Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC2
https://doi.org/10.5194/gmd-2022-163-RC2, 2023

Comment on gmd-2022-163

Anonymous Referee #2

Referee comment on "Deep learning for stochastic precipitation generation – deep SPG v1.0" by Leroy J. Bird et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-163-RC2, 2023

The manuscript by Bird et al. presents a deep learning based stochastically generation of daily and hourly precipitation time series in New Zealand. The Authors used a neural network framework and probabilistic mixture distributions (i.e., gamma and GPD) to simulate the precipitation intensity at daily and hourly timescales. A list of statistics is shown to examine the synthetic time series and model performance against the observations and future climate projection. The Authors also present a non-stationary version of their deep learning model that incorporates contemporary/future changes in the precipitation traces through a temperature covariate. Overall, this manuscript contributes by adapting a neural network approach to stochastically develop multiple ensembles of precipitation at the regional scale. My comments/questions are listed below that are mostly related to the model structure, performance assessment, and precipitation-temperature sensitivity parts:

1) What was the reason for making the location parameter of GPD zero (L200)? Also, were the parameters estimated at a seasonal scale, or any other specific considerations applied to not mix up different rain-generating mechanisms into the same distribution? I presume a threshold or set of thresholds should also be considered to distinguish between the mixture distributions (lower quantiles vs. heavy tail-like behavior); how did you set the model to learn about these limits?

2) I was expecting that the developed stochastic precipitation generator model could also produce traces of precipitation reasonably higher or lower the observed time series across different ensemble members. However, it seems any values greater than the observed maximums (L249-252; Table 5) are replaced with equal or smaller magnitudes than the recorded maxima at daily and hourly timescales.

3) I would strongly suggest providing a few quantitive measures for the model comparison (e.g., percent bias, index of agreement, and so on) in addition to showing the model

performance q-q plots in the "Stationary quality assessment" section for different precipitation characteristics, including moments, spells, extreme values of simulated traces. It is unclear how the stochastic model (e.g., different ensemble members) performs against the observations by only looking at these q-q plots.

4) How does the stochastic model preserve the year-to-year variability of seasonal/annual precipitation variation? Can the Authors compare the intra-annual variation between the simulated traces and observed time series?

5) precipitation-temperature sensitivity: When I look at the P-T sensitivity plots, it seems negative scaling rates are calculated in many stations. Note that the Clausius–Clapeyron equation demonstrates a 'positive' relation between an increase in temperature and rainfall when the atmosphere is (nearly) saturated. As an example, it seems there is a strong 'hook' structure in the P-T relation across your stochastic precipitation generator/observation/model precipitation time series, and it should be taken into account before calculating the rates.

6) "5.3 Post hoc addition of non-stationarity": It is unclear whether the established P-T relationship is valid given Figures 16-17 and Comment#5 above.

Additional comments:

L117-124: I doubt using a single weather@home grid cell is a robust approach here in general; how about using at least the four nearest grid cells to the site/station? This way, the failure rate of encountering a 'bad' cell will statistically decrease fourfold.

L183-186: What computational settings were used to execute this neural network and generate the traces? Please add some details about the logistics/computing configuration requirements.

L196-203: Were these 12 parameters estimated for each site without considering any seasonal influence on precipitation distributions? e.g., cold vs. warm season.

L221-222: It is unclear what "better numerical stability" means here.

L234-235: Is it causing a problem if the first eight days or 144 hours happen to be all zeros?

L312: What is your hypothesis on the hourly SPG that underestimates the seasonality in the proportion of dry hours for Auckland?

L321: Is this the entire section of the “4.5. Discussion” following the “4. Stationary quality assessment”? I feel this “4.5” section does not add any new information more than what is presented in “4.1” to “4.4” sections.

In general, it seems some figures can be merged, or removed, as they do not add any new information.