Reviewer 2

This study investigates the communication of probabilistic hydrological forecasts with different types of users based on phone survey and qualitative elaboration. They show some interesting findings, for example, users' responses to uncertainty of forecasting results, similarities and differences in visualization preferences of different users, their curiosity in hydrological forecasting methods and so on. This study also shows us a blueprint of forecasting visualization schemes from a holistic view of water depth, inundation area, discharge and the uncertainty according to wide suggestions from the users' end. The paper is generally well-organized and the structure is clear. Such study can improve hydrological early warning systems, thus, benefit flood risk management.

Thank you very much for reviewing our manuscript and for your valuable comments and suggestions. We would however like to emphasise that the interviews were not conducted on the phone. They were conducted via the online platform Zoom, and this is only because of the restrictions due to the pandemic in 2020 and 2021. The initial plan was to conduct all interviews in person. We think it makes a difference to conduct the interviews on an online platform with video rather than on the phone, because the phone would have removed more of the non-verbal language of the respondents. In addition, the vast majority of interviews were group interviews, which would have been very difficult on the phone but possible with an online video platform.

However, I have several major concerns that expect to authors to address:

1. The innovation of this study needs to be further addressed (i.e., things that has not been done by previous study). In the introduction, the authors fully reviewed previous investigates on the communication of flood risks and highlight the importance of survey on probabilistic forecasts. However, the difference from or increment to previous studies is not clearly pointed out. For example, previous studies may only investigate communication of deterministic forecasts or 1-D/2-D hydrological forecasts instead of inundation map, etc. Besides, this study only survey people living in south Québec, where floods are mainly caused by snow melt. However, the situation may be different for other regions and countries. It remains known to what degree the conclusion drawn from this study can be transferred to and referenced by other places of Canada and the world.

Thank you for pointing this out. First, as explained in our response to comment #7 From Reviewer 1, we are convinced that our conclusions are transferable to other places in Canada and in the world, as long as those places are culturally relatively similar. This is because the questions asked during the interview focussed on the visualisation and communication of the information, and not on the flood-generating mechanisms themselves. For instance, the fact that most floods in Quebec are generated by snowmelt is not relevant for the study and does not affect our results, as it was never question of flood-generating mechanisms during the interview. Some participants mentioned their specific concerns about snow, but it was never directly asked to them.

Regarding the innovation, at the beginning of the project (in 2019) and extensive literature review was performed. At that time, we did not find any study similar to ours, and based the design of the four visualisation prototypes on recommendations from other fields of studies (mainly the communication of hurricane forecasts, which have been studied for a long time). In fact, although operational probabilistic streamflow forecasts are now common, it is not yet the case with probabilistic flood maps. We will emphasise this in the

2. Survey should strictly take sample representativeness into account. The education background, gender and age of the participants and their living/working places may affect the results and the representativeness of samples. Thus, it will be essential to include statistics of these kind of information. For instance, a geographic distribution of the participants with flood risk map, proportion of people with/without hydrology or atmospheric education background, etc.

Thank you for pointing this out. We will add as much information as possible about participants background in the manuscript. However, note that participants were not asked to provide detailed information about themselves. For instance, we do not know their specific level of scolarity. However, none of the participants had a hydrology or atmospheric science background. We will also add more information about the geographic distribution of flood risk, possibly by improving Figure 1 and/or adding more details in section 3.1 (about the sampling strategy).

3. I also notice that the authors design different contents of phone survey for farmers and citizens from non-farmers or citizens (i.e., drop "the themes related to the nature of the information" for farmers and citizens) but did not explain the reason for doing this too much. I think the different treatment may cause the readers wondering whether the forecast maps should **originally** be designed differently for these two kinds of users (i.e., farmers and citizens & non-farmer or citizens). Since satisfying all kind of users with a single forecast map seems to be impossible. Therefore, why did not the authors design different kind of forecast maps for them at first and then do the survey?

The content of the Zoom interview was not very different between the different groups. Some questions were simply removed from the list for citizens and farmers because they were not applicable to them. Therefore, it is the same initial content, but slightly reduced for farmers and citizens. We will clarify that in the revised version of the manuscript.

One of the elements we wanted to verify was if a single forecast map would be sufficient to satisfy all kinds of users. Therefore, we originally wanted to present all groups with the same prototypes. We also consider this approach to be more objective, in the sense that we initially provided everybody with the same information instead of taking decisions based on our own a priori for certain groups. The decision to not present Prototype 3 to the farmers and citizens came later, after the interviews with the ministries and municipalities, during which it was strongly recommended.

4. The presentation is overall a bit too qualitative. Some quantitative descriptions and statistic plots are needed. For example, in L341-349, the authors can show the voting proportion of color scheme preferences with real numbers or a table or histogram. Table 7 offers too much unsorted information and words. Table 8-11 is the same without statistics and graph visualization.

Thank you for this comment, which is in agreement with comments # 6 and 8 from Reviewer 1, who also provided suggestions to transform Tables 2 to 4 in figures. We will modify Table 7 to make it less "wordy" and more orderly. In some cases, it will not be possible to provide quantitative information, simply because the design of the interviews followed a qualitative framework in which quantitative information was not demanded. However, whenever possible, we will strive to make it more quantitative. For instance, we

do not have quantitative information regarding Table 8-11. This is due to the initial methodological orientation of the research, and similar to other qualitative studies in the field of operation hydrological forecasting, for instance in Demeritt et al. (2012).

Demeritt D., Nobert S., Cloke H.L. and Pappenberger F. (2012) The European Flood Alert System and the communication, perception, and use of ensemble predictions for operational flood risk management, Hydrological Processes, 27(1), 147-157.

Minor comments:

5. The structure of the abstract need to improve. The background occupies almost half of the abstract, leaving little space for results and main conclusions. The conclusion is the only one sentence with "several" statement (L19-20). And the significance of the study needs to be further stressed.

We will modify the abstract according to your comments and suggestions.

6. Figure 1: The legend of the blue polygons and lines is needed. Also, please add coordinates for the map.

We will improve the map according to your suggestions and those of Reviewer 1.

7. As mentioned in L112, the investigation of color scale is one of the objectives of this study, however, there is no echo in the discussion or conclusion section.

Thank you for pointing this out. The blue colour scale was preferred by a majority of participants. We will add more details in the conclusion.

8. In the abstract and Figure 2, the number of the citizens and farmers are 37 in total, however, in Section 3.1.4, the author said 33 citizens plus 5 farmers. The numbers contradict. Please check. Besides, in L201, the number 11 is confusing.

The numbers will be corrected. As for the number (11) of focus groups for the citizen, we will rephrase this sentence to avoid confusion.

9. In Section 3.2, the authors said "except for citizens and farmers, one-to-one interviews are taken for the participants". However, in Figure 2, the interview number and respondents differ, which is confusing.

This is a mistake, thank you for pointing this out. Although the groups were small for ministries, municipalities and organizations, almost none of the interviews were one-to-one. This will be corrected.

10. The author should double check the upper and lower case of titles in the references. For example, the fifth and last reference in Page 29 use upper-case for the title, while others did not. The same problems can be found in Page 30.

This will be corrected.