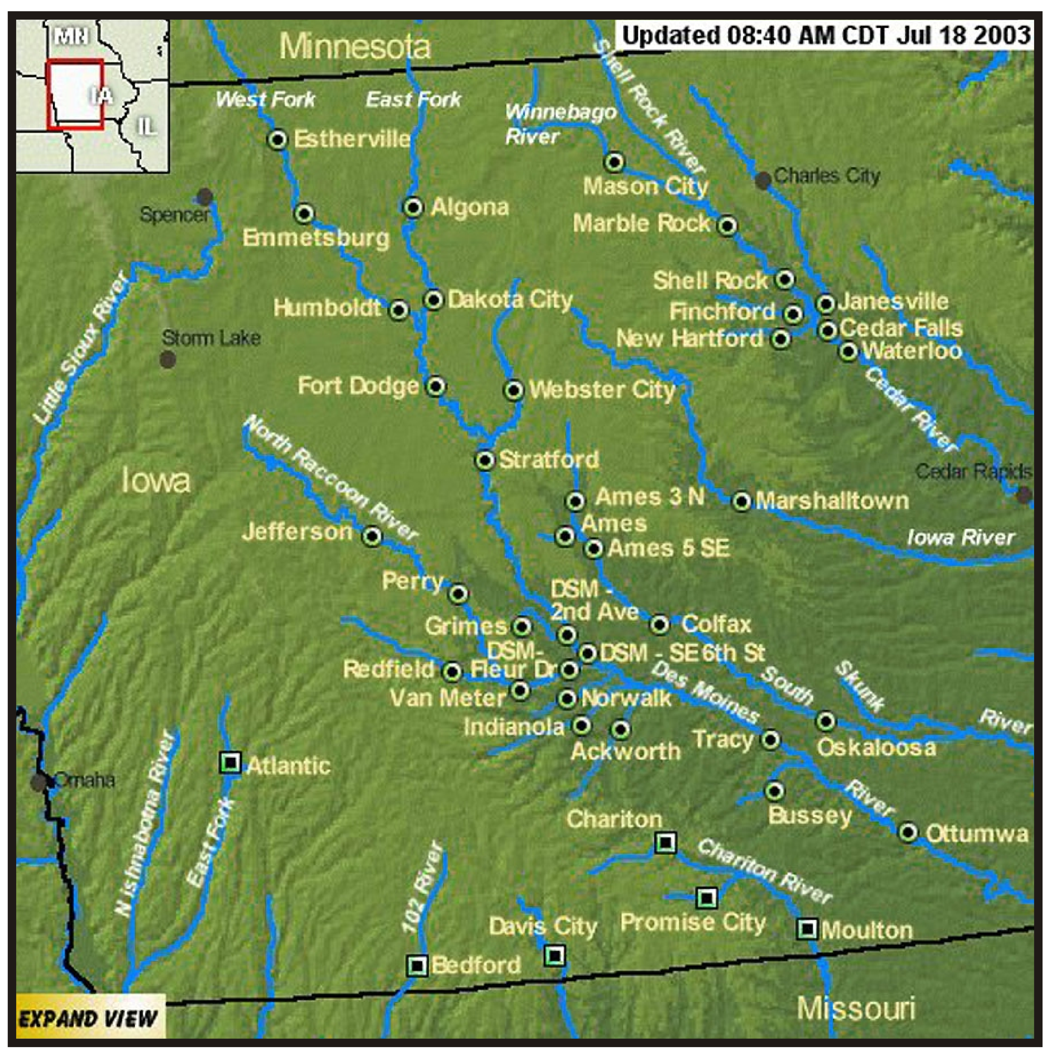
Verification of AHPS Ensemble Streamflow Predictions for the North Central River Forecast Center

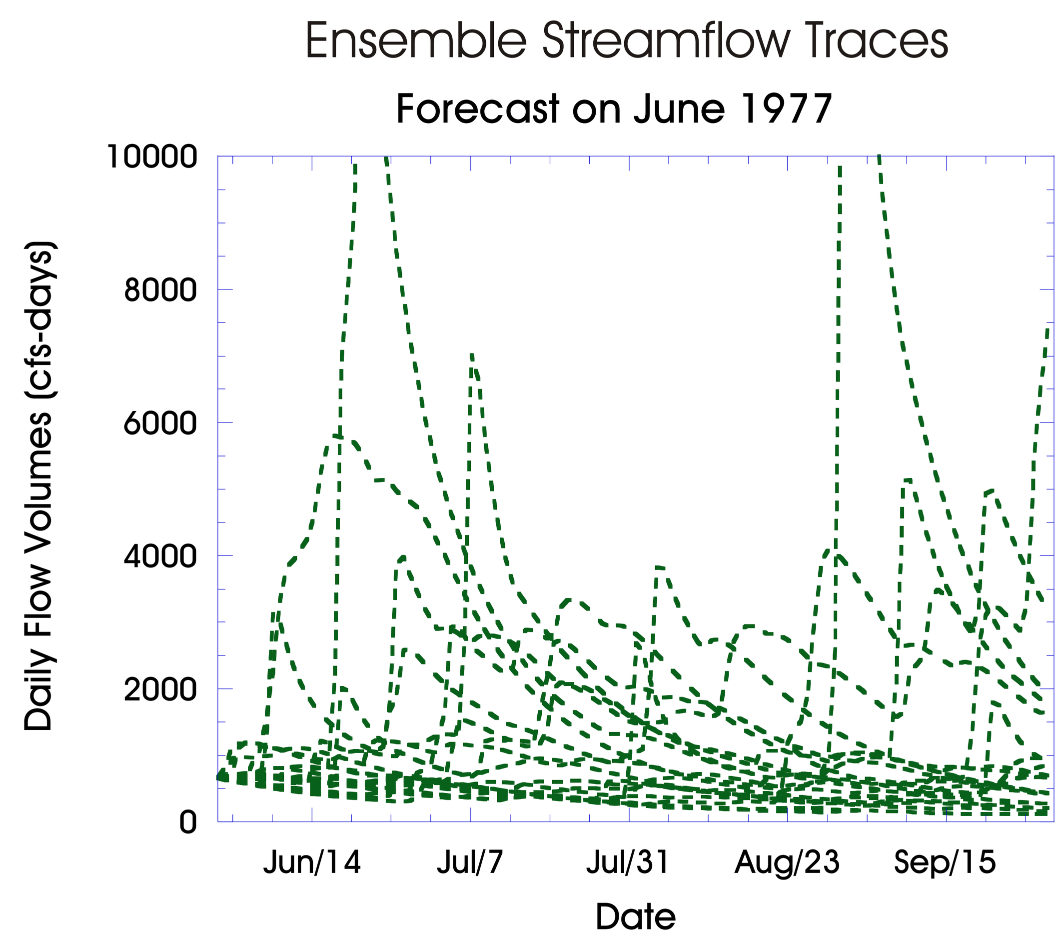
A. A. Bradley & A. Kruger
IIHR-Hydroscience & Engineering
The University of Iowa

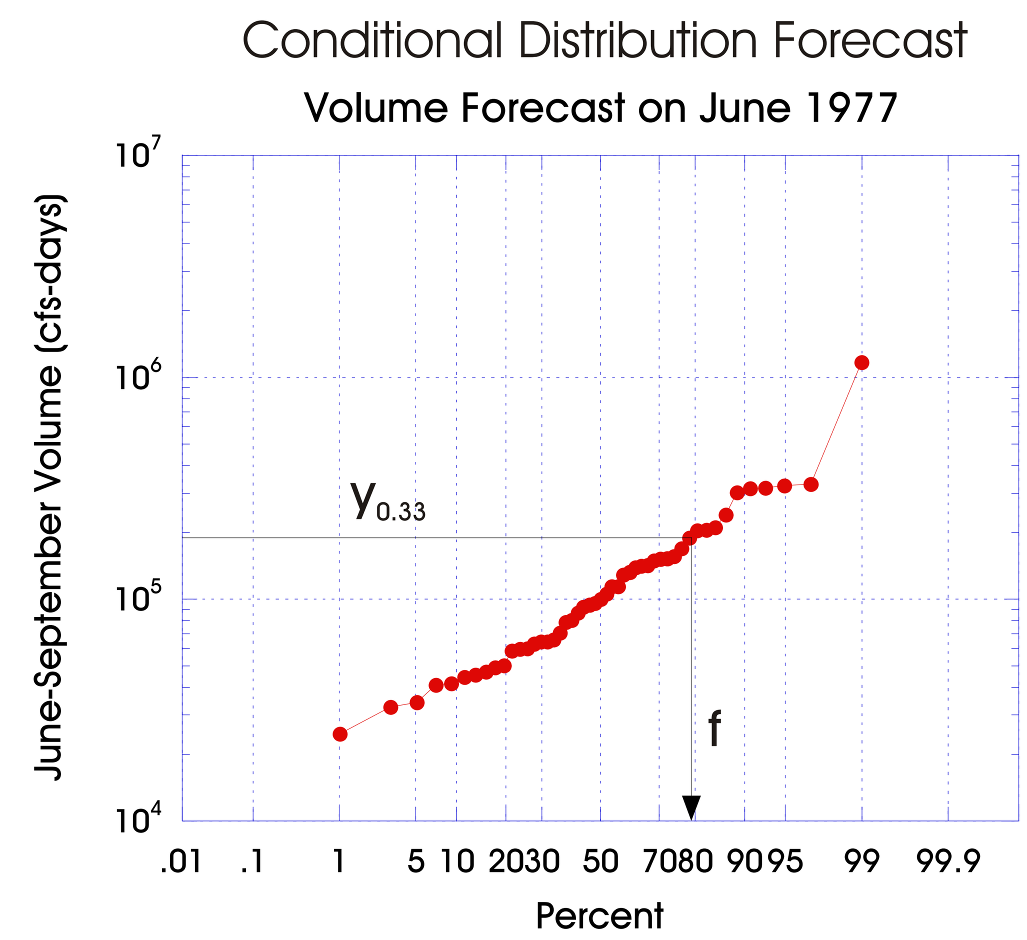
S. Schwartz
Center for Environmental Science, Technology, & Policy
Cleveland State University

Ensemble Streamflow Prediction

Des Moines AHPS







Ensemble streamflow predictions produce a probability distribution forecast for an continuous flow variable. For verification, probability forecasts (*f*) are constructed from the ensemble predictions for specific discrete events (flow above or below at threshold *y*).

Distributions-Oriented Forecast Verification

Verification of ensemble streamflow predictions (ESP) is inherently complex because the forecast is a *conditional probability distribution*.

A distributions-oriented approach is used to assess probability forecasts (*f*) and discrete observations (*x*) for specific flow levels (e.g., unconditional quantiles):

Skill: Accuracy of the probability forecast (relative to climatology forecasts)

Bias: Forecast are unbiased if the relative frequency of the observation ($_x$) equals the average forecast probability ($_F$).

Reliability: Forecasts are perfectly reliable if the relative frequency of the observations (conditioned on the forecast) equals the forecast probability (no conditional biases)

Resolution: Forecasts have resolution if the expected value of the observation (conditioned on the forecast) differs from climatology

Discrimination: Forecasts have discrimination if the characteristics of the forecasts differ for different outcomes

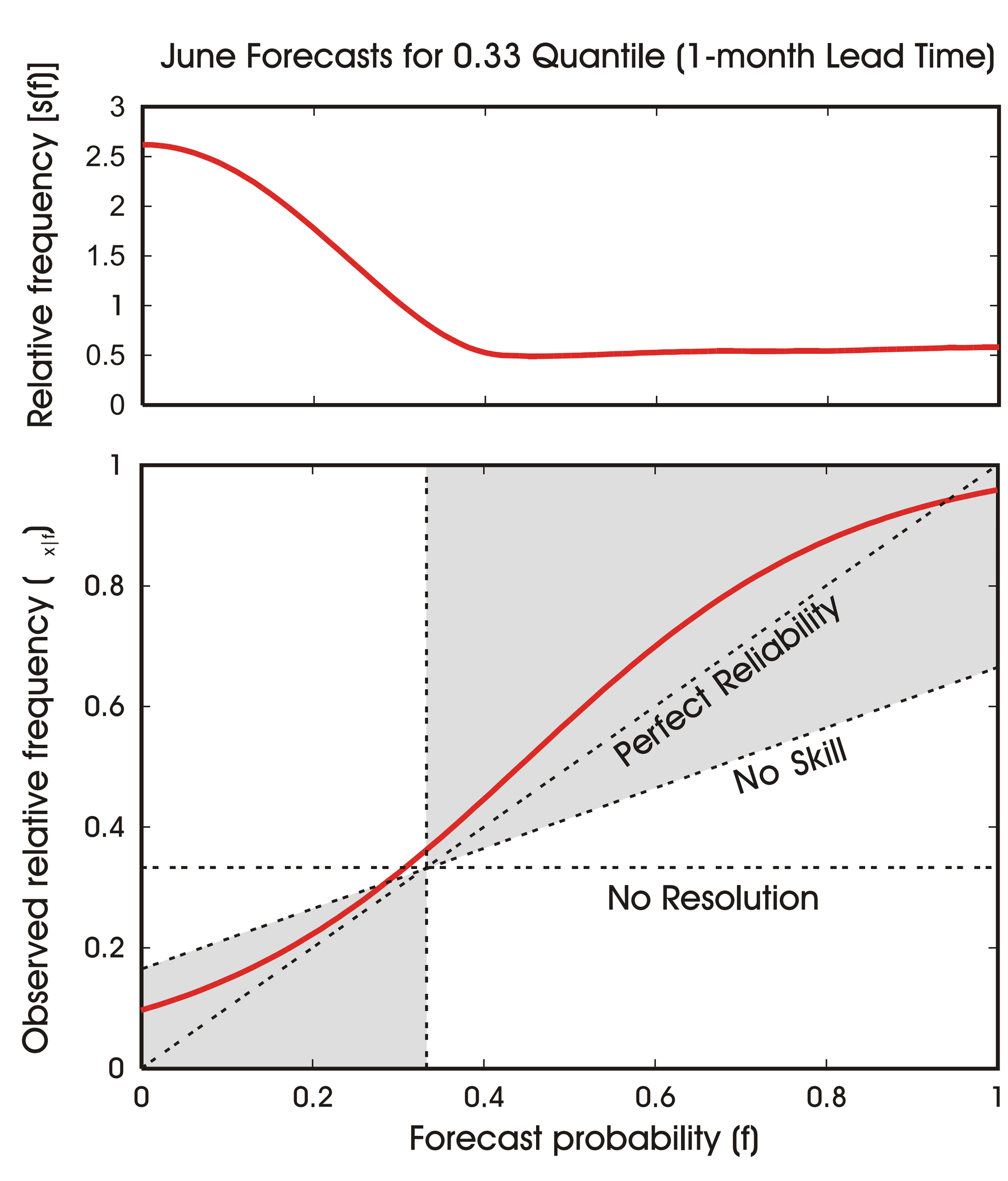
Forecast-Observation Pairs

Year	f	x	y
1949	0.630	1	*72320*
1950	0.394	1	*198440*
1951	0.005	0	*677080*
1952	0.046	0	*259610*
1953	0.051	0	*303070*
1954	0.058	0	*591810*
1955	0.935	1	*83020*
1956	0.856	1	*28160*
1957	0.188	1	*104240*
:	:	:	:

Probabilistic forecasts (f) and discrete observations (x) constructed from the ESP conditional distribution for the 0.33 quantile ($y_{.33}$=198,660 cfs-days).

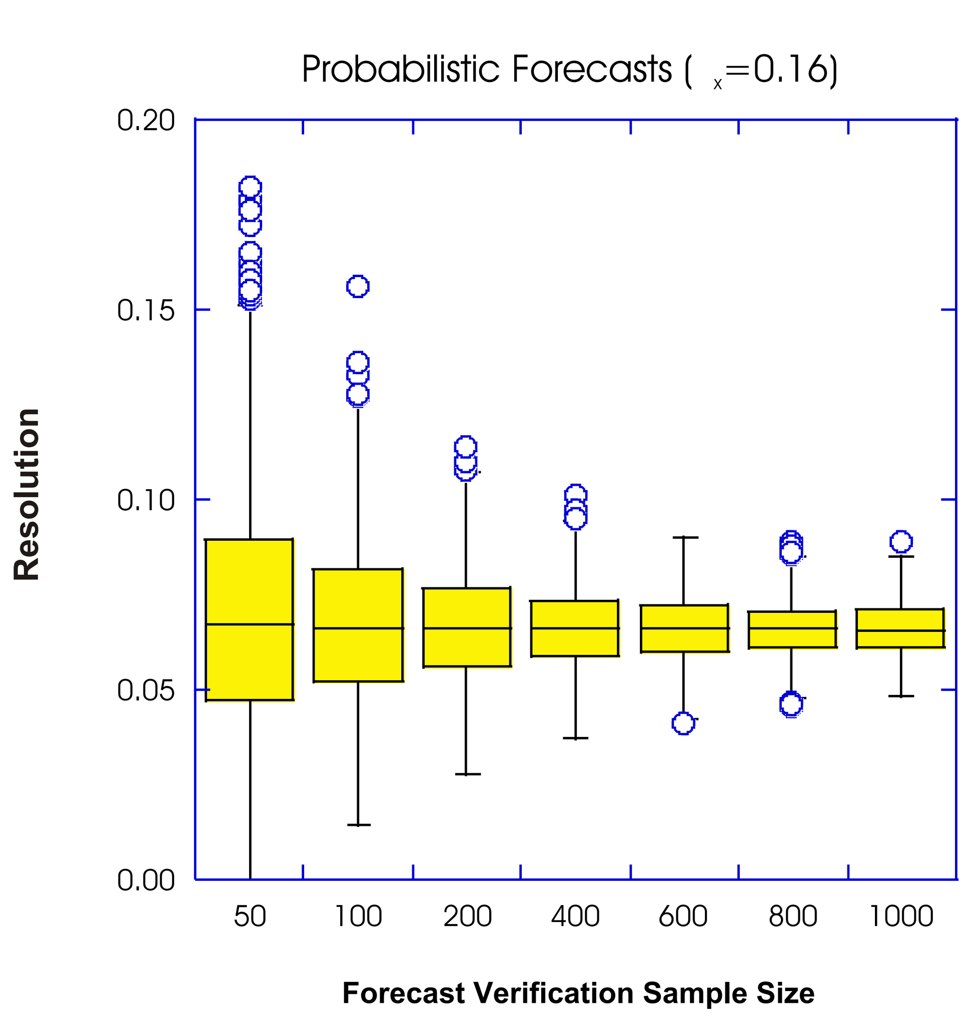
The forecast (f) is the probability that the observed flow (y) will be less than the $y_{.33}$ (i.e., $Y<y_{.33}$).

Reliability Diagram



Uncertainty Estimation

Effects of Sampling Variability on Forecast Quality Measures



Due to the small size of hydrologic forecast samples, forecast quality estimates have large uncertainties (see left figure).

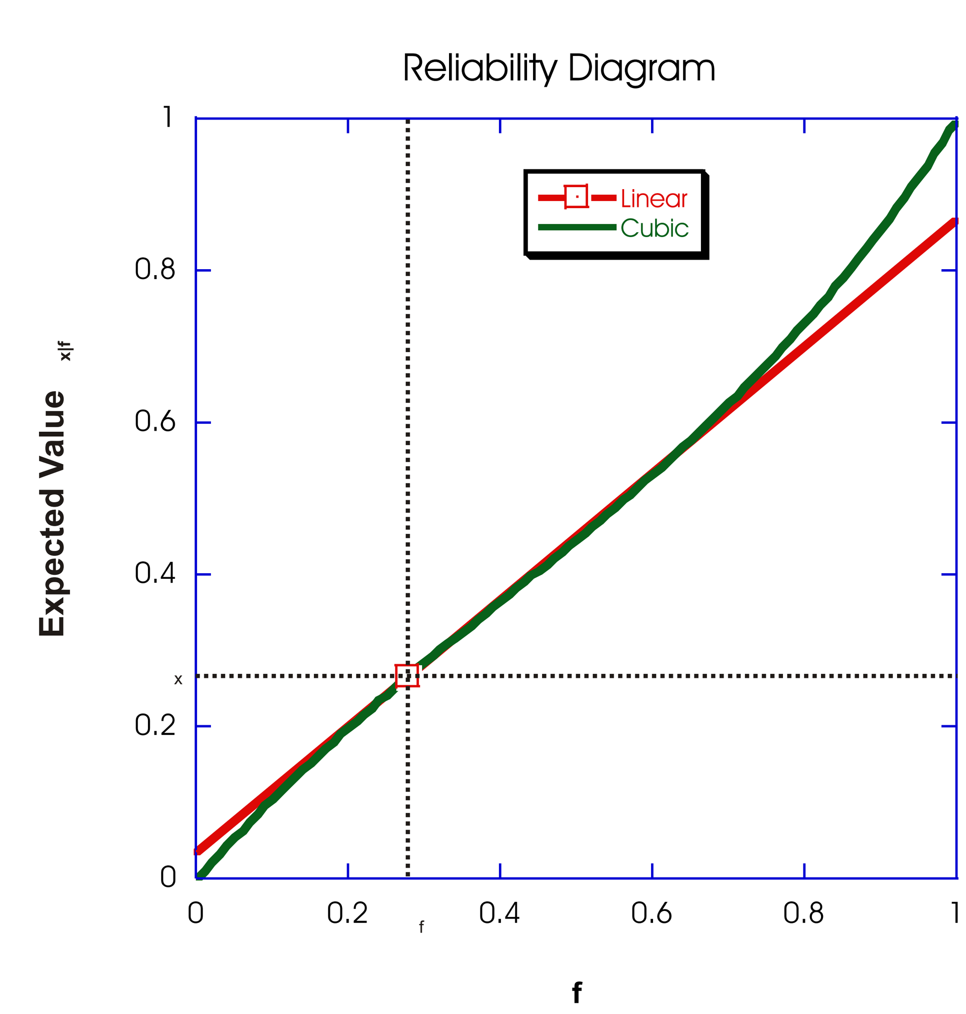
Standard error estimators for forecast quality measures were developed using sampling theory; a goal was to find analytical estimators based only on sample moments (rather than using Monte Carlo simulation):

Exact analytical expressions were derived for the mean error (*ME*) and the mean squared error (*MSE*).

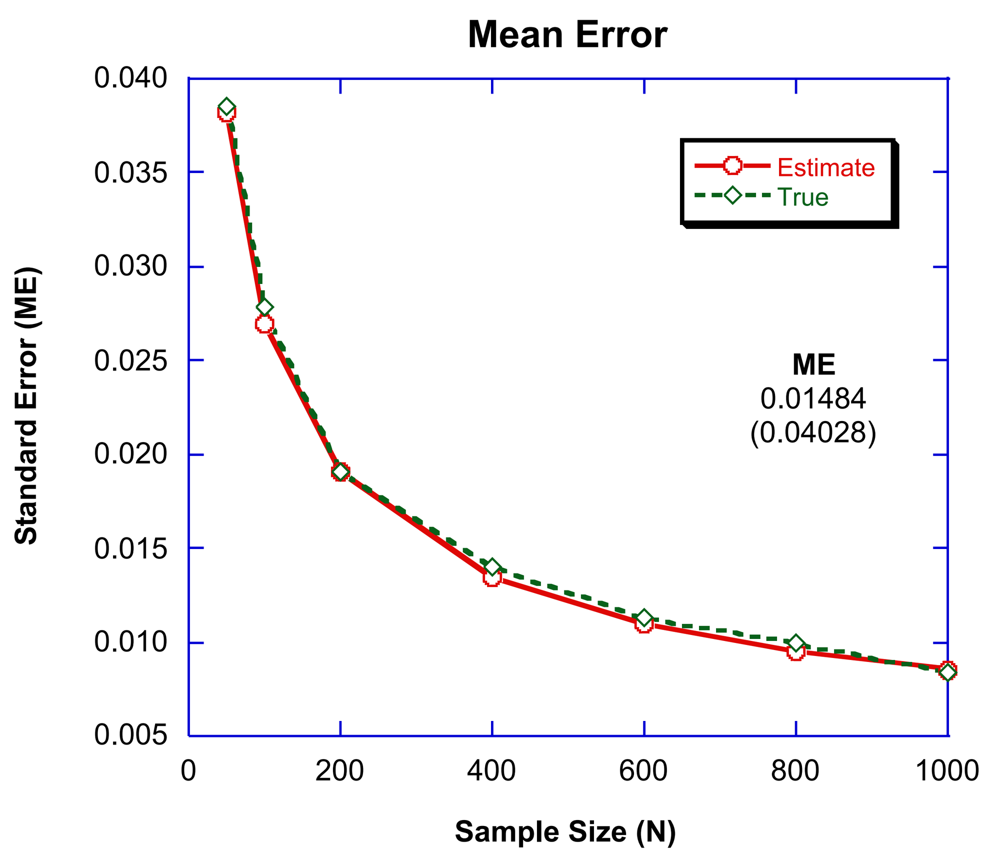
Approximate (2nd-order) analytical expressions were derived for likelihood base-rate measures of discrimination (*DIS*) and type 2 conditional bias (B_2).

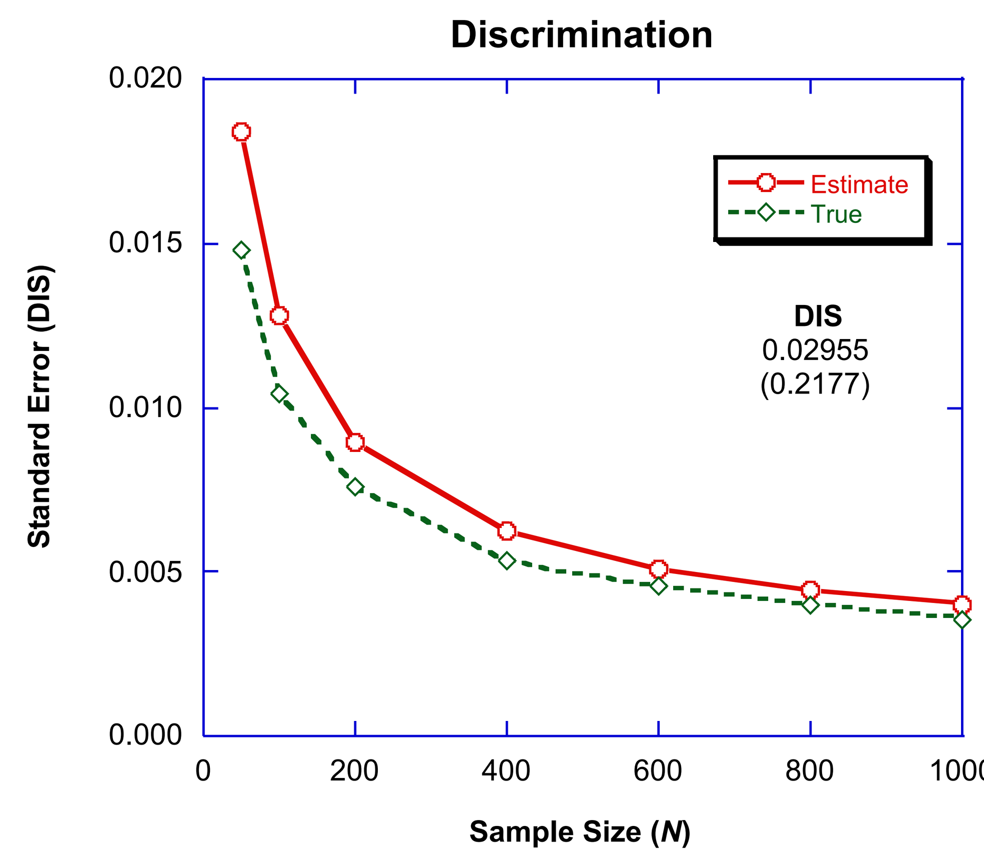
Uncertainties for calibration-refinement measures of resolution (*RES*) and reliability (*REL*) depend on the statistical model for $_{x|f}$. Approximate (2nd-order) analytical expressions were derived for linear and cubic polynomial models (see right figure).

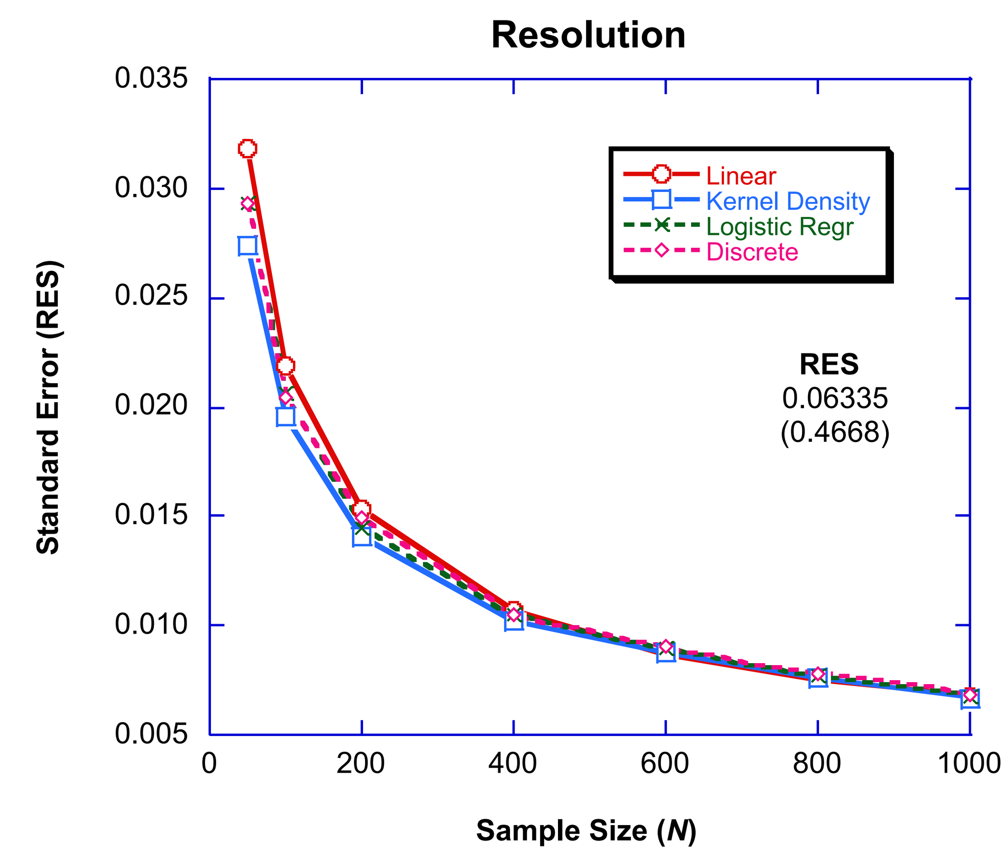
Approximation for the Reliability DIagram



Approximate Uncertainty Estimators





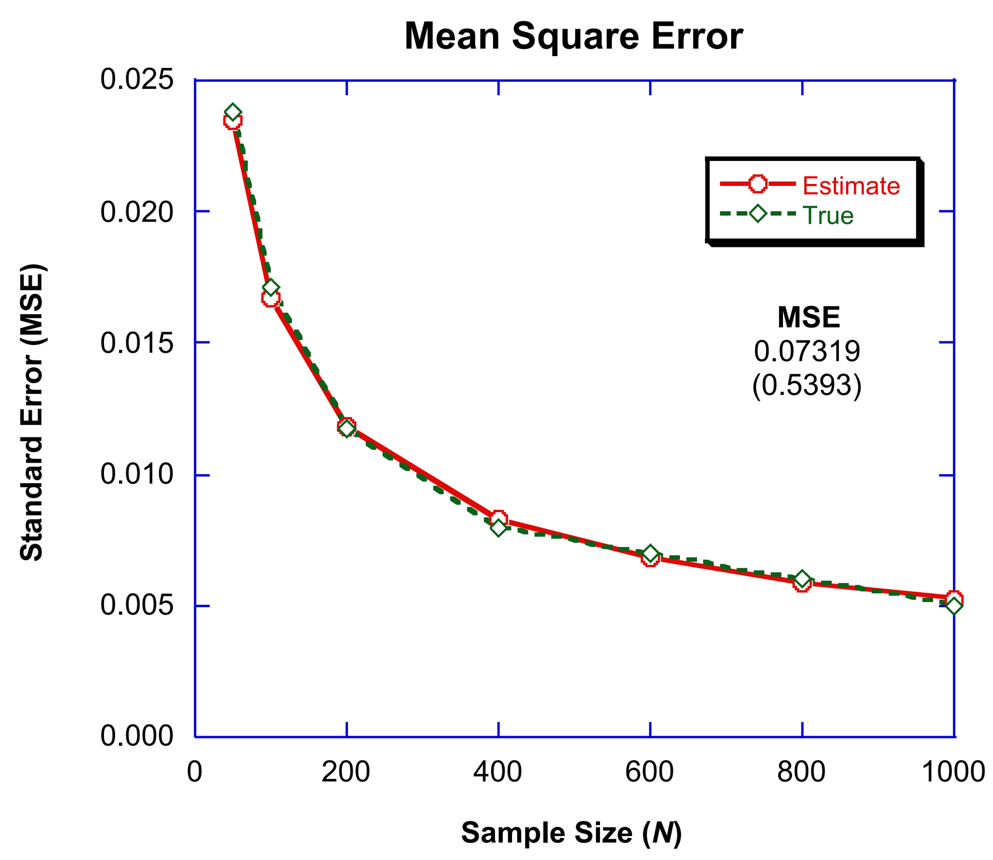


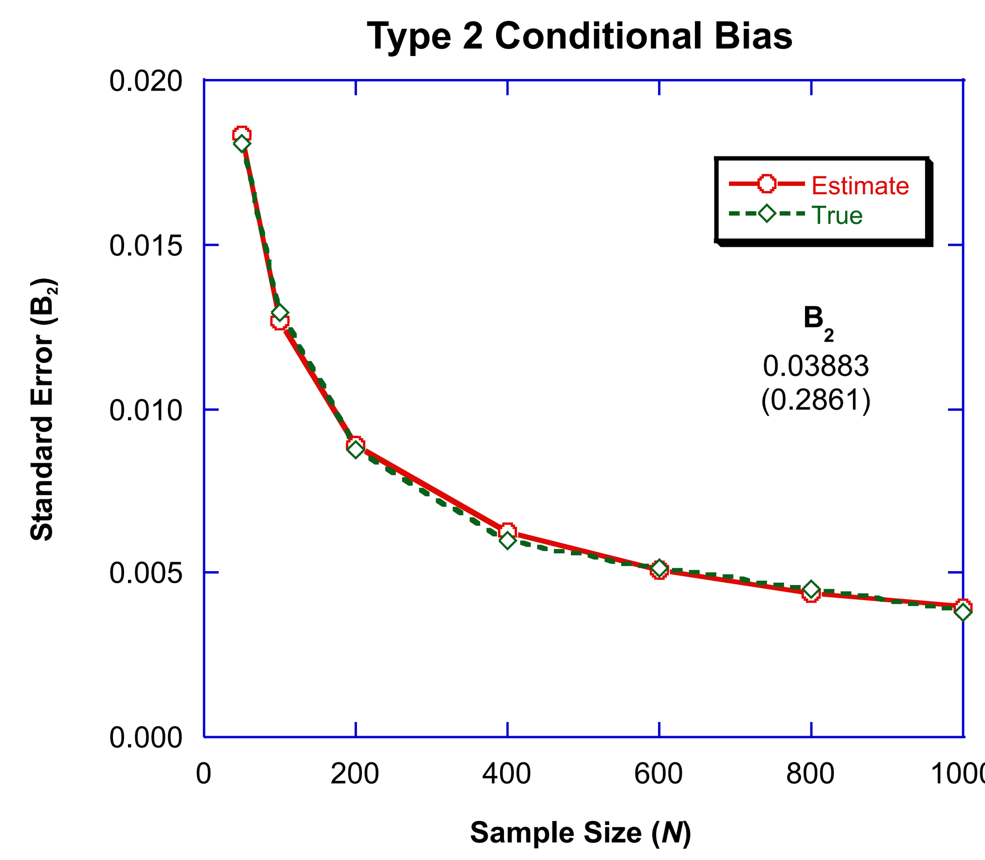
Summary Measures

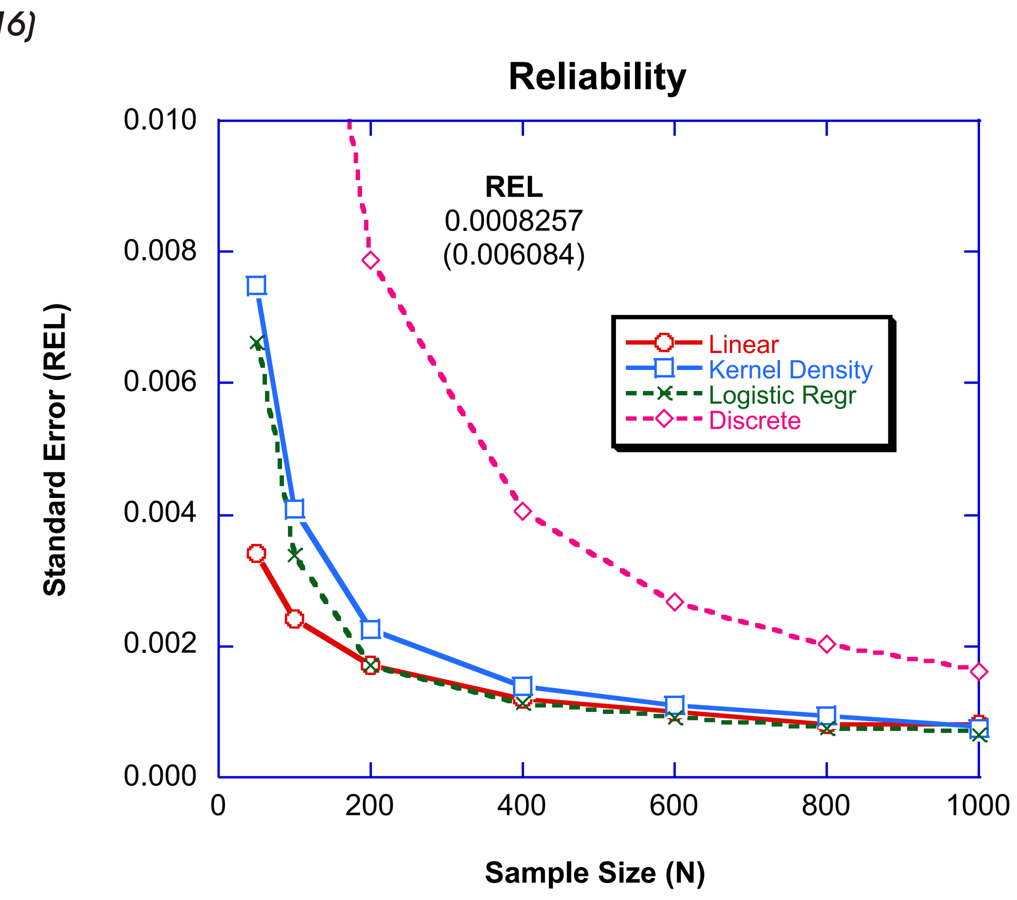
Comparison with Monte Carlo Estimates

Likelihood Base-Rate Measures

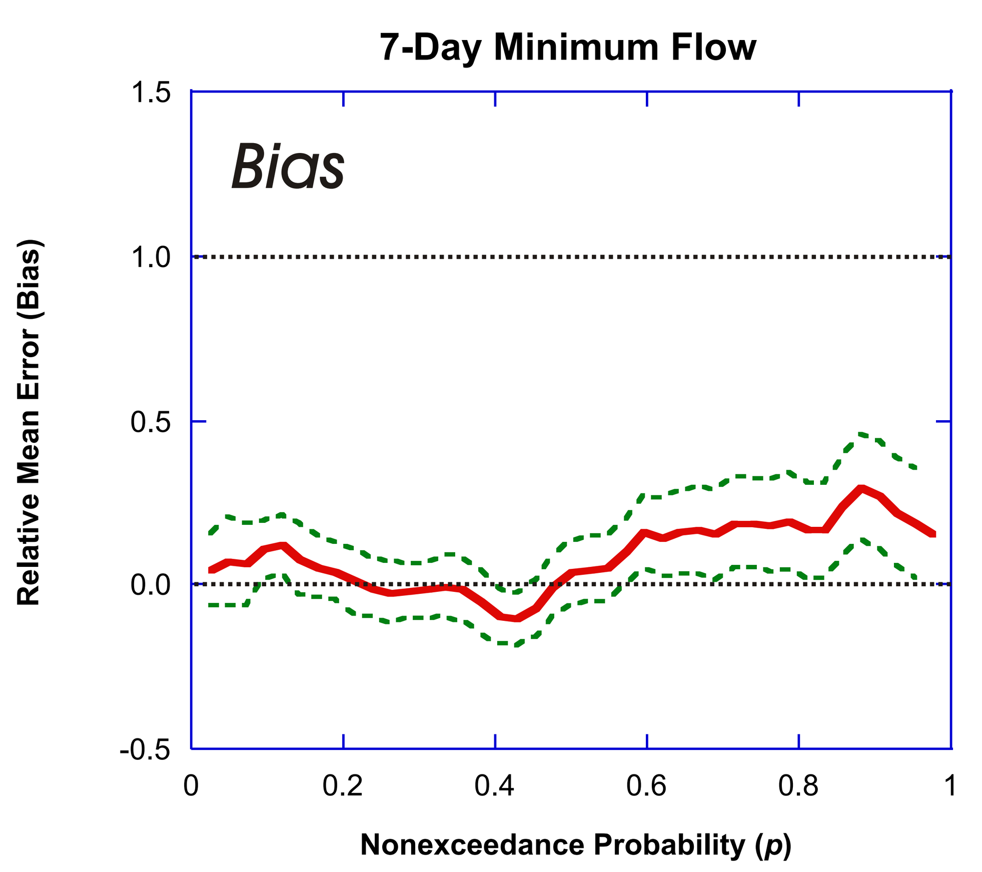
Probability Forecasts (x=0.16)

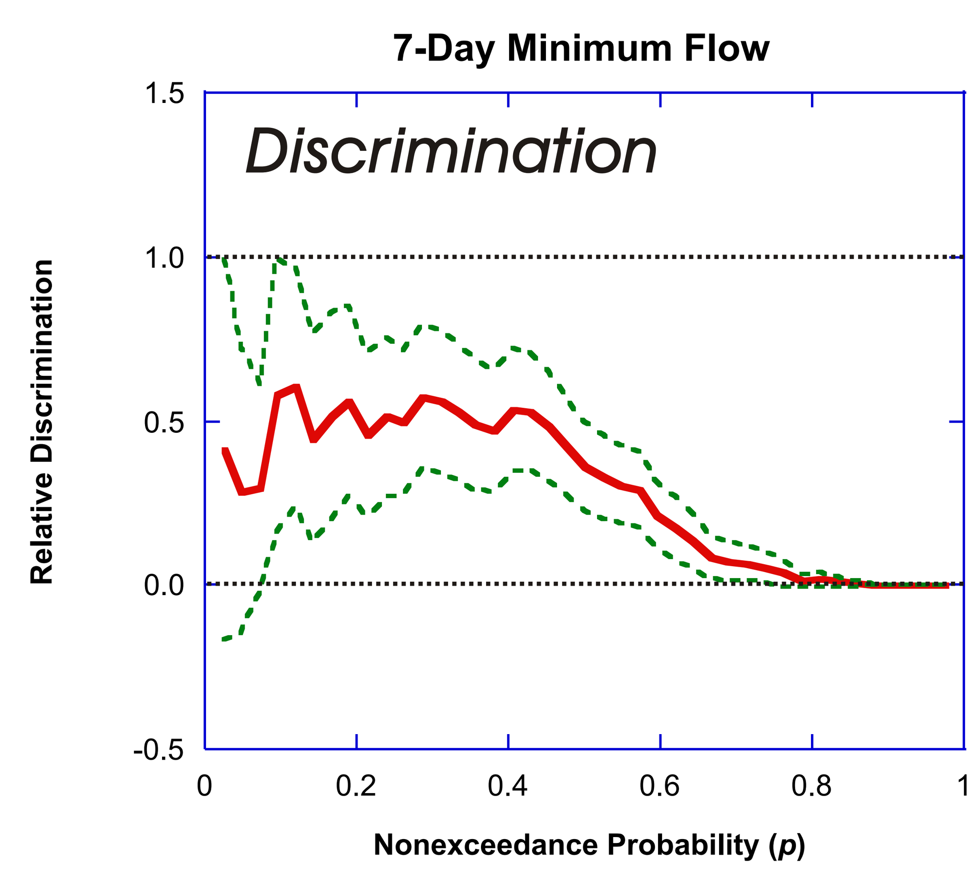


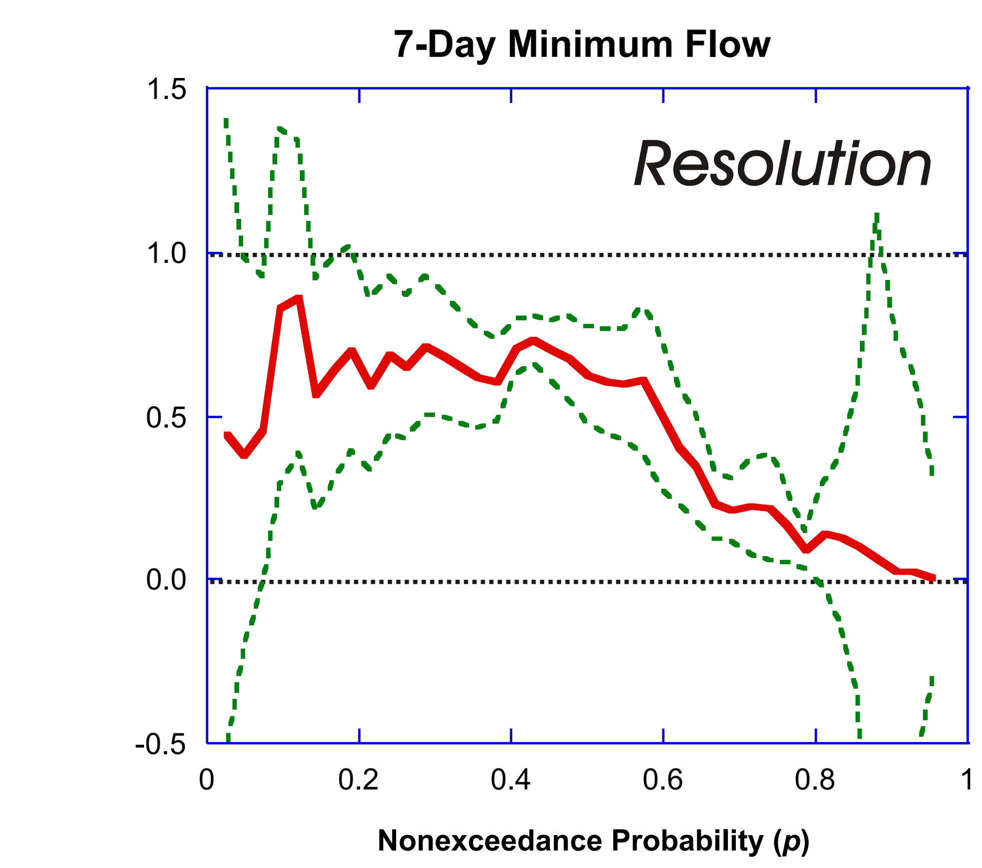




NWS Advanced Hydrologic Prediction System (AHPS)







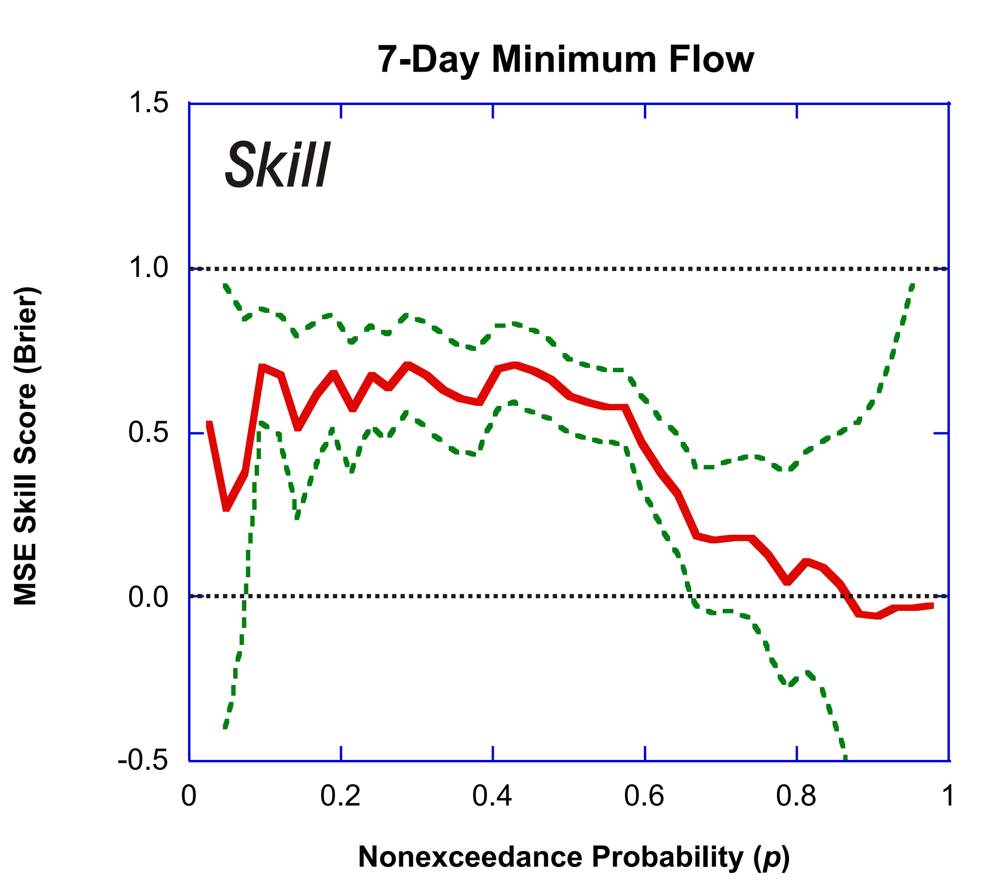
Summary Measures

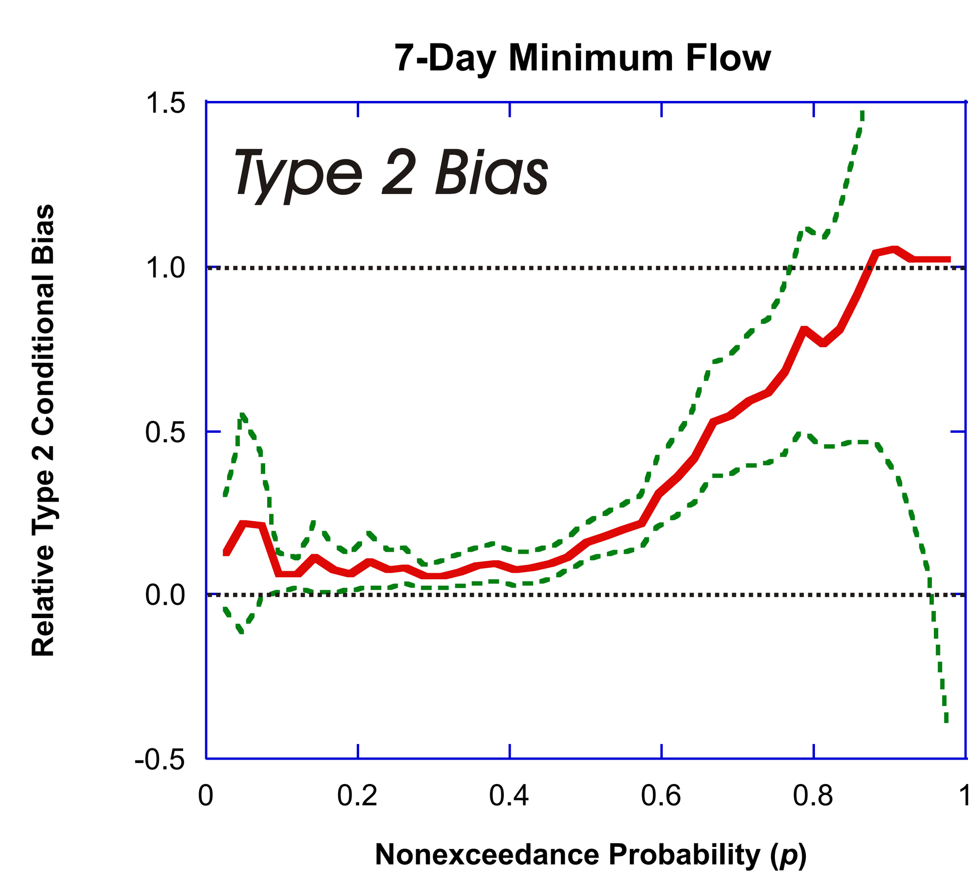
Des Moines River Basin (BAYI4)

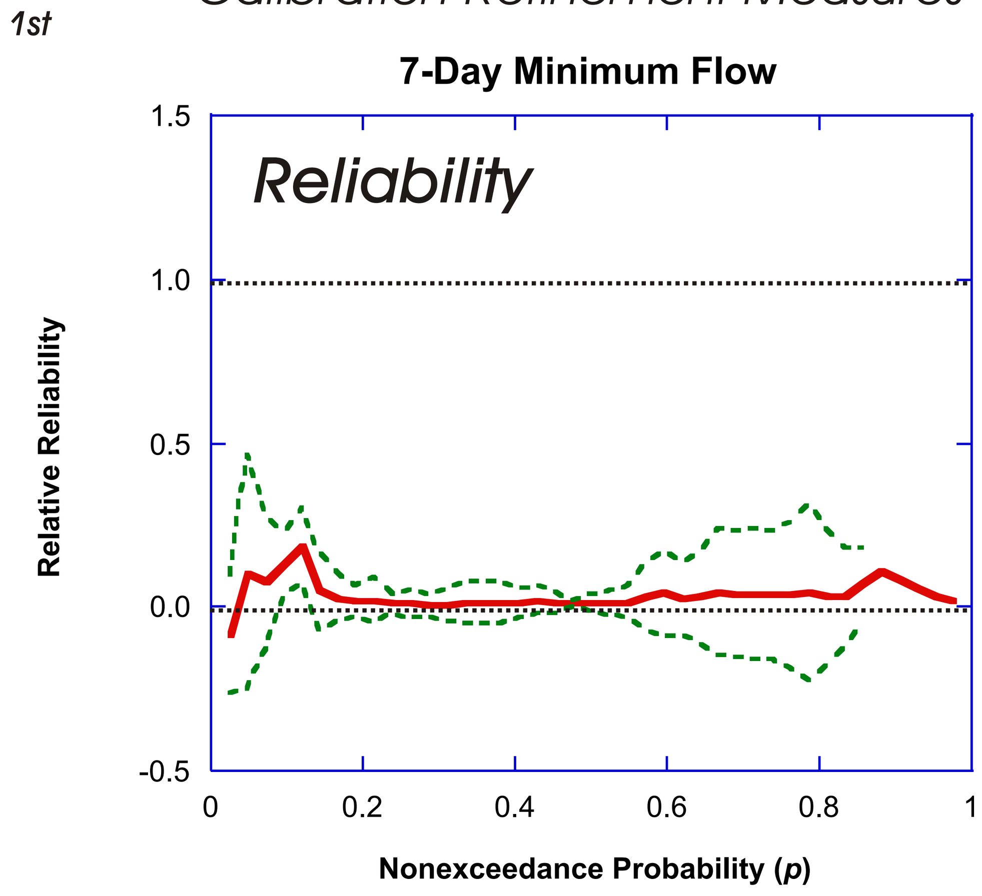
Likelihood Base-Rate Measures

Forecasts Made on April 1st

Calibration-Refinement Measures







Discussion

In comparisons with Monte Carlo estimates for three forecasting situations [Bradley et al. 2003], approximate uncertainty estimators for summary and likelihood base-rate measures are excellent.

Approximate uncertainty estimators for calibration-refinement measures are suitable for uncertainty assessments. A simple linear model for the reliability diagram appears to be sufficient. The linear model estimates are reasonable accurate for forecasts of common and rare events, although the model tends to underestimate standard errors for *REL* at small sample sizes.

Estimators were applied to April 1st forecasts of the 7-day minimum over a 90-day window for the Des Moines River AHPS. Results shows the tendency for higher uncertainties for extreme flow thresholds (low and high exceedance probabilities). Uncertainties also depend on the level (low or high) of the forecast quality measure (e.g., *skill* and *DIS*).

Conclusions

Since forecast quality measures are sample estimates, small sample sizes result in large sampling uncertainty. Assessment of uncertainty in the forecast quality measures needs to be an integral part of ensemble forecast verification.

Exact and approximate estimators of standard errors based on sampling theory are sufficient for summary and likelihood base-rate measures; calibration-refinement measures, which need to use a statistical model of the reliability diagram, may require more complex approaches.

Despite the high uncertainties in the forecast quality estimates, overall attributes of ensemble forecasts (and their significance) are apparent. For example, probabilistic forecasts for AHPS 7-day minimum flows are of good quality for lower flow thresholds (e.g., probability of dry conditions).