This paper evaluates the methane quantification of UAV-based mass-balance method using single-blind controlled-release experiment at two UK sites. The manuscript demonstrates strong practical value and provides useful experimental validation data for better characterizing the performance of UAV-based mass-balance methane emission quantification method and providing insights on methodological and environmental requirements to deliver robust measurements. This paper is generally well written, and the experimental design appears carefully executed. However, several methodological clarification and results discussions should be strengthened before publication.

The major comments are summarized below:

(1) While the manuscript investigates the influence of varying wind speeds on methane quantification, there is limited discussion regarding the impact of wind measurements obtained from different measurement platforms. Since the mass-balance flux calculation is directly proportional to wind speed, systematic differences between mast-based and UAV-mounted wind measurements could introduce additional uncertainty or bias into the methane emission estimates. The authors are encouraged to provide a more quantitative comparison of emission estimates derived using different wind datasets, or at minimum discuss how discrepancies between these wind measurement approaches may propagate into methane quantification uncertainty.

(2) The manuscript should include an additional caveat regarding the broader applicability of the controlled release experiments. In particular, caution should be exercised when extrapolating the reported uncertainties and performance metrics to other real-world applications, since the experimental sites, especially the Bedford site, were conducted on a former runway with relatively open and simplified terrain conditions, such environments may not fully represent the complexity of actual methane emission sites, which often involve buildings, equipment, surface roughness, and more complicated fields. A discussion of these limitations would help clarify the scope and transferability of the presented results.

(3) The manuscript currently provides relatively limited detail regarding the configuration of the controlled releases. Additional clarification would improve reproducibility and interpretation of the results. For example, it is unclear whether each site involved only a single emission source operating at a given time, or whether multiple emission sources were released simultaneously. The duration and timing of individual release events should also be described more explicitly. In addition, the physical arrangement, geometry, and locations of the release sources would be valuable to show directly in Figure 3, as these factors may influence plume structure and methane quantification performance.

(4) The description around line 201 ("In total, 24 flights were carried out, for seven different release rates and three different types of emission sources") requires additional

clarification. It is currently unclear whether the seven different release rates refer to temporal variations of the site-level aggregated emission plume, or to different emission sources operating within the site. Similarly, the "three different types of emission sources" appear to refer to different controlled-release geometries, but this should be explicitly stated. The authors should also provide justification for introducing these three different release configurations and explain what types of real-world methane emission scenarios they are intended to represent. Such discussion would help readers better understand the practical relevance of the experimental design. Similar concerns apply to the description around line 218.

(5) In the "Results and Discussion" section, the authors should consider adding additional subheadings to improve the organization and readability of the manuscript, such as subsections "Flight-level methane quantification" and "Analysis of factors affecting methane quantification"

(6) The statement around line 262 ("we see that there is no strong fundamental dependence on the release rate or the source type") should be revised for consistency with the rest of the manuscript. Based on Figure 5 and other discussions in the paper, the bias at lower release rates appears noticeably larger than that at higher release rates, and the authors also acknowledge this trend elsewhere in the manuscript. The authors are encouraged to revise this statement to more accurately reflect the observed trends and maintain consistency throughout the paper.

(7) The statement regarding the "largest bias of -60.55" is somewhat unclear. It is not evident whether this value refers to single flight-level observation or an average across flights 20-24. The authors should clarify the statistical basis of this reported bias value to avoid ambiguity.

(8) The manuscript concludes that extrapolating the lowest transect measurements to ground level improves agreement between estimated and known methane fluxes, based on the increased percentage of flights whose known release rates fall within the estimated uncertainty bounds. However, this analysis alone may not fully demonstrate improved quantification accuracy. In particular, it is unclear whether the apparent improvement arise from a genuine reduction in bias or simply from enlarged uncertainty intervals after extrapolation. The authors should provide additional analysis (changes in mean bias) to better support this conclusion.

(9) One key assumption of mass-balance method is that the entire plume cross-section is adequately sampled. The manuscript should provide more evidence that the UAV flight paths fully captured that methane plume under all experimental conditions. It would be helpful to include additional information, such as a simple schematic showing the flight paths relative to the plume, explanation of how the transect spacing was selected, discussion of possible underestimation if part of the plume was missed during measurements.

(10) The background CH4 determination method requires additional justification and discussion. Since both plume and background measurements are included within the same dataset, the resulting histogram may be influenced by plume-enhanced concentrations, potentially biasing the estimated background concentration. This issue may become particularly significant for low-emission cases where plume enhancements are comparable to background variability.

In addition, the manuscript states that the background range was "manually selected" based on visual identification of the distribution tail transition. This approach appears somewhat subjective and may introduce operator-dependent uncertainty. The authors should clarify the selection criteria in a more quantitative and reproducible manner and discuss the sensitivity of the results to the chosen background range.

Minor comments:

(1) Section "2.2 UAV-based mass balance flux method" currently contains only a single subsection. The authors may consider removing the subsection heading and integrating the content directly into the main section for improved formatting consistency and readability.

(2) Equation (3) may be difficult to read in its current format. The authors may consider separating the upper and lower uncertainty bounds into two lines to improve clarity and readability.

(3) In the right y-axis label of Figure 2, please correct the methane notation from "ch4" to the proper chemical format "$CH_4$"

(4) At line 311, the symbol $\Delta\theta$ appears for the time without definition or explanation. The authors should clearly define this parameter upon its first occurrence to improve clarity and readability.

(5) At line 291, the authors should provide a clear definition or explanation of the "dead sampling" when the term is first introduced, as its meaning may not be immediately clear to all readers.

(6) All abbreviations and acronyms should be defined in full upon their first appearance in the manuscript to improve clarity and readability for a broader audience.