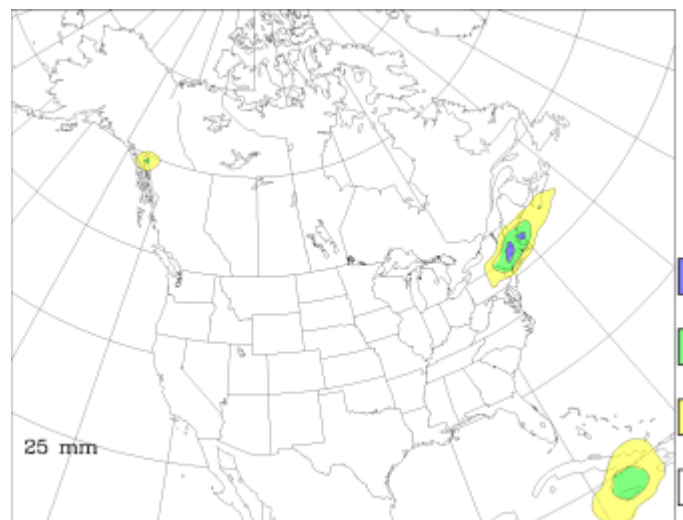
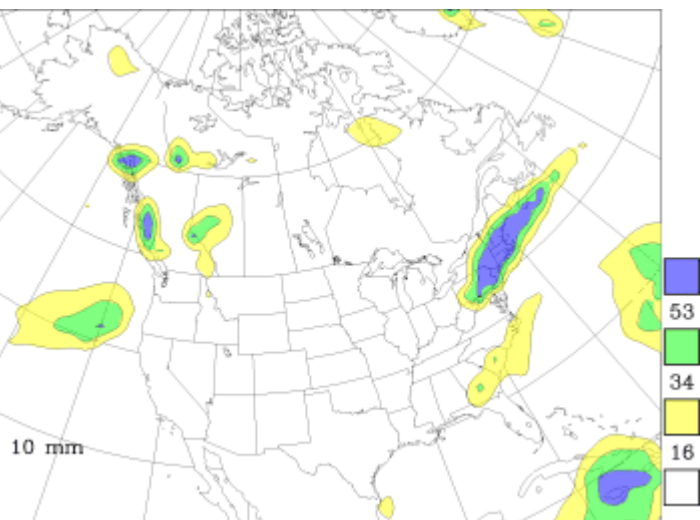
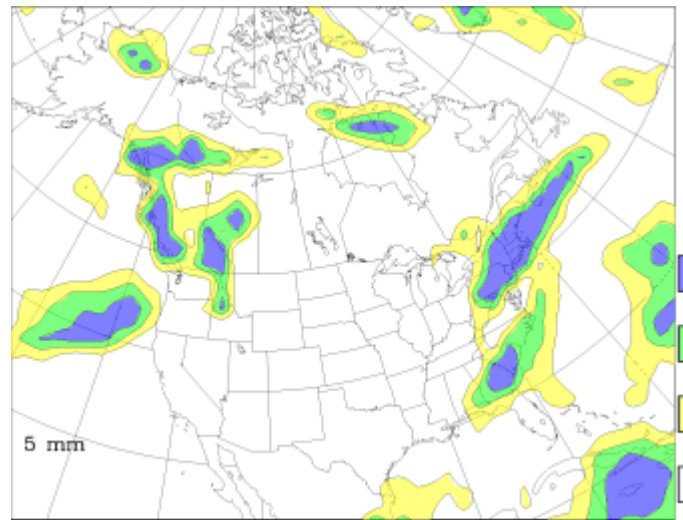
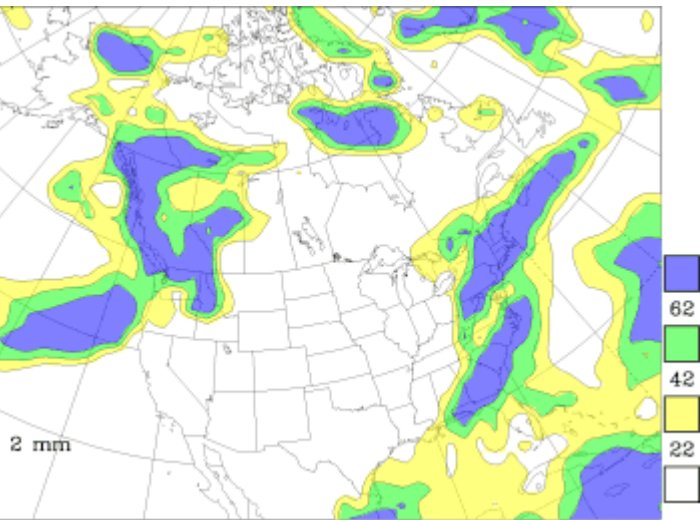


A diagnostic verification of
probability of quantitative
precipitation forecasts (PQPF)
from the Canadian Ensemble
system

Laurence Wilson and Syd Peel
Meteorological Service of Canada

Calibrated PQPF forecasts



weatheroffice.ec.gc.ca/ensemble/index_e.html

Maps available for this domain; data available for Download by ftp

Outline

BUT....

- This is about verification of the precipitation forecasts
 - using percentile thresholds rather than physical thresholds
 - Original uncalibrated forecasts
 - Focus on extremes (high impact weather)
- Data and method
- Results for 24h amounts
- Results for 3 day accumulations
- Results for 10 day dry periods
- Discussion

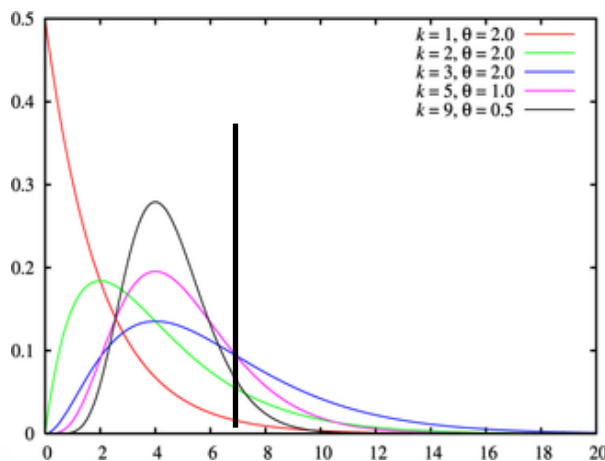
Data and method

- Data

- 3.5 years of ensemble forecasts of precipitation from 36 Canadian stations, 24h accumulations, 0 to 10 days
- Corresponding observations quality controlled without reference to models
- Verification sample stratified into warm and cool seasons
- Long-term precipitation climatology (~30 years) for all 36 stations as distribution

- Method

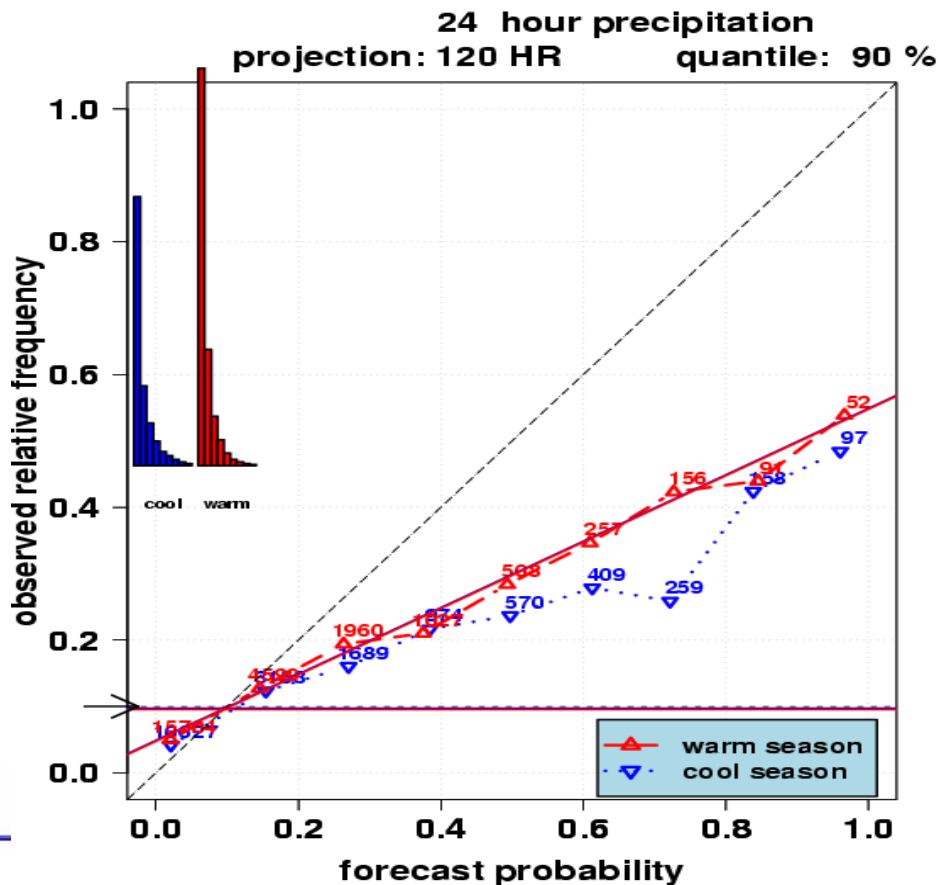
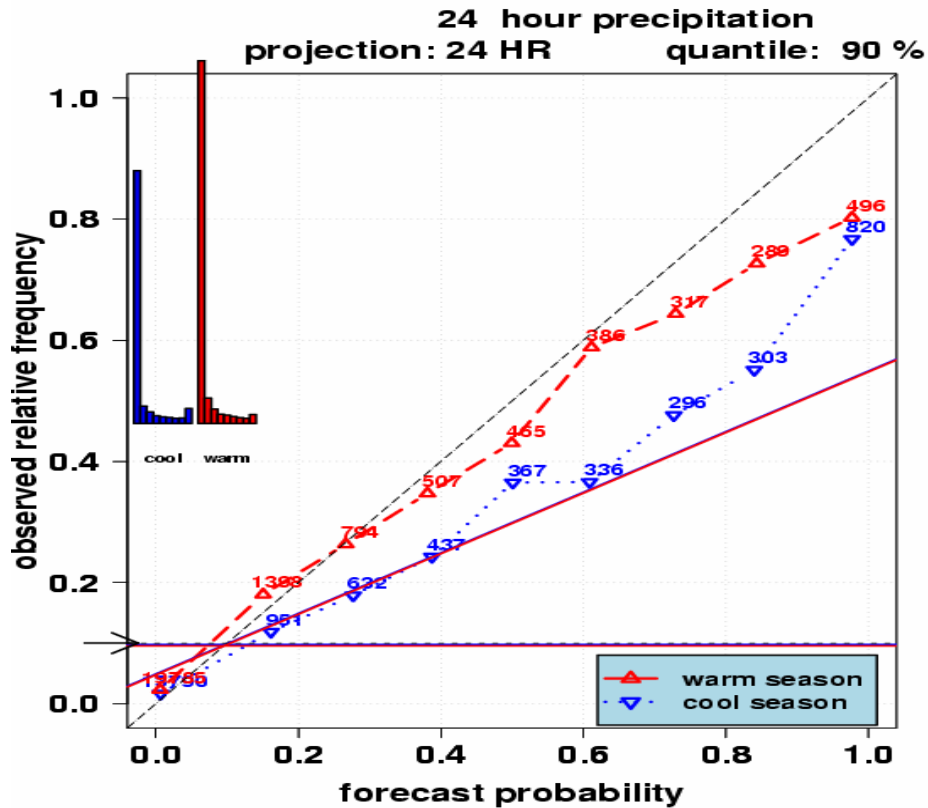
- Using the long-term climatology, find 90th, 95th and 99th percentile thresholds for each station.
- E.g. 90th percentile for Vancouver is 14.4 mm
- Probability of exceedence of these thresholds as estimated from the ensemble forecast distribution (gamma distributions)



Verification Methods

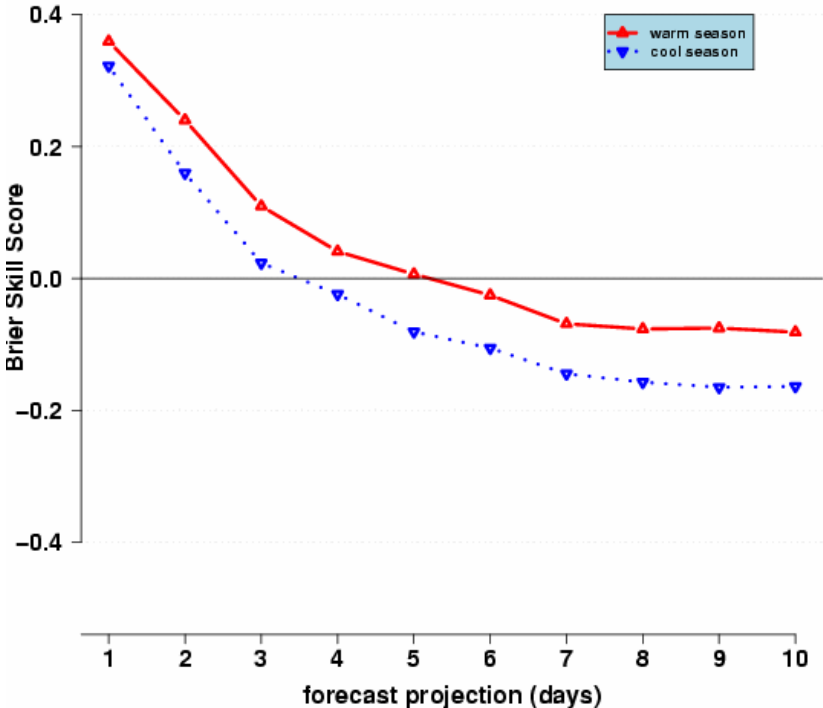
- Reliability Tables
 - Measures reliability (bias for probability forecasts)
 - Sharpness
- The Brier Skill Score
 - Measures Skill: The percent improvement of the forecast compared to a standard forecast (climatology in this case)
 - Can be partitioned into Reliability and Resolution
- The Relative Operating Characteristic Area
 - Measures Discrimination: The ability of the forecast system to accurately separate conditions preceding the event from conditions preceding the non-event
- These three, a score and two graphical measures make up a good basis for diagnosis of the quality of the forecasts

Reliability – 1 and 5 days – 90%

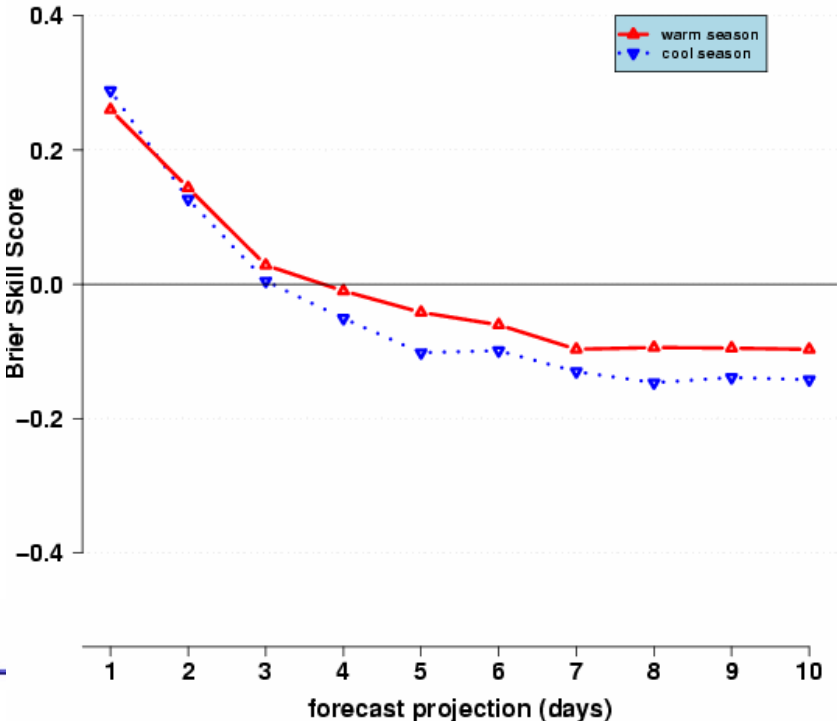


BSS for 90% and 95% threshold

Brier Skill Score
24-hour precipitation threshold: 90 %
seasonal comparison

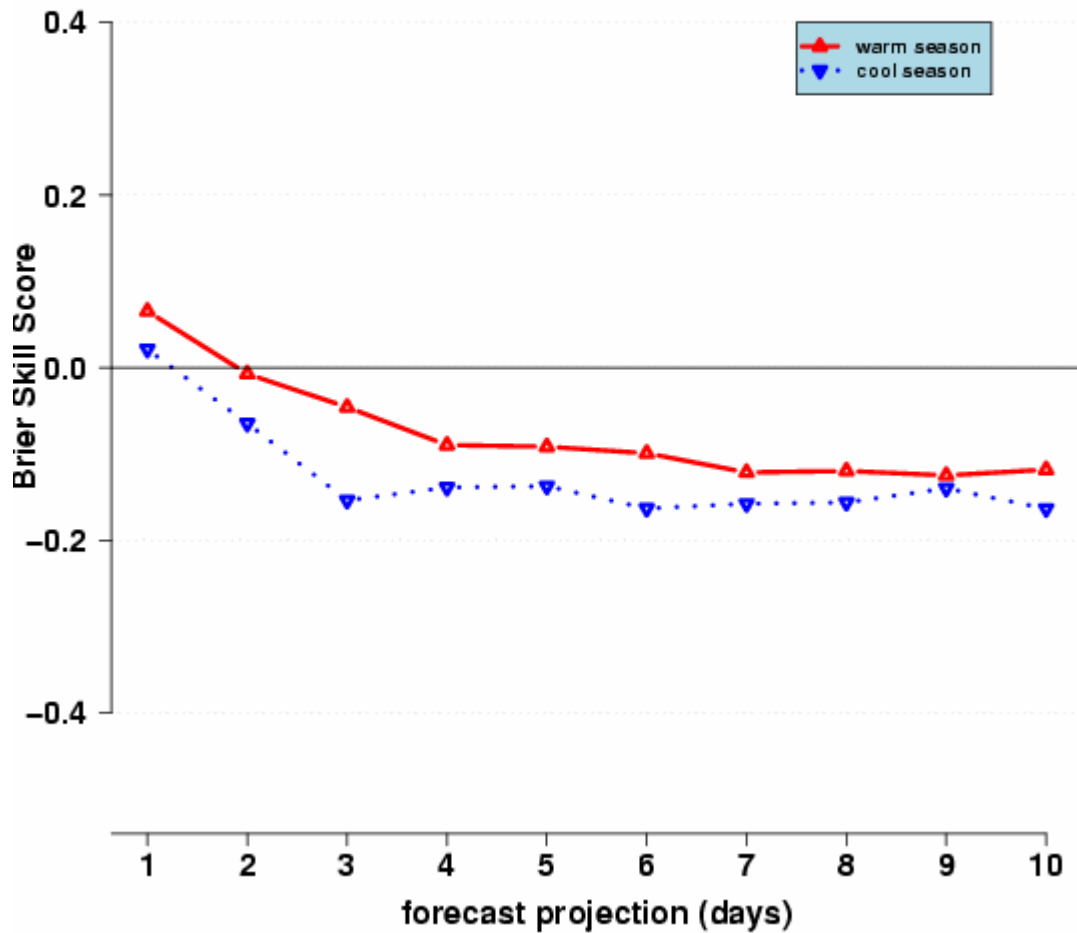


Brier Skill Score
24-hour precipitation threshold: 95 %
seasonal comparison

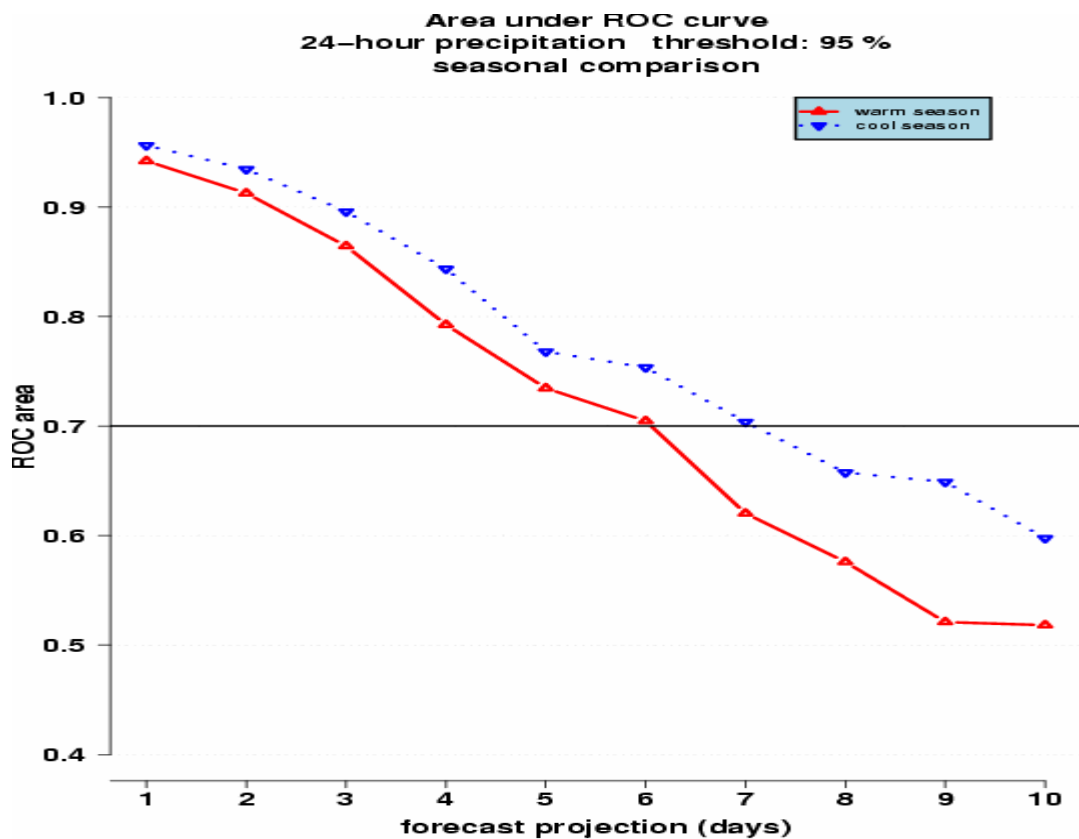
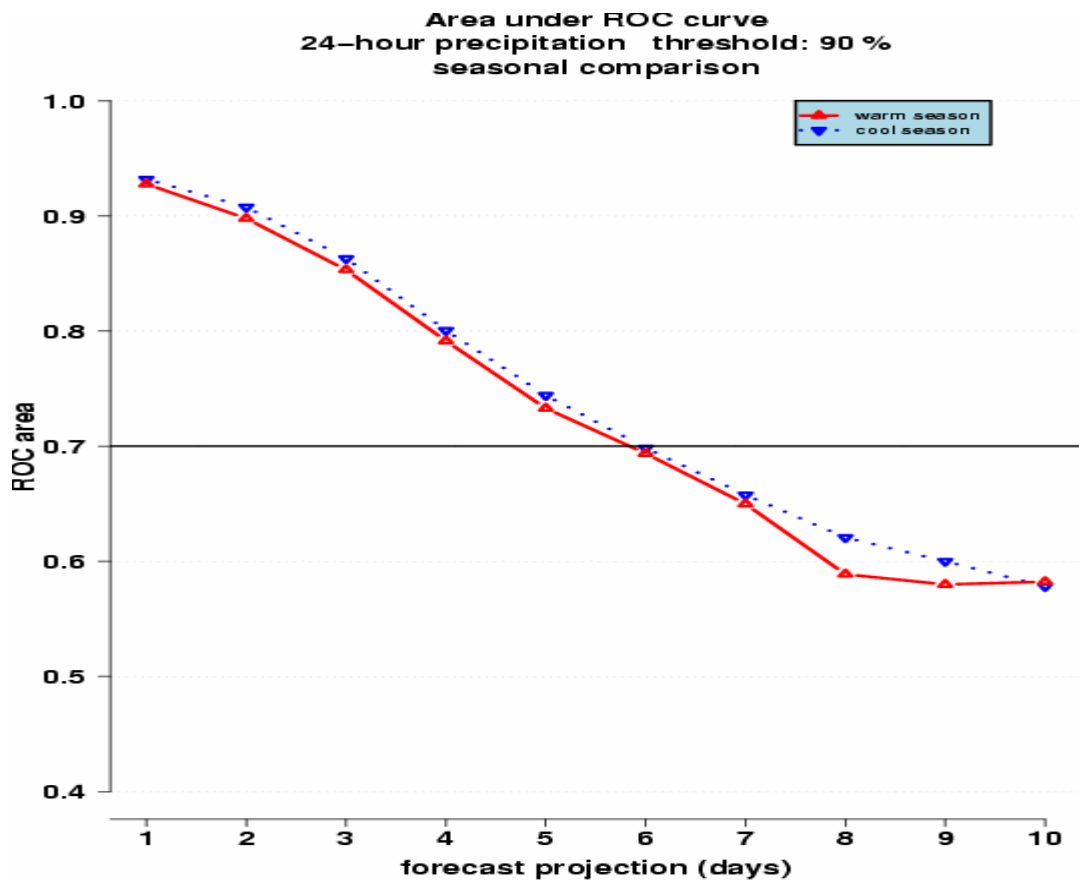


BSS – 99% threshold

Brier Skill Score
24-hour precipitation threshold: 99 %
seasonal comparison

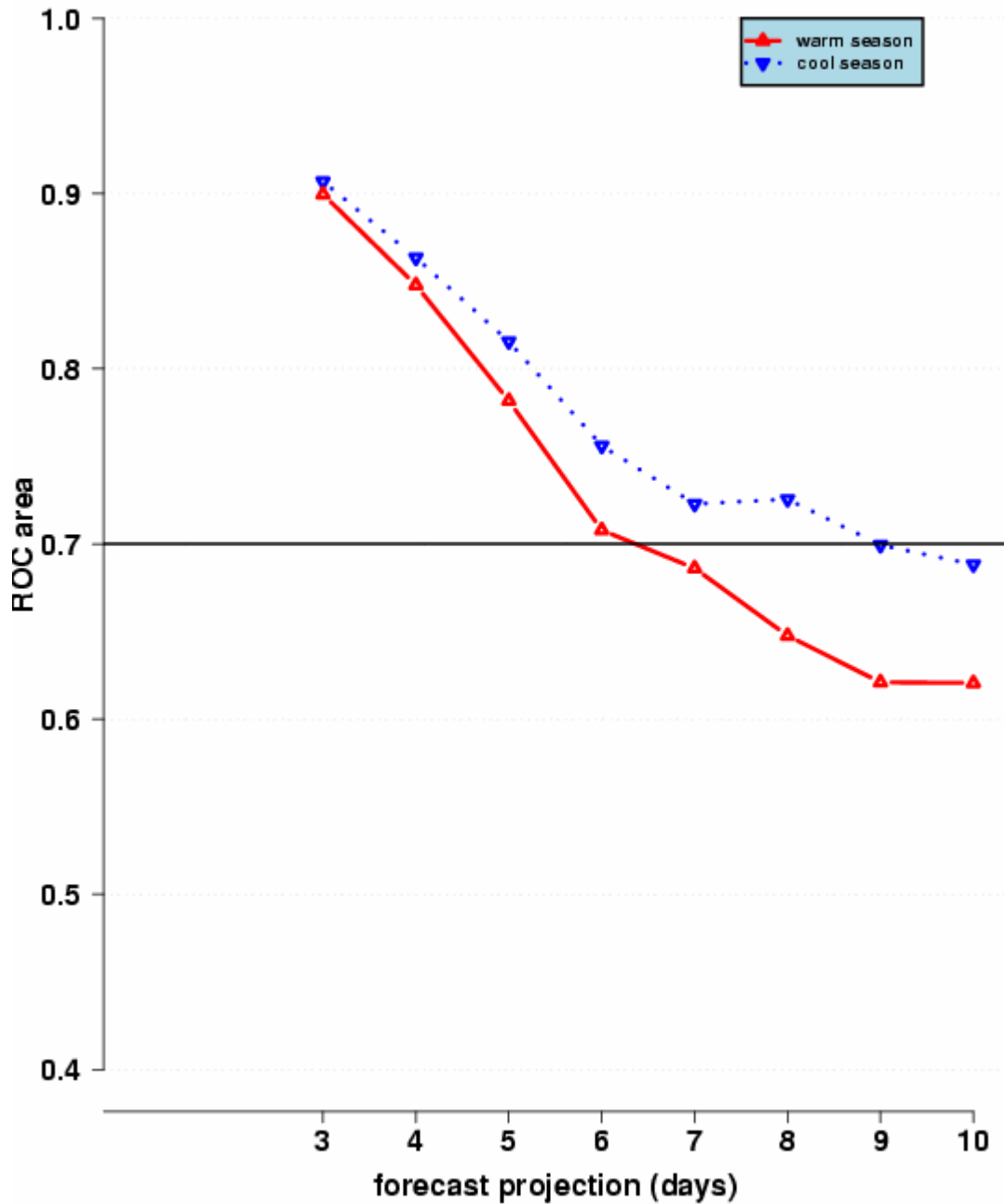


ROC area – 90% and 95% threshold

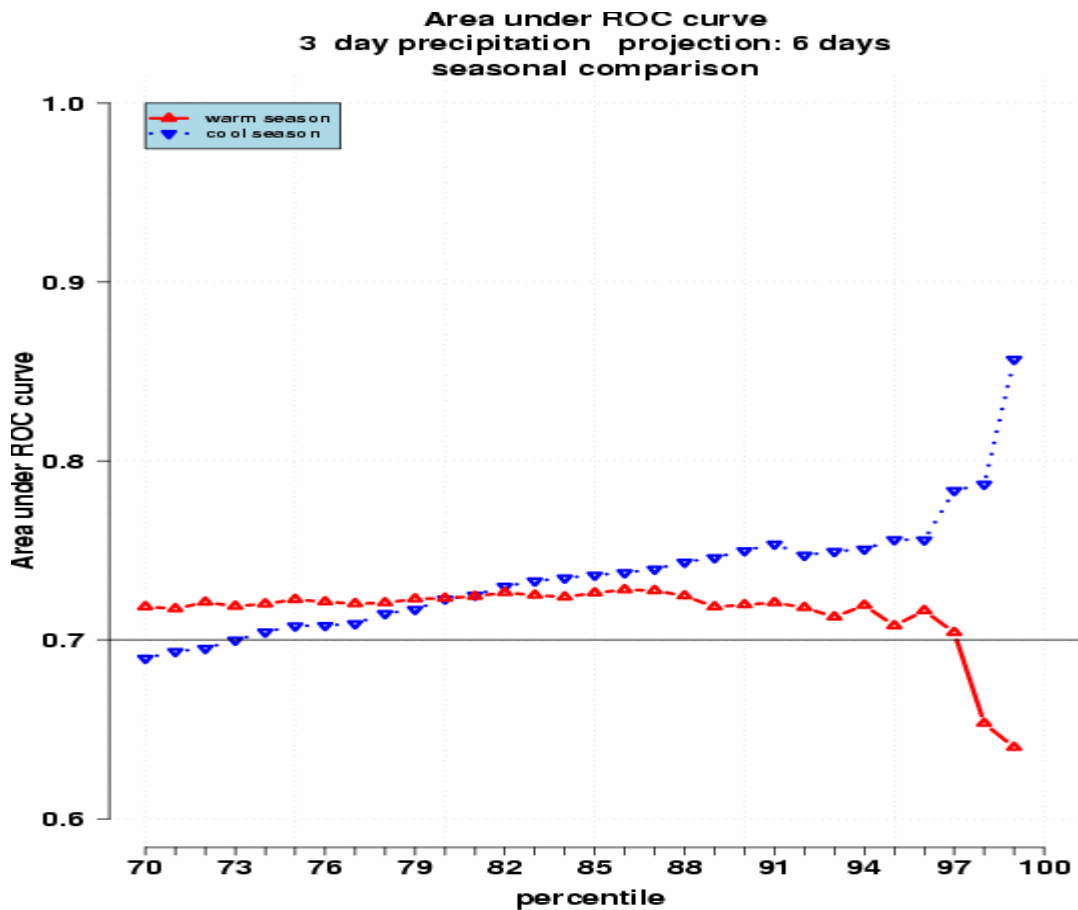
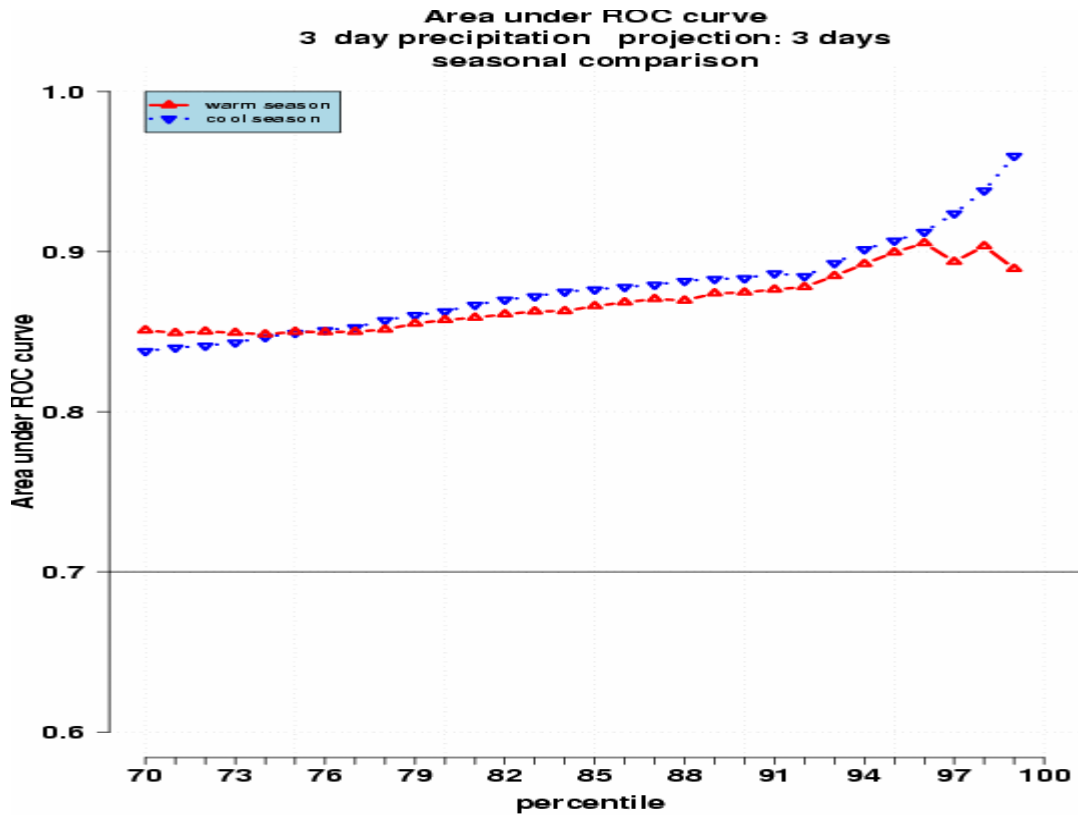


ROC area – 95% – 72h accumulation

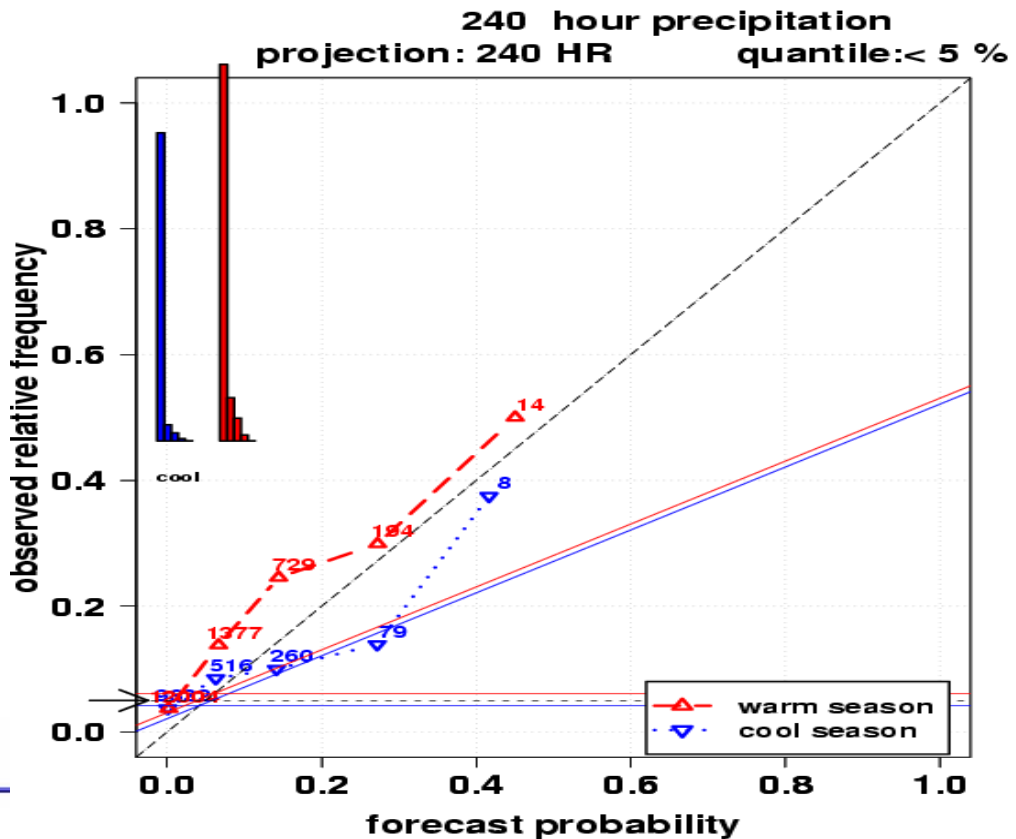
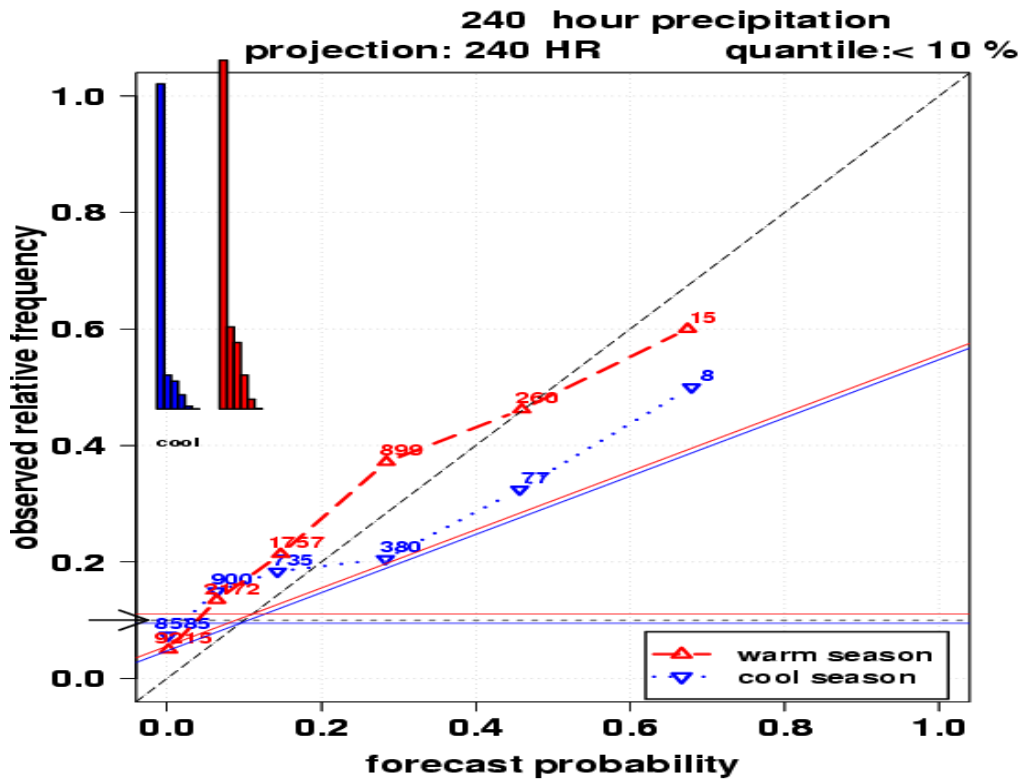
Area under ROC curve
3 day precipitation threshold: 95 %
seasonal comparison



ROC area – 72 h accumulation – as function of threshold



10 day dry period – 10% and 5%



Summary

- Diagnostic verification of probability forecasts
 - Brier Skill Score – SKILL
 - RELIABILITY and RESOLUTION
 - Graphical – Reliability table with SHARPNESS
 - Graphical – ROC and area DISCRIMINATION
- These results show that the eps has some skill in forecasting extremes: 5 days at 90%, 3 days at 95% and 1 day at 99%
- Forecasts of dry periods are reliable for the warm season
- Ensemble model recently changed
 - Results are a baseline for future eps