



EGUsphere, referee comment RC2
<https://doi.org/10.5194/egusphere-2022-497-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-497

Anonymous Referee #2

Referee comment on "DeepPrecip: a deep neural network for precipitation retrievals" by Fraser King et al., EGU Sphere, <https://doi.org/10.5194/egusphere-2022-497-RC2>, 2022

This manuscript describes an innovative machine learning (ML) approach to estimating surface precipitation from vertically pointing Doppler radar data, plus ancillary atmospheric state information from numerical reanalysis. One major strength is that the authors use the surface-based MRR to provide a vertically complete set of radar observations in the lower atmosphere, and then seek the key altitudes that provide the most information for this new approach, rather than immediately starting with the surface clutter-limited satellite radar data and struggling with what might or might not be hidden in the resulting blind zone. A second major strength is that the authors work to present the basis of their method's improvement, rather than simply reporting that the new scheme is better. The manuscript's major weakness is that terminology sometimes just appears without explanation; this is decidedly not reader-friendly, particularly for a study that pushes beyond "standard" ML practice; in some cases I'm not sure the descriptions even look expert-friendly. Also, while the code is accessible, and the manuscript gives useful insights into the choices made, it misses an opportunity to step the reader through the processing stream, exemplified by Fig. 2b, and explain where in the stream the various concepts in the text fit.

Major Items

- *Precip phase*: Sims and Liu (2015) show a simple scheme for estimating phase – it really depends on wet bulb temperature (T_w), not T , and you've already got the ERA5 data in hand to make the calculation.
- *Undefined magic operations*: These need a reference at a bare minimum, preferably with a 1-line "what they do" statement. I can cite: L.161: Adam optimization; L.181: squared errors elbow criterion method; L.214: using dropout; and I would encourage the authors to review the manuscript for other such mystery names.
- *Connect text to the flow chart in Fig. 2b*: In particular, it is not clear how Section 3.4 fits into the overall processing workflow.
- *Undercatch*: The text keeps alluding to gauge errors, particularly for snow, but never

really confronts the beast. Essentially all operational gauges bias low, most acutely for snow. This low bias affects the statistics, and there ought to be at least an organized, if short qualitative statement about this issue.

Minor Items

- *References in the text:* These are mostly in a form that is non-standard in my experience (no surrounding parentheses, such as L.16, but not always).
- *L.50-52:* Stating the references here is unnecessary; just point to Table 1.
- *L.57:* "NRT" is almost always "near-real-time". This is the only place where it's used, so just say "post-real-time" and don't give an acronym.
- *L.70-82:* The phrasing is awkward; I'd describe the standard situation first and then summarize the deviations.
- *Write out all acronyms:* TMP, WV, Ze, S/R, P, RF, MLP, CNN, perhaps plus others.
- *L.84-87:* The language can be simplified.
- *L.122-123:* If Cold Lake and Ny-Ålesund are excluded, what *is* done there?
- *L.183:* "event-types" would be more descriptive as something like "intensity classes". And, in subsequent discussions, the same terminology should be used whenever intensity classes are referenced.
- *L.203-206:* Kindly eliminate the redundancy.
- *L.214:* This is the only mention of mixed phase. If it's going to come up here, it needs to be introduced back when rain and snow events are introduced and related to Fig. 4.
- *L.219-221:* This feature isn't really discernable in Fig. 5.
- *Figs. 4a, 5a:* I would suggest bolder separation between the different sites so that it's easier to see them as separate time series. Also, the graphic needs to be larger. Finally, the horizontal time axis needs better labeling showing time increments.
- *L.276:* I think it's "In all other cases, ..."
- *L.280:* In addition to SPW, the abbreviations for reflectivity and Doppler velocity should be defined here.
- *Fig. 8:* In addition to labeling for bin, the Y axis also needs labeling for height, since this is referenced in the text.
- *L.301:* "assumption" isn't quite the right word; "speculation"?; "prior inference"?
- *L.305-308:* This last sentence needs to be more straightforward. I think you're suggesting more aggressive use of the un-blanked, or at least, less generously blanked, satellite radar data.
- *L.323-324:* The name usually includes "mission" as "Global Precipitation Measurement (GPM) mission".