This study presents a comprehensive assessment about the added value of 2m minimum and maximum temperature dynamically downscaled regional climate model (RCM) simulations from EURO-CORDEX initiative. To quantify and spatially characterize RCMs performance compared to the corresponding lower-resolution global scale driving fields, Authors take advantage of a distribution-based metric (DAV) previously introduced and presented in Soares and Cardoso (2018). The evaluation regards all the available ERA-Interim reanalysis and global climate models (GCM) driven RCM simulations corresponding to the Hindcast (1989-2009) and Historical (1971-2005) experiments respectively. All the simulations considered refer to the Iberian Peninsula domain and an observational-based Iberian Gridded Dataset (IGD).

As already mentioned for the precipitation-based part-I of the study, the present research involves a relevant research question namely if and eventually at what extent downscaled simulations can improve the large-scale forcing signal. This represents a very important point as RCMs are extensively used by a broad range of end users belonging to climate impacts and climate services communities.

The main value of the study is to consider the largest dataset of RCMs available and to consider a simple and straightforward metric identifying RCMs potential added value over the entire statistical distribution.

It follows some line-specific, minor remarks:

Line 33-35. This statement is not clear. Please rephrase.

Line 46. I would remove “or“.

Line 266. I would use “largest differences” instead of “most range”.

Line 301. “span” instead of “spam”.

Line 301. The statement: “all PDFs still reveal a close representation to each other” is not
clear. Do they present smaller variability across different RCMs?

Line 375. “The interpolation of the IGD causes a slight deterioration of the PDF, particularly for the extremes”. It is not clear to which is referred the deterioration mentioned.

Line 376. Please remove the comma.

Line 413. “For the heat and cold extremes, the results are more limited, namely for TASMIN.” This sentence is not clear, please rephrase it more clearly.

Lines 426-428. This is a very relevant point. The added value of downscaled RCMs seems, to some extent, clustering as function of driving GCM, though differently if we consider the entire distribution or only PDF’s tails. Here I would add a few further considerations about why the added value of RCMs can depend on the GCM considered. Here we are not considering the well-known dependency of RCM capability of reproducing observed values as function of the GCM “quality” but about the RCM capability of improving GCM signal. The fact that this feature is sometimes more GCM- than RCM-depending it is a relevant