Collaborative development of the Lidar Processing Pipeline (LPP) for atmospheric retrievals of aerosol and cloud

Juan Pallotta[1], Silvânia A. Carvalho[2], Fabio Lopes[3], Alexandre Cacheffo[4], Eduardo Landulfo[3], and Henrique M. J. Barbosa[5,6]

[1]Centro de Investigaciones en Láseres y Aplicaciones, UNIDEF (CITEDEF-CONICET), Buenos Aires, Argentina.
[2]Universidade Federal Fluminense, Volta Redonda, Brazil
[3]Instituto de Pesquisas Energéticas e Nucleares, São Paulo, Brazil
[4]Universidade Federal de Uberlândia, Ituiutaba, Brazil
[5]University of Maryland Baltimore County, Baltimore, United States
[6]Universidade de São Paulo, São Paulo, Brazil

Correspondence: Juan Pallotta (juanpallotta@gmail.com)

Abstract. ~~Lidars~~ Atmospheric lidars can simultaneously measure clouds and aerosols with high temporal and spatial resolution and hence help understand ~~their~~ cloud-aerosol interactions, which are the source of the largest uncertainties in ~~current~~ future climate projections. However, atmospheric lidars are typically custom-built, ~~so~~ and there are significant differences between them. In this sense, lidar networks play a crucial role as they coordinate the efforts of different groups, providing ~~the~~ guide-lines for quality-assured routine measurements, opportunities for side-by-side instrument comparisons, and enforce algorithms validation, all aiming to homogenize the physical retrievals ~~. With that in mind, this work describes an ongoing effort to develop a lidar processing pipeline~~ from heterogeneous instruments in a network. Here we provide a high-level overview of the Lidar Processing Pipeline (LPP)~~collaboratively~~, an ongoing, collaborative, and open-source coordinated effort in Latin America. The LPP is a collection of tools ~~developed in C/C++, python, and Linux script that handle~~ that have the ultimate goal of handling all the steps of a typical ~~lidar analysis . The~~ analysis of lidar measurements. The modular and configurable framework is generic enough to be applicable to any lidar instrument. The first publicly released version of LPP produces data files at levels 0 (raw and metadata), 1 (averaging and layer-mask), and 2 (aerosol optical properties). We ~~discussed the~~ assess the performance of LPP through quantitative and qualitative analyses of simulated and measured elastic lidar signals. For noiseless synthetic 532 nm 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 signals with a variable LR have an RMSE of 11%, mostly due to assuming a constant LR in the inversion. The application of LPP for ~~two case studies for Sao Pauloand Amazon, which shows the capabilities of the current release but also highlights the need for new features~~measurements in Sao Paulo, further constrained by co-located AERONET data, indicates a lidar ratio of 63.9 ± 6.7 sr at 532 nm, in agreement with reported values for urban aerosols. Over the Amazon, analysis of a 6-km thick multi-layer cirrus indicates a cloud optical depth of about 0.77, also in agreement with previous results. From this exercise, we ~~developed and presented~~ identify the need for new features and discuss a roadmap to guide future development, accommodating the needs of our community.

1 Introduction

Aerosols, clouds, and their interactions are the source of the largest uncertainties in current climate change estimates (IPCC, 2013, 2021). More frequent and higher quality measurements of aerosol, clouds, and the physical processes governing their link with climate are needed to reduce these uncertainties (Mather, 2021; National Academies of Sciences, Engineering, and Medicine, 2018), and lidars are a powerful instrument to accomplish this task (Reagan et al., 1989). This instrument can provide information on the optical and microphysical properties of aerosol particles and hydrometeors, the concentration of trace gases, and, more recently, the 3D structure of the vegetation and urban canopies, allowing ecologists to accurately estimate biomass content and engineers to develop self-driving cars and drones (Wang and Menenti, 2021).

Atmospheric lidars, specifically, measure the atmosphere's constituents, from the troposphere to the mesosphere. However, they are developed by individual groups for particular applications; hence their hardware characteristics differ in essential aspects, such as receiving optics, emitted and detected wavelengths, polarization capability, and signal-to-noise ratio, to name a few. Even in the realm of single-wavelength elastic lidars, typical differences between custom-built lidar systems are large enough to require a careful, dedicated analysis of their return signals (Wandinger et al., 2016). In this sense, lidar networks play a crucial role as they coordinate the efforts of different groups, providing the guidelines for quality-assured routine measurements on a regional scale (Antuña-Marrero et al., 2017). Moreover, the coordinated effort is of utmost importance to homogenize the physical retrievals from the highly non-uniform instruments in lidar networks, which typically involve comparing the retrievals based on the algorithms of different groups (Pappalardo et al., 2004) and the instruments themselves (Wandinger et al., 2016). This homogenization is only possible by developing a unified processing pipeline that accounts for the hardware heterogeneity in the pool of instruments, as it has been accomplished recently in the context of the European Aerosol Research Lidar Network (EARLINET) (D'Amico et al., 2015) and the Asian Dust and Aerosol Lidar Observation Network (AdNet) (Sugimoto and Uno, 2009). 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).

The Latin America Lidar Network (LALINET) (Landulfo et al., 2016) is a Latin American coordinated heterogenous lidar network to obtain extensive and intensive aerosol optical properties profiled in the atmosphere. This federated lidar network aims to establish a consistent and statistically sound database to enhance the understanding of aerosol distribution over Latin America and its direct and indirect influence on climate. There are currently 19 stations in 6 countries, most of which are equipped with tropospheric aerosol lidars measuring one or more elastic return signals; only a few systems can measure inelastic return signals, typically for N_2 and H_2O Raman scattering. Table 1 shows the wide distribution of emitted wavelengths and detection modes. Other important differences are found in the laser repetition rate (ranging from 10 to 30 Hz), beam expander factor (1 to 5x), mirror diameter (20 to 50 cm), telescope focal length (1 to 4 m), and width of the interference filters (0.25 to 1 nm). Finally, only a few stations have co-located

or nearby measurements of the aerosol optical depth and the thermodynamic profile. More details about the network can be found in ~~Landulfo et al. (2020)~~(Landulfo et al., 2020).

Table 1. Number of stations in LALINET for each combination of emitted wavelength and detection modes.

	Emitted Wavelength		
Detection	355 nm	532 nm	1064 nm
Total Elastic	5	6	12
Cross Elastic	10	10	1
Parallel Elastic	10	10	1
HSRL	-	3	-
N_2 Raman	7	4	-
H_2O Raman	6	-	-

In recent years, the ~~Lalinet~~ LALINET network has worked towards establishing routine quality-assurance tests and intercomparing the retrieval algorithms used by the different groups ~~(Guerrero-Rascado et al., 2016; ?)~~(Guerrero-Rascado et al., 2016; Barbosa et al. ... Here, ~~we describe~~ our first goal is to present a high-level overview of the Lidar Processing Pipeline (LPP), an ongoing and unfunded coordinated effort to homogenize the ~~physical~~ retrievals from different ~~instruments in the Lalinet network, only possible with the development of a processing pipeline that accounts for the hardware heterogeneity in the pool of instruments. This manuscript describes~~ lidar instruments in Latin America. Our second goal is to introduce the tools developed ~~in C/C++, Python, and Linux scripts~~ to handle all the steps of a typical lidar analysis. ~~The modular framework~~ We want to emphasize the modular framework that is generic enough to be applicable to any lidar instrument or network ~~. The code is open source and is available on GitHub. The first public release of LPP includes a layer mask (clouds or aerosol plumes without discrimination), molecular calibration, and~~ and, at the same time, also emphasize the open source character of the LPP development (see *Code availability*). Our third goal is to show how LPP performs, through quantitative and qualitative analyses of synthetic and measured lidar signals. We will discuss case studies based on synthetic and measured signals and analyze aerosol backscatter retrievals for elastic lidar signals ~~. This~~ and layer masking (clouds or aerosol), which are the focus of this first public release of LPP.

~~The rest of this~~ manuscript is organized as follows. Section 2 presents the overall concept of LPP, the algorithms, and the structure of the different levels of output files. Sample results from the application of LPP to actual lidar data from two ~~Lalinet~~ LALINET stations are presented in section 3. ~~Future perspectives~~Current limitations, future perspectives, and conclusions are given in section 4.

2 PROCESSING PIPELINE

The Lidar Processing Pipeline is being developed in a partnership between the lidar groups of the Latin American Lidar Network. The LPP reads a series of raw data files in the standard Licel format (Licel GmbH, 2023) and produces a NetCDF file containing its data levels products 0, 1, and 2. The processing pipeline has three main modules responsible for data processing at each level, all written in written in C/C++. 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.

The modules are driven by a single configuration file, and the input data for each module is the output file produced in the previous stage, as can be seen schematically in Figure 1.

Moreover, the output file of a given level (e.g., 2) contains all the content of the previous level (e.g., 1) in addition to the new information generated in that level of data processing. In other words, all the information used to process the data to a given level is available in the corresponding file, thus allowing its reprocessing if needed. Figure 2 shows the content of a level 2 data file. The following sections explain the concept of each data level.

2.1 Data level 0 (L0)

The main goal of the level 0 data processing is to convert all the data files containing the lidar profiles into the standard LPP format used for all data levels. At this stage, there is no smoothing, averaging, or processing of any kind. The level 0 module simply dumps all information from a series of raw lidar files into a single output file. This includes all the information from the header of the raw files that describe the measurements, the instrument, and the site, such as filename, site name, start and stop time of the measurement, altitude, latitude, and longitude of the site, zenith and azimuth angles of the lidar signals (in case of scanning lidars), accumulated laser pulses, laser repetition rate, and the number of channels acquired. A description of these parameters is found in user manual of the LICEL transient recorders (Licel GmbH, 2023). For each channel, the following information is saved: channel ID, polarization state, type (elastic/Raman), number of height bins, photomultiplier voltage, and wavelength. For analog channels, the number of bits and range of the
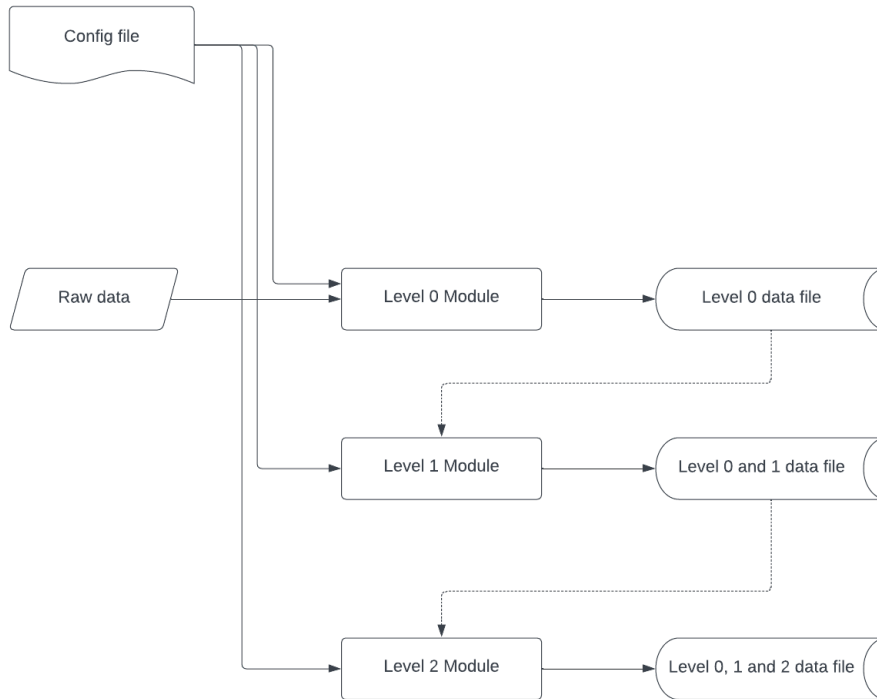
Figure 1. Flow diagram showing the structure of LPP version 1.1.0. Each module receives as input the data file produced as output in the previous module. A single configuration file is used for all the modules.

analog-to-digital converter (ADC) are recorded, while for photon counting channels, the discriminator level is recorded. For
110 each channel, the raw ADC ~~readings~~ counts are saved as 2-dimensional arrays indexed in height and time. ~~In the current version, the~~ Therefore, level 0 data consists of ADC counts for both analog and photon-counting channels, i.e., the raw values are not converted ~~from integer~~ to mV (analog) or ~~counts~~ MHz (photon counting). ~~All this information is saved in the root tree of the netCDF file (see Figure 2). In future versions, global information about the lidar system, such as the person in charge, contact, and hardware features, will be saved as global attributes. All this data will be kept in the files generated in the following~~
115 ~~data levels.~~ Figure 2 gives an overview of the file structure.

2.2 Data level 1 (L1)

The main goal of level 1 data processing is to apply the necessary corrections on the lidar profiles and compute a layer mask, which usually requires accumulating multiple profiles to increase ~~time averaging and~~ the signal-to-noise ratio. The level 1 module receives a single level 0 data file as input~~and~~, which contains all the raw signals and associated metadata, and produces
120 a single level 1 ~~NetCDF~~ file as output , with all new L1 information added to the subgroup L1 Data (see Figure 2). ~~L1 data includes updated start and stop times for each averaged profile and time-averaged values for zenith, azimuth, and lidar signal.~~
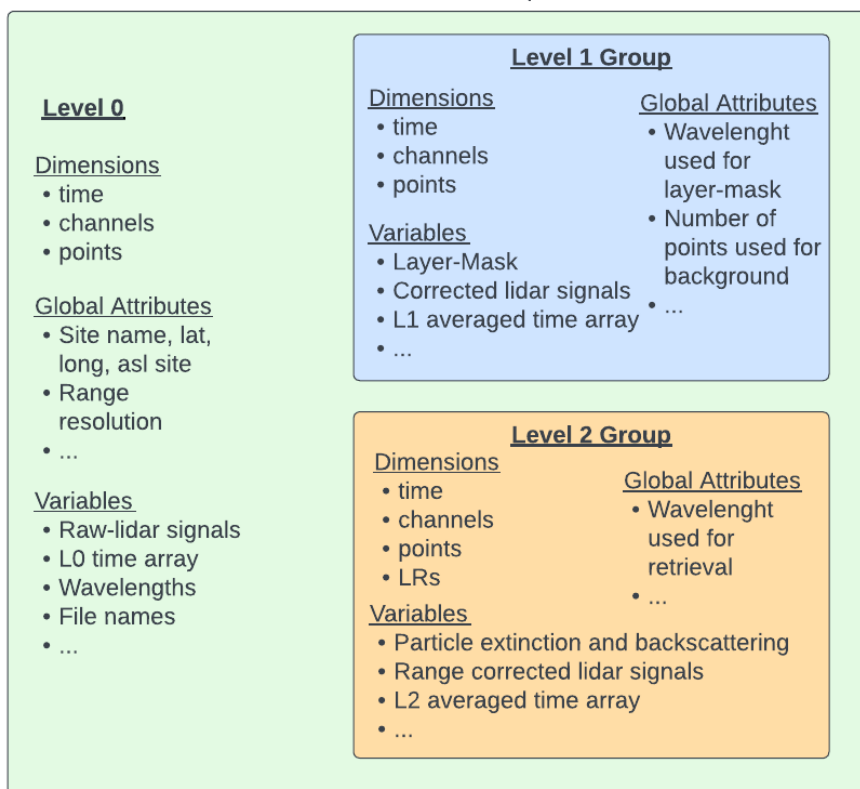
Figure 2. Output LPP's NetCDF file ~~inspected with Panoply software (Panoply)~~diagram. Product data levels 1 and 2 are stored in subgroups named ***LX Data***~~. Data~~, while data level 0 ~~information~~ data is stored in the root tree of the file. Dimensions and variables are explained in detail in the LPP's documentation on the GitHub repository (see section *Open Science Development*).

In the current version, the following corrections are implemented, based ~~one Guerrero-Rascado et al. (2016) and Freudenthaler et al. (201~~ on Guerrero-Rascado et al. (2016) and Freudenthaler et al. (2018). The first is the trigger delay correction. ~~There is typically a~~ This accounts for the possible delay between the emission of the laser pulse and the start of the data acquisition, resulting in

125 a vertical displacement of the measured signals, which affects both analog and photon-counting channels. ~~If the~~ The trigger delay can be measured by the so-called zero-bin (~~analog~~for the analog channels) and bin-shift (for the photon-counting ~~) tests have been performed, the number of bins that need to be shifted upward or downward can be~~ channels) tests. Each channel has a different time delay, which is given in terms of an integer number of range bins and is informed in the configuration file~~, and the correction will be applied. Second, there~~. The correction consists of discarding these first few bins, so that all

130 channels start in-sync with the laser pulse. The total length is also cropped so that all channels have the same length. Second is the dark current correction. ~~This accounts for~~, which accounts for the signal distortions that are ~~not~~ due to the ~~atmosphere~~

~~but the lidar system~~ ~~itself~~ 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. Third ~~, there~~ is the background correction ~~to remove the scattered solar~~, which accounts for sky radiation entering the telescope, which is unrelated to the lidar signal. ~~The background correction can be performed~~ This is due to scattered sunlight or moonlight, which produce a constant noise in the return signal, but would also include electronic noise if the dark-current correction was not applied. The background noise can be found by averaging the signal in ~~the far range~~ ~~or by a molecular fit~~ a high altitude range, defined in the configuration file, where the lidar signal is completely extinguished. Alternatively, it can be found by performing a so-called Rayleigh-fit. In this case, it corresponds to the constant term in a linear least-square-fit between the lidar signal and the molecular signal, as in (Grigorov and Kolarov, 2013) and (Barbosa et al., 2014b).

Besides the corrections and time averaging, L1 data also includes ~~a molecular density profile calculated from the input pressure and temperature profiles .~~ the temperature and pressure profiles provided by the user. The input thermodynamic profile can be obtained from any source, such as radiosondes, weather forecasts, or reanalysis. Alternatively, LPP includes and could use thermodynamic profiles from the US standard atmosphere (National Geophysical Data Center, 1992).

Finally, the L1 data processing creates a layer-mask ~~product based on dynamic thresholds calculated at the time-average resolution of level 1. Figure ?? gives an example of the range-corrected lidar signal (RCLS) and the corresponding layer mask for measurements at the lidar station of Sao Paulo on 30/07/2021.~~ to indicate the presence of aerosol and cloud layers. The method is based on the ideas proposed by (Vaughan et al., 2004), where the return signal is compared to the expected molecular signal. The threshold for detecting a layer is calculated dynamically, based on signal noise, hence it can be applied to the wide range of instruments in the LALINET network. The time-averaged resolution for this product can be different than the used in data level 2. The layer-mask shows if there is a presence of a layer (cloud or plume) by reporting a value of 1, and 0 otherwise. An example of the cloud-layer mask is given in the results section.

~~Color map of the range-corrected lidar signal (a) and its cloud-mask (b) from the L1 data processing for the Sao Paulo lidar station, on 30 July 2021.~~

2.3 Data level 2 (L2)

The main goal of level 2 data processing is ~~to obtain the optical~~ obtaining the profiles of aerosol ~~particles' optical properties.~~ optical properties, namely backscatter ($m^{-1}sr^{-1}$) and extinction coefficients (m^{-1}). Additional time and vertical averaging might be applied at L2 data, hence it might be different from that of L1 data. The level 2 module receives a single level 1 data file as input ~~and~~, which includes the thermodynamic profile and corrected lidar signals, and produces a single level 2 ~~NetCDF~~ file as output ~~, with all new L2 information added to the subgroup L2 Data (see Figure 2). LPP allows the time averaging at L2 data to be different from that of L1 data. Therefore, L2 data includes updated start and stop times for each average profile and time-averaged values for zenith, azimuth, and lidar signals. .~~

In ~~the current version~~ this first release of LPP, the optical properties ~~can be obtained from one of~~ are obtained from the analog elastic channels ~~at a time and include the aerosol extinction and backscatter profiles. The configuration file sets~~ using the Fernald method (Fernald, 1984). The value of the lidar ratio (LR)~~used for the inversion with the Fernald method (Fernald, 1984)~~, assumed to be constant, is set by the user in the configuration file. When multiple LR values are given, the inversion is performed for each ~~LR~~ value, producing a set of ~~profiles of~~ optical properties. ~~The aerosol~~ An example of the multi-LR retrieval is given in the results section. The reference height, z_0, is not determined automatically and must be set by the user as well. The reference signal, $P(z_0)$, which we assume contains no aerosol contribution, can be calculated either by averaging the signal or by taking the value of the molecular fit at the reference height. In both cases, this is evaluated in an altitude range defined by the user. The aerosol optical depth (AOD) is calculated assuming the extinction to be constant below a specific range (defined in the setting file), where the incomplete overlap precludes its calculation. ~~Figure 7 shows an example of multi-LR optical retrieval for a lidar signal from Sao Paulo lidar.~~

~~Extinction coefficient obtained using a set of pre-defined lidar ratios by the L2 data processing for the Sao Paulo lidar station, at 12:53 UTC on 14 September 2021. Extinction coefficient was assumed constant below 300 m.~~

~~In future releases, an option to automatically download Aeronet data will be implemented. This will allow using the column AOD to constrain the retrieval and obtain an estimation of the LR in the L2 datasets.~~

3 VALIDATION

Analyses of synthetic and measured lidar signals are carried out to demonstrate the usage of LPP and to provide quantitative and qualitative validation of our initial results. These analyses are based on elastic signals, which are the focus of this first public release of LPP. For a quantitative evaluation, we obtain the backscatter and extinction coefficients in the presence of aerosols or clouds and compare our results with either the input used for the simulations, AERONET retrievals of aerosol optical depth (Holben et al., 1998), or LR values reported in the literature. For a qualitative evaluation, we obtain the cloud-layer mask and compare it with the range-corrected lidar signal (RCLR) by visual inspection. The subsections below give the details of the four cases considered for validation.

3.1 ~~Open Science Development~~Synthetic Elastic Lidar Signals

~~The development of LPP follows an open-science approach. It is based on cooperative work and uses new digital technologies to diffuse knowledge and allow others to collaborate and contribute. This first released version reported here is ready to use. It can be obtained from GitHub (see *Code availability*)under an MIT license that enables the reuse, redistribution, and reproduction of all methods. 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~~We use two sets of synthetic lidar signals. The first is a simple case of an ideal lidar signal without noise and constant LR with altitude, provided by colleagues from Tropos, in Germany (Holger Baars, Personal communication, 2014). The aerosol profile has a constant extinction coefficient of 1100, 800, and $460\,\mathrm{Mm^{-1}}$ at 355, 532, and 1064 nm respectively, from the surface to the top of the

boundary layer at 1.5 km. Above that height, it decreases every 250 m, reaching almost zero at 2.5 km. The residual aerosol in the free troposphere has an extinction of 0.014, 0.01, and 0.0058 Mm^{-1} at 355, 532, and 1064 nm, respectively. The LR is fixed at 28, 39, and 77 sr at these wavelengths.

200   The second set of simulations corresponds to a more realistic case where the signals have noise, and the LR varies with altitude. These are the synthetic signals described and analyzed by Pappalardo et al. (2004) in the context of EARLINET's intercomparison of aerosol Raman lidar algorithms. This same dataset, which includes elastic signals at 355, 532 and 1064 nm and inelastic Raman signals at 387 nm and 607 nm, was used later to test the accuracy of SCC's optical products (Mattis et al., 2016). Three layers are clearly identified in this simulated atmosphere, denoted as a planetary boundary layer (PBL, from 0 to 1500

205 m), the free troposphere (FT, from 1500 to 3000 m), and a lofted layer (LL, 3000 to 7000 m). The dataset has 30 profiles with 2-min resolution, corresponding to 2400 laser shots per profile (20-Hz laser), and a total acquisition time of 30 min. Spatial resolution is 15 m. For our analysis, we consider the average 355 and 532 nm elastic signals only.

    Both sets of synthetic signals include the effect of incomplete overlap in the near-field to mimic a real measurement. We ignore this range for the inversions with LPP and analyze the profile only where the overlap is complete. Both simulations also

210 include information on the thermodynamic profile, which we use for calculating the molecular signal.

4 RESULTS AND DISCUSSION

3.1 Case Studies

We show LPP 's first results for two lidar systems To further validate LPP and demonstrate its application to real data, we analyze elastic return signals from the LALINET network: a station in São Paulo , lidar stations in Sao Paulo and Manaus,

215 both in Brazil. With over 21.5 million inhabitants, Sao Paulo is the largest metropolitan area in South America, and a station in Manaus, in the central Amazon rainforest. The case studies chosen for each site are based on each lidar characteristic, exploiting and highlighting the features of the current version of LPP. region in the Americas. One of the primary sources of air pollution there is vehicular emissions, and the city has struggled with high levels of traffic congestion for many years (Andrade et al., 2017). During the winter (June to September), this can be exacerbated by temperature inversions, which inhibit

220 mixing between the planetary boundary layer and the free atmosphere above. This well-stratified atmosphere shows high aerosol particle number concentrations within the boundary layer and a mostly clean atmosphere above it. While the air quality can vary depending on a number of factors, including weather patterns and traffic, we will evaluate measurements by the Sao Paulo lidar station on a typical winter day, 14 September 2020.

3.2 SPU Lidar - Brazil - 14 September 2020: Boundary Layer Aerosols

225 The lidar deployed at Sao Paulo (23$\circ$°56' S, 46$\circ$°74' W, 740 m above sea level) is a multiwavelength Raman LIDAR operated by the Lasers Environmental Applications Research Group at the Center for Lasers and Applications (CLA), Nuclear and Energy Research Institute (IPEN) (Landulfo et al., 2020). It is a monostatic coaxial system, vertically pointed to the zenith and

using a commercial Nd:YAG laser by Quantel, model Brilliant B at a repetition rate of 10 Hz. The output energy per pulse is 850 mJ for 1064 nm, 400 mJ for 532 nm, and 230 mJ for 355 nm. A 300 mm diameter telescope with a 1.5 m ~~and 0.1~~ focal distance and 1 mrad field-of-view (FOV) is used as a collection system, reaching a full overlap at 300 m above ground level. The detection box collects six different wavelengths: elastic 355 and 532 nm with the corresponding shifted Raman signals from nitrogen: 387 and 530 nm respectively. Also, the water vapor line at 408 nm and the elastic from 1064 nm. The electronic acquisition system is a Licel transient recorder model TR-20-160.

The second set of measurements is taken from the Manaus lidar station in the Amazon rainforest. The site is located about 20 km up-wind from the city, hence it is not affected by the significant urban emissions (Nascimento et al., 2022). Therefore, the atmosphere is mostly pristine throughout the year, with the exception being the dry season (June to October), when long-range transport of biomass burning affect the whole basin (Artaxo et al., 2013), and aerosols can be found up to 5 to 6 km (Baars et al., 2012). There is a marked diurnal cycle of convection, even during the dry season, with a peak in the late afternoon (Tanaka et al., 2014). Cirrus produced from the outflow of deep convective clouds are omni-present, with a frequency of occurrence much higher than other tropical regions (Gouveia et al., 2017b). Here, we will analyse a case of multi-layered cirrus clouds measured during the dry season, on 15 August 2011.

The lidar deployed in Manaus (2.89○ S, 59.97○ W, 100 m altitude above sea level) is a UV Raman Lidar operated by the University of Sao Paulo (Barbosa et al., 2014b). It is a bi-axial system pointed 5° from the zenith, which uses a commercial Quantel CFR-400 Nd-YAG laser at 355 nm with 95 mJ per pulse and a 10 Hz repetition rate. The receiving telescope has a 400 mm primary mirror, focal length of 4000 mm, and a field of view of about 1 mrad, reaching a complete overlap at 1.5 km. The detection box measures 3 wavelengths: elastic 355 nm and the corresponding Raman signals from nitrogen at 387 nm and water vapor at 408 nm. Data acquisition uses a Licel transient recorder model TR-20-160, with a raw resolution of 7.5 m.

4   RESULTS AND DISCUSSION

To validate LPP's first results, we performed the inversion of two sets of synthetic lidar signals and two sets of measurements from lidar systems in LALINET: a station in São Paulo, the largest metropolitan area in South America, and a station in Manaus, in the central Amazon rainforest. The simulations and the case studies exploit and highlight the features of LPP's first release.

4.1   Noiseless synthetic signals with constant LR

The range-corrected lidar signals (RCLS) from the first simulation are shown in Figure 3a for the three wavelengths. RCLS is suitable for plotting since it removes the inverse-squared range dependence in the raw-lidar signal, making it better for visualizations. The lack of noise in these signals makes finding the right reference value easier, reducing systematic errors related to this input parameter. Moreover, no vertical smoothing or time averaging was necessary. The inversions were performed using a reference altitude of 10 km and the true lidar ratio. Figure 3b shows the retrieved aerosol extinction coefficient and the input used for the simulation, showing an excellent visual agreement.

260   To quantify the small differences that might exist between the retrieval and the simulated profile, we computed the deviation as a function of altitude, and Table 2 reports the mean, maximum, minimum, and root-mean-square deviations in the boundary layer and in the free troposphere. There is a small negative bias within the boundary layer, where the relative deviations were always negative. The mean values are -0.045%, -0.089% and -0.901% for 355, ~~registering the lidar signals in its native lidar data files format, which is compatible with the LPP software~~532, and 1064 nm, respectively. Overall, the errors are greater for

265   1064 nm. In the free troposphere, where the aerosol loading is almost zero, there is a small positive bias. The mean deviations are 0.26, 0.092, and 0.003 Mm^{-1} for 355, 532, and 1064 nm, respectively.

Table 2. Mean, maximum, minimum, and root-mean-square deviations in the retrievals of the extinction coefficient (Mm^{-1}) at 355 nm, 532 nm, and 1064 nm. Values in the PBL (250 to 2500 m) are expressed as relative deviations (in %), while values in the free troposphere (FT, above 2500 m) are given in Mm^{-1}.

		355 nm	532 nm	1064 nm
PBL	mean	-0.045 %	-0.089 %	-0.901 %
	max	-0.036 %	-0.020 %	-0.298 %
	min	-0.071 %	-0.194 %	-1.623 %
	RMS	0.046 %	0.106 %	1.001 %
FT	mean	0.263 Mm^{-1}	0.092 Mm^{-1}	-0.003 Mm^{-1}
	max	0.345 Mm^{-1}	0.119 Mm^{-1}	0.005 Mm^{-1}
	min	-0.094 Mm^{-1}	-0.305 Mm^{-1}	-1.623 Mm^{-1}
	RMS	0.271 Mm^{-1}	0.098 Mm^{-1}	0.094 Mm^{-1}

   These errors are smaller than those reported by Böckmann et al. (2004) in a similar validation exercise in the context of EARLINET. Their case 2 considered an aerosol layer extending up to 4 km altitude, with constant LR, and the synthetic signals did not include noise. For stage 3 of their intercomparison, all 18 groups used the same LR and reference value at the

270   calibration height. The mean relative error within the aerosol layer was 1.87%, 1.48%, and 1.38% for 355, 532, and 1064 nm, respectively. While their synthetic profile was not the same as used here, this initial comparison gives confidence that LPP works well and can reproduce the simulations without biases if all input parameters are known.

4.2   Realistic synthetic signals with variable LR

The average signals from the set of realistic simulations, which includes signal noise and a variable LR, are shown in Figure

275   4. As this release of LPP considers a constant LR for the inversion of the elastic signals, we computed the mean of the LR profile below 7 km, where the simulated aerosols are. The values were 51 sr for 355 nm and 62 sr for 532 nm, which were the same constant values used by Mattis et al. (2016) to test the accuracy of SCC optical products. The reference height was set to 9 km for both wavelengths, with the molecular range between 7.5 to 10.5 km. Signals were vertically smoothed by applying a 5-point moving average, corresponding to an effective resolution of 75 m.
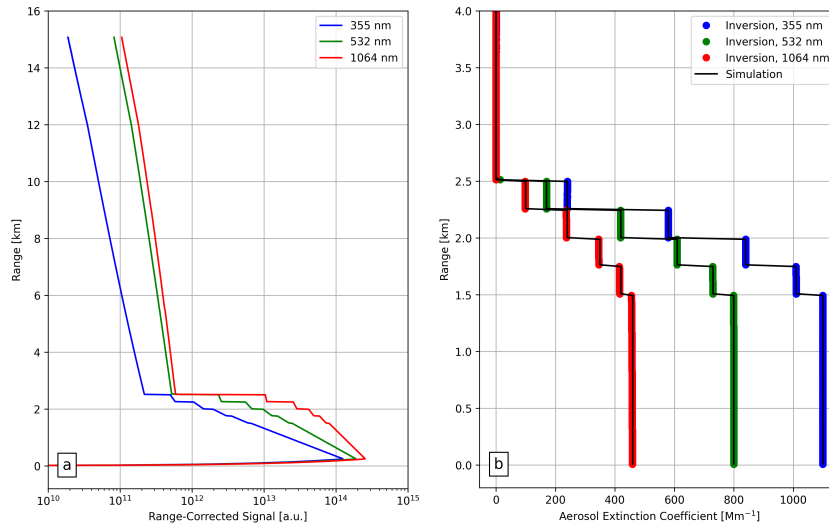
Figure 3. (a) Synthetic Range-corrected signals (a.u.) and (b) aerosol extinction coefficients (Mm^{-1}) retrieved with LPP (bullets) and used as input for the simulation (black line) are shown for 355, 532 and 1064 nm.
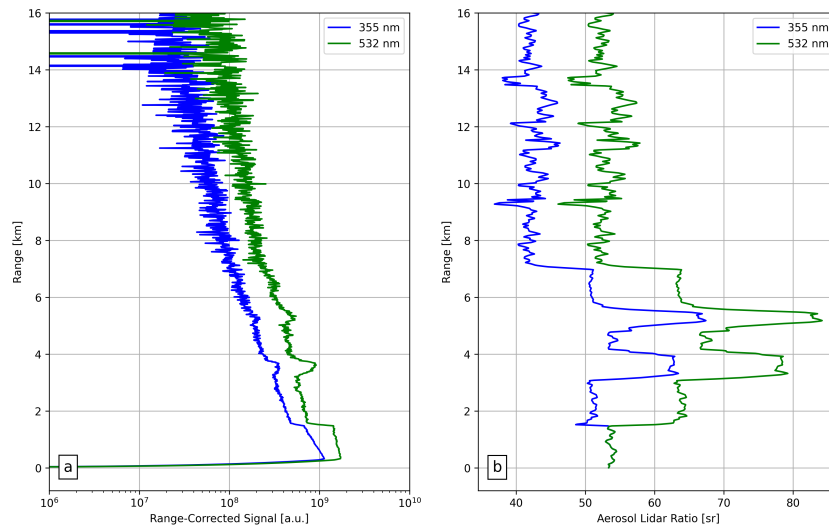
Figure 4. (a) Synthetic Range-corrected signals (a.u.) with raw vertical resolution of 15 m and (b) aerosol lidar ratio (sr) used as input for the simulation are shown for 355 nm and 532 nm. The retrievals shown in Fig. 5 assumed constant lidar ratios of 51 sr and 62 sr, respectively.

Figures 5a and 5b show the retrieved aerosol backscatter coefficients at both wavelengths. There is a very good agreement overall, despite our assumption of a constant LR. According to EARLINET requirements, relative errors in the optical retrievals at 355 and 532 nm should be below 20% or 0.5 Mm^{-1} sr^{-1} (Mattis et al., 2016). The relatively higher values obtained with LPP occur for relatively lower values of the backscatter coefficient, which indicate this is related to the signal noise and hence could

be minimized with stronger vertical smoothing. There are also large errors at the altitudes where the input LR makes a sudden

285 change, which can only be resolved by implementing the range dependent LR solution for the elastic signal (Klett, 1985), or by implementing the Raman solution (Ansmann et al., 1992).
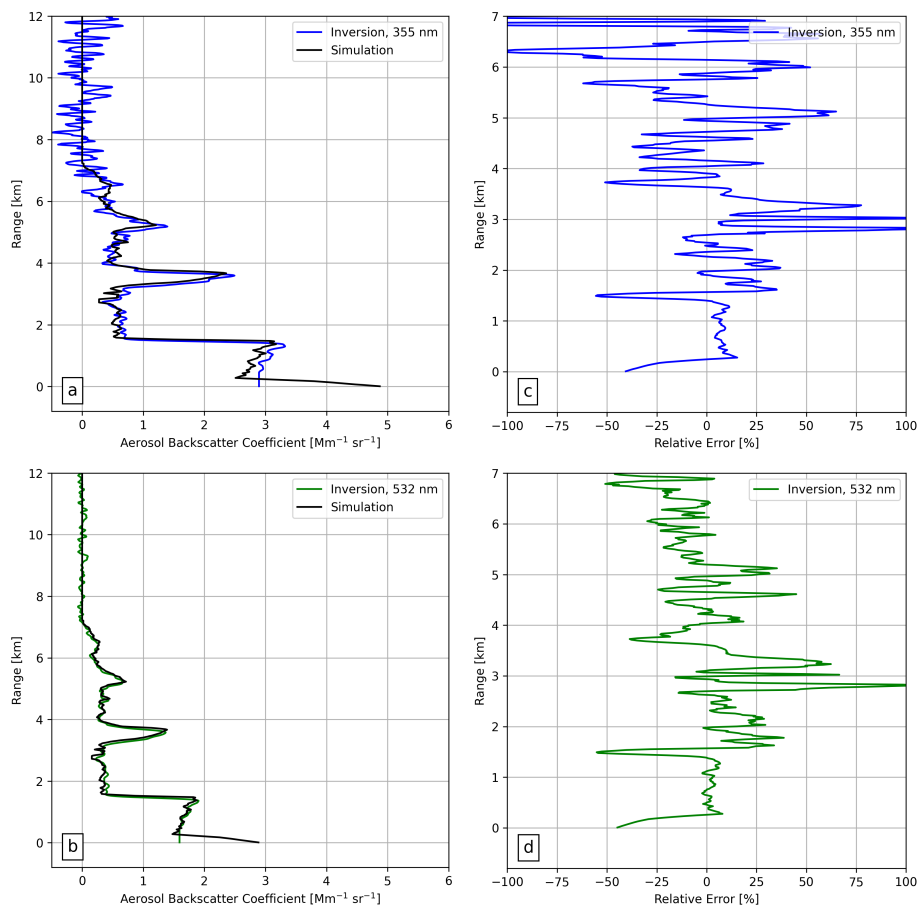
Figure 5. Left panels show a comparison of the aerosol backscatter coefficient (Mm^{-1} sr^{-1}) retrieved at (a) 355 nm (blue) and (b) 532 nm (green) with those used as input for the simulation (black line). The right panels show the respective relative errors (%).

The mean errors for the three aerosol layers and the two wavelengths are shown in Table 3. The layer-mean relative errors are smaller than EARLINET's limits, with the largest values (about 15%) found in the free troposphere, where the true LR profiles deviate the most from our assumed constant values. Root-mean-square relative errors are larger, reaching up to 41% in

290 the LL for 355 nm, however, these depend on the applied vertical smoothing as discussed above. Overall, the errors reported in table 3 are similar to those found for EARLINET's SCC (Mattis et al., 2016). This shows that LPP's retrievals do not have significant biases and can appropriately reproduce the realistic synthetic profiles.

Table 3. Mean, root-mean-square, and relative errors in the retrievals of the backscatter coefficient at 355 nm and 532 nm in the PBL (280 to 1500 m), free troposphere (FT, 1500 to 3000 m), and lofted layer (LL, 3000 to 7000 m). Values are reported in absolute ($Mm^{-1}\,sr^{-1}$) and relative (%) terms.

		355 nm		532 nm	
		$Mm^{-1}\,sr^{-1}$	%	$Mm^{-1}\,sr^{-1}$	%
PBL	mean	0.12	4.5	-0.013	-0.63
	RMS	0.37	13.	0.19	11.
FT	mean	0.048	15.	0.030	15.
	RMS	0.18	30.	0.10	29.
LL	mean	0.0026	-0.83	-0.0083	-3.3
	RMS	0.25	41.	0.10	22.

4.3 Case study: Sao Paulo

The day chosen for ~~the case analysis is~~ this case study was 14 September ~~2020.~~ 2020, a Monday near the end of the dry season. It was cloudless for most of the day, with some high clouds starting at 7 PM UTC ~~. For data level 1, the wavelength used for the layer mask was~~ (local time is UTC-3). We analyzed the elastic return signals in 532 nm, ~~and~~ with no time averaging ~~was applied for data level~~ (raw level-1 resolution was 1 ~~products (each raw lidar data has a 1~~ min~~acquisition time~~). Figure 6 shows the ~~colormaps~~ color maps of the range-corrected lidar signal (RCLS) and its cloud layer mask~~from the L1 output file of LPP:~~. Aerosols are trapped in the boundary layer, below 2.5 km. The diurnally forced convective boundary layer starts near the surface around 12 UTC (10 am LT), and grows until it is fully developed around 18 UTC (2pm LT). It takes over the nocturnal residual layer from the previous day, which can be seen in green colors during the first hours of the day.

~~Data level 2 analysis was performed for~~ Here we focus on the cloudless profiles between 15 h to 19 h (UTC) when AERONET~~'s~~ AOD data was available for comparison. ~~LPP was configured to use the 532 nm elastic channel, with an average of~~ The inversion of the elastic signals was performed by LPP level-2 module for 10 minutes ~~(10 files) and~~ time-averaged profiles, assuming constant LR values ranging from 30 to 100 sr, with steps of 5 sr (15 values). ~~LPP calculated the AOD values~~ Hence, LPP level 2 data files have the time series of extinction, backscatter profiles, and the AODs for each tentative LR. AOD is calculated integrating each of the extinction profiles, assuming it to be constant below 300 m ~~.~~

~~Figure 8a shows the AODs~~ to avoid the incomplete overlap region for this particular lidar station. Figure 7 shows the aerosol backscatter obtained from the multi-LR inversion ~~for each averaged profile~~ at 12:58 UTC. Figure 8a shows the time series of AOD values for the same LR values that can be compared with the AERONET retrievals shown in black.

Using 15 LRs allows searching for which LR value produces the closest AOD value measured by AERONET. Here, this is done a-posterior, during the analysis of the multi-LR inversion. The time series of the optimal LR values is shown in Figure 6b. Most of the values are between 50 and 75 sr, with a standard deviation of 6.9 sr, well within the range of values expected for urban aerosols (Ansmann and Müller, 2005). This initial evaluation shows that LPP is performing well for real
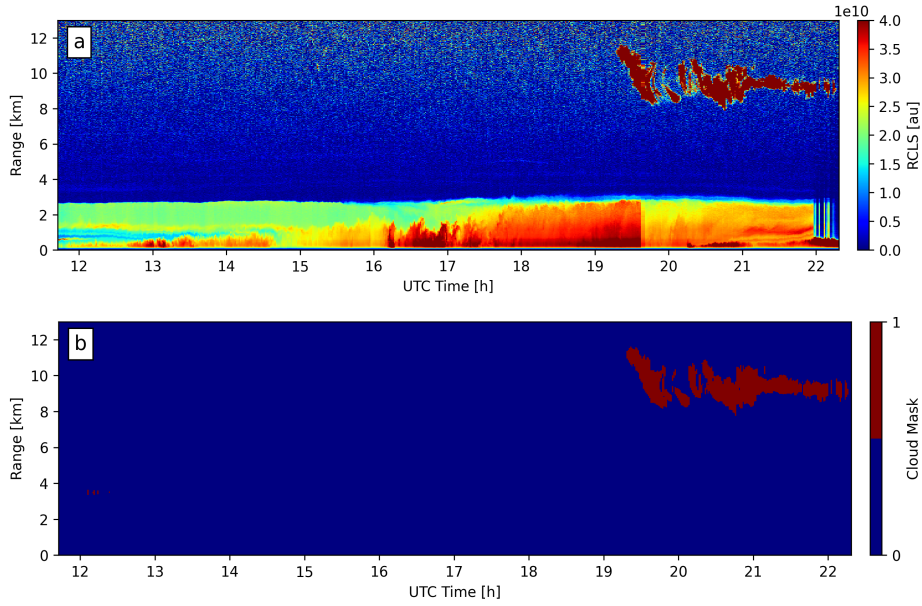
Figure 6. Level 1 data products showing the (a) RCLS and (b) cloud layer-mask from measurements at the Sao Paulo lidar station on 14 September 2020.
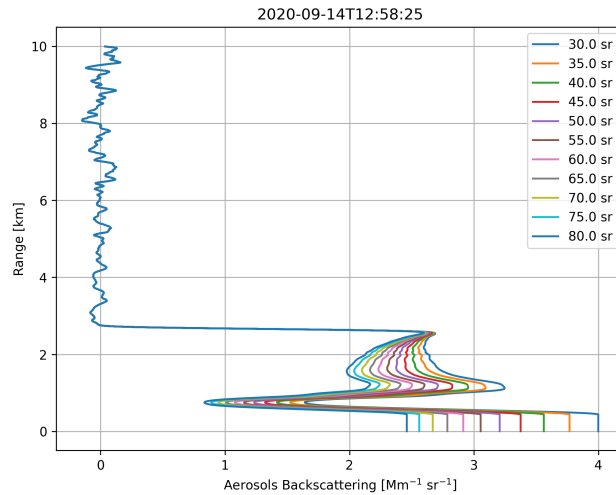
Figure 7. Aerosol backscattering coefficient obtained for the Sao Paulo lidar station on 14 September 2021, from 12:58 to 13:08 UTC, using a set of pre-defined lidar ratios. The extinction coefficient is assumed constant below 300 m.

315   lidar data, and demonstrate how the multi-LR retrieval can me used to constrain the optical properties obtained from elastic only measurements.
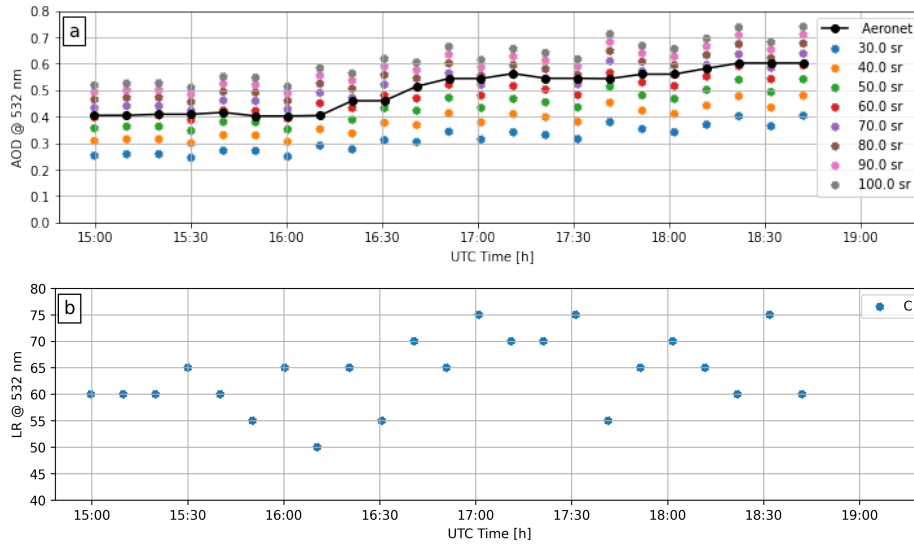
Figure 8. (a) AOD's multi-LR lidar product and with the closest Aeronet value and (b) LRs derived from each inversion.

The current version of LPP produces level 2 dataof the elastic multi-LR inversion, with the extinction, backscattering profiles, and the AODs for each LR. Future versions of LPP will include an optimization routine to find the LR value that gives precisely the reference AOD from AERONET. This could either be downloaded automatically from the AERONET website or informed by the user on the configuration file.

4.4   Case study: Manaus

4.5   Manaus Lidar - Brazil - 14 August 2011: Cirrus Clouds

The Manaus lidar site (Barbosa et. al., 2014b) is located inside the campus of Embrapa Amazônia (2.89∘ S, 59.97∘ W, 100 m altitude above sea level), approximately 30 km away from the city of Manaus-AM, Brazil. The instrument is a commercial lidar system model LR-102-U-400/HP by Raymetrics. The emitter is a Quantel CFR-400 Nd-YAG laser at 355 nm with 95 mJ per pulse and a 10 Hz repetition rate, with a beam expander producing a final divergence of 0.36 mrad. The reception optics consists of a Cassegrain telescope with a focal length of 4000 mm and a 400 mm primary mirror, resulting in an f/10 system. The field stop at the focal plane can be changed, allowing the telescope field of view to be adjusted between 0.25 and 3 mrad. The setup is bi-axial, with a 300 mm separation between the laser and the telescope axes. With a 7 mm field stop, the complete overlap of the system starts at 1.5 km. The detection box permits the acquisition of elastics and inelastic lidar signals in the UV using photon-current (ADC) and photon-counting (PC) methods. The lines acquired are 355 nm (ADC and PC), N_2's Raman line at 387 (ADC and PC), and the water vapor line, 408 nm (PC). Data acquisition uses a Licel transient recorder model TR-20-160, with a raw resolution of 7.5 m.

For this ~~lidar station~~second case study, we focus ~~here~~ on measurements of cirrus clouds ~~.~~performed on 15 August 2011.
Previous studies in the Amazon region showed that their frequency of occurrence is higher than in other tropical regions and that they can have an important radiative effect (Gouveia et al., 2017a). Figure 9 shows a color map of the RCLS and the corresponding cloud layer mask from the level 1 data, where a thin multi-layered cirrus between 10 and 15 km is clearly seen. The acquisition time of each lidar profile was 1 minute, and no time average was used for this layer mask retrieval, which clearly captured the whole cirrus cloud.

For the retrieval of the optical properties, we used a lidar ratio of 23.3 sr ~~for the cirrus clouds~~ following the measurements reported by ~~Gouveia et al. (2017a)~~(Gouveia et al., 2017b). Figure 10 shows the particle extinction profile ~~, corresponding to a~~ at 2:30 UTC (5-min average). While there seems to be three layers of cirrus clouds, the structure is quite complex and none of the layers are isolated. The corresponding cloud optical depth ~~of~~ is 0.77~~. These results might be affected by multiple scattering~~, which is ~~important for cirrus clouds and~~ large enough for multiple scattering to be important, but this has not been accounted for. Nonetheless, ~~values~~ the extinction values and the cloud optical depth (COD) are reasonable and in the typical range of previous ~~work (Gouveia et al., 2014)~~works (Gouveia et al., 2014, 2017b), and show what can be obtained with the current version of LPP.
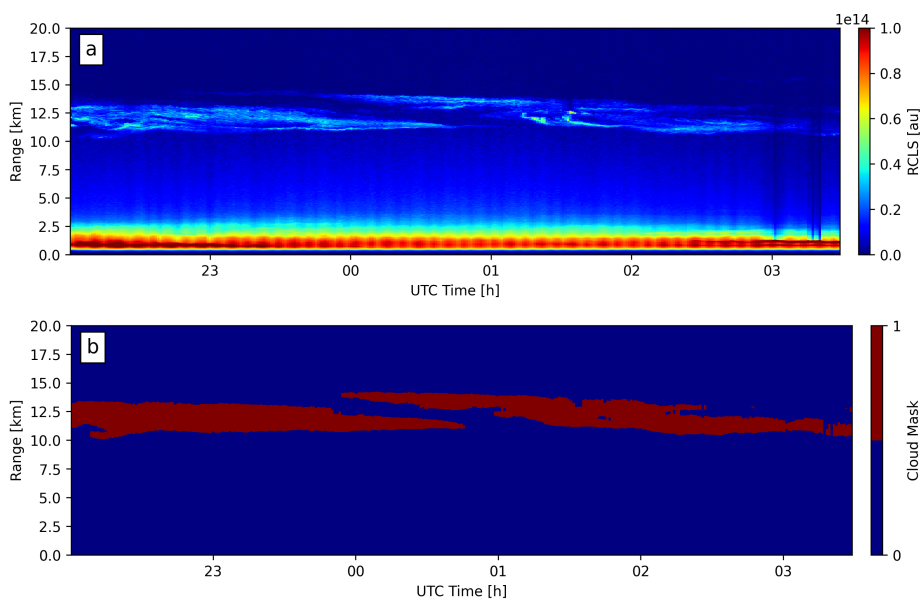
Figure 9. (a) RCLS (a) and layer-mask (b) from Manaus lidar for ~~21 September~~ 15 August 2011.

~~The next release of LPP will introduce the use of a profile of LR instead of assuming a fixed value. This will allow processing cases, as depicted in this example, with cirrus clouds above biomass burning. Further studies and developments are necessary for implementing an automatic algorithm for removing the multiple scattering effects in semi-transparent clouds.~~
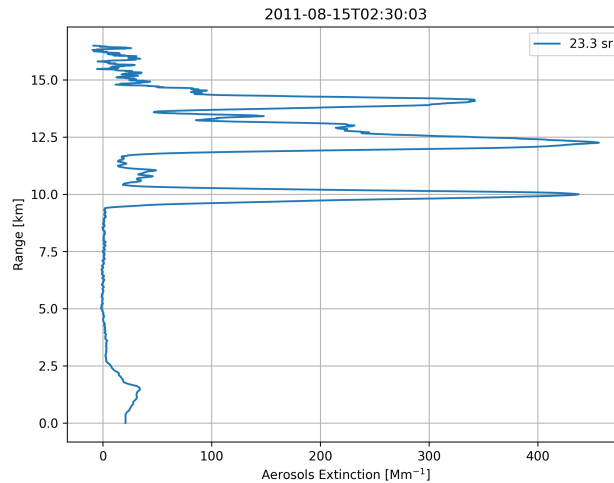
Figure 10. Extinction ~~profile taken~~ coefficient obtained using a constant lidar ratio for the Manaus lidar station, at ~~00~~02:~~15~~ 30 UTC ~~for 21 September~~ on 15 August 2011. The extinction coefficient was assumed constant below 300 m.

4.5 ~~Future roadmap~~FUTURE ROADMAP

With the first release of LPP and its use by the different groups in LALINET, we have identified the necessary improvements and built a roadmap to guide future development. ~~The benefit of LPP's highly modular concept is the possibility of different groups modifying and testing different modules without interfering with the rest of the pipeline.~~

355 An initial consideration is that LPP processed data files ~~need to~~ must be FAIR (Findable, Accessible, Interoperable, and Reusable) (Wilkinson et al., 2016) to be compatible with Open Science. In this sense, more information about the site, hardware, operation, files processed, and even the version of ~~the~~ LPP used needs to be added as metadata in the ~~NetCDF structure in future releases . The second important point is that not all~~ output files. Moreover, the benefit of LPP's highly modular concept is the possibility of different groups modifying and testing different modules without interfering with the rest of the
360 pipeline. To facilitate the customization of the pipeline to fulfill different needs and to allow more groups to contribute to LPP's development, future releases will include Python versions of all modules.

In terms of improvements in the physical retrievals, we have identified three priorities. First is to implement the (Klett, 1985) solution to the lidar equation with a range dependent LR. This would be useful, for instance, for the Sao Paulo station where the sea breeze frequently brings marine aerosols above the urban-polluted boundary layer (Rodrigues et al., 2013; Ribeiro et al., 2018)
365 . The second is to obtain the uncertainties in the extinction and backscatter coefficients by propagating the signal errors using a Monte Carlo approach (Press, 2007), following the work of (Alvarez et al., 2006) and (Mattis et al., 2016). Finally, we plan on implementing the Raman solution (Ansmann et al., 1992), but this might require an intercomparison effort of the existing algorithms in LALINET, as was done in EARLINET (Pappalardo et al., 2004). Moreover, in LALINET, the stations recording Raman return signals have photon-counting channels, which might be affected by dead-time effects (Johnson et al., 1966).

370 Hence, we need to implement the known dead-time corrections for paralyzable and non-paralyzable systems (Whiteman et al., 1992; Knoll, , which would also allow for 'gluing' the analog and photon-counting to extend the instrument dynamic range (Whiteman et al., 2006; News.

Regarding the automation of the pipeline, a few updates are planned. For instance, we noticed that only a few lidar stations in Latin America have a nearby radio-sounding site, and ~~even when it is available, it is only twice daily~~it is only once or
375 twice per day. To facilitate the processing of level 1 and level 2 data, an automatic "thermodynamic profile downloader" will be ~~needed. This should allow for downloading~~ developed to obtain a co-located thermodynamic profile ~~, either from a~~from a nearby radio-sounding, a forecast model, or a reanalysis. ~~Alternative methods of rescaling~~ The MPLNET data processing, for example, automatically retrieves meteorological profiles from the Goddard Earth Observing System, version 5 (GEOS-5), atmospheric general circulation model for all molecular calculations (Lewis et al., 2016). We also plan to implement a method
380 of re-scaling the standard atmosphere profiles based on co-located ground-based temperature and pressure measurements, which could also be retrieved automatically from meteorological databases~~, could be beneficial. Finally~~.

Moreover, a well-known problem with the inversion of elastic lidar data is the need to assume an a priori lidar ratio. The typical solution ~~for day-time measurements~~ is to choose a lidar ratio that brings the estimated AOD value closer to the reference value measured by AERONET, which can now be measured at day and nighttime (Perrone et al., 2022). This anal-
385 ysis can be made ~~a posteriori, but implementing~~ a-posteriori, as shown in Figure 8. Implementing an optimization routine would allow LPP to automatically identify the best LR for each profile, ~~given that the user provides~~ as has been done in previous studies (Córdoba-Jabonero et al., 2011; Román et al., 2018), and help reduce systematic errors in the retrieved profiles (Welton and Campbell, 2002). The user could provide the reference AOD value~~. Alternatively, the reference values~~, or it could be obtained by an "AOD data downloader" tool as part of the LPP framework.

390 **5 Conclusions**

Lidar networks are drivers of scientific advance as they coordinate the efforts of different groups, allowing for uniform quality-assurance procedures, as well as instrument comparison and algorithm validation. The development of LPP is a joint effort leveraging the expertise and manpower of different Latin-American ~~groups~~lidar groups, part of the Latin American Lidar Network. The goal ~~is to have~~ of LPP is to provide a set of open-source tools for each step of the typical lidar data analysis rou-
395 tine. ~~We described~~ Here, we provided a high-level overview of the first working version that is now released for the scientific community on our GitHub repository~~for open access to everyone who wants to be involved or use the software. The two~~.

The performance of LPP was evaluated through an analysis of synthetic and measured elastic lidar signals. For noiseless synthetic signals with a constant LR, the mean relative error in the aerosol extinction within the boundary layer was quite small, ranging from -0.005% to -0.9%, depending on the wavelength. For noisy synthetic signals with a variable LR, the
400 mean relative error in aerosol backscatter was larger, ranging from -0.63% to 4.5%, mostly due to assuming a constant LR in the inversion. For the case studies for urban aerosols in Sao Paulo and ~~Amazon~~ cirrus clouds in the Amazon, we found LR and COD values, respectively, in agreement with previous results. These analyses showed the capabilities of the current

release but also highlighted the need for new features. ~~From this exercise~~Hence, we have built a roadmap to guide future development, which includes~~adding additional metadata, developing a thermodynamic profile and AERONET data downloader tools, and LR optimization methods~~: (1) improvements in the physical retrievals (e.g. range dependent LR inversion of the elastic signal, or uncertainty propagation using Monte-Carlo), and (2) automation of the pipeline (e.g. optimizing the elastic LR by constraining the column AOD, or thermodynamic profile downloader tools). Future releases will bring these and other new features, accommodating the needs of our community.

Although the scientific community is moving towards open science, developing open-source code is still a hurdle, and the atmospheric lidar community has not yet embraced the idea. Consolidated networks have long developed their own algorithms and pipelines, which unfortunately remain mostly inaccessible to the community, hampering faster scientific advancement. We hope open-source efforts, as the one presented here, become the rule rather than the exception in the near future.

Code availability. The development of LPP follows an open-science approach, and is based on cooperative work and uses new digital technologies to diffuse knowledge and allow others to collaborate and contribute. It is distributed under an MIT license that enables the reuse, redistribution, and reproduction of all methods. The Lidar Processing Pipeline Version 1.1.0 reported here can be obtained from the GitHub repository https://github.com/juanpallotta/LPP/releases/tag/v.1.1.0. 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.

Author contributions. Conceptualization of this ongoing effort is shared by all authors. JP leads the software development, methods, and algorithms, with contributions from FL and HB. FL and HB prepared the example datasets analyzed in this study. SC performed the first test of the code across different machines. JP and HB were responsible for producing all figures and writing the original draft. All authors reviewed and edited the published version of the paper. HB supervised the research.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We would like to thank the financial support from CITEDEF used to acquire the ~~server PC~~ PC in which this project is developed. The authors also thank the financial support from Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) / (ARC: E-26/010/002464/2019 Ref.: 211.599/2019). We are grateful to Eng. Federico Verstraeten for sharing their knowledge in software development.

References

Alvarez, J. M., Vaughan, M. A., Hostetler, C. A., Hunt, W. H., and Winker, D. M.: Calibration Technique for Polarization-Sensitive Lidars, Journal of Atmospheric and Oceanic Technology, 23, 683 – 699, https://doi.org/https://doi.org/10.1175/JTECH1872.1, 2006.

430 Andrade, M. d. F., Kumar, P., de Freitas, E. D., Ynoue, R. Y., Martins, J., Martins, L. D., Nogueira, T., Perez-Martinez, P., de Miranda, R. M., Albuquerque, T., Gonçalves, F. L. T., Oyama, B., and Zhang, Y.: Air quality in the megacity of São Paulo: Evolution over the last 30 years and future perspectives, Atmospheric Environment, 159, 66–82, https://doi.org/10.1016/j.atmosenv.2017.03.051, 2017.

Ansmann, A. and Müller, D.: Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere, chap. Lidar and Atmospheric Aerosol Particles, pp. 105–141, Springer New York, New York, NY, https://doi.org/10.1007/0-387-25101-4_4, 2005.

435 Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., and Michaelis, W.: Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar, Appl. Opt., 31, 7113–7131, https://doi.org/10.1364/AO.31.007113, 1992.

Antuña-Marrero, J. C., Landulfo, E., Estevan, R., Barja, B., Robock, A., Wolfram, E., Ristori, P., Clemesha, B., Zaratti, F., Forno, R., Armandillo, E., Álvaro E. Bastidas, Ángel M. de Frutos Baraja, Whiteman, D. N., Quel, E., Barbosa, H. M. J., Lopes, F., Montilla-Rosero,

440 E., and Guerrero-Rascado, J. L.: LALINET: The First Latin American–Born Regional Atmospheric Observational Network, Bulletin of the American Meteorological Society, 98, 1255 – 1275, https://doi.org/10.1175/BAMS-D-15-00228.1, 2017.

Artaxo, P., Rizzo, L. V., Brito, J. F., Barbosa, H. M. J., Arana, A., Sena, E. T., Cirino, G. G., Bastos, W., Martin, S. T., and Andreae, M. O.: Atmospheric aerosols in Amazonia and land use change: from natural biogenic to biomass burning conditions, Faraday Discussions, 165, 203, https://doi.org/10.1039/c3fd00052d, 2013.

445 Baars, H., Ansmann, A., Althausen, D., Engelmann, R., Heese, B., Müller, D., Artaxo, P., Paixao, M., Pauliquevis, T., and Souza, R.: Aerosol profiling with lidar in the Amazon Basin during the wet and dry season: AEROSOL PROFILING IN AMAZONIA, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2012JD018338, 2012.

Barbosa, H., Lopes, F., Silva, A., Nisperuza, D., Barja, B., Ristori, P., Gouveia, D., Jimenez, C., Montilla, E., Mariano, G., Landulfo, E., Bastidas, A., and Quel, E.: The first ALINE measurements and intercomparison exercise on lidar inversion algorithms, Optica Pura y

450 Aplicada, 47, 99–108, https://doi.org/10.7149/OPA.47.2.99, 2014a.

Barbosa, H. M. J., Barja, B., Pauliquevis, T., Gouveia, D. A., Artaxo, P., Cirino, G. G., Santos, R. M. N., and Oliveira, A. B.: A permanent Raman lidar station in the Amazon: description, characterization, and first results, Atmospheric Measurement Techniques, 7, 1745–1762, https://doi.org/10.5194/amt-7-1745-2014, 2014b.

Böckmann, C., Wandinger, U., Ansmann, A., Bösenberg, J., Amiridis, V., Boselli, A., Delaval, A., Tomasi, F. D., Frioud, M., Grigorov,

455 I. V., Hågård, A., Horvat, M., Iarlori, M., Komguem, L., Kreipl, S., Larchevêque, G., Matthias, V., Papayannis, A., Pappalardo, G., Rocadenbosch, F., Rodrigues, J. A., Schneider, J., Shcherbakov, V., and Wiegner, M.: Aerosol lidar intercomparison in the framework of the EARLINET project. 2.Aerosol backscatter algorithms, Appl. Opt., 43, 977–989, https://doi.org/10.1364/AO.43.000977, 2004.

Córdoba-Jabonero, C., Sorribas, M., Guerrero-Rascado, J. L., Adame, J. A., Hernández, Y., Lyamani, H., Cachorro, V., Gil, M., Alados-Arboledas, L., Cuevas, E., and de la Morena, B.: Synergetic monitoring of Saharan dust plumes and potential impact on sur-

460 face: a case study of dust transport from Canary Islands to Iberian Peninsula, Atmospheric Chemistry and Physics, 11, 3067–3091, https://doi.org/10.5194/acp-11-3067-2011, 2011.

D'Amico, G., Amodeo, A., Baars, H., Binietoglou, I., Freudenthaler, V., Mattis, I., Wandinger, U., and Pappalardo, G.: EAR-LINET Single Calculus Chain – overview on methodology and strategy, Atmospheric Measurement Techniques, 8, 4891–4916, https://doi.org/10.5194/amt-8-4891-2015, 2015.

465 Dionisi, D., Barnaba, F., Diémoz, H., Di Liberto, L., and Gobbi, G. P.: A multiwavelength numerical model in support of quantitative retrievals of aerosol properties from automated lidar ceilometers and test applications for AOT and PM_{10} estimation, Atmospheric Measurement Techniques, 11, 6013–6042, https://doi.org/10.5194/amt-11-6013-2018, 2018.

Fernald, F. G.: Analysis of atmospheric lidar observations: some comments, Appl. Opt., 23, 652–653, https://doi.org/10.1364/AO.23.000652, 1984.

470 Freudenthaler, V., Linné, H., Chaikovski, A., Rabus, D., and Groß, S.: EARLINET lidar quality assurance tools, Atmospheric Measurement Techniques Discussions, https://doi.org/10.5194/amt-2017-395, 2018.

Gouveia, D. A., Barbosa, H. M. J., and Barja, B.: Characterization of cirrus clouds in central Amazon (2.89ºS, 59.97ºW): Firsts results from observations in 2011, Special Section: VII Workshop on Lidar Measurement in Latin-America, 47, 109–114, https://doi.org/http://dx.doi.org/10.7149/OPA.47.2.109, 2014.

475 Gouveia, D. A., Barja, B., Barbosa, H. M. J., Seifert, P., Baars, H., Pauliquevis, T., and Artaxo, P.: Optical and geometrical properties of cirrus clouds in Amazonia derived from 1 year of ground-based lidar measurements, Atmospheric Chemistry and Physics, 17, 3619–3636, https://doi.org/10.5194/acp-17-3619-2017, 2017a.

Gouveia, D. A., Barja, B., Barbosa, H. M. J., Seifert, P., Baars, H., Pauliquevis, T., and Artaxo, P.: Optical and geometrical properties of cirrus clouds in Amazonia derived from 1 year of ground-based lidar measurements, Atmospheric Chemistry and Physics, 17, 3619–3636,
480 https://doi.org/10.5194/acp-17-3619-2017, 2017b.

Grigorov, I. and Kolarov, G.: Rayleigh-fit approach applied to improve the removal of background noise from lidar data, in: 17th International School on Quantum Electronics: Laser Physics and Applications, edited by Dreischuh, T. N. and Daskalova, A. T., vol. 8770, p. 877010, International Society for Optics and Photonics, SPIE, https://doi.org/10.1117/12.2012998, 2013.

Guerrero-Rascado, J. L., Landulfo, E., Antuña, J. C., de Melo Jorge Barbosa, H., Barja, B., Álvaro Efrain Bastidas, Bedoya, A. E., da
485 Costa, R. F., Estevan, R., Forno, R., Gouveia, D. A., Jiménez, C., Larroza, E. G., da Silva Lopes, F. J., Montilla-Rosero, E., de Arruda Moreira, G., Nakaema, W. M., Nisperuza, D., Alegria, D., Múnera, M., Otero, L., Papandrea, S., Pallota, J. V., Pawelko, E., Quel, E. J., Ristori, P., Rodrigues, P. F., Salvador, J., Sánchez, M. F., and Silva, A.: Latin American Lidar Network (LALINET) for aerosol research: Diagnosis on network instrumentation, Journal of Atmospheric and Solar-Terrestrial Physics, 138-139, 112–120, https://doi.org/https://doi.org/10.1016/j.jastp.2016.01.001, 2016.

490 Holben, B., Eck, T., Slutsker, I., Tanré, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sensing of Environment, 66, 1–16, https://doi.org/https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, vol. 1535 pp, [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A.
495 Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/doi:10.1017/CBO9781107415324, 2013.

IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, vol. In Press, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896, 2021.

500 Johnson, F. A., Jones, R., McLean, T. P., and Pike, E. R.: Dead-Time Corrections to Photon Counting Distributions, Physical Review Letters, 16, 589–592, https://doi.org/10.1103/PhysRevLett.16.589, 1966.

Klett, J. D.: Lidar inversion with variable backscatter/extinction ratios, Appl. Opt., 24, 1638–1643, https://doi.org/10.1364/AO.24.001638, 1985.

Knoll, G. F.: Radiation detection and measurement, John Wiley, Hoboken, N.J, 4th ed edn., oCLC: ocn612350364, 2010.

505 Landulfo, E., da Silva Lopes, F. J., de Arruda Moreira, G., Marques, M. T. A., Osneide, M., Antuña, J. C., Arredondo, R. E., Guerrero Rascado, J. L., Alados-Arboledas, L., Bastidas, A., Nisperuza, D., Bedoya, A., Múnera, M., Alegría, D., Forno, R. N., Sánchez, M. F., Lazcano, O., Montilla-Rosero, E., Silva, A., Jimenez, C., Quel, E., Ristori, P., Otero, L., Barbosa, H. M., Gouveia, D. A., and Barja, B.: ALINE/LALINET Network Status, EPJ Web of Conferences, 119, 19 004, https://doi.org/10.1051/epjconf/201611919004, 2016.

Landulfo, E., Cacheffo, A., Yoshida, A. C., Gomes, A. A., da Silva Lopes, F. J., de Arruda Moreira, G., da Silva, J. J., Andrioli, V., Pimenta, A.,
510 Wang, C., Xu, J., Martins, M. P. P., Batista, P., de Melo Jorge Barbosa, H., Gouveia, D. A., González, B. B., Zamorano, F., Quel, E., Pereira, C., Wolfram, E., Casasola, F. I., Orte, F., Salvador, J. O., Pallotta, J. V., Otero, L. A., Prieto, M., Ristori, P. R., Brusca, S., Estupiñan, J. H. R., Barrera, E. S., Antuña-Marrero, J. C., Forno, R., Andrade, M., Hoelzemann, J. J., Guedes, A. G., Sousa, C. T., dos Santos Oliveira, D. C. F., de Souza Fernandes Duarte, E., da Silva, M. P. A., and da Silva Santos, R. S.: Lidar Observations in South America. Part II - Troposphere, in: Remote Sensing, edited by Hammond, A. and Keleher, P., chap. 2, IntechOpen, Rijeka, https://doi.org/10.5772/intechopen.95451,
515 2020.

Lewis, J. R., Campbell, J. R., Welton, E. J., Stewart, S. A., and Haftings, P. C.: Overview of MPLNET Version 3 Cloud Detection, Journal of Atmospheric and Oceanic Technology, 33, 2113–2134, https://doi.org/10.1175/JTECH-D-15-0190.1, 2016.

Licel GmbH: Licel Transient Recorder and Ethernet-Controller Programming Manual, Accessed: February 18, 2023, https://licel.com/manuals/programmingManual.pdf, 2023.

520 Mather, J. M.: Atmospheric Radiation Measurement (ARM) User Facility 2020 Decadal Vision, Tech. Rep. DOE/SC-ARM-20-014, 1782812, https://doi.org/10.2172/1782812, 2021.

Mattis, I., D'Amico, G., Baars, H., Amodeo, A., Madonna, F., and Iarlori, M.: EARLINET Single Calculus Chain – technical – Part 2: Calculation of optical products, Atmospheric Measurement Techniques, 9, 3009–3029, https://doi.org/10.5194/amt-9-3009-2016, 2016.

Nascimento, J. P., Barbosa, H. M. J., Banducci, A. L., Rizzo, L. V., Vara-Vela, A. L., Meller, B. B., Gomes, H., Cezar, A., Franco, M. A.,
525 Ponczek, M., Wolff, S., Bela, M. M., and Artaxo, P.: Major Regional-Scale Production of O_3 and Secondary Organic Aerosol in Remote Amazon Regions from the Dynamics and Photochemistry of Urban and Forest Emissions, Environmental Science & Technology, 56, 9924–9935, https://doi.org/10.1021/acs.est.2c01358, 2022.

National Academies of Sciences, Engineering, and Medicine: Thriving on Our Changing,Planet: A Decadal Strategy for Earth Observation from Space, National Academies Press, Washington, D.C., https://doi.org/10.17226/24938, pages: 24938, 2018.

530 National Geophysical Data Center: U.S. standard atmosphere (1976), Planetary and Space Science, 40, 553–554, https://doi.org/10.1016/0032-0633(92)90203-Z, 1992.

Newsom, R. K., Turner, D. D., Mielke, B., Clayton, M., Ferrare, R., and Sivaraman, C.: Simultaneous analog and photon counting detection for Raman lidar, Applied Optics, 48, 3903, https://doi.org/10.1364/AO.48.003903, 2009.

Pappalardo, G., Amodeo, A., Pandolfi, M., Wandinger, U., Ansmann, A., Bösenberg, J., Matthias, V., Amiridis, V., De Tomasi, F., Frioud,
535 M., Larlori, M., Komguem, L., Papayannis, A., Rocadenbosch, F., and Wang, X.: Aerosol Lidar Intercomparison in the Framework of the EARLINET Project. 3. Raman Lidar Algorithm for Aerosol Extinction, Backscatter, and Lidar Ratio, Applied optics, 43, 5370–85, https://doi.org/10.1364/AO.43.005370, 2004.

Perrone, M. R., Lorusso, A., and Romano, S.: Diurnal and nocturnal aerosol properties by AERONET sun-sky-lunar photometer measurements along four years, Atmospheric Research, 265, 105 889, https://doi.org/https://doi.org/10.1016/j.atmosres.2021.105889, 2022.

540   Press, W. H., ed.: Numerical recipes: the art of scientific computing, Cambridge University Press, Cambridge, UK ; New York, 3rd ed edn., oCLC: ocn123285342, 2007.

Reagan, J., McCormick, M., and SPinhirne, J.: Lidar sensing of aerosols and clouds in the troposphere and stratosphere, Proceedings of the IEEE, 77, 433–448, https://doi.org/10.1109/5.24129, 1989.

Ribeiro, F. N., Oliveira, A. P. d., Soares, J., Miranda, R. M. d., Barlage, M., and Chen, F.: Effect of sea breeze propaga-

545   tion on the urban boundary layer of the metropolitan region of Sao Paulo, Brazil, Atmospheric Research, 214, 174–188, https://doi.org/10.1016/j.atmosres.2018.07.015, 2018.

Rodrigues, P. F., Landulfo, E., Gandu, A. W., and Lopes, F. J. S.: Assessment of aerosol hygroscopic growth using an elastic LIDAR and BRAMS simulation in urban metropolitan areas, AIP Conference Proceedings, 1531, 360–363, https://doi.org/10.1063/1.4804781, 2013.

Román, R., Benavent-Oltra, J., Casquero-Vera, J., Lopatin, A., Cazorla, A., Lyamani, H., Denjean, C., Fuertes, D., Pérez-Ramírez,

550   D., Torres, B., Toledano, C., Dubovik, O., Cachorro, V., de Frutos, A., Olmo, F., and Alados-Arboledas, L.: Retrieval of aerosol profiles combining sunphotometer and ceilometer measurements in GRASP code, Atmospheric Research, 204, 161–177, https://doi.org/10.1016/j.atmosres.2018.01.021, 2018.

Sugimoto, N. and Uno, I.: Observation of Asian dust and air-pollution aerosols using a network of ground-based lidars (ADNet): Realtime data processing for validation/assimilation of chemical transport models, IOP Conference Series: Earth and Environmental Science, 7,

555   012 003, https://doi.org/10.1088/1755-1307/7/1/012003, 2009.

Tanaka, L. M. d. S., Satyamurty, P., and Machado, L. A. T.: Diurnal variation of precipitation in central Amazon Basin, International Journal of Climatology, 34, 3574–3584, https://doi.org/10.1002/joc.3929, 2014.

Vaughan, M. A., Young, S. A., Winker, D. M., Powell, K. A., Omar, A. H., Liu, Z., Hu, Y., and Hostetler, C. A.: Fully automated analysis of space-based lidar data: an overview of the CALIPSO retrieval algorithms and data products, in: Laser Radar Techniques

560   for Atmospheric Sensing, edited by Singh, U. N., vol. 5575, pp. 16 – 30, International Society for Optics and Photonics, SPIE, https://doi.org/10.1117/12.572024, 2004.

Wandinger, U., Freudenthaler, V., Baars, H., Amodeo, A., Engelmann, R., Mattis, I., Groß, S., Pappalardo, G., Giunta, A., D'Amico, G., Chaikovsky, A., Osipenko, F., Slesar, A., Nicolae, D., Belegante, L., Talianu, C., Serikov, I., Linné, H., Jansen, F., Apituley, A., Wilson, K. M., de Graaf, M., Trickl, T., Giehl, H., Adam, M., Comerón, A., Muñoz Porcar, C., Rocadenbosch, F., Sicard, M., Tomás, S., Lange, D.,

565   Kumar, D., Pujadas, M., Molero, F., Fernández, A. J., Alados-Arboledas, L., Bravo-Aranda, J. A., Navas-Guzmán, F., Guerrero-Rascado, J. L., Granados-Muñoz, M. J., Preißler, J., Wagner, F., Gausa, M., Grigorov, I., Stoyanov, D., Iarlori, M., Rizi, V., Spinelli, N., Boselli, A., Wang, X., Lo Feudo, T., Perrone, M. R., De Tomasi, F., and Burlizzi, P.: EARLINET instrument intercomparison campaigns: overview on strategy and results, Atmospheric Measurement Techniques, 9, 1001–1023, https://doi.org/10.5194/amt-9-1001-2016, 2016.

Wang, Z. and Menenti, M.: Challenges and Opportunities in Lidar Remote Sensing, Frontiers in Remote Sensing, 2,

570   https://doi.org/10.3389/frsen.2021.641723, 2021.

Welton, E. J. and Campbell, J. R.: Micropulse Lidar Signals: Uncertainty Analysis, Journal of Atmospheric and Oceanic Technology, 19, 2089–2094, https://doi.org/10.1175/1520-0426(2002)019<2089:MLSUA>2.0.CO;2, 2002.

Welton, E. J., Campbell, J. R., Spinhirne, J. D., and III, V. S. S.: Global monitoring of clouds and aerosols using a network of micropulse lidar systems, in: Lidar Remote Sensing for Industry and Environment Monitoring, edited by Singh, U. N., Asai, K., Ogawa,

575   T., Singh, U. N., Itabe, T., and Sugimoto, N., vol. 4153, pp. 151 – 158, International Society for Optics and Photonics, SPIE, https://doi.org/10.1117/12.417040, 2001.

Whiteman, D. N., Melfi, S. H., and Ferrare, R. A.: Raman lidar system for the measurement of water vapor and aerosols in the Earth's atmosphere, Applied Optics, 31, 3068, https://doi.org/10.1364/AO.31.003068, 1992.

Whiteman, D. N., Demoz, B., Rush, K., Schwemmer, G., Gentry, B., Di Girolamo, P., Comer, J., Veselovskii, I., Evans, K., Melfi,
580   S. H., Wang, Z., Cadirola, M., Mielke, B., Venable, D., and Van Hove, T.: Raman Lidar Measurements during the International H2O Project. Part I: Instrumentation and Analysis Techniques, Journal of Atmospheric and Oceanic Technology, 23, 157–169, https://doi.org/10.1175/JTECH1838.1, 2006.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-
585   Beltran, A., Gray, A. J., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B.: The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data, 3, 160 018, https://doi.org/10.1038/sdata.2016.18, 2016.