Comments on "Optimal estimation of cloud properties from thermal infrared observations with a combination of deep learning and radiative transfer simulation" by He Huang et al. (https://doi.org/10.5194/amt-2024-87)

The manuscript proposes a retrieval method to infer cloud optical thickness, cloud effective radius, and cloud top height from satellite thermal-infrared radiance measurements. The retrieval uses a synergetic method that combines optimal estimation (OE) and a convolutional neural network (CNN), and builds on the idea that the CNN provides the a-priori state on which the OE is applied.

The authors validate their retrieval techniques against MODIS products (daytime observations) and DARDAR products (nighttime observations). The authors claim a good agreement between their results and the MODIS or DARDAR retrieval products. However, I am not convinced about the effectiveness of the retrieval, nor the novelty and advantage of the proposed method over the existing MODIS and DARDAR retrievals. First, this may be a consequence of the imprecise and ambiguous presentation of the results. Second, the paper has flaws in the presentation of the algorithm, the technique, and the analysis of the results. Several, easy to avoid typos make the manuscript difficult to understand. It raises the impression that the manuscript was not well prepared.

After re-reading the manuscript several times, I have came to the conclusion that I do not recommend this manuscript for publication. The manuscript should only be considered for a second round of reviews after major revisions. The new manuscript version should explicitly highlight the advantages of the retrieval method compared to existing retrievals. The presentation, from typos over grammar to structure, must be improved. More specific comments are listed below

Major comment:

(1) Line 107-108: The retrieval results are verified with MODIS or DARDAR products. MODIS or DARDAR retrievals themselves are based on assumptions, are maybe biased, and are subject to uncertainties. I am not questioning the quality of the MODIS or DARDAR products, but by using these products as a reference, the proposed retrieval will, in my opinion, never be better than the existing retrievals.
Have you considered using "synthetic observations" based on large-eddy simulations (LES) coupled with radiative transfer (RT) simulations? Such LES and RT simulations would provide ground truth values of COT, CER, CTH, and radiances, and would allow a real evaluation of the proposed retrieval method.

(2) No information is provided about the convolutional neural network (CNN). In particular, the type of CNN, how it is configured, what data are used for training, how many MODIS granules are used during the training, etc? For the manuscript to be useful for the community, the CNN needs more explanation. Otherwise, the retrieval appears to be the typical "black box".

(3) The individual retrieval methods should be explained more clearly to highlight their differences, advantages, and potential disadvantages. In this context, Fig. 1 should be used.

(4) Regarding the general approach. It is not clear to me why CNN and OE need to be combined or what the physical rational supports this idea. CNN and OE are two different methods that could be run independently. Why run a CNN and add an OE approach? The authors state that CNNs are able to learn the complex non-linear relationships between observed radiance and cloud microphysical properties, and that CNN account for the spatial context. One could

question if CNNs are not as good as claimed in the introduction. Does adding OE onto a CNN retrieval really remove the criticism of CNN (Lines 81-85)?

(5) If CNN and OE are combined, would it not be better to go the other way and to improve the OE results with the CNN? This would be a well established approach, where CNN are used to improve existing retrieval methods.

(6) The analysis lacks a clear story that shows which method – OE-IR or OE-CNN-IR – is superior to the other and why.

(7) Line 262: “with those derived from CRTM” It is unclear to me what is happening here. Did you use the MODIS retrieved values of COT, CTH and CER, and put them into CRTM to make radiance simulations? Are these the values plotted on the y-axis with label “MYD06+CRTM” in Fig. 5 a,b,c ? If these are MODIS retrieval results and the resulting radiances / BT, do you have any explanation why the MODIS BT are higher than the simulations? Is there still a difference in the ice optical properties despite of the correction factor of 0.4.

Minor comments:

- Please check for consistent use of the serial comma.

- Please check all units. There are typos throughout the paper, e.g., use km instead of KM

- “Earth” is sometimes capitalized and sometimes not. Please be consistent.

Line 44: Instead of “signifies” do you mean “depends on”?

Line 45: You could abbreviate “cloud effective radius” as it is done before.

Line 55: Citation of Iwabuchi. Please replace uppercase letters with lowercase letters.

Line 63: The word “nears” might be replaced with “approaches”.

Line 64: Please give the wavelength range for your definition of far-infrared.

Line 65: “their limited presence” is misleading. I guess you want to say that these wavelength ranges are rarely measured / satellites are not capable of measuring in these wl ranges. Please consider to rephrase this sentence.

Lines 65-67: Plane-parallel computations are typically computationally efficient and widely applied. Contrarily, three-dimensional simulations are computational expensive. Later in the manuscript you talk about the spatial awareness of CNN for cloud retrievals. You could elaborate on the topic of 3D radiative transfer simulations if they play a role in the thermal infrared.

Lines 75-76: Are you saying that CNN methods are better than traditional look-up-table approaches? I would suggest rephrasing this sentence to better convey the main message.

Line 80: What is FY4A? Please explain.

Lines 81-85: Could you please elaborate on the meaning of “ lacking direct physical interpretation” as well as “offer explicit physical interpretations”. What do you mean by that?

Line 86: [..] integrate […] with. It might be “integrate […] into” or “combine […] with”. Please check your grammar.

Line 90 and elsewhere: Please check weather you mean “priory” or “a-priori”. In OE one usually refers to “a-priori”.

Line 98: Could you please specify “global data”?

Line 98: I suggest replacing “spacecraft” by “satellite”

Lines 100-102: Please check the grammar of this sentence. “For this research, […] for this study.”

Lines 100-101: It would be beneficial to list the MODIS products used, explicitly stating their spatial resolution and the quantities extracted from each product. While you do this in Table 1 it is missing in Lines 100-101.

Lines 103: Change “radiations” to “radiation” ?

Lines 108-109: “All parameters are aligned to a 5-kilometer spatial resolution grid, ensuring data and variable consistency”. Does this mean you have projected all parameters on a regular 5-kilometer grid. How do you deal with parameters at 1km resolution? Do you interpolate them or do you use nearest neighbor method for the selection?

Line 109: Suggestion: Replace “universality” with “applicability” ?

Line 109: What do you mean by “variable consistency”? This is contradictory, you could rephrase the sentence to be clearer.

Line 110: How do you choose a day from each month? By chance or by some specific method or criteria?

Lines 113-114: “By selecting days representative of each month, we aim to assess the algorithm's performance under varying seasonal and weather patterns.” You wrote this 2 sentences before. Please check, rearrange, or delete for more concise writing

Table 1: Could you provide the wavelength ranges for the bands? The unit of cloud water path should be “kg/m²”, using a lower case “k”. Spectral bands itself do not have a unit. They depict a wavelength range. What you mean is radiance instead of Band X? Please correct.

Line124: replace “adept” with “suited to”?

Line 131: Before writing about “our retrieval methods” could you please introduce the models that you are using? You could use Fig.1 to better introduce your general retrieval concept. Fig. 1 is only mentioned once and never explicitly explained.

Line 131: “… including temperature,….” Does this mean that you incorporate all ERA5 variables in your analysis? “including” suggests that you use all of them, not just a few. Please explicitly state which data from ERA5 you are using.

Line 131 / Fig1: After the introduction of Figure1 it would be very beneficial if the authors could explain at least one iteration of the model. Particularly highlighting the difference between the OE-IR and the OE-CNN-IR method.

Line 133: Do you construct look-up-tables based on the ERA5 data, i.e., a hypercube of ERA5 data, or do you construct look-up-tables of simulated radiance based on the ERA5 data?

Line 134 or earlier: When introducing ERA5 please cite Hersbach et al. 2020. The full citation is given at the end of the report.

Line 138: “This method iteratively adjusts parameters to reconcile observed data with model predictions.” Isn’t it the other way around? As I understand OE, the inputs to the forward model are iteratively adjusted such that the model output closely matches the observations. Please check.

Line 148-149: “ and is able to simulate the radiances observed by satellites. ” This has already been mentioned above. Please rephrase the paragraph and remove duplicate text.

Line 150: Suggestion: Replace “corroborated” with “validated”

Lines 150-154: Please check grammar. Please also specify which habit (mixture) you are using in your forward simulations.

Line 155: “various COT,..” do you mean “combinations of”. Please check and change if needed.

Table2: This table is difficult to understand because the column “Notes” mixes various information, i.e., products and parameter steps. You could also choose a better way to present the COT, CER, and CTH ranges, e.g., by giving intervals and step sizes. Please add spaces between the number and the unit (and check for this in the entire manuscript).
Please revise the table and make it more clear. Otherwise I suggest to remove the table.

Line 162: “observations in” to “observations of”

Lines 165ff: Please use sub figure labels [(a), (b),...] to guide the reader through the Figure. In that way the reader is directed to the correct sub panel, which facilitates the understanding of the text.

Line 173: In several instances the authors switch between “liquid clouds” and “water clouds”. Please revise the entire text to be consistent. All clouds consist of water. To be precise one should distinguish between “ice water” and “liquid water”.

Figures in general: Captions are not in bold. Only the figure number and the sub panel label, e.g., (a), is in bold.

Fig2: You changed the y-range of Fig2b but kept it constant everywhere else. It would be beneficial to keep the same range (220-280 K) for all panels. This helps to compare with Fig.3

Line 193: “standard MYD06 products” Please explicitly state which variables you are using. Or do you use all variables provided in the product?

Line 196: What is meant with “convolutions”? Please explain.

Lines 209-210: “model predictions” is ambiguous. Do you mean simulation results from the forward model?

Line 212: Suggestion: Replace “proficiency” with “capability” or “ability”.

Line 217: Equation 2: Typo. Xa should be X_a, I guess?

Line 221: Suggestion: Replace “encapsulated” by “implemented”. And it is more a gradient decent method instead of a “gradient decent iterative process”. You might rephrase the sentence.

Line 226: “ and is set to be the same for all three variables” Do you refer to the learning rate. If so please rephrase the text.

Line 242: If the results of the OE-CNN-IR remain at the a priori values, does this not mean that the iteration fails? It means that the information content of the measurements is not taken into account.?

Line 244: Please elaborate what you mean with “function effectively.”

Line 244: How do you know that the COT is correctly determined? You minimize the cost function between the observed and simulated radiances but this does not necessarily mean that the retrieved COT is correct. If you want to prove a correct COT retrieval you should plot the true / expected COT, CER, and CTH values together with your retrieval results in Fig4.

Fig4. Please add the COT values associated with “A thin cloud case” and “ A thick cloud case”. What do you mean by “thin” and “thick”? Optically or geometrically. Please specify.

Line 249: A feature of OE is the estimation of the errors that are associated with the a priori and a posteriori information, i.e., the uncertainties introduced by the measurements. In my opinion, this should be included, at least mentioned, in the section of OE. If you do not plan to use the OE uncertainty estimates from the OE technique please state why.

Line250: “.. forecast errors..” do you mean errors in the forward simulation. You are not making a forecast in the traditional sense.

Line 259: It is unclear which method you are referring to: OE-CNN-IR or OE-IR.

Line 261ff: Is there a particular reason, why you chose this particular MODIS granule?

Line 261: It would be more convenient to start with the figure that is currently Fig.6 and then showing Fig5. In this way the order would be more logical by showing the spatial distribution first and then the correlation.

Line 262. You are jumping between figures. Starting with 6, going to 5 and going back to 6 again. Please rearrange the entire section 3.1 and the figures.

Lines 273-275: Do you want to say that OE-IR is better than OE-CNN-IR? This would contradict your argumentation that the combination of CNN and OE is beneficial. Please check.

Lines 292-294: This is a trivial statement. Radiances or converted brightness temperatures are used for cloud property retrieval. This is the basis of all cloud property retrievals, e.g., of the MODIS retrieval that you use as a reference.

Lines 299-301: Please explain where the substantial differences come from.

Line 202: “below a certain threshold”. Please explicitly mention the threshold and give a number.

Figure6: The figure is difficult to read. Images and labels are too small, please enlarge.

Line 343: Please explain how selecting data between 60N and 60 ensures “consistency and reliability in these comparisons”.

Line 346: what do you mean with “all clouds”: ice and liquid water clouds? The caption of Fig 8 says ice clouds but over land and ocean. Please be more specific and rephrase.

Line 357: “ to be concentrated around the a priori value of 30 μm,” does this not mean that the OE retrieval is too much constrained by the a priori value and does not take the measurements into account. This would be an indication that the retrieval does not work for this value.

Line 362: “all types of clouds”. Please be more specific: ice, liquid water over land or ocean?

Line 368: “effective estimation of priori states” Does this not mean that you are getting the retrieved values from ERA5? Then why use satellite data?

Line 390: How many samples do you get from 2009? Please specify.

Line 393: Please state what you mean with “”meaningful assessment”.

Lines 399-400: “notably in the context of the challenges involved in accurately retrieving COT for ice clouds.” This contradicts your statement in Lines 365-366 where you argue that OE-CNN-IR is well suited to retrieve optically thin ice clouds.

Line 364: What does “thick water clouds” mean? Do you mean optically thick or geometrically thick?

References

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.