Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2020-396-RC2, 2021

Atmospheric Measurement Techniques Discussions

Interactive comment on "Evaluation of the New NDACC Ozone and Temperature Lidar at Hohenpeißenberg and Comparison of Results with Previous NDACC Campaigns" *by* Robin Wing et al.

Anonymous Referee #2

Received and published: 23 March 2021

General comments

This is a generally well written paper about an NDACC intercomparison campaign, augmented with results from other NDACC campaigns.

One of the main points in the paper is that the original Hohenpreissenberg lidar HOHO is compared to the NASA travelling reference lidar STROZ for the second time (2009 HOPE and 2018-2019 HOPS) and the new Hohenpreissenberg lidar HOH is now also compared to STROZ, as well as to the HOHO system. So the consistency of HOH with HOHO can be established on site for the overlapping parts of the profiles. It is

shown in Fig.2 that HOH has much greater performance in terms of range and SNR, but in the overlapping regions the systems are consistent. In order to validate HOH, STROZ is needed. Now, the general structure of the paper becomes sometimes a bit hard to follow, since the HOPS campaign are intertwined with LAVANDE results that have been published separately, which confronts the reader with a few storylines that have to be kept separate. The points brought in are certainly relevant, but it complicates the structure of the paper. In fact, this becomes clear in the conclusions: "The cross-comparison of NDACC campaign at Hohenpeißenberg Meteorological Observatory (HOPS) and at Observatoire de Haute Provence (LAVANDE) has allowed for the unique opportunity to assess potential biases in the NASA-STROZ reference lidar." My suggestion is to re-structure some of the sections to clarify this and move this material as much as possible to Sec.7.

Since the NDACC intercomparisons with a travelling reference lidar have been undertaken for some time (e.g. the references mentioned date back to 1995) it could be clearer described how the intercomparisons are generally carried out, according to an NDACC protocol, and perhaps explain how the HOPE and LAVANDE campaigns may be deviating from that protocol. There are some instances in the text that suggest there are different variants of the protocol. It would be interesting from the network design point of view to know why these variants exist.

Abstract - Remove the sentences "The previous 2017-2018 ... are reported in the companion article." - Add the main conclusions "The intercomparison exercise has confirmed that the original DWD lidar, HOHO continues to meet NDACC standards for ozone profiles at the 3% level between 16.5 and 43 km and at the 10% level between 10 and 44 km. The HOHO lidar meets the NDACC temperature standards for accuracy at the $\pm$1 K level between 18 and 70 km. The new DWD lidar, HOH, meets the 3% ozone standard between 17 and 41 km, the 10% ozone standard between 15 and 41 km, and the $\pm$1 K temperature standard 555 between 17 and 78 km." - Add "The cross-comparison of NDACC campaign at Hohenpeißenberg Meteorological Ob-

servatory (HOPS) and at Observatoire de Haute Provence (LAVANDE) has allowed for the unique opportunity to assess potential biases in the NASA-STROZ reference lidar. Possible biases may arise from algorithm initialisation choices and serve as strong motivation for another NDACC temperature algorithm paper."

Section 2 - The differences of the original Hohenpreissenberg lidar HOHO and the new lidar HOH are described. In the description of the travelling standard STROZ, it is not clear if instrument changes have been applied since the HOPE campaign in 2009. This is relevant since the consistency of the performance of HOHO is essentially compared again now in the HOPS campaign against the same travelling standard.

Section 6 - Earlier in the paper reference is made to Leblanc et al., 2016a, b, c. Are the results presented obtained using the methods described in those papers? Are results, following the blind intercomparison, processed by the proprietary processing algorithms of each group, or are they processed by a common processing code that is endorsed by NDACC? How would using a common code impact the intercomparison results for HOPE, HOPS and LAVANDE?

Small comments: - Not all readers may be familiar with Pearson's correlation coefficient. Pleaser briefly explain. - The names of the colours in the figures are a bit strange; e.g. "burnt orange", and "mustard". Why not just orange and yellow?

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-396, 2020.