

Response to Reviewer 2 Comments (RC2):

Reviewer comments are in italics, our responses are in plain text. Changes to the main text are included in red.

Top-down (atmospheric) and bottom-up (inventory) methods of determining greenhouse gas emissions are complementary approaches, with the goal of providing feedback as to the most effective mitigation strategies. There is a significant discrepancy between these approaches for methane, and in-home emissions are probably a major contributor. The EPA methane inventory now includes an estimate of these emissions, but the emission factors are all based on one study. This manuscript documents controlled release experiments to optimize the methods to determine these emission factors and to quantify the uncertainty using the mass balance method which quantifies whole-house emissions rates as a sum of all of the leaks in the house. This work is an important first step, laying the groundwork for future studies of different types of homes in different cities (with different types and ages of infrastructure) and in different climates. The manuscript is well-organized and well-written and I feel it should be published with very minor changes.

We thank the reviewer for their detailed reading and review of the manuscript.

Abstract: Don't need to specify "recently" AND "latest".

Yes --"recently" has been removed

Section 2: Confusing wording to discuss the overall 3-part plan in the previous paragraph and then specify measuring whole-house emissions rates of methane (part 2 of the plan) as the first sentence of the methods. Suggest changing "We measured whole-house emission rates of methane in a manufactured ..." to a more general statement along the lines of "We performed the controlled experiments described above in a manufactured house..."

This change has been made as suggested.

Section 2.1: "A flush pump pulled air at 10 L min⁻¹ through whichever inlet line the valve was set to." Awkward wording. Maybe "through the appropriate inlet line"

This change has been made as suggested.

I'm confused about the flowrates but I think it's just that there are three total tubes: indoor for CO₂/CH₄, outdoor for CO₂/CH₄, and indoor for SF₆. First two lines: 10 L/min flush, CRDS sipping. Third line: PTR-MS flowrate of 15 L/min with inlet flowrate of 120 mL/min. Should this be PTR-MS flush rate of 15 L/min with inlet flowrate of 120 mL/min?

This was poorly worded. The line has been edited regarding the PTR-MS flow as follows:
“The PTR-MS sampled SF₆ through 8 mm (inner diameter) PFA tubing with an inlet flowrate of 120 mL min⁻¹ from a line that was flushed at 15 L min⁻¹.”

The gas heating component of the furnace and the HVAC gas heating element are the same things, so the sentences in lines 113-115 in this section were confusing to me. I think the tracer experiments were in October(?) and then the heating element was turned on in December during quiescent emissions measurements.

Exactly. This has now been re-worded to be clearer. “While the HVAC fan was on for all experiments, the gas heating furnace was turned off for the tracer experiments, which ended in October. The gas heating was turned on in late December,…”

How far away is the weather station in Arlington, VA?

The weather station in Arlington, VA is about 27 km away. We have now obtained weather data from the local airport only 6.8 km away from NOAA ASOS, and are using that for pressure as well as for wind speed in this new draft. Noted now as “the outdoor pressure from a weather station at the Montgomery County Airpark 6.8 km away (NOAA, 2024) was used for both indoor and outdoor molar calculations.”

It's obvious, but consider specifying that the windows were closed.

Yes- now added to the methods: “All tests were performed with all windows and doors closed”.

Section 2.3: You used mean indoor, but median outdoor. I'm not sure which is correct here. The spikes are real, so isn't the mean more appropriate since some of that air exchanges with the indoor air? During the afternoon, it's hopefully a small difference. How different would the results be if you used mean for outdoor?

We thank the reviewer for this comment. We originally used the median in the assumption that the outdoor spikes were localized near the outdoor inlet and would not have infiltrated into the house. However, we had no evidence of that and now believe it is just as likely that the spikes were caused by elevated methane around the entire house that likely entered the space. We have revised our calculations using the mean of the 5-minute outdoor samples, and find a (very) slight improvement in our statistics, with the overall RMSE dropping to 19% from 20%. All figures and tables have been re-generated, and the text revised accordingly.

Section 2.6: For the first experiment, does the ERquiescent have to be taken into account? I see that you do mention that in Section 3.2 (This should be moved to the methods.)

We have now added a sentence in 2.6: “We note that the average $ER_{\text{quiescent}}$ from the second experiment was subtracted from the total ER in the first experiment to determine $ER_{\text{injection,MB}}$.”

Section 3.1 You didn't find any correlation with wind speed like Nabinger et al did?

We have now obtained wind data from a local airport and indeed we do find a correlation between wind speed and ACR now included here: “We also found that without any mechanical ventilation, the ACR showed a positive correlation with both the indoor-outdoor temperature difference ($R^2 = 0.44$) and wind speed at the nearby airport ($R^2 = 0.41$).”

Fig 6: maybe should expect error bars to be larger on the blue symbols (low ACR)

This was a suggestion we considered but ultimately we have decided to retain the total average RMSE (19%) for these error bars. While we calculated the errors (RMSE in Table 1) for the different ranges of ACR, we did not evaluate all possible ACRs because we could not generate ACRs smoothly across the covered range. Therefore, we hesitate to assign a different uncertainty for different ACRs arbitrarily because we do not really know the beginning or end of the range where the ACR is less certain. We hope to be able to better ascertain the relationship between ACR and error in the future.

Fig 6: The average seems lower in October than the rest of the analysis period.

This seems to be the case, but we do not really have enough statistics to analyze a possible reason for this, as the quiescent ER in October was only measured a few times.

Line 302 in Section 3.3 Consider rewording “The electric house cooling system was in operation with the heating system disabled for most of the study, including for all the tracer injection experiments.” to “The electric house HVAC system was in operation for all of the study, with the heating system disabled for most of the study, including for all the tracer injection experiments. “ or something along those lines.

This change has been made.