Application and verification of ECMWF products 2018

Met Office. Contributing authors: Dan Suri, Marion Mittermaier, Rob Neal, Ric Crocker, Helen Titley and Becky Stretton.

1. Summary of major highlights

ECMWF output is widely used by the Met Office at many timescales, especially the week 2 period where the Met Office is particularly dependant on ECMWF output. The heaviest users of ECMWF output are the Chief and Deputy Chief Meteorologists, who provide forecast guidance across the organisation for areas around the world and hydrometeorologists at the Flood Forecasting Centre. ECMWF output also forms part of 'BestData', a multi-model blend producing site-specific forecasts available, for example, on the Met Office public website. Meanwhile, the Met Office have been actively engaged in producing some novel and innovative tools to help operational meteorologists visualize ensemble output, creating and verifying regional weather regime forecasts via a tool known internally as the Decider system and developing a global hazard map, which includes a focus on tropical cyclones.

2. Use and application of products

2.1 Post-processing of ECMWF model output

2.1.1 Statistical adaptation

Nothing to report.

2.1.2 Physical adaptation

Nothing to report.

2.1.3 Derived fields

The end of February and first days of March 2018 saw the most severe outbreak of cold winter weather in the UK for several years, as well as over much of western Europe, compounded by Storm Emma which came from the south and generated exceptional blizzards and freezing rain in the south of the country. The cold outbreak had been well predicted in advance, firstly from long-range predictions of a Sudden Stratospheric Warming increasing the risk of blocking, and then two weeks ahead from the Decider system. Decider clusters ECMWF and other global ensembles according to weather patterns (Neal et. al., 2016), and identified very high probabilities of cold easterly flows up to two weeks ahead (Fig 1). In fact the real cold air came a few days later than first anticipated but by the end of the week before it was very clear that a major winter outbreak was coming. By a 7 day lead time the Met Office Global Hazard Map (Fig 2; Robbins and Titley, 2018), which post-processes ECMWF data, had very high probabilities of cold wave conditions across large parts of Europe. Forecasting throughout the week exploited a combination of ECMWF and Met Office models, and a few others including NCEP, to provide multi-model ensemble guidance at all time ranges.

ECMWF 00Z run Tue 13 Feb 2018 (30 regimes) – Updated 09:03:25 GMT, Tue 13 Feb 2018 [Latest | Run-1 | Run-2 | Run-3 | Run-4 | Run-5 | Run-6 | Run-7 | Run-8 | Run-9 | Run-10]

INTERACTIVE TABLE: Probability of each regime occurring at each lead time (30 regimes)



Fig.1 Probabilistic weather pattern forecast from the 0000Z ECMWF medium range run on 13th February 2018, showing high confidence forecast transition to weather pattern in week 2 (left image), which can be described as an anticyclonic easterly pattern with colder-than-average temperatures (top right image). Forecast confidence was unusually high for week 2 (bottom right image).



Fig.2 Probability of cold wave conditions at a 7 day lead time as shown by the Met Office Global Hazard Map based on the ECMWF 0000Z medium range run on 20th February 2018. Probabilities show cold wave conditions widespread over much of Europe.

2.2.1 Use of Products

At the Met Office, ECMWF output is used alongside the Met Office's own model suite. This suite includes an hourly-cycling, convection-permitting model (UKV), a 12 member convection-permitting ensemble (MOGREPSUK), the global model, run 4 times a day with horizontal resolution of approximately 10 km and an 18 member ensemble, MOGREPS-G.

Visualization of ECMWF output by Met Office operational meteorologists is via a number of mechanisms, the primary vehicles being ecCharts and ingestion of a subset of ECMWF HRES into the Visual Weather forecaster workstations. Visualization is supplemented by use of ECMWF's main products web page and some internal webbased tools and diagnostics, for example classifying ensemble members by synoptic type to objectively assess signals (this product is known as Decider), particularly for severe weather (Neal et. al. 2016) and for coastal flood risk (Price 2017), driving a number of downstream forecasting aids for fluvial, pluvial and tidal flooding and use of multi-model ensembles in, for example, tropical cyclone forecasting (Titley and Neal 2017). Meanwhile, ECMWF model data is blended with Met Office NWP to produce what is known in-house as 'BestData', a database of site-specific forecasts for thousands of sites used for many applications including, for example, forecasts available on the Met Office public web site. In BestData, ECMWF data is the primary source of data for forecasts for week 2.

Regarding Met Office forecasting operations, over-arching, authoritative guidance for the shorter term (Day 2) is provided by the Chief Operational Meteorologist, who has ultimate accountability for forecast output and is responsible for operational delivery of National Severe Weather Warning Service (NSWWS) impacts-based warnings. The Chief Operational Meteorologist is supported, on shift, by two Deputy Chiefs. One focuses on UK weather on time scales from Day 2 onwards through medium range and into monthly and seasonal time scales. The other focuses on providing guidance on global forecast matters across multiple time-scales, operating in what is known as the Global Guidance Unit; here customers include Met Office meteorologists working away from the UK, UK government departments, to whom support is provided during periods of severe weather such as the 2017 Atlantic Hurricane season, and cross-European projects such as ARISTOTLE.

Other forecast teams, for example aviation, media and marine, then use this guidance to influence their own more customer-specific output. These forecast teams are spread across the UK with some operational meteorologists based at UK military bases, some embedded with clients, for example at some UK airports and transport agencies. some at the main Operations Centre in Exeter or in the Aberdeen office. In addition, there are a number of operational teams at locations around the world (e.g. Falkland Islands). The Flood Forecasting Centre, meanwhile, is a joint partnership with the Environment Agency set up in the wake of severe flooding over parts of the UK in 2007 and comprises of a team of dedicated hydrometeorologists providing advice on fluvial, pluvial and tidal flooding.

EFI and meteograms plotted with M-Climate are very popular with the Global Guidance Unit who require an efficient means of visualization severe weather potential around the world and associated climatic context. Cyclone database products (Hewson and Titley 2010) remain an effective means of visualizing extratropical cyclone details. Objective fronts and Dalmatian plots showing low pressure centres and intensity are two of the most popular elements used, especially in medium range forecasting. Meanwhile, diagnostics mentioned above classifying ensemble members by synoptic weather type are immensely useful in decision-making, allowing the operational meteorologist to quickly draw out the main themes and trends shown in ensemble output.

Field modification software (Carroll 1997) is used by the Chief and Deputy Chief Operational Meteorologists to produce a graphical best estimate of weather conditions over the UK, a process which draws on a number of different elements of NWP, including ECMWF. Severe weather will be accompanied by the issue of National Severe Weather Warnings Service Warnings (NSWWS). This impacts-based warnings service is aimed at both the general public and resilience/emergency services and makes use of a colour-coded 4x4 matrix to communicate a sense of both likelihood and severity of weather conditions. The range of elements for which warnings are issued has been extended in summer 2018 from the original five elements (namely wind, rain, snow, ice and fog) to also include thunderstorms and lightning. Meanwhile, the capability to issue warnings from five to seven day lead times has also been recently introduced. ECMWF data is used alongside Met Office data in the consideration of NSWWS, and a web-based tool using ensemble output, known as EPS-W (Neal et. al. 2014), can be used to generate first-guess warnings.

On example of where ECMWF output proved particularly useful was during the severe cold spell in late February and early March 2018. Use of some ECMWF output during this period has been detailed in Section 2.1.3. Meanwhile, preceding the arrival of Storm Emma into the UK on 28th February 2018 were high confidence forecasts for significant snowfall across many areas. Also, unusually for the UK, was a significant forecast signal for freezing rain to affect parts of Southern England and Wales. The freezing rain risk was highlighted well by the ECMWF precipitation type bar plots (Fig 3), which were used by Met Office Chief and Deputy Chief Operational Meteorologists, this new diagnostic lending confidence to signals for freezing rain present in Met Office NWP.





Fig.3 ECMWF 1200Z run on 27th February 2018: Probability of precipitation type for Fishguard (left) and London (right) highlighting the significant possibility of freezing rain or ice pellets on Thursday night across parts of the SW.

2.2.2 Product requests

Feedback from the Met Office forecasting community is dominated by requests for faster performance of ecCharts. There has also been positive feedback regarding the recent introduction of outputs such as precipitation type diagnostics, lightning diagnostics, vertical profiles in ecCharts and IVT outputs.

One of the main foci for product development at the Met Office using ECMWF output is the use of multi-model ensembles and tools to identify impacts and significant work is being undertaken to develop a multi-model blended probabilistic post-processing system. Any products aiding this effort would clearly be of use.

Another area where input from ECMWF is sought is in developing tools and products to monitor ongoing and developing hazards, improved or novel methods to get better statistics of extremes to contextualise weather events and further developments to reanalysis and/or reforecasts for climate context and calibration.

Moving on to some more specific products, Flood Forecasting Centre Hydrometeorologists report that a storm/tidal surge ensemble would be useful and supplement information from MOGREPS. Meanwhile, with NWP being increasingly coupled to hydrological and other downstream/impact models, there may be an increasing need to forecast actual quantities and not just anomalies or probabilities so this may become a focus for reach back to ECMWF in the future. In a broader forecasting context, development of further diagnostics drawing attention to high-impact surface weather elements and features remain highly desirable, e.g. sting jets (perhaps DSCAPE fields to identify areas of possible CSI could be introduced).

Finally, a couple of suggestions regarding meteogram visualization are put forward. Firstly, the Meteogram view that can be generated on ecCharts has proved popular and incredibly useful but displaying meteograms can obscure large parts of the display on ecCharts - could an option to open the Meteograms in another browser session be developed? Secondly, the ability to see and highlight members in a meteogram to better understand the behaviour of individual or a cluster of members would be desirable. For example, highlighting the coldest 10 per cent of lowest air temperatures and then being able to see on an adjacent meteogram what the associated precipitation signal is would help understand if a cold signal is associated with snowy or dry conditions.

3. Verification of products

3.1 Objective verification

Describe verification activities and show related scores. 3.1.1 Direct ECMWF model output (both HRES and ENS) Nothing to report.

3.1.2 ECMWF model output compared to other NWP models

Nothing to report.

3.1.3 Post-processed products

<u>Decider</u>

Probabilistic weather pattern forecasts from Decider (mentioned in Section 2.1.3) are routinely verified for a range of global models (including ECMWF) against the Met Office Global Model analysis, with each daily forecast lead time verified separately. Results show that forecasts verified on a day-by-day basis have forecast skill out to around 10 days for all 30 weather patterns (according to the Brier Skill Score; Fig 4). Some weather patterns are also forecast with skill beyond this period. The observed and forecast frequency biases are shown in Figs 5, 6 and 7 for ECMWF medium range, ECMWF monthly, and Met Office GloSea5 monthly forecasts. The day 0 frequency is the observed frequency so drift away from that value represents the bias. The results indicate changes in forecast frequency for some weather patterns. For example, Pattern 13 (anticyclonic north-westerly; Fig 8) occurs more frequently at longer lead times, whilst Pattern 7 (small low pressure system close to the UK; Fig 8) is actually forecast less frequently at longer lead times, only becoming close to the observed frequency at days 1-3. Pattern 7, notably, has the worst brier score of all patterns at longer ranges (not shown) with both patterns having among the lowest brier skill scores (e.g. Fig 4 for ECMWF medium range). Results are generally consistent between models; however, some interesting behaviour has been identified when comparing forecast biases between the ECMWF medium-range and monthly systems.

The ECMWF medium range mean forecast biases increase smoothly with forecast lead time (Fig 5), which reflects a gradual increase in forecast uncertainty with lead time. However, the ECMWF monthly forecast biases (Fig 6) are very noisy in the first 10 days and actually reduce slightly with lead time, before settling down and increasing smoothly out to the maximum forecast lead time in Decider for ECMWF of 32 days. This unusual behaviour may be caused by the low number of forecast updates from the monthly system, with only two updates weekly from the monthly system compared to two updates daily from the medium-range system. This means that the monthly forecasts are likely to be very jumpy from run to run within the short to medium-range period, but remaining more consistent in the long range due to the model being more likely to favour the climatological patterns for the time of year. In contrast, the Met Office long range forecast system (GloSea5; Fig 7) has daily updates and produces a transition in forecast biases within the two week time range which look much more similar to the ECMWF medium range system compared to the ECMWF monthly system. This result suggests that having more updates and/or a time lagged approach improves forecast consistency at least within the shorter time ranges.







Fig.5 Probabilistic weather pattern forecast frequencies for all 0000 and 1200Z ECMWF 15 day forecast runs between 3rd January 2010 and 27th April 2018. Day 0 frequencies provide the observed frequencies for each weather pattern over the verification period.



Fig.6 Probabilistic weather pattern forecast frequencies for all 0000Z ECMWF monthly forecast runs between 6th August 2016 and 19th March 2018. Day 0 frequencies provide the observed frequencies for each weather pattern over the verification period.



Fig.7 Probabilistic weather pattern forecast frequencies for all 0000Z GloSea5 monthly forecast runs between 18th January 2014 and 19th March 2018. Day 0 frequencies provide the observed frequencies for each weather pattern over the verification period.



Fig.8 The complete set of 30 Met Office weather patterns (from Neal et. al., 2016).

Finally, it is worth noting that forecast skill is negative for nearly all weather patterns and models beyond around 15 days (e.g. Fig 4 for ECMWF medium-range). However, this is based on daily verification which is not how long-range forecasts are used in reality. To help extend the period of useful forecast skill, future verification for long-range forecasts could aggregate daily probabilities over rolling time-windows spanning several days, with these time windows possibly increasing with lead time.

Tropical Cyclones

At the Met Office, the MOTCTracker tropical cyclone tracker (Heming, 2017) is used to identify and track tropical storms in the MOGREPS-G, ECMWF ENS and NCEP GEFS ensembles. Ensemble-based products for named storms from each model, and from the multi-model ensemble are then created (see Fig 9). Other products are created for each tropical storm basin, including an animation of 24-hour tropical cyclone activity forecasts (including both named storms and those forecast to form during the forecast), and a product displaying forecast tropical cyclone centres coloured according to model or intensity.



Fig.9 Multi-model ensemble forecast for Hurricane Irma from 12z 21st August 2017, 138 hours prior to landfall in Barbuda. The tracks from each member, coloured according to the model, is on the left image. The centre image shows the multi-model ensemble probability that the storm will pass within 120km in the next 7 days, with the deterministic and ensemble mean tracks overlain. The right image shows a storm-following meteogram which shows the spread in intensity values (mean sea level pressure minima and 10m wind maxima).



Fig.10 Reliability diagram (left), and relative economic value plot (right) comparing MOGREPS-G, ECMWF ENS, NCEP GEFS and multi-model ensemble forecasts of named storm strike probability, from January to December 2017. There were 82 named storms in this time, across the six tropical cyclone basins.

Verification results for the named storm strike probability forecasts (the probability that the storm will pass within 120km within the next seven days) are shown in Figs 10 and 11, The results include 82 named storms in the 12 month period January to December 2017 for the three individual ensembles and the multi-model ensemble combination. The reliability diagram shown on the left in Fig 10 displays good reliability for all models, with ECMWF ENS showing near perfect reliability for all probabilities. MOGREPS-G and NCEP GEFS both show overforecasting for probabilities 50% and greater, with NCEP GEFS contrastingly showing slight under-forecasting for 0-20% probabilities. In the relative economic value plot on the right, the multi-model ensemble value curve fully encompasses the three individual models showing the multi-model ensemble combination gives the greatest potential economic value for all cost-loss ratios. All the models display the greatest relative economic value (over 0.7) for very small cost loss ratios (0 to 0.1).

UK



Fig.11 Brier Skill Score of MOGREPS-G, ECMWF ENS, NCEP GEFS and multi-model ensemble forecasts of named storm strike probability: All storms from January to December 2017 and split by tropical cyclone basin (left) and a comparison of storm-based verification for Hurricanes Irma and Matthew (right).

Fig 11 compares the Brier Skill Score for the strike probability forecasts between model, basin and storm. ECMWF ENS is the most skilful of the three included global ensembles in most basins, with the relative performance of MOGREPS-G and NCEP GEFS varying from basin to basin. However, in all basins additional skill is gained using the multi-model ensemble. The storm-based verification of two high-profile storms (Hurricanes Irma and Matthew) further demonstrates the value in the multi-model ensemble as for each storm a different ensemble displays the highest skill (ECMWF ENS for Irma and MOGREPS-G for Matthew). In both cases, the multi-model ensemble is of comparative skill to the strongest performing model (which would not be known at the time of the forecast). More details on the ensemble-based tropical cyclone products and verification can be found in Titley and Stretton (2018).

3.1.4 End products delivered to users

Nothing to report.

3.2 Subjective verification

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

Whilst no information concerning subjective scores is available, it is, perhaps, of interest to note that subjective verification by Met Office operational meteorologists paints ECMWF output in a favourable light. Some biases/weaknesses/challenging forecasts have been noted, including:

- Storm David/Friederike (17-19th January 2018) was a particularly difficult forecast as all models, including ECMWF struggled. The recollection of those on shift at the time is that HRES tended to flip/flop with its ENS and also continually underplayed the development. This led to some challenging forecasting and difficulty establishing potential impacts.
- Met Office meteorologists seconded to the British Antarctic Survey to work in the Antarctic during austral summer 2017-18 reported poor performance with low cloud and fog around the fringes of Antarctica. This may be linked to there appearing to be many sources of low-level atmospheric moisture not captured by ECMWF HRES and ENS, e.g. brackish sea ice cover, polynyas and large melt water pools on ice shelves and ice sheets (admittedly it would be hard to capture these explicitly). ECMWF output appeared too dry and hugely deficient in low cloud and fog in and around these features in particular. They postulate there may be similar summer time issues in the Arctic.
- Snow accumulations have not been good, though it is recognised that ECMWF recognise this as an issue.
- Tendency for the ensemble to jump with HRES.

Underestimation of maximum temperature, both in HRES and ENS during warm spells over the UK in April
and early May 2018 are good examples. During these events, even the most extreme tail on ENS plots fell
well short of actual values (Fig 12).



Fig.12 ECMWF ENS meteogram from early May 2018 for the Reading area. Note only one ENS member signals values in the low 20s Celsius for the 6th when in reality nearby Farnborough reached 25.3C.

It is, of course, recognised that, firstly, ECMWF output is not necessarily primed for forecasting some of these elements and that developments are taking place to tackle some of these issues, for example extreme point precipitation and visibility.

3.2.2 Case studies

See Sections 2.3.1 and 2.2.1 for a case during the the severe cold spell over the British Isles in late February and early March 2018.

4. Feedback on ECMWF "forecast user" initiatives

The known IFS forecast issues page is something the Met Office have pushed for strongly in the past, so its creation has been welcomed and this page continues to be widely used by our operational meteorologists. The severe event catalogue is less widely used by Met Office operational meteorologists and, in view of Met Office worldwide forecasting commitments perhaps something Met Office operational meteorologists may be able to contribute to. The use of this catalogue, meanwhile, to draw out common predictive elements on different time scales as the number of cases studies grows, as illustrated at recent UEFs, is likely to be of huge benefit. Meanwhile, the update to the Forecast User Guide is welcomed. A 'print version' is something that some users have comment would be useful.

5. References to relevant publications

Carroll, E.B., 1997: A Technique for Consistent Alteration of NWP Output Fields. Meteorol. Appl. 4: 171-178.

Heming, J.T., 2017: Tropical cyclone tracking and verification techniques for Met Office numerical weather prediction models. *Meteorol. Appl.*, 24, 1–8.

Hewson, T.D., and Titley, H., 2010: Objective Identification, Typing and Tracking of the Complete Life-Cycles of Cyclonic Features at High Spatial Resolution. *Meteorol. Appl.*, 17, 355–381.

Neal, R.A., Boyle, P., Grahame, N., Mylne, K. and Sharpe, M., 2014: Ensemble Based First Guess Support Towards a Risk-Based Severe Weather Warning Service, Meteorol. Appl. 21, 563-577.

Neal, R., Fereday, D., Crocker, R. and Comer, R., 2016: A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteorol. Appl.*, 23, 389-400.

Price, D., 2017: Using Weather Pattern Analysis to Identify Periods of Heightened Coastal Flood Risk in the Medium to Long Range. *ECWMF UEF 2017* available at https://www.ecmwf.int/en/learning/workshops/usingecmwfs-forecasts-uef2017.

Robbins, J. and **Titley**, H. 2018: Evaluating high-impact weather forecasts from the Met Office Global Hazard Map using a global impact database. Accepted for publication in *Meteorol. Appl.*

Titley, H.A. and **Neal, R.A.**, 2017: Processing ECMWF ENS and MOGREPS-G Ensemble Forecasts to Highlight the Probability of Severe Extra-Tropical Cyclones: Examples from the Recent Storm Doris. *ECWMF UEF 2017* available at <u>https://www.ecmwf.int/en/learning/workshops/using-ecmwfs-forecasts-uef2017</u>.

Titley, H. A. and **Stretton, R**. 2018: A probabilistic evaluation of global tropical cyclone forecasts from the upgraded Met Office MOGREPS-G ensemble, and the value of multi-model ensembles. *33rd Conference on Hurricanes and Tropical Meteorology*, Ponte Vedra, FL, US. Extended abstract.