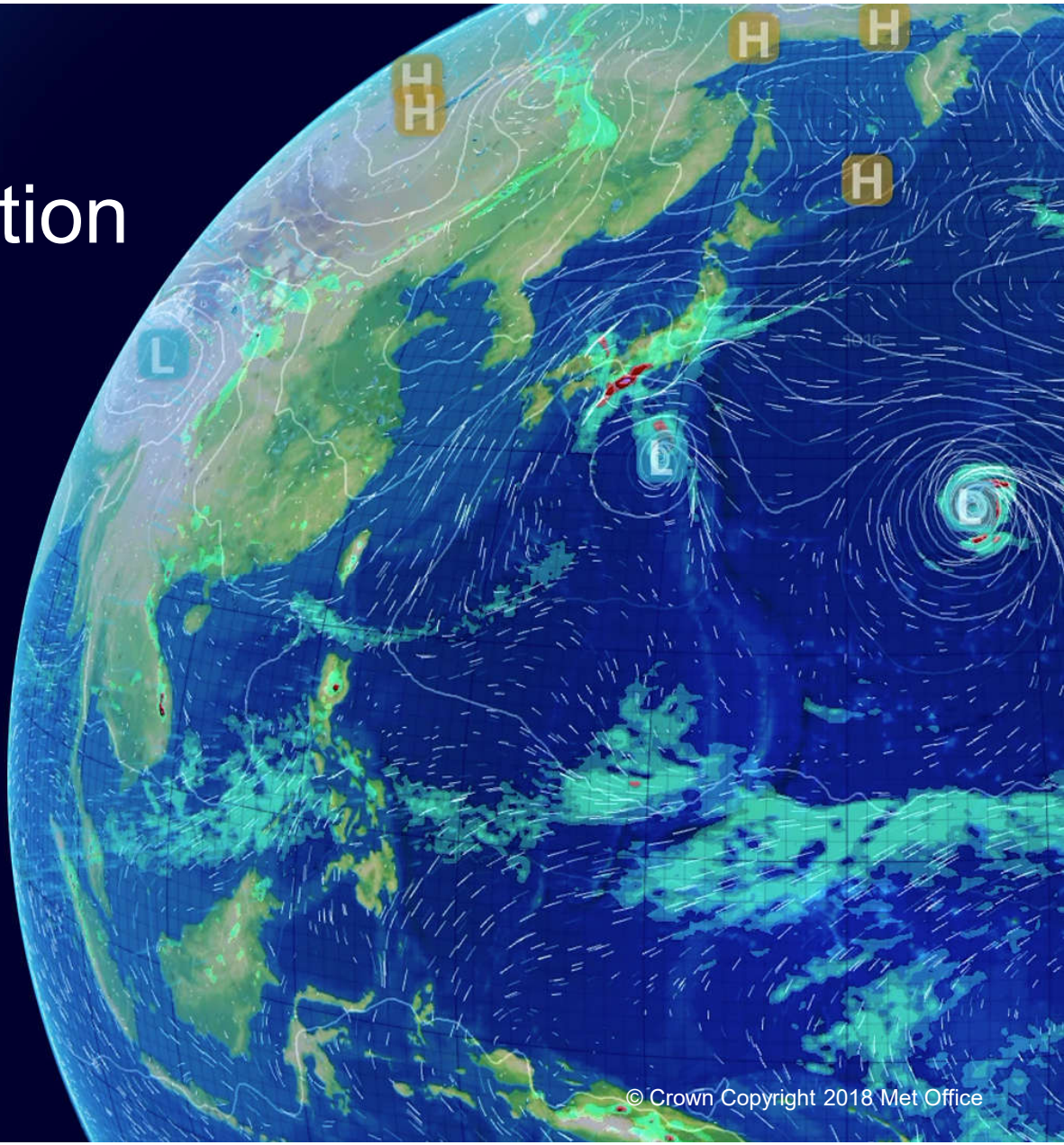
Predictability and verification of precipitation forecasts

Nigel Roberts



Predictability – what is it?

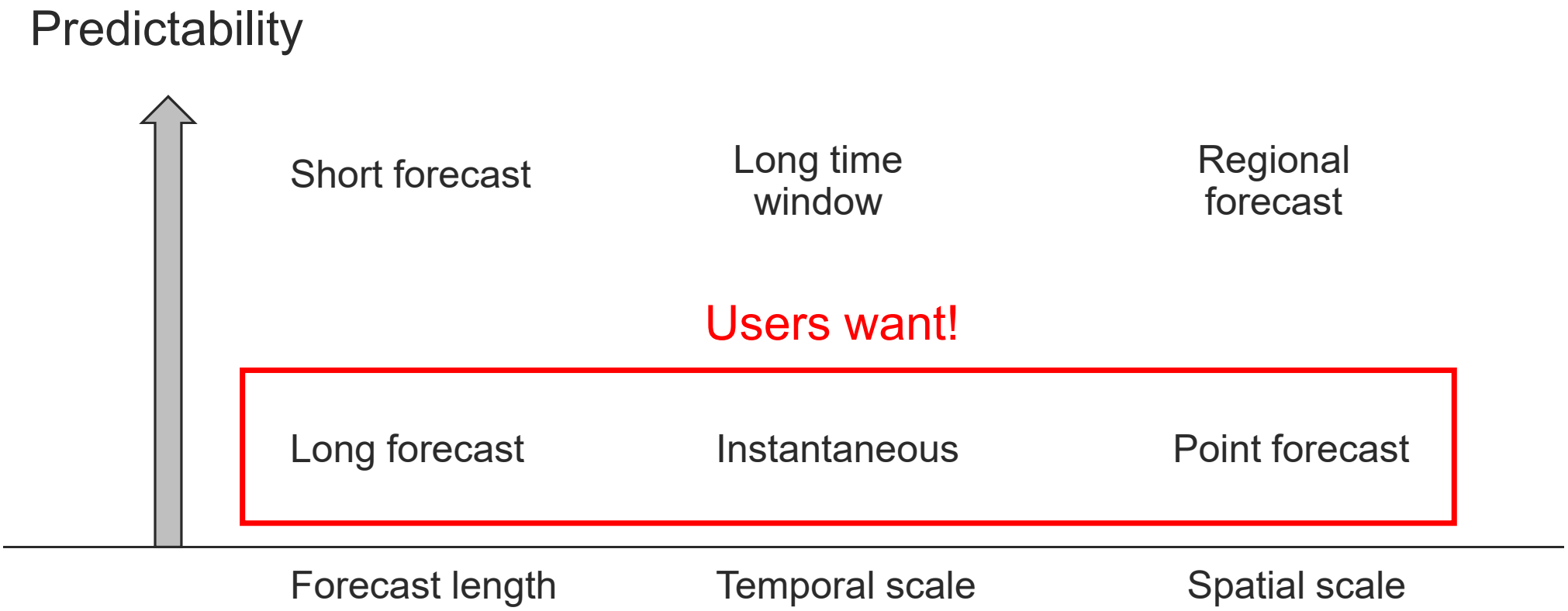
The ability to forecast something with confidence it will happen
(low certainty = low predictability)

Weather forecast predictability depends on:

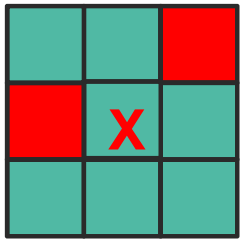
1. The phenomena – e.g. Icelandic low more predictable than a thunderstorm – inherent atmospheric uncertainty (volatility)
2. The model (or ensemble) and accuracy of initial state
3. Expectation – what does “something” mean exactly?

The expectation dilemma

Should we give users the bad forecasts they want or the good forecasts they don't want?



Probabilistic forecasts



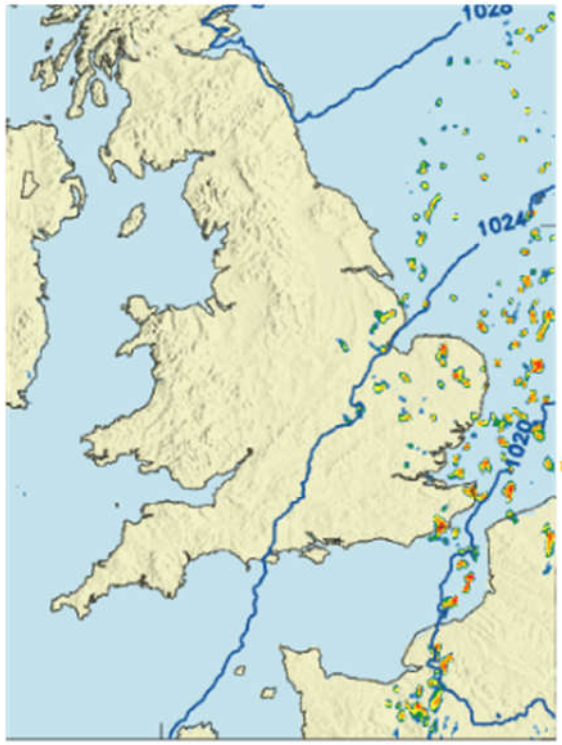
Neighbourhood processing or ensemble (or both)

More appropriate forecast.

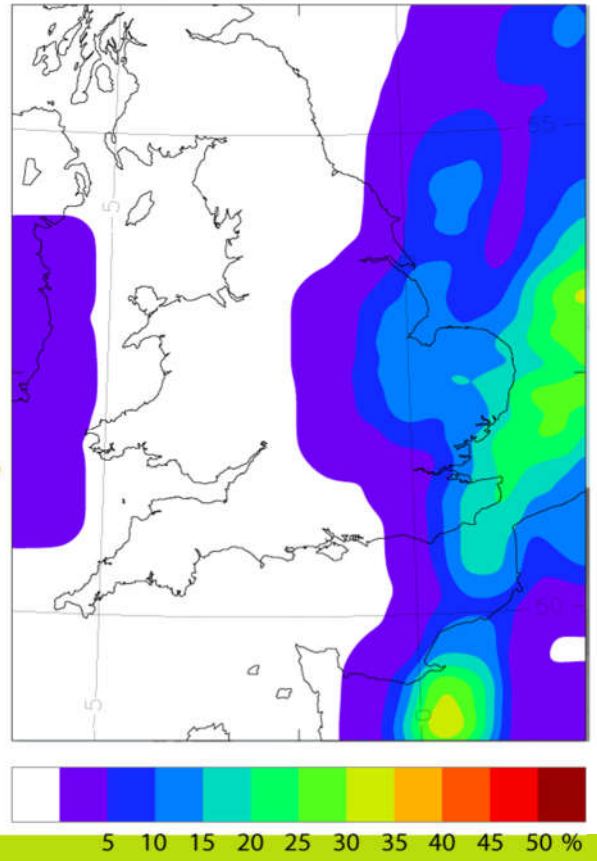
But what is the optimal neighbourhood or ensemble size?

Localised/extreme rain will give low probabilities. Is that useful?

UKV model forecast (1.5km)



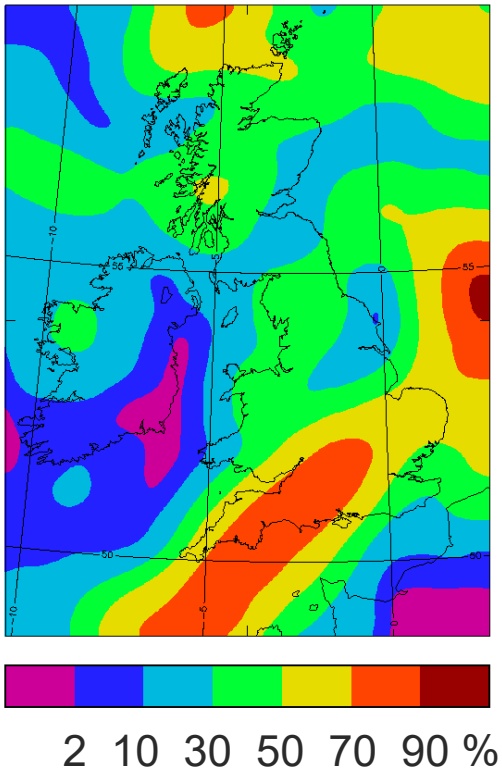
Radar for the same time



A compromise - upscaling in space or time

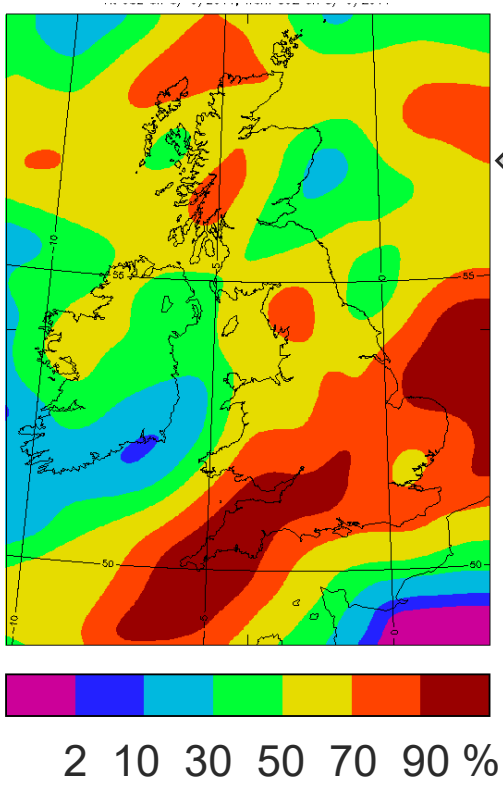
Probability

Rain at a particular time



Probability

Rain in a time window (or in the local area)



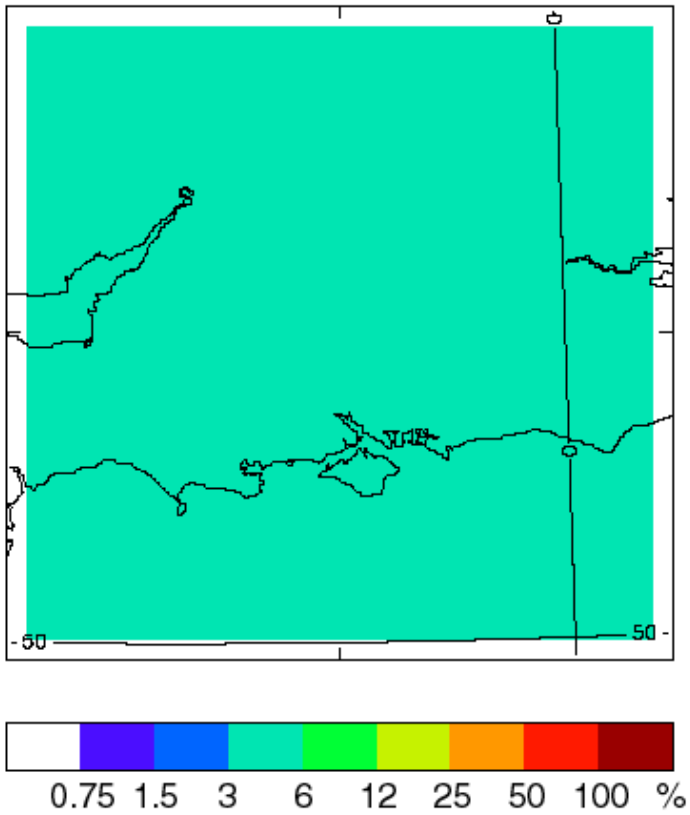
Should give better predictability (so long as the forecast system is capable of capturing the physical processes and uncertainty)

Different from standard neighbourhood processing!

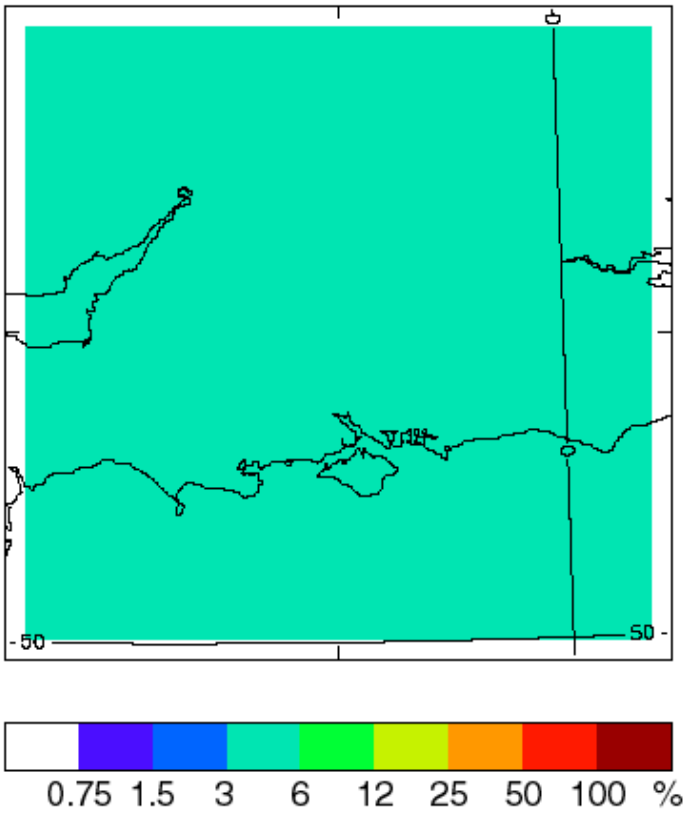
Schwartz, C.S. and R.A. Sobash, 2017: Generating Probabilistic Forecasts from Convection-Allowing Ensembles Using Neighborhood Approaches: A Review and Recommendations. *Mon. Wea. Rev.*, **145**, 3397–3418

At what scale does a forecast have skill?

Radar



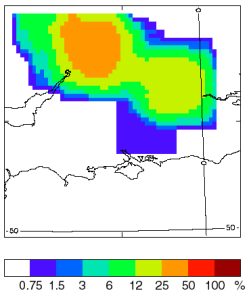
Forecast

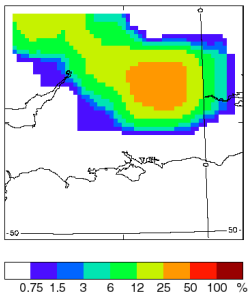


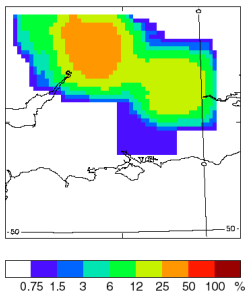
300 X 300 km

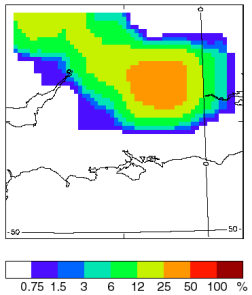
Fractions Skill Score (FSS)

$$FSS = 1 - \left\{ \frac{\text{Sum of } \left\{ \text{obs} - \text{fcst} \right\}^2}{\text{Sum of } \left\{ \text{obs} \right\}^2 + \text{Sum of } \left\{ \text{fcst} \right\}^2} \right\}$$









Fractions Skill Score (FSS)

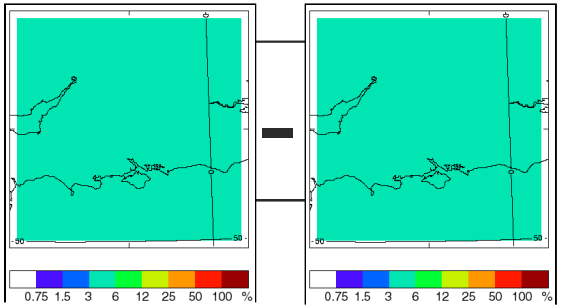
$$FSS = 1 - \frac{\sum_{j=1}^{N} (p_j - o_j)^2}{\left[\sum_{j=1}^{N} (p_j)^2 + \sum_{j=1}^{N} (o_j)^2\right]}$$

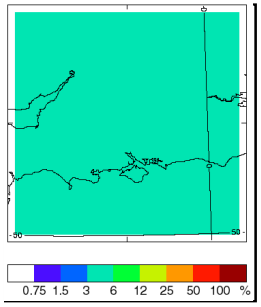
$0 \leq p_j \leq 1$ forecast fractions

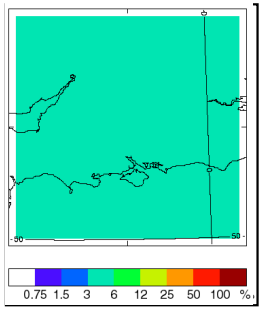
$0 \leq o_j \leq 1$ radar fractions

N number of points

$$FSS = 1 - \left\{ \frac{\text{Sum of } \left\{ \text{fcst} - \text{obs} \right\}^2}{\text{Sum of } \left\{ \text{fcst} \right\}^2 + \text{Sum of } \left\{ \text{obs} \right\}^2} \right\}$$

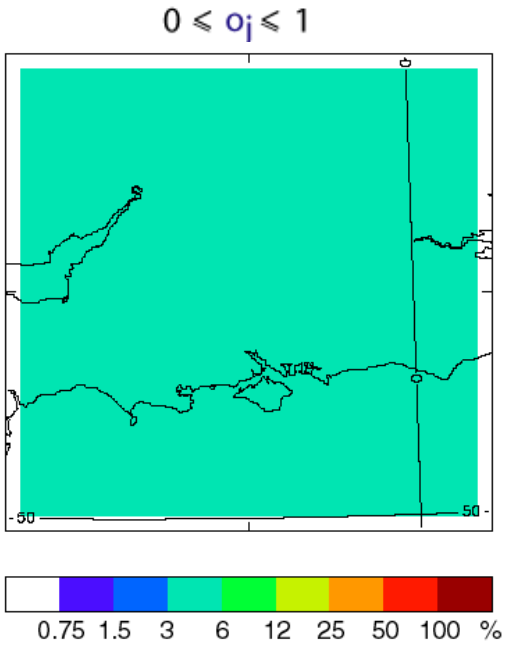




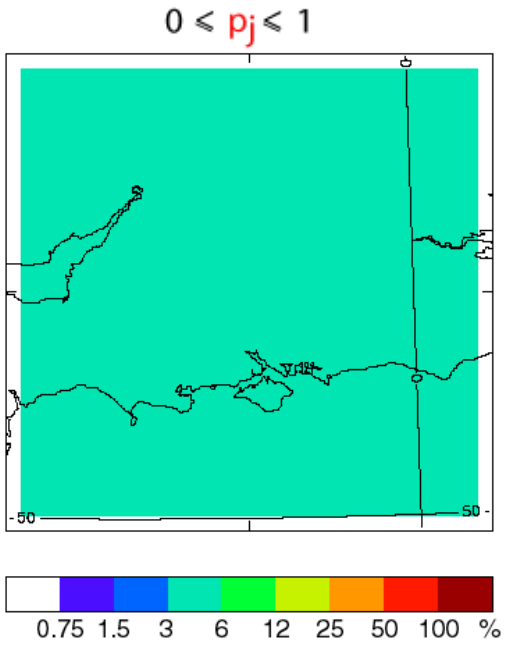


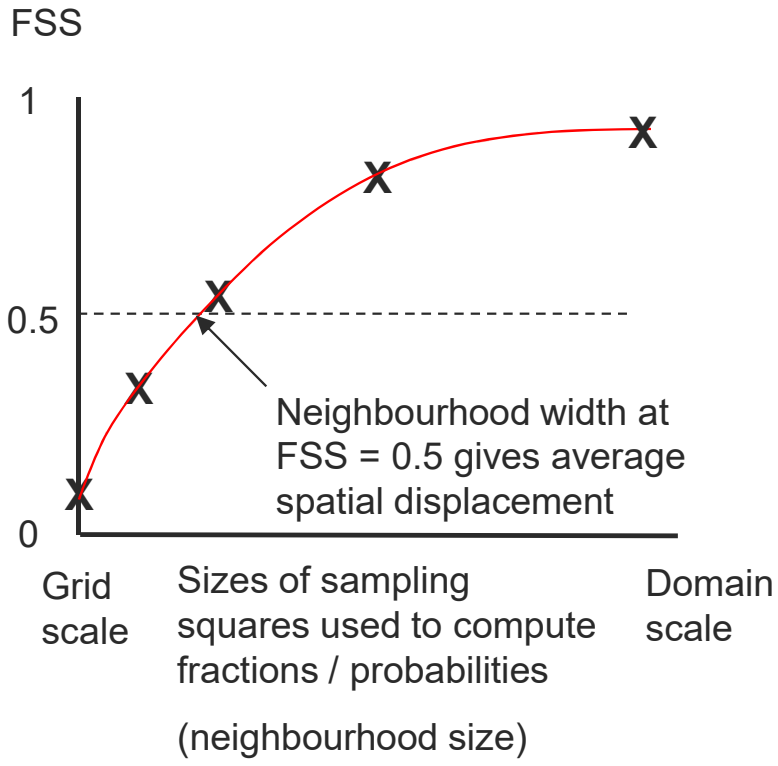
Behaviour of FSS with increasing neighbourhood

Observed fractions O



Forecast fractions P

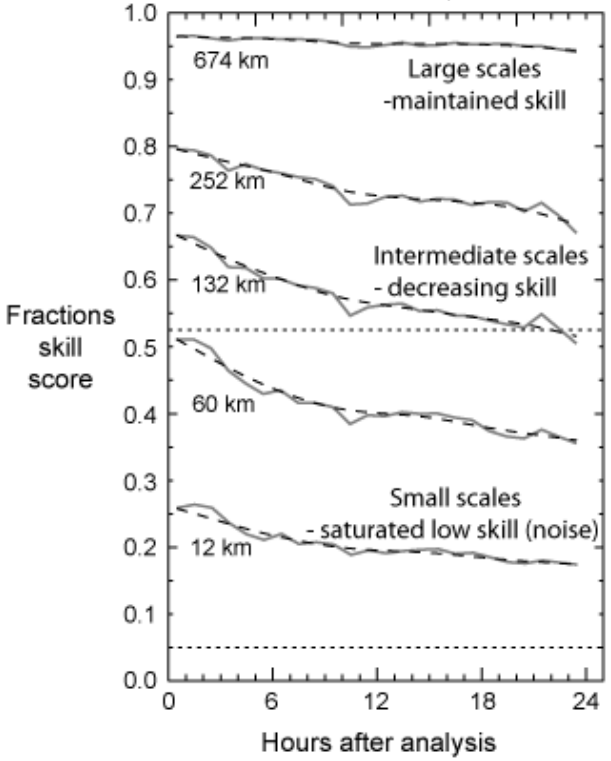


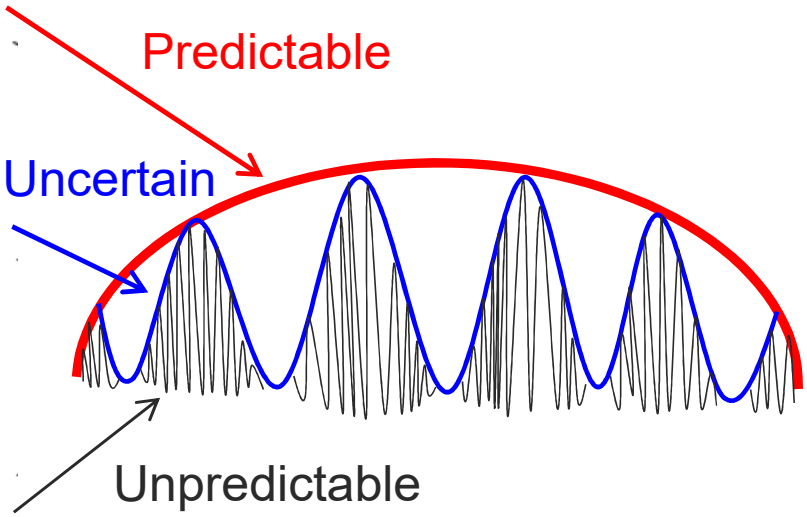


Roberts NM and Lean HW (2008)

Predictable scales

12km model forecast hourly accumulations





~ 600-1000 km

Unpredictable scales
- Small probabilities
- Poor resolution
- Even if good reliability

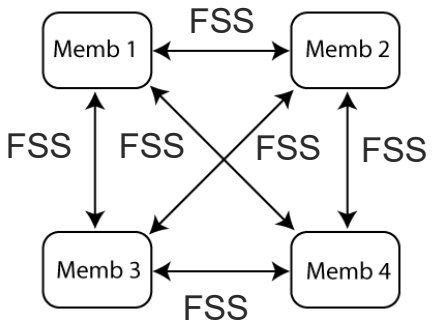
Predictable scales (large synoptic) – no need for an ensemble

Uncertain scales (mesoscale) – ensemble needed

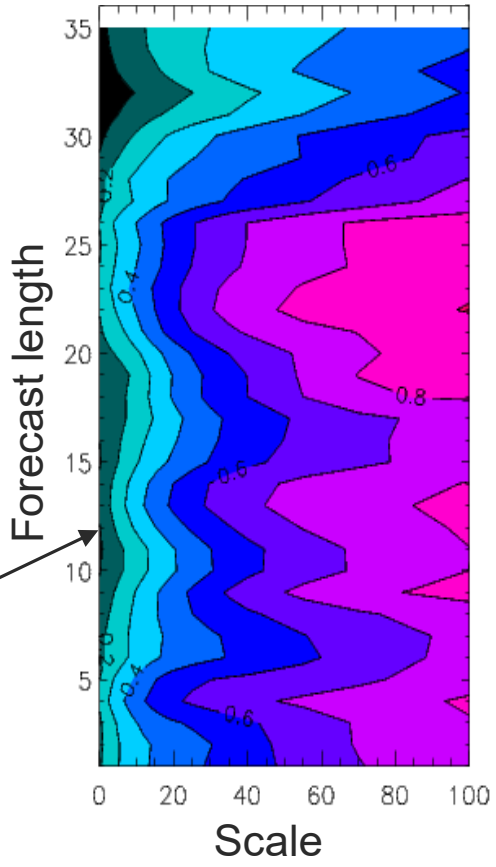
Unpredictable scales (individual showers) – large ensemble needed

Roberts N (2008)

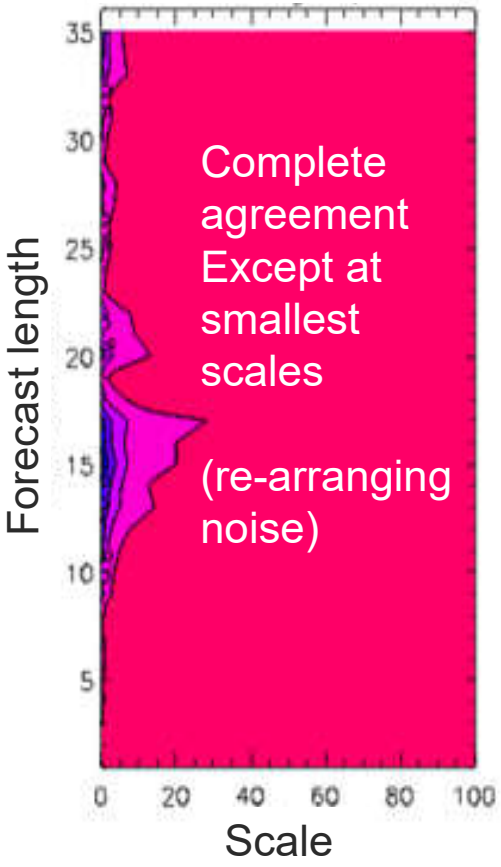
FSS for an ensemble

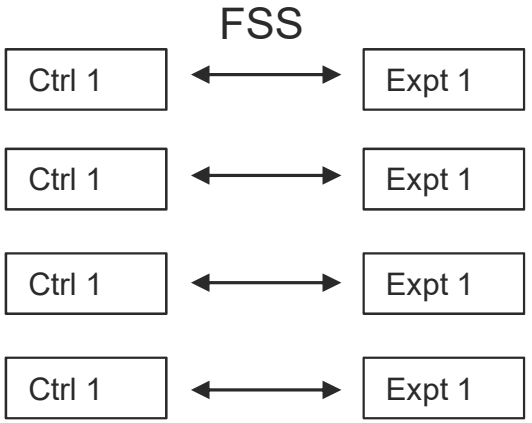


Ens mean FSS



Ctrl-Expt FSS





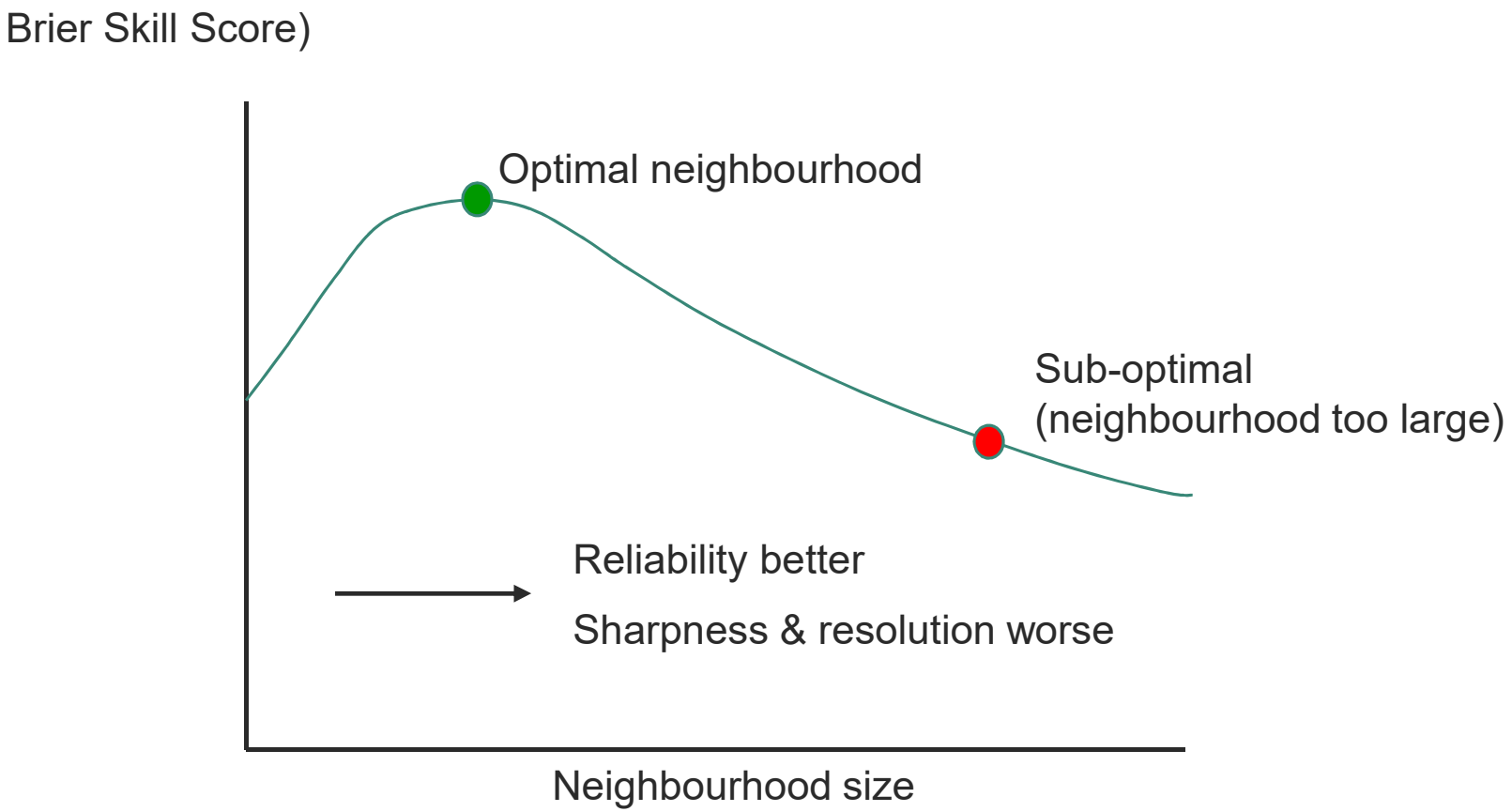
Little agreement at small scales
- low predictability

1 = ‘perfect match’

0 = ‘totally different’

Contours every 0.1, colours **black at 0.0** to **red at 1.0**

What if we only have point verification (no radar) ?



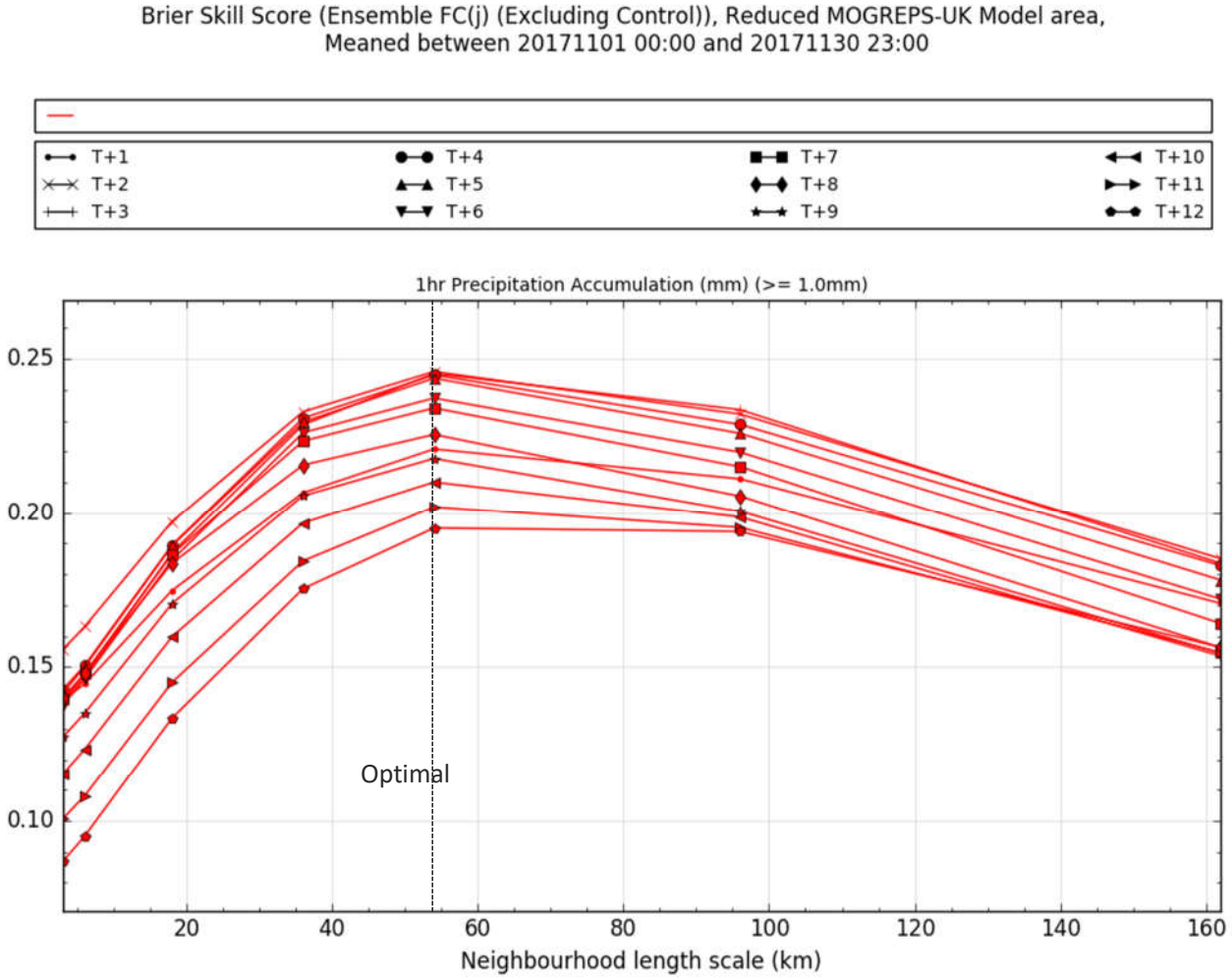
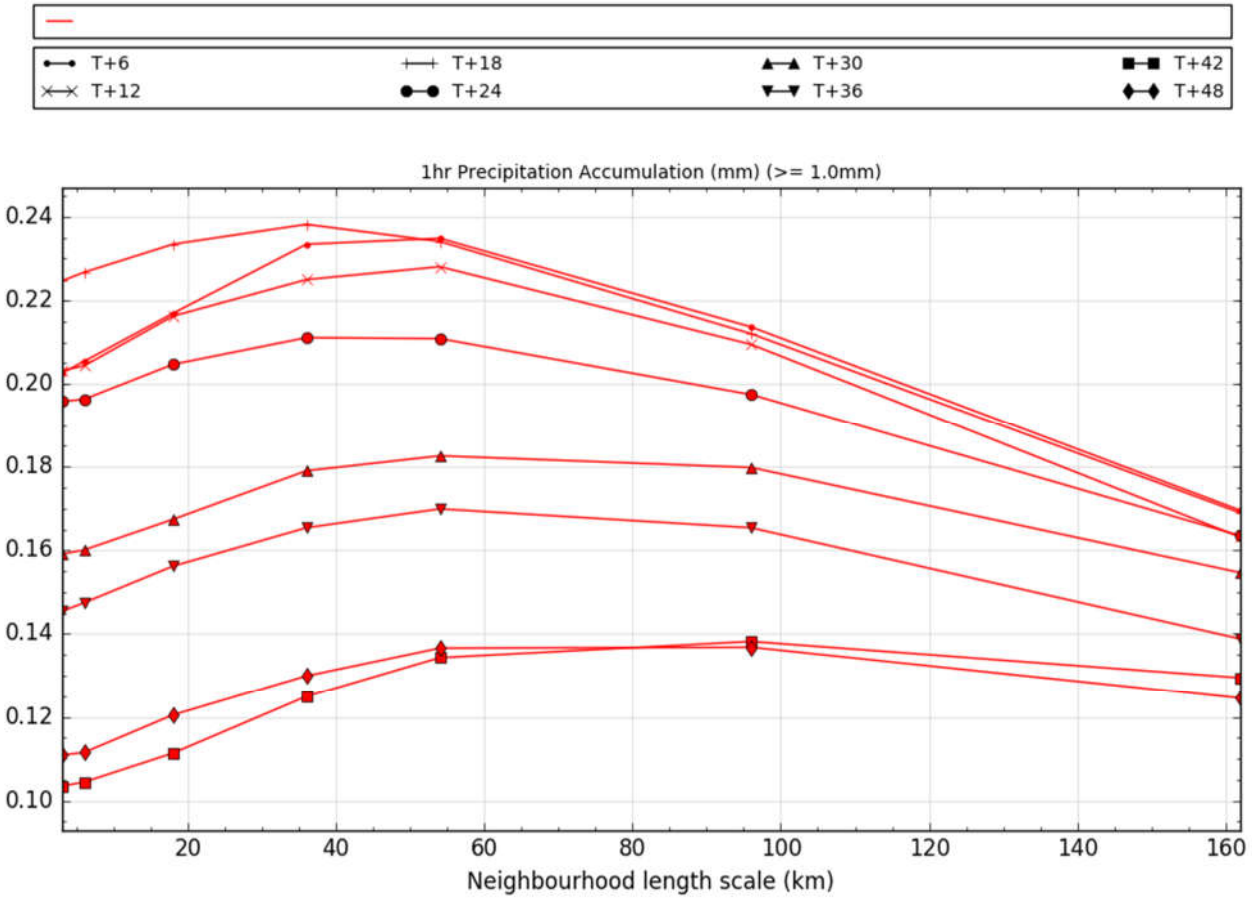


Figure 18 (a): BSS against neighbourhood radius for 1 hour precipitation accumulation with UKV, for a threshold of 1 mm.

Met Office

Brier Skill Score (Ensemble FC(j) (Excluding Control)), Reduced MOGREPS-UK Model area,
Meaned between 20171101 00:00 and 20171130 23:00



Can also use the FSS to find domain average displacement directly

displacement $d_{FSS} = (1 - FSS_{gridscale})\ L(1\text{-}FSS_L)$

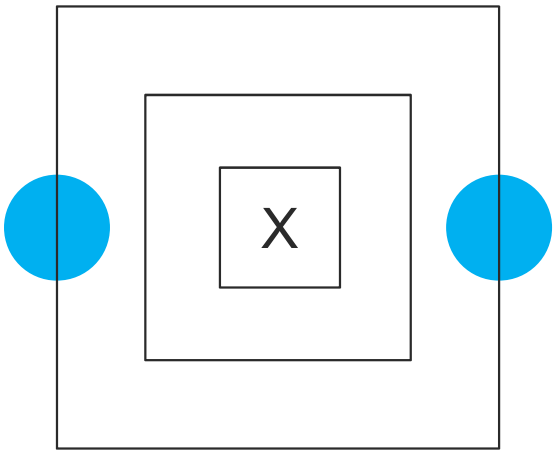
L = neighbourhood length for FSS_L ~0.5

Skok, G. (2015) “Analysis of Fraction Skill Score properties for a displaced rain band in a rectangular domain”, Met Apps 22 477-484 (2015)

Skok, G. and Roberts, N. (2018), Estimating the displacement in precipitation forecasts using the Fractions Skill Score. Q.J.R. Meteorol. Soc., 144: 414-425. doi:10.1002/qj.3212

A method for finding the local spatial agreement?

Rain in member 1



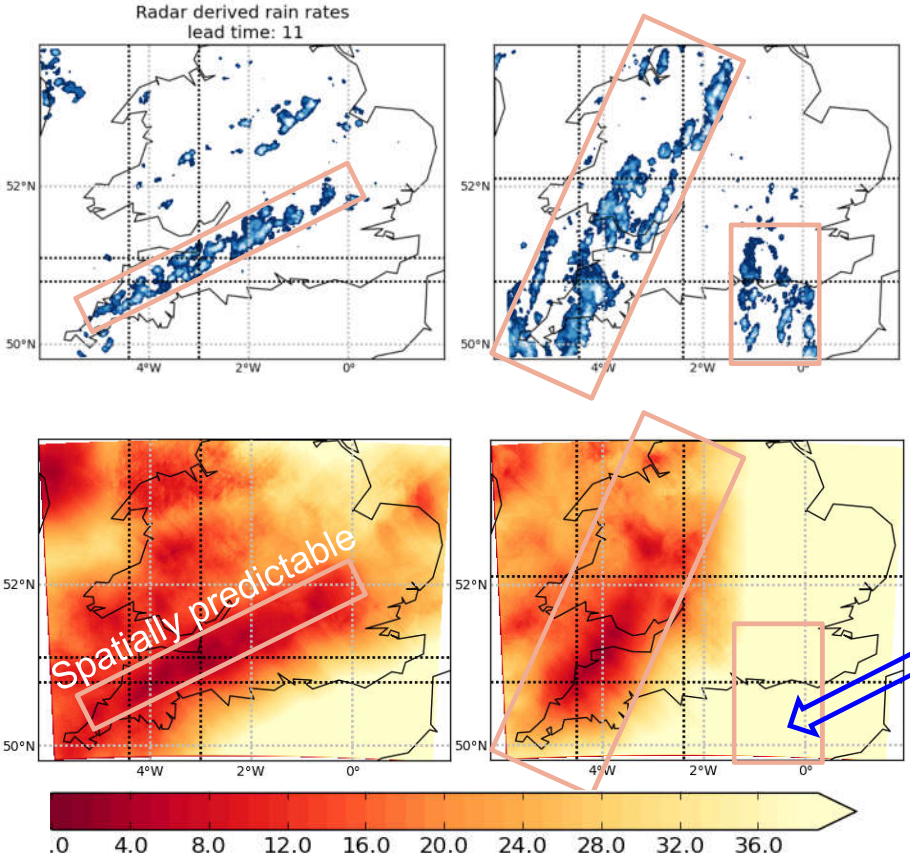
Rain in member 2

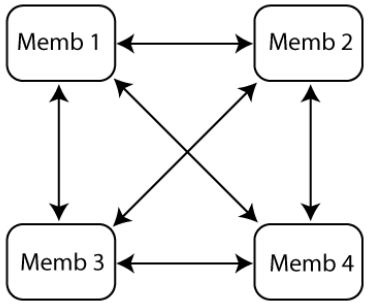
Consider different sized neighbourhood squares centred at point X

Do this for every grid square

FSS-type approach, see:
Dey et al, 2016, QJ “A new method for the characterization and verification of local spatial predictability for convective-scale ensembles.”

Use ensemble to give information about spatial uncertainty and verify spatially





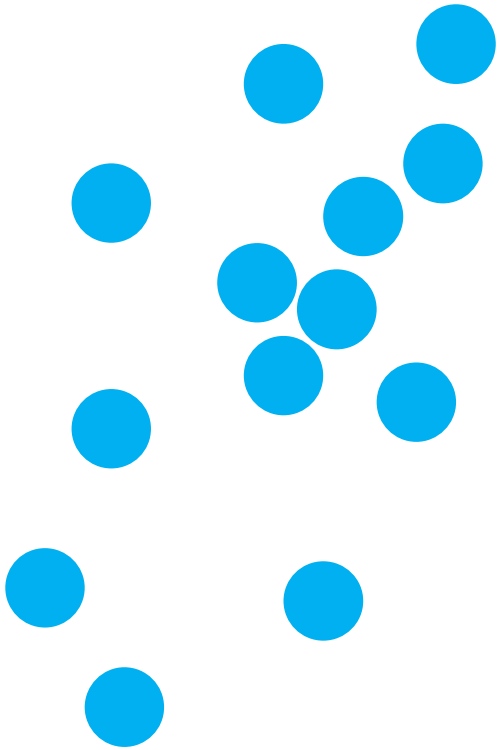
Not predictable (not forecast correctly on this occasion)

Small uncertainty — Neighbourhood size — Large uncertainty

Dey et al, 2016, QJ “A new method for the characterization and verification of local spatial predictability for convective-scale ensembles.”

Met Office

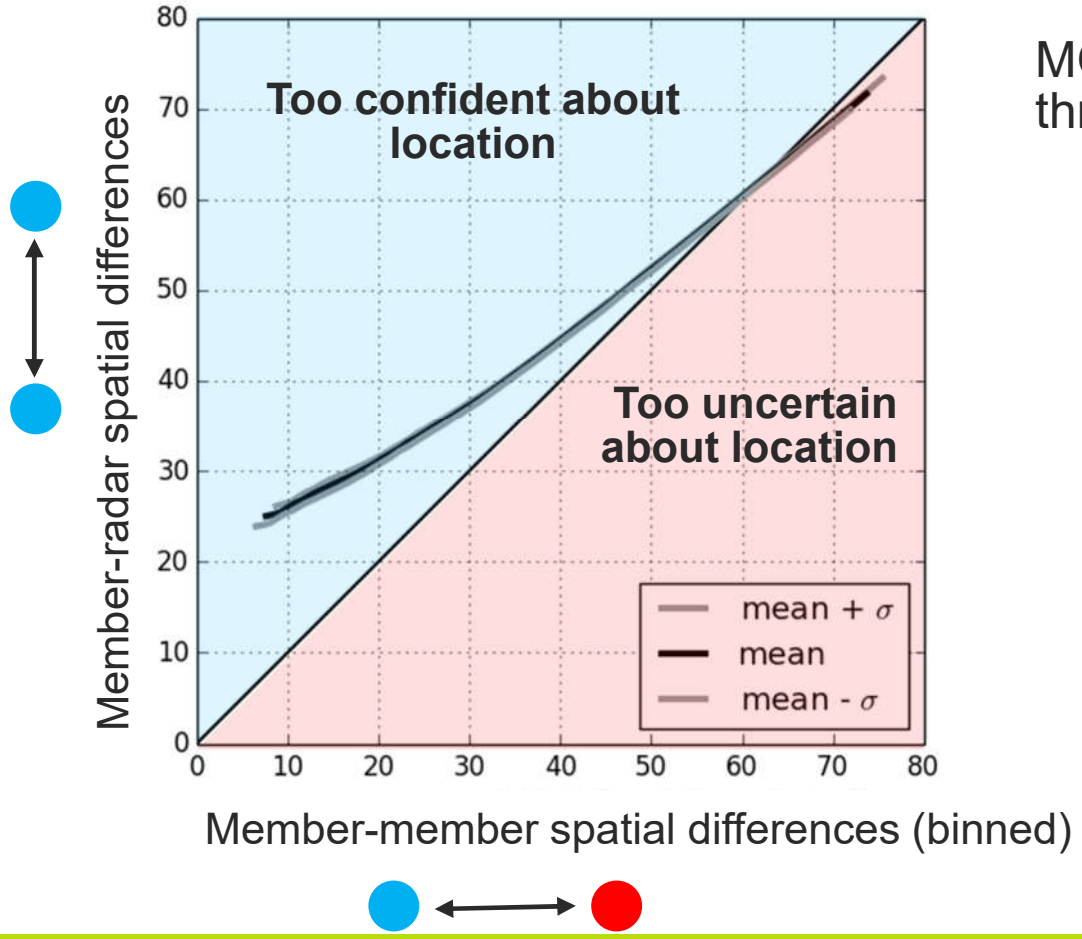
Does the ensemble indicate the true uncertainty?
- the skill-spread relationship
- think spatial!



Each blue circle represents rain from a different ensemble member.

The red circle represents the actual rain

Quantitative measure of ensemble spatial skill-spread



MOGREPS-UK, hourly instantaneous rain rates, three months of data (June, July, August 2013)

- Overall doing a reasonable job
- Somewhat too confident about where rain will occur
- Useful tool for evaluating spatial predictability from ensemble

Dey et al, 2016, QJ “Assessing spatial precipitation uncertainties in a convective-scale ensemble”

Final comments

Predictability is to do with what we try to forecast (spatial/temporal scales)

We need to forecast using probabilities with spatial/temporal filtering

We want to identify the smallest spatial/temporal scales that have useful skill

Use spatial measures like the FSS (or local FSS) or others to determine spatial skill, spatial ensemble spread or spatial skill-spread

Be wary of verifying at the grid scale if little predictability

Thanks for listening

Met Office

Expect the unexpected