

We thank the reviewers for their constructive comments, both general and specific. By addressing these comments, we believe the quality of the manuscript has improved. In the attached document, we have addressed the concerns raised by the reviewers, and have made the indicated changes to the manuscript where possible and appropriate. The reviewers' 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 ( $<0.1 \text{ m}^2 \text{ s}^{-2}$ ) 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 ‘shown’

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

# Reviewer 2

The authors compare different methods for calculation of turbulence parameters from measurements with a single Doppler lidar. The topic is up-to-date and is very important for the further development of different scientific disciplines and the further technical development, e.g., of wind power plants. The authors have put considerable work into the paper and AMT is the right journal to publish this study. I recommend publication if the following major issues are addressed:

Major issues:

1.) At many points in the document, quite subjective descriptions of a correlation or a match are given. Please look into this issue. You could, e.g., quantify what you mean with "good", "bad", "show skill" or "accurate" once within the document and connect it with proper numbers. It will really upvalue the paper, if you make it more quantitative. It will help to transport the message.

In several places throughout the manuscript, we have added quantitative descriptions (i.e., correlation coefficients, slopes to indicate bias) to the text where it seems appropriate, including in the abstract. However, to ensure ease of the narrative when describing these data sources in the discussion/results sections, we found it preferable to use more qualitative language to indicate whether we thought these values represented a good or bad fit with respect to the other techniques and measurements, and to avoid redundancy of the quantitative analysis already provided in the tables and figures themselves. Instead, the appropriate figures or tables are referenced in the text. In addition, Table 2 summarizes all the quantitative statistics for a quick look up by the reader. To further strengthen the quantitative analysis, we have also provided the RMSE values in the updated Table 2.

2.) p12 l10: -> This discussion must be mentioned earlier in the document. Is there really no way to estimate the uncertainties of these methods? I would doubt that. Understandably, analytic error calculation is extremely difficult for this kind of evaluations. However, numerical methods exist that can yield an error estimation for certain kinds of noise. You could, e.g. use Monte Carlo simulations, imposing small variations on the input data and then analyze in what range the results change. With such an approach it would be possible to discriminate between measurement errors and methodical uncertainties (e.g., incomplete overlap between the tower measurements and the lidar observations). That would greatly help the interpretation especially of Figs. 4/5 and 8/9. This topic is also connected with P.13 Line 8: "Approximately half of these outliers are negative TKE values, which were removed as discussed earlier..." -> Are those really outliers or just noisy values that happen to be close to zero. An uncertainty estimation or a more thorough description of the 6-beam technique would help here.

We agree with the reviewer that quantifying the uncertainty of the measurements is ideal for these intercomparisons and interpreting the results. Unfortunately, determining the magnitude of the uncertainty itself is not trivial. Entire studies have been devoted to addressing and quantifying sampling errors of turbulence and flux measurements from time-series analysis alone (see references in new discussion in paper). Since the measurement techniques presented here have are made using Doppler lidar observations over both time and space, no method has been developed yet to quantify errors with these techniques. Due to the intensive analysis required to properly attribute sampling errors from even one technique (as evidenced by multiple studies now referenced in the paper that evaluate sampling errors from time series analysis), entire studies could be done to determine the proper sampling errors for each of the techniques here separately. As such, properly quantifying and attributing sampling errors are out of the scope of this paper and is a topic for further studies. In lieu of quantifying these errors, we have added a more thorough discussion of sampling errors earlier in the paper before the results are presented (last paragraph before Sect. 4.1).

We do note that the input data (line-of-sight velocity measurements) are already contaminated with random error (i.e., noise). This is covered in Sect. 2 (see Eq. 1 & 2). However, using established techniques using the autocovariance of the line-of-sight velocity measurements, we have been able to quantify and remove this noise from all of the scans except for the RHIs (as is discussed in Sect. 2). As such, random errors are not anticipated to be a significant source of error except for the RHIs, as discussed in the manuscript. With this approach, we feel that a Monte Carlo simulation is not appropriate to assign errors, as these noise effects are already quantified and removed.

The negative TKE values are simply outliers based on the definition used in the manuscript, being more than 1 order of magnitude difference between sonic anemometer and observations (see p. 13, line 20). As mentioned above, quantifying the sampling error of six-beam measurements is difficult due to the spatio-temporal nature of the technique and out of the scope of the paper. However, we do believe the negative values are thought to be related to the sampling error of the measurement (since there is no other plausible explanation), and state this where the negative values are first discussed on p.12 line 2.

3.) p9 l18: "These erroneous echoes were removed using a discontinuity-based algorithm described by Bonin and Brewer (2016)" -> Maybe it is not so easy. Such a correction is never perfect and some artifacts always remain. The kind of signal folding you experience imposes spatially confined biases on the measured signal (spanning some range gates). Some techniques may be more susceptible to this influence than others, introducing an unknown bias into the intercomparison. E.g., the six beam technique will be affected differently by spatially confined shifts in a single beam than the RHI scans. Signal folding is also no necessity for Doppler lidar measurements. They can be avoided by reducing the pulse repetition rate, which should be mentioned. Several other questions arise and have to be discussed: (a) How is it possible to identify the folding effect unambiguously in the data? (b) What percentage of data is affected? (c) What is the remaining bias after correction?

We agree with the reviewer that the cited technique to remove range folded echoes is not perfect, largely since it relies on contextual information. This technique works better on RHI and PPI scans than it does on measurements at single beam positions, due to the necessity of having data nearby spatially and temporally. While showing and discussing the results of the intercomparison, we specifically point out in the manuscript that the large number of outliers in the six-beam TKE measurements are largely a consequence of range folded echoes not being removed due to short amount of time at each beam position (see p. 13, lines 22-28).

We acknowledge that range folding does not appear in all Doppler lidar measurements. However, the commercial Doppler lidar systems that have been increasingly used by a diverse set of users in the past 5 years (Leopshere and Halo systems) have PRFs ~10-20 kHz and have been documented to be susceptible to range folded echoes (see Päschke et al. (2015) and Bonin and Brewer (2017)). By clearly stating (see p. 9, line 17) that the high PRF (20 kHz) of the system is why range folding is an issue, it is implied that range folding is not an issue for low PRF systems. Thus, we do not see a need to expand upon the issue in the manuscript.

While these anomalous echoes do affect the data quality and we feel that they are important to mention, they are not the main focus of the manuscript and much of the data (as determined by manual inspection) is not affected after the described and referenced QC methods are applied. Going into details about these echoes, their characteristics, frequency of occurrence, etc. requires rigorous analysis itself and is outside of the scope of the paper. Instead, we have added a statement in the manuscript to refer the reader to Bonin and Brewer (2017) for these characteristics and how these anomalous echoes are identified.

We have provided a statement in the manuscript indicating that 5.6% of data points were flagged and removed, which we believe to be an overestimate (since we conservatively remove suspect data points). To really address this issue, a low-PRF system, such as our HRDL, needs to be synchronized with the high-PRF system scanning the same positions at the same time, to accurately quantify the percentage of data from that high-PRF system that is range folded through a direct intercomparison of measurements along each beam. This was not done during XPIA, and to our knowledge has not been done any other time before. Thus, the exact percentage of data affected by range folding is not known, until we can perform such as experiment.

Minor issues:

p2 l7: "good": -> As described above, please quantify...

We have decided to simply remove the word 'good', and provide a reference for the reader if they are interested in the details of the results of the cited study.

p2 l10 "the long time series of staring": -> Again, please quantify. It is actually a very good question what "long" means here. You correctly cite Lenschow et al. (1994) here, but leave the calculations to the reader. Please give a rough estimation of what "long" means in this context.

As we have rewritten much of the introduction, there is no obvious location to discuss this here. However, we have added a paragraph in the beginning of Sect. 4 discussing sampling errors, in which we give an estimation of how long a time series needs to be under convective conditions (1-2 hours) and stable conditions (5-10 min).

p1 l4: "trusted in situ instrumentation": -> I think I know what you mean, but please give a reference of what "trusted" means in this context. Do you mean something like "officially approved by a standardization institution"?

We have clarified in the introduction at p. 3 l 15 that sonic anemometers are a commonly used reference device, and have provided a reference to International Energy Agency (IEA) report that supports this claim. We have decided to keep the wording unchanged in the abstract (at p1 l4), as we feel it is not the appropriate location to give a reference for this.

p1 l12: "None of the methods evaluated were able to consistently accurately measure the shear velocity" -> Please discuss what accuracy is necessary to measure shear velocity "accurately". Which maximum error is allowed for which purpose?

As the reviewer states, the maximum error allowed for a measurement to be useful depends on the purpose or use of the measurement. We do provide a general discussion on how accurate measurements need to be depending on the application in the paragraph starting on p. 2 l. 33. But we do not believe that the essence of this study is to sort out whether a measurement is accurate enough for a given purpose. That is for users to decide based on their specific objectives. Instead, the objective here is to analyze the measurement techniques to determine systematic biases and if measurements are of any use at all to any audience. In the particular sentence referred to here, we state that stress velocity was not found to be accurate by any of the methods tried. This is true regardless of the use of the measurement as the correlation coefficient between the sonic and VAD or six-beam u-star was 0.171 and 0.147 respectively and the scatter was large across all values, indicating very little correlation, and that the lidar measured u-star was insufficiently accurate for any purpose.

p6 l7: Since data were collected at 2 Hz, two samples were collected 0.5 s apart -> Please decide between mentioning "2Hz" or "0.5s".

We have removed the 2 Hz statement and now are using only 0.5 s here.

p8 I4: SNR<-27 dB -> which definition of SNR is applied here?

A statement has been added here to clarify that the SNR values were taken as the carrier-to-noise ratio produced by the lidar manufacturer's processing algorithms. Since the processing algorithms are proprietary, we are unsure about the details of how the SNR/CNR is exactly calculated and the exact definition used. However, the values of SNR/CNR should be comparable for measurements from other Leosphere Doppler lidars.

p16 I1: "show skill" -> Please define "skill" together with the other descriptions.

We have added a definition of 'skill' here as showing correlation between the lidar and the sonic anemometer (reference instrument).

Typos:

p1 I11: Typo: "to biased" -> "to be biased"

Corrected

Table2: typo at "0.547nb"

Corrected

# Evaluation of Turbulence Measurement Techniques from a Single Doppler Lidar

Timothy A. Bonin<sup>1,2</sup>, Aditya Choukulkar<sup>1,2</sup>, W. Alan Brewer<sup>2</sup>, Scott P. Sandberg<sup>2</sup>, Ann M. Weickmann<sup>1,2</sup>, Yelena Pichugina<sup>1,2</sup>, Robert M. Banta<sup>2</sup>, Steven P. Oncley<sup>3</sup>, and Daniel E. Wolfe<sup>4</sup>

<sup>1</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado, USA

<sup>2</sup>Chemical Sciences Division, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA

<sup>3</sup>National Center for Atmospheric Research, Boulder, Colorado, USA

<sup>4</sup>Physical Sciences Division, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA

Correspondence to: T. A. Bonin (timothy.bonin@noaa.gov)

**Abstract.** Measurements of turbulence are essential to understand and quantify the transport and dispersal of heat, moisture, momentum, and trace gases within the planetary boundary layer. Through the years, various techniques to measure turbulence using Doppler lidar observations have been proposed. However, the accuracy of these measurements has rarely been validated against trusted *in situ* instrumentation. Herein, data from the eXperimental Planetary boundary layer Instrumentation Assessment (XPIA) are used to verify Doppler lidar turbulence profiles through comparison with sonic anemometer measurements. For 17 days at the end of the experiment, 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.

Measurements of turbulence kinetic energy, turbulence intensity, and ~~shear~~-stress velocity from these techniques are compared with sonic anemometer measurements at six heights on a 300-m tower. The six-beam technique is found to generally measure turbulence kinetic energy and turbulence intensity the most accurately at all heights ( $r^2 \approx 0.78$ ), showing little bias in its observations (slope of  $\approx 0.55$ ). Turbulence measurements from the velocity azimuth display method tended to be biased low near the surface, as large eddies were not captured by the scan. None of the methods evaluated were able to consistently accurately measure the shear velocity ( $r^2 = 0.15 - 0.17$ ). Each of the scanning strategies assessed had its own strengths and limitations that need to be considered when selecting the method used in future experiments.

## 1 Introduction

Turbulence within the planetary boundary layer (PBL) transports and disperses heat, moisture, momentum, and other quantities. Additionally, atmospheric turbulence affects several disciplines and industries, such as wind energy, aviation, and air quality. For example, wind turbines may perform poorly and have a lower power output when turbulence intensity is large (Wharton and Lundquist, 2012; Choukulkar et al., 2016), and turbulence can shorten the lifespans of wind turbines (Kelley et al., 2006). Pollutant dispersion from factories and other sources is primarily driven by advection and turbulent mixing within the PBL. Precise measurements are necessary to understand the role of turbulence within these disciplines, and to validate the turbulence



generated or parameterized in numerical weather prediction models and simulations. Scanning Doppler lidars are capable of addressing this need, by measuring vertical profiles of turbulent quantities throughout the entire PBL.

Turbulence has been measured using multiple Doppler lidars in recent years. Mann et al. (2009) performed the first intercomparison between triple-Doppler-lidar turbulence measurements with sonic anemometry, and investigated the effects of volume averaging. Fuertes et al. (2014) used three Doppler lidars that were staring at the same point for 1-hr to retrieve a 0.5 Hz timeseries of  $u$ ,  $v$ , and  $w$ . Many different approaches have used to measure turbulence quantities with Doppler lidar (Banta et al., 2013b; Sathe and Mann, 2013; S). Prior to the availability of commercial Doppler lidars, it was necessary to employ techniques for single lidars, which are the techniques evaluated in the present study. In the most simple case, data from fixed-beam vertical staring has been used to directly calculate vertical velocity  $w$ . From these measurements, all six components of the Reynolds stress tensor statistics in the unstable convective boundary layer (e.g., Mayor et al., 1997; Lenschow et al., 2000; Lothon et al., 2006, 2009; Pearson et al., 2009; Ba). More sophisticated techniques take advantage of scanning to provide other turbulence components and quantities. Eberhard et al. (1989) adapted a technique developed for scanning Doppler radar (Wilson, 1970; Kropfli, 1986; Frisch et al., 1989) that used 360° azimuth conical (plan position indicator, PPI) scans to measure profiles of turbulence kinetic energy (TKE), individual velocity variances ( $\overline{u'^2}$ ,  $\overline{v'^2}$ ,  $\overline{w'^2}$ ), and momentum fluxes ( $\overline{u'w'}$ ,  $\overline{v'w'}$ ,  $\overline{u'v'}$ ) were directly calculated, showing good agreement with sonic anemometer observations. Newman et al. (2016a) used a similar method to measure velocity variances over several days, for comparison with other turbulence measurements. Additionally, Newman et al. (2016a) extended the use of 'virtual towers', which are constructed using measurements from multiple lidar systems, to measure vertical profiles of turbulence quantities, whereas in previous work 'virtual towers' were used to make . Banta et al. (2002, 2006) used elevation (range-height indicator, RHI) scans, which provide a vertical cross section of radial wind-velocity data points, to calculate vertical profiles of the mean wind (e.g. Calhoun et al., 2006; Gunter et al., 2015). However, the long timeseries of staring data necessary to measure turbulence with meaningful significance (see Lenschow et al., 1994) precludes the possibility of making accurate turbulence profiles routinely, especially in the convective boundary layer where turbulent length scales are large. Doppler lidars are also expensive and many research groups do not have access to multiple systems. Hence, the focus in this paper will be on scanning strategies that provide vertical profiles of turbulence from a single Doppler lidar. streamwise variance, which was found to be approximately equal to TKE in the stable conditions studied. Pichugina et al. (2008b) evaluated these values against sonic anemometers at several levels on a 120-m tower and found good correlations. These scanning techniques will be further evaluated in the present study.

Over the years, several different techniques have been proposed to measure turbulence using measurements from a single Doppler lidar (Sathe et al., 2015a). A review of wind lidar turbulence measurements, through 2013, is provided by Sathe and Mann (2013). Since then, much of the recent work has focused on evaluating and improving turbulence measurements from the For instrument systems that lack full scanning capability, a simpler approach is the so-called Doppler beam swinging (DBS) technique , which is often used by low-powered Doppler lidars typically used in the wind energy field (e.g., Lundquist et al., 2015; Newman et al., 2016b; Kun). Herein, the DBS technique will not be evaluated since the exact methodology for the retrieval of turbulence statistics is still under development, with corrections being applied to reduce the effect of variance contamination and other error sources. Instead, three other scanning strategies were evaluated to determine their accuracies for measurement of turbulent quantities. The earliest proposed method to measure turbulence kinetic energy (TKE), individual velocity variances ( $\overline{u'^2}$ ,  $\overline{v'^2}$ ,  $\overline{w'^2}$ ), and

covariances ( $\overline{u'w'}$ ,  $\overline{v'w'}$ ,  $\overline{u'v'}$ ) from Doppler lidar observations is discussed by Eberhard et al. (1989), using the residuals of a velocity-azimuth-display (VAD) fitting on a conical plan position indicator (PPI) scan. Banta et al. (2002, 2006) used vertical-slice range-height-indicator (RHI) scans to quantify the streamwise velocity variance  $\overline{u'^2}$  within the PBL during nocturnal low-level jets. Vertical stares have been used in numerous studies (e.g. Pearson et al., 2010; Barlow et al., 2011; Bonin et al., 2015) to measure the boundary layer depth, vertical velocity spectra and variance, and dissipation. Recently, Sathe et al. (2015b) has proposed a method of using a typically used by radar wind profilers and Doppler sodar. In this method the transmitted beam cycles among (typically) five discrete fixed look angles, one vertical and four beams tilted at some elevation angle but aimed in four orthogonal horizontal directions, such as the four cardinal directions. While Sathe et al. (2011) identified problems with the computational procedure using these orthogonal beams to measure turbulence quantities, the technique has been modified in various ways to correct for these limitations (e.g., Lundquist et al., 2015; Newman et al., 2016b; Kumer et al., 2016). This problem can also be addressed using a six-beam strategy, which is essentially a modified DBS scan that typically uses five beams, to measure velocity variances, covariances, and TKE. Each of these strategies will be described in more detail within Sect. 2, and evaluated herein: five tilted beams instead of four plus a vertical (Sathe et al., 2015b). This six-beam technique will also be evaluated in this study.

These single-lidar, multiple look-angle techniques are potentially very powerful, given the importance of being able to measure vertical profiles of turbulence (Banta et al., 2013b), but a crucial assumption is that the mean flow and turbulence need to be homogeneous over the horizontal sampling footprint at each measurement height in space and time. When commercial Doppler lidars became available, it became possible to deploy three or more lidars to simultaneously and continuously sample a given sampling volume in space, thereby enabling all turbulence components to be measured directly without assumption. Mann et al. (2009), Fuertes et al. (2014), and Newman et al. (2016a) used triple-Doppler lidar arrays in this way to measure the six components of the Reynolds stress tensor. Although we do not evaluate multiple-Doppler techniques in this study, it is important to be aware of these capabilities in designing future multiple Doppler measurement programs, since they would be available as a component of a turbulence verification effort.

Turbulence measurements are needed to address a range of problems, which involve different breadths of the turbulence spectrum. All applications require accurate measurements of fluctuations by the largest energy-containing turbulent eddies and at least the lowest wavenumbers of the inertial subrange. Within the inertial subrange the magnitude of the fluctuations drops off quickly (exponentially) with increasing wavenumber, so high wavenumbers make correspondingly smaller contributions to the total variances (Taylor, 1938). Detailed studies of turbulence dynamics, which may include studies of inflows to wind turbines or turbulence generated by them, may require accurate representation of fluctuations over the entire turbulence spectrum from large-eddy to dissipation scales (e.g., Troldborg and Sørensen, 2014). For such studies, employing the best data acquisition strategies and understanding the errors involved is important. Other studies may not require this degree of precision. For example, evaluating the ability of NWP numerical weather prediction (NWP) models to predict TKE involves values of  $2\text{--}4\text{ m}^2\text{ s}^{-2}$  in convective conditions and  $1\text{--}2\text{ m}^2\text{ s}^{-2}$  in weakly stable conditions. Such accuracies are achievable without measuring the entire spectrum. Many field programs are employing scanning lidar remote sensing in arrays to investigate spatial and temporal variations of the mean wind, as recommended in Banta et al. (2013a). In such cases, the measurement of turbu-

lence is not the primary goal, so the data-acquisition and scanning approaches are not optimized for turbulence measurement. It is still desirable to obtain quantitative turbulence information (e.g., for NWP verification) from the scans that are performed. It is essential to understand the error properties of these techniques to know whether the calculated values are useful for the intended purpose.

5 To systematically evaluate these different turbulence measurement techniques, a Doppler lidar cycled each hour continuously through the methods during the last two weeks of the eXperimental Planetary boundary layer Instrument Assessment (XPIA) field campaign. These measurements are compared with ~~sonic anemometer measurements~~ measurements from sonic anemometers, a commonly used reference instrument (Sathe et al., 2015a), at six heights on a 300 m meteorological tower located 540 m from the lidar. Through this comparison, the following questions will be addressed in this study.

- 10
- How accurate are the various single-Doppler turbulence measurement strategies in determining turbulence characteristics? Does the accuracy vary depending on the measurement height?
  - What main caveats need to be considered when applying each technique? How should random errors and instrument noise be characterized and treated?
  - What is the optimal operational scanning strategy to derive turbulence estimates? Should different strategies be used for
- 15 different objectives?

To address these questions, the paper is organized as follows. ~~An overview of the experiment and the instrumentation used is detailed in Sect. 3.~~ The various scanning strategies and methods to measure turbulence, including specific details of implementation, are described in Sect. 2. An overview of the experiment and the instrumentation used is detailed in Sect. 3. Within Sect. 4, the techniques are statistically compared through validation with sonic anemometry. Implications for future studies and possible future research directions are discussed within Sect. 5. A summary and the conclusions are provided in Sect 6.

20

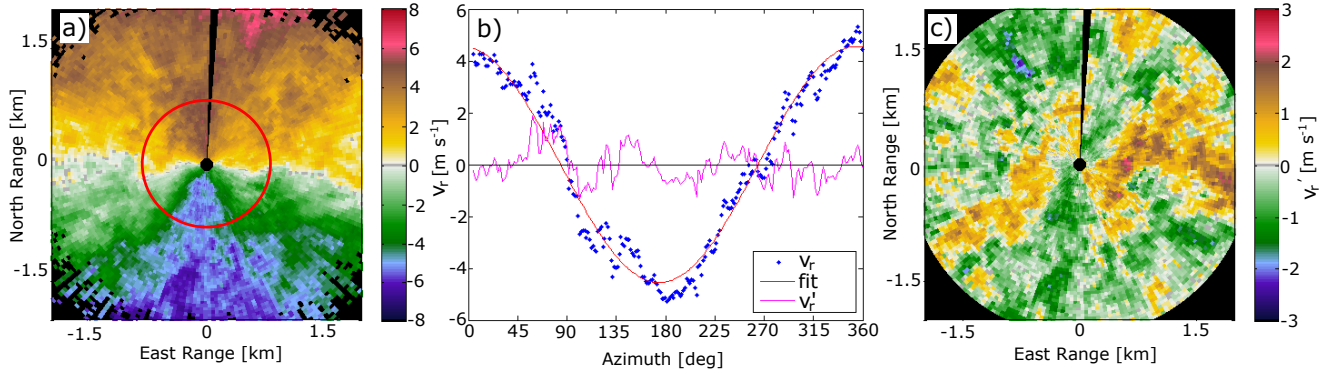
## 2 Turbulence measurement strategies

The scanning procedures used most often by Doppler lidars are azimuthal scanning, elevation scanning, and stares at ~~one a~~ particular look angle. Each of these approaches can be used to measure one or more of the velocity variances and covariances.

25 The theory for turbulence measurements is based on the relationship between the observed radial velocity  $v_r$  and the flow within the resolution volume given by

$$v_r = u \cos \theta \cos \phi + v \sin \theta \cos \phi + w \sin \phi + \epsilon, \quad (1)$$

wherein  $u$  is the streamwise horizontal velocity,  $v$  is the crosswise horizontal velocity,  $w$  is the vertical velocity,  $\phi$  is the elevation angle above the horizon,  $\theta$  is the angle between  $u$  and the azimuth of the lidar, and  $\epsilon$  is uncorrelated random error in



**Figure 1.** Sample PPI scan (a) during a turbulent time period, with the VAD fitting and its residuals to  $v_r$  observations at the range ring denoted by the red circle shown in (b). Turbulence structures can be visualized in the residuals across the entire scan (c).

the measurement. The value of  $\epsilon$  typically increases with range from the lidar, as the signal-to-noise ratio (SNR) decreases. By squaring Eq. 1 and removing the mean from each quantity, the radial velocity variance is given by

$$\overline{v_r'^2} = \overline{u'^2} \cos^2 \theta \cos^2 \phi + \overline{v'^2} \sin^2 \theta \cos^2 \phi + \overline{w'^2} \sin^2 \phi + 2\overline{u'v'} \sin \theta \cos \theta \cos^2 \phi + 2\overline{u'w'} \cos \theta \cos \phi \sin \phi + 2\overline{v'w'} \sin \theta \cos \phi \sin \phi + \overline{\epsilon^2}, \quad (2)$$

where the covariance terms involving  $\epsilon$  are zero since it is uncorrelated. All of the turbulence measurements techniques are ultimately based on Eq. 2. Brief derivations and details of how these measurements are made, in addition to modifications introduced within this study, are described here. Complete derivations for each method can be found in the works cited.

## 2.1 ~~Velocity-azimuth~~ Velocity-azimuth display

While PPI scans have been used to take accurate measurements of the mean wind through ~~VAD~~ velocity-azimuth display (VAD) analysis (Smith et al., 2006), these scans can also be used to quantify turbulence. Eberhard et al. (1989) details a technique for measuring turbulence from PPI scans, based on pioneering work by Wilson (1970) and Kropfli (1986) wherein turbulence is measured using Doppler radar observations. From PPI scans at two sufficiently different elevation angles, all six components of the Reynolds stress tensor can be retrieved using the residuals of the VAD fitting by utilizing a partial Fourier decomposition of Eq. 2. However, the covariances ~~and/or~~ momentum fluxes  $\overline{u'v'}$ ,  $\overline{u'w'}$ , and  $\overline{v'w'}$  can be measured from any single PPI scan, and TKE can be obtained from a single scan if  $\phi = 35.3^\circ$  (for mathematical basis, see Eq. 4a in Eberhard et al., 1989).

A sample PPI scan for a turbulent time period is shown in Fig. 1a. For each range ring, the mean wind speed and direction are determined using VAD analysis. The complete VAD analysis described by Browning and Wexler (1968) includes terms for the vertical velocity of the scatterers as well as horizontal divergence, stretching deformation, and shearing deformation. However, a more simplified variation of the VAD analysis is often used by neglecting divergence and the deformation terms.

For the results presented here, the simplified form is used since it yields more accurate estimates of the measured turbulent quantities when compared with sonic anemometer measurements. This may be due to variability from large turbulent motions being incorrectly partitioned into the divergence or deformation terms. However, in complex terrain or other locations these terms may not be negligible. An example of the fitting of this equation and its residuals  $v_r'$ , which are deviations from the expected mean  $v_r$ , is shown in Fig. 1b. If the mean flow (i.e.,  $\bar{u}$ ,  $\bar{v}$ , and  $\bar{w}$ ) is homogeneous over the scanning circle, then the residuals of the fitting are results of turbulent motions and  $\epsilon$ . This is visualized within Fig. 1c, wherein coherent areas of positive and negative  $v_r'$  represent turbulent eddies. Since turbulent structures are correlated spatially,  $\epsilon$  can be quantified and removed by applying a structure function fit to the autocovariance of  $v_r'$  across radials for a given range gate ~~using the technique outlined by~~ similar to the method outlined using Eq. 32 in Lenschow et al. (2000). To our knowledge, this is the first time that the autocorrelation technique has been used to remove noise variance from a scan, as it is typically used for a timeseries from prolonged stares (Mayor et al., 1997). This technique can lead to an overestimate of  $\epsilon$  when the inertial subrange is smaller than the distance between adjacent azimuths, which is more likely at long ranges from the lidar as the spatial separation between adjacent beams increases.

Previously, measurements using the technique described by Eberhard et al. (1989) have not been evaluated against *in situ* observations. Wang et al. (2015) used a variation of this technique by applying it to a  $30^\circ$  sector PPI and assumed isotropic turbulence to relate  $\overline{v_r'^2}$  to TKE. Estimates of TKE from the arc scan showed good agreement ( $r^2 = 0.89$ ) with those from sonic anemometer on a linear scale. Other studies (e.g., Berg et al., 2013; Sathe et al., 2015b) have used a loose variation of this VAD technique to quantify turbulence by using only a small number of beams (4 – 6) spaced around the entire  $360^\circ$ , which is substantially different from using more than 100 beams around the sampling ring.

## 2.2 Six-beam

Sathe et al. (2015b) propose a technique to measure all six components of the Reynolds stress tensor by continuously cycling between measurements at six different angles. One beam is vertical, and the other five are at a set elevation angle ( $45^\circ$  herein) and are equally spaced  $72^\circ$  apart in azimuth. For each beam, the time series of  $v_r$  are linearly detrended over a fixed time window, which is 20-min here, and  $v_r'$  is computed as its residual. ~~A~~ While the 20-min detrending window may filter out large convective eddies when the wind speed is small, a shorter window or a higher-order detrending ~~may result in effectively removing large-scale~~ would exacerbate this filtering effect for smaller eddies. Values of  $\overline{v_r'^2}$  are computed for each beam separately. Thus, there are six known values of  $\overline{v_r'^2}$ , one for each beam, and each is a function of differently weighted velocity

variances and covariances based on the scan elevation and azimuth, as in Eq. 2. This can be represented by the matrix relationship

$$M \begin{bmatrix} \overline{w'^2} \\ \overline{v'^2} \\ \overline{w'v'} \\ \overline{u'v'} \\ \overline{u'w'} \\ \overline{v'w'} \end{bmatrix} = \begin{bmatrix} \overline{v_{r1}'^2} \\ \overline{v_{r2}'^2} \\ \overline{v_{r3}'^2} \\ \overline{v_{r4}'^2} \\ \overline{v_{r5}'^2} \\ \overline{v_{r6}'^2} \end{bmatrix}, \quad (3)$$

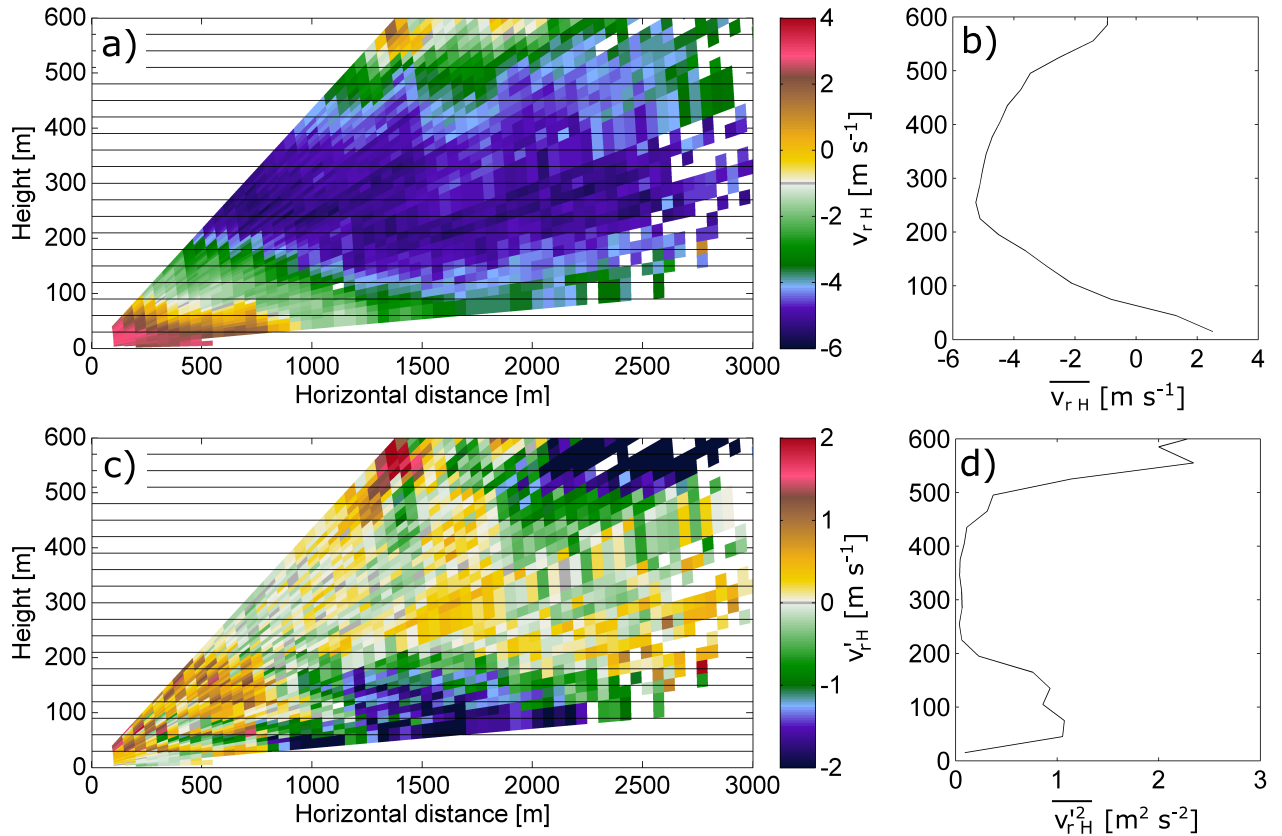
where  $M$  is a 6 by 6 matrix of coefficients based on different combinations of  $\theta$  and  $\phi$ , as in Eq. 2. Thus, it is possible to solve for the six unknown components of the Reynolds stress tensor through an inversion of Eq. 3.

For each beam, the lidar started at the given location for 1 s collecting two samples 0.5 s apart before advancing to the next position. ~~Since data were collected at 2 Hz, two samples were collected 0.5 s apart.~~ To remove uncorrelated noise  $\epsilon^2$  from the observed  $\overline{v_r'^2}$  for each beam, the autocovariance at the first lag for the samples that were 0.5 s apart was taken as  $\overline{v_r'^2}$ , following the technique presented by Lenschow et al. (2000). This likely results in a slight underestimate of  $\overline{v_r'^2}$ , since contributions from small eddies that are uncorrelated over short timescales are removed. In the future, it is recommended that more samples be collected along each beam so that a structure-function or linear fitting may be applied to the autocovariance for a more robust measurement of  $\overline{v_r'^2}$  for each beam. On average, the scanner took  $\approx 3.6$  s to slew between beam positions, so that the scanner returned to the same beam every  $\approx 27$  s.

The measured  $v_r'^2$  on the right-hand side of Eq. 3 include the desired turbulent fluctuations, but also variations due to spatial and temporal sampling by beams aimed at very different directions. Under some conditions, mostly when the turbulence was weak, the matrix-inversion calculation in Eq. 3 can lead to negative values for one or more of the calculated variances, which should be positive-definite quantities. This result, which has also been reported by Newman et al. (2016b), is a non-physical computational artifact, thought to be primarily due to sampling errors and nonturbulent variations within each beam-variance being propagated through the matrix inversion in Eq. 3, resulting in spurious values of the calculated variances. Non-physical negative variance estimates have been removed in the following analyses.

### 2.3 RHI scans and vertical stares

Shallow RHI scans have also been used as a means to measure horizontal velocity variances (e.g., Banta et al., 2006; Pichugina et al., 2008b). These scans are conducted by scanning from the horizon up to  $\approx 30^\circ$ , typically at two angles orthogonal to each other. Since the scans are mostly at low angles, it is assumed that the observed  $v_r$  are due to the horizontal wind and that the



**Figure 2.** Sample RHI scan (a) showing instantaneous values of  $v_{rH}$  over the scan plane. (b) Vertical profile of the mean  $v_{rH}$  for the scan, which is used to calculate  $v'_{rH}$  shown (c). For each height bin between the solid horizontal black lines on (a) and (c), the variance of  $v'_{rH}$  is calculated resulting in the profile shown in (d).

contribution of  $w$  is negligible. To ensure that measurements at different elevation angles are comparable, values of  $v_r$  are normalized by  $\phi$  by

$$v_{rH} = \frac{v_r}{\cos \phi}, \quad (4)$$

where  $v_{rH}$  is the radial velocity projected in the horizontal. For each RHI, observations are binned by height (30-m bins used herein) which are used to make a mean profile of  $v_{rH}$ . The profile of  $\overline{v_{rH}}$  is used to calculate deviations from the mean flow  $v'_{rH}$ . This is done by simply taking the difference between  $v_{rH}$  and  $\overline{v_{rH}}$  for the given height, where  $\overline{v_{rH}}$  is linearly interpolated between the center of mass of each height grid. The variance of  $v'_{rH}$  is calculated using the same height grid to produce a profile of  $\overline{v'^2_{rH}}$ , which is the horizontal wind variance within the RHI plane. An example of this process and each derived product is provided in Fig. 2.

**Table 1.** Summary of measured variables for each type of scanning strategy.

	$\overline{u'^2}$	$\overline{v'^2}$	$\overline{w'^2}$	$\overline{u'v'}$	$\overline{u'w'}$	$\overline{v'w'}$	TKE
VAD (single $\phi$ )				✓	✓	✓	✓ ( $\phi = 35.3^\circ$ )
VAD (two $\phi$ s)	✓	✓	✓	✓	✓	✓	✓
Six-beam	✓	✓	✓	✓	✓	✓	✓
RHI & vertical stare	✓	✓	✓				✓

When one of these scans is oriented with the mean flow and the other transversely, the two measured profiles of  $\overline{v_{rH}^{\prime 2}}$  can be treated as  $\overline{u'^2}$  and  $\overline{v'^2}$  respectively. If the scans are not oriented in such a way or if large directional shear is present, it is possible to rotate the variances to be aligned with the mean flow by

$$\overline{u'^2} = \overline{v_{rH1}^{\prime 2}} \cos^2 \Theta_1 + \overline{v_{rH2}^{\prime 2}} \cos^2 \Theta_2 - \overline{v_{rH1}^{\prime} v_{rH2}^{\prime}} \sin 2\Theta_1 \text{ and} \quad (5)$$

$$5 \quad \overline{v'^2} = \overline{v_{rH1}^{\prime 2}} \sin^2 \Theta_1 + \overline{v_{rH2}^{\prime 2}} \sin^2 \Theta_2 + \overline{v_{rH1}^{\prime} v_{rH2}^{\prime}} \sin 2\Theta_1, \quad (6)$$

wherein  $\Theta$  is the angle between the RHI scan azimuth and the mean flow, the subscripts denote the two different RHI scan planes, and the scans are orthogonal. Although the covariance term  $\overline{v_{rH1}^{\prime} v_{rH2}^{\prime}}$  cannot be measured with this method, it **becomes zero if turbulence is assumed to be isotropic** is typically small compared to the other terms and can be neglected. Thus, values of  $\overline{u'^2}$  and  $\overline{v'^2}$  can be computed directly through the rotation. The mean wind profile, including wind speed and direction necessary for the rotation, are directly computed using the two profiles of  $\overline{v_{rH}}$ . Using this technique, there is no straightforward way to remove contamination from  $\overline{\epsilon^2}$  in the variances. Thus, data were removed if the  $\text{SNR} < -27$  dB to reduce contamination from highly noisy data. The SNR values used for filtering were taken as the carrier-to-noise ratio (CNR) produced by the lidar manufacturer's processing algorithms.

To calculate TKE, values of  $\overline{w'^2}$  also need to be known. For quantification of  $\overline{w'^2}$ , vertical stares were used in conjunction with the shallow RHI scans. Vertical stares are the most straightforward method to measure any vertical turbulent quantity with a Doppler lidar. Since the  $w$  profile is continuously measured, it is simple to take the variance of time series of  $w$  to obtain  $\overline{w'^2}$ . However,  $\epsilon^2$  contaminates the measurement and needs to be removed to improve the accuracy of the measurement. As described earlier, the autocovariance technique described by Lenschow et al. (2000) is used to remove instrument noise. Herein, values of  $\sigma_w^2$  are taken as the extrapolated  $-2/3$  structure function fit to the autocovariance of the time series at lags 1-5. Using this technique removes contamination by  $\epsilon^2$  and mitigates volume averaging effects, which otherwise reduce the observed  $\overline{w'^2}$  (Bonin et al., 2016).

### 3 Experimental overview



~~The eXperimental Planetary-boundary-layer Instrument Assessment (XPIA) field campaign was conducted~~

~~A Leosphere WINDCUBE 200S<sup>®</sup> was operated at the now-defunct Boulder Atmospheric Observatory (BAO) in the spring of 2015. The primary goal of the XPIA experiment was to assess the capabilities for measuring flow in the planetary boundary layer with current technology such as wind-profiling radars, microwave radiometers, sonic anemometers, various Doppler lidar systems, and other additional instrumentation. Herein, observations from a single scanning Doppler lidar system are compared with those from sonic anemometers on a meteorological tower. Pertinent details about the sensors and site are described here, but a complete list of the instrumentation and a thorough description of the experiment along with its various objectives can be found in Lundquist et al. (2017).~~ during 15-31 May 2015, a total of 17 days. The sampling period for the present study immediately followed the eXperimental Planetary-boundary-layer Instrument Assessment (XPIA) field campaign during which a significant complement of remote-sensing instruments, including six Doppler lidars, were operated in the vicinity of the tower (Lundquist et al., 2017). The BAO featured a 300-m meteorological tower instrumented at multiple levels. The Doppler lidar system was deployed 540 m to the south-southwest of the 300-m tower, as shown in Fig. 3.

The BAO was located in Erie, CO, approximately 25 km east of the foothills of the Rocky Mountains, and was designed primarily for PBL research as well as testing and calibration of various atmospheric sensors (Kaimal and Gaynor, 1983). Within the immediate vicinity, the terrain is relatively flat with gently rolling terrain. ~~A~~The 300-m tower was located on the property. For this experiment, 3-D Campbell CSAT3 sonic anemometers were installed on northwest (NW, 334°) and southeast (SE, 154°) booms at six levels (50, 100, 150, 200, 250, and 300 m). Data were recorded at 20 Hz. A tilt-correcting algorithm that used a planar fit (Wilczak et al., 2001) was applied to the measurements after the experiment was finished. Data were filtered to remove time periods when the turbulence may be affected by the wake of the tower, following the results of McCaffrey et al. (2016). Specifically, data from the NW and SE sonics are removed when the wind direction is from between 100–170° and 300–20° respectively. Turbulence statistics from the sonic anemometers were averaged over 20-min blocks for comparison, similar to the averaging time for the various lidar scanning strategies discussed below. Any 20-min averages where the statistics between the two sonic anemometers at the same height differed by a factor of 2 or more were removed, to ensure the statistics were comparable and not affected by the tower.

~~The modest complexity of the terrain and the proximity of the site to the mountains present complications in calculating turbulence quantities from remote-sensing data using the techniques described here. Under these conditions, the flow can exhibit non-turbulent variability along a scan that contribute to unknown degree to the calculated variances and covariances producing larger variance and discrepancies between lidar and tower measured turbulent quantities. This variability is not expected over more homogeneous topography (e.g., Pichugina et al., 2008a) or the ocean (e.g. Tueker et al., 2009).~~

Measurements from only one Leosphere WINDCUBE 200S<sup>®</sup> are used in this study, although six scanning Doppler lidars were deployed during XPIA. At the end of the experiment from 15–31 May 2015, the system was deployed 540 m to the south-southwest of the 300-m tower, near the Visitor Compound, as shown in Fig. 3. The lidar was operating with a ~~The~~ lidar operated with 50-m pulse width and range gate size range-gate spacing and at a nominal pulse length of 50 m, meaning that most of the transmitted pulse energy lies within a 50-m window. The remnant of pulse energy outside this window is too weak to significantly affect the velocity calculation for normal atmospheric aerosol-backscatter conditions, although



**Figure 3.** Satellite imagery of the BAO site with the locations of the 300-m tower and Doppler lidar deployment indicated.

interceptions of the pulse by hard targets (such as clouds or wind turbines) can contaminate the return signal in adjacent range gates (see Banta et al., 2015). The accumulation time for each beam was 0.5 s and the pulse repetition frequency (PRF) was 20 kHz. Due to the high PRF, the maximum unambiguous range of the lidar was only 7500 m. Hence, range folded echoes from clouds or other strong targets were occasionally apparent in the signal. ~~These erroneous echoes were removed using~~

5 Characteristics of these erroneous echoes and a discontinuity-based algorithm used to remove them are described by Bonin and Brewer (2017). Over the 17 days, 5.6% of the data were conservatively detected as possible range folded echoes and removed. Each hour, a sequence of scanning strategies were conducted. For 20-min, the lidar cycled between PPI scans at a  $\phi$  of  $35.3^\circ$  and  $50.8^\circ$  scanning at  $3^\circ \text{ s}^{-1}$  in azimuth. Over this 20 min period, five scans were completed at each elevation angle. Following the PPI scans, three shallow RHIs were performed at perpendicular angles ( $\theta$  of  $330^\circ$  and  $60^\circ$ ), followed by a 10-min

10 vertical stare. For the rest of the hour, the six-beam scanning strategy was repeated, wherein each beam was sampled for 1 s before advancing to the next beam position as described in Sect. 2.2. This 1-hr scanning sequence was repeated continuously for the 17 days at the end of XPIA.

The modest complexity of the terrain and the proximity of the site to the mountains present complications in calculating turbulence quantities from remote-sensing data using the techniques described here. Under these conditions, the flow can

15 exhibit non-turbulent variability along a scan that contributes to an unknown degree to the calculated variances and covariances producing larger variance and discrepancies between lidar and tower measured turbulent quantities. This variability is not expected over more homogeneous topography (e.g., Pichugina et al., 2008a) or the ocean (e.g., Tucker et al., 2009).

#### 4 Turbulence statistics comparison

For most measurements, turbulent quantities measured by the Doppler lidar are not at precisely the same height as the sonic anemometers. This difference in measurement height is dependent on the type of scan. For instance, the range gate center for vertical scans is identical to the height the sonic anemometers at 100–300 m (50-m was below the minimum range). However, the closest range gate from the six-beam scan to the 300-m sonic was at 282.8 m. Thus, lidar-measured turbulent quantities are interpolated to the sonic anemometer heights.

Depending on the application or field of use, different turbulent quantities may be desired. Within the wind energy industry, turbulence intensity  $TI$  calculated as

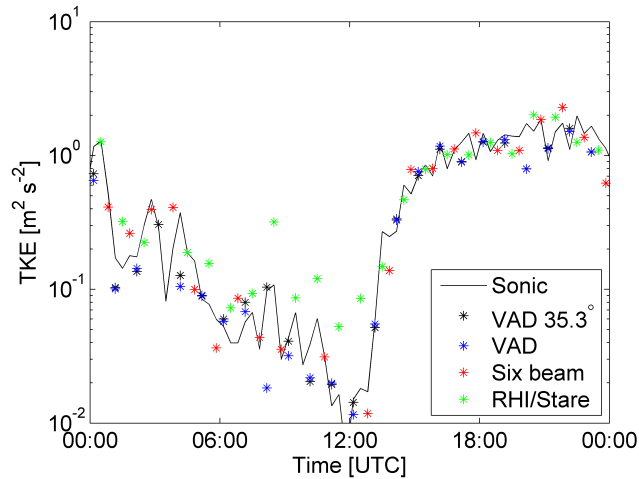
$$TI = \frac{\sqrt{\overline{u'^2}}}{U}, \quad (7)$$

where  $U$  is the mean horizontal wind speed, is most often used as it is a measure of the variability of the inflow into the turbine and affects the design requirements (International Electrotechnical Commission, 2005). In boundary-layer meteorology and air quality, TKE calculated as

$$TKE = \frac{1}{2}(\overline{u'^2} + \overline{v'^2} + \overline{w'^2}) \quad (8)$$

is often used as a measure of the turbulent mixing in the atmosphere (Stull, 1988; Arya, 1999). Additionally, the covariance terms  $\overline{u'w'}$  and  $\overline{v'w'}$  are the momentum flux and are necessary to test and validate models of atmospheric flow. Since these measures of turbulence are most commonly used, they are the focus of this section. A complete statistical comparison of each measured variable is provided in Appendix A.

When interpreting the intercomparisons presented here, it is necessary to consider the statistical uncertainty and representativeness of the measurements themselves. However, quantifying the sampling error of a turbulence measurement is not trivial. Numerous studies have been entirely focused on determining sampling errors of turbulence and flux measurements from time series analysis alone (e.g., Lenschow et al., 1994; Vickers and Mahrt, 1997; Mahrt, 1998; Finkelstein and Sims, 2001). Typically for time series analysis, the magnitude of these sampling errors largely depends on the record length with respect to the integral time scale of the measured variable. Sampling errors are reduced for a stationary time series when a longer record used for the calculation of the turbulence quantity. Still, the record length needs to be short enough to assume stationarity of turbulence fields. Thus it is difficult to identify a static record length that is appropriate for all conditions. During strongly convective conditions with weak winds, using a time series of 1–2 hours may be appropriate (e.g., Turner et al., 2014; Behrendt et al., 2015). Conversely, turbulence statistics may be more aptly calculated using a record length of 5–10 min during stable conditions (e.g., Sun et al., 2012; Mahrt et al., 2015). The scanning strategies evaluated here use a combination of spatio-temporal sampling to measure turbulent quantities. To date, no method has been developed to estimate sampling errors associated with these techniques. As such, and due to the intricate analysis necessary to determine the appropriate sampling error magnitude,



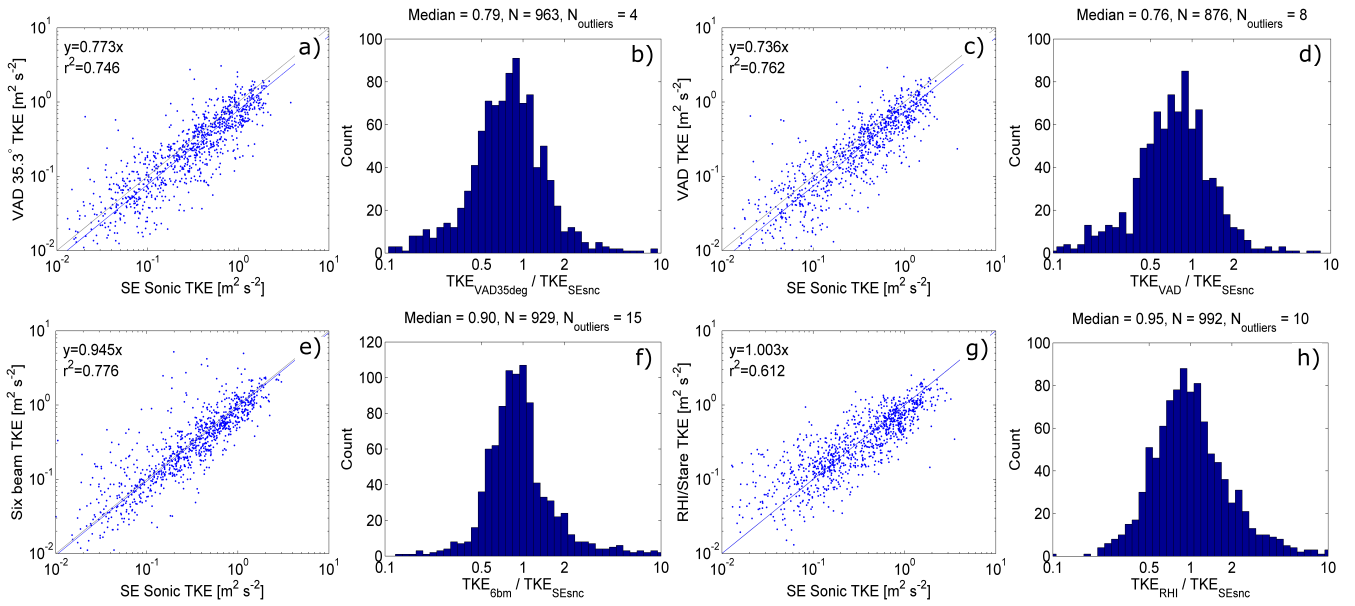
**Figure 4.** Sample time series of measured TKE on 30 May, 2015 at 200 m.

sampling errors are not quantified for any of the measurements presented. Still, it is necessary to interpret the results shown with the understanding that all of the presented measurements have statistical uncertainty of an unknown magnitude.

#### 4.1 Turbulence Kinetic Energy

For the six-beam, VAD with multiple  $\phi$ , and RHI/vertical stare techniques, TKE was directly computed as the sum of the measured velocity variances using Eq. 8. As discussed within Sect. 2.1, TKE can be directly computed from a  $35.3^\circ$   $\phi$  scan without measuring  $\overline{u'^2}$ ,  $\overline{v'^2}$ , or  $\overline{w'^2}$  directly. From here onward, measured quantities from two PPI scans at different  $\phi$  are referred to as ‘VAD’ measurements, and those from one PPI scan at  $35.3^\circ$  are ‘VAD  $35.3^\circ$ ’ measurements.

A sample 24-hr time series of TKE is provided in Fig. 4 to demonstrate the ability of the different methods to capture temporal changes. The diurnal pattern of TKE decreasing overnight between 03:00–12:00 UTC is visible in measurements from the lidar and sonic anemometer. Although all the lidar techniques capture the decrease in TKE in the evening and early night hours (00:00–06:00 UTC), the TKE measurement from the RHI and vertical stare is systematically overestimated later in the night when  $\text{TKE} < 0.1 \text{ m}^2 \text{ s}^{-2}$ . The lidar measured TKE from the VAD, VAD  $35.3^\circ$ , and six-beam techniques capture the trends in TKE well. When TKE measured by the sonic anemometer is small ( $< 0.1 \text{ m}^2 \text{ s}^{-2}$ ) such as at 12:00 UTC, measured TKE by the six-beam technique can be negative. Of all the TKE measurements during the experiment, 0.8% of the six-beam TKE values are negative. ~~This Since this result is unphysical, and is thought to be primarily due to sampling errors (see Lenschow et al., 1994) within each beam variance being propagated through the matrix inversion in Eq. 3, resulting in an underestimate of the variances. These sampling errors are difficult to quantify with the six-beam technique, since the time series is not continuous and the measurements are not equally spaced in time. Since negative TKE estimates are non-physical, they as discussed in Sect. 2.2, these values~~ have been removed. The other methods analyzed did not yield negative TKE values.



**Figure 5.** Scatter plots (a, c, e, g) and histograms (b, d, f, h) showing the relationship between the TKE measured by the lidar and southeast sonic anemometers at all heights. In a, c, e, g, the blue line is the best fit line given by the equation in the upper left and the black line indicates a 1-1 relationship. Histograms (b, d, f, h) show the ratio of the lidar measured TKE to the sonic measured TKE with the median ratio, number of points  $N$ , and number of outliers  $N_{\text{outliers}}$  that are more than 1 order of magnitude apart. Lidar measurements are from TKE at  $\phi = 35.3^\circ$  (a, b), TKE from two  $\phi$  (c, d), the six-beam technique (e, f), and the RHI and stare combination (g, h).

For a quantitative analysis of the TKE measurements, TKE values from the Doppler lidar and the southeast sonic anemometers are summarized in Fig. 5. These results are from measurements between 100–300 m for all the scanning strategies. No TKE measurements are available below 100 m since the first lidar range gate is 100 m so that no  $\overline{w'^2}$  values are available below this height from the vertical stares. Comparisons with the northwest anemometers are similar and provide little additional information, thus are not shown. The comparisons shown here (and throughout the manuscript) are on a logarithmic scale since values can range several orders of magnitude, generally from 0.01–10  $\text{m}^2 \text{s}^{-2}$  for TKE. If the analysis were conducted on a linear scale, large values of TKE would dominate the best fit lines and statistical comparison and the small values would not be important. In logarithmic space, the ability to differentiate values of  $\sim 0.01$  from  $\sim 0.1$  is equally as important as differentiating  $\sim 1$  from  $\sim 10$ , and values across all different orders of magnitude are weighted equally in determining the trend line. Within each of the scatter plots in Fig. 5, the best fit lines were determined fitting

$$\log_{10}(y) = \log_{10}(x) + b, \quad (9)$$

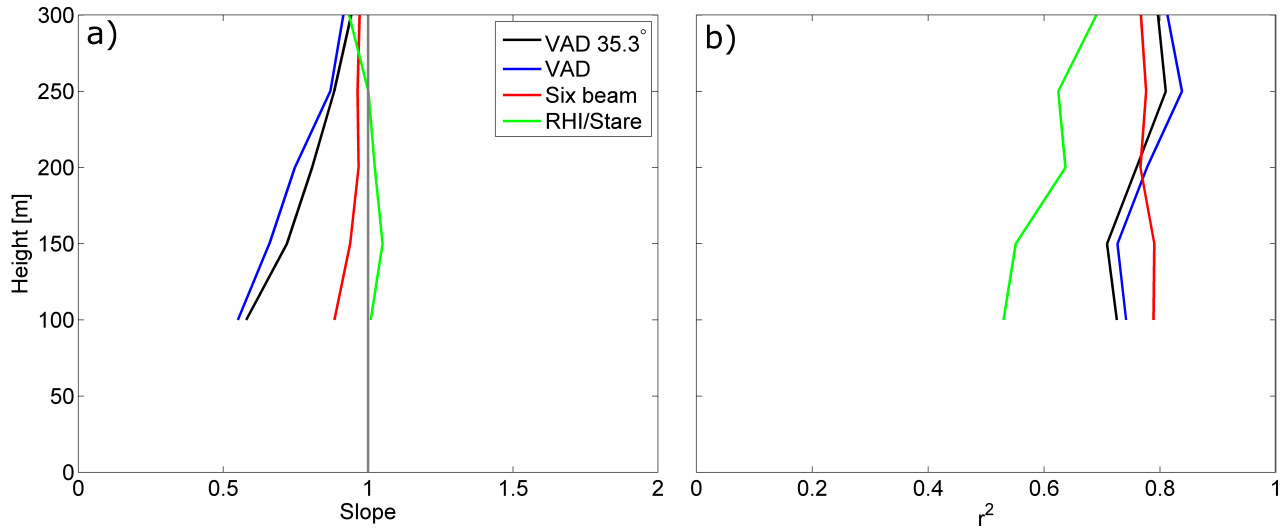
where  $x$  and  $y$  are data points on their respective axes in linear units and  $b$  is a constant. Transforming the equation back into linear space for ease of interpretation, the equation shown in the upper left of Fig. 5a, c, e, g is  $y = bx$ .  $y = 10^b x$ , where  $10^b$

is the slope of the regression. The slope is related to the bias of the measurement. For example, a slope less than one indicates that measured quantities are systematically smaller than the reference measurement over all scales.

Each of the techniques evaluated ~~herein~~ generally shows skill in measuring TKE, as indicated by  $r^2$  values greater than 0.6 in Fig. 5. Considering that the sonic anemometer and lidar measurements represent different spatial areas, which vary according to the scanning technique, and that each are subject to sampling error, the authors consider the correlation between the lidar and sonic TKE to be good. ~~Quantifying the sampling error. In the absence of sampling errors that~~ would allow a more statistical determination of ~~it whether~~ measurements are in agreement within their respective uncertainties. ~~However, no techniques that the authors are aware of can be used to estimate the sampling error from any of the scan types except the vertical stares (see Lenschow et al., 1994). Due to these limitations and for consistency, sampling errors are not estimated for any of the scans. Instead, the,~~ the relative correlations between sonic and lidar measurements, and their differences for various scan strategies, are used to understand biases and accuracy of each technique.

The six-beam technique demonstrates the best ability to measure TKE overall, as evident by the largest  $r^2$  and slope of 0.945, close to unity. Additionally, the histogram of the ratio of TKE measurements in Fig. 5f shows a distinct peak around 1 with reduced spread compared to Fig. 5b, d, h; 83.5% of six-beam TKE values are within a factor of 2 of the sonic measurement, the largest proportion of all the techniques analyzed. However, the six-beam technique also produces the largest number of TKE outliers, defined as being more than 1 order of magnitude different from the sonic observed TKE. Approximately half of these outliers are negative TKE values, which were removed as discussed earlier. The other outliers are when TKE is grossly overestimated, as visible in Fig. 5e. Upon manual inspection of these high outliers, many are due to contamination of range folded echoes. When range folded echoes appear intermittently, the  $v_r$  time series within each beam position changes erratically, resulting in an anomalously large variance, and spuriously increasing the observed TKE. The discontinuity-based algorithm used to detect range folded echoes largely relies on contextual information from proximate beams in time and space (Bonin and Brewer, 2017) not available from the six-beam technique. These anomalous echoes can typically be detected and removed in PPI, RHI, and stare scans, but these range folded returns persist through the quality-control process of the six-beam measurements and degrade the accuracy of the calculated variances .

As discussed in Sect. 2.1, the VAD technique can be used to measure TKE from either one scan at  $35.3^\circ$  or two scans at different  $\phi$ , herein  $35.3^\circ$  and  $50.8^\circ$ . Since the  $35.3^\circ$  scan is used in both approaches, the results from both methods are not independent of each other. This can be seen by the similar results from both approaches in Fig. 5a-d. Although the TKE from both techniques is highly correlated with sonic TKE, the VAD and VAD  $35.3^\circ$  measured TKE is systematically biased too low. The low bias is more pronounced for the two- $\phi$  technique, although the scatter is reduced slightly as evidenced by the larger  $r^2$  in Fig. 5c. This reduced scatter is attributed to smaller sampling errors, since twice the amount of data go into the measurement. The overall low biases may be due to the inability to capture all the scales of turbulence; the largest eddies may not be fully captured and resolved within the scanning circle. This effect would be more pronounced for higher elevation PPIs, such as at  $50.8^\circ$ , and explain the more significant low bias in TKE for the two PPIs used herein. If a lower  $\phi$  were used (i.e., at  $25^\circ$ ), the low bias may not be as pronounced for the two  $\phi$  VAD TKE. Unfortunately, no data are available from this experiment



**Figure 6.** The slope (a) and  $r^2$  (b) of the best fit line on a logarithmic scale (similar to those in Fig. 5) relating TKE measurements from the sonic anemometer and lidar for each measurement height.

to validate this hypothesis. Despite these biases, 73.3% (35.3° VAD) and 71.7% (two VADs) of the TKE measurements are within a factor of 2 of the sonic TKE.

Using the ~~RHI-RHIs~~ and vertical stares to measure TKE results in the largest scatter. Still, 74.7% of the lidar measured TKE are within a factor of 2 of the sonic measurements, indicating that the technique is still accurate. The six-beam and VAD techniques show similar scatter for all ranges of TKE, and the RHI/stare technique typically overestimates TKE when its value is small (i.e.,  $< 0.1 \text{ m}^2 \text{ s}^{-2}$ ), as apparent in Fig. 5g. The cause of the overestimate during weakly turbulent conditions is unclear, but it may be due to spatial variability of the flow that the sonic cannot detect ~~or the inability to remove instrument noise effects. These.~~ Mean or other non-turbulent variability along the horizontally oriented vertical bins, such as if the bin is sloped with respect to the underlying topography, can be a significant contribution to the overestimate of the calculated variances, especially evident in weakly turbulent conditions. Additionally, random errors are quantified and removed in the VAD and six-beam techniques as detailed in Sect. 2, but no established technique exists to remove these errors from RHI measurements, which may ~~be addressed in a future study~~ lead to this high bias in TKE as measured with the RHIs and vertical stares. Within Fig. 5h, this high bias under weakly turbulent conditions manifests itself as right-skewed distribution.

The results shown in Fig. 5 are for all measurement heights combined. The analysis can be further refined by comparing lidar TKE measurements at each sonic anemometer height separately. Figure 6 summarizes this analysis by showing the slope and  $r^2$  of the best fit line at each height. The accuracy of the six-beam technique is the most consistent at all measurement heights, as the slope and  $r^2$  are nearly constant with height. The value of  $r^2$  remains around 0.75 at every height, whereas the slope increases a small amount with height. This change in slope indicates that TKE is less-underestimated above 200 m than it is closer to the surface.

The bias of TKE measurements using the RHI and vertical stare method is independent of height and small (Fig. 5a), as the slope is generally around 1. The TKE measurement becomes more accurate with height, indicated by the increase of  $r^2$  with height in Fig. 5b. The cause for the increase in accuracy with height is unclear, but may be due to the mean flow becoming more homogeneous aloft. The low bias of VAD and VAD 35.3° TKE observations becomes less significant with height, as shown in Fig. 5a. Coincidentally, the VAD and VAD 35.3°  $r^2$  values increase with height, representing less scatter and more accurate values at higher altitudes.

To examine the decrease in low bias and increase in accuracy of VAD TKE measurements with height, the VAD circle diameter and is compared with the typical largest eddy size, the integral length scale of  $l$ , are compared. First, the integral time scale  $t_{int}$  is calculated from a linearly detrended 20-min timeseries of  $u$  from the sonic anemometer as

$$10 \quad t_{int} = \frac{1}{u'^2} \int_0^{\infty} \sum_{\tau=0}^{\tau(A=0)} A(\tau) d\tau, \quad (10)$$

where  $A(\tau)$  is the autocovariance of  $u$ , which is a function of the time lag  $\tau$ . Since the time series is discrete,  $d\tau = \Delta\tau$  is the sampling interval (0.05 s here, since the sonic data rate is 20 Hz). The median and various percentile values of  $l$ , computed as  $l = Ut_{int}$ , are shown as a function of height and with reference to PPI scan diameters in Fig. 7. Generally, individual 35.3° and 50.8° PPI scans do not fully sample the largest turbulent eddies close the ground, since the scan circle diameter is often less than  $l$ . The largest eddies are better captured by these scans at higher altitudes, especially for the 35.3° PPI scan. At 300-m, the integral scale is less than the 35.3° scan diameter over 90% of the time.

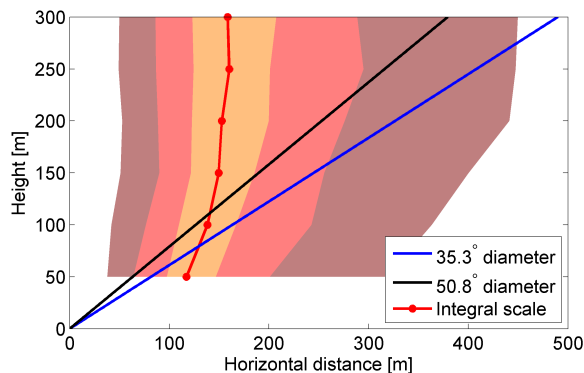
The results shown in Fig. 7 explain why VAD and VAD 35.3° TKE measurements become more accurate and less biased with height, as the largest turbulence scales are better more completely captured. These effects are not important to the RHI method, since the spatial extent of the average is typically several km, much larger than the typical eddy size. Since the vertical stare and six-beam techniques use time series analysis, the largest scales of turbulence are observed if the time window length exceeds the integral time scale, which is often  $\sim 10 - 100 \sim 10 - 100$  s during daytime (Lenschow et al., 2000; Bonin et al., 2016) and less than 10 s at night (Pichugina et al., 2008b).

## 4.2 Turbulence Intensity

Similar to the analysis of TKE presented in Sect. 4.1, measurements of TI measured from the six-beam, VAD (using 2  $\phi$  angles), and RHI techniques are compared and validated here. Since  $TI \rightarrow \infty$  as  $U \rightarrow 0$  following Eq. 7, TI is only calculated when  $U > 1 \text{ m s}^{-1}$ . A sample time series of TI is shown in Fig. 8. The diurnal trend in TI is clearly visible in the sonic measurements, as TI is generally low (3–10%) at night until 12:00 UTC and TI is larger (20–70%) during the day. During the morning hours (i.e., 12:00–18:00 UTC),  $U$  was less than  $2.5 \text{ m s}^{-1}$ , causing TI to become large. Despite some scatter, TI measurements from the Doppler lidar show a similar trend with smaller TI values at night and larger ones during the day.

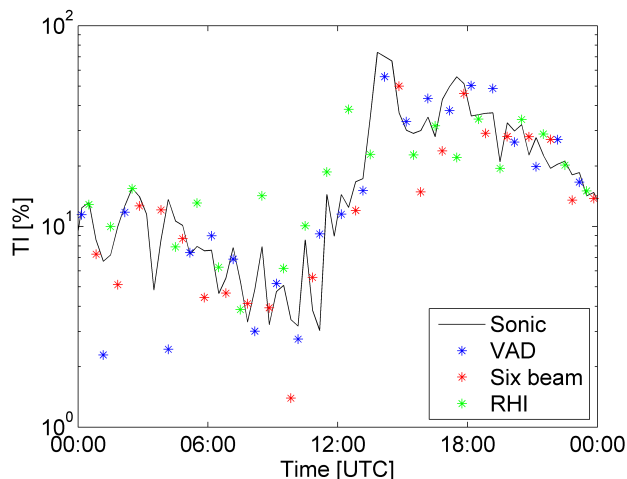
30 Nonphysical Non-physical negative  $\overline{u'^2}$  values due to computational artifacts as described previously by Newman et al. (2016b), caused by the same effect as the negative TKE measurements discussed in Sect. 4.1, have been removed from the



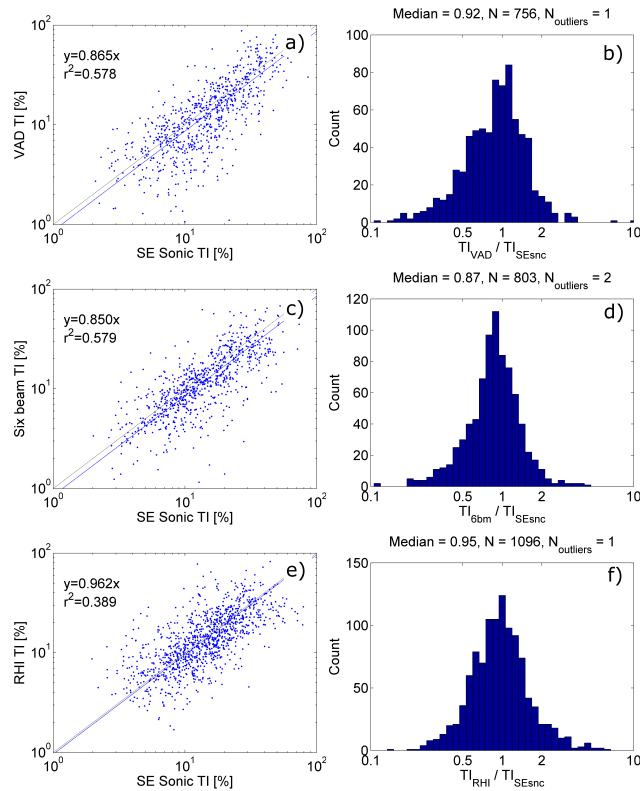


**Figure 7.** The diameter of the scanning circle of the two PPI scans is shown in comparison to  $l$ . The red line denotes the median  $l$ , while the progressively darker contours represent the 40–60%, 25–75%, and 10–90% percentile intervals of  $l$  over the entire 17-day experimental period.

analysis. Measurements of TI at all heights over the entire experiment are summarized in Fig. 9. For each of the three techniques analyzed, the  $r^2$  for TI is  $\approx 0.2$  lower than it is for TKE. This indicates the combined velocity variance components in TKE are more accurately measured than individual velocity variances separately (see also Table 2 for  $\overline{u'^2}$ ,  $\overline{v'^2}$ , and  $\overline{w'^2}$  comparison statistics). Still, the VAD, six-beam, and RHI techniques each show skill ([i.e., show correlation](#)) in measuring TI. The VAD and six-beam techniques perform comparably, having a similar  $r^2$  and slope indicated a low bias. Sathe et al. (2015b) also show that the six-beam technique tends to underestimate  $\overline{u'^2}$  [by a similar amount](#). The RHI TI measurements show more scatter than the other two methods, given the lower  $r^2$ , but showed little bias.



**Figure 8.** Sample time series of measured TI on 30 May, 2015 at 200 m.



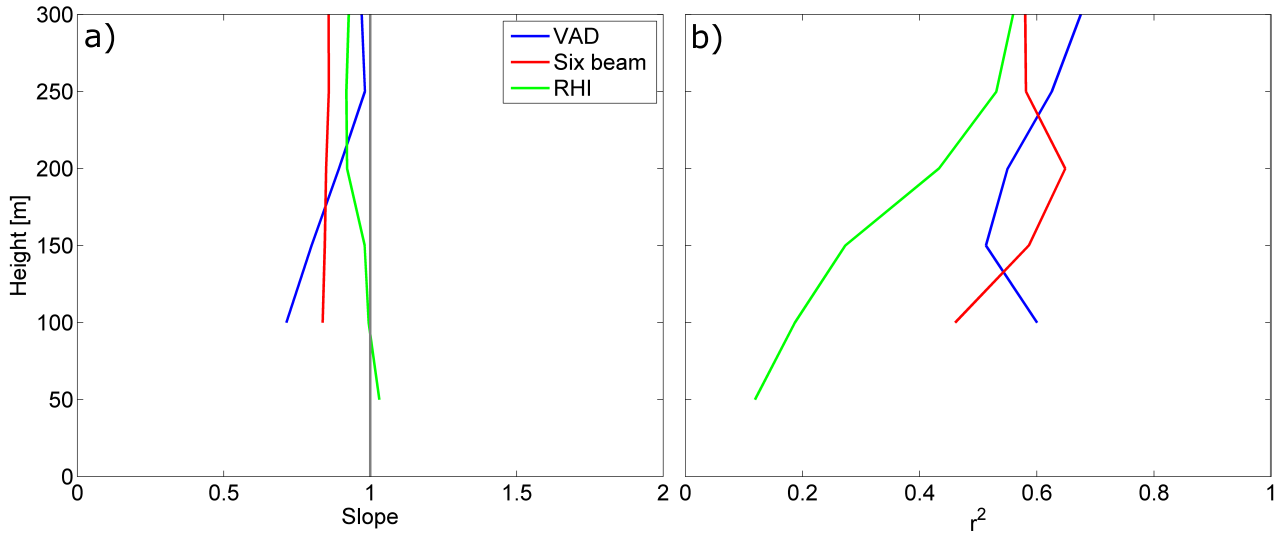
**Figure 9.** Same as Fig. 5 for TI instead of TKE. Lidar measurements are from TKE from two  $\phi$  (a, b), the six-beam technique (c, d), and the RHI scans (e, f).

The slope and  $r^2$  of the best fit line as a function of height is shown in Fig. 10. Similarly to TKE, VAD measurements of TI are biased low near the ground as indicated by the slope of  $\approx 0.7$  at 100 m. By 250 m, the bias becomes small as most of the turbulence scales are resolved, as discussed in Sect. 4.1. The accuracy of the VAD TI measurement does not change significantly with height, as  $r^2$  does not consistently depend on height. The six-beam TI measurement is biased consistently low regardless of height, as indicated by the slope of  $\approx 0.83$  at all heights, and the scatter of the measurements does not have a consistent trend with height. The slope of RHI TI tends to be larger near the ground and slowly decrease with height, as evidenced in Fig. 10a. Coincidentally, the scatter associated with these measurements decreases significantly with height, as the  $r^2$  increases from 0.12 to 0.56.

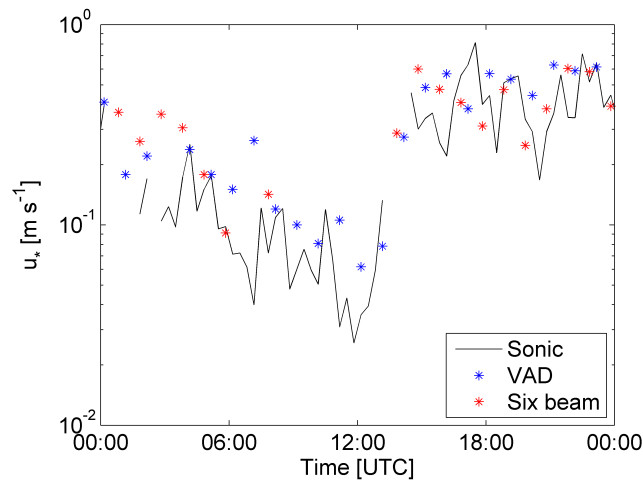
### 4.3 Shear Stress Velocity

10 The momentum flux terms  $\overline{u'w'}$  and  $\overline{v'w'}$  can be combined through the calculation of a stress velocity scale  $u_*$  by

$$u_* = \left[ \overline{u'w'^2} + \overline{v'w'^2} \right]^{1/4}. \quad (11)$$

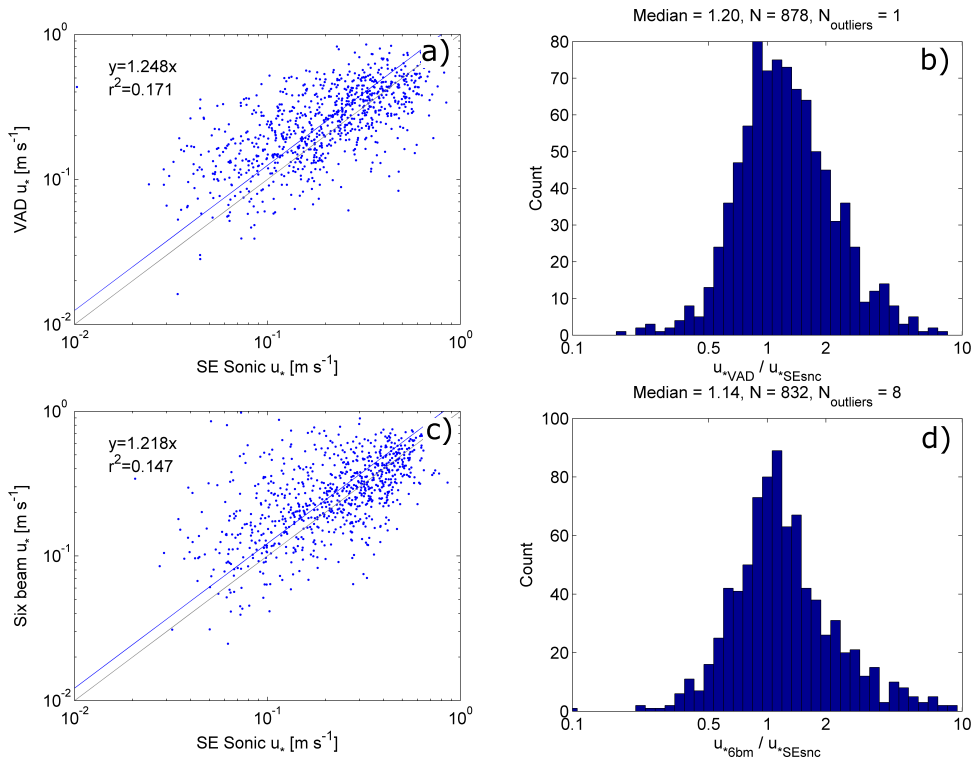


**Figure 10.** Same as Fig. 6 for TI instead of TKE.



**Figure 11.** Sample time series of measured  $u_*$  on 30 May, 2015 at 200 m.

Of the techniques analyzed, only the six-beam and VAD methods have a theoretical basis for measuring the covariances  $\overline{u'w'}$  and  $\overline{v'w'}$  necessary to compute  $u_*$ . Each PPI scan at any  $\phi$  can independently provide a measurement of the covariances, so the  $u_*$  values shown here are taken as the average of all PPI ~~scan-scans~~ scans at both  $35.3^\circ$  and  $50.8^\circ$ . An example of a 24-hr time series of  $u_*$  is shown in Fig. 11. The sonic data are not shown for 00:00–03:00 UTC and 14:00 UTC, since the  $u_*$  measurements on opposing booms were more than a factor of two different from each other, even though neither sonic was waked. Thus, neither is taken as a baseline measurement. For this sample period, the lidar and sonic data show a similar trend, values of  $u_*$



**Figure 12.** Same as Fig. 5 for  $u_*$  instead of TKE. Lidar measurements are from TKE from two  $\phi$  (a, b) and six-beam technique (c, d).

decreasing for 00:00–12:00 UTC, rapidly increasing for 12:00–15:00 UTC, and remaining nearly constant after 15:00 UTC. These trends are a result of  $u_*$  steadily decreasing overnight, increasing in the morning hours, and remaining steady over the day.

The [comparison-comparisons](#) between sonic and lidar  $u_*$  measurements are summarized in Fig. 12. Both the VAD and six-beam techniques generally overestimate  $u_*$ , as shown in both the histograms and scatter plots. During time periods when the [sonic-estimated-sonic-estimated](#)  $u_*$  is small (i.e.,  $< 0.1 \text{ m s}^{-1}$ ), the lidar techniques predominately overestimate  $u_*$  as indicated by the large number of data points above the one-to-one line in Fig. 12a, c. The small  $r^2$  for the best fit lines indicates that there is substantial scatter in the comparison of the lidar and sonic measurements. Thus, the six-beam and VAD methods show little skill in being able to accurately measure  $u_*$  and the covariances, as shown in Table 2. These results do not significantly change with height (not shown), as the  $r^2$  remains small for both methods at all measurement heights between 50–300 m. The accuracy of covariance and  $u_*$  measurements from the VAD and six-beam methods has not been evaluated in the past, but here the measurements are found to exhibit large error. Over simpler topography, Berg et al. (2013) present results from a DBS technique that produced more accurate measurements of  $\overline{u'w'}$  and  $\overline{v'w'}$ .

## 5 Discussion

From the results shown in Sect. 4, it is clear that TKE can be measured by each of the three techniques analyzed. However, measurements of each individual term of the Reynolds stress tensor are more difficult to accurately measure. The velocity covariances are particularly difficult to quantify, as the six-beam and VAD techniques show little skill in their measurement.

5 It is thought that the poor comparison for the covariance terms is due to the fact that the sampling error for the measurement exceeds the covariance typical dynamic range. Based on sonic anemometer observations, 80% of  $|\overline{u'v'}|$ ,  $|\overline{u'w'}|$ , and  $|\overline{v'w'}|$  were  $< 0.1 \text{ m}^2 \text{ s}^{-2}$ . Also, covariance terms having small correlations take much longer to converge to a stable value (Lenschow et al., 1994). [Since the individual covariance terms do not correlate 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](#)  
10 [\( \$r^2 = 0.14 - 0.17\$ \).](#)

### 5.1 Strengths and limitations of each strategy

Each of the scanning strategies evaluated herein has its own strengths and limitations. One of the biggest limitations for all of the techniques except vertical stares is that turbulence is assumed to be homogeneous over the area of each scan. Thus, these techniques do not always work well in complex terrain or differential land use where turbulence can significantly vary spatially [\(e.g., Maurer et al., 2016\)](#). For the VAD and RHI techniques in particular, spatial variations in the mean wind due to local drainage flows [\(e.g., Banta et al., 1997; Choukulkar et al., 2012\)](#) [\(e.g., Banta et al., 1997; Choukulkar et al., 2012; Adler and Kalthoff, 2012\)](#) result in large deviations from the spatially averaged mean wind. Since these methods are unable to differentiate turbulent deviations from mean deviations, turbulence is overestimated. In these situations, it may be possible to use arc segments from the PPI scans to compute TKE (Wang et al., 2015) over different radials where the mean flow is homogeneous. With the current technology, multi-Doppler measurements [\(i.e., Mann et al., 2009; Fuertes et al., 2014; Newman et al., 2016a\)](#) [\(e.g., Mann et al., 2009; Fuertes et al., 2016b\)](#) are the best able to quantify turbulence at specific locations in complex terrain.

The spatial height resolution for the PPI, six-beam, vertical stare, and RHI scans largely depends on scan geometry. Direct measurements of  $\overline{w'^2}$  from vertical stares can only be taken starting at the height of the lowest range gate and the spatial resolution is limited to the range gate size. Since the six-beam technique presented here and in previous studies (e.g., Sathe et al., 2015b; Newman et al., 2016b) uses a vertical beam, the spatial resolution is limited by that beam the same as vertical stares. Future studies may try removing the vertical beam and instead use six-beams all at  $\phi = 45^\circ$ , or another  $\phi$ , to [take-make](#) measurements at a lower altitude. The vertical resolution of a PPI scan is dictated by its  $\phi$ : a larger  $\phi$  results in a higher minimum measurement height, reduced vertical resolution, greater height coverage, and reduced horizontal scan footprint compared to typical eddy size. The residuals in the PPI scans are more sensitive to  $\overline{w'^2}$  for a larger  $\phi$ , and are more sensitive to  $\overline{u'^2}$  and  $\overline{v'^2}$  for a smaller  $\phi$ . The height resolution of an RHI scan is truly customizable, as  $\overline{u'^2}$  and  $\overline{v'^2}$  are computed by user-defined height bins. Since RHI scans typically start or end at the horizon,  $\overline{u'^2}$  and  $\overline{v'^2}$  can be calculated within a few meters of the surface. On the other hand, this technique is especially susceptible to non-turbulent horizontal variability along the scan due to complex terrain and other effects, especially since small  $\phi$  that cover large distances horizontally are used.

Although a 20-min averaging interval is used here for comparison, measurements for several of the techniques could be made much quicker. Turbulence statistics from vertical stares and the six-beam technique are computed through typical time series analysis; thus, the time series needs to be long enough to ensure that the largest turbulent eddies pass through the resolution volume to be captured (see Lenschow et al., 1994), yet short enough that the flow can be considered stationary (Banta et al., 2013a). Since the VAD and RHI methods compute turbulence through quantifying spatial variability, they are not subject to these same sampling error limitations when all scales of turbulence are captured in the scanning volume. With a data rate of 2 Hz, each PPI scan with 240 beams takes 2 min to complete; thus TKE can be measured in 2 min with a scan at  $\phi = 35.3^\circ$  and each velocity variance and covariance can be measured from two scans, taking 4 min. Since each RHI scan can be conducted in  $< 1$  min,  $\overline{u'^2}$  and  $\overline{v'^2}$  can be measured in  $\approx 2$  min.

The methods presented here measure velocity variances, but none currently are able to distinguish atmospheric turbulence from submeso motions, including waves. Since the value of TKE is calculated from  $\overline{u'^2}$ ,  $\overline{v'^2}$ , and  $\overline{w'^2}$ , the observed TKE may be a mixture of turbulent and submeso variances and not always a measure of pure atmospheric turbulence. Considering these submeso motions have been predominantly documented within the nocturnal stable PBL when turbulence is typically weak (Mahrt, 2014), the value of TKE defined as a measure of turbulent motion may be overestimated when waves are present. Numerically differentiating between non-turbulent and turbulent motions is difficult (Stewart, 1969) and is best done through multiresolution decomposition or wavelet analysis (e.g., Cuxart et al., 2002; Vickers and Mahrt, 2003; Viana et al., 2010). This requires a high-resolution ( $> 1$  Hz) time series. Thus, out of the methods analyzed, only the vertical stare has the data necessary to separate turbulent from non-turbulent motions using established techniques.

## 5.2 Future directions for improving turbulence estimates

One of the main limitations of the six-beam technique using current commercially available scanning Doppler lidars is the return time between samples at the same beam position ( $\approx 27$  s). When the six-beam technique was performed, the scanner spent 78% of the time slewing from one beam to the next. Thus a 2-axis hemispheric scanner is not the best option for running the six-beam technique. A wedge scanner that can quickly rotate between beam positions is more appropriate, as the time between beams could be minimized. This would yield a higher temporal resolution time series for each beam, enabling a better method of noise removal through a structure function fit (Lenschow et al., 2000) and possibly differentiating between turbulent and non-turbulent variances through multiresolution decomposition or wavelet analysis.

The RHI technique is best suited for measuring  $\overline{u'^2}$  and  $\overline{v'^2}$  near the surface ( $< 100$  m), since the measurements need to be made at a low angle. However, there is currently no method to remove random errors from RHI measurements. Thus, it may be better in the future to simply perform shallow horizontal stares where  $\phi < 20^\circ$  to measure the horizontal variances. Removing noise and correcting for volume averaging effects would be straightforward (Bonin et al., 2016), and it also may be able to distinguish turbulence from non-turbulent motions.

## 6 Conclusions

The XPIA field experiment was conducted at the Boulder Atmospheric Observatory in the spring of 2015. For 17 days at the end of the experiment, a Leosphere WINDCUBE 200S<sup>®</sup> continuously alternated between a PPI, RHI, vertical stare, and six-beam scanning strategy. Measurements from each scan type were used to calculate components of the Reynolds stress tensor and other measures of turbulence. These Doppler lidar turbulence measurements were compared to those from sonic anemometers on a 300-m tower located 540 m from the lidar to evaluate the accuracy of each technique.

Overall, TKE and velocity variances (i.e.,  $\overline{u'^2}$ ,  $\overline{v'^2}$ ,  $\overline{w'^2}$ ) were more accurately measured by the six-beam technique [than the other methods](#). Six-beam measurements showed the best agreement with the sonic-anemometer data across all ranges of turbulence magnitude ( $r^2 \approx 0.78$ ). Additionally, the error and bias of the six-beam turbulence measurements did not significantly change with height. On the other hand, the VAD measurements of TKE and velocity variances tended to become more accurate with height. VAD-measured turbulence tended to be biased low near the surface, and this bias decreased with height. This bias is attributed to the inability of the PPI scan to resolve all scales of turbulence near the surface, since the largest eddies extend beyond the scanning circle. The scanning volume geometrically becomes larger with height; thus, the PPI is better able to resolve all scales of turbulence and make more accurate measurements of turbulent quantities farther from the surface. Although the ~~sonic anemometer observations agreed most poorly with~~ RHI-measured TKE and TI [agreed most poorly with sonic anemometer observations](#), it showed little bias (slope of linear regression for TKE was 1.003) and still showed considerable skill in measuring turbulence. The inability to quantify and remove random errors from the RHI measurements led to an overestimate under time periods when turbulence was weak ( $\text{TKE} < 0.1 \text{ m}^2 \text{ s}^{-2}$ ). The methods evaluated herein showed little skill in measuring  $u_*$  and velocity covariances ( $r^2 = 0.15 - 0.17$ ).

When selecting a scanning strategy in future experiment, one needs to consider the desired turbulence measurements. While the RHI technique may be the least accurate of the three evaluated, it is the only method that can obtain measurements just above the surface. If a rapid update time is desired (i.e.,  $< 5 \text{ min}$ ), the VAD technique may best address these needs. Vertical stares and the six-beam technique use time series analysis to quantify turbulence. If the temporal resolution is sufficiently high, established techniques may be used to partition turbulent and non-turbulent variance, which no method currently exists for the RHI and VAD data.

### Appendix A: Comparison of all measured turbulent quantities

A complete statistical comparison of all measures of turbulence is provided for reference in Table 2. For brevity, only a selected portion of these results were closely evaluated within the main body of the manuscript. While the results in Table 2 summarize measurements at all heights, the accuracy and bias of velocity variances and covariances as a function of height are similar to those presented in Sect. 4. The results have been summarized through comparisons of values on both a logarithmic scale (except for the covariances which can be negative) and a linear scale. [Values of the root-mean-square error \(RMSE\) are also provided here, and the units of RMSE on a logarithmic scale are orders of magnitude difference.](#)

**Table 2.** Statistical comparison between sonic anemometer and lidar observations of all measured turbulence variables. Slope,  $r^2$ , and RMSE values are computed ~~seperately~~ separately for regression analysis on a logarithmic and linear scale. The column ‘% Non-physical’ indicates the % of measurements that are negative or non-real.

Variable	Method	Slope (log)	$r^2$ (log)	RMSE (log, order of magnitude)	Slope (linear)	$r^2$ (linear)	RMSE (linear)	% Non-physical
TKE	VAD ( $2\phi$ )	0.736	0.762	<u>0.269</u>	0.721	0.614	<u>0.245 m<sup>2</sup> s<sup>-2</sup></u>	0
	VAD ( $\phi = 35.3^\circ$ )	0.773	0.746	<u>0.272</u>	0.731	0.559	<u>0.275 m<sup>2</sup> s<sup>-2</sup></u>	0
	Six-beam	0.945	0.776	<u>0.246</u>	0.913	0.562	<u>0.359 m<sup>2</sup> s<sup>-2</sup></u>	0.8
	RHI/Vertical Stare	1.003	0.612	<u>0.272</u>	0.791	0.632	<u>0.252 m<sup>2</sup> s<sup>-2</sup></u>	0
$\overline{u'^2}$	VAD ( $2\phi$ )	0.940	0.524	<u>0.371</u>	0.876	0.335	<u>0.343 m<sup>2</sup> s<sup>-2</sup></u>	7.9
	Six-beam	0.872	0.620	<u>0.327</u>	0.835	0.342	<u>0.465 m<sup>2</sup> s<sup>-2</sup></u>	8.1
	RHI	0.918	0.355	<u>0.328</u>	0.627	0.294	<u>0.273 m<sup>2</sup> s<sup>-2</sup></u>	0
$\overline{v'^2}$	VAD ( $2\phi$ )	0.977	0.620	<u>0.334</u>	0.910	0.393	<u>0.316 m<sup>2</sup> s<sup>-2</sup></u>	4.9
	Six-beam	0.914	0.680	<u>0.319</u>	1.033	0.568	<u>0.315 m<sup>2</sup> s<sup>-2</sup></u>	8.8
	RHI	0.981	0.372	<u>0.339</u>	0.679	0.314	<u>0.281 m<sup>2</sup> s<sup>-2</sup></u>	0
$\overline{w'^2}$	VAD ( $2\phi$ )	0.739	0.542	<u>0.423</u>	0.539	0.268	<u>0.214 m<sup>2</sup> s<sup>-2</sup></u>	35.4
	Six-beam	0.996	0.789	<u>0.303</u>	0.927	0.648	<u>0.210 m<sup>2</sup> s<sup>-2</sup></u>	0
	Vertical Stare	0.971	0.790	<u>0.325</u>	0.979	0.591	<u>0.255 m<sup>2</sup> s<sup>-2</sup></u>	0
$\overline{u'v'}$	VAD				0.150	0.005	<u>0.164 m<sup>2</sup> s<sup>-2</sup></u>	0
	Six-beam				0.006	0.001	<u>0.310 m<sup>2</sup> s<sup>-2</sup></u>	0
$\overline{u'w'}$	VAD				-0.165	0.007	<u>0.115 m<sup>2</sup> s<sup>-2</sup></u>	0
	Six-beam				-0.020	0.005	<u>0.203 m<sup>2</sup> s<sup>-2</sup></u>	0
$\overline{v'w'}$	VAD				-0.010	0.001	<u>0.121 m<sup>2</sup> s<sup>-2</sup></u>	0
	Six-beam				0.001	0.002	<u>0.233 m<sup>2</sup> s<sup>-2</sup></u>	0
TI	VAD ( $2\phi$ )	0.865	0.578	<u>0.239</u>	0.967	0.535	<u>11.6%</u>	7.9
	Six-beam	0.850	0.579 <del>nb</del>	<u>0.206</u>	0.837	0.505	<u>8.2%</u>	8.1
	RHI	0.962	0.389	<u>0.223</u>	0.865	0.248	<u>10.1%</u>	0
$u_*$	VAD	1.248	0.171	<u>0.243</u>	0.992	0.133	<u>0.159 m s<sup>-1</sup></u>	0
	Six-beam	1.218	0.147	<u>0.258</u>	1.061	0.089	<u>0.194 m s<sup>-1</sup></u>	0

*Acknowledgements.* The authors acknowledge financial support for this work was provided by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, the NIST Greenhouse Gas Measurements and Climate Research Program, and NOAA’s Earth System Research Laboratory. The authors are also thankful to numerous individuals and organizations who assisted with the field deployment and processing of the data including Tom Ayers, Bruce Bartram, Duane Hazen, Paul Johnston, Jesse Leach, Katherine McCaffrey, Lefthand Water District, Erie High School, and the St. Vrain School District. We also appreciate NOAA/Earth Systems Research Laboratory/Physical Sciences Division for supporting the instrumentation at the BAO. We express appreciation to the National Science Foundation for supporting the CALB deployments of the tower instrumentation ([https://www.eol.ucar.edu/field\\_projects/cabl](https://www.eol.ucar.edu/field_projects/cabl)).



## References

- Adler, B. and Kalthoff, N.: Multi-scale Transport Processes Observed in the Boundary Layer over a Mountainous Island, *Bound.-Lay. Meteorol.*, 153, 515–537, doi:10.1007/s10546-014-9957-8, 2014.
- Arya, S. P.: *Air Pollution Meteorology and Dispersion*, Oxford University Press, 310 pp, 1999.
- 5 Banta, R. M., Shepson, P. B., Bottenheim, J. W., et al.: Nocturnal cleansing flows in a tributary valley, *Atmos. Environ.*, 31, 2147–2162, 1997.
- Banta, R. M., Newsom, R. K., Lundquist, J. K., Pichugina, Y. L., Coulter, R. L., and Mahrt, L.: Nocturnal low-level jet characteristics over Kansas during CASES-99, *Boundary-Layer Meteorol.*, 105, 221–252, 2002.
- Banta, R. M., Pichugina, Y. L., and Brewer, W. A.: Turbulent velocity-variance profiles in the stable boundary layer generated by a nocturnal  
10 low-level jet, *J. Atmos. Sci.*, 63, 2700–2719, 2006.
- Banta, R. M., Pichugina, Y. L., Kelley, N. D., Hardesty, R. M., and Brewer, W. A.: Wind energy meteorology: Insight into wind properties in the turbine-rotor layer of the atmosphere from High-Resolution Doppler Lidar, *Bull. Am. Meteorol. Soc.*, 94, 883–902, doi:10.1175/BAMS-D-11-00057.1, 2013a.
- Banta, R. M., Shun, C. M., Law, D. C., Brown, W., Reinking, R. F., Hardesty, R. M., Senff, C. J., Brewer, W. A., Post, M. J., and Darby, L. S.:  
15 *Observational Techniques : Sampling the Mountain Atmosphere*, in: *Mountain Weather Research and Forecasting*, edited by Chow, F. K., DeWekker, S. F. J., and Snyder, B. J., chap. Observatio, pp. 409–530, Springer Atmospheric Sciences, doi:10.1007/978-94-007-4098-3, 2013b.
- Banta, R. M., Pichugina, Y. L., Brewer, W. A., Lundquist, J. K., Kelley, N. D., Sandberg, S. P., Alvarez II, R. J., Hardesty, R. M., and Weickmann, A. M.: 3D Volumetric Analysis of Wind Turbine Wake Properties in the Atmosphere Using High-Resolution Doppler Lidar,  
20 *J. Atmos. Ocean. Tech.*, 32, 904–914, doi:10.1175/JTECH-D-14-00078.1, 2015.
- Barlow, J. F., Dunbar, T. M., Nemitz, E. G., Wood, C. R., Gallagher, M. W., Davies, F., O’Connor, E., and Harrison, R. M.: Boundary layer dynamics over London, UK, as observed using Doppler lidar during REPARTEE-II, *Atmos. Chem. Phys.*, 11, 2111–2125, 2011.
- Behrendt, A., Wulfmeyer, V., Hammann, E., Muppa, S. K., and Pal, S.: Profiles of second- to fourth-order moments of turbulent temperature fluctuations in the convective boundary layer : first measurements with rotational Raman lidar, *Atmos. Chem. Phys.*, 15, 5485–5500, 2015.
- 25 Berg, J., Mann, J., and Patton, E. G.: Lidar-Observed Stress Vectors and Veer in the Atmospheric Boundary Layer, *J. Atmos. Ocean. Tech.*, 30, 1961–1969, doi:10.1175/JTECH-D-12-00266.1, 2013.
- Bonin, T., Blumberg, W., Klein, P., and Chilson, P.: Thermodynamic and Turbulence Characteristics of the Southern Great Plains Nocturnal Boundary Layer Under Differing Turbulent Regimes, *Boundary-Layer Meteorology*, 157, 401–420, doi:10.1007/s10546-015-0072-2, 2015.
- 30 Bonin, T. A. and Brewer, W. A.: Detection of Range-Folded Returns in Doppler Lidar Observations, *IEEE Geoscience and Remote Sensing Letters*, 14, 514–518, 2017.
- Bonin, T. A., Newman, J. F., Klein, P. M., Chilson, P. B., and Wharton, S.: Improvement of Vertical Velocity Statistics Measured by a Doppler Lidar through Comparison with Sonic Anemometer Observations, *Atmos. Meas. Tech.*, 9, 5833–5852, doi:10.5194/amt-2016-134, 2016.
- Browning, K. A. and Wexler, R.: The determination of kinematic properties of a wind field using Doppler radar, *J. Appl. Meteorol.*, 7,  
35 105–113, 1968.
- Calhoun, R., Heap, R., Princevac, M., Newsom, R., Fernando, H., and Ligon, D.: Virtual Towers Using Coherent Doppler Lidar during the Joint Urban 2003 Dispersion Experiment, *J. Appl. Meteorol. Clim.*, 45, 1116–1126, 2006.

- Choukulkar, A., Calhoun, R., Billings, B., and Doyle, J.: Investigation of a Complex Nocturnal Flow in Owens Valley , California Using Coherent Doppler Lidar, *Boundary-Layer Meteorol.*, 144, 359–378, doi:10.1007/s10546-012-9729-2, 2012.
- Choukulkar, A., Pichugina, Y., Clack, C. T. M., Calhoun, R., Banta, R., Brewer, A., and Hardesty, M.: A new formulation for rotor equivalent wind speed for wind resource assessment and wind power forecasting, *Wind Energy*, 19, 1439–1452, doi:10.1002/we, 2016.
- 5 Cuxart, J., Morales, G., Terradellas, E., and Yagüe, C.: Study of coherent structures and estimation of the pressure transport terms for the nocturnal stable boundary layer, *Bound.-Lay. Meteorol.*, 105, 305–328, 2002.
- Eberhard, W. L., Cupp, R. E., and Healy, K. R.: Doppler Lidar Measurement of Profiles of Turbulence and Momentum Flux, *J. Atmos. Ocean. Tech.*, 6, 809–819, doi:10.1175/1520-0426(1989)006<0809:DLMOPO>2.0.CO;2, 1989.
- Finkelstein, P. L. and Sims, P. F.: Sampling error in eddy correlation flux measurements, *J. Geophys. Res.*, 106, 3503–3509, 2001.
- 10 Frisch, A. S., Martner, B. E., and Gibson, J. S.: Measurement of the vertical flux of turbulent kinetic energy with a single Doppler radar, *Bound.-Lay. Meteorol.*, 49, 331–337, 1989.
- Fuertes, F. C., Iungo, G. V., and Porté-Agel, F.: 3D Turbulence measurements using three synchronous wind lidars: Validation against sonic anemometry, *J. Atmos. Ocean. Tech.*, 31, 1549–1556, 2014.
- Gunter, W. S., Schroeder, J. L., and Hirth, B. D.: Validation of Dual-Doppler Wind Profiles with in situ Anemometry, *J. Atmos. Ocean. Tech.*, 15 32, 943–960, doi:10.1175/JTECH-D-14-00181.1, 2015.
- International Electrotechnical Commission: IEC 61400-1: Wind turbines - Part 1: Design requirements., 3rd PPUB edition, Geneva, Switzerland: International Electrotechnical Commission, 2005.
- Kaimal, J. C. and Gaynor, J. E.: The Boulder Atmospheric Observatory, *J. Clim. Appl. Meteorol.*, 22, 863–880, 1983.
- Kelley, N. D., Jonkman, B. J., and Scott, G. N.: The Great Plains Turbulence Environment : Its Origins , Impact and Simulation, Tech. rep., 20 NREL/CP-500-40176., National Renewable Energy Laboratory, Golden, CO, USA, 2006.
- Kropfli, R. A.: Single Doppler radar measurements of turbulence profiles in the convective boundary layer, *J. Atmos. Ocean. Tech.*, 3, 305–314, 1986.
- Kumer, V.-m., Reuder, J., Doringner, M., and Zauner, R.: Turbulent kinetic energy estimates from pro fi ling wind LiDAR measurements and their potential for wind energy applications, *Renewable Energy*, 99, 898–910, doi:10.1016/j.renene.2016.07.014, 2016.
- 25 Lenschow, D. H., Mann, J., and Kristensen, L.: How long is long enough when measuring fluxes and other turbulence statistics?, *J. Atmos. Ocean. Tech. Technol.*, 11, 661–673, 1994.
- Lenschow, D. H., Wulfmeyer, V., and Senff, C.: Measuring second- through fourth-order moments in noisy data, *J. Atmos. Ocean. Tech.*, 17, 1330–1347, 2000.
- Lothon, M., Lenschow, D. H., and Major, S. D.: Coherence and scale of vertical velocity in the convective boundary layer from a Doppler 30 lidar, *Boundary-Layer Meteorol.*, 121, 521–536, 2006.
- Lothon, M., Lenschow, D. H., and Major, S. D.: Doppler lidar measurements of vertical velocity spectra in the convective boundary layer, *Boundary-Layer Meteorol.*, 132, 205–226, 2009.
- Lundquist, J. K., Churchfield, M. J., Lee, S., and Clifton, A.: Quantifying error of lidar and sodar Doppler beam swinging measurements of wind turbine wakes using computational fluid dynamics, *Atmos. Meas. Tech.*, 8, 907–920, doi:10.5194/amt-8-907-2015, 2015.
- 35 Lundquist, J. K., Wilczak, J. M., Ashton, R., et al.: Assessing state-of-the-art capabilities for probing the atmospheric boundary layer: the XPIA field campaign, *Bull. Amer. Meteor. Soc.*, 98, 289–314, doi:10.1175/BAMS-D-15-00151.1, <http://journals.ametsoc.org/doi/10.1175/BAMS-D-15-00151.1>, 2017.
- Mahrt, L.: Flux Sampling Errors for Aircraft and Towers, *J. Atmos. Ocean. Tech.*, 15, 416–429, 1998.

- Mahrt, L.: Stably Stratified Atmospheric Boundary Layers, *Annu. Rev. Fluid Mech.*, 46, 23–45, 2014.
- Mahrt, L., Sun, J., and Stauffer, D.: Dependence of Turbulent Velocities on Wind Speed and Stratification, *Bound.-Lay. Meteorol.*, 155, 55–71, doi:10.1007/s10546-014-9992-5, 2015.
- Mann, J., Cariou, J.-P., Courtney, M. S., Parmentier, R., Mikkelsen, T., Wagner, R., Lindelöw, P., Sjöholm, M., and Enevoldsen, K.: Comparison of 3D turbulence measurements using three staring wind lidars and a sonic anemometer, *Meteorol. Z.*, 18, 135–140, doi:10.1127/0941-2948/2009/0370, 2009.
- Maurer, V., Kalthoff, N., Wieser, A., Kohler, M., and Gantner, L.: Observed spatiotemporal variability of boundary-layer turbulence over flat, heterogeneous terrain, *Atmos. Chem. Phys.*, 16, 1377–1400, doi:10.5194/acp-16-1377-2016, 2016.
- Mayor, S. D., Lenschow, D. H., Schwiesow, R. L., Mann, J., Frush, C. L., and Simon, M. K.: Validation of NCAR 10 . 6-  $\mu\text{m}$  CO<sub>2</sub> Doppler Lidar Radial Velocity Measurements and Comparison with a 915-MHz Profiler, *J. Atmos. Ocean. Tech.*, 14, 1110–1126, 1997.
- McCaffrey, K., Quelet, P., Choukulkar, A., et al.: Identification of Tower Wake Distortions Using Sonic Anemometer and Lidar Measurements, *Atmos. Meas. Tech. Discuss.*, in review, doi:10.5194/amt-2016-179, 2016.
- Newman, J., Bonin, T., Klein, P., Wharton, S., and Newsom, R.: Testing and validation of multi-lidar scanning strategies for wind energy applications, *Wind Energy*, 19, 2239–2254, doi:10.1002/we.1978, 2016a.
- 15 Newman, J. F., Klein, P. M., Wharton, S., Sathe, A., Bonin, T. A., Chilson, P. B., and Muschinski, A.: Evaluation of three lidar scanning strategies for turbulence measurements, *Atmos. Meas. Tech.*, 9, 1993–2013, 2016b.
- Pearson, G., Davies, F., and Collier, C.: An analysis of the performance of the UFAM pulsed Doppler lidar for observing the boundary layer, *J. Atmos. Ocean. Tech.*, 26, 240–250, 2009.
- Pearson, G., Davies, F., and Collier, C.: Remote sensing of the tropical rain forest boundary layer using pulsed Doppler lidar, *Atmos. Chem. Phys.*, 10, 5891–5901, 2010.
- Pichugina, Y. L., Banta, R. M., Kelley, N. D., Jonkman, B. J., Tucker, S. C., Newsom, R. K., and Brewer, W. A.: Horizontal-velocity and variance measurements in the stable boundary layer using Doppler lidar: Sensitivity to averaging procedures, *Journal of Atmospheric and Oceanic Technology*, 25, 1307–1327, doi:10.1175/2008JTECHA988.1, 2008a.
- Pichugina, Y. L., Banta, R. M., Kelley, N. D., Jonkman, B. J., Tucker, S. C., Newsom, R. K., and Brewer, W. A.: Horizontal-velocity and variance measurements in the stable boundary layer using Doppler lidar: Sensitivity to averaging procedures, *J. Atmos. Ocean. Tech.*, 25, 1307–1327, 2008b.
- 25 Sathe, A. and Mann, J.: A review of turbulence measurements using ground-based wind lidars, *Atmos. Meas. Tech.*, 6, 3147–3167, 2013.
- Sathe, A., Mann, J., Gottschall, J., and Courtney, M. S.: Can Wind Lidars Measure Turbulence ?, *J. Atmos. Ocean. Tech.*, 28, 853–868, doi:10.1175/JTECH-D-10-05004.1, 2011.
- 30 Sathe, A., Banta, R., Pauscher, L., Vogstad, K., Schlipf, D., and Wylie, S.: Estimating Turbulence Statistics and Parameters from Ground- and Nacelle-Based Lidar Measurements. IEA Wind Expert Report. DTU Wind Energy., 2015a.
- Sathe, A., Mann, J., Vasiljevic, N., and Lea, G.: A six-beam method to measure turbulence statistics using ground-based wind lidars, *Atmos. Meas. Tech.*, 8, 729–740, 2015b.
- Smith, D. A., Harris, M., Coffey, A. S., Mikkelsen, T., Jørgensen, H. E., Mann, J., and Danielian, R.: Wind Lidar evaluation at the Danish wind test site in Høvsøre, *Wind Energy*, 9, 87–93, 2006.
- Stewart, R. W.: Turbulence and Waves in a Stratified Atmosphere, *Radio Science*, 4, 1269–1278, 1969.
- Stull, R. B.: An Introduction to Boundary Layer Meteorology, Kluwer Academic Publishers, 666 pp, 1988.

- Sun, J., Mahrt, L., Banta, R. M., and Pichugina, Y. L.: Turbulence Regimes and Turbulence Intermittency in the Stable Boundary Layer during CASES-99, *J. Atmos. Sci.*, 69, 338–351, 2012.
- Taylor, G. I.: The spectrum of turbulence, *Proc. Roy. Soc.*, A164, 476–490, 1938.
- 5 Tonttila, J., Connor, E. J. O., Hellsten, A., Hirsikko, A., Dowd, C. O., Järvinen, H., and Räisänen, P.: Turbulent structure and scaling of the inertial subrange in a stratocumulus-topped boundary layer observed by a Doppler lidar, *Atmos. Chem. Phys.*, 15, 5873–5885, doi:10.5194/acp-15-5873-2015, 2015.
- Troldborg, N. and Sørensen, J.: A simple atmospheric boundary layer model applied to large eddy simulations of wind turbine wakes, *Wind Energy*, 17, 657–669, doi:10.1002/we, <http://onlinelibrary.wiley.com/doi/10.1002/we.1608/full>, 2014.
- Tucker, S. C., Brewer, W. A., Banta, R. M., Senff, C. J., Sandberg, S. P., Law, D. C., Weickmann, A. M., and Hardesty, R. M.:  
10 Doppler lidar estimation of mixing height using turbulence, shear, and aerosol profiles, *J. Atmos. Ocean. Tech.*, 26, 673–688, doi:10.1175/2008JTECHA1157.1, 2009.
- Turner, D. D., Wulfmeyer, V., Berg, L. K., and Schween, J. H.: Water vapor turbulence profiles in stationary continental convective mixed layers, *J. Geophys. Res.-Atmos.*, 119, 11 151–11 165, 2014.
- Viana, S., Terradellas, E., and Yagüe, C.: Analysis of gravity waves generated at the top of a drainage flow, *J. Atmos. Sci.*, 67, 3949–3966,  
15 doi:10.1175/2010JAS3508.1, 2010.
- Vickers, D. and Mahrt, L.: Quality control and flux sampling problems for tower and aircraft data, *J. Atmos. Ocean. Tech.*, 14, 512–526, 1997.
- Vickers, D. and Mahrt, L.: The cospectral gap and turbulent flux calculations, *J. Atmos. Ocean. Tech.*, 20, 660–672, 2003.
- Wang, H., Barthelmie, R. J., Clifton, A., and Pryor, S. C.: Wind measurements from arc scans with Doppler wind lidar, *J. Atmos. Ocean.  
20 Tech.*, 32, 2024–2040, doi:10.1175/JTECH-D-14-00059.1, 2015.
- Wharton, S. and Lundquist, J. K.: Assessing atmospheric stability and its impacts on rotor-disk wind characteristics at an onshore, *Wind Energy*, 15, 525–546, doi:10.1002/we, 2012.
- Wilczak, J. M., Oncley, S. P., and Stage, S. A.: Sonic anemometer tilt correction algorithms, *Boundary-Layer Meteorol.*, 99, 127–150, 2001.
- Wilson, D. A.: Doppler radar studies of boundary layer wind profiles and turbulence in snow conditions, in: *Proc. 14th Conference on Radar  
25 Meteorology*, pp. 191–196, Tucson, USA, 1970.