Technical note: Applicability of physics-based and machine-learning-based algorithms of geostationary satellite in retrieving the diurnal cycle of cloud base height – Wang et al. (egusphere-2024-1516)

This paper addresses several interesting but somewhat diverse research topics, and I appreciate the authors' efforts in conducting this work. However, the manuscript gives the impression that multiple studies are being combined, which can make it difficult to discern whether the primary focus is on the development of the ML algorithms or the accuracy of the CBH products from each algorithm. I believe that the organization and clarity of the presentation could be improved to better highlight the main objectives.

If the focus is on the diurnal cycles of CBHs, some sections, such as the MODIS product evaluation, might not be necessary and may detract from the central message of the paper. A more concise reference to this information, perhaps by adding appropriate citations, could suffice. Streamlining the content in this way could help avoid the level of detail typically found in a graduate thesis and keep the reader's attention on the core contributions of the paper.

Another main concern is, regarding line 251 and others:

I find the lack of discussions or additional explanations of CTH aspects in the CBH eval comparisons, even though those are mentioned in the description part. Probably the evaluation results from Seaman et al. 2017 or Noh et al. 2017 cited in this paper (e.g., line 407-) were conducted under the "within-spec" condition when CTH is a 2-km error range compared against ground truth data, which aimed to isolate CBH eval, decreasing CTH effects as the physics-based algorithms are highly dependent on CTH accuracy.

No sufficient discussions on multilayers: Multilayer cases (either limitations or future plans) as well as nighttime cases should be addressed in the conclusion.

Detailed comments and questions for further clarification are below.

Comments/questions:

Line 50 "sensor may be attributed to utilizing the same dataset …" : Not clear.

Line 130 "These methods aforementioned are prominent in retrieving CBH … space-based remote" : Not clear, which methods you are referring to?

Line 132 ...”The first method ...”: It seems like a starting sentence is missing. Clarification needed. What these first and second methods about?

Line 161: “... algorithm, achieving a high correlation coefficient (R) of 0.92 and a low root mean square error (RMSE): add compared against which data?

Line 167-168: It should be partially true, but Tana et al.'s study that the authors cited right above has used Himawari-8 data, which doesn't support this argument. It would be better to replace "mainly" with partially or something similar in line 166 and also better to rewrite line 164-168 to address the diurnal cycles haven't been well investigated in both GEO and LEO remote sensing research.

Line 217: For MODIS data, please add a couple of additional explanations why the authors describe MODIS here with such details (including the MODIS product evaluation in Appendix), in order to help readers' understanding, even though the reason appeared later but not clear yet here in this general Data section. Otherwise, it may distract the main topic of this paper.

Line 272-273: No need to address this here, out of the main focus of this study which adopts only the algorithm for the H8 application, not JPSS VIIRS, anyway.

Line 287: This “reliable” looks already quite deterministic. "Another" should be enough here.

Line 291-292: “... studies have also demonstrated a R of 0.569 and a RMSE of 2.3 km for the JPSS..” -> This is also unnecessary here (no info how the error statistics were obtains won't be helpful to readers, too), slightly out of this study's scope.

Line 311-317: It would be good if this is a thesis, but somewhat too much extra information for the paper. It looks already enough by citing Breiman and Min et al. Tan et al.'s papers.

Line 328: “based on” -> using additional information from NWP model data or similar sentences may be considered. Need to rewrite. It may give the impression that the algorithms rely solely on NWP data.

Line 336-337: but employs different view zenith angles and 337 azimuth angles. -> Not clear. Need more clarification.

Line 338: “matching method” what about parallax corrections between two sensors? It seems some technical details are missing here.

Line 407-412: Did the error statistics consider a similar factor for CTH eval with Seaman et al.'s "within spec" comparisons or just under all cloud conditions?

Line 421: "the CBHs lower than 2 km for"-> Is there any possibility of inversion in the low boundary layers as GFS NWP data may not have such high vertical resolution to resolve and thus CTH errors causing CBH errors in physics algorithms and also NWP input impact on ML algorithms maybe.

Line 477-478: Not necessary to describe all the general lidar observation theory. The paragraph can be trimmed.

Line 493: As the authors addressed, ground lidar observations tend to be quickly attenuated near lowest cloud base especially for thick clouds. If using solely lidar data in comparisons, cloud characteristics (type, depth, etc) related weather conditions would be good to be discussed as well.

Line 527: Line 538 and below details about the radar data should be placed here.

Fig. 1: Additionally, it will be good to mention these comparisons for all cloud conditions including single and multilayer cloud scenes, and something like the CTH accuracy (or evaluation) is not considered.

Figure 2 caption : CBHs -> lowest CBHs, 2017 -> the statistics for all cloud scenes including both single and multilayers? If so, please specify it.

Fig. 3: What happened to the "No CloudSat obs" part on the right end in (b)?

Line 534-538: It's not well organized, which seems like a jump in the context. It should be placed at the beginning of this ground-observation eval section.

Line 648-650: Specify what exactly the factor is.

Line 655: Too early conclusive remark with limited comparisons and without intensive case analyses.

Line 565: "it is more reasonable to opt for physics-based cloud base height algorithms." It seems like a too early conclusive remark with limited comparisons and without intensive case analyses.

Line 670: As well as "nighttime cases", multilayer cases should be mentioned in the conclusion.

Minors:

Line 35: remove 'one'

Line 123: remove "As well known"

Line 134: References in Line 227 should be put here, too, which is the first place for MODIS..

Line 162: the random forest -> RF. The acronym was already defined. Found the same errors in several places.

Line 204: The validation "is" -> has been

Line 335: remove "Global Forecast System" which acronym has been already defined.

Line 375: Not good to use two denotes COT and D_COT in one paper. Please use one consistently.