Responses to Editor for NHESS Article 2021-237: “Effective uncertainty visualization for aftershock forecast maps”.

Dear Drs. Kreibich, Malamud, Tarolli and Ulbrich,

Please find again enclosed our revised manuscript, “Effective uncertainty visualization for aftershock forecast maps” by Max Schneider, Michelle McDowell, Peter Guttorp, E. Ashley Steel, and Nadine Fleischhut, for publication in the journal *Natural Hazards and Earth Systems Sciences*.

The editor suggested to pay attention to two particular referee comments. Both of these were actually addressed in the revised manuscript; however, in our Author Response document, we unfortunately referred to the wrong (old) line numbers when mentioning where we made that revision. We do not believe we can actually expand more on these referee comments than we already have in the revised manuscript.

In the attached, please find our updated responses to the two referee comments that the editor spotlighted, with updated line numbers.

Thank you for considering our revisions.

Sincerely,

Max Schneider, Michelle McDowell, Peter Guttorp, E. Ashley Steel, and Nadine Fleischhut

Lastly, the results about the bounds/interval map are very interesting but there is limited discussion of how these might inform design in practice (e.g., what are the possible implications of using two separate boundary maps labelled optimistic and pessimistic in a high stakes crisis communication environment? Could these be misinterpreted or separated?)

We will add a sentence around line 703-704 discussing how the bounds maps might be useful in crisis management contexts.

What I see as a weakness in the experiment design is that task 3 depends on tasks 1 and 2 (reading accuracy). It should be ensured as a first experiment that all maps can be read with high accuracy. In a second experiment, only the maps fulfilling this criterion should be used for the judgment task. But maybe for this setup this is not a severe flaw. But the authors should add this to the limitations section.

We will include a discussion of this point in the paragraph of the Discussions section that starts around line 674-676. Because we did not separately test the map-reading accuracy of visualizations before testing their effects on judgment, we could not fix potential issues in a visualization's design that may impede map-reading, and we can discuss this concern as a limitation.