



Met Office

Verification of **operational** precipitation forecasts over the UK

Clive Wilson

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



Contents

- Aims
- Standard operational verification
- Radar composites and gridded gauges
- Fuzzy scores
- Conclusions



Aims

- Operational verification of 40km, 12km and 4km models uses synoptic stations and nearest gridpoint
- 12km and 4km also routinely verified against radar analyses using common 12km grid
- Can we find common signals and explain differences ?
- Compare radar and climatological gauge analyses to estimate uncertainty
- (Use idealised tests to understand
 - Domain size; Fractional pattern; sub-sampling)
- Is the higher resolution model better?



Primary Operational Forecast Systems

UK 4km

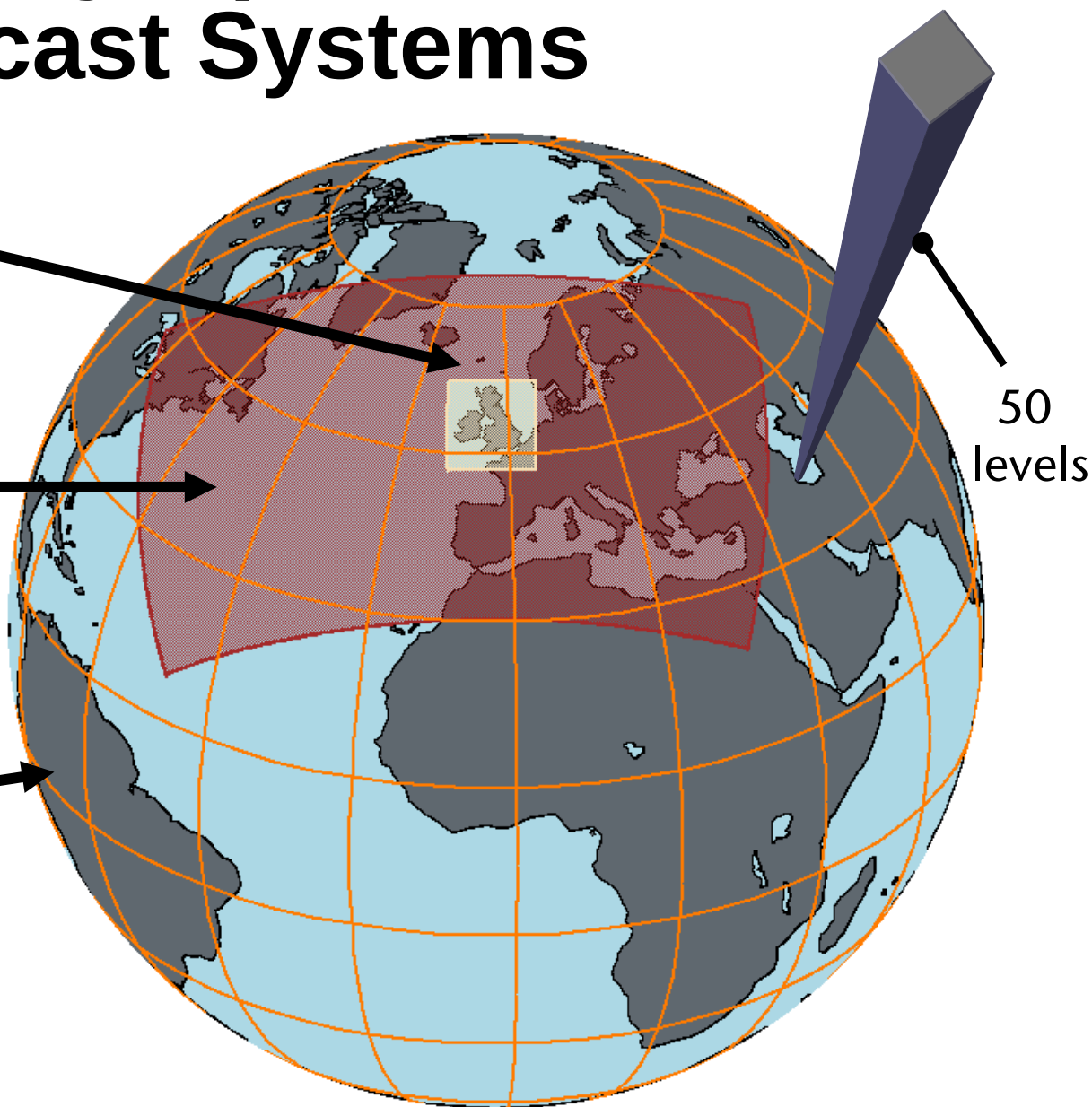
- 36 hour forecast
- 70 levels
- 4 times per day

Regional 12km

- 60 hour forecast
- 38 levels
- 4 times per day
- +EPS 24km, 24member

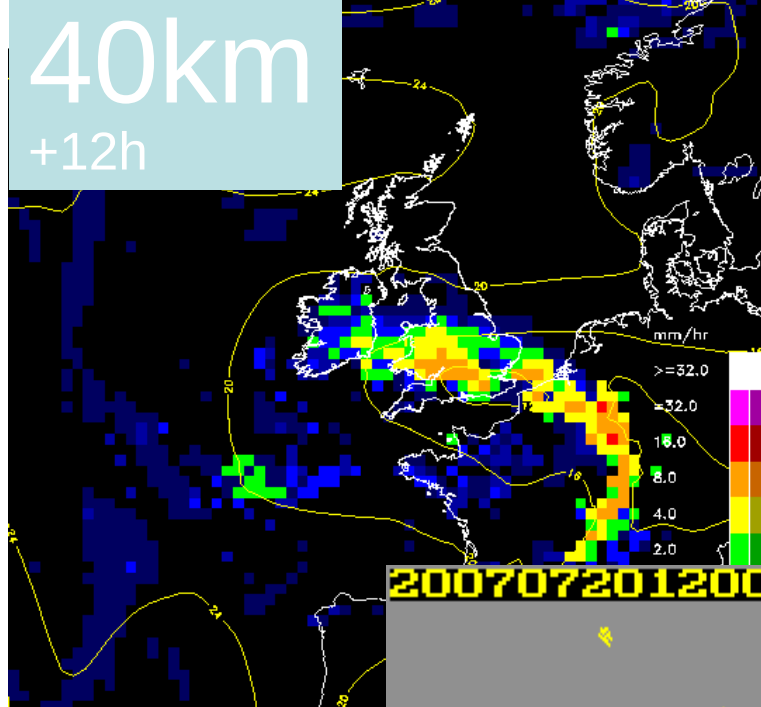
Global 40km

- 60 hour forecast twice/day
- 144 hour forecast twice/day
- +EPS 24member, 90km

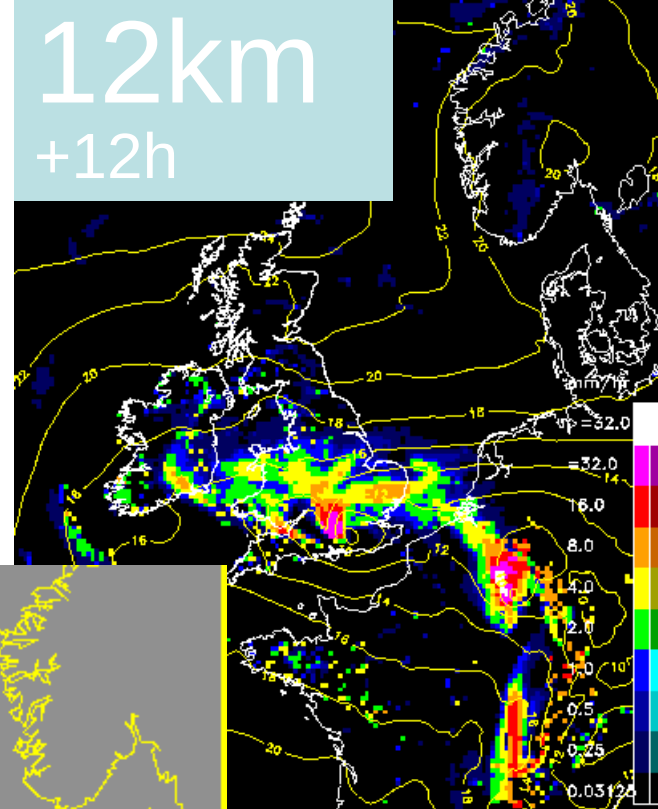




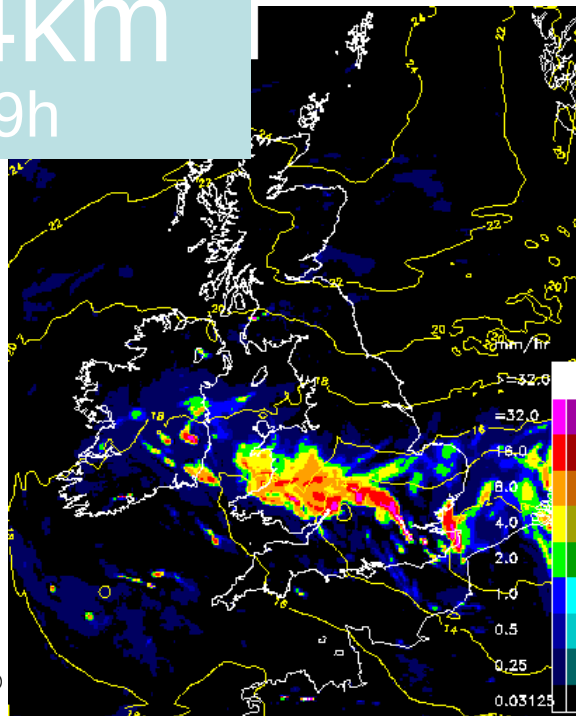
40km
+12h



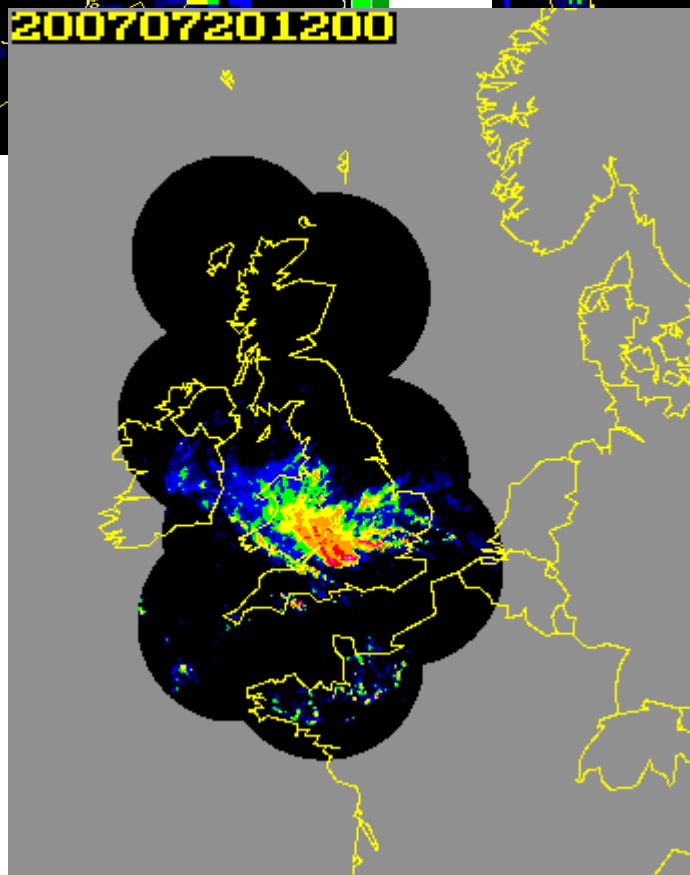
12km
+12h



4km
+9h



200707201200

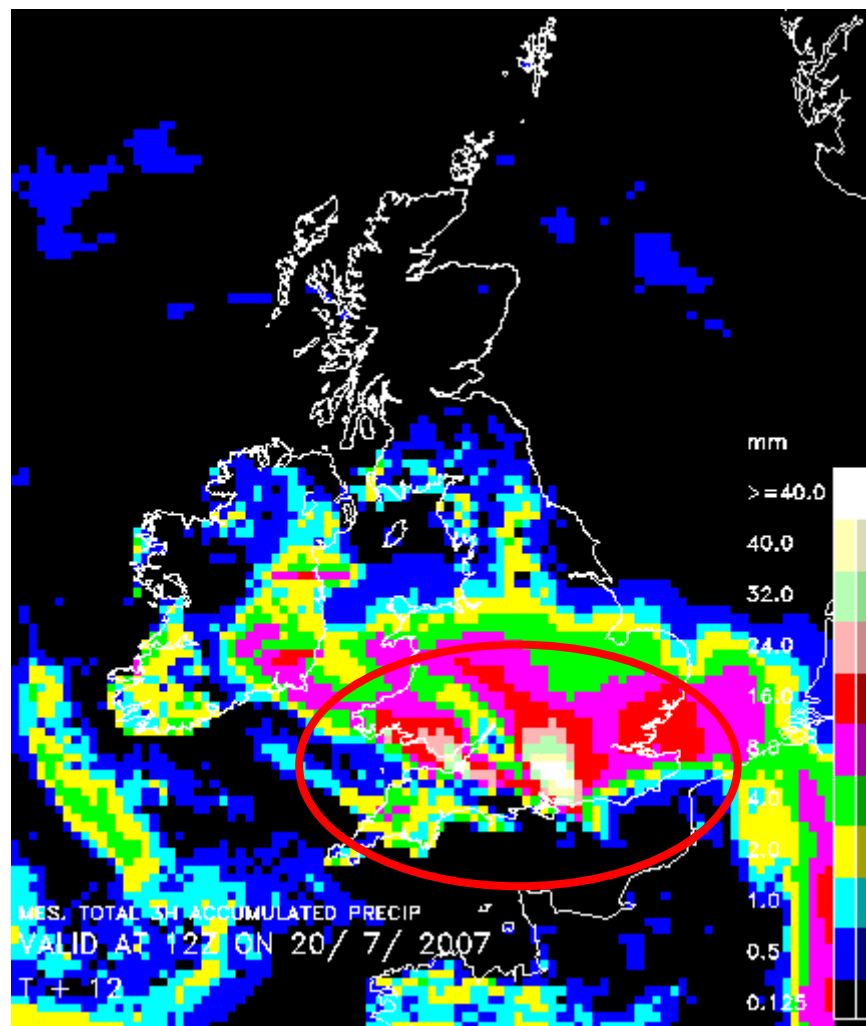
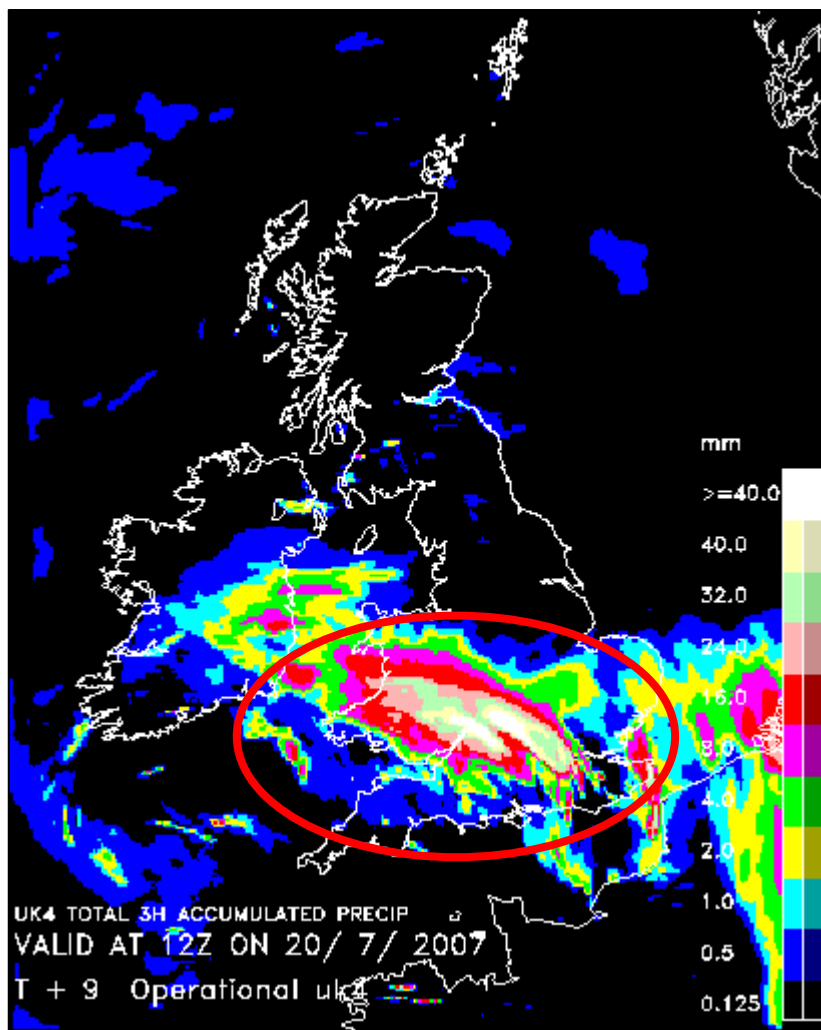


20 July
2007
Floods

3h accumulations

4km (6-9h)

12km (9-12h)





Met Office

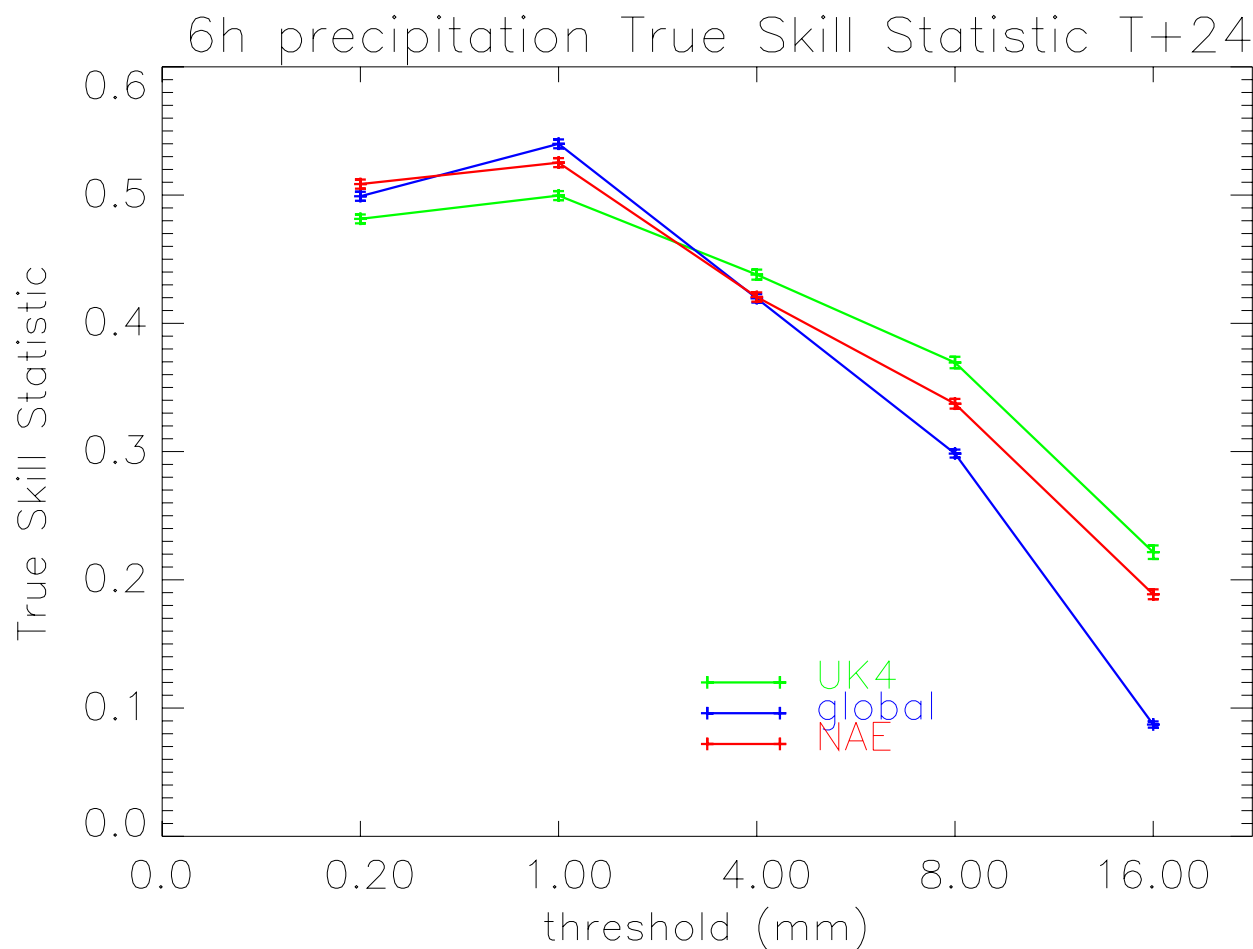


Standard Verification

Against synoptic station gauges – Inconsistent signals !

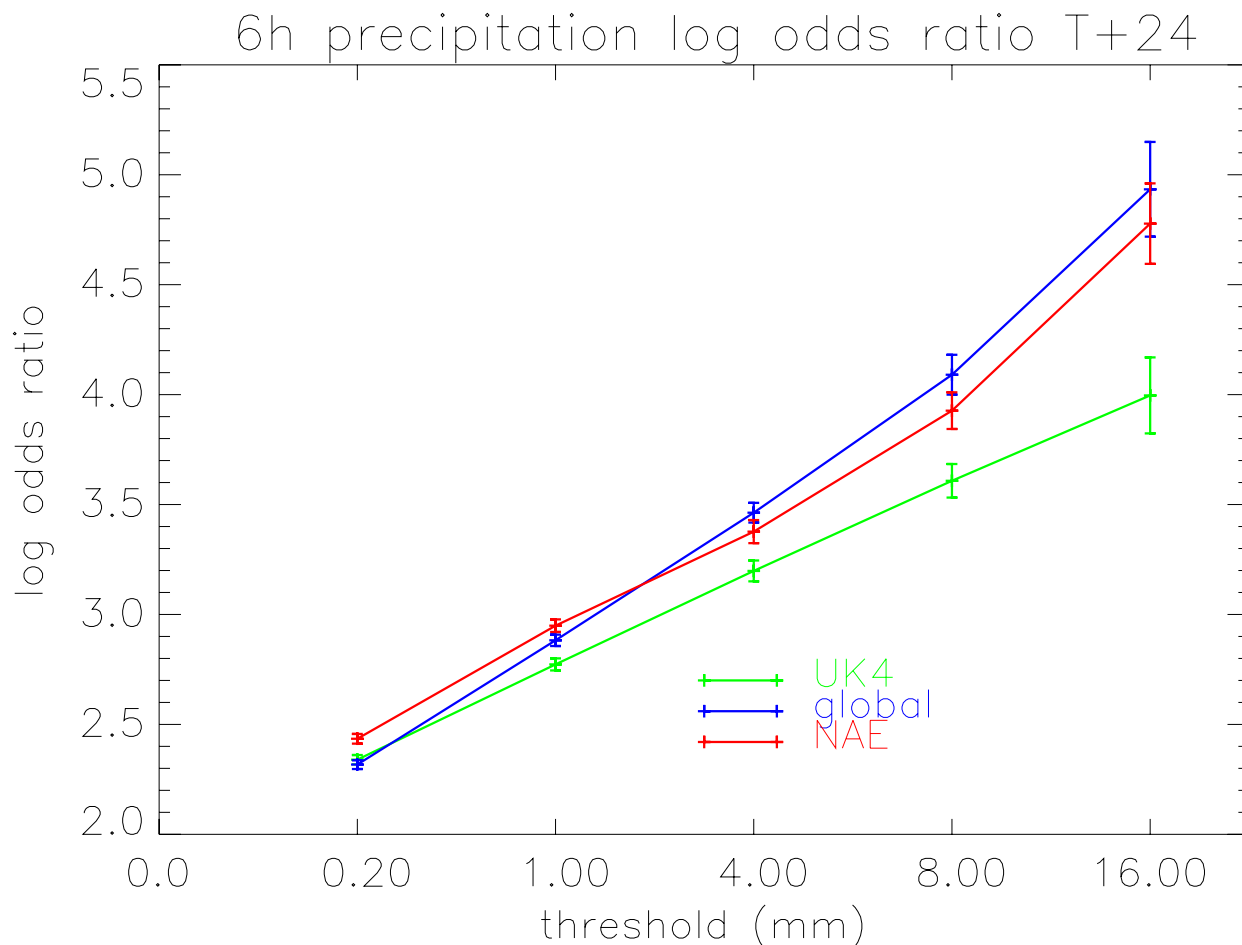
Standard UK area verification against surface stations- Nearest grid point – Dec06-Sep08

- True SS=Hit rate – False alarm rate
- UK4 better at higher thresholds



Standard UK area verification against surface stations- Nearest grid point – Dec06-Sep08

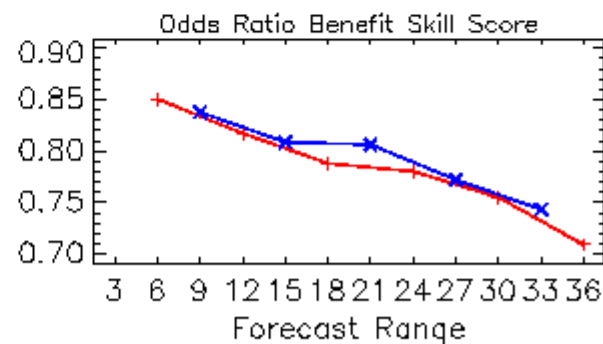
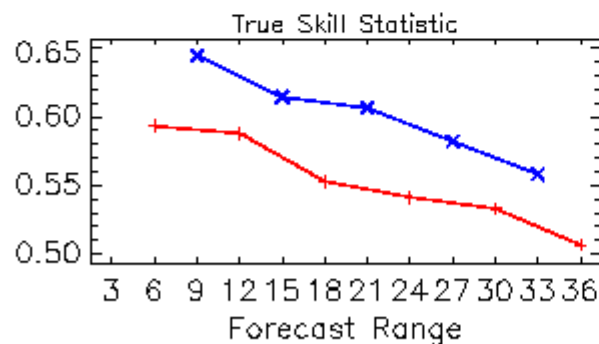
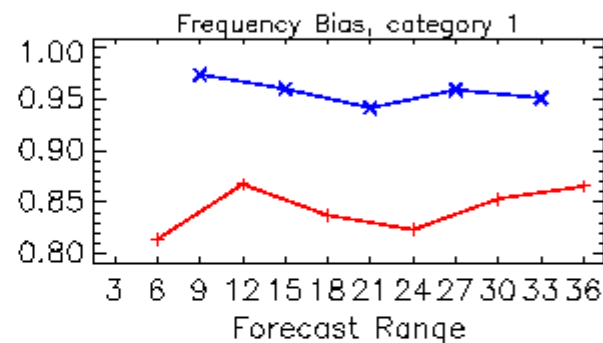
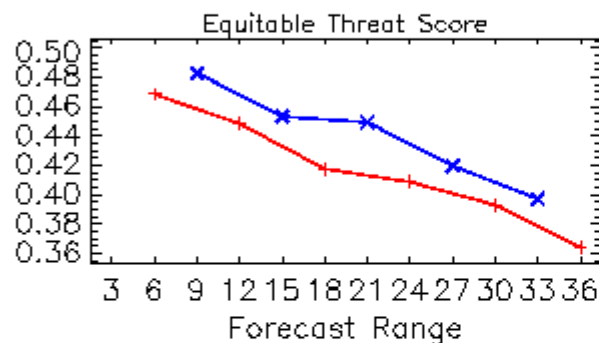
- Odds ratio increases as rarer events
- 40km model better at higher thresholds ?



Operational verification v synops (Dec 2006)

Combined times: 6hr Precip Accm ($\geq 1.0\text{mm}$) (Corrected obs)
Combined dates from 01/12/2006 to 31/12/2006: Combined stations: Surface Obs

Cases: + Operational NAE x Operational UK4





Met Office

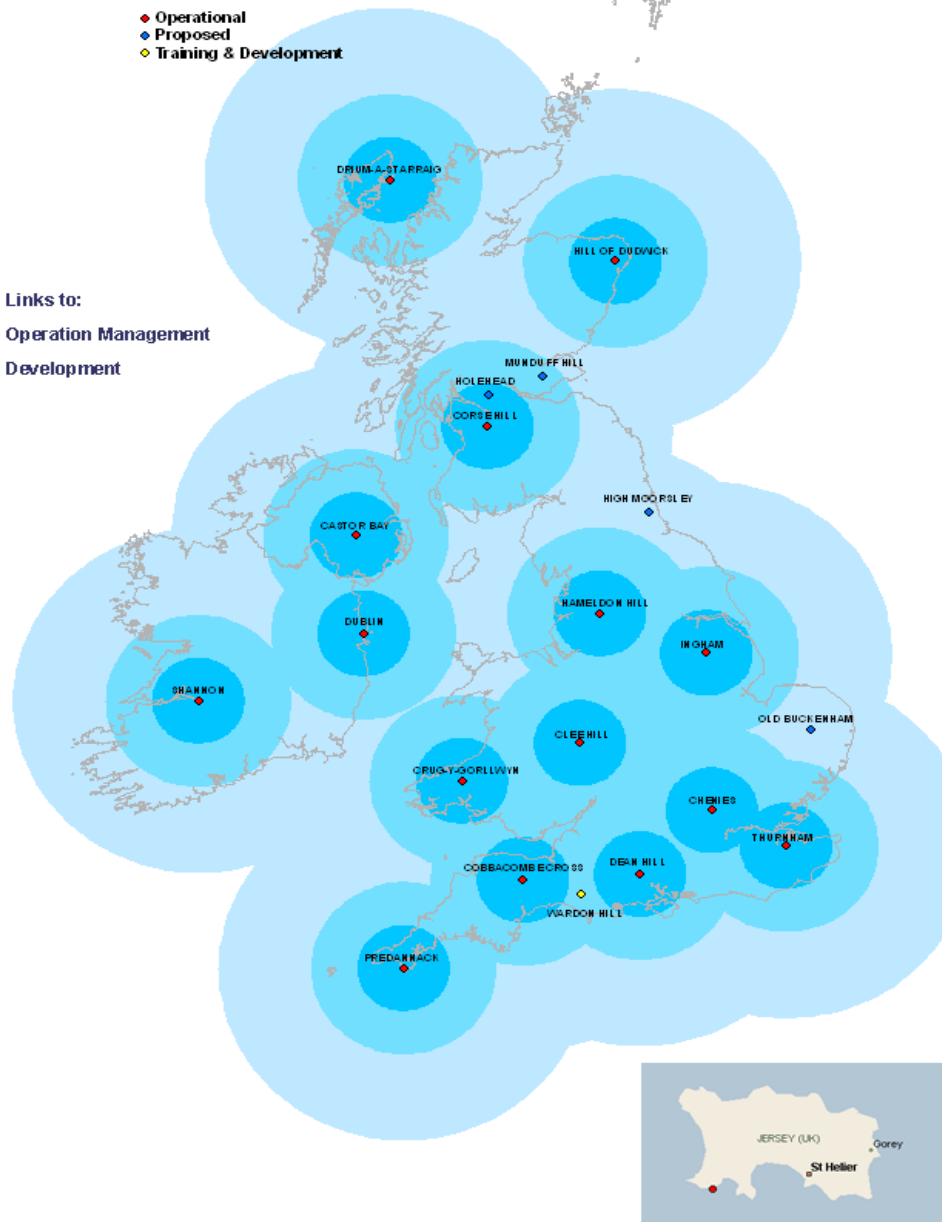


Radar composites & Gridded gauges

UK Weather Radar Network

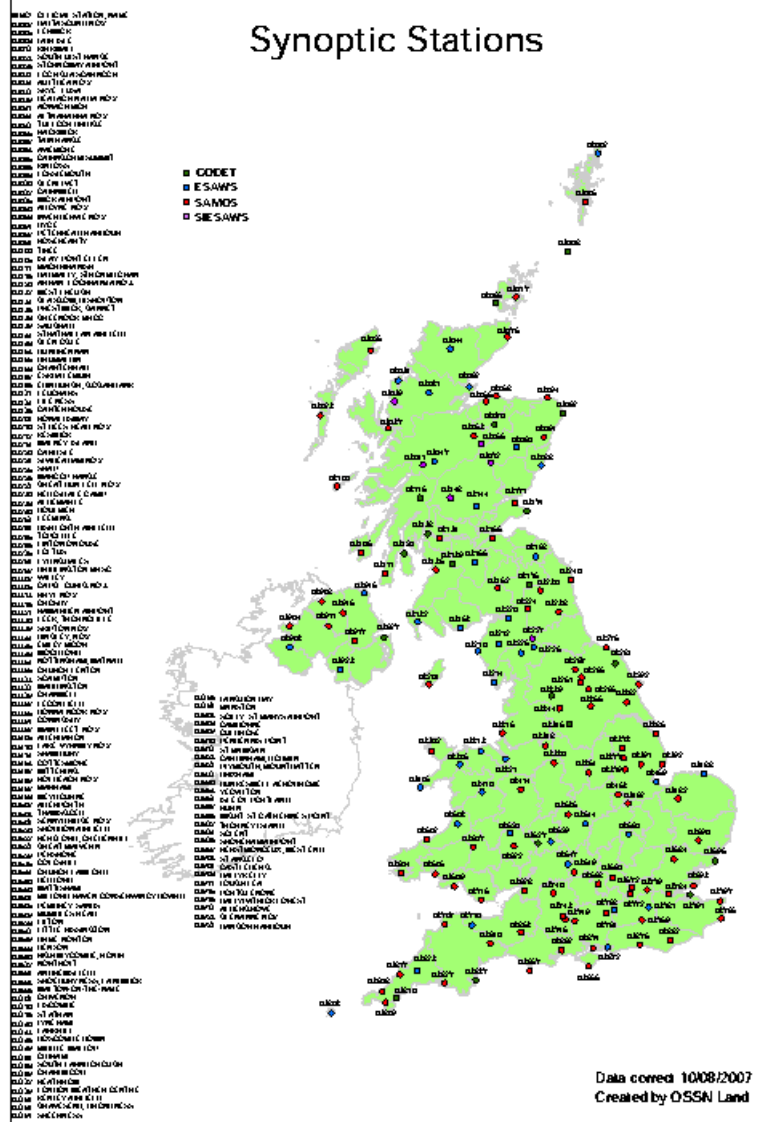
- Operational
- Proposed
- Training & Development

Links to:
Operation Management
Development



Synoptic Stations

- CODEY
- ESAMPS
- SAMOS
- SIE'SAMPS



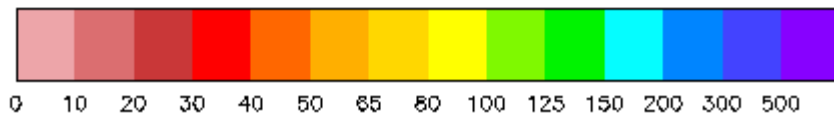
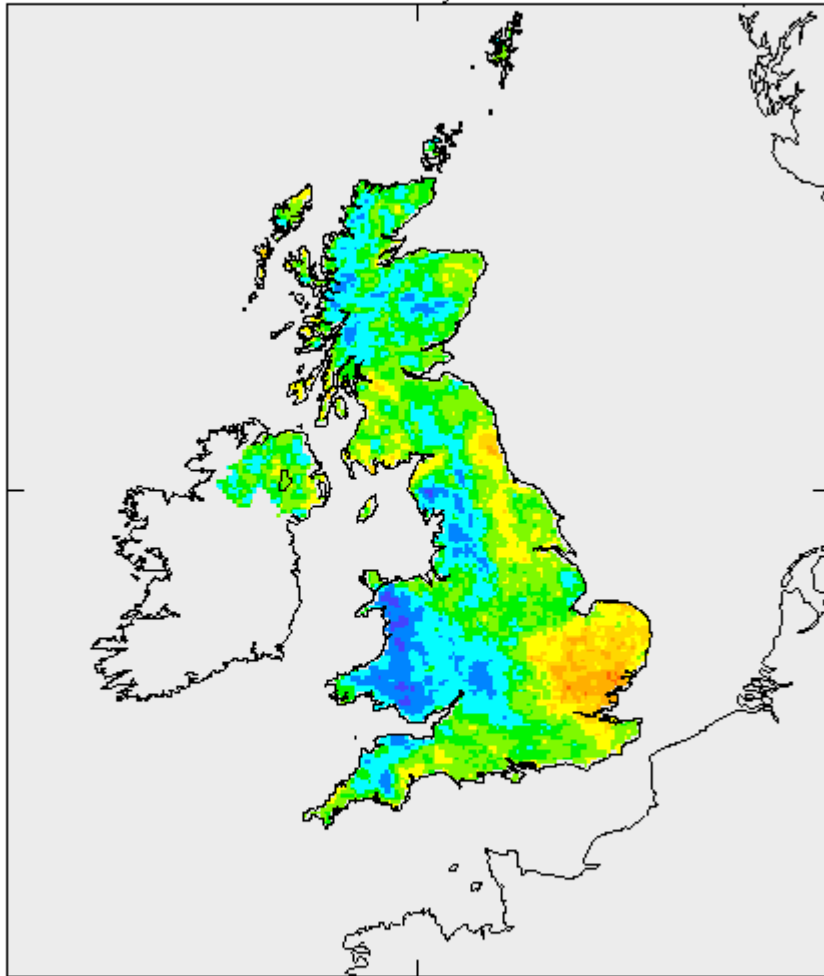
Data correct 10/08/2007
Created by OSSN Land

Uncertainties

- Synoptic network sparse ~25- 50km
 - Nearest grid point model forecast
- Radar estimates $\sim -50\%/+100\%$ error
 - Averaged to 3x grid length = 15km
- Climatological precipitation stations
 - ~4400
 - 7 x7 km typical spacing
 - Gridded analysis -5x5km (Perry & Hollis,2005,Int J Climatol)
 - Monthly
 - daily

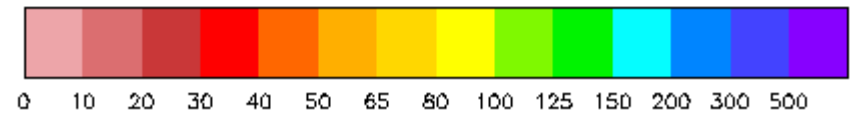
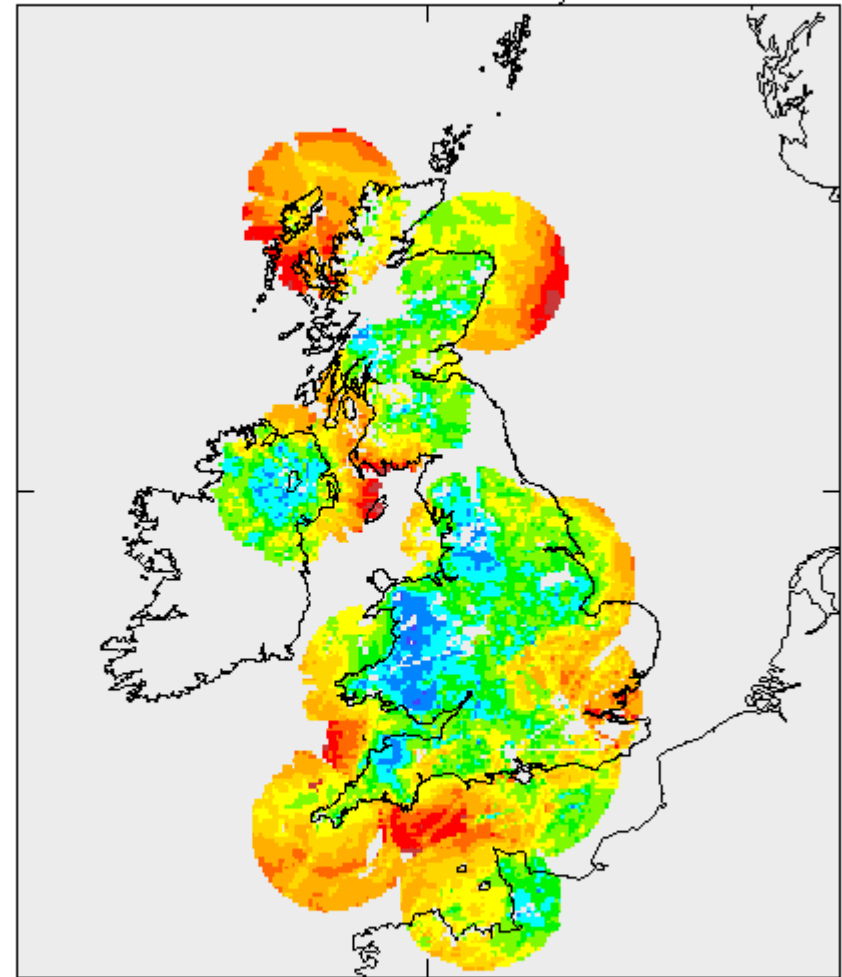
Gridded Gauges July 2007

total accumulation for July 2007_Actual_final.dat



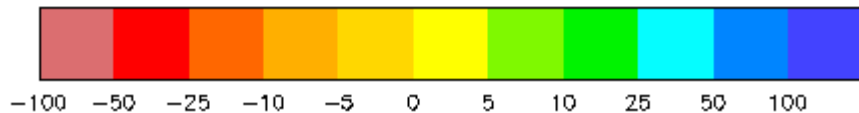
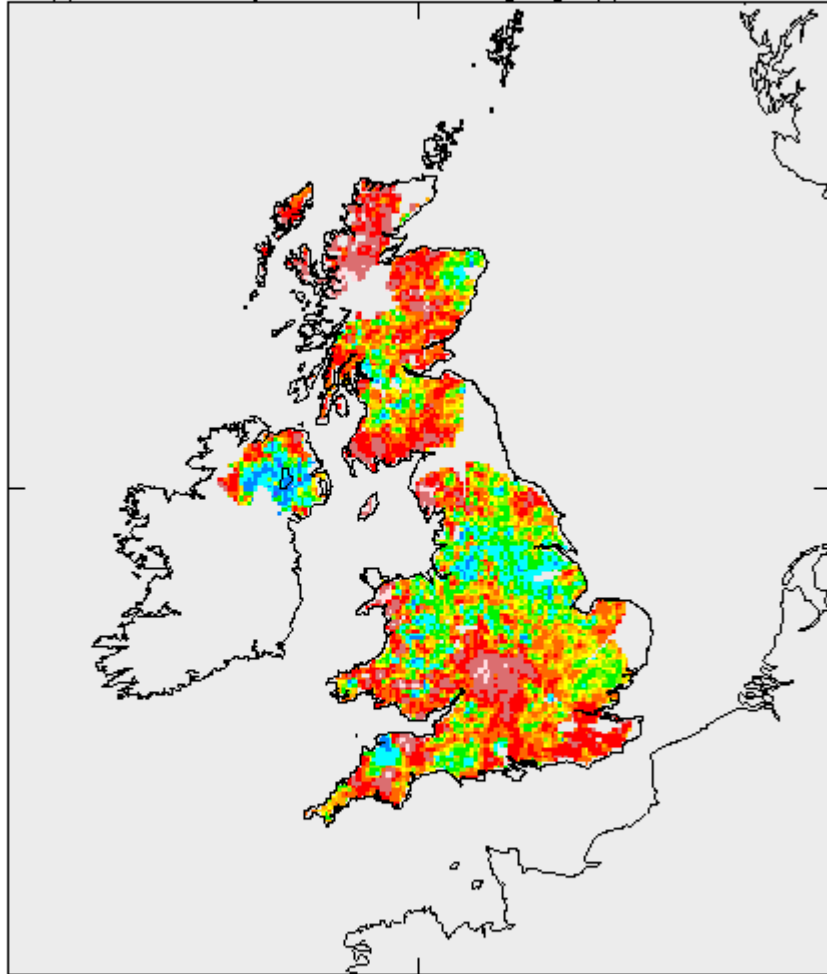
Radar

total accumulation for July 2007



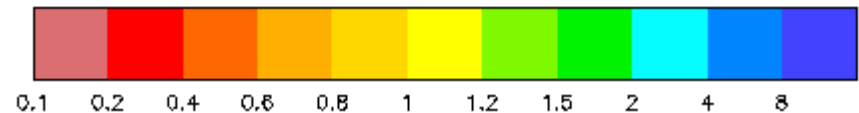
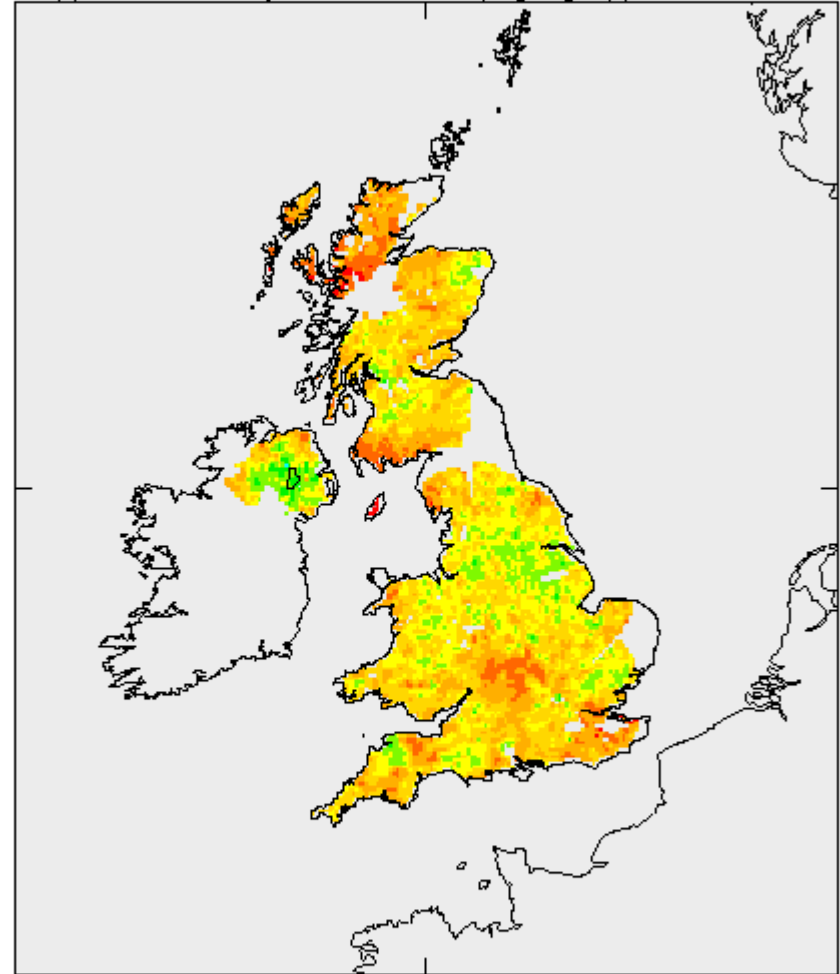
Radar – gauge (mm) July 2007

total ppn diff for July 2007nimaccu – gauge_ppn_Actual_final.dat



Radar/gauge

total ppn bias for July 2007nimaccu / gauge_ppn_Actual_final.dat





Met Office

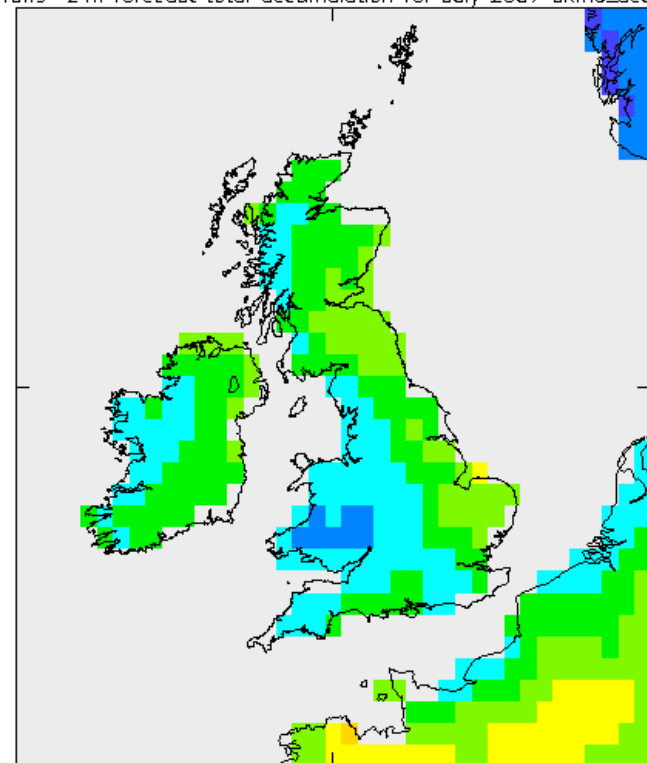
40 km

12km

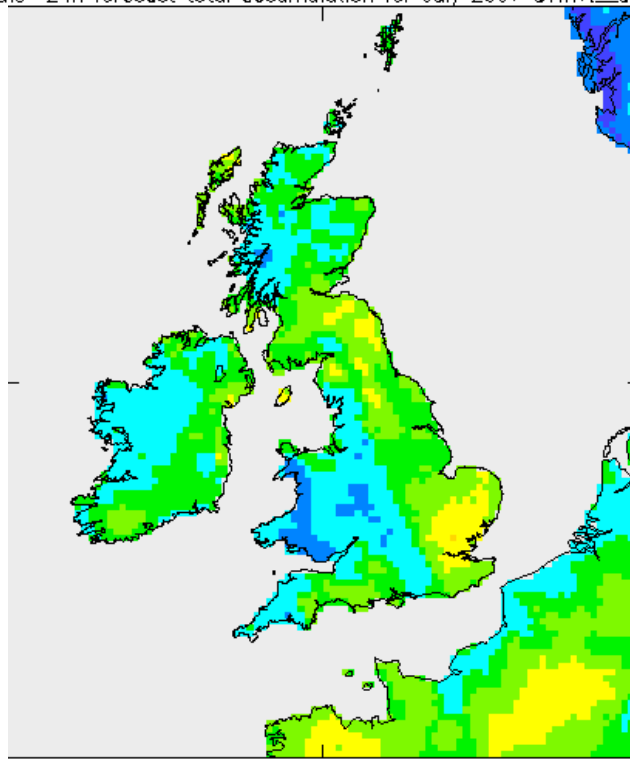
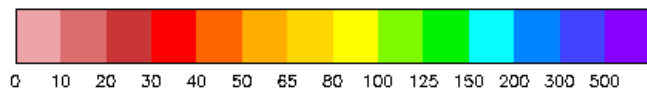
4km

Forecast monthly totals - July

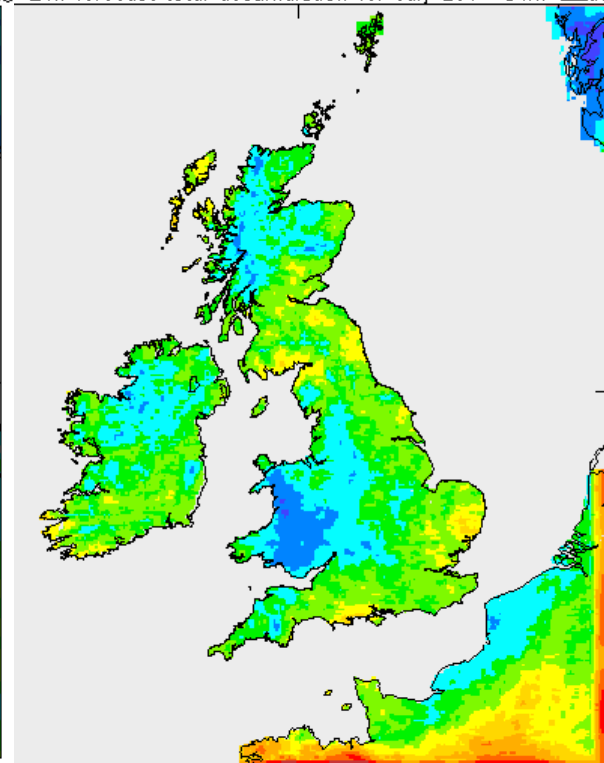
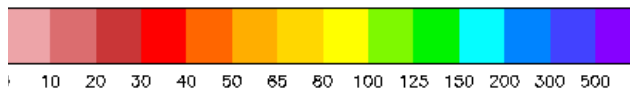
runs 24h forecast total accumulation for July 2007 ukmo_accus 24h forecast total accumulation for July 2007 QYmvt_acc 24h forecast total accumulation for July 2007 Q4mvt_acc



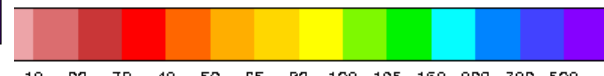
Mean=144.003 sd=39.0722 max=308.607 min=76.3321



Mean=144.485 sd=43.6954 max=412.872 min=79.4707



Mean=128.296 sd=50.7418 max=452.141 min=0.749023



Forecast monthly totals – July

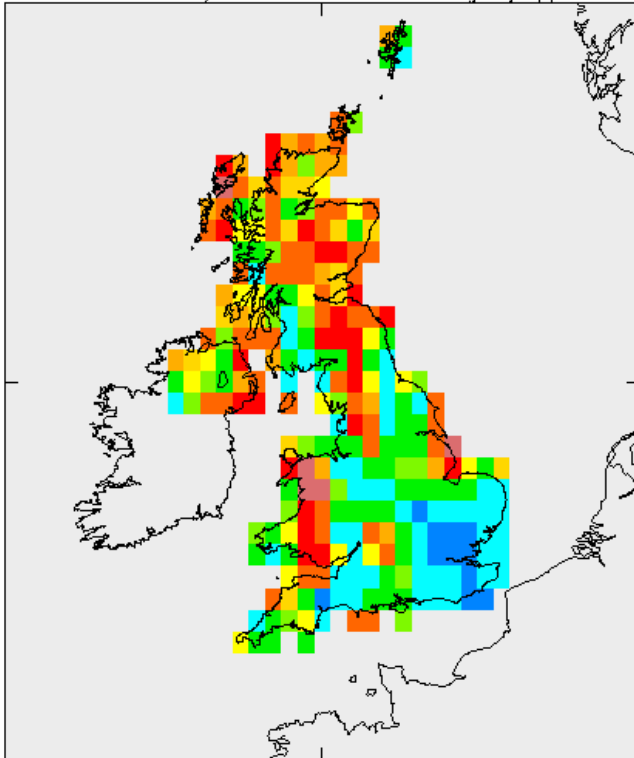
Errors v gauges

40 km

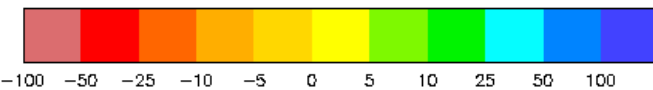
12km

4km

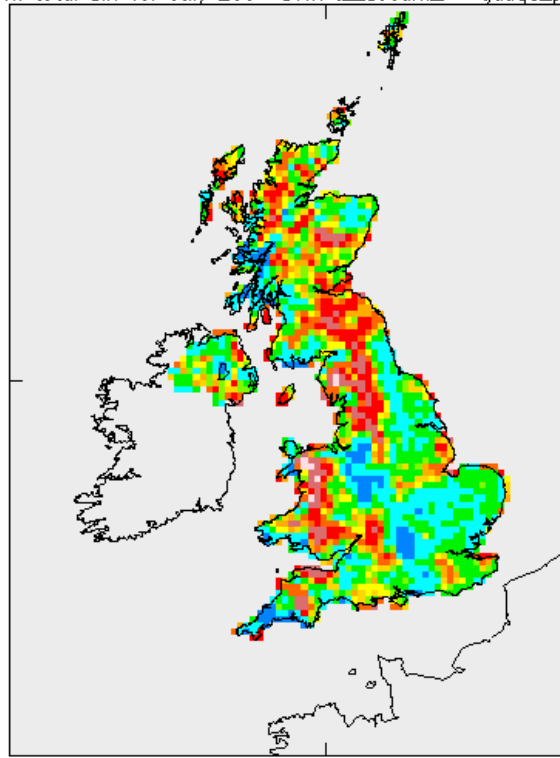
24h total diff for July 2007 ukmo_accum_ - gauge_ppn_Actu



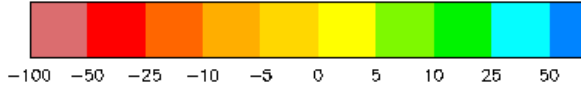
Mean=5.76088 sd=26.6418 max=70.6266 min=-86.9990



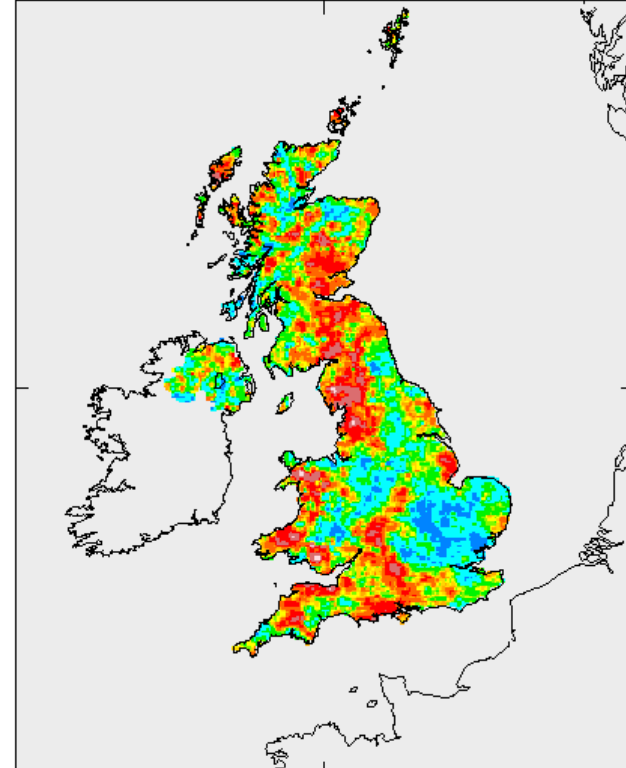
24h total diff for July 2007 OYmvt_accum_ - gauge_ppn_Actu



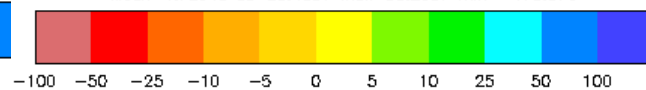
Mean=6.48054 sd=30.8096 max=90.8514 min=-178.874



24h total diff for July 2007 Q4mvt_accum_ - gauge_ppn_Actu



Mean=4.72948 sd=28.4397 max=93.2657 min=-175.076





Met Office

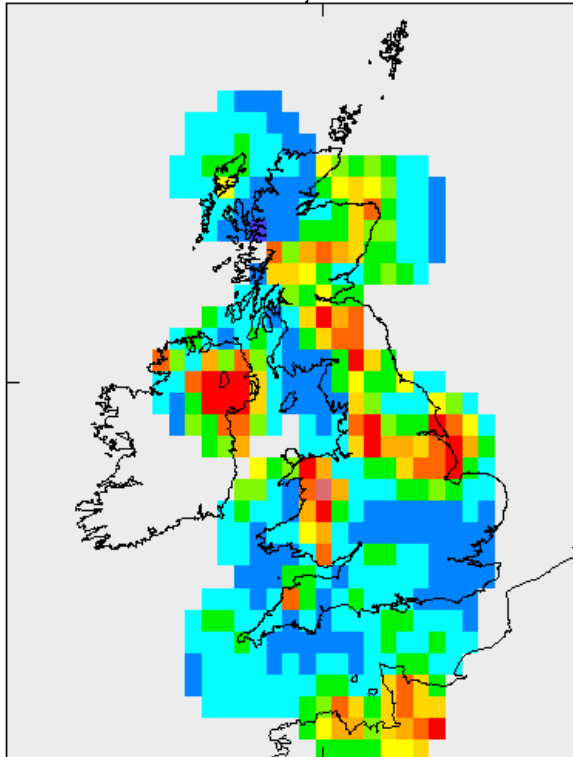
40 km

Forecast monthly totals – July Errors v radar

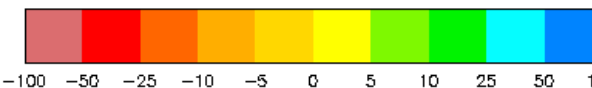
12km

4km

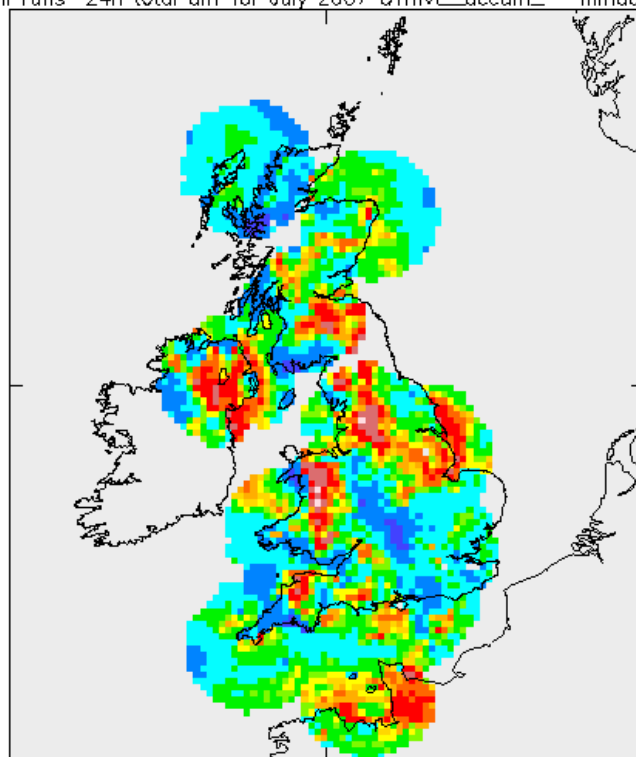
all runs 24h total diff for July 2007 ukmo_accum - r



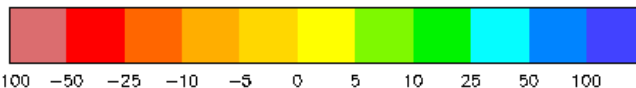
Mean=28.0653 sd=28.0808 max=107.091 min=-61.6284



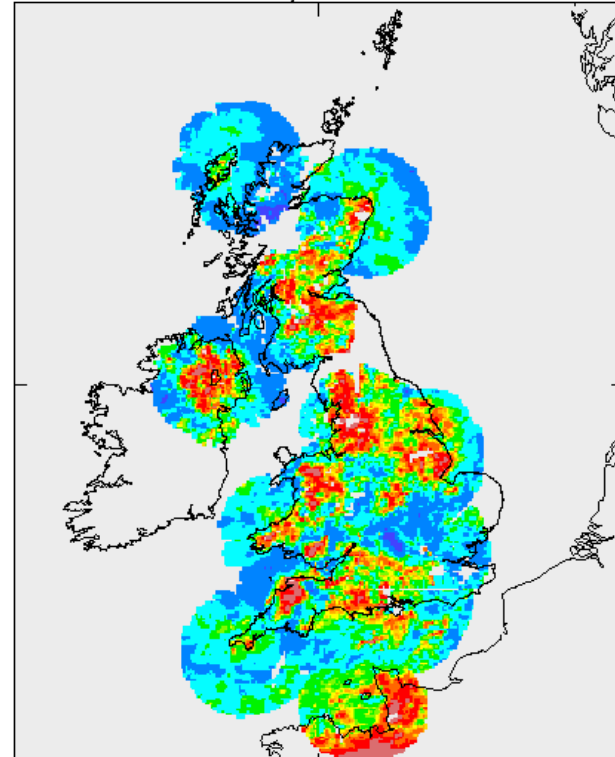
all runs 24h total diff for July 2007 Qymvt_accum - nimacc



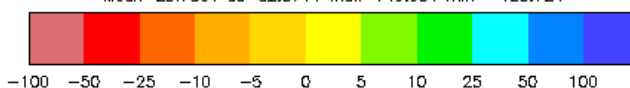
Mean=20.6159 sd=30.8577 max=148.634 min=-115.016



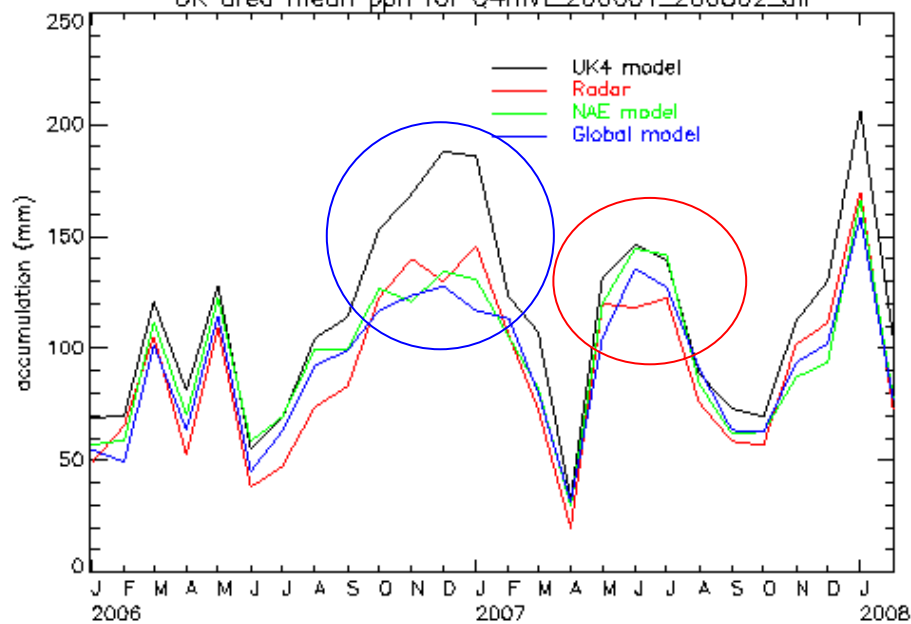
all runs 24h total diff for July 2007 Q4mvt_accum - nimacc



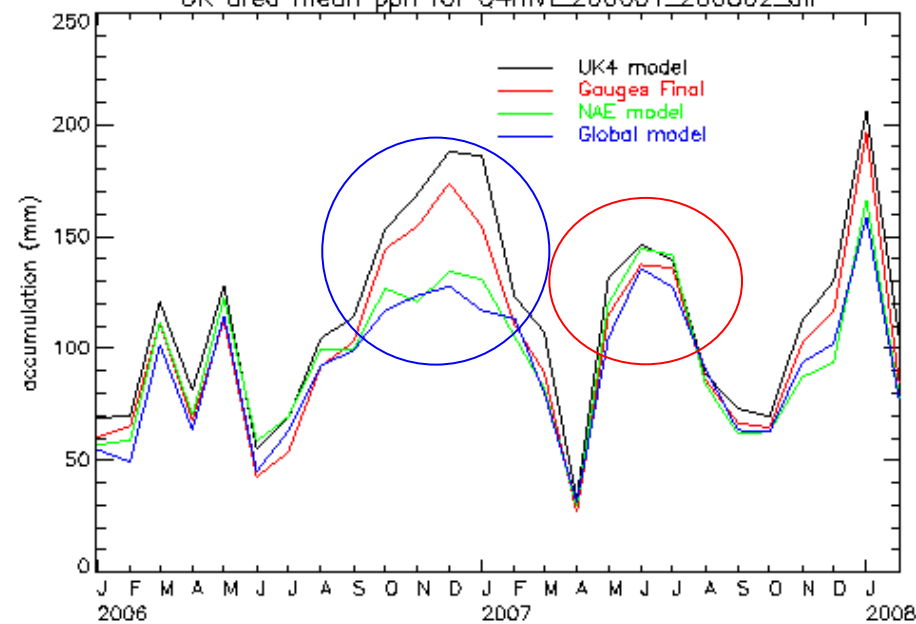
Mean=25.7861 sd=32.5711 max=149.051 min=-123.724



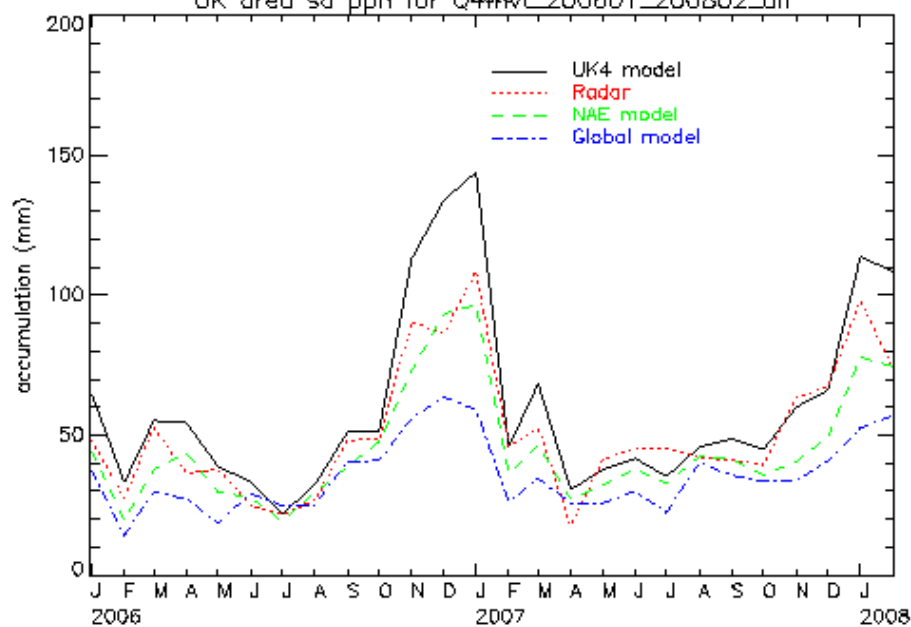
UK area mean ppn for Q4mvt_200601_200802_all



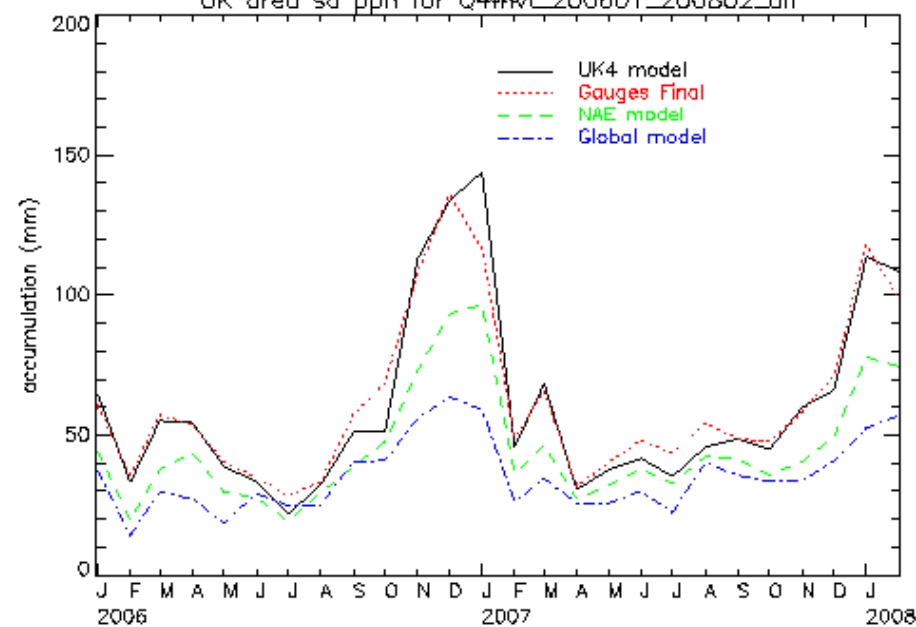
UK area mean ppn for Q4mvt_200601_200802_all



UK area sd ppn for Q4mvt_200601_200802_all

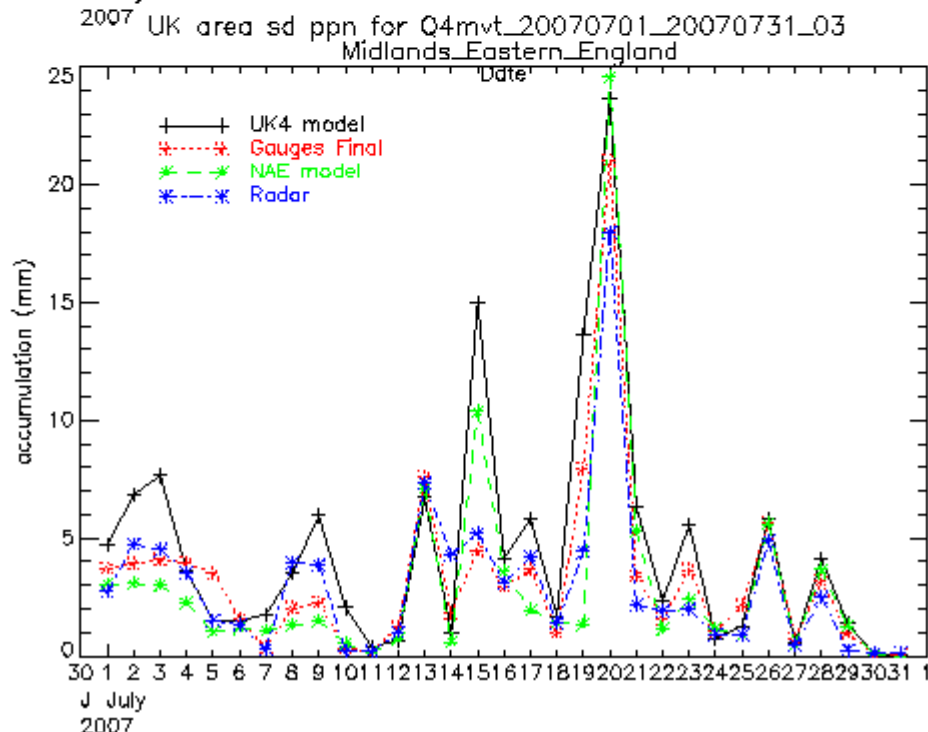
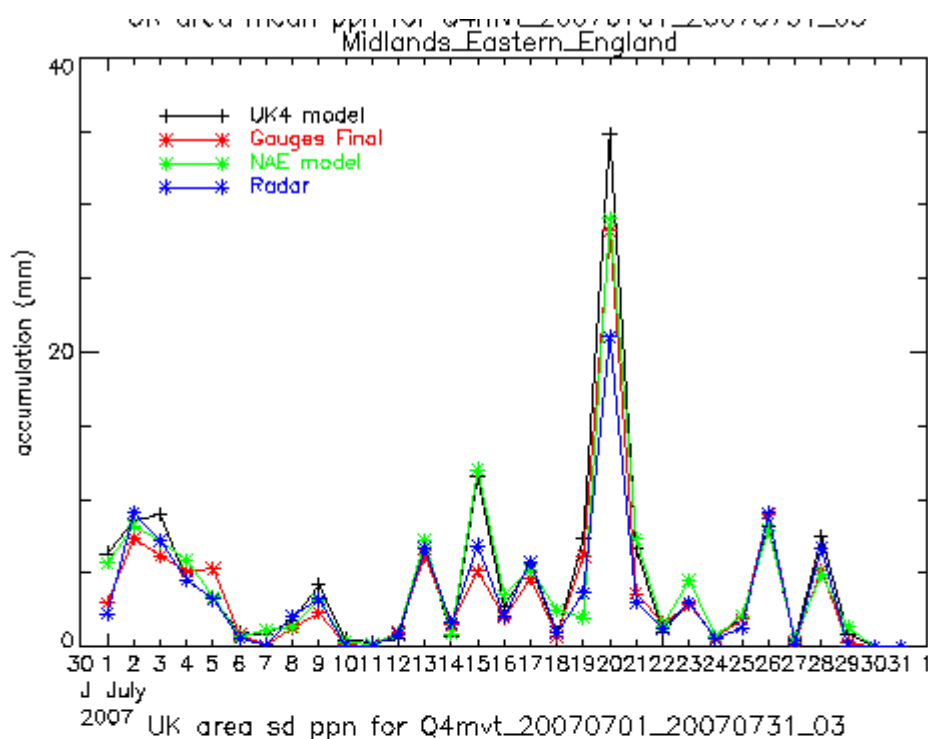


UK area sd ppn for Q4mvt_200601_200802_all





Daily rainfall July07





Met Office

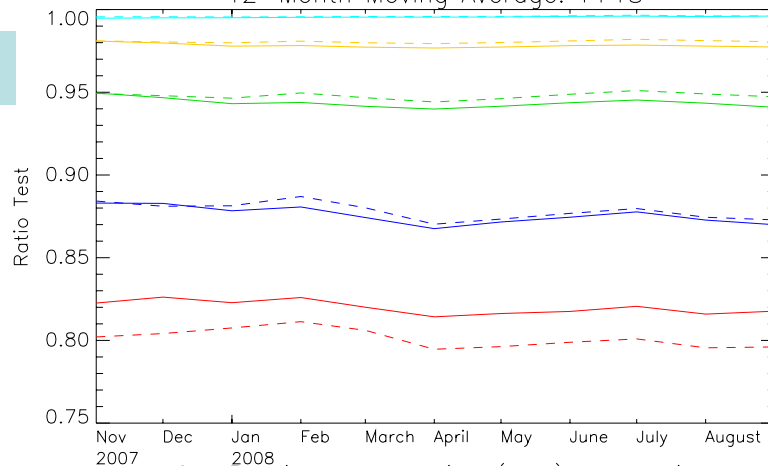
Comparison NAE & UK4 against stations (dashed) & radar composites (full) –Dec06- Sep08

NAE

Cases: — UK-NAE verified against Nirod analysis
- - - UK-NAE verified against surface obs

Anal Categories: — 0.2 — 1.0 — 4.0 — 8.0 — 16.0

6hr Precip Accumulation (mm): Analysis
Reduced UK 4km Model area
12-Month Moving Average: T+18

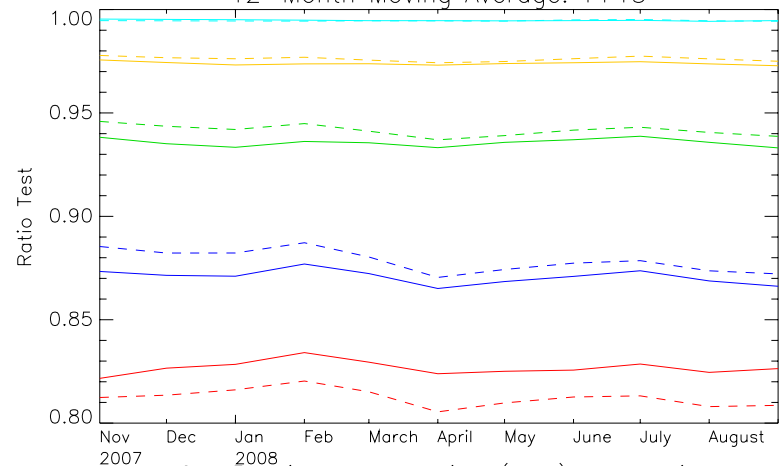


UK4

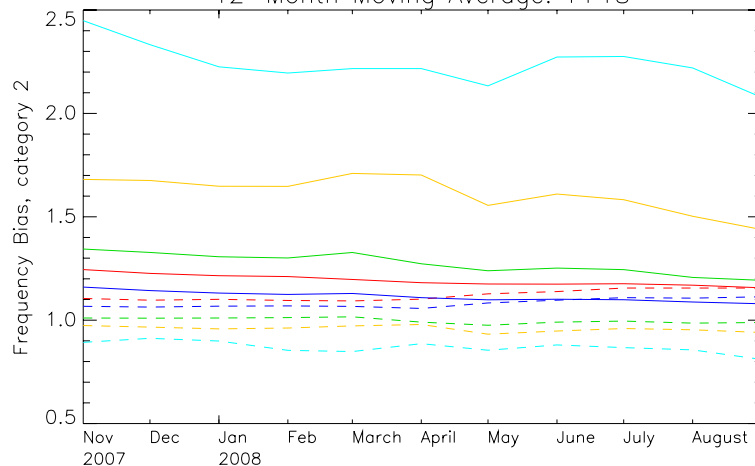
Cases: — UK-UK4 verified against Nirod analysis
- - - UK-UK4 verified against surface obs

Anal Categories: — 0.2 — 1.0 — 4.0 — 8.0 — 16.0

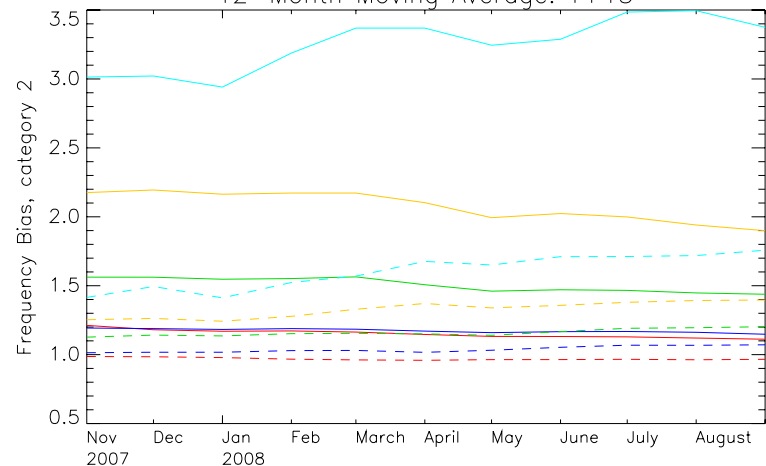
6hr Precip Accumulation (mm): Analysis
Reduced UK 4km Model area
12-Month Moving Average: T+15



6hr Precip Accumulation (mm): Analysis
Reduced UK 4km Model area
12-Month Moving Average: T+18



6hr Precip Accumulation (mm): Analysis
Reduced UK 4km Model area
12-Month Moving Average: T+15



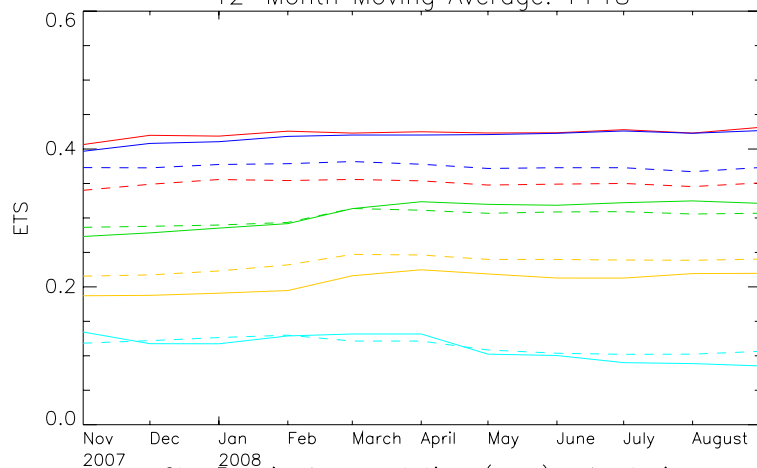


Comparison NAE & UK4 against stations (dashed) & radar composites (full) –Dec06- Sep08

NAE

Cases: — UK-NAE verified against Nimrod analysis
- - - UK-NAE verified against surface obs
Anal Categories: — 0.2 — 1.0 — 4.0 — 8.0 — 16.0

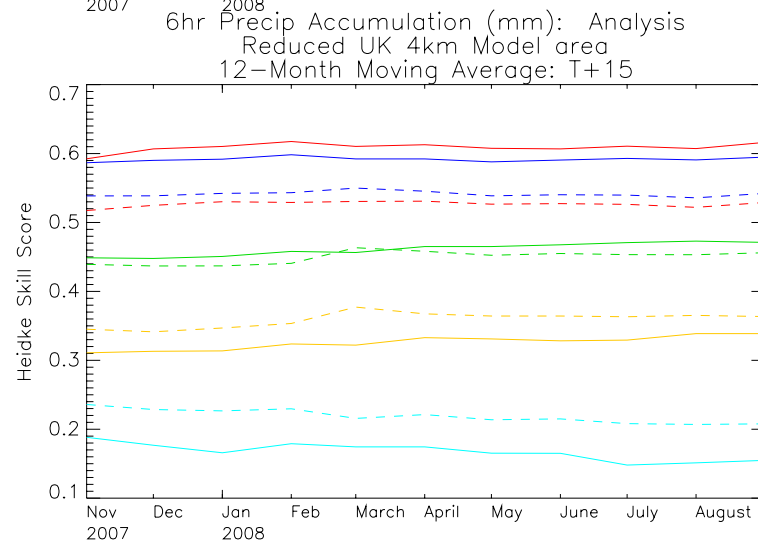
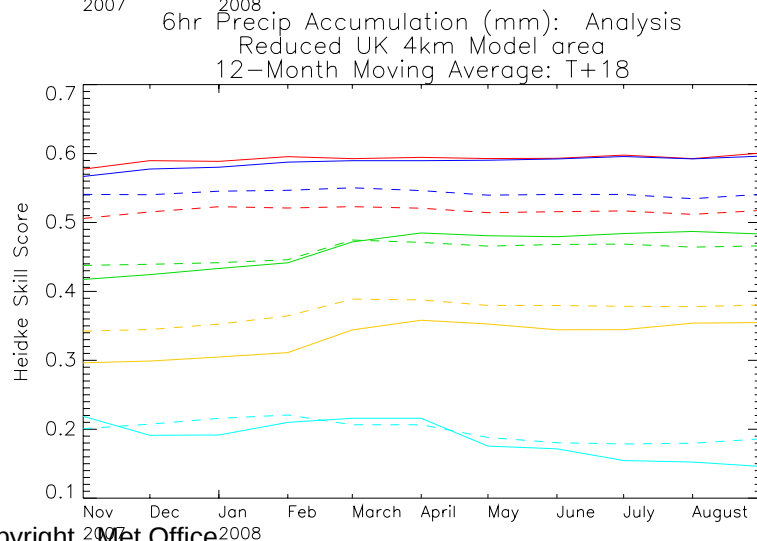
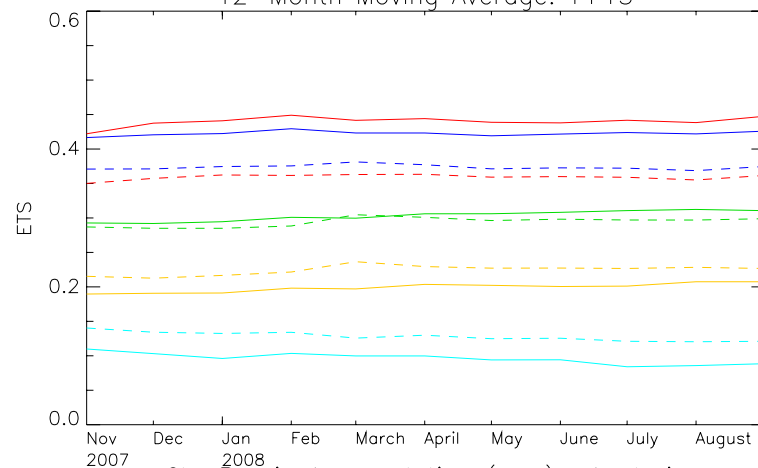
6hr Precip Accumulation (mm): Analysis
Reduced UK 4km Model area
12-Month Moving Average: T+18



UK4

Cases: — UK-UK4 verified against Nimrod analysis
- - - UK-UK4 verified against surface obs
Anal Categories: — 0.2 — 1.0 — 4.0 — 8.0 — 16.0

6hr Precip Accumulation (mm): Analysis
Reduced UK 4km Model area
12-Month Moving Average: T+15





Met Office

“Fuzzy” Verification

Against radar composites



Intensity/scale verification

Barbara Casati et al

- Threshold precipitation forecast and analysis $X > u$
 - Intensity, u
- Form Binary error & analyse with Haar wavelets (scale, l)

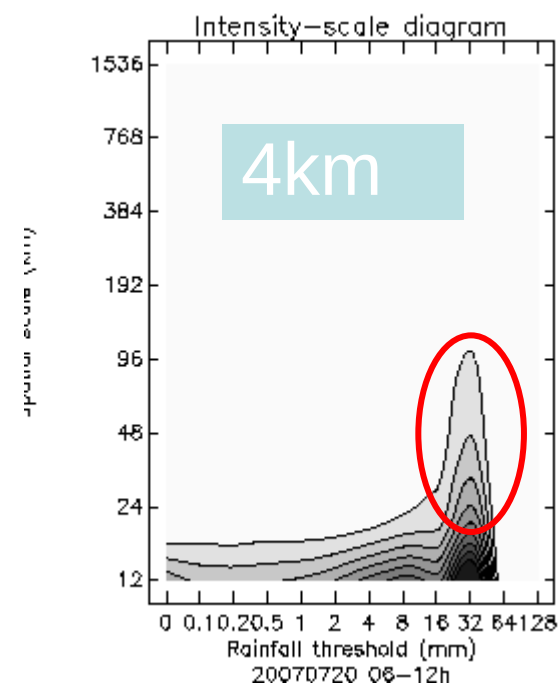
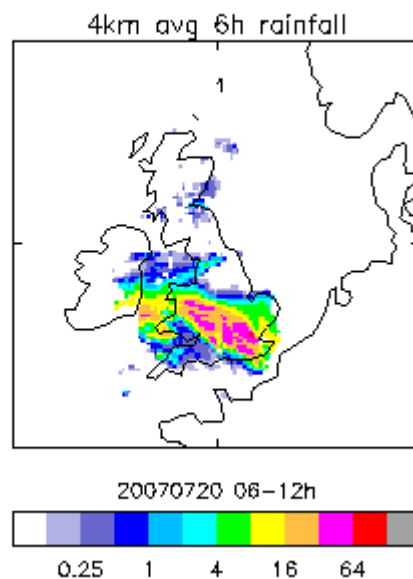
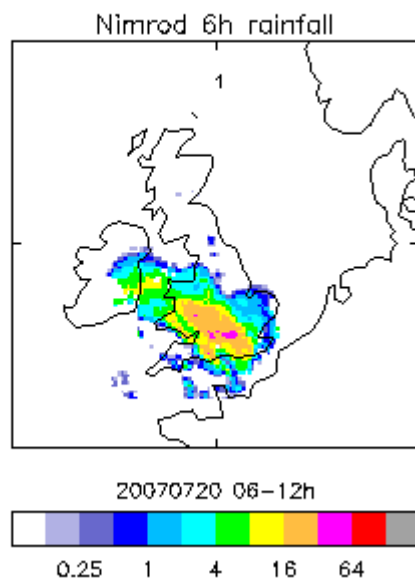
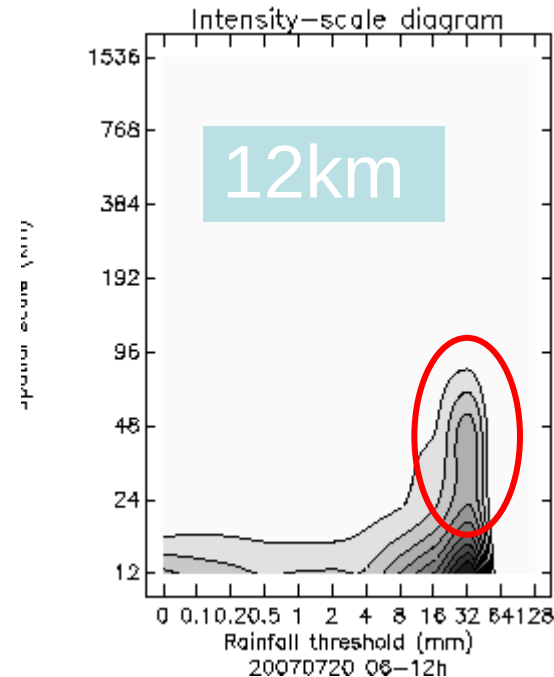
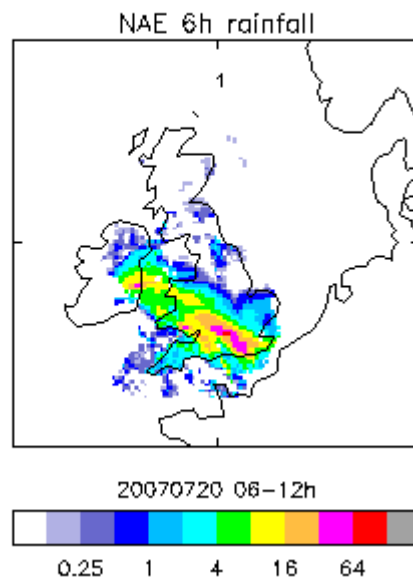
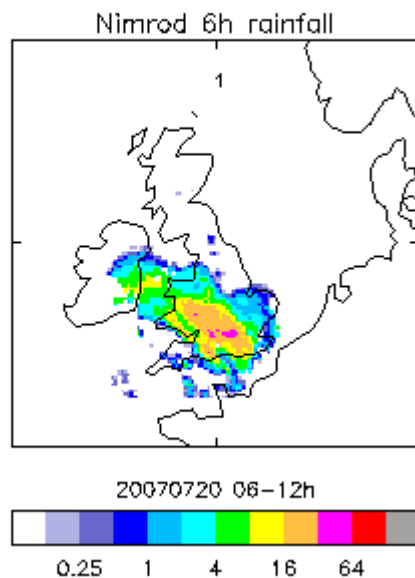
$$E_u = \sum_{l=1}^L E_{u,l}$$

- MSE of binary error sum of each length scale error (wavelet orthogonality)

$$MSE_u = \sum_{l=1}^L MSE_{u,l}$$

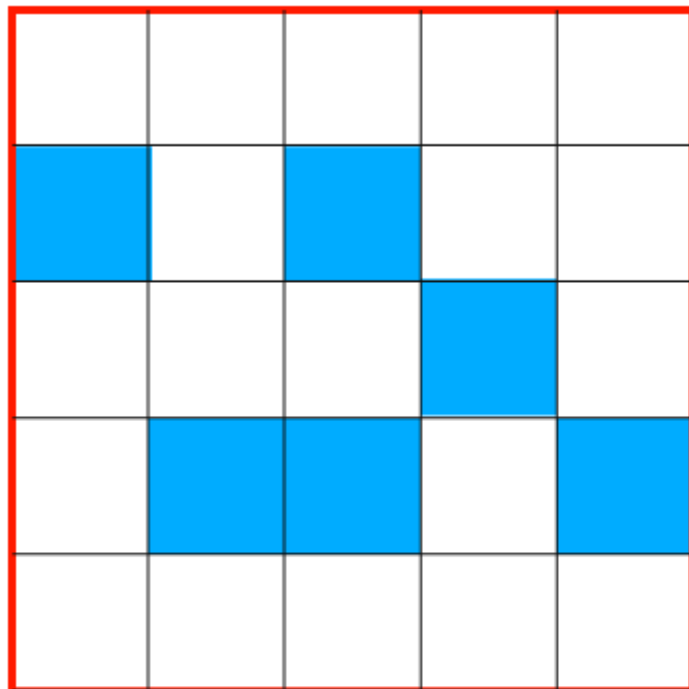
- Skill compared to random forecast with no scale dependency

$$SS_{u,l} = 1 - \frac{MSE_{u,l}}{2\varepsilon(1-\varepsilon)/L} \quad \varepsilon = \frac{a+c}{n} = P(X > u)$$



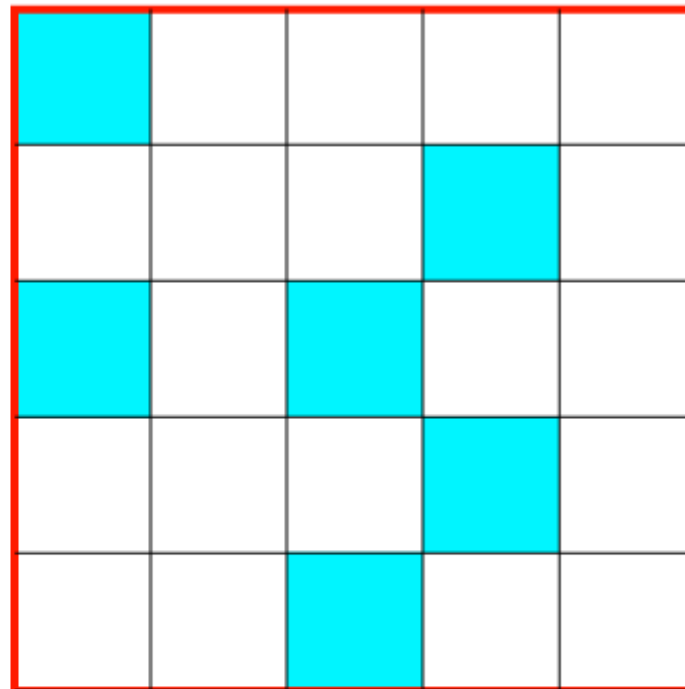
Schematic comparison of fractions (Nigel Roberts)

observed



Fraction = $6/25 = 0.24$

forecast



Fraction = $6/25 = 0.24$



A score for comparing fractions with fractions (Nigel Roberts)

Brier score for comparing fractions

FBS
(Fractions Brier Score)

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

$0 \leq p_j \leq 1$ forecast fractions

$0 \leq o_j \leq 1$ radar fractions

N number of points

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

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

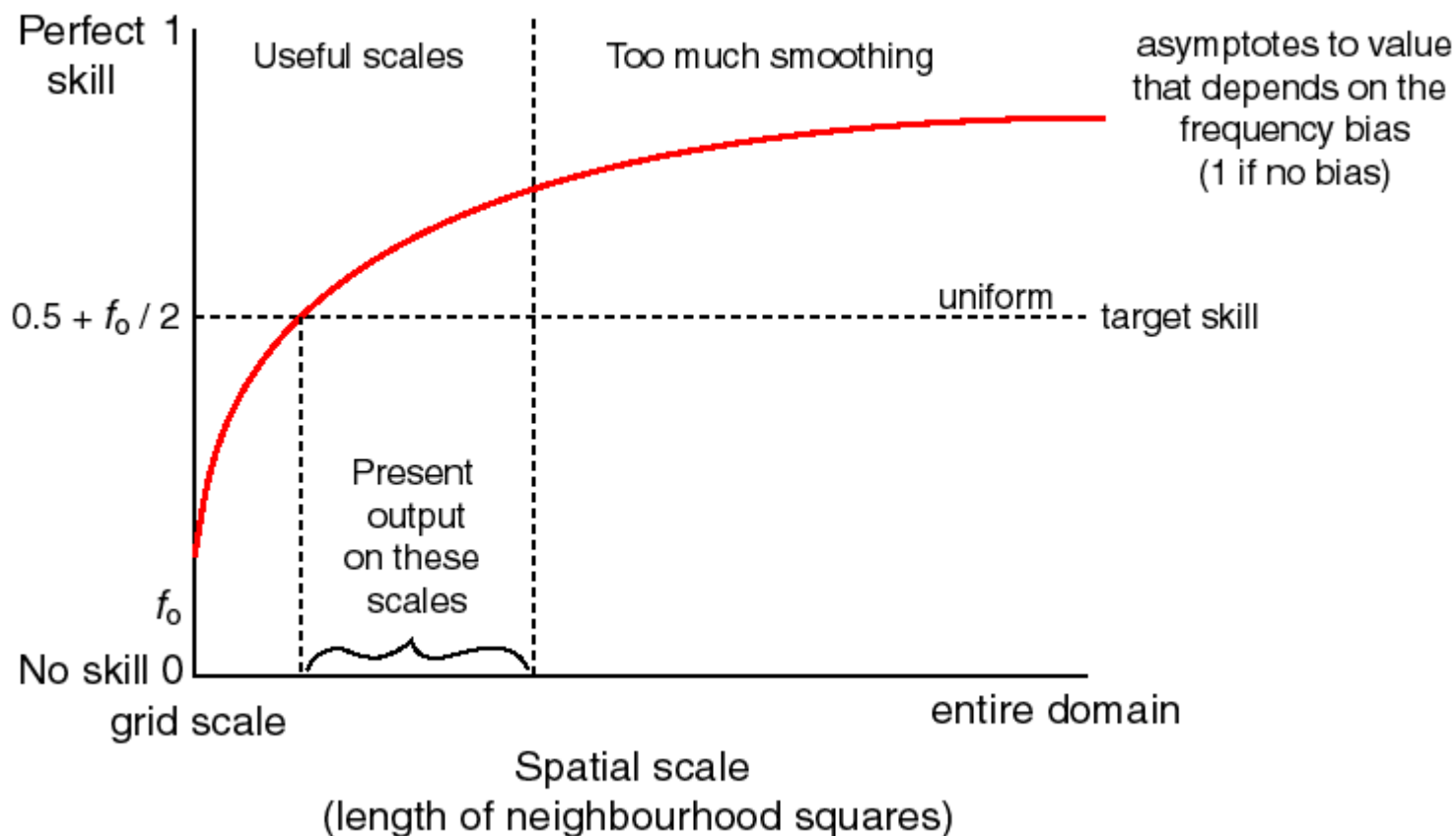
No overlap of fractions



Me

FSS

Example graph of FSS against neighbourhood size





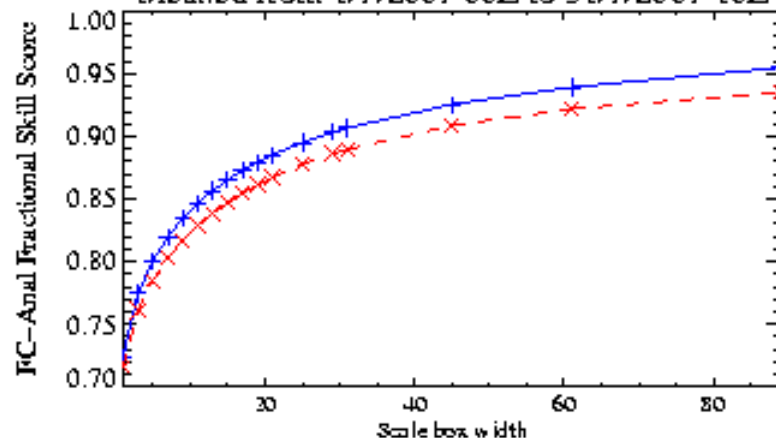
FSS JULY 2007

12km

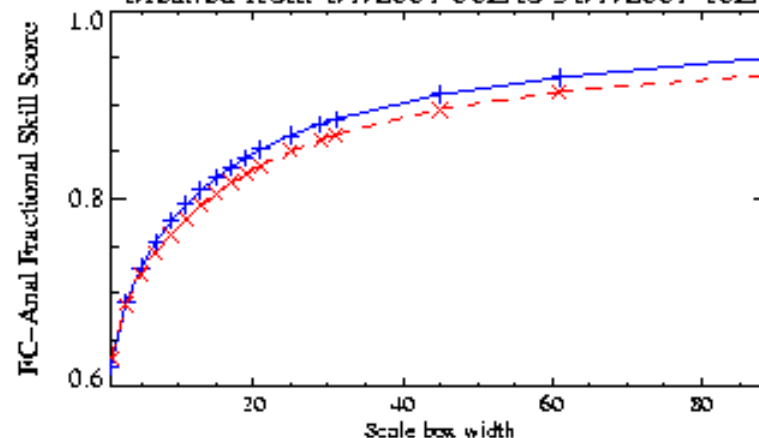
4km

FCRanges: +---+ T+9 x---x T+12

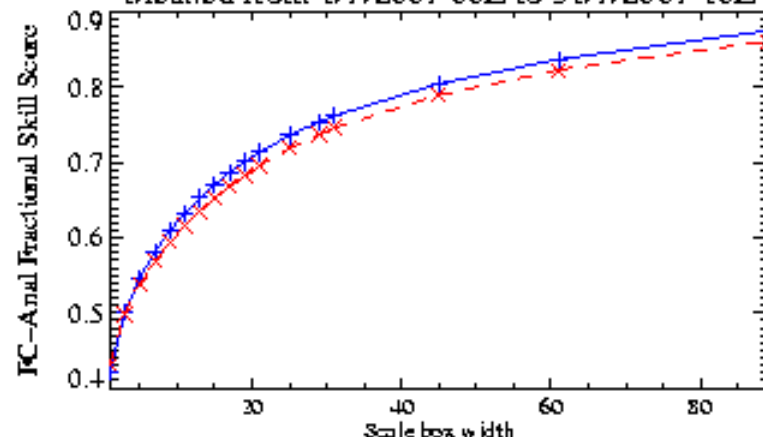
6hr Precip Accumulation (0.200mm): Analysis
UK area (scale rainfall)
Meaned from 1/7/2007 00Z to 31/7/2007 18Z



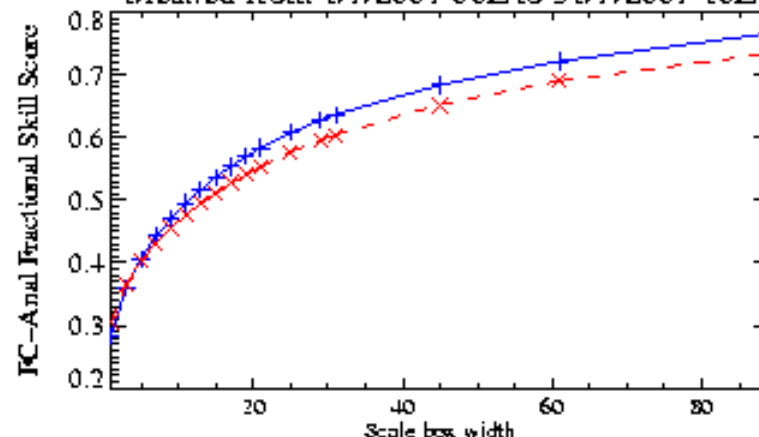
6hr Precip Accumulation (1.000mm): Analysis
UK area (scale rainfall)
Meaned from 1/7/2007 00Z to 31/7/2007 18Z



6hr Precip Accumulation (4.000mm): Analysis
UK area (scale rainfall)
Meaned from 1/7/2007 00Z to 31/7/2007 18Z



6hr Precip Accumulation (8.000mm): Analysis
UK area (scale rainfall)
Meaned from 1/7/2007 00Z to 31/7/2007 18Z





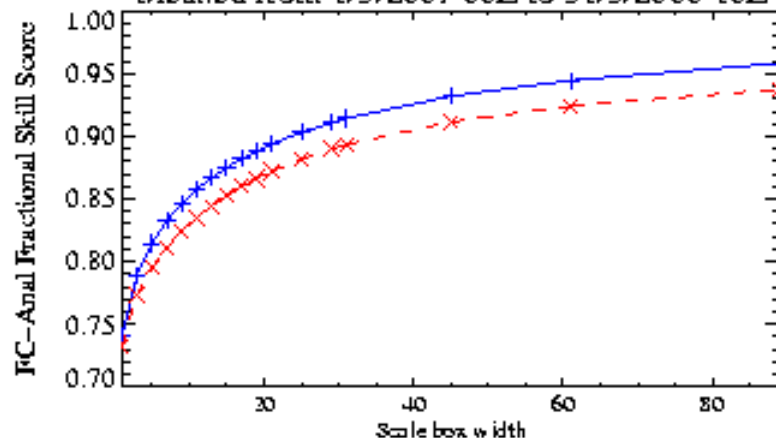
FSS Mar07-Mar08

12km

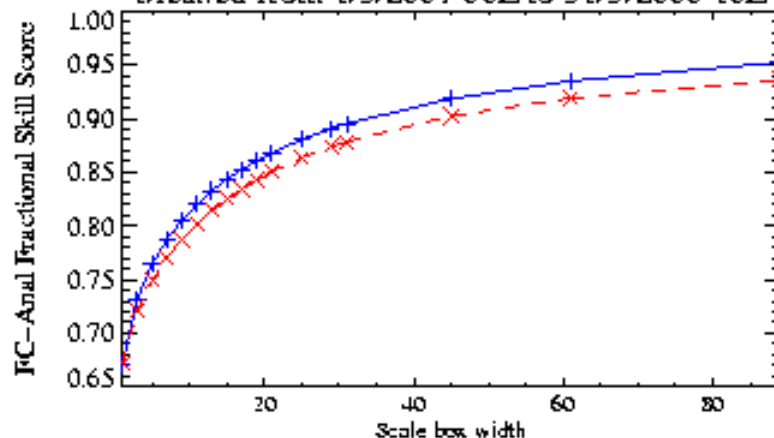
4km

FCRanges: + — + T+9 × — × T+12

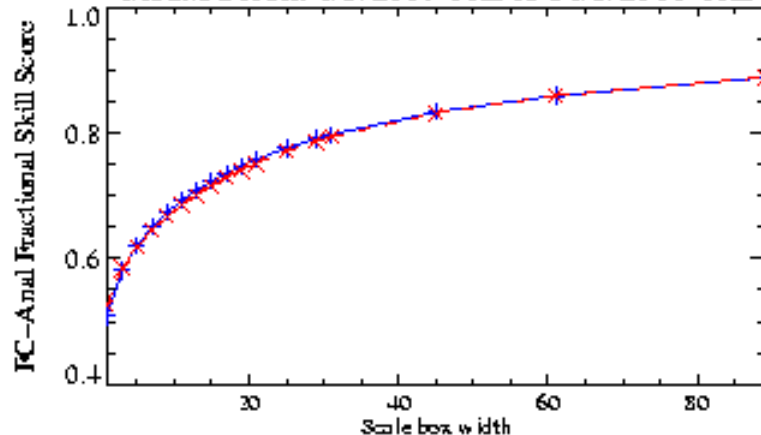
6hr Precip Accumulation (0.200mm): Analysis
UK area (scale rainfall)
Meand from 1/3/2007 00Z to 31/3/2008 18Z



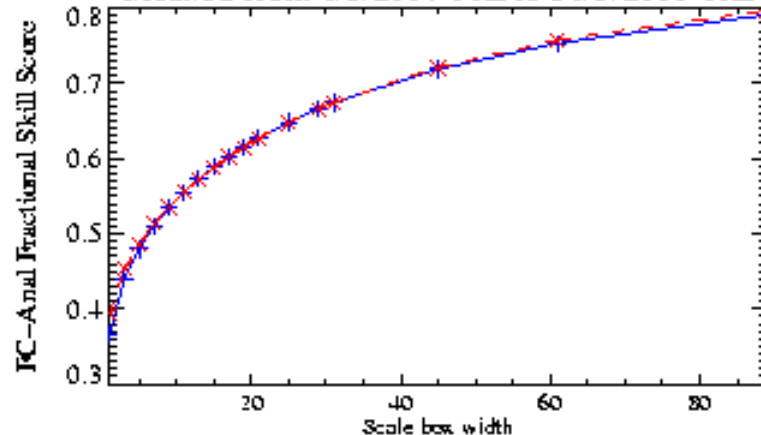
6hr Precip Accumulation (1.000mm): Analysis
UK area (scale rainfall)
Meand from 1/3/2007 00Z to 31/3/2008 18Z



6hr Precip Accumulation (4.000mm): Analysis
UK area (scale rainfall)
Meand from 1/3/2007 00Z to 31/3/2008 18Z



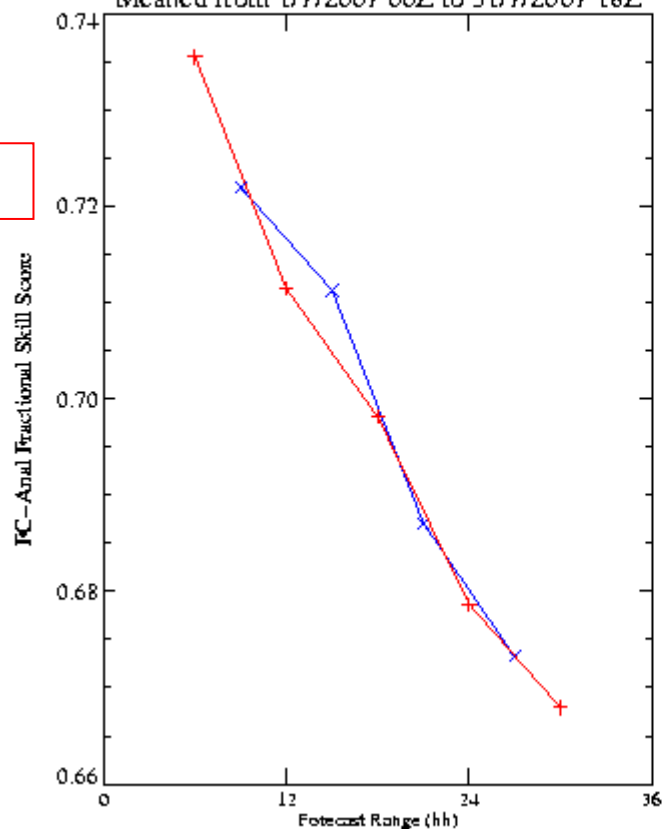
6hr Precip Accumulation (8.000mm): Analysis
UK area (scale rainfall)
Meand from 1/3/2007 00Z to 31/3/2008 18Z



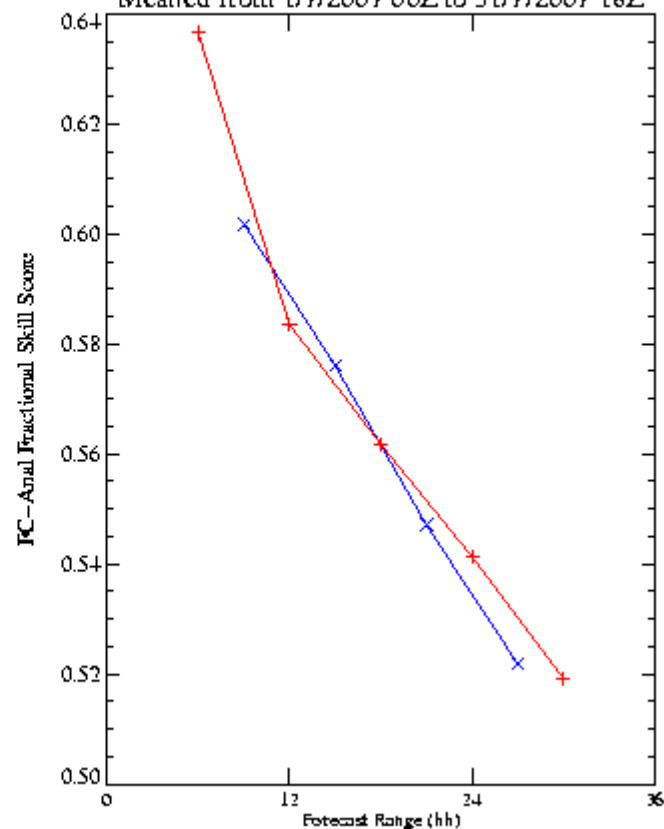
FSS – July 07 -0.2 & 1.0mm thresholds -scale=7*5km

Cases: + NAE x UK4

6hr Precip Accumulation (0.200mm): Analysis
UK area (scale rainfall), Scale: 7
Meaned from 1/7/2007 00Z to 31/7/2007 18Z



6hr Precip Accumulation (1.000mm): Analysis
UK area (scale rainfall), Scale: 7
Meaned from 1/7/2007 00Z to 31/7/2007 18Z



Conclusions

- Standard verification contradictory as resolution increases
 - Nearest grid point – “double penalty effect”
 - Unrepresentative
- Gridded gauges and radar composites can differ by 50-100%
- Daily precipitation compared to dense gauges or radar well forecast – better by 4km resolution
- Fuzzy scores – fractional skill score can identify scales at which forecasts show useful skill
- More work planned to show the effects of uncertainty in observations on verification



Met Office



Questions and answers