Performance of Frequency Estimators for real time display of high PRF pulsed fibered Lidar wind map

Laurent Lombard, Matthieu Valla, Guillaume Canat, Agnès Dolfi-Bouteyre
Onera, The French Aerospace Lab, Palaiseau, France
Laurent.lombard@onera.fr

Abstract: Long range wind lidars based on high pulse repetition frequency (PRF) fiber laser require Doppler mean frequency estimators (MFE) that are reliable up to very low CNR, and fast enough for real time display of long range wind maps. A comparison of MFEs has been performed on accumulated spectra computed on experimental data from a high PRF pulsed fibered Lidar. MFEs (maximum, centroid, matched filter, maximum likelihood, polynomial fit) have been compared in terms of average error versus CNR and in terms of processing time. The wind speed errors have been estimated by the wind data spectrum method for various estimators.

Keywords: coherent Doppler lidar, signal processing.

1. Introduction
Performance of coherent Doppler Wind lidars can be quantified, for a given measurement time (or averaging time) by its range limit and its wind measurement accuracy for each range gate. Those numbers depend on the laser power, the lidar optical architecture, electronics and acquisition/digitalization set up, and finally on signal processing. The most recent progresses have been obtained increasing laser output power. However, improvements can be also obtained by adapting signal processing software to high PRF laser.

Several types of algorithms are used to process coherent lidar wind experimental data. Because real time display of wind maps is mandatory, simple and/or quick mean frequency estimators are envisaged. In this paper, 5 different algorithms have been implemented and compared on the same set of lidar data (maximum, centroid, parabolic fitting, maximum likelihood, and adapted filtering). The Cramer-Rao bound (theoretical best performance) is also computed. Data were acquired with a middle range lidar (1.5-2km) but results can be applied to long range lidar.

2. Lidar data description
The following table gives the main useful lidar characteristics for the data set analysis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse duration</td>
<td>400 ns</td>
</tr>
<tr>
<td>Sampling rate $F_s$</td>
<td>500 MHz</td>
</tr>
<tr>
<td>Pulse repetition frequency PRF</td>
<td>14 kHz</td>
</tr>
<tr>
<td>Total number of acquired points per pulse</td>
<td>12 288</td>
</tr>
<tr>
<td>Intermediate frequency $F_{if}$</td>
<td>120 MHz</td>
</tr>
<tr>
<td>FFT length $N$</td>
<td>512</td>
</tr>
<tr>
<td>Maximum range</td>
<td>3500 m</td>
</tr>
<tr>
<td>Zero padding factor</td>
<td>1 or 4</td>
</tr>
<tr>
<td>Reduced bandwidth for frequency analysis</td>
<td>[90 150] MHz</td>
</tr>
</tbody>
</table>

3. Algorithms description
The backscattered signal is acquired for each pulse. A power spectral density (PSD) is computed for each distance and laser shot using a discrete Fourier transform (DFT) with zero padding. The PSD is the square of the DFT. PSD analysis includes an estimation of the CNR, defined as the ratio of the power contained in the signal to the power contained in the noise over the full detector bandwidth, and an estimation of the wind speed through one of the 5 estimators.
Zero Padding

Frequencies in the DFT are spaced at intervals of $F_s/N$, where $F_s$ is the sample rate and $N$ is the number of samples in each input time serie. Frequency estimation by peak search results in accuracy limited by frequency resolution. Zero-padding data before performing DFT artificially increases the frequency resolution and often improves the peak search estimate. While zero-padding does not improve the frequency resolution of the DFT, the resulting interpolation is smoother and the peak search more precise.

PSD accumulation

Fiber laser operates at high pulse repetition frequency (PRF), allowing an averaging over a large number of spectra in order to reduce speckle noise. For the analyzed lidar data, PRF is equal to 14 kHz and 1024 spectra are averaged on a current base corresponding to a measurement time of 70ms. This enables very efficient noise reduction. Figure 1 shows an example of averaged PSD at very low CNR (-23 dB) for an averaging over 32, and for an averaging over 1024. The average noise level is also displayed (yellow). Zero padding is equal to 4.

![Figure 1: Example of PSD accumulation](image)

Maximum estimator

Maximum estimator or peak search is the simplest estimator and consists in selecting the point in the average PSD with the highest power density. The accuracy of the algorithm directly depends on the zero padding level. All the following algorithms except for the adapted filter take as a starting point the result of the maximum estimator.

Centroid or barycenter:

Computing the moments of the average PSD gives access to the CNR (zero order moment), Doppler frequency (first order moment) and wind dispersion (second order moment after rectification from the frequency broadening due to the pulse length and DFT window size).

Parabolic fitting:

For Doppler lidar with short pulse (<500 ns) in laminar wind conditions (wind shear and wake vortex conditions are excluded), and assuming a Gaussian pulse shape and a Gaussian DFT window, the spectral shape of the theoretical PSD is close to Gaussian. The logarithm of the average PSD is a parabola that can be adjusted to the measured PSD. The fitted parabola center, integral and FWHM give the Doppler frequency, the CNR and the wind dispersion, respectively.

Maximum likelihood:

The maximum likelihood estimator (MLE) has been adapted to pulsed lidar [8] [1]. A model of the lidar signal is necessary for MLE. We have chosen a signal model based on the “feuilleté” model [2] for laminar wind condition, taking into account the shape of the lidar pulse. The signal model has three free parameters: the CNR, the Doppler shift and the spectral width. The MLE consists in finding the value of
the parameters which maximizes the joint probability of the signal. More details are available in [1] and [3].

Adapted filter:
An adaptive filter for frequency estimate of heterodyne Doppler lidar returns has been proposed by Zarader [4]. Further implementation was proposed by Dabas [5]. This adaptive filter removes the atmospheric contribution from noise. For fiber laser with a high PRF, PSD accumulation has already removed a large quantity of noise, and the algorithm is less effective than for low PRF lidar.

Outliers filtering:
All those algorithms are limited by outliers that appear when the highest peak in the spectrum becomes close to the noise level. An effective way to reduce the outliers occurrence is to increase the PSD accumulation. Figure 2 shows the evolution in limit range (maximum distance without outlier) when averaging 1024 (70 ms), 4096 (0.28s) and 65536 spectra (4.6 s). X axis is time and Y axis is range (m). However increasing the PSD accumulation decreases the data refreshment rate. Note that another commonly used method to reduce outliers occurrence is to search Doppler frequency within a limited bandwidth around heterodyne frequency.

As expected, the maximum range without outlier increases with number of accumulated PSD $M$ (figure 2). We have established a relation between CNR value and required $M$ to ensure a given acceptable outlier ratio: $\text{CNR} \cdot \sqrt{M}/2 = A$. $A$ is a constant that is experimentally evaluated. $A$ has been set to ensure zero outliers on our dataset. Figure 3 left shows the corresponding relation and the experimental maximum range versus $M$. The prediction fits well the data. Figure 3 right shows the increase of limit range with $M$ for a given outlier ratio.

Figure 2: Color velocity map evaluated with centroid algorithm, for 1024 (left), 4096 (center) and 65536 (right) accumulated PSD

Figure 3: Left : Theoretical number of accumulated PSD ($M$) required for outliers prevention for a given CNR (blue curve). Experimental CNR at limit range (red crosses). Right: outliers ratio versus range for different $M$
The increase in maximum range with $M$ is however limited by the coherence time of wind speed related to atmospheric stability. Indeed, wind velocity variation within the averaging time can degrade PSD contrast. The maximum accumulation time, for a given range and a given atmosphere, can be obtained by drawing the PSD of wind data (Figure 4). Indeed, wind data fluctuation is the consequence of wind turbulence, and follows a Kolmogorov law: its PSD is proportional to $f^{-5/3}$ [6]. Figure 4 shows the PSD of the wind data set at 1300m range for averaging times corresponding to Figure 2. On the PSD log-log graph we can visualize the $-5/3$ slope of the wind turbulence, and the flat level of the estimator noise [7]. To obtain the estimated wind speed error, the noise floor level is integrated over the full bandwidth of the measurement (equal to $PRF/M/2$). The position of the PSD elbow gives the maximum accumulation time (inverse of frequency cut-off) beyond which no accuracy is gained. It is dependent on the range and the atmosphere state. Information on wind is lost when averaging above this limit (as shown on figure 4 right).

![Figure 4: wind DSP for various accumulation time: 70 ms (left), 0.28 s (center) and 4.6 s (right)](image)

4. **Algorithm comparison**

The estimated wind speed error is computed with the above procedure for each range gate and for the 5 estimators using the PSD of the velocity data (see Figure 4) in order to compare the estimator accuracy.

![Figure 5: Estimators accuracy as a function of CNR without (left) and with (right) zero padding for 70 ms accumulation time](image)

The results are shown in Figure 5 without (left) and with (right) zero padding as a function of CNR. A turbulent event appears for the range gate corresponding to CNR -23 dB and degrades all estimators variance. The theoretical performance of the wind speed estimator, given by Cramér-Rao bound [1][3], has been computed from CNR estimated values and is illustrated by the dash blue line in Figure 5.
The parametric estimators (polynomial fit and MLE) do not require zero padding to reach good accuracy (within Cramér-Rao bound accuracy). On the contrary, non-parametric (maximum and centroid) estimators are improved with zeros-padding: centroid estimator can even reach the same level of accuracy as the parametric estimators. Filter estimator performances is comparable with MLE and polynomial fit performances.

Regarding computing time, maximum, centroid and polynomial fit estimators are much less computing intensive than MLE estimator. However, most of the time is dedicated to spectra computing and accumulation. The spectrum computing time even increases with zero-padding and with laser PRF (high in our case). In the following table, we reported performances and overall computing time (including spectra computing and wind estimation).

<table>
<thead>
<tr>
<th>Estimation</th>
<th>Computing time</th>
<th>Accuracy with zero padding</th>
<th>Accuracy without zero padding</th>
<th>Overall computing time for good accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>Short</td>
<td>medium</td>
<td>bad</td>
<td>N/A</td>
</tr>
<tr>
<td>Centroid</td>
<td>Short</td>
<td>Good</td>
<td>Medium</td>
<td>long</td>
</tr>
<tr>
<td>Polynomial fit</td>
<td>Short</td>
<td>Good</td>
<td>Good</td>
<td>shortest</td>
</tr>
<tr>
<td>MLE</td>
<td>Long</td>
<td>Good</td>
<td>Good</td>
<td>longest</td>
</tr>
<tr>
<td>Adapted filter</td>
<td>medium</td>
<td>Good</td>
<td>Good</td>
<td>medium</td>
</tr>
</tbody>
</table>

In our high PRF lidar case and for laminar wind field analysis, the best candidate is polynomial fit estimator on spectra computed without zero-padding.

5. Conclusion

For the coherent wind lidar, we have established a relation between the accumulation time and the CNR value to ensure a given acceptable outlier ratio. This relation has been verified experimentally and can be used to optimize the lidar data processing.

The performances of five common Doppler frequency estimators have been estimated and their total processing time compared. In our high PRF lidar case and for laminar wind field analysis, the best candidate is polynomial fit estimator on spectra computed without zero-padding. In the case of more complex wind field analysis (eg. Wake-vortex), zero-padding or more complex parametric estimators are required [9].

6. References