Below we reply to the reviewer comments point by point. The reviewer comments are shown in *italic*, and corresponding modifications and citations of the manuscript are quoted.

Referee #1

Liu et al. describes a new approach to inferring NO_x emissions from cities by combining two previously published methods. The first, the 1D MISATEAM approach described in the Liu et al. 2022 reference, is a whole-city mass balance approach that divides space-based NO_2 column data by wind speed and direction and finds the emissions that balance the transport and chemical removal of NO_2. The second, the divergence-based approach described in the Beirle et al. 2019 reference, is also a mass balance approach in essence, but one which is applied to individual grid cells. There the difference between horizontal flux of NO_2 into and out of a grid cell is taken to represent the sum of emissions and sinks in that grid cell, with the sink assumed to be the first-order chemical loss of NO_2. By combining these two methods, this paper is able to use lifetimes and background NO_2 columns derived from the whole-city analysis in the grid-cell level calculation.

This is an interesting evolution of our ability to directly constrain NO_x emissions from space without use of computationally expensive models. The paper generally does a good job of evaluating the veracity of this method using synthetic data, which demonstrates that this method has good skill in recovering known emissions assuming no systematic biases. The uncertainty estimates are reasonable and justified, though I have one suggestion for an additional test. There are a few points that can be strengthened, which I will detail below. However, this is already a strong paper and I recommend publication after the points below are addressed.

Response: We thank Dr. Laughner for the encouraging comments. All comments and suggestions have been considered carefully and addressed below.

- Point 1: the only limitation I saw in the validation with NU-WRF data was that possible systematic biases in the AMF were not tested. If I understood correctly, the synthetic NO_2 columns used in the validation were an integration of the NU-WRF profiles without any AKs from the NO_2 retrieval applied. Thus, this essentially assumes perfect AMFs. We know from Laughner et al. 2016 (https://doi.org/10.5194/acp-16-15247-2016) that AMF biases from the a priori profiles can lead to biases in the emissions and lifetime derived from methods similar to

the 1D MISATEAM approach. I suspect that such biases would be fairly small in this case, as the MINDS NO_2 retrieval used in this study does have reasonably high resolution a priori profiles (0.25 deg). But we also know from Valin et al. 2011 (https://doi.org/10.5194/acp-11-11647-2011) that even at ~25 km, chemical transport models don't capture the full nonlinearity of NO_x chemistry.

I think that there is a straightforward way to test whether any AMF biases present in the NO_2 retrieval are large enough to affect the 2D MISATEAM method. If you were to repeat the test where you derived emission by applying 2D MISATEAM to the synthetic NU-WRF columns, but this time apply MINDS AKs to the NU-WRF profiles rather than doing a simple column integration, then the emissions derived in this test should reflect the impact of an imperfect AMF. By comparing these imperfect AMF emissions against the emissions derived using the NU-WRF columns without AKs (that represent a "perfect" AMF case), that difference should reveal any systematic impact of systematic AMF biases on the 2D MISATEAM emissions.

Response: We acknowledge the limitation of using integrated NU-WRF columns as synthetic satellite observations, specifically the uncertainties arising from AMF errors. The recommendation to apply MINDS AKs to NU-WRF profiles to probe AMF impacts is well-received. While such an in-depth analysis aligns with the insights provided by Laughner et al. (2016) and would be valuable, our current study's focus is to present the optimal performance of 2D MISATEAM using "perfect columns". Given the extensive computational demands of such sensitivity analyses, we are going to perform a comprehensive investigation into the AMF's impact on 2D emissions in future study, following Laughner et al. (2016). We elaborate on the limitation in the revised Section 3.3, as follows:

"AMF accounts for the variable sensitivity of satellite observations to NO_2 at different atmospheric altitudes, which is informed by a priori knowledge of NO_2 vertical distribution as provided by chemical transport models. Laughner et al. (2016) demonstrated that urban NO_x emissions estimated via NO_2 VCDs with daily, high-resolution a priori profiles are considerably higher than those derived from retrievals using coarser resolution profiles. This presents a relevant challenge for our study. To refine the assessment of AMF influences on 2D MISATEAM-derived emissions, future work could include a sensitivity analysis where TROPOMI's AMFs are applied to NU-WRF profiles. This would generate NO_2 columns with AMF biases, which, when used to calculate emissions, can be contrasted with those derived from idealized columns. Such a comparison would reveal the extent to which AMF biases systematically affect the emissions determined by 2D MISATEAM."

- Point 2: there is one sentence at the end of Sect. 3.2 that could use additional justification - "The slopes of the linear regression lines in Fig. 5 decrease from 0.91 in 2019 to 0.85 in 2021. This can be attributed to the long-term trend of decreasing emissions in the US, primarily driven by the downturn trend in vehicular NO_x emissions (McDonald et al., 2018)." My concern is that 2021 may still include effects from the COVID-19 pandemic. In Laughner et al. 2021 (https://doi.org/10.1073/pnas.2109481118), we see that metrics for traffic and commercial flights (globally as well as in Los Angeles and San Francisco specifically), remain well below their Jan 2020 levels at the end of 2020.

If this conclusion (that the 2018 to 2021 decrease in NO_x emissions is part of the long term trend in the US) is an important part of your work, I'd strongly suggest looking at at least the Google mobility trends (https://www.google.com/covid19/mobility/) and possibly state/city level traffic metrics (e.g. CalTrans PEMS, https://pems.dot.ca.gov/) to check if the underlying traffic driving a substantial part of these emissions had returned to pre-pandemic levels to support this conclusion. If this conclusion isn't critical, then I would recommend adding a caveat that it could include some lingering effects of reduced traffic during the pandemic.

Response: We concur with your observation that the observed decrease in NO_x emissions from 2018 to 2021 was not solely caused by the long-term trend in the US. Since the emissions reduction from 2018 to 2021 is not our major conclusion in this study, we have added a caveat in Sect. 3.2 to underscore the possibility that the decrease in NO_x emissions during the mentioned period might be influenced by the lingering effects of reduced traffic due to the pandemic, in addition to the long-term trend, as follows:

"The slopes of the linear regression lines in Fig. 5 decrease from 0.91 in 2019 to 0.85 in 2021. This decline aligns with the long-term trend of decreasing emissions in the US, primarily driven by the downturn trend in vehicular NO_x emissions (McDonald et al., 2018). The reduced slope in 2021 (0.85) relative to 2019 (0.91) may also encapsulate lingering impacts of diminished traffic from the COVID-19 pandemic, since the traffic and commercial flight metrics at the end of 2020 were still substantially lower than their January 2020 levels (Laughner et al., 2021)."

- Point 3: unless I misunderstood, it seems like you should be able to check for closure of emissions between the 1D and 2D MISATEAM results. That is, the emissions which could be output by the 1D MISATEAM algorithm as in Liu et al. 2022 should represent the total city emissions, and so should be approximately equal to the sum of the gridded emissions derived in the 2D MISATEAM approach. In particular, I wonder if this could be a useful quality check to allow you to expect this method to more cities around the world without needing to validate each city with synthetic NU-WRF data. It would be interesting to see if the cities listed in Table S1 that failed NU-WRF validation also have these two emission estimates (from 1D MISATEAM and this method) differ by more than their uncertainty.

Response: We confirm that the 1D and 2D MISATEAM indeed achieve closure. Figure below compares the emissions estimates from both methods. They show a small relative difference of -5% on average. However, while this consistency might suggest a potential quality check, it cannot be used as such. The reason being that the observed consistency across all sources is inherent to both methods, as they are fundamentally based on the mass balance principle. In this way, the cities listed in Table S1 that failed NU-WRF validation demonstrate a similar consistency between the two methods, with a relative difference 1% ± 14% (mean ± STD).

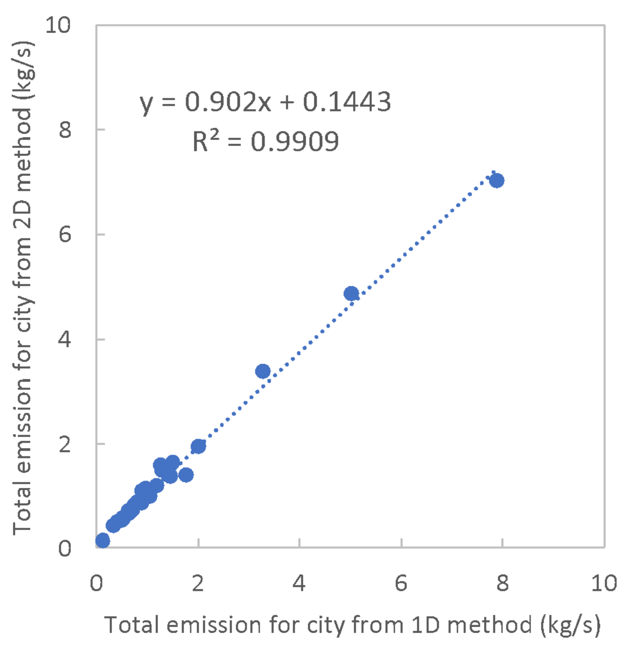

Figure: Scatterplot of the derived NO_x total emissions for the investigated cities based on the NO_2 tropospheric VCDs simulated by NU-WRF using 2D (y axis) as compared to those using 1D (x axis) methods. NO_x emissions from all grid cells within the domain of 70 km × 70 km around city center are summed up to derive the total emission for most cities; a 100 km × 100 km domain is used for New York, Chicago, Los Angeles and Houston.

We clarify this in Section 2.2, as follows:

"1D and 2D MISATEAM methods deliver consistent estimates for total emissions, displaying a small relative difference of -5% ± 9%. While this result demonstrates the internal consistency of the MISATEAM methods, it's important to note that such coherence is due to the shared fundamental principle of mass balance underlying both methodologies. Therefore, the similarity in emissions estimations should not be viewed as an independent validation metric."

- Point 4: it seems like the 2D MISATEAM method implicitly assumes that the background NO_2 is the NO_2 above the boundary layer. Otherwise, it doesn't make sense to me to use only the non-background NO_2 in the calculation of chemical loss (Eq. 3). Is this true? If so, it would be good to explicitly state that assumption.

Response: The 2D MISATEAM methodology does not make implicit assumptions regarding the vertical distribution of background NO_2. Instead, our approach utilizes non-background NO_2 specifically to compute the associated transport and chemical loss, with the primary objective of deriving non-background emissions. We have clarified this in Section 2.2, as follows:

"Note that we subtract the NO_2 background b from $\wedge$ in the calculation of the divergences and the sinks, because we aim to remove the natural and non-local contributions from the total emissions for each urban area in order to infer urban emissions."

- Point 5: I was initially confused by the discussion of the lifetime uncertainty in Sect. 3.3 (lines 221 to 225). The way the uncertainty analysis was presented made me think that the lifetime in Eq. (3) was a single lifetime used for all cities, rather than having unique lifetimes for each city but that does not change in time. On a second read, I found the sentence at line 109 that indicated that the lifetime and background were calculated for each city. Still, it might be good to restate in Sect. 3.3 that the constant lifetime over several years is different for each city. Also, I assume that the reason only 14 cities could be used for the year-by-year lifetime standard deviation in the uncertainty analysis is that they were the only cities with enough good quality data to derive robust lifetimes separately for each year? If so, please state that and list which cities those 14 were. That will be useful documentation in case it is later found that those 14

cities aren't representative of the trend in lifetime for the 39 cities for which emissions were estimated.

Response: Following the recommendation, we have clarified in Section 3.3 that the lifetime is inferred individually for each city. Additionally, we have provided an explanation regarding the selection of the 14 cities for the year-by-year lifetime analysis and highlighted these 14 cities in the revised Table S1. The updated Section 3.3 is as follows:

"Instead of using data spanning multiple years (2018-2021), we apply 1D MISATEAM to annual data to investigate the uncertainty introduced by presuming a consistent NO_x lifetime over several years. For each city, we infer τ by applying 1D MISATEAM to NO_2 VCDs, averaged from May to September for each individual year from 2018 to 2021. Among the cities analyzed, 14 cities (listed in Table S1) have valid NO_x lifetimes available for all four years."