Assessing high resolution forecasts using fuzzy verification methods

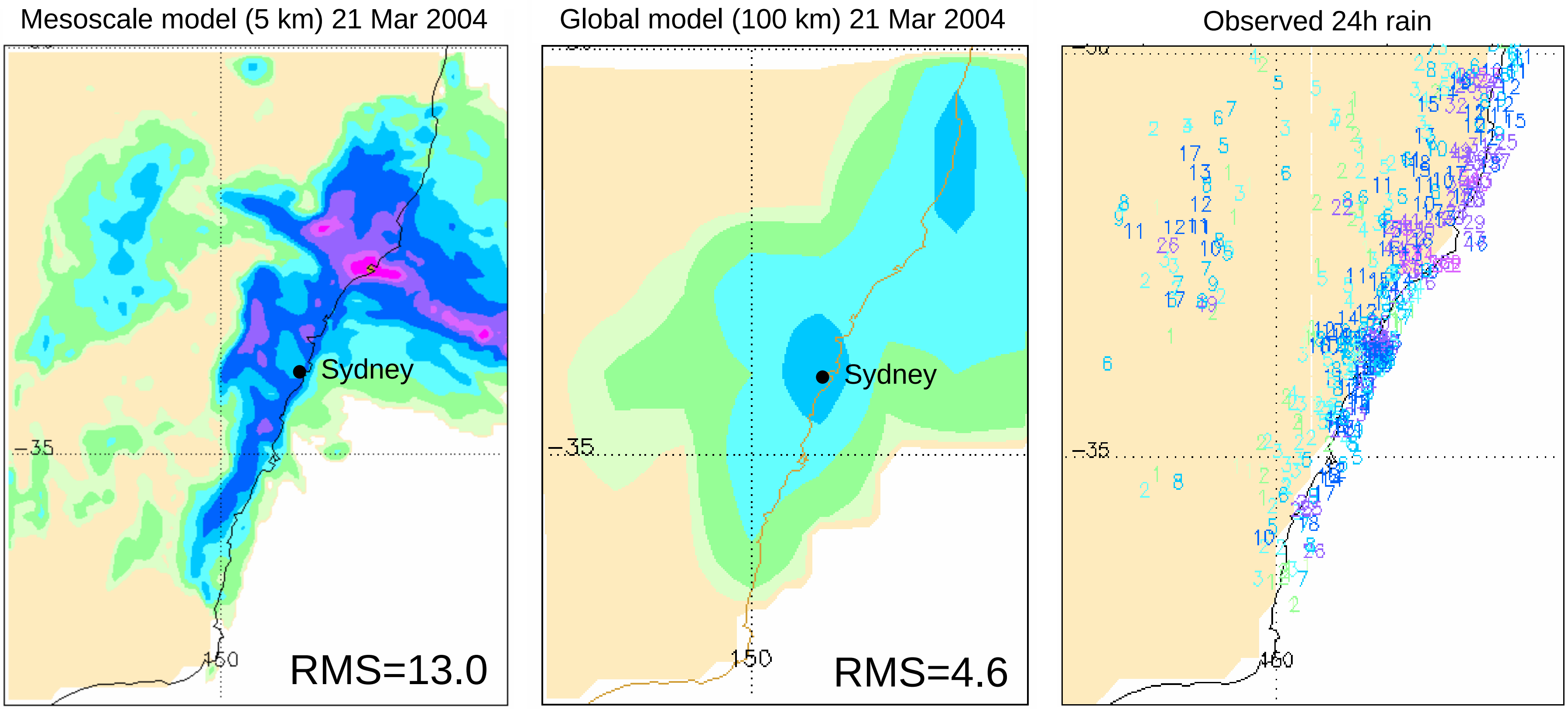
Beth Ebert

Bureau of Meteorology Research Centre, Melbourne, Australia

Thanks to Nigel Roberts, Barbara Casati, Frederic Atger, Felix Ament , Daniel Leuenberger, Urs Germann, Mike Kay, Susanne Theis, Ulrich Damrath, Daniela Rezacova

High vs. low resolution

Which rain forecast would you rather use?

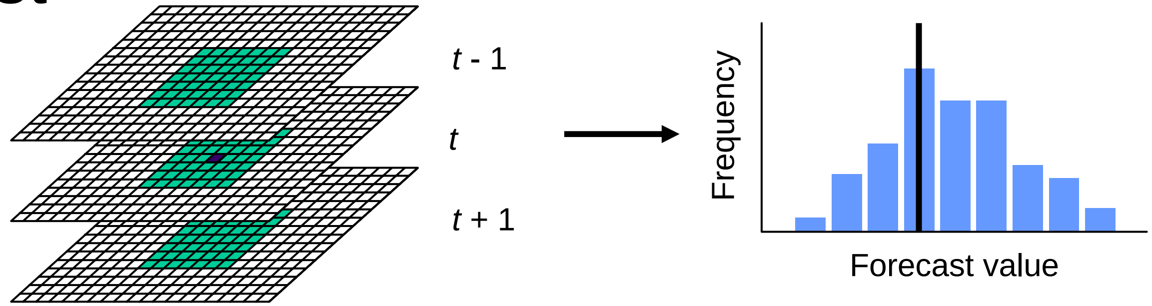


What makes a useful forecast?

- Resembles the observations on the broader scale
- Predicts an event somewhere near where it was observed
- Predicts the event over the same area (i.e., with the same frequency) as observed
- Has a similar distribution of intensities as the observations
- Looks like what a forecaster would have predicted if she'd had knowledge of the observations

"Fuzzy" verification methods

- Don't require an exact match between forecasts and observations
  - Unpredictable scales
  - Uncertainty in observations
- Look in a space / time neighborhood around the point of interest



  - Evaluate using categorical, continuous, probabilistic scores / methods

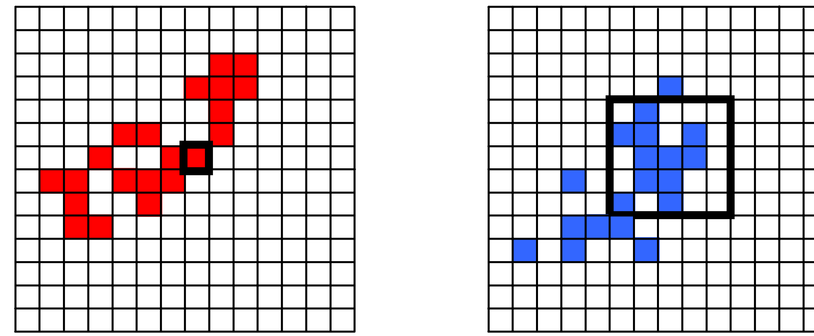
"Fuzzy" verification methods

- First (?) suggested by H. Brooks at 1998 Mesoscale Verification workshop
  - Brooks et al. (1998)
  - Zepeda-Arce et al. (2000), Weygandt et al. (2004)
  - Atger (2001)
  - Damrath (2004)
  - Casati et al. (2004)
  - Germann and Zawadski (2004)
  - Theis et al. (2005)
  - Roberts (2005)
  - Rezacova et al. (2006)

Fuzzy verification framework

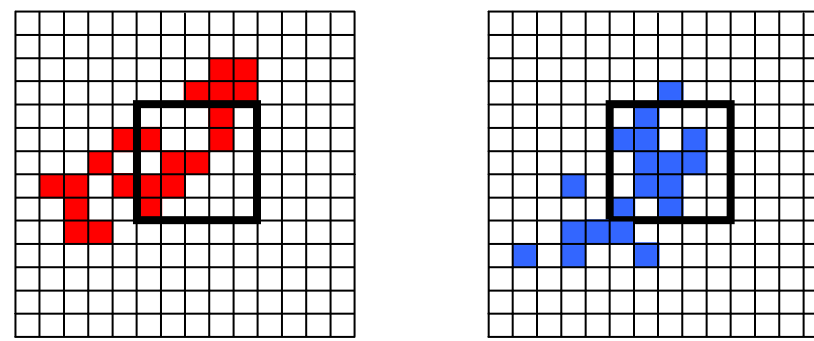
Fuzzy methods use one of two approaches to compare forecasts and observations:

single observation – neighborhood forecast



observation forecast

neighborhood observation – neighborhood forecast



observation forecast

Australian Government
Bureau of Meteorology

Fuzzy verification framework

Treatment of forecast data within a window:

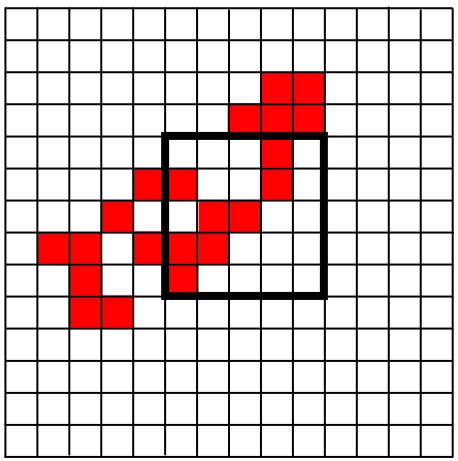
- Mean value (upscaling)
- Occurrence of event* somewhere in window
- Frequency of event in window → probability
- Distribution of values within window

May apply to observations as well as forecasts (neighborhood observation-neighborhood forecast approach)

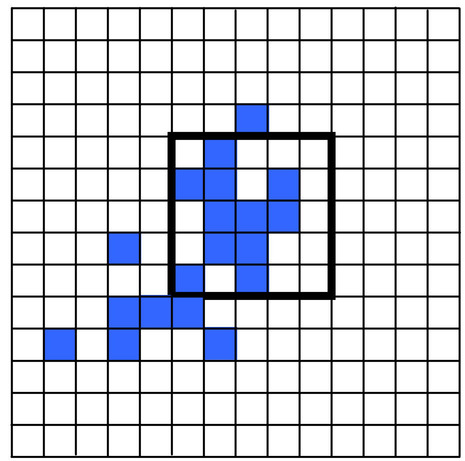
* *Event* defined here as a value exceeding a given threshold, for example, rain exceeding 1 mm/hr

Example: Fractions skill score
(Roberts and Lean 2005)

Compares fractional coverage in forecast with fractional coverage in observations



observation

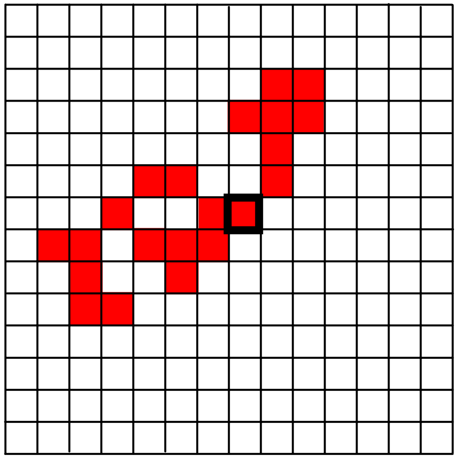


forecast

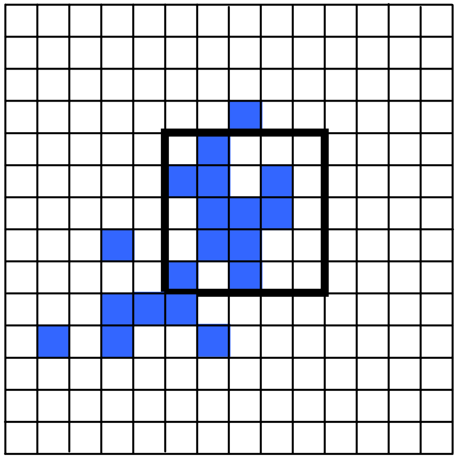
$$FSS = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(P_{fcst} - P_{obs})^2}{\frac{1}{N}\sum_{i=1}^{N}P_{fcst}^{\;2} + \frac{1}{N}\sum_{i=1}^{N}P_{obs}^{\;2}}$$

Example: Multi-category contingency table (Atger 2001)

Compares occurrence of event in forecast with observed occurrence of event



observation

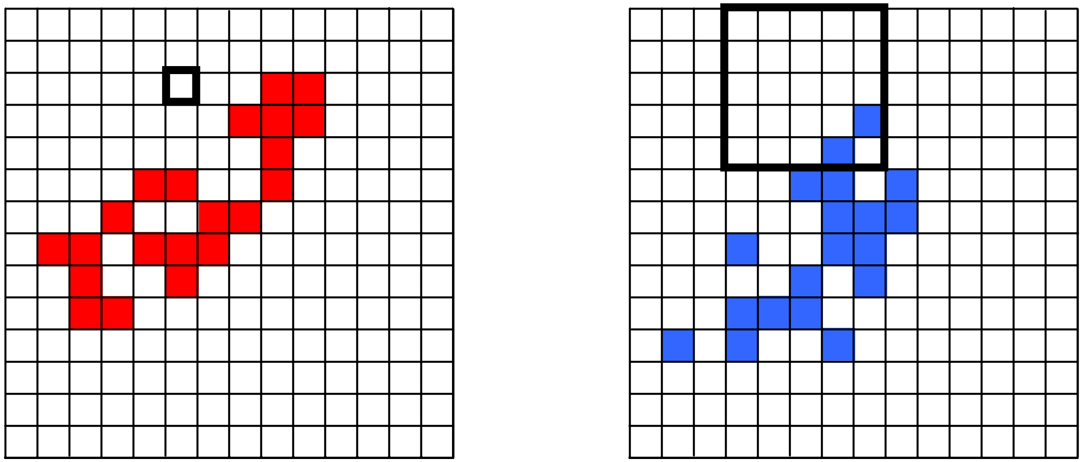


forecast

Hit = at least one forecast event in vicinity of observed event

Moving windows

Accumulate scores as windows are moved through the domain



observation forecast

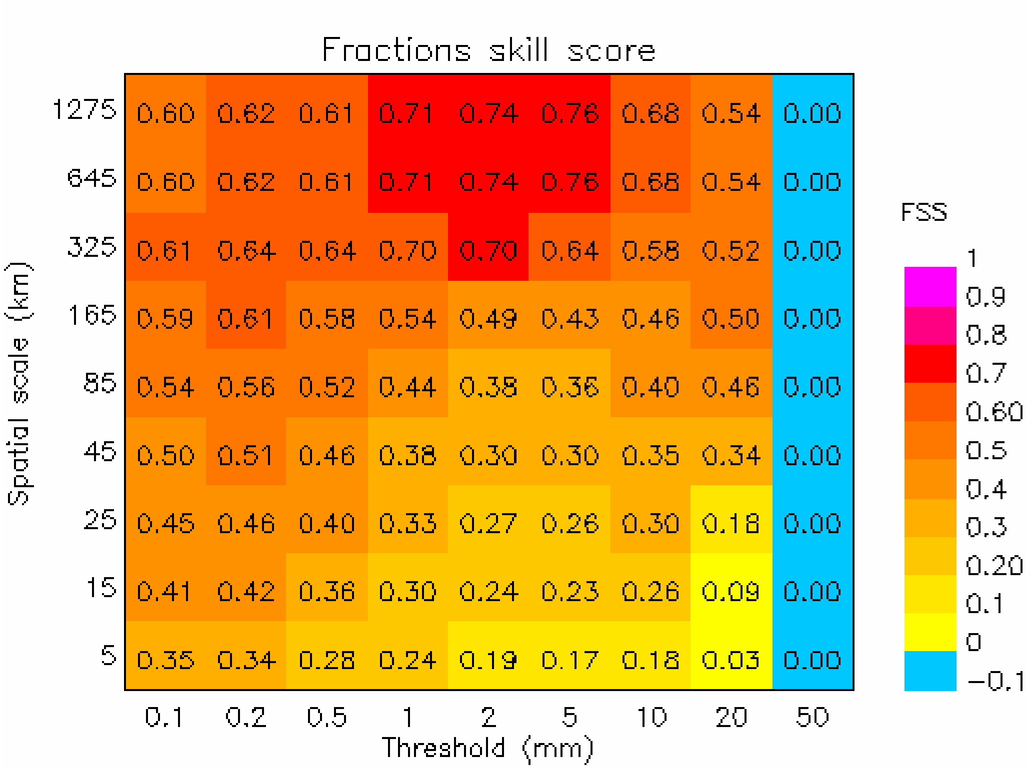
Decision models

Fuzzy method	Matching strategy*	Decision model for useful forecast
Upscaling (Zepeda-Arce et al. 2000; Weygandt et al. 2004)	NO-NF	Resembles obs when averaged to coarser scales
Minimum coverage (Damrath 2004)	NO-NF	Predicts event over minimum fraction of region
Fuzzy logic (Damrath 2004), joint probability (Ebert 2002)	NO-NF	More correct than incorrect
Fractions skill score (Roberts 2005)	NO-NF	Similar frequency of forecast and observed events
Pragmatic (Theis et al. 2005)	SO-NF	Can distinguish events and non-events
CSRR (Germann and Zawadzki 2004)	SO-NF	High probability of matching observed value
Multi-event contingency table (Atger 2001)	SO-NF	Predicts at least one event close to observed event
Practically perfect hindcast (Brooks et al. 1998)	SO-NF	Resembles forecast based on perfect knowledge of observations
Intensity-scale (Casati et al. 2004)	NO-NF	Lower error than random arrangement of obs
Area-related RMSE (Rezacova et al. 2006)	NO-NF	Similar intensity distribution as observed

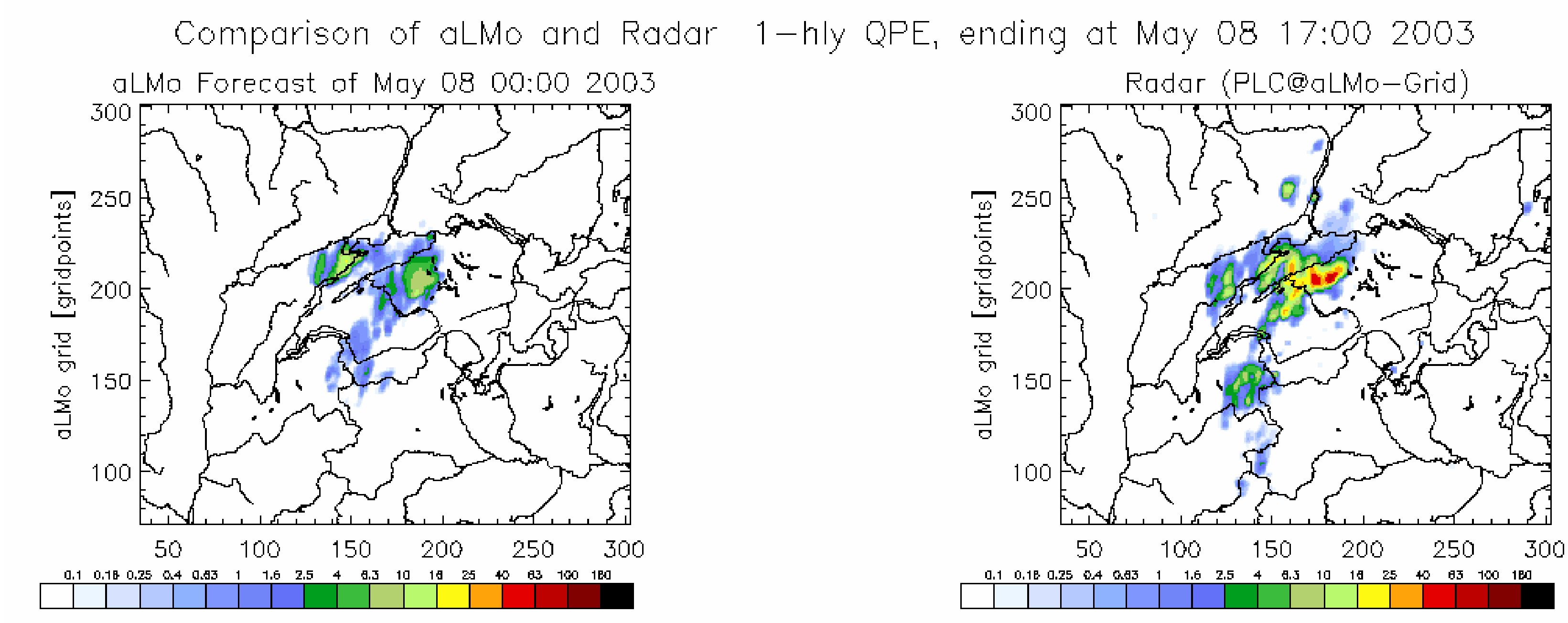
*NO-NF = neighborhood observation-neighborhood forecast,
SO-NF = single observation-neighborhood forecast

Multi-scale, multi-intensity approach

- Forecast performance depends on the scale and intensity of the event



Case study

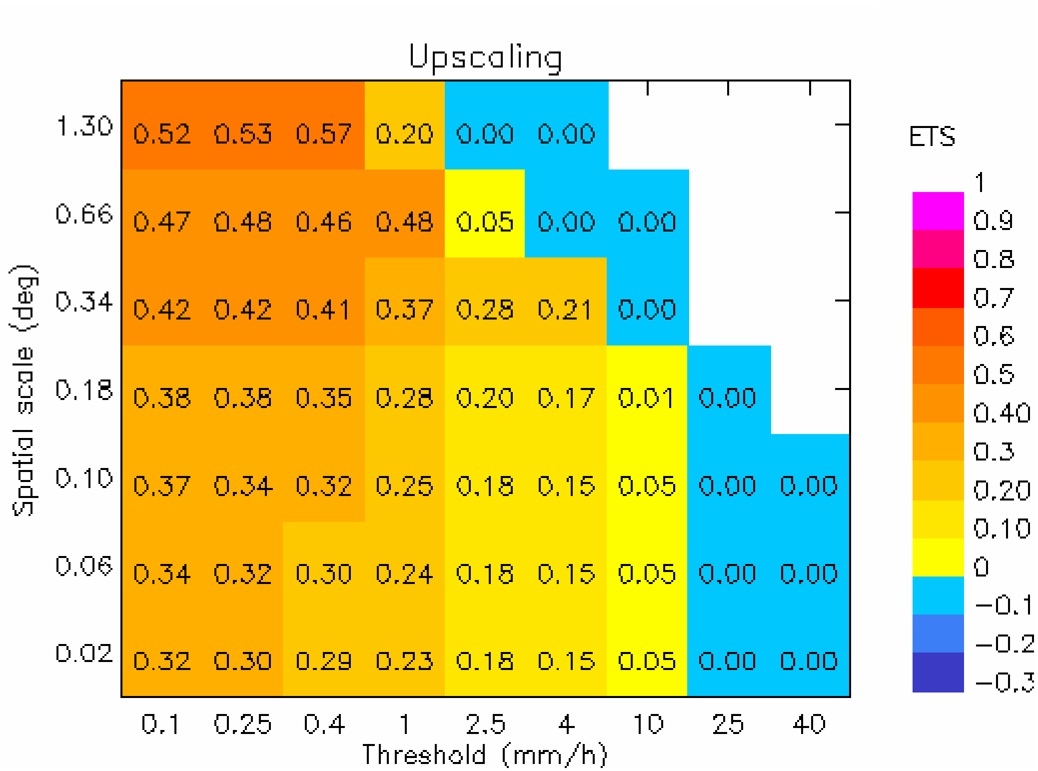


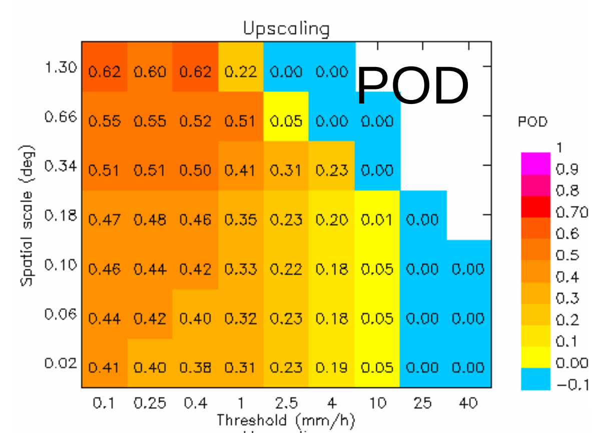
- Verification of 2 km resolution precipitation forecast of 1 hr rainfall in Switzerland using MeteoSwiss Alpine Model (aLMo)

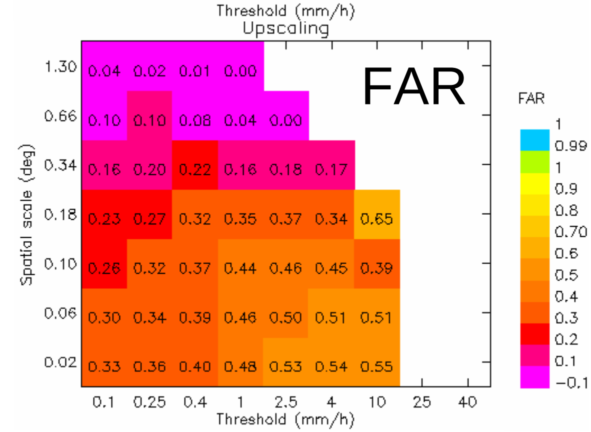
(data courtesy of Daniel Leuenberger, MeteoSwiss)

Upscaling

Decision model – Useful forecast resembles observations when averaged to coarser scales

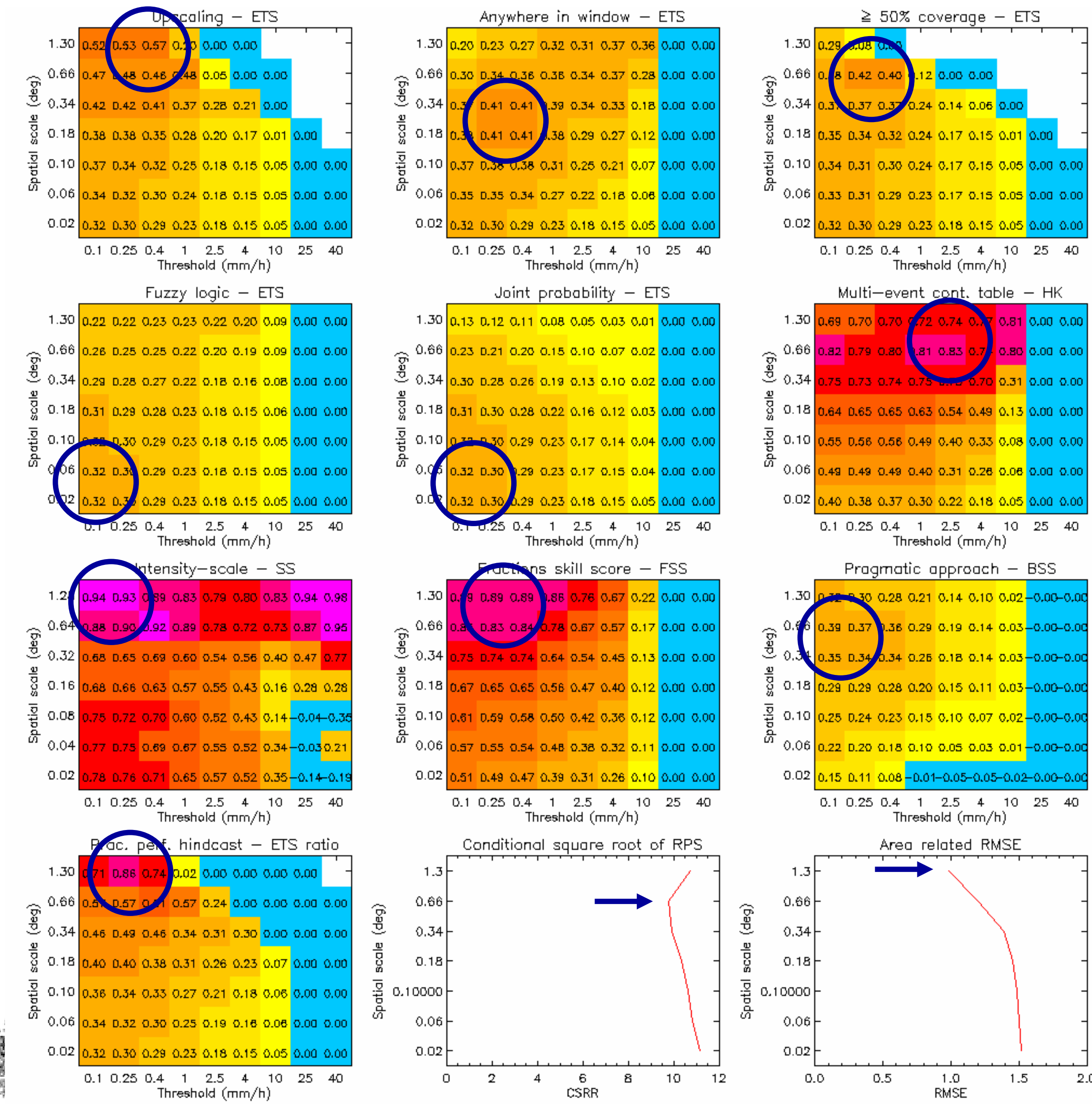




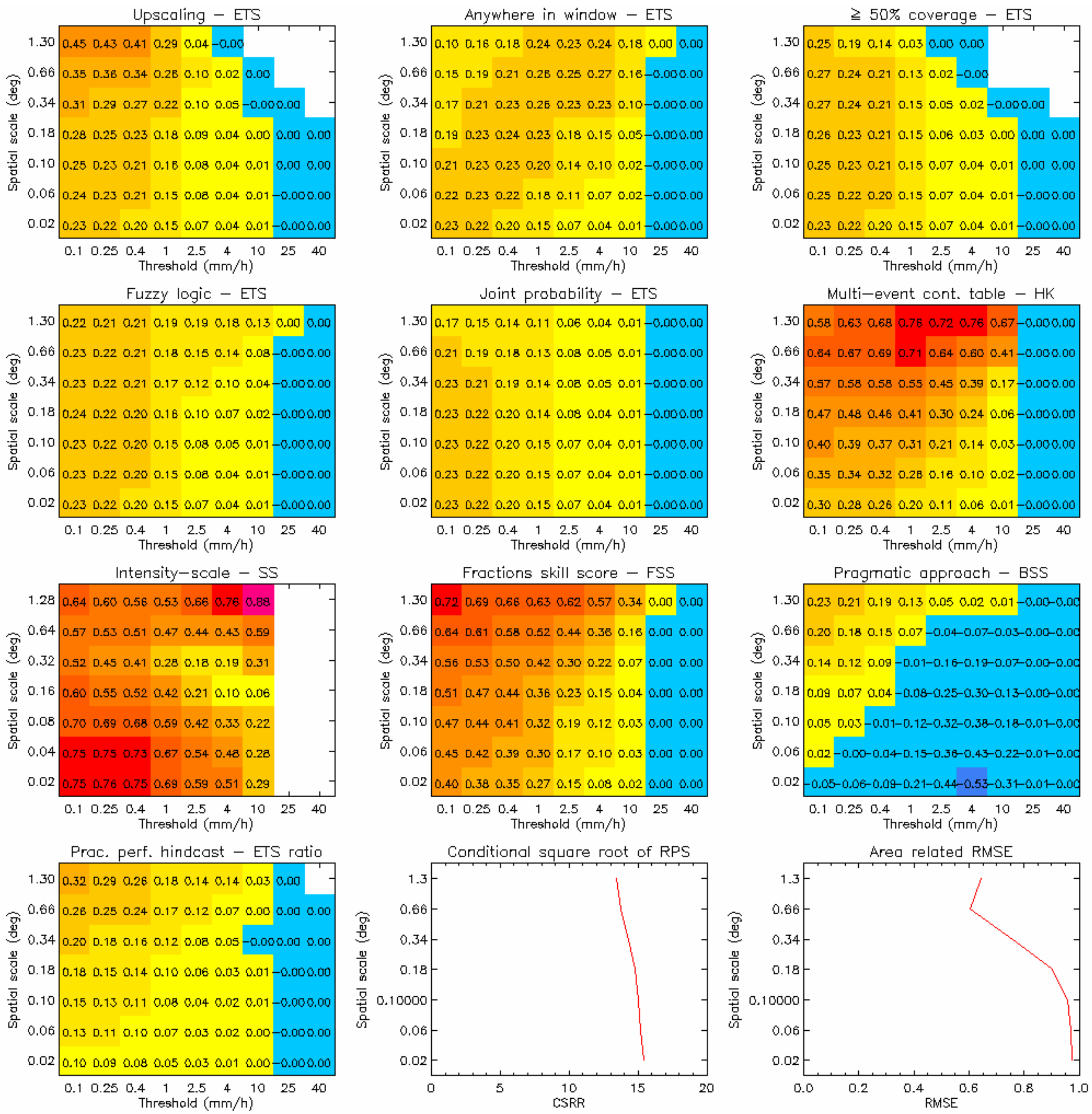


Australian Government
Bureau of Meteorology

Fuzzy verification framework



Aggregate results for 24 h period

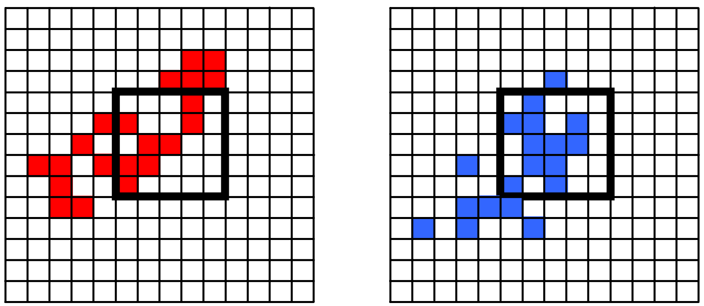


Advantages of fuzzy verification

- Knowing which scales have skill suggests the scales at which the forecast should be presented and trusted
- Suitable for discontinuous fields like precipitation
- Can give good results for forecasts that verify poorly using exact-match approach
- Results match with our intuition
- Can be used to compare forecasts at different resolutions
- Multiple decision models and metrics
  - Direct approach → verification of intensities
  - Categorical approach → verification of binary events
  - Probabilistic approach → verification of event frequency

Many verification possibilities

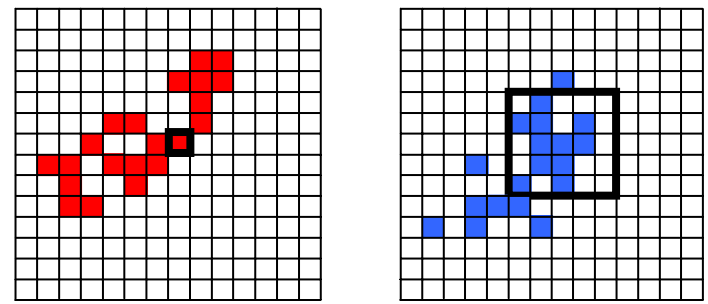
Neighborhood observation – neighborhood forecast (modeler viewpoint)



observation forecast

- categorical scores
  - POD, FAR, ETS, etc.
- probabilistic methods
  - BS, RPS, reliability, ROC, etc.
- continuous scores
  - RMSE, MAE, etc.

Single observation – neighborhood forecast (user viewpoint)



observation forecast

- categorical scores
  - POD, FAR, ETS, etc.
- probabilistic methods
  - BS, RPS, reliability, ROC, relative value, etc.

Thank you!