Anonymous Referee #1

General comments:
This manuscript compared 7 methods to diagnose PBL depth in the GEOS-5 model. Such investigation is important for applications that use the PBL depth as input. However there are some deficiencies in its current form (see comments below). Thus I recommend major revision before it can be accepted for publication.

Thank you for taking the time to review this manuscript. We are glad to hear that you found this work important. We have addressed your concerns in the comments below.

Major comments:

1. All the PBL depth diagnose methods are actually based on vertical profiles of variables such as potential temperature, wind, TKE etc. Vertical profiles of these variables describe/illustrate the boundary layer structure more clearly. So the comparison of those PBL depth diagnosed by different methods (e.g., Fig. 6) should be discussed with the aid of vertical profiles of those relevant variables. Good example are Fig. 3 of (Nielsen-Gammon et al., 2008), Fig. 1 of (Seibert et al., 2000)

An additional figure (new figure 6) has been added to the revised manuscript showing day and nighttime profiles of relevant variables and a discussion has been added concerning how they impact the PBL depth estimate.

2. Table 1 lists the PBL depth diagnose method for the model simulation. What is the method used to diagnose PBL depth based on radiosonde observations (particularly at night)? A separate section is needed in method section to describe the observations.

In the revised manuscript, a section on the radiosonde observations has been added to the methods section. The radiosonde observations were evaluated using the bulk Richardson number method (Method 4).

3. I would expect the nighttime PBL depth diagnosed by method 6 (Ricrit=0) is 0 in most places since there is always temperature inversion near the surface. I am a little surprised to see significant nighttime PBL depth from this method.

The minimum PBL depth obtained by the GEOS-5 AGCM coincides with the lowest model level at about 150 m. This has been added to the text in the revised manuscript.

4. Some of the text is repetitive from the figure captions, e.g., "The horizontal dashed lines indicate the PBL depth found using the total Kh (Method 1, Fig. 7a) and bulk Richardson number (Method 4, Fig. 7b).", which are unnecessary in the text.

Repetitive text has been removed.

Specific comments:

1. LN21, page 6593, The sentence of "estimates the PBL depth as the model level below that which Kh falls below" does not make sense.

This has been reworded in the revised manuscript.

2. "A spatial map of the JJA skin temperature (Fig. 3b) shows the same pattern as the PBL depth". This is not true for the Arabian Peninsula. Any reason why?

This is due to upslope winds opposing the overlying subsidence over the higher topography of the western part of the Arabian Peninsula. This has been added to the discussion in the revised manuscript.

3. "In general, both local Richardson number methods (Methods 5 and 6) estimate PBL depths that are lower than the other methods throughout the diurnal cycle." Please analyze vertical profiles of relevant variables such as potential temperature, wind, TKE, local Ri (pick up one time in the day and one at night)

We have updated Figure 7 (now figure 8) to include the vertical profile of wind speed. It also now includes both a daytime and nighttime profile.

4. "This has implications for estimating the shallow nocturnal boundary layer that has been shown to be relevant for constituent transport". This is not true for some air quality models that do not use the variable of PBL depth to compute transport of constituents. For these models, constituent mixing does not depend on the diagnosed PBL depth.

This has been reworded in the revised manuscript.

5. Many studies investigated/compared PBL depth diagnose method (e.g., (Helmis et al., 2012; Hu et al., 2010; Nielsen-Gammon et al., 2008)). They need to be better summarized in introduction.

The introduction has been expanded in the revised manuscript.

References:

Helmis, C.G., Sgouros, G., Tombrou, M., Schafer, K., Munkel, C., Bossioli, E., Dandou, A., 2012. A Comparative Study and Evaluation of Mixing-Height Estimation Based on Sodar-RASS, Ceilometer Data and Numerical Model Simulations. Bound-Lay Meteorol 145, 507-526.

Hu, X.M., Nielsen-Gammon, J.W., Zhang, F.Q., 2010. Evaluation of Three Planetary Boundary Layer Schemes in the WRF Model. J Appl Meteorol Clim 49, 1831-1844.

Nielsen-Gammon, J.W., Powell, C.L., Mahoney, M.J., Angevine, W.M., Senff, C., White, A., Berkowitz, C., Doran, C., Knupp, K., 2008. Multisensor estimation of mixing heights over a coastal city. J Appl Meteorol Clim 47, 27-43.

Seibert, P., Beyrich, F., Gryning, S.E., Joffre, S., Rasmussen, A., Tercier, P., 2000. Review and intercomparison of operational methods for the determination of the mixing height. Atmos Environ 34, 1001-1027.

Anonymous Referee #2

General comments:
This manuscript compared 7 different methods of defining the planetary boundary layer (PBL) depths in the GEOS-5 atmospheric general circulation model over land. It is very important to evaluate the PBL depths in models. However, there are some major issues in this manuscript. Please see comments below. Therefore, I recommend accepting this manuscript after major revisions.

Thank you for your time in reviewing this manuscript. We agree that it is important to evaluate the PBL depth in models. Your concerns are addressed in the comments below.

Major comments:

1. The introduction didn't cover most the PBL depths comparison works. It only referred to 3 papers related to PBL depths.

We have expanded the introduction to include studies by Helmis et al. (2012), Hu et al. (2010), Nielsen-Gammon et al. (2008), and Seibert et al. (2000).

2. Regarding Method 5 and 6, I am not surprised at "In general, both local Richardson number methods (Methods 5 and 6) estimate PBL depths that are lower than the other methods throughout the diurnal cycle." There are many inversions near the surface or at the low level, especially at night or in winter. It didn't get much meaningful information based on these two methods according to your results. Maybe consider to remove these two methods or just compare one of them or write some words on these methods?

In the new manuscript, the discussion about Methods 5 and 6 is expanded to include mean and percentage differences between the methods as well as the characteristic that the differences are maximal during the afternoon.

3. There are some comparisons between model results and the radiosonde-based results. Based on my understanding, the PBL depths provided in Seidel et al. (2012) is for the period 1981-2005, while the model results in this manuscript is from 1990 to 2013. In lines 17-27 Page 6601, it is said the differences between the model and radiosonde estimates were just 100 m. I don't think it gave us some convincible information since they used different study period and the differences could be changed a lot if using another study period.

We've included this comparison to radiosondes because both the model simulated PBL depths and those estimated from the radiosonde profiles represent climatological conditions and so provide an estimate of model uncertainty. The new manuscript acknowledges the difference in time period, but that both are climatological datasets in the subsection describing the radiosonde observations.

Specific comments:

1. In Section 2, please clarify the definition of the seasons used in this manuscript. In Figure 2, the "summer PBL depth" just popped out. Also, please describe briefly the observation results you used in this section.

We have moved our definition of summer and winter to section 2 and included a separate subsection within Section 2 describing the radiosonde observations.

2. Figure 2: It shows the four different climate classes. Why does the tropical forest show the annual result and others show the summer results? Additionally, please use the same scale of the y-axis and it would be easier to compare.

We show the annual result for the tropical rainforest because seasons near the equator are more dependent on precipitation than on temperature and there is no distinct summer or winter seasons. We have updated the figure to use the same y-axis for all four panels and explained why we show the annual mean diurnal cycle for the tropical rainforest.

3. Figure 3: You can't say "...PBL height variability is explained by skin temperature" only based on some correlation test or some diagram. The variation of PBL depth is complicated.

We have reworded this and added a section explaining that variability exists that is not explained by temperature.

4. Figure 5: The error bars are barely seen.

We've increased the size of the error bars and extended them to three standard deviations.

5. Figure 7: The x, y axes' labels are nearly overlapped with the values

This has been corrected.