Software for analyzing and visualizing laser remote sensing (LIDAR) data

D. POREH[*], L. FIORANI[a]

ENEA guest with ICTP fellowship, ENEA, Via Enrico Fermi 45, 00044 Frascati, Italy
[a]UTAPRAD-DIM, ENEA, Via Enrico Fermi 45, 00044 Frascati, Italy

This paper describes an IDL (Interface Description Language) script for analyzing and presenting laser remote sensing (LIDAR) data, which will enable scientists to monitor and visualize atmospheric phenomena like aerosol load, natural and pollutant gases concentration, and wind speed. The currently available versions of the script produce displays and analyses of the following types: (i) reading the LIDAR data in text and Microsoft Excel formats; (ii) visualizing available data, finding the backgrounds, and averaging the signals with different methods; (iii) applying different wise smoothing methods (like Savitzky-Golay smoothing filter) for ensure that reasonable information would not be lost; (iv) applying different mathematical method to extract the molecule concentration of gases like H_2O, NO, NO_2, CO_2, and SO_2. The output of this code helps scientists to interpret and visualize the natural (like volcanic plume) or human made (like industrial pollutant) aerosols and gases. The presented script is also capable of analyzing extinction coefficients, wind speed, and elevation data – like DEMs (Digital Elevation Models – by bridging with ENVI (a software application currently marketed by ITT for Visual Information Solutions, used to process and analyze geospatial imagery) for presenting the 3D view (topography) of the study area.

1. Introduction

The aim of this paper is to provide an up-to-date, comprehensive IDL (Interface Description Language) widget for LIDAR, focusing on applications to current topics in atmospheric science. First, we start with definition and usage of laser remote sensing of the atmosphere, and then we present the widget that has been developed for analyzing and processing laser data. At the next step, we explain in detail an example that has been carried out in the Stromboli Island, in Italy. A large number of methods are used to measure trace gas concentrations generated by natural or human activities. Among the different chemical and physical methods used for trace gas monitoring, optical methods have in recent years attracted a great deal of attention. In particular, LIDAR gives a unique possibility to measure four-dimensional (4D) trace gas concentrations with a temporal (fourth dimension) and spatial resolution (the other three dimensions) of direct interest for atmospheric modeling [Measures 1984]. This result is usually achieved by the differential absorption LIDAR (DIAL) technique. This method is based on the detection of the backscattered photons from laser pulses transmitted to the atmosphere at two different wavelengths. At one wavelength (OFF), the light is almost only scattered by air molecules and aerosols, whereas at the other one (ON), it is also absorbed by the trace gas under study. The difference between the two recorded signals is thus related to the trace gas concentration. For most molecular species such as H_2O, SO_2, NO_2, and NO, the wavelength separation between the

ON and OFF wavelengths can be smaller than 1/cm, which could be used to make significant differential cross sections for LIDAR environmental monitoring [Measures 1984]. Calculation of extinction coefficients and wind speed along the laser beams are also crucial parameters in laser remote sensing that are presented in detail in this paper [Fiorani 1999].

2. An IDL script for visualizing and analyzing LIDAR data

The IDL script presented in this paper is designed to read in and process text and Microsoft Excel files of LIDAR instruments as input matrixes. IDL needs to be installed on the computer. Prior to running the script, the text files are saved into the data folder. The script enables multiple files with different formats to be processed at the same time. As a default the plot file names and titles on plots are the same as the original data file names in the data folder. The software is also designed to pickup the files by *.txt and *.xls formats. While this is the default setup, it is relatively simple to modify the IDL script to be able to read in files from other instruments and customize where the data files are read from and written to. The IDL script enables the user to have full control of the plotting region e.g. title, axis and colors can all be modified by the user to suit different tastes. A range of 2D and 3D plotting functions (for laser and topography data of the study area) are produced, for visualize LIDAR and DEMs (Digital Elevation Models) simultaneously. Moreover, some buttons have been designed to access ENVI and Itools

(Iplot, Imap, Iimage and Icontour), in order to use the ability of these software applications for more specific usage. For instance, Icontour could be used for interpolation of data based on several methods (like Kriging, inverse distance, etc.). The script writes results (signals intensities, calculated backgrounds, concentrations, filtered signals, etc.) extracted from selected inputs into Itools. With this software, one can finalize all analyses that are necessary for LIDAR interpretations. The user is provided with a detailed tutorial on how to set up and run the IDL script, and some demonstration sample files to familiarize the user with the tool.

The script is also able to read in and process input files that are produced directly from the IDL software company, the only difference being that the script first manipulates the data produced by the instrument into the matrix format. It is relatively simple to modify the script presented in this paper to process non-standard LIDAR data which are not configured in a matrix format.

3. User interface

3.1 Concentration calculation

Once IDL is installed (this software also works with IDL virtual machines, that is free of charge and can be download from the IDL web pages) and started, the script is loaded into the editor window by browsing to the IDL script in the IDL graphical user interface (GUI): file, open script and browse to where the script was saved. OFF and ON signals are read with different buttons and automatically signals are plotted in an Itool (Iplot) with different colors. After reading the two inputs, the user can calculate the background in the following ways:

1- By averaging the signals between -5 and -1 μs.
2- By manually selecting on the graphs.
3- By averaging the signals for long range (> 4 μs).

After calculating the background value, the OFF and ON signals are plotted in a graph together.

For removing the possible noises, beside the conventional smoothing methods, a wise smoothing method, namely the Savitzky-Golay filter is implemented in the software [Press et al. 1992]. This filter was described in 1964 by Abraham Savitzky and Marcel J. E. Golay. In a nutshell, this filter performs a local polynomial regression of degree K on a series of values to determine the smoothed value for each point. Savitzky-Golay filter calculates also the first five derivatives. Despite the other methods, Savitzky-Golay filter preserve the features of the distributions such as relative maxima, minima and width, which are usually flattened by other smoothing techniques (like moving averages).

Fig. 1 shows the Savitzky-Golay filter with 33 point which is used in IDL. The smoothed data is generated by applying the convolution function of the Savitzky-Golay filter.
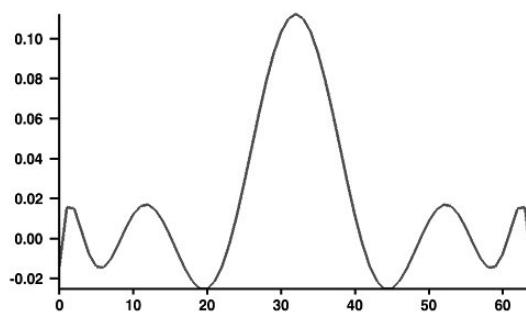

Fig. 1. Savitzky-Golay filter with 33 points which is used in IDL.

This filter is also known as least squares or digital smoothing polynomial (DSP) filter. The filter can use different number of points for this kind of weighted average. When the peaks of the signal are narrow (like for LIDAR data, for example), the Savitzky-Golay filter works very well.

After applying the smoothing methods, the critical information like max(OFF), max(ON), time interval, extremes of the signals (left and right boundary) and integral between extremes for OFF and ON are extracted automatically from graphs.

The concentration C of the targeted molecule at range R from the LIDAR system is given by [Measures 1984]:

$$C(R) = \frac{1}{2[\sigma(\lambda_{ON}) - \sigma(\lambda_{OFF})]} \frac{d}{dR} \ln\left[\frac{n(R, \lambda_{OFF})}{n(R, \lambda_{ON})}\right] \quad (1)$$

where n is the signal intensity measured for both the OFF and ON wavelengths (λ_{OFF} and λ_{ON}, respectively) and σ is the absorption cross section of the molecular species of interest. Lagrange-interpolation derivation is used for deriving the logarithm of the OFF/ON ratio in eq. 1 [Press et al. 1992].

After all analysis, plots can be written to file from the GUI: file, save as and choosing the output format (metafile, postscript, PDF, PNG, BMP, TIFF or JPEG). In the result menu, output data can be visualized and saved to proper formats. Fig. 2 shows the ENEA in operational mode (applying the Savitzky-Golay filter with 33 points to the original signal for smoothing).

3.2 Backscattering coefficient calculation

The strength of the LIDAR signal, i.e. the power registered by a LIDAR depends on the properties of the LIDAR and the properties of the atmosphere [Measures 1984]. The atmosphere can scatter, absorb, and depolarize laser light, shift its wavelength, and broaden its spectral distribution. The parameters that usually relate the measured LIDAR signal with the properties of the atmosphere are the extinction and backscattering coefficients. For the OFF signal, in the case of water vapor detection by CO_2-laser based DIAL, the backscattering coefficient can be calculated as follows.
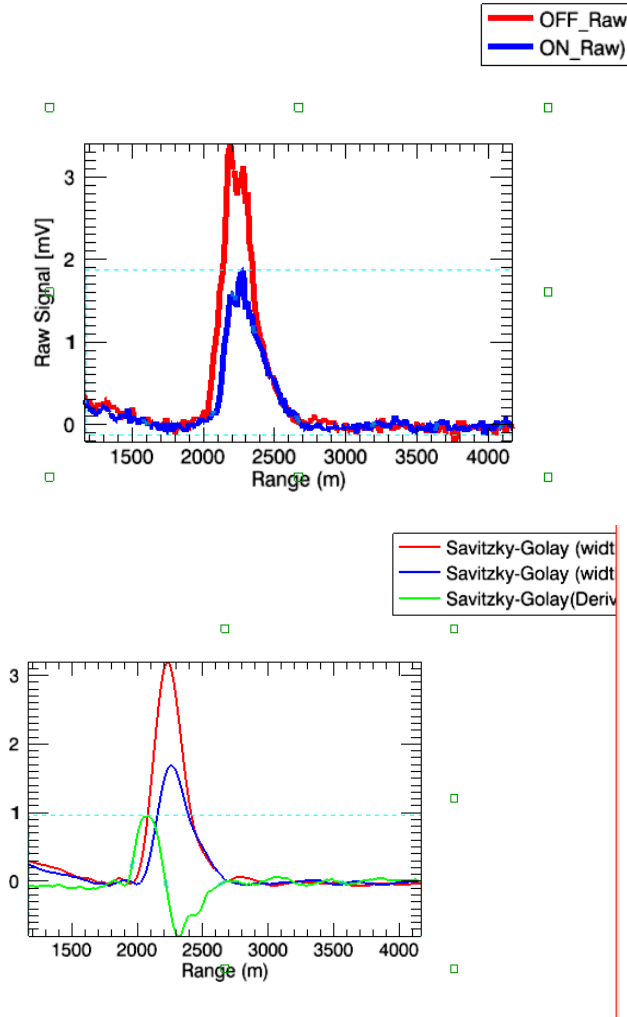
Fig. 2. ENEA software in operational mode.

At first, the range corrected signal is computed:

$$\mu(R) = \ln\left[n(R)R^2\right] \qquad (2)$$

Secondly, by an inversion of the LIDAR equation [Browell et al. 1985], one can find:

$$\frac{d\mu(R)}{dR} = \frac{1}{\beta(R)}\frac{d\beta(R)}{dR} - \frac{2}{P_\pi}\beta(R) \qquad (3)$$

where P_π is the ratio between the backscattering and extinction coefficients. Its value can be regarded as constant and can be found in the literature [Fiorani et al. 2009]. Eventually, by solving eq. 3 for β, one obtains:

$$\beta(R) = \frac{\exp\left[\mu(R) - \mu(R_M)\right]}{\dfrac{1}{\beta(R_M)} + \dfrac{2}{P_\pi}\displaystyle\int_R^{R_M}\exp\left[\mu(R') - \mu(R_M)\right]dR'} \qquad (4)$$

where R_M is selected manually by the user, to calculate the backscattering coefficient until this point, and $\beta(R_M)$ is estimated from values found in the literature [Fiorani et al. 2009].

3. 3 Wind speed calculation

There are several methods to calculate the wind speed based on LIDAR signal, like optical Doppler shift, but in this software we developed the method pioneered by Eloranta et al. [1975].

In order to simplify the formalism underlying the measurement of wind speed, let us consider a discrete distance variable: in the following an index i will be incremented over the spatial bins. First of all the signal is multiplied by the range squared and normalized by the

emitted energy on a shot-per-shot basis, to define a normalized signal $S(R_i)$ independent of the energy fluctuation and proportional to the backscattering coefficient:

$$S(R_i) = n(R_i)\frac{R_i^2}{E_0} \qquad (5)$$

The mean normalized signal $S'(R_i)$ is obtained by averaging N successively recorded individual signals at successive times t_n as:

$$S'(R_i) = \frac{1}{N}\sum_{n=1}^{N} S(R_i, t_n) \qquad (6)$$

The information about the shot-per-shot fluctuations of the signals with respect to $S'(R_i)$ is contained in the fractional deviations:

$$f(R_i, t_n) = \frac{S(R_i, t_n) - S'(R_i)}{S'(R_i)} \qquad (7)$$

In order to determine the displacement of the backscattering inhomogeneities, a cross-correlation function $\rho(L, z_a, z_b)$ is defined with respect to an altitude interval (z_a, z_b) and to a lag value L corresponding to a discrete number of spatial points as:

$$\rho(L, z_a, z_b) = \frac{1}{\sigma\sigma_L}\sum_{i=a}^{b}\sum_{n=1}^{N-1} f(R_i, t_n)f(R_{i+L}, t_{n+1}) \qquad (8)$$

where:

$$\sigma^2 = \sum_{i=a}^{b}\sum_{n=1}^{N-1} f^2(R_i, t_n) \qquad (9)$$

$$\sigma_L^2 = \sum_{i=a}^{b}\sum_{n=1}^{N-1} f^2(R_{i+L}, t_{n+1}) \qquad (10)$$

$$z_i = R_i \sin\theta \qquad (11)$$

$$\Delta R = R_{i+1} - R_i \qquad (12)$$

$$\Delta t = t_{n+1} - t_n \qquad (13)$$

where θ is the elevation angle of the telescope.

The radial wind speed v_R in the altitude interval (z_a, z_b) is determined from the lag value L_{max} for which the correlation is maximum:

$$v_R = L_{max}\frac{\Delta R}{\Delta t} \qquad (14)$$

In the absence of vertical transport, the horizontal wind velocity is then estimated as:

$$v = v_R \cos\theta \cos\Phi \qquad (15)$$

where Φ is the difference in the azimuthal orientation of the wind vector and of the light path.

4. Application example

The example was carried out in the Stromboli Island in southwestern Italy. Fig. 3a shows the SRTM (Shuttle Radar Topography Mission) shaded relief image of southwestern Italy. Stromboli (Fig. 3b) contains the caldera of an active volcano (one of the three active volcanoes in Italy). LIDAR data were collected in Stromboli for monitoring the concentration of H_2O in the volcanic plume. The direction of the laser beam was from east toward west (Fig. 3b). Fig. 3c shows two cross sections of the topography, to illustrate the elevation of the study area. Fig. 3d shows the average of 100 LIDAR returns for OFF and ON separately.

In the plume area, high scattering happens for either OFF and ON pulses. Fig. 3e shows a zoom of the signal in the plume area. For removing the noises, the Savitzky-Golay filter was applied to the original signal (Fig. 3f). Fig. 3g shows the H_2O concentration (in torr) for the study area.

Fig. 3h shows an example of calculated backscattering coefficient for the presented signal. After 2000 meters, the backscattering coefficient severely increases due to the plume particulate.
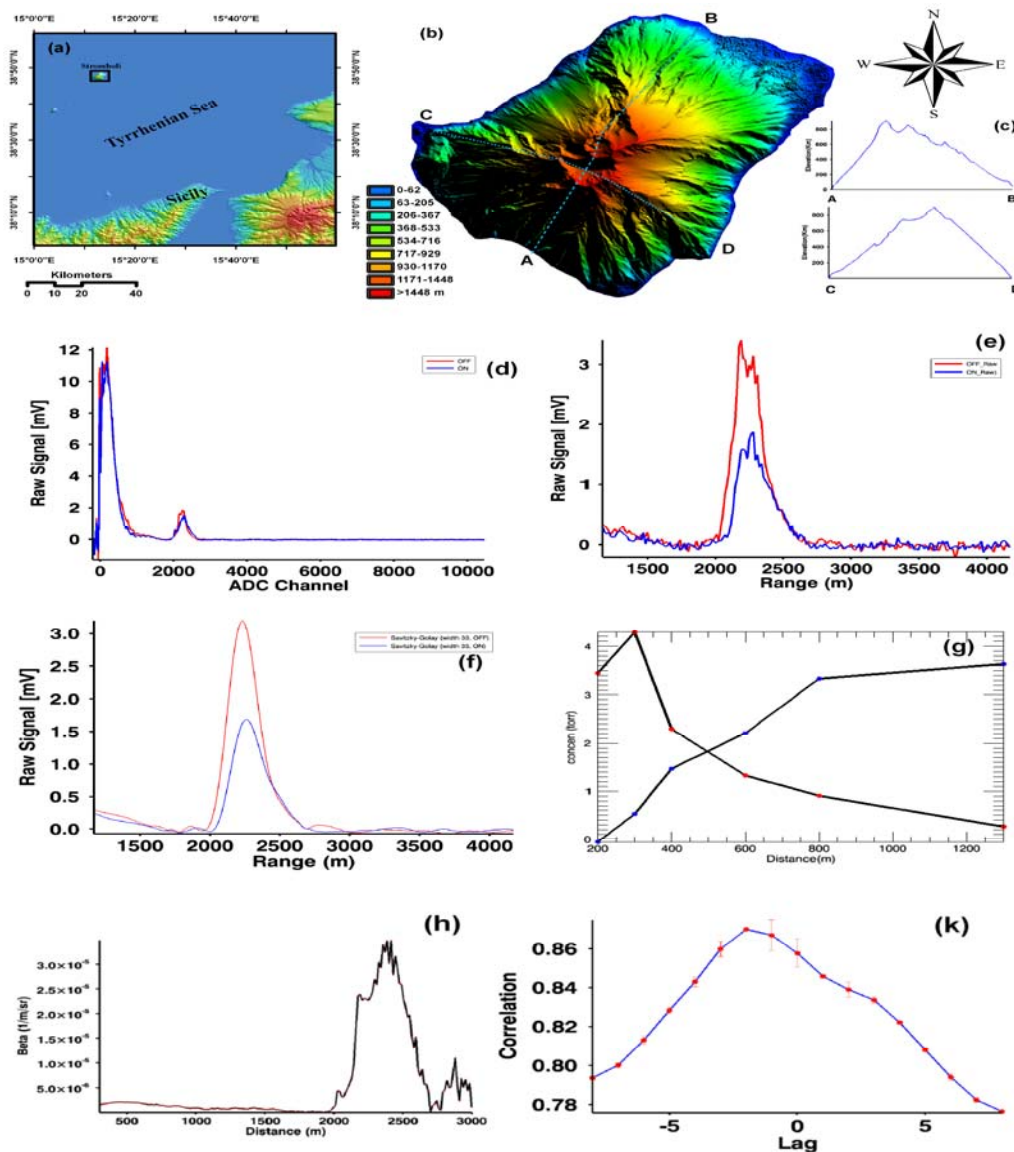
Fig. 3. (a) SRTM shaded relief for southwestern Italy. (b) Stromboli caldera (topographical shaded relief map with 10 m resolution data). (c) Two elevation cross sections perpendicular one to another. (d) LIDAR average signal for the OFF (red line) and ON (blue line) wavelength. (e) Zoom of the LIDAR signal in the plume area. (f) Application of the Savitzky-Golay filter with 33 points to the original signal for smoothing. (g) H_2O concentration calculation with two different methods, with (blue points) and without (red points) Savitzky-Golay filter. (h) Backscattering coefficient. (k) Correlation-lag graph for wind speed computation.

LIDAR wind profiles were computed from OFF signals for 1000 laser shots. Fig. 3k shows one of the correlation-lag results. In this case $L = -2$ corresponding to a velocity along the laser beam of 3 m/s toward the LIDAR system.

scattered signals, to visualize and analyze them in order to get useful information and produce outputs of publication quality. The outputs can be tailored to the user requirements and either visualized in the GUI or written directly to file for future use.

5. Conclusions

This paper has demonstrated the use of an open-source software to visualize and analyze laser remote sensing data. The IDL package is quick and easy to install, and the IDL script is able to rapidly help the user to assess

Software Availability

The last version of the software is available with permission of ENEA in Frascati (Rome) by contacting L. Fiorani (luca.fiorani@enea.it).

References

[1] E. V. Browell, S. Ismail, S. T. Shipley, Applied Optics **24**, 2827 (1985).

[2] E. W. Eloranta, J. M. King, J. A. Weinman, Journal of Applied Meteorology **14**, 1485 (1975).

[3] L. Fiorani, J. Optoelectron. Adv. Mater. **1**(3), 3 (1999).

[4] L. Fiorani, F. Colao, A. Palucci, , Optics Letters **34**, 800 (2009).

[5] R. M. Measures, 1984, Laser Remote Sensing. Fundamentals and Applications, Wiley, New York, US.

[6] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, 1992, Numerical Recipes in C. The Art of Scientific Computing, Cambridge University Press, Cambridge, UK.

_____

[*]Corresponding author: d.poreh@gmail.com