

General comments:

In this paper, the author built a statistical downscaling model based on (Convolutional Neural Network) CNN. And used the reanalysis data as a reference for the training phase. Then, the author employed the trained model to downscale the ensemble of CMIP5 models over Europe. The results were compared with the CORDEX RCMs for the historical (1975-2005) and the projection (2006-2100) period. The results from the deepSD algorithm are contributing to CORDEX initiative, which is a breakthrough since the statistical downscaling based on artificial intelligence was not trusted for climate studies in the past few years. In general terms, the manuscript is well structured, the methodology is well described, the figures and tables are well organized and the results are adequately discussed. This paper falls within the scope of this journal. In this sense, the article can be approved after minor revisions.

Recommendation: Accept after minor revisions.

Specific comments:

Abstract:

The sentence “To our knowledge, this is the first time that CNNs have been used to produce multi-model ensembles” is not that accurate since there are previous studies that employed CNN to downscale the model ensemble (e.g., Babaousmail et al. (2021)).

Introduction:

1. Introduction (third paragraph): “These methods are not computationally demanding...”
This sentence needs a citation.
2. The author should justify why he selected the RCP8.5 scenario out of the other scenarios?
Also, was there any method employed for the selection of the 8 GCMs?
3. The author didn't justify why E-OBS v20 was selected as an observation in this study.

Data and methods:

1. Since the author is comparing the ensemble resulting from 8 GCMs with the RCM ensemble projection, shouldn't be the number of GCMs equal to the number of RCMs?
2. Usually, when we train a neural net model, a validation phase is required after the training and it should be selected from the historical 25 years period, in this paper the author didn't mention it.
3. Concerning the CNN algorithm, we noticed that the CNN used to downscale precipitation has one more layer than the one for temperature (one output layer). Can the author explain the reason?