Expected Utility, a benefit for the forecaster

bruno.gillet-chaulet@meteo.fr

Introduction

In October 2018, the south of France was hit by a dramatic Heavy Precipitation Event that caused several fatalities. A red warning (the maximum level in the French weather warning system named “Vigilance”) was issued. Nevertheless, the timing of release was judged to have come too late to be useful and a sharp controversy broke out in the press. Events such as this have generated negative feedback therefore Public Authorities assigned non-meteorologist inspectors to investigate the forecast process. A report was published (**1**). One of the conclusions was:

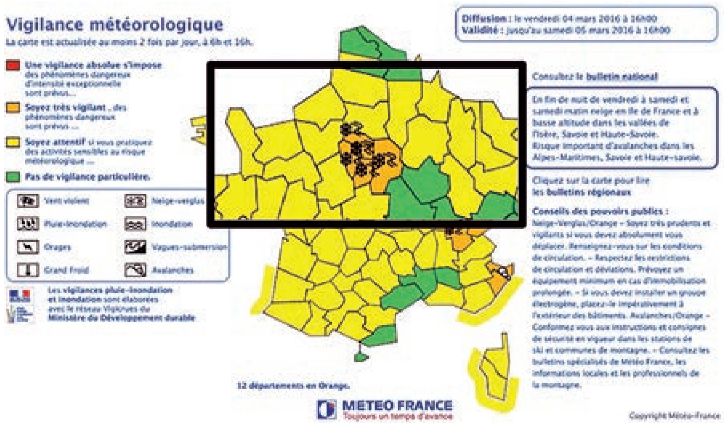
The event was present in the French fine mesh Ensemble Prediction System (EPS PEARO) available at that time, (as an extreme one in the sense of EFI Shift Of Tails). However forecasters were unable to properly use the information, because ***no doctrine of use had been established****.*

Such an assertion calls into question the role of the forecasters. As “experts” of weather warnings, how do they make decisions? Is it possible to follow **a doctrine** in this domain? These points are discussed in this short article.

To Warn or Not to Warn

This paper focuses on decision-making with high stakes outcomes but no time pressure, which means that forecasters have time to decide knowing the deadline. The question is straightforward: issuing a warning or not.

Let us look at an example less catastrophic than the one mentioned in the introduction. It is inspired by the weather conditions on 4th and 5th March 2016. A snowy episode is expected in the north of France. The 3 thumbnails of Figure 1 must be seen as the

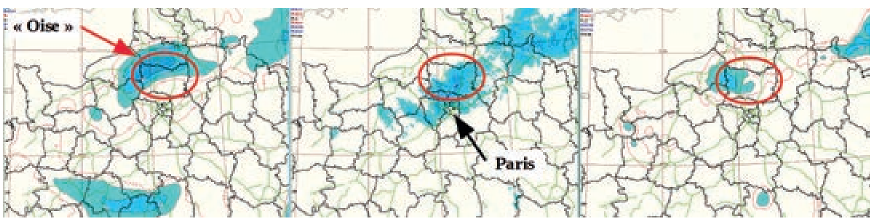


▲ *Figure 2: Vigilance watch map issued by Météo France on Friday 4th March 2016 at 16h00 local time valid for the next following 24 hours. Zoom in on Paris/Île de France.*

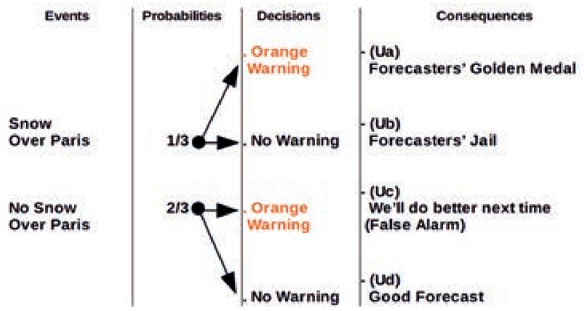
distribution of the possible outcomes in this situation: at the time of the forecast (the day before the event) the expertise does not allow one to determine the solution or rule out a scenario. These 3 possibilities are assumed equally probable (the ultimate goal of ensemble forecasting!). This is a short-sighted view but not detrimental to the reasoning. The displayed parameter is total snowfall over 24 H. The territorial zoning corresponds to the French administrative units thus the warnings apply to this geographical scale.

“Oise” (marked with a red ellipse) is affected in all cases and this is the most likely scenario. The middle thumbnail indicates that Paris might be impacted but this forecast is less likely, Paris being spared from troubles in the other options. Nevertheless, the Paris option was selected by the forecaster on duty at that time, as can be seen on the published Vigilance watch map (Figure 2).

Everything happens as if events in “Oise” are of no importance, whereas attention is only focused on Paris/Île de France. How do we explain this decision-making? The “Expected Utility” model (EU) seems truly relevant in this case (**von Neumann and Morgenstern, 1944**). It is illustrated in Figure 3.



◀ *Figure 1: 24 H accumulated precipitation (total snowfall). The 3 “equiprobable” scenarios for the situation on Saturday 5th March 2016 (Valid Time).*



▲ *Figure 3: Decision tree with probabilities and (imaginary!) utility values for the situation on Saturday 5th March 2016.*

The decision tree describes the possible events (snow or no snow over Paris) with associated probabilities of occurrence (respectively 1/3, 2/3). In each case, decisions are to issue an orange warning or not. Consequences (called utility values Ux) are assessed. For instance, had an event occurred whereas a warning wouldn't have been issued (Ub), the forecaster imagines the detrimental consequences of the decision: Capital City paralysed by huge gridlocks, dwellers blocked by snow, criticism from the Media… We could say, in a humorous way, that the forecaster could get thrown in jail!

Expected Utility is thus defined as the product of the probability of the weather event (Px) and the utility (Ux) previously described. The choice of the warning level derives from the comparison between expected utility values in the case of an orange warning and no warning. The option with the largest EU value is selected (here 1/3*Ua+2/3*Uc > 1/3*Ub+2/3*Ud).

We could then add that the expected utility values concerning "Oise" count as a negligible quantity given its close proximity to Paris. This "district" is then "forgotten" in the Vigilance watch map.

Make a decision!

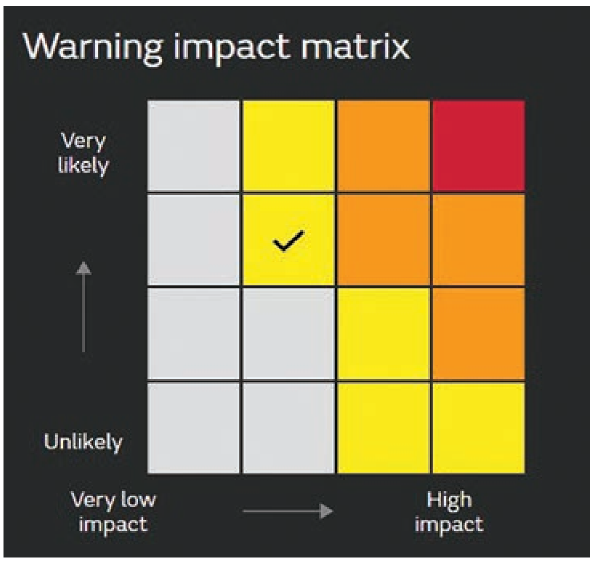
Generally speaking, it is recognised that decision making (to warn or not to warn) doesn't solely depend of the probability of occurrence of the meteorological event (otherwise the most likely scenario would systematically be forecast). The consequences of the decisions are often taken into account. Have you ever noticed how much easier the forecast is when you are not on duty!?

Expected Utility models give opportunity to explore numerous issues.

It is not always possible to quantify utility values especially in the domain of safety of life. Outcomes aren't straightforwardly measurable (fatalities, social and media impact etc) and there is thus a subjective aspect. Psychologists (**Cadet and Chasseigne, 2009**) speak of Subjective Expected Utility (SEU). Mental simulations (**Klein, 1999**) play an important role in this case and depend on the experience of forecasters. These simulations become more accurate with years of practice, training programs and sharing of feedback.

In the case of quantifiable utility values cost-loss models have been in effect for a long time (**Richardson, 2020**). They give the probability thresholds to use in order to minimize losses due to weather events and the costs of precautionary measures to prevent these later. Nevertheless they have limitations: They are only valid for users with a large number of similar repetitive cases e.g. they work very well for economic activities sensitive to weather conditions on a daily basis, where they will provide certain profit during a long period of application. For a specific need however, they are more difficult to use, EU or SEU can better applied to unique cases.

I remember, when I was younger, I was taught; that the aim of a forecaster was to provide users with weather conditions, what the user does with the information is not the forecasters business. This example shows that it is poor advice today. We could add that a "modern" forecaster (especially in the domain of protection of life and property) doesn't forecast the weather any-more but forecasts EU or a "risk". Indeed according to the International Or-



▲ *Figure 4: Example of Warning impact matrix, from the UK Met Office available at https://www.metoffice.gov.uk/weather/guides/warnings* ganisation for Standardization (ISO), risk would be defined as a "combination of the probability of an event and its consequences". The concept is thus very close to EU. The UK Met Office has made it explicit by displaying the famous “Warning impact matrix” (Figure 4). The forecast consists of assessing the probability of realization of a weather event (Hazard) AND its potential impacts (Stakes). A good forecaster must conduct a correct assessment in both aspects. Therefore it can be useful to be aware of cognitive biases involved within human thinking (already mentioned in European Forecaster newsletters, **Persson, 2014, Young, 2017**). It is unfortunate that these subjects are seldom studied at meteorological schools. Finally note that EU is a multiplication model and there is always a random factor in each part of the product!

Back to a doctrine

Let us look again at the example previously showed. Let us assume that the snow forecast is reliable (in the sense of EPS). A doctrine could be: Avoid Non-Detection. Taking 3 similar cases (*forecasting snow* when probability of snow is 1/3), we would obtain 1 correct forecast (detection) and 2 false alarms. Would the City accept these? Another doctrine could be: Avoid False Alarm. In the same way (*forecasting no snow* when probability of snow is 1/3), we would obtain 2 correct forecasts and 1 non-detection. Would the City tolerate it? Actually, the forecaster, as an expert of weather warnings, already follows a doctrine: Minimise non-detections AND false alarms. At Météo France, the Vigilance warning procedure sets figured target objectives, respectively 2% and 16%. Moreover, the forecasters share a strong consensus; in case of red warning, no false alarm would be accepted!

Finally, we would like to highlight an important aspect of the EU model. Looking at the decision tree, note how asymmetric the possibilities are: consider the branches Ua, Ub, Uc. The consequences of the decisions are immediately measurable. Examine the branch Ud. This is a good forecast but nothing happened and the forecaster predicted nothing. Who knows? Who cares? However, sometimes, from the forecaster’s perspective, isn’t it as difficult to choose path Ud as choosing path Ua? One decision is as tough as the other one. During a year of warnings, is it possible to count the number of these (Ud) good decisions which mainly go unnoticed?

Conclusion

EU (or SEU) models seem to be relevant to explain decision-making in high stakes cases when there is no time pressure (deadlines to issue a warning not having been set). They explain why the most likely weather scenario is NOT ALWAYS predicted. They can be applied to unique cases so from this point of view they are difficult to automate (contrary to repetitive cost-loss models). Even so, would we be ready to accept bad decisions of a machine? They thus represent a chance for the forecaster to develop their expertise. Actually they are very close to cost/benefit or cost/lost analysis. It is not clear whether EU (or SEU) models are very well known amongst the forecasters’ community but the paradox is that if you type “Expected Utility” in a search engine on the internet, you will probably find an example containing meteorology! We think they are worth discovering. We believe that detailing weather forecasts AND their possible consequences leads to more objective decision making. They thus become easier to discuss and share.

However, EU (or SEU) models are not applicable in case of sudden unexpected events when a quick reaction is required, when the forecaster is pressed for time. This subject remains to be better understood.

References

(1) Retour d'expérience des inondations du 14 au 17 octobre 2018 dans l'Aude, Isabelle GUION de MERITENS (IGA) - Dominique LACROIX (IGA) - Hugues AYPHASSORHO (CGEDD) - Gilles PIPIEN (CGEDD), available at https://www.interieur.gouv.fr/

Cadet B., Chasseigne G., 2009. Psychologie du jugement et de la décision, des modèles aux applications. De Boeck, Bruxelles, 480 p.

Klein G., 1999, Sources of Power, How People Make Decisions, The MIT Press, Cambridge, Massachusetts, 332 p.

Persson A., 2014. The weather forecaster as an « intuitive statistician ». The European Forecaster, Newsletter of the WGCEF, 19.

Richardson D., 2020, Applications of cost-loss models, ECMWF, available at https://www.ecmwf.int/

von Neumann J., Morgenstern O., 1944, Theory of games and economic behavior, Princeton University Press, Princeton, 774 p.

Young MV., 2017, The Human Element in Forecasting, a Personal Viewpoint, The European Forecaster, Newsletter of the WGCEF, 22.