To the editor,

The authors have really made considerable efforts to improve their manuscript. They present a very interesting study which is now much more readable and robust, with a stimulating discussion. The assumptions and caveats of the analysis are also more transparent and the limitations clearly stated. I consider sections 1 to 4 of excellent quality. I still believe that the discussion in section 5 has more to give. I recommend the manuscript to be accepted for publication after the following comments are satisfactorily addressed

1. Implications of this analysis in light of future satellite missions

There is a missed opportunity to connect this paper to forthcoming satellite missions, of which plume detection of NO2 will be a main focus, which is really a pity. GOSAT-GW (1 km), CO2M (4km), TANGO (300 m) and CO2Image (50 m) will provide a new suite of sensors with higher spatial resolution, increasing the capability of plume detection. I **strongly suggest** to the authors that they, at the very least, include a paragraph in the discussion on addressing the question of spatial resolution, and in doing so highlight the importance of their work to the science community.

The real question is: to what extent do subpixel variations in the surface albedo and elevation affect plume detectability? Such questions are of high value at this time, reflected in the literature (Weimer et al. 2026 for CO2M https://egusphere.copernicus.org/preprints/2026/egusphere-2026-1458/). While it is beyond the scope of this paper to provide an analysis on this issue, it should certainly be the focus of a follow-up paper. A comparable study with a sensor of higher spatial resolution would go a long way in solidifying the results of Figure 9. It may well be that the same top 5 features also govern plume detectability in these new missions (higher emissions would still lead to more detectability; higher albedo leads to more signal thus higher detectability). On the other hand, surface albedo and elevation are also highly correlated with retrieval biases which might lead to the scenario that these features are less important if they are better characterised at small scales. Thus how this work generalises to finer scales is of utmost importance. Then we could get rid of the 'would' and 'should' in the author response #6.

Furthermore, a sensor with better spatial resolution implies that the field of view is also smaller. This could further help reduce interference from neighbouring power plants/cities, which would allow more power plants to be characterised and, more importantly, a higher percentage of global emissions.

Based on their results can the authors give a statement in the discussion about what would most improve plume detectability? Better forecasting models, better satellite instruments (SNR, spatial/spectral resolution), more accurate emissions reporting?

Note: Please **do not** argue that any of the features are instrument-agnostic. This ignores the interdependece of information. Detectability, by definition, depends on the amount of plumes detected by a sensor. Thus, detectability is governed by how the features listed in Figure 9 impact satellite data. The authors themselves discuss this point several times (L443, L487, L511, section 5.1, L584).

2. L162: It is important to make a distinction between sensor properties and information obtained by the sensor. Table 1 can be a bit misleading in that, while some of the information may be taken from the TROPOMI product files, these parameters are themselves, like surface pressure, prior information that comes from other databases. For example I believe the surface albedo at 440 nm comes from OMI climatology. Perhaps a note at the start of section 2.4 to clarify this.
3. Appendix E is never referenced in the main text. The result shown in Figure E1 is very important and increase substantially the reliability of the results. This needs to be discussed in the main text and it would be appropriate in the context of the emission uncertainty discussion. Can you move appendix E to the discussion in section 5.3?

Other:

Appendix A is never referenced in the main text. Please include at least one refence in the main text so the reader knows what these refer to.

Figure 3: What is orange referring to? Should this say green?

L53: Please add the references. I would mention here the study by Kurchaba et al. 2024 who studied NO2 plume detectability from ships from TROPOMI.

L214 Please include that detectability accounts for the number of satellite overpasses, as explained in sec 4.2.

L344: Can the title of section 3.3 be changed to Training Sample Filtering

L434: Can it be clarified what P90 is referring to? Is this a metric of the detection frequency as discussed in the previous sentence, or are these now detectability values? I think the second.

L454: How does detection frequency correlate with detectability? Could this be quantified in the form of a correlation coefficient? Perhaps quoted for both US and global datasets?

Section 5.3: You could give here an example of how emission estimate affects the performance of the models. For example, take a power plant that over-reports it's emissions. This would lead to the situation that you expect the detectability of this power plant to be high, however the plume detection from TROPOMI is more difficult than expected because in reality the emissions from this plant are lower. This would certainly have bearing on the accuracy of predicting detectability, especially given the importance of NOx emissions to the models (L490).

L618: Please state that these results are characterised at 5 km.

L635: As in the first iteration of this manuscript material should not be presented in the conclusion that does not appear in main body text.