Royal Netherlands
Meteorological Institute
Ministry of Infrastructure and the Environment

Value assessment of extreme weather warnings in Holland

Kees Kok, Frits Koek and Ben Wichers Schreur

Issuing alerts for extreme events is the outcome of a tedious appraisal between avoiding misses and limiting the number of false alarms. In operational practice, many non-meteorological effects are taken into account. However, as demonstrated by Murphy, the most valuable basis for general decision strategies are probabilistic forecasts reflecting the 'true' meteorological skill of the forecasting system. Similarly, the intrinsic value of extreme weather forecasts can only be demonstrated by probabilistic verification.

1. Introduction

In the Netherlands KNMI is responsible for issuing so-called weather alerts and severe weather warnings to the general public. This is restricted to a limited number of predictands and only if a sufficiently large area (50x50km^2) is affected. On the basis of numerical output and observations the human forecaster assesses the probability of such events. These probabilities form the basis for the alert policy: in principle a probability of over 60% yields code orange, over 90% code red (Fig. A). But the actual alert level is decided by a so-called decision team and depends on the alleged vulnerability and compliance of 'the public' and the anticipated credibility damage in case of a miss or a false alarm. In the next sections an illustration is given of the way the evaluation of the probability forecasts is done over the 3 months winter period of 2010/11.

2. Data

For many extreme events the assessment whether an event happened cannot uniquely be determined from station observations and is therefore to some extent a subjective matter. The event occurrence then also must be expressed in terms of probabilities. This should be done using not only meteorological data but also information obtained from e.g. damage reports, eye-witness accounts and newspapers. However, in this experiment mostly station data are used.

The winter period contained about a dozen severe snow and freezing rain or sleet situations. In most cases it was not quite clear whether the severe weather criteria were met, and estimates had to be made.

Figure A. *Example of an extreme weather warning. Yellow indicates warning (for snow in this case); orange extreme weather warning; red weather alert.*

The criterion for snow is: at least 3cm/hour or 10cm/6 hours in an area of 50x50km^2. For freezing rain (or sleet) the criterion is met if an area of that size is affected.

3. Preliminary Results

In the verification we exclude cases in which the event already occurred somewhere in the Netherlands and the warning was subsequently issued for a different area. Although severe weather warnings are issued separately for the 12 Dutch provinces we disregard location mismatches. Results for the snow and freezing rain situations are combined.

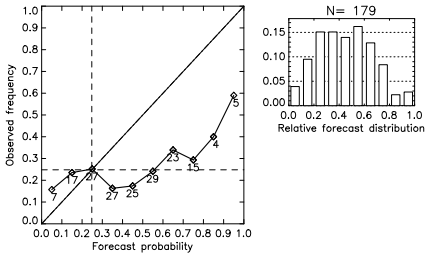

Figure B1. *Reliability plot for all probability forecasts with lead time of 12 hours or less.*

In the 2 reliability plots we have used all probability forecasts issued in a time range of 12 hours (Fig. B1) and 2 hours (Fig. B2) before the predicted event. Usually 3 or more forecasters assessed probabilities for the same event, sometimes at several lead times. So the cases are by no means independent.

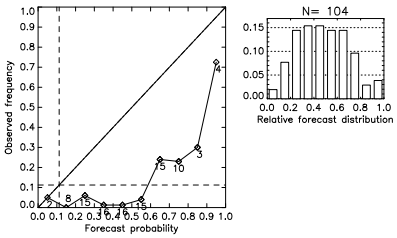

Figure B2. *Reliability plot for all probability forecasts with lead time of 2 hours or less. This is a subset of Figure B1.*

All 'no forecast – no event' cases are excluded. The observed frequency in the reliability plots is simply the mean per forecast bin of the estimated occurrence probabilities.

Event probabilities seem to be highly overestimated on average. There is little discrimination over a wide range of forecast values even for the very short lead times. The reliability seems poor. Only the high-confidence forecasts (low and high probabilities) seem to have some skill.

4. Conclusion

The very preliminary results seem rather disappointing but they may have been severely flawed by the rather crude first estimate of event occurrences. The number of events is too small to draw firm conclusions. Therefore we plan to extend the experiment to sub-extremes and to regional scales. This also provides highly necessary training opportunities for the forecasters.

* Royal Netherlands Meteorological Institute. PO Box 201, 3730 AE De Bilt, The Netherlands, E-mail: Kees.Kok@knmi.nl