ECMWF Newsletter

Number 144 – Summer 2015

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

www.ecmwf.int/publications/newsletter

CONTENTS

PUBLICATION POLICY

The *ECMWF Newsletter* is published quarterly. Its purpose is to make users of ECMWF products, collaborators with ECMWF and the wider meteorological community aware of new developments at ECMWF and the use that can be made of ECMWF products. Most articles are prepared by staff at ECMWF, but articles are also welcome from people working elsewhere, especially those from Member States and Co-operating States. The *ECMWF Newsletter* is not peer-reviewed.

Editor: Georg Lentze

Typesetting and Graphics: Anabel Bowen with the assistance of Simon Witter.

Any queries about the content or distribution of the *ECMWF Newsletter* should be sent to Georg.Lentze@ecmwf.int

Guidance about submitting an article is available at www.ecmwf.int/en/about/news-centre/media-resources

CONTACTING ECMWF

Shinfield Park, Reading, Berkshire RG2 9AX, UK

Fax:		+44 118 986 9450
Telephone:	National	0118 949 9000
	International	+44 118 949 9000
ECMWF website		www.ecmwf.int

New horizons

In our fortieth anniversary year we have had the opportunity to reflect on ECMWF's past and also to imagine its future. The recent publication of the WMO's book *Seamless Prediction of the Earth System: From Minutes to Months*, which brings together perspectives from the World Weather Open Science Conference in 2014, is another marker of where the science and practice of weather and its prediction is today and what the future might hold. In many respects the title of that book says it all. We are entering an era when, to predict the weather more accurately and reliably, it is necessary and advantageous to take into account many more attributes of the weather and climate system than was imagined back in 1975. For example, the role of the oceans via surface waves and deeper ocean circulation is very important in the development and evolution of weather systems. The properties of the land surface, of sea ice and of atmospheric composition are all now thought to provide potentially important sources of predictability of the weather. Accounting for these components in addition to the more traditional meteorological elements has become known as an Earth system approach.

Compared to 1975, what we aim to predict has changed out of all recognition. Despite the presence of chaos in the dynamical equations, we now believe that there are aspects of the weather – such as heat waves over western Europe – that may be predicted even as far ahead as a month. For large-scale anomalies such as El Niño, there are good indications that predictions may have some skill even out to months ahead. And for high-impact weather events we have the potential of predicting well into the second week ahead. So there is a real prospect of having predictions out to a rather wide range of timescales, and this is why the phrase 'seamless prediction' is coming into vogue.

Of course, another benefit of taking an Earth system approach is that the modelling capability can extend into analysing and predicting a wider range of environmental factors, such as air quality, smoke from fires, and greenhouse gas concentrations. These factors in and of themselves are of vital importance to people in living their lives. This goes far beyond the weather-focused information that has been the bread and butter of meteorological forecasting for many decades.

A final aspect of these trends is that this broadening and deepening of the science is increasingly bringing the weather and climate science communities together. Perhaps it is surprising to those outside the meteorological community that there should ever have been such a distinction to start with. The good news is that it is rightly beginning to disappear because of the growing recognition of the benefits of a seamless Earth system approach. The seasonal prediction problem sits at the interface between weather and climate and it offers great potential to advance our scientific understanding.

Alan Thorpe

Third OpenIFS user meeting held at ECMWF

GLENN CARVER (ECMWF),
PETER DÜBEN (Department of Physics, University of Oxford)

The OpenIFS programme at ECMWF provides a supported version of the operational Integrated Forecasting System (IFS) for education and research that is easy for universities and research organisations to use. Every year, existing, new and potential users of the OpenIFS model get together to present their work, hear about new developments and participate in hands-on exercises. This year, the OpenIFS user meeting was organised by the Department of Physics, University of Oxford, UK, and held at ECMWF. This was the first time the meeting took place at ECMWF and it gave participants an opportunity to visit and see the activities at the Centre and to present their work and speak to ECMWF scientists.

Each OpenIFS user meeting has a scientific theme. This year the theme was 'Uncertainty in Numerical Weather Prediction'. The meeting was opened with an address by ECMWF Director of Research, Professor Erland Källén, in the Council Chamber of ECMWF, the very room in which the initial OpenIFS project was proposed to the ECMWF Council in 2010. Two keynote lectures followed on the role of uncertainty and the ECMWF ensemble system: 'Recognizing the crucial role of uncertainty for developing high-resolution NWP systems' by Professor Tim Palmer (University of Oxford) and 'The ECMWF ensembles' by Dr Roberto Buizza (Head of Predictability, ECMWF). These were followed by two further presentations on stochastic methods in NWP by researchers from the University of Oxford. The use of OpenIFS in teaching courses was described by presenters from the universities of Helsinki and Oslo. The remaining talks illustrated the range of research topics now being studied with OpenIFS, such as the impact of tropical sea-surface temperatures on moisture transport; stratosphere-troposphere interactions; and 'small-planet' simulations.

Storm exercise

The workshop included some practical exercises based on a case study of the St. Jude storm in October 2013 (also called Simone and Christian) (see *Hewson et al.*, 2013, ECMWF Newsletter 139 for a more detailed description). This was an extreme wind storm that hit north-western Europe, killing a number of people. The storm produced the highest ever wind gust recorded in Denmark at 53 m/s. It was an interesting extreme event for a case-study exercise on ensemble prediction. ECMWF forecasts developed the storm too early and too far west, and forecast a too strong wind intensity for western England. However, a signal for the storm was seen in the Extreme Forecast Index (EFI) some five days before.

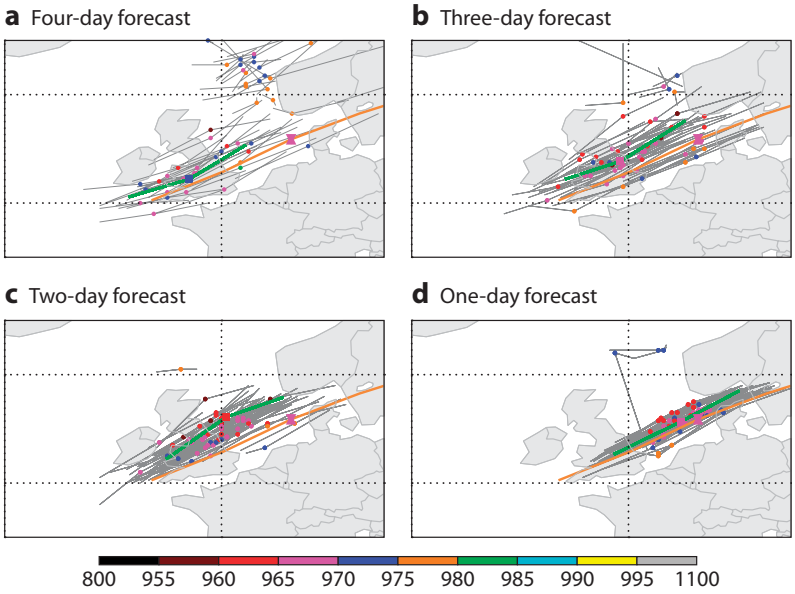
All the exercises took place in the ECMWF classroom and the participants were first given a short briefing on the characteristics of the storm and instructions on how to use the ECMWF Metview application. ECMWF analyses and forecasts of the storm from 24 to 27 October were provided and participants explored the storm development and forecast differences. They were then introduced to the ECMWF ensemble forecasts by visualising root-mean-square-error (RMSE) plumes, spaghetti maps, stamp maps and difference stamp maps and solved a series of tasks and questions designed to improve their understanding of how the ensemble performed; how the ensemble spread developed; and how the forecasts compared to the analysis for this event.

OpenIFS user meeting participants. Participants came from the universities of Oxford, Reading (UK); Kyoto, Tsukuba (Japan); Oslo (Norway); Helsinki (Finland); Twente (Netherlands); L'Aquila (Italy); and Stockholm (Sweden) as well as from Belarusian State University; Lappeenranta University of Technology (Finland); the Hungarian Meteorological Service; the Morocco National Meteorological Service; and ECMWF.

Practical exercises. The St. Jude storm, October 2013, was used as the basis for the exercises. Participants were shown how to work with Metview to visualise the ECMWF analysis and forecasts of this extreme wind-storm event, before exploring the ECMWF operational ensemble forecasts. As a class exercise, participants each ran a single ensemble member forecast using the OpenIFS model on the ECMWF Cray and were then able to analyse the 'classroom ensemble' of the storm.

Graphs of the cumulative distribution function for Reading, UK, were used to illustrate how probabilistic forecasts could be used in a hypothetical example of decision-making for an event at Windsor Castle hosted by the Queen!

The role of initial and model uncertainty was examined using forecasts with the OpenIFS model, where the stochastic parametrizations and variation in initial conditions (provided by the ensemble data assimilation system) were selectively disabled. This illustrated the importance of both types of errors in representing

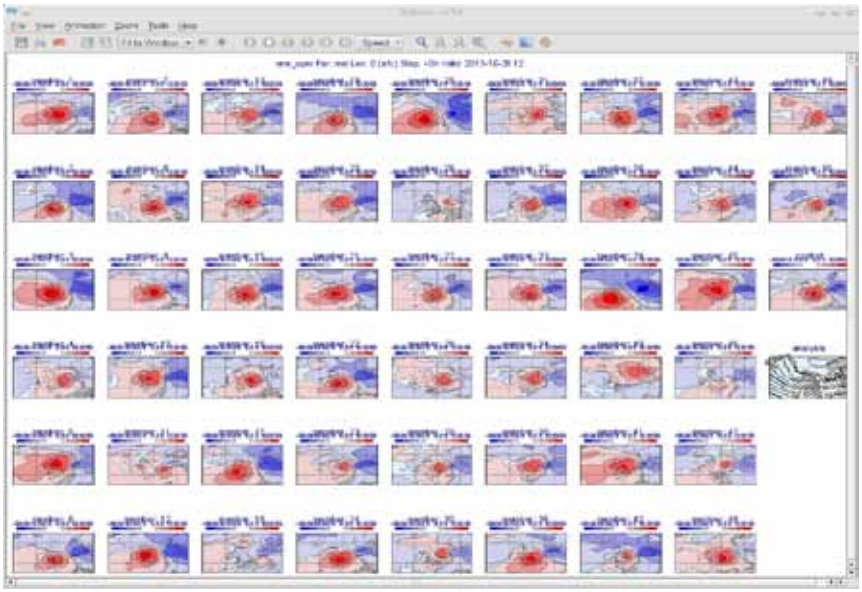


Storm track for different forecast lead times. The figure shows the track of the storm for the ECMWF operational ensemble for each member (grey) and the analysis (orange) for forecasts started from (a) 24, (b) 25, (c) 26, and (d) 27 October 2013 valid for 28 October at 1200 UTC. The symbols represent the position of the storm centre at 1200 UTC on 28 October, and the colour corresponds to the central pressure. The track is shown for 6 hours before and after 1200 UTC. The ensemble control forecast is shown in green with the square symbol. Forecasts from 24, 25 and 26 October illustrate the timing error in the forecasts as the storm was predicted to develop earlier than it did, with an improved forecast on the 27th. The case study was analysed using Metview macros prepared by Sándor Kertész and Linus Magnusson.

forecast uncertainty. Another classroom exercise was dedicated to running the OpenIFS model on the ECMWF Cray supercomputer, where each participant acted as a single 'ensemble member' and produced their own unique 5-day forecast for the storm. As each forecast started from the same initial state, the class ensemble was based on stochastic parametrizations only. These individual forecasts were combined into a single ensemble forecast so participants were able to also study the importance of the ensemble size.

Ideas for the future

The final session was an open discussion between participants and ECMWF staff about the workshop and future plans for OpenIFS. Participants found the workshop informative and interesting, and the opportunity to visit ECMWF and meet ECMWF scientists was appreciated. It was suggested that the user meeting could be held at ECMWF every two to three years though the importance of holding it at other European institutions was stressed. Some participants suggested that, with the growing number of OpenIFS users, the workshop should include an extra day to introduce new users to the models and tools and allow more time for exercises and scientific presentations. The use of e-learning and online videos for new users was also suggested. The growing use of the model for teaching led to discussions on collaboration between universities developing courses and sharing modules, possibly facilitated by ECMWF. The issue of the provision of data assimilation was also raised.



Stamp map of mean-sea-level pressure difference. The figure shows stamp plots of differences between the mean-sea-level-pressure (mslp) field of each ensemble member of the ECMWF operational ensemble and the ECMWF analysis for 28 October 2013 1200 UTC. The forecast was from 26 October 2013. Many ensemble members show the timing error associated with the St. Jude storm indicated by the dipole structure in the mslp difference plot. Only a few members in this case show a reduced error. The participants used visualizations like this to assess the ensemble performance for the different cases described in the text.

The success of the workshop was due to the contribution and support of individuals at the University of Oxford: Peter Düben, Aneesh Subramanian, Hannah Christensen, Peter Watson, Antje Weisheimer and Cathy Morrison, who helped with the organisation and practicals; and at ECMWF: Glenn Carver, Sándor Kertész, Linus Magnusson, Filip Váňa, Iain Russell, Roberto Buizza, Erland Källén, Martin Leutbecher and Karen Clarke, who designed the exercises and OpenIFS practicals and assisted in the organisation. We look forward to the next OpenIFS user meeting in 2016.

More information about OpenIFS can be found at:

http://www.ecmwf.int/en/research/projects/openifs

New model cycle launched in May

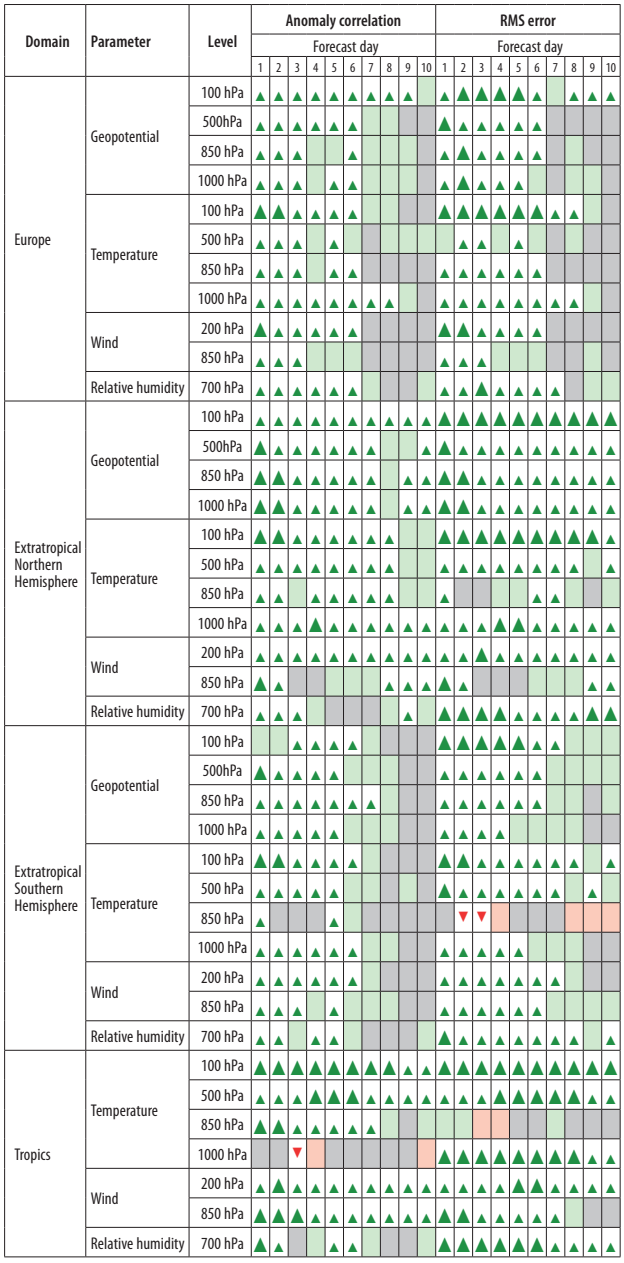
DAVID RICHARDSON,
PETER BAUER

ECMWF launched a new model cycle on 12 May bringing a range of improvements to its Integrated Forecasting System (IFS). Changes in the ways in which observations are assimilated and atmospheric processes are modelled have been shown to improve the representation of the initial state of the atmosphere as well as the skill of forecasts. IFS Cycle 41r1 replaces Cycle 40r1, which was introduced in November 2013.

The new model cycle improves both high-resolution forecasts (HRES) and ensemble forecasts (ENS) throughout the troposphere and in the lower stratosphere. Improvements are seen both in verification against the model analysis and verification against observations.

Cycle 41r1 brings consistent gains in forecast performance at the surface for total cloud cover and precipitation. Improvements in the modelling of cloud and precipitation reduce the predicted occurrence of drizzle in situations where large-scale precipitation dominates, and they increase the amount of rainfall in forecasts of intense events, leading to a better match with observations. Improvements are also seen for 2-metre temperature and 2-metre humidity in parts of the northern hemisphere and the tropics. Cycle 41r1 also introduces a number of new output parameters, such as precipitation type, including freezing rain (see *ECMWF Newsletter No. 141*, pages 15–21).

The average position error for tropical cyclones is slightly reduced, and tropical cyclones are generally forecast to be more intense. For example, IFS Cycle 41r1 performed better than Cycle 40r1 in predicting the track of tropical cyclone Pam, which devastated Vanuatu in the South Pacific in March 2015. In HRES, the sea level pressure minimum at the centre of tropical cyclones is on average slightly lower at all lead times. Up to and including day 3 this makes the forecast better, by reducing the slight positive bias. From day 5 onwards, however, the pre-existing



▲ Cy41r1 better than Cy40r1 – statistically highly significant

▲ Cy41r1 better than Cy40r1 – statistically significant

Cy41r1 better than Cy40r1 – not statistically significant

Little difference between Cy40r1 and Cy41r1

Cy41r1 worse than Cy40r1 – not statistically significant

▼ Cy41r1 worse than Cy40r1 – statistically significant

▼ Cy41r1 worse than Cy40r1 – statistically highly significant

Summary score card for Cycle 41r1. Score card for Cycle 41r1 versus Cycle 40r1 verified by the respective analyses at 00 and 12 UTC for 493 days in the period 2 January 2014 to 10 May 2015. Verification is also carried out against observations, but this is not shown.

bias towards over-deepening has increased slightly.

Cycle 41r1 introduces a lake model, FLAKE, which is applied to all resolved and sub-grid scale lakes. This improves 2-metre temperature forecasts in the vicinity of lakes not represented in the previous model and near coastlines.

Ocean wave forecasts benefit from the extension of the high-resolution wave model from the European and North Atlantic region to the whole of the globe.

New land–sea mask, orography and climate fields (glacier information, surface albedo) have been introduced, as well as new data for lake depth and other lake parameters. The new model also uses new CO_2, O_3 and CH_4 climatologies from the latest MACC-II reanalysis produced at ECMWF.

A revised vertical interpolation in the semi-Lagrangian advection scheme reduces gravity wave noise during sudden stratospheric warming events.

The inner-loop resolutions of the 4DVAR data assimilation system have been upgraded to T255 (80 km) for each of the three iterations of the outer loops to produce finer scale increments. The background error covariances are made more flow-dependent by reducing the sampling window and averaging the statistics over shorter past periods and a climatology. A range of additional satellite observations improves the representation of land surface, sea ice and ocean wave parameters.

Monthly ensemble forecasts and re-forecasts have been extended from 32 to 46 days. The extended forecasts should be used with care but results have shown that there is positive skill in some aspects of forecasts in the monthly to sub-seasonal range. The ENS re-forecast dataset is significantly enhanced, with re-forecasts run twice a week, for Mondays and Thursdays (previously just Thursdays), and with the size of each re-forecast ensemble increased from 5 to 11 members. This provides a substantial increase in the sample size for the model climates for the medium-range EFI/SOT and the extended-range (monthly) products.

The new model cycle is described in greater detail on ECMWF's website: http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/cycle-41r1.

EU approves scalability projects

PETER BAUER, NILS WEDI, TIAGO QUINTINO

The European Commission has given the green light to three international scalability-related projects in which ECMWF participates and which are funded under the EU's Horizon 2020 Research and Innovation programme.

ESCAPE

ESCAPE is a 4m-euro project coordinated by ECMWF to help prepare weather forecasting systems for the exascale era of supercomputing. Earlier this year, the Commission favourably evaluated the proposal. The grant agreement has now been signed and the project will start on 1 October with a two-day kick-off meeting at ECMWF.

The aim of ESCAPE (Energy-efficient Scalable Algorithms for Weather Prediction at Exascale) is to develop world-class, extreme-scale computing capabilities for European operational numerical weather prediction (NWP). It will do this by defining fundamental algorithm building blocks to run the next generation of NWP on energy-efficient, heterogeneous high-performance computing (HPC) architectures.

ECMWF's partners in the project are Danmarks Meteorologiske Institut; Deutscher Wetterdienst; l'Institut Royal Météorologique de Belgique; Météo-France; MeteoSchweiz; Instytut Chemii Bioorganicznej Polskiej Akademii Nauk; Loughborough University; National University of Ireland, Galway; Bull SAS; NVIDIA Corporation; and Optalysys Ltd.

ESiWACE

The European Commission has also approved ESiWACE (Excellence in Simulation of Weather and Climate in Europe), which aims to improve efficiency in computing and storage in the field of weather prediction and climate. ESiWACE will also support the end-to-end workflow of global Earth system modelling for weather and climate simulation in high-performance computing environments.

Partners include the German Climate Computing Center, Deutscher Wetterdienst, the Swedish Meteorological and Hydrological Institute, the UK Met Office, the University of Reading and a number of research organisations, other computing centres and companies.

NextGenIO

Another Horizon 2020 project which has been positively evaluated, NextGenIO (Next Generation I/O), aims to develop new hardware to accelerate I/O processes, using in particular upcoming Non-Volatile RAM (NVRAM) technologies. It includes several vendors such as Intel, Fujitsu and Allinea. The overall objective of the project is to design and prototype a new, scalable, high-performance, energy efficient computing platform designed to address the challenge of delivering the necessary I/O performance to applications at the exascale.

The energy question

NWP has been intimately connected with progress in supercomputing since the first numerical forecast was made about 65 years ago. But supercomputers are power-hungry and, with their current architectures, it will soon be impossible to deliver the required performance at a reasonable cost. Future, more energy-efficient systems with exascale capabilities (performing at least a billion billion calculations per second) will rely on parallel processing at levels to which current NWP codes are not adapted. There is therefore a need to make NWP processes and algorithms scalable so that they will work efficiently on tomorrow's high-performance facilities.

For more details on ECMWF's Scalability Programme, visit: www.ecmwf.int/en/about/what-we-do/scalability

ECMWF forecasts for tropical cyclone Pam

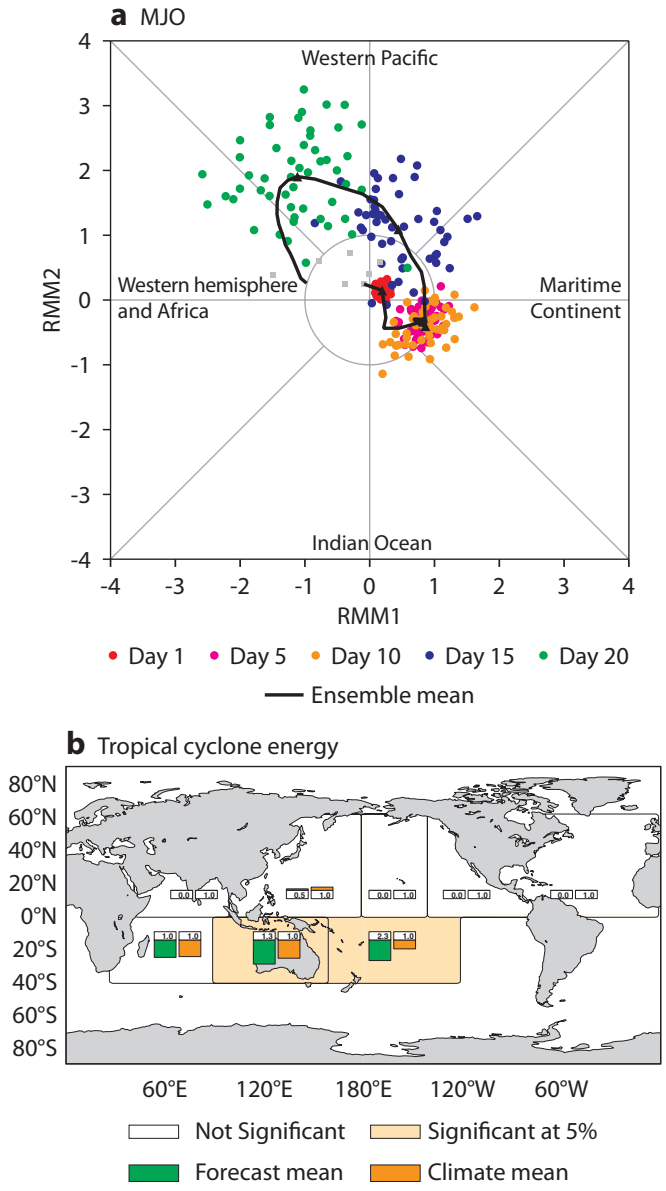
LINUS MAGNUSSON,
SIMON LANG,
FERNANDO PRATES,
FRÉDÉRIC VITART

On 13 March 2015 tropical cyclone Pam hit the islands of Vanuatu in the south Pacific, with devastating effect. Around 15 people were killed and many buildings were destroyed. The cyclone was the second strongest on record in the southern Pacific, second only to Zoe (2002). It is regarded as the worst natural disaster in Vanuatu's history.

The cyclone formed on 6 March east of the Solomon Islands and was classified as a tropical storm on 9 March. The formation is believed to have been influenced by a strong Madden-Julian Oscillation (MJO) event and a connected westerly wind burst in the western tropical Pacific. During the lifetime of Pam, three more cyclones formed in the western Pacific and the south-eastern Indian Ocean: Olwyn west of Australia, Nathan north-east of Australia, and Bavi north of the equator. The MJO event itself was one of the strongest in recent decades.

Tropical cyclones are more likely to occur during the 'active' phases of the MJO, when the frequency and intensity of deep convection are enhanced compared to the climatological mean. The impact of the MJO on tropical storms can be explained by the MJO changing the environment in multiple ways. For example, in an active MJO event relative humidity increases, and vorticity is increased by the westerly wind burst. Since the MJO is to some extent predictable, its modulation of tropical cyclone activity makes it possible to forecast changes in that activity on the intraseasonal (monthly) timescale.

In this case, the MJO and the westerly wind burst were predicted by the monthly forecast more than two weeks in advance. Several ensemble members in the MJO forecast from 26 February showed a strong event in the western Pacific 15 to 20 days into the forecast. The forecast of normalised tropical cyclone energy from 26 February for the week 9 to 15 March predicted cyclone energy levels 2.3 times the normal value in the south-western Pacific and 1.3 times the normal value north-west of Australia. In subsequent forecasts, the ensemble mean cyclone energy increased as the week during which Pam struck approached (to a factor of 3.5 on 2 March, 4.0 on 5 March, and 5.0 on 9 March for the south-western Pacific).



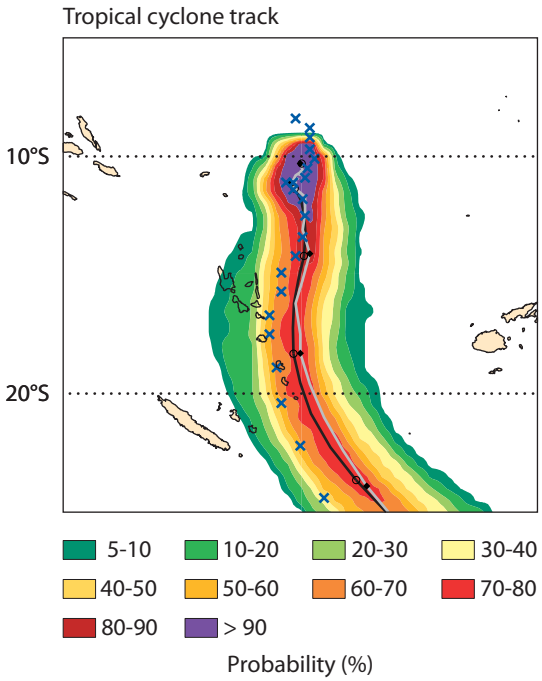
MJO and tropical cyclone energy forecasts. Monthly ensemble forecast from 26 February for (a) the MJO and (b) normalized accumulated tropical cyclone energy for 9–15 March. The dots in (a) show the progression of the active phase of the MJO over a period of 20 days as predicted by different ensemble members. The quadrants indicate the location of the active phase, and the distance of the dots from the centre represents its predicted strength. The solid line, which represents the ensemble mean forecast, suggests that the active phase of the MJO will strengthen as it moves from the maritime continent (comprising Indonesia, the Philippines and Papua New Guinea) into the western Pacific, before subsiding again as it moves over the western hemisphere.

The ensemble (ENS) and high-resolution (HRES) forecasts from 10 March 1200 UTC for the track taken by Pam were somewhat shifted to the east compared to the observed track. Still, a landfall on Vanuatu was inside the ensemble plume.

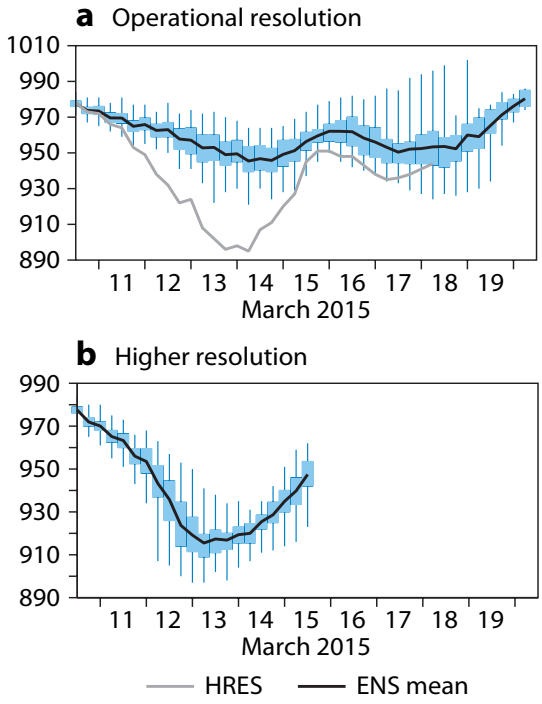
No in situ measurements of the intensity of the cyclone are available, but estimates from satellite images suggest a core pressure in the range of 896–915 hPa on 13 March. The HRES forecast from 10 March 1200 UTC predicted a minimum pressure below 900 hPa while the ensemble median minimum pressure was a much higher 950 hPa.

Tropical cyclone intensity forecasts are sensitive to the spatial resolution of the forecasting system. In the coming year the horizontal resolution of both HRES and ENS will be increased. An early test with the ENS at 17 km resolution for tropical cyclone Pam (the currently operational resolution of the ENS is 32 km) resulted in an increase in the predicted peak intensity of Pam. The median core pressure of the ensemble changes from 950 hPa to 915 hPa for this case. This corresponds quite well to the estimated observed intensity of Pam. The spread of the intensity among the ensemble members is also increased, especially during the phase of rapid intensification of the cyclone.

The example of Pam illustrates how tropical cyclones can be predicted on different timescales. In the extended range (two weeks ahead), increased tropical cyclone activity was forecast over the south-west Pacific in connection with a strong MJO event. For medium-range forecasts, we have given an example of the potential impact of increasing the resolution of the ensemble forecasts. Moreover, pre-operational experimentation with new IFS Cycle 41r1, introduced in May 2015, suggests that forecast performance for tropical cyclones will improve, both for the track and intensity. However, for this type of case, we have to appreciate that a small error in the track may drastically change the impact and that a reliable measure of forecast uncertainty is very important.



Tropical cyclone track forecast. The shaded areas show an estimate of the probability of Pam passing within a 120 km radius over the next 240 hours, starting from its position about 10° South and 170° East at 1200 UTC on 10 March. The crosses show the observed track, the black line represents the ensemble mean forecast and the grey line the high-resolution forecast.



Core pressure forecasts. Forecasts of Pam's central pressure at mean sea level from 10 March 1200 UTC, showing (a) the operational high-resolution forecast (HRES) and the operational ensemble mean forecast (ENS mean) with vertical lines indicating the extreme members and blue bars representing the 25th to 75th percentile of the ensemble distribution, and (b) a higher-resolution (17 km) ensemble mean forecast.

Rescuing satellite data for climate reanalysis

PAUL POLI (ECMWF),
ROGER SAUNDERS (UK Met Office),
DAVID SANTEK (SSEC Univ. of Wisconsin-Madison, USA)

Over the past year, ECMWF has released its first 20th century reanalysis, ERA-20C, developed under the EU project ERA-CLIM. This includes a 3-hourly description of the synoptic atmospheric, land-surface, and ocean wave states for the years 1900 to 2010, as well as monthly means and observation feedback. This reanalysis is not a definitive, final record of the 20th century meteorology and climate. ERA-20C assimilates only surface observations of pressure and marine winds, and not upper-air or satellite observations. Also, ERA-20C does not reanalyse observations of the deep ocean or the soil. Future steps towards a more complete reanalysis capability will be CERA-20C, a coupled atmospheric and ocean reanalysis (an ERA-CLIM2 initiative), and ERA5, due to overtake ERA-Interim and deliver a near-real-time service (supported by the Copernicus Climate Change Service, C3S).

Enriching the digital archives of observations requires historical observations to be recovered from original records held by meteorological organisations, satellite agencies, universities, and libraries. This was a key component of ERA-CLIM. It followed a first attempt in 2009 to produce an inventory of the satellite datasets of potential value to global reanalysis. Since then, several datasets have been located, rescued, or reprocessed. This has been achieved not just thanks to ERA-CLIM but also thanks to individuals, looking for data in their respective institutions, and space agencies, initiating programmes to secure their heritage data, such as the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC).

Why now?

Over the last 50 years, satellite agencies have gathered enormous volumes of environmental data characterizing Essential Climate Variables. The sheer volume of the data meant that at the time much of it was never publicly accessible. In parallel, the underlying media have aged, and so have the mission Principal Investigators and other pioneers of satellite meteorology, now retiring and taking with them sometimes unique and unpublished information.

The geostationary satellite data from NOAA over the US is an example. The University of Wisconsin-Madison Space Science and Engineering Center (SSEC) has recovered historical images from 1979 to 1996 from original tapes to new media, and is now moving to work on data from 1975 to 1979. The intent was to read the Geostationary Operational Environmental Satellite (GOES) image data before the tapes and the expert knowing how to operate the tape reader were both 'decommissioned'. The data now recovered will be available online, something unthinkable when GOES data transmission and archiving systems were first designed.

Such efforts also serve to raise awareness, so that no more old satellite data are lost. Examples of precursor environmental datasets seemingly lost forever include radiometer readings from an ozone spectrometer on Atmosphere Explorer-E (also known as Explorer 55) and from several infrared temperature and humidity sounders (part of the US Defense Meteorological Satellite Program), all in the 1970s and early 1980s.

Space Science and Engineering Center (SSEC), University of Wisconsin-Madison

Old archiving systems. A 9-track tape and 9-track drive used to read the 1975–1979 tapes from NASA.

What is reanalysis?

Climate reanalysis is increasingly seen as critical for understanding the processes associated with climate change and informing future climate change scenarios.

Reanalysis combines information from past meteorological observations with modern forecast models, using data assimilation techniques originally developed for numerical weather prediction. This helps improve the forecast model and also enables scientists to build robust climate models that accurately represent the evolution of the atmosphere during the past century, as described by vast numbers of observations available from numerous sources. A climate reanalysis must be continuously updated with new observations in order to monitor key atmospheric parameters that are responsive to climate change. In addition, the entire dataset needs to span at least a few decades because natural variations found in the climate records can hide some of the longer-term movements and trends.

Extracting information from unique data

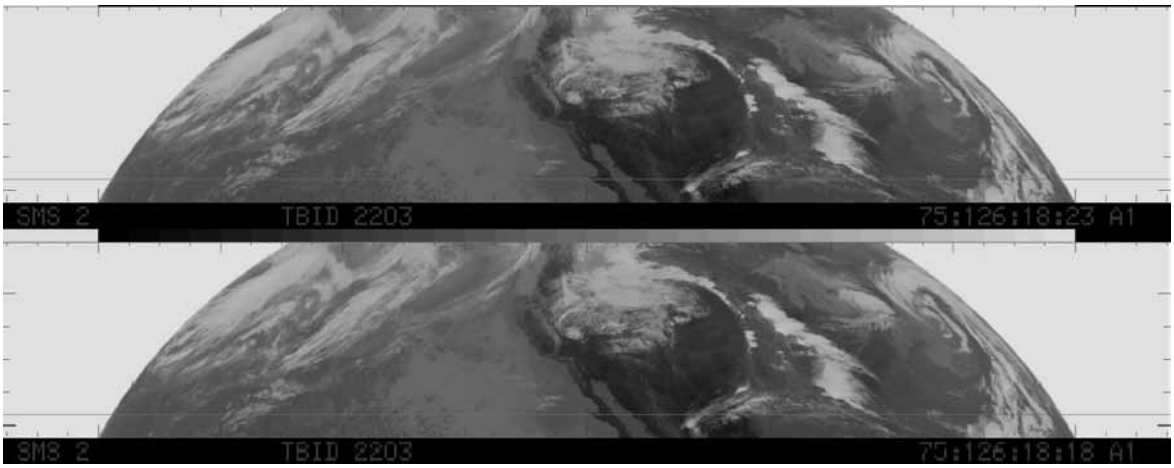
With near-global coverage, satellites often provide information in remote areas, unavailable from any other source. The first interferometer in space was the IRIS instrument built by NASA. On board Nimbus-4, IRIS collected thousands of spectra emitted by the Earth's atmosphere and surface in 1970/1971. These data, now available from NASA, are truly unique. For example, they will help refine polar stratospheric ozone estimates during the polar night in 1970, because the only other ozone satellite record at the time is daylight-only (measuring back-scattered ultraviolet).

The time for exploiting this early satellite data has now come. We have powerful scientific methods to work with such data, such as fast radiative transfer models. Also, thanks to increased computing capacity, such datasets can now be processed or compared with model simulations in a few hours. ECMWF's data assimilation system and models, as well as computing and archiving facilities, make it well placed to extract information in this manner. Key steps include adding ancillary information from various reanalyses or climate model integrations, running instrument simulations, carrying out quality control using state-of-the-art understanding of the observation physics, and eventually assimilating the data in climate reanalysis.

Data from the 1970s can extend the current satellite time series back another decade. Through data assimilation, these data can help enhance the value of reanalyses, by reducing the uncertainties where previously in situ was the only source of information, such as over southern oceans in the 1970s. However, for a single satellite mission to make a difference typically requires a long observation record, something that rarely characterized old missions.

Thankfully, reanalysis can benefit in more than one way from satellite records, such as those from Nimbus-4 IRIS. That mission lasted about 10 months, making it too short to make, by itself, a substantial difference in a 100-year reanalysis. Nevertheless, piecing together the various missions will help patch the record. Also, an immediate benefit comes from the high quality and high spectral resolution of Nimbus-4 IRIS data. The reference points they provide can help quantify uncertainties in reanalyses, which is useful to users developing downstream applications of reanalysis data. Finally, such findings can help spot deficiencies in models and data assimilation that need addressing for further progress.

Through reanalysis, the rescued satellite data will complement the new Sentinel satellite observations to be collected in the coming years by C3S. Once recovered, the early satellite observations will join the Sentinel and reanalysis data in the Climate Data Store, at the heart of the new C3S services.



Space Science and Engineering Center (SSEC), University of Wisconsin-Madison

Old satellite data. An example of data extracted from the 9-track tapes: a series of SMS-2 satellite images from 6 May 1975.

Over 100 attend NWP training programme

SARAH KEELEY

Another 120 students have completed numerical weather prediction (NWP) training on this year's courses run by the Research Department. Participants are usually researchers at meteorological services as well as academic institutions and private companies. The training courses have been developed to give a solid education in the main principles of current NWP as well as including the latest research being carried out at ECMWF. The ethos for the NWP training programme is that, alongside the teaching, there should be a dialogue between participants and lecturers, allowing the exchange of knowledge between ECMWF and the wider research community.

There are many opportunities for participants to discuss what they are working on both with Centre staff and other training course participants. Questions and discussions are encouraged in the classroom but also in more informal discussions, such as during the poster session and over dinner or a beer! Many participants commented that one of the aspects of the course they found very valuable was the interaction with peers and ECMWF scientists.

ECMWF's numerical weather prediction training programme is divided into modules covering different aspects of NWP. The 'Data assimilation' module considers how to make use of the millions of observations that are made each day of the atmosphere, oceans and land to produce the analysis. Modelling small-scale processes important for weather forecasting, such as clouds, soil processes and radiation, is covered in the 'Parametrization of subgrid physical processes' module. 'Predictability and ocean–atmosphere ensemble forecasting' looks at how we make probabilistic forecasting systems and how we can develop forecasting systems for longer timescales (monthly and seasonal forecasts). 'Advanced numerical methods for Earth-system modelling' explores how we can make the forecast models of the future, making the best use of computing resources to produce more accurate global forecasts at higher resolution.

Each module is made up of traditional classroom lectures as well as practicals and discussion sessions. Teaching sessions are given by experts in the various elements of the model and data assimilation system, with around 60 staff from the Research and Forecast Departments being involved.

NWP training programme session. Lectures are complemented by practicals and discussion sessions.

ECMWF hosts Eumetcal workshop on training

GLENN CARVER (ECMWF),
ANNA GHELLI (ECMWF),
SARAH KEELEY (ECMWF),
ALESSANDRO CHIARIELLO (Eumetcal, Finnish Meteorological Institute)

The 10th Eumetcal workshop was jointly organised by and held at ECMWF in June 2015. The workshop explored how training should advance in the next ten years as weather prediction science, models, technologies and services continue to improve and evolve. Education and training enable staff in hydrometeorological services to gain the skills needed to meet the continuously changing requirements and expectations of customers and stakeholders.

The workshop attracted 70 representatives from European national meteorological services, the WMO, EUMETSAT, the UCAR COMET Program, the Australian Bureau of Meteorology, European universities and ECMWF. Training is a key element of ECMWF's mission, and co-organising the workshop was an opportunity to strengthen links with the Eumetcal programme, Europe's virtual meteorological training organisation.

Participants explored current and future meteorological training requirements. They identified several developments in the provision of weather services to which training will have to respond. These include the increasing use of ensemble prediction products, forecasting extreme events, and nowcasting.

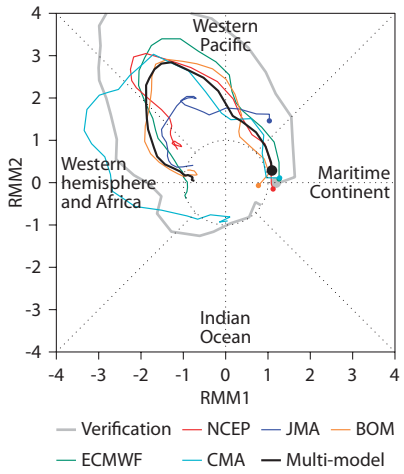
Eumetcal workshop. 70 participants from Europe and beyond attended the event.

New S2S database complements TIGGE archive

FRÉDÉRIC VITART,
MANUEL FUENTES,
LAURA FERRANTI

To bridge the gap between medium-range weather forecasts and seasonal forecasts, in November 2013 the World Weather Research Programme (WWRP) and World Climate Research Programme (WCRP) launched the Sub-seasonal to Seasonal prediction project (S2S). The main goal of this five-year project is to improve forecast skill and understanding of the sub-seasonal to seasonal timescale and to promote its uptake by operational centres and its exploitation by the applications community.

As part of this project, ECMWF has launched a data portal for S2S forecasts. The S2S database was released to the research community on 6 May 2015 following extensive consultations between ECMWF and its S2S partners.



MJO forecasts from five S2S systems. The figure shows MJO forecasts from 5 March 2015 for five forecasting centres, a multi-model forecast (black line) and the verification (grey line). The coloured lines show the progression of the active phase of the MJO over a period of 32 days as predicted by the different models. The position of the lines in the quadrants indicates the location of the active phase, and the distance of the lines from the centre represents its predicted strength. For instance, the line representing ECMWF's mean ensemble forecast suggests that, over the next 32 days, the active phase of the MJO will move from the region of Indonesia to the western Pacific, where it will become very intense before weakening again as it moves over the western hemisphere.

It contains sub-seasonal (up to 60 days) forecasts and reforecasts and thus complements the THORPEX Interactive Grand Global Ensemble (TIGGE) database for medium-range forecasts (up to 15 days) and the Climate-System Historical Forecast Project (CHFP) for seasonal forecasts.

The database will help address important questions for sub-seasonal to seasonal predictability, such as:

- the implementation and benefits of a multi-model approach to sub-seasonal to seasonal prediction
- the sub-seasonal to seasonal predictability of extreme events
- the representation of sources of sub-seasonal to seasonal predictability, such as the Madden-Julian Oscillation (MJO), sudden stratospheric warming, and soil moisture.

Different models

The S2S database includes near-real-time ensemble forecasts and re-forecasts up to 60 days from ECMWF and ten other forecasting centres: Australia's Bureau of Meteorology (BOM); the China Meteorological Administration (CMA); Environment Canada (EC); Italy's Institute of Atmospheric Sciences and Climate (CNR-ISAC); the Hydrometeorological Centre of Russia (HMCR); the Japan Meteorological Agency (JMA); the Korea Meteorological Administration (KMA); Météo-France; the US National Centers for Environmental Prediction (NCEP); and the UK Met Office. These models are generally different from the TIGGE models. Most of them are coupled to an ocean model, and some include an active sea ice model.

Since S2S is a research project, the near-real-time forecasts are available with a 3-week delay. About 80 fields are planned to be archived, including ocean variables, soil moisture and temperature. Pressure level fields are available in the stratosphere at

Some facts and figures

By the end of June, the S2S database had 90 registered users.

They had executed over 20,000 requests and had extracted 16 terabytes of data.

The size of the database was 22 terabytes.

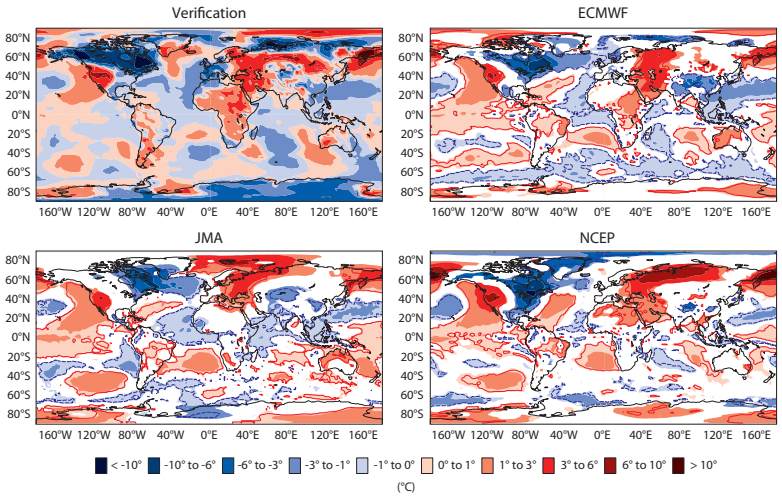
The database can be accessed at:

apps.ecmwf.int/datasets/data/s2s/

A description of the configuration of each S2S model is here: https://software.ecmwf.int/wiki/display/S2S/Models

The S2S prediction project page is available at:

www.s2sprediction.net



Multi-model comparisons. A possible use of the database is to make comparisons between the output of different forecasting centres. The image shows forecasts of 2-metre temperature anomalies from three S2S models and a verification panel based on observations. The forecast starting date is 22 January 2015 and the forecast range is days 12–18.

50 and 10 hPa to allow the diagnostic of sudden stratospheric events and their downward propagation. The data is archived in GRIB2 format, and NetCDF conversion will be made available.

There is much less consistency between the various S2S models than there is between those used in other databases (such as TIGGE, EUROSIP, CHFP). Some sub-seasonal forecasts are produced on a daily basis (e.g. UK Met Office, NCEP), others are produced on a weekly basis (e.g. ECMWF, Environment Canada) and yet others are produced on a monthly basis (Météo-France). The horizontal and vertical resolution of the models and the ensemble size vary greatly from one centre to another. The setup of the re-forecasts also varies. Sometimes re-forecasts are computed all at once and are used to calibrate the real-time forecasts for a number of years, but sometimes (as at ECMWF) they are produced routinely 'on the fly'. Other differences between re-forecast sets include the ensemble size, model resolution, frequency and also the number of years covered by the re-forecasts (from 12 to 30 years). Despite these differences, there are enough commonalities between the models to enable comparisons and multi-model combinations.

Near-real-time forecasts from four data providers (BOM, ECMWF, JMA and NCEP) have been ingested routinely since January 2015. The CMA and Météo-France models have now been added, and all 11 models are intended to be available by the end of 2015.

The Chinese Meteorological Administration (CMA) has expressed an interest in becoming a second archiving centre. ECMWF and CMA will cooperate to ensure the timely synchronisation of both databases.

S2S products

In order to monitor S2S data, a basic set of products, including ensemble mean anomalies for some meteorological parameters and atmospheric indices, has been developed. Products from each individual forecast system and for a multi-model ensemble are produced routinely. Examples of such products are shown in the figures. The multi-model Madden Julian Oscillation (MJO) forecast shown is of a record-strength MJO event that triggered the formation of twin tropical cyclones. In addition, the strong low-level westerly wind anomalies associated with this event are likely to enhance the chances of a strong El Niño developing later this year. Considering the importance of such an event, it is encouraging to see that all models predicted it more than two weeks in advance.

Week of events to explore visualisation in meteorology

STEPHAN SIEMEN

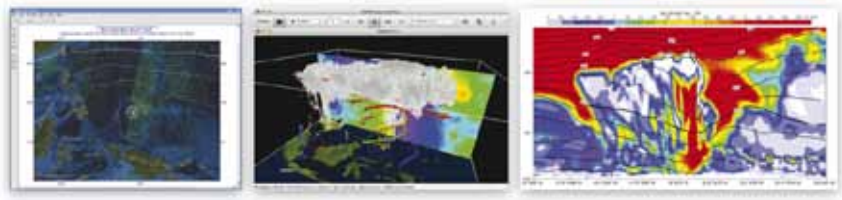
ECMWF will host the 'Visualisation in Meteorology Week' from 28 September to 2 October. Visualisation is important as it can help forecasters to analyse and communicate forecasts.

The visualisation week comprises four meetings: the biennial workshop on 'Meteorological Operational Systems' (MOS); the 'European Working Group on Operational meteorological Workstations' (EGOWS); a Royal Meteorological Society Seminar on 'The visualization of meteorological data'; and a Plugfest of the MetOcean Domain Working Group of the Open Geospatial Consortium (OGC) for web map and coverage services.

These meetings will review state-of-the-art ways of communicating forecasts in operational and research environments in graphical form. The presentations will be accompanied by an exhibition of visualisation systems. We hope to see interesting and novel forms of visualisation (animations, 3D/4D, etc.), especially making use of ensemble forecasts.

The event comes three months after the annual meeting on 'Using ECMWF's Forecasts' (UEF) looked at specific aspects of visualisation, especially how to quantify and communicate uncertainty. Outcomes of this meeting will be presented at the Visualisation in Meteorology Week.

Visualisation in Meteorology Week 28 September – 2 October 2015				
Monday	Tuesday	Wednesday	Thursday	Friday
	09:30 MOS Sessions Big data & Scalability	**09:30 MOS & EGOWS** Visualisation for forecasters & public on the web/desktop	**9:30 EGOWS** Technical challenges in developing forecaster tools	**09:30 OGC plugfest** Demonstration session
11:00 MOS Opening Keynotes		12:30 MOS Closure		12:00 Conclusion & Recommendations
Lunch break (13:00 – 14:00)				
14:00 MOS Session Cloud services & Visualisation	**14:00 MOS & EGOWS** Visualisation in operational meteorology	**14:00 RMetSoc** *The visualisation of meteorological data*	**14:00 EGOWS** Working groups *Challenges we face in developing forecaster systems*	
17:00 Reception	**16:00 MOS & EGOWS** Exhibition of Visualisation systems		15:00 EGOWS closure	
	19:30 Dinner			



Visualising data. State-of-the-art techniques will be showcased and discussed during visualisation week.

For more information on the visualisation week, visit www.ecmwf.int/en/learning/workshops-and-seminars/visualisation-meteorology-week-2015.

ECMWF-run Copernicus services get new websites

DICK DEE, RICHARD ENGELEN, MANUEL FUENTES, VINCENT-HENRI PEUCH, BAUDOUIN RAOULT, JEAN-NOËL THÉPAUT

ECMWF launched new websites on 2 July for the two Copernicus services it is implementing on behalf of the EU. Information and updates on the Copernicus Atmosphere Monitoring Service (CAMS) and the Copernicus Climate Change Service (C3S) can now be found at www.copernicus-atmosphere.eu/ and www.copernicus-climate.eu/, respectively.

Copernicus, the EU's flagship Earth observation programme, addresses six thematic areas: land and marine monitoring, emergency management, and security as well as atmosphere monitoring and climate change. An agreement for CAMS and C3S to be implemented by ECMWF until 2020 was signed on 11 November 2014.

C3S

C3S is currently in a two-year proof-of-concept phase, which will be followed by the pre-operational and operational phases. As part of the preparations for the service, ECMWF has organized and hosted a series of international Copernicus workshops during the first half of 2015. The first of these focused on the web-based Climate Data Store (CDS), which will give users full access to climate observations and derived data products. The results of this meeting, which took place at the Centre from 3 to 6 March, are summarised in the box.

Since then, further workshops on climate projections (20 and 21 April), communicating climate information (16 and 17 June, in Brussels) and climate observation requirements (29 June to 2 July) have taken place. These workshops are proving very useful in helping to tailor the service to the needs of its users.

For example, the climate observation requirements meeting brought together experts from fields including climate policy, water resource management, renewable energy and agriculture

The C3S website provides information and updates relating to the Climate Change Service.

CDS workshop findings

The main findings of the C3S workshop on the Climate Data Store were:

- User engagement is key to building a successful C3S software infrastructure. The C3S portal must provide different 'views' to different users, depending on their level of expertise and domain knowledge. A dedicated call desk will be set up to help users. The portal must allow users to discover, process and view C3S data and products, and to download these to their local systems. Data and products held within the CDS will be freely available without restriction (i.e. Open Data) and supplied with detailed and accurate metadata, which will include data quality information.
- In addition to basic data processing facilities, a collection of more advanced tools should be provided. Users should be able to download these tools, or invoke them from the data portal by selecting from a list of predefined workflows that can be parametrized. For more extensive processing of large amounts of data, cloud infrastructures should be used.
- The C3S software infrastructure must also be interoperable with other Copernicus services (such as CAMS) as well as with other relevant activities, including the Global Earth Observation System of Systems (GEOSS); the WMO Information System (WIS); the Global Framework for Climate Services (GFCS); and the World Climate Research Programme (WCRP). It will apply the relevant international standards to achieve this, such as the Infrastructure for Spatial Information in the European Community (INSPIRE).
- The performance of the C3S software infrastructure should be continually monitored and assessed, and the infrastructure should be designed in such a way that it can respond to the evolving needs of users.

The workshop presentations and proceedings are available at: www.ecmwf.int/en/copernicus-climate-data-store-workshop

applications to ensure the service provides information which is relevant, reliable and accessible.

CAMS

CAMS is making the transition to operations this year with the procurement of the various service elements well under way. The new website has news and event sections and provides instant access to 'today's forecasts' of reactive gases, aerosols, European air quality, the ozone layer and CO_2. It also links to the CAMS catalogue page, which gives access to the full range of available services and products.

The launch of the website coincided with the transition to a new funding regime for CAMS. Much of the research and development work for CAMS has taken place as part of the series of Monitoring Atmospheric Composition and Climate projects (MACC) which had received funding under the EU's FP7 and Horizon 2020 Research and Innovation programmes. On 1 July 2015, the last of the MACC projects ended. CAMS, the operational service that has grown out of MACC, is now fully funded under the Copernicus Earth observation programme.

The CAMS website provides instant access to 'today's forecasts'.

ECMWF makes its mark at geosciences conference

GEORG LENTZE

ECMWF delegates ran a series of well-attended sessions at the European Geosciences Union General Assembly in Vienna from 12 to 17 April and used the opportunity to catch up with the work of fellow scientists. Director of Research Erland Källén opened the numerical weather prediction (NWP) session with a solicited talk on 'Recent progress in global, medium-range numerical weather prediction'. He was joined by Thomas Haiden from the Forecasting Department, who talked about the 'Predictability of cloud fraction in global NWP models'. *"Both talks were well attended and the audience followed them up with good questions. ECMWF was mentioned in several of the talks, both as a benchmark for performance comparisons and as a provider of reanalysis data,"* Erland said.

According to Thomas Haiden, who has attended many EGU General Assemblies since they were launched in the 1980s, *"meteorology and NWP used to be a side show, but this year it was quite clear that this has changed and that the EGU has become an important event for meteorology in Europe."*

Others in ECMWF's dozen-strong delegation noted that ECMWF seems to have become much better-known in the broader geosciences community in recent years. Florian Pappenberger from the Forecasting Department organized sessions on 'ensemble hydro-meteorological forecasting' and on 'large-scale hydrology', which "had people standing as there were no chairs left". He said he was *"particularly pleased about the larger number of times ECMWF was mentioned in hydrology talks and presentations. This is quite different from the situation about 10 years ago, when ECMWF was much less well-known outside the meteorological community."* Emanuel Dutra from the Research Department, who convened the session on 'large-scale hydrology', pointed out that he received a lot of feedback from hydrologist and land surface modellers. *"It was clear that ECMWF has an important role to play in this area,"* he said.

The conference was attended by 11,837 scientists from 108 countries, according to the organizers. ECMWF participants praised the "intelligently organized programme schedule", which makes it easy to identify sessions of interest and to put together a personalised programme. The next EGU General Assembly is scheduled for 17 to 22 April 2016 in Vienna.

Dr Peter Janssen. The EGU General Assembly was a special occasion for the head of ECMWF's Marine Prediction Section, Dr Peter Janssen, who received the Fridtjof Nansen Medal for 2015. This medal was established in recognition of the scientific achievements of Fridtjof Nansen and is awarded annually by the EGU for distinguished research in oceanography.

CERA: A coupled data assimilation system for climate reanalysis

PATRICK LALOYAUX, DICK DEE

Climate reanalysis combines information from past meteorological observations with modern forecast models, using data assimilation techniques originally developed for numerical weather prediction. The resulting reanalysis datasets provide a comprehensive and coherent record of Essential Climate Variables over an extended period of time.

ECMWF recently completed the production of ERA-20C, an atmospheric reanalysis spanning the entire 20th century, based on observations of surface pressure and marine winds (Newsletter 141, p. 9). There are now plans for a new 20th century reanalysis, called CERA-20C, in which ocean and atmospheric observations are assimilated simultaneously into a coupled atmosphere–ocean model.

Interest in coupled data assimilation is growing. Several research groups and national weather services have begun to use coupled models for forecasting and are developing different ways to initialise the forecasts with observations. The innovative approach developed at ECMWF for the purpose of climate reanalysis is at the cutting edge of data assimilation research. As described in this article, the CERA system has been found to lead to a better fit with observations, an improved use of near-surface measurements and smaller initialisation shock effects.

From ERA-20C to CERA-20C

ERA-20C is the first major reanalysis product of an ambitious collaborative research and development programme in climate reanalysis, led by ECMWF and involving many institutions in our Member States and elsewhere. The overall aim is to improve our ability to produce consistent reanalyses of the climate system, reaching back in time as far as possible given the available instrumental record.

The European Commission has provided substantial financial support for climate reanalysis in the form of two successive collaborative research grants, starting with ERA-CLIM in 2011 and followed by ERA-CLIM2 in 2014. An important new goal for ERA-CLIM2 is to develop a second 20th-century reanalysis that uses a fully coupled atmosphere–ocean model. The tentative name for this coupled climate reanalysis, which we hope to complete by the end of 2016, is CERA-20C.

Technical development of the CERA system began in 2013 at ECMWF as part of a two-year project on coupled atmosphere–ocean data assimilation funded by the European Space Agency (ESA). The system is built around the same coupled atmosphere–ocean model used for ECMWF's ensemble forecasts (ENS), i.e. the Integrated Forecasting System (IFS) (atmosphere, land surface, waves) coupled with NEMO (ocean, sea ice). However, the data assimilation scheme implemented in the CERA system is new and fundamentally changes the way observations are used.

Stronger coupling

In the current operational configuration of the ENS, observations for the atmospheric component of the model are assimilated separately from ocean observations. On the one hand, the IFS model is used to produce the analysis for the atmosphere, waves and land. This analysis is commonly referred to as 'weakly coupled' as the IFS model computes model-observation misfits for the three components but relies on separate analysis schemes to correct them. Information from ocean observations can only affect atmospheric state estimates after forward integration of the coupled forecast model. On the other hand, the ocean analysis based on the NEMO model is influenced by the atmospheric analysis, allowing for some interaction between the atmosphere and the ocean. Coupled forecasts are initialised by combining the two analyses even though they may not be fully consistent with each other.

The CERA system introduces a stronger coupling between ocean and atmosphere in the analysis step, so that ocean observations can have a direct impact on the atmospheric analysis and, conversely, atmospheric observations can have an immediate impact on the analysed state of the ocean. Introducing a coupled analysis step makes it possible to obtain more consistent fluxes at the atmosphere–ocean interface and potentially to make better use of near-surface observations and to reduce initialisation shocks in coupled forecasting.

The initial configuration of the CERA system uses a low-resolution version of the coupled model (T159L91 for IFS and ORCA1Z42 for NEMO). The computation of the ocean and atmospheric analyses is based on an incremental (iterative) variational approach with a common 24-hour assimilation window shared by the two components. Figure 1 represents one assimilation cycle of the CERA system, which consists of a pair of nested loops that run the coupled model and compute corrections to the initial coupled model state.

The CERA system uses two outer iterations, which is the default for the uncoupled IFS data assimilation system at low resolution. At the beginning of the first outer iteration, a 24-hour coupled model integration produces a first guess and computes the misfit with ocean and atmospheric observations (outer iteration 1 in Figure 1). The ocean and atmospheric increments are then produced separately by two different inner loops that solve a linearised version of the variational formulation. The computation of the atmospheric increment is based on a 4DVAR approach using the tangent linear and adjoint of a simplified version

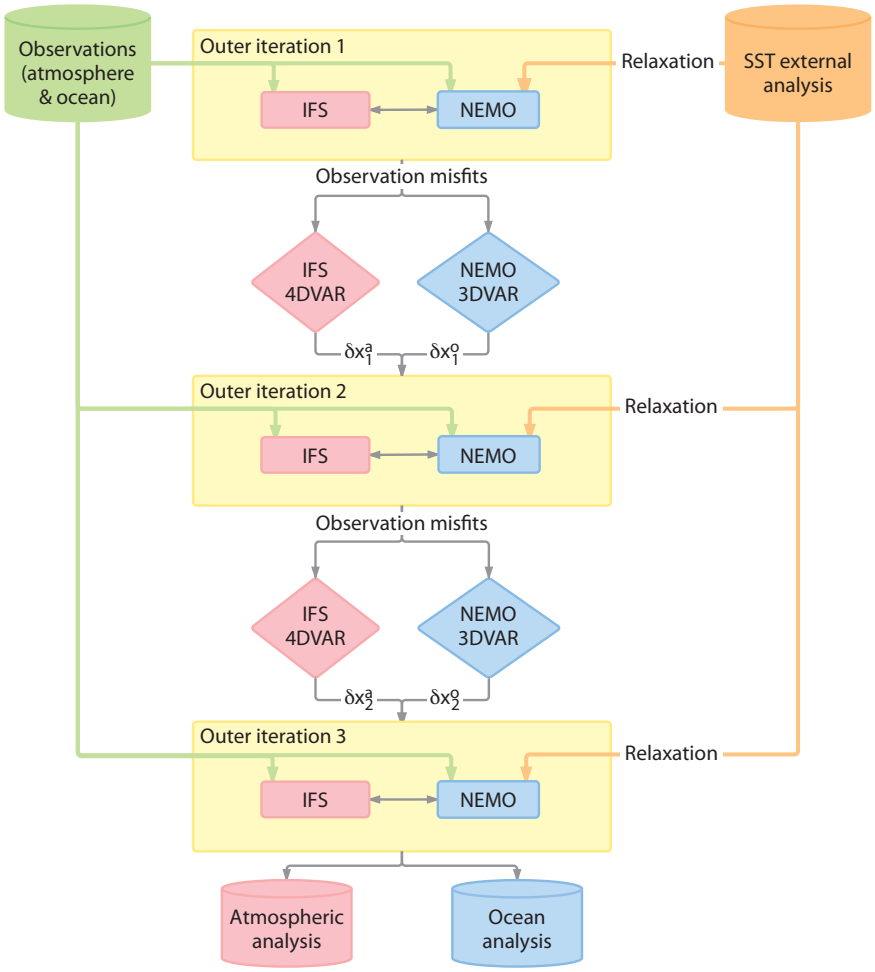


Figure 1 Schematic diagram of the CERA system.

of the atmospheric model at lower horizontal resolution (T95L91). The ocean increment is computed by the NEMOVAR system with its 3DVAR FGAT configuration, which does not require tangent linear or adjoint models. Each method uses its own background error covariance model, which means that correlations between the ocean and the atmosphere are not explicitly represented. The production of the two increments is represented in Figure 1 by the top IFS-4DVAR and NEMO-3DVAR diamonds. A Direct Initialisation (DI) technique updates the current coupled estimate by adding the increments to the initial coupled state defined at the beginning of the assimilation window.

The second outer iteration starts by integrating the coupled model from the new initial condition (outer iteration 2 in Figure 1). During this second 24-hour integration, ocean fields are calculated partly on the basis of fluxes and wind stresses that have been affected by the atmospheric increment computed in the first outer iteration, while atmospheric fields are calculated partly on the basis of sea-surface temperatures (SST) and surface currents that have been altered by the ocean increment. The second atmospheric and ocean increments are computed using the new available observation misfits (bottom IFS-4DVAR and NEMO-3DVAR diamonds in Figure 1) and the initial condition is updated accordingly. Finally, the ocean and atmospheric analyses are produced by the last coupled model integration, which ensures the computation of a dynamically consistent coupled state (outer iteration 3 in Figure 1). Rather than assimilating SST observational data directly, the CERA system uses a gridded SST analysis product during the coupled model integrations to strongly constrain the upper level ocean temperature. In practice this is implemented by a nudging scheme with a timescale of two to three days.

There are two fundamental differences in the design of the CERA system compared to the uncoupled operational ECMWF data assimilation system. The first difference is the use of a coupled model in the variational method and in the forecast that carries the analysis forward in time. The second difference is the treatment of the atmospheric boundary condition, which evolves dynamically in the coupled model by receiving the SST fields computed in the NEMO model through the coupled framework.

Benefits of a coupled analysis

A comparative study has been carried out to assess the differences between the analyses produced by CERA on the one hand and by an uncoupled system (UNCPL) based

on the uncoupled operational ECMWF scheme on the other. To isolate the effects of using the coupled system, UNCPL was run using the same resolution, model cycle and 24-hour assimilation window as CERA, and both systems assimilated the same ocean and atmospheric observational datasets with the same number of outer and inner iterations. Increments were applied by direct initialisation in both experiments. Finally, the SST relaxation performed in the NEMO model of the CERA and UNCPL systems used the same SST analysis product with the same relaxation coefficient. Comparing the behaviour of the two systems serves to highlight some potential benefits of the CERA system: producing a consistent coupled ocean–atmosphere analysis; making better use of near-surface measurements; and reducing initialisation shocks.

Better fit with observations

A first comparative study assesses the differences between the atmospheric analyses produced by the CERA and the UNCPL systems. As differences between both systems are expected to be located in the lowest layers of the troposphere above the ocean–atmosphere interface, the comparison focuses on heights of up to 700 hPa over sea. Figure 2 represents the spatial distribution of the conventional temperature observations assimilated by the CERA system between the Earth's surface and 700 hPa and located over sea or near coasts, for September 2010. The temperature observations are measured by aircraft taking off or landing from airports along coasts and by radiosondes launched from small islands or ships.

Figure 3 shows the background temperature root-mean-square error (RMSE) difference between the CERA and UNCPL systems with respect to the selected conventional temperature observations, for September 2010. Background RMSE is used as an indication of the quality of the analysis on which the background is based. The CERA system used about the same number of observations as the UNCPL system. The observations used to calculate the RMSE shown at 500 hPa include all the observations used above 500 hPa. The negative differences at 1,000 hPa mean that the CERA background RMSE is smaller at this level. The

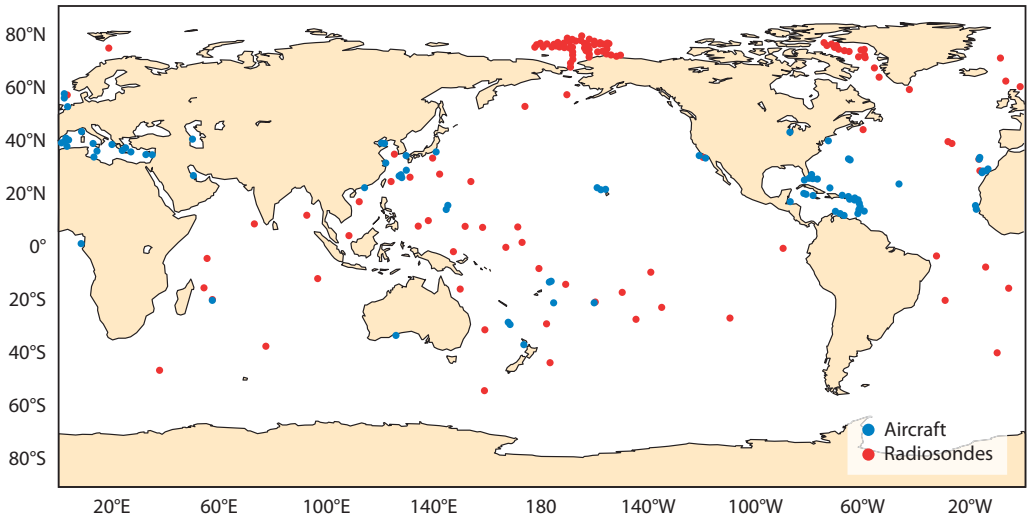


Figure 2 Location of conventional temperature observations by aircraft and radiosondes assimilated by the CERA system between the Earth's surface and 700 hPa over sea and near coasts in September 2010.

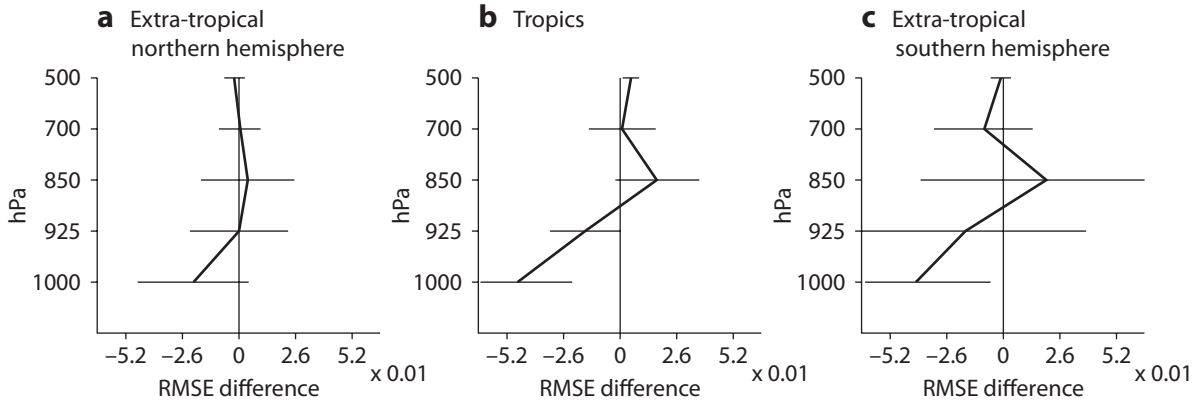


Figure 3 Vertical profiles of the background temperature RMSE difference between the CERA and the UNCPL systems with respect to the selected conventional temperature observations for September 2010. A negative difference means that the CERA background RMSE is smaller. The horizontal lines are error bars representing 95% confidence intervals.

error bars provide an indication of the significance of the results. CERA's better performance could be a result of using a coupled model to compute the first-guess estimate and the background departure vectors. By contrast, the UNCPL system uses persisted SSTs to constrain its atmospheric surface boundary condition in the first-guess computation.

Assimilation of near-surface observations

The CERA system is designed to make better use of near-surface measurements. This is because, in a coupled system such as CERA, any adjustment due to observations near the surface should have an impact on both atmospheric and oceanic variables. To assess the benefits greater coupling brings, we carried out Observing System Experiments (OSEs) in which the CERA and UNCPL systems were run over the period September–November 2013 with and without scatterometer near-surface wind measurements from ASCAT-A, ASCAT-B and OSCAT satellite instruments. In particular, we studied the impact of scatterometer measurements on the quality of the analysis during cyclone Phailin, which formed on 4 October 2013 over the Bay of Bengal and dissipated on 14 October 2013. Tropical cyclones are coupled phenomena with strong interactions between atmospheric wind and ocean temperature and the effect of scatterometer assimilation can be expected to be accentuated during such severe weather events.

The impact of scatterometer data on the analysis of ocean temperature was assessed with respect to conventional observations measured by one Argo buoy. The Argo observing system consists of a fleet of approximately 3,700 drifting probes deployed worldwide. In most cases probes drift at a depth of 1,000 metres and, every 10 days, by changing their buoyancy, dive to a depth of 2,000 metres and then move to the surface while measuring salinity and temperature profiles. The Indian National Centre for Ocean Information Services (INCOIS) operates several Argo probes in the Bay of Bengal. It has initiated a project to monitor upcoming severe weather conditions and to achieve a higher temporal resolution of temperature and salinity profiles for the upper ocean.

Figure 4 shows the observations at a depth of 40 metres from the Argo float 2901335 operated by INCOIS during the passage of cyclone Phailin. As this float was located on the track the cyclone was forecast to take, the probe setup was changed by a satellite transmission on 9 October to measure profiles approximately every 3 hours between the surface and a depth of 300 metres. This configuration was kept in place until 15 October. The strong winds produced over the ocean by the tropical cyclone led to cold water from the deep ocean rising to the surface. This generated a negative SST anomaly in the tropical cyclone's wake. The observed cold wake appears on 11 October with a drop in temperature by 3°C. Figure 4 shows the temperature analyses produced by the CERA and UNCPL systems at a depth of 40 metres with and without input from scatterometer observations.

Figure 4a shows that, in the CERA system, the assimilation of scatterometer data (CERA-SCATT) led to a consistent

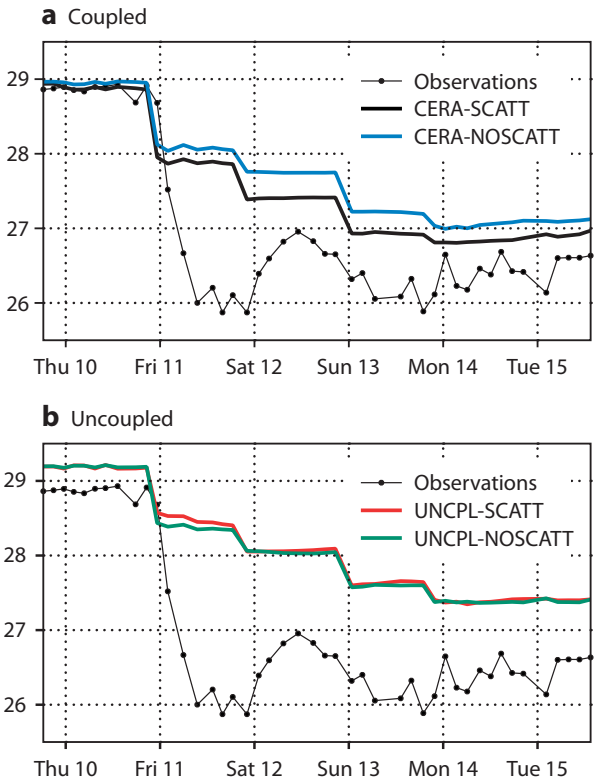


Figure 4 Time series of ocean temperature observations at a depth of 40 metres by the Argo float 2901335(Observations) with (a) the temperature analyses produced by the CERA system with scatterometer data assimilation (CERA-SCATT) and without scatterometer data assimilation (CERA-NOSCATT); and (b) the temperature analyses produced by the UNCPL system with scatterometer data assimilation (UNCPL-SCATT) and without scatterometer data assimilation (UNCPL-NOSCATT).

improvement in the temperature estimate, by up to 0.4°C, compared to the data assimilation without scatterometer data (CERA-NOSCATT), producing an analysis closer to observations. This illustrates that the use of the coupled model in the assimilation process can produce dynamical ocean and atmospheric feedbacks during the assimilation process. The wind observations using scatterometers affected the analysed state of the ocean in a way that was consistent with the phenomenon of a cold wake after the passage of a cyclone.

However, the analysis produced by the CERA-SCATT experiment and the Argo observations are far from a perfect match, with differences of up to 2°C. This can partly be explained by the coarse resolution of the ocean model, which has only 42 vertical levels and in which the thickness of each layer near the surface is around 10 metres. In addition, the OSTIA product used to constrain the SST analysis with a relaxation scheme is a daily-mean product. It produces the plateaus in the ocean analysis time series seen in Figure 4, which explains why the analysis does not follow the frequent fluctuations seen in the observations.

In the UNCPL system (Figure 4b), the time series of the analysis with scatterometer observations (UNCPL-SCATT)

is very similar to that without scatterometer observations (UNCPL-NOSCATT), showing that scatterometer assimilation had no significant impact on the ocean temperature analysis during the cyclone. A plausible explanation lies in the weaker interaction between atmosphere and ocean in the UNCPL system, where forcings come from instantaneous fields retrieved every 6 hours and accumulated fields over 24 hours, compared to the one-hour coupling frequency in the coupled approach.

The comparison of the analyses produced by CERA-SCATT and UNCPL-SCATT illustrates the better performance of the coupled assimilation system during the cyclone. The improvement appears to stem from the better use of scatterometer data as well as a smaller analysis departure before the cyclonic event.

Initialisation shocks

A major challenge of coupled ocean–atmosphere forecasting lies in the initialisation, which aims to incorporate information from ocean and atmospheric observations into the model components in an optimal manner. The initialisation method, particularly for relatively short-range coupled forecasts, should ensure that the ocean and atmospheric model components are consistent with one another at the beginning of the forecast, in order to avoid the generation of initialisation shocks. These shocks are imbalances in the vertical fluxes of heat, momentum or freshwater between the atmosphere and ocean initial states. They can occur when there is insufficient communication between the two model components during the calculation of the initial conditions.

Initialisation shock effects must not be confused with forecast errors due to model biases. The shocks that are discussed here are deviations of the forecast from observations that can demonstrably be reduced or eliminated through changes to the initialisation procedure.

To compare the role of initialisation shocks in CERA and UNCPL, two sets of 30 medium-range coupled forecasts spread over April–May 2008, December–January 2008/9 and August–September 2010 were run initialised by the CERA and the UNCPL analyses, respectively. These CERA and UNCPL forecast sets were run with the same coupled model (versions and resolutions) as used in the computation of the CERA and UNCPL analyses to avoid any effects on the forecast caused by using different models. The left-hand panel of Figure 5 shows the RMSE of CERA's 12-hour air temperature forecast at 1,000 hPa relative to the CERA analysis and averaged over all forecast start dates. Land areas are masked out, as the focus is on atmosphere–ocean imbalances. The errors present in the CERA forecasts are the result of biases in the models as well as any imperfections in the CERA initialisation method. The CERA forecasts are taken as a baseline case in the sense that any larger deviations of UNCPL forecasts from their reference analyses should represent a shock imparted by an initialisation procedure that differs from the CERA methodology.

The right-hand panel of Figure 5 represents the RMSE of UNCPL's 12-hour air temperature forecast at 1,000 hPa relative to the UNCPL analysis and averaged over all forecast start dates. Contours show 0.15°C differences between the RMSEs of CERA and UNCPL forecasts, with blue (green) contours marking increased (decreased) RMSE in UNCPL forecasts. UNCPL forecasts show small increases in RMSE in several areas. These are generally areas in which the difference between the SST in the ocean and atmospheric components of the UNCPL analysis is large.

This air temperature shock signal therefore appears to develop primarily due to the change in SST forcing felt by the atmosphere after the transition from the analysis, where the SST is prescribed by the OSTIA product, to the forecast phase, where the SST comes from the ocean component. These air temperature shocks are generally of magnitude 0.2°C or less, but compared to the small baseline RMSE

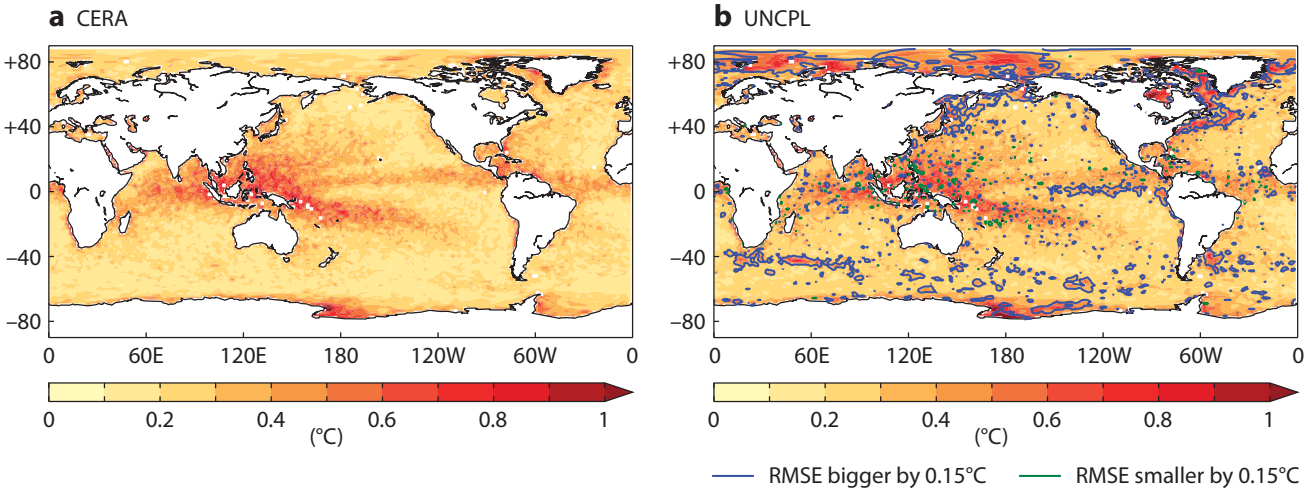
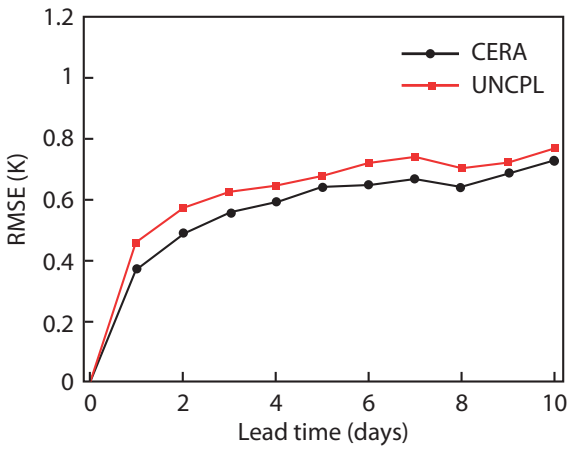


Figure 5 Average RMSE of 30 1,000 hPa temperature forecasts at 12-hour lead times, spread over April–May 2008, December–January 2008/9 and August–September 2010, for (a) CERA and (b) UNCPL , evaluated against their own corresponding analysis. Contours in (b) show 0.15°C RMSE differences between CERA and UNCPL forecasts.



seen in most areas in CERA, they represent substantial error amplifications: the RMSE is increased by 50% or more in the eastern equatorial Pacific, eastern tropical Atlantic, northern Pacific and across most of the Southern Ocean, and it is more than doubled in the Gulf Stream and Arctic regions.

Figure 6 shows the evolution of the RMSE of CERA and UNCPL air temperature forecasts at 1,000 hPa against their own analyses, averaged over the NINO3 region and over all forecast start dates. The larger errors in UNCPL forecasts compared to CERA result from the initial SST discrepancies between the ocean and atmospheric components. The effects of the shock are felt out to at least 10 days' lead time, showing that initialisation shocks can have an impact on short- and medium-range forecasts.

Figure 6 RMSE of thirty 1,000 hPa temperature forecasts spread over April–May 2008, December–January 2008/9 and August–September 2010 and averaged over the NINO3 region, for (a) CERA and (b) UNCPL, evaluated against their own corresponding analysis.

Future plans

The CERA system is now part of the ERA-CLIM2 project, which focuses on the production and assessment of multi-decadal reanalyses of the Earth's climate. The production of an extended climate reanalysis of the 20th century (CERA-20C) will be launched this summer. The use of a coupled model ensures spatial and temporal consistency in the production of long-term climate records.

The CERA system will constrain the coupled model by assimilating only conventional surface pressure and wind observations in the atmosphere, as well as salinity and temperature profiles in the ocean. The air–sea interface will be constrained by an SST relaxation towards the HadISST2 monthly analysis product. Ensemble techniques will be used to compute flow-dependent background error covariances and support uncertainty assessments. All observations used and associated quality feedback information will be provided to users via the ERA-CLIM Observation Feedback Facility. All reanalysis data products will be made available via ECMWF data servers.

The CERA system, originally designed for climate reanalysis, could pave the way for more advanced data assimilation used in weather forecasting. Performance of the coupled system at higher resolution, representation of the diurnal cycle at the air–sea interface and various other technical and scientific challenges will have to be addressed to test this possibility.

Improvements in IFS forecasts of heavy precipitation

RICHARD FORBES, THOMAS HAIDEN, LINUS MAGNUSSON

Good forecast skill for precipitation, and especially for heavy rainfall, is important for many applications. In particular, the accurate prediction of prolonged rainfall events combined with hydrological models can provide early warnings of flooding, which can have significant consequences for peoples' lives and infrastructure. Some of the most extreme precipitation events are associated with orography, where moist air is forced upward, leading to enhanced condensation and precipitation. As the upslope flow can persist over a period of time, considerable precipitation accumulations can occur within the surrounding river catchments, leading to increased levels of river flow and possible flooding both locally and downstream.

Here we evaluate the skill of quantitative precipitation forecasts in the Integrated Forecasting System (IFS) over the last 15 years, with an emphasis on heavy rainfall. The evaluation, which covers both high-resolution forecasts (HRES) and ensemble forecasts (ENS), shows significant improvements in skill over time measured by a number of different metrics and a skill increase equivalent to about one forecast day per decade.

Contributing factors include resolution upgrades and changes in data assimilation, ensemble perturbations, and forecast model numerics and physics, including the representation of cloud and precipitation processes. Recent changes to the cloud and precipitation physics in the latest operational cycle (41r1) have led to further improvements, particularly for high-impact precipitation events associated with orography.

Evolution of precipitation skill over time

ECMWF closely monitors the evolution of HRES and ENS precipitation forecast skill, using a number of different scores. Figure 1a shows the evolution of the headline skill score for precipitation from the HRES over the last 15 years for forecast days 1, 4, 7, and 10. The SEEPS skill

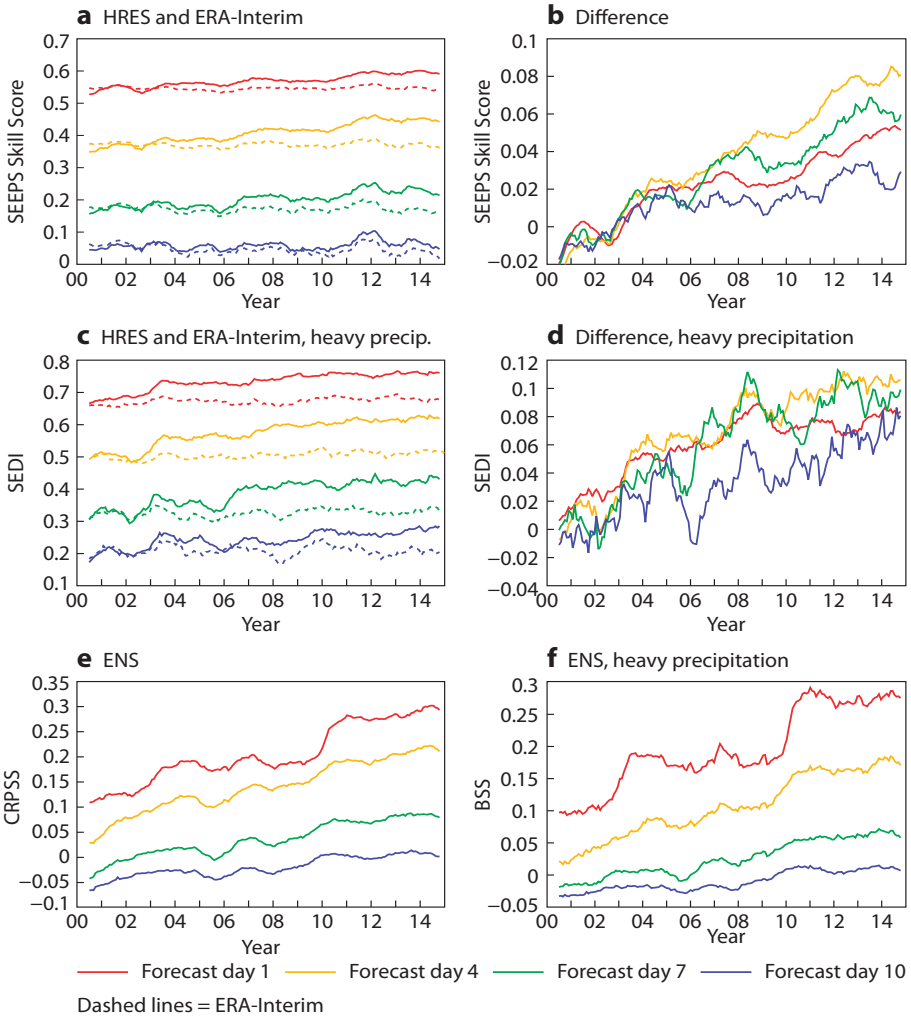


Figure 1 Evolution of extra-tropical 24-hour precipitation skill of the operational IFS over the last 15 years for forecast days 1, 4, 7 and 10 showing (a) the headline SEEPS skill score for the HRES (solid) and ERA-Interim (dashed), and (b) their difference; (c) the SEDI for the HRES (solid) and ERA-Interim (dashed) for 24-hour precipitation accumulations greater than 20 mm, and (d) their difference; (e) the CRPSS for ENS, and (f) the BSS for ENS for 24-hour precipitation accumulations greater than 20 mm. All figures show 12-month running averages of the skill scores.

score measures the ability of the forecast to distinguish between dry days, light precipitation, and moderate-to-heavy precipitation (Box A).

Figure 1a also shows the SEEPS skill score for forecasts from the ERA-Interim reanalysis over the same time period (dashed lines). Whereas the operational forecasting system has evolved over time, the ERA-Interim system is based on a 2006 release of the IFS. This provides a useful baseline, which removes the effect of year-to-year meteorological variations and changes to the observing system, and gives a more direct assessment of the evolution in skill due to changes to the forecasting system used for operations. Note that the resolution of the ERA-Interim forecasts is lower (TL255) than that of the operational HRES system at the time (TL799) and so the cross-over of skill is earlier than 2006.

The improvement over the last decade relative to the ERA-Interim system, shown in Figure 1b, is equivalent to a gain of about 1 forecast day, similar to the improvement in skill for synoptic-scale forecasts, as represented by the 500 hPa geopotential height. Note that jumps in skill due to the implementation of new model cycles are seen in the Figures as gradual increases over a year due to 12-month running averaging. Significant improvements for the operational HRES are associated with many of the IFS cycle changes over the 15-year period, including the resolution increase from TL399 to TL511 in model Cycle 23r3 (November 2000); data assimilation and cloud scheme changes in Cycle 25r3 (January 2003) and 31r1 (September 2006); and modifications during 2010 to 2013 including the change to prognostic rain and snow variables (*Forbes & Tompkins*, 2011) in Cycle 36r4 (November 2010). The magnitude of the SEEPS skill score change in Figure 1b is largest in the medium range for day 4 and then decreases with lead time.

The improvement of forecast skill specifically for heavy precipitation in the operational HRES is shown in Figures 1c and 1d using the Symmetric Extremal Dependence Index (SEDI) for a 24-hour precipitation accumulation threshold of 20 mm. SEDI is designed to be applicable to rare events and is a function of the hit and false alarm rates of a forecasting system (Box A).

Similar to precipitation in general, the skill of heavy precipitation forecasts has increased substantially over the last decade. In the case of heavy precipitation, it is more difficult to attribute periods of stronger improvement to individual model cycle changes because of inter-annual variations in the number of occurrences and smaller sample size. The ranking of improvements relative to ERA-Interim in Figure 1d is the same as for the SEEPS skill score (forecast days 4, 7, 1, 10) in Figure 1b, but the difference between them is smaller in recent years. The relatively big improvement seen in day-10 forecasts relative to ERA-Interim over the last five years suggests that recent developments in the IFS have been especially beneficial for heavy precipitation forecasts at longer lead times. This is an important result since one of the longer-term goals at ECMWF is to extend the range of useful high-impact weather forecasts.

Skill Scores

A

SEEPS

ECMWF developed the Stable Equitable Error in Probability Space (SEEPS) score to monitor the long-term trend in performance for forecasting precipitation (*Rodwell et al.*, 2011). Forecast precipitation accumulated over 24 hours is evaluated against observed precipitation amounts reported from SYNOP stations. At each observation location, the weather is partitioned into three categories: 'dry', 'light precipitation' and 'heavy precipitation'. The boundary between 'light' and 'heavy' is determined by the station climatology so that SEEPS assesses salient features of the local weather and accounts for climate differences between stations. The SEEPS score evaluates the performance of the forecast across all three categories with a value between 0 and 1. As a more accurate forecast gives a lower SEEPS score, it is useful to subtract the score from 1 to give the SEEPS skill score, which has higher values for improved skill.

SEDI

The Symmetric Extremal Dependence Index is a skill score appropriate for extreme events. It provides meaningful results in the case of rare events where the hit rate and false alarm rate decrease towards zero. It is defined for a binary event and thus requires a threshold to be set. Here we use 20 mm, which is a compromise between focussing on the more extreme precipitation events and yet having a large enough sample to reduce the level of noise (due to atmospheric variability) in the resulting scores.

CRPS/CRPSS

The Continuous Ranked Probability Score (CRPS) compares the probability distribution of the quantity forecast by the ensemble forecasting system to the observed value. Both forecasts and observations are expressed by cumulative distribution functions. The Continuous Ranked Probability Skill Score (CRPSS) then compares the CRPS of the forecast to a reference forecast, which in this case is the climatological probability distribution of the quantity.

BSS

The Brier Score (BS) is the most common accuracy measure of probabilistic forecasts of binary events. It is the squared difference between forecast probabilities and corresponding binary observations (0 for non-events, 1 for events). It requires a threshold to be set, and it is converted to a skill score (BSS) by comparing it to a reference forecast, which in this case is the climatological probability of the occurrence of the event.

The evolution of ENS precipitation skill in general is shown in Figure 1e using the Continuous Ranked Probability Skill Score (CRPSS). It measures both the reliability and sharpness of probabilistic forecasts for the whole range of precipitation amounts (Box A). One of the most prominent improvements seen over the last decade can be attributed to the increase in horizontal resolution from TL399 to TL639 in model Cycle 36r1 operational in

January 2010. However, in the very short range (day 1) the use of the ensemble of data assimilations (EDA) for initial perturbations, introduced with Cycle 36r2 in June 2010, has also contributed substantially to the increase in skill due to more spread in the initial conditions.

For heavy precipitation, improvements in the ENS are shown in Figure 1f using the Brier Skill Score (BSS) for a 24-hour accumulation threshold of 20 mm. It shows an evolution of skill over time which is similar to the evolution of the CRPSS score for all precipitation accumulations in Figure 1e, but with a substantial increase in the short range related to Cycle 25r3 data assimilation and ensemble changes in January 2003 and the changes in 2010. The overall similarity with the CRPSS is partly due to the fact that the CRPSS is also sensitive to errors in the magnitude of heavy precipitation.

New cloud and precipitation physics in Cycle 41r1

The latest IFS cycle, 41r1, became operational on 12 May 2015 and contains many modifications to the system, including specific changes to the precipitation physics to reduce the over-prediction of light rain and enhance the heavier rain. New formulations of rain-generation parametrizations were introduced. Changes were also made to the mixed-phase microphysics that increased the growth rate of snow particles falling through supercooled water clouds (Box B).

Precipitation generation is a very non-linear process, and the new physics slows down the initial formation (autoconversion) of rain when the cloud is shallow and liquid water content low, but as the cloud deepens, the amount of rain rapidly grows through the collection of cloud droplets (accretion) (Box B). If the cloud depth extends to temperatures significantly below freezing, ice and mixed-phase processes become important. In this situation, the growth of snow particles through deposition and collection, which subsequently melt before reaching the surface, can significantly enhance rain accumulations.

Evaluation shows improvements in precipitation forecasts due to the cloud physics changes in Cycle 41r1, reducing the occurrence of drizzle from shallow stratiform cloud (*Ahlgrimm & Forbes*, 2014) and increasing the amount of precipitation in forecasts of heavy rainfall (*Haiden et al.*, 2014). Figure 2 shows the SEDI skill score for 24-hour precipitation accumulations greater than thresholds of 20 mm and 50 mm, respectively, calculated over the 7-month period of the 41r1 experimental suite (e-suite) compared to Cycle 40r1, operational at the time. For the 20-mm threshold there is a small increase in skill in the first 5 days of the forecast and an improved frequency bias from 0.88 to 0.91 (a value less than 1.0 means there are fewer occurrences in the forecast than in the observations). For the higher 24-hour accumulation threshold of 50 mm, there is a larger relative increase in the skill further into the forecast, with an improved frequency bias across the forecast range from 0.48 to 0.55.

We expect the frequency of local extreme precipitation totals observed in SYNOP reports to be underestimated by the model due to sub-grid variability, which may for example be linked to unresolved orography. However, accumulations greater than 50 mm are observed at adjacent stations and in radar observations (not shown)

Precipitation physics

B

In the IFS, cloud and precipitation are represented with four separate prognostic variables (cloud liquid, cloud ice, rain and snow) and there are a number of parametrized microphysical processes that lead to precipitation generation and enhancement.

Autoconversion

The autoconversion process represents the collision-coalescence of cloud water droplets within the cloud to create larger droplets that start to fall. This is the process that initiates rain and is parametrized in the model as a rate of conversion from the cloud liquid category to the rain category, dependent only on the amount of cloud liquid water present and the number concentration of cloud droplets (currently a fixed value in the IFS). The more liquid water content there is, the larger the drops and the faster the collision-coalescence process to form rain.

Accretion

The accretion process represents the collection of cloud water droplets by falling rain drops which leads to the enhanced growth of the rain drops. The larger the rain drops, the faster they fall and the more cloud drops they collect, leading to a rapid increase in precipitation, particularly in deeper cloud systems. The parametrization of this process therefore depends on both the cloud liquid water and the rain water content.

Deposition

The deposition process represents the growth of frozen particles from water vapour in ice supersaturated regions. In mixed-phase cloud, the evaporation of supercooled cloud liquid water droplets can keep the air close to water saturation, which therefore maintains ice supersaturation and the continued growth of the ice particles (the Bergeron-Findeisen mechanism). Ice particles are transferred to the snow category as they grow, which can then enhance precipitation at the surface.

Riming

The riming process represents the collection of cloud water droplets by falling snow particles. For temperatures below 0°C, supercooled cloud droplets freeze as they collide with the snow particles, adding to the mass of the falling snow. The parametrized process therefore depends on both the cloud water content and snow water content. If the freezing level is above the surface, then the snow will melt, leading to enhanced rainfall at ground level.

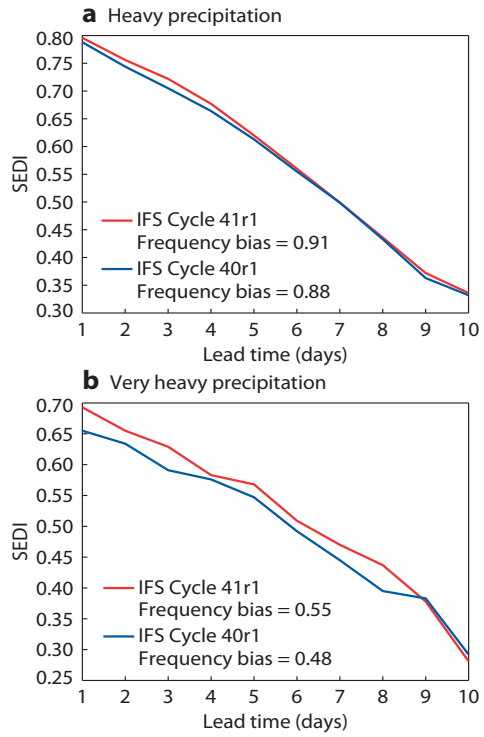


Figure 2 SEDI skill score for 24-hour precipitation accumulations greater than (a) 20 mm and (b) 50 mm for 40r1 (blue) and 41r1 (red) for the extratropics over the 7-month period from October 2014 to April 2015. The frequency bias has been calculated over the whole forecast range.

across scales larger than the model resolution, which suggests that a large part of the underestimated occurrence of high accumulations is the result of a real model error. The improvement in frequency bias is therefore a step in the right direction to improve forecasts of extreme precipitation events.

Impact of Cycle 41r1 on extreme precipitation forecasts

To further investigate the effects of changes in IFS Cycle 41r1 on the more extreme precipitation events, Figure 3 shows the 98th percentile of the 24-hour precipitation accumulation from a 20-year climatology for 40r1 and 41r1 and their difference, for a region of southern Europe centred on the Alps. The precipitation threshold corresponding to a 1-in-50 event in the model climate is based on re-forecast data from days 4 to 7 for the period February to April (for which data is available from the 41r1 e-suite period).

The results show that the predicted magnitude of the heaviest rainfall events has increased (during late winter to early spring) by more than 10% over the southern Alps, along the coast in northern Italy, and over the Balkans, and by over 20% in places. Figure 3d shows the orography for the region to highlight that the areas where the new cycle produces increased precipitation extremes are often associated with steep gradients in orography, where quasi-stationary forcing for deep precipitating cloud systems can lead to large local accumulations.

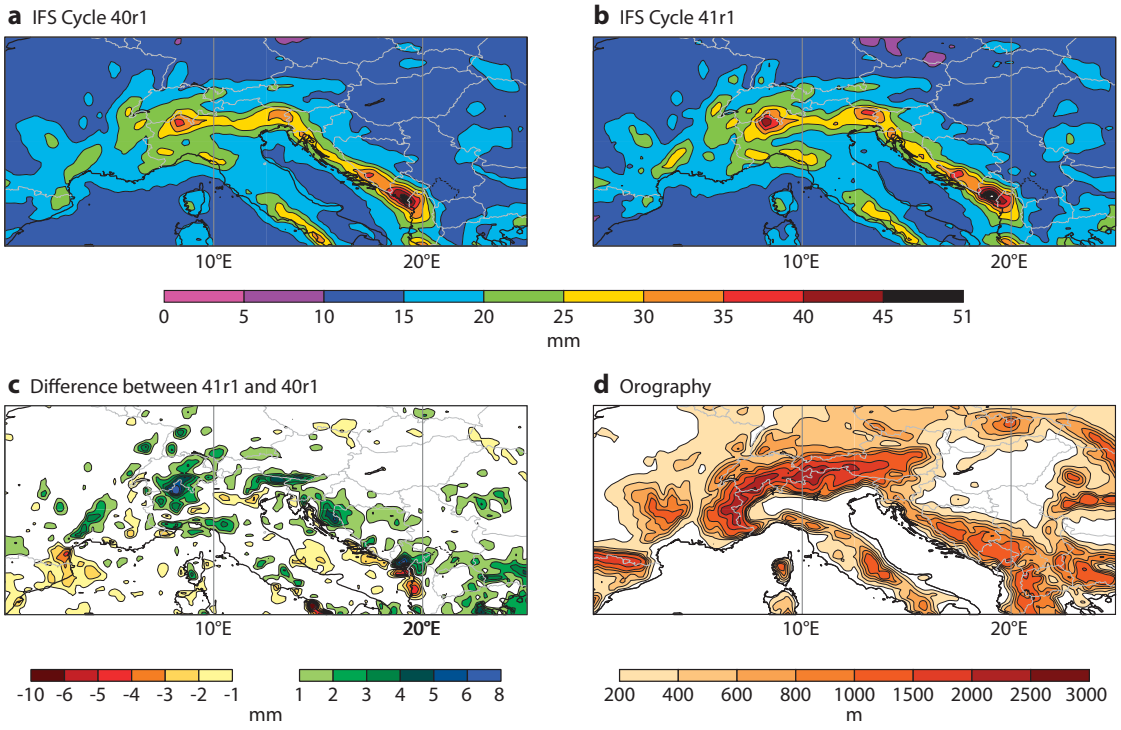


Figure 3 (a) 24-hour accumulated precipitation threshold for a 1-in-50 event (98th percentile) for the IFS Cycle 40r1 climatology over southern Europe, (b) the same as (a) but for the IFS Cycle 41r1 climatology, (c) the difference, Cycle 41r1 minus Cycle 40r1, and (d) the orography. The climatology is based on forecast days 4 to 7 from 4 perturbed ensemble members run once a week for the 3-month period February to April for each year in the past 20 years (4,160 forecasts).

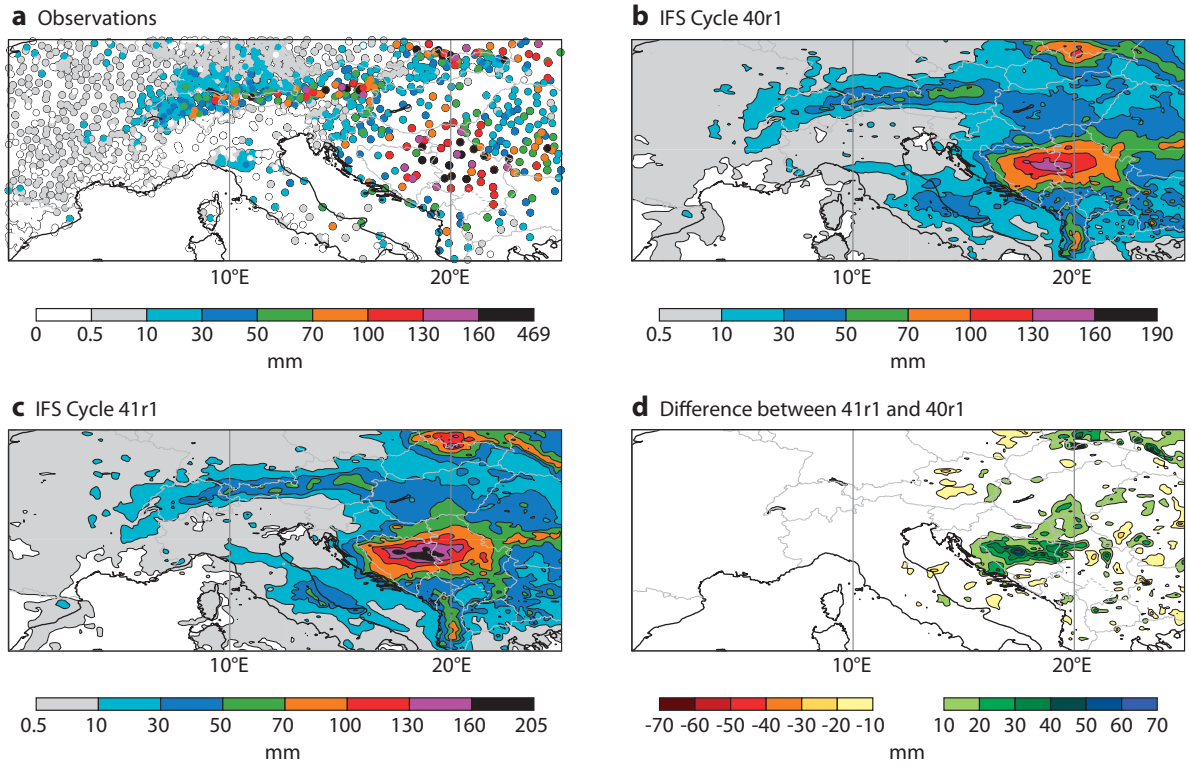


Figure 4 72-hour precipitation accumulation from 0600 UTC on 13 May 2014 according to (a) weather station observations (SYNOP reports), (b) 6 to 78-hour forecast produced by IFS Cycle 40r1 operational at the time, (c) 6 to 78-hour forecast produced by the new operational IFS Cycle 41r1 and (d) the difference, Cycle 41r1 minus Cycle 40r1.

To illustrate the effect of IFS Cycle 41r1 on forecasts of individual cases, Figures 4 and 5 show two extreme precipitation events in southern Europe. The first case occurred in May 2014 when severe floods affected the Balkans (see *Magnusson et al.*, 2014), and the second case is from November 2014 when flash floods affected southern France and northern Italy over a period of three days. In both cases there were fatalities and widespread damage reported from the flood-affected regions. The figures show the precipitation accumulated over 3 days for forecast hours 6 to 78. Figures 4a and 5a show the observed 72-hour precipitation, 4b and 5b show the 40r1 forecast, 4c and 5c show the 41r1 forecast, and 4d and 5d show the difference between the two forecasts.

In the first case study (Figure 4), there is increased precipitation in 41r1 over the central Balkans, where the observed precipitation accumulations are highest, bringing the model in closer agreement with the observations in this region. The maximum precipitation is on the upslope of the orography, with advection mainly from the east to north-east during the period. In the second case (Figure 5), the flow is from the south and the maximum increase in precipitation is along the upslope of the southern Alps and the north-Italian coast. Again 41r1 increases the forecast accumulations where the precipitation is heaviest, in closer agreement with the observations.

Although neither of these case studies is contained within the reforecast climatology period used for Figure 3, the magnitude of the precipitation increases in 41r1 in the two cases is consistent with the increases in the 1-in-50 event climatology based on re-forecasts.

Summary and outlook

Heavy precipitation events can have significant impacts on society and accurate forecasts of the location and magnitude of the precipitation are an important part of severe weather prediction. High-resolution limited-area models provide valuable information on the prediction of more extreme local precipitation from convection and smaller scale orographic features, particularly for the short range. The ECMWF global model plays a complementary role in providing precipitation forecasts for severe events on the larger scale and into the medium-range. Forecasts of precipitation from the IFS, as well as from higher-resolution limited-area models, are used to drive the hydrological model for the operational European Flood Awareness System (EFAS), which produces medium-range probabilistic flood forecasts for Europe. Improving the prediction of heavy precipitation, particularly for extreme events associated with orography, is therefore one of the priorities for improving the reliability of flood forecasting, where early warnings can give additional time for the mitigation of impacts on lives and infrastructure.

In this article, we have looked at the change in the skill of precipitation forecasts in the IFS over the last 15 years. A number of statistical measures all show a significant increase in skill of precipitation accumulations for both HRES and ENS, particularly for the heavier precipitation. The greatest increase in skill for HRES is in the medium range (days 4 to 7),

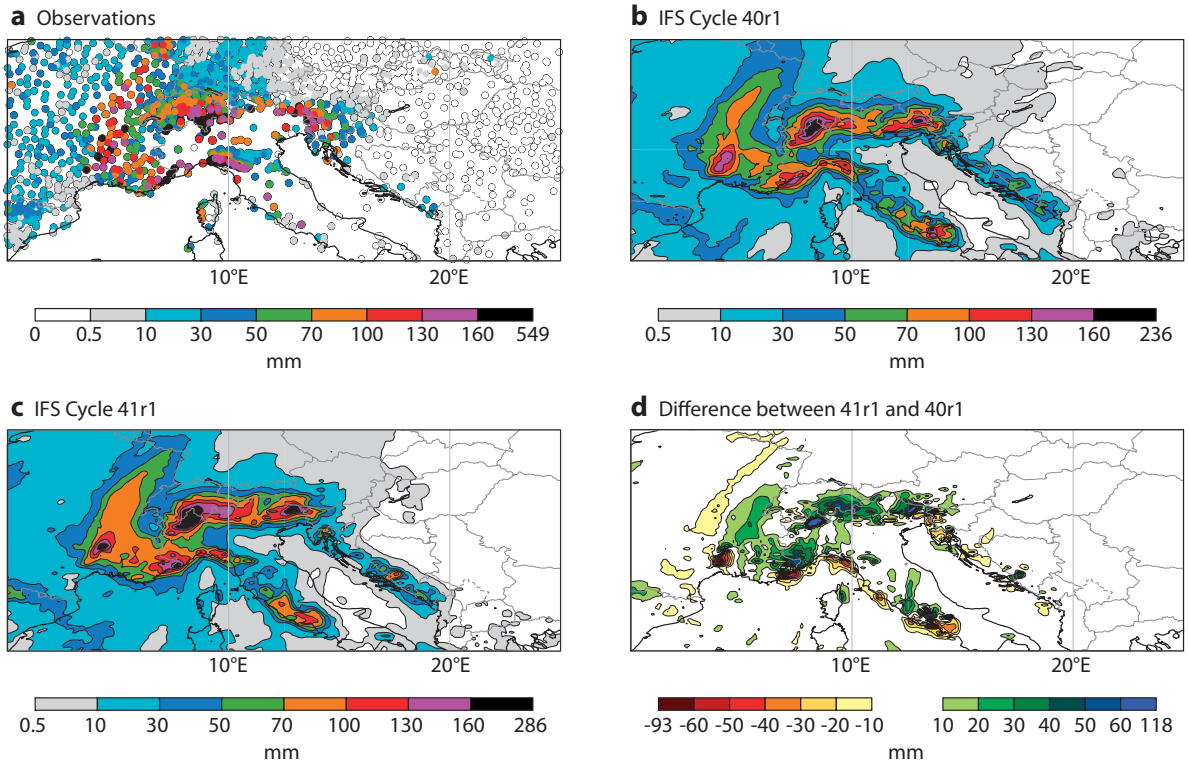


Figure 5 72-hour precipitation accumulation from 0600 UTC on 3 November 2014 according to (a) weather station observations (SYNOP reports), (b) 6–78-hour forecast produced by IFS Cycle 40r1 operational at the time, (c) 6–78-hour forecast produced by the new operational cycle IFS Cy41r1 and (d) the difference, Cycle 41r1 minus Cycle 40r1.

with a rate of skill increase of about 1 day per decade. These increases are a result of improvements in many aspects of the system, including changes to the data assimilation and use of observations, the forecast model, representation of uncertainty and resolution upgrades.

The new physics of warm rain and mixed phase precipitation processes in Cycle 41r1 has decreased the amount of predicted drizzle and increased the amount of predicted heavy precipitation. The biggest impact is in mountainous regions where the increased precipitation in Cycle 41r1 improves the prediction of the more extreme precipitation totals associated with orographic forcing.

Although a lot of progress has been made, work to improve quantitative precipitation forecasting in the IFS is continuing. The skill of precipitation forecasts is very much dependent on the predictive skill of the synoptic-scale forcing as well as the representation of cloud and precipitation physics and the resolution of the model. Further developments in data assimilation and model physics over the coming years should continue to feed into improved forecasts of precipitation.

High precipitation accumulations are usually associated with deep cloud systems extending to low-temperature altitudes. Anticipated improvements in ice and mixed-phase microphysics are expected to lead to further increases in predicted heavy precipitation accumulations, improving the low frequency bias of higher precipitation totals and bringing the model forecasts closer to observations. The grid resolution upgrade to approximately 9 km for HRES and 18 km for ENS planned for early 2016 will lead to an improved representation of the forcing over steep orography, with the potential for positive impacts on the predicted spatial distribution and magnitude of precipitation over complex terrain.

FURTHER READING

Ahlgrimm, **M.** & **R. Forbes**, 2014: Improving the representation of low clouds and drizzle in the ECMWF model based on ARM observations from the Azores. *Mon. Wea. Rev.*, **142**, 668–685.

Forbes, **R.** & **A. Tompkins**, 2011: An improved representation of cloud and precipitation. *ECMWF Newsletter* **No. 129**, 13–18.

Haiden, **T.**, **L. Magnusson**, **I. Tsonevsky**, **F. Wetterhall**, **L. Alfieri**, **F. Pappenberger**, **P. de Rosnay**, **J. Munoz-Sabater**, **G. Balsamo**, **C. Albergel**, **R. Forbes**, **T. Hewson**, **S. Malardel**, & **D. Richardson**, 2014: ECMWF forecast performance during the June 2013 flood in Central Europe. *ECMWF Tech. Memo.*, **723**, 34p.

Magnusson, **L.**, **F. Wetterhall**, **F. Pappenberger** & **I. Tsonevsky**, 2014: Forecasting the severe flooding in the Balkans. *ECMWF Newsletter* **No. 140**, 5–6.

Rodwell, **M. J.**, **T. Haiden** & **D. S. Richardson**, 2011: Developments in precipitation verification. *ECMWF Newsletter* **No. 128**, 12–16.

New EFI parameters for forecasting severe convection

IVAN TSONEVSKY

Forecasting severe convection is a challenging task for meteorological services as its prediction is inherently difficult. It is also a very important task since the impacts of severe convection on society can be substantial. For example, a series of convective storms affected western and central Europe during the first half of June 2014. At least 6 people died in Germany, and the storms are estimated to have caused economic losses totalling more than 2 billion euros in France, Germany and Belgium.

A strategic goal of ECMWF is to provide reliable forecasts of severe weather throughout the medium range (3–10 days) to national meteorological services. To this end, in 2003 ECMWF developed the Extreme Forecast Index (EFI), which provides indications of severe events and is based on the ECMWF ensemble forecast (ENS). The range of parameters to which the index is applied will soon be widened to include dedicated indicators of severe convection. The new EFI parameters have been shown to discriminate well between severe and non-severe convection in the medium range.

How the EFI works

The EFI provides specialised forecast guidance for severe weather events, such as heavy precipitation, strong winds, heavy snowfall, extreme temperatures, and, for the marine community, for unusually high ocean waves. The EFI varies from -1 to 1 and measures the difference between the ENS and model climate (M-climate) distributions. The latter are derived from a set of re-forecasts that comprises ensemble forecasts based on data going back 20 years. For more details see Box A. In addition, ECMWF has designed the Shift of Tails (SOT) index, which complements the EFI by providing information about how extreme an event could potentially be. It specifically compares the tails of the ENS and M-climate distributions (Box A). For the various weather parameters, the EFI and SOT are computed for intervals of various lengths up to day 15.

In the context of recent improvements in parametrizing deep convection in ECMWF's Integrated Forecasting System (IFS) (*Bechtold et al.*, 2014) and of a forthcoming increase in the horizontal resolution of the IFS, ECMWF has tested a number of options for providing guidance on the risk of severe convection. In particular, the aim was to select

Extreme Forecast Index (EFI) and Shift of Tails (SOT)

A

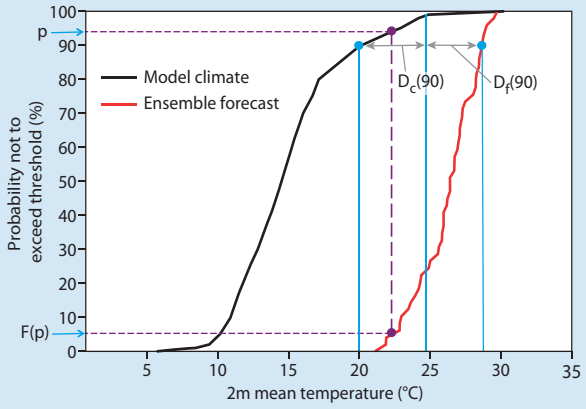
The Extreme Forecast Index (EFI) has been developed at ECMWF to inform users about how extreme an ensemble forecast is, by comparing the forecast distribution to the model climate (M-climate) distribution. It is computed as:

$$\mathrm{EFI} = \frac{2}{\pi}\int_0^1 \frac{p - F(p)}{\sqrt{p(1-p)}}\,dp$$

where $F(p)$ is the proportion of the ensemble members lying below the p-th percentile of the M-climate. Since IFS Cycle 41r1 implemented on 12 May 2015, the M-climate has been determined from an 11-member re-forecast ensemble that is run twice a week, on Mondays and Thursdays, from the same starting date in each of the last 20 years. A set of nine consecutive sets of the output of that re-forecast ensemble, spanning a five-week period, is used to create the M-climate. The middle week of the five is the week closest to the actual forecast run date. Extreme values of the EFI close to -1 or 1 denote a high probability of extreme weather. However, the EFI itself does not show how far beyond the M-climate extremes the ENS solutions go. The Shift of Tails (SOT) complements the EFI by specifically referencing this, comparing the tails of the ENS and M-climate distributions. It is given by:

$$\mathrm{SOT}(90) = -\frac{D_f(90)}{D_c(90)}$$

where $D_f(90)$ is the difference between the 99th M-climate percentile and the 90th percentile of the ENS distribution, and $D_c(90)$ is the difference between the 99th and 90th M-climate percentiles, as shown in the figure. Positive values of the SOT mean that at least 10% of the ENS members are beyond the M-climate extreme (i.e. greater than the 99th M-climate percentile). The bigger the SOT the further away these 10% are from the M-climate. More details about SOT are available on ECMWF's web pages showing EFI forecast charts and in Newsletter No. 133.



suitable parameters to which the EFI and SOT concepts could be usefully applied.

New EFI parameters

Deep moist convection (DMC) is the fundamental breeding ground for severe convective hazards such as hail, extreme rainfall, lightning, tornadoes and severe 'straight-line' winds produced by thunderstorm downdrafts. DMC requires three ingredients: conditional instability, moisture and a source of lift (*Doswell III et al.*, 1996). Convective available potential energy (CAPE) accounts for two of these: instability and moisture. Large CAPE is generally found where low-level moisture combines with steep lapse rates in the lower and middle troposphere. Indeed, CAPE is very sensitive to the temperature and dew point of the parcel that is ascending. When calculating CAPE, it is important to know which parcel is being notionally lifted in the computation. Details of the CAPE parameter that is computed and disseminated by ECMWF are given in Box B.

The likelihood of severe weather and its level of intensity tend to increase with increasing organisation of convection. Supercells are the most prominent example of organised DMC. They are chiefly found in mid-latitudes. They tend to form in the presence of strong vertical wind shear, which can occur even when CAPE is not extremely high. Strong vertical wind shear tilts the storm's updraught, allowing the downdraught and the updraught to occur in separate regions, which leads to long-lived storms. Most occurrences of large hail and tornadoes are associated with supercells.

Individual convective parameters do not discriminate well between severe and non-severe events, whereas considering both instability and shear simultaneously improves the results noticeably. Based on previous studies (*Rasmussen & Blanchard*, 1998; *Craven & Brooks*, 2004) showing that the product of CAPE and vertical wind shear yields better discrimination between severe convection and ordinary thunderstorms, a parameter referred to here as the CAPE-SHEAR Parameter (CSP), has been defined as follows:

$$\mathrm{CSP} = WS_{l_1}^{l_2}\sqrt{CAPE}$$

where WS_{l1}^{l2} is the wind shear between levels l_1 and l_2. The second factor in CSP is proportional to the maximum vertical velocity in convective updrafts, w_{max}, since $w_{max} = \sqrt{2CAPE}$. CSP is thus expressed in units of specific energy (energy per unit mass), m^2/s^2. The first factor is defined as the magnitude of the vector difference between the winds at two different levels. If these levels are far apart, we can refer to this as 'bulk shear'. In operational forecasting, 0–6 km wind shear is usually used. To calculate CSP, two standard pressure levels l_1=*925 hPa* and l_2=*500 hPa* have been selected, for reasons of availability (notably in the ECMWF archive) and close proximity to the levels used in calculating 0–6 km wind shear. CSP can be expected to be strongly correlated with the probability of severe weather related to convection.

CSP is computed for calendar days using the four forecast steps available (6-hour intervals). The maximum of these four CPS values is retained. This is consistent with the EFI computation for other parameters, such as maximum wind gusts. For example, to compute the EFI for a day-2 forecast (T plus 24–48 hours), CSP is computed for T+30h, T+36h, T+42h and T+48h steps and the maximum of these values is used.

The EFI for CAPE is computed in the same way, again retaining the maximum value out of four.

Convective available potential energy (CAPE) and Convective Inhibition (CIN) in the IFS

B

It is important to know what type of CAPE is provided for forecasting severe convection as air parcels at different heights have different CAPE. In the IFS, the most unstable CAPE in the lowest 350 hPa of the atmosphere is computed and provided as a model output parameter. For reasons of numerical efficiency, this CAPE parameter is computed using equivalent potential temperature θ_e instead of virtual temperature:

$$\mathrm{CAPE} = \int_{z_{LFC}}^{z_{EL}} g\left(\frac{\theta_{e,up} - \bar{\theta}_{e,sat}}{\bar{\theta}_{e,sat}}\right) dz$$

where g is the acceleration of gravity, $\theta_{e,up}$ is the updraught equivalent potential temperature, which is conserved during a pseudo-adiabatic ascent, $\bar{\theta}_{e,sat}$ is the environmental saturated equivalent potential temperature, which is a function of the environmental temperature only, z_{LFC} is the level of free convection where the air parcel becomes warmer than its environment and z_{EL} is the equilibrium level where the air parcel becomes colder than its environment.

CAPE is computed for parcels ascending from each model level in the lowest 350 hPa. For parcels in the lowest 30 hPa, mixed layer values of θ_e are used. Finally, it is the maximum value of CAPE from all these parcels that is retained.

By analogy with CAPE, CIN is defined as the negative part of the integral:

$$\mathrm{CIN} = -\int_{z_{dep}}^{z_{LFC}} g\left(\frac{\theta_{e,up} - \bar{\theta}_{e,sat}}{\bar{\theta}_{e,sat}}\right) dz\,;\ (\theta_{e,up} - \bar{\theta}_{e,sat}) < 0$$

where z_{dep} is the departure level from which the ascent starts. The minimum positive value of CIN from all the parcels is retained. This approximation generally overestimates CIN because for simplicity CIN is computed from z_{dep} without considering the lifted condensation level z_{LCL}. It is important to note that computing CAPE and CIN using θ_e instead of virtual temperature overestimates the water vapour contribution to the buoyancy, predominantly from the lower levels.

Verification

The skill of the EFI is normally assessed in terms of 'ROC area' (ROCA – the area under a curve representing the Relative Operating Characteristic). ROCA for the 10-metre mean wind EFI is one of several complementary headline skill scores that ECMWF uses to evaluate the performance of its forecasting system. ROCA values range from 0 to 1. For a skilful forecast, ROCA should be >0.5, and the higher the value the better the ability of the EFI to discriminate between anomalous and non-anomalous weather.

For Europe, EFI forecasts for CAPE and CSP have been verified against lightning data from the UK Met Office ATDnet lightning detection system from 1 April to 31 October 2014. ATDnet detects mainly cloud-to-ground strokes. ATDnet 'fixes' (radio atmospheric signals emitted by lightning and detected by the sensors) that occurred within 0.2° and within 1s have been grouped together as a single flash. This method of converting ATDnet 'fixes' into flashes is consistent with other studies. The data has been gridded onto a regular Gaussian grid N320 with the number of flashes per day represented at each grid point using the nearest-grid-point method. Each daily data file includes all detected lightning between 0300 UTC on the verifying day and 0300 UTC the following day to account in the best way for the validity period of the EFI forecast. The verification domain covers 35°N–65°N and 10°W–35°E. Studies (e.g. *Kaltenböck et al.*, 2009) suggest that higher lightning activity correlates well with the other severe weather events that are of interest here (for which remotely sensed data is not available), and especially with large hail and 'significant' tornadoes (F2 or higher).

The main finding of the EFI evaluation for Europe is that both CSP and CAPE are able to discriminate well between severe and non-severe convection. CSP appears to be particularly good at discriminating between events of different intensity. This is illustrated by Figure 1, where the difference in ROCA for different-intensity lightning events is higher for CSP than for CAPE, by an average of 0.042 over forecast days 1–7. At the same time, ROCA for CSP in cases of high-intensity lightning is higher than ROCA for CAPE, by an average of 0.018 over forecast days 1–7. This suggests that CSP can identify very severe convective hazards better than CAPE.

A dataset of severe weather reports over the USA has been used for an additional assessment of the ability of the new convective EFI parameters to discriminate between severe and non-severe convection. All the reports of tornadoes, large hail (diameter ≥ 2.5 cm) and severe wind gusts (≥ 26 m/s) for the same mid-year period as used for Europe were considered. Severe weather reports were gridded onto a reduced Gaussian grid N320 using the nearest-grid-point method. Only land points (land-sea mask value > 0.5) were considered. By analogy with Europe, each daily data file includes all severe weather reports in the 24-hour period starting at 0300 UTC. The EFIs for both CAPE and CSP show good discrimination of severe convection even in the medium range (Figure 2). Consistent with the results for Europe, the EFI for CSP highlights the 'very' severe

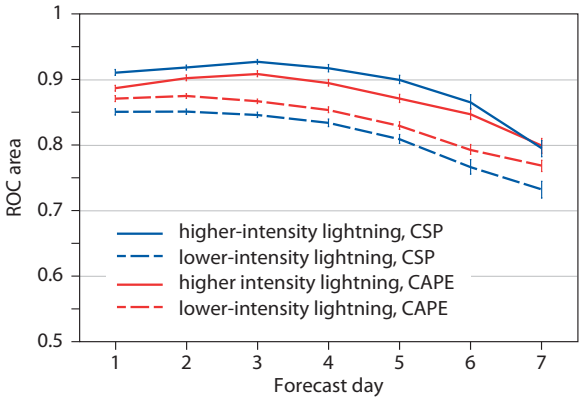


Figure 1 ROC area for the EFI for CSP (blue lines) and CAPE (red lines) over Europe (35°N–65°N, 10°W–35°E) for different intensities of lightning activity, with lower-intensity events corresponding to a frequency of 0.6% in the UK Met Office ATDnet lightning detection system dataset for the period 1 April to 31 October 2014, and higher-intensity events corresponding to a frequency of 0.04% in the same dataset. Error bars show 90% confidence intervals.

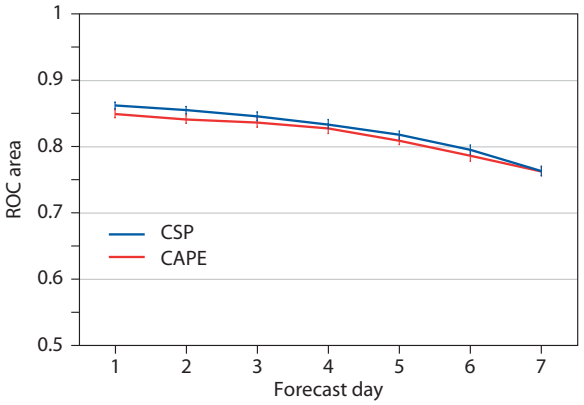


Figure 2 ROC area for the EFI for CSP and for CAPE over the USA. The EFI is verified against severe weather reports from the Storm Prediction Centre (SPC) database for 1 April to 31 October 2014. Error bars show 90% confidence intervals.

convective events better than the EFI for CAPE. On average, the values of the EFI for CSP are higher than the EFI for CAPE for severe thunderstorms producing hail at least 5 cm in diameter and/or wind gusts of 33 m/s or greater, while they are similar for tornadoes (Figure 3). The EFI values are, on average, higher for tornadoes than for hail and severe wind gusts. This is true for both CSP and CAPE.

Case study 1: 9 June 2014

On 9 June 2014, a hot air mass was moving north over Western Europe on the western fringe of a ridge. A quasi-stationary front could be found over the westernmost parts of the continent. Model and actual soundings showed steep lapse rates in the low to mid-troposphere with substantial low-level moisture (Figure 4). As a consequence of this, and owing also to the presence of some convective inhibition (CIN) (note the capping inversion in the lower troposphere on the tephigrams in Figure 4), very large CAPE built up. Model analysis fields suggested that CAPE values exceeded 3,000 J/kg. In addition, extreme CAPE overlapped with substantial vertical wind shear

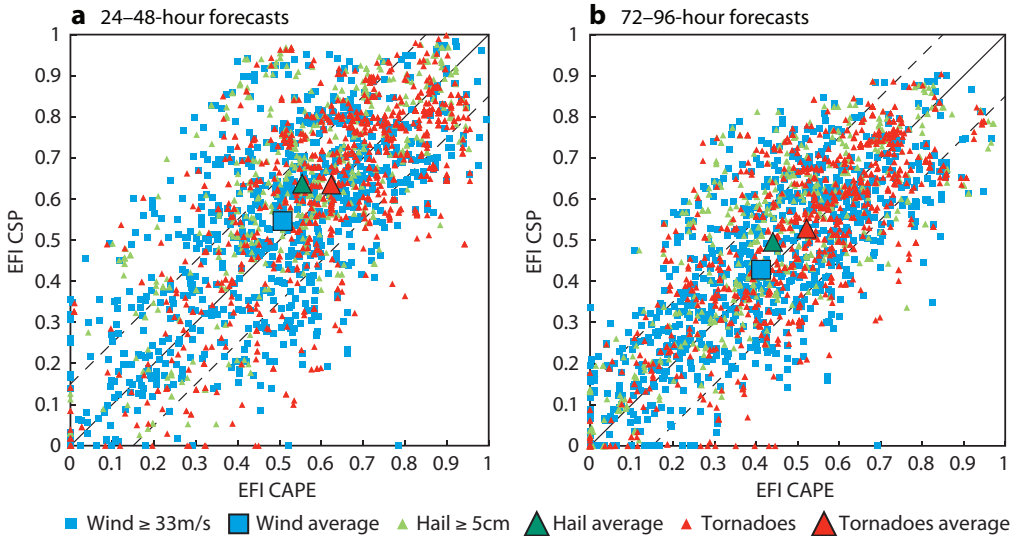


Figure 3 Scatter plot of EFI for CSP versus EFI for CAPE, for 'very' severe thunderstorms, including all reports of tornadoes, hail ≥ 5 cm in diameter and wind gusts ≥ 33 m/s from 1 April to 31 October 2014 over the USA for (a) 24–48-hour forecasts and (b) 72–96-hour forecasts. The EFI is represented by the maximum value within 100 km around each severe weather report. Dashed grey lines represent ±15% deviations from the diagonal as a threshold above which the EFIs differ significantly.

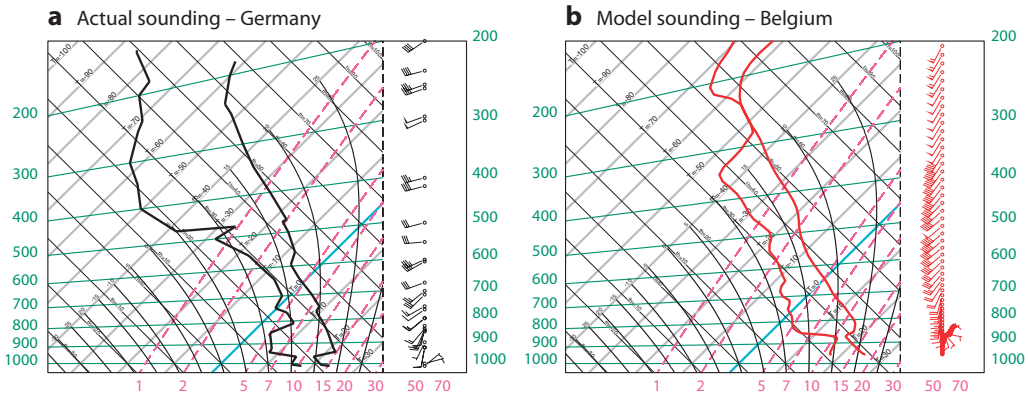


Figure 4 (a) Sounding from Bergen, northern Germany, at 1200 UTC on 9 June 2014, and (b) model tephigram for a location in eastern Belgium where the large values of CAPE were analysed at 1200 UTC on 9 June 2014 and where the severe storm that hit Germany (Figure 5a) originated.

that favoured the organisation of DMC into supercells and mesoscale convective systems (MCSs).

In the event, an outbreak of severe convection affected France, the Benelux countries and western Germany with many reports received, chiefly of strong wind gusts and large hail. A violent supercell developed over western Germany in the late afternoon (Figure 5a). It uprooted many trees and caused the death of six people as well as significant damage and disruption to transport. Gusts of up to 42 m/s were recorded at Düsseldorf Airport (Figure 5c).

The EFI forecast for 10-metre wind gusts (Figure 5c) gave no indication of severe wind gusts even in the short range. The EFI for CSP, however, reached extremely high values, in excess of 0.9, giving an indication of organised DMC and suggesting that supercells might develop that day. This signal appeared in the forecast many days in advance: it is quite uncommon to see EFI values close to 1 in a day-6 forecast (Figure 5d). In addition to severe wind gusts, large hail and damaging lightning were also reported, consistent with the high EFI values. The high values of the EFI for CSP cover the areas of the most intense lightning activity, from southern France through to northern Germany (Figure 5b).

Case study 2: 13 May 2015

The EFI for both CAPE and CSP is supposed to provide indications of potentially anomalous convection and in particular of severe convective outbreaks in mid-latitudes in the warm part of the year. The EFI shows the area where the given convective parameter is anomalous compared to the model climatology. On many occasions this area is much bigger than the area where severe convection actually occurs. In this sense the new EFI parameters show where the convection could be severe if it is initiated.

The example below aims to suggest to users how the area of severe convection might be specified more precisely

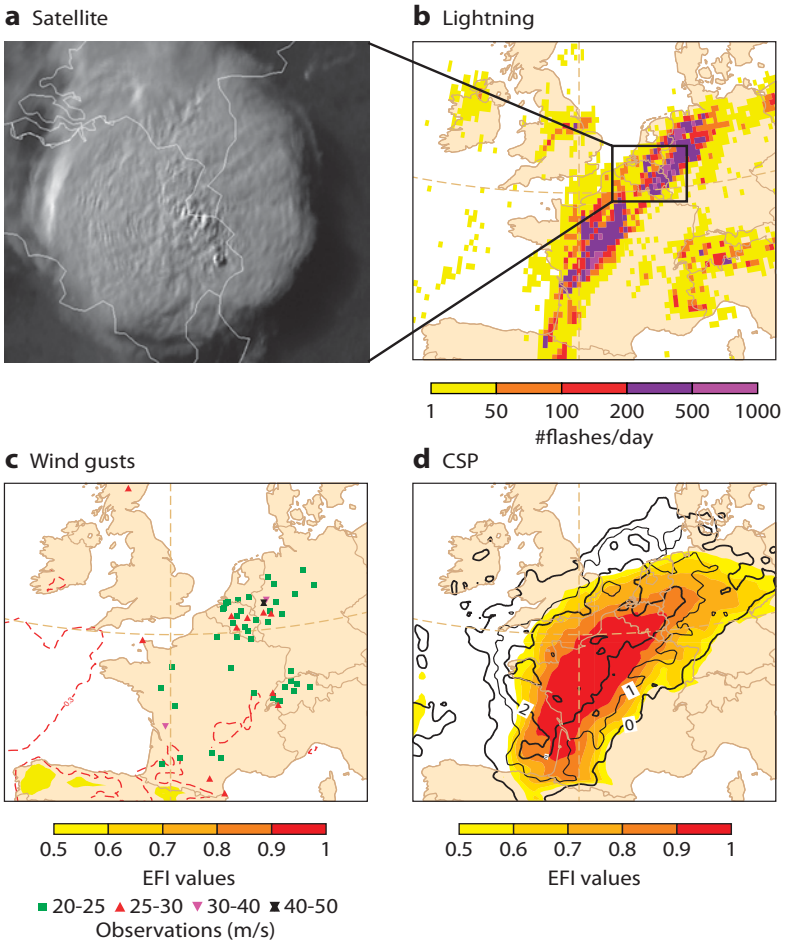


Figure 5 (a) Meteosat-10 HRV (High Resolution Visible) imagery at 1730 UTC on 9 June 2014 (source: EUMETSAT); (b) lightning activity (total number of flashes) on 9 June 2014 (source: UK Met Office ATDnet lightning database); (c) EFI for 10-metre wind gusts, T+0–24h forecast and reported maximum wind gusts (in m/s) both valid for 9 June 2014; and (d) EFI (shading) and SOT (contours) for CSP, T+120–144h forecast valid for 9 June 2014.

several days in advance, using the EFI and other forecast fields. The day-4 EFI forecast from 10 May 2015 0000 UTC shows a wide area, ranging from northern Iberia through France, Switzerland and southern Germany to Hungary, where CAPE and CSP reach extremely high values (Figure 6a and 6c). This suggests that convection, if it is initiated, could be well-organised and would be capable of producing large hail, severe wind gusts and even tornadoes. The positive SOT values show that at least 10% of ENS members forecast values of CAPE and CSP exceeding the 99th percentile of the M-climate (see Box A). The M-climate is shown for reference in Figures 6b and 6d.

Figure 6e shows the probability of convective precipitation greater than 1 mm from ENS, together with the large-scale flow forecast represented by the ensemble mean of 500-hPa geopotential. The highest probability of rain was forecast for easternmost France, Switzerland, southern Germany and parts of Austria and Hungary. For the rest of the area covered by the EFI the probabilities are very low, suggesting that DMC is very unlikely there. In the event, severe storms did develop over the aforementioned areas on 13 May, producing large hail and severe wind gusts in the evening hours with a tornado reported over southern Germany. Analysis of other cases similar to this one suggests that viewing the EFI for CAPE and CSP alongside the probability of convective rain can provide the forecaster with more precise guidance on where severe convection is likely to occur in the medium range – in this particular case that means 4 days in advance (T+72–96h).

Practical considerations

Neither of the convective parameters presented here considers Convective Inhibition (CIN – see Box B), which is a measure of the amount of energy an air parcel needs in order to reach the level of free convection (LFC). If CIN is large, DMC is unlikely to occur even if the EFI is high.

The EFI gives an indication of the likelihood of anomalous weather relative to a baseline distribution of what ordinarily occurs in the given location at the given time of year. So in areas where even the climatological extremes of the convective index are small, e.g. in continental areas in winter, severe convective weather such as tornadoes or large hail are unlikely even if the EFI is high. Therefore, to be sure of whether severe convection is possible or not, the forecaster should always view the EFI alongside the absolute values of the given convective parameter seen in climatology (provided by the M-climate), and then use their knowledge and experience of what levels of these parameters might be needed to lead to severe convection.

The results presented in this study suggest that the EFI can be applied successfully to forecasting severe convection in the medium range. EFIs for CAPE and the combined CAPE-SHEAR parameter CSP are now available on the ECMWF website in test mode and any feedback from users is welcome.

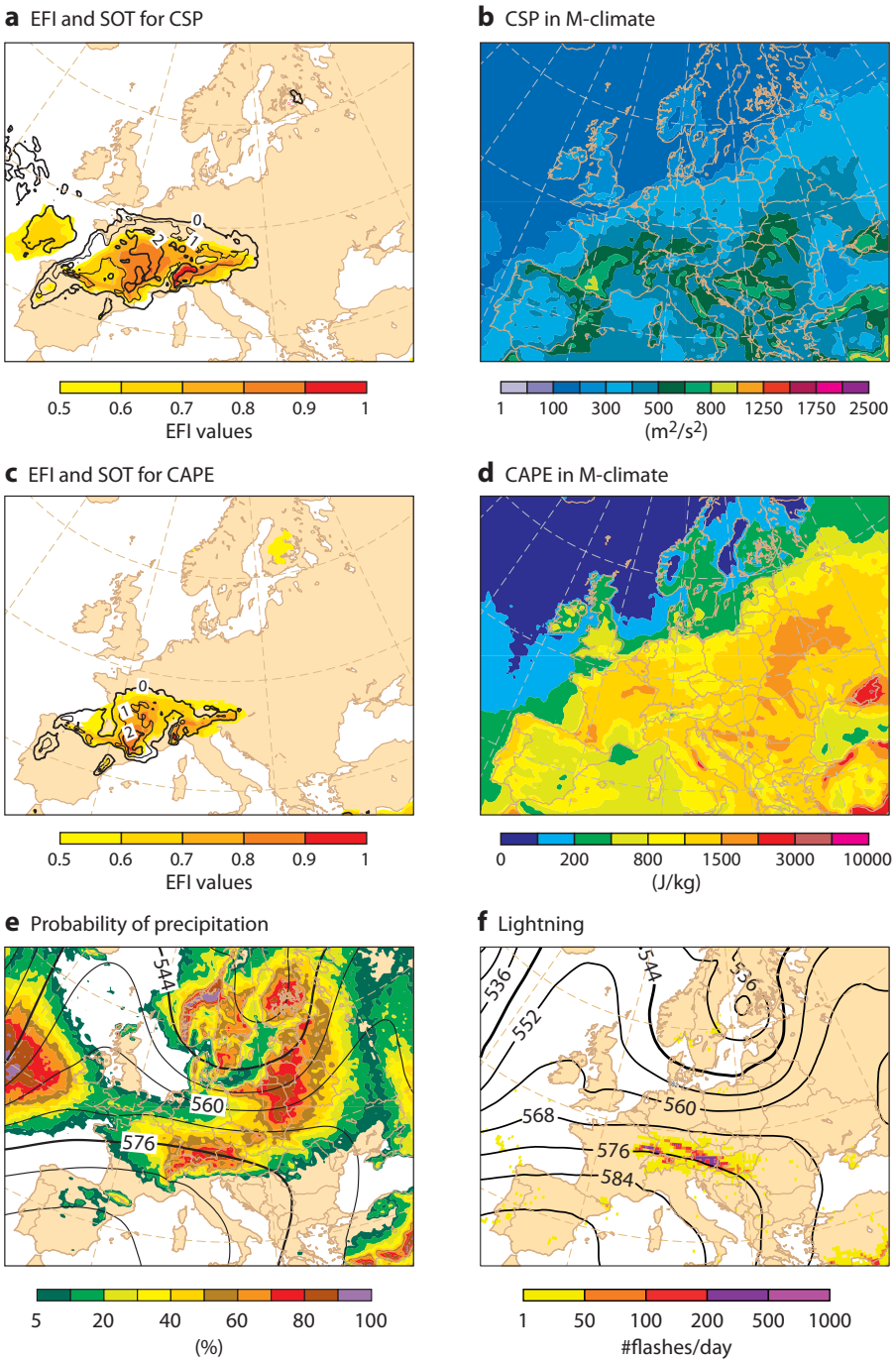


Figure 6 (a) EFI (shading) and SOT (contours) for CSP, T+72–96h forecast from 10 May 2015, (b) M-climate 99th percentile for CSP in m^2/s^2, (c) EFI (shading) and SOT (contours) for CAPE, T+72–96h forecast from 10 May 2015, (d) M-climate 99th percentile for CAPE in J/kg, (e) probability of convective precipitation ≥ 1mm, T+72–96h (in %, shading) and ENS mean of 500 hPa geopotential height (in geopotential decametres, contours), T+84h, from 10 May 2015, and (f) lightning activity (in number of flashes/day per grid box) valid for 13 May 2015 and ECMWF analysis of 500 hPa geopotential height (in geopotential decametres, contours) valid for 13 May 2015 1200 UTC.

FURTHER READING

Bechtold, **P.**, **N. Semane**, **Ph. Lopez**, **J.-P. Chaboureau**, **A. Beljaars** & **N. Bormann**, 2014: Representing Equilibrium and Nonequilibrium Convection in Large-Scale Models. *J. Atmos. Sci.*, **71**, 734–753.

Craven, **J.P.** & **H. Brooks**, 2004: Baseline climatology of sounding derived parameters associated with deep moist convection. *Natl. Wea. Dig.*, **28**, 13–24.

Doswell, **C.A. III**, **H.E. Brooks** & **R.A. Maddox**, 1996: Flash flood forecasting: An ingredients-based methodology. *Wea. Forecasting*, **11**, 560–580.

Kaltenböck, R., **G. Diendorfer** & **N. Dotzek**, 2009: Evaluation of thunderstorm indices from ECMWF analyses, lightning data and severe storm reports. *Atmospheric Research*, **93**, 381–396.

Rasmussen, **E.N.** & **D.O. Blanchard**, 1998: A baseline climatology of sounding-derived supercell and tornado forecast parameters. *Wea. Forecasting*, **13**, 1148–1164.

Promising results in hybrid data assimilation tests

MATS HAMRUD, MASSIMO BONAVITA , LARS ISAKSEN

Data assimilation systems for global numerical weather prediction (NWP) combine short-range forecasts (the 'background') with the latest observations to arrive at the best possible representation of the current state of the atmosphere. They are designed to ensure that the resulting analysis of the atmosphere is as accurate and realistic as possible, while taking into account the errors associated both with the background and with observations.

ECMWF has been a pioneer in the development and operational implementation of a data assimilation method called 4DVAR. '4D' stands for the three spatial dimensions plus time, as this method uses observations as they come in over a period of time, while 'VAR' refers to variational methods.

An alternative algorithm, called the Ensemble Kalman Filter (EnKF), is also suitable for operational use. In recent tests, an EnKF-based data assimilation system developed at the Centre has shown good forecast performance, and a hybrid 4DVAR/EnKF approach has been found to perform significantly better than the two systems in standard configuration individually. The EnKF algorithm is also highly scalable, making it particularly well-suited to future, massively parallel computer architectures.

The performance of the hybrid 4DVAR/EnKF system has been found to be comparable to that of a low-resolution version of the data assimilation system in operational use at ECMWF, which is a 4DVAR system combined with an ensemble of data assimilations.

The EnKF system

A Kalman Filter is an algorithm designed to estimate the state of a system based on predictions of the system's behaviour on the one hand and observations on the other. In NWP, the system is the atmosphere, the predictions are short-range forecasts, and the observations are the most recent set of atmospheric observations available. An Ensemble Kalman Filter makes direct use of an ensemble of short-range forecasts to estimate the errors associated with the background.

By contrast, ECMWF's operational data assimilation system is a 4DVAR system which uses error statistics provided by an Ensemble of Data Assimilations (EDA). The EDA is an ensemble of lower resolution 4DVAR assimilations using perturbed observations and perturbed model tendencies.

Despite these differences, the EnKF system developed at ECMWF uses the operational EDA component in a number of ways. The model propagating the ensemble is the same, and the operators calculating the observation equivalents are the same. Both use the ECMWF Integrated Forecasting System (IFS) code base. The observation Quality Control (QC) and data selection procedure are also similar, using the same kind of background checks and data thinning.

The EnKF system also uses large parts of the technical infrastructure of the operational system at ECMWF. The observation equivalents, H(x), are evaluated at the observation time by running forecasts over the analysis window and applying the observation operator to the evolving state. The post-processing of the analysis, including producing pressure level data, is also performed by the IFS.

Two main variants of EnKF have been implemented: the Ensemble Square Root Filter (EnSRF) (*Whitaker & Hamill*, 2002) and the Local Ensemble Transform Kalman Filter (LETKF) (*Hunt et al.*, 2007). Both these EnKF implementations do not require perturbing the observations, as in the EDA approach, which has the theoretical advantage of not introducing additional sampling errors. Early experimentation showed that there was no statistically significant difference in performance between the two EnKF schemes. The focus in this article will be on the LETKF (see Box A for more details on the LETKF).

The LETKF algorithm

A

The version of the EnKF used in the experiments is the Local Ensemble Transform Kalman Filter (LETKF). The LETKF algorithm can be thought of as a way of minimizing the standard 4DVAR cost function, i.e.:

$$J(\boldsymbol{x}) = \tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^b)^T\mathbf{B}^{-1}(\boldsymbol{x}-\boldsymbol{x}^b) + \tfrac{1}{2}(\boldsymbol{y}-\mathbf{H}(\boldsymbol{x}))^T\mathbf{R}^{-1}(\boldsymbol{y}-\mathbf{H}(\boldsymbol{x})) \quad (1)$$

in the space spanned by the ensemble. This means that in Equation 1 the background error covariance matrix is directly sampled from the ensemble background forecasts $\boldsymbol{x}_i^b$:

$$\mathbf{B} = \frac{1}{N_e-1}\sum_{i=1}^{N_e}(\boldsymbol{x}_i^b-\overline{\boldsymbol{x}^b})^T(\boldsymbol{x}_i^b-\overline{\boldsymbol{x}^b}) = \frac{1}{N_e-1}\mathbf{X}^b\,(\mathbf{X}^b)^T \quad (2)$$

For linear observation operators, the analysis is the minimum of the cost function (1) in the space spanned by the ensemble background forecasts $\boldsymbol{x}_i^b$ and it can be expressed in closed form as:

$$\boldsymbol{x}^a = \overline{\boldsymbol{x}^b} + \mathbf{X}^b\boldsymbol{w}^a \quad (3)$$

$$\boldsymbol{w}^a = \left((N_e-1)\boldsymbol{I} + (\mathbf{Y}^b)^T\mathbf{R}^{-1}(\mathbf{Y}^b)\right)^{-1}(\mathbf{Y}^b)^T\mathbf{R}^{-1}(\boldsymbol{y}-\overline{\mathbf{H}(\boldsymbol{x}_i^b)}) \quad (4)$$

where $\mathbf{Y}^b$ are matrices whose columns are formed by $\mathbf{H}(\boldsymbol{x}_i^b) - \overline{\mathbf{H}(\boldsymbol{x}_i^b)}$. From Equations (3) and (4), it is apparent that the analysis essentially determines which linear combination of the ensemble members is the best estimate of the current state, given the current batch of observations. The error covariance of the analysed state is given by:

$$\mathbf{P}^a = \mathbf{X}^b\left((N_e-1)\boldsymbol{I} + (\mathbf{Y}^b)^T\mathbf{R}^{-1}(\mathbf{Y}^b)\right)^{-1}(\mathbf{X}^b)^T \quad (5)$$

Taking the symmetric square root of Equation (5) provides the perturbations to add to the analysis (3) to form the full analysed ensemble, which can then be integrated forward in time.

The analysed atmospheric variables in the ECMWF EnKF are temperature, wind vector components (u, v), specific humidity, surface pressure and surface pressure tendency.

Cap on observations

In the LETKF context, it has been found useful to limit the number of observations affecting the analysis of any grid point. This number is currently limited to 30 for each combination of different report type and observation variable, with the observations closest to the analysed grid point being selected. This was found to produce a significant advantage in terms of forecast scores in the northern hemisphere compared to using all available observations in the analysis update. At the same time it enables massive computational savings. Results are remarkably insensitive to the precise number of selected observations (for example, doubling or halving the number of locally selected observations has a marginal impact on analysis quality).

Covariance inflation

As in most EnKF implementations, two complementary types of covariance inflation have been implemented, namely multiplicative and additive inflation (*Whitaker & Hamill*, 2012). The multiplicative inflation relaxes the analysis variance towards the background variance to prevent it from dropping to unrealistically low values in densely observed regions. Experiments have shown that the EnKF skill is not very sensitive to the relaxation factor α; an α of 0.9 has been used in the experimentation. The additive inflation is based on the difference between the 48-hour and 24-hour IFS forecasts verifying at the same time, sampled over a 6-year period. A random sample of these forecast differences is added to each ensemble member at each analysis cycle. These are scaled by a factor of 0.25, which converts the forecast difference to a nominal 6-hour difference.

Bias correction

In the ECMWF 4DVAR implementation, a variational bias correction method is used for satellite radiance data. The computational cost of the variational bias correction is insignificant in terms of the whole 4DVAR cost. In the EnKF context, an equivalent scheme can be implemented but at a non-negligible computational cost. We have not implemented such a scheme; instead we have relied on the stored bias corrections produced by the 4DVAR system. In the context of the hybrid system presented below, the bias correction can be performed by the 4DVAR component of the system.

Controlling analysis noise

It is well documented that EnKF analyses tend to have problems in representing the model dynamical balance constraints, leading to the excitation of spurious inertia-gravity waves in the short-range forecast. This is normally attributed to the effect of covariance localization disrupting the mass-wind field geostrophic balance. Diagnostics within the ECMWF EnKF system do show evidence of significant small-scale noise associated with inertia-gravity waves in forecasts based on the EnKF analysis. To counter this problem, EnKF systems often apply some form of Digital Filter Initialization (DFI).

Experimentation with different versions of DFI applied to the EnKF analysis did in fact reduce the imbalances, but always degraded forecast scores. We have therefore developed a new method which adjusts the column horizontal wind divergence to match the analysed surface pressure tendency. This 'divergence adjustment' (see *Hamrud et al.*, 2014) is performed for all ensemble members. The wind increments caused by this procedure are normally very small (much smaller than the analysis increments) and the details of their distribution in the vertical do not seem to be very important. The significant change in the initial pressure tendency achieved by this procedure is illustrated in Figure 1. This method effectively imposes an additional dynamical balance constraint on the analysis and thus removes a large portion of the excess inertia-gravity waves from the short-range forecast.

The impact of this 'divergence adjustment' on the quality of the analysis is significant and positive. The 6-hour forecast fit for land-based surface pressure observations improves by around 5% and by around 10% for marine surface pressure observations. The effect on the observational fit of other variables is negligible.

Scalability and computational cost

There are no fundamental algorithmic constraints on the scalability of an EnKF data assimilation system. The forecasts and the computation of the observation equivalents can be done independently for all ensemble members. The EnKF analysis itself is local. In principle, the limit on dividing the computations into independent pieces of work is given by the size of the forecast state.

There are, however, a number of practical limits to the scalability of the EnKF system. One issue is the Input/Output (IO) bottleneck. In the analysis stage, all forecasts are needed to describe the background state, thus they have to be read in and distributed. At the end of the analysis, all the analysis states have to be written out to provide the initial states for the next analysis cycle's forecasts. In addition, all the observational data, including the observation departures for all ensemble members, have to be read in. This IO load is distributed over tasks in the ECMWF EnKF implementation, but for large applications the overall throughput of the IO subsystem still limits the scalability.

The other practical scalability issues for the analysis stage are quite different for the EnSRF and LETKF analyses. Limiting our analysis to the LETKF (the interested reader can find a broader discussion in *Hamrud et al.*, 2014), where the state is distributed in contiguous areas and the observations are distributed so that each task has local access to the observations it needs, the problem becomes that of load balancing. The cost of analysing a grid point is dependent on the number of nearby observations, and as the observation density is not homogeneous, the cost varies for different parts of the globe. Thus a distribution

with an equal number of state points on each processor would lead to a poor load balance.

To prevent this, we measure the cost of performing the analysis of each state column as a function of the number of close observations. Then, in subsequent analysis steps, we find out the number of nearby observations for each state profile in the global view and combine this information with the empirical relation between observation density and computational cost to compute weights for the distribution of the state space variables. The grid points are then distributed so that each processor ends up with approximately the same computational weight, rather than with the same number of grid points. This method normally gives a reasonable load balance, with a ratio of the fastest to the slowest analysis time on an individual processor of typically around 0.85.

In terms of computational cost, the natural comparison for an EnKF implementation is ECMWF's EDA system, based on multiple independent copies of 4DVAR (*Isaksen et al.*, 2010). Here an EnKF implementation has an obvious advantage: the cost per member of an EnKF system is considerably lower than for the EDA, mainly due to the fact that the Kalman Gain is only computed once in the EnKF.

Sensitivity experiments

EnKF systems are relatively easy to set up and tend to produce acceptable results with a broad range of 'reasonable' choices of the main filter parameters. A few examples of parameters that affect LETKF performance will be discussed here.

A larger ensemble size is desirable as it reduces the sampling errors in the background error covariance estimates. A set of experiments was carried out in which the ensemble size was set to 60, 120 and 240. The experiments were run with a triangular spectral truncation of 159 for the IFS model (~120 km linear grid spacing) and 91 vertical levels. All available observations were used, with the exception of radiances from satellite imagers and scatterometer winds, which, at the time of the experiments, had not yet been implemented in the EnKF. The sample ensemble covariances were localized in order to reduce sampling errors, with a localization radius of 2,800 km in the horizontal and 2.5 scale heights in the vertical (scale height is here the vertical distance over which pressure falls by a factor of 1/e).

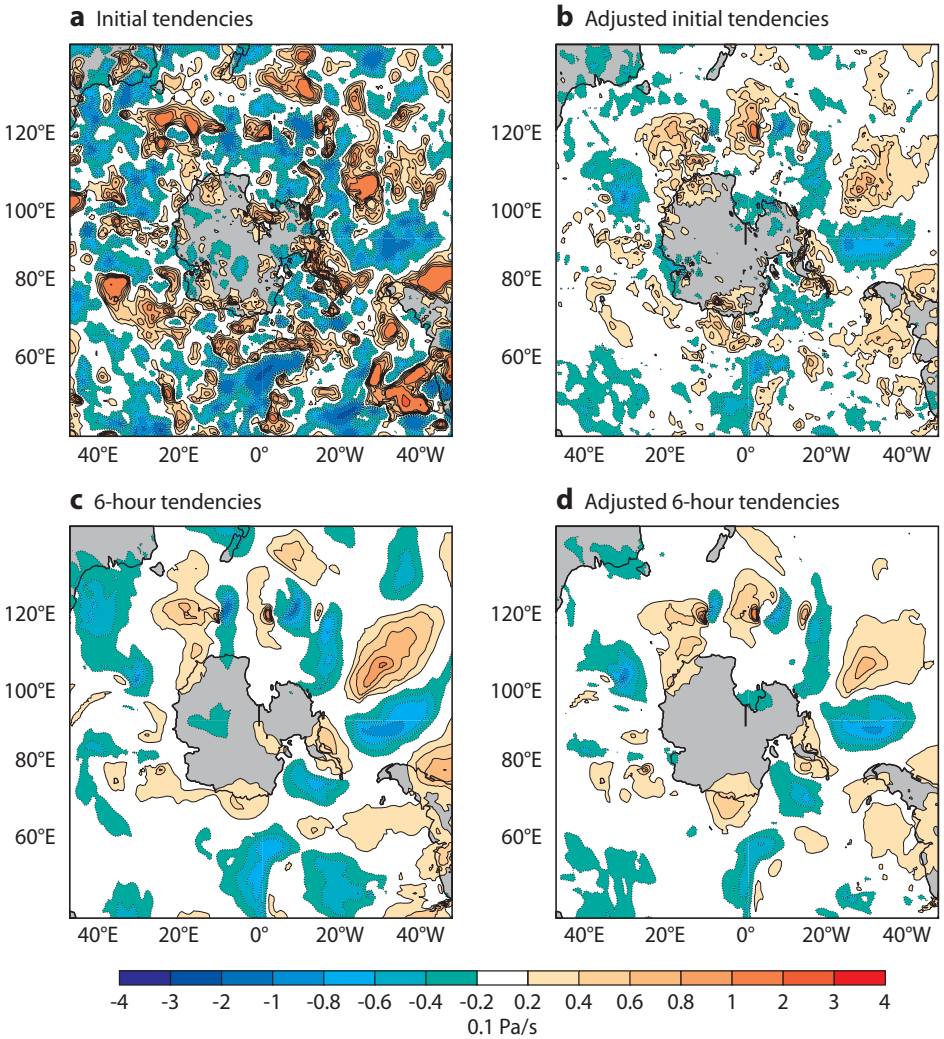


Figure 1 Surface pressure tendency fields over the southern hemisphere in forecasts based on the ensemble mean analysis for (a) initial tendencies without divergence adjustment, (b) initial tendencies with divergence adjustment, (c) tendencies after 6 hours without divergence adjustment, and (d) tendencies after 6 hours with divergence adjustment.

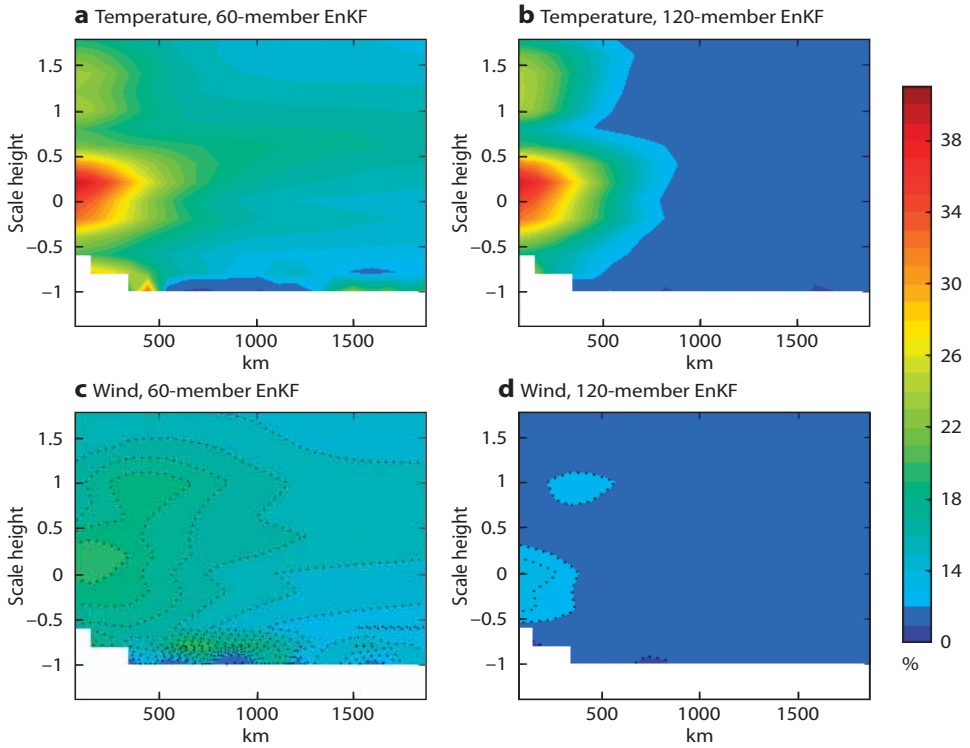


Figure 2 Azimuthally-averaged absolute correlation between short-range (3-hour) forecasts of (a) temperature at 500 hPa and AMSU-A brightness temperature model equivalents for a 60-member EnKF; (b) temperature at 500 hPa and AMSU-A brightness temperature model equivalents for a 120-member EnKF; (c) zonal wind at 500 hPa and AMSU-A brightness temperature model equivalents for a 60-member EnKF; and (d) zonal wind at 500 hPa and AMSU-A brightness temperature model equivalents for a 120-member EnKF.

Figure 2 shows how the increase in ensemble size helps to better characterize sampled error correlations from the EnKF. Figures 2a–b show, as a function of horizontal and vertical distance, the average of the sample absolute correlations between 3-hour forecasts of 500 hPa temperature and observations, in the form of AMSU-A (Advanced Microwave Sounding Unit A) brightness temperature model equivalents, for the 60-member EnKF (Figure 2a) and the 120-member EnKF (Figure 2b). Figures 2c–d show the average of the sample absolute correlations between 500 hPa zonal wind 3-hour forecasts and AMSU-A brightness temperature model equivalents for the 60-member EnKF (Figure 2c) and the 120-member EnKF (Figure 2d). In both instances it is apparent how the structure of the spatial correlations is much better defined in the 120-member ensemble, which presents lower values of spurious long-range correlations.

Figure 3 shows forecast skill scores, averaged over the month of February 2011, of forecasts based on the EnKF ensemble mean analyses for a 60-member EnKF, a 120-member EnKF and a 240-member EnKF in the northern extra-tropics (Figure 3a) and the southern extra-tropics (Figure 3b). The improvement from a 60-member to a 120-member ensemble is consistent and statistically significant throughout the forecast range. On the other hand, gains are more limited when moving from a 120-member ensemble to 240 members. Part of the reason is the fact that covariance localization parameters were not changed in these experiments. Re-running the 240-member EnKF experiment with double the vertical covariance localization factor shows a clear improvement.

Higher spatial resolution is beneficial for the EnKF as the short-range forecast fields behave more realistically in the

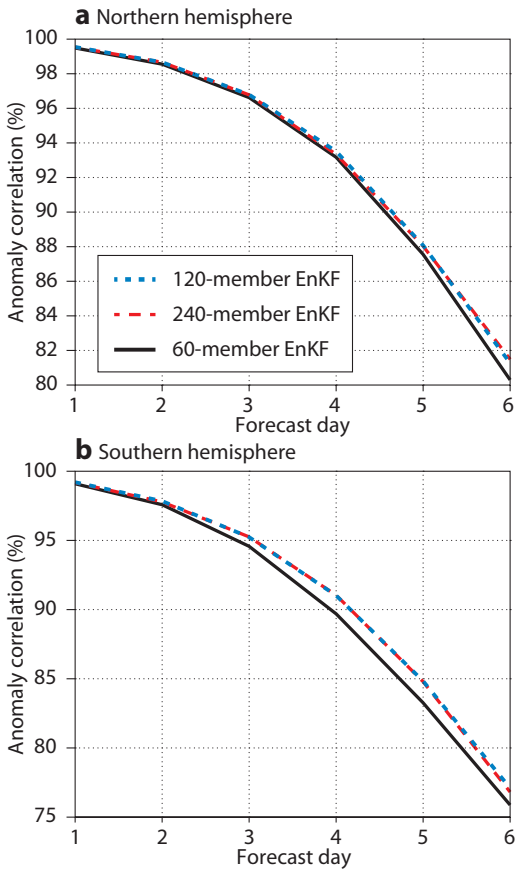


Figure 3 500-hPa geopotential forecast anomaly correlation scores for the 60-member EnKF, the 120-member EnKF, and the 240-member EnKF computed with respect to ECMWF's operational analysis and averaged over the month of February 2011 in (a) the northern hemisphere and (b) the southern hemisphere.

high wavenumber part of the spectrum, and they benefit from a more accurate representation of the physiographic constraints. This is shown in Figure 4, where we present the deterministic skill scores for forecasts, over the months of February and March 2011, based on ensemble mean analyses in 60-member EnKF assimilation experiments run at three different resolutions: spectral truncations TL159 (approximate grid-point spacing 120 km), TL319 (approximate grid-point spacing 60 km) and TL639 (approximate grid-point spacing 30 km). The positive impact of increasing horizontal resolution is evident going from TL159 to TL319, while diminishing returns are seen from TL319 to TL639. The increase in horizontal resolution also has the important side effect of generating more active ensemble background fields with a larger spread. This in turn helps to reduce the reliance of the EnKF on the adaptive covariance inflation algorithm that is used to prevent the filter collapse.

EnKF performance

The development activities described in the previous sections have led to the availability of a stable EnKF-based data assimilation system at ECMWF. This system is able to assimilate all the available observation types, it has good scalability properties and it is computationally efficient. It is thus possible to perform a detailed comparison with the 4DVAR system available at ECMWF under controlled conditions.

An extended assimilation experiment has been run to evaluate the skill of the ECMWF EnKF analysis in a deterministic sense, i.e. when the ensemble mean is used as the initial condition for a single, deterministic forecast. This experiment has been run with triangular spectral truncation TL399 (corresponding to approximately 50-km grid spacing) and 137 model levels (near surface to 0.01 hPa), which is the current resolution of the EDA, using IFS model cycle 40r1.

An example of the results is given in Figure 5, where the forecast skill scores (geopotential height anomaly correlation at 500 hPa) of the EnKF system and a 4DVAR system using climatological background errors and covariances are presented. It is apparent that the two lines are hardly separable. The general conclusion that can be drawn is that, at this stage of development, the ECMWF EnKF-based data assimilation system has similar deterministic forecast performance to a 4DVAR system using static background error covariance estimates. The EnKF is at TL399 resolution and 4DVAR at TL399 outer loop and TL95/TL159 inner loop resolution.

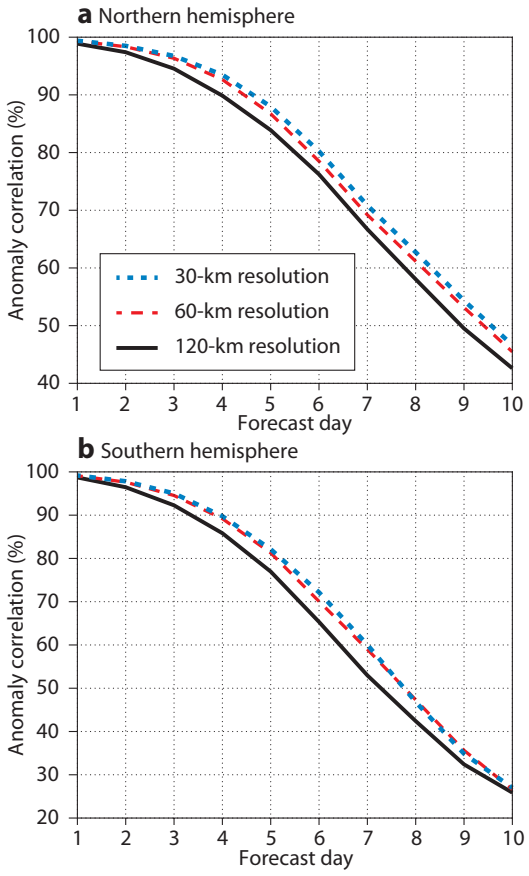


Figure 4 1,000-hPa geopotential forecast anomaly correlation scores for ensemble mean forecasts with 120-km resolution (T159) EnKF, 60-km resolution (T319) EnKF, and 30-km resolution (T639) EnKF for (a) the northern hemisphere and (b) the southern hemisphere. Scores computed with respect to ECMWF operational analysis and averaged over the months of February and March 2011.

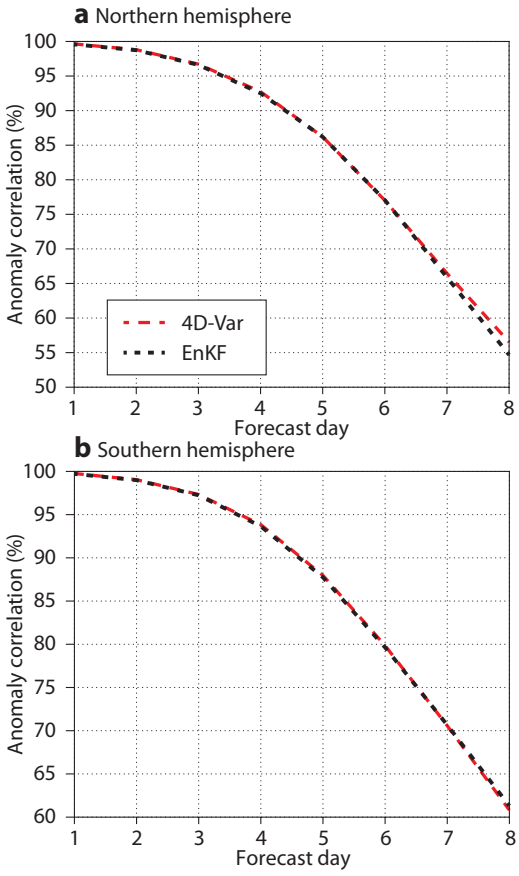


Figure 5 500-hPa geopotential forecast anomaly correlation for a 100-member EnKF and 4DVAR using climatological background errors and covariances, both at 50-km resolution (TL399), in (a) the northern hemisphere and (b) the southern hemisphere. All scores computed with respect to ECMWF's operational analysis and averaged over the period from 1 May to 31 August 2012.

Hybrid Gain EnDA

Recent work by *Penny* (2014) proposed, in an idealized setup, a new hybrid variational/EnKF formulation that involves taking a weighted mean of the EnKF and variational analysis and using this 'control' state to re-centre the EnKF analysis ensemble (Box B). This is equivalent to blending the Kalman gains of the EnKF and 4DVAR systems and is thus called Hybrid Gain Ensemble Data Assimilation (HG-EnDA). At ECMWF, such a system produced a considerable gain in forecast skill compared to that of the EnKF and a 4DVAR system using climatological background errors and covariances (Figure 6, similar results are obtained for most variables and vertical levels). Where does this improvement originate?

An indication is given by Figure 7, where we show the mean sea level (MSL) pressure control background forecast and its standard deviation for an HG-EnDA experiment valid on 25 February 2014 at 0000 UTC (Figure 7a); the EnKF analysis increment to a single surface pressure observation located at (58.5°N, 30.3°W) and 1 hPa lower than the background forecast (Figure 7b); and the corresponding analysis increment of the 4DVAR analysis (Figure 7c).

It is apparent that the EnKF analysis increments have a smaller footprint due to the applied horizontal covariance localization; 4DVAR analysis increments have a considerably larger spatial extension as they are based on climatological correlations which have not been localized. On the other hand the EnKF increment shows interesting, flow-dependent, short-range covariance structures while the corresponding 4DVAR increments present a more isotropic structure (due to the climatological sampling of the 4DVAR

The Hybrid Gain EnDA

B

In the Hybrid Gain EnDA scheme, in addition to a standard EnKF analysis update as described in Box A, an incremental 4DVAR analysis update is also performed. The 4DVAR analysis uses the background forecast from the previous ensemble mean analysis as background and starting linearization trajectory. Formally, the incremental 4DVAR cost function is minimised, i.e. in standard notation:

$$J(\delta x) = \frac{1}{2}\delta x^T \mathbf{B}^{-1}\delta x + \frac{1}{2}(\mathbf{H}\delta x - \mathbf{d})^T \mathbf{R}^{-1}(\mathbf{H}\delta x - \mathbf{d}) \quad (1)$$

where $\mathbf{d} = \mathbf{y}^o - H(\mathbf{x}^b)$ is the innovation vector and $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^b$ is the incremental control vector , with the background provided by a short (t+3h) nonlinear model integration of the Hybrid Gain mean analysis valid at the previous analysis update time:

$$\mathbf{x}^b = M\left(\overline{\mathbf{x}^a_{iEnKF}}(t_{k-1})\right) \quad (2)$$

The EnKF and 4DVAR analysis fields valid in the middle of the assimilation window (i.e., 0000, 0600, 1200, 1800 UTC in the 6-hour cycling setup used in the experiments) are then linearly combined to produce a 'hybrid gain' analysis around which the EnKF analysis ensemble is re-centred:

$$\mathbf{x}^a_{Hyb} = \alpha\overline{\mathbf{x}^a_{iEnKF}} + (1-\alpha)\mathbf{x}^a_{4DVar} \quad (3)$$

The weight given to each of the contributing analyses (denoted as α) is a free tuning parameter, reflecting the expected accuracy of the EnKF and 4DVAR analyses. A value of $\alpha = 0.5$ has been used in the experiments reported in this work. This is thought to be not far from the optimal value, as later experimentation with $\alpha = 0.75$ and $\alpha = 0.25$ has generally produced inferior results.

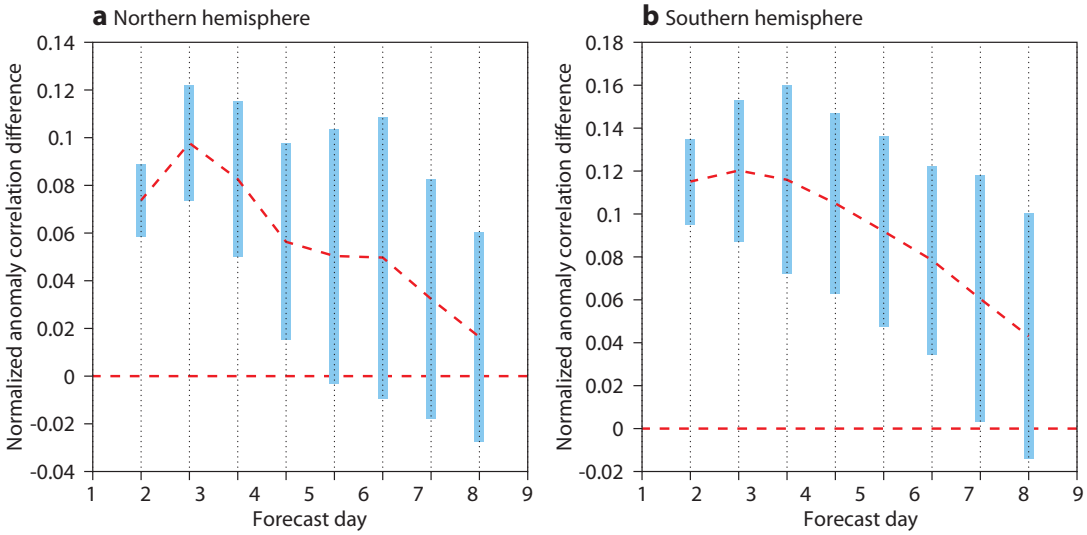


Figure 6 Normalised difference of the 500-hPa geopotential height forecast anomaly correlation between Hybrid Gain EnDA and 4DVAR using climatological background errors and covariances in (a) the northern hemisphere and (b) the southern hemisphere. Positive values indicate superior skill of the Hybrid Gain EnDA. Error bars show 95% confidence intervals. All scores computed with respect to ECMWF's operational analysis and averaged over the period from 1 May to 31 August 2012.

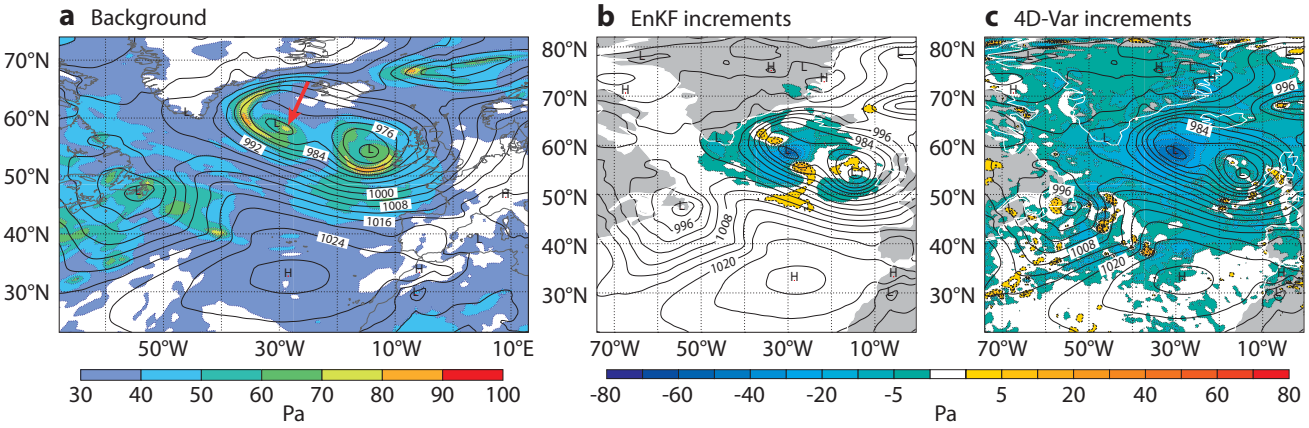


Figure 7 (a) Mean sea level (MSL) pressure control background forecast (black isolines, units: hPa) and standard deviation of MSL pressure ensemble background forecasts (shading) from a Hybrid Gain EnDA experiment valid on 25 February 2014 at 0000 UTC, (b) analysis increments for the single observation experiment with a synthetic surface pressure observation located at 58.5°N, 30.3°W (red arrow in left-hand panel) and 1hPa lower than the background forecast, in the EnKF mean analysis, and (c) analysis increments for the same observation experiment in the 4DVAR analysis.

covariances and the locally isotropic structure of the background error model used in 4DVAR).

The HG-EnDA, as a mixture of these two increments, is thus able to retain some of the flow-dependent aspects of the EnKF analysis update but also to partially correct for its two main apparent drawbacks, i.e. the underestimation of longer-range correlations and the underdispersiveness of the ensemble background in specific meteorological situations.

The use of a static 4DVAR in the re-centring step of the Hybrid Gain EnDA can be seen as a regularization procedure of the EnKF analysis which has two main benefits: a) it reduces the effect of sampling noise and localization; and b) it introduces climatological information into the background error covariance estimates, which is important to reduce the effect of model biases in the analysis and is not easy to do in a standard EnKF. The results show that, similar to the operational hybrid 4DVAR/EDA system, the Hybrid Gain algorithm clearly improves on the accuracy of its individual EnKF and 4DVAR components while its computational costs are only marginally higher than those of a standard EnKF algorithm. But it does require the development and maintenance of both the EnKF and the 4DVAR analysis systems.

Further tests not detailed here suggest that the hybrid 4DVAR/EnKF system is comparable in performance to a low-resolution version of the operational hybrid 4DVAR/EDA system (see *Hamrud et al.*, 2014).

Outlook

The wish to carry out detailed comparisons between variational and more scalable ensemble-based assimilation systems has led to the development of a state-of-the-art EnKF system at ECMWF. The results of these comparisons are promising in terms of both forecast accuracy and computational cost, especially in the Hybrid Gain configuration. Research is continuing to further evaluate the system and also to verify whether these results, obtained at the resolution of the currently operational EDA, can be confirmed at the higher horizontal resolutions that are planned to be introduced later this year. Another important aspect to evaluate is the impact of using the Hybrid Gain analysis ensemble to initialize ECMWF's ensemble forecasts. The Hybrid Gain is one of several hybrid 4DVAR configurations that will be explored at ECMWF over the coming years as potential alternatives to the current hybrid 4DVAR/EDA system.

FURTHER READING

Hamrud, **M.**, **M. Bonavita** & **L. Isaksen**, 2014: EnKF and Hybrid Gain Ensemble Data Assimilation. *ECMWF Tech. Memorandum* **No 733**.

Hunt, **B. R.**, **E. J. Kostelich** & **I. Szunyogh**, 2007: Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D*, **230**, 112–126.

Isaksen, **L.**, **M. Bonavita**, **R. Buizza**, **M. Fisher**, **J. Haseler**, **M. Leutbecher** & **L. Raynaud**, 2010: Ensemble of data assimilations at ECMWF. *ECMWF Tech. Memorandum* **No 636**.

Penny, **G. S.**, 2014: The Hybrid Local Ensemble Transform Kalman Filter. *Mon. Wea. Rev.*, **142**, 2139–2149.

Whitaker, **J. S.** & **T. M. Hamill**, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.

Whitaker, **J. S.** & **T. M. Hamill**, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089.

ECMWF Calendar 2015

Aug 31–Sep 4	Annual Seminar
Sep 28–Oct 2	Visualisation Week: • Workshop on Meteorological Operational Systems • European Working Group on Operational Meteorological Workstations (EGOWS) • RMetS Seminar on Visualisation in Meteorology • OGC MetOcean Interoperability Session
Oct 5–7	Training Course: Use and Interpretation of ECMWF Products
Oct 12–14	Scientific Advisory Committee
Oct 15–16	Technical Advisory Committee
Oct 21	Policy Advisory Committee
Oct 22–23	Finance Committee
Oct 26	Advisory Committee for Co-operating States (ACCS)
Nov 2–5	Workshop on Sub-Seasonal Predictability
Dec 8–9	Council

Contact information

ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, UK

Telephone National 0118 949 9000

Telephone International +44 118 949 9000

Fax +44 118 986 9450

ECMWF's public website http://www.ecmwf.int/

E-mail: The e-mail address of an individual at the Centre is firstinitial.lastname@ecmwf.int. For double-barrelled names use a hyphen (e.g. j-n.name-name@ecmwf.int).

Problems, queries and advice	Contact
General problems, fault reporting, web access and service queries	calldesk@ecmwf.int
Advice on the usage of computing and archiving services	advisory@ecmwf.int
Queries regarding access to data	data.services@ecmwf.int
Queries regarding the installation of ECMWF software packages	software.support@ecmwf.int
Queries or feedback regarding the forecast products	forecast_user@ecmwf.int

ECMWF publications

(see http://www.ecmwf.int/en/research/publications)

Technical Memoranda

757 **Lee**, **S.**, **K. Salonen** & **N. Bormann**: Assessment of AMV's from COMS in the ECMWF system. *July 2015*

756 **Bormann**, **N.**, **M. Bonavita**, **R. Dragani**, **R. Eresmaa**, **M. Matricardi** & **T. McNally**: Enhancing the impact of IASI observations through an updated observation error covariance matrix. *July 2015*

755 **Janssen**, **P.**: Notes on the maximum wave height distribution. *June 2015*

754 **Buizza**, **R.** & **M. Leutbecher**: The Forecast Skill Horizon. *June 2015*

753 **Lupu**, **C.**, **C. Cardinali** & **A.P. McNally**: Adjoint-based forecast sensitivity applied to observation error variances tuning. *June 2015*

752 **Semane**, **N.** & **P. Bechtold**: Convection and waves on Small Earth and Deep Atmosphere. *March 2015*

751 **Shi**, **W.**, **N. Schaller**, **D. MacLeod**, **T.N. Palmer** & **A. Weisheimer**: Impact of hindcast length on estimates of seasonal climate predictability. *March 2015*

750 **Di Giuseppe**, **F.** & **A.M. Tompkins**: A parameterization of cloud overlap as a function of wind shear and its impact in ECMWF forecast. *June 2015*

749 **Lupu**, **C.**, **A.J. Geer** & **N. Bormann**: Revision of the microwave coefficient files in the IFS. *March 2015*

748 **Lupu**, **C.** & **A.J. Geer**: Operational implementation of RTTOV-11 in the IFS. *March 2015*

ERA Report Series

20 **Poli**, **P.**, **H. Hersbach**, **P. Berrisford et al**: ERA-20C Deterministic. *July 2015*

18 **Hersbach**, **H.**, **P. Poli** & **D. Dee**: The observation feedback archive for the ICOADS and ISPD data sets. *May 2015*

EUMETSAT/ECMWF Fellowship Programme Research Reports

37 **Lawrence**, **H.**, **N. Bormann**, **Q. Lu**, **A. Geer** & **S. English**: An Evaluation of FY-3C MWHS-2 at ECMWF. *June 2015*

36 **Salonen**, **K.** & **N. Bormann**: Atmospheric Motion Vector observations in the ECMWF System: Fourth year report. *May 2015*

Index of newsletter articles

This is a selection of articles published in the *ECMWF Newsletter* series during recent years.
Articles are arranged in date order within each subject category.
Articles can be accessed on the ECMWF public website – http:// www.ecmwf.int/en/research/publications

ECMWF celebrates its 40th anniversary

ECMWF marked its 40th anniversary on 25 June with a series of six afternoon talks in the Centre's lecture theatre. The speakers were (from left to right) Professor Lennart Bengtsson, the Director of ECMWF from 1981 to 1990; Michel Jarraud, the Secretary-General of the World Meteorological Organization (WMO); Professor Sarah Jones, the Head of Research and Development at the German Meteorological Service (DWD); Dr Florence Rabier, ECMWF's Director of Forecasts; Dr Nils Wedi, the Head of ECMWF's Numerical Aspects Section; and Professor Alan Thorpe, the Director-General of ECMWF. Bottom left: Professor Erland Källén, ECMWF's Director of Research, Florence Rabier and Lennart Bengtsson.

For a full report, see: www.ecmwf.int/en/about/media-centre/news/2015/ecmwf-40th-anniversary-speakers-hail-progress-nwp

ECMWF

Newsletter | Number 144 – Summer 2015

European Centre for Medium-Range Weather Forecasts

www.ecmwf.int