Application and verification of ECMWF products 2011

Met Office – compiled by Marion Mittermaier

1. Summary of major highlights

ECMWF products continue to be widely used across the Met Office. Input has been requested from many parts of the organisation: Forecast Service Delivery (Guidance Unit), model development groups for the short- and medium-range as well as ensembles.

2. Use and application of products

2.1 Post-processing of model output

2.1.1 Statistical adaptation

None.

2.1.2 Physical adaptation

None.

2.1.3 Derived fields

None.

2.2 Use of products

Forecaster input has been solicited. Two questions were posed:

1. What products have you started using in the last year in particular that you like and found really useful? Why?

2. Any specific aspects re the EC resolution upgrade that stand out for you, that you may want to comment on (subjective verification).

Edited response:

"Products I've started using in the last year include the WREP display system, the ECMWF cyclone database Tim Hewson has developed and the TC genesis ensemble products.

WREP is really useful. It's level of interactivity is useful (e.g. allowing the user to zoom in and determine which fields they see), as is viewing the products in 3 hourly steps. Tim's products are good, especially the Dalmatian plots which allow a quick assessment of ensemble spread."

Not sure I've got much to say about subjective verification. It'd be useful if EC made some model characteristics available." – Dan Suri

3. Verification of products

3.1 Objective verification

3.1.1 Direct ECMWF model output (both deterministic and EPS)

The Met Office is the lead centre in an SRNWP-V (verification) initiative to compare surface weather parameters from regional forecast models over Europe. The ECMWF deterministic forecast is included as an independent reference forecast. As ECMWF forecasts are assessed as part of a comparison results are shown in the next subsection.

3.1.2 ECMWF model output compared to other NWP models

i) Surface parameters *(Marion Mittermaier)*

It is not appropriate that we show full SRNWP comparison results here, but restrict ourselves to showing monthly verification results of surface parameters between the 12 km Met Office North Atlantic European (NAE) Unified Model (UK-EU) and the equivalent area from the operational deterministic ECMWF forecast (EC-GM).

For 2 m temperature, pmsl and 10 m vector wind four lead times: t+12, 24, 36 and 48h are shown for the 00Z runs only. There is a big difference in run lengths between the NAE and ECMWF deterministic forecasts, the NAE runs extending to only 48 hours, with emphasis on the short-range. This difference is reflected in the scores which show that at t+48h ECMWF is more evidently superior, with the main emphasis on the medium range beyond day 2.

Fig.1 shows the monthly *rmse* for 2 m temperature. All the lead times reflect the difficulties models experienced with the temperature forecasts in December 2010, with the ECMWF *rmse* worse at 3 of the four lead times. Whilst cold season temperature errors are more comparable there are clear signs that the NAE has struggled with both midnight and midday temperatures in the warmer months especially. It is worth noting though that the verifying domain is very influential on the results. Over the common SRNWP-V domain (which happens to be the Aladin domain and is a sub-region of the NAE domain), the signal is quite different.

The same sequence is shown for pmsl in Fig. 2. The impact of lateral boundaries is evident with increasing lead time, where the NAE pmsl is much degraded compared to ECMWF which has no artificial boundaries. Only at t+12h does the NAE show signs of being better, or at least comparable.

Fig.3 shows the 10 m vector wind *rmse*. Whilst there is no clear seasonal cycle in the midday *rmse* at t+12h and t+36h, there is more evidence for midnight forecasts to have worse *rmse* in the winter months. Overall the NAE has lower *rmse*, except at t+48h where ECMWF shows signs of lower *rmse*.

Overall the NAE is close to retirement with recent upgrades probably sub-optimal for this configuration. For short-range regional deterministic modelling the emphasis has shifted to the km-scale over the UK, whilst the NAE will morph into the control member of the revised MOGREPS-R regional ensemble system.

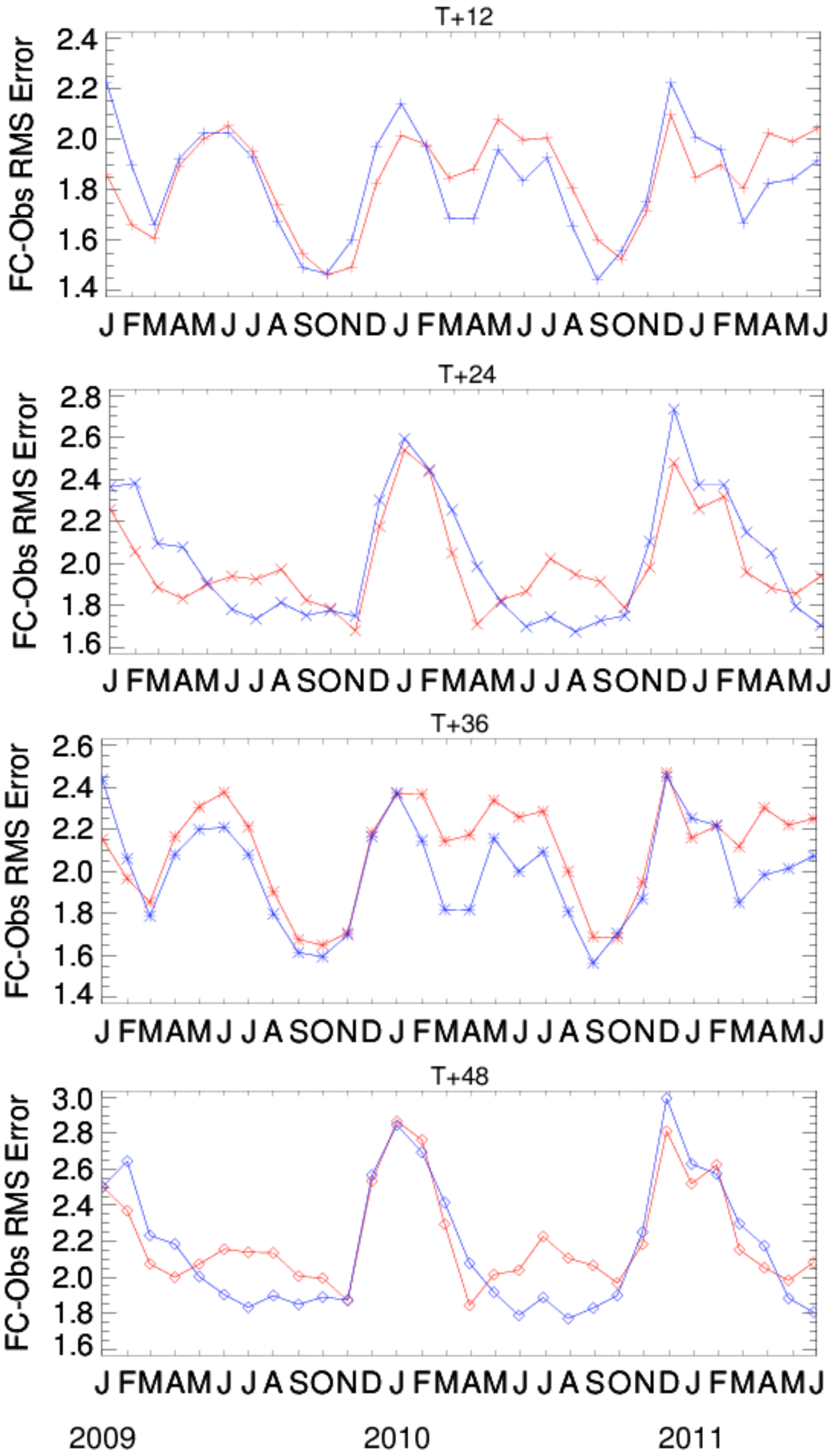


Figure 1 Comparison of a 2 m temperature (K) from the 00Z ECMWF deterministic and NAE forecasts. **Red = NAE**, **blue = ECMWF**.

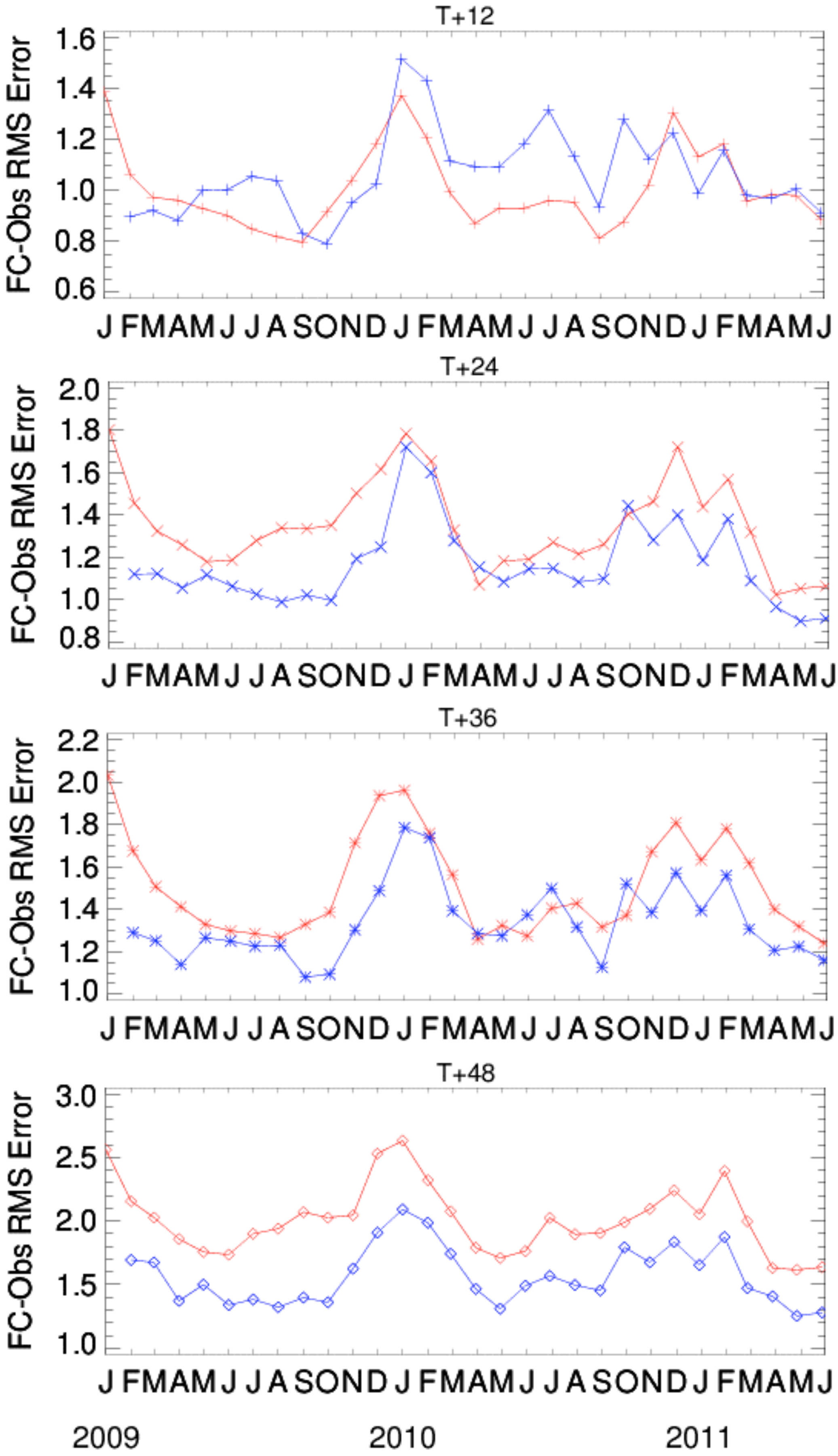


Figure 2 Comparison of a pmsl (hPa) from the 00Z ECMWF deterministic and NAE forecasts. **Red = NAE**, **blue = ECMWF**.

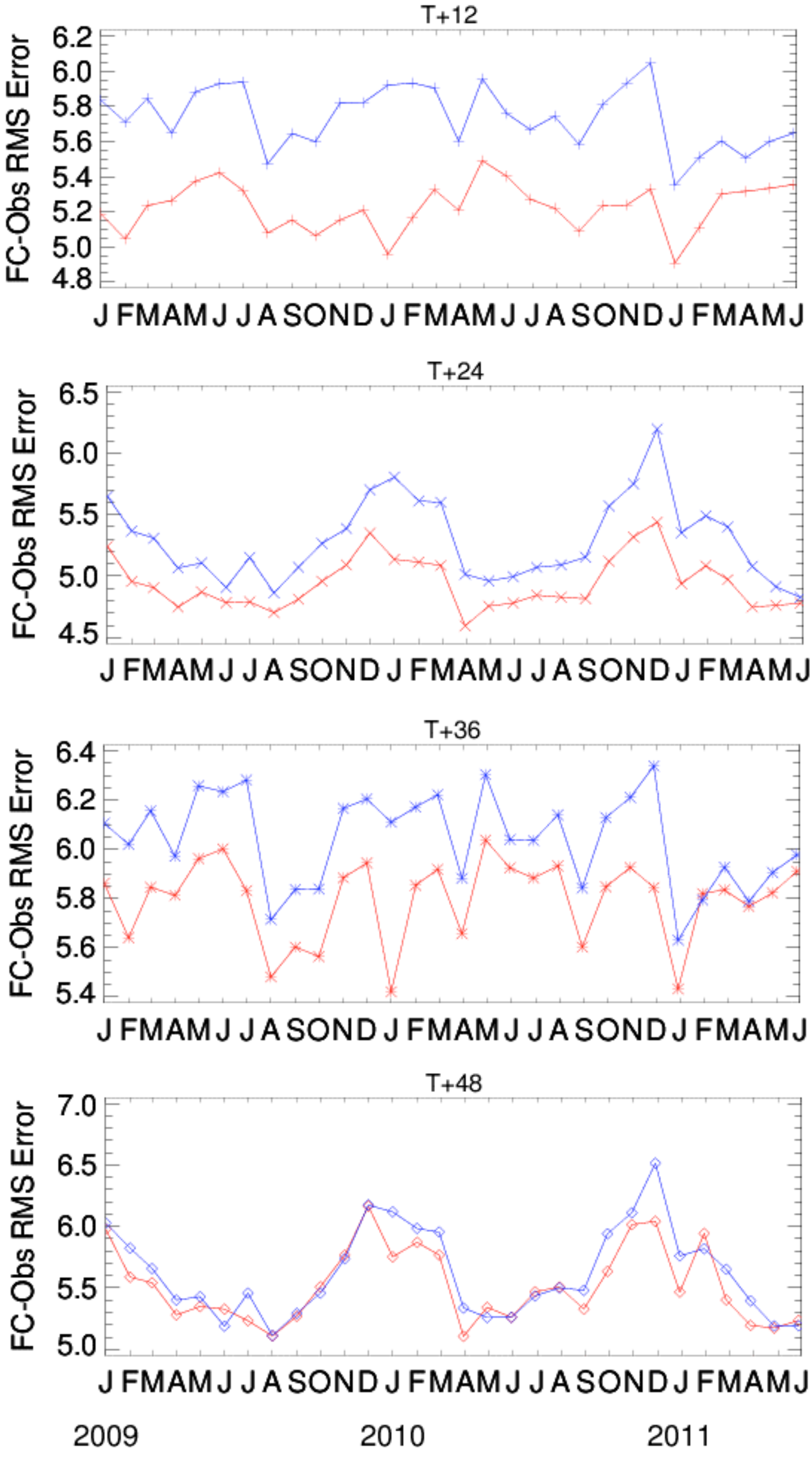


Figure 3 Comparison of a 10 m vector wind (knots) from the 00Z ECMWF deterministic and NAE forecasts. **Red = NAE**, **blue = ECMWF**.

In addition to what was shown last year, six-hourly precipitation is also considered categorically for two thresholds: >= 0.5 and 4 mm/6h in terms of the Equitable Threat Score (ETS) and the frequency bias. Two lead times are shown: t+12h and t+48h. This highlights the differences in run length, where ECMWF is clearly focusing on skill beyond day 2 whilst the NAE is purely for the short-range with error growth strong at t+48h.

At t+12h Fig. 4 shows that the both models have a positive frequency bias that fluctuates with the seasons, with biases increasing in the warmer months. In terms of the low thresholds at t+12h NAE performs better than the ECMWF forecasts.

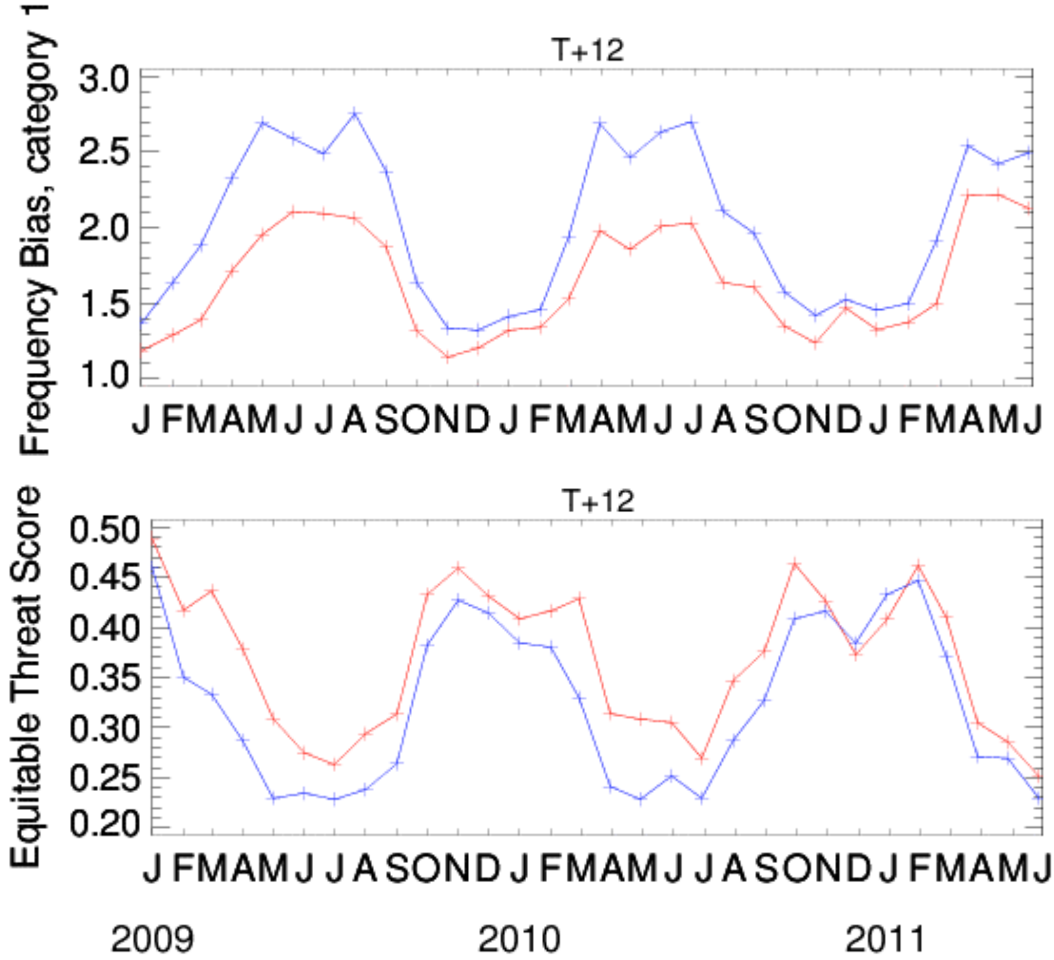


Figure 4 Comparison of a 0.5 mm/6h frequency bias and ETS at t+12h from the 00Z ECMWF deterministic and NAE forecasts. **Red = NAE**, **blue = ECMWF**.

At t+48h (Fig. 5) the situation changes to one where the biases are much reduced for the 0.5 mm/6h threshold, with the NAE bias better, but prone to slight under-forecasting in the summer months. The ETS seems to suggest that the ECMWF forecasts are now better. However, given the level of over-forecasting, the ECMWF ETS is probably inflated, as the score is sensitive to the bias and can be optimised through over-forecasting.

For the higher threshold of 4 mm/6h at t+12h Fig. 6 shows that, on the whole the NAE has a fairly neutral bias in the past year with ECMWF exhibiting more pronounced seasonal variations in the bias. Despite the differences in the bias, the ETS are far more similar for both models.

Fig. 7 shows that at t+48h the 4 mm/6h frequency bias tends towards under-forecasting for the ECMWF forecasts, with the NAE more neutral to under. At this stage the ETS suggests that the ECMWF forecasts are more skilful but the bias may again favour ECMWF scores.

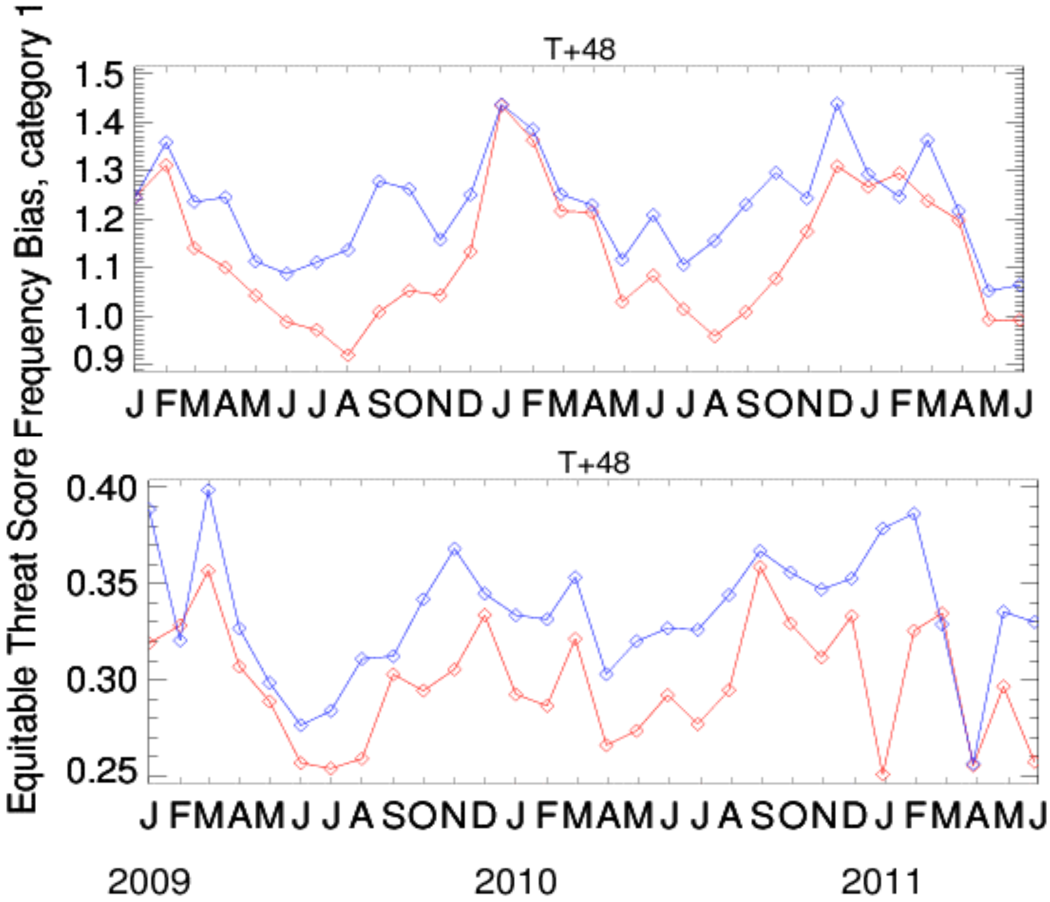


Figure 5 Comparison of a 0.5 mm/6h frequency bias and ETS at t+48h from the 00Z ECMWF deterministic and NAE forecasts. **Red = NAE**, **blue = ECMWF.**

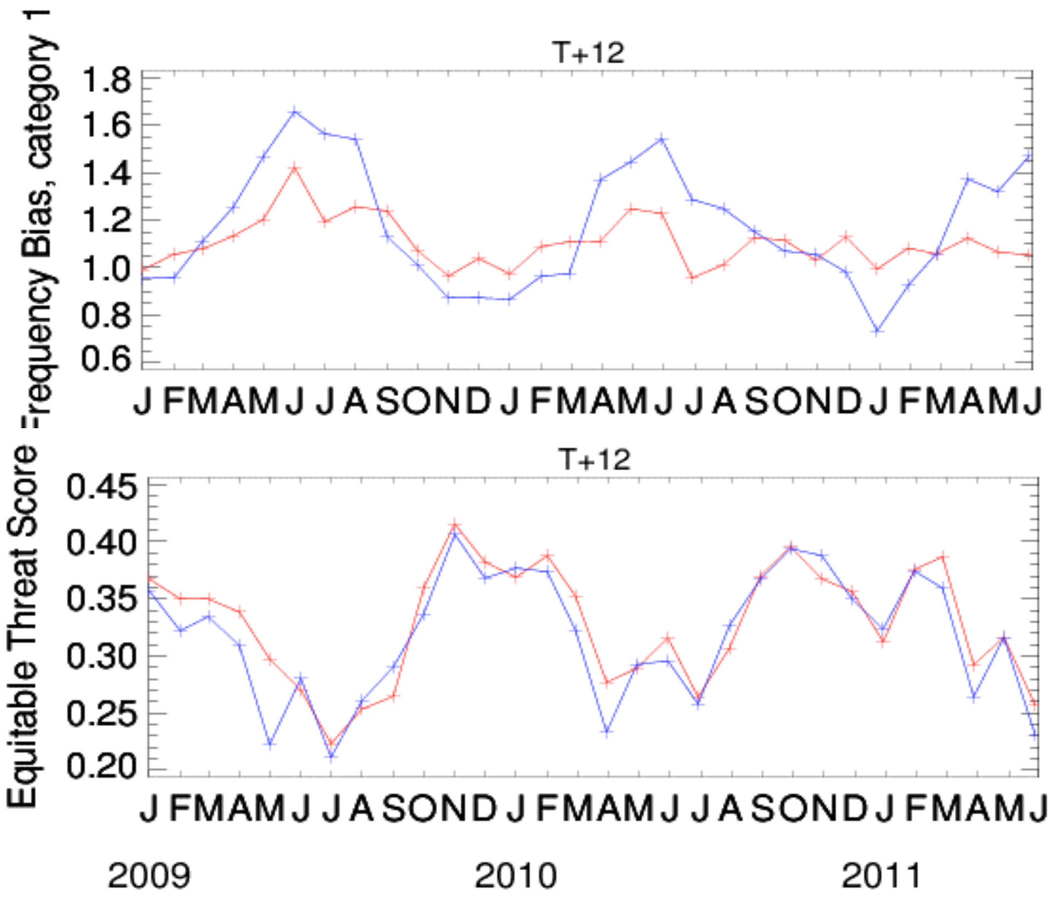


Figure 6 Comparison of a 4 mm/6h frequency bias and ETS at t+12h from the 00Z ECMWF deterministic and NAE forecasts. **Red = NAE**, **blue = ECMWF.**

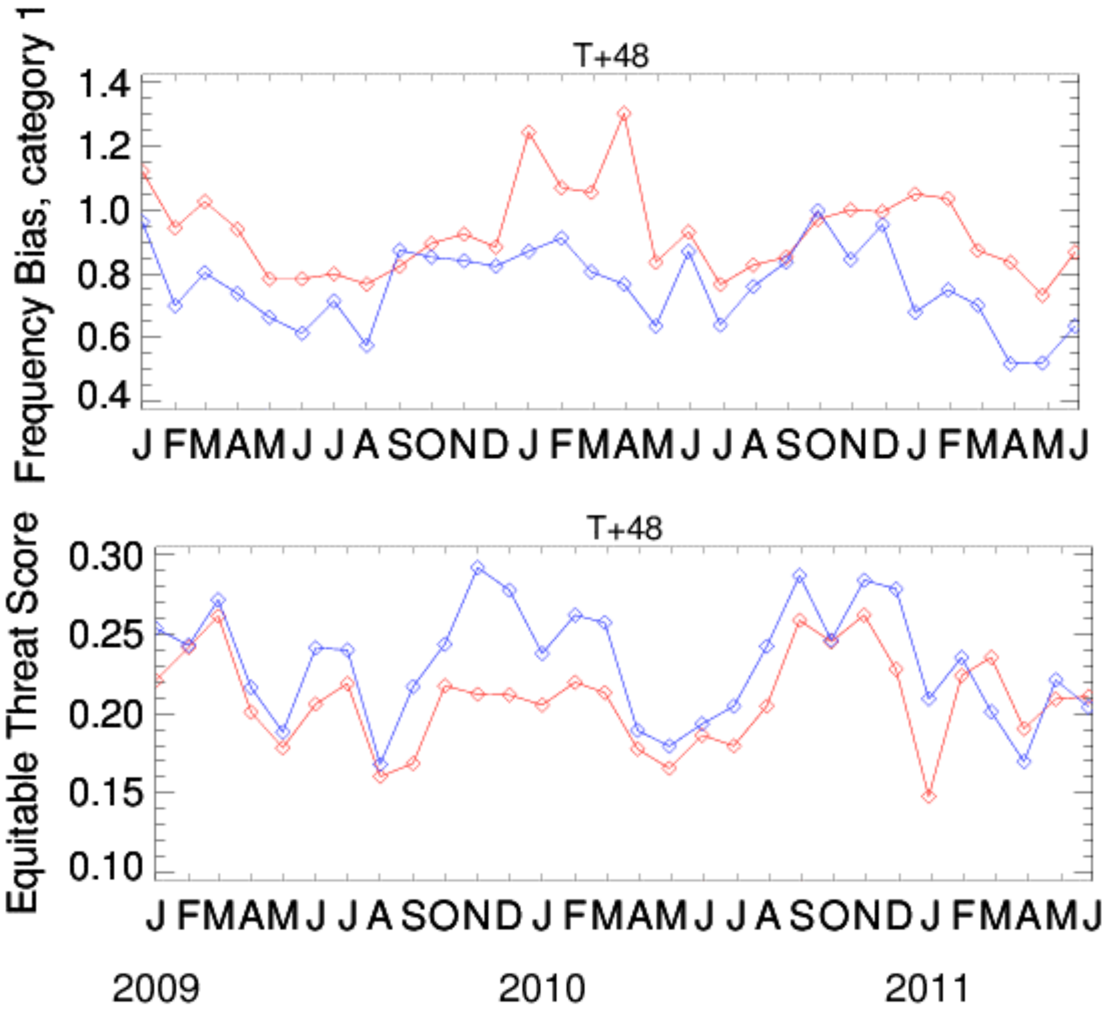


Figure 7 Comparison of a 4 mm/6h frequency bias and ETS at t+48h from the 00Z ECMWF deterministic and NAE forecasts. **Red = NAE**, **blue = ECMWF.**

ii) Verification and intercomparison of ECMWF Tropical Cyclone Forecasts *(Julian Heming)*

The Met Office has objectively verified tropical cyclone (TC) forecast tracks from its global model since 1988 using the verification scheme described on the Met Office web site at http://www.metoffice.gov.uk/weather/tropicalcyclone/method. This scheme has also been used to verify TC forecast tracks from the ECMWF model since 1994. This has enabled intercomparisons of the performance of the two models to be made since then. The first two figures show the differences between Met Office and ECMWF track forecast errors and skill scores against CLIPER. The graphs are for a global homogeneous sample of forecasts from the two models during the period 1994-2010. CLIPER is a benchmark statistical forecast based on a combination of climatology and persistence. In both graphs, positive values indicate the Met Office forecasts were better than ECMWF and vice versa for negative values.

Previous reports have shown how the ECMWF track errors have reduced markedly and skill scores increased markedly relative to the Met Office global model at short lead times since 2004 and at longer lead times since 2007. This was due to large reductions in ECMWF errors rather than increase in Met Office global model errors. The results in the last two years indicate a stabilisation in the gap in performance of the two models. For track errors (Fig. 8) the gap in performance actually closed significantly at long lead times (96 to 144 hours) in the last year. At short lead times the gap in 2010 was a little wider than in 2009, but similar to that seen in 2008. This is also reflected in the skill scores (Fig. 9) which show a gap larger than in 2009, but smaller than in 2008.

Fig. 10 shows the difference in analysis positional errors between the ECMWF and Met Office global models. Since the Met Office uses a TC initialisation procedure its analysis errors have always been smaller than ECMWF. The gap in analysis errors had been closing for many years such that by 2008 ECMWF analysis errors were only 18% larger than the Met Office. However, the gap suddenly widened again in 2009 with ECMWF analysis errors being more than twice as large as the Met Office. This large gap was maintained in 2010 and may explain why the gap in forecast errors has stabilised or even reduced since 2009. The mean analysis positional error in the Met Office global model in 2010 was in fact the smallest it has ever been since this intercomparison started in 1994.

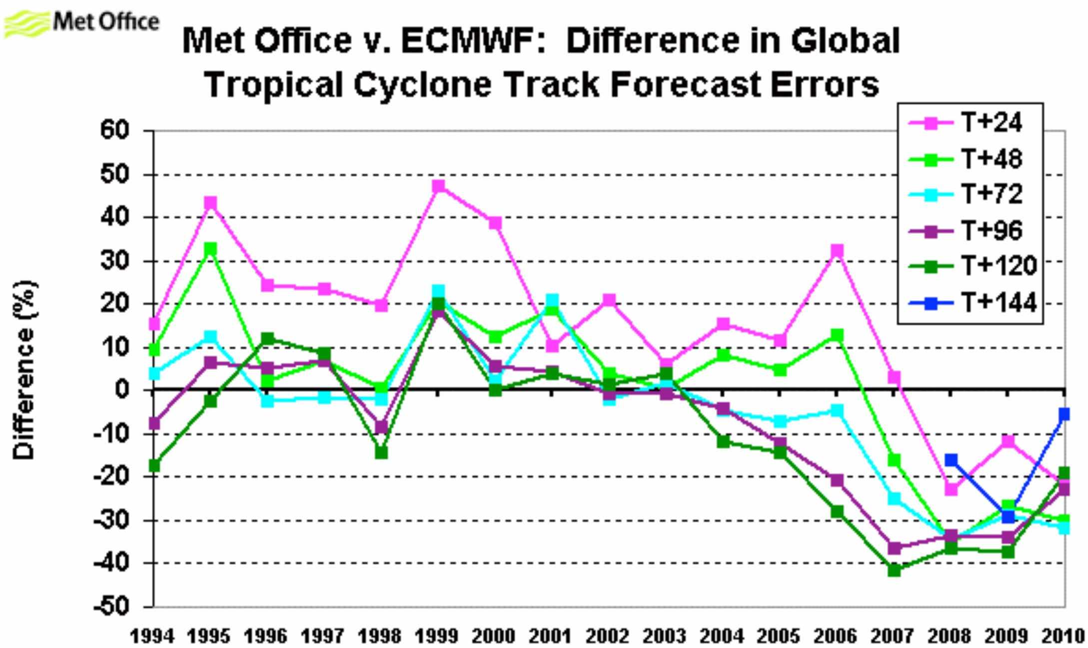


Figure 8 Comparison of TC track errors between the ECMWF and Met Office models.

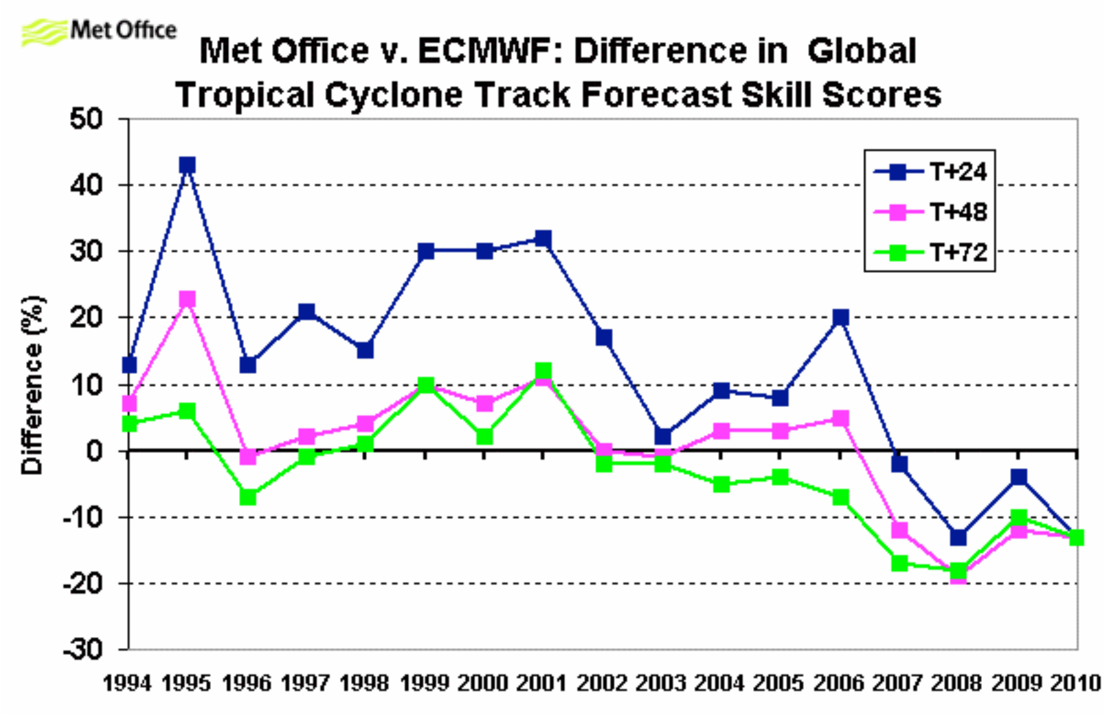


Figure 9 Comparison of TC track forecast skill scores between ECMWF and Met Office models.

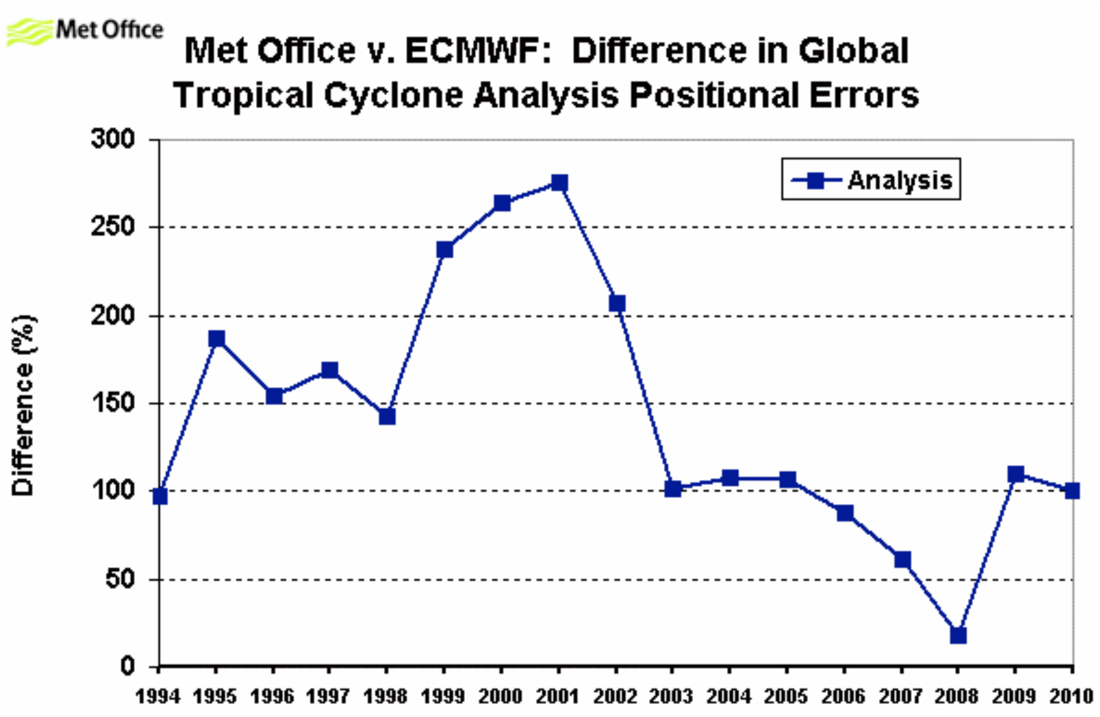


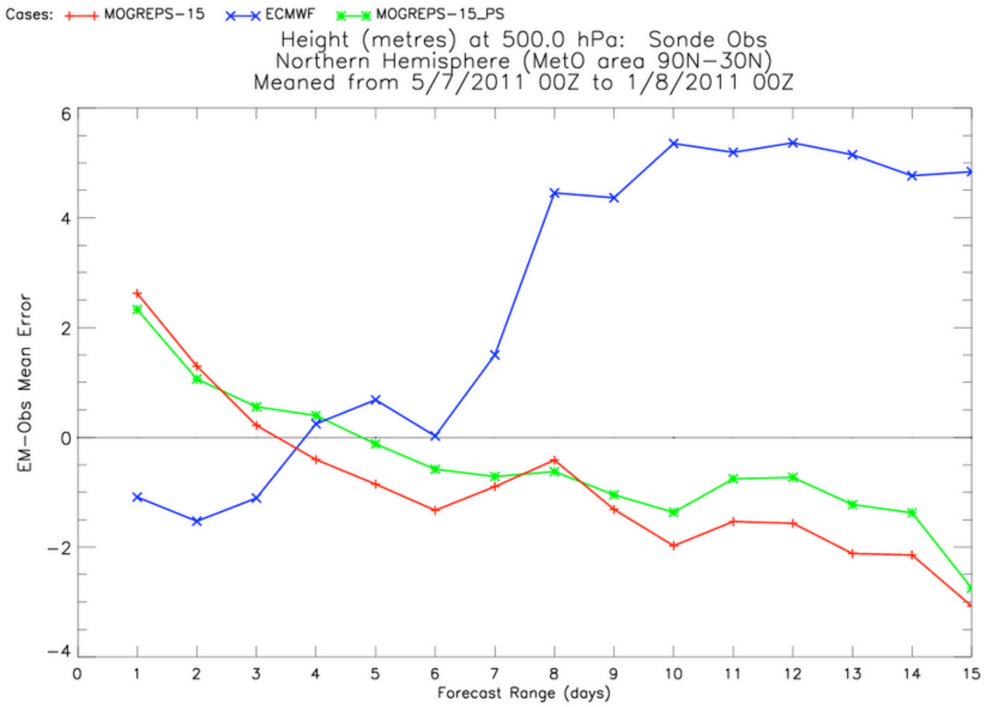
Figure 10 Comparison of TC analysis position errors between ECMWF and Met Office models.

iii) Comparison of MOGREPS-15 with ECMWF *(Simon Thompson)*

On 20 July 2011, the Met Office 15-day 24-member ensemble system, MOGREPS-15, was upgraded following the implementation of a parallel suite (PS27). The upgrade included several improvements to the forecast such as the inclusion, for the first time, of perturbations to sea surface temperatures. The main source of improvement to verification statistics however, was the inception of the hybrid DA system whereby ensemble error structures were blended with covariances derived from climatological-averaged statistics. This results in hybrid static/ensemble covariances that are more sensitive to the meteorological situation.

Figs. 11 and 12 show the comparison of the statistics derived from MOGREPS-15 during the parallel suite, with the red lines denoting the existing system, and the green lines the upgraded system. We have also included statistics from the ECMWF EPS system shown in blue.

(a)



(b)

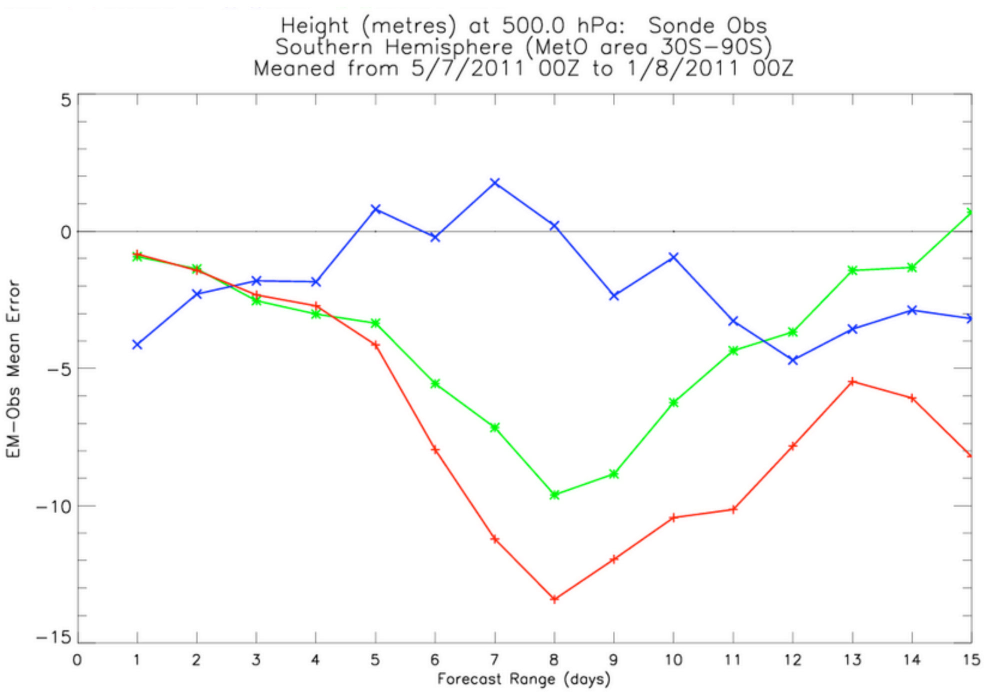
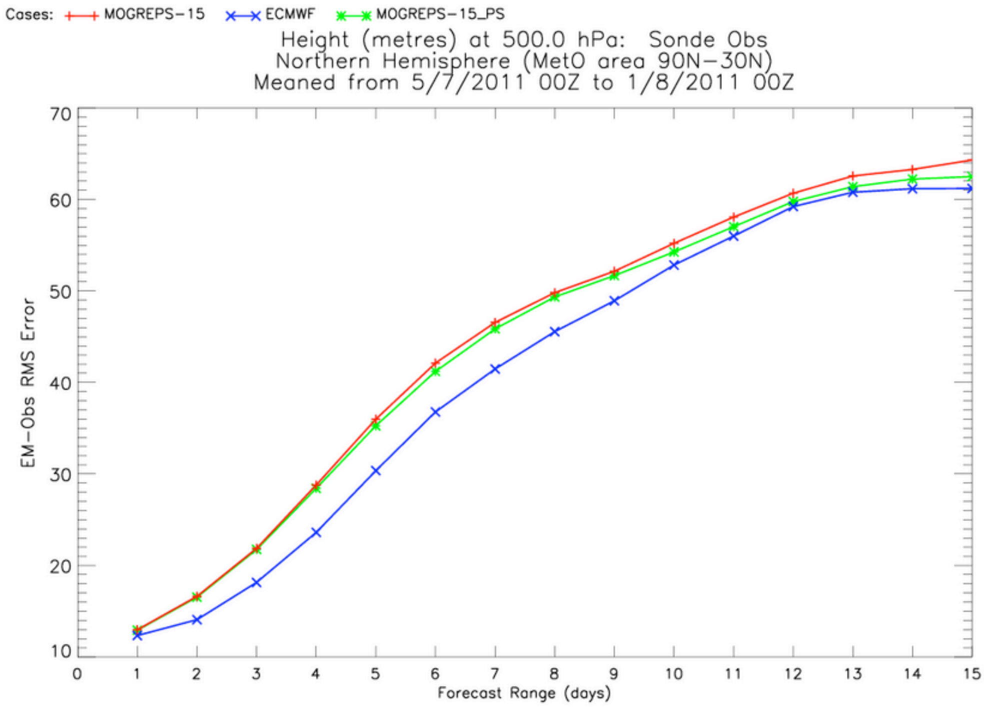


Figure 11 Comparison of 500 hPa GPH mean error for PS27 MOGREPS-15: (a) NH, (b) SH.

The parameter shown is height at 500hPa in the northern and southern hemipheres, and statistics are calculated versus sonde observations. Both comparisons for the bias (Fig. 12) and RMS error (Fig. 13) in the ensemble mean are presented. Improvements following the upgrade to MOGREPS-15 can be seen at virtually all forecast ranges for both the bias and RMSE, although the ECMWF EPS forecasts maintain an advantage of approximately 1 day of lead-time when considering the RMSE in the forecast range 6-10 days.

(a)



(b)

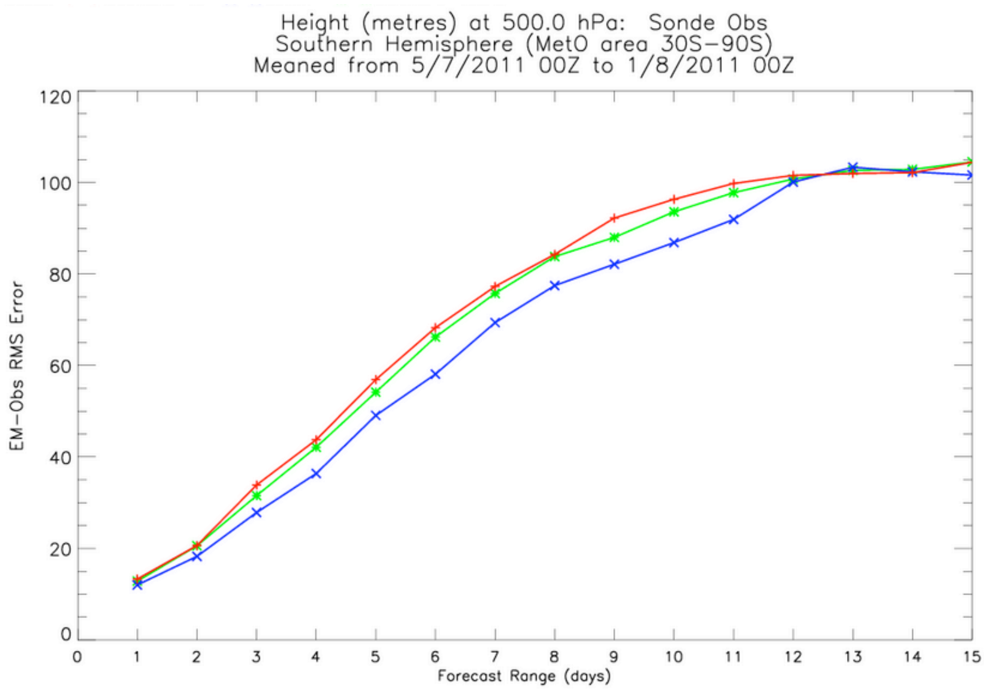


Figure 12 Comparison of 500 hPa GPH rmse for PS27 MOGREPS-15: (a) NH and (b) SH.

iv) Use of ECMWF analyses for verification of data assimilation trials *(Peter Jermey)*

Verification against (own) analyses provides a useful measure of forecast skill, but the method must be used with caution when investigating the impact of Data Assimilation (DA) changes which alter the analysis since the 'truth' is dependent on the change. For this reason, verification against (own) analyses can misrepresent the impact of DA changes on the forecast. This may include large differences with verification against observations. Such differences were seen when verifying the inclusion of hybrid ensemble-climatological background covariance in the Met Office DA scheme (see Fig. 13 A and B, which show estimated impacts on skill scores when measuring against observations and (own) analyses, respectively). In this case, ECMWF analyses provided a trustworthy verification 'truth' which was independent of the DA change and led to skill scores which largely agreed with those obtained when measuring against observations (see Fig. 13 A and C, which show estimated impacts on skill scores when measuring against observations and ECMWF analyses, respectively). It is anticipated that verification against ECMWF analyses will become a routine measure of the impact of DA changes at the Met Office.

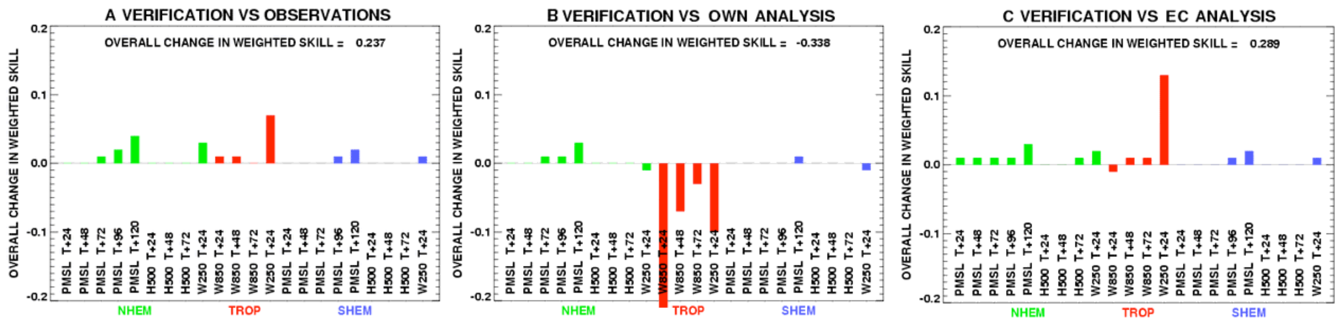


Figure 13 Impact of verification truth on DA trials.

3.1.4 End products delivered to users

None.

3.2 Subjective verification

3.2.1 Subjective scores (including evaluation of confidence indices when available)

None.

3.2.2 Synoptic studies

Some comments on subjective verification are included in the forecaster feedback listed in Section 2.2.

4. References to relevant publications

None.