

***Interactive comment on “Understanding high wintertime ozone pollution events in an oil and natural gas producing region of the western US” by R. Ahmadov et al.***

**Reply to anonymous Referee #1**

We thank the reviewer for the valuable comments on our paper. Below we provide answers (highlighted in blue) describing how we addressed the comments.

The authors present a thorough analysis of the modeling of ozone in the Uintah Basin. The central outcome is that modeling of ozone in 2013 with the use of a top down emissions estimate (from measurements) is better able to reproduce episodic ozone than with bottom up emissions estimate (from inventory calculations). This is an important point in particular given the apparent lack of difference in 2012. The authors explore the importance of model setup besides emissions and provide useful observations. The paper could be improved by some re-working of sections as noted below:

Abstract reads like a summary and could be shorter to ensure main points are not too diluted.

We shortened the Abstract a bit. However we tried to keep the major statements in the Abstract, as this paper has several important conclusions.

Also the issue of under or over estimation is inferred from the performance of ozone modeling. Emissions are not validated directly against measurements of the given emission parameters. So the abstract has the cart before the horse. Summary is too long and seems like an extension of the discussion of results.

In Table 3a and b we present the model evaluation statistics for multiple gaseous species including ozone. As we discuss in the paper these comparisons support our judgment about the accuracy of the bottom-up and top-down emission inventories. Using atmospheric models we can't verify emission profiles for individual sources in the oil/gas sector. The profiles and speciation of the emissions for different sources can be validated by sampling individually emissions of those sources. We think the following statements in the Summary section address the raised comments: “Although using emissions based on observations offers significant improvement over the bottom-up emissions inventory, there are potentially important details missing in the top-down emission estimates. Speciation profiles for oil vs. NG wells, and other

stationary sources, are known to be different (Warneke et al., 2014). Thus, a synergy between bottom-up and top-down approaches will be needed in order to develop more accurate emission inventories for the oil/gas sector in the US.”

We shortened the Summary section substantially and made some statements more succinct and clear as the reviewer suggested.

It also explicitly states some important points for the first time (e.g. p20322 line 9 Air quality models.....; p20323 line 21 Since the analysis.....; p20324 line 12 Our results suggest.....). The latter example is a critical point that is somewhat buried by surrounding material. Perhaps a conclusions section would help bring the main points into clearer focus?

We realize that the sentence “Air quality models...” makes a statement about all the air quality models, while this study is based on the WRF-Chem model. So we removed this sentence, also to shorten the Summary section in order to address the reviewer’s suggestion.

We removed the sentence “Since the analysis...”, as indeed the box and 3D models can’t be compared in a straightforward manner even if they use the same parameterizations.

The statement “Our results suggest...” is based on our finding that the emissions of aromatic VOCs contribute to daytime ozone at the surface nearly as much as alkane species, which are the dominant fraction of the VOC emissions. These sensitivity simulations are presented in Section 4.2.3. Consequently, reducing the aromatic VOC emissions would be more efficient in ozone reduction than other species.

Finally letters labels in Figure 2 seem to be invisible.

We updated Figure 2 with letters and labels to illustrate it better.

## **Reply to anonymous Referee #2**

We are grateful to the reviewer for the thorough comments on our paper. Below are our answers (highlighted in blue) to all the questions and comments raised by the reviewer.

This paper deals with a WRF/CHEM modeling study to simulate and explain wintertime ozone

events in the Uintah Basin (UB), Utah. When I first saw the paper, I was excited and was curious to see how a 3D model would contribute to better understanding of these O<sub>3</sub> events, but eventually my disappointment grew as I came across several obscure parts and inconsistencies, which need substantial explanation (for specific details see below).

General comment: The authors make a reference to Baker et al. (2011). This paper actually mentions that the occurrence of a cold pool is a frequent observation in the western US, including urban areas such as Salt Lake City. Salt Lake City is not too far away from the UB and has quite similar climatological conditions. The authors should explain, why the occurrence of wintertime ozone events is confined to oil and gas producing regions.

This is a good question. The total NO<sub>x</sub> and non-methane VOC emissions for Salt Lake County are 78 and 64 tons per day (weekday), respectively based on the NEI-2011 emission inventory. From Table 2 these are a factor of ~7 higher for NO<sub>x</sub> and ~8 lower for VOC compared to the top-down inventory for the oil and gas producing region of the Uinta basin. As pointed out several times in the paper, the high VOC emissions in the UB are the driving factor in high O<sub>3</sub> formation, so if all Salt Lake County emissions were placed in the Uinta Basin we would expect at least a factor of 8 less O<sub>3</sub> produced based on VOC emission differences. Additionally, the factor of 55 difference in the VOC/NO<sub>x</sub> emissions ratio would force the photochemistry into a regime that is even more NO<sub>x</sub> saturated than the bottom-up results, further suppressing O<sub>3</sub> formation. Therefore, Salt Lake City (or a city/town with a similar emission profile) does not experience high wintertime ozone events when cold pool conditions are present. We added this statement in the Summary section to explain the contrast between urban and oil/gas sector VOC vs. NO<sub>x</sub> emissions: “It should be noted that total VOC emissions, especially VOC/NO<sub>x</sub> emission ratios over urban areas (e.g. Salt Lake City) are much lower than those in the UB, therefore such urban areas are not expected to experience high wintertime ozone production when cold pool type wintertime conditions are present.”

There is no doubt that Figure 3b clearly shows enhanced O<sub>3</sub> levels in 2013 compared to 2012. Also, the observations of ozone show a gradual increase from day to day in those episodes the authors define (29 Jan - 8 Feb; 12-17 Feb; 19-21 Feb). It looks to me that in the observations the ozone background in the UB continuously increases from about 40 ppbv to about 80 ppbv

in the first period (similar in the other periods). Daily increases are superimposed and are about 40 ppbv - 50 ppbv. This feature is different from O<sub>3</sub> events in the Upper Green River Basin (UGRB), where daily increases of about 100 ppb can be observed, while the background values remain about the same level (Carter and Seinfeld, 2012; Rappengluck et al., 2014; Oltmans et al., 2014). The authors show that using the top-down emission scenario the model is able to capture the peak O<sub>3</sub> values. However the model fails to replicate the nighttime ozone levels. The authors mention that this is caused by stronger katabatic winds at night in the model, which advect cleaner background air from outside the UB. This raises a few questions:

In response to the reviewer's concerns about clarity, and with regard to the strong katabatic influence in particular, the revised version of the paper contains updated Figures 5a, b, which are more illustrative of the meteorological setting and impact on the residual layer O<sub>3</sub> than before.

#### Chemistry:

1) It looks like the model simulates larger daytime O<sub>3</sub> increases than are observed. For instance, on 6 Feb it seems there is an O<sub>3</sub> increase of 70 ppbv in the model, while the observations show an increase of about 30 ppbv over the background values. A similar feature can be seen in Figure 6. Does the model overestimate chemical processes vs meteorological processes?

For daytime conditions, the answer to this question depends on how well the model reproduces the impact of recirculated O<sub>3</sub> versus recently produced O<sub>3</sub>. For nighttime conditions the answer is certainly no. The model under-predictions start near sunset and are most severe just before sunrise. O<sub>3</sub> (and O<sub>x</sub>) photochemical loss is essentially zero after sunset when radicals responsible for both O<sub>3</sub> loss and production are absent. The text has been modified to better explain the interpretation of the model results with respect to the low O<sub>3</sub> values predicted by the model in the early morning hours. We address the reviewer's comment in the text by pointing out that the strong diurnal cycle is dynamically as well as photochemically induced.

In the main text we have these updated paragraphs about the nighttime low bias for O<sub>3</sub>:

“One feature of the base case model O<sub>3</sub> time series, not as pronounced as in the observations, are the sharp decreases in late afternoon, with minimums typically occurring between the hours of

03:00–06:00 MST. One could easily misinterpret the  $O_3$  increase from early morning to late afternoon as being due exclusively to model photochemistry. Model  $O_3$  typically peaks between 15:00–17:00 MST, and sharply decreases after sunset. From sunset to sunrise both photochemical loss and production are essentially zero and aside from minor titration,  $O_3$  is controlled by dynamics and surface deposition. Simulated time series for primary emitted species ( $NO_y$ ,  $CH_4$ , and toluene) are presented in the SI that also show sharp decreases between sunset and the 03:00–06:00 MST time period. These strong decreases in model  $O_3$  and other species during nighttime at Horse Pool are caused by katabatic flows within shallow layers near the model surface that advect polluted air down the Green River drainage and replace it with cleaner background air from outside the UB. As discussed in more detail further in the text, with reference to the transport and photochemistry occurring on 5 February 2013, the 3-D model predicts a residual  $O_3$  layer from the previous days' photochemistry, effectively trapped within the UB, which also impacts daytime levels.”

“Figure 5a and b illustrates the simulated  $O_3$  distribution within the UB during a high ozone event. Modeled  $O_3$  mixing ratios and horizontal wind vectors over the vertical plane across the Ouray and Horse Pool sites (Fig. 1) are shown at night (05:00 MST) and in the afternoon (15:00 MST) on 5 February 2013, a typical stagnation day when  $O_3$  levels reached their February maximum at Horse Pool (Fig. 3b). The plots show the very low wind speed inside the UB, demonstrating the wind regime during a VCP type situation. Figure 5a shows the high  $O_3$  from the previous days' photochemistry in early morning above the western slope within a residual layer. The highest  $O_3$  mixing ratio is simulated at ~ 100 m above ground at night. The downslope terrain flows converge at the bottom of the UB at night. Later the layer enriched with  $O_3$  and other pollutants aloft, is moved to the eastern part of the UB during the morning hours. As we show in the SI, the wind speed of the nocturnal drainage flows over the UB are overestimated at Horse Pool and possibly elsewhere. This overestimation likely causes a stronger and faster drainage of the air masses near the ground at night, than what was observed. These strong nighttime drainage flows also displace primary emitted species with background concentrations at night (shown in Fig. S8).”

We also direct the reviewer to the new time series plots of  $NO_y$ ,  $CH_4$  and toluene in the supplement (also suggested by the reviewer). We added a new paragraph (Time series of observed and model  $NO_y$ ,  $CH_4$  and toluene for 2013) in SI to describe these time series. Also, this

section in SI discusses the nighttime low bias in primary species such as  $\text{NO}_y$  as follows:

“As best shown in Figure S8a, all primary emitted species in the bottom-up inventory show a regular peak occurring between 16:00 and 18:00 MST, and a minimum typically occurring between 03:00 am and 06:00 MST. The sharp decrease in primary species for the bottom-up emission scenario mimics the diurnal pattern of  $\text{O}_3$  predicted in the top-down emission scenario (main text, Figure 3b). In both cases the model effectively transports Basin wide accumulations down the Green River valley floor, with the compensating katabatic circulation bringing in diluted and cleaner air to the Horse Pool site from up-wind. For the primary emitted species in the top-down emission scenario (constant diurnal profile) the effect of nighttime dilution on the diurnal profile is significantly reduced but still apparent.”

In order to prove that the multi-day buildup of ozone and other ingredients of the photochemistry are necessary in the model, here we included a time series plot illustrating modeled and observed ozone time series at Horse Pool for 5 February 2013, when the highest ozone was observed at the site during the stagnation episode. Here we performed a test simulation where all chemical species were reinitialized using a simulation output for a previous day, but the oil/gas emissions set to zero. The time series from this test simulation in Figure R1 (added at the end of this text) shows that in this case the model ozone at Horse Pool reaches ~90 ppb maximum during daytime. As Figure 3b in the paper shows, without the oil/gas emissions the model does not predict any significant ozone enhancement over the background levels. Therefore the high ozone values (~90 ppb) are simulated solely due to the recirculated air masses containing previous days' pollution. The difference between our test simulation and base case during daytime is ~20 ppb, which is an indication of additional ozone produced during February 5<sup>th</sup> due to the oil/gas emissions. We'd like to stress that due to non-linear chemistry it's hard to accurately quantify the  $\text{O}_3$  production in a given day compared to the carry-over from the previous day in the model. The test simulation started at 00:00 UTC using the same initial conditions as the base case on Figure R1, but we plotted the time series from 00:00 MST, therefore there is a slight difference between the two model  $\text{O}_3$  time series at 00:00 MST. First 6 hours of the simulations are assumed a spin-up time for the meteorological fields.

2) It is unclear whether the authors used a chemical mechanism, which is adjusted to colder temperatures similar to the approach by Carter and Seinfeld (2012). In case they did not adopt this approach, the authors should explain the reasoning for their different approach.

First, we also would like to bring to the reviewer's attention a fact noted by Carter and Seinfeld (2012): "At present there is no information available to estimate how aromatic mechanisms may change at lower temperature, since uncertain portions of the mechanism have to be adjusted to fit chamber data (Carter, 2010), and there are insufficient data available to derive mechanisms for lower temperatures. Therefore, the individual or lumped mechanisms for the aromatics were not modified except for a few sensitivity calculations that are discussed below."

The underlying gas chemistry mechanism in our WRF-Chem model configuration is primarily based on the RACM scheme as described by Stockwell et al., (1997). Stockwell et al., (1997) state that "The RACM mechanism was created to be capable of simulating the troposphere from the Earth's surface through the upper troposphere and to be valid for simulating remote to polluted urban conditions". Hence, our gas chemistry mechanism should be suitable for wide ranging temperature conditions throughout the atmosphere. We aren't aware of any version of the RACM scheme that was developed for particularly wintertime conditions.

The Carter and Seinfeld (2012) study is based on the SAPRC gas chemistry scheme and a box model, where individual VOC precursors were treated explicitly. In a fully lumped mechanism, such as RACM, it's impossible to adapt the scheme to different temperature regimes. Such adjustment needs to be conducted at the stage of building a lumped mechanism, which is suitable for 3D air quality models. Moreover, even if we had a different gas chemistry scheme for a different temperature regime, it would have been infeasible to utilize it in a 3D model framework. Very diverse meteorological and chemistry conditions varying in space and time need to be represented within a unified modeling framework by 3D air quality modeling tools. Our WRF-Chem model results for 2012 and 2013 with sharply different temperature conditions show that the RACM gas chemistry mechanism is quite flexible and suitable in capturing the different temperature conditions.

Nevertheless, we recognize that improvements in gas chemistry mechanisms used by 3-D models may be needed to better characterize photochemistry pertinent to the oil/gas emissions

during wintertime, which is quite different than the summertime photochemical smog problem. In fact, we discuss this point in the Summary section. We removed the sentence in the Summary section (p. 20325) about adapting gas chemistry mechanism to cold conditions, as this might imply that we suggest using different (adapted) gas mechanisms for different meteorological conditions in a 3D modeling framework. Instead, we added this statement: “Further refinements in gas-phase chemical mechanisms may be needed to better characterize the conditions pertinent to oil/gas emissions in wintertime beyond the chemistry associated with summertime urban photochemistry.”

3) On page 20306 the authors state that the top-down oil/gas emission estimates are based on regression slopes of individual compounds with CH<sub>4</sub> observed between 10-16 MST. How did the authors account for the different photochemical decay of these compounds in the calculation of the regression slopes?

We added this text to the end of the first paragraph in Section 2.2: “As explained in that study, VOC emission ratios determined by this regression technique do not account for removal by the OH radical. Thus, emission estimates, particularly for more reactive VOCs, should be considered lower limits.”

4) Tables 3a and b show almost the same NO<sub>x</sub> levels in 2012 and 2013, while hydro- carbon levels in 2013 are significantly enhanced. Why is this case?

Due to much stronger photochemistry a lot of NO<sub>x</sub> is converted to other nitrogen species in 2013. NO<sub>z</sub>(=NO<sub>y</sub>-NO<sub>x</sub>) values increase by a factor of 7.3 in 2013 compared to 2012. We have modified the related text in the paper as follows:

“The statistical comparisons in Table 3a and b for reactive nitrogen species show a marked difference between 2012 and 2013 for observed levels of NO<sub>x</sub> and NO<sub>y</sub>. Median NO<sub>y</sub> levels are more than a factor of 3.5 higher in 2013 compared to 2012, yet median NO<sub>x</sub> levels are nearly identical. The differences in NO<sub>x</sub>/NO<sub>y</sub> partitioning between the two years is qualitatively reproduced in the top-down inventory results, but difference for individual NO<sub>y</sub> species are apparent.”



5) While tables 3a and b provide a general overview of the model performance for a range of compounds for daytime hours of the time frames shown in Figures 3a and b, it would be good to see specific time series of some selected other compounds than O<sub>3</sub>, e.g. NO<sub>x</sub>, methane, toluene, HCHO, HNO<sub>3</sub>. In particular toluene would be nice, as the different emission inventories are shown in Figure 2.

Per reviewer's suggestion we added NO<sub>y</sub>, methane and toluene time series at the Horse Pool comparisons for 2013 in Supplemental Information (SI).

#### Meteorology:

1) Recirculation processes might be of importance as shown by Rappengluck et al for the UGRB. At some point (page 20315) the authors mention recirculation processes in the west-east direction. However, I have some problems finding this feature in the data shown. Figure S2d shows little diurnal variation in the wind direction (mainly ranging from SE to SW) and would actually indicate some persistent transport in the south- north direction. Unfortunately, Figure 5 only shows horizontal wind vectors and no upslope/downslope vectors, which could otherwise help to understand, whether there would be circulation processes in the vertical in the basin.

The reviewer's point is well taken. We updated Figure 5a and b to show vertical wind component as well. In addition, the new figures better illustrate the wind regime and O<sub>3</sub> distribution in the UB than the previous ones.

The cold pool meteorology is quite complex and its modeling is a very challenging task, because its evolution can be dependent on both the microscale and synoptic scale flows. This paper is focused on the emissions and atmospheric chemistry modeling for the UB in winter when meteorological conditions are reasonably simulated. A separate publication is under preparation by Banta et al. to discuss in much more detail the mesoscale flows and vertical mixing in the UB during the cold pool conditions.

We agree that Figure S2d (previous version) might be misleading, as wind direction averaging convolutes the wind regime. Therefore we removed that figure, and instead we included Figure S4 in SI to show wind roses from both observations and the model for afternoon and nighttime

hours. These figures provide insight into the wind flow near the surface (9 m above ground) and changes during the transition from night to day at the Horse Pool site. The observed wind flow was from the NE at Horse Pool at night, however in the afternoon the wind was from different directions with the SW direction dominating. The model winds at night are also predominantly easterly at Horse Pool. In the afternoons the model wind vectors are from different quadrants, but dominated by southeasterlies. It should be noted that as Figure S3c and Figure S4 show, the measured wind speeds at Horse Pool during the stagnation episode were quite low, in most cases around 1 m/s, and the accuracy of wind direction measurements decreases in such weak wind situations. Also, the wind roses in Figure S4 show wind near the surface not aloft. The updated version of Figure 5a and b clearly demonstrates model recirculation of air within the UB, as they show the wind vectors aloft.

2) In table S2 the authors mention that 18 layers within the lowest 500 m were used in WRF-CHEM. A recent WRF/CHEM study employing various PBL schemes (Cuchiara et al., 2014) showed ambiguous results for the different PBL schemes. How do the modeled vertical profiles of meteorological parameters agree with observations in the UB?

We provided Figure S5 to show boundary layer height comparisons, which is the most critical meteorological variable related to cold pool and mixing depth. To determine the PBL height (PBLH), we used vertical profiles of virtual potential temperature, which is determined by using both potential temperature and moisture mixing ratios. As noted above our paper is focused primarily on explaining the emissions and photochemistry related to high O<sub>3</sub> events, and not the exact details of the meteorology within the UB. We think adding comparisons for vertical profiles of individual temperature and moisture measurements does not necessarily add new information to our paper. Other paper(s) will present those measurements in more detail. Moreover, the tether sonde measurements that we used didn't include an accurate way of measuring wind.

3) The PBL height is a very critical quantity. Figure S5 shows a very poor relationship between observed and modeled daytime PBL height, even when averaged over 3 hours. Different PBL schemes may yield different results (see Cuchiara et al., 2014); this is most

likely more challenging in snow-covered areas. What was the reasoning to pick MYNN level 2.5? Have you employed other PBL schemes and what were the results? It would be helpful to see a time series of modeled vs observed PBL heights to see whether the model performance is different on individual days and/or hours of the day.

We think the model doesn't exhibit "very poor relationship" in capturing the boundary layers in cold pools. First of all, it's very challenging to accurately simulate cold pool conditions, as the development involves the formation of drainage flows and the intermittent turbulence associated with them while the maintenance of cold pools requires an accurate depiction of the overlying synoptic flow interacting with the stable inversion. Second, the daytime PBLH values were 50-200 m deep, which agreed favorably with the modeled values (Figure S5). There are uncertainties in determining the PBLH values from measurements. For example, the paper by Cuchiara et al., 2014 (pointed out by the reviewer) reports that "We estimate the uncertainty of the PBL height to be on the order of  $\pm 50$  m." In our case we applied different threshold numbers for vertical gradient of virtual potential temperature to calculate the PBLH. By using 0.5 and 1.5 K as threshold numbers, we get different PBLH estimates. The standard deviation of the difference between these two estimates is 80 m, while the average PBLH for these estimates are 59 m and 117 m when 0.5 and 1.5 K threshold values were used, respectively. Thus, the uncertainty in the PBLH estimation from the tethered sonde measurements is comparable to the absolute values of the PBLH estimates. During daytime in summer, the PBLH values are typically in the range of 1000-2000 m and sometimes even deeper. In such cases the uncertainties in PBLH calculations are much smaller compared to the absolute PBLH values. In Figure S5 we used the same 1 K threshold for both the model and observation to determine the PBLH values. However, as discussed above these values are subject to uncertainties. The purpose of presenting Figure S5 is to show that we don't see systematic bias in the mixing layer depth; rather, on the average the simulated PBLH is very close to the observed values. The fact that the mean error is often less than the estimated uncertainty and that the differences were never greater than 300 m in this region of complex flow implies that the simulated meteorology is sufficiently skillful for this air quality study.

Per the reviewer's request we show here the time series of the PBLH values for all three sites (Figure R2 in the end of this text). We didn't include these plots in SI, as they don't provide much help with regard to the results discussed in this paper. The time series in R2 show that

the tethersonde measurements are confined mostly to daytime. The vertical profile measurements were taken as tethersonde soundings up and down with a short time lag. The “random” PBLH variability between the individual soundings is quite apparent from the time series plots. Unlike the observation the model doesn’t show such variability, rather smooth diurnal profiles of the PBLH evolution. Therefore, we averaged the PBLH values within 3 hourly intervals in order to display them in Figure S5.

The MYNN level 2.5 scheme is based on local mixing approach like other TKE-based schemes, but it has been tuned to a database of Large-Eddy Simulations (LES) to overcome some of the typical biases noted in many other local mixing PBL schemes (for example, Coniglio et al., 2013 WF, Zhang et al., 2011 MWR, Ngan et al., 2013, JAMC, among others). In WRF-Chem the exchange coefficients calculated by the PBL schemes are used to perform vertical mixing of chemical variables. Therefore, using the MYNN PBL scheme makes vertical mixing of meteorological and chemical species more consistent. Additionally, in our WRF-Chem simulations the MYNN scheme has traditionally demonstrated a good performance, e.g. Ahmadov et al., 2012. The operational models, Rapid Refresh (RAP) and High-Resolution Rapid Refresh (HRRR), both have switched to the MYNN from the MYJ PBL schemes due to improved boundary layer profile verification (see conference posters & presentations from the NOAA/ESRL/GSD group, e.g. Alexander et al.:

<http://www2.mmm.ucar.edu/wrf/users/workshops/WS2014/WorkshopPapers.php>)

We also recognize that the improved simulations of the PBL structures for such conditions over complex terrain may require higher spatial resolution (e.g. 500 m), better snow cover data and improvement of the PBL schemes for such resolutions and conditions. For example, as shown (as suggested by the reviewer) in the updated version of Figure 6a, the PBLH over Horse Pool at night was as high as during the daytime (~ 100 m). Since the boundary layer during daytime in the Basin was driven by very weak surface heat fluxes, it is unclear why the PBLH can remain relatively high at night. Wind shears generated by nocturnal drainage flows and wave mixing induced by the surrounding complex topography are the likely reason, but only ultra-high resolution modeling (e.g. to capture better the rough terrain in the UB) will uncover the relative importance of these processes. Therefore we don’t have a clear picture how well different mesoscale models and PBL schemes can simulate the mixing in stable cold

pool conditions over complex terrain. We hope that these research questions will be thoroughly addressed in future studies.

4) On page 20312 the authors mention that biases in the model meteorological fields could be affecting either year's statistics for observed and simulated chemical species (i.e. table 3a and b). In order to support this statement it would be helpful to include a table similar to tables 3a and b for meteorological parameters.

We included the statistics for meteorological variables during 2012 and 2013 at the surface site in the updated version of SI in Table S3. Here we chose the same time periods in 2012 and 2013 as in Tables 3a,b. Unlike Tables 3a,b here we present RMSE instead of MMO, as the latter isn't applicable to air temperature comparison.

Additional remarks:

p 20299, lines 13-14: It should be shortly described, what the differences between the VOC mixture from oil and NG production operations vs vehicle emissions are, as there are also similarities.

The sentences referring to these differences have been refined as follows:

“Gilman et al. (2013) analyzed extensive VOC measurements at the same site in Colorado during the winter of 2011, confirming an abundance of alkanes associated with the oil and NG production operations in northeastern Colorado based on differences in chemical signatures (e.g. i-pentane/n-pentane ratios and relative amounts of acetylene to propane) between urban sources and oil/gas well emissions.”

p 20299, line 25: It is not correct to mention that in both years, 2005 and 2008, very high O<sub>3</sub> values were measured. Schnell et al (2009) report O<sub>3</sub> values of above 140 ppbv for the year 2008, only. In 2005 O<sub>3</sub> values were much lower (above 75 ppbv) according to Schnell et al (2009).

We corrected it. The new sentence is: “Contrary to the traditional view, in the rural Upper Green River Basin (UGRB), Wyoming in winters of 2005 and 2008 high levels of O<sub>3</sub> were measured,

with hourly levels exceeding 140 ppb in 2008 (Schnell et al., 2009).”

p 20299, line 27-28: Rappengluck et al (2014) report O<sub>3</sub> values well above 160 ppbv, not ~160 ppbv.

We modified that sentence and also added another statement:

“Very high O<sub>3</sub> mixing ratios, with hourly values reaching as high as 166 ppb, were observed in the same region during the winter of 2011, well above estimated surface background levels of 40 ppb (Rappenglück et al., 2014).”

p 20299, line 28 and following lines: It would be good to mention the altitudes of the UB and the UGRB, as this might also have an impact on photolysis rates.

This is a good point that the photolysis rates depend on elevation as well. We added this sentence in the text: “We note the photolysis scheme in WRF-Chem takes into account surface elevation in calculating the effect of Rayleigh scattering. This could be important for high elevation sites in the western US, such as the UB and UGRB, which are located about 1400 m and 1800 m a.s.l, respectively.”

Also, it is worth to mention that the UGRB is considered a non-attainment area by the EPA with regard to the 8 hr O<sub>3</sub> standard.

We added a statement: “As of July 2012, EPA declared the UGRB as a non-attainment area for the 2008 ground level 8 h ozone standard (Rappenglück et al., 2014).”

p 20300, line 9: 160 ppbv of O<sub>3</sub> were not reported by Schnell et al (2009). However, values of more than 160 ppbv for the 2011 UGRB field study are reported by Carter and Seinfeld (2012) and Rappengluck et al (2014). The latter publication also shows O<sub>3</sub> sonde data with more than 170 ppbv in the boundary layer.

That statement has been changed: “Sunlight passing through the trapped pollutants and reflecting back off the high albedo snow drives diurnal photochemical production of O<sub>3</sub> that peaks soon after solar noon (Schnell et al., 2009; Rappenglück et al., 2014).”

p 20301, lines 11-13: It should also be mentioned that PBL heights may have been quite different in these specific studies.

The new statement is: “During January and February 2012 the UB experienced relatively warm weather (above freezing temperatures during daytime) and the ground was mostly snow free. Thus, the meteorological conditions, including wind speed and mixing layer heights, were not favorable for O<sub>3</sub> production.”

p 20301, lines 24-28: Tethersonde data for the UB are not presented in Oltmans et al. (2014).

We fixed that. In the updated text we refer to two papers by Oltmans et al. 2014a and 2014b. The 2<sup>nd</sup> paper by Oltmans et al., (2014b) is based on the aircraft measurements conducted in 2013 and it also discusses some tetheredsonde measurements.

p 20302, lines 27-28: The statement that valley cold pools are caused by an approaching upper-level trough and midlevel cooling is not correct and is not presented this way in Reeves and Stensrud (2009).

We corrected the statement: “The VCPs form as an upper-level ridge approaches the western US and in response to strong midlevel warming, and they can exist several days until being removed usually by advection or turbulent erosion processes (Reeves and Stensrud, 2009).”

p 20303, lines 2-4: The original reference should be cited, which is Zhong et al (2001) and not Baker et al (2011), who are citing Zhong et al.

We added the reference to Zhong et al. in the Introduction section.

p 20306, line 7: It is not clear what the authors mean, when they say "when the boundary layer is active".

We changed that statement: “...a broader, more regional representative sample is available during daylight hours when the boundary layer is well mixed”

p 20306, lines 27-28: The authors should justify why they assume that the VOC speciation is the same for both oil and NG wells.

Figure R3, shown below, gives an example of the toluene versus methane scatter plot, showing the high correlation between the two species ( $r^2=0.90$ ). Table S1 shows that nearly all primary VOCs have  $r^2$  correlation values as high as or higher than toluene. So from the observations, it is difficult, if not impossible, to discern a distinct relationship between emissions of VOCs vs. methane for oil versus NG wells. But as we noted in the paper (the Summary section):

“VOC speciation profiles for oil vs. NG wells, and other stationary sources, are known to be different (Warneke et al., 2014), and the treatment of  $\text{NO}_x$  emissions in the top-down approach does not consider the heterogeneity of the various source categories. Thus, a synergy between bottom-up and top-down approaches will be needed in order to develop more accurate emission inventories for the oil/gas sector in the UB.

p 20309, lines 11-13: The chemical species for which idealized vertical profiles were used should be explicitly mentioned. This can also be done in table S2. How were initial and boundary conditions for  $\text{O}_3$  represented?

To answer to the reviewer’s question about the initial and boundary conditions for  $\text{O}_3$  we show here (Figure R4 at the end of this text) the vertical profile of  $\text{O}_3$  mixing ratio that is used in the default WRF-Chem code. We didn’t change these values in our simulations. In the main text we added this text:

“Idealized vertical profiles were used as initial and boundary conditions to assign the background mixing ratios for some of the long lived chemical species (McKeen et al., 2002). These idealized profiles are based upon results from a NOAA’s Aeronomy Laboratory Regional Oxidation Model. The detailed description of the initial and boundary conditions can be found in WRF-Chem User’s Guide available from [http://ruc.noaa.gov/wrf/WG11/Users\\_guide.pdf](http://ruc.noaa.gov/wrf/WG11/Users_guide.pdf).”

The WRF-Chem user’s guide lists all the long-lived chemical species, for which idealized profiles are available.



p 20309, line 18: What albedo values were used for 2012?

In the original version of the text we had this statement: “The default values of the surface albedo appropriate for bare ground conditions in the TUV scheme lies within the range of 0.05 - 0.15 depending on the wavelength.”

In snow free conditions such as in 2012, the model uses the default albedo values.

p 20313, lines 6-7: It would be good to mention the value for the  $\text{HNO}_3$  deposition velocity explicitly here. From the statements on page 20308, lines 19-25, I had the impression that apart from  $\text{O}_3$ , also dry deposition for other species were adjusted to wintertime conditions with snow cover. What were the "other species"?

We didn't adjust the dry deposition of individual species to snowy conditions. We modified the dry deposition scheme, so the land use category in a given grid cell changes if a snow cover is present in the model. Thus, all the dry deposition velocities change somewhat in snowy conditions. The modified text:

“The dry deposition of gas species is parameterized following Erisman et al. (1994). As the default WRF-Chem code was intended to simulate summertime  $\text{O}_3$ , the deposition scheme uses “snow or ice” type of land use category solely for the areas permanently covered by ice, e.g. the Arctic. The deposition scheme was modified to treat model grid cells as “snow or ice” whenever they are covered by snow. This helps to properly simulate dry deposition fluxes of  $\text{O}_3$  and other species to the snow-covered ground for the 2013 simulations.”

Also, please see the answer to the next question.

p 20313, lines 8-9: It should be stated that  $\text{HNO}_3$  dry deposition velocity to snow is not only low at cold temperatures, but increases with increasing temperature (Johansson and Granat, 1986), and also depends on atmospheric stability (Bjorkman et al., 2013), which means it displays a diurnal variation, with a maximum during daytime, when also ambient  $\text{HNO}_3$  tends to have maximum values.

The model calculates the dry deposition velocities of all the chemical species depending on the atmospheric stability. The details can be found in Erisman et al. (1994), which is also cited in the paper. Regarding  $\text{HNO}_3$  dry deposition velocities temperature dependence, we

reformulated the sentence: “Measurements of  $\text{HNO}_3$  deposition to snow indicate a low deposition velocity at cold temperatures but increases with increasing temperature (Johansson and Granat, 1986), and even  $\text{HNO}_3$  emissions from snow due to heterogeneous conversion of other reactive nitrogen species to nitrate at the snow surface (Dibb et al., 1998).”

p 20314, lines 14-15: Only horizontal vectors are shown in Figure 5, but no downslope (or upslope) flows.

We fixed that. As discussed above the updated Figure 5a and b show the wind patterns in much more detail.

p 20314, lines 15-17: Figure S2d predominantly shows SE winds during the morning hours which cannot lead to a transport from west to east.

As we discussed above in more detail in answering to the question about meteorology and recirculation, in the revised version of SI, we included wind rose plots (Figure S4) instead of Figure S2d in the previous version. The wind roses illustrate the wind patterns in more detail for both the measurements and model during nighttime and afternoon hours.

p 20317, lines 16-18: I assume enhanced photolysis rates were also important for  $\text{NO}_x$  sensitive regimes studied by Carter and Seinfeld (2012).

We mentioned only the VOC sensitive cases, as in our modeling study we didn’t find  $\text{NO}_x$  sensitive regimes.

p 20319, line 26: I assume the authors mean "...less than or equal to 5 carbon atoms" instead of "...greater than or equal to 5 carbon atoms".

We mean greater than or equal to 5 carbon atoms, as we talk about more reactive alkanes here. The aromatic to  $\geq \text{C}_5$  alkane emission ratios are taken directly from Table S1.

p 20319, lines 27-29: It should be added that the magnitude of the disproportionate impact of

the oil/gas sector aromatics also depends on the amount of HONO found in ambient air according to Carter and Seinfeld (2012) and thus may be variable.

In the Introduction we already have this sentence related to the study by Carter and Seinfeld (2012): “The study found that one site for the 2008 episode was highly NO<sub>x</sub> sensitive and insensitive to VOCs and nitrous acid (HNO<sub>2</sub>), while the other 2008 case and both 2011 cases were highly sensitive to changes in specified VOC and HNO<sub>2</sub> concentrations.” So stating the sensitivity of the results by Carter and Seinfeld, (2012) again in the paper seems to be unnecessary.

p 20324, line 19: I would suggest to add "...in snowy conditions in the UB" Tables:

We aren't sure which tables are meant by the reviewer, but in the text we clearly state that UBWOS-2013 case occurred mostly in snowy conditions. In addition we'd like to refer the reviewer to paper by Edwards et al., 2014, which discusses the HONO measurements during the different UBWOS field campaigns and associated issues with the measurements.

Table 1: Abbreviations used in table S2 should be included (MYNN, NAM, RRTMG).

In the updated version we spell out all these abbreviations in Table S2.

Table 2a: Metric units should be used.

We changed the units in Table 2.

Table 3a/b: The root mean square error should be presented.

Here we show 3 statistical parameters for each species such as correlation, median bias and median of model over observation. We aren't sure what value would add showing RMSE values as well.

Table S1: The table caption should mention that regressions with CH<sub>4</sub> were performed. NO<sub>y</sub> is not a VOC. I would suggest to replace VOC in the header row by a more appropriate term.

We fixed it.

Table S2: I would suggest to include the following information: initial and boundary chemistry conditions, information about advection schemes, photolysis and dry deposition treatment. Would be good to include a map which shows the domains.

We provided all the references to all the physics and chemistry options used in our WRF-Chem model configuration. Please also see the answer to the question about the boundary conditions above.

As suggested we added the domain maps in the SI (Figures S1a,b).

We'd like to stress that in our study the boundary conditions don't make any noticeable contribution on the simulated ozone enhancements in the UB. The high values of chemical species are dominated by the oil/gas and other anthropogenic emissions within the model domains. As Figure S1a shows the boundaries of the outer 12 km resolution domain is quite far from the UB.

With respect to the advection scheme and other options used in the WRF-Chem model, Table S2 provides all the information, and we include two references (WRF and WRF-Chem web-sites), where User's Guides provide detailed information about all the physics and chemistry options in the WRF and WRF-Chem models.

In the paper we provide references to the dry deposition (Erisman et al., 1994) and photolysis schemes (Madronich, 1987) used in our simulations.

Figures:

Figure 2: Letters for the measurement sites are extremely tiny.

We have fixed this in the revised version.

Figure 5: The authors state in table S2 that 18 layers within the lowest 500 m were used. However, wind vectors are shown for less layers.

In the previous version of Figures 5a,b we plotted wind vectors for every 2<sup>nd</sup> layer so the figure doesn't become too obscure with too dense wind vectors. In the updated version of Figures 5a,b the visualization tool NCL ([www.ncl.ucar.edu](http://www.ncl.ucar.edu)) chooses automatically the density of the wind vectors for visualization. We think the density is sufficient to highlight major features of the wind regime.

Figure 6: Information about the measured and modeled PBL height should be included. Also, wind vectors should be included.

The tethersondes during the UBWOS-2013 didn't measure wind speed and direction accurately, so wind vectors aren't shown on these plots. We added PBLH values on Figures 6a,b as suggested.

#### References:

The Edwards et al (2014) paper has been submitted. It is up to the editor, but personally I feel it is not helpful, since the reader has no access to this literature:

[This paper has been published in Nature. We have updated the reference.](#)

#### Literature:

Bjorkman et al. (2013): Nitrate dry deposition in Svalbord, *Tellus B*, 65, 1-18.

Cuchiara et al. (2014): Intercomparison of planetary boundary layer parameterization and its impacts on surface ozone formation in the WRF/Chem model for a case study in Houston/Texas, *Atmos. Env.*, 96, 175-185

Zhong et al. (2001): Meteorological processes affecting the evolution of a wintertime cold air pool in the Columbia Basin, *Mon. Weather Rev.*, 129, 2600-2613

[The following references are used only in the responses to the reviewer's comments. All other references mentioned above are already included in the paper.](#)

[Coniglio, M.C., et al.: Verification of Convection-Allowing WRF Model Forecasts of the Planetary Boundary Layer Using Sounding Observations, \*Weather Forecast.\*, 28, 842-862, 10.1175/waf-d-12-00103.1, 2013.](#)

[Zhang C. et al.: Improved Representation of Boundary Layer Clouds over the Southeast Pacific in ARW-WRF Using a Modified Tiedtke Cumulus Parameterization Scheme. \*Mon. Wea. Rev.\*, 139, 3489–3513, 2011.](#)

Ngan F. et al.: A Study of Nocturnal Surface Wind Speed Overprediction by the WRF-ARW Model in Southeastern Texas. *J. Appl. Meteor. Climatol.*, 52, 2638–2653, 2013.

Figure R1. Time series of the observed and modeled O<sub>3</sub> mixing ratios at the Horse Pool site.

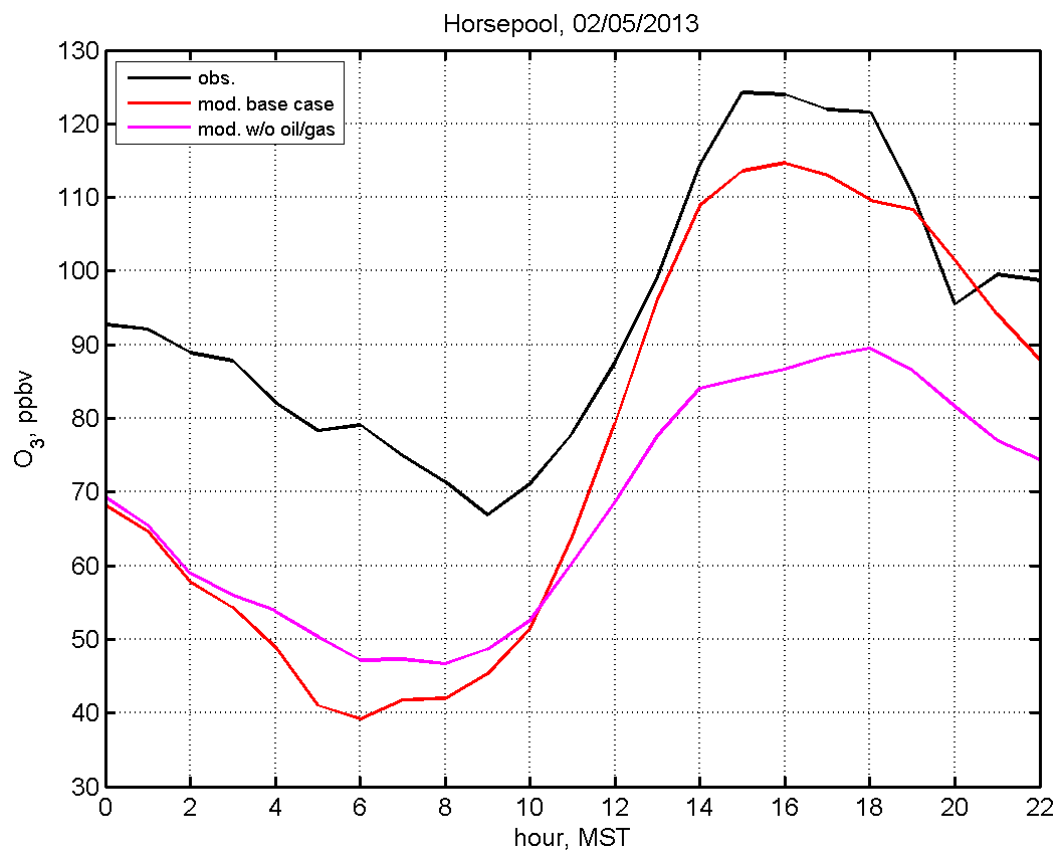


Figure R2. Time series of PBL heights determined from the tethersone measurements and the model output for three surface sites during UBWOS-2013.

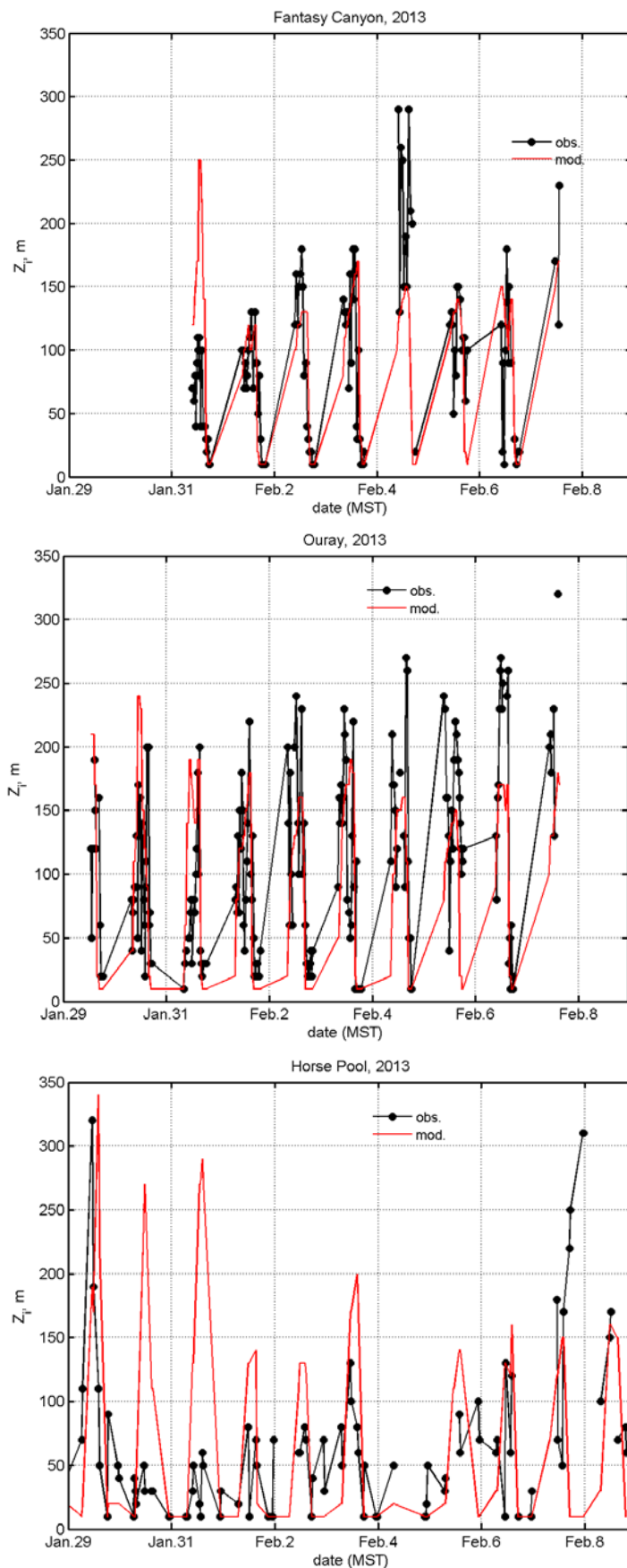




Figure R3. Toluene versus methane mixing ratios measurements at Horse Pool, during the UBWOS-2012 and -2013 fields campaigns.

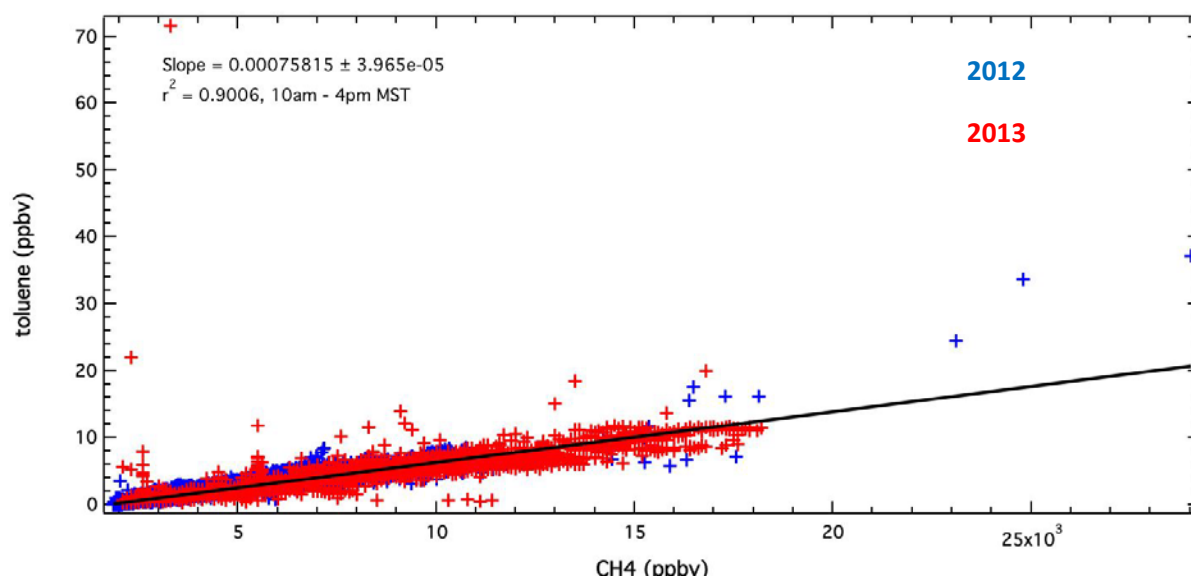


Figure R4. Vertical profile of O<sub>3</sub> mixing ratio used as initial and boundary conditions in the default WRF-Chem model.

