Methodology for error analysis and simulation of lidar aerosol measurements

Philip B. Russell, Thomas J. Swissler, and M. Patrick McCormick

We present a methodology for objective and automated determination of the uncertainty in aerosol measurements made by lidar. The methodology is based on standard error-propagation procedures, a large data base on atmospheric behavior, and considerable experience in processing lidar data. It yields algebraic expressions for probable error as a function of the atmospheric, background lighting, and lidar parameters. This error includes contributions from (1) lidar signal; (2) molecular density; (3) atmospheric transmission; and (4) lidar calibration. The validity of the algebraic error expressions is tested by performing simulated measurements and analyses, in which random errors of appropriate size are injected at appropriate steps. As an example, the methodology is applied to a new airborne lidar system used for measurements of the stratospheric aerosol. It is shown that for stratospheric measurements below about 25 km, molecular density uncertainties are the dominant source of error for wavelengths shorter than about 1.1 \textmu m during nonvolcanic conditions. Because the influence of molecular scattering (relative to particulate scattering) decreases with increasing wavelength, stratospheric measurements with a Nd:YAG lidar can thus be more accurate than those made with a ruby lidar, provided that a suitable detector is used.

1. Introduction

Accurate measurements of atmospheric aerosols are of considerable importance to our understanding of the physics of the atmosphere, including the possible effects of aerosols on climatic processes. Remote measurements of the vertical distribution of aerosol backscatter by lidar contribute significantly to this understanding. Many such measurements have been made, both in the stratosphere\textsuperscript{1-4} and in the troposphere\textsuperscript{5-7}; moreover, some of these measurements have been used in radiative models to infer possible climatic significance.\textsuperscript{3,8} Also, lidar measurements of stratospheric aerosols have been compared with direct in situ measurements and favorably correlated.\textsuperscript{3,10}

As with any experimental technique, lidar measurements are subject to a number of experimental uncertainties, arising both from the measurement \textit{per se} and from the assumptions or uncertain values that enter into the data analysis. These uncertainties depend on a number of factors, including the laser wavelength and other lidar system parameters, the background light, the aerosol concentration, the accuracy, resolution, and proximity of the nearest molecular density measurement, the validity of lidar calibration procedures, and the uncertainty of the atmospheric transmission profile at the lidar location. In general, the total measurement uncertainty depends on these factors in a complicated and often counterintuitive way. Nevertheless, it is important that quantitative estimates of these uncertainties be derived, both as a means of assigning error bars to lidar-derived data products, and as a tool in designing future lidar systems for improved measurements or different operating conditions. The need for improved assessments of the uncertainties in all aerosol measurements was recently emphasized by an international group of experts.\textsuperscript{11}

In this paper we present a methodology for objective and automated determination of the uncertainty in lidar aerosol measurements. The methodology is based on
standard error-propagation procedures, a large data base on atmospheric behavior, and long experience in processing lidar data. It yields algebraic expressions for probable error as a function of the relevant parameters. The validity of these expressions is then tested by performing simulated measurements and analyses in which random errors of appropriate size are injected at appropriate steps of the measurement and analysis process. As an example, the methodology is applied to a new lidar system now being used for airborne measurements of the stratospheric aerosol.

II. Deriving Particulate Backscattering from Lidar Measurements

Lidar aerosol measurements are usually made by elastic backscattering, in which the received signal is analyzed at the same wavelength as the transmitting laser. For this case, the number of charge carriers generated by the detector, for light backscattered in a range bin of length $\Delta z$, from a single laser pulse, can be expressed as

$$N_S(\lambda,z) = \frac{\Delta z E(\lambda) A_r T_o(\lambda) T_d(\lambda) \left( \frac{\lambda}{\lambda_0} \right) F(\lambda,z)}{(z-z_L)^2} \exp \left[ -2 \int_{z_L}^{z} \sigma_e(\lambda,z') dz' \right],$$

(1)

where $F(\lambda,z)$ is the backscattering cross section per unit volume at wavelength $\lambda$ and altitude $z$, $E(\lambda)$ is the laser energy per pulse, $c$ is the speed of light, $h$ is Planck’s constant, $A_r$ is the area of the receiver, $z_L$ is the lidar altitude, $T_o(\lambda)$ is the system optical efficiency, and $T_d(\lambda)$ is the photodetector quantum efficiency. The exponential term describes the transmission of radiation on the two-way path from the transmitter to the scattering volume and back to the receiver; $\sigma_e(\lambda,z')$ is the atmospheric extinction coefficient at wavelength $\lambda$ and height $z'$.

In addition to the signal photoelectrons described by Eq. (1) (i.e., photoelectrons resulting from laser light backscattered at height $z$), the detector also generates noise photoelectrons, which result from skylight, ground light (for a downward-viewing lidar), laser fluorescence, and internal thermal effects. This noise, or backscattered light (for a downward-viewing lidar), laser fluorescence, and internal thermal effects. This noise, or backscattered light (for a downward-viewing lidar), laser fluorescence, and internal thermal effects. Hence, $F(\lambda,z)$

$$F(\lambda,z) = F_M(\lambda,z) + F_A(\lambda,z),$$

and $\sigma_e(\lambda,z) = \sigma_{eM}(\lambda,z) + \sigma_{eA}(\lambda,z)$.

It is often convenient to analyze lidar aerosol data in terms of the backscatter ratio $R(\lambda,z)$, which is defined as the ratio of total backscatter to molecular backscatter, i.e.,

$$R(\lambda,z) = \frac{F(\lambda,z)}{F_M(\lambda,z)} = \frac{F_M(\lambda,z) + F_A(\lambda,z)}{F_M(\lambda,z)}.$$

(4)

Therefore, via Eq. (2), $R(\lambda,z)$ can be derived from the lidar measurement as

$$R(\lambda,z) = \frac{(z-z_L)^2 N_S(\lambda,z)}{K(\lambda) Q^2(\lambda,z_L,z) F_M(\lambda,z)}.$$

(5)

In solving this equation, $N_S(\lambda,z)$ is provided by the lidar measurement, $Q^2(\lambda,z_L,z)$ is usually provided by a model (which may be updated using lidar data), and $F_M(\lambda,z)$ is determined from current radiosonde or satellite density data, or from a model. However, $K$, a calibration factor, is usually not well known, as it depends on transmitted energy and the efficiencies of the detector and of all optical components in the lidar system. Hence, $K$ is usually determined by normalizing the ratio $R(\lambda,z)$.

This normalization is operationally performed by searching for the minimum value of $R(\lambda,z)$ in the altitude range where $N_S(\lambda,z)$, $Q^2(\lambda,z_L,z)$, and $F_M(\lambda,z)$ are sufficiently well known. This height at which $R(\lambda,z)$ attains its minimum is designated $z^*$ (the normalization height). The parameter $K(\lambda)$ is then adjusted to force $R(\lambda,z^*)$, as defined by Eq. (5), to equal a value $R_{\text{min}}(\lambda)$, expected on the basis of previous aerosol measurements and models. Hence,

$$K(\lambda) = \frac{(z^*-z_L)^2 N_S(\lambda,z^*)}{R_{\text{min}}(\lambda) Q^2(\lambda,z_L,z^*) F_M(\lambda,z^*)}.$$

(6)

[A refinement of this procedure for the practical case of a noisy profile of $R(\lambda,z)$ is given in Appendix A.]

Using this in Eq. (5) yields

$$R(\lambda,z) = \frac{(z-z_L)^2 N_S(\lambda,z) Q^2(\lambda,z_L,z) F_M(\lambda,z^*)}{(z^*-z_L)^2 N_S(\lambda,z^*) Q^2(\lambda,z_L,z^*)} F_M(\lambda,z^*) R_{\text{min}}(\lambda).$$

(7)

This normalization process is one source of uncertainty in derived data products and is discussed further in Sec. III and by Russell and Livingston. In Sec. III we show that the commonly used value, $R_{\text{min}}(\lambda) = 1$, can sometimes lead to significant error, and we present more suitable values.

Once the backscatter ratio profile is determined from Eq. (7), one can derive from it the aerosol particle backscattering coefficient $F_A(\lambda,z)$. From the definition (4) we have

$$F_A(\lambda,z) = [R(\lambda,z) - 1] F_M(\lambda,z).$$

(8)

Substituting from Eq. (7) yields

$$F_A(\lambda,z) = \left[ \frac{(z-z_L)^2 N_S(\lambda,z) Q^2(\lambda,z_L,z) F_M(\lambda,z^*)}{(z^*-z_L)^2 N_S(\lambda,z^*) Q^2(\lambda,z_L,z^*) F_M(\lambda,z^*) R_{\text{min}}(\lambda) - 1} \right] F_M(\lambda,z).$$

(9)

This equation shows explicitly that, when the normal-
ization procedure is used, the derived aerosol backscatter coefficient at any height depends not on the signal at that height per se, but on the ratio of the signals at that height and at the normalization height. Derived \( F_M \) depends on two-way transmission through a similar ratio (which depends only on extinction between \( z \) and \( z^* \)—see below). Also, much of the dependence of derived aerosol backscatter coefficient at any height depends not on the signal at that height and at the normalization height, but on molecular density occurring at different heights.

For purposes of error analysis and also to simplify computation, it is useful to rewrite Eq. (9) as

\[
F_M(\lambda,z) = h(z - z_L, z^* - z_L) s(\lambda,z,z^*) q(\lambda,z,z^*) F_M(\lambda,z^*)
\]

where

\[
h(z - z_L, z^* - z_L) = \frac{(z - z_L)^2}{(z^* - z_L)^2},
\]

\[
s(\lambda,z,z^*) = \frac{N_S(\lambda,z)}{N_S(\lambda,z^*)},
\]

\[
q(\lambda,z,z^*) = \frac{Q(\lambda,z,z^*)}{Q^2(\lambda,z,z^*)} = \exp\left[-2\int_z^{z^*} \sigma_\epsilon(\lambda,z')dz'\right].
\]

### III. Error Analysis

For a function \( \chi \), derived from several measured variables \( u,v,\ldots \), the uncertainty in \( \chi \) can be approximated as

\[
\delta \chi^2 = (\delta u)^2 \left( \frac{\partial \chi}{\partial u} \right)^2 + (\delta v)^2 \left( \frac{\partial \chi}{\partial v} \right)^2 + 2C_{uv}^2 \left( \frac{\partial \chi}{\partial u} \right) \left( \frac{\partial \chi}{\partial v} \right) + \ldots,
\]

where \( \delta u \) is in the uncertainty in measured variable \( u \), and \( C_{uv}^2 \) is the covariance between measured variables \( u \) and \( v \). This covariance can be expressed mathematically as

\[
C_{uv}^2 = \lim_{n \to \infty} \frac{1}{n} \sum[(u_i - \bar{u})(v_i - \bar{v})]
\]

where \( u_i, \bar{u}, v_i, \bar{v} \) are the measured mean values for \( u \) and \( v \), and \( u_i, v_i \) are the values obtained in a particular measurement \( i \). \( C_{uv}^2 \) vanishes when measurement errors \( u_i - \bar{u} \) are uncorrelated with errors \( v_i - \bar{v} \) but is nonzero if these errors are correlated.

In applying Eq. (14) to Eq. (10) we first consider which, if any, covariances are nonzero. Because of the different methods of determination (or assumption), there is no correlation among \( h, s, q, R_{\text{min}} \); hence, the corresponding covariances are all zero. However, \( F_M(\lambda,z) \) values at different heights \( z, z^* \) will be correlated if \( |z_1 - z_2| \) is less than the height resolution of the density measurement (or model) used to derive \( F_M(\lambda,z) \). [This follows because the profile \( F_M(\lambda,z) \) at the height resolution of the lidar analysis is usually obtained by interpolating between measurements (or a model) having coarser resolution.] Thus, applying Eq. (14) to Eq. (10) yields

\[
(\delta F_M)^2 = (\delta h)^2 \left( \frac{\partial F_M}{\partial h} \right)^2 + (\delta s)^2 \left( \frac{\partial F_M}{\partial s} \right)^2 + (\delta q)^2 \left( \frac{\partial F_M}{\partial q} \right)^2 + (\delta R_{\text{min}})^2 \left( \frac{\partial F_M}{\partial R_{\text{min}}} \right)^2 + 2C_{hs}^2 \left( \frac{\partial F_M}{\partial h} \right) \left( \frac{\partial F_M}{\partial s} \right) + 2C_{hq}^2 \left( \frac{\partial F_M}{\partial h} \right) \left( \frac{\partial F_M}{\partial q} \right) + 2C_{qs}^2 \left( \frac{\partial F_M}{\partial q} \right) \left( \frac{\partial F_M}{\partial s} \right) + \ldots
\]

where we have assumed that errors in determining height are negligible (i.e., \( \delta h \approx 0 \)) and have defined \( F_M = F_M(\lambda,z^*) \).

Substituting the appropriate partial derivatives yields

\[
(\delta F_M)^2 = (\delta h)^2 \left( \frac{RF_M}{s} \right)^2 + (\delta s)^2 \left( \frac{RF_M}{q} \right)^2 + (\delta q)^2 \left( \frac{RF_M}{R_{\text{min}}} \right)^2
\]

\[
+ (\delta R_{\text{min}})^2 \left( \frac{RF_M}{R_{\text{min}}} \right)^2 + 2C_{hs}^2 \left( \frac{RF_M}{s} \right) \left( \frac{RF_M}{q} \right) + 2C_{hq}^2 \left( \frac{RF_M}{s} \right) \left( \frac{RF_M}{R_{\text{min}}} \right) + 2C_{qs}^2 \left( \frac{RF_M}{q} \right) \left( \frac{RF_M}{R_{\text{min}}} \right),
\]

where we have used Eqs. (7) and (11)–(13). After rearranging, the relative uncertainty in particulate backscattering becomes

\[
\frac{(\delta F_M)}{F_M} = \frac{(\delta h)}{s} \left( \frac{RF_M}{s} \right) + \frac{(\delta s)}{q} \left( \frac{RF_M}{q} \right) + \frac{(\delta q)}{R_{\text{min}}} \left( \frac{RF_M}{R_{\text{min}}} \right) + \frac{(\delta R_{\text{min}})}{R_{\text{min}}} \left( \frac{RF_M}{R_{\text{min}}} \right) + 2C_{hs} \left( \frac{RF_M}{s} \right) \left( \frac{RF_M}{q} \right) + 2C_{hq} \left( \frac{RF_M}{s} \right) \left( \frac{RF_M}{R_{\text{min}}} \right) + 2C_{qs} \left( \frac{RF_M}{q} \right) \left( \frac{RF_M}{R_{\text{min}}} \right).
\]

Similarly, the relative uncertainty in scattering ratio may be obtained as

\[
\frac{(\delta R)}{R} = \frac{(\delta h)}{s} \left( \frac{RF_M}{s} \right) + \frac{(\delta s)}{q} \left( \frac{RF_M}{q} \right) + \frac{(\delta q)}{R_M} \left( \frac{RF_M}{R_M} \right) + \frac{(\delta R_{\text{min}})}{R_{\text{min}}} \left( \frac{RF_M}{R_{\text{min}}} \right) + 2C_{hs} \left( \frac{RF_M}{s} \right) \left( \frac{RF_M}{q} \right) + 2C_{hq} \left( \frac{RF_M}{s} \right) \left( \frac{RF_M}{R_{\text{min}}} \right) + 2C_{qs} \left( \frac{RF_M}{q} \right) \left( \frac{RF_M}{R_{\text{min}}} \right).
\]

Thus, the total relative error in particulate backscattering, or in scattering ratio, contains contributions from signal-measurement error, two-way transmission error, density errors, and error in the assumed value of \( R_{\text{min}} \). In the following paragraphs we obtain more explicit expressions for these terms.

### A. Signal Measurement Error

From Eqs. (12) and (14) we have

\[
\frac{(\delta h)}{s} = \frac{(\delta N_S)}{N_S} \left( \frac{1}{N_S^*} + \frac{1}{N_S^*} \right), \quad \frac{1}{N_S^*} \neq z^*,
\]

\[
\frac{(\delta h)}{s} = \frac{(\delta N_S)}{N_S}, \quad \frac{1}{N_S^*} = z^*.
\]

where \( N_S^* \equiv N_S(\lambda,z^*) \). The result (21) occurs because the ratio (12) is identically equal to one at \( z = z^* \); equivalently, at this height the covariance term in Eq. (14) exactly cancels the two preceding terms. A covariance term does not appear explicitly in Eq. (20) if \( \delta N_S \) is taken to be the random part of the signal-measurement error. (We will follow this definition.) Additively systematic errors in the measured signal are canceled by the covariance term. The size of the random signal-measurement error is given by

\[
\frac{\delta N_S}{N_S} = \frac{(N_S + N_B + N_1)^{1/2}}{N_S},
\]
where \(N_S\) is given by Eq. (1),
\[
N_B = \text{number of charge carriers resulting from viewed background radiation, and}
\]
\[
N_I = \text{number of charge carriers resulting from internal (electronic) detector noise.}
\]

Equation (22) assumes that \(N_B\) and \(N_I\) are determined accurately enough that their uncertainty is negligible compared with that in Eq. (22); it also assumes that errors caused by imperfect transmit-receive alignment, gain switching, digitizing, or other signal processing are either negligible compared with Eq. (22) or of such a nature that they cancel in the ratio (12). If this is not the case, additional terms must be included in Eq. (22).

B. Transmission Errors

Equation (13) may be rewritten as
\[
q(\lambda,z,z') = \exp[-2r_A(\lambda,z,z') - 2TM(\lambda,z,z') - 2T_3(\lambda,z,z')],
\]
where \(r_A(\lambda,z,z')\) is the optical thickness of constituent \(x\) between heights \(z\) and \(z'\), and \(A, M, \) and 3 designate aerosol particles, molecular (Rayleigh) scatterers, and ozone (\(O_3\)), respectively. [If other absorbers are significant at the lidar wavelength, corresponding terms must be added to Eq. (23).] Since errors in determining these three optical thicknesses are uncorrelated, we have from Eq. (14)
\[
\delta q = 4\{[\delta r_A(\lambda,z,z')]^2 + [\delta TM(\lambda,z,z')]^2 + [\delta T_3(\lambda,z,z')]^2\}. \tag{24}
\]

We note that this equation shows correctly that, when the normalization procedure is used, transmission uncertainties in \(P_A(\lambda,z)\) result only from uncertainties in optical thickness between \(z\) and \(z'\) and not at any other heights.

The optical thickness uncertainties in Eq. (24) depend on typical atmospheric variability and the amount of care used in choosing an extinction model. Studies of ozone variability indicate that, over height ranges of 5 km or more near the stratospheric ozone peak, ozone column content can be predicted to within a standard error of about ±20% by using a model for the appropriate latitude, longitude, and month. Hence,
\[
\delta T_3(\lambda,z,z') = 0.2T_3(\lambda,z,z'). \tag{25}
\]
In height intervals where \(T_3(\lambda,z,z')\) is small (i.e., away from the O3 peak, or where \(|z - z'|\) is less than a few km), Eq. (25) probably underestimates \(\delta T_3\). However, in such regions \(\delta T_3\) is usually sufficiently small as to make a negligible contribution to the over-all uncertainty [Eq. (19)]. [Exceptions to this rule may occur in photochemically polluted boundary layers.]

Similar reasoning on aerosol transmission models and molecular density models (or measurements) yields the estimates
\[
\delta r_A(\lambda,z,z') = 0.5r_A(\lambda,z,z'), \tag{26}
\]
\[
\delta T_M(\lambda,z,z') = 0.1T_M(\lambda,z,z'). \tag{27}
\]
In cases where a molecular density measurement is made, Eq. (27) is a very conservative estimate; however, at wavelengths of 0.6 \(\mu\)m and longer, the result [Eq. (27)] is already negligibly small in comparison to other sources of error. The estimate [Eq. (26)] can sometimes be reduced by iteratively using the lidar data themselves and a careful assessment of the uncertainty in particulate backscatter-to-extinction ratio; however, we do not consider this procedure here.

C. Molecular Density Errors

Since the molecular (Rayleigh) backscattering coefficient per mass (or number) of molecules is known to high accuracy, the uncertainty in \(F_M(\lambda,z)\) results essentially from uncertainties in the molecular density profile \(D(z)\). At the location and time of a radiosonde measurement, molecular density can typically be determined with an uncertainty of less than 1\%.\(^{15,16}\) Other studies of spatial and temporal variability above the boundary layer\(^{18}\) indicate that, within about 100 km and 6 h of the radiosonde measurement, this uncertainty increases only to about 1\%, provided there are no intervening frontal air mass movements. (That is, \(\delta F_M/F_M = 0.01\).) When density models or interpolations must be used (i.e., above about 30 km and at great distances from a radiosonde), much larger errors can result unless great care is taken to use the most appropriate model or interpolation procedures. If these precautions are taken, the density errors can be limited to about 3\%.\(^{18}\) (That is, \(\delta F_M/F_M \approx 0.03\).) In sum,
\[
\begin{align*}
\delta F_M(\lambda,z) &\approx 0.01 \text{ (density measured),} \tag{28} \\
\delta F_M(\lambda,z) &\approx 0.03 \text{ (density modeled or interpolated).} \tag{29}
\end{align*}
\]

We note that Eqs. (10) and (18) correctly show that, when the normalization procedure is used, the lidar-derived particulate backscattering at height \(z\) includes uncertainties arising from molecular density uncertainties at two heights: the height of analysis \(z\), and the normalization height \(z^*\). Hence, reducing \(\delta F_M^*\) improves the accuracy of \(F_p\) at all heights; thus, if possible, normalization should be performed within the height range of a current radiosonde (or rocketsonde) measurement.

To evaluate Eq. (18) we must first evaluate the covariance term \(C_{F_p}^2\), for molecular backscattering values at two different heights. If the two heights, \(z\) and \(z^*\), differ by more than the height resolution \(\Delta z_D\) of the density model or measurement from which \(F_M\) is derived, and if the model or measurement is free of systematic error, \(C_{F_p}^2\) must vanish, since \(F_M\) and \(F_M^*\) are then derived from independent values and hence have uncorrelated errors. At the other extreme, if \(z\) and \(z^*\) are the same, \(C_{F_p}^2\) is given by [see Eq. (15)]
\[
C_{F_p}^2 = \lim_{n \to \infty} \frac{1}{n} \sum_{n=1}^{n} [F_M(\lambda,z^*) - F_M(\lambda,z*)]^2. \tag{30}
\]
which is by definition equal to \(\delta F_M^2(\lambda,z^*)\).\(^{14}\) For \(|z - z^*| \leq 0\) and \(\Delta z_D\), \(C_{F_p}^2\) must vary smoothly between \(\delta F_M^2\) and 0. From the form of Eqs. (15) and (30) it is reasonable to expect a quadratic dependence
on \(|z - z^*|\); hence, for our purposes we feel that a sufficiently accurate approximation is
\[ C^3_{p_3} = \begin{cases} (\delta F_{M*})^2 \left[ 1 - \left( \frac{|z - z^*|}{\Delta z_D} \right)^2 \right], & |z - z^*| \leq \Delta z_D, \\
0, & |z - z^*| \geq \Delta z_D, \end{cases} \]
which is exact in the limits \(z = z^*\), and \(|z - z^*| \geq \Delta z_D\). We note that, if the density model or measurement contains a systematic error, an additional term must be added to Eqs. (31) and (32). This term tends to cancel part of the \(\delta F_{p_3}\) term in Eq. (18), but the cancellation is not complete. This can best be seen by inspection of Eq. (9): if, for example, \(F_M\) is too small by 3% at all heights, derived \(F_r\) will also be too small by 3% at all heights, in spite of the cancellation in the factor multiplying \(R_{\text{min}}\). Such systematic errors do not, however, affect derived \(R\) values, as shown by Eq. (7).

As a test of the validity of Eq. (31) in the limit \(z = z^*\), it is useful to evaluate Eq. (18) for this case. This yields
\[ (\delta F_{A*})^2 = \left( \frac{F_{A*}}{F_{M*}} \right)^2 \left[ \frac{(\delta R_{\text{min}})^2}{R_{\text{min}}^2} + \frac{(\delta F_{M*})^2}{F_{M*}^2} \right], \tag{33} \]
[The two zeros in Eq. (33) follow from Eqs. (21) and (24).] The result [Eq. (34)] is the same as that obtained by first letting \(z = z^*\) in Eq. (10) and then applying Eq. (14). It shows correctly that, when the normalization procedure is used, the particulate backscattering derived at the normalization height \(z^*\) is not affected by errors in signal measurement or transmission but only by errors in the molecular density estimate at the normalization height and by errors in the assumed value \(R_{\text{min}}\).

\[ (\delta F_{A*})^2 = \left( \frac{\delta R_{\text{min}}}{R_{\text{min}}} \right)^2 + \left( \frac{\delta F_{M*}}{F_{M*}} \right)^2 = \left( \frac{\delta R_{\text{min}}}{R_{\text{min}}} \right)^2 + \left( \frac{\delta F_{M*}}{F_{M*}} \right)^2, \tag{34} \]

D. Errors in Assumed \(R_{\text{min}}\)

The problem of determining the best value for \(R_{\text{min}}(\lambda)\) is attacked by Russell and Livingston,\(^{12}\) hereafter referred to as RL. As noted by RL, \(R_{\text{min}}(\lambda)\) cannot be determined from conventional lidar measurements, because \(R_{\text{min}}(\lambda)\) is an assumed input to the conventional lidar data processing procedure. However, for the case where the lidar profile includes the upper troposphere and part of the stratosphere, there is a historical data set that allows the most probable value of \(R_{\text{min}}(\lambda)\) to be estimated to within a very useful uncertainty; the same data set can also be used to evaluate this uncertainty.

This data set is the large collection of particle number profiles measured by the University of Wyoming (UW) at eleven globally distributed sites over a period of several years.\(^{19}\) The sensor used by UW is a dustsonde (photoelectric single-particle counter) flown on a balloon. Each standard dustsonde flight measures vertical profiles of the number of particles larger than two size cutoffs (0.15-\(\mu\)m and 0.25-\(\mu\)m radius) as well as temperature and pressure profiles (from which a gas density profile can be computed). As shown by RL, these profiles of particle number—\(N_{0.15}(z), N_{0.25}(z)\)—and gas density—\(D(z)\)—can be converted to profiles of scattering ratio—\(R(\lambda,z)\)—by using an optical model that assumes a particle size distribution and a refractive index. Since the dustsonde data provide two pieces of information on particle number \([N_{0.15}(z)\) and \(N_{0.25}(z)\)], they can be used to constrain the model size distribution to be usefuly constrained. In this way RL have converted a large set of dustsonde-measured particle-number profiles to corresponding profiles of scattering ratio. In each case, the model size distribution was varied with height to agree with the measured \(N_{0.15}(z)\) and \(N_{0.25}(z)\).

The resulting profiles of \(R(\lambda,z)\) can be grouped by latitude band, height range, atmospheric condition (volcanic or nonvolcanic), and assumed refractive index; and individual profile minima, \(R_{\text{min}}(\lambda)\), can be determined. The average of \(R_{\text{min}}(\lambda)\) within each group then provides the expectation value of \(R_{\text{min}}(\lambda)\) for lidar measurements made under corresponding conditions; similarly, the standard deviation of \(R_{\text{min}}(\lambda)\) gives the probable \((1\sigma)\) variation about this expectation value.

Representative results obtained in this manner are shown in Table I(a). In each case, the expectation value \((R_{\text{min}}(\lambda))\) was obtained by taking the mean of all \(R_{\text{min}}(\lambda)\) values in the appropriate group of dustsonde-derived scattering ratio profiles. Similarly, \(\delta R_{\text{min}}(\lambda)\), the \(1\sigma\) deviation of \(R_{\text{min}}(\lambda)\) from \(\langle R_{\text{min}}(\lambda) \rangle\) to be expected in any given measurement, is the standard deviation about \(\langle R_{\text{min}}(\lambda) \rangle\) of all \(R_{\text{min}}(\lambda)\) in the group. Also shown for comparison is \(\delta R_{\text{min}}(\lambda)\), the probable deviation of \(R_{\text{min}}(\lambda)\) from the commonly assumed value of 1.0. This is the value that should be substituted for \(\delta R_{\text{min}}(\lambda)\) in Eqs. (18) and (19) if \(R_{\text{min}}(\lambda) = 1.0\) is assumed; comparison of columns (3) and (4) of Table I shows that considerable reduction in normalization error can be achieved by using \((R_{\text{min}}(\lambda))\) in place of unity.

Table I(b) presents an alternate way of describing \((R_{\text{min}}(\lambda)), \delta R_{\text{min}}(\lambda), \text{and } \delta R_{\text{min}}(\lambda), \text{i.e., in terms of } R_{\text{max}}(\lambda), \text{the maximum scattering ratio observed in a particular profile. This form has the advantage that, within broad bounds, virtually all dependence on wavelength and refractive index is contained in } R_{\text{max}}(\lambda), \text{which is provided by the lidar measurement itself [although some iteration may be required to derive a sufficiently accurate } R_{\text{max}}(\lambda) \text{from the lidar measurement]. Thus, Table I(b) can be used for wavelengths and refractive indices not included in Table I(a) [see Table I footnote for bounds on wavelength and refractive index; note also that refractive index must be essentially independent of height for either Table I(a) or I(b) to apply].
The latitude dependence of \( \langle R_{\text{min}}(\lambda) \rangle \) for nonvolcanic conditions shown in Table I is significant and discussed in more detail by RL. In brief, it arises in the following manner. During nonvolcanic conditions, at latitudes south of about 60° N, \( R_{\text{min}}(\lambda) \) usually occurs near or below the tropopause (within \(+2\) and \(-5\) km; note, however, the seasonal description by RL). The relative absence of light-scattering aerosols (on a particle-to-gas mixing ratio basis) in the upper troposphere in these latitudes evidently results from relatively large vertical or horizontal separation from aerosol source regions (especially land areas), or from effective scavenging by clouds and rain. (Different mechanisms, or combinations, are effective in different latitude bands.) However, the upper troposphere north of about 60° N is subject to greater aerosol contamination than are other latitudes because of the combined effect of lower tropopause heights and of northerly transport from the large land areas of the northern hemisphere. These land areas, which are much larger than the southern hemisphere land areas, contribute both man-made aerosols and desert dust.\(^\text{20}\) Thus, relatively large aerosol mixing ratios in the upper troposphere in northern high latitudes result in correspondingly large \( \langle R_{\text{min}}(\lambda) \rangle \).

We note briefly the choice of refractive indices and height regions shown in Table I. The lower height cutoff, \( z = 5.0 \) km, is convenient for lidar stratospheric aerosol measurements, while still including the clean region usually found in the upper troposphere. The upper height cutoff, 24–34 km, is imposed by the dustsonde data and is also appropriate for many lidar stratospheric aerosol measurements. (Note that normalization of lidar data above 30 km is not usually recommended because there is usually a lack of radiosonde density data there; see Sec. III.C. Further, there can be significant amounts of dust above 30 km, especially in tropical regions.) The refractive index \( m = 1.425–0.0i \) applies to a mixture of 75% sulfuric acid and 25% water, a composition revealed by several measurements and models of the nonvolcanic stratospheric aerosol.\(^\text{21–23}\) (Actually, the relative proportions of sulfuric acid and...
water are expected to vary with height, but the resulting
effect on derived stratospheric scattering ratios has been
shown to be negligibly small.)

Thus, the values of \( R_{\text{min}}(\lambda) \), \( \delta R_{\text{min}}(\lambda) \), and \( \delta R_{\text{min}}(\lambda) \) given in Table I can be used to assess normalization
errors for nonvolcanic lidar measurements that cover
the height range from 5 km to about 29 km. For lidar

\[
\begin{array}{|c|c|c|}
\hline
\text{Table II. Parameters of Airborne Lidar Used for Simulations} \\
\hline
\text{Transmitter} & \text{Ruby} & \text{Nd:YAG} \\
\hline
\text{Wavelength (\( \mu \text{m} \))} & 0.6943 & 1.06 \\
\text{Energy per pulse (J)} & 1.0 & 0.5 \\
\text{Rep rate (ppps)} & 1.0 & 20 \\
\text{Pulse width (nsec)} & 30 & 20 \\
\text{Beam divergence (mrad)} & 1.0 & 1.0 \\
\text{Beam diameter (cm)} & 8 & 7.6 \\
\hline
\text{Receiver} & & \\
\text{Diameter (cm)} & 36 & 36 \\
\text{Field of view (mrad)} & 2 & 2 \\
\text{Filter bandwidth (A)} & 0.35 & 0.35 \\
\text{Optical eff. to PMT} & 0.10 & 0.02 \\
\text{PMT quant. eff.} & & \\
\hline
\end{array}
\]

IV. Application and Measurement Simulation

In this section we demonstrate the use of the expressions
derived above to calculate error bars (uncertainties)
for particular experimental situations. For this purpose
we assume the lidar system parameters
shown in Table II. The values shown are the design
parameters of a new airborne lidar being used to make
relative measurements of the stratospheric aerosol
in support of two satellite experiments: SAM II
(Stratospheric Aerosol Measurement, flown on the
Nimbus-7 satellite) and SAGE (Stratospheric Aerosol
and Gas Experiment, flown on the AEM-B satellite).
(For details of both experiments see Refs. 28–31.)

The lidar was originally built with a ruby laser (\( \lambda = 0.69 \mu \text{m} \))
and corresponding detector. A Nd:YAG laser (\( \lambda = 1.06 \mu \text{m} \))
and corresponding detector are to be added in the
near future. One purpose of the calculations shown
here was to determine whether the Nd:YAG system
should be used to yield smaller uncertainty in derived
stratospheric particulate backscatter coefficients
than did the ruby system. Since SAM II and SAGE are

![Fig. 1. Simulation procedure for evaluating lidar measurement errors. Circles symbolize random number generators that inject simulated errors into derived quantities at appropriate steps of the computation. Numbers in parentheses indicate equations in the text.](image-url)
Table III. Assumed Sizes of Error Sources in Lidar Data Analysis

<table>
<thead>
<tr>
<th>Source</th>
<th>Relative uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detector signal $N_S$</td>
<td>$\frac{\delta N_S}{N_S} = \frac{(N_S + N_B + N_f)^{1/2}}{N_S}$</td>
</tr>
<tr>
<td>Two-way transmission $q$</td>
<td>$\frac{\delta q}{q} = 2 [(0.2 \tau_2)^2 + (0.5 \tau_A)^2 + (0.1 \tau_M)^2]^{1/2}$</td>
</tr>
<tr>
<td>Molecular density $D$</td>
<td>$\frac{\delta D}{D} = 1%$ below 30 km and $3%$ above 30 km</td>
</tr>
<tr>
<td>Minimum scattering ratio $R_{\text{min}}$</td>
<td>See Tables I and II and Russell and Livingston (1979)</td>
</tr>
</tbody>
</table>

Notes: All error sources are explained more fully in Sec. III. $N_B$, $N_f$ = Detector output resulting from background light and internal noise, respectively. $N_B$ was calculated assuming the receiver characteristics of Table II and a skylight radiance of $2 \times 10^{-4}$ W (m$^2$ sr A)$^{-1}$ and $1.3 \times 10^{-6}$ W (m$^2$ sr A)$^{-1}$ at 0.69 $\mu$m and 1.06 $\mu$m, respectively. This radiance applies to a zenith-viewing lidar flying near 4 km, with the sun near the horizon and a receiver polarization analyzer aligned perpendicular to the skylight polarization. $N_f$ was obtained from photomultiplier specifications and is small compared with $N_B$. 

$\tau_2, \tau_A, \tau_M$ = One-way optical thickness of ozone, aerosol particles, and molecular scattering, respectively, between normalization altitude and altitude of analysis.

* Assumes radiosonde density profile available within about 100 km and 6 h of lidar measurement and no intervening frontal activity (see Sec. III and Ref. 18 for further discussion).

Expected to determine particulate extinction (at $\lambda = 1.0$ $\mu$m) to an accuracy of 10-15% wherever particulate 1-$\mu$m extinction exceeds approximately $5 \times 10^{-5}$ km$^{-1}$, a useful design goal for the lidar is a 10-15% uncertainty in derived particulate backscattering in such regions (roughly speaking, all heights below 20-30 km, depending on latitude).

As a check on the validity of the algebraic expressions for uncertainty, we have also simulated the lidar measurement and data analysis procedure. That is, in addition to calculating expected signal levels, we have used random number generators to inject errors of the appropriate size at appropriate steps of the measurement and analysis sequence. Figure 1 shows the flow of the calculations and the points at which random errors are injected. These random numbers simulate errors in signal measurement, density measurement, and transmission estimation; the standard deviation of the probability distribution from which each random number is drawn is shown in Table III. [Compare Eqs. (20)-(22), (24)-(27), and (28)-(29); see also Appendix B for the details of simulating density and transmission errors.] Note that no random error is added to simulate the error in assumed $R_{\text{min}}$; this error is set in each simulation by the choice of a model scattering ratio profile [which has an $R_{\text{min}}$ value that in general differs from the value ($R_{\text{min}}(\lambda)$) assumed in the analysis and listed in Table I].

The model atmosphere used for the first simulations is shown in Fig. 2. The molecular density and particulate extinction profiles were derived from a dustsonde measurement (designated B-10) made at Pt. Barrow, Alaska, in November 1973, during nonvolcanic conditions. This location and date were chosen because the lidar is to be used frequently at high latitudes (to which the SAM II measurements are confined) and
because nonvolcanic conditions are highly likely and pose a difficult test of lidar measurement accuracy. The ozone profile is from an average of winter highlatitude observations.\textsuperscript{15} Figure 3 shows ruby ($\lambda = 0.69$-\(\mu\)m) and Nd ($\lambda = 1.06$-\(\mu\)m) profiles of scattering ratio $R$ derived from the B-10 dustsonde measurement by fitting a two-parameter size distribution to the two-channel particle-number data and assuming a composition of 75\% H$_2$SO$_4$ and 25\% H$_2$O (refractive index 1.42–0i). Three points of Fig. 3 are worthy of note. First, the major peak in scattering ratio occurs at about 17 km, which is rather low compared with typical mid-latitude peak heights (~21 km). (Decreasing peak height with increasing latitude is a normal characteristic of the stratospheric aerosol.)\textsuperscript{19} Second, because of this low height and the denser air there, the peak value of $R - 1$ for each wavelength is only about two-fifths as large as it would be if the peak were at the typical midlatitude height of 21 km. And, third, all values of $R - 1$ for the Nd profile are about three times as large as for the ruby profile. (This occurs because of the $\lambda^{-4}$ dependence of Rayleigh backscattering and the weaker dependence of stratospheric particulate backscattering, which can be approximated as $\lambda^{-1}$ to $\lambda^{-1.5}$ in the wavelength region considered here.) As will be seen, this latter factor tends to make the Nd measurement considerably more accurate than the ruby measurement, at least for the lidar parameters assumed in Table II and for high-latitude nonvolcanic model atmospheres.

Figure 4 shows the expected scattering-ratio error bars derived from Eq. (19) for the assumed ruby and Nd lidar parameters and the B-10 model atmosphere. The calculations assume that both lidars are flying at a height of 4 km, a skylight background (Table III) appropriate for zenith viewing from 4 km with the sun on the horizon (as it is during SAM II and SAGE scans), and an integration time of 1 min. Vertical range bins of 0.25 km are used, except at high altitudes where combining bins is necessary to reduce signal-measurement uncertainty.

Figure 5 shows the corresponding errors in particulate backscattering coefficient derived from Eq. (18). In addition, the dots show simulated measurements with all four sources of random error included, as a check on the error-bar computation. The scatter of simulated data points is approximately the same as the range of error bars, indicating that the error-bar computation is at least approximately correct. However, a closer look at the model (solid line), the simulated measurements (dots), and the error bars reveals two noteworthy points. First, it can be seen that the error bars actually span a range that is slightly greater than the ±1σ scatter of data points. This occurs because the error bars include a contribution caused by uncertainty in the value of $R_{\text{min}}$; error in the assumed $R_{\text{min}}$ does not introduce scatter among different range bins but instead affects all range bins systematically, as shown by Eqs. (7) and (9) (see also the second paragraph in this section). Second, the simulated data points for each wavelength are systematically underestimated, such that most data points are less than the model. This type of systematic underestimation originates at the normalization height $z^*$ and can have one of four causes, as shown by Eqs. (7) and (9): a positive error in $N_S(\lambda,z^*)$; a negative error in $F_M(\lambda,z^*)$; a negative error in the assumed $R_{\text{min}}(\lambda)$; and an error in the extinction model that is positive between the lidar and $z^*$ and negative beyond $z^*$. Close inspection of the random errors used in this simulation revealed that most of the systematic underestimation in Fig. 5 was caused by the second possibility mentioned above, namely, a relatively large negative error (~1.1\%) in molecular density at the normalization height ($z^* \approx 28$ km). We note these points to emphasize some of the
subtleties in lidar-error analysis and the value of introducing random errors into simulated measurements as a way of checking the validity of analytical error expressions.

It can be seen from Fig. 5 that the relative errors in the Nd backscattering coefficient profile are about half as large as the relative errors for the ruby profile. The relative size and the height dependence of these errors can be seen more clearly in Fig. 6, which also shows how the four sources of uncertainty (density, transmission, signal measurement, and $R_{\text{min}}$) contribute to the total. Identifying the sources of uncertainty aids in roughly predicting the effect of varying certain lidar parameters. Note, for example, that the major source of uncertainty for the ruby measurement, at all altitudes, is the uncertainty in the molecular density profile. This uncertainty cannot be reduced by using more laser energy, a larger receiver, a more efficient detector, longer counting, coarser vertical resolution, or flying closer to the aerosol peak. In fact, the only way of reducing this uncertainty is by determining the relative molecular density profile in the lidar-measurement range to an accuracy of better than 1%. The measurement errors and horizontal drift of radiosondes make reduction of this error very unlikely if radiosonde measurements are the source of density information. However, there is hope that simultaneous measurements with a frequency-doubled ruby transmitter ($\lambda = 0.347 \, \mu\text{m}$) could provide significantly reduced density errors, especially when an iterative scheme is used to correct for particulate backscattering in the frequency-doubled channel.

As shown by Fig. 6, density errors are also the dominant source of error for the Nd backscattering measurement, at least below 25 km. (These errors could possibly be reduced by using the frequency-tripled Nd:YAG output—$\lambda = 0.355 \, \mu\text{m}$—to measure density.) Above this height, signal-measurement errors do become significant, and the expected total uncertainty is sensitive to transmitted power, detector efficiency, aircraft height, vertical resolution, and received signal integration time. In this connection we note that Nd signal-measurement errors would be reduced somewhat by flying at a higher altitude, say 8 km. However, we do not recommend this procedure, because normalization errors increase significantly if the lidar profile excludes the height of minimum mixing ratio, which is frequently found several km below the tropopause. (See RL. The minimum mixing ratio does not occur below the tropopause in the B-10 model—possibly because of the upper tropospheric dust discussed by RL and in Sec. III.D—and it usually does not occur there after volcanic eruptions. However, analyses of dustsonde data have shown that the minimum mixing ratio does tend to occur below the tropopause for nonvolcanic conditions in most latitude bands.)

Figure 7 shows calculated uncertainties and simulated measurements for P-10 (tropical nonvolcanic) model atmosphere. This model is tropical and nonvolcanic, based on a tropical ozone model and on dustsonde and molecular density measurements made in Panama in November 1973 by the University of Wyoming. As expected, the
Aerosol peak is at a higher altitude (23 km) than the peak in the Arctic nonvolcanic model (17 km). Partially as a result of this greater height and the smaller molecular density there, maximum scattering ratios are also greater (~1.13 (ruby) and 1.35 (Nd), as compared with 1.09 (ruby) and 1.23 (Nd)). Because of the larger scattering ratios, density-induced uncertainties at the peak height are smaller at both wavelengths than they were for the Arctic nonvolcanic model (Fig. 6). However, because of the greater peak height, signal-measurement errors do become significant within the peak. As with the Arctic nonvolcanic case, density-induced errors are smaller at the Nd wavelength (1.06 m) than they are at the ruby wavelength (0.69 m). Note that uncertainties arising from the density ratio are larger at 1.06 m than at 0.69 m, reflecting the larger density ratio uncertainties shown in Table I. However, at nearly all heights the total uncertainty for the ruby (0.69-m) measurement exceeds that for the Nd (1.06-m) measurement, because of the strong effect of density-induced uncertainties.

Another point that bears special mention is the large uncertainty in particulate backscattering below 17 km shown in Figs. 7(b)–7(d). The large relative uncertainties result from the small scattering ratios in this region. [Hence, the leading \( F_M/F_A \) factor in Eq. (18) is large.] However, because molecular density below 17 km is large compared with that at the peak height (23 km), even the small scattering ratios below 17 km [Fig. 7(a)] yield absolute particulate backscattering [Fig. 7(b)] comparable with that at the height of the mixing ratio peak. Thus, large relative uncertainties in particulate backscattering below 17 km correspond also to large absolute uncertainties. The net result of this is that vertically integrated backscatter (often used as an indicator of column stratospheric aerosol behavior) can be rather accurately measured (on both an absolute and a relative basis) over the 17–30-km height range (in this example); however, over the 10–30-km height range, the measurement is relatively inaccurate, because the inaccurate values from 10 km to 17 km contribute a major part of the 10–30-km integral. That is, the low-altitude values can strongly contaminate the integral, even though their scattering ratios are small. Thus, the quality of a lidar-derived backscattering column content is a very strong function of the height limits used for the column. Optimum height limits can be revealed by using standard error-propagation procedures to compute the effect on \( \int F_A(z)dz \) of the \( \Delta F_A(z) \) values given here. However, without an explicit error analysis these very important subtleties can easily be missed. This could lead, for example, to the reporting of poor-quality data products when a choice of different height limits could result in high-quality data products.

As a final example, Fig. 8 shows results for a third model, representing mid-latitude moderate volcanic conditions. (The model is based on lidar, rawinsonde, and dustsonde measurements made at Kansas City, Missouri, July 1975, eight months after the major 1974 eruption of Volcán de Fuego in Guatemala.) Because of the volcanic particles, scattering ratios are much larger than in either of the previous two models. As a result, density-induced errors and signal-induced errors are greatly reduced on a relative basis. Transmission-induced errors are strongly dependent on height because of the competing effects of increasing scattering ratios [reflected through the \( F_M/F_A \) factor in Eq. (18)] and increased aerosol particle optical thickness [the \( \tau_A \) factor in Eqs. (24) and (26)]. The loss of the normally clean upper troposphere for normalization has produced a significant increase in the uncertainty caused by uncertainty in \( F_M/F_A \), as shown by RL. In fact, uncertainty in \( R_{\min} \) has become the dominant source of uncertainty for both wavelengths at most heights of interest. Because of this, and because \( \Delta R_{\min} \), is larger for wavelength 1.06 m than it is for 0.69 m (Table I), the ruby measurement is more accurate at most heights than it is for the Nd measurement. This is in strong contrast with the situation for nonvolcanic conditions (those most common), as shown by comparison of Figs. 8(c) and 8(d) with Figs. 6, 7(c), and 7(d). Note, however, that for most heights and both wavelengths, the...
total relative error for volcanic conditions is less than for nonvolcanic conditions. This is because of the over-all $F_M/F_A$ factor in Eq. (18).

V. Conclusions

We have presented a methodology for analyzing the error in lidar aerosol measurements that use the normalization technique for lidar calibration. Although the method was demonstrated for stratospheric aerosol measurements, it is also applicable to tropospheric aerosol measurements and yields usefully small errors, provided the lidar data include the clean region usually found in the free troposphere (see RL). We have derived algebraic expressions for uncertainty and tested them by introducing random errors in simulated measurements and analyses.

The error in lidar-derived particulate backscattering coefficient includes contributions from errors in (1) lidar signal, (2) molecular density, (3) atmospheric transmission, and (4) lidar calibration (which is determined by the assumed minimum scattering ratio $R_{\text{min}}$ when the normalization technique is used for lidar calibration). Any lidar experiment designed to measure aerosol properties should examine the effects of these four error sources in a sensitivity analysis of what accuracies are possible in a proposed lidar experiment. For example, it is seen from Eq. (18) that basic to the error analysis of the aerosol backscatter is the ratio of particulate backscatter to molecular backscatter. As this ratio increases, as is the case for the Nd wavelength (compared with ruby or dye wavelengths) or for the case of volcanic conditions where large amounts of aerosol are present in the stratosphere, the total relative error in determining the particulate backscatter decreases.

This is an important design consideration for lidar experiments intending to measure stratospheric aerosols. Various trade-off studies can be performed such as the ones illustrated above weighing wavelength selection vs detector efficiency and so forth.

Furthermore, the analysis of aerosol backscatter data from lidar experiments frequently includes a column (vertically integrated) backscattering quantity, which is used to describe the time history of the stratospheric aerosol loading. The uncertainty in column backscatter depends strongly on column height limits and aerosol mixing ratio (see discussion of Fig. 7). Erroneous values for derived atmospheric parameters such as stratospheric aerosol decay rates could be deduced if the larger uncertainties in the column aerosol backscatter over certain height regions and during periods of low volcanic activity were not properly acknowledged. Therefore, we recommend a detailed error sensitivity analysis such as described for all lidar aerosol measurements that are used to derive such implications.

We are grateful to John Livingston for assistance in programming and computer runs and to James Laver, Roderick Quiroz, and Fred Finger for advice on density measurement errors.

Appendix A: Estimating the True Minimum from a Profile of Noisy Data

We assume that there exists a true vector (profile) $x$ of the quantity of interest and that the measurements yield a vector $\hat{x}$ that is composed of the true vector and an error vector $\epsilon$ such that

$$\hat{x} = x + \epsilon.$$  \hspace{1cm} (A1)

Each error $\epsilon_i$ is assumed to be drawn from a normal distribution with mean zero and standard deviation (or probable error) $\sigma_i$. After a given measurement, $x$ and $\epsilon$ are unknown, and only $\hat{x}$ and the probable error vector $\sigma$ are known. (Appendix A.E describes how $\hat{x}$ and $\sigma$ are obtained in our simulations.) From these two vectors, $\hat{x}$ and $\sigma$, we wish to obtain the best estimate of $x_{\text{min}}$, the minimum in the true vector $x$.

Simply choosing the minimum measured element, $\hat{x}_{\text{min}} = \hat{x}_m$, as the best estimate of $x_{\text{min}}$ is unsatisfactory if the vertical spacing of the height vector $z$ (elements $z_i$) is small compared with the usual vertical spacing of significant aerosol backscatter variations in the height region of interest. (Roughly speaking, a significant backscatter variation is one that exceeds $\sigma$ in the height region of interest.) This is often the case with lidar vertical profiles, and, when it is, the reason for a particular element $\hat{x}_m$ being less than all other elements $\hat{x}_i$ is likely to be that $\hat{x}_m$ contains a large negative error $\epsilon_m$ [Eq. (A1)], rather than because the true element $x_m$ is less than all other $x_i$. Under such circumstances, choosing $\hat{x}_{\text{min}}$ as the best estimate of $x_{\text{min}}$ is likely to introduce significant underestimation, that is, $x_{\text{min}}$ is probably less than $\hat{x}_{\text{min}}$.

This problem of obtaining the best estimate of the true minimum in a noisy profile of scattering ratios has been confronted as an integral part of lidar data analysis by many groups over the past years. We have been unable to find a rigorous solution to the problem, even after consulting statisticians. Nevertheless, because of the importance of the problem, we present here a method (weighting data points by the normal error integral) that is somewhat satisfying heuristically and that makes use of the two available sets of information, $\hat{x}$ and $\sigma$. As an introduction to the method, we briefly review two other methods that have been used in the past.

A. Degrading the Vertical Resolution

In this method, groups of adjacent bins of lidar data (elements $\hat{x}_i$) are averaged together, and the minimum of the resulting coarse-resolution profile is taken as the best estimate of $x_{\text{min}}$. The rationale is that the errors $\epsilon_i$ within each group are likely to differ in sign and to tend to average to zero (i.e., a central limit is approached). Thus, one is less susceptible to the danger of finding a profile minimum that is a minimum primarily because of a large negative error. This technique has been used previously with a coarse vertical resolution of 0.5 km. (The original fine vertical resolution is retained in all subsequent steps of analysis after normalization by the coarse-profile minimum.) A vertical
resolution coarser than 0.5 km is not used because of the danger of averaging over significant aerosol vertical structure.

This method has the advantage of simplicity but also has several shortcomings. First, if the errors $\epsilon_i$ are correlated over the averaging distance, the error in the coarse profile does not approach zero. [This is the case in height regions where density errors are the primary source of error, and the density measurements have a vertical spacing of 1 km or more. Note, for example, the ripple structure of the dots below 17 km in Figs. 7(a) and 7(b).] Second, no explicit use of the probable error vector $\sigma$ is made, even though one knows that larger values of $\sigma_i$ warrant more averaging. (That is, averaging is warranted whenever $\sigma_i$ is large compared with $(x_i - x_{i+1})$, the expected vertical variation in true aerosol content.)

B. Averaging of Statistically Equivalent Data Points

An alternative method avoids some of the above difficulties by making use of the probable-error vector $\sigma$. This method begins by stating that all $\tilde{x}_i$ for which

$$\tilde{x}_i - \sigma_i \leq \tilde{x}_m$$  \hspace{1cm} \text{(A2)}

are, in an approximate sense, statistically equivalent to $\tilde{x}_\text{min}$. (The index of $\tilde{x}_\text{min}$ is $m$.) That is, for each such $\tilde{x}_i$, the difference, $\tilde{x}_i - \tilde{x}_\text{min}$, is less than the combined uncertainty in the two data points. The best estimate of $x_{\text{min}}$ is then taken to be the weighted mean of all data points in this equivalent set. Each data point is weighted by its information content, i.e., the inverse square of its probable error. Hence, one uses for the best estimate of $x_{\text{min}}$

$$\frac{\sum x_i w_i}{\sum w_i},$$  \hspace{1cm} \text{(A3)}

where

$$w_i = 1/\sigma_i^2,$$  \hspace{1cm} \text{(A4)}

and the primed summations are restricted to all $i$ that satisfy (A2). This procedure has been used at SRI International (formerly Stanford Research Institute) for many years. It has the advantages that: The number of averaged data points tends to increase as the relevant $\sigma_i$ increase.

Even errors that are correlated on a fine vertical scale (e.g., density errors) tend to average to zero, because the set (A2) often extends over a height range of several km and may even be split into several regions separated by many km.

The weighting (A4) insures that the mean is most strongly influenced by the least error-prone data points. Data points $\tilde{x}_i$ that appear to exceed $\tilde{x}_\text{min}$ because of obvious aerosol structure are automatically excluded from the sum (A3) because of (A2). (This last exclusion does not automatically occur when the method in App. A.A is used.)

This method thus achieves a vertical resolution that is influenced both by the probable errors $\sigma_i$ and the apparent aerosol structure revealed by the measurements $\tilde{x}_i$. Nevertheless, it is unsatisfactory in two ways. First, the cutoff defined by (A2) is somewhat arbitrary. [It is difficult to say why the $\sigma$'s in (A2) should not be replaced by $0.67\sigma$, $2\sigma$, $3\sigma$, or some other common measure of probable error.] Second, the weighting (A4) can (and often does) give equal weight to $\tilde{x}_\text{min}$ and to other data points $\tilde{x}_i$ that exceed $\tilde{x}_\text{min}$ by as much as $\sigma_i + \sigma_m$, whereas it seems clear that the probability of $\tilde{x}_\text{min}$ being close to the true $x_{\text{min}}$ is greater than the probability of $\tilde{x}_i (= \tilde{x}_\text{min} + \sigma_i + \sigma_m)$ being close to $x_{\text{min}}$. (That is, one seeks a procedure that assigns decreasing weight as $\tilde{x}_i - \tilde{x}_\text{min}$ increases.)

C. Weighting Data Points by the Normal Error Integral

As an introduction to this method, we recall that each measured data point $\tilde{x}_i$ is assumed to be normally distributed with mean equal to the true value $x_i$ and standard deviation equal to $\sigma_i$. In this method, the estimate of the true-profile minimum is again taken as the weighted mean of a set of measured data points, but the weight given to point $\tilde{x}_i$ is now taken as $P(\tilde{x}_i - \tilde{x}_m | x_i \leq x_m)$, that is, the conditional probability of obtaining the measured difference $\tilde{x}_i - \tilde{x}_m$, under the condition that the true value $x_i$ is less than or equal to the true value $x_m$. Hence, the best estimate for $x_{\text{min}}$ is taken to be

$$\frac{\sum x_i w_i}{\sum w_i},$$  \hspace{1cm} \text{(A5)}

where

$$w_i = P(\tilde{x}_i - \tilde{x}_m | x_i \leq x_m).$$  \hspace{1cm} \text{(A6)}

The conditional probability (A6) may be written as an integral,

$$P(\tilde{x}_i - \tilde{x}_m | x_i \leq x_m) = k \int_0^\infty P(\tilde{x}_i - \tilde{x}_m | x_i = x_m - h) dh,$$  \hspace{1cm} \text{(A7)}

where $k$ is a normalization constant to assure that all probabilities sum to unity. By a shift in the x-axis origin we have

$$P(\tilde{x}_i - \tilde{x}_m | x_i = x_m - h) = P(\tilde{x}_i - \tilde{x}_m + h | x_i = x_m).$$  \hspace{1cm} \text{(A8)}

The conditional probability [Eq. (A8)] may be obtained by using a theorem from sampling theory which states that, when two variables, $\tilde{x}_i$ and $\tilde{x}_m$, are normally and independently distributed with equal means [or true values, the condition stated on the right side of Eq. (A8)], the difference, $\tilde{x}_i - \tilde{x}_m$, is also normally distributed, with mean zero and standard deviation $\sigma = (\sigma_i^2 + \sigma_m^2)^{1/2}$. Hence we have

$$P(\tilde{x}_i - \tilde{x}_m + h | x_i = x_m) = f((\tilde{x}_i - \tilde{x}_m + h)/(\sigma_i^2 + \sigma_m^2)^{1/2}),$$  \hspace{1cm} \text{(A9)}

where $f(y)$ is the normal distribution function,

$$f(y) = \frac{1}{(2\pi)^{1/2}} \exp(-y^2/2),$$  \hspace{1cm} \text{(A10)}

$$y = (\tilde{x}_i - \tilde{x}_m + h)/\sigma.$$  \hspace{1cm} \text{(A11)}

Using Eqs. (A8)--(A11) in Eq. (A7) then yields

$$P(\tilde{x}_i - \tilde{x}_m | x_i \leq x_m) = ke \int_{(\tilde{x}_i - \tilde{x}_m)/\sigma}^\infty f(y) dy.$$  \hspace{1cm} \text{(A12)}
Hence, via Eq. (A6),
\[ w_i = E[(\hat{x}_i - \hat{x}_m)/\sigma], \tag{A13} \]
where \( E(a) \) is the (normalized) normal-error integral from \( a \) to infinity. (Note that the normalization factor \( k \) is absorbed into \( E \) by virtue of \( E \)'s normalization.)

This method avoids the problem of the arbitrary cutoff in the previous method and achieves a weight that decreases as \( \hat{x}_i - \hat{x}_m \) increases. It does sacrifice the advantage of assigning decreasing weight to points with large \( \sigma_i \), but we feel that this is not so important as achieving a smoothly decreasing weight as \( \hat{x}_i - \hat{x}_m \) increases. This method of estimating \( x_{\text{min}} \) was used for the simulations shown in this paper.

D. Which Method is Best?

We emphasize that none of the above methods, including the last, has been rigorously demonstrated to be correct. We feel that the last method has the most desirable features, that it is somewhat satisfying heuristically, and that it has some support in classical sampling theory. Nevertheless, selection of the optimum method should be empirical, based on simulations that test which method consistently gives the best results in finding a known \( x_{\text{min}} \). Our efforts to date have provided some guidance in this regard, but far more simulations, aimed at this specific question, are required before one particular method can be recommended unequivocally.

E. Values Used for \( \hat{x} \) and \( \sigma \) in Practice

As noted in Sec. II, the search for the minimum measured scattering ratio is made before the profile has been normalized in terms of \( (R_{\text{min}}(\lambda)) \). Hence, \( \hat{x} \) is given by Eq. (5), with an arbitrary value for \( K(\lambda) \), and model or measured (i.e., inexact) values are given for \( N_S \), \( Q^2 \), and \( F_M \). More explicitly,
\[
\hat{x}_i = \hat{R}(\lambda, z_i) - \frac{(z_i - z_L)^2}{(z_i - z_L)^2} N_S(\lambda, z_i) \frac{K(\lambda) Q^2(\lambda, z_i) F_M(\lambda, z_i)}{K(\lambda) Q^2(\lambda, z_i) F_M(\lambda, z_i)}. \tag{A14}
\]

Since any value of \( K \) can be used in Eq. (A14) and since this value will later be canceled in the normalization, only uncertainties in \( N_S \), \( Q^2 \), and \( F_M \) contribute to \( \sigma \), the uncertainty that is of concern in estimating the minimum of the true vector \( \hat{x} \) approximated by Eq. (A14). Hence, by Eq. (A14),
\[
\sigma_i = \delta \hat{x}_i = \frac{\left[ \frac{\delta N_S(\lambda, z_i)^2}{N_S(\lambda, z_i)} + \frac{\delta Q^2(\lambda, z_i)^2}{Q^2(\lambda, z_i)} \right]^{1/2}}{F_M(\lambda, z_i)} \tag{A15}.
\]

Appendix B: Simulating Density and Transmission Errors

In analyzing an actual lidar profile, the density errors made at adjacent heights are usually correlated to some extent, as are the transmission errors. For density errors this correlation occurs because the density profile has typically been measured (or modeled) at a vertical resolution coarser than that of the lidar data, and values at the lidar data-bin heights \( (z_i) \) are typically obtained by interpolation between the density-measurement heights \( (z_j) \). For transmission errors this correlation occurs because a single optical-thickness profile is guessed, and the guessed optical thickness at all heights between the lidar and the height of analysis influences the transmission value used at the analysis height \( z_i \). For these reasons it is not realistic to simulate either density or transmission errors by drawing an independent sample from an error distribution at each lidar analysis height \( z_i \). In our simulations we have used the following procedures to simulate these correlation effects.

The coarse-density profile was given a vertical resolution of 1 km below 30 km (to simulate a radiosonde measurement) and 5 km above 30 km (to simulate a model). At each coarse height the density \( D(Z_j) \) was set equal to the exact model value \( D(Z_j) \) plus a random error \( \epsilon_D(Z_j) \) drawn from a distribution with mean zero and standard deviation given by Eq. (28) or Eq. (29), that is,
\[
D(Z_j) = D(Z_j) + \epsilon_D(Z_j). \tag{B1}
\]

Errors at adjacent coarse heights, \( Z_j, Z_{j+1} \), are drawn independently of each other. Density values at the lidar analysis heights \( z_i \) are then obtained by interpolation between values of the coarse profiles \( D(Z_j) \). The resulting profile of density errors is correlated over distances \( |z - z'| \) that are less than the coarse spacing \( |Z_j - Z_{j+1}| \). This results in the ripple structure with about 1-km vertical period that can be seen in Fig. 7(a) between 5 km and 15 km.

The choice of an incorrect transmission profile was simulated in the following manner. At the height of analysis farthest from the lidar \( z_{\text{far}} \), the optical thickness \( \tau \) of each constituent was set equal to the exact model value \( \tau(z_{\text{far}}) \) plus a random error drawn from a distribution with mean zero and standard deviation given by Eq. (25), (26), or (27), as appropriate. That is,
\[
\tau(z_{\text{far}}) = \tau(z_{\text{far}}) + \epsilon(z_{\text{far}}). \tag{B2}
\]

where \( \tau \) applies only to a single constituent in this equation. At every other height the error was set to a fraction of \( \epsilon(z_{\text{far}}) \) that increased in proportion to optical thickness. That is,
\[
\tau(z) = \tau(z) + \frac{|\tau(z) - \tau(z)|}{|\tau(z_{\text{far}}) - \tau(z)|} \epsilon(z_{\text{far}}). \tag{B3}
\]

References