Dear Reviewer#1,

we thank you for your comments and the valuable suggestions on our manuscript. You will find below the answers to your comments. Concerning the location of figures and tables, we rearranged them close to the text where they are mentioned, following your suggestion. Finally, we appreciate your accuracy in reading the manuscript; we fixed all the spelling and grammar corrections you pointed out directly on the revised version of the manuscript.

- **Figure 1: Case study area. Panel (a) shows the Lambro catchment, the partitioning in 15 sub-basins (HRUs), and the position of the sensors, while panel (b) reports the scheme of HRU interaction. I suggest adding "in a network". Also, referring to Fig. 1, see line 230, where this image is first referred to – above 100 lines of text.**

  Thank you, we added "in the network". Moreover, we may split Figure 1 into two figures and move the scheme of HRU interaction close to line 230.

- **Figure 3: There are no 30 Gz data mentioned in the text [see lines 106 - 108] nor shown on the figure please move this Figure 3.**

  Our CML network did not have links with frequency around 30 GHz. We will disregard the corresponding curve in the figure.

- **288:... On the other side, CMLs tend to return higher estimates than RGs during High rain rate events (circles), even though the trend is not as evident.**
  Comment: the 4 rain-rates are very nearly the same ...

  We agree with you, this comment is improper. We disregarded such sentence.

- **301: We therefore focused on the CML hourly wet-dry (see Sec. 2)**
  'please see Section 2' - but note that it is 170 lines above (I eventually found it at line129!) where it should read: 'An hour is considered dry when the detected rainfall depth is lower than 1 mm and wet otherwise.

  As you suggest, it could be useful to recall the definition of wet-dry hours here, adding the following sentence in line 302:

  We recall that we defined as dry those hours in which the detected rainfall depth is lower than 1 mm and viceversa for wet hours.

- **305 . . . ., whereas the occurrence of a false positive is relatively rare in both cases.**
  Comment: then you should junk data which are false negative - by the way, what is the proportion of false negative?

  We did not junk false negatives as to make it possible to understand which is the impact of those false negatives on the performances of the hydrological model, especially during *Low rain rate* events. Moreover, if we disregarded false negatives we should replace with RG data, making the two sets of data not independent and therefore their comparison would be unfair.

  The box plots are referred to the proportion of hours identified as false negative, as well as false positive, with respect to the total amount of hours in the events. The two box plots on the left report the distribution of the percentages related to the 8 *High rain rate* events for each HRU (8 events x 15 HRUs = 120 values) while the two box plots on the right refer to the percentages of the 4 *Low rain rate* events, again in each HRU (4 events x 15 HRUs = 60 values). For example, the maximum observed percentage of false negative is 60%, which refers to event 7 (*Low rain rate* event) and HRU 2.

  To make it clearer, we modified the paragraph from line 301 as follows:
  We therefore focused on the CML hourly wet-dry classification, inferred in HRU centroids, again considering RG estimates as benchmark. We recall that dry hours are those in which the detected rainfall depth is lower than 1 mm and vice versa for wet hours (see also Sec. 2). Figure 8 depicts box plots of the percentage of false negatives and false positives for low rain rate and high rain rate events. In contrast to a false negative, a false positive occurs when an hourly slot is classified as wet by CMLs and dry by RGs. The two box plots on the left were obtained from a population of 120 samples (8 high rain rate events x 15 HRUs), while the two on the right were computed from 60 samples (4 low rain rate events x 15 HRUs). For example, the maximum percentage of false negatives is 60%, which corresponds to HRU 2 during the low rain rate event 7 of Table 1. From a general point of view low rain rate events exhibit a higher median and a larger dispersion of false negatives than high rain rate events, whereas the occurrence of a false positive is relatively rare in both cases. These results confirm the inability of CMLs in detecting low rain rates, which depends on the quantization error issue discussed in Sec. 3.

- **Figure 6. Relative difference ΔE between CML and RG hourly rain depths against RG rain depths. X-axis has a logarithmic scale.**
  - **What is the lowest rain depth? I guess 1... Also, what is the meaning of ΔE being -1.**
  - **You should make your figure captions more informative. Also, it's difficult for the reader the way the text is presented relative to the figures, as this mismatch causes the reader to hunt desperately for the linkage – please fix this irritation. I have got around the problem by splitting the screen of the pdf - but it's a pain!**

Yes, the lowest rain depth is 1 mm as we set to 0 all the values lower than 1 mm.

Moreover, ΔE is defined as:

$$\Delta E = \frac{R_{CML} - R_{RG}}{R_{RG}}. \tag{1}$$

So, when the CML estimate, R_{CML}, is 0, the formula reduces to $\Delta E = -\frac{R_{RG}}{R_{RG}} = -1$. We hence obtain $\Delta E = -1$ in the case of false negative, and for this we observe in Figure 6 a high density of values corresponding to $\Delta E = -1$. We explicity added all these clarifications in the revised manuscript and in the caption.

The paragraph from line 290 has been modified as follows:
We further assessed CML and RG rainfall estimates on the hourly time scale and on the sub-basin spatial scale by calculating the relative error of CML estimates with respect to RG ones, assuming the latter as benchmark, for the hourly rainfall depths inferred in the 15 HRU centroids. The relative error (ΔE) is evaluated as:

$$\Delta E = \frac{R_{CML} - R_{RG}}{R_{RG}}, \tag{2}$$

where R_{CML} and R_{RG} are the 1 hour rainfall depths estimated in each HRU centroid, respectively from CMLs and RGs. For the calculation, we only considered wet hours ($R_{RG} \geq 1$ mm in HRU centroids), relying on a dataset of 2061 values. Hence, when the CML estimate yields 0 and the RG estimate is greater than 0 (false negative), the relative error is -1. Figure RC1.1 shows a 2D histogram representing the count of rain hours falling in a given range of RG-estimated rainfall depths and in a given range of relative errors. The increasing spread of ΔE values with respect to the decrease of the RG-based rainfall depths is due to the greater uncertainty of CMLs in detecting low rain rates. If the RG-based rainfall depth is smaller than 3 mm, only 30% of ΔE values falls in the range [-0.4, 0.4], whereas if it is larger than 3 mm, the percentage increases up to nearly 70%. Moreover, for the lowest rainfall 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. The high count related to $\Delta E = -1$ and RG-based rainfall depths < 5mm is due to the occurrence of false negatives.

Please, note that we replaced Figure 6 in the old version with Figure RC1.1, as suggested by Reviewer 2, to make it clearer.

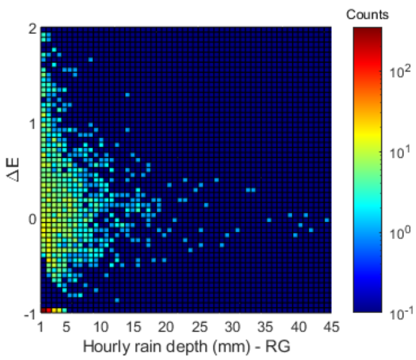

Figure RC1.1: 2D histogram of hourly rain depths and ΔE. The colour of each equally spaced 2D bin represents its height, which is the count of data falling in the bin. Note that the scale bar has a logarithmic scale and the dark blue bins correspond to 0 counts. Values of ΔE equal to -1 represent false negative hours.

- **Figure 8. Box plots of ΔE for the 12 storm events grouped by HRU. What is the range of the box plots? the usual ? Minimum (Q0or 0th percentile): the lowest data point excluding any outliers.**
  Maximum (Q4or 100th percentile): the largest data point excluding any outliers. Median (Q2or 50th percentile): the middle value of the dataset.
  First quartile (Q1or 25th percentile): also known as the lower quartile qn(0.25), is the median of the lower half of the dataset.
  Third quartile (Q3or 75th percentile): also known as the upper quartile qn(0.75), is the median of the upper half of the dataset.

  Yes, the ranges of the box plots are those reported by you. We clarified it in all the figures' captions were box plots are present.

- **6 Conclusions**
  In the conclusion, as well as I suggested in the introduction, please reintroduce the full meaning of the acronyms as many readers might skip, via figures to the conclusion...

  We agree with your suggestion, and we reintroduced the meaning of the acronyms.

- **403: The hydrographs simulated by the hydrological model highlight better performances in terms of NSE and Dv.**
  Insert before Dv: 'the relative error on flow volume,'

  We fixed it directly in the revised manuscript.

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 modified the subsection **Commercial Microwave Links**, from line 150, adding 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 (Fig. RC2.1). The new version of the subsection also integrates clarifications of your comments from **L152** to **L187** which will be individually addressed even later.

   We also added an Appendix where is reported a local comparison on precipitation time series, between some selected CMLs and their nearby RGs.

   Eventually, Fig. RC2.2 (not reported in the manuscript) 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.

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.**

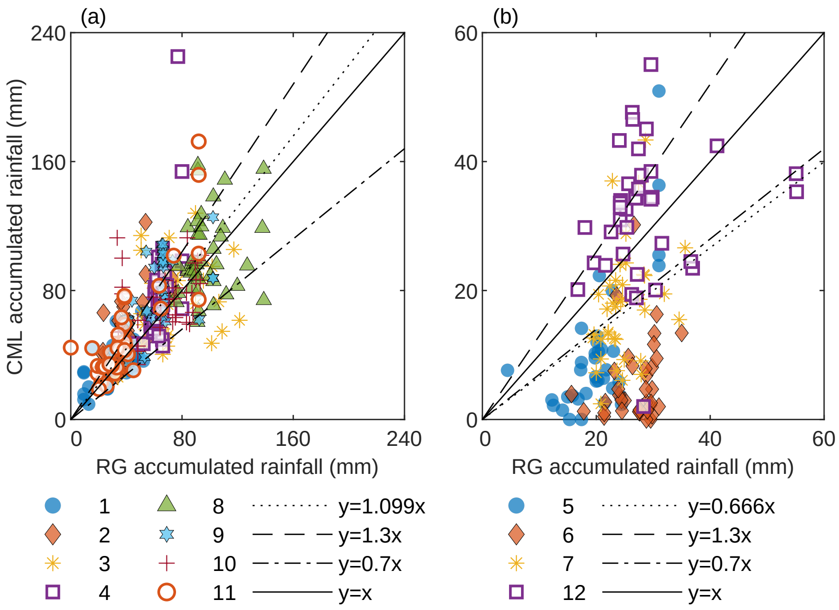

Figure RC2.1: Accumulated rainfall during (a) *high rain rate* events and (b) *low rain rate* events: CMLs against nearby RGs. The best fit of data, the $\pm 30\%$ bounds, and the 45° line are shown as well.

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 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 reported 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**

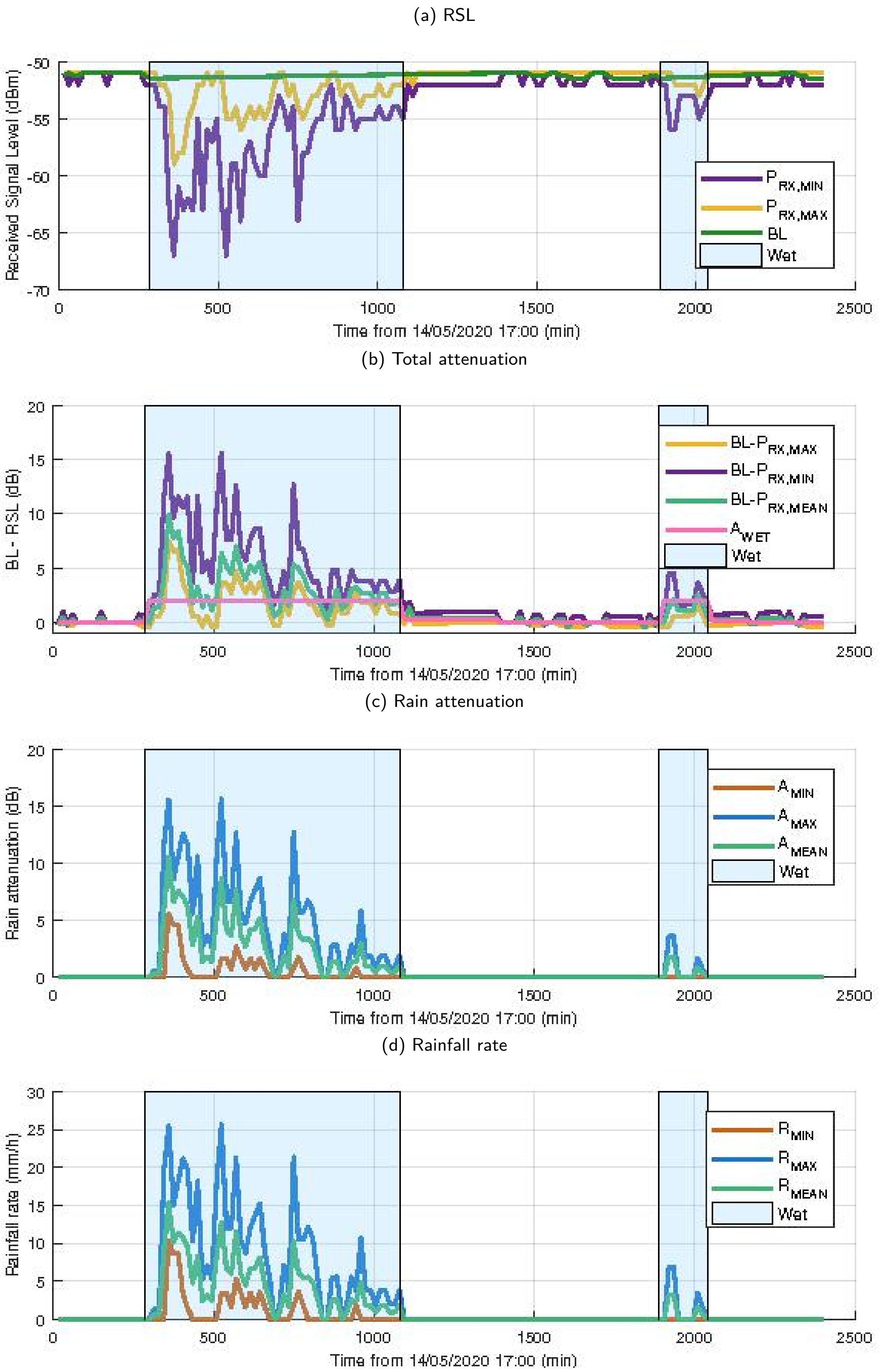

Figure RC2.2: Example of CML data processing: (a) Min and max received signal level (over 15-min slots), wet slots and baseline level, (b) total attenuation and wet antenna component, (c) rain attenuation, (d) rainfall rate.

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 or 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. We partially modified the subsection **Calibration and validation of the hydrological model** to make it clearer.

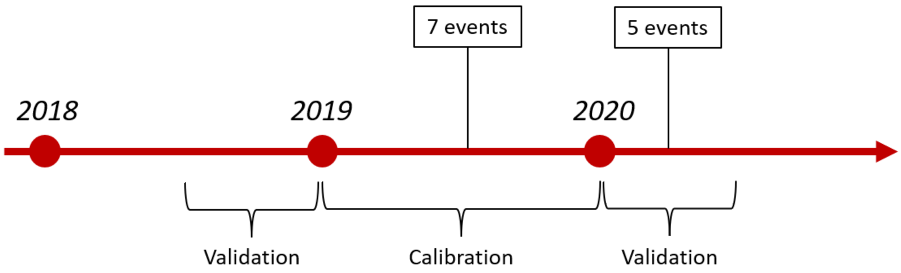

Figure RC2.3: Timeline with calibration and validation periods.

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 revised manuscript we reported 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..."**

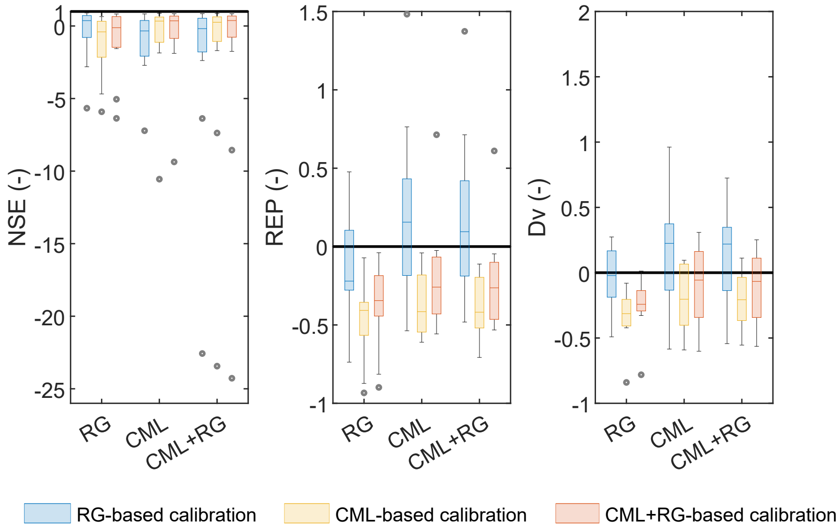

Figure RC2.4: Box plots of the performance indices for the 12 flood events obtained with three different calibration sets (drawn in as many colours) and with three different input types (on the x-axis). The ranges of the box plots are those reported in Fig. 8.

We agree with you, here "even if" is improperly used. We modified line 23 as follows:
However, because of economic or geographical factors, an adequate density of rainfall sensors is often not ensured.

- **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.*

- **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 added in line 41 the following text:

  Recently, the dual-polarization upgrade on radars (Zhang et al. 2019, Chen et al., 2021) has added information about shape, composition, and phase of the hydrometeors. 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 L115 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 added 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 added 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

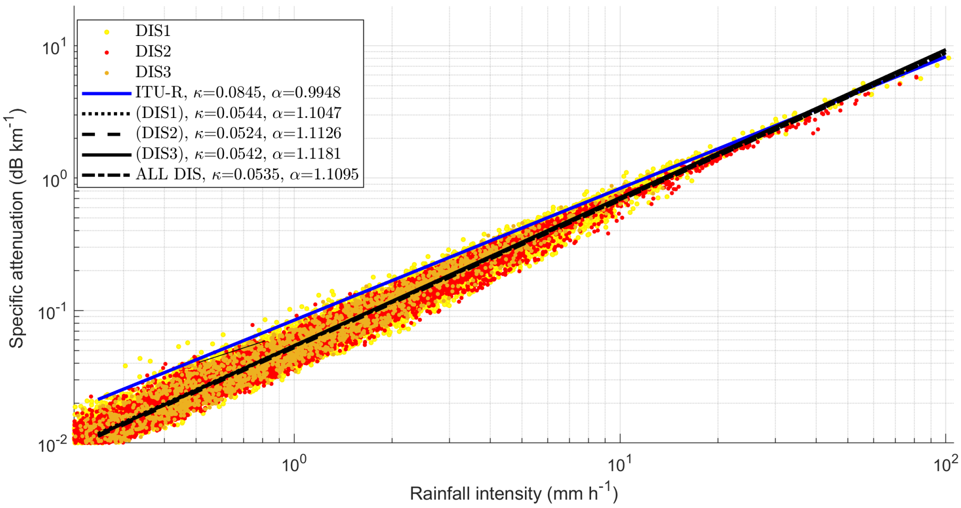

Figure RC2.5: Regression lines (in log-log scale) for k and α coefficients estimation.

  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 comment" n. 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 algorithm, 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 disregarded 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.

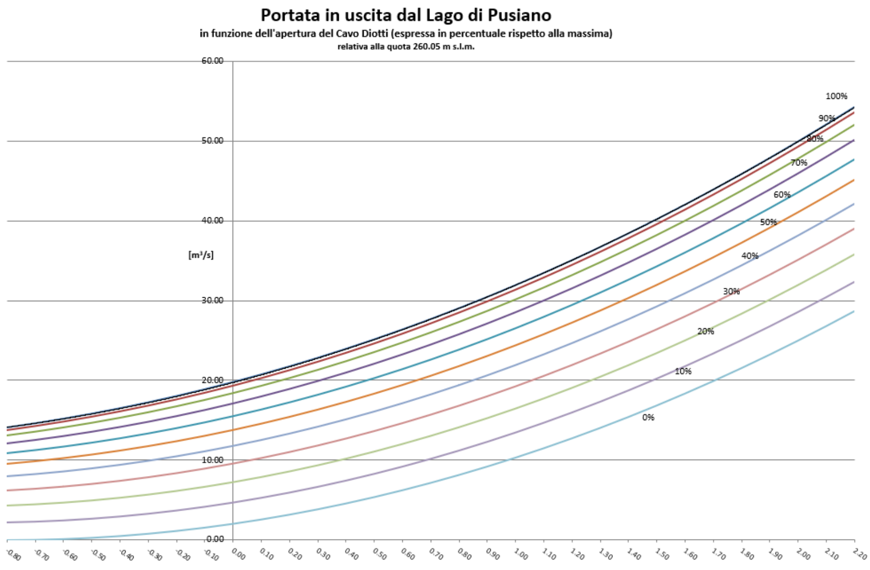

Figure RC2.6: Dam outflow in m^3 s^{-1} (y-axis) with respect to hydrometric level in m above the hydrometric 0 (x-axis) and the percentage of gate-dam opening (different colors of curves).

- **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 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.

  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

  We modified the text in line 131:
  According to the maximum observed rain rates, we classified events in low rain rate and high rain rate, adapting the classification reported in Met Office (2007) to our specific case study.

  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 disregarded 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 modified the paragraph in line 285 as follows: *Figure 6 shows the scatter plot of the rainfall accumulated at the end of each of the 12 storm events and averaged over the entire catchment area. Yellow markers are CML against RG rainfall depths, while in orange are CML+RG against RG rainfall depths. On the one hand, for all the low rain rate events (squares), estimates from CMLs and from CMLs+RGs are lower than the ones from RGs. On the other hand, CML (and CML+RG)-estimates of high rain rate events, are very nearly the same (with either lower and higher values) to the RG-based ones, with the only exception of event 3. From a more general perspective, the two regression lines indicate a good agreement between the two sets of sensors.*

  In addition, we changed the caption of Figure 5 (Fig. 6 in the revised manuscript).

- **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 added the equation of ΔE in line 290 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.

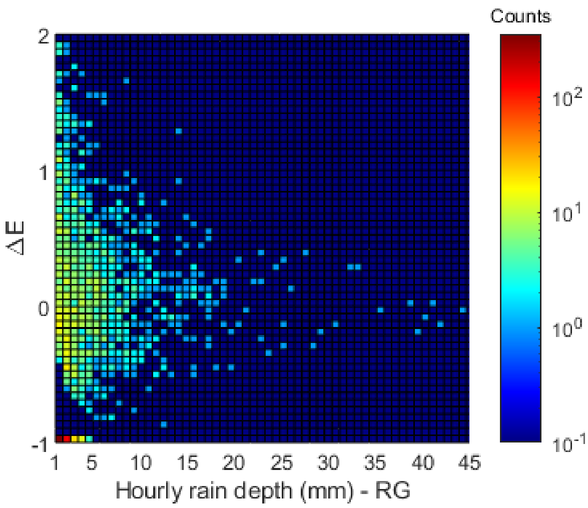

Figure RC2.7: 2D histogram of hourly rain depths and ΔE. The colour of each equally spaced 2D bin represents its height, which is the count of values falling in the bin. Note that the scale bar has a logarithmic scale and the dark blue bins have 0 counts. Values of ΔE equal to -1 represent false negative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

- **Fig 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.

- **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.

Figures/fig08.pdf

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 fixed 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).

- **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.