Responses to Reviewer #1:

We are very thankful to the reviewer for his/her valuable advice and comments, which helped us improve the manuscript significantly. We have addressed these comments and revised the manuscript accordingly. The detailed responses to the queries are below:

Major comments

1. *Page 4051, lines 7-9: This part is unclear: if the radiosondes are at 15 minutes past the hour then the sounding during the next hour can always be used. So not sure what the authors mean by these lines.*

   Response: The radiosondes are not available for every hour. As mentioned in the manuscript, "during the experiment, radiosondes were released 2-4 times a day (around 05:15, 11:15, 17:15, and 23:15 LST)." As such, if the radiosondes are at 15 minutes past this hour, it is likely that there is no radiosonde data during the next hour. These lines (Page 4051, lines 7-9) have been modified to be:

   "**To ensure accuracy, only soundings released within 15 minutes around the hour were used in this study, yielding a total of 168 records.**"

2. *Equation 2, second line: I see that the condition on the second gradient is taken from a reference but the authors should elaborate on the physical rationale of such condition since it seems rather ad-hoc.*

   Response: The elaboration of the second gradient is added after page 4052, line 20, as follows:

   "**For Type I SBL, PTG decreases with height and the inversion near the surface is relatively strong. There is always a sudden decrease of PTG at the PBL top (Fig. 1a1). As such, the derivative of PTG with respect to *z* should be negative, that is,** $d^2\theta / dz^2 < 0$**.While for Type II SBL, PTG increases with height and the inversion is relatively weak. No sudden change of PTG at the PBL top is seen (Fig. 1a2) and thus** $d^2\theta / dz^2 \geq 0$**.**"

3. *Page 4052, lines 16-18: The selection of the value of δ seems rather arbitrary and seems due more to measurement accuracy. Also according to this picture there is an abrupt transition when H goes from -1 to 1 from a stable to an unstable PBL, but physically it is unclear if that actually happens. In a modeling framework, that would suddenly alter the height of the PBL by potentially hundreds of meters as the Ri_cr is switched from the SBL to UBL in the proposal model at $H \approx 0$.*

Response: We agree with the reviewer that the value of δ is specified through trial and error, which depends on measurement accuracy as well as surface properties. We have tested many different δ values from 0.1 to 10, and the most reasonable values based on the observations are used in this paper. According to the equation 2, a boundary layer is classified as an UBL when $H \geq \delta$ and a SBL when $H < \delta$. As such, when $\delta = 1$ as adopted in our study over land, cases with -1<H<1 (i.e., near-neutral conditions) is actually considered to be SBLs.

The reason that δ is specified as a small positive number instead of zero is to allow for near-neutral conditions to be handled by the methods for SBLs. Since under near-neutral conditions, stable stratification usually prevails above the boundary layer and wind shear is the only source of turbulence. Both of these features are similar to those of a stable boundary layer, and as a result, the near-neutral cases are treated as the SBL cases (Serbert et al., 2000), to be exact, Type II SBL cases, as mentioned in page 4055, line 21-22.

We again agree with the reviewer that indeed there is possibly an abrupt change of Ri_{bc} between 0.30 (type II SBL) and 0.39 (UBL) when H crosses the threshold $\delta = 1$. However, we note that such change of Ri_{bc} has little effect on the PBL height determination, because under near neutral condition the Ri_b increases drastically with height at the PBL top, and using 0.30 or 0.39 as Ri_{bc} only changes the PBL height by about 15 m (or 3%). Figure S1 shows the calculated boundary layer height from the LLJ method (the black arrow) and the bulk Richardson number (the green and purple arrows) with Ri_{bc} = 0.3 and 0.39. As can be seen, Ri_{bc} does not affect the calculated PBLH significantly.

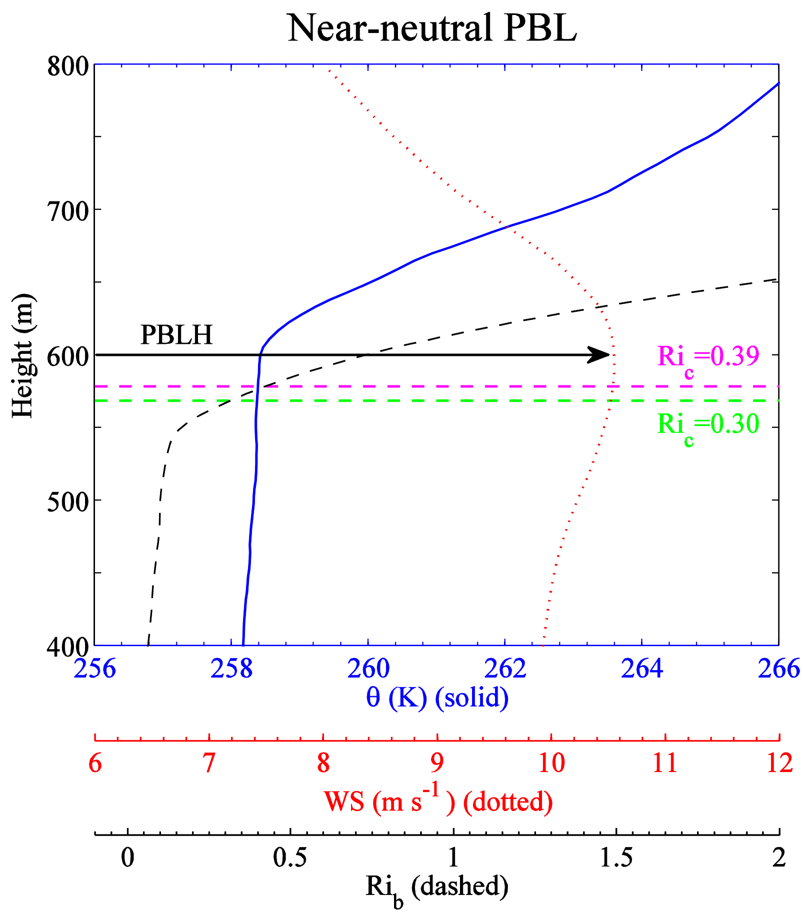

Figure S1: Typical profiles of potential temperature (blue), wind speed (red), and Ri_b (black) for a near-neutral PBL (from SHEBA on 19 October, 1997 0515 LST). The PBLH indicated by the black arrow is calculated by the LLJ methods and the PBLH indicated by the green and purple arrows are calculated by the bulk Richardson number method with Ri_{bc}= 0.30 and 0.39, respectively.

The relative discussion has been added in section 3.1, as following:

"In our study, cases with $-\delta < H < \delta$ (i.e., under near-neutral conditions) are treated as Type II SBL cases. Since under near-neutral conditions, stable stratification usually prevails above the boundary layer and wind shear is the only source of turbulence. Both of these features are similar to those of a stable boundary layer, and as a result, the near-neutral cases are treated as the SBL cases (Serbert et al., 2000). It appears there is an abrupt change in the calculation of PBLH at H = δ, the threshold above which the boundary layer is classified as an UBL and below which the boundary layer is classified as an SBL. However, changes of Ri_{bc} at H = δ from UBLs to SBLs have little effect on the PBL height determination, because under near neutral condition the Ri_b increases drastically with height at the PBL top and thus using Ri_{bc} for either UBLs or SBLs gives reasonable estimates of PBLH."

4. Page 4053: line 4: what is the magnitude of the drop, particularly that the drop in *the figure seem to be of different magnitudes? Is it automated?*

Response: In the turbulence method, continuous wavelet transform is applied to the absolute magnitude of turbulent fluctuations of each velocity component. The PBLH is automatically determined to be the level at which the absolute magnitude of these velocity fluctuations shows the most rapid decrease with height. This is similar to the methodology as detailed in Dai et al. (2014).

It is true that the heights determined by u', v', w' are usually different, so we did a weighted average using the absolute magnitude of the reciprocal velocity fluctuations as weights.

The related discussion has been modified in page 4053, lines 3-5, as follows:

"**In the turbulence method, continuous wavelet transform is applied to the absolute magnitude of turbulent fluctuations of each velocity component. The PBLH is automatically determined to be the level at which the absolute magnitude of these velocity fluctuations shows the most rapid decrease with height (Dai et al., 2011; 2014). The PBLHs determined by u', v', w' are then averaged using the absolute magnitude of the reciprocal velocity fluctuations as weights.**"

References:
Dai, C., Wang, Q., Kalogiros, J. A., Lenschow, D. H., Gao, Z, and Zhou, M.: Determining boundary-layer height from aircraft measurements, Bound.-Lay. Meteorol., 152, 277-302, doi:10.1007/s10546-014-9929-z, 2014.

5. *Page 4054, line 20: defining the lowest level as the PBLH seems ad-hoc and maybe these periods should instead not be used.*

Response: Following the suggestion of the reviewer, we have removed these cases in the latest results. Since the number of these removed cases is small and the PBLHs of these cases are also small, removing these cases has little impact on the error analysis and there is no visible change in the results of Ri_{bc}.
The related figures have been modified, and page 4054, line 20-21 has been modified as follows:

"…**if there is a LLJ, the case is reclassified to a Type II SBL; if not, the case is removed.**"

6. *Page 4059, last line: The absolute bias the authors use should in fact be able to reflect the dispersion since negative and positive errors would not cancel out as with the regular bias (by the way this should be called absolute bias rather than bias). So the first part of the line should be removed.*

Response: This suggestion is adopted. "…because the bias cannot reflect the dispersion of data" in page 4059, last line has been removed. In page 4059, line 12-13 has been modified, as follows:

"**Bias, SEE, and NSEE are the absolute bias, standard error, and normalized standard error of**…"

7. *It seems the model performance is in general sensitive to zs, so why not optimize for the value of zs also?*

   Response: Thanks for the constructive suggestion. We indeed tested many values of z_s, as can be seen from Figure 9 and 10 and Table 1. It is found that the model performance is not significantly sensitive to z_s in the stable boundary layer. In Figure 9, we can see the errors for z_s=40 and 80 m are close, especially for ARM with a large number of samples. In Figure 8, the better performance for z_s=40 m than z_s=80 m is mainly due to that the sample size for z_s=40 m is much larger. However, the model performance is indeed sensitive to z_s in unstable boundary layer. We did many different tests with z_s as 40, 80, 120, 160 m, 0.1 PBLH, and z_{SAL} (the level of the first minimum potential temperature from surface). As shown in Figure 10, we found z_s=z_{SAL} was optimal among these tests, and the impact of z_s on Ri_{bc} was also not significant as compared to the impact of thermal stratification on Ri_{bc}. The optimal Ri_{bc} with the total sample are close for different z_s, as shown in Table 1.

Minor comments:

1. *Page 4048, lines 1-2: The statement is very generic and I do not recall seeing it in Stull stated in that way. For example, when the turbulence diminished to what? It should be revised.*

   Response: Page 4048, line 1-2 has been revised, as follows:

   "**The PBL is characterized by the presence of continuous turbulence, while turbulence is lacking or sporadic above the PBL. Therefore PBLH is usually defined as the level where continuous turbulence stops (Wang et al, 1999; Seibert et al., 2000).**"

2. *Page 4048, lines 13: in the SBL the buoyancy force can be positive or negative, depending on whether the parcel is displaced upwards or downwards from its equilibrium position, so please remove the word "negative" (it is the buoyancy TKE term that is on average negative in the SBL).*

   Response: The word "negative" has been removed.

3. *Page 4049, line 26: delete "there is even" or fix the next line to be grammatically*

correct.

Response: “there is even” has been removed.

4. *Page 4051, line 8: add “a” before “time”*

   Response: Revised.

5. *Page 4052, line 19: replace “noises” by “variability”*

   Response: Revised

6. *Page 4054, line 20: replace “classified” by “reclassified”*

   Response: Revised

7. *Page 4056, line 18: replace “replaced by” by “estimated as the”*

   Response: Revised

8. *Figure 10,11 and related figures: It would be good if the authors can homogenize the y-scales and make them similar for a given metric*

   Response: Figure 8-11 have been revised.

9. *The legend of Figure 11 seems wrong. For example it is unclear which part of the figures or lines correspond to the SBLs and UBLs mentioned in the caption.*

   Response: For this part, we gathered all types of soundings instead of distinguishing them. Our goal was to get an optimal Ri_{bc} for all types of soundings by error analysis. So in this figure the comparison between estimation and observation did not distinguish boundary layer type.

   The caption of Figure 11 has been modified to emphasize this, as follows:

   “**Figure 11. Comparison between PBLH estimated using the Bulk Richardson number method and PBLH estimated using the PTG, LLJ, and modified parcel method for all types of PBLs. The correlation coefficient (a), bias (b), standard error (c), and normalized standard error (d) are shown. The sounding data are taken from Litang (plus sign), ARM Shouxian (diamond), and SHEBA (pentacle). The curved lines are obtained by quadratic curve-fitting, the black vertical dashed lines indicate a representative Ri_{bc} for all three sites, and the error bars indicate the range of Ri_{bc} across the three sites.**”