

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

K. Mylne (Referee)

ken.mylne@metoffice.gov.uk

Received and published: 5 February 2013

General Comments This discussion paper addresses a very important area in the need for understanding of decision processes with forecasts including uncertainty. Forecasting systems, usually based around ensemble prediction, increasingly provide estimates of uncertainty alongside best-estimate or expected-value forecasts. Scientifically these provide a much better representation of forecast capability, and potentially allow better decision-making, but there is a strong perception among users and service managers that the information is too complex for most decision-makers to make use of. Experiments such as this, which aim to understand better whether people can indeed make better decisions when presented with uncertainty information, are therefore extremely

C6605

valuable. This paper provides a useful addition to the literature on this topic, and it also references many of the recent contributions in this field so fits well within the literature.

Having said that, and having had some personal experience of presenting results of similar work (some of which is referenced) I think there are some significant drawbacks to the simple experiment conducted here which will limit its influence. The authors fully acknowledge the limitations of their small sample size and the simplicity of the experiment conducted within a 15 minute time slot in a conference. The approach of doing this within a conference has the great benefit of raising awareness in the audience of the need for understanding decision-making in this way, but it also means that the participants in the experiment are very far from typical of either the public or even the business decision-maker. Made up of scientists, many of whom are probably working in related fields, a large proportion of participants will have previously thought about a how to make rational decisions with uncertainty so are likely to perform better in Game 1 than a sample of the general population. I would therefore not expect the results of the experiment to be accepted as useful evidence by those considering whether to include uncertainty information in services to their end-users. I know that the value of the Roulston and Kaplan reference have been challenged as not being “representative” because they used undergraduate students, and the sampling in this paper is unfortunately even more selective. The authors should at the least acknowledge the fact that they are sampling from a very expert set of decision-makers.

Another general problem with the paper is that, despite trying to design an experiment which could be run in 15 minutes with a small sample, the authors nevertheless attempted to put in several different tests and I feel that the end result is that none of the experiments were properly “clean”. There is undoubtedly some useful information in the results, and some broad conclusions can be justified, for example that people did make better decisions and saved more money using the forecasts with uncertainty. However I feel that many of the more detailed analyses of results in section 3.4 are over-analysed and there are too many competing influences to justify some of the con-

C6606

clusions drawn. For example, I don't believe that the influence of earlier action-event pairs can be adequately inferred from only two orders of questions when this is also mixed up with questions with and without uncertainty information. Similarly the effect of how much money they had left would require a much larger sample with different orders of questions. I also think they over-complicated the analysis by providing uncertainty in two of the cases in Game 2.

Thus overall I think this was a useful awareness-raising exercise, both in the conference and as a published paper, but I would suggest that the authors shorten the analysis and concentrate on the key overall results rather than trying to squeeze out too many conclusions from a very limited dataset.

Finally, I applaud the authors for making the game materials available for use in training activities. It would be worth considering whether they also might invite trainers to return the results to a central database, so that they could build up a much larger dataset of results from which more statistically significant results might be drawn.

Specific Comments 1. I found reading the paper I became confused about the differences between "cases" and "rounds", especially around Fig 3. In the first sentence of Section 3.2 it says "...for each of the six rounds played in both games" – I think this should be "...six cases...". I found myself comparing results from Game 1 with the wrong ones in Game 2, so it should be emphasised that the same Case is shown across each row in the figure. It would also be helpful to repeat the forecast data for each case (eg Case 1: 3.81+/- 0.07; 9.74%) on each row of Fig 3, just to emphasise that this is consistent across the figure. It would also be helpful to highlight somehow on the figure in Game 2 which were the two cases where uncertainty info was provided, and also what the optimal decision was.

2. Closely related, in first para of p13580, 5th line and again 9th line: Rounds 4 and 6 should be Cases 4 and 6. Because these were wrong I became confused and misinterpreted Fig 3.

C6607

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

C6608