

Lidar inversion of atmospheric backscatter and extinction-to-backscatter ratios by use of a Kalman filter

Francesc Rocadenbosch, Cecilia Soriano, Adolfo Comerón, and José-María Baldasano

A first inversion of the backscatter profile and extinction-to-backscatter ratio from pulsed elastic-backscatter lidar returns is treated by means of an extended Kalman filter (EKF). The EKF approach enables one to overcome the intrinsic limitations of standard straightforward nonmemory procedures such as the slope method, exponential curve fitting, and the backward inversion algorithm. Whereas those procedures are inherently not adaptable because independent inversions are performed for each return signal and neither the statistics of the signals nor *a priori* uncertainties (e.g., boundary calibrations) are taken into account, in the case of the Kalman filter the filter updates itself because it is weighted by the imbalance between the *a priori* estimates of the optical parameters (i.e., past inversions) and the new estimates based on a minimum-variance criterion, as long as there are different lidar returns. Calibration errors and initialization uncertainties can be assimilated also. The study begins with the formulation of the inversion problem and an appropriate atmospheric stochastic model. Based on extensive simulation and realistic conditions, it is shown that the EKF approach enables one to retrieve the optical parameters as time-range-dependent functions and hence to track the atmospheric evolution; the performance of this approach is limited only by the quality and availability of the *a priori* information and the accuracy of the atmospheric model used. The study ends with an encouraging practical inversion of a live scene measured at the Nd:YAG elastic-backscatter lidar station at our premises at the Polytechnic University of Catalonia, Barcelona. © 1999 Optical Society of America

OCIS codes: 010.0010, 010.1290, 010.3640.

1. Introduction

In an atmospheric lidar, the emission of a short laser pulse is followed by the reception of some radiation scattered from atmospheric constituents such as molecules, aerosols, and clouds. The interaction of the incident radiation with these constituents changes the intensity, the wavelength, or both, depending on the strength of this optical interaction and on the concentration of the interacting species. Consequently it is possible to retrieve information about the physical state of the atmosphere along the exploration beam path. In this

study we consider the particular case for which there is no wavelength shift in reception, and we focus on the inversion of elastic-backscatter lidar returns under the simplifying assumptions that the scattering processes are dominated by aerosols, that the spectral shape and composition of the scattering aerosols are spatially invariant, and that there are no wind components along the line of sight and under the approximation that there is no multiplicative noise (transmitter intensity fluctuation, speckle, or turbulence).

The inversion of atmospheric optical parameters,¹ namely, extinction and backscatter, based on pulsed elastic-backscatter lidars (i.e., with no wavelength shift in reception) has usually been treated by classic procedures such as the slope method,² exponential curve fitting,³ and Klett's method.⁴ These procedures all depart from the single-scattering elastic-backscatter lidar equation⁵

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

F. Rocadenbosch and A. Comerón are with the Antennas, Micro-waves, Radar, and Optics Group, Department of Signal Theory and Communications, Universitat Politècnica de Catalunya, C/Sor Eulalia de Anzizu s/n, 08034 Barcelona, Spain. The e-mail address for F. Rocadenbosch is roca@tsc.upc.es. C. Soriano and J.-M. Baldasano are with the Department of Engineering Projects, Universitat Politècnica de Catalunya, 647 Diagonal Avenue 10th, 08028 Barcelona, Spain.

Received 26 October 1998; revised manuscript received 8 March 1999.

0003-6935/99/153175-15\$15.00/0
© 1999 Optical Society of America

where $P(R)$ is the range-received power (W), $\beta(R)$ is the range-dependent volume backscatter coefficient of the atmosphere ($\text{m}^{-1} \text{sr}^{-1}$), $\alpha(R)$ is the range-dependent extinction coefficient (m^{-1}), R is the range (m), and A is the system constant (W m^3).

For the slope method and the least-squares and exponential-curve fitting algorithms we use the assumption of a homogeneous atmosphere to retrieve constant values (α , β) over the whole inversion range as estimates of the sought-after functions $\alpha(R)$ and $\beta(R)$. In a different category, the inversion of the range-dependent function $\alpha(R)$ in a turbid or moderately turbid inhomogeneous atmosphere is well solved by Klett's backward stable solution.^{4,6} This algorithm needs a boundary calibration at the far end of the inversion range interval in terms of $P_m = P(R_{\text{max}})$ and $\alpha_m = \alpha(R_{\text{max}})$,⁷ and we choose γ in the backscatter-to-extinction relationship:

$$\beta(R) = B_0 \alpha(R)^\gamma, \quad (2)$$

where B_0 is the backscatter-to-extinction ratio and γ is a fitting coefficient ($0.67 < \gamma < 1$).

The standard power law approach of Eq. (2) with B_0 and k constants has been shown by experiment to be reasonably good, so great precision in the functional description of this relation is not required for useful results to be obtained.⁴ Physically speaking, however, the most accurate approach assumes $k = 1$, a variable backscatter-to-extinction ratio $B(R)$,⁶ and even a distinction between aerosol and molecular ratios, $B_a(R)$ and B_m , respectively.^{6,8} In practice, it is difficult to determine $B(R)$ *a priori* as a function of the range because this value depends on physical properties of the scatters. We assume the simplifying hypotheses that $B(R)$ is constant along the inversion range of interest, that $B(R) = B_0$ (if necessary, we can achieve this equality by partitioning the inversion interval into appropriate subintervals with similar optical characteristics), that $k = 1$ as mentioned, and that gaseous absorption effects are negligible,⁹ as is multiple scattering (we note, however, that the last-named assumption may well not be true in the IR, where molecular absorption tends to dominate). In other words, these assumptions are equivalent to our previous assumptions that fluctuations in the backscatter figures from aerosols are due only to changes in the aerosols number density⁸ and that the spectral shape and composition of the scattering aerosols are spatially invariant. Here we also define the extinction-to-backscatter ratio $C = 1/B_0$ for mathematical convenience, notwithstanding the customary notation introduced in Eq. (2) by Klett.⁶

In spite of the fact that Klett's algorithm is significantly superior to the slope and that exponential-curve-fitting algorithms for the homogeneity approximation are no longer assumed, the accuracy of the inverted profile $\alpha(R)$ is limited by a set of user calibrations. In particular, the inverted extinction $\alpha(R)$ is nothing else but a representative of the family $\alpha(R, \gamma, \alpha_m)$ linked to the calibration pair (γ, α_m) , which is thought to be close to unknown atmospheric

values. In other words, the inversion of a range-dependent extinction profile from the return power is a many-to-one inversion problem (see App. B of Ref. 10), which can be solved only with the help of appropriate *a priori* information (e.g., calibrations along the observation path, physical constraints).

Furthermore, a common denominator in all these inversion methods is their lack of memory. Thus, for each power return received, a new inversion, which is completely independent of past ones, is performed (here we note that the meaning of "power return" or "power observable to the filter" must be understood as the result of averaging a number of different pulse returns according to some estimation of temporal correlation length of the atmosphere, i.e., the time interval for which one can talk of stationarity given a degradation criterion in the expected pulse integration gain). Yet one of the things that distinguishes the Kalman filter¹¹⁻¹³ from nonmemory estimators is the convenient way by which it accounts for any prior knowledge or inversions by means of a recursive process. As long as there are different incoming powers the filter updates itself weighted by the imbalance between the *a priori* estimates of the optical parameters (i.e., past inversions) and the new ones. As a result, the new estimation of the optical parameters, or the project-ahead step (*a posteriori* estimate), is improved based on a statistical minimum-variance criterion.

It is important to stress that the Kalman filter offers key advantages, such as averaging (so-called pulse integration), compared with classic approaches and that these advantages refer not only to the way in which observable data (optical power returns) are treated but also to the way in which the sought-after optical parameters are dealt with. Thus, from the point of view of the incoming data to the filter, the filter is able to combine the observable power returns in a better way than simple averaging, provided that *a priori* information about the observable noise, such as its covariance matrix R_k , is available to the filter. Stationarity is no longer necessary either, because it is perfectly possible to update the noise covariance matrix at each succeeding step t_k , for example by cancellation of synchronized flash-lamp noise or other sources of nonstationary noise, based on simultaneous noise characterizations at each shot. This is not true for classic pulse averaging, for which stationarity is indeed a difficult requirement. Therefore, one must adjust the integration time of the system according to the temporal correlation length of the scene or an estimate of it. Thus, for example, a scene with moving storm clouds may well have a correlation length of a few seconds, so longer integration intervals will probably lead to severe deterioration of the signal component itself.

From the point of view of estimating the optical parameters or the state vector x_k , one can tell the filter to improve its estimation guided by a suitable noise projection model and all the *a priori* information available from the user, namely, the *a priori* error covariance matrix for its initial guess x_0 and the

state-noise covariance matrix Q_k (discussed below), which refers to the driving noise or engine of the projection model. (In Appendix A the reader will find a more detailed explanation of the vocabulary just introduced.) It is this convenient way of assimilating *a priori* information that distinguishes the Kalman filter from any other classic technique such as data averaging, data trimming, and data fitting. Of course, if no *a priori* knowledge is available, the advantage of the filter is lost and its estimates will certainly coincide with those of the classic deterministic least-squares estimator.

Rye and Hardesty¹⁴ and Lainiotis *et al.*¹⁵ have found applications of the Kalman filter for estimation of the return power for incoherent backscatter lidar with multiplicative noise.¹⁶ Here we introduce an application of the filter to the solution of the inverse problem of joint estimation of the backscatter profile and the extinction-to-backscatter ratio from the return power in an elastic-backscatter lidar. Theoretical development is described in Sections 2 and 3, simulation results in Section 4, and first inversion of a live scene in Section 5.

2. Formulation of the Problem

(At this point the reader is advised to refer to Appendix A, which introduces both linear and nonlinear filters and notation used. See also Brown and Hwang¹³ for further insight.)

In Section 1 we outlined the key advantages of the extended Kalman filter (EKF) compared with non-memory-inversion algorithms. However, we shall see that its estimation trajectory depends on the observables or measurements at each sample run, so there is not a precomputed trajectory but a linearized one [see Eqs. (A20) and (A21) of Appendix A]. As a result, the main disadvantage of the EKF is that it is risky and stability becomes of concern, particularly when there are large initial uncertainties. For this reason, a design goal is to find a good choice of system-state variables that permits both reasonable modeling of the physical situation under consideration and rough performance of the filter even with poor initial guesses. Stability analysis of the filter is quite involved, if not impossible, for most practical situations; nearly always, one must resort to experimentation with different designs.

In this section we describe a fairly stable filter that permits simultaneous derivation of the backscatter profile and of the extinction-to-backscatter ratio from range-corrected power returns. As *a priori* information it is assumed that this ratio is spatially and temporally constant over the inversion range of interest along the lines discussed in Section 1. This ensures validity of the approximation $\gamma = 1$ in Eq. (2). Thus we wish to find the functions $\beta(R, t_k)$ and C that, under a minimum-mean-square-error criterion, best fit the observable power $P(R, t_k)$ at every discrete time t_k . The term "mean" refers here to the ensemble average over time t_k .

A. State Vector

First let us define f_s as the acquisition sampling rate of the system. Considering the two-way path of the lidar signal, the optical power time samples P_i [Eq. (1)] correspond to the acquisition system's spatial sampling period

$$\Delta R = c/2f_s, \quad (3)$$

and the spatial sampling points over the inversion interval become

$$R_i = R_{\min} + (i - 1)\Delta R, \quad i = 1 \dots N, \quad (4)$$

where R_{\min} is some predetermined minimum range of the system (for example, the minimum range of full overlap between the laser and the field of view of the receiving optics or some other inversion starting point). Equation (4) defines the observation cells of the system.

The state vector x_k to be estimated at discrete time t_k is a decimated discrete-time version of the backscatter function $\beta(R)$ formed by N/M inversion cells uniformly distributed along the whole lidar range plus one component that models the extinction-to-backscatter ratio $C = 1/B_0$ [after Eq. (2)]. It turns out that

$$x_k \equiv (\beta_1 \quad \beta_2 \dots \beta_{N/M} \quad C)^T, \quad (5)$$

where the backscatter at the i th inversion cell can be expressed as

$$\begin{aligned} \beta_i &= \beta(R_i), \\ R_i &= R_{\min} + (i - 1)M\Delta R, \\ i &= 1 \dots (N/M), \end{aligned} \quad (6)$$

where M is the decimation ratio. This means that, whereas the equivalent sampling period of the optical parameters in the filter's model is $M\Delta R$, the sampling period of the observable power returns is ΔR . As a result, we have more observables (N power samples or observation cells) than variables (N/M backscatter variables or inversion cells plus constant C) to estimate. This problem yields to an overdetermined vector problem with enhanced observability.^{13,17} Note, however, that the true atmospheric spacing $\Delta R'$ is differential in nature and that discretization involves the simplification $\Delta R' = \Delta R$ in the optical atmospheric profile, so such modeling errors are neglected.

B. Measurement Equation

The measurement model equation relates the observation vector z_k to the state vector x_k introduced above as [see also Eqs. (A2) and (A19) of Appendix A]

$$z_k = h_k(x_k) + v_k, \quad (7)$$

where h_k is a nonlinear function of the state x_k and v_k is zero-mean white Gaussian measurement noise with covariance matrix

$$R_k = E[v_k v_k^T]. \quad (8)$$

In the lidar inversion problem the measurement function $h_k(x_k)$ is a vector function whose i th component is defined as the R^2 -corrected power returns of the i th cell:

$$F_i = R^2 P(R_i), \quad (9)$$

after the lidar equation [Eq. (1)]. As a result, we can identify $h_k(x_k)$ in Eq. (7) as $F_k(x_k)$. Observation vector z_k is formed by time samples of the measured range-corrected power return, each of which corresponds to an observation cell along the exploration path, so at each time t_k , z_k represents a measurement realization. Hence

$$z_k \equiv [F_1(x_k) \ F_2(x_k) \ \dots \ F_N(x_k)]^T + v_k. \quad (10)$$

By using $F(R)$ instead of $P(R)$ we reduce the dynamic margin of z_k and, consequently, the numerical errors as we cycle through the Kalman loop [see Eqs. (A30)–(A34) in Appendix A]. At the far ranges, however, amplification of quantization noise generated during the analog-to-digital conversion might become significant and must be accounted for in a description of the statistics of the observation noise.

$$F_{M+1} = A\beta_2 \exp(-2C\{\beta_1[R_{\min} + (M-1)\Delta R] + \beta_2\Delta R\}), \quad (13)$$

$$F_N = A\beta_{N/M} \exp\left(-2C\left\{\beta_1[R_{\min} + (M-1)\Delta R] + \sum_{i=2}^{N/M} \beta_i M\Delta R\right\}\right). \quad (14)$$

Equations (11)–(14) relate measurement vector z_k presented to the filter to the *a priori* estimate of the state vector \hat{x}_k^- , and hence we can linearize h_k about \hat{x}_k^- to build the N -by- $N/M + 1$ observation matrix H_k [see Eq. (A25) of Appendix A]. Its elements can be split into

$$H_{ij}^{(1)} = \left. \frac{\partial F_i}{\partial \beta_j} \right|_{x=\hat{x}_k^-}, \quad H_i^{(2)} = \left. \frac{\partial F_i}{\partial C} \right|_{x=\hat{x}_k^-}, \quad (15)$$

where $H_k = [H_k^{(1)} H_k^{(2)}]$ and $H_k^{(1)}$ and $H_k^{(2)}$ are N -by- N/M and N -by-1 submatrices that represent the vector measurement equations for the backscatter and the extinction-to-backscatter ratios, respectively. It follows that

$$H_1 = \begin{bmatrix} \left(-2CR_{\min} + \frac{1}{x_1}\right)F_1 & 0 & 0 \dots 0 \\ \left[-2C(R_{\min} + \Delta R) + \frac{1}{x_1}\right]F_2 & 0 & 0 \dots 0 \\ \dots & \dots & \dots \\ -2C[R_{\min} + (M-1)\Delta R] + \frac{1}{x_1} \Big\} F_M & 0 & 0 \dots 0 \\ -2C[R_{\min} + (M-1)\Delta R]F_{M+1} & \left(-2C\Delta R + \frac{1}{x_2}\right)F_{M+1} & 0 \dots 0 \\ \dots & \dots & \dots \\ -2C[R_{\min} + (M-1)\Delta R]F_N & -2CM\Delta R F_N & -2CM\Delta R F_N \dots \left(-2CM\Delta R + \frac{1}{x_{N/M}}\right)F_N \end{bmatrix}_{N \times N/M}, \quad (16)$$

$$H_2 = \begin{bmatrix} -2R_{\min}x_1 F_1 \\ -2(R_{\min} + \Delta R)x_1 F_2 \\ \vdots \\ -2[R_{\min} + (M-1)\Delta R]x_1 F_M \\ -2\{[R_{\min} + (M-1)\Delta R]x_1 + \Delta R x_2\} F_{M+1} \\ \vdots \\ -2\{[R_{\min} + (M-1)\Delta R]x_1 + M\Delta R x_2 + \dots + M\Delta R x_{N/M-1} + M\Delta R x_{N/M}\} F_N \end{bmatrix}. \quad (17)$$

If a rectangular approximation is used in computing the transmittance term in Eq. (1), one can write

$$F_1 = A\beta_1 \exp(-2C\beta_1 R_{\min}), \quad (11)$$

$$F_M = A\beta_1 \exp\{-2C\beta_1[R_{\min} + (M-1)\Delta R]\}, \quad (12)$$

The subscript k is a reminder that, at each time t_k , these matrices must be recomputed at the *a priori* estimate \hat{x}_k^- such that the filter updates the trajectory that is used for linearization at each recursive step. At first sight, this would seem to make sense because, after all, there is no need to use the old trajectory when a better one is available. In the

words of Brown and Hwang,¹³ the flaw in this argument is that the “better trajectory” is better only in a statistical sense. There may be a chance that the updated trajectory is poorer than the nominal trajectory. In that event the estimates will be poorer and in turn will lead to further errors in the trajectory, leading to eventual divergence of the filter. As we had advanced, it must be cautioned a divergence of the filter is a potential risk that is common to all EKF’s, especially in situations when the initial uncertainties and measurement errors are large.

Inspection of Eqs. (16) and (17) (see the first M rows of each matrix) reveals some interesting points that have an effect on the filter performance:

In Eq. (16), which relates the observables to the filter and the backscatter component, the weight that each backscatter component cell has in the measurement equation increases linearly with ΔR along each observation cell that belongs to the same block of M cells (we read observation cells by rows). Eventually that weight becomes basically proportional to $-2C[R_{\min} + (M - 1)\Delta R]$ for the first inversion cell and to $-2CM\Delta R$ for the others (we read inversion cells by columns). As a result, it follows that, if R_{\min} is large compared with $M\Delta R$, the weights will tend to be unbalanced in favor of the first cell and the trajectory of the filter will be dominated completely by the estimation of this cell. Here a trade-off arises for parameter M : If $\Delta R \ll R_{\min}$, as is usually the case, increasing M such that $M\Delta R/R_{\min} \approx 1$ will cause all inversion cells to have similar weights, and therefore all those cells will contribute to the estimations of the filter (note that the contribution depends also on the choice of the covariance matrices). This situation also conveys the extra advantage of reducing the computational load because the number of inversion cells available is lower. However, increasing M will always degrade the atmospheric model presented to the filter because increasingly larger homogeneous atmospheric intervals will be unrealistically assumed. Likewise, if we compare the EKF formulation with the classic homogeneous exponential curve fitting, we see that the latter works with M equal to the length of the inversion interval, which consists of a single inversion cell. All things considered, this trade-off is a particular choice for each inversion scene at hand.

Equation (17), which relates the observables to the filter and the extinction-to-backscatter ratio component, reveals that state variable C will be sensitive to power-return contributions from all the observable cells along the exploration path, as evidenced by weights of the form $[R_{\min} + i\Delta R]$ along all the matrix rows. Qualitatively, this is a good point in favor of the robustness or stability of the filter because the trajectory of C is conditioned on the fluctuation of all power observables F_i rather than on the fluctuation of a few. The result is a smoothing effect on the trajectory.

Observation noise v_k is modeled as range-dependent stationary electronic thermal noise (i.e.,

Gaussian additive noise) whose variance can be written as³

$$\sigma_r(R)^2 = a[P(R) + P_{\text{back}}] + b, \quad (18)$$

where $P(R)$ is the range-return power defined in Eq. (1), P_{back} is the background power from any other interfering source (for instance, the Sun), and a and b are noise constants that depend only on specific parameters of the receiving system. The first term accounts for the contributions of the signal-induced shot noise to the total noise, and the second one merges into variable b , which includes contributions from both dark-current shot noise and thermal noise.

Therefore the observable-noise covariance matrix R_k can be expressed as the diagonal matrix

$$R_k = E[v_k v_k^T] = \begin{bmatrix} \sigma_r(R_1)^2 R_1^4 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_r(R_N)^2 R_N^4 \end{bmatrix}. \quad (19)$$

If other measurement noise sources (e.g., R^2 -amplified quantization noise) are present, one can increase pertinent terms along the major diagonal of R_k to accommodate such an extra variance. Off-diagonal terms make room for nonwhite measurement noise sources. We can address nonstationary noise by recomputing R_k at each succeeding step of the filter, and colored noise (such as synchronized flash-lamp interference) can be modeled by use also of elements off the main diagonal of R_k (see also Ref. 13 for further insight).

3. Atmospheric Model for the Extended Kalman Filter

The dynamics of both the optical backscatter and the extinction-to-backscatter ratio are difficult to define in terms of Mie’s scattering theory¹⁸ and its underlying physical variables. In addition, this is a many-to-one problem because there is not a unique relation between atmospheric backscatter and extinction.^{19,20} To overcome this difficulty a more convenient alternative is to try to model the macroscopic effects that these variables have on the optical parameters of interest. This is done in terms of the state-vector recursive model [see also Eqs. (A1) and (A18) of Appendix A]:

$$x_{k+1} = \Phi_k x_k + w_k, \quad \Phi_k = \begin{bmatrix} \Phi_\beta^{(N/M \times N/M)} & 0 \\ 0 & \Phi_C^{(1 \times 1)} \end{bmatrix}, \quad (20)$$

where Φ_k is the transition matrix from time t_k to time t_{k+1} , x_k is the state vector introduced in Section 2, and w_k is the state-noise vector with covariance matrix

$$Q_k = E[w_k w_k^T]. \quad (21)$$

Equations (20) and (21) combine all the key information that the filter has about the atmospheric model, as sketched in Fig. 1. In the β model (top) the backscatter of each inversion cell is modeled as a stochastic process that has both temporal and spatial correlation. The state-vector noise process w_k ($k = 1 \dots N/M$) is formed by spatially correlated components at the output of linear system A, which is, in

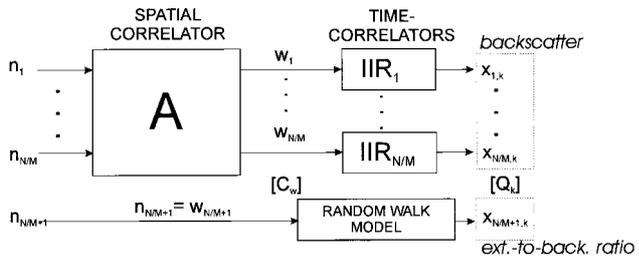


Fig. 1. Time-space atmospheric correlation model for the EKF. Both backscatter and extinction-to-backscatter ratio models, β ($x_{i,k}$, $i = 1 \dots N/M$ components) and C ($x_{N/M+1,k}$), are shown. The former is a Gauss-Markov vector process; the latter, a scalar random walk. For the backscatter components the spatial correlator is implemented by linear system A ; the time correlators, by an array of digital filters (IIR₁-IIR_{N/M}). The system is driven by an array of $N/M + 1$ white-noise uncorrelated components. w_i is the same array but with N/M correlated components for the β model plus another one uncorrelated for the C model ($n_{N/M+1}$).

turn, driven by an array of white-noise uncorrelated processes. In the scalar C model (bottom), C is modeled as a random-walk model^{21,22} ($\Phi_C = 1$) but with a low variance for the driving noise component $W_{N/M+1}$ ($\sigma_C^2 \approx 10^{-6}$). This random-walk model is preferable to a random-constant model ($\sigma_C^2 = 0$) because in the real world few things remain absolutely constant with time (for example, instrument biases, even though they are called biases, have a way of slowly changing with time, and recalibrations are necessary from time to time). A remedy for this problem is always to insert some driving noise into the state variable, even though doing so may cause some degree of suboptimality.

Finally, the atmospheric model assumes that there are no wind components along the line of sight, so the temporal and spatial correlation processes become independent.

A. Temporal Correlation

Temporal correlation is one of the most attractive advantages of the EKF formulation compared with nonmemory inversion algorithms. Temporal correlation comes from telling the filter that its projection steps can be improved based on the fact that the atmosphere usually has long correlation times and that, therefore, swift changes in any optical parameter are not possible. Temporal correlation is achieved by modeling of each component of state vector x_k (with k the discrete time) as a Gauss-Markov process.¹³ To simplify the notation, let us next redefine the i th component of state vector x_k [i.e., the backscatter component $\beta_i(t_k)$] as the scalar process y_k , so $y_k = x_{i,k}$ (here recall that subscript k is a reminder of discrete time t_k and that subscript i stands for the i th component, $i = 1 \dots N/M$). For the time being, let us concentrate our attention on the scalar Gauss-Markov process y_k .

The discrete-time equation of y_k takes the form of

an autoregressive moving-average discrete-time equation:

$$y_{k+1} = \exp(-1/L_c)y_k + w_k, \quad (22)$$

where y_k and w_k are Markovian and white sequences, respectively, and L_c is the temporal correlation length (in sample units). Spectrally, this is equivalent to saying that a Gauss-Markov process consists basically of low-pass-filtered white noise, so the 3-dB cutoff frequency of the coloring infinite impulse response (IIR) filter equals $1/L_c$ (see the IIR boxes in the block diagram of Fig. 1). In the limit when $L_c \rightarrow \infty$, a random-walk model¹³ is encountered. Based on Eqs. (19) and (22), the transition matrices for the backscatter and C subvectors of x_k become

$$\Phi_\beta^{(N/M \times N/M)} = \exp(-1/L_c)I^{(N/M \times N/M)}, \quad \Phi_C^{(1 \times 1)} = 1. \quad (23)$$

Note that Φ_k is constant over time t_k and that L_c is the same for all cells along the lidar exploration path.

In practice, Gauss-Markov processes are useful for telling the filter how to model time drifts in the magnitude of the atmospheric backscatter from driving Gaussian noise. For this reason it would be advantageous for us to link the Gaussian noise standard deviation σ_w with that of the Markovian noise, σ_m . After Eqs. (23) the two quantities can be related by

$$\sigma_w = \sigma_m[1 - \exp(-2/L_c)]^{1/2}. \quad (24)$$

Now we can return to the vector-state model of Eq. (20) and include again the subscript $i = 1 \dots N/M$ to refer to each component of the state vector. Figure 1 enables us to conclude that at the input of the IIR time correlators we have a set of N/M white-noise processes with standard deviation $\sigma_{w,i}$ and at their output a set of N/M Gauss-Markov processes (representative of the backscatter fluctuations) with standard deviation $\sigma_{m,i}$.

For the i th backscatter component it is assumed that the amplitude of the fluctuation is some percentage p of the backscatter magnitude itself. In addition, it is well known that, given $\sigma_{m,i}$, more than 95% of the Markovian amplitudes are distributed between $\pm 2.5\sigma_{m,i}$. Thus

$$\sigma_{m,i} = \frac{p}{2.5} \beta_i, \quad (25)$$

and the driving white-noise strength (at the input of the i th time correlator) $\sigma_{w,i}$ needed to cause a p -per-1 change in the amplitude of β_i over a correlation L_c becomes [from Eq. (24)]

$$\sigma_{w,i} \rightarrow \sigma_{\beta_i} = \frac{p}{2.5} \beta_i \left[1 - \exp\left(-\frac{2}{L_c}\right) \right]^{1/2}, \quad i = 1 \dots (N/M). \quad (26)$$

In what follows, p will be referred to as the strength parameter and $\sigma_{w,i}$ as σ_{β_i} , as indicated above.

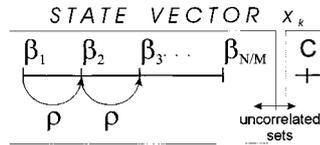


Fig. 2. Spatial correlation.

B. Spatial Correlation

Nonmemory algorithms, such as those introduced in Section 1, assume tight analytical correlation relations such as homogeneity in the case of the slope and exponential-curve-fitting algorithms or the power law of Eq. (2) in the case of Klett's method. Here we supersede these analytical constraints with loose stochastic ones based on the spatial correlation graph of Fig. 2 (this assumes that the pulse duration of the emitting laser source is so short that successive returns are correlated only by the atmospheric model). The underlying philosophy is that from a macroscopic point of view it seems sensible to guess that fluctuations in the backscatter component of some cell will certainly induce similar ones in its neighbor cells and related extinctions but with decaying intensity. At this stage it is assumed that instantaneous percent variations in both extinction and backscatter fluctuations, ε_α and ε_β , respectively, are quite similar ($\varepsilon_\alpha \approx \varepsilon_\beta$), so they cancel out when their ratio is computed. If, as discussed in Section 1, the aerosol composition is space invariant, one can assume that C is approximately the same for all cells along the inversion range (if this is not so, the inversion interval must be partitioned accordingly). As a result, the extinction-to-backscatter ratio C becomes uncorrelated with the backscatter set, so the following state-vector covariance matrix yields

$$C_w = \begin{bmatrix} C_\beta & 0 \\ 0 & C_C \end{bmatrix}, \quad (27)$$

where $C_C = \sigma_c^2$ and σ_c is a low value that models the random constant C to ensure that C_w is a positive-definite matrix. C_β is the backscatter covariance matrix, constructed as depicted in Fig. 2 as

$$C_\beta = \begin{bmatrix} \sigma_{\beta_1}^2 & \rho\sigma_{\beta_1}\sigma_{\beta_2} & \dots & \dots & \rho^{n-1}\sigma_{\beta_1}\sigma_{\beta_n} \\ \dots & \sigma_{\beta_2}^2 & \dots & \dots & \rho^{n-2}\sigma_{\beta_2}\sigma_{\beta_n} \\ \dots & \dots & \dots & \dots & \vdots \\ \dots & \dots & \dots & \dots & \sigma_{\beta_n}^2 \end{bmatrix}, \quad (28)$$

where ρ is the correlation coefficient between one cell and the next one along the beam path and σ_i has already been defined in relation (26). On the condition that $|\rho| < 1$ and using standard algebraic techniques, we can show that Eq. (28) does represent a covariance matrix.

Finally, with the simplification that there are no wind components along the line of sight, temporal and spatial correlation processes become independent and the filter's state-vector covariance matrix

for the atmospheric model can be computed from expressions (26)–(28) as

$$Q_k = \begin{bmatrix} Q_\beta & 0 \\ 0 & Q_C \end{bmatrix}, \quad Q_\beta = \frac{C_\beta}{1 - \exp(-2/L_c)}. \quad (29)$$

It emerges that Q_k is the same as C_w except for a scaling factor that we can skip by redefining an equivalent intensity parameter p' .

The fact that the standard deviation of the driving noise process for β_i , σ_{β_i} in relation (26), is proportional to β_i suggests a recursive update for Q_k . This solution is, however, ill advised when we take into account that the EKF is a linearized filter. Unless the user's initial guesses are quite close to the solution and the model parameters for Q_k (particularly p and ρ) are well matched, which is not often the case, experimental results have corroborated that a static Q_k approach (i.e., no updating) yields the best results for nonwindy atmospheres (equivalently, scenes with low backscatter drift). Even though it is far beyond the scope of this paper, for windy atmospheres recursive updating offers a solution that has yet to be investigated. Thus, in instances when radial wind strongly correlates both space and time fluctuations along the line of sight, one might consider a variant Q_k such that observation cells along the beam path become progressively affected by different correlation links as time goes on. In addition, boosting elements off the main diagonal would tell the filter that there is a significant increase in the correlation among neighbor cells.

Usually the state-noise covariance matrix of the EKF Q_k is the most difficult input to assess because its atmospheric counterpart $Q_{k,a}$ (the subscript a stands for atmospheric) is unknown (the problem of finding good models for Q_k has sometimes been addressed from a partitioned approach.¹⁵ These large model uncertainties justify that the initial *a priori* error covariance matrix be much larger than the state-vector covariance, so

$$P_0^- = \mu Q_0, \quad \mu \geq 1. \quad (30)$$

4. Simulation Results

Now we present two simulation sets to test EKF performance with different modeling errors and visibilities. Synthetic atmospheric scenes were generated by means of the stochastic simulator of Fig. 1 and the input of a set of atmospheric model parameters C_a [$1/B_0$ in Eq. (2)], $L_{c,a}$ [Eq. (22)], p_a [relation (26)], and ρ_a [Eq. (28) and Fig. 2] that are different from those given to the EKF (C , L_c , p , and ρ), so $Q_{k,a} \neq Q_k$. In particular, we set $C_a = 25$ sr, $C = 22.5$ sr (90% of C_a), $L_{c,a} = 10$ samples, $L_c = 5$ samples, $p_a = 0.4$, $p = 0.5$, $\rho_a = 0.6$, and $\rho = 0.3$. This choice of values is based on comparisons of time-space plots of synthesized power-return signals with nonwindy time-space real observations. Typically it has been found that ρ_a lies in the range 0.3–0.7 and that values close to unity are ill advised because they yield stiff spatial profiles that are too correlated, thus making it diffi-

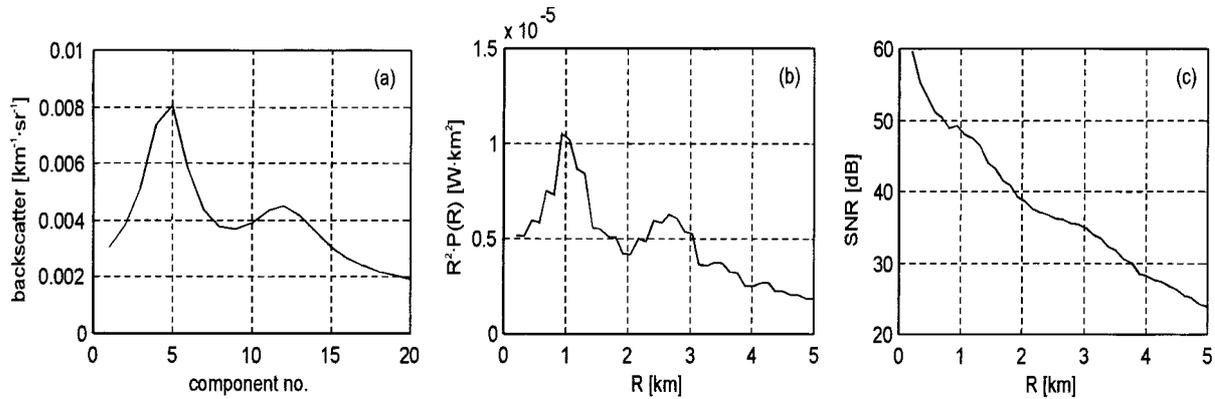


Fig. 3. First simulation set: (a) input backscatter profile, (b) synthesized range-corrected return power, (c) associated SNR.

cult to accommodate heterogeneities along the exploration path. At this simulation stage, p and L_c have turned out to be the most critical parameters. Finally, we use eigenvalue decomposition²³ to compute the linear correlator (A in Fig. 1), which is driven by white-noise uncorrelated sequences $[n_1 \dots n_{N/M}, n_{N/M+1}]$.

Both simulation sets have worked with a lidar observation range from $R_{\min} = 200$ m to $R_{\max} = 5$ km and a spatial resolution $\Delta R = 123.1$ m, which translates into $N = 40$ observation cells. Return-power and noise parameters have also been simulated according to typical system parameters ($A = 2.35 \times 10^{-3} \text{ W km}^{-3}$ in Eq. (1); $a = 1.8 \times 10^{-10} \text{ W}$, $b = 5 \times 10^{-18} \text{ W}^2$ and $P_{\text{back}} = 2 \text{ nW}$ in Eq. (18)] of the elastic-backscatter lidar of the Polytechnic University of Catalonia. The atmospheric extinction-to-backscatter ratio and the mean value of the synthesized hump-shaped inhomogeneous backscatter profile were computed from the pairs ($\alpha = 0.1 \text{ km}^{-1}$, $\beta = 4 \times 10^{-3} \text{ km}^{-1} \text{ sr}^{-1}$) for the first set and ($\alpha = 1 \text{ km}^{-1}$, $\beta = 3 \times 10^{-2} \text{ km}^{-1} \text{ sr}^{-1}$) for the second one. These pairs of optical parameters are representative of “exceptionally clear” and “moderate” visibility, respectively, according to the terminology of Kosmieder²⁴ and Kruse *et al.*²⁵

Because the simulator generates inhomogeneous backscatter profiles for each pair (α, β) (explained in the examples below), we find it convenient to describe the atmospheric condition in terms of the optical thickness or path-integrated extinction coefficient as

$$\tau(R_{\min}, R_{\max}) = \int_{R_{\min}}^{R_{\max}} \alpha(r) dr = C \int_{R_{\min}}^{R_{\max}} \beta(r) dr, \quad (31)$$

where a constant extinction-to-backscatter ratio along all the inversion region has been assumed. This hypothesis is assumed for the sake of simplicity in the following inversion examples. In a practical inversion, the inversion range should be partitioned adequately to ensure the validity of this hypothesis for each cooperative EKF and, in turn, the validity of Eq. (2).

In all cases the filter has an estimated 20 inversion cells ($M = 2$) during 150 iterations. To test EKF

performance we initiated filter action in the simplest way by means of the homogeneous profile $\hat{x}_0^- = 0.9[\beta \dots \beta \ C_a]$, which represents a 10% underestimation from the figures mentioned above, and $\mu = 10^3$ in Eq. (30).

Figures 3–6 illustrate simulation results for the nonturbid atmosphere of the first set ($\alpha = 0.1 \text{ km}^{-1}$, $\beta = 4 \times 10^{-3} \text{ km}^{-1} \text{ sr}^{-1}$), whose optical thickness is 0.25 [Eq. (31)]. Figures 3(b) and 3(c) represent the synthesized range-corrected return power and its associated signal-to-noise ratio (SNR), which is computed as a ratio of voltages at the receiver output, in response to the hump-shaped backscatter profile of Fig. 3(a). Note that the mean value of the backscatter over the inversion range equals $\beta = 4 \times 10^{-3} \text{ km}^{-1} \text{ sr}^{-1}$.

Figures 4 and 5 compare the simulated atmospheric scene with the EKF inversion estimates from the power-return observable fed to the filter at each iteration. For Figs. 4(a) and 4(c), time-fluctuating backscatter profiles and extinction-to-backscatter ratio C were generated according to the simulator of Fig. 1. In particular, Fig. 4(c) shows the random

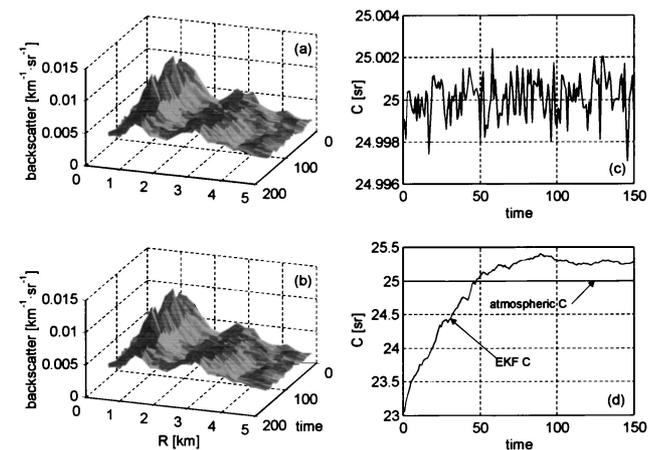


Fig. 4. First simulation set. Comparison of atmospheric and inverted optical parameters: (a) time animation of the atmospheric backscatter profile; (c) time animation of the atmospheric extinction-to-backscatter ratio; (b), (d) EKF inversion results for (a) and (c), respectively.

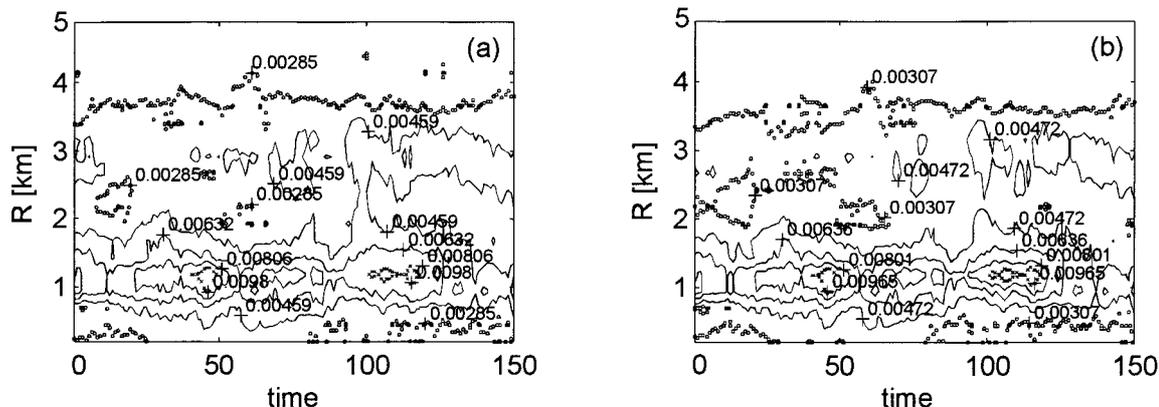


Fig. 5. First simulation set: contour plots showing (a) time-space evolution of the atmospheric backscatter and (b) the filter's backscatter estimates.

constant behavior of ratio C as described by Eqs. (23) and (1). Comparing Figs. 4(b) and 4(d) with Figs. 4(a) and 4(c) shows that, whereas the filter is able to retrieve the shape of the backscatter profile virtually from the first iteration, it takes some time (75 iterations) to estimate the profile for extinction-to-backscatter ratio C . Misestimation of C causes, in turn, misestimation of the magnitude of the backscatter's profile, even when its shape is fairly well recovered. Iso-backscatter contour curves for both the atmospheric backscatter and the inverted profiles are shown in Fig. 5.

In this example fast retrieval of the backscatter profile is achieved not only because the SNR is high but also because the filter has both a reasonably good estimation and a suitable model for extinction-to-backscatter ratio C from the beginning. The key here is that *a priori* information based on physical knowledge that C must remain approximately constant with time is fed into the C model to tell the filter that fast fluctuations for the backscatter components are possible, whereas they are not for C . At this point one might well argue that, if the SNR were high, one could perform appropriate least-squares inversions for the unknown state variables on reception of each observable. Even though doing this seems feasible in theory, in practice with such large fluctuations in the estimate of C with time it may result that the inversion of the backscatter components will lack any physical meaning.

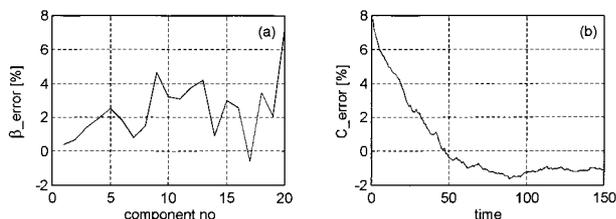


Fig. 6. First simulation set. Relative inversion errors after 150 iterations: (a) atmospheric backscatter error indexed by number of inversion cells, (b) time evolution of the extinction-to-backscatter inversion error.

As a result of this first simulation we can conclude that, in spite of the user-induced modeling errors and the modest initialization, the filter manages to keep excellent track of the atmospheric counterparts all along the 5-km inversion range. Filter performance is best seen from Fig. 6, where inversion errors in both the backscatter and the extinction-to-backscatter ratio are computed. We see that after a training stage consisting of some 75 inversions, C error dwindles from an initial 10% underestimation of \hat{x}_0^- (8% just in the 1st iteration) to virtually nil (-1%) during the tracking stage (75th to 150th iterations). The same thing happens with the backscatter estimates if one compares Fig. 5(b) with Fig. 5(a) for the same time interval. Even though the filter's performance is fairly good all along the exploration path, β errors in Fig. 6(a) tend to increase for cells that are located far from the lidar, suggesting a decreasing trend in the SNR of Fig. 3(c). All things considered, slightly poorer performance is observed in the upper range (see Fig. 5). The effect of larger optical depths on the SNR and on the subsequent range-limited filter performance is corroborated in the next simulation.

The second simulation set ($\alpha = 1 \text{ km}^{-1}$, $\beta = 3 \times 10^{-2} \text{ km}^{-1} \text{ sr}^{-1}$) corresponds to a moderately turbid atmosphere with an optical thickness of 2.62 [Eq. (31)] and comprises Figs. 7–10. This deeper optical thickness now translates into a steeply-decreasing range-corrected return power in Fig. 7(c) and to a SNR in Fig. 7(d) that ranges from 67.5 dB at R_{\min} to -34.3 dB at R_{\max} . Under these circumstances, the filter inversion is acceptable only for the first 10 cells of the inversion range, where the SNR is still above a reasonable 15 dB, as illustrated in Fig. 8, where both simulated and inverted scenes are presented as three-dimensional plots. In the upper range, filter performance becomes dramatically reduced, and even some backscatter components incur negative values, which lack any physical significance.

Likewise, Fig. 9 illustrates comparative time plots of the atmospheric and the inverted backscatter magnitudes in four cells along the inversion path, which have quite different SNR's. For the first cell [Fig. 9(a)] the

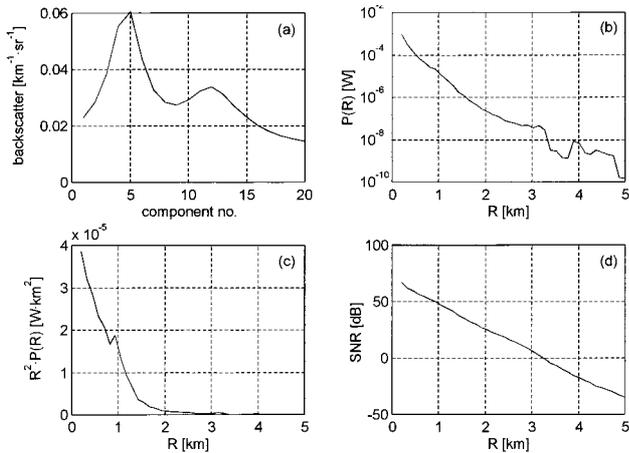


Fig. 7. Second simulation set: (a) input backscatter profile; (b), (c) synthesized return and range-corrected return powers, respectively; (d) associated SNR.

SNR is ~ 65 dB, and the filter's estimates perfectly follow the Markovian drift of the atmospheric backscatter. This excellent behavior degrades in the 6th cell, whereas for the 13th cell, for which the SNR is less than 0 dB, the filter is lost. Horizontal solid lines in Fig. 9 recall the backscatter value of each cell at $t = t_0$, which is inherited from its location in the initial hump-shaped backscatter profile of Fig. 7(a). As time goes on, the profile is distorted, and each cell follows a Markovian walk, as illustrated in Fig. 9.

Inversion errors follow trends similar to those shown in Fig. 6, but now the convergence time (i.e., the time to reach steady tracking) is much faster, as evidenced by the fast-decreasing slope of the C error in Fig. 10(b). This enables us to reach a steady-state C error of -1% (as in the first simulation) in fewer than ten iterations. Note that, in spite of the extremely poor SNR at high ranges, because the observable noise covariance matrix R_k describes the quality of the observables to the filter in an efficient way the EKF is able to adjust its gain such that it can compute the extinction-to-

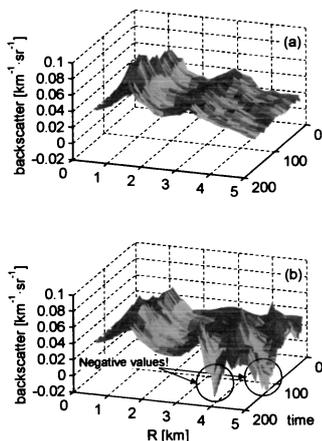


Fig. 8. Second simulation set. Time-space evolution of the backscatter state vector: (a) synthesized atmospheric backscatter, (b) EKF estimates.

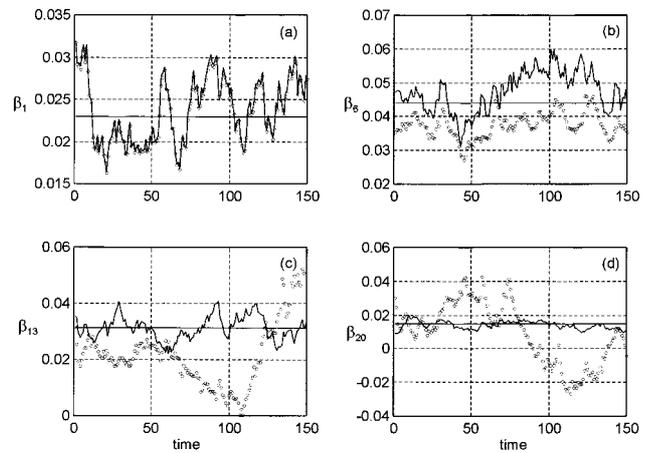


Fig. 9. Second simulation set. Temporal evolution of the backscatter in observation cells 1, 6, 13, and 20. Solid curves, atmospheric backscatter; circles, EKF estimates; solid horizontal lines, backscatter values at the beginning of the simulation.

backscatter ratio from the first section of the observation range, where the SNR is much higher. In the first simulation set, however, where the observables were not much corrupted by measurement noise, we considered the whole inversion range when we computed C . In the estimation of the backscatter profile there is no way to counteract noise, and, as a result, only cells that belong to the lower range are estimated reasonably well (inversion errors are 0–30%).

A comprehensive collection of extra simulations has shown that filter performance improves as long as $Q_k \rightarrow Q_{k,\alpha}$ and $SNR \rightarrow \infty$, as expected, and that for reduced uncertainty spans in the model parameters of Q_k or in the filter's initialization, steady convergence is always achieved in more or less time. Yet, if the user's uncertainty span is too large so unstable errors occur, one can easily identify this situation, chiefly by monitoring the trace of the β and C block matrices of *a posteriori* error covariance matrix P_k and the convergence behavior of extinction-to-backscatter ratio C . For ill-advised input values, the C trace increases without bound and, consequently, the estimation of C tends to plus or minus infinity. This result is further discussed in Section 5, where we treat the inversion of a real scene.

5. Live-Scene Inversion

Power returns measured with the vertical-exploration elastic-backscatter lidar system of the Polytechnic University of Catalunya at Barcelona

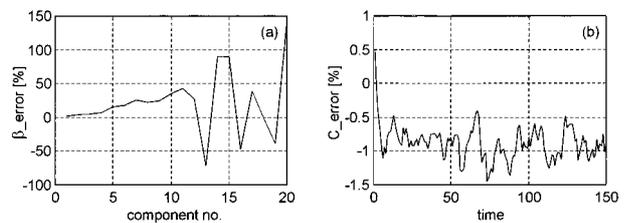


Fig. 10. Same as Fig. 6 but for the second simulation set.

Table 1. Elastic-Backscatter Lidar System Specifications

Laser		Receiver		System	
Gain medium	Nd:YAG	Focal length	2 m	Configuration	Vertical biaxial
Energy	0.5 J/532 nm	Aperture \varnothing	20 cm	System noise-equivalent power	70 fW Hz ^{-1/2}
Divergence	0.1 mrad	Detector	Avalanche photodiode (EG&G C30954)	Minimum detection	<5 nW
Pulse length	<10 ns	Net	$6 \times 10^1 - 3 \times 10^6$	Power acquisition	20 Msps/12 bits
Pulse-repetition frequency	10 Hz	Responsivity	V/W	Spatial resolution	7.5 m
		Bandwidth	10 MHz		

Mps, mega samples per second; 20×10^6 samples/s.

(Table 1) were inverted by use of the EKF formulation presented above.

Figure 11(a) illustrates the measurement scene from the campaign of 25 September 1996. The visibility could be described as very clear, and two cloud layers were detected, at 6.4–7.7 and 8.6–9.4 km. The scene corresponds to a stationary range interval of 2.28–2.65 km cut from the original one, which ranges from 0.5 to 15 km, to reduce the filter order and therefore the computational load (0.12- and 8.4-s/iteration for filter orders equal to 50 and 250, respectively, with MATLAB running on a Pentium 166 MMX platform). The inversion interval contains 50 observation cells ($\Delta R = 7.5$ m) and 50 signal packets, where a packet is defined as the averaging or integration of a certain number of return signals in response to an emission pulse. In this example 15 return pulses were integrated. According to the pulse-repetition frequency of the system, this integration time is virtually nil compared with the long temporal coherence length of the atmosphere in this case, which can be assumed to be approximately stationary. Because longer time records were not available, we performed the inversion by feeding the filter with those 50 packets in a round-robin cycle of as many as 500 iterations ($N = 50, M = 2$) to ensure steady-state tracking. In other words, each packet constituted an observable vector z_k to the filter, and the feeding sequence could be written as $(z_1, z_2, \dots, z_{50}), (z_1, z_2, \dots, z_{50}), \dots, (z_1, z_2, \dots, z_{50})$.

This procedure can be understood as a kind of fixed-interval smoothing. In this case the time in-

terval of measurements (i.e., a data span of 50 packets) is fixed, and we seek optimal estimates at some or, perhaps, all interior points. The round-robin approach is poorer (yields a higher smoothed error variance) than a two-way smoothing,¹³ but it is conceptually simpler.

The lidar was operated at 390-mJ output energy, 532-nm wavelength, and 1.8×10^4 V/W net responsivity, yielding $A = 1.81 \times 10^{-3}$ W km⁻³ in Eq. (1) and observation noise parameters $a = 3.7 \times 10^{-9}$ W and $b = 3.4 \times 10^{-16}$ W² in Eq. (18), which we used to compute the noise covariance matrix R_k .

Modeling the scene into the state-noise covariance matrix Q_k of the filter was the most difficult part of the process and was done after completion of these two experimental steps: First, the correlation length L_c was estimated. As no significant slopes were evident in the time series of the scene in Fig. 11(a), which looks like a plateau, we conjectured that the scene should be associated with a correlation length greater than or equal to the time record length; hence L_c was set to 50 samples. Second, the strength parameter p was estimated. Because p is directly linked to backscatter fluctuations, the fluctuations were directly estimated from those in the range-corrected power realizations because the optical depth along the field of sight was so low. Thus the strength parameter was computed as the quotient between the standard deviation of the range-corrected return power in each observation cell and the time-averaged value in each cell. Formally,

$$\hat{p}_i = \frac{[E(\{F_i(t_k) - E[F_i(t_k)]\}^2)]^{1/2}}{E[F_i(t_k)]}, \quad (32)$$

where \hat{p}_i represents the estimation of the strength parameter of the i th cell and $E(\cdot)$ is the expectancy operator, which in practice is computed as a time average over power realizations. This simulation yielded an estimate of p of 6–12%, and 10% was input to the filter. As for correlation coefficient ρ , we set $\rho = 0.3$ as in the previous simulations, because it turned out to be a noncritical parameter. The filter was initialized again with a homogeneous profile of the form $\hat{x}_0^- = [\beta \dots \beta \ C]$, now with $\beta = 4 \times 10^{-3}$ km⁻¹ and $C = 33.3$ sr, which are typical values for a clear atmosphere.^{24,25}

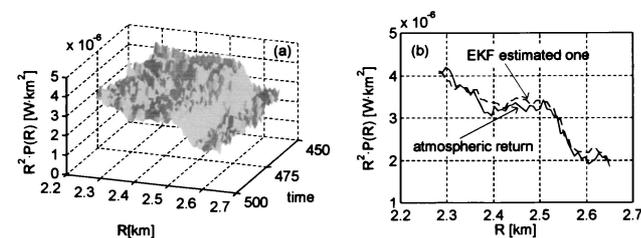


Fig. 11. Actual application example. (a) Range-corrected received power as observables to the filter (z_k). Fifty signal packets are shown in the range 2.28–2.65 km. (b) Comparison of the range-corrected measured power and the filter's estimated power, showing good tracking.

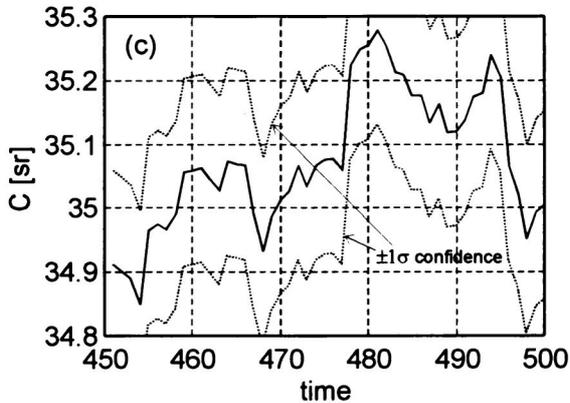
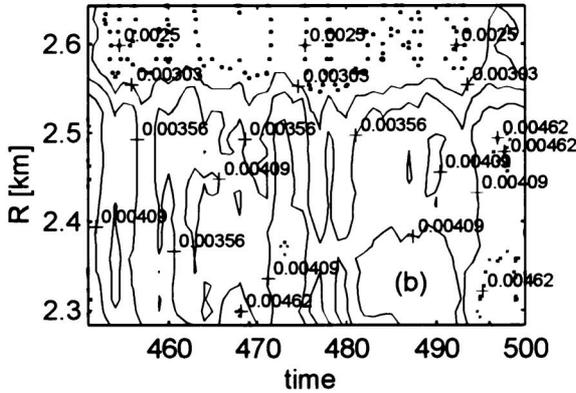
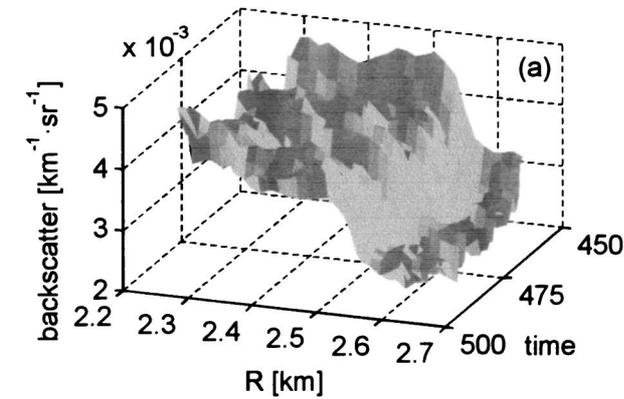


Fig. 12. Actual inversion. Inversion results in response to the scene of Fig. 11: (a) time-space inversion of the atmospheric backscatter, (b) contour plot of (a), (c) inversion of the extinction-to-backscatter ratio as a function of time (solid curve) and $\pm 1\sigma$ confidence margins.

Inversion results are shown in Figs. 11(b)–13. Figure 11(b) shows good agreement between the live atmospheric power returns and the estimated ones, whereas Fig. 12 plots the sought-after inversion, which is in accordance with values reported in the literature.⁵

Key performance of the filter is well illustrated by examination of Figs. 11(b) and 12(b): As time goes on, the filter reaches excellent balance between fitting the power observables of Fig. 11(b) and ensuring

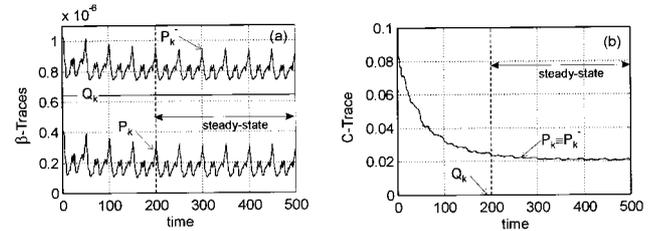


Fig. 13. Actual inversion. Convergence and tracking indicators: (a) time evolution of P_k , P_k^- , and Q_k β traces, (b) C -trace time plot. Steady convergence is indicated by a virtually constant figure from the 200th iteration onward.

physically meaningful time-space correlation in Fig. 12(b). Note that for each actual measurement datum received on a particular run as in Fig. 11(b), power matching is good but not excellent. This result, however, gives continuity to the traces of the isobackscatter lines of Fig. 12(b). Recall that the filter works not under an instantaneous minimum-least-squares error philosophy but under a minimum-mean least-squares one over time t_k , which is a statistical criterion.

Interpretation of Fig. 12(c) warrants some comments: First, recall that the starting value for C is $C(t_k = 1) = 33.3$ sr. After some 450 iterations (i.e., 10 periods later), $C(t_k = 450) = 34.9$ in Fig. 12(c), which represents the final estimation for C . To interpret this difference we must distinguish between the training interval (approximately 1–100 iterations) and the tracking interval of the filter (100 and more iterations). The filter is estimated to be a quasi-periodic signal with a period of 50 iterations because of the round-robin feeding sequence to the filter [an example of such a signal is shown in Fig. 13(a)]. Thus, if we read the filter estimates at the beginning of each period (not shown) we would get $C(t_k = 1) = 33.33$, $C(t_k = 50) = 34.5$, and $C(t_k = 100, 150, \dots, 450, \dots) = 34.9$ and stable, the same values as in Fig. 12(c). Note also that the initial value $C = 33.33$ sr is only an initial guess, not the correct value. As a result, we can interpret any period that belongs to the tracking interval of the filter (in this example, the tenth period, corresponding to iterations 450 to 499) as the final estimation of the EKF in response to the set of 50 observables measured.

Estimations of the inversion errors in the EKF are illustrated in Fig. 13: From the first iterations, *a posteriori* and *a priori* estimation errors, which are represented by β and C traces of P_k and P_k^- , respectively, are progressively reduced to a residual that is close to the trace of Q_k . Qualitatively, this suggests that the filter is wise enough to counteract measurement noise in the inversion estimates through convenient averaging with past estimations until it reaches some lower bound given by the state noise owing to atmospheric mobility, represented by the trace of Q_k . Figure 13 is also useful for derivation of an estimation of the inversion error in both β and C estimates. For scalar C this derivation is easy because Fig. 13(b) yields the variance of the *a posteriori*

inversion error ($\sigma_c^2 = 0.02$). If we interpret C as a stochastic process such that for each point in time t_k we have a stochastic variable C_k with mean \bar{C}_k and standard deviation σ_k , then, assuming a Gaussian error distribution for C_k , it is well known that the probability will be

$$p[|C - \bar{C}| \leq n\sigma] = \text{erf}\left(\frac{n}{\sqrt{2}}\right). \quad (33)$$

Because in Fig. 13(b) $\sigma_k \approx \sigma_C$ for all t_k in the steady-state time interval of the filter, the error margin for C in Fig. 12(c) becomes ± 0.14 (1σ) for a 68.3% confidence level [as plotted in Fig. 12(c) by superimposed traces] or ± 0.42 (3σ) for 99.7%. This behavior is in accordance with the physical model given for C that assumes that it is approximately constant with time. Yet the significance of Fig. 12(c) lies in the fact that the filter is able to depart from the user's initial guess of $C = 33.3$ sr and estimate C within the confidence margin given. Similar computations can be derived for the backscatter components, which turn out to be some $\pm 2.5 \times 10^{-4} \text{ km}^{-1}$ for 99.7% confidence. This is an extremely thin error margin compared with the vertical scales of Figs. 12(a) and 12(b).

Inversions with different model parameters were also carried out for the same scene. For a reduced span of input values, a few more convergent sets were found. In these cases, slightly different inversion results arose but always in the backscatter component. Yet these uncertainties can be fitted inside the error span of the estimation error inferred from β and C traces. As a result, once can conclude that, for a reasonable input set, the EKF yields coincidentally good inversions.

6. Conclusions

Inversion of the atmospheric backscatter and the extinction-to-backscatter ratio has been achieved with an EKF (Section 2) and an appropriate stochastic model (Section 3) that describes the macroscopic effects of the atmosphere over the backscatter component in terms of spatial and temporal correlation. As times goes on, the filter keeps track of the atmospheric fluctuations by projection steps that conveniently average actual estimations with past ones under a statistical minimum-mean-square error criterion.

Through different simulations (Section 4) and a prototype inversion of a real scene (Section 5), the filter has shown excellent inversion performance and robustness to modeling errors; the *a posteriori* error covariance matrix trace and the asymptotic trend of the extinction-to-backscatter ratio are good convergence indicators. The encouraging results achieved by this adaptive technique promise that the EKF will be a long-lasting alternative algorithm for lidar inversion.

Appendix A

To present a self-contained paper and to clarify the notation used in this paper, we present here a theoretical summary of information on the linear Kalman

filter (Section A) and the extended Kalman filter (Section B) based on Refs. 11 and 13.

A. Linear Discrete Kalman Filter

Let us consider the stochastic discrete time vector process x_k to be estimated (dimensions are indicated in parentheses), as modeled by

$$x_{k+1} = \Phi_k x_k + w_k, \quad (A1)$$

where x_k is the system's time-state vector at time t_k ($n \times 1$), Φ_k is the transition-state matrix from time t_k to t_{k+1} ($n \times n$), and w_k is the state-noise vector ($n \times 1$).

The measurement or observation of the process is assumed to take place at discrete times t_k , according to the following linear relationship:

$$z_k = H_k x_k + v_k, \quad (A2)$$

where z_k is the measurement vector at time t_k ($n \times 1$), H_k is the measurement matrix ($m \times n$), and v_k is the measurement error vector ($m \times 1$).

Before proceeding further, let us illustrate the meaning of these variables by considering the example of xy -speed monitoring of uniform two-dimensional motion. In this case, x_k would be the mobile's xy -speed components (v_x, v_y) and F_k a two-dimensional matrix that describes the mobile's speed components from one sample time to the next (as the motion is expected to be uniform for all t_k , $F_k = I$ can be assumed). Finally, w_k would be a driving noise process that models velocity drift in the mobile's motion. As for the observation system, H_k would be the linear equations that relate the mobile's xy coordinates to its vector speed components. Were these equations not linear, it would be necessary to resort to the extended Kalman filter introduced in Section B. Finally, v_k would be the measurement noise.

Refer to Eqs. (A1) and (A2); the noisy vectors w_k and v_k must be white sequences with known covariance matrices. In addition, w_k and v_k must be uncorrelated. In instances when these conditions are not fulfilled, one can augment the state vector and estimate the correlated noisy samples w_k' and v_k' from other variables w_k and v_k that form an uncorrelated, orthogonal base. Doing so would enable the Kalman filter to work with colored noise, even if it were nonstationary. Noise covariance matrices are given by

$$E[v_k v_i^T] = \begin{cases} R_k & i = k \\ 0 & i \neq k \end{cases}, \quad (A3)$$

$$E[w_k w_i^T] = \begin{cases} Q_k & i = k \\ 0 & i \neq k \end{cases}, \quad (A4)$$

$$E[w_k v_i^T] = 0 \quad \forall k, i. \quad (A5)$$

At this point it is assumed that an initial estimate of the process at the same point in time t_k is known. This estimation, called an *a priori* estimate, is denoted \hat{x}_k^- , where the superscript minus is a reminder that it is the best estimate before the measurement is assim-

ilated at t_k and the circumflex over the vector means "estimate of." The estimation error can be written as

$$e_k^- = x_k - \hat{x}_k^-, \quad (\text{A6})$$

and its associated error covariance matrix is

$$P_k^- = E[e_k^- e_k^{-T}] = E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T]. \quad (\text{A7})$$

Now a linear combination of the measurement z_k is sought to improve the prior estimate \hat{x}_k^- :

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H_k \hat{x}_k^-), \quad (\text{A8})$$

where \hat{x}_k is the updated estimate and K_k is the Kalman gain (yet to be determined).

The problem now is to find the particular blending factor K_k that yields an optimal updated estimate under some reasonable criterion. Just as in the Wiener solution, a minimum-mean-square error criterion is used. Toward this end, let us write the error covariance matrix associated with the *a posteriori* (updated) estimate:

$$P_k = E[e_k e_k^T] = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]. \quad (\text{A9})$$

If Eq. (A2) is substituted into Eq. (A8) and then into Eq. (A9), the result is

$$P_k = (I - K_k H_k) P_k^- (I - K_k H_k)^T + K_k R_k K_k^T. \quad (\text{A10})$$

The optimization problem is now equivalent to finding the optimum gain at each time t_k , K_k that minimizes the error variances for the elements of the state vector being estimated. As the individual terms along the major diagonal of P_k do represent these error variances, we can solve the optimization problem by minimizing the trace of P_k . If we follow straightforward matrix differentiation analysis and the next two formulas are considered:

$$\frac{d[\text{tr}(AB)]}{dA} = B^T \quad (AB \text{ must be square}), \quad (\text{A11})$$

$$\frac{d[\text{Tr}(ACA^T)]}{dA} = 2AC \quad (C \text{ must be symmetric}), \quad (\text{A12})$$

we have that

$$\frac{d[\text{Tr}(P_k)]}{dK_k} = -2(H_k P_k^-)^T + 2K_k(H_k P_k^- H_k^T + R_k), \quad (\text{A13})$$

where $\text{Tr}()$ is the trace operator. Now the derivative can be set equal to zero and the optimal gain, the Kalman gain, found. The result is

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1}. \quad (\text{A14})$$

Once the Kalman gain is known, if it is substituted into Eq. (A10) the *a priori* and *a posteriori* error covariance matrices can be related as

$$P_k = (I - K_k H_k) P_k^-. \quad (\text{A15})$$

Now note that both Eq. (A8) and the Kalman gain of Eq. (A14) provide a means to assimilate the measurement at t_k in a recursive relation that makes use of the *a priori* variables \hat{x}_k^- and P_k^- . For this reason a similar need for \hat{x}_{k+1}^- can be anticipated at the next step to assimilate the next measurement z_{k+1} . This new *a priori* estimate can be computed if we project ahead the old estimate \hat{x}_k through a transition matrix:

$$\hat{x}_{k+1}^- = \Phi_k \hat{x}_k. \quad (\text{A16})$$

Finally, the new *a priori* error covariance matrix P_{k+1}^- is computed as

$$\begin{aligned} P_{k+1}^- &= E[e_{k+1}^- e_{k+1}^{-T}] \\ &= E[(\Phi_k e_k + w_k)(\Phi_k e_k + w_k)^T] \\ &= \Phi_k P_k \Phi_k^T + Q_k. \end{aligned} \quad (\text{A17})$$

Given the needed quantities at time t_k , the measurement z_{k+1} can be assimilated just as in the previous step. Thus it is a routine matter to cycle through the Kalman loop that comprises Eqs. (A8) and (A14)–(A17). The loop begins with the *a priori* estimates of the state vector \hat{x}_0^- and the error covariance matrix P_0^- .

B. Extended Discrete Kalman Filter

Some of the most successful applications of the Kalman filter arise in situations when nonlinear dynamics and measurements have to be considered. This is the actual case when we have a nonlinear measurement model (i.e., the lidar equation). Yet the expressions given next are generalized to potential nonlinearities in both the system and the measurement model. In these instances the stochastic process to be estimated and the measurement relation can be written in their most general forms as

$$x_{k+1} = f_k(x_k) + w_k, \quad (\text{A18})$$

$$z_k = h_k(x_k) + v_k, \quad (\text{A19})$$

where f_k and h_k are nonlinear functions. The same constraints as those given for w_k and v_k in Section A also apply here. In the lidar inversion problem treated here, the system model is linear, and f_k is equivalent to Φ_k [Eq. (20)].

To begin, let us approximate these functions by their Taylor series expansions about the *a posteriori* \hat{x}_k and *a priori* \hat{x}_k^- estimates, respectively, and let us retain first-order terms only. We find that

$$f_k(x_k) \approx f_k(\hat{x}_k) + \left. \frac{\partial f_k(x)}{\partial x} \right|_{x=\hat{x}_k} (x_k - \hat{x}_k), \quad (\text{A20})$$

$$h_k(x_k) \approx h_k(\hat{x}_k^-) + \left. \frac{\partial h_k(x)}{\partial x} \right|_{x=\hat{x}_k^-} (x_k - \hat{x}_k^-). \quad (\text{A21})$$

Note that these developments assume the following conditions:

$$|x_k - \hat{x}_k| \ll 1, \quad (\text{A22})$$

$$|x_k - \hat{x}_k^-| \ll 1, \quad (\text{A23})$$

as it is expected that there will not be large differences among the three variables. Yet careful attention should be drawn to the fact that use of the EKF is risky, as the linearization process takes places about the filter's estimated trajectory of the state vector rather than about a precomputed nominal trajectory (see text). Thus the gain sequence is not predetermined by the process model assumptions as in the linear Kalman filter.

Following a mathematical analysis similar to that in Section A, one can bridge gulfs with the classic linear filter if the equivalent matrices F_k and H_k are defined in the following way:

$$F_k = \left. \frac{\partial f_k(x)}{\partial x} \right|_{x=\hat{x}_k}, \quad (\text{A24})$$

$$H_k = \left. \frac{\partial h_k(x)}{\partial x} \right|_{x=\hat{x}_k^-}. \quad (\text{A25})$$

If Eqs. (A24) and (A25) are identified with the first-order terms of Eqs. (A20) and (A21), they yield

$$x_{k+1} \approx f_k(\hat{x}_k) + F_k(x_k - \hat{x}_k) + w_k, \quad (\text{A26})$$

$$z_k \approx h_k(\hat{x}_k^-) + H_k(x_k - \hat{x}_k^-) + v_k. \quad (\text{A27})$$

Relations (A26) and (A27) represent the linearized version of the filter and look much like Eqs. (A1) and (A2), except that rather than presenting total quantities to the filter they present incremental quantities, which, correlated with Eqs. (A26) and (A27), are

$$\Delta x_k = x_{k+1} - f_k(\hat{x}_k), \quad (\text{A28})$$

$$\Delta z_k = z_k - h_k(\hat{x}_k^-). \quad (\text{A29})$$

In summary, the EKF's recursive equation loop becomes

(1) Update the *a priori* estimate with measurement z_k :

$$\hat{x}_k = \hat{x}_k^- + K_k[z_k - h_k(\hat{x}_k^-)]. \quad (\text{A30})$$

(2) Compute the error covariance for the updated estimate:

$$P_k = (I - K_k H_k) P_k^-. \quad (\text{A31})$$

(3) Project ahead:

$$\hat{x}_{k+1}^- = f_k(\hat{x}_k), \quad (\text{A32})$$

$$P_{k+1}^- = F_k P_k F_k^T + Q_k. \quad (\text{A33})$$

(4) Compute the filter gain:

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1}. \quad (\text{A34})$$

We acknowledge the sponsorship of the Comisión Interministerial de Ciencia y Tecnología (Intermin-

istry Committee for Science and Technology) under grant AMB96-1144-C02-C01.

References

1. A. I. Carswell, "Lidar remote sensing of atmospheric aerosols," in *Propagation Engineering: Third in a Series*, L. R. Bissonnette and W. B. Miller, eds., Proc. SPIE **1312**, 206–220 (1990).
2. G. J. Kunz and G. de Leeuw, "Inversion of lidar signals with the slope method," *Appl. Opt.* **32**, 3249–3256 (1993).
3. F. Rocadenbosch, A. Comerón, and D. Pineda, "Assessment of lidar inversion errors for homogeneous atmospheres," *Appl. Opt.* **37**, 2199–2206 (1998).
4. J. D. Klett, "Stable analytical inversion solution for processing lidar returns," *Appl. Opt.* **20**, 211–220 (1985).
5. R. M. Measures, *Laser Remote Sensing: Fundamentals and Applications* (Krieger, Malabar, Fla., 1992), Chap. 4, pp. 138–145.
6. J. D. Klett, "Lidar inversion with variable backscatter/extinction ratios," *Appl. Opt.* **24**, 1638–1643 (1985).
7. J. D. Klett, "Lidar calibration and extinction coefficients," *Appl. Opt.* **20**, 514–515 (1983).
8. F. G. Fernald, "Analysis of atmospheric lidar observations: some comments," *Appl. Opt.* **23**, 652–653 (1984).
9. Y. Sasano and H. Nakane, "Significance of the extinction/backscatter ratio and the boundary value term in the solution for the two-component lidar equation," *Appl. Opt.* **23**, 11–13 (1984).
10. G. J. Kunz, "Probing of the atmosphere with lidar," in AGARD Conf. Proc. **23**, 1–11 (1992).
11. R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.* 35–46 (1960).
12. H. W. Sorenson, "Kalman filtering techniques," in *Theory and Applications*, Vol. 3 of *Advances in Control Systems*, IEEE Press Selected Reprint Series, H. W. Sorenson, ed. (IEEE Press, New York, 1985).
13. R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering* (Wiley, New York, 1992).
14. B. J. Rye and R. M. Hardesty, "Nonlinear Kalman filtering techniques for incoherent backscatter lidar: return power and log power estimation," *Appl. Opt.* **28**, 3908–3917 (1989).
15. D. G. Lainiotis, P. Papapaskeva, G. Kothapalli, and K. Plataniotis, "Adaptive filter applications to lidar: return power and log power estimation," *IEEE Trans. Geosci. Remote Sens.* **34**, 886–891 (1996).
16. R. J. McIntyre, "Multiplication noise in uniform avalanche photodiodes," *IEEE Trans. Electron Devices* **ED-13**, 164–168 (1966).
17. P. S. Maybeck, *Stochastic Models, Estimation and Control* (Academic, New York, 1977), Vol. 1.
18. F. G. Smith, *Atmospheric Propagation of Radiation*, Vol. 2 of *The Infrared and Electro-Optical System Handbook* (SPIE, Bellingham, Wash., 1993).
19. R. W. Fenn, "Correlation between atmospheric backscattering and meteorological visual range," *Appl. Opt.* **5**, 293–295 (1966).
20. S. Twomey and H. B. Howell, "The Relative merits of white and monochromatic light for the determination of visibility by backscattering measurements," *Appl. Opt.* **4**, 501–506 (1965).
21. F. Spitzer, *Principles of Random Walk* (Van Nostrand, Princeton, N.J., 1964).
22. A. Papoulis, *Probability, Random Variables and Stochastic Processes* McGraw-Hill, New York, 1991), pp. 345–354.
23. J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Monographs on Numerical Analysis (Clarendon, Oxford, 1965).
24. H. Koshmieder, "Theorie der horizontalen Sichtweite," *Beitr. Phys. Freien Atmos.* **12**, 33–53 (1924).
25. P. W. Kruse, L. D. McGlauchlin, and R. B. McQuiston, *Elements of Infrared Technology: Generation, Transmission and Detection* (Wiley, New York, 1962).