir+"SCRIP_NEI2011_curvilinear.nc"
    fil_csv = "gridFile.csv"
    lonlat = asciiread(scripts_dir+fil_csv,(/mlon,nlat,2/),"float")
    lat2d = (/ transpose(lonlat(:,:,1)) /) ; Thanks Lee!
    lon2d = (/ transpose(lonlat(:,:,0)) /)
    lon2d = where(lon2d.gt.180,lon2d-360,lon2d)

    Opt               = True
    Opt@ForceOverwrite = True
    Opt@PrintTimings   = True
    Opt@Title         = "NEI 2011 curvilinear grid"
    curvilinear_to_SCRIP(srcGridName,lat2d,lon2d,Opt)
    delete(Opt)

    ; destination grid
    dstGridName = scripts_dir+"SCRIP_0.1x0.1_rectilinear.nc"

    ; 0.25 x 0.25
    ;lon1 = -139.875
    ;lon2 = -50.125
    ;nlon = 360
    ;lat1 =  20.125
    ;lat2 =  59.875
    ;nlat = 160
    ;Regional = True
    ;Title = "0.25x0.25 rectilinear"
```

```
  ; 0.1 x 0.1
  lon1 = -139.95
  lon2 = -50.05
  nlon = 900
  lat1 =  20.05
  lat2 =  59.95
  nlat = 400
  Regional = True
  Title = "0.1x0.1 rectilinear"

  lat1d = fspan( lat1, lat2, nlat )
  lon1d = fspan( lon1, lon2, nlon )

  Opt              = True
  Opt@ForceOverwrite = True
  Opt@PrintTimings   = True
  Opt@Title          = Title
  rectilinear_to_SCRIP(dstGridName,lat1d,lon1d,Opt)
  delete(Opt)

 ; create weights
  Opt                = True
  Opt@InterpMethod       = "conserve"
  Opt@SrcRegional        = True
  Opt@DstRegional        = Regional
  Opt@PrintTimings       = True

  ESMF_regrid_gen_weights(srcGridName, dstGridName, wgtFile, Opt)
  delete(Opt)

 end if

; ************************************************************
; write netCDF
;************************************************************
  outdir  = "/fischer-scratch/ztzompa/NEI2011ek/"+source+"/output/"
  newname = "NEI2011ek_0.1x0.1_"+datestr(split_number)+"_"+source+".nc"
  ncPath  = outdir+newname
  if (isfilepresent(ncPath)) then
     continue
  end if
  ;print ("stopped line 227")
  setfileoption("nc","Format","NetCDF4Classic")
  setfileoption("nc","CompressionLevel",5)
  ncdf    = addfile(ncPath ,"c")     ; open output netCDF file
```

```
; Create global attributes
  ncdf@title      = "NEI2011 version 6.3 ek from EPA"
  ncdf@source_file  = infile       ; copy some file attributes
  ncdf@history   = "Created by Zitely Tzompa-Sosa using Katie Travis & Christoph
Keller ncl script; " + systemfunc("date")
  ncdf@date     = datestr(split_number)

; Define some general variable properties
  VarTname = "time"
  VarYname = "lat"
  VarXname = "lon"
  VarZname = "lev"

; Options for regridding
  Opt             = True
  Opt@PrintTimings = True
  ncdf->time=time
; ***************************************************************
; loop over all species in input data
; ***************************************************************
  VARLIST = VARLIST
 do i=0,dimsizes(VARLIST)-1

  ; Read in species and set to Geos-Chem tracers
    varIn      = fd->$VARLIST(i)$

  ; Remove 25th value and extraneous layer dimension
    varTmp   =  varIn(0:23,0,:,:)

  ; Convert from unit/s to unit/m2/s (area = 12 km x 12 km)
    varTmp  = varTmp /(12.0E3*12.0E3)  ; moles/m2/s or g/m2/s

  ; remap data
    varRd = ESMF_regrid_with_weights(varTmp,wgtFile,Opt)
    varRd@_FillValue = 0.0
    delete(varRd@_FillValue)
  ; Create variable
    varRd!0 = VarTname ;change TSTEP to time
    varRd!1 = VarYname
    varRd!2 = VarXname
    varRd@long_name=varIn@long_name

  ; fix missing values
    varRd@_FillValue = varRd@missing_value
    delete(varRd@missing_value)
```

```
; convert from moles to g and attach GEOS-Chem tracername
  if (.not. ismissing(str_match(varRd@long_name,"CO"))) then
    varRd = varRd *28.0/1.0E3  ; convert from moles to kg
    varRd@NEI11_name = "CO"
    varRd@units = "kg/m2/s"
    varRd@molecular_weight_g_mol= "28"
  else if (.not. ismissing(str_match(varRd@long_name,"HONO"))) then
    varRd = varRd *47.0/1.0E3  ; convert from moles to kg
    varRd@NEI11_name = "HNO2"
    varRd@units= "kg/m2/s"
    varRd@molecular_weight_g_mol= "47"
  else if (.not. ismissing(str_match(varRd@long_name,"NO2"))) then
    varRd = varRd *46.0/1.0E3  ; convert from moles to kg
    varRd@NEI11_name = "NO2"
    varRd@units= "kg/m2/s"
    varRd@molecular_weight_g_mol= "46"
  else if (.not. ismissing(str_match(varRd@long_name,"NO"))) then
    varRd = varRd *30.0/1.0E3  ; convert from moles to kg
    varRd@NEI11_name = "NO"
    varRd@units= "kg/m2/s"
    varRd@molecular_weight_g_mol= "30"
  else if (.not. ismissing(str_match(varRd@long_name,"ISOP"))) then
    varRd = varRd *12.0*5/1.0E3  ; convert from moles to kg
    varRd@NEI11_name = "ISOP"
    varRd@units= "kgC/m2/s"
    varRd@molecular_weight_g_mol= "12*5"
  else if (.not. ismissing(str_match(varRd@long_name,"SO2"))) then
    varRd = varRd *64.0/1.0E3  ; convert from moles to kg
    varRd@NEI11_name = "SO2"
    varRd@units= "kg/m2/s"
    varRd@molecular_weight_g_mol= "64"
  else if (.not. ismissing(str_match(varRd@long_name,"SULF"))) then
    varRd = varRd *98.0/1.0E3  ; convert from moles to kg
    varRd@NEI11_name = "SO2"
    varRd@units= "kg/m2/s"
    varRd@molecular_weight_g_mol= "98"
  else if (.not. ismissing(str_match(varRd@long_name,"NH3"))) then
    varRd = varRd *17.0/1.0E3  ; convert from moles to kg
    varRd@NEI11_name = "NH3"
    varRd@units= "kg/m2/s"
    varRd@molecular_weight_g_mol= "17"
  else if (.not. ismissing(str_match(varRd@long_name,"HNO4"))) then
    varRd = varRd *79.0/1.0E3  ; convert from moles to kg
    varRd@NEI11_name = "HNO4"
    varRd@units= "kg/m2/s"
```

```
  varRd@molecular_weight_g_mol= "78"
else if (.not. ismissing(str_match(varRd@long_name,"ACROLEIN"))) then
  varRd = varRd *56.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "MACR"
  varRd@units= "kg/m2/s"
  varRd@molecular_weight_g_mol= "56"
else if (.not. ismissing(str_match(varRd@long_name,"ALD2"))) then
  varRd = varRd *12.0*2.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "ALD2"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*2"
else if (.not. ismissing(str_match(varRd@long_name,"ALDX"))) then
  varRd = varRd *58.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "RCHO"
  varRd@units= "kg/m2/s"
  varRd@molecular_weight_g_mol= "58"
else if (.not. ismissing(str_match(varRd@long_name,"BENZ"))) then
  varRd = varRd *12.0*6.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "BENZ"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "72"
else if (.not. ismissing(str_match(varRd@long_name,"CH4"))) then
  varRd = varRd *16.04/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "CH4"
  varRd@units= "kg/m2/s"
else if (.not. ismissing(str_match(varRd@long_name,"ETHA"))) then
  varRd = varRd *12.0*2.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "C2H6"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*2"
else if (.not. ismissing(str_match(varRd@long_name,"ETH"))) then
  varRd = varRd *12.0*2.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "C2H4"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*2"
else if (.not. ismissing(str_match(varRd@long_name,"ETOH"))) then
  varRd = varRd *46.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "EOH"
  varRd@units= "kg/m2/s"
  varRd@molecular_weight_g_mol= "46"
else if (.not. ismissing(str_match(varRd@long_name,"FORM"))) then
  varRd = varRd *30.03/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "CH2O"
  varRd@units= "kg/m2/s"
  varRd@molecular_weight_g_mol= "30"
else if (.not. ismissing(str_match(varRd@long_name,"IOLE"))) then
```

```
  varRd = varRd *12.0*4/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "PRPE"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*4"
else if (.not. ismissing(str_match(varRd@long_name,"OLE"))) then
  varRd = varRd *12.0*2/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "PRPE"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*2"
else if (.not. ismissing(str_match(varRd@long_name,"PAR"))) then
  varRd = varRd *12.0*4/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "ALK4"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*4"
else if (.not. ismissing(str_match(varRd@long_name,"MEOH"))) then
  varRd = varRd *32.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "MOH"
  varRd@units= "kg/m2/s"
  varRd@molecular_weight_g_mol= "32"
else if (.not. ismissing(str_match(varRd@long_name,"TOL"))) then
  varRd = varRd *12.0*7.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "TOLU"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*7"
else if (.not. ismissing(str_match(varRd@long_name,"XYL"))) then
  varRd = varRd *12.0*8.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "XYLE"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*8"
else if (.not. ismissing(str_match(varRd@long_name,"POC"))) then
  varRd = varRd/1.0E3  ; convert from g to kg
  varRd@NEI11_name = "OC"
  varRd@units= "kg/m2/s"
else if (.not. ismissing(str_match(varRd@long_name,"PEC"))) then
  varRd = varRd/1.0E3  ; convert from g to kg
  varRd@NEI11_name = "BC"
  varRd@units= "kg/m2/s"
else if (.not. ismissing(str_match(varRd@long_name,"PSO4"))) then
  varRd = varRd/1.0E3  ; convert from g to kg
  varRd@NEI11_name = "SO4"
  varRd@units= "kg/m2/s"
else if (.not. ismissing(str_match(varRd@long_name,"PRPA"))) then
  varRd = varRd *12.0*3.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "C3H8"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*3"
```

```
else if (.not. ismissing(str_match(varRd@long_name,"ACET"))) then
  varRd = varRd *12.0*3.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "ACET"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*3"
else if (.not. ismissing(str_match(varRd@long_name,"KET"))) then
  varRd = varRd *12*4/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "MEK"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*4"
else if (.not. ismissing(str_match(varRd@long_name,"NH3_FERT"))) then
  varRd = varRd *17.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "NH3_FERT"
  varRd@units= "kg/m2/s"
  varRd@molecular_weight_g_mol= "17"
else if (.not. ismissing(str_match(varRd@long_name,"ALD2_PRIMARY"))) then
  varRd = varRd *12.0*2.0/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "ALD2_PRIMARY"
  varRd@units= "kgC/m2/s"
  varRd@molecular_weight_g_mol= "12*2"
else if (.not. ismissing(str_match(varRd@long_name,"FORM_PRIMARY"))) then
  varRd = varRd *30.03/1.0E3  ; convert from moles to kg
  varRd@NEI11_name = "CH2O_PRIMARY"
  varRd@units= "kg/m2/s"
  varRd@molecular_weight_g_mol= "30"
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
end if
```

```
            end if
            end if
            end if
            end if
            end if
            end if
            end if
            end if
            end if
            end if
            end if
            end if

            varRd@_FillValue = 0.0
            delete(varRd@_FillValue)
            ncdf->$VARLIST(i)$ = varRd
            delete(varRd)
            delete(varIn)
            delete(varTmp)
          end do ; end looping over VARLIST

        counter = counter+1

        print("------------------------------------------------------------")
        print("  The number of "+source+ "files regridded is "+counter)
        print("------------------------------------------------------------")

          end do ; end looping over file


       ; add attributes to netCDF
          ;ncdf->time=time

end do ; end looping over sources
end
```

Python script to regrid the 2011NEIv6.3ek point emission source data from Lambert conformal conic projection at a 12 km x 12 km grid to a rectilinear 0.1x0.1 degree grid that can be read by GEOS-Chem

```
#!/usr/bin/env python
import os
import shutil
import argparse

from collections import defaultdict

import numpy as np
from netCDF4 import Dataset
from datetime import datetime

parser = argparse.ArgumentParser()
parser.add_argument('--template', default = 'NEI11_0.1x0.1_template.nc', help = 'An existing
gridded file')
parser.add_argument('--date', help = 'Date in YYYYMMDD format for output file')
parser.add_argument('stackpath', help = 'Point file meta-data input path')
parser.add_argument('inlnpath', nargs = '+', help = 'Point file emissions input path')
args = parser.parse_args()

renamer = dict(BENZ = 'BENZ')
lon = np.array([-139.94999695, -139.84999084, -139.75      , -139.6499939 ,
                -139.55000305, -139.44999695, -139.34999084, -139.25      ,
                -139.1499939 , -139.05000305, -138.94999695, -138.84999084,
                -138.75      , -138.6499939 , -138.55000305, -138.44999695,
                -138.34999084, -138.25      , -138.1499939 , -138.05000305,
                -137.94999695, -137.84999084, -137.75      , -137.6499939 ,
                -137.55000305, -137.44999695, -137.34999084, -137.25      ,
                -137.1499939 , -137.05000305, -136.94999695, -136.84999084,
                -136.75      , -136.6499939 , -136.55000305, -136.44999695,
                -136.34999084, -136.25      , -136.1499939 , -136.05000305,
                -135.94999695, -135.84999084, -135.75      , -135.6499939 ,
                -135.55000305, -135.44999695, -135.34999084, -135.25      ,
                -135.1499939 , -135.05000305, -134.94999695, -134.84999084,
                -134.75      , -134.6499939 , -134.55000305, -134.44999695,
                -134.34999084, -134.25      , -134.1499939 , -134.05000305,
                -133.94999695, -133.84999084, -133.75      , -133.6499939 ,
                -133.55000305, -133.44999695, -133.34999084, -133.25      ,
                -133.1499939 , -133.05000305, -132.94999695, -132.84999084,
                -132.75      , -132.6499939 , -132.55000305, -132.44999695,
                -132.34999084, -132.25      , -132.1499939 , -132.05000305,
                -131.94999695, -131.84999084, -131.75      , -131.6499939 ,
```

-131.55000305, -131.44999695, -131.34999084, -131.25 ,
-131.1499939 , -131.05000305, -130.94999695, -130.84999084,
-130.75 , -130.6499939 , -130.55000305, -130.44999695,
-130.34999084, -130.25 , -130.1499939 , -130.05000305,
-129.94999695, -129.84999084, -129.75 , -129.6499939 ,
-129.55000305, -129.44999695, -129.34999084, -129.25 ,
-129.1499939 , -129.05000305, -128.94999695, -128.84999084,
-128.75 , -128.6499939 , -128.55000305, -128.44999695,
-128.34999084, -128.25 , -128.1499939 , -128.05000305,
-127.94999695, -127.84999847, -127.75 , -127.6499939 ,
-127.54999542, -127.44999695, -127.34999847, -127.25 ,
-127.1499939 , -127.04999542, -126.94999695, -126.84999847,
-126.75 , -126.6499939 , -126.54999542, -126.44999695,
-126.34999847, -126.25 , -126.1499939 , -126.04999542,
-125.94999695, -125.84999847, -125.75 , -125.6499939 ,
-125.54999542, -125.44999695, -125.34999847, -125.25 ,
-125.1499939 , -125.04999542, -124.94999695, -124.84999847,
-124.75 , -124.6499939 , -124.54999542, -124.44999695,
-124.34999847, -124.25 , -124.1499939 , -124.04999542,
-123.94999695, -123.84999847, -123.75 , -123.6499939 ,
-123.54999542, -123.44999695, -123.34999847, -123.25 ,
-123.1499939 , -123.04999542, -122.94999695, -122.84999847,
-122.75 , -122.6499939 , -122.54999542, -122.44999695,
-122.34999847, -122.25 , -122.1499939 , -122.04999542,
-121.94999695, -121.84999847, -121.75 , -121.6499939 ,
-121.54999542, -121.44999695, -121.34999847, -121.25 ,
-121.1499939 , -121.04999542, -120.94999695, -120.84999847,
-120.75 , -120.6499939 , -120.54999542, -120.44999695,
-120.34999847, -120.25 , -120.1499939 , -120.04999542,
-119.94999695, -119.84999847, -119.75 , -119.6499939 ,
-119.54999542, -119.44999695, -119.34999847, -119.25 ,
-119.1499939 , -119.04999542, -118.94999695, -118.84999847,
-118.75 , -118.6499939 , -118.54999542, -118.44999695,
-118.34999847, -118.25 , -118.1499939 , -118.04999542,
-117.94999695, -117.84999847, -117.75 , -117.6499939 ,
-117.54999542, -117.44999695, -117.34999847, -117.25 ,
-117.1499939 , -117.04999542, -116.94999695, -116.84999847,
-116.75 , -116.6499939 , -116.54999542, -116.44999695,
-116.34999847, -116.25 , -116.1499939 , -116.04999542,
-115.94999695, -115.84999847, -115.75 , -115.6499939 ,
-115.54999542, -115.44999695, -115.34999847, -115.25 ,
-115.1499939 , -115.04999542, -114.94999695, -114.84999847,
-114.75 , -114.6499939 , -114.54999542, -114.44999695,
-114.34999847, -114.25 , -114.1499939 , -114.04999542,
-113.94999695, -113.84999847, -113.75 , -113.6499939 ,
-113.54999542, -113.44999695, -113.34999847, -113.25 ,

-113.1499939 , -113.04999542, -112.94999695, -112.84999847,
-112.75 , -112.6499939 , -112.54999542, -112.44999695,
-112.34999847, -112.25 , -112.1499939 , -112.04999542,
-111.94999695, -111.84999847, -111.75 , -111.6499939 ,
-111.54999542, -111.44999695, -111.34999847, -111.25 ,
-111.1499939 , -111.04999542, -110.94999695, -110.84999847,
-110.75 , -110.6499939 , -110.54999542, -110.44999695,
-110.34999847, -110.25 , -110.1499939 , -110.04999542,
-109.94999695, -109.84999847, -109.75 , -109.65000153,
-109.54999542, -109.44999695, -109.34999847, -109.25 ,
-109.15000153, -109.04999542, -108.94999695, -108.84999847,
-108.75 , -108.65000153, -108.54999542, -108.44999695,
-108.34999847, -108.25 , -108.15000153, -108.04999542,
-107.94999695, -107.84999847, -107.75 , -107.65000153,
-107.54999542, -107.44999695, -107.34999847, -107.25 ,
-107.15000153, -107.04999542, -106.94999695, -106.84999847,
-106.75 , -106.65000153, -106.54999542, -106.44999695,
-106.34999847, -106.25 , -106.15000153, -106.04999542,
-105.94999695, -105.84999847, -105.75 , -105.65000153,
-105.54999542, -105.44999695, -105.34999847, -105.25 ,
-105.15000153, -105.04999542, -104.94999695, -104.84999847,
-104.75 , -104.65000153, -104.54999542, -104.44999695,
-104.34999847, -104.25 , -104.15000153, -104.04999542,
-103.94999695, -103.84999847, -103.75 , -103.65000153,
-103.54999542, -103.44999695, -103.34999847, -103.25 ,
-103.15000153, -103.04999542, -102.94999695, -102.84999847,
-102.75 , -102.65000153, -102.54999542, -102.44999695,
-102.34999847, -102.25 , -102.15000153, -102.04999542,
-101.94999695, -101.84999847, -101.75 , -101.65000153,
-101.54999542, -101.44999695, -101.34999847, -101.25 ,
-101.15000153, -101.04999542, -100.94999695, -100.84999847,
-100.75 , -100.65000153, -100.54999542, -100.44999695,
-100.34999847, -100.25 , -100.15000153, -100.04999542,
-99.94999695, -99.84999847, -99.75 , -99.65000153,
-99.54999542, -99.44999695, -99.34999847, -99.25 ,
-99.15000153, -99.04999542, -98.94999695, -98.84999847,
-98.75 , -98.65000153, -98.54999542, -98.44999695,
-98.34999847, -98.25 , -98.15000153, -98.04999542,
-97.94999695, -97.84999847, -97.75 , -97.65000153,
-97.54999542, -97.44999695, -97.34999847, -97.25 ,
-97.15000153, -97.04999542, -96.94999695, -96.84999847,
-96.75 , -96.65000153, -96.54999542, -96.44999695,
-96.34999847, -96.25 , -96.15000153, -96.04999542,
-95.94999695, -95.84999847, -95.75 , -95.65000153,
-95.54999542, -95.44999695, -95.34999847, -95.25 ,
-95.15000153, -95.04999542, -94.94999695, -94.84999847,

-94.75 , -94.65000153, -94.54999542, -94.44999695,
-94.34999847, -94.25 , -94.15000153, -94.04999542,
-93.94999695, -93.84999847, -93.75 , -93.65000153,
-93.54999542, -93.44999695, -93.34999847, -93.25 ,
-93.15000153, -93.04999542, -92.94999695, -92.84999847,
-92.75 , -92.65000153, -92.54999542, -92.44999695,
-92.34999847, -92.25 , -92.15000153, -92.04999542,
-91.94999695, -91.84999847, -91.75 , -91.65000153,
-91.54999542, -91.44999695, -91.34999847, -91.25 ,
-91.15000153, -91.04999542, -90.94999695, -90.84999847,
-90.75 , -90.65000153, -90.54999542, -90.44999695,
-90.34999847, -90.25 , -90.15000153, -90.04999542,
-89.94999695, -89.84999847, -89.75 , -89.65000153,
-89.54999542, -89.44999695, -89.34999847, -89.25 ,
-89.15000153, -89.04999542, -88.94999695, -88.84999847,
-88.75 , -88.65000153, -88.54999542, -88.44999695,
-88.34999847, -88.25 , -88.15000153, -88.04999542,
-87.94999695, -87.84999847, -87.75 , -87.65000153,
-87.54999542, -87.44999695, -87.34999847, -87.25 ,
-87.15000153, -87.04999542, -86.94999695, -86.84999847,
-86.75 , -86.65000153, -86.54999542, -86.44999695,
-86.34999847, -86.25 , -86.15000153, -86.04999542,
-85.94999695, -85.84999847, -85.75 , -85.65000153,
-85.54999542, -85.44999695, -85.34999847, -85.25 ,
-85.15000153, -85.04999542, -84.94999695, -84.84999847,
-84.75 , -84.65000153, -84.54999542, -84.44999695,
-84.34999847, -84.25 , -84.15000153, -84.04999542,
-83.94999695, -83.84999847, -83.75 , -83.65000153,
-83.54999542, -83.44999695, -83.34999847, -83.25 ,
-83.15000153, -83.04999542, -82.94999695, -82.84999847,
-82.75 , -82.65000153, -82.54999542, -82.44999695,
-82.34999847, -82.25 , -82.15000153, -82.04999542,
-81.94999695, -81.84999847, -81.75 , -81.65000153,
-81.54999542, -81.44999695, -81.34999847, -81.25 ,
-81.15000153, -81.04999542, -80.94999695, -80.84999847,
-80.75 , -80.65000153, -80.54999542, -80.44999695,
-80.34999847, -80.25 , -80.15000153, -80.04999542,
-79.94999695, -79.84999847, -79.75 , -79.65000153,
-79.54999542, -79.44999695, -79.34999847, -79.25 ,
-79.15000153, -79.04999542, -78.94999695, -78.84999847,
-78.75 , -78.65000153, -78.54999542, -78.44999695,
-78.34999847, -78.25 , -78.15000153, -78.04999542,
-77.94999695, -77.84999847, -77.75 , -77.65000153,
-77.54999542, -77.44999695, -77.34999847, -77.25 ,
-77.15000153, -77.04999542, -76.94999695, -76.84999847,
-76.75 , -76.65000153, -76.54999542, -76.44999695,

-76.34999847, -76.25 , -76.15000153, -76.04999542,
-75.94999695, -75.84999847, -75.75 , -75.65000153,
-75.54999542, -75.44999695, -75.34999847, -75.25 ,
-75.15000153, -75.04999542, -74.94999695, -74.84999847,
-74.75 , -74.65000153, -74.54999542, -74.44999695,
-74.34999847, -74.25 , -74.15000153, -74.04999542,
-73.94999695, -73.84999847, -73.75 , -73.65000153,
-73.54999542, -73.44999695, -73.34999847, -73.25 ,
-73.15000153, -73.04999542, -72.94999695, -72.84999847,
-72.75 , -72.65000153, -72.54999542, -72.44999695,
-72.34999847, -72.25 , -72.15000153, -72.04999542,
-71.94999695, -71.84999847, -71.75 , -71.65000153,
-71.54999542, -71.44999695, -71.34999847, -71.25 ,
-71.15000153, -71.04999542, -70.94999695, -70.84999847,
-70.75 , -70.65000153, -70.54999542, -70.44999695,
-70.34999847, -70.25 , -70.15000153, -70.04999542,
-69.94999695, -69.84999847, -69.75 , -69.65000153,
-69.54999542, -69.44999695, -69.34999847, -69.25 ,
-69.15000153, -69.04999542, -68.94999695, -68.84999847,
-68.75 , -68.65000153, -68.54999542, -68.44999695,
-68.34999847, -68.25 , -68.15000153, -68.04999542,
-67.94999695, -67.84999847, -67.75 , -67.65000153,
-67.54999542, -67.44999695, -67.34999847, -67.25 ,
-67.15000153, -67.04999542, -66.94999695, -66.84999847,
-66.75 , -66.65000153, -66.54999542, -66.44999695,
-66.34999847, -66.25 , -66.15000153, -66.04999542,
-65.94999695, -65.84999847, -65.75 , -65.65000153,
-65.54999542, -65.44999695, -65.34999847, -65.25 ,
-65.15000153, -65.04999542, -64.94999695, -64.84999847,
-64.75 , -64.65000153, -64.54999542, -64.44999695,
-64.34999847, -64.25 , -64.15000153, -64.04999542,
-63.95000076, -63.84999847, -63.75 , -63.64999771,
-63.54999924, -63.45000076, -63.34999847, -63.25 ,
-63.14999771, -63.04999924, -62.95000076, -62.84999847,
-62.75 , -62.64999771, -62.54999924, -62.45000076,
-62.34999847, -62.25 , -62.14999771, -62.04999924,
-61.95000076, -61.84999847, -61.75 , -61.64999771,
-61.54999924, -61.45000076, -61.34999847, -61.25 ,
-61.14999771, -61.04999924, -60.95000076, -60.84999847,
-60.75 , -60.64999771, -60.54999924, -60.45000076,
-60.34999847, -60.25 , -60.14999771, -60.04999924,
-59.95000076, -59.84999847, -59.75 , -59.64999771,
-59.54999924, -59.45000076, -59.34999847, -59.25 ,
-59.14999771, -59.04999924, -58.95000076, -58.84999847,
-58.75 , -58.64999771, -58.54999924, -58.45000076,
-58.34999847, -58.25 , -58.14999771, -58.04999924,

```
               -57.95000076,  -57.84999847,  -57.75      ,  -57.64999771,
               -57.54999924,  -57.45000076,  -57.34999847,  -57.25      ,
               -57.14999771,  -57.04999924,  -56.95000076,  -56.84999847,
               -56.75      ,  -56.64999771,  -56.54999924,  -56.45000076,
               -56.34999847,  -56.25      ,  -56.14999771,  -56.04999924,
               -55.95000076,  -55.84999847,  -55.75      ,  -55.64999771,
               -55.54999924,  -55.45000076,  -55.34999847,  -55.25      ,
               -55.14999771,  -55.04999924,  -54.95000076,  -54.84999847,
               -54.75      ,  -54.64999771,  -54.54999924,  -54.45000076,
               -54.34999847,  -54.25      ,  -54.14999771,  -54.04999924,
               -53.95000076,  -53.84999847,  -53.75      ,  -53.64999771,
               -53.54999924,  -53.45000076,  -53.34999847,  -53.25      ,
               -53.14999771,  -53.04999924,  -52.95000076,  -52.84999847,
               -52.75      ,  -52.64999771,  -52.54999924,  -52.45000076,
               -52.34999847,  -52.25      ,  -52.14999771,  -52.04999924,
               -51.95000076,  -51.84999847,  -51.75      ,  -51.64999771,
               -51.54999924,  -51.45000076,  -51.34999847,  -51.25      ,
               -51.14999771,  -51.04999924,  -50.95000076,  -50.84999847,
               -50.75      ,  -50.64999771,  -50.54999924,  -50.45000076,
               -50.34999847,  -50.25      ,  -50.14999771,  -50.04999924])

lat = np.array([ 20.04999924,  20.14999962,  20.25      ,  20.34999847,
               20.44999886,  20.54999924,  20.64999962,  20.75      ,
               20.84999847,  20.94999886,  21.04999924,  21.14999962,
               21.25      ,  21.34999847,  21.44999886,  21.54999924,
               21.64999962,  21.75      ,  21.84999847,  21.94999886,
               22.04999924,  22.14999962,  22.25      ,  22.34999847,
               22.44999886,  22.54999924,  22.64999962,  22.75      ,
               22.84999847,  22.94999886,  23.04999924,  23.14999962,
               23.25      ,  23.34999847,  23.44999886,  23.54999924,
               23.64999962,  23.75      ,  23.84999847,  23.94999886,
               24.04999924,  24.14999962,  24.25      ,  24.34999847,
               24.44999886,  24.54999924,  24.64999962,  24.75      ,
               24.84999847,  24.94999886,  25.04999924,  25.14999962,
               25.25      ,  25.35000038,  25.44999886,  25.54999924,
               25.64999962,  25.75      ,  25.85000038,  25.94999886,
               26.04999924,  26.14999962,  26.25      ,  26.35000038,
               26.44999886,  26.54999924,  26.64999962,  26.75      ,
               26.85000038,  26.94999886,  27.04999924,  27.14999962,
               27.25      ,  27.35000038,  27.44999886,  27.54999924,
               27.64999962,  27.75      ,  27.85000038,  27.94999886,
               28.04999924,  28.14999962,  28.25      ,  28.35000038,
               28.44999886,  28.54999924,  28.64999962,  28.75      ,
               28.85000038,  28.94999886,  29.04999924,  29.14999962,
               29.25      ,  29.35000038,  29.44999886,  29.54999924,
               29.64999962,  29.75      ,  29.85000038,  29.94999886,
```

30.04999924, 30.14999962, 30.25 , 30.35000038,
30.44999886, 30.54999924, 30.64999962, 30.75 ,
30.85000038, 30.94999886, 31.04999924, 31.14999962,
31.25 , 31.35000038, 31.44999886, 31.54999924,
31.64999962, 31.75 , 31.85000038, 31.94999886,
32.04999924, 32.15000153, 32.25 , 32.34999847,
32.45000076, 32.54999924, 32.65000153, 32.75 ,
32.84999847, 32.95000076, 33.04999924, 33.15000153,
33.25 , 33.34999847, 33.45000076, 33.54999924,
33.65000153, 33.75 , 33.84999847, 33.95000076,
34.04999924, 34.15000153, 34.25 , 34.34999847,
34.45000076, 34.54999924, 34.65000153, 34.75 ,
34.84999847, 34.95000076, 35.04999924, 35.15000153,
35.25 , 35.34999847, 35.45000076, 35.54999924,
35.65000153, 35.75 , 35.84999847, 35.95000076,
36.04999924, 36.15000153, 36.25 , 36.34999847,
36.45000076, 36.54999924, 36.65000153, 36.75 ,
36.84999847, 36.95000076, 37.04999924, 37.15000153,
37.25 , 37.34999847, 37.45000076, 37.54999924,
37.65000153, 37.75 , 37.84999847, 37.95000076,
38.04999924, 38.15000153, 38.25 , 38.34999847,
38.45000076, 38.54999924, 38.65000153, 38.75 ,
38.84999847, 38.95000076, 39.04999924, 39.15000153,
39.25 , 39.34999847, 39.45000076, 39.54999924,
39.65000153, 39.75 , 39.84999847, 39.95000076,
40.04999924, 40.15000153, 40.25 , 40.34999847,
40.45000076, 40.54999924, 40.65000153, 40.75 ,
40.84999847, 40.95000076, 41.04999924, 41.15000153,
41.25 , 41.34999847, 41.45000076, 41.54999924,
41.65000153, 41.75 , 41.84999847, 41.95000076,
42.04999924, 42.15000153, 42.25 , 42.34999847,
42.45000076, 42.54999924, 42.65000153, 42.75 ,
42.84999847, 42.95000076, 43.04999924, 43.15000153,
43.25 , 43.34999847, 43.45000076, 43.54999924,
43.65000153, 43.75 , 43.84999847, 43.95000076,
44.04999924, 44.15000153, 44.25 , 44.34999847,
44.45000076, 44.54999924, 44.65000153, 44.75 ,
44.84999847, 44.95000076, 45.04999924, 45.15000153,
45.25 , 45.34999847, 45.45000076, 45.54999924,
45.65000153, 45.75 , 45.84999847, 45.95000076,
46.04999924, 46.15000153, 46.25 , 46.34999847,
46.45000076, 46.54999924, 46.65000153, 46.75 ,
46.84999847, 46.95000076, 47.04999924, 47.15000153,
47.25 , 47.34999847, 47.45000076, 47.54999924,
47.65000153, 47.75 , 47.84999847, 47.95000076,
48.04999924, 48.15000153, 48.25 , 48.34999847,

```
            48.45000076,  48.54999924,  48.65000153,  48.75     ,
            48.84999847,  48.95000076,  49.04999924,  49.15000153,
            49.25     ,  49.34999847,  49.45000076,  49.54999924,
            49.65000153,  49.75     ,  49.84999847,  49.95000076,
            50.04999924,  50.15000153,  50.25     ,  50.35000229,
            50.45000076,  50.54999924,  50.65000153,  50.75     ,
            50.85000229,  50.95000076,  51.04999924,  51.15000153,
            51.25     ,  51.35000229,  51.45000076,  51.54999924,
            51.65000153,  51.75     ,  51.85000229,  51.95000076,
            52.04999924,  52.15000153,  52.25     ,  52.35000229,
            52.45000076,  52.54999924,  52.65000153,  52.75     ,
            52.85000229,  52.95000076,  53.04999924,  53.15000153,
            53.25     ,  53.35000229,  53.45000076,  53.54999924,
            53.65000153,  53.75     ,  53.85000229,  53.95000076,
            54.04999924,  54.15000153,  54.25     ,  54.35000229,
            54.45000076,  54.54999924,  54.65000153,  54.75     ,
            54.85000229,  54.95000076,  55.04999924,  55.15000153,
            55.25     ,  55.35000229,  55.45000076,  55.54999924,
            55.65000153,  55.75     ,  55.85000229,  55.95000076,
            56.04999924,  56.15000153,  56.25     ,  56.35000229,
            56.45000076,  56.54999924,  56.65000153,  56.75     ,
            56.85000229,  56.95000076,  57.04999924,  57.15000153,
            57.25     ,  57.35000229,  57.45000076,  57.54999924,
            57.65000153,  57.75     ,  57.85000229,  57.95000076,
            58.04999924,  58.15000153,  58.25     ,  58.35000229,
            58.45000076,  58.54999924,  58.65000153,  58.75     ,
            58.85000229,  58.95000076,  59.04999924,  59.15000153,
            59.25     ,  59.35000229,  59.45000076,  59.54999924,
            59.65000153,  59.75     ,  59.85000229,  59.95000076])

xres = 0.1
hxres = xres / 2
late = np.append(lat - hxres, lat[-1] + hxres)
lone = np.append(lon - hxres, lon[-1] + hxres)
Re = 6375000.0
latr = late * np.pi / 180
nlon = 360. / xres
area = 2. * np.pi * Re * Re / (nlon) * ( np.sin( latr[1:] ) - np.sin( latr[:-1] ) )
AREA = area[:, None].repeat(lon.size, 1)
print('Total Area', AREA.sum())
print('Mean Area', AREA.mean())
print('Mean Length', AREA.mean()**.5)
LATE, LONE = np.meshgrid(late, lone)
leve = np.array([-0.006, .123, .254, .387, .521, .657, .795, .934, 1.075]) * 1000. # in meters

oldstack = Dataset(args.stackpath, 'r')
```

```
stklats = oldstack.variables['LATITUDE'][0, 0, :, 0]
stklons = oldstack.variables['LONGITUDE'][0, 0, :, 0]
stkhgts = oldstack.variables['STKHT'][0, 0, :, 0]

kji_cells = defaultdict(lambda: [])
for stkidx, (stkhgt, stklat, stklon) in enumerate(zip(stkhgts, stklats[:], stklons[:])):
    levk = np.where((stkhgt >= leve[:-1]) & (stkhgt < leve[1:]))[0]
    latj = np.where((stklat >= late[:-1]) & (stklat < late[1:]))[0]
    loni = np.where((stklon >= lone[:-1]) & (stklon < lone[1:]))[0]
    assert(latj.size == 1)
    assert(loni.size == 1)
    assert(levk.size == 1)
    kji_cells[levk[0], latj[0], loni[0]].append(stkidx)

stackks = [k for k, j, i in kji_cells.keys()]
kji_cells = dict([(k, np.array(v)) for k, v in kji_cells.items()])
#kji = set(zip(stackks, stackjs, stackis))
print('Unique Locations', len(kji_cells))
#kji_cells = {}
#for k, j, i in kji:
#    idx = (stackks == k) & (stackjs == j) & (stackis == i)
#    kji_cells[k, j, i] = idx

print('Found', sum([idx.size for idx in kji_cells.values()]), 'of', len(oldstack.dimensions['ROW']))
molwt_outunit = {'HCL': ('36.45', 'kg/m2/s'), 'NAPHTHALENE': ('12*10', 'kgC/m2/s'), 'ETHY':
('12*2', 'kgC/m2/s'), 'TERP': ('12*10', 'kgC/m2/s'), 'CH4': ('16', 'kg/m2/s'), 'CL2': ('35.45*2',
'kg/m2/s'), 'ACROLEIN': ('56', 'kg/m2/s'), 'BUTADIENE13': ('12*4', 'kgC/m2/s'), 'PRPA':
('12*3', 'kgC/m2/s'), 'ISOP': ('12*5', 'kgC/m2/s'), 'KET': ('12*4', 'kgC/m2/s'), 'ACET': ('12*3',
'kgC/m2/s'), 'PAR': ('12*4', 'kgC/m2/s'), 'ETH': ('12*2', 'kgC/m2/s'), 'NH3': ('17', 'kg/m2/s'),
'ALD2': ('12*2', 'kgC/m2/s'), 'ETOH': ('46', 'kg/m2/s'), 'XYL': ('12*8', 'kgC/m2/s'), 'OLE': ('12*2',
'kgC/m2/s'), 'NO2': ('46', 'kg/m2/s'), 'ALDX': ('58', 'kg/m2/s'), 'SO2': ('64', 'kg/m2/s'), 'BENZ':
('12*6', 'kgC/m2/s'), 'NO': ('30', 'kg/m2/s'),  'IOLE': ('12*4', 'kgC/m2/s')  , 'ETHA': ('12*2',
'kgC/m2/s'), 'CO': ('28', 'kg/m2/s'), 'FORM': ('30.03', 'kg/m2/s'), 'MEOH': ('32', 'kg/m2/s'),
'HONO': ('47', 'kg/m2/s'), 'TOL': ('12*7', 'kgC/m2/s'), 'SULF': ('98', 'kg/m2/s'), 'POC': ('1',
'kg/m2/s'), 'PEC': ('1', 'kg/m2/s'), 'FORM_PRIMARY': ('30.03', 'kg/m2/s'), 'ALD2_PRIMARY':
('12*2', 'kgC/m2/s'), 'NH3_FERT': ('17', 'kg/m2/s')}
variables = {}
units = {}
nlat = lat.size
nlon = lon.size
nlev = len(np.unique(stackks))
for inlnpath in args.inlnpath:
    oldinln = Dataset(inlnpath, 'r+s')
    inkeys = set(getattr(oldinln, 'VAR-LIST').split())
    for vark in inkeys:
        invaro = oldinln.variables[vark]
```

```
        if not vark in molwt_outunit:
            if invaro.units.strip() == 'g/s':
                print('Adding', vark)
                molwt_outunit[vark] = ('1', 'kg/m2/s')
            else:
                continue
        print('Gridding', vark)
        if vark not in variables:
            variables[vark] = np.zeros((25, nlev, nlat, nlon), dtype = 'd')
            units[vark] = invaro.units.strip()
        outvaro = variables[vark]
        indata = invaro[:, 0, :, 0]
        for (stkk, stkj, stki), idx in kji_cells.items():
            #import pdb; pdb.set_trace()
            thisdata = indata[:, idx].sum(1)
            outvaro[:, stkk, stkj, stki] += thisdata
    regridkeys = set(molwt_outunit.keys())
    print('Skipping', inkeys.difference(regridkeys))
    print('Missing', regridkeys.difference(inkeys))
    if args.date is None:
        sdate = datetime.strptime('%07d %06d' % (oldinln.SDATE, oldinln.STIME), '%Y%j
%H%M%S')
    else:
        sdate = datetime.strptime(args.date, '%Y%m%d')
    rdate = datetime(1970, 1, 1)
    idt = (sdate - rdate).total_seconds() / 3600.

    newpath = 'GRIDDED_' + os.path.basename(inlnpath)
    newneigridded = Dataset(newpath, 'w')
    newneigridded.createDimension('time', None)
    newneigridded.createDimension('lev', nlev)
    newneigridded.createDimension('lat', nlat)
    newneigridded.createDimension('lon', nlon)
    timev = newneigridded.createVariable('time', 'd', ('time',))
    timev.units = 'hours since 2000-01-01 00:00:00'
    timev.long_name = 'Time'
    timev[0:23] = idt + np.arange(23)
    print('this is the date', timev)
    latv = newneigridded.createVariable('lat', 'd', ('lat',))
    latv.units = 'degrees_north'
    latv[:] = lat
    lonv = newneigridded.createVariable('lon', 'd', ('lon',))
    lonv.units = 'degrees_east'
    lonv[:] = lon
    levv = newneigridded.createVariable('lev', 'd', ('lev',))
    levv.units = 'model levels'
```

```
levv.long_name = "GEOS-Chem levels"
levv[:] = np.arange(nlev)
areav = newneigridded.createVariable('area', 'd', ('lat', 'lon'))
areav.units = 'm**2'
areav[:] = AREA

for vark, vardata in variables.items():
   print('Processing', vark, end = ' ')
   molwt, outunit = molwt_outunit[vark]
   inunit = units[vark]
   if inunit in ('moles/s', 'g/s'):
      factor = eval(molwt) / 1000 # g/mol -> kg/mol
      factor = factor / AREA # kg/mol/m2
   else:
      raise ValueError('inunit = ' + inunit)
   print('factor', factor.mean())
   outk = renamer.get(vark, vark)
   varo = newneigridded.createVariable(outk, 'f', ('time', 'lev', 'lat', 'lon'))
   varo.long_name = outk
   varo.units = outunit
   varo.molecular_weight_g_mol = molwt
   varo[:]  = vardata * factor
newneigridded.sync()
```

Table B1: C_2-C_5 alkane % contribution to total PAR based on top VOC emissions in the U.S.

Top 50 VOC emissions*	PAR (CB05)	Tons day^{-1}*	M (g mol^{-1})	TonsC day^{-1}	ALK4 (GEOS-Chem species)	A**	B***
toluene	NA	3.49E+03	NA	NA	NA		
n-butane	4	2.88E+03	58.12	2.38E+03	2380	17.07	19.61
mineral spirits	NA	2.68E+03	NA	NA			
xylenes	NA	2.57E+03	NA	NA			
iso-pentae	5	1.62E+03	72.15	1350	1350	9.69	11.12
ethene	0	1.48E+03	NA	NA	NA		
acetone	3	1.45E+03	58.08	901		6	
benzene	1	1.28E+03	78.11	197		1	
ethane	0	1.22E+03	30.07	NA			
ethanol	0	1.15E+03	46.06844	NA			
3-methyl-1butene	3	1.01E+03	70.14	521	521	4	4.29
propane	1.5	9.55E+02	44.1	390		3	
2-methylpentane	6	9.11E+02	86.18	761	761	5	6.27
n-pentane	5	8.88E+02	72.15	739	739	5	6.09
dicholomethane	0	8.81E+02	84.93	NA			
propene	1	8.44E+02	42.08	241	241	2	1.99
n-hexane	6	8.21E+02	86.18	686	686	5	5.65
ethyl benzene	1	7.84E+02	106.17	88.6	88.6	1	0.73
formaldehyde	0	7.54E+02	30.03	NA			
Unknown	NA	7.54E+02	NA	NA			
acetylene	1	7.17E+02	26.04	331	331	2	2.73
acetaldehyde	0	6.35E+02	44.05	NA			
isobutane	4	6.20E+02	58.12	512	512	4	4.22
2,2,4-trimethylpentane	7	6.13E+02	114.23	451	451	3	3.72
3-methylpentane	6	5.68E+02	86.18	475	475	3	3.91
methanol	NA	4.87E+02	32.04	NA			
methyl ethyl ketone (MEK)	4	4.64E+02	72.11	309		2	
n-heptane	7	4.49E+02	100.21	377	377	2.71	3.11
ethyl-3-ethylbenzene	1	4.05E+02	120.2	40.4	40.4	0.29	0.33
isopropanol	3	3.97E+02	60.1	238	238	1.71	1.96
1,2,4-trimethylbenzene	1	3.82E+02	120.2	38.2	38.2	0.27	0.31
1-butene & isobutene	3	3.60E+02	56.11	231	231	1.66	1.90
2-methyl-1-butene	5	3.60E+02	70.14	246	246	1.77	2.03
unidentified	NA	3.45E+02	NA	NA			

Top 50 VOC emissions*	PAR (CB05)	Tons day^{-1}*	M (g mol^{-1})	TonsC day^{-1}	ALK4 (GEOS-Chem species)	A**	B***
2,3-dimethylpentane	7	3.30E+02	100.21	277	277	1.99	2.28
methylcyclopentane	6	3.10E+02	84.16	270	270	1.94	2.22
2-methylhexane	7	3.08E+02	100.21	258	258	1.85	2.13
2,3-dimethylbutane	6	3.00E+02	86.18	251	251	1.80	2.07
2-methyl-2-butene	3	2.63E+02	70.14	135	135	0.97	1.11
2,3,4-trimethylpentane	8	2.63E+02	114.23	221	221	1.59	1.82
3-methylhexane	7	2.56E+02	100.21	214	214	1.54	1.76
n-octane	7	2.41E+02	114.23	177	177	1.27	1.46
2-butoxyethanol	6	2.41E+02	118.18	147	147	1.06	1.21
undefined VOC	NA	2.33E+02	NA	NA			
2,4-dimethylpentane	7	2.41E+02	100.21	202	202	1.45	1.66
trans-2-pentene	1	2.11E+02	70.14	36.1	36.1	0.26	0.30
1,3-butadiene	0	2.11E+02	54.09	NA			
propyl acetate	5	2.11E+02	102.13	99.2	99.2	0.71	0.82
2,2-dimethylbutane	6	2.04E+02	86.18	142	142	1.02	1.17
trichlorotrifluoroethane-F113	0	1.89E+02	187.38	NA			
			Sum	13931	12136	100.0	100.0

*: from Simon et al. (2010).
A**: % contribution to total PAR
B***: % contribution without acetone, benzene, propane, MEK (CB6-like consideration)

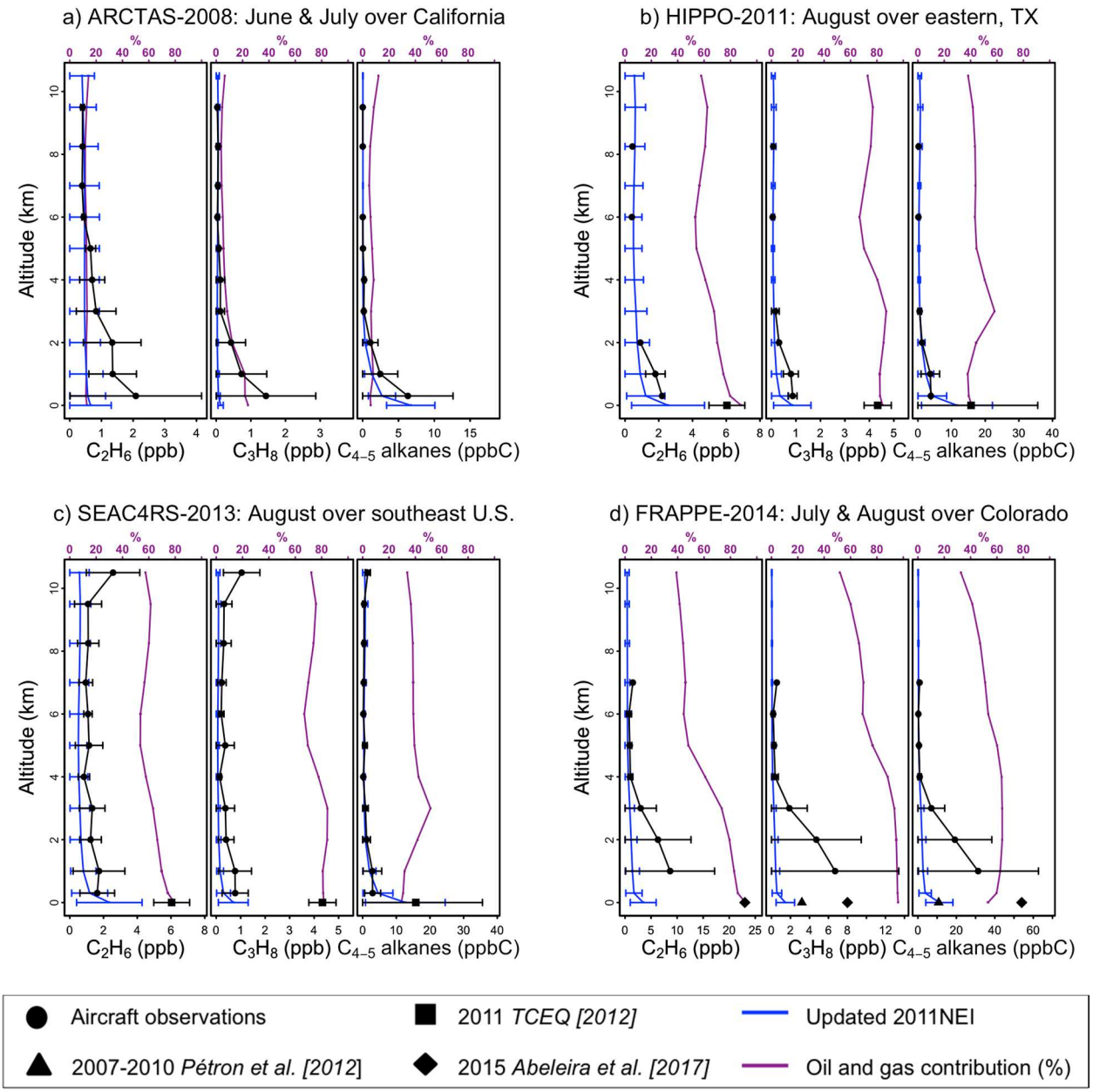


Figure B1: Vertical profiles and surface observations of C_2H_6, C_3H_8, and C_4-C_5 alkanes for regions and time periods shown in Table 3.3 and Figure 3.4. Black lines with circles represent mean observed abundances. Black triangles, rhomboids, and squares represent available mean summer surface observations. Blue lines represent 2011 modeled abundances and purple lines represent the percentage contribution from the oil and gas sector to total abundances of light alkanes. Horizontal lines denote the standard deviation of the observations in each vertical bin.

APPENDIX C

Table C1: Differences between 2011NEIv1 and updated 2011NEI U.S. emission totals for O_3 precursors emitted by anthropogenic sources.

O_3 precursor species	**Tg[1]**	**% change[2]**
NO	-1.5	-16.1
C_2H_6	0.3	24.8
C_3H_8	0.7	215.7
C_4-C_5 alkanes (36% of PAR)	2.1	30.9
Formaldehyde	-0.3	-61.7
Acetaldehyde	-0.1	-31.6
Acetone	-0.4	-58.2
≥C_3 alkenes	0.2	114.5
Methyl ethyl ketone (MEK)	-0.1	-46.1

Notes:
1. Updated 2011NEI – 2011NEIv1
2. % change with respect to 2011NEIv1: (updated 2011NEI *100/2011NEIv1)-100