We would like to thank reviewer 1 for his/her comments, we have addressed these in the new version of the manuscript. We addressed each comment below and highlighted our answers in blue, the reviewer's comments are black.

Reviewer 1:

This paper investigated fire NOx emissions using TROPOMI NO2 observations. The authors also explored the impact of aerosol on TROPOMI NO2 observations and thus emissions by comparing against results derived from aircraft measurements. They concluded that a correction factor of 1.3 to 1.5 shall be applied to correct NOx emissions inferred from satellite NO2 observations. I would recommend some revisions before the publication.

General comments:

1. AMF calculation. Is there any specific reason to use both GEOS-Chem and GEM-MACH to calculate AMF? The authors mentioned that free tropospheric NO2 is not well represented in GEM-MACH. If so, does it make more sense to use GEOS-Chem for all layers? I'm worried that the usage of two models will introduce additional uncertainties.

   Only GEOS-Chem is used in the a priori to calculate the AMF, output from the GEM-MACH model is not used for the AMF estimate. Inside the plume, in the boundary layer a constant profile is used, based on the enhancement of the VCDs. Outside the plume simply the GOES-Chem profile is used representing background concentrations. Inside the plume the contribution from GEOS-Chem is actually very small.

   Our colleagues are currently developing the emissions for the free troposphere to input in GEM-MACH and hopefully in the near future we can use the upper tropospheric profile from just the GEM-MACH model. In this study, GEM-MACH is only used for the model sensitivity test of the two emission estimate methods to test if the methods can determine the true emissions (input emissions).

   We included the following in the manuscript (Sect. 3.1) to make this point clearer:

   "The amount in the free troposphere is on the order of 10^{14}molec/cm^2 and small compared to the total tropospheric column inside fire plumes (~10^{16} molec/cm^2). It is better to assume even a small amount of NO2 in the free troposphere when estimating AMFs than assuming 0."

2. Uncertainties of EMG method by assuming constant lamda and sigma. Please clarify the uncertainties in the manuscript.

   We added a little more details and discussion on the sensitivity of the EMG to the lifetime and plume spread in the appendix:

A1. Sensitivity to lifetime and plume spread

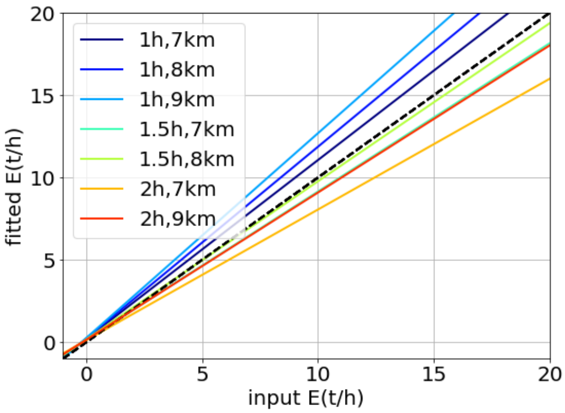

Figure A1. The impact of changing the lifetime and plume spread parameter on the slope of best fit (under a (i) scenario) using the EMG method to obtain the fitted emissions.

Using a different lifetime and plume spread does have an impact on the bias to the true emissions however the correlation is not affected by this. Note that changes in lifetime and plume spread can compensate each other. For the previous cases, discussed in Sect. 4.2 we use a plume spread and lifetime of 6 km and 1 h (note that this does not represent the true chemical lifetime).

Figure A1 shows the variation of the slope of best-fit of the fitted emissions to the true emissions: a lower lifetime will increase the emissions and a lower plume spread decrease the emissions. Thus, the emissions are almost identical when using σ = 9 km, τ= 2 h, and σ= 7 km, τ= 1.5 h. Based on this analysis, the uncertainty is about 25% within the associated spread of lifetimes and plume spreads. This is a major contributor of uncertainty, and thus it is important to find a realistic lifetime to reduce the overall uncertainties of the emissions estimate, which is not always easy.

3. “the difference of the lifetime between the model and the TROPOMI observations are expected, since the chemical lifetime of NO2 is shorter in the model compared to reality.” I’m concerned about the robustness of this conclusion. It’s likely that the winds in the model and reality differ significantly, which also causes the different plume shapes.

   Yes the winds could impact the plume spread and consequently the lifetime (as those go together – we have seen various examples using EMG for point sources that a high wind speed reduces the plume spread, but changes the lifetime to a lesser extend), however it is unlikely to impact a large sample of different locations and days. The winds between ERA5 (this not really measured, but the best we can get for a large variety of locations) and GEM can be different, but they do not have a systematic bias..

   There are a few reasons why the lifetime in GEM-MACH is shorter than in reality. However, this is not part of this study and needs further investigation as to what exactly is going on. Our colleagues are working on improving and developing the GEM-MACH model:

One possibility might be that the NOx cycling is shorter than it should be, since the number of organic reactions in the breakdown sequence is smaller than in reality. Another possible reason could be numerics and advection: if the source of emissions is less than 10 km in extent (the case for a 10km model), then there will be a tendency for the model's advection code to "smooth out" the plumes in the horizontal direction, at least until they grow a bit in size, not just due to the grid cells being much larger than the source size, but due to the inability of the model to resolve spatial gradients. The mass conservation algorithm will tend to "flatten" peaks in concentration.

Specific comments:

1. Page 2, line 31, I suggest reorganizing this paragraph, since the key message is not clear. I’m not sure whether the authors would like to emphasize the advantage or limitation of satellite observations.

   We have changed the paragraph to the following:

   “A few species can be observed by satellite instruments and used to estimate fire emissions. Satellite-remote sensing observations have the advantage of continuous, near global coverage, if meteorological conditions are favorable (e.g., clear-sky) and the emissions are above the instrument's detection limit.

   Ground-based and aircraft measurements are difficult to obtain near the fire source (due to Temporary Flight Restriction zones) and field campaigns are infrequent with limited spatial coverage, while satellite-borne observations can be used to constrain wildfire emissions and can provide emission estimates for fires missed by measurement campaigns.”

2. Page 3, line 13. The 2011 work is based on OMI observations.

   Thank you for noticing this error. We have corrected the text accordingly.

“As satellites improved so did space-borne emission estimates, and in 2011 NOx emissions were derived for the first time on a city-wide scale (Beirle et al., 2011) using observations Ozone Monitoring Instrument (OMI; 2004–present; 13_24 km2; at nadir; Levelt et al., 2006; Krotkov et al., 2016).”

3. Page 3, line 19. Are there any differences between biomass burning investigated by Jin et al. (2021) and wildfire in this study? If not significantly, I would recommend a discussion or comparison with Jin’s work in the manuscript since both studies use TROPOMI NO2 to infer NOx emissions. I notice the authors tries to do the comparison in the introduction by listing the topics covered by both studies. But I would appreciate some descriptions/clarification about differences, because it may be difficult for readers who are not familiar with Jin’ work to understand the differences by just reading the list.

   As suggested, we have included further details in the introduction:

"Recently, TROPOMI-derived NOx emissions have been reported (Jin et al., 2021), which focused on TROPOMI-derived global NOx emissions and NOx emission factors. Our study explores the derivation of top-down NOx emissions from wildfires using TROPOMI NO2 observations and assesses its accuracies, with a focus on (1) the methods used for the emission estimates, (2) the conversion of retrieved NO2 to estimates of NOx, and (3) the explicit aerosol correction, and (4) validation of the TROPOMI-derived emissions using aircraft observations."

4. Page 4, line 5. please correct the typo of " the he".

   Done.

5. Page 4, line 30. What is the resolution after 6 Aug, 2019?

   3.5kmx5.5km. To make it clearer we changed the text to: "(3.5kmx5.5km after August 6, 2019; 3.5kmx7km prior to August 2019)"

6. Page 5, line 17. Is RPRO for the whole year of 2018? Please clarify here.

   We included the dates: "(RPRO; April to November 28, 2018)"

7. Page 5, line 27. What does the under script of EC stand for? Is it Environment Canada? It will give readers the impression that these are the official NO2 products from Environment Canada. Please consider renaming the products if it is not the case and the products are investigational. But this is just my feelings. Other readers may have different opinions about this. I would suggest ask around and make the final decision about the name.

   In previous studies the same terminology has been used. While this is not the official product of Environment Canada, but it's unofficial product of Environment Canada. We decided to keep this name.

8. Page 5, line 33. please correct the typo of ". hourly".

   Done

9. Page 6, line 12. It is not clear to me how the model setup simplifies determining the accuracy of emissions estimation method.

   In the typical set up, the input fire emissions have a prescribed diurnal cycle. If the emissions change every hour it is harder to know the exact input emissions as these are variable and will impact the VCDs downwind of the fire (the VCDs are the result of ~3h of emissions, depending on wind speed and lifetime). If the input emissions are constant for the entire day, we know exactly the true emissions. We included the following in the text to make this point clearer:

“This removes the prescribed diurnal variability (used in the standard model run) and thus simplifies determining the accuracy of the emission estimation methods, as the input emissions are constant and known; concentrations downwind were emitted at the same rate as those close to the fire.”

10. Is there any specific reason for only showing the flight track for AOSR, but not other three campaigns?

    The flight tracks (at least for the flights used in this study) of the other three campaigns are shown in Figs. 9 and 10 together with the TROPOMI observations.

    We have modified the text to point to those figures.

    “Further details, including the flight path are presented in Sect. 5.3 (Fig. 10).”

    “Further details, including the flight path are presented in Sect. 5.2 (Fig. 9).”

11. Figure 4. Please make the sizes of panels consistent.

    The panels have the same size, however, the size on the left is determined by the ratio of the longitudes and latitudes.