Reply on RC1

Douglas Finch et al.

Author comment on "Automated detection of atmospheric NO2 plumes from satellite data: a tool to help infer anthropogenic combustion emissions" by Douglas Finch et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-177-AC1, 2021

General comments

The manuscript entitled 'Automated detection of atmospheric NO2 plumes from satellite data: a tool to help infer anthropogenic combustion emissions' by Finch et al. presents a convolutional neural network to identify plumes of NO2 from TROPOMI. The authors claim that the algorithm can be used for detecting individual plumes (urban, oil and gas production, power plants) and the distribution of the detected plumes was compared with an anthropogenic CO2 emission inventory. I found the approach is appealing, especially with regard to combining VIIRS data to sort out locations of open biomass burning. Furthermore, the attempt to correlate NO2 plumes with anthropogenic CO2 emission can be an interest of many readers in AMT. Notwithstanding the possible global application of such algorithms, I would suggest drawing conclusions more carefully by stating possible false detections of plumes caused by either the proposed model itself or the TROPOMI retrieval algorithm.

This is a good point. Since receiving this comment from the initial review, we have increased the text dedicated to false positive detection. We have now added this caveat to the concluding remarks to ensure the reader appreciates this point.

Specific comments

Line 89. I am not sure whether this 'active fire' (VNP14 data) can be identical to 'open biomass (or fossil fuel) burning. Maybe further explanations or rationales may be useful.

The reviewer is correct in stating that the active fire product is not identical to open biomass or fossil fuel burning. We use this product as a proxy for fire and have amended the text to clarify this point.

Line 114. Can you please elaborate regarding the random drop of 50% of the intermediate features (not the data)? As far as I understand, this random dropout is not learnable. This means that you could actually reduce the number of convolutional layers before dropping the features randomly.

The reviewer is correct that this should be "randomly drop 50% of the features (or layers)" not the data. This has now been corrected in the text.

Randomly dropping a subsample of the features is a common and simple way of preventing overfitting of the model. This has the effect of thinning the network which in turns requires the subsequent layer to apply more or less weight to the inputs of that layer. This forces each node in the layer to specialise, helping the model to become more general and prevents nodes from co-adapting which can lead to overfitting. Although reducing the number of layers would also thin the network, this would give less opportunity for the model to learn features in the image. We have now included a reference to this paper (https://jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf) in the manuscript if readers wish to know more about the drop out stage. There are a huge number of different configurations possible when creating the model and we came to this configuration through multiple iterations and changes. It may be possible to refine the model further in both accuracy and efficiency, however this would be a task for future work.

Line 122. What does this mean 'individually normalized'? Does it mean it was normalised per image? If it is true, doesn't it increase the possibility for false detections? What happens if you normalise TROPOMI data globally? What is the benefit of normalizing per image?

The data was normalized per image. This has the potential to increase false detections if the background noise resembles a plume shape. However, the alternative is to normalize the data relative to the highest possible value in all the data and then the model will struggle to detect small sources, increasing the likelihood of false negatives. This point has now been made clearer in the text.

Line 125-140. A machine learning algorithm is basically 'training data' itself. It seems that the training data were selected by 'crowd sourcing', and then by authors. Is it correct? If yes, why is that? Why not using actual distribution of plumes (or several known plumes)? Please discuss this point.

The dataset used to train the model was not crowd sourced. We tried to gather a dataset via crowdsourcing and found that it introduced a lot of uncertainty into what was considered a plume (which would then be passed into the model). In the end, we developed our own training set for these data as we are unaware of any dataset that matches the requirements of 1) having enough data of the correct pixel dimensions across the globe with an equal number of plumes/no plumes in the image, and 2) an assortment of different shapes and orientations.

Line 232-234. '2019 ODIAC emissions were used for January-June 2020' – How about the effect of COVID lockdowns during 2020? Can you also mention about this?

We have added the following to the manuscript: "We do not anticipate that the COVID-19 related lockdowns of 2020 will significantly impact our results as the reduction in CO$_2$ emissions were less than expected \cite[]{Tollefson2021}"

Line 189-290, 287-290, and Figure 7. I would suggest to examine carefully these detected clusters with other data sources (even Google maps). Couldn't these be possible errors from TROPOMI retrieval algorithms? For instance, reflection from salt lakes, solar panels,,,,

etc.. ?

We examined these data sources using Google Earth Images as mentioned on line 275, and present our hypothesis about what they could be although the satellite imagery on Google Maps can be several years out of date. Reflections from features such as salt lakes and solar panels may have made it through the TROPOMI quality control process, however it is unlikely that these features would show up as plume shaped anomalies and therefore would be less likely to be picked up by the model. Further examination with more up to date sources (e.g. Sentinel-2 imagery) is outside the scope of this paper but this demonstrates the potential of this method to locate new sources. This has also now been emphasised in the modified text.

technical corrections

Line 83. 'the the Copernicus'

This has now been corrected.

Line 85. the spatial resolution of TROPOMI has been changed since 6 August 2019 (https:/ /sentinels.copernicus.eu/web/sentinel/data-products/-/asset_publisher/fp37fc19FN8F/cont ent/sentinel-5-precursor-level-2-nitrogen-dioxide)

Although the spatial resolution of TROPOMI was increased to 3.5 x 5.5 km as of the 6th of August 2019, the original resolution is still available to use for all dates. We decided to keep with this for consistency across the study period as the model may need retraining on different resolution data. This is something to consider for future developments. We have now included this information in the manuscript.

Line 106. Is '(progressively incomprehensible to human)' part necessary?

We believe this informs the reader that the features found by the model are not necessarily obvious or comprehensible and therefore makes connections which go beyond what a human might be able to do.

Line 186. 'overpass time of 1330' to 'overpass time of 13:30'

This has been changed

Line 301. 'and then and then displays'

This has now been corrected.