Precipitation paper, Anonymous Referee #2

Comments on: Generating spatial precipitation ensembles: impact of temporal correlation structure

The manuscript presents a methodology for generating ensemble rainfall fields using conditional sequential simulation. The manuscript fits the scope of Hydrology and Earth System Sciences. However, I cannot recommend publication in its current form. My main concern is that the presented approach is not validated rigorously. The manuscript could be published pending a major revision by the authors.

Answer:
We thank Anonymous Referee #2 very much for her/his constructive review. Below you will find our response to the comments raised by Anonymous Referee #2. Additionally, we agree that validation of the presented approach is very important and therefore will include this validation in the revised version of the manuscript. The validation is elaborated more in detail in our answer to the comment #1.

More specific comments are available below:

1) The simulated ensembles and the conditional simulations are not validated rigorously. Presenting the mean and standard deviations of observations and simulations cannot be considered as a thorough validation. The authors could leave a certain number of gauges out of simulations and use them to cross-validate the conditional simulation. This is a very common practice in most geostatistical papers, some of which cited in the manuscript. Alternatively, the bootstrap technique can be used to evaluate the representativeness of the simulated ensemble.

Answer: Thank you for your comment. We will include a validation of the presented results on three real world events in Section 3.3 as follows:
To verify the accuracy of the presented method, validation was carried out in the terms of the mean error. The mean error is defined as the difference between the rain gauge observation and the corresponding across ensemble mean (mu_J,t, Eq. 5). The rain gauge observations employed in the validation were independent from the data used for simulation. To simulate the precipitation fields, 14 rain gauges out of the complete observation network of 27 rain gauges were used (Fig. 6b in the manuscript). The remaining 13 rain gauges were kept for validation and their mean errors were calculated for all time steps and for all eight simulation memories (dashed histograms in Figure I).
Additionally, we compared those validation mean errors with the simulation mean errors at the same 13 locations. These simulation mean errors were obtained by simulating precipitation fields using all 27 rain gauges (grey histograms in Figure I). Figure I shows that the validation mean errors at the unobserved locations are unbiased and have consistent behaviour over all simulation memories (dashed histograms in Figure I). Furthermore, the histogram of validation mean errors have smaller peaks than the simulation mean errors. This increase in uncertainty can be expected, because in the validation only half of the rain gauge data are used to simulate spatial precipitation fields. Note that the spread in histograms agrees well with the corresponding standard deviations shown in Table 1.

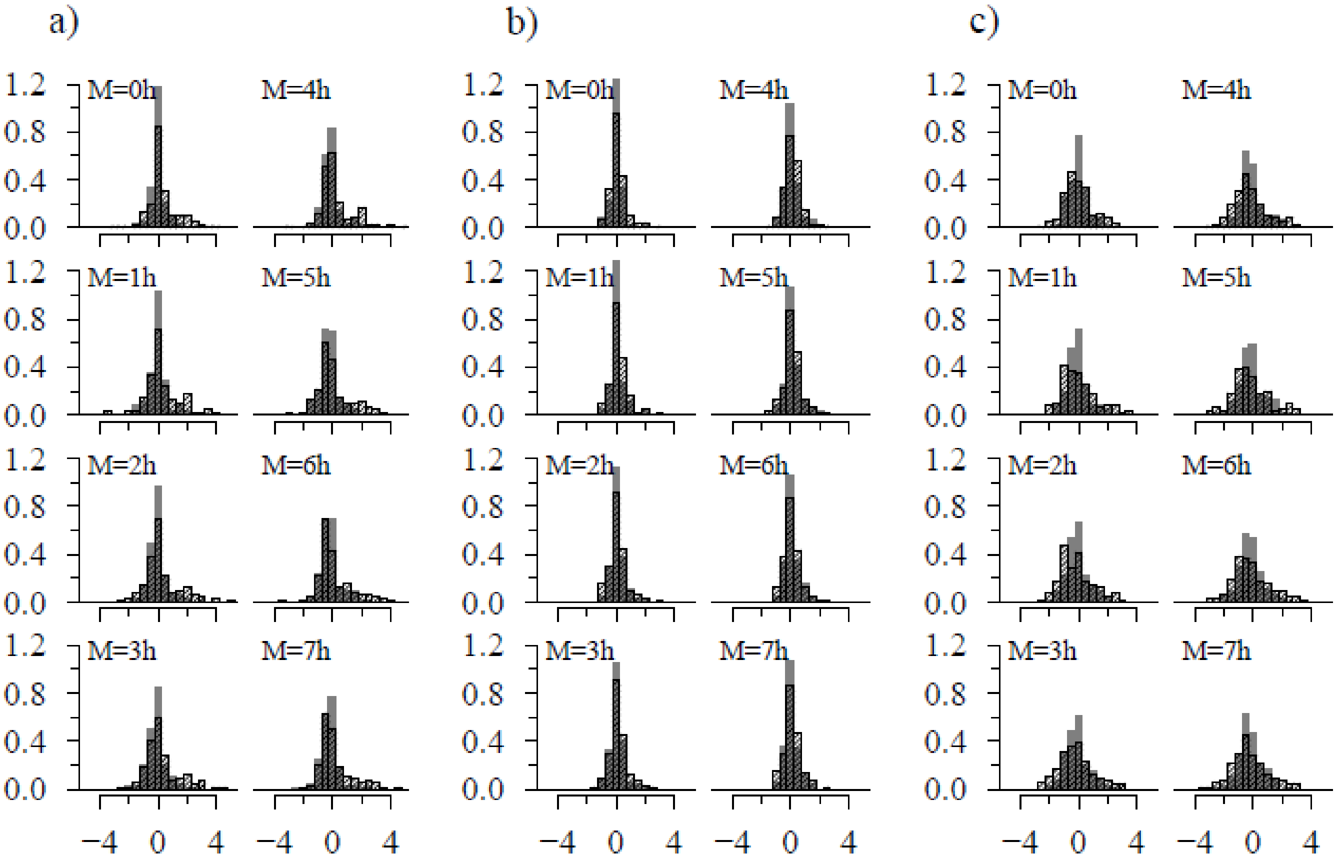

Figure I: Validation for three real world events: a) 22 October 2002, b) 22 December 2002 and c) 1 January 2003. Histograms of the validation mean errors at 13 rain gauges (dashed histograms). Histograms of the simulation mean errors for the same 13 rain gauges (grey histograms).

2) a) Parameter estimation: no information on the representativeness of the proposed variogram is provided. b) Also, it is not clear how the parameters are estimated. This is an important part of the manuscript, as simulations depend on the choice of variogram. Include a discussion on parameter estimation and explain how the goodness-of-fit is evaluated.

Answer: a) Our main interest was to see how the temporal aspect affects the uncertainty in simulated fields and therefore we decided to keep the parametrization as simple as possible (simple variogram model with no anisotropy). The spherical variogram model is a very widely employed variogram model for rainfall (e.g. Berne et al., 2004; Schuurmans, et al., 2007; van de Beek et al., 2011; Verworn and Haberlandt, 2011). Certainly, the choice of variogram model involves some subjectivity. The results are not expected to change radically if another variogram model would have been chosen. We will explicitly mention this in the revised manuscript. (See our answer to Reviewer #1, Major comment #3.)

b) We agree that the parameter estimation of the applied spherical variogram models was not described properly. Therefore, in agreement with Anonymous Referee #1 and Martinus van den Berg (Referee #3), we will make a clear link between Sect. 2.2 and Sect. 2.3 and we will add a clear description on how those model parameters were obtained/fitted. The gstat R package function fit.lmc() was employed to fit the variogram models using the weighted least square fit as a goodness-of-fit. The default method uses weights N_h/h^2, where N_h the number of point pairs and h the distance (Pebesma, 2004). The initial model parameters of the spherical variogram model were set as follows: partial sill = 0.1mm, range = 70000m and nugget=0.02mm.

Function fit.lmc() also ensures as well that the system is positive definite (Pebesma, 2004).

The purpose of the synthetic simulations is not clear at all. In my opinion, the authors should focus on the “real-world” experiments and conduct a rigorous validation. That, in my opinion, would make the manuscript acceptable for publication. I urge the authors to eliminate the synthetic simulation all together (see also the next comment).

Answer: We fully agree with the Review #2 that the expression “synthetic experiment” was overstated. We employed the synthetic simulations as sensitivity analysis to understand relevant space/time scales of precipitation as well as to demonstrate their effect on the cross-variogram. It was not meant at all to come to the conclusions of the real world experiment. Therefore we will change the name of the analysis from “synthetic experiment” into “**sensitivity analysis**” in response to this relevant comment by Reviewer#2.

We think that the “synthetic sensitivity analysis” is an important part of the current paper and therefore we would prefer to keep it. However, we will justify it more clearly in Sect. 2.5.2.: by carrying out sensitivity analysis of a very simplistic rainfall cell with known spatial statistics, constant rainfall intensity and known advection speed, we are able to eliminate several sources of uncertainty which would arise with the real world data.

Moreover, the other two reviewers supposedly did not suggest removing this part.

3) Apparently, the authors use the spherical variogram to generate the synthetic scenarios. Then, the same model is used for simulations. In a sound synthetic analysis, the synthetic rainfall generator should be independent of the model used for simulations. Such synthetic analyses cannot be used for validation and verification purposes as the same model is used for both data generation and simulation. For this reason, I think the presented synthetic analyses do not provide much insight into the methodology. Neither do the analyses lead to more information on applicability and/or validity of the model. For the reasons mentioned in the above two comments, I suggest removing the synthetic analyses part.

Answer: Apparently this part of the paper was not explained thoroughly enough. We agree completely with the reviewer that in a sound synthetic experiment the rainfall generator should be independent of the model used for simulations. The current paper indeed made this distinction, but apparently this was not described properly in the manuscript. We used a moving cicular rain cell as the synthetic rainfall generator with constant precipitation intensity, whereas the model simulations precipitation estimates were generated based on using the parameters of a spherical variogram model.

As explained in our previous answer, we would prefer to keep the synthetic sensitivity analysis in the paper, because we employed the synthetic simulations as sensitivity analysis to understand relevant space/time scales of precipitation and to aid the interpretation of the cross variogram. Moreover, the other two reviewers supposedly did not suggest removing this part.

4) Figure 2 is not discussed in sufficient details. Explain the presented boxplot including the solid black lines and the bounds.

Answer: Figure 2 shows that the time/memory scales of the precipitation are highly relevant and that the memory follows an exponential decay. We will refer back to it when we will discuss the recommended simulation memory of 2-3h for the real world simulations.

The bold black line in the boxplot is the median, the body of a boxplot shows the interquantile range (Q75-Q25) and the whiskers represent the sample minima and sample maxima, unless the extreme value occurs further than the distance 1.5 times the interquartile range from the box. Then is the outlier shown by a dot. This is a default R configuration.

5) Indicate how the data in Figure 2 is used in simulations, and discuss the limitations of deriving autocorrelation coefficients from just few events.

Answer: Figure 2 is shown to illustrate the autocorrelation coefficients of the real areal data. These data were not used in simulations at all. Figure 2 shows that the time/memory scales of the precipitation are highly relevant and that the memory follows an exponential decay. We will refer back to it when we will discuss the recommended simulation memory of 2-3h for real world simulations.

6) Figure 3 seems to present a general example of experimental and spherical variograms. General examples are already available in standard textbooks. Either eliminate the figure or revise the example and use real data from the gauges used in this study (identify the events, and discuss representativeness of the fit).

Answer: We would like to keep this figure because it illustrates an example of both direct and cross-variograms. However, we will derive it from real data from gauges used in the study.

7) References:

(I) Reference to Krajewski and Smith 2002 on radar bias and error.
Krajewski WF; Smith JA, 2002, Radar hydrology: rainfall estimation, Advances in Water Resources, 25 (8-12), 1387-1394.

Answer: We will add this reference to the introduction.

(II) P3: Germann et al. 2009 present an ensemble radar simulator and not gauge adjusted radar precipitation fields for bias. Germann, et al, 2009: REAL-Ensemble radar precipitation. . .., Q. J. Roy. Meteorol.Soc.

Answer: Yes, we are sorry, you are correct, Germann et al. (2009) do not present a study on rain gauge adjusted radar estimates to correct for bias. We will properly acknowledge their work in the revised version of this manuscript.

(III) P4 Parag. 10: Acknowledge studies on rainfall ensemble simulators using conditional simulation (e.g., Grimes and Pardo-Iguzquiza, 2010, AghaKouchak et al., 2010, Clark and Slater, 2006).
Grimes D., Pardo-Iguzquiza E., 2010, Geographical Analysis, 42 (2), 136-160.
AghaKouchak A., Bardossy A., Habib E., 2010, Conditional Simulation of Remotely Sensed Rainfall Fields Using a Non-Gaussian V-Transformed Copula, Advances in Water resources, 33 (6), 624-634.
Clark M.P., Slater A.G., 2006, Probabilistic quantitative precipitation estimation in complex terrain Journal of Hydrometeorology, 7(1), 3-22.

Answer: We will acknowledge these reference in the revised version.

8) P8: “additionally on a number of previously simulated hours of the corresponding realization.” The statement “A number of previously simulated hours” requires more explanation. It is not quite clear what the simulation is conditioned on.

Answer: We will reformulate this statement in the revised version (similar comment raised by Martinus van den Berg, Referee #3):
The multivariate simulation for realization j and time t is conditioned on:
a) rain gauge observations at time t and b) previously simulated realizations j at times (t-1,t-2, …,t-M), where M is the memory of the system (Page 3094, line 18). This means that the observed precipitation by rain gauges at times (t-1,t-2, …,t-M) is substituted by the simulated fields, which

encounter all the points of the whole simulation grid.

9) Figure 6: 10x10km (not km2) or a 100km2 grid

Answer: We would prefer to use “10km x 10km”, because it shows that the spatial dimension is squared, whereas 10x10km might confuse the reader.

10) P11: 100×100 km with a 10×10 km (instead of 100×100 km2 with a 10×10 km2)

Answer: Similar to answer 9): we will change it to”100km × 100 km with 10km × 10km”

11) P12: “The empirical variogram is estimated very well by the fitted spherical variogram model.” It is not clear on what basis the empirical variogram is estimated “very well” by the fitted spherical variogram. Avoid descriptive statement and include quantitative measures of goodness-of-fit.

Answer: We agree that this sentence is rather confusing and we will better completely leave it out. We will describe the evaluation of the goodness-of-fit (weighted least square fit, see Answer #2) in Sect. 2.2. However, we think that providing qualitative measures of the weighted sum of square error would not help the reader to demonstrate the differences between Fig.8a vs. 8b and Fig.8c vs. 8d, because the absolute value of the goodness-of-fit depends on the binning, which is different for the sampling from all grid points as compared to only rain gauge grid points.