Chers Collègues,

I have received three short reviews of the revised version of your paper. They have been submitted by the same three referees of the first version, who had all recommended minor revisions of the paper. The referees are identified by the same numbers as before (Referee #2 has in the meantime let his name known, and is Ken Mylne, from the British Met Office).

All three referees recommend acceptance of the paper, with two suggestions from Referees #1 and #2, which I suggest you follow.

I have in addition as Editor a number of comments and suggestions.

1. My first comment has actually to do, at least in part, with science. You apparently use the words *resolution* and *sharpness* as if they corresponded to different properties of a probabilistic prediction system. They actually correspond to the same property, namely (as you write) the *ability to separate a priori the probability classes*, or to distinguish *a priori* between different outcomes (see Broecker, 2014). Please use only one word or, if you use both, say they refer to the same property.

I add that the ROC curve, shown on your Figure 7, is a diagnostic of that property. It shows the degree to which the system under consideration is able to distinguish *a priori* between 'hits' and 'false alarms', *i.e.* between occurrence or non-occurrence of the considered events. That is exactly sharpness. I suggest you replace the words *'good' prediction* in the caption of the figure with the words *sharp prediction system* (that will also remove the uncertainty implied by the quotation marks in *'good'*).

2. A number of acronyms are not expanded, at least not the first time they are used (*e.g.* PEARP-S2M on l. 31). Please check systematically that all acronyms are expanded on their first occurrence, and give appropriate references whenever necessary.

3. Figure 6. Was the number of intervals used for building the histograms arbitrary, or did it correspond to anything imbedded from the start in the prediction system. If yes, to what does it correspond (that is not clear to me) ?

4. L. 202, … *at the end of* ~~*March*~~ *the period*.

5. Table 1, l. 7. *PR0 Raw probability of HN > 0*

6. L. 216, ~~*The second column*~~ … → *The right panel* …

7. I find the caption of Figure 3 somewhat confusing. I suggest … *lead time (orange full lines)* …, *QRF (purple dashes) and EMOS (green points)*. And next

sentence *For each of the three prediction systems, the lower and upper curves represent the 10th and 90th percentiles respectively*.

8. Figure 2. Say more precisely what the vertical coordinate on the figure is.

9. L. 139, … *equals 0* (or *is equal to 0*)

10. L. 219 (and Table 2). I presume *CI* means centiles?

Please revise your paper along the suggestions of Referees #1 and #2, as well as along my own. Should you disagree with a particular suggestion and decide not to follow it, please state precisely your reasons for that.

I am looking forward to receiving the revised version of your paper.

REFERENCE

Broecker, J. (2014) *Resolution and discrimination - two sides of the same coin*. Quarterly Journal of the Royal Meteorological Society, 141 (689). pp. 1277-1282. ISSN 1477-870X doi: https://doi.org/10.1002/qj.2434