Warnings verification

Clive Wilson & Marion Mittermaier

30th EWGLAM/15th SRNWP meetings – Madrid 6-9 Oct 2008

Contents

- Flash warnings
- Operational verification
- Theory
- Conclusions

Flash warnings
Issued for local authority regions

Type	Criteria
Severe Gales	Repeated gusts of 70 mph or more over land areas *ie 2 or more gusts of 70 mph or more at separate hours within the period of the warning*
Heavy rain	Heavy rain expected to persist for at least 2 hours and to give at least 15mm in 3 hours *Or a period of rainfall of sufficient intensity to cause flooding on already saturated ground*
Must also be at least 80% confident (ie FAR < 0.2)	

Scores

		Observed		
		Event	No event	Total
Forecast	Event	a=hits	b=false alarms	a+b=B*(a+c)
	No event	c=misses	(d=correct no)	
	Total	a+c		

$$\text{Hit rate, } H = \frac{a}{a+c} \quad \text{False alarm ratio, } FAR = \frac{b}{a+b}$$

$$\text{Threat, } TS = \frac{a}{a+b+c} \quad \text{Bias, } B = \frac{a+b}{a+c}$$

Deterministic limit (Hewson 2006)

- More forecasts correct than either missed or false alarms

  $a > (b + c)$

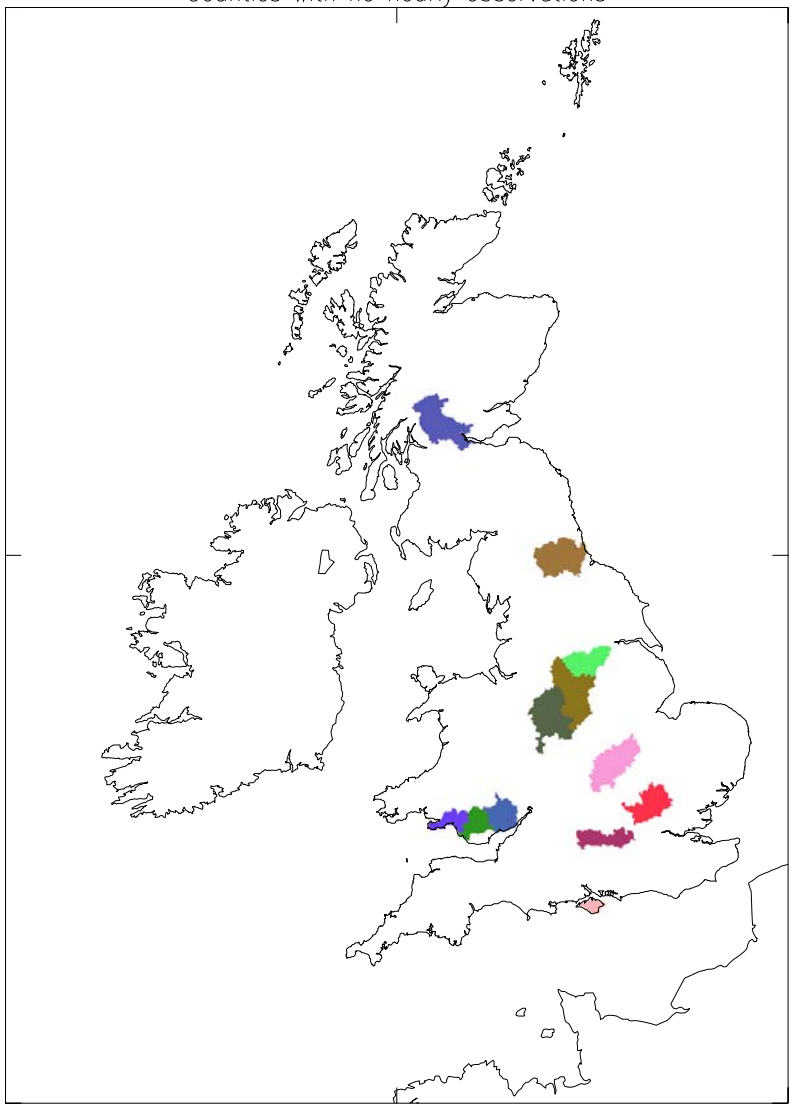
  $2a > (a + b + c)$

  $a/(a + b + c) = TS > 0.5$

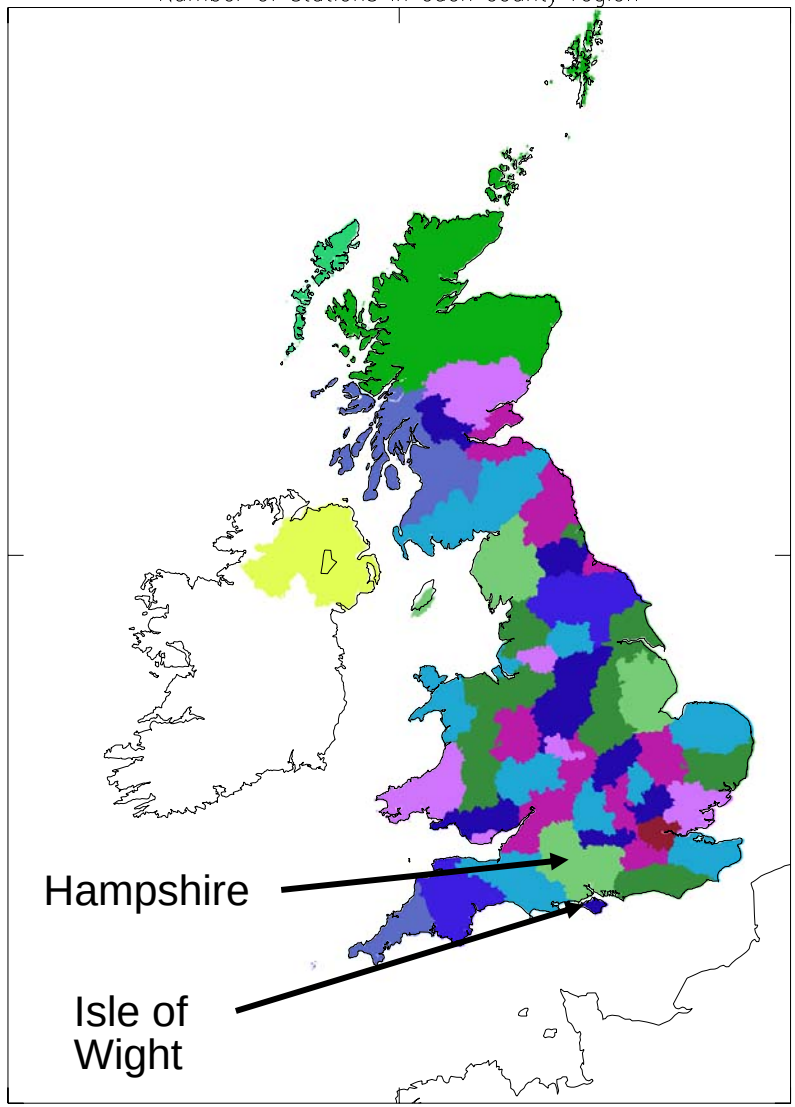
Verification regions & Truth types

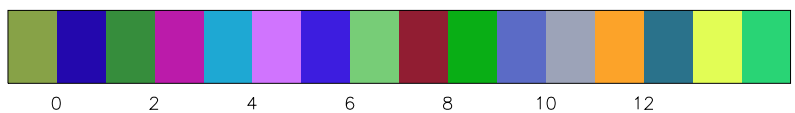
- Amalgamate some small areas into 65 “county” regions
- Must have criteria for gale or heavy rain exceeded for *at least one location*
- Truth:
  - Observations
    - Some counties have none or few
  - Virtual observations
    - UKPP= post-processed UK4 model+radar (2km)
    - Locally adjusted UKPP for site location
    - At least 2 per region
  - Nimrod (nowcast) analyses (15km grid)
  - UKPP analyses (nominally 2km grid)

Counties with no hourly observations



Number of stations in each county region





Variation with H and FAR

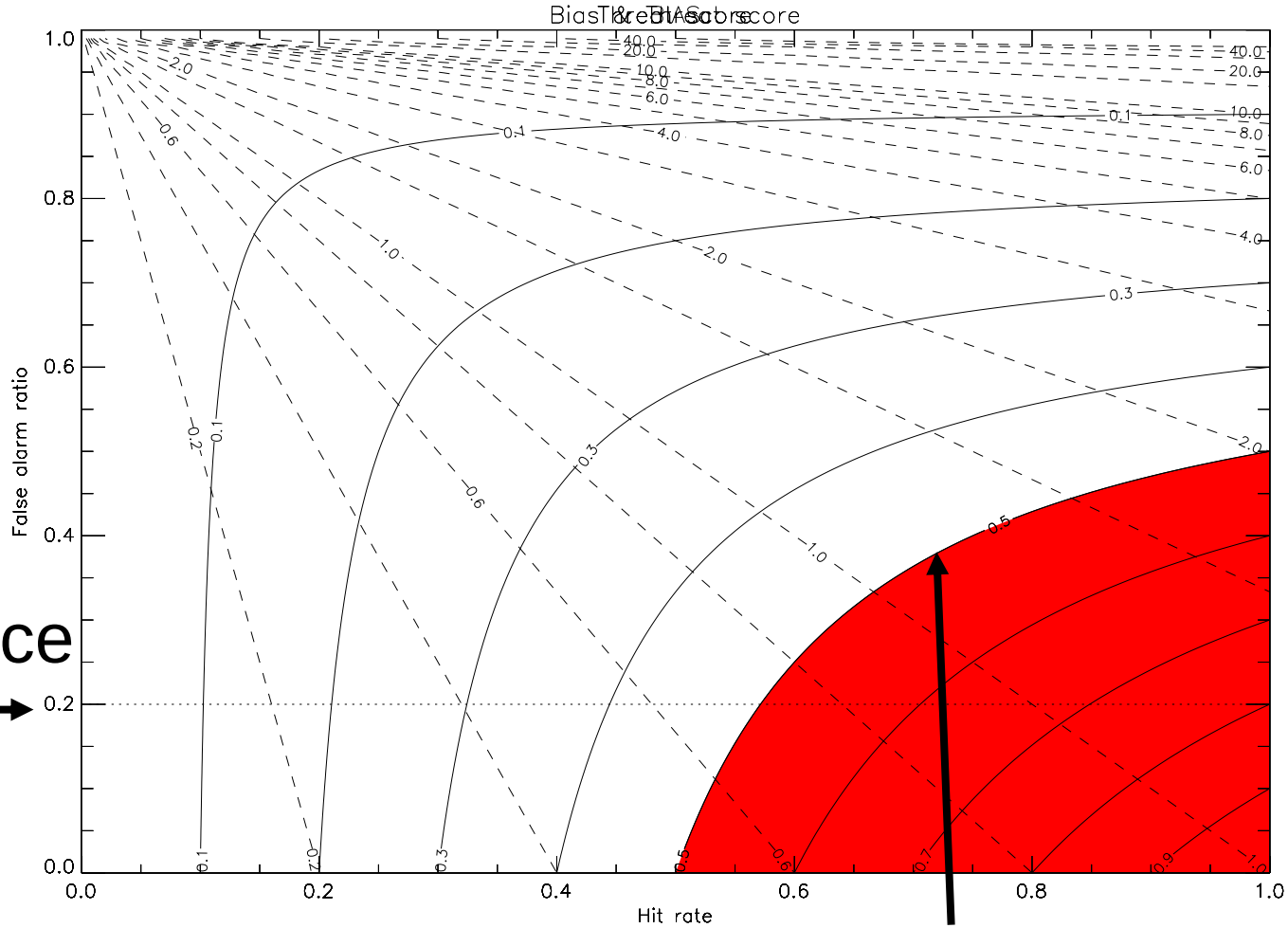
$$H = \frac{a}{a+c} \Rightarrow c = a\frac{(1-H)}{H} \qquad FAR = \frac{b}{a+b} \Rightarrow a+b = \frac{a}{1-FAR}$$

$$TS = \frac{a}{a+b+c} = \frac{a}{a(1/(1-FAR)) + a(1-H)/H}$$

$$\Rightarrow \boxed{TS = \frac{(1-FAR)H}{1-FAR(1-H)}}$$

$$B = \frac{a+b}{a+c} \Rightarrow \boxed{B = \frac{H}{1-FAR}}$$

Hit rate v False alarm ratio plots



Confidence limit →

Deterministic limit

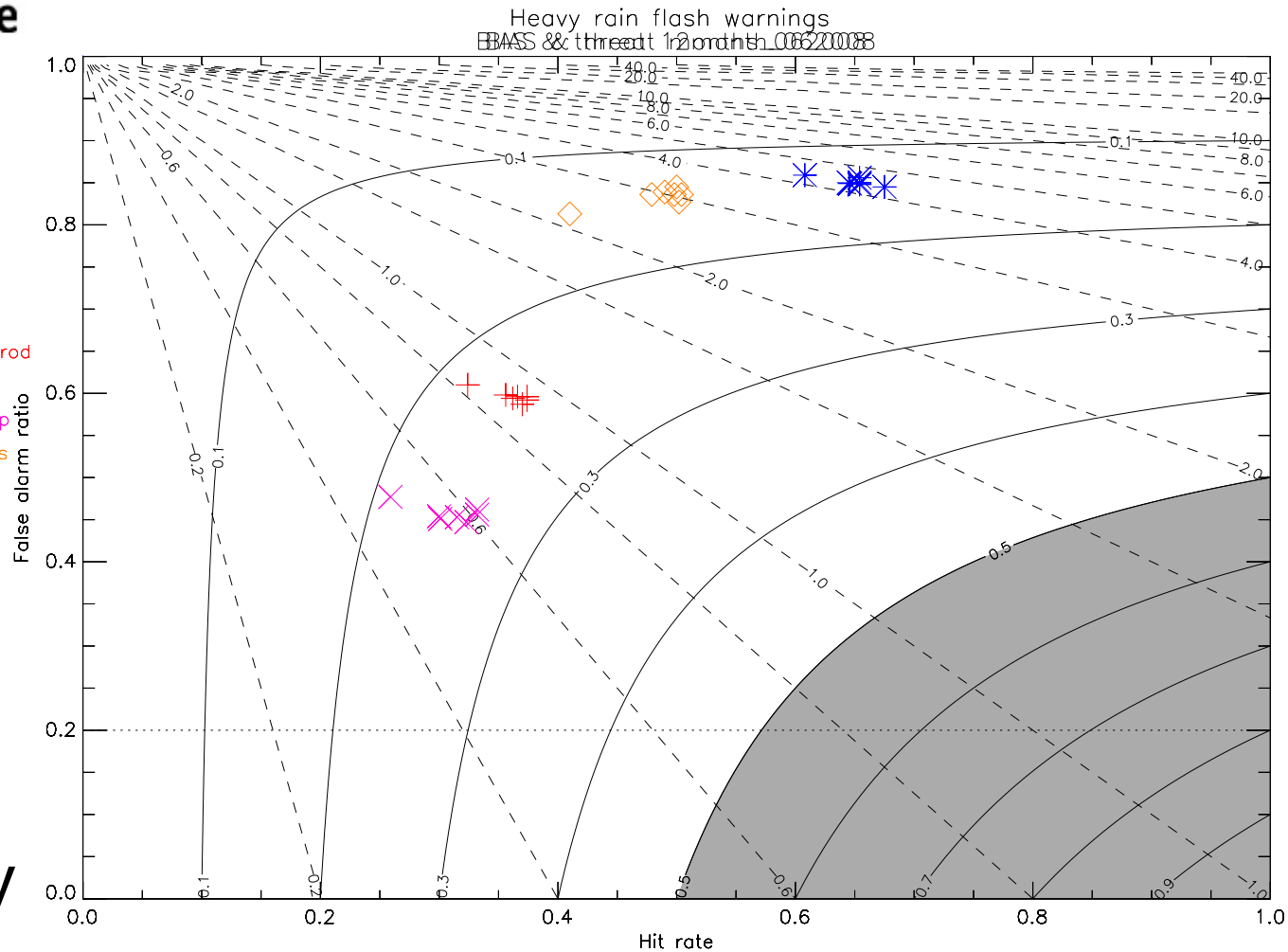
Heavy Rain - forecasters

Nimrod

Obs

UKPP

Virtual Obs



months

12-monthly

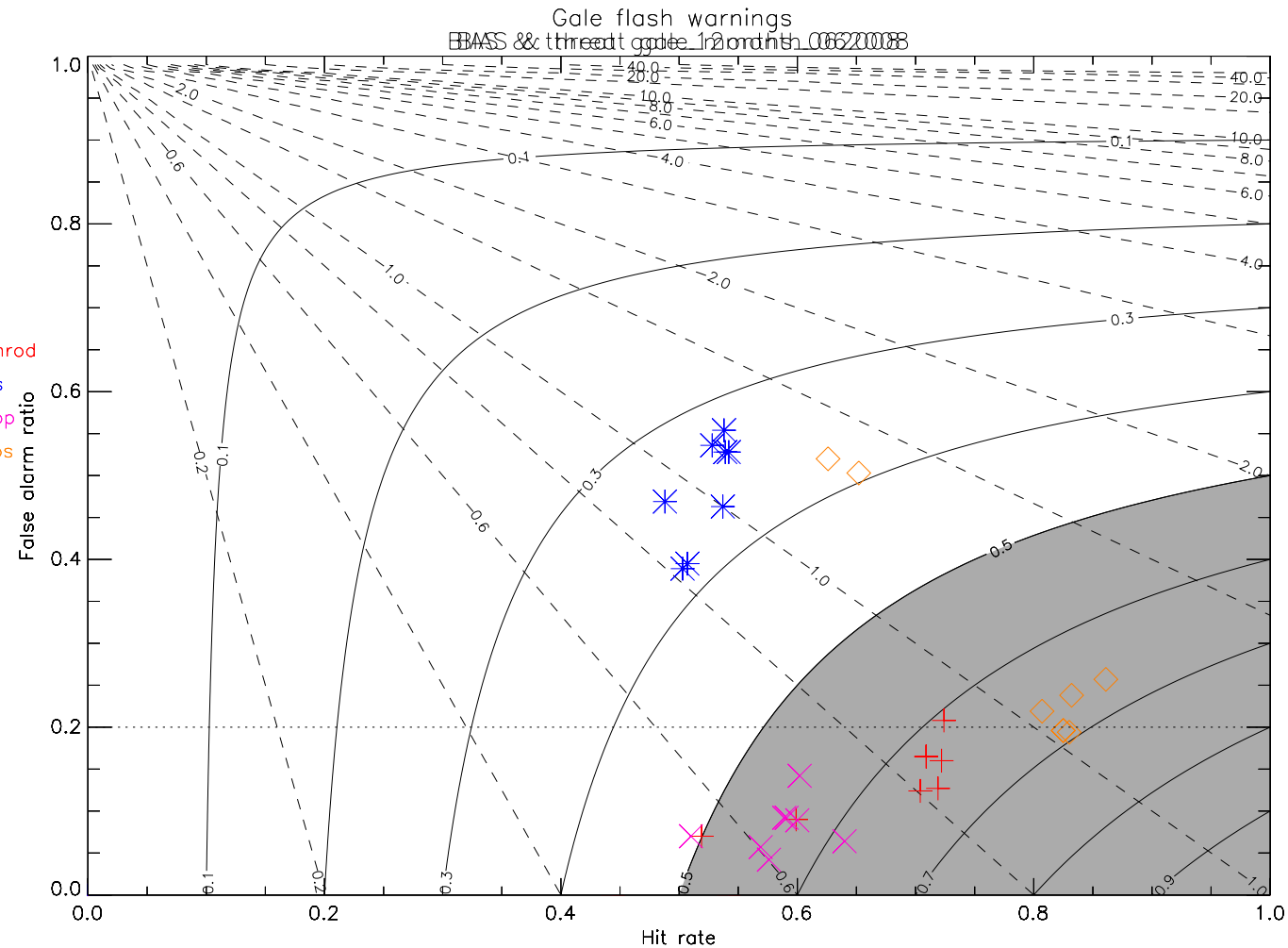
Severe gales - forecasters

Nimrod

Obs

UKPP

Virtual Obs



months

12-monthly

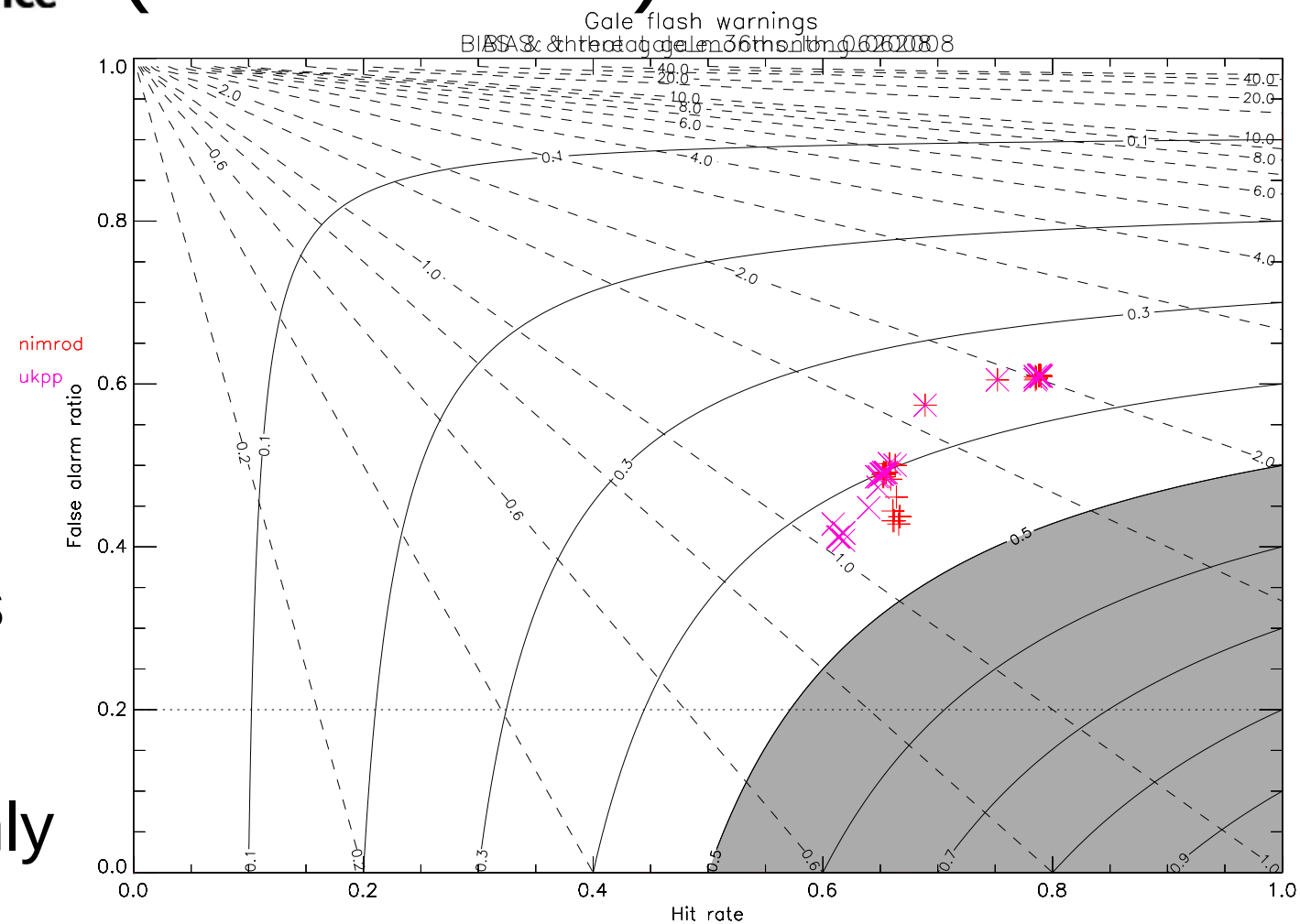
Severe Gales 36-month (forecasters)

Nimrod

UKPP

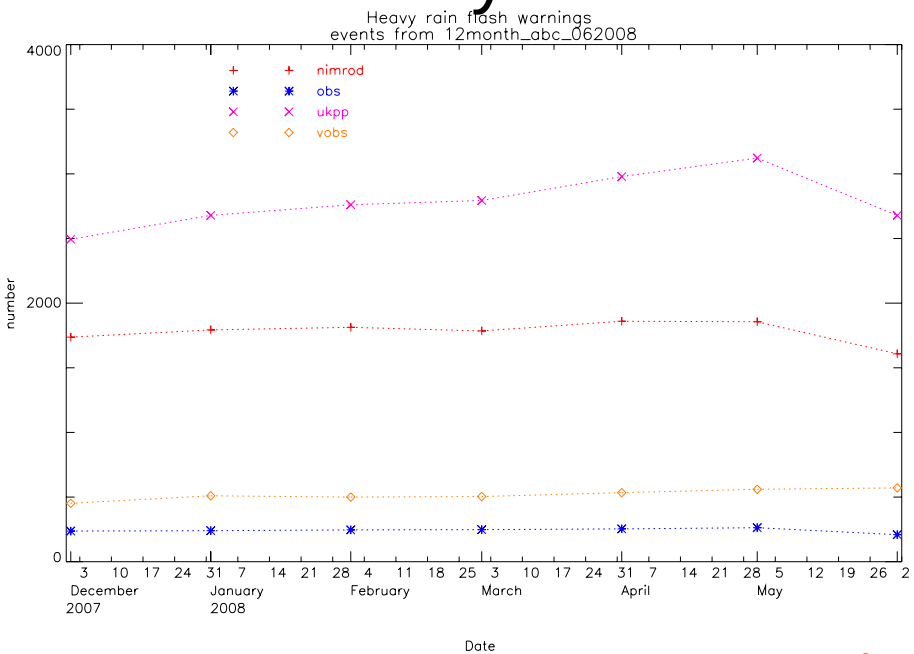
months

36-monthly

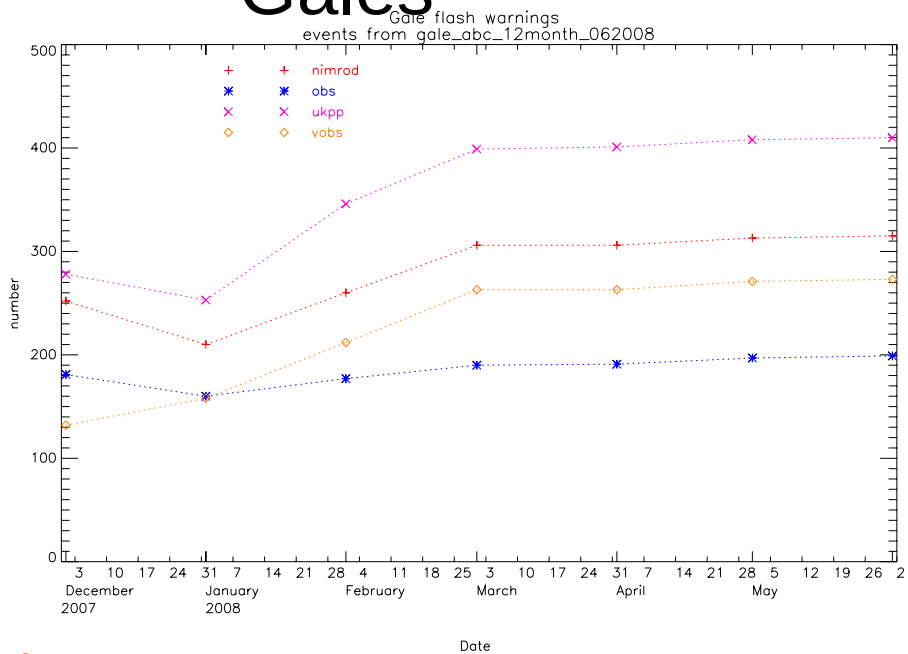


Detection of events

Heavy rain

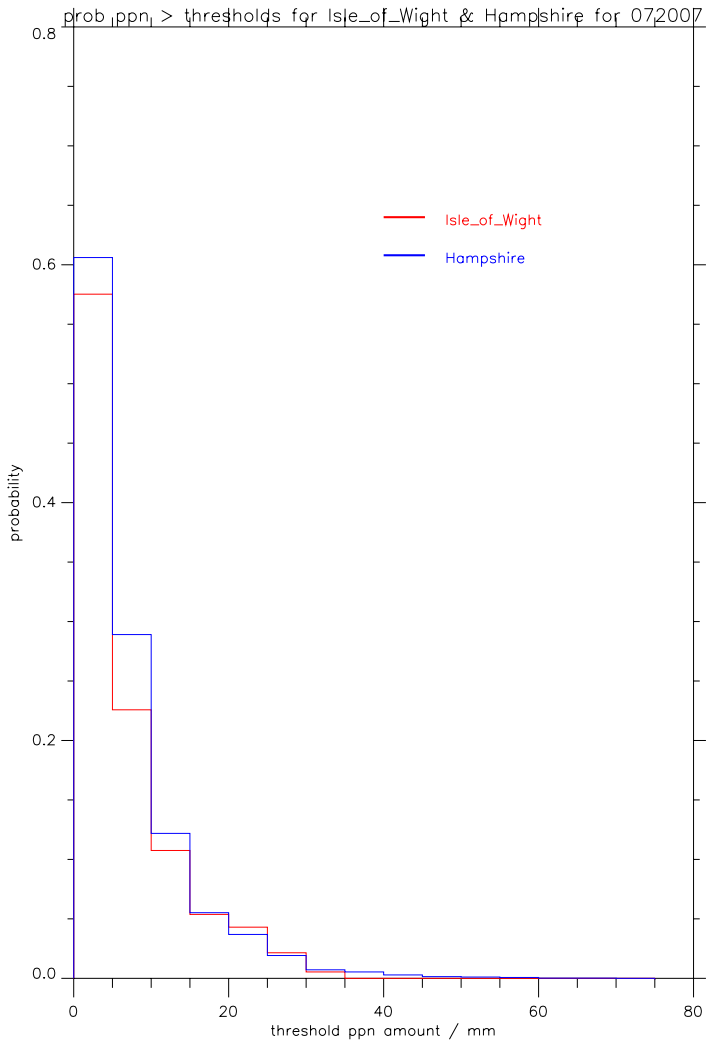


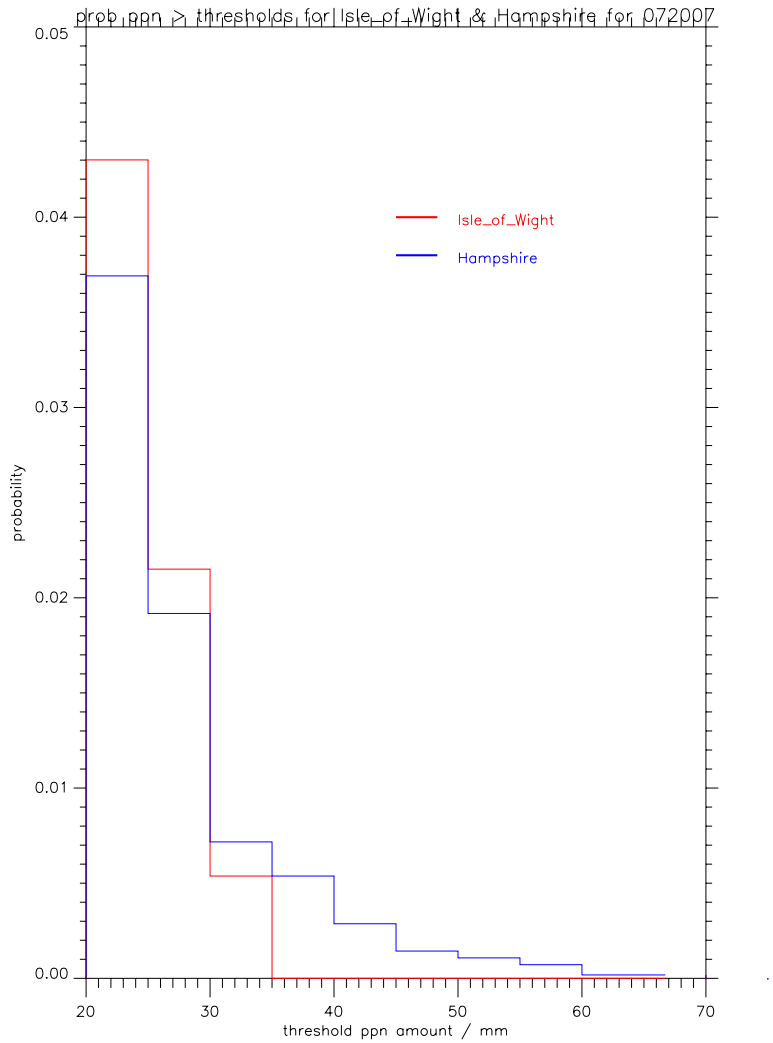
Gales



Nimrod
Obs
UKPP
Virtual Obs

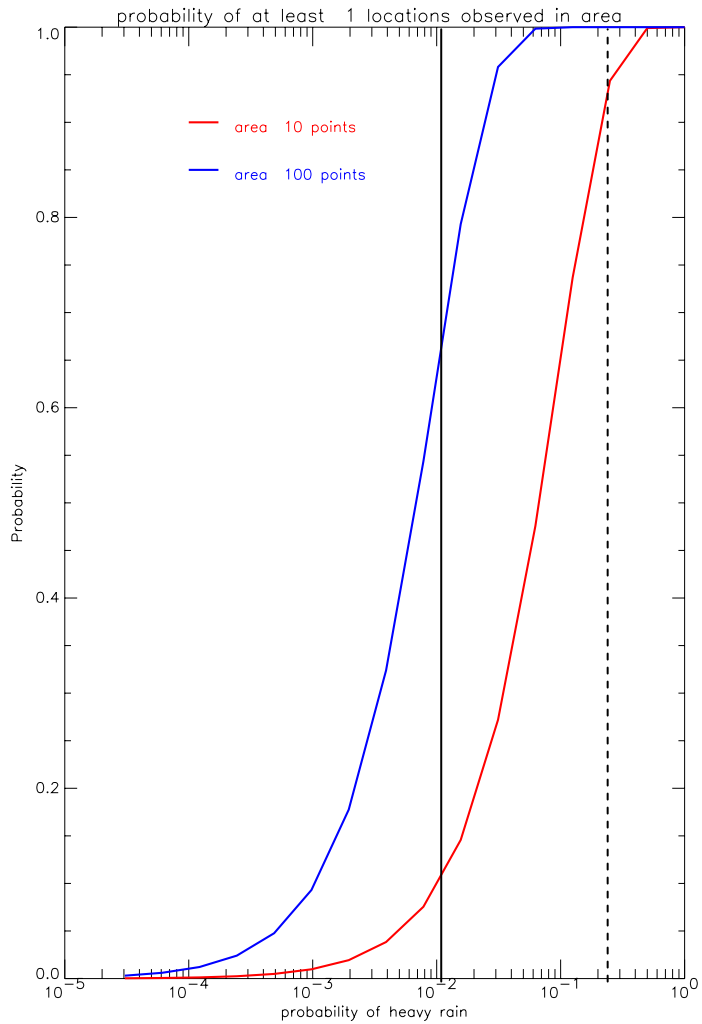
Probability of heavy rain depends on region size





Variation of detecting heavy rain at 1 location with base rate probability

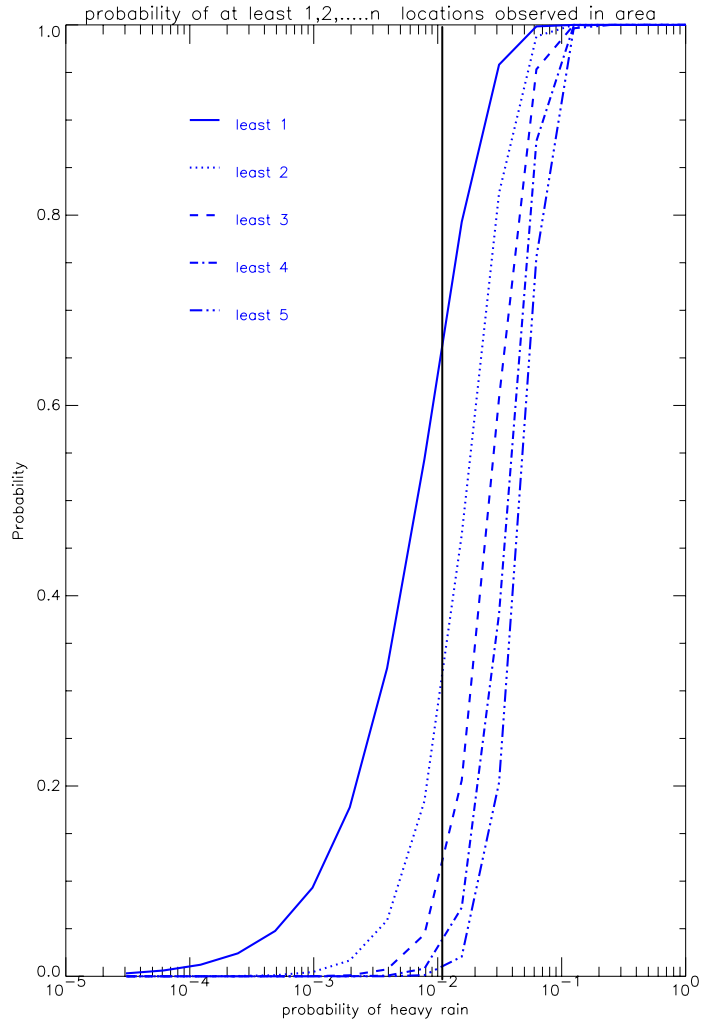
- 2 regions
  - 10 grid points
  - 100 grid points
  - Same base rates p
  - 6-10x more likely to detect for larger region with typical p



Base rate, p

Variation of probability of detecting heavy rain at more than 1 point

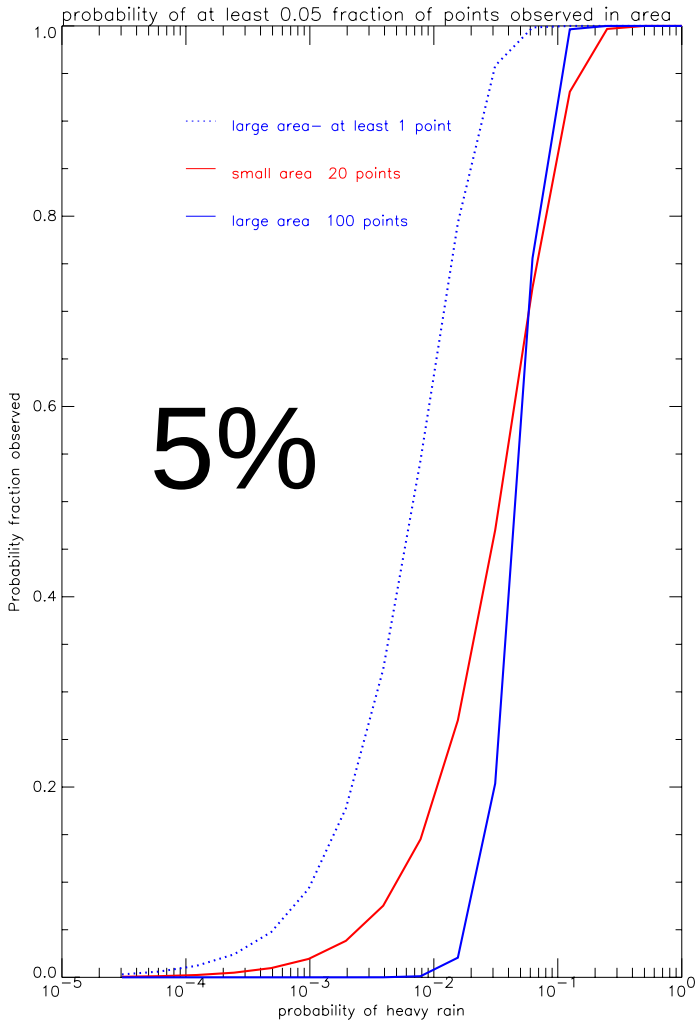
- 1 region
  - 100 grid points
  - base rate p
  - At least 1,2,3,4,5 locations simultaneously



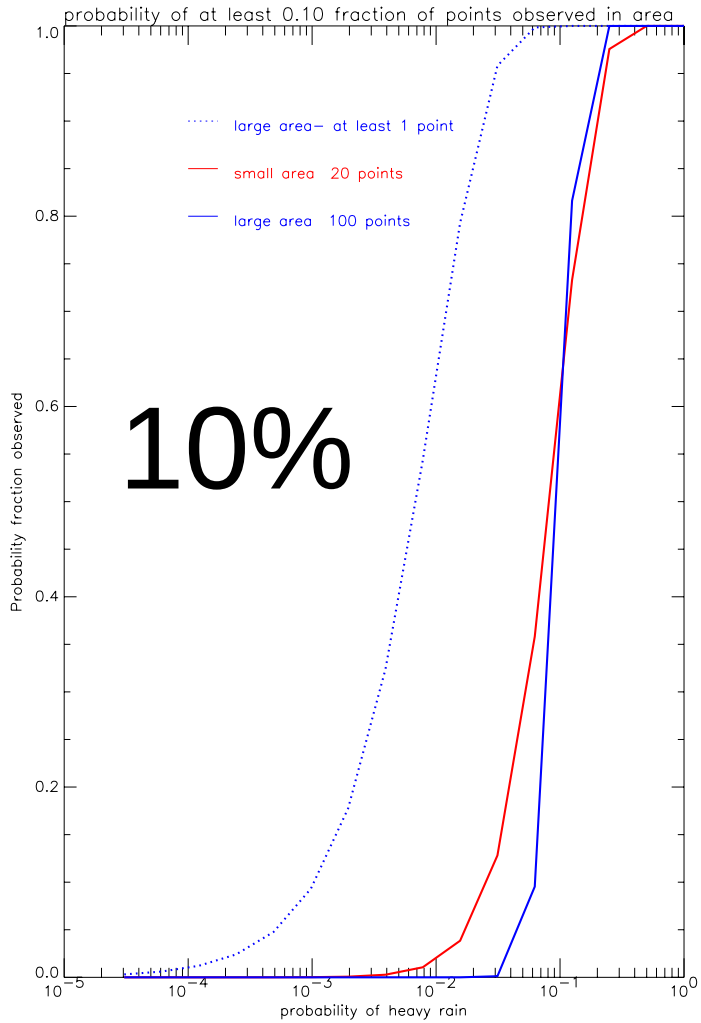
Base rate, p

Probability of detecting a fixed %age of points per county region

More likely to detect over smaller area for rare events

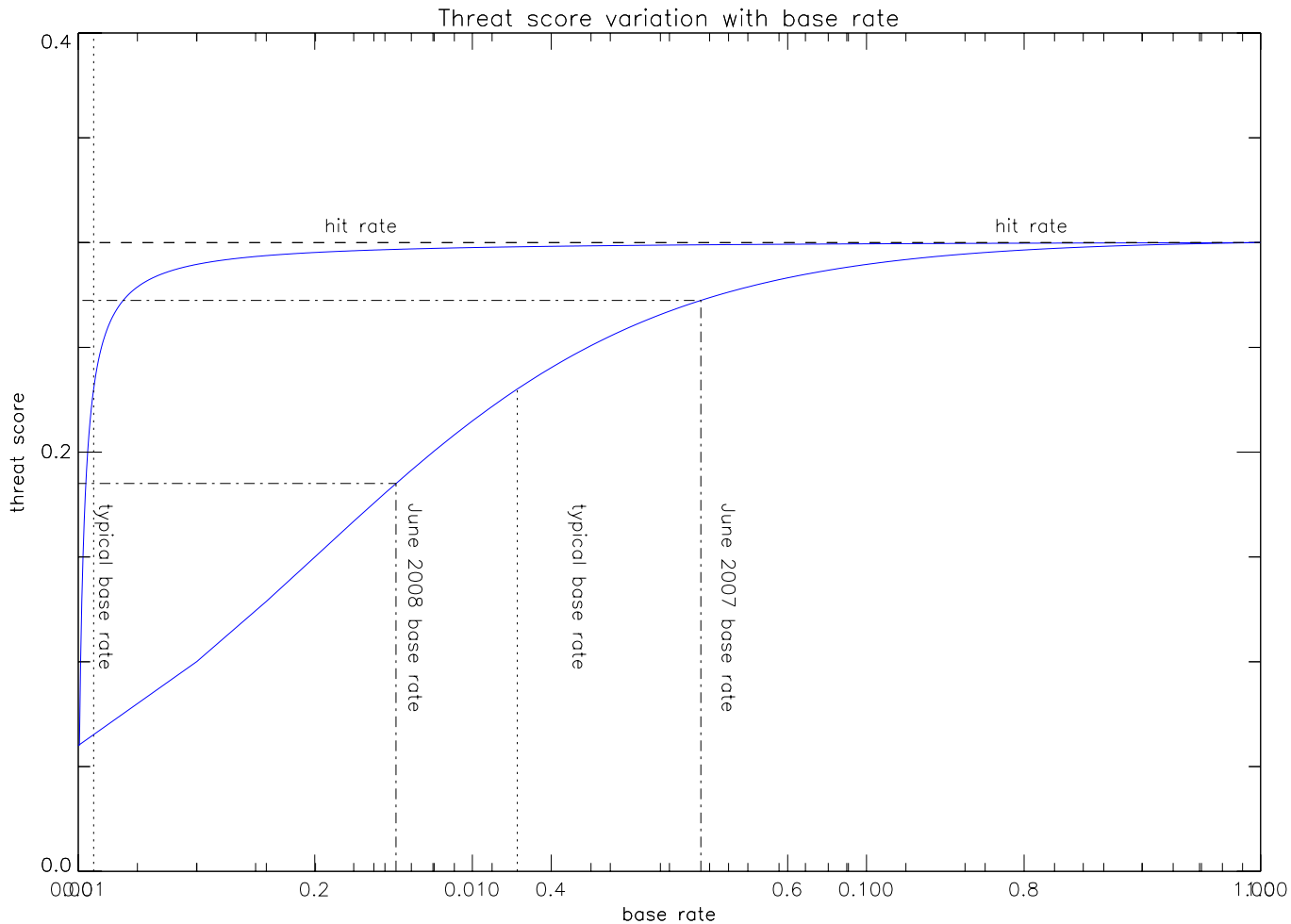


Base rate, p



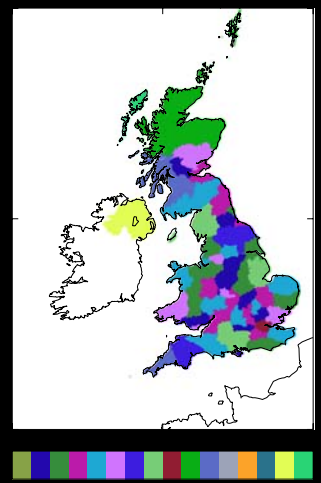
Base rate, p

Variation of threat score for heavy rain with base rate



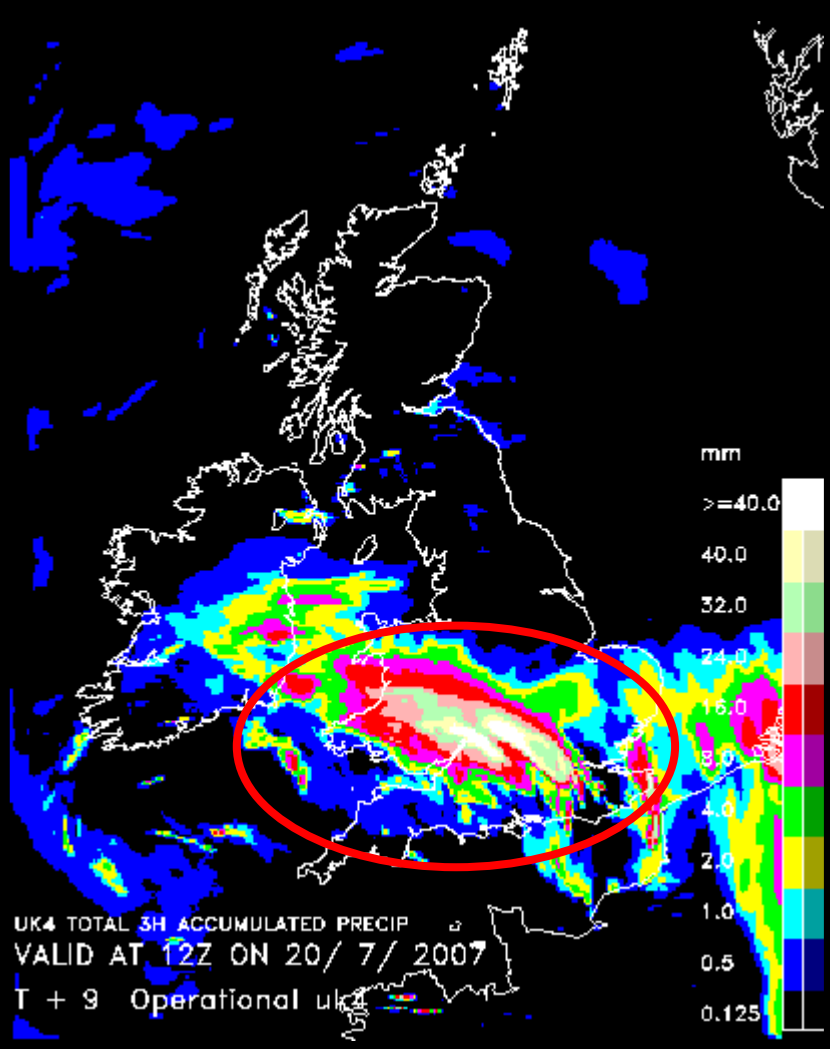
Model forecasts for heavy rain flash (15mm/3h)

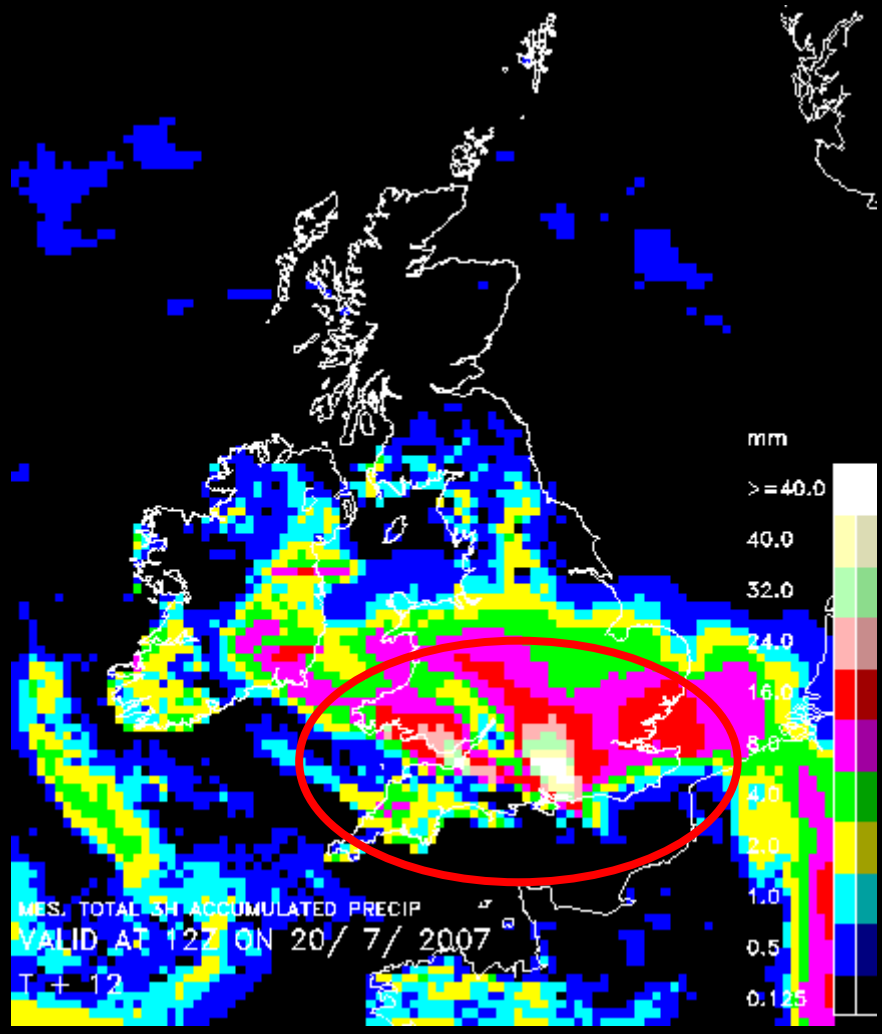
- NAE (12km) and UK4 4km forecasts over UK for July 2007
- Compared to radar composites (5km)
- Verified at all 3h periods 0-3, 1-4, 2-5 … 33-36
- Model forecasts verified at 12km and 5km (UK4)
- thresholds 5mm, 10mm, 15mm (/3h)
- Verified
  - At all grid points with radar ppn
  - for “county regions”
  - at least one location per county



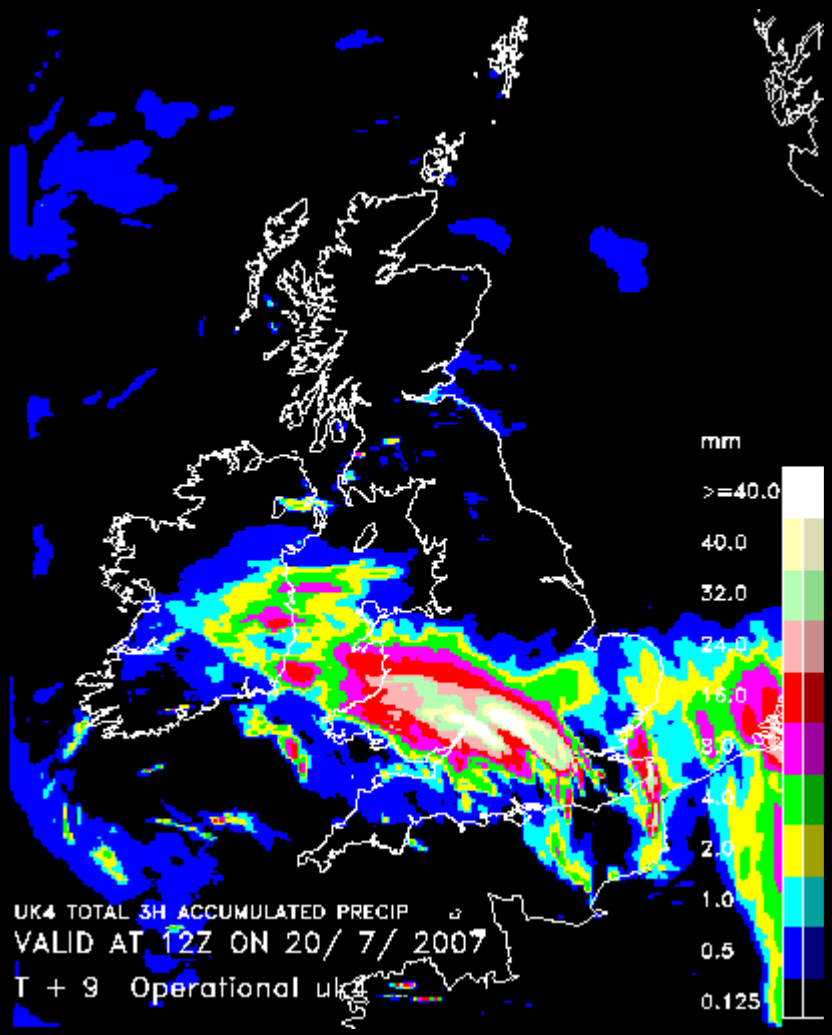
3h accumulations -20 July 12Z

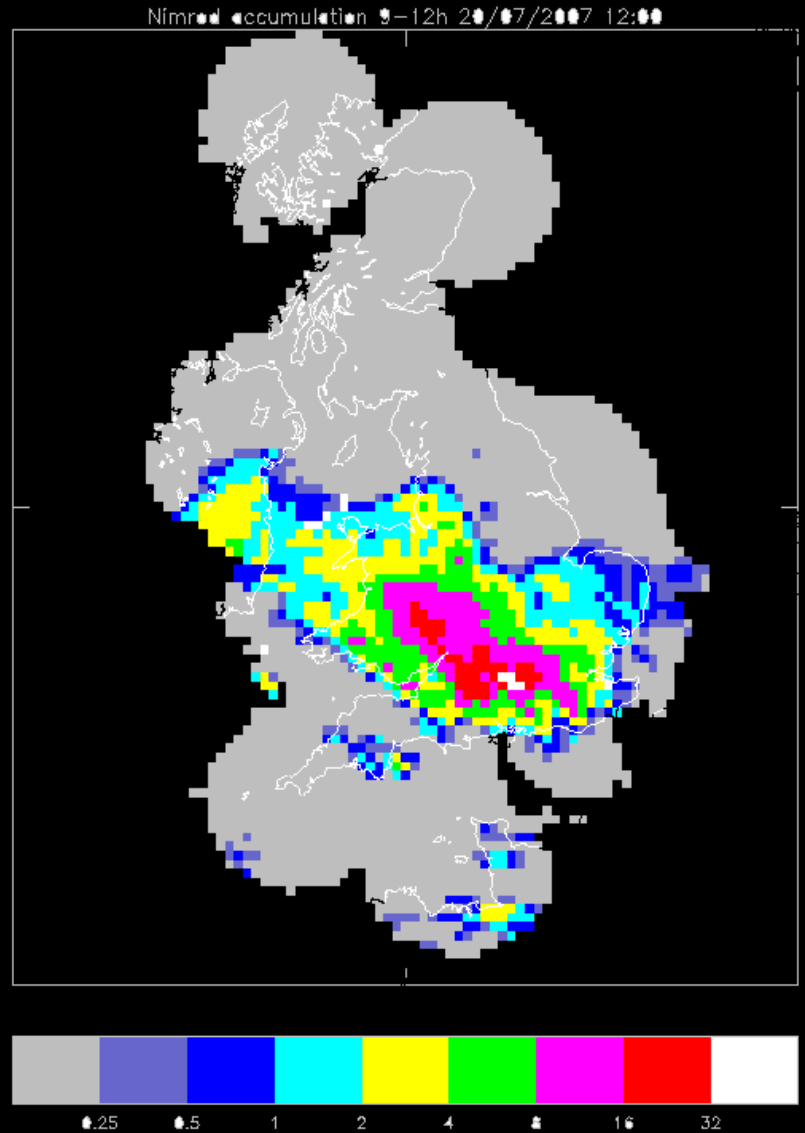
4km (6-9h) 12km (9-12h)



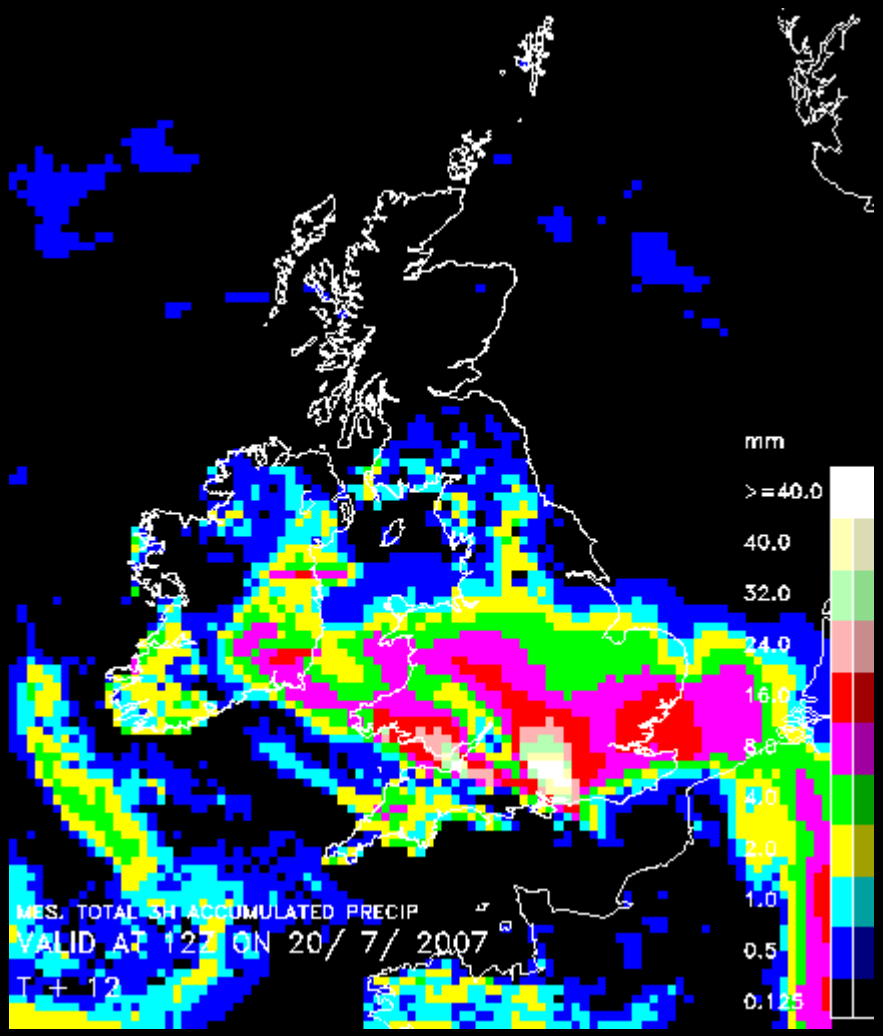


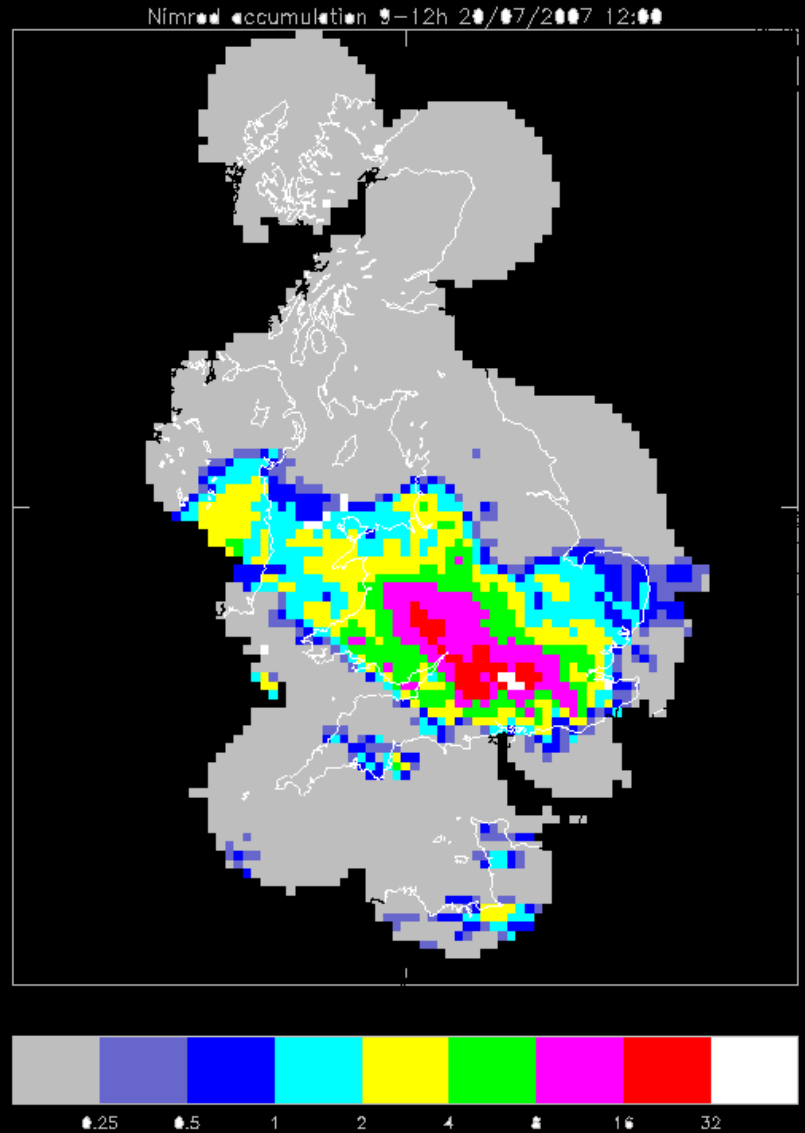
Radar 3h accumulation on 12km grid





Radar 3h accumulation on 12km grid

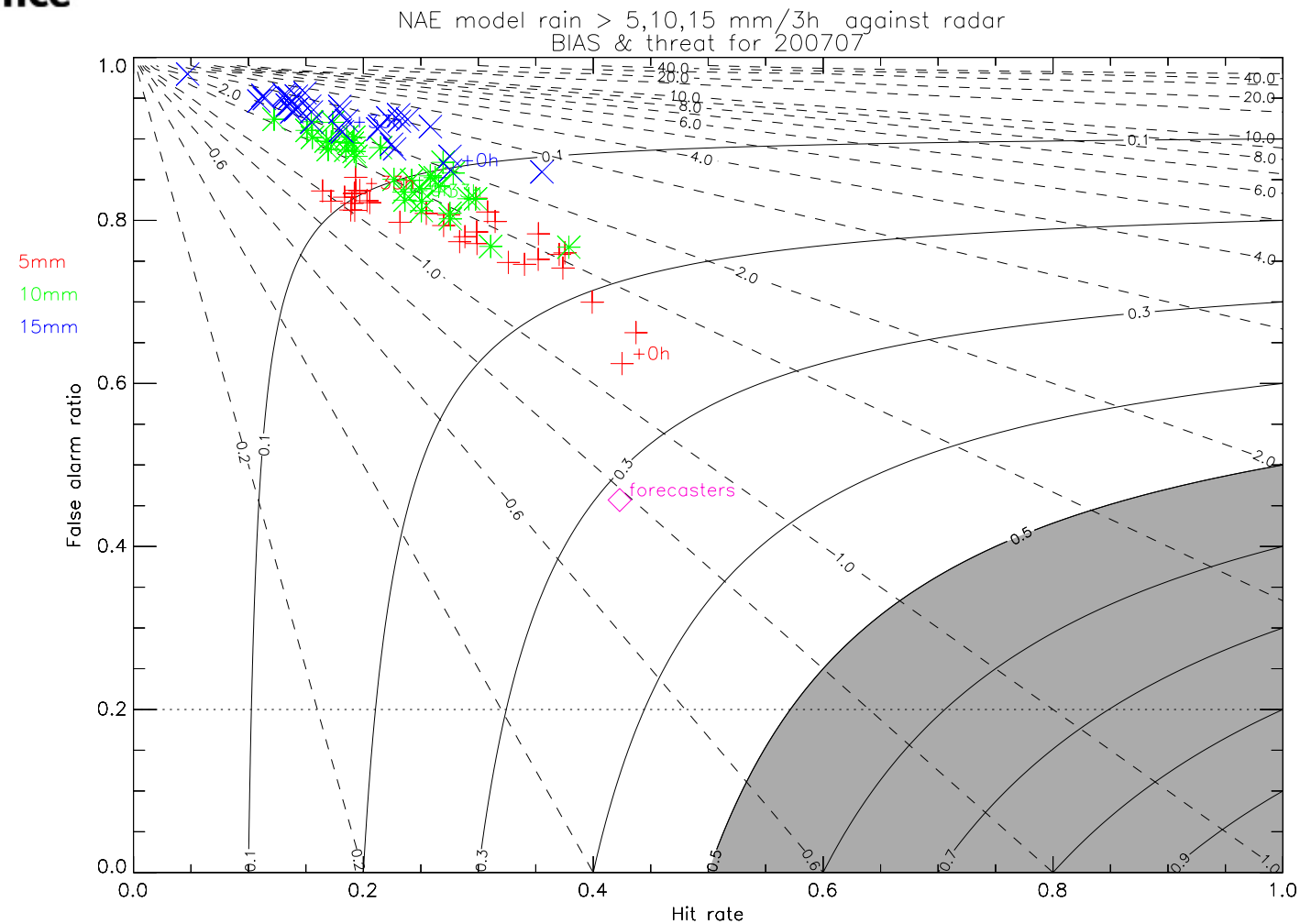




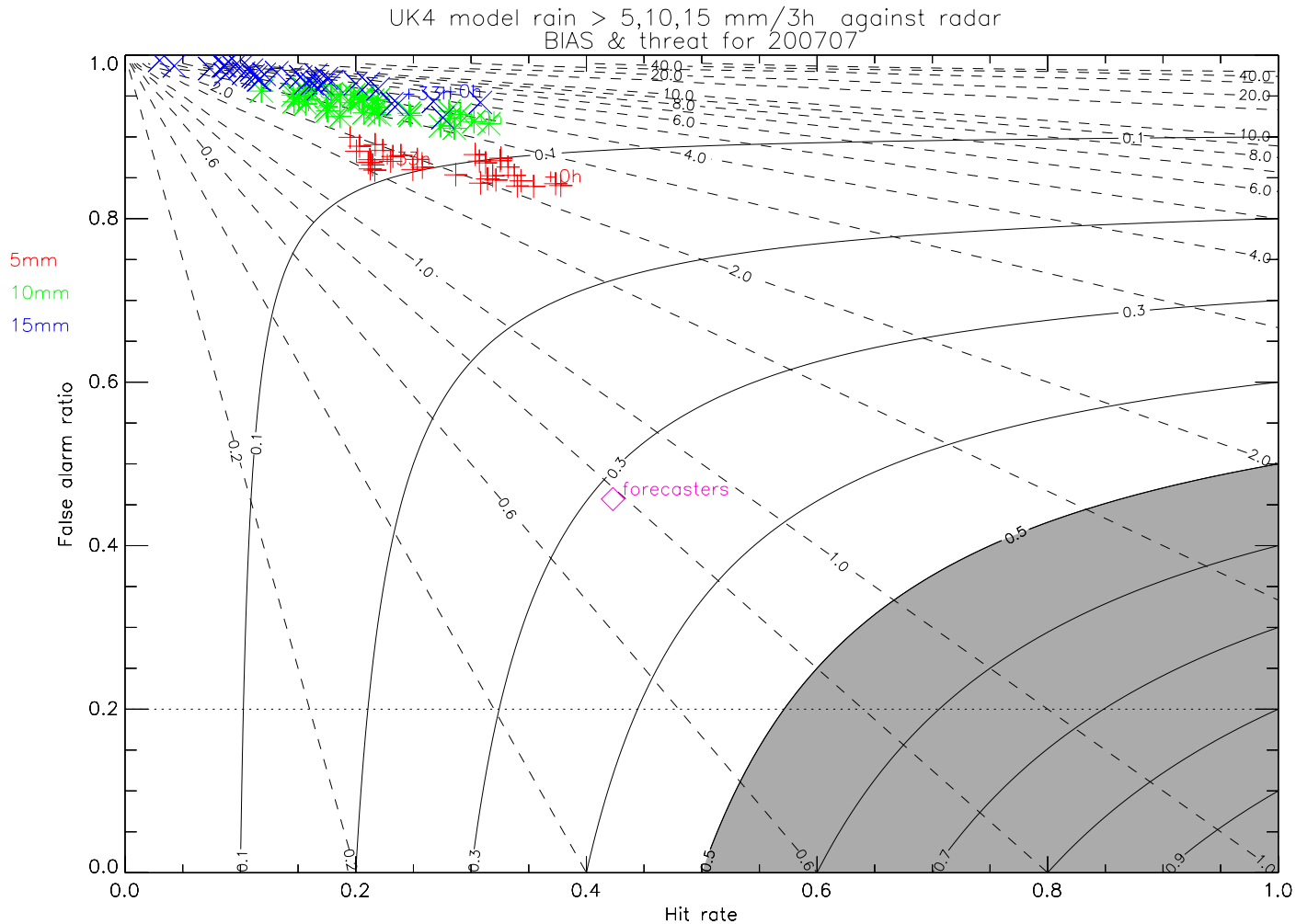
NAE(12km) & UK 4km models (12km grid verification) July 2007

NAE

12km



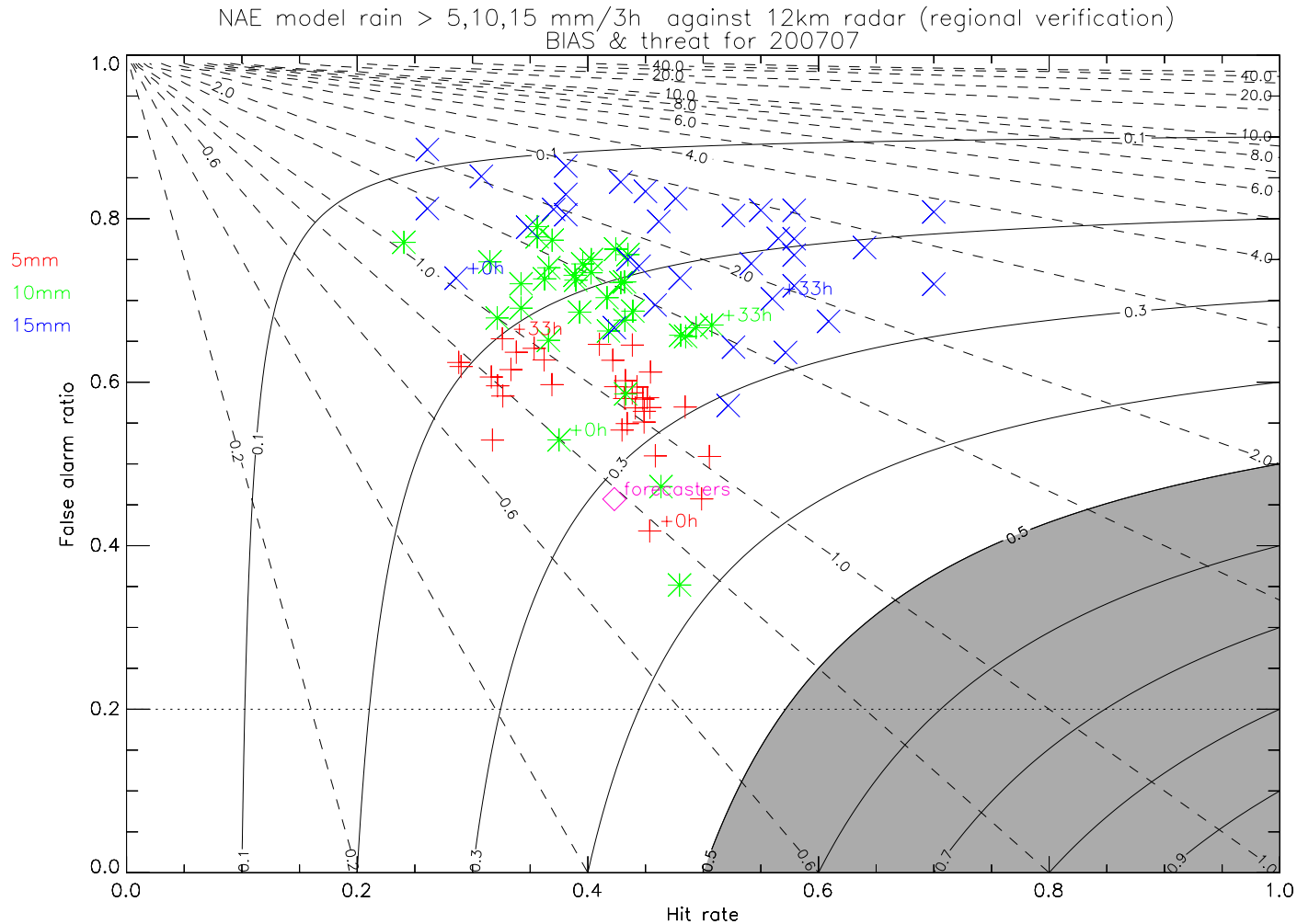
NAE(12km) & UK 4km models (12km grid verification) July 2007



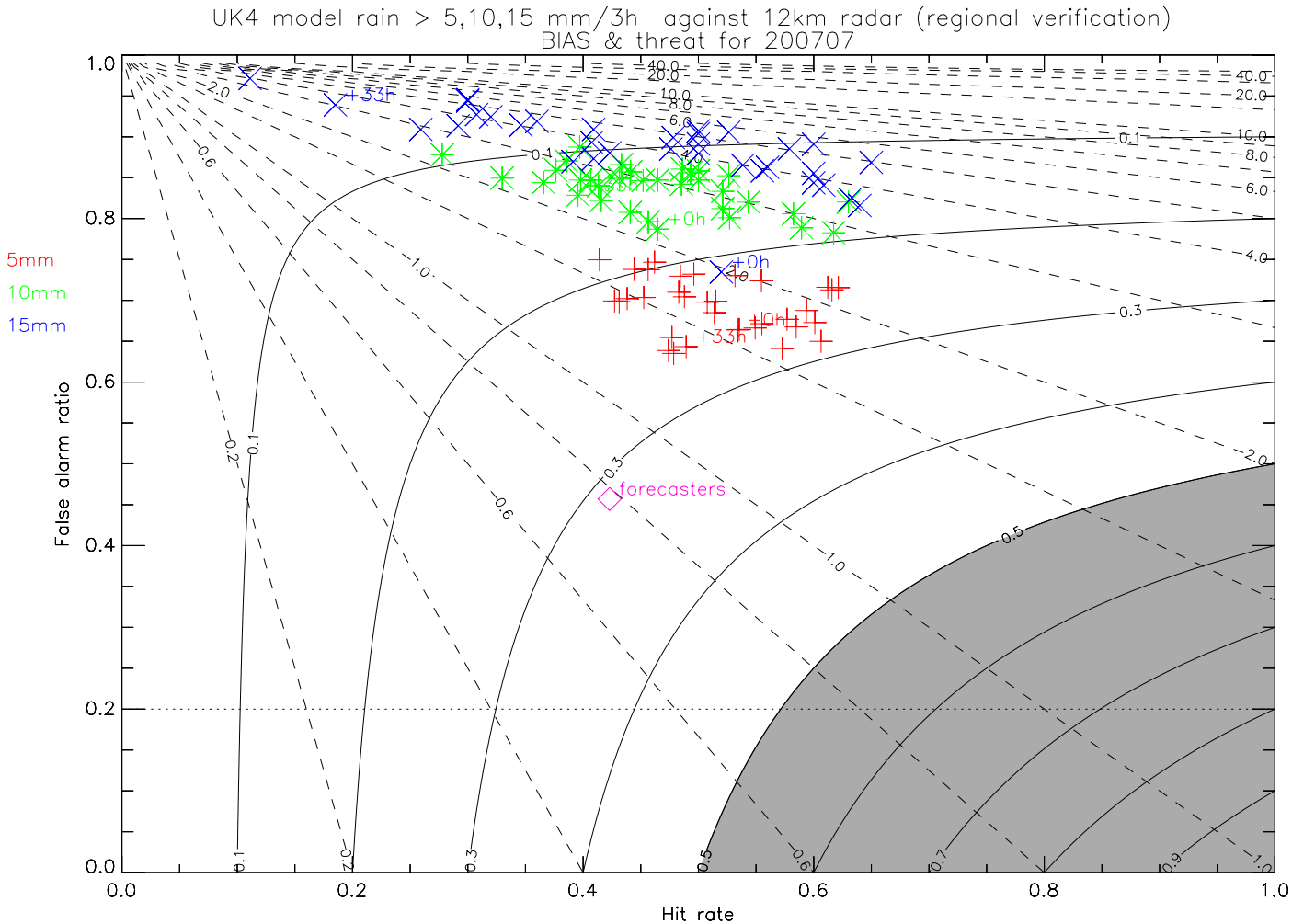
4km

NAE(12km) & UK 4km models (regional verification) July 2007

NAE
12km



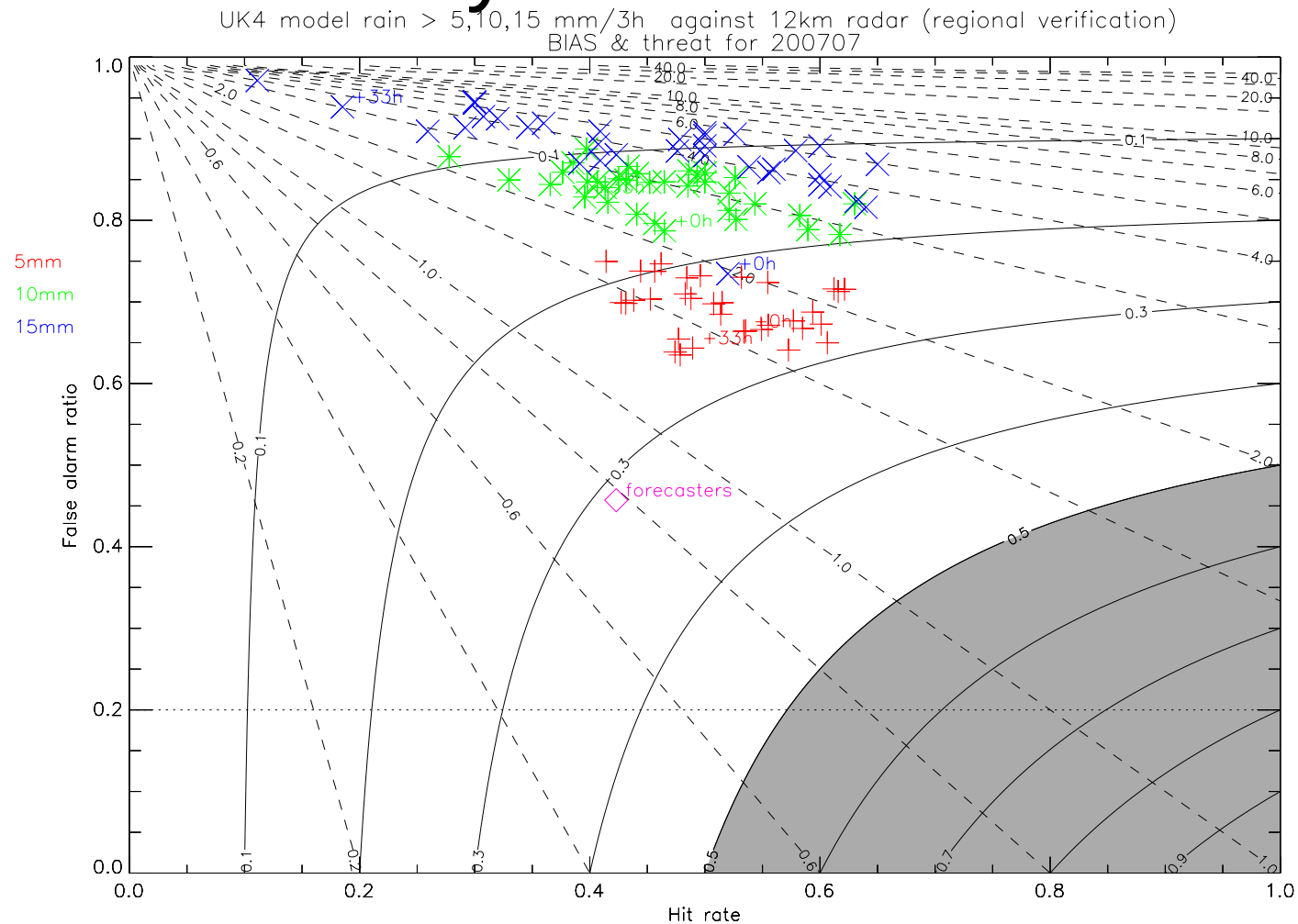
NAE(12km) & UK 4km models (regional verification) July 2007



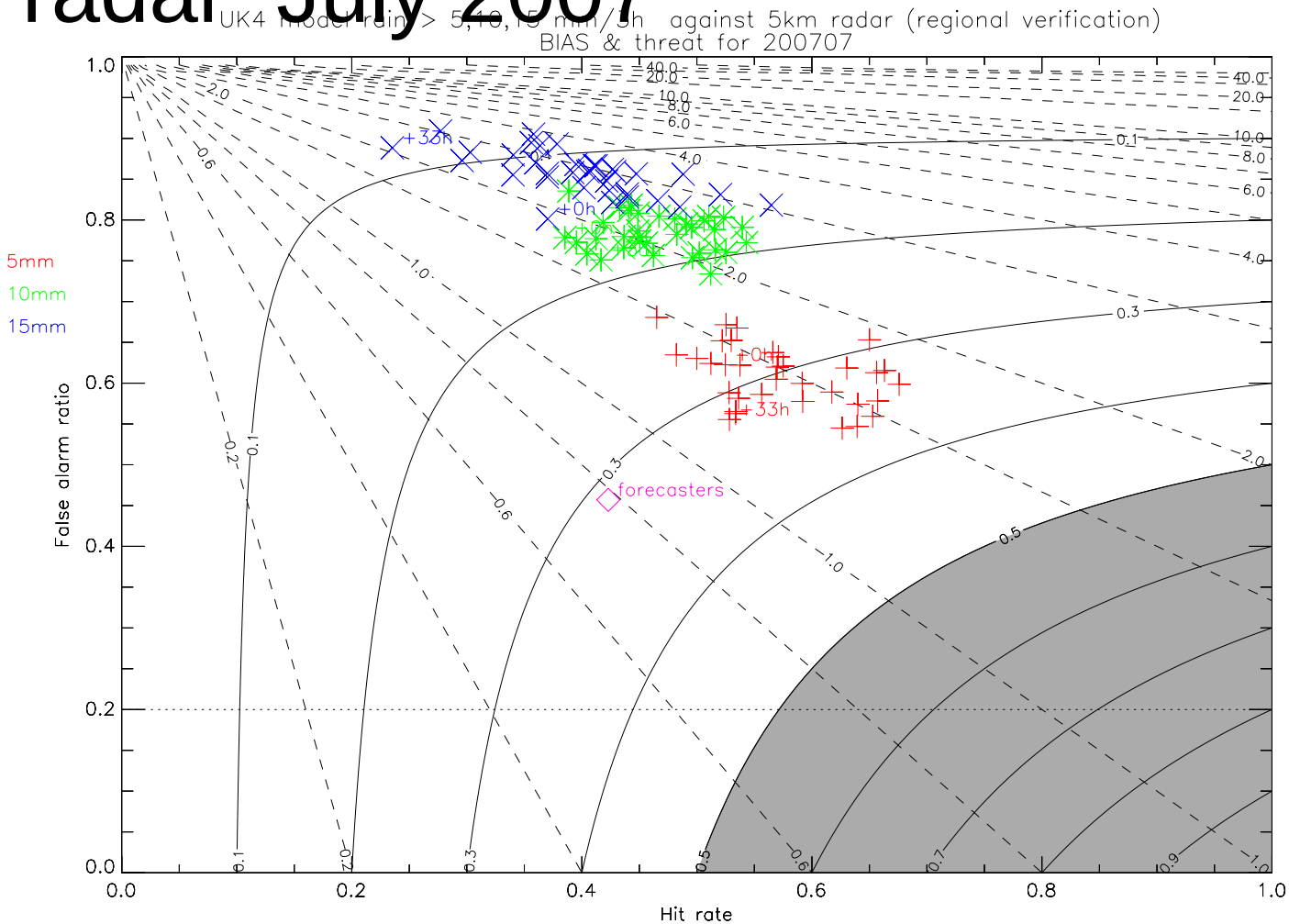
4km

UK 4km model (regional verification) v 12km & 5 km radar July 2007

12km radar

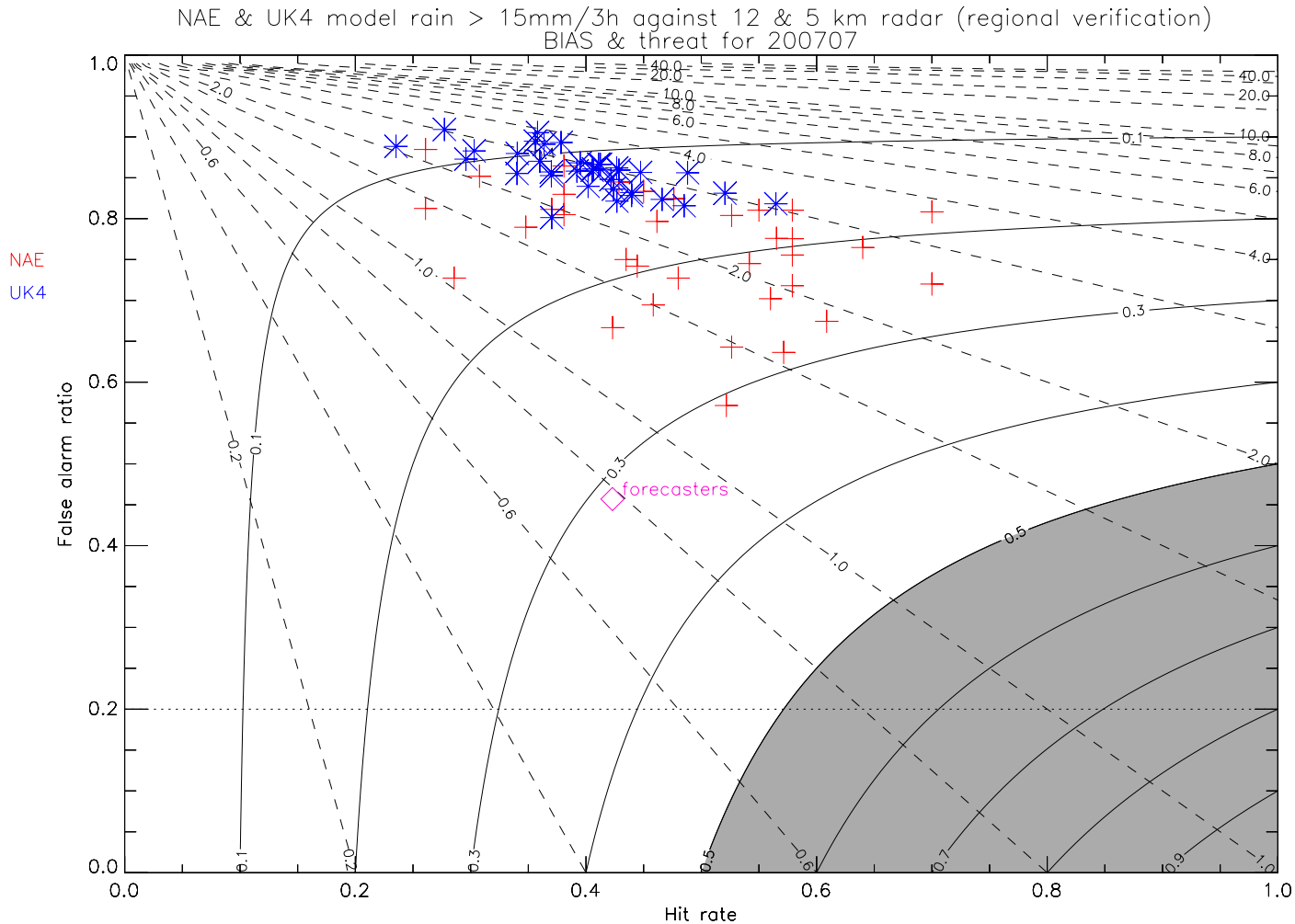


UK 4km model (regional verification) v 12km & 5 km radar July 2007

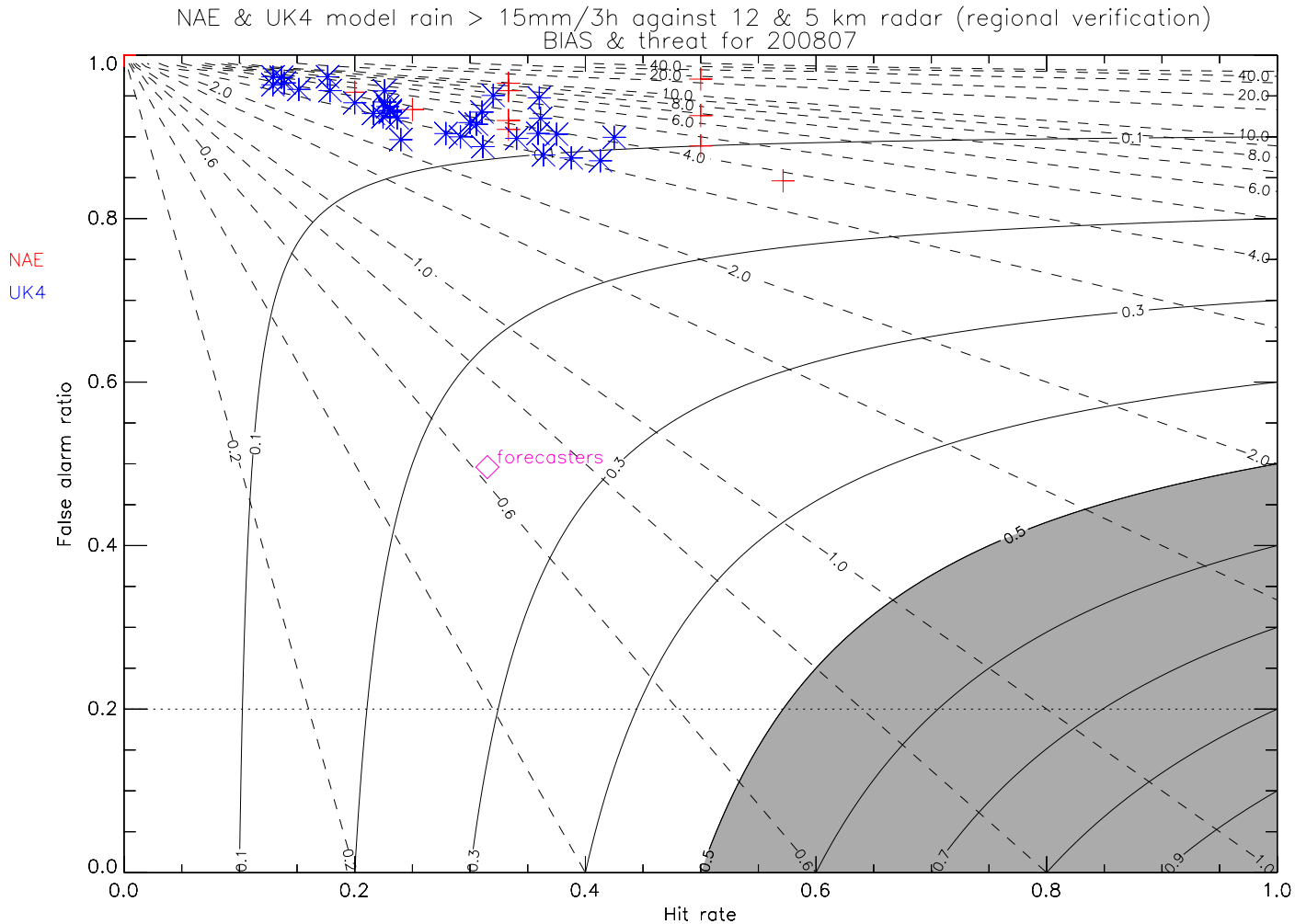


5km radar

Comparison NAE and UK4 15mm/3h July 2007



Comparison NAE and UK4 15mm/3h July 2008



Conclusions -1

- Useful summary plots
  - False alarm v hit rate with Bias, threat score contours
- Single (threat) score inadequate
- Always show bias – scores may be hedged
- Scores depend on “truth” type
- Regional verification problems
  - Variation in area
  - Obs missing
  - Detection depends on no. of locations for event & frequency

Conclusions -2

- Confidence (80%) generally not achieved by forecasters
- Deterministic limit – not generally satisfied
- Forecasters improve on raw model guidance
- Threat score very dependent on base rate
  - Perhaps use Extreme dependency score (EDS) -need “d”
- Models – heavy rain
  - Better performance July 2007 than July 2008
    - Larger base rate
  - NAE 12km better than UK 4km ?
    - Need to look at more months