Responses to RC2:

We would like to thank the reviewer for his sharp comments and suggestions, which have improved significantly the quality of our manuscript. We have addressed all reviewer's points and detailed point by point replies are provided below.

General remarks:

We thank the reviewer for his suggested references on previous comparisons or campaigns. We could not find any Papayannis et al. 2015 publication, but we assume a 2005 ANGEO paper was meant to be referred instead. We added text at the end of the manuscript, where our results are compared to previous work, namely, Wang et al 2017, Kuang et al., 2011, and Papayannis et al., 2005. We are not sure which "Trickl et al., 2015" paper the reviewer is referring to. Trickl et al. (ACP, 2015) do not report on a specific campaign but on long range transport of ozone and smoke. Consequently, we did not refer to this paper in the discussion on comparisons with ozonesondes and other lidars. However, we think this paper is relevant to the discussion on the expected impact by aerosols, and therefore we refer to it in the revised text.
Finally, we also addressed the issue of the significance of using the GLASS, raised by the reviewer. Indeed, the main reason the GLASS results are shown here is that the GLASS is used as a reference transfer. We tried to make it clearer throughout the text (see detailed responses below).

Detailed remarks and questions:

P. 2, line 11: Add TOAR reference
Reference (Gaudel et al., 2018) added.

P. 3, line 28: Merge Tables 1 and 2 and provide meaning for ON/OFF
Following the reviewer's suggestion, we merged table 1 and table 2. We also clarified ON/OFF by specifying "*DIAL wavelengths*" (ON/OFF meant on-line wavelength and off-line wavelength)

P. 5, line 17: Clarify sentence starting with "The background subtraction…"
The LMOL data acquisition system during SCOOP consisted of 2400 bins of 7.5 m each, which sums up to 18 km range. As mentioned in the article, it is now upgraded so that background subtraction can be done more safely at higher altitudes where no signal remains. For clarification, we re-wrote the sentence as follows:
"*The background subtraction value is determined from approximately the last 2 km of the data collection window*."

P. 10, line 20: why not use AMOLITE?
The AMOLITE full record is shown as a concluding figure at the end of the manuscript. We thought it would be more representative of the campaign's efforts to show non-automated measurements instead. Also, showing curtain plots with gaps, as in Fig 2, emphasizes/justifies clearly the need for more automated measurements within the network.

P. 10, line 22: change 8-9 km to 8 km
Done.

P. 15, line 15: Consider lidar differing vertical sampling resolution into homogenized SCOOP VR?
Yes, the “SCOOP” effective vertical resolution is computed based on the NDACC-standardized resolution definition described in (Leblanc et al., 2016a), and therefore takes into account the initial resolution of each lidar.

P. 12, line 3: Is fig. 5 necessary?
We understand the reviewer’s point of view here, and we agree that Figs. 4 and 5 both address the same problem of sampling. However we think keeping both figures is still very useful: Fig. 4 illustrates sampling variability by instrument, and Fig. 5 allows a clear quantitative estimate of the impact of these sampling differences.

P. 13, line 13: Why use GLASS in Sect. 4 if GLASS is not used for TOLNet processing?
Actually, as mentioned in the text (P. 13, l. 14-15) the GLASS uses all NDACC-standardized features recommended in the Leblanc et al. (2016a-b) papers. The effect of vertical smoothing and the propagation of uncertainty are indeed important contributors to the data processing evaluation. For this reason, GLASS can be used as a reference transfer across all TOLNet instruments, which minimizes the impact of the forward model in the comparisons of the simulation exercise. Documentation for the GLASS is currently under development, with probable publication in 2019. We added the reference (Leblanc, 2019, manuscript in preparation) to the text.

P. 14, line 22: little impact from aerosols
The combination of high elevation site (above most of the ABL) and short wavelength differential in the Hartley band contributes to reduce the impact of aerosols, although some effect is still expected. Also, Trickl et al., 2015 show that ozone changes due to transported smoke are mainly associated with long-range transport of enhanced ozone rather than the result of measurement contamination. To clarify all these points, text now reads as follows: *“…mostly above the boundary layer, with reduced impact from aerosols considering the wavelength differentials considered (Trick et al. 2015)*”. We also added a short paragraph on aerosol correction in the conclusion section. It reads: “*Finally, additional coordinated efforts within TOLNet are planned to provide improved ozone retrievals including aerosol corrections. Several groups (e.g., TOPAZ) have previously implemented an optional correction, and future efforts within TOLNet will concentrate on the possible homogenization of aerosol corrections across the network.*”

P. 15, line 4: Fig 8
We thank the reviewer for his comments on this figure. We are not sure if any change is being requested. Fig 8 summarizes the uncertainty budget of 4 TOLNet lidars, and illustrates the validation efforts made by comparing them to the “expected” GLASS standardized budget.

P. 16, line 24-25 : TMTOL bias at 10 km
Following the reviewer’s comment, we change the text to the following:
“*For example, the positive difference of 10-15% and 8-9% observed at around 10.5 km and 5 km altitude respectively on the TMTOL panel….*”

P. 16, line 27: 10% at 3 km
We changed the text, i.e., replaced “*at 3 km*” by “*around 3 km*” (the bias is much less than 10% at 3.5 km).

P. 16, line 27: Similarity in the tunable vs. YAG laser biases
We cannot really find a clear physical explanation for it, and therefore we decided to not comment on it. This is likely just a coincidence.

P. 17, line 8: poor sampling statistics
Yes, lower number of coincidences. To clarify, we replaced "…to *a poorer sampling statistics*" by "…to *the reduced number of coincidences*"

P. , line : text
Following the reviewer's suggestion, we added the following:
"*Addition*"

P. 16-17, Sect. 5: discussion of feature 1) 8% bias at 7 km
We actually do not observe a +8% bias for all lidars at 7 km. Does the reviewer instead refer to a +5% bias of the lidars w.r.t. the ozonesondes at 11 km?
To account for the reviewer's remark, we added the following sentence: "*Nevertheless, it is interesting to note that all five lidars exhibit a positive bias of about 5% at 11 km with respect to ozonesondes, which points out to either a negative ozonesonde measurement bias, or to a co-location error between the sondes and the lidars*"

P. 16-17, Sect. 5: discussion of feature 2) TMTOL noisy yet more powerful
Following the reviewer's suggestion, we added the following sentence:
"*Note again the noisy profile for TMTOL below 6 km, especially during daytime, as the instrument suffered from reduced signal-to-noise ratio in 2016 (see earlier discussion).*"

P. 16-17, Sect. 5: discussion of feature 3) shorter wavelength balance lower laser power
This is an interesting comment indeed, although after making a quick comparison of ozone absorption and atmospheric extinction, the use of slightly shorter wavelengths does not seem to improve the measurement compared to the YAG-laser-based systems. Instead we think the reason for compensating the lower laser power might be a higher overall receiver efficiency (including optical transmission and quantum efficiency of the PMT), the choice of spectral filters (e.g., using cut-off filters as opposed to notch filters), combined with the use of higher laser repetition rate. To reflect this, we added the following sentences: "*Finally, the good performance of the tuneable laser instruments (LMOL and TOPAZ) with respect to the other lidar instruments is noteworthy. However, it is not clear what actually balances the lower power of the tuneable lasers. Likely candidates are the overall transmission of the receivers (including optical and electronic/quantum efficiencies), the choice of the spectral filters, a higher laser repetition rate, and possibly the shorter wavelengths used, although for this latter, a quick calculation of ozone absorption and atmospheric extinction differentials yields little difference with the YAG-based systems.*"

P. 17, line 16: Fig 12, distinguish random from systematic curves, and replace profile by differences in left row
Following the reviewer's comment, we thickened all curves to make sure random and systematic components are easy to distinguish.
Regarding the left row plots: Although the reviewer has a good suggestion, and we agree that comparing regions of higher uncertainty and higher differences is important (see our reply to next comment), we

decided to keep the figure as is, for the following reasons: the ozone profiles are shown in the left column with the main purpose of documenting the shape of the ozone profile, so that changes in relative uncertainty (right column) can be easily associated with changes in ozone values. This way, it allows the separation of uncertainty changes associated with the ozone profile itself, and uncertainty changes inherent to the measurement technique (i.e., independent from the shape of the ozone at that precise moment and location). To clarify our motivation, we added to the following text: "*Showing the ozone profiles (left panels) allows to distinguish between uncertainty changes associated with ozone changes, and uncertainty changes inherent to the measurement technique itself (i.e., independent of the ozone content at a precise time and location). For example, the localized uncertainty peak for TMTOL at 5.5 km altitude (top right panel, green curve) is essentially due to the dip in ozone mixing ratio (top left panel) rather than a change in the uncertainty (top middle panel, green curve, shows a nearly constant absolute systematic uncertainty).*"

P. 17, line 31: Uncertainty at 8 km overestimated for TMTOL? Underestimated for AMOLITE at 6.5 km?
The reviewer raised a very good point here, and several sentences were added to the discussion to clarify or emphasize this.
For the case of TMTOL large uncertainty at 8 km, the reported uncertainty is an estimate of the error made when correcting the high-intensity channel for saturation (pile-up effect). The correction equation uses the photon-counter dead-time, which has an uncertainty provided by the manufacturer. The trueness of the correction equation must also be taken into account. For the two reasons above, it is difficult to provide an accurate estimate of this uncertainty, and it is not uncommon to see this component overestimated (for the sake of being conservative). To take this into account, we added the following sentence: "*The estimation of this uncertainty component depends on the photon-counter's dead-time (provided by the manufacturer), and on the trueness of the correction equation. It is often overestimated for the sake of choosing a conservative sid*e."
For the AMOLITE "contrary" case, the limited number of available daytime coincidences with ozonesondes during SCOOP mainly contributes to exhibit larger than expected differences, which does not necessarily imply erroneous uncertainty estimation. The consequence is that a co-location (or more generally, sampling) uncertainty should probably be inferred, which takes into account atmospheric variability and possibly other external factors. This is unfortunately beyond the scope of this paper, but surely worth mentioning. For this reason, we added the following text in paragraph 5.1 "…*(e.g., AMOLITE at 6.5 km),….*".

To emphasize it, we added several sentences:
"*Addition*"

P. 18, line 2: AMOLITE, LMOL, TMTOL lidar processing afterwards
In fact, the "optimization" in the lidar data processing simply results from the fact that they were somewhat "de-optimized" for Level 2 using a common "SCOOP" vertical resolution. As for the other changes, they indeed occur either before, during or just after the campaign. To avoid confusion associated with the timing of the changes, the sentence "*The new data version also incorporates some data processing refinements resulting from the algorithm validation exercise*" was deleted.

P. 19, line 1: "outstanding"
Following the reviewer's comment, a new sentence is used: "*Taken as a whole, the TOLNet lidars show excellent agreement with the ECC ozonesondes, with an overall mean bias of 0.7 ppbv or 1.7% for the*

altitude range 3-10 km, and with a root-mean-square deviation of 1.6 ppbv or 2.4%, although Table 5 shows sometimes larger bias or RMS for a single system"

P. 19 Summary: Add sentence about QA/QC and previous campaigns
Following the reviewer's suggestion, we added the following text: *"When compared to previous intercomparison campaigns, these estimates are fully consistent with the 1.2%-4% differences found by Wang et al. (2017) and 2% lidar-sonde differences found by Papayannis et al. (2005), and they are smaller than the 10-20% estimates reported in Kuang et al. (2011).*"

P. 20, conclusion : future of TOLNet and measurement below 3 km
We are not sure if the reviewer means 3 km above sea level, or 3 km above ground level. All our results were presented above sea level, knowing that TMF is at an elevation 2.3 km above sea level. In our paper, results below 3 km therefore refer to 700 m above ground. We agree with the referee that extending the measurements downward to ranges shorter than 700 m is essential for studies of the boundary layer (TOPAZ can do it thanks to their scanning transmitter mirror). Indeed, several TOLNet groups, including TMTOL and LMOL, have worked towards extending the measurements downward to about 100 m above the surface (2 manuscripts submitted to AMT this summer).
We therefore added the following sentences: "*In addition to automated measurements, TOLNet is working towards extending the lidars' measurement range downward to about 100 meter above ground. The TOPAZ system can already measure ozone at such low range thanks to their scanning transmitter mirror, and assuming homogeneous ozone field in the vicinity of the instrument deployment site. Other TOLNet groups (e.g., TMTOL, LMOL) just started to provide valid measurements down to 100 meter above ground (Chouza et al., 2018; Farris et al., 2018).*"