We thank the reviewers for the thorough revisions and for providing constructive comments on our manuscript, "Trends of ground-level O_3 in Mexico during 1993-2014: Comparison of Monterrey with Mexico City and Guadalajara". We are pleased that the editor and reviewer's perspective on addressing O_3 long-term trends in Mexican urban areas is in agreement with our own views on the issue. We have addressed the concerns and recommendations received, and we believe that these helped to improve significantly the quality of our manuscript. Please find below our detailed response to the comments received, which are also highlighted in red in the revised version of the manuscript, submitted along with this response.

Reviewer #1:

Hernandez et al present trends in ozone precursor emissions and measured ozone levels in three urban areas in Mexico: Monterrey, Mexico City, and Guadalajara. This is an important research topic because, while there has been a long history of ozone trends analysis in the EU and US, there has been relatively little published on trends in other parts of the world. The paper itself needs some revisions before it is suitable for publication in ACP. Please see comments below.

Overarching comments:

Trends in emissions of ozone precursor: The authors need to more fully explore trends in ozone precursor emissions and discuss how the trends were derived. They provide some citations but don't address how reliable these sources are and whether there have been methodological changes over time in the emissions estimates that might impact the calculated emissions trends. Since these trends are later used to explain resulting ozone trends, they are a fundamental basis of the paper and need more discussion and exploration. In addition, Duncan et al., 2016 analyzed NO_x trends in these three metro areas based on satellite NO_2 column measurements between 2005-2014. The NO_x trends reported by Duncan et al do not match those reported by the authors in Fig 1a. For instance, Duncan et al (Table S9) found that NO_2 had decreased in Guadalajara in this period while Fig. 1a suggests that the increased. Additionally, Duncan found that NO_2 in Monterrey increased 8x more than NO_2 in Mexico City while Fig 1a shows them increasing at similar rates.
The authors should compare their results with Duncan et al and use this to explore uncertainties and limitations in the emissions trends shown in Fig 1a.

Response: The reviewer is right, no details regarding the methodology used to obtain the estimates of emissions and their uncertainty were included previously. The source of the emission estimates reported here and the methodologies used to obtain them were included in section 2.2 NEI data. See Lines: 181-205. We have also modified Fig. 1, now Fig. S1 to include more concise information, and discussed in the introduction section the uncertainties in the emission estimates reported in existing studies. See lines 108-122, 124-133.

Text modified:

"2.2 NEI data

Estimates of NO_x and VOCs emissions have been made at the national scale for the 1999-, 2005- and 2008-base years and reported in the NEI, and were obtained from the SEMARNAT website (http://sinea.semarnat.gob.mx). The data set is provided by emission source (mobile, point, area and natural), air pollutant, and at national, state and municipality scales. The NEI emission estimates are developed in accordance with the Manual for the Emission Inventories Program of Mexico (Radian, 2000), which is based on the US EPA AP-42 emission factors categorisation (EPA, 1995). The emission factors are regionalised for each Mexican state, based upon on-site measurements and survey information. Updates to the emission factors have been conducted for each released NEI, although no

changes in the methodology were implemented between the 1999- and 2008-base years. Overall, the mobile emissions were estimated using the MOBILE6-Mexico model (EPA, 2003). The emissions from point sources were derived using the annual operation reports submitted to the Environment Ministry. The emissions from area sources were obtained using the categorisation of Mexican area sources and the regionalised AP-42 emission factors.

The MCMA emissions inventories have been developed with a 2-year frequency since 1996, and were obtained from the MCMA Environment Secretariat website (http://www.aire.cdmx.gob.mx/). The methodology used to construct the MCMA inventories estimates is consistent with that used in the NEI (SEDEMA, 2016a), which is based on the AP-42 EPA emission factors. However, more speciated emission factors have been developed in each released version, considering updates in the local industrial activity, survey information and field measurement campaigns. To date, the only significant change in the methodology is the replacement of the Mobile6-Mexico model with the MOVES model to obtain the 2014-base year mobile emissions (SEDEMA, 2016b). As for the MCMA inventories, more speciated emission factors than those contained in the NEI were developed to produce the MMA emissions inventory 2013-base year (SDS, 2015), although, mobile emissions estimates were obtained with the Mobile6-Mexico model (EPA, 2003).".

Incomplete coverage of past trends work: In the introduction and throughout the paper the authors have a haphazard presentation of past trends work. One of the largest long-term ozone monitoring networks is located in the United States and yet the authors fail to cite any of the numerous studies looking at trends of US ozone (a subset of US trends references are listed at the end of the review). Rather, the authors inexplicably try to understand Monterrey O_3 trends by comparing them to studies from London, Tokyo and other far off places with little in common meteorologically or emissions-wise to Mexico. While it is worth discussing broadly the ozone trends across the Northern Hemisphere, the authors have a huge gap in this exploration because they don't include any work from the US. Additionally, when trying to explain/understand O_3 phenomena in Mexico, the authors should try to make comparisons to locations that have similar meteorological or emissions change drivers. Instead, the comparisons and reported trends from the literature are discussed in a disjointed way and don't provide an overall picture or provide context for the Mexican trends work presented here.

Response: The reviewer is right, no data regarding O_3 trends in US urban areas were included previously. As requested, we re-wrote the introduction section to include relevant information of O_3 trends in the US. We thank the reviewer for the list of references provided. See lines: 81-92.

Lack of transparency of O_3 metrics discussed: In the introduction, the authors cite numerous trends studies and say that ozone has changed by XX ppb but their description leaves out what metrics are being used. A 5 ppb change in annual average O_3 would mean something completely different than a 5 ppb change in 5th percentile or 95th percentile O_3. Additionally, O_3 calculated using all hours versus O_3 calculated using daily max (1-hr or 8-hr) will behave quite differently. In order for the reader to fully understand the literature that is being cited, the authors must provide information on which metrics the studies investigated. In addition, while the results in this paper do generally state the metric used, the authors switch between metrics (monthly avg – all hrs, annual avg – all hrs, 1-hr daily max values) without providing the reader with any information on why different metrics were used or how they might relate to each other. The authors need to provide more context in their own results about the meaning of each metric and what it reveals about O_3 changes.

Response: The reviewer is right, there was no description of the metrics used to derive the cited O_3 long-term trends. As requested, the metrics used to assess the changes in O_3 reported in the introduction were included in the text. See lines: 65, 71, 72, 74, 75, 80, 83-84, 88 and 89. Regarding the relevance of the metrics described in the current study, a sentence describing this was included before the discussion of each metric addressed. See lines: 135, 307-309, 367-371 and 402-403.

Specific comments:

Line 43: add ", methane" between "CO" and "and volatile organic compounds".
Response: ", methane" was added. See line: 43.

Line 87-92: It would be helpful if the authors provided some basic background information on the relationship between emissions of NOx and VOC and O_3. For instance explaining the conditions under which NO_x increases versus decreases O_3 concentrations.
Response: A brief description of the O_3 production regimes has been added. Text modified: "The system of O_3 production is not linear, being VOC-limited whether it responds to the input of VOCs, or NO_x-limited, whether O_3 production increases in response to increasing NO_x emissions (Monks et al., 2015; Pusede et al., 2015).". See lines: 45-47.

Line 142-143: But didn't the authors state that previous trends work had been conducted for Mexico City and Guadalajara?
Response: We have clarified in the text that the existing studies have focused mostly on long-term trends in O_3 within the MCMA. Additionally, we have stated that to date, only Benitez-Garcia et al. (2014) have considered changes in ground-levels of O_3 within the GMA and the MMA, however their results were obtained using the non-robust, simple regression analysis of annual averages, which could result in significant misestimations of the actual trends. See lines: 102-106.

Lines 176-178 and 180-186: These appear to be results which are stuck in the middle of the methods section. I suggest moving these to the results section.
Response: As suggested, lines 176-175 and 180-186 were moved to the results section, which are now part of section "**3.1 Wind occurrence at the MMA**". See lines: 272-279.

Lines 195-199: Were new and old instruments ever co-located to inter-compare the measurements? Just following QA procedures is probably not sufficient to control for changes in O_3 data due solely to different measurement techniques.
Response: Unfortunately, no simultaneous measurements of O_3 were performed using the 49 and 49C instruments; since the analysers model 49 reach their recommended operative life by early 2003, when were replaced with the analysers model 49C. However, to rule out the impact of different instrumentation and calibration methodologies studies have recommended the use of 3-yr O_3 averages, when no intercomparing measurement period was conducted. For instance, Akimoto et al. (2015) used 3-yr averages of O_3 when assessing long-term changes in ground-levels of O_3 at 4 large metropolitan areas of Japan. Similarly, Parrish et al. (2011) assessed the decreases in the O_3 4^{th} highest annual maximum mixing ratio within Los Angeles and the MCMA, using data calculated as 3-yr averages. This was noted in Lines: 171-173. Additionally, long-term trends in O_3 annual averages were compared with those derived using the methodology as above in Supplementary information S1.1 (Fig. S2).

Text modified:
S1.1 Comparison of long-term trends in O_3 annual averages with 3-yr averaged data
Linear trends were tested both for O_3 annual averages and 3-yr average O_3 data with the non-parametric Theil-Sen approach. Although, slight larger O_3 growth rates are determined for the smoothed data than for the annual averages as shown in Fig. S2, non-significant differences (p>0.05) were observed between both Sen slopes. Considering this, and that the smoothing of O_3 annual averages could lead to miss significant features in the current trends (Carslaw et al., 2007; Carslaw, 2015), in the current study, O_3 annual averages with no smoothing were used to determine the long-term trends reported at the 3 metropolitan areas.

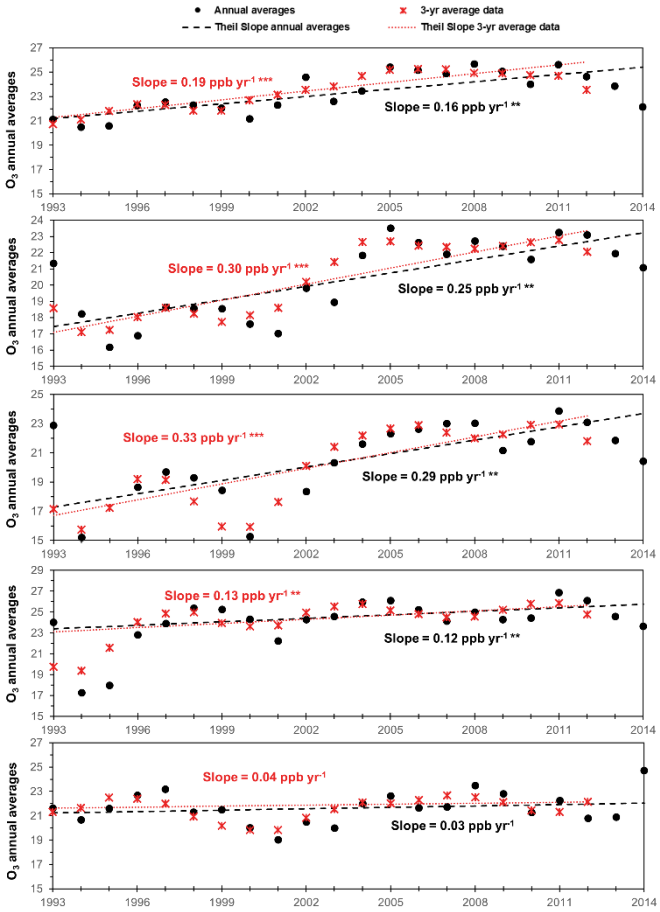

Fig. S2. Comparison of long-term trends in for O3 annual averages (1993-2014), and 3-yr average O_3 data (1993-2012). The dashed lines represent the Sen slopes. Statistical significance is expressed as p<0.1 =$^+$, p<0.05 = *, p<0.01 = ** and p<0.001 = ***.

Carslaw, D. C., and Carslaw, N.: Detecting and characterising small changes in urban nitrogen dioxide concentrations, Atmos. Environ., 41, 4723-4733, doi:10.1016/j.atmosenv.2007.03.034, 2007.

Parrish, D. D., Singh, H. B., Molina, L., and Madronich, S.: Air quality progress in North American megacities: A review, Atmos. Environ., *45*, 7015-7025. doi:10.1016/j.atmosenv.2011.09.039, 2011.

Akimoto, H., Mori, Y., Sasaki, K., Nakanishi, H., Ohizumi, T., and Itano, Y.: Analysis of monitoring data of ground-level ozone in Japan for long-term trend during 1990-2010: Causes of temporal and spatial variation, Atmos. Environ., 102, 302-310, doi:10.1016/j.atmosenv.2014.12.001, 2015.

Carslaw, D. C.: The openair manual - open-source tools for analysing air pollution data, Manual for version 1.1-4, King's College London, 2015.

Line 267: GPE had a higher max value than STA according to numbers reported in the following paragraphs.

Response: The reviewer is right. There was a mistake in the sentence, we have specified the site where the highest O_3 1-h average mixing ratio was measured during the studied period. Text modified: "The highest O_3 1-h average was observed at SNB,". See lines: 283-285.

Line 276: It would be more accurate if this sentence read: "Reaction with O3 rapidly converts NO to NO_2".

Response: The sentence was modified as suggested. Text modified: "Reaction with O_3 rapidly converts NO to NO_2,". See line: 294.

Line 287: Here the authors switch from data using all hours (and daily averages) to daily max 1-hr O_3 values. They should note the switch and explain the importance of the different metrics.

Response: As requested by the reviewer, the relevance of the long-term trend assessment for maximum O_3 1-h averages was stated. Text added: " A study conducted among asthmatic children resident in the MCMA revealed an increase in coughing and wheezing rates, associated with cumulative exposure to high 1-h averages mixing ratios of O_3 and NO_2 (Escamilla-Nuñez et al., 2008). To assess changes in cumulative exposure to O_3 and O_x within the MMA, long-term trends of de-seasonalised maximum daily 1-h averages in O_3, O_x and NO_x were calculated, using annual averages filtered with the STL technique (Fig. 4). ". See lines: 307-311.

Escamilla-Nuñez, M. -C., Barraza-Villarreal, A., Hernandez-Cadena, L., Moreno-Macias, H., Ramirez-Aguilar, M., Sienra-Monge, J. -J., Cortez-Lugo, M., Texcalac, J.-L., del Rio-Navarro, B., and Romieu, I.: Traffic-related air pollution and respiratory symptoms among asthmatic children, resident in Mexico City: The EVA cohort study, Respir. Res., 9, doi:10.1186/1465-9921-9-74, 2008.

Line 288: Is this significant? If so state *p*-value.

Response: As requested by the reviewer, *p*-values were written along the text where required. See line: 312.

Line 289: Here you state that changes 0.79 ppb/yr are “large” but on line 38 you referred to a change of 0.76 ppb/yr as “gradual”. Be consistent with characterization of these trends.

Response: The reviewer is right. We have changed "gradual" in Line 38 to "large" in order to be consistent with the trends characterisation. See line: 38.

Line 290: the authors should state the magnitude and direction of the trend at STA is before discussing causes.

Response: As requested by the reviewer, the magnitude of the trend and significance value were stated. Text added: " By contrast, the non-significant (p>0.05) trend of -0.01 ppb O_3 yr^{-1} observed at STA is may be masked by local import of O_3…”. See lines 315-318.

Line 315: What are the daily O_3 profiles normalized to? It is not clear what calculations were performed here.

Response: The reviewer is right. We have stated how the normalised cycles were constructed. Text added: " To compare the O_3 diurnal cycles by season, normalised daily profiles were constructed by subtracting daily averages from hourly averages in order to remove the impact of the long-term trends (Fig. 6; Hernández-Paniagua et al., 2015), with daily amplitude values (AV_d; calculated by subtracting the lowest normalised values from the highest normalised values) used to assess diurnal variations in O_3 among seasons.”. See lines: 345-349.

Lines 315-326: It would be interesting if the authors could discuss whether AV_d has changed over time.

Response: As requested by the reviewer, the long-term trends in AV_ds from 1993 to 2014 were determined for the 5 sites within the MMA. Figure 7 shows the long-term trends in AV_ds, which are discussed in the manuscript. See lines: 361-371.

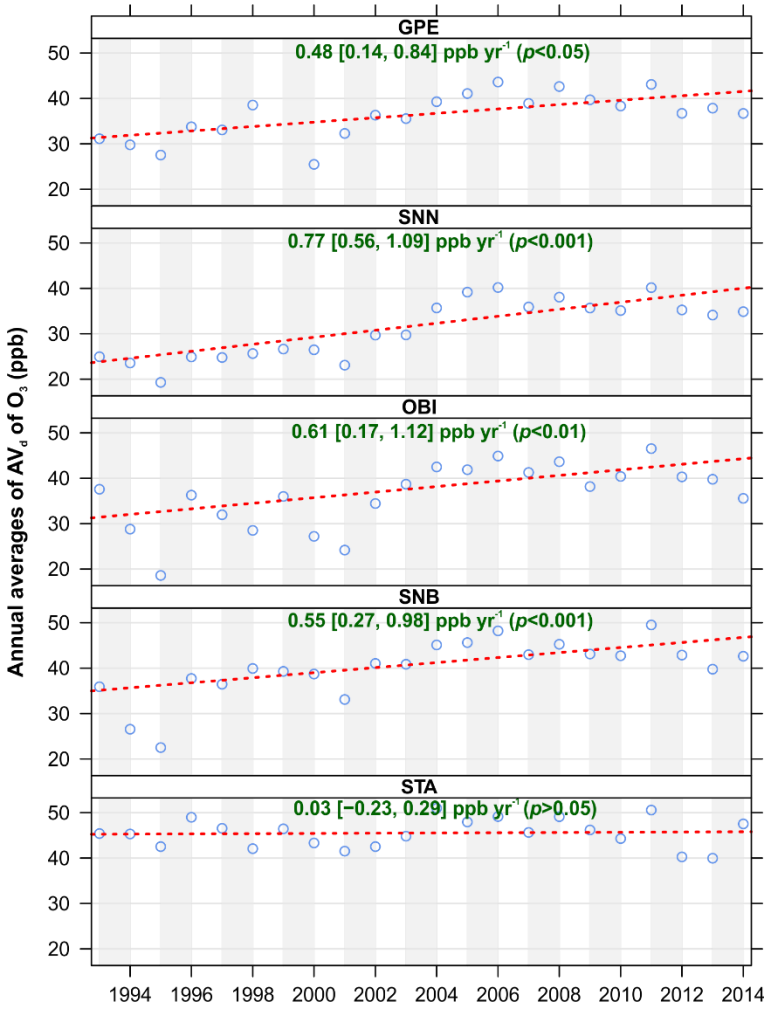

Fig. 7. Long-term trends of AV_d O_3 annual averages at the 5 sites within the MMA during 1993-2014. The dashed lines represent the Sen slopes. Statistical significance is expressed as p<0.1 =$^+$, p<0.05 = *, p<0.01 = ** and p<0.001 = ***.

Lines 329-338: In contrast, the maximum O_3 concentrations in the US usually occur in June-August. It would be good to note this difference.

Response: As requested, O_3 seasonal cycles within the MMA are compared with those reported for several regions of the US, including the southeast. This as the MMA is also influenced by air masses from the Gulf of Mexico. Text modified: "Figure 8b shows the seasonal cycles of O_3, with spring-time maxima and winter minima, in strong correlation with SR (Lelieveld and Dentener, 2000).

This behaviour agrees well with the O_3 spring maxima and winter minima characteristic of the US southeast regions (Strode et al., 2015), and follows the NH mid-latitudes O_3 cyclic pattern (Monks 2000; Vingarzan, 2004). However, it differs with the O_3 seasonal cycles observed over the US west coast regions (particularly in California), where the maxima occur between June-August, in response to the local influence of precursor emissions upon O_3 production and photochemical conditions (Vingarzan, 2004; Strode et al., 2015). By contrast, downward spikes in the seasonal cycles of O_3 within the MMA are observed recurrently between July-August (Fig. 8b), which likely result from high wind speeds (>6 km h^{-1} in average) that disperse O_3 precursors and increase the boundary layer height (ProAire-AMM, 2008), and high day-time temperatures (>40° C) that could suppress the O_3 formation. Steiner et al. (2010) reported that within VOC-limited areas, temperatures >38° C may lead to decreases in O_3 formation, in response to a decrease in the peroxyacetyl nitrate lifetime (NO_x sink). The peak in O_3 observed in September is characteristic of humid regions, and can be ascribed to an increase in OH radicals derived from the increment in RH during the rainy season (Lee et al., 2014). Zheng et al. (2007) reported that this O_3 secondary peak became less noticeable since 2000 over the mid-western and

eastern US regions. Indeed, the O_3 secondary peak is characteristic of the Asian summer monsoon, which transports maritime clean air to land with constant rainfall, thereby increasing RH (Xu et al., 2008).". See lines: 381-400.

Line 355: The authors state that AVs are similar to those recorded in the US but they have provided no information about the US with which to make this comparison.
Response: Data of the seasonal cycles over the US were included in order to discuss AV_s with those observed within the MMA. Text modified: " AV_s for the MMA are similar to those calculated using dynamic linear models by Zheng et al. (2007), over the mid-western US region between ca. 12 ppb O_3 in 2004 and 18 ppb O_3 in 1999, but lower than those between ca. 19 ppb O_3 in 2004 and 27 ppb O_3 in 1999 determined for the eastern region. When compared with European regions, the AV_s determined within the MMA are slightly lower than those calculated at the North Kensington site in London, which ranged from ca. 7.0 ppb O_3 in 2000 to ~25.5 ppb O_3 in 2005 (Bigi and Harrison, 2010), presumably due to lower emissions of NO_x and VOCs within the MMA (SDS, 2015). It is striking that the average AV_s for the MMA agrees well with that of 10.5 ppb O_3 recorded during 2004-2005 at the Pico Mountain Observatory in Portugal, which is a receptor of exported NA air pollution (Kumar et al., 2013). Thus, despite trends of increasing O_3 precursor emissions within the MMA, AV_s lie within the range of those recorded at sites in the mid-west US, but are slightly lower than those determined for more populated and urbanised sites in the east US and Western Europe.". See lines: 407-418.

Line 371: Are monthly averages calculated using all hours or just daytime max values?
Response: The O_3 monthly averages were derived from daily averages of all 1-h data, as described in section 2.3, which was clarified in the manuscript. Text added: "The long-term trends were constructed from de-seasonalised annual data derived from monthly averages filtered with STL, which were calculated from daily data of all 1-h averages, as described in Methodology (Sect. 2.3). ". See lines: 440-442.

Lines 381-389: Duncan et al. can provide NO_2 trends at many more locations than just Toronto. Also the US EPA publishes trends reports which include trends in emissions which could be used for comparison.
Response: As requested, we have discussed the O_3 trends observed in terms of response to changes in NO_x determined within the MMA, with contrast with the NO_2 trends reported by Duncan et al. and economic indicators. See: Figure 10. Additionally, studies of relevance conducted within the MMA and from the list provided by the reviewer were used to discuss and explain the observed trends in O_3 within the MMA. See lines: 459-465, 469-479, 481-489, 491-503.

Lines 444-448: This explanation does not fit with the current literature. The most dramatic weekend/weekday effects have been observed in Southern California under VOC limited conditions, so VOC limitation would not explain the lack of a weekend/weekday effect.
Response: As requested by the reviewer, the discussion of the O_3 weekly cycles section was re-written. We discussed the non-significant changes between weekdays and weekends reported in our study with those reported by Wolff et al. (2013) for urban areas across the US. Moreover, we provide a plausible explanation for the O_3 weekly patterns observed, based on i) the assessment of ambient levels of O_3 as reported by Torres-Jardon (2004) for the MCMA and, ii) the vanishing effect in urban areas of the Southern California reported by Wolff et al. (2013). Text modified: "No significant differences (p>0.05) were observed at any of the metropolitan areas between O_3 AV_d during weekends and weekdays. This lack of a weekend effect in O_3 was reported previously at the MCMA for 1987-2007 by Stephens et al. (2008), who attributed it to weekday O_3 production being limited by VOCs and inhibited by NO_x; this was also observed by Song et al. (2010). By contrast, simultaneous decreases in emissions of VOCs and NO_x mostly from vehicle sources during weekends could have counteracting effects on the O_3 production rates, leading to similar levels of O_3 during weekdays at the 3 metropolitan areas. This behaviour was

reported previously by Wolff et al. (2013) for US urban areas of the Northeast, Midwest and Coastal California regions, which exhibited similar or even higher (±5 %) O_3 levels during weekdays than at weekends, despite lower O_3 precursor emissions during weekends. Moreover, Wolff et al. reported that from 1997-1999 to 2008-2010 the sites studied exhibiting a weekend effect decreased from ca. 35 % to less than 5 %, which was attributed to an increase in the VOC/NO_x emission ratio derived from a greater decline in NO_x than in VOCs emissions (Pusede et al., 2014).

It is likely that the O_3 weekly patterns observed at the metropolitan areas arise from reduced traffic activity during weekends, leading to increases in ratios of VOCs/NO_x. Within the MMA, this would be confirmed by lower NO_x mixing ratios (on average 5 %) during weekends, changing to a transition O_3 production between VOC- and NO_x-limited during weekends. Moreover, a change to a NO_x-limited O_3 production derived from the reduction in NO_x seems unlikely since this would result in lower O_3 levels during weekends, not observed at any of the studied urban areas (Torres-Jardon et al., 2004).". See lines: 531-542 and 544-550.

Lines 451-456: Past work (Simon et al, Cooper et al) has shown that O_3 trends are much more pronounced at high percentiles than at average levels, so an annual average may not be a very good metric to use to see long-term trends.
Response: As suggested by the reviewer, we have included for the 3 metropolitan areas, the analysis of long-term trends at the annual 5th and 95th percentiles, median and averages. The observed trends are discussed with those reported in the references provided. Text modified: "Long-term trends of the annual 5th and 95th percentiles (%ile), median and average of O_3 during 1993-2014 were calculated using the Mann-Kendall test and Sen's estimate for the 5 sites within the MMA (Salmi et al., 2002; Carslaw and Ropkins, 2012), and are shown in Fig. 9. The long-term trends were constructed from de-seasonalised annual data derived from monthly averages filtered with STL, which were calculated from daily data of all 1-h averages, as described in Methodology (Sect. 2.3). Overall, O_3 shows significant increasing trends (p<0.05) mostly in the annual averages ranging from 0.11 ppb O_3 yr^{-1} at SNB to 0.31 ppb O_3 yr^{-1} at OBI, and in the 95th %ile, which ranged from 0.39 ppb O_3 yr^{-1} at OBI and SNB to 0.75 ppb O_3 yr^{-1} at SNN. The 5th %ile increased significant only at OBI in 0.08 ppb yr^{-1}, while the median increased at SNN by 0.14 ppb O_3 yr^{-1} and at OBI by 0.23 ppb O_3 yr^{-1}. Note that if trends are segmented and considered only after the decline in 1994-1995, the only significant change is that the O_3 growth rate at SNN would increase to 0.31 ppb O_3 yr^{-1} and GPE would decrease to 0.14 ppb O_3 yr^{-1}, while in the 95th %ile the trends would decline slightly at GPE and SNB to 0.27 ppb O_3 yr^{-1}, and at OBI to 0.42 ppb O_3 yr^{-1}. Despite exhibiting the highest O_3 mixing ratios within the MMA, STA did not exhibited significant trends in any of the tested metrics." See lines: 438-451.

"Long-term trends of de-seasonalised O_3 annual median, 5th and 95th percentiles at the 3 urban areas were determined following the same methodology as for annual averages (Fig. S10). Overall, the linear trends observed in O_3 annual averages for the MMA and MCMA are also seen in the other tested metrics, with significant (p<0.05) increases at MMA ranging from 0.05 ppb O_3 yr^{-1} (5th percentile) to 0.41 ppb O_3 yr^{-1} (95th percentile), and decreases at MCMA between 0.37 ppb O_3 yr^{-1} (5th percentile) and 2.32 ppb O_3 yr^{-1} (95th percentile). As for the O_3 annual averages, the GMA shows non-significant (p>0.05) trends in the other tested metrics. Notably, only the tropospheric CO decreased significantly (p<0.05) at the 3 urban areas studied, with the largest decrease rate of 0.12 ppm CO yr^{-1} detected at the MCMA and the lowest one of 0.02 ppm CO yr^{-1} calculated at the MMA. Thus, whereas O_3 precursors have decreased linearly within the MCMA and the GMA during the studied period, within the MMA those have increased during the same period despite the introduction of emission control policies (SDS, 2015)." See lines: 566-576.

Lines 458-464: Zheng et al and Camalier et al have analyzed the impact of inter-annual meteorological variation on O_3 trends. These studies should be cited and discussed.

Response: As requested by the reviewer, the references provided were cited and discussed in the seasonal cycles analysis section. Briefly, since O_3 time-series contain a significant seasonal component as reported in the literature, several methodologies have been developed to remove it and filter the influence of meteorology when determining long-term trends. In the present study, the STL technique (Cleveland et al., 1990) was used to filter out the seasonal component from the O_3 data, as the seasonality accounts for the year-to-year variation caused by changes in SR, RH, temp. As described along the manuscript, all annual data used to determine long-term trends for all pollutants analysed were derived from de-seasonalised data. Therefore, it is expected that the reported trends have no significant influence of the year-to-year variations in meteorology. See lines: 374-383, 393-400 and 582-585.

Lines 466-475: The explanation linking O_3 trends to emissions trends does not follow logically and is in contrast to results presented by Duncan et al.

Response: The results presented in our manuscript were revised and contrasted with the trends reported by Duncan et al. (2016). Additionally, we conducted an exhaustive revision of the data reported in the NEI and local emission inventories to verify consistency in methodologies used, which is described in section 2.2. See lines: 181-193, 195-205, 459-465, 474-479 and 572-576.

Tables 3 and 4: Are O_3 statistics based on hourly O_3 data or some other averaging period/daily max period. Please clarify in text and table headings.

Response: As requested by the reviewer, the resolution of O_3 data reported in Table 3 and 4 (now Table S1 and S2) was added. See: Table S1 and Table S2. See line: 287.

Fig 1a: Text should describe how this figure was created from the data sources listed. Do different data sources/years use consistent methodologies?

Response: The methodologies used to obtain the emission estimates are included in section 2.2. Fig. 1 was moved to Supplementary information (now Fig. S1). Fig. S1 only shows NEI emission data of VOCs and NO_x as described in the caption.

Fig 3: The label for panel d is missing.

Response: Label (d) was included in the Fig. S5.

Figure 8: How were 95% CIs constructed? Were they based on all daily values? Or on variation among sites in annually averaged profiles? In either case, these confidence intervals look VERY small, I think there is an error in the plotting. It is hard to believe that there would be so little day to day or site to site variability.

Response: The 95 % confidence intervals shown in Fig. 12 were calculated through bootstrap re-sampling (Carslaw et al. 2015), since it provides a better estimation compared with calculations based on normal data distributions. This explains the small confidence intervals compared with those constructed using a parametric test.

References provided by the reviewer:

Camalier, L.; Cox, W.; Dolwick, P. (2007) The effects of meteorology on ozone in urban areas and their use in assessing ozone trends, Atmospheric Environment, 41, 7127-7137.

Duncan, B. N., L. N. Lamsal, A. M. Thompson, Y. Yoshida, Z. Lu, D. G. Streets, M. M. Hurwitz, and K. E. Pickering (2016), A space-based, high-resolution view of notable changes in urban NOx pollution around the world (2005–2014), J. Geophys. Res. Atmos., 121, 976–996, doi:10.1002/2015JD024121.

Zheng, J.; Swall, J.L.; Cox, W.M.; et al. (2007) Interannual variation in meteorologically adjusted ozone levels in the eastern United States: A comparison of two approaches, Atmospheric Environment, 41, 705-716.

**US O3 trends references (partial list):*

Berlin, S. R.; Langford, A. O.; Estes, M.; Dong, M.; Parrish, D. D. Magnitude, decadal changes, and impact of regional background ozone transported into the greater Houston, Texas, Area. Environ. Sci. Technol. 2014, 47, 13985−13992.

Butler, T. J.; Vermeylen, F. M.; Rury, M.; Likens, G. E.; Lee, B.; Bowker, G. E.; McCluney, L. Response of ozone and nitrate to stationary source NOx emission reductions in the eastern USA. Atmos. Environ. 2011, 45 (5), 1084−1094.

Chan, E., Regional ground-level ozone trends in the context of meteorological influences across Canada and the eastern United States from 1997 to 2006. J. Geophys.Res.-Atmos. 2009, 114.

Chan, E.; Vet, R. J. Baseline levels and trends of ground level ozone in Canada and the United States. Atmos. Chem. Phys. 2010, 10(18), 8629−8647.

Cooper, O. R.; Parrish, D. D.; Stohl, A.; Trainer, M.; Nedelec, P.; Thouret, V.; Cammas, J. P.; Oltmans, S. J.; Johnson, B. J.; Tarasick, D.; Leblanc, T.; McDermid, I. S.; Jaffe,D.; Gao, R.; Stith, J.; Ryerson, T.; Aikin, K.; Campos, T.; Weinheimer, A.; Avery, M. A. Increasing springtime ozone mixing ratios in the free troposphere over western North America. Nature 2010, 463 (7279), 344−348.

Cooper, O. R.; Gao, R. S.; Tarasick, D.; Leblanc, T.; Sweeney, C., Long-term ozone trends at rural ozone monitoring sites across the United States, 1990−2010. J. Geophys. Res.-Atmos. 2012, 117.

Hidy, G.M.; Blanchard, C.L. (2015) Precursor reductions and ground-level ozone in the Continental United States, Journal of the Air & Waste Management Association, 65, 1261-1282.

Hogrefe, C.; Hao, W.; Zalewsky, E. E.; Ku, J. Y.; Lynn, B.; Rosenzweig, C.; Schultz, M. G.; Rast, S.; Newchurch, M. J.; Wang, L.; Kinney, P. L.; Sistla, G. An analysis of long-term regional-scale ozone simulations over the Northeastern United States: Variability and trends. Atmos. Chem. Phys. 2011, 11 (2), 567−582.

Lefohn, A. S.; Shadwick, D.; Oltmans, S. J. Characterizing changes in surface ozone levels in metropolitan and rural areas in the United States for 1980−2008 and 1994−2008. Atmos. Environ. 2010, 44 (39), 5199−5210.

Lin, C. Y. C.; Jacob, D. J.; Munger, J. W.; Fiore, A. M. Increasing background ozone in surface air over the United States. Geophys. Res. Lett. 2000, 27 (21), 3465−3468.

Lin, M., Horowitz, L.W., Cooper, O.R. et al (2015) Revisiting the evidence of increasing springtime ozone mixing ratios in the free troposphere over western North America, Geophysical Research Letter, 42, 8719-8728.

Parrish, D. D.; Law, K. S.; Staehelin, J.; Derwent, R.; Cooper, O. R.; Tanimoto, H.; Volz Thomas, A.; Gilge, S.; Scheel, H. E.; Steinbacher, M.; Chan, E. Long-term changes in lower tropospheric baseline ozone concentrations at northern mid-latitudes. Atmos. Chem. Phys. 2012, 12 (23), 11485−11504.

Pusede, S.E.; Steiner, A.L.; Cohen, R.C. (2015) Temperature and Recent Trends in the Chemistry of Continental Surface Ozone, Chemical Reviews, 115, 3898-3918.

Sather, M. E.; Cavender, K. Update of long-term trends analysis of ambient 8-h ozone and precursor monitoring data in the South Central U.S.; encouraging news. J. Environ. Monit. 2012, 14 (2), 666− 676.

Sather, M.E. and Cavender, K. (2016) Trends analyses of 30 years of ambient 8 hour ozone and precursor monitoring data in the South Central US: progress and challenges, Environmental Science-Processes & Impacts, 18, 819-831.

Simon, H.; Reff, A.; Wells, B.; et al. (2015) Ozone Trends Across the United States over a Period of Decreasing NOx and VOC Emissions, Environmental Science & Technology, 49, 186-195.

Souri, A.H.; Choi, Y.; Li, X.; et al. (2016) A 15-year climatology of wind pattern impacts on surface ozone in Houston, Texas, Atmospheric Research, 174, 124-134.

Strode, S.A.; Rodriguez, J.M.; Logan, J.A.; et al. (2015) Trends and variability in surface ozone over the United States. JGR-Atmospheres, 120, 9020-9042.

Tanner, R. L.; Bairai, S. T.; Mueller, S. F. (2015) Trends in concentrations of atmospheric gaseous and particulate species in rural eastern Tennessee as related to primary emission reductions, Atmospheric Chemistry and Physics, 15, 9781-9797.

U.S. EPA, Our Nation's air: Status and trends through 2015. Office of Air Quality Planning and Standards: RTP, NC, 2015, https://gispub.epa.gov/air/trendsreport/2016/