Deep-learning statistical downscaling of precipitation in the middle reaches of the yellow river: A residual in Residual Dense Block based network:

The authors have used a residual based approach to downscale precipitation over the reaches of China. The residual-based approach significantly improves on traditional machine learning approaches when compared to observational data.

While there are many aspects of the research that are useful in this manuscript, there is little novelty in this approach and many aspects of the research have been widely investigated elsewhere. The authors have many redundant figures, and the manuscript could be condensed significantly. Additionally, for your evaluation for extremes you only use 4 years of data. I suggest doing another experiment, with a larger test set.

More epxeriments (i.e larger ensemble).

Concerning how much noise there are in the downscaled outputs.

Title:

I suggest changing your title to the following or something similar:

Deep learning-based downscaling of precipitation in the middle reaches of the yellow river using residual networks

The title is a bit clunky.

Abstract:

Line 5: “good performance on precipitation simulations”, I’d be explicit and provide some indication of improvement in %. Also, I’d say “when evaluated against observations we obtain X % of improvement relative to linear methods”.

Introduction

Line 20:” Over simplified parameterizations” I’d remove this text. The main reason for downscaling is the resolution issue. Your research is not on parameterizations.

Line 35 & Line 40: “tricube methods”- what is tricube methods?

Line 45: “Say that capturing non-linearities is important for downscaling variables such as precipitation”. This is a good reference (https://www.sciencedirect.com/science/article/pii/S2212094722001049).

Line 50: Bano Medina et al., (2020) is not about forecasting it is about downscaling. If you use that reference, how might need to clarify this.

Line 50: (Pan et al., 2019, Bano Medina 2020) add Sun et al., (2021) and Rampal et al., (2022) . These are also important references. https://www.sciencedirect.com/science/article/pii/S2212094722001049 and https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/joc.6769. You also need to add other recent papers by Bano Medina et al. (e.g. 2021, 2022). I don’t think you’ve summarized the recent literature too well.

Lines 55-60: Here you need to state what contribution your work makes to the overall literature. It's not very convincing about the value added of your research here. Something like "Our work expands on the existing literature by incorporating X y and Z". You could say that we also downscaling the temporal variability which has not been considered before.

Study Area

Line 70: Maybe describe why you chose to coarsen ERA5 to 2° resolution (i.e. to be consistent with Medina et al., (2020).

Methodology:

Lines 85-90: You need to cite other work that as used similar methodology here. For example, Cannon et al., (2008), Rampal et al., (2022), Sun et al., (2021), Bano Medina et al., (2020, 2021, 2022).

Table 1: I personally think this should go in a supplementary section).

Equation 1: Also should go in a supplementary section.

Figure 4: This should go in a supplementary or be combined with Figure 3.

Lines 125: Clarification: is the paper downscaling to daily precipitation? If you are discussing the BG distribution, make sure you cite Rampal et al., (2022) and Sun et al., (2021).

Lines 135: Again are you performing temporal disaggregation? Where for a daily input you are predicting hourly or sub-daily output?

Equation 6: This could be combined in one expression with equation (2).

Lines 170: Are you using any regularization, as these methods can overfit very easily? Please clarify if not.

Lines 180: These metrics are commonly known, so we do not need these in the text. If you'd like to include them add them to the supplementary section or appendix.

Results:

Lines 185: This is a very small validation sample (4 years), which makes it challenging to examine the performance on extremes. I'd recommend having a supplementary analysis with a longer validation dataset.

Lines 185: "downscaling projections" – you are not downscaling projections?

Table 2: I think some measure of uncertainty is required perhaps. I'd suggest repeated the experiment 20 times (with different random seeds) and investigate whether your results are statistically significant. This is importance, as this is the premise of your entire results.

Table 4: I don't think table 4 is useful, this could easily go in the supplementary section.

Figure 6: This seems excessive, I suggest making a (2 x 5) plot and combining this with Figure 5. Where the columns are the models (e.g. GPM, GLM) and the row is the climatology and bias.

Figure 5 & 6. It seems a little concerning that there is so much noise in your predictions from the GLM, CNN and other ML models. Other papers do not show such “noise”. Are you training your models enough or too much and that your models are overfitting? Some clarification on why there is more noise is needed.

Figure 7: Plot the bias instead of the raw amounts. Reorder axes from DJF, MAM, JJA, SON.

Figure 8: Not useful, this could go in the supplementary instead.

Lines 230: Use either R99P or R95P in the analysis, not both. If you’d like you could keep one in the supplementary section.

Figure 9: Frequency at 100mm is very large in Figure 9b and 9d, is there something wrong in your analysis, this is very concerning?

Figure 10: Again only one of the R95 and R99 plots should be plotted. You should also compute the bias of the R95 fields against the observations (the percentage bias) in one single plot. It seems strange that you have so much noise in your plots.

Figure 11: Again combine with Figure 10.

Table 5: This could be in the supplementary section or combined with information in Figure 10.

Figure 12: Not required, and again very interesting why the outputs are so noisy.

Figure 13: Not required, could be supplementary.

Table 6: Not required, could be supplementary.

Figure 16, should be in the supplementary section.

Figure 15: This is not a commonly used validation metric. I would validate against the RX1Day (wettest day of the year).

Figure 14: This should be the bias instead of the average precipitation.

Table 7 & 8, against too much information, this should be in the supplementary section.