Authors' responses to interactive comment RC2 from Anonymous Referee #2

Dear authors,

Thank you for the interesting read. To me this is a valuable insight in the reality of flood forecasting and (early) warning in practice, provided first hand, by the duty officers of the UK Environment Agency. I appreciate the choice of the authors to present many quotes from the interviews performed, which helps in their attempt to make an authentic account of the current practice, as well as the expectations of the upcoming introduction of a probabilistic flood forecasting system.

We thank the reviewer for their positive comments with regards to this paper's added value in published literature and format, as well as their valuable feedback which will help improve the paper for final publication.

I have the following general comments:

1) The experience with the recent (2013/2014) transition from single forecast to 2-scenarios may be quite relevant for the perception of the interviewees and prospects of the upcoming change from two scenarios to probabilistic forecasts. Did you discuss this with the interviewees and could you perhaps elaborate more on this aspect in the paper?

We did not discuss this explicitly during interviews as not all interviewees had experienced this transition. However, a few interviewees compared the single forecast to the two scenarios and overall seemed positive about the added value of the scenarios. This reinforces the need for this paper to be published. Indeed, if similar work/interviews had been done for the previous transition to the two scenarios there could have been written records of the challenges and opportunities it presented to help the current transition to probabilistic forecasts.

2) From reading the paper I get the impression that there is little known yet about what will be the procedures for preparing, communicating internally, and use in warning decision making of the probabilistic forecasts. Could you, for example in the Context section, elaborate on what is known, and what was known to the interviewees at the time of the interviews, about the upcoming transition to probabilistic forecasts? If nothing is known yet, the good thing is that the recommendation of co-design (recommendation 3) can still be taken up, and at the same time it might explain some of the perceived challenges associated with the upcoming transition to using probabilistic forecasts.

This is a very good point, which we will elaborate on in the Introduction-Context Section. There was indeed very little known about the communication and forecast use procedures and no internal procedures in place at the time of the interviews. This motivated this work and interviews. Regarding the interviews, while some interviewees knew about the transition and were involved in the technical design of the new forecasting system (as part of their daily job at the EA, when not on duty), a few interviewees had just learnt about the transition a few days/hours prior to the interviews.

3) The presentation and discussion of the answers of the interviewees to the last question, about the upcoming introduction of probabilistic flood forecasts, seems to me to be somewhat limited. This impression is fed by the sudden change in reporting format from in-line quotes to a wordcloud,

summarising table (Table 1), and a reference to an Appendix, none of which are discussed in the manuscript text (lines 460-464). I may be overlooking something, and if not, you may well have chosen this approach for good reasons. If by design, then I would recommend to explain the reasons in the same section (Section 4.3.2). If possible, however, I would recommend continuing with the reporting format of the previous sections, or at least including a discussion of the wordclouds and Table 1.

We fully agree with the reviewer on this point. We had chosen to summarise this information in a table and wordclouds in order to present results succinctly and to shorten the overall length of the paper. However, in hindsight, we agree that this is one of the main highlights of the paper and needs more elaboration. We will incorporate this section into the same format as the other result sections, with a mix of text and supporting quotes.

4) Also the Discussion and Recommendation section leaves me with a feeling that more reflection on the interview results can be done. It would be, for example, interesting to reflect on whether the interviewees' answers to the first questions are in-line with, help explain, or not, their perceived opportunities and challenges (last question). The 10 recommendations in section 5.2 seem somewhat disconnected from the interview results (only Refs to literature are given). I would recommend to put in, in Section 5, more references to findings reported in earlier sections of the paper.

In order to link more clearly the interview results and recommendations, we will:

- Link Table 1 topics with 1 or 2 recommendations.
- Rewrite Discussion section to combine Sections 5.1 and 5.2, where each paragraph will present: interview finding – literature finding – recommendation.

Specific questions and comments:

5) You focus on the benefit of probabilistic forecasts of increasing lead time (e.g. p1 l14 and l36, p3/4 l124/125, while other benefits include the potential of increasing the probability of detection of floods (reducing missed events), and supporting risk-based decision making. Could you reflect on this in the text? E.g. adding advantages or explaining why you refer mainly to increasing lead time. Interview findings indicate that in the current practice final decision of issuing a warning is often based on nowcasts with lead times of only a couple of hours and/or on observations (e.g. page 7 line 267 and page 11 line 397/398). Do the interviewees and/or you think that the introduction of probabilistic forecasts is going to change this practice? Could you reflect on this in the paper?

This is a very good point which we have overlooked and not explicitly mentioned here. We will add these additional benefits of probabilistic forecasts to the paper.

The authors' hope is indeed that this transition to probabilistic forecasts will be reflected in the EA's decision-making practice (e.g. the lead time at which warnings are issued). This is reflected in Recommendation 7 (Page 7 starting line 590). We will develop this point further in the Discussion section. It was not explicitly mentioned by any of the interviewees.

6) Some of the duty officers (DOs) seem to be concerned about how the probabilistic forecasts will be received by the action response units. Does this mean that in the current practice, forecast hydrographs are send along with the warning to action response units? Please clarify in the text.

To our knowledge the information shared with emergency responders is only textual. Probabilistic information would however also affect verbal communication, which some duty officers indeed expressed worries about. This will be clarified in the text.

7) I may have missed it, but could you include more information (and refs if available) on the probabilistic forecasting system that will be used? Is it based on meteorological ensembles or on another probabilistic forecast method? Will the MFDOs be responsible of running it through the hydrological models (as they seem to be now)? Will hydrological uncertainty also be included in the hydrometeorological ensemble, and how? etc. Please also reflect on whether this information on the features of the new forecasting system was known to the interviewees.

This was not known at the time of the interviews. Together with the EA co-authors of this paper, we will try to provide some more information about the new probabilistic forecasting system, if possible.

8) When reading Table 1, I do not perceive a strong concern about the upcoming introduction of probabilistic forecasts, while when reading the quotes of Annex C I do sense a strong concern among the Duty Officers interviewed. This concern seems mainly to be that probabilistic forecasts will put all the responsibility of taking a decision with themselves (rather than with the forecasters or with the action response units). Could you note and discuss this in Section 4.3.2.? And then elaborate on recommendations on how to prevent/manage that? For example, Recommendation 9), setting guidelines on '..the forecast confidence at which certain decisions and actions should be made..', may also not be the answer, because the DOs indicated in the present-day practice the value of local expert judgement in issuing warnings and seem to appreciate having the freedom of applying such expertise. Prescribing decision making rules, may, therefore, be a step too far in taking away forecast interpretation responsibility from them.

We will reword the content of Table 1 and reformat it as per comment 3) to capture the quotes more adequately.

We understand and agree to an extent with the reviewer's comment about taking away forecast interpretation responsibility from the duty officers. We however believe that this could perhaps be a starting point (i.e. recommended decision), from which duty officers could be allowed to deviate when needed. This will be discussed further with the EA paper co-authors in order to propose an adapted (set of) recommendation(s) to tackle this.

9) Could you reflect on whether the interviewees see the probabilistic forecasts as an additional input to their flood forecast confidence assessment? The DOs are already communicating a confidence level with the warnings they send-on down the line. One quote in Appendix 3 confirms seeing this as an opportunity, but it would be interesting to read from you what is your impression on this for the other interviewees.

A few interviewees mentioned the fact that probabilistic forecasts would reveal uncertainty otherwise hidden with the flood scenarios, as is indeed reflected in Quote O2, and some words of

the left wordcloud (Fig. 5; e.g. “apparent”, “displays”, “reveal”, “hidden”, etc.). This will be expanded on in the text.

10) Page 4 line 130/131 refers to the EA already using probabilistic coastal flood forecasts. Why not learn from the experiences of that earlier transition (perhaps too long ago), or at least from the user experiences. Could you elaborate a bit more, e.g. whether or not you think that would be interesting for other researchers and the EA to pick up.

This was not explored during these interviews as not all interviewees were familiar with coastal flood forecasts (given the non-proximity of some of the EA centres where interviews were carried out to the coast). A few lines will be added to the Discussion section, suggesting this as future research to learn from this past transition to coastal flood probabilistic forecasts for the current transition for fluvial floods.

11) Page 5 lines 182-184: Consider referring back to these research questions in your Conclusions section.

This will be addressed.

12) Page 6 lines 222-227: Choosing the What-if scenario could be perceived as quite a responsibility. A responsibility that might be (partly) taken away with the introduction of probabilistic forecasts. Did you discuss this with the MFDOs and what are their and your thoughts on this? Consider elaborating on this in the paper.

This was indeed mentioned in Table 1, under “The forecasting system”: “Some interviewees mentioned that the two scenarios, and the What If scenarios used to produce them, were sometimes challenging to play with and required a lot of expert judgment, thus making them inconsistent nation-wide.” It will be expanded on when the table is adapted into text (as per comment 3) above).

13) Are all the warnings issued being archived (including alerts, issue time, updates, etc.)? Are actual flood occurrences being documented? And if so, are the archives being compared and analysed? Could you reflect on this, and do you think such analysis could be/has been helpful for identifying challenges in the current forecasting system and warning practices, as well as for analysing in the near future the impact (or lack thereof) when introducing the probabilistic forecasts and identifying persistent and potentially new challenges?

The ‘Flood Intelligence Files’ compile information (e.g. highest events on record, what rainfall led to them, what the catchment state was at the time and any known impacts) for every gauge the EA is providing forecasts for. Further information (e.g. whether the warnings are being logged as well for post-event analysis) will be added to the paper after discussions with the EA paper co-authors. The use of such a system to monitor the transition’s performance in practice will be elaborated on in the Discussion section. If currently non-existent, such a system would be very valuable indeed and will be added to the recommendations.

14) Page 11 line 391/392: Could these differences perhaps also be a consequence of differences in catchment size/rainfall-runoff response time/land use and differences in flood management actions

that follow the warnings and the time these measures take? Consider mentioning/reflecting on this (at this point) in the paper.

This is indeed true and the historical differences are also caused by catchment response differences as per our interview discussions, and possibly by the amount of time it takes to prepare in anticipation for a flood (partly controlled by the catchment size too). This will be mentioned in the text.

15) Page 11 line 400: Are the DOs being scored, and if so, how are they scored, and what are these scores used for? If possible, would be interesting to comment on below this citation, and perhaps consider to reflect on how such scoring may have an encouraging or discouraging impact on the uptake of probabilistic forecasts.

This could be a figure of speech used by the interviewee and will be explored further with the EA paper co-authors. It will be clarified in this section and subsequently discussed in the Discussion section, alongside discussion about FWDOs' worries of the transition to probabilistic forecasts moving "the burden of making a decision further down the tree" (Page 14 line 539).

16) The paper concerns the upcoming transition from a 2-scenario forecast to a probabilistic forecast, but the Supplementary material seems to focus on the recently completed transition from a single forecast to the two scenario's (following 2013/2014). Please clarify, e.g. in the author response, not necessarily in the manuscript.

The Supplementary material indeed displays examples of the current system (two scenario-based), as the future probabilistic system is still in the making. This will be clarified in the paper in the Appendix caption.

Detailed comments and editorials:

We will make the following changes

p1 l12 Consider ..inclusion of uncertainty information in..

p1 l13 Consider ..potential upcoming floods.. instead of 'future' to avoid confusion with climate change. Also consider for other occurrences of 'future'.

p1 l18 Consider ..understand their perception on how this transition..

p1 l24 This sentence is rather broad and in my view not necessary. Consider leaving it out. Instead consider putting some of the key findings and recommendations (similarly to what is written in the Conclusions section)

We will leave this sentence in the abstract as it is important to highlight at this stage of the paper that a glossary of terms is available, as this is a paper for Geoscience Communication and should hence be able to communicate these findings for readers from a range of disciplines.

We will however include a few more specific key findings to the Abstract, and thank the reviewer for this suggestion.

p1 l38/39 Consider ..given the explicit provision of uncertainty information.. Single forecasts are as uncertain as probabilistic forecasts. The uncertainty is just not shown.

p2 l42/43 ..designed to capture scenarios that may not always realise.. That does not sound quite right/the point of probabilistic forecasting. Consider just leaving that whole sentence out, or reformulate to something like "Warning on the basis of low probabilities of flood, for example, will reduce the chance of missing an event, but will also lead to more false alarms."

p2 l44 Consider .. when a pre-defined threshold (e.g. river stage) is reached..

p3 l88 Consider ..whilst local flood authorities..

p3 l113 I would suggest including here an explanation on how the two scenarios are prepared. (I realise this is described later, but it left me curious from this point onward, especially because it matters for the context of the interviews to what extend these two scenarios are a step towards probabilistic forecasting or not).

This will be addressed and clarified at this stage when the Introduction and Context sections are merged.

p3 l114-118 I am not sure if I now understand correctly how the two scenarios should be used. This might be due to me not being a native speaker, but if you could reformulate to further clarify, that would be appreciated.

This is vocabulary used by the EA around these scenarios and cannot be paraphrased. We will however expand on this further (as per the comment above) to clarify this point.

p3 l120 ..potential risks of impacts.. Please reformulate.

This will be rephrased to: “Allowing to quantify the potential impacts of upcoming floods and their associated likelihood”

p4 l129 I do not think you can ’ensure’ the appropriate use. Consider to reformulate, e.g. ’support’.

p5 l171 ..advance..

p5 l186 ..communicated by several interviewees,..

p6 l234 ’waiting for the forecast to be confident’ Please explain how the forecast can become confident in the context referred to here (present practice).

As stated in the glossary of technical terms (Appendix A), “confident” here refers to “A forecaster’s expert judgement of how certain they are that the forecast is right”. The forecast is composed of two scenarios, which might sometimes show very diverging outcomes. This may lead to the duty officer being less “confident” about the signal shown by the forecast and the decision to make. Furthermore, combining different sources of information (highlighted in Section 4.1.1, but also in Section 4.2; e.g. national/county scale forecasts, mode performance information, river level correlations), the FWDO will add some expert judgment to gauge whether they can “trust” what the two forecast scenarios show. This will be clarified both in the text and in the glossary.

p7 l257 please add who has the 'Expert knowledge'.

The FWDOs' expert knowledge and the knowledge contained in the 'Flood Intelligence Files' (see Page 10 lines 373-374). This will be clarified here.

p7 l264 please also describe how/for what/when the 'reasonable worst case' should be used.

It is "used for preparation, information and response to flooding" (Page 3 line 116). As stated on Page 9 lines 343-345: "For certain types of events, such as convective rainfall events*, for which the duty officers know models are still limited, they might decide to issue a warning based on the 'Reasonable Worst Case', although it is "technically against procedure" [MFDO2]." Section 4.2 will be merged with Section 4.1 and this will thus become clearer at this stage (as per responses to RC1).

p9 l345 please consider to add also the procedure for using the reasonable worst case again.

See comment above.

P10 l380 - 383 seem to me a bit too personal. Kindly double-check.

We believe that it is a valid and important perspective to quote. The interviewee will however be contacted to verify that they are fine with this quote being in the paper.

p12 l453 In my view this is an important finding that can also be used in the upcoming transition (and gives reason for a positive prospect). I cannot recall whether you clearly refer back to this finding in the Discussion and Conclusion sections at the end of the paper, but if not, I recommend including it.

We absolutely agree and will mention it again and discuss it in the Discussion section.

p13 l463/464 Unclear, and appears as a stand-alone sentence. I suggest making this part of a discussion to be added, of Fig. 5, Table 1, and Appendix C. (See also my third general comment)

We agree - see our answer to comment 3) above.

p13 l463 consider ..sound extreme.. or alternative formulation.

We will reformulate to: "sound extreme"

p13 l469/470 I do not understand this sentence. If we achieve increased confidence levels before moving up lead time, why is the second part of the sentence, about the chaotic system, posing problems? Consider clarifying or leaving out the sentence.

This will be rephrased to: "However, despite improvements in flood forecasting at increasing lead times, the predictability is still inherently limited by the chaotic nature of the system we are trying to model."

p13 l472/473 ..'uncertain' science.. Not sure I understand. Do you mean new discoveries being at first 'uncertain' (or not trusted), until the experiment has been reproduced with the same results (or tested in pilots, practice, etc.)? Or do you mean, more specifically, the science of quantifying uncertainty associated with predictions? Consider reformulating.

We refer here to the latter, the apparent uncertainty displayed by probabilistic predictions. We will rephrase to: “of new and probabilistic science”.

p13 l490/491 consider ..decision makers operate, and where the forecast..

p13 l497 consider ..uncertainty information.. or ..information on uncertainty..

p14 l512/513 ..crucial to develop a methodology.. Not sure if a single methodology is the ’crucial’ solution of the challenge of using probabilistic forecasts. It may be that case-dependent and user-community dependent ways have to be found, by scientists and users together, on how to effectively use confidence (uncertainty) information.

We agree and will rephrase this.

p14 l518 I do not think this reflects the main idea of probabilistic forecasts, nor that the provision of scenarios causes more false alarms. I would argue the other way around, that having the scenarios, the probabilistic forecasts, gives the opportunity to the decision maker (and its beneficiaries) to balance the number of missed events and false alarms to their needs. Please reformulate.

We will reformulate to reflect the main point of this paragraph: “A transition to probabilistic flood forecasts should be reflected in an institution’s wider flood management priorities.”

p14 l525 This seems a bit out of the blue and not clear to what extend you think this relates specifically to EA communication pathways and warning procedures.

This entire paragraph will be reformulated to capture the main point better (as per the comment above).

p14 l531 Consider adding a brief explanation on what is a ’post-factual society’.

p15 l547 consider ..we made a list..

We used the present tense here as these recommendations are relevant and actionable now.

p15 l548 consider something like ..The recommendations concern actions we think the EA should take with high priority..

p15 recommendation 1: Before this campaign, should there first be known a bit more on how the change will be done, or not?

We agree to an extent, but think that the EA shouldn’t wait for the entire system to be set up before starting to tell people about the changes to come. This will give key actors a chance to be involved for a more successful transition. This will be reflected in the text.

p15 recommendation 2: consider ..to all players..

p15 l577 is this recommendation specifically for EA, or for the flood forecasting and early warning research community. If the latter, consider moving this elsewhere.

This recommendation is both general and specific to the EA. This will be clarified and make more sense when the discussion is merged (see response to comment 4) above).

p16 l582/583 given the reported differences in catchments, forecast performance, impacts, and warning response actions, etc., double-check this recommendation. Consider 'customised' rather than 'homogeneous'.

We like the word "customised" as it reflects inherent differences amongst centres and areas (as stated in the paper).

p16 l601 consider ..should be collected and used to update design and procedures..

p16 l602 consider ..To handle situations..

p16 l613 stand-alone sentence, consider moving somewhere else or elaborating.

p17 l623 because of the mentioned concern of the responsibility of decision making being pushed down to the DOs, I would not write 'lie mostly outside their role', rather just something like 'the main perceived challenges concern..'. In the sentence before, perhaps it would be more clear to write something like '..concerns about impacts of this transition on the communication and interaction between them.'

We agree with this point and will rephrase this sentence.

p17 l624 consider replacing 'translating uncertain information to a binary decision' (because it is a challenge that they already perceive in the present situation) for the worry of the responsibility of decision making being pushed further to them.

p17 l625 consider reformulating to something like ..High priority actions were recommended to the EA.. to support a successful..

p23 Caption figure 4: consider ..Complex flood forecast interpretation landscape.. because the decision making landscape includes more elements, such as external pressure.

p23 Caption figure 5: please add from what the opportunities and challenges arise (e.g. ..from the introduction of probabilistic flood forecasts)

p24 Title table 1: consider ..A sample of supporting quotes..

This will be adapted into text and the comments below will be incorporated.

p24 Table 1 line 2: 'improve long-term communication' please clarify.

p24 Table 1 line 2: add a full-stop at the end of the sentence.

p24 Table 1 line 5: consider ..contain information on forecast uncertainty..

p24 Table 1 line 22: ..it is worth noting that..