The reviewer's comments are given in green, while our response is provided in black.

Reviewer 1

General comments: The paper presents observations from the eXperimental Planetary boundary layer Instrumentation Assessment (XPIA). The observations were used to verify Doppler lidar turbulence profiles through comparison with sonic anemometer measurements. During a 17-days period, a single scanning Doppler lidar continuously cycled through different turbulence measurement strategies: velocity azimuth display, six-beam, and range height indicators with a vertical stare. The investigation focused on turbulence kinetic energy, turbulence intensity, and shear velocity data. For evaluation, sonic anemometer measurements at six heights on a 300-m tower were available. The paper is well written and of general interest for the lidar community (scientists and users alike). I therefore recommend the paper to be published with minor revisions.

Specific comments
Section 1: the introduction is quite general and could be more concise on the topic tackled (more focused) in this investigation.
We have rewritten the paragraphs in the Intro that review the previous work and alternative approaches to turbulence calculation, to make it more evident how they are related to the current work.

Page 5, line 2: Specify which of the methods mentioned in Lenschow et al. (2000) is used to remove noise.
We have clarified here that we are fitting a structure function of the form Eq. 32 in Lenschow et al. (2000) to the autocovariance of the residuals.

Page 5, line 17: could you make a statement concerning the time window for detrending (15, 20, 30, 60 min?) on the turbulence results. Applying a 20-min window could filter out large convective cells.
We have added a statement here clarifying that the 20-min window could filter out large convective eddies when the wind speed is small, and that the effect would be exacerbated for a shorter time window or higher-order (i.e., non-linear) filtering.
For this study, the 20-min window is naturally chosen since the 6-beam technique was used for 20-min at a time at the end of every hour, as later discussed in Sect. 3.

Page 9, line 14: '... 50-m pulse width...': does that mean that physically independent measurements are (physical resolution of the lidar) 50 m?
For practical applications, yes, physically independent measurements are spaced by 50 m (the pulse width), as long as the pulse does not intercept a hard target. This has been clarified in the paper here.

Page 12, line 5: y = bx: Transformation of equation (9) gives y = x 10b
We regret this error in the previous version, and have modified the equation accordingly here.

Figure 5 and 9: a zero line would be helpful.
Since the plots are on a logarithmic scale, values at and below zero are not shown on these plots as they are off the scale. Thus, no changes are made.

Page 13, line 22 ("..... may be due to the inability to capture all the scales of turbulence"): Spectra should be included in order to see which scales are not captured (to prove the statement would be good)
While we agree that showing spectra would be helpful to understand the limitations of the measurements, however it is not possible to calculate spectra from the VAD technique since the measurements from a PPI scan are over a large spatial area where the pointing angle is constantly varying. Since spectra cannot be used to see which scales are captured, the integral time scale is calculated from the sonic anemometer data to determine the characteristic

maximum eddy size (see Fig. 7 and discussion), which is used to show that the PPI scan area near the lidar is often smaller than the largest eddies to support the statement made here.

Page 15,line 18: (negative u-variance values). Is this the same effect as for TKE mentioned on page 11,line 5?
Yes, this is the same effect as mentioned for TKE earlier. We have added a statement to the manuscript to explicitly affirm this.

Page 16, line 7: "the bias becomes small as most of the turbulence scales are resolved". Once again, please prove that by providing a spectrum!
Again, while we agree that spectra would be the clearest way to support this statement, we are unable to produce power spectra using time-series analysis from the measurements using the VAD technique (see response above). That is why the analysis comparing the VAD scanning circle area to the integral scale of turbulence (typical largest eddy size) is given at the end of Sect. 4.1 to support the statement made here instead of using spectrum.

Section 4.3, Figure 9: As the comparison of u-star shows a huge scatter, sample time/spatial series and spectra from periods when the sonic and lidar data agree and disagree, respectively could provide more insight into the differences. Have you looked at the data in more detail? Any additional information would be good?
We assume that the reviewer is referring to Fig. 12, not 9, as 12 shows u-star measurements and shows the large scatter pointed out by the reviewer. Again, due to the nature of how the 6-beam and VAD methods operate (see descriptions in Sect. 2.1, 2.2), it is not possible to calculate spectra or show a time series of the u, v, or w themselves. For these reasons, it is difficult to pinpoint the reasons for why the u-star measurements are often so poor using either the 6-beam or VAD technique from the Doppler lidar.

However, since u-star is calculated based on the measurements of u'w' and v'w', the accuracy of u-star itself is reliant on the validity of u'w' and v'w' measurements. As discussed in Sect. 5, u'w' and v'w' are difficult to measure since their magnitudes are typically small (<0.1 m^2 s^-2) and are not measured accurately by either the 6-beam or VAD techniques (see Appendix A for quantitative analysis). Since neither of these covariance terms can be accurately measured with the 6-beam and VAD techniques, it is unsurprising that u-star, the combination of these terms, measured by the lidar shows poor correlation with u-star calculated from sonic measurements. We have added the following statement in the first paragraph of Sect. 5 to point out this relationship: 'Since the individual covariance terms do not correlate well with sonic anemometer measurements, it is unsurprising that the u* values computed from either the six-beam or VAD techniques also show little correlation with u* from the sonic anemometer (r^2=0.14--0.17).'

Page 15, line 8: 'the largest scales of turbulence are observed if the time window length exceeds the integral time scale ... ': Although this is correct a discussion about the error should be added, i.e. what is the error due to poor statistics if the time window is 5 or 10 min only.
We have added a paragraph providing a discussion of the sampling errors (i.e., statistical representativeness, related to the time window) at the beginning of Sect. 4 before any of the results are shown. By placing the discussion here before the intercomparison of turbulence measurements, the material and possibility of these errors is introduced to the reader before results are interpreted. In this discussion, we provide several sentences about the error due to poor statistics if the time window is short compared to the integral time scale.

Page 19, line 13: examples that turbulence can significantly vary spatially is shown in Maurer et al. (2016) doi:10.5194/acp-16-1377-2016.
A reference to Maurer et al. (2016) has been added here.

Page 20, line 1: examples for spatial variations in the mean wind due to local flow (valley wind) is also demonstrated in Adler et al. (2014) doi: 10.1007/s10546-014-9957-8.
A reference to Adler and Kalthoff (2014) has been added here.

Page 20, line 9: (spatial resolution):what is the physical resolution? See comment above.
See response to comment above. The physical resolution of the lidar, generally speaking, is related to the pulse width (here 50 m).

Page 21, line 28: Please rewrite the sentence “Although the sonic anemometer observations agreed most poorly with RHI-measured TKE and TI” into “Although the RHI-measured TKE and TI agreed most poorly with sonic anemometer observations” because sonic observations are considered to be the “truth”.
This sentence has been rewritten as indicated.

Typing errors:
Figure caption Figure 2: ‘..... shown (c)’: delete âAŽshown’
Corrected
Figure 7: ‘u integral scale’ should be ‘l integral scale’
Changed simply to ‘Integral scale’ to prevent confusion.
Page 15, line 9: ‘10 - –100 s’ should be ‘10 –100 s’
Corrected
Page 20, line 4: "e.g., Mann et al...“ instead of "...i.e.“
Corrected