Anton Beljaars elected as an AMS Fellow

XUBIN ZENG
(University of Arizona),
PETER BAUER (ECMWF)

In October 2014 the American Meteorological Society (AMS) elected 28 of its members to the prestigious rank of AMS Fellow. Anton Beljaars was included in recognition of his fundamental contributions to the observation, understanding and model parametrizations of atmospheric turbulence, and land and ocean surface processes.

As Principal Scientist and Head of the Physical Aspects Section, Anton has played a major role in developing parametrization of atmospheric and surface processes used in ECMWF's Integrated Forecasting System (IFS). In particular, he has developed new ideas about turbulence under very stable and unstable conditions and the interaction between turbulence and land surfaces. These insights enabled him to develop improved

turbulence parametrizations that are widely used.

Anton has played a major role in international research programmes. His studies of land–surface interactions provided one of the scientific underpinnings for the success of GEWEX (Global Energy and Water Exchange Experiment). Also, Anton's work on ocean–atmosphere interactions inspired some aspects of TOGA-COARE (Tropical Ocean Global Atmosphere-Coupled Ocean Atmosphere Response Experiment) that brought together meteorologists and oceanographers

Due to his international reputation, Anton has served on a variety of key international committees (e.g. GEWEX Scientific Steering Group and Mission Advisory Group of the ESA Earth-CARE satellite).

The AMS has about 14,000 members from academia, government and the private sector in various countries. Each year, the Council of the AMS elects as Fellows no more than 1 of every 500 of the Society's members. Fellows are nominated by their peers for outstanding contributions over many years to the atmospheric and related oceanic and hydrologic sciences and applications.

The new Fellows will be honoured formally in January 2015 at the 95th AMS Annual Meeting in Phoenix, Arizona.

Decisions, decisions…!

TIM PALMER (University of Oxford),
DAVID RICHARDSON (ECMWF)

"Forecasts possess no intrinsic value. They acquire value through their ability to influence the decisions made by users of the forecasts" (Allan Murphy)

As indicated by the quote above, the sole purpose of making weather forecasts is to aid decision-making. As a daily commuter, should I take my umbrella to work? As a regional governor, should I order the evacuation of a coastal city ahead of some possible hurricane? As an aid worker, should I prepare for relief measures ahead of an ongoing drought? But are forecasts any good for aiding these types of decisions? If we think that they are, how would we actually go about measuring this quantitatively?

In this note, we outline the reasons why one of ECMWF's principal headline scores – the continuous ranked probability skill score (CRPSS) – is just such a measure. For many readers this might come as a surprise; when defined explicitly, the CRPSS looks like a rather arcane probabilistic skill score which only ensemble-forecast experts are able to understand well.

The continuous ranked probability score (CRPS) compares the forecast probability distribution of a quantity to its analysed value. Both forecast and analysis are expressed as cumulative distribution functions. The CRPS is the squared difference between these distributions, integrated over the range of the quantity being assessed. The CRPSS then compares CRPS of the verified forecast to that of a reference unskilled forecast.

However, it turns out that the CRPSS has a direct and very practical interpretation in terms of decision-making. This is due to the work of Allan Murphy, professor at Oregon State University, who was a pioneer in the field of probabilistic weather forecast verification and devised methods for assessing the value of probabilistic weather forecasts (Allan died in 1997).

To see how to interpret CRPSS we need to discuss how we might go about measuring the value of weather forecasts. To do this, it makes sense to try to generalise and idealise the examples given above, so that the notion of value can be discussed independent of the minutiae of the practical details which are important in individual real-world situations.

Hence, imagine a hypothetical user of weather forecasts, who stands to make a loss L if some adverse weather event occurs (e.g. freezing temperatures, winds exceeding a given speed, rain exceeding some chosen threshold). The loss need not be purely monetary – in the examples related to disaster mitigation and relief, the loss includes human suffering. However, in order to define a quantitative measure of value we have to assume that L can, in principle, be given a numerical value. We will also assume that the user can take precautionary action at cost C to avoid these losses. If the user is the regional governor in the example above, then C denotes the cost of evacuation.

When should such precautionary action be taken?

If L is sensitive to weather, then a good weather forecast is clearly of value in deciding whether or not to take action. However, we need a strategy on how to use the weather forecast for such decisions. A particularly simple strategy might be this: take precautionary action if the forecast predicts the event will occur. This strategy makes sense if the user only has available a single deterministic forecast. However, if the user is an ECMWF customer, then another (generally superior) decision strategy is available.

In economic language, the risk associated with some event is equal to the probability p of that event occurring multiplied by the damage L associated with that event. The ECMWF ensemble forecast (ENS) allows users to directly estimate this all-important probability, without which a proper assessment of risk is impossible. If ENS estimates that the event will occur with probability p (which means that the frequency of occurrence of the event in ENS, at the relevant lead time, equals p), then the risk of the event is equal to pL. The superior strategy referred to above is to take precautionary action when the risk pL exceeds C. Put another way, the strategy is to take precautionary action if the forecast probability of the event exceeds C/L. To make this more concrete, suppose the cost C of precautionary action is one tenth of the unmitigated loss L if the weather event occurs (so that $C/L = 0.1$), then the user should take precautionary action when the forecast probability of the event, according to ENS, exceeds 0.1.

Now if users were to pursue either of these decision strategies, they could assess, let us say after a season of forecasts, whether the forecasts were valuable in making decisions. How are we to measure value? Rather than present value in euros, dollars or pounds, we can measure value by comparison with two standard benchmarks.

- An upper bound on value is associated with that of a hypothetical oracle. This hypothetical oracle can, by definition, forecast the weather perfectly. If one had access to the oracle, one would take precautionary action only when the oracle said the weather event will occur. We will assign a normalised value of 1 to this perfect hypothetical oracle.
- A lower bound on value is associated with knowledge of the climatological frequency, p_c of the weather event of interest. A decision strategy based on a knowledge of p_c alone, is to take precautionary action when $C/L<p_c$. In this (low-value) strategy, one should always take precautionary action if the cost is sufficiently cheap compared to the potential loss, and never take preventative action if the cost is sufficiently high. We will assign a normalised value of 0 to this 'climatological' decision strategy.

We can calculate a corresponding normalised value for the ECMWF forecasts – we call this the 'potential economic value'. The potential economic value of ECMWF forecasts can never be greater than 1; hopefully it will be greater than 0. If it is less than 0, a user should instead base decisions on the climatological strategy mentioned above. The user can readily convert from this normalised measure to financial value: if they know the financial benefit that a perfect forecasting system would bring, the potential economic value shows what fraction of this would be realised by the available forecasting system.

Figure 1 compares the potential economic value of decision strategies based on the ENS and ECMWF's high-resolution forecast (HRES). The x-axis in Figure 1 describes the user cost-loss ratio (C/L), and the y-axis describes the normalised potential economic value. It can be seen that ENS has greater value than HRES for all cost-loss ratios. What may be

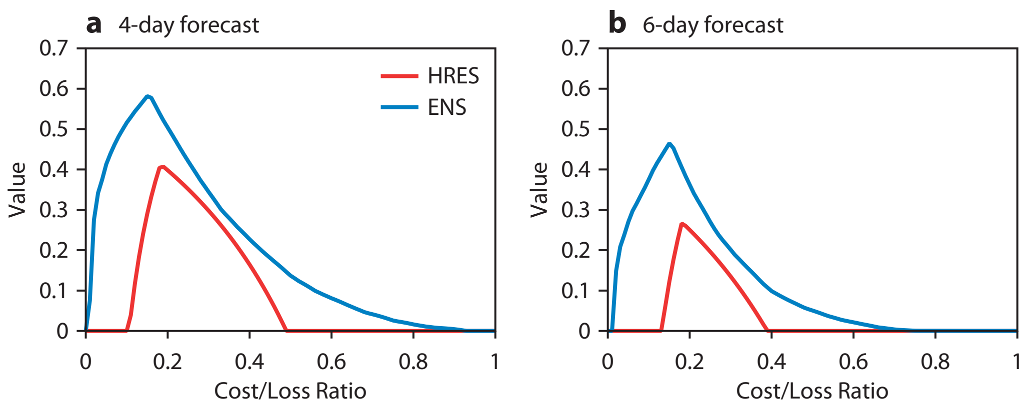

Figure 1 Potential economic value of the high-resolution forecast (HRES) and the ensemble forecast (ENS) in predicting 24-hour precipitation amounts greater than 5 mm over Europe for the summer (June to August) 2014: (a) four-day forecast and (b) six-day forecast.

surprising is that HRES has no value (over decisions made with climatological information only) for users with either high or low cost-loss ratios. This means that decisions using the climatological strategy are superior to those using HRES, for low and high cost-loss ratios. The ENS is valuable for a much larger range of users than the HRES.

Note that the value of HRES can be improved by 'dressing' it using a probability distribution function based on climatological errors. However, ENS is still superior to this dressed HRES in the medium range, and ENS can be further improved by statistical calibration of the ensemble itself, *Gneiting* (2014).

Let us denote by $\overline{V}$ the average potential economic value of ENS overall cost-loss ratios between 0 and 1. As shown by *Murphy* (1966), $\overline{V}$ is equal to a standard skill score for evaluating probability forecasts called the Brier Skill Score.

But we can go further. So far we have imagined decisions based on the occurrence of a particular weather event. For the sake of argument, let us suppose that this event is associated with the daily rainfall total exceeding some threshold T_0 (Figure 1 shows the results for $T_0 = 5$ mm). The Brier Skill Score depends on the choice of threshold T_0. However, we can ask what is the average Brier Skill Score as T_0 varies across all values of threshold, where the average is weighted with the climatological probability of occurrence of T_0. Such an averaged Brier Skill Score is none other than the CPRSS.

So, putting all this together, the CRPSS is simply a normalised measure of the potential economic value of a forecast system (typically an ensemble forecasting system) for a family of users which span the possible range of cost-loss ratios and for weather events which span the range of possible rainfall thresholds. That is to say, CRPSS is perhaps the simplest single measure of the overall value of a forecasting system for decision-making!

By contrast, the traditional Anomaly Correlation Coefficient (ACC) of say the 500 hPa height cannot be directly related to decision strategies in this way. ECMWF's set of headline scores and the wide range of additional evaluation measures provide essential feedback to both users and model developers on the quality of the forecasting system. Together, they assess the underlying quality of the forecasting system as well as the potential benefits to users.

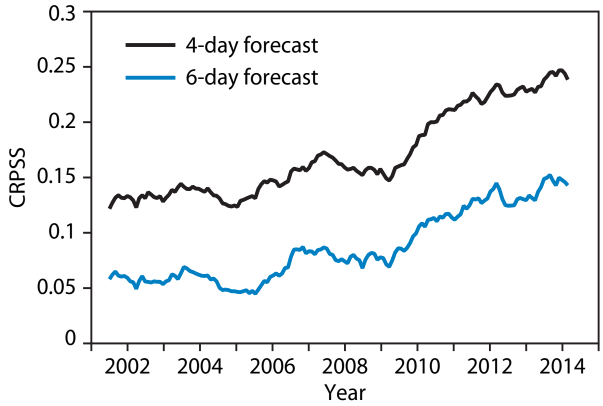

Figure 2 CRPSS for ENS probabilistic precipitation forecasts at day 4 and day 6 over Europe. Each curve is a centred 12-month mean.

Insofar as the primary metrics of forecast quality should reflect their use in the real world, CRPSS is a much more relevant metric than ACC. As such, CRPSS deserves to have much higher visibility amongst ECMWF's customers than it currently does! Hence, to close, we show a timeseries of CRPSS for the ENS (Figure 2). The ENS skill has improved substantially over the years. More directly, in the language of the simple economic decision-making framework that we have used in this note, the potential economic value of the ECMWF ensemble has doubled over the last 15 years. In this idealised framework, if a perfect forecasting system (the oracle) would save the European economy €100 billion, then a forecasting system with a CRPSS of 0.2 would realise €20 billion of this saving. Doubling the CRPSS from 0.1 to 0.2 means doubling the savings from €10 billion to €20 billion.

FURTHER READING

The quote at the beginning of this note is from an article by Alan Murphy published in 1993 entitled "What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting" (*Wea. Forecasting*, **8**, 281–293).

The potential economic value shown in Figure 1 is regularly updated on the ECMWF web site, as are the CRPSS headline scores for the ENS. For these and other routine verification see: http://www.ecmwf.int/en/forecasts/tools-and-guidance/quality-our-forecasts

For more information on the cost-loss model and the potential economic value of the ECMWF ensemble (including limitations and extensions) see:

Murphy, **A.H.**, 1966: A note on the utility of probabilistic predictions and the probability score in the cost-loss ratio decision situation. *J. Appl. Meteorol.*, **5**, 534–537.

Richardson, **D.S.**, 2012: Economic value and skill. In *Forecast Verification: a Practitioner's Guide in Atmospheric Science (2nd edition)*, Jolliffe, I.T. and Stephenson, D.B., Eds., Wiley, 249 pp.

Richardson, **D.S.**, 2001: Measures of skill and value of ensemble prediction systems, their interrelationship and the effect of ensemble size. *Q. J. R. Meteorol. Soc.*, **127**, 2473–2489.

Richardson, **D.S.**, 2000. Skill and relative economic value of the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.*, **126**, 649–668.

Calibration of ensemble forecasts is reviewed in:

Gneiting, **T.**, 2014: Calibration of medium-range weather forecasts. *ECMWF Tech. Memo. No. 719.* http://old.ecmwf.int/publications/library/do/references/show?id=91014