CHAPTER 7
LIDAR INVERSION METHODS WITHOUT MEMORY

This chapter is aimed at inverting the sought-after parameters of extinction and backscatter from the lidar equation using algorithms without memory. This is to say, that every lidar return signal will be processed independently from past shots. Rather than tackling the problem in its most general form, where both optical parameters can have different values in different cells along the lidar beam-path, a bottom-up approach is followed. At the outset, the chapter begins by reviewing the simplest straightforward classic approaches based on linear regression (slope-method) and least-squares algorithms. At this level, an homogeneous atmosphere with identical cell features will be assumed. The main contributions here encompass the blend of all this background into efficient hybrid algorithms as well as their evaluation in terms of mean-square error in the optical parameters vs. key variables such as signal-to-noise ratio, visibility margin, etc. In addition, the development of mathematical models for the slope-method have enabled to identify inversion error sources, thus ending a long-lasting controversy between the two methods that has arisen in the literature.

Taking a step further, the void of lidar inversion in inhomogeneous atmospheres using non-memory algorithms is filled by Klett's method. Likewise, an error assessment study has been conducted that focus on how calibration errors (in addition to those mentioned above) translate into a final inversion error in the optical parameters. Working on these grounds, the inverted profiles are within known error bounds, as the inversion error can never be known exactly.

To sum up, and with a view to integrating all the worthy algorithms, a third software package named *link-detect* has been built. All the inversion plots of this investigation along with the live-data inverted profiles have been processed and evaluated with it.

1. LIDAR INVERSION GROUNDS

As it has been presented in *Chap.3*, the basic single-scattering LIDAR equation may formally be expressed as [9]

$$P(R) = \frac{A}{R^2}\beta(R)\exp\left[-2\int_0^R \alpha(r)\,dr\right] \qquad (1)$$

where:

$P(R)$ is the range-received power *[W]*,

$\beta(R)$ is the range-dependent volume backscatter-coefficient *[$km^{-1}\,sr^{-1}$]*,

α is the range-dependent extinction-coefficient *[km^{-1}]*,

R is the range *[km]* and,

A is the system constant [179],

which is defined as follows

$$A = 10^{-12}\frac{Ec}{2}A_r\xi(\lambda)\xi(R) \qquad [Wkm^3sr] \qquad (2)$$

where

E is the transmitted energy [mJ],
c is the speed of light [ms^{-1}],
A_r is the effective receiver area [m^2],
$\xi(\lambda)$ is the receiver's spectral transmission factor [] and,
$\xi(R)$ is the overlap factor or geometrical form factor [].

The goal of LIDAR inversion is to retrieve the optical atmospheric parameters $\alpha(R)$, $\beta(R)$ from the return signal $P(R)$. ([175][176][188])

All the algorithms discussed here lie on the basis of single scattering situations. This is always true except in instances of very low visibility, where multiple scattering methods have to be considered [168].

Next, this section primarily describes an *homogeneous method* to solve the lidar eq.(1) for α and β. In addition, it is of interest to know the feasibility and accuracy of optical-to-physical conversions. Usually, the received signal $P(R)$, is range-normalized and corrected for instrumentation transfer anomalies. System constants (e.g. receiver area) and pulse-to-pulse variations in lidar performance (e.g. output power) are eliminated by dividing the profile by its range-normalized value at R_o *(reference range)*. This leads to

$$S(R) = 10\log\frac{\beta(R)T^2(R)}{\beta(R_o)T^2(R_o)} \tag{3}$$

where

$$T(R) = \exp\left(-\int_0^R \alpha(r)\,dr\right) \tag{4}$$

is the one-way path transmittance. Expressed in differential form, this equation becomes

$$\frac{dS}{dR} = \frac{4.34}{\beta}\frac{d\beta}{dR} - 8.7\alpha \tag{5}$$

and from this, given:

a) an assumption on the relationship between α and β, and
b) a boundary value of an appropriate parameter,

one can derive the two optical parameters, or, given additional relationships, certain physical parameters [183][163]. [9] provides a general form from which solutions for various aerosol parameters such as *particle number, or mass concentration* may be derived. Adopting the letter Φ for the aerosol parameter in question, the solution is

$$\Phi(R) = \frac{\exp[C_1 S(R)]}{\frac{1}{\Phi^{-1}(R_o)} - C_2\int_{R_o}^{R}\exp[C_1 S(r)]\,dr} \tag{6}$$

Specific examples of Φ with appropriate values for C_1 and C_2, and the assumed relationships for the aerosol quantities are listed in Tab.1. Except for the special circumstances noted in the bottom rows of it, *propagation errors* are a major penalty. This means that to guarantee reasonable error levels in the estimated physical quantities, one would have to ensure even much smaller ones in the optical parameters first inverted, which would be unattainable for Mie lidars under typical homogeneous atmospheres.

SOLUTION FOR (Φ=)	BASIC RELATIONSHIP MEASURED OR ASSUMED	C_1	C_2
β (backscatter-coef)	$\beta = k_1 \alpha$	0.23	$2/k_1$
α (extinction-coef)	$\frac{d\ln\beta}{d\ln\alpha} = k_2$	0.23	$2/k_2$
N_p (*) (number concentration)	relative size distribution, $n(r)$ invariant with range R	0.23	$2\int_0^\infty \pi r^2 Q_{ext}(r) n'(r)\, dr$
ρ_m (mass concentration)	α/ρ_m and β/ρ_m invariant with range R	0.23	$2\alpha/\rho_m$
(*) Refer to Chap.2, Sect.3.2.1. Note also that $n(r) = N_p n'(r)$.			

Tab.1 Lidar equation solution possibilities for some physical parameters [9].

Ultimately, a major result would be to infer physical variables such as temperature, pressure and humidity from optical measurements. As it has been discussed in Chap.2 such physical magnitudes have repercussions on the refractive index and on the concentration of scatterers, let alone humidity growing effects on them. These phenomena are particularly difficult to pinpoint in the Mie's case because not only do the efficiencies $Q_{ext}(x,m)$, $Q_{back}(x,m)$ directly depend quite complicatedly on the optical parameters (Chap.2, Sect.3.2) but also because there is no one-to-one relation between the efficiencies and the couple (x,m). This is equivalent to say that *scatterers of different sizes and refractive indexes may yield the same efficiency.*

The Rayleigh case, being the only one attainable in this way, has brought into the market some LIRAs (lidar Rayleigh) that are capable of remotely sensing the high atmosphere by monitoring the concentration of some permanent gas [113]. Note that in the lower atmosphere, Mie scattering will outweigh Rayleigh's one, since the former is some orders of magnitude higher. Compare, for instance, Rayleigh values in Fig.15 and Fig.16 (Chap.2) to the Mie's ones in Fig.30 (Chap.2) or to the haze M model in Fig.31 (Chap.2).

To shed more light on the issue, Collis and Russell [9] derived typical variations in the interrelation of optical and physical parameters of atmospheric aerosol particles.

There were considered characteristics such as particle number density, N, mass density, M, type of distribution function, shape, refractive index, etc. Thus, for instance, a change in the refractive index from $m=1.5\text{-}j0$ to $m=1.33\text{-}j0$ results in a $x0.3$ change in β/N. Changes in shape are also very important. In this way, a change in shape from homogeneous sphere to an *onion shaped artifact* leads to a $x0.7$ change in the backscatter-to-extinction ratio, β/α, or even up to a $x0.1$ change, if irregular plates are considered. This does not take into account errors due to misrepresentation of the distribution function of the scatterers [193].

In spite of the complexity, large propagation errors and a number of unknowns, it is possible to derive approximate or qualitative information from lidar measurements, and since this information is obtained remotely, it is often of great value [204]. Thus, estimation of coarse aerosol statistics of relevance to weather forecast and ceilometry studies such as cloud and boundary layer height determination are often possible from backscatter maps and synchronism information [179] from the hardware. Wind vector determination is also possible as a result of template window correlation on two-dimensional backscatter maps [170][171][191]. Extra processing of this information can contribute to wind-shear early warning systems, which are precedent in aviation [133].

Inversion algorithms are classified in this PhD thesis into *memory and non-memory algorithms*. The words non-memory algorithms refer to case where each return lidar signal is processed separately from others belonging to different lidar shots. This chapter will thoroughly concentrate on them.

They have both pros and cons. On the one hand, non-memory algorithms are fast but if processing speed is not an end (e.g. off-line processing) they suffer from the basic fault that they do not take advantage from the correlation amongst lidar shots. This results in an additional error source and in a lack of adaptability. On the other hand, as it may well be guessed, memory algorithms can counteract these penalties at the expense of large CPU times.

In addition, the different possibilities of the *algorithms have also been classified depending on whether the atmosphere is assumed homogeneous or not*. If it considered to be constant over the entire lidar range, which is clearly an unrealistic assumption, one talks of an homogeneous atmosphere. To simplify, extinction and backscatter-coefficients will be extrapolated from the visibility margin, V_M and the tabulated margins in [9][25], as it has been described in *Chap.3,* when talking about the Koshmieder's relation. Thus,

$$(\alpha,\beta) = f(V_M)\ ; \qquad \alpha(R)=\alpha, \quad \beta(R)=\beta \tag{7}$$

Homogeneity is the case of slope-method and exponential-curve-fitting algorithms. On the contrary, if at least one of the two optical parameters, α or β, is range-dependent, one talks of inhomogeneous atmospheres. Klett's method tackles the inversion problem in such instances.

2. ALGORITHM RANKING POINTS. SOME CRITERIA

When testing and grading different algorithms it is important to assess what are their sensitivity to noise by means of some SNR criterion, as each one performs a different translation of noise into α- and β-rms error. This idea has prompted a plan to develop testing *software (link-detect), that accounts for the main noise sources described in Chap.3.* All these sources have been modelled according to the parameters of the APD-based receiver. The following points must be observed:

As far as *SNR* versus range is concerned, it is in agreement with the definitions given in Sect.2.2 of Chap.3. This is to say that

$$SNR(R) = \frac{R_v LP(R)}{\sigma_{eq}\sqrt{B}} = \frac{R_v LP(R)}{\sqrt{\sigma_{sh,s}^2 B + \sigma_{sh,d}^2 B + \sigma_{th}^2 B}} \qquad \frac{[V]}{[V]} \qquad (8)$$

The text will be addressed in terms of linear units rather than in decibels *(dB)*. This avoids the conflict of using $20 \cdot log(SNR)$ at the receiver output (SNR appears as a ratio of voltages) or $10 \cdot log(SNR)$ at the optical input of the system (where it appears as the ratio of two optical powers).

As it was done with the *stochastic hardware simulator* of Chap.3, noise modelling lies in finding the equivalent gaussian noise variance $\sigma_{eq}^2(R)$ that merges into a single body singlan-induced shot, dark shot and thermal variances ($\sigma_{sh,s}^2$, $\sigma_{sh,d}^2$ and σ_{th}^2, respectively). Noise has also been filtered and limited to simulate a *9.2-MHz* bandwidth, ± 5-*V* receiver. Tab.2 summarizes the system default values used in the simulations. Note that sometimes the values shown do not perfectly match those of Chap.3. This is because Chap.3 tables advanced measurements from subsequent chapters and Tab.2 is based on *pessimistic* link-budget estimations, which were made before the construction of the lidar receiver. All considered, the simulations fulfil their purpose if they are interpreted in terms of SNR. A moderate *background power figure,* $P_{back} = 2 \cdot 10\text{-}7\ [W \cdot cm^{-2} \cdot nm^{-1} \cdot sr^{-1}]$, has been used.

PARAMETER		VALUE	REMARKS
Quantum efficiency	η	36 %	Tab.2 Chap.3; Sect.2.2 Chap.5
Current responsivity	R_i	37.1 A/W	Tab.2 Chap.3; Sect.4 Chap.5
Multiplication gain	M	120	Tab.2 Chap.3
Excess noise factor	F	25.6	Tab.6 Chap.5
Dark surface current	I_{ds}	49 nA	Tab.2 Chap.3; Tab.4 Chap.5
Dark bulk current	I_{db}	8.4 pA	Tab.2 Chap.3; Tab.4 Chap.5
Equivalent transimpedance gain	G	1 MΩ	Tab.3 Chap.3
Equivalent input noise density	σ_{th}	0.6 pA·Hz$^{-1/2}$	Tab.2 Chap.3
Equivalent noise bandwidth	B	50 MHz	Tab.2 Chap.3
Optical system losses	L	50 %	Tab.1 Chap.3

Tab.2 System default values used in link-detect

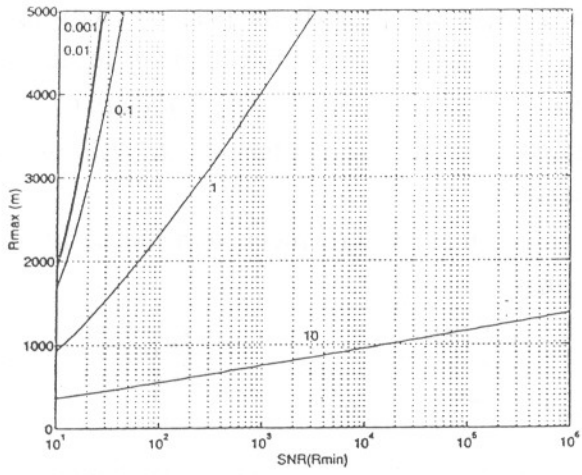

Fig.1 R_{max} LIDAR range for different extinctions in [km^{-1}].

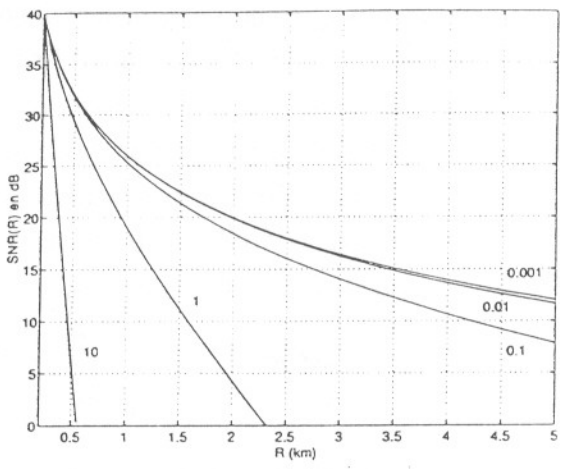

Fig.2 SNR(R) in dB for the extinctions shown.

Whenever the term SNR is addressed in the text, SNR values should be understood to be *SNR at* R_{min}, as it is a range-dependent function. If the atmosphere is assumed to be homogeneous, $SNR(R_{min})$ unequivocally defines the signal-to-noise ratio at every point R. Were the atmosphere to be inhomogeneous, the same holds true, provided the atmospheric test profile is known (see comments in Sect.4.2).

Minimum and maximum system ranges also take into consideration the *overlap factor (ovf)* assuming uniform ilumination. According to Fig.3 in Chap.3 and Ap.2 it can be said that ranges between 200 m and 5 km are well within the full overlap interval and that, therefore, overlap factor losses can be neglected in the test study.

These points settled, it is time to define some criteria regarding system ranges. As for the *minimum system range,* the only limiting factor, provided the receiver is far enough from saturation, is again the overlap factor. For this reason, according to Fig.3 (Chap.3), a value of *200 m* has been chosen for R_{min}.

On the other hand, the *maximum system range* has been defined as the minimum between these two ranges:

- *5 km*, which has been considered to be a sensible figure according to overlap factor considerations (Fig.3 in Chap.3),
- and the range R, where $SNR(R) = 1$.

$$R_{\max} \triangleq \begin{cases} SNR(R_{\max}) = 1 \\ 5\ Km \end{cases} \tag{9}$$

Fig.1 shows R_{max} criterion vs. *SNR* for different visibility margins, V_M.

A further point to consider is the *visibility margin,* V_M. Its definition is based in the Koschmieder's relation introduced in eq.(21) of Chap.3, Sect.1.3.

Tab.3 lists interpolated figures for the pairs α, β, given a wide set of visibility margins margins and according to different references. In Tab.3, the wavelength dependent behaviour is indicated for the extinction and backscatter in the columns $\alpha_{532(1064)}$ and $\beta_{532(1064)}$, respectively. Wavelength dependent behaviour has been computed for the extinction according to eq.(21) of Chap.3 and [25]. Since no mathematical extrapolation has been found for the backscatter-coefficient, it has been computed according to [9], where empirical intervals are given for both the extinction- and backscater-coefficient. The simulations have always used *1064-nm* values in ref.[9], which are indicated in Tab.3 as α_{sim} and β_{sim}. Both extinction and visibility will indistinctly parameterize the simulations.

VISIBILITY (V_M)	EXTINCTION		BACKSCATTER		ATMOSPHERIC CONDITION based on	
	α_{sim}	$\alpha_{532(1064)}$ *	β_{sim}	$\beta_{532(1064)}$ *	ref.[9]	ref.[25]
39.12 m	10^{2}	1(0.87)	5	5(5)	Dense water cloud	---
391.2 m	10	1(0.75)	$5\cdot10^{-1}$	5(5)	Light water cloud	Moderate fog
3.912 km	1	1(0.54)	$3\cdot10^{-2}$	3(3)	Moderate/light haze	Haze
39.12 km	10^{-1}	1(0.42)	$3\cdot10^{-3}$	4(3)	Clear air	Standard/very clear
391.2 km	10^{-2}	1(0.42)	$4\cdot10^{-4}$	10(4)	Rayleigh gaseous	Exceptionally clear
(*) This column compares the optical parameter at $\lambda=532$ *nm* to the one at $\lambda=1064$ *nm*. Values are times the order of magnitude given.						

Tab.3 Estimated optical parameters vs. visibility margins.

Above all, *rms relative error* in an inverted optical parameter, let us say, α, is computed in a root mean-square sense based on the following relation

$$\epsilon_{r,\bar{\alpha}} = \sqrt{\frac{1}{N}\left[\bar{1} - \frac{E(\bar{\alpha})}{\bar{\alpha}}\right]^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[1 - \frac{E(\hat{\alpha}_i)}{\alpha_i}\right]^2} \qquad (10)$$

where N is the vector dimension and $E()$ means the inverted value of its argument. Notice the vector notation in use: In instances where the algorithms work with inhomogeneous atmospheres, each component of the vector represents the value of the variable at every atmospheric cell *(7.5 m)*.

3. SLOPE-METHOD vs. LEAST-SQUARES

In this section two important algorithms for homogeneous atmospheres are presented along with a comparative study of their *rms-errors*. The error-simulations have considered the same sets of visilibity margins, V_M, and signal-to-noise ratios, *SNR*, as those given in Tab.3.

3.1 Slope-method

Mathematically speaking, the basis of this algorithm, first introduced by Kunz and Leeuw [187], lies in finding an operator that linearizes the lidar equation in such a way that the optical parameters, α and β, can easily be inverted. If that operator is taken to be

$$S(R) = \ln[R^2 P(R)] \tag{11}$$

where $P(R)$ is the lidar eq.(1), it takes the form of a straight line whose slope and y-cut provide direct information about the sought-after parameters. Next equation accounts for the noisy term contribution in $P(R)$:

$$S(R) = \ln(R^2[P(R)+n(R)]) = \ln(A\beta) - 2\alpha R + \ln\left(1+\frac{n(R)}{P(R)}\right) \tag{12}$$

Note that were it not for the noisy term containing $n(R)$, the first two terms would represent a line of such characteristics, wherefrom

$$\hat{I}(R) = mR+b; \qquad \hat{\alpha} = -\frac{m}{2}, \qquad \hat{\beta} = \frac{\exp(b)}{A} \tag{13}$$

The third term in eq.(12) is of special significance in view of the inversion error, since it results in a biased term in the form of the natural logarithm. This means that if the argument of the logarithm is in a zero-centred signal, $n(R)$, the result will be a non-zero mean signal whose offset will be able to change the parameters that should come out from the ideal line during the inversion process. The slope-method algorithm has consisted in obtaining the best fitting line, despite the offset problem, from a linear regression applied to $S(R)$. In the future, $S(R)$ will be called the *range-corrected function.*

Clearly, each operator performs a different translation of noise into a final inversion error in each parameter. As it might be guessed for the slope-method algorithm, the lower the signal-to-noise ratio, the higher the offset added. Therefore, recalling from Fig.2 and Chap.3 that the SNR is a range-dependent function, the offset added to each sample will increase as long as we move towards R_{max}.

As $S(R)$ fluctuates due to the noisy term, the inversion problem will have to consider the slope and y-cut expectations of the regression line, rather than their instantaneous values, which would be equivalent to consider a single realization. By applying the expectancy operator to both sides of eq.(12), it can be written that [23]

$$E[S(R)] = \ln(A\beta) - 2\alpha R + E\left\{\ln\left[1+\frac{n(R)}{P(R)}\right]\right\} = I(R) + bias(R) \tag{14}$$

To compute the expected bias, one can resort to the expectancy theorem [7][27]. Recalling the probability density function (in the future, *pdf*) of a gaussian noise

$$f_N(n) = \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{n^2}{2\sigma^2}\right) \tag{15}$$

the bias can be found as

$$bias(R) = \int_{-\infty}^{+\infty} \ln\left(1+\frac{n(R)}{P(R)}\right) f_N(n)dn \tag{16}$$

Surprisingly, the above integral diverges. The problem arises from the fact that $n(R)$ has some likelihood to span as far as $-P(R)$, where the logarithm argument reaches its zero value. This can be overcome including hardware constraints that limit the maximum noise excursion (e.g. saturation limits).

What is more, the development will only exist if

$$z = \frac{n(R)}{P(R)} \rightarrow |z| < 1 \tag{17}$$

This is equivalent to set $SNR > 1$. If one takes advantage of this condition, the McLaurin's series expansion of eq.(14) exists and it yields

$$E\left[\ln\left(1+\frac{n(R)}{P(R)}\right)\right] = \frac{E[n(R)]}{P(R)} - \frac{1}{2}\frac{E[n^2(R)]}{P^2(R)} + \frac{1}{3}\frac{E[n^3(R)]}{P^3(R)} + \ldots \tag{18}$$

Now, notice that the odd order moments vanish as $n(R)$ is an even function. To a certain extent, it can be assumed to be true provided receiver saturation effects do not represent a serious constraint to this affirmation. Obviously, noise can never exceed the sampling card dynamic range, but even though this is to say that its *pdf* is a gaussian one with truncated tails, it represents a minor effect (recall that within a 3σ-interval, there is a 99.7 % area enclosed). To make matters simpler, the analysis has been carried out using an uniform *pdf*. (In fact, the study has been done for other sorts of *pdf* including a Hamming's, without significant differences). The n-th order moments can be computed using the concept of characteristic function of a *pdf* and its properties [27]. If the symbol $\Pi(x/a)$ denotes an uniform *pdf* in variable x centred around zero with span a, its associated characteristic function $M_N(\omega)$, is given by

$$\frac{1}{2\sigma}\Pi\left(\frac{n}{2\sigma}\right) \Leftrightarrow M_N(\omega) = \frac{\sin\omega\sigma}{\omega\sigma} \tag{19}$$

Their n-th order moments can be computed by

$$m_i = E[n^i] = \frac{1}{j^i}\frac{d^i M_N(\omega)}{d\omega^i}\bigg|_{\omega=0} \tag{20}$$

$$m_k = \begin{cases} \dfrac{\sigma^{2n}}{2n+1} & k=2n \\ 0 & otherwise \end{cases} \tag{21}$$

Finally, substituting eq.(20) and eq.(21) into eq.(18), the following expression is found for the bias term

$$bias(R) = -\sum_{i=1}^{\infty} \frac{1}{2n(2n+1)} \frac{\sigma^{2n}}{P(R)^{2n}} = -\sum_{i=1}^{\infty} c_n \frac{1}{SNR(R)^2} \tag{22}$$

where the SNR definition has been used.

Summing up, as it had been advanced, the offset added to the ideal line *I(R)* given in eq.(14) increases for low SNRs and viceversa. A second conclusion would be that this algorithm has to deal with a range-dependent offset as SNR is also range-dependent.

The next algorithm presented is a classical one, but counteracts these problems.

3.2 The least-squares algorithm

This classical viewpoint of the inversion problem not only exhibits better performance but also seems more familiar to work with.

The problem can be formulated as minimization of the mean-square error between the range-corrected function (now, without the logarithm) and the lidar equation (eq.(1)). Mathematically,

$$\begin{cases} \min_{a,b} \| R^2 P(R) - b e^{-aR} \|^2 = \min_{a,b} \sum_{i=1}^{N} \left[R_i^2 \cdot P(R_i) - b e^{-aR_i} \right]^2; \\ \hat{\alpha} = \dfrac{a}{2} \quad \hat{\beta} = \dfrac{b}{A} \end{cases} \tag{23}$$

where the optical parameters, α and β can be inverted as indicated.

The exponential fitting has been programmed in MATLAB [1][2][3], as have been all the simulations presented here. Least-squares optimization begins with an initial search, which has shown to be quite sensititve to the initial guesses. So as to offset this effect, the final implementation of the algorithm uses initial guesses from slope-method to achieve the best results. The hybrid algorithm outweighs stand-alone application of slope-method by far.

The improvement can easily be understood if we keep an eye on the role noise plays in the least-squares algorithm. Contrary to what happened to slope-method, noise affecting *P(R)* in eq.(23) becomes multiplied by R^2, which results in an unbiased operation

$$E[R^2 n(R)] = R^2 \cdot E[n(R)] = 0 \tag{24}$$

To exemplify the unbalance between these two algorithms let us consider Fig.3 where an atmospheric situation with $V_M = 3.91\ km$ and $SNR = 1000$ is considered. Slope-method regression line, hybrid least-squares fitting line and solution line are shown. In the figure, it can be seen how least-squares is in good agreement with the ideal line. This backs the hypothesis that ***unbiased operators entail less inversion errors.***

As is often the case, *cooperative algorithms (least-squares plus slope-method initialization) exhibit better performance that their stand-alone counterparts.*

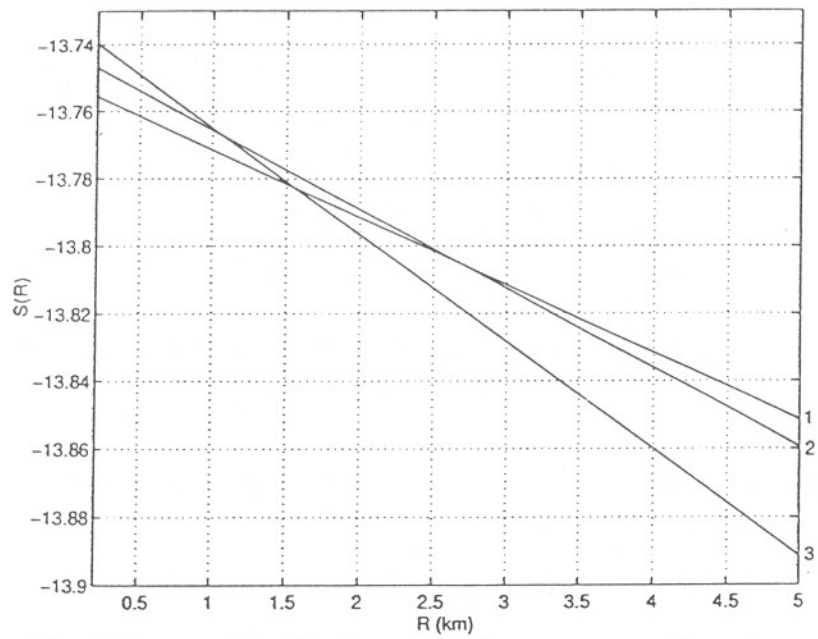

Fig.3 True profile (1) along with least-squares (2) and slope-method (3) inverted profiles.

3.3 Evaluation and sensitivity studies

3.3.1 Simulation procedure overview

This section is intended to assess homogeneous algorithms performance in terms of their mean-square errors as described by eq.(10).

To do so, testing software has been developed which is capable of synthesizing lidar return-signals based on the procedure of Fig.4: It starts by choosing an extinction-coefficient within the $[10^{-3}, 10]$ km^{-1} interval and a *SNR* within the $[10, 10^6]$ interval. For each pair of values, the program computes maximum lidar range (e.g. based on a unity SNR criterion (see *Chap.3*)), system constant, *A,* and adds the computed range-dependent filtered noise *(B=10 MHz)* to the lidar signal. Once multiple stochastic lidar signals, $P(R)+n(R)$, have been synthesized, they are fed into the inversion algorithms, from which α and β are obtained. α and β rms errors have been calculated averaging results over 10 simulations. Each set of simulations yields a single point $(SNR, RMS_\alpha, RMS_\beta)$.

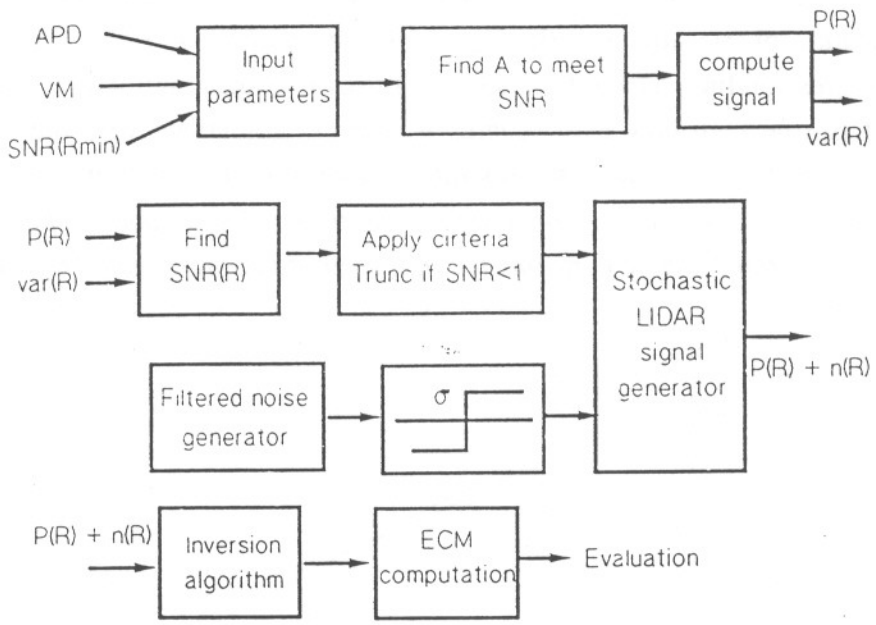

Fig.4 Link-detect block-diagram showing error evaluation procedure for homogeneous algorithms.

Range-dependent noise variance has been modelled very much like it was done with the stochastic hardware simulator discussed in *Chap.3*.

3.3.2 Noise-sensitivity study

During the evaluation, the extinction-coefficient has been allowed to logarithmically vary within the interval $[10^{-2}\text{-}10]\ km^{-1}$. Based on Tab.3, very low values of α represent unrealistically large V_M under the assumption of homogeneous atmosphere. This would illustrate asymptotic situations where extinction becomes negligible.

Fig.5 to Fig.8 illustrate rms errors for both α and β using both slope-method and least-squares. The latter is an hybrid form of least-squares since it has always been initialized with slope-method. In the figures, abscissae represent SNR whereas ordinates indicate rms error in the sought-after optical parameters. The change in slope coincides with those appearing in Fig.1, in agreement to the *criterion about* R_{max} stated in Sect.2. Note that for each atmospheric extinction, the interval wherein the maximum range is below *5 km* is ruled by the *5-km* criterion (the limiting one) rather than by a $SNR = 1$. Consequently, SNR is above unity outside this interval, which results in lower inversion errors.

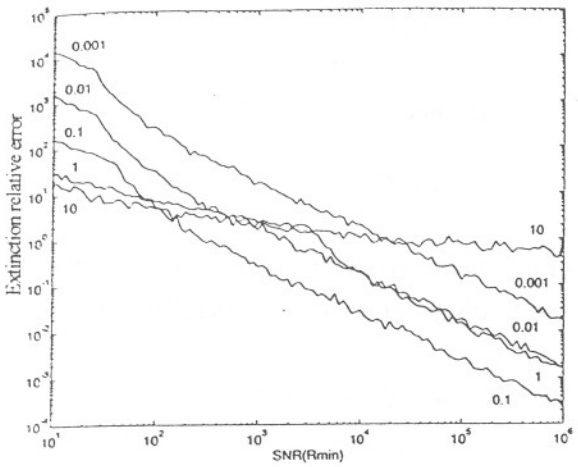

Fig.5 Extinction relative error with slope-method.

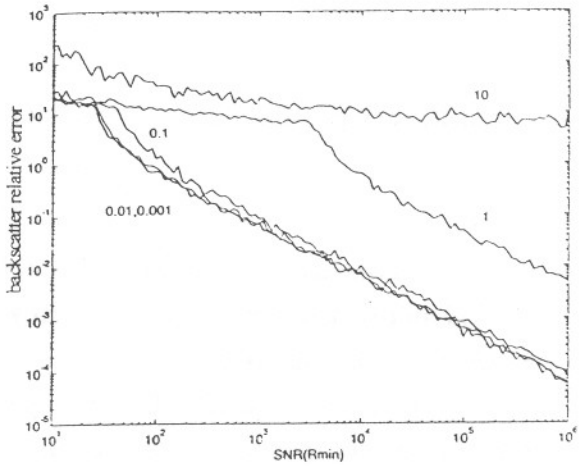

Fig.6 Backscatter relative error with slope-method.

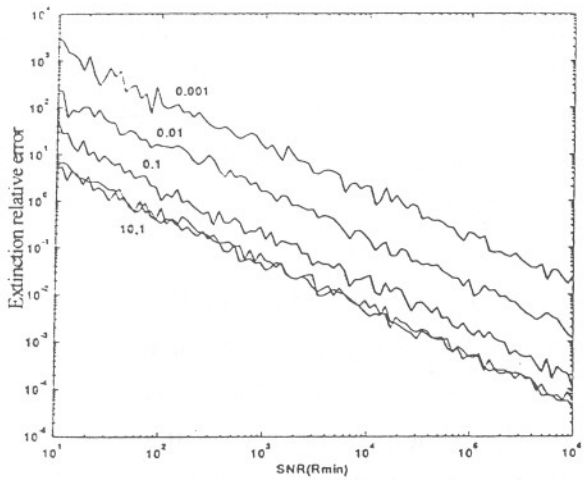

Fig.7 Extinction relative error with least-squares.

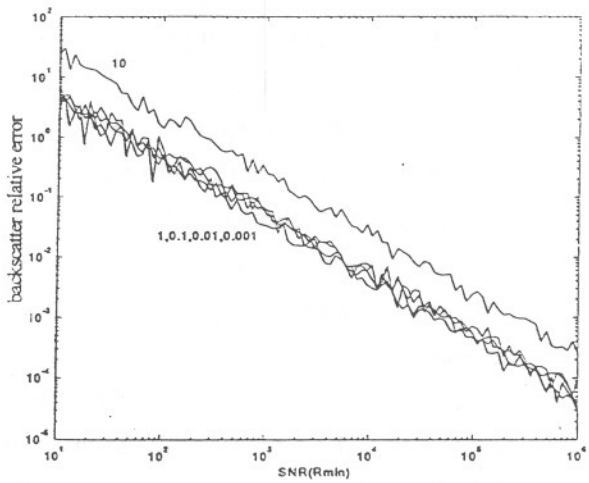

Fig.8 Backscatter relative error with least-squares.

NOTE: The plots above are parameterized by the extinction-coefficient [km^{-1}].

3.4 Hybrid methods: LANL live scene inversion

If one of the two optical parameters, α or β, is known under homogeneous atmosphere assumptions by any of the algorithms presented, an inhomogeneous version of the other can be inverted from the received power. Let the homogeneous extinction, α_h, be known, then eq.(1) reduces to

$$P(R) = \frac{A}{R^2}\beta(R)\exp(-2\alpha_h R) \tag{25}$$

From it, it is possible to retrieve a range-dependent backscatter-coefficient,

$$\hat{\beta}(R) = \frac{1}{A}R^2 P(R)\exp(2\alpha_h R) \tag{26}$$

In conclusion, so as to obtain an inhomogeneous backscatter profile, it is necessary to find an extinction estimate, α, by any of the homogeneous algorithms presented, *slope-method or least-squares.*

Obviously, the end all these algorithms pursue is to invert live data scenes. In this section, hybrid methods have been used to invert data from a portable lidar belonging to Los Alamos National Laboratory (LANL) (New Mexico, USA). LANL along with ITEMA carried out the *Barcelona Air Quality Initiative (BAQI) campaign* [166][170][171] during the Olympic Games held in Barcelona in 1992. (The acronym ITEMA stands for ***Institut de TEcnologia i Modelització Ambiental.*** Branching from the Polytechnic University of Catalonia, this institute is concerned with environmental modelling techniques). The aim then was to remotely sense wind speed vector field around the city, since it is of prime importance for the determination of pollution transport models.

LANL lidar, though emmitting less power than our Nd:YAG laser, has the advantage of portability and thus it is capable of scanning the atmosphere at different elevation and azimuth angles. The lidar is positioned by computer-controlled servo-engines. The scan introduced here comprises elevation angles between *-3°* and *40°* at *0.5°* intervals (the lidar was situated on top of a 284 m-height hill). This means a total of *87 laser-shots*. Under exceptional visibility conditions, the maximum system range is *7696.5 m*. Fig.9 shows the lidar-return signal versus range in arbitrary units for an elevation angle of *37°*. The analog-to-digital converter is also a *20-Msps card,* which means an spatial sampling rate of *7.5 m*. On the whole, there are 1200 power samples from which only 1027 are of interest. The 173 first samples prior to the lidar shot provide a measurement of the atmospheric *background noise*. Next, the whole inverting procedure would be described as follows:

The pre-processing steps of this data involves two steps:

1) *Estimation of the background noise.* This has been done by computing the mean of the *173* first samples and,

2) *Obtaining the useful data interval.* Once the background component has been subtracted from the raw data vector, not all of it contains useful data yet. Note in Fig.9 that the receiver saturation has clipped part of the peak of the return-signal. These data discarded, there will only be 993 useful samples left to be processed.

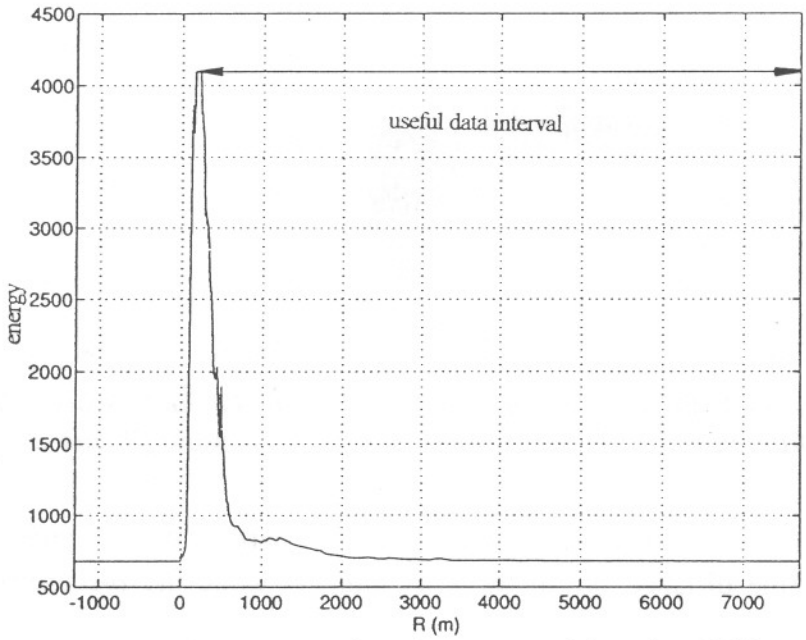

Fig.9 LANL-lidar-return signal vs. range (elevation 37°)

The processing steps are:

3) *Inversion of the inhomogeneous backscatter profile as described.* Fig.10 illustrates $A \cdot \beta(R)$*-profile* (in [W·km^2]) inverted with hybrid least-squares processing. First, note that the inverted backscatter-profile is scaled by the system constant, A, since its value was not supplied from LANL.

4) *Filtering of the inverted profile [28].* As it is noticed in Fig.10, range-dependent noise variance (eq.(24)) distorts backscatter information from the high atmosphere quite a lot. The filtering here has consisted in ideal low-pass filtering. Though there is room for improvement using bidimensional picture filtering techniques, unidimensional low-pass filtering of the backscatter spectrum has released challenging results using an optimum filter length of 460 samples (Fig.11).

5) *Graphical representation.* Rather than one by one filtering of the β*-profile,* as it has been done, image processing techniques [1] should be considered at this step. Fig.12 and Fig.13 show final results using hybrid *slope-method* and hybrid *least-squares,* respectively.

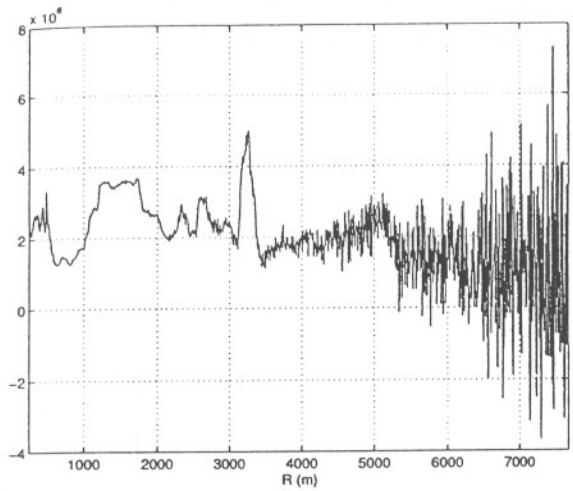

Fig.10 A·β(R) [W·km^2] *profile using non-filtered hybrid least-squares.*

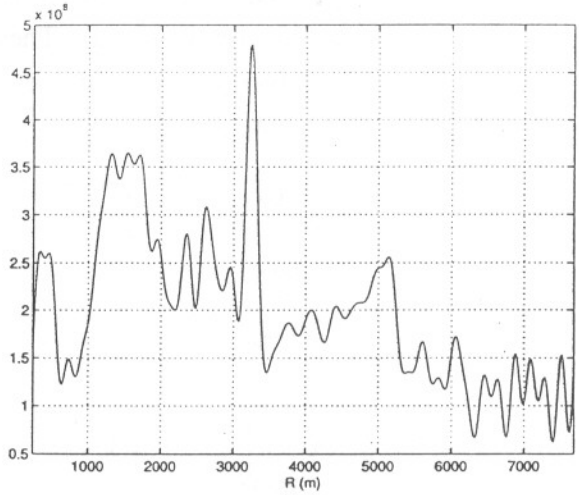

Fig.11 Final A·β(R) [W·km^2] *profile using filtered least-squares.*

If polar coordinates are considered to interpret these figures, the ρ-coordinate represents the distance, *R,* to the laser emitter, while the *θ*-coordinate represents the elevation angle.

The backscatter magnitude $A \cdot \beta(R)$ is shown in a colourmap fashion. Dark colours like black o red indicate high backscatter areas, while bright ones, like white or yellow denote low backscatter areas.

Least-squares inversion in Fig.13 retrieves basic characteristics of the atmosphere more realistically than in Fig.12, where slope-method inversion has been used. Thus, for instance, the boundary layer is shown by a red layer near 2300 m height in Fig.13. Underneath this layer, aerosols virtually have constant concentrations with height. This is shown by prevailing yellow tones above 2300 m. The red patch right below the layer may well correspond to a cloud extending between 800 m and 1300 m in height and 4800 m and 7500 m horizontally (recall that the red color indicates an increase in backscatter, as it is the case). The fact that aerosol concentration decreases with height above the boundary layer (Sect.4.3, Chap.2) is also clearly depicted in the figure. Note that above the boundary layer, colors fade progressively up to 5000 m height, indicating lower and lower concentrations.

On the other hand, the whole characteristics retrieved form Fig.13 are difficult to be inferred from Fig.12, except for the identification of the cloud. This corroborates Fig.13 for it cannot be attributed to picture saturation effects of the image. Note that in Fig.12 it is of yellow colour, which is a mid-scale colour of the colourmap.

Summing up, least-squares inversion outweighs slope-method one.

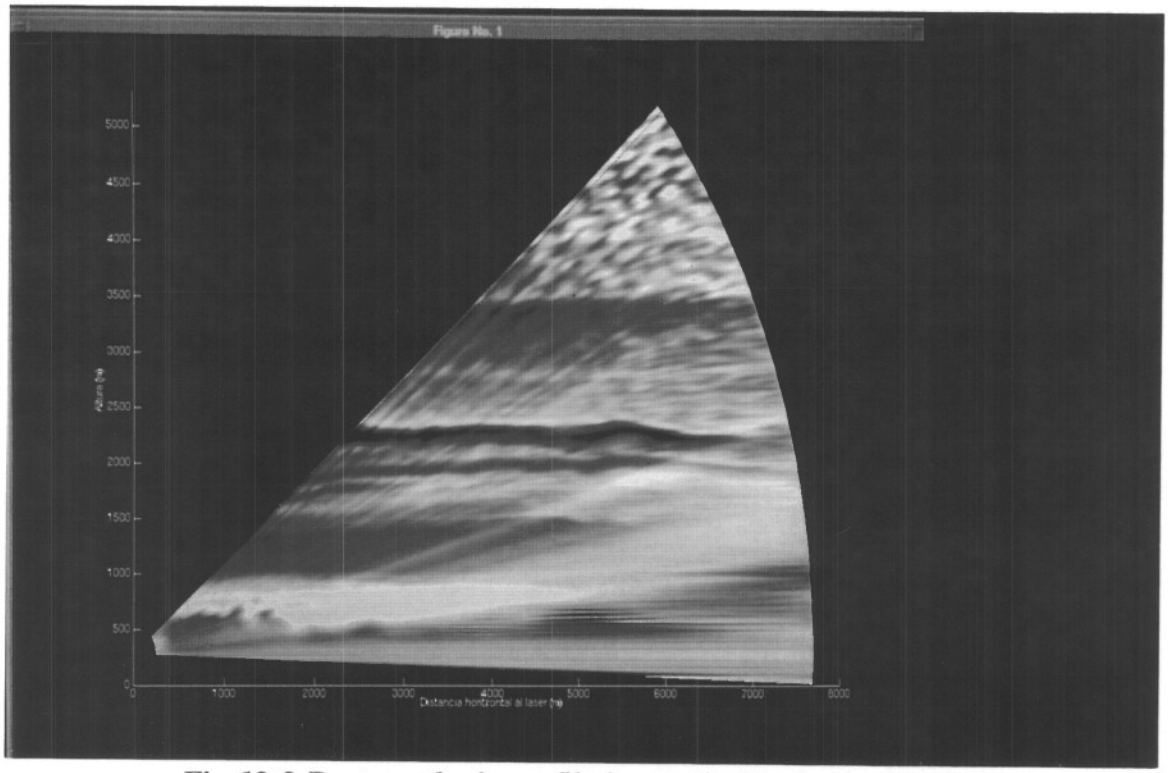

Fig.12 2-D atmospheric profile inverted with slope-method.

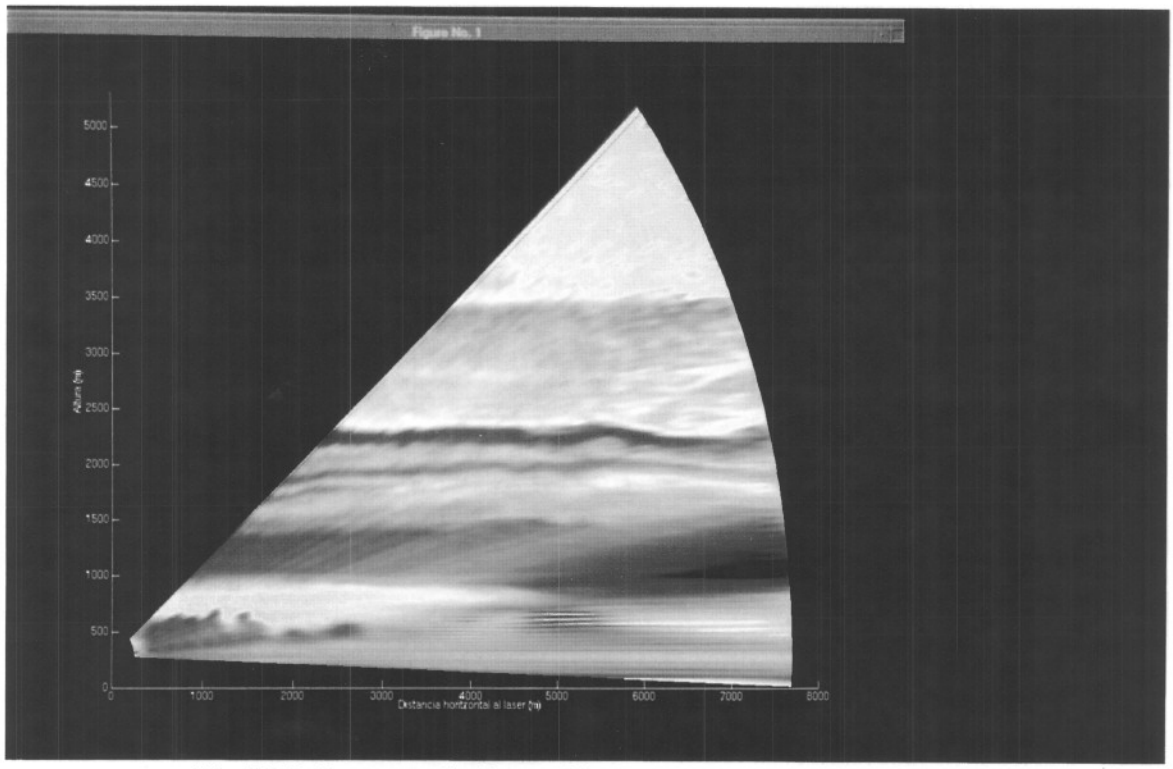

Fig.13 2-D atmospheric profile inverted with least-squares.

4. KLETT'S METHOD

4.1 The algorithm

Klett's method most prominent feature is that of solving the lidar inversion problem in inhomogeneous atmospheres, as is usually the case. Yet, some extra information is needed for the algorithm to pay off. Firstly, it is necessary to give a correlating hypothesis between the two optical parameters; these are no longer constants but vectors. Owing to this, they will be denoted $\alpha(R)$ and $\beta(R)$ or, equivalently, in vector form. The correlating hypothesis can be expressed as [185]

$$\beta(R) = C \cdot \alpha(R)^k \qquad 0.67 \le k \le 1.0 \tag{27}$$

where k is a constant within the margins given [186] and C can be estimated form the Koschmeider's relation and Tab.3.

Second, a calibrating measurement of the already defined range-corrected function must be given. It must be done at the maximum system range, R_{max}, rather than at R_{min}, otherwise numerical round-off errors would make the algorithm highly unstable:

$$\begin{cases} S(R) \triangleq \ln[R^2 \cdot P(R)] \\ S(R_{\max}) = S_m \leftrightarrow \alpha(R_{\max}) = \alpha_m \end{cases} \tag{28}$$

Since it may be especially difficult to calibrate at R_{max} [184], unless atmospheric probe balloons or other radar equipment worked in cooperation with the lidar [152] (see also airborne and shipboard sounding in [177][178][192]), algorithm sensitivity studies versus calibration errors are often needed.

Working on the basis of a differential equation formulation for the LIDAR relation, it can be written as

$$\frac{dS(R)}{dR} = \frac{k}{\alpha}\frac{d\alpha(R)}{dR} - 2\alpha(R) \tag{29}$$

Noting that it has the form of Bernouilli's or Ricatti's homogeneous differential equation, its general solution takes the form

$$\alpha^{-1}(R) = \exp\left(-\int_0^R \frac{1}{k}\frac{dS}{dr}dr\right)\cdot\left[D - 2\int_0^R \frac{1}{k}\exp\left(-\int_0^r \frac{1}{k}\frac{dS}{dr'}dr'\right)dr\right] \tag{30}$$

Now, considering that k is constant over the entire range (this does not represent a severe constraint regarding the small variation interval of the variable) and joining eqs.(25) and (26), it arises what is known as the *backward solution*

$$\alpha(R) = \frac{\exp[(S - S_m)/k]}{\alpha_m^{-1} + \frac{2}{k}\int_R^{R_m} \exp[(S - S_m)/k]\,dr} \tag{31}$$

Contrary to what happens to its minus sign counterpart (wherein integration from R_o to R is performed), this solution is highly stable. Its stability stems from the fact that as long as R decreases from R_m, the extinction is determined as the ratio of two numbers that each become progressively larger. If no external calibration, α_m, can be provided, α_m can be estimated by either of these two methods

$$\alpha_{homog} = -\frac{1}{2}\frac{dS(R)}{dR} \Rightarrow \alpha_m = \frac{1}{2}\frac{S_o - S_m}{R_m - R_o} \tag{32}$$

where an homogeneous atmosphere approximation has been considered for eq.(29), or

$$\alpha_m = \frac{\exp[(S_b - S_m)/k] - 1}{\frac{2}{k}\int_{R_b}^{R_m} \exp[(S - S_m)/k]dr} \tag{33}$$

where homogeneity is assumed within the (R_b, R_m) range by imposing $\alpha(R_b) = \alpha_m$ in eq.(31).

In order to interpret the calibration condition, let us consider Fig.14, where a synthesized atmospheric profile is depicted along with the inverted one. In the simulation there was assumed a signal-to-noise ratio, $SNR \to \infty$, $V_M = 10\ km$ and true knowledge of the correlating constant, k, (k=1). The inversion algorithm was miscalibrated with a wrong extinction ($\alpha_{cal} = 0.2531\ km^{-1}$) rather than with its true value ($\alpha_m = 0.5061\ km^{-1}$). Predictably, the inversion accommodates the calibration as a wrong boundary value. ***An important conclusion is that in spite of the miscalibration, the atmospheric profile is retrieved fairly well up to the approximate end of the inhomogeneity.***

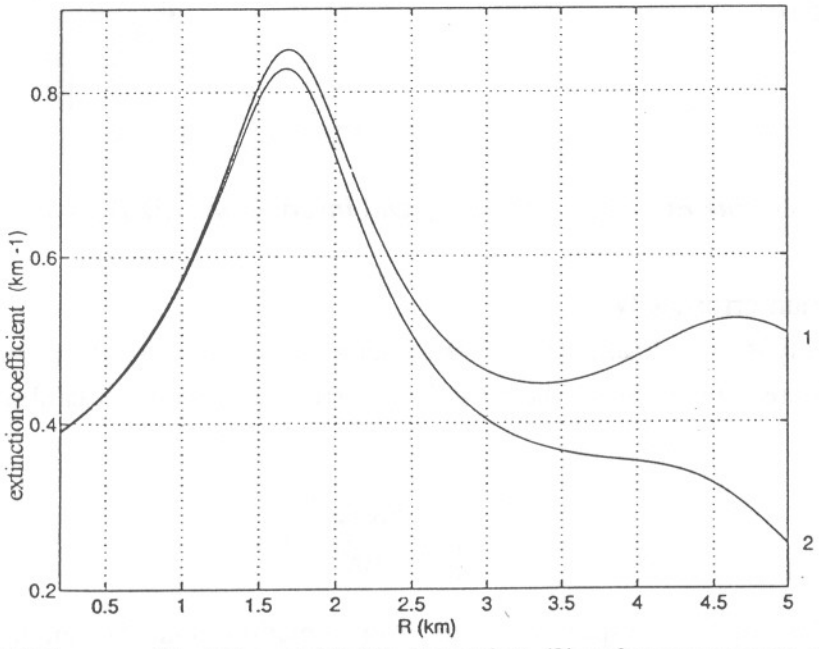

Fig.14 True profile (1) and Klett's inversion (2) using erroneous calibration.

4.2 Evaluation and sensitivity studies

This section is committed to doing sensitivity studies that quantify rms inversion errors in Klett's algorithm. Since the studies are parameterized, all the study parameters have taken their ideal default values, except for the one under evaluation. These are $\alpha_m=\alpha(R_{max})$, $k=1$, $SNR\rightarrow\infty$.

Trapezium-shaped atmospheric profiles have been synthesized in order to model inhomogeneous situations. They have been built in this way: bottom base with extinction α and top base with 2α. These profiles are very much like the ones used by Klett in [186]. *α-values* agree with V_M chosen from the set: *100 m, 1 km, 10 km.* The trapezium has been built from samples *7.5 m* apart (spatial sampling rate) up to a total of *641* samples, which enable to reach the maximum range of *5 km* (see criteria in Sect.2). The correlating constant, *k,* is assumed *unity*. Fig.15 depicts the synthesized profile for V_M = *10 km* along with the inverted one under exact values of calibration and *k,* given *SNR = 3000 (70 dB).*

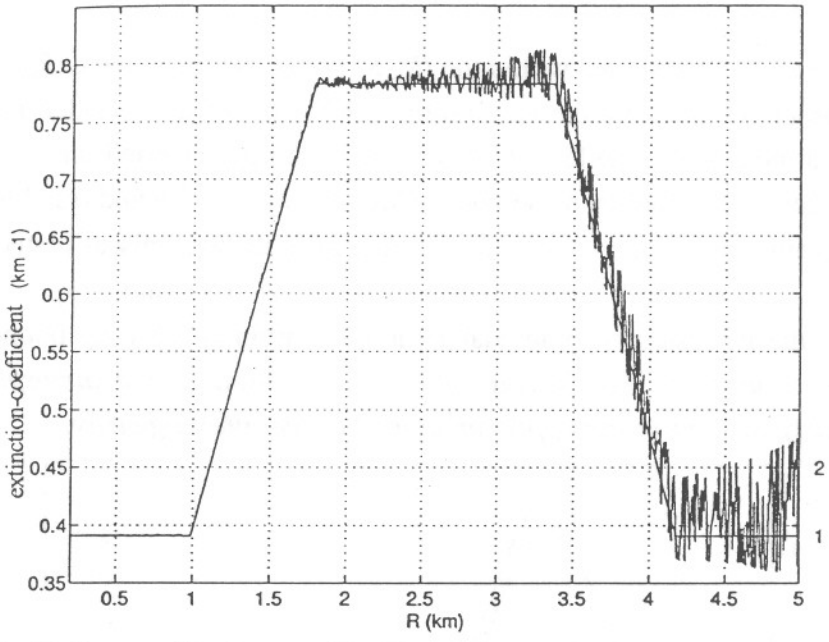

Fig.15 True extinction profile (1) and inverted one (2) (V_M=10 km).

4.2.1 Calibration error study

The study is twofold as both over and underestimation errors of the true calibration, α_m, are considered. Once given the percentage error, the estimated calibration, $\alpha_{m,e}$, is computed as

$$\alpha_{m,e} = \left[1 \pm \frac{e_r(\alpha_m)}{100}\right] \cdot \alpha_m \qquad (34)$$

where the plus sign corresponds to an overestimation and the minus sign to an underestimation.

Overestimation errors vary from 0 to 1000 % whereas underestimation errors range from 0 to 99 %.

Considering SNR infinite, maximum system range becomes *5 km* (Sect.2). Laser energy has been set to 1 mJ (note that this is irrelevant, the SNR being unlimited).

Fig.16 and Fig.17 depict rms inversion errors (refer to Sect.5 for discussion).

4.2.2 Correlation error study

Correlation errors between the two optical parameters, extinction and backscatter, arise from the value chosen for k in eq.(27). Simulation values for k are within the interval $(0.2, 1.5)$ which range from an 80%-overestimation to a 50%-underestimation. See Fig.18 to Fig.21.

4.2.3 Noise-sensitivity study

This study is very much like the one performed for the homogeneous atmosphere algorithms (Sect.3). To lessen the dependency on other parameters, exact calibrations (α_m, S_m) and correlation constant, k, have been input to Klett's algorithm according to the trapezium-sized extinction profile. This has left noise as the only error source.

Once the procedure starts, an initial SNR ($SNR \equiv SNR(R_{min})$) and visibility margin, V_M, are chosen. From this latter value, a trapezoidal atmospheric profile is created in the way explained in the introduction of this section. Maximum system range, R_{max}, is computed according to the diagram of Fig.4 and the criteria of Sect.2. Range-received power and range-dependent noise variance vectors are truncated to meet system ranges. Next, lidar signals can be synthesized by adding to the range-received power the computed noise realizations. Then, after a set of ten lidar inversions have been processed, relative rms-errors, $e_{r,\alpha}^{\,i}$, in the inverted optical parameters are available for each realization. Eventually, what is called *average extinction inversion error* (that takes into account the whole set of ten realizations) is computed. It is defined as

$$e_{r,\alpha} = \sqrt{\frac{1}{10}\sum_{i=1}^{10}\left(e_{r,\alpha}^{\,i}\right)^2}\cdot 100 \quad \% \tag{35}$$

Average extinction error is displayed in Fig.22. From the equation above and the correlation relation given in eq.(33), both average extinction and average backscatter errors can be related noting that

$$e_{r,\beta}^{\,i} = 1 - \left[1 - e_{r,\alpha}^{\,i}\right]^k \tag{36}$$

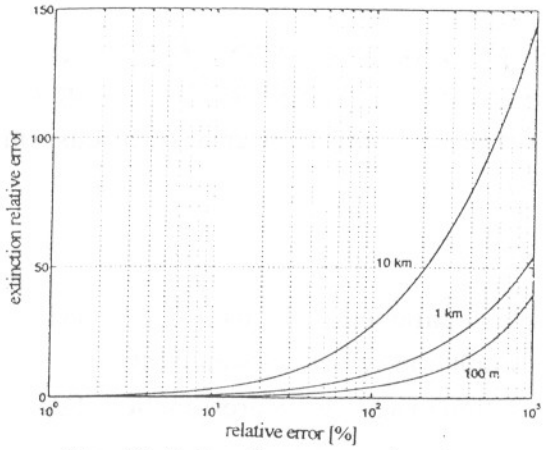

Fig.16 Extinction error due to overestimated calibration.

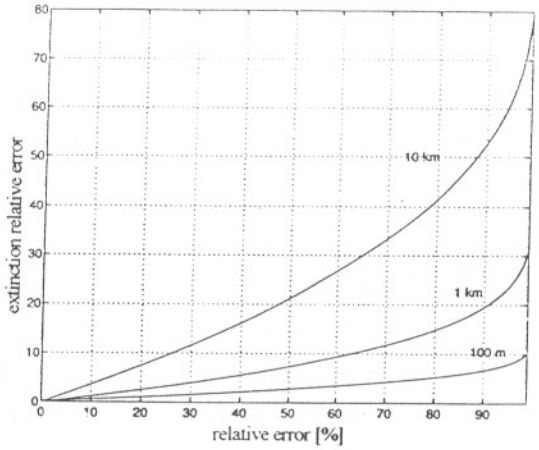

Fig.17 Extinction error due to underestimated calibration.

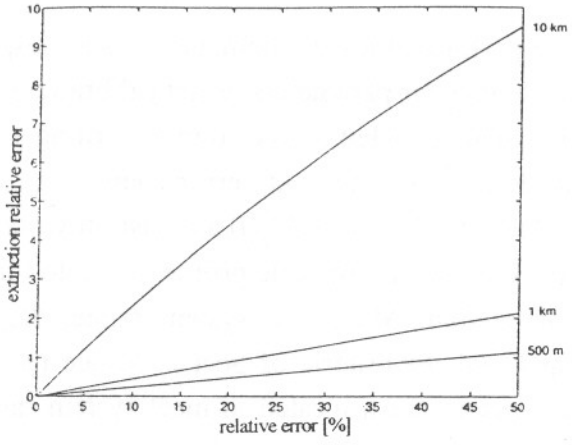

Fig.18 Extinction error due to overestimated correlation factor.

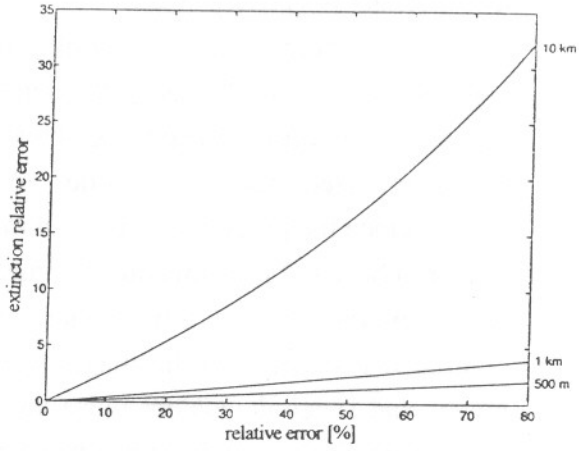

Fig.19 Extinction error due to underestimated correlation factor.

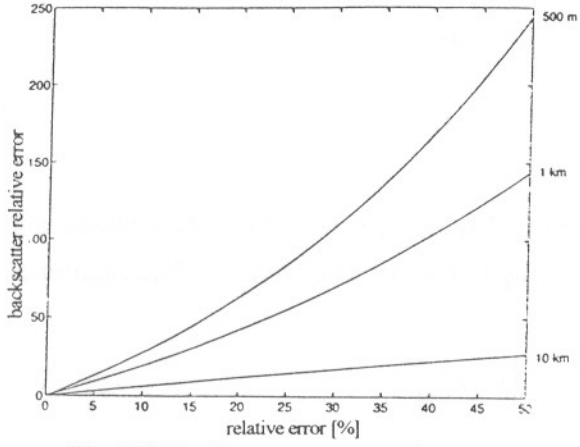

Fig.20 Backscatter error due to overestimated correlation factor.

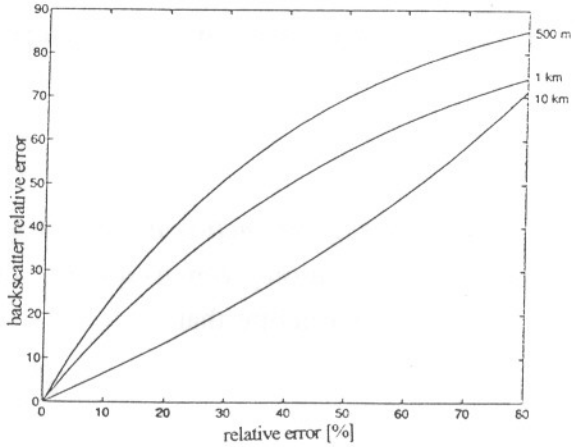

Fig.21 Backscatter error due to underestimated correlation factor.

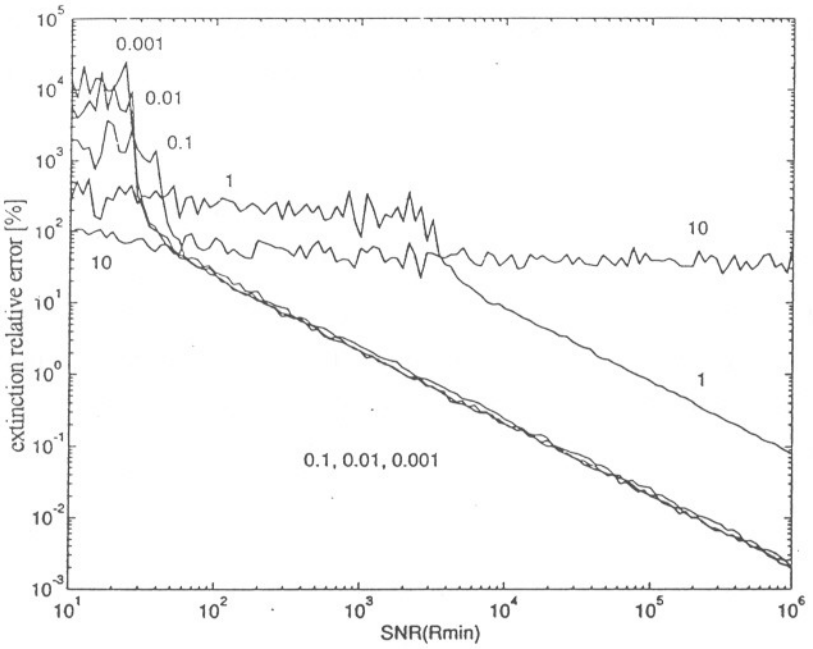

Fig.22 Extinction error vs. SNR and α in Klett's method.

4.2.4 Homogeneous atmosphere study

All the points settled for the time being, one may well wonder whether it is worth applying Klett's method to homogeneous atmospheres.

The study has managed a perfect calibration and a *k-constant* set to unity. Following the same procedure as for the homogeneous algorithms, the extinction has spanned the interval $(10^{-3},10)\ km^{-1}$ and the SNR the interval $(10,10^{6})$. *C-constant* in eq.(27) has been computed as the ratio β/α. These parameters can be known from extrapolation of Tab.3, once given a visibility margin, V_M. If inversion errors are cross-examined, a similar behaviour arises, though overall Klett's quantitative error performance is worse. This can be explained if we realize that the correlation relation remains unexploited. This is because the optical parameters are constants, and their relation, therefore, a constant. Except for the plots undergoing higher errors, no qualitative differences appear with respect to the inhomogeneous simulations.

4.3 LANL live scene inversion

As in Sect.3.4, LANL raw data has been inverted using Klett's algorithm. The most important parameters here are the calibration, α_m, and the correlation constant, k.

The first one was taken equal to the one yielded by a least-squares inversion. This option seemed more reliable because the use of eq.(32) derives from slope-method, which exhibits higher inversion errors than least-squares.

As for the second parameter, it was assumed $k=1$, as there was no prior knowledge from the correlation between the optical parameters.

The use of hybrid least-squares has also helped to find the C-constant of eq.(33) (note that C and k are needed to relate extinction and backscatter). It has been estimated using

$$C' = \frac{\hat{\beta}(R_{\min})'}{\hat{\alpha}(R_{\min})} \qquad \beta(R)' = A\beta(R) \tag{37}$$

Again, recall that the estimated backscatter $\beta(R)'$, includes the system constant, A.

Fig.23 compares a 36.5°-elevation-profile inverted by Klett's method to the same one using least-squares. As for the backscatter-coefficient, there are no significant differences.

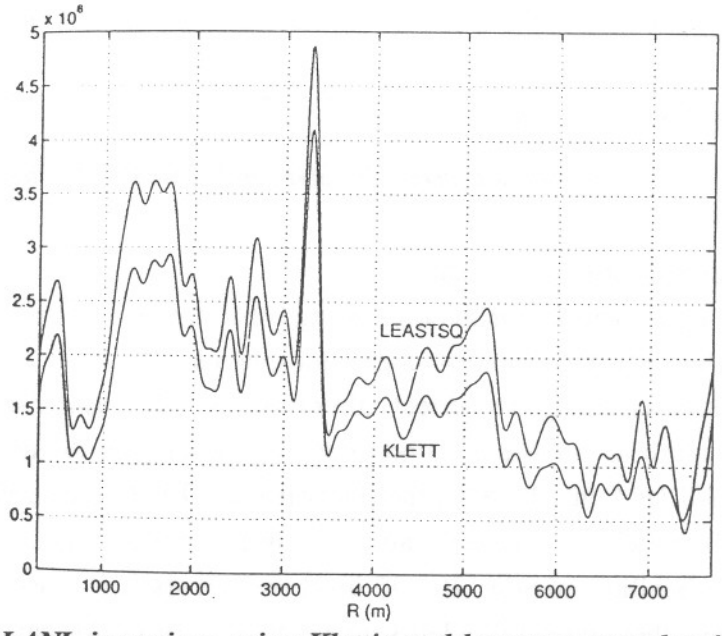

Fig.23 LANL inversions using Klett's and least-squares algorithms.

5. CONCLUSIONS

5.1 About slope-method and least-squares

The studies conducted to spotlight error dependency on SNR have shown two important points:

First, ***α-inversion error increases as the visibility margin does.*** To illustrate this with an example, consider a high V_M in an homogeneous atmosphere. Then, the noisy range-received power can be approximated by

$$P(R) = \frac{A}{R^2}\beta e^{-2\alpha R} + n(R) \approx \frac{A}{R^2}\beta(1-2\alpha R) + n(R) \tag{38}$$

Obviously, as V_M increases, the extinction tends to be masked by the noisy term. Thus, low noise spikes added to a power sample must be offset by large deviations in the extinction value, during the inversion process. This results in large inversion errors.

Secondly, and contrary to what might have been expected at first sight, ***β-inversion error decreases as the visibility margin increases.*** This trend reversal in relation to α can be accounted for by looking into lidar SNR: Given a V_M, *SNR(R)* becomes higher, at a same range, for higher transmittance atmospheres (higher visibilities).

Lastly, a further remarkable effect of these algorithms is that *β-errors vs. SNR plots tend to coincide for small extinctions.* Prevailing this condition, SNR(R) decreases as R^{-2}, no matter the atmospheric extinction.

About the question of which is the best algorithm amongst them, the plots shown in Fig.5 to Fig.8 appear to point towards the least-squares algorithm though slope-method is needed to provide sensible initialization guesses.

5.2 About Klett's method

To begin with, Klett's error inversion study vs. calibration proves that ***the larger the visibility range, the higher the calibration error.*** A first explanation to this would be that for high V_M, the extinction α becomes small and thus, its inverse in eq.(31) grows up to a large figure. As a result, this equation becomes more sensitive to the V_M. A second explanation is that as V_M dwindles, so does the range-corrected function, the calibration term being more and more important. To prove the latter affirmation one can resort to the integral version of eq.(29)

$$S - S_m = k \cdot \ln \frac{\alpha}{\alpha_m} + 2 \int_R^{R_m} \alpha(r) dr \qquad (39)$$

As the quotient in the logarithm argument can be considered a relative value independent of V_M, the extinction in the integral (equivalently its counterpart, V_M) is in direct proportion to the range-corrected function *S(R)*.

To follow on, when Klett's error versus SNR is studied, it has been shown that ***inversion error in both optical parameters increases as so does the visibility margin.*** In the same way as for the homogeneous atmosphere algorithms, low power-to-extinction sensitivity has to be blamed for (eq.(38)). Note that in Klett's method one can talk indistinctly about α-inversion error or β-inversion error as they look alike because of the correlating relation given in eq.(33).

Summing up evaluation results, it can be said that ***Klett's overall behaviour is quite different from that of the slope-method and least-squares algorithms. With them,*** it has been shown that the extinction error followed a similar trend, that is, ***α-inversion error increased as the visibility margin did*** **but** ***β-inversion error decreased.***