

Response to Gabriele Pfister (TOAR Steering Committee Member) comment on “Ground-based Tropospheric Ozone Measurements: Regional tropospheric ozone column trends from the TOAR-II/ HEGIFTOM homogenized datasets” by Van Malderen et al.

This review is by Gabriele Pfister, member of the TOAR-II Steering Committee. The primary purpose of these reviews is to identify any discrepancies across the TOAR-II submissions, and to allow the author teams time to address the discrepancies. Additional comments may be included with the reviews. While members of the TOAR Steering Committee may post open comments on papers submitted to the TOAR-II Community Special Issue, they are not involved with the decision to accept or reject a paper for publication, which is entirely handled by the journal’s editorial team.

Comments regarding TOAR-II guidelines:

TOAR-II has produced two guidance documents to help authors develop their manuscripts so that results can be consistently compared across the wide range of studies that will be written for the TOAR-II Community Special Issue. Both guidance documents can be found on the TOAR-II webpage: <https://igacproject.org/activities/TOAR/TOAR-II>. The TOAR-II Community Special Issue Guidelines: In the spirit of collaboration and to allow TOAR-II findings to be directly comparable across publications, the TOAR-II Steering Committee has issued this set of guidelines regarding style, units, plotting scales, regional and tropospheric column comparisons, and tropopause definitions.

The TOAR-II Recommendations for Statistical Analyses: The aim of this guidance note is to provide recommendations on best statistical practices and to ensure consistent communication of statistical analysis and associated uncertainty across TOAR publications. The scope includes approaches for reporting trends, a discussion of strengths and weaknesses of commonly used techniques, and calibrated language for the communication of uncertainty. Table 3 of the TOAR-II statistical guidelines provides calibrated language for describing trends and uncertainty, similar to the approach of IPCC, which allows trends to be discussed without having to use the problematic expression, “statistically significant”.

General Comments

The study uses the homogenized ground-based and in-situ tropospheric ozone measurements from the TOAR-II Focus Working Group HEGIFTOM (Harmonization and Evaluation of Ground-based Instruments for Free-Tropospheric Ozone Measurements) to conduct an in-depth analysis of regional trends of total and partial tropospheric ozone columns over the 1995-2022 time period. It builds on the previous Van Malderen et al. (2024) analysis which presented trends for the individual sites based on the HEGIFTOM data set.

The overarching objective of TOAR-II is to provide an updated state of science of the ozone global distribution and trends relevant to climate, human health and vegetation. As such this paper makes an important contribution to the TOAR-II objective and also provides valuable input for evaluating chemical transport models as well as guidance on calculating trends.

The study is very detailed and carefully conducted and limitations in both the measurements and statistical analysis are characterized. The authors introduce new approaches to define regions for the trend calculation (a re-analysis based correlation analysis), include the Trajectory-mapped Ozone-sonde dataset for the Stratosphere and Troposphere (TOST) in their analysis for comparison and present a linear mixed-effects modelling approach to calculate synthesized trends. Where possible, the authors compare their calculated trends to previous studies and if discrepancies arise, they are investigated and if possible explained.

While the trends in many parts could be explained by changing precursor emissions, I think the results do give motivation for further studies to understand what other drivers besides emissions might be contributing to the trends.

Thank you for your careful assessment of our study, Gabi.

Specific Comments

The paper requires careful reading because of the amount of detail discussed and the different advanced statistical methods involved. It would help for understanding to make more clear what the different methods were used for and why the different methods, specifically as it relates to LMM and DLM. Otherwise readers who are not fully familiar with advanced statistics might find it challenging to follow the paper.

This is really a common comment raised by all the reviewers. So, we included a table that summarizes the most important differences between the LMM and TOST approach, and why different statistical tools are used for either the TOST or the HEGIFTOM dataset. In the response to Referee #1, you will find the new Table 4.

The figures contain a lot of information and this makes them hard to read. Maybe there is a way to simplify some of them and keep the current complex figures in a supplement.

We are aware that the figures contain a lot of information. We invest quite some time in making them clearer in terms of font/symbol sizes, line thickness, etc. and to make them more consistent between the different techniques.

E.g.,: Figure 6 & 7: It is not clear to me what the difference between filled and open symbols is. Maybe it would help to discern the different regions (which I found challenging because some of the colors look very similar) and the region-specific colors and symbols could be used for all the plots in a consistent manner (at the moment squares are used for median trends and circles for the difference plots).

We did not use filled and open symbols in those plots anymore. We also consistently use the same colours and figures for similar regions for TOST and LMM in the manuscript's figures now.

Figure 8: Increase the font size for the labels

Done. Thank you for pointing this out.

Figure 10&11&12: Maybe increasing the font size and making the symbols larger could help with the readability of the graphs?

Done. Thank you!

Figure 3: It might be of interest to also include sites that were not used in the analysis in a different symbol.

On Figure 3, we only included the sites that actually have been used in the analysis, which complies with the listing in Table 1. The figure caption now reads as "Map showing the 19 different regions (squares, numbered) and the individual HEGIFTOM ground-based sites (black dots) that are used for the regional synthesized trends calculation using the linear mixed-effects modelling (LMM) approach. An overview of those different regions and included HEGIFTOM sites is given in Tables 1 and S1."