ROBUST EXTRACTION OF RAIN-INDUCED ATTENUATION FROM MICROWAVE LINK OBSERVATIONS USING LOCAL REGRESSION

by

A. Schatzmann[(1)], A. Scheidegger[(2)], J. Rieckermann[(2)] and A. Ruckstuhl[(1)]
[(1)] *Institute for Data Analysis and Process Design, Zurich University of Applied Sciences, Winterthur, Switzerland (anders.schatzmann@zhaw.ch, andreas.ruckstuhl@zhaw.ch)*
[(2)] *Dept. of Urban Water Management, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland (joerg.rieckermann@eawag.ch, andreas.scheidegger@eawag.ch)*

ABSTRACT
Recently, the rain-induced attenuation of microwave links (MWL) has been suggested as a novel tool to monitor rainfall in urban areas. However, this requires robust methods to identify and remove baseline attenuation during dry weather, which is not trivial. In this paper, we present a novel statistical method to identify the baseline attenuation based on robust local regression (REBS) offline and in real-time. To test the correct baseline removal, we were able to test the algorithm on an extraordinary dataset of a dense network of rain gauges and disdrometers that provides an unrivalled ground truth reference rain fall and avoids uncertainties due to scaling effect between the path-average MWL measurements and point rain gauge observations. Our results show that the robust regression algorithms predict the path-average rainfall intensities from MWL signal levels that are practically unbiased and match the reference rainfall well. Although the scatter for 1-min rainfall intensities can vary up to 80% of the reference measurements, this it reduced to about 20% for 10min average intensities.

Keywords: telecommunication microwave links, rainfall retrieval, robust regression

1 INTRODUCTION

Recently, microwave links (MWL), which are commonly used in telecommunication networks for wireless data transmission, have been suggested as a novel tool to monitor rainfall in urban areas (Upton et al., 2005; Messer et al., 2006). MWL seem to complement traditional rainfall sensors nicely because they provide rain rate measurements almost at ground level at an intermediate scale between point measurements from rain gauges and weather radars. This tool is based on the well known relationship $A=KR^{\alpha}$ between microwave attenuation (*A, in db* km^{-1}) and rainfall rate (*R, in mm* h^{-1}), where K is a link characteristics reliant constant and α depends on the wavelength. Atlas and Ulbrich (1977) have found that the *A*-*R* relationship, is essentially independent of the drop size distribution of rainfall for electromagnetic signals with wavelengths around 1 cm.

Prior to rainfall estimation, it is very important to correctly identify the rain-induced attenuation and remove the baseline (Wang et al., 2012). Depending on the link characteristics, baseline removal can be very difficult, because dry-weather signal attenuations can exhibit significant variability from changes in water vapour, wind effects on the antennas, birds or insects crossing the beam, losses during transmission or reception, interferences, wet-antenna and multi-path effects.

In this paper, we therefore present a novel statistical method to identify the baseline attenuation. Specifically, we adapted a baseline extraction algorithm based on robust local regression (REBS) that was originally developed for extracting background concentrations from atmospheric trace species measurements (Ruckstuhl et al., 2010). We then modified this algorithm to allow rain retrieval in near real-time. This is innovative, because, in contrast to existing algorithms, the modification considers different variability patterns in the received signal to discriminate rain from other sources of attenuation, such as wet antenna effects. Hence, our tool requires no other information than the real time signal strength of the MWL to identify the baseline attenuation and thereby estimate the rain intensity.

2 MATERIAL

To test the correct baseline identification, most studies used “ground truth” reference precipitation measurements from a nearby rain gauge, or even weather radars. This introduces additional uncertainties due to a mismatch in scale between the path-average MWL measurements and point rain gauge measurements (Fencl et al., 2012). In our study, we were able to test the algorithm on an extraordinary dataset from the COMMON experiment, a dedicated experimental campaign in Dübendorf, Switzerland. It consists of observations from a 1.85 km long commercial dual-polarization microwave link (Mini Link, Ericsson, dt=4s). In addition, 5 disdrometers (Parsivel, OTT, dt=30s) and 3 tipping bucket rain gauges (R13029, Précis Mechanique) were placed at equal distances along the path of the link. The entire dataset used in this study comprises two months of rainfall and MWL observations. A two-day period of the obtained data is shown in *Figure 1 (a)*. It illustrates the co-fluctuation of A and R during several several rain events. The labelled area shows signal attenuation during dry weather that is not caused by rain and should therefore be removed by the baseline estimator.

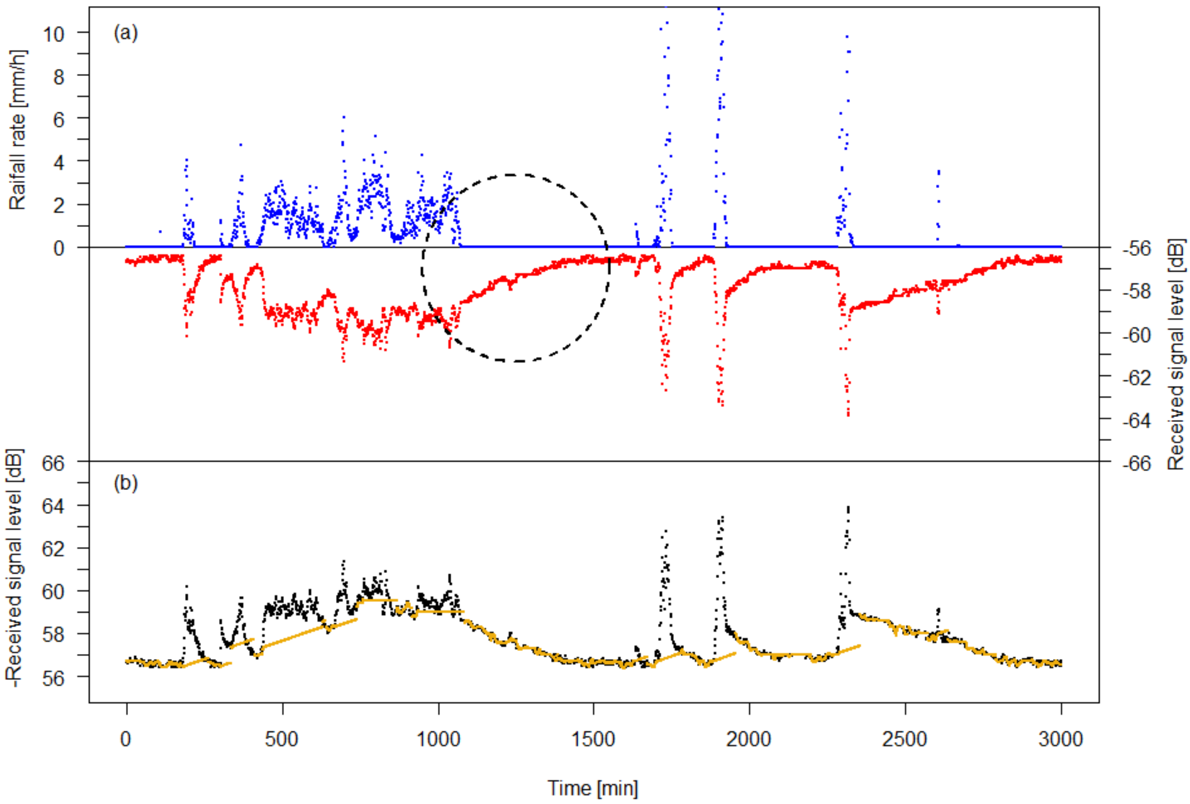

Figure 1 – (a): Co-fluctuations of MWL received signal levels (*red*) and rain intensities (*blue*) from a dedicated experiment with a commercial microwave link, 5 disdrometers and 3 tipping bucket rain gauges. The long tails in the MWL signal during post-rainfall periods (*marked area*) clearly demonstrate the importance to correctly classify antenna wetting during baseline removal. *(b):* Inverted received signal levels (black) and baseline estimated with the AREBS algorithm (*orange*). The baseline clearly follows the attenuation during post-rainfalls periods. This provides the possibility of estimating the rainfall rate more precisely.

3 METHODS

3.1 Data pre-processing

As mentioned above, there are considerable differences in the temporal resolution of the rainfall data and the received signal levels (RSL) of the MWL. To use the data for further analysis, the time series were aggregated into equivalently long data sets with a temporal resolution of 1min. Specifically, average values for each minute were computed for both data sets.

3.2 Removal of dry-weather baseline by robust regression

The first step to remove the baseline attenuation is the analysis of the variability on the MWL received signal levels (RSL). As the variability of the RSL is considerably larger during wet periods, we use the variability of the past 10 minutes of the signal for classification. Specifically, we use the REBS algorithm by Ruckstuhl et al. (2010), which is based on robust local regression and was originally designed for extracting background concentrations from atmospheric trace species measurements. The REBS algorithm was chosen for this problem due to its robust nature. As the final goal is to retrieve rain intensities from MWL attenuation in real time, the baseline extraction method needs to be stable and reliable, also in case of bad or erroneous input data. REBS uses robust methods in two directions in order to identify the baseline. In the time domain, a local regression applying a kernel weighting function is used. In addition, an asymmetric approach is used on the observations, that down-weights far ouliers and leads to the required stability of the baseline. If the variability exceeds a predefined threshold value, it is classified as a rainy period. Using this information the REBS algorithm can be tuned towards the correct baseline attenuation, which usually uses all available observations in an offline-fashion.

In addition to this, we also constructed a real-time version of the algorithm that uses an asymmetric weighting function and thus only takes into account past observations. As the original REBS algorithm is symmetric on the time axes it is to be adapted as future values can not be used for the local regression. The asymmetric algorithm (AREBS) uses asymmetric Kernel weights for the local regression that only consider the past observations

$$K\left(\frac{t_i - t_0}{h}\right) = \begin{cases} \left[\max\left\{1 - \left|\frac{t_i - t_0}{h}\right|^3, 0\right\}\right]^3 & if\ t_0 - t_i \geq 0 \\ 0 & otherwise \end{cases} \quad (1)$$

where t_0 is the present time point, h is the onsided bandwith of considered neighbor points for the regression and t_i is a point in time.

For the AREBS algorithm, one particular challenge is to predict the evolution of the baseline during a rain event, which is not only unobserved, but is also complicated by the fact that baseline levels after a rain event do not necessarily equal those before an event (*Figure 1a*). In a first approach, we therefore chose a simple linear baseline model during rainy periods, which has a constant slope during all rainy periods and a fixed maximum for the baseline (cf. *Figure 1b)*. Exploratory analysis of several rain events results in an average slope of about - 0.005 [dB/min] and a maximum received signal level of about -59 dB,

$$b^{(i)} = \begin{cases} b^{(i-1)} + s & for\ b^{(i)} \leq 59\ dB \\ -59 & otherwise \end{cases} \quad (2)$$

where $b^{(i)}$ is the baseline in minute i and s is the constant slope.

3.3 Rainfall retrieval from MWLs with a double-logarithmic model

In literature, rainfall (R) is usally estimated from the specific rainfall-induced attenuation of a MWL (k, [dB/km]) with a power law model:

$$R = \alpha k^{\beta} \quad (3)$$

with empirical parameters α and β. The validity of this model mainly depends on a low variability of the drop size distribution along the link, which is not always the case. For our dataset, in-depth investigations based on residual analysis suggested that a double-logarithmic linear model provides the best fit to the data (not shown). To formally avoid negative rain-induced attenuation values, we re-formulated the power-law to

$$\ln(R_i) = \beta_0 + \beta_1 \ln(A_i - b_i) + E_i \tag{4}$$

where $\ln(R_i)$ is the natural logarithm of the measured rainfall rate, $\ln(A_i - b_i)$ is the logarithm of the rain-induced attenuation (i.e., difference between RSL and the baseline) and β_0, β_1 are the model parameters. Model estimation with classical least square regression leads then to the following model for rainfall retrieval:

$$\hat{R} = \begin{cases} e^{\hat{y}+\frac{\hat{\sigma}^2}{2}} & for \quad \hat{y} > 0 \\ 0 & otherwise \end{cases} \tag{5}$$

with

$$\hat{y} = \begin{cases} 0.12 + 1.17\ln(A - b) & for \quad A - b > 0 \\ 0 & otherwise \end{cases} \tag{6}$$

where $\hat{R}$ [mm/h] is the model estimated rainfall intensity. It is straight forward to apply this model in realtime estimation, because the linear model is simple and computationally fast.

4 RESULTS AND DISCUSSION

If we compare the path-average rainfall intensities predicted from MWL received signal levels by the REBS algorithm to the reference values from the rain gauges and disdrometers, they are practically unbiased (-0.0035 mm/h) and approximate the observations well. As depicted in *Figure 2a* and *2b,* the predictions on a one minute basis (*a*) show some variability, which increases with higher intensities and has a maximum range of the prediction variability is around 40 mm/h. That means, that the predicted rainfall per minute can vary strongly, but that the total model prediction bias is negligible. If average intensities over 10 minutes are predicted from the MWL signals (*b*), the scatter in the predicted intensities is reduced to a range of 17 mm/h while the model bias remains approximately the same.

Interestingly, the performance of the AREBS algorithm is comparable to the offline analysis. Especially for longer rain events (*Figure 1b, 0 -1500 min*), the constant slope seems to work well. However, it is not steep enough for the short rain events (>1500 min). Furthermore, certain discontinuities are introduced after rain events. Due to the asymmetric nature of the algorithm such, jumps occur if the rain event is longer than the number of used points for the baseline. In this case, AREBS estimates the baseline from the past 25 observations. Therefore, if the duration of the rain-event is longer, then the algorithm has no more past data for the baseline estimation. This results in a jump to the first observation that is classified as baseline when the variability of the received signal drops to dry weather levels after the rain event. Despite the discontinuities the constant slope, the results of AREBS (*Figure 1b*) are very similar to those of REBS (not shown). This is also illustrated in *Figure 2c* and *2d,* which shows that the prediction accuracy of AREBS is only slightly worse, wich a variability of about around 48 mm/h and a bias of 0.016 mm/h. The prediction error of 0.162 mm/h is the same as in the non-averaged case. In general, the AREBS predictions are very comparable to those of REBS. Both mean residuals are far below the threshold of 0.2 mm/h, which is the detection limit of tipping bucket rain gauges, and the predictions improve when averaged over multible data points.

The results are satisfactory, however, this might to some degree explained by the fact that the same data are used for both, model estimation and model testing. To genuinely verify the model, an out of sample test should be applied to new data. Therefore, the results should be taken with care. With the available data we were not able to test how the model is reacting during long rain events over several days or during snowfall. Furthermore, seasonal climatic effects might also introduce additional variability.

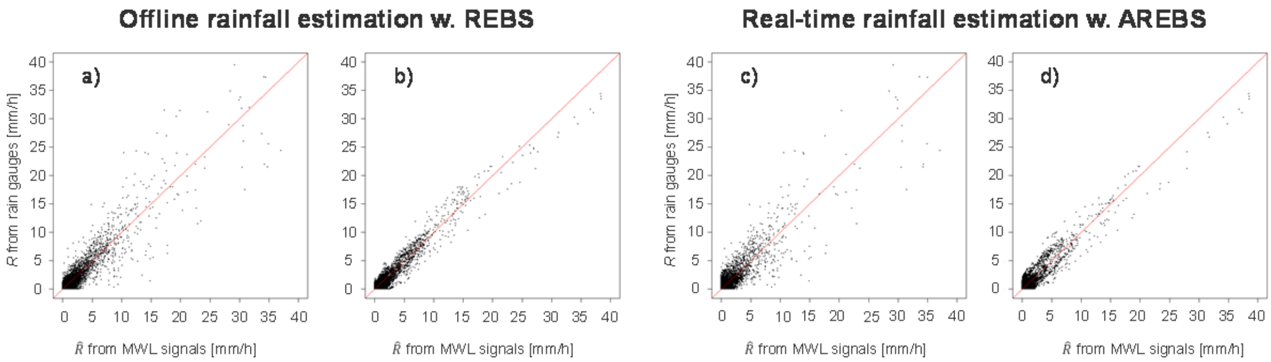

Figure 2 – Comparison of path-avergage rainfall intensities from MWL data to those observed with a rain gauge network. *(a)*: Three months of rainfall intensities with a temporal resolution of one minute. The observations from the offline analysis with the robust regression scatter regularly around the bisectrix. *(b)*: Predicted average rainfall intensities over 10min intevals are more precise and show less variability than the 1min values. *(c)*: The results of the offline analysis with the AREBS algorithm for 1min resolution and *(d)* 10min averages.

5 CONCLUSIONS

In this study, we developed a novel statistical method to identify the rain-induced attenuation from commercial microwave links and directly estimate path-average rainfall intensities from received signal levels. Our results show that rainfall retrieval is possible by only using MWL signals. Interestingly, the online version of the robust regression algorithm (AREBS), which only uses past observations, performs nearly as good as the offline version (REBS) which uses all the available information. The results demonstrate that a single predicted value can vary up to 80% of the observations from a rain gauge network with a temporal resolution of 1 min. For average intensities over 10 min, the scatter in predicted intensities is reduced to about 20%. With this study we could show that the modified baseline extraction for real time rainfall estimation shows satisfying results. Due to a lack of data no out-of-sample test have been performed yet. If MWL should be used on a large-scale for precipitation measurements, robust models are crucial. Therefore, further studies should investigate the uncertainty of the prediction on longer data sets and in situations where no calibration is possible.

6 REFERENCES

Atlas, D. and C.W. Ulbrich, (1977) Path and area integrated rainfall measurement by microwave attenuation in the 1-3 cm band. *J. Appl. Meteor.* 16, 327–332.

Fencl, M., J. Rieckermann, D. Stránský, D. Fidranská and V. Bareš, (2012) Telecommunication microwave links as a source of rainfall information for urban rainfall-runoff modelling. Presented at the 9th INTERNATIONAL WORKSHOP on PRECIPITATION IN URBAN AREAS Urban Challenges in Rainfall Analysis, St. Moritz, Switzerland.

Messer, H., A. Zinevich and P. Alpert, (2006) Environmental monitoring by wireless communication networks. *Science* 312, 713–713.

Ruckstuhl, A.F., S. Henne, S. Reimann, M. Steinbacher, B. Buchmann and C. Hueglin, (2010) Robust extraction of baseline signal of atmospheric trace species using local regression. Atmospheric Measurement Techniques Discussions 3, 5589–5612.

Upton, G.J.G., A.R. Holt, R.J. Cummings, A.R. Rahimi, A.R. and J.W.F. Goddard, (2005) Microwave links: the future for urban rainfall measurements? *Atmos. Res*. 77, 300–312.

Wang, Z., M. Schleiss, J. Jaffrain, A. Berne and J. Rieckermann, (2012) Using Markov switching models to infer dry and rainy periods from telecommunication microwave link signals. Atmospheric Measurement Techniques Discussions 5, 411–445.