Comment on egusphere-2022-282
Anonymous Referee #2

Referee comment on "CCdownscaling: an open-source Python package for multivariable statistical climate model downscaling V1.0 " by Andrew D. Polasky et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-282-RC2, 2022

The authors proposed an open source python package (namely CCdownscaling) for climate model downscaling. The package includes the self-organizing map (SOM), two random forest-based methods, and the quantile mapping approach, which have all been used for downscaling for many years. All the methods including SOM was trained and tested at one location in Chicago, Illinois. In general, the manuscript presents a case of software package development, but not geoscientific model development, which seems to not match with any one of the six aims of the GMD journal. It does not improve current downscaling methods or models, or present a detailed, rigorous descriptions, and evaluations of any new methods. I provide specific comments as follows.

Major issues:

- The novelty of this study is not justified. This study focuses on develop software package for downscaling by packing up a few existing empirical downscaling methods, but does not improve any downscaling techniques, or presenting any new method. We know SOM is an existing downscaling method that have been used for downscaling for many years. It is also not clear about the advantages of SOM compared to the other downscaling approaches. Is it worthwhile to develop such software package?

- There are several existing downscaling packages that did not consider in the manuscript, such as ClimDown in R: https://rdrr.io/cran/ClimDown/, and SBCK in both R and Python: https://github.com/yrobink/SBCK. These two packages include many downscaling and bias correction approaches that have been fully evaluated in the upstream research papers. What is the added value from this CCdownscaling package?

- It is not clear about the advantage of SOM method compared with other approaches for climate downscaling. The authors need to justify the value for developing a software package for SOM.

- The validation does not include any extreme events or indices. The validation results can be highly uncertain since it relies on the data from a single site. Sufficient number
of sites covering different conditions are needed for robust evaluations.
- The downscaling method only considers four variables: relative humidity at 850 hPa, air temperature at 850 hPa, geopotential height at 850 hPa, two wind speed at 700 hPa, while so many physical covariates of precipitation are ignored. Note there are many precipitation covariates available in different reanalysis datasets as well as GCMs simulations.

**Minor issues:**

- In Section 4, the authors mentioned the SOM advantage of providing insight into the weather patterns giving rise to specific downscaled outcomes through pattern detection. But the explanation about pattern detection is not clear. Figures 2 and 3 are lacking context. The node example (0,4) is not clear. More notation and explanation are needed for pattern detection.
- The performance of SOM approach is similar to the basic quantile mapping approach (see Figure 6 and Figure 7 as well as the two tables). What could be the reason for that and the potential avenues to improve SOM performance?
- Computer code should be put in the appendix, not be part of the main text.
- Figure 6 is not a histogram