Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2020-43-AC1, 2020

Atmospheric Measurement Techniques Discussions

Interactive comment on "Commercial Microwave Links as a tool for operational rainfall monitoring in Northern Italy" *by* Giacomo Roversi et al.

Giacomo Roversi et al.

federico.porcu@unibo.it

Received and published: 20 May 2020

Anonymous Referee 1

Review of the paper "Commercial Microwave Links as a tool for operational rainfall monitoring in Northern Italy" submitted to AMTD by Roversi et al. in 2020.

This manuscript is the first one to analyze the performance of commercial microwave links (CMLs) with a large data set (350 CMLs) in Italy. The applied processing algorithm is based on an open-source state-of-the-art method (RAINLINK) which has only been modified slightly. The main contribution of the manuscript hence is the analysis and discussion of the unique CML data set for the target region. The authors also give some new insights into how CML

rainfall estimates perform under certain circumstances. Since CML data is not easily available for analysis in most countries the current manuscript provides an important contribution to the scientific community working on rainfall estimation and is certainly interesting for readers of AMT.

The manuscript is well structured and the applied processing methods are sound. Writing should be improved, though. I also have two major general comments regarding the analysis of the results which will require a major revision of the manuscript. I also have some minor and specific comments.

We thank the reviewer for the positive introductory remarks and the careful review of the manuscript. We addressed all her/his comments point-by-point in italic, outlining how we are modifying the manuscript.

Major general comments and suggested major changes:

1. The main limitation of this study is that it lacks comparability to other studies because the quantitative analysis of the skill of the produced CML rainfall maps is only carried out for a subset of the data, namely the data pairs where the reference and the CML rainfall is > 0.1 mm/h. None of the other studies that use the RAINLINK algorithm and similar CML data (15 minute min/max) use this approach (see Table A1 in de Los et al. 2019 DOI: 10.1175/JTECH-D-18-0197.1). This also limits the interpretability of the results in this manuscript since the effect of bad FAR and POD, which lead to overestimation (high FAR) and underestimation (low POD), cannot be studied in the resulting rainfall fields. I strongly suggest to carry out the analysis of the rainfall fields for different subsetting variations. The most commonly used ones for comparing rainfall maps seem to be: 1. No threshold 2. Reference > 0.1 mm. This does not mean that all the plots have to be done several times, but at least the main skill metrics should be provided for the different subsets.

We agree that for the comparison with the previous studies a set of indicators with the *same (or similar) settings is needed. We will compute continuous indicators for the set with no threshold and with Reference > 0.1 mm/h and we will add them to Table 2 in the new manuscript. For the categorical scores instead no filtering was performed. The threshold of 0.1 mm/h was used there to discriminate wet and dry samples in the confusion matrix. We acknowledge that it could have been unclear from the text and we will rephrase the paragraph for better understandability.*

2. Since a large part of the quantitative analysis is based on interpolated rainfall maps, I strongly suggest to show several examples of interpolated CML rainfall maps, e.g. of one or two specific events and e.g. accumulated over the whole period.

We welcome this suggestion and we will add a section where two case studies are reported (e.g. the best and the worst), including the discussion of meteorological conditions, reporting interpolated maps and indicators.

Minor general comments:

- Choice of POD and FAR: I assume (since it is not specified in the manuscript I looked at other papers that use RAINLINK) that POD is hits/(hits+misses) and FAR is falsealarm/(hits+falsealarm). If this is the case, POD is the true positive rate (TPR). Wouldn't it than be better to use the false positive rate (FPR), like it is used in the ROC curve, instead of FAR. FPR and TPR are both normalized by the reference conditions. FAR instead is normalized by the predicted positive conditions. Can you elaborate on this choice?

The reviewer understood correctly: POD=hit/(hits+misses) and FAR=falsealarm/(hits+falsealarm). This choice to favour FAR over FPR derives from its common use in deterministic precipitation forecast/estimate validation (Tang et al., 2020; Petracca et al., 2018; Puca et al., 2013, McBride and Ebert, 1998, among others). The FPR= falsealarm/(falsealarm+correctnegatives), more common in probabilistic forecast verification, is heavily influenced by the most populated category

(correctnegatives): in case of small scale (or rare) rain pattern, FPR can decrease without any skill in the forecast since the no-rain condition is the most common in the target area. FPR, for the same reason, could also be misleading when different seasons/climates, with different rain occurrence, are compared.

- The writing needs improvement throughout the manuscript, in particular the introduction and conclusion. Hence, I stopped very early to note down technical corrections and suggestions for stylistic improvements when reading the manuscript.

We will improve the revised manuscript through a careful review of the language.

Specific comments:

L28: It would be good to have another or additional reference for the claim that the "last generation polarimetric systems have only partially mitigated“ the radar QPE problems. The book by Ryzhkov and Zrnic, 2019 is certainly a very valuable textbook, but is is hard to find this conclusion in a 450 page reference. Access to it might also be limited.

We agree with the Reviewer and we'll include few references more focused on the QPE of polarimetric radar performance evaluation (Figueras i Ventura et al., 2012; Gou et al., 2019; Cocks et al., 2019).

Figueras i Ventura J, Boumahmoud A-A, Fradon B, Dupuy P, Tabary P. 2012. Long-term monitoring of French polarimetric radar data quality and evaluation of several polarimetric quantitative precipitation estimators in ideal conditions for operational implementation at C-band. Q. J. R. Meteorol. Soc. 138: 2212–2228. DOI:10.1002/qj.1934

Gou, Y.; Chen, H.; Zheng, J. Polarimetric Radar Signatures and Performance of Various Radar Rainfall Estimators during an Extreme Precipitation Event over the Thousand-Island Lake Area in Eastern China. Remote Sens. 2019, 11, 2335. https://doi.org/10.3390/rs11202335

Cocks, S., L. Tang, P. Zhang, A. Ryzhkov, B. Kaney, K. L. Elmore, Y. Wang, J. Zhang, and K. Howard, 2019: A prototype quantitative precipitation estimation algorithm for operational S-band polarimetric radar utilizing specific attenuation and specific differential phase. Part II: Performance verification and case study analysis. J. Hydrometeor., 20, 999–1014, https://doi.org/10.1175/JHM-D-18-0070.1.

L32: ": : :accuracy is still under evaluation (Tan et al., 2018): : :". This statement is a bit weak. In addition, there are many studies that evaluate the performance of IMERG, also on a broader level than Tan et al., 2018.

Of course, many papers are dealing with satellite product validation, but very few of them deal with high resolution (hourly) data, mostly focusing on daily to annual integrals. We rewrite more precisely the sentence "their accuracy is difficult to assess at high spatial and temporal scales, depending on local climatology (Tang et al., 2020)", and included a more recent and pertinent reference.

Tang, G., M. P. Clark, S. M. Papalexiou, Z. Ma, Y. Hong, Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets, Remote Sensing of Environment, Volume 240, 2020, 111697, https://doi.org/10.1016/j.rse.2020.111697.

L34: I think "broad diffusion" is the wrong term here. Something like "ubiquity" would fit better.

"ubiquity" means everywhere, including ocean and desert, we'll cut the adjective, leaving: "...in the last decades with the diffusion of microwave...".

L36: "Accurate algorithms were introduced to measure : : : drop size distribution : : : water content". Since the sentence before talks about CMLs, the used references do not fit here, since they did not use, or only partly used, CML data. Dual-frequency and dual-polarization data, as used in the references, is mostly not (yet) available in operational CML networks.

The reviewer is right: the sentence was not correctly contextualized, and we reworded to: "Accurate experiments and numerical simulation were used to assess the capability of microwave links to measure average rainfall rates (Rahimi et al., 2003), drop size distribution (Rincon and Lang, 2002; van Leth et al., 2020) and water content (Jameson, 1993). On the same token, the possibility to have a spatially continuous rainfall pattern depends on the density and distribution of the links, making this approach of particular interest for urban areas... ".

L38: ": : :a spatially continuous rainfall path: : :" It is not clear to me what that means. Please rephrase.

The word "path" is replaced by "pattern".

L50: This is a very long and confusing sentence.

We'll remove this sentence in the new version of the manuscript.

Section 2.1: What is the power quantization of Pmin and Pmax? Please specify.

The quantization for Power is 1dB, we'll report this number in the revised manuscript.

Fig1. and section 2.1.1: Are there pixels without a CML, i.e. LC = 0. This is not clear from the map and the text. Please clarify. If yes, what are the implications. E.g. if you would have to interpolate a rainfall field over two empty pixels in the west of Parma that would decrease the performance a lot compared to pixels that at least have one CML.

We changed the colour scale of Figure 1, to make clear the presence of few LC=0 grid boxes. Nevertheless, we do not think that cells with LC=0 represent an issue because we aim to evaluate an interpolated product whose goal is precisely filling the empty gaps between separate measurements. Previous CML papers also show rainfall maps interpolated at a finer scale (1 km) and with sparser and more inhomogeneous CML networks (e.g. Overeem et al., 2016). Besides, we agree that better results are likely to be expected from regions with higher coverage. We already address the matter *throughout the analysis of the LC dependency.*

L170: It would be good to know what "set of consistency checks" has been used. Is everything done as in Overeem et al 2016? Even if yes, a short summary (2-3 sentences) would be good so that the reader does not have to go through the explanation in the reference.

The consistency criteria require that: the frequency is inside a specified range; there are no multiple occurrences for the same ID and DateTime, every ID has always the same geographical coordinates, not-available (NA) entries are not present. We will add a sentence that clarifies this point and possibly add some other algorithm information in the Supplement.

L174: What is "a comparable decrease"? Please be more specific.

Wet-Dry Classification is described in Appendix C of Overeem et al. 2016 and we used exactly their procedure. The description was here treated only qualitatively on purpose: we will add an explicit reference to the appendix and we will specify that the values inside of the NLA classification are left unchanged.

L178: Also here, it would be good to get more info on the outlier filters. What exactly was done? And even more important. How much data was removed?

We will add some additional details on the procedure. Also, statistics on outliers will be added to the revised version.

L186: Did you use specific a and b values from van Leth et al, 2018? If not, it is not clear why this is cited here. Please cite the source of the a and b values.

Van Leth et al.(2018) is cited here only to support the assertion about which variables the a and b parameters are sensitive to. A proper description of which parameters are utilized in our work is reported in the following Section 3.2. We will add some internal references to make that clearer.

L199: What was the length of the CMLs below 10 GHz? Even at 30 km (the maximal length in your data set according to section 2) a CML with 5 GHz is very insensitive to rainfall (approx. 0.05 dB at 1 mm/h path-averaged rainfall) so that light to moderate rain might not cause a detectable signal. Can you make sure that this does not have negative effects on the rainfall fields for light and moderate rainfall events (in the range 1-10 mm/h)? Couldn't it be that CMLs with zero rain rate are introduced in the interpolation method, which would better be left out? How much CMLs would you loose if you do not include CMLs below 10 GHz and how much does the "spatial coverage" decrease?

This is a very important point and we thank the reviewer for having it highlighted. In our network, we have only five links between 5 and 10 GHz. We will remove them and we do not expect any major change in the results. We also are investigating the sensitivity for all links, and we'll discuss this in the revised paper.

L209: My feeling as a non-native speaker is that "delineate" is the wrong term here.

We replaced "delineate" with "detect".

Section 3.3: It is not clear from which reference you took which skill indicator. In my opinion, it would be best to define the skill indicators here to avoid any misconceptions.

We'll add the description of the indicators to be clearer.

L214: Complicated sentence and unclear formulation. I guess you are trying to say that your CML and reference products have a lot of rain rate that are so low that they can be neglected in any application.

The reviewer understood correctly. However, we reworded the sentence to "Both interpolated CML and reference field have a large number of very low positive values (below 0.1 mm h-1) that are not of interest in any application, but which are potentially *very influential in normalized error metrics".*

Fig. 2: It would be good to know the minimal distance from the individual CMLs to the reference rain gauge.

We added, as further information on the figure, the minimum distance between gauge and link. Discussions around this distance are also going to be added to Section 4.1 and following.

Fig. 3: This figure contains a lot of useful information. It is a bit unstructured, though. It could be cleaned up by aligning the x-axis of each column and by sharing the legend in row 1 and 2. In row 3 two columns for "back" and "forward" could be used. Reusing the colors from row 2 in row 3 for different variables is also not ideal. The x-axis tick labels are also different in row 3 from row 1 and 2, so that it is not clear if the depicted periods are exactly the same. Hence, in particular the alignment of the x-axis would help. If you redo the plot, which is what I would suggest, than you could also reconsider the order of the rows. I feel that starting with the raw data (now row 2) would make more sense since this follows the CML data processing workflow. The meaning of the pink horizontal band, explained in the text, should also be nna explained in the figure caption.

We are considering the re-design of Figure 3, following the reviewer's advice. We were aware that this figure could appear unclear, and thank the reviewer for the suggestions.

L262: Can the overestimation of the one CML be explained by the spatial distance between this CML, the other CMLs and the gauge? If not, what is your explanation?

We'll include the distance information in the revised manuscript, and consider also other causes of uncertainties in the matching.

L269: Does this CML show these differences between Pmax and Pmin during the whole period? If yes, are there other CMLs that show something similar?

Do you have any explanation or mitigation strategy? Regarding an operational application of CMLs there should be a way to deal with this kind of signals.

This is to our understanding a common feature to all CMLs and, more generally, to the 15 minutes MinMax sampling strategy. We will say that more clearly in the revised manuscript. We will also address the connection of this feature to the manual ATPC correction which involves only Pmin.

L281: If I understand correctly this data set is not part of the data set for the main analysis of the paper, correct? Please clarify in the text.

Section 4.1.3: It would be important to know the height of the antennas and the estimated height of the melting layer or zero-degree level. Form the fact that the data is from the month of March is cannot be concluded that the CML measured mixed-phase or solid precipitation.

L288: ": : :"bright band in the radar reflectivity maps and is thus easily detected". If you have a dual-pol radar with a working hydrometeor classification, then yes, it can be detected. If not, than this is quite hard to do for smaller scale precipitation events and on short temporal durations. I suggest to add some more details to the explanation in the text.

Of course, we have polarimetric data observations and assessed the presence of bright band without any doubt. However, after the suggestion of the other reviewer also, we decided to drop this section, being a little out of the main objective of the work.

L296: Is this underestimation due to missed event or to a general underestimation of the CML rain rates? And why didn't you try to adjust the wet antennae compensation to compensate this underestimation? Overeem et al. 2013 and 2016 calibrated the wet-antenna compensation for a specific subset of their data, so it might neither be optimal nor applicable to your data. Please explain.

The first question could be addressed by comparing the overall values ME (-26%)

indicating the relative deficit of measured rain amount with the MB (0.7), the relative occurrence of estimated wet samples with respect to the real number of wet samples. The underestimate seems to affect the 30% the number of "events" and a little bit less (26%) the amount of water. From the conceptual point of view, however, the two things are tightly connected: the underestimation of the rainrate results in an underestimation of rain occurrence, as soon as the underestimate affects rainrate values just above the threshold. We will add a comment on this.

As for the second issue, our feeling is that the reference data we considered (used in operational offices) are not suitable to be used as calibrator, in term of quality and spatial and temporal characteristics, as also the other reviewer remarked. Anyway, we are running some trials with decreasing Aa, and will report and discuss their results in the revised manuscript. We already show how an overestimation of Aa could affect the algorithm (figure 3), and indicate the Aa high value as a possible issue to address, once an experimental facility with the necessary accuracy would become available.

L298: Since your study and the two other studies all use a different "Filter" (see your Table 2) the results are not really comparable. In particular your choice of "Ref. AND Product > 0.1 mm h-1" neglects the negative effects of false positives and false negatives. See my major comment above.

We agree with the reviewer and we will provide a new set of indicators' values suitable for comparisons with previous works, and added a column in Table 2.

L318: "The accuracy in the estimates is reached at the expense of POD, ETS and BIAS: around 50% of the rainfall duration is lost in this area". I understand that when FAR is lower (mentioned in the sentence before) and POD is lower there are less rain events, both correct and incorrect ones, in the resulting CML rainfall time series. That would explain that there is even more tendency to underestimate here. But, if I understood correctly, the bias is only calculated from values where both CML and reference are above 0.1 mm/h, so that false and missed CML

rain events have no impact on the calculation of the bias. Can you elaborate on that?

First, we made a mistake: "BIAS" (undefined in this work) stands for Multiplicative Bias (MB), i.e. numberofestimatedwet)/(numberofobservedwet). Hence a Multiplicative Bias of 0.47 indicates that only half of the wet samples are found, where for wet we mean rain depth > 0.1 mm/h (both estimated and observed). However, the amount of rain lost in this area (given by ME) is similar to other areas, and the indicators of numerical accuracy of the estimates (CV and CC), computed on wet-wet samples, are quite high. This indicates that in this area the rainrate is estimated with higher accuracy, while the discrimination wet/dry is worse. Categorical scores are calculated on the unfiltered datasets, around the filtering threshold.

L324: Remove the "For" at the beginning of the sentence

Ok, thanks.

L327: ": : :this suggests that LC is probably not the only variable at play there". This is good to know, since that would have meant that regions with high CML density perform bad with the used algorithm. The CML data set of Overeem et al 2016 also has regions with a very dense network and regions with a coarser network. Hence, a strong dependence of the RAINLINK algorithm on LC should have already been noticed by them. Could it be that there is one CML in this area that shows "strange" behavior, e.g.strong fluctuations, that negatively affects the POD of the many surrounding CMLs by not letting RAINLINK do the detection of rain events?

Overeem et al. (2016) showed how the CML performance varies against the mean link density (our "LC") by analysing normalized variance and correlation on a 74km2 grid. Our results for CV and CC estimated at 25 km2 are in good agreement with them. In addition to their work, we show also the effect of LC on the categorical indicators, providing some interesting results for the FAR especially and giving more insight into the *topic in general. However, we were not able to isolate all the sources of uncertainties and to gauge the performances of the single links individually. We'll better express this peculiarity of our work in the revised manuscript.*

L355: Since your reference data set ERG5 is an interpolated rain gauge product, it might miss small scale rain events compared to the radar. Assuming that clutter removal was done in a sufficiently good way, the radar should not have a high FAR in general. Couldn't the fact that ERG5 might miss some real rain events explain the high FAR of the radar product?

The clutter is removed through a static map of clutter, a beam trajectory simulation and an anomalous propagation cancellation, (see Fornasiero et al., 2006). Moreover, WiFi/WiMax signals are filtered through a decision tree and a fuzzy logic techniques which expoitZ, Zdr, W, V and Z and Zdr variance. We do not think therefore that the clutter is the reason for the high FAR. We suppose instead, as the reviewer pointed out, that a reason for high false alarms ratio could be that ERG5 misses some small scale events. We'll modify the sentence: "...while rain gauge (as well as the reference product ERG5) and CML networks...".

L362: ": : :making CML a more robust sensor." Robust in what sense? Please explain in more detail in the text.

We want to point out that to fully exploit radar capabilities a proper Z-R relationship should be used, while the CML k-R relation is almost not dependent on DSD due to its linearity in this frequency range (Leijnse et al., 2008). We added few words to the sentence: "thus making CML a more robust sensor, in the sense that the coefficients to retrieve rainrate do not depend on the DSD".

bf L365: When speaking about the "operational context" and the advantages of CMLs it should be discussed how the low POD, found in this study, affects the CML's potential for operational applications. This should be part of this paragraph.

We'll specify better the possible role of CML in the operational context, modifying the sentence to: "In an operational context, where several precipitation products (each one with its proper error structure) are available to the forecaster, it is of great relevance the latency of the precipitation product, i.e. the time taken from the acquisition of the basic data (the occurrence of the event) and the delivery of the product in a ready-to-use form." And at the end of the paragraph, we added the comment: "It is to the operators' preference, based on product accuracy and current meteorological conditions, to make use of the most suitable product".

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-43, 2020.