Statistics of the slope-method estimator

Francesc Rocadenbosch, Adolfo Comerón, and Lorena Albiol

The slope method has customarily been used and is still used for inversion of atmospheric optical parameters, extinction, and backscatter in homogeneous atmospheres from lidar returns. Our aim is to study the underlying statistics of the old slope method and ultimately to compare its inversion performance with that of the present-day nonlinear least-squares solution (the so-called exponential-curve fitting). The contents are twofold: First, an analytical study is conducted to characterize the bias and the mean-square-estimation error of the regression operator, which permits estimation of the optical parameters from the logarithm of the range-compensated lidar return. Second, universal plots for most short- and far-range tropospheric backscatter lidars are presented as a rule of thumb for obtaining the optimum regression interval length that yields unbiased estimates. As a result, the simple graphic basis of the slope method is still maintained, and its inversion performance improves up to that of the present-day computer-oriented exponential-curve fitting, which ends the controversy between these two algorithms. © 2000 Optical Society of America
OCIS codes: 010.0010, 010.1290, 010.3640.

1. Introduction

The slope method first proposed by Collis[1] has been used for more than three decades and is still used as a simple straightforward algorithm for estimating the atmospheric extinction and backscatter optical components in lidar systems under the assumption of a homogeneous atmosphere. This method departs from the simple single-scattering monostatic lidar equation[2]

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

where $P(R)$ represents the range-return power (in watts), $\beta(R)$ is the range-dependent atmospheric backscatter coefficient (in inverse kilometers per steradian), $\alpha(R)$ is the range-dependent atmospheric extinction coefficient (in inverse kilometers), R is the range (in kilometers), and K is the system constant (in watts times cubic kilometers) (a function of the laser energy and effective receiver area).

Under the assumption of a homogeneous atmo-

The authors are with the Department of Signal Theory and Communications, Antennas, Microwaves, Radar, and Optics Group, Universitat Politecnica de Catalunya, C/Sor Eulalia de Anzizu s/n, 08034 Barcelona, Spain. The e-mail address for F. Rocadenbosch is roca@tsc.upc.es.

sphere [$\alpha(R) \approx \alpha$, $\beta(R) \approx \beta$], the sought-after optical coefficients (α and β) can be estimated from slope m and intercept c of a linear regression applied to the logarithm of the range-corrected lidar signal:

$$S(R) = \ln\{R^2[P(R) + n(R)]\}$$

$$= \underbrace{mR + C}_{\text{ideal line}} + \underbrace{\ln\left[1 + \dfrac{n(R)}{P(R)}\right]}_{\text{noisy term}}, \qquad (2)$$

with

$$m = -2\alpha, \qquad c = \ln(K\beta). \qquad (3)$$

The utility of the slope method increases with increasing optical depth, and, in particular, for optical depths greater than unity, a safe estimate can usually be obtained[3] from Eq. (2) [if the signal-to-noise ratio (SNR) is high enough, the noisy term can be made redundant]. However, we must warn the reader that *a priori* information confirming the existence of a homogeneous atmosphere is strictly necessary. Otherwise, the algorithm may yield to situations in which an inhomogeneous reflection (e.g., a cloud reflection) cannot be distinguished from the returns of a homogeneous atmosphere (see Appendix B in Ref. 4). Other applications of the slope method in inhomogeneous atmospheres encompass boundary calibration of Klett's method[3,5] and the so-called slice method of inversion, which is merely an extremely close variant of the slope method but applied to successive range intervals.[6]

As originally reported by Kunz and Leeuw,[7] although the simplicity and straightforwardness of the slope method is advantageous, its biased estimation for low SNR's, caused by samples of the return signal approaching the noise floor [$n(R) \approx -P(R)$ in Eq. (2)], emerges as the main drawback of the algorithm. This effect is illustrated in Section 2, and in Section 3 we study the bias and the mean-square error (MSE) associated with the regression operator used in the slope-method algorithm from a statistical point of view.

Currently, with the computational tools available, direct application of the slope method itself has been superseded (in terms of inversion errors) by the application of nonlinear least-squares (NLSQ) procedures to exponential-curve fitting, as demonstrated in Ref. 8. Formally, we wish to solve

$$\min[\|S(R) - b\exp(-aR)\|^2]_{a,b} =$$
$$\min\left\{\sum_{i=1}^{N}[S(R_i) - b\exp(-aR)]^2\right\}_{a,b}, \quad (4)$$

where R_i is the range of the ith observation cell and a and b are unbiased estimates of the terms 2α and $K\beta$, respectively.

The largest advantage of the exponential-curve-fitting method over the customary slope method, in which all the available samples are considered for regression, is that the exponential fitting of Eq. (4) introduces no bias and therefore there is a lower inversion error result. This is because Eq. (4) gives more weight to the largest terms of the summation, which correspond to the first samples of $S(R)$ and for which the SNR is higher. However, its main disadvantages are that a computational approach is needed to solve nonlinear Eq. (4) by means of some numerical iterative recipe[9] and that a reasonably good initialization is also needed for the algorithm to pay off. In such iterative solutions, having a good first guess, which often comes from the slope method as in the present case or from a look-up visibility table, is 90% of the battle; the inversion results are independent of the initialization method used, provided that the quality of the guess is good enough. Otherwise, unrealistic divergent results occur.[7]

In Section 4 we consider the optimum slope method, in which the optimum interval length for the minimum MSE is selected for regression, and its performance is compared with that of the NLSQ exponential-curve fitting introduced above [Eq. (4)]. Computation of the optimum length is presented in the form of universal plots suitable for most short- and far-range tropospheric backscatter lidars.

For the sake of simplicity in this study, we have assumed full overlap between the laser beam and the receiver's field of view in the regression interval of Eq. (2) and we have neglected the influence of the receiver bandwidth limitation on the estimated optical parameters. The former topic is covered in detail in Ref. 10, the latter in Ref. 11, and a practical rule for choosing the regression interval, given the range of
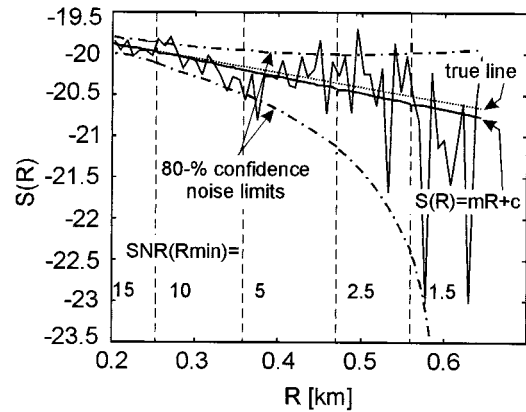

Fig. 1. Illustration of the bias effect for the slope method: solid curve, noisy realization of the log- and range-corrected signal $S(R)$ in Eq. (2); dotted line, unbiased ideal line built from the true atmospheric parameters in the simulator ($\alpha = 1$ km^{-1}, $\beta = 3 \times 10^{-2}$ km^{-1} sr^{-1}); solid line, regressed line from the noisy $S(R)$ in application of the slope method. The line becomes tilted downward as a result of the estimator bias; dashed–dotted curves, 80% noise confidence limits for N_i' [Eq. (8)].

full overlap and the receiver bandwidth, is given in Ref. 8.

2. Noise Performance

If the noise contribution $n(R)$ at each observation cell R_i along the lidar exploration path is interpreted as a random variable N_i, then the range-corrected function $S(R)$ of Eq. (2) can be expressed as a set of random variables of the form

$$S_i = \ln[R_i^2(P_i + N_i)] = c + mR_i + N_i', \quad (5)$$

where the log-noise and the observation-noise random variables N_i' and N_i, respectively, are related by

$$N_i' = \ln\left(1 + \frac{N_i}{P_i}\right). \quad (6)$$

(Hereafter, capital letters represent random variables except for P_i and R_i, which denote the deterministic return power and the range, already defined; lowercase letters represent feasible values of the random variable in its existence domain).

A. Bias

If the signal strength $P(R)$ is typically more than 50 photons over the integration time, as is always the case, the statistics of the observation-noise random variable N_i, which merges into single-body signal-induced dark-current shot noises and electronic thermal noise, can be assumed continuous Gaussian ones[8] (computation of the system's total noise variance is also given in the same reference). Consequently there is some likelihood, particularly higher for decreasing SNR's, that a noise realization of N_i approaching the noise floor $-P_i$ at the ith observation cell R_i causes either a large noise spike in N_i' or even a singularity in Eq. (6). This is illustrated in Fig. 1, in which it is evident that noise spikes at the end
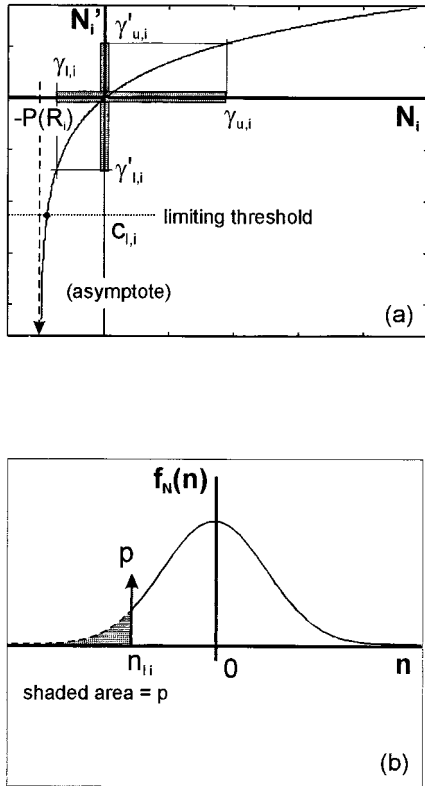
Fig. 2. (a) Equation (6) functional relationship between N_i' and N_i random variables. Note that the symmetric interval $[\gamma_{l,i}, \gamma_{u,i}]$ for the Gaussian variable N_i translates into an asymmetric one $[\gamma_{l,i}', \gamma_{u,i}']$ for variable N_i'. As a result $E[N_i'] \neq 0$ and the estimator becomes biased. (b) PDF of the threshold-limited Gaussian random variable N_i.

range are so large that the regressed line computed from direct application of the slope method [Eq. (2)] becomes biased (i.e., slightly tilted downward) from the ideal line built with the true atmospheric parameters ($\alpha = 1 \text{ km}^{-1}$, $\beta = 3 \times 10^{-2} \text{ km}^{-1} \text{ sr}^{-1}$) in the simulator.[8] This is how inversion errors arise.

The gist of the role bias plays on the slope method for different SNR's can be inferred from the limiting locus

$$[S_0(R) - \gamma_l'(R), \qquad S_0(R) + \gamma_u'(R)] \qquad (7)$$

in Fig. 1, where $S_0(R)$ is the noiseless ideal regression line indicated in Eq. (2) and $\gamma_l'(R)$ and $\gamma_u'(R)$ are two functions, still to be computed, that ensure at each range R_i that

$$p(N_i' \leq \gamma_{l,i}') = 10\%, \qquad p(N_i' \geq \gamma_{u,i}') = 10\%, \qquad (8)$$

where the subscript l indicates the lower limit, u indicates the upper limit, and i indicates R_i.

To solve Eq. (8) for $\gamma_{l,i}'$ and $\gamma_{u,i}'$, we use the fact that Eq. (6) is a monotonic increasing function [see Fig. 2(a)] and we use the basic properties of probability density function (PDF) transformations.[12,13]

By these means, the probability of N_i' can be related to that of the Gaussian random variable N_i as

$$p(N_i' \leq n_i') = p(N_i \leq n_i) = \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{n_i}{\sqrt{2}\sigma_i}\right), \qquad (9)$$

so that solution of the first equation of Eqs. (8) becomes

$$p(N_i' \leq \gamma_{l,i}') = 0.1 \Rightarrow \gamma_{l,i}' = \ln\left[1 - \frac{1.28\sigma(R_i)}{P(R_i)}\right]. \qquad (10)$$

The procedure is analogous for $\gamma_{u,i}'$.

In addition, by using Eq. (8) and relation (10), we can establish a one-to-one relation between the 80% probability intervals in N_i and N_i' as follows:

$$p(\gamma_{l,i}' \leq N_i' \leq \gamma_{u,i}') = 80\% \Leftrightarrow$$
$$p(-1.28\sigma_i \leq N_i \leq 1.28\sigma_i) = 80\%. \qquad (11)$$

Finally, after introducing the SNR definition, we can write the solutions of Eqs. (8) as

$$\gamma_{l,i}' = \ln\left[1 - \frac{1.28}{\text{SNR}(R_i)}\right], \qquad \gamma_{u,i}' = \ln\left[1 + \frac{1.28}{\text{SNR}(R_i)}\right]. \qquad (12)$$

Figure 1 shows the Eq. (7) locus for $\alpha = 1 \text{ km}^{-1}$, $R_{\min} = 0.2 \text{ km}$, and $\text{SNR}(R_{\min}) = 15$ (in linear units). Because for a given backscatter lidar system we can always relate the product $K\beta$ in Eq. (1) [or, alternatively, K if the pair (α, β) is related by means of some look-up visibility table[14,15] to the SNR at the starting point $R_{\min}$ of full overlap between the laser beam and the telescope field of view[7,8] [hereafter $\text{SNR}(R_{\min})$], we can indistinctly parameterize our plots in terms of $\text{SNR}(R_{\min})$'s or K's. Selection of $\text{SNR}(R_{\min})$ instead of K in Fig. 1 is preferred because the former is not related to any particular set of system parameters.

For $\text{SNR}(R) < 5$, a simple visual inspection of Fig. 1 is enough to reveal that the N_i' 80% confidence limit starts to distribute asymmetrically. In other words, relation (11) translates a symmetric probability distribution (a Gaussian one) for N_i into a nonsymmetric one for the log-random variable N_i'. This effect is at the core of the slope-method algorithm, and it is responsible for unwanted biased estimation ($E[N_i'] \neq 0$, where $E[\,.\,]$ is the expectancy operation).

B. Singularities

In theory, noise values n_i equal to or exceeding the asymptotic noise floor $-P(R_i)$ in Fig. 2(a) lead to undefined figures in the logarithm of the range-compensated signal $S(R)$. In practice, however, these singularities are avoided by means of some limiting criterion. Two of the most commonly used criteria consist of either discarding or resetting negative peaks of $S(R)$ larger than some predefined threshold v_l. In the former case, discarded points are not used for regression and the observation points R_i become a nonuniform set. The latter case yields a uniform set R_i, and it is equivalent to limiting the domain of the

N_i PDF $f_{N_i}(n_i)$ to $[-|n_{l,i}|, \infty)$ [see Fig. 2(b)]. Formally, after Eq. (6)

$$S_i = v_l \Rightarrow n_{l,i} = P_i[\exp(c_{l,i}) - 1], \quad (13)$$

where

$$c_{l,i} = v_l - c - mR_i = \ln\left(1 + \frac{n_{l,i}}{P_i}\right). \quad (14)$$

Note that in Eq. (13), the same limiting threshold v_l yields to different limits, $n_{l,i}$, one for each cell along the exploration path. A good proposal for v_l is to locate it approximately one unit below the expected $S(R_{max})$. Formally,

$$v_l = S_1 - 2\hat{\alpha}_{max}(R_{max} - R_1) - 1, \quad (15)$$

where S_1 is the first sample of $S(R)$ and $\hat{\alpha}_{max}$ is a user estimation of the largest expected extinction. This ensures that the limiting threshold v_l is well below the ideal regression line [Eq. (2)].

Of course, each decision rule (resetting or discarding) leads to different statistics for the truncated Gaussian N_i and therefore a different bias performance should, in principle, be expected. However, extensive simulation has shown that either a rule or the v_l figure plays a significant role in the final estimator performance only if $\mathrm{SNR}(R_{min}) < 30$. Taking into account that, for most tropospheric backscatter lidars $\mathrm{SNR}(R_{min})$ is approximately between 10^2 and 10^4 for atmospheric extinctions in the 10^{-2}–10-km^{-1} range,[8] we can assume that the statistical results that are derived next are not restricted to any decision rule, even though the resetting decision rule of Eq. (13) is assumed for simplicity reasons.

3. Statistics of the Regression Operator

Close inspection of Fig. 1 suggests that if we removed the far-end samples from the regression interval, bias would be considerably reduced because it is precisely the end part of the regression interval where the 80% confidence noise limits are most asymmetric. However, if the regression interval were too short, fewer useful samples would be used and therefore the inversion error would have a larger variance. As a result, it seems sensible to guess that an optimum regression interval length for minimum MSE exists.

In this study, we commit to the statistics of slope estimate $\hat{m}$ [Eqs. (3)] only because the procedure is completely similar for the intercept estimate $\hat{c}$, although mathematically it is much more involved and does not shed any new light on the concluding results.

A. Bias and Mean-Square Error

The estimates of slope $\hat{m}$ and intercept $\hat{c}$ for a straight-line fit to the log- and the range-corrected lidar set of data $\{x_i = R_i, y_i = S_i\}, i = 1 \ldots N$ takes the form[16]

$$\hat{m} = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}, \qquad \hat{c} = \bar{y} - \hat{m}\bar{x}, \quad (16)$$

where $\hat{m}$ and $\hat{c}$ yield the sought-after atmospheric optical estimates after Eqs. (3) and $x_N = R_N$ is defined as the maximum detection range; equivalently, N represents the maximum number of available samples up to that range where the received power equals the noise-equivalent power of the receiver [$\mathrm{SNR}(R_N) = 1$]. A plot of R_{max} versus $\mathrm{SNR}(R_{min})$ is shown in Fig. 1 of Ref. 8, and a mathematical approximation of R_{max} can be found in Appendix A of Ref. 7.

Because for each incoming measurement, Eqs. (16) yield the pair of estimates $\hat{m}$ and $\hat{c}$, which are themselves random variables and are expected to be close to their atmospheric counterparts m and c, we can compute the bias and the MSE associated with them. The estimator bias associated with $\hat{m}$ is defined as

$$\mathrm{bias} = E[\hat{m} - m] = E[\hat{m}] - m, \quad (17)$$

and the MSE takes the form

$$\varepsilon_{\hat{m}}^2 = E[(\hat{m} - m)^2] = E[\hat{m}^2] - 2mE[\hat{m}] + m^2. \quad (18)$$

The solutions of Eqs. (17) and (18) require computation of the first and the second moments of $\hat{m}$ as follows.

We can easily compute the first moment of $\hat{m}$ in two steps, first by applying the expectancy operator to both terms of Eqs. (16) and second by expressing the expectancy of S_i, $E[S_i]$ as a function of the log-noise expectancy $E[N_i']$ by means of Eq. (5). This yields

$$E[\hat{m}] = m + \frac{1}{NQ}\sum_{i=1}^{N}(R_i - \bar{R})E[N_i'], \quad (19)$$

where Q will be defined by Eq. (25) for contextual reasons. The computation of $E[\hat{m}^2]$ is more involved, but it follows a similar procedure that finally reduces to the computation of $E[N_i']$, $E[N_j']$. Assuming independence if $i \neq j$, we obtain

$$E[\hat{m}^2] = E[\hat{m}]^2 + \frac{1}{N^2Q^2}\sum_{i=1}^{N}(R_i - \bar{R})^2(E[N_i'^2]$$
$$- E[N_i']^2). \quad (20)$$

Finally, substitution of Eq. (19) into Eq. (17) and of Eqs. (19) and (20) into Eq. (18) yields the sought-after bias and MSE, respectively. It follows that

$$\mathrm{bias} = \frac{1}{NQ}\sum_{i=1}^{N}(R_i - \bar{R})E[N_i'], \quad (21)$$

$$\varepsilon_{\hat{m}}^2 = \frac{1}{N^2Q^2}(A + B), \quad (22)$$

where

$$A = \left(\sum_{i=1}^{N} (R_i - \bar{R}) E[N_i'] \right)^2, \qquad (23)$$

$$B = \sum_{i=1}^{N} (R_i - \bar{R})^2 (E[N_i'^2] - E^2[N_i'])$$

$$= \sum_{i=1}^{N} (R_i - \bar{R})^2 \sigma_{N_i'}^2, \qquad (24)$$

$$Q = \frac{1}{N} \sum_{i=1}^{N} R_i^2 - \bar{R}^2. \qquad (25)$$

Interpretation of the bias results of Eq. (21) becomes evident after our discussion in Subsection 2.A and identification of the term $E[N_i']$.

For the MSE results of Eq. (22), it follows that the MSE is proportional to the sum of two terms, A and B, which are related to different noise characteristics. The former depends on the estimation bias $E[N_i']$, whereas the latter has to do with the log-noise power (i.e., the variance of N_i'). However, mean and variance of the N_i' are not the only key factors that determine the final MSE performance because the noise contribution from each cell to the final MSE is weighted by the distance from that cell to the center of the regression interval $(R_i - \bar{R})$. As a result, noise from cells located at the inversion interval boundaries is magnified. The definition of Q in Eq. (25) is also related to the interval length, and it represents the variance of the discrete set R_i.

B. Computation of the Optimum Regression Interval Length

It is now time to justify our qualitative reasoning given in the introduction of Section 3 from a mathematical point of view by using Eqs. (22)–(25) in this way: Assume that we consider a subinterval of length $k < N$ samples, which begins at $R_{\min}$ as above so that we substitute k by N in Eqs. (22)–(25). If we increase the interval length k toward N, which is equivalent to including new samples from the right-hand side of Fig. 1 to the regression process, this not only increases Q but also A and B terms in Eqs. (23) and (24), respectively. This is so because the new samples represent not only additional terms in the series but higher bias and variance contributions, which in turn are magnified by the fact that these samples are located farther from the interval center point $\bar{R}$. As a result, Eq. (22) becomes the ratio of two numbers progressively larger. As long as the increment of Q offsets that of the term $A + B$, the MSE decreases and adding new samples to the regression interval is advantageous. However, for some $k \geq N_{\mathrm{opt}}$ (note that $N_{\mathrm{opt}} = N$ is also possible for large SNR's) this trend reverses and the MSE worsens. Consequently, we are interested in solving

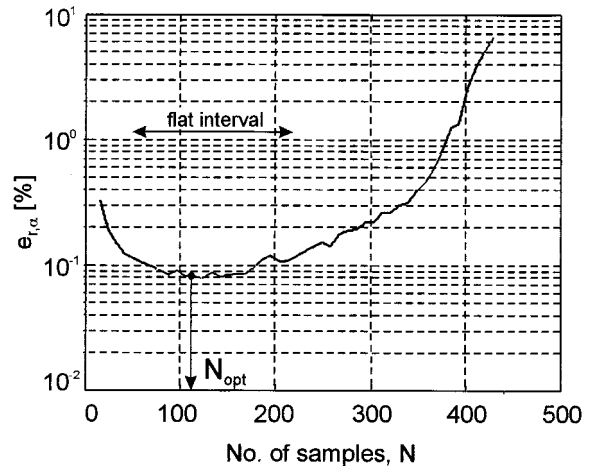$$\frac{\mathrm{d}\varepsilon_{\hat{m}}^2}{\mathrm{d}N} = 0 \Rightarrow N = N_{\mathrm{opt}}. \qquad (26)$$

Fig. 3. Equation (22) plot of the relative inversion error [as defined in Eq. (27)] versus N. The study parameters are $\alpha = 1$ km^{-1}, SNR($R_{\min} = 100$, $R_{\min} = 0.2$ km, $\Delta R = 7.5$ m (sampling rate 20×10^6 samples/s), $\rho = 2 \times 10^{-5}$ Hz$^{-1/2}$.

The best way to do this is to find the minimum MSE from a MSE versus N plot of Eq. (22), as shown in Fig. 3. Note that Fig. 3 plots the relative inversion error defined as

$$e_{r,\alpha} = \frac{\sqrt{\mathrm{MSE}_{\hat{m}}}}{2\alpha} \times 100\%. \qquad (27)$$

The Fig. 3 plot requires successive evaluation of Eqs. (22)–(25) and, particularly, computation of the first- and the second-order moments of N_i', $E[N_i']$, and $E[N_i'^2]$. In the most general case, such computation implies evaluation of the integral expressions in Eqs. (A5) and (A6) in Appendix A, but for SNR(N_{opt}) > 9 it has been found by experiment that approximations (B7) and (B8) are good, easy-to-use approximations that yield ± 1 sample errors for most extinction ranges ($\alpha = 0.1$–10 km^{-1}). In general, for those interested in either of these two mathematical procedures (Appendix A or Appendix B), we advise rounding the N_{opt} figure upward because the flat zone of the MSE versus N plot always tend to be larger to the right of N_{opt}, as in Fig. 3. The procedure is relatively simple if the SNR enables us to use the approximated expressions given in Appendix B.

4. Optimum Slope Method and Exponential-Curve Fitting

The main problem with the procedure just described to solve Eq. (26) is that the whole computational procedure must be repeated for each particular case study, which comprises the following parameters: the (simulated) atmospheric extinction α under study, the minimum system range $R_{\min}$, the spatial sampling ΔR (equivalently, the system acquisition rate), the SNR at the minimum range SNR($R_{\min}$), and, finally, a set of different system noise parameters, namely, the signal-induced and the dark-current shot-noise spectral densities and the thermal-noise spectral density $\sigma_{\mathrm{sh\_s},i}^2$, $\sigma_{\mathrm{sh\_d},i}^2$, and

Table 1. Parameter Sets Used in the Compilation of the Universal Plots

α (km^{-1})	0.1, 1, 10
SNR($R_{\min}$)	30, 100, 251, 630, 1584, 3961, 10.000
$R_{\min}$ (km)	0.2, 0.6, 1
ΔR (m)	7.5, 15, 35, 50
ρ (Hz$^{-1/2}$) [Eq. (28)]	7×10^{-5}, 2×10^{-5}, 6×10^{-6}

$\sigma_{\text{th},i}{}^2$, respectively.[17,18] With so many parameters, the computation of Eq. (22) becomes cumbersome and time-consuming.

Working on the basis of simplicity has suggested to us the goal of drawing universal plots to solve N_{opt}. These plots should be parameterized by a subset of independent parameters so that the remaining parameters from those given above should play a secondary role in the determination of N_{opt}. The universality of these plots stems from the fact that we have considered typical high-, medium-, and low-performance front-end lidar receivers, either avalanche photodiode or p-i-n-based receivers.

The struggle to find the most appropriate subsets has led us to an advanced computer-oriented statistical study that solves Eqs. (22)–(26) for all combinations of the typical system parameters of Table 1. For SNR($R_{\min}$) $\geq$ 100 in Table 1, approximations (B7) and (B8) have been used, but for SNR($R_{\min}$) = 30 we have had to resort to the integral computation of Eqs. (A5) and (A6). In addition, decorrelation studies have suggested the definition of an auxiliary noise parameter ρ, which merges into a single-body detector and system-noise parameters as follows:

$$\rho = \frac{2qFM}{(\sigma_{\text{sh}\_d,i}{}^2 + \sigma_{\text{th},i}{}^2)^{1/2}} \ (Hz^{-1/2}), \qquad (28)$$

where q is the electron change, M is the avalanche photodiode multiplication factor, F is its related excess-noise factor,[19] $\sigma_{\text{th},i}{}^2$ is the system's equivalent thermal-noise spectral density (in square amperes

per hertz) and $\sigma_{\text{sh}\_d,i}{}^2$ is the receiver's dark-current spectral density (in square amperes per hertz), both of which are referred to as the transimpedance amplifier's input (see Appendix A of Ref. 8 for further insight). [One can easily extend Eq. (28) to include p-i-n-based receivers by taking $F = 1$, $M = 1$ in Eq. (28).]

As a first result of the statistical simulation, we found that all the study cases of the Table 1 output curves are similar in shape to those in Fig. 3, all of which consist of a plateau-like interval around N_{opt}, which spans approximately ± 0.1 N. In the second step, this morphological feature was put at the core of a regression study, as in the example of Fig. 4(a), which enabled us to identify N_{opt}/N, SNR($R_{\min}$), and the atmospheric extinction α as the primary subset and $R_{\min}$, ΔR, and ρ as the secondary subset. In all cases, the quality of the regressed data has been checked with the exact computation of N_{opt}, again by use of Eqs. (22)–(25). The relative regression error $\varepsilon_{\text{fit}}{}^\alpha$ is typically less than 5% for $\alpha = 10$ km^{-1} [see the Fig. 4(b) histogram] and ~1% for $\alpha = 0.1$–1 km^{-1} (not shown).

The universal plots, which enable us to solve the ratio N_{opt}/N from an available estimation of the atmospheric extinction α (its order of magnitude) and SNR($R_{\min}$), are shown in Fig. 5. Note that N is defined as that range in which the received power is equivalent to the noise-equivalent power of the receiver [SNR(R_N) = 1] and hence it must be known to solve N_{opt} from the ratio N_{opt}/N.

An additional comment with regard to Fig. 5 is the asymptotic behavior of N_{opt} with increasing SNR($R_{\min}$). Thus, as long as SNR($R_{\min}$) increases, for each atmospheric extinction α and system parameter subset ($R_{\min}$, ΔR, ρ), N_{opt} tends toward a saturation figure, which we determined as SNR($R_{\min}$) = 500 for $\alpha = 10$ km^{-1}, SNR($R_{\min}$) = 1500 for $\alpha = 1$ km^{-1}, and SNR($R_{\min}$) = 10^4 for $\alpha = 0.1$ km^{-1}. Therefore there is no need to recompute N_{opt} for SNR($R_{\min}$)'s greater than these SNR limits.
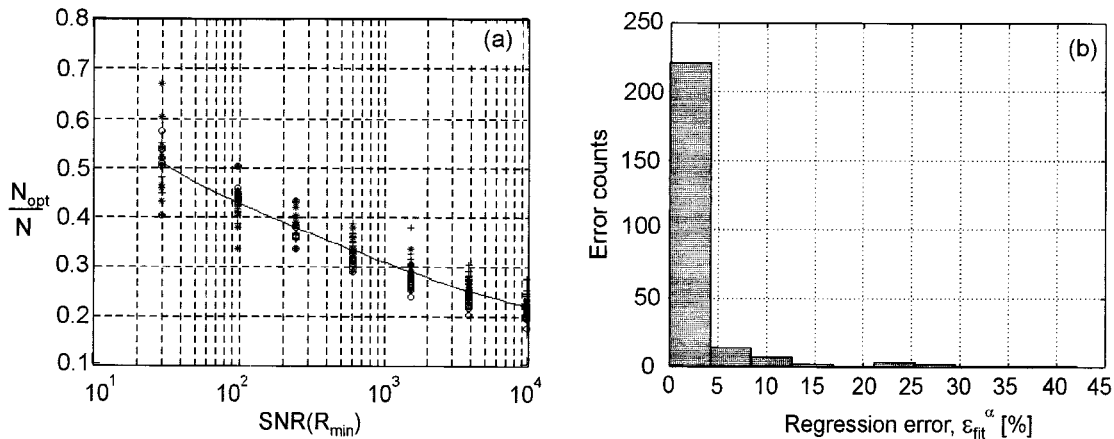

Fig. 4. (a) Example of a regression plot used to compute the universal plots. This plot solves N_{opt}/N given SNR($R_{\min}$) for a simulated atmospheric extinction ($\alpha = 10$ km^{-1} here). In this example, the regressed data sweep only ρ from Table 1: +, $\rho = 6 \times 10^{-6}$ Hz$^{-1/2}$; *, $\rho = 2 \times 10^{-5}$ Hz$^{-1/2}$; and $\bigcirc$, $\rho = 7 \times 10^{-5}$ Hz$^{-1/2}$. Note that for most of the trials N_{opt}/N is within a ± 0.1 interval, which can be fitted into the flat zone of Fig. 3. (b) Binned data used to estimate the regression error quality, which is typically ~5% for $\alpha = 10$ km^{-1}.
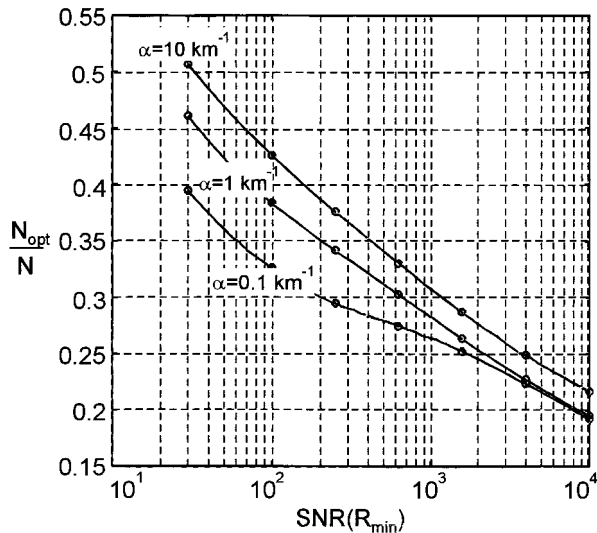
Fig. 5. Universal plots to solve N_{opt} for minimum MSE in the slope method given $SNR(R_{min})$, an estimate of the order of magnitude of the atmospheric extinction α, and N, which is defined as that range where the received power equals the noise-equivalent power of the receiver [i.e., $SNR(R_N) = 1$, as explained in Subsection 3.A].

Finally, Fig. 6 compares the error performances [Eq. (27)] achieved by the next three algorithms: Fig. 6(a), the customary slope method, i.e., by using all the available samples N with $SNR(R) > 1$ for regression and by limiting $S(R)$ to a constant threshold v_l, as described in Subsection 2.B; Fig. 6(b), the optimum slope method discussed here, i.e., by using $N_{opt} \leq N$ samples for regression and the same limiting threshold as in Fig. 6(a) to cope with singularities; and Fig. 6(c), the exponential-curve fitting described in Section 1, Eq. (4), initialized from the results of the
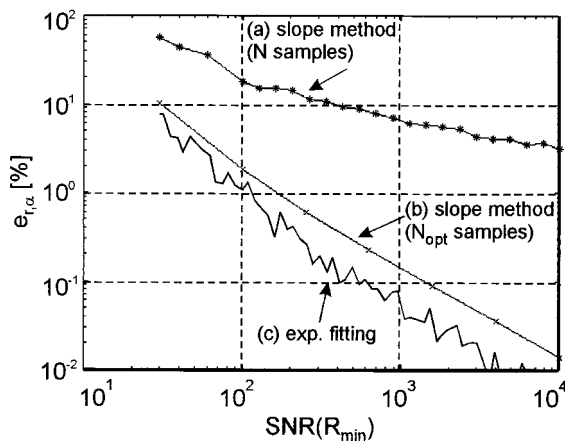

Fig. 6. Comparison of the relative inversion errors [Eq. (27)] among (a) the customary nonoptimized slope method, which uses all the available samples for inversion; (b) the optimized slope method, which uses the optimum interval length N_{opt}; (c) the NLSQ exponential fitting described in Section 1. It is shown that the performances of (b) and (c) algorithms unify in similar inversion errors, which ends an old controversy between the slope method and the exponential-curve fitting.

slope method of Fig. 6(a) to prevent divergence of the NLSQ numerical routine.[9]

If we compare the plots in Figs. 6(a) and 6(b) and note the log scale in the grid, it becomes obvious that the most significant improvement with regard to the optimum slope method in lieu of the customary method is for moderate and poor $SNR(R_{min})$'s. Thus, for low SNR's, such as $SNR(R_{min}) = 50$, $e_{r,\alpha}(N) = 40\%$ and $e_{r,\alpha}(N_{opt}) = 5\%$ represent a 35% gain; for large $SNR(R_{min})$'s, such as $SNR(R_{min}) = 10^4$, $e_{r,\alpha}(N) = 3\%$ and $e_{r,\alpha}(N_{opt}) = 0.02\%$ represent a 3% gain, although this is not significant in absolute terms. Similar conclusions follow for the rest of the atmospheric extinctions of Table 1.

Finally, a comparison of Figs. 6(b) and 6(c) proved that the slope method with optimum inversion length and the exponential-curve fitting converge to similar inversion error figures, with the optimum slope method being the simplest because the regression can even be performed graphically with the help of Fig. 5. A comparison of Figs. 6(a) and 6(c) is also in good agreement with the error figures reported in Ref. 8. As a result, we can conclude that selection of the optimum regression interval length for the slope method yields a virtually unbiased estimator whose performance in a minimum MSE is similar to that of the exponential-curve fitting of Eq. (4).

5. Conclusions

A statistical and computer-oriented simulation study has been conducted to characterize the statistics of the well-known slope method, in particular, its bias and mean-square inversion error, and to compare its inversion performance with that of the exponential-curve fitting by use of a previous study.[8]

It has been shown that the slope method is dominated by its inherent biased estimation and that inversion errors can largely be reduced if one chooses an optimum interval length for regression. To compute this length, analytical expressions, either exact or approximated, have been derived [Eqs. (22)–(25) plus Eqs. (A5) and (A6) for $SNR(R_{min}) = 30$ or approximations (B7) and (B8) for $SNR(R_{min}) \geq 100$] as well as the quick-reference universal plots of Fig. 5, which are valid for most typical elastic-backscatter lidar systems according to Table 1.

Finally, we found that, when such an optimum regression length is considered, the MSE of the slope method reduces to that of the unbiased exponential-curve-fitting estimator [Eq. (4)], and hence both algorithms are equivalent in terms of minimum bias and MSE performance.

Appendix A: Integral Computation of $E[N_i']$, $E[N_i'^2]$

We depart from the Gaussian-limited PDF random variable N_i, $f_{N_i}(n_i)$, described in Subsection 2.B and depicted in Fig. 2(b) and from application of the expectancy theorem[12] to Eq. (6), which states that the expectancy of a random variable N_i' related to another random variable N_i by means of a functional

relationship [Eq. (6) in the present case] can be computed as

$$E[N_i'] = \int_{-\infty}^{\infty} \ln\left(1 + \frac{n}{P_i}\right) f_{N_i}(n) dn. \qquad (A1)$$

Based on the fact that the existence domain of N_i is $[n_{l,i}, \infty)$ and that the interval that has been removed $(-\infty, n_{l,i})$ is equivalent to a delta function with area (probability)

$$p = \int_{-\infty}^{n_{l,i}} \frac{1}{\sqrt{2\pi}\sigma_{n,i}} \exp\left(-\frac{n^2}{2\sigma_{n,i}^2}\right) dn = \frac{1}{2} \operatorname{erf}\left(\frac{-n_{l,i}}{\sqrt{2}\sigma_{n,i}}\right), \qquad (A2)$$

Eq. (A1) can be computed as

$$E[N_i'] = \int_{-\infty}^{n_{l,i}} \ln\left(1 + \frac{n}{P_i}\right) p\delta(n - n_{l,i}) dn$$

$$+ \underbrace{\int_{n_{l,i}}^{\infty} \ln\left(1 + \frac{n}{P_i}\right) \frac{1}{\sqrt{2\pi}\sigma_{n,i}} \exp\left[-\frac{1}{2}\left(\frac{n}{\sigma_{n,i}}\right)^2\right] dn}_{\text{main term}}.$$

$$(A3)$$

After identifying $c_{l,i}$ [Eq. (14)] in the first integral term and introducing the following substitutions in the main term for enhanced numerical processing,

$$x = \ln\left(1 + \frac{n}{P_i}\right), \qquad n = P_i(e^x - 1), \qquad dn = P_i e^x dx, \qquad (A4)$$

we finally obtain

$$E[N_i'] = pc_{l,i} + \int_{c_{l,i}}^{\infty} x e^x \frac{P_i}{\sqrt{2\pi}\sigma_{n,i}}$$
$$\times \exp\left\{-\frac{1}{2}\left[\frac{P_i(e^x - 1)}{\sigma_{n,i}}\right]^2\right\} dx. \qquad (A5)$$

Computation of $E[N_i'^2]$ follows the same procedure and yields

$$E[N_i'^2] = pc_{l,i}^2 + \int_{c_{l,i}}^{\infty} x^2 e^x \frac{P_i}{\sqrt{2\pi}\sigma_{n,i}}$$
$$\times \exp\left\{-\frac{1}{2}\left[\frac{P_i(e^x - 1)}{\sigma_{n,i}}\right]^2\right\} dx. \qquad (A6)$$

Equations (A5) and (A6), together with Eqs. (22)–(25), enable exact computation of the MSE. This requires knowledge of only the measured return power samples P_i and the sampling points R_i and characterization of the range-dependent system-noise standard deviation $\sigma_{n,i}$ from system parameters.[8]

Appendix B: Practical Computation of $E[N_i']$, $E[N_i'^2]$

In nearly all practical cases, the range-dependent SNR associated with each particular sampling cell R_i is high enough to enable considerable simplification of Eqs. (A5) and (A6) into the expressions that are presented next. Validity of this approximate method is limited to $\text{SNR}(R) \gg 5$ along all the inversion range with $\text{SNR} > 9$ being a practical lower bound found by experiment.

For most practical SNR's,[7,8] the role of the limiting threshold ν_l in Eq. (13) or, equivalently, $n_{l,i}$ is to avoid singularities caused by noise figures that exceed the noise floor $-P_i$ rather than clip noise. This is particularly true for any conservative choice of ν_l well below the expected $S(R_{\max})$, as indicated in Subsection 2.B. Under these circumstances, the $n_{l,i}$ threshold is not restrictive, and hence we can write

$$n_{l,i} \ll -5\sigma_{n,i} \Rightarrow p \approx 0. \qquad (B1)$$

As a result, the first term of Eq. (A5) can be neglected and $E[N_i']$ can be approximated by

$$E[N_i'] \approx \int_{n_{l,i}}^{\infty} \ln\left(1 + \frac{n}{P_i}\right) \frac{1}{\sqrt{2\pi}\sigma_{n,i}} \exp\left[-\frac{1}{2}\left(\frac{n}{\sigma_{n,i}}\right)^2\right] dn, \qquad (B2)$$

which, in turn, we can compute by means of a Taylor series expansion by noting that

$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + O(x^5), \qquad |x| \ll 1, \qquad (B3)$$

where $x = n/P_i$. Using relation (B1) above, we can rewrite the convergence radius constraint as

$$\left|\frac{n_{\max}}{P_i}\right| \ll 1 \Rightarrow \frac{5\sigma_{n,i}}{P_i} \ll 1 \Rightarrow \text{SNR}_i \gg 5, \qquad (B4)$$

which indicates that the SNR must be much higher than 5 to ensure series convergence.

Now, if Eq. (B3) is substituted into approximation (B2) and the integration limits are expanded from $-\infty$ to ∞, again based on the fact that under the constraint of relation (B1) the contribution of the $(-\infty, n_{l,i})$ interval to the integral is negligible, then $E[N_i']$ in approximation (B2) can be expressed in terms of the moments of Gaussian random variable N_i as

$$E[N_i'] \approx \frac{E[N_i]}{P_i} - \frac{E[N_i^2]}{2P_i^2} + \frac{E[N_i^3]}{3P_i^3} - \frac{E[N_i^4]}{4P_i^4}. \qquad (B5)$$

Note that the nth-order moments of a Gaussian random variable $X = N_i$ of zero mean and σ standard deviation[12] are

$$E[X^n] = \begin{cases} \dfrac{\sigma^n n!}{2^{n/2}} & n = \text{even} \\ 0 & n = \text{odd} \end{cases}. \qquad (B6)$$

By substituting these variables into approximation (B5), we finally obtain

$$E[N_i{}'] \approx -\frac{1}{2}\frac{1}{\mathrm{SNR}_i{}^2} - \frac{3}{2}\frac{1}{\mathrm{SNR}_i{}^4}. \qquad (B7)$$

Derivation of an approximate expression for $E[N_i{}'^2]$ follows the exact same procedure, and the final result can be expressed as

$$E[N_i{}'^2] \approx \frac{1}{\mathrm{SNR}_i{}^2} + \frac{11}{2}\frac{1}{\mathrm{SNR}_i{}^4} + \frac{137}{8}\frac{1}{\mathrm{SNR}_i{}^6}. \qquad (B8)$$

At this point, it is important to warn that validity of final approximations (B7) and (B8) is conditioned to that of relation (B4), irrespective of the number of terms used for series expansion, because relation (B4) constraint is due to singularitylike behavior of the series for values close to $x = -1$. Cross examination of the exact and the approximate equation sets given by Eqs. (A5) and (A6) and approximations (B7) and (B8), respectively, has shown that the approximate set of approximations (B7)–(B8) can perfectly be used for the computation of the optimum interval length (N_{opt}, as described in Subsection 3.B) and its associated MSE [Eqs. (22)–(25) in Subsection 3A], provided that $\mathrm{SNR}(N_{\mathrm{opt}}) > 9$.

References

1. R. T. H. Collis, "Lidar: a new atmospheric probe," Q. J. R. Meteorol. Soc. **92,** 220–230 (1966).
2. R. T. H. Collis and P. B. Russell, "Lidar measurement of particles and gases by elastic backscattering and differential absorption," in *Laser Monitoring of the Atmosphere*, E. D. Hinkley, ed. (Springer-Verlag, New York, 1976), Chap. 4, pp. 71–102.
3. J. D. Klett, "Stable analytical inversion solution for processing lidar returns," Appl. Opt. **20,** 211–220 (1981).
4. G. J. Kunz, "Probing of the atmosphere with lidar," in *Proceedings of the Remote Sensing of the Propagation Environment Conference*, AGARD-CP-502, (AGARD, Neuilly sur Seine, France, 1992), Vol. 23, pp. 1–11.
5. J. D. Klett, "Lidar calibration and extinction coefficients," Appl. Opt. **22,** 514–515 (1983).
6. R. T. Brown, "A new lidar for meteorological application," J. Appl. Meteorol. **12,** 698–708 (1973).
7. G. J. Kunz and G. de Leeuw, "Inversion of lidar signals with the slope method," Appl. Opt. **32,** 3249–3256 (1993).
8. F. Rocadenbosch, A. Comerón, and D. Pineda, "Assessment of lidar inversion errors for homogeneous atmospheres," Appl. Opt. **37,** 2199–2206 (1998).
9. J. J. More, "The Levenberg–Marquardt algorithm: implementation and theory," in *Numerical Analysis*, Vol. 630 of Springer-Verlag Lecture Notes in Mathematics Series, G. A. Watson, ed. (Springer-Verlag, New York, 1977), pp. 105–116.
10. R. Velotta, B. Bartoli, R. Capobianco, L. Fiorani, and N. Spinelli, "Analysis of the receiver response in lidar measurements," Appl. Opt. **37,** 6999–7007 (1998).
11. G. J. Kunz, "Effects of detector bandwidth reduction on lidar signal processing," Rep. PHL 1977-31 (Physics and Electronics Laboratory, TNO-FEL, The Hague, 1977).
12. A. Papoulis, *Probability, Random Variables and Stochastic Processes* (McGraw-Hill, New York, 1991), pp. 345–354.
13. A. B. Carlson, "Signal transmission and filtering," in *Communication Systems*, 3rd ed. (McGraw-Hill, Singapore, 1986), Chap. 3, pp. 177–178.
14. H. Koschmieder, "Theorie der Horizontalen Sichtweite," Beitr. Phys. Freien Atmos. **12,** 33–53 (1924).
15. P. W. Kruse, L. D. McGlauchlin, and R. B. McQuiston, *Elements of Infrared Technology: Generation, Transmission and Detection* (Wiley, New York, 1962).
16. R. J. Barlow, "Least squares," in *Statistics: A Guide To The Use Of Statistical Methods In The Physical Sciences*, F. Mandl, R. J. Ellison, and D. J. Sandiford, eds. (Wiley, New York, 1989), Chap. 6.
17. W. B. Jones, *Introduction to Optical Fiber Communication Systems* (Holt, Rinehart & Winston, New York, 1988), Chaps. 7 and 8.
18. R. M. Measures, *Laser Remote Sensing: Fundamentals and Applications* (Krieger, Malabar, Fla., 1992), Chap. 4, pp. 138–145.
19. R. J. McIntyre, "Multiplication noise in uniform avalanche photodiodes," IEEE Trans. Electron Devices **ED-13,** 164–168 (1966).