

**Review of “Influence of oil and gas field operations on spatial and temporal distributions of atmospheric non-methane hydrocarbons and their effect on ozone formation in the winter” by Field et al. for publication in *Atmospheric Chemistry and Physics***

The author’s present a valuable dataset detailing measurements of speciated non-methane hydrocarbons (NMHC), methane, CO, NO<sub>x</sub>, and ozone in the Upper Green River Basin (UGRB) in Wyoming for 3 consecutive winters starting in late 2010. The aim of this study is to “better understand emission sources and to determine which compounds are most active in promoting [winter] ozone formation.” This work shows that the UGRB in the vicinity of the Jonah-Pinedale Anticline natural gas field experienced very high wintertime ozone (e.g., O<sub>3</sub> > 100 ppbv) in February 2011. This is in accordance with previous studies (e.g., Schnell et al., 2009 and Edwards et al. 2014) have shown that these wintertime ozone events are dependent upon large emissions of ozone precursors (volatile organic compounds and nitrogen oxides) from oil and natural gas operations. In order for O<sub>3</sub> buildup to occur, VOC and NO<sub>x</sub> emissions must be trapped in a shallow, stagnant surface layer of air that is exposed to increased actinic flux from the presence of snow. Documenting and understanding the wintertime ozone phenomenon is a critical step towards improving the local air quality in affected regions. To this end, the authors use in-situ measurements they conducted from October 2010 to March 2011 at the South Boulder Road site located south of Pinedale, Wyoming, as input to a positive matrix factorization (PMF) analysis. The PMF analysis was solved for 3 primary source categories (i.e., factors): “combustion/traffic,” “fugitive natural gas”, and “fugitive condensate” that were positively identified by comparison to known source profiles. Additionally, the author’s identify water treatment and well completion activities as a potentially significant sources of BTEX (benzene, toluene, ethyl benzene, and xylenes) that are not included in current emission inventories.

This work represents an important dataset that should be published. I recommend publication after major revisions.

Major comments:

1. A brief, but better, ‘connection to’ or ‘distinction of’ this study and previous studies within Wyoming’s UGRB should be described. For example, were the meteorological conditions of this study similar or different compared to Schnell et al. [2009]? Does this study provide measurements of NMHCs that were not included in the photochemical modeling of Carter and Seinfeld [2012]? Field et al. provides valuable information on the source apportionment of NMHC emissions in the UGRB, but the question remains if these source categories are broadly applicable to or independent of previous studies.

2. There is little discussion of the meteorological conditions for the winters of 2011 and 2012. The author's observed lower levels of NMHCs in 2012, but I would assume that has more to do with meteorology than changes in emissions. As noted in the introduction, previous studies have shown that meteorology (shallow temperature inversions, low wind speeds, presence of snow to increase actinic flux, etc.) is a key variable in determining the buildup of O<sub>3</sub> within western shale basins [Schnell 2009 and Edwards 2014]. Figure 4 clearly shows that high NO<sub>x</sub> (between 6-20 ppbv NO<sub>x</sub>) and NMHC (>0.2 ppmC) values will be present yet O<sub>3</sub> may still be below 40 ppbv indicating that there is another important variable that has not been adequately addressed.
3. One interesting observation is the occasional differences in O<sub>3</sub> profiles and peak O<sub>3</sub> levels measured at the Boulder South Road (BSR) and the Boulder (BLDR) sites when they are only 5 km apart. The authors point to the "importance of background, production, and transport [of O<sub>3</sub>];" however, not enough information is currently provided to show that (a) the O<sub>3</sub> levels *should* be similar but aren't or (b) the air mass trajectories were affected by different source regions of O<sub>3</sub> precursors, etc. For example, the observed ozone gradients between the sites could suggest inhomogeneity in O<sub>3</sub> precursor sources; however, the authors note that the NMHC composition is quite similar throughout the basin with the exception of the water treatment facility. If this is an important source of O<sub>3</sub> precursors, then it should be possible to show that the emissions from this facility had a larger impact on one site versus the other. Does geography or meteorology effect air flow over the 2 sites? Is there a significant difference in elevation between BSR and BLDR sites? If one site is lower than the other, does that site experience higher O<sub>3</sub> levels?
4. By comparing the total NMHC response and the sum of the speciated NMHC measurements, the authors show that the reported NMHCs accounted for 57% of the total NMHC and all detected NMHCs account for 88% of total NMHC. The authors state that the "missing mass is related to heavier NMHC" and that the "contributions of oxygenated VOCs...not detected by the OPA can only be small." This leads to the following questions:
  - a. Do temporal/diurnal trends of this "missing mass" give any indication that the source is primary vs. secondary? Based on the work of Edwards et al. (2014), oxygenated VOCs (e.g., carbonyls) have been shown to be photochemically produced in significant quantities from NMHC precursors and that these species are an important photolytic source of HO<sub>x</sub> radicals that contribute to the photochemical formation of winter O<sub>3</sub> in the Uintah Basin. I would expect oxygenated VOCs to play a non-negligible role in O<sub>3</sub> formation in Wyoming even if they may be a small fraction of the carbon mass.
  - b. Can this "missing mass" residual be included in the PMF analysis? If it is related to "heavier NMHC," then it may have a strong contribution from Factor 3.

- c. Can the measurements or PMF analysis identify a crude oil or raw natural gas signature/factor? Presumably each of these hydrocarbon sources have a different chemical composition based on volatility differences. If the two sources cannot be distinguished perhaps because they are co-produced at each well site (e.g., associated oil) then this should be specified in the discussion.
5. Could NO<sub>x</sub> measurements be included in the PMF analysis? It would be interesting to see if the predominate source is traffic dependent or if a signature from power generation at the well sites or processing facilities could be determined.

**Other comments (by order of occurrence):**

P24944 L5: Define an "ozone episode" (as done in section 3.2).

P24944 L6: What is the significance of the two sites experiencing different O<sub>3</sub> levels?

P24944 L7-8: "In 2012 the lack of O<sub>3</sub> episodes coincided with a reduction in ambient levels of NMHC." What else was different between 2011 and 2012? Did the emissions change? What about NO<sub>x</sub>?

P24944 L18: Define "condensate range NMHC."

P25945 L18: Start the discussion of winter O<sub>3</sub> in a new paragraph ("Until recently..." would be the start of paragraph #3 of introduction).

P24946 L25: A large portion of the first paragraph in the discussion section (Section 3; P24950 L15-26) should be moved to the beginning of the methodology section in order to accurately describe the measurement site locations and proximity to oil and natural gas wells. These details are critical to the study's methodology and do not belong in the results section.

P24947 L1: A more detailed description of the total NMHC instrument should be included (manufacturer, analytical range, duty cycle, etc.). How well does it measure OVOCs, halogens, nitrogen containing VOCs, etc.? The analytical capabilities of this instrument should be clearly stated for ease of comparison to the more detailed description of the OPE.

P24947 L4-6 and L8-12: All references to speciated NMHC measurements should be moved to Section 2.1.

P24947 L6-7: It is unclear if "similar measurements" refers to "discontinuous speciated NMHC measurements" or the full suite of "air quality" measurements.

P24947 L10: There is no description of the diffusive samplers included. Is the data from these samplers not included in the analysis?

P24950 L9: State the range of "extremely high outlier values."

P24951 L4: Be more specific regarding "numerous high hourly O<sub>3</sub> values." How many and what level constitutes "high O<sub>3</sub>?"

P24951 L9-10: A brief discussion of why NMHCs and O<sub>3</sub> are higher in 2011 compare to 2012 but NO<sub>x</sub> levels are similar should be included somewhere in this manuscript. Is it related to changes in emissions, oil and gas activities, meteorology, regulations, and/or chemistry?

P24951 L16: The discussion relates to “O<sub>3</sub> episodes” which are arbitrarily determined to be hourly averages greater than 85 ppbv, but this is a non-standard O<sub>3</sub> metric. How many O<sub>3</sub> exceedances have there been in WY (8-hr average > 75 ppbv)?

P24951 L16: I’m not convinced that these plots indicate the importance of background, production, and transport. Perhaps the sentence should be amended to state, “These plots indicate the variability in the background levels, local production, and transport of O<sub>3</sub> to the different measurement sites.”

P24954 L2-4: It is not clear which VOC measurements at BSR are used in the PMF analysis as it is not stated until P24955 L9. The data set used for the PMF analysis should be made clearer in this section as well as section 2.3.

P24955 L7-15: This portion of the discussion could be improved by simply stating the fact that Factor 2 dominates for methane through cyclohexane, Factor 3 dominates for the heavier VOCs heptane to o-xylene. Benzene and ethyne are the only 2 NMHCs that have a roughly equivalent contributions from all 3 factors. CO is the only species that is dominated by factor 1.

P24955 L11 and L15: Why these 14 NMHCs specifically? Why are others, e.g. nonane, excluded? This isn’t answered until L16.

P24955 L10: It isn’t appropriate to identify the factors until you discuss how you determined the identities (L15-20 of same page).

P24955 L14: “e.g.,” changed to “e.g.”

### **Figures/Tables:**

Figure 1: Add markers of all sampling sites to show the locations throughout the basin.

Figure 2: The left axes should be aligned vertically so that it is easier to compare the time series for O<sub>3</sub>, NO<sub>x</sub>, and NMHC for a specific year.

Figure 5: Define CO-b and methane-b in the figure caption. It is not defined until P24955 L22 when discussing Figure 6. What does “normalized mass profiles between 14 NMHCs” mean? Why only the stated 14 NMHCs when 22 are presented? Are you are normalizing all percentages to 100% in order to avoid showing circumstances when the calculated factors may be over/under 100%? Why not show the residuals to give an idea of how well the PMF analysis represents the actual measurements?