The paper describes the use and improvements of the GEMS formaldehyde retrievals. Also, it shows the results of comparison between GEMS and TROPOMI, as well as ground observations. The paper is well written and provides useful information for the scientific community. Therefore, I recommend its publication, after the minor comments below are addressed.

Specific comments

- Line 27: Is the agreement substantially higher in terms of correlation coefficient only ? or the bias is much lower in this case? Any reason why the agreement is better over Northeast Asia?

- Line 41: Could you provide the ranges for the emissions and the uncertainties?

- Line 100: How do you define clean?

- Line 110: I was wondering how sensitive the background correction is to the selection of the transport chemical model?

- Line 127: Why did you select 0.4 as a threshold for the cloud radiance fraction for the clear-sky pixels? What is the % of the pixels taken into account after masking in Fig 1?

- Line 153: There is an updated version for CEDS available. Are the emissions much different compared to the version used in the study?

McDuffie, E. E., Smith, S. J., O'Rourke, P., Tibrewal, K., Venkataraman, C., Marais, E. A., Zheng, B., Crippa, M., Brauer, M., and Martin, R. V.: A global anthropogenic emission inventory of atmospheric pollutants from sector- and fuel-specific sources (1970–2017): an application of the Community Emissions Data System (CEDS), Earth Syst. Sci. Data, 12, 3413–3442, https://doi.org/10.5194/essd-12-3413-2020, 2020.

-Line 165: the meaning of the sentence is not very clear to me.

-Figure 5&7&8: The orientation of the labels in these plots is different from the others. The labels in the colorbal are written from top to bottom, whereas in these plots they are written from bottom to top.

- Figure 5: panel c shows the relative or the absolute difference (a-b as indicated in the caption)?

- Line 177-182: Could you explain better and elaborate about the determination of the fitting window?

- Line 256 and line 257 : Any reference for the reason for the highest HCHO? The peak for Hanoi by GEMS is not in spring but around September.

- Line 283: Why was the comparison done only for this site? Why in the comparison with TROPOMI you used the averaged values over pixels within a 20 km x 20 km grid box centered on the center of each city, while now you set a grid of 0.4 0.4 degrees?

-Line 295: is the agreement better just because GEMS and ground obs use the same info (same a priori profile)?

- Line 342 and 345: How much is the high positive bias?

- Line 358: I would give numbers for the correlation coefficient and the biases in the conclusions to summarise the main findings.

-Figure S4: The figure can be read easier if the sites are ordered by increasing longitude in the legend.

-Figures: The use of italics in the figures labels and units is not consistent in the main paper and the supplement.