Review of manuscript hess-2021-56, "Simulation of rainfall fields conditioned on rain gauge observations and radar estimates using random mixing" by Jieru Yan et al.

Remko Uijlenhoet

General remarks

Quantitative Precipitation Estimation (QPE), i.e. our ability to estimate space-time rainfall fields using data collected by weather radars and rain gauges, is important for obtaining hydrological process understanding as well as for water management applications. As the space-time sampling properties and other instrumental characteristics of radar and gauges are complementary to some extent, several methods have been proposed in the scientific literature to optimally combine ("merge") radar and gauge data to obtain spatial precipitation estimates over areas such as river catchments.

Conditional simulation has great potential to be used in spatial precipitation estimation as it is able to account for uncertainties associated with QPE through a Monte Carlo framework. The authors claim that one obstacle hampering the application of conditional simulation in QPE is obtaining the marginal distribution function of the rainfall field with sufficient accuracy. They propose a method to obtain this marginal distribution function based on rain gauge observations and radar estimates, namely Random Mixing (RM). Spatial rainfall fields estimated using RM are compared with those from well-known radar-gauge merging methods: Ordinary Kriging, Kriging with External Drift, and Conditional Merging.

The subject of radar QPE through radar-gauge merging is definitely suitable for publication in HESS. However, apart from the specific remarks and editorial remarks provided below, this study has two important limitations, which should be stressed more clearly in the paper: 1) it is a pure simulation study – in spite of what is suggested by the title of this paper, there are no radar or rain gauge data employed in this study; 2) the study only considers the estimation of *spatial* rainfall fields, completely neglecting the *temporal* aspect of QPE, which is so important for hydrological applications. These two limitations should be reflected more clearly in the title, abstract and introduction of the paper, as far as I am concerned. Moreover, a proper Discussion section appears to be missing from the paper.

Overall, this paper proposes an interesting and relevant stochastic radar-gauge merging method, which is quantitatively tested in a controlled environment through stochastic simulations. Although I found the presentation of the proposed algorithm sometimes challenging to follow (see remarks below), I do find it an interesting and relevant addition to the hydrological engineering literature. Thanks very much in advance for taking my suggestions as provided above and below into account.

Specific remarks

- L.45: Hitschfeld and Bordan (1954) is a classical reference for (physical) radar attenuation correction. I would not cite this reference as a typical example of (statistical) mean field bias correction. Perhaps you could refer to some of the papers by D.-J. Seo (formerly at NOAA-NWS) on this issue.
- L.91–97: If I understand correctly, you are forcing the intermittency in the (transformed) radar data to be the same as that of the rain gauge data. However, since both sensors (radar and gauges) have very different space-time sampling properties, the probability of finding zeroes in the radar rainfall field may be quite different from the probability of finding zeroes in the rain gauge field. In particular, the spatial aggregation associated with radar rainfall observations and the temporal aggregation associated with rain gauge observations make that their respective capacities to detect intermittency in space and time are quite different. To what extent does this affect your approach?
- L.118–124: Here, "correlation function" and "variogram" are used interchangeably. Please choose one of the two terms as a measure of spatial dependence. Moreover, it may be true that "it can be seen that the empirical and the true variograms have very similar patterns", but that does not mean that the truncated Gaussian fields necessarily have the appropriate (i.e. a realistic) spatial intermittency structure. Is that a problem for your approach or not? If not, why not?

- L.192: “George Dantzig’s simplex algorithm, the BFGS method, etc.” – please add literature references to these methods.
- L.135–204: I am able to follow the general reasoning of Section 2.3 (“Random Mixing”), however, I am not able to grasp all the intricacies of the proposed method. It would be good if one of the referees would be familiar with the RM method, or at least if he/she would have a solid background in stochastic processes.
- L.212–213: The authors use again the notion of the correlation function (also see Fig. 3 and caption), whereas previously they employed the variogram to express the degree of spatial dependence. For the sake of consistency, it would be good to stick with the notion of variogram and use that throughout the paper. Alternatively, the correlation function could be employed instead of the variogram. However, please do not use these notions interchangeably.
- L.217–218: “rainfall rates over an area, or the objective of this paper: short time rainfall over an area” – what do you mean exactly with “short time rainfall”? To what extent is it different from “rainfall rates”?
- L.232–233: “A random error is therefore introduced to mimic this kind of error”, namely “factors […] that can reduce the representativeness of the radar estimates for the rainfall pattern on the ground, such as evaporation, complex terrain effects and anthropogenic influences” (L.231–232). However, many of the representativeness errors in radar rainfall estimation are not random, but systematic, e.g. range-dependent errors associated with the increasing height of the radar beam above the ground the further one goes away from the radar antenna, or the associated increase of the radar sampling volume. How would one account for such systematic effects in the proposed simulation? How important would it be to incorporate such systematic effects?
- L.238 – 240: “It should be notified that the introduction of random error in Step 1 differs the radar estimates from the true rainfall field in terms of the field pattern, yet the changes in the statistical properties are tiny” – see previous remark: would these “changes” still be “tiny” if more realistic *systematic* errors would be accounted for in the simulation framework, rather than *random* errors alone?
- L.252–253: “The choice of the two parameters - factor 0.87 and exponent 0.83 - is quite arbitrary. We have modeled a case when radar underestimates the precipitation. Surely one could model other cases.” – this approach does not appear to be based on sound assumptions (“quite arbitrary”). How did you determine these parameter values? Have you considered other values? What process are you trying to mimic here, the Z–R relation? With the typical Z–R exponent of 1.6, one would perhaps have expected an exponent of 1 / 1.6 = 0.625, rather than 0.83. In any case, a clear motivation for using a value of 0.83 seems to be lacking. The same holds for the employed value for the factor (0.87). In addition, how representative are your results, if these parameter values were chosen in such an arbitrary manner?
- L.265: “This gives an approximate coverage of one rain gauge for every 256, 178, and 131 km^2, respectively.” – why these rain gauge densities? How do they compare to rain gauge densities encountered in practice? Please motivate these choices, which (again) seem to be quite arbitrary.
- Fig. 5: These are certainly interesting simulation results, however, the generated random fields do not really resemble actual radar-estimated rainfall fields, which often display a strong sense of directionality (anisotropy), associated with the prevailing movement direction of the rainfall field. So, how representative do you think your simulation results are for practical purposes, e.g. when applied to radar and rain gauge data from the German national weather service (DWD)? This also refers to the lacking temporal aspect of radar QPE, which was mentioned already under “general remarks”.
- L.295–296: “to a certain degree” – to what degree? Please try to be as concrete as possible.
- L.306–307: “Instead, one can obtain an infinite number of realizations for the same true rainfall field by RM” – but Kriging and other geostatistical methods can also be employed for (conditional) simulations, can’t they? Or would these methods lead to the issues associated with the inaccurate representation of the marginal distribution of the generated rainfall fields, which the authors referred to in the abstract and introduction of their paper? If yes, please mention this explicitly.
- L.314: “the mean realization is more helpful in identifying the locations of the rain cells” – OK, but this apparently goes at the expense of a realistic spatial variability. Can you get the locations *and* the variability right simultaneously with the proposed method? That would be relevant for practical applications, it seems to me.
- Fig. 9, Table 1: Have “G25” (presumably 5x5 rain gauges), “G36” (6x6) and “G49” (7x7) been defined before? If not, please do so.

- L.405: After presenting the Sensitivity Analysis (Section 4.2), which is part of the Results (Section 4), you immediately jump to the Conclusions. However, a true Discussion section, where one puts one's own results into perspective, by critically discussing assumptions and the associated limitations and by comparing obtained results with results reported elsewhere in the scientific literature, appears to be lacking. However, this is an important element of any scientific study. Therefore, I urge the authors to include such a Discussion in a revised version of this paper (unless I missed it).

Editorial remarks

- L.1, 19: "temporospatial" → "spatiotemporal".
- L.5, 38: "lack" → "lacks".
- L.10, 62, 81: "accuracy" → "sufficient accuracy".
- L.43, 48: "generic" → "class".
- L.43: "have" → "has".
- L.60: "hamper" → "hampers".
- L.78: "cdf" → "CDF" (here and elsewhere in the paper).
- L.80: "cdf-rainfall" → "rainfall CDF" (here and elsewhere in the paper).
- L.84: replace full stop (".") after "as follows" with colon (":").
- L.85–86: "the pixels" → "pixels".
- L.87: "Uniform" → "Transform".
- L.88, 90, 254: "uniformed" → "transformed".
- L.99: "Due to" → "For".
- L.102: "distribute unbiasedly" → "are distributed without bias".
- L.107: "in the marginal" → "in terms of the marginal" (if this is what you mean).
- L.109: "In specific" → "In particular" / "Specifically".
- L.109: replace full stop (".") after "constraints" with colon (":").
- L.110, 116: "in the standard normal marginal" → "in terms of the standard normal marginal" (if this is what you mean).
- L.118: "A problem" → "The problem".
- L.119: "namely" → "where".
- L.119: "And due" → "Due".
- L.125: "to a certain degree" → "as close as possible".
- L.129–140: Remove "the formula of the".
- L.137: "an estimate for" → "an estimate of".
- L.141–142: "distribution" → "distributions".
- L.142, 180: "L-2" → "L2".
- L.145: "(Bárdossy and Hörning, 2016)" → "Bárdossy and Hörning (2016)".
- L.146: "function, and we …" → "function. We …".
- L.157: "subject" → "subject to".
- L.159: "equation" → "equations".
- L.159: "it" → "this".
- L.161: Insert "L2" before "norm".
- L.162, 169, 237: "subject" → "imposed" (if that is what you mean).
- L.191: Remove "the" at the end of this line.
- L.195–196: "as the following" → "as follows".
- L.209: "is representative" → "is assumed to be representative".
- L.224: "Namely, …" → "Hence, …".

- L.228: replace full stop (".") after "applied" with colon (":").
- L.239: "differs" → "changes" / "affects".
- Figs. 4, 9, 10: The y-axis labels read "number of frequencies". This should simply be "frequency" (which means "number of occasions").
- L.260: If you use "gauge" rather than "gage", you should also employ "ungauged" instead of "ungaged".
- L.260: "from the rain gauge" → "to the rain gauge".
- L.271: "in details" → "in detail".
- L.272, 273, 308, 385: "on Panel(s)" → "in Panel(s)".
- L.278, 280, 283, 294, 297: "extreme" → "extremes".
- L.282: "scenario is radar" → "radar scenario" (if this is what you mean).
- L.296, 411: replace full stop (".") after "as follows" with colon (":").
- L.310: "capture" → "captures".
- L.317: "uniformly distribute" → "are uniformly distributed".
- L.322: "from" → "to".
- L.330, 332, 333: "sited" → "located".
- L.336, 341, 342, 343: do you mean "extreme" or "maximum"?
- L.338: Insert "for" before "stormwater management".
- L.349, 352, 422–423: "field extreme" – do you mean "field maximum"?
- L.391–392: "And for …" → "For …".
- L.437 – 532: Please carefully check your reference list. Some articles titles are capitalized, others not; book titles need to be capitalized, whereas some are not; the price of Battan (1973) does not have to be included in the reference (L.444).