Response R2: “Intercomparison and Evaluation of Ground- and Satellite-Based Stratospheric Ozone and Temperature profiles above Observatoire Haute Provence during the Lidar Validation NDACC Experiment (LAVANDE)”

The study presents a detailed comparison of the stationary lidars at OHP with a series of other data sets, in particular a mobile lidar system from NASA. Also ECC, MLS, SABER ozone and temperature profiles are considered. The comparison has been performed “blind” by an impartial expert, which is a very interesting approach. For the comparison, a wide range of visualizations is presented reaching from mean profiles, over time series and scatter plots to correlation profiles. This gives a detailed insight in the behavior of each data set but makes the paper a bit lengthy. The study also evaluates in detail the uncertainties of each data set, which is a very interesting aspect and should be done more in validation studies. Validations studies such as the one presented here are important contributions to understand instrumental differences and to obtain consistent long term data sets. Methods and results are well explained and the paper follows a logic structure. I recommend the paper for publication in AMT and provide below minor comments for the authors to consider.

Thank you for your comments regarding the intercomparison of satellite and ground based lidar measurements. This is a very important (and difficult) problem for all of us. It is our hope that NDACC data, in particular our long data set of lidar temperatures and ozone measurements can be presented in a way that is useful for long-term comparisons with space based instruments. If we can develop robust procedures for determining coincidence with satellites perhaps near-real time comparisons can be made and automatically uploaded to the NDACC website.

Minor comments In the effort to identify co-located profiles between lidar and satellite the authors allow a time difference of up to 12 hours. Given the diurnal cycles in temperature and ozone in the stratosphere and mesosphere, this seems too tolerant. Have any effects related to tides and diurnal cycles been corrected? An analysis of the distribution of the time differences would be helpful to convince the reader that systematic biases are not a consequence of diurnal cycles.

The 12 hour window mostly applies to SABER measurements. It is difficult to achieve very close spatio-temporal matching and have a sufficient number of coincident measurements to have a statistically meaningful comparison with a non-sunsynchronous satellite. We only have 28 nights of lidar measurements with which to conduct the intercomparison exercise. Wing et al. 2018 discusses the pros and cons of this trade off in more detail. Another study in JGR:Atmospheres by Dawkins et al. 2018 “Validation of SABER v2.0 operational temperature data with ground-based lidars in the mesosphere-lower thermosphere region (75–105 km)” Also has a very good discussion

on this topic (however given the upper mesospheric and lower thermospheric focus it is not directly applicable to the work in LAVANDE)

For MLS there are generally one or two overpasses included in the nightly coincidence criteria. Generally around 1:40 am local time plus or minus approximately 99 minutes. With the sun synchronous satellite the geographic constraint is more strict than the temporal constraint.

The weighted mean on p6, l16 takes into account the typical wind speed in the stratosphere. What is the justification for this? Diurnal cycles in temperature and ozone are not driven by advection but by photochemistry and tidal waves. Please comment.

Our objective here is to try and weight profiles by both distance and time. From a simple mathematical standpoint we need to assume a wind speed to make the units agree. From a geophysical standpoint, we can have cases where there are two or more satellite overpasses in our geographic coincidence box at different times. For example, we could have a lidar measurement centered around local midnight with a satellite overpass 1000 km to the west and a second satellite overpass 100 minutes later 700 km to the east. However, during the period between the two satellite overpasses the air is generally advected from west to east over the lidar site. (See our reply to R1 with respect to tropospheric variability and local frontal systems). The use of an assumed stratospheric wind speed is an attempt to correct for the relative motion of the atmosphere with respect to the lidar station over these short timescales.

We recognise that this is not a perfect correction but given our recent experiences with validating Aeolus wind measurements using the Doppler wind lidar at OHP approximately 10 m/s is not an unreasonable assumption for late spring and summer. Khaykin, Sergey M., et al. "Doppler lidar at Observatoire de Haute-Provence for wind profiling up to 75 km altitude: performance evaluation and observations." ***Atmospheric Measurement Techniques*** **13.3 (2020).**

P6, l11: do you mean 10 to 20 matching profiles per night? With the chosen wording this is not absolutely clear

Replaced with:
“It results in between 10 to 20 coincident profiles for MLS and SABER, which are generally divided between one or two satellite overpasses, for a given night during the LAVANDE campaign.”