Answer to **RC1**: 'Comment on nhess-2023-40', Anonymous Referee #2, 9 Sep 2023

*** General comments ***

(1) Although the authors clarified that they do not claim to have developed a calibrated model, I think that this fact should be clearly stated in the text. Hence I suggest to add a sentence in Section 5 and/or Section 6 like "Although AR-ChaMo produces probabilities for large hail and lightning, these probabilities are not calibrated. This implies that the computed probabilities do not necessarily coincide with the observed frequency of the phenomena but the model might nevertheless give a valuable guidance on occurrence of these hazards."

Authors: Thank you for pointing this out. We added this point to Section 6: "Finally, although AR-CHaMo produces probabilities for large hail and lightning, these probabilities are not calibrated. This implies that the computed probabilities do not necessarily coincide with the observed frequency of the phenomena.

Apart from these limitations, the models give a valuable guidance on occurrence of these hazards and represent an improvement compared to state-of-the-art composite parameters in Europe."

(2) I appreciate the interpretation of the meteorological situation at the end of Section 4.1 and in Section 4.2, however one should be careful to explain the behaviour of an ensemble by analyzing the deterministic run (or a singe ensemble member). The issue here is that each ensemble members has by definition its own realization and a vertical profile might look quite different between the members. In my opinion this should be communicated to the reader to provide a word of caution with this sort of analysis. A more probabilistic way of doing the same analysis would be to analyse the mean or median profile together with an indication of the ensemble spread of the profiles (e.g. by plotting the standard deviation or a difference of quantiles).

Authors: We clarified this in the text in Section 4.2: "Although this analysis provides a physical explanation of the different lightning probabilities, it is important to note that here we considered vertical profiles from the two deterministic runs only while the profiles of the single ensemble members could differ significantly, for instance in the realization of mid-level moisture, to the deterministic ones"

*** Specific comments ***

Line 83:
I think it should read "grid" instead of "grid box".

Authors: Adapted accordingly.

Line 115:
Maybe it should read "grid" instead of "box"?

Authors: Adapted accordingly.

Lines 131 to 134:
Indeed, but a common strategy is to train separate models per leadtime. From this sentence I think that you use a single model and apply it for all leadtimes. Please add a sentence that clarifies that to the reader.

Authors: The sentence has been rephrased to the following: "By training a single model for all lead times, if for instance two hail reports occurred on Tuesday at 12:00 UTC and on Wednesday at 18:00 UTC, it would not be possible to qualitatively compare the environmental conditions at these two time steps. This is because for a forecast initialized on Tuesday at 00 UTC, the environmental conditions associated with the Wednesday at 18:00 UTC report would be subject to larger uncertainty due to the larger lead time.