Probabilistic Forecasts in Meteorology

Caren Marzban
Dept. of Statistics, University of Washington
and
Center for Analysis and Prediction of Storms
Univ. of Oklahoma
http://www.nhn.ou.edu/~marzban

Performance is a multifaceted thing.

A lot has been done. A book?

Murphy and Winkler in atmospheric sciences.

Diagrams, diagrams, diagrams.

Event (t) Observation (t) Target (t)	Prediction (f) Forecast (f) Output (f)	Example Method
categ	categ	C-table
categ	**prob**	ROC, reliability diagram, etc.
cont	cont	Residual plot
cont	prob	Rank histogram

$$\begin{array}{ll} \text{Forecast} & f_i : i = 0, 1, 2, ..., I-1 \\ \text{Target} & t_j : j = 0, 1, 2, ..., J-1 \end{array}$$

$$p(f_i, t_j) \qquad \text{has it all.}$$

Two useful decompositions of $p(f, t)$:

$$\begin{array}{lll} p(f_i, t_j) = p(f_i|t_j)p(t_j) & & \text{Likelihood-Base Rate} \\ p(f_i, t_j) = p(t_j|f_i)p(f_i) & & \text{Calibration-Refinement} \end{array}$$

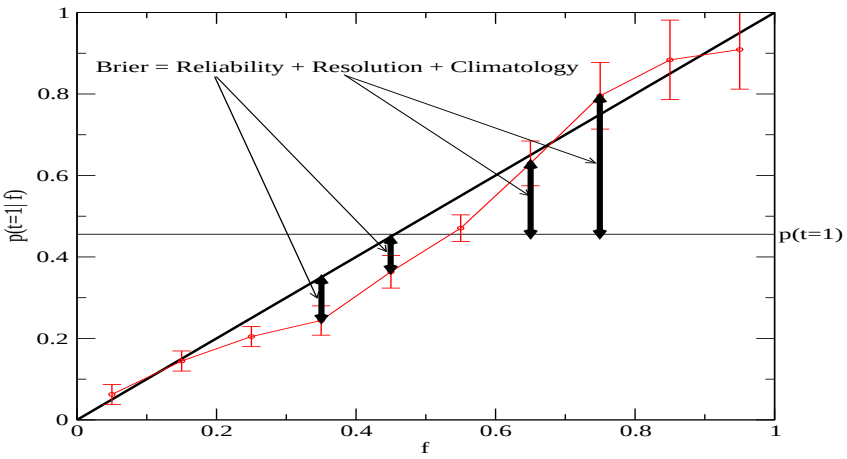
Connections:

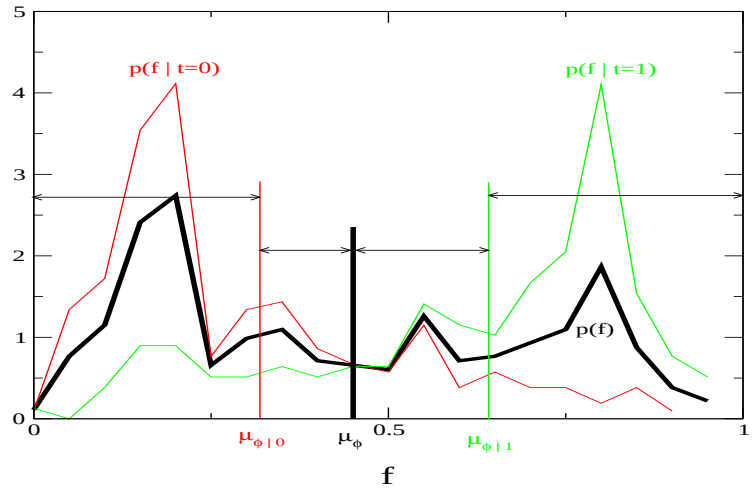
MSE(f,t)	Accuracy
$p(f = 1\|t = 1)$ vs. $p(f = 1\|t = 0)$	ROC
$p(t = 1\|f)$ vs. f	Reliability/Calibration
Plot of $p(f)$	Refinement/Sharpness
Plot of $p(f\|t = 0)$ and $p(f\|t = 1)$	Discrimination
Plot of ????????	Resolution
Plot of $p(t)$	Base rate (not an attribute of f)

Need at least TWO. E.g. ROC and Reliability.

Decompositions of MSE:

$$
\begin{aligned}
\text{Brier Score} &= \Sigma(t_i - f_i)^2 && \text{MSE(f,t)} \\
&= \sigma_t^2 && \text{MSE}(\mu_t\text{,t), prior} \\
&+E_f(\mu_{t|f} - f)^2 && \text{Reliability} \\
&-E_f(\mu_{t|f} - \mu_t)^2 && \text{Resolution} \\
\\
&= \sigma_f^2 && \text{Refinement} \\
&+E_t(\mu_{f|t} - t)^2 && \text{Discrimination+} \\
&-E_t(\mu_{f|t} - \mu_f)^2 && \text{Discrimination-}
\end{aligned}
$$



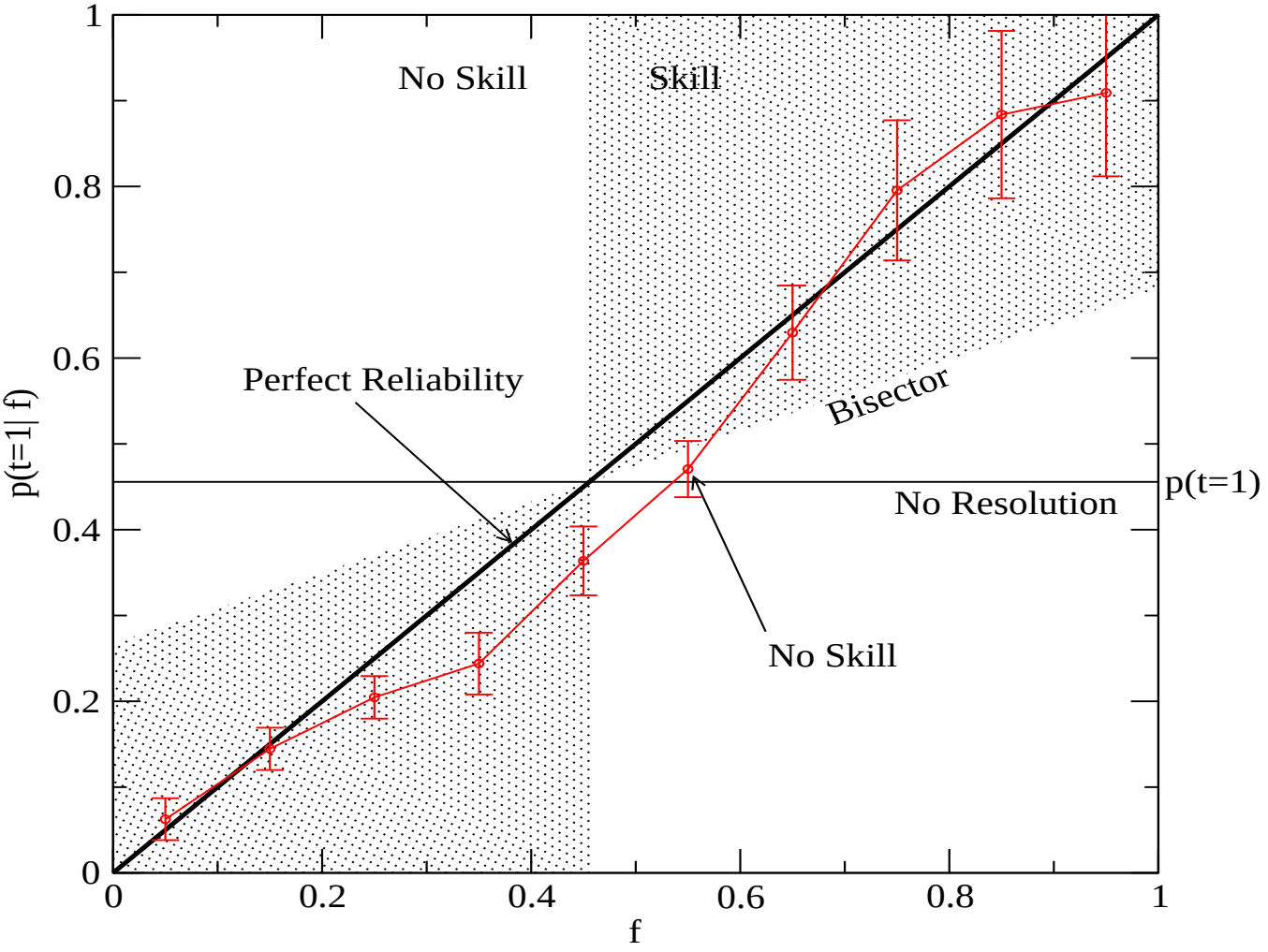


One more concept: Skill

Performance above and beyond a standard (e.g. prior prob)

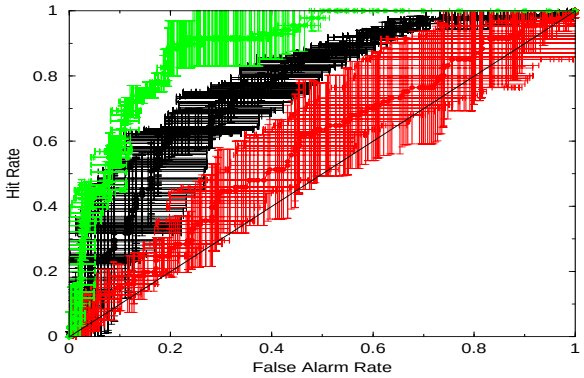
$$\text{Brier Skill Score} = 1 - \frac{MSE(f,t)}{MSE(p(t=1),t)}$$

All together:

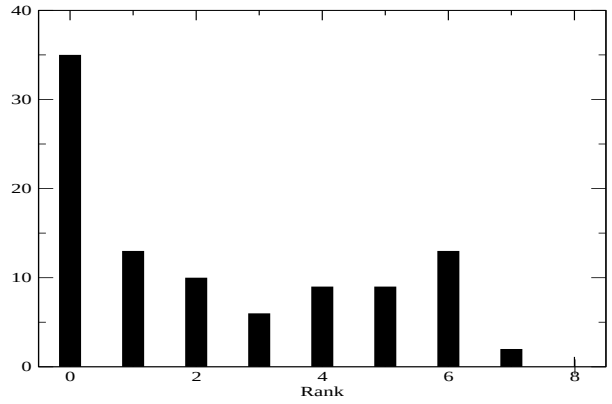


Summary

Performance is multifaceted.
Probabilistic forecasting even more so.
$p(f, t)$ contains all relevant information.
Accuracy, Reliability, Refinement, Resolution, R(D)iscrimination
Diagrams faithful to multidimensionality.
Brier has problems.
Try Ranked Probability Score (MSE in cumulative prob space).



Error-bars on attribute diagram and ROC.



Continuous target. Rank Histogram

Useful Links

An excellent FAQ directed at the atmospheric community:
www.bom.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.html

A Book:
www.met.rdg.ac.uk/cag/publications/Tableofcontents.pdf

Probabilistic forecasts of continuous quantities:
www.stat.washington.edu/tilmann/

Interpreting ROC (and AUC) in terms of the parameters of $p(f, t)$:
www.nhn.ou.edu/~marzban/roc.pdf

An R package for verification:
www.r-project.org/ [install.packages(verification)]