

The reviewer's comments are in bold type and our responses are in normal text. We thank the reviewer for his/her careful reading of our manuscript and for the comments.

General comments:

1. I think the analysis of the data in some cases has been stretched to the limit of credibility; in particular the comparison of weekday / weekend effects from such a limited data set. The authors should put their 15 days worth of data into context using the longer record of data from the near road site.

The authors recognize that 4 days on weekends are not statistically robust enough to show the difference between weekdays and weekends, especially for O₃. Therefore, we extracted the summer O₃ data in 2013, 2014, and 2015 from a regular NAPS station near the highway. Please see the details in point 12 below, the revised manuscript, and the supplemental material.

Minor comments:

2. P3. It is in the interest of brevity, the section on page 3 describing how PBL dynamics can impact surface concentration of pollutants is probably unnecessary for the readership of this journal. I found this introductory material unnecessary and I think it adds to manuscript bloat.

We believe this section introduces an essential aspect of our study, and would like to keep it in. We have shortened it by removing 3 sentences. We have also moved most of section 2.2 to the supplementary material.

3. P3. "... first direct comparison of this kind..." It would be good to check the publication of M. Grutter at Centro de Ciencias de la Atmósfera, UNAM, Ciudad Universitaria, Mexico City. He also uses OP FTIR and there had been some big field international air quality field experiments in Mexico City over the last 15 years that would have likely produced opportunities for OP FTIR / fixed point measurement comparisons. I know he has done this for formaldehyde.

We thank the reviewer for the references. Yes, Grutter et al. (2005) did a comparison of op-FTIR with a point measurement at the same site for formaldehyde mixing ratio in downtown Mexico City in 2003. They also did a comparison of op-FTIR measurement of CO with a point measurement of CO at a different site.

We have included Grutter et al. (2005) in the main text. At the end of 5th paragraph in the Introduction, "Grutter et al. (2005) measured the formaldehyde mixing ratio by op-FTIR in downtown Mexico City in 2003 and compared it with a point measurement at the same site." And in the second last paragraph of Introduction, "To our knowledge, (Grutter et al., 2005) presents the first of very few direct comparisons of this kind to be published."

4. P4. Experimental section should list dates of the study period.

The dates of the study have been added to the first paragraph in the Experimental section.

5. P8. It is not clear why the GEM-MACH model results for CO was averaged over 3 hours (1 hour period on each side of the h1-hr period of interest) to get a running average to compare with the 1-hr averages of the data?

This is a misunderstanding. GEM-MACH model outputs were not averaged over 3 hours. This version of GEM-MACH produced concentration of pollutants every hour hh:00 (a snapshot), and we wanted an averaged concentration over 60 minutes, so we took two consecutive outputs from GEM-MACH, averaged these two outputs, and used this as the average concentration for the 60 minutes between these snapshots. This has been described in the second paragraph of Section 2.4.

6. P8. WindTrax. The discussion didn't make clear how the concentration at the measurement site was apportioned to the source area (highway lanes) of interest. Wouldn't the back trajectory model need a high resolution emission model to determine what mass of CO measured at the site was from the emission area of interest? This needs to be clarified for the reader who hasn't used WindTrax. Why is this model needed for equation (9) if the denominator is being determined by another model (the bLS model)? I found this section confusing.

WindTrax is a software tool which incorporates a backward Lagrangian Stochastic (bLS) model (Flesch et al., 1995) with a graphical interface. In this interface, we defined the paved surface of Highway 401 as the only source of CO and other primary pollutants listed in Table 2. The bLS model then releases a large number of virtual particles from the point of measurement and calculates individual trajectories backward in time, based on the input meteorological conditions (wind direction, u_* , L , temperature). The fraction of trajectories that originate from the defined source area is then determined, and used to calculate the simulated relationship between source strength and concentration, as given in the denominator of equation (1). (Please note that equation (9) in the previous version of the manuscript is now equation (1) in the revised manuscript, since we decided to put the theory and equations for the scintillometer in the supplementary material.) This equation can then be used to calculate the actual source strength, given the observed and background concentrations.

To make it clear to readers, Section 2.5 has been revised. “We used a backward Lagrangian Stochastic (bLS) model (WindTrax, <http://www.thunderbeachscientific.com> ; Flesch et al., 1995) which calculates the emission rate Q through formula (1) where C is the concentration of a pollutant at the measurement site, C_b is the background concentration, and $(C/Q)_{sim}$ is the simulated ratio of concentration at the site to the emission rate upwind.”

7. Figure 2. I can't tell the difference between the line for z/L and the line for u_* .

The color of the line for u_* has been changed to red, to distinguish it from z/L .

8. Fig 4. This is a nice figure but it isn't clear from the text what is actually plotted – the image looks smoothed to color code difference ranges rather than being a collection of individual data points.

Figure 4 shows the difference of CO concentrations (as the color code) between FTIR measurements and NAPS measurements, as well as GEM-MACH model results and NAPS measurements. Yes, it is correct that the color has been smoothed. The angle is the wind direction, and the radius is the wind speed. We used this figure to show that the difference of concentration from these two sources depends on the wind direction (i. e. across the highway or not). Since the color smoothing may be confusing to other readers as well, we have changed these figures back to show individual data points.

9. P11. Ambient temperature. It is well known that traffic emission of CO can be influenced by temperature but this is primarily due to start emissions when catalytic converters are still cold (< 200 C). Vehicle running emissions of CO are not strongly influenced by ambient temperature. This section has an odd reference “Choi pdf” accessed from the internet. It would be better to cite an actual EPA report on MOVES temperature parameterization of vehicle emissions. One suggested reference is “MOVES2010 Highway Vehicle Temperature, Humidity, Air Conditioning, and Inspection and Maintenance Adjustments”, EPA-420-R-10-027.

Thank you! This EPA report has been added to the references of the revised manuscript.

10. P12. Why was a background value of 256 ppbv used for CO – what is the reasoning for this as a “background” values for the airshed or for upwind of the FTIR beam? Do you get the same value for a CO vs NOx regressions? Air entering the urban airshed or crossing the highway will contain NH3 and CO – shouldn’t these background values be subtracted from both to reveal increase due to local traffic emissions? This background value is an important number as it is later used in the WindTrax calculations so it deserves better definition.

In the revised supplemental material, we have added the plot of [NH₃] vs. [CO] from FTIR results (Fig. S5). From the linear relationship analysis, 265 ppb is the intercept of [CO] ([CO] = 256 ppb when [NH₃] = 0 ppb). Our reasoning was that we assumed that traffic emissions were the only source of CO above background at this spatial scale, and all the NH₃ from traffic emissions were associated with CO. The slope of the linear regression between [NH₃] and [CO] has also been reported by Livingston et al. (2009) and Perrino et al. (2002). We have compared our slope with those slopes, as discussed in Section 3.3 of our manuscript.

From our FTIR observations, [NH₃] reached 0 ppb sometimes (see time series of [NH₃] in Fig. 6), so we did not consider NH₃ background from other sources in this analysis. We used a linear relationship between ([CO]-[CO_background]) and [NH₃], to estimate NH₃ from traffic emissions, which is shown in Fig. 5. We stated “ Overall, there is no indication of a background offset of NH₃, and most measured NH₃ at this site can be accounted for by traffic emissions ” in the manuscript.

As the reviewer suggested, we did another regression analysis: [CO] vs. [NO_x] from NAPS measurement. The intercept of CO on weekdays and weekends are 196 and 186 ppb. This is discussed in point 11 below.

11. P14. It is more common in the literature to report CO vs NOx regressions and to discuss CO-to-NOx molar ratios (cf. the papers by D.D. Parrish or Wallace et al Atmos Environ. 2012). A

ratio of ~5 would be expected for running emissions at your site. I think it would be better to show Figure 10 in the traditional way (NO_x vs CO) so that your slopes could be compared with the literature and vehicle emission inventory.

As mentioned in Section 3.4, we did not retrieve reliable mixing ratios of NO and NO₂ from the FTIR, so we could not do the same regression analysis of [CO] vs. [NO_x] for FTIR measurements. However, we did a regression analysis of [NO_x] vs. [CO] from NAPS measurements as shown in Fig. 10. These slopes have been compared to previous values in Kim et al. (2016) in Section 3.4 in the discussion manuscript. There is another recent publication, Hassler et al. (2016) (D. D. Parrish is one of the authors) showing a long-term observation of [NO_x] and [CO]. In that paper, they showed [NO_x] / [CO] from LA basin observations as well as from near road monitoring stations from Paris and London. Our slopes of 0.1- 0.2 agree well with observation data from LA Basin after 2010.

Figure 10 is in the orientation of [NO_x] vs. [CO] already. Perhaps the reviewer meant a plot of [CO] vs. [NO_x]? We did another regression analysis with NAPS measurement [CO] vs. [NO_x]. The slopes are 3.14 and 7.75 for weekdays and weekends, respectively, i.e., bracketing the expected slope of 5. In Parrish et al. (2002), the slopes are in the range of 6.3 to 18.9, given data from 1987 to 1999. In Wallace et al. (2012), the slope is 4.2 from morning rush hours in Meridian, Idaho, USA in 2008- 2009.

The ratio of [NO_x] / [CO] depends on the mix of vehicle types, fuels and engines, driving cycles, mileage, and emission control techniques. Figure 10 also shows the difference of [NO_x] / [CO] between weekdays and weekend, due to decreased numbers of heavy-duty diesel vehicles on the weekends (as discussed in Section 3.4). It is also consistent with Hassler et al. (2016) showing ratios of [NO_x] / [CO] in Paris and London are higher than ratios in the LA Basin after 2010, due to larger fractions of diesel vehicles in Paris and London than in the LA Basin. Since there are recent publications of [NO_x] vs. [CO], and our slopes can be compared to them, we think either way is acceptable.

We have added this discussion on [CO] vs. [NO_x] in the main text at the end of the third paragraph in Section 3.4: “ Hassler et al. (2016) showed the trend of [NO_x] / [CO] in the Los Angeles Basin, and the ratio is between 0.1 and 0.2 after 2010, which agrees well with our results. There are also some previous studies showing ratio of [CO] / [NO_x] from regions near heavy traffic emission. Parrish et al. (2012) reported the slope of [CO] vs. [NO_x] was in the range of 6.3 to 18.9 for the measurement from 1987 to 1999. Wallace et al. (2012) reported the slope of [CO] vs. [NO_x] was 4.2 in the morning rush hours in 2009. The slopes of [CO] vs. [NO_x] in this study are 3.14 and 7.75 for weekdays and weekends, respectively. Therefore, our results on NO_x and CO are comparable with these previous studies.”

Hassler, B., McDonald, B.C., Frost, G.J., Borbon, A., Carslaw, D.C., Civerolo, K., Granier, C., Monks, P.S., Monks, S., Parrish, D.D., Pollack, I.B., Rosenlof, K.H., Ryerson, T.B., von Schneidemesser, E. and Trainer, M.: Analysis of long-term observations of NO_x and CO in megacities and application to constraining emissions inventories, *Geophys. Res. Lett.*, 43, 9920-9930, doi: 10.1002/2016GL069894, 2016.

12. P14. The analysis of the weekend / weekday comparison of ozone is perhaps more than what the data can support. There were only 2 weekend periods. Is this really enough data to statistically demonstrate that weekends have different ozone production rates that weekdays? Isn't the production and accumulation of ozone in the airshed also affected by meteorology (irradiance,

dispersion)? How were these factors accounted for? You state poor statistics in explaining HCHO patterns. The weekday /weekend difference of vehicle emission on ozone production is interesting but you do not have a statistically relevant difference with 15 days of data. This should be recognized in this section. I would recommend you can place the campaign data into context with ozone data from the NAPS site for a multi-year summer period.

Yes, we agree that only 16 days of results, including 4 weekend days, may not be statistically robust enough to show the difference on ozone production. We also think that the different levels of NO_x between weekdays and weekends affect the ozone production. O_3 production actually goes down at high NO_x (on weekdays), because the OH is reacting with NO_2 to form HNO_3 and not with VOCs to form O_3 and HCHO. Weekend NO_x levels more efficiently form O_3 and HCHO.

Yes, the production and accumulation of O_3 in the surface air probably was also affected by meteorology. As has been discussed in the CO and NH_3 sections, turbulent mixing was much stronger during the day, and sensible heat flux H reached maximum during 12 to 14:00 EDT (Fig. 2). Turbulent mixing would “dilute” the pollutants accumulated in the surface air, so they would not contribute to the peak of O_3 in the early afternoon.

The NAPS station which provided measurement results for this manuscript was a recently installed station, so it was there only for that summer. However, there is another nearby NAPS station which has been in operation for several years. This station is situated about 175 m south of Highway 401, and about 200 m southwest of the new NAPS research trailer. We have extracted the O_3 data from that NAPS station for July and August from 2013 to 2015, to show some multi-year statistics.

Diurnal plots of O_3 on weekdays and weekends in the summer of 2013, 2014 and 2015 are shown in Fig. S7. For 2014 and 2015, the median of the diurnal average on weekends from 7 am to midnight are higher than median on weekdays, which is consistent with the results from the new research NAPS trailer shown in the main text and Fig.9. There is no resolvable difference between weekdays and weekends in 2013. Therefore, the difference between weekdays and weekends we observed in our 16-day measurement agrees with longer observations in 2014 and 2015, which suggests it is representative. We are not expecting the results from these three summers to be same as results from the research NAPS trailer, since this NAPS station is considerably farther from the highway. It is also interesting to see that for 2015, the maximum median [O_3] on weekends (50 ppb) from the regular NAPS station was lower than weekends median [O_3] (54 ppb) from the highway NAPS research trailer. Both stations show differences of 20 – 24 ppb of maximum median [O_3] between weekdays and weekends.

We recognize we do not have enough weekend data to be statistically robust, so as suggested, we have included this discussion above in the supplemental material, and in Fig. S7. We also have added two sentences to the main text in Section 3.4 to summarize these results and thoughts. “To evaluate how representative the contrasts between weekends and weekdays based on this 16-day data set are compared to longer timeframes, 3 summers of O_3 measurements from a nearby NAPS station in Toronto West were extracted and analyzed (Supplementary material Section 3 and Fig. S7). Similar diurnal patterns and differences were observed in 2 of the 3 years, suggesting that the analysis presented above is representative of longer terms as well.”

13. P15. If the gas phase mechanism in the GEM-MACH model does not explicitly represent HCHO then it shouldn't be portrayed as HCHO in Figure 6, that is somewhat misleading if one doesn't read the fine print. What other compounds are included with HCHO, methacrolein and methyl vinyl ketone? If this is the case then I suggest leaving out the model data in Fig 6 for "HCHO".

Methacrolein is partially included in the HCHO lumped model species. As this reviewer suggested, HCHO GEM-MACH results have been removed from Fig. 6.

14. P15. The HCN section is very brief. Any idea why it is so variable; most data appear below DL of instrument except for 3 days at the end of the campaign. If HCN is from vehicle exhaust why isn't it elevated when CO was elevated? It is hard to tell from the figure, but it doesn't seem to follow CO.

As the reviewer noted, most HCN mixing ratios were below detection limit. HCN also did not follow CO, since it was below 1 ppb for most of the period we measured. The variable HCN mixing ratio from FTIR spectrum analysis is probably due to operating so close to its detection limit, which makes it hard to fit HCN's absorption features among much more abundant interfering gases. However, indications are that we did detect a real HCN signal for at least a few short periods. We are very interested in HCN because it has adverse health effects and our colleagues have observed and modeled HCN in emissions from vehicles in the past (Moussa et al., 2016). One sentence has been added to the text in this section: "Only on July 28, 29 and 30th, the HCN was observed above its detection limit".

15. P15. I don't understand the reasoning behind the statement " ...flat on weekends, indicating that a large component of CH₃OH may have come from traffic emissions". Methanol doesn't covary with CO from examination of the figures. I find it hard to believe all CH₃OH in an urban area is due to vehicles. What are other sources of methanol? As far as I know methanol is not included as a compound in vehicle emission inventories by the US EPA but perhaps this is different in Canada? Trees emit methanol. You would probably measure similar levels of methanol outside of the Toronto urban area as a result. Are there other urban sources of methanol that are relevant, solvent use for example?

The reasoning here is comparing the diurnal cycles on weekdays and weekends; weekends did not have the early morning peak which corresponds to rush hour traffic on weekdays, so the early morning peak of CH₃OH on weekdays was probably from traffic emissions. In addition, in the early morning on a few weekdays, we found linear relationship between CO and CH₃OH, and the slopes have been shown in Table 2 to estimate emission rate of CH₃OH from traffic. The linear relationship in the early morning on weekdays also suggests at least some CH₃OH was from traffic emissions.

Besides the two references in the manuscript, there are more studies on urban VOC fluxes. Rantala et al. (2016) studied urban VOC fluxes in urban Helsinki and found methanol fluxes were correlated with traffic and with CO fluxes. Traffic could partially explain the observed methanol. Sahu and Saxena (2015) also reported CH₃OH mixing ratios at Ahmedabad (an urban site in India), and both traffic emission and the transport from biomass burning and biogenic sources outside the city contributed to

CH₃OH. Reyes et al. (2006) reported vehicle emission of non-regulated pollutants, including methanol, by using local gasoline and driving conditions in Mexico City.

Other sources of methanol include biogenic sources, such as trees and forest plants, residential wood combustion, and biomass burning. Our measurement could not directly distinguish different sources. We found no relationship of [CH₃OH] with ambient temperature (Fig. S6c), and we observed a linear relationship between [CH₃OH] and [CO] in the early morning on some weekdays. Therefore we concluded that at least some of the observed CH₃OH was from the traffic.

We have revised Section 3.6: “As shown in Fig. 6, mixing ratios of CH₃OH from the FTIR were between 2 and 20 ppb most of the time, with some high spikes. Figure 12 presents the corresponding average weekday and weekend diurnal cycles of CH₃OH for the study period. This plot shows the mixing ratio reached a peak (maximum of 20 ppb at 7:30) from 7:00 to 9:00 on weekdays whereas there was no peak in the mornings on weekends. In addition, a linear relationship between [CH₃OH] and [CO] was observed during the early morning rush hours on some weekdays (Table 2). These results suggest that at least a fraction of observed CH₃OH was from traffic emissions. Observations of methanol associated with traffic have been reported in other studies. Rogers et al. (2006) reported CH₃OH in the diluted pipeline exhaust of a mobile laboratory. CH₃OH may also come from non-engine sources, such as windshield wiper fluid. Durant et al. (2010) measured gas and particle pollutants near Interstate 93 in Massachusetts. They reported CH₃OH was above 20 ppb at 7:20 50 m downwind of the highway, possibly with contributions from some other local sources. Reyes et al. (2006) reported vehicle emission of non-regulated pollutants, including methanol, by using local gasoline and driving conditions in Mexico City. Rantala et al. (2016) studied urban VOC fluxes in urban Helsinki and found methanol fluxes were correlated with traffic and with CO fluxes, traffic could partially explain the observed methanol. Sahu and Saxena (2015) also reported CH₃OH mixing ratios at Ahmedabad (an urban site in India), and both traffic emission and the transport from biomass burning and biogenic sources outside the city contributed to CH₃OH. The mixing ratio of CH₃OH we observed did not correlate with ambient temperature (Fig. S6c), so there was no strong indication of biogenic sources.”

Rantala, P., Järvi, L., Taipale, R., Laurila, T.K., Patokoski, J., Kajos, M.K., Kurppa, M., Haapanala, S., Siivola, E., Petäjä, T., Ruuskanen, T.M. and Rinne, J.: Anthropogenic and biogenic influence on VOC fluxes at an urban background site in Helsinki, Finland, *Atmos. Chem. Phys.*, 16, 7981-8007, doi: 10.5194/acp-16-7981-2016, 2016.

Sahu, L.K. and Saxena, P.: High time and mass resolved PTR-TOF-MS measurements of VOCs at an urban site of India during winter: Role of anthropogenic, biomass burning, biogenic and photochemical sources, *Atmos. Res.*, 164-165, 84-94, doi: 10.1016/j.atmosres.2015.04.021, 2015.

Reyes, F., Grutter, M., Jazcilevich, A. and González-Oropeza, R.: Technical Note: Analysis of non-regulated vehicular emissions by extractive FTIR spectrometry: Tests on a hybrid car in Mexico City, *Atmos. Chem. Phys.*, 6, 5339-5346, 2006.