We appreciate the thorough reviews and helpful comments. In the following, we present our answers and explain how we integrated changes in the final manuscript.

Reviewer 1:

In their article "Sodium Thiosulfate-Coated Ceramic Denuders for Ozone Removal in Ultrafine Particle Sampling," manuscript number egusphere-2025-6287, the authors present a device for removing ozone upstream of particle sampling. The approach - sodium thiosulfate-based removal - has been previously demonstrated thoroughly, but the form factor employed here of a multichannel denuder is novel and enables application to particle sampling. Overall, the authors present a thorough and well-developed validation and testing scheme, including both its efficacy and evaluation of the possibility of artifacts. The approach is well-founded, well-tested, and found to be highly effective. I have generally very few comments and believe the article is suitable for publication after addressing a few minor presentation points described below.

Specific comments:

1) A diagram of the field sampling setup would be helpful in Section 2.4. Especially given that the label of Figure 4A is "O3 mixing ratio Container" - it is not clear to me what the "container" is in this context.

We thank the reviewer for this helpful comment. We agree that the previous labelling in Figure 4 was ambiguous. We have therefore revised Figure 4 accordingly and replaced the label “Container” with “Sampling site”. In addition, we moved the legend to the top of the 4B to improve the visibility. To better illustrate the field deployment, we also added a photograph of the sampling setup to the Supplement as Figure S2.

2) In Figure 2, what is the "measurement point"? Based on the trend in O3_in, I assume these are just sequential points across time, so why not just label it with sample time?

Figure 2 does not represent a continuous time series, but rather a sequence of discrete measurement intervals at different inlet O_3 mixing ratios. The purpose of this experiment was not to assess denuder performance as a function of time, but to evaluate whether ozone breakthrough occurred under increasing inlet O_3 levels. For this reason, we consider a sequential representation more appropriate than a time axis.

To avoid ambiguity, we have revised the figure caption and corresponding text to clarify that the x-axis refers to consecutive measurement intervals rather than sample time. Long-term denuder performance under ambient conditions is addressed separately in the field experiments presented in Section 3.2.

L 294-296 Thus, the inlet O3 mixing ratio was incrementally increased to up to 200 ppbV in a sequence of discrete measurement intervals. As depicted in Figure 2, the outlet O_3 mixing ratio (4.4–5.0 ppbV) were indistinguishable from this zero-air baseline, indicating that ozone downstream of the denuder was below the practical detection limit of the measurement setup

Figure 2 Average O3 mixing ratios upstream (“$O_{3,\,in}$”, red) and downstream (“$O_{3,\,ou}t$”, orange) of the TSOD during consecutive laboratory measurement intervals with stepwise increasing inlet O_3 mixing ratios (40–200 ppbV). RH was simultaneously monitored and remained constant. Lines are included to guide the eye.

3) In the discussion of Figure 3, it is never discussed by the concentration is so much lower at 4 lpm, though it is shown that the difference between with and without the TSOD is not substantially. I assume this is due to diffusive losses to walls of the sampling inlet due to the longer residence time, but no such explanation is discussed or given. Would such losses be expected based on the diffusion timescale? It does not look there is a preferential loss of smaller particles though. Or is it just because of changes in particle concentrations over time?

We thank the reviewer for this comment. The purpose of Figure 3 was not to compare absolute particle number concentrations across the three flow-rate experiments, but to assess the relative difference between measurements with and without TSOD at each individual flow rate. The experiments at 30, 11, and 4 L min^{-1} were conducted sequentially rather than simultaneously, and the chamber was only loaded initially with ambient air, seed particles, and α-pinene, as described in Section 2.3. Therefore, temporal changes in chamber aerosol concentration, as well as reconfiguration steps between the measurements, may explain the lower absolute concentrations observed at 4 L min^{-1}. This does not affect the main conclusion from Figure 3, namely that the differences between measurements with and without TSOD remained small and within instrumental uncertainty at all tested flow rates. To avoid this possible misinterpretation, we have clarified in the revised manuscript that Figure 3 is intended for pairwise comparison of wTSOD and woTSOD at each individual flow rate rather than for direct comparison of absolute concentrations between the separate flow-rate experiments.

L319-L321 Since the measurements at 30, 11, and 4 L min^{-1} were conducted sequentially, the absolute particle number concentrations between the different flow-rate experiments vary and are not directly comparable.

4) In Figure 4, a legend on panel B would be helpful. Also, I note that the authors state on line 294 that worst performance is expected at lower RH (side note: I believe the R is usually also capitalized unlike in the manuscript), but it looks like the opposite is true in this panel (though indeed, performance remains excellent)

We have revised Figure 4B by improving the visibility of the legend, and we have changed the notation from rH to RH consistently throughout the manuscript. Regarding the reviewer's observation on panel B, we are of the opinion that it is not suited to assess a quantitative RH dependence of denuder performance. Our statement in Section 3.1 refers to the expected behaviour based on previous literature and was intended to explain why the laboratory tests were performed under comparatively dry conditions. In contrast, Figure 4B shows ambient field operation rather than a controlled humidity experiment. Since the O_3 mixing ratio downstream of the TSOD remained consistently at or below the detection limit during the field deployment, the figure demonstrates robust performance over the encountered ambient RH range, but it does not allow interpretation of a detailed RH dependence. We have clarified this point in the revised manuscript.

L287-291 We therefore intentionally conducted our laboratory experiments under distinctly drier conditions (reported as RH = 52.4–53.3 %), in order to test TSOD performance near the lower end of the humidity range expected during field operation. Previous studies indicate that thiosulfate-based ozone removal decreases under drier conditions, consistent with reduced water availability at the denuder surface (Ernle et al., 2023; Rynek et al., 2025)

5) Figure 4 might be clearer in a square form, since it is showing 1:1 comparisons. A few thoughts on the discussion: the 6PPD discussion is very interesting and I appreciated the quantitative yield discussion; why is there so much scatter in the PAH comparison? In particular, there are substantially differeces in the sum Chry BaA scatter - is this because one compound is more reactive than the other? Or can you correlate the deviation from the 1:1 line with ozone concentration, which would be an interesting plot to see?

We thank the reviewer for this helpful comment. We assume that the reviewer refers to Figure 5 rather than Figure 4. We agree that this figure benefits from a square aspect ratio and have revised it accordingly. Due to space limitations in the revised layout, the regression slopes were removed from the figure legends. However, as these values are reported in Table 1, we consider this a reasonable compromise.

Regarding the PAH comparison, we agree that the PAH data are less tightly constrained than the corresponding bOA data. We attribute this primarily to the substantially lower ambient PAH concentrations on the UFP filters (in the tens of pg m^{-3} range), for which variability from sampling, extraction, blank correction, and chromatographic quantification becomes proportionally larger than for the more abundant bOAs or 6PPD/6PPDq. Additional variability may arise from day-to-day differences in aerosol matrix and particle-phase accessibility under ambient conditions. In addition, the analytical recovery varied among the PAHs. For Σ(Chry +

BaA) the recovery from standard reference material was for example lower than for some of the other PAHs, which may further contribute to uncertainty in this parameter.

We appreciate the reviewer’s suggestion to correlate the deviation from the 1:1 line with ambient ozone concentration. We have tried this approach during our investigations and data analysis as well. However, because each data point represents a 24 h integrated sample collected under changing ozone, RH, temperature, and aerosol-composition conditions, we found that the present dataset does not allow a robust interpretation of ozone concentration as a single controlling variable. From our view, such an analysis would require a dedicated controlled-exposure experiment and would go beyond what can be robustly derived from the present dataset, which is beyond the scope of the manuscript.

6) This Data Availability statement is outdated. Though there is likely not much demand for validation data such as this, it is generally more accepted to include the data, at least those used to create the figures, as supplemental data, tables, or a published dataset.

The study was conducted within a project funded by the Bavarian State Ministry of the Environment and Consumer Protection (StMUV). Therefore, the release of the underlying project data requires prior approval by the funding authority.

We are happy to provide the data used to generate the figures upon reasonable request. Should public deposition of the figure data or a dedicated dataset be required by the editor or the journal, we will submit a corresponding request to the StMUV. We kindly note that obtaining this approval may require additional time.