This work (GPROF-NN: A neural network based implementation of the Goddard Profiling Algorithm) presented a retrieval algorithm for two passive microwave sensors (GMI and MHS). The paper is pretty well written and easy to follow. The results are quite impressive. However, I do have two major concerns and some minors comments.

Majors:

1. The validation scheme is not quite convincing. What you did is: using part of the training as the validation dataset (near L255, *first three days of every month from the retrieval database*). This can be a major issue since it is shown that GPROF-NN and GPROF-3D is better than GPGORF-Bayesian. The better performance from GPROF-NN and 3D may result from the over-fitting of the Neural network. I am particularly concerned about the over-fitting issue for surface precipitation from GPROF-NN-3D (Fig. 6, bottom left panel, it seems that the vast majority of the pixels are on 1-by-1 line from 0.1 to 10 mm/hr)

Why not use 1-yr independent data (say, 2020 DPR) to validate your results? Based on Fig. 15, it takes about 120 ~ 250 seconds per orbit to get the results. I highly recommend to redo the validation.

2. The most noticeable improve from NN method is for the very light precipitation (<0.1 mm/hr to 0.01 mm/hr, Fig. 6, 1st column). Then the question is: such light precipitation is really beyond the detection capability of both GMI and MHS. Many previous studies showed that the detection threshold value is around 0.2 mm/hr (e.g*., Munchak, S. Joseph, and Gail Skofronick-Jackson. "Evaluation of precipitation detection over various surfaces from passive microwave imagers and sounders." Atmospheric Research 131 (2013): 81-94*.).

In other words, even if GPROF-NN and GPROF-NN-3D can make this light surface precipitation retrieval better, it is difficult to justify physically you did correctly since these light precipitation are beyond the GMI/MHS detection capability.

Minors:

Line 3: “at such high temporal resolution” to “at three hours temporal resolution”, because the temporal resolution from PMWs is rather low (even with the constellation), compared with IR (can be 10 minutes or less).

Line 23: “can be expect” to “can be expected”

Line 33: “3 hours” to “three hours” to be consistent with what you have used in the abstract.

Line 34: “GPM level 3 retrieval products” probably need to change to “GPM level 3 retrieval product”. My understanding is that: there is only one Level 3 product (ie.., IMERG). Also, it may be better to briefly introduce IMERG via one sentence since IMERG is more widely used and known. But not so many studies realized that PMWs form the foundation for IMERG.

Line 134: I believe there are two typos in the multiple-variate normal distribution: (1) ni should be 1; and (2) 2pi, should be (2pi)^n (n is the variable number, should be 13 TBs). Please double check.

Line 157: “as well” to “as well as”

Fig. 5. I don’t understand what is the color squares. In the caption, it is mentioned “Grey squares mark equilaterals with …”, what are the colored squares? I guess grey and color squares are the same??

Fig. 5: Please explain more what are the Fig. 5(c), and why you are doing this way?

Line 250: *To obtain two-dimensional training scenes that are sufficiently wide to train a CNN, we make use of an intermediate CNN based model to 'retrieve' simulated brightness temperatures across the full GMI swath*. Please explain in more details how you did this (i.e., extend from DPR swath to the whole GMI swath).

Both Figure 6 and Figure 7 are over all surface types (i.e., land, ocean, coast, ect)? Please clarify.

Throughout the paper, I did not find which MHS you used (maybe I missed it). Please specify MHS onboard which satellite (there are 5 MHSs, I think).

Line 440: *we are not aware of any other operational PMW algorithms that incorporate structural information using CNNs.* Yes, you are probably correct that nobody is using structural information via CNN. However, structure information has long been used for retrieval from the TRMM era. The land algorithm did by Ferrao group used quite a bit structural information (spatial information) before GPROF transitioned into all Bayesian technique. (see "Estimation of convective/stratiform ratio for TMI pixels" in Gopalan, Kaushik, et al. "Status of the TRMM 2A12 land precipitation algorithm." *Journal of Atmospheric and Oceanic Technology* 27.8 (2010): 1343-1354.) A more recent paper to use the spatial information (Guilloteau, Clément, and Efi Foufoula-Georgiou. "Beyond the pixel: Using patterns and multiscale spatial information to improve the retrieval of precipitation from spaceborne passive microwave imagers." *Journal of atmospheric and oceanic technology* 37.9 (2020): 1571-1591.).

It will be good to briefly discuss how previous studies are using the structural information.