New Developments of the Intensity-Scale technique within the Spatial Verification Methods Inter-Comparison Project

Barbara Casati

OURANOS

NCAR

Acknowledgements: Barb Brown, Eric Gilleland, John Halley-Gotway, Dave Ahijevych, Matt Pocernich, Beth Ebert, Marion Mittermaier, Louis Lefaivre

Talk outline:

1. The Intensity-Scale verification technique
2. Sensitivity to displacement and intensity error
3. Scale of the error, single-band spatial filter
4. Tiling – Aggregating – Confidence Intervals
5. Conclusions

Intensity-scale verification technique

Casati et al (2004) Met Appl, vol 11

→

Casati (2009) Wea. & Forecasting, submitted

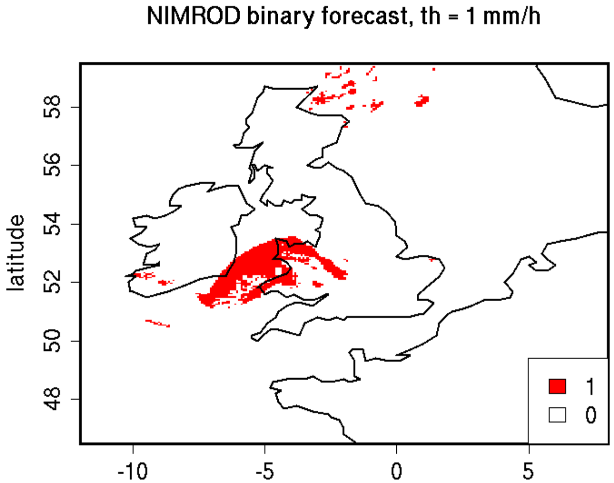
“Evaluate the forecast skill as a function of the precipitation intensity and the spatial scale of the error”

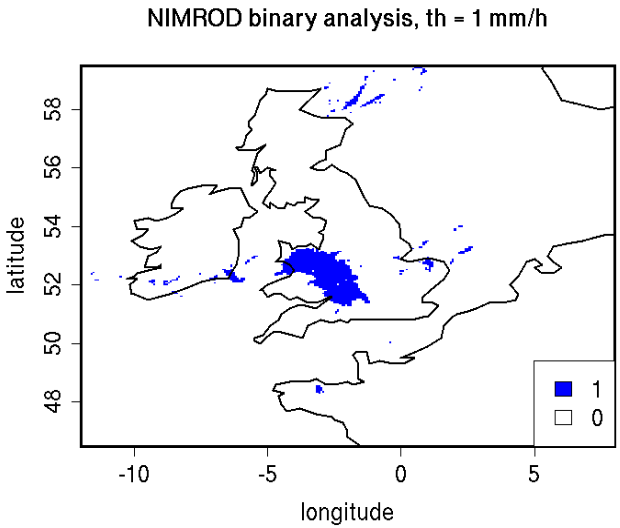
Changes: no more recalibration, biased forecasts. Energy enables the assessment of bias for different thresholds and scales. Strategies to tackle dyadic domain constraints: tiling. Aggregation of IS statistics, bootstrap Confidence Intervals.

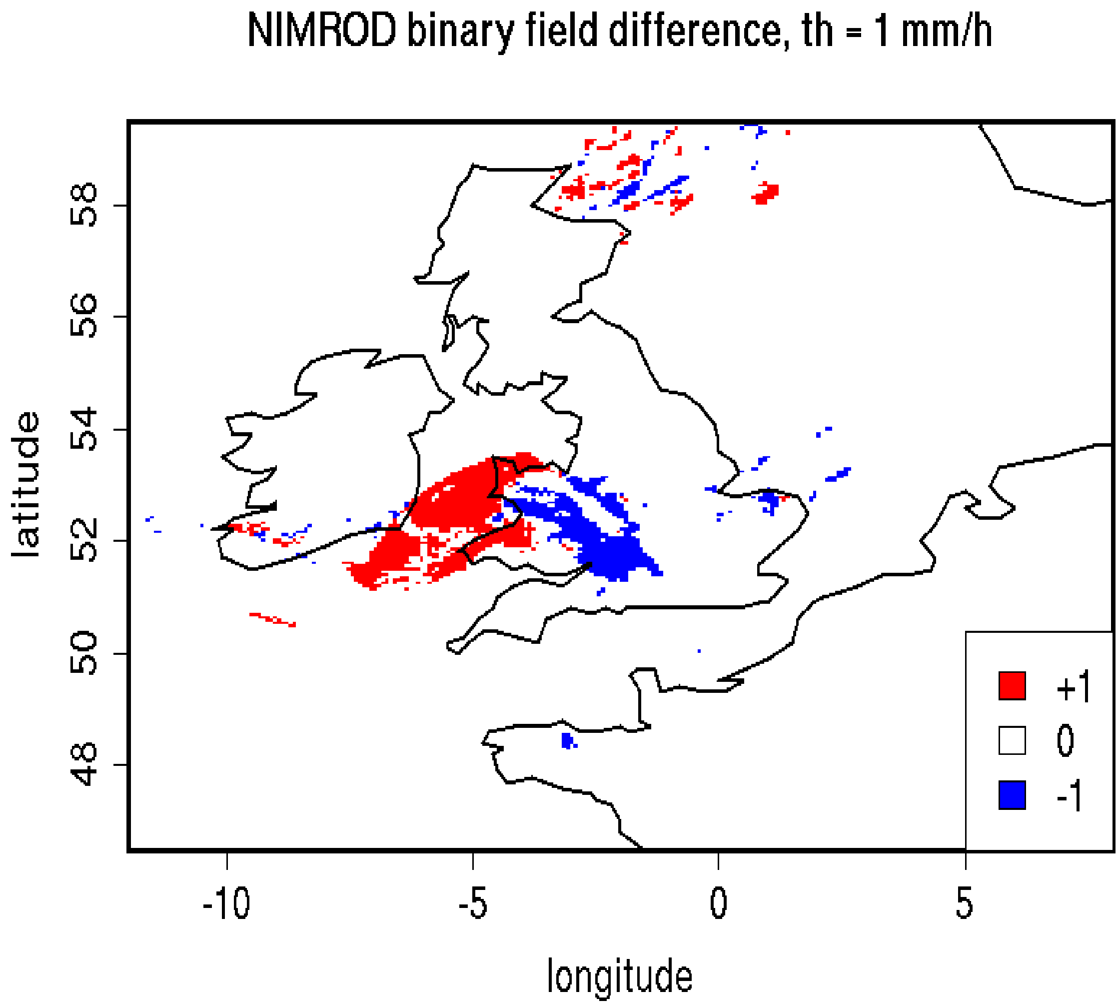
NOTE: **scale** = single-band spatial filter → separate features of different scales → feedback on different physical processes

In the neighborhood based (fuzzy) verification, the **scale** is the neighborhood size (low band pass filter): as the scale increases the exact positioning requirements are more and more relaxed

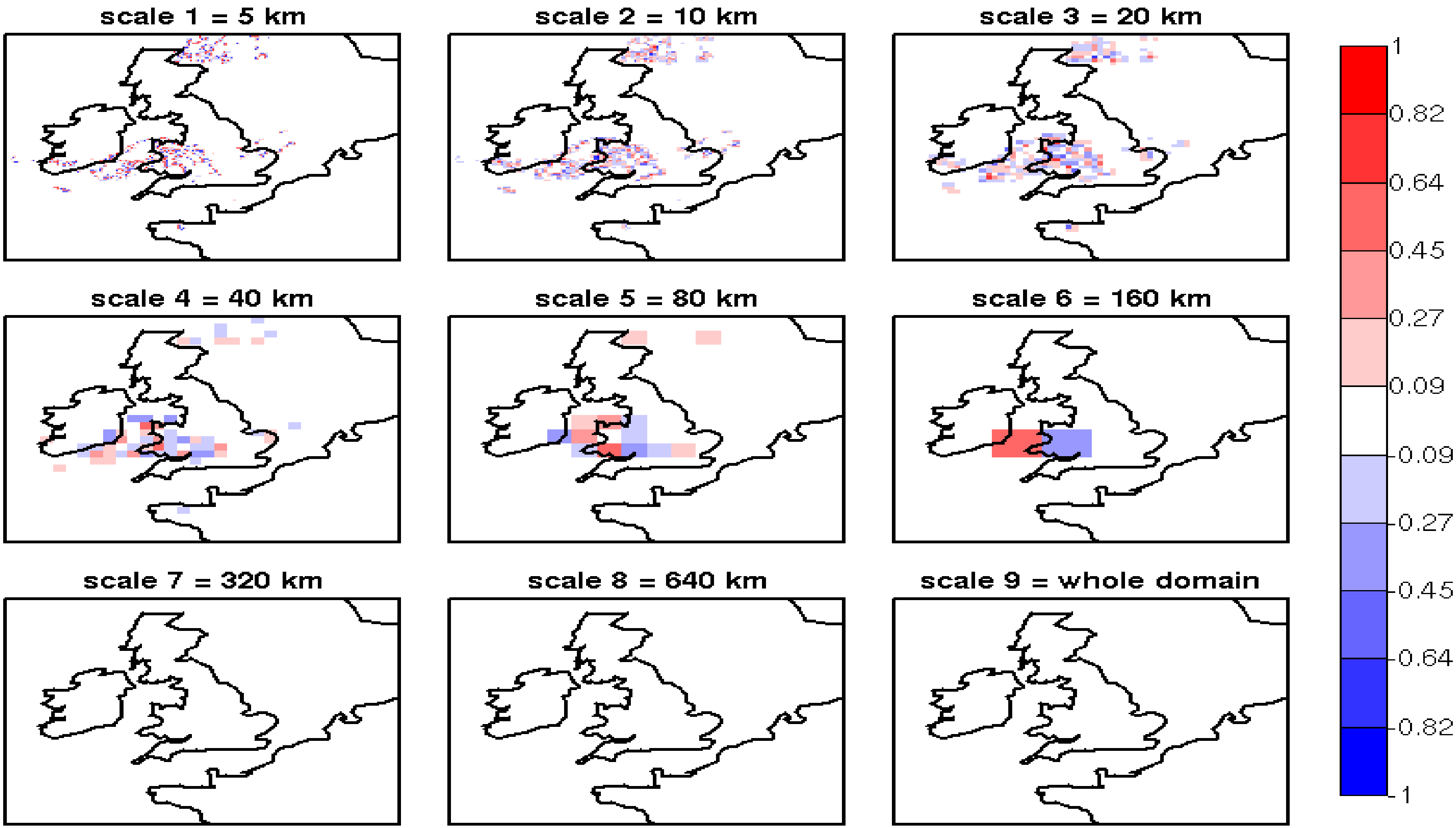
Intensity: threshold to obtain binary images







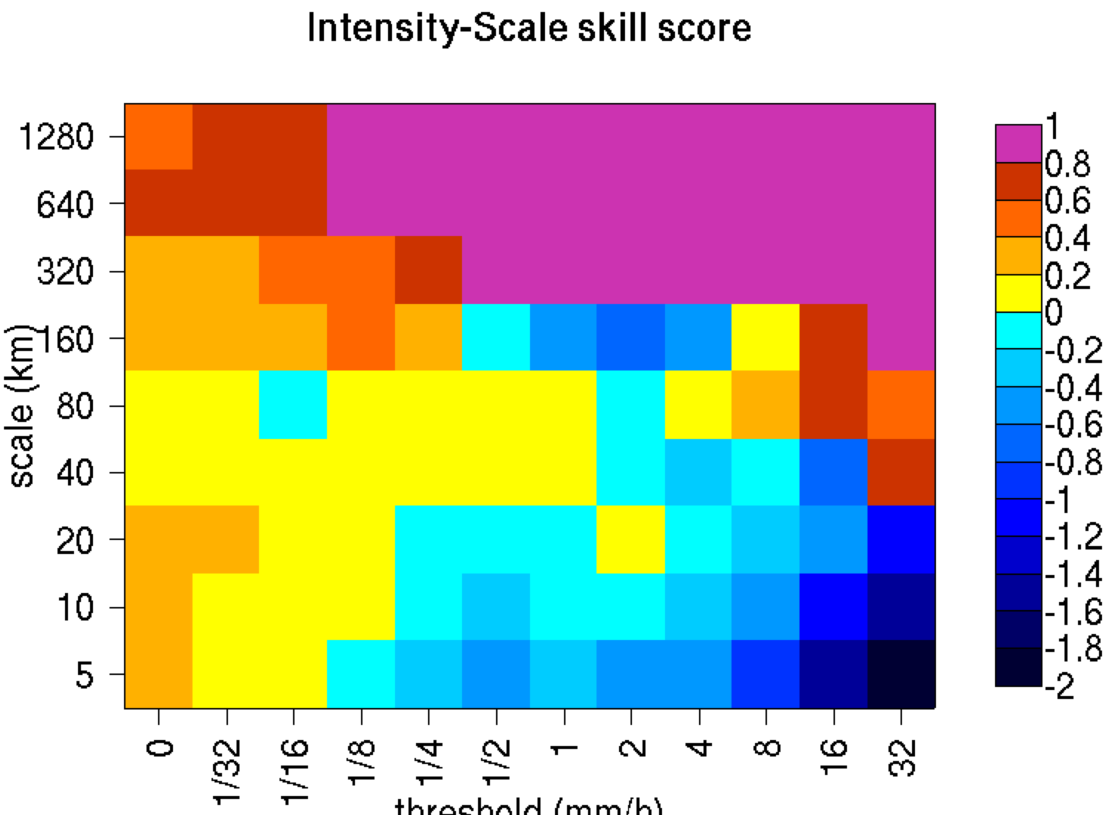
Scale: 2D Haar wavelet decomposition of the binary images



Intensity-scale skill score

For each threshold and scale component: skill score associated to the MSE of binary images (= HSS)

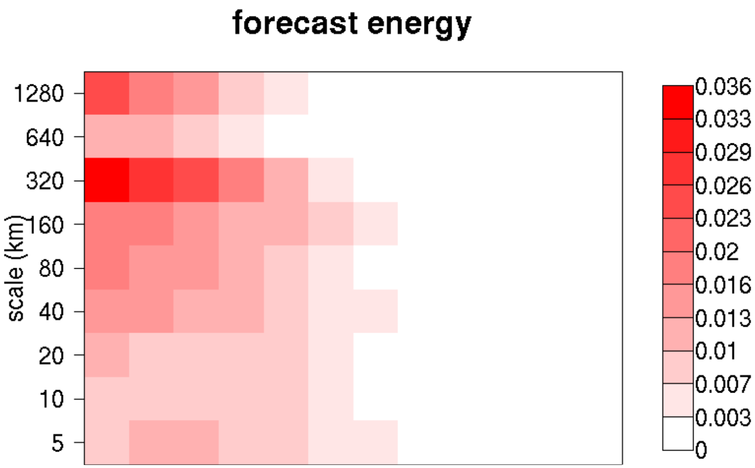
Skill versus random chance, equally partitioned across the scales

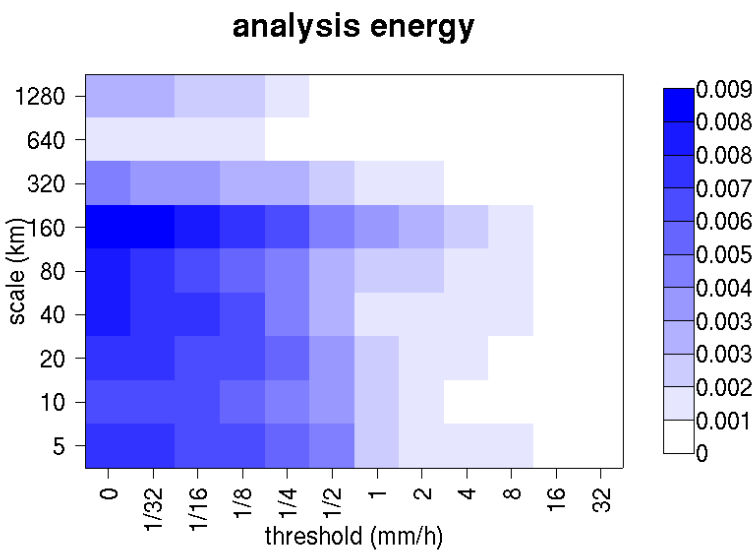


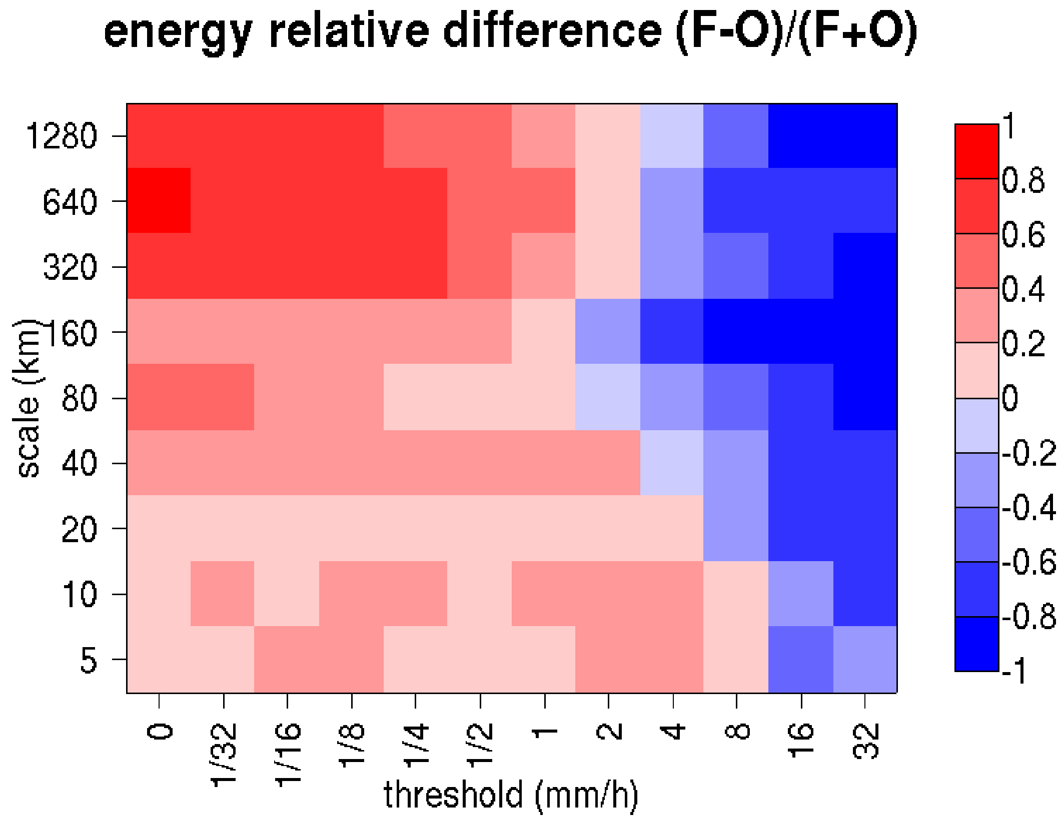
The IS skill score is capable of isolating specific scale-dependent errors: the displaced storm exhibits negative skill for the 160km scale

In general, small scales exhibit negative skill, whereas large scales exhibit positive skill

Energy and Bias





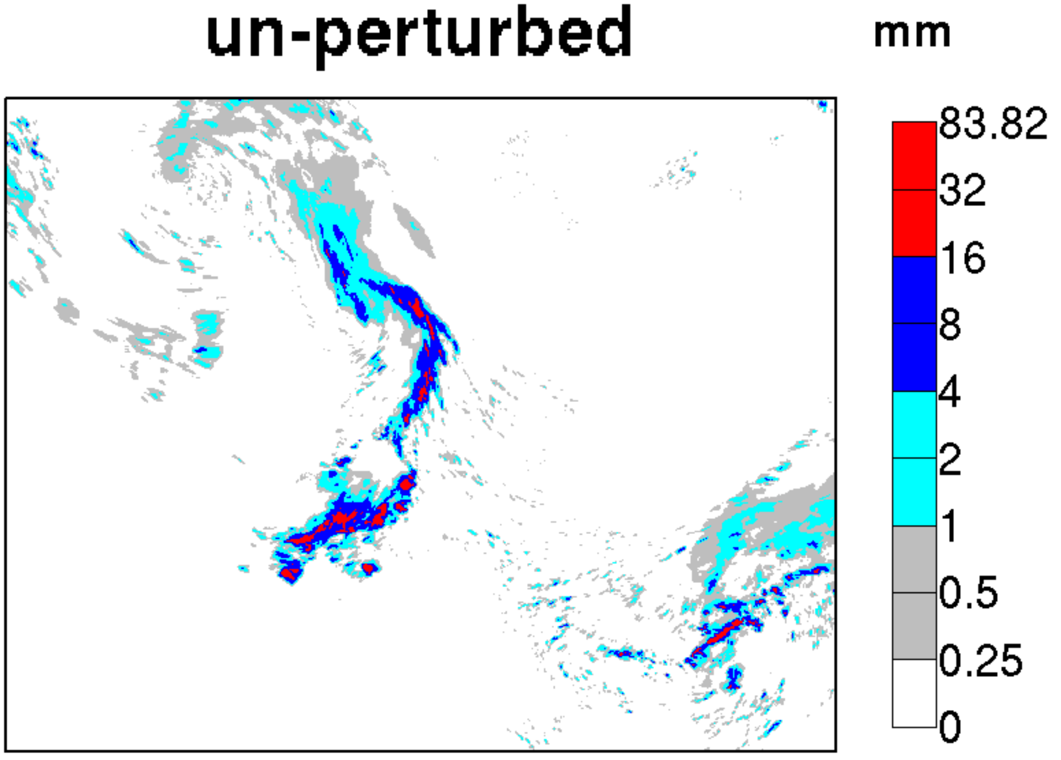


For each threshold and scale:

the energy informs on the amount of events (energy = mean(X^2))

the energy relative difference measures the bias = (B-1)/(B+1)

Perturbed cases



S-E displacement:

case1: **(3,5)** gpt

case2: **(6,10)** gpt

case3: **(12,20)** gpt

case4: **(24,40)** gpt

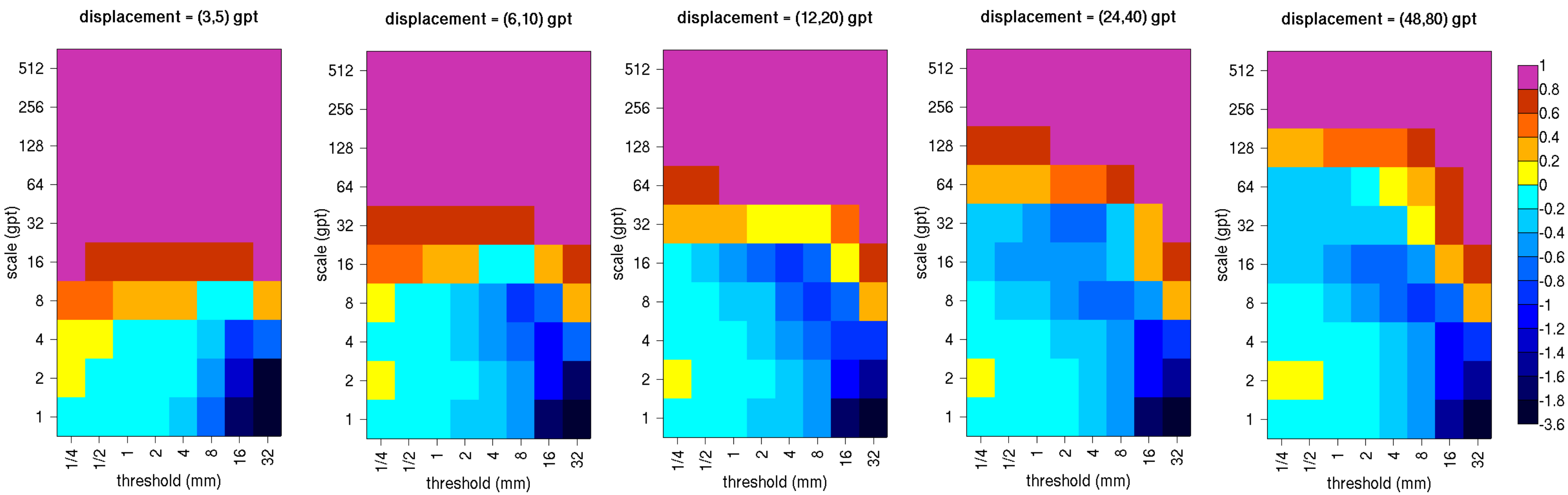
case5: **(48,80)** gpt

Intensity perturbation:

case6: **pcpn x 1.5** (same displacement as case3)

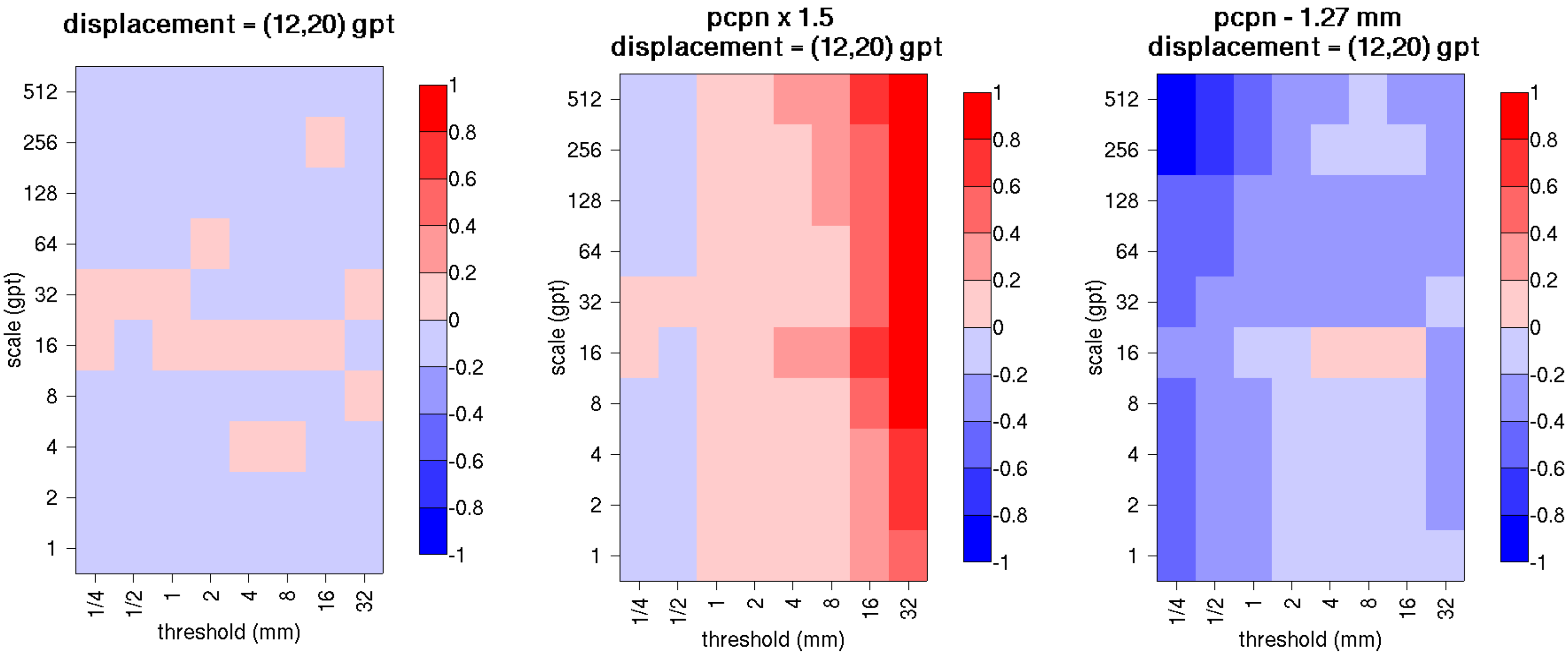
case7: **pcpn – 1.27mm** (same displacement as case3)

Sensitivity to displacement errors



As the displacement gets larger, the no-skill to skill transition scale gets larger: the IS skill score is sensitive to displacement errors

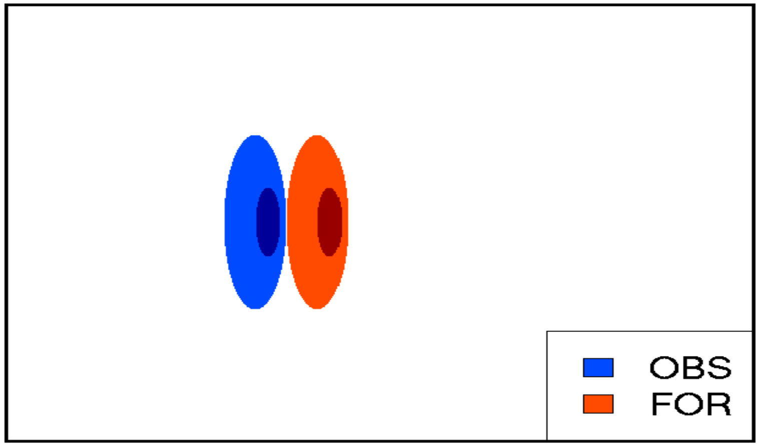
Sensitivity to intensity errors



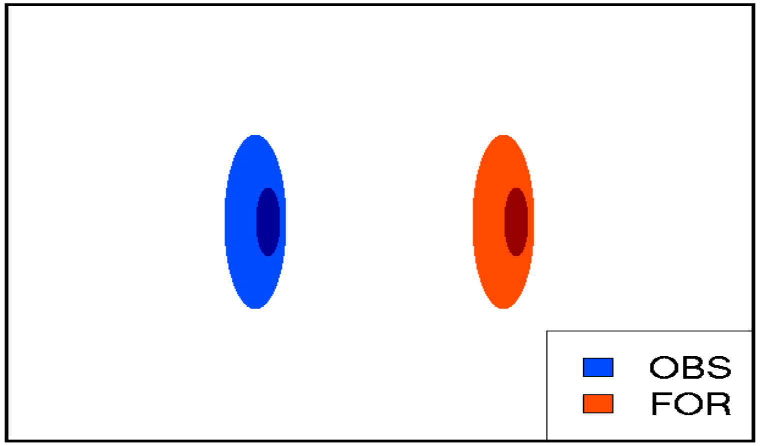
pcpn x 1.5 affects mostly large intensities, overforecast; pcpn - 1.27mm affects mostly small intensities and large scales, underforecast. The energy bias is sensitive to intensity errors

Geometric cases

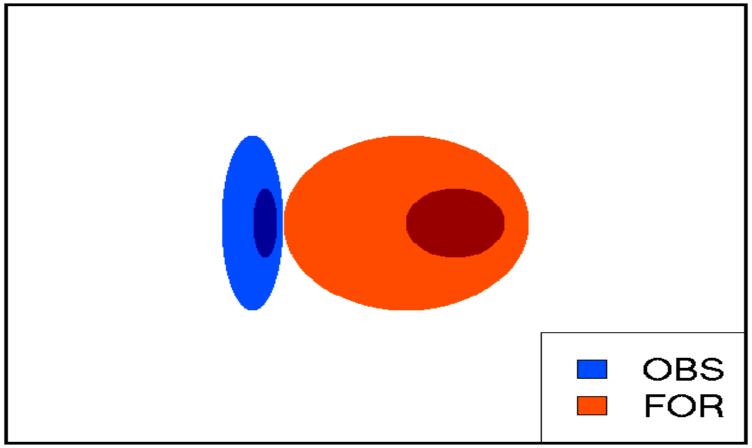
50 gpt to the right



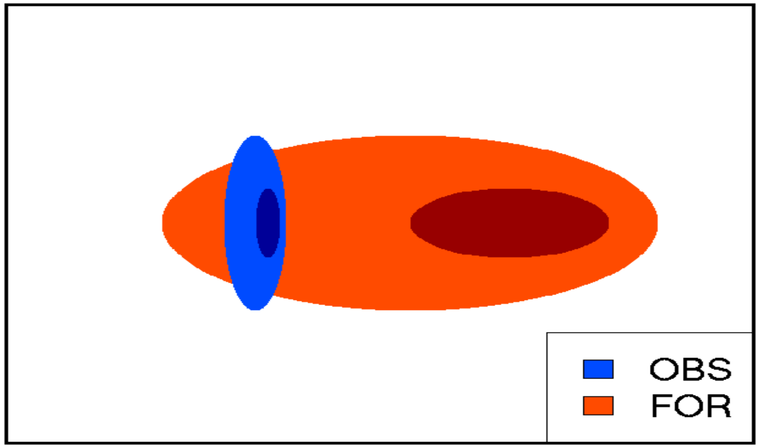
200 gpt to the right



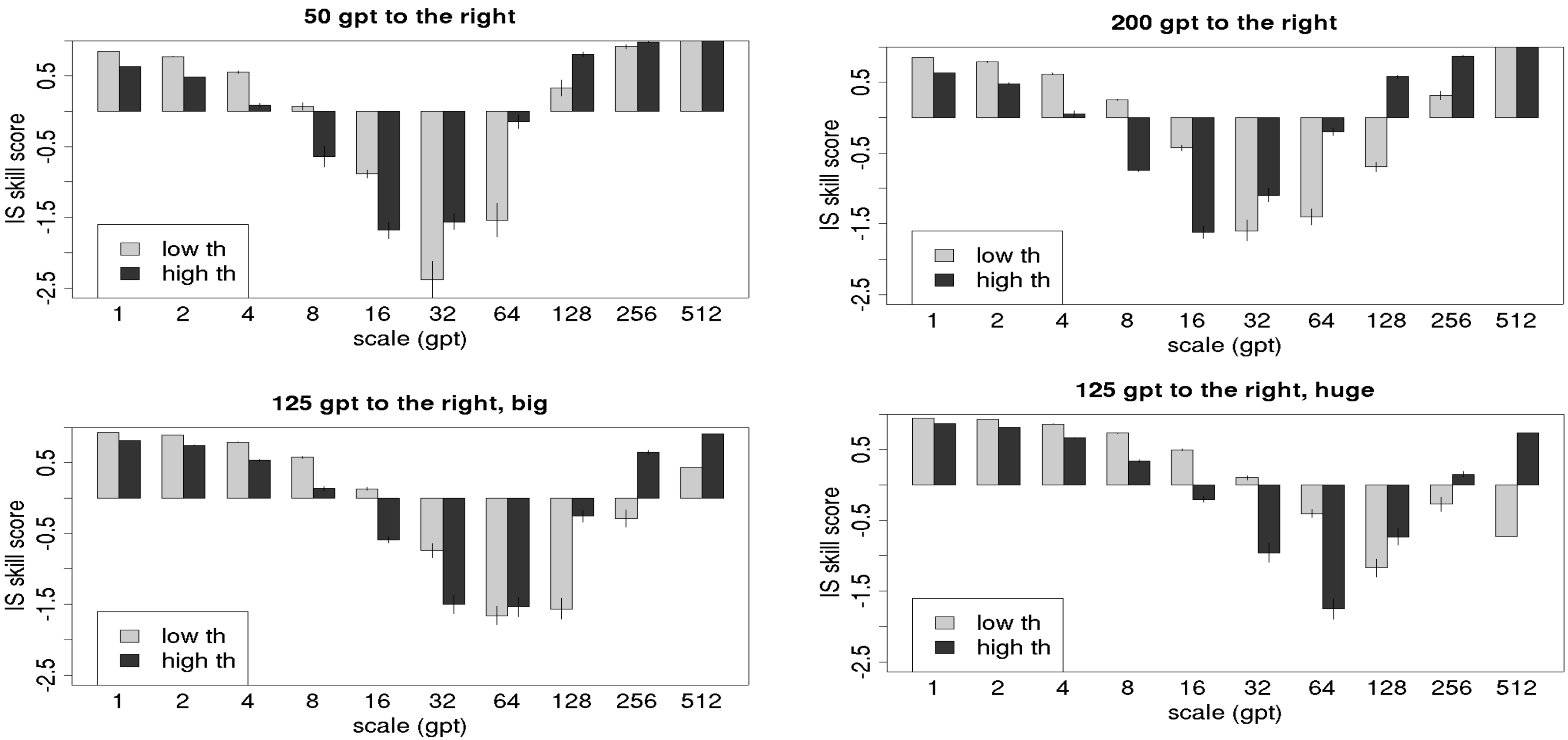
125 gpt to the right, big



125 gpt to the right, huge

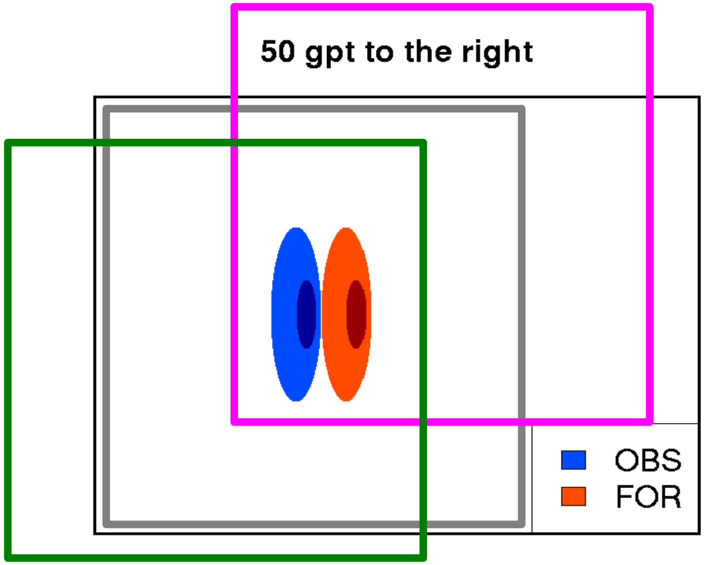


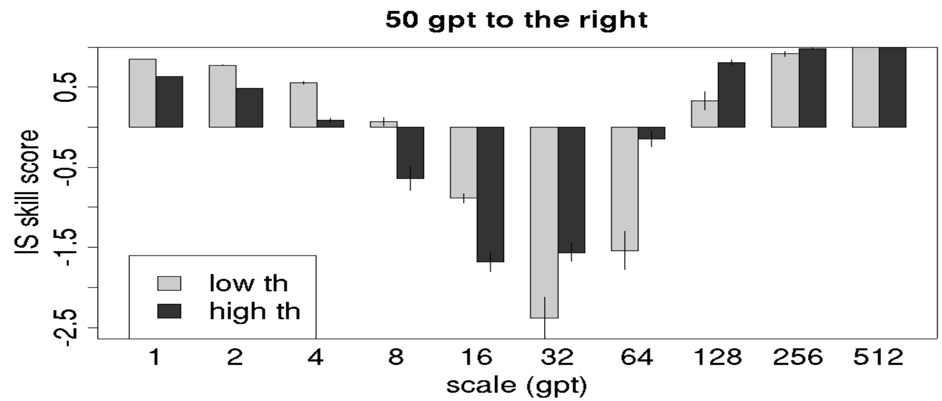
IS skill score for the geometric cases

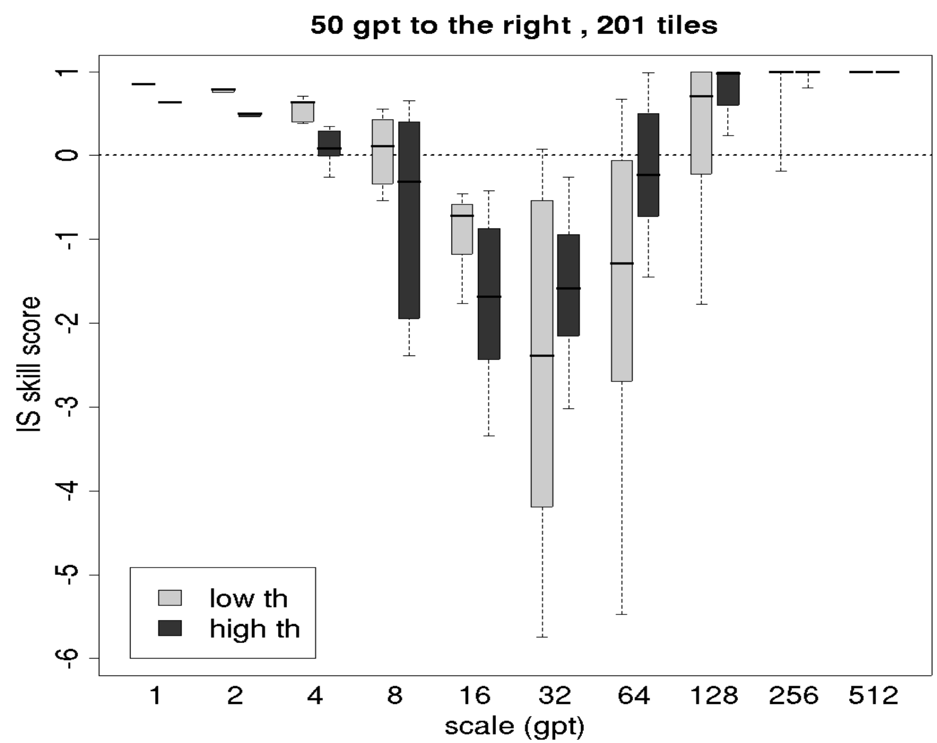


- Scale of the error = size of the feature and its displacement
- Positive skill on small scales? Smooth ellipses, no error (variability) on small scales (NOTE: IS verification relies on a single-band spatial filter).

Tiling – Aggregating – Bootstrap CIs



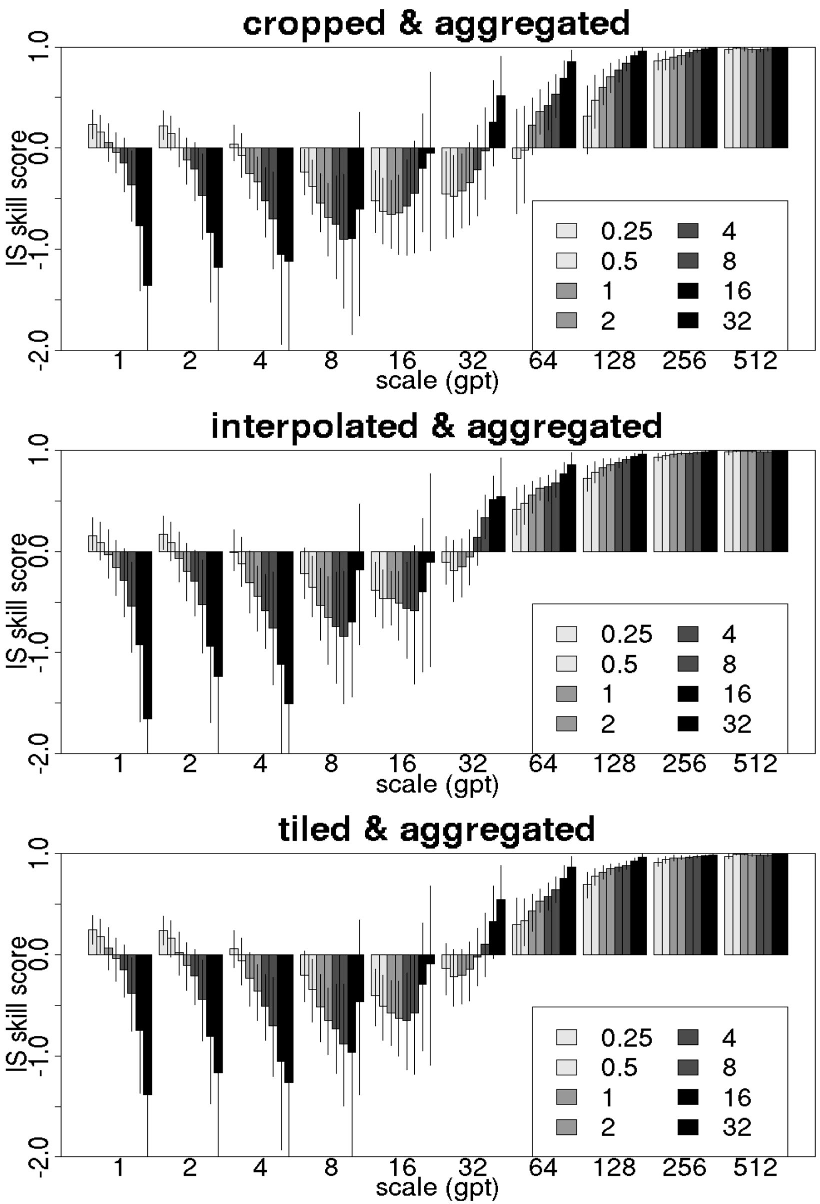




<u>Tiling eliminates the effects of the discrete wavelet support:</u>

1. aggregate IS statistics for all tiles (aggregation ≠ average);
2. Confidence Intervals evaluated by bootstrapping

Multiple precipitation forecasts



Spring 2005 data-set, WRF 4km NCAR

Precipitation fields require less tiles than geometric cases to eliminate discrete support effects

Aggregation of multiple forecasts implicitely eliminates the discrete support effects, since weather moves wrt the wavelet support

Strategies to address dyadic square domain constraints: **cropping, padding, interpolation, tiling**. For

Conclusions

The intensity-scale verification approach evaluates bias and skill for different precipitation intensities and spatial scales.

Precipitation forecasts in general exhibit negative skill on small scales, and positive skill on large scales (predictability). However the IS skill score is capable of identifying specific scale-dependent errors (e.g. Nimrod displaced storm → negative skill at 160km scale).

The scale of the error is associated with both the features' size and their displacement; the IS statistics are sensitive to displacement and intensity errors.

Tiling smooths the effects of the discrete wavelet support: appropriate for single cases verification. For aggregated precipitation forecasts tiling, cropping, padding or interpolating provide similar results.

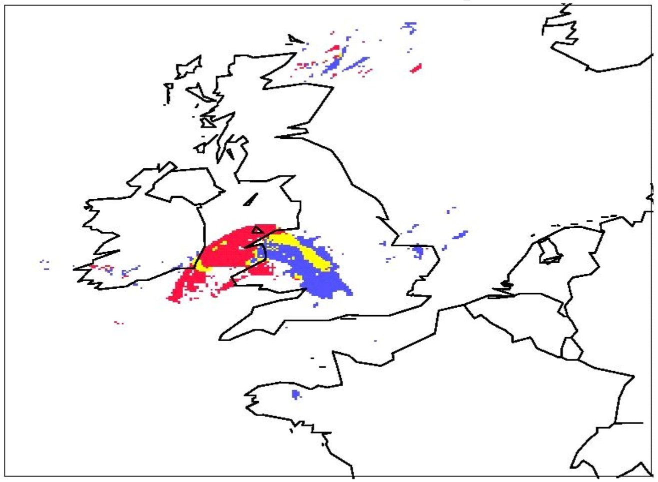
THANK YOU!

Intensity-scale verification technique
Casati et al. (2004), Met App, vol. 11

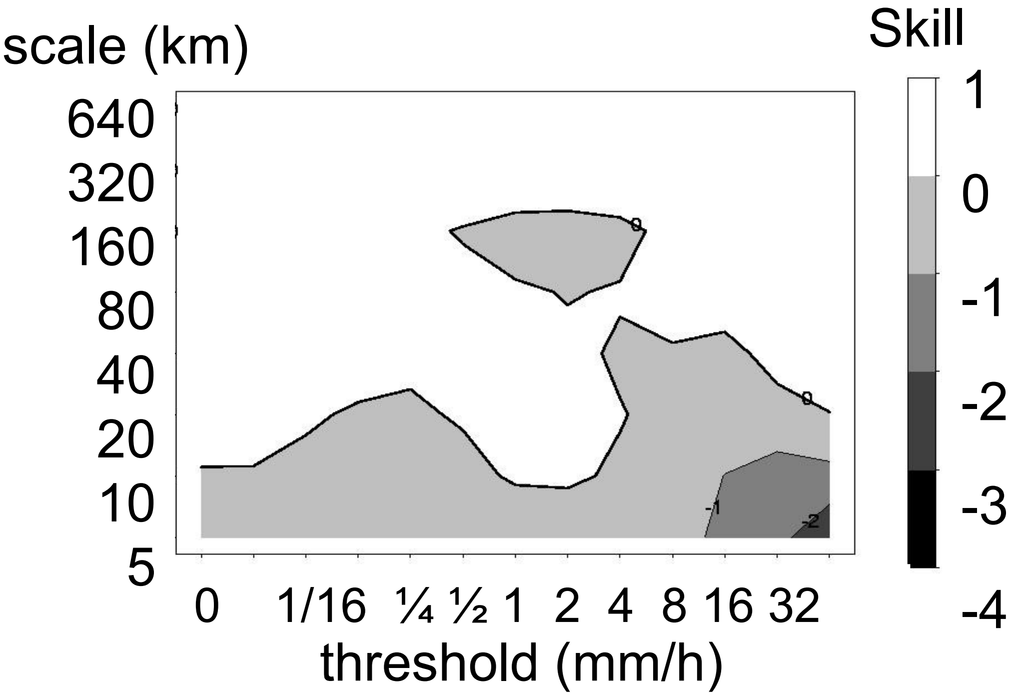
The intensity-scale verification approach measures the skill as function of precipitation intensity and spatial scale of the error

- ✓ **intensity: threshold** → binary images (categorical approach)
- ✓ **scale: 2D Wavelets** decomposition of binary image difference
- ✓ For each threshold and scale: skill score associated to the MSE of binary images = Heidke Skill Score

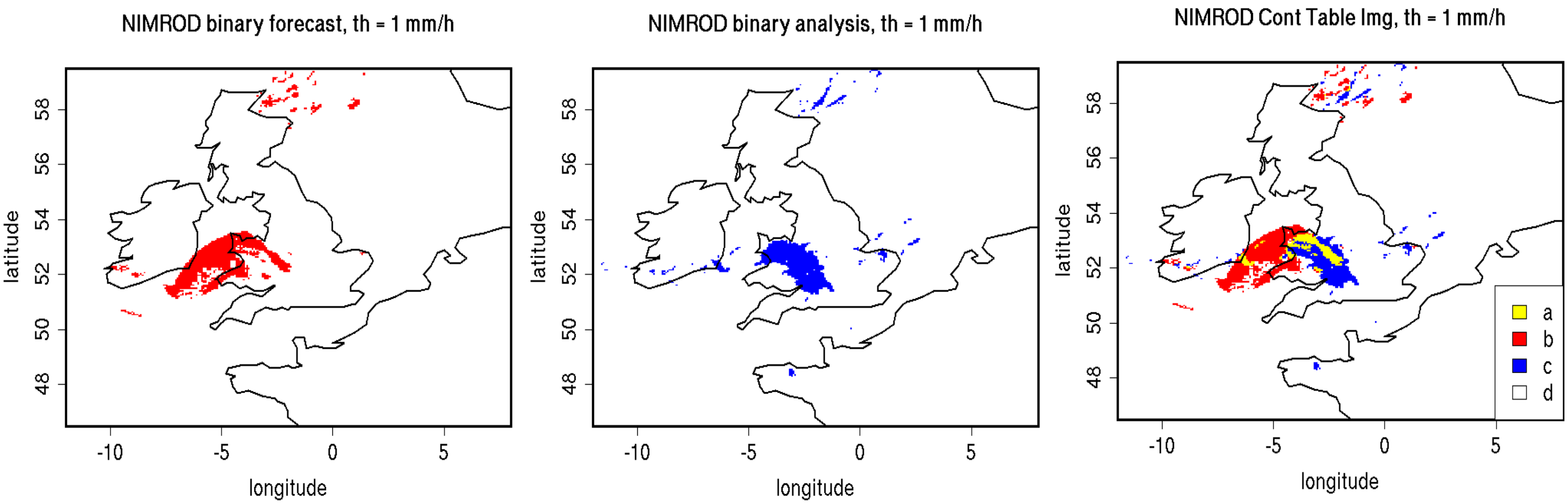
Intense storm displaced



threshold = 1mm/h



Links with categorical verification



$$MSE_u = \frac{b+c}{n}$$

$$SS_u = HSS$$

	X > u	X < u	
Y > u	Hits a	False Alarms b	a+b
Y < u	Misses c	Correct Rejections d	c+d
	a+c	b+d	a+b+c+d=n

Energy bias for the geometric cases