Dear Dr. Cléo Quaresma Dias-Junior ,

we have revised the manuscript, now titled: "Turbulent Enhancement Ratios used for Characterizing Local Emission Sources in a Complex Urban Environment", according to suggestions from both reviewers. We have added new figures and revised the corresponding sections on methodology. Below is a detailed response to both reviewers.

Best regards,

Thomas Karl and Christian Lamprecht

Reviewer 1:

Summary statement:

First, I would like to apologize to the authors for the late turn-in of the review. I hope this report is still of use to you.

This study proposes a statistical quantity termed the 'Turbulent Enhancement Ratio (TER)' to detect and evaluate different scalar sinks and sources of reactive trace gases in urban airflows. The terminology TER is chosen in analogy to a commonly used quantity NER, which has been used in atmospheric chemistry studies when the background concentrations needed to compute excess mixing ratios (EMRs) are unknown. The difference between NER and TER is that NER are computed from slow-response analyzers or time-averaged quantities from fast-response analyzers, while the instrumentation for quantifying the TER can resolve the turbulent motions and hence the variability on shorter timescales. In a first step, TERs from long-term observations are compared against a third quantify termed 'flux ratio' FR for validation, before they are used to study bulk statistics and case studies for the observations in Innsbruck partly dedicated to separating the effects of the anomalous covid-19 lockdown to 'normal' conditions.

I find the current study already has some merit, but to tap into its full potential and merit full publication it requires a much more thorough presentation and discussion of the definitions, similarities, and differences across the statistical flow and flux quantities. Since this journal is concerned with 'techniques', these questions need to be answered unambiguously. Based on the current draft I cannot tell whether the authors are aware that mathematically the TER is identical to the NER, or if it is just a poor explanation/ presentation of the statistics or an oversight. What is correct is that our physical interpretation of these quantities may be different because these quantities may represent different portions of the turbulence spectrum and/ or the mean flow, and hence the processes contained in these statistical quantities may be different. I explicitly say 'may' because the authors do not define the meaning of their triangular brackets usually indicating some spatial or conditional averaging in Eq. 3, and hence I cannot tell if true physical or mathematical differences exist. To me, the TER is rather a spectral similarity ratio rather than a novel quantity separating sink and sources since it is almost identical to the FR, but again, the authors need to improve its explanation. The later part dedicated to bulk statistics and case studies is informative, I have some minor questions about specific statements listed below.

In summary, I believe that the current draft may offer substantial merit after the statistical questions are clarified. The study fits well into the scope of the journal. I recommend reassessment after major revisions.

REPLY:
We thank the reviewer for his valuable assessment. Below we address all comments grouped by their relevance. We fixed some errors that were pointed out. We agree that putting the discussion more in context of spectral analysis will benefit the overall concept description of the paper.

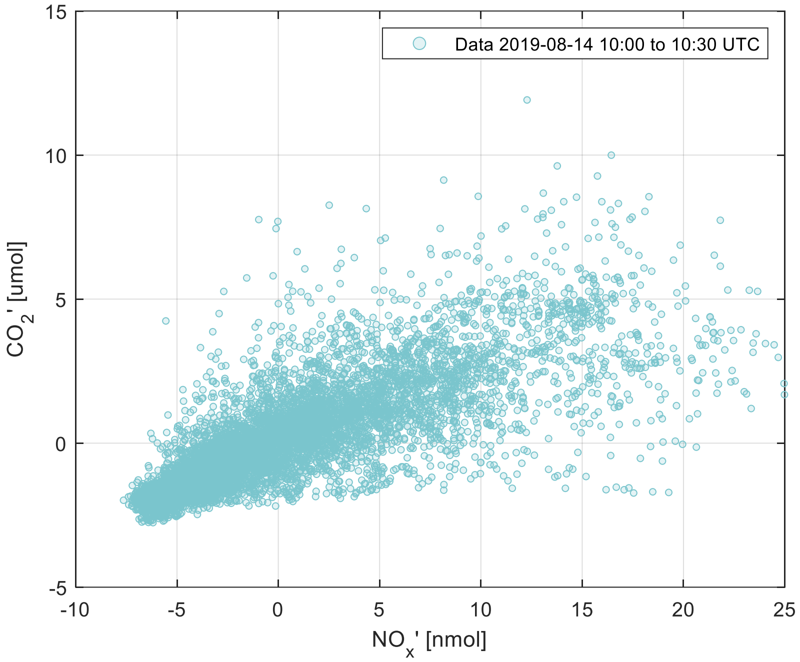
Major comments:

COMMENT 1:
Definition and novelty of the TER: As mentioned above, mathematically the TER in Eq. 4 is identical to the NER in Eq. 3. What remains unclear, why and at what timescales you apply the averaging. Recall that Reynold's first postulate states that the average over all perturbations is zero by definition (I save the time to type this simple equation in the processor), so your triangular brackets cannot mean averaging over the length of the perturbation time scale to derive the perturbations indicated by the primes. So what do they mean? Some physical averaging in the analyzers because it does not capture the full turbulence spectrum down to the Kolmogorov length scale? And what do the triangular brackets mean in Eq. 3? A longer time scale? Please clarify. Deriving Eq.3 from Eq. 2 is not trivial and involves some differential calculus operations, so you need to walk the reader through this process or reference an appropriate source, since this a technical 'techniques' journal and at the heart of your supposedly novel quantity. In addition, it is unclear to me why you claim the validity of Reynold's second postulate to lead from the LHS to the RHS of Eq. 4. This would imply that $\overbar{Y}$ and $\overbar{X}$ are zero, which is difficult to imagine given the supposedly shorter averaging time scales indicated by the triangular bracketing $ \langle \rangle$ (see earlier argument). Only at the averaging time scale indicated by the overbar (i.e. the perturbation time scale), the advective term becomes zero as $\overbar{w} \equiv 0$ because of the rotation. I am confused, please explain all steps and assumptions of the derivation clearly. Similarly, in Eq. 5 you need to explain, if the triangular and overbar averaging are identical to the one used in Eqs. 3 and 4. You may also link your derivation of Eq. 3 to the set of equations for the Relaxed Eddy Accumulation technique, which uses a very similar definition of the b-coefficient as the slope ratio of plotting $w\prime$ versus $c\prime$ in a quadrant analysis plot. Actually, explaining your perturbations and averages using a set of quadrant plots of $w\prime$ versus $CO2\prime$, and $w\prime$ versus $NO_x\prime$, and a scalar-scalar plot of $CO2\prime$ versus $NO_x\prime$' for the different analyzers may be very illustrative to explain the differences.

REPLY +CHANGES:
We thank the reviewer for pointing out this issue. In fact there was a small but consequential typo in equations (3-5) that caused confusion. The brackets should have been round instead of angle brackets. In fact as pointed out by the reviewer, the key of the method is to experimentally resolve turbulent scales fast enough and capture the entire variance. As such mathematically the different approaches are equivalent. We have fixed the equations as following:

$$TER_{X/Y} = \frac{\overline{(\bar{Y}+Y')(\bar{X}+X')}}{\overline{(\bar{Y}+Y')^2}}$$

The angular brackets should have been denoted as round brackets. This should clarify most of the discussion in this context. As pointed out by the reviewer, mathematically NER and TER are identical, just that with TER we are arguing that by resolving the entire turbulent spectrum an unbiased result can be obtained. Similarily, for the flux ratios the co-variance should have not been in brackets. In eq.4 we also identified another typo (the original formula would have represented $TER_{Y/X}$ and not $TER_{X/Y}$, which we also fixed. As noted the analysis of TER in context could be illustrated as the fluctuating parts i.e. $\frac{\overline{Y'X'}}{\overline{(Y')^2}}$. A typical example is shown here:



By attempting to resolve turbulent scales, the approach more likely is similar to true eddy accumulation. Either way we agree with the reviewer, that a more precise and clearer discussion is necessary, which we incorporated in the revised manuscript. We note that the timescale used here was for averaging periods of 30 minutes, which is widely applied in flux calculations, but could potentially be shortened depending on the location. The present data were obtained ~40 m above the street canyon where larger eddies are already influencing the analysis. We now also show typical spectra and ogives (see comment below). In light of these comments we re-interpret TER as the unbiased spectrally resolving enhancement ratio and NER depending on the filter scale, which is subsequently defining what can be resolved. Unlike conventional NERs, which are often derived from ensemble averages (e.g., Parrish et al., 2002; Ehrnsperger and Klemm, 2021), the TER method retains high-frequency turbulent information, making it better suited for complex source separation in dynamically evolving urban environments. This is also now illustrated with a more extensive analysis and new figure, which is shown below.

COMMENT 2, 3 and 5:
Following the comment in A, I think the authors need to include the sampled turbulence spectra for their quantities to make any inferences about which portion of the power-/ cospectrum is resolved and their physical interpretation. Section 4.1: It is difficult to truly understand the very

close to 1:1 relationship of the TER versus FR for NO_x and CO$_2$ without the information requested in comment A. It would suggest that the denominator in the RHS term of Eq. 4 is identical to the denominator in Eq. 5, assuming that the numerators are identical. Hence, I think the TER can rather be interpreted as a spectral similarity ratio rather than a novel quantity representing differences in sink / source. The use of TER reminds me of the triple decomposition (Antonia, R.A., Browne, L.W.B., Bisset, D.K., Fulachier, L., 1987. A description of the organized motion in the turbulent far wake of a cylinder at low Reynolds numbers. J. Fluid Mech. 184, 423–444.) often used in turbulence analysis, which decomposes the excursions from a mean into two different time scales, which are subject to different forcings. I think the authors want to root their statistical quantities in the existing turbulence literature and point out similarities and important differences. Please add the data density isopleths to the plots, these scatter plots with most datapoints overlapping each other centered around the line of unity may give a false representation of the variability. Bars indicating variability (not uncertainty) need to be added to both axes (ordinate and abscissa), I suggest using an bin averaging operator of variable width such that the number of data points included on the a-axis are identical (and hence the standard error defined by $sigma \sqrt{N}^{-1}$), since N varies dramatically across bins because of the uneven pdf. Page 8, line 5ff: not sure what you call the 'bias', but you essentially evaluate the loss of co-variance from 5s to 30min, compared to 0.2s to 30min. Factor of 1/0.43 approx. 2.2 is reasonable. Again, if you show the turbulence cospectra, or even better its cumulative Ogives, then this ratio (and not bias) can be explained.

REPLY:

We appreciate the idea of putting the discussion more in context of an analysis of power spectra and decomposition approaches used in the micrometeorolgical literature (e.g. triple decomposition). Briefly, our urban data are typically biased towards an unstable boundary layer, we also do not detect a major influence of waves (e.g. gravity waves) in the spectra. In addition due to the urban heat island conditions are mostly unstable, even during night we rarely see strongly stratified conditions. In this context we now provide a detailed analysis of power spectra and ogives in context of the importance for resolving parts of the turbulence spectrum related to the inertial subrange. From this analysis we can more accurately infer loss terms due to different scale averaging. For example, spectral damping for 5 min data would represent a loss of 32%. Due to the large roughness in the urban area and high measurement point (42m above street level), the peak in the turbulent spectrum is shifted substantially towards longer scales, which relaxes the requirement of sampling speed quite a bit. Conceptually the spectral analysis though can give an important insight on how fast measurements should be conducted. E.g. more closely to the source one would expect more contributions in the higher spectral range. For the current dataset, we tested the suggested low pass filtering for a range of different values and modify the original figure 2. In this context we found an error in our original analysis, where the regression was conducted between NO and NOx, rather than for NOx vs NOx. We also realized that the bulk regression analysis (shown in the original figure) is not very meaningful, since the correlation between NER and TER is best for background values, but is heavily skewed towards times with high variation in the data. The new figure now plots NER and TER for each half hour showing that the uncertainty of large filter width significantly decreases the accuracy during times of high fluxes, large turbulence and significant covariance signal between NOx and CO2. We are now also more precise about the definition of NER vs TER. We define TER as the filter scale that resolves at least >99.9% of the entire turbulent spectrum, and NER as a filter dependent average. Naturally this will depend on location and turbulence characteristics. Our analysis should give an insight into a real urban test-case on 'how fast is fast enough' for this type of observation.

CHANGES:

We revised the manuscript accordingly, adding new figures of spectra and ogives. We also revised the original Fig. 2 and include references suggested by the reviewer (Antonia, R.A., Browne, L.W.B., Bisset, D.K., Fulachier, L., 1987; C. J. Nappo, Atmospheric Gravity Waves, AP, BOOK citation).

Panel B (Fig. 2) provides a more detailed comparison of TER and NER calculated for different filter timescales ranging from 1s to 1800s throughout the diurnal cycle. The observed pattern aligns well with the expected behavior, displaying background levels during the night time and elevated values during the day, predominantly influenced by traffic emissions. During night time, the agreement between different averaging periods improves as high-frequency fluctuations are sup-pressed, allowing larger-scale eddies to dominate, which are less impacted by low-pass filtering. Additionally, since the tur-bulence measurements are taken at approximately 42 m above ground, they are more influenced by larger eddies, further improving correlation for longer averaging periods. However, when turbulence is strongly pronounced and boundary layer dynamics increase, particularly during the morning and evening, substantial variability is evident, and in some cases, the enhancement ratios may not be accurately represented. This effect becomes especially relevant during critical time periods such as rush hour, where short-term emission peaks play a crucial role in accurately characterizing emission sources. Since NER's ability to capture these variations depends on the averaging interval, excessively long sampling times could lead to an information gap, potentially underestimating peak emissions and their temporal variability. To further investigate the extent to which different averaging intervals influence the agreement between these two metrics, Panel C quantifies this relationship by depicting the coefficient of determination (R^2) as a function of averaging time. The results indicate a strong correlation (R^2 > 0.9) for short averaging intervals up to 60 s, confirming that both metrics capture the turbulent dynamics effectively at these timescales. However, as the averaging period increases, the agreement deteriorates rapidly. For standard air quality monitoring stations (in Austria) that typically employ 600 s (10 min) sampling intervals, most of the turbulence spectrum is already lost, leading to a significant drop in correlation (R2 < 0.4). This suggests that longer averaging intervals may be insufficient to capture the influence of turbulence-driven fluctuations in urban environments, particularly during high-emission periods. The inability to resolve these rapid variations may introduce uncertainties in emission assessments and hinder the accurate representation of short-term pollution dynamics.

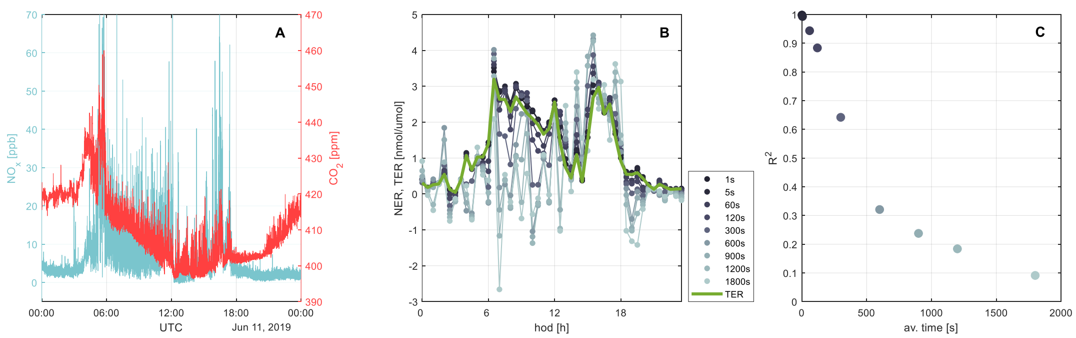


Figure 2: (A) Volume mixing ratios for NO_x (turquoise line) and CO_2 (red line) on June 11, 2019, with a sampling rate of 5 Hz. (B) The TER (green line) and the NER (grey to black lines) for different averaging intervals (1s to 1800s) for each half-hour beginning from midnight. (C) Comparison of the coefficient of determination (R^2) between TER and NER as a function of averaging time

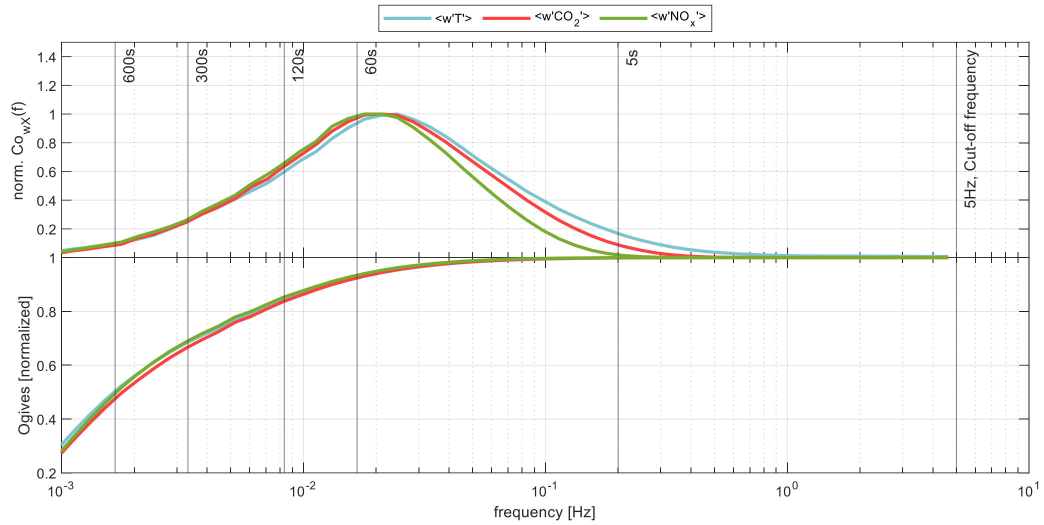


Figure 3: The upper panel shows the normalized co-spectra for sensible heat flux (blue), CO_2 flux (red) and NO_x flux (green) as function of frequency (log-scale). The lower panel shows the corresponding normalized ogives.

Fig. 3 presents an analysis of the frequency-dependent behavior of flux co-spectra and their corresponding cumulative contributions (i.e. ogives) at the field site. The upper panel displays the normalized co-spectra of the sensible heat flux, CO_2 flux, and NO_x flux, with the sensible heat flux serving as a reference, as it is expected to be minimally affected by damping. The peak of the co-spectra typically occur at around 60s for the NO_x flux. The high surface roughness in the urban inertial sublayer and measurement height shift the peak of the turbulent spectrum towards longer timescales, as larger-scale eddies dominate the transport. At higher frequencies, a systematic damping effect is observed for the CO_2 and NO_x fluxes, indica-tive of low-pass filtering effects likely caused by instrumental response and averaging methodology. This phenomenon relaxes the stringent requirements for high-frequency sampling at this location (42 m above ground), as turbulent transport occurs on relatively larger scales.

The lower panel of Fig. 3 summarizes the corresponding ogives, which quantify the cumulative flux contributions across different frequencies. Scale dependent analysis is a common approach in micrometeorology to separate important contributions to the co-spectrum. In this context we interpret TER as a spectral similarity ratio, similar to approaches used to filter data in turbulent flows (Antonia et al., 1987, Nappo, 2012) This provides further insight into the degree of flux loss associated with various averaging periods. The results demonstrate that for averaging intervals down to 60 s, the damping effect is minimal, with only a 7 % loss of the total flux. However, typical air quality observations are often conducted on time-scales between 10 min to 1 h. A NER analysis in these cases would miss > 50 % of the fluctuation, confirming that longer averaging intervals lead to a significant underestimation of fluxes and corresponding TER. Due to the high roughness at the current locations, observations on the order of 5 s would retain 99.9 % of the total spectrum. These findings highlight the critical role of sampling frequency relative to sampling location for accurately resolving turbulent fluxes in urban environments and emphasize the importance of maintaining sufficiently high sampling rates.

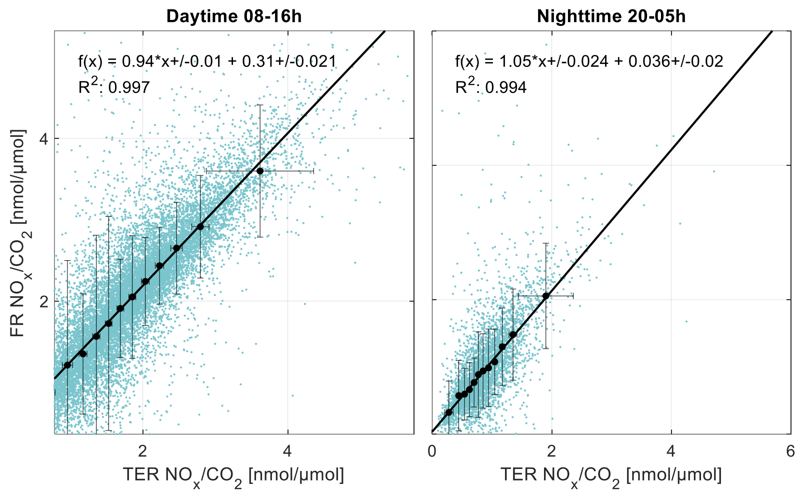


Figure 1: Comparison of TER and FRs for day (a) and nighttime (b). The turquoise points are QA/QC-filtered half-hourly ratios for the entire campaign. Black points (with error bars representing the standard deviation) are the medians binned into intervals with an equal number of data points per bin. The solid line is the regression line for the black points.

According to the reviewer, we redesigned the comparison figure between TER and FR. We added variation bars in both the x- and y-directions and calculated the median for variable bins, as recommended.

Minor detailed comments:

COMMENT 4:
Page 4, line 18ff: Please briefly add the most important EC processing steps, this information is important to understand the behavior of the FR and the results of evaluating TER vs. FR in Section 4.1.Do you mean a correlation coefficient $r\geq$ 0.5 or its magnitude? Please clarify.

REPLY:
Good point. We have extended the sentence where we refer to the paper by Striednig et al. (2020), clarifying the methodology used to calculate the fluxes.

CHANGES
The fluxes used for calculating the flux ratios were determined following the methodology outlined by Striednig et al. (2020) and included Sonic tilt correction, lag time correction, de-trending and despiking. We also clarified the correlation coefficient by adding the R^2 *to the sentence.*

COMMENT 6:
Figure 3: It may be misleading to express the ratio in percent, please use fractions. Since the number of trucks on Sundays is so small and the bar invisible, please use relative scaling in the y-axis.

REPLY:
We have redesigned Figure 4 (E & F) (previous Figure 3) to improve readability. The y-axis has been changed to a logarithmic scale, allowing for a clearer comparison of total traffic load between workdays and weekends while better visualizing the lower number of trucks on

weekends. Additionally, we replaced percentage-based ratios with absolute vehicle counts, which can now be interpreted directly as a fraction (Cars/Trucks).

COMMENT 7:
Page 10, line 22ff: I recommend checking for excursion from common wind patterns when nocturnal winds are up-valley, and daytime winds are down-valley (which must exist) to separate differences in sinks/ sources from their advective distribution.

REPLY:
We have checked this and found the same conclusion that TER in the eastern sector is higher than in the western sector due to differences in surface emissions (a higher fraction of roads in the eastern flux footprint sector). On average TER is about 38% lower in the western sector compared to the eastern sector. We put the figure below also to the supplement and added a hint to the paper.

CHANGES
Figure 4S (Supplement) further supports this finding through a wind-sector-dependent analysis of TER, showing that, on average, TER is 38% lower in the west sector compared to the east sector.

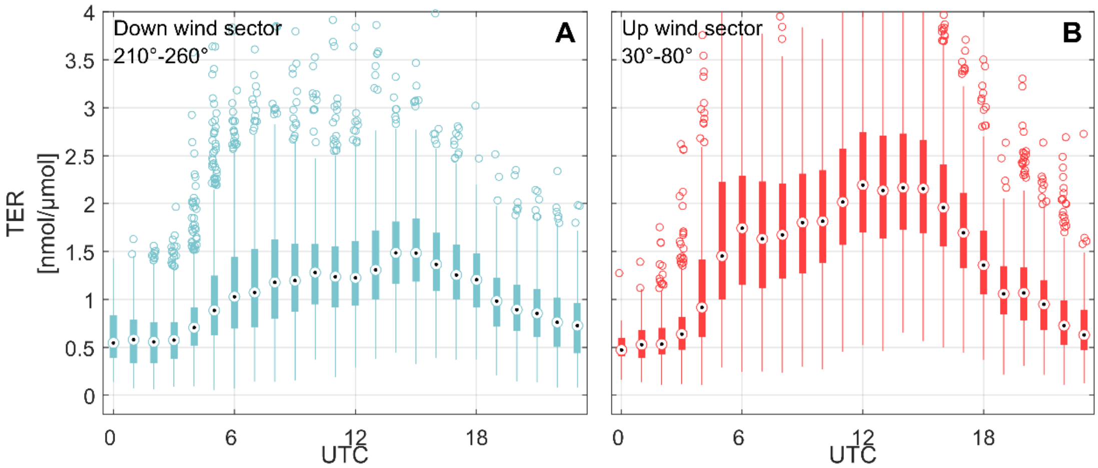


Figure 4S: Boxplot of the diurnal TER on a half-hourly basis for the valley downwind sector (A: 210°–260°) and the valley upwind sector (B: 30°–80°) over the entire campaign duration. Black dots represent median values, circles indicate outliers, and the bars the interquartile range and the overall distribution.

COMMENT 8:
Section 4.3.1: it is unclear to me why F_{CO_2} by depending solely on temperature? The net CO_2 flux integrates over all sinks and sources including plant uptake (photosynthesis), and plant release /respiration) and release from combustion etc. You had mentioned in the introductory section that you see CO_2s uptake by plants during daytime (leading to negative CO_2 fluxes which is surprising given your height well above the buildings), so why is the net flux always positive here? In urban environments it usually is, I am confused. I think this section needs to be improved. Similarly, is the 'heizgrenze' temperature visible only in the transitionary seasons (spring, fall), or also during the summer? Even in the summer the daytime/ daily temperatures may drop down to 10 deg C. This analysis would lend better support to your claimed explanations.

REPLY:

We agree that this section was not clearly formulated and have revised it accordingly. At our measurement site, the net flux primarily results from traffic and residential heating activities. The mention of photosynthetic activity and the resulting CO_2 decrease, as stated on page 7, line 13, refers to the CO_2 concentrations shown in Fig. 2A and not to the CO_2 fluxes. While local processes dominate CO_2 fluxes, larger-scale processes play a crucial role in CO_2 concentrations. In this context, the surrounding vegetation on the mountain slopes along the valley is responsible for the observed CO_2 minimum.

The positive CO_2 fluxes at our flux site can be explained by the dominant emission sources located below the flux tower, meaning that the urban canopy layer acts purely as a source. Consequently, the observed temperature dependence is triggered by methane consumption from gas heating systems, which are the predominant heating source within the flux footprint. As analyzed in Stichaner et al. (2024), methane consumption exhibits a strong temperature dependence below the heating threshold. We have extended this analysis with a detailed investigation of CO_2 fluxes, which follow the same pattern.

Fig. 6B (before Fig. 5B) illustrates the occurrence frequency of the respective mean temperatures in relation to the seasons. Mean temperatures of 10°C occurred a total of four times in summer. In Austria, heating regulations define the official heating period from November 1 to May 30, meaning that central heating systems are generally switched off outside this period, regardless of whether the mean temperature drops below 12°C. During the transitionary seasons the variability is much higher as the heating period with activated heating systems does not totally cover this season.

CHANGES (p12f):

Fig. 6 provides a detailed analysis of the average NO_x and CO_2 fluxes as a function of the daily mean temperature. Panel A shows that CO_2 fluxes exhibit a distinct negative correlation ($dCO_2/dT = -0.49$ µmol/K) at temperatures below 12 °C, consistent with Ward et al. (2022). Above this threshold, the CO_2 flux stabilizes ($dCO_2/dT = +0.02$ µmol/K), indicating that emissions remain relatively constant at higher temperatures. This shift aligns well with the Austrian norm ÖNORM H 7500-3, locally known as the Heizgrenze (heating threshold), which regulates heating system activation when outdoor temperatures drop below 12 °C.

The strong temperature dependence of CO_2 fluxes can be directly attributed to methane consumption from gas heating systems, which dominate the heating sector in the flux footprint. As outlined in Stichaner et al. (2024), methane consumption follows a pronounced temperature dependency below the heating threshold, a pattern that is mirrored by the CO_2 fluxes analyzed here. Furthermore, photosynthetic activity plays only a minor role in this footprint (see Supplement Fig. 2S B) and does not significantly affect the site as a CO_2 source. This further supports the conclusion that heating-related emissions, rather than biogenic processes, drive the observed temperature dependence of CO_2 fluxes.

REMOVED:

The temperature dependency of CO_2 fluxes aligns closely with the natural gas consumption analysis conducted by Sti-chaner et al. (2024), which also reveals a negative correlation at a specific temperature threshold. Natural gas is predominantly used for heating in the RCP (ie. residential/domestic, commercial and public) sectors. Since many heating systems within the footprint are powered by gas and oil, sources with lower NOx/CO_2 ratios become more prominent during the heat-ing season

COMMENT	9:
Page 13, Line 6: The term RCP for me is taken by the IPCC's 'representative concentration pathway', please check if there is alternative terminology for your field.	

REPLY:
You are right; this could cause misunderstandings. Therefore, we have replaced all instances of "RCP" with "urban energy sector" for clarity.

COMMENT	10:
Page 13, line 28: please only note significant digits.	

REPLY:
adapted according to the suggestion

COMMENT	11
Page 14: Lines 6-11: I would like to see wind speed and dynamic stability / cross wind variance be included in the discussion, as wind direction alone oversimplifies the interpretation. The flux footprint will vary also with the additional quantities.	

REPLY:
We have added a discussion on flux footprint, which takes into account stability and cross wind variance by referring in more detail to the flux footprint shown in the supplement. In the urban area the stability expressed by Monin-Obhukov Length shows that the site is mostly characterized by unstable to neutral conditions, which is due to the urban heat island effect.

CHANGES:
We added a new figure (Fig. 3S) to the supplement showing the stability and crosswind for the lockdown and non-lockdown period for the site and extended the description of Fig. 7 (before Fig. 6) with:

To further assess potential differences in meteorological conditions between the two periods, we also analyzed stability conditions (represented by the Monin-Obukhov length) and crosswind patterns (Supplement Fig. 3S). The results indicate that no significant differences were found between the lockdown and non-lockdown periods. This further reinforces the robustness of the observed pollutant variations, confirming that changes in air quality were primarily driven by emission reductions rather than shifts in meteorological conditions.

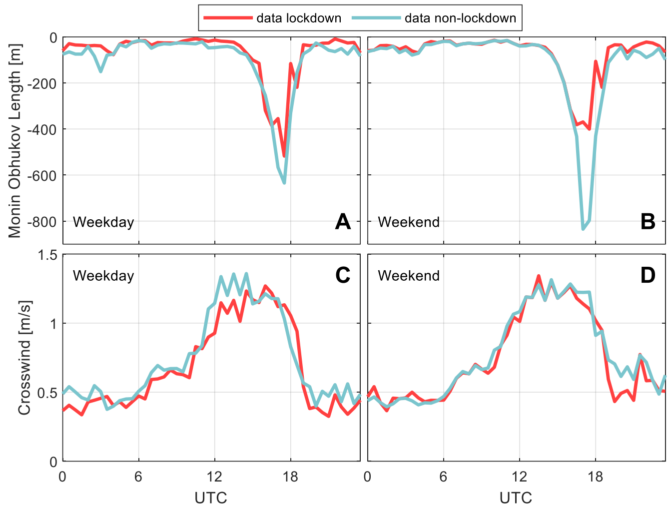


Figure 3S: (A, B) Comparison of the average stability (represented by the Monin-Obukhov length) between the lockdown period (red) and non-lockdown period (turquoise) on a half-hourly basis for weekdays (left panel) and weekends (right panel). (C, D) Similar to (A, B), but showing the crosswind for the same periods**.**

COMMENT 12:
Section 5: I think some portion may need to be rewritten after addressing the major comments.

REPLY:
Based on the comments, we have adjusted parts of the abstract in accordance with the reviewers' recommendations

Reviewer 2:

Summary statement:

The paper presents a novel method of examining sources of pollutants (in this case NOx) in an urban environment using a 'Turbulent Enhancement Ratio (TER)'. The quantity is calculated from fast response analysers, thus resolving the turbulent motion of the pollutant. The TER is compared to a flux ratio (FR) of the pollutants (NOx / CO2), calculated using the widely used eddy covariance technique for each species. Once it has been shown that the agreement between the TER and FR is good, the authors use TER to examine sources of NOx in the Innsbruck area, in particular using the difference between COVID19 lockdown and 'normal' times.

This is an interesting paper describing a novel technique that could lead to important insights into air pollutant emissions in urban environments. I recommend publication subject to the following revisions and additions.

REPLY:
We thank the reviewer for his valuable assessment. Below we address all comments grouped by their relevance. We fixed some errors that were pointed out.

General comments

COMMENT 1 + 2:
I think the manuscript needs more discussion as to why the TER is preferable to standard Eddy covariance fluxes? It seems that they essentially give the same thing (especially when looking at ratios), and both require high time resolution measurements. So what are the main benefits of TER.

On page 8 line 10 it is stated that TER is the preferred methodology (compared to the more widely used normalised emission ratio (NER)). It is not clear to me why this is the case, and this section would benefit from an expansion to better explain this.

REPLY:
We revise this statement to clarify that TER should be seen as an extension to NER by taking into account the turbulent part of the correlation between two tracers. As noted by reviewer one, mathematically the approach should be more seen as a spatial filtering approach to NER. In this context we interpret TER as a spectral similarity ratio, similar to approaches used to filter data in turbulent flows (Antonia et al., 1987, Nappo, 2012) This is particularly relevant in urban areas and other complex environments, where the background can not be easily separated from plume enhancements. This leads to biases and has been discussed in the literature (e.g. Yokelson, 2013). To overcome these limitations, we leverage the advantages of turbulence to distinguish between local and non-local sources, allowing for a more accurate representation of enhancement ratios in dynamic and complex environments. We have also expanded the theory and discussion sections to provide greater clarity and detail.

CHANGES:
As suggested by Reviewer 1, we have expanded the discussion to provide a more detailed context on spectral analysis. Accordingly, we have revised Fig. 2 and introduced Fig. 3, which presents the co-spectra and ogives of CO_2 and NO_x. In this context, we also emphasize that the

importance of experimentally resolving TER becomes particularly pronounced when the measurement site is close to an emission source, as turbulent mixing plays a dominant role in determining enhancement ratios. This effect is especially critical during morning and evening hours, when the planetary boundary layer (PBL) undergoes strong transitions, leading to shifts in vertical mixing conditions that directly influence the observed ratios.

COMMENT 3:
Page 13 line 8. This paragraph would benefit from expansion. It says that higher CO_2 fluxes are seen in the colder seasons the result of increased domestic heating in these seasons. This is almost certainly true but why do the NOx fluxes not also increase in the colder season? Presumably there is also a NOx emission from heating – is this not observed at all? Please comment on this.

REPLY:
It is true that NO_x emissions also originate from heating systems. However, within our footprint, most heating systems are gas-powered, emitting approximately 1,000 times more CO_2 than NO_x per TJ of energy (based on the emission factor inventory of Austria). Since more than 90% of the NO_x emissions at the field site area are traffic-related - remaining relatively stable between summer and winter - the impact of temperature or the "heating period" on NO_x fluxes is minimal. Instead, other factors, such as slight shifts in the footprint, play a more significant role in influencing NO_x flux variations.

CHANGES:
In contrast, the NO_x flux demonstrates a different pattern. It exhibits only a very moderate positive temperature dependence (dNO_x/dT = +0.14 nmol/K), which remains relatively constant across the entire temperature range examined. This suggests that NO_x emissions, primarily influenced by traffic (Lamprecht et al. 2021), do not fluctuate significantly with temperature changes. A key factor in this weak temperature dependence is the composition of heating systems within the turbulent footprint. Most heating systems in this area are gas-powered, which produce negligible NO_x emissions compared to the traffic sector, thus the impact on NO_x fluxes are largely decoupled from temperature-driven heating demand.

The interplay of these two trends is reflected in the Turbulent Enhancement Ratio, which decreases at lower temperatures. This decline can be attributed to the increasing dominance of CO_2 sources, particularly from heating systems, as temperatures drop. Consequently, the relative contribution of NO_x to the total emissions diminishes, altering the NO_x/CO_2 ratio.

COMMENT 4:
In section 4.3.2, the authors could comment on whether their data shows a change of dominant source during the lockdown period (i.e. does the NOx / CO2 ratio suggest a change from traffic dominated to domestic heating dominated emissions), as has been seen in other studies (e.g. Cliff et al., ACP, 2023).

REPLY:
we agree that the NO_x/CO_2 ratio shifts more towards other sources (e.g. residential) exhibiting a lower ratio during the lockdown, has been observed for other locations (e.g. London). As pointed out by the reviewer these findings are similar to the study by Cliff et al. from central London, which we now consider in our discussion.
Significant shifts from road- to non-road sources have also been observed in London during pandemic restrictions (Cliff et al. 2023); (p13)

COMMENT 5:
The authors state that they have a data series from mid 2018 to early 2022 but do not comment on any longer term trend. It would be nice to see if the NOx / CO2 ratio has changed from the start to the end of their dataset (when COVID restrictions had been lifted) and what any change could be attributed to

REPLY:
In Innsbruck the vehicle fleet still includes a significant proportion of diesel-powered cars, and unlike cities such as London, there are no low-emission zones where high-emitting vehicles exceeding predefined thresholds are restricted. While the transition to lower-emission vehicles with improved Euro standards is currently underway, longer and more extensive measurement campaigns at our field site will be necessary to accurately quantify the impact of this transformation.

Specific comments

COMMENT 6:
2.2 Instrumentation section: It is not clear which of the two NO2 measurements is used for the calculations (of TER, NER and fluxes). If it is the Moly converter NO2, have the authors investigated any potential issues with the residence time of air in the converter and the validity of 5Hz measurements?

REPLY + CHANGES:
We added a hint that the CLD899Y was used to perform the calculations for Flux and TER.

We confirm that the converter is a molybdenum converter, and its performance was assessed during an instrumentation characterization after the instrument was delivered. Our tests determined that the residence time of air in the converter is less than 1s. We have previously tested the response time for measuring NOx EC fluxes using the CLD899 and found a damping timescale of 0.8 s (e.g. Karl et al., 2017. doi: s41598-017-02699-9). This represents a high frequency loss of about 13%. From our analysis, here we find that in an urban area a 1s response time is still sufficient to capture the relevant turbulent co-variant parts.

COMMENT 7:
Table 1: what do 'dd' and 'ff' represent in the measured parameters from the sonic anemometer? It is probably obvious but should be made clearer.

REPLY:
Thank you. More specified by adding a footnote below the table. The abbreviation is commonly used in the field of meteorology.

COMMENT 8:
P 13 line 24: It says ‘the TER remains higher’ but higher than what. Please better explain this sentence.

REPLY:
Replaced the sentence for greater clarity. "Remain higher" referred to the 55% traffic reduction, whereas the TER did not decrease by the same proportion: Compared to the approximately 55% reduction in weekday traffic during the lockdown, the decrease in TER is less pronounced (p13)

COMMENT 9:

Section 2.3: I think it would benefit from a slightly more detailed description of how the eddy covariance fluxes are calculated, including the key parameters and filtering methodology used. I realise it is in the literature but just a few sentences would greatly benefit this paper.

REPLY:

good point, the key parameters for filtering are given now in the section "2.3 Dataset": We used the following quality assurance/quality control (QA/QC) criteria: a signal-to-noise ratio > 3, a steady-state criterion ≤ 0.5, a noise RMSE ≤ 20 for NOx and CO2 fluxes and a correlation coefficient ≥ 0.5 for TER while TER and FR were calculated for 30 min averages.