Application and verification of ECMWF products 2013

Met Office response; compiled by Marion Mittermaier, Steve Manktelow & Mark Hodkinson

1. Summary of major highlights

The use of ECMWF forecast products continues to increase across the Met Office: in operational forecasting, weather and applied science (research), climate and other applications and work is ongoing to encourage further use and applications of ECMWF forecast products at the Met Office.

Feedback was received from many parts of the organisation: Global Guidance Unit, Flood Forecasting Centre, the climate applications group and model development groups for short and medium-range NWP.

2. Use and application of products

2.1 Post-processing of model output

2.1.1 Statistical adaptation
The forecasts used in 2.1 are subject to Kalman Filtering in the blending process.

2.1.2 Physical adaptation
None.

2.1.3 Derived fields
We have the capability for processing ECMWF ensemble members in clustering algorithms.

2.2 Use of products

Ensemble forecasting & 'Best Data'
The Met Office in-house system PREVIN (Predictability Visualisation) Ensemble Web Pages continues to receive ECMWF data via DART.

The Met Office has set up a feed of ECMWF data to produce what it calls 'Best Data' for 5000 cities. The data being used includes: precipitation amount, screen temperature, 10m wind-speed and direction, wind-gust, dew point temperature, surface temperature, snowfall, downwards solar radiation, total cloud amount, MSLP, min/max temperature.

The Best Data group receives and is looking to make use of other parameters such as convective/dynamic rain amounts, 100m U and V or model level data and wind speed forecasts at height. It is developing a gridded best data prototype using the following ECMWF data out to 7 days: 10m wind speed and direction, PMSL, maximum 10m gusts, total precipitation, total snowfall, 2m temperature, maximum and minimum temperatures.

In weather regime work the Met office is using ECMWF 500 hPa GPH, 10m U and V, PMSL, surface and 1.5m temperature, T850hPa out to 15 days.

Forecasters (general) & guidance unit
ECMWF products remain well-respected and are widely used along with output from other centres as a comparator against our own model output. The following ECMWF products are accessed via EC Charts: ECMWF Global GRIB - Td, gph, height, Tmax, Tmin, WBPT, MSLP, P, RH, Specific Humidity, T, T 2m, Cloud cover, total precip, total snow, Vert Velocity, wind u, wind v (T+0 - T+240). Ensemble GRIB - Sig wave height, total precip, wind gust (T+0 - T+120). EC wave data both deterministic and ensemble is also used.

EC EFI output is frequently used along with EFI EPS meteograms and the basic EPS meteograms for a wide variety of locations globally. Deterministic data from the EC website for various global regions provides useful information beyond T+144. EC tropical cyclone pages and database and 'Dalmation plots' widely used for the guidance assessments.

EC Meteograms (both 10 and 15 day) containing max/min temperature, 10m wind speed and total precipitation (mm/6hr) continues to be used for a small number of Open Road customers.

The Guidance Unit look at a number of ECMWF products via the main website but also EC Charts (available to member states). There is also access to post-processed info via Met Office systems such as PREVIN. EC gridded data is also available and well used on HORACE - Visual Weather tool. All of these are visualised but some are included in guidance. The actual data does not drive any products directly in the Guidance Unit.

Flood Forecasting/extreme weather
The EC Charts application is used extensively in the Flood Forecasting Centre (FFC) in the Met Office where ECMWF data in visual weather used as comparison against Met Office UM data. Precipitation and wind fields are of particular interest. This year the FFC has also started to use EFAS products

The ECMWF EFI output is also used and accessed both from the ECMWF website and EC Charts. ECMWF ensemble precipitation forecasts, wave data (deterministic and ensemble are used. ECMWF ensemble postage stamp forecasts are viewed from PREVIN pages and longer range anomaly maps. EPS meteograms are used when required alongside the EFI output to provide additional confirmation of severe weather.

Climate applications
The Met Office Climate Applications team are using products from the following ECMWF systems: VarEPS/monthly system, the seasonal system and EUROSIP. From these monthly rainfall and temperature products for East, West and southern Africa are made available to African collaborators. ECMWF and EUROSIP visualised seasonal forecast products are also used to help generate consensus seasonal outlooks at Regional Climate Outlook Forums in East, West and southern Africa. The main products used are probabilistic rainfall products and ENSO plumes. Digital forecast and hindcast data from the ECMWF system is also used in capacity training and forecast generation.

Monthly and seasonal forecasting: Public Weather Service (PWS)
ECMWF data is used to generate products (e.g. the monthly outlook) for the UK and other regions globally with predictions made weekly. Parameters used: Temp mean, Temp max, Temp min, precipitation, windspeed (10m), PMSL and Z500.

The EC Var EPS/monthly system is utilised for monthly forecasts for Africa (International Aid Programme) Seasonal, 3 month outlook uses data for the UK and other regions, including Africa etc. products (related to tropical storm activity over the next 6 months) used in multi-model combination (with our model), for various products, incl re-insurance industry and for PWS. The EC seasonal pages is also used to compare with other Met Office outputs.

Future requests for new and enhancement of existing products and data
Met Office users of EC data and services have the listed the following items which they believe would add value to the data and products they currently receive from the Centre:

- Severe weather products to help provide longer lead times (3-5 days)
- Access to higher resolution EC Website data
- Development of a ‘Regime change’ tool: **eg Change of weather type, end of block (winter) freezing conditions or Summer Heat wave;** 6-15 day lead times.
- EC Charts - Thete-w, CAPE, Omega equation components
- More EC Fields on SWIFT/VISUAL weather
- Training for forecasters on how EC products can be used to assist their work
- Extend windstorm concept to cover rainstorms and snowstorms
- Postage stamps type displays for other parts of the world

3. Verification of products

3.1 Objective verification

3.1.1 Direct ECMWF model output (both deterministic and EPS)
See sections below.

3.1.2 ECMWF model output compared to other NWP models

<u>a. Tropical Cyclone Tracking</u> *(Helen Titley, Piers Buchanan and Julian Heming)*

The Met Office has an in-house real-time capability for tracking tropical cyclones and potential storms using MOGREPS-15, EC-EPS and NCEP GEFS data and their 4 different multi-model combinations (M-E-N, M-E, M-N and E-N). Tracks from the individual ensemble members are produced and plotted. An ensemble mean track was calculated from each model and multi-model combination for every named storm and verified against each storm's observed position for the period January to December 2012. Figure 1 shows the mean track error from each model combination. The full multi-model combination of EC-EPS, NCEP GEFS and MOGREPS-15 has the lowest track errors from T+120, while at short lead times it is slightly bettered by the mean track from the EC-EPS and NCEP GEFS combination, although the 3 model combination is close behind in 2nd place. The single model with the lowest track error is the EC-EPS, while MOGREPS-15 has the highest track error. However, interestingly MOGREPS-15 still brings a significant benefit in a multi-model combination. By T+168 the ECMWF-MOGREPS-15 multi-model combination gives around a 12-hour benefit over just using ECMWF and a 24-hour benefit over just using MOGREPS-15.

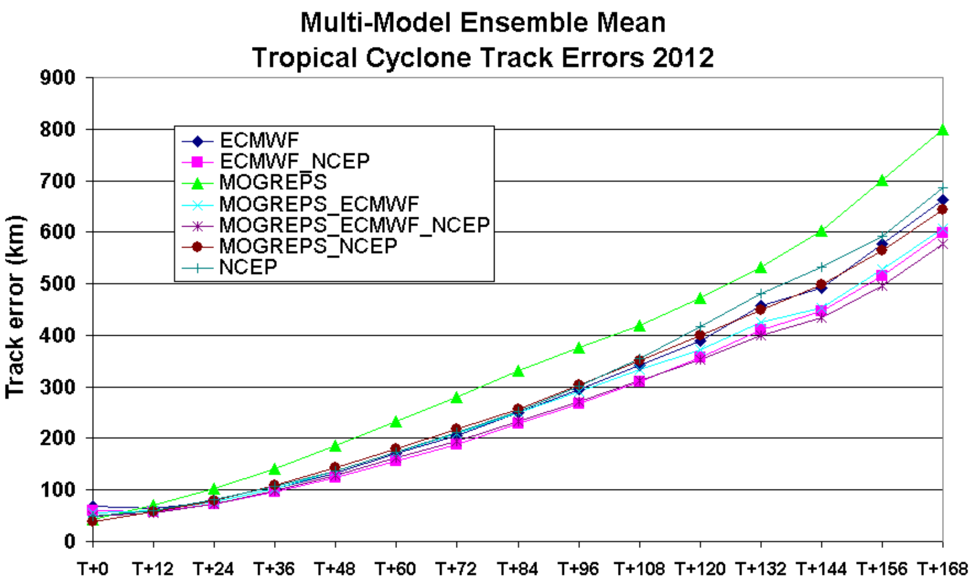

Figure 1 Track errors in km from the mean track from the MOGREPS-15, EC-EPS and NCEP GEFS ensembles, together with the mean track from the various multi-model combinations. The verification period is from January-December 2012.

3.1.3 Post-processed products

<u>a. Verification of probabilistic severe weather forecasts</u> *(Rob Neal)*

The Met Office Ensemble Prediction System First Guess Warnings tool (EPS-W) is an ensemble-based first-guess support tool for severe weather. The system post-processes ensemble data into a format which mimics the National Severe Weather Warning Service (NSWWS) colour states. Ensemble members from both MOGREPS and ECMWF are used providing guidance from just a few hours ahead up to six days. Verification results presented here are for ECMWF only.

For verification purposes, UK county probabilities are derived, whereby gridded probability fields are overlaid by a UK county map. The 95th percentile grid point probability in each county is then taken as the area probability. County probabilities are verified against a Met Office 2 km resolution analysis. An event is considered to occur in a county if ≥5% of the analysis grid-points in the area exceed the event threshold.

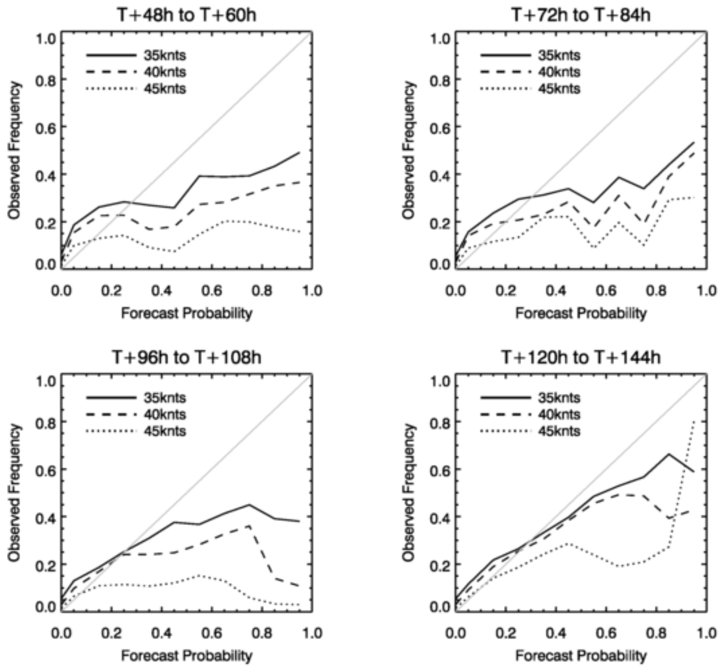

Figure 2 10m wind gust reliability diagrams. Verification period: 01/03/2012 to 31/07/2013.
The wind gust reliability diagrams (Figure 2) predominantly show over-forecasting for all wind gust thresholds and forecast lead times. However, over-forecasting does reduce with forecast lead time. The lower wind gust thresholds have the best reliability.

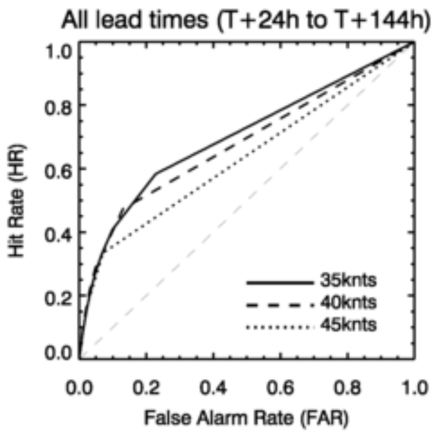

Figure 3 10m wind gust ROC curves. Verification period: 01/03/2012 to 31/07/2013.

The wind gust ROC curves (Figure 3) show that the lower thresholds have better ROC areas. Forecasts for these lower thresholds are therefore better at discriminating between events and non-events.

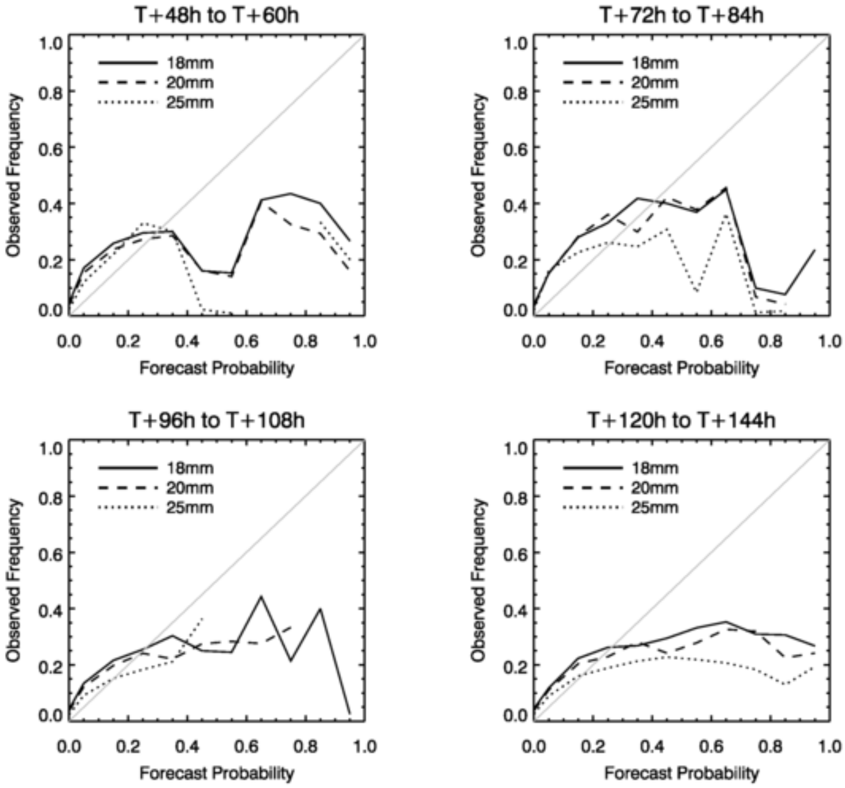

Figure 4 24hr precipitation reliability diagrams. Verification period: 01/03/2012 to 31/07/2013.

The 24hr precipitation reliability diagrams (Figure 4) predominantly show over-forecasting for all precipitation thresholds and lead times. However, unlike with wind gusts, reliability does not improve with forecast lead time. The 24hr precipitation reliability diagrams are also noisy (particularly in the earlier forecast lead times), suggesting that even the 17 months worth of data used here may not be enough when verifying rare events.

The 24hr precipitation ROC curves (Figure 5) show a similar pattern to those shown for wind gusts (Figure 2), whereby the lower thresholds have better ROC areas.

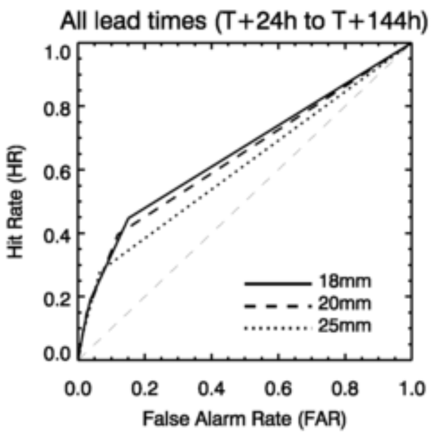

Figure 5 24hr precipitation ROC curves. Verification period: 01/03/2012 to 31/07/2013.

b. Initial verification of site-specific probabilistic forecasts with focus on week 2 *(Rachel North and Marion Mittermaier)*

Currently there are two probabilistic data streams which cover the week 2 forecast period: ECMWF and MOGREPS-15. In terms of resolution these models are vastly different up to day 10 with ECMWF operational ensembles now running at T639 (31 km), and T319 (60 km) beyond. MOGREPS-15 is still at ~60 km resolution at UK latitudes. However for the main period of interest, the resolutions are similar. It could be expected that forecast evolution errors will dominate the total error in week 2 and therefore even probabilistic site-specific forecasts are not expected to be very accurate, unless the weather regimes remain fairly constant with little variation from day to day.

At present only three main parameters are being considered: temperature, wind speed and precipitation. These are assessed using the Continuous Ranked Probability Score (CRPS), Ranked Probability Score (RPS) and Brier Score (BS) respectively. All these scores are focusing on accuracy, not skill. A decision on what to use as the reference for a skill score has yet to be made. First we want to see the value added by post-processing, in this case the application of the Kalman Filter (KF). The second comparison is to check how the KF has improved the relative performance of MOGREPS-15 and EC. In essence, this is the fair comparison between different models, where post-processing mitigates against the inequalities in the native grid resolution of EC and MOGREPS-15. Unfortunately at present only the raw comparison can be considered as post-processed site-specific forecasts are not available for MOGREPS-15. The final comparison will come about when the MOGREPS-15 KF and EC KF forecasts are blended together, to see what benefit there is in blending.

These initial results are based on 4-months of data, from October 2012 to January 2013 (the dates which are available in the archive). UK station data have been processed to calculate CRPS, RPS and the Brier Score for screen temperature, 10m wind speed and precipitation amount respectively.

Figure 6 shows the CRPS for screen temperature with forecast lead time for the raw ECMWF forecasts (EC), the raw MOGREPS-15 (M15) forecasts and the Kalman-filtered ECMWF forecasts (KF-EC). Only the 0Z forecasts have been analysed, in order to highlight any diurnal effects. Data plotted are valid at 0Z (night) and 12Z (day).

The day/night variation is apparent, with the day-time scores better than the night-time scores (as expected). In the earlier lead times (days 1-5) in autumn there is a much slower error growth than at night. The behaviour in winter is a more constant growth across the range of lead times.

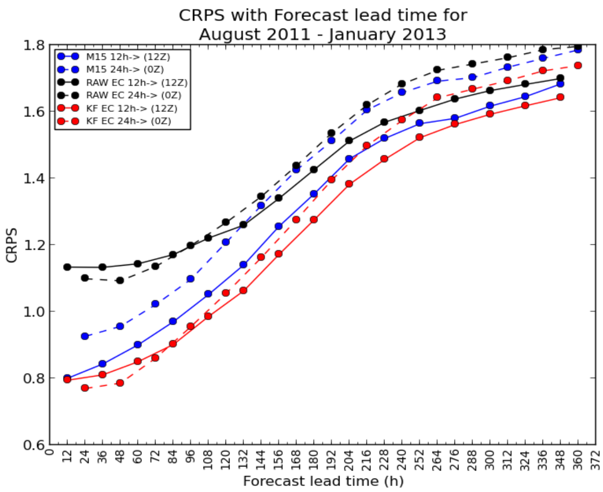

Figure 6 Continuous Ranked Probability Score (CRPS) for screen temperature calculated from the 3 components. The night (dashed lines) and day (solid lines) signals have been plotted for each data set.

The daytime errors from each model merge together after day 10 for the autumn data, and to a lesser extent the winter data. The day-night contrast is greater in autumn than winter at the longer forecast ranges. The night-time error growth appears to flatten off more quickly in the autumn data (at 9 days as opposed to 11.5).

The more interesting aspects happen to be at the shorter forecast ranges. Days 1 to 2 in both the night-time and day-time CRPS scores show that the ECMWF model verifies worse than MOGREPS-15. However, applying the Kalman Filter improves the scores by some margin, especially for the night-time temperatures. It will be interesting to see whether this behaviour remains with a longer time series. The raw MOGREPS-15 model outperforms the raw ECMWF model for the daytime temperatures at those earlier forecast ranges. However, the error growth in the ECMWF raw model is slower than that in MOGREPS-15 over days 1 to 3, both at night and in the day.

There does not appear to be any obvious sign of the resolution change in the ECMWF raw model in the autumn data, but there may be in the winter data. For winter, there is a marked increase in the CRPS between day 11.5 and day 12.5 (day), and day 12 and day 13 (night). This is mirrored in the Kalman-Filtered data.

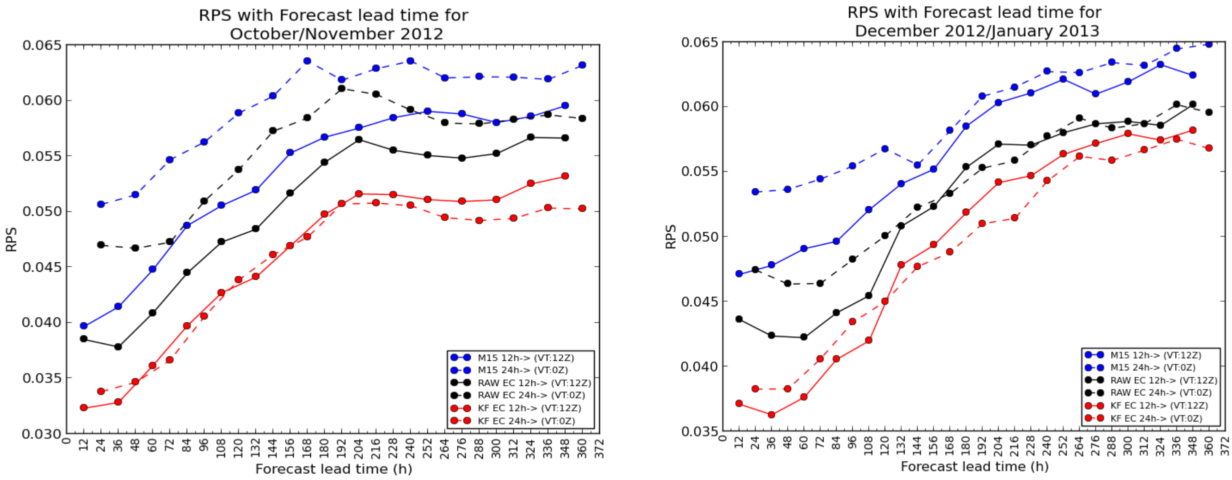

Figure 7 Ranked Probability Score for wind speed calculated from the 3 components. The night (dashed lines) and day (solid lines) signals have been plotted for each data set.

Figure 7 shows the Ranked Probability Score for 10-m wind speed with forecast lead time for the raw ECMWF forecasts (EC), the raw MOGREPS-15 (M15) forecasts and the Kalman-Filtered ECMWF forecasts (KF-EC).

There is a notable flattening of the error growth beyond day 8 in the autumn data, for both daytime and nighttime wind speed. What is interesting is that there is a score difference between the raw model day and night wind speeds (day better than night), but the Kalman-Filtered ECMWF wind speeds produce scores which are similar between day and night, in both the autumn and winter data.

MOGREPS-15 raw model scores worse for wind speed than both ECMWF raw model and Kalman-Filtered ECMWF; at the early forecast ranges there is a large performance difference, but beyond day 8 the difference gets slightly smaller (winter).

Figure 8 shows the Brier Score for the 6h precipitation accumulations for all the components under consideration.

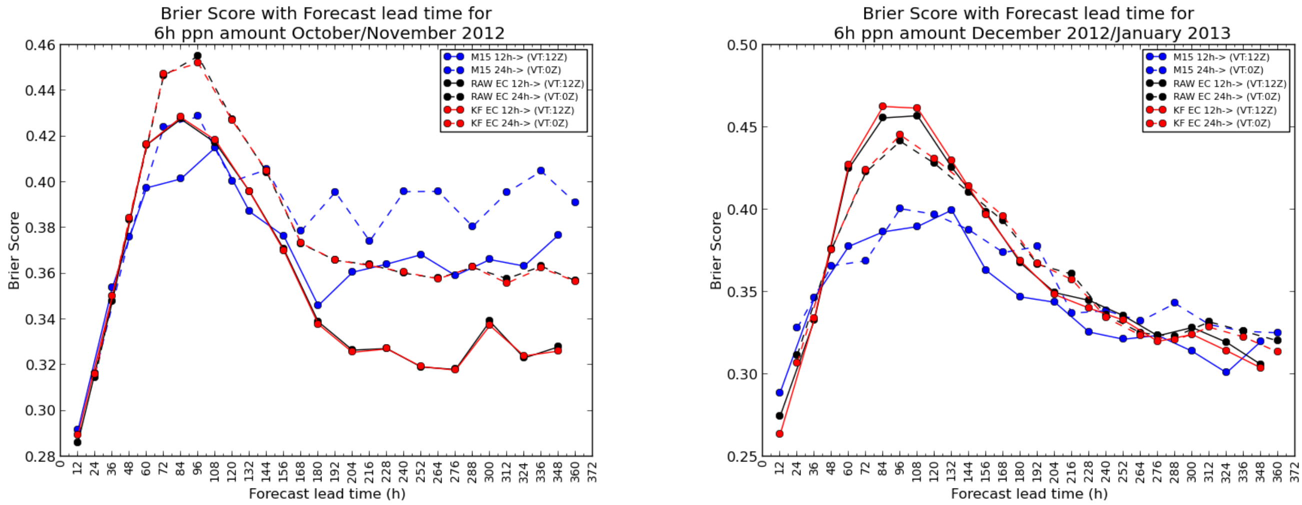

Figure 8 Brier Scores for 6h precipitation accumulations calculated from the 3 components. The night (dashed lines) and day (solid lines) signals have been plotted for each data set.

The main point to note is that post-processing does not seem to help with precipitation. There is no discernible difference between the raw ECMWF scores and the Kalman-Filtered ECMWF scores. This is the case for both the 24-hour totals and the 6-hour totals.

Interestingly, at days 8 to 15, the winter scores appear to cluster together, yet the autumn scores don't. The scores tend to a limiting value at the longer forecast ranges though, in both cases; the error growth is limited beyond day 8.

The 6-hour precipitation error appears to peak at day 4, and decrease from then onwards. This could purely be a signal that due to the nature of the model bias, at the longer forecast ranges the model is just plastering precipitation over the domain, and that in this time period the weather has helped the score by being generally wet. Whether this feature remains in the longer timeseries will help to answer that question. However, MOGREPS-15 has a lower peak score, than the raw ECMWF data.

Most obvious is the steep error growth over days 1 to 3.5. This is to be expected as the models exhibit 'spin-up' problems at early forecast ranges.

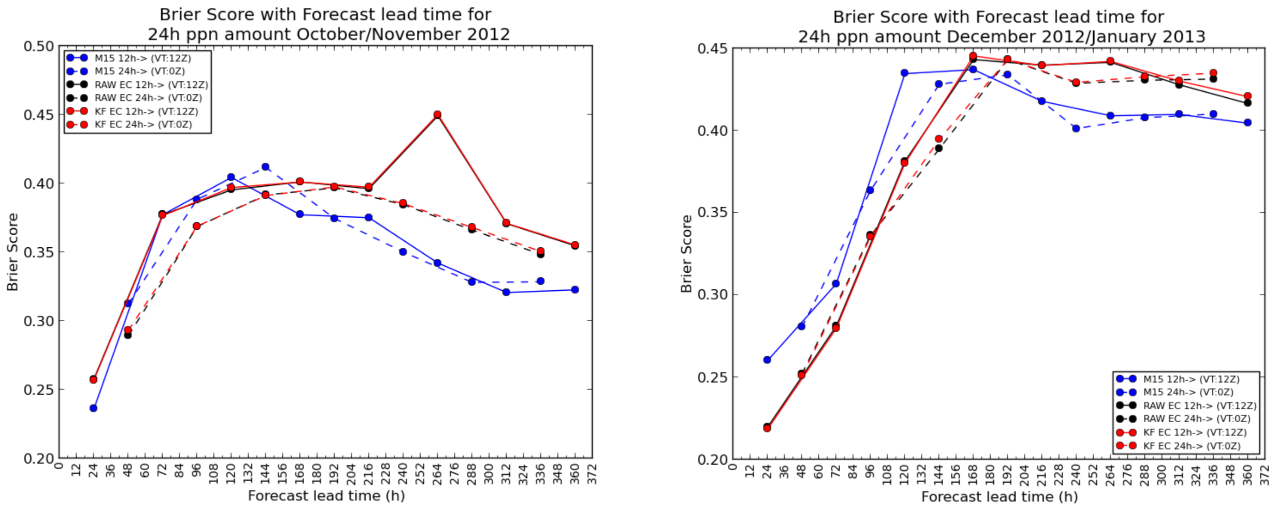

Figure 9 Brier Score for 24hr precipitation amount calculated from the 3 components. The night (dashed lines) and day (solid lines) signals have been plotted for each data set.

The 24-hour precipitation totals have a behaviour more consistent with the other scores for the other parameters. At early lead times for the error growth is steeper than longer lead times. After day 7 the errors start to diminish, the behaviour is the same for both time periods and day and night temperature scores.

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

a. Case studies *(David Walters)*

When developing upgrades to global configurations of the Met Office Unified Model, model developers routinely run sets of standard case studies initialised from Met Office operational deterministic global analyses. The impact of each change is then assessed via objective verification against observations and operational analyses as well as subjective assessment of various diagnostic fields and metrics. In the past, these case studies have usually been initialised from Met Office operational analyses. When we believed that a change was likely to have a significant impact on the analysis cycle itself, such that the initial and verifying analyses are likely to be more relevant to the control than the test, we replaced these Met Office analyses with deterministic analyses from the ECMWF IFS.

Our current upgrades are based around a new dynamical core (Wood et al., submitted), which we know changes some properties of the model analysis. For this reason, we have decided to use ECMWF analyses as standard in our case study suite. To ensure that the surface fluxes are consistent with those from a data assimilation, we use ECMWF atmospheric analyses, but operational Met Office analyses for surface parameters such as soil moisture content and snow amount. This use of ECMWF T+0 as an independent analysis has allowed an objective assessment of model upgrades, prior to testing in full data assimilation trials.

b. Evaluating storm tracks using ECMWF analyses *(Claudio Sanchez)*

There are clear improvements made by the new dynamical core EndGAME on the simulation of mid-latitude storms. Figure 10 shows the distance and intensity errors of MetUM extra-tropical cyclones against EC analysis for various MetUM resolutions and configurations. Cyclone tracks and intensity are computed using TRACK algorithm (Hoskins and Hodges, 2002) and then simulated storms are paired to those in the analysis if these satisfy certain conditions (Froude *et al*, 2007). EndGAME produces stronger cyclones helping to reduce a long withstanding issue in the MetUM and many other models, the excessive diffusivity of mid latitude variability caused amongst other things by the implicit diffusivity of the numerical core (i.e interpolation to the departure point). Distance errors are also slightly reduced by EndGAME (figure 10a and 10b) and when horizontal resolution increases.

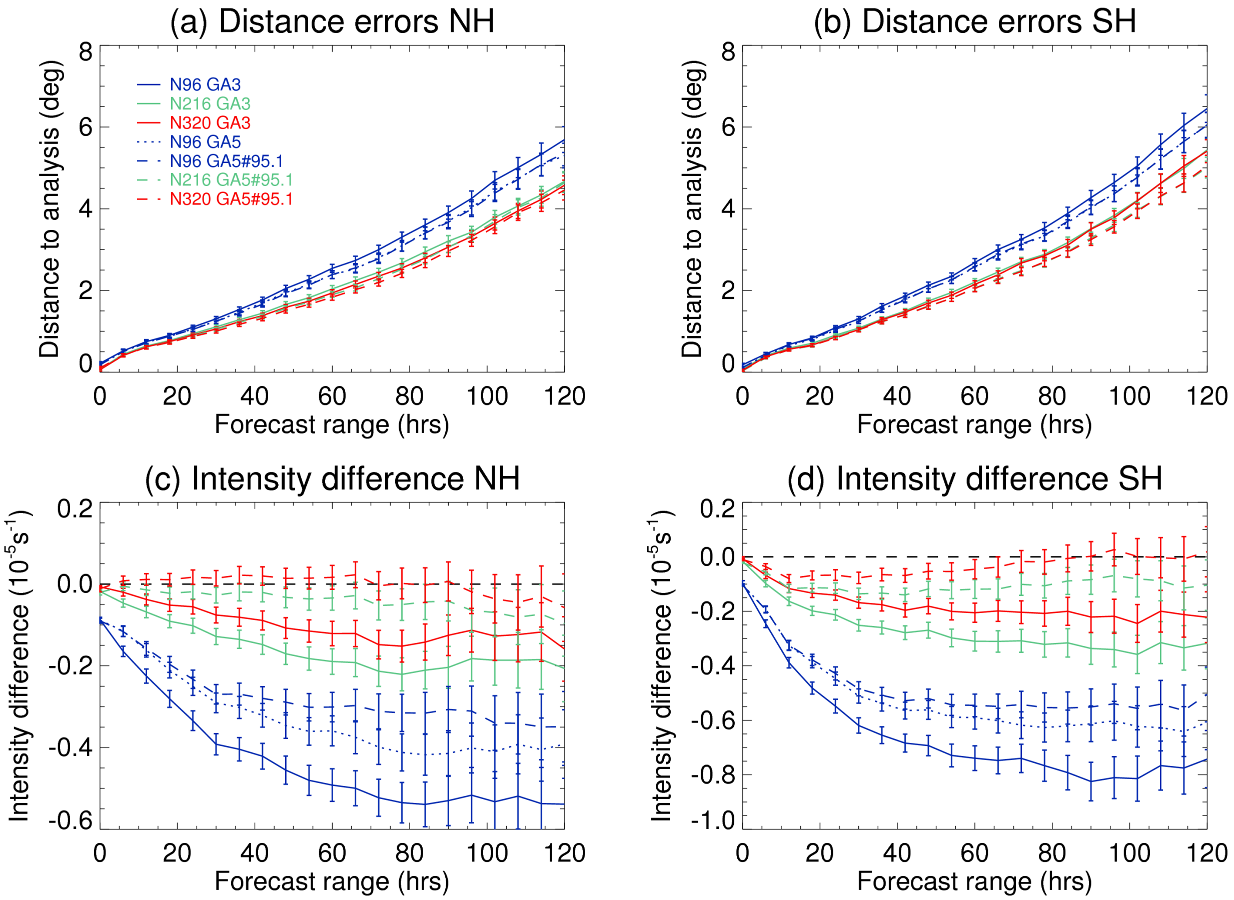

Figure 10 Assessment of the mid-latitude storm track in the Northern and Southern Hemispheres at different resolutions against the ECMWF analyses.

4. References to relevant publications

Froude, L. S. R., L. Bengtsson, and K. I. Hodges, 2007: The prediction of extratropical storm tracks by the ECMWF and NCEP Ensemble Prediction Systems. Mon. Wea. Rev., 135, 2545–2567.

Hoskins, B. J., and K. I. Hodges, 2002: New perspectives on the Northern Hemisphere winter storm tracks. J. Atmos. Sci., 59, 1041–1061.

Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., Melvin, T., Smith, C., Vosper, S., Zerroukat, M., and Thuburn, J.: An inherently mass-conserving semi-implicit semi-Lagrangian discretisation of the deep-atmosphere global nonhydrostatic equations, *Q. J. R. Meteorol. Soc.*, subm