Responses to comments posted by Referee #1

We thank the referee for reviewing our manuscript and providing constructive feedback. It helped us to clarify unambiguous phrasings and the presentation of the analysis. We answer all comments in the following text. Our answers are in blue.

In this work the authors apply information theory to rainfall retrieval from commercial microwave links. Though the theory has been applied to other areas in hydrology, it has not been applied to rainfall retrieval from CML before, and the approach, therefore, is an interesting one. The authors apply the information theoretic framework to two CML processing steps, namely the quantification of precipitation estimates, and the wet-dry classification. They show that considering additional CML variables, as well as external environmental variables can highlight the importance, or weight, of the different parts in the CML processing chain, as well as improve the prediction of the target variables, which is promising.

The article is generally well structured with clearly described sections. It is written to-the-point, though sometimes at the expense of being too brief, particularly on the discussion side. The article does include a comprehensive overview on information theory with references for further consideration, which is nice.

General comments:

Wet-dry classification

My main comments on the article refers the methodology of the wet-dry classification analysis. As the target variable the authors use manually identified wet and dry timesteps at 1 minute resolution, and the analysis is based on a single CML. In my opinion this strongly hampers the reproducibility and the generalizability of this analysis.

It unclear to me how the authors choose the current link the analysis is based on. Can they show how this framework would behave if a different link is chosen? In other words, how generalizable is this analysis, and how sensitive are the results to the specific link chosen?

Thank you for this comment. First of all, we would like to clarify that dry-wet analysis is meant as a demonstration study illustrating the feasibility and the potential of a non-parametric model based on information theory. As such, it does have limited representativeness. We have revised accordingly the introduction section on L77 - 78 in the track-changes manuscript, where the analysis is first introduced. In the revised manuscript, we make this also clear in the abstract (L20 - 22 in the track-changes manuscript) and conclusions (L583 - 586).

The link was chosen from Špačková et al. (2023), we now mention it in the data section (L314 - 316). Specifically, it is a link from the group 1a. We highlighted the selected link in Fig. 1 (L341).

Špačková, A., Fencl, M., and Bareš, V.: Evaluation of error components in rainfall retrieval from collocated commercial microwave links, Atmos. Meas. Tech., 16, 3865–3879, https://doi.org/10.5194/amt-16-3865-2023, 2023.

The authors also state that the wet and dry timesteps are separated by visually inspecting the total loss attenuation. Is this visual inspection not based on an implicit threshold? And would it then not be possible to apply such a threshold programmatically? For reproducibility it would be important if the authors can show how the analysis would turn out when the wet-dry timesteps are identified programmatically, using a certain condition. Because up-scaling this information theoretic framework in its current form to multiple CMLs, let alone an entire network, does not seem feasible manually.

Following up on your first question, the visual inspection does not rely on an implicit threshold. Increased attenuation can result from non-rainfall factors like wet antenna attenuation or dew.

CMLs are opportunistic sensors, and current methods are limited to specific datasets and hardware, with few intercomparisons, which is outside our scope. While programmatic methods are useful and included in our comparisons, they also have limitations, particularly in distinguishing light rain from noise at high temporal resolution. Using semi-automatically identified rainfall events as a reference at 1 min resolution is a practical available reference in the absence of a precise ground truth. We agree that upscaling remains a challenge for many CML methods, and our aim is to identify variables that carry relevant information without formulation of empirical relations rather than propose a network-wide solution.

Referee 2 raised a similar comment, and we kindly recommend referring to our detailed response to comment 3 for further clarification.

If this should become too similar to the alternative wet-dry approaches discussed in the article, I suggest using the weather radar data as reference, as is done for the QPE analysis, even though this would change the temporal resolution to 5 minutes. Alternatively a nearby gauge could be used. This would also ensure that wet and dry periods are determined based on actual rainfall sensors.

As noted in the first response, the analysis serves as an illustration of the information-theory framework to design a non-parametric model for CML processing. The semi-automatic dry-wet classification was a pragmatic choice and despite limited reproducibility, it serves this purpose well in our opinion. Moreover, due to restricted access to radar data in Czechia (we have only data for relevant wet-weather periods), an automated programmatical approach would have to rely on a near-by rain gauge only, which is at 1 min time scale insufficient.

The rain-gauge was actually used in our classification as a visual reference, and we make this now clear in the L316 - 322. We also did additional analysis to evaluate consistency between our semi-automatic classification and data from the nearest rain gauge being approx. 3 km far from the CML path. We refer to the response to the second reviewer comment no. 3 for more details and changes in the manuscript.

Finally, regarding the alternative wet-dry classification approaches the authors suggest two common approaches. Regarding approach B, there are other common approaches in the references listed by the authors that either build on top of the approach by Schleiss and Berne

(2010), namely Graf et al., 2020, or that use completely different methodology (Overeem et al., 2013). Therefore to call the approach by Schleiss and Berne (2010) state-of-the-art seems premature, and it would be worthwhile to see how the results from the information theoretic framework compares to other common wet-dry classification approaches, that apparently have been preferred in those publications over the Schleiss and Berne (2010) method.

Agreed. We have selected Shleiss and Berne (2010) because it is often used as a benchmark. We now call their approach "commonly used" instead of "state-of-the-art". The presented analysis does not aim to address the research gap in method intercomparisons, which is certainly needed in some other methodological paper. The alternative approaches selected in this analysis are the most straightforward and interpretable as well as being the most common.

General textual comments

To me the term 'external variables' is not very intuitive and rather vague. On several occasions the authors use 'environmental' or 'atmospheric' variables. In my opinion exchanging the word external with either environmental or atmospheric would improve the readability of the article.

For the sake of clarity and readability, we have followed your advice and modified the term to "environmental variable". Please see the changes of 'external variable' to 'environmental variables' in the revised manuscript.

In the same way, I would consider changing the term 'internal variables' to sensor or system variables, as this more directly describes the list of so-called internal variables.

In contrast to the previous, we agree that the term 'sensor variables' is more clear to the reader than the term 'internal variable'. The term was replaced in the respective parts of the text where it appeared. Please see the changes of 'internal variable' to 'sensor variables' in the revised manuscript.

In my understanding the term 'synoptic' often refers to large scale weather states, though granted, the term can have different meaning. To avoid confusion, and since in this case the study area is 35x35km, I would suggest simply using the term weather types.

You are right, the synoptic types are large-scale weather states valid beyond the domain of our study area. The data source of synoptic types is unfortunately only in the Czech language (CHMI: Typizace povětrnostních situací pro území České republiky: https://www.chmi.cz/historicka-data/pocasi/typizace-povetrnostnich-situaci). The synoptic type describes large-scale atmospheric patterns (cyclones, fronts, etc.) driving the weather, while the weather type is about specific local weather conditions experienced on the ground. Synoptic types provide the context for different weather types. We rely on the translation of the term 'synoptický typ' provided by the Czech Meteorological Society (http://slovnik.cmes.cz/fulltext/synoptick%C3%BD%20typ). To be consistent with the expertise of the Czech Meteorological Society, we prefer to keep the term 'synoptic type'.

The goal of the article and title

As I read it, the goal of the article is twofold. It is to give insights into the relative importance of different CML processing steps, and subsequently to see how the uncertainty can be reduced by using (qualitative) sensor or environmental variables. For example, in the abstract the authors state they "propose to evaluate CML processing" (L11) and end with the goal to "ultimately facilitate the improvement of CML rainfall estimates" (L24).

In my opinion the title is a little bit vague and does not cover these two goals entirely. I would also add the term "information theory" in the title to immediately make explicit the method used, or make the term "additional variables" more explicit, i.e. additional sensor and environmental variables.

Some suggestions:

- Information-theoretic analysis of processing steps for rainfall retrieval from commercial microwave links.

- Using environmental variables and information theory to gain insights into processing steps for rainfall retrieval from commercial microwave links

Moreover, if my interpretation of the goal of this article is indeed correct I would accentuate that dual goal more clearly in the introduction.

Thank you for the suggestions for improving the title. We propose combining both your suggestions in the following title: "*Information-theoretic analysis of commercial microwave link and environmental variables in rainfall estimation*".

Specific comments:

L54-L56: It is not entirely clear to me what the authors want to say. By "interpretation of the CML data using deterministic models" do they mean estimating rainfall from attenuation suffers from many assumptions in the currently available models? And which CML empirical relations do the authors refer to, the k-R relation? And which variables, total loss? It would be good to be explicit in this paragraph, or give some examples of what you mean.

The changes in the track-changes manuscript (L59 - 62) are in red:

In summary, despite significant progress in the retrieval of CML QPE over the past decade, the interpretation of the CML precipitation data using deterministic models is difficult. The precise processing of observed total loss to rainfall intensity requires accurate description of conditions along the CML path as total loss is composed from free space loss, atmospheric absorption, and losses and gains at the antennas (ITU, 2009). CML empirical relations, e.g. $k - R$ relation with its empirical parameters or models for wetting of antenna radomes, are

designed to reduce bias by utilising variables that can be obtained from the opportunistic sensors. …

ITU-R: Handbook - Radiowave propagation information for designing terrestrial point-to-point links, ITU, Geneva, Switzerland, 2009.

L59: It is commendable that the authors mention that there are uncertainties associated with gauge-adjusted weather radar too. Since they use this as their reference, a short description in the discussion of the effects this has on their results would be appropriate.

This is addressed in the Discussion section under 'Accuracy of reference rainfall', which provides helpful details and explanations.

L132: It is appreciated that in the context of this work the authors later (in Fig. 3) explore what a "large enough dataset" is.

Thank you.

L184: Applying a threshold of 0.5mm/h means there are no dry timesteps. Though it makes sense to apply this processing step is isolated like in this analysis, but in a near real-time processing chain there will be dry timesteps present too. Hence a comment, here, or in the discussion, on the implications of this threshold on the results would be appropriate.

The changes in the track-changes manuscript (L216 - 218) are in red:

… To prevent skewness, caused by a large number of dry timesteps, the dataset is subset by exceedance over a threshold of 0.5 mm h^{-1} rainfall intensity of the adjusted radar at the corresponding CML path. The results reflect this consideration because the choice of threshold can influence precipitation occurrence statistics, as light precipitation below the detection thresholds is common.

L190: As mentioned here and in L117, binning is a subjective choice. It would help to mention based on what user requirements you made your choice, and how your choice reflects the size, distribution and precision of the data.

We suggest elaborate on this more in Discussion in the 'Binning' section, where it is already partly covered for rainfall intensities. Specific comments on the binning choice in this study is in L295.

The changes in the track-changes manuscript (L550 - 553) are in red:

…defragment the occupancy of the bins. The binning was selected to capture the variability of rainfall data and skewness of its distribution, while maintaining a reasonable sample size

within the bins. Even though regular binning is applied wherever possible to keep distribution variability, for some variables, e.g. TL gradient, the wider bin size is applied for data close to the data range limits to preserve the data range and limit the number of empty bins. The selection of bin size…

Section 3.2: please see general comments on wet-dry classification.

Please see the response above.

L213: "..greater than the threshold." Which threshold? The detection threshold you are trying to determine, making this an iterative process?

The changes in the track-changes manuscript (L254 and 256) are in red:

… The detection threshold is determined iteratively as follows: …

… greater than the detection threshold, it is labelled as wet. …

L221: So in the end you use one optimal threshold?

The changes in the track-changes manuscript (L267) are in red:

The optimal threshold is then applied on the testing subset of the data. The results of the proposed…

L229: From the text it is unclear to me if, and where, approach A has been used in literature before, or whether it was designed specifically for this analysis. This would be helpful to show how established this method is.

The changes in the track-changes manuscript (L279 - 280) are in red:

It first estimates the baseline on a data-driven basis as a 1 % quantile of total loss baseline, a 7 day baseline of 15 min averages, a constant median baseline, or as 7 day moving quantile window baseline (Overeem et al., 2016; Fencl et al., 2020; Fencl et al., 2017).

Fencl, M., Dohnal, M., Rieckermann, J., and Bareš, V.: Gauge-adjusted rainfall estimates from commercial microwave links, Hydrol. Earth Syst. Sci., 21, 617–634, https://doi.org/10.5194/hess-21-617-2017, 2017.

Fencl, M., Dohnal, M., Valtr, P., Grabner, M., and Bareš, V.: Atmospheric observations with E-band microwave links – challenges and opportunities, Atmos. Meas. Tech., 13, 6559–6578, https://doi.org/10.5194/amt-13-6559-2020, 2020.

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network, Atmos. Meas. Tech., 9, 2425–2444, https://doi.org/10.5194/amt-9-2425-2016, 2016.

L241: Please elaborate why you chose for the different temporal resolutions of 15 and 1 minute. This is currently not clear to me from the text.

The changes in the track-changes manuscript (L290 - 293) are in red:

… The data resolution for rainfall estimation analysis is 15 min. As the raw 5 min weather radar sampling suffers from significant bias, even when adjusted by municipal rain gauges, the data were aggregated to 15 min balancing sufficient size of the dataset with bias reduction (see the subsection 5.3.2 Accuracy of reference rainfall in Discussion). ~~and~~The data resolution for wet-dry classification analysis ~~it~~ is 1 min thus preserving the resolution of the CML data. …

L280: What is meant by 'visual shift'?

The changes in the track-changes manuscript (L337 - 338) are in red:

… CMLs with no visual ~~shift~~ artifacts, e.g., sudden changes in total loss caused by hardware malfunction, are accepted. This results in~~, thus making~~ 117 CMLs available for the analysis. …

L303: What is meant with aligned? Aligned temporally?

Thank you for highlighting this ambiguousness. The changes in the track-changes manuscript (L361 - 362) are in red:

… To relate the CML and rain gauge time series, the ~~minimum~~ distance between the centre point of the CML and all rain gauge stations is ~~found~~ determined, and the data from the ~~corresponding~~ closest rain gauge ~~data~~ are ~~aligned~~used as a predictor. The calibrated tip records …

L312: I recommend adding a table in an appendix with the synoptic types, and how frequently each of the types occurred during the studied period to show how applicable this framework is to the range of different weather types. Also mention what is the temporal resolution of these synoptic types.

We complemented the text. The table is in the appendix (L593 - 597) and the changes in the track-changes manuscript (L371 - 372) are in red:

… Lastly, the daily data of synoptic type is determined based on the expertise of the Czech Hydrometeorological Institute (CHMI, 2024). The synoptic types are listed in the Appendix A.

L322: What is meant by “scales”? As in: puts it in perspective?

Thank you, that is a better term. The changes in the track-changes manuscript (L382) are in red:

The solid vertical line in Fig. 2 ~~scales~~ puts in perspective the other results, …

L327: Regarding the selected results, please state why only these are selected? Are these the most successful combinations of predictors? Perhaps worthwhile to add the other combinations as an appendix.

The changes in the track-changes manuscript (L386 - 387) are in red:

… (bars in Fig. 2, note that only selected results, which show interesting predictor combinations and also address at least one predictor related to the CMLs, are shown) …

L349: If the sample size is that important please also add the sample size used for generating Fig. 2 in the caption.

The changes in the track-changes manuscript (L412) are in red:

Figure 2: Conditional entropy of models built from full dataset using one to three predictors.…

L409-410: The fact that more training data leads to better results, is that the case for all predictors? In other words, could the need for a large sample size not simply be inherently related to the temporal scale of predictors like synoptic type or season?

The model relies highly on its dimensionality, as the inclusion of each additional predictor exponentially increases the number of possible state combinations. Consequently, an increase in the number of predictors leads to a corresponding rise in combinations.

The analysed model is built using the most informative combination of predictors. As demonstrated in Fig. 5, the simplest (one-predictor) model does not perform better with a larger training sample. Conversely, the model with more predictors requires a larger sample.

L379: This should be surprising since TL is used to manually identify the wet-dry timesteps right?

This is surprising as it demonstrates that there remains considerable uncertainty that can be explained by the other variables. Even though the total loss provides a lot of information about the presence of dry or wet conditions, the additional variables are required to identify the specific conditions more accurately.

L382: The entire wet-dry classification analysis relies on this one CML. As mentioned in the general comments this is a strongly limiting factor in the generalizability of this analysis.

Indeed. We state now clearly that it serves as an illustration of the information-theory framework to design a non-parametric model for CML processing. Please see the General comment section above.

L419: See general comments on the wet-dry classification and the use of other established wet-dry classification methods. Furthermore, these established methods were often calibrated using weather radar or gauges. It would at least be appropriate to discuss this difference and the effect that has on this comparison, as well.

The changes in the track-changes manuscript (L487 - 490) are in red:

… however, it works well when identifying stronger events. This can be explained by the origin of their approach, which was developed using weather radar data as a reference with a spatial resolution 1 km x 1 km and a temporal resolution of 5 min. The approach is thus influenced by (and reproduces) the bias and inaccuracies related to weather radar observation technique (see the subsection 5.3.2 Accuracy of reference rainfall in Discussion). In contrast, ...

L436: In case of rainfall events, subsequent timesteps can hardly be considered independent.

True, that is the reason why total loss data with temporal shifts t ± 1 and t ± 2 min were also tested as predictors. This approach was adopted to take into consideration the local variations in the course of the total loss and its efficacy in predicting the occurrence of the wet or dry conditions at timestep t. To be more specific, the changes in the track-changes manuscript (L510) are in red:

… by using the variable shifted in time, ~~as~~ e.g. total loss used as predictor in the second analysis in this study.

L429-444: An additional comment further elaborating on the interpretability of the results would be beneficial. For example, when listing the percentage reduced uncertainties (L329 3%, L332 6%, L337 1.5%, etc.) is this a statistically significant decrease? Also, the percentage decrease

in conditional entropy in Fig. 4 is much larger but the scale (x-axis) is much smaller. For the article to be self-contained, also for readers unfamiliar with information theory, a note on how conditional entropy in bits, for example, relates to the number of correctly labeled wet and dry timesteps would help. Such a comment can maybe best be made in the results section.

The advantage of the information theory framework is that it sets the minimum and maximum rate of uncertainty and the significance of the uncertainty reduction using predictors can be scaled to those worse and best uncertainty scenarios. The quantification of entropy in bits can be interpreted as the average number of binary questions required to determine a value of a variable. The distribution of the variable affects the measure of entropy.

This enables the definition of maximum and minimum entropy. The minimum uncertainty occurs when the variable concentrates in one bin, i.e. the entropy (uncertainty about the variable) is zero. On the other hand, the uniform probability distribution will require a maximum number of questions. The entropy is maximised, which sets the upper boundary of uncertainty.

The significance of the results is analysis-specific; therefore, the relative entropy is also presented instead of only measures of entropy in bits.

We suggest elaborating on this in the Information theory section as it fits better with the general information about the approach, instead of the result section. The changes in the track-changes manuscript (L153 - 158) are in red:

, where X is a discrete variable. Intuitively, the higher the entropy, the higher the uncertainty. Moreover, entropy cannot be negative, which would lead to counter-intuitive negative uncertainty. It measures the spread of the distribution, but in contrast to variance, it is not as influenced by data at large distances from the mean. On the contrary, it is mainly influenced by higher probability outcomes. Consequently, the distribution of qualitative variables can be incorporated in the models. Compared to variance it has ~~an upper~~ boundar~~y~~ies ~~(see the next subsection)~~. The minimum entropy is 0 bit, i.e. probability of the outcome is 1. The maximum entropy is associated with uniform probability distribution. …

L440: the comment on climate-specific constraints is appreciated since week or month of year may be of little significance in regions where the intra-annual variability is a lot smaller.

Thank you.

L459-460: It is appreciated that the authors acknowledge this lack of discussion.

Thank you. We also believe that this issue should be given greater prominence in CML research outputs.

L464: What is meant with “independent”? As in, an additional study?

‘Additional’ is a better term, in fact it was both independent and additional. The changes in the track-changes manuscript (L538) are in red:

… An ~~independent~~ additional study on the cross-validation of radar adjustment…

L478-484: It would be fair to acknowledge the availability of the data as well. For example, synoptic type has a different latency than CML attenuation. How readily available are all these environmental variables and are they able to be incorporated in near real-time?

The approach does not allow to incorporate near real-time data as the underlying assumption of the framework is the system stationarity and the data are transformed into multivariate, discrete distribution. However, we comment on the different latency of the data. The changes in the track-changes manuscript (L558 - 560) are in red:

… seasonality and synoptic type. Despite the demonstrated benefit of environmental predictors, it is relevant to acknowledge that the rainfall retrieval models based exclusively on CML data are not dependent on data with longer latency such as synoptic type. ~~It~~The presented approach concurs…

In addition, please also comment on/acknowledge the dependence of the variables used as predictors, and the influence this has on adding predictors. For example, temperature and month and week of year are not entirely independent variables. Is independence accounted for in the current framework?

If the target and the predictor are independent, the predictor will not decrease the uncertainty about the target, i.e. the conditional entropy will equal to the unconditional entropy. Therefore, potential dependence of variables is inherently embedded in the framework.

Technical corrections:

L293: Conversed as in processed?

The changes in the track-changes manuscript (L352) are in red:

…, the raw and ~~conversed~~ processed data are used as predictors.

L301: The sampling …? area?

The changes in the track-changes manuscript (L360) are in red:

…The sampling area is 500 cm^2 and the resolution is 0.1 mm…

L445: This sentence is not clear to me.

We have removed the entire paragraph as it was deemed unnecessary.