Reviewer 1

We thank the reviewer for her/his efforts and time to evaluate our manuscript and we are grateful for the provided feedback.

General Assessment

This manuscript presents systematic investigation into the potential of using commercial microwave communication links, first presented David et.al., and suggest machine-learning methods to detect fog where rain may be the cause for signal attenuation. The topic is highly relevant for atmospheric monitoring, transportation safety, and applied meteorology.

The work consists of the following steps:

1. Construct a reference data set for training and validation.

2. Construct the appropriate McFly/InceptionTime architecture

3. Train it on high temporal-resolution experimental dataset from Wageningen

4. Apply its reduced version to a national-scale commercial network

The work represents a comprehensive study with many details. However, in terms of novelty, its claimed contribution is in the use of the McFly/InceptionTime architecture for fog detection. Considering the constraints and the limitations of the available data, the results are reasonable and justify publication. However, there are several methodological and interpretational issues that should be addressed before publication.

Major Comments

1. Reference Dataset Construction. The paper aims at distinguishing between rain and fog events. However, dry periods can also be identified as fog, especially with high levels of humidity. It is therefore important that the reference dataset consist of rain and fog events, as well as dry periods. Moreover, for meaningful accuracy assessment of any ML-algorithm, the reference dataset should be balanced.

We thank the reviewer for the remark about the balanced data. We originally performed an exercise to balance the dataset, however doing this tended result in an overestimation of the occurrence of fog.

After balancing the dataset and feeding the raw data into the McFly algorithm, it gave a 49% accuracy. For a balanced two-class dataset, this is not better than just random guessing. It appears that the raw data for each individual timestep was not enough to create a neural network that detects fog. However, the attenuation signal in the CML data shows a pattern associated with fog. This pattern however changes over time, something McFly does not automatically take into consideration. To account for the temporal aspects, a rolling rate of change of the CML attenuation signal over 15 minutes was included for every time step as input to McFly. On its own, this input value could be used to gain a validation accuracy of 67%, but combining it with the raw data brought it up to 73%. In addition, we included a form of data to represent the time of the day and a time of the year, to allow the neural network to learn the climatology of fog. On its own, this could correctly predict fog 60% of the time steps, but when combined with the other observations the validation accuracy is 77%.

To further account for the imbalance of the used dataset, we use the F1 score as measure of predictive performance. The F1 score is particularly relevant in applications which are primarily concerned with the positive class and where the positive class is rare relative to the negative class. In the revised version of the manuscript we have also added the “balanced accuracy” and “Matthews correlation coefficient” as metrics. These metrics are particular suitable for unbalanced datasets, and confirm the earlier findings based on the F1 score only. We added a reference to the book chapter that underlines this insight, i.e.

Diallo, R., Edalo, C., Awe, O.O. (2025). Machine Learning Evaluation of Imbalanced Health Data: A Comparative Analysis of Balanced Accuracy, MCC, and F1 Score. In: Awe, O.O., A. Vance, E. (eds) Practical Statistical Learning and Data Science Methods. STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health. Springer, Cham. https://doi.org/10.1007/978-3-031-72215-8_12

2. Reference Dataset Validation. Existing validation of the reference data set is qualitative. A quantitative analysis, indicating on misdetection vs. false identification of fog in rainy/dry periods via an e.g., ROC , is required.

The reference dataset qualitative since direct visibility observations in the path are missing. Note that the experiment has never been set up to test fog detection, but we use the existing data explore its potential for fog detection. Visibility observations in the path are of course recommended for future analogue work.

3. The McFly/InceptionTime architecture. Since the architecture has been chosen by comparing performance of several options, the performance measure is crucial. As mentioned, imbalanced data leads to optimistic accuracy performance. But, what really counts as a performance measure is the accuracy. Therefore, it is recommended to balance the data, as often done in similar applications (1 See, for example, Habi, Hai Victor, and Hagit Messer. "Recurrent neural network for rain estimation using commercial microwave links." *IEEE Transactions on Geoscience and Remote Sensing* 59.5 (2020): 3672-3681.). Alternatively, using the imbalanced data, indication on false detection may indicate on the effect of the long no-fog periods.

-> See rebuttal to point 1. The suggested reference has been cited.

4. Wet/Dry classification. As mentioned, previous studies of Fog and CML excluded rain periods, while this study does not. However, a discussion of the possibility of combining wet/dry classification with fog/rain classification is missing.

We have added a short discussion about this aspect in the Discussion section of the revised manuscript.

5. Misclassification. Treating 30-second segments as independent samples does not adequately represent the temporal evolution of fog. This likely contributes to misclassification of drizzle and weak precipitation and even high humidity. Consideration of temporal sequence models or multi-scale temporal inputs is strongly recommended.

We truly agree that fog undergoes a typical life cycle. However the trend (time aspect) of the Received Signal Level (RSL) was already accounted for the in original manuscript (original section 2.4.3).

6. Generalisation and Dataset Mismatch. While application of the method of country wide network is important, and is probably the major novelty in the paper, it needs consideration. In particular, the reduced model is trained on a single-link dataset but applied to a national network with different temporal coverage and climatological conditions. The manuscript should more clearly acknowledge and discuss the limitations of this generalisation.

Based on the reviewer's suggestion, and the feedback by reviewer #2, we have move this piece to the Discussion section, indicating this part of the study is relatively more explorative.

7. Interpretation of Physical Mechanisms Several explanations (e.g., droplet-size effects, Rayleigh attenuation assumptions, cold fog behaviour) are presented without sufficient direct support from the literature. These interpretations should be more cautiously framed and supported with appropriate references.

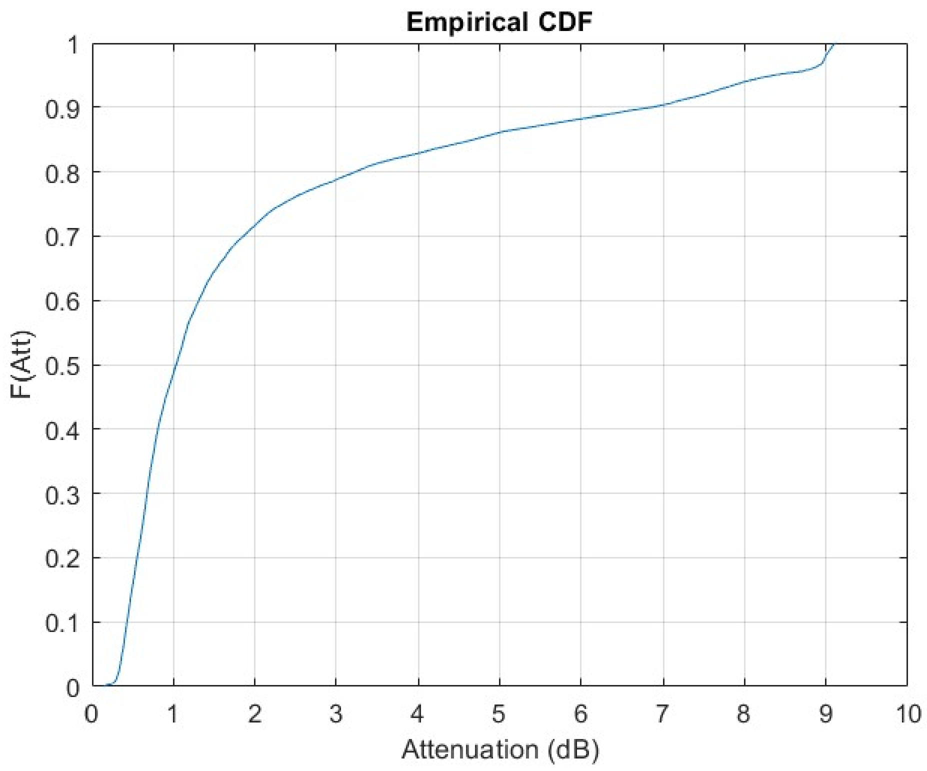
The particular paragraph has been removed.

Minor Comments

1. **Construction and validation of the fog reference dataset**. This is the most consequential methodological step, yet the description is somewhat subjective and potentially yields systematic biases. It raises few concerns, including:

- The BLS900 15th percentile threshold is climatologically arbitrary. The text states "there is a clear divide," but no histogram or statistical justification is shown.

The CDF of the BLS signal is provided below.



- Fog detection is partially based on VK humidity ≥ 90%, even though the weather station is 3 km away and not colocalized. Fog is extremely local; this may introduce classification errors.

We agree with the local nature of the fog, and it was exactly therefore that we did not include the records from the visibility sensor at VK.

- The exclusion of precipitation from the fog definition removes the possibility to train the model on precipitation fog, later causing systematic false negatives.

- The qualitative validation via time-lapse imagery is not sufficiently extensive to quantify accuracy or bias.

To deal with these concerns, it is recommended that the following will be Included:

- ROC or sensitivity analysis for percentile selection.

- A confusion table comparing the reference against manual labeling from a subsample of time-lapse images.

- A discussion of bias induced by using VK humidity for local fog inference.

All these points consider the validation of the reference dataset. However, as mentioned in the manuscript, an additional independent dataset that can act as ground truth is not available (apart from the time lapse). This is partly because the experiment was not set up to a priori detect fog (but rainfall as in VL18), but in the current study we explore additional potential of its use. The validation of the reference dataset was therefore performed using the camera time lapse. In the revised manuscript we added a sentence that developing a confusion matrix, ROC and related metrics is not possible.

2. **Experiments and Results.** Some interpretations mix meteorological reasoning with speculative hypotheses. For example, the drop in microwave attenuation during dense fog is attributed to droplet size changes, but no microphysical evidence is presented. Also, claims about Rayleigh theory and droplet spectra require stronger linkage to literature. It is therefore recommended to state explicitly when results are hypotheses vs validated explanations.

-> the particular paragraph about the Rayleigh theory has been removed in the revised version. Note in the original manuscript we already mentioned this piece was beyond the precise scope of the study.

Two fundamentally different failure modes appear: "fog not detected due to physics" vs "fog not detected due to model design". This distinction between physics-limited detection (e.g., microwave attenuation negligible for precipitation-fog or cold-fog due to droplet sizes) and model architecture limitations (e.g., nighttime drizzle misclassified as fog due to learned temporal priors) should be emphasized.

->We thank the reviewer for this suggestion, which we have implemented. Note that we removed the snow case from the manuscript so the cold-fog case failure is less prominent.

Please elaborate on the reduced model – how did you relate the 15 minutes min-max measurements with the instantaneous measurements in the reference data set?

->We clarified this point that fog should be present in the 15 min interval through a mean 15-min signal in the BLS.

Please clarify the country-wide results as presented in e.g., Fig. 14. The captions says that it presents "spatial maps comparing fog detection" while the reference map (b) indicates in rain.

Figure 14 has been removed from the revised manuscript.

3. Discussion

- Since both fog and humidity cause similar attenuation, the challenge of distinguishing between them in the reference dataset should be discussed, given that relative humidity higher than 90% is required.

One of the challenges in fog detection with CMLs lies in the fact that water vapour itself also generates some attenuation. While the strongest attenuation appears in the 22 GHz band, even some attenuation can be present in at 38 GHz that we utilize. D13b shows that for high RH, the attenuation amounts to 0.18 dB for a 1 km path and 0.82 dB for a 5 km path. Moreover they show that in practise (their figures 5 and 8) the attenuation difference between fog and high water vapour loads is clear enough for a 2.2 km path. For their setup, they concluded that the difference in attenuation between a foggy and a humid (but non-foggy) night, at 38 GHz, is in the range between 0.05 and 0.3 dB for links of 1 and 5 km lengths, respectively. This appears to be substantially lower with respect to the expected attenuation from high concentration fog.

- Discussion of the potential of using other ML methods (LSTM, GRU, Temporal CNN with causal convolutions) is missing, especially since treating windows as independent loses crucial information and encourages overfitting to instantaneous patterns.

As mentioned before, in the original manuscript we already took into account the RSL trend (time evolution effect) in the ML approach. In addition, while we agree that other ML approaches may have potential for fog detection, one of the original aspects of the study is that the McFly model has been employed to select from the potential pool of ML techniques the best one for the dataset at hand. Adding more ML techniques would make the already relatively long manuscript even longer.

4. Clarity/Structure

- Figures 7–10 are information dense; consider adding small labeled boxes showing fog/non-fog intervals visually.

We thank the reviewer for this suggestion. However, such a time line indicating the presence of fog was already present in the figures in the original manuscript (grey bar at the bottom of the mentioned figures).

- The introduction reviews the fog prediction/remote sensing literature well, but ML-for-atmospheric sensing needs additional contextualization.

We have included a paragraph and references accordingly:

Machine learning has been more and more used in fog forecasting and detection. Sun et al (2026) applied random forest, logistic regression, K-nearest neighbours, Gaussian Naive Bayes, and decision trees to ECMWF forecast data as inputs over land (China). All five machine learning algorithms demonstrated skill in short-term heavy fog forecasting, with the random forest algorithm achieving the highest threat score. Water vapor, thermodynamic conditions, and upstream–downstream effects were identified as the most important predictors. Xiang et al, (2025) presented a hybrid forecasting approach for the Yangtze River Estuary that combines ML with NWP. The approach appeared suitable for various locations and seasons, and can predict sea fog events with visibility below 1 km up to 60 h in advance. This study demonstrates the potential of ML to improve the accuracy of sea fog forecasts. Peláez-Rodríguez, C., et al, implemented successfully forecasted visibility for 1 h, 3 h and 6 h horizons for Galicia (Spain), leading to a 17.3% increase in score over the reference prediction models. Such an approach was confirmed to be successful for a 5-year period for the Indo-Gangetic Plains (Deshpande et al., 2026).

Relatively less studies focus on for detection using ML. Kim et al;. (2020) developed a fog detection algorithm based on Lidar Ceilometer, but they report fog detection was difficult with an F1 score of 0.10. ApinayaPrethi and Nithya (2023) used dashcam images and estimated fog density and visibility to detect using machine learning approaches (SVM). Finally Nasichah et al. (2026) reviewed deep learning approaches for fog detection and prediction using satellite and CCTV imagery, and maps datasets, architectures, preprocessing, and evaluation strategies. Model credibility was shaped by data provenance and evaluation design than by model complexity. Their review provides a reproducible foundation for operational reliability in future fog-detection.

- Section 4.3 (Methods) is long and mixes speculation, results, and suggestions. Consider splitting into "Limitations" and "Future Work."

We have split the section (despite we do not consider it as long with 298 words only) into Limitations and Future Work accordingly.

5. References

Relevant references are from 3 groups: meteorological (e.g., footnote 2), relevant machine learning, and relevant CMLs studies (e.g., footnote 1). Since the claim contribution is mostly in the use of ML for fog detection using CML, it is recommended also to mention the first3 to use ML with CMLs measurements.

2 For example, see Gultepe et al. (2017). *Fog Microphysics and Visibility.*

3 Habi, Hai Victor, and Hagit Messer. "Wet-dry classification using lstm and commercial microwave links." *2018 IEEE 10th sensor array and multichannel signal processing workshop (SAM)*. IEEE, 2018.

Thanks for the suggested references. They have been included in the revised version.