Response to reviewer 1:

We thank the reviewer for their comments and thorough evaluation of our manuscript. The aim of our paper has been to introduce a new method for wet and dry detection that is able to detect wet and dry periods with a higher temporal resolution as compared to existing methods. This allows for more precise identification of rainy time steps, and as a consequence, making identification of short intermittent periods in the CML time series possible. Our main take-away from the two reviewers has been that our paper should have more clearly stated the study's goal and limitations.

Please find our reply in blue to the issues raised.

Title: Review of Manuscript on the Development and Testing of Feedforward Neural Networks for Classifying Wet and Dry Periods Using Commercial Microwave Links

Reviewer: Anonymous

Manuscript ID: egusphere-2024-647

Journal: HESS

Recommendation: Accept with Minor and Major Revisions
This manuscript presents the development and testing of two simple feedforward neural networks (MLPs) designed for classifying wet and dry periods using signal attenuation from commercial microwave links (CMLs), comparing the performance against existing methods. The focus on high temporal resolution is crucial for enhancing the accuracy of rainfall measurements, representing the manuscript's central novelty. As a technical note, the manuscript sufficiently describes the relevance of the study within the existing literature and highlights the existing gaps it aims to bridge.

Minor Comments:

1. The last paragraph of the introduction (paragraph 45) could be more explicit in stating the study's objectives.

We agree that this paragraph can be stated more explicitly. A rephrasing of the paragraph could be "In this study we present two methods for detecting rainy time steps in CML time series data. The goal of both methods is to detect rainy time steps in the time series of a CML where the signal attenuation is provided every 1 minute. This is done with a higher temporal resolution compared to existing methods so that short dry spells during rainy periods can be identified. One method is trained on radar reference data and the other method is trained on rain gauge reference data. Both methods are tested against rain gauge and radar data, highlighting their difference. We also examine the performance of the developed methods in comparison to existing approaches, aiming to gain a clearer understanding of the differences between the two alternative methods."

2. The first sentence in paragraph 120, subsection 3.1, should be rephrased for enhanced clarity.

We agree with this. The rephrased sentence could be "The performance (MCC) of MLP_RA and MLP_GA for the training and test dataset as a function of the number of neurons and hidden layer sizes is shown in Figure 1. For each hidden layer configuration, the optimal regularization and initial learning rate that yielded the highest mean MCC was selected and plotted together with the minimum and maximum of all 5 folds obtained from k-fold cross-validation.

3. It would be beneficial for the authors to elaborate on why using total signal loss from both sublinks, as opposed to one, results in improved classification outcomes in paragraph 70.

This is indeed an interesting aspect worth elaborating further. However, we do not know exactly why this is the case, but we do know that there is an improvement. Adding too much speculation in the methods section is not desirable, but we could for instance add:

- Paragraph 70: There was also an improvement from using both sublinks rather than one, possibly because two sublinks include more information than one, which could help the MLP filter out noise. Note that the CNN also uses two sublinks whereas the sigma_80 method just uses one. As this topic was not the focus of this study we do not show these findings in this note.
- Paragraph 80: We note that, similar to our MLP, the CNN method is also trained to use two sublinks, whereas the sigma_80 method just uses one.

Major Comments:

1. The dataset comprising 3901 CMLs covers only a single month (01-07-2021 to 31-07-2021), which may not adequately represent different rainfall periods and seasonal variations. Expanding the temporal coverage or discussing the potential limitations and implications of this scope on the study's conclusions would strengthen the paper's validity.

It is a good point that our dataset may not fully represent different rainfall periods and seasonal variations. The choice of using a temporally limited dataset was driven by several factors:

a) Since our dataset covers the whole of Germany using 395 CML-rain gauge pairs we believe that the dataset still captures sufficiently different rainfall events. For instance, in addition to several smaller rainfall events the dataset also captures the large precipitation event that happened in Germany between the 13th and 15th of July 2021.

b) As stated in the paper, in order for the training to converge properly, the training data consisted of the top 26 CML-rain gauge pairs that showed good correlation when the sigma_80 method was used for CML wet dry classification. We could not find any improvement when going beyond 26 pairs. This implies that the training dataset consisting of 26 CML-rain gauge pairs contains a sufficient number of different rainfall events to generalize to the 369 other CML-rain gauge pairs. Thus, we do not expect that expanding the dataset to include several more months of data will impact the results significantly. A very interesting case could be to use data from winter months as this data typically consists of different precipitation types such as dry snow and sleet. However, as the rain gauges and weather radar do not distinguish between snow and rain it would be very hard to train a MLP to detect rain, as there is no way of knowing the ground truth.

We propose to reformulate the discussion and conclusion to highlight the points discussed above.

a) Paragraph 90 delete section 2.5. Add the following to paragraph 60 section 2.1: “Our study focused on CML-rain gauge pairs located closer to each other than 5 km. This resulted in 395 pairs of CMLs and rain gauges spread out across Germany. Even though there are many CMLs in our dataset, we only have 249 unique rain gauges serving as references. This means that some CMLs use the same rain gauge for reference.”
b) Paragraph 96: specify that in the training data we used data from 26 unique rain gauges.
c) Paragraph 195: Add a discussion on the implication of only using 1 month of data. "Our study comprises CML, weather radar and rain gauge data from 395 CML-rain gauge pairs over one month. A possible limitation is that one single month might not adequately represent the different rainfall types associated with other months or different geographical locations. On the other hand, since our dataset covers the whole of Germany the dataset contains widely different precipitation events. For instance, in addition to several smaller events, the dataset also captures the large precipitation event that happened in Germany between the 13th and 15 of July 2021. Moreover, in order to ensure convergence of the MLPs the training data used only 26 CML-rain gauge pairs. Including more pairs, however, did not improve the results on the validation dataset, indicating that the MLPs in fact generalize to several different precipitation events.”

2. The manuscript would benefit significantly from a brief discussion regarding the limitations of the study. Such a discussion may include the potential biases introduced by the dataset's temporal limitations (1-minute gauge data versus 5-minute Radar data), the generalizability of the MLP models to other geographic contexts, and the implications of the methodological choices made (e.g., neural network configuration).

We suggest to extend the results and discussion chapter with another subsection named general discussion containing the discussion of major comment 1 and the other issues in major comment 2:

> Our results indicate that MLP_RA provide rainfall estimates that are more continuous, and more consistent over time, compared to the more intermittent estimates generated by MLP_GA. This could come from the fact that the rain gauges have a 1 minute resolution while the weather radar has a 5 minute resolution, making the radar rainy periods more continuous. Another explanation could be that at low rainfall rates, the rain gauge will not record any rainfall before the droplets has been transported to the weight, making the period seem more intermittent than it actually is. Further, while the rain gauges measure point rainfall close to the CML, the weather radar measures average rainfall along the CML. This path averaging blurs the rainy periods, making the rainy period more continuous with less intermittent breaks. An interesting finding is that even if the rain gauge do not represent the average rainfall along the CML, the ground truth is still precise enough so that MLP_GA is able to capture more of the underlying intermittency as compared to MLP_RA. This is also reflected in the neural network configuration where the MLP_GA benefits from a more complex network architecture as compared to MLP_RA.
>
> Both MLPs were trained using the 26 CML-reference pairs that showed the highest MCC estimated using the sigma_80 method. This can be thought of as a pre-processing step, where the goal was to ensure training data with a good match between the reference and the CML. In our case this was important for making the MLPs converge to approximately the same weights every time we trained the model. These particular pairs might, since they by selection have a good correlation with its reference, also contain little or no noise. Thus, the MLP training datasets might lack exposure to noisy CML time series, and as a consequence, the MLPs might not very well handle noisy periods. On the other hand, from Figure 2 we know that the MLPs still outperforms the sigma80 and CNN method on the 369 CMLs used in the test dataset, which was not subject to any noise filtering, suggesting that the MLPs at least to some extent are able to handle noise. Moreover, very noisy CMLs are typically handled using pre-processing methods such as filtering out CMLs with strong diurnal cycles or plateaus such as done in (Graf2020) and (blettner_2023).

Overall it must be noted that while the MCC is a useful and balanced metric, its score must be seen in relation to the reference chosen for evaluation. As weather radar provides average rainfall intensities for the entire radar grid cell, we expect that the radar rainfall estimates are less intermittent than what is observed by a rain gauge. This is supported by the findings in Figure 3, Figure4 and Figure 5 where the weather radar rainfall events are less intermittent than what is the case for the rain gauges. The CML, like the weather radar, also measures spatially averaged rainfall. However, the CML measures rainfall closer to the ground and might thus be able to better capture the intermittency as seen by the rain gauge. In this study MLP_GA was able to better detect rainfall events as seen by the rain gauge than MLP_RA. This suggests that there is no single best reference or method for evaluating CML rainy periods. Rather, the CML rain event detection method must be seen in relation to its application."

We further suggest to rewrite the conclusion to:

In this technical note, we introduced two simple feedforward neural networks (MLPs) trained to detect rainy time steps in signal attenuation data from commercial microwave links (CMLs). The MLPs are trained and tested using reference data from rain gauges (MLP_GA) with a temporal resolution of 1 minute and gauge-adjusted radar (MLP_RA) with a temporal resolution of 5 minutes. Whereas existing methods tend to estimate longer continuous rainy periods, the MLPs estimates shorter rainy periods that more closely resembles the intermittent rainfall patterns that is observed by the rain gauges and weather radar. The performance of the MLPs are evaluated by comparing the MLPs estimates with estimates produced by two existing methods using Matthews correlation coefficient. Our results show that the MLPs outperforms existing methods in almost all cases.

Interestingly, even if the rain gauges do not resemble the path averaged rainfall as observed by the CML, MLP_GA was still able to learn the rainfall pattern in the CML time series. Moreover, MLP_GA better estimates rainy periods as recorded at the nearby rain gauges than what is the case for MLP_RA, while both methods perform equally well when radar data is used as reference.

Both MLPs tend to estimate rainy periods after the CML total loss as started to increase. Thus, if the MLPs are used for baseline estimation the user should, similar to (pastorek2022), consider using dry time steps at least 5 minutes away from the identified rainy time step for baseline estimation.

Future work may involve further refining the model architecture and testing its robustness in generalization to other datasets. Another interesting topic could be to better understand how different wet and dry classifications affect the resulting baselines and the effect this has on rainfall rate estimation from CML data. Overall, both MLPs showed successful skill for the challenge of rainfall event detection in CML attenuation time series.