Response to Referee #1

In the following, our responses are in black, reviewer comments in *blue italics.*

Review of the paper "A neural network-based method for generating synthetic 1.6 µm near-infrared satellite images" by Florian Baur et al., MS No.: egusphere-2023-353

This paper focuses on the development of a neural network to estimate the reflectance emerging from the atmosphere at 1.6 micron from SEVIRI on Meteosat Second Generation. This approach should also be suitable for other near-infrared channels on instruments such as AHI, ABI or FCI. This is a nicely written paper discusses the neural network performance when trained on reflectances calculated using DOM with both IFS and ICON-D input profiles.

I believe that this paper would be suitable for publication subject to minor revision addressing the specific comments detailed below.

We are pleased to hear that and want to thank the reviewer for the constructive and helpful comments, which will try to address below.

Specific comments

L37: here you should explain what you mean by 3D effects, e.g. reflection on complex topography etc. A reference could also be useful.

We agree that an explanation should be added here, e.g. something like "By 3D effects we mean effects not taken into account in the plane-parallel RT equations, e.g. everything related to horizontal photon fluxes. One of the most important effects is the impact of inclined cloud tops on reflectance, for which a fast approximation was developed in Scheck et al. 2018. Other effects are related to cloud shadows, complex topography or photon transport through the cloud sides. A full discussion can be found in *Marshak & Davis (2005): 3D Radiative Transfer in Cloudy Atmospheres*".

L50: here do you mean little information on discriminating the cloud phase?

Yes. We will clarify that in the revised version.

L66: It would be really interesting if you could show here an example of the water and ice cloud jacobians for the 1.6 micron channel, possibly by comparing to those for the visible (and/or thermal IR) channels.

The derivatives with respect to the cloud variables depend strongly on the cloud structure. It is not obvious for which cloud profiles it would be interesting to see the Jacobians. Moreover, Jacobians are often complex and there are many different aspects that could be discussed. We think this information would be more relevant for a publication about the assimilation of near-infrared channels with a variational or hybrid assimilation method and we would thus prefer to postpone the discussion of Jacobians to a future study. As we plan also to add two more figures and several paragraphs of text in response to other comments of the two reviewers, we think the additional topic of the Jacobians would make the paper too long. What we would like to do, however, is to mention in the outlook that performing variational assimilation tests including a discussion of the Jacobians of the neural network should be one of the next steps.

L84: How do you assess that the statement "are not very important" is true? Please add a reference to a paper where this is discussed and/or add a few explanations.

We cannot find such a statement in line 84. You probably refer to the sentence "MFASIS makes use of the fact that for non-absorbing visible channels, the cloud top height and details of the vertical cloud structure are not very important." in lines 88/89. In MFASIS the cloud top height and details of the vertical cloud structure are not taken into account for computing the reflectance, and still the reflectance errors with respect to a reference solution are small. In this sense, these properties of the input profiles can be considered to be "not very important". We will clarify this in the revised version.

L97: Do you mean interpolate the reflectance for the specific value of the albedo at a given location? If so you should expand your text here as you discuss this interpolation only later in the paper.

Here we mean the linear interpolation in the optical depth, effective radii and angle dimensions of the LUT. In fact, interpolation in the albedo dimension is not necessary and we can exactly compute the reflectance as a function of the albedo, as we discuss later in the manuscript. We will clarify this in the revised version.

L133: How many hidden layers, and did you test the effects of having more or fewer layers? how did you initialize the weights?

We used 8 hidden layers and 15, 25 or 32 nodes per layer. We did some tests with fewer or more hidden layers (but roughly the same total number of parameters in the network) that did not show relevant differences. A more detailed discussion of this question can be found in Scheck 2021, where it was found that networks with between 4 and 8 hidden layers gave similar results. The weights were initialized using random numbers. We will add this information to the revised manuscript.

L136: What distribution did you use for the random numbers? If uniform, which intervals?

We used uniform random numbers between 0 and 1 for the normalized input parameters. The normalization was performed using the parameter ranges given by Table 2, as we will discuss more clearly in the revised version.

L138: How many epochs were used in the training?

We trained the networks for 4000 epochs, as specified in line 406. We can provide this information already in 2.1.3.

Fig 2 caption: Please specify units of effective particle radius

We will add the unit micrometers in the revised version.

L234 ("exceeds a threshold value of 1"): Is this ok also for low optical depth clouds? From Fig 3 there are quite a few profiles with log(tau_w) ~ 0.1. Did you test having a threshold dependent on cloud optical depth categories? And did you test the radiative effects of the use of different thresholds?

For low optical depth clouds the vertical cloud structure should not be very important any more, as reflectance is dominated by single-scattering and for this case it does not matter at which height the scattering process takes place. We tested different values for the threshold and found that smaller threshold values did not significantly improve the results.

L235 ("where z_sfc is the height…"): do you mean here the height of the highest model level below the bottom cloud layer?

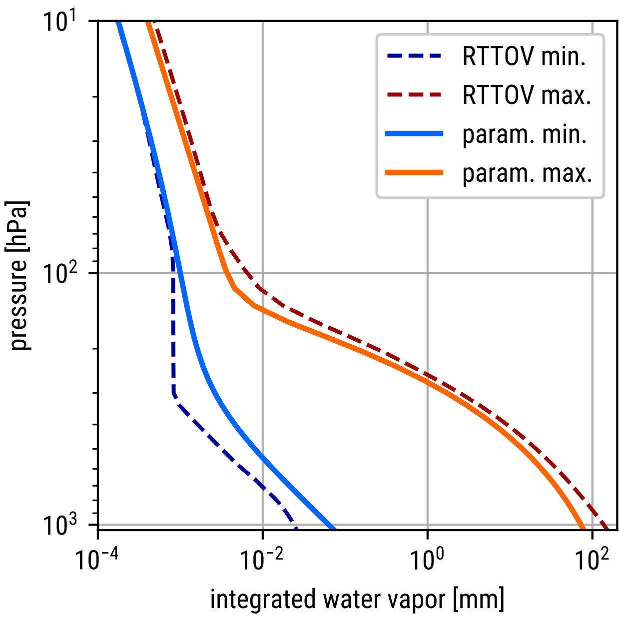
For the sake of simplicity, the lower bound $z^{l,bot}_{w}$ of the integral in Eq. 3 is set to the surface. We could provide a better definition for the bottom of the cloud layer, but this is just not necessary – if we integrate over cloud-free levels below the bottom cloud layer this will not change the integral. We will make this clear in the revised version.

L244 ("exceeds a threshold value…"): I guess this at least partially answers my previous question. But I don't understand if this is done also for cth or only for ctp. And did you test other values (i.e. tau_t/5 etc.)

The cloud top is defined as the level where the cloud optical depth exceeds the value given in the manuscript and this level defines also cloud top pressure and cloud top height.
The aim of the cloud top detection is to quantify the air mass above the cloud, as this quantity determines how strongly absorption by water vapor and trace gases influences the reflectance. We tested different threshold values. We found that the threshold proposed in the manuscript is a good compromise to prevent on the one hand that very thin, high clouds above thicker clouds trigger the cloud top detection and to avoid on the other hand that the cloud top is detected too deep inside the cloud.

L255: It is not clear how these parameterizations are consistent with minimum and maximum values of IWP accepted by RTTOV

We agree that this point was not discussed in sufficient detail. We will add the plot displayed below, which shows the minimum and maximum allowed IWV values above the pressure level given on the y-axis. The dashed lines are integrals over the hard humidity limits imposed by RTTOV (Table 1 in the RTTOV 13.2 user guide). The parameterizations (solid curves) lie between these limits. The focus was on using smooth, differentiable, not overly complicated functions rather than following the hard limits as closely as possible.



L262: If I understand correctly, please replace “full vertical profiles” with “full idealised vertical profiles

Correct, we will add “idealised”.

Fig 12 caption (“8 hidden layers with 25 nodes”): These numbers are inconsistent with those in the figure. Please correct the caption or the figure. Also, please check the total number of weights and biases is as stated.

This is indeed inconsistent and will be changed in the revised version.