The Signal Corps Meteorological Service (. . .) in WWI issued forecasts that included a statement as to the probable accuracy of the forecast, (. . .) expressed in terms of the odds in favour of the forecast. (. . .) the inclusion of this information made it possible “to make the forecast absolutely definite and such qualifications as ‘probable’ or ‘possibly’ have never been used”

Murphy (1998)

Due to the inherent uncertainties, any forecast (in particular for weather and climate) should include a statement as to its probable accuracy. One option are probabilities.

Scoring rules

Problem: Evaluation

How to compare forecasts with observations, as these are two unlike objects?

- Observations $\in \{1 \ldots K\}$
- Forecasts $q = (q^{(1)} \ldots q^{(K)})$, with $q^{(k)}$ 0, $\sum_k q^{(k)} = 1$.

A Scoring Rule $S(q; \;)$ assigns “points” to q based on the observation .

Convention: A smaller score indicates a better forecast.

[redacted]

Reliability

Reliability means: forecast probabilities should agree with actually observed relative frequencies.

> There should be rain on 20% of those days where the forecast for rain was 0.2.

More rigorous definition:

$$P(\quad = k | q) = q^{(k)}; \qquad k = 1 \ldots K$$

Proper scoring rules decomposition

- $^{(k)}(q) = P(\quad = kjq)$
- $\overline{}^{(k)} = P(\quad = k)$
- Reliability: $^{(k)}(q) = q^{(k)}$
- No resolution: $^{(k)}(q) = \overline{}^{(k)}$

Proper scoring rules decomposition

Part II

How to misinterprete probability forecasts

Probabilities in climate

Question Will hurricanes become more frequently in the future?

Forecast Probabilities $p^{(0)} \ldots p^{(K)}$ that there will be $0 \ldots K$ hurricanes per year in the North Atlantic.

An example
multiple shots

Let $_n \in \{0;1\}; n = 1 \dots N$. Your reward on average per n is $R = a + b\frac{1}{N}\sum_n \;_n$. Forecaster says: “The probability of $= 1$ is q”, but he probably means: “The long–term frequency of $_n = 1$ is about q”.

We make this precise: $q \in [0;1]$ is a random variable, and given q, the $_n$ are iid with expectation q.

Then

$$ER = a + b\bar{q}$$

$$\mathrm{Var}R = b^2 \quad \bar{q}$$

Second order probabilities?

A rigorous way to deal with these problems co

References I

Jochen Bröcker. Reliability, suf ciency, and the decompositio n of proper scores. Quarterly Journal of the Royal Meteorological Society, 135(643):1512 – 1519, 2009.

Jochen Bröcker and Leonard A. Smith. Scoring probabilistic forecasts: The importance of being proper. Weather and Forecasting, 22(2):382–388, 2007.

Thomas A. Brown. Probabilistic forecasts and reproducing scoring systems. Technical Report RM–6299–ARPA, RAND Corporation, Santa Monica, CA, June 1970.

Allan H. Murphy. The early history of probability forecasts: some extensions and clari cations. Weather and Forecasting, 13:5–15, 1998.

References II

Allan H. Murphy and Robert L. Winkler. A general framework for forecast veri cation. Monthly Weather Review, 115: