Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-397-RC1, 2016

Hydrology and Earth System Sciences Discussions

Open Access

HESSD

Interactive comment

Interactive comment on "Gauge-Adjusted Rainfall Estimates from Commercial Microwave Links" *by* Martin Fencl et al.

H. Leijnse (Referee)

hidde.leijnse@knmi.nl

Received and published: 27 September 2016

This paper describes a method for incorporating accurate rain gauge measurements in commercial microwave link (CML) rainfall estimation through on-line parameter adjustment of the CML retrieval model. The idea of adjusting those model parameters that we know are most uncertain based on rain gauges is very appealing. This means that the accuracy of the gauges is used where it is most needed. The authors test their method on two different datasets, with different algorithm settings and different distances to the gauges used for adjustment. I think that the paper is interesting and certainly appropriate for HESS. I also have some issues that I think the authors should deal with before the paper is ready for publication. The most important of these issues are: 1) How well does the presented method work when gauges are even further away from the links (i.e., how well can this method be employed in sparsely gauged regions)? 2) The

Printer-friendly version

Discussion paper

model is claimed to be linear, but this is not the case (see specific comments below). and 3) The evaluations presented here are likely to be heavily influenced by the very high correlation (perfect in the case of one of the datasets) between the gauges used for adjustment and those used for validation. More specific remarks are given below.

Specific remarks

1. On p.3, line 24 the units of α are incorrect (should be mm h^{-1} km$^{\beta}$ dB$^{-\beta}$).
2. On p.6, lines 10-12 it is mentioned that four links are selected. It's not clear to me what this selection was based on. I'm guessing that they were selected because these links were in (or close to) the catchment. Or were there more links in the area that were not selected. Can you provide a short statement in the paper about why these links were selected?
3. Section 2.3 seems redundant to me, and its contents can simply be put in Sections 2.1.1 and 2.1.2.
4. On p.7, lines 2-3 the authors claim that using the power law of Eq.(1) could result in overfitting. However, this power-law relation has been shown to be robust and relatively insensitive to variations in raindrop size distributions. So the parameters of this relation can be safely taken from literature without fitting them within a retrieval algorithm. The key to getting good rainfall estimates is to properly take effects of a variable baseline and wetting of antennas into account. So while I can certainly understand that the authors want to use an as simple equation as possible for the analyses presented in this paper, I think that the risk of overfitting should not be stated as a reason here.
5. On p.7, line 7 it is stated that k is the specific attenuation after baseline separation. It would be good to specify here which method is used for determining and

separating this baseline.

6. On p.7, line 7, I suggest stating that you can use this simplification because b is very close to 1 for the frequencies that are often used in CML networks.

7. On p.7, lines 20-21, as first glance I didn't think that it is necessary to state how the optimization is carried out because of the linearity of Eq.(2) and the fact that aggregation over time is a linear operation. Hence minimizing L in Eq.(3) is a linear regression problem that has an analytical solution (even if you force the line to go through zero). However, I'm assuming that the authors are setting resulting rainfall estimates to zero if $k < k_w$ (which would yield $R < 0$ mm h^{-1}). This effectively means that although Eq.(2) is linear, the model that the authors are using is not. It should be expressed as

$$R = \begin{cases} \gamma\,(k - k_w) & \text{if} \quad k > k_w \\ 0 & \text{if} \quad k \leq k_w. \end{cases}$$

I think that it should be clearly stated in the text that the model is effectively not linear. I also think that the implications of this nonlinearity should be discussed in the text. Furthermore, this means that the reason for using this linearized form that is stated by the authors is not valid (because they're using a nonlinear model). In fact, one could argue that Eq.(1) could be kept as a basis for the equation that is optimized, with a provision for correcting for wet antennas and baseline variations. Something like

$$R = \begin{cases} \alpha\,(k - k_w)^{\beta} & \text{if} \quad k > k_w \\ 0 & \text{if} \quad k \leq k_w, \end{cases}$$

where k_w includes wet antenna and baseline variation effects, and hence should then be the only parameter that is fitted (and α and β taken from literature).

8. On p.7, line 31 a description is given on how the second parameter optimization run is carried out. It is stated that this run uses the parameter *distribution* of the first optimization run. However, I don't understand how the first run can yield a distribution of parameters. Or is it the distribution of parameters over all time steps in the entire dataset? In that case, the method cannot be used in a real-time setting.

9. On p.8, lines 7-12 it is stated that the effect of temporal aggregation is studied by comparing the gauge-adjusted CML rainfall product with the same gauges that were used to adjust the CML data. I expect the fact that the gauges are not dependent to have a large effect on the outcome of the analyses. Am I correct in assuming that this is only the case for the Dübendorf dataset, and that in Prague you're using the municipality gauge network as a reference? I think that the fact that the gauges in Switzerland are not independent should be discussed in the paper.

10. On p.8, line 24, a reference rainfall measurement is mentioned. It is not clear to me what this reference is. It this the average of the six (p.6 line 16) or four (Fig.1) rain gauges operated by the municipality for the Prague datset and the rain gauges and disdrometers for the Dübendorf dataset?

11. In Section 3.1 the authors discuss the reasons why parameter fitting for shorter intervals yield better results than for lnger intervals. I don't really agree with this discussion. What effectively happens when the length of the aggregation is increased is that the CML data receive more weight in determining the temporal evolution of the rainfall signal (relative to the gauges). Because either the same gauges (Dübendorf) or a gauge dataset that is well-correlated to the gauges that are used for the parameter fitting (Prague; see top-right panel of Fig.4) are used for verification, it is expected that the results are best if the weight of the gauges is largest (i.e., for the shortest accumulation intervals). So I don't think that you

Printer-friendly version

Discussion paper

can actually draw conclusions about which accumulation interval is best suited for this method based on these analyses.

12. On p.9, lines 20-22 the authors state that using daily rainfall accumulations to fit the model parameters would minimize the effect of diurnal fluctuations in baseline level. I think the converse is true: in order to minimize the effect of diurnal fluctuations, the model parameters should be fitted on a time scale that is significantly shorter than a day so that this variability is actually captured.

13. On p.13, Section 4.2 the authors discuss how the distribution of the γ parameter changes with aggregation interval. This is then related to the fact that the proposed model includes the effect of wet antennas. However, this effect should be more related to the k_w paramter of the model, and not so much to γ. Of course, the two model parameters can compensate, and this would result in wider distributions of γ, but this is a purely an effect of the fitting procedure.

14. On p.13, lines 17-18 the authors state that they've found a connection between the observed systematic errors and the degree of preservation of rainfall space-time structure through averaging. I don't really see this connection, and I think this should be better explained.

15. On p.14, line 9 the use of CML networks in sparsely gauged regions is mentioned. However, the method persented in this paper probably won't work in sparsely gauged reasons because rain gauges located close to the links are essential (see Figures 1, 4, and 5). So I think this statement needs to be altered.

16. On p.15, line 18 it is stated that CML networks can provide rainfall data on a (sub-)kilometer scale. However, I really don't think that this will be attainable with the method presented here. This is because of the fact that the CML data are adjusted to a (point) rain gauge somewhere in the vicinity, which will effectively

smooth out much of the variability captured by the individual links. So this statement should also be put into perspective.

17. In Figure 1, right panel, there seem to be white letters over the figure that are partly over the disdrometers.

18. In Figures 2, 5, and 6 the coefficient of determination (R^2) becomes negative. It would be good to give the definition of R^2 that was used in the paper in Section 2.6 (there are different versions of R^2, some of which cannot become negative).

19. In Figures 3 and 4 the slope of the regression line $y = ax$ (i.e., with fixed offset) is given. It should be noted here that the correlation between the two variables affects this slope. The slope will always be lower with a low correlation coefficient (you can try this by switching the x- and y-axes; see also the right-hand panels of Fig.4).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-397, 2016.