Comments from anonymous Referee #1:

We would like to thank the reviewer for his/her helpful comments. We hope that we could address all questions and unclear points satisfactorily.

In the course of the revision, we have made the following important changes to the analysis: We updated our wind and ozone data and are now using ERA5 reanalysis data averaged about the boundary layer instead of a fixed height to include possible seasonality in these input data. This led to changes in all our estimates but not in any fundamental way. We also applied our method to a power plant for which hourly NOx emissions are reported, and analyzed the temporal variability over the course of the year in more detail. From this comparison we conclude that the EMG method applied to TROPOMI data at least for this power plant can reproduce the temporal variability reasonably well and does not show a clear seasonal bias. To get a better representation of source areas from Southern Hemisphere, we included five new target areas. The results fit well into the overall result of the study.

Legend: Referee comments in black, author comments in blue

This manuscript estimates the NOx emissions from cities and large sources around the globe using TROPOMI NO2 satellite data. I appreciate that the manuscript is thorough and well-written.
I am concerned about the derived NO2 lifetimes by season and month. The magnitude of the seasonal cycle of emissions reported here are in disagreement (and sometimes strong disagreement) with previous literature. Please see Crippa et al., 2020 which shows that CO2 emissions (which can be used as a rough surrogate for NOx in this instance) don't have a very strong seasonal cycle. Using data reported in Crippa et al., 2020… CO2 emissions are maybe 10% lower in summer in the US and China as compared to winter. In Spring and Fall, CO2 emissions are lowest, perhaps 15-20% lower than the winter peak. The summer-to-winter ratios reported in this manuscript, such as ~0.5 in New York City, ~0.4 in Chicago, ~0.25 in Wuhan are simply not reasonable. Even in Europe, Crippa et al., report ratios of ~0.8, and the values reported here are significantly lower.
My guess is that the EMG method is having trouble discerning the true NO2 lifetime during winter. This is a known issue and is why previous literature focus on summer time emissions (Lu et al., 2015; Goldberg et al., 2019). To prove whether your method is reasonable during the winter, I suggest two potential strategies: 1.) You can apply the method to a model simulation that has data during both the winter and summer; since the emissions in the model are known, the method should hopefully reproduce the seasonal cycle of the emissions that are input into the model (whatever they are). 2.) You can apply the method to power plants which report their NOx emissions for all seasons such as all the large power generation facilities in the US. This will hopefully provide insight on whether the seasonal effects on the NO2 lifetime and therefore NOx emissions are being correctly accounted for.

We understand the concerns of the Reviewer about the strong seasonality seen for several cities that also was a surprise to us. We checked the literature from Crippa et al. (2020). Especially Figure 7, showing a time series of monthly fossil CO_2 emissions by sector in the world, shows a strong variability over the course of the year. This is a global picture but driven by the top CO_2 emitting countries located in the Northern Hemisphere (China, USA, Europe, Russia). The annual variability is strongly influenced by the power generation and residential combustion sectors, which are showing lower emissions from May to August and higher emission during winter. The residential sector is showing emissions more than three times higher during colder months than during summer. Figure 8 shows more regional results, showing summer to winter ratios from approximately 0.6 for Russia, 0.7 for Europe and 0.85 for the USA, driven by combustion of fuels in the power and residential sectors during cold months. Still, these are very broad areas and it could be expected that results differ on smaller scales (in cities) due to composition of emission sources, policy regulations, and people's behavior.

Nevertheless, we also see the point that the deviations between the seasonality from Crippa et al. (2020) and our results are large and should therefore be questioned. We therefore followed the referee's suggestion and applied our method to a power plant for which hourly NOx emissions are reported, and analyzed the temporal variability over the course of the year in more detail to get more confidence in the method especially during different seasons.

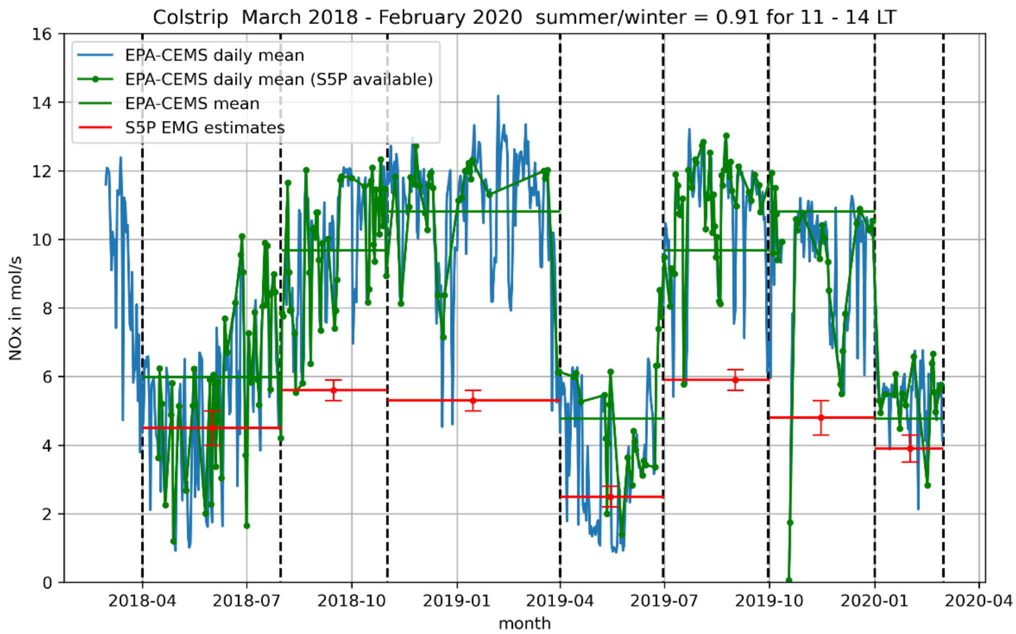

Figure 1: Comparison of CEMS NOx data for the Colstrip (Montana, USA) power plant with emissions estimated by the EMG method, based on TROPOMI data from February 2018 to March 2020. CEMS data are daily mean values, averaged between 11-14 local time, shown in blue and filtered for days with TROPOMI measurements shown in green. Means over periods, marked by the dashed black lines, are shown by the green lines for the CEMS data and red lines for the TROPOMI EMG estimates.

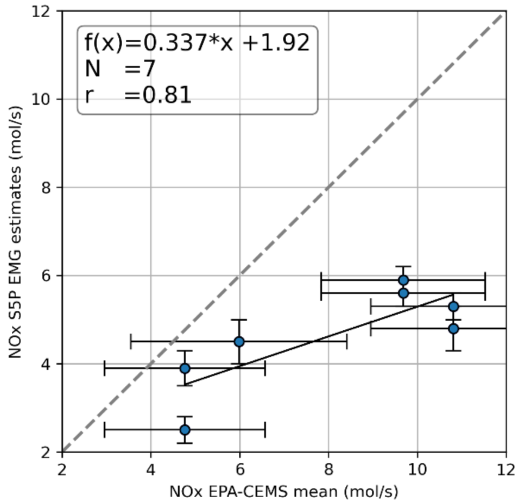

Figure 2: Scatter plot of the CEMS NOx measured emission vs S5P EMG NOx estimates for seven time periods of a range between two to five months during February 2018 to March 2020. The 1:1 line is indicated with the grey dashed line. The solid black line indicates the linear regression, with a slope of 0.34 and a correlation of 0.81.

From this comparison we conclude, that a) power plant emissions can vary strongly over time which was also confirmed by a check of power generation reported by German power plants, b) that the EMG method applied to S5P data at least for this power plant can reproduce the temporal variability reasonably well and does not show a clear seasonal bias, c) that the S5P based NO_2 emissions can be lower by about a factor of 2 for power plants as already discussed in previous publications (Beirle et al., 2019, supplement) and usually explained by a mismatch of the model based a priori profile used and the power plant plume which is not well mixed in the proximity of the source.

With the aforementioned said, I think that Sections 4.1 and 4.4 are useful advancements to the literature. I recommend that Sections 4.2, 4.3 and 4.5 are excluded, and included in a follow-up manuscript that more rigorously evaluates the top-down NO2 lifetime by season.

Since we could reproduce the temporal variability of the NOx emissions of the power plant reported by EPA-CEMS in a satisfying way and also regardless of the season, we decided to keep the Sections 4.2, 4.3 and 4.5 where seasonality is analyzed and add the new data to make things a little clearer.

In response to the reviewer's comment, in addition to the new section, in which we show that the method can reproduce temporal variability over the course of the year, we also added comments in which we highlight potential problems when looking at seasonality and included remarks at several corresponding passages throughout the text. Furthermore, we updated our wind and ozone data and are now using averages about the boundary layer instead of a fixed height to include possible seasonality in these input data. This lead to changes in all our estimates but not in any fundamental way.

All suggestions:

Line 39: in-homogeneous —> heterogeneous

Changed as suggested.

Line 43: Missing comma between winter and residential

Added a comma.

Line 62: states —> countries

Done.

Line 147: Availability of data in March and April?

We added a sentence that the data before 30 April is only available on the S5P Copernicus Expert Hub and therefore not publicly accessible.

Figure 1: Suggestion to remove lat/lon lines on the plot. The lines obscures a few cities.

Lat/lon grid is removed from the plot, which gives more visibility for some sources. Some are still partially obscured by coastlines and borders, which cannot be changed without losing information about country affiliation.

Line 163: Wind direction likely similar between heights, but wind speed can be different, especially if comparing near surface to 1000 m. Maybe something to note.

We followed the suggestion and changed our wind data from 100m to wind speed and direction averaged over the boundary layer, extracted from ERA5 re-analysis model level data. Therefore, we adapted the section about the wind data, and included that the chosen height of wind data is not only critical for wind speed but also for wind direction.

Line 163: Might also want to mentioned ease of use of 100-m wind speed since it is a standard variable output in the ERA5 re-analysis.

This is true, that is why we had chosen the 100-m wind data in the beginning. To exclude the possibility that the strong seasonality in the emissions is due to a lack of seasonality in the wind data (seasonal variability of boundary layer height -> higher/lower wind speed), we decided to use an average now over the boundary layer instead of the fixed height of 100 m.

Line 172: Can you clarify? Do you mean every three hours (0Z, 3Z, etc.)? I originally interpreted three hourly estimates to mean three different outputs each hour.

Thank you for pointing out this unclear formulation. It was meant that every three hours a value is available.

As part of the update to the new wind data, we also updated the ozone data which we now also derive from the ERA5 reanalysis data which have a better spatial resolution and much better temporal resolution. Ozone data are averaged over the boundary layer just as wind data and thereby provide a better representation of seasonal variability.

We now use hourly ozone volume mixing ratios with a horizontal resolution of 0.25° in model levels averaged over the boundary layer and interpolated to the TROPOMI overpass time and oversampled to the same 0.01° resolution as the TROPOMI data.

Line 173: Presumably the interpolation is location specific (Europe may use 12Z model data but China may use 3Z for example). Please clarify.

Yes, that is right, the interpolation is location specific. We updated the section about the ozone data and added that the ozone data are interpolated to the TROPOMI overpass time.

Line 207: earth —> globe

Changed.

Line 213: Modify "low cloud coverage" to "few clouds". "Low cloud" could be interpreted to mean low in the atmosphere.

We changed it accordingly in the text.

Line 318: Verhoelst et al., 2021 and Judd et al. 2020, which you already cite, show a low bias in polluted areas of ~20-30%. I see no problem with using this to scale the final estimates by the value, while also noting the large uncertainty in this conversion factor.

We chose not to scale our results due to the difficulty of determining which region needs scaling and which would be the appropriate factor. Different factors are discussed in the literature and due to the different regions and sources discussed, different factors are likely to be needed. However, we now point out in the corresponding sections that we know about the low bias and also that it may be reduced with the reprocessed TROPOMI data version, which is unfortunately not available yet.

Line 336: Early studies, published in 2018 & 2019 used an algorithm that has since been re-processed. Also there was a different rotation in Goldberg et al., 2019; it was done manually.

We checked the Goldberg et al. (2019), which we used for comparison to our results and found that data of "operational v1 varies from v1.00.01 to v1.01.00" were used, these are the same versions used also in our analysis.

We have added the point about the additional manual rotation to the text.

Line 350: I am not convinced that these ratios should be this small. Traffic pollution and industrial/manufacturing pollution, is relatively constant year-round and represents a large fraction of emissions. If anything traffic and industrial emissions would be biased to be larger in the summer (traffic peaks in July and lowest in February: https://www.fhwa.dot.gov/policyinformation/travel_monitoring/tvt.cfm). NO2 lifetime, in theory, should be varying much more than you found. So I'm thinking the summer/winter ratios are an artifact of an erroneous lifetime fit. See Zheng 2018 as an example. They report that residential emissions in China (the sector causing the most intraseasonal variation in emissions) is ~5% of Chinese emissions. Even if this percentage is off by a factor of two and power genration varies much more than we suspected (Crippa et al. 2020 show this doesn't vary much by season in China), it would be very unlikely to get ratios larger/small than 0.8-1.2.

Thank you for making us aware of the study of Zhen et al. (2018). They showed that residential emissions have only a small contribution to Chinese emissions. However, this is an average value for the whole country and probably does not reflect the specific situation in cities. We checked the seasonality of the residential sector for China in Crippa et al. (2020) shown in Figure 1, which show only a small contribution of the residential sector but strong seasonal variability with more than two times higher CO_2 emissions in winter than in summer. This would of course only be an explanation for a strong seasonal emission ratio when it is a bigger part of emission source composition in Chinese cities than for China as a whole, which would be expected.

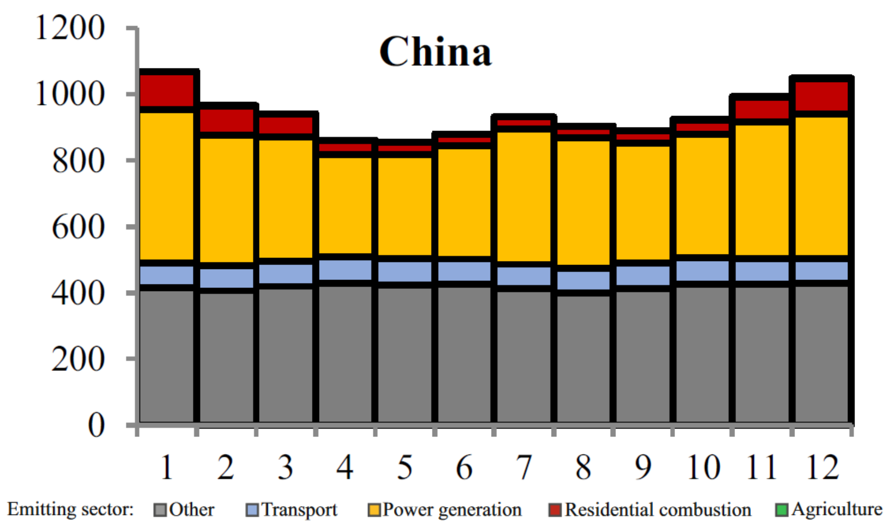

Figure 3 From Crippa et al., 2020 Seasonality of regional fossil CO_2 emissions in 2015 (expressed in Mt/month).

As described above, the updated wind and ozone data and the fact that our method reproduced the variability seen in the CEMS data have given us more confidence in our analysis and the results. Nevertheless, we now point out more strongly in the manuscript, that this is a widely used method, but the first time it is used also extensively on winter data and that there are more uncertainties involved. We have also added that our results to some part differ from previous literature, which however is mainly based on emission inventories, and that further detailed investigations and comparisons would be helpful to understand the discrepancies.

Line 381: I'm glad you are transparent with this, but these don't seem to be reasonable. NO2 lifetime should be shortest in summer, similar in Spring & Fall, and longest in Winter.

Yes, this is true, and in most cases, we do see this behavior, the shortest lifetimes in summer, the longest in winter, with spring and fall in-between and almost similar. Nevertheless, we also see, as in this case for Madrid, that there are deviations from the average and expected behavior. We see that lifetimes are not deviating a lot between seasons and already commented on possible reasons in the text (not yet sufficient statistics, clear-sky bias, midday observation time of S5P, which all lead to more balanced lifetimes). We are now stating this more clearly in the text, especially for the case of Madrid in Figure 5a.

Line 391: Can't you go a step further and estimate OH for a few cities, perhaps from a model or previous literature, and see if the NO2 lifetimes you derive are approximately similar?

Unfortunately, we do not have OH model data available, but we checked previous literature. Lu et al. (2013) measured daily OH concentration maxima of 4-17e6 molec/cm^3 in Yufa, a suburban site south of Beijing, in summer 2006. With the assumption that the decay of NOx is determined by the termolecular reaction of OH with NO_2 (with the rate constant kOH+NO_2+M for 298K) this results in lifetimes of 1.5 h (17e6 molec/cm^3) to 6.3 h (4e6 molec/cm^3).

Dusanter et al. (2019) measured during the MCMA (Mexico City Metropolitan Area) field campaign in March 2006, maximum median OH concentrations of 4.6e6 molec/cm^3 during the day. The maximum OH concentration observed from day to day varied between 2e6 and 15e6 molec/cm^3. On a median basis OH peaked near 4.6e6 molec/cm^3 at noon, which would result in a lifetime of 5.5 h (1.7 h to 12.6 h).

This shows only two examples, one for summer measurements and one closer to winter in two different cities. Dusanter et al. (2019) compared their results to several other campaigns and concluded that OH measurements from winter are rare, that results for summer campaigns indicate that OH measurements are in a relatively small range of concentrations from 2e6 to 9e6 molec/cm^3, except for two campaigns where OH concentrations up to 20e6 molec/cm^3 were observed on some days. It shows how variable the OH concentrations can be between different measurement sites and even within one season.

The lifetimes estimated by OH concentrations from literature are similar to our lifetime estimates. It must be considered that the calculation of lifetime from OH concentrations is only a simplified assumption since NOx is not only lost through oxidation by OH in daytime (although this is the main process, especially during sunny conditions).

Line 415: Often, Saturday is different than Sunday… Saturday emissions can be quite larger than Sunday. See Goldberg et al., 2021 and Crippa et al., 2020 as an example

We added the following:
"Often one weekend day can be different than the other, for example, emissions on Saturdays are often not as low as those on Sundays, even if they are already lower than those from Monday to Friday (Crippa et al., 2020; Goldberg et al., 2021). If the weekend would be limited to one day in the analysis, this would lead to a strengthening of the weekend-to-week ratio for some cities."

Line 420: Also see Goldberg et al., 2021

Added Goldberg et al. (2021) as well as the explanation for deviations due to the different choice of weekend days.

Section 4.5 and Figure 7: If seasonal emissions are biased high in winter, then these drops attributed to COVID will be overestimated. Sections 4.2 & 4.3 need to revised in order for me to have more confidence in the results presented in this section.

In Section 4.5 we focus on comparing only identical months (for example April 2020 with April 2019). The different biases in different seasons should not affect this comparison much, since each month should have the same bias.

Section 4.6: Please more explicitly differentiate between random errors, which should mostly cancel out since you are averaging over many days of observations, and systematic errors, which would not cancel out.

We have changed the text to make it clearer which errors are systematic and which are random.

Line 564: Perhaps may even want to say that re-gridding TROPOMI data to a resolution of 0.1 degrees (~10 km) might be better than a higher resolution (~1 km) for this specific purpose only.

Yes, the Four Corners/San Juan power plant is a special case. We added an additional sentence:
"In this particular case, re-gridding the TROPOMI data to a coarser resolution would potentially allow an analysis"

References:

Crippa, M., Solazzo, E., Huang, G., Guizzardi, D., Koffi, E., Muntean, M., Schieberle, C., Friedrich, R. and Janssens-Maenhout, G.: High resolution temporal profiles in the Emissions Database for Global Atmospheric Research, Sci. Data, 7(1), 121, doi:10.1038/s41597-020-0462-2, 2020.

Goldberg, D. L., Lu, Z., Streets, D. G., De Foy, B., Griffin, D., Mclinden, C. A., Lamsal, L. N., Krotkov, N. A. and Eskes, H.: Enhanced Capabilities of TROPOMI NO2: Estimating NOx from North American Cities and Power Plants, Environ. Sci. Technol., 53(21), doi:10.1021/acs.est.9b04488, 2019.

Goldberg, D. L., Anenberg, S. C., Kerr, G. H., Mohegh, A., Lu, Z. and Streets, D. G.: TROPOMI NO2 in the United States: A detailed look at the annual averages, weekly cycles, effects of temperature, and correlation with surface NO2 concentrations, Earth's Future, e2020EF001665, doi:10.1029/2020EF001665, 2021.

Lu, Z., Streets, D. G., de Foy, B., Lamsal, L. N., Duncan, B. N. and Xing, J.: Emissions of nitrogen oxides from US urban areas: Estimation from Ozone Monitoring Instrument retrievals for 2005-2014, Atmos. Chem. Phys., 15(18), 10367–10383, doi:10.5194/acp-15-10367-2015, 2015.

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C., Geng, G., Li, H., Li, X., Peng, L., Qi, J., Yan, L., Zhang, Y., Zhao, H., Zheng, Y., He, K. and Zhang, Q.: Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions, Atmos. Chem. Phys., 18(19), 14095–14111, doi:10.5194/acp-18-14095-2018, 2018

BEIRLE, Steffen, et al. Pinpointing nitrogen oxide emissions from space. *Science advances*, 2019, 5. Jg., Nr. 11, S. eaax9800.
DUSANTER, S., et al. Measurements of OH and HO 2 concentrations during the MCMA-2006 field campaign–Part 1: Deployment of the Indiana University laser-induced fluorescence instrument. *Atmospheric Chemistry and Physics*, 2009, 9. Jg., Nr. 5, S. 1665-1685.
LU, K. D., et al. Missing OH source in a suburban environment near Beijing: observed and modelled OH and HO 2 concentrations in summer 2006. *Atmospheric Chemistry and Physics*, 2013, 13. Jg., Nr. 2, S. 1057-1080.