# Second Response to Referee #1

We thank the reviewer for carefully looking into our manuscript and providing valuable suggestions, which helped to improve the manuscript considerably. Below are our responses to each one of the remaining minor suggestions and/or questions.

We took the opportunity to carefully revise the English writing one more time and fixed some additional typos and grammar errors. Moreover, we have worked to improve the quality of the figures and their captions. In this process, we noticed a minor issue with the figures for Manaus. For the cirrus case over the Amazon, the figure showing the signal was for Sep-21st while the cirrus extinction profile was for Aug-15th 2011. For the previous round of review, we changed the date to show a day with cirrus in multiple layers, but we inadvertently changed only one of the figures. Now all the data shown for this case is from Aug-15th, without any vertical smoothing.

**About the new title recommendation written in the Copernicus platform**
Title should be in my opinion:
"Collaborative development of a Lidar Processing Pipeline (LPP) for retrievals of atmospheric aerosol and cloud properties"

Thanks for the suggestion, the new title now reads:

> *Collaborative development of the Lidar Processing Pipeline (LPP) for retrievals of atmospheric ~~retrievals of~~ aerosols and clouds*

**Page 1 - Comment #1:** Well, this is not really true. There are also other processing chains which could be adapted to be used for every lidar and also yours need some certain requirements, eg, the LICEL data acquisition. But nevertheless, your efforts are well appreciated.

Thank you for the very fruitful discussion.

Yes, there are other processing chains and they might even have greater capabilities than our LPP. However, as far as we know, they are not publicly available. If there is something similar to LPP: configurable, open-source, and free to use by anyone, we can't find it. If you know of other processing codes that fit this description, please let us know and we will gladly cite it.

About the Licel software: According to the information on their web page: https://licel.com/software.htm, they only provide code to read the data and to do basic calculations, like adding files together or gluing the AN and PC channels.

**Page 2 - Comment #1 on:** Please, state that this holds for the Klett/Fernald method / or for elastic signals only.

Thanks for the suggestion. It definitely improves the accuracy of the statement. Now, the highlighted paragraph in the abstract reads as

> *For noiseless synthetic 532 nm elastic signals with a constant LR, the root-mean-square error (RMSE) in aerosol extinction within the boundary layer is about 0.1 %. In contrast, retrievals of aerosol backscatter from noisy elastic signals with a variable LR have an RMSE of 11 %, mostly due to assuming a constant LR in the inversion.*

**Page 2 - Comment #2 on:** A cloud optical depth is in fact not really a quality indicator because as intensive quantity it can be much lower or much higher depending on cloud strength. Probably you can report extinction values within the cloud and or the lidar ratio or you compare your COD/AOD from Lidar with the AOD from Aeronet (level 1.0).

We understand the reviewer's concern. However, extinction is also an intensive quantity and will vary depending on cloud strength as well. Moreover, this retrieval was performed assuming an average cirrus LR reported in previous studies, hence giving the LR is not relevant. Finally, while comparing the COD with AERONET's level 1.0 is in principle a good idea, this data is for nighttime measurements and this Aeronet station did not have a lunar-capable photometer at that time.

**Page 3 - Comment #1:** Great! But for the paper a Persistent Data Source is needed. Yet, on github you can always delete the code if you like and then the reference becomes worthless. So please try again to become a DOI (by the way, at zenodo you always can add authors not having a github account) or submit you release version as supplementary material so that it will be always available!.

The persistence of the code related to the paper is a good point. We followed the reviewer's suggestion and created the LPP's DOI for the version described in the manuscript, which is 10.5281/zenodo.7982889. We mention this in the **code availability** section, which now read:

> *The Lidar Processing Pipeline Version 1.1.2 reported here can be obtained from the persistence Zenodo repository https://zenodo.org/record/7982889. Up-to-date LPP versions can be obtained from the GitHub repository https://github.com/juanpallotta/LPP/. Besides the three LPP main modules, this repository also includes sample configuration files, shell scripts for automating the operation, sample lidar data files, and detailed instructions on using LPP.*

**Page 4 - Comment #1:** The implementation of the Raman does not need a algorithm intercompanson as you can build on the heritage from your cited paper. You can just implement one method which seems to be the most promising. WRT the photon counting, I do not agree. Usually, the Raman channels are photon counting only, thus no need for glueing. Dead time correction might be indeed needed, but could be gathered from the

manufacturers (of the PMT and DAQ) or measured. The glueing should be needed for elastic only, but I guess this is needed anyway for a proper Klett solution?..

Thank you for the suggestion.
While it is true that implementing a new algorithm does not require an intercomparison, doing so is the only way to learn where the different lidar groups in LALINET might have problems in their existing algorithms.

About photon counting, we have a different experience. At least two LALINET sites have N2-Raman channels recording AN+PC, while the H2O-channels are only PC. Hence, we need to implement and test a gluing routine for the Raman analysis, which uses both the Elastic and N2-Raman channels. Before gluing, however, we need the deadtime correction. We remind the reviewer that what we can get from the manufacturer is the "System dead-time" constant. However, for each measured PC signal, we would still need to apply the correction. For instance, for the non-paralyzable systems:

$$True\ Count\ Rate = Observed\ Count\ Rate \frac{1}{1 - Observed\ Count\ Rate * System\ dead\_time}$$

This is very simple and straightforward to implement, but it is mentioned in the paper because we still have to do it. Of course, using a glued elastic signal will also help the Klett inversion, as it has a greater dynamic range.

**Page 4 - Comment #2:** Open science: Please consider to develop together with other LPP's: This might a way to optimized work load.

Definitely. See the answer below as well.

**page 4 - Comment #3:** Great sounds promising. Any efforts planned to share expertise between MPLNet EARLINET, tAILNET and others? In terms of coding, it could bring benefit to all involved parties.

Thanks, for encouraging us about the Aeronet Downloader idea. This module is almost done, and soon we'll start integrating it within the LPP's workflow.

About sharing the expertise: yes, definitely! We always win with these collaborations, and we are glad to have collaborated (and to continue collaborating) with EARLINET, MPLNet and ADNet, and also ESA. During the LALINET bi-annual meetings we always invite and/or bring these scientists to participate. Part of our meetings is a Lidar School for graduate students and young researchers, and having these international lecturers over the last 20 years has certainly had a huge impact on the development of the lidar community in Latin America (as recognized in our BAMS paper: doi:10.1175/BAMS-D-15-00228.1).

However, we definitely need to be more aggressive in terms of engaging with these networks for code co-development. We hope this manuscript will be a big push in this direction.

**Page 5 - Comment #1:** Indeed, you are right. I was not aiming for a direct inter-comparison but a more general, "top-Lever descriptions of the differences as you have done just now: PollyNET for PollyXT only, SCC for ..., MPLNet for MPL only with a very specific configuration.....ADNet...Meaning: to highlight why you need an own LPP and can (want) not to use an existing one.

We are glad that the changes we made provided the comparison level you deemed necessary.

**Page 5 - Comment #2:** please cite.

Thanks for noticing that. We should have indeed cited PollyNET.
It is now added in the introduction:

> *(...) In contrast, homogeneous networks have the advantage of uniform calibration and data processing procedures, like those performed by the NASA Micro Pulse Lidar NETwork (MPLNET) (Welton et al., 2001), ~~or~~ the Italian Automated LIdar-CEilometer network (ALICEnet) (Dionisi et al., 2018), and the Raman and Polarization Lidar Network (PollyNET) (Baars et al., 2016).*

**Page 5 - Comment #3:** fully.

Thanks! Added.

**Page 6 - Comment #1:** Okay, but what about the height dependent "background noise" of analog signals. What about dark measurements? I.e. Sec 4 and 5 in http.//amt.copernicus.org/preprints/amt-2017-395/ .

Dark correction is already included. We explain this correction in the second paragraph of section 2.2 Data level 1 (L1), on line 104 of the manuscript:

> *"(...)Second is the dark current correction, which accounts for the signal distortions that are due to the acquisition system. Typical examples are transient peaks from firing the laser flashlamp and time-dependent electronic noise in analog channels, which are measured in the absence of light entering the telescope. If a dark current test has been performed, a file with this information can be provided, and the dark current for each channel will be subtracted from the corresponding measurements. (...)".*

**Page 8 - Comment #1:** (status 2023).

Thanks for noticing this missing information. Table 1 caption now reads:

*Number of stations in LALINET for each combination of emitted wavelength and detection modes (status 2023). More details about the network can be found in (Landulfo et al., 2020).*

**Page 9 - Comment #1:** is it possible to make the code also platform-independent, i.e. for windows and Mac? This is just a question out of interest….

Thank you for asking. It made us realize that we mention "Linux shell script" and "Linux terminal", which gives the wrong idea that the code is meant for Linux only. However, this is not the case. LPP works on all three platforms:

- Linux (tested on Ubuntu, but should work on all flavors),
- Windows (using the WSL - Windows Subsystem for Linux), and
- MAC (tested on both Intel and ARM architectures)

We modified this sentence to read:

*The processing pipeline has three main modules responsible for data processing at each level, all written in C/C++, which can run on Linux, Mac or Windows. These modules are independent, and the whole pipeline can be automated with a ~~Linux shell~~ script, or each module can be run directly in a ~~Linux~~ terminal.*

**Page 10 - Comment #1:** Same comment as above: Great thanks, but for the paper a Persistent Data Source is needed. Yet on github you can always delete the code if you like and then the reference becomes worthless. So please try again to become a DOI (by the way, at zenodo you always can add authors not having an github account) or submit you release version as supplementary material so that it will be always available!.

Agreed and solved. See the previous answer for **Page 3 - Comment #1**.

**Page 11 - Comment #2:** Okay, but maybe bring a reference to LICEL again.

Thanks for the suggestion, added.

**Page 12 - Comment #2:** Just for my understanding: Averaging the zenith angle is not really useful in case you really scan, right?.

Definitely, it is not useful, and not recommended to average a wide zenith angle. However, if LPP is fed with multiangle lidar signals, and the user asks to average the data, LPP would have to average angles as well. Whether that makes sense for the analysis, it might depend on the time resolution, how fast the scanning is performed, etc.

**Page 12 - Comment #3:** stored.

Thanks for the suggestion. We changed: "informed" => "set".

**Page 12 - Comment #4:** which might be height dependent - see comment for analog signals.

The signal noise is composed of at least two parts: the dark current and the background light. Dark current in a lidar system might depend on the time (i.e. range gate) because of the large transients introduced by firing the laser. On the other hand, background noise is either from the Sun/Moon or from thermal effects in the electronics, and both are independent of time (and hence range gate).

2.2. Data Level 1 (L1)
Line 104 - Explains Dark Current correction
Line 108 - Explains Background correction

**Page 12 - Comment #5:** The phrasing could be misleading. I guess you mean the last 500 or 1000 bins of the signal which refer to a height of xxx to yyy. Maybe rephrase to Typically the last ~500 or ~1000 bins out of xxx the lidar profile are used for determination of the background, this accounts for height ranges between yyy and zzz km.

Agreed, it can be misleading, thanks for pointing this out. Anyway, this is not a text in the manuscript, but only in our previous comment.

Moreover, it is not easy to give the actual range in altitude used for the BG calculation, because that will depend on the system setup. Some lidar stations might be recording the whole LICEL memory (16384 bins), while others might be recording only 4000 bins. And there are systems with 7.5 m and 15 m resolution.

**Page 13 - Comment #1:** Okay, well understood! Still the height depended electronic noise might be an issue….

Agreed. It would be wrong to try to find the constant BG level without having first corrected for the Dark Current range-dependent noise.

If the height dependent noise is particularly strong in a lidar system, then one needs to perform the Dark Current tests and use that data to do the proper correction. After that is corrected, the remaining BG (constant) level could be found by either of the methods mentioned here.

**Page 14 - Comment #1:** I think you should add a note that a paper for the feature mask is under preparation: I.e. Author xx, Title, under preparation. This means, if you fix the title and first author now, an interested person can later find the respective publication.

Thanks for the suggestion. However, the journal guidelines do not allow for citing unpublished work:

https://www.geoscientific-instrumentation-methods-and-data-systems.net/submission.html

*Works cited in a published manuscript should be published already, accepted for publication, or available as a preprint with a DOI.*

**Page 14 - Comment #3:** okay, I missed this information before. Please state this also in the mansucript.

In section **2.3 Data Level 2 (L2)**, line number 132 was changed and now looks like this:

> *The reference height, $z_0$, is not determined automatically and must be set by the user ~~as well~~ in the configuration file.*

**Page 15 - Comment #2:** Thanks, it looks much better now. I still see a little bias, but this might be attributed to a height-dependent electronic noise (dark signal).

Thank you for carefully looking at our results. For this particular lidar station, the operators perform a dark-current acquisition before and after every measurement. For the analysis shown in the paper, we used the closest dark to make the correction. Hence, there shouldn't be any time-dependent background. We will further investigate if the dark-current could be changing within a few hours, but this is beyond the scope of this manuscript.

**Page 16 - Comment #1:** this is ok as long as you clearly state that it is yet a user input!.

This point was answered in our response to Comment #3, page 14.

**Page 17 - Comment #1:** symbol for degree

Fixed