TECHNICAL MEMORANDUM

797

Use and Verification of ECMWF products in Member and Co-operating States (2016)

Tim Hewson

Forecast Department

February 2017

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/research/publications

Contact: library@ecmwf.int

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

Abstract

Each summer ECMWF invites Member States and Co-operating States to submit updated reports on the application and verification of ECMWF's forecast products. This report summarises the findings from the 2016 reports, in combination with findings from official visits made to certain member and co-operating states during the preceding 12 months, and in combination with feedback provided during the "Using ECMWF Forecasts" (UEF) workshop held in June 2016. As such this report provides a snapshot, for 2016, of the way in which ECMWF forecasts are being used, of how these forecasts are performing relative to forecasts from other models, notably Limited Area Models (LAMs), and of aspects of ECMWF output that users would particularly like to see improved. The report does not refer to ECMWF's in-house verification of its own forecasts.

Key findings regarding usage are as follows. At short lead times, up to about day 3, ECMWF IFS (Integrated Forecast System) output provides boundary conditions required for most operational LAM runs in Europe, with forecasts from those LAMs typically being used alongside the IFS forecasts themselves. From day 3 onwards direct IFS output is the primary source of forecast information, and for lead times of 10-30 days, and beyond, it is the only source for most users. Interest in monthly and seasonal time ranges continues to grow, fuelled in particular by requirements of the energy sector.

Regarding verification, most reports focused on comparing HRES (ECMWF's high resolution forecast) with LAMs, for sensible weather parameters. Relative performance varies, but overall HRES performance is as good as LAM performance. However LAMs have better precipitation biases than HRES, versus point totals, and also usually perform better over complex topography. Some improvement in HRES output was noted, as expected, when resolution increased from 16 to 9km in March 2016, though only limited evidence was available at the time of the reports.

Model problems highlighted by users include ongoing issues with incorrect sea ice cover, and, partly related to this, near-coast sea surface temperature errors. As in previous years ECMWF also received reports of problems with inversions and low cloud cover, notably in winter. There was also reference to underestimation of peak totals in orographic precipitation events, and to poor handling of summer-time convection in some areas near the Alps. There was positive feedback regarding improvements to near-coast 2m temperatures following the change ECMWF made to the radiation scheme in spring 2016.

1. Introduction

Each summer ECMWF invites Member States and Co-operating States to submit updated reports on the application and verification of ECMWF's forecast products. The NMSs (national meteorological services) submitted their reports (21 out of 34), which can be accessed on the ECMWF website.

A summary of the NMS reports is presented below. Content has been combined with (i) feedback from the "Using ECMWF Forecasts" (UEF2016) workshop held at ECMWF from 6 to 9 June 2016 (primarily from breakout groups), and (ii) feedback from official triennial Member State/Co-operating State visits undertaken by ECMWF between July 2015 and June 2016.

Note that the results of ECMWF's own objective verification are not included directly here, but can be found in Haiden et al (2016).

Section 2 provide background information for this report whilst Section 3 discusses how ECMWF forecasts are used. Sections 4 and 5 deal, respectively, with objective and subjective verification aspects. Section 6 looks at monthly and seasonal forecasts whilst section 7 outlines the results of a small survey into utilisation of ECMWF's online resources for forecasters.

2. Background

Reports received from the member and co-operating states are now available on the ECMWF website at:

http://www.ecmwf.int/search/elibrary?year=2016&secondary_title=Green%20Book%202016

Reports came from 21 states: Belgium, Croatia, Czech Republic, Finland, France, Greece, Hungary, Iceland, Ireland, Israel, Italy, Latvia, Macedonia, Montenegro, Portugal, Romania, Serbia, Spain, Sweden, Switzerland and Turkey. Contributions were invited under the following headings:

a) Summary of major highlights

b) Use and application of products

c) Verification of products (objective and subjective)

d) Feedback on ECMWF "forecast user" initiatives

e) References to relevant publications

This document presents a summary of those contributions, combined with feedback from the "Using ECMWF Forecasts meeting" (UEF2016), held at ECMWF 6–9 June 2016, and also from official ECMWF visits to Member and Co-operating States undertaken in the past 12 months. In chronological order these were to: Bulgaria, Hungary, Slovakia, Czech Republic, Italy, Latvia, Lithuania, Portugal and Israel.

The ECMWF IFS undergoes improvements each year, which naturally affect aspects of performance, so summary information presented here has to be read with this in mind. For example in March 2016 the grid structure was changed and the resolution of HRES was increased from 16km to 9km whilst the two resolutions used by the ENS (ECMWF's ENSemble prediction system) were increased from 32km and 64km, to 18 and 36km respectively. These changes facilitated better representation of small scale meteorological features, notably those affected by coastal and topographic details. Moreover the jump in ENS resolution used to occur at the end of day 10 but now it is later, at the end of day 15, and so only affects the monthly forecast.

3. Use of IFS Output

In the shorter ranges (typically to 48–72h ahead), ECMWF IFS products are commonly used in conjunction with products from other sources, notably limited-area deterministic systems, but to an increasing extent limited area ensemble systems too. In most cases ECMWF IFS data provides boundary conditions (BCs) for these limited-area runs, four times per day.

In the medium, extended and longer ranges, ECMWF products are generally the main or the only output used by Member States/Co-operating States. It seems that all Member States/Co-operating States provide forecasts of some sort up to lead times of 10 days. When looking beyond, to longer and longer lead times, the number of organisations with formal forecast commitments slowly diminishes but it is clear, as in previous years, that this is a growth area in many countries. Growth is commonly being driven by requests from customers such as energy suppliers.

There seems to be clear evidence that, relative to other model output, ECMWF output is used more now than hitherto. Frequent usage is on the basis of (i) its relatively high skill, and (ii) the fact that the resolution upgrade allows for a better identification of some weather details that were traditionally the remit of LAMs. With regard to (i) Ireland note that when their forecasters can choose whether to drive a short range road ice model with forecasts from ECMWF or HIRLAM or HARMONIE it is "the ECMWF model, by some distance, that is the most popular choice". Regarding (ii), Norway are now phasing out use of their local HIRLAM configurations (at 8 and 12 km resolution), illustrating that our resolution upgrade has provided a direct cost saving for them, whilst Portugal are similarly anticipating the phasing out of ALADIN, as its function transfers over to HRES, although they are retaining 3 AROME domains at 2.5 km resolution. It should also be stressed that our changes are allowing the Member and Co-Operating States to focus more on modelling at even higher resolution, as this Portuguese example illustrates.

Whilst the main operational focus for all Member and Co-operating states is on local weather, many also have international commitments, for which ECMWF forecasts are very regularly used. Norway for example has undertaken capacity-building activities in three countries in Southeast Asia, and for this HRES and ENS data are being promulgated.

3.1 Local post-processing of model output

3.1.1 Statistical adaptations

Approximately 80% of states apply some local statistical 'recalibration' procedures to post-process ECMWF forecasts, especially to make forecasts of sensible weather for specific locations. HRES has always provided the main input, though ENS distributions are also being modified in some countries. Activity in these fields has continued in spite of the fact that raw model output becomes ever more accurate, and, as resolution increases, ever more representative of point locations. It will be interesting to see in 2017 if there has been much of a change following a full year with the resolution upgrade. Research studies suggest that one can continue to make noteworthy improvements via statistical post-processing even as forecast quality increases.

The main post-processing methods are labelled Model Output Statistics (MOS), in which model bias is catered for, and Perfect Prog (PP), in which it is not. Statistical techniques include height mismatch adjustment and bias removal, which although simple can dramatically reduce the mean error (as reported by Iceland). Other more advanced techniques include Kalman filtering, multiple linear regression, polynomial regression, logistic regression, linear discriminant analysis, percentile matching and non-homogeneous Gaussian regression. Indeed many countries report using a mixture of approaches.

Most calibration techniques are based on real-time forecast performance, as measured using standard synoptic stations. There is also continued use of analogues by Serbia and the Czech Republic for monthly forecasts.

Statistical combinations of forecasts from ECMWF and other models are common, employing weighting according to lead time. The other models are usually limited-area models with higher horizontal resolution than HRES (usually in the range 1–7 km).

ECMWF reforecasts are increasingly seen as being a very convenient and important dataset for re-calibration activity, as noted in Norway's report. Indeed there were strong requests from many users at the last two UEF meetings to provide a full year's worth with each new cycle. However, such an initiative would delay implementations and require substantial computer resources that are not available at the current time.

3.1.2 Physical adaptation

ECMWF output, in one form or another, is used very widely to provide boundary conditions (BCs) for running limited area models, and in some instances initial conditions (ICs) for those models too. Mostly this happens via the ECMWF Optional Programme "Boundary Conditions for Limited Area Modelling" (BC Optional Programme), which provides additional forecasts from 06 and 18Z data times. The limited-area model suites most widely used in this way are ALADIN, AROME, HIRLAM, HARMONIE, COSMO and WRF.

Initially only a single dataset was available, to help drive deterministic systems, but now ENS BCs are also available and are steadily being adopted (e.g. Hungary, Italy, Switzerland, and soon also Sweden). Indeed Hungary manage to run one of their high resolution ensemble systems eight times per day, using ECMWF BCs.

BC sampling times vary between one and six hours, with one country reporting using 3h or 1h temporal resolution depending on whether the nested model they use is respectively higher or much higher resolution than HRES. In any case the trend to shorter sampling times continues, based on evidence of performance improvement. To provide further assistance in this regard ECMWF introduced hourly output from standard 00 and 12UTC ENS runs also, with cycle (43r1) that became operational in November 2016.

"Multiple nesting" approaches seem to have reduced in the last year, with Switzerland for example retiring the use of COSMO-7 (7 km) as the intermediate nest of 3, the view being that HRES and ENS with their upgraded resolutions are now good enough to use as the sole drivers for COSMO-1 (1 km deterministic) and COSMO-E (2.2 km ensemble) respectively. Removing the need to run a third model is clearly an efficiency saving for MeteoSwiss.

ECMWF's model suite also provides ICs and BCs for Portugal's SWAN wave model. Increased BC and IC use for wave forecasting on various scales in Portugal is envisaged as ALADIN, their other driving model, is retired. Finland and Greece also use ECMWF BC data for wave modelling, whilst France's internationally competitive wave model is driven directly by IFS winds.

Even NMSs that have the capacity to provide BCs from their own global models still use ECMWF output of certain types for BCs and ICs. For example IFS and CAMS (Copernicus Atmospheric Monitoring Service) data are being used in collaborative fashion in France in the evolving MOCAGE suites, for pollutant modelling.

HRES and ENS fields continue to also be used or adapted to drive trajectory and dispersion models in numerous countries, with Italy using ECMWF Metview FLEXTRA code for the trajectory calculations themselves. And in several countries there are application models of other types (e.g. road state, hydrological and oil spill models) that are or can be driven by ECMWF data.

3.2 Derived fields

3.2.1 For forecasters

Many countries perform additional post-processing to provide derived products that historically ECMWF has not provided, and also tailored versions of those products that ECMWF has been providing.

For non-bespoke fields ECMWF continues to provide efficiency-savings to its users, by adding more derived fields to its range of web products, and to the ecCharts platform, and by fulfilling specific requests for new products whenever possible. As an example, in their report the Czech Republic note that they themselves have been computing CAPE and CIN fields, precipitation type, wind shear between different pressure levels and fog (visibility). Now, in the last two years, ECMWF have added all of these parameters to ecCharts. Portugal have also been computing two types of vertical wind shear, as a well as a “Fog Stability Index”. It would be interesting to see how the latter compared with the IFS visibility products. Meanwhile Hungary have taken our new precipitation type diagnostic from the ENS, and created a “precipitation type probability diagram” which has engendered “much attention and very positive feedback from forecasters and internal users”. An example is shown in Figure 1. ECMWF is now building on this idea with a view to developing new products: for example, a new component for meteogram displays.

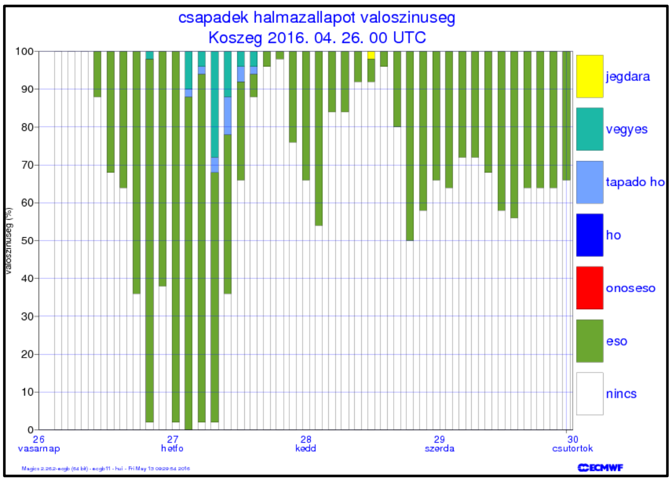

Figure 1: Precipitation type probability chart example from Hungary. Legend colour meaning, from top down is: ice pellets, sleet, wet snow, snow, freezing rain, rain, dry.

The primary forecasting challenge for which derived products are developed seems to be severe convection. This is because (i) ECMWF cannot explicitly represent the related hazardous weather - hail, lightning, tornados, extreme gusts – even with the improved global model resolution, (ii) this weather can be very destructive, (iii) warnings of such are required of NMSs, (iv) indices from global models can provide useful pointers. Convection-related indices thus generated in Member States and Co-operating States for forecasters include Jefferson, Lifted, SWEAT, Showalter, KO, wind shear and helicity. Many countries also compute total precipitable water. Finland and Portugal compute thunder/lightning probabilities from such indices. Meanwhile Hungary displays ENS vertical profiles in ‘spaghetti’ format. Whilst the range of convection-related products that ECMWF provides directly is somewhat limited (which is another reason for the Member and Co-operating State activity) last year EFI and SOT-style products were introduced on ECMWF's website, for CAPE and "CAPE-SHEAR". Evidence suggests these were welcome and are now being actively used in the Member and Co-operating states. ECMWF plans to further improve these particular diagnostics in the coming year, and also plans to investigate whether better guidance on hail potential can be provided (noting that many countries invest in hail suppression programs).

Flash floods can be devastating and commonly also relate to active convection, and this is another very challenging area for Member and Co-operating State forecasters as noted in various communications (Hungary for example have this year investigated 100 extreme rainfall events, whilst Greece report that two recent, destructive flash flood events were not well handled by IFS). In recognition ECMWF is now working to downscale gridbox forecasts from its ENS to provide probabilistic forecasts of rainfall at points, recognising that sub-grid variability and biases vary hugely according to situation (e.g. slow moving convection, orographic rainfall). Initially the resulting products will provide input to the GLOFAS initiative, to help provide early warnings of flash flood potential. In due course these pointwise forecasts may be made more widely available as a product.

Other products derived for forecaster use include clustering of ENS members for specific areas and parameters (Hungary), sea state codes (Italy), a rime predictor (Czech Republic), a daylight index (Latvia), a thermal front parameter and Q-vector convergence (Portugal), and monthly and seasonal meteograms (Hungary).

On the technical front, Serbia, Hungary and Latvia generate cross-sections for forecasters. Various interpolation procedures and grids are in use, whilst Sweden apply smoothing to IFS output on the basis of 20 km radius circles.

3.2.2 For the public

Various miscellaneous public-oriented fields continue to be generated across the Member and Co-operating States, for example "characteristic weather" symbols (e.g. Italy, Portugal), for mobile phone "Apps" for example, and "traffic light" indicators for uncertainty (Norway). Meanwhile Belgium create a "Wet Bulb Globe Temperature" index from the ENS, to focus on human discomfort from heat stress.

Interestingly, Portugal report that their weather symbol generation tool has benefited greatly from incorporating also ENS data: the resulting map-based output now looks smoother and more meteorologically reasonable.

3.2.3 For customer applications

Application and adaptation of ECMWF model forecasts to specific societal issues continues. Many countries have reported a continuation of products they cited in reports they provided prior to 2016.

Examples, in use or under development, include indices for potato blight and slurry-spreading (Ireland), for aviation an icing index (Portugal and Italy), a rime parameter (Czech Republic), condensation on railway equipment (Belgium), UV indices (Israel) and climate indices at seasonal time ranges (Switzerland). HRES output is also used in tandem with scatterometer retrievals, to give wind speed analyses with high spatial resolution, especially near to coasts (Portugal).

Iceland post ECMWF SST forecasts on their website for the marine community, though whether these are from the uncoupled HRES (where any evolution is based on climatology) or the ENS (which is coupled) is not clear. There may be more benefit to be had from using the ENS now, since the ocean model moved to much higher resolution in November 2016, with the implementation of IFS Cycle 43r1.

Rapid growth in energy provision from renewables (wind and solar power in particular) continues to drive product development in certain spheres. This continues to be a key area, presenting security issues for society due to an innate increase in our dependence on the weather for electricity provision. Some countries now have a sharper focus on related IFS quantities; for example helpful feedback from Israel facilitated the reduction of unwanted biases in a UV product from CAMS.

3.3 Direct use of ECMWF products, including severe weather prediction

3.3.1 Timescales

ECMWF products are primarily used for days 3 to 10, though most services also reference products for shorter ranges (and as discussed above *indirect* use for short ranges, via provision of BCs for LAMs, is commonplace). Use of and demand for forecasts for days 10 to 30 and seasonal ranges continues to grow (Section 6 below).

3.3.2 Severe weather

IFS forecasts are very widely used for official severe weather warnings, for lead times up to about 5 days, with participants at the UEF meeting re-iterating that "helping us provide early warnings of high impact weather" should continue to be a key target for ECMWF. Also stressed at the UEF meeting was a "special interest in worst-case scenarios", which melds well with the ensemble approach at the core of ECMWF's new 10-year strategy.

The longest lead time at which a warning will be issued varies by country (e.g. for Italy it is only about 24h), and so as with everyday forecasts the extent to which ECMWF output is actively used also varies, with nowcasting tools and LAMs being utilised more at short leads.

Warning systems are slowly becoming more probabilistic in nature, though from a purist's perspective many warnings are arguably still too "deterministic". In tandem, some countries continue to place more reliance on HRES than ENS (e.g. Croatia/Ireland, which may relate also to the focus on shorter leads), though overall ENS usage is probably increasing. There also appears to be a gradual evolution from threshold-based warnings to ones that are more impact-based. This focus on impacts relates closely to the return period philosophy that underpins the ECMWF EFI and SOT products.

Indeed the usefulness of the EFI and SOT for alerting forecasters to severe weather events (rain, wind, heat, cold, snow) is now very well established. Favourable comments are regularly received by ECMWF. Finland and the UK have also highlighted how the EFI can be especially useful beyond the local region, when forecasters' knowledge of local climate is lacking. In the coming year ECMWF is planning to further enhance the EFI and SOT product suite, extending the daily products beyond day 7.

Other ENS products, such as probabilities, are widely used in conjunction with the EFI to provide additional uncertainty estimates. These can be particularly valuable in countries where warnings continue to be threshold-based, and several countries report that the ENS is tailored accordingly, via ecCharts or other mechanisms. Norway and the UK have also highlighted how products from the "cyclone database" (extra-tropical cyclone tracker) assist in the warning process, by clearly highlighting uncertainties in the behaviour of cyclones that are responsible for severe weather, whilst Greece report the strong desire of forecasters to attach precipitation events to specific “weather systems”.

As highlighted earlier, the prediction of severe convective storms and related hazards are not as well-served by ECMWF as the prediction of most other types of severe weather, though the situation here has improved, with the resolution upgrade, and with more enhancements in the pipeline for our relatively new convective EFI fields.

Tropical cyclone-tracking products are used by some Member and Co-operating States, due to certain NMSs having specific responsibilities in tropical regions (notably France, from La Réunion).

Whilst many warnings focus on standard meteorological parameters such as wind and rain that are well served by ECMWF, warnable phenomena in some countries are more diverse, including, for example avalanches and landslides (Iceland and Norway), and also flooding related to ice break-up in spring (Latvia) which is not addressed by the EFAS setup at present. For these more diverse categories IFS data is also used, but only indirectly.

3.3.3 ecCharts and other web products

ecCharts was being actively used in 5 of the 9 states visited over the past year, and these users greatly appreciated its functionalities, together with the special training sessions that ECMWF provided during some of the visits. In the other 4 states use was limited or nil. Reasons for this varied. In some countries (e.g. Hungary) ECMWF data is ingested into advanced local forecaster workstations, which provide all the tools needed, as well as the additional facility to overlay other models and observational data (note also that ecCharts’ WMS capabilities permit product transfer into such a platform). More concerning was the fact that other countries had some issues with ecCharts. Complaints about slow speed continue, at the user meeting also, whilst Italy initially described ecCharts as “not fit for operational use”. ECMWF has continued to work hard to address these concerns, for example by proposing ways to tackle local IT/firewall issues, which can sometimes explain the lack of speed in a particular country. In the case of Italy it transpired that there was a training requirement, so ECMWF ran a special webinar at the end of May, which attracted 81 participants. Via IT checks it has also been ascertained that ecCharts was running at normal speed within Italy.

ECMWF continues to add parameters, display schemes and extra functionality, on a regular basis (key updates being in early summer and early winter). Following requests, with cycle 41r2 in March 2016ECMWF introduced a new facility to allow e-suite fields to also be examined within ecCharts. This has proved popular and was praised at the UEF meeting. ECMWF will work to further improve this, by for example reducing ‘random’ non-availability of data for some time steps.

One concern raised by users (e.g. Bulgaria) was the demise of the “clickable EFI” charts when the old-style ECMWF web site was retired in late 2015. This facility has now been replaced using new and much improved technology on the new web site, which presently offers 4 types of clickable chart (with more to be added; previously there was just 1), more display options, and far more domains, including ocean-centric regions for marine forecasting. Italy described this facility as “highly useful”.

Extensive use is still made of the more traditional single-image web products, with Meteograms (hitherto called EPSgrams) still apparently being the most widely used of ECMWF products. Ireland report that usage of web site products continues to increase.

3.4 User Requests

There were many requests again for new output from ECMWF, right across the product range. These are not discussed in this document. As regards more general requests relating to ECMWF operating policy, the recurring themes in 2016 were similar to 2015. At the UEF meeting for example many users wanted higher temporal resolution in HRES and ENS output (e.g. hourly) partly because of the more exacting requirements of customers in renewable energy, some want more low level wind heights, for wind power applications, most want earlier delivery of everything, some also requested four IFS runs per day, some want output to week 6 in the monthly forecast, and as discussed above a number want a full year of reforecasts to be made available with each new cycle. Mostly these depend both on the recommendations of the Committees that provide governance for ECMWF , and on computational resources.

There have also been requests for post-processed data (in which known biases are removed, for example). Traditionally post-processing has been done downstream, but ECMWF is now investigating the possibility of in-house post-processing, involving external collaboration where appropriate. This could potentially bring substantial efficiency savings for Member and Co-operating states, by reducing duplication of effort.

ECMWF continues to be asked how forecasters should deal with “forecast jumps”. A module in the products training courses now addresses this; in summary, statistics suggest that on average jumpy forecasts are no less accurate than non-jumpy ones. At the same time there is a lot of interest in forecasts of regime changes, and this is reflected in ECMWF’s new 10-year strategy 2016–2025.

4. Objective Verification of Products

4.1. Overview

Most countries have reported results from the verification of ECMWF forecasts, generally by comparison with observations in the local area of interest, and often also comparing with forecasts by their own high resolution models. Of relevance to interpretation are the dates of the most recent upgrades to the IFS:

Cycle 40r1 became operational 19 Nov 2013

Cycle 41r1 became operational 12 May 2015

Cycle 41r2, with its resolution upgrade, became operational 8 Mar 2016

This means that in reports submitted in 2016 verification periods correspond mainly to cycle 41r1, but (depending on the period chosen) some will also span cycles 40r1 and 41r2. It is reasonable therefore to expect to see some improvements, overall, in performance relative to 2015 (when 40r1 was the main cycle referred to). However note that, as always, year-on-year changes in IFS performance depend also on the prevalence of different synoptic patterns (that can have different error characteristics associated).

Biases and errors in sensible weather parameters can depend strongly on the prevailing weather types, which differ each year. And when considered alongside the (fixed) geographical differences between regions, and the different impacts that those can have in a given synoptic pattern, it is no surprise to sometimes see different findings from different countries.

Another issue, affecting inter-comparability between different countries' results for HRES, and indeed comparisons between HRES and LAM forecasts, is the range of "interpolation" and "site-selection" techniques that can be used. Sometimes, for example, full resolution IFS output is not being exploited. Moreover in some reports received the method(s) of extraction and interpolation used were not entirely clear.

Many reports focussed on comparing HRES with LAMs, and for this reason usually centre on the shorter ranges (up to about 48h).

A common finding, seen in virtually every verification result, for almost every sensible weather parameter, was that biases in IFS forecasts have a diurnal cycle.

Overall, as in 2015, HRES seems to perform just as well as LAMs, a result probably bolstered by the recent resolution increase. However, precipitation biases in LAMs, for both small and large totals (versus point observations), are mostly smaller. It is also clear that handling of surface weather parameters by different models can vary greatly according to synoptic situation, geographical region and parameter in question, and it seems that in mountainous regions LAMs do, in relative terms, tend to perform a bit better than HRES for some other parameters too. Ultimately all models have their strong and weak points, and the impression one gets from the wealth of statistics provided in this year's reports is that in the short ranges at least (where the bulk of the comparisons were performed) a multi-model approach to forecasting has considerable merit. This would be particularly true if one could vary weightings according to known synoptically-varying performance characteristics. There is little evidence that such a strategy is being applied automatically at the moment, though it is undoubtedly used in subjective fashion by forecasters.

Details, by parameter, are given in the following sub-sections. Some of the IFS issues raised here are known, and most of these are also listed in the ECMWF 'Forecasting Issues' web page at: https://software.ecmwf.int/wiki/display/FCST/Known+IFS+forecasting+issues which is updated a number of times each year.

4.2 2m temperatures

Combining the results from all reports one can conclude that for 2 m temperature HRES performs slightly better than LAMs. By better we mean smaller MAE and RMS errors, and these are usually accompanied by smaller biases. Evidence of better HRES performance can be found in reports from Hungary, Romania, Serbia, Turkey, Sweden, Czech Republic and Iceland. In some instances multi-model results are presented. Romania for example show HRES to perform better than ALARO and ALADIN and COSMO-7. The only evidence of (slightly) better LAM performance was provided by Sweden (2.5 km AROME model only), Iceland (2.5 km HARMONIE model only) and Greece (3 and 7km COSMO models). Better representation of complex terrain (mountains, lakes, small islands) undoubtedly contributes to the relatively high skill of high resolution LAMS in such countries. For Portugal, for 2m temperature, the performance of AROME and HRES, in the months since the resolution upgrade, has been about the same.

Where reference is made to equivalent statistics from previous years, two countries (Croatia and Serbia) suggest similar performance in the last two years, whilst Ireland show that biases have reduced over the long term. However, both Norway and Iceland provide positive feedback regarding sites near coasts in winter. Historically, in situations favouring radiative cooling, there were significant negative biases at these sites (especially where the land edge in plan view is convex). Even the general public had noticed this in Iceland! The error was because the larger-scale grid used for radiation calculations allowed emissions over land according to (SST)[4]. Now with a fix in place in cycle 41r2, and with higher resolution anyway, the problem has reduced (see also Seierstad et al, 2016).

As regards HRES biases and diurnal variations in those results were mixed. Sometimes they seem contradictory, but within Turkey's report one can see how such behaviour is very site-dependant. Also noteworthy in Turkey's report was the prevalence of higher RMS errors in the northeast of the country than elsewhere (3-5˚C versus 2-3˚C, at day 2 for VT=12Z).

Finland describe in detail a number of different types of temperature error, with some striking examples of missing inversions in radiation conditions, in both summer and winter (e.g. Figure 2), which implies positive biases in HRES that are sometimes very large.

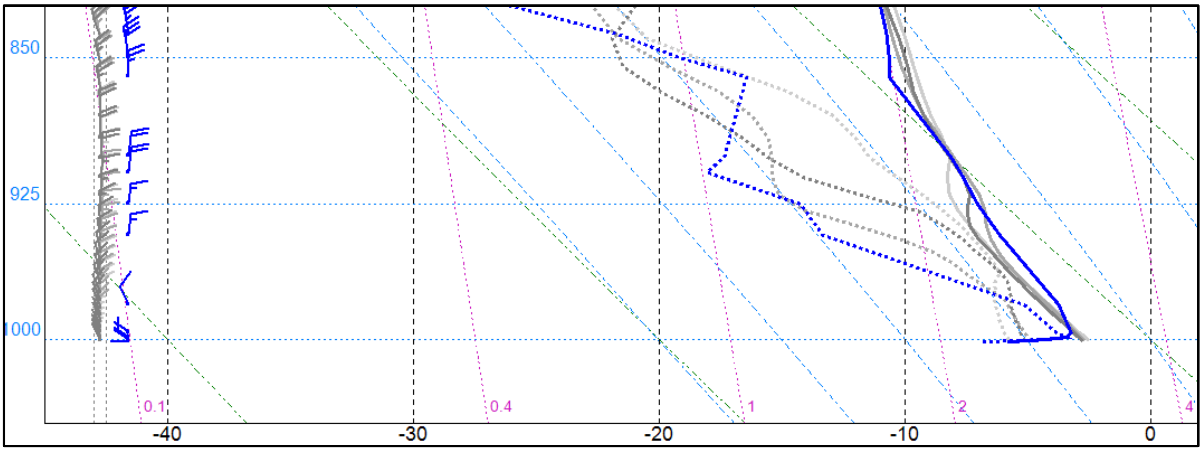

Figure 2: Example sounding from Jyväskylä 16/12/2015 18UTC (blue), with corresponding HRES profiles at T+6,18,30 (grey). The forecast soundings lack the inversion.

On the other hand Finland have also reported, again, that HRES temperatures during spring evenings can be much too low, in spite of related model changes ECMWF made last year. In turn this error can causes under-prediction of flood risk in the snowmelt season. And from bulk statistics it seems Sweden may experience similar spring biases. Temperature prediction and representation over Finland is particularly problematic because of: (i) the large proportion of lakes (now modelled in IFS using FLAKE = "Fresh shallow water LAKE model"), which have very different thermal properties to land, where observations are based. (ii) the prevalence of forested areas, when observations will usually be in a clearing with different thermal characteristics. (iii) the prevalence of low sun angles for which radiative transfer calculations are more challenging, due to intermittent shading. Nonetheless ECMWF is fully aware of these issues and continues to work towards improved performance.

Croatia have also noted that temperature prediction is much more problematic in stable winter situations, which is indeed an issue right across Europe.

4.3 10m wind

Combining results from different reports, as for temperature, it seems that on balance HRES is currently providing mean wind speed forecasts that are as good as those from LAMs. France, Hungary, Serbia and Norway all found HRES to be better, Sweden found performance to be similar, whilst Turkey, Iceland and Portugal found HRES to be worse.

Croatia provide subjective feedback that our “wind forecasts have improved”. Whether this relates specifically to the increased resolution is not clear, though it is not unreasonable to expect that better representation of topographic features will have a notable positive benefit on wind forecasts. Indeed in an ECMWF case study it seems that Bora winds are somewhat better represented at 9 than at 16 km (though still significantly under-predicted).

Notably, Iceland continue to highlight a marked under-prediction bias in HRES winds, which is particularly marked during cyclonic windstorms in winter. For them it is key to get these dangerous events right. Their warnable threshold is remarkably high, at 40 m/s. Their 2.5 km HARMONIE model continues to perform much better than HRES, as reported last year. Norway also comment and illustrate a substantial under-prediction bias at mountain sites, and much better performance here of AROME (2.5 km resolution) even if overall across Norway HRES is slightly better than AROME and HARMONIE.

Windstorms related to mountain waves can be completely missed by HRES, but captured by a LAM; Norway provide a nice example (Figure 3).

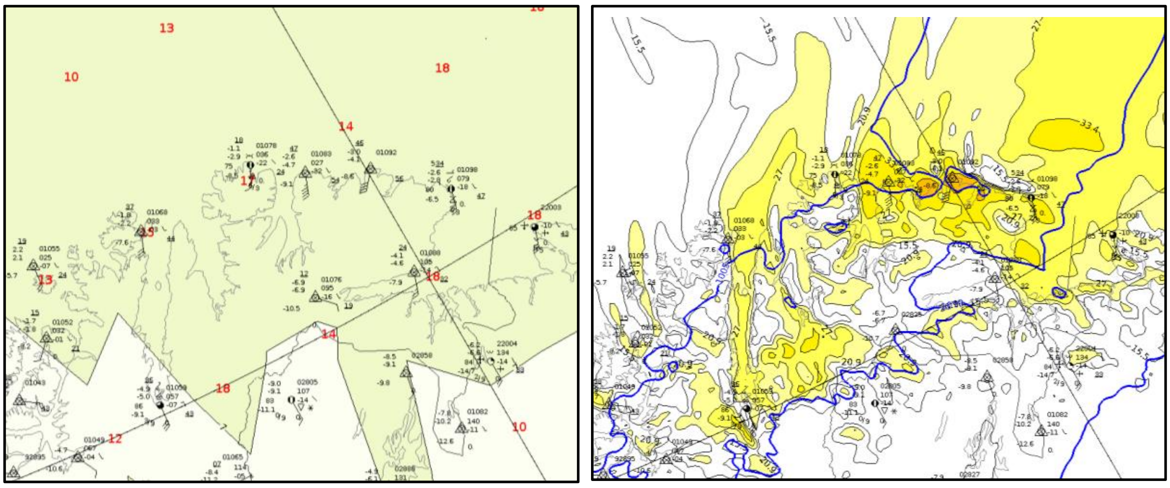

Figure 3: North Norway 10 mean wind forecasts from HRES (left) and a 2.5km AROME configuration (right) at 18UTC 7 November 2014. Observed values at the north coast reached 51kts, vs 44kts in AROME and only 24kts in HRES.

4.4 Rainfall

For rainfall forecasts many different skill metrics have been quoted, including, this year, SEDI (Symmetric Extremal Dependence Index), FSS (the Fractions Skill Score) and SEEPS (Stable Equitable Error in Probability Space). This wide range, the sometimes disaggregated nature of precipitation fields, and the different interpolation methods used all together make result intercomparison rather more difficult than it is for wind or temperature.

Firstly it should be noted that no country has compared forecast totals, which innately apply to a gridbox, with the average observed over a gridbox (although Sweden go some way towards this). This would be the correct approach, but is avoided because of technical challenges and lack of observations.

Many countries quote the frequency bias index, or FBI, for point measurements. For HRES plots are fairly consistent between countries. Small totals tend to be overestimated, and large totals underestimated, which is a virtually inevitable consequence of a point versus gridbox approach, so this is not necessarily a model problem. Meanwhile for LAMs FBI profiles are almost always more horizontal, and relatively close to 1, which is also what one expects, for smaller gridboxes. For HRES the "crossover" FBI=1 is quoted as 10 mm/24h by Switzerland, 10 mm/12h by Italy and 4 mm/24h by Hungary. Differences here may relate to topographical differences – in-house verification shows that areas in and around the Alps can experience very different biases and skill levels for summer rainfall. It should also be re-iterated here that ECMWF is currently investigating ways of converting forecast gridbox totals into point total pdfs (probability density functions). The adopted procedure can take account of both sub-grid variability and different types of bias, and in the future new products may result from this work.

When comparing HRES with LAMs, for relatively modest totals, up to about 5mm in 24hr, Romania, Hungary, Serbia, France, Italy and Greece all report HRES as being in some general sense better, though for Turkey WRF forecasts are subjectively rated as better, and for Sweden, using the Fractions Skill Score, AROME is better.

When examining totals over periods of less than 24h diurnal variations in bias become apparent. In results from Croatia it is clear that such variations in HRES were lower in 2015 than in any of the previous 3 years, whilst on plots from Italy it is clear that in HRES the diurnal variations in bias are lower than they are in COSMO-7.

The Czech Republic compare a number of meteorological parameters between models, and in this comparison it is clear that for HRES "areal coverage of precipitation", in summer, exhibits the worst performance, relative to other models, of any of the parameters examined. This will relate to representivity issues for convective precipitation.

Meanwhile Iceland and Norway report that substantial convective outbreaks can be under-represented by HRES. And at the UEF meeting some reported that over Europe there was too much convective activity (notably over mountainous areas), whilst others reported that there was too little (notably over flat areas). Meanwhile France highlight that in the tropics rainfall forecasts are more prone to errors than they are in the extra-tropics (in HRES and their LAMs), probably because tropical rainfall is on average more convective, and this is a more difficult forecasting problem, even in models for which parametrisation is not needed.

Finland point to a specific problem with "persistent showers" focussed on lakes, which should just be drizzle. This is a known model issue, which has now been documented, and which relates to the handling of aerosol. Aerosol are not advected but are a function of the underlying surface, and in turn droplet growth and capacity to rain out relate. A fix to this problem has been trialled, and looks successful, though there may also be adverse impacts on the tropics which need further investigation.

Again this year hardly any verification of ENS precipitation has been reported.

4.5 Ice-phase precipitation

Less feedback has been received this year regarding snow, perhaps because of the mild European winter, although there were comments again at the user meeting about the melting of lying snow being too slow (a known issue that will be addressed in due course).

Meanwhile Croatia highlighted how the relatively new precipitation type diagnostic, and associated cloud physics changes, introduced in 2015, helped them to accurately predict a freezing rain event on 5 January, as much as 5 days in advance. Forecasts at shorter leads were consistent and a red warning was issued. Hungary also provided a reasonably favourable appraisal of IFS guidance for a localised freezing rain event on 1 December 2015. Whilst Germany welcomed the model and diagnostic developments related to freezing rain they nonetheless discovered some issues with an event on 22 and 23 January. Although onset was well predicted by the IFS, the event lasted longer than forecast. This was attributed to cold air damming in valleys that were not resolved in the HRES topography.

4.6 Cloud and screen-level humidity

There were fewer reports in these categories, but HRES seems to perform slightly better than LAMs for low cloud (according to Hungary, Finland and Sweden), whilst for Greece HRES performs much better than their COSMO models for total cloud cover (see Figure 4, this may relate to lack of vertical levels in COSMO).

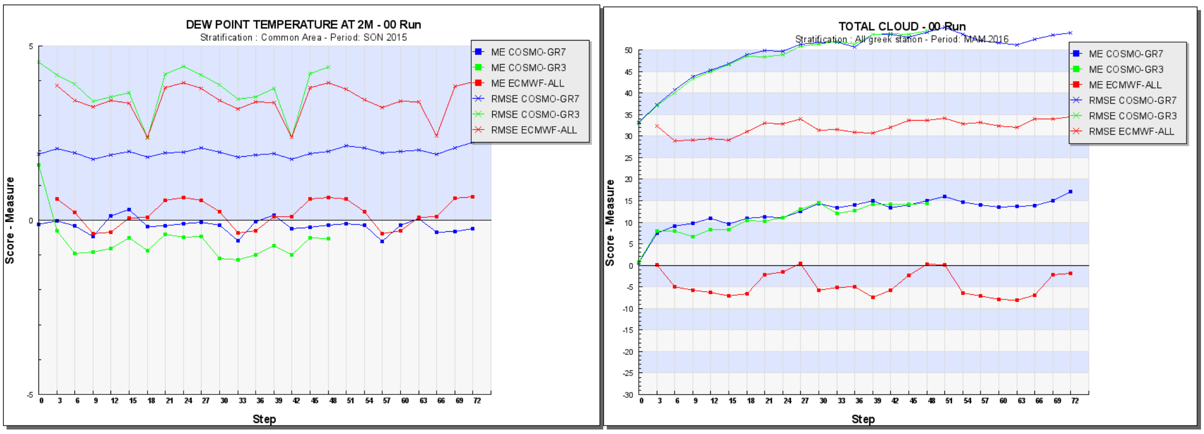

Figure 4: Multi-model verification graphs from Greece, based on 50 stations. On each panel the lower 3/upper 3 curves show bias/RMS error respectively; red is HRES, green 3km COSMO and blue 7km COSMO. Left panel is for 2m dewpoint (°C), autumn 2015; the COSMO 7km COSMO is clearly the best. Right panel is for total cloud cover (%), spring 2016; HRES is clearly the best – this probably relates to HRES' greater number of levels, 137 versus 35.

Both Finland and Sweden say that low clouds are under-predicted by HRES; for Sweden this is during windy weather. However, Sweden also say that they have seen some small improvements in low cloud cover in Cycle 41r2. Conversely Bulgarian forecasters have reported that stratus erosion is too rapid; this probably relates to the long-standing winter-time difficulties with handling inversions referenced above.

For cloud base forecasts, which are of particular relevance for aviation, Sweden show that AROME gives slightly better forecast than HRES.

For screen-level humidity Hungary, Sweden and Portugal all show HRES to be slightly better than their LAMs. For Greece results are mixed and counter-intuitive. Whilst the 3km version of their COSMO model had larger errors and biases than HRES, most notably in autumn, the 7km version out-performed HRES quite dramatically (see Figure 4) This curious behaviour probably warrants further investigation. Firstly, it would of course be helpful to understand the reasons, and secondly because, in several IFS case studies this summer, poor dewpoint forecasts have been implicated in poor forecasts of surface-based convection. So there is evidently potential for ECMWF convective rainfall forecasts to benefit from a better understanding of the accurate dewpoint forecasts in COSMO (7km version).

France and Hungary are consistent in highlighting an under-prediction bias for low level humidity in HRES, of order 5-10%. Hungary also illustrate that there was a systematic downward drift in RH values with forecast lead time, during summer 2015. At 0, 24 and 48h leads the mean bias at stations across the country was -6, -14 and -16% respectively. This is concerning. Meanwhile AROME showed a similar but even more marked downward trend (-1, -15, -18%). For Portugal, AROME had larger biases than HRES during spring.

Although screen-level dewpoint and (equivalently) relative humidity are of less direct relevance to most customers across Europe, compared to say temperature and precipitation, these are very important parameters for ECMWF, because of close links with surface-based convection as highlighted above. Because of this ECMWF has encouraged Member and Co-operating States to provide more feedback on these parameters in reports in future years.

4.7 Visibility

With cycle 41r1 in May 2015 ECMWF introduced in a new visibility diagnostic, which is now also available in ecCharts. Whilst Croatia provided some subjective comments, questioning the validity of values in fog and rain, little other feedback was received. Sweden continue to ‘calculate’ visibility themselves from IFS data, though it is not clear why.

4.8 SST and Sea Ice

Though no routine verification of these parameters has been performed some issues have been recorded. Croatia indicate that there can be problems in the Adriatic, with errors in the ECMWF SST field, which is supposed to replicate the Met Office’s OSTIA analysis, of up to 10˚C compared to station reports. At the same time they report differences between ECMWF and OSTIA. This may possibly be a resolution issue in an area complicated by many islands, but it needs further investigation.

Finland also report issues with SST, near to coastlines, notably just south of Helsinki where it seems that SST in summer 2016 was much too low. It is possible that this relates to interpolation, the way the SST field is stored, and the change in grid structure that came with cycle 41r2, but this needs further investigation. As regards interpolation, ECMWF are working on a new interpolation package.

Finland have also re-stated the problems they have had with spurious sea ice. This has been a recurring issue for ECMWF, and there has been much related correspondence and actions to rectify. The issue relates mainly to the inability of satellite-derived sea ice cover to correctly delineate in areas with a high freshwater content, such as the Baltic, though there are other non-trivial complicating factors, such as summer meltponds, which are made of freshwater but which should be kept as sea ice. The new IFS sea ice model, to come in with cycle 43r1 late in 2016, should help.

4.9 Miscellaneous

Some countries verify or at least comment upon forecaster performance compared to models. Whilst performance gaps are small, there is evidence from Hungary, Turkey and the Czech Republic that forecasters continue to add a little value, on average. For Hungary the added value is primarily in maximum and minimum temperature forecasts, whilst for the Czech Republic it is in sunshine amounts. For modellers such results can provide pointers to where model output can and should be improved.

Meanwhile Hungary were one of the few countries to examine ENS skill alongside that of HRES (ECMWF would keen to see more verification of this type). They show that the ENS mean has more skill, on average, than HRES from about day 4 onwards, though for minimum temperature it looks better even on day 1. Meanwhile they also show that *daily* temperature values in ENS (strictly from just the monthly forecasts) cease to have value from about day 12.

5. Subjective verification

5.1 Subjective conditional verification

Feedback has also been received concerning forecast biases related to specific synoptic situations, most of which were inferred subjectively. This can be called ‘subjective conditional verification’. Caution is needed given that these impressions have not been subject to objective testing, but some of the comments are recorded below.

Secondary frontal waves were reported to not develop enough near to Iceland. In cycle 43r1 ECMWF expects the more extreme examples of this to be *marginally* deeper, on average, due to a reduction in friction over very high waves. Meanwhile Finland and Estonia have also suggested that there may be a slight under-development bias for small low pressure systems, which leads in turn to under-estimation of the related strong winds. Conversely Greece have praised the ability of HRES to predict small-scale cyclones in the Mediterranean.

Different southern European countries have reported that the tracks of slow-moving upper lows can be poor. There are also reportedly problems with precipitation on the northern flank of Mediterranean cyclones. In such setups small changes in synoptic pattern can have a disproportionately large impact on surface weather, so error perception may be partly a reflection of this. Greece report a general overestimation of precipitation totals, notably snow, in the northeast of mainland Greece, and one can imagine that this too could be closely linked to Mediterranean cyclone handling.

Winds and waves during the Indian monsoon, and in the South China Sea, can be underestimated, even at short ranges. These are aspects that ECMWF is less familiar with, so more focussed investigations may be warranted.

“MJO propagation is too slow through the Maritime Continent”. MJO is something that attracts a sharp focus at ECMWF; we continually to try to improve representation. It should also be highlighted that ECMWF handling of the MJO, relative to other global models, is extremely good.

Customer “complaints” about jumpiness, and particularly slave-like behaviour between HRES and ENS, continue on a semi-regular basis. This is something that ECMWF monitors, but at short ranges it is an inevitable consequence of ensemble design, because perturbations tend to keep HRES mid-range relative to ENS.

5.2 Synoptic studies

Several centres report on specific severe weather events that have affected them recently – notably Italy (many cases!), Norway, Croatia, Montenegro, Sweden, Iceland, Germany, UK, Greece and Bulgaria.

Very brief summaries of a few of these are given below but for more details please refer to the Member State/Co-operating State reports, and to the UEF2016 meeting presentations and posters at: http://www.ecmwf.int/en/learning/workshops/using-ecmwf's-forecasts-uef2016.

Orographic rainfall and flooding related to cyclonic storm "Synne" in southwest Norway ("Desmond" in the UK and Ireland), on 4–6 December 2015, was highlighted well in advance in EFI products, although the absolute peak totals were more than predicted, as is invariably the case in orographic rainfall situations. Statistical downscaling can help to overcome biases in such situations - ECMWF is currently investigating this. Snowmelt was an additional factor here; one would not easily recognise this in current IFS products.

Three more flooding events, in Iceland and Montenegro, in spring and summer, exhibited similar under-prediction characteristics due again to an orographic component; in each case broad-scale guidance from IFS was good, it seems.

Italy mainly discuss convective cases, re-advertising the under-prediction "bias" relative to point totals. Some such cases have a strong dynamical component to forcing, and these are generally handled better. They also discuss several recent cases since the resolution upgrade, pointing out that maximum 6h totals in HRES are now more than they used to be. This is as expected. Relative to point totals there was under-prediction in 2 of their 6 tabulated cases. With the old ECMWF HRES model this probably would have been zero. Clearly forecasters need to become accustomed to new model versions, particularly when resolution changes, to avoid incorrectly compensating for bias characteristics that have also changed.

6. Monthly and seasonal forecasts

Demand for, and use of, monthly and seasonal forecasts continues to grow year on year, despite there being no clear perception that these forecasts are getting better. Indeed feedback regarding accuracy is again, understandably, much less positive than it is for forecasts at shorter ranges, with some states again commenting on the large frequency of 'no signal' forecasts. ECMWF is now investigating ways to portray the full ensemble distribution in monthly forecasts, in for example cdf format; potentially this can turn a 'no signal' forecast into one that contains information useful for certain customers, describing the probability of an extreme, for example. Of course one also needs to test how reliable spread signals actually are, so the resource requirement is non-trivial.

11 out of 18 reports contain some reference to use of ECMWF monthly and/or seasonal products, though some of the remaining 7 states may also use them. In many instances a key motivation is demand from the energy sector, for example for hydropower. To satisfy this Switzerland have invested considerable effort in post-processing to extract as much value as possible. At monthly lead times this has been helpful, though for seasonal ranges attempts to predict rarer scenarios, as might be requested

by customers, have yielded lower skill in tests. France also post-process monthly forecasts for energy suppliers, whilst Serbia complement ECMWF guidance with an analogue method. This aside there was very little reference to use of other approaches or products from other centres at either monthly or seasonal leads. Indeed only Romania and Portugal made direct mention of EUROSIP (EUROpean Seasonal and Interannual Prediction – a multi-model system).

Israel have a requirement to predict, in September, the winter-time water level of the Sea of Galilee (effectively a reservoir), in order to minimise desalination costs. They are investing in the downscaling of seasonal forecasts for this purpose.

Croatia, Hungary, Switzerland and France provide some verification data. Forecasts of temperature are always found to be more accurate than forecasts of precipitation. Croatia found it difficult to show much skill at seasonal ranges, though results seemed more positive when employing just a two category system. Switzerland indicate that monthly forecasts are more accurate in autumn and winter than in spring and summer, but in any event skill in precipitation forecasts stops at week 2. Meanwhile Hungary show that there is only skill in month 1 of seasonal precipitation forecasts. For temperature some skill is seen out to month 6, although it is possible that this is a climate change signal. It is also apparent that when verifying and indeed forecasting there is merit in referencing individual members (as UK do), because means are commonly very under-dispersive.

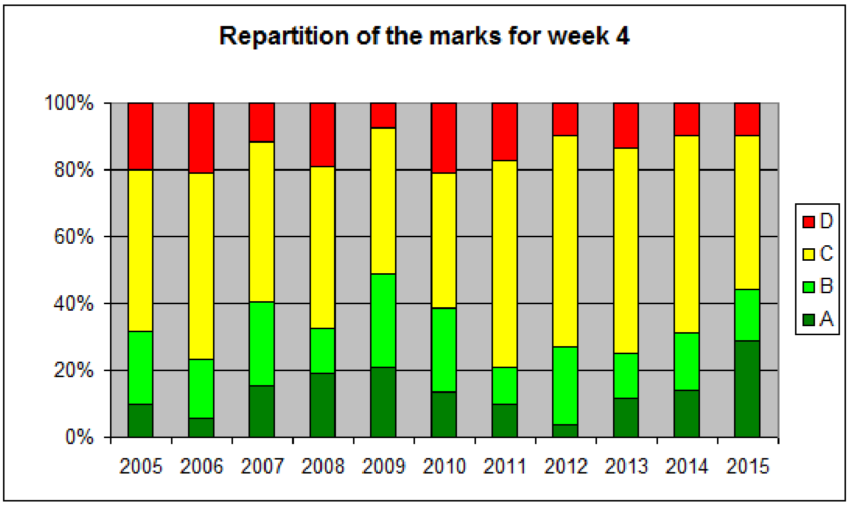

Figure 5: Meteo-France subjective verification for temperature anomalies over France in week 4 of the Thursday monthly forecast. In short: A is good, B OK, C neutral and D opposite occurred.

Since 2005 France have subjectively verified the monthly 2 m temperature anomaly forecast for their country, for weeks 1, 2, 3 and 4. Interestingly week 4 in 2015 brought the highest number of "good" forecasts of all 11 years (Figure 5). It seems that this was largely due to a run of accurate forecasts (of mild) late in the year, which probably related, in turn, to the strong El Nino. So this was perhaps not a sign of a long term trend.

7. Interactive feedback facility for forecast users

In June 2014 ECMWF introduced a new web portal for interaction with its forecast users (https://software.ecmwf.int/wiki/display/FCST/Forecast+User+Home). The portal contains two main sections, one about known forecasting issues and one about severe weather events, which are both regularly updated through the year. In 2016 ECMWF asked whether users were aware of these facilities, and how useful they found them. The responses in the reports from member and co-operating states are summarised in Table 1.

Table 1: Member and co-operating state responses to questions about ECMWF web-based facilities for forecasters.

Facility	No response	"Didn't know about this"	"Not so useful"	"Useful"	"Very useful"
"Known IFS forecast issues"	9	4	0	1	6
"Severe event catalogue"	10	4	1	2	3

Those that "didn't know" were pleasantly surprised to discover the web facilities, suggesting that uptake will increase. The other main message in associated comments was that these pages can be hard to find, and that there is no link on the main ECMWF web page. These issues are being addressed.

References

Haiden, T, Janousek, M, Bidlot, J, Ferranti, L, Prates, F, Vitart, F, Bauer, P and Richardson, DS. 2016. Evaluation of ECMWF forecasts, including the 2016 resolution upgrade. ECMWF Technical memorandum no. 792. Copy here: www.ecmwf.int/en/elibrary/16924-evaluation-ecmwf-forecasts-including-2016-resolution-upgrade.

Seierstad, I, Kristiansen, J and Nipen, T. 2016. Better temperature forecasts along the Norwegian coast. ECMWF Newsletter. No 148. Copy here: http://www.ecmwf.int/sites/default/files/elibrary/2016/16523-newsletter-no148-summer-2016.pdf.