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.

Anonymous reviewer
27 March 2024

In this study, the authors propose a synergistic method to retrieve cloud properties from infrared radiometric measurements by Moderate Resolution Imaging Spectroradiometer (MODIS). This method is based on the idea that the retrievals obtained by a machine learning algorithm are used "*to provide the priori state*" for the optimal estimation algorithm. The authors claim that "*the cloud properties retrieved by the new algorithm show an overall better performance by optimizing the initial values*". In the sentences above, I presented direct citations from the abstract of the paper in order just to pay attention to the importance of using rigorous terminology. In my opinion, it is necessary to distinguish between a priori information and an initial guess because the choice of an initial guess can only speed up an iteration process of the optimal estimation method but can not affect retrieval results. Also, I would like to note that the term "*better performance*" is vague and can be misleading for potential readers. Obviously, the final goal of any retrieval algorithm is not to optimize the initial values, but to obtain the information products with highest possible accuracy. Well, below I will return to this discussion once again.

After careful reading of the manuscript I came to the conclusion that I can not recommend this paper for publication due to the fact that, first of all, the main idea of the study seems to me very questionable. Second, the paper has flaws in terminology, formulation of the problem, and analysis of the obtained results. I point at all these flaws below in Section "General critical issues". Minor comments are given in Section "Specific remarks".

General critical issues:

My major concern refers to the question that remains without an answer in the manuscript: where from the additional information is taken by the neural network which might help to improve the retrieval results for cloud properties? The authors do not even mention in the manuscript any training procedure for CNN. By the way, I have the general recommendation to the authors to explain every acronym when it is used for the first time. CNN is not explained in the text. Does it mean Convolutional Neural Networks? How was it trained? What is the physical basis for using the CNN output instead of using values from a priori statistics as it is done in a classical optimal estimation method?

My second concern refers to terminology and formulation of the problem. The authors present the cost function of the classical optimal estimation method as Eq. 3. However, after that they write

> *"Then we employ an iterative process to find an optimal solution based on observed data and prior information."*

and present Eq. 4. which in fact describes the Levenberg–Marquardt method. I do not see a connection between Eq. 3 and Eq. 4. In Eq. 4, the a priori information in the form of a priori covariance matrix is missing and regularisation is provided by the damping parameter only. As a result, the obtained results can not be called as obtained by the optimal estimation method. By the way, in Eq. 4, the transposition sign is missing at $K_{F/Xi}$ before the square brackets. The authors after that write:

'During each iteration, the cost function is evaluated."

If the authors imply the cost function described by Eq. 3, then I see here an evident inconsistency, because Eqs. (3) and (4) do not correspond to each other. I have an impression that the authors mix up the Levenberg–Marquardt method and the classical optimal estimation method and such mixture can be misleading for potential readers.

Figure 1 does not help much to understand the formulation of the problem and the structure and logic of the retrieval algorithm. How are the Jacobians calculated for initial guess and recalculated at the iteration steps? How are the LUTs connected with the calculation of Jacobians? Why do we see only one-way arrow connecting parts (a) and (b) of the Figure? Does it mean that the Jacobians are recalculated without using LUTs?

I would like to discuss once again the difference between the a priori information and the initial guess. In the classical optimal estimation method, the a priori information in the form of mean value and the covariance matrix of the state vector is used for the regularisation of the so-called ill-posed (or improperly posed) inverse problem and therefore provides the unique solution. It is this information that makes the retrieval "optimal". In order to eliminate the errors caused by nonlinearities, the iteration procedure can be applied. In this iteration procedure, one can choose any initial guess, i.e. starting point for iterations. The successful choice of the initial guess can help to reduce the number of iterations but has no influence on the solution which is the result of contributions from measurements and a priori information. Obviously, one can take the a priori mean state vector as the initial guess. In this case the terms "a priori information" and "initial guess" would refer to the same state vector, but in general these terms describe different state vectors. The authors in their manuscript use both terms without rigorous clarifying the meaning of each of them. The above mentioned mixture of the optimal estimation and the Levenberg–Marquardt algorithms makes the description of the retrieval procedure in the manuscript very unclear. Also, I would like to notice that nothing is said in the manuscript about the convergence of the iterative process.

I have questions also to the test retrievals and analysis of the retrieval results presented by the authors. I do not quite understand the idea to take the results of the standard MODIS algorithm as a reference. In the ideal case, even if the results of the proposed algorithm and the results of the standard MODIS algorithm are the same, the conclusion about the superiority of the proposed algorithm could not be made. Then, what is the purpose of such comparison? Of course, the authors also present the comparison with the independent data (DARDAR), however the information about the quality of independent data is missing, and taking these data as a reference is not justified.

When the authors analyse the results, they repeatedly claim that the proposed method has advantages with respect to other considered methods. In my opinion, these advantages are exaggerated. I do not see any considerable improvement in retrievals provided by the proposed method. It should be also mentioned that the authors do not present clear descriptions of other methods which they use for comparison: TIR-CNN, OE-IR.

Summarising mentioned above concerns, I would like to notice that the manuscript makes an impression of a study that tries to apply the neural network approach in a so-called "mechanical" manner: without paying enough attention to analysis of the information content of the "additional" data in the form of the CNN results and without the justification of the physical basis for the improvement of retrieval results provided by IR observations.

Specific remarks:

Line 10
What is meant by the term "effectively"? Please specify.

Lines 64-65
The reference is missing in the list of references. May be, this reference is the following:
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016JD026423
If so, please specify that the referenced study considers only ice clouds but not all clouds with "*substantial optical thickness*".

Lines 83-85
I can not just understand what the authors want to say in these lines. In my opinion, it is evident that any used approach (either physical, or machine learning) should have physical basis in the form of information content of measurements.

Lines 86-88
I think that it is necessary to specify that the method is developed for ice clouds.

Line 90
Please, take into account the difference between the a priori state and the initial state for iterative process. I can not understand what "*iterative optimisation*" means. What namely is optimised?

Line 93
Please explain what is meant by "*passive and active*" cloud products.

Lines 111-112
What geographical regions are taken? Please specify.

Line 122
Were the products of the original MODIS algorithm validated by DARDAR?

Line 134 and Fig. 1
As it follows from Figure 1, TIR-CNN provides not initial estimate but a priori state.

Line 147
Typo: simulate

Line 150
Typo: provides

Table 2
Typo: Geometries

Line 160
Typo: m x n

Eq. 2
This equation is unclear since *K* should be a matrix, but we see a vector here.

Lines 171-172
This fact is evident and known.

Caption to Fig. 2
Why were the presented geographical coordinates selected for retrievals?

Figure 3
The results presented in Figure 3 are obvious. I think that there is no need to show them.

Lines 196-197
I would say that this is not the additional but the main feature of the method.

Line 200
Typo: S_a

Figure 5
It is hard to make any conclusions from Figure 5.

Lines 294-295
What is the algorithm to estimate vertical extent of clouds from the top height? Please describe where from the cloud bottom height is taken.

Lines 326-329
I see no physical basis for the improvement. Where from does the neural network get the information on COT?

Lines 332-333
It is not surprising since BTs are informative with respect to CTH.

Lines 334-336
This statement is trivial and it does not justify the approach that is based on the integration of CNN into the retrieval process.

Line 370
This conclusion does not seem to be justified. The correlation coefficients are very low and do not indicate sufficient accuracy.

Lines 380-381
This conclusion does not seem to be justified.

Lines 383-384
I cannot understand the sentence "Subsequently…". The initial guess can not be adjusted. It is the solution that is adjusted in the iterative process.

Line 395:
The statement "Prospects…" was not justified in the study.

Lines 401-403
The statement "In the future…" is vague and can be misleading for potential readers.