

## Reply to anonymous referee 1

We thank the reviewer for the time and efforts she/he spent reading our manuscript carefully and providing valuable comments and suggestions. In the revised manuscript, we have tried to accommodate all the suggested changes. Please find below our responses. The reviewer's comments are in italic. Changes/additions made to the text are given in quotes.

*This article presents the possibility to use of deep neural network (DNN) to retrieve cloud properties (optical thickness and effective radius) accounting for the horizontal photon transport. This is so far not accounted for in the classical algorithm, that use the homogeneous cloud assumption. This is a move in the way to improve remote sensing algorithm and account for 3D radiative effects. However, the presentation and explanations do not put the paper in favor. Consequently, several precisions and corrections need to be added in the paper before publications. They are indicated below.*

### *Major comments*

*1) The originality of the paper seems to be more related to the possibility to make a multi-pixel inversion of cloud properties than to use DNN. It should appear in the title. I suggest “Feasibility study of multi-pixels retrieval of cloud optical thickness and effective radius using deep neural network”*

**Response:** We appreciate the suggestion. Indeed, this is a feasibility study of such a multi-pixel inversion. It is also important that the deep learning techniques are applied to retrieval of inhomogeneous clouds. Thus, by combining the reviewer's suggestion and our points, we have changed the title as

“Feasibility study of multipixel retrieval of optical thickness and droplet effective radius of inhomogeneous clouds using deep learning”

*2) Abstract is too succinct and need to be completed.*

**Response:** We have rewritten the abstract, adding explanations about the two DNNs

*3) In the introduction, the authors described similar works, they did previously to retrieve COT and CEDR accounting for neighboring pixels (Iwabuchi et Hayasaka, 2003). Through the paper, the disadvantages of this previous method comparing to the new one are not sufficiently explained. I did not understand “which was an obstacle to generalizing the algorithm (p2, li10)”. Which obstacles? Does it not the same problem*

*with the NN method ? The authors should add a discussion about the advantages/disadvantages and about the implementation of each method in the introduction or in the conclusion A comparison of previous method and DNN method in terms of results will also be valuable.*

**Response:** We have rewritten explanations about limitations in the previous study by Iwabuchi and Hayasaka (2003), as follows:

"Since 3D radiative effects vary with COT, CDER, cloud geometrical thickness, cloud top roughness, and sun–cloud–satellite geometry, Iwabuchi and Hayasaka (2003) had to construct different sets of fitting coefficients, thus limiting the applicability of the technique in practice. In addition, their method was based on linear regression, which is not flexible to account for any nonlinear 3D radiative transfer effects."

A comparison with the previous method in Iwabuchi et Hayasaka (2003) is technically difficult. One reason is that the previous method was developed for a feasibility study assuming a particular type of cloud model (Fourier transform cloud), which is quite different from that in the present study. In addition, the assumed cloud inhomogeneity was in one horizontal dimension while the present study treats two horizontal dimensions.

We have added an explanation of one of advantage of DNN, ease of addition of input and output variables, as follows:

" In addition, input and output parameters can easily be added, and structures can be modified - another advantage of deep learning"

*4) li 15-22: Some important references are incorrectly cited: Faure et al. (2001) is about the retrieval of mean cloud properties accounting for the sub-pixel heterogeneities while Faure et al. (2002) concerns the retrieval of cloud parameters from high-resolution data using adjacent pixels which is a different study. This second one is the closest to the current study. Correct it also in section 4.2. These two papers are major because they are the first papers in the fields but they are limited to fluxes and not applicable to real data. Following the paper of Faure et al., 2001, which is for medium spatial resolution, Cornet et al. (2004) present ways to apply to real data heterogeneous cloud retrieval using NN. It is finally tested on real data in Cornet et al. (2005) on MODIS data. The paragraph citing these studies about cloud neural network retrieval needs to be clarified.*

**Response:** As the reviewer requested, we revised Section 1 and clarified the description about previous studies. The main discussion in Faure et al. (2001) was the application of NN to retrieval of mean cloud properties, taking sub-pixel inhomogeneity into account, but they also

tested a NN that uses the radiances of target pixel (0.8km x 0.8km region) and eight adjacent pixels (0.8km x 0.8km regions) for retrievals, showing a promising result. Our DNN method is compared to the latter one in the Faure et al.'s (2001) study. We have added descriptions about this point in the Section 4.2, as follows:

" Originally, this NN had two hidden layers with 10 units each and it used the  $0.8 \text{ km} \times 0.8 \text{ km}$  area-averaged radiances at four wavelengths (0.64, 1.6, 2.2,  $3.7 \mu\text{m}$ ) at the target pixel and eight adjacent pixels (called as "with ancillary data"). It is described in the section 3.3 (2) in the original paper."

Faure et al. (2002) uses high resolution but one-dimensional cloud data, and we cannot directly compare our results with that of Faure et al. (2002). We have added the reference Cornet et al. (2005) in Section 1.

*5) To my knowledge, this is the first time in atmospheric science that Deep Neural Network (DNNs) are used. More explanations are needed in a specific section explaining clearly how it works and allowing to understand some affirmations and vocabulary used in the text. For example, in the introduction, why "a DNN is more suitable for approximating complex non linear functions" than a classical NN? What is "automatic feature extraction", can the authors give an example? For the same reasons, Section 3.1 are confuse and consequently not very clear for a non-expert in deep learning. It needs to be separate with generality on the DNN in the specific section rewritten with more explanation and in a pedagogical way. Some schemas may also help to understand. Another section should specify to the choices made (see major comment 6 below). In the specific section about DNN should appears what is "shortcuts DNN"(p5, li 17) or what is convolutional layer ? How the filter weights are obtained? Can the authors also explain in few lines the paper of He et al. (2015) in order that the readers understand?*

**Response:** We appreciate this comment. We have moved explanations on fundamental DNN techniques to Section 3.1. As the reviewer commented, some important explanations were missing in the previous manuscript. So we have added some more explanations (ex. "a DNN is more suitable for approximating complex problem", "automatic feature extraction", and "shortcuts"). We hope the revised manuscript is easier to read.

We have not added very long explanations on the techniques that are really technical and not essential to the conclusions of this paper. We did discuss essential characteristics of each deep learning technique in the current manuscript. We think it is better to leave the technical details of each optimization and deep learning technique for readers to consult the textbooks or references cited in the current manuscript. Indeed, the deep learning is applied to a remote sensing problem of atmospheric target for the first time, to our best knowledge. However, it is a rapidly growing technique in broad areas of sciences, upon many successes in engineering and

applications for the artificial intelligence. In the atmospheric sciences, too, we know at least a few research groups working on applications of deep learning techniques. Nowadays, it is easy to find a book for practical use of this technology even for undergraduate students.

*6) p5 . There is also no enough explanation about the choice of the input vector and the architecture of the DNN. Li- 5-10: why these two input vectors? The paragraph should start with an explanation of the philosophy. The first input vector is built in order to correct IPA retrieval and the second to retrieve directly cloud properties. I'm wondering also why four wavelengths and not only two as for the bi-spectral method? Does the authors test this last configuration with two wavelengths? I'm wondering also how the architecture of the DNN was chosen (convolutional layer or not, activation function or not), does the authors made test to find the best architecture?*

**Response:** As the reviewer commented, we tested numbers of different DNNs with different number of wavelengths, the use of convolutional layer, different activation function, and so on. The best two DNNs are shown in this paper. We have added the description about this in the manuscript, as follows:

"The above two DNN structures were obtained from various trial-and-error experiments. Different DNN structures were also tested. For example, we tested a DNN similar to DNN-2r but with four wavelengths, and one similar to DNN-4w but with only two wavelengths. However, DNN-2r and DNN-4w performed best. There is room for improvement in DNN structures, which should be investigated in the future."

*7) P7- li 28-30: I am not completely agree with the assertion "the DNN retrieves COT values that are close to the true values assumed in the test, successfully corrected the phase lag." In Figure 6-a, near 11.7km, DNN-2r retrieval shows also large differences and near 17km clearly the DNN retrievals overestimate the COT and the phase lag is not completely cancelled. Can the authors be more precise in the description of the figures? In addition to cross-sections, could also the authors add the relative errors transects and the RMSE of the different retrievals to have more qualitative idea of the improvements. Same remarks concerning Re retrieval.*

**Response:** We have rewritten the sentence as

"Compared to the IPA retrieval, the DNNs retrieve COT values that are closer to the true values assumed in this test. The DNN-4w successfully corrects the phase lag as shown in Fig. 6c."

As the reviewer suggested, we have added figures of cross-sections of the relative errors and a table of RMSEs of the different retrievals in Figure 6.

*8) p8, li 1-6: Concerning Re retrieval with the homogenous cloud assumption, it is not*

*really surprising to obtain large differences between homogeneous assumption and true results. The overestimation with IPA is not only related with shadowing effects. Indeed, homogeneous cloud assumption involve homogeneous  $Re$  profile. From satellite remote sensing, the upper part of the cloud is retrieved (See for example Platnick et al. 2000). Therefore, if the effective radius is vertically increasing in the cloud, the retrieved  $Re$  is larger than the mean  $Re$ . For the DNN, training with heterogeneous clouds allows to learn the relation between vertically averaged  $Re$  and radiances. Discussion about this issue (p8, li 20-24) is too late in the paper and should be moved here. Try also to highlight better the shadowing by reporting for example the difference between true COT and homogeneous COT as in Cornet et al. (2015) or Marshak et al., (2006).*

**Response:** Figures 5 and 6 are good for demonstrating spatial characteristics of retrieval errors, but the mean bias error for all test data are shown in Fig. 7. Differences between vertically homogeneous and inhomogeneous assumptions are well related to the mean bias in  $Re$ . We think it is better to see the discussion about the issue as in the current manuscript. We have modified the sentences as follows:

"IPA-retrieved CDER thus tends to be larger than column-mean CDER, whereas DNNs are by design trained to learn the relationship between the column-mean CDER and radiances by taking into account the vertical inhomogeneity. However, this vertical inhomogeneity effect on the IPA-retrieved CDER does not seem to be the main cause of the large positive bias."

The issue of difference between true COT and homogeneous COT as in Cornet et al. (2015) should be important, but the sub-pixel horizontal inhomogeneity is not considered in this paper.

*9) P9, section 4.2: Authors made comparisons with previous works of Faure et al., 2002 but the settings are exactly the same. First, only pairs of wavelengths were used and not the four wavelengths mentioned. In addition, in the study of Faure et al., (2002), 15 neighboring pixels of each side of the target pixels were used (62 components in the input vector) and here only 3. This can change a lot the results. The comparisons have to be done again with the same parameters than the one used in Faure et al. (2002) or at least the same conditions that the DNN, that is 10 pixels for each side, otherwise, it is not possible to conclude the comparisons and to know really why retrieval is better (DNN or neighboring pixels?)*

**Response:** As explained in the reply to the comment 4, we compared our results with that of Faure et al. (2001).

*Minor comments*

*1) p. 1, li 17: why the bispectral method follows the IPA assumption? The authors should add reasons why in the text (time computation, simplicity, others?) and also insist about the independence of each cloudy columns which is considered infinite.*

**Response:** We have rewritten the sentences as follows:

" The method is based on the independent pixel approximation (IPA) assuming plane- parallel, horizontally and vertically homogeneous cloud for each pixel of the satellite image because of high computational cost for simulation of three-dimensional (3D) radiative transfer. The observed cloud radiances result from three-dimensional (3D) radiative transfer in the cloud field, ..."

*2) p.2, li 5: Until which distance, have the neighboring pixels to be considered ? can the authors here or further in the text give some values and references ?*

**Response:** We have added an explanation about the scales, as follows:

"The 3D radiative effects operate on horizontal scales that are determined mainly by cloud thickness and solar zenith angle. When the sun is oblique (i.e., with a solar zenith angle of 60° or larger), the maximum horizontal scale for 3D radiative effects is roughly 15–20 times larger than cloud thickness (Marshak and Davis, 2005).

*3) p3, section 2.1: what are the resolution and dimension of the generated cloud fields?*

**Response:** We added a description about the resolution of the generated cloud fields as "The resolution for the x- and y-axis of the cloud field is originally 35 m. For the z-axis, the resolution is 5 m at the bottom of the atmosphere, and it is coarse (less than 60 m) for the upper layers."

The dimension sizes are written in the manuscript.

*4) p3, li 17: IPA (Independent Pixel Approximation) does not mean that the vertical profiles is homogeneous but only that each pixel is considered independently of his neighbors. Authors should speak about the homogeneous cloud assumption horizontally as well as vertically*

**Response:** The IPA retrieval usually assumes vertically and horizontally homogeneous cloud in each pixel. We modified the introduction part as "the independent pixel approximation (IPA) assuming plane-parallel, horizontally and vertically homogeneous cloud for each pixel of the satellite image".

*5) p3, eq. 3: Why the authors used the square of the usual definition of the inhomogeneity parameter defined in the others study. For comparisons, it seems to be better to use the same definition.*

**Response:** This is because the 3D radiative effects (e.g. 3D minus IPA radiances) are approximately linear to the square of the usual definition of the inhomogeneity. In this paper, we tried to show that the open cell cloud was much more inhomogeneous.

*6) P5, li 9: Radiances at 3.75micron is used, I suppose that is only the solar part. It should be precise in the text that thermal correction need to be done before using this wavelength.*

**Response:** We have added an explanation about this, as follows:

"It should be noted that only solar radiation is considered in the present study, which requires that the thermal radiation at 3.75  $\mu\text{m}$  be corrected during pre-processing."

*7) p5: explain why the number of pixels considered in the input vector (10x10) is larger than those considered in the output vector (8x8 or 6x6) and why it is not the same for the two DNN. 8) P6, li10: add the URL for the chainer framework*

**Response:** We have added a sentence for the reason as follows:

"The reason for including margin pixels in the input field is to take into account the 3D radiative effects from the surroundings of the cloud field. The DNN-2r network consists of several fully connected layers."

*9) P6, eq 8 and9: is there a justification for the choice of these functions?*

**Response:** The functions in these equations are less nonlinear to radiances. This kind of simple transform helps better performance of DNN retrievals. We have added an introductory sentence as follows:

"Because the radiances are highly nonlinear with respect to the COT and CDER, it is convenient to transform the COT and CDER by some simple functions. "

*10) P7, li 14: DNN and IPA are not really comparable: the first is an inversion tools as look-up tables and the second one is a direct model. It seems better to write multipixelsDNN inversion versus IPA-LUT inversion.*

**Response:** We agree, but since we described our terms, we prefer to stick with the terms of "DNN and IPA retrievals". We modified the manuscript to use "IPA retrieval" in a consistent

way.

*11) Figure 5 and 6: Precise data corresponds to only the test data set or to a mix between the training and test dataset*

**Response:** These are for test data.

*12) p7, li 22 and Figure 5 and 6: precise the geometry of the observation: view zenithal and azimuthal angles?*

**Response:** We added the geometry information to Figure 5 and 6, and in the manuscript.

*13) p7: li 26: illumination and shadowed effects are well-known under IPA assumption: please add some references*

**Response:** Várnai and Davies (1999), Várnai and Marshak (2002), and Marshak et al. (2006) are added as references about illuminating and shadowing effects.

*14) p7: li 26: Large errors are due to the flattening of the relation-ship between radiances and COT due to saturation effects: a small difference in radiance lead to a quite large difference in COT*

**Response:** We appreciate this comment and have added a sentence as "This can be ascribed to weaker sensitivity of radiance to COT when COT is larger; a small difference in radiance leads to large difference in retrieved COT."

*15) p8, li 9, figure 7 : I agree that the bias (mean error) is particularly large only for COT less than 1. For COT > 1, the difference in errors is not so important. For COT>10, the standard deviation is larger for IPA meaning that dispersion (roughening) is more important.*

**Response:** Large IPA error dispersion for COT > 10 are shown for SZA of 60 degrees, which suggests that the roughening cause the error. This point is mentioned in the manuscript.

*16) Figure 8: Could the authors indicate the COT and Re associated with the filters and be more precise in the description of the figure? Which filters patterns are "symmetrical around the center" and how is distributed the optical thickness? Also on which figures does appear "the feature related to the solar direct beam"? Comments also the difference between wavelengths.*



**Response:** In general, this kind of analysis of NN coefficients is difficult to understand when the NN becomes deeper. To study the associations of these filters with COT and CDER, we should investigate the coefficients in the DNN-4w layers subsequent to the convolutional layers (as shown in Fig. 4). It would be hard to interpret the relationships between the two destination fully connected layers in the DNN-4w. In this paper, we just tried to check that the DNN-4w indeed learned meaningful patterns of 3D radiative transfer.

As suggested by the reviewer, we have added a few explanations about specific pattern shown in Fig. 8. We have also added an introductory note as follows:

"It is noted that the convolutional layer is designed to correct the 3D radiative effects appeared in radiances (Fig. 4)."

*17) Section 4.2: are the same training set and generalization set used for all the NN trainings?*

**Response:** Yes, the original dataset are the same as those used for DNNs. We have added a sentence as "Data used for training and test for the NN are from the same original datasets used for DNNs."

*18) p9, li 30-33: add the issue concerning the vertical profiles for Re in the conclusion.*

**Response:** As shown in Section 4.1, the vertical profile for Re seems to be not a main reason for the IPA retrieval bias in Re, at least, in the present cases. We think there is no strong need to mention this point in the conclusion.

*19) P10: in the conclusion, can the authors insist on the limitations of using NN methods such as the one related to database used and extrapolation issues. In other words, how will work the DNN is the cloud is quite different to those used for the training dataset?*

**Response:** It is a well-known issue of NN as the reviewer suggested. Training of NN should includes enough variety of realistic cloud. This study is just a feasibility study, but practical applications definitely require training for various types of cloud. We have added a few words as "...will require training using realistic cloud fields for various types of cloud."

*20) P10: Following the previous points, can the authors speaks about the steps needed in order to develop an operational multi-pixels algorithm?*

**Response:** An expansion of cloud variety is one step. An appropriate DNN architecture for addition of input parameters should be investigated. For example, the sun-cloud-satellite geometry parameters are very different from the radiance image data that are used as input data in this study. There should be appropriate DNN structure to add the geometrical parameters. We have added the last sentence in the conclusion, as "An appropriate DNN architecture for addition of input parameters should be investigated in the future."