More on ACCESS-GE Rainfall Verification

D. Smith

Model Evaluation Workshop
May 19 2015

ACCESS-GE rainfall verification

- Review domain averaged error/spread results for members and mean, with ACCESS-G comparison
- Look at some real ensemble verification scores, ROCA and Brier
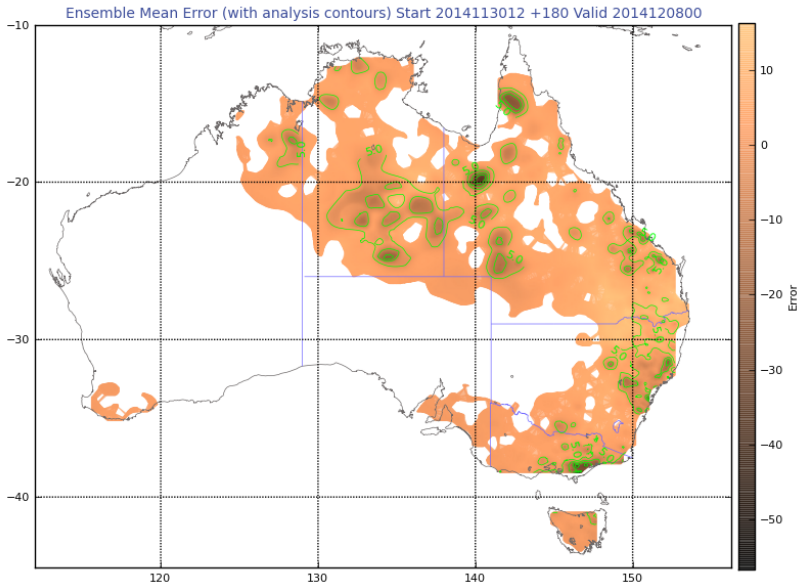- 10 day forecast, 24 member global ensemble
- N216 (60km) resolution,

Previous efforts calculated domain RMS errors for all
ensemble members, and the mean, comparing
results with ACCESS-G

For November 2014, this looked like

Rainfall Verification Against AWAP November 2014
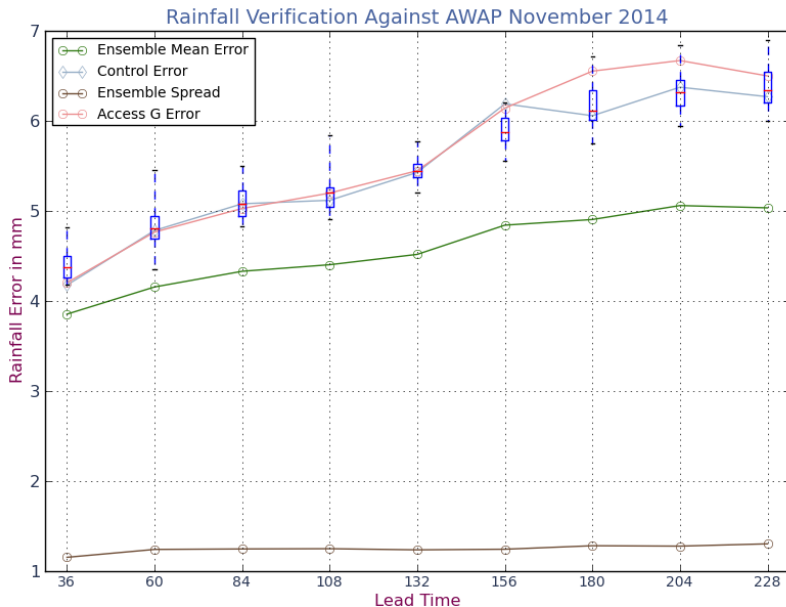
Error snapshot at 2014113012+180

Ensemble Mean Error (with analysis contours) Start 2014113012 +180 Valid 2014120800

Largest errors are negative, in heavy rain areas

Choose a rainfall (exceedance) threshold, x_{mm}

At any grid point:

What does the analysis say?

<div align="center">Yes (1) or No (0)</div>

What do the members say?

<div align="center">Yes (1) or No (0) 24 times</div>

Average the members' opinions to give a 'probability' fraction, between 0 and 1.
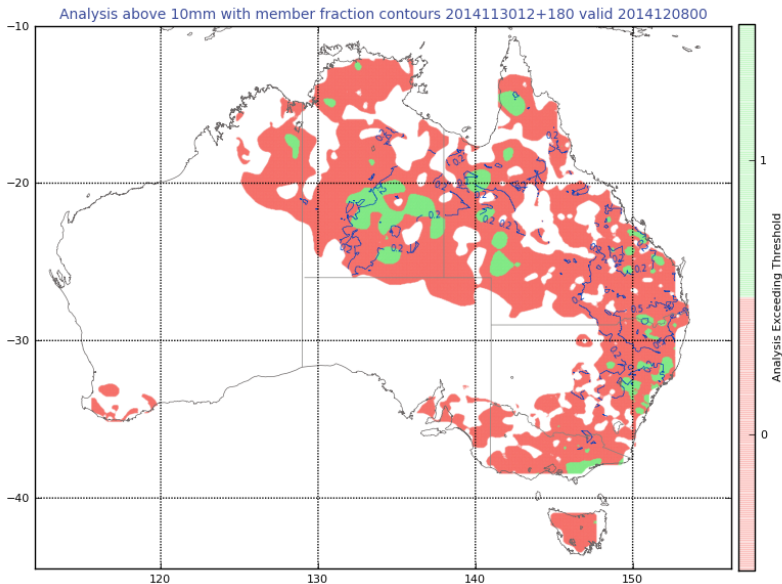
Define 'error' as

$$\text{error} = \text{member fraction} - \text{analysis}$$

and take domain RMS (weighted).

Low(good) Brier scores are achieved by high member consensus at analysis *Yes* gridpoints and low member consensus at analysis *No* gridpoints
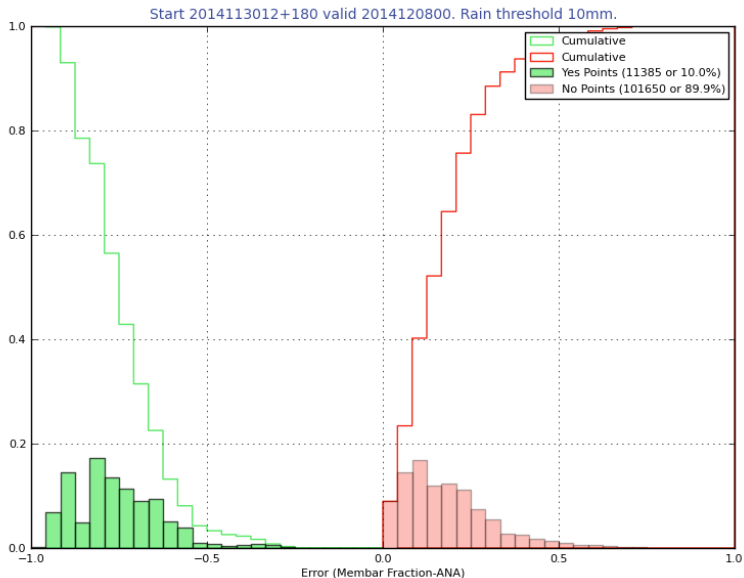
Take a picture for 10mm threshold on November 30 forecast, +180 lead time

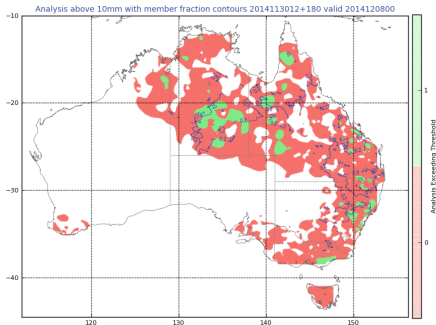# *Yes*,*No* Points, 10mm Threshold, and Member Contours

Analysis above 10mm with member fraction contours 2014113012+180 valid 2014120800

Note 50% contour in *No* area, S QLD

Start 2014113012+180 valid 2014120800. Rain threshold 10mm.

Cumulative
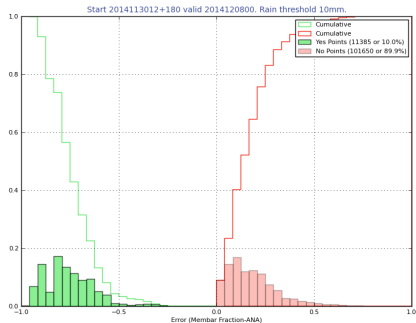Cumulative
Yes Points (11385 or 10.0%)
No Points (101650 or 89.9%)

Error (Membar Fraction-ANA)

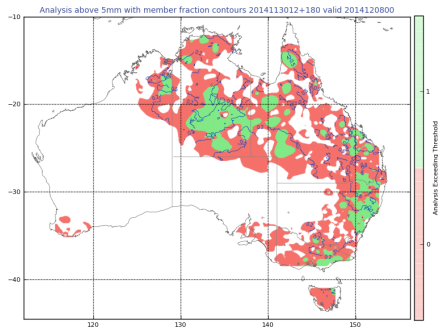Yes,No Points and member contours

Error Histograms

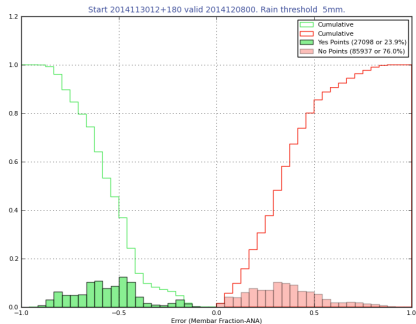- Almost 90% *No* points at this threshold
- Error histogram for *No* points appears far healthier

Repeat at 5mm threshold

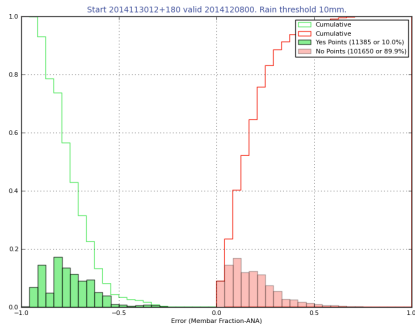Yes,No Points and member contours
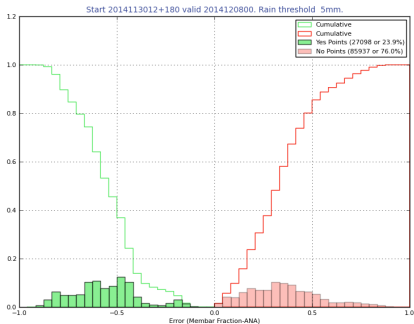
Error Histograms

Analysis above 5mm with member fraction contours 2014113012+180 valid 2014120800

Start 2014113012+180 valid 2014120800. Rain threshold 5mm.

- More *Yes* points, just under 25%
- *Yes* histogram creeping towards zero

Compare histograms 5/10mm thresholds

10mm (score=0.102, better)   5mm (score=0.2, worse)
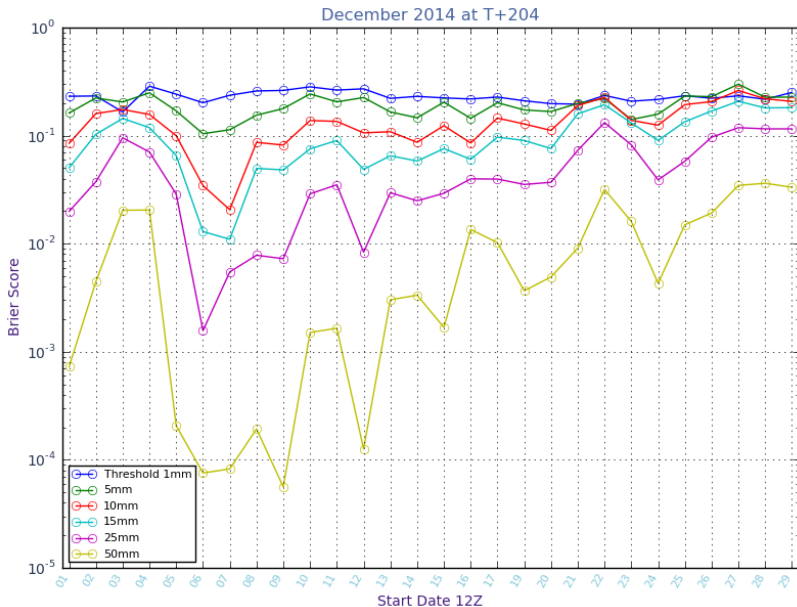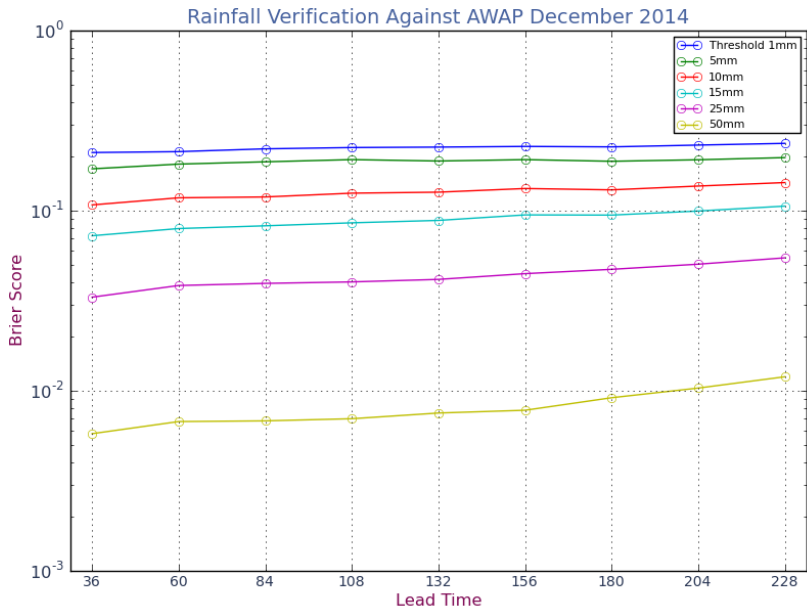

On decreasing the threshold:

- *No* points shifted away from 0 (worse)
- *Yes* points shifted towards 0 (better)

Improvement at 10mm appears to come from the *No* points

December 2014 at T+204

December Brier Results


Rainfall Verification Against AWAP December 2014

The ROCA score is concerned with ability to *discriminate*

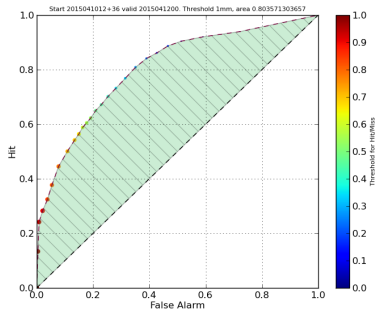Once again, consider analysis *Yes* and *No* points.

- Construct ROC curve (parametrically) by taking a series of warning(decision) thresholds
- Proportion of *Yes* points exceeding warning threshold is *hit rate*, similarly for *false alarm rate* at *No* points
- Area under ROC curve is the ROCA score
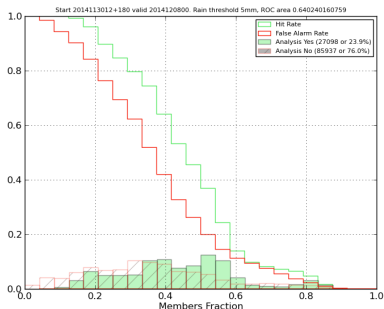- ROCA above $\frac{1}{2}$ means skill, perfect score is 1
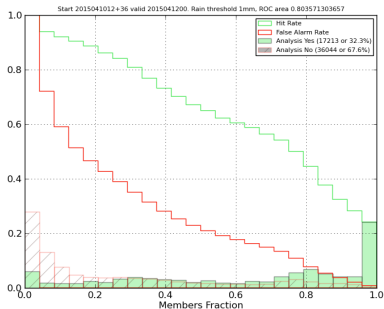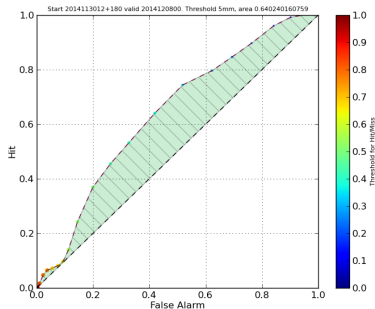
Histograms

ROC Curve

- *Yes* histogram heavy right, *No* heavy left
- Reverse cumulative summation = 'hit', 'false alarm' rates
- Rapid FA rate decrease and slow HR decrease → gap
- ROC well above diagonal means positive skill
- ROCA Score = 0.80, by trapezoid approximation

Histograms
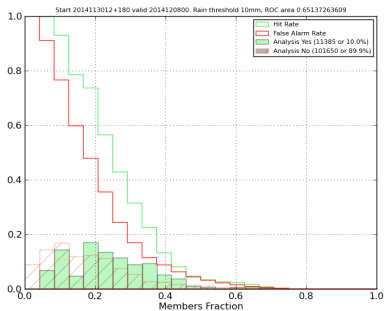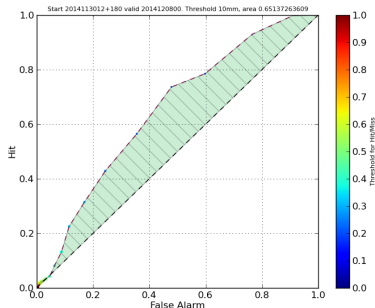
ROC Curve

- *Yes* histogram not so good, *No* histogram ok
- Colourbar on ROC picture indicates warning threshold
- Calculated score = 0.64
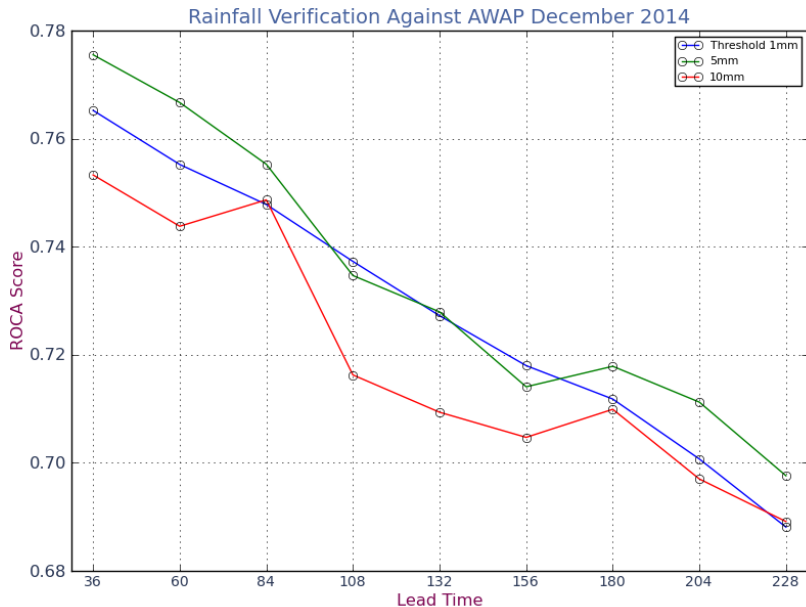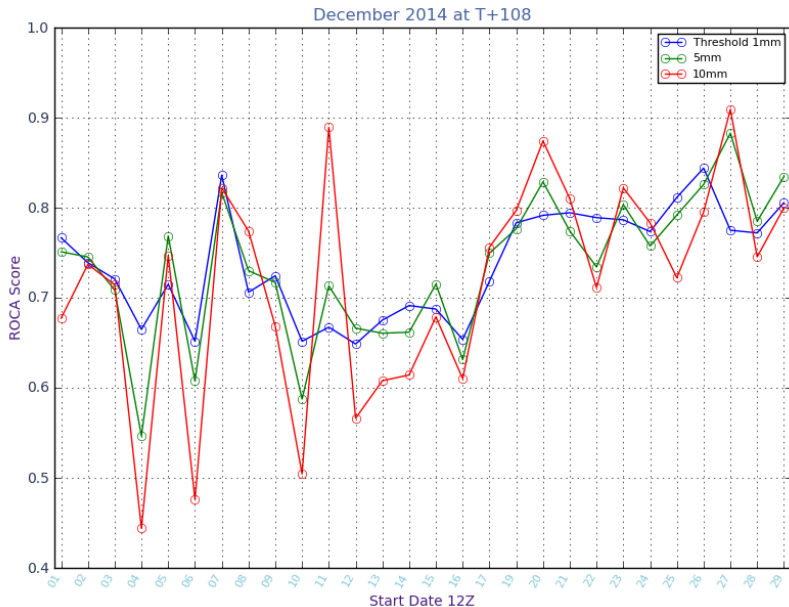
And 10mm threshold

Histograms

ROC Curve

- Calculated score $= 0.65$, slightly better than 5mm case
- Histogram for *Yes* points (10%) not looking healthy
- ROC stayed *above* diagonal ($+$ skill)

December ROCA Results

Rainfall Verification Against AWAP December 2014

December ROCA Results

December 2014 at T+108

- Also need ensemble CRPS score
- Apply rainval to members, mean and compare ACCESS-G for various other scores
- Compare with other ensemble(s)