Review of the manuscript titled "Geostatistical Interpolation by Quantile Kriging" by H. Lebrenz and A. Bardossy

The proposed manuscript presents a new geo-statistical interpolation method (Quantile Kriging – QK) that is able to relax three of the main assumption/limitations of the most used Ordinary Kriging: 1) spatial stationarity of the process mean, 2) Gaussianity of the interpolated variable and 3) independence of the uncertainty on the estimation value. The work extends the formulations of other well-known kriging methods with logic and statistical rigour. Although the presented technique still has a major limitation in the ability to handle the presence of many zero values (as often happens when dealing with rainfall, especially at finer scales than the presented one), it can be considered an improvement on the state of the art and a contribution to the advancement of the field. Additionally, although the authors do not mention it in the manuscript (and should) there are many applications to a variety of environmental variables where the presence of zeros is not a problem and the presented technique could be better applied. The manuscript is very well written and easy to follow. I suggest the following improvements:

1. The introduction explains a lot about the evolution of kriging techniques. However a little bit more discussion about applications (especially to rainfall) their limitations in hydrology, the main challenges, etc… could help defining the framework.
2. P.4 l. 18, many of the presented geostatistical techniques were developed in geological sciences, where the temporal evolution of the studied variables is often irrelevant. I would mention this to explain why the temporal variability is often ignored in kriging.
3. I would mention spatio-temporal kriging and other similar techniques as attempts to incorporate the temporal variability. How is this method better/different (e.g. Gaussianity)? Examples:
    - Snepvangers, J. J. J. C., Heuvelink, G. B. M., & Huisman, J. A. (2003). Soil water content interpolation using spatio-temporal kriging with external drift. Geoderma, 112, 253–271. https://doi.org/10.1016/S0016-7061(02)00310-5
    - Sideris, I. V., Gabella, M., Erdin, R., & Germann, U. (2014). Real-time radar-rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. Quarterly Journal of the Royal Meteorological Society, 140(April), 1097–1111. https://doi.org/10.1002/qj.2188
4. Eq. 5: I am not sure why you fit a Beta distribution to the quantiles: isn't the Normal Score Transformation (NST) designed to work with empirical distributions?
5. Eq.6 and Eq.7: You applied the NST, so isn't this E[F(U)] = m and same for Variance? Maybe I'm missing something
6. Nowhere is explained how you calculate the variograms for all the interpolations you do. Maybe worth mentioning it somewhere.
7. Pg 7 top: you introduce the elevation dataset, but you don't explain why. Mention you use it for both EDK of the parameters and for the reference EDK of the rainfall process.
8. is the dry ratio the number of stations that recorded zero rainfall over the whole month over the total number of stations? Can you state this a bit more explicitly?

9. P.7, l. 18: if you fit a PDF for each month for each station, you have only 22 points to do it, it seems a very little number to be statistically robust. maybe one of the reasons why you need to fit mean and variance rather than the parameters?
10. Eq. 8 and eq. 9: you here present both the distributions but don't explain why. Do you want to compare their performance? How did you select Gamma and Weibull distributions? Nowhere in the paper you comment on which one performs best overall.
11. P.8, l.27: You need to state that you do EDK with elevation as the drift. One of the problems I have in this comparison is that often EDK is performed with radar data, which probably would do better than elevation in defining the spatial pattern of rainfall. Can you comment on this?
12. P.9, l. 19: One of the drawbacks I observe is that QK does not estimate a higher uncertainty where there are less rain gauges, eg. top left corner of Figure 5f.
13. Explain what rho (eq. 12) represent, why you use it, what is its range, and what the optimal value)
14. P 13: I find the explanation about chi squared a bit confusing. I could not understand what had to be uniform and why, until later on you introduce the histogram. Maybe worth introducing the histograms first? or at least explain more in details.
15. Conclusions: You need to write more here, and remove one of the two paragraphs that are repeated (l. 20-26 or 27-3).
16. I feel in general a little bit more discussion of the overall results could be introduced either in the Results and Discussion or the Conclusion section, including many of the comments I mentioned before.