Review of the SRNWP-V programme

Clive Wilson –Expert Team on diagnostics, validation & verification

35th EWGLAM/20th SRNWP meetings –Turkey , 30 Sep- 3 Oct 2013

Contents

- SRNWP-V programme
  - 1st phase 2009-2010
  - 2nd Phase 2011-2012
- Main outcomes and conclusions
- Comparison results
- New scores SEEPS
- Additional parameters
  - Cloud/gusts
- Spatial and scale selective verification of precipitation
  - OPERA radar quality
- Inclusion of Severe/High impact Weather verification
- Full documentation . Portable methods

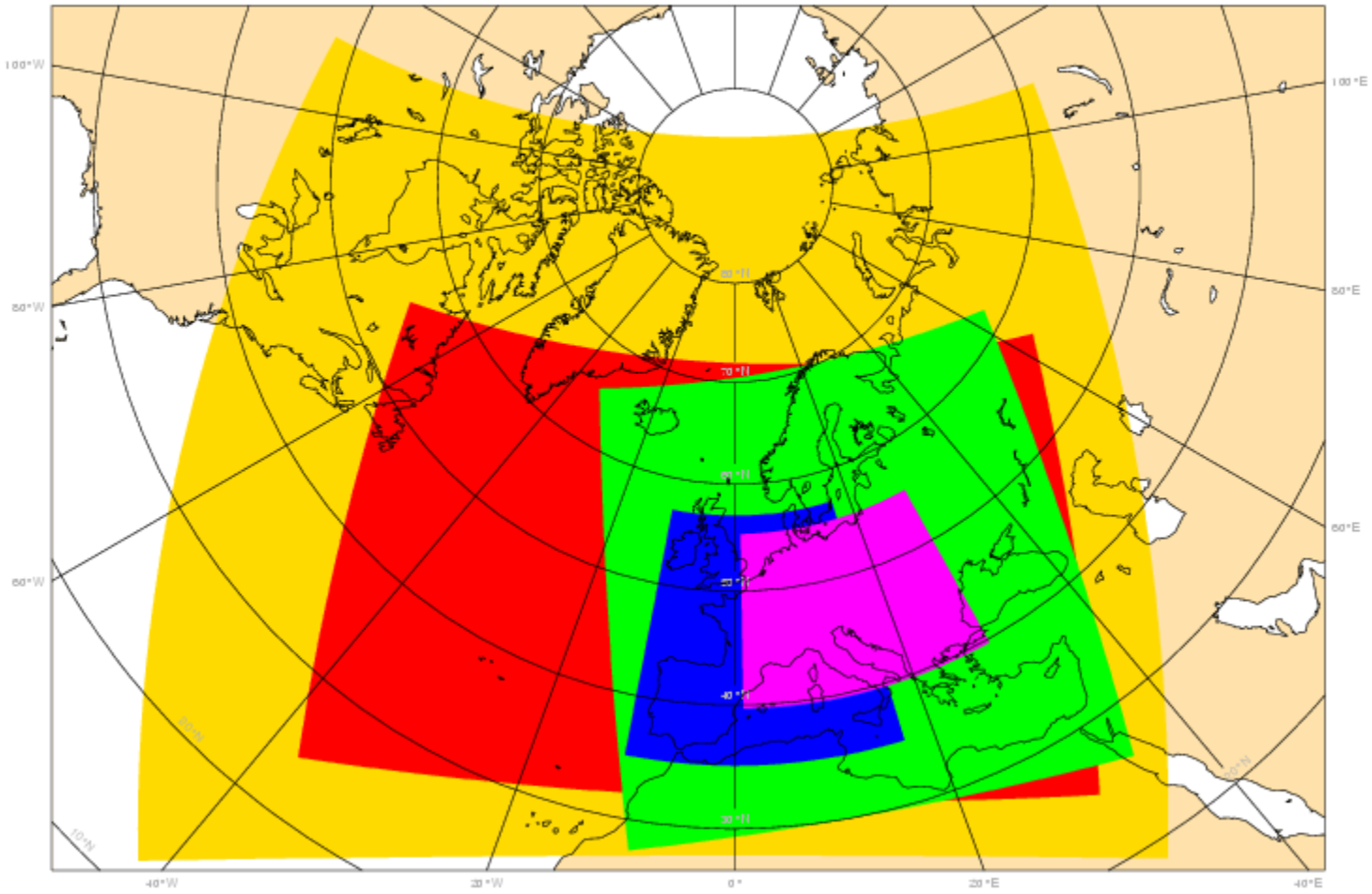
Outcomes & Conclusions

- First (published) extended intercomparison of operational LAM forecasts
- Demonstrated additional value of LAMS
  - NB in practice actually larger due to timeliness (need to use older global model)
- Identified strengths & weaknesses in main consortia models
- Evaluated new scores and scale selective methods
- Extended scale methods to cloud
- Severe/High impact Weather verification – Extreme dependency scores
- Evaluation of OPERA composites & QC
- No longer appropriate to continue intercomparison
  - Smaller / high resolution domains
  - Short range EPS now focus
  - But scale selective verification increasingly important

Intercomparison

continuation
higher resolution of future operational models

Domains of 5 consortia reference models



15km

12km

7km

10km

5 km

(10km)

Hirlam UM COSMO ALADIN Aladin-Lace

Key

Model	Label
Hirlam reference run by FMI	UK-FI
Aladin-France run by Meteo-France	UK-FR
COSMO Europe run by DWD	UK-GE
The North Atlantic European configuration of the Unified Model run by the Met Office	UK-EU
Aladin-Czech (LACE) run by CHMI	*UK-LC*
ECMWF high resolution global model	EC-GM

Comparison over ALADIN-France domain unless otherwise stated

Sea level Pressure

RMSE

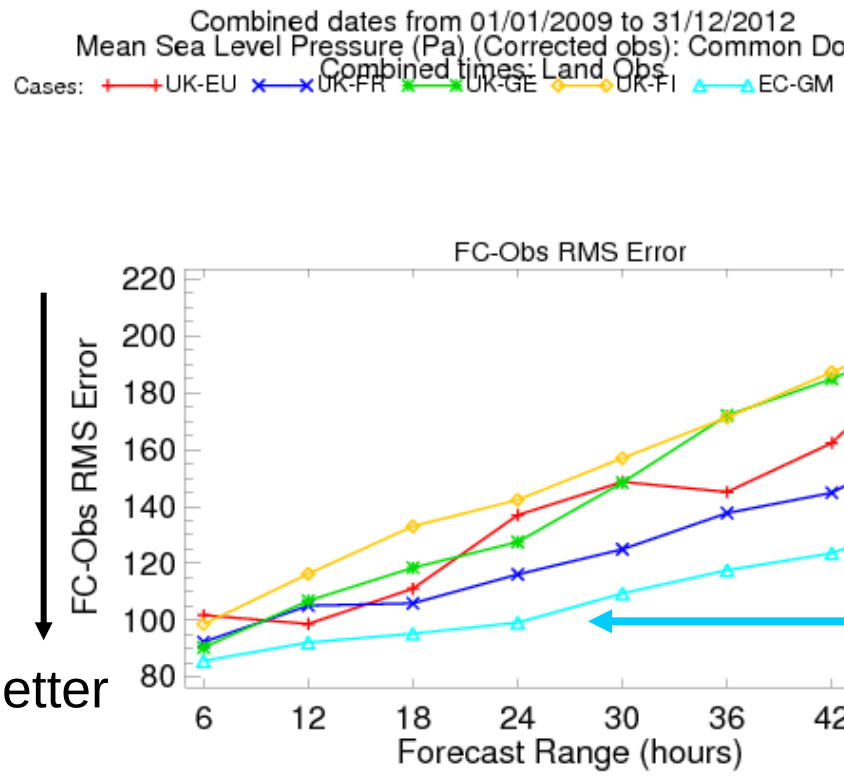
4 year mean
01/2009-12/2012

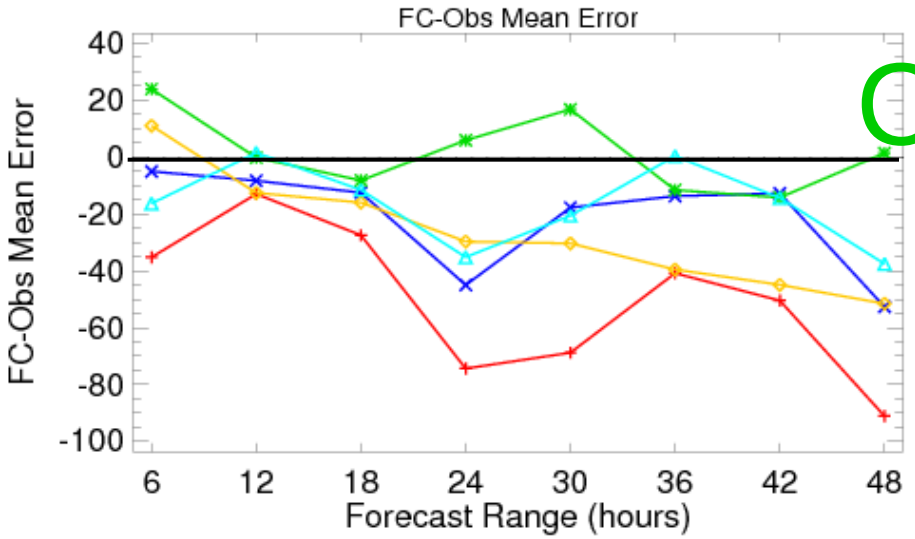
Bias

All LAMS inferior to EC global

Most have negative bias except COSMO

UM largest bias





Met Office

Screen temperature

RMSE

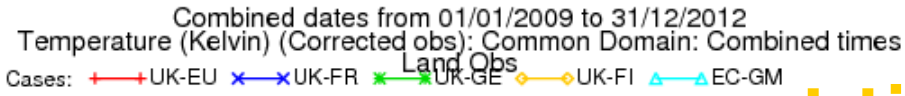
4 year mean
01/2009-12/2012

BIAS

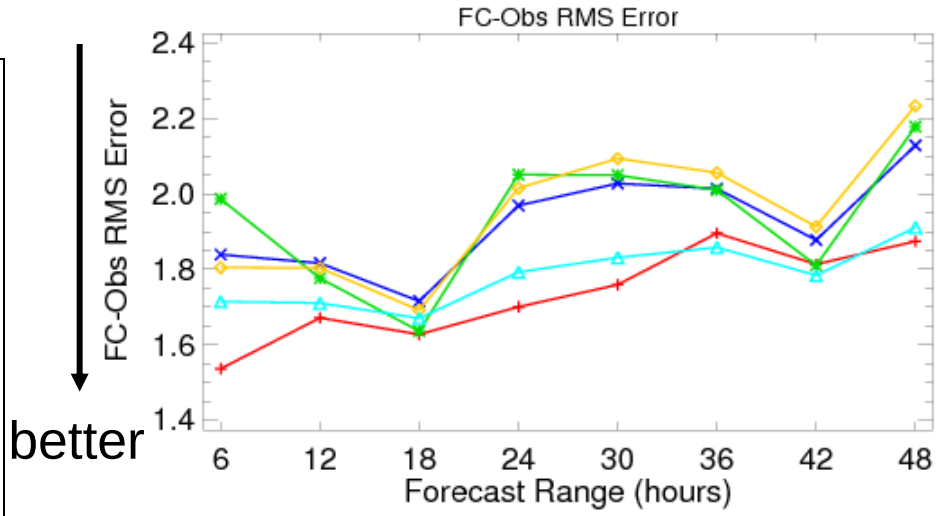
Warm night, cold day bias tendency

COSMO and Hirlam warm bias

UM better than EC till T+36

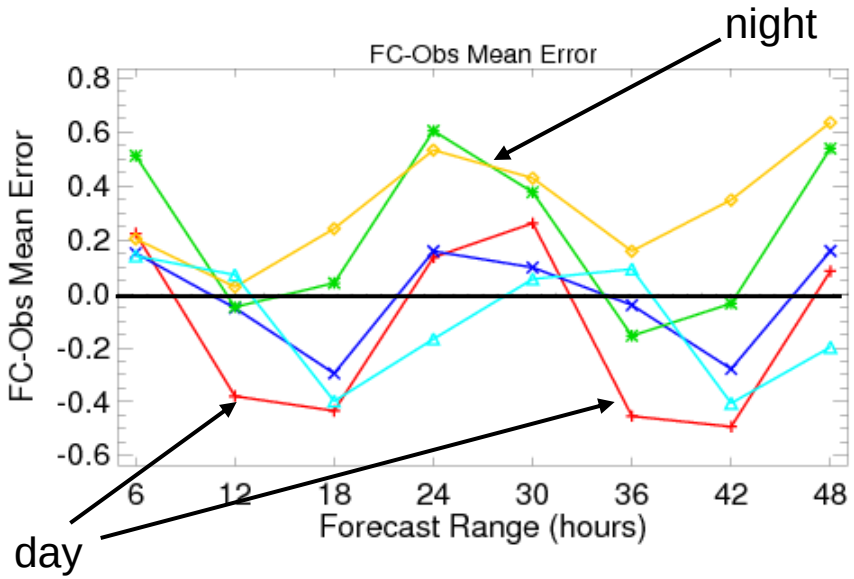


Hirlam



better

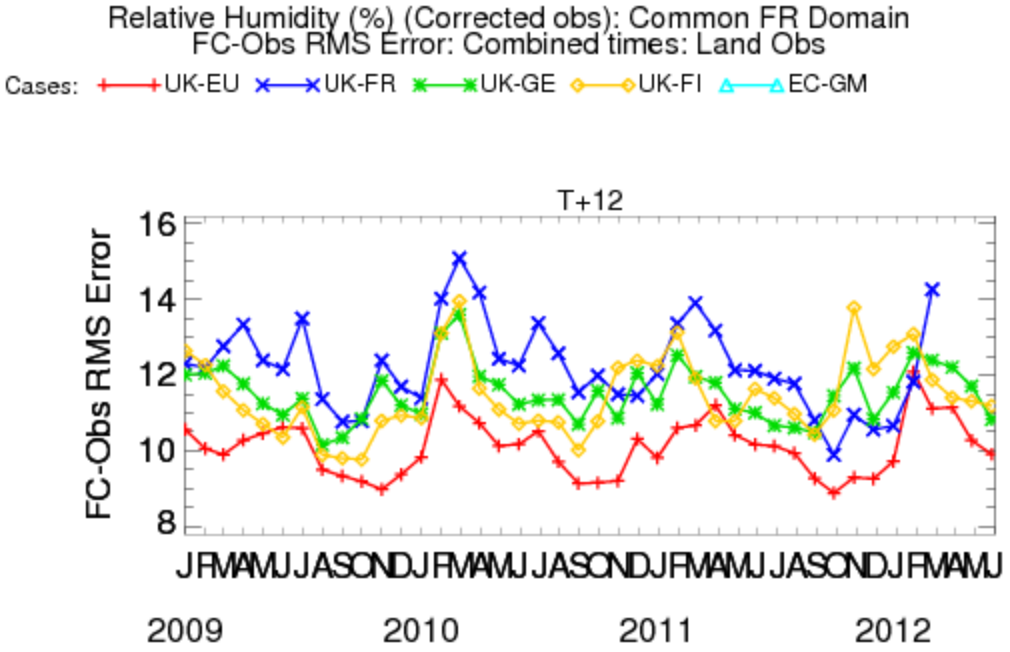
EC
UK

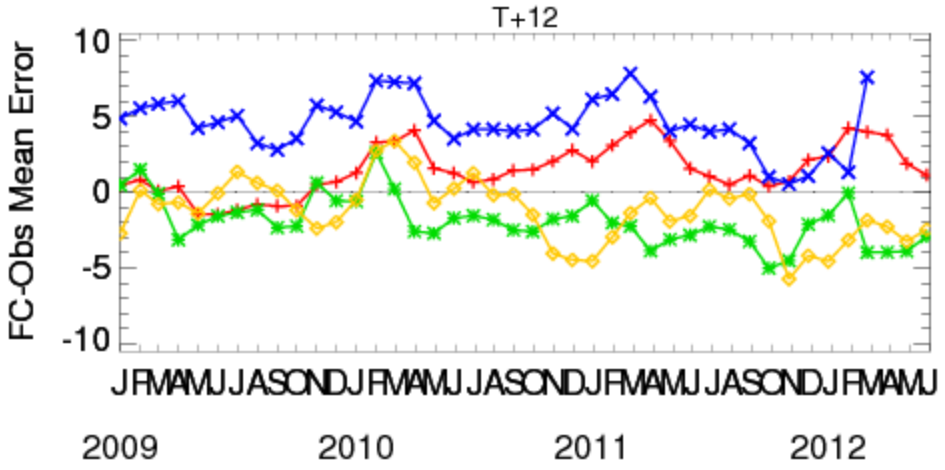


Relative humidity

NB screen temperature and humidity assimilated in UK-NAE

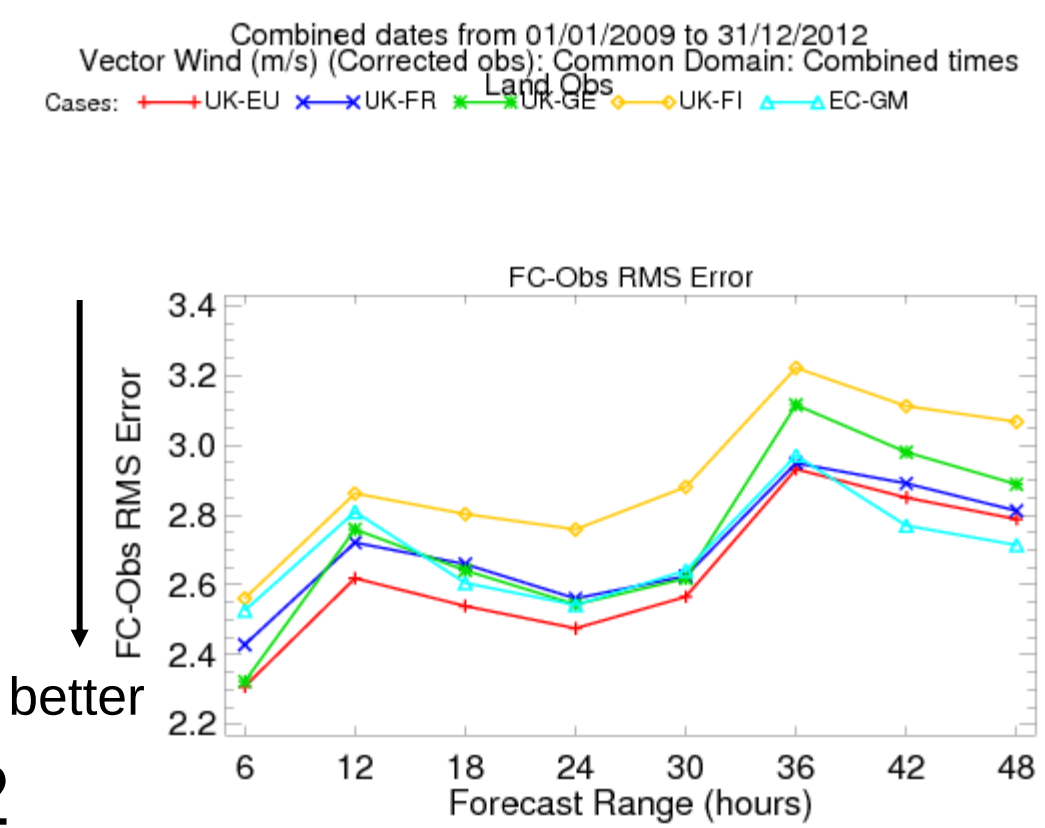
COSMO and Hirlam slight dry bias cf warm bias





10m vector wind rmse

4 years
01/2009-12/2012

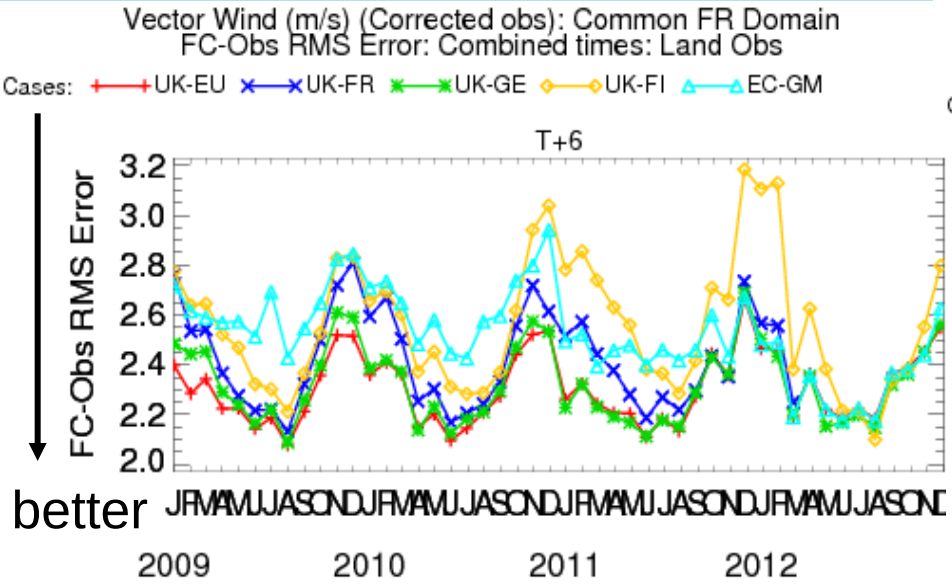


Regional models better/= than global in short range to T+30,

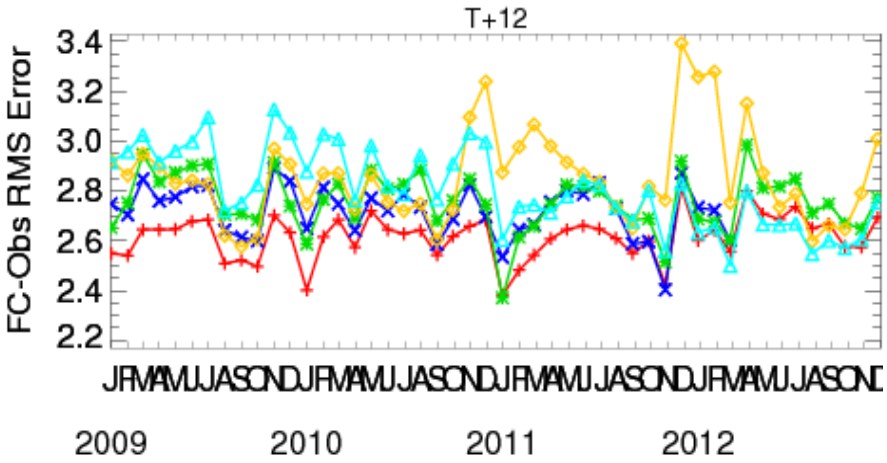
UM better to T+36

Hirlam (coarsest resolution) inferior to EC

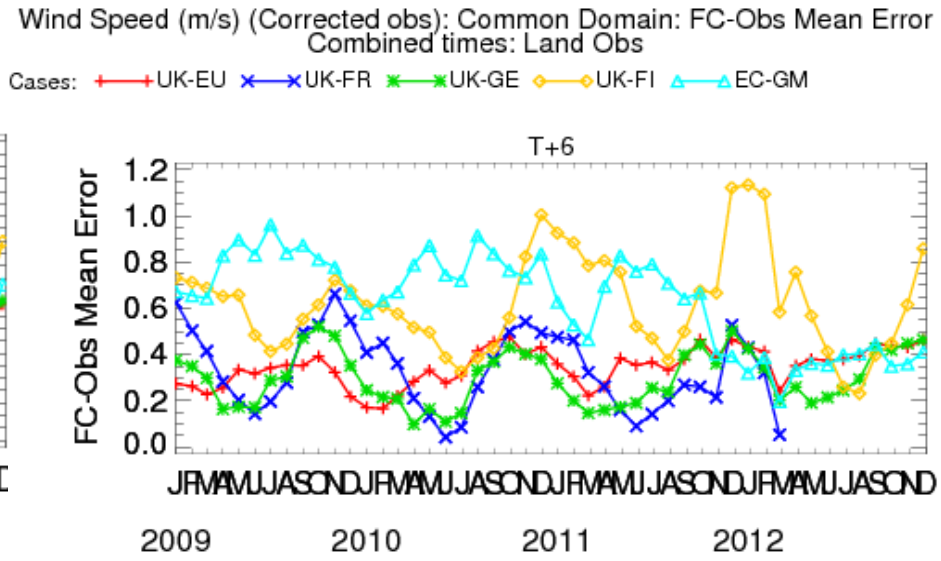
Vector wind error

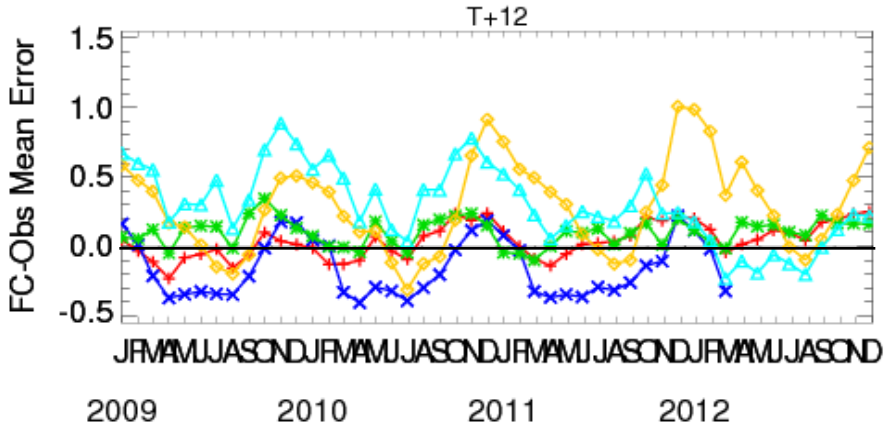


better



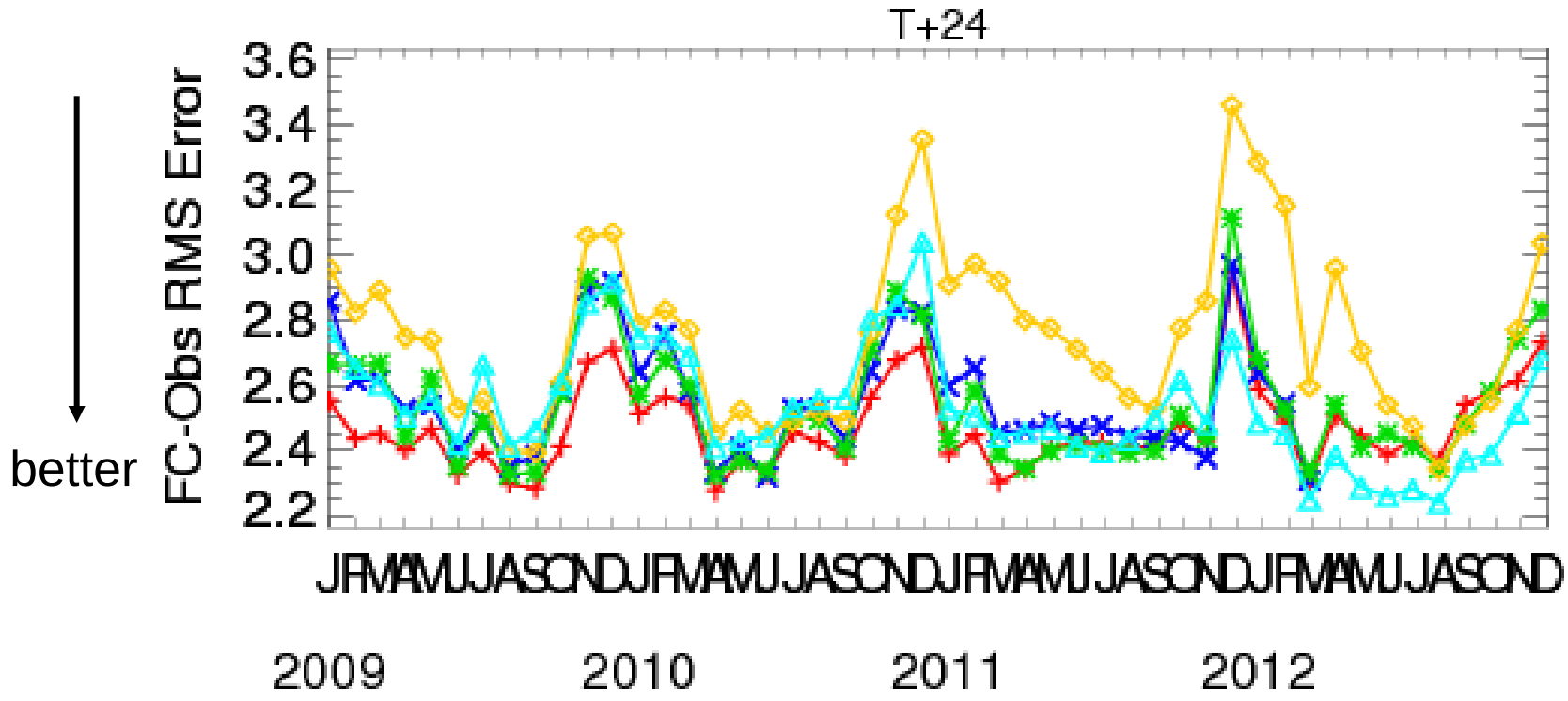
Speed Bias





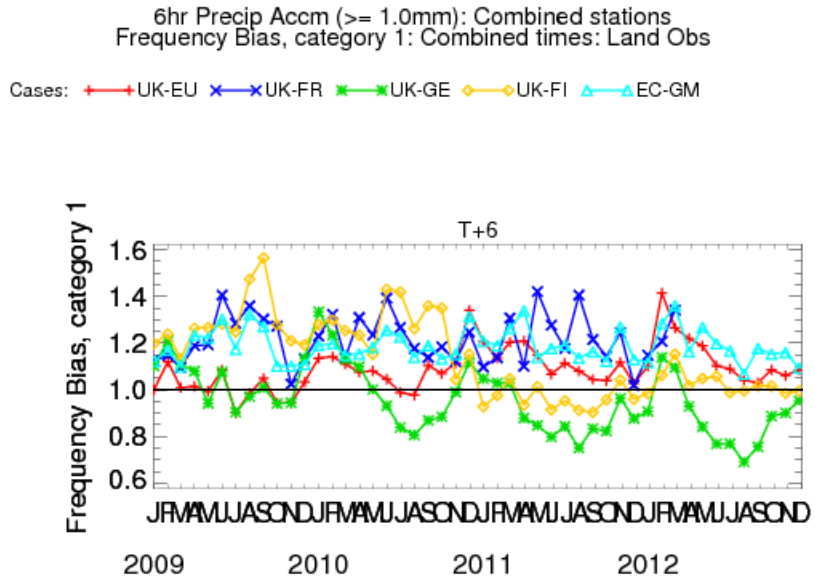
Models closer in 2012; EC steady improvement, Hirlam relatively poorer since 2011

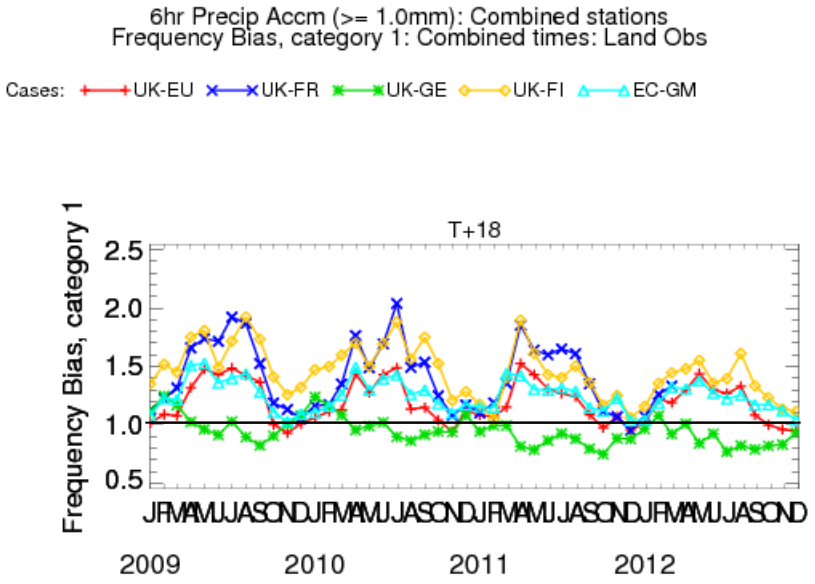
Vector wind error

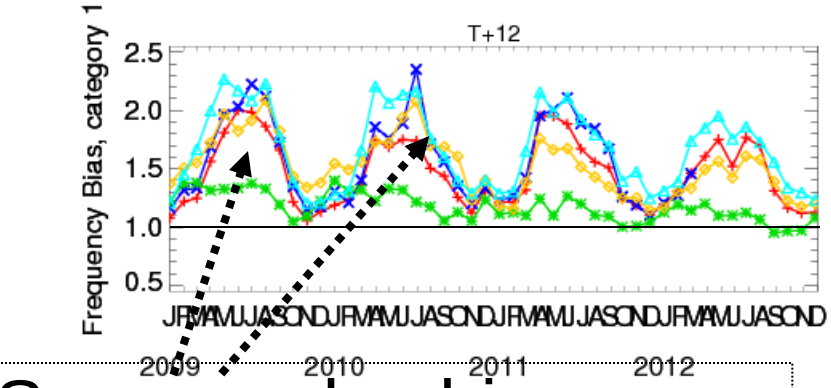


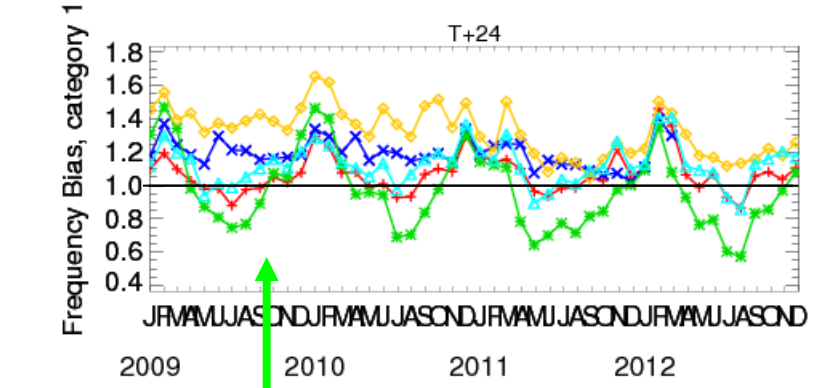
Models closer in 2012; EC steady improvement, Hirlam relatively poorer since 2011

ppn frequency bias >1mm/6h





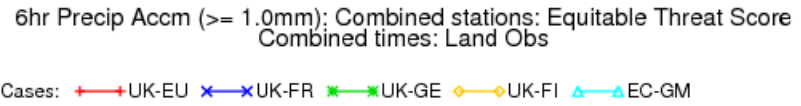


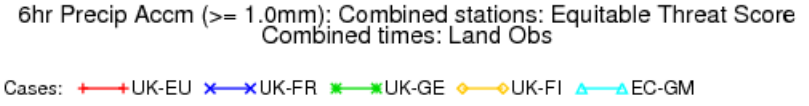


Summer day bias, convection

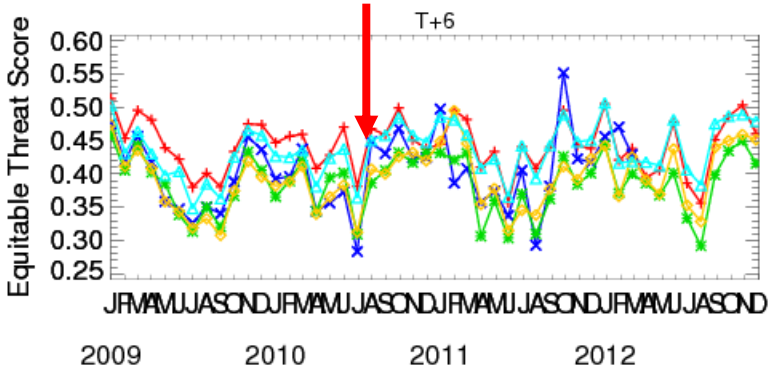
Night negative bias

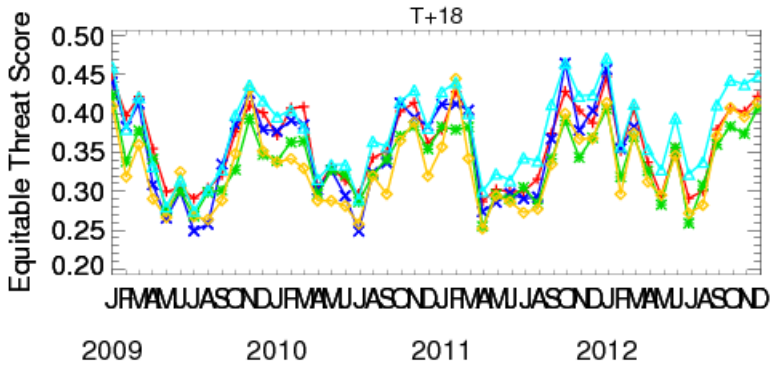
ppn ETS >1mm/6h



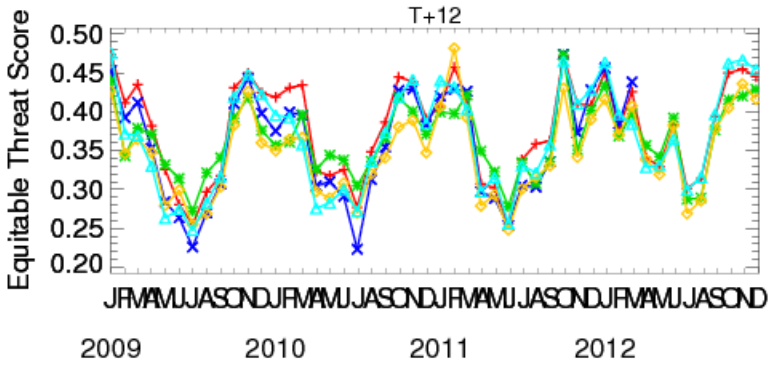


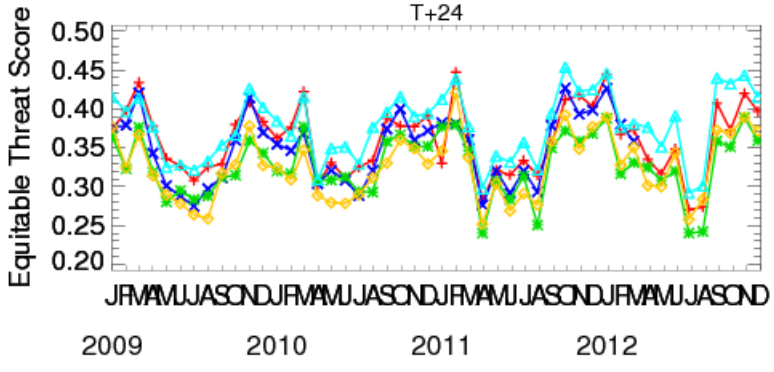
DA – radar/cloud

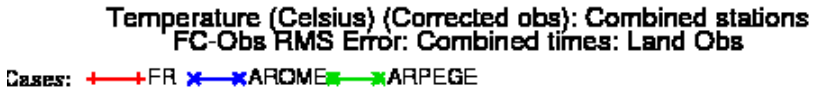


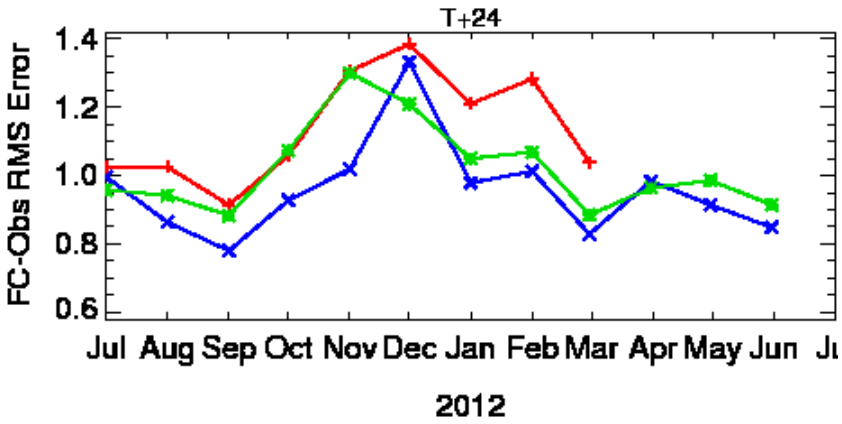


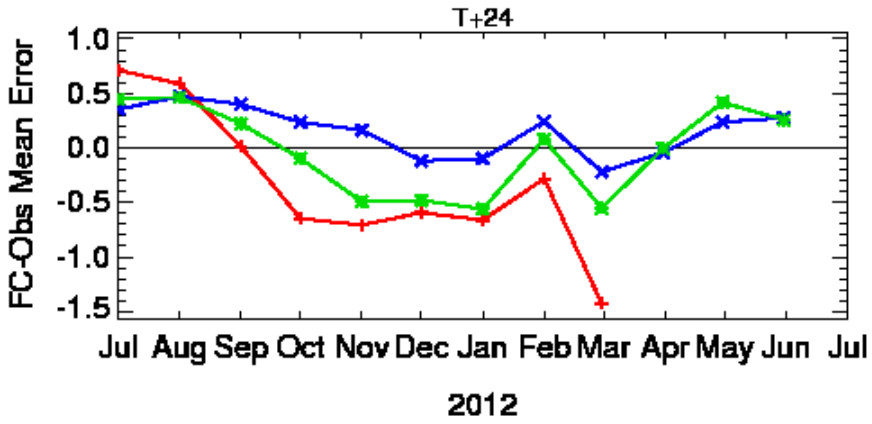
better

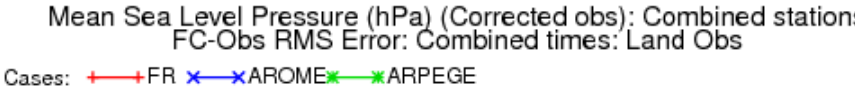


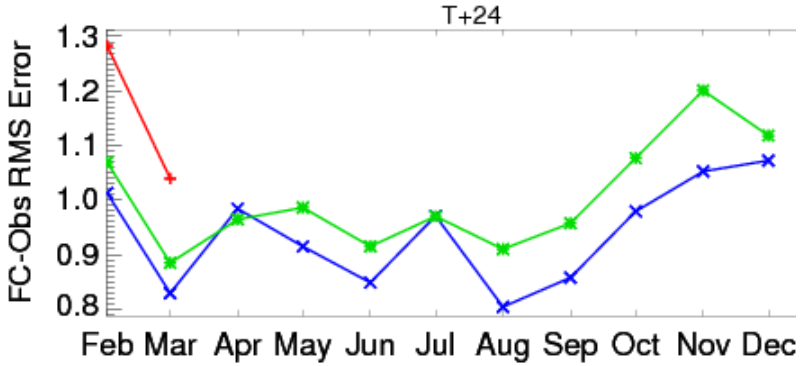


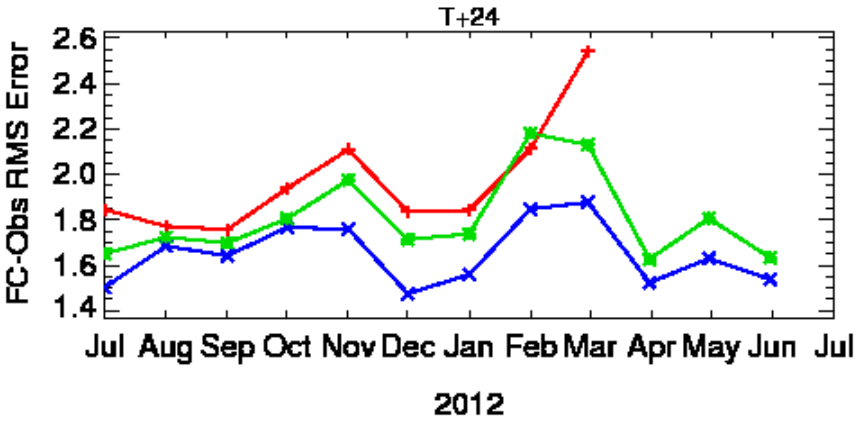










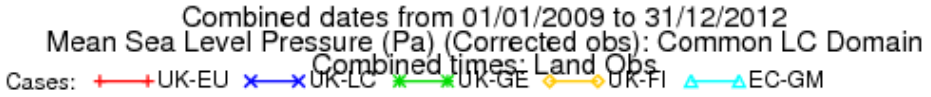


ALADIN v AROME v ARPEGE

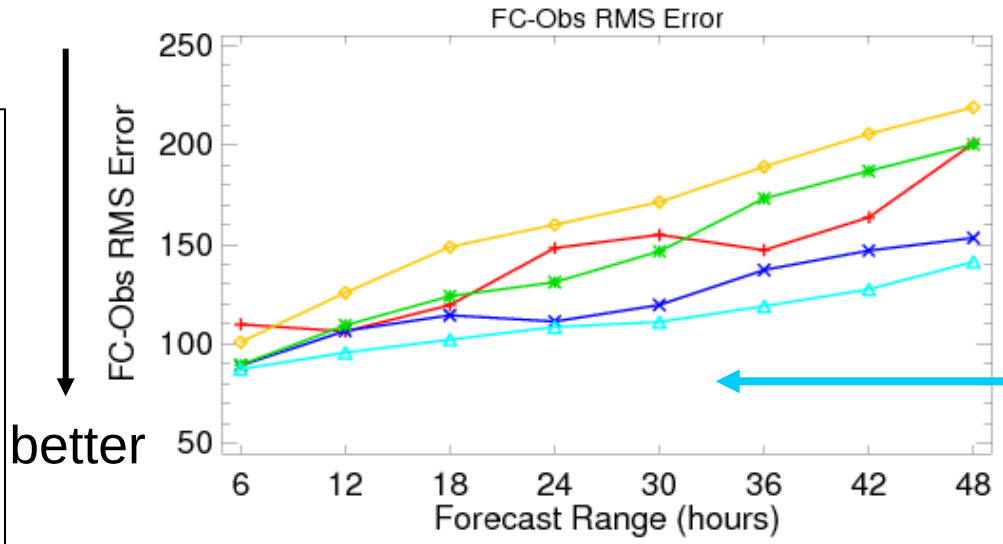
Sea level Pressure

ALADIN-LACE domain

Mean 01/2009-12/2012

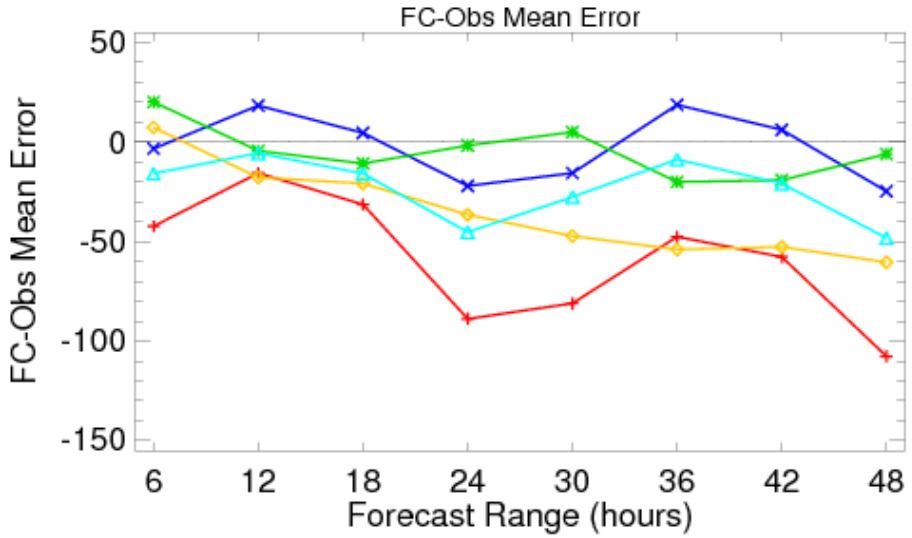


LC



better

EC



AlLADIN LC closer to EC than ALADIN FR

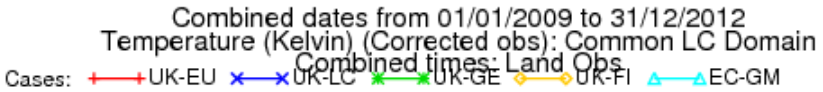
Met Office

Screen temperature

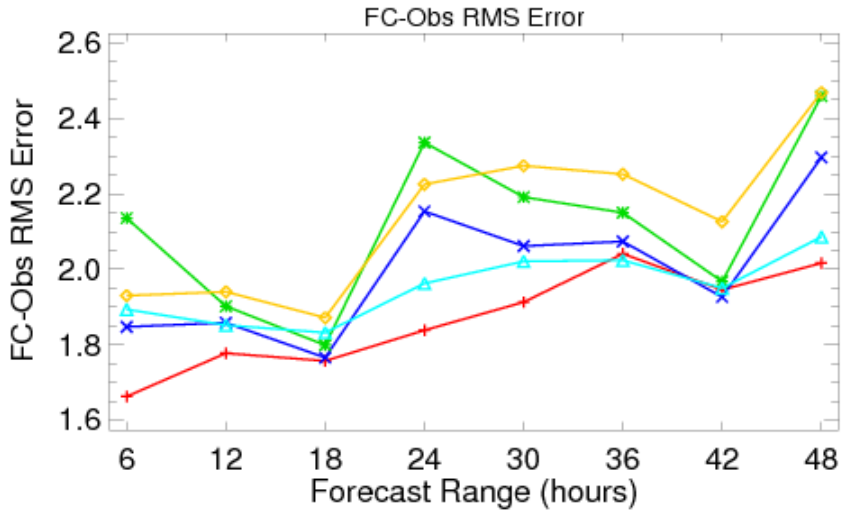
ALADIN-LACE domain

Mean 01/2009-12/2012

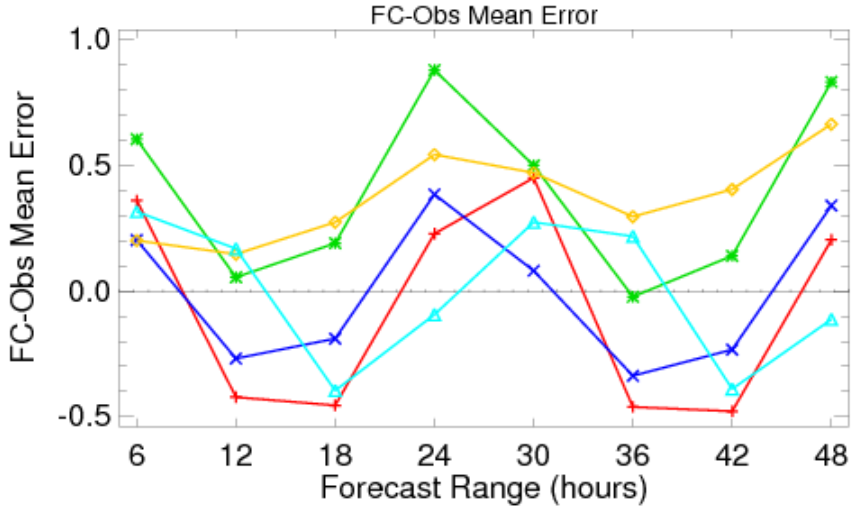
ALADIN LC closer to EC than ALADIN FR



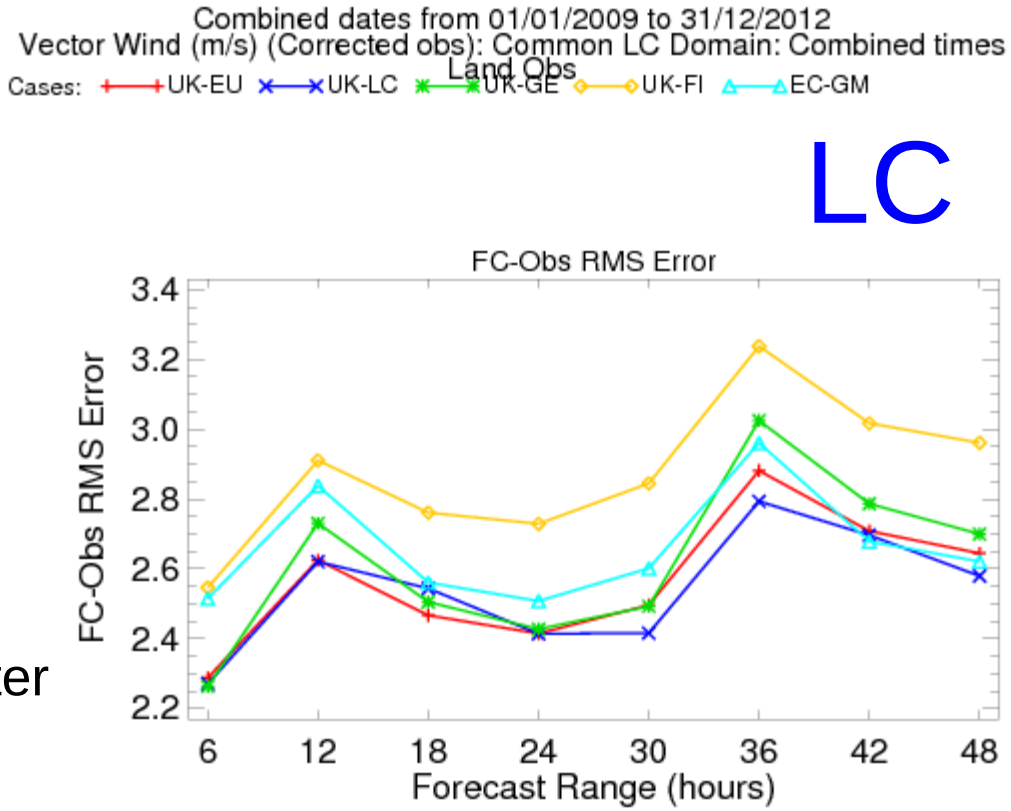
LC



better



Met Office



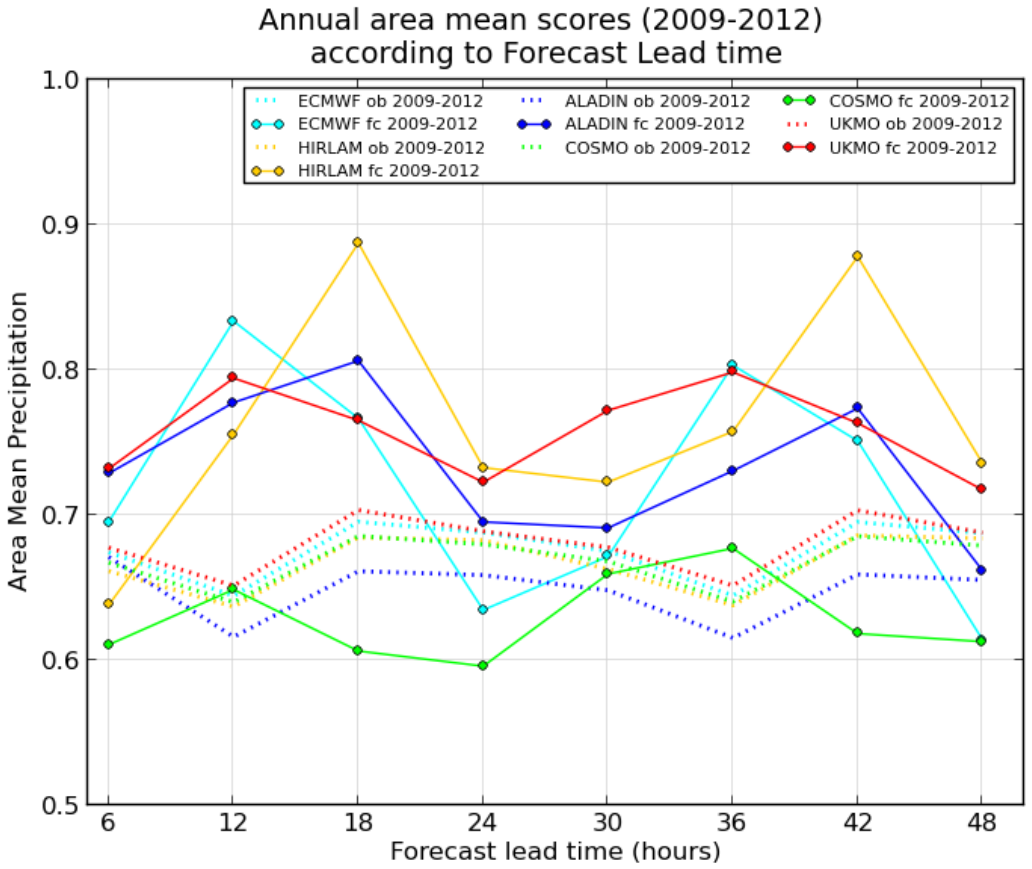
10m vector wind rmse ALADIN-LACE domain
Mean 01/2009-12/2012

better

Regional models better/= than global in short range to T+36,
ALADIN LC better to T+48
Hirlam (coarsest resolution) inferior to EC

Annual mean area precipitation, 2009-2012

- Most models overforecast
  - COSMO exception
  - under but less
- Too strong diurnal variation
  - esp. ECMWF
  - UM best amplitude
- Aladin and Hirlam better phase diurnal (+18, +42)
- UM 12km and ECMWF peak too early (+12,+36)
- COSMO anti-phase



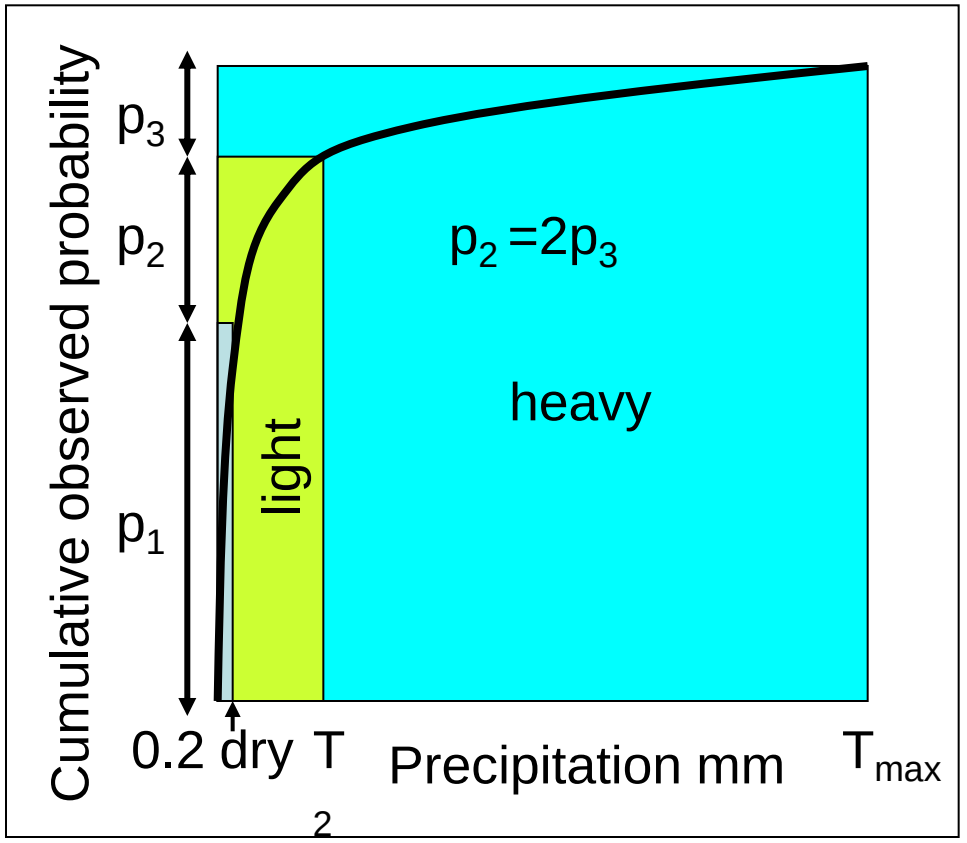
New scores

SEEPS

SEEPS=Stable Equitable Error in Probability Space

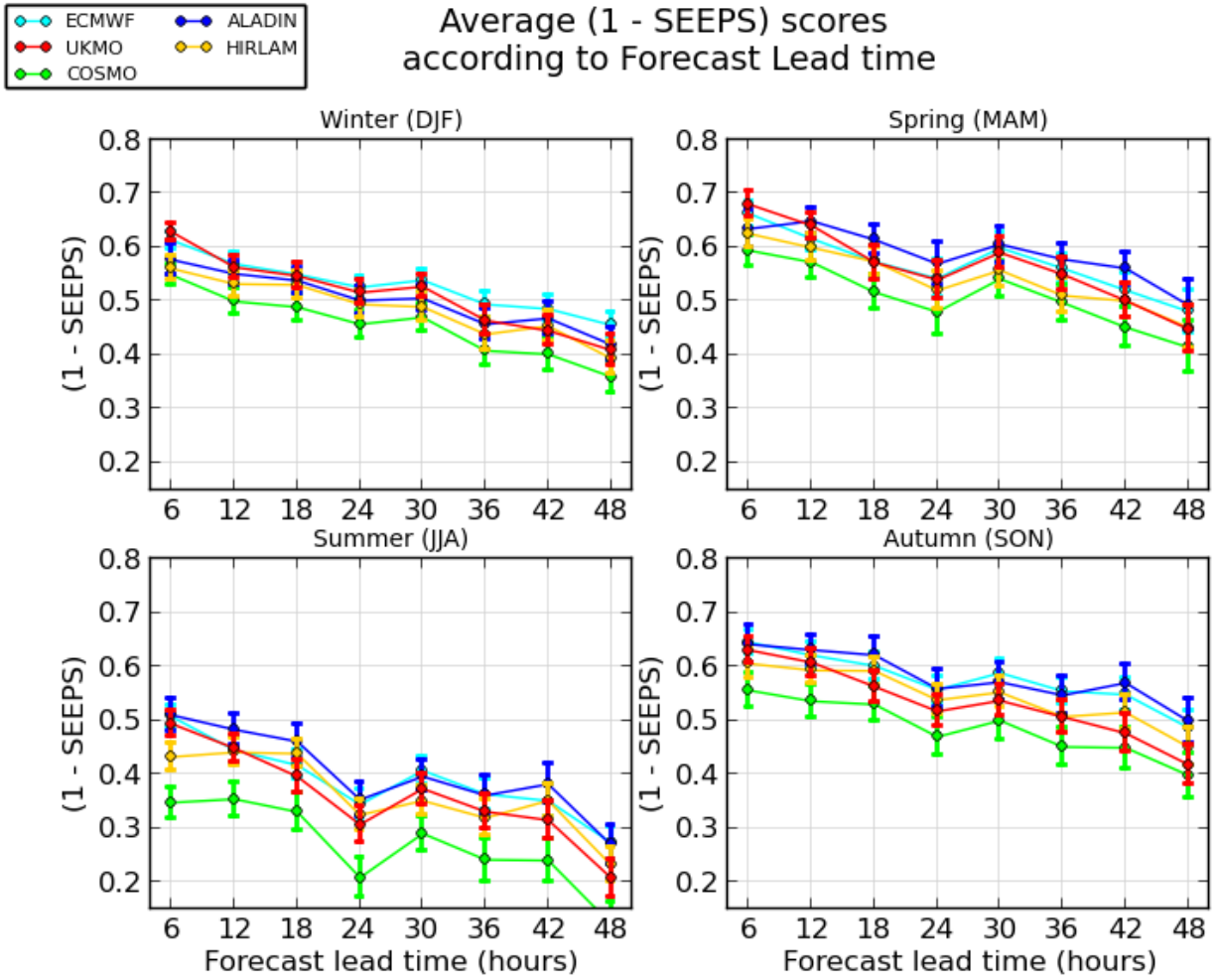
Rodwell et al, 2010, QJRMS 136

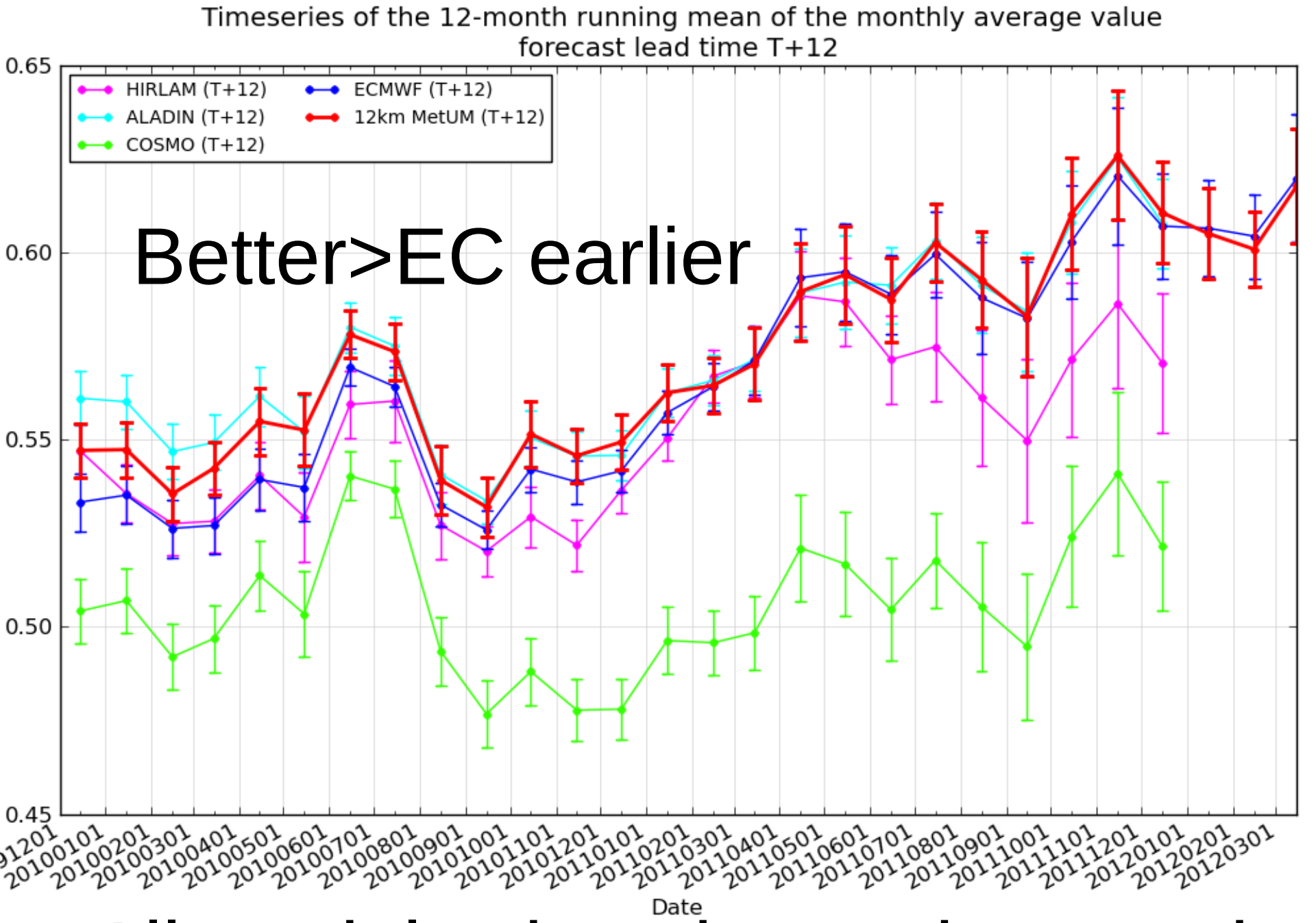
- Dry, light , heavy based on observed climatology (24h) at station – p_1 , p_2 , p_3
- Contingency table probabilities based on these categories
- Scoring matrix – stable, equitable
  - SEEPS=0 (perfect) , =1 (no skill - , eg constant)
- Applied to 6h accumulations in SRNWP-V
  - 6h climatology (courtesy Mark Rodwell)



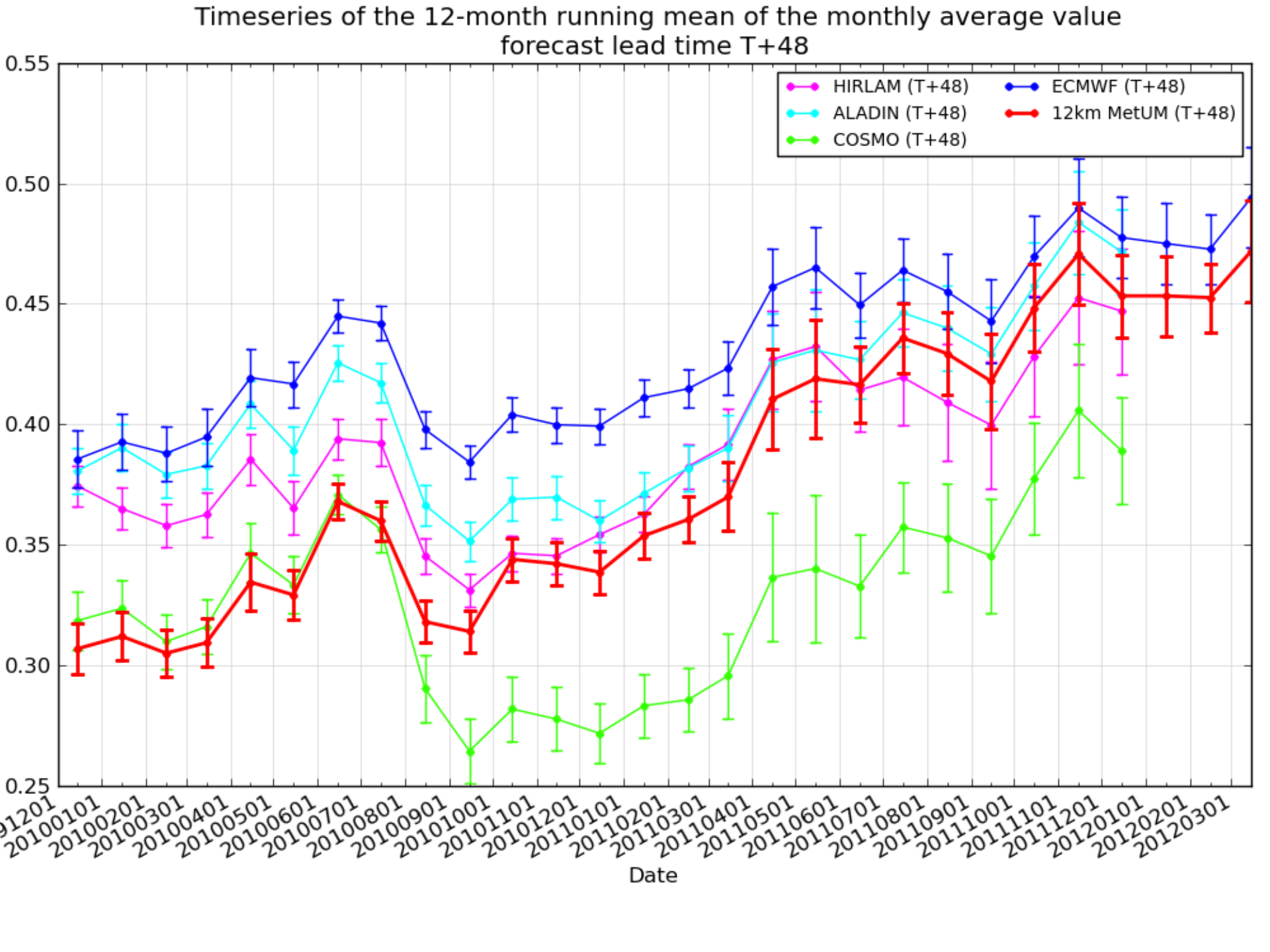
1-SEEPS with error bars, (Higher = better)

- SEEPS uses DRY, LIGHT & HEAVY categories based on climate at each station
- Best autumn & spring, worst summer – dip at midnight
- COSMO least skilful





All models show improving trend


Timeseries of the 12-month running mean of the monthly average value
forecast lead time T+48
HIRLAM (T+48)
ALADIN (T+48)
COSMO (T+48)
ECMWF (T+48)
12km MetUM (T+48)
0.55
0.50
0.45
0.40
0.35
0.30
0.25
91201
20100101
20100201
20100301
20100401
20100501
20100601
20100701
20100801
20100901
20101001
20101101
20101201
20110101
20110201
20110301
20110401
20110501
20110601
20110701
20110801
20110901
20111001
20111101
20111201
20120101
20120201
20120301
Date

Severe/High Impact weather verification

Extreme Dependency Scores

Extremal dependency scores

	Observed event	Observed nonevent	Totals
Forecast event	$a \rightarrow 0$ Hits	$b \rightarrow 0$ False Alarms	$qn \rightarrow 0$
Forecast nonevent	$c \rightarrow 0$ Misses	$d \rightarrow n$ Correct Rejections	$(1-q)n \rightarrow n$
Totals	$pn \rightarrow 0$	$(1-p)n \rightarrow n$	$n \rightarrow n$

Hit rate $H = a/(a+c)$ False Alarm rate $F = b/(b+d)$

Base rate $=p$ Forecast rate $=q$

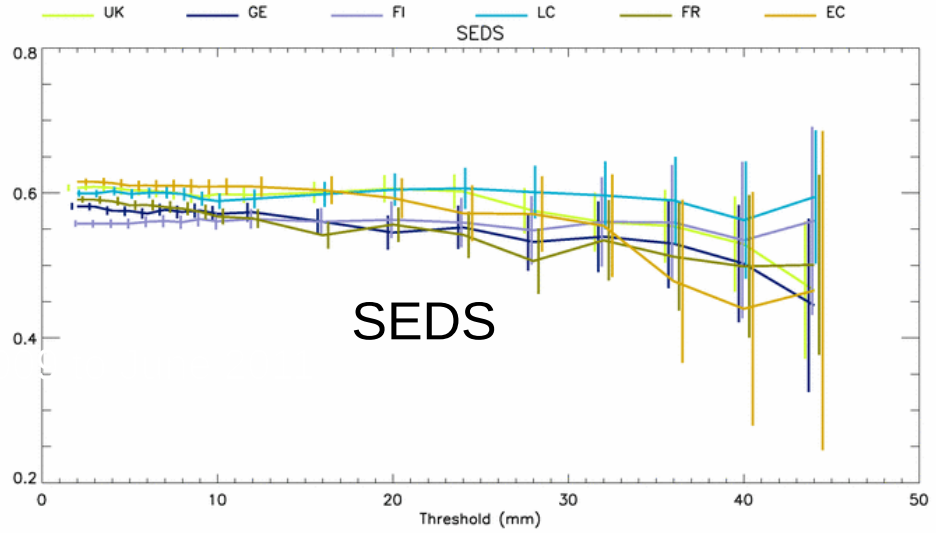
$$\mathbf{SEDS} = \frac{\log(q) - \log(H)}{\log(p) + \log(H)} \qquad \mathbf{SEDI} = \frac{\log(F) - \log(H) - \log(1-F) + \log(1-H)}{\log(F) + \log(H) + \log(1-F) + \log(1-H)}$$

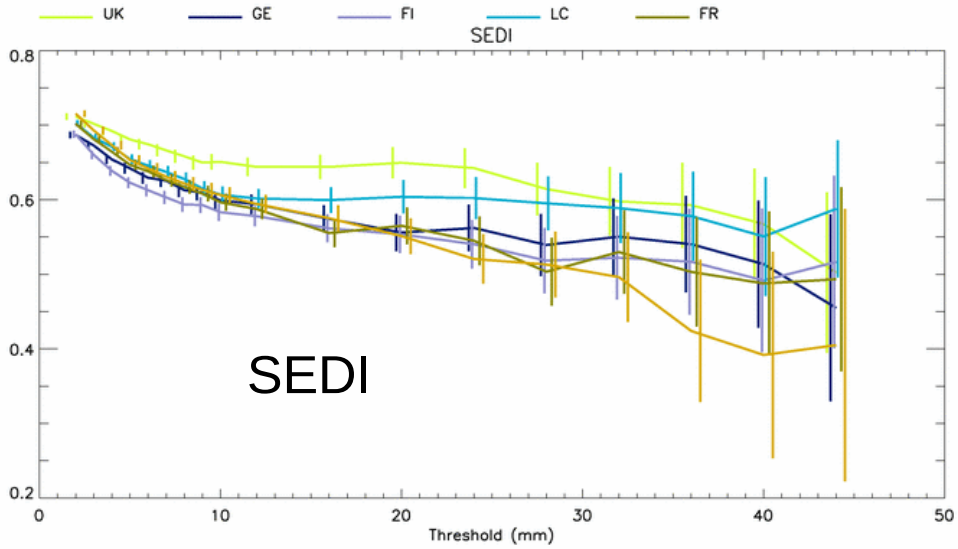
Matthew Trueman

Extremal dependency scores

$$SEDS = \frac{\log(q) - \log(H)}{\log(p) + \log(H)} \qquad SEDI = \frac{\log(F) - \log(H) - \log(1-F) + \log(1-H)}{\log(F) + \log(H) + \log(1-F) + \log(1-H)}$$

Jan 2009 – June 2011



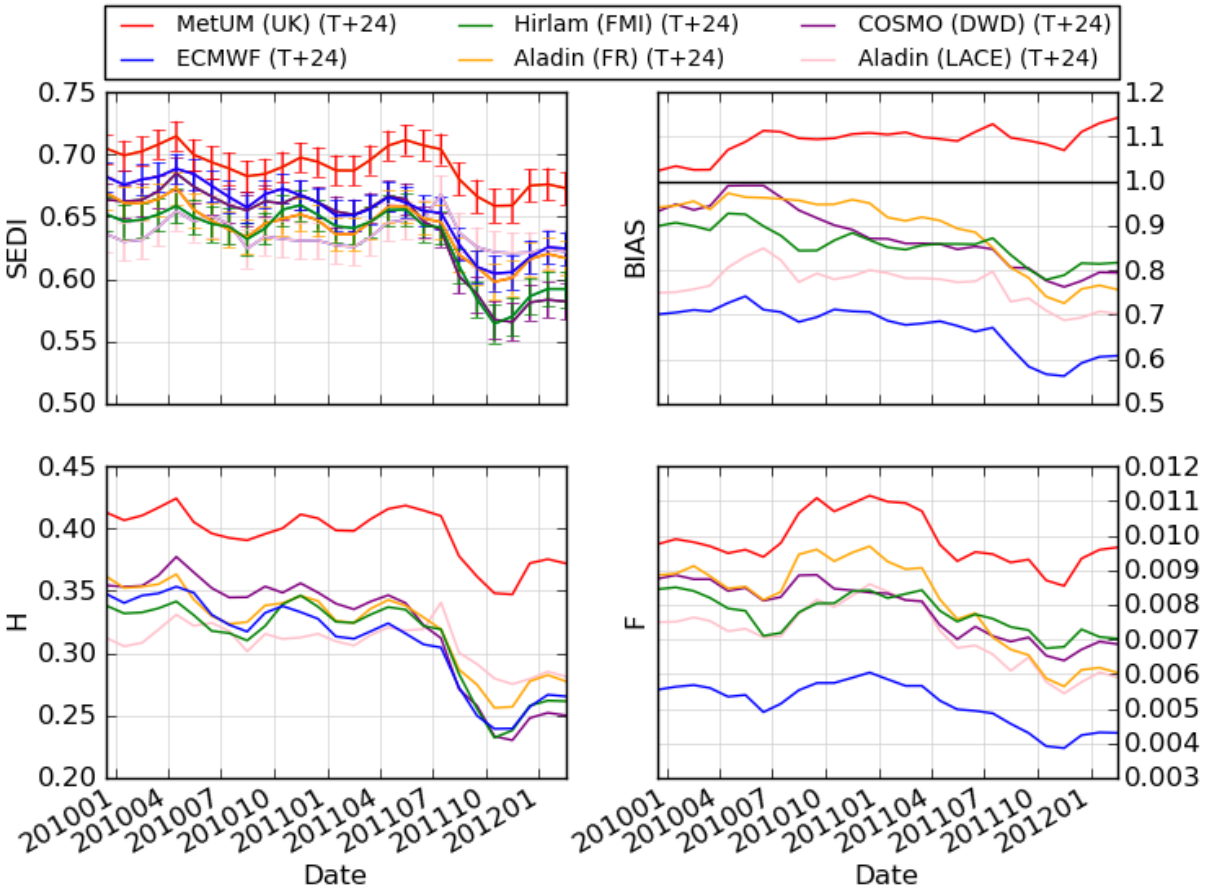


Matthew Trueman

Extremal dependency scores

$$SEDI = \frac{\log(F) - \log(H) - \log(1-F) + \log(1-H)}{\log(F) + \log(H) + \log(1-F) + \log(1-H}$$

- 99th percentile threshold +95% error bars
- LAMS have better bias than ECMWF
- EC under-forecast extremes
- H, hit rate LAMS better



Rachel North

Additional parameters

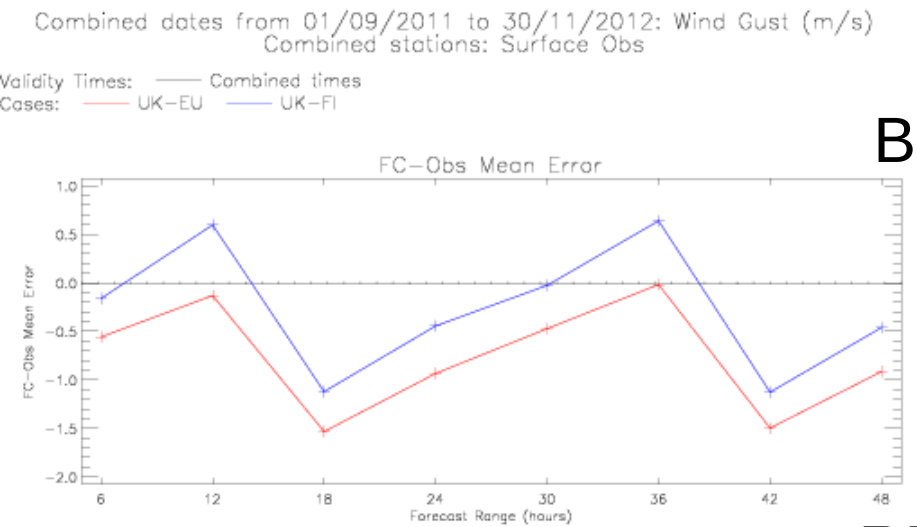
- cloud
- wind gusts

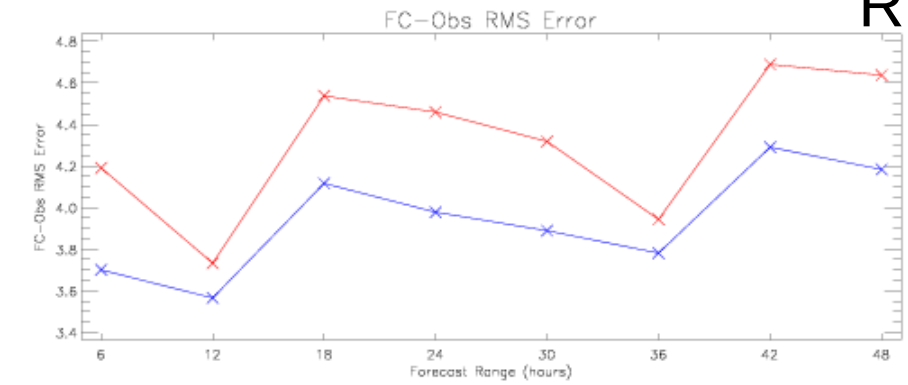
Gusts 14 months

UM and COSMO too low

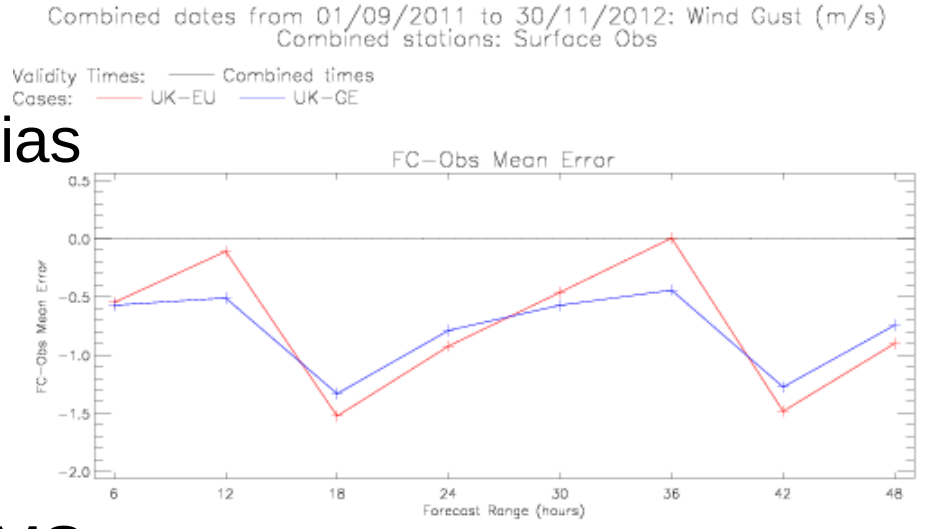
UM worst RMS, Hirlam and COSMO very similar

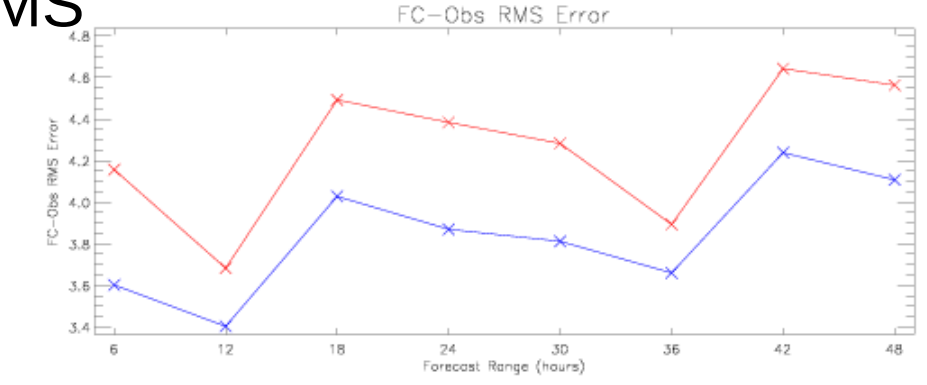
UM v Hirlam





UM v COSMO





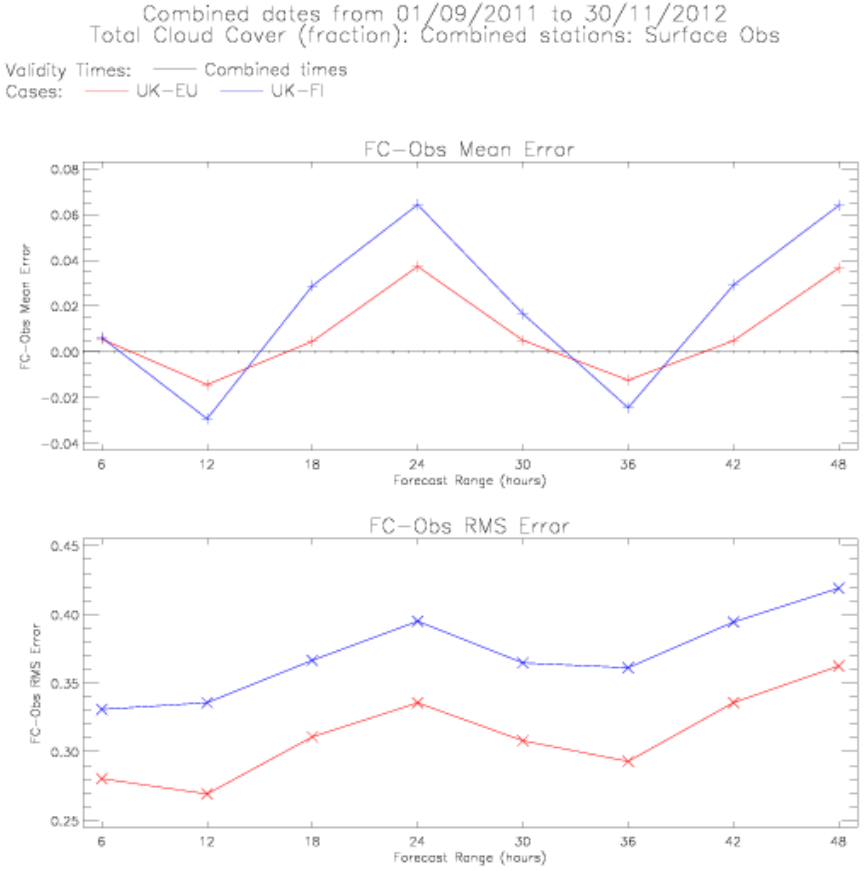
Bias

RMS

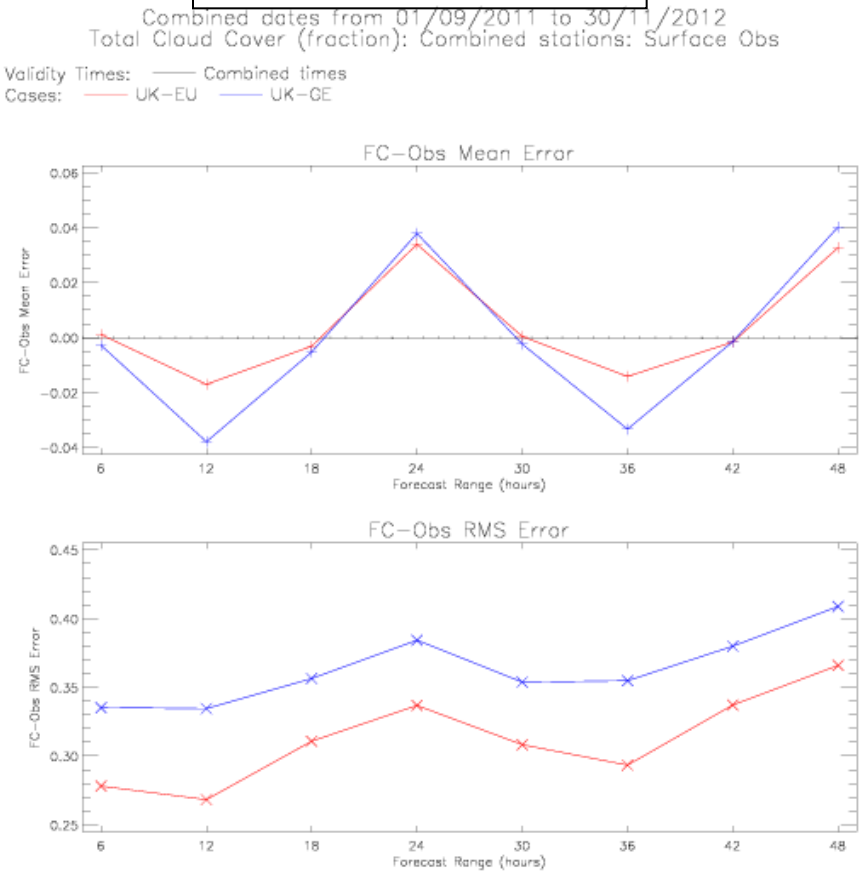
Total cloud 14 months

Day/night -/+ bias
UM best RMS, Hirlam and COSMO very similar

UM v Hirlam



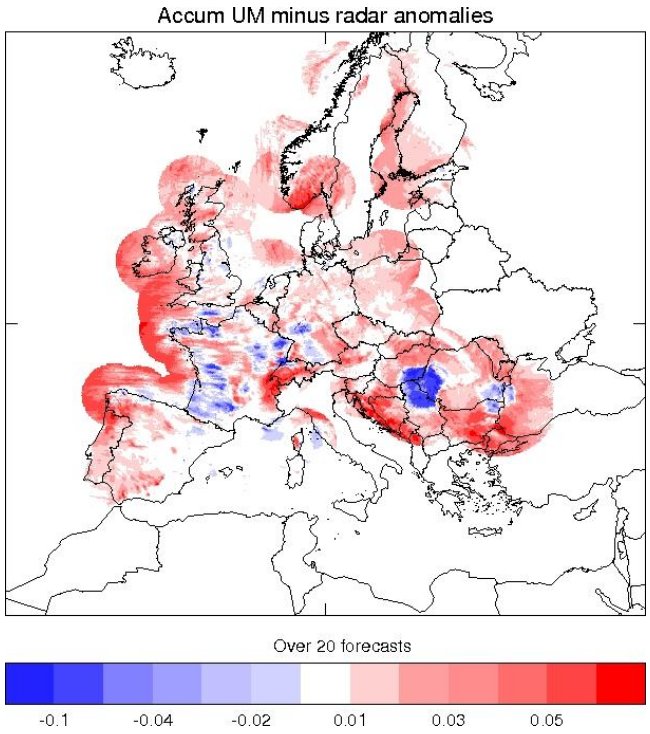
UM v COSMO

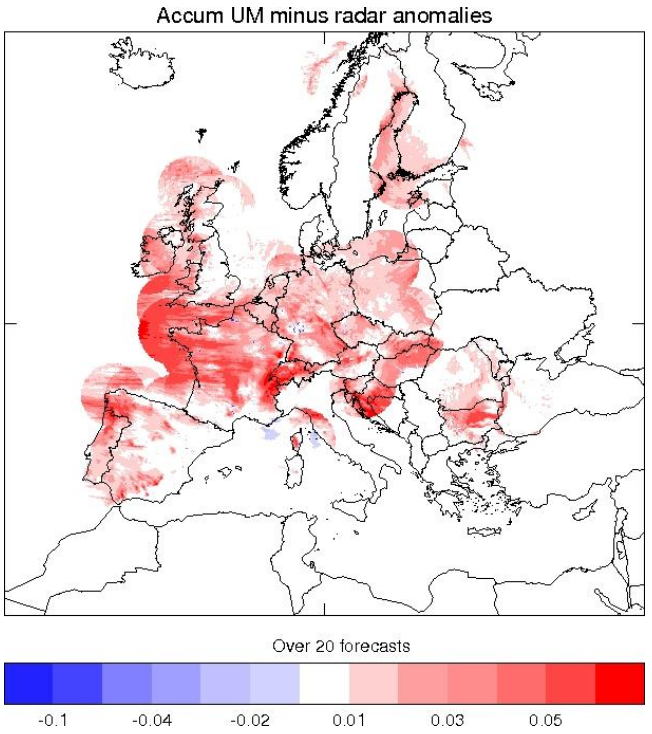


OPERA Odyssey evaluation – improved QC

comparison against 4km Euro UM

With further QC (VPR) should be suitable for verification use





ND3: Spatial & scale selective verification of precipitation and cloud

- Verify against
  - Gridded analyses- ECMWF, Meteo-France, Met Office (UK only)
    - Other national gridded sets ?
  - High resolution radar (5 min,1-2 km) -UK
  - OPERA radar composite QC / gauge bias correction

- Methods
  - Fractional skill (Roberts & Lean)
  - Upscaling
  - Intensity scale (Casati)
  - Structure,amplitude,location (SAL) (Wernli et al)
  - Contiguous rain areas (Ebert & McBride)

Met Office

Fuzzy verification

using SEVERI cloud mask as truth binary 0 or 1

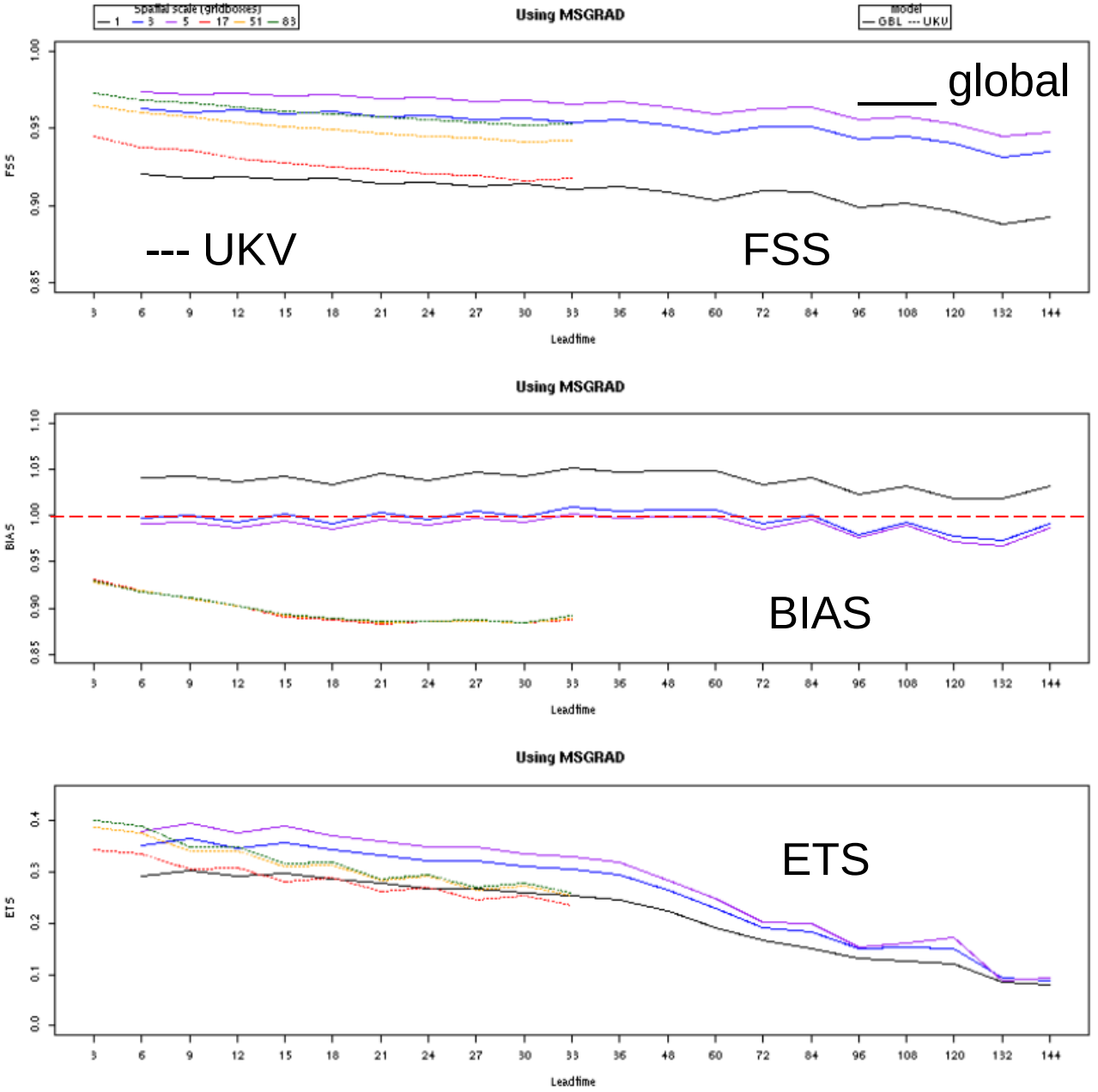
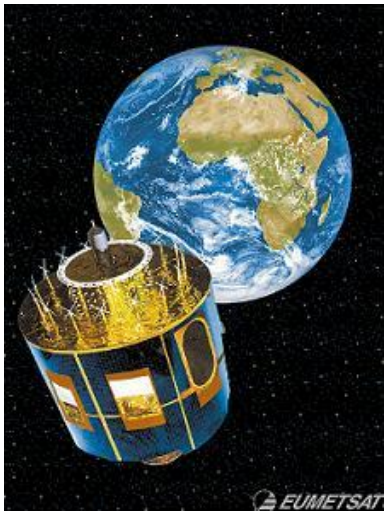
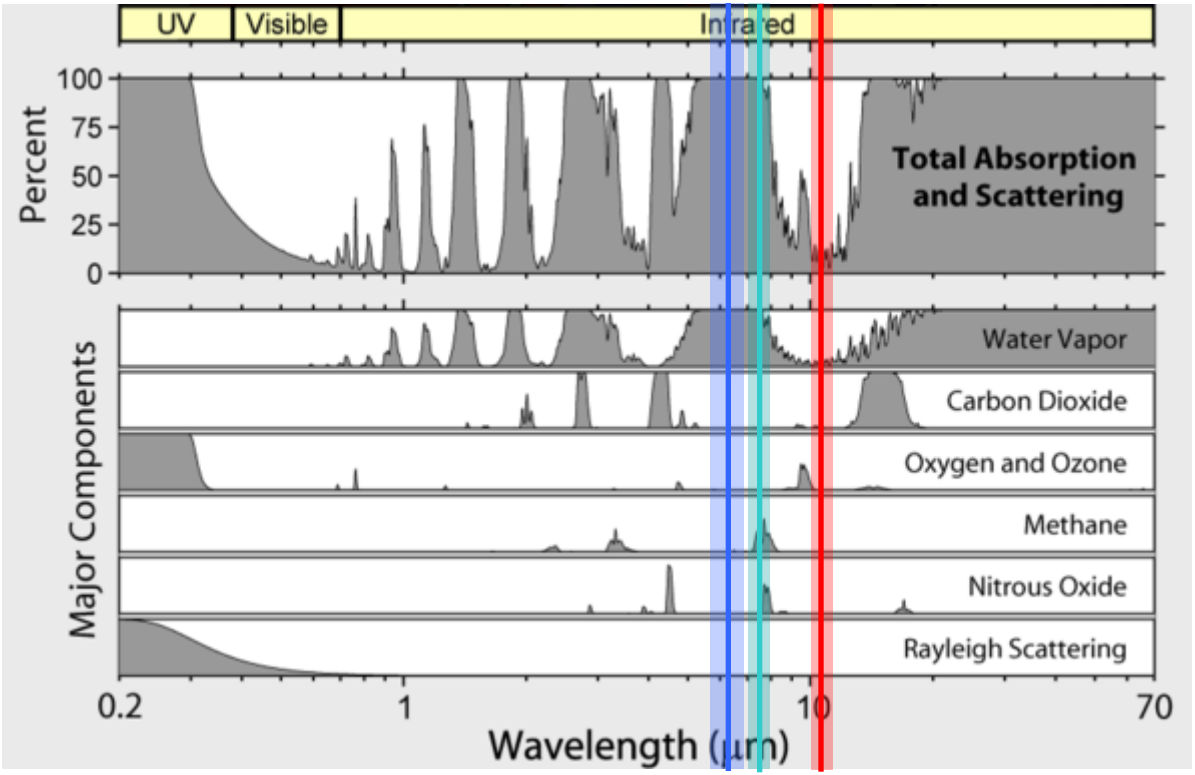


Figure 1 - FSS, bias and ETS scores for a cloud threshold >0.375 when verified against the MSGRAD cloud mask data. All spatial scales of the GM and UKV models are displayed.

Simulating images to compare to SEVIRI[1]



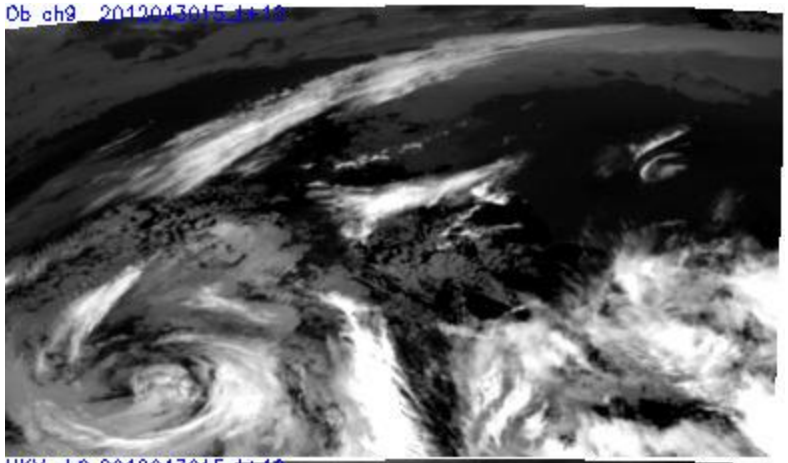


Ch05 6.2 µm - upper water vapour sensitive

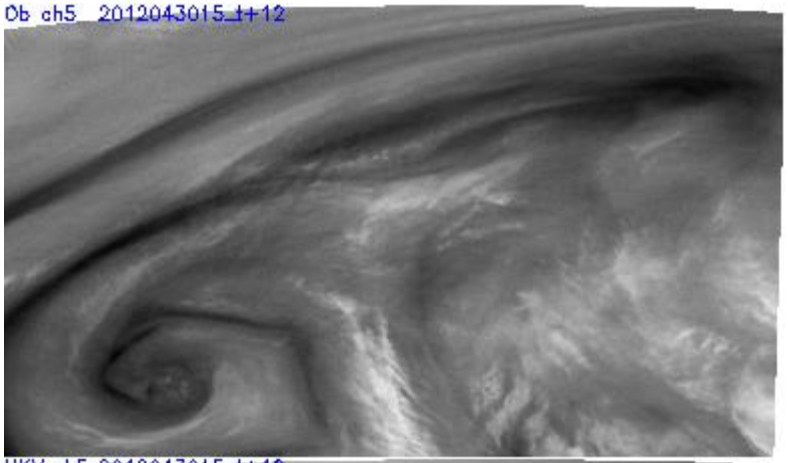
Ch06 7.3 µm

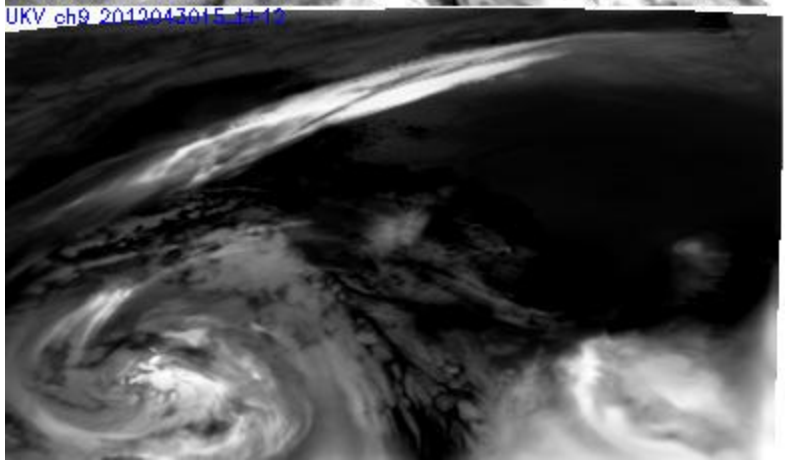
Ch09 10.8 µm- detects cloud and surface

1. Spinning Enhanced Visible and Infrared Imager



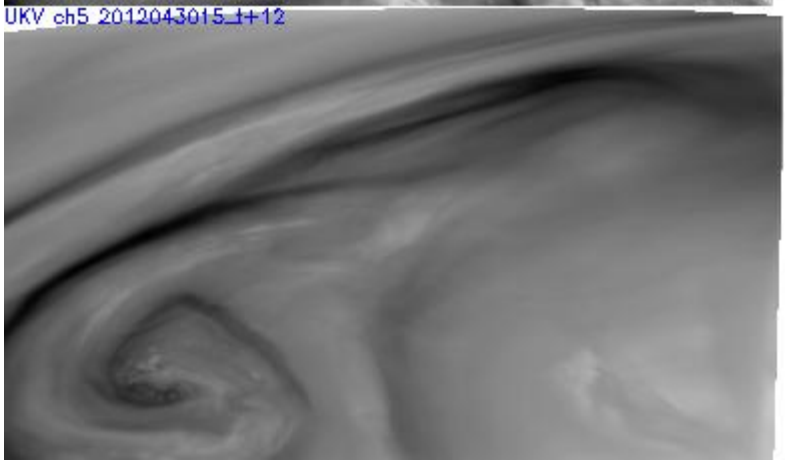
SEVIRI





UKV model

T+12

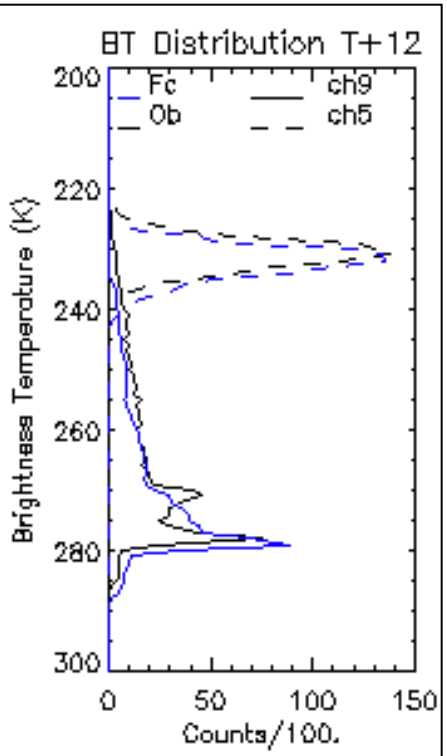


ch5

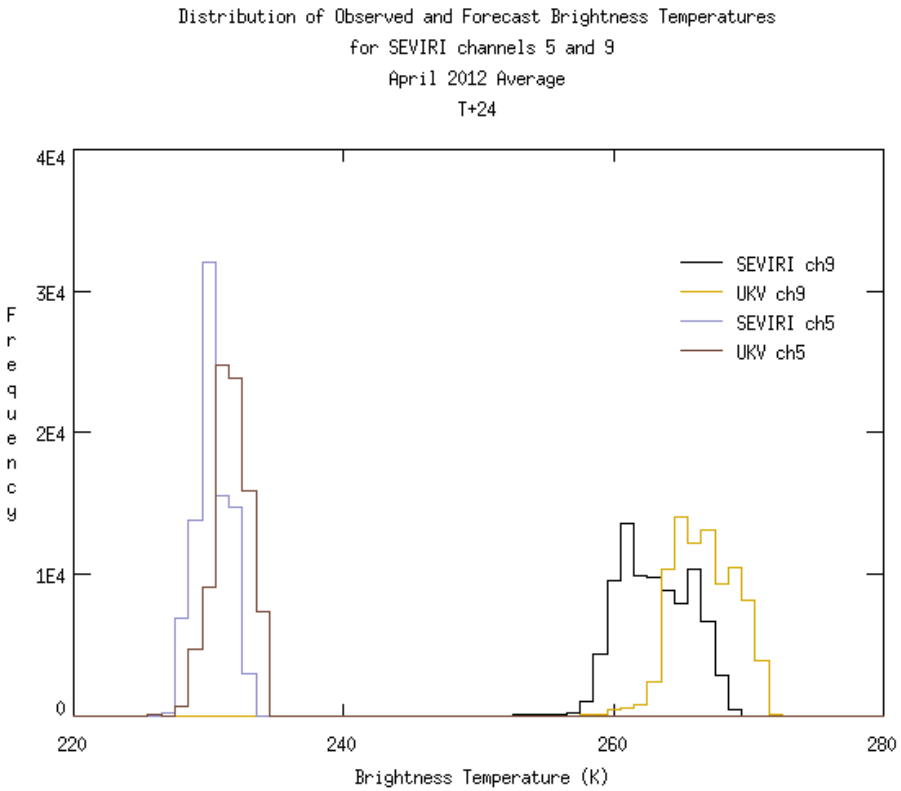
ch9

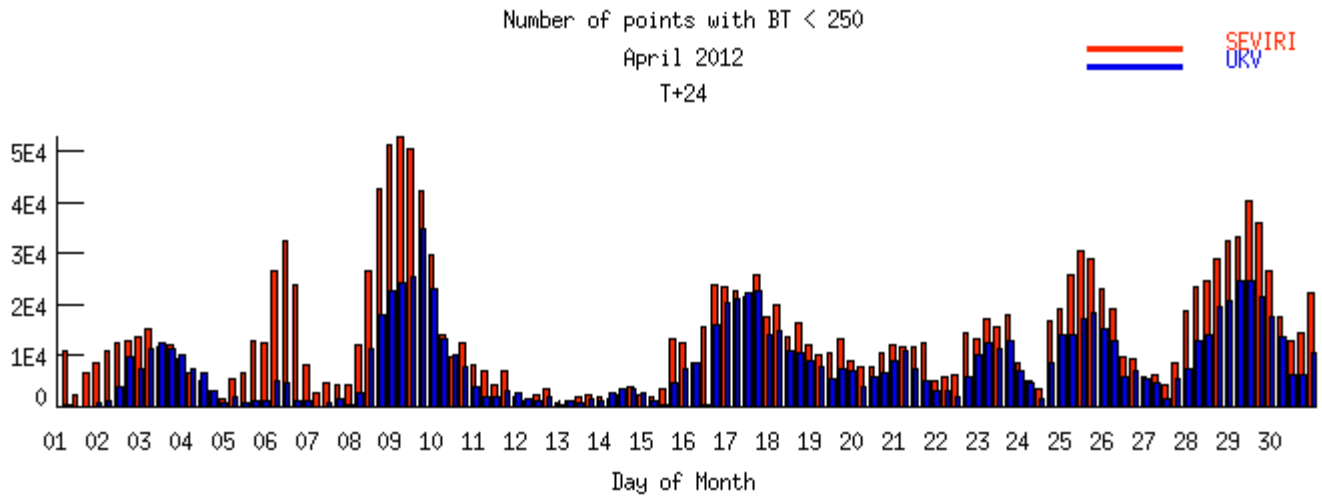
15UTC 30th April 2012

Bob Tubbs



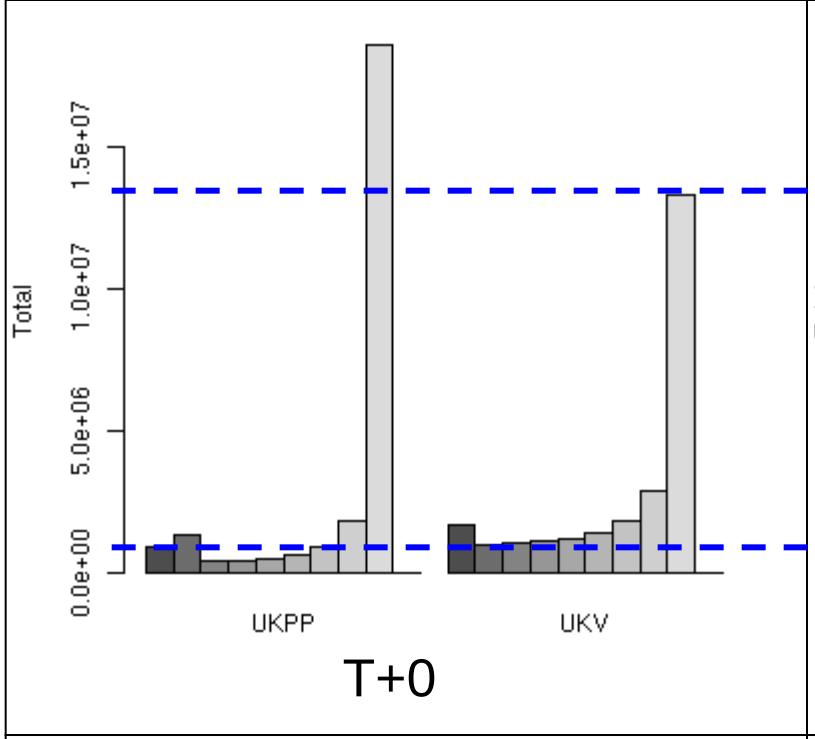
15UTC 30th April 2012

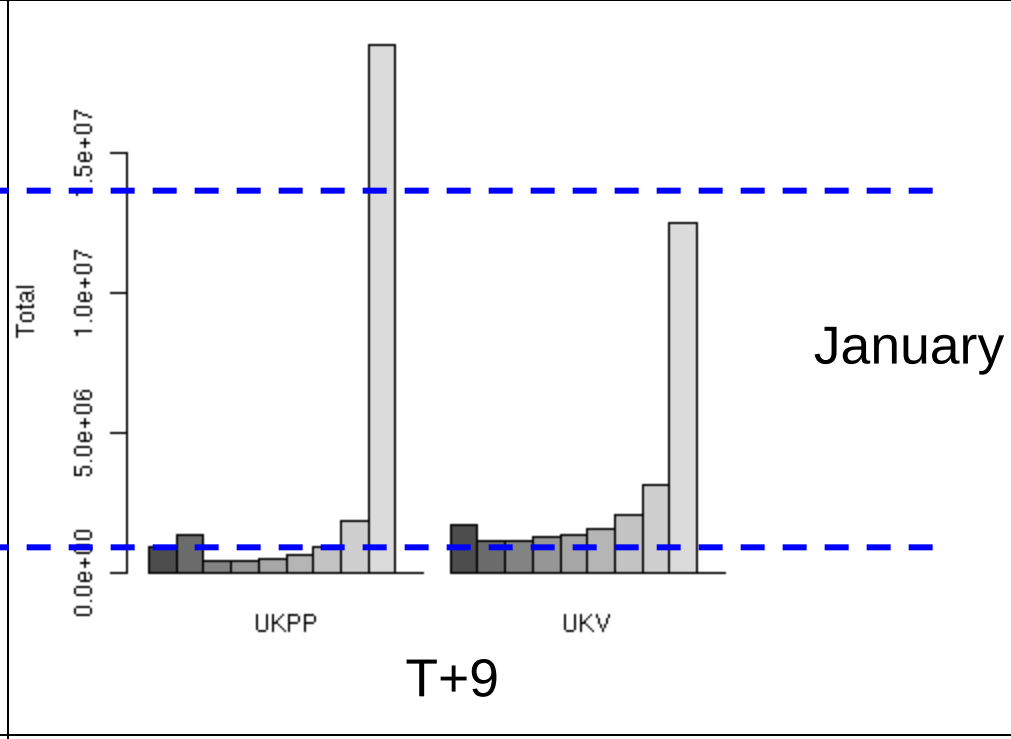


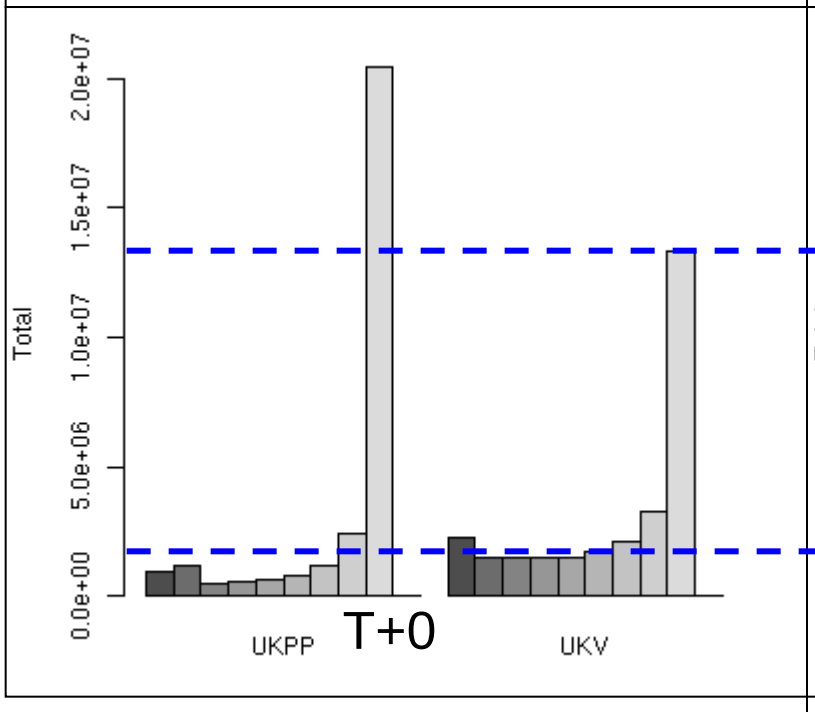


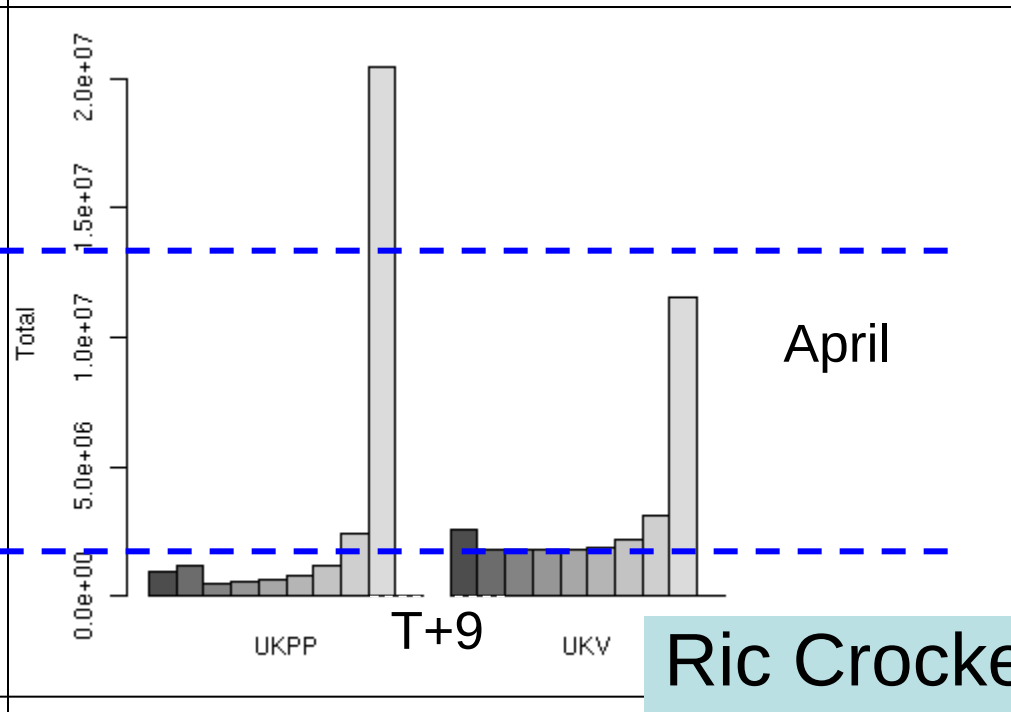
Ric Crocker

E


Total
0.0e+00
5.0e+06
1.0e+07
1.5e+07
UKPP
UKV
T+0


Total
0.0e+00
5.0e+06
1.0e+07
1.5e+07
UKPP
UKV
T+9
January

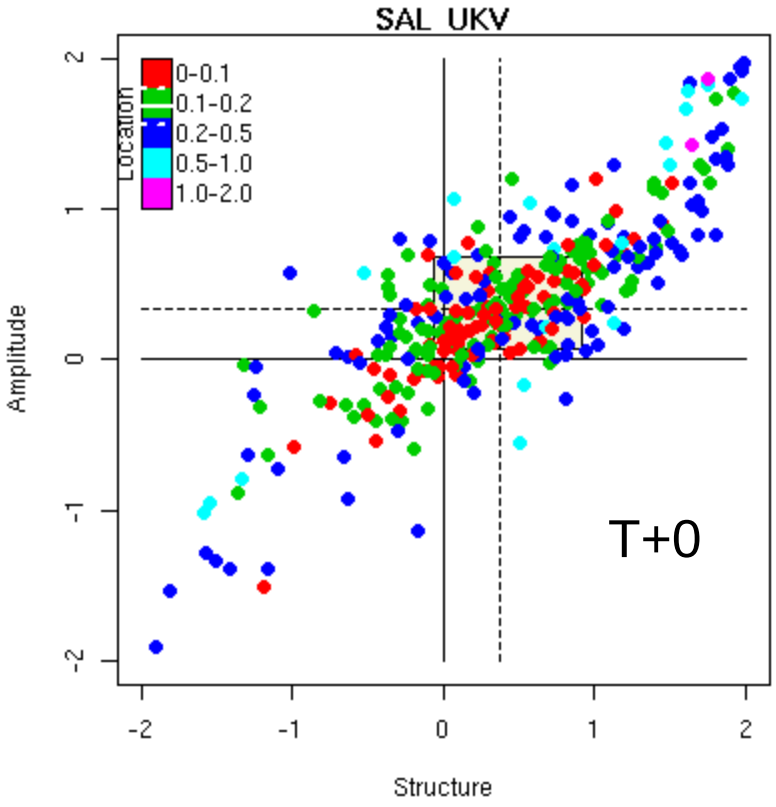

Total
0.0e+00
5.0e+06
1.0e+07
1.5e+07
2.0e+07
UKPP
UKV
T+0

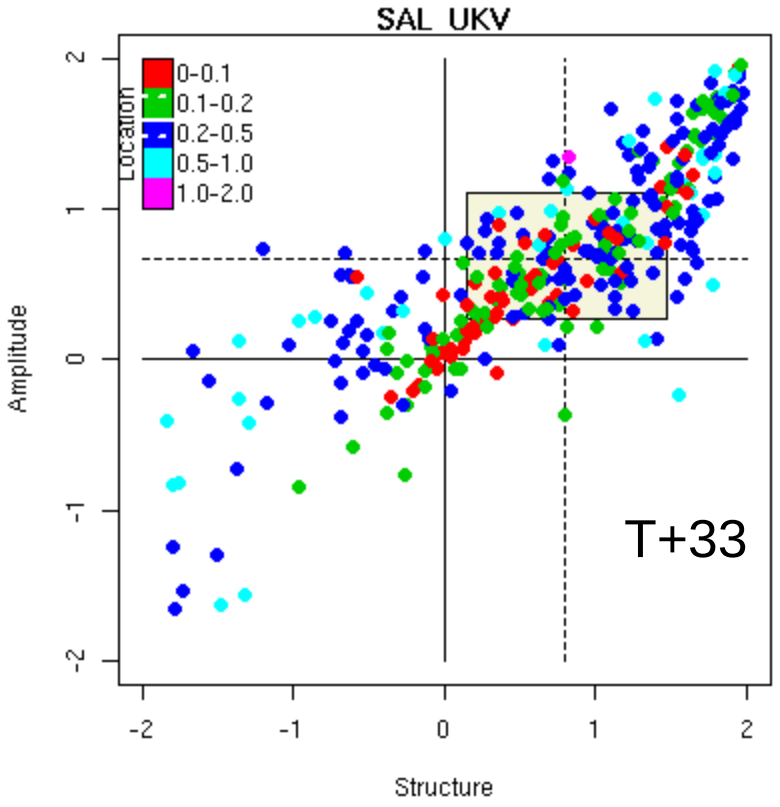

Total
0.0e+00
5.0e+06
1.0e+07
1.5e+07
2.0e+07
UKPP
UKV
T+9
April

Ric Crocker

SAL for clear area (negative cloud threshold). UKV model compared to cloud analysis

- Positive S,A – too little cloud
- Increase with forecast – even less cloud and bigger holes





Other deliverables

- Full documentation of the methods used in the intercomparison.
- Newer spatial methods code to be portable.

List of reports on EUMETNET portal

- Verification Results Jan 2009 to December 2012
- Stable Equitable Error in Probability Space (SEEPS) score analysis for SRNWP-Verification – Rachel North
- Deliverable ND3: Assessment of OPERA/Odyssey composites - Marion Mittermaier and Robert Scovell
- Cloud Verification methods and Met Office experience – Ric Crocker (& Bob Tubbs)
- *Final report and Recommendations*

Thanks - Questions

SRNWP-V Phase 2
Deliverable completed in RED

- ND1 Continue & expand comparison
  - Longer more robust results – up-to-date, publication of results on EUMETNET Portal
  - higher resolution of future operational models
    - AROME and ARPEGE results processed since start of 2011
  - SEEPS scores calculated over common domain
- ND2 Additional products verified
  - Cloud amount/base
    - truth ? Auto/manual different biases
    - Distribution approach
  - Wind gust –short validation /compariosn