Dear Editor,

we have revised our manuscript on “A new method for estimating megacity NO_x emissions and lifetimes from satellite observations”, taking into account the comments raised by the two reviewers during discussion phase.

The main change compared to the initial submission is that we now compare our seasonal emission estimates to the respective seasonal EDGAR emissions, which we compiled from monthly emissions provided by EDGAR. Interestingly, this did not improve the overall correlation of the comparison. While our results (TROPOMI based emission/lifetime estimates) have not changed, the comparison to EDGAR (correlation coefficient and ratio of means) is now slightly different due to the comparison performed on seasonal EDGAR data.

We now also present the seasonal cycle of emissions for 9 cities where our method works for all season. However, these 9 examples do not allow us to draw clear conclusions.

In addition, we have revised the manuscript according to suggestions and questions raised by the reviewers, as detailed in the one-to-one-response to the reviewers (attached below) as well as marked in the tracked-changes version of the revised manuscript.

Note that in contrast to the statement of intent we gave during initial submission, we did not change the tables and figures showing seasonal results to hemispheric seasons, since we would consider this to be potentially more confusing than clarifying. Instead, we clarified the definition we applied for seasons in section 3.2, and (following the recommendation given by reviewer 2) we repeat the definition of seasons in each table and figure caption or label, respectively.

Kind regards,

Steffen Beirle

> The authors present a new method for estimating NOx emissions and effective lifetimes from large cities. They combined observations for calm as well as two opposite wind directions to shorten the period of data needing for estimation. The method is sound, and the paper is well written. I would recommend minor revisions before publication.

We would like to thank the reviewer for the postive evaluation of our study and for the valuable feedback. Below we refer to the reviewer's comments one-by-one.

> General comment

> 1. Additional details for the fitting procedure would be helpful for readers to follow up. For instance, Line 113, "E(x) has the same unit as L(x) (amount per length unit) and corresponds to the line density that would be observed if no wind transport would occur. " Does it indicate that E(x)=L(x) when w=0? Does this condition be used for the fit? If not, in Line 120, "the distribution of emission densities E(x), lifetime τ and backgrounds b are fitted simultaneously." It looks like too many parameters, each item in E(x), are required to be fit here. How it could be achieved?

We thank the reviewer for the in-depth discussion of our method. Concerning the raised questions:

- E(x) is the distribution of emission as function of x. In case of w being exactly zero, the *calculated* line density L(x) would indeed equal E(x), if there would be no additional smoothing (sigma=0) nor background (b=0). The *measured* line densities for calm wind conditions (blue lines in Fig. 4) differ from E(x) (red lines) due to the additional smoothing (sigma=7 km), additional background b, and wind speeds for calm conditions not being exactly 0, but below 2 m/s.
- E(x), tau and b are fitted to match the observed line densities for forward, backward and calm wind conditions at the same time. The number of parameters (tau, b_calm, b_forward, b_backward, and each item in E(x) (29), i.e. 33 fit parameters in total) are outnumbered by the simultaneous fit of 3 different wind conditions (up to 3 x 29 = 87 contraining data points), resulting in an overdetermined system allowing for a least-squares fit.

In the revised manuscript, we have clarified this procedure by providing more detailed explanations about e.g. the meaning of E(x) and the number of fit parameters vs. number of measurements used for the fit.

> 2. Seasonal analysis. The big improvement of the method is to shorten the required data period from annual to seasonal. A seasonal analysis of derived emissions will help to clarify this improvement.

The seasonal results for fitted emissions and lifetimes are provided in Table 1. As there are seasons missing for most cities, a combined seasonal analysis (like the mean seasonality of all cities, or cities within a latitude range) is not meaningful. Thus, in the revised manuscript, we present and discuss the seasonality of emissions for those 9 cities where the fit worked for all seasons. This is complemented by EDGAR emission data which is now processed on seasonal basis as well. However, the resulting seasonal cycles differ with respect to amplitude and patterns amongst the cities as well as between TROPOMI and EDGAR and do not allow for clear conclusions.

> 3. Section 5.3.1. I'm wondering how much the bias of TROPOMI NO2 over urban areas contribute to the lower NOx emissions?

For v1 of the TROPOMI NO2 column densities, a low-bias was reported over urban areas (e.g. Judd et al., 2020, https://doi.org/10.5194/amt-13-6113-2020; Lange et al., 2023, https://doi.org/10.5194/amt-16-1357-2023). In the PAL dataset, based on the

NO2 processor version v2.3.1, many retrieval steps have been improved. In particular an improved cloud algorithm has been used, resulting in higher cloud altitudes, thus lower air-mass factors and higher NO2 columns. Thus, the reported low-bias is expected to be reduced. However, any remaining bias in the NO2 column would directly affect the presented (as well as any other TROPOMI-based) emission estimate. In the revised manuscript, we have extended section 5.2 accordingly.

> Specific comment

> 1. Line 74. Too many full stops here?

Lines 74-77 contain a direct quote, where the triple dot is meant to indicate an omission. We will clarify with Copernicus how this should be formated properly.

> The original ERA5 data is 0.3 degree, but the intermediate meteorological dataset is 1 degree. I suppose additional errors have been introduced by this interpolation. A case study using both the original and interpolated datasets for a few cities would be useful to quantify the level of uncertainties.

We agree that 1° is quite coarse, and we are currently preparing the implementation of high-resolution wind data for future studies. So far, we applied hi-res wind data for the calculation of advection of SO2 fluxes (Adrian Jost et al., in preparation) based on the algorithm developed for NOx (Beirle et al., 2023). For most locations, the resulting emission estimates only change slightly for high vs. low resolution wind fields, but for some it can be significant (up to 20%). On average, emission estimates for high resolution wind fields are higher by about 5%. This systematic difference can be understood, as the analysis considers the wind speed w_model in wind direction alpha_model, while the actual transport in direction alpha_model is w_true * cos(alpha_model-alpha_true). I.e. an error in input wind data (systematic or statistic) results in an underestimation of the actual wind speed, thus an overestimation of tau, and an underestimation of E.

In the revised manuscript, we discuss the impact of uncertainties of wind fields in detail within section 5.2 about errors and limitations. In addition, we add the aspect of horizontal wind fields taken at one selected altitude as general simplification of our method to the conclusions.

> 1. Figure 2. The period used for averaging is missing in the caption.

We have added the respective time period to the figure caption.

> 2. Line 133. How many fitted lifetimes are not plausible? A ratio of the failed values is helpful for readers to understand the results.

We thank the reviewer for raising this issue.

From the 11200 combinations (700 cities, 4 seasons, 4 axes), a fit was performed for 2154 cases with sufficient data coverage and appropriate mean wind speeds. After applying the filter for chi2, 1862 valid fit results remain. From these results, 151/34 fits are discarded due to too short/too long lifetime.

While checking the number of skipped cases, we noticed that a further selection criterium was not described in the manuscript so far: Cities are skipped if interfering emissions (E(x) for |x| > 50 km) are found.

In the revised manuscript, we have revised and extended section 3.6 accordingly.

> 3. Figure 5. The period used for comparison is missing in the caption.

We have added the respective time period to the figure caption.

> 4. Section 5.3.1. It is surprised to see the large power plants are missing from TROPOMI NO2 maps. Could you confirm the locations of power plants by Google map to support the guess that emissions are shift?

We thank the reviewer for this suggestion. We had a closer look to the locations of enhanced NOx emissions in EDGAR shown in Fig. 7, and could indeed identify potential NOx sources at the respective coordinates:

- The enhanced emissions West from Algiers correspond to the Hadjret en Nouss gas-fired combined cycle power plant at 2.0797°E, 36.5767°N, with a capacity of 1.2 GW.
- The enhanced emissions South from Chelybinsk correspond to power plants in the cities Yuzhnouralsk (54.4552°N, 61.2358°E, https://www.gem.wiki/Yuzhnouralskaya_GRES_power_station) and Troitsk (54.0361°N, 61.6512°E, https://www.gem.wiki/Troitskaya_GRES_power_station).

It is unclear to us why these power plants are not visible in the TROPOMI NO2 data, while several other power plants show up as point sources in Algeria as well as in Russia (Beirle et al., 2023). This finding might indicate that those facilities are not running on full capacity for the last years.

Due to these findings, however, we cannot any longer support the hypothesis that the observed mismatch is just caused by a simple displacement of point sources in EDGAR, and skip this part in the revised manuscript.

> This paper by Beirle and Wagner presents a new method for estimating NOx emissions and lifetimes for 100 large cities from TROPOMI observations. The estimate is based on the downwind plume evolution for different wind directions separately. A simultaneous fit of downwind patterns for opposing wind directions makes the estimate more robust than previous studies using a single fit.

> Results are compared to EDGAR. Seasonal NOx emissions and lifetimes are provided. The assumption of point-source like cities, the background dependency on wind direction, and the benefits of the described method are discussed.

> The paper is well-written, falls into the scope of AMT and is scientifically relevant. Therefore, I recommend publication in AMT after addressing the following comments.

We would like to thank the reviewer for the postive evaluation of our study and for the valuable feedback. Below we refer to the reviewer's comments one-by-one.

> General comments:

> Line 91: You wrote: "As the NO2 photolysis rate is driven by the SZA, which shows a seasonality with minimum and maximum close to the solstices, seasons are defined accordingly as winter (NDJ), spring (FMA), summer (MJJ) and autumn (ASO) in this study." This definition of seasons is not in line with the meteorological and more common definition of seasons as winter (DJF), spring (MAM), summer (JJA), and autumn (SON) and might cause confusion when comparing the presented results with other studies. It may lead to wrong seasonality in the presented emissions and lifetimes. I understand that adapting the seasons to the meteorological and more common definition is probably a lot of work and not easily possible. Therefore, I would recommend pointing out your definition of seasons at important points. This would hopefully avoid confusion and inconsistency in future comparisons for readers not reading the manuscript in detail.

We agree that the unusual definition of seasons in our study is potentially confusing. In order to clarify this point, we further motivate our definition of seasons in subsection 3.2 in the revised document. In addition, we repeat our definition of seasons within all figure captions or labels.

> It is possible to do your analysis on a seasonal basis, even if for many cities only one or two seasons are possible to analyze. But is it also necessary to do the analysis on a seasonal basis, as wind directions and speed have a seasonal dependence, or would your "Selection and averaging of fit results", that is described in Sect. 3.6, consider this? Please comment on this.

The TROPOMI data is sorted for wind direction before averaging. For each season, the TROPOMI pixels are averaged separately for each wind condition. The subsequent fits are based on the seasonal mean TROPOMI columns combined with the respective seasonal mean wind properties. Thus potential seasonal dependencies of wind conditions are accounted for in the gridding & fitting process.

In the revised manuscript, we have now explicitely specified that seasonal means are calculated for NO2 **and wind fields** in section 3.2.

> If you are not separating the data set into seasons, you wouldn't need to work on such long-term temporal averages (Line 231).

This is true, but our intention was to check for potential seasonalities in emissions and/or lifetimes, whis is now explicitely stated in section 3.2. Considering total means might be misleading in cases where some seasons (in particular winter for Northern mid-latitudes) are not contributing valid measurements. Thus we would like to stick to the seasonal analysis.

> Specific comments:

> Abstract: Add that your study is based on satellite data, maybe even more specifically, 3.5 years of TROPOMI NO2 data.

We have added the following information to the first sentence of the abstract:

"…based on NO2 measurements from TROPOMI (PAL dataset, May 2018-November 2021)."

> Line 65: There is a newer, consistently reprocessed data product than the used v2.3.1. Can you comment on how the latest reprocessed data product v2.4 may influence your results?

To our understanding, the processor version 2.3 is a major improvement compared to v1.x, as cloud altitudes are now more realistic. Thus, the reported low-bias of NO2 columns (e.g. Judd et al., 2020, https://doi.org/10.5194/amt-13-6113-2020; Lange et al., 2023, https://doi.org/10.5194/amt-16-1357-2023) is expected to be reduced. This aspect is now discussed in section 5.2 of the revised manuscript.

Compared to this major improvement, the updated from v2.3 to v2.4 have only minor impact on troposheric NO2 columns, thus we would expect that the emission estimates based on v2.4 would be similar to the results presented in this study.

> Line 66: How are the good viewing conditions determined? Have you tested for larger SZA? You mentioned that this may influence the lifetime estimates.

This study is based on the same data processing as in Beirle et al., 2023. Therein, the quite strict selection criteria were applied for two reasons:

- SZA: NO2 is scaled to NOx based on a parameterization of the photolysis frequency as function of SZA proposed by Dickerson et al., 1982, which is recommended to be applied for SZA < 65°.
- VZA: Large VZAs (> 56°) are skipped in order to avoid the large ground pixels at the TROPOMI swath edges.

In the revised manuscript, we provide additional information about why the angles where selected that strictly for this study.

> Line 69: Are ozone concentrations from ERA5 as well or only the temperature data?

Ozone concentrations were not taken from ERA5. In the revised, we avoid this potential misunderstanding by changing the order of the listing of temperature (ERA5) and ozone.

> Line 76: You wrote ERA5 wind data were interpolated on a regular horizontal grid with a resolution of 1° and stored in intervals of 6 hours. This means the closest interval to the TROPOMI overpass is chosen? Can it still represent the wind conditions around the TROPOMI overpass? ERA5 data are available in intervals of one hour. Is horizontal variability a problem when interpolating the original 0.25° resolution of ERA5 to 1° resolution? This is significantly larger than TROPOMI pixel sizes. The study would probably benefit from using finer temporal and spatial grids.

We have compiled an intermediate ERA5 dataset on 1° grid and 6 hour sampling. During data processing, these intermediate wind fields are interpolated in space and time matching to the TROPOMI measurements. We have clarified this procedure in the revised manuscript.

We agree that 1° is quite coarse, and we are currently preparing the implementation of high-resolution wind data for future studies (compare the reply to reviewer 1 on this issue).

In addition, we add the aspect of horizontal wind fields taken at one selected altitude as general simplification of our method to the conclusions.

> Fig. 2/3: Please add the used data period. Are these figures based on data from May 2018 to November 2021 or a specific season?

As specified in the caption, the figures are showing seasonal means for winter (Fig. 2) and summer (Fig. 3). In order to clarify, we have modified the caption of Fig. 2 to “Mean NOx distribution for Riyadh depending on wind conditions for winter months (NDJ) in the PAL period (May 2018-November 2021)”, and Fig. 3 accordingly.

> Some wind directions show distinct outflow patterns (Riyadh: North-West-wards and South-East-wards), some do not (Riyadh: East-wards) or are not very pronounced. Can you comment on this?

We thank the reviewer for this very detailed investigation of the presented data. The East-ward transport for Riyadh in winter, as shown in Fig. 2, is indeed less pronounced than for other wind directions. This corresponds to the findings shown in Fig. 4 (left): The observed line density for West to East (straight green line) cannot be reproduced well by the forward model (dashed green line). Actually, the reduced chi2 for this wind axis is quite high (4.8) such that this wind axis is skipped from the calculation of mean emissions and lifetimes for Riyadh in winter (see section 3.6). In the revised manuscript, we now motivate the chi2 selection criterium in section 3.6 for the Riyadh case.

The reason for the poor match of the fitted line densities is probably that ERA5 wind speeds and/or directions are not matching well to the actual transport for the observations sorted into the West-to-East direction.

> How is the data availability? Is there a way to provide information on data availability for the different wind sectors?

We considered the question how to best provide information on data availability, but since the processing is quite complex, there are different levels of availability (available data per wind sector, available wind conditions per axis, available fit results per city, available mean results per season) and it is difficult to condense this information e.g. in a table. In the revised manuscript in section 3.6, further information is now provided on the net data reduction due to the different selection criteria. Concerning individual cities, we would like to refer to the Supplement: In the plots for mean columns and line densities for each city and each season, the different levels of data availability can be assessed. In order to allow for a simple identification of valid fit results, we changed the labels for discarded wind axes to grey in Fig. 4 and the respective figures in the Supplement.

> Line 100: You wrote you consider distances of +/-50km in all directions. I agree that this is enough in the across-wind direction, but is it enough in the outflow direction? For Riyadh, plumes cover more than 50 km in the outflow direction.

We apologize for the confusion of the 50 km vs. 150 km distance:

- The +/-50km are applied for the *calculation of emissions*, i.e. spatial integration.
- For the *fit of the outflow pattern*, larger distances are considered, as stated in line 122.

We have clarified the usage of 50 km for emission estimates vs. larger distances in upwind/downwind direction for the fit in the revised manuscript.

As the line densities are binned with 10 km resolution, the selection of pixels within 150 km actually keeps the bin from -145 km to -135 km as first element, and 135 km to 145 km as last. Thus, we actually considered distances from -145 km to 145 km, which can also be seen in the line density plots (Fig. 4). We have modified this in the revised document.

> Line 104: You wrote "the mean line density is calculated for calm, forward and backward wind direction". Maybe replace with: calculated for calm wind conditions and forward and backward direction for windy conditions. Or replace direction in condition. Are windy conditions considered as >= 2? I think this was not specified yet, maybe obvious, but it is better to add it when introducing the criterion for calm conditions.

We thank the reviewer for this proposal and now use the term "condition" instead of "direction" whenever we refer to the set of forward/backward/calm wind conditions.

Windy conditions are considered as wind speeds >= 2 m/s, which is now explicitely specified in section 3.1.

> Fig. 4: "The line densities for calm (blue), forward (green) and backward (purple) wind directions" See comment above "calm" is not a wind direction.

We changed "calm" into "calm conditions" in the figure caption.

> Line 115 & Fig 4: You wrote: "E(x) represents the spatial density of emissions. It is considered to be the same for all 3 wind conditions." Is this the case, I see that it is sometimes very similar to one of the wind conditions. Which E(x) is plotted in Fig. 4, the result from the simultaneous fit?

E(x) represents the spatial distribution of emissions and is thus considered to be the same for all wind conditions (i). We have clarified this in the extended description of the forward model. For calm conditions, L(x) is generally most similar to E(x), as almost no transport is taking place.

> Line 122: You wrote: "observed line densities within 150 km of the city center". This is not in line with the statement in line 100 that you consider distances of +/-50km in all directions, please clarify. The resulting emission estimates are determined within 100 x 100 km2, but line densities are fitted within 150 km of the city center?

Yes (see also comment/reply above). We have clarified this in the revised manuscript.

> Line 132: You wrote: "Fit results for a wind axis are considered only if at least 2 directions have sufficient data". This is not clear to me. One wind axis, e.g., North-South, only has three wind conditions: calm, forward, and backward. Do you mean the wind conditions, or is it something else?

Yes, meant is wind condition, which is now modified in the manuscript.

> Line 138: You wrote "weighted by the number of contributing directions for each axis". If the weighting considers calm, forward, and backward wind conditions, I would prefer to replace "directions" maybe with "wind conditions". The same is true for the word direction in lines 140 & 141

We changed "directions" into "conditions".

> Fig. 5: Please provide error bars for the TROPOMI based emission estimates.

As discussed in section 5.2, we estimate the uncertainty of the TROPOMI based emission estimates to be about 30%-50%, which is still considerably large. Including these uncertainties as error bars would degrade readability of Fig. 5. Instead, we state the uncertainty of the emission estimates in the figure caption and refer to section 5.2 for details.

> Line 153: You mentioned that for most cities, valid emission estimates could only be derived for 1 or 2 seasons. How do you think this influences this comparison? For most cities, emission estimates are only derived for summer and autumn, but EDGAR considers data from all seasons.

We agree that the comparison of the mean of seasonal values to an annual mean is problematic due to the many gaps. In the revised manuscript, we thus now provide a comparison of seasonal emission estimates, based on the monthly gridded EDGAR NOx emissions. The comparison to seasonal mean EDGAR yields a correlation R=0.72 and a ratio of means of 0.86, which is similar to the former values based on annual mean EDGAR emissions (R=0.76, ratio 0.84). Interestingly, correlation did not improve by accounting for seasonality in EDGAR.

> Do the estimated emissions show seasonality for cities for which estimates are available for all seasons?

In the revised manuscript, we now present and discuss the seasonality of emissions for those 9 cities where the fit worked for all seasons. This is complemented by EDGAR emission data which is now processed on seasonal basis as well. We have extended section 4.1 by a new figure showing the seasonality of emissions. Some cities (e.g. Tripoli) show hardly any seasonality, while for others (e.g. Algiers) the seasonal emission estimates vary considerably.

> See your discussion in Line 200. If higher emissions are expected for winter months, TROPOMI based estimates are low biased because most of them do not consider winter months, and EDGAR does. In your comparison, most cities are close to the 1:1 line.

We agree that the comparison of mean of seasonal means vs. annual means is problematic and could cause a bias in case of gaps, mostly affecting winter. In the revised manuscript, we account for seasonal mean EDGAR emissions, thus this issue is resolved. Interestingly, the ratio of seasonal mean emissions from TROPOMI vs. EDGAR is not much different from that derived for annual means before.

> Line 214: Which seasons are considered for Jeddah, Vishakhapatnam, and Copenhagen? Could this explain the overestimation of EDGAR compared to the TROPOMI based emissions? I think it is not given that ERGAR is wrong (too high) here.

For Jeddah and Vishakhapatnam, emission estimates could be derived only for summer and autumn, respectively. For Copenhagen, summer and autumn are available. With the updated comparison based on seasonal emissions, EDGAR values are still considerably larger than the TROPOMI based estimate (for Jeddah by a factor >5). As the respective fit results look reasonable, and the discrepancy is already evident in the spatial maps of EDGAR emissions vs. mean VCD, we still conclude that for these locations the EDGAR emissions are probably too high. We have updated the discussion in section 5.3 accordingly.

> Line 219: Add mean lifetime with standard deviation.

Done.

> Technical corrections:

> Line 51: I think m/s should be in exponential notation, please check throughout the manuscript

We have formatted units according to Copernicus guidelines in the revised document.

> Line 68: Change Ozone to ozone

Done.

> Line 83: Add a reference/link and add that you use NOx data

We added a link to EDGAR and specify that we use NOx data.

> Line 86: Change ERA-5 to ERA5 to be consistent

Done.

> Line 94: Change Winter into winter.

Done.

> Line 105/6: Change calm to calm wind conditions

Done.

> Line 110: Delete “the patterns”

Done.

> Equation 1: multiplication should be a dot and not x.

Modified accordingly.

> Line 146: Add NOx

Done.

> Data availability: You could add links to used data like EDGAR, ERA5 wind data, ozone data, and the World Cities Database.

We will provide links and references to the external datasets used in this study in the revised manuscript.