Comment on gmd-2022-14
Anonymous Referee #1

Referee comment on "Repeatable high-resolution statistical downscaling through deep learning" by Dánnell Quesada-Chacón et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-14-RC1, 2022

This paper evaluates a suite of related Deep Learning models for downscaling precipitation data over the Ore Mountain region in Germany. In addition to evaluating U-Net and U-Net++ architectures with a number of different options, the authors also explore the use of containerization to enable repeatability of the experiment, and provide a Singularity container with the associated code and environment.

This paper is well-written and of interest to the readers of GMD, and the topic is highly relevant and of value to the scientific community. I have only minor revisions to suggest. It's a good paper, and I enjoyed reading it.

Good work!

Suggested revisions:

The manuscript has too many novel abbreviations. I recommend replacing all of the following with their full expansions, both in the text and in figures:
- CC - climate change
- TF - transfer function
- OM - Ore Mountains
- EOM - Eastern Ore Mountains? (This abbreviation is never defined)
- SN - Saxony
I recommend changing the following abbreviations:
IB - international borders - these are obvious from context and can be omitted from the legend
m.a.s.l - meters above sea level - I would call this simply "elevation (m)"

Consider also expanding:
SD - statistical downscaling
DL - deep learning
Although these abbreviations are not unknown, I think the text would read more clearly without them.

Line 107: change "the Fichtelberg with 1215 m.a.s.l and the Kahleberg (905 m.a.s.l.)" to "the Fichtelberg (elevation 1215 m) and the Kahleberg (elevation 905 m)"

---

URLs should be listed as part of a reference, not included in the text of the paper. For example, on line 132, you could simply write "The code needed to recalculate these results can be found on GitHub (Quesada-Chacon, 2022), with all the modifications..."

Please replace the URLs on lines 113, 132, 198, 212, 214, and 217 (and any others I may have missed) with citations.

---

In section 3.2, please add a definition of the Bernoulli-Gamma loss function used.

---

Line 165: A sentence or two discussing why batch normalization and spatial dropout were included as options (i.e., what effect they have and when one would want to use them) would be valuable to the reader.
Although many writers use the passive voice in scientific writing, active writing using the first person is much clearer and easier to understand. The authors could improve the paper by switching to active voice throughout; for example, on line 123, replacing "The reanalysis dataset employed as a predictor is ERA5..." with "For the predictor, we used the ERA5 reanalysis..." makes the text much easier to follow. Since this would require extensive editing, I do not expect the authors to make this change, but mention it merely for the sake of future papers.

---

Figure 4:

I think this figure would be easier to understand if the green-to-orange color bar were reversed, with orange indicating lower values (worse performance) and green indicating higher values (better performance).

---

Line 294: rather than giving Lat & Lon coordinates, I think it would be clearer to point to this location as the southeast corner of the region.

---

Minor corrections:

Line 59: change "1 500" to "1500"

Line 120: change "which leads to a number of 1916 pixels" to "giving a region with 1916 pixels"

Line 123: change "with a spatial resolution" to "which has a spatial resolution"
Line 209: I think you want "interface" here, not "interfere"

Line 225: change "functions" to "function"

Line 227: change "improved significantly" to "significantly improved"

Line 252, 289: change "yet," to "although" (no comma)

Line 323: change "observations" to "failures" (not a correction, but clearer)

Line 334: change "Besides" to "In addition"

Line 380: change "Yet" to "However"

Line 500: "https://doi.org/" has been doubled in this URL