Dear Reviewer #2,
many thanks for your comments and the valuable suggestions on our manuscript. You will find below the answers to each comment you provided.

Major comments

1. **Lack of validation of new CML processing: The authors introduce a new CML data set of 15-minute min-max data. They apply and own processing which is described in an already published conference paper (Nebuloni et al, 2020b). The referenced paper is not very detailed in describing the method and in analysing the methods performance. It also uses a CML dataset from a different region. Hence, in this manuscript more details should be provided on the actual derivation of rain rates from the CML raw data and on analysing the resulting performance. The presented analysis only shows aggregated results, e.g. in Fig 5. This is not enough to understand how individual CMLs perform and if specific CMLs impact the spatial rainfall estimates in the surrounding negatively because of challenging signal fluctuations in the raw data. Showing individual time series of CML raw data together with derived rain rates and data from nearby gauges (or from a spatially interpolated gauge product) would help. Better would be a more quantitative analysis. More information about the pure rainfall estimates at the individual CMLs is important because errors there will propagate through the whole analysis. The fact that insensitive CMLs perform badly during low rain rates might not be the only reason for degraded performance of CML-rainfall in the hydrological simulation.**

   Following your suggestion, we added further information about CML data processing with respect to the ones provided in Nebuloni et al, 2020b, as well as a quantitative comparison between CML-based rainfall estimates and rain gauge measurements. The text below will be the modified subsection **Commercial Microwave Links** of the revised manuscript. It also integrates clarifications of your comments from **L152** to **L187** which will be individually addressed even later.

   CML raw data, collected by a network monitoring tool, are minimum and maximum values of the transmitted and received power levels (TSL and RSL, respectively) every 15 min. Microwave links of mobile networks are usually two-ways and provide dual-frequency operation. The CML data set used here has two to four channels available for every link, which usually permits to deal with missing or invalid data appearing sometimes over a certain channel. Procedures for the conversion of RSL into rainfall rate have been detailed by several authors (e.g., Schleiss and Berne, 2010; Fenicia et al., 2012; Overeem et al., 2016). As the format of the available CML data is the same as in Overeem et al., 2016, we built from the procedure outlined there. Specifically, data processing went through the following steps: (1) identification and removal of outliers and artifacts (i.e., occasional spikes, which are not caused by rain); (2) classification of each 15 min time slot into dry or wet (i.e., rainy) by thresholding the difference between maximum and minimum RSL values; (3) estimation of the baseline, i.e., the RSL in absence of rain; (4) calculation of total signal attenuation as the difference between the baseline and the actual RSL; (5) identification and subtraction of the components of total attenuation not due to rainfall (e.g. wet antenna attenuation); (6) conversion of rain attenuation into rainfall rate.
   Some details of the major processing steps are discussed in the following. Figure RC2.1 shows an example of time series of raw and processed data in the case of a 8.55 km link operating at 18.03 GHz during the event of 14-16 May 2020. Dry/wet classification at step (2) is required by subsequent steps (3) and (5). First, the RSL is thresholded by an hystheresis method (see Nebuloni et al., 2020b). Then, each CML is given a score, which is the product of the binary outcome of the thresholding by the inverse of its sensitivity to rainfall, the latter depending on CML frequency and path length. Finally, a CML is flagged as wet if the aggregate score of the CML itself and of all its neighbors exceeds 0.5, otherwise it is dry. Two CMLs are neighbors if they fulfill any of the following conditions: (1) they have a terminal in common, (2) their paths intersect, and (3) their distance is within a defined maximum threshold. The baseline on step (3) is obtained through a windowing algorithm. An N-sample window is centered around each sample of the RSL time series. If enough samples in the window are dry, the baseline value in the center of the window is the average of minimum and maximum RSL. Once the entire time series has been processed, the baseline missing points are obtained by linear interpolation. In step (5), it is assumed that wet antenna attenuation is the only relevant component of path attenuation not due to rain. This wet antenna contribution is subtracted using the model proposed by Schleiss and Berne (2010), which predicts an exponential increase of attenuation during the wetting transient, a constant value while raining and an exponential decrease during the drying transient. The input parameters to the model, that are the duration of the initial transient, and the maximum value of wet antenna attenuation are 900 s and 2 dB, respectively. They were determined analyzing a set of RSL and TSL time series sampled every 10 s, which were made available over a few CMLs. The relationship between rain attenuation per unit path length γ_R (dB km^{-1}) and rainfall intensity R (mm h^{-1}) is usually modelled by the following power-law function:

   $$\gamma_R = \kappa R^{\alpha}. \tag{1}$$

   The coefficients κ and α have been tabulated by the International Telecommunication Union as a function of *signal frequency and polarization (ITU-R P.838-2, 2005). In principle, the $\gamma_R - R$ relationship is dependent on the microphysics of rain as well, hence κ and α should be calibrated, provided the characteristics of precipitation are known in the climatic area where CMLs are deployed. In this work, raindrop size distribution data gathered from disdrometers were used to calculate the optimum value of κ and α coefficients following the procedure outlined in Luini et al. (2020). In the available CML data format, only the two extreme values of TSL and RSL are saved in every 15 min window. Therefore, if the average rainfall rate is to be estimated, for instance to calculate hourly accumulations, it is necessary to derive it from the extremes. To this aim, TSL and RSL time series sampled each 10 s were made available for a subset of CMLs during some of the events considered here and processed as shown in Nebuloni et al. (2020a). Average, min and max rainfall rate within 15-min windows were calculated from the 10 s time series and the following unbiased estimator of the average rainfall rate was derived:*

$$R_{\mathrm{MIN-MAX}} = \frac{1}{1.14}\,\frac{R_{\mathrm{MIN}} + R_{\mathrm{MAX}}}{2}. \tag{2}$$

Two aspects of the above procedure deserve more discussion. First, the available RSL (and TSL) has a coarse 1 dB quantization step. That is, the time series of power have a random zero-mean error superimposed, with rectangular distribution and limiting values equal to ± 0.5 dB, that is, ± 12% when the power is measured on a linear scale. It descends that it is impossible to distinguish between rain and quantization-induced noise below a certain rainfall intensity threshold. Figure 3 shows the minimum detectable rainfall intensity without ambiguity as a function of the CML path length with the CML frequency as parameter. The square markers correspond to the 50 CMLs in the study area divided in three groups according to their frequency band. Continuous lines are drawn at four reference frequencies as well. Moreover, quantization affects the accuracy of rainfall intensity estimates. The accuracy of instantaneous measurements (at the 95% confidence level) is within 20% if the rainfall intensity exceeds 3 mm h^{-1} for the link with the most favorable combination between length and frequency. However, in the worst case, the above accuracy is achieved only if the rainfall intensity is above 10 mm h^{-1}. The only way to mitigate quantization effects is to average in time.
Second, it is assumed that rain attenuation measured over a CML of length L, is L times the attenuation per unit path length in Eq. (1), that is rain is considered uniform along the path. The effect of the inhomogeneity of precipitation can be relevant as CML paths range from about 1 km to nearly 9 km (Fig. 3). Some authors proposed to retrieve the spatial distribution of the rainfall field across the measurement area by processing all the CML data together, for instance through tomographic techniques. In this work, a simpler approach is used. Each CML is considered independently of the other and the corresponding rainfall measurement is given a weighting coefficient dependent on CML length, as discussed in Sec. 3.3.
Please note that, apart from the initial calibration of the $\gamma_R - R$ relationship, carried out through disdrometer data, which is made before the above CML processing steps, the CMLs are a fully independent network of rainfall sensors, as no external information is used. In order to validate the rainfall estimates provided by CMLs, let us now compare the accumulated rainfall during each of the events in Table 1 with RG direct measurements. In order to carry out a fully fair comparison, an ad-hoc array of RGs should be deployed along the CML path. This is seldom feasible though. Here, CMLs and RGs are associated according to their mutual distance. Each RG is given a different weight depending on its position with respect to an associated CML as follows: the CML path is divided into short segments, the distance between the RG and each CML segment (approximated by its midway point) is calculated, and all the above distances are averaged. The number coming out of this calculation takes into account the relative position of the CML and of the RG as well as CML length. Finally, the e accumulated rainfall from the set of RGs associated with a given CML is calculated by the IDW method using the average CML–RG distance. The events in Table 1 have been divided into two sets according to their intensity. The scatter plot between CML– and RG–based accumulated rainfall is plotted in Figure RC2.2a and RC2.2b for the eight high-intensity and the four low-intensity events, respectively. Only RGs within 5 km (average distance) from a CML are considered. During high-intensity events, there is a good match between CMLs and RGs estimates, whereas CMLs exhibit an evident underestimate (more than 30% on the average) in the case of low-intensity events. This pattern can be explained by the lack of sensitivity of CMLs to low rainfall intensities due to signal quantization.

2. **Lack of validation of modified IDW method: The authors introduce an adjustment factor for the IDW interpolation method, namely an additional weight that decreases with increasing length of a CML. The reasoning is that the localisation of rainfall information is worse for longer CMLs because of inhomogeneity of rainfall along the path and because, for IDW, the rainfall information is represented by a virtual gauge in the centre of the CML path.**
2a. While this is true, there are also drawbacks to this approach. Longer CMLs have higher sensitivity to rainfall (as shown in this manuscript) and hence can be more accurate at lower rain rates than short CMLs with the same frequency. Furthermore the orientation of the CMLs in relation to their direction towards an HRU centroid should have an effect. A long CML that points directly towards a HRU centroid should have a higher weight there because it observes rainfall closer to this centroid than a perpendicularly

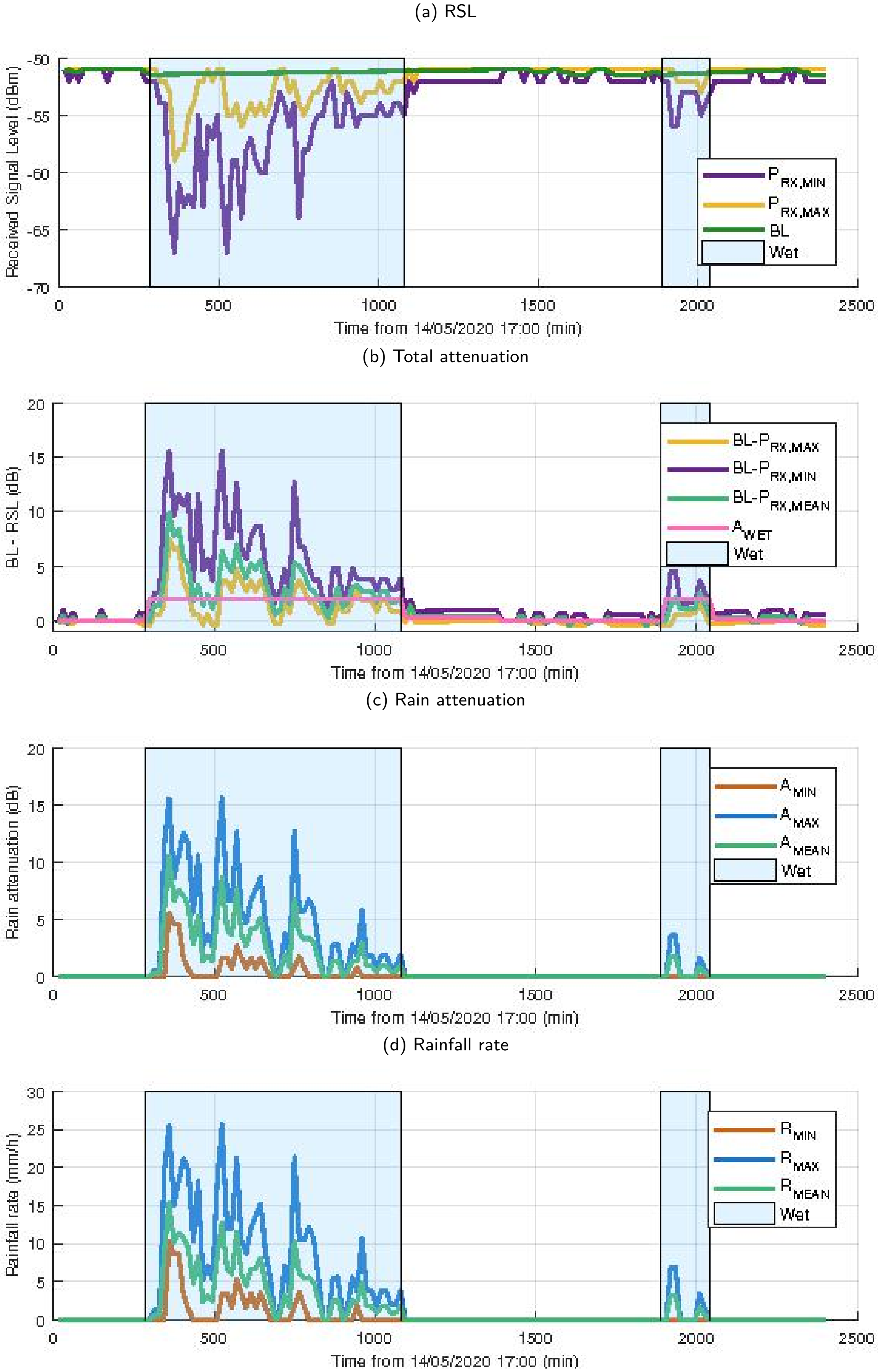

Figure RC2.1: Example of CML data processing: (a) Min and max received signal level (over 15-min slots), wet slots and baseline level, (b) total attenuation and wet antenna component, (c) rain attenuation, (d) rainfall rate.

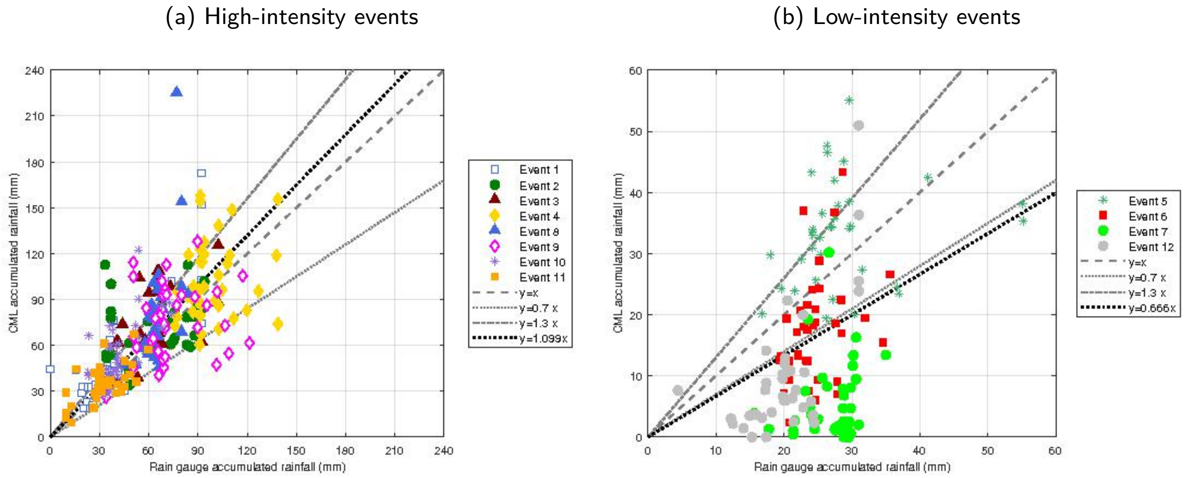

Figure RC2.2: Examples of accumulated rainfall during (a) an high-intensity event and (b) a low-intensity event: CMLs against nearby RGs. The 45° line, ±30% bounds and a best fit of data are shown as well.

oriented CML would. This does not mean that the proposed modified IDW method does not provide an advantage. But this advantage should be shown empirically, at least by providing individual examples of the modified IDW vs the standard IDW.

2b. After thinking a bit more about the modified IDW, I want to add the following comment. If a CML's length is accounted for in IDW by assigning length-dependent weights, shouldn't there be a distance dependent adjustment of this length-dependent weight? The further away a target point (to be interpolated) is from a CML, the less the length of the CMLs plays a role. I can see the reasoning of the length-dependent weight in close vicinity of the HRU centroids because longer CMLs tend to observe rainfall also outside the HRU, but if the CML density inside the HRU is low or zero the length dependent weight might not be optimal. A long close-by CML could be "overruled" by shorter CMLs which are further away. Please comment.

Following your comment, we carried out an extensive comparison between the standard and the modified IDW method, using the entire database of events and the rain gauges as benchmark. Specifically, we estimated the hourly rainfall depth from CML data in each rain gauge location, by the two IDW methods (useing $\gamma = 3$ as exponent for the distance in both cases) as well as the rainfall depth from rain gauge data. As performance metrics, we computed the root mean square error (RMSE) between CML and rain gauge estimates. In particular, we calculated a single RMSE grouping together all the wet hours (where RG rain depths are > 1 mm) of the 12 events for all the locations. In Table RC2.1 we report results of the validation, considering several values of the coefficient β. The standard and the modified IDW methods turn out to give very similar RMSE values. From this, we found that the modified IDW does not provide a significant enhancement to the rainfall estimates respect to the standard IDW. Thus in the revised manuscript, we will report the analyses of rainfall spatial interpolation with the standard IDW. It is worth to notice that new results are quite similar to the original ones. So, there is no need for a new description of results and same conclusions can be given.

Table RC2.1: RMSE values (in mm) between CML-based rainfall estimates and rainfall observations from rain gauges.

Modified IDW			Standard IDW
$\beta = 4$	$\beta = 5$	$\beta = 6$	-
1.70	1.68	1.67	1.65

3. **Overlap of calibration period with analysed rain events: From how I understand the section about model calibration (starting in L233) the 12 analysed rain events are within the calibration period or within the validation period and hence are not independent from the calibration procedure. Since it is not exactly clear how the calibration was performed, I cannot comment on how problematic this is. But the authors should comment on that.**

   In Figure RC2.3 we show the periods used for calibration and validation and where the analysed events are located in time, to make it clearer. Despite there is a partial overlap between events and calibration periods we think that this is not problematic for the aim of the study as we only want to lead a comparison between CML-based and RG-based results. Eventually, this overlap would advantage RG-driven performances, with respect to CML ones and for this we provided a new calibration based on CML inputs. It is also worth to notice that the selected events in 2019 are not those providing the best performances on river discharge simulations but are the most significant ones in the year considering different typologies of precipitation (either convective and stratiform) and for which we asked for CML raw data. Regarding calibration, it was led over the year 2019 by testing different combinations of the two parameters, K_{sat} and z, and selecting the combination which maximized the NSE. As reported in the manuscript, for K_{sat} we tested the values reported in literature (different for each HRU with respect to the type of soil) and multiplied by several powers of 10. Concerning the z we tested all the values inside the range [10 cm, 300 cm], with a 10 cm step. Then, we simply checked that, for the chosen validation period, the performance index exceeded the threshold of 0.5 defined by Moriasi et al., 2007.

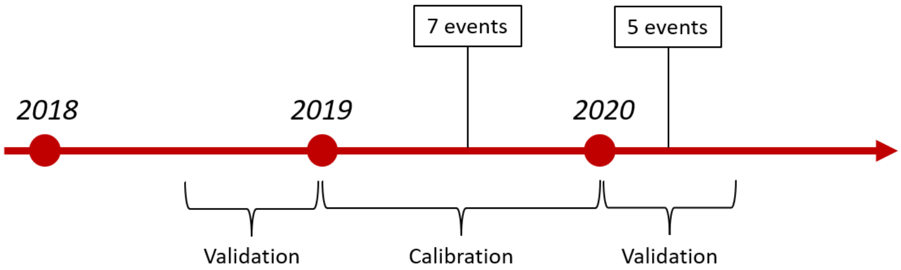

Figure RC2.3: Timeline with calibration and validation periods.

4. **Recalibration of hydrological model only done for CMLs: The idea to recalibrate the hydrological model for the CML-derived input is good and important because it is hard to beat the original forcing which was used for calibration. However, the CML-derived rainfall forcing then has an unfair advantage because its setup is calibrated only to the 12 days used for validation while the original setup was calibrated with a much longer period. Wouldn't it be more fair if there would be a recalibrated setup for the 12 days with the gauge-derived forcing for the analysis shown in Fig 12? One can only speculate if and how much this would improve the performance of the gauge-derived forcing. Please justify or reconsider your approach.**

   We agree with your comment that it would be more fair to make a comparison with the same type of calibration and in the manuscript we will report the new results (Fig. RC2.4), where the RG-driven model has been calibrated based on the same 12 events of the remainder two cases. As results are similar, the description of results will be the same.

Specific comments

- **L13 I find the formulation "may lead to benefit in hydrological modelling" a bit vague for the last sentence of the abstract. It should be made clearer what these benefits (better understanding of hydro processes, better streamflow forecasts, etc) are and why they could be expected from the finding that "CML-driven outputs performances are comparable with RG-driven ones" which only indicates equal performance of CMLs and gauges (which is still an important finding, given the dense rain gauge network)**

  Our results showed that the exploitation of CML-based rainfall data into a hydrological model may be useful when no or few rain gauges are present over the case study area, as we observed that performances are comparable to those gained using a dense traditional rainfall monitoring network. We hence thought to replace the discussed sentence with the following one:
  ...confirming that the exploitation of a CML network may be a great support to hydrological modelling in those area lacking of a well designed and dense traditional monitoring system.

- **L23 I do not understand the second half of the sentence starting at "..., even if..."**

  We agree with you, here "even if" is improperly used. A more correct formulation of the sentence could be:
  ..., however it may happen that some economic and geographical circumstances prevent the presence of an adequate density of rainfall sensors.

- **L32 It should be noted here that the temporal aggregation time of the rainfall data is also an important factor. The longer the aggregation time the less critical the rain gauge density is.**

  Thank you for having pointed this out, we will add this sentence in line 33: *...In particular, the shorter is the aggregation time the more critical is the rain gauge density.*

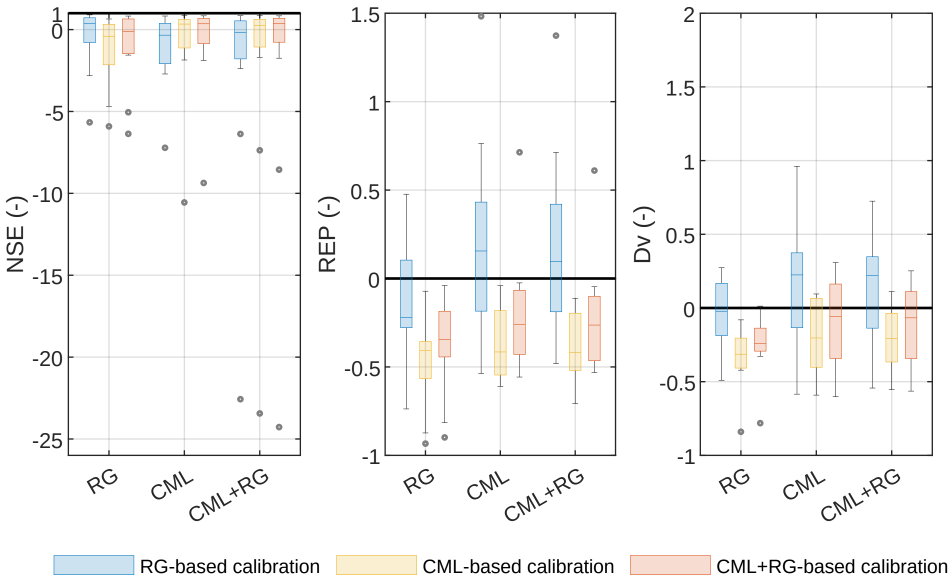

Figure RC2.4: Box plots of the performance indices for the 12 flood events obtained with three different calibration sets (drawn in as many colours) and with three different input types (on the x-axis).

- **L35 and following: This short review of Radar QPE could mention some publications on recent progress with dual-pol radars, e.g. Zhang et al 2020 (https://doi.org/10.1175/JHMD-19-0194.1), Chen et al 2021 (https://doi.org/10.1175/JHM-D-20-0299.1) or this overview from Zhang et al 2019 (https://doi.org/10.1007/s00376-019-8172-4).**

  Thank you, we will add these recent references in the revised manuscript:

  Recently, the dual-polarization upgrade on radars (Zhang et al. 2019, Chen et al., 2021) brought considerable improvements as further information about shape, composition, and phase of hydrometeors can be provided. Hence, the quantitative precipitation estimation (QPE) could greatly benefit from such advancements.

- **L46: The usage of microwave links for rainfall estimation was proposed earlier, e.g. by Atlas and Ulbrich 1977 (https://doi.org/10.1175/1520-0450(1977)016%3C1322:PAAIRM%3E2.0.CO;2). Giuli et al. maybe where the first to propose a mesh of link for tomographic reconstruction.**

  We will add the reference changing lines 45–48 as reported below:

  The use of microwave links as opportunistic rainfall detectors was firstly proposed by Atlas and Ulbrich (1977). The method exploits the relationship between the rainfall intensity and the attenuation (i.e., the loss of power) experienced by the electromagnetic wave along the propagation path from the transmitter to the receiver. Later, Giuli et al. (1991) made use of a mesh of microwave links for the 2D reconstruction of the rainfall field.

- **Fig 2 and related text: How were the CML paths treated in the calculation if a CMLs crosses several HRUs and how was the distance to the HRU centroid calculated for CMLs? In general, how was the mean distance of the sensors to the HRU centroid calculated? Are all sensors taken into account or only the closest ones?**

  The CML-HRU distance was calculated between the CML middle point and the HRU centroid. The mean distance, related to each HRU, was then calculated only considering those CMLs which middle point have a distance from the centroid not higher than 10 km. We would like to add this clarification in the manuscript changing the text from line L112 as here reported:

  a different number of sensors was exploited, according to a defined maximum distance of 10 km from the HRU centroid. Figure 2 shows some features of the rainfall sensors used for spatial interpolation in each HRU: the number of exploited RGs, CMLs, and their sum, the ratio between CMLs and RGs number, the mean distance between rainfall sensors and HRU centroids, and the mean length of CMLs. Please, note that the CML-HRU centroid distance is calculated considering the CML middle point.

- **Table 1: Would be good to specify what the averaging period for the max. rain rates are because only then one can interpret them. I assume these are 10-minute maxima from the raw data, but it should be stated here**

  They are 1 hour maximum rain rates. We will add a clarification in the revised manuscript. We chose such period in order to classify events taking inspiration by the classification reported in Met Office, 2007.

  Met Office, August, 2007. Fact Sheet No. 3: Water in the Atmosphere. Crown Copyright. p.6. Archived from the original on 2012-01-14. Retrieved 2011-05-12.

- **L145: Why did you chose to fill missing RG values with spatially interpolated data? Why not leave the gaps and treat missing values accordingly when the rainfall at the HRU centroid is derived?**

  This choice comes from the fact that we intended to rely on the same number of rain gauges for the rainfall spatial interpolation over the entire selected period.

- **L152: The frequency separation of CMLs is small, typically around 1 GHz for CMLs operating below 40 GHz. I do not agree with the statement that this is "adding a certain degree of redundancy when it comes to rainfall estimates". Or maybe I do not understand what is meant here. Please explain.**

  Our statement was not very clear indeed. The redundancy here helps us in the case a channel is not available (it happens) or if there have missing or invalid data somewhere. As for rainfall estimates, we averaged (attenuation) among the available channels. We modified the unclear statement in the manuscript. Please see our answer to "Major comment" n. 1.

- **L155: Nebuloni et al., 2020b does not provide much detail of the proposed processing method and in particular no quantitative analysis of the resulting CML rain rates, e.g. scatter plots, Pearson correlation, RMSE, bias, false positive rate, etc.. Because the processing of the CML data can have a large impact on the derived QPE it would be important to provide more information on the CML-derived rainfall data here. The choice not the use the available RAINLINK methods should also be explained.**

  We followed the procedure outlined in one of the papers published by the research group who is behind the RAINLINK methods. As for rainfall estimates, in the original paper we showed the values calculated at the HRU centroids, which were obtained by aggregating data from different CMLs according to an IDW method. Based on your comments we added more details about CML processing steps and we added a comparison of rainfall estimates gathered by CMLs and RGs. Please see our answer to your "Major comment" n. 1.

- **L159: Nebuloni et al. 2020b explains that no wet antenna compensation was applied (end of section 3). What was done for the analysis in this manuscript?**

  We will add the explanation about the analysis for wet antenna compensation (which is not referred to in Nebuloni, 2020b) in the subsection **Commercial Microwave Links** as reported in the answer to "Major comment" n. 1.

- **L162: I do not understand why the RSL time series is considered in Watt here to point out to the +-12% uncertainty range in relation to the measured value. All relevant calculations to derive rain rates from raw RSL data are carried out in dB and the uncertainty in dB directly translates to the uncertainty in rain rate.**

  Yes, rainfall rate is related to the attenuation expressed in dB units. We just pointed out that a 1 dB uncertainty on attenuation is equivalent to a 12% uncertainty when we consider quantities measured on a linear scale. It is a message intended for the readers not very familiar with dB units. We changed the statement. Please see our answer to "Major comment" n. 1

- **L167: Can you explain how these numbers were calculated?**

  We moved from the uniform probability density function (PDF) of the quantization error on the RSL. We assumed, for simplicity, that the same distribution of the quantization error holds for rain attenuation. Then, by virtue of the power-law relationship between attenuation and rainfall rate, we computed the PDF of the quantization error on the rainfall rate and the associated 95% confidence level. For every frequency-path length pair in our CML network, the calculation is repeated over a set of nominal values of attenuation (i.e. with no quantization error), i.e. for a corresponding set of nominal values of the rainfall rate.

- **L178: How much did the derived k-R parameters differ from the ITU recommendations? And how relevant is this difference in comparison to the uncertainty from quantization and wet antenna effect?**

  About the estimation of k and α coefficients, we performed a calibration based on the knowledge of the drop size distribution (DSD) of rainfall measures by 3 disdrometers in Valmalenco, a small catchment located in North Italy. We retrieved the optimal coefficients by regressing the specific attenuation calculated from DSD data against the rain rate calculated from DSD and raindrop velocity data again provided by the disdrometers. Calculations were repeated over the CML frequencies. For a given frequency, the differences among best-fits from different disdrometers and from different events are small. Hence, it is reasonable to calculate a single set of (α, k) parameters. For instance, at 18.80 GHz, we obtained $\alpha = 1.1095$ and $k = 0.0535$ against the following ITU-R values: $\alpha = 0.9948$ and $k = 0.0845$. Figure RC2.5 shows: 1) three best fit curves (DIS1, DIS2, DIS3) obtained considering all the events for each disdrometer, 2) a best fit curve (ALL DIS) considering all disdrometer data and all events, and 3) the ITU-R coefficients. In Table RC2.2 we report, for $\gamma_R = 10^{-1}, 10^0, 10^1, 10^2$ dB km^{-1} the relative difference, ΔR, on rain

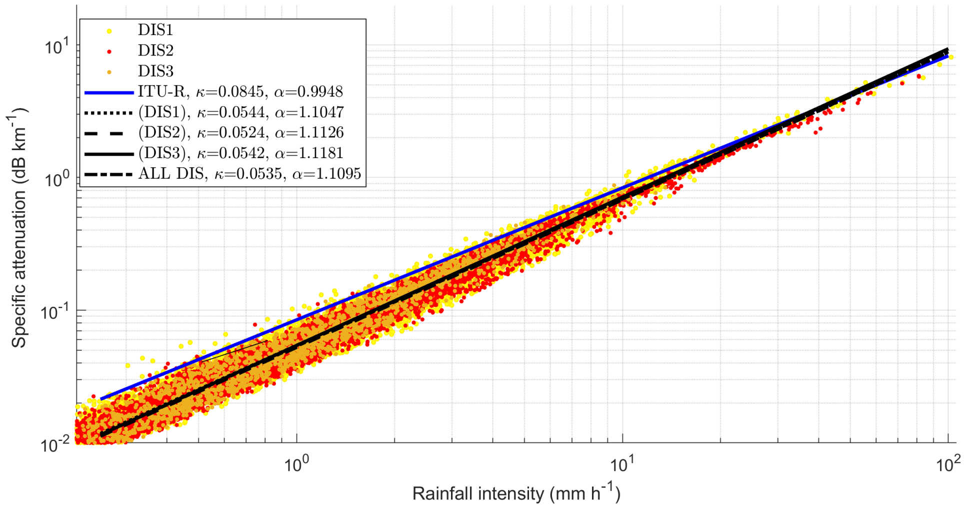

Figure RC2.5: Regression lines (in log-log scale) for k and α coefficients estimation.

estimation (R) calculated as $\Delta R = 100 \cdot (R_{cal} - R_{ITU-R})/R_{ITU-R}$, where R_{cal} is the rain intensity estimated from calibrated coefficients k and α and R_{ITU-R} is the rain intensity estimated from ITU-R parameters.

Table RC2.2: Comparison between rainfall estimates relying on calibrated and ITU-R coefficients.

γ_R (dB km^{-1})	10^{-1}	10^0	10^1	10^2
R_{cal} (mm h^{-1})	1.76	14.00	111.55	888.70
R_{ITU-R} (mm h^{-1})	1.18	11.99	121.33	1228.02
ΔR (%)	48.36	16.79	-8.07	-27.63

Concerning the uncertainty on rain intensity estimation from quantization, it is reported in line 167. On the other hand, we cannot make statements on the uncertainty due to the wet antenna effect as we should make a comparison with the real attenuation given by rain, which we do not have.

- **L181: It is not clear if and how the 10-second data was used in this study and why it was required to estimate the average rainfall from the min-max data. Several studies from Israel and the Netherlands have done rainfall estimation from 15-minute min-max data, e.g. by using a calibrated weighting factor of the min and max values. Please elaborate on this.**

  10-second data were used for calibration of the weighting factor of min and max rainfall rate to derive the average rainfall rate in every 15-min window, as described in Nebuloni et al. (2020a). 10-second data were also used to estimate the parameters of the wet antenna attenuation model. This information has been added to our answer to your major point 1.

- **L187: Why was this simpler approach chosen? The fairly dense CMLs network would presumably provide a good basis for algorithms that derive uneven distributes of rainfall along the CMLs by taking into account measurements from nearby CMLs. I am not saying that simple approaches provide inferior results, but it would be interesting to know the reasoning behind this choice.**

  We decided to use a simple approach since we are dealing with a semi-distributed hydrological model which is not as sensible as a distributed one to get the spatial variation of the rainfall field. A tomographic algortihm, for example, would provide a detailed reconstruction of the rainfall field but in our case it would not be very useful as we should re-aggregate the estimates at the HRU scale calculating an average value.

- **L189: I cannot comment on the scientific soundness of the hydrological model, but I would like to understand the choice of this particular model. Was this model applied already in the region, is it applied frequently by the authors or regional water authorities, is it developed by the authors or collaborators?**

This model has been developed by authors as no semi-distributed model was developed in this area. However a further challenge, that we are working on, is the exploitation of an already existing distributed model, the FEST-EWB model (Ravazzani, et al. 2008). To this aim, we intend to rely on advanced methods for reconstruction of the 2D rainfall field, as the tomographic algorithm.

Ravazzani G, Rabuffetti D, Corbari C, Mancini M. 2008. Validation of FEST-WB, a continuous water balance distributed model for flood simulation. Proceedings of XXXI Italian Hydraulic and Hydraulic Construction Symposium. Perugia

- **L232 and following: I am surprised that many of the 12 selected rain events are within the calibration period, namely the year 2019. The other selected rain events are within the validation period of the parameter calibration in the year 2020. Doesn't that carry the risk of overfitting the model on these rain events, assuming that the selected ones are the most prominent ones in these years?**

  Please, refer to the answer to your "Major comment" n. 3.

- **L242: What does "trial and error calibration" mean here? From the sentences above I thought I understood that all the listed parameters and their combinations have been used to run the model. What other "optimum combination" is there that is not covered by sweeping over the defined parameter range?**

  The term "trial and error calibration" is misleading indeed. In fact, as reported above, we only tested the listed parameters and selected the set providing the best NSE. Hence, we will disregard that sentence.

- **L245: Was there any information available for how the outflow of the dam was regulated? If not, and if the model is not able to account for that, how much sense does it makes to study events with high flow which might be severely affected by the dam outflow?**

  Despite we did not directly report on the manuscript, we did take into account the dam outflow modelling. As Figure 1 in the manuscript shows the HRU involved is the 6, where the lake is located in the southernmost part and its outflow enters in HRU 7. We modeled the area not covered by the lake as an HRU in series with the lake, so that the output of the HRU is the input of the lake. Then we modeled the dam outflow, relying on the reservoir curves and the abacus of dam outflows with respect to (1) the hydrometric level of the lake and (2) the opening of the gate-dam (in Figure RC2.6). This information was provided to us by the Cavo Diotti manager and Parco Regionale della Valle del Lambro.

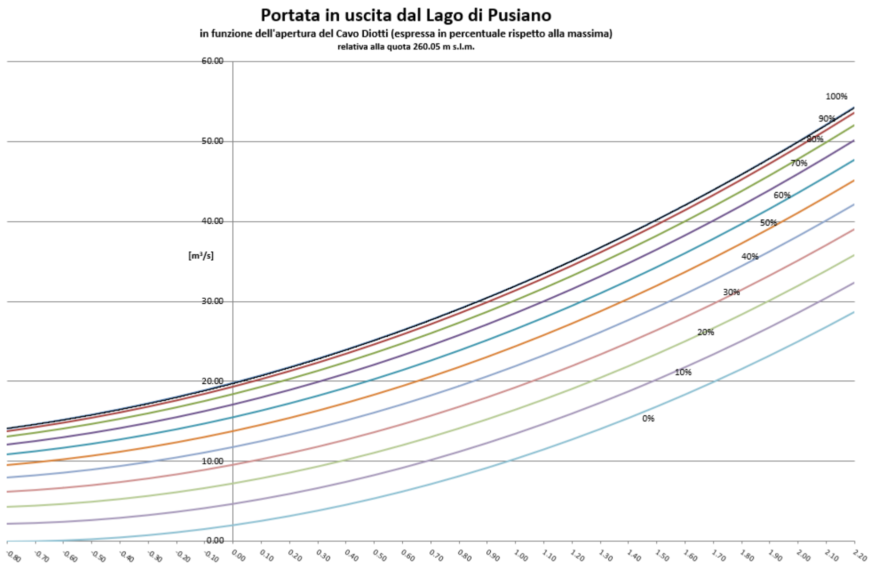

Figure RC2.6: Dam outflow in m^3 s^{-1} (y-axis) with respect to hydrometric level in m above the hydrometric 0 (x-axis) and the percentage of gate-dam opening (different colors of curves).

- **L272: How was the selection of the parameter beta done?**

  As reported in the answer of "Major comment" n. 2, we disregarded the use of the modified IDW method.

- **L287: How are "low rain rates" and high ones defined?**

  To define the threshold we firstly referred to the classification proposed by Met Office, 2007, which is reported in Table RC2.3. Then we selected the threshold of 15 mm h^{-1}, as it was close to the central threshold (10 mm h^{-1}) reported in Met Office and within the largest gap between consecutive maximum rain rates (excluding the gap

Table RC2.3: Rain intensity classification.

Term	Criteria (II)
Light rain	$<$ 2.5 mm per hour
Moderate rain	2.5 - 10 mm per hour
Heavy rain	10 - 50 mm per hour
Violent rain	$>$ 50 mm per hour

between the two highest values). The highest value included in category *Low rain rate* is 12.6 mm h^{-1}, which is really close to 10 mm h^{-1}, the central threshold from Met Office, 2007.

Met Office, August, 2007. Fact Sheet No. 3: Water in the Atmosphere. Crown Copyright. p.6. Archived from the original on 2012-01-14. Retrieved 2011-05-12.

- **L288: Since HRU 8, 2, 9 and 4 do not show an overestimation of CML rainfall compared to the gauge-derived data, I do not agree with the conclusion that "CMLs tend to return higher estimates... during high rain rate events...".**

  We agree with you and we will disregard such sentence.

- **Fig 5: It is not clear to me why there is only one marker type, i.e. either high or low rain rates per HRU. I expected that the event rainfall accumulation is done two times for each HRU, once for events (or maybe even hours) where a certain threshold rainfall rate is exceeded and once for the events (hours) where it is not exceeded. Please clarify how the split into high and low rain rates was done. In case more data points will be added, the plot could be split up into two subplots, one for high and one for low rainfall rates.**

  The accumulation values are not reported for each HRU, rather they are areal-averaged over the entire basin area. Each yellow marker, as well as each orange one, represents the rainfall accumulation calculated for a single event. The number next to each marker refers to the event ID, reported in the manuscript in Table 1, and not to the HRU. In fact, there are 12 markers (as the events are 12) for each color. The circles correspond to those events classified as *High rain rate* while the squares refers to *Low rain rate* events. To provide a more explanatory description of the Figure we would like to modify line 285 as follows: *Figure 5 shows the scatter plot of the rainfall depths accumulated at the end of each storm event and averaged over the entire catchment area.*

  In addition, we would also modify the caption related to Figure 5: *Rain depths, averaged over the catchment area and accumulated at the end of each event. The number next to the markers refer to the ID event. The two different markers, circles and squares, respectively stand for High rain rate and Low rain rate events. The black line represents the 1:1 line of perfect matching between rain depth estimates from RGs and from CMLs (yellow) or from the combination of RGs and CMLs (orange). Yellow and orange lines are the corresponding regression lines.*

- **Fig 6: I do not understand (but maybe that is my fault) why there are no negative values smaller than -1. If the CML rainfall at a HRU centroid yields 0 because of a false negatives, then ΔE should be -1 times the RG value, shouldn't it?**

  Δ_E is defined as:

  $$\Delta E = \frac{R_{CML} - R_{RG}}{R_{RG}}. \tag{3}$$

  where R_{CML} is the 1 hour rain depth from CMLs and R_{RG} is the 1 hour rain depth from RGs. We considered only rainfall data for which $R_{RG} \geq 1$ mm. Hence, when the CML estimates R_{CML} yields 0 because of a false negative, the formula reduces to $\Delta E = -\frac{R_{RG}}{R_{RG}} = -1$. We will add the equation of ΔE in the text so to make it clearer.

- **Fig 6: Maybe a 2D histogram, e.g. a hexbin plot, would be easier to interpret than the current figure. The distribution of points can also be conveyed in a 2D histogram and loglog scales can also be used.**

  In Fig. RC2.7 you can find the new 2D histogram, as suggested by you, which will replace Figure 6 in the revised manuscript.

- **Fig 6: Besides the negative values of the mean of ΔE for small rain rates, I find it worth mentioning that the spread of ΔE in the positive range increases significantly for small rain rates. Is this the effect of false-positive CML rain rates?**

  The increasing spread of ΔE values with respect to the decrease of the hourly rain depth is due to the greater uncertainty of CMLs in detecting low rain rates. However, we can observe that for the lowest rain depths there are fewer negative values of ΔE as we set to zero all the CML rain rate estimates lower than the sensitivity of the link itself.

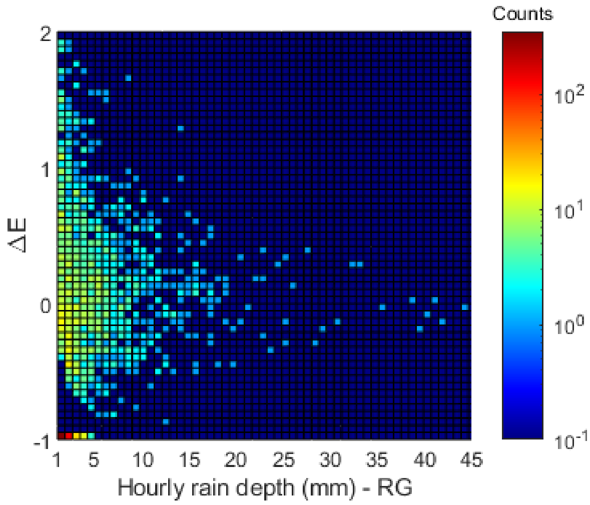

Figure RC2.7: 2D histogram of hourly rain depths and ΔE. The colour of each equally spaced 2D bin represents its height, which is the count of values falling in the bin. Note that the scale bar has a logarithmic scale and the dark blue bins have 0 counts. Values of ΔE equal to -1 represent false negative hours.

- **Fig 8: Since the distinction between high and low rain rates is made in most other plots, it would also be valuable to show it here, e.g. by having two differently coloured box plots for each HRU.**

  We followed your suggestion and prepared Fig. RC2.8, which will replace Figure 8 in the revised manuscript. As expected, the boxplots related to low rain rates are those with the major dispersion and have mostly negative median values.

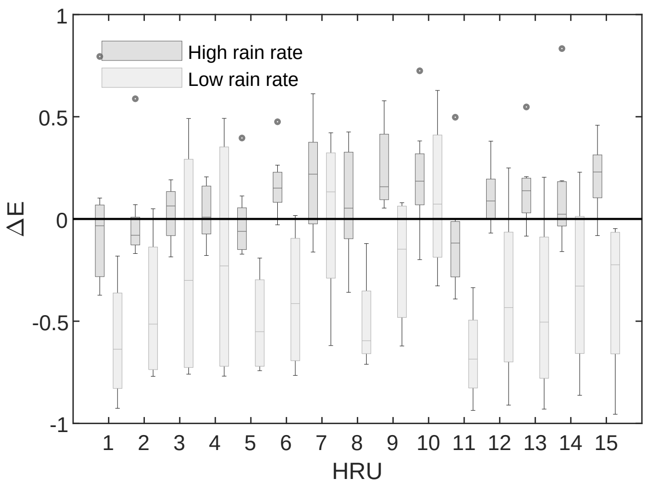

Figure RC2.8: Box plots of ΔE for the 12 storm events grouped by HRU and by event intensity.

- **L315: It would be good to show these direct comparisons of CMLs and closest rain gauge so that the reader can judge himself or herself how the different CMLs behave and perform. Maybe these could be added to the Appendix. See also my main comments.**

  Please, refer to the answer to "Major comment" n.1.

- **L341: It would be better to write "...Fig 11b shows an example for which the CML-driven simulation better represents..." because otherwise this reads like a general statement that CML-driven simulations**

perform better.

We will fix this issue directly on the manuscript.

- **L351: "the major drawback of the present work is definitely that we did not rely on a large and real-time CML-based dataset" Why didn't you use longer periods of CML data?**

  Unluckly, we could ask the phone company for a limited number of hours of power data. So we could only select some specific rainy events.

- **L355: To have a fair competition between RG and CML, you would have to recalibrate the RG-driven setup also only using the 12 selected events. By calibrating the CML-driven simulation exactly to the event that you analyse you might given them a significant advantage. (See also my major comment)**

  Please, refer to our answer to "Major comment" n.4.

- **L364 Discussion section: This section reads more like a summary of the results, not like a discussion. Potential causes and consequences of the results, as well as limitations of the chosen approach should be discussed here. I suggest to add some subsections to structure that.**

  We created a new subsection entitled "Limitations and improvements", to be included in the Discussion section. Here we report 1) part of the conclusions that mainly focus on the limitations of our work and on which it could be enhanced in future and 2) a new part on the benefit that such work could give in modelling the Areal Reduction Factor (ARF), as suggested by CC1 (https://doi.org/10.5194/hess-2021-389-CC1). Here below we report the new subsection:

  Limitations and improvements

  One of the major difficulties encountered during analyses was the small amount of CML data, as we relied on only 458 hours of CML raw data grouped in 12 events. On the other hand, the database from rain gauge observations was much more wider and we disposed of real-time data. This led to a rather unfair competition between the two of them. An extension of the CML-based dataset of events, or better yet, to have access to real-time CML raw data would definitely bring great benefits to the present work. Firstly, it would allow the development of a more robust statistical analysis on storm/flood events. Secondly, it would enable a proper calibration, and a validation as well, of the hydrological model based on CML data as rainfall input.

  To enhance this work, it would also be useful to resort to the implementation of a CML-driven distributed model, which is expected to provide a more accurate description of the spatial variability of the precipitation field with respect to a semi-distributed one. In such a case, the CML measurements would be better exploited by the use of advanced methods for spatial reconstruction of the rainfall field. For instance, techniques as the tomographic reconstruction algorithm (D'Amico et al., 2016) or the stochastic reconstruction based on copulas (Haese et al., 2017; Salvadori et al., 2007), take advantage of the path integrated nature of CML measurements.

  It is also worth to notice that, although we showed that CML rainfall data can be successfully assimilated into hydrological models, their integration into real-time operational platforms (e.g. early warning systems) remains challenging. A number of aspects should be still considered including:

  - *generation of CML raw data formats suitable for rainfall estimation;*
  - *real-time collection of raw data, which should be transparent to network operation;*
  - *data transfer to a control centre;*
  - *data reduction process, especially if a large number of CMLs are managed.*

  The above mentioned issues suggest a systematic cooperation with mobile operators, who are the owners of CML network infrastructure.

  Up to now, we mainly focused on the exploitation of CML-based rainfall estimates with the purpose to test their impact on the hydrological simulations of river discharge, with respect to the use of RG data. However, still important hydrological issues could be addressed by dealing with CML data. One of this is definitely the modelling of Areal Reduction Factor (ARF), which represents the factor transforming a point rainfall, for a given duration and return period, into the areal average value, for the same duration and return period (Natural Environmental Research Council (NERC), 1975). In last decades, great efforts have been put for the modelling of the ARF, useful in the design of hydraulic and hydrologic infrastructures, for flood risk evaluations, and rainfall threshold estimations in early warning systems (e.g., De Michele et al. 2001, Kim et al., 2019, Biondi et al., 2021). As we dealt with a semi-distributed hydrological model, we needed to transform point (from RG) and linear (from CML) precipitation measurements into areal values, over the HRU areas. Therefore, from a different perspective, this work could be also seen as a first step in order to test the modelling of ARF by using a combination of conventional and unconventional sensors.

- **L423 and following: Regarding the first three points mentioned here it shall be noted that continuously operating CML real-time data collection and processing systems exist or are in a final implementation phase in Sweden (https://www.smhi.se/en/services/professionalservices/memo-microwave-based-environmental-monitoring/), Czech Republic (http://www.tel4rain.cz/) and Germany (https://amt.copernicus.org/articles/9/991/2016/). Furthermore I do not agree with point 4. Why should "heavy data reduction" be require? Data storage, transfer and processing of TBs of data is not a problem with today's computer resources, in particular at met services or research institutions.**

  We are still in contact with the company who designed and currently manages the network monitoring tool, which generates the CML data used in this work. According to them, it would be challenging to generate, transfer and process data in real-time for hundreds (or thousands) of CMLs because a) this operation would demand network resources that are allocated for user traffic, and b) an upgrade of the firmware would be necessary, which could be done only upon agreement with the network operator. Moreover, a CML-based rainfall monitoring system would be really beneficial where conventional rainfall sensors are not in place, that is in developing countries. Here available computational power and, in general, TELECOM infrastructure are probably not as powerful as in countries like Germany. We removed the word "heavy" from the last point because it may suggest that the critical step is this one.

Technical corrections

Thank you for being so accurate in reading the text, we will definitely fix in the text all the errors you pointed out.