



HAL
open science

STRAT: an automated algorithm to retrieve the vertical structure of the atmosphere from single channel lidar data

Yohann Morille, Martial Haeffelin, Philippe Drobinski, Jacques Pelon

► **To cite this version:**

Yohann Morille, Martial Haeffelin, Philippe Drobinski, Jacques Pelon. STRAT: an automated algorithm to retrieve the vertical structure of the atmosphere from single channel lidar data. *Journal of Atmospheric and Oceanic Technology*, 2007, 24, pp.761-775. hal-00147430

HAL Id: hal-00147430

<https://hal.science/hal-00147430v1>

Submitted on 17 May 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

STRAT: an automated algorithm to retrieve the vertical structure of the atmosphere from single channel lidar data

^{1,*}Morille Y., ¹Haeffelin M., ²Drobinski P., ²Pelon J.

1. Laboratoire de Météorologie Dynamique, Institut Pierre-Simon Laplace, Palaiseau, France.
2. Service d'Aéronomie, Institut Pierre-Simon Laplace, Paris, France.

JAOT: manuscript No. A-846-H

() Corresponding author.*

Laboratoire de Météorologie Dynamique, Institut Pierre Simon Laplace,
Ecole Polytechnique, 91128 Palaiseau Cedex, France.

Tel.: +33 169 33 31 94; fax: +33 169 33 30 05.

E-mail address: morille@lmd.polytechnique.fr

ABSTRACT:

Today several lidar networks around the world provide large data sets that are extremely valuable for aerosol and cloud research. Retrieval of atmospheric constituent properties from lidar profiles requires detailed analysis of spatial and temporal variations of the signal. This paper presents an algorithm called STRAT (STRucture of the ATmosphere) designed to retrieve the vertical distribution of cloud and aerosol layers in the boundary layer and through the free troposphere and to identify near particle free regions of the vertical profile and the range at which the lidar signal becomes too attenuated for exploitation, from a single lidar channel. The paper describes each detection method used in the STRAT algorithm and its application to a tropospheric backscatter lidar operated at the SIRTa observatory, in Palaiseau, 20 km south of Paris, France. STRAT retrievals are compared to other means of layer detection and classification; retrieval performances and uncertainties are discussed.

1. Introduction

Remote sensing of the atmosphere by Lidar dates from the 1960's. Lidars have since become very popular remote sensing tools. They are used extensively both in ground-based and airborne configurations to get detailed vertical profiles of cloud or aerosol properties, such as the extinction coefficient to learn more about the nature of particles present aloft. Lidars are commonly deployed during intensive observation field experiments well as for routine long-term monitoring.

Today several global and regional networks of atmospheric observatories exist that use lidars as their main monitoring instrument. The European Aerosol Research Lidar Network (EARLINET) monitors aerosol transport over Europe based on 21 Rayleigh/Mie and Raman lidar stations (Boesenberg et al. 2003). Similarly the National Institute for Environmental Studies of Japan coordinates a network of 12 automatic dual wavelength polarization lidars continuously to study Asian dust transport (Shimizu et al., 2004). The continent between Europe and Asia is covered by the CIS-LiNet project that aims at aerosol and ozone monitoring through a network of 6 lidar stations since the beginning of 2004. Building from the research capabilities already established at a number of Eastern North America lidar facilities, a Regional East Atmospheric Lidar Mesonet (REALM) has been proposed to monitor air quality in the vertical from multiple locations in that region (Hoff et al., 2002). The National Aeronautics and Space Administration funds a Micro-Pulse Lidar Network (MPLNet ; Welton et al., 2001) to monitor tropospheric aerosols based 20 micropulse lidars located at climatologically diverse locations on the planet and coordinated with the Aerosol Robotic Sunphotometer Network (AERONET). The Network for the Detection

of Stratospheric Change (NDSC) monitors stratospheric ozone using lidars that range through the stratosphere.

Continuous or routine operations of such systems produce upwards of 1000 vertical profiles per day for each channel of each system. All systems produce fundamental information about the structure of the atmosphere, that is the vertical distribution of the particle layers from near the ground to the top of the lidar range. Those are basic parameters, yet our current understanding of the vertical distributions of cloud and aerosol layers in the atmosphere remains limited by lack of large-scale analysis of available data sets. Furthermore, in the context of new satellite missions carrying active remote sensing payloads, such as the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO, Winker et al., 2003), ground-based lidar observatories have an increasingly important role to play in establishing regional climatologies that will tie in the temporally sparse global-scale satellite measurements.

The need for robust algorithms designed to process large lidar data sets has been established for over a decade (e.g. Platt et al., 1994). Several authors have exploited long-term lidar datasets to derive statistics and climatologies of cloud and aerosol macrophysical properties (e.g. Cadet et al., 2003; Comstock et al., 2002; Sassen et al., 2001). In such analyses, particle layers, being clouds or aerosols, are typically detected using slope change in the lidar profile (e.g. Flamant et al., 1997; Shimizu et al., 2004), or comparing the lidar power return to an expected clear-sky value (e.g. Clothiaux et al., 1998). Several authors developed methods using wavelet transforms to identify particle layers in the lidar profile (e.g. Cohn and Angevine, 2000; Brooks, 2003), but those studies were limited to boundary layer height analyses.

In this paper we present an end-to-end algorithm designed to retrieve an ensemble of basic parameters of the particulate atmosphere from single channel lidar profiles. The algorithm is designed with modules that retrieve the height of the boundary layer, the vertical distribution of particle layers (clouds and aerosols), and that identify layers that are near particle free. Each retrieval taken independently is not new, however, the added value comes from the use of multiple analyses on each lidar backscatter profile to derive a self-consistent classification. The specifications of this algorithm are:

- To detect all particle layers from near ground to the top of the lidar range, that is, cloud and aerosol layers in the free troposphere and the vertical extent of the boundary layer from the raw uncalibrated lidar backscatter profile.
- To identify pristine layers predominantly populated by molecules, and hence considered as near particle free, so that the lidar signal can be calibrated (with respect to a calculated molecular backscatter) in an automated manner for further inversion of the signal.
- To be widely applicable to any single- or multi-wavelength ground-based lidar system (micro- to milli-joule pulse energy, high or low pulse frequencies, U.V., visible, near-infrared wavelength) by using the signal-to-noise ratio as only threshold reference.
- To be automatic and robust so that long lidar time series can be processed in an operational environment.

The objective of this algorithm development is to provide a tool that can be used to process large ground-based lidar datasets and hence give access to a very detailed

macrophysical classification of the particulate atmosphere. This is the first and necessary step towards extending regional statistics of cloud and aerosol macrophysical properties.

In Section 2, we present the lidar measurement data set used to develop the STRAT algorithm. Section 3 is dedicated to the description of the algorithm and Section 4 to the evaluation of the algorithm. Conclusions are provided in Section 5.

2. Data description

To develop the STRAT algorithm we use a database produced by the SIRTA¹ observatory, a facility dedicated to observing the atmosphere in support of cloud and aerosol research. SIRTA is located in Palaiseau, 20 km south of Paris, France; the data can be retrieved through its web site (www.sirta.fr). SIRTA gathers active and passive remote sensing instruments to retrieve optical, radiative and dynamic properties of the atmosphere and its constituents. This includes a dual wavelength polarisation lidar LNA Lidar for Lidar Nuages Aérosols, that is Cloud Aerosol Lidar), a millimetre wave Doppler radar, a near-IR ceilometer, a surface broadband flux station, and standard weather measurements (Haeffelin et al., 2005). The SIRTA database includes also radiosonde profile data produced by Météo-France 15 km from the site as part of their national operational network (00 and 12 UT).

The main dataset used to develop STRAT is that of the LNA lidar. Table 1 lists its technical characteristics. LNA lidar profiles corrected for electronic noise, atmospheric background signal and range divergence are available in the SIRTA data base in netcdf format for the period starting March 2002 through today (3000 hours).

¹ Site Instrumental de Recherche par Télédétection Atmosphérique

These level-1 files contain lidar backscatter profiles for 6 different channels (see Table 1) along with a quality flag that indicates when signal-to-noise ratio drops below 3 (see Section 3.1). Several studies (Hodzic et al., 2004; Naud et al., 2004; Chiriaco et al., 2005) are based on this lidar dataset and the output of the STRAT algorithm described in this paper. The lidar signal analysis by STRAT makes use of density profiles derived from radiosounding data (temperature and pressure) to simulate the theoretical lidar backscatter solely due to molecules. Two radiosonde profiles per day are enough to capture most density profile variations.

In section 4 we use ceilometers measurements collocated with the LNA lidar to perform comparisons of cloud detection and cloud base height retrievals. The ceilometer uses a laser diode with an emitted wavelength of 855nm, its pulse repetition frequency is 6494 hz with maximum emitted light power of $50 \mu\text{W}\cdot\text{m}^{-2}$ and its optical divergence is 1.2 mrad. Cloud information is retrieved through a Vaisala proprietary algorithm that looks for high backscatter values in the profile.

In section 4 we also compare STRAT retrievals of the boundary layer height with estimations based on potential temperature profiles. For that we derive high-resolution (10 m) temperature and pressure profiles from the standard significant points provided in the Météo-France file.

3. Algorithm description

The STRAT algorithm includes four successive detections carried out on individual lidar profiles:

- noise detection, to identify where the lidar signal becomes too weak (signal-to-noise ratio) for further exploitation, so that data is not misinterpreted.
- molecular layer detection, to identify regions that are near particle free, i.e. populated by molecules only or with a negligible concentration of particles, because lidar calibration is carried out in those regions by comparison with computed molecular backscatter profiles. This will enable lidar data users to automate processes that require calibration such as inversion of extinction profiles or calculation of layer optical depth or investigation of cloud thermodynamic phase using depolarization ratio.
- particle layer detection with separate cloud and aerosol layer identification, so cloud and aerosol processes can be studied
- boundary layer detection, so boundary layer process studies can be addressed.

Finally, a flag variable is derived with value 0 for noise, 1 for molecular layers, 2 for boundary layer, 3 for aerosol layers, 4 for cloud layers and 10 for unidentified layers.

Figure 1 shows a diagram of the STRAT algorithm data processing.

a. Noise detection

1) Method

A simple signal-to-noise ratio threshold allows us to determine where the signal is too noisy to extract information from the lidar measurement. For each profile, it has been shown (Durieux et al. 1998) that the noise level $\sigma(r,t)$ of the measured signal, assuming all its components are independent, can be written as

$$\sigma(r, t) = \sqrt{\sigma_{sig}(r, t)^2 + \sigma_b(r, t)^2 + \sigma_d(r, t)^2},$$

1

where $\sigma_{sig}(r, t)^2$ is the shot noise induced by and proportional to the backscattered lidar signal, $\sigma_b(r, t)^2$ is the shot noise resulting from background light and $\sigma_d(r, t)^2$ is the shot noise resulting from dark current.

Assuming that $\sigma_{sig}(r, t)^2$, $\sigma_b(r, t)^2$ and $\sigma_d(r, t)^2$ are proportional to the backscattered lidar power $P(r, t)$, to the background noise power $B_b(r, t)$ and to the dark current noise power $B_p(r, t)$, respectively, equation 1 becomes

$$\sigma(r, t) = C \sqrt{P(r, t) + B_b(r, t) + B_p(r, t)},$$

2

where C is a proportionality coefficient to be determined. Assuming that $B(r, t)$, defined as the sum of $B_b(r, t)$ and $B_p(r, t)$, is constant along the profile, the noise level can be estimated at the altitude range where there is no lidar return ($P(r, t)=0$) by computing the signal standard deviation $\sigma_{P=0}(r, t)$ of $P(r, t)$ at this range.

The proportionality coefficient C can thus be estimated by averaging the ratio derived from equation 2 over a given number of points, as

$$C = \frac{\overline{\sigma_{P=0}(r, t)}}{\sqrt{B(r, t)}} = \frac{1}{N} \sum_1^N \frac{\sigma_{P=0}(r, t)}{\sqrt{B(r, t)}}.$$

3

The noise level can then be expressed as

9

$$\sigma(r, t) = \frac{\sigma_{P=0}(r, t)}{\sqrt{B(r, t)}} \sqrt{P(r, t) + B(r, t)} .$$

4

So the signal-to-noise ratio $SNR(r, t)$ can be written as

$$SNR(r, t) = \frac{P(r, t)}{\frac{\sigma_{P=0}(r, t)}{\sqrt{B(r, t)}} \sqrt{P(r, t) + B(r, t)}} .$$

5

Note that C can also be derived from pre-trigger data if available. This retrieval can be checked against the result of equation 3 for consistency.

2) Threshold determination

In the STRAT algorithm, the signal is considered to be too noisy for further analysis when the SNR falls below a threshold of $T_i=3$. Indeed, for a gaussian noise 99% of values are contained in the interval $\pm 3\sigma_{P=0}(r, t)$. Boundary layers, molecular layers, and cloud and aerosol layers will be detected on the part of the signal that is above that threshold. For systems with very low signal to noise ratio, the algorithm must be applied to time averaged profiles.

Note that if $B(r, t)$ is large (e.g. daytime) $\sigma_{P=0}(r, t)$ can be used as the signal noise for convenience. This simplification does not introduce significant error in determining if $SNR(r, t)$ is less or greater than 3, as for small SNR values $\sigma(r, t)$ is quite close to $\sigma_{P=0}(r, t)$ because $P(r, t)/B(r, t)$ tends toward 0.

b. Molecular layer detection

1) Method

Identification of particle-free or near particle-free layers is of particular importance as they are often used in lidar calibration algorithms (e.g. Sassen et al. 1989; Platt 1979). For simplicity, these layers will be labelled as molecular layers although they may contain aerosols in small quantities (see Section 4.a. for discussion). The identification algorithm for molecular layers is based on analysis of the variability of the lidar signal around a theoretical molecular backscatter profile computed from pressure and temperature profiles. Thermodynamic profiles can be obtained from collocated atmospheric sounding measurements or extracted from model analysis data.

A normalization coefficient $K(r,t)$ is estimated at each range of the lidar signal as

$$K(r,t) = \frac{\sum_{r'=r-N}^{r+N} \frac{\beta_{mol}(r',t)}{\beta_{lidar}(r',t)}}{2N+1}$$

6

where

$$\beta_{lidar}(r,t) = P(r,t)r^2$$

7

is the non-normalized attenuated backscatter coefficient, $\beta_{mol}(r,t)$ is the computed theoretical molecular backscatter coefficient and $2N+1$ is the size of the averaging window.

The variability, $V(r,t)$, of the normalized lidar signal around the molecular backscatter profile at a range r is determined as

$$V(r, t) = \frac{\sum_{r'=r-N}^{r+N} \left(\frac{1}{r'^2} \left[\beta_{lidar}(r', t) - \frac{1}{K(r, t)} \beta_{mol}(r', t) \right] \right)^2}{2N+1},$$

8

where K is the coefficient used to normalize the lidar profile to the molecular profile in the averaging window. A sensitivity study showed 20 gates (e.g. 300m for the LNA lidar) to be a wide enough averaging window to detect only atmospheric variations.

In molecular layers, the lidar backscatter signal $\beta_{lidar}(r,t)$ can be expressed as the sum of molecular backscatter $P'_{mol_lidar}(r,t)$ and an additional zero mean noise $M(r,t)$:

$$\beta_{lidar}(r, t) = \left(P'_{mol_lidar}(r, t) + M(r, t) \right) r^2.$$

9

Hence, in molecular layers Eq. 8 becomes

$$V(r, t) = \frac{\sum_{r'=r-N}^{r+N} \left(M(r', t) \right)^2}{2N+1}.$$

10

The variability $V(r,t)$ is only due to the noise variability and so can be compared to the noise variance $\sigma(r,t)^2$. So, a threshold value $V_{thr}(r,t)$ can be defined with respect to the noise variance $\sigma(r,t)^2$ as

$$V_{thr}(r, t) = T_2 \sigma(r, t)^2$$

11

12

where T_2 is the molecular layer threshold coefficient. Thus, if $V(r,t)$ is below the threshold value $V_{thr}(r,t)$, we consider that the lidar backscatter power is characteristic of a molecular layer.

2) Threshold determination

Here again we use $\sigma_{P=0}(r,t)$ as a substitute for $\sigma(r,t)$ because in molecular layers the two values are quite close. Indeed, in our case $P(r,t)$ is typically lower than the background backscatter $B(r,t)$ and hence $\sigma(r,t)$ tends toward $\sigma_{P=0}(r,t)$.

To determine this threshold, values of $V(r,t)/\sigma_{P=0}(r,t)^2$ (Eq. 8 and 11) have been computed from one profile per day over the available database (October 2002 - October 2005) only on signal values with SNR greater than 3.

Figure 2 illustrates the probability density function (PDF) of $V(r,t)/\sigma_{P=0}(r,t)$ at two wavelengths (532 and 1064 nm) with $2N+1=21$ gates = 315m. The PDF of the function $V(r,t)$ expressed in Eq. 10 is also represented in Figure 2 with a solid line. The simulated noise used to estimate this PDF is a Gaussian noise similar to the real one.

Distributions are divided in two separate regions. The first one is a narrow Gaussian distribution ranging between 0 and 3. It can be associated with molecular layers where the variability is smaller than in particle layers. The second one is a very broad distribution of $V(r,t)/\sigma_{P=0}(r,t)^2$ with values greater than 3. Because the objective is to use molecular layers for calibration, it is important not to falsely detect stable particle layers as molecular layers. For our application, a threshold value $V_{thr}(r,t)$ of $3\sigma_{P=0}(r,t)^2$ is a good compromise to separate molecular layers from other layers. The difference

in distribution width between the 532 and 1064nm curves of Figure 2 can be attributed to the signal quality of the two different channels.

c. Cloud and aerosol layer detection

1) Method

The majority of particle layer detection techniques described in the literature use thresholding tests on the first derivative of the backscatter intensity (e.g. Pal et al. 1992). Such methods give satisfying results as long as the signal-to-noise ratio remains high. Others techniques use algorithms that depend on cloud type (e.g. Chazette et al. 2001). While they are suited for case studies, they cannot be used for automated detection. In the STRAT algorithm, a combination of wavelet transform and Pr^2 ratio thresholding is used to identify particle regions in lidar profiles.

The Continuous Wavelet Transform (CWT) is used to detect discontinuities in the lidar signal as the base, the top and the peak backscatter of individual particle layers. This method, based on seeking high correlation between the lidar signal and the wavelet characterized by the “Mexican hat” shape for each range and for each scale, is inspired on studies by Mallat et al. (1992) and an algorithm developed by Brooks (2003). The “Mexican hat” wavelet $\Psi(r)$, shown in Figure 3, is the second derivative of a gaussian. It is used because its shape is very similar to the shape of the lidar signal backscattered by cloud or aerosol layers. Additionally, “Derivates of Gaussians are most often used to guarantee that all maxima lines propagate up to the finest scales”, which is not the case of the Haar wavelet (Mallat et al., 1992).

Firstly, the CWT is computed for each $P(r,t)$ profile as

$$CWT_{a,b}(r,t) = \sum_r P(r,t) \psi_{a,b}(r)$$

12

where

$$\psi_{a,b}(r) = \frac{1}{\sqrt{a}} \Psi\left(\frac{r-b}{a}\right),$$

13

where a is the wavelet dilation (or scale) and b is the location of its center. CWT coefficients can be interpreted as a correlation coefficient between the wavelet (centred on b and scaled by a) and the signal $P(r,t)$.

Secondly, the modulus of CWT coefficients are determined to extract the lines of modulus maxima of the $CWT_{a,b}(r,t)$ that are lines (or ridges) formed by all maxima found at all dilations. This skeleton of the CWT, formed by all ridges, represents the highest correlation and anti-correlation between the signal $P(r,t)$ and the wavelet from the largest to the finest scale.

Figure 4 illustrates the performance of this method for an ideal cloud or aerosol backscatter profile (Figure 4a) and a real one (Figure 4d). Figure 4b and Figure 4e show CWT coefficients for the ideal profile and the real one, respectively. Figure 4c and Figure 4f show the corresponding maxima lines. Ridges of highest correlation and anti-correlation coefficients propagate to the finest scale at the base and top of each particle layer, as well as at the location of the maximum backscatter.

So, each ridge shown in Figure 4c is associated with a discontinuity of the $P(r,t)$ signal. The value $M_{CWT}(i_{ridges})$ of the average CWT coefficients along this ridge allows us to discriminate a backscatter peak from a layer base or top as

15

$$M_{CWT}(i_{ridg}) = \overline{CWT}_{a,b}(r,t) \Big|_{a,b \in i_{ridg}} \begin{cases} \dot{=} 0: \text{ layer peak} \\ \dot{=} 0: \text{ layer base or top} \end{cases}$$

14

For each identified backscatter peak, the base (top) of the same layer can be found by looking for the first base (top) detected below (above). If the top of one layer is the base of the next one, the STRAT algorithm is designed to link these two layers into a single one with a peak defined as the maximum P value of the two original peaks.

Finally, we apply a threshold value R_{thr} on the difference of backscatter power between peak height and base height defined as

$$R = P(r_{peak}) - P(r_{base}) > R_{thr} .$$

15

This threshold removes over-detections that are due to noise variations such as the discontinuities detected between points 200 and 350 shown in Figure 4f. The R_{thr} threshold implemented in the STRAT algorithm is derived with respect to the noise level $\sigma(t)$ as

$$R_{thr} = T_3 \sigma(t)$$

16

where T_3 is the particle layer threshold coefficient.

2) Threshold determination

As this threshold is used to identify false peak/base detections in layers with low backscatter signal (i.e. molecular layers) we use $\sigma_{P=0}(r,t)$ as a substitute for $\sigma(r,t)$.

A PDF of $R(r,t)/\sigma(t)$ values is derived from the LNA lidar database (10 profiles per day) to determine the R_{thr} . The PDF is shown in Figure 5. Because of noise-related

16

signal variations, discontinuities can also be detected in molecular layers but corresponding $R(r,t)$ values are smaller than for particle layers. The PDF of $R(r,t)/\sigma(t)$ values for discontinuities identified in molecular layers are also represented in Figure 5 (solid line). This curve is obtained by processing simulated noisy molecular profiles derived from radiosonde data. The distribution is divided in two separate regions. The gaussian distribution between 0 and 10 that contains 86% of the detections is due to noise variations. This effect corresponds to the many short CWT ridge lines shown in Figure 4f. Values $R(r,t)/\sigma(t)$ for true particle layers are logically greater. Picking a threshold value R_{thr} of $10\sigma(t)$ allows us to remove over-detections.

d. Cloud and aerosol distinction

1) Method

Aerosol and cloud layers can have similar signatures in lidar backscatter profiles. However, for near I.R, I.R, visible and U.V. wavelengths, the lidar backscatter power is generally greater for liquid-water and optically thick ice for clouds than for aerosols. The cloud and aerosol distinction algorithm is based on the study by Wang et al (2001) who applied a threshold on the peak Pr^2 to base Pr^2 ratio. The ratio is expressed as

$$dPr^2 = \frac{P(r_{peak}, t)r_{peak}^2}{P(r_{base}, t)r_{base}^2} .$$

17

Ratios greater (less) than a threshold T_4 classify a layer as cloud (aerosols). Figure 6a shows a 7-hour time series of 532nm backscatter power profiles. The

17

measurements show significant backscatter between ground and 2500m. After 11.30 UT one can see several occurrences of very strong extinction that are characteristic of dense water clouds. Figure 6b shows profile-by-profile dPr^2 ratios for the areas identified as particle layers. It reveals a large profile-by-profile variability of dPr^2 values.

To improve this method, we derive average dPr^2 values for a given object (cloud or aerosol layer). To obtain this averaged value on a vertically and temporally consistent particle layer, range-time processing is required.

An average dPr^2_{layer} value is computed for each identified particle layer and the T_4 threshold is applied to this average value to separate cloud from aerosol layers as

$$\begin{cases} dPr^2_{layer} > T_4 & \text{then layer is cloud layer} \\ dPr^2_{layer} < T_4 & \text{then layer is aerosol layer} \end{cases} .$$

18

Figure 6c shows the dPr^2_{layer} values for cloud and aerosol layers observed on 26 May 2003. Some particle layers appear with significantly stronger dPr^2_{layer} than others.

2) Threshold determination

The PDFs of dPr^2_{layer} values are shown in Figure 7 for three different vertical range intervals. The distribution based on the complete vertical range (0 to 15 km) is represented by a solid line, the distribution of dPr^2_{layer} for layers below 7.5 km is shown in dashed line with square markers and the distribution of dPr^2_{layer} for layers above 7.5km is drawn with a dashed line and diamond markers. Those intervals of altitude are used because except for exceptional events like volcanic eruptions, we assume

18

that aerosol concentrations are not sufficient to be detected above 7.5km whereas cloud layers extend from 0 to 15km. So, distributions of values dPr^2_{layer} are due to different contributions: under 7.5km a combination of cloud and aerosol contributions, above 7.5 km only cloud contributions. Thus, the distribution of dPr^2_{layer} for aerosols layers is located between 1 and 4 where the two dashed lines are distinct whereas the distribution of dPr^2_{layer} for clouds layers is wider. To separate cloud from aerosol layers based on these distributions, we select a threshold value $T_4=4$. So equation 18 becomes

$$\begin{cases} dPr^2_{layer} > 4 \text{ then layer is cloud layer} \\ dPr^2_{layer} < 4 \text{ then layer is aerosol layer} \end{cases} ,$$

19

for particle layers below 7.5km. Above 7.5km we assume that 100% of particle layers correspond to clouds.

e. Boundary layer height detection

1) Method

The atmospheric boundary layer (ABL) is the lowest part of the troposphere which is directly influenced by the Earth's surface and responds on short time scales to surface forcing. This is the region which is well-mixed due to convectively driven mixing. Several boundary layer height (BLH) detection methods are described in the literature. Methods using a simple signal threshold (e.g. Melfi et al. 1985; Boers et al. 1988) are not appropriate for cases with varying aerosol extinction. Methods based on gradient properties at the top of the boundary layer (e.g. Flamant et al. 1997) need

19

averaged or smoothed signals and so lose resolution. In the presence of boundary layer clouds, all methods, including wavelet analysis methods (e.g. Cohn et al. 2000; Brooks 2003) are likely to identify the top of the cloud as BLH because the strongest gradient (or correlation) will occur in that part of the profile. In the STRAT algorithm we use the output of the molecular layer module and particle layer module to help distinguish the low-altitude clouds from the boundary layer below them. The boundary layer height detection method used in the STRAT algorithm, is similar to the particle layers detection method described in section c. It is inspired by the work of Mallat et al. (1992) and Brooks (2003).

The wavelet used here is the first derivative of a gaussian $\Psi'(r)$, shown in Figure 8 because its shape is very similar to the negative gradient of the backscatter signal at the top of the boundary layer during daytime. A standard boundary layer backscatter signal is shown in Figure 9a. As for particle layer detections, the CWT is computed for each $P(r,t)$ profile as

$$CWT'_{a,b}(r,t) = \sum_r P(r,t) \Psi'_{a,b}(r)$$

20

where

$$\Psi'_{a,b}(r) = \frac{1}{\sqrt{a}} \Psi'\left(\frac{r-b}{a}\right)$$

21

a is the wavelet dilation (or scale) and b is the location of its center.

Then modulus maxima lines of the $CWT'_{a,b}(r,t)$ are also determined to detect all gradients in the backscatter signal. Because of the wavelet shape, negative gradients

20

can be discriminated from positive ones using average values of the CWT' coefficients along this ridge, as follow:

$$M'_{CWT}(i_{ridg}) = \overline{CWT'_{a,b}(r,t)}|_{a,b \in i_{ridg}} \begin{cases} > 0: \text{ positive gradient} \\ < 0: \text{ negative gradient} \end{cases}$$

22

The detection of negative gradients combined with the altitude of the lowest molecular range, H_{min_mol} (shown in Figure 9), and the base height of the lowest particle layer (H_{min_part}) allows us to estimate the boundary layer height. Four different cases, illustrated in Figure 9a,c,e, must be considered:

- if $H_{min_mol} < H_{min_part}$ (molecular layer below the lowest identified particle layer)
 - o there exists a ridge with $M'_{CWT} < 0$ that propagates up to a range $r < H_{min_mol}$: BLH is the range r (if there is more than one ridge, only the ridge with the minimum M'_{CWT} value is kept). This case is illustrated in Figure 9a with an example of a standard lidar backscatter signal and in Figure 9b with the corresponding wavelet coefficients.
 - o there does not exist a ridge with $M'_{CWT} < 0$ that propagates up to a range $r < H_{min_mol}$: BLH is undefined
- if $H_{min_part} < H_{min_mol}$ (molecular layer above the lowest identified particle layer)
 - o there exists a ridge with $M'_{CWT} < 0$ that propagates up to a range $r < H_{min_part}$: BLH is the range r (if there is more than one ridge, only the ridge with the minimum M'_{CWT} value is kept). A cloud or aerosol layer is located near the top of the BL. This case is illustrated in Figure 9c with an example of a backscatter signal where the molecular layer is above

21

the lowest identified particle layer. Figure 9d shows with the corresponding wavelet coefficients.

- there does not exist a ridge with $M'_{CWT} < 0$ that propagates up to a range $r < H_{min_part}$: BLH is H_{min_part} . A cloud or aerosol layer is located at the top of the BL. This case is illustrated in Figure 9e with an example of a backscatter signal where a cloud layer is present at the top of the BL.

Figure 9f shows with the corresponding wavelet coefficients.

After daytime convection ceases, aerosol layers become stratified and multiple layers can form near the surface (boundary and residual layers). In such situations, the STRAT algorithm is not able to distinguish the top of the boundary layer and the top of the residual layer.

4. Evaluation of the STRAT algorithm

a. Evaluation of the molecular layer detection

Here we evaluate if layers identified by STRAT as molecular layers contain any additional extinction due to the presence of some quantity of aerosols. To do so we apply a classic approach of optical thickness estimation (Platt, 1973) that is based on the ratio of the lidar power attenuation from the base to the top of the molecular layer to a theoretical molecular attenuation. Analysis of 4 years of SIRTALidar profiles containing molecular layers extending more than 1 km reveals that these layers exhibit attenuation uncertainties of $\pm 2.10^{-5} \text{ m}^{-1}$ in terms of equivalent extinction. Hence the parameters used in the molecular layer detection module (see Table 2) imply that STRAT will allow layers whose attenuation is somewhat different from that

of theoretical molecular layers to be identified as particle free. As a result such layers could contain up to $2 \cdot 10^{-5} \text{ m}^{-1}$ particle extinction, equivalent to a 0.02 optical depth for a 1 km deep layer. This uncertainty can be reduced by increasing the test range (e.g. from 300 to 500 m) or by reducing the variability threshold (T_2).

b. Cloud and aerosol layer detection

1) Performance evaluation based on simulated data

Figure 10 shows results obtained by the STRAT algorithm cloud and aerosol layer detection with a simulated backscatter profile containing a cloud (Figure 10a). Two slopes S_1 and S_2 are used to describe the majority of cases, where S_1 is the molecular slope and S_2 is the slope of the backscatter profile in the cloud between the base and the peak. Figure 10b and c illustrate results obtained on CBH and CTH detection, respectively. We describe results obtained for slopes between $-0.5 \cdot 10^{-10}$ and $-2 \cdot 10^{-10} \text{ m}^{-1} \cdot \text{sr}^{-1} / \text{m}$ for S_1 and between $1 \cdot 10^{-8}$ and $7 \cdot 10^{-8}$ for S_2 . The CBH detection is sensitive to both slopes but the maximum resulting error is - 3 gates (- 45m for LNA lidar profiles). CTH detection depends on the slope of backscatter in the particle layer. CTH errors are biased high between 0 and 5 gates (0 to 75m for LNA lidar data) for the largest S_2 values.

2) Comparison with retrievals from ceilometer

We compare cloud base height retrievals derived by applying STRAT to LNA lidar data to those derived by a VAISALA ceilometer located nearby (100m). Figure 11a shows the PDF of cloud base height for the two systems based on 12 months of

observation. We limit our comparisons to situations where each retrieval is consistent for 10 minutes. We note that the LNA misses clouds below 1300m, due to the large overlap function. Above 5000m, the ceilometer data becomes unreliable due to limited power. Comparisons of cloud and cloud free occurrence detections by the two systems for the 1300-5000m vertical range are shown in Table 3. In situations labelled as cloud free by the ceilometer, we find 92% agreement, and 8% cloud detection by LNA/STRAT. In situations where the ceilometer detects a cloud between 1300 and 5000m, LNA/STRAT detects a particle layer 93% of the time with 74% clouds and 19% aerosols. The cloud vs aerosol discrepancy can result from the simple cloud/aerosol threshold used in STRAT as well as possible aerosol detection by the 855nm ceilometer. Next we compare CBH retrieved by both systems when they agree that a cloud is present in the 1300-5000m range. Figure 11b shows a scatter plot of LNA CBH versus ceilometer CBH and Figure 11c shows the PDF of the difference between the two retrievals, based on 12 months of observation. The VAISALA CBH detection method (Vaisala proprietary algorithm) is based on the detection of high backscatter in the profile so the retrieved CBH is frequently placed between the base of the cloud and the altitude of maximum backscatter in the cloud. The position of the ceilometer CBH in the lidar backscatter is illustrated in Figure 12. The mean difference between LNA/STRAT and ceilometer CBH is -178 m which is consistent with the result of Figure 12. The standard deviation of the comparison is 265 m. The PDF can be divided into three zones: zone 1 contains 80% of detections and gather the situations with most consistent retrievals, zone 2 includes 5% of the situations for which ceilometer CBH are lower than corresponding STRAT retrievals,

and zone 3 contains 14% of the distribution for which STRAT retrievals are lower than the ceilometer CBH.

C. Boundary layer height detection, comparison with radiosounding retrievals

Figure 13 shows a comparison of the BLH estimated from radiosoundings (launched every day at 1200 UT 15 km from SIRTA) and BLH processed by the STRAT algorithm on LNA lidar data. The method used to extract BLH from soundings is a threshold method applied on the Richardson number $R_{ib}(z)$ (Menut et al 1999) calculated as

$$R_{ib}(z) = \frac{g(z-z_0) (\theta(z) - \theta(z_0))}{\theta(z) (u(z)^2 + v(z)^2)},$$

23

where θ is the potential temperature, g is the acceleration due to gravity, z is the height, z_0 is height of the surface and u and v are the zonal and meridian wind components. The BLH is estimated with a threshold value of 0.21 (Vogelezang 1996).

The lidar derived BLH is the median BLH extracted between t_0 and $t_0 + 5'$ (Menut et al. 1999) where t_0 is the radiosonde launch time.

We study 200 temporally and spatially collocated radiosonde and lidar profiles. 125 situations correspond to clear-sky events without clouds below 5000 m and without aerosol layers above the boundary layer in a 20-min window around the RS launch. The mean difference between lidar and RS derived BLH estimates is 99 m with a standard deviation σ of 452 m, hence the standard error in the mean is 62 m for a

95% confidence interval. We find that 83% of the population is within ± 500 m (close to 1σ).

We assume that points beyond ± 500 m are outliers and restrict the comparison to situations when the difference is in the interval $[-500 \text{ m}, +500 \text{ m}]$, the mean difference becomes 21 m, the standard deviation 200m, and the standard error is 30 m. When we further restrict the comparison to clear-sky situations, we find very similar statistics (see Table 4). The population of BLH differences contains two subgroups, one representing 83% of the situations where BLH retrievals agree within $20\text{m} \pm 30\text{m}$ (95% confidence) and the other (17% of the population) representing cases with very large discrepancies (between 500 and 1500 m). The inconsistency between the two retrieval methods can be due to lack of mixing or entrainment of aerosols, or poor collocation of radiosonde and lidar profiles. Table 4 shows that the presence of cloud does not introduce additional discrepancies.

5. Conclusion

The STRAT algorithm has been developed to analyze large data sets of lidar backscatter profile and retrieve the vertical structure of particle layers in the atmosphere. The algorithm is based on four successive detections carried out on individual profiles. The signal noise level is a key parameter in the algorithm as the detection thresholds at each step of the process are determined with respect to it. Hence the algorithm automatically adjusts to varying levels of signal noise. Molecular or (near) particle free layers are determined with a conservative approach to minimize false detections so that those layers can effectively be used for automated normalization processes. Identification of particle layers is done by using continuous wavelet transforms to identify discontinuities in the lidar profile and choosing those that effectively correspond to cloud or aerosol layer boundaries. We find good consistency between cloud-base heights retrieved by STRAT and those provided by a commercial ceilometer analysis. The height uncertainty inherent to the method is evaluated to be less than 3 times the vertical resolution (e.g. less than 45 m for the LNA lidar). Similarly the transition from the boundary layer to the free troposphere is analyzed with wavelet transforms. When compared to boundary layer heights retrieved from radiosondes, we find no significant bias in the STRAT retrievals, but the comparison reveals a large scatter due to the inconsistency between the aerosol-based and the thermodynamic-based BLH definition. Even though a few test cases have been carried out with the STRAT algorithm on 355-nm, 532-nm and 1064-nm

lidar systems, with both analogue and photon counting detection systems, the true portability of the STRAT algorithm to diverse large lidar data sets is still under study.

Acknowledgments

This work has been carried out with the financial support of the French National Space Agency CNES. The authors would like to acknowledge Stéphane Mallat for stimulating discussions on our application of the CWT method. We extend our acknowledgments to the reviewers for their comprehensive reviews and useful inputs.

References

- Boers, R., E. W., J. D. Spinhirne, and W. D. Hart, 1988: Lidar observations of the fine-scale variability of marine stratocumulus clouds. *J. Appl. Meteor.*, **27**, 797–810.
- Bösenberg, J., V. Matthias, A. Amodeo, V. Amoiridis, A. Ansmann, J. M. Baldasano, I. Balin, D. Balis, C. Böckmann, A. Boselli, G. Carlsson, A. Chaikovsky, G. Chourdakis, A. Comeron, F. De Tomasi, R. Eixmann, V. Freudenthaler, H. Giehl, I. Grigorov, A. Hagard, M. Iarlori, A. Kirsche, G. Kolarov, L. Komguem, S. Kreipl, W. Kumpf, G. Larcheveque, H. Linné, R. Matthey, I. Mattis, A. Mekler, I. Mironova, V. Mitev, L. Mona, D. Müller, S. Music, S. Nickovic, M. Pandolfi, A. Papayannis, G. Pappalardo, J. Pelon, C. Perez, R. M. Perrone, R. Persson, D. P. Resendes, V. Rizi, F. Rocadenbosch, J. A. Rodrigues, L. Sauvage, L. Schneidenbach, R. Schumacher, V. Shcherbakov, V. Simeonov, P. Sobolewski, N. Spinelli, I. Stachlewska, D. Stoyanov, T. Trickl, G. Tsaknakis, G. Vaughan, U. Wandinger, X. Wang, M. Wiegner, M. Zavrtnik, and C. Zerefos (2003): EARLINET: A European Aerosol Research Lidar Network to Establish an Aerosol Climatology. *MPI-Report*, **348**, ISSN 0937-1060, Max-Planck-Institut für Meteorologie, Hamburg, 192 pp.
- Brooks, I. M., 2003: Finding Boundary Layer Top: Application of a Wavelet Covariance Transform to Lidar Backscatter Profiles. *Journal of Atmospheric and Oceanic Technology*, **20**: 1092 – 1105.

- Cadet, B., V. Giraud, M. Haeffelin, P. Keckhut, A. Rechou and S. Baldy, 2005: Improved retrievals of cirrus cloud optical properties using a combination of lidar methods. *Applied Optics*, **44**, 1726-1734.
- Chazette, P.; J. Pelon, and G. Mégie, 2001: Determination by Spaceborne Backscatter Lidar of the Structural Parameters of Atmospheric Scattering Layers. *Applied Optics*, **40**, 3428-3440.
- Chiriaco, M., H. Chepfer, V. Noel, A. Delaval, M. Haeffelin, P. Dubuisson and P. Yang, 2004: Improving retrievals of cirrus cloud particle size coupling lidar and three-channel radiometric techniques, *Monthly Weather Review*, **32**, 1684-1700.
- Clothiaux, E. E., G. Mace, T. Ackerman, T. Kane, J., Spinhirne, V. Scott, 1998: An Automated Algorithm for Detection of Hydrometeor Returns in Micropulse Lidar Data. *Journal of Atmospheric and Oceanic Technology*, **15**: 1035-1042 CloudNet Project. URL: <http://www.met.rdg.ac.uk/radar/cloudnet/>
- Cohn, S. A., and W. M. Angevine, 2000: Boundary-layer height and entrainment zone thickness measured by lidars and wind pro-filing radars. *J. Appl. Meteor.*, **39**, 1233–1247.
- Comstock, J. M., T. P. Ackerman, and G. G. Mace, 2002: Ground-based lidar and radar remote sensing of tropical cirrus clouds at Nauru Island: Cloud statistics and radiative impacts, *J. Geophys. Res.*, *107(D23)*, 4714, doi: 10.1029/2002JD002203.
- Durieux, E. and L. Fiorani, 1998: Measurement of the lidar signal fluctuation with a shot-per-shot instrument, *App. Opt.* **37**, 7128–7131

- Flamant, C., J. Pelon, P. H. Flamant, and P. Durand, 1997: Lidar determination of the entrainment zone thickness at the top of the unstable marine atmospheric boundary layer. *Bound-Layer Meteor.*, **83**, 247–284.
- Haeffelin, M., L. Barthès, O. Bock, C. Boitel, S. Bony, D. Bouniol, H. Chepfer, M. Chiriaco, J. Cuesta, J. Delanoë, P. Drobinski, J-L. Dufresne, C. Flamant, M. Grall, A. Hodzic, F. Hourdin, F. Lapouge, Y. Lemaître, A. Mathieu, Y. Morille, C. Naud, V. Noël, B. O’Hirok, J. Pelon, C. Pietras, A. Protat, B. Romand, G. Scialom and R. Vautard, 2005: SIRTA, a ground-based atmospheric observatory for cloud and aerosol research. *Annales Geophysicae*, **23**, 253-275.
- Hodzic A., Chepfer H., Vautard R., Chazette P., Beekmann M., Bessagnet B., Chatenet B., Cuesta J., Drobinski P., Goloub P., Haeffelin M., Morille Y., 2004: Comparison of aerosol chemistry transport model simulations with lidar and Sun photometer observations at a site near Paris. *J. Geophys. Res.*, **109**, D23201, doi:10.1029/2004JD004735
- Hoff, R. M., and K. J. McCann, 2002: A Regional East Atmospheric Lidar Mesonet (REALM). *Eos Trans. Amer. Geophys. Union*, **83** (Fall Meeting Suppl. A22C-0147).
- Mallat, S. G. and W. L Hwang., 1992: Singularity Detection And Processing With Wavelets. *IEEE Transactions on Information Theory*, **38**: 617 – 643.
- Mathieu, A., J.-M. Piriou, M. Haeffelin, P. Drobinski, F. Vinit and F. Bouysse, in press 2006: Identification of error sources in planetary boundary layer cloud forecast using SIRTA observations. *Geophys. Res. Lett.*

- Melfi, S. H., J. D. Sphinhirne, S. H. Chou, and S. P. Palm, 1985: Lidar observations of the vertically organized convection in the planetary boundary layer over the ocean. *J. Climate Appl. Meteor.*, **24**, 806–821.
- Menut, L., C. Flamant, J. Pelon and P.H.Flamant, 1999: Urban boundary layer height determination from lidar measurements over the Paris area. *Applied Optics*, **38**, 945-954.
- Naud, N., Haeffelin, M., Muller, P., Morille, Y., Delaval, A., 2004: Assessment of MISR and MODIS cloud top heights through comparison with a back-scattering lidar at SIRTA. *Geophys. Res. Lett.*, **31**, L04114, doi:10.1029/2003GL018976.
- Vogelezang, D. H. P., and A. A. M. Holtslag, 1996: Evaluation and model impacts of alternative boundary-layer height formulations, *Boundary Layer Meteorol.*, **81**, 245-269.
- Pal, S. R., W. Steinbrecht, A. I. Carswell, 1992: Automated method for lidar determination of cloud base height and vertical extent. *Appl. Opt.*, **34**, 2388-2399.
- Platt C.M., 1979: Remote Sounding of High Clouds: I. Calculation of Visible and Infrared Optical Properties from Lidar and Radiometer Measurements. *Journal of Applied Meteorology*, **18**: 1130-1143
- Platt C.M., S.A. Young, A.I. Carswell, S.R. Pal, M.P. McCormick, D.M. Winker, M. DelGuasta, L. Stefanutti, W.L. Eberhard, M. Hardesty, P.H. Flamant, R. Valentin, B. Forgan, G.G. Gimmestad, H. Jäger, S.S. Khmelevtsov, I. Kolev, B. Kaprieolev, Da-ren Lu, K. Sassen, V.S. Shamanaev, O. Uchino, Y. Mizuno, U. Wandinger, C. Weitkamp, A. Ansmann and C. Wooldridge, 1994: The

- Experimental Cloud Lidar Pilot Study (ECLIPS) for Cloud-Radiation Research. *Bulletin of the American Meteorological Society*, **75**, No. 9, pp. 1635-1654.
- Sassen, K, M. Griffin, G. Dodd, 1989: Optical Scattering and Microphysical Properties of Subvisual Cirrus Clouds, and Climatic Implications. *Journal of Applied Meteorology*, **28**: 91-98
- Sassen, K., and J. R. Campbell, 2001. A midlatitude cirrus cloud climatology from the Facility for Atmospheric Remote Sensing: I. Macrophysical and synoptic properties. *J. Atmos. Sci.*, **58**, 481-496.
- Schneider, J., Balis, D., Böckmann, C., Bösenberg, J., Calpini, B., Chaikovsky, A.P., Comeron, A., Flamant, P., Freudenthaler, V., Hagard, A., Mattis, I., Mitev, V., Papayannis, A., Pappalardo, G., Pelon, J., Perrone, M.R., Resendes, D.P., Spinelli, N., Trickl, T., Vaughan, G., Visconti, G., 2000: A European Aerosol Research Lidar Network to Establish an Aerosol Climatology (EARLINET). *J. Aerosol Sci.*, **31**, S592-S593.
- Shimizu, A., N. Sugimoto, I. Matsui, K. Arai, I. Uno, T. Murayama, N. Kagawa, K. Aoki, A. Uchiyama, and A. Yamazaki, 2004: Continuous observations of Asian dust and other aerosols by polarization lidars in China and Japan during ACE-Asia. *J. Geophys. Res.*, **109**, D19S17, doi:10.1029/2002JD003253.
- Wang, Z., Sassen, K., 2001: Cloud Type and Macrophysical Property Retrieval Using Multiple Remote Sensors. *Journal of Applied Meteorology*, **40**, 1665-1682.
- Welton, E. J., J. R. Campbell, J. D. Spinhirne, and V. S. Scott, 2001: Global monitoring of clouds and aerosols using a network of micro-pulse lidar systems,

in Lidar Remote Sensing for Industry and Environmental Monitoring, U. N. Singh, T. Itabe, N. Sugimoto, (eds.). *Proc. SPIE*, **4153**, 151-158.

Winker D., J. Pelon and P. McCormick, 2003: The CALIPSO mission : Spaceborne Lidar for observations of aerosols and clouds. *Proc. SPIE*, **4893**, 1-11.

Figure captions:

FIGURE 1 STRAT ALGORITHM DIAGRAM.....38

FIGURE 2 PDF OF $V(R,T)/\sigma_0(T)$ VALUES ESTIMATED ON 1 PROFILE PER DAY FOR THE ENTIRE DATABASE AT TWO WAVELENGTH : 532NM (GREY SOLID LINE) AND 1064NM (GREY DOTTED LINE) AND PDF OF SIMULATED $V(R,T)/\sigma_0(T)$ VALUES IN MOLECULAR LAYERS (BLACK SOLID LINE) WITH $2N+1 = 21$ GATES = 315M.....39

FIGURE 3 SECOND DERIVATIVE OF A GAUSSIAN WAVELET CALLED "MEXICAN HAT" WAVELET.....40

FIGURE 4 (A) SIMULATED AND NOT NORMALIZED BACKSCATTERING POWER RECEIVED FOR AN IDEAL CLOUD OR AEROSOLS CASE IN FUNCTION OF ALTITUDE, (B) CORRESPONDING CWT COEFFICIENTS CALCULATED FOR DIFFERENT DILATION A (FINEST HIGH UP) AND DIFFERENT LOCATION OF WAVELET'S CENTER B WITH HIGHEST COEFFICIENTS IN WHITE AND LOWEST IN BLACK, (C) SKELETON (MAXIMA LINES) OF THE CWT MODULUS. (D,E,F) SAME AS (A,B,C) FOR A REAL BACKSCATTERING.....41

FIGURE 5 PDF OF $R(R,T)/\sigma(T)$ VALUES ESTIMATED ON 10 PROFILES PER DAY FOR ALL THE DATABASE AND PDF OF SIMULATED $R(R,T)/\sigma(T)$ VALUES IN MOLECULAR LAYERS (SOLID LINE).....42

FIGURE 6 (A) LIDAR LNA DATA 532 NM WFOV TELESCOPE ON MAY 26 2003, (B) PROFILE BY PROFILE DPR22 RATIO (17), (C) LAYER BY LAYER DPR2LAYER RATIO, (D) FLAG OBTAINED WITH PARTICLE LAYERS DISTINCTION, CLOUD LAYERS ARE IN RED, AEROSOL LAYERS IN ORANGE (18). CEILOMETER CTH DETECTIONS ARE REPRESENTED WITH BLACK POINTS.....43

FIGURE 7 PDF OF DPR2LAYER VALUES ESTIMATED ON THE ENTIRE DATABASE FOR ALL THE RANGE (SOLID LINE), RANGE BELOW 7.5 KM (DASHED LINE WITH SQUARES) AND RANGE ABOVE 7.5 KM (DASHED LINE WITH DIAMONDS).....44

FIGURE 8 FIRST DERIVATIVE OF A GAUSSIAN WAVELET.....45

FIGURE 9 (A), (C), (E) NOT NORMALIZED RANGE CORRECTED BACKSCATTERED SIGNAL, (B), (D), (F) CORRESPONDING CWT' COEFFICIENTS CALCULATED FOR DIFFERENT DILATION A (FINEST HIGH UP) AND DIFFERENT LOCATION OF WAVELET'S CENTER B WITH HIGHEST COEFFICIENTS IN WHITE AND LOWEST IN BLACK. (A), (B) A CLEAR CASE, (C), (D) A CLOUDY CASE WITH A CLOUD NEAR THE BL, (E), (F) A CLOUDY CASE WITH A CLOUD AT THE BL.....47

FIGURE 10 (A) SIMULATED BACKSCATTERED PROFILE WITH AN ADDED NOISE, WHERE S1 AND S2 ARE TWO SLOPES THAT ALLOW US TO DESCRIBE THIS PROFILE. (B) DIFFERENCE BETWEEN DETECTED CBH BY STRAT AND THE REAL CBH (C) DIFFERENCE BETWEEN DETECTED CTH BY STRAT AND THE REAL CTH.....48

FIGURE 11 (A) VERTICAL DISTRIBUTION OF CLOUDS SEEN BY THE LIDAR LNA (DARK LINE) AND THE CEILOMETER LD40 (GREY DASHED LINE) (B) SCATTER PLOT OF CBH DETECTED BY STRAT ALGORITHM ON 532 NM NFOV TELESCOPE DATA AND CBH DETERMINED WITH CEILOMETER DATA, (C) PDF OF CBHLIDAR - CBHCEILOMETER.....50

FIGURE 12 CBHLIDAR ESTIMATED ON THE RANGE CORRECTED BACKSCATTERED PROFILE (DARK SOLID LINE) AND THE CORRESPONDING CBHCEILOMETER51

FIGURE 13 (A) SCATTER PLOT OF BLH RETRIEVED BY STRAT ALGORITHM ON 532 NM NFOV TELESCOPE DATA AND BLH ESTIMATED BY RADIOSONDES, FOR ALL CASES WITH CROSS MARKERS AND FOR CLEAR CASES WITH RING MARKERS (B) PDF OF BLHLIDAR-BLHRADIOSONDES FOR ALL CASES.....52

Table captions:

TABLE 1 LIDAR LNA DESCRIPTION.....	53
TABLE 2 PARAMETERS OF STRAT ALGORITHM TO PROCESS LIDAR LNA DATA.....	54
TABLE 3 COMPARISON BETWEEN STRAT DETECTION AND CEILOMETER DETECTION.....	55
TABLE 4 COMPARISON BETWEEN STRAT BLH RETRIEVALS AND BLH ESTIMATED BY RADIOSONDES.....	56

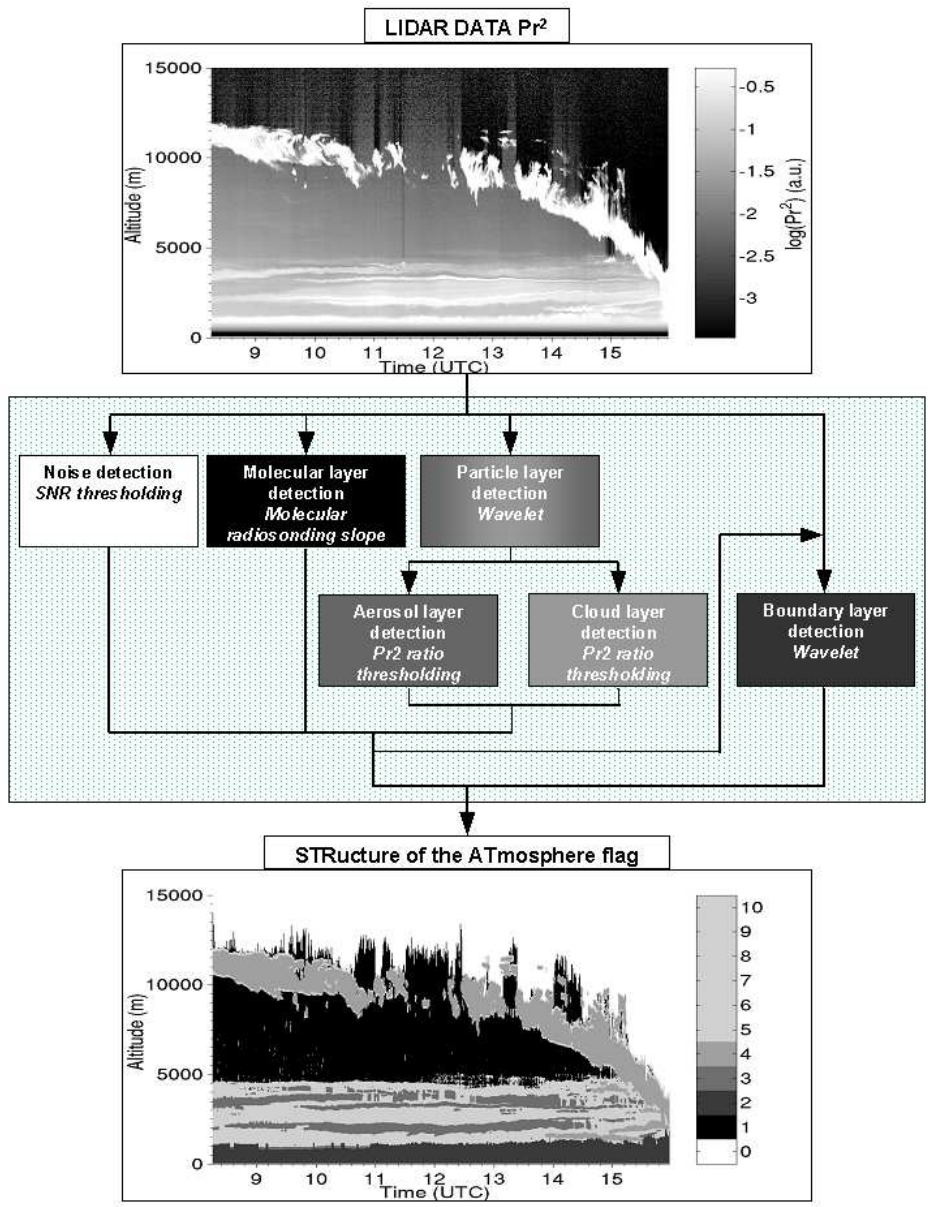


Figure 1 STRAT algorithm diagram.

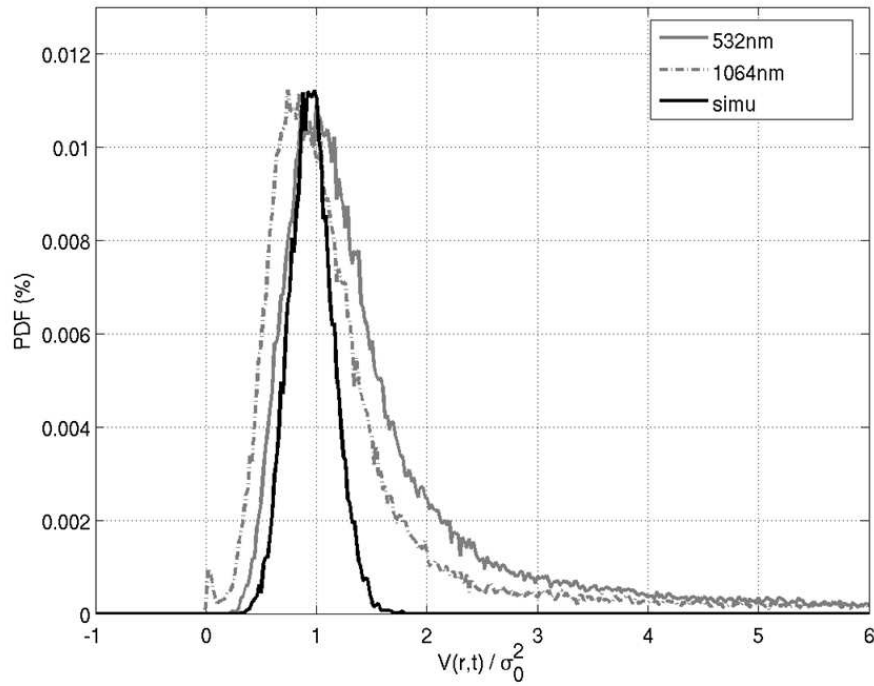


Figure 2 PDF of $V(r,t)/\sigma_0(t)$ values estimated on 1 profile per day for the entire database at two wavelength : 532nm (grey solid line) and 1064nm (grey dotted line) and PDF of simulated $V(r,t)/\sigma_0(t)$ values in molecular layers (black solid line) with $2N+1 = 21$ gates = 315m.

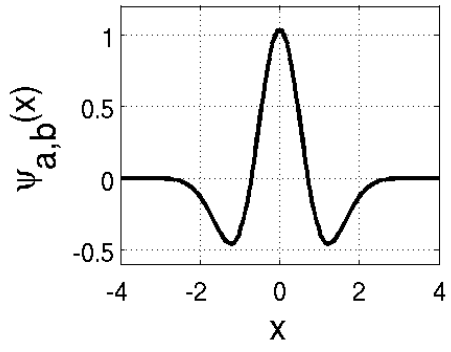


Figure 3 Second derivative of a Gaussian wavelet called "Mexican hat" wavelet

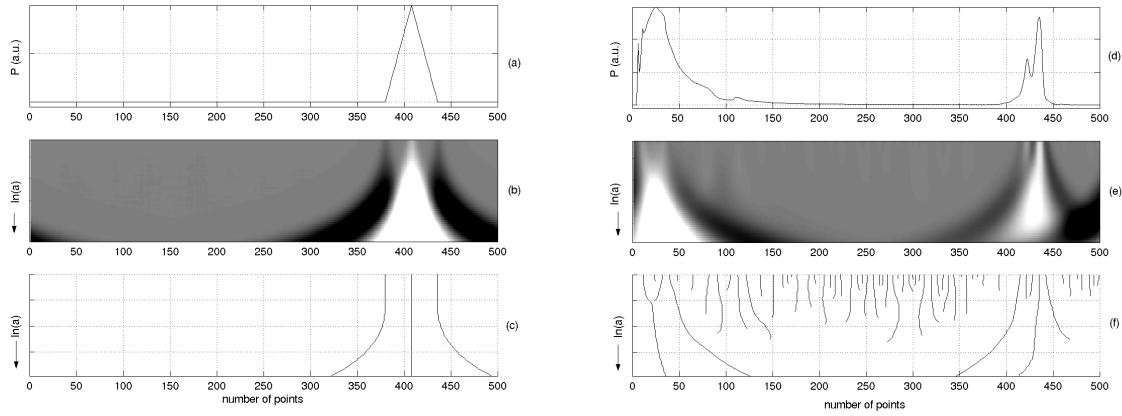


Figure 4 (a) simulated and not normalized backscattering power received for an ideal cloud or aerosols case in function of altitude, (b) corresponding CWT coefficients calculated for different dilation a (finest high up) and different location of wavelet's center b with highest coefficients in white and lowest in black, (c) Skeleton (Maxima lines) of the CWT modulus. (d,e,f) same as (a,b,c) for a real backscattering.

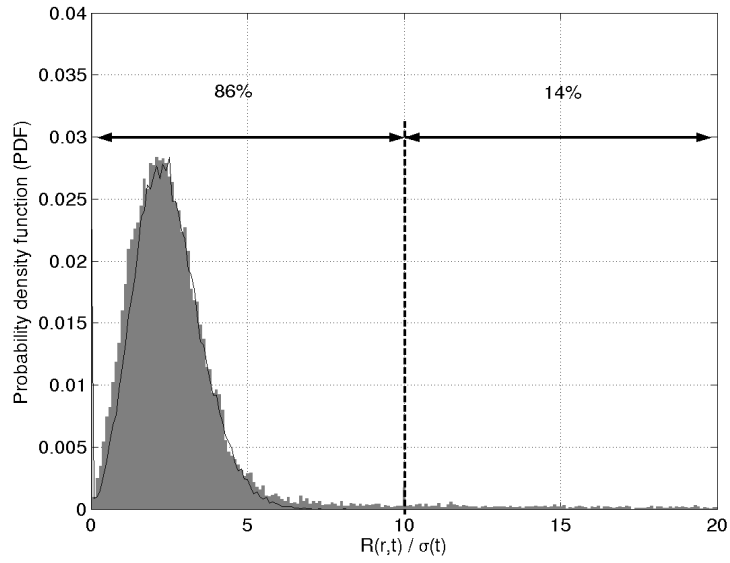


Figure 5 PDF of $R(r,t)/\sigma(t)$ values estimated on 10 profiles per day for all the database and PDF of simulated $R(r,t)/\sigma(t)$ values in molecular layers (solid line).

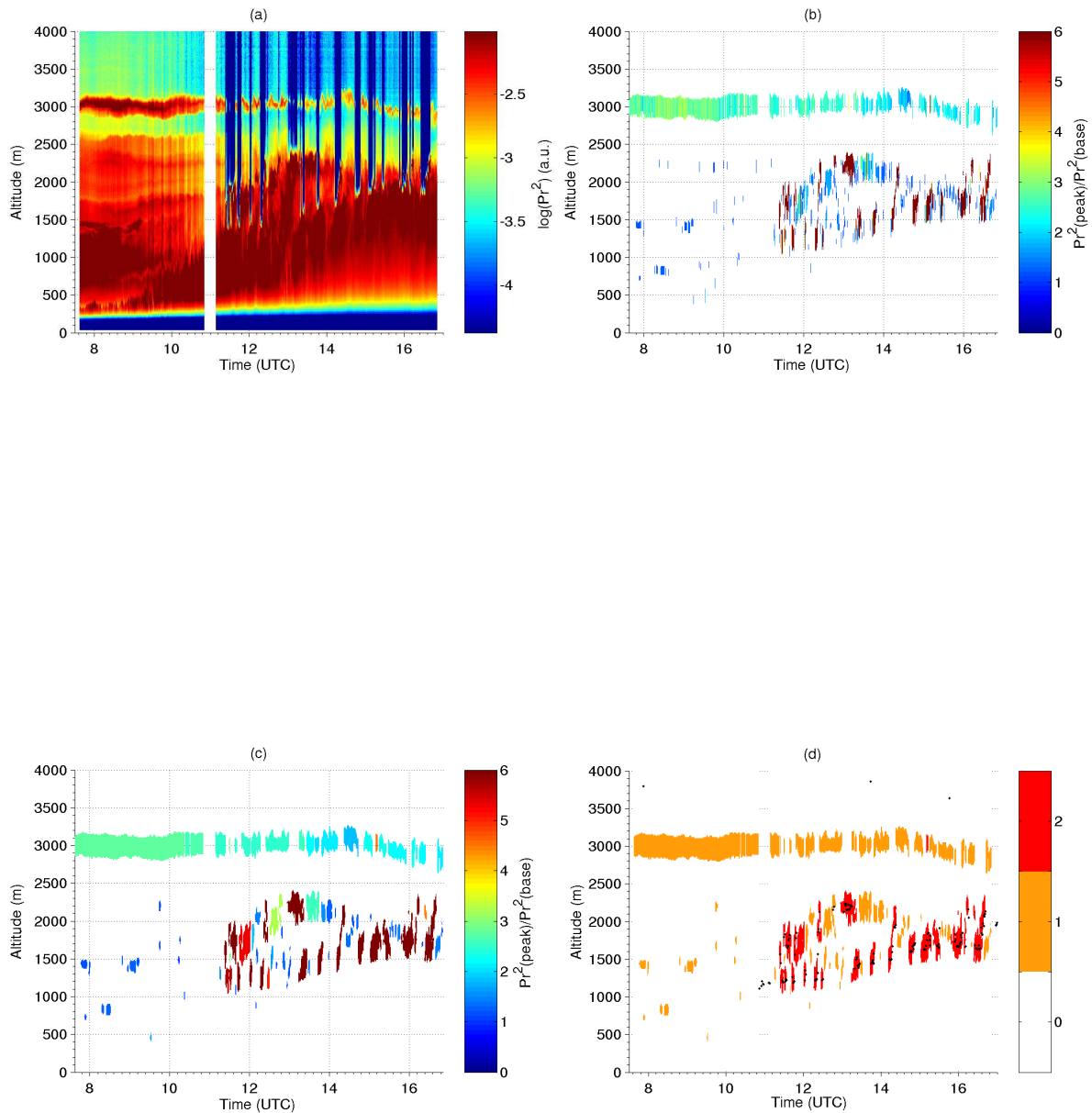


Figure 6 (a) lidar LNA data 532 nm WFOV telescope on May 26 2003, (b) profile by profile dP^2 ratio (17), (c) layer by layer dP^2_{layer} ratio, (d) flag obtained with particle layers distinction, cloud layers are in red, aerosol layers in orange (18). Ceilometer CTH detections are represented with black points.

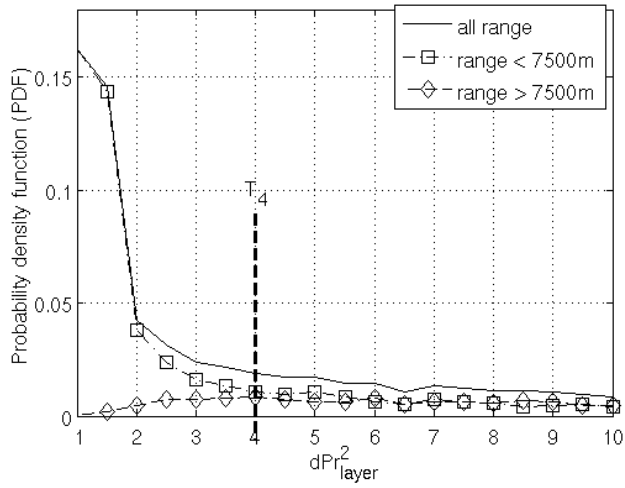


Figure 7 PDF of dPr^2_{layer} values estimated on the entire database for all the range (solid line), range below 7.5 km (dashed line with squares) and range above 7.5 km (dashed line with diamonds).

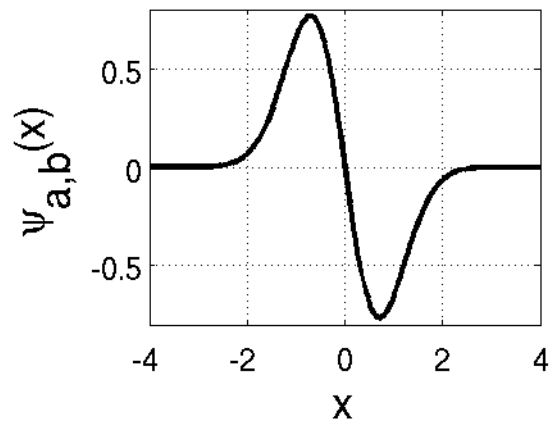
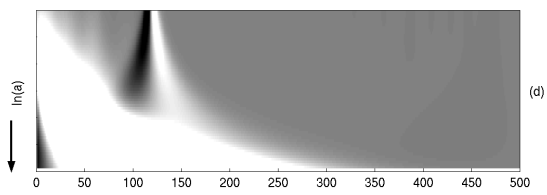
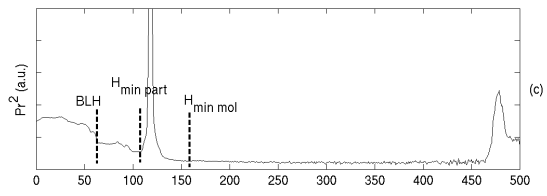
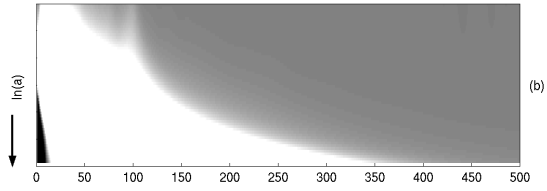
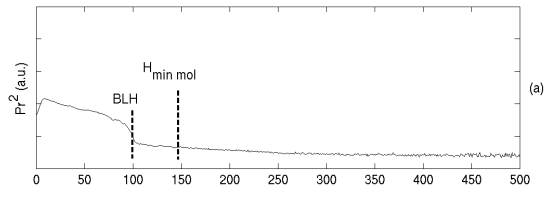


Figure 8 First derivative of a Gaussian wavelet.



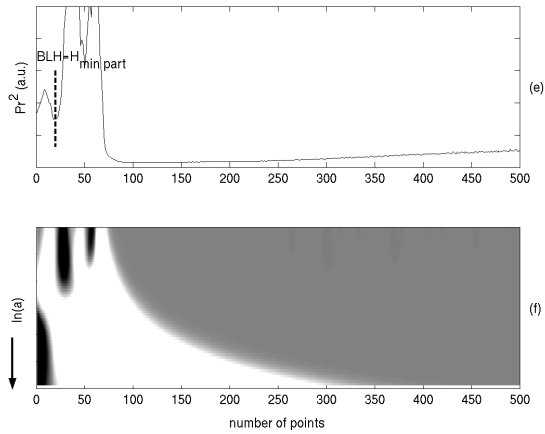


Figure 9 (a), (c), (e) not normalized range corrected backscattered signal, (b), (d), (f) corresponding CWT' coefficients calculated for different dilation a (finest high up) and different location of wavelet's center b with highest coefficients in white and lowest in black. (a), (b) a clear case, (c), (d) a cloudy case with a cloud near the BL, (e), (f) a cloudy case with a cloud at the BL.

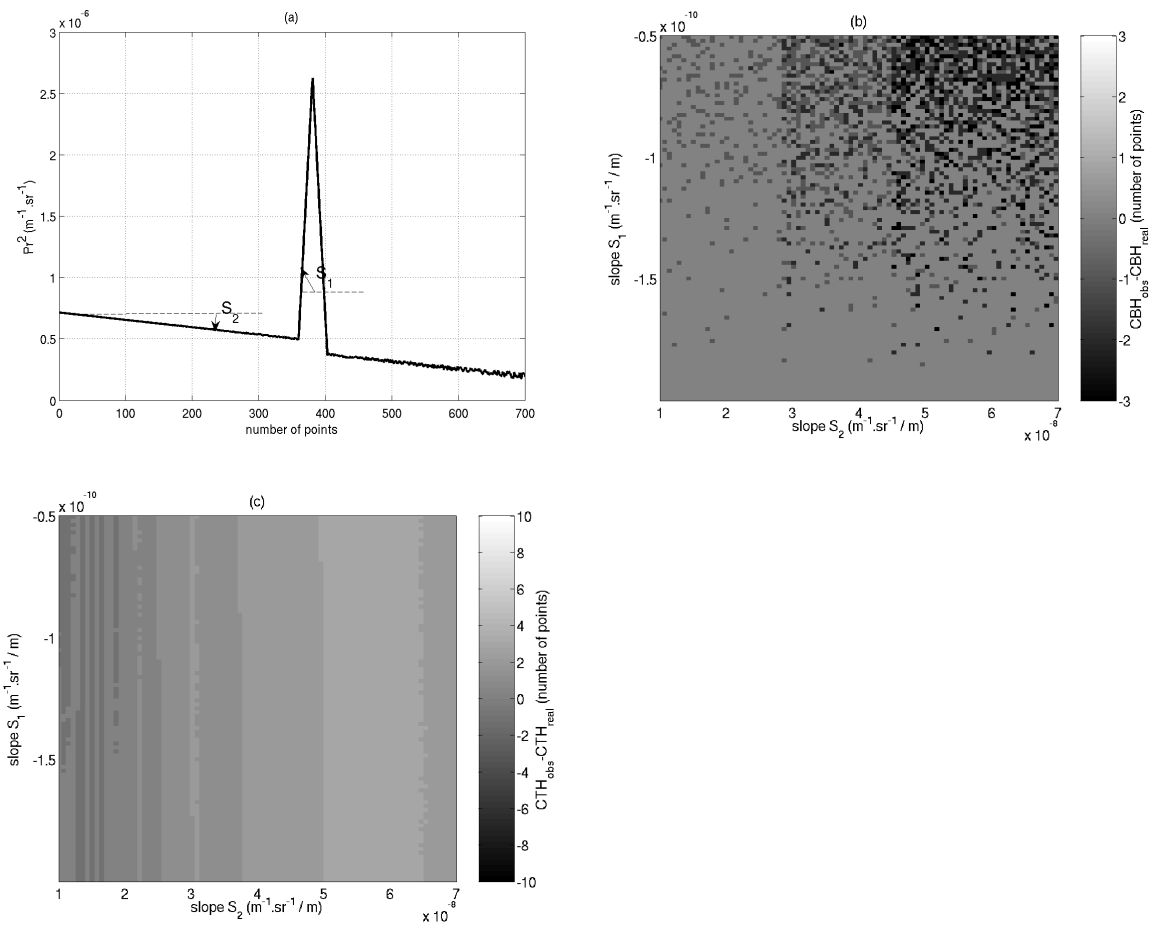
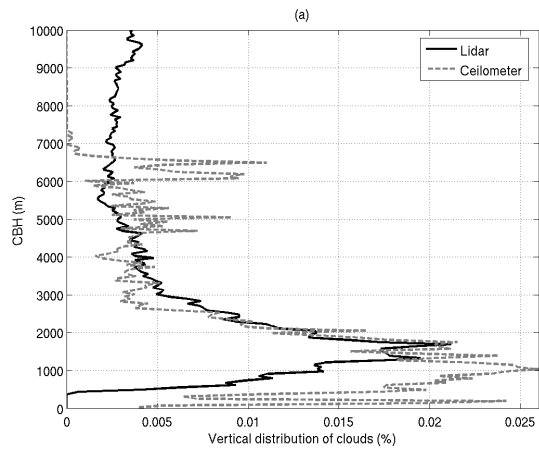


Figure 10 (a) Simulated backscattered profile with an added noise, where S_1 and S_2 are two slopes that allow us to describe this profile. (b) Difference between detected CBH by STRAT and the real CBH (c) Difference between detected CTH by STRAT and the real CTH



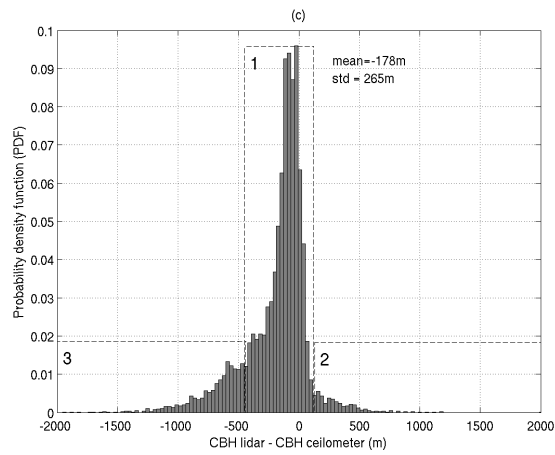
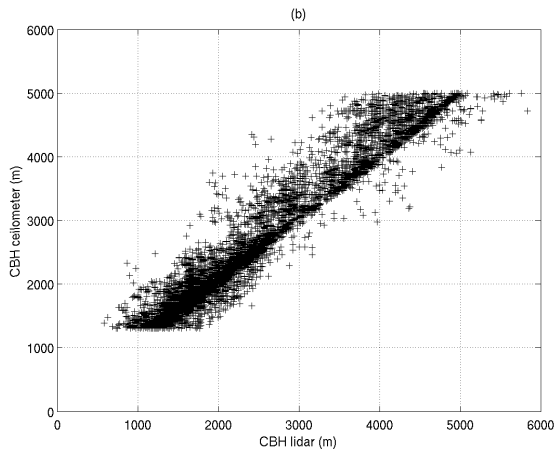


Figure 11 (a) Vertical distribution of clouds seen by the lidar LNA (dark line) and the ceilometer LD40 (grey dashed line) (b) Scatter plot of CBH detected by STRAT algorithm on 532 nm NFOV telescope data and CBH determined with ceilometer data, (c) PDF of $CBH_{\text{lidar}} - CBH_{\text{ceilometer}}$.

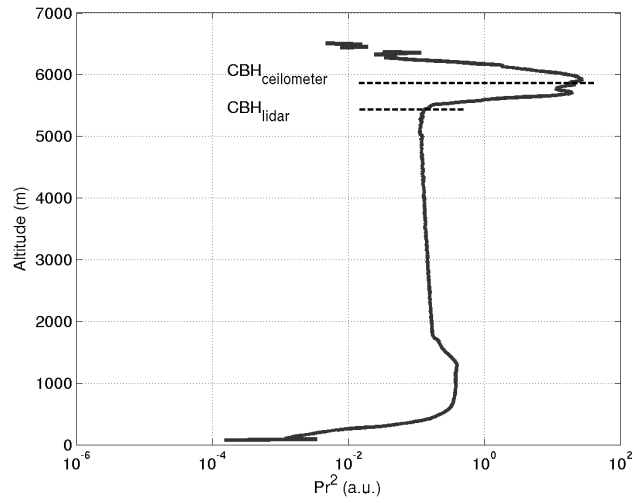


Figure 12 CBH_{lidar} estimated on the range corrected backscattered profile (dark solid line) and the corresponding $CBH_{\text{ceilometer}}$

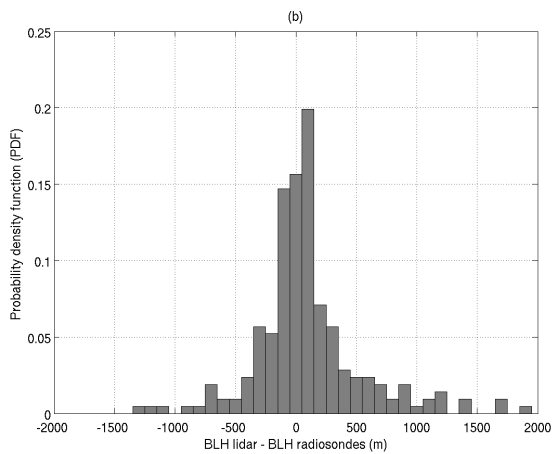
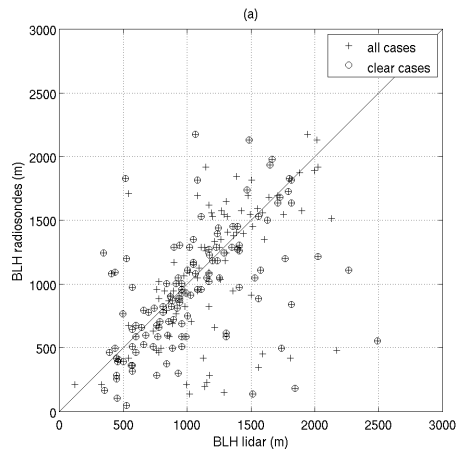


Figure 13 (a) scatter plot of BLH retrieved by STRAT algorithm on 532 nm NFOV telescope data and BLH estimated by radiosondes, for all cases with cross markers and for clear cases with ring markers (b) PDF of $BLH_{lidar} - BLH_{radiosondes}$ for all cases.

Table 1 Lidar LNA description

Laser type	Nd-Yag
Emitted wavelengths	532 and 1064 nm parallel polarized
Pulse energy	160 – 200 mJ
Repetition rate	20 Hz
Range resolution	15 m
Detected wavelengths	532 nm parallel polarized 532 nm cross polarized 1064 nm
Telescopes	Narrow field of view Ø = 60 cm 0.5 mrad Wide field of view Ø = 20 cm 5 mrad

Table 2 parameters of STRAT algorithm to process Lidar LNA data

	Noise detection	Molecular layer detection	Cloud and aerosol detection	Cloud and aerosol distinction	Boundary layer height detection
Threshold	$T_1=3$	$T_2=3$	$T_3=10$	$T_4=4$	/
Window length	5 gates or 75m	21 gates or 315m	/	/	/

Table 3 Comparison between STRAT detection and ceilometer detection

* cloud and aerosol free

LNA lidar Ceilometer	Cloud free*	Cloud	Aerosol
No detection	92	8	--
Cloud	7	74	19

Table 4 Comparison between STRAT BLH retrievals and BLH estimated by radiosondes

Situations	ALL	CLEAR-SKY	ALL (+/-500m)	CLEAR-SKY (+/- 500m)
Number of cases	211	125	173	105
Mean (BLH_{lidar} – BLH_{RS}) (m)	99	89	24	51
Standard Dev. (BLH_{lidar} – BLH_{RS}) (m)	452	445	202	199
Standard Error of the Mean (BLH_{lidar} – BLH_{RS}) (m) (95% conf. Int.)	62	80	30	38