We thank the reviewer for valuable comments and suggestions to improve the manuscript *Uncertainty of eddy covariance flux measurements over an urban area based on two towers*. Please find our point-by-point responses below.

General comments

The Manuscript AMT-2018-89 by Leena Järvi et al. discusses a paired tower approach to assess the representativeness of measurements of vertical momentum, sensible heat, latent heat and carbon dioxide fluxes with the Eddy Covariance technique in a densely built urban environment. The two identical instrument systems were installed on the same building and only 10m apart, therefore they are virtually sampling the same source area outside of flow distortion angles that need to be excluded. The study is relevant to the scientific community since it will (1) help to better understand the relative and absolute magnitudes of measurement errors over an urban, heterogeneous surface area. Furthermore, the approach also allows to (2) assess the representativeness of measurements at sites equipped with a single Eddy Covariance system, specifically when measurements over flow distortion areas need to be excluded which can result in systematic biases in temporally cumulated flux sums. The study was conducted with micrometeorological and statistical rigor; however, a list oftechnical corrections need to be addressed before publication in Atmospheric Measurement Techniques. Most of the corrections can easily be fixed by quite simple edits and efforts to make the text slightly more concise, predominantly in the Abstract and the Conclusion sections.

———————-

Specific comments

(1) Given the extensive experimentation, subsequent processing and analysis of the data, the abstract has to be edited and refined quite substantially. In my view, this will ensure the accessibility of the technically dense study to a wider, perhaps even non-technical audience. Some of the sentences are worded in a way that is too vague, so they could be misinterpreted. I am providing more specific feedback on the abstract in the Technical Corrections section below.

We thank the reviewer for pointing out that the abstract is not necessarily meeting the right audience for the paper. Detailed responses can be found from below.

(2) The following result that is stated in the abstract appears to be not stated explicitly anywhere throughout the manuscript, or at least I was not able to find it: “the random uncertainties of the two systems are between 10 and 40 %.” I suppose the values are simply “read off” of Fig. 6? If this is the case, it looks to me that really what was deducted here is that the interquartile range of the random uncertainties is between 10% and 40 %? I suggest to either provide the range of average (or median) uncertainties, or to add the statistical significance of the uncertainty assessments (e.g., at the 75% or 95% significance level). If the authors feel that would be too much detail for the abstract, then please incorporate this detail into the text of the manuscript. Otherwise the reader may be left guessing where this quite important result came from. Finally, the wording of this sentence as it stands also allows for some misinterpretation that the random uncertainties are calculated by looking at both towers together (which is ultimately one key objective of the study, cf. the key word “between”), when really the numbers 10%-40% simply represent the random uncertainties of one system at a time.

Thank you for pointing this out. We agree that it is more meaningful to have here the range of mean values rather than the interquartile range. We changed the range to 12-28% (also reported in the main text). With the suggested modification to the sentence it now reads “the median random uncertainties of the studied fluxes measured by one system are between 12 and 28 %” (P1, L17).

(3) One thought I kept pondering about while reading the article was the downstream implication of this study for future experimental EC studies in urban environments. Specifically, the results on the representativeness and sensitivities of the measurements as obtained by the paired tower approach. I.e., is the result conclusive of noconfirmation being needed at other locations? Are the results truly transferable to other urban locations with fairly homogeneous flux source areas as pointed out at the end of the abstract? Or, is further experimental validation needed? May the authors please discuss this in some more detail, perhaps at the end of section 3.4 in a short paragraph.

The referee raises here a good point. In addition to the abstract this is also shortly mentioned in the conclusions. As suggested we now added further discussion to the end of the Section 3.5 (P20, L25-30):

"The outcome of our study is that a single EC measurement point can produce reasonable estimations for surface fluxes above relatively homogeneous urban surface, but the next question naturally is that how applicable this result is for other urban EC sites. Each urban measurement location is unique and in order to get a final answer, each site should be separately evaluated with more than one measurement setup. Nevertheless, the obtained uncertainties from this study can be used as a first approximation for urban EC measurements in a same way as the few two or multiple tower studies made in vegetated ecosystems are used to give general guidelines for the uncertainties."

(4) It is really encouraging to see that the random uncertainties decreased by applying the paired tower approach in an urban environment. Even better, the relative magnitude of the uncertainties appears to be in the same range as reported by previous studies with more homogeneous terrain. To me, this is one of the key results of the study, and could perhaps even be highlighted in the Abstract since it is highly relevant to future studies conducted in urban or other heterogenous terrain without "directional deviations" in the source area.

We cannot really say that the random uncertainties of the EC observations decrease by applying the two-tower approach as no joint value for the combined dataset is calculated. We only show that the systematic uncertainty decreases and this already mentioned in the abstract. However, we added a sentence "The obtained random and systematic uncertainties are in the same range as observed in vegetated ecosystems." (P1, L21-22) to emphasize the correspondence of the random and systematic uncertainties with those obtained in vegetated ecosystems.

Technical Corrections

Abstract, LL3-4: "Often one ecosystem is monitored using only a single EC measurement station bringing uncertainties to the ecosystem-level flux values." I would re-write this to: "Typically an ecosystem is monitored by only one single EC measurement station at a time, making the ecosystem-level flux values subject to random and systematic uncertainties"

Corrected as suggested.

Abstract, LL12-14: are "measurement location" and "measurement structures" used synonymously in this sentence? Might not be clear to a wider audience.

Yes. To clarify this, we removed measurement structures from the sentence and now it simply reads "The momentum flux is the most sensitive to the measurement location whereas scalar fluxes are less impacted" (P1, L13-14).

Abstract, L18: I suggest writing: "Combining the data from two EC systems also increases the percentage of usable half-hourly carbon fluxes from 45% to 69% at the annual level."

Modified to “Combining the data from two EC systems also increases the fraction of usable half-hourly carbon fluxes from 45 % to 69 % at the annual level.” (P1, L19-21)

Abstract, LL17-19: I suggest to also give absolute values for the underestimation in grams of Carbon p.a., next to the 12% and 5-8%.
Added as suggested. We also added the values in units g C m-2 to the manuscript main text.

Abstract, L22: Which uncertainties are you referencing here? Random, systematic, or both? Please specify. (If I understood correctly, you are referencing both systematic uncertainties due to excluding flow-distorted wind sectors, and, random uncertainties due to turbulent sampling errors as assessed by the relative random uncertainty (RRE) metric.)
We refer to systematic uncertainty and this is now clarified also in the sentence (P2, L4).

Abstract, L22: I suggest changing “The same results can be assumed to apply in similar dense city locations [...]” to “Comparable results can be expected in similarly dense city locations [...]”
Corrected as suggested (P2, L4).

Pg2, L26: please add a reference.
Reference added (P2, L31).

Pg2, L27: “to reject large amount of data”: I suggest writing “to reject a relatively large fraction of the data”.
Corrected as suggested (P2, L31).

Pg3, L2: I suggest writing “On top of that, any statistical gap-filling technique can be biased [...]” instead of “Either way, statistical gap-filling techniques can be biased [...]”
Corrected as suggested (P3, L7).

Pg3, L8: I would add the study of Hollinger & Richardson (2005) to this list of paired tower approaches, since it was the first of its kind.
Added as suggested (P3, L13).

Pg3, L19: Figure reference is missing. (???)
This was fixed.

Section 2.1: Can the authors please add one sentence in section 2.1 (Site description) on the representativeness of the flux source area as surveyed by the tower with respect to the “Helsinki city centre” that is referenced further down in the Conclusions
We added sentence “The two systems have a separation distance of 10 m and thus measure virtually the same source area” to P4, L2-3.

(Pg21, L4)? Perhaps simply by referencing information in the original citation for this site. Since the results of the study at hand are discussing the “representativeness” of measurements in a sampling sense, it may be helpful to the reader to be able to put things into perspective. It would also illustrate how essential it is to understand the fluxes extremely well at a fine spatial scale, to then use these measurements as the basis for accurate assessments of larger neighborhood or city level scales.
We modified the sentence to (P23, L4-8) “This result is naturally location-specific for this highly built-up site with vegetation fraction only 22% and relatively homogeneous roof level

(Nordbo et al. 2013). The same result could be considered to apply also in other dense city centers with similar relatively homogeneous surface characteristics."

Pg5, L26: Spelling mistake in the word "square", please run a spell check before submitting the revision.
Done.

Pg8, L1: "the median R2 between the two measurement systems is 0.85"
Added as suggested (P8, L11).

Pg8, L25: this discussion may be more meaningful if an equation for R2 was provided. There are different equations for R2 in the statistical literature. Also, I suggest to re-write this same sentence and the following to: "Both statistical variables RRE and R2 should theoretically be a measure of random uncertainty. When RRE between the two systems are larger, R2 is expected to be smaller. Furthermore, we expected the two resulting uncertainty rankings (according to RRE and R2) across the different fluxes to be consistent."
We added sentence "calculated as the square of the Pearson correlation coefficient" to P8, L6. We do not calculate joint RRE for the two setups but rather separate RRE's for the two systems. We modified the change to "Both statistical variables RRE and R2 should theoretically be a measure of random uncertainty. When RREs measured with the two systems are larger, R2 between the two systems is expected to be smaller. Furthermore, we expected the two resulting uncertainty rankings (according to RRE and R2) across the different fluxes to be consistent." (P9, L25-28).

Pg12, L10: typo "1decreased"
Typo fixed.