

**Interactive comment on “Rainfall retrieval with commercial microwave links in São Paulo, Brazil”
by Manuel F. Rios Gaona et al.**

Anonymous Referee #1

Received and published: 11 October 2017

Summary:

The paper presents an interesting topic. The authors analyze CML data in sub-tropic climate, namely Sao Paulo (Brazil), to derive rainfall information and validate it via a fairly dense network of rain gauges. This seems to be the first time a CML data set from this part of the world is analyzed in this sense, making the manuscript a potentially valuable scientific contribution in AMT. However, in my opinion the analysis is far from complete and misses out a lot of potential. As the authors state, and I acknowledge their honesty, they neglected the majority of the available CML data sets in their analysis, because, either their existing processing code cannot cope with it, or because comparison with nearby rain gauges was not possible or showed low correlation. This is a major shortcoming (see the list of my main concerns below). In general the paper is well structured and the writing is okay. Given the number of major concerns that I have and owing to the fact that this manuscript is already in open discussion, I recommend a major revision. Completely redoing the analysis with a new direction (focusing more on the CML data quality issue) and resubmitting would maybe be easier if the manuscript would not be openly available already.

R/. We thank the reviewer for the constructive assessment of our manuscript.

Main concerns:

- It has already been shown in numerous publications, among them many from the authors of this manuscript, that CML data can be used to derive reliable rainfall information. Hence, the result, that the authors can derive meaningful rainfall information from CML data is not very exciting news. The fact that the rainfall climate is different for the data set presented here, is relevant, however, the impact on the resulting rain rate seems to be negligible in comparison to the other uncertainties (e.g. the considerable differences of the relative bias for the 5 CML-gauge pairs, or the known uncertainties due to wet antenna, quantization, etc.).

R/. The fact that for this dataset the impact of a different rainfall climate indeed seems limited can also be interpreted as an import finding. It gives an indication that the rainfall retrieval from CML data is relatively insensitive to the rainfall climate, which confirms its general applicability. In addition, obtaining CML data is generally regarded as difficult. In this study CML data from, Brazil have been obtained, which shows that cellular communication companies are willing to provide data for yet another region.

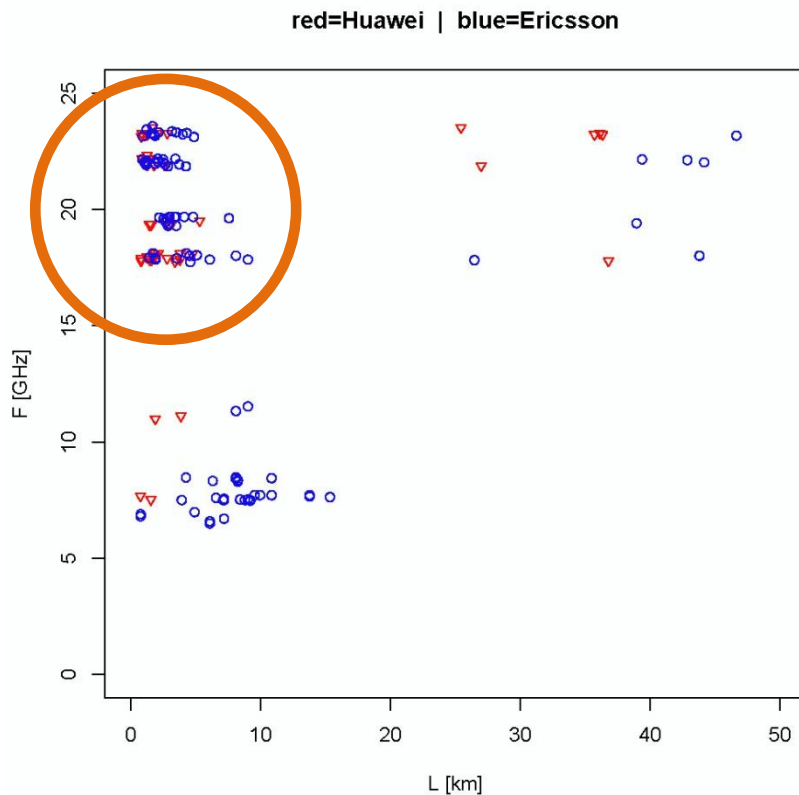
- Only being able to derive meaningful results for 5 out of 250 CMLs indicates that either the methods used by the authors are lacking or the technique of using CML data for rainfall estimation in general is less promising than expected.

R/. We acknowledge that our rainfall retrieval algorithm, developed for a temperate climate, may not be optimal for another climate, although we expect that the main principles, e.g. those behind the wet-dry classification, still hold. Rainfall estimation using CML data has already been successfully demonstrated for a variety of climates (Mediterranean, temperate, mountainous, and a tropical climate), although not for a subtropical one and not always on large datasets. For some seasons the rainfall characteristics from these climates will show similarities with a subtropical climate. Hence, we do expect that CML data are also promising for rainfall estimation in a subtropical climate. Please note that 35 out of 101 HU CMLs

could not be used because of lacking received signal levels due to issues in the log-file of the attenuation measurements. Moreover, 149 CMLs (ER) were not used because only minimum received powers were provided (which we will analyze in a revised version of this paper; see our response to the next comment). As far as we know, such a sampling strategy has not been encountered before by the microwave link rainfall estimation community, and RAINLINK has not been optimized for this. So actually we end up with 5 out of 66, instead of 250, CMLs providing reasonable estimates. Only 17 CMLs are within 1 km of a rain gauge, of which 6 CMLs were discarded due to dubious metadata (a strange microwave frequency and path length combination), and 6 other ones showing low correlation with rain gauge data perhaps because of erroneous metadata. So the number of CMLs revealing a clear rain signal could be (much) higher than 5 if those 149 CMLs only recording minimum received signal levels and those 49 more than 1 km away from a rain gauge would have been analyzed. Data availability and erroneous metadata certainly play a role. In a revised version of the manuscript we will greatly extend the analyses to include as many CML data as possible. Firstly, we will analyze the 149 Ericsson links as well (see our response to the next reviewer comment). And secondly, we will show rainfall accumulations for all links (also those that do not have a gauge within 1 km). Note that we will use a simple quality check on the gauges before we decide to use them in further analyses, and that we will remove all links that have unrealistic length-frequency combinations from the analyses.

- The fact that the majority of the CML data, the Ericsson data which only provides the minimum signal levels, cannot be used with the existing codebase of the authors (RAINLINK) should not be an excuse for not analysing it. Rather this calls for adjusting or extending the existing code.

R/. We agree with the reviewer that it would be valuable to also obtain rainfall estimates from CMLs providing only minimum received powers. It appears that the RAINLINK package itself does not need any modifications to do so. We only need to set $P_{\max}$ to $P_{\min}$ in a preprocessing step, and then use a conversion factor from maximum rainfall intensity to average rainfall intensity as a postprocessing step. Hence, we will show the analyses of the Ericsson dataset from its processing through RAINLINK in a revised version of our manuscript, as the reviewer encourages us now. The number of useful CMLs is expected to be much lower than 149. For instance, out of those 149 CMLs, many will be discarded because of the selection stated on page 4 of our manuscript “Our experience tells us that CML with both lengths above 20 km and frequencies above 15 GHz are not common in CML networks (they are highly unlikely from a network design perspective: long links experience more attenuation in rain, and should hence operate at low frequencies to limit this attenuation).”. A preliminary and equivalent figure for the Ericsson dataset (see below) suggest that the same proportion of reliable CMLs will be drawn from this dataset. Nevertheless, it is certainly worthwhile to also employ the Ericsson links.



* The orange circle indicates the region where reliable measurements can be drawn from. The region where $L > 20$ km and $F > 15$ GHz is clearly where the unrealistic length-frequency combinations are. We also choose to not include those links with frequencies below 15 GHz because these frequencies are less optimal for CML rainfall retrieval.

- The final analysis is based only on short or very short CMLs, but the authors do not state if they applied a wet antenna correction method, even though they note themselves that the effect of wet antenna can strongly impact shorter CMLs. This makes all the reasoning about biases arbitrary.

R/. We do apply a fixed wet antenna attenuation correction as described in Overeem et al. (2016a), using the default value of 2.3 dB. We will add this to the main text.

- The authors state that gauge records can also be unreliable, nevertheless they use low correlation with gauge records as indicator to neglect CML data.

R/. We agree with the suggested concept of “false negative” that the reviewer implies. Nevertheless, and given that rain gauges are the only available source we could refer our CMLs rain retrievals to, we considered that “it is likely that the high values of r^2 indicate that both types of observations contain a true rain signal.” (page 8). With lower correlations, it’s actually not possible to know whether the inaccurate rainfall estimate comes from the CML or gauge measurements. In the revised version of our paper, we will use a basic quality check on the gauge data before using the gauges in the analyses. This quality check is based on comparison of (long-term) accumulations with those from its nearest neighbors.

Recommendations:

- I recommend an extensive major revision, i.e. a real extension of the current analysis (see my points below)

R/. We will present analyses on a greatly extended datasets in the revised version of the paper.

- Given the seemingly very heterogeneous quality of the raw data set (which is fine for an opportunistic sensing technique like the one used here), the scientific focus should in my opinion be to describe how to cope with this data quality issue.

R/. To provide suggestions on how to cope with the data quality issue is rather difficult given the errors in the metadata and the lack of a study confirming the quality of our reference rain gauge data. However, we do give recommendations on using link length and frequency information for link metadata quality control. And the RAINLINK package also includes several quality control steps. The analyses presented here are also an indication to what degree these quality control procedures work.

- The constraint to neglect CMLs which are further than 1 km away from a rain gauge should be weakened. One can argue about what a “reasonable” threshold distance for comparing two rainfall measurements is. But, 1 km is really very strict, in particular, since the CMLs integrate over hundreds of meters or several kilometers anyway. The increased distance between CML and gauge will add additional uncertainty for sure, but when I look at the presented results and the relative biases from Table 1, having more data for the analysis seems to be more important than absolute accuracy of rain rates and/or rainfall sums.

R/. We agree with the reviewer that the 1-km constraint removes a large part of the links. On the other hand, we would like to limit the influence of sampling errors on the analyses. In order to meet both of these requirements we present an additional global analysis where we show time series of rainfall retrievals from all links, compared to the averaged gauge accumulations. For the computation of statistics we will keep the 1-km constraint. In this way, we will show the potential of this method on a large dataset, as well as limiting the effect of sampling differences on the more quantitative analyses.

- The Ericsson data should be included, i.e. RAINLINK should be extended to be able to process this data, or other code should be written or reused.

R/. We'll include the analysis of the Ericsson dataset (see replies above).

Other major comments and questions:

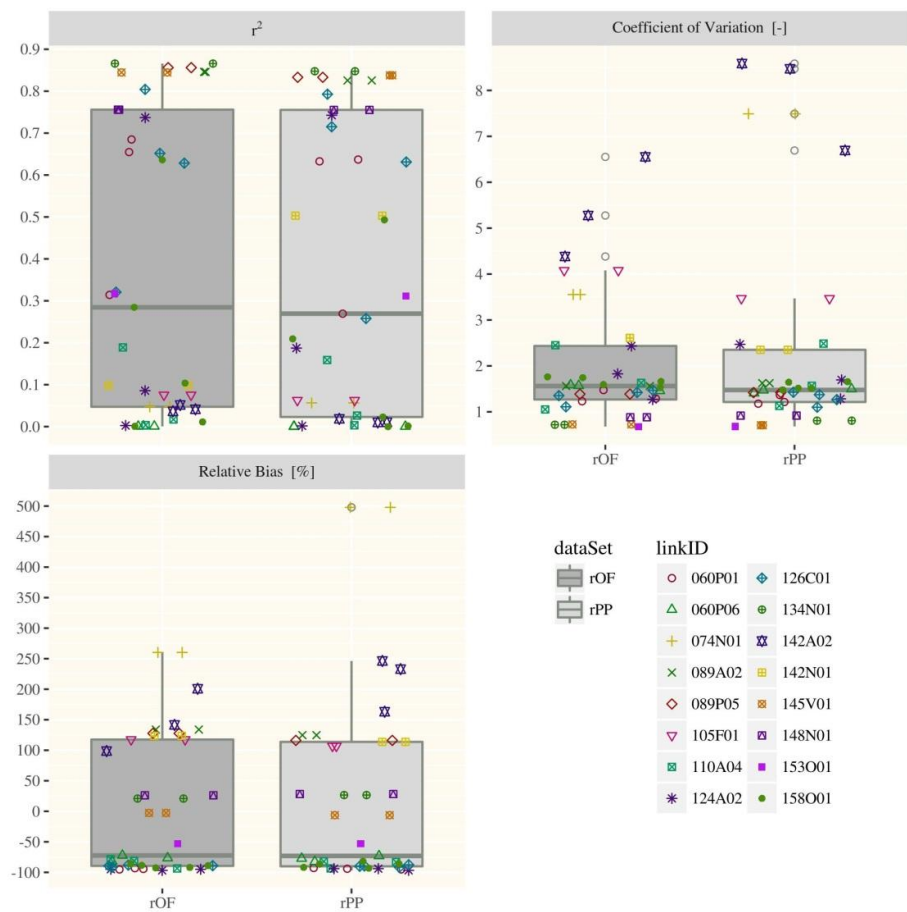
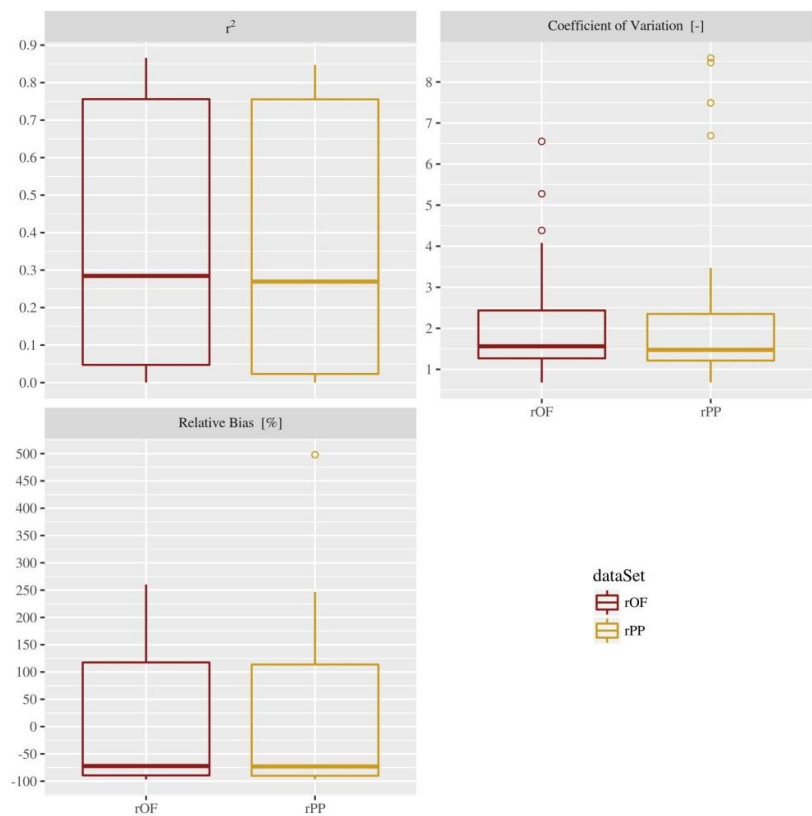
Page 4, line 22: What were the actual lengths and frequencies of the “long” CMLs? If the transmit power is high enough or large antennas are used, “uncommon” combination are possible. From Fig 1. some of the very long CMLs look strange indeed, though.

R/. Please see Figure 1, where frequencies against link-lengths are shown for all possible CMLs in our dataset. Not shown in this figure are several 100-km CMLs with frequencies close to 0. Personal communication with network design engineers from a telecommunication company in the Netherlands, confirms that the discarded microwave frequency - path length combinations should be erroneous. Having a larger transmit power to compensate for longer path lengths is generally not used because of the greatly increased probability of interference with other systems in the same band.

Page 6, line 13: A 50 km radius to look for CMLs with jointly decreasing power levels seems a bit large, in particular since, as the authors write in section 2.1 and 3.1, there is a lot of convective spatially very variable rainfall in the study region. Hence, is this radius of 50km too big? And how sensitive are the RAINLINK processing results on this threshold?

R/. Figures 2 and 3 (below) show the “in-sensitivity” of the analyses to either the use a radius-threshold (“rOF” approach) or the no use of it at all (“rPP”) approach. Figure 3 is equal to Figure 2 but all 16 CML in the analysis are plotted in the box-plots. In the revised version of the manuscript we will consider using a

shorter radius (15 or 20 km). Given the fact that we intend to use more links in the analyses, this should be feasible.



Page 8, line 7: Limiting the analysis to CML-gauge pairs were both show a rainfall depth above 0 mm, neglects the validation of the challenging step of detecting rain events in the CML time series, which, to my understanding, is the first step in RAINLINK. Wrong detections, i.e. missed rain events or artificially generated rain, may considerably add bias to the accumulations. Hence, this effect should be included in the validation or added in a separate validation.

R/. We agree with the reviewer that this is indeed an important aspect of CML rainfall retrieval. We will therefore include analyses of this aspect in the revised version of the paper.

Page 8, line 31ff: Given that this is the result for 1 out of 250 CMLs, I would recommend not to draw that optimistic conclusions based on the current state of the analysis.

R/. We will rephrase our conclusions if needed based on the analyses of the greatly extended CML dataset.

Fig 1: As it is mentioned in the text, the very long CMLs indeed look strange since they do not even end on one of the summit of the mountains in the north and north-east. Wouldn't it be possible to check via GoogleMaps satellite images if there is a relay or cell phone tower there? It would be nice to have a more solid basis for neglecting these CMLs. At least give more details in the text. Maybe it would also be good to show two or three maps, one with all CMLs, one with "reasonable" CMLs and in addition only the CMLs used for analysis (which hopefully will be much more in the next revision of the manuscript: : :).

R/. Checking the locations of the antennas of links on e.g. Google Maps could indeed be a valuable addition. However, the effort of manually checking antenna locations is not feasible for the large dataset we're dealing with here. It is also important to realize that antennas that were previously used for other links could have been re-used without having changed the location metadata in the database (a likely error; personal communication with representatives from a cellular communication company in the Netherlands). This means that there are likely still antennas at that location, but the specific antenna will have been moved. Hence, checking for the presence of antennas on Google Maps is likely not to yield the necessary information.

Technical and minor comments (this is a uncomplete list, since I assume that the manuscript will considerably change with the next iteration):

Fig 2: I only see 4 crosses not 5 as indicated in the caption. Also the red circles and red crosses seem not to add up to 11. Maybe overplotting is an issue here. If yes, this should be mentioned. Furthermore, no CMLs longer than 8 km are shown, even though the caption states that all HU CMLs are plotted, for which, according to Fig 1., some are definitely longer than 8 km.

R/. This is indeed due to overplotting, which will be mentioned in the caption.

Huawei CMLs with no useful data (or not close to gauges) could be larger than 8km (see Figure 1 of this rebuttal). But Figure 2 only shows the 11 selected CMLs, which are all shorter than 8 km. In the revised version of the manuscript, we will extend the axes of this figure to include all links that were used in the analyses.

Fig 5: The two yellowish colors are hard to distinguish. Anyway, if colors are different, markers could maybe be the same to make the graph easier to read. Or even better, have separate scatter plots for the CMLs, or at least for selected ones, if the number of CMLs increases with an extended analysis.

R/. The reason for using different markers for the different links in this graph is indeed that some of the colors could be difficult to distinguish. We feel that the use of different markers is sufficient to make the graph readable. We will consider making separate plots if this is required due to the extended dataset.

Table 1 and Table 2: The relative biases are exactly the same in both tables. As far as I understood, Table 2 is based only on a subset of the rain events from Table 1. Hence, I assume there is something wrong with either Table 1 or Table 2.

R/. The relative biases presented in Tables 1 and 2 should be exactly the same. Table 1 presents the statistics for all 30-minute accumulations, whereas Table 2 presents the same statistics for event accumulations (where an event is a contiguous period in time with nonzero rainfall). Because the bias is the difference (between link estimates and gauge measurements) in total accumulated rain (normalized by the gauge accumulation), this bias should be exactly the same in both tables (because the total accumulations are also exactly the same). The difference between the two tables is in the other statistics.

Table 1 and Table 2: Is CML 12 and 13 along the same path, but just the two directions?

R/. Exactly. We will mention this in the revised version of the paper.

Anonymous Referee #2

Received and published: 15 October 2017

The paper proposes to analyze an important topic : possible use of CMLs data for quantitative rainfall estimation in one of the largest city under tropical climate, i.e Sao Paulo. As reminded by the authors the area is prone to intense rainfall, leading to flash floods and other natural hazards such as land slides. The authors and various other groups have already demonstrated the potential of the CMLs based method under a range of climate and weather situations (from widespread systems in the Netherland to intense convection in Africa, through Mediterranean areas and even mountainous regions). This new data set in Brazil is an opportunity to test the CMLs method in a more challenging context then in previous studies: the quality of the CMLs data set is not homogenous, the validation network is sparse. The authors seems to have partially avoided this challenge by focusing only on a very limited subsample of the data set (where and when it works: : :); unfortunately this also limits the scientific impact of the study and its interest as a demonstrator of CMLs potential for hydro-meteorological monitoring over Sao Paulo: : :.

Given the existing literature on the CMLs topic and the extensive data set available here, the present study should be taken a step further and provide a more robust and extensive analysis of the available data set, including issues related to data quality, sparse GV and data format variation among CML providers..

R/. We thank the reviewer for the constructive review. We will extend the analyses by including the 149 Ericsson links, from which only minimum received powers are available. And we will present the rainfall retrievals of the links further than 1 km from a gauge in a separate analysis. In many cases it will be difficult to investigate the data quality in more detail, because of erroneous metadata or the lack of information regarding the quality of the rain gauge network used as a reference. We will, however, use a basic quality check on the gauge data to remove malfunctioning gauges. Please see our replies to reviewer #1 for more details.

A major limitation of the paper in its present form is that conclusions are drawn from a very limited subset of the available data set : only a few links (5 out of a possible total of above 200) are exploited and for theses links the analysis is restricted to time steps where both the link and the nearby gauge

detect rainfall. Doing so the authors miss a major issue : capability of the method to detect rain and not generate false alarms, and so over the whole network.

R/. *In the revised version of the paper we will present analyses of a greatly extended dataset, and we will adapt the conclusions if necessary. Furthermore, we will add information about the effectiveness of estimating zero rain (see our replies to reviewer #1 for more details).*

This is a major forthcoming of an otherwise very well written paper, which also provides a good review of the state of the art in CMLs based rainfall estimation. I can only encourage the authors to take the necessary time to submit an improved version of their work and take the analysis a step further.

Detailed major/minor recommendations :

-One important feature of sub-tropical rainfall is the occurrence of intense (and possibly extreme) rainfall rates associated with convective cells. This is very important for some of the applications the authors put forward in their introduction . No information is provided on the actual rain rate distribution (at the 30 minutes time step for instance) observed over the study period in Sao Paulo by the gauges and how well (or not) the CML method retrieves it. The global statistics provided in Table 2 and 3 do not inform us on the performance of the Rainlink/CML data according to rain rate classes . This is an important question, for hydrological applications for instance.

R/. *We agree with the reviewer that this is indeed a relevant question. However, we feel that this is outside of the scope of the present paper and a topic for future research. We will state this as a recommendation for future work in the revised version of the paper.*

-Selection of the time steps and 'events'. The authors should provide statistics covering the whole analysis period and not solely on a selected number of 30' times steps. Time step where one OR the other sensor detected rain should be included and a contingency table provided. The definition of 'events' , as presented in Fig 5 is not clear. Does it include some non rainy time steps or is it based on the same selection as the 30 ' (both CMLS and gauge > 0)? Daily statistics would be useful and would allow comparisons with other studies : : .

R/. *We decided to provide conditional statistics (i.e. $R > 0$ mm/h) on CML rainfall retrieval performance. In the revised version of the paper we will also include information about the effectiveness of estimating zero rain (see our responses to the comments by reviewer #1 for more details).*

-The authors mention wet antenna as a possible source of bias : this should be explored further - The order of magnitude of wet antenna attenuation is known, is it compatible with the observed bias ?

R/. *We already apply a fixed wet antenna attenuation correction of 2.3 dB. This is just an average value. For a chosen rainfall event, the wet antenna attenuation may differ a lot since one, two or no antennas can become wet. Moreover, the amount of attenuation can also depend on rainfall intensity, and biases can also be caused by other phenomena. Hence, it is rather difficult to assess whether remaining biases are caused by wet antenna attenuation. At least part of the wet antenna attenuation has been corrected for.*

-CMLS data selection : the authors should extend the analysis to other CMLs links even if they keep the present 5 links to illustrate the best case— This is important to assess the actual potential of the method in a context representative of reality. Given that the analysis is carried out at the 30' and 'event' time step, 1 km maximum distance from the gauge seems very severe.

R/. *In the revised version of this paper we will present analyses with a greatly extended dataset of CMLs. More details of the additional analyses can be found in our replies to the comments of reviewer #1.*

The conclusions should be revised once a truly extensive assessment has been done on this data set. I am looking forward to see a revised version that will investigate further this rich data set acquired in Brazil !

R/. We will modify our conclusions based on the analyses of the much larger dataset if necessary.

Anonymous Referee #3

Received and published: 17 October 2017

General comments:

The manuscript Rainfall retrieval with commercial microwave links in Sao Paulo Brazil aims to evaluate potential of commercial microwave links (CMLs) as rainfall sensors in the subtropical climate. The authors collect data from several microwave links, process them using RAINLINK R package and compare them to rain gauges operated by CAMADEN. Moreover, disdrometer observations are used to estimate parameters of attenuation-rainfall power-law model and these parameters are compared to those from ITU recommendations and from Dutch case studies.

Although the topic of CML rainfall retrieval in subtropical climate is relevant and the presented dataset is valuable the study has several major drawbacks: i) the authors select for evaluation only well performing CMLs. This is a reasonable approach if the selection procedure is independent of a reference rainfall data. However, this is not the case, as one of the selection criteria is correlation of CMLs to the reference rain gauges ii) the results are presented and discussed very briefly without sufficient attempt to investigate the causes of good/bad performance of particular CMLs. The influence of drop size distribution to the attenuation-rainfall model is analyzed in more detail, however, this effect can explain only small fraction of total errors. Especially spatial representativeness of reference rain gauge data should be more properly analyzed to avoid interpreting discrepancy between path-integrated CML rainfall observation and point RG rainfall observation as an inaccuracy of a CML iii) The conclusions are not sufficiently supported by the data. The authors claim that CMLs are very promising source of rainfall data only based one very good and two relatively well performing CMLs. Also the suitability of RIANLINK package for CML processing in subtropical regions is not proofed. The data rather indicate that constant WAA correction used in the RAINLINK package is inappropriate for CMLs.

Given the above mentioned shortcomings the reviewer does not recommend the manuscript for a publication, however, encourage the authors to improve the data analysis, rewrite especially the results, discussion and conclusion sections and resubmit the manuscript. Some suggestions for revisions are given in the specific comments bellow.

Specific comments:

The reviewer suggests changing the structure of the manuscript: i) moving the descriptions of the evaluation procedure (event definition) from the Results and Discussion section to the Data and Methods section, ii) considering moving the results of DSD analysis from the Rainfall retrieval algorithm section.

R/. We feel that the description of the evaluation procedure is an important part of the Results and Discussions section, and that moving this part would not increase the readability of the paper. The same holds for the DSD analyses and the Rainfall Retrieval Algorithm section. We will therefore keep the structure of the paper as it was.

P4L26: Is the threshold value $r^2 \geq 0.7$ chosen arbitrary? Why not 0.5 or 0.9? In any case, the selection of CMLs for evaluation based on reference data does not enable to evaluate potential of CMLs without having reference rainfall. This is one of the major drawbacks of the whole analysis. Moreover, it might be valuable keeping the bad performing CMLs in the analysis and identify the causes of the bad performance.

R/. We agree with the reviewer that the value of 0.7 is indeed arbitrary. We will now implement a very basic quality control to the gauge data to remove erroneous gauges, so that this issue will be less important. We will also remove the constraint that r^2 should be above 0.7 for presenting the analyses. An attempt will be made to identify the causes of bad performance, keeping in mind that we will not get additional information about the link network (limiting the amount of explanation that we can give).

P5L5: Given the CML paths lengths from several hundreds of meters up to several km the criterion of 1 km distance from link path seems to be too strict and not always reasonable. E.g. for CML 14 it might be more representative to use average of two RGs even though the second RG is several km far away. In any case, the reviewer suggests presenting at least some basic analysis of RG correlation and set the criterion based on this analysis. Such analysis would also support the results and enable to distinguish between discrepancy of path and point measured rainfall and between errors due to inaccuracy of CMLs.

R/. In the revised version of the paper we will employ a basic quality control on the gauge data based on nearby gauge data. We will not extend the radius of 1 km in order to limit the effect of differences in sampling. Rather, we will present CML rainfall retrievals for all links (also those that are more than 1 km from a link), and compare this to the average of all gauges. See our replies to reviewers #1 and #2 for more details.

P6L21: The section describes rather in detail generally well known performance metrics, however does not provide complete information about evaluation procedure. E.g. it should be explained here how the event based evaluation is performed (metrics are calculated for each event and then averaged as presented in Tab 2?).

R/. The statistics of the events are not averaged and then presented on Table 2. Events are established per link (as Table 2 clearly indicates) and statistics computed for all events altogether. We will improve the explanation of the “event” evaluation and how statistics are computed in the revised version of the paper.

P7L15: why -90 dB and not some other value?

R/. We use this value because from our experience this value indicates an extreme value that is not caused by rain.

P7L17: Both overall evaluation and event based evaluation is presented here. This is very good idea, as one could learn e.g. during which types of events CMLs perform well. However, at the end the event based results are presented in overall statistics (Tab. 2) except results presented in the Fig. 5. It might be very interesting to see how stable the CML performance is (e.g. in terms of variance of the metrics). This could be presented as boxplots or scatter plots of metrics, similarly as on Fig. 5. This would also enable more proper discussion of the results with potentially answering to questions like these: Do CMLs perform better during strong rainfalls than light rainfalls?

Do they better reproduce rainfall temporal dynamics (r^2) during light or heavy rainfalls?

R/. We agree with the reviewer that this is indeed interesting to know. However, we feel that this is outside of the scope of this paper. Furthermore, computing r^2 for classes of rainfall intensity is likely to

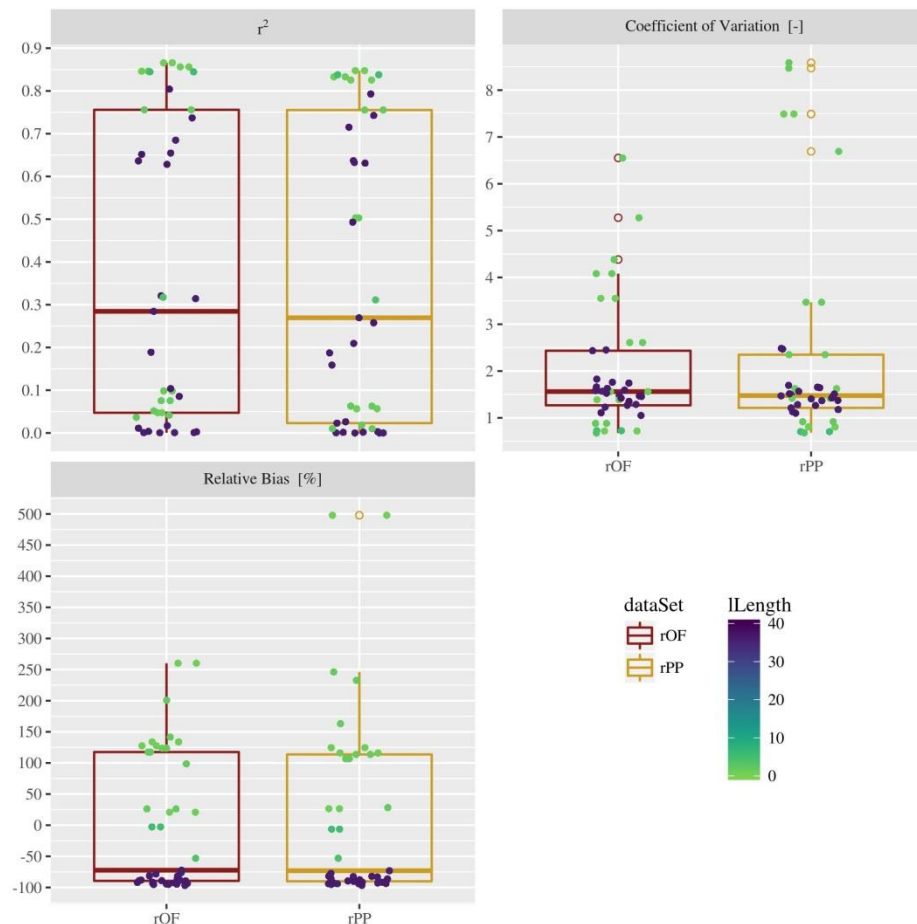
heavily influence the values of r^2 themselves (by taking classes of rainfall intensities, the variance is greatly reduced, yielding a likely increase in the fraction of explained variance).

P8L6-L11: The event definition might be rather in the method section

R/. We will keep the structure of the paper as it is. See our reply to the earlier comment about this.

P8L14-18: It seems that shorter CMLs are substantially more biased than longer CMLs. This indicates that the bias arises from wet antenna attenuation. Thus, RAINLINK's representation of baseline (constant) seems not working very well.

R/. The fact that short CMLs have large biases may indeed be caused by wet antenna attenuation. However, the shortest link in our analyses (link 06) does not have a large bias. Also, the fact that these short links (links 12 and 13) have a very high r^2 (mostly >0.8) coupled with such high biases indicates that the errors are likely to be multiplicative (wet antenna attenuation is mostly additive). A possible explanation can be that for these links the metadata is also incorrect (either the frequency, inducing a multiplicative bias in the prefactor of the k-R relation; or the link length, inducing a multiplicative bias in the specific attenuation itself). We will add a more thorough discussion on this in the revised version of the manuscript.



P8LL35 – P9L2: The performance was clearly very good only for one CML whereas the other experience relatively high bias. This is not really proving the good performance of RAINLINK in subtropical regions.

R/. We agree with the reviewer that this doesn't prove that RAINLINK performs well in subtropical regions. We will provide a more thorough discussion of the observed statistics (see also our previous

reply). And we will adjust the conclusions drawn in the paper if this is needed based on the analyses of the extended dataset.

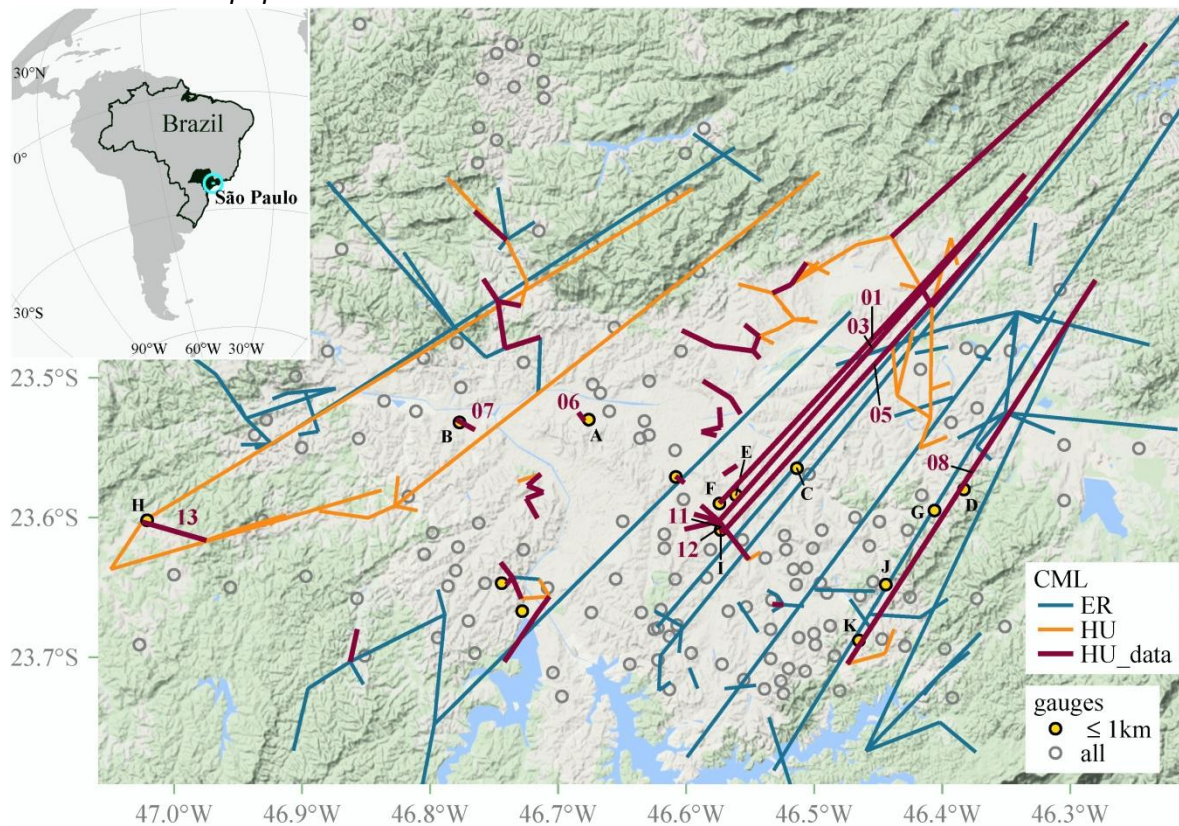
P9L18-20 and P10L4-6: Only three CMLs out of 17 resp. 11 were identified (based on reference rainfall) as well performing. The suitability of RAINLINK for processing such data should be, therefore, discussed more critically. Similarly, the authors claim that the potential of CMLs would be great if the data and metadata are properly stored. This is unfortunately not happening in the reality as demonstrated by the presented results.

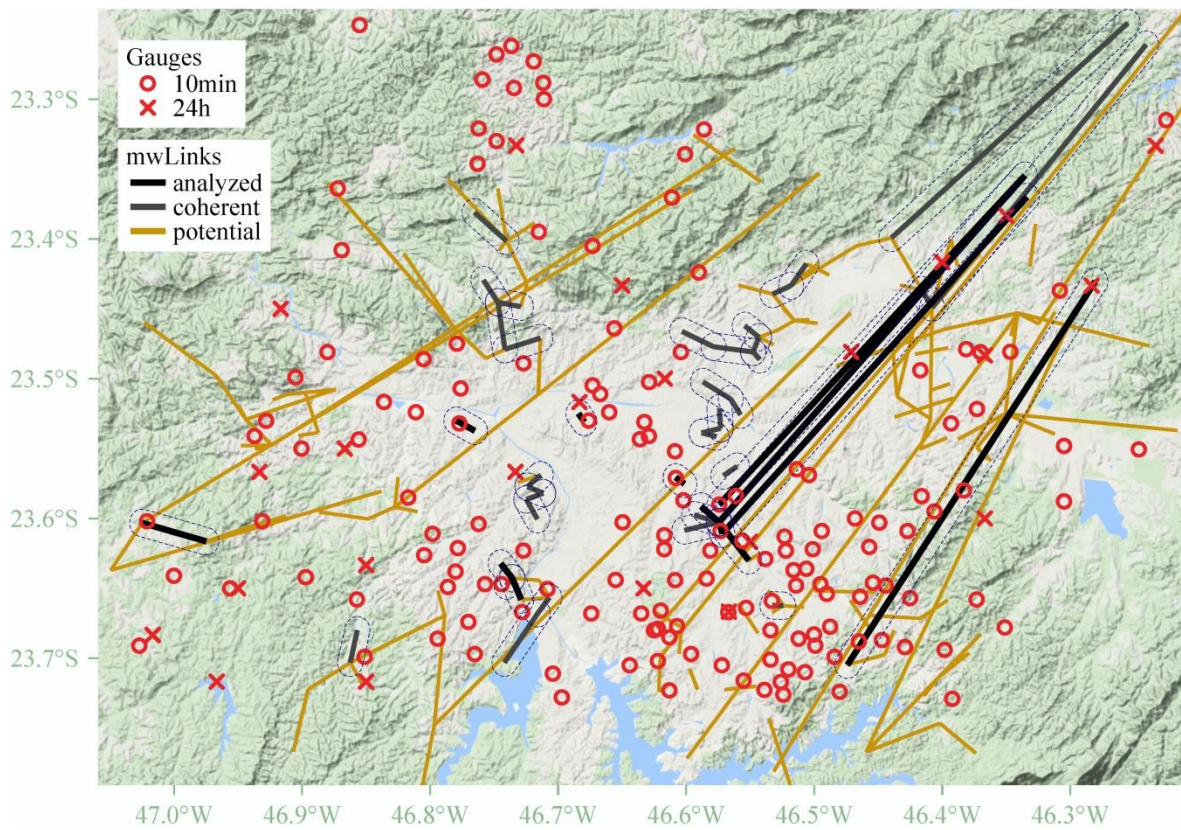
Thus, use of CMLs for subtropical regions is still rather big challenge. The dataset presented in this paper might, however, contribute to coping with this challenge. Thus, the reviewer highly encourages the authors to invest more work into its analysis and resubmit the improved manuscript.

R/. The challenge seems more related to (meta)data issues in CML data and the gauge data. As replied to Reviewers #1 and #2, we intend to increase our analysis. And we give suggestions on how to cope with metadata issues. See our responses to the comments of reviewers #1 and #2 for more details.

Fig.1: CMLs selected for the analysis are really tiny in the figure. Maybe cropping and resizing the figure would help (long CMLs aiming to the north-west could be cropped as they are not used for the analysis).

R/. We will keep the figure as it is, partly also because we want to stress the importance of having correct metadata in this paper.





Tab. 2: It seems to be there is no distinctive difference in the effect of DSD when evaluated over the whole dataset (tab. 1) and event based. It might be, therefore, reasonable to present here only results for fitted DSD (i.e. best performing a , b parameters) and instead one value (Mean of a metric?) present e.g. mean and standard deviation of a metric.

R/. We think that presenting the effect of using different k - R relations is very relevant here because this is a different climate than has been presented before (with corresponding differences in DSDs). We will therefore keep the analyses as they are.