Response to reviewer 2

We thank the reviewer for the valueable comments. We provide a detailed point by point answer to the reviewers remarks.

The authors propose a new kriging technique developed to regionalize non-normally distributed spatio-temporal variables. The approach, as I understand it, involves (i) fitting time series observations to a predetermined distribution type at each gauge independantly; (ii) using the fitted distribution to assign a probability of non-exceedance to each observation at each gauge; (iii) fit the subset of these probabilities corresponding to each observation time to (different) beta distributions; (iv) map the fitted probabilities to normal quantiles; (v) use the (now normally distributed and *assumed* second-order stationary) outcomes as a basis to apply OK or EDK, as described (rather cryptically) in section 2.1.2. While the approach is intriguing, its description in the paper lacks statistical rigor and minimal proofs and intuitions. Its application in cross validation suggests that it does a decent job at predicting some measure of rainfall, but so would most properly fitted interpolation methods. This does not mean that the estimator is BLUE (best or even efficient unbiased linear estimator) of the considered stochastic process. Actually, the apparent correlation between Z*(x) and sigma2(x) (Figure 4) suggests that the process has some degree of heteroskedasticity. Under these conditions, linear models (like most kriging estimators) are not necessarily efficient (see, for instance, the Gauss Markov Theorem for ordinary least squares). Again, the approach is promising, but its exposition needs a major overhaul to be convincing.

The reviewer described the procedure we applied reasonably well. However note that the distributions fitted for each location are the same type (gamma for example). The parameters of the distribution have to be interpolated, this step is missing from the reviewers description.

An intuitive description of the procedure is based on the following properties of precipitation fields:

- The monthly (and daily) precipitation amounts for a given month often follow a skewed distribution.
- Monthly (even daily) precipitation amounts cannot be considered as stationary in space. Differences in expected precipitation amounts become clear if one considers long time accumulations.
- The precipitation generating meteorological processes are usually of large spatial extent. This means if there is heavy rainfall at one location it likely that other locations also have heavy rainfall.
- Correlations between time series of precipitation indicate a strong spatial dependence, while the spatial dependence of precipitation on a given time accumulations (day, month) usually show a much weaker spatial dependence.

A possible process model reflecting the above properties can be described as follows:

Let $Y_0(x,t)$ be independent (for each different t) normal stationary spatial fields with $E[Y_0] = 0$ and $D^2(Y_0) = 1$ for each t.

In order to reflect large scale meteorological processes the process $M(t)$ is introduced. High $M(t)$ values correspond to heavy rainfall covering the area - while low correspond to dry conditions. This M modifies the spatial process:

$$Y_1(x,t) = Y_0(x,t) + M(t) \tag{1}$$

Were $M(t)$ is a process (in time) with zero mean. We may assume that the distribution of $M(t)$ is normal. In this case for each x $Y_1(x,t)$ is normally distributed with $N(0,d)$ with $d = \sqrt{1+\sigma_M^2}$.

For each fixed t the distribution of $Y_1(x,t)$ is $N(M(t),1)$. Now Y_2 is temporal non-exceedence probability at location x - formally:

$$Y_2(x,t) = \Phi_{0,d}(Y_1(x,t)) \tag{2}$$

where $\Phi_{0,d}$ is the distribution function of $N(0,d)$. (By definition $0 \leq Y_2(x,t) \leq 1$.)

The rainfall is then *generated* as:

$$Z(x,t) = F_x^{-1}(Y_2(x,t)) \tag{3}$$

where F_x is the distribution function of rainfall at location x. The F_xs can be different for different x locations due to topography and other influencing factors. (These F_x-s can be interpolated - example see also in Mosthaf and Bardossy (2017)).

We use $Y_2(x,t)$ for each t and assume that it follows a beta distribution. In fact its distribution depends on $M(t)$. If $M(t) = 0$ for all t-s then monthly rainfall is fully characterised by independent realizations over space. In this case the distribution of Y_2 is uniform for each t.

This however is not the case with observed data. The reason is that wet and dry conditions occur simultaneously over the whole area. This is controlled by $M(t)$. One can take $M(t)$ for example as independent random variables or to follow an ARMA process. If $M(t) \neq 0$ then the distribution of $Y_2(x,t)$ for this t is not uniform. The reason for assuming it as beta was due to the fact that beta distributions are very flexible and can well describe distributions in $[0,1]$. The exact form of the corresponding distribution would be something like:

$$G_t(v) = \Phi_{0,1}\left(\Phi_{M(t),1}^{-1}(v)\right)$$

However the use of this would require the estimation of $M(t)$ for each t. We decided to use a simple beta distribution instead.

The introduction of $M(t)$ is reasonable as it explains the difference between the correlation between stations and the spatial correlation calculated using a variogram type approach for a given time. The later correlations are usually lower (smaller ranges) which are increased by the common large scale weather described with $M(t)$. Note that the introduction of $M(t)$ leads to a correlation of the precipitation time series even if the individual *snapshots* $Y_0(x,t)$ are independent in space.

In our procedure we start with $Z(x,t)$, estimate and interpolate F_x. Than calculate Y_2 for the observation locations. We interpolate Y_2 and come back to $Z(x,t)$.

Spatial variograms are calculated for Y_2 for each t, and Y_2 is stationary in space. Non-stationarity and non-Gaussian distributions occur only for Z. That is the reason why we concentrate on Y_2.

Non-Gaussianity: In its canonical form, Ordinary Kriging is based on the method of moment (i.e. variance minimization subject to unbiasedness) and so is not technically restricted to normally distributed processes. Gaussiannity is, however, required for maximum likelihood (ML) type estimations, which some studies have shown can be more efficient for kriging-based regionalization (e.g., Lark 2000). ML approaches become necessary to have unbiased prediction of both the mean and the variance when it comes to EDK or Universal Kriging, particularly when subject to more intricate error correlation structures (e.g., Muller 2015). Since non-gaussianity appears to be a key rationale for the proposed approach, it is important to be specific on that.

Non-Gaussianity is considered because of the usually skewed distribution of precipitation amounts for a given time step. The suggested model should enable a simulation of the precipitation amounts. Non-Gaussianity is only in the sense of the marginal distribution at a given time-step. The spatial dependenceis considered to correspond to a multi-Gaussian copula. This kind of transformation is frequently used - for example for Lognormal Kriging.

Beta-distribution: The use of the beta distribution is intriguiging, but more intuition is needed on the assumed underlaying stochastic process. Please describe clearly the properties of the (space-time) stochastic process that you assume gives rise to the observed sample and use that to demonstrate Eqn 6 and 7 in a rigorous mathematical proof. I find the use of the beta distribution promising because a common interpretation is that it describes the distribution of the probabilities associated to a binomial process observed over a finite sample. Lets say that the binomial process in question is the exceedance of a given threshold (as eluded to in the manuscript). Then, if the underlaying point process is identically distributed in space and if an identically sized sample is taken at each observation point, the proportion of observations lower than a given threshold across all gauges will be beta-distributed. Perhaps thats a start?

The above description shows that the beta distribution is only a convenient tool, not a statistically rigorous approach. As the beta distribution is very flexible it provided an easy and quick approximation of the distribution of the non exceedence probabilities.

Stationarity: More fundamentally, a main issue that I have is that your approach involves fitting distributions independantly at different points in space (gamma or weibull) and time (beta), which implies that the underlaying random point process follows a different distribution at each point in space. Granted, EDK and Universal kriging allow the first moment of the underlaying distribution to vary through space, as allowed by the scaling properties of the expectated value estimator. However, I am not aware of any existing geostatistical approach that allows for higher order moments to vary through space. This may be fine, but please demonstrate that your approach does not violate the second order stationarity assumption (i.e that the variogram is constant through space), which is critical (and arguably more important than gaussianity) for kriging.

The distributions fitted to the individual locations are supposed to have a spatial dependence. Further they are assumed to follow the same distribution. These distributions are then interpolated. Here we use the assumption that the distributions show a much clearer effect of the large scale rainfall generating meteorological processes than a sigle mothly (or daily) realization would show. Therefore the use of extrenal covariates , such as topography is more appropriate for this interpolation. The use of these distributions transforms the process to a staionary one. The stationary process is interpolated using the beta distribution of the non-exceedance probabilities. The reason for this is that we intended to avoid problems with interpolated probabilities being outside the $[0, 1]$ interval.

We intend to to revise the paper and to include the above description and discussions.

Mosthaf, T. and A. Bárdossy, Regionalizing non-parametric precipitation amount models on different temporal scales, *Hydrology and Earth System Sciences*,**21** , 2463-2481, 2017