**Response to RC2 comments on "A new skill score for ensemble flood maps: assessing spatial spread-skill with remote sensing observations"**

Many thanks to reviewer 2 for your time reviewing our paper. Our responses to numbered comments in black are detailed below in *blue italics.*

1. Details pertaining to the modelling solution used are insufficient and not shown in the workflow figure for some reason.

   *The pre-computed flood map creation process is not strictly part of the Flood Foresight system. However, we will add additional details to the flow diagram such as the inclusion of the hydraulic model used.*

2. As the AEP based map library is described, I struggled to understand how the authors arrived at the number 36 for the maps, or the interpolation done to derive these maps

   *We will make it clear that each of the 6 main return period flood maps are interpolated linearly to make 5 additional intermediate flood maps between each of the main RP flood maps plus an additional 5 below the lowest return period flood map.*

3. The large tracts of quoted text from Hooker et al., 2022 seem rather unusual if not unacceptable, I would summarize and rephrase with a citation, or at least check that the quoted text stays under 10% of the manuscript and reformat correctly. Please consult the APA/MLA guidelines for verbatim quotes >40 words as they should be formatted as block quotes and page numbers from the original article must be provided which currently is not the case. https://research.wou.edu/apa/apa-block-quote

   *Most of the quoted text is mathematical material such as definitions of terms used in the equations, which we feel are useful to include to aid understanding of the ensemble application. We will rephrase and summarise the text so that this will read differently to the previous paper.*

4. The authors choose splines to aggregate the observation data to the model scale, this is strange in my view. Spline interpolation is typically used when we have only a few measured points based on which the underlying random field must be estimated. What the authors try to do here is simply upscaling, thus, the use of splines is not really justified to me. In fact when upscaling one aggregates not interpolates and thus I would expect to see averaging or majority or max or other aggregation techniques or some justification as to why this was chosen.

*This was coded as 0-order spline interpolation, which is in fact bilinear interpolation (equivalent to averaging for aggregation). We will update the manuscript to make this correction.*

Otherwise great effort from the authors, I am particularly happy to see the annually devastating Assam floods receive more research attention! I look forward to seeing this manuscript published online :)