Comments from anonymous Referee #2:

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

Review "Variability of nitrogen oxide emission fluxes and lifetimes estimated from Sentinel-5P TROPOMI observations", by Lange et al.

The study by Lange et al. applies the EMG method to TROPOMI NO_2 observations to estimate NOx emissions and lifetimes in various urban areas and sources. They study in detail the variability, seasonality, trends and geographical dependence of the NOx emissions and lifetimes. The method is robust and has been applied to satellite data in the past, with the added value of detailed analysis and other sources being analyzed in this study. The paper is well written and well structured, with quality figures supporting the methods and discussion provided in the text.
Although results and conclusions are presented in a clear way, there are some points that need to be addressed to make the conclusions more substantial and robust. So, I would recommend for publication once the comments below are addressed by the authors.

General comments
In general, the introduction could be improved. Currently the concepts are mixed, and it is not to the point. After reading it, it is not clear what the study is going to be about, and what the research adds to the current literature that is discussed is lacking. It starts with some chemistry, then lifetime, emission estimates, seasonal, day to day, trends, then covid and then back to lifetimes to then introduce the satellite measurements that have already been mentioned at different points in the text before. In the last paragraph you explain what you do in the study but still go back to literature, so not to the point. I think the text and ideas are there, but somewhat disorganized.

Thank you for this feedback. We went through the introduction and structured it more clearly in the different parts and better highlighted what we do in this study. We hope it is now more to the point.

Sect. 4.5 and the discussion on the Covid effects on NOx emissions is vague. You argue that EMG method accounts for meteorology and other possible effects on NOx emissions, but the arguments do not hold equally for the three cities that are discussed and different months. Only months where results are supported by the covid hypothesis are brought forward, which makes conclusions less solid. Many cities have been analyzed in literature to study covid effects on air pollution, so the fact that only three are presented here is also a sign that the method does not work in other cities. If this is not the case, then this should also be stated in the manuscript.

Yes, this is right. We cannot use our method for many cities. We also discussed this in the manuscript. Since the EMG method can only be used to investigate isolated point sources, the range of possible study areas is limited. In addition, we are also much more limited compared to the rest of the study, because we decided to compare only monthly means from the two different years with each other to minimize influence by meteorology. Even if the method takes wind conditions into account, other things like temperature still play a role. In addition, Covid regulations changed quite quickly, and changes can be better detected when shorter time periods are compared.
However, monthly averages bring the disadvantage that sometimes only a few days are available for averaging and analysis due to cloud cover. If only a few days are available in a month, the statistics are simply missing and, for example, weekday effects also have significant influence. This is also the reason why we cannot consider all months even for the cities selected here, and it has to be considered that the statistics may not be ideal. In this context, we would like to point out that some earlier studies use data with less strict quality filtering, including also measurements under cloudy conditions which have intrinsically much larger uncertainties.
We are also discussing months which are not fitting in the expected behavior, as for example for Buenos Aires the June with clearly higher emissions in 2019 than 2020, in this case temperature can be a possible explanation (see also comment below).

The results discussed in other sections may be affected by Covid, but there is no mention to that. How are the summer to winter ratios affected by covid? How is the comparison to EDGAR affected by the supposedly lower emissions when restrictions due to covid where active?

All results with exception of the COVID chapter are based on data from March 2018 to February 2020, basically excluding data for which COVID regulations changed emissions. However, we agree that we should be more careful with our results for Chinese cities since restrictions due to COVID started in Wuhan on 23 January 2020 followed by more Chinese cities and then next northern Italy on 8 March. Therefore, of the regions we analyzed, the cities of Wuhan and Xian in China are potentially already affected by COVID regulations when analyzing data from March 2018 to February 2020. Therefore, we repeated the analysis for these two cities and limited the data period from March 2018 to 22 January 2020.
For Wuhan this resulted in a change of estimated winter emissions from 151.1 $\pm$ 20 mol/s to 254.9 $\pm$ 39 mol/s but of course also to reduced statistic due to a reduction of available days from 16 to 9 days in average over the target area during winter. With this, the summer-to-winter ratio changed from 0.5 to 0.3.
For Xian we could not perform a seasonal and weekday analysis of emission due to insufficient data availability caused by filtering the data for clouds and quality.

Sect. 4.6: The 43-63% uncertainties are higher than the 1-sigma uncertainty provided by the EMG method (e.g. Table 1 uncertainties are ~ 5 %). The high uncertainties need to be reflected in the uncertainty values you provide for the emissions and lifetimes, and discussed in this section.

We commented in Section 3.3 on the fact that our given uncertainties and error bars for emission and lifetime estimates are only based on 1-sigma uncertainties provided by the EMG method and that uncertainties in general are discussed in more detail in section 4.6. We have added the following, to clarify that the uncertainties discussed in section 4.6 are much higher than the 1-sigma from the fit:
"Since the estimates are influenced by additional error sources, dominated by the accuracy of the TROPOMI NO_2 tropospheric vertical column and the wind field, uncertainties in general are significantly higher, ranging from 43-63%. More details can be found in section 4.6."
In addition, we added some small comments to the manuscript when critical results are discussed.

Specific comments

Introduction
Line 101: Is not only urban areas that are the target of the study, right? Same at the beginning of Sect. 5.

Yes, indeed not only urban areas were investigated. For the first case we just deleted the "urban" in the text, for the second case we changed it to "NOx source areas".

Why you do not discuss Beirle et al. (2021) in the Introduction? Is it because those are only point sources?

Thank you for pointing out that Beirle et al. (2021) is missing, we included it now in the newly structured introduction.

Data
Line 154 when referring to Fig. 1: "The red circles mark regions with higher NO_2 than their surroundings and are analyzed in this study." Is this the reason why you choose these regions? Now it reads as it would be like that, but there are many more places with "higher NO_2 than the surroundings".

We agree that this was misleading. We skipped the reference in this section and added the figure to section 3.1., where also the guideline for the selection of targets is described, and changed the text to "we selected 45 target regions, which are marked with red circles".

Wind data and ozone mixing ratio: one is interpolated to the exact overpass time and the other to "typical mean early afternoon overpass time". Why this difference?

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

Line 191: were “new sources” found in Georgoulias et al. (2019)? Or was it positive trends as well as negative ones, that also reversed during the last decades? Could you use the trends from 2015-2017 from this study to extrapolate EDGAR inventory?

You are right, this formulation was a bit misleading. The text now is:
“2015 is the most recent year available in EDGAR v5.0, which cannot reflect either the recent negative as well as positive trends, which were found in trend analysis of NO_2 column satellite data (Georgoulias et al., 2019). Since a large part of the regions analyzed in this study is located in industrialized and highly populated regions, where NO_2 has generally decreased according to Georgoulias et al. (2019), we anticipate that the TROPOMI estimates for the majority of the analyzed regions are lower than the EDGAR inventory estimates for 2015”

Georgoulias et al. found positive as well as negative trends during the last 2 decades (April 1996 – September 2021). Tropospheric NO_2 has generally decreased during the last 2 decades over the industrialized and highly populated regions of the heavily industrialized world and increased over industrially more developing. It is suggested in the reference that linear trends cannot be used efficiently worldwide. Tropospheric NO_2 is very sensitive to socioeconomic changes which may cause either short-term variations or even a reversal of the trends. Different and recent trends would be needed for each city, which is not available.

Sect. 3.1.: why you do not choose any region in Australia? Southern Hemisphere is underrepresented in your selection.

We checked for more possible source areas from Southern Hemisphere and added the following to our study:

Adelaide (Australia)
Brisbane (Australia)
Hwange (Zimbabwe)
Melbourne (Australia)
Perth (Australia)

Pleases mention here the > 2m/s applied to the overall filtering criteria.

Done: “Regions having few clouds are preferred to maximize the number of satellite observations. In addition, a local meteorology for the target region, having rather homogeneous wind patterns, facilities the detection of the outflow patterns, which is further enhanced by filtering out data with wind speeds of less than 2m/s (Beirle et al., 2011).”

Have you looked at boundary layer height information? 100 m is probably too low, so you might introduce a systematic error here.

We changed our wind data to wind speed and direction averaged in 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 can also be critical for wind direction. 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 over the boundary layer instead of the 100 m data.

Results and discussion

Sect. 4.1 & Sect. 4.2:

How do your summer estimates agree with Goldberg et al. (2019)?

We added results for the same time period (April to September 2018) in brackets to the table and some additional comments to the text. Using the same time periods gives a better agreement with Goldberg et al. (2019), deviations are much smaller and could be caused by different wind and ozone data which have quite a big influence. An additional difference is that data are manually rotated in Goldberg et al. (2019) while we use ERA5 wind data for this purpose.

Are there any temperature anomalies in the areas you study that could point to an over/under estimation of the emission inventories based on the threshold temperatures assumed for heating/air conditioning?

Unfortunately, we do not have enough information about temperature anomalies in the different areas and for the relevant years (2015,2018-2020), as well as how exactly heating and cooling degree days are considered for the individual countries in EDGAR, to analyze this.

Sect. 4.3: Lifetimes need to be validated in order to discard unrealistic effects on the NOx emissions and ratios shown in Sect. 4.2. Also, in view of the last paragraph of Sect .4.4. Or at least make clear that the uncertainties associated to the lifetimes is much higher than 1-sigma given by the method, which results in important uncertainties in the emission estimates.

We already compare our lifetimes to existing studies in the manuscript. In response to the reviewer's comment, we now discuss this more and also put more emphasis on the uncertainties in the lifetime estimates and classify what that means for our results.

Sect.4.4: line 420: what if you only take into account summer data as in Goldberg et al. 2019?

We found out that weekend in Goldberg et al. was defined only as Sunday and we defined weekend as Saturday and Sunday which explains the big discrepancies between Goldberg et al. (2019) and our results as Goldberg et al. (2021) and Crippa et al. (2020) showed that emissions on Saturdays are often not as low as those on Sundays, even if they are already lower than from Monday to Friday.

Sect. 4.5: line 469: this can be known by looking at meteorology anomalies in this period.

We checked ERA5 re-analysis temperature data averaged over the boundary layer in the Buenos Aires target area for the measurement days used in our study of the two years 2019 and 2020 for the months May, June and July. The month of June 2020 was 3°C colder than the year before. For May and July 2019 and 2020 the temperatures are very comparable. The colder temperatures in June 2020 than in June 2019 are still not a clear causal factor for the unexpectedly higher emissions in June 2020 compared to June 2019 despite the COVID pandemic but support this possible explanation. We added a comment to the text.

	Mean temperature (K)	
	2019	2020
May	288.4 K	288.3 K
June	**285.7 K**	**282.8 K**
July	281.8 K	281.2 K

Line 475: two years cannot be considered a “trend”. Has the economy changed so much that is noticeable in emissions?

Agreed. We have now formulated this more carefully:

“In January and February, the calculated NOx emissions are higher in 2020 than in 2019. There is no impact of Covid-19 yet in this period, and India's fast-growing economy is probably the explanation for the upward trend in NOx emissions.”
Is changed to:
“There is no impact of Covid-19 yet in this period, and the higher emissions in 2020 compared to the year before match the fact that India has a fast-growing economy. Georgoulias et al. (2019) showed a positive trend of 3.1$\pm$0.5 % per year for the period April 1996 to September 2017.”

Line 483: This would indicate that the lockdown was much stricter in May compared to October, so these estimations might be affected by other factors rather than covid, which could point to the lifetime that is estimated being too low.

We assume you meant it the other way round, because we found higher reductions in October than in May.
The strictest lockdown for Madrid was from 28 March to 12 April, then some restrictions were lifted from 13 April to 1 May, and in May the government followed a plan for easing lockdown restrictions back to normal. On 1 October the government ordered a partial lockdown for Madrid again. It therefore might be the case that the lockdown was stricter in October than in May but of course, this is difficult to quantify.
It is difficult to exclude that other factors besides Covid have had an influence. The influence of seasonal lifetime issues should be small as we look at relative changes between the same months only (October 2020 to October 2019 and May 2020 to May 2019, respectively). Statements about absolute values should be viewed with more caution, which is why we do not make any.

What is the explanation for higher emissions in Feb. 2020 compared to 2019?

Do you still refer to Madrid, there we see it the other way round than in your comment, but we assume you meant it, because this strong deviation between the two months without the influence of regulations due to COVID are striking as they are not expected. We think that the significantly higher emissions in February 2019 compared to 2020 are mainly caused by two factors, first a strong synoptic meteorological variability in Europe and second February is also a month with typically persistent cloud cover, resulting in a reduced number of TROPOMI observations which will reveal more natural variability. Similar findings are also described in Bauwens et al. (2020) and Levelt et al. (2021).

Technical corrections

Throughout the manuscript, please modify “nitrogen oxide” by “nitrogen oxides” and “emission” by “emissions” when necessary.

We checked our manuscript for this and changed it accordingly.

NOx is plural, so NOx are emitted (e.g. page 1, line 15).

We checked our manuscript for this and changed it accordingly.

Please define acronyms first time they are mentioned (e.g. abstract lines 1-2 TROPOMI, NOx)

Done.

Abstract: “is ECMWF ERA 5” relevant? Just “wind data” could be sufficient, as there is no need to go into details in the abstract.

We deleted ECMWF ERA5 from the abstract.

Line 90: “but analyses with GEOS-Chem…”: reader does not necessarily know what GEOSChem is, so please specify.

Thank you for the comment, we have added in the text that it is a chemical transport model.

Line 92: do you mean “measurements of NO_2” instead of “measurements of NOx”?

Changed NOx to NO_2.

Line 116: “(“ missing

Corrected.

Line 152: “removes part of the scenes” -> “removes the scenes”

Thank you for pointing out, this is misleading, the sentence was corrected to:
“A qa_value of 0.75 removes problematic retrievals, errors, partially snow/ice covered scenes and measurements with cloud radiance fractions of more than 50%, which roughly corresponds to a geometric cloud fraction of 0.2.”

Line 215: facilities - > facilitates

Done.

References:

BAUWENS, M., et al. Impact of coronavirus outbreak on NO2 pollution assessed using TROPOMI and OMI observations. *Geophysical Research Letters*, 2020, 47. Jg., Nr. 11, S. e2020GL087978.

GEORGOULIAS, Aristeidis K., et al. Trends and trend reversal detection in 2 decades of tropospheric NO 2 satellite observations. *Atmospheric Chemistry and Physics*, 2019, 19. Jg., Nr. 9, S. 6269-6294.

GOLDBERG, Daniel L., et al. Enhanced Capabilities of TROPOMI NO2: Estimating NO X from North American Cities and Power Plants. *Environmental science & technology*, 2019, 53. Jg., Nr. 21, S. 12594-12601.

LEVELT, Pieternel F., et al. Air quality impacts of COVID-19 lockdown measures detected from space using high spatial resolution observations of multiple trace gases from Sentinel-5P/TROPOMI. *Atmospheric Chemistry and Physics Discussions*, 2021, S. 1-53.