RC2:

The manuscript, while providing and overview over several BTEX datasets in one place to me appears to be lacking depth in the analysis and truly novel results. To me it appears that the paper is more of a measurement report paper. Hence while I am not convinced about novel findings related to atmospheric chemistry, I do believe it can be considered as a measurement report.

However, I have several concerns regarding the statistical analysis presented:

It is clear from the box plots in Figures in the supplement and main text that the data does not follow a normal distribution at any of the sites. The mean is always significantly higher than the median. It is very likely that all these distributions will fail a normality test when subjected to one hence statistical parameters used to represent the data must be robust parameters.

My statistics and data analysis related concerns are as follows:

1)Average and standard deviation are not the correct statistical parameters to report when data fails the normality test. Instead, the median would be more appropriate than the mean. Sigma can be approximated in a robust manner from the median median deviation MAD (1 sigma = 1.48*MAD). Authors can also report the median and the range of values observed instead.

Response: We acknowledge that reporting the mean and standard deviation may not be appropriate given the non-normal distribution of our data. In the revised manuscript, we have replaced these metrics with the median and the median absolute deviation (MAD) as robust statistical measures. We have added the following sentence to the method section (line 182-184):

Therefore, in this study, BTEX levels are presented as median mixing ratios (MED) ± median absolute deviation (MAD) in parts per trillion by volume (ppt) to reflect the non-normal distribution of the data.

2) The difference of median between seasons must be assessed with a robust test such as Mann-Whitney test since the data is non-normal. The robust substitute of ANOVA would be a Kruskal-Wallis test for equal median although a Mann-Whitney pairwise comparison between season may turn out to be more revealing and interesting.

Response: Thank you for your suggestion. We agree that a robust test is necessary given the non-normality of the data. Therefore, we have replaced the current analysis with the Mann-Whitney test to assess the differences in BTEX levels between seasons. Additionally, we incorporated a Kruskal-Wallis test for assessing equal medians and conducted pairwise Mann-Whitney comparisons between seasons, as suggested. The results still indicate significant differences in BTEX levels across the seasons. Although the conclusions are consistent with those obtained using the previous ANOVA approach, we followed the reviewer's recommendation to apply the Mann-Whitney test to ensure a more robust analysis of the non-normally distributed data.

Test Statistics[a]

	Benzene	Toluene	Ethylbenzene	m,p-Xylene	o- Xylene
Mann-Whitney U	11141561.500	22785960.500	19782083.000	19342025.000	20020547.500
Wilcoxon W	24333377.500	36501663.500	33211236.000	32349575.000	33387582.500
Z	-55.716	-13.506	-22.179	-21.629	-20.893
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000
a. Grouping Variable: Spring-Summer					

Test Statistics[a]

	Benzene	Toluene	Ethylbenzene	m,p-Xylene	o- Xylene
Mann-Whitney U	41042737.000	37480442.500	35590465.000	34944741.000	35669694.500
Wilcoxon W	88510633.000	87846108.500	83497831.000	81762067.000	83420572.500

Z	-14.329	-27.996	-27.245	-26.466	-26.513
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000
a. Grouping Variable: Spring-Autumn					

Test Statistics[a]

	Benzene	Toluene	Ethylbenzene	m,p-Xylene	o- Xylene
Mann-Whitney U	29642393.000	41652262.500	39365751.500	39224692.500	39342415.000
Wilcoxon W	77110289.000	92017928.500	87273117.500	86042018.500	87093293.000
Z	-43.891	-17.123	-17.238	-15.662	-16.540
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000
a. Grouping Variable: Spring-Winter					

Test Statistics[a]

	Benzene	Toluene	Ethylbenzene	m,p-Xylene	o- Xylene
Mann-Whitney U	9224353.000	15511278.000	12996645.000	12702955.500	13457629.000
Wilcoxon W	22416169.000	29226981.000	26425798.000	25710505.500	26824664.000
Z	-62.518	-39.336	-46.748	-46.051	-44.463
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000
a. Grouping Variable: Summer-Autumn					

Test Statistics[a]

	Benzene	Toluene	Ethylbenzene	m,p-Xylene	o- Xylene
Mann-Whitney U	4764916.000	17959649.000	15280818.000	15304223.000	15663183.000
Wilcoxon W	17956732.000	31675352.000	28709971.000	28311773.000	29030218.000
Z	-80.742	-29.284	-37.256	-35.495	-35.121
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000
a. Grouping Variable: Summer-Winter					

Test Statistics[a]

	Benzene	Toluene	Ethylbenzene	m,p-Xylene	o- Xylene
Mann-Whitney U	34862699.500	42758108.000	40750515.000	39671945.000	40395430.000
Wilcoxon W	80602529.500	89440061.000	84878830.000	83219723.000	84045926.000
Z	-28.606	-10.573	-9.253	-9.948	-9.198
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000
a. Grouping Variable: Autumn-Winter					

Test Statistics[a,b]

	Benzene	Toluene	Ethylbenzene	m,p-Xylene	o- Xylene
Kruskal-Wallis H	7514.139	1791.224	2398.635	2256.100	2181.433
df	3	3	3	3	3
Asymp. Sig.	0.000	0.000	0.000	0.000	0.000
a. Kruskal Wallis Test					
b. Grouping Variable: Seasons					

3) All the fits in Figure 5 appear to be ordinary least square regressions despite the fact that both x and y axis contain measured data that carries measurement uncertainties. OLR is only suitable when the quantity on the X-Axis is absolute and error free. Please use major axis regression (MA) also

known as orthogonal regression to fit your data. Also: You need to give the error of the slope and intercepts. Figure 5B urban panel the line fit for summer doesn't appear to have much to do with the data at least I can't see the corresponding circles anywhere near the line. The slope may be driven by some extreme values outside the x and y axis boundaries. Please check your data

Response: Thank you for your insightful comments. We appreciate your input regarding the use of ordinary least squares regression (OLR) and its suitability for our data. In response, we have employed major axis regression (orthogonal regression) for our analyses, which better accounts for measurement uncertainties in both x and y variables. We have also included the error estimates for the slope and intercept in the revised Figure 5. We have updated the Figure 5 to reflect this change. Furthermore, we have re-evaluated the data points in Figure 5B and ensured that the regression line corresponds appropriately with the observed data points, removing any misleading extremes.
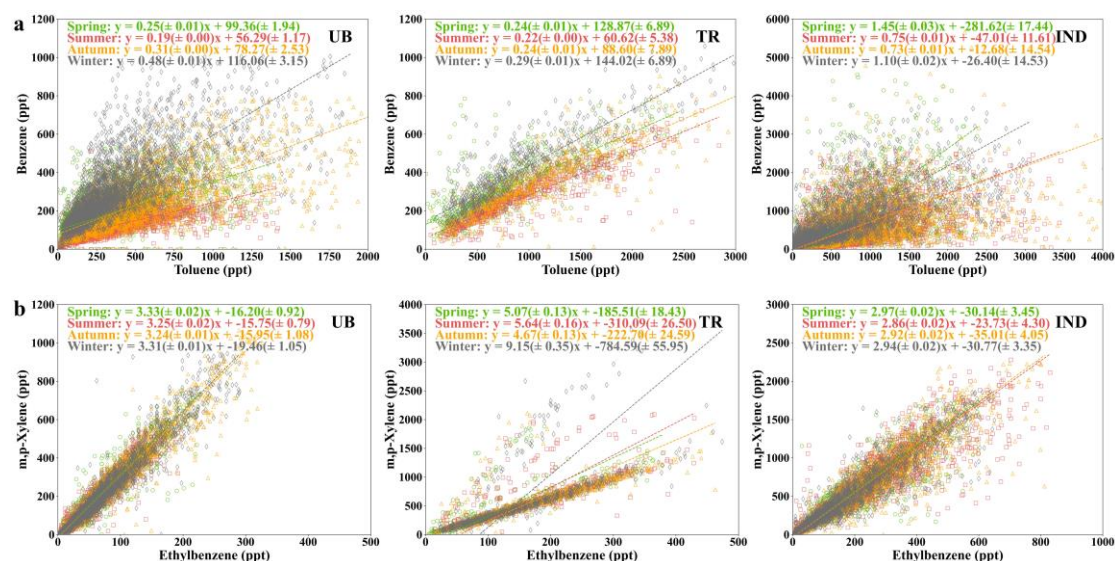

Figure 5. Regression of benzene to toluene (a) and m,p-xylene to ethylbenzene (b) ratios in different urban locations (UB, urban background; TR, traffic; IND, industry) across different seasons (green, spring; red, summer; gold, autumn; gray, winter).

My other concerns are es follows

1) Referring to sites with the site code instead of the name of the city makes the manuscript hard to follow. It would be nicer to refer to them with the name of the city and nature of the site even if it costs a few more words.

Response: Thank you for your comments. Regarding the concern about referring to the sites with codes instead of city names: The cities mentioned in the manuscript are part of the RI-URBANS project, and the site codes (i.e., the first three capitalized letters of the city name followed by the site type: UB, TR, IN) are standardized across all related documentation. This format ensures consistency throughout the project and allows for easy reference to the sites across different studies. Therefore, while we understand that using city names might make the text more intuitive for some readers, we believe that adhering to this standard is essential for maintaining uniformity within the context of the broader project. To assist readers, we have included Table 1, which provides a clear reference linking the site codes to the corresponding cities and site types, thereby helping to clarify

the manuscript.

2) Please make all comparison qualitative: E.g. a statement like

"Nonetheless, compared to the period before 2000, the levels of benzene and other BTEX compounds have shown a decreasing trend due to the successful implementation of air quality measures in Greece, such as the extension of metro lines and the use of catalytic converters in cars" Implies a time series trend analysis was done which I can't find in the results section. So it appears to be a comparison between values reported in this study and a previous study. Then why not state clearly what the was then, and is now.

Response: Thank you for your valuable suggestion. Following your suggestion, we have now added the specific values to clearly indicate the comparison between the results from this study and previous studies.

"Nonetheless, compared to the period before 2000, the levels of benzene and other BTEX compounds have shown a decreasing trend due to the successful implementation of air quality measures in Greece, such as the extension of metro lines and the use of catalytic converters in cars (Panopoulou et al., 2021). For instance, benzene levels at traffic sites have decreased significantly, dropping by as much as eightfold, from approximately 12520 ppt in 1994 to about 1565 ppt in 2016 (Panopoulou et al., 2021). Similarly, at urban monitoring stations, benzene concentrations fell sharply, from around 4695 ppt during the period of 1993–1996 to between 313 and 1565 ppt in 2016 (Panopoulou et al., 2021)."

3) Please highlight and explain interesting data instead of only focusing on the data that meats the expectations of BTEX=traffic & Industry. Why is the B/T rations in urban Helsinki and suburban Paris so high?

Response: Thank you for your insightful comments. We acknowledge that the B/T ratios in urban Helsinki and suburban Paris warrant further discussion. In Helsinki (HEL_UB), the monitoring data is limited to February 2016, which may lead to a comparability bias and explain the relatively high B/T ratio observed. We have emphasized this limitation in the revised manuscript. For suburban Paris (PAR_SUB), the elevated B/T ratio can be attributed to seasonal factors. Benzene concentrations typically increase from September to April, driven by more active sources during the winter months, such as residential wood burning, and limited dispersion due to a lower boundary layer. This trend is further supported by the presence of wood-burning tracers like furfural and benzenediol, which exhibit similar seasonal patterns (https://doi.org/10.5194/essd-15-1947-2023). Although toluene does not show as strong a seasonal variation as benzene, it also has higher levels during autumn and winter. The primary source of toluene, traffic, remains important throughout the year, and more stagnant conditions in these seasons contribute to the accumulation of pollutants. We have revised the manuscript to include these explanations for a clearer understanding of the data.

Specific modifications can be found in Lines 347-350: At HEL_UB, the B/T ratio is relatively high, which can be attributed to the limited monitoring data available only from February 2016, introducing a potential comparability bias. For PAR_SUB, the elevated B/T ratio can be attributed to seasonal factors (Simon et al., 2023). Benzene concentrations typically increase from September to April, driven by more active sources during the winter months, such as residential wood burning (Languille et al., 2020), and limited dispersion due to a lower boundary layer (Simon et al., 2023). This trend is further supported by

the presence of wood-burning tracers like furfural and benzenediol, which exhibit similar seasonal patterns. Although toluene does not show as strong a seasonal variation as benzene, it also has higher levels during autumn and winter. The primary source of toluene, traffic, remains important throughout the year, and more stagnant conditions in these seasons contribute to the accumulation of pollutants.

4)Why are B/T ratio's higher at industrial sites. While toluene is known to be an industrial solvent significant toluene emission should actually lower the ratio. Industrial benzene use has supposedly been phased out due to the carcinogenic nature. Can the authors comment on the industrial benzene source? Alternatively, is this actually representative of industrial sites in Europe in general since 2/3 industrial sites are in Lyon. This could be related to local emissions from a specific industry.

Response: Thank you for your valuable comments on the B/T ratio observed at industrial sites. Although toluene is commonly emitted in greater quantities from industrial sources, its atmospheric lifetime is much shorter than that of benzene (toluene: 2.1 days, benzene: 9.5 days). As a result, even though there may be significant toluene emissions, the shorter lifetime means that it degrades more rapidly in the atmosphere compared to benzene. This leads to an accumulation of benzene relative to toluene, particularly near industrial sources, which results in a higher B/T ratio despite toluene's greater initial emissions (Atkinson and Arey, 2003; Liu et al., 2008). We believe this factor, combined with potential localized sources of benzene, could explain the relatively higher B/T ratios observed at some industrial sites.

In addition, the observed B/T ratios may not be fully representative of industrial sites across Europe, as the industrial emissions are highly dependent on specific processes at each location. In this study, two-thirds of the industrial monitoring sites are in Lyon, which may reflect local industry characteristics rather than a broader pattern across European industrial sites. Emissions can also be episodic, influenced by particular operational activities, and may vary with changes in wind direction, further complicating comparisons across regions. These factors suggest that the higher B/T ratios observed may be more indicative of localized conditions rather than a general trend for European industrial areas.

Minor corrections

Line 109-113 BTEX became TEX please correct

Response: Thank you for pointing this out. We have reviewed the original text and confirm that "TEX" is correct in this context, as the proportion of benzene in traffic emissions is relatively small compared to other TEX compounds. Therefore, the use of TEX, instead of BTEX, accurately reflects the findings of the referenced study.