Scale selective verification of precipitation forecasts

Nigel Roberts

Background

Operational 1.5 km model by end of decade

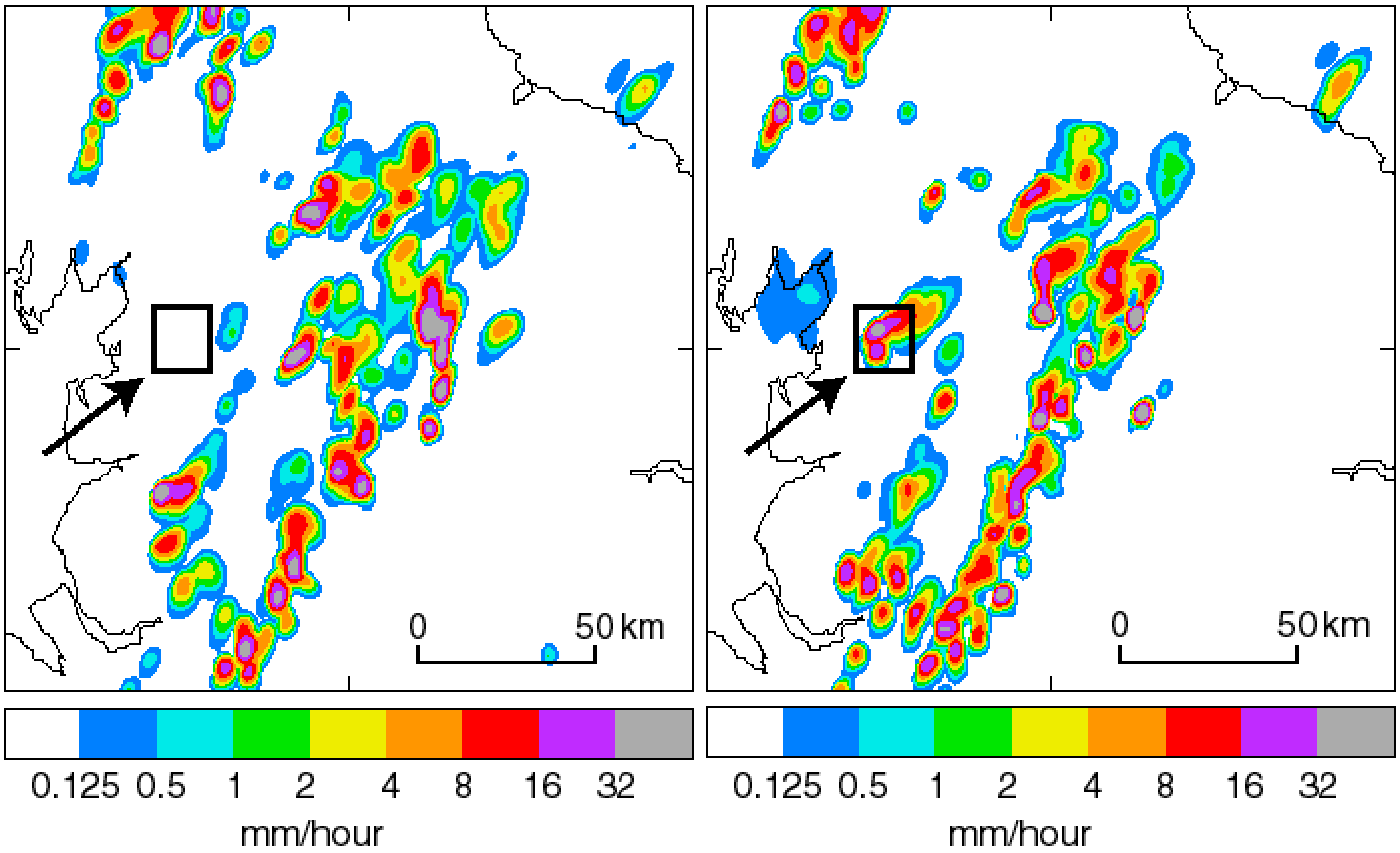
Higher-resolution precipitation forecasts look more realistic, but are computationally expensive. Are they more accurate?

Traditional point verification is inappropriate. We shouldn't believe the small scales.

Can a ~1-km model provide more accurate rainfall forecasts on useful scales (e.g. over river catchments)?

How should we interpret and present the model output.

Forecast divergence



High-resolution model predictability

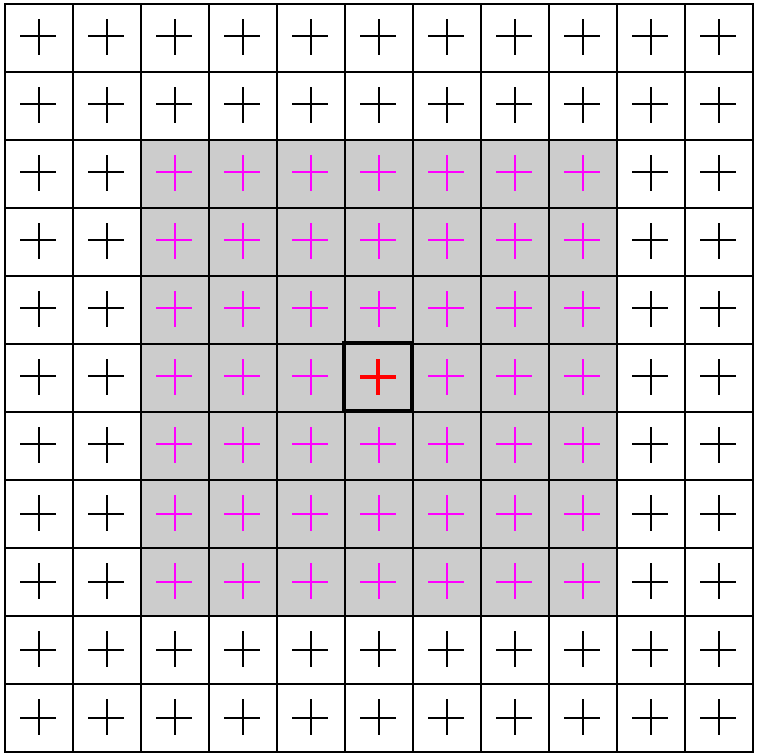
1. We can never get the initial conditions exactly right (worse at higher resolution) and forecast errors grow
2. Finer resolution introduces faster growing errors at the smaller scales. Small scales become unpredictable quickly. (Lorenz 1969, Zhang et al 2003)

- Each forecast is one possible solution from a pdf of alternatives
- We must use probabilities (or other indicators of uncertainty) to deal with unpredictable scales when presenting high-resolution precipitation forecasts. But what scales are unpredictable?

Probabilities from nearest neighbours

Probabilities come from fractions of occurrences within neighbourhoods

What neighbourhood size should we use?



Theis et al 2005

RUC model - Weygandt and Benjamin, 2005

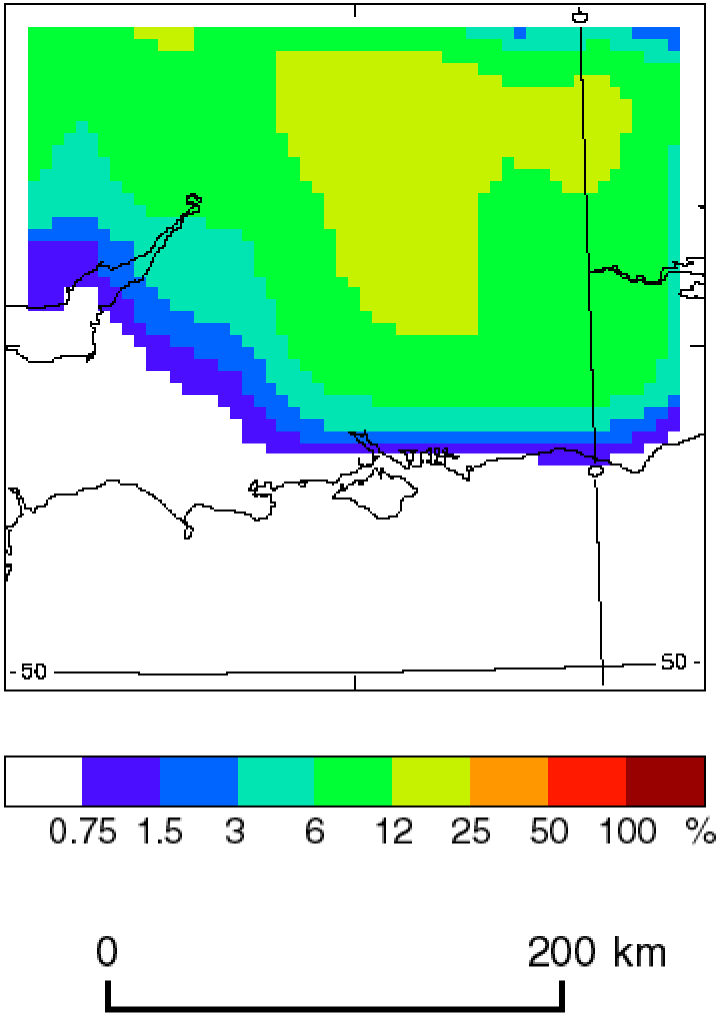
Nimrod nowcasting system, UK Met Office

Example of neighbourhood probabilities/fractions

Probabilities/fractions from a square neighbourhood of length 125 km

More appropriate for a much less accurate forecast system.

Much smoother – this would not be a cost effective use of a 1-km model.



Verification approach

We want to know

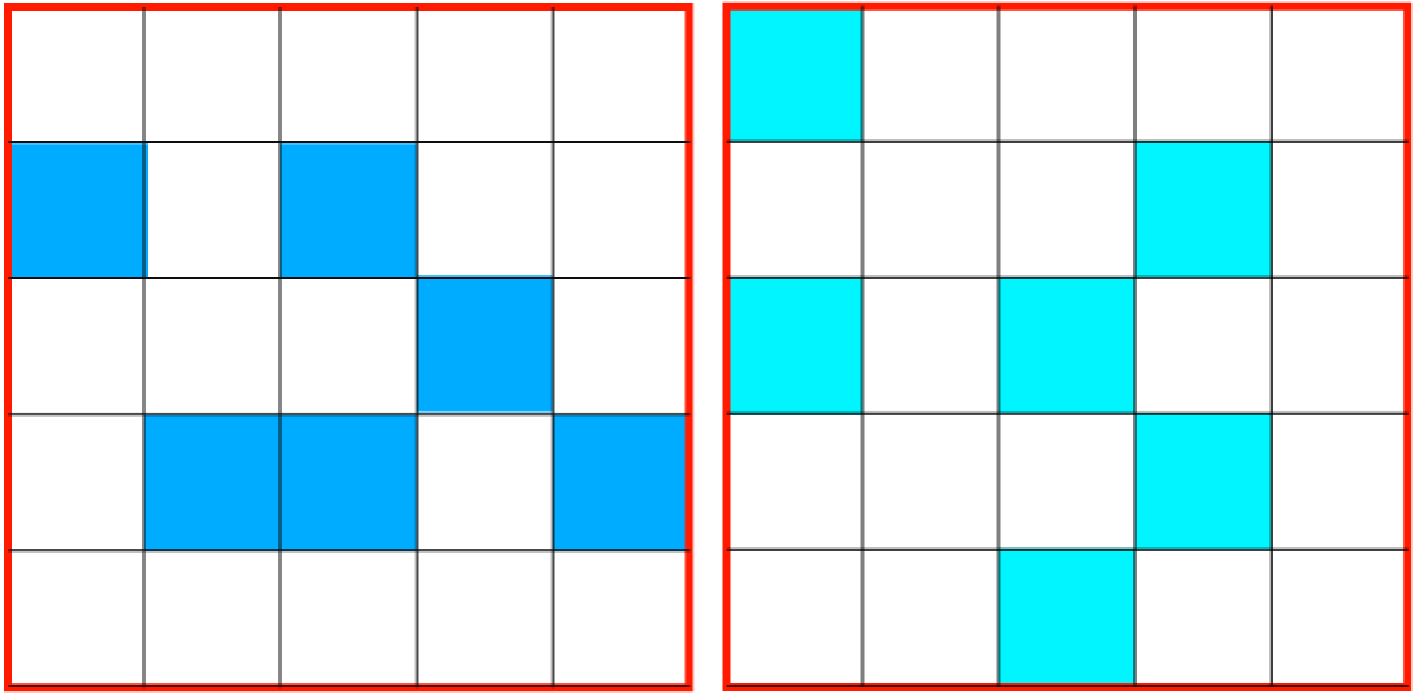
1. How forecast skill varies with neighbourhood size.
2. The smallest neighbourhood size that can be can be used to give sufficiently accurate forecasts.
3. Does higher resolution provide more accurate forecasts on scales of interest (e.g. river catchments)

Compare forecast fractions with fractions from radar over different sized neighbourhoods (squares for convenience)

Use rainfall accumulations to apply temporal smoothing

Schematic comparison of fractions

observed forecast



Fraction = 6/25 = 0.24 Fraction = 6/25 = 0.24

Threshold exceeded where squares are blue

A score for comparing fractions with fractions

Brier score for comparing fractions

$$\text{FBS} = \frac{1}{N}\sum_{j=1}^{N}(p_j - o_j)^2$$

FBS (Fractions Brier Score)

$0 \leqslant p_j \leqslant 1$ forecast fractions

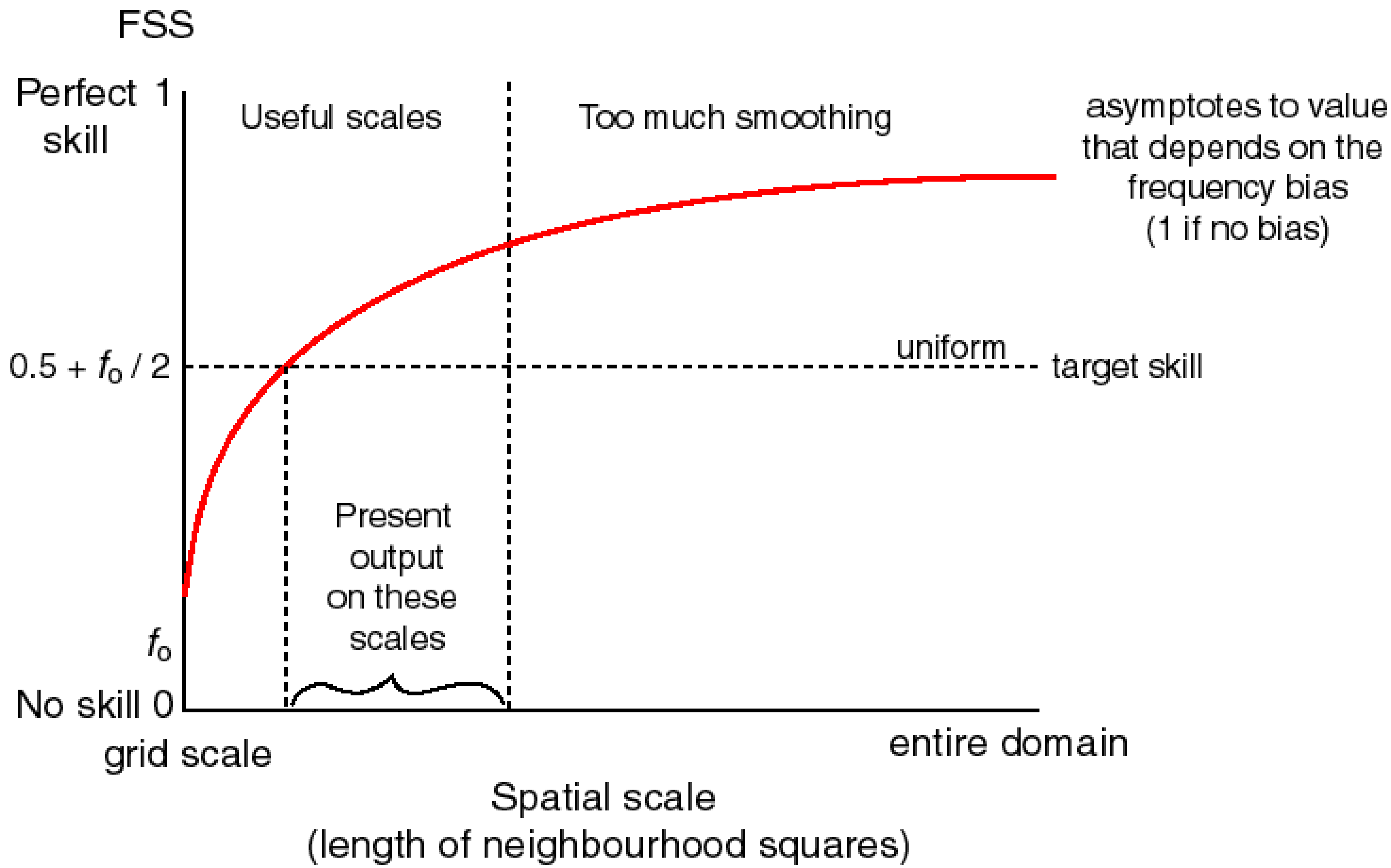
$0 \leqslant o_j \leqslant 1$ radar fractions

N number of points

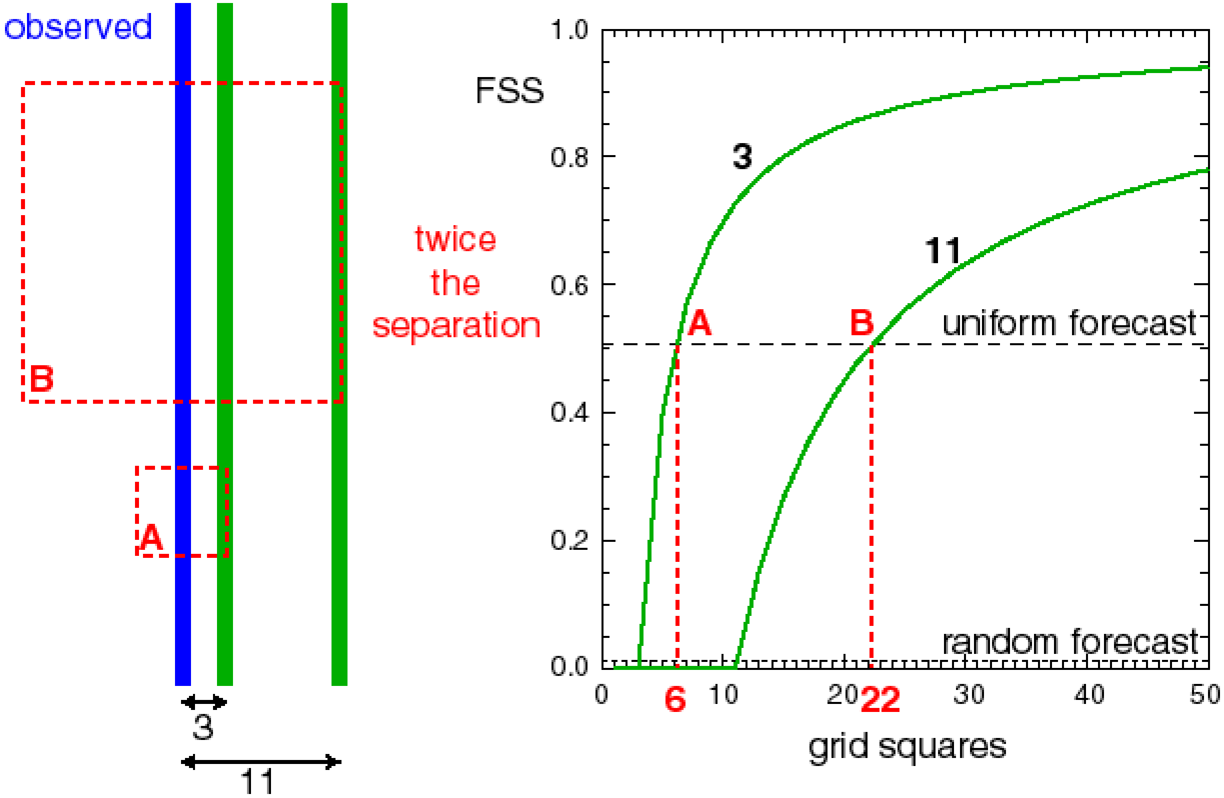
Skill score for fractions/probabilities - Fractions Skill Score (FSS)

$$\text{FSS} = 1 - \frac{\text{FBS}}{\frac{1}{N}\left[\sum_{j=1}^{N}(p_j)^2 + \sum_{j=1}^{N}(o_j)^2\right]}$$

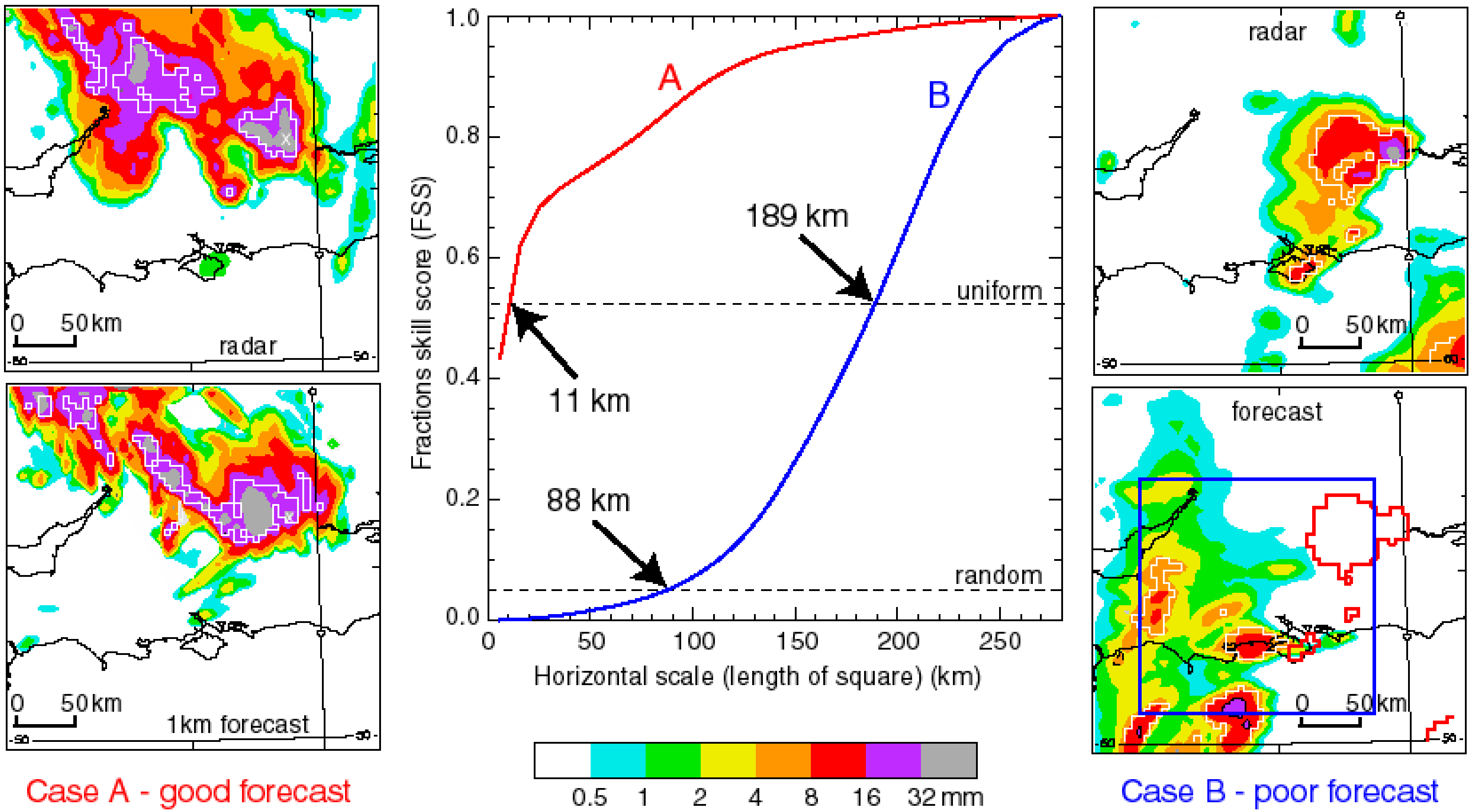
Example graph of FSS against neighbourhood size



Idealised example



Real examples



Case A - good forecast

Case B - poor forecast

Conclusions

This verification method has provided a way of answering **some** important questions about forecasts from ‘storm-resolving’ NWP models.

How does forecast skill vary with spatial scale?

At what scales are higher resolution forecasts more skilful (if any)?

At what scales are forecasts sufficiently accurate?...........

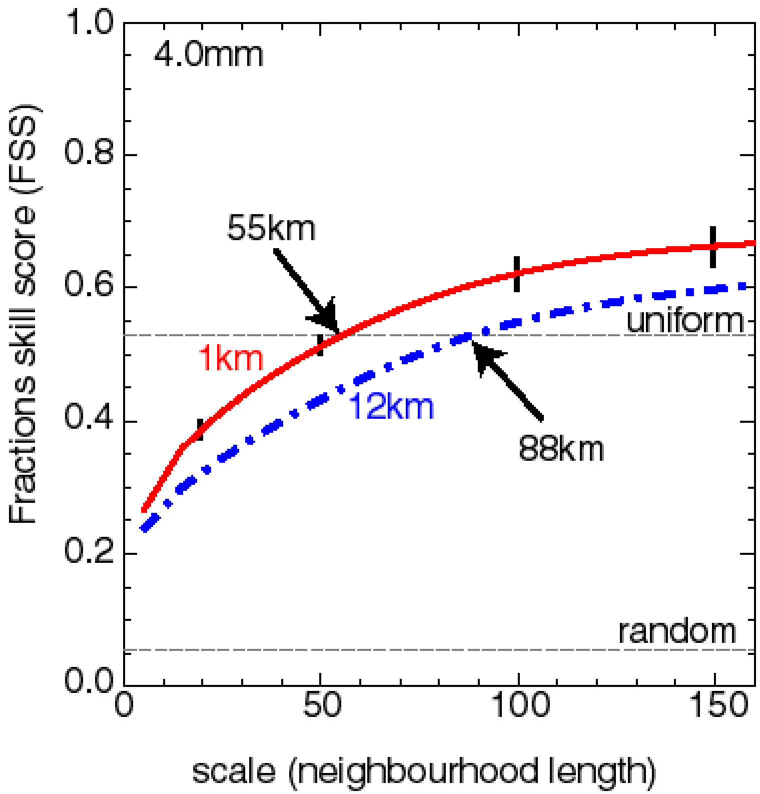
(There are other questions that need different approaches)

Directly relates forecast skill to forecast presentation – adaptive probabilistic forecasting is a possibility

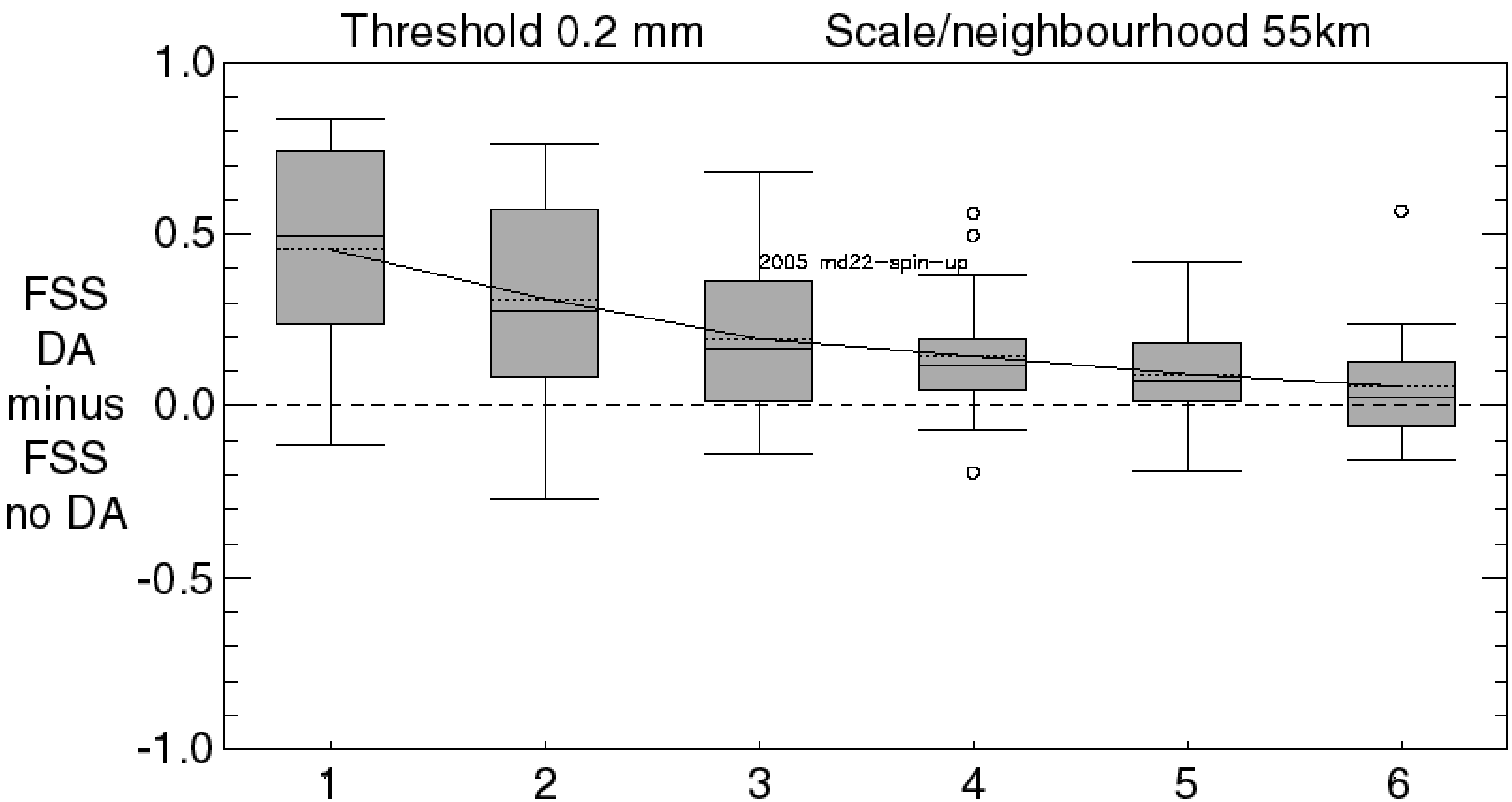
Thanks

4-hour accumulations T+2 to T+6

From a sample of 40 forecasts of convective rainfall events



Impact of data assimilation on hourly accumulations



Courtesy of Mark Dixon