We would like to thank all reviewers for their valuable insights and comments made to improve this work. We believe we have addressed all comments made by the reviewers as shown below, and the suggested changes have greatly improved this manuscript.

Our response to all comments is structured so that the reviewers comments are shown in "**bold**", our responses are shown in *"italics"*, prior text in the manuscript is shown in "blue", added text is shown in "red", and any deleted text in the manuscript is shown in ~~"blue strikethrough"~~.

We would once again like to thank the reviewers for their valuable insights and comments regarding this manuscript.

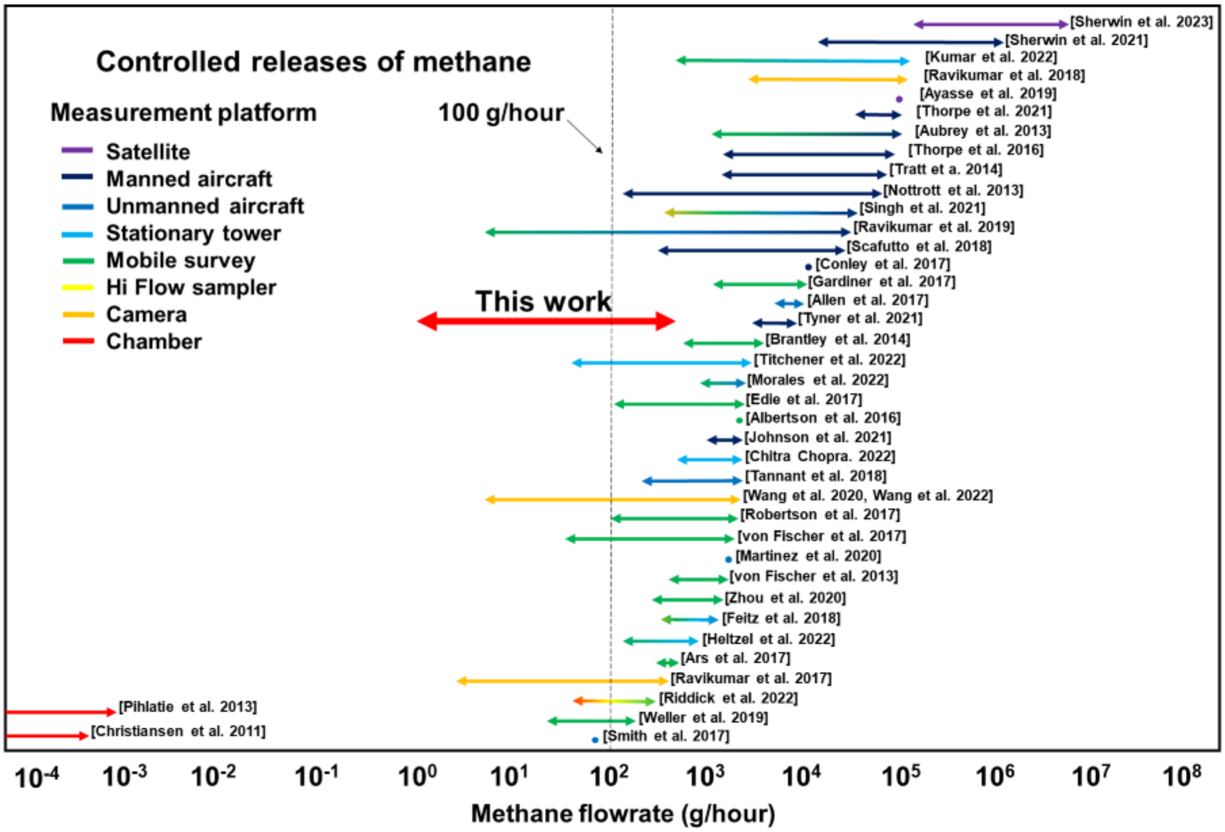
In addition to the changes outlined below for this reviewer, we also performed additional edits that we believe improve on the current version of the manuscript without significantly altering any results. First, we address some previously unmentioned methods we used for the linear regressions for controlled releases where methane concentrations were expected to rapidly reach steady-state. We also modified our presentation of the accuracy of measurements to encapsulate the median of the over- and underestimates to better represent the bias in our measurements. Finally, we incorporated two new studies on prior controlled releases into Figure 2 and the subsequent analysis section. The changes are outlined as follows:

Line 140: "For some experiments, methane concentrations within the chamber were expected to rapidly reach steady-state. Steady-state is reached when methane concentrations no longer increase over time in the chamber and the concentration of methane within the chamber is equal to the concentration of the released gas. The residence time, or time to reach steady-state, is defined by (2): where τ is the residence time, V is the volume of the chamber, and Q is the volumetric flowrate of gas (i.e., methane and balance gas combined) into the chamber. For any controlled releases where the expected residence time was two minutes or lower, we only used the initial ten data points for the linear regression to avoid the period of exponential decay as methane concentrations approach steady-state (Pihlatie et al., 2013).

(3) $\tau = V/Q$

For each factor being investigated, we grouped the results depending on whether the measurement was an under- or overestimate of the true methane flowrate. We calculated the accuracy of measurements as a range spanning from the median of the over- and underestimated methane flowrates, respectively. We determined the bias of measurements as the average of the raw percentage errors to determine whether tests were biased more towards the under- or overestimation of methane flowrates. "

New Figure 2:



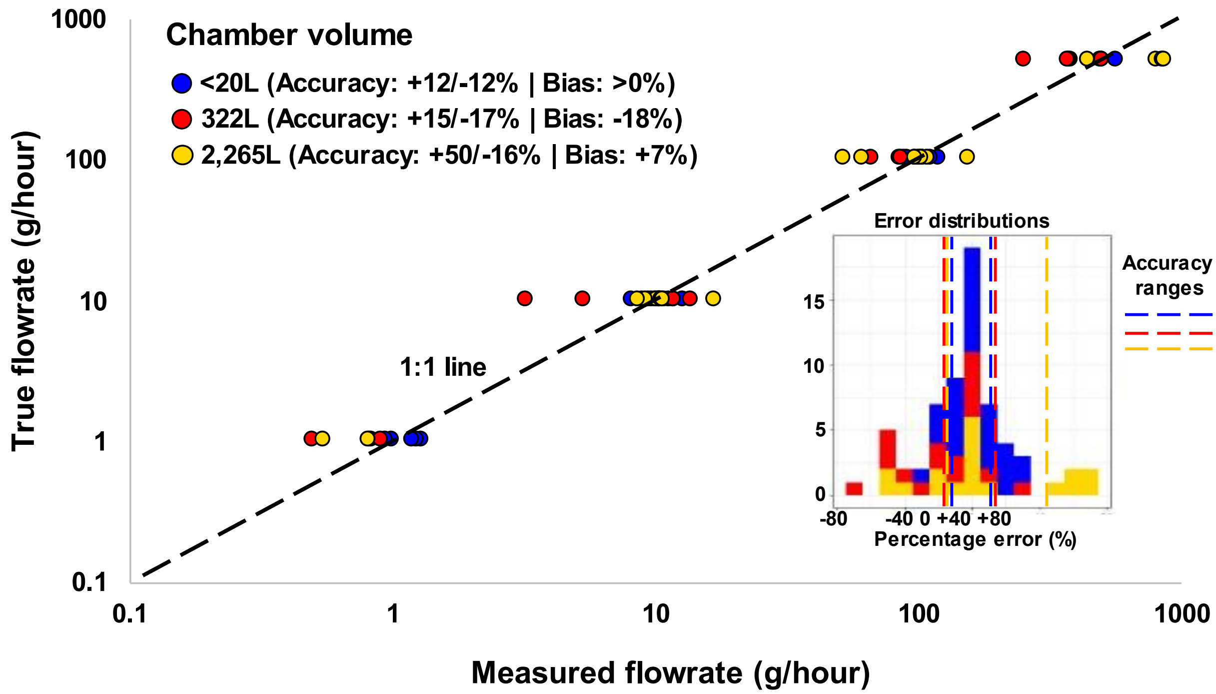
Reviewer #2

This work builds upon previous research on the assessment of the performance of static chamber methods for methane emission quantification. Such controlled release testing is important in interpreting methane emissions measurement data needed to assess the magnitude of emissions and progress toward emission reductions. The focus on component-level measurements is also important, given the methane policy implications. In addition, the detailed assessment of the various factors that could influence measurement accuracy contributes to the novelty of the work presented here. A few comments and suggestions for revisions are included below:

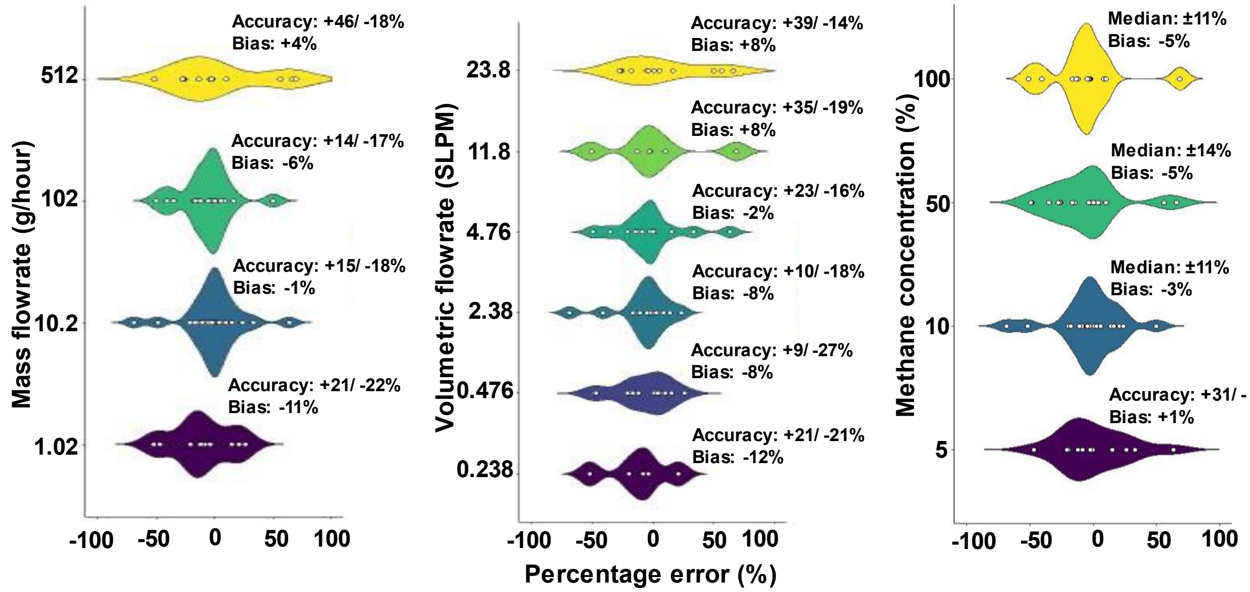
- **Some additional statistical assessment/visualization could help improve the data interpretation here. A regression analysis/parity plot of the performance of the static chamber method against metered flow rates is standard in these kinds of controlled release experiments, and it is not clear why the authors excluded this in their presentation. These parity plots (and accompanying goodness of fit tests) are easily accessible to**

the lay person than, say Figure 5, showing the percentage error correlation with metered flow rates. The authors should consider including such statistical analysis and visualization in the revised manuscript.

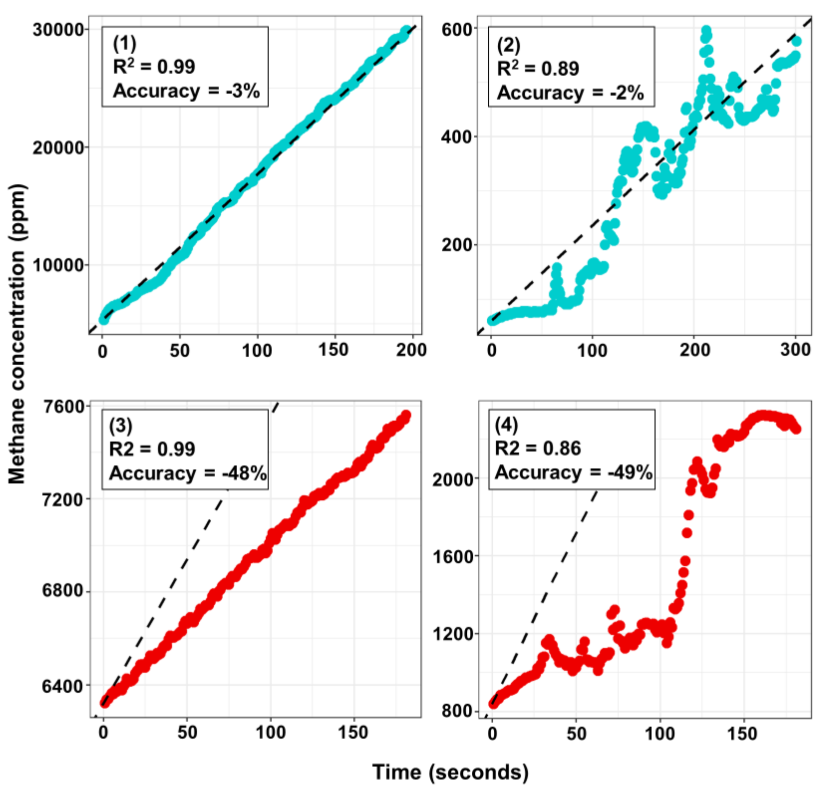
We agree with the reviewer and have since replaced Figure 3 with three new parity plots as suggested. The parity plots also contain histograms of the error distributions. We have also modified Figure 4 to remain a violin plot but instead show the actual measurement errors rather than the absolute measurement errors. Figure 5 has been removed and replaced with a figure showing examples of the raw data from the controlled release experiment for discussion on the large variability in accuracy observed in some of the controlled releases. Examples of the changed figures and revised text are shown below:



Parity plot example



Revised Figure 4



New Figure 8 that replaces old Figure 5

Line 178-185: "Our analysis of chamber volume with respect to quantification accuracy showed that the accuracy of measurements increased with smaller chamber volumes (Figure 4). The <20 L chambers had the highest accuracy at +12/-12% with an error standard deviation of 12%. The 322 L chamber had a lower accuracy of +15/-17% with a standard deviation of 23%. Our highest errors were measured from the largest 2,265 L chamber with an accuracy of +50/-16% and a standard deviation of 26%. We analyzed all three chamber sizes for bias and found that the <20 L chambers showed a slight tendency for underestimation of flowrates with an average bias of >0%, the 322 L chamber showed a stronger tendency towards the underestimation of flowrates at -18%, and the 2,265 L chamber showed a slight bias towards overestimating flowrates at +7% (Figure 4)."

Line 187-193: "Our comparisons of different chamber shapes showed that the cylindrical chambers were more accurate than the rectangular chambers, showing an accuracy of +5/-14% and a standard deviation of 18% (Figure 5). We found that the rectangular chambers showed a lower accuracy of +17/-14% with a standard deviation of 22%. Similar to the chamber volume, the median percentage error was smaller than the average error for both chamber shapes, which indicates an extreme distribution in percentage errors. We analyzed both chamber shapes for bias and found that the cylindrical chambers were biased towards the underestimation of methane flowrates with an average bias of -13% whereas the rectangular chambers showed a small bias towards the overestimation of methane flowrates with an average bias of +6% (Figure 5)."

Line 195-199: "The most impactful physical factor we observed on chamber measurement accuracy was the presence of fans, where chambers with fans present had a median percentage error of +6/-5% and a standard deviation of 17% (Figure 6), which was higher than chambers without fans which had an accuracy of +17/-17% and a standard deviation of 22%. For both data-sets we observed median values lower than the mean indicating a skewed data-set. We analyzed both data-sets for bias and found that both chambers with and without fans showed slight biases towards the underestimation of methane flowrates at -2% and -4% respectively (Figure 6)."

Line 202-208: "We tested four different mass flowrates for our controlled release tests: 1.02 g/hour, 10.2 g/hour, 102 g/hour, and 511 g/hour (Figure 7). The lowest errors were measured from the 10.2 and 102 g/hour mass flowrates each with accuracies of +8/-11% and +7/-13% respectively. The lowest accuracy of +56/-15% was attributed to the highest mass flowrate of 512 g/hour. We found that the 1.02, 10.2, and 102 g/hour mass flowrates all had negative biases of -11%, -1%, and -6% respectively. The mass flowrate of 512 g/hour had a slight bias of +4% towards the overestimation of mass flowrates, and also the highest upper accuracy estimate of +46% we observe among the different factors we analyzed."

Line 210-215: "We analyzed six different volumetric flowrates for the range of methane flowrates we tested: 0.238 L/min, 0.476 L/min, 2.38 L/min, 4.76 L/min, 11.9 L/min, and 23.8 L/min (Figure 7). We found that the lowest accuracies were attributed to both the highest and lowest volumetric flowrates with accuracies of +50/-15% and +21/-14% respectively, whereas higher accuracy was observed with the mid-level volumetric flowrates of 11.8, 4.76, 2.38, and 0.476 SLPM with accuracies ranging from +26/-3% to +9/-11%. Similar to the mass flowrates, we also found the highest accuracies were associated with the mid-level volumetric flowrates while the lowest accuracies were observed at the upper and lower volumetric flowrates."

Line 217-222: "We analyzed four different percentages of methane in the leaking gas for the controlled releases (Figure 7). The lowest accuracies were associated with the 5% methane gas with an accuracy of +31/-16%, whereas the highest accuracies were observed with the 10% methane at +15/-8%. The three highest percentages of methane in the leaking gas all had small negative biases ranging from -5% to -3%, whereas the 5% methane leak had a slight positive bias at +1%."

Line 238-250: "At low mass flowrates of methane (i.e., ≤ 100 g/hour), we found that smaller sized chambers were more accurate than larger chambers, with ~~median percentage errors of ±10% and ±16%~~ with accuracies of +12/-8% and +15/-19% respectively. ~~We found that t~~The usage of fans had little impact on the accuracy of smaller sized chambers at these low flowrates, with smaller chambers with fans producing ~~a median percentage error of ±9%~~ an accuracy of +16/-8% and smaller chambers without fans having an accuracy of +7/-13% ~~producing median percentage errors of ±13%~~. In contrast, ~~we found that~~ the usage of fans was important for the accuracy of larger chambers at ~~these~~ lower mass flowrates.~~, with~~ Larger chambers with fans ~~producing a median percentage error of ±10% and larger chambers without fans producing a median percentage error of ±27%~~ had an accuracy of +4/-30% and larger chambers without fans had an accuracy of +48/-19%. In terms of chamber shape, ~~we found that~~ at low flowrates smaller cylindrical chambers had an accuracy of +1/-11% compared to small rectangular chambers which produced an accuracy of +15/-3% ~~produced a median percentage error of ±8% compared to small rectangular chambers which produced a median percentage error of ±14%~~. For larger chambers at low mass flowrates, we ~~found~~ observed a contrasting result with large rectangular chambers producing an accuracy of +6/-16% and large cylindrical chambers producing a median percentage error of +24/-48%. ~~a median percentage error of ±14% and large cylindrical chambers producing a median percentage error of ±33%. Overall, we found that at these smaller mass flowrates of methane, small cylindrical chambers with fans~~

~~produced the lowest median percentage error of ±10% compared to ±16% from large chambers."~~

Line 251-258: "We observed similar results for optimizing chamber configurations for high methane mass flowrates (i.e., ≥100 g/hour). We found ~~observed~~ that the usage of fans was critical for measurement accuracy for larger sized chambers at these higher mass flowrates of methane. Larger chambers with fans had an accuracy of +4/-4% ~~a median percentage error of ±4%~~ compared to larger chambers without fans which had had an accuracy of +66/-35% ~~a median percentage accuracy of ±50%~~. For chamber shapes, ~~we found that~~ cylindrical chambers were more accurate than rectangular chambers with an accuracy of +6/-14% compared to +26/-15% from rectangular chambers ~~with median percentage errors of ±13% compared to ±17% from rectangular chambers~~. At ~~these~~ higher mass flowrates of methane, we found that large cylindrical chambers were highly accurate at +2/-3% of the true methane flowrate ~~with fans produced the lowest average percentage errors of ±3%~~. "

- **In the discussion section (e.g., Page 13, lines 289-296), some more direct comparison with previous studies could be useful. Is a median percentage error of 14% consistent with similar previous studies? Similarly, how does the quantification performance of the static chamber method as assessed here compares with other indirect quantification methods?**

This is an excellent suggestion, and we have now added some additional text in the discussion that comments on previous controlled release experiments and the expected quantification errors. We chose one paper for mobile surveys and one for aircraft based surveys, which gives a rough idea of where the accuracy of indirect methods lies.

Line 290: "Our results show that the static chamber methodology can quantify methane emissions ranging from 1.02 g/hour to 512 g/hour with a median percentage error of ±14%. ~~Overall, we found a small bias towards the underestimation of methane flowrates which is similar to prior studies (Lebel et al., 2020, Pihlatie et al., 2013).~~ In comparison to indirect methods, Johnson et al., 2023 state that their aircraft-based method has a multi-pass uncertainty range of -46/+54%, which roughly corresponds to an absolute error of ±50%. In von Fischer et al., 2017, they state an uncertainty range of -24/+32% after five mobile survey passes, which roughly corresponds to an absolute error of ±28%. With regards to other controlled release tests on static chambers, we do find that our median uncertainty of ±14% falls within the 10-20% range reported by Lebel et al. 2020 and Pihlahtie et al. 2013."

- **On Page 13, lines 297 to 307, it is not clear, and was not tested here, whether static chamber methods can accurately quantify emission rates beyond e.g., greater than 1,000 g/h. Is there a threshold beyond which this method does not work? If so, there needs to be a discussion of that limitation here, otherwise the reader is left with the impression that this method can be used to quantify all kinds of emission rates > 100 g/h.**

Theoretically there is no upper limit, although practically the upper limit would be somewhere in the range of 250 kg/hour based on the largest chamber ever used in literature (~32,000 L by Lebel et al., 2020). Given a methane flowrate of 250 kg/hour, it would take ~5 minutes for a 32,000 Litre chamber to reach steady-state. We have since added additional text to the manuscript to reflect this limitation.

Line 300-303: "Methane concentrations within a smaller chamber can also rapidly reach explosive levels which can pose safety concerns if the environment is not intrinsically safe (Riddick et al., 2022), but these risks can be minimized at little cost to accuracy if fans are omitted. Furthermore, intrinsically safe methods of chamber mixing such as external pumps could be used to mix air within chambers, regardless of the size of chamber. Theoretically, there is no upper methane flowrate limitation of the static chamber method, and utilizing large chambers such as the ~32,000 L chamber used in (Lebel et al., 2020) could theoretically quantify methane flowrates in the 100-200 kg/hour range. However, there are practical limitations to directly measuring components emitting methane at these high levels, the most notable being safety concerns and access issues (e.g., measuring flare stacks and liquid storage tank unloadings). Another factor to consider is the time to reach steady-state. Enclosing a high methane emitting source within a smaller chamber causes methane concentrations within the chamber to rapidly reach steady-state, essentially creating a dynamic chamber, which we do not test in this work (Pedersen et al., 2010, Levy et al., 2011)."

- **The visualization of the percent error (Equation 2, Figures 3-5) is potentially misleading. Because these are presented on an absolute basis using equation 2, Figures 3-5 could be interpreted that all quantified emission**

rates are greater than metered rates. One could assume that there were quantified emission rates that were less than metered rates, which would lead to a negative percent error, in some cases. The authors should consider revising Figures 3-5 to include those values that were quantified both above (positive percent errors) and below (negative percent errors) the actual metered emission rates.

We have since changed these figures as mentioned in an earlier comment to this reviewer. We hope the new changes address these concerns.

- **Line 19-20, please include a reference for the Global Methane Pledge.**

Done, reference added.

- **In the introduction, suggest including specific examples of "components" that can be quantified using the static chamber method (e.g., wellheads at oil and gas well sites). Also, can static chamber methods quantify total facility level emissions, which could be thought of as an aggregation of emissions from individual methane emitting "components" at the facility?**

Agreed, we have added examples of different components in the Introduction. We have also added a sentence that the static chamber method would not be able to quantify facility-level emissions through a single measurement, but could if enough measurements from sites/components are made. Although, there is always a potential to "miss" an emitting component/site using direct measurement techniques, which we have already highlighted in the text.

Lines 45-48: "Methane sources can be classified as component, site, facility, regional, and global level sources in order of increasing spatial scales (NACEM, 2018). As an example, a valve on an abandoned oil and gas well would constitute a component level source whereas all abandoned oil and gas wells in the Appalachian basin would comprise a regional methane source. The advantages of methane inventories created from component level measurements are high resolution and easy comparisons to regional inventories ,which are predominantly made using component level data (U.S. GHGI, ECCC GHGI), where specific discrepancies can be identified (Rutherford et al., 2021). Indirect measurements can be used to measure ~~a large number of sources quickly and efficiently~~ methane emissions at site/facility/regional levels, but have higher limits of detection when compared to direct methods and additional challenges related to source attribution at the component level. On the other hand, direct measurement methods are labour intensive and ~~typically limited to measuring emissions at a smaller scale~~ can omit methane sources when scaling up measurements to facility/regional/global levels , but can quantify and attribute methane emissions at the component level.”