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

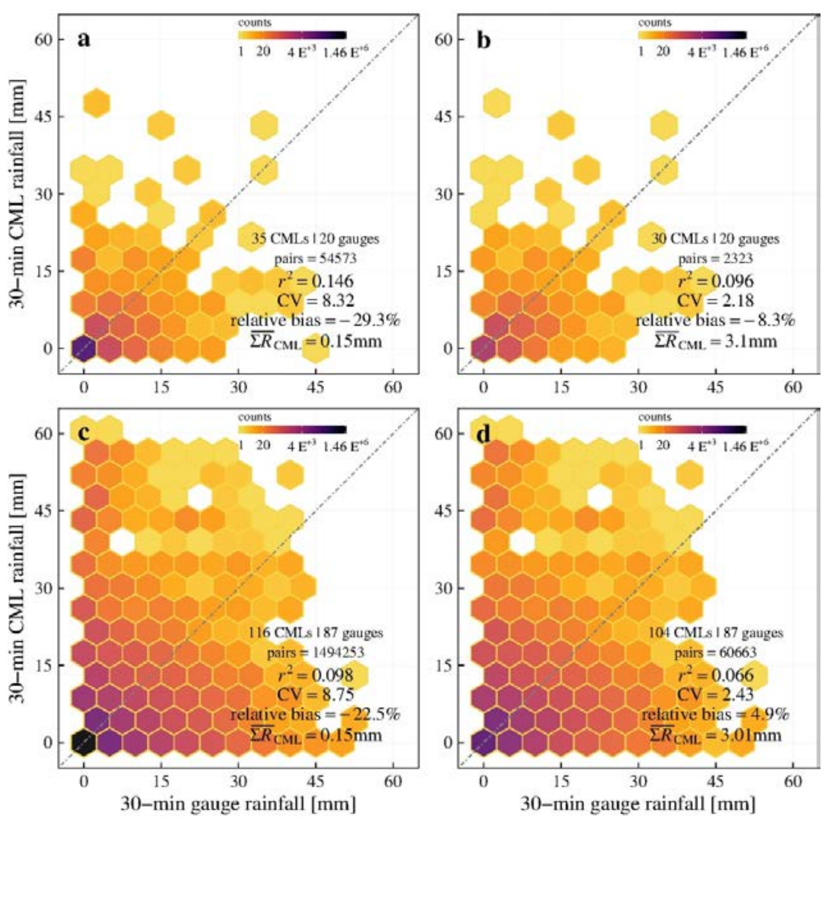
Anonymous Referee #1
Received and published: 11 October 2017

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/. *Through a more comprehensive analysis we found a suitable application of RAINLINK for a subtropical climate like the one of São Paulo, despite of RAINLINK being calibrated for a typical Dutch climatology. The suitable applicability of RAINLINK for such (subtropical) climatologies comes from the fact that we were able to identify its "alpha" parameter as 0.38, which is rather similar to the default one of 0.33 (which was obtained from Dutch rainfall data). Thus, as the reviewer suggests, this has indeed a negligible impact on the RAINLINK output. Nevertheless, following the same comprehensive analysis, we were able to identify the three best performing CMLs (Commercial Microwave Links), one of which is an ER (Ericsson) CML, which is three times more than we were able to identify for the first version of the manuscript. The particularity of these three best CMLs is that they are shorter than 1.7 km, where representativeness errors play a smaller role. We also found overestimations in CML estimates (Figs. 4 and 6 of the updated version of the manuscript). Such overestimations may be related to the fact that rain-induced attenuation along the link path may be relatively small compared to the attenuation caused by wet antennas, i.e., the wet antennas could contribute to some of the overestimations.*
Hence, and in the updated version, we now reflect on this stating that "The results of Fig. 5 are obtained for short links (< 1.7 km), where representativeness errors will play a smaller role. Overestimations by CMLs may be related to the fact that rain-induced attenuation along the link path may be relatively small

compared to the attenuation caused by wet antennas, i.e., the wet antennas could contribute to some of the overestimations." (for the case of rainfall overestimation over short-link paths); and "The 1-min rainfall intensities from the 3 disdrometers from the region of São Paulo are also employed to estimate the value of α used to convert the minimum and maximum rainfall intensities from the HU CML to mean 15-min intensities. The found value, 0.30, is close to the default one in RAINLINK, 0.33, based on Dutch data and used in this study. This confirms the usefulness of the default value of α for application in a subtropical climate." (for the case of the suitability of RAINLINK for subtropical climates). Please find in our reply to comment "P8L14-18" of reviewer #3 details with regard to the insertion of this new text in the updated version of the manuscript.

• 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 have now substantially improved the derivation of meaningful results. By means of implementing a gauge validation, and a modification in the RAINLINK code to derive rainfall intensities from minimum received power levels only, we now present meaningful results for a maximum of 116 CMLs. The analysis for 116 CMLs corresponds to a case in which we compared CML-derived rainfall against gauges up to a distance of 9 km in the vicinity of CMLs. If the vicinity is reduced to a 1 km, the results are meaningful for 35 CMLs. Still, from those 35 CMLs, we deemed as best performing CMLs those for which the relative bias is within $\pm$25%, and for which the coefficient of correlation is above or equal to 0.6.*
Figure 6 of the updated version of the manuscript (please see figure below) shows the results (and metrics) of these improved analyses. Overall, the presented metrics may suggest a poor performance of the network and of the RAINLINK estimates. Nevertheless, and throughout the updated version of the manuscript, we demonstrate how this technique is very promising, despite all the inconveniences found in the datasets.



• 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 have now extended the functionality of RAINLINK to be able to retrieve rainfall intensities from minimum received powers only. Thus, we have added 91 ER CMLs that were not previously analyzed. We now even have one best performing CML that came from this dataset. In the updated version of the manuscript, we indicate how this was done: "The ER CMLs only provide minimum power levels. RAINLINK is designed to retrieve rain rates from minimum and maximum power levels. Thus, in order for RAINLINK to compute mean rainfall estimates only from minimum power levels, two steps extra are required: 1) in the input file(s) for RAINLINK, the column with maximum power levels has to receive the values of the column with minimum power levels; 2) the mean path-averaged rainfall intensity, i.e. the output from RAINLINK, is now a maximum rainfall intensity and needs to be multiplied by a conversion factor to obtain the actual mean intensity…". Please find in the replies to your comments the complete updated text (and its placement in the manuscript).*

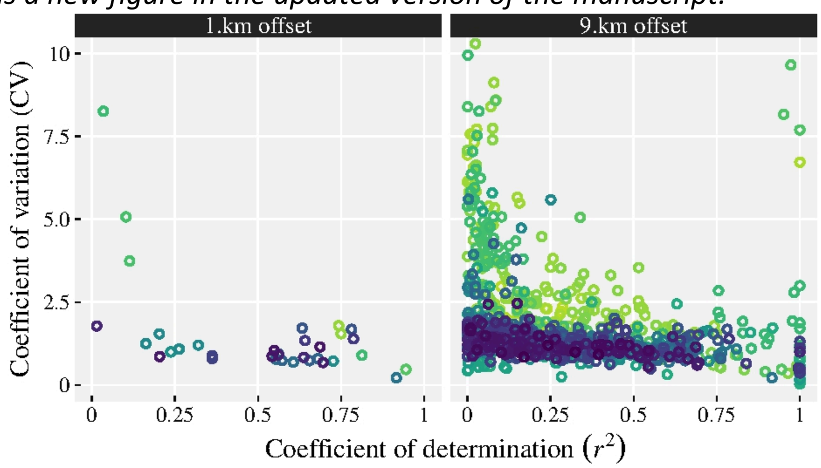
• The final analysis is based only on short or very short CMLs, but the authors do not state if they applied a wet antenna correction method, even though they note themselves that the effect of wet antenna can strongly impact shorter CMLs. This makes all the reasoning about biases arbitrary.
R/. *We do apply a fixed wet antenna attenuation correction as described in Overeem et al. (2016a), using the default value of 2.3 dB. Thus, in the last paragraph of sub-section "2.3 Rainfall Retrieval Algorithm", the sentence "4) rainfall retrievals -- once attenuation estimates are obtained from the difference between RSL and the reference signal level, 15-min average rainfall intensities are computed from a weighted average of minimum and maximum rainfall intensities obtained by the (inverse) power-law of Eq. (1);" was rephrased as* ***"4) rainfall retrievals -- once attenuation estimates are obtained from the difference between RSL and the reference signal level, a fixed wet antenna attenuation correction is applied (2.3 dB), and subsequently 15-min average rainfall intensities are computed from a weighted average of minimum and maximum rainfall intensities obtained by the (inverse) power law of Eq. (1);"****. Please see our remarks to your very first bullet (above) with regard to the effect of wet antenna on short CMLs.*

• 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/. *As suggested in our reply to the second bullet of the reviewer, we have now implemented a gauge validation procedure in order to validate gauges, and CML retrievals more consistently. The validation procedure is as follows: 1) For every gauge (152 in total) the closest two gauges were selected for comparison; 2) For the entire period, 30-min rainfall pairs (dry periods included) were evaluated throughout the relative bias and the coefficient of correlation for both closest gauges; 3) If the metrics of at least one of the two closest gauges are within $\pm$25% for the relative bias, and $\geq$0.6 for the correlation coefficient, the gauge under evaluation was deemed reliable.*
We describe this procedure in the updated version of the manuscript by replacing the sentences "Stations located within 1 km distance from the evaluated link paths were selected. Hence, only 11 stations were used to evaluate CML rainfall estimates in Sao Paulo." (at the end of the second paragraph of sub-section "2.2 Data") by ***"A gauge validation procedure was necessary due to availability issues and doubts about the quality of the rainfall observations from the CEMADEN gauge network. The validation procedure is as follows: 1) For every gauge (152 in total) the closest two gauges were selected for comparison; 2) For the entire period, 30-min rainfall pairs (dry periods included) were evaluated through the relative bias and the coefficient of correlation for both closest gauges; 3) If the metrics of***

at least one of the two closest gauges are within ±25% for the relative bias, and ≥0.6 for the correlation coefficient, the gauge under evaluation was deemed reliable. This selection results in 96 valid gauges out of 152. Comparisons of city-averaged rainfall were carried out among data from valid (96), and all (152) gauges, and all (145) CMLs (Fig. 4). For comparisons of individual path-averaged estimates of CMLs against gauges, only gauges within 1 or 9 km from the evaluated link paths were selected. For the 1-km case 35 CMLs were compared against 20 gauges, whereas for the 9-km case 116 CMLs were compared against 87 gauges."

Given that rain gauges are the only available source we could refer our CMLs rain retrievals to, we still consider that high values of r^2 indicate that both types of observations contain a true rain signal (with lower correlations, it is actually not possible to know whether the inaccurate rainfall estimate comes from the CML or from the gauge measurements). We reflect on this in the updated version of our manuscript in the new paragraph ***"Figure 7 shows the performance of individual CMLs by plotting the values of CV against r^2, based on CML-gauge pairs both above 0.0 mm (for the studied period). Many CMLs have fairly high values of r^2. For instance, 43% of the CMLs have a value of r^2 larger than 0.5 (for CML-gauge pairs within 9 km). Here, CML and gauge measurements are totally independent. Thus, it is very likely that the high values of r^2 for a large minority of CMLs indicate that both types of observations contain a true rain signal."****, which was added before the end of the sub-section "3.1 Evaluation of 30-min Rainfall" (now sub-section "3.2 Evaluation of 30-min Rainfall").*

Figure 7 (see below) is a new figure in the updated version of the manuscript.



Recommendations:

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

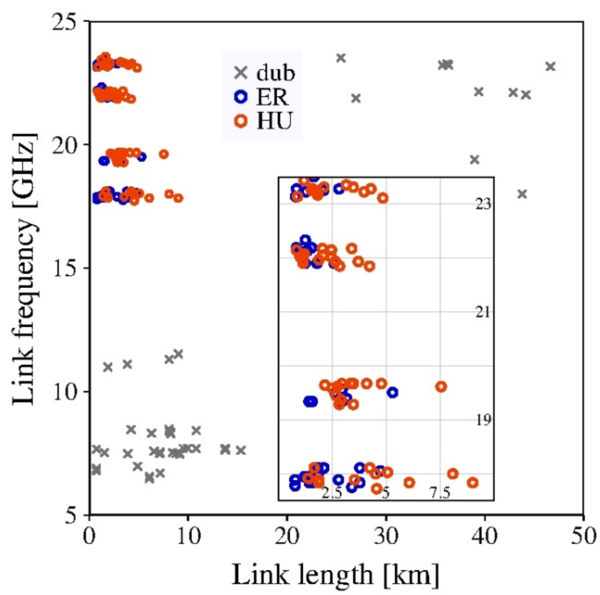
R/. *As the reviewer can see from the updated version of the manuscript, we carried out a much more comprehensive evaluation of the performance of CMLs in the city of São Paulo, implementing almost of the suggestions from the reviewers.*

• 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, the RAINLINK package also includes several quality control steps, and we now explicitly mention a couple of recommendations on using link length and frequency information for link metadata quality control. Such recommendations are made in the sub-section "2.2 Data" of the manuscript, as follows:*

- *In the first paragraph, "From the 66 HU CML, we selected 17 CML given their proximity to rain gauges (1 km or less)." was replaced by* ***"Figure 1 shows the location of these CMLs. [new paragraph] Figure 2 shows the scatter plot of link frequency against link length for all CMLs. In Fig. 2 the CMLs with uncommon or dubious (dub) combinations of length and frequency are denoted by gray markers (grey paths in Fig. 1)."****.*
- *At the end of the first paragraph, the sentences "Hence, we discarded 6 CML as dubious and did not consider them in our analyses, which reduced the number of CML to 11. Finally, from the 11 remaining CML, we only kept the 5 CML which showed clear rainfall signals as compared to nearby rain gauges, i.e. for which $r^2 \geq 0.7$. The other 6 CML practically showed no correlation with nearby gauges ($r^2 \sim 0.3$ for one CML-gauge pair, and $r^2 < 0.1$ for the other 5 CML-gauge pairs), due to malfunctioning gauges and/or CML data issues. Figure 2 shows the scatter plot of frequency against length for all HU CML." were replaced by* ***"The group of markers in the left bottom corner of Fig. 2 is also considered as dubious. Nevertheless, some CMLs around 7 GHz, having link lengths above 10 km, could still be realistic. We decided to only use the group of CMLs with path lengths shorter than 20 km and microwave frequencies above 15 GHz. Hence, 91 ER CMLs (40 link paths) and 54 HU CMLs (55 link paths) are left for the analyses, i.e., 145 CMLs in total (95 link paths)."****.*
- *Footnote 4 (previously and wrongly in Pag. 5) "We received CML data from a third party. It was not possible to verify the topology of the network, shown in Fig. 1 on-site, which we suspect not always to be accurate given the orientation of the long links." was rephrased as* ***"We received CML data from a third party. It was not possible to verify on-site the topology of the network shown in Fig. 1, which we suspect not to be accurate given the orientation of the long links."****, and now it is inserted on Pag. 16 (where it should have been placed).*

The updated Figure 2 is:



• 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 now present a comparative analysis for two CML-gauge distances (1 and 9 km), and a global analysis of rainfall time series for all CMLs against average gauge accumulations.*
For the comparative analysis of two CML-gauge distances (threshold distance) we decided to keep the one of 1 km, and use an alternative one of 9 km. This latter is based on the de-correlation distance (9.1 km) for 30-min rainfall in the city of São Paulo (value obtained from the gauge validation procedure). For each of these two distances (1 and 9 km), we also carried out two types of analyses: 1) where all possible paired rainfall depths were used; and 2) where only paired rainfall depths above 0.0 mm was used, i.e., to only account for significant/rainy events. Thus, in the updated version of the manuscript paragraphs 4 and 5 of sub-section "3.1 Evaluation of 30-min Rainfall" (now sub-section "3.2 Evaluation of 30-min Rainfall") were removed from the manuscript, and the following two paragraphs were inserted instead: ***"Figure 6 shows an overall assessment of the CML performance to retrieve 30-min rainfall depths (over the studied period). Scatter density plots are for CML-gauge pairs within 1 km (top panels, a and b) and within 9 km (bottom panels, c and d). The left column (panels a and c) is for all CML-gauge pairs, whereas the right column (panels b and d) only includes rainy intervals, i.e., CML-gauge pairs where both rainfall depths are above 0.0 mm. The rainfall estimates for CML-gauge pairs within 1 km are somewhat better than the ones for 9 km in terms of r^2 and CV, but the relative bias of the latter is smaller than that of the former. If all CML-gauge pairs are used, on average CMLs underestimate rainfall by 23-29%, with high values for CV and low values for r^2. Assuming that the gauges provide reliable measurements, this performance indicates that the applied wet-dry classification could be sub-optimal. Perhaps a sensitivity analysis of the threshold values in the wet-dry classification could improve this classification. If only rainy intervals are used, i.e., CML-gauge pairs both above 0.0 mm, these lead to a strong reduction in the value of CV, a decrease in the r^2, and a much smaller relative bias. [new paragraph] A reason for the large discrepancies among the statistics of the scatter density plots (Fig. 6) could be the fact that only minimum (and also maximum for HU CMLs) RSL data is used to compute 15-min rainfall intensities, i.e., a limited temporal sampling. Rios Gaona et al. (2015) compare CML (actual) and gauge-adjusted (simulated) path-average rainfall depths for a 12-day dataset from the Netherlands, based on rainfall pairs for which at least one depth exceeds 0.1 mm. The most prominent difference is their much higher value for r^2 (0.437), which was found for 15-min rainfall. Hence, the sampling strategy is not necessarily the main reason for the low values of r^2. Given the erroneous metadata found in the CML dataset (Sec. 2.2), which led to discarding CMLs with dubious combinations of path length and frequency, there could be errors in the metadata from selected CMLs too, i.e, wrong location of one of the antennas or wrong frequency. In addition, although a basic assessment of gauge quality has been performed, even records from gauges classified as valid could still contain measurement errors."***.
For the global analysis of cumulative rainfall series for all selected CMLs and gauges, averaged over the city of São Paulo, a new sub-section ("3.1 City-average Rainfall", right at the beginning of section "3 Results and Discussion") was added to the manuscript. This new sub-section focuses on the city-average performance of gauges and CMLs. The new paragraph reads as follows: ***"For each dataset we compute the cumulative city-average rainfall for the studied period (Fig. 4). According to the reference, i.e., the 96 valid gauges, the cumulative rainfall depth in this ~3-month period is ~600 mm. The differences in cumulative rainfall depths between the valid and all (152) rain gauges are small. Such a small difference suggests that the gauge dataset is reliable. For the "PreProcessed" CML dataset no wet-dry classification and no outlier filter are applied. This contributes to cumulative rainfall depths being roughly twice as large as the gauge-based ones. Moreover, the dynamics do not often correspond with that of the gauges, for instance around 1 December 2014. For the "OutFiltered" dataset of 145 CMLs,***

which includes a wet-dry classification and outlier filter, a much better correspondence is found. The dynamics of the cumulative series agree reasonably well, and an overall underestimation is found, ~200 mm at the end of the period, albeit much smaller than the difference between the "PreProcessed" dataset and the reference. The separate performance of the HU and ER CMLs shows that the HU dataset performs quite well with some overestimation, whereas the ER dataset gives a huge underestimation, despite roughly capturing the rainfall dynamics.".
Also, the last paragraph of sub-section "2.4 Error and Uncertainty Metrics" "The metrics were systematically computed on 30-min paired rainfall depths, both above 0 mm (to only account for significant rainfall events), and for which their equivalent 15-min minimum received powers (i.e., "min PRx ..." in Fig. 4) were larger than -90 dB. 30-min aggregation was necessary given the temporal resolutions of the datasets, i.e., 10 min for gauge and 15 min for link-retrieved data." was rephrased as **"The metrics were systematically computed on 30-min paired rainfall depths, using either all rainfall pairs or only pairs where both CML and gauge depths are above 0.0 mm. The latter to account only for significant rainfall events. 30-min aggregation was necessary given the temporal resolutions of the datasets, i.e., 10 min for gauge and 15 min for CML-retrieved data."**.

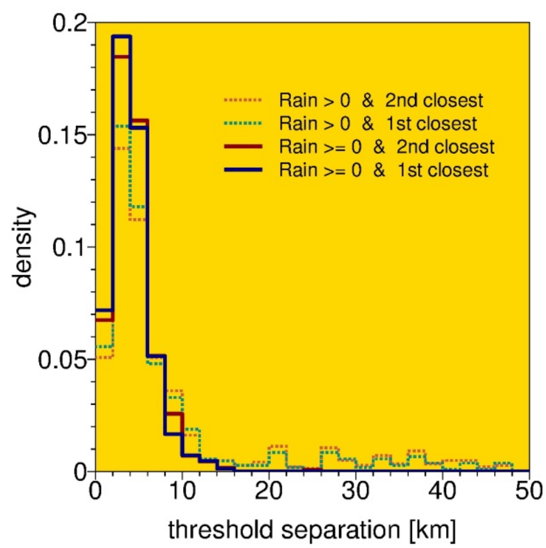
• The Ericsson data should be included, i.e. RAINLINK should be extended to be able to process this data, or other code should be written or reused.
R/. *The ER CMLs were included in the current analyses, and the methodology for RAINLINK to retrieve rainfall from minimum received powers only is clearly described in the revised version of the manuscript. Hence, the following two paragraphs were inserted at the end of sub-section "2.3 Rainfall Retrieval Algorithm" to provide a background on the retrieval of rainfall depths (from RAINLINK) for CML-measurements of only minimum power levels:* ***"The ER CMLs only provide minimum power levels. RAINLINK is designed to retrieve rain rates from minimum and maximum power levels. Thus, in order for RAINLINK to compute mean rainfall estimates only from minimum power levels, two steps extra are required: 1) in the input file(s) for RAINLINK, the column with maximum power levels has to receive the values of the column with minimum power levels; 2) the mean path-averaged rainfall intensity, i.e. the output from RAINLINK, is now a maximum rainfall intensity and needs to be multiplied by a conversion factor to obtain the actual mean intensity. This conversion factor needs to be determined by means of a calibration dataset. Here, we use the 1-min rainfall intensities from the three disdrometers from the region of São Paulo to obtain an estimate of such a conversion factor. For each 15-min interval, the minimum rainfall intensity is selected from the lowest intensity of the 15 1-min intensities. 0.38 was found as the conversion factor, by comparing this minimum rainfall intensity against the mean 15-min rainfall intensity from the same disdrometers. ER-CML maximum rainfall intensities are then multiplied by this factor to obtain (actual) mean rainfall intensities. [new paragraph] The 1-min rainfall intensities from the three disdrometers from the region of São Paulo are also employed to estimate the value of α used to convert the minimum and maximum rainfall intensities from the HU CMLs to mean 15-min intensities. The found value, 0.30, is close to the default one in RAINLINK, 0.33, based on Dutch data and used in this study. This confirms the usefulness of the default value of α for application in a subtropical climate."***.

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 the updated Figure 2 (please see the figure above in reply to the second recommendation of the reviewer), where frequencies against link lengths are shown for all possible CMLs in the revised*

dataset (both HU and ER). In this figure there are several 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/. *From the gauge validation procedure we found that the de-correlation distance of 30-min rainfall for the São Paulo are is 9.1 km. The figure below (not shown in the updated version of the manuscript) is a histogram of the distances at which the paired gauge evaluations comply with the thresholds of a relative bias within ±25% and a coefficient of correlation above 0.6. From this figure one can see that almost all of the distribution is within a 'radius' of 10 km (9.1 km being the arithmetic average). Hence, for the re-analysis presented in the revised version of the manuscript we modify the RAINLINK radius parameter to 9 km.*



For the revised version of the manuscript, we use the "OutFiltered" RAINLINK-approach, i.e., the approach including wet-dry classification and outlier filter.

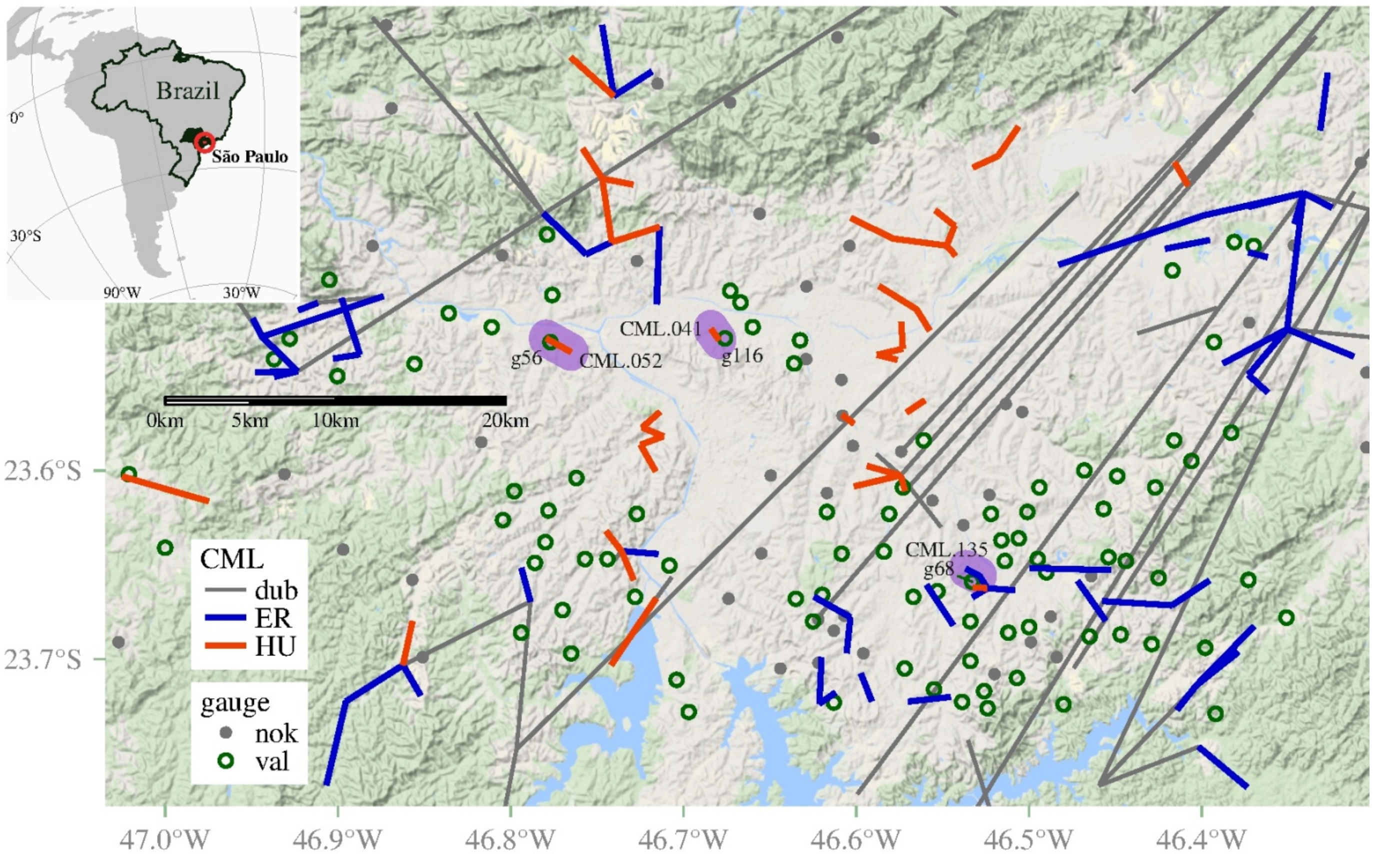
Page 8, line 7: Limiting the analysis to CML-gauge pairs were both show a rainfall depth above 0 mm, neglects the validation of the challenging step of detecting rain events in the CML time series, which, to my understanding, is the first step in RAINLINK. Wrong detections, i.e. missed rain events or 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 have included such analyses in the revised version of the manuscript. Please see the first part in reply to the third recommendation of the reviewer, in which we explain in detail how these analyses were carried out. Please see the reply to the second "main concern" of the reviewer in which the support figure (new figure in the updated version of the manuscript) of the analyses of rainy events and all events with dry spells is presented.*

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/. *For the updated version of the manuscript, we have now included the ER dataset, we have carried out more consistent and comprehensive analyses, such as city-average rainfall and wet and wet-dry spells, and we have even tripled the amount of "outstanding" results we have gathered for the first version. Such good and promising findings were updated accordingly in the conclusion section.*
Thus, in the section "Summary, Conclusions and Recommendations", the second paragraph "30-min rainfall estimates from CML were evaluated against rainfall measurements from the nearest rain gauge for the period from 20 October 2014 to 9 January 2015. We focused our analyses on the 5 CML for which r^2*>0.7. Three out of five CML gave good results in terms of CV. One CML also had a low relative bias. Subsequently, the quality of rainfall estimates from these 5 CML was also evaluated in terms of cumulative rainfall from 272 events. The good results indicate that RAINLINK can successfully be applied to CML data from a subtropical climate, even though most parameters have been optimized for the temperate climate of the Netherlands." was rephrased as* ***"30-min rainfall estimates from CMLs were evaluated against rainfall measurements from rain gauges for the period from 20 October 2014 to 9 January 2015. Despite the mixed results, the potential of CML technology for rainfall estimation in subtropical climates is confirmed. Especially, given the rainfall dynamics captured by the city-average rainfall (Fig. 4), the good performance of some individual CMLs (Fig. 5), and a high correlation for a large minority of CMLs (Fig. 7). This gives an indication that the RAINLINK package is suitable to retrieve rainfall via CML data from a subtropical climate, even though many of its parameters have not been optimized for such a climate. Since biases propagate in hydrological model predictions, given the low relative bias found for rainy periods (Fig. 6), CML rainfall estimates could be considered as an alternative input in hydrological models."****.*

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 are 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 moved. Hence, checking for the presence of antennas on Google Maps will likely not yield the necessary information.*
Please also see our reply to the second recommendation of the reviewer, in which we gave more precise arguments on how to discard dubious or erroneous link paths.
We have managed as well to implement in only one figure the suggestions of the reviewer concerning the display of used and discarded CMLs for the respective analysis. The figure below is the updated Figure 1 in the revised version of the manuscript.



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/. *The scatter plot the reviewer refers to has been updated in the revised version of the manuscript (please see our reply to the second recommendation of the reviewer, in which the updated figure is presented). The axes have been extended to show the characteristics of all possible CMLs. Now, Figure 2 is completely consistent with Figure 1 (figure immediately above).*

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/. *This figure (and section) has been removed from the manuscript. It does not appear in the revised version of the manuscript.*

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/. *Given that we now carried out different analyses (focused on the suggestions of all the reviewers), such tables are not needed anymore. Thus, these tables have been removed from the manuscript. They do not appear in the revised version of the manuscript.*

Table 1 and Table 2: Is CML 12 and 13 along the same path, but just the two directions?
R/. *Exactly. Nevertheless, as mentioned in the previous reply these tables do not appear in the revised version of the manuscript.*

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 have now extended the analyses by including all the ER CMLs. We have implemented the capabilities of RAINLINK to retrieve rainfall depths from only minimum received powers (which is the case of the ER dataset). We have also carried out analyses of rainfall retrievals not only for gauges within 1 km in the vicinity of selected CMLs but also for vicinities up to 9 km. The data quality is still difficult to investigate in more detail because of erroneous metadata or the lack of information regarding the quality of the rain gauge network used as a reference. Nevertheless, we have implemented a basic quality check on the gauge data to remove 'malfunctioning' gauges. Please see our replies to reviewer #1, especially those for "main concern" # 5 (fifth bullet), and recommendations # 2, 3, and 4 (second, third, and fourth as the reviewer actually does not provide numbers).*

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/. *As mentioned in the previous reply, in the revised version of the paper we have now carried out analyses on the whole dataset, i.e., HU + ER CMLs. We have also analyzed CML retrievals from wet, and dry-and-wet spells, to account for effectiveness of estimating zero rain.*
Please see our replies to reviewer #1, especially those of "main concern" # 2, and third recommendation (first part), in which we addressed specifically these issues.

This 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 a 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.*

This recommendation (jointly with others) was taken into account in the revised version of the manuscript, as follows:

In the previous to the last paragraph the sentences ***"We did not evaluate the performance of CML-RAINLINK retrievals based on rain rate classes. Nevertheless, this evaluation is highly encouraged as it would shed some light on the suitability of CMLs for hydrological applications, for instance. [new paragraph]"*** *were inserted after the period in "... paths. We...".*

Also, the sentences ***"Note that the value of α, estimated from local 1-minute disdrometer rainfall intensities, was close to the default value from RAINLINK. Especially the value of A_a and the threshold values for the wet-dry classification and the outlier filter should be investigated."*** *were inserted after the period in "... regions. Missing...".*

In its last sentence (paragraph previous to the last one) "This shows that accurate metadata, such as link coordinates for instance, are essential." was rephrased as ***"This shows that accurate metadata, such as link coordinates for instance, are essential, as well as the feedback about obtained CML and reference datasets.".***

-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 not to include daily statistics in the updated version of the manuscript. Nevertheless, we now compute statistics for the entire period of study (~3 months). For this updated version of the manuscript, we removed our analyses on 'rain events'.*

Please see our replies to the two reviews (previous to "Detailed major/minor recommendations:") of the reviewer.

-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 (please see our reply to the fourth "main concern" of reviewer #1). This is just an average value. For a given 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 asses 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/. *As also noted in previous replies to the reviewer, we have extended our analyses by adding the ER dataset. We have also carried out analyses within 1 and 9 km in the vicinities of selected CMLs (116 in total).*
Please see our replies to reviewer #1, especially those concerning the third recommendation (first part) and the fourth recommendation.

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/. *As noted in all the previous replies to reviewers #1 and #2, the conclusions of the revised version of the manuscript have been updated accordingly.*
Also, the fourth paragraph of the section "4 Summary, Conclusions and Recommendations" was removed.

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 r2 _ 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 have now implemented a basic quality control to the gauge data to remove erroneous gauges, so that this issue will be less important. We have also removed the constraint that r^2 should be above 0.7.*
Please see our replies to reviewers #1 and #2, especially that of reviewer #1 in his/her third recommendation (first part).
It was really difficult to identify the causes of bad performance as we did not get additional information about the link network.

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/. *As noted in previous replies to the reviewer, we have extended our analyses by adding the ER dataset. We have also carried out analyses within 1 and 9 km in the vicinities of selected CMLs (116 in total). Please see our replies to reviewer #1, especially those concerning the third recommendation (first part) and the fourth recommendation.*

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/. *As also noted in previous replies, we decided not to include daily statistics in the updated version of the manuscript. Nevertheless, we now compute statistics for the entire period of study (~3 months). For this updated version of the manuscript, we removed our analyses on 'rain events'. Tables 1 and 2 were also removed.*

P7L15: why -90 dB and not some other value?
R/. *Because now we present analyses based on the "OutFiltered" RAINLINK approach, we use the default RAINLINK value, i.e., -32.5 dB* km^{-1} h^{-1}.

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 (r2) during light or heavy rainfalls?
R/. *We agree with the reviewer that this is indeed interesting to know. However, we feel that this is beyond the scope of this paper. Nevertheless, in the previous to the last paragraph, after the period in "... paths. We...", we have introduced the following recommendation:* ***"We did not evaluate the performance of CML-RAINLINK retrievals based on rain rate classes. Nevertheless, this evaluation is highly encouraged as it would shed some light on the suitability of CMLs for hydrological applications, for instance. [new paragraph]"****.*
For more details, please see our reply to the first comment in "Detailed major/minor recommendations" of reviewer #2.

P8L6-L11: The event definition might be rather in the method section
R/. *We removed the analyses on 'rain events'. Please 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/. *As noted in our reply to the first "main concern" of reviewer #1, we were able to identify the three best performing CMLs. The particularity of these three best CMLs is that they are shorter than 1.7 km, where representativeness errors play a smaller role. We also found overestimations in CML estimates (Figs. 4 and 6 of the updated version of the manuscript). Such overestimations may be related to the fact that rain-induced attenuation along the link path may be relatively small compared to the attenuation caused by wet antennas, i.e., the wet antennas could contribute to some of the overestimations.*
The above discussion is inserted in the third paragraph of sub-section "3.1 Evaluation of 30-min Rainfall" (now sub-section "3.2 Evaluation of 30-min Rainfall"), where the sentences "The results of Fig. 4 are obtained for the longest link (5.3km), where representativeness errors could play a larger role. The worst results are found for the shortest links (<1.0km). This may be related to the fact that rain-induced attenuation along the link path may be relatively small compared to the attenuation caused by wet antennas, i.e., the wet antennas may explain the large overestimation found for CML 13 and 12 (see Tables 1 and 2)." were rephrased as ***"The results of Fig. 5 are obtained for short links (< 1.7 km), where representativeness errors will play a smaller role. Overestimations by CMLs may be related to the fact that rain-induced attenuation along the link path may be relatively small compared to the attenuation caused by wet antennas, i.e., the wet antennas could contribute to some of the overestimations.".***

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 have now demonstrated the suitability of RAINLINK for a subtropical climate to a greater extent. Please see our reply to the first "main concern" of reviewer #1.*

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/. *As noted in our replies to reviewers #1 and #2, we have substantially improved our manuscript.*

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/. *Figure 1 has been updated in the revised version of the manuscript. Please see our reply to the last comment of "Other major comments and questions" from reviewer #1.*

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/. *As noted in our previous replies, Tables 1 and 2 have been removed from the manuscript.*

OTHER COMMENTS.

The following are some other changes that have been implemented in the revised version of the manuscript:

The amount of CMLs and its distribution was updated accordingly throughout the whole manuscript.

In the abstract, "Results were found to be promising and encouraging, especially for short links, for which high correlations (>0.9) and low biases (~30% and lower) were obtained." was rephrased as ***"Results were found to be promising and encouraging when it comes to capturing the city-average rainfall dynamics. Mixed results were obtained for individual CML estimates, which may be related to erroneous metadata."****.*

"Uijlenhoet et al. (2018) give a non-expert summary of the history, theory, challenges, and opportunities toward continental-scale rainfall monitoring via CMLs of cellular communication networks." *was added as the last sentence at the end of the first paragraph of Pag. 3 (submitted version).*

"Because our CML retrieval algorithm RAINLINK (Sec. 2.3) only retrieves rain rates from minimum and maximum power levels, we discarded the ER CML. Due to issues in the log-file of the attenuation measurements, it was only possible to correctly and unequivocally assign power levels to 66 HU CMLs and

147 ER CMLs." was replaced by ***"The ER CMLs are assumed to have constant transmitted power levels."*** *in the first paragraph of sub-section "2.2 Data".*

The sentence "For the remaining 5 CML evaluated here, the mean difference between 15-min transmitted power levels is ~0.0 dB, with a maximum of 0.5 dB (for the 81 days considered)." was removed from the manuscript.

Sub-section "3.2 Evaluation of Event Rainfall Accumulations" was erased from the manuscript and replaced by the new sub-section "3.2 Evaluation of 30-min Rainfall".

The paragraph ***"For the studied period, we evaluate the quality of 30-min path-averaged rainfall estimates from individual CMLs against gauges by: 1) time series from rainfall events for the three best performing CMLs; 2) scatter density plots based on data from all CMLs; and 3) metrics for each CML."*** *was added at the beginning of sub-section "3.1 Evaluation of 30-min Rainfall" (now sub-section "3.2 Evaluation of 30-min Rainfall").*

At the end of the first paragraph of sub-section "3.1 Evaluation of 30-min Rainfall" (now sub-section "3.2 Evaluation of 30-min Rainfall"), the sentence "The figure presents the two longest rainfall events for CML 14." was rephrased as ***"The figure shows that these three CMLs capture reasonably well two of the rainiest events of the studied period."****.*

The first half of the last paragraph of sub-section "3.1 Evaluation of 30-min Rainfall" (now sub-section "3.2 Evaluation of 30-min Rainfall") "The results for different R-k relations are quite similar, indicating that differences in DSD climatologies play a smaller role. For CML 12 and 13 the relative bias becomes less severe for the R-k relation derived from São Paulo data." was rephrased as ***"The presented results are based on the R-k relation derived from São Paulo data, which is representative for the local rainfall climatology. The results (not shown here) for the different R-k relations are quite similar (Sec. 2.3), which indicates that differences in DSD climatologies play a smaller role."****.*

At the end of sub-section "2.2 Data", the sentences "The DSD recorded by the Parsivels were corrected by the method of Raupach and Berne (2015a, b). We use updated correction factors trained from French disdrometer data. Due to instrumental, climatic, and location differences, these correction factors are taken as approximations." were removed from the manuscript.

Summary
15/04/2018 20:44:34

Differences exist between documents.

New Document:	Old Document:
04_paper_v08d3	04_paper_v07d2
22 pages (1.93 MB)	21 pages (1.82 MB)
15/04/2018 20:44:11	15/04/2018 20:44:11
Used to display results.	

Get started: first change is on page 1.

No pages were deleted

How to read this report

Highlight indicates a change.
~~**Deleted**~~ indicates deleted content.
indicates pages were changed.
indicates pages were moved.

Rainfall retrieval with commercial microwave links in São Paulo, Brazil

Manuel F. Rios Gaona[1], Aart Overeem[1,2], Timothy H. Raupach[3], Hidde Leijnse[2], and Remko Uijlenhoet[1]

[1]Hydrology and Quantitative Water Management Group, Department of Environmental Sciences, Wageningen University, 6708 PB Wageningen, the Netherlands.

[2]R&D Observations and Data Technology, Royal Netherlands Meteorological Institute, 3731 GA De Bilt, the Netherlands.

[3]Environmental Remote Sensing Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Correspondence to: Manuel F. Rios Gaona (feliperiosg@gmail.com)

Abstract. In the last decade there has been a growing interest from the hydrometeorological community regarding rainfall estimation from commercial microwave link (CML) networks. Path-averaged rainfall intensities can be retrieved from the signal attenuation between cell phone towers. Although this technique is still in development, it offers great opportunities for the retrieval of rainfall rates at high spatiotemporal resolutions very close to the Earth's surface. Rainfall measurements at high spatiotemporal resolutions are highly valued in urban hydrology, for instance, given the large impact that flash floods exert on society. Flash floods are triggered by intense rainfall events that develop over short time scales.

Here, we present one of the first evaluations of this measurement technique for a subtropical climate. Rainfall estimation for subtropical climates is highly relevant, since many countries with few surface rainfall observations are located in such areas. The test bed of the current study is the Brazilian city of São Paulo. The open-source algorithm RAINLINK was applied to retrieve rainfall intensities from (power) attenuation measurements. The performance of RAINLINK estimates was evaluated for 145 of the 213 CML in the São Paulo metropolitan area for which we received data, for 81 days between October 2014 and January 2015. We evaluated the retrieved rainfall intensities and accumulations from CML against those from a dense automatic gauge network. Results were found to be promising and encouraging when it comes to capturing the city-average rainfall dynamics. Mixed results were obtained for individual CML estimates, which may be related to erroneous metadata.

1 Introduction

Rainfall is the key input in environmental applications such as hydrological modeling, flash-flood and crop growth forecasting, landslide triggering, quantification of fresh water availability, and waterborne disease propagation. Because it is a natural process with a high spatiotemporal variability (Hou et al., 2008; Sene, 2013b), its accurate estimation is a demanding task.

The most common technologies that are currently used to measure rainfall at larger scales are rain gauges, radars and satellites. Each technology presents advantages and drawbacks with regard to the accuracy of rainfall estimates and the spatiotemporal coverage. Rain gauges directly measure the quantity of precipitation that falls on the ground. They offer accurate estimates of rainfall collected at temporal scales from minutes to days. Nevertheless, their rainfall estimates are only representative of their direct vicinity. In addition, in most cases the gauges within a network are unevenly distributed in space. Weather radar (RAdio

Detection And Ranging) offers indirect estimates of rainfall, with horizontal resolutions of $\sim$1 km (or even less depending on the radar settings) every $\sim$2 to $\sim$5 min. They scan distances of $\sim$100$-$300 km, which represent an area of $\sim$125,000 km^2, if issues of beam blockage are not present. The accuracy of rainfall estimates from radar depends on how well the measurements of backscattered power from hydrometeors are transformed into rain rates. Satellites offer also indirect estimates of rainfall at several spatiotemporal resolutions. For instance, Geostationary Earth Orbit (GEO) satellites (orbiting the Earth at $\sim$36,000 km) provide observations at resolutions of $\sim$10$-$60 min, and 1$-$4 km (Sene, 2013a; Wang, 2013), whereas Low Earth Orbit (LEO) satellites (orbiting the Earth at $\sim$800 km) can provide observations at resolutions of $\sim$1 km or less. Gridded rainfall products from the Global Precipitation Measurement mission (GPM) offer precipitation estimates between 60°N$-$60°S at a spatial resolution of 0.1°$\times$0.1° every 30 min. The main advantage of satellites above radars and gauges is that they provide global rainfall estimates (oceans included).

Commercial microwave links (CML) represent a technology that in the past decade has gained momentum as an alternative means for rainfall estimation. CML rainfall estimates are more representative of rainfall at the ground surface than those offered by satellites and/or weather radars. Networks of CML are more dense than gauge networks given their worldwide deployment for telecommunication purposes (Overeem et al., 2016b; Kidd et al., 2017). This worldwide spread of CML potentially offers rainfall estimates in places where rain gauges are scarce or poorly maintained, or where ground-based weather radars are not yet deployed or cannot be afforded. The spatiotemporal resolution of rainfall estimates from CML varies from seconds to minutes, and from hundreds of meters to tens of kilometers. For instance, Messer et al. (2012), and Overeem et al. (2016b) use maximum and minimum Received Signal Level (RSL) measurements over 15-min intervals, for CML with (spatial) densities of 0.3 to 3 links per km^2, and 0.1 to 2.1 km per km^2, respectively. Messer et al. (2012), and Fencl et al. (2015) provide 1-min rainfall estimates, whereas 1-s retrievals are obtained by Doumounia et al. (2014), and Chwala et al. (2016).

The interaction between attenuation and rainfall has long been studied by the electrical engineering community (from the attenuation perspective), and since the last two and a half decades by the hydrological community (from the rainfall perspective). Hogg (1968), and Crane (1971) review the influence of atmospheric phenomena on mm- and cm-wavelength based satellite communication systems. Later, Hogg and Chu (1975), and Crane (1977) focus exclusively on the role of rainfall in satellite communication, as rainfall is the major source of propagation issues for frequencies above 4$-$10 GHz. Recently, Chakravarty and Maitra (2010), and Badron et al. (2011) study rain-induced attenuation in satellite communication at tropical locations, where the attenuation is severe. Even more recently, Barthès and Mallet (2013), and Mercier et al. (2015) retrieve high-resolution rainfall fields (0.5$\times$0.5 km every 10 sec) from 10.7- and 12.7-GHz Earth-space links used in satellite TV transmission, even though at Ku band the estimation of weak rainfall rates is not optimal.

Our main interest here is rainfall estimation from terrestrial links. The idea of rain rate retrieval from attenuation measurements via tomographic techniques is presented by Giuli et al. (1991). Cuccoli et al. (2013), and D'Amico et al. (2016) present reconstructed 2D-rainfall fields from operational ML networks via tomographic techniques. Ruf et al. (1996) use a 35-GHz dual-polarization link for rainfall estimation at 0.1$-$1 km horizontal resolutions. Holt et al. (2000), Rahimi et al. (2004) and, Upton et al. (2005) estimate path-averaged rainfall from the differential attenuation of dual-frequency links. Minda and Nakamura (2005) use a 50-GHz link of 820 m to estimate rainfall. At such frequencies (or higher) rainfall estimation is sensitive to the raindrop size distribution and raindrop temperature. The synergistic use of ML, gauges and radars for rainfall estimation is proposed by Grum et al. (2005), and Bianchi et al. (2013). The first references to rainfall estimates from CML are Messer et al. (2006), and Leijnse et al. (2007). Berne and Uijlenhoet (2007), Leijnse et al. (2010), and Zinevich et al. (2010) study sources of uncertainty in rainfall estimates from CML. Methods for country-wide rainfall fields from CML are developed in Zinevich et al. (2008), and Overeem et al. (2013). Uijlenhoet et al. (2018) give a non-expert summary of the history, theory, challenges, and opportunities toward continental-scale rainfall monitoring via CMLs of cellular communication networks.

In the last decade the use of CML has broadened its spectrum to several other environmental applications beyond rainfall estimation, for instance, melting snow (Upton et al., 2007), water vapour monitoring (David et al., 2009), wind velocity estimation (Messer et al., 2012), dense-fog monitoring (David et al., 2013), urban drainage modelling (Fencl et al., 2013), flash flood early warning in Africa (Hoedjes et al., 2014), and air pollution detection (David and Gao, 2016).

We evaluate the performance of 145 CMLs located in the city of São Paulo, Brazil, in terms of their capacity to retrieve rainfall for the period between 20 October 2014 and 9 January 2015 (∼3 months). Rainfall evaluation against data from nearby gauges was coherently possible for 116 links from a network of 213 CMLs. Previously, da Silva Mello et al. (2002) studied the attenuation along ML due to rainfall for São Paulo. They used 6 links (7−43 km) with frequencies between 15 and 18 GHz. Here, we invert the problem by considering the attenuation suffered by such signals to be a valuable source of rainfall information instead of considering rainfall to be a nuisance for the propagation of radio signals. Since CMLs were not intended for rainfall estimation purposes, these devices can be considered a form of opportunistic sensors. They are potentially cost-free as the retrieved rain rates can be regarded as a by-product of power measurements.

As subtropical and tropical regions are the ones most deprived of radar (Heistermann et al., 2013) and gauge networks (Lorenz and Kunstmann, 2012; Kidd et al., 2017), CMLs could serve as complementary (or even alternative) networks for rainfall monitoring. Most of the recent studies concerning rainfall retrieval from CMLs have focused on temperate and Mediterranean climates, e.g., Messer and Sendik (2015); Overeem et al. (2016b). Thus, our evaluation is one of the first which focuses on a subtropical climate, complementing the study of Doumounia et al. (2014), which focused on a semi-arid climate. Focus on accurate rainfall estimation within the subtropics is of high relevance given that in such regions (e.g., São Paulo) intense events develop more often into flash floods and mud slides, which cause damage to property, disruption of business, and occasional casualties (Pereira Filho, 2012).

This paper is organized as follows: Section 2 describes the study area, the datasets (CML, rain gauges, disdrometers), the retrieval algorithm, and the evaluation metrics. The results and related discussion of our major findings are presented alongside in section 3. Summary, conclusions and recommendations are provided in section 4.

2 Study Area, Data and Methods

2.1 Description of Study Area

The city of São Paulo is located ∼60 km from the Atlantic Ocean at ∼770 masl, where sea breeze fronts commonly push from the SE against prevailing continental NW winds (cold fronts). In general, the incoming sea breeze interacts with the warmer and drier (urban) heat island of São Paulo, producing very deep convection with heavy rainfall, wind gusts, lightning and hail (Pereira Filho, 2012; Machado et al., 2014; Vemado and Pereira Filho, 2016). de Oliveira et al. (2002) characterize the local climate as typical of subtropical regions of Brazil, with a dry winter (June-August) and a wet summer (December-March). With regard to the climatology of São Paulo[1], February is the warmest month with 22.4° C, and July the coldest with 15.8° C. Climatological averages for temperature and humidity, for November and December (the full two months of the studied period), are 20.2 and 21.1°C, and 78% and 80%, respectively. August is the driest month with 39.6 mm of precipitation, and January the wettest with 237.4 mm, on average. The (climatological) yearly accumulated rainfall is 1,441 mm. Overeem et al. (2016b) report winter time issues in rainfall estimates from CML, i.e. solid and melting precipitation. However, for the subtropical climate of São Paulo, such winter issues are not expected to play a role, which is advantageous for accurate rainfall estimation.

2.2 Data

We received power measurements from two brands of CML: Ericsson (ER) and Huawei (HU). Power levels were registered every 15 min from 01:00 UTC 20 October 2014 to 00:45 UTC 8 January 2015, i.e., 81 days exactly. Their quantization level was 0.1 dB. Minimum and maximum levels of received and transmitted power were available for 66 HU CMLs, whereas only minimum received powers were available for ER 147 CMLs. The ER CMLs are assumed to have constant transmitted power levels. Figure 1 shows the location of these CMLs.

Figure 2 shows the scatter plot of link frequency against link length for all CMLs. In Fig. 2 the CMLs with uncommon or dubious (dub) combinations of length and frequency are denoted by gray markers (grey paths in Fig. 1). Our experience tells us that CMLs 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). The group of markers in the left bottom corner of Fig. 2 is also considered as dubious. Nevertheless, some CMLs around 7 GHz, having link lengths above 10 km, could still be realistic. We decided to only use the group of CMLs with path lengths shorter than 20 km and microwave frequencies above 15 GHz. Hence, 91 ER CMLs (40 link paths) and 54 HU CMLs (55 link paths) are left for the analyses, i.e., 145 CMLs in total (95 link paths). For RAINLINK to work, it is necessary that the power level of the transmitted signal is essentially constant.

Rainfall depths from 152 stations were retrieved from the National Early Warning and Monitoring Centre of Natural Disasters (CEMADEN), Brazil[2]. These 152 stations offer 10-min rainfall depths for the period and region under study (Fig. 1). A gauge validation procedure was necessary due to availability issues and doubts about the quality of the rainfall observations from the CEMADEN gauge network. The validation procedure is as follows: 1) For every gauge (152 in total) the closest two gauges were selected for comparison; 2) For the entire period, 30-min rainfall pairs (dry periods included) were evaluated through the relative bias and the coefficient of correlation for both closest gauges; 3) If the metrics of at least one of the two

[1]The climatological data presented here cover the period from 1961 to 1990 and correspond to the station "Mir. de Santana" located in the heart of São Paulo city (−46.6 lon, −23.5 lat, 792 masl). These data are freely available at the INMET (METeorological National Institute) web portal: http://www.inmet.gov.br/portal/index.php?r=home2/index.

[2]Gauge data from Brazil is freely available at http://www.cemaden.gov.br/mapainterativo/.

closest gauges are within ±25% for the relative bias, and $\geqslant 0.6$ for the correlation coefficient, the gauge under evaluation was deemed reliable. This selection results in 96 valid gauges out of 152. Comparisons of city-average rainfall were carried out among data from valid (96), and all (152) gauges, and all (145) CMLs (Fig. 4). For comparisons of individual path-averaged estimates of CMLs against gauges, only gauges within 1 or 9 km from the evaluated link paths were selected. For the 1-km case 35 CMLs were compared against 20 gauges, whereas for the 9-km case 116 CMLs were compared against 87 gauges.

Thanks to the CHUVA project (Machado et al., 2014), we retrieved 1-min drop size distributions (DSD) from three Parsivel disdrometers located in the region "Vale do Pariba", ∼93 km east of the study area[3]. These DSD data were collected from 1 November 2011 to 14 March 2012.

2.3 Rainfall Retrieval Algorithm

Rainfall estimation from CMLs is based on power measurements from the electromagnetic signal along a link path, i.e., between transmitter and receiver. Rainfall rates can be retrieved from the decrease in power, which is largely due to the attenuation of the electromagnetic signal by raindrops along the link path. The power-law relation between attenuation and rainfall (along a link path) was established by Atlas and Ulbrich (1977), and Olsen et al. (1978) as:

$$k = aR^b, \tag{1}$$

where k is the specific attenuation [dB·km^{-1}] along the link path attributed to rainfall and R is the rainfall rate [mm·h^{-1}]. The coefficient a and exponent b depend on the frequency and polarization of the electromagnetic signal, the DSD, and (to a much lesser extent) on the raindrop temperature. In the frequencies at which CML commonly operate, the exponent b in Eq. (1) is ~ 1.0. Atlas and Ulbrich (1977) state that the near-linearity between rain rates and specific attenuation (in the $20-40$ GHz band) "makes it possible to use the total path loss as a direct measure of $\overline{R}$ [average rain rate] independent of the form of the distribution of R [rain rate] along the path".

Both the degree to which Eq. (1) holds and the values of a and b are determined by the DSD. In order to study how strongly this relation deviates from other relations found in the literature, we determine values of a and b based on measured DSDs from the São Paulo region. For each 1-min DSD, we compute the corresponding rainfall intensity and specific attenuation of the common frequencies in São Paulo, i.e., from 8 to 23 GHz. Specific attenuation is computed for vertically polarized signals (most CML operate using this polarization) using T-Matrix scattering computations (e.g. Mishchenko, 2000), assuming raindrop oblateness as a function of its volume-equivalent diameter given by Andsager et al. (1999), and an average raindrop temperature of 298.36 K. The values of a and b are subsequently determined in log–log space (orthogonal regression) by a linear fit of $R = ak^b$ to the computed values of R and k. Note that conversely to Eq. (1) $\log(R)$ is the dependent variable in this case.

[3] DSD data from Parsivel and other disdrometers for the region of São Paulo (and other regions across Brazil) are freely available at http://chuvaproject.cptec.inpe.br/soschuva/

Figure 3 shows the power-law relations of 3 microwave frequencies (8, 15, and 23 GHz). This figure also shows the power-law relations derived for rainfall in the Netherlands (Leijnse, 2007, p. 65), and those recommended by the International Telecommunication Union (ITU-R Recommendation P.838-3). It is clear from this figure that there certainly are differences, and that such differences are largest for 8 GHz at high rainfall intensities. For the higher frequencies, such differences are more limited, especially at high rainfall intensities. This is in line with what has been found earlier (e.g. Berne and Uijlenhoet, 2007; Leijnse et al., 2008, 2010).

RAINLINK (Overeem et al., 2016a) is an R package (R Core Team, 2017) in which rain rates and area-wide rainfall maps can be derived from CML attenuation measurements. A very brief description of the algorithm is as follows: 1) wet-dry classification – a link is considered for non-zero rainfall retrievals if the received power jointly decreases with that of nearby links (9-km radius for this study); 2) reference signal level estimation – the median signal level of all dry periods in the previous 24 h is considered as the representative level of dry weather; 3) outlier removal – exclusion of links for which the specific attenuation (accumulated over 24 h) deviates too much from that of nearby links; 4) rainfall retrievals – once attenuation estimates are obtained from the difference between RSL and the reference signal level, a fixed wet antenna attenuation correction is applied (2.3 dB), and subsequently 15-min average rainfall intensities are computed from a weighted average of minimum and maximum rainfall intensities obtained by the (inverse) power law of Eq. (1); and 5) rainfall maps – rainfall intensities are interpolated into rainfall maps through Ordinary Kriging. This latter step was not implemented in this study. Overeem et al. (2016a) give a more detailed and in-depth review and description about all the technicalities within the RAINLINK package.

The ER CMLs only provide minimum power levels. RAINLINK is designed to retrieve rain rates from minimum and maximum power levels. Thus, in order for RAINLINK to compute mean rainfall estimates only from minimum power levels, two steps extra are required: 1) in the input file(s) for RAINLINK, the column with maximum power levels has to receive the values of the column with minimum power levels; 2) the mean path-averaged rainfall intensity, i.e. the output from RAINLINK, is now a maximum rainfall intensity and needs to be multiplied by a conversion factor to obtain the actual mean intensity. This conversion factor needs to be determined by means of a calibration dataset. Here, we use the 1-min rainfall intensities from the three disdrometers from the region of São Paulo to obtain an estimate of such a conversion factor. For each 15-min interval, the minimum rainfall intensity is selected from the lowest intensity of the 15 1-min intensities. 0.38 was found as the conversion factor, by comparing this minimum rainfall intensity against the mean 15-min rainfall intensity from the same disdrometers. ER-CML maximum rainfall intensities are then multiplied by this factor to obtain (actual) mean rainfall intensities.

The 1-min rainfall intensities from the three disdrometers from the region of São Paulo are also employed to estimate the value of α used to convert the minimum and maximum rainfall intensities from the HU CMLs to mean 15-min intensities. The found value, 0.30, is close to the default one in RAINLINK, 0.33, based on Dutch data and used in this study. This confirms the usefulness of the default value of α for application in a subtropical climate.

2.4 Error and Uncertainty Metrics

We evaluated the rainfall estimates from RAINLINK through: 1) the relative bias, 2) the coefficient of variation (CV), and 3) the coefficient of determination (r^2).

For a given CML (dataset), the relative bias is a relative measure of the average error between the RAINLINK estimates $R_{\text{RAINLINK},i}$ and the rain gauge measurements $R_{\text{gauge},i}$ (the latter considered as the ground truth):

$$\text{relative bias} = \frac{\overline{R}_{\text{res}}}{\overline{R}_{\text{gauge}}} = \frac{\sum_{i=1}^{n} R_{\text{res},i}}{\sum_{i=1}^{n} R_{\text{gauge},i}}, \quad (2)$$

where $R_{\text{res},i} = R_{\text{RAINLINK},i} - R_{\text{gauge},i}$ and n represents all possible time steps for the period under consideration. $R_{\text{res},i}$ are the residuals, i.e., the difference between $R_{\text{RAINLINK},i}$ and $R_{\text{gauge},i}$. $\overline{R}_{\text{res}}$ and $\overline{R}_{\text{gauge}}$ are the average of the residuals and gauge rainfall measurements (in mm), respectively. The relative bias ranges from -1 to $+\infty$, where 0 represents unbiased rainfall estimates.

The coefficient of variation is a dimensionless measure of dispersion (Haan, 1977), defined in this case as the standard deviation of the residuals $\sqrt{\widehat{\text{Var}}(R_{\text{res}})}$ divided by the mean of the rain gauge measurements, for the evaluated CML:

$$\text{CV} = \frac{\sqrt{\widehat{\text{Var}}(R_{\text{res}})}}{\overline{R}_{\text{gauge}}}. \quad (3)$$

The CV is a measure of uncertainty. It ranges from 0 (a hypothetical case with no uncertainty) to ∞.

The coefficient of determination is a measure of the strength of the linear dependence between two random variables, RAINLINK estimates and rain gauge measurements in this case. It is defined as the square of the correlation coefficient between $R_{\text{RAINLINK},i}$ and $R_{\text{gauge},i}$:

$$r^2 = \frac{\widehat{\text{Cov}}^2(R_{\text{gauge}}, R_{\text{RAINLINK}})}{\widehat{\text{Var}}(R_{\text{gauge}}) \cdot \widehat{\text{Var}}(R_{\text{RAINLINK}})}, \quad (4)$$

where $\widehat{\text{Var}}(R_{\text{gauge}})$ and $\widehat{\text{Var}}(R_{\text{RAINLINK}})$ are the variance of rain gauge measurements and RAINLINK estimates, respectively; and $\widehat{\text{Cov}}^2(R_{\text{gauge}}, R_{\text{RAINLINK}})$ the squared covariance between these two variables. r^2 ranges from 0 to 1, this latter the case of perfect linear correlation, i.e., all data points would fall on a straight line without any scatter. Perfect linearity does not imply unbiased estimates because the regression line does not have to coincide with the 1:1 line, even if it captures all variability.

The metrics were systematically computed on 30-min paired rainfall depths, using either all rainfall pairs or only pairs where both CML and gauge depths are above 0.0 mm. The latter to account only for significant rainfall events. 30-min aggregation was necessary given the temporal resolutions of the datasets, i.e., 10 min for gauge and 15 min for CML-retrieved data.

3 Results and Discussion

3.1 City-average Rainfall

For each dataset we compute the cumulative city-average rainfall for the studied period (Fig. 4). According to the reference, i.e., the 96 valid gauges, the cumulative rainfall depth in this ∼3-month period is ∼600 mm. The differences in cumulative rainfall depths between the valid and all (152) rain gauges are small. Such a small difference suggests that the gauge dataset is reliable. For the “PreProcessed” CML dataset no wet-dry classification and no outlier filter are applied. This contributes to cumulative rainfall depths being roughly twice as large as the gauge-based ones. Moreover, the dynamics do not often correspond with that of the gauges, for instance around 1 December 2014. For the “OutFiltered” dataset of 145 CMLs, which includes a wet-dry classification and outlier filter, a much better correspondence is found. The dynamics of the cumulative series agree reasonably well, and an overall underestimation is found, ∼200 mm at the end of the period, albeit much smaller than the difference between the “PreProcessed” dataset and the reference. The separate performance of the HU and ER CMLs shows that the HU dataset performs quite well with some overestimation, whereas the ER dataset gives a huge underestimation, despite roughly capturing the rainfall dynamics.

3.2 Evaluation of 30-min Rainfall

For the studied period, we evaluate the quality of 30-min path-averaged rainfall estimates from individual CMLs against gauges by: 1) time series from rainfall events for the three best performing CMLs; 2) scatter density plots based on data from all CMLs; and 3) metrics for each CML.

Figure 5 shows minimum and maximum received powers and the derived CML rainfall rates at 15-min resolution, as well as the rain rates from the nearest gauge (< 1 km) at 10-min resolution. The upscaled 30-min rainfall rates from both CML and gauges are also shown in Fig. 5. It can be seen that the minimum and maximum received powers are strongly negatively correlated with the gauge rainfall rates. The figure shows that these three CMLs capture reasonably well two of the rainiest events of the studied period. One can see that the stronger the rainfall event is, the larger is the attenuation registered by the CMLs.

Uncertainties in gauge and attenuation measurements themselves are the two sources of error that mainly constrain our evaluation. Our work compares CML rainfall estimates against rain gauge measurements, which are considered here as the “ground truth”. Nonetheless, a gauge is only representative of its surrounding area and does not account for the spatial variability of rainfall along the link path. Representativeness errors will increase for longer link paths and for more intense rainfall events. For subtropical regions where intense rainfall is associated with small convective raincells, da Silva Mello et al. (2014) showed that due to smaller raincells only a part of the link-path contributes to the attenuation, which causes an effective link-rain rate smaller than the one(s) measured by gauges.

The results of Fig. 5 are obtained for short links (< 1.7 km), where representativeness errors will play a smaller role. Overestimations by CMLs may be related to the fact that rain-induced attenuation along the link path may be relatively small compared to the attenuation caused by wet antennas, i.e., the wet antennas could contribute to some of the overestimations.

Figure 6 shows an overall assessment of the CML performance to retrieve 30-min rainfall depths (over the studied period). Scatter density plots are for CML-gauge pairs within 1 km (top panels, a and b) and within 9 km (bottom panels, c and d). The left column (panels a and c) is for all CML-gauge pairs, whereas the right column (panels b and d) only includes rainy intervals, i.e., CML-gauge pairs where both rainfall depths are above 0.0 mm. The rainfall estimates for CML-gauge pairs within 1 km are somewhat better than the ones for 9 km in terms of r^2 and CV, but the relative bias of the latter is smaller than that of the former. If all CML-gauge pairs are used, on average CMLs underestimate rainfall by 23 – 29%, with high values for CV and low values for r^2. Assuming that the gauges provide reliable measurements, this performance indicates that the applied wet-dry classification could be sub-optimal. Perhaps a sensitivity analysis of the threshold values in the wet-dry classification could improve this classification. If only rainy intervals are used, i.e., CML-gauge pairs both above 0.0 mm, these lead to a strong reduction in the value of CV, a decrease in the r^2, and a much smaller relative bias.

A reason for the large discrepancies among the statistics of the scatter density plots (Fig. 6) could be the fact that only minimum (and also maximum for HU CMLs) RSL data is used to compute 15-min rainfall intensities, i.e., a limited temporal sampling. Rios Gaona et al. (2015) compare CML (actual) and gauge-adjusted (simulated) path-average rainfall depths for a 12-day dataset from the Netherlands, based on rainfall pairs for which at least one depth exceeds 0.1 mm. The most prominent difference is their much higher value for r^2 (0.437), which was found for 15-min rainfall. Hence, the sampling strategy is not necessarily the main reason for the low values of r^2. Given the erroneous metadata found in the CML dataset (Sec. 2.2), which led to discarding CMLs with dubious combinations of path length and frequency, there could be errors in the metadata from selected CMLs too, i.e, wrong location of one of the antennas or wrong frequency. In addition, although a basic assessment of gauge quality has been performed, even records from gauges classified as valid could still contain measurement errors.

The presented results are based on the $R-k$ relation derived from São Paulo data, which is representative for the local rainfall climatology. The results (not shown here) for the different $R-k$ relations are quite similar (Sec. 2.3), which indicates that differences in DSD climatologies play a smaller role. In general, local parameters (i.e., SP) are the best approach for RAINLINK. Nevertheless, the RAINLINK default parameters offer (subtropical, São Paulo) CML estimates of reasonable quality despite the local (temperate) climate for which they were obtained, namely the Netherlands. The ITU parameters often lead to a much higher value of CV, and always to a larger overestimation.

Figure 7 shows the performance of individual CMLs by plotting the values of CV against r^2, based on CML-gauge pairs both above 0.0 mm (for the studied period). Many CMLs have fairly high values of r^2. For instance, 43% of the CMLs have a value of r^2 larger than 0.5 (for CML-gauge pairs within 9 km). Here, CML and gauge measurements are totally independent. Thus, it is very likely that the high values of r^2 for a large minority of CMLs indicate that both types of observations contain a true rain signal.

4 Summary, Conclusions and Recommendations

CML networks are an opportunistic technique for rainfall estimation, with the potential to be used worldwide given the spread of CML-based telecommunication systems during the last two decades. Here we presented one of the first evaluations of CML rainfall retrievals for a subtropical climate. Subtropical regions could benefit from this technique given that rainfall events are often more extreme, and usually fewer surface rainfall measurements are collected. We evaluated rainfall retrievals from power measurements for 145 CML from a network located in the city of São Paulo. We used RAINLINK (Overeem et al., 2016a) to retrieve rainfall from these CML.

30-min rainfall estimates from CMLs were evaluated against rainfall measurements from rain gauges for the period from 20 October 2014 to 9 January 2015. Despite the mixed results, the potential of CML technology for rainfall estimation in subtropical climates is confirmed. Especially, given the rainfall dynamics captured by the city-average rainfall (Fig. 4), the good performance of some individual CMLs (Fig. 5), and a high correlation for a large minority of CMLs (Fig. 7). This gives an indication that the RAINLINK package is suitable to retrieve rainfall via CML data from a subtropical climate, even though many of its parameters have not been optimized for such a climate. Since biases propagate in hydrological model predictions, given the low relative bias found for rainy periods (Fig. 6), CML rainfall estimates could be considered as an alternative input in hydrological models.

In order for RAINLINK to capture the rainfall characteristics from the region of São Paulo, we derived a-b coefficients of power-law $R-k$ relations from local DSD data. The a and b coefficients are a function of the polarization and frequency of the link, DSD and raindrop temperature. These local DSD parameters gave the best results, whereas the ITU/NL parameters proved to be very useful and accurate enough when local a-b coefficients cannot be derived. The NL parameters are characteristic for the hydroclimatology of the Netherlands, and are the default set in RAINLINK. They also outperform the ITU parameters.

A more thorough evaluation could be done to study and explain differences between CML and gauge rainfall estimates. For instance, the influence of rainfall variability along link paths could be studied. This can be achieved if local radar measurements are compared against CML estimates, which would allow to better track the rain events and their incidence over the link paths, especially relevant for longer link paths. We did not evaluate the performance of CML-RAINLINK retrievals based on rain rate classes. Nevertheless, this evaluation is highly encouraged as it would shed some light on the suitability of CMLs for hydrological applications, for instance.

We also encourage future work on sensitivity analyses focused on the optimization of RAINLINK parameters to improve the accuracy of rainfall estimates in subtropical regions. Note that the value of α, estimated from local 1-minute disdrometer rainfall intensities, was close to the default value from RAINLINK. Especially the value of A_a and the threshold values for the wet-dry classification and the outlier filter should be investigated. Missing maximum signal level data, and unexpected combinations of link lengths and microwave frequencies, forced us to remove many CMLs from the original dataset. This shows that accurate metadata, such as link coordinates for instance, are essential, as well as the feedback about obtained CML and reference datasets.

CMLs are not the replacement of current standard technologies such as radars, rain gauges (and even satellites), but their opportunistic use is rather valuable as complementary networks for high-resolution rainfall estimation. To conclude, we were able to obtain good results for a minority of CMLs, which confirms the great potential of this technique if the data and metadata are properly stored.

Author contributions. M.F. Rios Gaona sorted, analysed and plotted the data, and wrote most of the paper. T. Raupach processed the Parsivel DSD data. A. Overeem, H. Leijnse, and R. Uijlenhoet analysed the results, gave valuable feedback, and wrote parts of the paper.

Competing interests. The authors manifest not to have competing interests with the Planetary Skin Institute / Italia Mobile, which provided us the CML data.

Acknowledgements. We thank Juan Carlos Castilla-Rubio (from Planetary Skin Institute); the Brazilian Ministry of Science, Technology and Innovation; and Italia Mobile for providing the CML data. This work was financially supported by The Netherlands Organisation for Scientific Research NWO (project ALW-GO-AO/11-15).

References

Andsager, K., Beard, K. V., and Laird, N. F.: Laboratory Measurements of Axis Ratios for Large Raindrops, J Atmos Sci, 56, 2673–2683, https://doi.org/10.1175/1520-0469(1999)056<2673:LMOARF>2.0.CO;2, 1999.

Atlas, D. and Ulbrich, C. W.: Path- and Area-Integrated Rainfall Measurement by Microwave Attenuation in the 1–3 cm Band, J Appl Meteorol, 16, 1322–1331, https://doi.org/10.1175/1520-0450(1977)016<1322:PAAIRM>2.0.CO;2, 1977.

Badron, K., Ismail, A. F., Din, J., and Tharek, A.: Rain induced attenuation studies for V–band satellite communication in tropical region, J Atmos Sol-terr Phy, 73, 601–610, https://doi.org/10.1016/j.jastp.2010.12.006, 2011.

Barthès, L. and Mallet, C.: Rainfall measurement from the opportunistic use of an Earth–space link in the Ku band, Atmos Meas Tech, 6, 2181–2193, https://doi.org/10.5194/amt-6-2181-2013, 2013.

Berne, A. and Uijlenhoet, R.: Path–averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability, Geophys Res Lett, 34, https://doi.org/10.1029/2007GL029409, 2007.

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

Chakravarty, K. and Maitra, A.: Rain attenuation studies over an earth–space path at a tropical location, J Atmos Sol-terr Phy, 72, 135–138, https://doi.org/10.1016/j.jastp.2009.10.018, 2010.

Chwala, C., Keis, F., and Kunstmann, H.: Real-time data acquisition of commercial microwave link networks for hydrometeorological applications, Atmos Meas Tech, 9, 991–999, https://doi.org/10.5194/amt-9-991-2016, 2016.

Crane, R. K.: Propagation phenomena affecting satellite communication systems operating in the centimeter and millimeter wavelength bands, P IEEE, 59, 173–188, https://doi.org/10.1109/PROC.1971.8123, 1971.

Crane, R. K.: Prediction of the Effects of Rain on Satellite Communication Systems, P IEEE, 65, 456–474, https://doi.org/10.1109/PROC.1977.10498, 1977.

Cuccoli, F., Baldini, L., Facheris, L., Gori, S., and Gorgucci, E.: Tomography applied to radiobase network for real time estimation of the rainfall rate fields, Atmos Res, 119, 62–69, https://doi.org/10.1016/j.atmosres.2011.06.024, 2013.

da Silva Mello, L., Pontes, M. S., Fagundes, I., Almeida, M. P. C., and Andrade, F. J. A.: Modified rain attenuation prediction method considering the effect of wind direction, J Microw Optoelectron Electromagn Appl, 13, 254–267, https://doi.org/10.1590/S2179-10742014000200012, 2014.

da Silva Mello, L. A. R., Costa, E., and de Souza, R. S. L.: Rain attenuation measurements at 15 and 18 GHz, Electron Lett, 38, 197–198, https://doi.org/10.1049/el:20020105, 2002.

D'Amico, M., Manzoni, A., and Solazzi, G. L.: Use of Operational Microwave Link Measurements for the Tomographic Reconstruction of 2-D Maps of Accumulated Rainfall, IEEE Geosci Remote Sens Lett, 13, 1827–1831, https://doi.org/10.1109/LGRS.2016.2614326, 2016.

David, N. and Gao, H. O.: Using Cellular Communication Networks To Detect Air Pollution, Environ Sci Technol, 50, 9442–9451, https://doi.org/10.1021/acs.est.6b00681, 2016.

David, N., Alpert, P., and Messer, H.: Technical Note: Novel method for water vapour monitoring using wireless communication networks measurements, Atmos Chem Phys, 9, 2413–2418, https://doi.org/10.5194/acp-9-2413-2009, 2009.

David, N., Alpert, P., and Messer, H.: The potential of commercial microwave networks to monitor dense fog–feasibility study, J Geophys Res, 118, 11,750–11,761, https://doi.org/10.1002/2013JD020346, 2013.

de Oliveira, A. P., Machado, A. J., Escobedo, J. F., and Soares, J.: Diurnal evolution of solar radiation at the surface in the city of São Paulo: seasonal variation and modeling, Theor Appl Climatol, 71, 231–249, https://doi.org/10.1007/s007040200007, 2002.

Doumounia, A., Gosset, M., Cazenave, F., Kacou, M., and Zougmoré, F.: Rainfall monitoring based on microwave links from cellular telecommunication networks: First results from a West African test bed, Geophys Res Lett, 41, 6016–6022, https://doi.org/10.1002/2014GL060724, 2014.

Fencl, M., Rieckermann, J., Schleiss, M., Stránský, D., and Bareš, V.: Assessing the potential of using telecommunication microwave links in urban drainage modelling, Water Sci Technol, 68, 1810–1818, https://doi.org/10.2166/wst.2013.429, 2013.

Fencl, M., Rieckermann, J., Sýkora, P., Stránský, D., and Bareš, V.: Commercial microwave links instead of rain gauges: fiction or reality?, Water Sci Technol, 71, 31–37, https://doi.org/10.2166/wst.2014.466, 2015.

Giuli, D., Toccafondi, A., Gentili, G. B., and Freni, A.: Tomographic Reconstruction of Rainfall Fields through Microwave Attenuation Measurements, J Appl Meteorol, 30, 1323–1340, https://doi.org/10.1175/1520-0450(1991)030<1323:TRORFT>2.0.CO;2, 1991.

Google Maps: São Paulo, SP, http://maps.googleapis.com/maps/api/staticmap?center=-23.513238,-46.632861&zoom=10&size=640x640&scale=2&maptype=terrain&style=feature:road|element:all|visibility:off&style=feature:administrative%7Celement:labels%7Cvisibility:off&style=feature:poi.park%7Celement:all%7Cvisibility:off&sensor=false, (accessed Aug 5, 2017), 2017.

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

Haan, C. T.: Statistical Methods in Hydrology, Iowa State University Press, Ames, USA, 1st edn., iSBN 8-8138-1510-X, 1977.

Heistermann, M., Jacobi, S., and Pfaff, T.: Technical Note: An open source library for processing weather radar data (wradlib), Hydrol Earth Syst Sci, 17, 863–871, https://doi.org/10.5194/hess-17-863-2013, 2013.

Hoedjes, J. C. B., Kooiman, A., Maathuis, B. H. P., Said, M. Y., Becht, R., Limo, A., Mumo, M., Nduhiu-Mathenge, J., Shaka, A., and Su, B.: A Conceptual Flash Flood Early Warning System for Africa, Based on Terrestrial Microwave Links and Flash Flood Guidance, ISPRS Int J Geo–Inf, 3, 584–598, https://doi.org/10.3390/ijgi3020584, 2014.

Hogg, D. C.: Millimeter-Wave Communication through the Atmosphere, Science, 159, pp. 39–46, http://www.jstor.org/stable/1723412, 1968.

Hogg, D. C. and Chu, T.-S.: The role of rain in satellite communications, P IEEE, 63, 1308–1331, https://doi.org/10.1109/PROC.1975.9940, 1975.

Holt, A. R., Goddard, J. W. F., Upton, G. J. G., Willis, M. J., Rahimi, A. R., Baxter, P. D., and Collier, C. G.: Measurement of rainfall by dual–wavelength microwave attenuation, Electron Lett, 36, 2099–2101, https://doi.org/10.1049/el:20001468, 2000.

Hou, A. Y., Skofronick-Jackson, G., Kummerow, C. D., and Shepherd, J. M.: Precipitation: Advances in Measurement, Estimation and Prediction, chap. Global precipitation measurement, pp. 131–169, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-77655-0_6, 2008.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D. B.: So, How Much of the Earth's Surface Is Covered by Rain Gauges?, Bull Amer Meteor Soc, 98, 69–78, https://doi.org/10.1175/BAMS-D-14-00283.1, http://dx.doi.org/10.1175/BAMS-D-14-00283.1, 2017.

Leijnse, H.: Hydrometeorological application of microwave links : measurement of evaporation and precipitation, Ph.D. thesis, Wageningen, the Netherlands, http://edepot.wur.nl/121940, Wageningen University thesis nr. 4353, 2007.

Leijnse, H., Uijlenhoet, R., and Stricker, J. N. M.: Rainfall measurement using radio links from cellular communication networks, Water Resour Res, 43, WR005631, https://doi.org/10.1029/2006WR005631, 2007.

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

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

Lorenz, C. and Kunstmann, H.: The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis, J Hydrometeor, 13, 1397–1420, https://doi.org/10.1175/JHM-D-11-088.1, 2012.

Machado, L. A. T., Silva Dias, M. A. F., Morales, C., Fisch, G., Vila, D., Albrecht, R., Goodman, S. J., Calheiros, A. J. P., Biscaro, T., Kummerow, C., Cohen, J., Fitzjarrald, D., Nascimento, E. L., Sakamoto, M. S., Cunningham, C., Chaboureau, J.-P., Petersen, W. A., Adams, D. K., Baldini, L., Angelis, C. F., Sapucci, L. F., Salio, P., Barbosa, H. M. J., Landulfo, E., Souza, R. A. F., Blakeslee, R. J., Bailey, J., Freitas, S., Lima, W. F. A., and Tokay, A.: The Chuva Project: How Does Convection Vary across Brazil?, Bull Amer Meteor Soc, 95, 1365–1380, https://doi.org/10.1175/BAMS-D-13-00084.1, 2014.

Mercier, F., Barthès, L., and Mallet, C.: Estimation of Finescale Rainfall Fields Using Broadcast TV Satellite Links and a 4DVAR Assimilation Method, J Atmos Oceanic Technol, 32, 1709–1728, https://doi.org/10.1175/JTECH-D-14-00125.1, 2015.

Messer, H. and Sendik, O.: A New Approach to Precipitation Monitoring: A critical survey of existing technologies and challenges, IEEE Signal Proc Mag, 32, 110–122, https://doi.org/10.1109/MSP.2014.2309705, 2015.

Messer, H., Zinevich, A., and Alpert, P.: Environmental Monitoring by Wireless Communication Networks, Science, 312, 713, https://doi.org/10.1126/science.1120034, 2006.

Messer, H., Zinevich, A., and Alpert, P.: Environmental sensor networks using existing wireless communication systems for rainfall and wind velocity measurements, IEEE Instru Meas Mag, 15, 32–38, https://doi.org/10.1109/MIM.2012.6174577, 2012.

Minda, H. and Nakamura, K.: High Temporal Resolution Path–Average Rain Gauge with 50–GHz Band Microwave, J Atmos Oceanic Technol, 22, 165–179, https://doi.org/10.1175/JTECH-1683.1, 2005.

Mishchenko, M. I.: Calculation of the amplitude matrix for a nonspherical particle in a fixed orientation, Appl Opt, 39, 1026–1031, https://doi.org/10.1364/AO.39.001026, http://ao.osa.org/abstract.cfm?URI=ao-39-6-1026, 2000.

Olsen, R., Rogers, D., and Hodge, D.: The aR^b Relation in the Calculation of Rain Attenuation, IEEE T Antenn Propag, 26, 318–329, https://doi.org/10.1109/TAP.1978.1141845, 1978.

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Country-wide rainfall maps from cellular communication networks, P Natl Acad Sci Usa, 110, 2741–2745, https://doi.org/10.1073/pnas.1217961110, 2013.

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

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Two and a half years of country-wide rainfall maps using radio links from commercial cellular telecommunication networks, Water Resour Res, 52, 8039–8065, https://doi.org/10.1002/2016WR019412, 2016b.

Pereira Filho, A. J.: A mobile X-POL weather radar for hydrometeorological applications in the metropolitan area of São Paulo, Brazil, Geosci Instrum Method Data Syst, 1, 169–183, https://doi.org/10.5194/gi-1-169-2012, 2012.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2017.

Rahimi, A., Upton, G., and Holt, A.: Dual–frequency links–a complement to gauges and radar for the measurement of rain, J Hydrol, 288, 3–12, https://doi.org/10.1016/j.jhydrol.2003.11.008, 2004.

Rios Gaona, M. F., Overeem, A., Leijnse, H., and Uijlenhoet, R.: Measurement and interpolation uncertainties in rainfall maps from cellular communication networks, Hydrol Earth Syst Sci, 19, 3571–3584, https://doi.org/10.5194/hess-19-3571-2015, 2015.

Ruf, C. S., Aydin, K., Mathur, S., and Bobak, J. P.: 35–GHz Dual–Polarization Propagation Link for Rain–Rate Estimation, J Atmos Oceanic Technol, 13, 419–425, https://doi.org/10.1175/1520-0426(1996)013<0419:GDPPLF>2.0.CO;2, 1996.

Sene, K.: Flash Floods: Forecasting and Warning, chap. Precipitation Measurement, pp. 33–70, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-007-5164-4_2, 2013a.

Sene, K.: Flash Floods: Forecasting and Warning, chap. Rainfall Forecasting, pp. 101–132, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-007-5164-4_4, 2013b.

Uijlenhoet, R., Overeem, A., and Leijnse, H.: Opportunistic remote sensing of rainfall using microwave links from cellular communication networks, WIREs Water, https://doi.org/10.1002/wat2.1289, 2018.

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

Upton, G. J. G., Cummings, R. J., and Holt, A. R.: Identification of melting snow using data from dual–frequency microwave links, IET MICROW ANTENNA P, 1, 282–288, https://doi.org/10.1049/iet-map:20050285, 2007.

Vemado, F. and Pereira Filho, A. J.: Severe Weather Caused by Heat Island and Sea Breeze Effects in the Metropolitan Area of São Paulo, Brazil, Advances in Meteorology, 2016, 13 pages, https://doi.org/10.1155/2016/8364134, Article ID 8364134, 2016.

Wang, P. K.: Physics and Dynamics of Clouds and Precipitation, Cambridge University Press, https://doi.org/10.1017/CBO9780511794285, Cambridge Books Online, 2013.

Zinevich, A., Alpert, P., and Messer, H.: Estimation of rainfall fields using commercial microwave communication networks of variable density, Adv Water Resour, 31, 1470–1480, https://doi.org/10.1016/j.advwatres.2008.03.003, 2008.

Zinevich, A., Messer, H., and Alpert, P.: Prediction of rainfall intensity measurement errors using commercial microwave communication links, Atmos Meas Tech, 3, 1385–1402, https://doi.org/10.5194/amt-3-1385-2010, 2010.

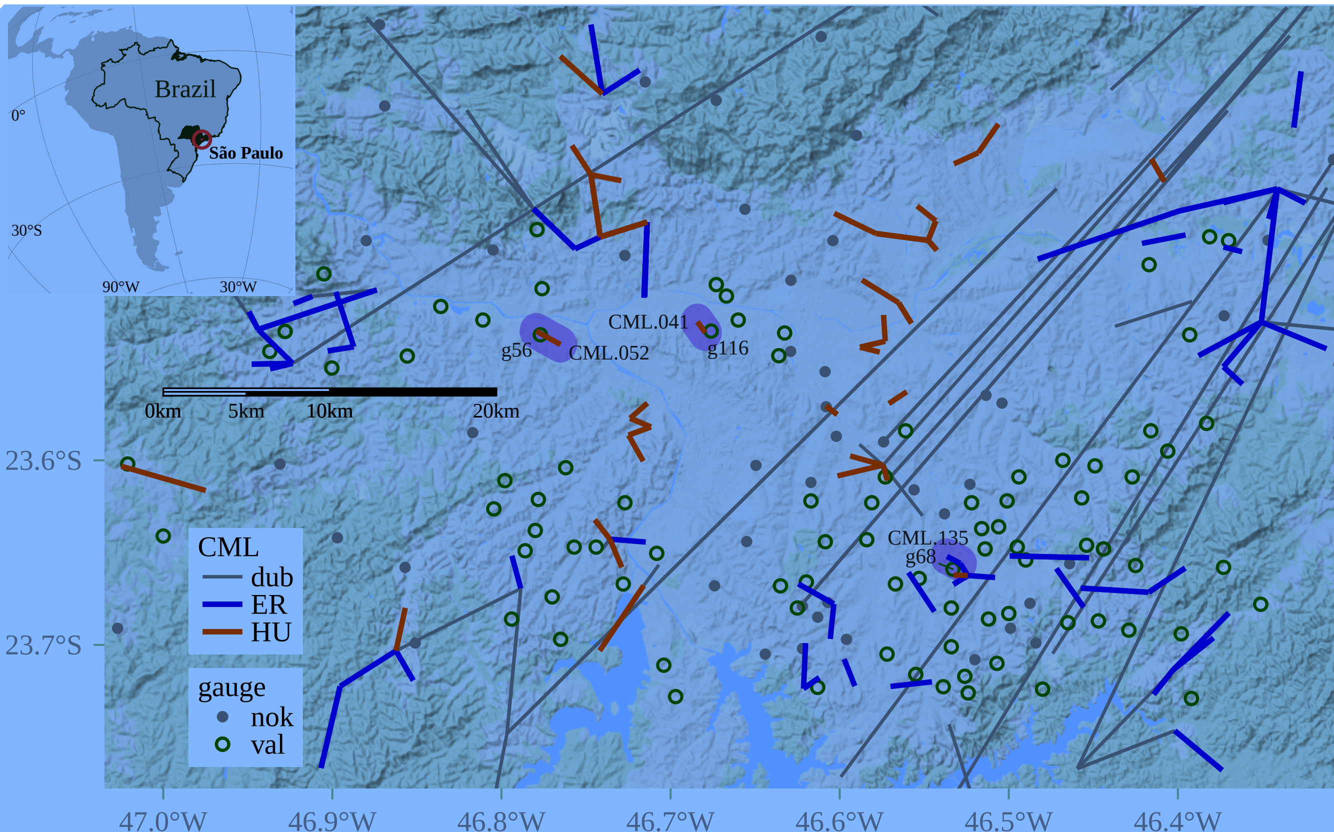


Figure 1. Topology of one CML network in the city of São Paulo, Brazil. 54 Huawei (HU; orange lines) CMLs (40 link paths), and 91 Ericsson (ER; blue lines) CMLs (55 link paths) are shown. The grey link paths (dub) are the 68 CMLs (HU and ER) with frequencies below 15 GHz and link-lengths above 20 km. Such CMLs are not analyzed here due to their dubious configuration. CMLs with frequencies above 15 GHz and link-lengths below 20 km (blue and orange link paths) suggest a very likely power level assignment. The circles are 152 CEMADEN gauges with 10-min resolution available for the studied period (20 October 2014 to 9 January 2015). The 96 green circles (val) represent the valid gauges. A gauge is deemed valid if its coefficient of correlation (r^2) is larger than 0.6 and its relative bias (rB) is lower than $\pm 25\%$ for at least one of the two closest gauges for which it is compared against (see Sec. 2.2). The dots in grey (nok) are the gauges that do not satisfy such thresholds. The 3 CMLs surrounded by a purplish shadow are those CML for which $r^2 \geqslant 0.6$ and rB $\leqslant \pm 25\%$ against their respective closest gauge (see also Fig. 5). CML data was provided by the Planetary Skin Institute / Italia Mobile[4]. The geographical location of São Paulo is given in the upper left corner. The DEM was extracted from Google Maps (Google Maps, 2017).

4

[4] We received CML data from a third party. It was not possible to verify on-site the topology of the network shown in Fig. 1, which we suspect not to be accurate given the orientation of the long links.

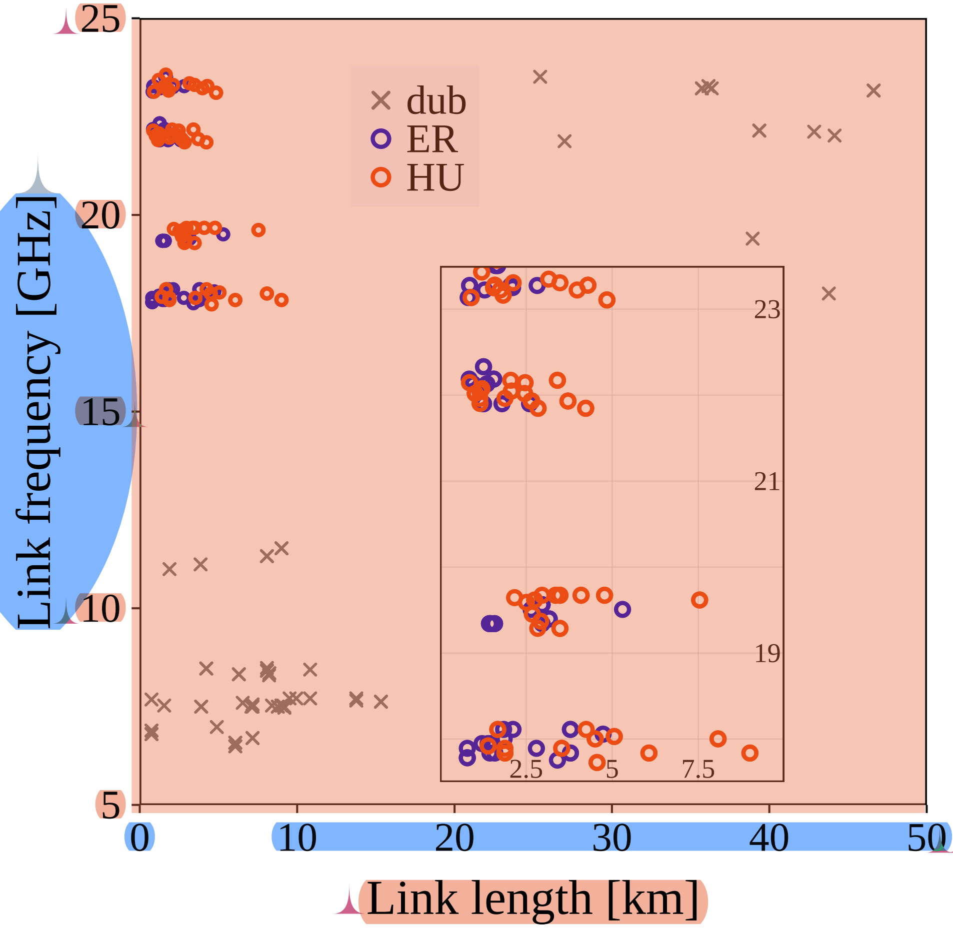


Figure 2. Scatter plot of frequency against link length for all 213 CMLs (149 ER, and 66 HU) shown in Fig. 1. The orange circles are the 54 HU CMLs, the blue circles are the 91 ER CMLs, and the grey markers are those (68) CMLs with frequencies below 15 GHz or link-lengths above 20 km. Inset, there is a zoom into the not-dubious region of frequency vs. link-length commonly found in commercial link networks worldwide.

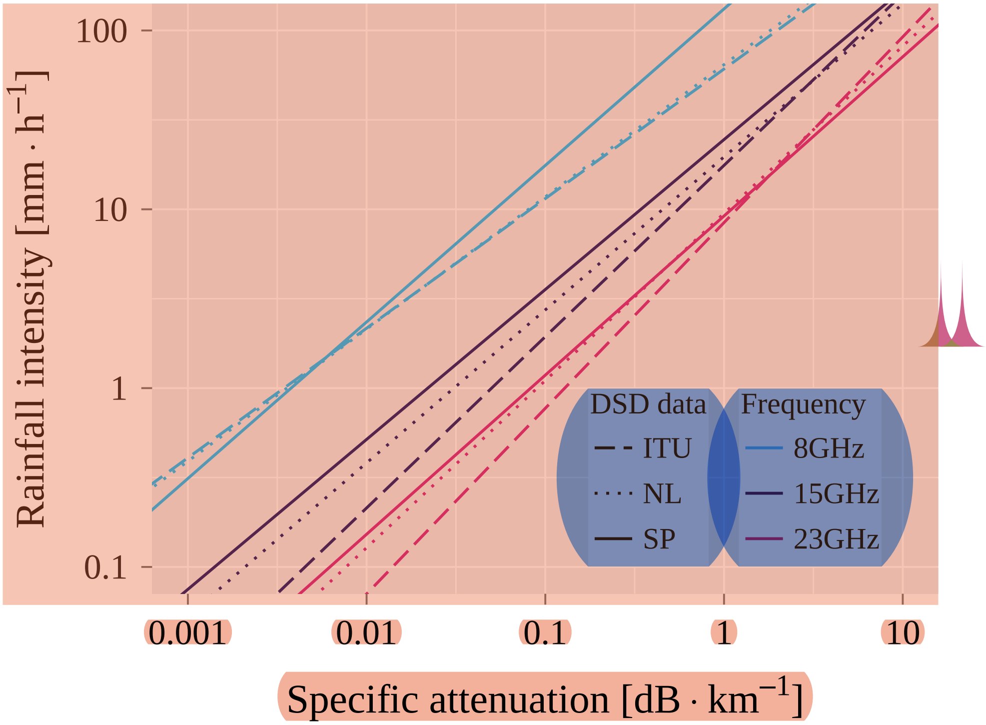


Figure 3. Rainfall intensity against specific attenuation for the a and b parameters of 3 DSD models, i.e, local (SP - continuous line), suggested by ITU-R Recommendation P.838-3 (ITU - dashed line), and RAINLINK's default (NL - dotted line). The $R-k$ relations are presented for 3 frequencies: 8 (cyan), 15 (blue), and 23 GHz (pink).

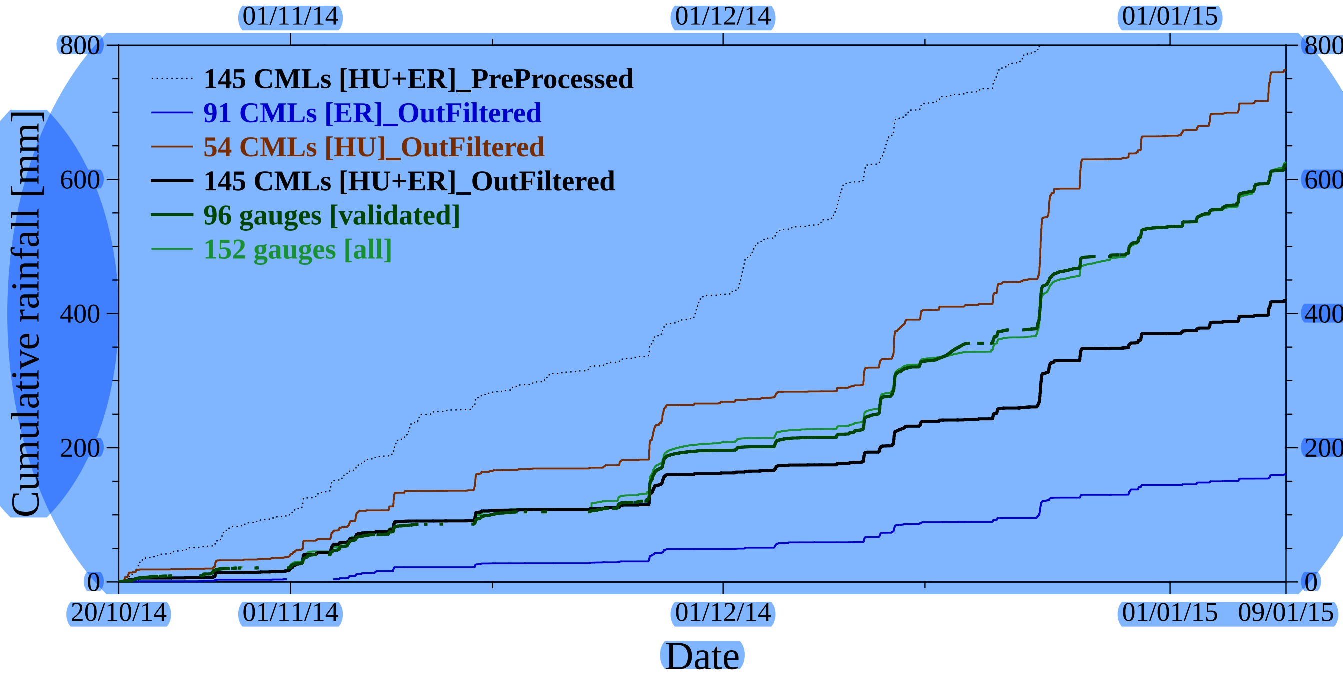


Figure 4. Cumulative time series of 30-min rainfall averaged over the city of São Paulo, Brazil (see Fig. 1) for the period between 20 October 2014 to 9 January 2015. The dotted black line is for the "PreProcessed" RAINLINK approach, i.e., without wet-dry classification and outlier filter; whereas the continuous black line is for the "OutFiltered" RAINLINK approach (with wet-dry classification and outlier filter). Both results (black lines) are obtained from the joint retrieval of Huawei and Ericsson CMLs ([HU + ER]). The blue line is for the ER CMLs only; whereas the orange line is for the HU CMLs. The dark green line is for only the valid gauges (a gauge is deemed valid if its coefficient of correlation (r^2) is larger than 0.6 and its relative bias is within $\pm 25\%$ for at least one of the two closest gauges for which it is compared against (see Sec. 2.2)). The light green line is for all gauges in the CEMADEN network (in the vicinity of São Paulo). In the legend, the numbers indicate the amount of devices (gauges or CML) used in the average. The blank spaces in the cumulative series indicate that no average was possible for that particular time step. It was assumed that no rainfall occurred in such blank spaces; therefore, the curve resumes with its immediate previous value.

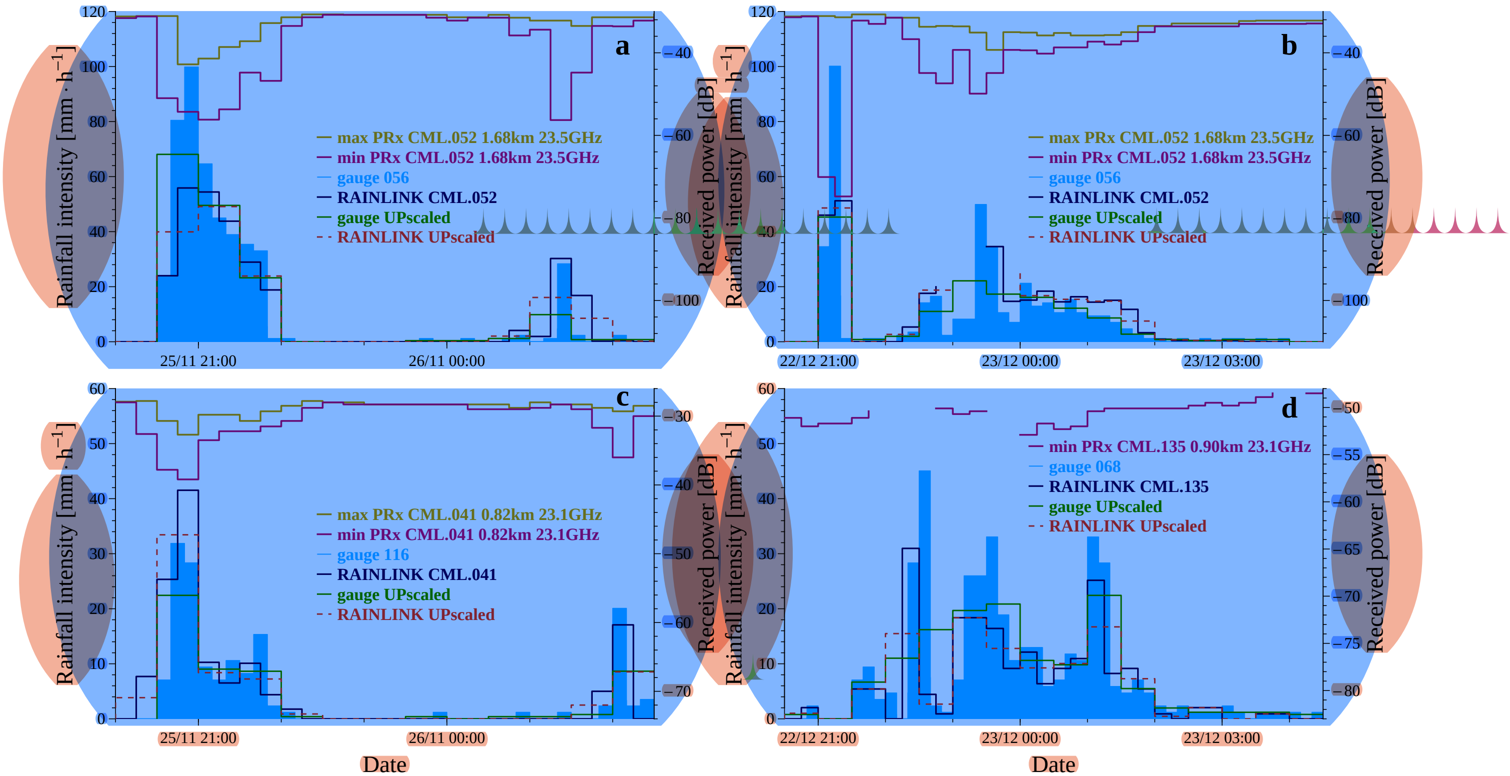


Figure 5. Time series of two rainfall events (November and December 2014, panels a and c, and b and d, respectively) for the three best performing CML, i.e., CML 052 (panels a and b), CML 041 (panel c), and CML 135 (panel d). Their evaluation is done against gauges 56, 116, and 68. CML 135 is an ER link, whereas CML 052 and 041 are HU links (see Fig. 1). Cyan is for 10-min gauge measurements, blue is for 15-min RAINLINK estimates, green is for 30-min upscaled gauges, red is for 30-min upscaled RAINLINK, and pink and gold are for 15-min minimum and maximum received powers, respectively. ER CML only sampled minimum received power. The RAINLINK series are computed for the local DSD parameters (Fig. 3, SP $R-k$ relation).

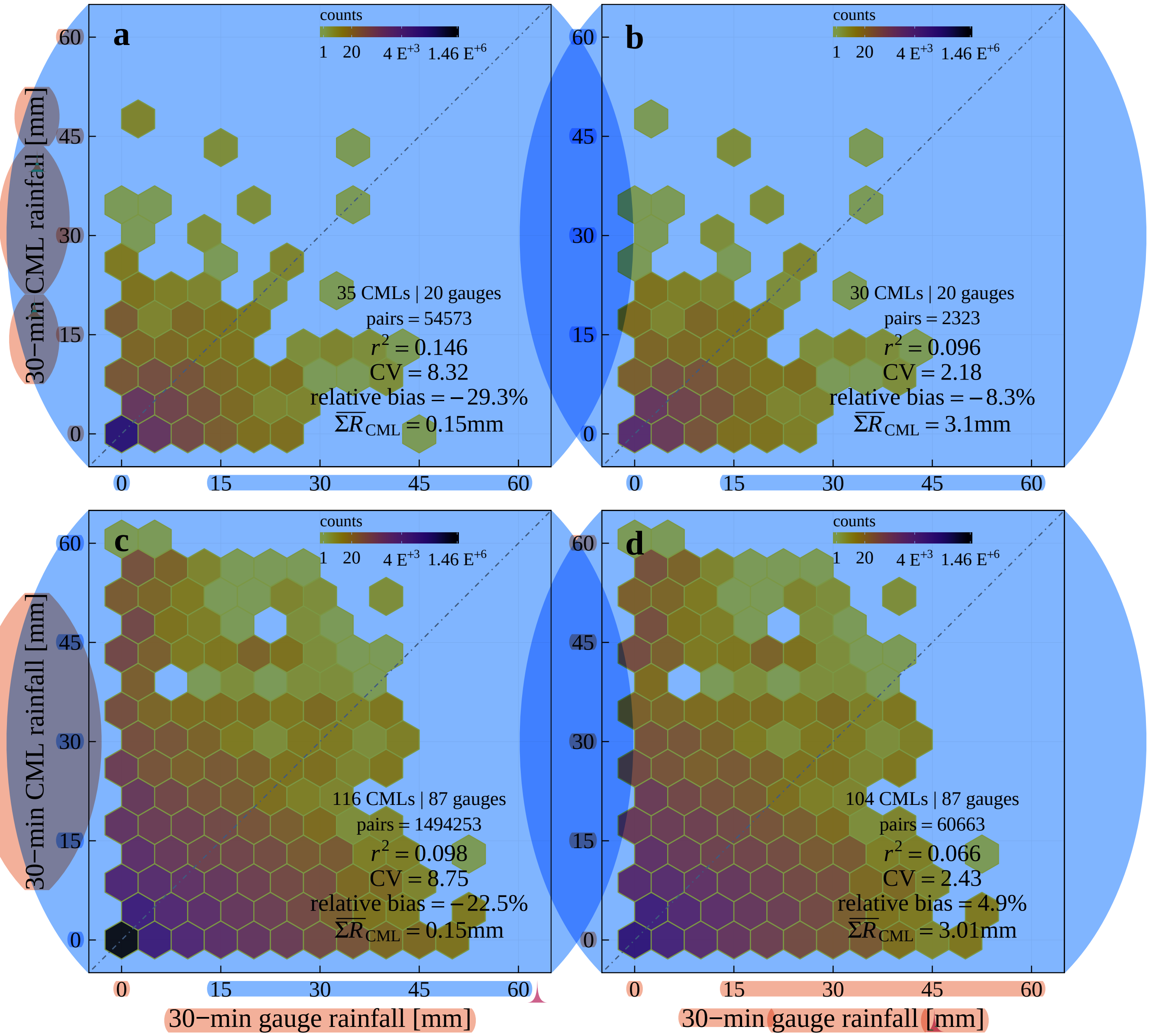


Figure 6. Scatter density plots of half-hourly CML rainfall depths vs. gauge rainfall depths from 20 October 2014 to 9 January 2015. Top row (panels a and b) is for the analysis up to 1 km in the vicinity of the selected CMLs. Bottom row (panels c and d) is for the analysis up to 9.1 km in the vicinity of the selected CMLs. As noted in the inset metrics, the number of CMLs vary given the selection of the vicinity/radius. Left column (panels a and c) is for the analysis of all rainfall pairs, i.e., zeros included. Left column (panels b and d) is for the analysis of those pairs in which both rainfall depths are above zero (i.e., rainy events). The color scale is logarithmic.

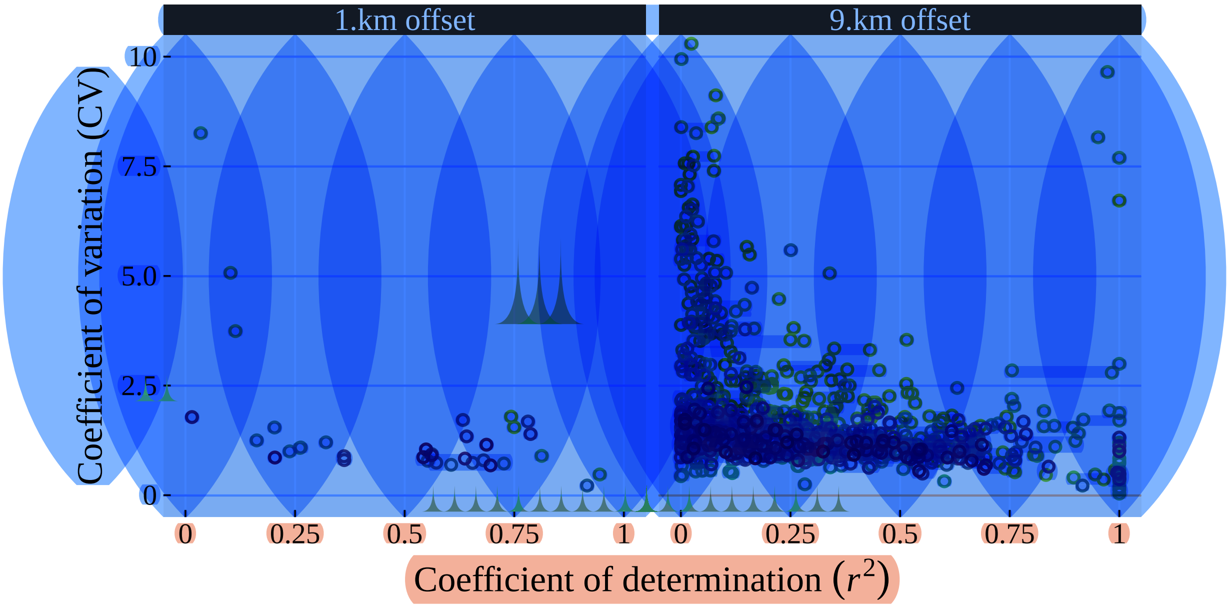


Figure 7. Scatter plot of the performance of individual CMLs against gauges (coefficient of variation against coefficient of determination). The left panel (1.km offset) is for gauges within 1 km from the selected CMLs. The right panel (9.km offset) is for gauges within 9 km from the selected CMLs. Each distinguishable color in the plots represents the metrics of an individual CML, i.e., one color per evaluated CML (regardless its gauge comparison). The metrics are for cases in which both CML and gauge rainfall depths are above 0.0 mm (Fig. 6, right column). RAINLINK estimates are computed for the SP R–k relation.