Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-181-RC2, 2018

Geoscientific Model Development Discussions

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

Anonymous Referee #2

Received and published: 31 December 2018

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). While, stochastic weather models (like the one deployed in this study) are commonly deployed in various applications, it would be preferable to give some physical justification to the application and comprehend the results obtained. This would bring more confidence

into the purely statistical methods which otherwise may not have captured any physical relationships/behavior of the system been dealt. This is particularly significant in the present study, since multi-site occurrences might be directed by many climatic feedbacks and also controlled by many local factors also. Absence of any such physical explanation may leave the methods sound robotic and put doubts in its generic applicability. 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.

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.

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?

3. Line 142: "multisite occurrence X and the observed multisite occurrence x". Aren't both these variables multi-dimensional and of same size? It would be ideal to denote both in capitals then.

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

GA mixing in the procedure?

5. Line 178-179: It is mentioned later in the manuscript that the changes in the mutation and cross-over probabilities may be carried out to adapt to the changes in the transition and marginal probability distributions (See lines 187-188). Considering that, would it be ideal to fix these as 0.01, following Lee et al (2010b). Shouldn't this be case specific? If not then, the later statement (lines 187-188) are questionable.

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.

7. How tuning of crossover and mutation probabilities could handle the non-stationarity in the time series of multiple stations? Can the model change these parameters in between the time frame of the simulation, so as to incorporate the parameter change(s) in the probability distributions?

8. Section 4: Please provide more details about the precipitation data used, its seasonality, rainy day characteristics etc. Are the stations selected meteorologically homogenous?

9. Section 5: This may go into the results section, if it sounds fine.

10. Line 222: " ....., since a synoptic scale weather system could result in lagged cross-correlation" – Can this statement be generalized for all locations?

11. Figure 2-4: Ensemble means from MONR are close to the observed mean, than those of DKNNR model. Is MONR better in that sense? Please clarify.

12. Line 254-255: "Even though the transition probabilities were not employed in simulating rainfall occurrence, the DKNNR model preserved this statistic fairly well" – Is it merely by chance? Please provide justification to build confidence. Do you expect the

results to vary, when deployed in different regions?

13. Line 273-274: "Precipitation is not significantly correlated with more than one day" – Please provide reference. The statement may not hold well globally, as Box-Jenkins models of higher order are often applied for simulating precipitation events.

14. It would be better to number the stations considering its proximity. It will help in analyzing the results.

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.

16. Disadvantage of the simple KNNR model is the inability to simulate different patterns from the observed series. Do the stations selected exhibit significant non-stationarity? If not, will the KNNR model also serve the purpose?

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?

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-181, 2018.