Responses to comments from Review # 2

■ General Comment

I thank the authors for a well-organized and written discussion of their new ozone profile product. They have clearly laid out the changes from the previous version and how these have led to an improved product. I would, however, like to know a bit more about the calibration choices that were made. I will appreciate if the authors can address my questions below, and where appropriate modify the manuscript text to clarify points regarding their approach.

Reply to general comments: Thank you very much for your thoughtful and encouraging comments on our manuscript. We also appreciate your interest in the calibration choices. We will carefully address each of your questions in detail.

■ Specific Comments

Comment1 Section 2.5.2 It's clear that using a measured irradiance that includes seasonal and long-term errors will result in significant ozone errors. What is not clear is the best approach to manage this problem. The authors choice to use daily solar irradiance measurements more or less forces them to find a correction for Working diffuser irradiance measurement errors. Have the authors performed a trade study that indicates this is really the optimum solution? An often-heard claim is that normalization using daily solar measurements is required in order to adequately account for detector variations and anomalies. Such an assertion must really be demonstrated for each instrument, and actually for each product. What would the ozone product performance be if the authors utilized a GEMS irradiance fixed in time near the start of the mission? Does all the extra effort creating daily corrections really improve the product compared with simply normalizing by a Day 1 solar?

➔ **Reply1.1** We appreciate the reviewer's insightful comment, and we believe the reviewer has deep expertise in irradiance calibration issues. We believe, adopting a single Day 1 reference would require assurance that radiance degradation does not occur over time and that no unintended cancellation effects with the common degradation behavior of the optical elements in radiance and irradiance measurements. Most importantly, such an approach would only be feasible if the BTDF-induced geometry effects were negligible. In our view, the use of a fixed irradiance could be a viable option once the BTDF angular-dependent issue is fully resolved and temporal soft calibration is reliably applied. The L1B team is currently preparing the development of a BTDF correction, will help make such an approach practically achievable. Based on this comment, we plan to first test the fixed irradiance approach with long-term OMI ozone profile retrievals with dense ozonesonde measurements (e.g., Uccle, with two to three launches per week). Depending on the outcome, we will then consider applying this approach to GEMS in preparation for the next version.

➔ **Reply1.2**. To optimize ozone profile retrievals, we extensively performed a trade study to evaluate different radiometric correction methods. In version 3, we ultimately adopted the use of a scaling factor for irradiance together with a soft calibration spectrum

(dependent only on the CCD dimension) applied to normalized radiance. In earlier tests, we also experimented with polynomial fitting to reduce offsets in irradiance:

$$F = F_G - \sum_{0}^{N-1} P_b(i)\left(\lambda - \bar{\lambda}\right)^{i-1} \ (N = 0 \ldots 3)$$

As a result, the first-order correction was considered most effective (Fig1.). However, its performance was not consistent across seasons. In particular, the middle-stratospheric ozone is severely underestimated across seasons (Fig.2). In preparation for version 3.0, we also tested the multiple irradiance beta data provided by the GEMS L1C team, but its application further degraded our retrievals. Therefore, we ultimately decided to apply a combined approach of scaling and soft calibration. In the manuscript, the polynomial experiment results were omitted to maintain a clear focus on the GEMS v3.0 operational version. It briefly mentions "*This decision was made because applying the baseline polynomial Pb directly to the irradiance introduced artificial structures into the spectral fitting of the normalized radiance, resulting in a significant underestimation of stratospheric ozone retrievals*."

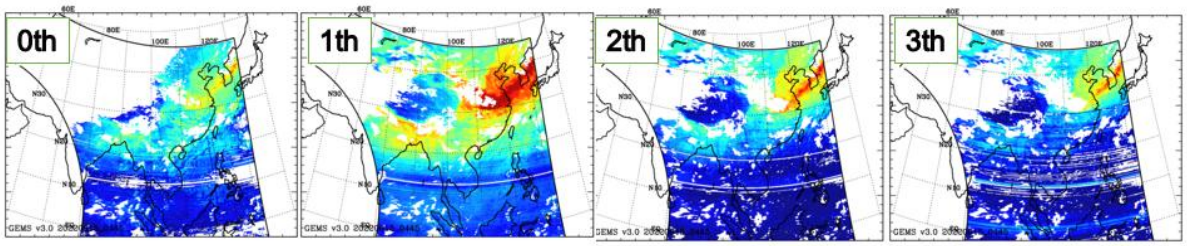

Fig 1. Tropospheric ozone retrievals using different polynomial orders for irradiance correction (note that soft calibration was not applied). Daytime is 20220615_0445.

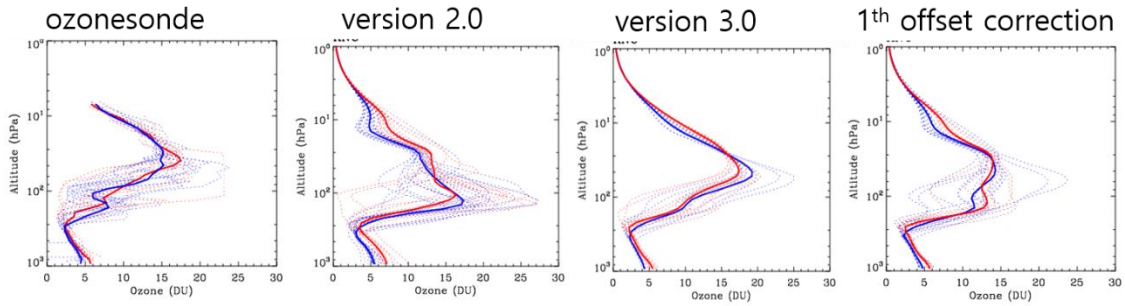

Fig.2. Ozonesonde observations during the Asia-AQ 2024 campaign (Feb/blue–Mar/red) and the corresponding GEMS version 2.0 and version 3.0 retrievals. A first-order offset correction is applied as the traditional approach.

Comment2 Regardless of which method is used to generate the solar irradiance, there is no discussion of how long-term radiometric changes in the instrument are accounted for. This is not something soft calibration is capable of dealing with. The authors fail to discuss which of the two GEMS solar measurements they are using, Reference diffuser or Working diffuser. But given the temporal density of Fig. 3, I can assume they are using Working. Why not use the Reference instead and avoid most of the diffuser degradation? BTDF issues can be dealt with by choosing Reference measurements at similar solar incidence angles. This will typically yield two useful solar measurements per year, which can then be used to interpolate in time. I can understand that this approach may not work as well for the ozone product as the method the authors have chosen. But have the authors considered such alternatives? The authors should discuss alternative approaches and why they believe the chosen approach works

best. They should also the discuss the drawbacks (i.e. long-term trends) with their chosen approach.

Rely3

- Thank you for this comment. We had no option to choose between the reference and working irradiance, as the reference diffuser data was not shared with the L2 team. Instead, the L1C team is preparing an update to the irradiance calibration that will address both geometry dependence and degradation
- However, applying a scaling correction to the daily irradiance measurements is expected to help mitigate the degradation issue. Figure 3 (manuscript) shows that the correction value dynamically varies with time and geometry. Demonstrating the improvement between version 2 and version 3 with respect to degradation is challenging due to the limited validation data from ozonesondes. Nevertheless, our companion paper (Hong et al., under review) demonstrates that GEMS version 3.0 ozone profiles, when integrated to total ozone, show reduced impact compared with GEMS version 2.1 total ozone, for which no calibration was applied, as validated against Pandora data.

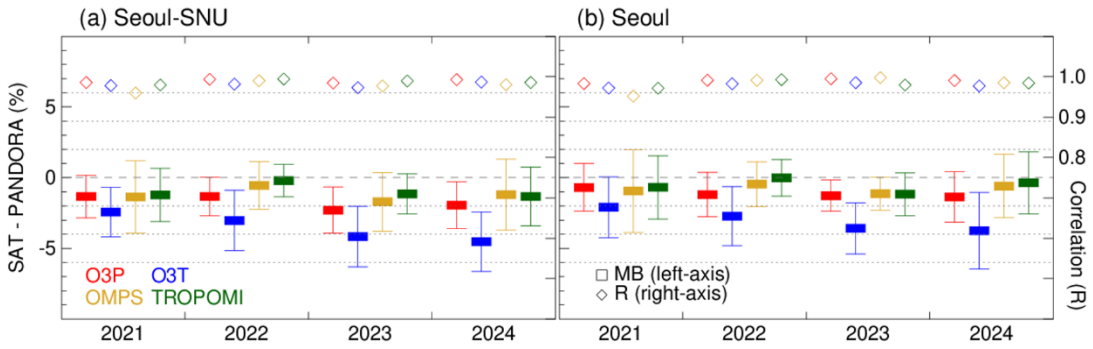

Fig.3 Comparison between satellite total ozone products and pandora observations (Hong et al. under review). O3P and O3T indicates GEMS v3.0 ozone profile and GEMS v2.1 total ozone products.

Comment3 Section 3, Line 334 I had to read this line several times before I understood that the authors are referring to a specific viewing condition when they use the phrase "sideways solar irradiance". I recommend using more precise terminology (e.g. involving SolZA) so as not to confuse the readers.

Rely3. According to this comment, we have revised Line 334 as “The offset decrease under more oblique solar and viewing geometry, although this is accompanied by a reduction in tropospheric retrieval sensitivity.”

The offset is reduced for more sideways solar irradiance and observation of the troposphere, although it has to be taken into account that the tropospheric retrieval sensitivity is at the same time reduced as well (see above).

Comment 4 Figure 13. This figure is somewhat confusing. It is not possible to see any data points from OMPS, yet there are some faint points with an unknown color (in the 330-390 DU range) that are not assigned to any mission. Please provide a better figure or clear up the confusion by describing in the figure caption.

Rely4.

- We have revised Figure 13 to clearly display the OMPS data points. During February and March 2024, there were unprecedentedly frequent rain events, and the available Pandora data cover 44 days. After quality assurance, 42 valid days were retained. As shown in Figure below, no data points fall within the 370–380 DU range.

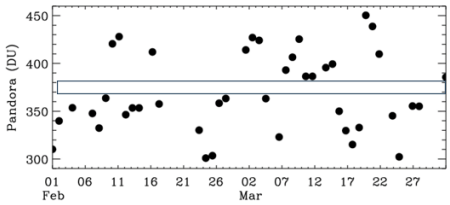

Figure 3. Pandora observations in Feb-Mar 2024.

- In caption of Fig .13, "*A total of 44 Pandora observation days was available, of which 42 remained after quality control.*" has been added.