Reply to Anonymous Referee #2:

We appreciate the reviewer's comments on the manuscript. We have studied them carefully. Below are point-by point responses to the referee's comments.

Comments from the editors and reviewers:
General comments:

This manuscript reports climatological values of mixing layer height (MLH) over Beijing for from Lidar measurements, their evaluation against radiosonde based MLH calculation and argues that MLH values derived from Lidar is better than that from RS for estimation of PM2.5 at surface. The manuscript needs a lot of improvement before publication. Please see comments and suggestions below:

1. The first part of the results and discussion, where the comparison of the various approaches to estimate MLH is presented is not satisfactory. First, the English language is poorly written which makes it hard to understand what is being conveyed. Also, the text corresponding to Figures 2-4 lacks any discussion of the features seen in them. More detailed discussion and language improvement is needed.

   A1: This part has been rephrased, please see the revised manuscript.

2. Fig 3d: why is MLH_RS detected at 0.6 km? it should be 2km as seen in Fig 3c.

   A2: The vertical profile present is actually indicated by "the white edge triangle in the upper picture", also the measurement time can be seen on the top of Fig. 3c (20170606-14) and Fig. 3d (20170606-20). Fig. 3d show MLH_{RS} (0.6 km) at 2000 LST, while Fig.3c indicates (2 km) at 1400 LST. Due to the different time, MLH_{RS} varies.

3. Actually, the whole context of the first 4 figures is not understood. Why are these days shown here? why not any other day? Is it meant to show seasonal variability i.e one for one season? If the figures are introduced to show various types of differences between the 3 methods, then the discussion show be organized in that manner and the inter comparison figure 5 should be discussed in context of features seen in these figures 1-4. Better organization is needed.

   A3: Fig. 1-Fig. 4 present as case study is aimed to show the evolution of the MLH and comparisons of lidar measurement (MLH_L and MLH_L') and RS measurement (MLH_{RS}). The four cases is almost all cloud free to present continuous retrievals. Fig.1 shows obvious separated layers in the evening (spring without 1400LST measurement). Fig.2 shows low aerosol load, in which condition lidar can still capture the small gradient (summer with 1400 LST data). Fig. 3 relative high aerosol load and MLHL and MLHL' show obvious difference in the afternoon (summer with 1400 LST data).Fig. 4 keep stable though the whole day. The information delivered of the figure are discussed in the following context in the revised manuscript.

4. Please give more details about the cloud screen/flag used (Line 111). How was cloud detected, at what resolution etc?

   A4: The content is added that "a threshold is selected to distinguish between clouds and aerosol layers." and the lidar "can detect a long range profile up to 30 km every 1 second. For the enhancement of signal noise ratio, 60 profiles are averaged to restore as one with the time resolution of 1 minute." "For convenient comparison with air quality and meteorological parameters, all MLH results are one hour averaged."

5. Fig 1a, 2a,3a,4a all are stretched, the fonts are of different sizes compared to other panels and should be made consistent. Figure 1b have an undefined variable MLH2 in the figure label?

   A5: The layout of the sub figure is based on requirement of clarity, expressiveness and artistry. In our view of Fig. 1a-4a, the information conveyed changes nothing. In Figure 1b, MLH2 has been revised to MLH_L', which indicates the first local maximum of lidar MLH.

6. Evaluation of MLH_lidar is essential for second half of this manuscript. Hence, more detailed comparison and discussion is necessary. scatter plots in supplementary should be presented in main and discussed in detail. Moreover, as seasonal variability is significant on MLH, the comparison in Figure 6 should include and discuss scatter plots for all the seasons, separately.

   A6: The comparisons of MLH from lidar (MLH_L and MLH_L') and MLH_{RS} is the one of objective of the study. From Fig. 5, we can see that about 35% SBL is not at the lidar detection range for 0800 and 2000 LST. It can be concluded that the agreement of MLH from lidar (MLH_L and MLH_L') and MLH_{RS} is poor, even though the scatter plots is not shown. The comparisons between MLH_L' and MLH_{RS} in the afternoon can be seen from the case study and the diurnal cycle, which indicates MLH_L' is much lower than MLH_{RS}. Please see the revised manuscript. Actually, the data of radiosonde is only available in summer, so the comparisons of the other seasons cannot be conducted.

7. From Fig 1-5, MLH from Lidar is termed as MLH but in Fig 6, it is termed as MLH_Lidar. Please be consistent.

   A7: All the biggest local maximum MLH of lidar is represented by MLH_L, while the first local maximum is indicated by MLH_L' and radiosonde result by MLH_{RS}. They are consistent in through the revised manuscript.

8. Fig 9: Is the MLH from RS also showing yearly variability similar to the MLH derived from Lidar? Please include the same from RS also in this figure and discuss.

   A8: From 2014 to 2018, MLH_{RS} at 0800 LST is 0.402, 0.412, 0.453, 0.444, 0.451 km, respectively. MLH_{RS} at 2000 LST is 0.445, 0.501, 0.512, 0.515, 0.480 km, respectively. Both of the two measurement time show increasing trend. Due the availability of RS measurement is only in summer, it cannot compare with the annual mean value of lidar. We can see the annual variation of RS is not exactly the same as lidar. Actually, in the revised manuscript, the compassion of lidar

and RS measurement is discussed it the prior section, and in the section, we focus to present the annual variation of diunal cycle of MLH from lidar, while the RS measurements is just time points result.

9. In Figure 10d shows that correlations are highest when RMSE is also highest, this is non intuitive, please describe this feature. Also, the comparison in Figure 10 should be presented separately for each season as MLH has strong seasonal variations.

   A9: The comment is added in the revised manuscript that "The smaller RMSE is related to the limited samples." And "larger RMSE is associated to the larger amount of samples into statistic." In the revised manuscript, there is already a lot of content. The seasonal variations of calculated PM2.5 and in-situ can be shown in the next paper.

10. Throughout the manuscript standard deviation and RMSE is used interchangeable, which should be corrected. They are not same.

    A10: Thank you for reminding, and the standard deviation has been revised to RMSE.

11. Please check carefully for the typos in the manuscript. The title itself has one "airpollution"

    A11: Thank you for reminding and I have noticed the typos. Hope the staff member of ACP could help to correct it. Actually, in my manuscript, it is right "air polllution".