ASSESSING THE QUALITY AND ECONOMIC VALUE OF WEATHER AND CLIMATE FORECASTS

Rick Katz

Institute for Study of Society and Environment
National Center for Atmospheric Research
Boulder, CO USA

Email: rwk@ucar.edu

Web site: www.isse.ucar.edu/HP_rick.html

Lecture: www.isse.ucar.edu/HP_rick/pdf/csmkatz.pdf

QUOTES

- "*You don't get points for predicting rain. You get points for building arks.*"

Lou Gerstner

- "*Do not judge human action by what happens.*"

Jacob Bernoulli

OUTLINE

(1) Forecast Verification

(2) Decision-Making Framework

(3) Economic Value of Forecasts

(4) Cost-Loss Decision-Making Model

(5) Quality-Value Relationships

(6) Valuation Puzzles

(7) Resources

(1) FORECAST VERIFICATION

(1.1) History

- Finley's Tornado Forecasts (1884)

	Observed	
Forecast	Tornado	No Tornado
Tornado	$n_{11} = 28$	$n_{10} = 72$
No Tornado	$n_{01} = 23$	$n_{00} = 2680$

-- 96.6% correct

(1) FORECAST VERIFICATION

(1.1) History

- Finley's Tornado Forecasts (1884)

	Observed	
Forecast	Tornado	No Tornado
Tornado	$n_{11} = 28$	$n_{10} = 72$
No Tornado	$n_{01} = 23$	$n_{00} = 2680$

-- 96.6% correct (But 98.2% correct if never forecast tornado!)

- **Peirce's Skill Score ("The Numerical Measure of the Success of Predictions", *Science*, 1884)**

-- Proposed "science of the method" measure for forecast evaluation

-- Model as if produced by two idealized observers

(i) Science of the method (Infallible observer)

(ii) Random guessing (Ignorant observer)

$$\mathrm{PSS} = n_{11} / (n_{11} + n_{01}) - n_{10} / (n_{10} + n_{00}) \approx 0.549 - 0.026 \approx 0.523$$

Estimates "Hit Rate" – "False Alarm Rate"

PSS = 0 if never forecast tornado

(1.2) FRAMEWORK

- **Joint Distribution of Observations & Forecasts**

  Observed Weather $\Theta = \theta$

  Forecast Weather $Z = z$ **("categorical")**

-- Factorizations (Murphy & Winkler)

***Calibration - Refinement*:** $f(\theta, z) = g(\theta \mid z)\, h_Z(z)$

***Likelihood - Base Rate*:** $f(\theta, z) = L(z \mid \theta)\, h_\Theta(\theta)$

(1.3) PROBABILITY FORECASTS

Forecast probability p (e.g., induced by Z)

- ***Reliability Diagram***

Observed weather:

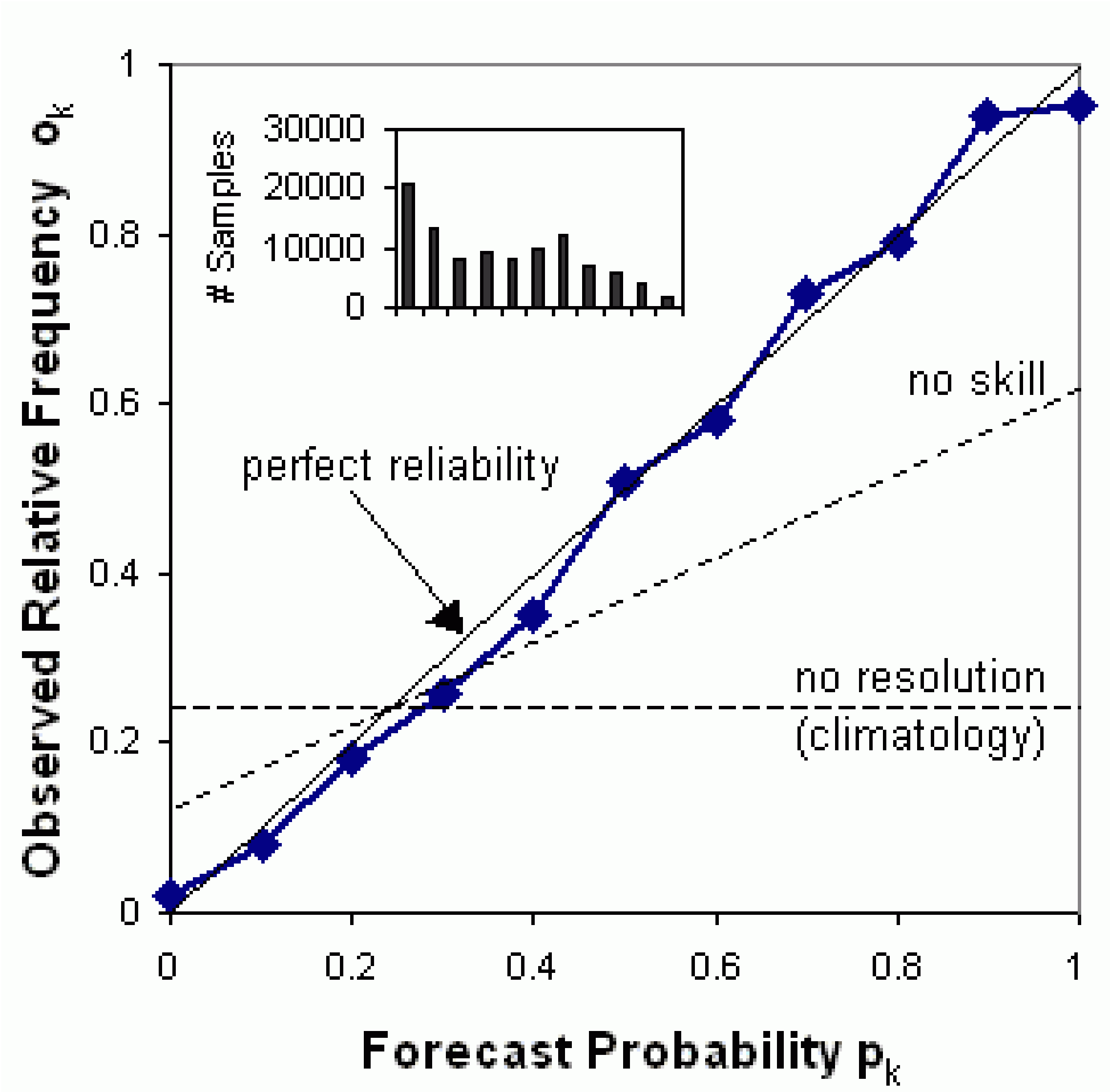
$\Theta = 1$ **(Adverse weather occurs)**

$\Theta = 0$ **(*No* adverse weather)**

$h(p) = \Pr\{\Theta = 1 \mid p\}$ **(Cond. prob. of averse weather given forecast p)**

Plot estimated $h(p)$ versus p

"Perfect Reliability": $h(p) = p, \ 0 \leq p \leq 1$



Observed Relative Frequency o_k
Forecast Probability p_k
Samples
30000
20000
10000
0
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
perfect reliability
no skill
no resolution
(climatology)

- **“Sharpness” or “Resolution”**

-- Reliability necessary, but not sufficient

-- Also need measure of spread of forecast probability

(Ideally issue forecast probabilities close to zero or one)

- **Recent Statistical Literature**

-- “Strictly Proper Scoring Rules, Prediction, and Estimation”

(Gneiting & Raftery, *JASA*, 2007)

-- “Probabilistic Forecasts, Calibration and Sharpness”

(Gneiting et al., *JRSSB*, 2007)

(2) DECISION-MAKING FRAMEWORK

(2.1) ELEMENTS OF DECISION MAKING

- **Events ($\Theta = \theta$)**

-- Adverse weather event (such as "Rain")

- **Actions (*a*)**

-- Protection (to prevent or reduce impact of event)

- **Consequences [$w(a, \theta)$]**

-- Action-Event pair

(2.2) FORECAST INFORMATION SYSTEM

- **Forecast**

-- Conditional probability distribution for event

$Z = z$ indicates forecast for particular occasion

Specifies conditional distribution of weather $(\Theta \mid Z = z)$

- **Climatological Information ("Climatology")**

-- Unconditional probability distribution of Θ

(3) ECONOMIC VALUE OF FORECASTS

(3.1) OPTIMAL USE OF PROBABILITY FORECASTS

- **Criteria for Optimality**

-- Maximize expected utility

Utility function *u*, payoff *u*[*w*(*a*, θ)]:

***Risk neutral* (*u* linear function; maximize expected return or minimize expected expense)**

***Risk aversion* (*u* concave function)**

***Risk taking* (*u* convex function)**

-- **Optimization procedure**

(i) For each possible action *a*, determine expected utility

(weighted average with respect to conditional probability distribution

for future weather as specified by forecast)

$$E_{\Theta | Z}\{u[w(a, \Theta)] \,|\, Z = z\}$$

(ii) Select action *a* for this expected utility is a maximum

(3.2) ECONOMIC VALUE OF PROBABILITY FORECASTS

- **Standard of Comparison**

-- e. g., climatological information

Other examples ("persistence", alternative forecasting system)

- **Concept of Value of Imperfect Information (VOI)**

-- Measure of how much better off decision maker is

(with vs. without forecasting system)

-- e. g., increase in expected return (or reduction in expected expense)

for forecasting system as compared to climatology

- **General Definition of VOI ("Demand Value")**

-- Maximum amount, measured in same units as those in which consequence of decision measured, that decision maker would be willing to exchange for information system

-- Formal Definition of Demand Value

$$E_Z\left(\max_a E_{\Theta|Z}\{u[w(a, \Theta) - VOI] \mid Z\}\right) = \max_a E_\Theta\{u[w(a, \Theta)]\}$$

-- Consistent with concept of "contingent valuation" in economics

(4) COST-LOSS DECISION-MAKING MODEL

(4.1) EXPENSE MATRIX [$w(a, \theta)$]

Action	Weather Event: Adverse ($\Theta = 1$)	Weather Event: Not Adverse ($\Theta = 0$)
Protect ($a = 1$)	C	C
Do Not Protect ($a = 0$)	L	0

If take protective action, then incur cost C

If do *not* protect and adverse weather event occurs, then incur loss L

(4.2) OPTIMAL USE OF PROBABILITY FORECASTS

- **Decision-Making Criterion**

-- Minimize expected expense

- **Optimal Policy**

-- Forecast probability of adverse weather event p

-- Expected Expense (*Protect*): C

-- Expected Expense (*Do Not Protect*): $(1 - p)\ 0 + p\ L = p\ L$

-- So protect if $p > C / L$

(4.3) ECONOMIC VALUE OF PROBABILITY FORECASTS

- **Climatology**

-- Climatological probability of adverse weather $p_A = \Pr\{\Theta = 1\}$

-- Minimal Expected Expense E_{CLIM}

$$E_{\text{CLIM}} = p_A L, \text{ if } p_A \leq C / L$$

$$E_{\text{CLIM}} = C, \text{ if } p_A > C / L$$

- **Forecasts (with Minimal Expected Expense** E_{FORE}**)**

-- Value of Forecasts: $E_{\text{CLIM}} - E_{\text{FORE}}$

(5) QUALITY-VALUE RELATIONSHIPS

(5.1) MODEL FOR PROBABILITY FORECASTING SYSTEM

- **Beta Distribution (for probability forecast p)**

-- Natural distribution for probabilities, $0 < p < 1$

-- Parameters r, s ($0 < r < \infty, 0 < s < \infty$)

-- Mean: $r / (r + s)$

-- Assume perfectly reliable probability forecasts

In particular, $p_A = r / (r + s)$

-- Climatology (i. e., no skill)

$r \to \infty, s \to \infty$

-- Perfect information

$r \to 0, s \to 0$

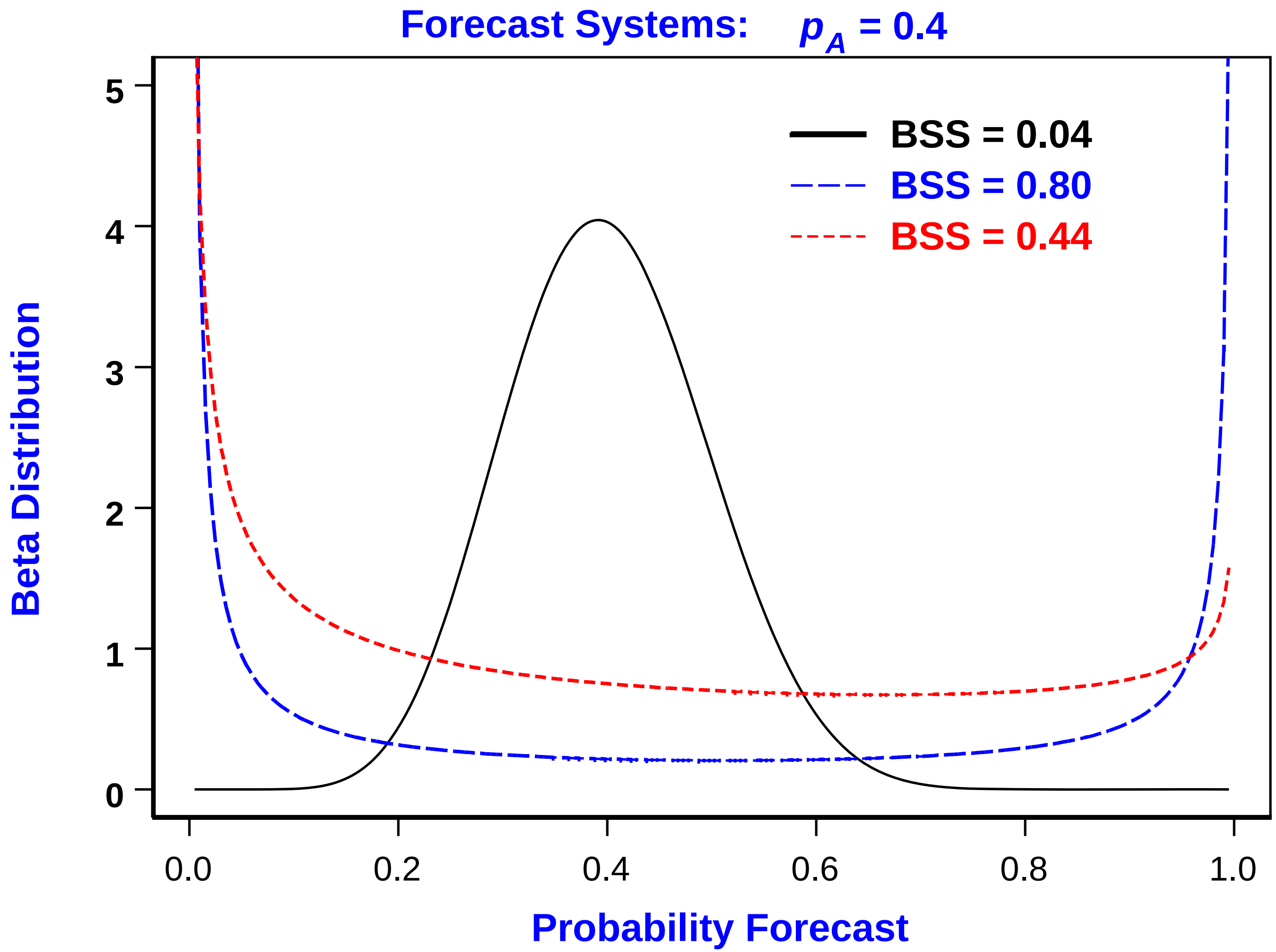
-- Brier skill score (BSS)

Brier score: $BS = E[(\Theta - p)^2]$

$BSS = 1 - [BS / Var(\Theta)]$

(BSS = 0 for climatology, BSS = 1 for perfect information)

Beta distribution: $BSS = 1 / (r + s + 1)$



Forecast Systems: p_A = 0.4
BSS = 0.04
BSS = 0.80
BSS = 0.44
Beta Distribution
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
Probability Forecast

(5.2) EXAMPLE

- **Cost-Loss Decision-Making Model**

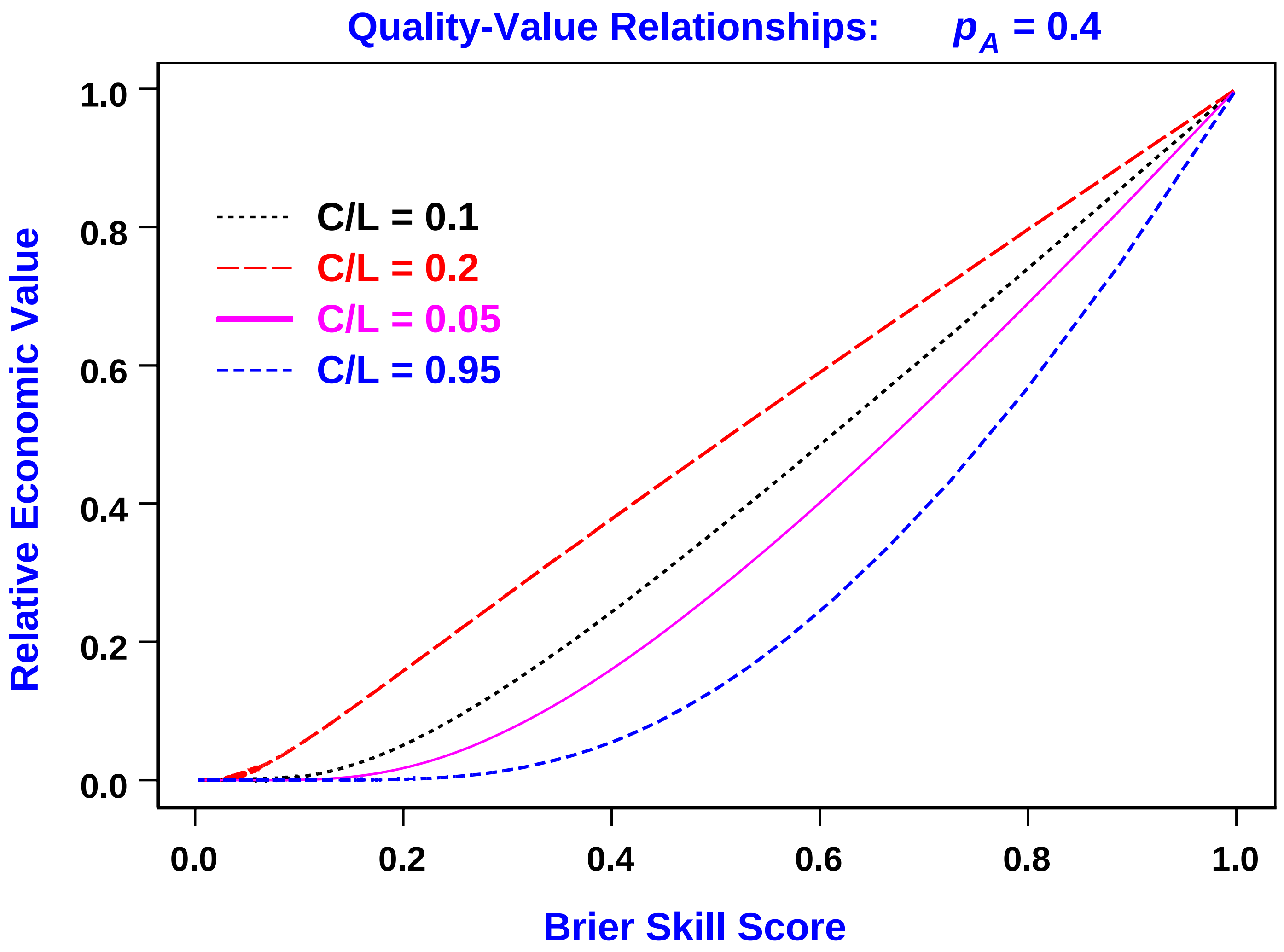
-- Fix climatological prob. of adverse weather p_A

-- Fix cost-loss ratio C / L

-- Vary skill of forecasting system (through parameters r & s of beta distribution) from no skill (i. e., climatology) to perfect information

-- Examine how economic value of forecasting system changes (on relative scale from zero for climatology to one for perfect)

-- Use expressions in Katz & Ehrendorfer (2006)


Quality-Value Relationships: p_A = 0.4
Relative Economic Value
Brier Skill Score
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
C/L = 0.1
C/L = 0.2
C/L = 0.05
C/L = 0.95

- **General Result**

-- Concept of "sufficiency" (as in "sufficient" statistic)

Criterion for comparison of "quality" of two forecasting systems

Stochastic transformation ("Randomization")

Partial ordering

-- Theorem (David Blackwell, "Equivalent Comparisons of Experiments", *Annals of Math. Stat.*, 1953)

Quality measure consistent with sufficiency

Quality-value relationships must be non-decreasing

(6) VALUATION PUZZLES

(6.1) QUALITY-VALUE REVERSALS

- **Cost-Loss Decision-Making Model**

-- Ensemble prediction system (assume perfect numerical model)

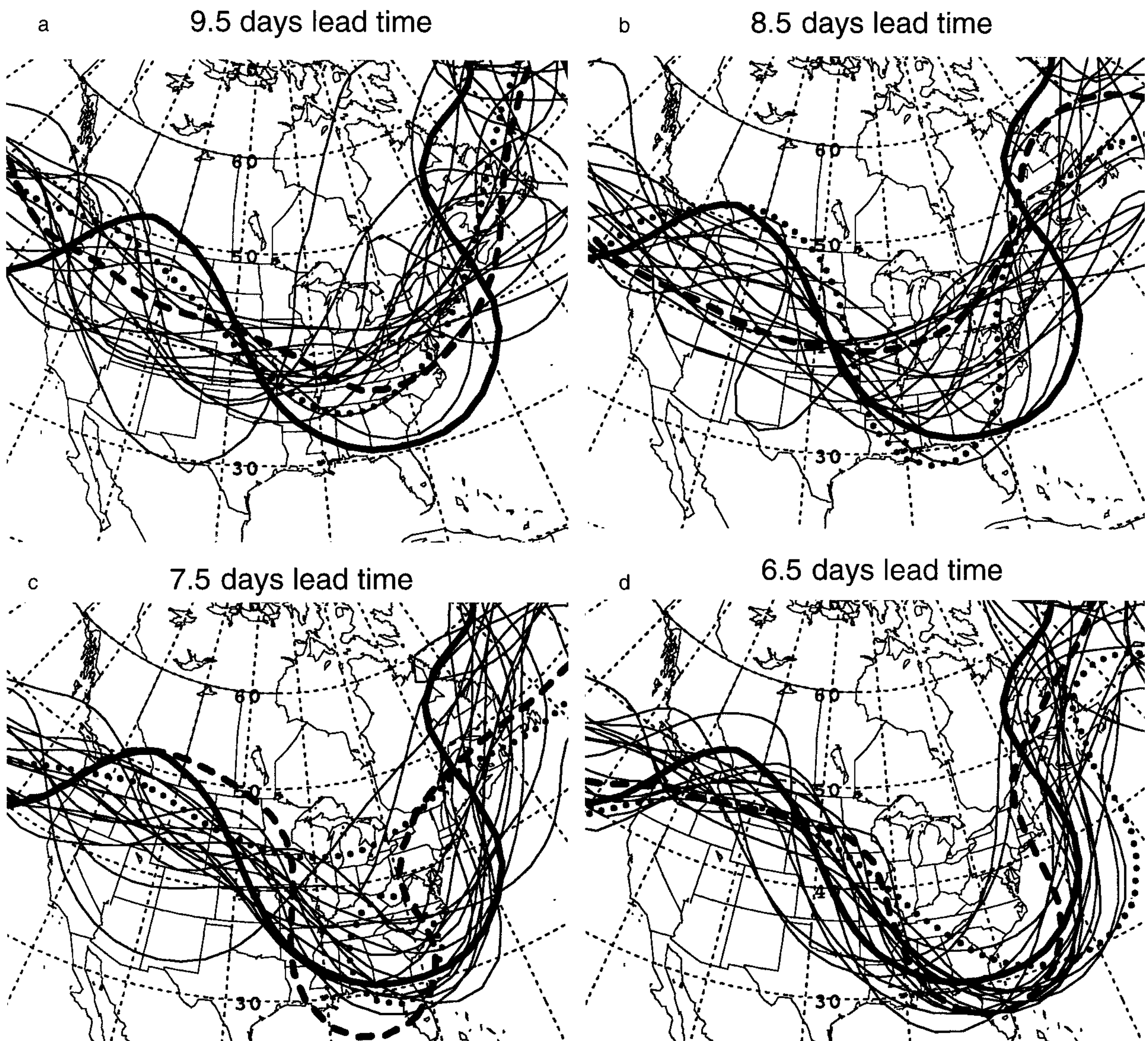
Finite number of ensembles n

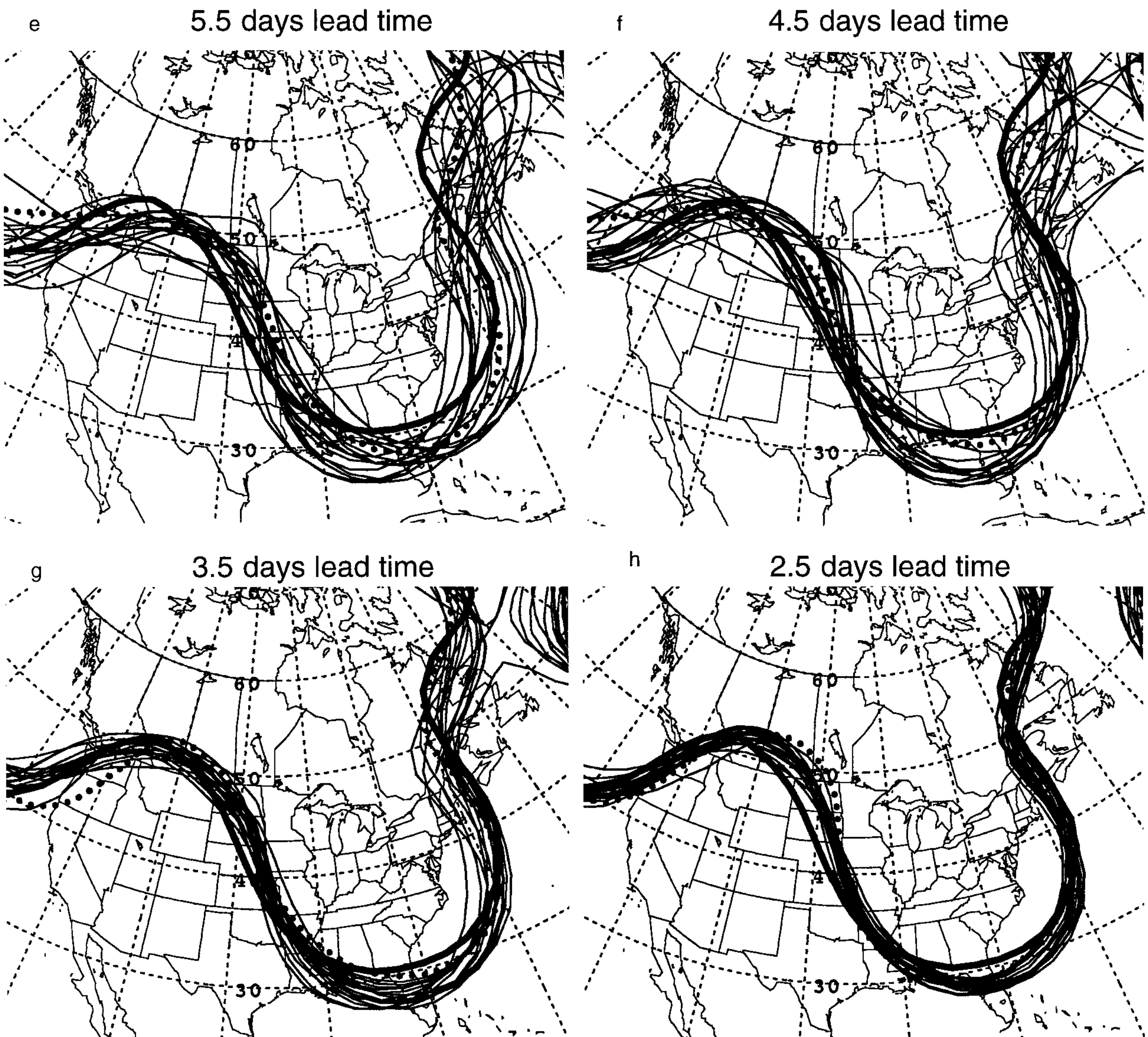
Take ensembles at "face value":

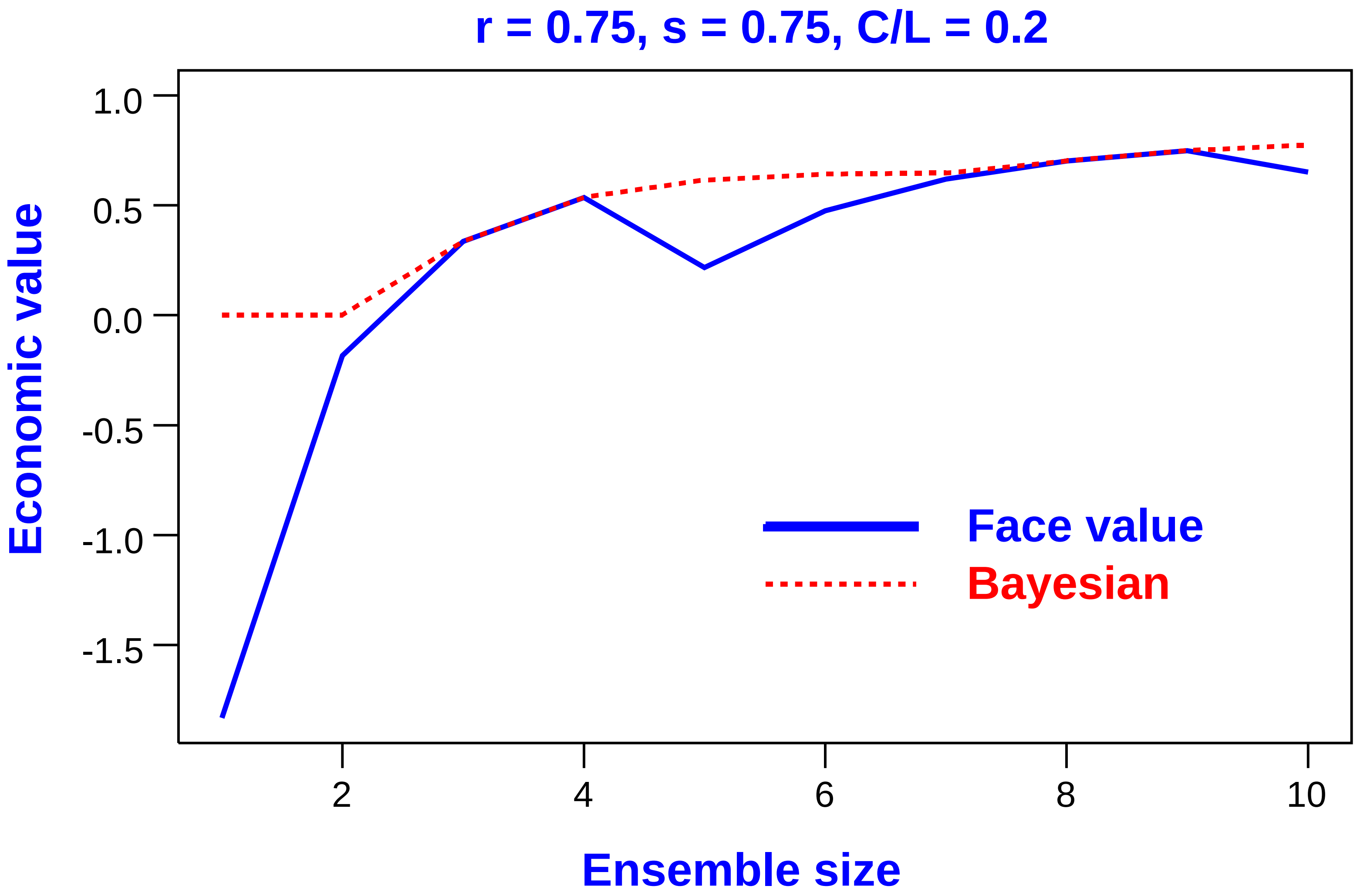
Suppose event occurs for k **out** n **ensembles**

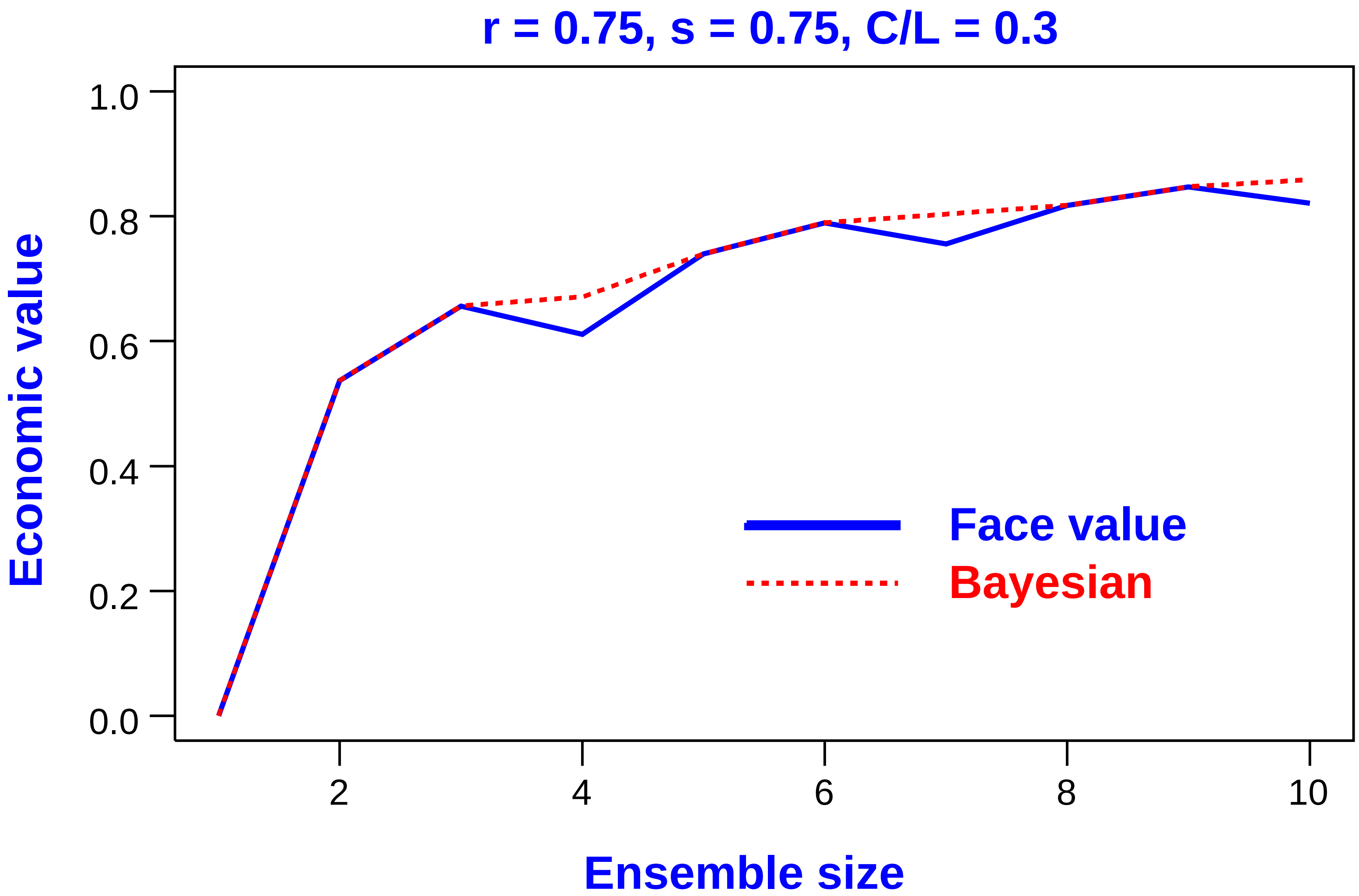
Estimate forecast probability as: $\hat{p} = k/n$

Results from Katz & Ehrendorfer (2006)


a
9.5 days lead time
60
50
30
b
8.5 days lead time
60
50
30
c
7.5 days lead time
60
30
d
6.5 days lead time
60
50
30


e
5.5 days lead time
f
4.5 days lead time
g
3.5 days lead time
h
2.5 days lead time
60
50
30


r = 0.75, s = 0.75, C/L = 0.2
1.0
0.5
0.0
-0.5
-1.0
-1.5
Economic value
2
4
6
8
10
Ensemble size
Face value
Bayesian



r = 0.75, s = 0.75, C/L = 0.3
Economic value
1.0
0.8
0.6
0.4
0.2
0.0
2
4
6
8
10
Ensemble size
Face value
Bayesian

(6.2) DEARTH OF GENERAL RESULTS

- **Attributes of decision setting & decision maker**

-- **Flexibility**

-- **Risk aversion**

-- **Wealth**

-- **Prior uncertainty**

- **No monotonic relationship between information value and any of these attributes**

(7) RESOURCES

- **Recent Case Studies of Economic Value (Self)**

`www.isse.ucar.edu/HP_rick/esig.html`

- **Forecast Verification**

-- Web page (Beth Ebert, BMRC)

`www.bom.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.html`

-- Software package in `R` (Matt Pocernich, NCAR)

`cran.r-project.org/doc/packages/verification.pdf`