

# MONITORING OF DISPERSED SMOKE-PLUME LAYERS BY DETERMINING LOCATIONS OF THE DATA-POINT CLUSTERS

Vladimir Kovalev\*, Cyle Wold, Alexander Petkov, and Wei Min Hao

*Missoula Fire Sciences Laboratory, MT 59808, USA, \*E-mail: vladimirakovalev@fs.fed.us*

## ABSTRACT

A modified data-processing technique of the signals recorded by zenith-directed lidar, which operates in smoke-polluted atmosphere, is discussed. The technique is based on simple transformations of the lidar backscatter signal and the determination of the spatial location of the data point clusters.

The technique allows more reliable detection of the location of dispersed smoke layering. Examples of typical results obtained with lidar in a smoke-polluted atmosphere are presented.

## 1 INTRODUCTION

The worldwide increase of the number of wildfires is considered as a result of global warming. On the other hand, wildfires further influence climate change. Therefore, monitoring wildfires and predicting their possible distribution is an imperative task of atmospheric science. However, the determination of the exact location of the smoke plume is a significant challenge, especially when monitoring dispersed smoke layers and plumes is made in the areas where the dispersion process creates an extended transition zone between regions of polluted and clear air.

Lidar is a good tool for monitoring smoke layer behavior. In principle, it can easily distinguish regions with high levels of backscatter from regions of clear atmosphere and detect the boundary between different atmospheric layers. When determining the boundaries of the aerosol formations with lidar, generally, some relative characteristics of the backscatter signal from areas of interest are used [1–3]. In particular, when using the gradient method, one can find the smoke-plume boundary location as the range where the derivative of the square-range-corrected lidar signal either is maximum or decreases from the maximum down to a fixed, user-defined level [1]. Unfortunately, there is no way to establish an optimal value for this

level, which would be acceptable for different atmospheric situations. The use of other techniques, including the wavelet technique [2], also requires the selection of concrete parameters, and might be a significant issue.

The situation becomes especially challenging when one should define the upper height of the area of increased backscattering, created by dispersed smoke. The boundary of such layering is difficult to define as the conventional criteria generally assume a sharp boundary between the clear sky and the area of increased backscattering. The data-processing method which allows determining the upper height of the region of increased backscattering with poorly defined boundaries is considered in [3].

In our current study, the method in [3] is modified and used together with the application of the cluster principle, which allows more reliable determination of the locations of dispersed smoke boundaries. This methodology can be used both for one-directional and scanning lidar; however, below this method is considered only for profiling the atmosphere with a zenith directed lidar.

## 2 METHODOLOGY

In the study [3], the recorded lidar signal,  $P_{\Sigma}(r)$ , which is the sum of the variable backscatter signal,  $P(r)$  at the range  $r$  and the range-independent offset,  $B$ , is transformed in the auxiliary function  $Y(x)$ . For the zenith-directed lidar, the latter is defined as the function of independent  $x$  in the form,

$$Y(x) = P_{\Sigma}(x)x = [P(x) + B]x, \quad (1)$$

where  $x = h^2$ , and  $h$  is the height. The sliding derivative of this function,  $dY/dx$ , is calculated and the intercept point of each local slope fit of the function with the vertical axis is found as the difference of two terms,

$$Y_0^*(x) = Y(x) - \frac{dY}{dx} x. \quad (2)$$

By using the intercept function instead the derivative of the square-range-corrected backscatter signal, the determination of the systematic offset,  $B$ , can be avoided. The retrieval technique is based on the normalized intercept function, which is defined as,

$$Y_0(x) = \frac{Y_0^*(x)}{x} = \frac{Y(x)}{x} - \frac{dY}{dx}. \quad (3)$$

For determining the upper boundary of the dispersed smoke plume of interest, the absolute value of this difference is calculated,

$$|Y_0(x)| = \left| \frac{Y(x)}{x} - \frac{dY(x)}{dx} \right|, \quad (4)$$

and then transformed into the function of height,  $|Y_0(h)|$ . The maximum value of this function,

$$Y_{0,\max} = \max |Y_0(h)|, \quad (5)$$

within the total height interval from  $h_{\min}$  to  $h_{\max}$  is found, and used for calculating the data points of interest,

$$Y_{0,\chi}(h) = |Y_0(h)| - \chi Y_{0,\max}; \quad (6)$$

here the independent variable  $\chi$  can be chosen in the range  $0 \leq \chi < 1$ , so that  $Y_{0,\chi}(h) \geq 0$ .

The same as in [3], the consecutive values of  $\chi$  with the fixed step  $\Delta\chi = 0.05$  are used in this study; that is,  $\chi_{\min} = 0$ ,  $\chi_1 = 0.05$ ,  $\chi_2 = 0.1$ , etc. For each discrete  $\chi$ , the corresponding height of a smoke layer,  $h_{\max}(\chi)$  was determined; the latter was found as the maximum height where  $Y_{0,\chi}(h)$  was a non-zero value.

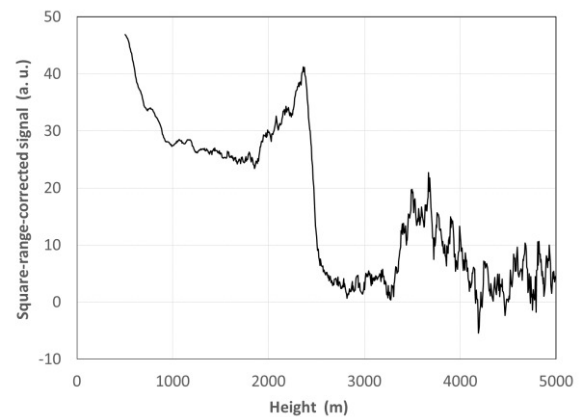
The main challenge is the selection of the optimal level,  $\chi_{opt}$ , and the corresponding height of interest,  $h_{\max}(\chi_{opt})$ . In the previous study [3], such a selection was based on the calculation of the differences between the adjacent heights  $h_{\max}(\chi_{k-1})$ ,  $h_{\max}(\chi_k)$ , and  $h_{\max}(\chi_{k+1})$ . The value of  $\chi_{opt}$  was determined using the following two conditions. First, the difference  $h_{\max}(\chi_{opt} - \Delta\chi) - h_{\max}(\chi_{opt})$  should be maximal. Second, the next consecutive increase of  $\chi$ , that is, selection of  $\chi$  equal to  $\chi_{opt} + \Delta\chi$ , then  $\chi_{opt}$

+  $2\Delta\chi$ , etc., should result in an invariable, or at most, slowly decreasing height  $h_{\max}(\chi)$ .

As mentioned, the method in [3] works properly when the boundary between the layer with increased backscattering and clear air can be properly defined. The situation becomes much more challenging when lidar data are noisy and are obtained in smoke-polluted layers where the dispersion processes create an extended transition zone between the regions of polluted and clear air. In such situations, the use of the modified data-processing technique based on the analysis of the cluster locations of the data points,  $h_{\max}(\chi)$ , and their temporal changes is extremely effective.

### 3 ANALYSIS OF EXPERIMENTAL DATA

To clarify the modified data processing technique, let us consider Figs. 1 and 2, in which real square-range-corrected lidar signal and corresponding dependence of  $h_{\max}(\chi)$  on  $\chi$  are respectively shown. Here and below we will consider data, obtained from the lidar backscatter signals measured close to the zenith direction ( $80^\circ$ ); these signals were obtained on August 4, 2016 near Missoula, Montana, during lidar profiling of the atmosphere influenced by three wildfires: Roaring Lion Fire [4], Cedar Fire [5], and Observation Fire [6].



*Fig. 1. The square-range corrected signal measured in the atmosphere influenced by three closely located wildfires.*

The square-range corrected lidar signal shown in Fig. 1 is not extremely noisy, so that the use of gradient methods for determining the height of smoke layering could yield acceptable results within the height up to  $\sim 3000$  m. However, it could be difficult to discriminate the useful spikes caused

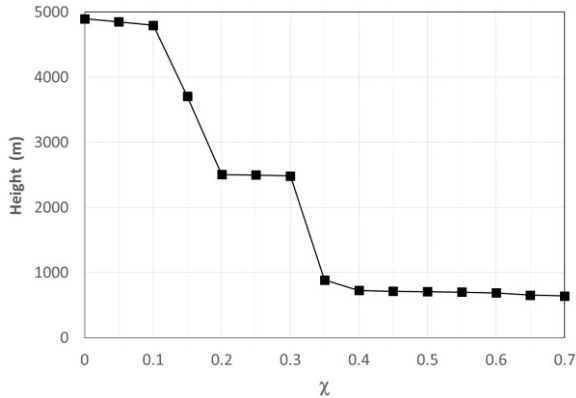


Fig. 2. Dependence of the height  $h_{max}(\chi)$  on  $\chi$  extracted from the signal in Fig. 1. The data points close to the height 5000 m are originated by noise.

by the actual change in the backscattering signal from the spikes originated by random noise over the larger heights. Using the criteria given in [3] and analyzing the corresponding dependence  $h_{max}(\chi)$  on  $\chi$  in Fig. 2, one can overcome this issue, but only for heights up to  $\sim 3000$  m. In particular, one can easily distinguish two distinct layers, which upper heights are located at  $\sim 700 - 800$  m and  $\sim 2500$  m. However, the origin of the data point at  $h \approx 3700$  m at  $\chi = 0.15$  is unclear. Moreover as mentioned, the dependencies of  $h_{max}(\chi)$  on  $\chi$  that obey the criteria in [3] are not generally obtained when the smoke

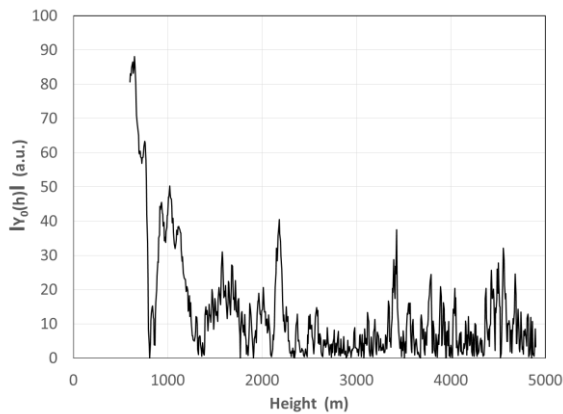


Fig. 3. The typical function  $|Y_0(h)|$  versus height obtained from a noisy signal.

plume is strongly dispersed and the backscatter signals from that area are too noisy. Such a very typical case is shown in Figs. 3 and 4, where the

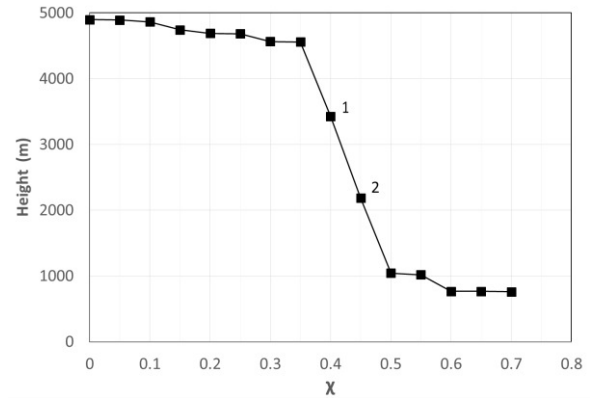


Fig. 4. Dependence of the height  $h_{max}(\chi)$  on  $\chi$  extracted from the function  $|Y_0(h)|$  in Fig. 3.

function,  $|Y_0(h)|$  and the corresponding dependence  $h_{max}(\chi)$  on  $\chi$ , obtained on the same day, August 4, 2016, are presented. Unlike Fig. 2, in which one can easily discriminate two separated layering heights, the same criteria do not allow reliable discrimination of smoke-layer heights in Fig. 4. In particular, it is unclear whether points 1 and 2 which correspond to  $\chi = 0.4$  and  $\chi = 0.45$  are originated by the noise or by the backscatter signal. Note that eight points, located in the vicinity of maximum height,  $h_{max} \approx 5000$  m, are originated by increased noise level at the far end of lidar range. This noise is more intense than the noise in Fig. 2, where the noise data points are obtained within the range of  $\chi$ , from 0 to 0.15, whereas in Fig. 3, this range extends from 0 to 0.35.

Thus, in a common case, the data points in the dependencies  $h_{max}(\chi)$  on  $\chi$  can be originated from three sources: (1) the data points originated from the random noise on the signal far end; these data points are located at or close to the maximum height,  $h_{max}$ ; (2) the data points originated from the change in the backscatter signal on the upper boundary of smoke area, and (3) the data points, whose source is unclear; they can originate from either the first or the second source.

To solve the problem of discriminating the backscatter and noise data points, the auxiliary cluster principle can be used. To clarify this principle, let us initially consider dependencies  $h_{max}(\chi)$  on  $\chi$ , obtained for three consecutive time periods, at local times 15:30, 15:31, and 15:34 (Fig. 5). In all these curves, the boundary of the increased

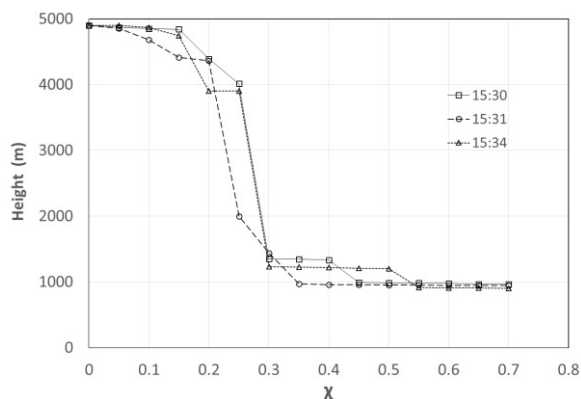


Fig. 5. The dependence of the height  $h_{max}(\chi)$  on  $\chi$  obtained from the signals measured at 15:30, 15:31, and 15:34 local time (MST).

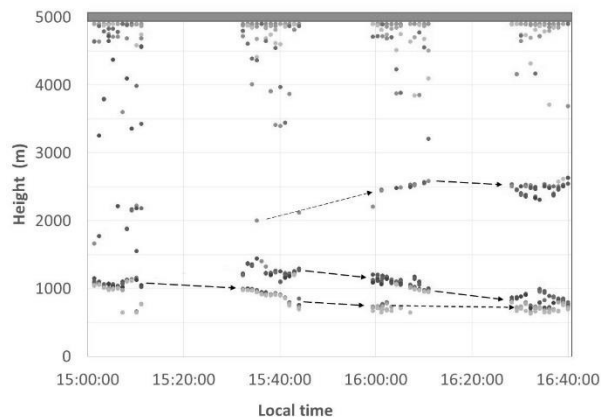


Fig. 6. Data points  $h_{max}(\chi)$  versus time for the period from 15:00 to 16:15 of local time.

backscatter layer within the height interval ~1000-1300 m is clearly seen. The same as in the previous figures, the far-end data create noise data points, located close to maximum height,  $h_{max} = 5000$  m. The data point at the height 2000 m can be seen only at 15:31; therefore, this point also most likely originated from random noise. The origin of the three data points close to ~ 4000 m is unclear and cannot be clarified without additional analysis of the data points obtained before 15:30 and after 15:34. This principle of separation between the casual and systematically appearing data points in the two-dimensional dependence of data points  $h_{max}(\chi)$  on time and locating clusters of such data points is the basis of the cluster method.

To illustrate this method, in Fig. 6 the set of data points  $h_{max}(\chi)$  versus time for the extended period from 15:00 to 16:15 of local time is given. One can discriminate three separate layers in this figure: (1) the lower layer, which height decreases with time from 1000 m to ~600 m; (2) the layer, which during the period 15:32 – 15:44 is located at the heights, ~ (1200 - 1400 m) and then goes down to ~700 m; and (3) the upper layer which appears at 15:30 at the height ~2000 m and then rises. After 16:30 and until 16:40 (the termination of our measurements), this layer at the height ~2500 m was reliably fixed. The scattered data points at the heights larger than 3000 m originated from signal random noise.

#### 4 SUMMARY

The modified lidar data-processing methodology is presented which allows more reliable determination of the upper heights of significantly dispersed smoke layering in presence of increased signal noise.

#### References

- [1] L. Menut, C. Flamant, J. Pelon, and P. H. Flamant, 1999: Urban boundary-layer height determination from lidar measurements over the Paris area, *Appl. Opt.* **38**, 945-954.
- [2] I. M. Brooks, 2003: Finding boundary layer top: Application of a wavelet covariance transform to lidar backscatter profiles, *J. Atmos. and Oceanic Technol.* **20**, 1092-1105.
- [3] V. Kovalev, A. Petkov, C. Wold, and W. M. Hao, 2011: Lidar monitoring of the regions of intense backscatter with poorly defined boundaries”, *Appl. Opt.* **50**, 103-109.
- [4] <http://inciweb.nwcg.gov/incident/4913/>.
- [5] <http://inciweb.nwcg.gov/incident/4874/>.
- [6] <http://inciweb.nwcg.gov/incident/4819/>.