BAYESIAN VERIFICATION MEASURES for FORECASTS of CONTINUOUS PREDICTANDS

By

Roman Krzysztofowicz
University of Virginia

*Presented at the*
18th Conference on Probability and Statistics
in the Atmospheric Sciences
Atlanta, Georgia

29 January – 2 February 2006

Acknowledgments:

Work supported by the National Science Foundation under Grant No. ATM–0135940.
Data provided by the Meteorological Development Laboratory of the National Weather Service.
BAYESIAN THEORY
OF FORECAST VERIFICATION

Forecast
• provides information for decision making
• reduces uncertainty about a predictand
• has economic value

Attributes

“Can the forecast be taken at its face value?”

CALIBRATION – necessary for consistent interpretability
  – attainable through a transformation (re-calibration)

“Does the forecast reduce the uncertainty?”

INFORMATIVENESS – necessary for positive economic value
  – intrinsic to the forecast system
DECISION-THEORETIC EVALUATION

VARIATES: Predictand $W$, $w \in \mathcal{W}$
Forecast $X$, $x \in \mathcal{X}$

FORECAST-DECISION SYSTEM

likelihood $f(x|w)$
prior $g(w)$
utility $u(a, w)$

BAYESIAN RATIONALITY

Bayes utility of forecaster for decision maker

$$U(g, u) = \int \left[ \max_{\mathcal{X}} \int_{\mathcal{W}} u(a, w) f(x|w) g(w) \, dw \right] dx$$

COMPARISON

likelihood Bayes utility
Forecaster $X$ $f_X(x|w)$ $U_X(g, u)$
Forecaster $Y$ $f_Y(y|w)$ $U_Y(g, u)$
THEORY OF SUFFICIENT COMPARISONS

DEFINITION

X is more informative than Y if

\[ U_X(g, u) \geq U_Y(g, u) \quad \forall g, u \]

(for every rational decision maker)

DEFINITION

X is sufficient for Y if there exists a stochastic transformation \( \psi \) such that

\[ f_Y(y|w) = \int \ldots \int \psi(y|x)f_X(x|w) \, dx \quad \forall y, w \]

THEOREM (Blackwell, 1951)

If X is sufficient for Y,
then X is more informative than Y.
META-GAUSSIAN LIKELIHOOD

If $f_X, f_Y$ meta-Gaussian (or Gaussian), then

• analytic solution for $\psi$

• algebraic condition for existence of $\psi$

• if deterministic forecast $x = (x_1)$
  or probabilistic forecast $x = (x_1, x_2)$
  then there exists
  
  sufficiency characteristic, $SC$, …

THEOREM (Krzysztofowicz, 1987; Long, 1990)

$SC_X > SC_Y \iff X$ sufficient for $Y$

$\implies X$ more informative than $Y$
DETERMINISTIC FORECAST

DISTRIBUTIONS

Predictand $G(w)$
Forecast $K(x)$

NORMAL QUANTILE TRANSFORM (NQT)

$$V = Q^{-1}(G(W))$$
$$Z = Q^{-1}(K(X))$$

LINEAR REGRESSION

$$E(Z|V = v) = av + b$$
$$Var(Z|V = v) = \sigma^2$$

SUFFICIENCY CHARACTERISTIC

$$SC = \frac{a^2}{\sigma^2}$$

“signal”

“noise”

(uninformative) $0 \leq SC \leq \infty$ (clairvoyant)
**DETERMINISTIC FORECAST**

**INFORMATIVENESS SCORE**

\[
IS = \left[ \frac{1}{SC} + 1 \right]^{-\frac{1}{2}}
\]

\[
IS = \left[ \frac{\sigma^2}{a^2} + 1 \right]^{-\frac{1}{2}}
\]

(uninformative) \(0 \leq IS \leq 1\) (clairvoyant)

**PROPERTIES**

1. Interpretable as “Rank Correlation” between \(X\) and \(W\)

2. Orders forecasters consistently with *Bayes Utility* (or *Economic Value*)

3. Meaningful for cross-comparisons
KUIL 12-36h, Cool

W - Precip. Amount, when W > 0

X - Model Estimate, when W > 0

\[ \rho = 0.613 \]

\[ IS = 0.631 \]

\[ a = 0.681 \]

\[ b = 0.028 \]

\[ \sigma = 0.829 \]
PROBABILISTIC FORECAST

FORECAST: \( \mathbf{X} = (X_1, X_2) \)

\( X_1 \) — median of \( W \)

\( X_2 \) — width of the 50% central credible interval

INFORMATIVENESS SCORE

\[ IS = \ldots \]

PROPERTIES

1. Interpretable as “Rank Correlation” between \((X_1, X_2)\) and \( W \)

2. Orders consistently forecasters: deterministic versus probabilistic

\[ IS_{X_1} \leq IS_{(X_1, X_2)} \]