Author response and changes in the revised manuscript of "Commercial Microwave Links as a tool for operational rainfall monitoring in Northern Italy" by Giacomo Roversi et al. (2020)

We would like to thank both Anonymous Referees for the careful review, the comments, and the suggestions on how to improve our work. We report the responses to all their comments point-by-point in italic, also indicating how we modified the manuscript. The integral revised manuscript marked-up with track changes is attached below the answers.

Response to Anonymous Referee #1

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 computed continuous indicators with the filter set as Reference > 0.1 mm/h we added them to Table 2 and the sentence "To make easier the comparison with past works, we computed continuous indicators with the filter set as Reference > 0.1 mm and with no filtering at all. Results with the first setting yields slightly worse indicators, increasing the ME to -0.41 and the CV to 0.95, with a second digit increase of R2, around 0.5. The no-filter run shows values of ME and R2 in line with our original results, while CV is greatly affected by very small rainrates." in the reviewed manuscript.

For the categorical scores, instead, no filtering was ever performed: the threshold of 0.1 mm/h was used to discriminate wet and dry samples in the confusion matrix. We stated this explicitly in the revised document: "Thus we have set a wet-dry threshold equal to the minimum rain quantity detected by the tipping bucket rain gauge, i.e. 0.1 mm h^{-1}, for both estimate and reference". More information is provided answering the specific comment on L298.

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 added a section (4.2.2 in the revised manuscript) where three case studies are reported, including the discussion of meteorological conditions, reporting interpolated maps and indicators. Moreover, in the Supplement, we discuss the cumulated map.

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 false_alarm/(hits+false_alarm). 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 Referee understood correctly: POD = hits / (hits + misses) and FAR = false_alarms / (hits + false_alarms). 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= false_alarm/(false_alarm+correct_negatives), more common in probabilistic forecast verification, is heavily influenced by the most populated category (correct_negatives): in case of small scale, sparse or intermittent rain phenomena, 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 improved 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 Referee, and we included few references more focused on the QPE of polarimetric radar performance evaluation: Figueras i Ventura et al., 2012; Gou 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

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 rewrote the sentence to be more precise:"… their accuracy is difficult to assess at high spatial and temporal scales (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 the 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.

We agree and replaced "broad diffusion" with "ubiquity".

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 Referee is right: the sentence was not correctly contextualized, and we reworded to: "Accurate experiments with high-quality links 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 reconstruct a spatially continuous rainfall field relies on a sufficiently high density of the links, making the CML 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 "field".

L50: This is a very long and confusing sentence.

We rewrote the sentence to "Even if the general relationship between signal attenuation and rain rate is already well established, the successful use of CML data to quantitatively monitor precipitation still depends on the quality and technical characteristics of the transmitted power data and the fine-tuning of the algorithms."

Section 2.1: What is the power quantization of P_min and P_max? Please specify.

We reported the quantization value (1dB) in the revised manuscript: "…are measured by the provider with the resolution of 1dB at a frequency of ten…"

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 and we already addressed the matter throughout the analysis of the LC dependency in Section 4.2.1. Near the end of that Section, we added: "It seems that the sensitivity to LC could explain the improvement in the FAR of the RRB area, but not the sharp decline in the POD, suggesting that LC is probably not the only variable at play there. In Reno basin. These results integrate the findings of Overeem et al. (2016b), that highlighted the positive impact of higher LC on CV and CC at lower spatial resolution."

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 always has the same geographical coordinates, not-available (NA) entries are not present. We added a sentence that clarified this point and added some statistics about the rejected data in the Supplement. Paragraph 3.1.1 now reads: "**1. Preprocessing:** the raw input goes through three consistency checks concerning data formatting and labelling. Any multiple observations for the same LinkID and DateTime are discarded, each LinkID is verified to maintain the same metadata throughout the whole dataset (Frequency, PathLength and antenna coordinates) and rows with NA values in any of the columns except for Polarization (which is supposed vertical if not indicated) are discarded as well.."*

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, but we now added quantitative references to increase clarity.

We modified the point 3.1.2 as follows:

*"2. **Wet-Dry Classification**: the samples are discriminated in wet and dry periods by assuming that rainfall is correlated in space, through the so-called Nearby Links Approach (NLA), which works as follows. For each link, a time interval with a decrease in the received power is labelled as wet if at least half of the links in the vicinity (within 15 km radius) experience a comparable reduction, i.e. if the medians of the attenuation and the specific attenuation of the nearby links are below – 1.4 dB and – 0.7 dBkm-1 respectively. This is the second most computationally time-consuming step of the algorithm.*

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 added some additional details on the procedure. Also, statistics on outliers will be added in the supplementary material.

We reworded the point 3.1.4 as follows:

*"**4. Outliers filter and power correction:** outliers due to malfunctioning links can be removed again by assuming that rainfall is correlated in space. The filter discards a time interval of a link for which the cumulative difference between its specific attenuation and that of the surrounding links over the previous 24 h (including the current time interval) becomes lower than the outlier filter threshold, which is fixed at -32.5 dBkm-1h. After removing the outliers, the classification information is used to clean the receiving powers of the noise over the dry periods. The corrected powers P_i^{Cor} will be equal to P_{ref} on dry periods and P_i on wet ones."*

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. To avoid misunderstanding, we removed the Van Leth et al. citation and completed the sentences mentioning the source of a and b: "are finally calculated using the k-R relationship $R = ak^b$, where the coefficients a and b were from Leijnse 2007 and Lejinse et al 2010 for vertical and horizontal polarization respectively."

> *Leijnse, H., 2007: Hydro-meteorological application of microwave links - Measurement of evaporation and precipitation. PhD thesis, Wageningen University, Wageningen. See page 65. Provided for frequencies from 1 - 100 GHz.*
>
> *Leijnse, H., R. Uijlenhoet, and A. Berne, 2010: Errors and uncertainties in microwave link rainfall estimation explored using drop size measurements and high-resolution radar data. J. Hydrometeorol., 11 (6), 1330–1344, doi:https://doi.org/10.1175/2010JHM1243.1.*

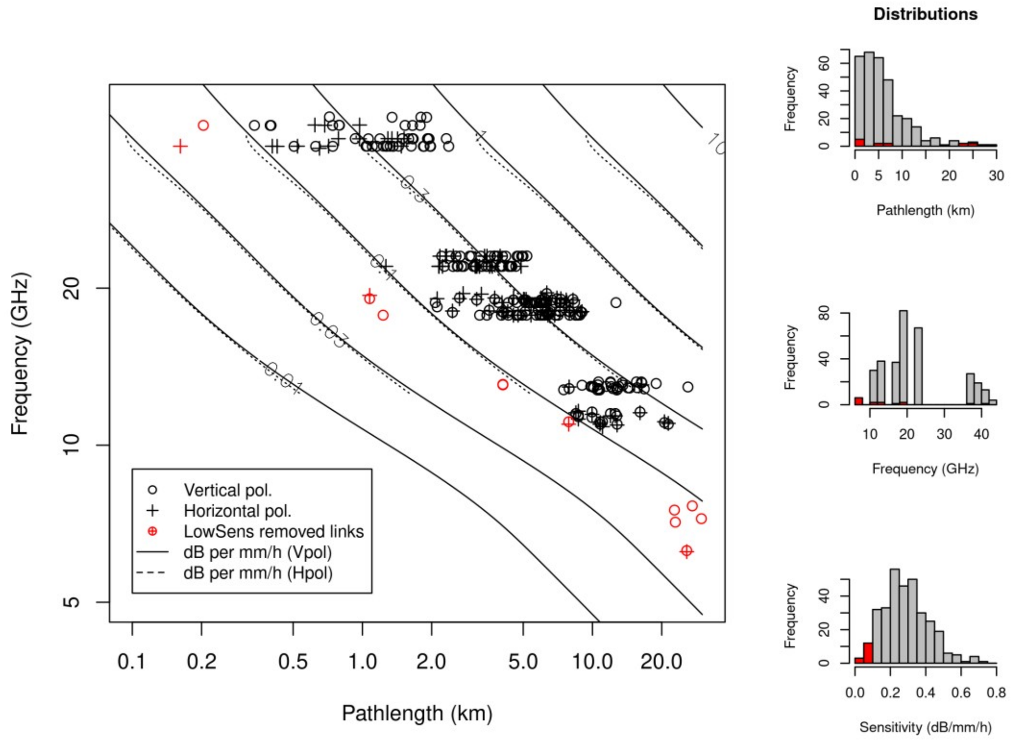
L199: What was the length of the CMLs below 10 GHz? Even at 30 km (the maximal length in your data set according to section 2) a CML with 5 GHz is very insensitive to rainfall (approx. 0.05 dB at 1 mm/h path-averaged rainfall) so that light to moderate rain might not cause a detectable signal. Can you 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 Referee 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. A complete statistic on the CML characteristics is now presented in the Supplementary Material.

Moreover, we went deeper into this analysis, following also the comments of the Anonymous Referee #2: we investigated the sensitivity for all links, calculating the theoretical sensitivity through the inversion of the kR relationship at a fixed 1 mm/h rainrate. It has to be remembered here that the manipulations within the algorithm (especially the Aa threshold) do not allow a direct translation of the theoretical sensitivities into actual instrumental uncertainties or error bands.

The analysis is presented by means of the plot below, where all links are distributed according to their length (x-axis) and frequency (y-axis) and where the theoretical sensitivity field is showed as contour lines of equal sensitivity with small differences for the two polarizations.

We decided to remove from the dataset all the 15 links with a sensitivity below 0.1 dB per mm h^{-1} (among which are the 5 low-frequency ones), here highlighted in red. We added the following sentence to the revised manuscript: "The CMLs' operational frequency in our region spans between 5.0 and 45.0 GHz. We decided to extend the default frequency allowance window from 12.5 - 40.5 GHz (as was in the Netherlands) to 10.0 - 45.0 GHz, leaving out five low-frequency CMLs. We also removed from the dataset 10 other links with higher frequencies but with sensitivities below 0.1 dB per mm-1 (see Supplement for more details). This is done to avoid contamination by coarse low sensitivity signals."



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 added the description of the indicators to be more explicit, with the following sentence: "In the present work, we selected two sets of classical skill indicators, broadly used in the validation community (Nurmi, 2003): the first one is to assess the capability of the product to detect rainfall occurrence (categorical indicators), and the second one is to evaluate the skill in estimate correctly the quantitative precipitation rate (continuous indicators). The first set is computed after a definition of a confusion matrix by counting the number of samples where both estimate and observation agree on classifying wet (hits, H), or dry (correct negatives, CN) samples, and where there are misses (M, observed wet and estimated dry) or false alarms (F, observed dry and estimated wet). Namely, Probability of Detection is defined as POD=H/(H+M), the False

Alarm Ratio as FAR=F/(H+F), the Multiplicative Bias as MB=(H+F)/(H+M) and the Equitable Threat Score as ETS=(H−Hrnd)/(H+M+F), where Hrnd represents the number of hits obtained by chance. Given ei *and* oi *as estimated and observed values resp., continuous indicators are the normalized Mean Error, defined as* ME = Σ(ei − oi)/ō, *the normalized Mean Absolute Error, defined as* MAE = Σ|ei – oi|/ō), *the Coefficient of Variation (CV) defined as the root mean square error divided by the mean of the observed values* ō, *and the Pearsons' Correlation Coefficient (CC), as the covariance of observed* oi *and estimated values* ei *divided by the product of the two standard deviations (Nurmi, 2003, Overeem et al., 2016b).”*

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 Referee understood correctly. However, we reworded the sentence to improve clarity: “Both interpolated CML and reference field have a large number of very low positive values (below 0.1 mm h^{-1} that do not have any physical relevance, 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.

In the legend of the revised Figure 2 are now indicated the shortest distances between the link paths and the respective raingauges. The distance is always shorter than 3 km (well below any decorrelation length of precipitation in our region) and still under 70% of the respective link length (see figure below). Further inspection showed that there is no relation between the distance from the raingauge and the link performance. We added a sentence in the revised manuscript: “We have selected links in rural areas and different terrains with an active raingauge close to the link: the distance between link and raingauge, reported in Figure 2, is always below 3 km (significantly lower than the correlation distance of precipitation in Italy (Puca etal., 2014)) and always lower than the length of the link itself. In general, no dependence of the link performance on the distance from the raingauge is found.”

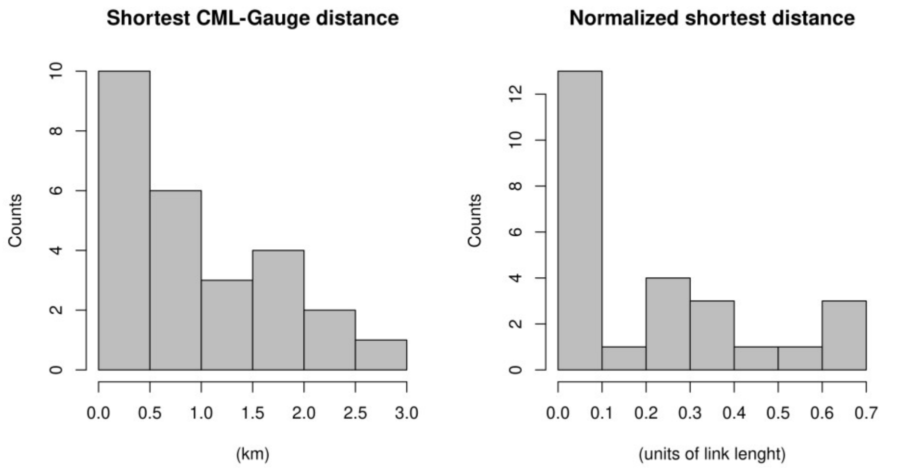
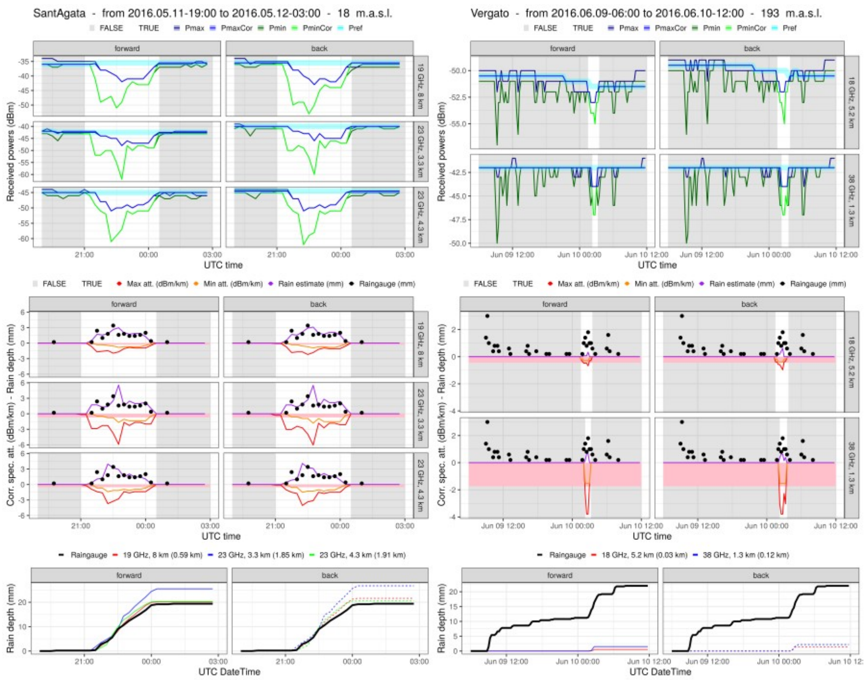


Fig. 3: This figure contains a lot of useful information. It is a bit unstructured, though. It could be cleaned up by aligning the x-axis of each column and by sharing the legend in row 1 and 2. In row 3 two columns for "back“ and "forward“ could be used. Reusing the colors from row 2 in row 3 for different 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 re-designed Figure 3, following the Referee's advice, which was much welcome. We also removed the bright-band case following comments on L288 and on Section 4.1.3 by the Anonymous Referee #2.



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?

The spatial distances between the Sant'Agata links and gauge are all similar and very short. Many factors could concur in poor performances, but, unfortunately, we are not able, with the current reference dataset, to dig deeper into this issue. For example, different links orientations could lead to different wetting rates of the antenna's radomes in case of winds, or very local phenomena (as hails or showers) could perhaps hit one link and not the others nor the raingauge. Specifically, the mentioned overestimation seems generated by a slightly stronger peak in the middle of the event, common to both sublinks of the CML, located 1.85 km away from the raingauge. The three CML are so close together that the issue is probably wholly wiped out by the smoothing of the interpolation process. We added this comment: "cannot be certainly related to any macroscopic characteristics of the three links".

L269: Does this CML show these differences between P_max and P_min 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.

The gap between Pmin and Pmax is to our understanding a characteristic common to all CMLs and to the MinMax sampling strategy. It happens every time there is a variation of receiving power within the time *interval. If the gap persists in time, then there are probably some power fluctuations that must have a frequency higher than 15 min^{-1} and an amplitude equal to the gap. This seems to us more a feature of the MinMax sampling strategy rather than an issue to be mitigated or dealt with. We added it more clearly in the revised manuscript: "Looking at Figure 3d, Pmin does show a decrease coupled to the missed rainfalls, but Pmax does not. This behaviour of Pmax is not an issue itself, as the NLA classification relies on Pmin only, but it indicates that there are power fluctuations which happen faster than 15 min-1. Rapid fluctuations, in turn, suggest irregular and scattered precipitation patterns, that actually could be a factor that affects the correct classification, since NLA relies on the spatial correlation of the rain field. Therefore, a Pmax signal always near the baseline could be a precursor of local NLA issues."*

Nevertheless, we briefly verified that, if rain is sensed from the algorithm, Pmin and Pmax are likely to be distinct, while the opposite is not true (there are cases of Pmin-Pmax differences, sometimes very relevant, not associated to rain or outliers). There is also no particular correlation between the gap width in dB km^{-1} and the estimated rain amount in mm.

Sensitivity analyses (see also answer to comment on L296) showed optimal values for alpha (the parameter which weights between retrievals from Pmin and Pmax) very similar to the default one (0.28 vs 0.33).

The ATPC corrections (see the answer to comment on Section 2.1.2 by Anonymous Referee #2) mostly amplifies already existing Pmin-Pmax differences, which are moreover sensibly bigger than the correction itself (10-40 dB against 4-6 dB respectively). Lastly, the increased spread will not directly affect NLA, as the classification exploits only Pmin observations.

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.

That is correct, the Referee understood correctly. See answer the following two comments for changes in the revised manuscript.

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.

The geographical height of the ground under the antenna is known to the authors (not the one of the pylons themselves) but confidentiality restrictions with the data provider prevent us from being more precise. Independent instrumental measurements of the freezing level in the atmosphere are known too, and they were carefully taken into consideration before discussing the melting layer case, but we chose not to show them to avoid cluttering the Section with data not related to the main discussion. For similar reasons, we decided eventually to remove the melting layer case form the revised paper. See also the following answer to comment on L288.

L288: ": : :âA˘ Žbright 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.

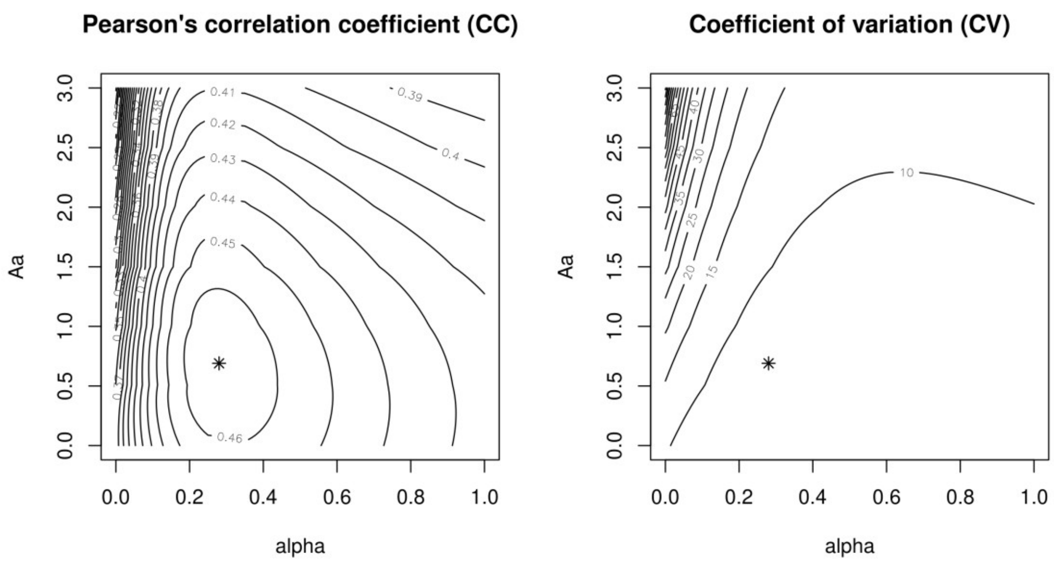
We based our statements on polarimetric data observations and assessed the presence of bright band without any doubt. However, after the suggestion of the other Anonymous Referee also, we decided to drop this section about the melting layer entirely, being a little out of the main direction of the work.

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 (0.77), the relative occurrence of estimated wet samples with respect to the real number of wet samples. The underestimation seems to affect for a 23% the number of "events" and a little bit more (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.

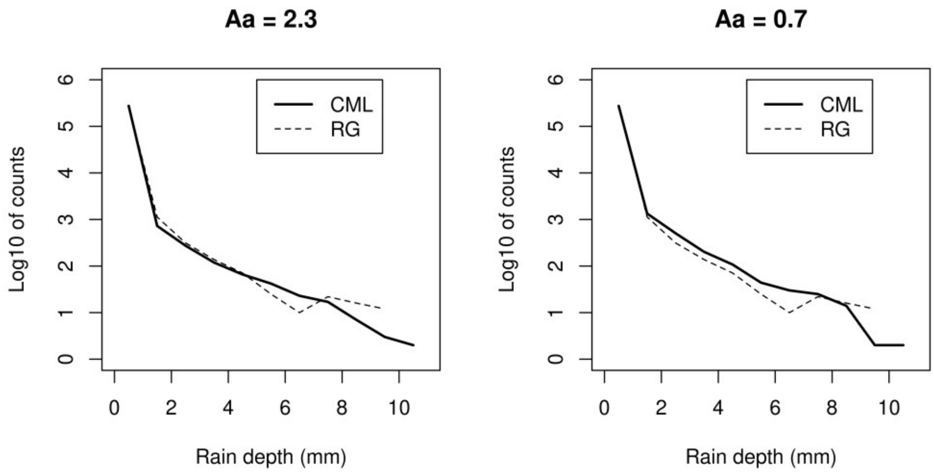
As for the second issue, our feeling is that the reference data we considered (used in operational offices) are not suitable to be used as a calibrator, in term of quality and spatial and temporal characteristics, as also the other Referee remarked. Anyway, we performed some trials with decreasing Aa, for the single-link analysis.

First, we checked for the 27 links with a close-by 15-min raingauge in which Aa and alpha values were producing the best overall performances, and the results are summarized in the following figures.



The CC surface (left) shows a clear maximum at: alpha = 0.3, Aa = 0.7, while CV (right) reaches no local minimum in the examined domain but has a plateau-like area of fair performance in which falls the best match for CC. This analysis suggests an Aa value much smaller than the one used by RAINLINK and in our work (2.3 dB), while the alpha parameter remains almost unchanged. However, looking at the PDF of the estimated rainfall emerges that the physical representativeness drops down with the lower Aa. Transferring these results on the whole datasets also leads to worse overall results: CV worsens of +0.05 with no sensible improvement on R^2, the same for FAR and ETS, which change resp. of +0.08 and +0.01.

In our opinion, this indicates that the major criticalities of the algorithm are to be found somewhere else (probably in the classification process or in the outliers filter), and that they should be addressed first, before fine-tuning these retrieval parameters.



We added a comment: "However, a simple sensitivity test, carried out to assess the impact of a decrement in the Aa value on the single link scores, did not show any substantial improvement, especially when its results are extended to the whole dataset. More information are provided in the supplementary material. "

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 agreed on this and run the calculations with the filter on the Reference only and without any filter too. As expected from the fairly good categorical scores (which evaluates all the four values of the confusion matrix), false positives and negatives do not affect the overall picture too much. We added specific columns to Table 2 to allow comparisons with other studies.

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. number_of_estimated_wet)/(number_of_observed_wet). Hence a Multiplicative Bias of 0.47 indicates that only half of the wet samples are found. Then, we have to remind that categorical scores in our work are always calculated over the unfiltered dataset, around the threshold, which would later be used in the filter. This was not sufficiently clear, and we already presented the changes in answer to the first general comment and the one on L298.

Addressing now this case specifically, 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), are quite high. This indicates that in this area, the rainrate is estimated with higher accuracy, while the discrimination wet/dry is worse.

We modified the paragraph to: "The higher accuracy in the estimates is reached at the expense of POD, ETS and MB: around 50% of the rainfall duration is lost in this area. The main peculiarity of the RRB area is the high LC, which is 50% higher than the rest of the regions. We can infer that the higher coverage led to a more selective NLA classification, which reduced FAR and POD. The marked improvement of continuous indicators suggests that the quantitative matching between estimated and reference could be positively related to LC."

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 74km ^{2} grid. Our results for CV and CC estimated at 25 km^2 are in good agreement with them. In addition to their work, we also show 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 added a sentence in the revised manuscript: "These results integrate the findings of Overeem et al. (2016), that highlighted the positive impact of higher LC on CV and CC at lower spatial resolution".

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 fuzzy logic techniques which exploit Z, 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 Referee pointed out, that a reason for high false alarms ratio could be that ERG5 misses some small scale events. We modified 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 emphasize that to fully exploit radar capabilities a customized Z-R relationship should be used for each type of precipitation. At the same time, the k-R relation of the CML retrieval is almost independent on the 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 same coefficients for the retrieval can be effectively applied regardless the type of precipitation".

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 high relevance also their latency, i.e. the time taken from the acquisition of the primary data (the occurrence of the event) and the delivery of the product in a ready-to-use form.."

If the error structure of the products is well assessed, and the operator is well aware of it, he can decide to use a product that usually underestimates (low POD, as CML in our case), or a product that tends to overestimate (Radar, in our examples). At the end of the paragraph, we added the comment: "It is to the operators' preference, based on product error structure, current meteorological conditions, and user's requirements, to make use of the most suitable product".

Response to Anonymous Referee #2

Main weakness of the manuscript is related to CML data processing and data analysis which is based on open-source package RAINLINK applied on CML data in northern Italy. I am missing the definition and answering the important research questions which can provide new insights in CML rainfall retrieval. The overall scientific significance of the manuscript is fair.

Therefore, the manuscript needs major revisions. I see several aspects that can be studied using such data set. The quality of CML product is questionable and it show systematic underestimation. Then one way could be to test/develop other processing methods of CML data to reduce this bias and improve the quality of the product. Other interesting point could be an orographic aspect which is mentioned in the manuscript, but not studied in detail.

We agree with the Referee that in this paper we do not address fundamental research questions, such as to set-up advanced algorithms or tackle challenging issues. Still, we think that one important task in the research activity is to communicate to potential users possible applications of the research itself.

Moreover, in our opinion, the data available to us was simply not accurate and complete enough to develop and test new algorithms or to analyse the impact of orography or other known critical aspects of the rain retrieval from CMLs. Longer data time series over wider regions and a more reliable and representative reference dataset is needed to do such studies, which was not accessible to us at the time. However, we believe that this work demonstrates a good potential of the technology even at its most basic implementation, and gives valuable hints for future regional improvements.

The objective of our study, indeed, was to test the possible role of CML retrievals in an operational environment without any previous research on the characteristics of the available CMLs. We performed an "out of the box" approach as we aimed to test the performances obtainable without specific calibration (whose related effort could be not sustainable in many places). We assessed that a robust and freely available algorithm (such as RAINLINK) provides a product with spatial and temporal characteristics comparable to products routinely available to the operators in our region. Moreover, we highlighted how the performances of RAINLINK could be improved, addressing the few parameters that could benefit from a calibration/validation campaign (with proper instruments), once it will become possible.

To clarify, we'll modify the sentences in the revised Introduction as follows:

".The first objective of the present work is to make a validation of precipitation amounts and distributions estimated only from CML attenuation data, using a well-established, freely-available algorithm (i.e. RAINLINK), over two areas of interest in the Po Valley (provinces of Bologna and Parma), where CML data have been obtained from Vodafone (direct purchase). Both areas contain river basins of considerable local interest, which will be explicitly addressed. Moreover, we consider for intercomparison only precipitation products routinely available at Meteorological Service of the Regional Agency for Environmental Protection and Energy (Arpae-SIMC): this, from one side, prevents us from performing a proper calibration of the algorithm, but, on the other hand, allows us to understand how CML product behaves with respect to other operational prod- ucts. The further aim of the validation study is thus to test the potential of the technology even at its most basic implementation, indicating where to direct the tuning efforts, to set the background for possible inclusion of CML data in the operational routine procedures for precipitation monitoring."

General comments:

1. The results show systematic underestimation of QPE derived from CMLs. RAINLINK package contains several strong assumptions (constant WAA of 2.3 dB, constant k-R parameters etc.) which can influence the results significantly. Recent knowledge shows that WAA is complex process with many unknowns (e.g. Leth et al., 2018). The dataset probably contains a certain portion of sensors with low sensitivity (this is reviewer assumption since the CML statistic is not provided) to rainfall where WAA can play dominant role in resulting rainfall retrieval. I would recommend to make at least sensitivity analysis of the results to most significant parameters.

It is well known (van Leth et al., 2018 among many others) that antenna wetting is one of the main problems in microwave estimation of precipitation, and for this reason, we think we cannot address this issue with our operationally oriented verification system. We remark that to address this issue properly, van Leth et al. (2018) deployed a unique experimental setting, with extremely controlled antenna conditions (time-lapse camera pointing the antennas) and accurate reference measurements (five disdrometers along link path). Even with this unique experimental setting, neither van Leth et al. (2018) could definitively address this issue in a general way.

Anyway, to give some more hints to the reader interested in the use of RAINLINK, we performed a sensitivity study to test the impact of the Aa (Antenna attenuation) parameter on the results, though bearing in mind that Overeem et al. (2016) itself reports that " applying Aa should be seen as a pragmatic approach towards correcting for wet antennas". The study is performed on the sample dataset of the 27 CML against the nearby raingauges, exploring the sensitivity to Aa and alpha parameters paired and using CC and CV as loss functions. The results are summarized in response to the Anonymous Referee #1 about L296 and added to the Supplementary Material.

For the discussion about the dataset characteristics and the low sensitivity links, please refer to the reply to the specific comment on L90-94 instead.

2. Spatial interpolation is based on assumption the path-integrated rainfall is represented as a point measurement. This assumption can be used for rough grid 5x5 km and shorter CMLs. However, it is weak for single link comparison (section 4.1) including single event comparison. Here, spatial-temporal structure of rain together with the layout of given RG and CMLs can play significant role. Then it is impossible to compare single point measurements and CMLs observations.

We thank the Referee for pointing out this issue, and we agree that the comparison between single link and single rain gauge is affected by many uncertainties. However, similar shortcomings should also apply when comparisons with other instruments are considered. The only way to proceed correctly seems to be to follow Van Leth et al. (2020)'s approach or similar (i.e. many instruments along the link path), which unfortunately is impossible to replicate in an operational scenario.

Still, we believe that our single-link vs raingauge analysis gives the reader some valuable hints to understand the behaviour of the interpolated RAINLINK product when presented.

To corroborate the robustness of the analysis, we added in the new version of Figures 2 and 3 the information about the shortest distance between the CML and the respective raingauge. These distances are distributed as discussed in answer to the comment on Fig. 2 by the Anonymous Referee #1, who raised a similar question, and it can be seen that a sufficient level of consistency is always guaranteed.

We, therefore, added the following sentence in the new manuscript: "We have selected links in rural areas and different terrain with an active rain gauge close to the link: the distance between link and raingauge, reported in Figure 2, is always below 3 km (significantly lower than the correlation distance of precipitation in Italy (Puca et al., 2014)). In general, no dependence of the link performance on the distance from the raingauge is found."

Basing on the various studies from the literature where space integrated rainfall is compared with raingauges measurements, we are not convinced by the last sentence of the comment.

3. Since rainfall maps are the key product of the presented study, I would expect to show visually CML rainfall maps – event-based or cumulative rainfall compared to reference.

According to also the hints of the Anonymous Referee #1, we added three case studies in the revised manuscript, with maps and skill indicators. We will show also the cumulated map for the whole period in the supplementary material.

4. I am missing relevant discussion section in the paper

We would prefer not to change the structure of the paper, but we added deeper discussion in many parts of the manuscript. Moreover, we added a new chapter, supplementary material, and completely rewrote the conclusions, where more discussion is reported.

5. I am not satisfied with the conclusions which do not provide novel information beyond the state of the art in the field of CML rainfall retrieval.

We better specified in the Conclusions that the focus of the paper remains on the operational implementation of the CML-based rain retrievals: the novelty, therefore, consists in evaluating the RAINLINK performance in a situation different from the original one, exploiting different metrics, techniques, and operational products as benchmarks, gauging in the process also the implementation effort and indicating the principal strength and weaknesses while suggesting ways to benefit from CML products the most. We completely rewrote the Conclusions also to include comments on the newly presented event-scale analysis. Now the conclusions reads as:

"An assessment of the 445 rainfall retrieval capability of CML opportunistic sensors over heterogeneous terrain in northern Italy is conducted at different spatial and temporal scales for two months of data. We implemented the open source RAINLINK algorithm in a new area and context, where no regional CML studies had previously been performed. We evaluated its performance through a complete validation scheme which involves operational precipitation products as benchmark, gauging in the process also the implementation effort and identifying major strengths and weaknesses to make a profitable use of CML products.

First, 26 CMLs (out of the total 308) are compared with the closest raingauges at a 15 min scale. Overestimation and underestimation of rain amount are both present, though the latter appears dominant. A marked variability among different links does not prevent to achieve a generally acceptable skill (CC from 0.50 to 0.88). The wet-dry classification approach and the value of the wet antenna correction may generate loss of rain amount in case of small scale and/or intermittent episodes.

Finally, higher elevation CMLs show in general worse performances. Interpolated products obtained from the full sample of 308 links confirm that a non-negligible quantity of rain is missed (normalized Mean Error is -0.26, overall CC is 0.68 and overall CV is 0.78), but also show that the rain retrieval capability is suitable for operational application, especially if the product is integrated over large areas (CC rises to 0.92). Higher link densities increase the quality of the CML estimates at both gridbox and basin scales, mostly in terms of decreased FAR.

Performances at event scale show enhanced skill in case of heavy precipitation, even in case of small scale rain episodes, while problems arise when light/moderate rainrates challenge the algorithm in the ways we already identified in the single-link analysis. Negative impact on the overall results comes from areas with poor sensor coverage, especially near the border of the areas, but it should be considered that also reference rainfall fields can be affected by shortcomings of the same nature.

Furthermore, when compared to other products currently available for operational real-time exploitation, CML sensors show similar or better abilities than their counterparts, especially if latency is also taken into account. Hence an integration of microwave links sensors in an operational service is highly desirable, even without a proper calibration of the algorithm to the local climatology and CML network characteristics.

When a more complete dataset would become available the validation scheme implemented for this work could be promptly used to tune the RAINLINK parameters (NLA radius, Aa,) on a training sample specific of the study area."

Specific comments:

L. 33-45. I don0t agree with this paragraph since the first sentence refer to CMLs. The provide references are partly based on experimental microwave link setup, not CMLs. I wonder we know accurate algorithms for DSD, water content etc. based on CML observation.

The Referee is right: the sentence was poorly structured. We rewrote as: "Accurate experiments with high-quality links 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 field depends on the density and distribution of the links, making the CML approach of particular interest for urban areas… "

L. 70. observation period – since later in the manuscript some analysis are event based I would add into the Supplementary material information and data about precipitation events during observation period. For selected rainfalls and locations used later in section 4.1 some detailed rainfall metrics would be welcome.

We added Supplementary material where a section with a description of rainfall characteristics is presented. Moreover, in the new section 4.2.2, three case studies are shown, also discussing specific precipitation features.

L. 90-94 The usage of CMLs with low operating frequencies 6 – 15 GHz is questionable for QPE because of low sensitivity of those devices to rainfall even with longer path lengths. It would be useful to provide statistic evidence of different frequency bands in the data set including calculated theoretical sensitivity to rainfall. Then the effect of constant WAA to the results would be much clearer.

We thank the Referee for having pointed out this topic. In our network, we have only five links between 5 and 10 GHz, and a few more around 12 GHz. We investigated the sensitivity for all links, as presented in the answer to comment on L199 by the Anonymous Referee #1. Following that analysis, we decided to remove from the dataset the 15 links with a sensitivity below 0.1 dB per mm h^{-1} and to add a sentence to the revised manuscript: "The CMLs' operational frequency in our region spans between 5.0 and 45.0 GHz. We decided to extend the default frequency allowance window from 12.5 - 40.5 GHz (as was in the Netherlands) to 10.0 - 45.0 GHz, leaving out five low frequency CMLs, but also to remove from the dataset 10 other links with higher frequencies but with sensitivities below 0.1 dB per mm h^{-1}. This is done to avoid contamination by coarse low sensitivity signals."

L. 104 Spatial distribution of LC – could you explain why the LC is lower in the main regional cities (Parma and Bologna) than in countryside – Figure 1?

In Italy and generally in the world, most of the CMLs in urban areas are being substituted by underground optical fibre cables. See also the reply to the comments for L228.

Section 2.1.2 Transmitting power levels I found this paragraph a little bit confusing. I would ask to rephrase it to provide clear information about ATCP processing.

We rephrased the paragraph as follows, describing the manual correction of the ATPC in detail and hopefully improving clarity and exhaustiveness.

"CMLs are usually equipped with Automatic Transmit Power Control devices (ATPC) which modulate the transmit power to guarantee a constant power level at the receiving end of the link, cancelling minor

fluctuations of the total attenuation along the path. ATPC works at a higher frequency than 15 min^{-1} and in a power window spanning from 0 to +6 dB. With ATPC active, attenuation measurements should, therefore, be performed subtracting receiving to transmitting powers and are not possible from receiving powers only. The CMLs analysed in this work are equipped with ATPC, but we do not have access to the transmitting powers, due to confidentiality restrictions. Luckily, provider engineers gave us instead some statistics of the functioning of the ATPC devices (specifically, the modulation maxima (in dB) in the time interval), through which we are able to correct the receiving power levels, compensating for the power modulation effects, simulating CML data with constant transmitting powers and thus allowing RAINLINK to estimate attenuations from receiving powers only. The correction intervenes on minimum received powers (Pmin), which are with no doubt affected by the ATPC: they are manually lowered by the maximum ATPC modulation applied within the respective 15 min time window. Maximum receiving powers (Pmax) instead are left untouched as the ATPC working frequency and the 15 min^{-1} sampling frequency does not coincide and there was no way to infer a reasonable compensation. This could result in a broader gap between Pmin and Pmax.. This could result in the broader gap between Pmin and Pmax."

Implications of the ATPC corrections on the Pmin-Pmax gap are discussed in the answer to the comment about L269 by Anonymous Referee #1.

L. 193 Interpolation – please explain how path-averaged rainfall depth from each CML is implemented into spatial interpolation. This not very clear from provided description Section 4.1 Single link verification – see my general comment about point and pathaveraged rainfall estimates. This is difficult to understand especially when we don't see detailed information about precipitation metrics during observation period. The data also does not correspond with previous statement in Section 2, that in higher altitudes are higher amount of rains.

We modified the description of interpolation procedure that now reads: "CML path averaged precipitation estimates are assigned to the mid points of the links like point measurements ("virtual raingauges"). Interpolation of the point-like measurements is performed at hourly scale with ordinary kriging on a spherical semivariogram on the ERG5 grid. Sill and Range parameters are estimated from the available raingauge stations of three consecutive years. The interpolated field is truncated if it gets smaller than 0.05 mm, which is half of the minimum detectable rain from a raingauge."

More details on the precipitation characteristics involved in the analyses are now provided in the Supplementary Material.

The sentence in Section 2 is related to the rainfall climatology of the region and states only that the highest rainfall amounts are located on the hills. No direct correlation or proportionality between rain amount and height is suggested. From the additional information about the total cumulated rain depth over the two months, now provided in the supplementary material, it is possible to verify that the precipitation maxima are indeed on the meridional border of the region, which coincides with the Apennines ridge (see Fig 1 of the manuscript).

To improve clarity, the sentence in Section 2 now reads:

"..., with the maxima of the rainfall amounts located on the Apennines ridge (see Supplement)."

L. 228 I suggest this statement as weak and confusing "They have been chosen in areas with different terrain and network density and far from the cities, as CMLs in urban areas are already well studied and

also the most eligible to be replaced by optic fibres.“ I don0t see why CMLs in cities should work in different way than in country side. Is there evidence that CML in cities are already well studied and in the countryside not? Network development is not relevant for this paper and this sentence is speculation.

CML network characteristics are different going from cities to the countryside. Specifically, in the cities, there are fewer CMLs since most of them have been already replaced with optical fiber (see also the answer to the comment on L104 and the recent The Netherlands’ situation reported for example in Overeem et al. (2016), Introduction section, 4^{th} paragraph). We provide many references for metropolitan CML studies with short links. Moreover, implications of the network’s developments on operational retrieval capabilities are among the most relevant topics in the CML field.

We removed the sentence and modified the previous sentence to: “We have selected links in rural areas and different terrains, that have an active rain gauge close to the link...”

Sections 4.1.1.-4.1.2. - Best and Worst Case Example – I do not understand why there is no text information and results interpretation with respect to rainfall intensity and rainfall characteristics. 4.1.2 represents light rain when the sensitivity to rainfall of CMLs is low. WAA is significant here anyway. Also, data provided from NMS system in form Pmin Pmax are limiting factor. This shows clear limits of CML for light rainfalls and Pmin Pmax approach.

We thank the Referee for this comment. As also suggested in the comment on L70, we added some Supplementary Material to the revised manuscript in which we describe the precipitation characteristics. We added a more detailed discussion on these results, and we included three new event-scale validations, addressing primarily the type of precipitation and the Pmax-Pmin approach, and also considering previous Referee’s comments.

However, regarding the case treated in the Section 4.1.2 specifically, we would first like to point out that the information about the precipitation characteristic for the duration of the event is provided in Fig.3 in the form of 15min raingauge measurements (see Fig.3, solid circles and black cumulated profiles). Secondly, the links over Vergato have an average theoretical sensitivity of around 0.3 dB/mm/h, which is roughly the same as the links over Sant’Agata, so the theoretical sensitivity alone could not explain the different performances and behaviours between the two case studies. Nor could the wet antenna attenuation parameter Aa, as it is clear from the figure that the issues there are mainly related to the classification, which misses most of the wet intervals.

Lastly, we do not see how the Pmin Pmax sampling should be a limiting factor here, as there are no evident anomalies within the presented signals. The fact that Pmax rarely moves from the baseline value, as discussed in the response to comment on L269 by Anonymous Referee #1, has no direct implications on the the classification, as NLA does not consider Pmax in its calculations. As stated in the original manuscript, Pmax near the baseline indicates only that the power fluctuations are faster than 15 min^{-1}, which suggests irregular and scattered precipitation patterns. It is this irregular precipitation, and not the constancy of Pmax itself, that could be a factor that affects the correct classification, since the NLA algorithm is based on the spatial correlation of the rain field.

We modified the sentences to be more clear on the topic:

“Looking at Figure 3d, Pmin does show a decrease coupled to the missed rainfalls, but Pmax does not. This behaviour of Pmax is not an issue itself, as the NLA classification relies on Pmin only, but it indicates that there are power fluctuations which happen faster than 15 min-1. Rapid fluctuations, in turn, suggest

irregular and scattered precipitation patterns, that actually could be a factor that affects the correct classification, since NLA relies on the spatial correlation of the rain field. Therefore, a Pmax signal always near the baseline could be a precursor of local NLA issues."

Section 4.1.3 – I do not think that this melting layer story fits to this story. First, the data set is presented as spring – summer period. The article is focused on liquid precipitation, this is another story.

The melting layer episode does not belong to the 2-month dataset used for the main study, but it was a standalone dataset obtained from Vodafone for preliminary checking. This event occurred in early March when the freezing level could reach the ground, especially on the hills. Since liquid precipitation at midlatitude originates from frozen hydrometeors, the bright band is a rather common feature in our regions and introduces errors in the radar estimates often difficult to correct. Anyway, we understand that this issue is a bit far from the mainline of the work, so we decided to delete this subsection.

L. 320-330 I do not fully agree with those statements about LC. Different LC often means different frequency bands distribution. In the region with high LC one can expect higher frequency bands with higher rainfall sensitivity.

This would be true if LC measured the density of the antennas, but since it is an estimate of the cumulated link path lengths crossing a square of 5x5 km, LC also benefits of long links overpasses, which work at lower frequencies. In fact, we did not find any correlation between frequency and LC to date, but we thank the Referee for the hint.

Figures general – I found inconsistency when using brackets for units – none, () or [] in different figures

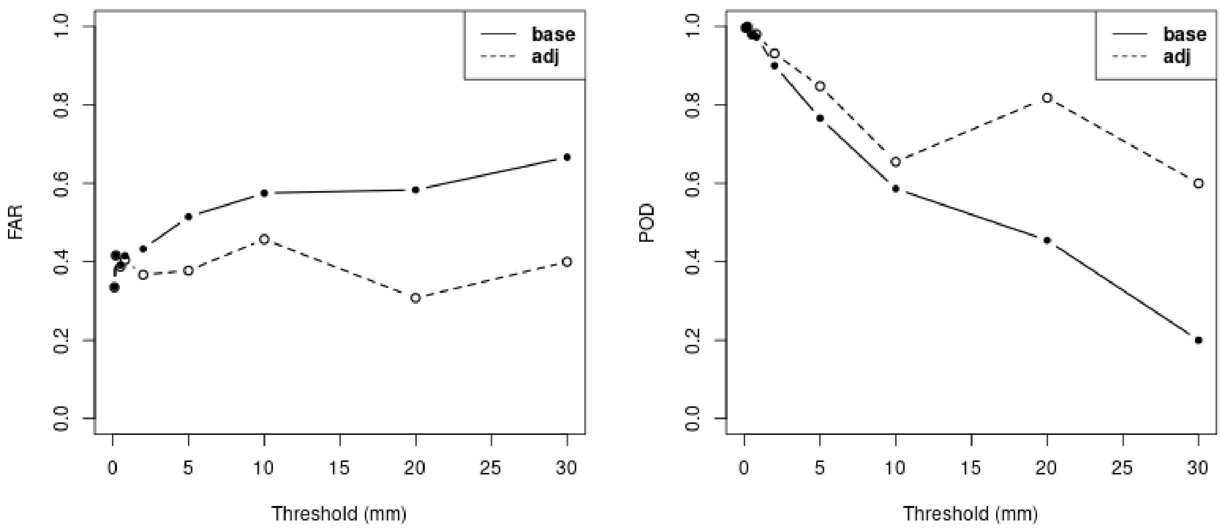
Thank you for noticing this, we have fixed this inconsistency.

Figure 7. I do not understand the "bad" results of adjusted radar in comparison to the reference which was used for radar adjustment. The results are comparable to unadjusted radar data. Could you explain that?

The adjustment is performed with gauges and not with the interpolated reference, as specified in Section 2.2.3. The procedure matches the rainrates estimated over the gauge locations but does not ensure the consistency of the whole radar field with the gauge interpolated one, mostly because of the high spatial variance of the radar field (as already discussed in Section 4.2.2, L356).

Therefore, discrepancies in the areal averages are not only to be tolerated but also expected. Moreover, the spatial autocorrelation of the G/R adjustment factor is even lower during convective events, leading to a less effective correction. The following figure shows an independent analysis of the radar adjustment performances in the same two months of 2016 and it is visible that around the low thresholds used in our paper (0.1 mm) practically no improvement is expected to come from the adjustment procedure.

We added a sentence in the revised manuscript: "Radar product shows, in this metric, almost the same performance both with and without the gauge adjustment. This can be expected since the radar adjustment happens above the raingauge locations but does not ensure the consistency of the areal average of the *whole rain field. The adjustment also affects mainly higher rainrates than our 0.1 mm threshold and has lower performances as the spatial variance increases, e.g. in cases of small scale convection..”*



Marked-up version of the revised manuscript:

Commercial Microwave Links as a tool for operational rainfall monitoring in Northern Italy

Giacomo Roversi[1], Pier Paolo Alberoni[2], Anna Fornasiero[2], and Federico Porcù[1]

[1]Department of Physics and Astronomy, University of Bologna, Bologna, 40100, Italy
[2]Arpae-SIMC, Bologna, 40100, Italy

Correspondence: Federico Porcù (federico.porcu@unibo.it)

Abstract. There is a growing interest in emerging opportunistic sensors for precipitation, motivated by the need to improve its quantitative estimates at the ground. In this work, a preliminary assessment of the accuracy of Commercial Microwave Links (CMLs) retrieved rainfall rates in northern Italy is presented. The CML product, obtained by the publicly available RAINLINK software package, is evaluated at different scales (single link, 5 km×5km grid, river basin) against the precipitation products operationally used at Arpae-SIMC, the Regional Weather Service of Emilia-Romagna, in northern Italy. The results of the 15 min single-link validation with close-by raingauges show high variability, with the influence of the area physiography and precipitation patterns and the impact of some known issues ~~(e. g. melting layer).~~. However, hourly cumulated spatially interpolated CML rainfall maps, validated with respect to the established regional gauge-based reference, show performances (R^2 of ~~0.47~~ 0.46 and CV of ~~0.77~~ 0.78) which are very similar, when not even better, to satellite- and adjusted radar-based precipitation gridded products. This is especially true when basin-scale total precipitation amounts are considered (R^2 of ~~0.85~~ 0.83 and CV of ~~0.63~~ 0.48). A diffuse underestimation is evident at both grid box (Mean Error of -0.26) and ~~basin scale~~ basin-scale (Multiplicative Bias of 0.7), while the number of false alarms is generally low and gets even lower as coverage increases. Taking into account also delays in the availability of the data (latency of 0.33 hours for CML against 1 hour for the adjusted radar and 24h for the ~~quality controlled~~quality-controlled raingauges), ~~CMLs appear~~ CML appears as a valuable data source in particular from a local operational framework perspective. Finally, results show complementary strengths for CMLs and radars, encouraging ~~a~~ joint exploitation.

1 Introduction

Precipitation is one of the most difficult geophysical observables to measure and monitor, given its very high temporal and spatial variability. Its accurate measurement would benefit a wide range of applications in meteorology, hydrology, climatology, agriculture, just to mention the most direct fields where precipitation plays a key role. The precipitation rate can be measured or estimated directly at the ground or ~~by means of~~using different remote sensing approaches. ~~Rain gauge~~Raingauge networks provide point-like measurements of the amount of rain fallen within the instrument's sampling area, cumulated over time intervals which usually range from one minute to one day, with well known instrumental constraints (Lanza and Stagi, 2012) and representativeness limitations (Porcù et al., 2014). ~~Ground based~~ Ground-based weather radars, often deployed in large scale networks (Serafin and Wilson, 2000; Huuskonen et al., 2014; Saltikoff et al., 2019), are widely used by hydro-meteorological services to quantitatively monitor precipitation fields, being an effective trade off between spatial-temporal coverage and accuracy in the measurements. However, radar estimates are affected by ~~a number of~~ several errors, which the last generation polarimetric systems have only partially mitigated ~~(Ryzhkov and Zrnic, 2019)~~(i.Ventura et al., 2012; Gou et al., 2019). Satellite estimates received a renewed boost in the last decade from the full exploitation of the Global Precipitation Measurement mission (GPM, Skofronick-Jackson et al. (2017)) that operationally releases a new suite of precipitation products with a high temporal and spatial resolution (Mugnai et al., 2013; Grecu et al., 2016). Despite the undoubted potentials of satellite products to provide estimates over open oceans and regions not equipped with ground instruments, their accuracy ~~is still under evaluation (Tang et al., 2020) and~~ difficult to assess at high spatial and temporal scales (Tang et al., 2020), and their latency hinders the use in ~~real time~~ real-time monitoring of rain patterns.

A relatively new and independent approach to the estimates of precipitation at the ground became available in the last decades with the ~~broad diffusion~~ ubiquity of microwave links for cellular communication (or Commercial Microwave Links, CMLs): integral precipitation content along a straight path between two antennas can be estimated by measuring the attenuation of the microwave signal travelling down the same path (Turner and Turner, 1970; Harden et al., 1978). Accurate ~~algorithms were introduced to~~ experiments with high-quality links 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., 2019) and water content (Jameson, 1993)~~along the link. The~~ . On the same token, the possibility to have a spatially continuous rainfall ~~path~~ field depends on the density and distribution of the links, making ~~this~~the CML approach of particular interest for urban areas (Upton et al., 2005; Overeem et al., 2011; Fenicia et al., 2012; Fencl et al., 2013; Rios Gaona et al., 2017; de Vos et al., 2018) with also direct hydrological use in combination with conventional instruments (Grum et al., 2005; Fencl et al., 2013). A further application of CML approach could be in regions where other instruments are lacking or absent at all (Mulangu and Afullo, 2009; Abdulrahman et al., 2011; Doumounia et al., 2014). However, as it happens for conventional precipitation instruments, the quality of the retrieval is sensitive to ~~a number of factors~~several factors, often difficult to control (Leijnse et al., 2008), and to the precipitation ~~micro-physical~~microphysical structure (Berne and Uijlenhoet, 2007; Leijnse et al., 2010). Given these limitations intrinsic to the measurement geometry and to the nature of precipitation, possible synergistic approaches are considered, to minimize the uncertainties of the different instruments, suggesting the blending of CML measurements with conventional precipitation estimates, such as ~~rain gauges~~raingauges (Fencl et al., 2017; Haese et al., 2017), radar (Cummings et al., 2009; de Vos et al., 2019), or both (Grum et al., 2005; Bianchi et al., 2013).

~~If~~ Even if the general relationship between signal attenuation and rain rate is already well established, ~~with an obvious dependence on precipitation type and general climatological settings (e.g. air humidity), the success of the~~ the successful use of CML data to ~~effectively monitor precipitation~~ quantitatively monitor precipitation still depends on the quality and ~~general~~

technical characteristics of the transmitted power data and ~~on the fine tuning~~ the fine-tuning of the algorithms~~to the different physiography, climatic regimes and network topologies. The rather~~ . The somewhat standardized policies of acquisition and storage of the different companies in different countries make the use of CMLs feasible all around the world, but there is no standard way yet to access them as scientific data. As they consist mostly of confidential maintenance data, major obstacles to face are the widespread unwillingness of releasing them cost-free and the ~~bad~~ inadequate data-quality standards (Chwala and Kunstmann, 2019).

The first objective of the present work is to make a validation of precipitation amounts and distributions estimated only from CML attenuation data, ~~by means of a well established~~ using a well-established, freely-available algorithm (i.e. RAINLINK, Overeem et al. (2016a)), over two areas of interest in the Po Valley (provinces of Bologna and Parma), where CML data have been obtained from Vodafone (direct purchase). Both areas contain river basins of considerable local interest, which will be ~~addressed specifically.~~ explicitly addressed. Moreover, we consider for intercomparison only precipitation products routinely available at Meteorological Service of the Regional Agency for Environmental Protection and Energy (Arpae-SIMC): this, from one side, prevents us from performing a proper calibration of the algorithm, but, on the other hand, allows us to understand how CML product behaves with respect to other operational products. The further aim of the validation study is ~~to~~ thus to test the potential of the technology even at its most basic implementation, indicating where to direct the tuning efforts, to set the background for ~~a~~ possible inclusion of CML data in the operational routine procedures for precipitation monitoring ~~in the Meteorological Service of the Regional Agency for Environmental Protection and Energy (Arpae-SIMC)~~.

In Section 2 we will describe the area of interest and the different rainfall datasets (CML, radar and rain gauges), including data quality and coverage. In Section 3 we will ~~briefly~~ describe the RAINLINK algorithm and the minor modifications performed to adapt it to the Emilia-Romagna area. The comparison – at single link and gridded map scales – between the rainfall estimates from the different data sources is presented and discussed in Section 4, while conclusions are provided in Section 5.

2 Data

We have considered ~~a period of~~ 57 days from 5 May to 30 June 2016. The two target areas for which we have available CML data are the provinces of Bologna (BO, 3702 $\mathrm{km^2}$) and Parma (PR, 3447 $\mathrm{km^2}$), both in the Po Valley in Emilia-Romagna, northern Italy (coloured areas in Figure 1). The physiography of the two regions is similar: the highest peaks (about 1,500 m a.s.l.) are located on the southern border, in the Apennine chain, while the central and northern part of the two areas are flat land. The two river basins (thick lines in Figure 1) are both located in the hilly ~~part~~ region and have their closing sections located near the cities, in densely populated and assets-rich areas.

Precipitation climatology in the Po Valley during the late spring season is characterized by both stratified structures and small scale convection, with ~~highest~~ the maxima of the rainfall amounts located ~~in the hilly part~~ on the Apennines ridge (see Supplement). We divided the whole area into square boxes of $5\mathrm{km}\times5\mathrm{km}$ (see also Section 2.2.2) and this grid will be used to carry out rainfall interpolation and products intercomparison.

The validation has been carried out comparing, at different spatial and temporal scales, the rain amount obtained by CMLs, through the RAINLINK algorithm (Overeem et al., 2016a) with other rainfall estimates operationally available over the target domain. In particular, CML product has been compared with radar surface rain rates, both raw and gauge-corrected, rain gauges measurements and the operational precipitation analysis (ERG5) made available by Arpae-SIMC.

2.1 CMLs

Microwave attenuation data and metadata were purchased as a single dataset of two months, from Vodafone Italia S.p.A. within the Life EU project called RainBO LIFE15 CCA/IT/000035 (Alberoni et al., 2018). Received powers are measured by the provider with the resolution of 1 dB at a frequency of ten times per second for maintenance purposes, but only maxima (P_{max}) and minima (P_{min}) readings in a 15 minutes time window are stored for backup. Therefore, data is in the format of 15 minutes $[P_{min}, P_{max}]$ pairs. All the available 357 CMLs are "duplex" links, so that two sub-links (back and forth) are present for the same link (although not always simultaneously active). Signal polarization is vertical for 259 CMLs, horizontal for the remaining 98, while carrier-signal frequencies span from 6 to 42.6 GHz, with an average frequency ($\bar{f}$) of 22.1 GHz. Sublinks of the same CML ~~share always~~ always share the same polarization and differ only in frequency by a small gap of around one GHz. Path lengths of the links ~~varies~~ vary from 162 m to 30 km, the interquartile range extends between 2.4 and 8 km, and the average length ~~(LL)~~ is 6 km. As expected, the carrier frequency is anti-correlated with path length since high frequencies, while allowing a wider transmission band, are more subject to attenuation compared to the lower frequencies (Leijnse et al., 2008).

2.1.1 Coverage and data quality

The number of working CMLs varies slightly across the months: it grows from 348 at the beginning of May to the maximum of 357 in June. The number of valid CMLs for rain retrieval is lower because of the quality and sensitivity filtering performed by the pre-processor of the algorithm (see Section 3.1), resulting in ~~an average number of 319 valid CMLs almost constant in time~~a median number of 308 valid CMLs with very small fluctuations. Most of the rejected data is empty or incomplete (P_{min} or P_{max} missing), probably due to failures in reading or storing raw data. More details on the rejected data are presented in the Supplement.

Four parameters are utilized to summarize the topological structure of the CML network: the link density LD (defined as the total number of link paths divided by the ~~total~~ whole area, in km^{-2}), the average link length LL (in km), the bulk link coverage BC (defined as the sum of the link path lengths divided by the total area, in km km^{-2}) and the local link coverage LC (calculated as BC but for each gridbox, in km km^{-2}). Due to Vodafone confidentiality restrictions, we are not allowed to show the exact location of the available links, so we show ~~instead~~, instead, in Figure 1 the spatial distribution of LC.

Since the RAINLINK original settings depend on the network characteristics, we compared the Emilia-Romagna network (ER) with the one from The Netherlands (NL), which is included in the RAINLINK software package as test sample (Overeem et al., 2016a), and with other datasets on which the algorithm was employed (Overeem et al., 2013, 2016b). The datasets properties are summarized and compared in Table 1. ER has comparable link density and higher average link length, resulting in a higher bulk coverage with respect to the NL network. The province of Bologna hosts more than half of the links (~~211 against 146~~195 against 113) and thus has a higher LD .

2.1.2 Transmitting power levels

~~All the links are equipped with an Adaptive Transmission Power Control device~~ CMLs are usually equipped with Automatic Transmit Power Control devices (ATPC) which ~~adapts the transmitting powers to the environmental conditions, in order to guarantee a minimum receiving~~ modulate the transmit power to guarantee a constant power level at the ~~other~~ receiving end of the link~~. This means that~~ , cancelling minor fluctuations of the total attenuation along the path. ATPC works at a higher frequency than 15 min^{-1} and in a power window spanning from 0 to +6 dB. With ATPC active, attenuation measurements should~~not be possible without knowing the transmitting powers too. That information is unfortunately not available to us again for confidentiality issues. To preserve the retrieval capabilities, data were corrected by the mobile operator itself, subtracting to the receiving powers the actual shift applied to~~ , therefore, be performed subtracting receiving to transmitting powers and are not possible from receiving powers only. The CMLs analysed in this work are equipped with ATPC, but we do not have access to the transmitting powers, due to confidentiality restrictions. Luckily, provider engineers gave us instead some statistics of the functioning of the ATPC devices (specifically, the ~~transmitting levels, thus~~ modulation maxima (in dB) in the time interval), through which we are able to correct the receiving power levels, compensating for the power modulation effects, simulating CML data with constant transmitting powers and~~ so matching RAINLINK requirements~~thus allowing RAINLINK to estimate attenuations from receiving powers only. The correction intervenes on minimum received powers (P_{min}), which are with no doubt affected by the ATPC: they are manually lowered by the maximum ATPC modulation applied within the respective 15 min time window. Maximum receiving powers (P_{max}) instead are left untouched as the ATPC working frequency and the 15 min^{-1} sampling frequency does not coincide and there was no way to infer a reasonable compensation. This could result in a broader gap between P_{min} and P_{max}.

2.2 Reference rainrate fields

2.2.1 ~~Rain gauges~~Raingauges

~~Rain gauges hourly~~ Raingauges hourly and 15 min data are provided by Arpae Rirer (regional hydro-meteorological network), established in 2001 by bringing together existing hydrological and meteorological station networks, managed at the time by various public bodies and ~~regional~~ local authorities. The network of the whole Region is composed of 285 stations, equipped with tipping bucket ~~rain gauges~~raingauges: 110 of them are divided ~~in~~ between the Bologna (54) and Parma (56) provinces. ~~Rain gauges~~ Raingauges have different sampling intervals (from 10 to 60 minutes), they undergo a process of homogenization and quality control and are released as an hourly point-like product.

2.2.2 ERG5 rainfall analysis

The ERG5 gridded meteorological data set has been developed by Arpae-SIMC, in order to support agricultural activities in the region of Emilia-Romagna. ERG5 data are operationally produced since 2001, interpolating the hourly station measurements of the main meteorological variables (air temperature, relative humidity, precipitation, wind, solar irradiance) onto a $5\,\mathrm{km} \times 5\,\mathrm{km}$ grid covering the Emilia-Romagna region Region. The interpolation method used for hourly precipitation consists of a Shepard (1968) modified scheme using topographic distances instead of Cartesian distances. This allows the interpolation to take into account the influence of topography on precipitation, by making locations separated by orographic obstacles more distant than they would be if Cartesian distances were used (Antolini et al., 2016). Data are stored and distributed freely in the form of GRIB2 files, which were imported in an R environment thanks to the rNOMADS package (Bowman and Lees, 2015). Among all the variables included in ERG5, we consider here only the hourly accumulated precipitation. Its input is based on the same Rirer network described in the previous Section, no longer limited to the two areas of study, but extended to the whole Region. Some discrepancies are therefore expected between the two products, mainly near the borders and in areas where the distribution of the instruments is less uniform.

2.2.3 Radars

Radar data set is based on hourly precipitation estimates obtained from the composite of the regional radar network managed by Arpae-SIMC. The regional network is composed by of two C-Band systems, located in San Pietro Capofiume and in Gattatico (eastmost and westmost red crosses in Figure 1, respectively). For each instrument the equivalent radar reflectivity factor close to the ground is extracted and interpolated from polar coordinates to a 256×256 Cartesian grid of $1\,\mathrm{km} \times 1\,\mathrm{km}$ resolution, then merged to obtain a composite of both radars.

Raw radar images are affected by various non-meteorological echoes that are removed before computing the Quantitative Precipitation Estimation (QPE). The current scheme used at Arpae-SIMC during operational service includes many steps: the ground clutter is removed at first statically through the map of signal-free elevations recorded in dry conditions, then dynamically by combining a beam trajectory simulation at the current atmospheric state (as measured by radio soundings) and a Digital Elevation Model (Fornasiero et al., 2006). The beam blocking reduction and correction is performed based on a geometric optic approach (Bech et al., 2003), while anomalous propagation is detected after the analysis of the echo coherence in the vertical direction (Alberoni et al., 2001). The final conversion between reflectivity and rainfall rate is performed on the corrected data set using the classic relationship $Z = aR^b$, with $a = 200$ and $b = 1.6$.

Rain rates are obtained every 5 minutes and the final rain total over a one-hour period is computed by an advection algorithm which takes into account the movement of the precipitating systems. The algorithm is based on the computation of maximum cross-correlation between consecutive maps, leading to the estimate of the displacement vector for each precipitating system. The rainfall field is then reconstructed every minute between the observations and cumulated over each hour. Finally, radar QPE is adjusted with rain gauges data, via the spatial analysis of the ratio G/R between rain gauges (G) and radar (R) rainfall rates over the station locations. The spatial analysis is obtained as the weighted mean of the G/R values where the weight is a function of both the distance of the grid point from the station and the mean spacing between observations (Koistinen and Puhakka, 1981; Amorati et al., 2012). In this work we will compare the CML product with both adjusted and unadjusted radar QPEs.

3 Methodology

The process chain which takes CML signals and returns rainfall maps is governed by the RAINLINK algorithm (Overeem et al., 2016a) published open source on GitHub (https://github.com/overeem11/RAINLINK) as an R package. We used the 1.14 version of the RAINLINK algorithm, available online from July 2019, and we added some minor modifications and optimizations (forked version available here: https://github.com/giacom0rovers1/RAINLINK).

3.1 CML rain retrieval algorithm

The algorithm works for both instantaneous power measurements and $[P_{min}, P_{max}]$ pairs: for the present work we use the latter, on 15 minutes intervals. The algorithm treats P_{min} and P_{max} separately (we will then use P_i to refer to both alternatively). Two separate rain estimates R_{min} and R_{max} will thus be obtained. The retrieval process ~~works as follows:~~ is summarized below, while we show more details of the implementation in the Supplementary Material.

1. **Preprocessing**: the raw input goes through three consistency checks concerning data formatting and labelling. Any multiple observations for the same LinkID and DateTime are discarded, each LinkID is verified to maintain the same metadata throughout the whole dataset (Frequency, PathLength and antenna coordinates) and rows with NA values in any of the columns except for Polarization (which is supposed vertical if not indicated) are discarded as well.

2. **Wet-Dry Classification**: the samples are discriminated in wet and dry periods by assuming that rainfall is correlated in space, through the so-called Nearby Links Approach (NLA), which works as follows. For each link, a time interval with a decrease in the received power is labelled as wet if at least half of the links in the vicinity (within 15 km radius) experience a comparable reduction, i.e. if the medians of the attenuation and the specific attenuation of the nearby links are below – 1.4 dB and – 0.7 $\mathrm{dB\,km^{-1}}$ respectively. This is the second most computationally time-consuming step of the algorithm.

3. **Baseline determination**: a 24 h moving-window median of the quantity $\frac{1}{2}(P_{min} + P_{max})$ over the dry time intervals defines a reference level P_{ref} (baseline). This is the computationally time-consuming operation of the algorithm.

4. **Outliers filter and power correction**: outliers due to malfunctioning links can be removed again by assuming that rainfall is correlated in space. The filter discards a time interval of a link for which the cumulative difference between its specific attenuation and that of the surrounding links over the previous 24 h (including the current time interval) becomes lower than the outlier filter threshold, which is fixed at -32.5 $\mathrm{dB\,km^{-1}\,h}$. After removing the outliers, the classification information is used to clean the receiving powers of the noise over the dry periods. The corrected powers P_i^{Cor} will be equal to P_{ref} on dry periods and P_i on wet ones.

5. **Rainrate retrieval**: attenuation A_i is computed as $A_i = P_{ref} - P_i^{Cor}$. A fixed quantity $Aa = 2.3$ dB is subtracted from the attenuation A_i in order to take into account the wet-antenna effect, which is independent on path length and it is assumed independent also on frequency and rain intensity. If $A_i - Aa > 0$ then the specific attenuation k_i (dB km^{-1}) is calculated as $k_i = (A_i - Aa)/L$, otherwise 0 is returned. Path-averaged mean rain intensities R_i (mm h^{-1}) are finally calculated using the $k - R$ relationship $R_i = a\,(k_i)^b$, where the coefficients a and b were from Leinse (2007) and Leijnse et al. (2010) for vertical and horizontal polarization, respectively.

6. **Path-averaged rainfall depth**: to obtain a single path averaged rain depth, R_i are combined through a weighted mean: $R = \frac{1}{4}[\alpha R_{min} + (1-\alpha)R_{max}]$. The factor $\frac{1}{4}$ transforms rain rates in 15 min rain depths. The weight α varies between 0 (estimate derived from P_{max} only) and 1 (estimate derived from P_{min} only); we adopted the default value ($\alpha = 0.33$). We specify that, unlike Overeem et al. (2016a), we chose to keep the subscripts related to the original receiving powers, thus in our notation the rainrate R_{min} is higher than R_{max} because it is obtained from the most attenuated signal P_{min}.

7. **Interpolation**: CML path averaged precipitation estimates are assigned to the mid points of the links like point measurements ("virtual raingauges"). Interpolation of the point-like measurements is performed at hourly scale with ordinary kriging on a spherical semivariogram on the ERG5 grid. Sill and Range parameters are estimated from the available raingauge stations of three consecutive years. The interpolated field is truncated if it gets smaller than 0.05 mm, which is half of the minimum detectable rain from a raingauge.

3.2 RAINLINK implementation in northern Italy

The implementation of RAINLINK required some technical and conceptual considerations. The main differences between Italian and Dutch case studies concern ~~links length, links~~ link length, link frequency and orography. ~~Our CMLs' operational frequencies span over a wider range than The Netherlands' ones, so~~ The CMLs' operational frequency in our region spans between 5.0 and 45.0 GHz. We decided to extend the default frequency allowance window ~~was extended~~ from 12.5 - 40.5 GHz ~~to 5.0~~ (as was in the Netherlands) to 10.0 - 45.0 GHz~~in order not to miss any available information,~~ , leaving out five low-frequency CMLs. We also removed from the dataset 10 other links with higher frequencies but with sensitivities below 0.1 dB per mm^{-1} (see Supplement for more details). This is done to avoid contamination by coarse low sensitivity signals.

The NLA radius has to be consistent with the typical spatial correlation of rainfall and with the density of links available. In our case, the link density (~~0.045) is slightly higher than the~~ 0.043) is the same as the one used for the original setup of the algorithm (~~0.043,~~ see Table 1), and the spatial pattern of precipitation is expected to be similar in Italian and Dutch sites (Caracciolo et al., 2006). So, we let the settings for the wet-dry classification unaltered. Similarly, the k-R relationship is maintained, because northern Italy and the Netherlands share similar climates and overall differences in drop size distributions between the two countries are expected to be negligible (Caracciolo et al., 2006) and certainly lower than variations of the same distribution along the link path and during the 15 minutes time intervals (Tokay et al., 2017). Finally, spherical variogram parameters are calculated for three years of local climatology. Range and sill are 36.12 km and 1.12 mm^2, respectively, which are very well in agreement with what results from the "ClimVarParam" subfunction of Overeem et al. (2016a) calibrated over 30 years of Dutch climate, confirming the previous assumptions. Similarly to them, Nugget is fixed at one tenth of the Sill.

3.3 Error metrics

In the present work, we selected two sets of classical skill indicators, broadly used in the validation community (Puca et al., 2014), the first (Nurmi, 2003): the first one is to assess the capability of the product to delineate detect rainfall occurrence (categorical indicators), and the second one is to evaluate the skill in estimate correctly the quantitative precipitation rate (continuous indicators). The first set includes Probability of Detection (POD), False Alarm Ratio (FAR) , Equitable Threat Score (ETS) is computed after a definition of a confusion matrix by counting the number of samples where both estimate and observation agree on classifying wet (hits, H), or dry (correct negatives, CN) samples, and where there are misses (M, observed wet and estimated dry) or false alarms (F, observed dry and estimated wet). Namely, Probability of Detection is defined as $POD = H/(H+M)$, the False Alarm Ratio as $FAR = F/(H+F)$, the Multiplicative Bias as $MB = (H+F)/(H+M)$ and the Equitable Threat Score as $ETS = (H - H_{rnd})/(H+M+F)$ where H_{rnd} represents the number of hits obtained by chance.

Given e_i and Multiplicative Bias (MB), while belong to the second set Normalized Mean Error(ME), Normalized o_i as estimated and observed values respectively, continuous indicators are the normalized Mean Error, defined as $ME = \sum_i (e_i - o_i)/\bar{o}$, the normalized Mean Absolute Error(MAE),, defined as $MAE = \sum_i \|e_i - o_i\|/\bar{o}$), the Coefficient of Variation (CV) defined as the root mean square error divided by the mean of the observed values $\bar{o}$, and the Pearsons' Correlation Coefficient (CC)(Nurmi, 2003; Overeem et al., 2016b)., as the covariance of observed o_i and estimated values e_i divided by the product of the two standard deviations (Nurmi, 2003; Overeem et al., 2016b).

Due to both interpolation and retrieval methodology, the CML precipitation rate and most of the reference fields we have used here can reach positive values that are sometimes very low and have no interest for any application. Both interpolated CML and reference field have a large number of very low positive values (below 0.1 $\mathrm{mm\ h^{-1}}$) that do not have any physical relevance, but which are potentially very influential in normalized error metrics. Thus we have set a wet-dry threshold equal to the minimum rain quantity detected by the tipping bucket rain gauge, i.e. 0.1 $\mathrm{mm\ h^{-1}}$, for both estimate and reference. Categorical indicators are calculated with respect to this threshold for the whole dataset, while all the continuous indicators are computed only for the product-reference pairs where both values exceed the threshold (i.e. wet-wet). ME, MAE and CV are normalized with the averaged reference rain depth.

4 Comparison between CML and conventional precipitation products

We carried out the validation of CML product at three different levels. First, we compared single link estimates with the measurements of a nearby rain gauge raingauge, at the shortest temporal scale available (15 minutes), to discuss success and failure cases, trying to understand the latter. Secondly, we compared the interpolated 5 $\mathrm{km \times 5 km}$ CML hourly rainfall maps versus the ERG5 product at grid box scale, also analysing three case studies. In the third step, the map comparison is carried out at a basin scale including also even the other precipitation products available at Arpae-SIMC.

4.1 Single link verification

We have selected links in rural areas and different terrains with an active ~~rain gauge at a distance of at least an order of magnitude~~ raingauge close to the link: the distance between link and raingauge, reported in Figure 2, is always below 3 km (significantly lower than the correlation distance of precipitation in Italy (Puca et al., 2014)) and always lower than the length of the link itself. In general, no dependence of the link performance on the distance from the raingauge is found. ~~They have been chosen in areas with different terrain and network density and far from the cities, as CMLs in urban areas are already well studied and also the most eligible to be replaced by optic fibers.~~ Selected links had to be active for all the analysed period. In many cases more than one link was selected for one ~~rain gauge~~raingauge. Temporal sampling is kept at the highest frequency, which is a measurement every 15 minutes for both the CML and the ~~rain gauges~~ raingauges. 12 rain gauges and ~~28~~ 26 CMLs have been ~~selected~~chosen, 14 of which are in the northern part of the domain and the other ~~14~~12 on the hilly region at elevations between 193 m and 960 m a.s.l..

The rain depths of the ~~28~~ 26 CMLs are reported in Figure 2 for the whole study period, grouped accordingly to the closest rain gauge and ranked by its altitude. As a general comment, a large variability is found (ranging from ~~near perfect~~near-perfect agreement to discrepancy of a factor of 2 or 3 in the worst cases). ~~Nonetheless, CC for the 28 links are between 0.67 and 0.86,~~ The 75% of the 26 links CC is between 0.5 and 0.88, with overall median value 0.68, proving an acceptable overall skill. We relate this variability to the heterogeneity of CML sensitivity, the small scale of the meteorological events (see Supplement) and different site exposure and elevation. In most cases, CMLs underestimate the ~~rain gauge~~raingauge values: the links located in the lowlands (Figure 2a, 2b, 2d and 2e) show a better correspondence than those in the hilly regions, where underestimation is ~~larger~~more significant.

In some cases (Figure 2f, 2k and 2l) the discrepancies between CMLs ~~of the same group~~ close to the same raingauge (but different in location, frequency and length~~are low compared to the discrepancies between each of them and the respective rain gauge~~) are much lower than the CML-raingauge differences: all these CMLs are in good mutual agreement and share the same classification issues, resulting in a systematic underestimation which therefore seems to be caused by the algorithm setup. In other cases (Figure 2b, 2d and 2g) some links clearly ~~outperforms~~ outperform others members of the same group. This second kind of discrepancies ~~are~~ is more likely related to real differences, like inhomogeneous rainy structures which crossed the link paths or different hardware setups~~(we do not know antennas specifications),~~ while there is no evidence of a correlation with frequency or path length. ~~In rare cases the disagreement takes place even~~The difference between the two directions of the same link is generally below 10%, except for the Ostia Parmense site (see Figure 2g).

To gain a deeper understanding of better and worse performances of the single links, we performed ~~some~~ a more detailed analysis ~~on~~ of case studies at the rain-event scale (Figure 3). We show a case when the link retrievals ~~properly~~ accurately match the measurements of the close-by raingauge, and a case with markedly low performance~~, and lastly a case that shows the impact of a well know problem in any microwave-based precipitation retrieval: the melting layer~~. In Figure 3, graph panels are organized in columns by CML and in rows by sub-link. In the top ~~panel rain gauge measurements (blue dots)~~ panels are shown all the signals managed by the algorithm: the reference power P_{ref}, the raw received powers P_{min} and P_{max} and the filtered received powers P^{Cor}_{min} and P^{Cor}_{max}. In the middle panel raingauge measurements are compared with CML estimates ~~(purple line)~~ and also the minimum and maximum attenuation signals are plotted A_{max} and A_{min} respectively). The grey background indicates when the classification detects a dry period. The pink background indicates the band inside which attenuation is considered as caused by a wet antenna (Aa parameter) and is discarded for rain retrieval. ~~In the middle panel all the signals managed by the algorithm are shown: the reference power P_{ref}, the raw received powers P_{min} and P_{max} and the filtered received powers P^{Cor}_{min} and P^{Cor}_{max}. The bottom panel shows~~ The bottom panels show the cumulated rainfall ~~depth~~ depths in the same time frame.

4.1.1 Best cases example

Between 11 and 12 May 2016 ~~the~~ an extensive convective system covered the Bologna Province area almost entirely, with a maximum rainrate of 23 $mm\ h^{-1}$, and widespread precipitation around. For this case, the NLA classification on the three links near Sant'Agata (Bologna Province, 18 m a.s.l.) works properly: in Figure 3 ~~a most blue dots (rain gauge measurements) are~~ b most of the measured rain is on white background. In Figure 3~~d~~ a, after the attenuation event, the noisy signal is correctly filtered, and a very small amount of rain (just above the gauge threshold) is neglected. The agreement is qualitatively very high between each pair of sub-links and good among the different links, in terms of specific attenuations and retrieved quantities (see Figure 3b). Quantitative retrievals give some overestimation for one of the CMLs, whose effect is evident on the accumulation plot (Figure 3~~g~~c) where the total rain depths are compared, but cannot be certainly related to any macroscopic characteristics of the three links. During the two months the Sant'Agata links are generally in good agreement with the close-by rain gauge, with CC ranging between ~~0.61 and 0.86~~ 0.66 and 0.88 and CV between ~~0.4 and 0.7~~ 0.47 and 0.96.

4.1.2 Worst cases example

On Vergato (Bologna Province, 193 m a.s.l.) between 8 and 10 June 2016 many ~~events~~ precipitation spots are missed due to wet-dry misclassification (Figure 3~~b~~e) which result in a 20 mm loss in the rain accumulation (Figure 3~~h).~~ f). The event was characterized by intense precipitation peaks (rainrate up to 14.6 $mm\ h^{-1}$) and scattered moderate precipitation, that hit the Vergato site in different times. Looking at Figure 3~~e~~ d, P_{min} ~~(orange) does in fact sense~~ does show a decrease coupled to the missed rainfalls, but P_{max} ~~(blue)~~ does not. This ~~suggests~~ behaviour of P_{max} is not an issue itself, as the NLA classification relies on P_{min} only, but it indicates that there are power fluctuations which ~~develop and completely end during a single time interval of~~ happen faster than 15 ~~min. This high variability could prevent the NLA to see the required spatial correlation to identify wet periods.~~ min^{-1}. Rapid fluctuations, in turn, suggest irregular and scattered precipitation patterns, that actually could be a factor that affects the correct classification, since NLA relies on the spatial correlation of the rain field. Therefore, a Pmax signal always near the baseline could be a precursor of local NLA issues. The POD over the whole period for these two links is between ~~0.34 and 0.36~~ 0.22 and 0.29.

When the NLA classification works correctly, there is still a general quantitative underestimation. It could be seen that half of the signal is hidden from the wet antenna attenuation threshold: in this case we can ~~suggest~~ suppose that the antenna could be dry, due to wind or no rain directly on it, so the Aa threshold is too high (also noted by de Vos et al. (2019)). However, a simple sensitivity test, carried out to assess the impact of a decrement in the Aa value on the single link scores, did not show any substantial improvement, especially when its results are extended to the whole dataset. More information is provided in the supplementary material. The continuous scores for the wet-wet sample on the ~~whole~~ entire period show a good correlation with gauges but are poor in statistical relevance because of the high number of misses. They nevertheless confirm the tendency to ~~quantitatively~~ underestimate, by around ~~20%~~ 40% (ME=-0.40).

~~For this pair of links located in the hilly Reno river basin there is therefore both a (minor)issue in quantitative underestimation and a (major) issue in the detection capability of the NLA classification.~~

4.1.3 ~~Melting layer case~~

~~The presence of melting layer is a critical issue in retrieving precipitation at the ground by microwave signal (i.e. radar sensors), so we take advantage of a winter time subset of data obtained from Vodafone for testing, but not numerous enough to perform a systematic analysis.~~

~~When the melting layer of a precipitating cloud crosses the link path (e.g. the antennas are on two hill crests and the link goes a few hundred meters high above the valley) the attenuation signal grows abruptly and leads to widespread overestimation. In Figures 3c, 3f, 3i we see it happening over Vergato on 3 March 2016. One of the two CMLs close to Vergato station has no estimates (it was probably rejected by the algorithm during initial filtering) but the signal is evident nonetheless.~~

~~This issue could be mitigated via the joint use of radar and CML measurements: the melting layer which produces strong attenuation is the same that causes the "bright band" in the radar reflectivity maps and is thus easily detected.~~

4.2 Gridded product verification

The verification of the RAINLINK gridded product (1 h cumulated on the 5 km × 5 km grid) with respect to the ERG5 product is first performed at the highest available resolution , ~~therefore~~ (grid box by grid box), since the two products intentionally share the same interpolation grid. Secondly, the comparison is carried out at the ~~basin-scale~~ basin scale by matching spatially averaged time series over areas of different size, in parallel with other operational precipitation products available at Arpae-SIMC.

4.2.1 Highest ~~resoluton~~ resolution matching

Figure 5 shows a scatter density plot for the whole dataset over the entire period. CML estimates from RAINLINK in northern Italy over uneven ground have an overall underestimating performance of -26% on the accumulated rain over the two months. The CV is ~~0.77~~ 0.78 and R^2 (the square of the already defined Pearson's correlation coefficient CC) is ~~0.47~~0.46, based on a sample of ~~10638~~ 10672 total wet hours. To make easier the comparison with past works, we computed continuous indicators with the filter set as Reference > 0.1 mm and with no filtering at all. Results with the first setting yields slightly worse indicators, increasing the ME to -0.41 and the CV to 0.95, with a second digit increase of R^2, around 0.5. The no-filter run shows values of ME and R^2 in line with our original results, while CV is greatly affected by very small rainrates. These performances are in good agreement with similar studies (Overeem et al., 2013, 2016b) despite the differences in the products involved: comparison between our results, with both filters, and the ones presented in the mentioned works are shown in Table 2.

~~As mentioned~~For the rest of the analysis, the data set was filtered so that only the rainfall depths of the grid boxes in which both CMLs and ERG5 reported more than 0.1 mm were used for the ~~quantitative~~ continuous indicators. The performances of the rain detection capabilities with respect to this threshold are evaluated separately with the already presented categorical scores (see Section 3.3).

The ~~means~~ mean values of categorical and continuous indicators are computed ~~on~~ in five areas, with a different extension (S) and averaged Link Coverage ($\overline{LC}$)~~, and~~. They are reported in Table 3, ranked according to the $\overline{LC}$ value: Parma Province (PP), Total Area (TA), Parma ~~river~~ River Basin (PRB), Bologna Province (BP), Reno River Basin (RRB). The total area and the two provinces do not have any specific hydrological meaning, but could be seen as a good foretype of larger river basins with heterogeneous terrain (see Figure 1). All normalized indicators are relative to the average reference (ERG5) rain rate. Numbers in bold (italics) are the best (worst) value in the column.

The general tendency of the RAINLINK product is also confirmed also for the sub-areas (excluding RRB for now) to underestimate the rain occurrence (MB < 1), with a relatively low value of POD (0.48 to 0.57) in all areas. The FAR is also rather ~~low~~small, (0.28 to 0.32), resulting in ETS values (0.38 to 0.43) comparable to the values obtained in the validation of other precipitation estimates, e.g. as the ones available from satellite observations (Puca et al., 2014; Feidas et al., 2018). Mean Error confirms the underestimation of rain amount (ME between ~~-0.19~~ 0.18 and -0.34), CV ranges between 0.73 and 0.80, CC between 0.62 and 0.74.

The averages over the Reno River Basin stand out for all the indicators, either positively or negatively~~,~~; therefore they need a separate discussion. As highlighted in Table 3 by bold and italics fonts, RRB has half the FAR the other samples have (0.16), almost ten points less CV (~~0.63) and almost~~ 0.62) and nearly fifteen points better CC (0.8, which is unexpectedly high), with the mean errors aligned to the other samples. The higher accuracy in the estimates is reached at the expense of POD, ETS and ~~BIAS~~MB: around 50~~%~~ % of the rainfall duration is lost in this area. The main peculiarity of the RRB area is the high $\overline{LC}$, which is 50% higher than the rest of the ~~field~~regions. We can infer that the higher coverage led~~, via the nearby links approach, to a very strict~~ to a more selective NLA classification, which reduced FAR and ~~produced some misses. This is in agreement with what is seen in the single link analysis of the Vergato CMLs (Section 4.1.2), which are located inside the Reno river basin.~~ POD.

~~Since a sensitivity to LC seems likely~~The marked improvement of continuous indicators for RRB suggests that the quantitative matching between estimated and reference could be positively related to $\overline{LC}$. Thus, we further investigate its effect on scores by grouping each grid box by LC quartiles, regardless of the actual geographical location, and reported the results in Figure 6. ~~For four~~ Five out of six indicators improve as LC increases (FAR, MAE, ETS, CC and CV), among which the most striking is the FAR, ~~as it had already been noticed in the case of the Reno river basin. The~~ while POD remains mostly unchanged, ~~therefore the ETS improves only slightly, but we do not , however, see the POD sharp decline that occurred in the Reno basin: this suggests~~ allowing the ETS improvement.

It seems that the sensitivity to LC could explain the improvement in the FAR of the RRB area, but not the sharp decline in the POD, suggesting that LC is probably not the only variable at play ~~there.~~ in Reno basin. These results integrate the findings of Overeem et al. (2016b), that highlighted the positive impact of higher $\overline{LC}$ on CV and CC at lower spatial resolution. Other studies will be conducted in the future to ~~better investigate~~ investigate better these problems, as well as a local calibration of the algorithm parameters (e.g. Aa and α).

The values presented above are fully comparable, and in many cases better, than the ones obtained for the main satellite based rainfall products in similar regions. Petracca et al. (2018) analysed over Italy the instantaneous estimate of the Global Precipitation Measurement - Dual-frequency Precipitation Radar (GPM-DPR), considered as the most reliable and accurate instrument to measure precipitation from space. Over a footprint of a size comparable to the one used in this paper, the best value of CC is 0.57, while the CV was between 1 and 2. Other validation studies of GPM-DPR products in the alpine region (Speirs et al., 2017) obtained relatively good POD (up to 0.78), FAR (below 0.08) and CC (up to 0.63) over flat terrain, with a dramatic drop of the skill indicators when areas with complex topography ~~is considered~~ are considered.

4.2.2 Case studies

To assess the performance of RAINLINK with respect to the structure of rainfall fields we focus the analysis on three one-day events with different characteristics for which the RAINLINK provided results of varying quality. The best performance was achieved on May 19 (see Figure 4, left), when an intense event was characterized by a few convective episodes on the Apennines, in the Parma Province. Precipitation peaks were around 90 $mm\ day^{-1}$ (see Figure 4c), maximum and mean hourly rainrates were about 24 and 2.6 $mm\ h^{-1}$, respectively (see Table 4). A large area of widespread moderate precipitation over the Bologna Province (Figure 4a) is also present. RAINLINK is able to localize precipitation local maxima (Fig. 4b), even if it occurred in areas where link coverage is relatively poor (see Figure 1), providing also accuracy in the peaks intensity. Estimated PDF matches closely the ERG5 curve, indicating that all rainrates are represented in the estimates (see Figure 4d), even if underestimation is always present, more marked for highest rainrates. Numerical indicators confirm the goodness of the estimate, in terms of wet area detection (ETS=0.59) and relative error (CV=0.69), while the fractional amount of rain lost by the estimate is low (ME=-0.29).

The second case (11 May) shows a more patchy rainfall field (Figure 4e), which resulted from a series of storms that occurred in the area during the day. Maximum and mean rates are lower with respect to the first case (Figures 4g, 4h), as well as the wet fraction of overall samples (see Table 4). Some local peak is correctly located (especially in Bologna Province), as shown in Figure 4f, and some other, in Parma Province and particularly on the Apennines, is missing. In this case the underestimate is marked for all rainrates resulting in higher ME (-0.40) and lower POD (0.66).

A completely different scenario is represented by case three (May 12), when ERG5 measured light to moderate precipitation (see Figure 4i), with peaks on the Apennines, and a much lower fraction of wet samples. RAINLINK (Figure 4j) is not able to estimate the highest rainrates and neither to locate the area with higher intensity. Moreover, it find a spurious peak in the northern area of Bologna Province, not detected by ERG5. Here the fractional amount of rain loss is -65%, the POD is low, and an increase of FAR is also to be remarked, indicating that underestimation dominates at all rainrates (see Figures 4k, 4l), but in case of light rain, overestimation could also take place.

This analysis points out that RAINLINK is undoubtedly able to resolve small size, short-living episodes, even providing quantitatively accurate estimates. Also in case of widespread moderate precipitation the overall rain pattern is effectively represented, with some underestimation of the numerical values. On the other side, the detection of light and intermittent rainfall seems to be the main challenge, probably due to the impact of the wet antenna attenuation and NLA approach, as already seen in previous sections. Finally, we note also that discrepancies can be related to the discontinuity of link distribution across the borders of the considered areas.

4.2.3 Area-average matching

In this Section, the matching between estimate and reference field is performed at basin (and Province) scales, comparing hourly rain amounts averaged over areas of different sizes. The areas selected for this evaluation are the ones introduced in the previous Section: two of them are selected chosen because of direct hydrological interest (RRB and PRB), while the other three (BP, PP and TA) are chosenselected to assess the impact of the increasing target area.

The categorical (rain/no-rain In Table 5 we present the categorical indicators calculated around the 0.1 mm h^{-1} threshold) and continuous (and the continuous indicators calculated on wet-wet occurrences) indicatorsonly, for the five mentioned areas are reported in Table 5, listed this time in order of increasing area size. In general, best performances are found for the largest areas (BP and TA), while the smallest ones (PRB and RRB) show the worst values. CML product underestimatunderestimates precipitation occurrence (MB between 0.41 and 0.70) and amount (ME between -0.23 and -0.43-0.18 and -0.34) at all scales. Due to the areal averaging, CC is markedly higher than the high-resolution values reported in Table 3. The characteristic behaviour of RRB (lowest FAR and POD, highest CC) remains alsoalso remains in this case.

The same areal-averaged statistical indicators have been computed also for also been computed for all the operational products available at Arpae-SIMC for operational routine use and described in Section 2.2, all reported to a reported to an hourly scale and compared with the ERG5 product. We show in Figure 7 the values of the statistical indicators as a function of the target area.

The raingauge product, obtained by averaging the measurements of the raingauges in the area, performs similarly to its interpolated version ERG5, as expected, and diverges only for small areas, where the impact of a single sensor in disagreement with neighbours is the highest.

Radar product shows, in this metric, almost the same performance with or both with and without the gauge adjustment. This can be expected since the radar adjustment happens above the raingauge locations but does not ensure the consistency of the areal average of the whole rain field. The adjustment also affects mainly higher rainrates than our 0.1 mmthreshold and has lower performances as the spatial variance increases, e.g. in cases of small scale convection. Both have very good detection capabilities (POD is almost 1) but high rates of false alarms (FAR around 0.5) and marked quantitative discrepancies (MAE around 10.9, CV between 1.5 0.75 and 2). Radarhowever, , however, can see finer precipitation structure given its spatially continuous coverage, while rain gauge (as well as the reference product ERG5) and CML networks (with point-like and line-integrated observations) ~~,~~ are both prone to miss some information from small scale events, often observed between meteorological spring and summer in Italy (see Supplement).

The CML product outperforms both radar products in terms of CC, CV, MAE and FAR, while it ~~clearly~~ lacks in detection capability (CMLs POD between 0.4 and 0.6), as discussed in ~~the previous Section~~ previous Sections. CML retrieval process, being based on electromagnetic attenuation instead of back-scattering, does not share the radar's high sensitivity to the size distribution of the hydrometeors (Leijnse et al., 2008), thus making CML a more robust sensor , in the sense that the same coefficients can be effectively applied regardless the type of precipitation. Figure 8 shows ~~the overestimating behaviour of both radar products and the average underestimation of the CML product. The~~ that the overestimating and underestimating behaviours of radar and CML products, respectively, can be seen as complementary. For radar, the spread is more relevant ~~for radars than~~ than for CML, but it has to be remarked that the latter has a smaller sample size due to the already mentioned low POD issues.

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 latencyof the precipitation product~~ high relevance also their latency, i.e. the time taken from the acquisition of the ~~basic~~ primary data (the occurrence of the event) and the delivery of the product in a ready-to-use form. In Table 6 are reported the latency and sampling characteristics of the four precipitation products we took for comparison, along with CML product. CMLs operational specifications refer to an implementation of the RAINLINK algorithm as part of a real-time service, tested in 2019 by MEEO S.r.l. within the RainBO project (LIFE15 CCA/IT/000035). It can be seen that the combination of ~~short latency~~ short-latency and good coverage provided by CMLs is unmatched by all products except the raw radar, which though lacks the required quantitative accuracyIt is to the operators' preference, based on product error structure, current meteorological conditions, and user's requirements, to make use of the most suitable product or combination of them.

5 Conclusions

An assessment of the rainfall retrieval capability of CML opportunistic sensors over ~~heterogeneous~~ complex terrain in northern Italy is conducted at ~~various levels~~ different spatial and temporal scales for two months of data. We implemented the open-source RAINLINK algorithm in a new area and context, where no regional CML studies had previously been performed. We evaluated its performance through a complete validation scheme which involves operational precipitation products as a benchmark, gauging in the process also the implementation effort and identifying major strengths and weaknesses to make profitable use of CML products.

First, ~~a sample of 28~~ 26 CMLs (out of the total ~~319) is~~ 308) are compared with the closest ~~raingauge~~ raingauges at a 15 min scale. Overestimation and underestimation of rain amount are both present, though the latter appears dominant ~~, with a~~ . A marked variability among different links ~~, with~~ does not prevent to achieve a generally acceptable skill (CC from ~~0.67 to 0.86). Higher~~ 0.50 to 0.88). The wet-dry classification approach and the value of the wet antenna correction may generate a loss of rain amount in case of small scale and/or intermittent episodes. Finally, higher elevation CMLs show in general

~~worst performances, being also prone to known shortcomings, such as melting layer across the microwave path, which led to significant performance drops~~worse performances.

Interpolated products obtained from the full sample of ~~319~~ 308 links confirm that a non-negligible quantity of rain is missed (normalized Mean Error is -0.26, overall CC is 0.68 and overall CV is ~~0.77~~0.78), but also show that the rain ~~detection~~retrieval capability is suitable for the operational application, especially if the product is integrated over large areas (CC rises to 0.92). Higher link densities increase the quality of the CML estimates at both gridbox and basin scales, mostly in terms of decreased FAR.

Performances at event scale show enhanced skill in case of heavy precipitation, even in case of small scale rain episodes, while problems arise when light/moderate rainrates challenge the algorithm in the ways we already identified in the single-link analysis. Negative impact on the overall results comes from areas with poor sensor coverage ~~(i.e. low density of link), especially in terms of increased false alarms. The results show a high variability of responses, with an overall skill in agreement with other CML studies.~~

~~Furthermore, the validation scheme implemented for this work, if a more complete dataset would become available, would~~ be ~~used to improve the performance of RAINLINK by tuning the algorithm parameters (NLA radius, Aa, α) on a more complete training sample specific of the study area~~especially near the border of the areas, but it should be considered that also reference rainfall fields can be affected by shortcomings of the same nature.

~~Finally~~Furthermore, when compared to other products currently available for ~~operational~~ real-time operational exploitation, CML sensors show similar or better abilities than their counterparts, especially if latency is also taken into account. Hence an integration of microwave links sensors in an operational service is highly desirable, even without a proper calibration of the algorithm to the local climatology and CML network characteristics.

When a more complete dataset would become available the validation scheme implemented for this work could be promptly used to tune the RAINLINK parameters (NLA radius, Aa, α) on a training sample specific of the study area.

Code and data availability. CML data were provided by Vodafone Italia S.p.A.. via direct purchase from MEEO S.r.l. and are not publicly available. Gauge data from Emilia-Romagna are freely available at https://simc.arpae.it/dext3r/. Radar reflectivities in near real-time are freely available at https://www.arpae.it/sim/?osservazioni_e_dati/radar, while derived rain products and ERG5 analyses are available upon request at Arpae-SIMC (https://www.arpae.it/sim/). The core algorithm is available (open source) at https://github.com/giacom0rovers1/RAINLINK and was forked from https://github.com/overeem11/RAINLINK on the 26th of August 2019 (RAINLINK version 1.14).

Author contributions. GR adapted the RAINLINK code to Italian data, ran the analysis, plotted the data and contributed to the interpretation of the results and to the writing of the manuscript. PPA and AF performed the reference data pre-processing and contributed to data analysis. FP contributed to the design of the validation strategy, to the interpretation of the results and to the writing of the paper.

Competing interests. The Author declare that no competing interests are present.

Acknowledgements. This work has been partially funded by the Life EU Project RainBO (LIFE15 CCA/IT/000035). The Authors thank Stefania Pasetti and Marco Folegani of MEEO S.r.l. (www.meeo.it) for their support, and are grateful to D. Vecchiato and A. Viaro of Vodafone Italia S.p.A. for the technical assistance with the data. We also thank Aart Overeem, for having developed and released open source the RAINLINK algorithm and for the kind feedback and support he provided to this research.

References

Abdulrahman, A., Bin Abdulrahman, T., Bin Abdulrahim, S., and Kesavan, U.: Comparison of measured rain attenuation and ITU-R predictions on experimental microwave links in Malaysia, International Journal of Microwave and Wireless Technologies, 3, 477–483, https://doi.org/10.1017/S1759078711000171, 2011.

Alberoni, P. P., Andersson, T., Mezzasalma, P., Michelson, D. B., and Nanni, S.: Use of the vertical reflectivity profile for identification of anomalous propagation, Meteorological Applications, 8, 257–266, https://doi.org/10.1017/S1350482701003012, 2001.

Alberoni, P. P., Fornasiero, A., Roversi, G., Pasetti, S., Folegani, M., and Porcù, F.: Comparison between different QPE based on: Microwave Links, Radar adjusted and Gauges, in: 10th European Conference on Radar in Meteorology and Hydrology, pp. 851–860, KNMI, https://doi.org/10.18174/454537, 2018.

Amorati, R., Alberoni, P., and Fornasiero, A.: Operational Bias Correction of Hourly Radar Precipitation Estimate using Rain Gauges, in: Proceedings of the Seventh Eropean Conference on Radar in meteorology and Hydrology, http://www.meteo.fr/cic/meetings/2012/ERAD/extended_abs/QPE_007_ext_abs.pdf, 2012.

Antolini, G., Auteri, L., Pavan, V., Tomei, F., Tomozeiu, R., and Marletto, V.: A daily high-resolution gridded climatic data set for Emilia Romagna, Italy, during 1961-2010, International Journal of Climatology, 36, 1970–1986, 2016.

Bech, J., Codina, B., Lorente, J., and Bebbington, D.: The Sensitivity of Single Polarization Weather Radar Beam Blockage Correction to Variability in the Vertical Refractivity Gradient, Journal of Atmospheric and Oceanic Technology, 20, 845–855, https://doi.org/10.1175/1520-0426(2003)020<0845:TSOSPW>2.0.CO;2, 2003.

Berne, A. and Uijlenhoet, R.: Path-averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL029409, 2007.

Bianchi, B., Jan van Leeuwen, P., Hogan, R. J., and Berne, A.: A Variational Approach to Retrieve Rain Rate by Combining Information from Rain Gauges, Radars, and Microwave Links, Journal of Hydrometeorology, 14, 1897–1909, https://doi.org/10.1175/JHM-D-12-094.1, 2013.

Bowman, D. and Lees, J.: Near real time weather and ocean model data access with rNOMADS, Computers and Geosciences, 78, 88 – 95, https://doi.org/https://doi.org/10.1016/j.cageo.2015.02.013, 2015.

Caracciolo, C., Prodi, F., and Uijlenhoet, R.: Comparison between Pludix and impact/optical disdrometers during rainfall measurement campaigns, Atmospheric Research, 82, 137 – 163, https://doi.org/https://doi.org/10.1016/j.atmosres.2005.09.007, 14th International Conference on Clouds and Precipitation, 2006.

Chwala, C. and Kunstmann, H.: Commercial microwave link networks for rainfall observation: Assessment of the current status and future challenges, WIREs Water, 6, e1337, https://doi.org/10.1002/wat2.1337, 2019.

Cummings, R., Upton, G., Holt, A., and Kitchen, M.: Using microwave links to adjust the radar rainfall field, Advances in Water Resources, 32, 1003 – 1010, https://doi.org/https://doi.org/10.1016/j.advwatres.2008.08.010, weather Radar and Hydrology, 2009.

de Vos, L., Raupach, T., Leijnse, H., Overeem, A., Berne, A., and Uijlenhoet, R.: High-Resolution Simulation Study Exploring the Potential of Radars, Crowdsourced Personal Weather Stations, and Commercial Microwave Links to Monitor Small-Scale Urban Rainfall, Water Resources Research, 54, 10 293–10 312, https://doi.org/10.1029/2018WR023393, 2018.

de Vos, L. W., Overeem, A., Leijnse, H., and Uijlenhoet, R.: Rainfall Estimation Accuracy of a Nationwide Instantaneously Sampling Commercial Microwave Link Network: Error Dependency on Known Characteristics, Journal of Atmospheric and Oceanic Technology, 36, 1267–1283, https://doi.org/10.1175/JTECH-D-18-0197.1, 2019.

Doumounia, A., Gosset, M., Cazenave, F., Kacou, M., and Zougmore, F.: Rainfall monitoring based on microwave links from cellular telecommunication networks: First results from a West African test bed, Geophysical Research Letters, 41, 6015–6021, https://doi.org/10.1002/2014GL060724, 2014.

Feidas, H., Porcu, F., Puca, S., Rinollo, A., Lagouvardos, C., and Kotroni, V.: Validation of the H-SAF precipitation product H03 over Greece using rain gauge data, Theoretical and Applied Climatology, 131, 377–398, https://doi.org/10.1007/s00704-016-1981-9, 2018.

Fencl, M., Rieckermann, J., Schleiss, M., Stránský, D., and Bareš, V.: Assessing the potential of using telecommunication microwave links in urban drainage modelling, Water science and technology : a journal of the International Association on Water Pollution Research, 68, 1810–8, https://doi.org/10.2166/wst.2013.429, 2013.

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

Fenicia, F., Pfister, L., Kavetski, D., Matgen, P., Iffly, J.-F., Hoffmann, L., and Uijlenhoet, R.: Microwave links for rainfall estimation in an urban environment: Insights from an experimental setup in Luxembourg-City, Journal of Hydrology, 464-465, 69 – 78, https://doi.org/https://doi.org/10.1016/j.jhydrol.2012.06.047, 2012.

Fornasiero, A., Bech, J., and Alberoni, P. P.: Enhanced radar precipitation estimates using a combined clutter and beam blockage correction technique, Natural Hazards Earth System Science, 6, 697–710, https://doi.org/10.5194/nhess-6-697-2006, 2006.

Gou, Y., Chen, H., and 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 Sensing, 11, 2335, https://doi.org/10.3390/rs11202335, 2019.

Grecu, M., Olson, W., Munchak, s., Ringerud, S., Liao, L., Haddad, Z., Kelley, B., and McLaughlin, S.: The GPM combined algorithm, Journal of Atmospheric and Oceanic Technology, 33, 2225–2245, https://doi.org/10.1175/JTECH-D-16-0019.1, 2016.

Grum, M., Kraemer, S., Verworn, H.-R., and Redder, A.: Combined use of point rain gauges, radar, microwave link and level measurements in urban hydrological modelling, Atmospheric Research, 77, 313 – 321, https://doi.org/https://doi.org/10.1016/j.atmosres.2004.10.013, precipitation in Urban Areas, 2005.

Haese, B., Hörning, S., Chwala, C., Bardossy, A., Schalge, B., , and Kunstmann, H.: Stochastic Reconstruction and Interpolation of Precipitation Fields Using Combined Information of Commercial Microwave Links and Rain Gauges, Water Resources Research, 53, 10 740–10 756, https://doi.org/https://doi.org/10.1002/2017WR021015, 2017.

Harden, B., Norbury, J., and White, W.: Attenuation/rain-rate relationships on terrestrial microwave links in the frequency range 10- -40 GHz, Electronics Letters, 14, 154 – 155, https://doi.org/10.1049/el:19780103, 1978.

Huuskonen, A., Saltikoff, E., and Holleman, I.: The Operational Weather Radar Network in Europe, Bulletin of the American Meteorological Society, 95, 897–907, https://doi.org/10.1175/BAMS-D-12-00216.1, 2014.

i Ventura, J. F., Boumahmoud, A., Fradon, B., Dupuy, P., and Tabary, P.: 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, Quarterly Journal of the Royal Meteorological Society, 138, 2212–2228, https://doi.org/10.1002/qj.1934, 2012.

Jameson, A. R.: Estimating the Path-Average Rainwater Content and Updraft Speed along a Microwave Link, Journal of Atmospheric and Oceanic Technology, 10, 478–485, https://doi.org/10.1175/1520-0426(1993)010<0478:ETPARC>2.0.CO;2, 1993.

Koistinen, J. and Puhakka, T.: An improved spatial gauge-radar adjustment technique, in: Preprints of the 20th Conference on Radar Meteorology, p. 179–186, Am. Meteorol. Soc., 1981.

Lanza, L. and Stagi, L.: Non-parametric error distribution analysis from the laboratory calibration of various rainfall intensity gauges, Water science and technology : a journal of the International Association on Water Pollution Research, 65, 1745–52, https://doi.org/10.2166/wst.2012.075, 2012.

Leijnse, H., Uijlenhoet, R., and Stricker, J.: Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation, Advances in Water Resources, 31, 1481–1493, https://doi.org/10.1016/j.advwatres.2008.03.004, 2008.

Leijnse, H., Uijlenhoet, R., and Berne, A.: Errors and Uncertainties in Microwave Link Rainfall Estimation Explored Using Drop Size Measurements and High-Resolution Radar Data, Journal of Hydrometeorology, 11, 1330–1344, https://doi.org/10.1175/2010JHM1243.1, 2010.

Leinse, H.: Hydro-meteorological application of microwave links - Measurement of evaporation and precipitation, Ph.D. thesis, Wageningen University, 2007.

Mugnai, A., Casella, D., Cattani, E., Dietrich, S., Laviola, S., Levizzani, V., Panegrossi, G., Petracca, M., Sanò, P., Di Paola, F., Biron, D., De Leonibus, L., Melfi, D., Rosci, P., Vocino, A., Zauli, F., Pagliara, P., Puca, S., Rinollo, A., Milani, L., Porcù, F., and Gattari, F.: Precipitation products from the hydrology SAF, Natural Hazards and Earth System Sciences, 13, 1959–1981, https://doi.org/10.5194/nhess-13-1959-2013, 2013.

Mulangu, C. and Afullo, T.: Variability of the propagation coefficients due to rain for microwave links in southern Africa, Radio Science - RADIO SCI, 44, https://doi.org/10.1029/2008RS003912, 2009.

Nurmi, P.: Recommendations on the verification of local weather forecasts, ECMWF Technical Memorandum, 430, 2003.

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Measuring urban rainfall using microwave links from commercial cellular communication networks, Water Resources Research, 47, https://doi.org/10.1029/2010WR010350, 2011.

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Country-wide rainfall maps from cellular communication networks, Proceedings of the National Academy of Sciences, 110, 2741–2745, https://doi.org/10.1073/pnas.1217961110, 2013.

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

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

Petracca, M., D'Adderio, L., Porcu, F., Vulpiani, G., Stefano, S., and Puca, S.: Validation of GPM Dual-Frequency Precipitation Radar (DPR) Rainfall Products over Italy, Journal of Hydrometeorology, 19, 907–925, https://doi.org/10.1175/JHM-D-17-0144.1, 2018.

Porcù, F., Milani, L., and Petracca, M.: On the uncertainties in validating satellite instantaneous rainfall estimates with raingauge operational network, Atmospheric Research, 144, 73–81, https://doi.org/10.1016/j.atmosres.2013.12.007, 2014.

Puca, S., Porcu, F., Rinollo, A., Vulpiani, G., Baguis, P., Balabanova, S., Campione, E., Erturk, A., Gabellani, S., Iwanski, R., Jurašek, M., Kaňák, J., Kerényi, J., Koshinchanov, G., Kozinarova, G., Krahe, P., Lapeta, B., Labo, E., Milani, L., and Gattari, F.: The validation service of the hydrological SAF geostationary and polar satellite precipitation products, Natural hazards and earth system sciences, 14, 871–889, https://doi.org/10.5194/nhess-14-871-2014, 2014.

Rahimi, A. R., Holt, A. R., Upton, G. J. G., and Cummings, R. J.: Use of dual-frequency microwave links for measuring path-averaged rainfall, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002JD003202, 2003.

Rincon, R. and Lang, R.: Microwave link dual-wavelength measurements of path-average attenuation for the estimation of drop size distributions and rainfall, Geoscience and Remote Sensing, IEEE Transactions on, 40, 760 – 770, https://doi.org/10.1109/TGRS.2002.1006324, 2002.

Rios Gaona, M., Overeem, A., Raupach, T., Leijnse, H., and Uijlenhoet, R.: Rainfall retrieval with commercial microwave links in São Paulo, Brazil, Atmospheric Measurement Techniques Discussions, pp. 1–21, https://doi.org/10.5194/amt-2017-287, 2017.

Ryzhkov, A. and Zrnic, D.: Radar Polarimetry for Weather Observations, Springer Atmospheric Sciences, https://doi.org/10.1007/978-3-030-05093-1, 2019.

Saltikoff, E., Haase, G., Delobbe, L., Gaussiat, N., Martet, M., Idziorek, D., Leijnse, H., Novák, P., Lukach, M., and Stephan, K.: OPERA the Radar Project, Atmosphere, 10, 320, https://doi.org/10.3390/atmos10060320, 2019.

Serafin, R. J. and Wilson, J. W.: Operational Weather Radar in the United States: Progress and Opportunity, Bulletin of the American Meteorological Society, 81, 501–518, https://doi.org/10.1175/1520-0477(2000)081<0501:OWRITU>2.3.CO;2, 2000.

Shepard, D.: A Two-Dimensional Interpolation Function for Irregularly-Spaced Data, in: Proceedings of the 1968 23rd ACM National Conference, ACM '68, p. 517–524, Association for Computing Machinery, New York, NY, USA, https://doi.org/10.1145/800186.810616, 1968.

Skofronick-Jackson, G., Petersen, W. A., Berg, W., Kidd, C., Stocker, E. F., Kirschbaum, D. B., Kakar, R., Braun, S. A., Huffman, G. J., Iguchi, T., Kirstetter, P. E., Kummerow, C., Meneghini, R., Oki, R., Olson, W. S., Takayabu, Y. N., Furukawa, K., and Wilheit, T.: The Global Precipitation Measurement (GPM) Mission for Science and Society, Bulletin of the American Meteorological Society, 98, 1679–1695, https://doi.org/10.1175/BAMS-D-15-00306.1, 2017.

Speirs, P., Gabella, M., and Berne, A.: A comparison between the GPM dual-frequency precipitation radar and ground-based radar precipitation rate estimates in the Swiss Alps and Plateau, Journal of Hydrometeorology, 18, 1247–1269, https://doi.org/10.1175/JHM-D-16-0085.1, 2017.

Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., and Hong, Y.: Hydro-Meteorological Assessment of Three GPM Satellite Precipitation Products in the Kelantan River Basin, Malaysia, Remote Sensing of Environment, 240, 11 697, https://doi.org/10.1016/j.rse.2020.111697, 2020.

Tokay, A., D'Adderio, L. P., Porcù, F., Wolff, D. B., and Petersen, W. A.: A Field Study of Footprint-Scale Variability of Raindrop Size Distribution, Journal of Hydrometeorology, 18, 3165–3179, https://doi.org/10.1175/JHM-D-17-0003.1, 2017.

Turner, D. J. W. and Turner, D.: Attenuation due to rainfall on a 24 km microwave link working at 11, 18 and 36 GHz, Electronics Letters, 6, 297–298, https://doi.org/10.1049/el:19700208, 1970.

Upton, G., Holt, A., Cummings, R., Rahimi, A., and Goddard, J.: Microwave links: The future for urban rainfall measurement?, Atmospheric Research, 77, 300–312, https://doi.org/10.1016/j.atmosres.2004.10.009, 2005.

van Leth, T. C., Leijnse, H., Overeem, A., and Uijlenhoet, R.: Estimating raindrop size distributions using microwave link measurements, Atmospheric Measurement Techniques Discussions, 2019, 1–27, https://doi.org/10.5194/amt-2019-51, 2019.

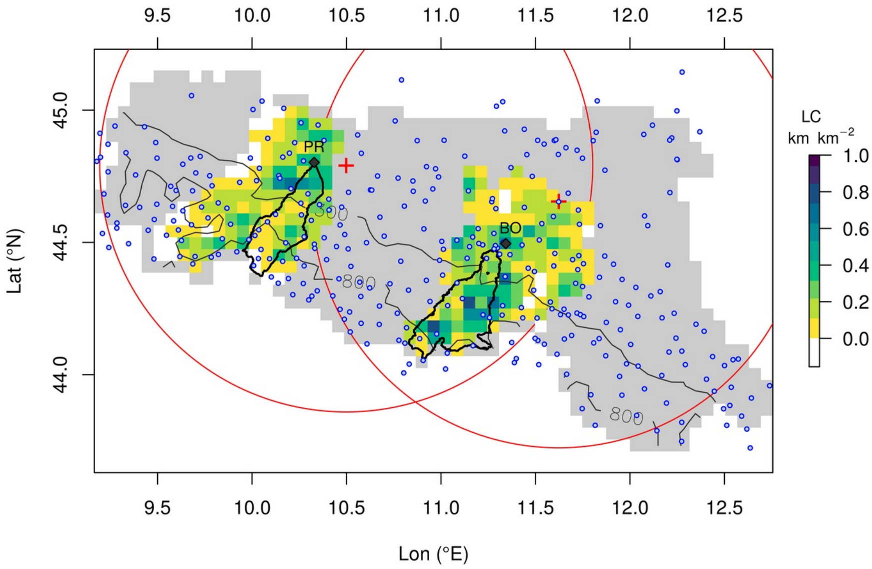


Figure 1. Map of the Emilia-Romagna region in Northern Italy (grey area). The coloured areas are the two Provinces where the CML estimates are computed (the color scale represents the Link Coverage, LC) and black thick lines delimit the two river basins (Parma, to the east, and Reno). Blue dots and red crosses indicate operational raingauge and weather radar locations, respectively, while red circles are the 100 km radar coverage. Thin black lines show two elevation contours (300 and 800 m a.s.l.). The main cities in the area Bologna and Parma are indicated with the black diamonds.

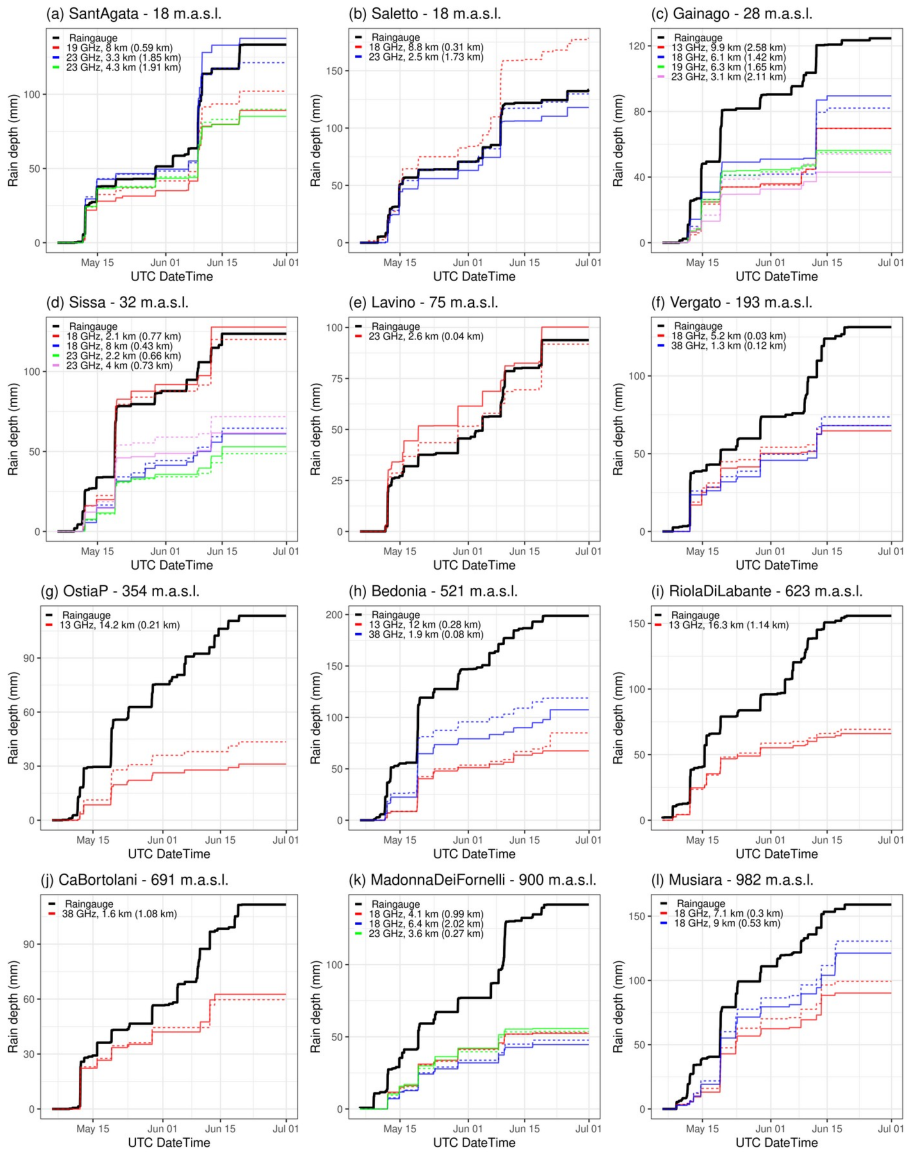


Figure 2. Accumulated rain depths over the whole period for the ~~28~~26 CMLs selected for the single-link analysis. Each tile is named by the corresponding rain gauge, whose accumulated rain depth is shown by the black thick line. Solid and dashed lines represent the two directions (if both active) for every CML (distinguished by colour).

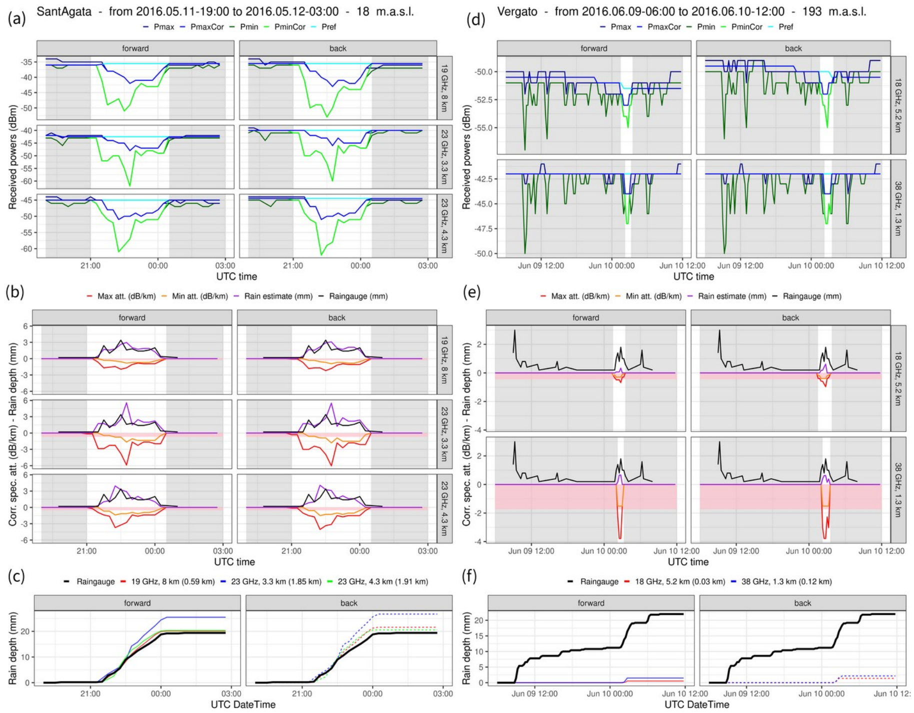


Figure 3. ~~Case studies of the single~~ Single link analysis ~~;~~ for Sant'Agata (from 11.05-19:00UTC to 12.05-03:00UTC) and Vergato (from 09.06-06:00UTC to 10.06-12:00UTC): (a) and (d) show the received signals (P_{max}, blue; P^{Cor}_{max}, light blue; P_{min}, green; P^{Cor}_{min}, light green; P_{ref}, cyan); (b) and ~~e~~(e) show ~~the specific~~ maximum attenuations (red), minimum attenuation (orange)~~and the rain estimates ,~~ estimated rainrate (purple)~~compared with the~~, and gauge measurements (~~blue dots~~black)~~, d~~; in (c) ~~, e)~~ and (f) the ~~received signals, g), h) and~~cumulated raingauge rainrate (~~i~~black) is plotted with the ~~cumulated rainfall amounts~~link estimates. Grey background corresponds to intervals labelled as dry by the NLA classification. ~~Y axes resolutions~~ Y-axes ranges are specific for each CML as received powers differ between different ~~pathlengths~~path lengths.

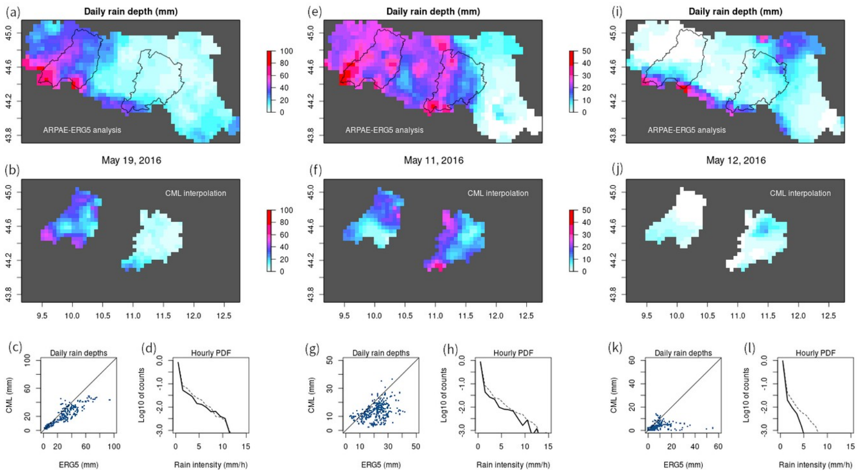


Figure 4. Analysis three one-day case studies (May 19, left; May 11, center; May 12, right): (a), (e) and (i) daily cumulated ERG5 precipitation; (b), (f) and (j) daily cumulated RAINLINK precipitation; (c), (g) and (k) scatterplot between the two daily precipitation; (d), (h) and (l) PDF of hourly rain rates.

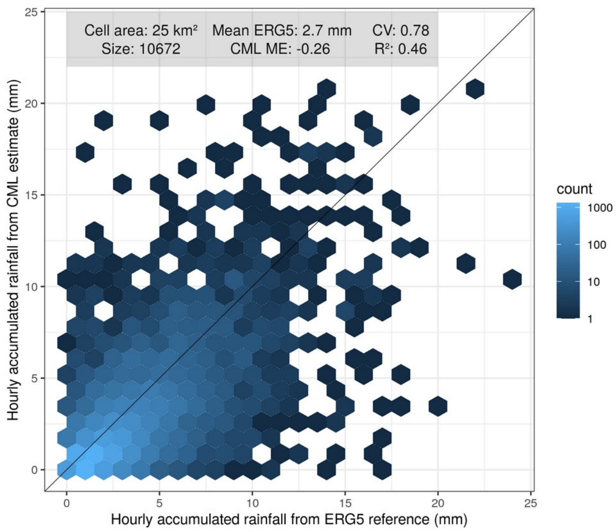


Figure 5. Hourly validation of link rainfall maps against ERG5 rainfall maps at grid box scale (highest resolution). Only the rainfall depths in which both CMLs and ERG5 measured > 0.1 mm were used. The black line is the y=x line.

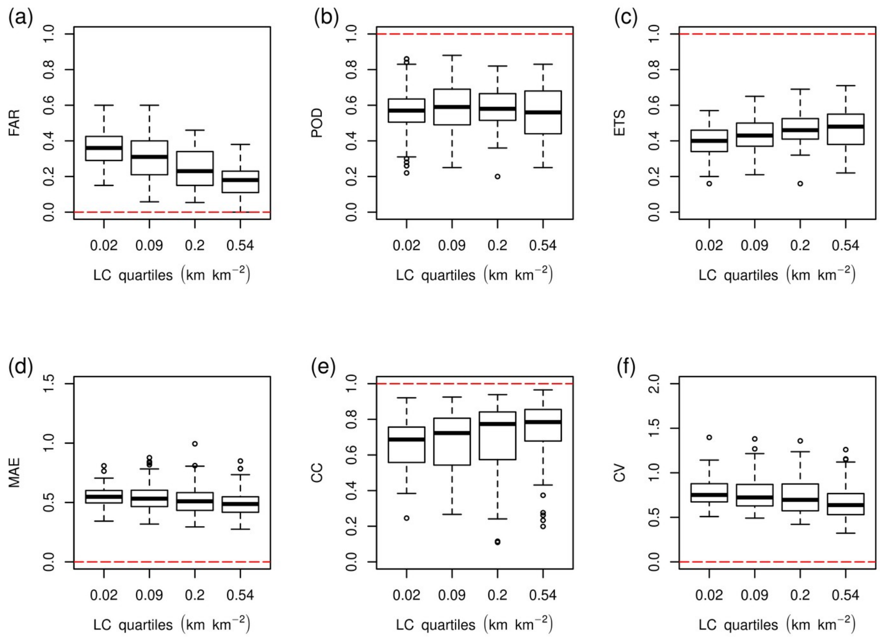


Figure 6. Distributions of four statistical indicators computed for every grid box and grouped in boxplots by quartiles of the link coverage *LC* (labelled by the quartiles centre). Red dashed lines are the optimal values for each score.

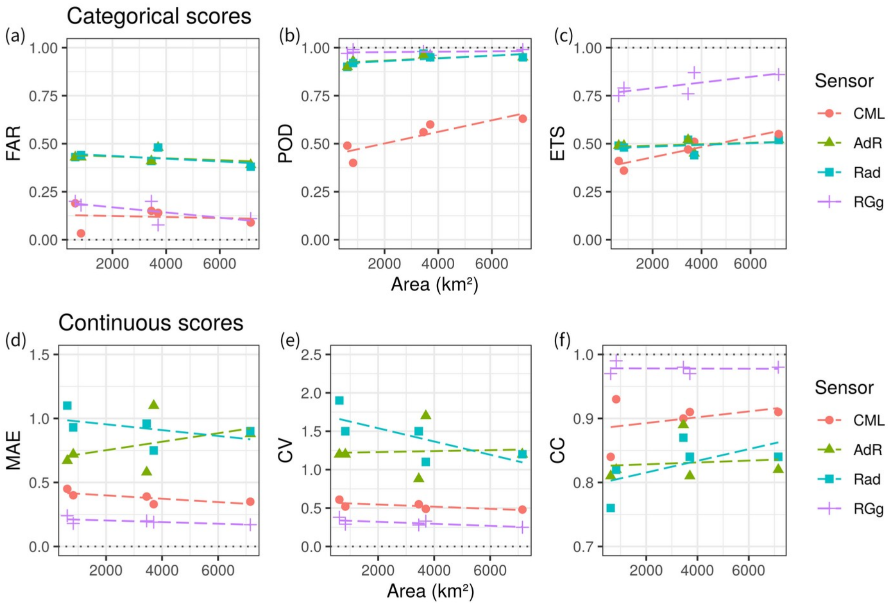


Figure 7. Scores of the areal-averaged rainfall amounts grouped per sensor and plotted against basin area. Linear fits are highlighted with dashed lines. The CML scores are ~~indicated~~ also indicated numerically in Table 5.

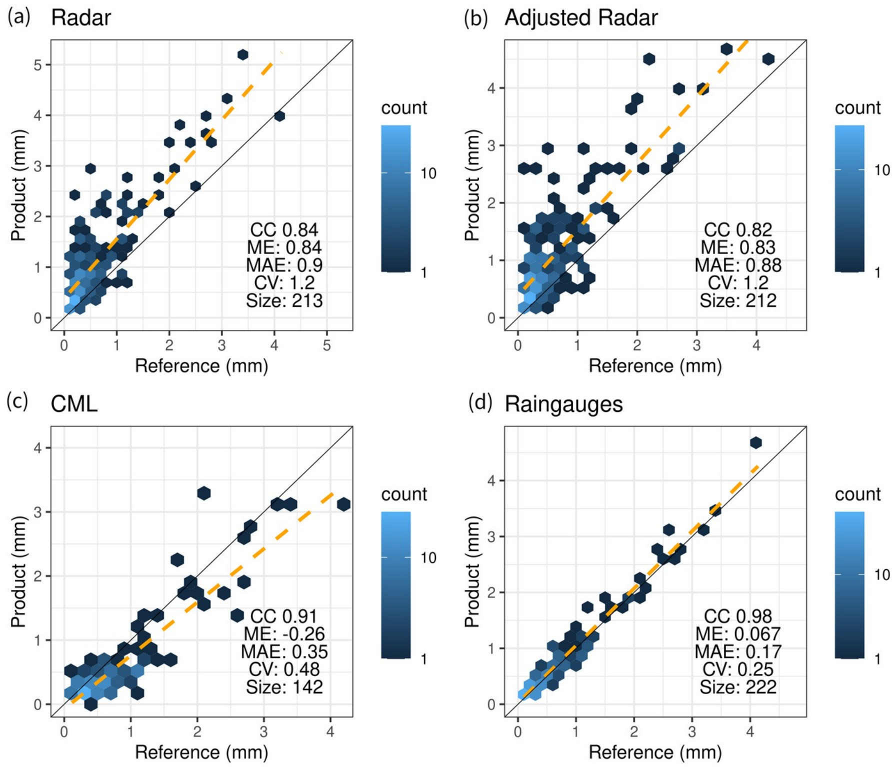


Figure 8. Comparison of hourly areal-averaged rainfall depths from the four products against the ERG5 reference. The total area (TA) wet-wet hours are considered.

Table 1. CML datasets comparison.

Variable	Unit	ER	NL	Overeem et al. (2013)	Overeem et al. (2016b)
Total area	km^2	7149	35500	35500	35500
CMLs	counts	~~319~~ 308	1527	1514	2044
sub-links	counts	~~625~~ 606	2473	2902	3383
LD	km^2	~~0.045~~ 0.043	0.043	0.043	0.058
LL	km	~~6.0~~ 5.8	2.9	3.1	3.6
BC	$km\ km^{-2}$	~~0.27~~ 0.25	0.13	0.13	0.21
$\bar{f}$	GHz	22.1	37.11	37-40	37-40

Table 2. Comparison with previous studies.Ref isa the reference rainrate, Pr is the product

Variable	Unit		ER (present work)		Overeem et al. (2013)	Overeem et al. (2016b)
Total time window	-		2 months		3 months	2.5 years
~~Filter - Ref. AND Product > 0.1 Ref. OR Product > 0.1 Ref. > 0.1~~ Time scale	min		60		15	60
Gridbox area	km^2		25		81	74
Reference	-		interpolated rain gauges		Gauge-adjusted radar	Gauge-adjusted radar
Filter	-	Ref. *and* Pr > 0.1 mm	Ref > 0.1 mm	none	Ref. *or* Pr > 0.1 mm	Ref > 0.1 mm
ME	-	-0.26	-0.41	-0.33	0.02	-0.16
CV	-	0.77	0.95	4.6	1.13	0.64
R^2	-	0.47	0.50	0.53	0.49	0.49

numbers in bold are obtained by performing the validation with the filter Ref. > 0.1 mmmm only

Table 3. Statistical indicators for each considered area, considering the highest resolution information (grid box scale), shown in ascending order of $\overline{LC}$. Continuous indicators are normalized and fractional. Values in bold (italics) are the best (worst) values in the column.

Area	$\overline{LC}$ (km km^{-2})	S (km^2)	FAR	POD	ETS	MB	ME	MAE	CV	CC
PP	0.17	3447	0.28	0.51	~~0.40~~ 0.41	0.71	*-0.34*	*0.55*	*0.80*	*0.62*
TA	0.18	7149	0.30	0.54	0.42	0.77	-0.26	0.52	0.77	0.68
PRB	0.19	624	0.30	0.48	0.38	0.69	-0.31	0.50	0.76	0.67
BP	0.19	3702	*0.32*	**0.57**	**0.43**	**0.83**	**-0.18**	~~0.49~~ 0.48	0.73	0.74
RRB	0.29	828	**0.16**	*0.39*	*0.35*	*0.47*	-0.31	**0.45**	**0.62**	**0.80**

Table 4. Rainfall characteristics and performance indicators for the three one-day case studies

Date	mean R (mm)	max R (mm)	wet fraction	FAR	POD	ETS	ME	CV	CC
19.05	2.60	24.0	0.37	0.10	0.77	0.59	-0.29	0.69	0.78
11.05	2.50	21.0	0.35	0.10	0.66	0.49	-0.40	0.76	0.82
12.05	1.80	14.0	0.16	0.20	0.58	0.46	-0.65	1.10	0.46

Table 5. Values of the statistical indicators for the mean rain amounts over each considered area, shown in ascending order of surface area S. Values in bold (italics) are the best (worst) values in the column.

Area	S (km^2)	$\overline{LC}$ ($km\ km^{-2}$)	FAR	POD	ETS	MB	ME	MAE	CV	CC
PRB	624	0.19	*0.18*	0.51	0.43	0.63	~~-0.36~~ -0.34	*0.45*	*0.61*	*0.84*
RRB	828	0.29	**0.03**	*0.40*	*0.36*	*0.41*	*-0.34*	~~0.48~~ 0.40	~~0.67~~ 0.52	**0.93**
PP	3447	0.17	~~0.15~~ 0.14	~~0.55~~ 0.57	~~0.47~~ 0.48	~~0.63~~ 0.66	~~-0.37~~ -0.34	~~0.44~~ 0.48	~~0.67~~ 0.56	~~0.91~~ 0.98
BP	3702	0.19	0.14	0.60	0.51	**0.70**	**-0.18**	**0.33**	~~0.66~~ 0.49	~~0.92~~ 0.91
TA	7149	0.18	~~0.09~~ 0.10	**0.64**	**0.55**	**0.70**	~~-0.32~~ -0.26	~~0.41~~ 0.35	**0.48**	~~0.92~~ 0.91

Table 6. Latency and spatial and temporal sampling of the considered precipitation products.

Product	Reference time step (min)	Latency (min)	Spatial resolution (km)
CML	15	20	5
Radar raw	5	15	1
Radar adj.	60	60	1
Raingauges raw	60	60	-
ERG5	60	1440	5