# Wind retrieval error characterisation in best-case conditions for the study of a spaceborne infrared coherent lidar in Japan.

P. Baron[1], S. Ishii[1], K. Okamoto[2], K. Gamo[3], K. Mizutani[1], C. Takahashi[3], T. Itabe[1],
T. Iwasaki[4], T. Kubota[5], R. Oki[5], D. Sakaizawa[5], M. Satoh[6], Y. Satoh[5] and M. Yasui[1]

[1]NICT, [2]MRI, [3]Fujitsu FIP Corporation, [4]Tohoku University, [5]JAXA, [6]The Univ. of Tokyo

email: baron@nict.go.jp

**Abstract**

A feasibility study for wind observations with a spaceborne infrared coherent lidar is conducted in Japan. Here we present a study to characterise the line-of-sight wind retrieval errors for the considered instrument characteristics and for the best case conditions, e.g. homogeneous atmosphere and a spectrum broadening dominated by the laser pulse width. The model for simulating time sequences of the atmospheric signal is presented as well as the method for retrieving the Doppler frequency. A numerical expression of the retrieval errors is derived for various signal-to-noise ratios (SNR), spectral and spatial resolutions. The SNR and spatial resolution trade-off for good quality retrievals is shown. Future studies for a comprehensive error analysis are discussed in conclusion.

## 1 Introduction

Wind profile measurements from space have mostly been performed in the middle and upper atmosphere (altitude >20 km) using passive techniques (see references in [1]). In order to respond to the strong demands for global observations in the troposphere [2], the European Space Agency (ESA) will launch in 2015 the Atmospheric Dynamics Mission Aeolus equiped with an UV lidar [3]. In Japan, a working group is studying a spaceborne coherent Doppler lidar operating in the infrared region [4] while the National Aeronautics and Space Administration (NASA) is looking at a dual system using both infrared and visible lidars [2, 5].

The Japanese effort includes the implementation of the Integrated Satellite Observation SIMulator for Coherent Doppler Lidar (ISOSIM) in order to support with numerical simulations, the definition of the instrument and of the observation strategy. The simulator allows us to compute realistic measurements and wind retrievals along a series of orbits using wind and aerosol 3-D fields from atmospheric models. The impacts of the observations on atmospheric models are assessed using the simulated measurements and an Observing System Simulation Experiment (OSSE) [6]. In parallel, a simplified model has been implemented which does not include the atmosphere and observation details has been implemented. This model simply reprduced time series of the signal with the expected statistical and spectral properties. It is well suited for assessing the uncertainties arising from observational parameters on a case-by-case basis.

The objectives of this paper are twofold. First the algorithms for generating a time sequence of the atmospheric signal and for retrieving the Doppler frequency induced by the line-of-sight wind are described. These algorithms are shared by ISOSIM and the simplified model. Second, the retrieval error are characterised for the considered instrumental settings and for the best observational conditions: homogeneous atmsophere and signal spectrum dominated by the transmitted pulse width.

## 2 Observational characteristics

The line-of-sight wind ($v$) is derived from the Doppler shift ($F_d$) of the narrow aerosol spectral line (width $\approx 1$ MHz): $F_d = 2v/\lambda$ with $\lambda$ the signal wavelength. The use of the heterodyne technique allows us to resolve the spectral line and to measure its position. The target is a precision of 1 ms$^{-1}$ for the retrieval of the horizontal wind. Note that because of the nadir angle between 30–40°, the line-of-sight velocity is about half that of the horizontal wind. Two lidar technologies, both operating at infrared wavelengths are being considered. The main instrumental and observational characteristics are given in the tables here below.

| Laser $\lambda = $ **2.1 (1.5)** µm | |
|---|---|
| Frequency $\Leftrightarrow$ velocity | 1 ms$^{-1}$ $\equiv$ 1 (1.3) MHz |
| Pulse energy | 125 (10) mJ |
| Frequency repetition (PRF) | 30 (2500) Hz |
| Pulse power profile FWHM ($\tau$) | 200 ns |
| Power spectrum FWHM ($\varpi = 2\ln 2/\pi\tau$) | 2.2 MHz |

**Observation characteristics**

| | |
|---|---|
| Telescope diameter | 40 cm |
| Satellite altitude | 180 to 400 km |
| Nadir angle | 30–40 ° |
| Satellite velocity wrt surface | 7.7 km s$^{-1}$ |
| Number of Pulses over | ≈32500 for PRF= 2500 Hz |
| 100 km ground-track | ≈390 for PRF= 30 Hz |
| Line-of-sight velocity precision (target) | <0.5 m s$^{-1}$ |

**Processing characteristics for M=256**

| | |
|---|---|
| Sampling frequency, $F_s$ | 400 MHz |
| Spectral Resolution, $\Delta_M$ | 1.5625 MHz |
| Range gate observation time | 640 ns |
| Vertical resolution (35° nadir angle) | 63 m |
| $\Omega = M\varpi/(2\sqrt{2\ln 2}\,Fs)$ | 0.6 |

In order to detect a line-of-sight wind between $\pm 100$ ms$^{-1}$, a receiver bandwidth ($B$) of 200 MHz (sampling freqency $F_s = 2B = 400$ MHz) at $\lambda = 2.1$ µm is required. The local oscillator and the transmitted pulse frequencies are separated by 100 MHz so that the frequency of the signal without Doppler shift is in the middle of the receiver bandwidth (100 MHz). Figure 1 shows the atmospheric signal over a time window of 1 µs. A series of speckles due to the time correlation between signal samples are seen. The correlation time (coherence length) is twice that of the laser power pulse: $\tau_c = 2\tau = 400$ ns. An elementary range gate composed of $M$ consecutive samples is used to derive a spectrum. The single range spectrum is most of the time noisy and requieres to be averaged with other ones to obtain good quality retrievals. The observation time for $M = 64$ and 256 is 155 and 620 ns, respectively. The set of spectra to be averaged are taken along the line-of-sight (same pulse) and along the orbit track (different pulses). The former case defines the vertical resolution while the latter one defines the horizontal one. In this work we consider that the line-of-sight wind measurements performances should be better than a precision of 0.5 ms$^{-1}$ and a spatial resolution (horizontal,vertical) of 100x0.5 km.



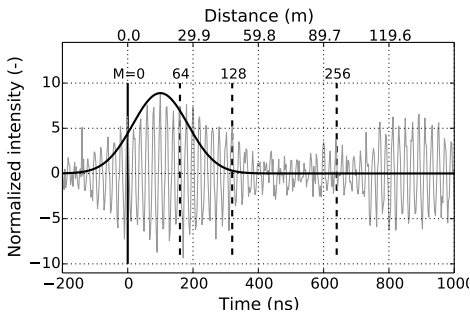Figure 1: Simulated signal with a sampling frequency $F_s$=400 MHz, a Doppler frequency of $F_d$=50 MHz and a SNR=10. The intensity is normalised in order to get a noise power equal to one. The thick black line is the shape of the transmitted pulse power profile (FWHM $\tau$=200 ns).

## 3 Model description

The heterodyne signal is computed using a widely used model [7–9]. The discrete real-valued time series is the $M$-elements vector $z_i$:

$$z_i = \sqrt{2}s_i \cos(2\pi i \frac{F_d}{F_s} + \phi_i) + \Xi_i. \qquad (1)$$

Here $i/F_s$ is the elapsed time between the transmission and the measurement, $s_i$, $\phi_i$ and $\Xi_i$ are zero-mean random and mutually independent parameters with a log-normal, uniform and Gaussian distribution, respectively. The parameter $\Xi_i$ describes the detector noise, and $s_i$ and $\phi_i$ results from the integration of the light backscattered by the large number of scatterers randomly distributed in the illuminated volume at time $i$. The phase $\phi$ varies between 0 and $2\pi$.

The Doppler frequency is assumed to be constant in the analysed time window as well as the ensemble average of the atmospheric signal and of the noise power: $\langle s_i^2\rangle=$ S and $\langle\Xi_i^2\rangle=$ N. Considering that the detector noise is due of the shot noise of the local oscillator, the signal to noise ratio (SNR) can be written as [8]:

$$\text{SNR} = \frac{\text{S}}{\text{N}} = \eta\frac{\lambda\text{P}_L}{Bhc},$$

where $\text{P}_L$ is the power of the atmospheric signal which is propotional to the laser pulse energy, $\eta$ is the heterodyne efficiency, $h = 6.6261\times10^{-34}$ Js is the Planck constant and $c = 2.9979 \times 10^8$ ms$^{-1}$ is the speed of light in vacuum. The theoretical Power Spectral Density (PSD) is given by:

$$\text{P}(\nu_i) = \frac{N}{F_s} + \frac{\sqrt{\ln 2}\,S}{\sqrt{\pi}\,\varpi}\,(g(\nu_i + F) + g(\nu_i - F)) \quad (2)$$

where $g$ is a Gaussian function with a FWHM of $\varpi$. The power spectrum is estimated from M temporal samples using the equation:

$$\widehat{P}_M(\nu_i) = \frac{|\text{FFT}(z)|^2}{M\,F_s} \qquad (3)$$

The signal $z_i$ is generated in three steps. (**1**) A theoretical power spectrum is generated (Eq 2) with a spectral resolution $F_s/(2M)$. (**2**) A complex valued Fourier spectrum is produced with $2M$ random coefficients $\epsilon_i$ so that $\langle\mathbb{R}(\epsilon_i)^2\rangle=\langle\mathbb{I}(\epsilon_i)^2\rangle=$ P$(\nu_i)/(2MF_s)$ and $\langle\mathbb{R}(\epsilon_i)\rangle = \langle\mathbb{I}(\epsilon_i)\rangle = 0$. Here $\mathbb{R}$ and $\mathbb{I}$ denote the real and imaginary components of the complex number. Coefficients with positive frequency are independent and the ones in the negative frequency range follow the symmetrical properties associated with a real-valued temporal signal. (**3**) The temporal sequence of M elements is extracted from the $2M$ elements of IFFT($\epsilon$) (IFFT is the inverse Fourier transform). Finally the mean intensity of the time sequence is set to zero.

## 4 Retrieval method and error characterisation

The wind retrieval is performed in several steps. First the analysed time windows (M samples each) are extrapolated with zero values and the power spectrum is

computed (Eq. 3). The number $M_p$ of zero-padding is determined by the spectral resolution needed to reach the desired precision for the retrieved line-of-sight velocity (better than $0.5 \text{ m s}^{-1}$). The spectra are averaged to obtain a smooth spectrum $\widehat{P}_{M+M_p,N_a}(\nu_i)$, with $N_a$ the number of accumulated spectra.

In a second step, an estimate $\widehat{F_d}$ of the Doppler shift is derived using the periodogram maximum likelihood estimator [8]. The log-likelihood function $L(\nu_k)$ is computed in the positive frequency domain $((M + M_p)/2$ frequencies) as:

$$L(F_k) = \frac{-2}{(M + M_p)} \sum_{i=0}^{(M+M_p)/2-1} \frac{\widehat{P}_{M+M_p,N_a}(\nu_i)}{\mathsf{P}(F_k, \nu_i, \mathsf{S}, \mathsf{N}, \varpi_M)}. \tag{4}$$

Here the width $\varpi_M$ is the pulse spectral width ($\varpi$) multiplied by $\eta$, a constant to account for the spectral broadening due to the limited observation time. We use $\eta = 2.0, 1.6$ and $1.2$ for $M$=64, 128 and 256, respectively.

In the final step of the processing, the frequency retrieval is improved with a small computational cost, by using a second-order polynomial fit of the maximum of $L_p$ and its two adjacent points:

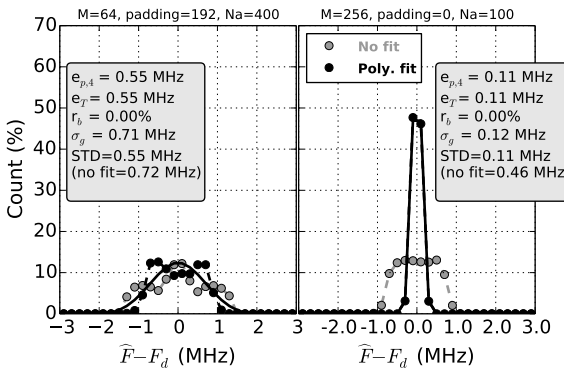$$\widehat{F_d} = \frac{F_s}{M+M_p} \left( p + \frac{L_{p-1}-L_{p+1}}{2(L_{p-1}+L_{p+1}-2L_p)} \right). \tag{5}$$



Figure 2: Histogram of the Doppler frequency retrievals for cnr=0.05 and M=64 (left panel) and M=256 (right) panel. The total number of samples ($N_a \times M$) and the frequency sampling (1.5 MHz) are the same. No significant differences are found between M=128 and 256 (not shown).

An ensemble of retrievals is composed of bad estimates uniformly distributed over the positive frequency range ($B$), and of good estimates with a Gaussian probability function centered at the expected value $F_d$ [8]. Because of the bad estimates, the ensemble average of the estimates is biased toward the center of the spectral band:

$$e_a = \left\langle \widehat{F_d} \right\rangle - F_d = (B/2 - F_d) \times r_b, \tag{6}$$

where $r_b$ is the ratio of the bad estimates. The total random error $e_T$ is derived from the mean-square retrieval errors (same as Eq. 36 in [8] but for the spectral range $B = Fs/2$):

$$e_T^2 = \frac{r_b}{3} \left( B^2 - 3F_d(B - F_d) \right) + (1 - r_b)\sigma_g^2, \tag{7}$$

with $\sigma_g$ the standard deviation of the good estimates. The total error can also be written as $e_T^2 = \sigma_e^2 - F_d(2e_a - Fd)$ where $\sigma_e$ is the standard deviation of all the retrievals.

If the outliers outside the range $F_d \pm \Delta/2$ can be filtered out, the retrievals become unbiased and the precision $e_{p,\Delta}$ is:

$$e_{p,\Delta}^2 = \frac{r_b\Delta^3}{12B} + (1 - r_b)\sigma_g^2. \tag{8}$$

This equation is similar to Eq. 7 but with $B = \Delta$ and $F_d = B/2$ (the spectral band is centered on the expected value of the Doppler frequency). Such filtering is done for the OSSE conducted in this project since the retrieved winds further than $\pm 2 \text{ ms}^{-1}$ from the background wind are rejected.

In this work, we characterise the precision with $e_{p,4}$ ($\Delta = 4$ MHz) and the retrievals with $r_b$ <50% are considered as good retrievals. The number of 0-padding is defined such that M+$M_p$=256. Under such conditions, the standard deviation of good estimates $\sigma_g$ is smaller that 1 MHz and well approximated by $e_{p,4}$ ($\pm 10\%$). Figure 2 shows the histogram of the retrieved Doppler frequencies. The Doppler frequency varies between $50 \pm 10$ MHz from one set of averaged spectra to another. The histogram is computed before and after the polynomial fit (Eq. 5). The other simulation parameters are given in the figure legend. The values of $\sigma_g$ is derived from fitting the histogram with the theoretical distribution function and the number of bad estimates is derived from the equation $r_b = \frac{N_b/N_T}{B/(B-\Delta_f)}$ with $N_T$ is the total number of retrieved winds, $N_b$ is the number of retrieved values outside the range $50\pm\Delta_f$ MHz and $\Delta_f = 15$ MHz.
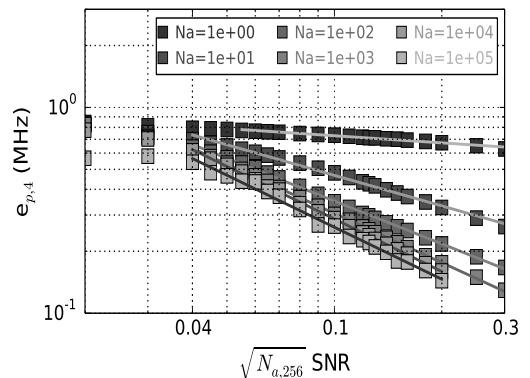


Figure 3: Retrieval precision wrt SNR and the number of accumulated spectra. Computations are for M=256. The fit with straight lines where $r_b < 50\%$ and for same $N_a$ are shown (full line).
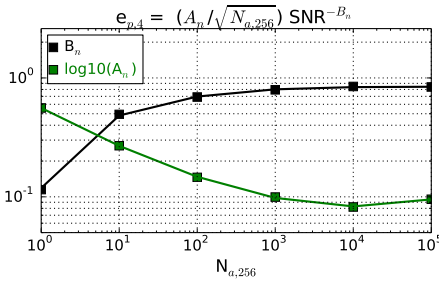
Figure 4: Precision model parameters $A_n$ and $B_n$ wrt the number of accumulated spectra ($M$=256).

## 5.  Results and future works

Figure 3 shows the precision for various SNR and number of averaged spectra. The results are shown for $M$=256. For $N_a > 10$, the number of bad retrievals is smaller than 50% for $\sqrt{Na}$ SNR$> 0.035$. In this range the precision is well approximated with the function shown in Fig 4:

Error model parameters: $\sum_{i=0}^{3} K_i \left(\log_{10} N_{a,256}\right)^i$

| | $M$=**64**, $Mp$=**192** | | | |
|---|---|---|---|---|
| $\log_{10} A_n$ | -2.143e-01 | -2.037e-01 | -2.137e-02 | 9.050e-03 |
| $B_n$ | 1.539e-01 | 3.976e-01 | -7.279e-02 | 3.755e-03 |
| | $M$=**128**, $Mp$=**128** | | | |
| $\log_{10} A_n$ | -2.563e-01 | -3.947e-01 | 3.935e-02 | -4.358e-04 |
| $B_n$ | 1.336e-01 | 4.790e-01 | -1.040e-01 | 8.010e-03 |
| | $M$=**256** | | | |
| $\log_{10} A_n$ | -2.526e-01 | -3.400e-01 | 1.727e-02 | 3.992e-03 |
| $B_n$ | 1.185e-01 | 4.529e-01 | -9.517e-02 | 6.712e-03 |

The trade-off between the single range SNR and the number of averaged spectra for a Doppler frequency precision of 0.5 ms$^{-1}$ is shown in Fig. 5. The number of averaged spectra needed to reach the horizontal (along track) and vertical resolutions of 100x0.5 km are indicated. The SNR lower bound for good retrievals (precision better than 0.5 ms$^{-1}$) is $10^{-3}$ and $2\times10^{-4}$ with a PRF=30 and 2500 Hz, respectively. ISOSIM first simulations indicate that this range corresponds to the SNR variability in the boundary layer but it is only a part of that in the free troposphere. Further analyses have to be performed to estimate the expected percentage of good retrievals within the troposphere and all latitudes.

These results are for the most favorable conditions where we assume that the signal spectrum width is that of the transmitted pulse and the atmosphere is homogeneous. The impacts of the variability of the SNR and of the spectrum widths due to atmospheric inhomogeneities will be assessed. For that, both the simplified model (for small horizontal scales) and ISOSIM (for horizontal scales larger that 100 km and for vertical variabilities) will be used. The simplified model will also be used to investigate other retrieval approaches. In particular, we will first test a more efficient computational approach for the maximum likelihood function calculation where the signal histograms (auto-correlation function) are averaged instead of the spectra. Methods for smoothing and averaging the signal before retrievals as well as for bad retrievals filtering will also be investigated.
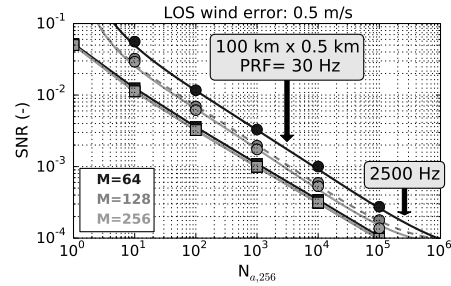


Figure 5: Single range SNR and spatial resolution trade-off for line-of-sight wind retrieval error of 0.5 ms$^{-1}$ (circles). The squares indicate the limit where the number of bad retrievals is 50%.

# References

[1] P. Baron, D. P. Murtagh, J. Urban, H. Sagawa, S. Ochiai, Y. Kasai, K. Kikuchi, F. Khosrawi, H. Körnich, S. Mizobuchi, K. Sagi, and M. Yasui. Observation of horizontal winds in the middle-atmosphere between 30S and 55N during the northern winter 2009–2010. *Atmospheric Chemistry and Physics*, 13(13):6049–6064, 2013.

[2] W. E. Baker, R. Atlas, C. Cardinali, A. Clement, G. D. Emmitt, B. M. Gentry, R. M. Hardesty, E. Klln, M. J. Kavaya, R. Langland, Z. Ma, M. Masutani, W. McCarty, R. B. Pierce, Z. Pu, L. P. Riishojgaard, J. Ryan, S. Tucker, M. Weissmann, and J. G. Yoe. Lidar-measured wind profiles: The missing link in the global observing system. *Bull. Amer. Meteor. Soc.*, 95(4):543–564, September 2013.

[3] A. Stoffelen, J. Pailleux, E Källén, J.M. Vaughan, L. Isaksen, P. Flamant, W. Wergen, E. Andersson, H. Schyberg, A Culoma, R Meynart, M Endemann, and P Ingmann. The atmospheric dynamics mission for global wind field measurement. *Bull. Amer. Meteor. Soc.*, 86:73–87, 2005.

[4] S. Ishii, T. Iwasaki, M. Sato, R. Oki, K. Okamoto, T. Ishibashi, P. Baron, and T. Nishizawa. Future Doppler lidar wind measurement from space in japan. In Michio Kawamiya, Tiruvalam N. Krishnamurti, and Shamil Maksyutov, editors, *Remote Sensing and Modeling of the Atmosphere, Oceans, and Interactions IV*, volume 8529 of *Proceedings Vol. 8529*, 2012. ISBN: 9780819492685.

[5] D. Wu, J. Tang, Z. Liu, and Y. Hu. Simulation of coherent doppler wind lidar measurement from space based on calipso lidar global aerosol observations. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 122(0):79–86, June 2013.

[6] K. Okamoto, S. Ishii, P. Baron, M. Yasui, Y. Satoh, D. Sakaizawa, R. Oki, T Kubota, C Takahashi, K Gamo, T. Ishibashi, and T. Y. Tanaka. Simulation and impact study of future spaceborne Doppler wind lidar in Japan. In *94th American Meteorological Society Annual Meeting*, 2014.

[7] D. S. Zrnić. Estimation of spectral moments for weather echoes. *Geoscience Electronics, IEEE Transactions on*, 17(4):113–128, 1979.

[8] R. G. Frehlich and M. J. Yadlowsky. Performance of mean-frequency estimators for Doppler radar/lidar. *J. Atmos. Ocean. Technol*, 11:1217–1230, 1994.

[9] R. Frehlich. Velocity error for coherent doppler lidar with pulse accumulation. *J. Atmos. Oceanic Technol.*, 21(6):905–920, June 2004.