**Comments on "Discrete k-nearest neighbor resampling for simulating multisite precipitation occurrence and adaption to climate change" by Taesam Lee and Vijay P. Singh**

Authors have addressed some of the comments satisfactorily. However, clarifications are needed on a few responses. I am highlighting those below.

**1.**

Present study attempts to develop a novel simulation method for multi-site precipitation occurrence, combining the k-nearest neighbor sampling technique and genetic algorithm. The coupled model has been applied in precipitation occurrence simulation in single sites. The (only) novelty probably lies in the application of this coupled technique in generating the multi-site precipitation occurrence. Authors may clarify these and may specify whether the novelty lies in the method deployed or in the application (See line 35 in the abstract and further such claims in the manuscript body).

*Reply: The authors appreciate this reviewer's insightful comment. The novelty of the current study is to propose the discrete version of KNNR-GA model in simulating multisite occurrence. The KNNR-GA model has been developed for multisite simulation of streamflow for continuous variables. The novelty of the current study is how to handle the multisite discrete binary process which is the main difference between the continuous version and the discrete version of the current study. The authors have improved the abstract and manuscript to emphasize this point. Hope this modification is satisfactory.*

In response to the general comment of highlighting the novelty of the work, modified abstract says "Multisite occurrence model with standard normal variate (MONR) has been used preserving key statistics and contemporaneous correlation in literature, but it cannot reproduce lagged crosscorrelation between stations and long stochastic simulation is required to estimate its parameters. Employing a nonparametric technique, k-nearest neighbor resampling (KNNR), and coupling it with Genetic Algorithm (GA), this study proposes a novel simulation method for multisite precipitation occurrence overcoming the shortcomings of the existing MONR model." **This sounds as if the focus of the study itself is only to overcome the limitations of MONR model. The novelty (if any) is still not brought out clearly.**

**2.**

In addition, the present method is compared with a method (MONR) which is developed almost two decades back. Is MONR a frequently used method for multi-site precipitation occurrence simulation? It would be convincing to compare the present technique with more recent methods deployed for multi-site precipitation occurrence simulation. More specific comments are provided below for the kind consideration of the authors.

*Reply: The authors appreciate the reviewer's insightful comment. Even if MNOR model is rather old-fashioned, this model has been popularly employed in this field and its performance is more comparable to the Markov Chain model especially in multisite occurrence cases of precipitation dataset.*

**A few recent studies are given below on the same topic, which focus on the same topic – multi-site precipitation occurrence.**

Evin et al., HESS, 2018: Stochastic generation of multi-site daily precipitation focusing on extreme events

Mehrotra et al., JH, 2006: A comparison of three stochastic multi-site precipitation occurrence generators

**3.**

1. Line 68 – 74: Wilks (1998) model assumes standard normal variate and underestimates the lagged cross correlation. As mentioned before, is it really worth to compare the present method to this model, which works on an entirely different hypothesis? As mentioned by the authors in the next paragraph (lines 75-81), KNNR and KNNR-GA are proved to be efficient. Won't it be better to compare the present model (DKNNR) to compare with the above model, to highlight its applicability in multi-site precipitation occurrence, given that the novelty of the study is claimed to be in this application.

*Reply: The authors appreciate the reviewer's insightful comment. The MONR model is the model of Wilks (1998) and it has been popularly employed in the literature. The present study compared the discrete version of KNNR-GA with the model of Wilks (1998), named as MONR here. See the first line of the section 2.2 as the following:*
*"Wilks (1998) suggested a multisite occurrence model using a standard normal random number (here, denoted as MONR) that is spatially dependent but serially independent."*

**Please clarify how the results would be different for DKNNR and KNNR models?**

**4.**

2. Line 78-81: It is mentioned that KNNR model cannot produce different patterns and coupling with GA solves this drawback. Please provide more details on how GA could possibly solve this. And how the application of GA could ensure generation of similar populations. It would be interesting if some physical sense can also be provided here – how possibly GA could simulate those system behavior?

*Reply: The authors appreciate the reviewer's detailed comment. Further explanation is added in the manuscript to improve the clarity in the result section.*

**The authors have explained the need for GA in the methodology, to simulate the patterns different from the historical patterns. This is understood. However, it is not clear how GA will be trained to generate those patterns specific to the study area. I am sure that GA might generate many unwanted patterns also, which is not physically possible in the study region. How GA is supposed to avoid this unwanted patterns?**

**5.**

4. Line 158: When the algorithm will select the GA mixing? What is the criterion for GA mixing in the procedure?

*Reply: The authors appreciate the reviewer's insightful comment. It is subjective. If one wants to simulate the dataset as the same observed pattern, this procedure can be skipped. Otherwise, the GA procedure gives the benefit of generating new patterns that we already discussed under comment 2. The sentence is modified accordingly.*
*"Execute the following steps for GA mixing if GA mixing is subjectively selected. Otherwise, skip this step."*

**So, is it up on the user to opt for GA mixing? It should have been based on the properties of the time series and study region. If the rainfall exhibits more or less an unchanging pattern across the stations, then the future pattern can be found in the historical patterns too. In that case GA mixing could be avoided. The algorithm should have the criterion for that.**

**6.**

6. Section 3.2: Authors must be pointing towards "Dealing with Non-stationarity" than "Adaptation to climate change". It is clear that only changes in marginal and transition probabilities are been considered, by tuning the crossover and mutation probabilities? "Climate change" may refer to a larger phenomenon, which might not be addressed directly in the present study. Please explain.

*Reply: The authors totally agree with the concern of the reviewer. Tuning the crossover and mutation probabilities only affected the marginal and transition probabilities. This limitation must be addressed as this reviewer commented. We added the following to address the*

**Thanks for agreeing to this comment. In that case, there is an over-emphasis in the title regarding the "adaptation to climate change". If the methodology is not addressing the climate change, please remove the section or modify it accordingly. Section 5.4 still claims "Adaptation to climate change". This can be addressed along with the next comment (7th comment), where again authors justify the changing of these probabilities to address the climate change. It is not clear, how tuning of crossover and mutation probabilities could handle the non-stationarity (or climate change according to authors) in the time series of multiple stations?**

**7.**

15. It would be interesting to see the results generated by the simple KNNR model in this application. Also, it would be helpful, if you may please explain how the incorporation of GA possibly helped in modeling the physical laws of the precipitation system.

*Reply: The authors appreciate the reviewer's insightful comment. We produced the results without the GA process as presented in the following (See Figure S2-Figure S6). The presented results show that no significant difference from the one with the GA mixing can be found. The following is discussed in the manuscript right before the results of the probability selection (section 6.1).*

**I could not find much difference between simple KNNR model and KNNR model with GA mixing (Figures s2-s6 and Figures 5-9). Both produce almost same results. Does that mean, the incorporation of GA has not added much value?**

**8.**

17. Section 6.3: I am a little confused here. How can the parameters be changed in the future, for the model to adapt to the future changes, given that we may not clear information about these changes?

*Reply: The authors appreciate the reviewer's comment. The authors did not fully investigate the specific changes required to be made for specific climate change assessment at this stage. As mentioned under comment 7, the focus of the current study is to propose a novel approach that simulates multisite occurrence process through nonparametric approaches. Further development for adopting to climate change and its application are partially presented as a possible improvement of the proposed model in the near future and will be presented as a separate work as explained in the conclusion. This limitation and possible development are discussed in the last section.*

**Please see comment 6 in this document, regarding the adaptation to climate change.**

Kind Regards.