The manuscript entitled “Downscaling Multi-Model Climate Projection Ensembles with Deep Learning (DeepESD): Contribution to CORDEX EUR-44” applied a Convolutional Neural Networks (CNN) to downscale present-day and future GCM simulations, with a focus on precipitation and temperature simulation. The paper falls within the scope of the journal. The manuscript is well structured, the method is generally clearly presented, and the results fully support its conclusion. I have a few specific comments regarding the method and analyses.

One common problem I found in the manuscript is the use of specific terms that might be well known in the Deep Learning area but is unfamiliar to me, a GCM and RCM modeler, and people like me. I suppose the authors may want to broaden their impact not only in the DP but also on people working on dynamic modeling. Below I have listed a few:

Line 32: “perfect prognosis” please explain the term in detail

Lines 35-36: please include 1-2 sentences to introduce dynamic downscale as a comparison to statistical downscale

Line 61: Please provides more information about “E-OBS v20”

Line 62: I assume “—” is a typo?

Line 70: can you explain the “harmonization process” further?

Lines 75-77: The method of DL should be explained in further detail. “They consist of … (one per each gridpoint in E-OBS)”. I found many terms in these sentences that might be barriers to fully understand the method. Can you rephrase it?

Line 97-98: Merge the single sentence to the following paragraph

Line 120: It is interesting that the DeepESD has the smallest ensemble spread over the historical period (Fig. 2) but has the largest one over the future. Any explanation for that?

125: what do “these differences” refer to?

Line 130: “These differences are quite systematic for the case of precipitation indicating a robust CNN extrapolation fingerprint.” Can you explain in further detail?

Figure 4: How do you produce Row 3? Is it the difference (DeepESD minus E-OBS v2.0) shown in each model in row 2 minus the mean of the difference? In that case, the mean of 8 panels in row 3 should be 0 ($\overline{Diff} - \overline{\overline{Diff}} = 0$), but clearly they are not.