

[General comments]

Basically, this study looks interesting and meaningful because recently the intrusion of stratospheric ozone into the troposphere has been significantly discussed related to the issue of high tropospheric ozone (particularly in Asia). Since vertical ozone profile datasets are very rare, it is time to think about the active usage of satellite ozone profiles. While it has some uncertainty, validated satellite ozone profile can provide useful information to the ozone research community, I believe. In this context, I would like to put some weight on the value of this study. However, there are unclear parts in this manuscript. Thus, sincere responses from authors about suggested comments below are very required and proper revision looks very necessary. In spite of the meaningful topic, data analysis part does not look qualified.

[Specific comments (L: line)]

- Introduction: It is not clear in this manuscript, "Why MIPAS and IASI are used in this study"? There are longer measurements of ozone in both limb (e.g., MLS) and nadir (e.g., OMI and TROPOMI) satellite missions. What can we get distinguished information from MIPAS+IASI combination?

- L47: It is not clear the meaning of "adequate horizontal resolution". Here, is it really necessary for addressing the resolution issue related to the STE?

- L49: Well, it seems that authors would mention 'relatively' higher vertical resolution of limb sounding than nadir-viewing. It is OK, but the question for "Is the vertical resolution of limb sounding is enough for detecting the STE well?" is separated one (3km vertical resolution is very rough actually to see the STM, compared to the ozonesonde measurement). Thus, the 'high vertical resolution' in line 47 looks not proper to me. I do not think that the 'resolution' is the most important motivation for combining limb and nadir viewing satellite data for the STM study. Since limb-sounding does not have the tropospheric profile, combining limb and nadir looks meaningful for the analysis of STM in the UTLS.

- L55-59: Then what is the distinguished motivation of this study compared to suggested previous studies? Especially, I do not know what the main difference is from Tirelli et al. (2025), which seems performed the limb-nadir merging.
- L64: "to evaluate and enhance" => not understanding the meaning of 'enhance' here.
- L67: The full name of 'UTLS' was already suggested in line 46.
- L69-70: Again, then what is the distinguished merit of this study different from Ceccherini (2016), that looks performed the MIPAS + IASI using the CDF method?
- L78: The full name of 'MIPAS' was already suggested above.
- L78: Please correct this sentence to indicate either the launched year, or total duration of MIPAS operation
- L126: It seems necessary to show the portal link here.
- L127: (FORLI-O3) => (FORLI)
- L227: no need to show full name of CM once more again here.
- L241-242: What is the meaning for "to tune the method"? How is it associated with Table 1? What is it tuned, why is it needed, and is it a reliable approach? Although it is OK, is it enough only with ozonesonde data in 2008, just a specific one year?
- L249: Is there any previous work showing that the latitudinal difference of ozonesonde

measurements affect the validation result? It seems that some references are needed related to the statement "while avoiding outliers due to large latitudinal differences".

- L271: The statement "For the validation over the 2009–2011 period," here does not look connected to the statement "data from the 2009-2012 period were used to validate the entire dataset" in L242.

- L276: Why here the ozonesonde data from -75 S to 75 N are used, different from the usage of data from 5 to 55 N above. I cannot get the explanation in this last paragraph.

- L278-283: What is the difference between the validation with ozonesonde (2009-2011) and comparative analyses with ozonesonde (2008)? It is not clear to me.

- L288-289: Why does the peak bias occur around 10 km height?

- L291 (and in this part entirely): What does the negative value mean? negative bias or negative ozone? The meaning of negative value, and the reason to consider this negative value should be better described in this part.

- L298-319: This paragraph is too long. Please make two paragraphs.

- L320-328: In contrast, here are 4 paragraphs having short statements. Proper length of each paragraph and smooth connection (flow) of all paragraphs are important for readers' understanding. Please revise this part carefully.

- L320: Again, there is no explanation of high error around 10 km. The reason should be explained. Also, I am not sure if the word 'bias' and 'error' is differently used or has same meaning. It is better to use consistent and accurate wording.

- Chapter 3: Most of statements are just reading figures, not including 'real discussion'. In other words, it is difficult to find the reason why errors are high at certain altitudes.