GPROF-NN: A neural network based implementation of the Goddard Profiling Algorithm

Response to reviewer comments

1 Comments from reviewer 2

We want to thank the reviewer for taking the time to read our manuscript and provide valuable feedback.

1.1 Principal changes

Many comments from both reviewers were related to the description of the generation of the training data and the implementation of the retrievals. In order to address these comments and avoid excessive growth of the main text of the manuscript, we have added appendices that explain the generation of the training data, the training itself and the application of the retrievals. We have also rewritten much of Sec. 1 and 2 to improve the presentation of both the scope of the manuscript and the design of the GPROF-NN retrievals.
In addition to that, an error in the simulated brightness temperatures for MHS was discovered and corrected. We have updated the results for MHS. While this improved the overall accuracy of the retrieval, it did not affect the study's main findings. Furthermore, we identified and corrected a minor issue in our evaluation of GPROF that caused a slight overestimation of its accuracy. Again, this correction did not affect the conclusions of the paper.

1.2 Major comments

Reviewer comment 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.

Author response:

It seems that the reviewer has misunderstood our evaluation scheme. We have, of course, not evaluated the model on a sub-set of the data that was used for training. Instead, only days 6 until 31 of every month have been used for training, while days 1 until 3 were used for the evaluation. We will revise this section to make this more clear.

The alternative validation proposed by the reviewer is not really suitable for this study. Firstly, it is not clear whether one year of DPR data would provide sufficiently many collocations with MHS. Secondly, the use of independent validation data introduces an additional error source into the evaluation. Since the declared aim of the study was to assess only the impact of the retrieval method, we consider the validation against independent measurements outside the scope of this study.

We will extend the introduction of the manuscript to highlight these difficulties and better define the scope of the manuscript.

Changes in manuscript:

- We will add a paragraph to the introduction that discusses the difficulties of evaluating precipitation retrievals and explains the motivation for our evaluation scheme.

> **Changes starting in line 102:**
>
> Before a retrieval can replace the current operational version of GPROF, it is imperative to establish its ability to improve the retrieval accuracy to avoid degradation of the GPM products. A balanced evaluation of the accuracy of precipitation retrievals is difficult because it depends on the statistics of the data used in the assessment. Data-driven retrievals generally yield the most accurate results when evaluated on data with the same distribution as the data used for their training. At the same time, evaluation against independent measurements may distort the evaluation when these measurements deviate significantly from the training data. In this study, the retrieval performance of the GPROF-NN algorithms is evaluated and compared to that of GPROF using a held-out part of the retrieval database~~and compared to that of the upcoming version of GPROF. This new version of GPROF~~. This provides the most direct estimate of the benefits of the neural network based retrievals because it avoids the distorting effects of using test data from a different origin. Moreover, the nominal accuracy of both the GPROF and GPROF-NN algorithms provides a reference for future validation against independent measurements.

- We will add a paragraph that clearly states that the data we use for evaluation is not used during the training of the neural network retrievals.

> **Changes starting in line 285:**
>
> The held-out test data comprises observations from the first three days of every

month from the retrieval database. ~~It should be noted that we have deliberately limited this evaluation to data from the retrieval database in order to isolate the effect of the retrieval algorithm from that of the database. We conclude this section with a case study of overpasses of Hurricane Harvey. These results are based on real observations and thus provide an indication to what extent the performance on the retrieval database can be expected to generalize to real observations~~ This data has not been used for training the neural network retrievals. It is, however, derived from the same data sources and thus stems from the same distribution as the training data.

Reviewer comment 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.

Author response:

We do not agree with the reviewer on this point. The findings from Munchak and Skofronick-Jackson (2013) are themselves based on a retrieval. It is therefore possible that a more advanced retrieval method can improve the detection threshold of the sensors. In fact, when we apply the technique from Munchak and Skofronick-Jackson (2013) but instead of the cost function of their variational retrieval use the probability of precipitation retrieved by GPROF, we obtain the graph shown in Fig. 1.1. The detection thresholds for GPROF, GPROF-NN 1D and GPROF-NN 3D are about 0.15, 0.08 and $0.04\,\mathrm{mm\,h^{-1}}$, respectively, as can be seen from the graph. This indicates that the GPROF-NN 1D (3D) retrieval increases the minimum sensitivity of GMI by a factor of 2 (4) and that there is a precipitation signal even at precipitation rates below $0.1\,\mathrm{mm\,h^{-1}}$ Moreover, the simple fact that the neural network based retrievals can improve the retrieval of weak precipitation indicates the presence of a signal from that precipitation. If that wouldn't be the case, there would be no way for the neural network based retrievals to make better predictions than GPROF.

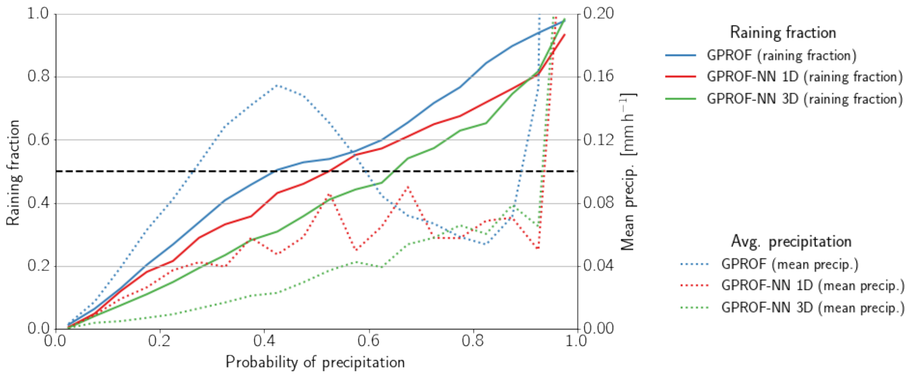

Figure 1.1: Factional occurence of rain (solid lines, left y-axis) and corresponding mean precipitation (dotted lines, right y-axis). This figure is similar to Fig. 6 in Munchak and Skofronick-Jackson (2013) but uses the retrieved probability of precipitation instead of the OEM cost.

1.3 Minor comments

Reviewer comment 1

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).

Author response:

We will reformulate this first part of the abstract to improve the description of the role of PMW observations.

Changes in manuscript:

- We will reformulate the first paragraph of the abstract.

> **Changes starting in line 1:**
>
> The Global Precipitation Measurement (GPM) mission ~~aims to provide global measurements of precipitation with~~ measures global precipitation at a temporal resolution of ~~three hours in order to allow~~ a few hours to enable close monitoring of the global hydrological cycle. ~~To achieve global coverage at such high temporal resolution, GPM combines~~ GPM achieves this by combining observations from a space-borne precipitation radar, a constellation of passive microwave (PMW) sensors and geostationary satellites.

Changes in manuscript:

Reviewer comment 2

Line 23: "can be expect" to "can be expected"

Author response:

We will reformulate the corresponding paragraph and corrected the mistake.

Changes in manuscript

> **Changes starting in line 23:**
>
> Application of the ~~retrieval algorithm to real observations from the GMI and MHS sensors of Hurricane Harvey suggest that these improvements can be expect to~~ retrievals to GMI observations of hurricane Harvey shows moderate improvements when compared to co-located GPM combined and ground-based radar measurements indicating that the improvements at least partially carry over to ~~operational application.~~ assessment against independent measurements.

Reviewer comment 3

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

Author response:

We will replace 'three' with 'few' in the revised version of the manuscript because IMERG actually achieves a temporal resolution of 30 minutes.

Changes in manuscript

> **Changes starting in line 21:**
>
> The Goddard Profiling Algorithm (GPROF, Kummerow et al. (2015)) is the operational precipitation retrieval algorithm for the passive microwave (PMW) observations from the ~~constellation of satellites of~~ radiometer constellation of the Global Precipitation Measurement (GPM, Hou et al. (2014))~~mission~~, whose objective is to provide consistent global measurements of precipitation at a temporal resolution of ~~3 hours. In addition to being used directly by meteorologists and climate scientists, the precipitation that is retrieved using~~ a few hours .

Reviewer comment 4

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.

Author response:

Although, officially, there are many GPM level three products it is true that IMERG is probably the most popular one. We will therefore reformulate the sentence in the revised version of the manuscript to mention IMERG.

Changes in manuscript

> **Changes starting in line 36:**
>
> ~~In addition to being used directly by meteorologists and climate scientists, the precipitation that is retrieved using~~ a few hours. The precipitation retrieved by GPROF serves as input for ~~GPM level 3 retrieval products~~the Integrated Multi-Satellite Retrievals for GPM (IMERG), which can be considered the state-of-the-art of global precipitation measurements.

Reviewer comment 5

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

Author response:

We would like to thank the reviewer for pointing out this mistake. However, instead of removing n_i from the Eq. (2), we will remove it from Eq. (1) and move the 2π inside the determinant.

Equation (1), which has been renamed to (A1), will look as follows in the revised version of the manuscript:

$$\int_{\mathbf{x}} \mathbf{x} p(\mathbf{x}|\mathbf{y})\, d\mathbf{x} = \int_{\mathbf{x}} \mathbf{x}\, \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}\, d\mathbf{x} \approx \frac{\sum_i p(\mathbf{y}|\mathbf{x}_i)\mathbf{x}_i}{\sum_i p(\mathbf{y}|\mathbf{x}_i)}. \tag{1.1}$$

Equation (2), which has been renamed to (A2), will look as follows in the revised version of the manuscript:

$$p(\mathbf{y}|\mathbf{x}_i) = \frac{n_i}{\sqrt{\det(2\pi\mathbf{S})}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{y}_i)^T\mathbf{S}^{-1}(\mathbf{y}-\mathbf{y}_i)\right\} \tag{1.2}$$

Reviewer comment 6

Line 157: “as well” to “as well as”

Author response:

We will correct this in the revised version of the manuscript.

Changes in manuscript:

> **Changes starting in line 179:**
>
> For the GPROF-NN retrievals, the predicted CDF is used to derive most likely and mean surface precipitation (the latter of which is identical to the solution that would have been obtained with common mean squared error regression), the terciles of the posterior distribution as well as the probability of precipitation.

Reviewer comment 7

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??

Author response:

The shading in the background just shows the GMI brightness temperatures. Grey squares are drawn on top to better show the distorting effect of the conical viewing geometry. We will update the figure caption to hopefully make the figure easier to understand.

Changes in manuscript

- The caption of Fig. 5 in the manuscript will be updated. The updated caption is shown in Fig. 1.2

Reviewer comment 8

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).

Author response:

We will add a section to the newly added appendix which describes the process of generating the GPROF-NN 3D training data for sensors other than GMI.

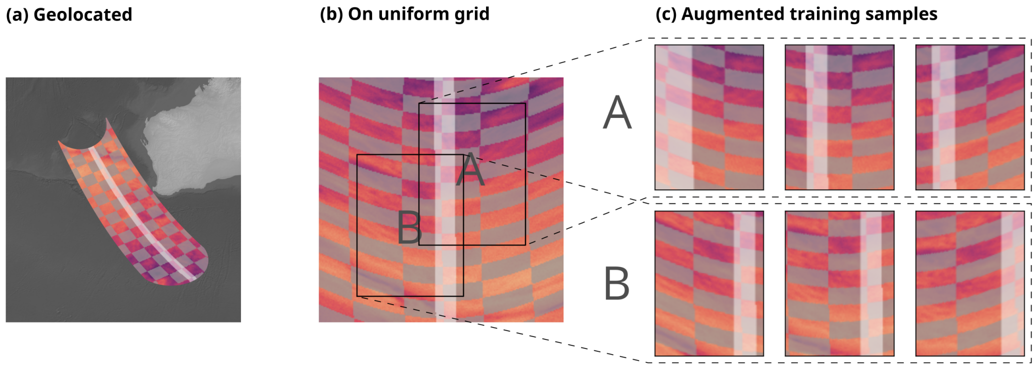

Figure 1.2: The effect of GMIs conical viewing geometry on observed features. Panel (a) displays geolocated observations of the 10.6 GHz channel (colored background). Grey squares mark equilaterals with a side length of 200km oriented along the swath. The highlighted stripe located at the swath center marks the region where the values of the retrieved variables are known. Panel (b) shows the same observations viewed as an image on a uniform grid. Panel (c) shows six synthetically generated training inputs based on two input regions marked in Panel (b). The first row shows three synthetic samples that simulate the effect of viewing the input in region A at a different position across the GMI swath. The second row shows the corresponding transformations for the input in region B.

Changes in manuscript:

- A description of the generation of the training data will be added to Sec. B1 of the revised manuscript.

Reviewer comment 9

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

Author response:

Yes, both plots use all surface types. We will add the clarification to the manuscript.

Changes in manuscript:

> **Changes starting in line 297:**
>
> Scatter plots of the retrieval results for these five quantities over all surfaces are displayed in Fig. 7 for GMI and Fig. 8 for MHS.

Reviewer comment 10

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).

Author response:

The GPROF database doesn't distinguish between the different instances of the MHS sensors, which is why the platform is not stated in the manuscript. For the observations of hurricane Harvey the platform is l. 432.

Reviewer comment 11

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.

Author response

We would like to thank the reviewer for this suggestion and the provided references. We will extend our discussion of the use of spatial information in previous retrievals.

Changes in manuscript:

- We will add a paragraph to the introduction that discusses machine learning and the use of spatial information in remote sensing retrievals.

> **Changes starting in line 64:**
>
> While GPROF is currently based on a data-driven method to solve Bayesian inverse problems, more general machine learning techniques have recently gained popularity for application in precipitation retrievals. Deep neural networks ~~have led to~~ (DNNs), which have enabled a number of ~~important break-throughs in the fields of computer vision, natural language processing and artificial intelligence. They have also gained popularity for remote sensing retrievals of precipitation .~~ significant breakthroughs in different scientific fields (Silver et al., 2016; Jumper et al , have in recent years been explored for retrieving precipitation from satellite observations. Especially convolutional neural networks (CNNs) are appealing for this application because of their ability to leverage spatial patterns in image data. This property sets them apart from traditional retrieval methods and shallow machine-learning techniques, which are limited in their ability to use this information by computational complexity (Duncan et al., 2019) or the need for feature engineering or manual incorporation of spatial information through techniques such as convective-stratiform discrimination (Gopalan et al., 2010) .

- We will also reformulate the discussion of the use of spatial information to include the reference provided by the reviewer.

> **Changes starting in line 509:**
>
> ~~The use of structural information for precipitation retrievals is common practice in algorithms based on infrared observations (Sorooshian et al., 2000; Hong et al., 2004) and the potential benefits of CNN based retrievals have been shown in Sadeghi et al. (2019). While basic structural information has been used in earlier PMW precipitation retrieval algorithms, as e.g. by Kummerow and Giglio (1994), we are not aware of any other operational PMW algorithms that incorporate structural information using CNNs~~ Because precipitation exhibits distinct spatial patterns in satellite observations, many algorithms make use of this information to improve precipitation retrievals (Kummerow and Giglio, 1994; Sorooshian et al., 2000; Hong et al., . Our results confirm that CNNs learn to leverage this information directly from the satellite imagery and that it can notably improve the retrieval accuracy, which is in agreement with the findings from other precipitation retrievals that employ CNNs (Tang et al., 2018; Sadeghi et al., 2019; Gorooh et al., 2022; Sanò et al., 2018)
> .

Bibliography

Duncan, D. I., Eriksson, P., and Pfreundschuh, S.: An experimental 2D-Var retrieval using AMSR2, Atmospheric Measurement Techniques, 12, 6341–6359, https://doi.org/10.5194/amt-12-6341-2019, 2019.

Gopalan, K., Wang, N.-Y., Ferraro, R., and Liu, C.: Status of the TRMM 2A12 Land Precipitation Algorithm, Journal of Atmospheric and Oceanic Technology, 27, 1343 – 1354, https://doi.org/10.1175/2010JTECHA1454.1, 2010.

Gorooh, V. A., Asanjan, A. A., Nguyen, P., Hsu, K., and Sorooshian, S.: Deep Neural Network High Spatiotemporal Resolution Precipitation Estimation (Deep-STEP) Using Passive Microwave and Infrared Data, Journal of Hydrometeorology, 23, 597 – 617, https://doi.org/10.1175/JHM-D-21-0194.1, 2022.

Hong, Y., Hsu, K. L., Sorooshian, S., and Gao, X. G.: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system, J. Appl. Meteor., 43, 1834–1852, 2004.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T.: The Global Precipitation Measurement Mission, Bull. Amer. Met. Soc., 95, 701–722, https://doi.org/10.1175/BAMS-D-13-00164.1, 2014.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al.: Highly accurate protein structure prediction with AlphaFold, Nature, 596, 583–589, 2021.

Kummerow, C. and Giglio, L.: A Passive Microwave Technique for Estimating Rainfall and Vertical Structure Information from Space. Part I: Algorithm Description, Journal of Applied Meteorology and Climatology, 33, 3 – 18, https://doi.org/10.1175/1520-0450(1994)033<0003:APMTFE>2.0.CO;2, 1994.

Kummerow, C. D., Randel, D. L., Kulie, M., Wang, N.-Y., Ferraro, R., Joseph Munchak, S., and Petkovic, V.: The Evolution of the Goddard Profiling Algorithm to a Fully Parametric Scheme, J. Atmos. Oceanic Technol., 32, 2265–2280, https://doi.org/10.1175/JTECH-D-15-0039.1, 2015.

Munchak, S. J. and Skofronick-Jackson, G.: Evaluation of precipitation detection over various surfaces from passive microwave imagers and sounders, Atmospheric Research, 131, 81–94, 2013.

Sadeghi, M., Asanjan, A. A., Faridzad, M., Nguyen, P., Hsu, K., Sorooshian, S., and Braithwaite, D.: PERSIANN-CNN: Precipitation estimation from remotely sensed information using artificial neural networks–convolutional neural networks, Journal of Hydrometeorology, 20, 2273–2289, 2019.

Sanò, P., Panegrossi, G., Casella, D., Marra, A. C., D'Adderio, L. P., Rysman, J. F., and Dietrich, S.: The passive microwave neural network precipitation retrieval (PNPR) algorithm for the CONICAL scanning Global Microwave Imager (GMI) radiometer, Remote Sensing, 10, 1122, 2018.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al.: Mastering the game of Go with deep neural networks and tree search, nature, 529, 484–489, 2016.

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B., and Braithwaite, D.: Evaluation of PERSIANN system satellite based estimates of tropical rainfall, Bull. Amer. Meteor. Soc., 81, 2035–2046, 2000.

Tang, G., Long, D., Behrangi, A., Wang, C., and Hong, Y.: Exploring Deep Neural Networks to Retrieve Rain and Snow in High Latitudes Using Multisensor and Reanalysis Data, Water Resources Research, 54, 8253–8278, https://doi.org/https://doi.org/10.1029/2018WR023830, 2018.