Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2022-288-RC1, 2023

Comment on gmd-2022-288

Anonymous Referee #1

Referee comment on "Segmentation of XCO_2 images with deep learning: application to synthetic plumes from cities and power plants" by Joffrey Dumont Le Brazidec et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-288-RC1, 2023

This manuscript presents a work of developing a deep-learning-based model for plume segmentation of XCO2 over cities or power plants in European countries. They evaluated the model for model generalization on new data from the same region and model extrapolation on unseen data from another region. The results indicate the proposed segmentation model outperforms the usual segmentation technique based on thresholding.
In general, the presentation of the paper is clear, and the potential of this technique is well-suggested. However, further explanation is needed on how this technique can be applied to estimate emissions from satellite imagery.

Detailed comments:

1)In the introduction section,
- The additional reference is needed for that NO2 can be a proxy to CO2 and with NO2, the plume detection capabilities are significantly improving.

- Since CO2M is a satellite mission, the author is considering applying this technique; a more detailed explanation of CO2M is needed, such as the spatial resolution, channel information, etc.

2) In the 2.2 section, page 5.
- The data for Paris are selected for Jan., Mar., and Aug. Is there any specific reason to use these three month?

- How much has the results performance improved using data augmentation techniques? The following paper introduced the data augmentation technique for weather applications considering major wind direction. Like this, have you considered the domain

characteristics in data augmentation methods?
"Seo, Minseok, et al. "Domain Generalization Strategy to Train Classifiers Robust to Spatial-Temporal Shift." arXiv preprint arXiv:2212.02968 (2022)."

3) In the 3.4 section,
- The results showed when the concentration is low or signal-to-noise is small, the performance is significantly degraded. The author mentioned NO2 is helpful for that in the introduction section. Then, why is NO2 data not used as an additional input to solve this problem?

- In the deep-learning approaches, the data split is important. Generally, the training and validation dataset are randomly split, while the test is separated from the training and validation. It would be best if you used separate datasets, not days in the middle of the same month used in the training dataset. And please indicate how many datasets are in each training, validation, and test dataset.

4) In the results,
- Most plume smoke shapes are long-tailed, and when the smoke does not spread and gathers in the middle, the segmentation results are not as good as those from long-tailed shapes. There has been a bias towards the plume shape. It seems necessary to analyze whether the result of having a higher wbce score was influenced by the shape of the plume.

- How you get the emission amount in the Figure 13.