

# COMPARISON BETWEEN THE KALMAN AND THE NON-LINEAR LEAST-SQUARES ESTIMATORS IN LOW SIGNAL-TO-NOISE RATIO LIDAR INVERSION

Francesc Rocadenbosch, Michaël Sicard, Adolfo Comerón, M. N. Md. Reba

Remote Sensing Lab (RSLAB), Dep. of Signal Theory and Communications (TSC), Univ. Politècnica de Catalunya (UPC)/IEEC CRAE, C/Jordi Girona, 1-3, D4-016, 08034 Barcelona, SPAIN.  
Email: roca@tsc.upc.edu, Phone: +34 93-401-60-85, Fax: +34 93-401-72-32

## ABSTRACT

This work departs from previously published results of the authors and focus on joint estimation and time evolution of the atmospheric backscatter profile and a range-independent lidar ratio by means of 1) adaptive extended Kalman filtering (EKF) and 2) non-linear least-squares (NLSQ), under moderate-to-low signal-to-noise ratios (SNR < 100 at the starting sounding range). A Rayleigh/Mie atmosphere and a calibrated lidar system are considered. Performance parameters studied are data sufficiency, tracking of the optical parameter time fluctuations, inversion errors, power estimation, and noise impact. The EKF inversion solution is, in turn, compared with Klett's method as a reference. Finally, it is shown that the EKF outweighs the NLSQ in noisy environments.

**Index Terms**—Lidar, inversion, Kalman filter, least-squares.

## 1. INTRODUCTION

Micro-pulse lidars usually operate with 5-40- $\mu$ J energy and kHz repetition rates to achieve 30-to-60-s time resolution and 30-to-75-m spatial resolution using photon-counting detection. In comparison to classic 0.1-1-J energy, 10-50-Hz repetition-rate laser sources this could represent a 2-to-30 reduction factor in the SNR [5]. This motivates research on moderate-to-low SNR inversion methods.

So far, independent inversion of the optical atmospheric parameters of interest, namely the aerosol extinction, the aerosol backscatter, and the lidar ratio has only been tackled by combining at least one elastic and one inelastic Raman channel, multiple zenith-angle elastic returns under the assumption of a homogeneously horizontally stratified atmosphere or by means of a HSRL (High Spectral Resolution Lidar).

In the case of the elastic lidar equation, inversion of the sought-after optical parameters requires both the introduction of "a priori" correlation hypotheses between

the extinction and the backscatter profiles such as a range-dependent aerosol lidar ratio and a boundary calibration (Klett-Fernald-Sasano's (KFS) method [1]). Besides, a temperature/pressure radiosonde measurement is used to estimate the molecular component.

The fundamentals of data sufficiency, i.e., the classical question of how to retrieve *two vector unknowns* (the range-dependent aerosol backscatter and a range-independent lidar ratio) from one single elastic equation are at the basis of all inversion algorithms. Alternatively to KFS, we ensure data sufficiency [3] by introducing the concept of *data decimation* of the backscatter profile and estimate the optical parameters using an EKF [2][4]. This is presented in Sect. 2. Cross-examination of EKF and NLSQ estimation results (and Klett's solution) are discussed in Sect. 3, and conclusions remarks are given in Sect. 4.

## 2. THE ESTIMATION PROBLEM

### 2.1. Problem formulation and data sufficiency

We depart from the single-scattering elastic lidar equation,

$$P(R) = \frac{A}{R^2} [\beta_{aer}(R) + \beta_{mol}(R)] \times \exp\left[-2 \int_{R_{min}}^R [\alpha_{aer}(r) + \alpha_{mol}(r)] dr\right] \quad (1)$$

where  $\alpha$  and  $\beta$  stand for the extinction and backscatter, subscripts "aer" and "mol" stand for "aerosol" and "molecular" components,  $R$  is the range,  $R_{min}$  is the minimum sounding range (e.g., the range of full overlap in a biaxial system), and  $A$  is the *equivalent* system constant,

$$A = A_0 \exp\left\{-2 \int_0^{R_{min}} [\alpha_{aer}(r) + \alpha_{mol}(r)] dr\right\}, \quad (2)$$

which is defined as the product of the system constant ( $A_0$  [ $W \cdot m^3$ ]) times the total transmissivity from  $R=0$  to  $R_{min}$ . The two-way path molecular transmissivity is defined as

$$T(R) = \exp\left[-2 \int_{R_{min}}^R \alpha_{mol}(r) dr\right]. \quad (3)$$

Under the approximation of a range-independent aerosol lidar ratio,  $C$  ( $C = \alpha_{aer}(R)/\beta_{aer}(R)$ ), the sought-after quantities to estimate are the aerosol backscatter,  $\beta_{aer}(R)$ , and the lidar ratio,  $C$ . For data sufficiency (Sect. 1) there must be fewer unknowns than measurements and a decimated version of the aerosol backscatter is introduced. This is to say that the backscatter is estimated in less inversion cells than observation cells (i.e., the power-measurement cells) following a M-to-1 ratio. Thus, if M is the decimation factor and  $\Delta R$  is the spatial resolution of the power observables, only N/M aerosol backscatter components (each of them with an inversion spatial resolution  $M\Delta R$ ) are estimated.

The *state vector* (vector of unknowns) to estimate is

$$x_k \equiv [\beta_1, \beta_2, \dots, \beta_{N/M}, C]^T, \quad (4)$$

where  $k$  is a reminder of discrete (estimation) time  $t_k$ .

The *observation vector* (or measurement vector) is the range-corrected lidar return power (Eq.(1)), at discrete spatial observation cells  $R_i$ ,  $i = 1..N$ ,

$$z_k \equiv [F_1, F_2, \dots, F_N]^T + v_k \quad (5)$$

where

$$F_i = R^2 P(R_i), \quad (6)$$

and  $v_k$  is the range-corrected observation noise.

From Eqs.(1)-(6) above, the observation vector can be related to the state vector (unknowns) in terms of a N-equation set containing N/M+1 unknowns (we drop subscript ‘‘aer’’ and replace ‘‘mol’’ by ‘‘m’’):

$$\begin{cases} F_1 = A(\beta_1 + \beta_{m,1})T_1 \\ \dots \\ F_M = A(\beta_1 + \beta_{m,M})\exp\{-2C[\beta_1(M-1)\Delta R]\}T_M \\ F_{M+1} = A(\beta_2 + \beta_{m,M+1})\exp\{-2C[\beta_1 M\Delta R]\}T_{M+1} \\ \dots \\ F_N = A(\beta_{N/M} + \beta_{m,N})\times \\ \times \exp\left\{-2C\left[\sum_{i=1}^{\frac{N}{M}-1}\beta_i M\Delta R + \beta_{N/M}(M-1)\Delta R\right]\right\}T_N \end{cases} \quad (7)$$

Note that the aerosol backscatter is assumed constant over adjacent spatial inversion cells of length  $M\Delta R$ .

Next, two different approaches are presented to solve Eq.(7) set above at each time  $t_k$  for the state vector,  $x_k$ , given as input the observation set,  $z_k$ ,  $k = 1..K$ .

## 2.2. The Extended Kalman Filter (EKF) approach

In the EKF adaptive approach [2], as long as the filter iterates and assimilates a new measurement,  $z_k$ , the state vector,  $x_k$ , the estimated ‘‘a posteriori’’ error covariance matrix,  $P_k$ , and the Kalman gain,  $K_k$ , are recomputed,

which enables the filter to correct its projection trajectory and to enhance its current estimation of the optical components,  $\hat{x}_k$ . The filter takes into account the actual observables,  $z_k$  (Eq.(5)), a linearized observation model ( $H_{ij}^{(1)} = \partial F_i / \partial \beta_j |_{x=\hat{x}_k^-}$ ,  $H_i^{(2)} = \partial F_i / \partial C |_{x=\hat{x}_k^-}$ ), and ‘‘a priori’’ information provided by the user. The latter is cast into the filter in the form of an initial guess  $\hat{x}_0$ , a state-vector transition matrix  $\Phi_k$  and, three covariance matrices, namely, the state-noise covariance matrix,  $Q_k$ , the initial ‘‘a priori’’ error covariance matrix,  $P_0^-$ , and the observation-noise covariance matrix,  $R_k$ . Details of a similar EKF are found in [3] with the following remarks/corrections: 1) The filter’s state-vector transition model has now been simplified to a random walk ( $\Phi_k = I$ ), 2) the state-noise covariance matrix model assumes now uncorrelated backscatter components, and, therefore, becomes the diagonal matrix,

$$Q = \text{diag}[\sigma_{\beta_1}^2, \dots, \sigma_{\beta_{N/M}}^2, \sigma_C^2]. \quad (8)$$

This new EKF model formulation is simpler and requires just basic inputs from the user’s side.

## 2.3 The non-linear least-squares (NLSQ) approach

In NLSQ, each succeeding inversion of Eq.(7),  $\bar{x} = \hat{x}_k$ , in response to an incoming observation,  $\bar{z} = z_k$ , is completely independent from the others. Formally, we form the objective function  $\Psi(\bar{x})$  and minimise its norm using a Levenberg-Marquardt’s algorithm in Matlab7. In short,

$$\Psi(\bar{x}) = \bar{z} - \bar{h}(\bar{x}) \xrightarrow{lsq} \min_{\bar{x}} \left\{ \|\Psi\|^2 \right\}, \quad (9)$$

where  $\bar{h}(\bar{x}) = [F_1(\bar{x}), F_2(\bar{x}), \dots, F_N(\bar{x})]^T$ .

## 3. SIMULATION RESULTS

### 3.1. EKF vs. NLSQ performance

The simulation considers  $R_{\min}=0.2$  km,  $R_{\max}=4$  km, planetary boundary layer,  $R_{\text{PBL}}=3$  km, and a 532-nm clear atmosphere (typ. aerosol lidar ratio,  $C_a=25$  sr, typ. aerosol backscatter,  $\beta_a=4 \times 10^{-3}$  km<sup>-1</sup>sr<sup>-1</sup>, molecular components following a US standard model). The simulated SNR goes from 100 at  $R_{\min}$  down to 2 at  $R_{\max}$ . With a spatial resolution,  $\Delta R=96.6$  m, and e.g., decimation M=10, the range settings above yield 4 inversion cells (N/M=4) ranging from [0.20- 1.17], [1.17-2.15], [2.15-3.12], and [3.12-4] km (note that the first three backscatter cells,  $\beta_1$ - $\beta_3$ , lay on the boundary layer (PBL) while the last one,  $\beta_4$ , lies on the pure molecular layer above). The initial state is

shown in Fig. 1. A total of 250 iterations are generated, where both the aerosol backscatter and the lidar ratio vary dynamically with time following a Gauss-Markov stochastic model similar to [3], Sect.3.

Both the EKF and NLSQ estimators are initialized with  $\hat{x}_0^- = [\beta_0, \dots, \beta_0, C_0]$ , ( $\beta_0 = \beta_a = 4 \times 10^{-3} km^{-1} sr^{-1}$ ,  $C_0 = 0.5C_a = 12.5 sr$  (this simulates a 50% underestimation from the user's side), and  $A = 6.65 \times 10^{-4} [W \cdot km^3]$  (Eq. (2)), a relatively modest figure for a short-range tropospheric lidar system. Additional EKF parameters are: 1)  $\sigma_{\beta_i} = 10^{-1} \beta_0$ ,  $i = 1..3$ ;  $\sigma_{\beta_4} = 10^{-5}$  in Eq.(8). This tells the filter that we expect 1- $\sigma$  backscatter fluctuations roughly about 10% of the initialization backscatter figure inside the PBL and virtually nil for  $\beta_4$  (molecular region); 2)  $\sigma_C = 10^{-1} C_0$ . These values are not critical within less than a factor 10 approximately provided they impose reasonable "search" limits to the filter.

For the "a priori" error covariance matrix,  $P_0^-$ , we use a similar model to Eq.(8), with initial uncertainties  $\sigma_{P, \beta_i} = \beta_0$ ,  $i = 1..4$ ,  $\sigma_{P, C} = 10 sr$ . Finally, the observation-noise covariance matrix has been initialized with an estimate of the simulated noise variance,  $R_k = diag[\sigma_{n_1}^2, \dots, \sigma_{n_N}^2]$ .

Figs. 2-3 illustrate inversion results. They are very poor for the NLSQ estimator, particularly for the lidar ratio, which is stranded at the initialization value. In clear atmospheres, the lidar ratio is the most difficult parameter to estimate for the two-way path optical thickness leads to transmissivities close to unity. This causes low lidar-ratio sensitivity on the projected observables,  $\hat{z}_k$ . In contrast, the EKF shows outstanding tracking capabilities for all optical parameters. Consider e.g.,  $\beta_4$  failing in the molecular level. After the first 50 iterations the filter tends to zero indicating that  $\beta_4$  is in an aerosol-free layer. It is the convenient way by which the filter combines past estimates and estimated errors with new incoming observables that yields excellent tracking results as time goes on. Now, if  $\beta_1$  is considered, both EKF and NLSQ estimates coincide, for  $\beta_1$  is the nearest inversion cell, therefore, with a high SNR.

Likewise, Figs. 4-5 depict power estimation (i.e.,  $\hat{z}_k$ , Eq.(5)). Again, the de-noising capability of the EKF becomes fairly evident in Fig.4b and Fig.5b.

### 3.2. Comparison with Klett's method

Fig.6 compares last-iteration EKF estimates with Klett-Fernald-Sasano's (KFS) method (calibration range,  $R_{cal} = 3.8 km$ ) [1]. EKF errorbars have been computed from the diagonal of the error covariance matrix,  $P_k$  while KFS errorbars assimilate two error sources: 1) noise propagation

at the calibration range, and 2) 50% user uncertainty in the assumed lidar ratio ( $C = 12.5 sr$  instead of  $C_a = 25 sr$ ). In spite of the slowly time-varying atmosphere all 250 observations ( $z_k$ ) have had to be time-averaged to enable KFS inversion, otherwise virtually unlimited errorbars are obtained.

## 4. CONCLUSIONS

Data sufficiency has been demonstrated in joint dynamic estimation of the aerosol backscatter profile and range-independent lidar ratio from elastic lidar returns using a M-to-1 decimated version of the aerosol backscatter profile.

In clear atmospheres and low SNRs (<100) a simple random-walk EKF outweighs NLSQ estimation and enables good tracking of the sought-after optical parameters, inclusive of the lidar ratio. Main limitations are: 1) the need of a calibrated lidar system (work is under way to supersede this) and 2) an inversion spatial resolution M times poorer than the raw observation resolution. Here we note that Klett's method also requires time averaging (if the scene is stationary) and/or spatial averaging to boost the SNR to suitable levels apt for inversion.

## 5. ACKNOWLEDGEMENTS

European Commission under the EARLINET-ASOS (EC Coordination Action) contract n° RICA-025991, and (EU Specific Support Action) contract n° 011863 (RIDS); European Space Agency (ESA) under the contract no. 21487/08/NL/HE, MCYT (Spanish Ministry of Science and Technology) and FEDER funds under the projects TEC2006-07850/TCM, Complementary Actions CGL2007-28871-E/CLI, CGL2006-26149-E/CLI, and CTM2006-27154-E/TECNO. MCYT is also thanked for the Ramón y Cajal position hold by Dr. M. Sicard, and Generalitat de Catalunya/AGAUR for Mr. Md. Reba's predoctoral fellowship.

## 6. REFERENCES

- [1] J.D.Klett, "Lidar Inversion with variable backscatter/extinction ratios," *Appl. Opt.* **24**, 1638-1643 (1985).
- [2] R.G. Brown, P.Y.C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering* (Wiley, New York, 1992).
- [3] F. Rocadenbosch, C. Soriano, A. Comerón, J.M<sup>a</sup>. Baldasano, "Lidar Inversion Of Atmospheric Backscatter And Extinction-To-Backscatter Ratios By Use Of A Kalman Filter," *Appl. Opt.* **38** (15), 3175-3189 (1999).
- [4] 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).
- [5] S. Lolli, L. Sauvage, I. Stachlewska, R. Coulter, R. Newsom, "Assessment of EZ lidar and ARM/SGP MPL lidar performances for qualitative and quantitative measurements of aerosol and clouds", *Geos. Rsch. Abstracts*, **10**, EGU2008-A-11091, (2008).

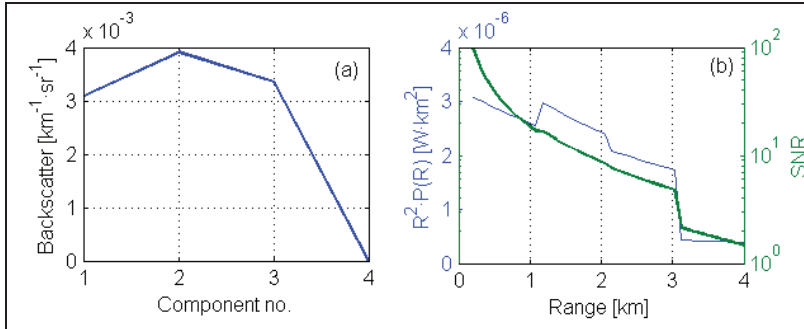


Figure 1. Initial state. (a) Atmospheric backscatter indexed by inversion cell no. (b) Synthesized range-corrected power return and related range-dependent SNR.

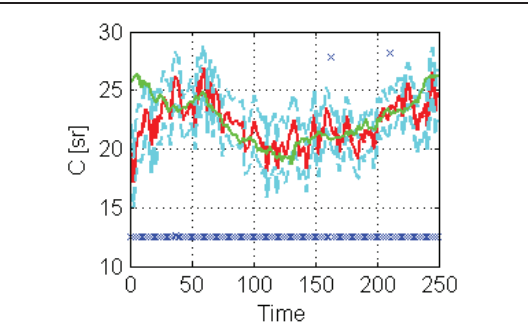


Figure 2. Lidar ratio estimation. (Green) Simulated atmosphere, (red) EKF inversion, (cyan) EKF 1- $\sigma$  errorbars, (blue crosses) NLSQ estimation.

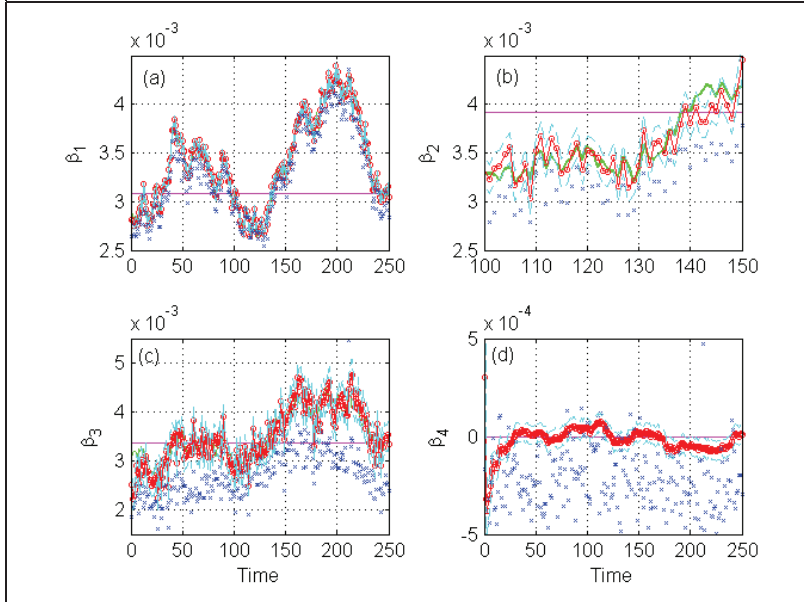


Figure 3. Aerosol backscatter estimation. (a-d) (Green) Simulated atmospheric backscatter dynamics, (magenta) initial atmospheric state as a reference (Fig. 1a), (red) EKF estimates, (blue crosses) NLSQ estimates.

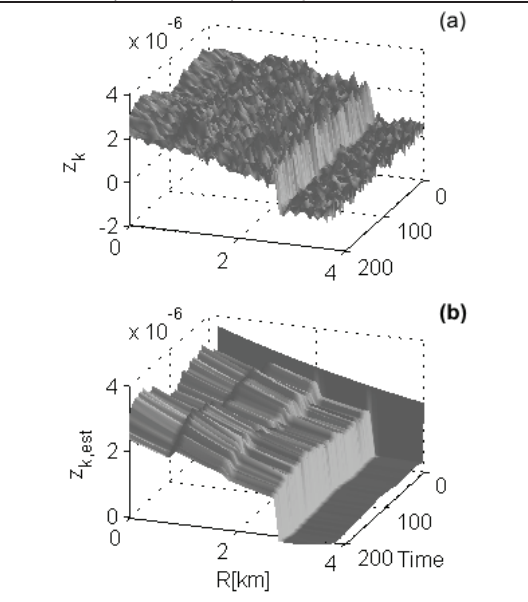


Figure 4. Power return estimation (I). (a) Simulated range-corrected power returns, (b) De-noised EKF estimates (first estimates around time  $t=0-10$  clipped for representation purposes).

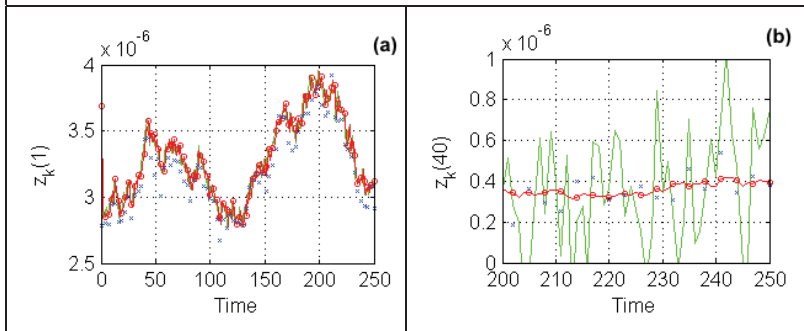


Figure 5. Power return estimation (II). (a-d) (Green) Simulated atmospheric range-corrected power returns, (red) EKF (range-corrected) power estimates, (blue crosses) NLSQ (range-corrected) power estimates.

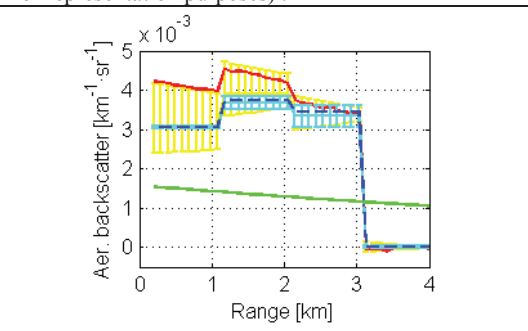


Figure 6. Comparison with Klett-Fernald-Sasano's (KFS). (Red) KFS inversion using the 250 time-averaged returns as input, (yellow) associated errorbars. (Dashed blue) last EKF inversion estimate ( $t=250$ ), and (cyan) associated 1- $\sigma$  errorbars.