The referee provided excellent feedback that we have addressed in the responses given below. Their suggestions have helped to improve the clarity of our manuscript.

Major comments:

1. (a) 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]?

<Response> 1. (a) The last paragraph of the introduction has been modified to provide improved connection and distinction with other studies in the UGRB. A complete, thorough, in depth comparison analysis of the prevailing meteorological and other conditions during this study to those in other years would constitute a separate stand-alone paper, that would parallel the related statistical study carried out for the UGRB and Uintah basins by Mansfield and Hall (Air Quality, Atmosphere & Health, December 2013, Volume 6, Issue 4, pp 687-699). We have added some extra detail later in the paper (section 3.1) regarding meteorology, in particular as it pertains to 2011 to 2012 differences, as well as strengthening the paper by including the matrix used by Wyoming DEQ to define appropriate meteorological conditions for ozone production.

1. (b) Does this study provide measurements of NMHCs that were not included in the photochemical modeling of Carter and Seinfeld [2012]?

<Response> 1. (b) This study presents 2,389 hourly measurements at BSR of speciated NMHC that were not included in the photochemical modeling of Carter and Seinfeld [2012]. The latter paper only used data from 1 of 12 spatial surveys that are presented here.

1. (c) Field et al. provides valuable information on the source apportionment of NMHC emissions in the UGBR, but the question remains if these source categories are broadly applicable to or independent of previous studies.

<Response> 1. (c) The principal aim of the paper is to provide NMHC source apportionment information. This analysis has not been attempted in previous work that has utilized a limited subset of the BSR dataset. The results of the analysis presented here, while original and independent, are indeed applicable to other studies that consider fugitive natural gas emissions and/or fugitive condensate emissions both in the UGRB and elsewhere in the US and

beyond. This paper provides clear confirmation of the dominance of fugitive emissions, in particular "wet natural gas." In particular, the analysis related to water treatment emissions is unique and demonstrates the influence of such emissions on the production of ozone in the Boulder region of the Pinedale Anticline. Emissions from water treatment facilities have not been recognized previously as being so significant, in terms of their effect on ozone production. The results presented here are applicable to previous work and address some of the data needs noted by Carter and Seinfeld [2012]. The novel findings of the paper therefore add to the previously published work, in particular by highlighting the importance and impact of water treatment emissions, and, by inference, by providing information that enhances our understanding of air quality impacts resulting from produced water.

2. (a) 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_3 within western shale basins [Schnell 2009 and Edwards 2014].

<Response> 2.(a) Additional discussion is now given to meteorological differences between the two years of measurements at BSR in section 3.1, with additional supporting data from elsewhere (i.e. UGWOS reports, in particular MSI 2014). We add information regarding meteorology in 2012 compared to 2011. We have also added information that emission inventories show reductions from 2011 to 2012 for VOC and NO_x. Also other corroborating ambient measurement data that supports the different behavior of total NMHC in 2011 vs. 2012 is now included. Specifically, the differences of total NMHC and NO_x between 2011 and 2012 are now quantified in section 3.1. Changes for NO_x and total NMHC are better explained with reduction of peak NMHC more likely are due to changes of emissions. Year-to-year changes are now noted as probably being driven by emissions, but we note that meteorological influences cannot be entirely discounted.

2. (b) Figure 4 clearly shows that high NO_x (between 6-20 ppbv NO_x) and NMHC (>0.2 ppmC) values will be present yet O_3 may still be below 40 ppbv indicating that there is another important variable that has not been adequately addressed.

<Response> 2.(b) Figure 4 shows that elevated precursor levels are sometimes observed when ozone is not significantly elevated. This is not surprising, because as stated in the text other factors such as meteorology and actinic flux levels must also be favorable. We have now noted this important condition in the text with respect to Wyoming DEQ predictions, thereby clarifying the last paragraph of section 3.3 with respect to the coupling of high ozone with and elevated NMHC. The importance of high NMHC is now placed in a better context.

3. One interesting observation is the occasional differences in O_3 profiles and peak O_3 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_3];" however, not enough information is currently provided to show that (a) the O_3 levels should be similar but aren't or (b) the air mass trajectories were affected by different source regions of O_3 precursors, etc. For example, the observed ozone gradients between the sites could suggest inhomogeneity in O_3 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_3 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_3 levels? experience higher O_3 levels?

<Response> 3. In section 3.2 we have now included reference to the lack of difference between the two Boulder sites for data collected outside of the time period of ozone episodes during 2010/2011. We have added comments that show that under these conditions ozone levels are similar with for example hourly values with 10 ppbv, compared to the observations shown in Figure 3. Over the distance between the two sites the modeling of air flow is not practical for this paper, but we have now included reference to wind direction information. During the two days with the highest difference in ozone with high levels at BLDR we note the possible influence of factor 3 (Water treatment facility). This point is noted in the text, but the influence of factor 3

upon BLDR is inferred rather than measured. We have added notation of the geography of the two sites as BLDR is elevated at the top of a hill and BSR is in a river valley. However the difference in elevation is only 20m. The question of pooling was one of the reasons for the site location at BSR and there was no evidence for this behavior with ozone levels comparable, excluding episodes when differences were evident. We have added the elevation information of the Boulder sites.

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,b,c)

4. (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_x radicals that contribute to the photochemical formation of winter O_3 in the Uintah Basin. I would expect oxygenated VOCs to play a non-negligible role in O_3 formation in Wyoming even if they may be a small fraction of the carbon mass.

<Response> 4. (a) The median and mean difference between NMHC and VOC is not statistically significantly different by time-of-day. There are 6 to 7 high outliers in the afternoon hours, but the change in the median is definitely not statistically significant. The periods when differences are >500 ppbC do not correlate with those when ozone > 75 ppb. Thus unreported and unidentified mass is not correlated with high ozone, residuals aren't changing significantly with time-of-day, and do not explain the high ozone. Additional insight could be obtained if speciated residual mass were to be available, but the lump sum isn't explanatory/predictive. While carbonyls are not measured by the OPA system their efficiency of trapping is unknown on the 55i instrument. We add discussion related to the role of carbonyls and add reference to data collected WDEQ and reported in MSI 2014 that includes measurements of oxygenates and also

carbonyls. The point that while carbonyls have low contribution with respect to mass but a significant role in ozone production is now included.

4. (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.

<Response> 4. (b) The "missing mass" is already included in the PMF, as total NMHC is included as a bulk species. And discussion of the difference between this parameter and speciated NMHC is included. It would not be appropriate to add unidentified peaks or rather a lump sum of detected but undefined material from the chromatogram of the PE OPA system. It is plausible to include "missing mass" or an unidentified NHMC component to the PMF analysis. However, that lumped sum may be very different chemically at different times. So we have not added a summation of unidentified material that would not be possible to define. We will be better able to address the identity of unreported VOC in a future campaign with PTR-TOF-MS measurements anticipated for 2015. However we prefer to only use only validated specific parameters in the PMF analysis.

4. (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.

<Response> 4. (c) The referee comment is right that our factor 2 is a mixed source and we have re-emphasized this is the text. PMF analysis reflects the ambient data and the BSR site is downwind of a wet gas field. Consequently, in the absence of crude oil such emissions are absent. As the text states, three factors are produced from the PMF analysis, the second of which has a signature of fugitive wet gas with a small added contribution from inferred dehydration activities. The match with wet gas signatures measured by the WOGCC is striking; additional BTEX is believed to originate from dehydration units at well pads. Figure 6 shows the weighted combined signature of these two sources.

5. Could NO_x 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.

<Response> 5. When NO_x was included in the PMF analysis, it generated its own factor that was unrelated to other parameters. Since our focus was source apportionment of NMHC, it was therefore subsequently excluded. A brief discussion of and reference to this analysis has been included in the text but not added to the Figure.

Other comments (by order of occurrence):

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

<Response> Text is now corrected, and includes “(hourly O_3 ≥85 ppbv) .”

P24944 L6: What is the significance of the two sites experiencing different O_3 levels?

<Response> Text is now corrected as follows: “During 2011 wintertime O_3 episodes at two sites near Boulder Wyoming, situated ~5 km apart, were observed to sometimes differ, indicating the potential for measured O_3 at these sites to be differentially influenced by localized ozone precursor emission sources”.

P24944 L7-8: “In 2012 the lack of O_3 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_x?

<Response> The text in this section is now modified as follows: “In 2012 levels of oxides of nitrogen and meteorological parameters were similar. The four key parameters used by regulators to predict episodic ozone; snow cover, 700 mb height, temperature and wind speed, were comparable to previous years. However in 2012 the lack of O_3 episodes coincided with a reduction in ambient levels of total non-methane hydrocarbons (NMHC)”.

P24944 L18: Define “condensate range NMHC.”

<Response> This sentence has been revised as follows to avoid unnecessary complications of definitions in the abstract: “A water treatment and recycling facility was found

to be a significant source of NMHC abundant in condensate, in particular toluene and *m*+*p*-xylene".

P25945 L18: Start the discussion of winter O_3 in a new paragraph ("Until recently…" would be the start of paragraph #3 of introduction).

<Response> Corrected in the revised text.

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.

<Response> The entire first paragraph is now re-positioned as recommended.

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.

<Response> A more detailed description of the total NMHC measurement is added.

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

<Response> These sentences are 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.

<Response> This sentence is clarified and also moved to section 2.1

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

<Response> This data is not included. It will be published in a separate paper regarding diffusive sampling, that is currently being prepared. The omission is clarified.

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

<Response> The sentence is corrected and edited to reflect the 7 outliers removed due to elevated toluene or CO values.

P24951 L4: Be more specific regarding “numerous high hourly O3 values.” How many and what level constitutes “high O3?”

<Response> This section is revised to reflect exceedances of the NAAQS for ozone.

P24951 L9-10: A brief discussion of why NMHCs and O_3 are higher in 2011 compare to 2012 but NOx 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?

<Response> We have added discussion of this important point in section 3.1 and have added data reference to additional data from WDEQ.

P24951 L16: The discussion relates to “O_3 episodes” which are arbitrarily determined to be hourly averages greater than 85 ppbv, but this is a non-standard O_3 metric. How many O_3 exceedances have there been in WY (8-hr average > 75 ppbv)?

<Response> In this paper we define an ozone episode as $\geq$85 ppbv, this value is not totally arbitrary, it is selected as a value that clearly reflects episodic production of ozone for hourly values. As such it should be above 75 ppbv but should not be so high as to exclude any days that exceeded the standard value, while including some days with production values indicative of strong wintertime production, even if limited in temporal extent. This is important as we consider and contrast diurnal variations at two sites based upon hourly values to understand behavior rather than compliance with regulatory values. However we agree that reference to standard values should be included. We have edited the text in section 3.1 to refer to the standard 8-hour average and number of exceedances in specific years.

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_3 to the different measurement sites.”

<Response> The text is edited as recommended.

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.

<Response> The text is now revised to clarify the data that is used for the PMF analysis. The section is also edited for major comment no. 5.

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.

<Response> The suggested text is added and the text with L7 -15 is reduced. This text is important to set the transition between Fig. 5 and Fig. 6. In particular as factor 3 loses many species that are important constituents.

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

<Response> The explanation of presence in source profiles is now earlier in the text making the shift from Fig. 5 to Fig. 6 clearer.

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

<Response> Yes. The pre-identification sentences are deleted.

P24955 L14: "e,g,", changed to" e.g."

<Response> This typo is deleted through text edits noted above.

Figures/Tables:

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

<Response> Fig. 1 now has all additional canister sampling sites marked.

Figure 2: The left axes should be aligned vertically so that it is easier to compare the time series for O3, NOx, and NMHC for a specific year.

<Response> Fig. 2 has the axes re-aligned as requested.

Figure 5: Define CO-b and methane-b in the figure caption. It is not defined until P24955

<Response> The figure capture now has "–b signifies minus background"

L22 when discussing Figure 6. What does "normalized mass profiles between 14 NMHCs" mean? Why only the stated 14 NMHCs when 22 are presented?

<Response> This is a typo in Fig. 5 and is now corrected. The reason for 14 NMHC with respect to Fig. 6 is clarified.

Are you are normalizing all percentages to 100% in order to avoid showing circumstances when the calculated factors may be over/under 100%?

<Response> No as this is the standard PMF output display. The problem is that an average is just that. Any given sample and parameter is a linear regression fit with a slope and intercept. Residuals are evident when considering hourly data. Low concentrations may be over-predicted while high concentrations could be under-predicted or vice versa on a sample-by-sample basis. As noted in the text the fit for the analysis is excellent. Table R1 shows below shows how the data were fit. We could calculate the average species concentration under/over-prediction, display that as a percentage bar, then normalize the fractions of the contribution relative to that if necessary. That would require an extra figure. We believe that adding additional text, tables and or figures is not needed for PMF residuals.

Table R1 PMF Analysis Data Fit Parameters

Species	Intercpt	Slope	SE	r^2	KS Test Stat	KS Test P Value
CO_b	6.51	0.87	10.02	0.94	0.1	0
methane_b	9.59	0.92	129.69	0.94	0.07	0
NMHC	-3.45	1.02	23.42	0.96	0.05	0
ethane	0.2	0.98	4.82	0.97	0.02	0.29
ethene	0.13	0.75	0.32	0.8	0.03	0.07
propane	-0.56	1.08	1.63	0.98	0.03	0.01
i-butane	-0.1	1.05	0.39	0.98	0.08	0
n-butane	-0.15	1.07	0.38	0.98	0.04	0
ethyne	0.28	0.6	0.4	0.8	0.11	0
cyclopentane	0.01	0.91	0.03	0.93	0.18	0
i-pentane	0.03	0.97	0.24	0.97	0.07	0
n-pentane	0.05	0.93	0.22	0.96	0.07	0
2-methylpentane	0.09	0.73	0.17	0.89	0.11	0
3-methylpentane	0.06	0.73	0.11	0.88	0.12	0
n-hexane	0.13	0.72	0.26	0.88	0.08	0
benzene	0.05	0.91	0.26	0.94	0.04	0
cyclohexane	0.12	0.76	0.27	0.91	0.07	0
i-octane	0.01	0.9	0.07	0.89	0.11	0
n-heptane	0.1	0.78	0.22	0.94	0.06	0
toluene	-0.24	1.04	1.13	0.89	0.11	0
n-octane	0.08	0.76	0.21	0.93	0.05	0
ethylbenzene	-0.01	0.95	0.09	0.88	0.06	0
m+p-xylene	0.1	0.88	0.51	0.94	0.04	0
o-xylene	0.01	0.92	0.11	0.92	0.06	0
nonane	0.09	0.62	0.27	0.81	0.17	0

Why not show the residuals to give an idea of how well the PMF analysis represents the actual measurements?

<Response> The success of the 3-factor solution is already indicated in the text with the R^2 values, and this is reflective of relatively low residuals. Our current figure selections are not amenable to residual plots. Adding residual information in the body of the manuscript would not add to the current direction of the paper. Furthermore each compound has it's own residual for a given hour. So we would need to show all 25 plots of model and actual time series. An example is given for propane in Figure R2, these plots could be added in supplementary material but we do not feel that this is necessary

Figure R2 Fit of observed vs predicted (PMF) mixing ratios for propane