Major comments:

1. It is meaningless to do sensitivity experiments if the base case (Control) failed to capture the basic meteorological fields (e.g., surface sensible and latent heat flux, temperature, etc.).

In my view, these sensitivity experiments should be done at the beginning for choosing the best configuration for the base case, so there is no need to present your initial testing results. Moreover, even the author didn’t conduct the sensitivity test (e.g., the surface heat flux effect); we can expect the corresponding results.

2. Once the author obtained the best configuration, the results from base case should be carefully validated with the observation. Note that you choose the observation as a true value to evaluate the performance of the model, so you can’t doubt the reliability of observational data. Unfortunately, the author attributes the huge discrepancy between simulation and observation to that the observed sensitive and latent heat flux might be underestimated (Lines 215-216). The author should find the reason from the model itself, rather than observational data!

3. Even for the observational data, the author didn’t mention how the surface temperature, 2m temperature, wind speed and wind direction at 10m comes from. Moreover, the author introduced the eddy-covariance system “was a R3-50 supersonic anemometer developed by Gill Company, UK……. (lines 174-175)”, that is totally wrong, how can you obtain SH and LE only from supersonic anemometer? There are so many related references on it, the Data part should be carefully described.

4. In order to explain the huge discrepancy between simulated and observed SH and LE, the author should examine every component of surface energy balance equation, like incoming downward shortwave radiation, etc., to better find which factor make this discrepancy happens.

5. The author states that they have verified PBL parameterization schemes in their previous study of Xu et al. (2018) in the response to the 3rd question of reviewer 3. One conclusion from Xu et al. (2018) is that the vertical structure simulated with the Shin-Hong scheme was closer to that in both the WRF-LES (large-eddy-simulation) and observations than that simulated with the YSU. So why the authors still choose YSU scheme in current study rather than Shin-Hong scheme?

6. I am very shocked that the lowest level of the WRF is set to 1130.473m (lines 136). If it is true, how the MOST applied in such height between the surface and the first model level? Is it reasonable and why? I guess because of this reason, the values in the vertical cross section shown in Figure 6 and Figure 8 is missed below around1km.

7. It is commonly accepted in the boundary layer community that the WRF need to spin up, usually the model output from the first 8 hours or more need to exclude before comparing with observation. However, the author just run the case with 12 hours and didn’t exclude any data in spin up period, is it reasonable?

8. if the author want to evaluate the impact of different land surface model, please explicitly explain the difference between them. Also the land use of two different dataset should be also given.

9. Current version of this paper are failed to check and correct the grammar mistake throughout the whole manuscript, lots of issues are still existed. The author state that this revised paper have rescrutinized to improve the English by a native English speaker, however it is obvious not.

Specific comments

1. Line 45, full name of GPS, explains it.

2. line 70, Wang et al?, the year lost.

3. line 77, “development of a near-surface thermal low pressure system,”, reorganize the sentence.

4. line 81, “fundamentally restrict the development of understanding desert and surrounding area”, it is like translate directly from Chinese

5. line 86, “To fill in the gaps of Taklimakan desert”, oh, my god!! How can you fill in the gaps of Taklimakan desert, there are so many similar issues.

6. line 97, PBL can heavily impacted???

7. line 98, “One way to tackle complex turbulent flows in weather forecast models is Large eddy simulation (LES)…..”, LES is a way??

8. line 112, “Thus, the LBCS of can significantly alter high-resolution LES status through inflow boundaries(Rai et al. 2017).”, LBCS of what???

9. line 118, “this paper is to examine assess the skillfulness….”, my god, examine assess??

10. line 136-138, what is the unit of the heigh?

11. line 140, “The sizes of model grids are 411 ×321 791x651 211x201 and 403x406 respectively.”, it should be numbers, not sizes.

12. line 216, “Researchers found that”, which researcher? Please cite the reference correctly.

13, line 258-260, how can you attribute the reason to the potential temperature lapse rate?? If the initial temperature profile is failed to agree with observation, how can you say your model result is reliable?

14, line 266, CBLH? Explain it!

15. line 318, “LES simulation”, LES has contain simulation, so the simulation after LES is needless.

16. line 330-331, “However, the comparison results reveal that discrepancies among different experiments are large for CBL”, the discrepancy of which variable is large?? Please specify.

17. line 350, “CBL, the instantaneous vertical velocity fields for the horizontal are displayed in”, for the horizontal? Are displayed in what?

18. line 353, 357, the figure number is missing.

19. line 375, the surface-land schemes??

20. line 398-399, “This indicates that SH may not the dominant factor for the deep CBL over the Taklimakan desert.” Here you say SH is not the dominant factor, but in conclusion (line 461) you mention SH is dominant, can you convince yourself before present to readers?

21. line 414, large LES experiments? What is that?

22. line 446-447, “The overestimation of CBL profile may be caused by discrepancy between model and measurement initially”, Does this sentence make sense?

23. Figure captions in page 23 didn’t match with following one.