Referee #1:

General comments:

This manuscript presents the results of wind measurements by coherent Doppler lidar from a ship in the Yellow Sea. The authors give a description of the algorithm for processing lidar data, which makes it possible to compensate for the measurement error associated with the motion of the ship. The results of joint measurements of height wind profiles by lidar and radiosonde are analyzed. The paper may be of interest to the readers of AMT. However, when describing the experiment and the data processing procedure, excessive attention is paid to secondary issues, and important details are ignored. Sometimes the terminology used by the authors makes it difficult to understand what they mean and how they obtained results presented in the manuscript. Some results raise doubts about their correctness:

1) The authors assume that the bias of lidar estimate of the wind velocity is associated only with errors in determination of the ship speed and direction and with the pointing angle knowledge errors (Section 3.3). One can agree with this, if lidar estimates of the radial velocity are obtained at a sufficiently high signal-to-noise ratio (SNR, ratio of the signal spectrum peak to the standard deviation of noise component of the spectrum estimate), when the probability (or fraction) of a bad (unreliable) estimate of the radial velocity is practically zero. However, results shown in Fig.9 for heights above 2 km were obtained at SNR =2 dB when the probability of bad estimate b = 0.3. As shown in Fig.4, true wind speed V = 5 m/s at a height of 2 km. According to the theory (Frehlich, R.G. and Yadlowsky, M.J.: Performance of mean-frequency estimators for Doppler radar and lidar, Journal of Atmospheric and Oceanic Technology 11(5), 1217-1230, 1994), the bias of velocity estimate BIAS = <V^> - V, where <...> is ensemble averaging and V^ is the velocity estimate, is determined by the equation: BIAS = -b*V. Therefore at b = 0.3 and V = 5 m/s the bias equals -1.5 m/s. Nevertheless, in Fig. 9(Ñ ¸ A) we see that the bias is about zero at SNR = 2.

R: I did not explain it clearly. The random error of radial velocity e_V in an individual Doppler Lidar velocity estimate is dependent on the signal-to-noise ratio (SNR) of the measurement. It can be evaluated based on the frequency spectrum of the retrieved velocity, which is applied for vertical velocity random error estimation in this paper, and we can also use the velocity differences from even- and odd-numbered pulses to estimate the random error (Frehlich 2001). As for the random error of horizontal velocity, we need firstly obtain the random error of each radial velocity that is used for 4-DBS wind profile. Then the radial velocity error should be scaled into the horizontal velocity error based on the relationship between horizontal velocity components U, V and the radial velocity.

As for the bias in this paper, we deal with the derivation of systematic errors (bias) to the horizontal wind retrieval. The error sources from the knowledge of the ship velocity and the lidar pointing angle are systematic part, and it is assumed that the random error part of the ship velocity and the lidar pointing angle is zero, which is reasonable and robust for horizontal wind retrieval according to the specific parameters of lidar, GNSS and INS. The small bias at SNR=2 in Fig.9 actually represents the bias from the contribution of knowledge error of ship velocity and lidar pointing angle.

However, the bias of velocity estimates BIAS= <Vˆ> - V, as the referee mentioned, is different from the definition in this paper. The definition of bias in this paper is $bias_v = \hat{V} - V_{truth} - e_V$, where $\hat{V}$ is the measured velocity estimate, V_{truth} is the desired or true wind measurement and e_V is the random error. It can be seen that the BIAS referee mentioned is the sum of $bias_v$ and e_V. As for the BIAS = -b*V=-0.3*5=-1.5 m/s, the referee mentioned, actually I really wonder how the b=0.3 is determined. It is mentioned in Frehlich’s paper that the empirical model for the fraction of bad estimates b as a function of Φ for fixed Ω and M is: $b(\Phi) = [1 + (\frac{\Phi}{b_0})^{\alpha}]^{-\gamma}$. It is noted that the in this paper SNR is defined as the ratio of the peak value of FFT spectral signal in each range bin to the Root-Mean-Square (RMS) of background noise signal, which is different from Frehlich’s paper definition and need to be treated carefully when determining those parameters. It would be possible to compare bias of two systems if we know the details of Frehlich's empirical model.

2) Fig.9 (b) shows the random error of wind velocity. On the other hand, the second term on the right-hand side of Eq. (11) is defined as the random error with zero mean. It is unclear how the result shown in Fig.9 (b) was obtained. It is necessary to describe in more detail the procedure for obtaining this result. The results shown in Fig. 4 and Fig. 9 are obtained from the same lidar data (measurements from 15:52 to 16:02 on May 9, 2014)?

R: Various methods of estimating the magnitude of the random error of Doppler Lidar velocity measurements have been introduced (Frehlich 2001). The measurements of error from velocity spectrum are used in this paper. A 50 % window overlap factor, a Hamming window is used in order to reduce the leakage in the spectra. A zero-padding of the missing values were applied to each window for each spectrum calculation to improve the frequency resolution. The constant high-frequency region of velocity spectrum higher than 0.2 Hz, shown in Figure 1 below, represents uncorrelated random error contribution, which is departing from the Kolmogorov’s -5/3 law. The random error of vertical wind velocity is estimated as the standard deviation of the measured signal after high-pass filter.

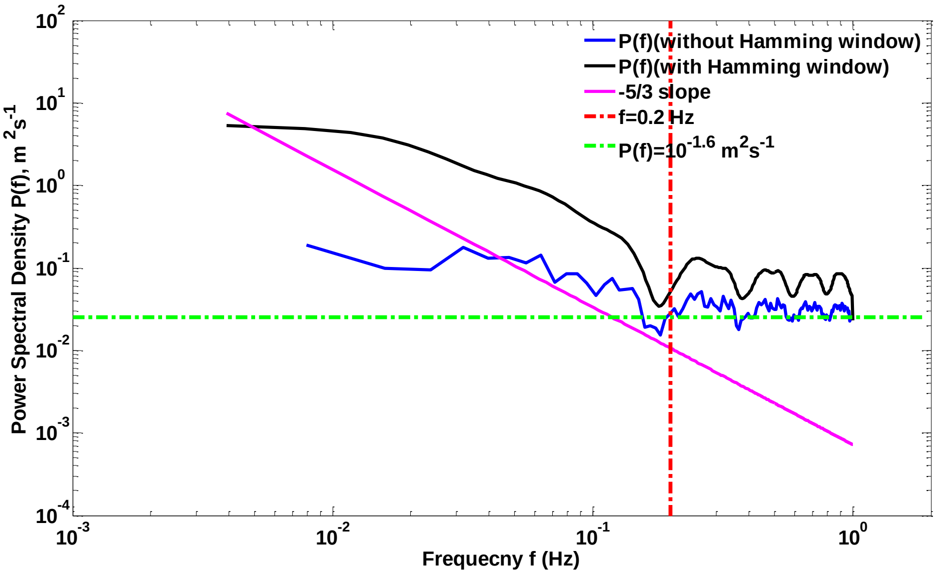

Figure 1: Power spectral density P(f) without and with Hamming window for the CDL measured vertical speed between 15:52 and 16:02 LST on 09 May and for an altitude of 1495 m (blue and black solid line, respectively). The expected spectral behavior according to the Kolmogorov's−5/3 law (pink solid line), the noise frequency threshold (red dotted line) and the derived noise floor for the CDL (green dotted line) are shown.

The results shown in Fig. 4 and Fig. 9 are obtained from the same lidar data (measurements from 15:52 to 16:02 on May 9, 2014).

3) How can SNR be determined below 2 dB, if in this case with a high degree of probability the peak in the measured spectrum is associated with the noise, but not with the signal?

R: The SNR in this study is defined as the ratio of the peak value of FFT spectral signal in each range bin to the Root-Mean-Square (RMS) of background noise signal. Figure 1 shows the array of the spectral $S(l\Delta f;k\Delta R)$, where $l=0,1,2,3,...,L-1$ is the spectral channel number and $L=100$. In this case the frequency resolution $\Delta f \approx 0.98$ MHz and the corresponding velocity resolution is $\Delta V = 0.76\ \mathrm{ms}^{-1}$. The bandwidth $B_{100} = (L-1)\Delta f = 97.68$ MHz, and the corresponding radial velocity measurement range is $\pm 37.5\ \mathrm{ms}^{-1}$. Figure 1a shows the last 10 range gates raw array of spectral in green line. We estimate the averaged background noise spectrum

$$\overline{S}_N(l\Delta f) = \frac{1}{10}\sum_{k=94}^{103} S(l\Delta f;k\Delta R) \tag{8}$$

Subtracting the background noise spectral $\overline{S}_N(l\Delta f)$ from the raw spectral array $S(l\Delta f;k\Delta R)$, the unnoisy array of spectral $S(l\Delta f;k\Delta R)$ can be obtained and shown in red line in Fig. 1. The peak value index l_{peak} from the $S(l\Delta f;k\Delta R)$ can be firstly obtained and thus the absolute signal power $P_s(k\Delta R)$ at various ranges $k\Delta R$ can be represented as:

$$P_s(k\Delta R) = S(l_{peak}\Delta f; k\Delta R) - \frac{1}{12}\left(\sum_{l_{peak}-20}^{l_{peak}-15} S(l\Delta f; k\Delta R) + \sum_{l_{peak}+15}^{l_{peak}+20} S(l\Delta f; k\Delta R)\right) \tag{9}$$

Replacing integration by summation and taking into account that the zero velocity point in one channel is $l_{zero} = 50$, we estimate the noise power P_N as

$$P_N = \frac{1}{10}\sum_{k=94}^{k=103}\sqrt{\frac{1}{21}\sum_{l=l_{zero}-10}^{l_{zero}+10} \hat{S}_N(l\Delta f; k\Delta R)^2} \tag{10}$$

Finally, we obtain the range profile of the $SNR(k\Delta R)$ using the equation

$$SNR(k\Delta R) = 10\log_{10}\left(\frac{P_s(k\Delta R)}{P_N}\right) \tag{11}$$

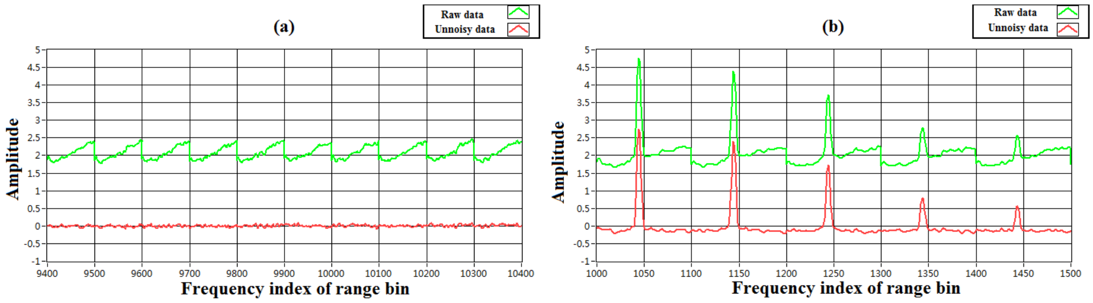

Figure 1: The CDL measured array of the FFT spectra (a) the last 10 range gates spectra for background noise spectrum estimation (b) the 1st – 5th range gates (150 m – 270 m, range resolution is 30 m) spectrum.

The SNR from Banakh et al. 2013 is defined as the ratio of the averaged heterodyne signal power P_s to the average detector noise power P_n in a 50-MHz bandwidth. The power P_s and P_n are integrals of the spectral densities $S_s(f)$ and $S_n(f)$, respectively, in frequency f within a band of width B_{50}, that is:

$$P_s = \int_{B_{50}} S_s(f)df \tag{5}$$

$$P_n = \int_{B_{50}} S_n(f)df \tag{6}$$

Comparing the definition from Banakh et al. 2013, the SNR in this paper is simpler and also indicates the CDL detection capability, data accuracy and atmospheric tracer particle relative intensity. In this sense, the SNR threshold value in this paper is higher than the one in previous studies (Banakh et al. 2013; Achtert et al 2015) for the same signal power spectrum.

Specific comments

1) Page 3, lines 27-30: The pulse energy depends on the pulse width? If so, what is the pulse energy (and pulse repetition rate) for pulse durations of 100, 200 and 400 ns?

R: I did not explain it clearly. The pulse energy is fixed and the pulse width is configurable. Pulse width is the full width at half maximum of the laser pulse waveform. A wider pulse width results in a larger measurement blind spot, but increases the average power and detection distance of the laser. Narrower pulse width can reduce the measurement of blind spots, but must also reduce the average power of the laser in order to control the laser peak power within the maximum range of the fiber, that is, will reduce the detection range.

The description in the revised manuscript is: “The achieved pulsed energy is approximately 150 μJ and the pulse repetition frequency is 10 kHz.”

2) In section 2 the following information should be added:
a) width of the time window (T) for obtaining the lidar signal power spectrum (T equals probing pulse duration of 200 ns?);
R: T=200 ns.
“The pulse width produced by the modulation, which is also the width of time window for obtaining the lidar signal power spectrum, is adjustable from 100 ns to 400 ns, thus the spatial resolution can be varied from 15 m to 60 m. We typically operate the CDL with a pulse width of 200 ns in this study.”

b) width of the frequency band (B) within which the radial velocity was estimated from the lidar signal power spectrum (B = 50 MHz?);
R: The bandwidth $B_{100} = (L-1)\Delta f = 97.68$ MHz is used for radial velocity estimation. The specific introduction can be seen in the answer to general comment question 3.
c) number of laser shots used for the spectral accumulation;
R: N=5000
d) number of radial velocity estimates (for each range) that were obtained from lidar measurement for 10 minutes and then they were used for obtaining one estimate the wind vector.
R: Both the determination of the ship-induced Doppler shift and the radial velocity have the same temporal resolution of 0.5 s. Figure 4 in the revised manuscript shows the flowchart of shipborne CDL data processing. Specifically, the LOS velocity and Signal to Noise Ratio (SNR) can be firstly determined using lidar data and FFT analysis. After the data pre-processing including the quality control based on SNR threshold, the attitude transformation is then used to obtain the azimuth and elevation in each LOS vector in Earth coordinate system with temporal resolution of 0.5 s. The LOS velocity detected by lidar is the atmosphere motion relative to ship coordinate system, thus the removal of the along-beam platform velocity due to ship motion is needed. In this study, the horizontal wind profile with 2-min temporal resolution will be retrieved for vertical velocity correction. Basically, the LOS velocities from N, S, E, and W direction after SNR quality control during the chosen 2-min interval are collected firstly. Then the procedure of filtration of reliable estimates of each radial velocity based on SNR threshold is used to obtain “good” speed estimates. The selected radial velocities and corresponding ship condition information in each radial direction are averaged and the averaged ship condition will be used for the removal of platform velocity effect. Finally, the horizontal with 2-min temporal resolution can be retrieved using modified 4-DBS mode. The vertical wind measurement has a temporal resolution of 0.5 s, the horizontal wind whose retrieved time is closest to vertical wind measured time will be used for vertical velocity correction.
3). Add the tele-scope diameter and beam diameter (1/e**2) to Table 1

R: The parameters have added to the Table 1 in revised version.

Telescope diameter	3 inches
Beam effective diameter	60 mm
Focal length	290 mm

4) Page 8, lines 25-28: “It can been seen that the discrepancies in wind profile above 1 km between the radiosonde and lidar measurement are significant due to the multipath effect at the ship platform and decrease in collocation of the measurement.” Another reason for the discrepancy between the results of the measurement of the wind by the lidar and the radiosonde at heights above 1 km is quite possible: the bias of the corrected lidar estimate of the wind due to the low SNR. It would be nice to add high profiles of the SNR in Figures 4 and 5. By the way, using some known procedure of filtration of good (reliable) estimates of the radial velocity obtained from 10-min lidar (4-DBS) measurements, the authors could obtain an unbiased wind speed estimate even in the case when the SNR is about 0 dB (if the percentage of good estimates is not below 20%).

R: It can be seen that the discrepancies in wind profile above 1 km between the radiosonde and lidar measurement are significant. On the one hand, the random error of the corrected CDL estimation of the wind due to the low SNR shown in Fig. 6a contributes to this discrepancy. On the other hand, the drift of radiosonde is affected by atmospheric turbulence perturbations and the CDL detection volume is changing during cruising observation. The spatial separation between radiosonde and CDL which can be called multipath effect, can cause larger discrepancy with increasing altitude.

Thanks for your suggestions, the SNR profile has added to the Fig 5 and 6 in the revised version, as shown below:

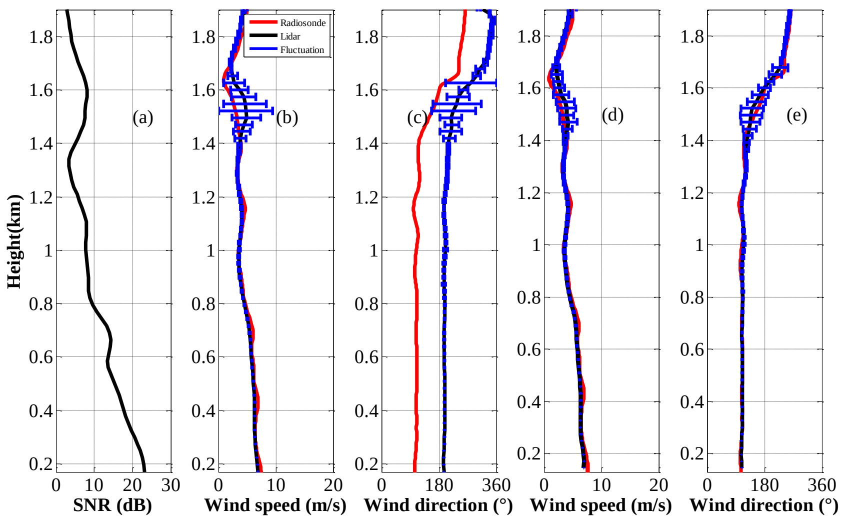

Figure 1: Anchored observation: (a) SNR profile (b) (c) wind speed and (d) (e) wind direction measured by CDL (blue line) before and after attitude correction, respectively. The simultaneous

radiosonde data is shown in red line. The blue bars represent the sampling fluctuations from 15:52 to 16:02 LST, 09 May, 2014.

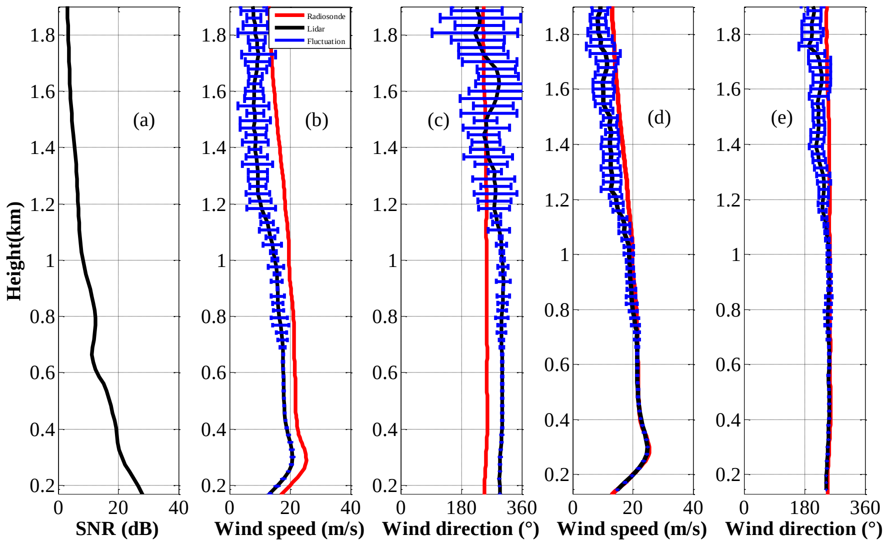

Figure 2: As Fig. 1, but for 07:44 to 07:54 LST 13 May, 2014 in cruising observation.