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 #2

This manuscript presents an improved top-down NO_x emission estimate methodology using TROPOMI for select US cities and discusses the method validation and outputs. The improved methodology is the combination of two previously published and widely accepted methods, developed respectively by the two leading coauthors of the manuscript. While I like the manuscript is concise and generally well-written, my main concerns are lack of explanation in some key places and also the lack of details on the derived emissions.

Response: We greatly appreciate your insightful review and the positive remarks regarding the conciseness and clarity of our manuscript. We will expand the sections to provide a more in-depth explanation of our methodology and the derived emissions. These enhancements will address the current gaps in explanation and detail as noted in your comments.

1. *The improved emission mapping algorithm, 2D MISATEAM, is the foundation of the paper. I found Section 2.2 as written does not provide a sufficient justification and motivation for it. Line 93-94 simply states that 2D MISATEAM "is capable of mapping NOx emissions over urban areas.". This statement is not followed by any justification, making it a speculation. The authors need to provide more details on the precedent methodologies, namely 1D MISATEAN and 2D divergency method, as to their respective pros and cons that motivate the development of 2D MISATEAM and how the presented 2D MISATEAM method overcomes the shortcomings of the precedent methods.*

   Response: We appreciate the opportunity to clarify the advancement 2D MISATEAM represents. The 1D approach effectively quantifies total city-level NO_x emissions, treating urban areas as point sources, yet it lacks the ability to provide spatial distribution details. The 2D divergence method improves upon spatial resolution but is dependent on additional sources for NO_x lifetime estimations, either through a constant prescribed lifetime or external data. 2D MISATEAM synergistically combines these two approaches to infer spatial emissions distributions directly and independently, deriving NO_x lifetimes internally from NO_2

measurements and winds info without relying on external lifetime data. We update Section 2.2 accordingly as follows:

"We couple our 1D CTM-Independent SATellite-derived Emission estimation Algorithm for Mixed-sources (MISATEAM; Liu et al., 2022) with the 2D divergence method of Beirle et al. (2019). The 1D MISATEAM quantifies the magnitude of city-level NO_x lifetime and emissions by conceptualizing urban areas as point sources and thus does not capture the spatial variability within the urban areas. Conversely, the 2D divergence method allows for the resolution of finer spatial details in NO_x distributions but relies on additional, often external, sources for determining NO_x lifetimes, which can be a significant limitation. The coupled algorithm (hereafter referred to as 2D MISATEAM for simplicity) leverages the strengths of both: it maps NO_x emissions over urban areas with enhanced spatial detail and does so independently by deriving lifetimes directly from NO_2 measurements, thereby overcoming the need for prescribed or externally sourced lifetime constraints."

2. *The limitation of 2D MISATEAM and its applicability to outside of US should be discussed better. The paper uses the new methodology to large US cities based on population. If the community wants to adopt 2D MISATEAM to other countries/regions which have different population sizes than the US, what should they use to select the suitable places? Does it require that the city has a well-defined urban core with concentrated emissions so that it means certain assumptions in the shape of the urban plumes, etc?*

   Response: Urban NO_x emissions are closely linked to population size, as densely populated areas often have higher fossil fuel consumption, leading to greater NO_x emissions. We thus choose the populations exceeding 200,000 to select US cities, as it is a categorization that corresponds to medium to large urban areas as designated in Organization for Economic Co-operation and Development (OECD) countries. For non-OECD countries, local demographic and urban characteristics may necessitate different population thresholds. However, the application of the 2D MISATEAM is not stringently tied to this population parameter. Following this study, we recommend excluding cities where the background to mean VCD ratio is too high (above 50%) to limit the uncertainties associated with background determination. The urban area's size or the specific shape of the urban plumes are not essential to MISATEAM. We clarify this in Section 4 as follows:

"When applying the 2D MISATEAM to cities globally, particularly in non-OECD countries, it may be necessary to adjust population thresholds to reflect local demographics and urban profiles. We advise removing cities with too weak emissions signals, i.e., b_{calm}/mean VCDs > 50% (Text S2 of the Supplement), as such a high ratio can introduce significant uncertainties associated with determined background b_{calm}."

3. *The paper states that TROPOMI NO_2 columns from May – Sep of each year during 2018-2021 were used to derive top-down emissions. It is not clear what's the temporal time step of 2D MISATEAM when it derives top-down emissions. Does it apply to monthly-averaged TROPOMI NO2 using monthly-averaged winds to derive monthly mean emissions per city, or does it apply to May-Sep mean of those quantities and estimate May-Sep averaged emissions? Also, what determines the temporal resolution suitable for 2D MISATEAM? For example, if one wants to use to derive weekly or even daily emissions, assuming TROPOMI has plenty of good pixels for such a short time period, is there anything assumed within 2D MISATEAM that prevents such application from being successful?*

   Response: The 2D MISATEAM has been applied to May-Sep mean of NO_2 columns and wind data to estimate May-Sep averaged emissions. We have clarified this in Section 2.2 as follows:

   "We average both NO_2 VCDs and reanalysis wind data from May to September each year. We then use those averaged data to infer NO_x emissions E by summing the divergence of the NO_x flux D with the NO_x sink S based on the continuity equation for steady state."

   2D MISATEAM is contingent on 1D MISATEAM for NO_x lifetime estimation, which is a determining factor for the temporal resolution of derived emissions.1D MISATEAM relies on NO_2 observations under calm wind conditions to infer lifetimes. We perform a sensitivity analysis for all US cities investigated in this study, using 1 to 5 months of NU-WRF data. we have identified trends in data coverage. The resulting data, as detailed in the subsequent table, show that a one-month period provides complete data coverage for less than 60% of cities. Coverage improves to approximately 80% with a three-month data span and does not significantly increase with longer data spans. Given that actual satellite observations necessitate cloud filtering, which further reduces data availability, we generally observe that a 3-6 month period of TROPOMI observations is needed to ensure sufficient coverage over urban areas. Therefore, we do not advise the use of 2D MISATEAM for emission estimations over periods shorter than 3 months if we do not want to use prescribed or externally sourced lifetimes. Nonetheless, should lifetime estimates be obtained from external sources, 2D MISATEAM could theoretically be adapted to calculate NO_x emissions over shorter intervals.

Table Perecentage of cities with complete data coverage of NO_2 VCDs under calm-wind situations over the urban areas.

length of data used for averaging	percentage of cities
1 month	58%
2 month	68%
3 month	77%
4 month	80%
5 month	80%

We clarified this limit in the conclusion, as follows:

"2D MISATEAM is contingent on 1D MISATEAM for NO_x lifetime estimation. 1D MISATEAM relies on NO_2 observations under calm wind conditions to infer lifetimes, which in turn influences the temporal resolution of the emissions data we can confidently derive. Our investigation indicates that typically 3 to 6 months of TROPOMI data are required to ensure comprehensive data coverage of calm-wind NO_2 observations for urban emissions analysis. Therefore, we advise caution when considering the use of 2D MISATEAM for emission estimations over periods shorter than three months, unless we want to use prescribed or externally sourced lifetimes."

4. *The manuscript does not provide a good amount of details on the derived top-down emissions over the US cities and how they compare to the NEI inventory. In the abstract, the last sentence states that "there are noticeable differences in the spatial patterns of emissions in some cities" between the TROPOMI-derived and NEI inventory. I don't find where the manuscript elaborated on this main point. Figure 1 is the only place I saw the spatial pattern with a city is presented, but that's only for NYC. To prove the point for "some cities", at least two more cities should be presented. Is the within-city spatial pattern resolution a key strength of 2D MISATEAM? What's the key innovation in it that makes it outperform 1D MISATEAM and 2D divergency method in achieving this?*

Response: We add a new figure (Figure 5) to compare emission maps between the TROPOMI-derived and NEI inventory for two more cities, Dallas and Tucson. We also use the intracity spatial correlation $R_{intracity}$, the correlation coefficient of emissions at grid level over the city domain between TROPOMI-derived emissions E and NEI emissions E_{NEI}, to compare the spatial patterns of both inventories in Figure 4.

We extend the discuss in Section 3.2 as follows:

“Figure 5 compares the NO_x emission patterns from TROPOMI NO_2 with those reported in the NEI, using Dallas and Tucson as case studies. Consistent with observations in New York (as shown in Fig. 1), TROPOMI-derived emission maps reveal several more pronounced point sources as compared to NEI. Notable emissions from the Dallas/Fort Worth International Airport, Perot Field Fort Worth Alliance Airport, and three major cement factories—TXI, Holcim, and Ash Grove—are distinctly evident (Fig. 5a), whereas these details appear diffused in NEI (Fig. 5b). Similarly, emissions from the Asarco Mission copper mine are clearly discernible in the TROPOMI data (Fig. 5c) but are not as apparent in the NEI data (Fig. 5d).

We use $R_{intracity}$ to compare the intracity spatial distribution of emissions for more cities in Fig. 4. We upscale E to the same spatial resolution of E_{NEI} to calculate their $R_{intracity}$ (Fig. S3). $R_{intracity}$ between E and E_{NEI} is 0.57 ± 0.16, which is smaller than that between $E_{NU\text{-}WRF}$ and $\boldsymbol{E'_{NU-WRF}}$ in the evaluation using model data (0.88 ± 0.06; Fig. 2). The generally smaller values of $R_{intracity}$ are likely caused by the uncertainties of both TROPOMI-based and NEI emissions. Compared to $E_{NU\text{-}WRF}$ inferred from perfect NO_2 columns and wind fields, the uncertainties of TROPOMI NO_2 retrievals (25%; van Geffen et al., 2022) and GEOS FP-IT wind reanalysis (30%; Liu et al., 2022) are propagated into the uncertainties of TROPOMI-based emissions E and may result in incorrect spatial patterns. More details about the uncertainties are discussed in Section 3.3. Uncertainties in E_{NEI} also contribute to the disagreement. NEI uses spatial-distribution proxies, such as maps of population densities or road networks, to allocate country-level emissions from non-point sources onto a grid. This procedure may be associated with biases due to either a spatial mismatch between the locations of emissions and spatial proxies or incorrect emission magnitudes. Some hotspots shown in E are missing from E_{NEI}, for instance, JFK airport (Fig.1) and Asarco Mission mine (Fig.5), indicating missing sources or misallocation of sources.**”**

The capacity for within-city spatial pattern resolution is indeed a principal advantage of all 2D methodologies, including both the 2D MISATEAM and divergence methods. The key innovation of the 2D MISATEAM, which enables it to surpass the 1D MISATEAM in this regard, has been detailed previously in our response to your Comment 1.

Minor comments:

- *Line 73: the TROPOMI footprint changes over the study period. Specify the changes.*

  Response: We specify it in the revised manuscript as follows:

  "It has a ground pixel size at nadir of 7.5 km×3.5 km before August 6, 2019, and improved to 5.5 km ×3.5 km afterwards".

- *Line 79: Specify what official product of TROPOMI NO_2 is and provides a reference.*

  Response: We specify it in the revised manuscript as follows:

  "We selected TROPOMI NO_2 retrieved by NASA Goddard Space Flight Center (GSFC), TROPOMI Multi-Decadal Nitrogen Dioxide and Derived Products from Satellites (MINDS) NO_2 product (Lamsal et al., 2022), in this study."

- *Line 88: There is no cloud screening applied? Why?*

  Response: We use the TROPOMI quality assurance value filter (qa_value > 0.75) to remove low-quality observations. This filter removes cloud-covered scenes with cloud radiance fraction > 0.5. We add this in the revised manuscript as follows:

  "Only high-quality pixels with a quality assurance value (qa_value) above 0.75 are considered for averaging, which excludes cloud-covered scenes with cloud radiance fraction > 0.5."

- *Line 245: The uncertainty in the derived NOx emissions is 47%. How does this affect the comparison with NEI? The abstract last sentence attributed all the discrepancy to the NEI. Will the uncertainty in the top-down emissions explain some of the discrepancies?*

  Response: We have discussed the impact of the uncertainties in the TROPOMI-derived emissions on the comparison with NEI in Section 3.2, as follows:

  "We compare TROPOMI-based NO_x emissions E with NEI estimates E_{NEI} for 2019 in Fig. 4…. The relative difference of the total emission between E and E_{NEI} is within the uncertainty range of E (47%; see Section 3.3) for 31 out of 39 cities. The comparison for all cities shows a bias with

NMB of -0.24. The bias is likely associated with uncertainties in the TROPOMI NO_2 retrievals, which have been reported to be biased low by 23% on average (van Geffen et al., 2022).

The comparison of intracity spatial distribution of emissions $R_{intracity}$ shows more disparity in Fig. 4. $R_{intracity}$ between E and E_{NEI} is 0.57 ± 0.16, which is smaller than that between $E_{NU\text{-}WRF}$ and E'_{NU-WRF} in the evaluation using model data (0.88 ± 0.06; Fig. 2). The generally smaller values of $R_{intracity}$ are likely caused by the uncertainties of both TROPOMI-based and NEI emissions. Compared to $E_{NU\text{-}WRF}$ inferred from perfect NO_2 columns and wind fields, the uncertainties of TROPOMI NO_2 retrievals (25%; van Geffen et al., 2022) and GEOS FP-IT wind reanalysis (30%; Liu et al., 2022) are propagated into the uncertainties of TROPOMI-based emissions E and may result in incorrect spatial patterns."

We also revise the last sentence of the abstract to reflect the impact of the uncertainties in the TROPOMI-derived emissions, as follows:

"There are noticeable differences in the spatial patterns of emissions in some cities. Our analysis suggests that uncertainties in TROPOMI-based emissions and potential misallocation of emissions and/or missing sources in bottom-up emission inventories both contribute to these differences."