HEFS workshop, 03/12/2015
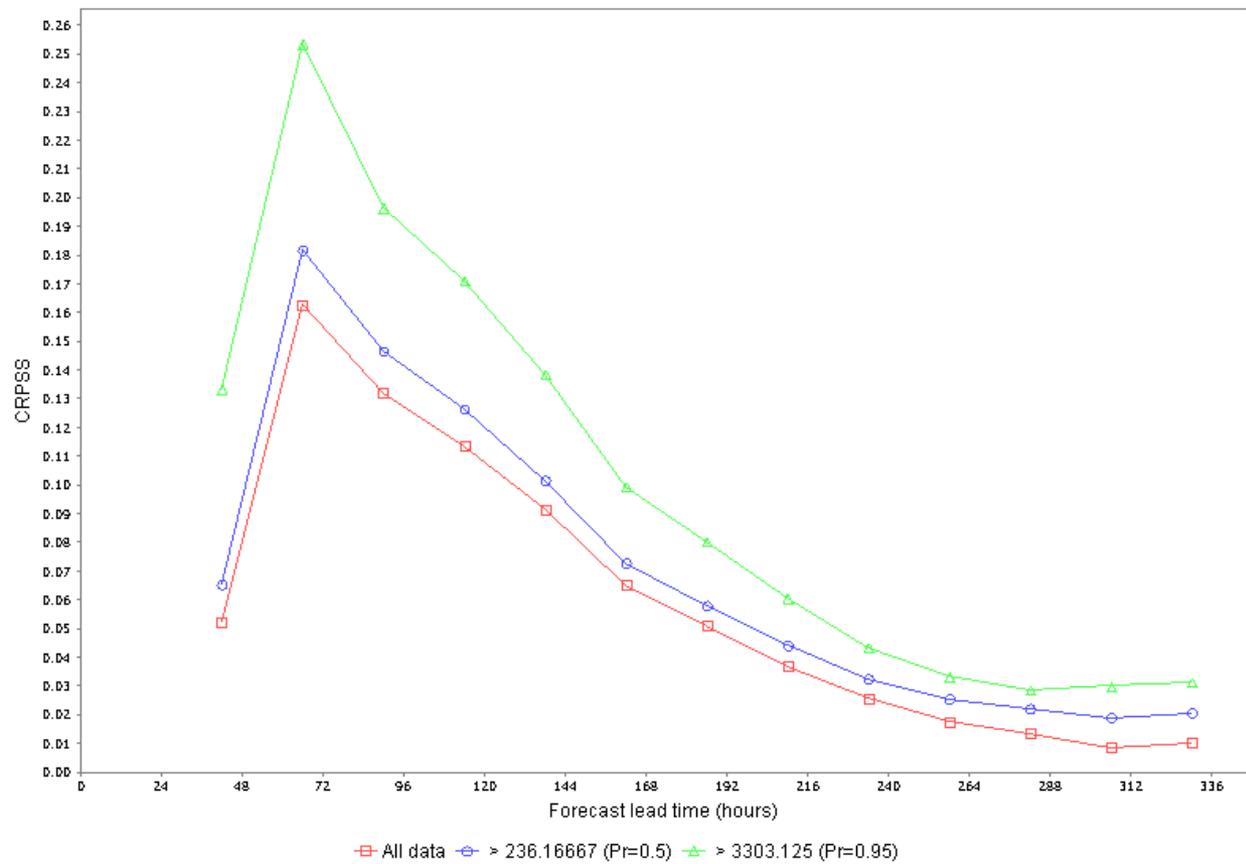
Seminar E: review results from Exercise 4

James Brown

james.brown@hydrosolved.com

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
BLKO2.Streamflow.GEFS (reference forecast: BLKO2.Streamflow.RCLIM)

Q1: Are the MEFP-GEFS forecasts more skilful than climatology?

Exercise 4: Q1 (BLKO2) answer
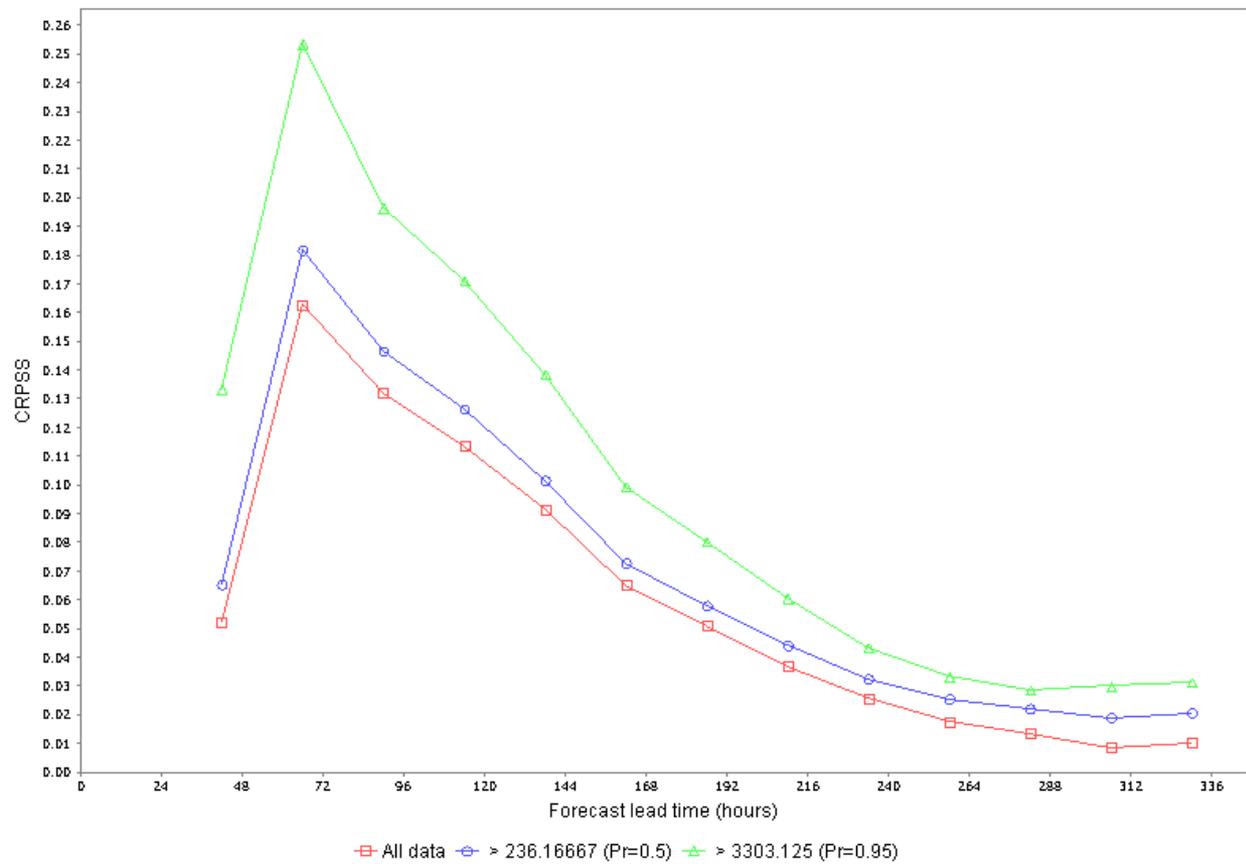
Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
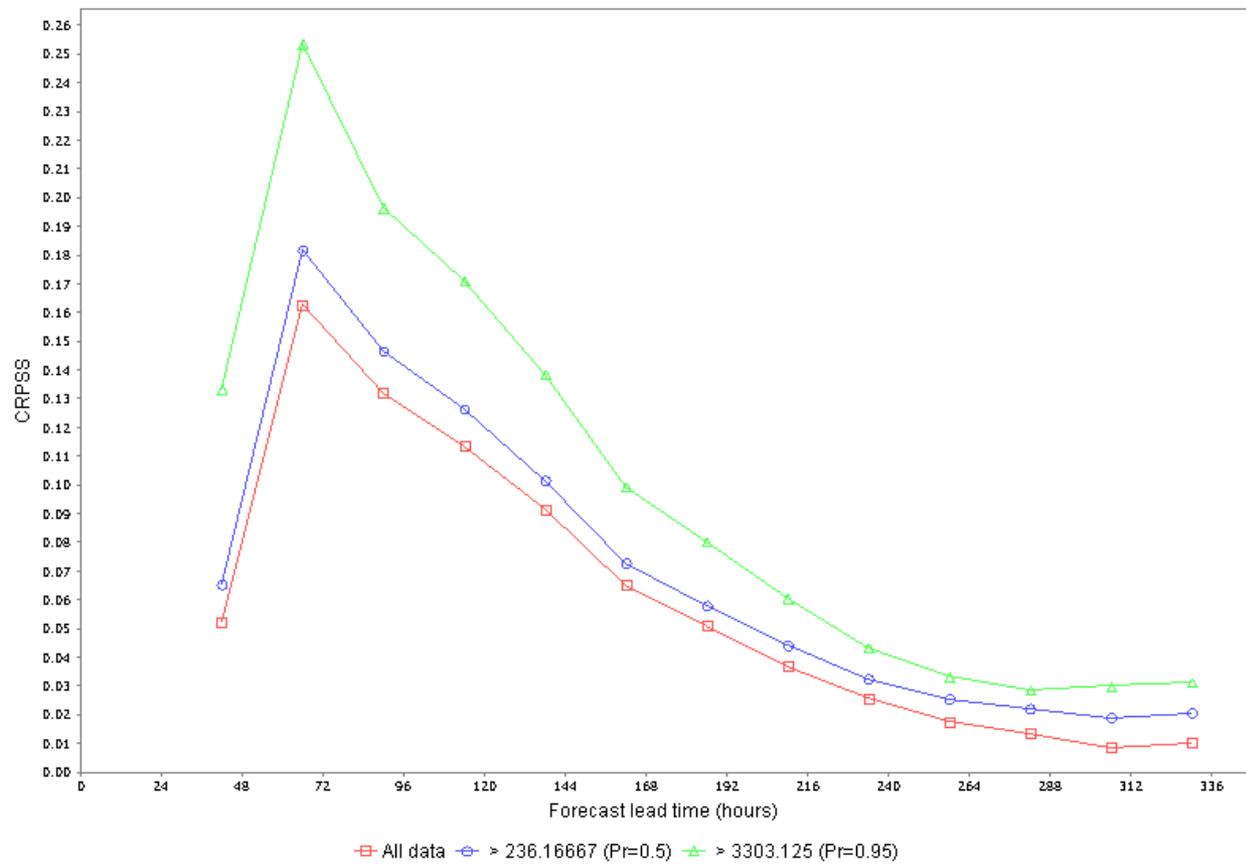BLKO2.Streamflow.GEFS (reference forecast: BLKO2.Streamflow.RCLIM)

Legend: □ All data   ○ > 236.16667 (Pr=0.5)   △ > 3303.125 (Pr=0.95)

Q1: Are the MEFP-GEFS forecasts more skilful than climatology?

A1: In general, substantially so at moderate lead times and higher flows

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
BLKO2.Streamflow.GEFS (reference forecast: BLKO2.Streamflow.RCLIM)

Q2: Why does the skill increase rapidly from 42-66 hours then decline?
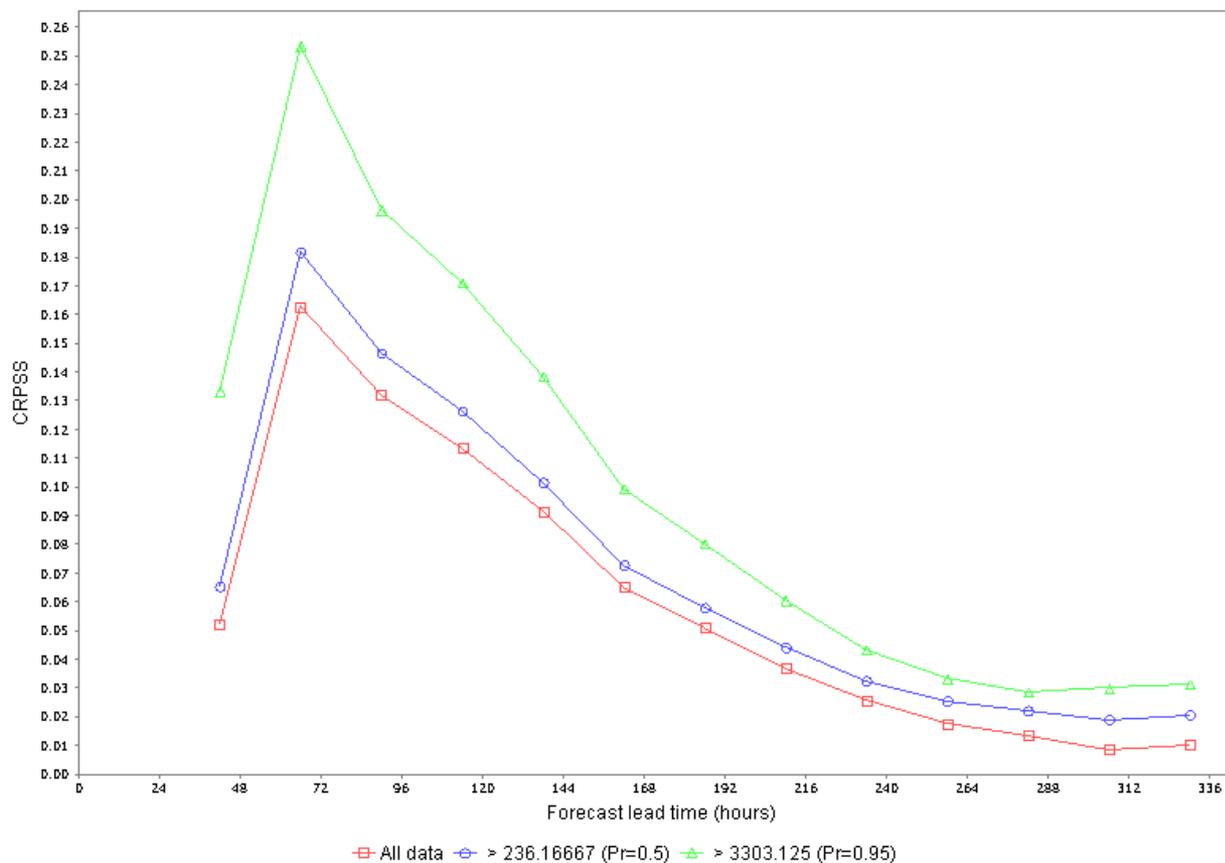
Exercise 4: Q2 (BLKO2) answer

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
BLKO2.Streamflow.GEFS (reference forecast: BLKO2.Streamflow.RCLIM)
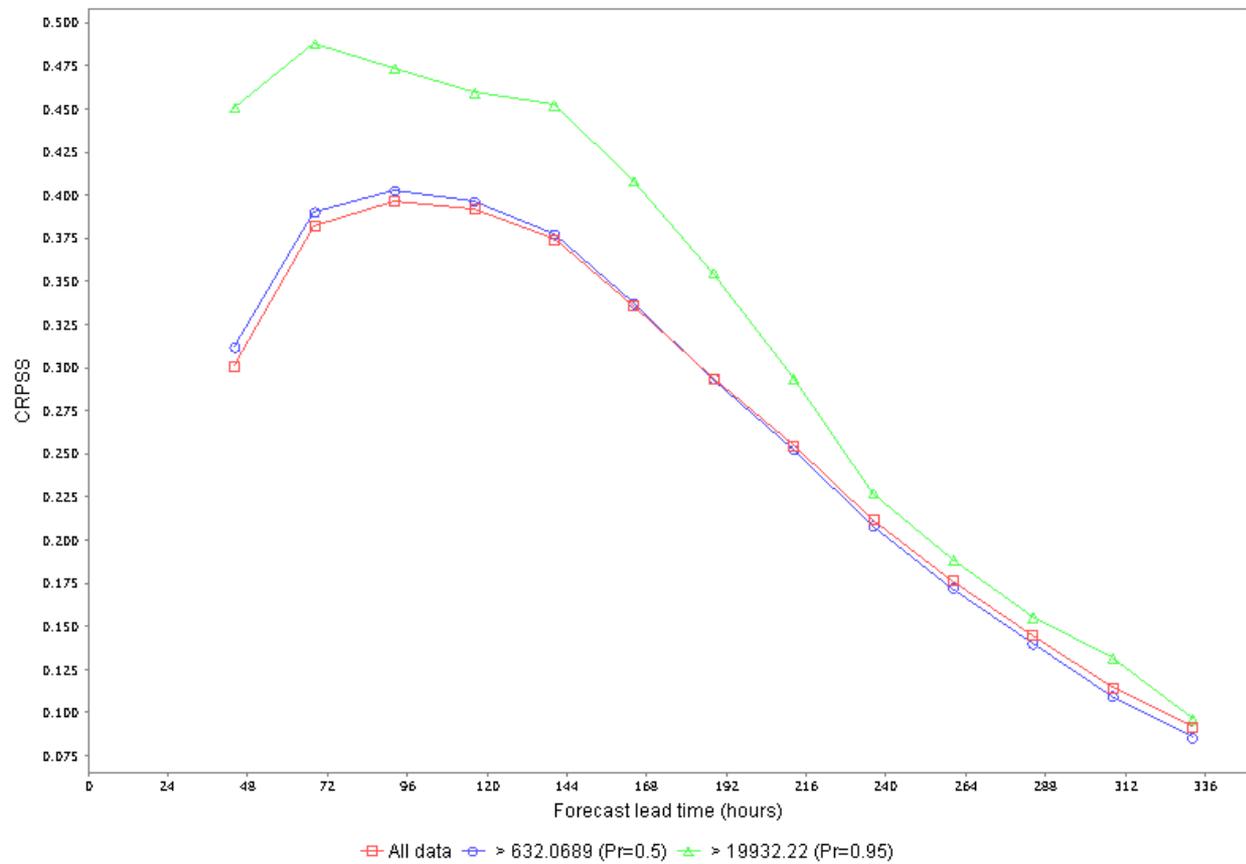
Q2: Why does the skill increase rapidly from 42-66 hours then decline?

A2: Early lead times driven by persistent model states (also in reference)

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
FTSC1.Streamflow.GEFS (reference forecast: FTSC1.Streamflow.RCLIM)

Q3: Are the forecasts more skillful at higher or lower flows? Why?

National Oceanic and Atmospheric Administration's
National Weather Service

SE.6

Office of Hydrologic Development
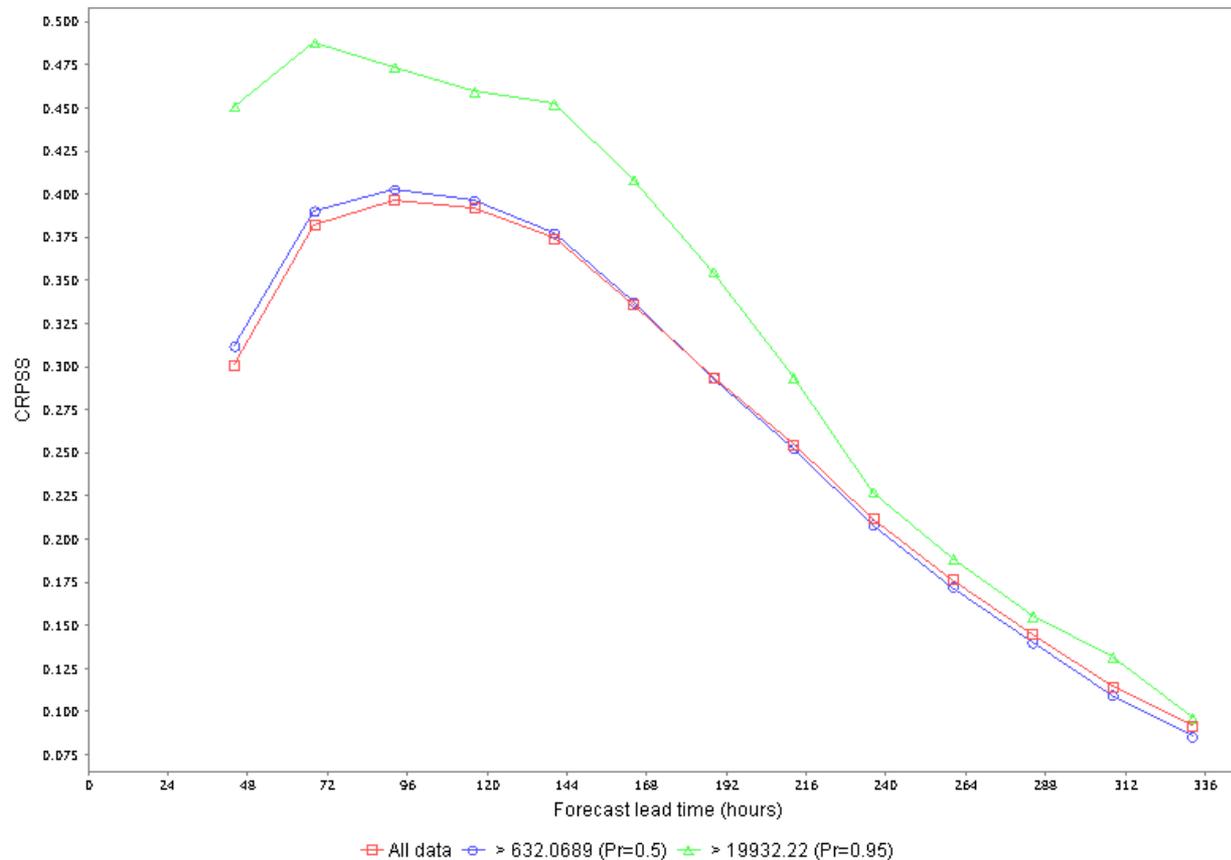Silver Spring, MD

Exercise 4: Q3 (FTSC1) answer

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
FTSC1.Streamflow.GEFS (reference forecast: FTSC1.Streamflow.RCLIM)
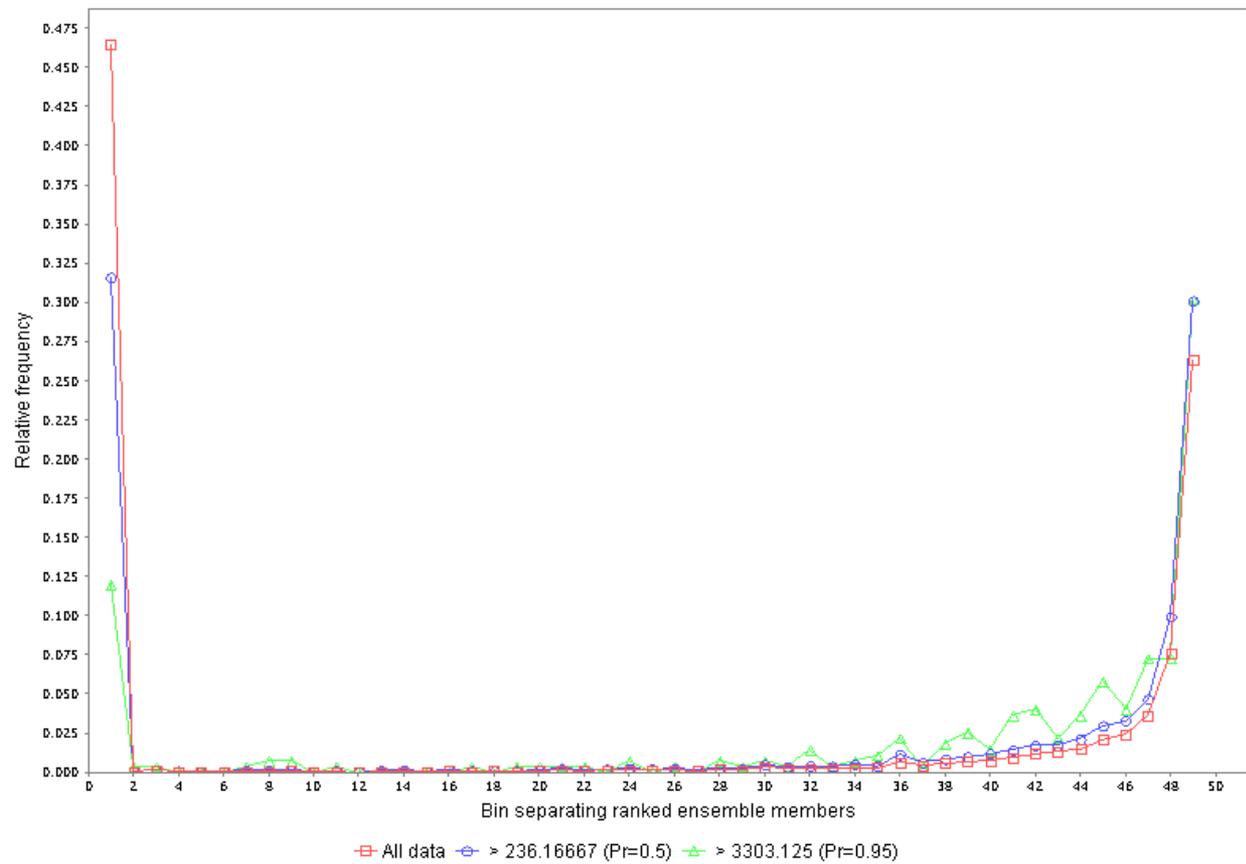
Legend: —□— All data   —○— > 632.0689 (Pr=0.5)   —△— > 19932.22 (Pr=0.95)

Q3: Are the forecasts more skillful at higher or lower flows? Why?

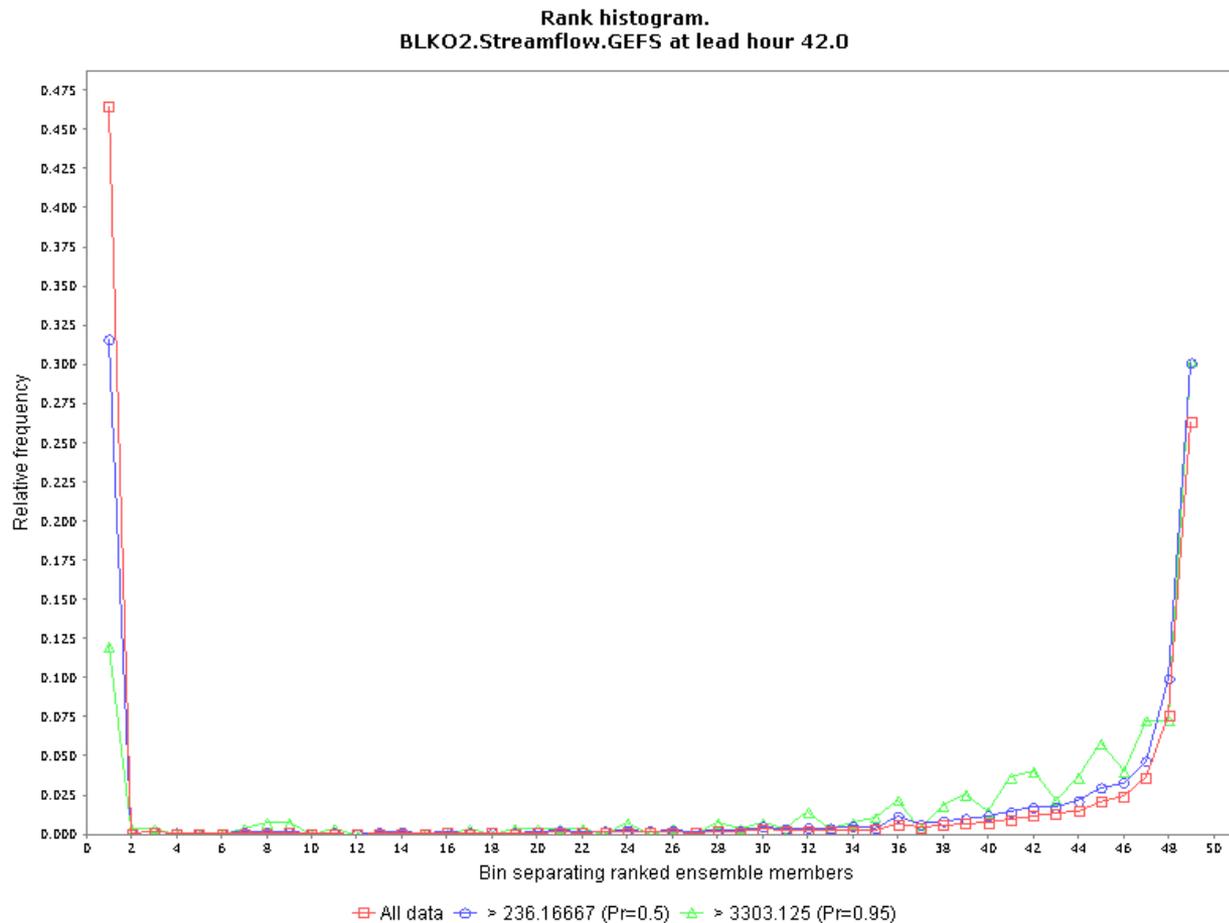**A3**: High. Good forcing & hydro calibration, plus climatology lacks skill.

Rank histogram.
BLKO2.Streamflow.GEFS at lead hour 42.0

Q4: Focusing on "all data" (red), are the forecasts broadly reliable? How?

Rank histogram.
BLKO2.Streamflow.GEFS at lead hour 42.0

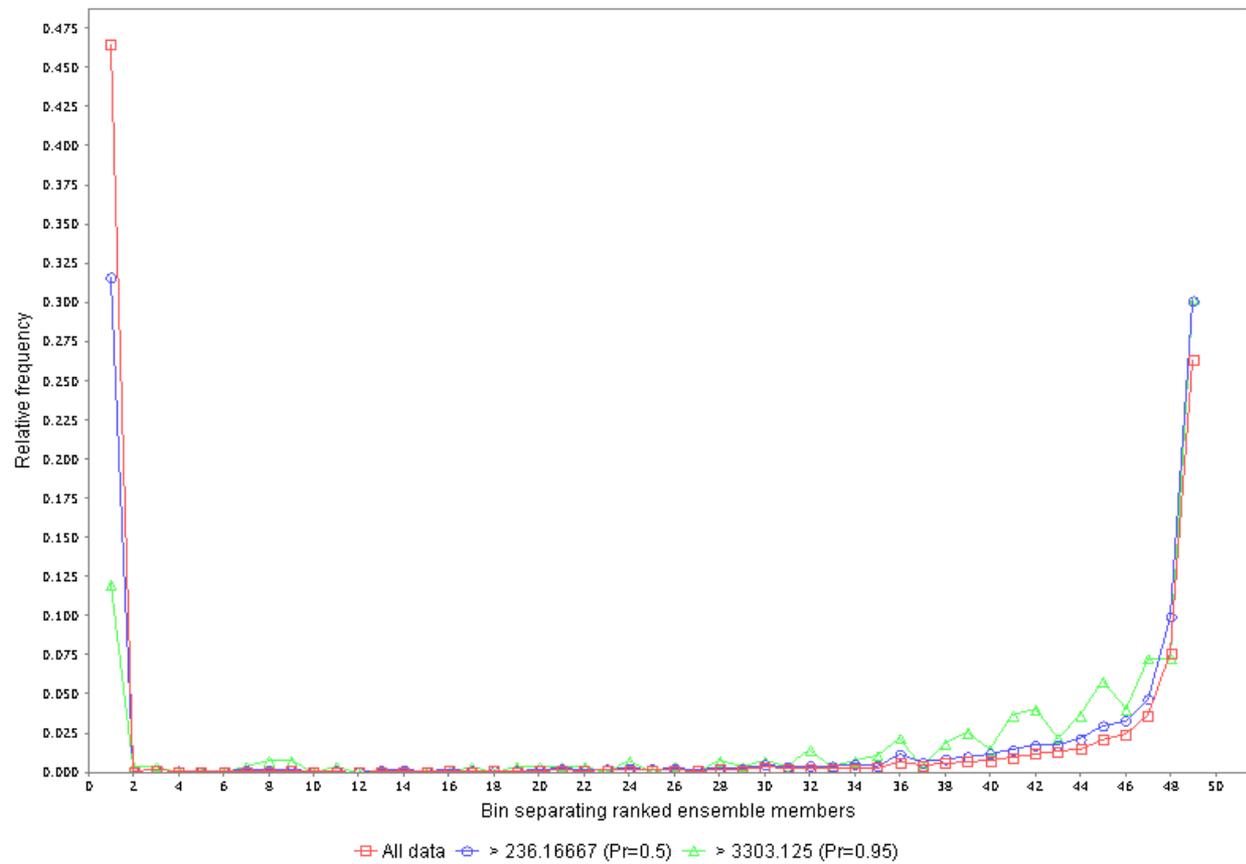Legend: — All data · > 236.16667 (Pr=0.5) · > 3303.125 (Pr=0.95)

Q4: Focusing on "all data" (red), are the forecasts broadly reliable? How?

A4: Not really. Too many observations fall at high and low ends of forecast
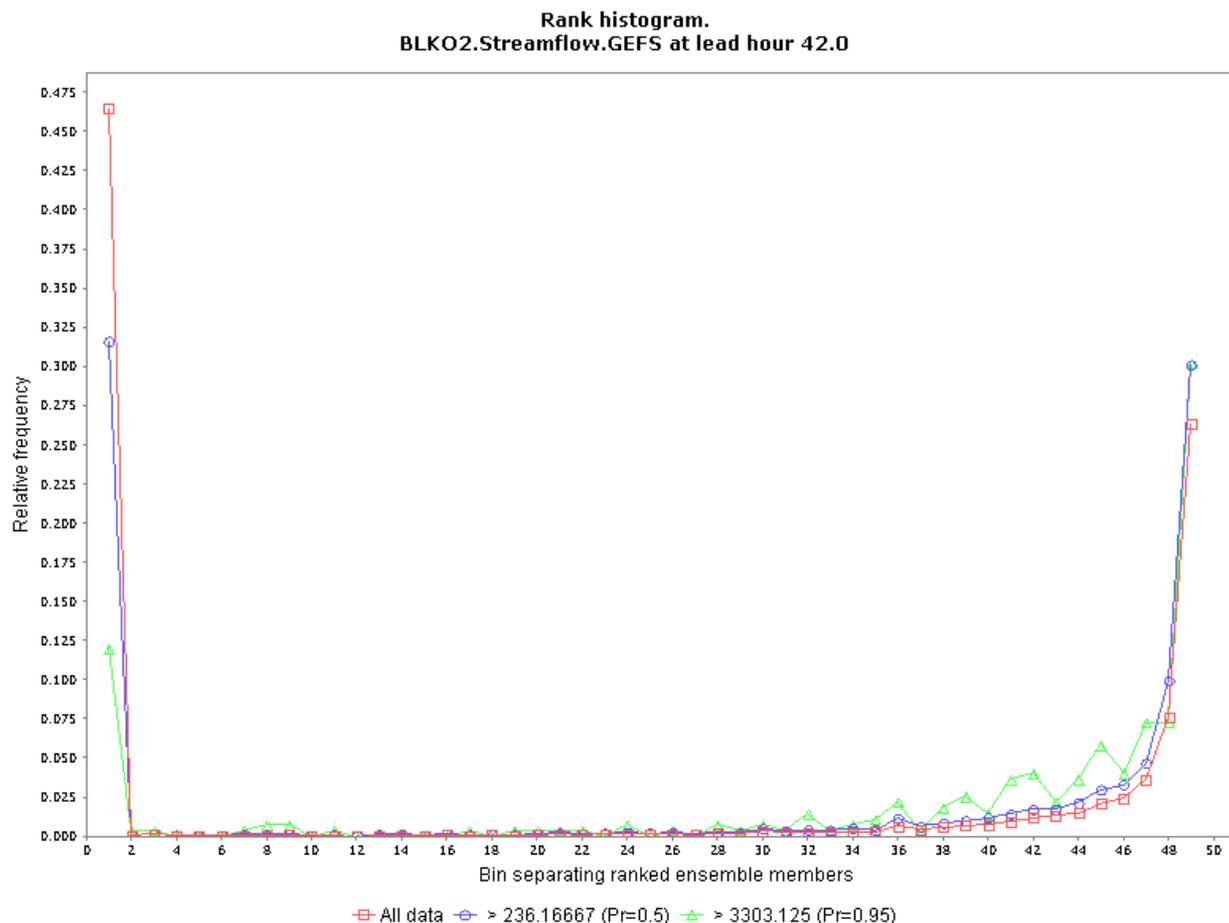
Rank histogram.
BLKO2.Streamflow.GEFS at lead hour 42.0

Q5: What might explain the "U-shape"?

Rank histogram.
BLKO2.Streamflow.GEFS at lead hour 42.0
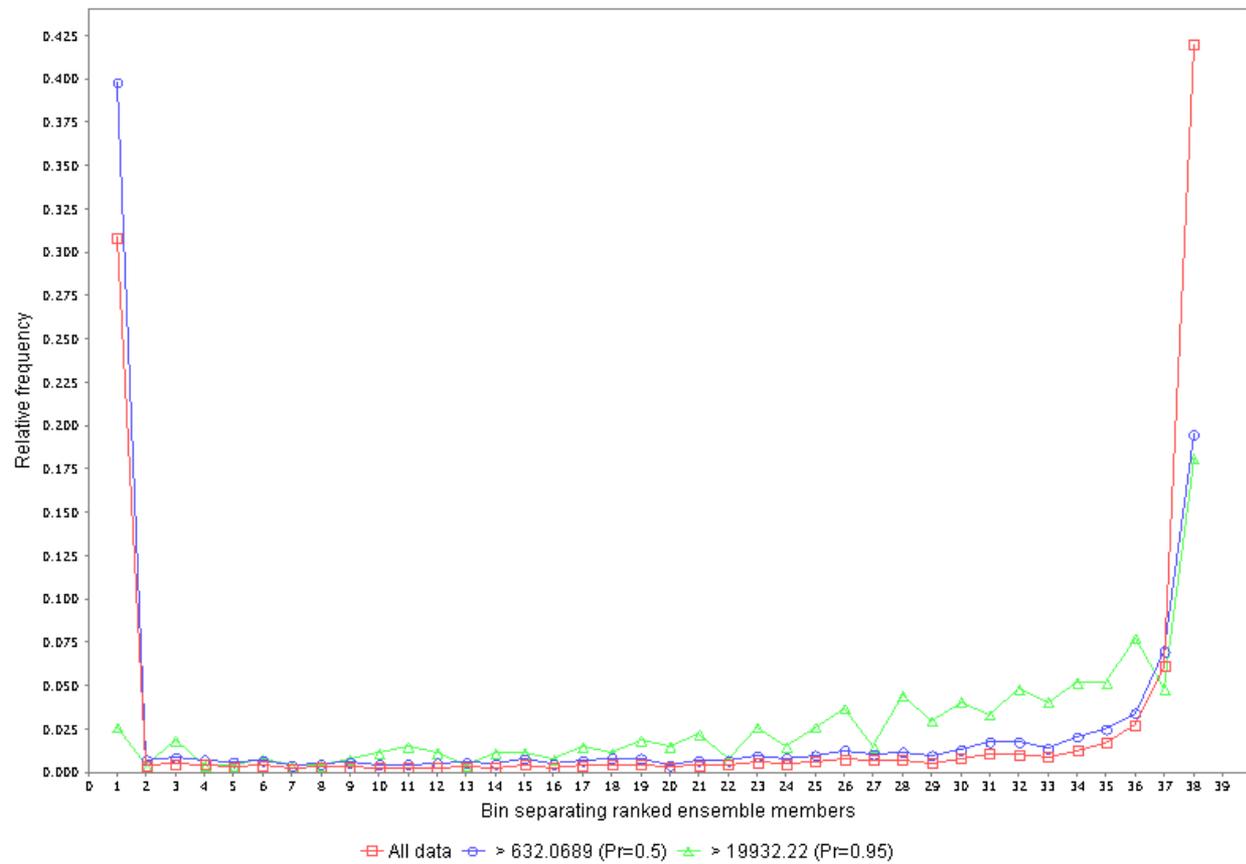
Q5: What might explain the "U-shape"?

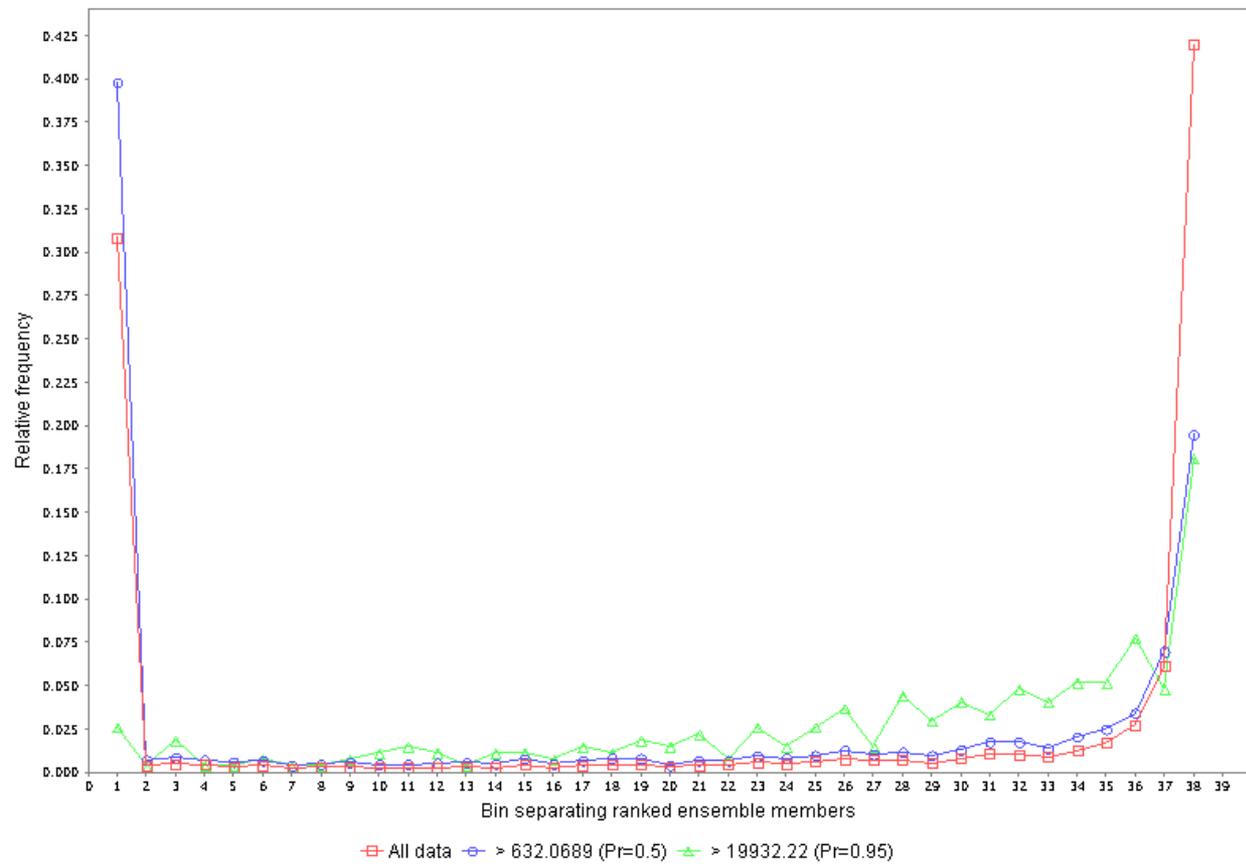**A5:** This is classic behavior indicative of lack of spread (over-confidence)

Rank histogram.
FTSC1.Streamflow.GEFS at lead hour 44.0

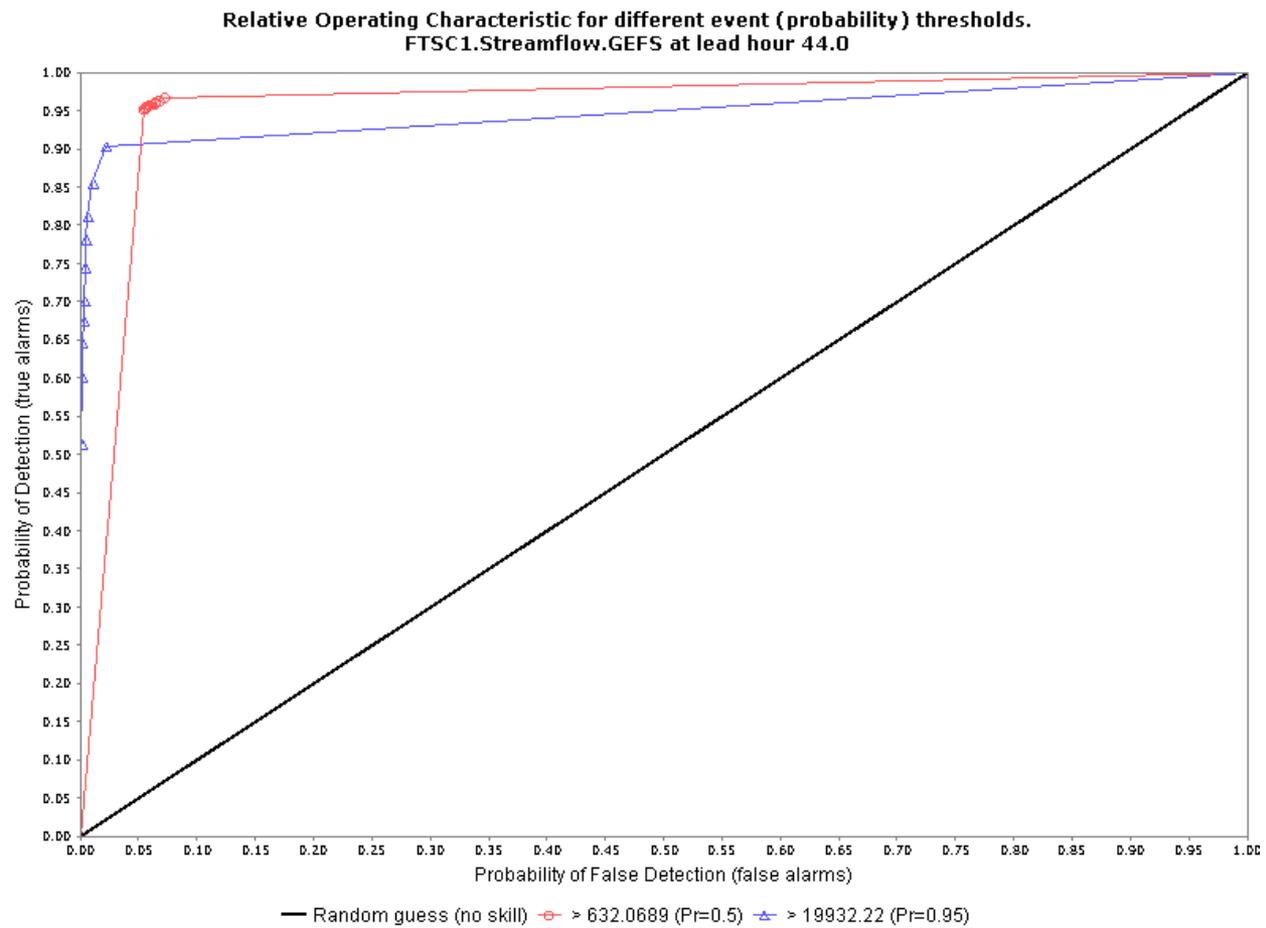**Q6:** What might explain the tendency for higher frequencies in upper tail?

Exercise 4: Q6 (FTSC1) answer

Rank histogram.
FTSC1.Streamflow.GEFS at lead hour 44.0

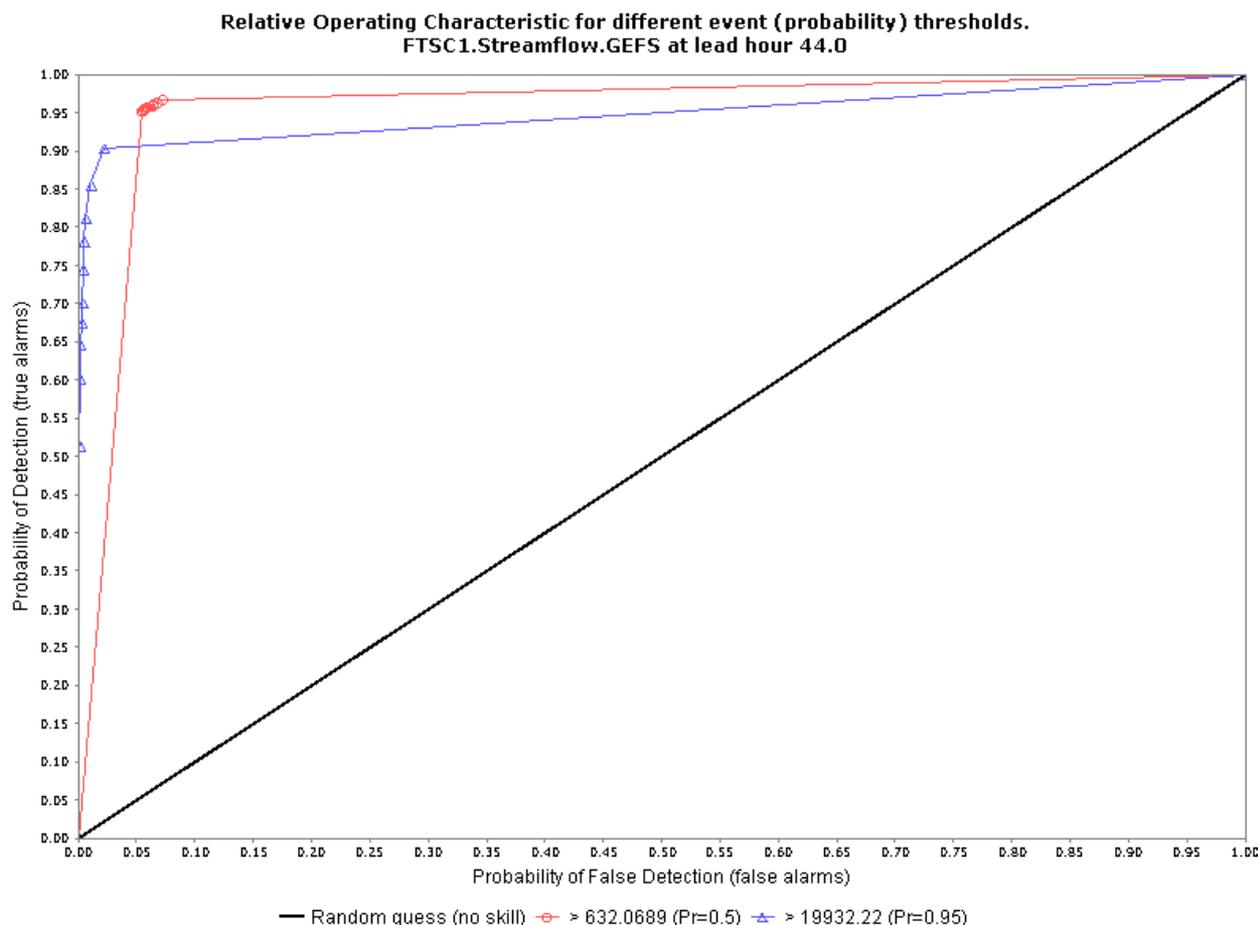Legend: All data — ≥ 632.0689 (Pr=0.5) — ≥ 19932.22 (Pr=0.95)

Q6: What might explain the tendency for higher frequencies in upper tail?

A6: Two things. Lack of spread (both tails) and a conditional bias (too low)

National Oceanic and Atmospheric Administration's
National Weather Service

SE.13

Office of Hydrologic Development
Silver Spring, MD

Exercise 4: Q7 (FTSC1)

Relative Operating Characteristic for different event (probability) thresholds.
FTSC1.Streamflow.GEFS at lead hour 44.0

Legend: Random guess (no skill) — ≥ 632.0689 (Pr=0.5) — ≥ 19932.22 (Pr=0.95)

Q7: Can the forecasts discriminate occurrences from non-occurrences?

Relative Operating Characteristic for different event (probability) thresholds.
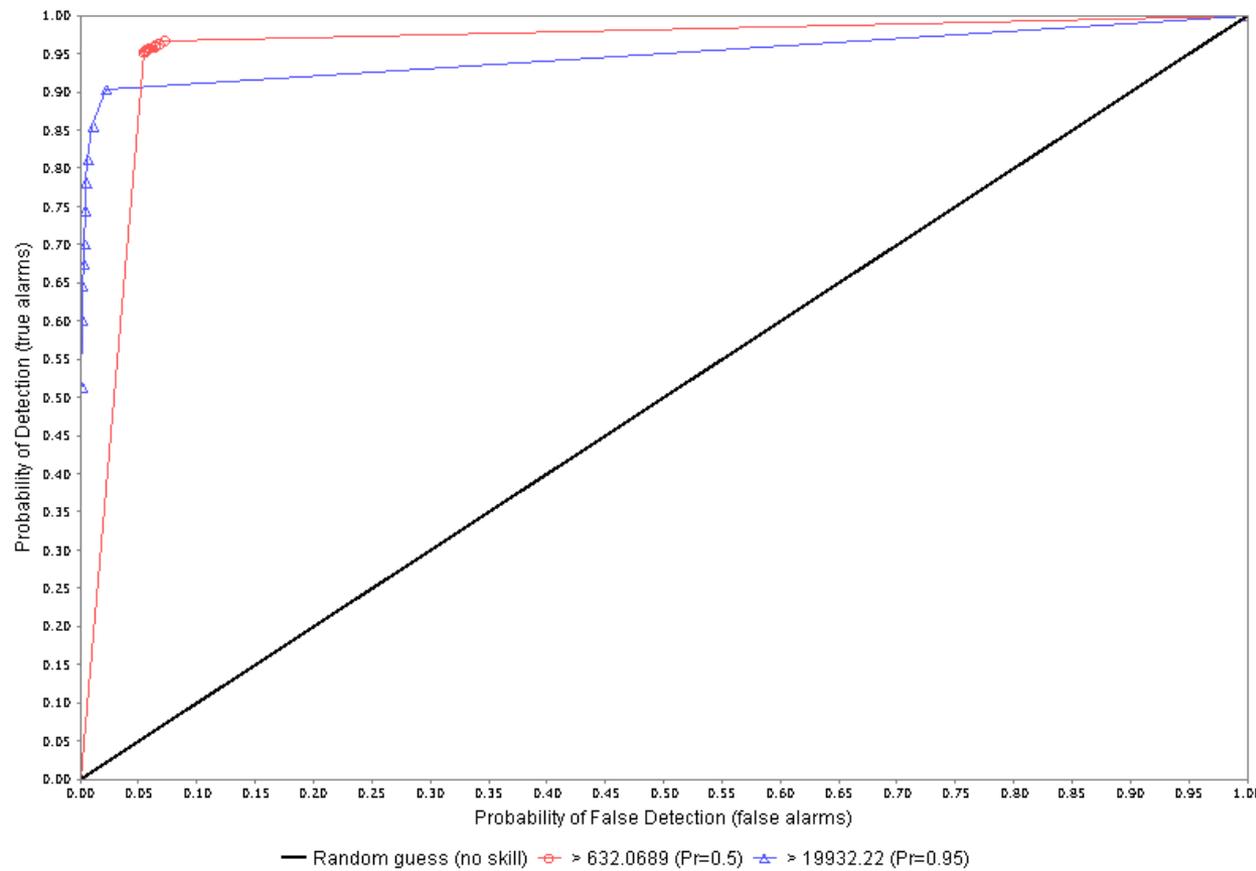FTSC1.Streamflow.GEFS at lead hour 44.0

Q7: Can the forecasts discriminate occurrences from non-occurrences?

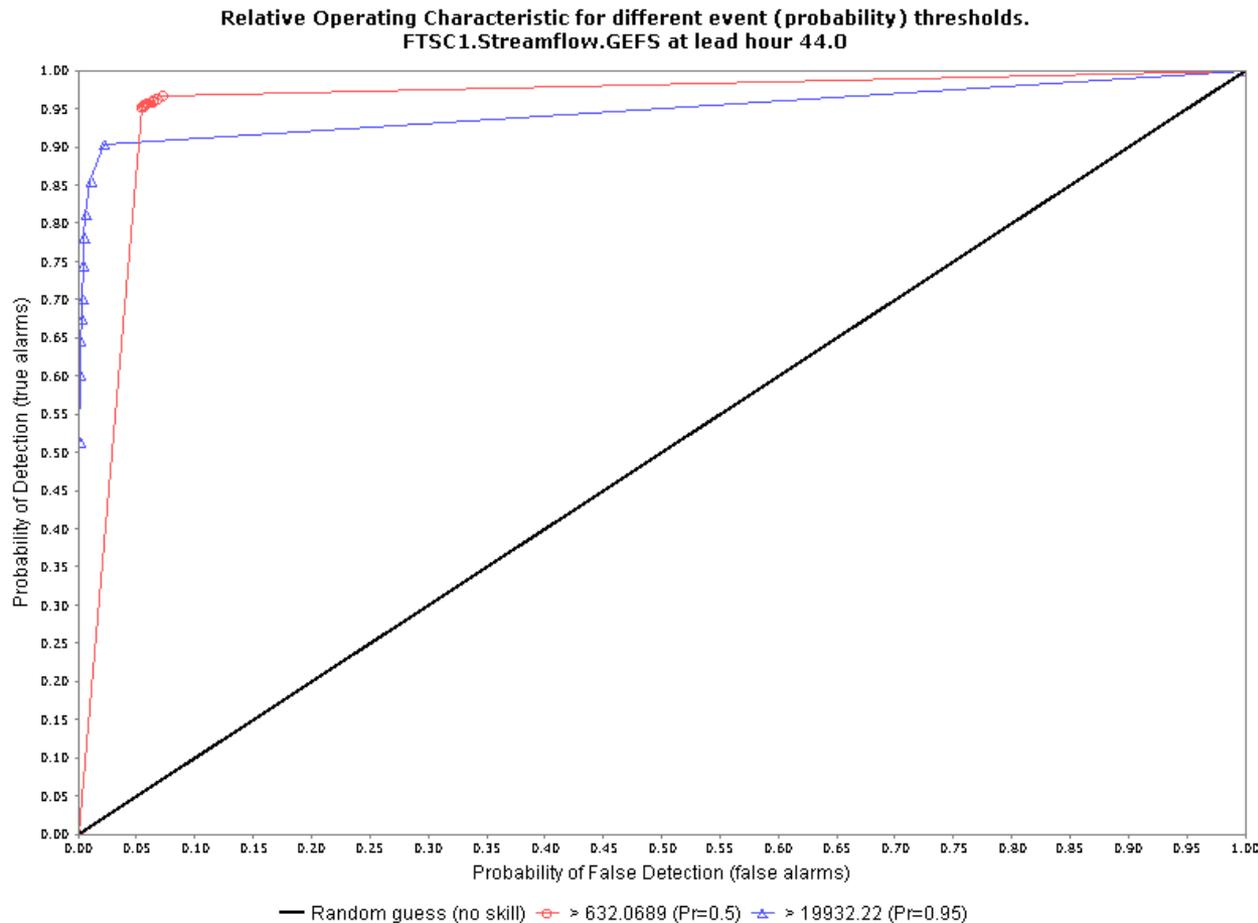**A7:** Yes, much better than climatology. PoD much higher than PoFD.

Relative Operating Characteristic for different event (probability) thresholds.
FTSC1.Streamflow.GEFS at lead hour 44.0

Q8: Decision maker: accept 5% PoFD for flows > 19,932 CFS. What PoD?

Relative Operating Characteristic for different event (probability) thresholds.
FTSC1.Streamflow.GEFS at lead hour 44.0

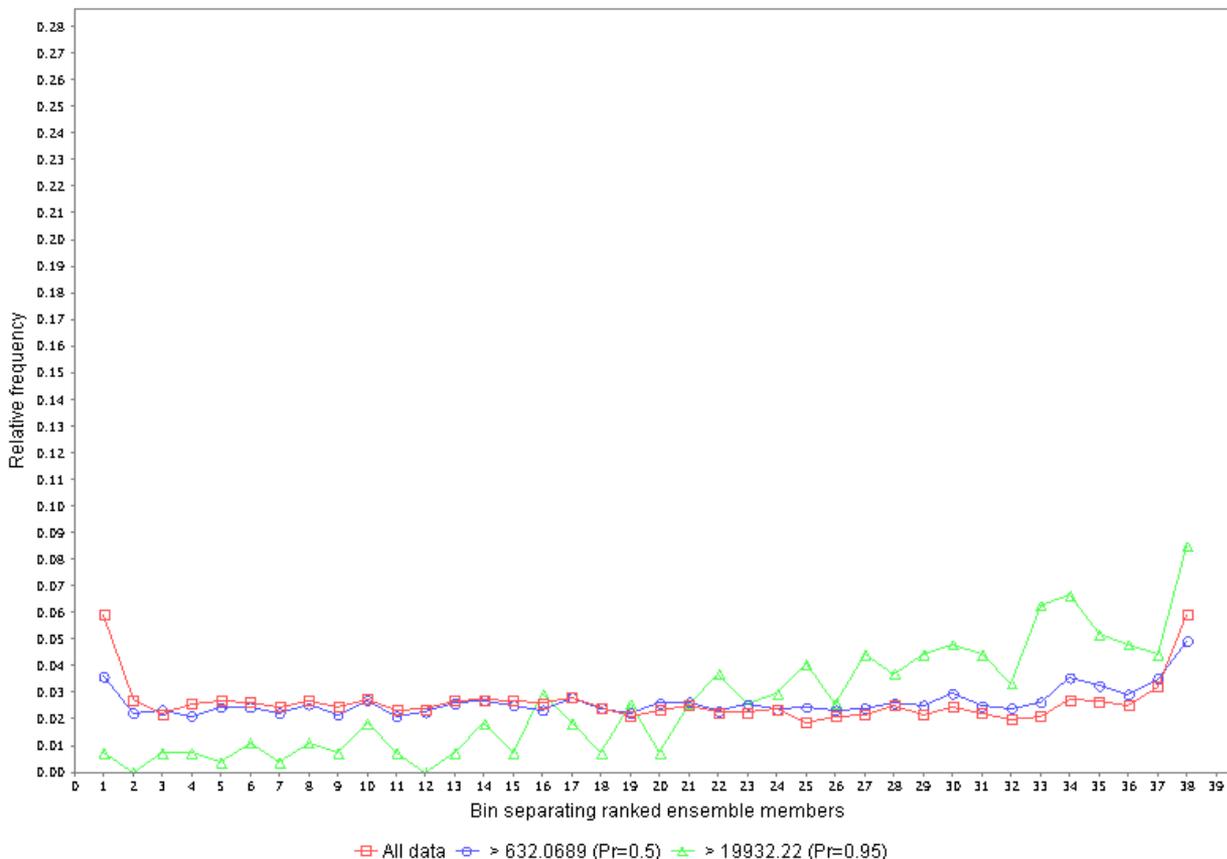—— Random guess (no skill) —○— > 632.0689 (Pr=0.5) —△— > 19932.22 (Pr=0.95)

Q8: Decision maker: accept 5% PoFD for flows > 19,932 CFS. What PoD?

A8: About 90%, i.e. when event occurred, warnings were correct ~90%
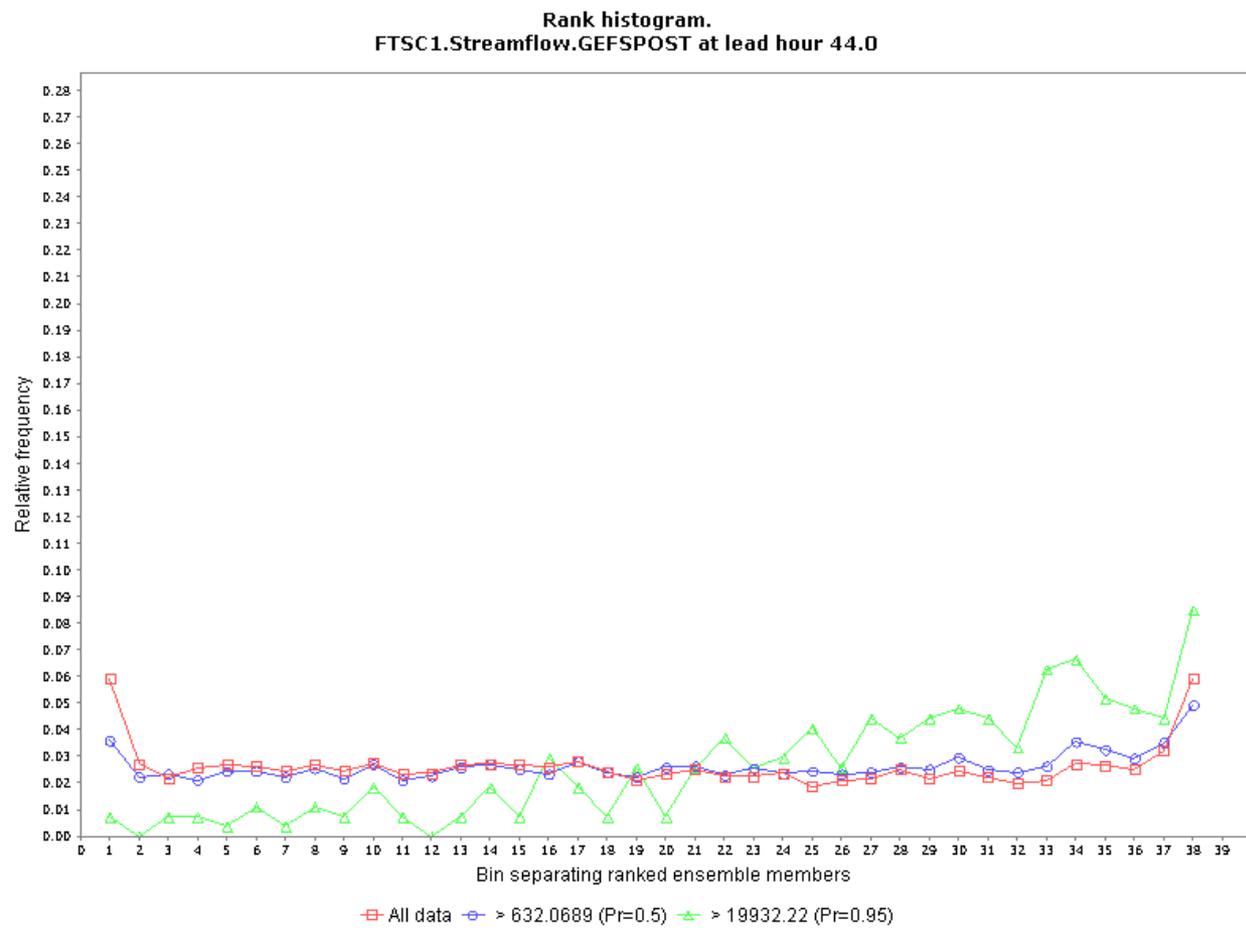
Additional questions

Rank histogram.
FTSC1.Streamflow.GEFSPOST at lead hour 44.0

Q9: Compare with histogram before EnsPost (SE.13). More reliable after?
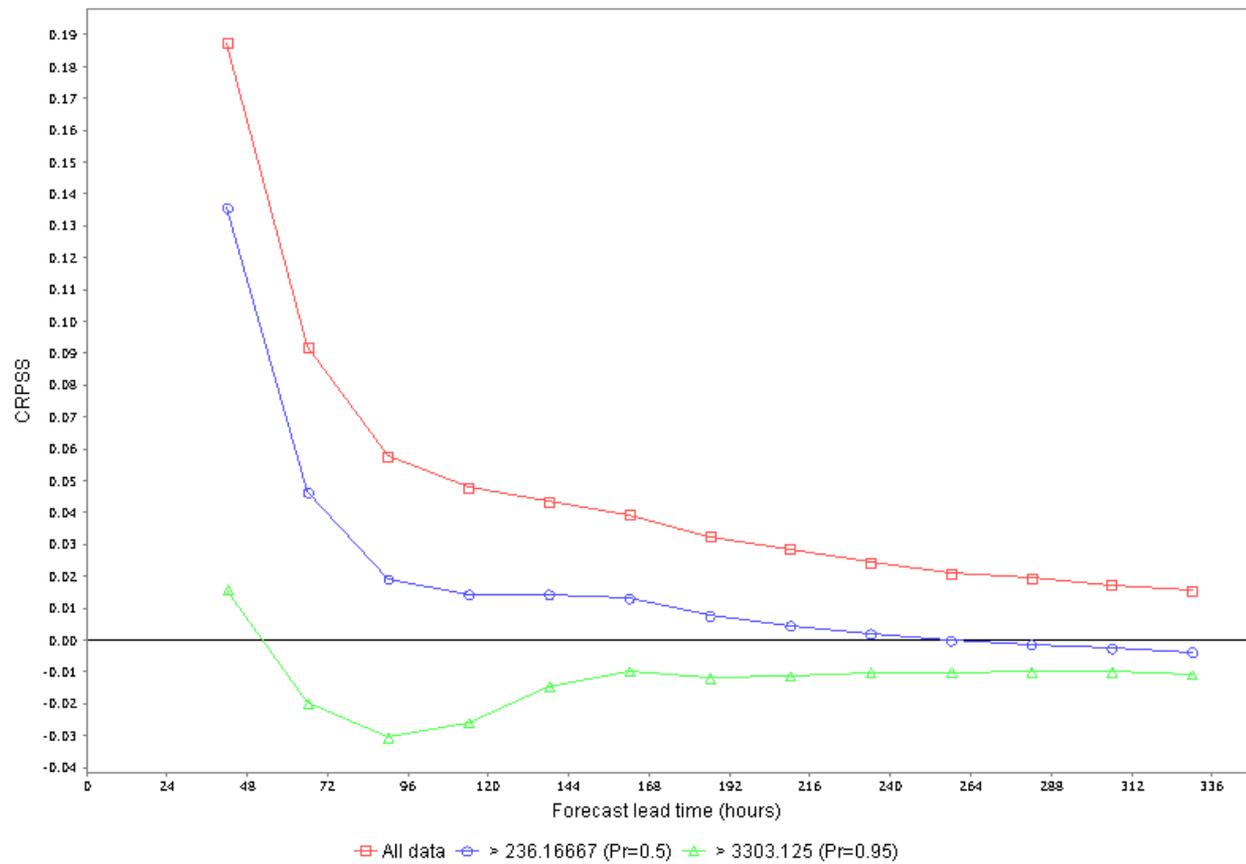
Exercise 4: Q9 (FTSC1) answer

Rank histogram.
FTSC1.Streamflow.GEFSPOST at lead hour 44.0

Q9: Compare with histogram before EnsPost (SE.13). More reliable after?

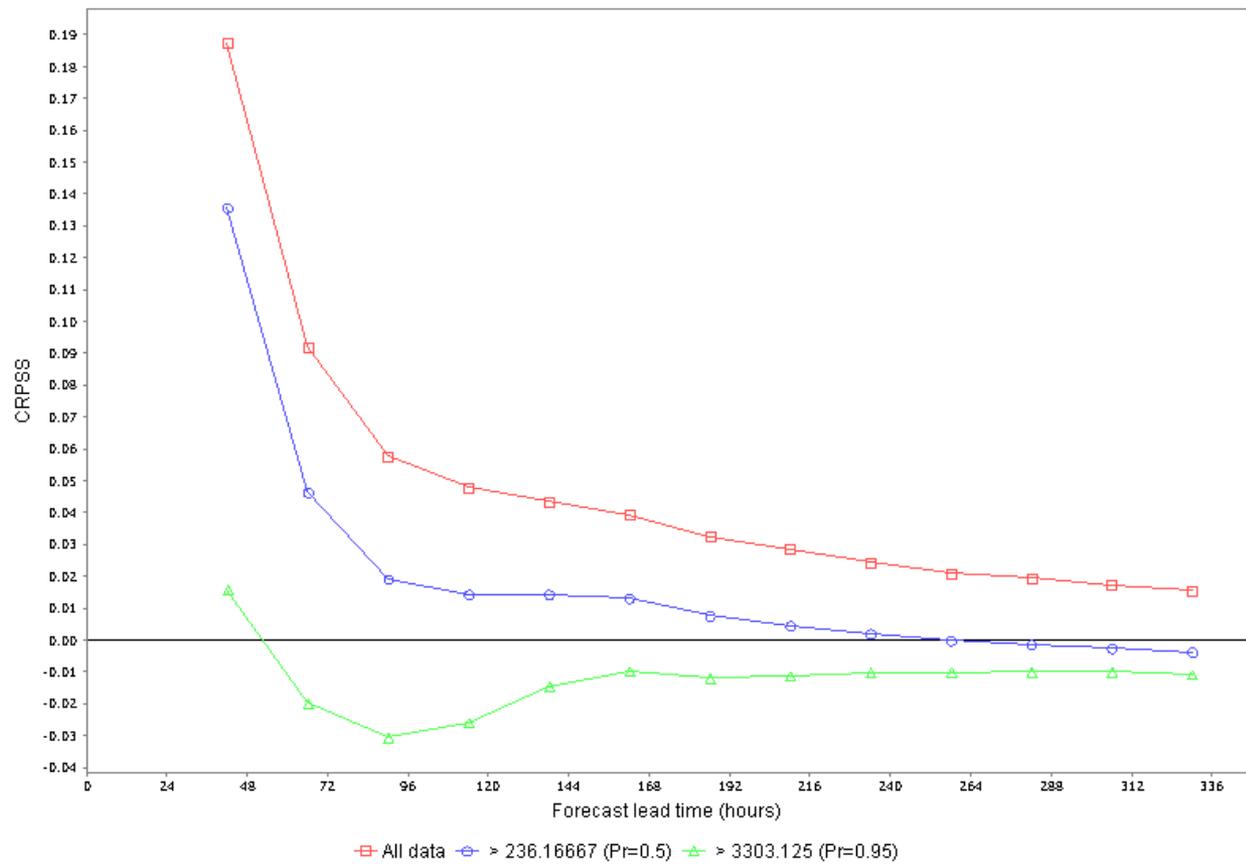**A9:** Yes, for "all data", the EnsPost substantially increases reliability

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
BLKO2.Streamflow.GEFSPOST (reference forecast: BLKO2.Streamflow.GEFS)

Q10: To what extent has EnsPost increased skill of raw forecasts?

Exercise 4: Q10 (BLKO2) answer

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
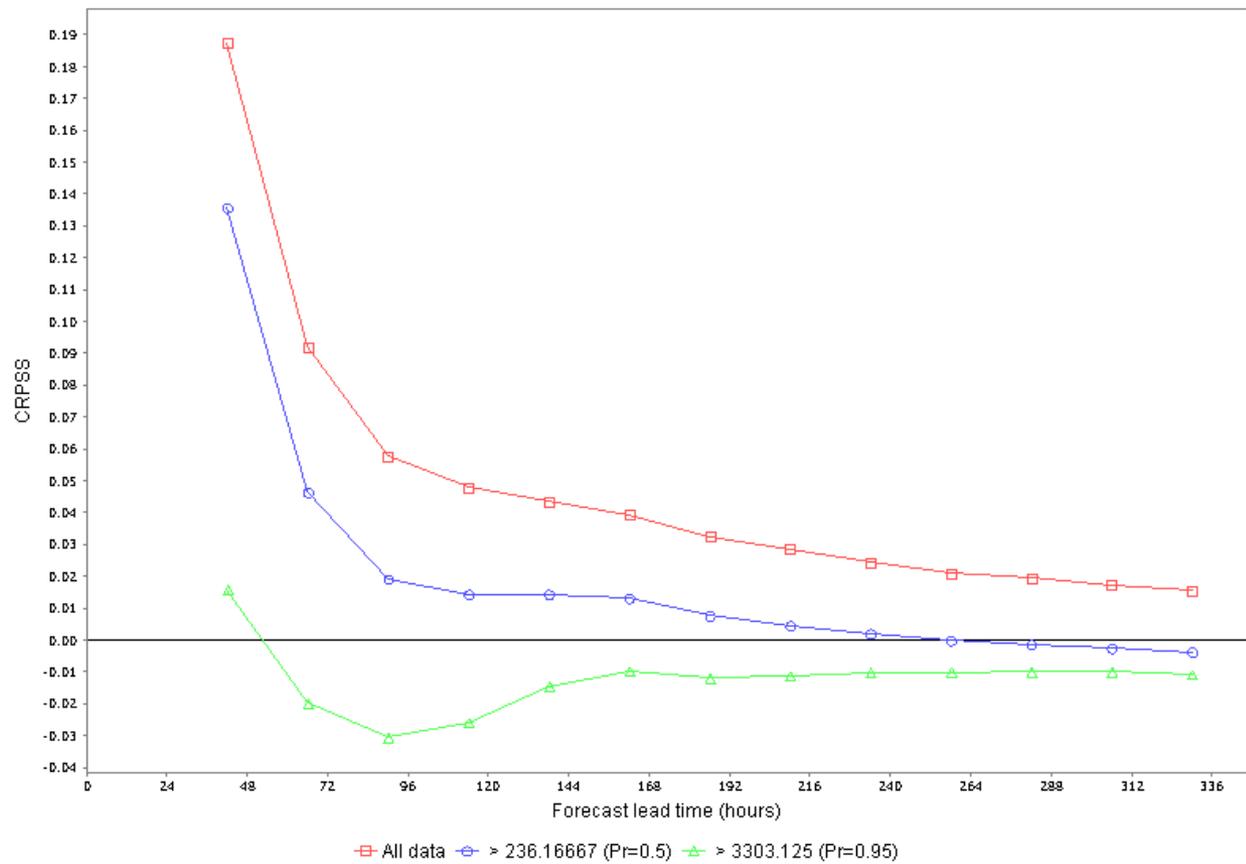BLKO2.Streamflow.GEFSPOST (reference forecast: BLKO2.Streamflow.GEFS)

Q10: To what extent has EnsPost increased skill of raw forecasts?

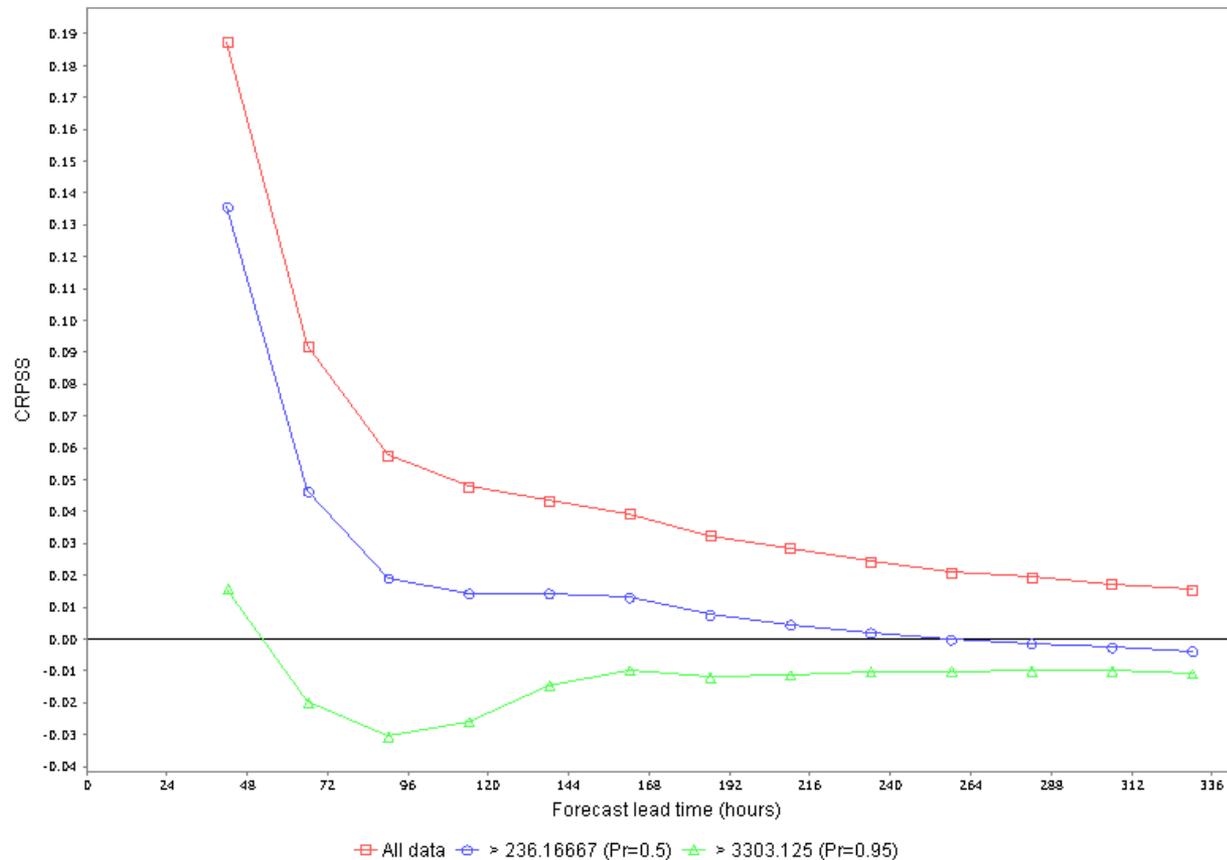**A10:** Notably, for low/moderate flows. Not for high flows after ~72 hours.

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
BLKO2.Streamflow.GEFSPOST (reference forecast: BLKO2.Streamflow.GEFS)

Q11: Why would EnsPost perform better at early lead times?

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
BLKO2.Streamflow.GEFSPOST (reference forecast: BLKO2.Streamflow.GEFS)
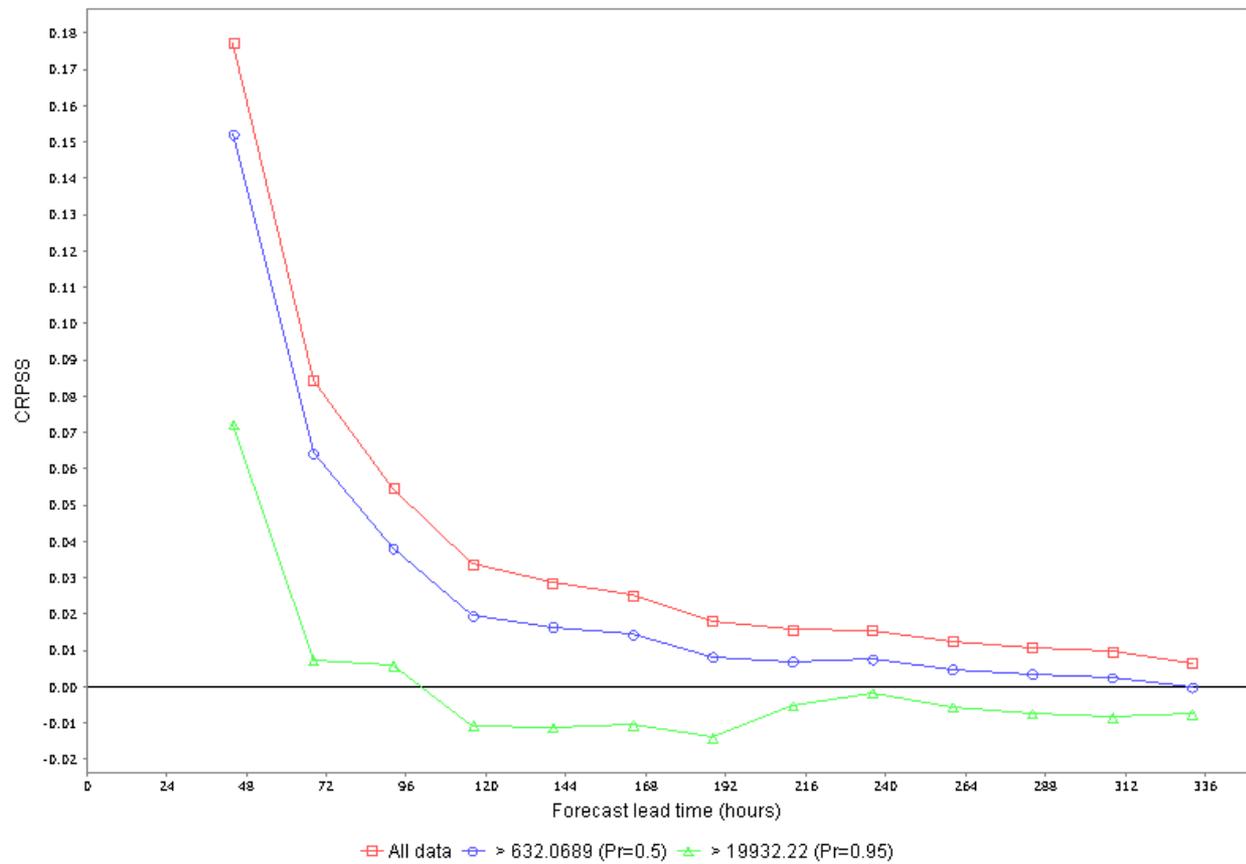
Q11: Why would EnsPost perform better at early lead times?

A11: Because EnsPost uses prior observation as a predictor (~Adjust-Q).

Exercise 4: Q12 (FTSC1)

Q12: To what extent has EnsPost increased skill of raw forecasts?

National Oceanic and Atmospheric Administration's
National Weather Service

SE.25

Office of Hydrologic Development
Silver Spring, MD

Continuous Ranked Probability Skill Score (CRPSS) by forecast lead time.
FTSC1.Streamflow.GEFSPOST (reference forecast: FTSC1.Streamflow.GEFS)

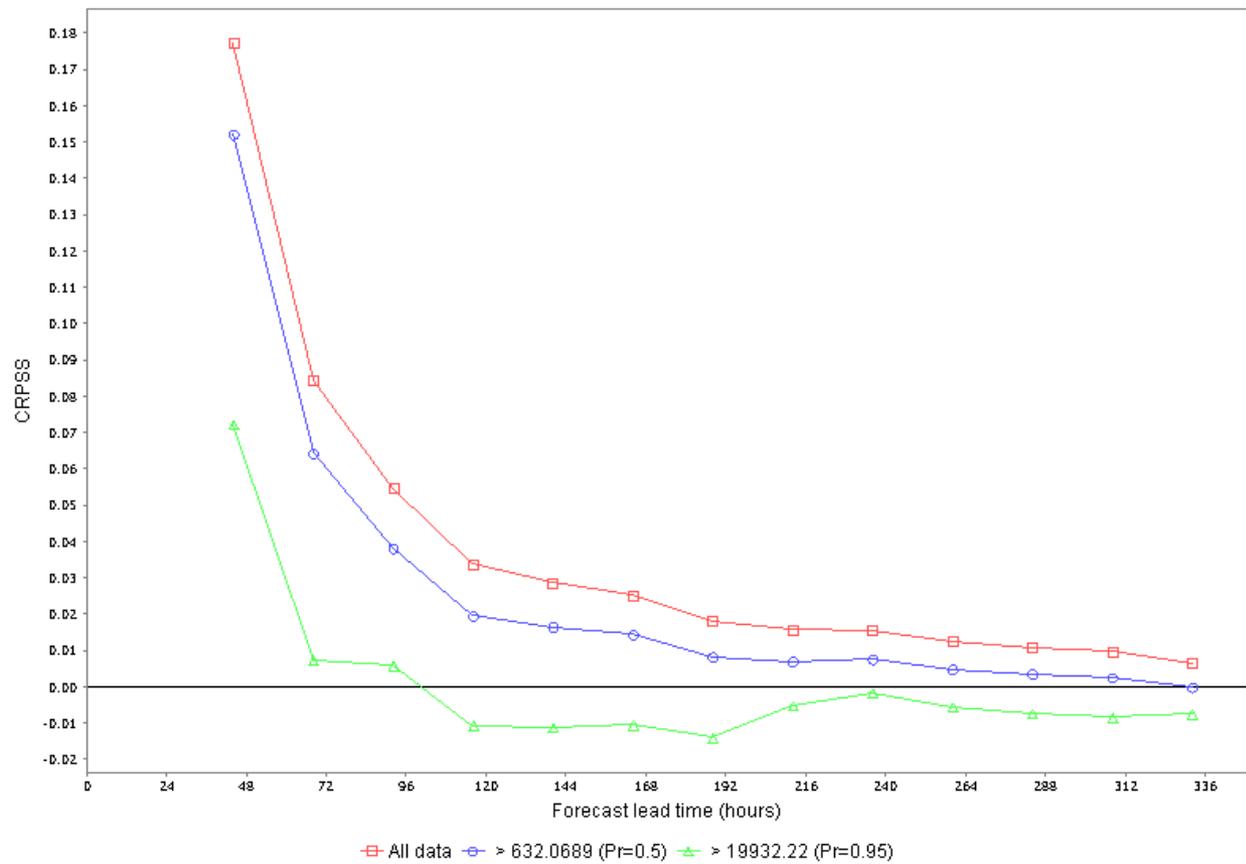**Q12:** To what extent has EnsPost increased skill of raw forecasts?

A12: As with BLKO2, notably, except for high flows at long lead times.