Reviewer 1:

General and Major Comments

1) The manuscript focusses only on the statistical comparison of the ship-borne wind-lidar to radiosondes and sonic anemometer. A short chapter on a measurement case for an interesting or typical situation needs to be added to demonstrate the new measurement capabilities of this lidar for atmospheric research. Figures of time-height cross-sections for horizontal wind speed and direction could be shown and discussed, together with complementary measurements (e.g. ceilometer obs., surface wind, radiosonde profiles). This would add much more confidence in the performance evaluation of this new ship-borne Doppler lidar than showing only statistical results.

We added an example figure in Section 3.1 showing the Ceilometer backscatter and Halo co-backscatter coefficient, aerosol depolarization determined from the Halo co- and cross-backscatter coefficient, vertical and horizontal wind as well as wind direction and precipitation measured on Oden:
A typical measurement day of 17 September 2014 is presented in Figure 7. Marine stratocumulus was present for the entire day with cloud base heights descending from 400 m to 200 m during the course of the day. The depolarisation ratio from the Doppler lidar reveals that ice (high values) is precipitating from liquid water clouds (low values) between 06 and 18UTC. This is also in agreement with the measurements of vertical wind speed (negative values equal downward motion of due to precipitation) and precipitation rate. Averages of 20 minutes of Doppler lidar measurements around the times of radiosonde launches at 06, 12, and 18 UTC are used for a profile-to-profile comparison of wind speed and direction in Figure 7e and f. No major differences are visible for the height range covered by the Doppler lidar apart from the wind speed in the lowermost 100 m for the 18-UTC sonde.

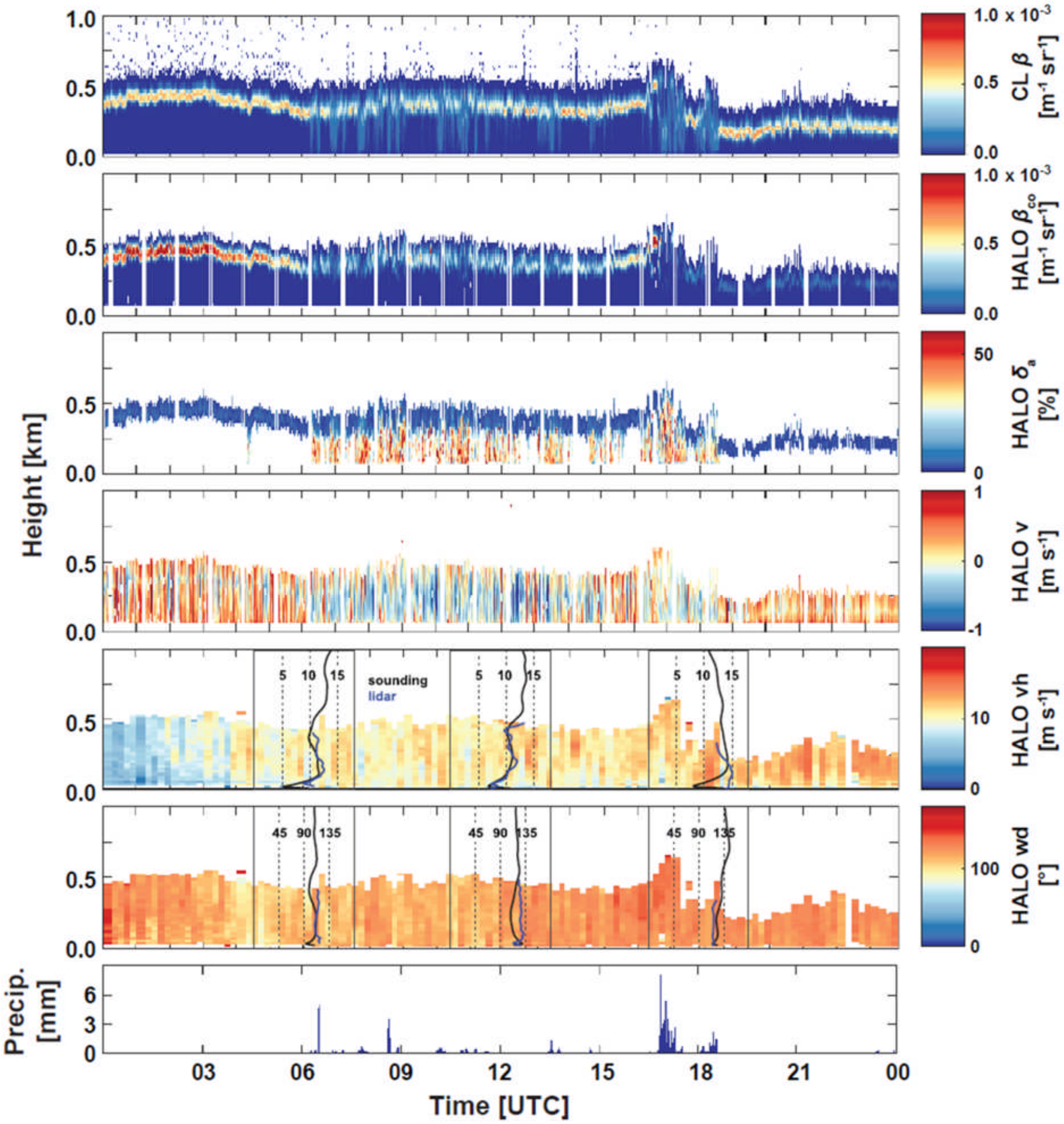


Figure 7: Example measurement of 17 September 2014: calibrated backscatter coefficients from ceilometer (a) and Doppler lidar (b), depolarisation ratio (c), vertical (d) and horizontal (e) wind speed, wind direction (f) from Doppler lidar, and precipitation amount from a precipitation sensor on the ship (g). The profiles in (e) and (f) show the comparison of horizontal wind speed and direction to the radiosondes launched at 06, 12, and 18 UTC. The ships average speed during that day was XX m s^{-1}.

2) The innovative part of the manuscript wrt measurement methodology is related to the ship-borne DWL. Only few details w/o any equations describing the correction method are provided in ch. 2.2. Thus more details (e.g. equations or at least citation of the equation numbers in the references, simplified sketch of correction scheme) should be provided here, e.g. about the derivation of the remaining Doppler shift from the pointing or the filtering approach. A figure could be shown with AHRS data and the low-frequency ship data and how they are combined.
Also more technical details about the accuracy of the AHRS sensor should be provided; also the weight of the lidar, which is stabilised is missing. In addition the data source for deriving the platform motion is unclear: Does the AHRS sensor include a GPS for that, or is the horizontal motion derived

from acceleration sensors? Where does the low-frequency horizontal velocity correction from the ship´s navigation data come from (GPS?) and what is the accuracy of this? Typical values for the residual correction, e.g. horizontal velocity, ship-induced LOS velocity should be provided, in order to assess the magnitude of the correction in comparison to the atmospheric wind speed.
I would also propose to write 1-2 paragraphs about the performance of the wind-lidar wrt. ship vibrations and other environmental conditions during this Arctic cruise? Any problems/degradations of the lidar performance (e.g. laser) due to the ship vibrations? Any issue with the harsh environment (sea spray, temperature)?

Details of accuracy of AHRS measurements have been added to the text. The ship navigation data is derived from a combination of a Javad GPS and Funuro GPS-gyro compass. The ship velocity is derived from the change in GPS position over time. GPS positional accuracy is of the order of 4m; here the velocity is low-pass filtered and estimated over a period of order 100s giving a velocity error of order cm s-1. The heading (RMS) accuracy is 0.5°.

The weight of the lidar was added to Table 1.

There were no significant problems with the lidar – the scanning head motor jammed on a handful of occasions, we think following sudden violent shocks during ice breaking. The laser performance was stable and within the specifications.

Details on the motion-stabilisation algorithm have been added as Appendix A.

2) The statistical comparison provided in the manuscript lacks some important quantities as the mean of the difference (bias, systematic error) and the standard deviation of the difference (random error) between the lidar and radiosonde/sonic anemometer. The linear fit coefficients do not cover all aspects of a performance evaluation; the intercept of the linear fit is only equal to the bias, if the slope is 1. Also statistical quantities for all altitudes (e.g. in Table 2 and 4) should be given. On the other hand the line “Pearson´s r” could be deleted, because it is already covered by the quantity “R_2”. The bias and standard deviation of the comparison should be discussed in chapter 3.1, and also included in the conclusion. I consider the discussion in ch. 3.1 as too much limited to the squared correlation coefficient R_2; it is only 1 parameter to show correspondence and is dependent on the range of values for the measured quantity. Thus high values are achieved for wind direction ranging from 0° to 360°; I consider the systematic and random error as more important. Thus more room should be given in Ch. 3.1 to discuss these quantities.

The numbers for the bias and standard deviation have been added to Tables 3 and 4. Both revised tables are given below. The text has been revised accordingly to:
“The lidar and radiosonde wind directions agree very well for the 5-point geometrical wind solution (R2=0.96, not shown) and the 4-point sinusoidal fit (R2=0.99, see Fig. 4a). Both methods show negligible bias while a slightly better standard deviation is found for the 4-point sinusoidal fit (Table 2).”
And
“Both bias and standard deviation for wind direction increase with altitude (Table 3). This is primarily the result of the drift of the radiosonde and the resulting decrease in collocation of the measurements.”
And
“The squared correlation coefficient for the solution for wind speed improves from 0.86 at 100 to 0.94 at 400 m. In addition, the absolute value of the bias decreases from 0.5 to 0.3 ms −1 (Table 3). The change from positive to negative bias with height marks systematically larger wind speed from the radiosoundings compared to the lidar measurement.”

Table 2: Statistics of the comparison between lidar and radiosounding for a height of 75 m.

		Lidar vs. Radiosonde	
		75m	
		Wind speed	Wind direction
Geometrical wind solution	N	175	163
	Std	**1.3 m s^{-1}**	**20 º**
	Bias	**0.5 m s^{-1}**	**-1 º**
	R^{2}	0.85	0.96
	Intercept	0.71	4.35
	Slope	0.85	0.98
	Normelised RMSE	**1.7%**	**4.2 %**

Sinusoidal fit	N	229	220
	Std	**1.2 m s^{-1}**	**10 º**
	Bias	**0.4 m s^{-1}**	**0 º**
	R^{2}	0.86	0.99
	Intercept	0.48	3.53
	Slope	0.89	0.98
	Normelised RMSE	5.9%	3.5%

Table 3: Statistics of the comparison between lidar and radiosounding at heights of 75, 100, 400, 600, and 700 m.

	Wind speed						Wind direction					
Height (m)	75	100	200	400	600	700	75	100	200	400	600	700
N	229	226	182	90	47	37	220	219	181	90	43	35
Std (m s^{-1}) / (º)	**1.2**	**1.3**	**1.0**	**1.0**	**0.9**	**0.9**	**10**	**10**	**10**	**12**	**13**	**16**
Bias (m s^{-1}) / (º)	**0.5**	**-0.3**	**-0.4**	**-0.3**	**0.0**	**-0.3**	**0**	**0**	**1**	**-1**	**2**	**5**
R^{2}	0.86	0.88	0.94	0.94	0.96	0.96	0.99	0.99	0.99	0.98	0.97	0.97
Intercept	0.48	0.64	0.45	0.14	-0.48	0.15	3.53	5.12	-0.46	0.43	-5.78	-12.95
Slope	0.89	0.96	1.00	1.01	1.06	1.03	0.98	0.94	0.99	1.00	1.01	1.03
Normelised RMSE (%)	**5.9**	**3.7**	**3.2**	**3.4**	**4.4**	**6.8**	**3.5**	**3.2**	**1.1**	**1.4**	**4.8**	**6.4**

3) Abstract: A measurement random error of 0.2 m/s is stated in the abstract. This low value was not derived from the comparison with the radiosonde/sonic anemometer for which most of the statistical comparison values and Figures are included in the manuscript. The value of 0.2 m/s is derived from the lidar measurement itself derived using the autocorrelation technique. I would propose to add (instead of the 0.2 m/s or in addition) at least the main parameters from the statistical comparison from lidar to the radiosonde in the abstract: I found a systematic difference of 0.3 m/s (p. 9), the correlation coefficient and the standard deviation of the comparison for all altitudes needs to be derived (see my comment 3)

We added the following sentences to the abstract: The comparison of lidar measured wind and radiosoundings gives a mean bias of 0.3 ms −1 (2°) and a mean standard deviation of 1.1 ms −1 (12°) for wind speed (wind direction). The agreement for wind direction degrades with height.

4) I was confused about Section 3.2. It is not clearly stated in the first sentence, which measurements were used for the auto-correlation approach (although it becomes clear, when reading the full section). Was it the horizontal wind speed or the vertical velocity as usually taken in these approaches (e.g. Pearson et al 2009, O`Connor et al. 2010) This needs to be clearly stated. If it is the vertical velocity, then the derived value of 0.2 m/s is representative for the lidar random error in radial direction and not for the horizontal wind speed. This needs to be corrected in the formulation of the abstract, the discussion of Section 3.2 and conclusion. This derived value cannot be taken as the measurement error for the horizontal wind speed, because here several LOS measurements are combined (reducing the error) but horizontal homogeneity has to be assumed (increasing the error). Also it needs to be stated, if the vertical velocity is taken from the vertical LOS measurement (5th beam direction) or if it is an output parameter after the sinusoidal fit (or geometrical solution).

Section 3.2 deals with measurements of vertical wind speed. We added the suggested clarification to Section 3.2, the abstract, and the conclusions.

Specific Comments

p. 9340: There is a more recent WMO 2014 reference for the SoG for global NWP; also a reference for ADM-Aeolus (see below) could be added for documenting the need for additional wind observations.

We updated the WMO reference (http://www.wmo.int/pages/prog/www/OSY/SOG/SoG-Global-NWP.pdf) and added the ADM-Aeolus science report (2008) as reference.

p.9341: I would propose the replace the reference for ship-borne lidars (Baker et al. 2014) by a more appropriate (e.g. Tucker et al. 2009), because Baker et al. 2014 is focussed on spaceborne wind lidar.

We included Tucker et al. (2009), but left the Baker et al. (2014) reference to emphasize the need for measurements over the ocean.

p. 9341: It is stated that: “Only a few studies have used Doppler lidar on ships, and fewer still have actively stabilised the system against ship motions.” I am wondering, if there is a reference of a Doppler lidar on a ship w/o active stabilisation, which is suggested by the sentence. If not, I would propose to re-rewrite the sentence.

We changed the sentence to: This is why only a few studies have used Doppler lidar on ships (Hill et al., 2007; Wolfe et al., 2007; Pichugina et al., 2012). Improvements in technology along with decreasing costs make this an increasingly attractive approach.

p.9341: The following reference should be added to the introduction and also referenced in ch. 2.2. Hill, R. J., W. A. Brewer, and S. Tucker, 2007: Platform-motion correction of velocity measured by Doppler lidar. J. Atmos. Oceanic Technol., 25, 1369–1382.

We included the reference in the introduction.

p.9341: The provided values for the random errors (e.g. 0.3 m/s) are given in Pichugina et al. (2012), which should be added here.

We included the reference at that point.

p.9342: It would be useful to introduce the objectives of ACSE in chapter 2 with 1-2 sentences.

We added the following statement to summarize the ACSE objectives:
The primary objective of ACSE was to study Arctic clouds and their relation to tropospheric vertical structure, meridional transport and the surface energy balance for a variety of surface conditions, from open ocean, through marginal ice and into dense pack ice.

p. 9345: Ch. 2.3.: The calculation of the SNR is ambiguous, because of the used bandwidth of the signal. The bandwidth for the calculation of SNR or some more details should be given. The different SNR threshold values are only comparable, if the same bandwidths and the same calculation approach is used.

We used the same approach as in the literature on measurements with the same instruments (HALO Photonics Stream Line scanning micro-pulsed Doppler lidar. SNR is a parameter that is provided by the instruments. This means that we performed calculations identical to previous studies.
We changed the sentence to clarify that we used a commercial instrument: Previous studies using a HAlO Doppler lidar suggest a SNR threshold ranging from …

p.9346, 2nd paragraph: Is there a reference to the 5-point geometrical wind solution? It is also necessary to mention here (or in section 3.1), if the vertical wind is set to zero (w=0) for the sinusoidal fit, or if the vertical wind is retrieved as an independent parameter.

Both methods are described in Werner et al., (2005). We changed the sentence to: Wind speed and direction were obtained from the motion corrected HALO Doppler lidar measurements using the five-point geometrical wind solution and the 4-point sinusoidal fit method both described in Werner (2005)…
Further we added a sentence to describe the retrieval of the vertical wind: The vertical wind is retrieved as an independent parameter for the sinusoidal fit (Werner et al., 2005).

p.9346, 2nd paragraph: More information should be given, which QA is applied, e.g. by given the equations or by citing the corresponding equations in Päschke et al. (2015). Currently it is too vague, which QA is applied.

We revised the text to point to the respective equations given Päschke et al. (2015): For the 4-point sinusoidal fit method, we applied the quality assurance criteria described in Päschke et al. (2015), i.e. we tests for horizontal homogeneity of the wind field (Eq. 13 in Päschke et al. (2015), we applied a threshold of $R^2 > 0.95$*) and the collinearity of the Doppler velocity measurements used within one scan (Eqs. 14,15 in Päschke et al. (2015).*

p.9347: The radiosonde uncertainty is provided: It should be clearly stated, if this is a random or systematic error or a combination of both.

According the RS92 datasheet it is the total uncertainty. We added the information to the manuscript.

p.9347/9348 paragraph about the lidar data coverage and the boundary layer height determination: It would be very useful to provide some data on the observed boundary layer height and the corresponding lidar coverage. The typical boundary layer height during the cruise could be provided within a Figure or a Table (e.g. daytime maximum, night-time) from the Ceilometer. This would be important to assess coherent wind-lidar performance in Arctic regions. It is only stated on p.6, that the lidar coverage is comparable in the boundary layer to previous studies, but no values for typical boundary layer height is provided.

We added the average daytime maximum and minimum boundary layer height in Sec. 2.3: The average daytime maximum and minimum atmospheric boundary layer depths during the cruise were 576.3 and 81.4 m, respectively.

p.9348, ch. 3.1, last paragraph and Table 2: What is the reason to show the data for different BLD in Table 2? The results for different BLD are quite similar. Also the results are not discussed (only mentioned in the text). Thus I see no need to show Table 2; one could mention that this was investigated but no significant differences were found.

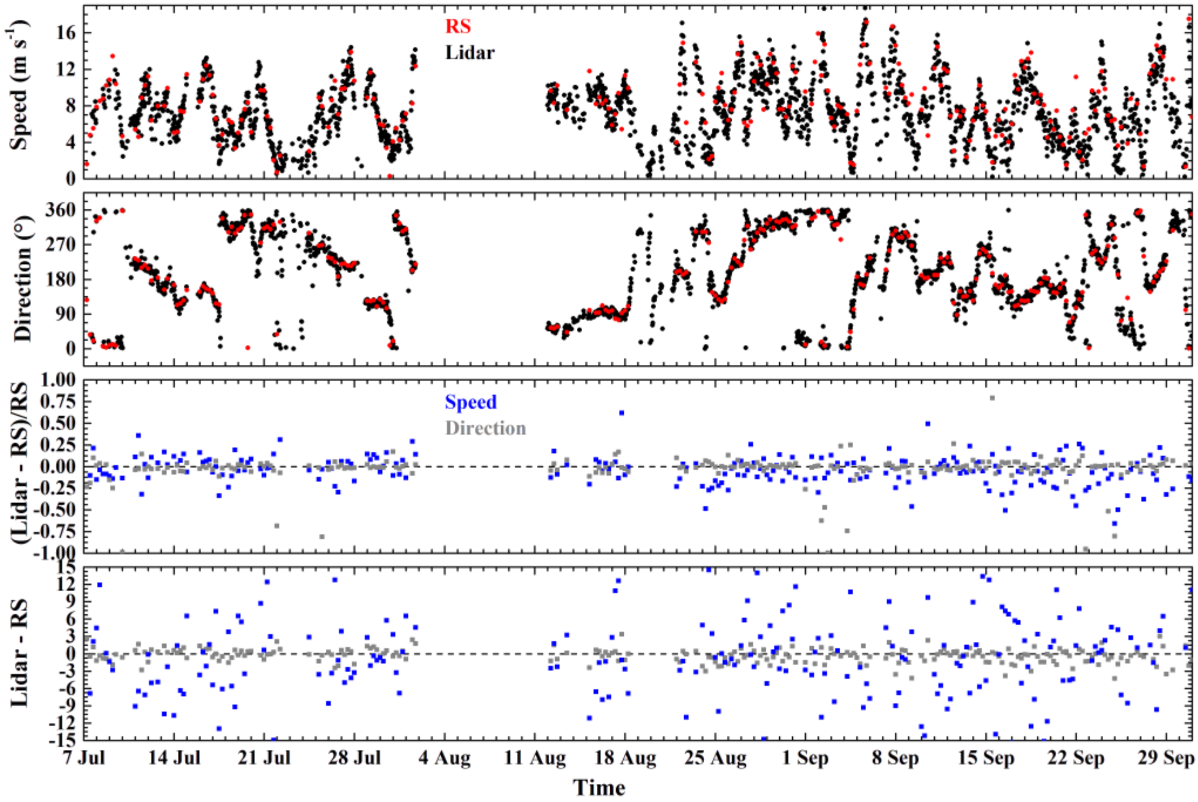
We omitted the table and the discussion of the comparison to flux tower measurements in the revised manuscript.

p.9350: The effect of the ship on the radiosonde lowest level winds is discussed here. Is this really relevant for the comparison of the 75 m wind? I would expect that the arguments here are relevant for the lowest 10-20 m after launch of the sonde. Is there some evidence in the radiosonde data that it is influenced by the ship until 75 m?

As discussed later in Section 3.3 the effect of the flow distortion over Oden are significant for the lidar wind measurements below 75m. The radiosoundings should be similarly influenced by the flow distortion. However, the influence depends strongly on how quick and far the radiosonde drifted away from the Oden after launch.

p. 9351 and Fig. 6: A panel showing the absolute difference (Lidar-RS) in [m/s] and [°] should be provided to provide an indication of the time dependency (or not) of the bias. The relative difference is only relevant for the slope error.

A panel that shows the absolute difference has been added to Figure 6. Please find the revised figure below. We also added a respective statement to the text:
A time series of the absolute differences in wind speed and direction reveal slightly increased scatter in the second part of the cruise during which higher wind speed has been observed.
New caption: Time series of lidar (black, every 50 min) and radiosonde (red, every 6 h) wind speed (top) and wind direction (middle) at 100 ma.s.l. The lower two panels show the relative and absolute difference for wind speed (blue) and direction (gray).



p. 9352: equation number from O`Connor, which was used should be given.

We added the equation number from O'Connor et al. (2010) to the manuscript.

p. 9352: I do not think that the larger std. dev. (compared to Pearson et al. 2009?) are due to lower aerosol load. The lower aerosol load leads to lower SNR; thus only more values with lower SNR should be present. The same SNR should result in the same standard deviation. I see the main difference in the additional contribution of the ship-motion. It seems that the ship-motion adds a constant random error of about 0.1 m/s to the lidar random error.

We agree with the reviewer. The statement about the effect of aerosol load on the measurement error has been removed.

p.9353: I would use the word "projection" instead of "aliasing" for the horizontal wind error influence on the vertical. "Aliasing" is more commonly used for components of a frequency spectrum folded into another frequency part.

We changed the word aliasing to projection.

p.9353: I would clearly distinct here error sources, which add a systematic error to the vertical velocity to explain the derived mean value of 0.1365 m/s and components which add a random error (as visible in Fig. 7 of about 0.1 m/s). Ship motions and projections of the horizontal wind during 2 s would add a random component to the vertical velocity, because it could be positive or negative (and not as stated a bias). Constant biases could be due to flow distortions of the ship (as mentioned and shown).

We added the following statements to the beginning and end of the paragraph, respectively:
In addition to the random error we also investigated the measurements of vertical wind speed for systematic errors.
This constant bias could be due to flow distortions of the ship that are discussed below.

p. 9354, ch. 3.3: It is mentioned that the horizontal wind speed is biased by 2% from the ship. It should be added, that this is a positive bias, meaning that the horizontal wind speed above the ship is higher than in the background flow. I am wondering, how this influences the horizontal wind speed retrieval from the lidar, which depends certainly on the direction of the LOS winds wrt the axis of the ship. I am also wondering, how this influences the comparison with the radiosonde and sonic. If both the lidar and the radiosonde winds are influenced by the ship, than this is not visible in the comparison, which certainly depends on the horizontal distance of the sonde wrt the ship.

That is valid point. The wind retrieval of the lidar does not take into account the orientation of the axis of the ship with respect to wind direction. We lack a truly independent reference measurement as all wind observations are affected by the flow distortion over the Oden. We can only raise the reader's attention of the error but we cannot correct for it.
We added the horizontal wind speed bias is a positive bias.

p.9355: It must be added here that the measurement error of the lidar wind speed is relevant for the radial velocities (and not for hor. Wind speed) and was derived from vertical pointing measurements.

We added that the measurement error is only relevant for the vertical wind speed.

p. 9356: A reference for the ADM-Aeolus mission should be provided, e.g. ESA 2008, available from http://www.esa.int/Our_Activities/Observing_the_Earth/The_Living_Planet_Programme/Earth_Explorers/ADM-Aeolus

Reference was added to the manuscript.

Fig 7: It would be nice to add the median of the measured occurrence in the plot as a line; also it needs to be added in the caption, that this was derived from vertical velocity measurements.

We added the median as red line to Figure 7 and revised the caption accordingly: Standard deviations of the Doppler horizontal velocity determined from the zero leg impulse in the auto-covariance and the theoretical standard deviation (black line) for a signal spectral width of 2 ms $^{-1}$. The median standard deviation per SNR bin is shown as red line.

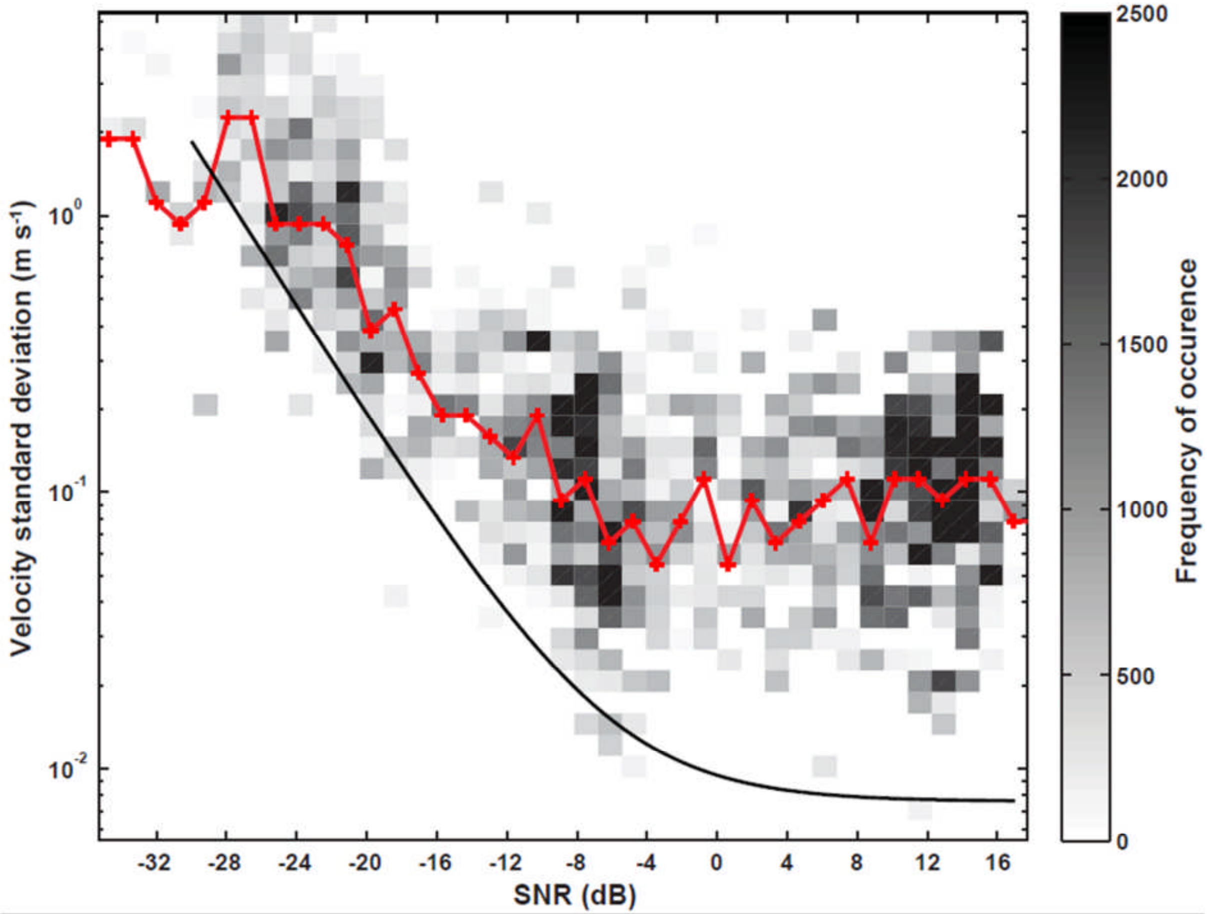


Fig. 8, caption. The horizontal wind speed for the CFD simulation should be mentioned in the Fig. caption.

We added the information on wind speed used for the CFD simulation to the figure caption.

Editorial

p.9348, ch. 3.1, L22: The 75-m results are reported in Table 3 and not as stated in Table 2.
p.9349, ch. 3.1, L11: The BLD results are reported in Table 2 and not as stated in Table 3; thus it would be good to change the orders of the Tables
p.9349, L7: no parentheses for the reference within the sentence.
p.9349, L26: must be 400 m ASL (instead of 40 m ASL).
p.9356, L361: less than “5%” instead of “0.5%”

We corrected all proposed editorial changes.