Quality Assessment of the AC SAF GOME-2 gridded ozone profile data records – author responses

Olaf Tuinder, Peggy Tesche-Achtert and Andy Delcloo

16 Febraury 2026

In this author response, the text in black are the comments from the reviewer, in green are the answers from the authors.

1   Response to RC#2, from Juseon Bak

We thank the reviewer for her comments and have attempted to answer them below as best as possible.

1.1   Specific comments

1. L298 : Please better specify about "a seasonal dependency present in the dataset" and this was also the case in the level 2 product".

We describe the behaviour of the level-2 ozone profiles in many validation reports of the AC SAF concerning the ozone profile product (https://acsaf.org/valreps.php). There is a significant seasonal behaviour present in the dataset, which makes it necessary to do the validation over a full year, in order to exclude a bias because of this behaviour.

2. L313 said "30-50 km range, where the application of the averaging kernel is essential due to the limited vertical sensitivity of nadir-viewing instrument". In general, the nadir-UV observations better perform for retrieving stratospheric profiles than tropospheric profiles (OMI/Liu et al 2010, TROPOMI/Keppens et al. 2024), due to the Rayleigh scattering-induced vertical sensitively. Liu et al. (2010) also show the less importance of applying averaging kernels for evaluating the stratospheric column ozone.

-Keppens et al.: 5 years of Sentinel-5P TROPOMI operational ozone profiling and geophysical validation using ozonesonde and lidar ground-based networks, Atmos. Meas. Tech., 17, 3969–3993, https://doi.org/10.5194/amt-17-3969-2024, 2024.

-Liu, X et al. Validation of Ozone Monitoring Instrument (OMI) ozone profiles and stratospheric ozone columns with Microwave Limb Sounder (MLS)

measurements, Atmos. Chem. Phys., 10, 2539–2549, https://doi.org/10.5194/acp-10-2539-2010, 2010.

We agree that nadir-viewing UV instruments generally provide good sensitivity to stratospheric ozone due to Rayleigh scattering, and that averaging-kernel effects are less critical for evaluating broad stratospheric column ozone, as shown by Liu et al. (2010) and Keppens et al. (2024). In our study, the application of averaging kernels is motivated by the comparison of vertically resolved ozone profiles and partial columns in relatively narrow altitude ranges (30–40 km and 40–50 km) with high-vertical-resolution ground-based observations. In this context, differences in vertical smoothing between satellite and ground-based measurements can significantly affect the comparison, even in the stratosphere. The manuscript has been revised accordingly to clarify this point and to better reflect the role of averaging kernels in profile-based validation (Line 351 to 358).

1. Figure 6. The correlation between L3 and ozonesonde is negative for column amount of 40-50 km, consistently from GOME A/B/C. The time-series clearly show some shift. This result should be deeper addressed. I believe, it could be rooted from satellite a priori data.

We indeed observe a systematic phase shift between the satellite and ozonesonde time series in the 40–50 km altitude range, which leads to a negative correlation despite a generally good agreement of the mean vertical profiles. This behavior can be explained by the combination of (i) narrow altitude-band integration, (ii) occasional crossings of the satellite and ground-based profiles in this region, and (iii) the increased influence of the apriori profile in the upper stratosphere. A seasonal phase mismatch in the apriori representation can therefore result in an apparent anti-correlation when compared to high-vertical-resolution ozonesonde observations. We have updated the test on Line 361 to 372 to better explain the observed negative correlation.

2. We believe that a L3 product intended for long-term applications should primarily be evaluated in terms of its capability to reproduce robust trends, rather than focusing on systematic mean biases. Please include a quantitative assessment of trend consistency between the ozonesonde observations and the L3 product, to see how well the L3 product reproduces ozonesonde-derived trends.

In this paper we are mainly focus on assessing the quality of this product to verify if the product is within the committed error bounds (conform the ACSAF ozone profile ATBD). A trend analysis requires much more detail and investigation in different ozone-relevant regions, and this can be a topic for a follow-up paper. For this study we feel that it is out of scope. Having said that, we observe no significant trends in relative difference between the L3 ozone data and the reference data, so this gives confidence that the L3 product can be used for trends in a future study.

3. This work reports validation statistics based on high-resolution reference data that been convolved to the GOME vertical resolution. While this approach is useful for assessing other retrieval errors excluding the smoothing error, the resulting statistics do not represent the actual accuracy and precision experienced by end users. The authors should therefore primarily provide vali-

dation metrics without vertical smoothing, and additionally (as supplementary information) report validation metrics with vertical smoothing.

The application of averaging-kernel smoothing is essential when comparing measurements with substantially different vertical resolutions. The L3 satellite profiles represent retrieval products with limited vertical sensitivity and are influenced by their a priori profiles. Direct comparison with high-vertical-resolution ground-based observations therefore introduces representativeness errors related to vertical smoothing differences. In our analysis, this effect is clearly demonstrated quantitatively. The correlation coefficients (r) shown in the time-series panels (Figure 5, 6 and 7) are consistently higher when the satellite data are compared with averaging-kernel-smoothed ground-based observations than when compared with the original unsmoothed data. This behavior is observed across stations and altitude ranges and is further confirmed by the station-to-station consistency analysis, which shows reduced bias spread and improved agreement after smoothing. These results demonstrate that the smoothing procedure is a necessary step to ensure physically meaningful and vertically consistent comparisons between satellite and ground-based ozone profile data.

1.2 Technical comments

L216: please revise "4 % Stubi et al. (2008)"

Fixed, the references are now in brackets

L223: Please revise 13 hPa, 30-32 km.

Fixed.

L241: Please revise (Garcia et al. (2022))

Fixed the double brackets.

L294 Corresponding levels derived from ozonesonde data, following (Table 4) ==> "ozonesonde data" for improving readability.

What was meant is that the ozone at the corresponding pressure levels are compared. We added the word 'pressure' to the text to make this clear.

L294 The averaging kernels are applied *to high-resolution ozone profiles* for this evaluation.

Your text suggestion was adopted in the text.

L295. Table 5 summaries the validation statistics as function of latitude and vertical regimes specified in Table 5 for GOME-2 A/B/C, respectively. In addition, Figure 5 supplements the validation results from the mid-latitude stations in terms of time-series approach.

Your text suggestion was adopted in the text.

L319. Inconsistence of terminology between "retrieved ground-based" (L319) and sat adj (Figure 6-7). And, "retrieved" is not proper.

We have changed the terminology througout the paper to smoothed ground-based profile, groundbased and L3