Please find our comments (non-italicized) to the following author comments (italics) addressed below:

Overall this paper is interesting but not entirely convincing. However, it is the authors staking their reputation on this and I will not stand in their way. However, I hope that the many co-authors are sufficiently convinced that they will not also agree to be coauthors any paper using their data that comes to a different conclusion about high NOx photochemistry. There are three major issues the authors should address:

We thank the reviewer for his or her time spent on providing a review for this manuscript. We agree that the results presented here are not entirely convincing and have conducted further laboratory tests, improved our data analysis, done more model runs, and re-written the manuscript to focus on the evidence for ozone production rates and its implications.

We will ignore the reviewer's second sentence – perhaps the reviewer should be gently reprimanded by the editor for it? The third sentence is interesting in that it implies that doing science is staking a position and sticking to it, even if previous evidence is found to be wrong or is overwhelmed by the weight of new evidence. In our view, science advances as evidence accumulates and scientists should change their thinking in light of new evidence, even if it contradicts their previous conclusions. So we respectfully disagree with the reviewer on these two sentences.

On the other hand, the reviewer makes many good points in the rest of the review and we will do our best to answer them.

1) They dismiss large negative values but not large positive values as unphysical. It is not clear that both types of errors shouldn't be treated as representative of the random error of the method used.

As pointed out in Baier et al. (2015), this technique seems to produce a large negative differential ozone signal that is roughly correlated with temperature, relative humidity, or actinic flux. This negative ozone differential is not related to negative chemical ozone production because it occurs even when the reference chamber film has been removed from the reference chamber (so that ozone production is the same in both chambers) and on low ozone production days with low levels of actinic flux and precursor gases. It should be noted that the raw $P(O_3)$ data show positive deviations from this negative pattern between 0800-1200 LT on most days. Thus, this large negative is related to something occurring in the chambers or in the Thermo ozone monitor.

After we received the reviews, we investigated in the laboratory that sensitivity of the Thermo 49C ozone monitor to relative humidity and to temperature. We did not find a strong correlation between differences in temperature between the sample and the reference chamber, but we did find a rather large shift in the Thermo instrument zero as a function of relative humidity equivalent to roughly 70% ambient relative humidity. The shift is almost binary, with an offset that is steady up to within a few percent of a certain relative humidity and then a sudden shift of ~2 ppbv within a few percent around that value. Thus, we restrict our analysis to days with lower relative humidity, when we are confident that a relative humidity initiated ozone offset change did not happen.

Four zeros were applied to the raw $P(O_3)$ data during the campaign. Two were by removing the ultem (reference chamber) film for an entire 24-hour period and two were using known low ozone production days as zeros. To avoid the problems with the sudden shift in the instrument zero, we have restricted our analysis to days when the ambient relative humidity was below 70%. In addition, we use zeros only from days that have similar relative humidity values and variations as the days for which the ozone production rate is being measured.

We include a figure of the resulting $P(O_3)$ offset that we subtracted from the observed $P(O_3)$ measurement in the Supplemental Information Figure S1 in order to get the corrected $P(O_3)$ that is shown in subsequent figures within the manuscript.

In order to ensure that the reported $P(O_3)$ values were not sensitive to our choices, we varied the threshold relative humidity and the choice of zero days that we used in the analysis over wide ranges. While the peak ozone production rate varied from 8 to 15 ppbv/hr and the minimum early evening values varied from 2 to -4 ppbv/hr, the overall shape of the measured ozone production rate was the same (i.e. MOPS corrected $P(O_3)$ data increase around 0800 LT, peak around 9-11 LT, and then decrease later in the day. We added a figure in the supplemental material to show this (Figure S2).

We used a Thermo 49C instrument to make these measurements, and at the level of the expected ozone difference from chemical ozone production (i.e. an ozone differential of 0.3 ppbv is equivalent to 8 ppbv/hr), the data are noisy and need to be integrated over time to reduce the noise. In Figure 1 that shows the measured values as a function of day of the year, we chose a one-hour average as opposed to the 30-minute averages chosen in the original manuscript submission. It is clear that the negatives are no longer as prominent – mainly because of the improved zero correction and partly because of the increased integration time, but the positive feature at ~1000 LT remains quite robust.

Additional information has been added on Pg. 6 lines 4-22 outlining these comments above, as well as in Section 1 in the Supplemental Information. We have also added the following lines to section 3.2 Pg. 11 lines 6-11:

"Third, when different relative humidity thresholds are used to correct the raw $P(O_3)$ data, measured $P(O_3)$ consistently exhibits the same diurnal behavior with a positive deviation from modeled $P(O_3)$ around 1000 LT. Fourth, deviations from the O_3 differential baseline derived from zeroing methods are observed between 0900-1100 LT even before correcting the MOPS measurements. Thus, we have confidence in the positive MOPS $P(O_3)$ signatures, which are greater than the modeled $P(O_3)$ during the morning hours."

2) Much is made of the linear NOx dependence of the model observation mismatch. A figure showing this should be presented.

We should correct ourselves in that the model-data mismatch increases monotonically for increasing NO above 1 ppbv. A figure showing the NO dependence of the model observation mismatch now been added to the manuscript (now Figure 4) along with the following text on page 10 line 31 to page 11 lines 1-2:

"The missing modeled $P(O_3)$ appears to be monotonically increasing with NO for NO values greater than roughly 1 ppbv (Fig. 4). The difference between measured and modeled $P(O_3)$ is near zero up to 1 ppbv NO and almost 15 ppbv/h at 5 ppbv NO. This unexpected increase in $P(O_3)$ with increasing NO provides a clue as to what might be causing the difference between measured and modeled $P(O_3)$."

3) The authors dismiss HONO and VOCR as a potential source of discrepancies too easily. It would be better to instead assume the entire discrepancy is due to HONO or to VOCR that is correlated with NO and show the reader the relationships between NO, HONO and VOCR that would be required to explain the results. Then the reader can judge whether those ideas are plausible.

This is a good suggestion. We have added to Sections 3.3.3 and 3.3.4, which use model calculations to assess whether HONO or RO_2 sources could potentially explain the model-data mismatch during the morning hours. We have further described the effect that these two additional $P(O_3)$ sources have on

the HO_2/OH ratio as a function of NO, but have not been able to reproduce the measured HO_2/OH NO_x dependence. (Please see the following comments below that discuss the plausibility of HONO as a morning $P(O_3)$ source.)

An additional HO_2 source via the generic reaction RO_2 + NO -> HO_2 + NO_2 + R'O was added to the MCMv331 using CH_3O_2 as the organic peroxy radical. However, the addition of this reaction elevates the entire modeled $P(O_3)$ curve throughout the day and does not alter morning $P(O_3)$ patterns relative to the MOPS. Additional RO_2 added to the MCMv331 proportional to approximately 0.05% NO_x was added to the MCMv331, which more closely replicates the MOPS morning $P(O_3)$, but underestimates the MOPS afternoon $P(O_3)$. Please see Figure 6, which was taken from the supplemental material and added to the main text. We have added text beginning on Pg. 13 line 28 to Pg. 14 lines 1-15 to discuss the plausibility of missing model RO_2/VOCR.

Other comments

Page 2 Line 27: Only the formation of peroxyacyl nitrates, and not alkyl nitrates, is a form of Ox loss. The reaction RO2 + NO -> RONO2 does not produce or consumer either O3 or NO2.

Fixed in Eq. 3. Pg. 2 lines 26 – Pg. 3 line 1 now read:

"Equation 3 describes the chemical O_3 (or NO_2) destruction rate or rate of removal to reservoir species as the fraction of O1D molecules resulting from O_3 photolysis that react with H_2O to form OH; reactions of O_3 with HO_x (HO_2 + OH); the net production of peroxyacyl nitrates (PANs); the reaction of OH and NO_2 to form nitric acid (HNO_3) and O_3 loss through reactions with alkenes and halogens."

Page 2 Line 27: Equation 3 includes the production of nitric acid via OH + NO2 - this should also be included in the description of ozone chemical loss processes.

Fixed in Eq. 3 and on Pg. 2 lines 26- Pg. 3 line 1.

Page 4 Line 1- 3: It would be useful to understand which of these references are from ambient field studies and which (if any) are from chamber studies. Some of the proposed explanations of the MOPS v. Modeled PO3 in this paper should exist in both field and chamber studies (e.g. OH + NO2 + O2) while others (e.g. ClNO2) are unlikely to be a factor in chamber studies.

All of the studies listed in the last paragraph on Pg. 4 lines 5-7 are field studies. To our knowledge, no chamber studies have been published that explain the measured vs. modeled $P(O_3)$ discrepancy or the HO_2/OH vs. NO discrepancy also mentioned. Thus, to our knowledge, there have not been any chamber studies conducted that could propose an explanation to the results shown in this manuscript. The proposed reaction (OH + NO_2 + O_2) has been removed as a possible solution to the observed model-data mismatch.

Page 4 Line 13 - 20: It is unclear what this paragraph is actually describing. Are these results all from a single study, or are these similar results from a variety of MOPS deployments?

We clarify this sentence to state which deployments (and when) were MOPS deployments, and to note that all of the other studies referenced were field studies. Pg. 4 lines 28-30 now read:

"Due to the aforementioned discrepancies between measured and modeled radicals, $P(O_3)$ calculated from measured peroxy radicals can routinely be more than double the $P(O_3)$ calculated from modeled HO_2 or RO_2 at high NO_x according to several field studies conducted during the past decade… "

Page 4 Lines 33-37 now read:

"For example, in 2010 the first version of the MOPS (Cazorla et al., 2012, Ren et al., 2013 compared directly-measured O_3 production rates to both modeled $P(O_3)$ and that calculated from measured peroxy radicals. Observed $P(O_3)$ was approximately equal to that modeled for NO levels up to 1 ppbv, but was significantly larger for higher values of NO."

Page 8, Line 22: Were time periods selected for removal based on the humidity or of the MOPS measurement itself? It is unclear from the description in this paragraph and should be clarified.

Please read the description of the analysis method given earlier in our responses.

Page 9, Line 17-19: While Figure 1 shows several days where the MOPS peak PO3 value is between 10-20 ppb/hr greater than the modeled value, it also shows several days where the MOPS minimum PO3 value is between -10 and -20 ppb/hr (e.g. Jul 21, 23, 24, Aug 6). Since these large negative values are explained as non-physical, why is it not equally plausible that the large positive values are a measurement artifact as well?

Please read the description of the analysis method given earlier in our responses. Many of these large negative deviations have been reconciled with a) averaging to 1-hour time periods and b) through revising the correction of the raw $P(O_3)$ data to account for only days when the relative humidity was less than 70%.

Page 11 Line 12 & Page 12 Line 5: The claim that the difference between MOPS and Modeled PO3 is linear with NO is difficult to evaluate from the figures provided. Since the NO dependence of the MOPS-Model disagreement is an important part of the analysis, the authors should include a figure that directly shows the difference in PO3 during the morning as a function of NO.

Please see our comment addressing this above. Figure 4 has been added in the text.

Page 12, Line 8-9: This sentence appears to contradict section 2.2, where median diurnal values of VOCs are said to be used in the model calculation. The authors should clarify how the model includes the dependence of VOC reactivity on NO.

We should clarify that the model includes the average VOC reactivity dependence on NO. VOC reactivity is linked to the median diurnal values of the VOCs used to constrain the model. Thus, the whole-air sampled VOCs are the available sources providing an average dependence of VOC reactivity on NO in the chemical mechanisms used in this study. We have stated this on Pg. 13 lines 24-26:

"Although measurements of VOC reactivity were not available during the field campaign time period and thus are unavailable for comparison to modeled VOC reactivity, the suite of VOCs measured in Golden should sufficiently capture the average VOC reactivity dependence on NO in the mechanisms used here."

Page 12, Line 15-17: As written, the proposed OH + NO + O2 -> HO2 + NO2 reaction seems to be an OH loss as well as an HO2 source. The authors should clarify why the argument here against a missing VOC source do not also apply to this proposed reaction.

This proposed reaction has been removed from the analysis, as preliminary laboratory testing and theoretical work has shown that this reaction is highly unlikely. Please see our comment above about a missing VOC source.

Page 12, Line 27-Page 13, Line 5: I am uncertain whether the mechanisms listed in this paragraph are candidates for the OH + NO + O2 reaction mechanism or not. Rewording this paragraph to clarify what some possible forms of this reaction are, and whether it might involve additional species such as acetylene would be useful.

This paragraph has been omitted.

Section 3.3.6: Given the known difficulties models have in capturing HONO concentrations, this section feels overly brief in dismissing HONO as a HOx source. I would appreciate at least brief consideration of what magnitude HONO source would be necessary to make the modeled PO3 values match the MOPS values.

Section 3.3.4 has been thoroughly modified on Pg. 14 starting on line 16. We have conducted additional model tests in order to describe the magnitude of a HONO source that would allow the modeled $P(O_3)$ and MOPS $P(O_3)$ to match. We have also described model sources of ambient HONO here. An additional HONO source is one probable missing element from our model analyses. As HONO production should be linearly dependent upon NO_x levels, additional HONO is added to MCMv331 as 10% of NO_x values, yielding peak HONO levels at 1000 LT of 0.9 ppbv. While the modeled $P(O_3)$ with added HONO best replicates the MOPS diel $P(O_3)$ pattern shown in Figure 6 out of all of the model case studies conducted for this analysis, HONO levels needed to do so are almost twice levels observed at a nearby observation site in Colorado as well as in other, more polluted areas. Modeled peak $P(O_3)$ is shifted to approximately an hour later than the MOPS and is slightly overestimated in the afternoon. We have added the following text on Pg. 15 lines 3-10 to indicate this:

"A HONO source proportional to NO_x was added to the MCMv331, resulting in average HONO levels of 0.5-0.9 ppbv between 0700 and 1200 LT, with peak HONO levels of 0.9 ppbv at 1000 LT when MOPS $P(O_3)$ exhibits its diel peak. This case study approximately replicates the observed morning $P(O_3)$ and observed OH within uncertainty levels (Fig. 6). However, while added HONO in the MCMv331 improves the agreement between observed and modeled diel $P(O_3)$,mid-morning HONO levels needed to do so are over a factor of two higher than those observed in other areas within Colorado (Brown et al., 2013 Vandenboer et al., 2013) and in environments with much higher NO_x levels (Vandenboer et al., 2015). Thus, the HO_2/OH ratio and the abnormally high HONO required to match the observed $P(O_3)$ provide evidence that at most only a part of the observed $P(O_3)$ can be explained by atmospheric HONO. "

Figure 5: If the disagreement between measured and modeled PO3 is strictly a function of NOx, why is the disagreement so much worse on plume days, even though plume and non-plume days both have average morning NO concentrations of approximately 2 ppb? Are there other significant differences between plume and non-plume days, or is there significantly more variability in the amount of NOx on plume than on non-plume days?

The difference between measured and modeled $P(O_3)$ is dependent upon NO_x. We have shown that the chemical mechanisms may be missing a hydroperoxy radical source co-emitted with NO relative to measurements and have explored several model case studies to try to reconcile this model-measurement disagreement.

 It is noted in Section 2.2 Pg. 7 lines 27-30 that NO_x and VOCs are both higher, on average, on plume days than on non-plume days. We have also shown that there can be significant differences (~5-10

ppbv/h) between the modeled and MOPS P(O_3) at NO levels greater than approximately 1 ppbv. Thus, the difference between the MOPS and models on plume vs. non-plume days is consistent with this result.

We have added the following text to Pg. 17 lines 10-12:

 "Measured O_x maxima are 2-7 ppbv greater on these ``plume" days than when air is advected from elsewhere, and higher P(O_3) is measured by the MOPS than is modeled by the RACM2 and MCMv331 (Fig. 7). This result is roughly consistent with the difference between measured and modeled P(O_3) as a function of NO shown in Fig. 4."

Technical Corrections Page 12, Line 12: If this number is a rate of HO2 production, then the units on this are incorrect (perhaps intended to be radicals cm-3 s-1?)

Corrected.

Page 13, Line 1-2: In this sentence, the phrase "the formation of" appears to be extraneous. The accent on FRAPPÉ is inconsistent. Sometimes the campaign is incorrectly written as FRAPPÈ.

Corrected.