Reply to reviewer 2

September 5, 2026

A deep learning-driven emission estimator utilizing a mixture of experts for local wind speed situations applied to high-resolution methane imagery

Plewa et al.

The authors would like to sincerely thank the reviewer for their expertise and time in reviewing the first version of the manuscript and their helpful comments. We considered every comment carefully to improve the revised version of the manuscript. In the following we respond to the comments and suggestions of reviewer 2. The reviewer comments are marked as "RC2" and appear in boxes. The authors' responses are indicated with "AC".

Reviewer 2

RC2:

> The abstract and conclusion state that the MoE network achieves a MAPE of 5.65% above 100 kg $^{-1}$ and a 40% improvement over the previous implementation. This is a clear improvement relative to Plewa et al. (2025), which reported MAPE values of 10.29% above 40 kg $^{-1}$ and 9.48% above 100 kg h^{-1}. However, the present evaluation uses the same synthetic data. They do not yet support broader claims such as "new state-of-the-art emission estimation" or robust applicability across a wide range of real-world settings. The authors should either narrow this language or add other validation.

AC: We are sorry for the lack of clarity of our statement and added a clarification that explicitly states that the statement describes the performance on the realistic synthetic data such as the data presented in Jongaramrungruang et al. [2022]. We further included a statement regarding future work that should make the separation between results on synthetic data and on real measurements clearer and emphasizes that there are difficulties contained in real-world data that are not represented in a study with synthetic data.

RC2:

> Figure 3 shows predictions and ground truth strikingly close to the 1:1 line. While encouraging, this result is also likely to raise concerns that the model is evaluated within the assumptions of the LES simulator and augmentation pipeline. Real methane plume quantification is affected by retrieval noise and bias, plume masking, heterogeneous surfaces, and so on. I strongly recommend adding a controlled-release test set, ideally including a Stanford-organized controlled-release experiment such as the TADI/AVIRIS-4 campaign, or another appropriate airborne imaging dataset. Even a limited blind external evaluation would substantially strengthen the manuscript for both academic and operational audiences. The key question is whether the router and the emission experts remain informative when the plume originates in the real atmosphere rather than in the same simulation used for training.

AC: We added results for three controlled release experiments to the manuscript (page 17 line 335 to 353) that were previously used in Jongaramrungruang et al. [2019] and Jongaramrungruang et al. [2022] to provide a first small test of the performance on real data. However, we would like to acknowledge that these scenes do not replace a full study to test the performance under the much more demanding circumstances real measurements provide and the added statements in our conclusion hopefully reflect that. Further we would like to point out that in this demonstration of the methodology we did not include preprocessing steps that might be necessary or beneficial for a successful application to real-world measurements as demonstrated in Ouerghi et al. [2025].

RC2:

Release code, models, and reproducibility materials. This is the most important practical revision. The manuscript proposes a machine-learning estimator whose value depends on independent reproduction and comparison. The data-availability statement currently says only that data are available upon request, and I did not find a code-availability statement. The authors should release public code, trained model weights, inference and training scripts, configuration files, and enough synthetic-data generation information to reproduce the principal figures and tables.

AC: We are releasing the data publicly together with this study. Furthermore, we are working towards a study on real-world applications and there is also work being done for an operational release. This upcoming work will contain a public and broadly usable release of the code for operational users. However, until those more user-friendly releases we would rather encourage people to get in touch with us, as the focus of the current study is on methodological development and not operational deployment.

References

S. Jongaramrungruang, C. Frankenberg, G. Matheou, A. K. Thorpe, D. R. Thompson, L. Kuai, and R. M. Duren. Towards accurate methane point-source quantification from high-resolution 2-d plume imagery. *Atmospheric Measurement Techniques*, 12(12):6667–6681, 2019. doi: 10.5194/amt-12-6667-2019. URL https://amt.copernicus.org/articles/12/6667/2019/.

Siraput Jongaramrungruang, Andrew Thorpe, Georgios Matheou, and Christian Frankenberg. MethaNet – an AI-driven approach to quantifying methane point-source emission from high-resolution 2-D plume imagery. *Remote Sensing of Environment*, 269:112809, 02 2022. doi: 10.1016/j.rse.2021.112809.

E. Ouerghi, T. Ehret, G. Facciolo, E. Meinhardt, R. Marion, and J.-M. Morel. Tightening up methane plume source rate estimation in enmap and prisma images. *Atmospheric Measurement Techniques*, 18(18):4611–4629, 2025. doi: 10.5194/amt-18-4611-2025. URL https://amt.copernicus.org/articles/18/4611/2025/.