Responses to referee comments:

An initial paragraph or section evaluating the overall quality of the discussion paper ("general comments"):

The paper presents an overview of the comprehensive wind lidar intercomparison field test XPIA conducted around the 300 m BOA met tower in Colorado April 2015.

The paper present and address wind lidar measurement capability and uncertainties with reference to sonic anemometry installed and operated simultaneously the met tower every 50 meter.

In total, five pulsed wind lidar systems were co-operated and their measurement capabilities are described and data retrieval algorithms and data quality including uncertainty with the reference to the Sonics intercom pared.

The paper shows the capability with the state of the art commercial available Leosphere scanners 200S and compare them with scientific grade pulsed when operated in staring mode and in standard PPI scanning mode.

The authors thank the referee for taking the time to review this manuscript and providing their valuable inputs.

a section addressing individual scientific questions/issues ("specific comments"),

The lidar measured wind fields are presented using several data retrieval methodologies, using single lidars with interpolation techniques and also multiple, but all lidar beams and measured line of sight Doppler velocities were unsynchronized and un-coordinated while measuring.

This manuscript investigates both temporally synchronized and un-synchronized measurement uncertainty. The authors would like to note that while it is true the lidar measurements discussed here were not coordinated using techniques described in Vasiljevic et al (2016), temporal coordination was achieved. As summarized in Table 2 and 3, the Doppler lidars were able to maintain pointing uncertainty of less than 0.15^o (+/- 2.5m @ 1km range) and time uncertainty of less than 0.4 s. Since, all our averaging periods are much greater than 0.4 seconds and the common volume is much bigger than the pointing uncertainty, we feel that the lidars were able to make synchronized measurements (in time and space) when we aimed to do so. We have renamed the sections to make it clearer where synchronized scanning was used and where the measurements were allowed to be unsynchronized.

The quality of the resulting wind field measurements seems, in addition to large probe volumes, signal-to noise issues and laser beam pointing accuracy seems challenged also by select tradeoff between relative low sampling rate and the quest to cover large spatial coverage.

The measurement techniques investigated span the range of highest sampling frequency with lowest spatial coverage (virtual tower stare scans) and highest possible spatial coverage but lower sampling frequency (multi-Doppler volume scans) due to mechanical and data rate limitations.

The logistical setup of the lidars around the BOA tower seems not to be ideal but impaired by the BOA towers surroundings land use, resulting in that the experimental setup, including angles of vertical inclination and distances were not ideal for demonstrating the full potential capabilities of the pulsed lidars. Nevertheless, given these constraints, which seems to be imposed by the surroundings, this paper documents to my knowledge the first lidar-to lidar intercomparison study between different groups lidar systems, and as such, it presents a study on capability, including data retrieval methodologies, and uncertainties, that have not previous been described using lidars.

The authors agree that the setup is non-ideal. However, even though two of the lidars have close to 180 difference in azimuth, they have substantially different elevation angles when interrogating volumes near the BAO tower and hence do provide some unique velocity information.

As such although not ideal, this paper should be published as one of the first references for multiple joint lidar wind measurement capability and intercomparison.

The lidars applied all seem to be limited in their scanning operation to perform either PPI or RHI scanning. This is unfortunate and limits the best use of the lidars. Unfortunately, at the time of XPIA, limited scan mode represented the state of the art. However, if user defined multiple lidar coordinated trajectory scanning had been available the trajectory scanning over large volumes of air could be more flexible and efficient.

The authors agree with the reviewer about this as the operation of the lidars was limited by the GUI provided by the manufacturer. In addition, as many commercially available lidar systems currently are not capable of performing more complicated scanning trajectories, this paper discusses the possibilities available to such lidar systems. We have added text to the manuscript to make this point.

Listing of purely technical corrections

It is suggested to amend the references in the paper with previous and recent publication about other staring lidar field tests with the following publications:

Line 95: I suggest to amend the ref list re triple lidars by adding a ref to Simley et al[1]

Added

Line 96: regarding long-range triple scanning lidars, add ref to Vasiljevic et al. 2016 [2]

Added

Line 120: regarding triple staring lidars, add ref to Mann et al. 2009 [3]

Added

Line 159: the ref is not found in the ref list?

Fixed

Line 161: foot note: need to explain why this particular lidar needed additional accumulation time- it's the lidar closest to the BAO tower hence it should have the highest CNR?

Added: "… to ensure coverage during multi-Doppler volume scans"

Line 177 for helping the reader it could be stated that this corresponds to approximately +/- 1 meter pointing uncertainty at the 1 km range

Added

Line 191: add a ref to Mann et al. (2009) again following the ref to Calhoun 2006

Added

Line 311: 2 the meaning of the statement "from lags 1 through 4" is not clear to me, is 1-4 range gates, if so pls specify?

The lags refers to the lags of the autocovariance of the LOS velocity. Modified the statement to make this clearer.

Line 320: Stating: "To determine the additional uncertainty due to range averaging" The word "uncertainty" may not be the most correct word to use here, the lidar's range-average results in a deterministic filtering effect, that filters variance at short length scales, but this is not an uncertainty issue, rather an instrument filtering effect, which is common to all probe volume averaging instruments, lidars included. Could it be stated as: To quantify the variance reduction due to spatial averaging of the lidars Line-of-sight probe volume, …

We agree with the reviewer that the effect of the range-gate volume is to filter the variations at smaller scales. However, this can result in differences in the estimate of the LOS velocity (and its derivatives) when comparing with a “point” measurement such as a sonic anemometer. Hence, in this context we refer to the effect of the volume averaging as contributing to the uncertainty.

[1] Simley E, Angelou N, Mikkelsen T, Sjöholm M, Mann J and Pao L Y 2016 Characterization of wind velocities in the upstream induction zone of a wind turbine using scanning continuous-wave lidars *J. Renew. Sustain. Energy* **8** 13301

[2] Vasiljevi N, Lea G, Courtney M, Cariou J, Mann J and Mikkelsen T 2016 Long-range WindScanner system *Remote Sens.* **Dec 2016** 1–24

[3] Mann J, Cariou J-P, Courtney M S, Parmentier R, Mikkelsen T, Wagner R, Lindelöw P, Sjöholm M and Enevoldsen K 2009 Comparison of 3D turbulence measurements using three staring wind lidars and a sonic anemometer *Meteorol. Zeitschrift* **18** 135–40