

## ***Interactive comment on “Do probabilistic forecasts lead to better decisions?” by M. H. Ramos et al.***

**M. H. Ramos et al.**

maria-helena.ramos@irstea.fr

Received and published: 14 March 2013

We thank Referee #2 for his/her positive statements and evaluation of our paper. The thoughtful and encouraging review undoubtedly contributes to improve it. The points raised were addressed as explained below.

Referee general comment #1: "(...) the limited number of cases in each round makes the overall analysis not robust. The authors rightfully caution about the conclusion by stating clearly the weak points of the overall experiment: untrained population sample, lack of training on this particular system, restricted number of cases, etc., and finally conclude on the benefit on this experiment to start developing training on probabilistic forecasts."

C7088

Answer: We agree that 6 rounds for each game is a limited number of cases. Unfortunately we could not have more than the 12 rounds played due to time constraints. This limited number was partially compensated by the fact that we were able to reach a large group (98 participants) by playing the game during the EGU assembly. The analyses carried out from the results pointed to general conclusions, which were similar to the ones drawn from other applications the authors did of the same game (students and other conferences). We think therefore that the exercise can be seen as robust enough for an awareness-raising exercise, as it was capable of illustrating some of the main aspects involved in decision making under uncertain forecasts. Limitations exist indeed and are carefully discussed to prompt new game designs and initiatives to overcome them.

Referee general comment #2: "But the paper could benefit from assessing what drives the most the value of the probabilistic forecasts, using gains and loss for example. What is driving the value of the probability forecast in this experiment:

Referee point i) learning from previous decisions: authors analyze transitions probabilities, but adding transition gains in the analysis might help"

Answer: With the limited number of rounds and a study that was not focused on the relationship between cost/loss ratio and decisions (the focus of the paper was on discussing the added value – not necessarily economic – of probabilistic forecasts in making decisions), the analysis of gains regarding transitions from one case to another was limited to a simplified one, presented in Fig. 8. In this figure, money left in purse was separately analyzed at the transitions to a “yes” or a “no” answer. The aim was to investigate if money left might influence the decisions. The conditional distributions showed up to be different, and the analysis gave indications that higher (lesser) money in purse prompted to more “no” (“yes”) risk-seeking (risk-averse) attitudes (see section 3.5). However the point raised by the referee about the analysis of transition gains is a very interesting one and future studies could benefit from planning such analysis from the beginning of the set up of a laboratory-based study. We will add a comment on this

C7089

in the revised version of the paper, in Sect.4, Discussion.

Referee point ii) "how much money left in the purse: this seems like an over-analysis based on the sample population. A few would think of it in the audience while many others (students for example) might just see at the end in how much trouble they would have been. Starting with 30,000 tokens seemed like we could afford a couple of mistakes; We were allowed to flood the town 5 times out of 6. We could just play and see what we ended up with. If we could afford only 2 mistakes for example, how different would it be? This measure of money left in the purse seems to be a proxy for a more general measure "how much variation in the cost/loss ratio would affect the decisions" which could not be answered by the experiment. This might be worth discussing though. In particular, as mentioned, having only one dimension in this experiment is simplifying the real system tremendously. It is difficult to bring the current experimental conclusion to application or generalization"

Answer: as mentioned, our intention was not to make participants focus on monetary issues (in the case of the flood protection of a city, by the way, this can be a sensitive issue), but instead on the use of the information available to them. Despite this, the analyses of money left in purse (Fig. 6 and Fig. 7) were very informative. On the basis of subsequent experiences (not reported in this paper), they seem to be robust: from the results of the application of the game to a group of 18 participants from the conference FloodRisk held in Rotterdam in 2012, only 3 participants ended with more money when decisions were made without uncertainty. This number was 17 from the application of the game to a group of 65 undergraduate students in China in 2012, and 14 after the application of the game to a group of 42 undergraduate students in France on February 2013. The main results obtained from these applications, even if slight differences appear among the groups, show a general tendency similar to the results presented in our paper concerning the application of the game during EGU 2012 Assembly.

We agree that it is worth discussing the influence of how many mistakes one can afford

C7090

and how that affects decision making in order to raise awareness of the importance of the issue pointed out by the referee: "how much variation in the cost/loss ratio would affect the decisions". It is impossible for us to fully evaluate the attitude of each participant within our game design, and this is probably irrespective of the sample size. We thank again the referee for raising an interesting point for discussion and we will add a short comment on this in the revised version of the paper.

Referee point iii) "patterns of the uncertainty and behaviors: how close to the threshold with narrow uncertainty range or further from the threshold with large uncertainty range are associated with corresponding risk-proned or averse behaviors. That sounds very common-sense although this is based on a very limited number of cases"

Answer: we fully agree that this is maybe common sense, but, even for such statements, an evidence-based discussion usually needs to be built. Given our sample size and the clearly articulated limitations of the study, our paper provides a puzzle piece for this issue (although it should never be used in isolation, we fully agree).

Referee general comment #3: The paper could benefit from further discussing the question of "How much can learning and training drive the value of the probabilistic forecast?" (...)As the authors conclude, only training on probabilistic forecast can bring operational staff to an equivalent level. But what are we then evaluating; quality of the training or the added value of assumed sharp and reliable- information on uncertainty?"

Answer: These are clearly interesting points that deserve discussion. The question of whether one evaluates the value of the training or the added value of the forecast is an interesting one, but surely not restricted to the setting discussed in our paper (it could be applied to any training course with similar aims). From our paper, we are surely not proposing a training or education plan for probabilistic forecasting, as this will highly depend on many other aspects as the forecasting system used and its purposes. However, we strongly believe that good training will increase the appreciation for the value of the forecasts. Applying probabilistic forecasting to an operational chain cannot be

C7091

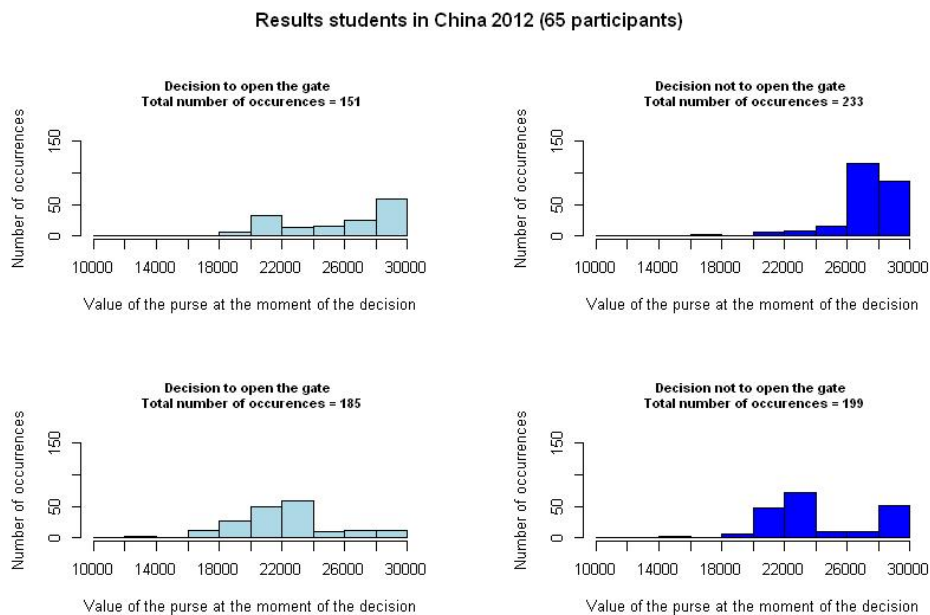
done “suddenly”, as the referee recognizes. However, is the implicit understanding of forecasters regarding the uncertainties of their deterministic forecast systems really equivalent to the full representation of uncertainties through a probabilistic forecasting system? One could argue on the same basis against any forecasting system and simply say that forecasters’ implicit understanding of river behaviour is enough to predict the next flood. We promote and believe in an explicit and transparent representation of uncertainties, which, together with forecasters’ knowledge, will lead in the end to better decisions (see also Faulkner et al., 2007). An in-depth discussion of this issue goes considerably beyond the scope of our paper, but, given the importance of the point raised by the referee, we will add a sentence on it in the discussion section of the revised paper.

Referee minor comments:

- Fig. 3: it will be changed following the suggestions of all the three referees
- Fig. 6: legend will be rephrased
- Fig. 8: each histogram is based on a large number of occurrences (all participants and rounds are considered), as indicated in the figure. Besides, as mentioned above, the results from other applications of the game indicate the same tendencies (see figure below for an example). We think that this analysis, together with the other analyses shown in the paper, provides additional evidence to the general conclusions pointed out by the study. Results from a game with more rounds and applied to more participants may surely contribute to the discussion and such studies are highly encouraged by the authors.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 13569, 2012.

C7092



**Fig. 1.** Identical to Fig. 8 but for results from 65 undergraduate students in China in 2012

C7093