

Responses to referee 1:

We would like to thank the referee for the useful comments and constructive suggestions. In the following, we address the referee’s comments and describe corresponding changes we have made to the manuscript. The referee’s comments are listed in *italics*, followed by our response in **blue**. New/modified text in the manuscript is in **bold**.

The paper by Lonsdale and Sun presents an interesting analysis of the effects of COVID-19 on NOx emissions in 54 cities over North America, Europe and East Asia, as well as the region- and city-specific features of such effects. The calculation is based on their previously developed “directional derivative approach” which considers the mass conservation under the influence of horizontal advection, emissions, chemical loss and terrain effect. The paper is well written and easy to read. I applaud their efforts to use a new emission inversion algorithm to study the COVID effects on air quality. There are a few major issues on the methodology and results in this paper that should be addressed.

We thank the referee for recognizing the novelty of the directional derivative approach and strive to address the referee’s concerns below.

Based on Eq. 1 (and Sun et al., 2022, GRL), the “wind divergence” term ($\Omega \langle \nabla \cdot \vec{u} \rangle$) is assumed to be equal to vertical flux at $z1$. This assumption might be overly simple. Consider a flat terrain, in which the terrain term becomes zero, Eq. 1 apparently misses the wind divergence.

We thank the referee for bringing up the wind divergence term, which is the most critical difference between the directional derivative approach from this work and the recently established “flux divergence” approach (Beirle et al., 2019, 2021; de Foy and Schauer, 2022; Dix et al., 2022; Liu et al., 2021; Veeffkind et al., 2023). We argue that the wind divergence term is not “missed”, instead, it should be excluded, otherwise the divergence of horizontal wind vector will cause spurious positive/negative emissions which, depending on the value of Ω , interferes with or completely overwhelms the true emission signals. Adopting the referee’s suggestion of assuming flat terrain, and assuming no chemistry to focus on the transport terms, the flux divergence approach includes the wind divergence when calculating the emission:

$$\begin{aligned} \langle E \rangle &= \langle \nabla \cdot (\Omega \vec{u}) \rangle \\ &= \langle \vec{u} \cdot (\nabla \Omega) \rangle + \langle \Omega (\nabla \cdot \vec{u}) \rangle \end{aligned}$$

The consequence of including the wind divergence term (highlighted by red in the equation above) is clearly demonstrated by Fig. 1a in Sun (2022), where large negative emissions are found over coastal land, positive emissions are found over coastal ocean, and the Gulf Stream appears as a strong sink due to its wind convergence. All of those are highly unrealistic and apparently not real sources/sinks. Intuitively, the wind divergence should not correlate with emission. On the contrary, the directional derivative term $\langle \vec{u} \cdot (\nabla \Omega) \rangle$ makes intuitive sense.

The gradient of column amount $\nabla\Omega$ should align with the horizontal wind vector \vec{u} where there is emission, as higher column amount should appear downwind.

We add a new Appendix section to the manuscript to explicitly discuss assumptions. See the response below.

In Sun et al. (2022), there are a few key assumptions for which the validity is not (well) tested. For example, assuming no horizontal gradient of NO₂ above PBL to derive Eq. 3, assuming surface concentration to be derived from the column using a single inverse scale height across a large domain, assuming the wind divergence to be equal to vertical flux at z₁ (Eq. 10), assuming the chemical loss in the PBL to be the same as the loss for the whole column (Eq. 12). All these assumptions are necessary to finally derive the “directional derivative approach”, but these assumptions are subject to large uncertainties and should be tested rigorously to ensure the results for 54 cities here are robust. These tests are especially necessary, because there are few tests in Sun et al. (2022).

We appreciate the referee’s efforts to review the methodology in Sun (2022). To begin with, we would like to make two factual clarifications. First, Eq. 3 in Sun (2022) neglects the horizontal gradient of NO₂ subcolumn above the PBL at the spatial scale of satellite pixels, which is only 3–6 km for TROPOMI. The horizontal gradient above the PBL surely exists, e.g, from dispersion of lightning NO_x or long-range transport of wildfire plumes, but the spatiotemporal scale of mixing in the free troposphere should be much larger than emissions in the PBL. The mixing time scale through the troposphere is ~ 1 month whereas it is ~ 1 day in the PBL (Brasseur and Jacob, 2017). More importantly, these gradients above the PBL happen over large areas and at varying locations and would be largely averaged out at monthly/annual scales in this study.

Second, Eq. 12 in Sun (2022) does not assume the chemical loss in the PBL to be the same as the loss for the whole column. That equation is copied here:

$$\frac{\Omega_b}{\bar{\tau}} = \frac{\Omega}{\tau}.$$

For NO_x, Ω is the tropospheric column, and Ω_d is the subcolumn from surface to z_1 and can be understood as the PBL subcolumn. Consider a relatively clean condition where $\Omega_d = 0.5\Omega$, i.e., the PBL and free troposphere have an equal subcolumn amount, the chemical lifetime over the troposphere, τ , will be twice the value in the PBL, $\bar{\tau}$, i.e., $\tau = 2\bar{\tau}$. This should also be a major reason that τ in this study appears longer than what people unusually see for NO_x lifetime; it represents the lifetime of the tropospheric column, not just the PBL and certainly not at the surface.

To rigorously test individual assumptions and quantify their uncertainties, one would need a high resolution, spatiotemporally comprehensive, and validated CTM simulation. That is beyond the scope of this study and a significant project by itself. However, we would like to emphasize that Sun (2022) already compared NO_x emissions in four cities over the

CONUS (New York, Los Angeles, Dallas, and Houston) with the JPL chemical reanalysis, which includes global 3D transport and chemistry and assimilates many observations (Section 3.3 and Fig. 4 of that paper). We would consider these as more direct tests to the final results—the aggregated city-scale emissions.

We do agree that the assumptions leading to the directional derivative equation and the potential consequences should be more clearly stated. The following is included in the newly added Appendix A to discuss these assumptions and the (often implicit) assumptions in its precursor, the flux divergence approach:

“The flux divergence approach (e.g., Beirle et al., 2019, 2021; de Foy and Schauer, 2022; Dix et al., 2022) is based on the following equation, expressed in terms defined in this work.

$$\begin{aligned}\langle E \rangle &= \langle \nabla \cdot (\Omega \vec{u}) \rangle + \frac{\langle \Omega \rangle}{\tau} \\ &= \langle \vec{u} \cdot (\nabla \Omega) \rangle + \langle \Omega (\nabla \cdot \vec{u}) \rangle + \frac{\langle \Omega \rangle}{\tau}.\end{aligned}$$

Here the second step makes it clearer to compare with the counterpart of the directional derivative approach (i.e., Eq. 1). The key implicit assumptions of the flux divergence approach are discussed below.

1. The emission includes the divergence of horizontal flux and chemical loss. Without the chemical loss, the emission equals the horizontal flux divergence, as shown by studies applying flux divergence to methane (Liu et al., 2021; Veefkind et al., 2023). The problem is that the divergence of horizontal flux is also driven by the divergence of wind ($\nabla \cdot \vec{u}$), which can have positive or negative values climatologically for different locations. This leads to spurious emission values seen in the flux divergence literature that often need empirical correction (Liu et al., 2021; Dix et al., 2022; Veefkind et al., 2023).
2. The topography does not contribute to the flux divergence. In reality, the wind vector usually partially aligns with the gradient of surface altitude even over a long-term average, resulting in terrain-dependent artifacts.

The directional derivative approach (Sun, 2022, this work) addresses these assumptions by explicitly considering the wind divergence and topography effects. The assumptions that lead to the directional derivative approach are detailed in Sun (2022) and discussed below.

1. There exists an altitude z_1 where emissions, as observed by satellites, are confined within. We equate z_1 as the PBL height for ease of conceptualization, but it does not have to be explicitly defined to derive Eq. 1.
2. The horizontal gradient of subcolumn amounts above z_1 is negligible compared to that below z_1 at the spatial scale of adjacent satellite observations.

3. The vertical flux of observed species at z_1 is only due to divergence/convergence of wind below z_1 and is thus not sensitive to emissions. This assumption is a consequence of assumption 1 and the assumption that air flow is incompressible. Conceptually, the upward flux of the observed species at z_1 would not be due to emissions, as z_1 is chosen not to “feel” the emission impact; the only cause of this flux is the convergence of air in the column below that squeezes air upwards or the divergence of air below that draws air downwards.
4. The scale height of the observed species is a constant through the domain. This is necessary to relate the surface concentration to the column amount in the topography term.
5. The column-integrated chemical lifetime of the observed species is a constant through the domain. This is necessary to simplify the chemical loss term, and it is the same for the flux divergence approach.

Assumptions 1–3 are from reasoning. We encourage future testing of these assumptions, presumably through high-quality model simulations. Assumptions 4–5 are apparently significant simplifications. The following two paragraphs discuss their implications.

The scale height is expected to be lower over polluted regions than clean regions. We fit the scale height over rough terrains in each subregion, which are inherently cleaner than the urban areas. Therefore, the scale height applied to urban areas is likely overestimated, and the topography term is hence underestimated as it scales with the inverse of scale height. Fortunately, the urban areas are generally situated over flat terrains. The median value of monthly term $|\langle \Omega \vec{u}_0 \cdot (\nabla z_0) \rangle|$ for all 54 cities averaged in each city is $1.3 \times 10^{-7} \text{ mol m}^{-1} \text{ s}^{-1}$. That means neglecting the topography effect resulting from a 1000 m scale height would only give rise to an emission error of $1.3 \times 10^{-10} \text{ mol m}^{-2} \text{ s}^{-1}$, which is below the noise floor. However, there are two caveats. First, this does not mean that the topography term is unimportant. It might be small over the flat city, but it is large over rough terrains that are close to many cities. Second, some emission sources do appear over rough terrains.

The column-integrated chemical lifetime is a complicated and challenging parameter to obtain. A wide range of values and strategies exist in the literature. Two main factors determine its value, the chemical lifetime within the PBL and the partition of column amounts in the PBL vs. in the free troposphere. The PBL chemical lifetime is highly nonlinear. In the “NO_x-limited” regime, it decreases with increasing NO_x, whereas in the “NO_x-suppressed” regime, the relationship is reversed. The range of variation is within a factor of two (Valin et al., 2013; Laughner and Cohen, 2019). The PBL vs. free troposphere partition may have a larger impact given the high urban-rural column amount contrast and significant

free tropospheric contribution in the clean regions (Silvern et al., 2019). Overall, we expect the column-integrated lifetime determined over relatively clean regions to be higher than the true value over urban areas. This is also consistent with the longer lifetimes shown by Fig. 1 than literature values of urban PBL NO_x lifetime. Consequently, the chemical loss term is likely underestimated in polluted regions.

As such, both topography and chemical loss terms are expected to be underestimated for NO_x over urban areas. This under-correction is preferred to over-correction. Directions of future improvements include using model simulations to inform the spatiotemporal variations of scale height and lifetime and fitting more complex functions (e.g., as polynomial functions of column amount) of the scale height and lifetime. The current constant scale height and lifetime are just the special case of zeroth order polynomial. This will require even higher signal-to-noise ratio, more observations, and/or finer spatial resolution than TROPOMI.”

A few major assumptions are taken in addition. For example, the fixed NO : NO₂ ratio, the fixed lifetime (for monthly climatology and for a large domain), and the fixed scale height (for a large domain). The NO_x chemistry is highly nonlinear and its chemical lifetime varies greatly in space and time. The vertical mixing, PBLH and convection, which are related to the “scale height”, also vary substantially in space and time. How the assumptions will affect your city-specific results need clarification. There exist fast algorithms in the literature that have considered varying NO : NO₂ ratio, as well as concentration-dependent lifetime to account for nonlinear chemistry, and these studies should be discussed and/or compared.

Please see the previous response about scale height and lifetime. Regarding the $\text{NO}_x:\text{NO}_2$ ratio, we have devoted section 3.3 to discuss this issue. Basically, we could not find a reliable, public, and global $\text{NO}_x:\text{NO}_2$ dataset that can match the timeliness and resolution of TROPOMI. This $\text{NO}_x:\text{NO}_2$ ratio will have to be column-integrated, so any error in the PBL/free troposphere partition will propagate to the resultant $\text{NO}_x:\text{NO}_2$ and directly to the emission estimates, even when the $\text{NO}_x:\text{NO}_2$ ratios were individually correct in the PBL and in the free troposphere.

We are not aware of any “fast” algorithms that satisfy our needs for $\text{NO}_x:\text{NO}_2$, chemical lifetime, and scale height. Emission estimation from the directional derivative approach only needs the open-source Python code and TROPOMI and ERA5 data, which are consistent across the globe and throughout the data periods and are available near real time. One does not need a computing cluster to run the algorithm. Adding model simulation to the approach negates its advantage of usability and timeliness, and opens up to more uncertainties due to chemical mechanism, transport, and emission inventory selection. The issue of emission inventories is especially critical due to their time lags and inconsistency among regions.

Therefore, we made choices based on the principle of parsimony. Despite these compromises, the levels of consideration on atmospheric physics and chemistry surpass previous works on

multiple fronts. For example, although the scale height and lifetime are fixed over the subregion and at climatological/calendar months, the spatiotemporal granularity is significantly improved over previous studies. We add another Appendix section in the revised manuscript to compare this work and previous ones:

“Table B1: Considerations of physical and chemical processes by this work and previous studies. The flux divergence and directional derivative approaches are distinct by whether wind divergence is included or excluded.

Study	Wind divergence	Topography	Lifetime	NO _x :NO ₂
Beirle et al. (2019)	Included	None	4 h	1.32
Beirle et al. (2021)	Included	None	None	Photo-stationary state
Dix et al. (2022)	Included	Empirical background correction ^a	Calculated based on OH	1.32
de Foy and Schauer (2022)	Included	None	9 h	1.32
Goldberg et al. (2022)	Included	None	Fitted using EMG ^b	1.32
Chen et al. (2023)	Included	None	Calculated using surface measurements	1.32
Sun (2022)	Excluded	Fitted monthly over the CONUS	Fitted over the CONUS after aggregating 2018–2022	1.32
This work	Excluded	Fitted monthly over subregions with similar climate	Fitted over subregions with similar climate for each climatological month	1.32

^a This may compensate both topography and chemical loss effects. ^b EMG = exponentially modified Gaussian function.”

The resulting scale height is often very high (up to 5 km). It becomes difficult to interpret the exact physical meaning – it is too high for PBLH – and how it can be assumed constant across a large domain without causing major uncertainties.

We thank the reviewer for pointing this out. First, the scale height is not necessarily driven by the PBLH. The scale height for CO derived from this approach is 5–7 km, and the scale height for methane is 7–9 km. When the vertical profile is more evenly distributed across the column, the scale height would approach the atmospheric scale height. Consider a clean PBL with enhanced free tropospheric NO_x due to lightning or long-range transport. The scale height may even be higher than the atmospheric scale height.

Moreover, the PBLH in the southwest US is high and up to 4 km in summer months according to reanalysis. See Fig. 6 for spatial distribution and Fig. A1 for time series averaged over the entire southwest US climate region in Ayazpour et al. (2023). The less latent and more sensible heat flux at the surface in drier areas and the larger surface roughness in high mountain areas should contribute to the higher PBLH.

Furthermore, we second the comment by referee 2: *“My impression is that the height scale and the lifetime are at least partly numerical tuning parameters that have a loose connection with a physical interpretation. This would explain the particularly large values: the beta values are the inverse of the parameters, so large values suggest smaller than expected impact of the corresponding terms in the equations.”* The values are just driven by the data by minimizing “ghost” positive/negative emissions where/when there should not be any. Although the scale height and lifetime in the fitting correspond to physical terms, we would not over-interpret the results. We expand the discussion about this in the second last paragraph in the conclusions and discussion section:

“In this study we fit scale heights at monthly resolution and fit chemical lifetimes for each climatological month to strike a balance between the quality of the fitting results and temporal resolution. However, we assume spatially homogeneous scale heights and chemical lifetimes within each subregion. Considering that the fitting is conducted over cleaner locations where free tropospheric NO_2 subcolumn is expected to take a larger fraction of the tropospheric column, the fitted scale heights and chemical lifetimes are likely overestimated for urban areas. Additionally, the NO_x chemical lifetime is highly nonlinear with respect to NO_x concentration (Valin et al., 2013; Laughner and Cohen, 2019). Therefore, although some aspects of the fitted results are consistent with the expected spatial and temporal variation of PBL height and NO_x chemical lifetime, we caution that the inverses of scale heights and chemical lifetimes are fundamentally linear fitting parameters and suggest against over-interpreting the results. Future investigations might be helpful to achieve higher spatial granularity and/or considering the dependencies of scale height and chemical lifetime on the column amount.”

The derived lifetime is also too high (tens of hours or even more than 100 hours), much longer than what one would expect for a city. What is the exact meaning of this derived lifetime? Why is it OK to use this high value to obtain a reasonable estimate of emission? Since the lifetime for each sub-region and climatological month is constant, what is the impacts on the

derived interannual changes in emissions?

See the previous responses about adding Appendix A to the revised manuscript and about the fitted scale height.

The derived emissions often show very strange seasonal patterns, e.g., emissions in winter could be a few times those in summer. The only anthropogenic source of NO_x that exhibits strong seasonality of residential heating, but this source only contributes a small fraction of total emissions for a city. Better interpretation and explanations of the resulting seasonality should be given.

We thank the reviewer for pointing this out. We expand the discussion in Section 4.1 by adding NO_x:NO₂ ratio as another possible contributor to the seasonality:

“Strong seasonal variations with higher emissions in winter months are observed in some cities (e.g., all cities in the northern subregion, Dallas, Houston, San Diego, and Juarez), which are inconsistent with flatter seasonalities often given by bottom-up emission inventories (Sun et al., 2021). These observed seasonal variations might be caused by seasonally varying artifacts, such as retrieval biases, vertical sensitivity of the retrieval at the surface, and the uncertainties in the wind vectors. In addition, because we use a global constant NO_x:NO₂ ratio, its seasonality that is unaccounted for will propagate to the NO_x emission seasonality. One would expect higher PBL NO_x:NO₂ ratio in winter than summer, but in the summer relatively more NO_x is in the free troposphere, where NO_x:NO₂ ratio is higher than the PBL (Seinfeld and Pandis, 2016). As a result, the exact impact of NO_x:NO₂ ratio on each city is inconclusive. However, we note that no clear seasonality can be identified in Tijuana, whereas the adjacent San Diego shows a much more prominent seasonal pattern. This is inconsistent with the potential impacts by the aforementioned factors, because they should have impacted the estimated city emissions similarly at such a close distance. Moreover, similar seasonalities are not so common in the regions of Europe and East Asia to be shown in the following sections. Further validation of the emission values and seasonality will be the subject of future studies.”

We also add the following sentences to last paragraph in the conclusion to be cautious about the absolute emission data:

“The main focus of this work is the relative emission changes for each city in the pre- and post-COVID-19 years. The absolute emission values of one city compared to another and absolute estimates of emissions month-by-month would be subject to larger uncertainties than the relative values, given the assumptions and simplifications discussed above. We expect future evaluations of spatiotemporal variations of derived emissions against known emission rates of point sources and bottom-up emission inventories.”

There are relatively detailed descriptions of the resulting emissions for each sub-region. How-

ever, it is not clear how the derived emissions are robust for each sub-region and city, given the abovementioned methodological weakness, and the fact that no independent data are used to compare with the emission results. There exist many social indices and other quantitative (proxy) data that can be used to represent the mobility change and lockdown stringency during COVID. Many studies of COVID have used these data, as should be used here.

We did not compare with mobility change/lockdown stringency indices, as we did for Po Valley in Italy in Sun et al. (2021), for two reasons. First, cities included in this study cover a wide range of major NO_x sources, such that we do not think mobility is representative collectively. Additionally, industrial regions and sea ports in many urban areas appear to be much larger than conventional “city” emissions. Second, the emission-derivation approach in this work uses globally consistent observation datasets and builds upon first principle-based equations. The inter-country/inter-continent consistency, data quality, and traceability of proxies and indices are not deemed robust enough to support the robustness of our derived emission. See the previous responses about weaknesses identified by the referee.

The interpretation that the difference from 2019 is the effect of COVID should be cautious. For developing countries, their emissions often change from one year to another greatly, even without COVID, for example, due to economic growth and end-of-pipe control. Additional efforts should be made to assess (quantitatively or qualitatively) the non-COVID factors. There have been many COVID related studies that address this issue, for example, by taking advantage of the pre-COVID trend or variability.

We thank the reviewer for this comment. Practically, the TROPOMI NO_2 record started in May 2018, so consistent data are limited (part of 2018 and 2019) to infer pre-COVID trend or variability. More importantly, COVID-19 has been constantly evolving in the studied regions. Every country and even every city have been impacted differently. Many COVID related studies that have been published focused on limited regions and/or short periods (usually months) and treated COVID as a one-off event. In reality, the COVID-19 impacts linger to this day and continue their diffusion to many aspects of human societies. We do not think it makes much sense to extrapolate the trends from 2019 and before to 2023 and to cut a boundary between “COVID” and “non-COVID” factors for each city. As such, we replace all appearances of “**COVID-19-induced variations/perturbations**” to “**post-COVID-19 variations/perturbations**” We also add the following sentence to the second paragraph of Section 4.2:

“Some of the post-COVID-19 reductions relative to 2019 may extend from a pre-existing decreasing trend, as indicated by consistently higher 2018 emissions than 2019 in some cities in the northern subregion (Fig. 9).”

And add the following to the second paragraph of the Section of conclusions and discussion:

“We caution that the relative differences between post-COVID-19 months in 2020–2023 and the corresponding months in 2019 may exist even without COVID-19. These non-COVID-19 factors include the Military World Games impact in

Wuhan and pre-existing long-term decreasing trend in many cities in the northern Europe subregion, as indicated by higher emission in 2018 than 2019 (Fig. 9).”

References

- Ayazpour, Z., Tao, S., Li, D., Scarino, A. J., Kuehn, R. E., and Sun, K.: Estimates of the spatially complete, observational-data-driven planetary boundary layer height over the contiguous United States, *Atmospheric Measurement Techniques*, 16, 563–580, <https://doi.org/10.5194/amt-16-563-2023>, 2023.
- Beirle, S., Borger, C., Dörner, S., Li, A., Hu, Z., Liu, F., Wang, Y., and Wagner, T.: Pinpointing nitrogen oxide emissions from space, *Science Advances*, 5, eaax9800, <https://doi.org/10.1126/sciadv.aax9800>, 2019.
- Beirle, S., Borger, C., Dörner, S., Eskes, H., Kumar, V., de Laat, A., and Wagner, T.: Catalog of NO_x emissions from point sources as derived from the divergence of the NO₂ flux for TROPOMI, *Earth System Science Data*, 13, 2995–3012, <https://doi.org/10.5194/essd-13-2995-2021>, 2021.
- Brasseur, G. P. and Jacob, D. J.: *Modeling of atmospheric chemistry*, Cambridge University Press, 2017.
- Chen, Y.-C., Chou, C. C. K., Liu, C.-Y., Chi, S.-Y., and Chuang, M.-T.: Evaluation of the nitrogen oxide emission inventory with TROPOMI observations, *Atmospheric Environment*, p. 119639, <https://doi.org/https://doi.org/10.1016/j.atmosenv.2023.119639>, 2023.
- de Foy, B. and Schauer, J. J.: An improved understanding of NO_x emissions in South Asian megacities using TROPOMI NO₂ retrievals, *Environmental Research Letters*, 17, 24006, <https://doi.org/10.1088/1748-9326/ac48b4>, 2022.
- Dix, B., Francoeur, C., Li, M., Serrano-Calvo, R., Levelt, P. F., Veefkind, J. P., McDonald, B. C., and de Gouw, J.: Quantifying NO_x Emissions from U.S. Oil and Gas Production Regions Using TROPOMI NO₂, *ACS Earth and Space Chemistry*, 6, 403–414, <https://doi.org/10.1021/acsearthspacechem.1c00387>, 2022.
- Goldberg, D. L., Harkey, M., de Foy, B., Judd, L., Johnson, J., Yarwood, G., and Holloway, T.: Evaluating NO_x emissions and their effect on O₃ production in Texas using TROPOMI NO₂ and HCHO, *Atmos. Chem. Phys.*, 22, 10875–10900, <https://doi.org/10.5194/acp-22-10875-2022>, 2022.
- Laughner, J. L. and Cohen, R. C.: Direct observation of changing NO_x lifetime in North American cities, *Science*, 366, 723 LP – 727, <https://doi.org/10.1126/science.aax6832>, 2019.

- Liu, M., van der A, R., van Weele, M., Eskes, H., Lu, X., Veefkind, P., de Laat, J., Kong, H., Wang, J., Sun, J., Ding, J., Zhao, Y., and Weng, H.: A New Divergence Method to Quantify Methane Emissions Using Observations of Sentinel-5P TROPOMI, *Geophysical Research Letters*, 48, e2021GL094151, <https://doi.org/https://doi.org/10.1029/2021GL094151>, 2021.
- Seinfeld, J. H. and Pandis, S. N.: *Atmospheric chemistry and physics: from air pollution to climate change*, John Wiley & Sons, 2016.
- Silvern, R. F., Jacob, D. J., Mickley, L. J., Sulprizio, M. P., Travis, K. R., Marais, E. A., Cohen, R. C., Laughner, J. L., Choi, S., Joiner, J., and Lamsal, L. N.: Using satellite observations of tropospheric NO₂ columns to infer long-term trends in US NO_x emissions: the importance of accounting for the free tropospheric NO₂ background, *Atmospheric Chemistry and Physics*, 19, 8863–8878, <https://doi.org/10.5194/acp-19-8863-2019>, 2019.
- Sun, K.: Derivation of Emissions from Satellite-Observed Column Amounts and Its Application to TROPOMI NO₂ and CO Observations, *Geophysical Research Letters*, 49, e2022GL101102, <https://doi.org/https://doi.org/10.1029/2022GL101102>, 2022.
- Sun, K., Li, L., Jagini, S., and Li, D.: A satellite-data-driven framework to rapidly quantify air-basin-scale NO_x emissions and its application to the Po Valley during the COVID-19 pandemic, *Atmospheric Chemistry and Physics*, 21, 13311–13332, <https://doi.org/10.5194/acp-21-13311-2021>, 2021.
- Valin, L. C., Russell, A. R., and Cohen, R. C.: Variations of OH radical in an urban plume inferred from NO₂ column measurements, *Geophysical Research Letters*, 40, 1856–1860, <https://doi.org/10.1002/grl.50267>, 2013.
- Veefkind, J. P., Serrano-Calvo, R., de Gouw, J., Dix, B., Schneising, O., Buchwitz, M., Barré, J., van der A, R. J., Liu, M., and Levelt, P. F.: Widespread Frequent Methane Emissions From the Oil and Gas Industry in the Permian Basin, *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037479, <https://doi.org/https://doi.org/10.1029/2022JD037479>, 2023.

Responses to referee 2:

We would like to thank the referee for the useful comments and constructive suggestions. In the following, we address the referee’s comments and describe corresponding changes we have made to the manuscript. The referee’s comments are listed in *italics*, followed by our response in **blue**. New/modified text in the manuscript is in **bold**.

This manuscript applies the flux divergence method to numerous cities on 3 continents in order to identify the changes in emissions over the 3 years impacted by COVID-19 lockdowns. The method includes refinements on the handling of terrain and lifetime that were reported in a prior publication. A clustering algorithm was applied to show that cities in North America, Europe and Asia had very different annual variability in NO_x emissions over the last 4 years.

I believe that the method is sound and the results are valuable. The paper is clear and well written. I am happy to recommend it for publication.

We appreciate the positive feedback from the referee. Strictly speaking, the method in this work does not calculate flux divergence. Therefore, it is more accurate to refer to the method as the “directional derivative” method. The newly added Appendix B compares this work and previous works applying the flux divergence method in details.

My impression is that the height scale and the lifetime are at least partly numerical tuning parameters that have a loose connection with a physical interpretation. This would explain the particularly large values: the beta values are the inverse of the parameters, so large values suggest smaller than expected impact of the corresponding terms in the equations. I wonder if the question of the parameters would merit some more discussion and caveats in the analysis.

Thanks for this comment. We expand the discussion about this in the second last paragraph in the conclusions and discussion section:

“In this study we fit scale heights at monthly resolution and fit chemical lifetimes for each climatological month to strike a balance between the quality of the fitting results and temporal resolution. However, we assume spatially homogeneous scale heights and chemical lifetimes within each subregion. Considering that the fitting is conducted over cleaner locations where free tropospheric NO₂ subcolumn is expected to take a larger fraction of the tropospheric column, the fitted scale heights and chemical lifetimes are likely overestimated for urban areas. Additionally, the NO_x chemical lifetime is highly nonlinear with respect to NO_x concentration (Valin et al., 2013; Laughner and Cohen, 2019). Therefore, although some aspects of the fitted results are consistent with the expected spatial and temporal variation of PBL height and NO_x chemical lifetime, we caution that the inverses of scale heights and chemical lifetimes are fundamentally linear fitting parameters and suggest against over-interpreting the results. Future investigations might be helpful to achieve higher spatial granularity and/or considering the dependencies of scale height and chemical lifetime on the column amount.”

My other impression is that for each site the time series is robust in a relative sense. However, I think there are probably larger uncertainties in the absolute emission values of one city compared to another and of absolute estimates of emissions in the winter compared with the summer. Because the purpose of the paper is to look at lockdown-induced variability, I don't think this is a major problem. However, I do think it should be discussed to prevent over-interpreting the data. A more detailed comparison of emission totals by city with published emission inventories is beyond the scope of this study, but would be interesting in the future.

We add the following sentences to last paragraph in the conclusion to be cautious about the absolute emission data:

“The main focus of this work is the relative emission changes for each city in the pre- and post-COVID-19 years. The absolute emission values of one city compared to another and absolute estimates of emissions month-by-month would be subject to larger uncertainties than the relative values, given the assumptions and simplifications discussed above. We expect future evaluations of spatiotemporal variations of derived emissions against known emission rates of point sources and bottom-up emission inventories.”

References

- Laughner, J. L. and Cohen, R. C.: Direct observation of changing NO_x lifetime in North American cities, *Science*, 366, 723 LP – 727, <https://doi.org/10.1126/science.aax6832>, 2019.
- Valin, L. C., Russell, A. R., and Cohen, R. C.: Variations of OH radical in an urban plume inferred from NO₂ column measurements, *Geophysical Research Letters*, 40, 1856–1860, <https://doi.org/10.1002/grl.50267>, 2013.