ECMWF

European Centre for Medium-Range Weather Forecasts

Newsletter

No. 166 | Winter 2020/21

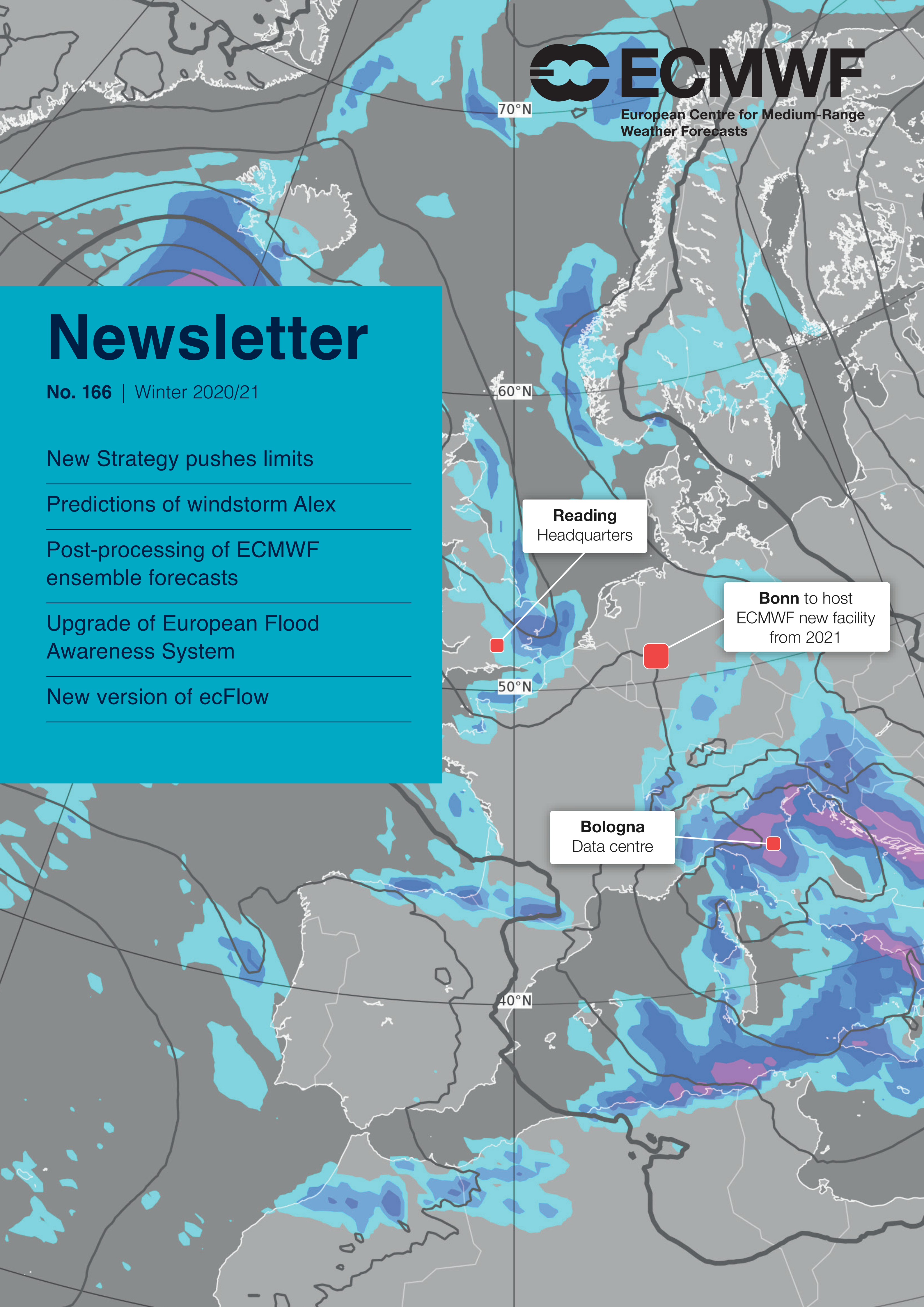
New Strategy pushes limits

Predictions of windstorm Alex

Post-processing of ECMWF ensemble forecasts

Upgrade of European Flood Awareness System

New version of ecFlow



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The ECMWF Newsletter is not peer-reviewed.

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

Guidance about submitting an article and the option to subscribe to email alerts for new Newsletters are available at *www.ecmwf.int/en/about/media-centre/media-resources*

Reading, Bologna, and now Bonn

ECMWF's Council in December chose Bonn in Germany as the site for a new location, where we will focus on our role in the EU's Copernicus Earth observation programme but will also include other research and forecasting activities. Bonn spreads our wings more widely in terms of locations as our next supercomputer will be housed in Bologna, Italy, while our headquarters and most scientific and forecasting activities will remain in Reading, UK. ECMWF thus turns into a multi-site organisation. The move to Bonn became necessary after Britain decided to leave the EU, but it brings fresh opportunities: the new location is in the centre of Europe, with a lot of world-class scientific institutions within easy reach.

Copernicus also features in this Newsletter with a feature article about a major upgrade of the European Flood Awareness System. The latter is part of the Copernicus Emergency Management Service, for which ECMWF provides computational services. The article illustrates well the way in which weather forecasting contributes to flood forecasting. Conversely, flood forecasting provides useful feedback on the quality of weather forecasts. The ultimate goal is to develop a consistent Earth system model with a realistic water budget.

There are many articles in this Newsletter which provide insights into the activities in our core business of medium-range weather forecasting. They range from an article reviewing wind and rain forecasts when storm Alex affected Europe to a review of progress in near-surface forecast biases. An overview of the direction ahead is given by Director of Research Andy Brown in an interview on the new Strategy to 2030. He makes it clear that we have now set our sights on global ensemble forecasts with three- to four-kilometre resolution – a revolution compared to the 18 kilometres still operational today. This ever-finer grid spacing will be accompanied by many developments, among them the advance of machine learning. In this field, we will be seeking to usefully combine data-driven approaches with the strengths of our physics-based forecasting systems. Other key areas are advances in data assimilation, with a lot of new weather observations, and moves to open data.

As the Strategy makes it clear, our numerical weather prediction in the medium and extended range benefits in different ways from running the Copernicus Climate Change Service and the Copernicus Atmosphere Monitoring Service. The reanalysis of past weather, for example, is fundamental to improving weather forecasts. And some of the processes in atmospheric composition are relevant for numerical weather prediction, too. It is this kind of synergy which we will seek to develop, including by moving part of our activities to Bonn.

Florence Rabier
Director-General

Contents

Editor Georg Lentze • **Typesetting & Graphics** Anabel Bowen

New Strategy pushes limits of weather science

The new ECMWF Strategy to 2030 foresees a pushing of the boundaries of weather science to achieve global ensemble forecasts at kilometre-scale resolution. In this interview, ECMWF Director of Science Andy Brown explains how research, forecasting and computing must work together to achieve the goals set out in the Strategy.

What is the purpose of ECMWF's Strategy?

The societal needs for better weather forecasts and climatology are huge, but the science challenges are very big and significant as well. That is why we need a vision on a ten-year timescale in order to prioritise the science we're doing. In terms of the evolving ECMWF products and services for our users and ways of distributing data like the European Weather Cloud, we also need to share an agreed long-term vision about where we're trying to get to.

The Strategy covers ten years but a new one is drawn up every five years. Why?

Very deliberately, the Strategy doesn't change every year. We're setting a multi-year direction, which we hope is robust. Yet, at the same time, to think we have a perfect crystal ball for ten years ahead would be a bit arrogant. So, every five years it's good to take stock, and very often large parts of the previous Strategy will still be valid. That's also the case this time, but there are also exciting new directions where the science, technology or user needs have moved on.

"Operationally we are talking about three- to four-kilometre resolution, in research mode we'll be experimenting with higher still. That's challenging both computationally and scientifically."

What are the main innovations in the new Strategy?

Beyond the continuity regarding Earth system science, seamless modelling, and ensembles, I think there's a real pushing of the boundaries, both scientific and technical. We are moving towards global kilometre-scale modelling. Operationally we are talking about three- to four-kilometre resolution, in research mode we'll be experimenting with higher still. That's challenging both computationally and scientifically.

The strategy also contains a lot on the use of observations. New observations are coming, like EUMETSAT MTG (Meteosat Third Generation), EPS–SG (EUMETSAT Polar System–Second Generation) and further Copernicus Sentinel satellites. We also have to learn how to get even more value out of satellite data. An exciting opportunity is the application of machine learning in our field, where we will seek to combine the best of what data-driven approaches can provide with the huge strengths and physical understanding encapsulated in our existing forecasting systems. We are also developing the tools for sharing the vast amounts of data that we use and produce, through the European Weather Cloud and moves to open data.

"We'll need to move to non-hydrostatic forecasting models. We'll have to do some of the physics differently as we start resolving convection."

What are the main targets for research?

As I mentioned, there will be moves towards higher-resolution ensembles and we will work to overcome the scientific and technological challenges in order to get full benefit from those directions. We'll need to move to non-hydrostatic forecasting models. We'll have to do some of the physics differently as we start resolving convection. And we will continue to increase the range of processes modelled where additional Earth system complexity can be demonstrated to lead to improved forecast outputs.

To get even more value from observations, we need to continue to develop data assimilation and take a more consistent approach across different Earth system components. Also, we are launching a new, big initiative called 'from all-sky to all-surface'. We did a ten-year research programme to be able to use microwave data in cloudy conditions. This all-sky initiative delivered a much more extensive use of microwave data and demonstrated improvements in the quality of the forecast. But we're still less able to use satellite data over land, snow and sea ice. We're envisaging a similar multi-year effort to extract more value out of those observations.

"For extreme temperature anomalies and hydrological impacts, the Strategy calls for skilful predictions up to three weeks ahead. It's a huge scientific challenge."

What are the main targets for forecasts?

The current Strategy is continuing to focus on user-relevant improvements in forecasts. There is thus a continued focus on ensemble forecasting, near-surface weather measures, and extremes. We're continuing to push for more accurate forecasts of high-impact weather into the second week to give better lead-time warning of severe weather. For extreme temperature anomalies and hydrological impacts, the Strategy calls for skilful predictions up to three weeks ahead. It's a huge scientific challenge.

In this Strategy, while medium-range forecasting remains our bread and butter, and while we do seasonal forecasts as well as climate monitoring, we particularly emphasise the extended range, up to 46 days ahead. Doing better in the extended range is also an exacting test of whether we're doing the right thing at the medium range. Hence, we're interested in the extended range as an end in itself because the users are interested in it, but the way we'll make the extended range better will also have knock-on benefits for the medium range.

What is ECMWF's strategic interest in the Copernicus Services?

The Copernicus Services are very valuable to society: the Copernicus Atmosphere Monitoring Service provides air quality forecasts and reanalyses; and the Copernicus Climate Change Service provides weather reanalysis over many decades and the multi-model seasonal forecast. We want to build on that in future Copernicus agreements. We are also looking to expand into monitoring and verification support for CO_2 emissions. And we operate the global and European flooding services, GloFAS and EFAS, and the fire danger forecast service as part of the Copernicus Emergency Management Service.

There's huge value to us scientifically in the synergies between the Copernicus work and the work we do for numerical weather prediction in the medium and extended range. The reanalysis of past weather, for example, is fundamental to our efforts to make the weather forecast better. Over the period of this Strategy, more and more of the processes that are of interest to Copernicus, for example in atmospheric composition, become relevant for numerical weather prediction, too.

"There are likely to be more hybrid architectures in the future, CPUs and GPUs for example. There's an opportunity for us to exploit these new technologies."

What developments in computing are important for the next ten years?

One of the reasons why our forecasts have got better over many years is bigger computers to let us run higher-resolution models and use more complex algorithms. That is why we continue to need more computing capacity. There are likely to be more hybrid architectures in the future, CPUs and GPUs for example. There's an opportunity for us to exploit these new technologies, but there's also a challenge with adapting our codes to take advantage of them.

We are very actively working to be ready for a future procurement to use these hybrid architectures. We also try and develop our codes to be more flexible. As well as adapting to the known changes, we've got a longer-term strand of work trying to make ourselves resilient to as yet unknown future evolutions in computing.

"We'll be thinking over the coming years of how to take advantage of the opportunities offered by remote meetings."

What are the main changes in education and communication?

It's clearly part of our role to educate and communicate, so we have a very active programme of seminars, and workshops and courses. And we will continue to expand that. I think there are interesting opportunities. The recent COVID crisis has forced us all to work remotely, but through many of our events being entirely virtual, we have seen an increased reach.

We'll be thinking over the coming years of how to take advantage of the opportunities offered by remote meetings when we can come back together. I think a hybrid approach is called for, with a central hub and people on site while making sure that a wider audience is fully engaged. This will also help to improve environmental sustainability by decreasing our carbon footprint.

What about partnerships and collaborations?

We have huge numbers of partnerships across everything we do, both at the user end and on the science and technology side. We can't have all the skills and expertise ourselves, so we work with partners in many universities and research institutes around Europe and the world. Many of our Member and Co-operating States have strong and active research units. We work, for example, very closely with some of the limited-area model consortia, with whom we share code.

We have the Fellows programme, which gives us access to some of the top expertise worldwide. And we work very closely with EUMETSAT, ESA (European Space Agency) and the WMO (World Meteorological Organization). For example, we're reliant on the exchange of data that's coordinated through the WMO, and we're strongly represented in the science side of the WMO.

"People are at the heart of what we do. I think the Strategy rightly elevates people to one of its three headings."

What are the Strategy's key messages about organisation and people?

We know that our forecasts have improved for three reasons: new and better observations, bigger computers, and scientists developing new techniques and new code. It's not all about hardware and infrastructure. People are at the heart of what we do. I think the Strategy rightly elevates people to one of its three headings.

How do we attract, recruit and train a diverse workforce while overcoming the challenges of multiple sites that we're moving to? I think an absolute commitment to our people, finding efficient ways of working that are flexible, in a diverse organisation across multiple sites, is crucial to our success.

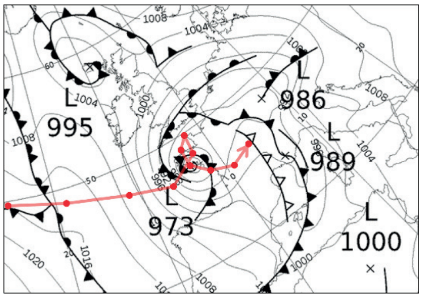
ECMWF's new Strategy. The Strategy document is available for download from *https://www.ecmwf.int/en/about/what-we-do/strategy*.

Windstorm Alex affected large parts of Europe

Linus Magnusson, Tim Hewson, David Lavers

At the beginning of October 2020, storm Alex affected large parts of western Europe, bringing wind gusts above 50 m/s to northwest France and precipitation locally more than 500 mm/24 hours to northern Italy and southeast France that resulted in multiple fatalities. As Alex's centre decayed over France on 3 October, new cyclonic developments on the northern flank brought the wettest day on record for the UK.

Storm Alex developed from a warm front wave moving east-southeast across the North Atlantic into the circulation of a large-scale upper trough. The surface cyclone rapidly intensified south of Ireland on 1 October below the 'left-exit' of a jet stream. Subsequently, a sting jet feature seems to have formed just south of the cyclone centre, hitting the coast of Brittany in France around 00 UTC on 2 October with extreme wind gusts (see top right chart). While the main cyclone centre stayed near Brittany over the next 24 hours, its cold front progressed toward the southeast and caused local flooding and strong wind gusts in southwest France and Spain. Ahead of the front, strong water vapour flux in the form of an atmospheric river was situated across the Mediterranean Sea from the Straits of Gibraltar to southeast France and northwest Italy; this led to torrential rainfall a little way inland from the French/Italian coast and over the southern Alps, primarily overnight from 2 to 3 October. In this article we will discuss the prediction of both the wind gusts over Brittany and the rainfall over southern France and northern Italy.



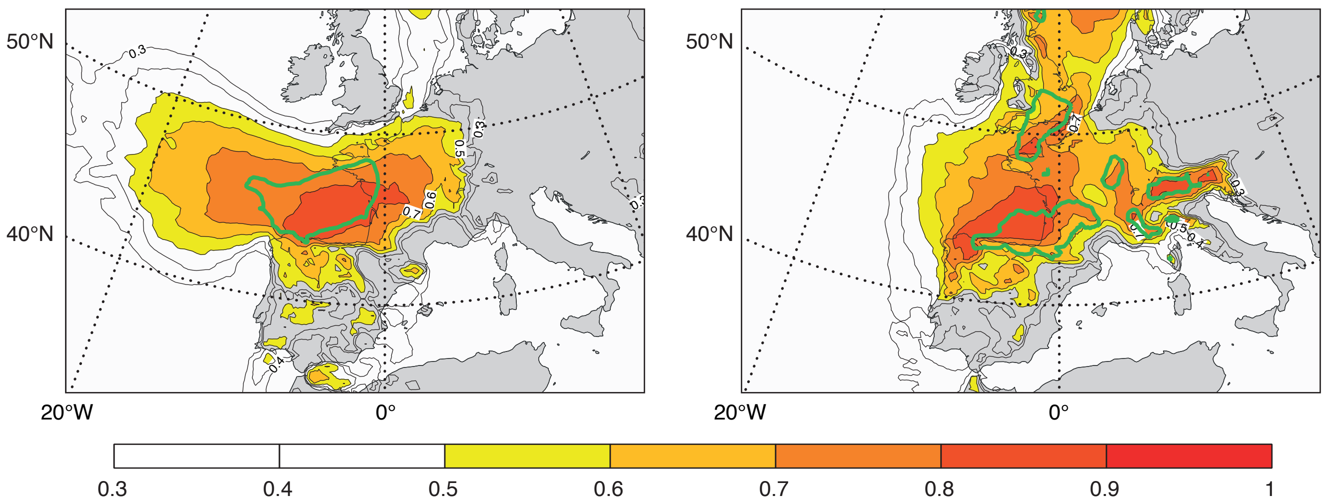
Surface analysis. This is the UK Met Office surface analysis for 18 UTC on 2 October 2020. Alex is the 973 hPa cyclone. The red line shows the complete track of Alex as inferred from these charts. It starts at 00 UTC on 1 October and ends at 12 UTC on 3 October, with dots at 6-hour intervals.

Wind gust predictions

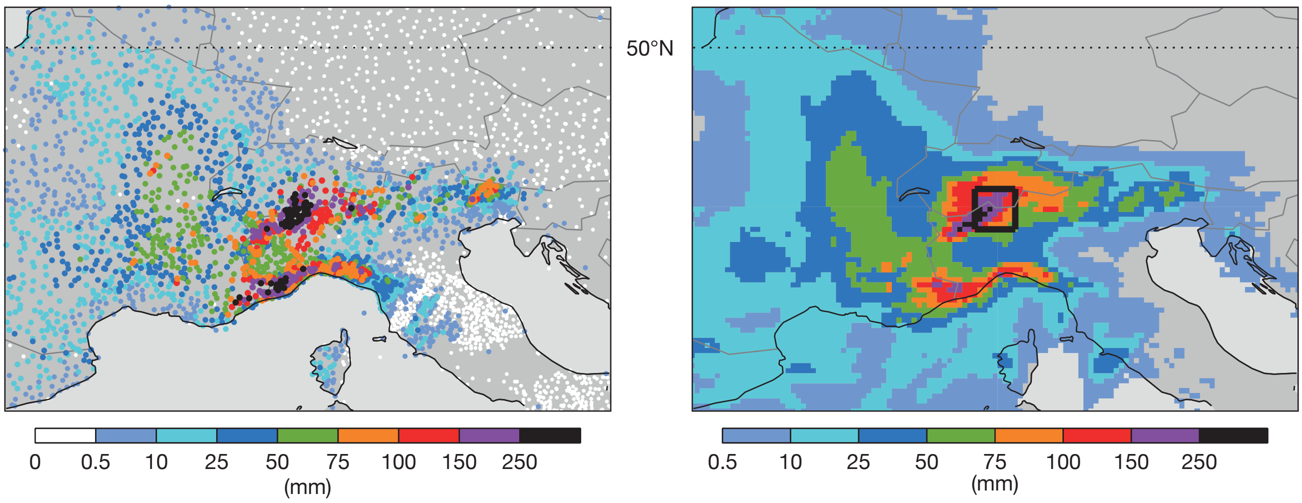
Starting with the wind over Brittany, the most severe band of gusts was associated with the tip of the back-bent front, and as such had some of the hallmarks of a sting jet. The cloud structure associated with the feature was visible in satellite images and was also present in ECMWF's short-range high-resolution (HRES) forecast. The maximum wind gusts in forecasts 3 days before (29 September 12 UTC) and onwards ranged from 44–51 m/s, compared to 51 m/s in observations. Such predictions give confidence that the model can simulate this type of fine-scale feature. Going further back to forecasts from earlier dates, one sees the forecast signal sharply decrease. In those earlier forecasts, strong wind gusts were predicted, but not at the observed levels. Indeed, those gust forecasts had more broadscale characteristics, more reminiscent of a cold jet phenomenon, and linked to a larger cyclonic circulation over the Bay of Biscay which lacked the intense gradients found near the core of Alex (see the Extreme Forecast Index in the left-hand panel below).

Precipitation predictions

For the extreme precipitation over southeast France and northern Italy, even the short-range forecast had problems in capturing the absolute magnitude of the event, especially in the mountains along the coast. This was also the case for the southern Alps. The average 24-hour precipitation up to 3 October 06 UTC



Extreme Forecast Index. Extreme Forecast Index (EFI) for 24-hour maximum wind gusts on 1 October (left) and precipitation from 00 UTC on 2 October to 00 UTC on 5 October (right) in forecasts from 27 September. The plots also show areas as green lines where the values are greater than 0.9 in the one-day EFI maximum wind gust forecast on 1 October and the three-day EFI precipitation forecast starting on 2 October.



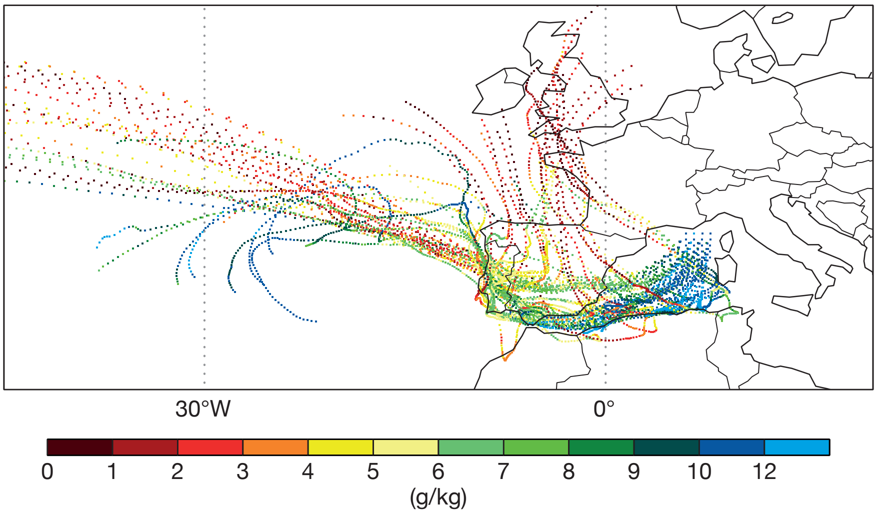
Precipitation observations and forecast. 24-hour accumulated precipitation up to 3 October 06 UTC from observations (left) and the corresponding HRES run from 2 October 00 UTC (right). The black rectangle shows the 1x1 degree box referred to in the text.

among 79 stations (using a high-density network provided by the ECMWF Member and Co-operating States) inside a 1x1 degree box was 223 mm (but individually ranging from 22 to 615 mm). This compares with 163 mm predicted by the corresponding HRES forecast starting on 2 October 00 UTC, averaged over the corresponding grid boxes (see observations and forecast above). Comparing all forecasts in the 4 days leading up to this event, we find that HRES and ensemble (ENS) control forecasts systematically predicted more precipitation than the ensemble median. One could speculate that this is a result of model perturbations becoming asymmetric under near-saturation conditions, as it is not possible to sustain a perturbation that results in a super-saturated state.

Looking at the evolution of the forecast with decreasing lead-time, we find that the signal for this extreme rainfall event gradually became stronger, even if we see a slight jump on 29 September that coincides with the much clearer jump for the wind gust event (see the right-hand Extreme Forecast Index chart on the previous page). The extreme rainfall was due to strong advection of moisture from the Mediterranean Sea towards the southern Alps. Many analysed airmass trajectories had in fact a maximal sea passage, from Gibraltar eastwards and then on into France/Italy (see chart below). In the 6-day forecast, the atmospheric river was captured but the main flow towards the Alps was from the west instead of the south. This meant the airmass mainly passed over land instead of over sea. One could suspect that forecasting the position and shape of the trough was critical to capture this maximal sea passage, and any shift in direction would have decreased the moisture uptake.

Outlook

To summarise the predictability aspects for this event, the medium-range signal for the wind event originated from a broad-scale upper trough, while it was much more challenging to capture the rapidly developing small-scale cyclone that seems to have resulted in a sting jet. For the precipitation event, the shape of the upper-level trough is believed to have played a critical role both for the flow over the Alps and for the path the air took over the Mediterranean to pick up moisture. Important model aspects here include sea surface evaporation and other boundary layer processes as well as physical processes associated with orographic rainfall. An open question remains regarding whether this type of precipitation event over the Mediterranean can have predictability on the sub-seasonal time scale, a topic currently under research in the H2020-CAFE project.



Specific humidity analysis. Backward air trajectories ending off the Mediterranean coast on 2 October 12 UTC between 925–850 hPa, based on ECMWF analyses and calculated with LAGRANTO, a software package to calculate parcel trajectories in the atmosphere provided to ECMWF by ETH Zurich. The colour of the trajectories indicates the specific humidity.

EUMETNET convection-permitting ensemble database hosted at ECMWF

Alfons Callado-Pallarès (AEMET, Spanish Meteorological Agency), Chiara Marsigli (DWD, Deutscher Wetterdienst), Francesca Marcucci (COMET–Italian Air Force Operational Center for Meteorology)

An archive of convection-permitting limited-area model ensembles of European countries has been established at ECMWF, as part of the cooperation within the SRNWP–EPS (Short-Range Numerical Weather Prediction – Ensemble Prediction System) project of EUMETNET. The main goal of the SRNWP-EPS project is to foster and coordinate research and collaborations on limited-area model ensembles (LAM-EPS) to improve convection-permitting ensembles. The focus of these collaborations is on improved prediction of high-impact weather events (HIW) in the context of a European framework bringing together all European LAM-NWP consortia (ALADIN, HIRLAM, COSMO, LACE, Met Office partners, etc.).

Participants in the LAM-EPS archive

MOGREPS (UK Met Office)

MEPS (MetCoOp of the Finnish Meteorological Institute, Norwegian Meteorological Institute and Swedish Meteorological and Hydrological Institute)

γSREPS (AEMET – Spanish State Meteorological Agency)

IT-EPS (ItAF-REMET – Italian Air Force, Department of Meteorology)

IREPS (Met Éireann – Irish National Meteorological Service)

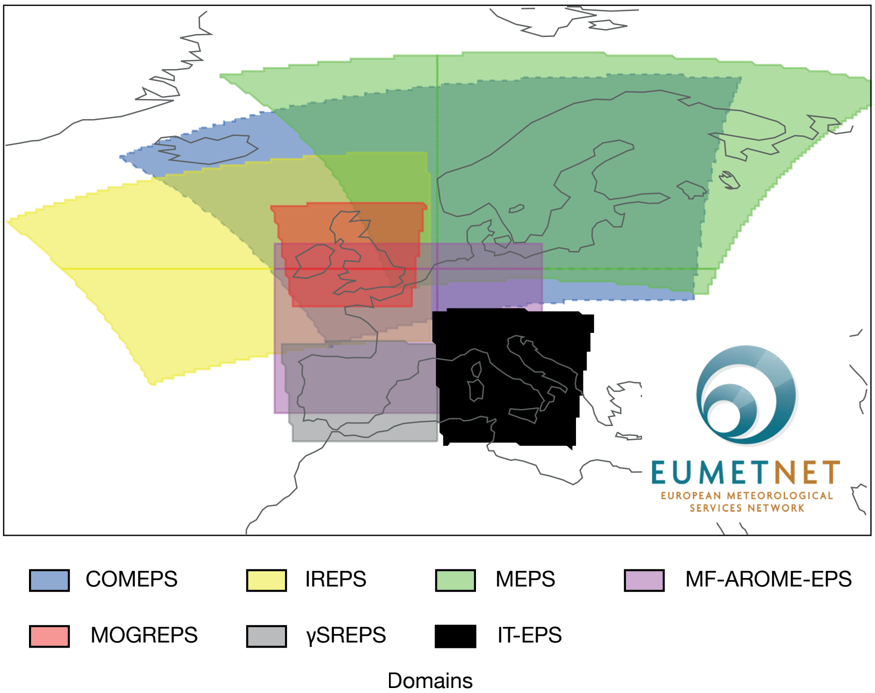
COMEPS (DMI – Danish Meteorological Institute)

MF-AROME-EPS (Météo-France)

RMI-EPS (RMI – Belgian Royal Meteorological Institute) – not running currently

ICON-D2-EPS (DWD – German Weather Service) – not yet archiving

ECMWF has been chosen for hosting the archive since it provides many facilities to access its archiving system (ECFS), which is accessible to all the SRNWP-EPS participants. Furthermore, ECMWF provides user support for its Member and Co-operating States, and access to all the necessary tools and support for archives and retrievals as well as other useful tools, such as clipping and re-gridding onto common overlapping areas and grid types.



Domains archived at ECMWF. The seven convection-permitting LAM-EPS domains archived into the database hosted at ECMWF.

Operational aspects

The decision to create an off-line LAM-EPS database was proposed during the annual workshop of SRNWP-EPS in October 2019 by the Expert Team. The main requirements for this common, flexible set of archives are: i) complete domains to maximise overlapping; ii) GRIB format; iii) two daily cycles with all members; and iv) the most relevant near-surface meteorological parameters (see table). The archive was first set up to cover a three-month period (June–July–August 2020) plus some HIW selected cases.

At the time of writing, nine participants have agreed to archive their ensembles (see list). The first seven participants have already successfully archived the first three-month period. The figure shows and highlights the different domain locations and sizes of the LAM-EPSs.

The database will facilitate different research activities to be carried out: benefits of multi-model ensembles on ensemble spread, inter-comparison, diagnostic studies of physics perturbations, and case study analysis. At the last Joint EPS-Postprocessing Workshop in October 2020, it was decided that research groups with common domains and/or common research interests would coordinate their research activities and form sub-groups. Two main sub-projects were presented in the workshop. They are the first pieces of research using the database: a) the Met Office, Météo-France, DMI and Met Éireann identified 11 high-profile and difficult cases to test the value of using multi-model ensemble data; and b) Météo-France performed an ensemble verification inter-comparison between the French AROME-EPS and the UK's MOGREPS-UK in their overlapping zone.

This EUMETNET SRNWP-EPS convection-permitting LAM-EPS database is a much-simplified continuation of the former TIGGE-LAM archive, with two main differences: the LAM-EPS is based on high-resolution (around 2.5 km) non-hydrostatic convection-permitting NWP models, and the archive has been set up only for European LAM-EPSs. The database has been made possible due to the very lively cooperation in the field of ensemble forecasting in Europe. Since 2013, EUMETNET has supported LAM-EPS cooperation through different phases. It formally started the SRNWP-EPS project in July 2015, where tasks dealing with ensemble calibration and post-processing are performed. On top of the plans for this research database, additional coordinated experiments aim to understand the behaviour of the different LAM ensembles, particularly with respect to model perturbations. Annual meetings offer the opportunity to exchange results and open new lines of research. The archive is aimed to be open to other EUMETNET projects, such as C-SRNWP, for verification, post-processing and nowcasting, and to the European and international research communities. For instance, a collaboration with the WMO HIWeather project is envisaged.

Main near-surface parameters archive

Pmsl	Mean sea level pressure
Psurf	Surface pressure
T2m	Screen level temperature
T2mMAX and T2mMAX	Daily maximum and minimum 2 m temperature
U10m and V10m	The two components of 10 m wind
G10m	Wind gust speed at 10 m
RH2m and/or Q2m	Relative and/or specific humidity at 2 m
AccPcp	Total accumulated precipitation from lead time 0
TCC	Total cloud cover
VIS	Visibility
LTG	Lightning

Additionally, there are convection related parameters, such as hail and graupel, and constant ones, such as orography and land-sea mask.

A look to the future

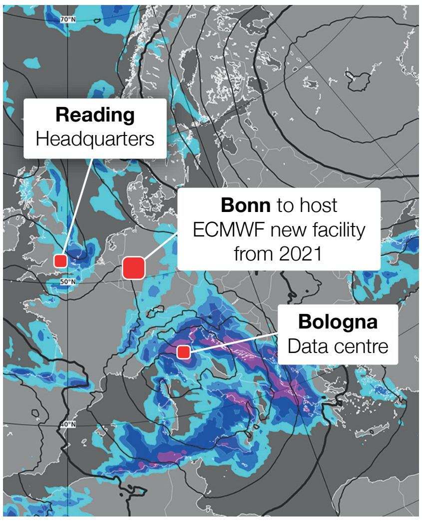
The LAM-EPS database is being extended to archive continuously from 1 June of 2020 to 31 December 2023 (i.e. until the end of the current EUMETNET 2019–2023 third phase). On top of relevant surface variables, and fostered by the first research activities, a few isobaric vertical levels will be archived at the discretion of each centre to diagnose differences in upper-level forcings. These could drive on some occasions distinct LAM-EPS in a different way. It is expected that this database will keep on growing in the future, for instance by more European LAM-EPSs participating. Finally, other research activities are expected to take shape, such as common verification guidelines and benchmarking new LAM-EPS configurations to improve ensemble spread for HIW phenomena such as fog, storms, winter cases, etc.

Bonn to host new ECMWF premises in 2021

ECMWF will start to welcome staff in new offices in Bonn, Germany, from the summer of 2021 while its headquarters will continue to be in Reading, UK, and its data centre will be in Bologna, Italy. The new location will focus on activities that ECMWF conducts in partnership with the EU, but some of the Centre's core activities will also move to the new facility.

Whilst the first wave of staff will be accommodated in temporary offices, it is expected that new and permanent ECMWF offices will become available in Bonn in 2026. This move to a multi-site operation aims to foster even more collaboration across Europe. The new offices to open in Bonn will provide ECMWF with a very favourable central location in Europe as well as a high density of world-class scientific institutions in the immediate surroundings, not only in Germany.

With ECMWF's partnership with the European Union growing over the years, and especially with its role in the EU's ground-breaking Copernicus Programme, the associated activities had to be moved to a location compatible with EU funding policies relevant to them.



Progress in investigating near-surface forecast biases

Irina Sandu, Thomas Haiden, Gianpaolo Balsamo

In autumn 2017, ECMWF started an internal project focusing on 'Understanding uncertainties in surface–atmosphere exchange' (USURF). The aim was to investigate the systematic forecast biases in near-surface weather parameters, disentangle their sources and identify ways to reduce them in the future. Three years on, as this investigation has drawn to a close, its main conclusions and recommendations for future model development, verification and diagnostic work have been summarised in a recent ECMWF Technical Memorandum (https://bit.ly/387SXyu). They were also presented to the autumn sessions of ECMWF's Scientific and Technical Advisory Committees and received positive feedback from representatives of our Member and Co-operating States.

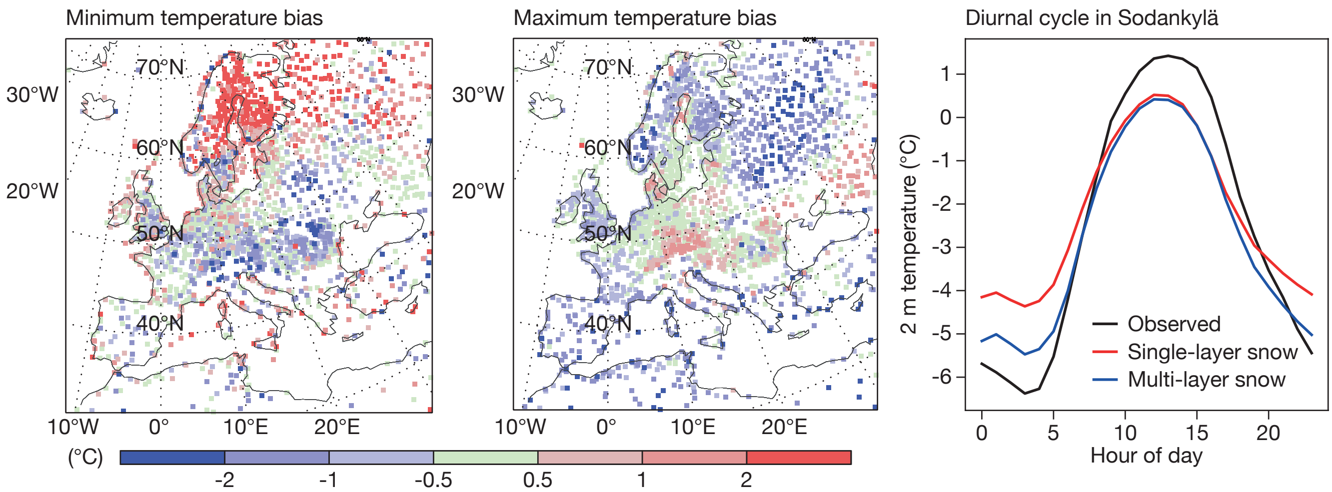
Progress made

The work carried out in USURF is of interest to ECMWF's Member and Co-operating States not only because it aims to improve forecasts of near-surface weather parameters, but also because systematic biases in these parameters are one of the major open issues in the wider weather and climate modelling communities. These biases, which manifest themselves at all forecast ranges, limit predictive skill from hours to seasons ahead. Eliminating or at least reducing them is becoming increasingly important in the context of an enhanced demand for more accurate near-surface weather forecasts. This demand is driven by various interests, such as renewable energy applications or the occurrence of more intense and frequent extreme events.

Systematic biases in near-surface temperature, humidity or winds are the result of a complex interplay between (i) processes parametrized in the atmospheric and surface columns of the forecasting system, which can lead to locally generated errors, and (ii) advection, which constitutes a non-local source of errors. Understanding the leading causes of these systematic errors, which often have complicated geographical patterns and temporal structure, is a necessary step to improve the realism of the model and enhance the near-surface forecast accuracy.

Substantial progress in this direction has been made in the USURF project by using a combination of methods, including conditional verification, process-based diagnostics at observational supersites and model sensitivity experiments. It was shown that, although near-surface forecasts have gradually improved in the past decades, systematic biases with often complicated spatio-temporal patterns remain. As detailed in the Technical Memorandum, answers were found to a number of questions related to the sources of biases in temperature (e.g. cold/warm bias over southern/northern Europe during wintertime), dew point and winds. A work plan to address some of these issues was defined. This will include, among other developments, an improved representation of the snow (figure) and a revision of the land cover and vegetation maps, accompanied by a retuning of uncertain parameters in the surface–atmosphere coupling.

USURF has also provided further evidence that increases in near-surface forecast skill in a complex Earth system model critically depend on two things: the availability of comprehensive observations, and in-depth studies using process-based



Temperature biases. The first two figures show spatial maps of March–April 2014 daily minimum and maximum temperature biases for the ECMWF operational system at a lead time of 2 days. The third panel shows the March–April 2014 mean diurnal cycle of 2-metre temperature at Sodankylä (Finland) in observations and in the ECMWF forecasting system with single-layer and multi-layer snow. It shows that the underestimation of the amplitude of the diurnal cycle of 2-metre temperature is due to a lack of sensitivity to changes in radiation, which is partly the result of using a single-layer snow model. (Figure from the ECMWF science blog, https://bit.ly/3i7AR4a) diagnostics that can correctly attribute model error. Ongoing improvements to the diagnostic and verification tools used at ECMWF are therefore an important contribution towards further enhancements of forecast skill, alongside model developments.

Outlook

Advances in the surface–atmosphere interactions and hydrometeorological processes at increasingly high resolutions towards kilometric scales are relevant not only to ECMWF but equally to all Member and Co-operating States. To accelerate progress, work on related topics could be carried out in partnership with colleagues from the Member and Co-operation States and could for example be an area of interest for secondments at ECMWF.

First Alexander von Humboldt research fellowship starts at ECMWF

On 1 October 2020, Tobias Becker joined ECMWF for a period of two years after being awarded a Feodor Lynen Research Fellowship by the Alexander von Humboldt Foundation (https://www.humboldt-foundation.de/en/). He is a guest of Irina Sandu, an ex-Humboldt fellow at the Max Planck Institute for Meteorology, after having written a successful proposal on the interaction of deep convection with its environment in global km-scale simulations (1–4 km). The prestigious Humboldt Foundation supports research in Germany by promoting academic cooperation between excellent scientists and scholars from Germany and from abroad. One of their initiatives, the Feodor Lynen Research Fellowship, funds young post-doctoral scientists from Germany to carry out a research project abroad as a guest of one of more than 30,000 Humboldt Foundation alumni worldwide – the Humboldtians.

Deep convective systems

Tobias will work with Peter Bechtold, Irina Sandu, Nils Wedi and others to improve the representation of moist physics, and in particular of deep convective systems, in the Integrated Forecasting System (IFS). He will specifically look at km-scale resolutions that will be used for ensemble forecasting at ECMWF in the second half of this decade. This work will build on the ideas developed during Tobias' visit at ECMWF in the summer of 2019, while he was a post-doc at the Max Planck Institute for Meteorology.

During his visit in summer 2019, Tobias looked at the intensity and propagation of deep convective systems over Africa. He used IFS simulations with 4 km resolution, with and without a parametrization for deep convection, and identified shortcomings in both model setups. While rain events are too strong with explicit deep convection, they are generally too weak and too frequent with parametrized deep convection. This is partly because convective inhibition (CIN) barely increases after a rain event. With parametrized deep convection, deep convective systems also often propagate eastward over West Africa. This behaviour is much less common in observations, where the great majority of deep convective systems propagate westward. These results, summarized in a recent ECMWF science blog (see https://bit.ly/2VYf7xb), led to the idea to include the total advective moisture convergence in the convective instability closure. With this revised parametrization, the characteristics of deep convection are in between the two initially studied model versions: deep convective systems get more intense and propagate more realistically westward over West Africa. This revision to the deep convective parametrization is now part of the major moist physics update in an upcoming IFS cycle and was discussed in more detail in the Summer 2020 ECMWF Newsletter.

Open questions

Building on this work during his Alexander von Humboldt fellowship, Tobias will use IFS simulations at 1 km resolution as a virtual laboratory (see the recent ECMWF science blog entry https://bit.ly/3gr7wkg). He will investigate how misrepresentations in deep convection feed back both on the local environment and on the large-scale circulation in IFS simulations at coarser resolution (4 to 9 km). His work will focus on three research questions:

Tobias Becker. The Alexander von Humboldt Foundation fellow will work on the interaction of deep convection with its environment in global km-scale simulations.

1. What are the reasons for the overestimated intensity of precipitation in km-scale simulations with explicit deep convection, and how can this be linked to misrepresentations in convective organisation and in the interaction of deep convection with its environment?
2. To what extent are systematic errors in the propagation of mesoscale convective systems and the diurnal cycle a result of misrepresented interactions between deep convection and the environment at 4 to 9 km resolution?
3. What changes to moist physics parametrizations are needed to better represent deep convection at future resolutions used for ensemble forecasts at ECMWF (1–4 km)?

Forecast performance 2020

Thomas Haiden, David Richardson

ECMWF maintains a comprehensive range of verification statistics to evaluate the accuracy of its forecasts. Each year, a summary of verification results is presented to ECMWF's Technical Advisory Committee (TAC). Their views about the performance of the operational forecasting system in 2020 are given in the box.

ECMWF's headline scores are computed as 12-month running averages to filter out the annual cycle and better identify trends in forecast performance. This means that the beneficial effect of new model cycles is fully visible only 12 months after implementation. The first figure shows the significant improvement of upper-air ensemble forecast (ENS) skill due to Cycle 46r1 of the Integrated Forecasting System (IFS), which was implemented in June 2019. The second figure shows that the beneficial effect of this cycle includes surface parameters, specifically a further reduction of the fraction of large ENS errors in 2 m temperature. Substantial improvements are also seen in the precipitation forecast. Compared to forecasts from other global modelling centres, ECMWF has been able to maintain the overall lead in the medium range, both for upper-air and surface parameters. It is worth noting that the medium-range forecast performance of the IFS did not show any obvious degradation due to reduced aircraft observations from March 2020 onwards as a result of COVID-19. The signal was apparently sufficiently small to get masked by natural performance variations within the annual cycle, year-to-year atmospheric variability, as well as the positive effects of new and additional observations and the most recent model upgrade.

Parallel pre-operational testing showed that Cycle 47r1, which was implemented on 30 June 2020, brings substantial improvements in the stratosphere as well as slight improvements in the troposphere. These will be fully visible in the operational scores by June 2021.

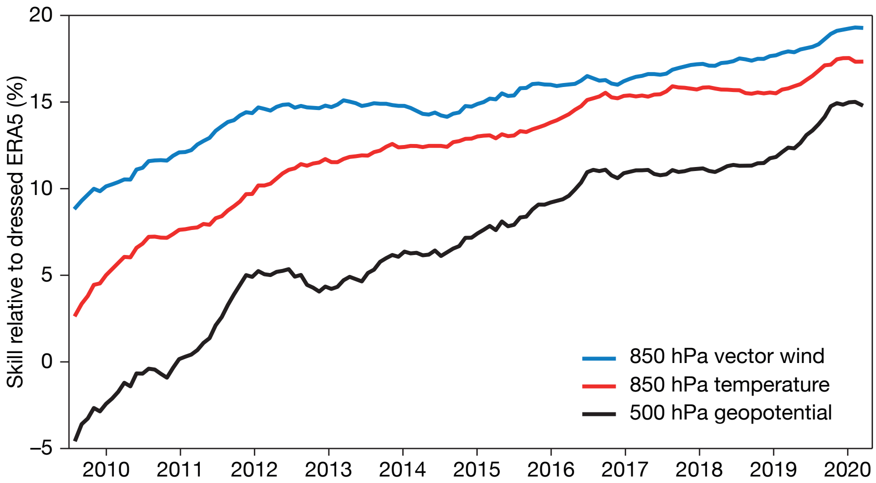
The position error for forecasts of tropical cyclones increased compared to the previous year due to atmospheric variability, as indicated by forecasts based on the ERA5 reanalysis system, which show a very similar increase. High-resolution (HRES) tropical cyclone intensity errors (as measured by the error in central pressure) have reached their lowest value so far. However, the decrease relative to the previous year is also seen in ERA5.

For ocean waves in the extratropics, ECMWF leads other global wave forecasting systems in terms of significant wave height but ranks closer to average for peak period. In the tropics, ECMWF leads in terms of peak period.

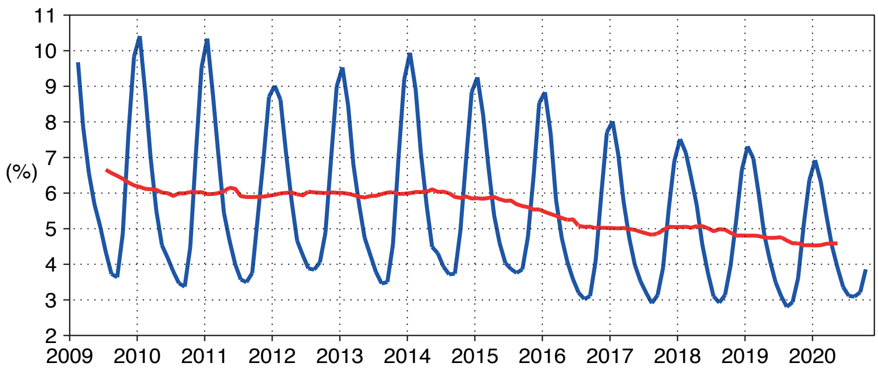
Verification of ensemble forecasts of 2-metre temperature anomalies in week two in the northern extratropics shows a statistically significant positive trend. There has also been a significant improvement in the headline score which monitors ENS probabilistic skill for weekly mean 2-metre temperature in week three. In order to increase the signal-to-noise ratio, this score is based on the verification of re-forecasts. The initialization of the re-forecasts moved from ERA-Interim to ERA5 with the implementation of Cycle 46r1 in 2019, which means that the re-forecast verification now gives results that are closer to the actual real-time forecast skill.

Because of the lack of a strong El Niño or La Niña signal, a very strong positive phase of the Indian Ocean Dipole (IOD) peaking towards the end of 2019 became the main tropical driver of global long-range forecast skill. As a result, 2 m temperature anomaly patterns in boreal winter (DJF 2019–20) were reasonably well predicted over ocean areas, including the North Atlantic. In mid- and high-latitude regions of the American and Eurasian continents, however, forecast skill was lower. An extreme positive anomaly over Siberia (exceeding 1.5 standard deviations in some areas) was hinted at but not captured in magnitude, and the cold anomaly in Alaska and northern Canada was missed.

In spring 2020, both the IOD as well as temperature anomalies in the tropical Pacific returned to close-to-neutral values, leaving the 2020 boreal summer without two strong drivers on seasonal timescales. Forecast skill was accordingly low in many areas, including Europe, where a cold anomaly in the Mediterranean



Upper-air ENS skill improvements. Skill of the ENS at day 5 for three upper-air parameters in the northern extratropics, relative to a Gaussian-dressed ERA5 forecast as a reference. Values are running 12-month averages, and verification is performed against own analysis.



Reduction in the occurrence of large ENS 2 m temperature errors. Evolution of the fraction of large 2 m temperature errors (CRPS > 5K) in the ENS at day 5 in the extratropics. Verification is against SYNOP observations. The 12-month running mean is shown in red, the 3-month running mean in blue.

region and higher than normal temperatures in much of Scandinavia (forming the western extension of a strong Siberian warm anomaly) were missed in the forecast.

The complete set of annual verification results is available in ECMWF Technical Memorandum No. 880 on 'Evaluation of ECMWF forecasts, including the 2020 upgrade', downloadable from https://www.ecmwf.int/en/publications/technical-memoranda.

The following are other sources of information about verification and forecasting system changes.

- Verification pages: http://www.ecmwf.int/en/forecasts/charts
- Inter-comparison of global model forecast skill: http://apps.ecmwf.int/wmolcdnv/
- Ocean wave model inter-comparison results: https://confluence.ecmwf.int/display/WLW/WMO+Lead+Centre+for+Wave+Forecast+Verification+LC-WFV
- A list of 'Known IFS Forecasting Issues': https://confluence.ecmwf.int/display/FCST/Known+IFS+forecasting+issues
- All IFS forecasting system cycle changes since 1985: http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model

Assessment of ECMWF's Technical Advisory Committee, 15–16 October 2020

With regard to its overall view of the performance level of ECMWF's operational forecasting system, the Committee:

a) noted that ECMWF headline scores continue to show high and improving skill, especially in the light of the introduction of 46r1 and 47r1, and particularly for upper air, precipitation and 2 m temperature (away from higher latitudes);

b) noted the lead over other centres has been maintained and acknowledged that scores for a number of elements were the highest ever;

c) recognised that ECMWF maintains an overall lead compared to other centres in terms of ensemble spread and error and acknowledged that there remains under-dispersion in the summer;

d) noted the highest ever medium-range verification scores for EFI wind and acknowledged the recovery of EFI precipitation scores following a recent drop in predictability as shown in ERA5;

e) recognised that ECMWF has an overall lead in verification scores for extratropical ocean wave height but not for peak period;

f) noted that the winter 2019/2020 SEAS5 forecast was good for many, but not all, regions and appreciated attribution studies looking into these errors; recognised the difficulties in producing good summertime seasonal forecasts;

g) noted that ECMWF tropical cyclone track forecast errors were larger than in recent years but central pressure errors were lower, a signal also present in ERA5; appreciated improvement in the tropical cyclone maximum wind–central pressure relationship following 47r1 and welcomed further investigation into extratropical transition;

h) noted the highest ever skill for the northern extratropics in week 2 extended-range output relative to persistence; acknowledged the lack of positive trend for weeks 3 and 4 whilst recognising ECMWF's strategy to improve this;

i) appreciated the introduction of new verification metrics and encouraged the development of further forecaster-relevant verification metrics;

j) welcomed the move to ecCharts-2 and Dashboard-2 and appreciated the addition of new diagnostics as well as both the continuing support for this service and commitment to further improve its efficiency;

k) welcomed the introduction of 47r1, including the extended range of web charts, improvements to convective diagnostics and new tropical cyclone diagnostics;

l) appreciated the opportunity to join the Early Availability of ECMWF Data Pilot;

m) welcomed proposed improvements at 47r2 and 48r1 to the ensemble vertical and horizontal resolution and changes to the extended range to run daily with 100 members;

n) recognised an absence of any obviously detectable degradation in forecast performance in the face of loss of observations due to COVID-19; appreciated efforts to exploit additional, new and novel observations to address possible shortcomings in observational networks;

o) appreciated the continued very good support ECMWF provided to Member and Co-operating States over the last year, particularly in the face of COVID-19 when training and events such as the annual UEF continued;

p) appreciated the training, documentation and feedback processes provided by ECMWF and welcomed future training opportunities and webinars introducing new products and developments.

Update on latest options in ecCharts

Cihan Sahin, Ivan Tsonevsky, Tim Hewson

ecCharts is a suite of web-based applications to inspect, explore and visualise ECMWF forecast data in an interactive way. It now offers nearly 300 layers representing surface and upper-air parameters from ensemble (ENS), high-resolution (HRES), extended-range (ENS Extended) and wave forecasts. Content and functionalities are updated regularly based on requests from users. The latest and previous updates can be accessed on ECMWF wiki pages: https://confluence.ecmwf.int/display/ECCHARTS/Updates.

Among many new additions, the November 2020 update includes medium-range deviations/anomalies, extended-range Cumulative Distribution Function (CDF) diagrams, new vertical wind shear options, and powerful tools to share ecCharts content between users.

Deviations/anomalies

The 'Deviation/Anomaly' layers provide a spatial representation of the deviation of 24-hour forecast fields from model climate (M-Climate) metrics, at all available lead times. Four parameters are available: 2-metre temperature and 10-metre wind speed (both being 24-hour means, computed from four values at 6-hour intervals ending at the selected time) and also total precipitation and total snowfall (both being 24-hour accumulations ending at the selected time).

The layers display the absolute difference, for a user-selected parameter and model run, between forecast values and a user-selected M-Climate percentile. The model run options are: ENS control, ENS mean and HRES, whilst the M-Climate percentile options are: 1, 10, 25, 50, 75, 90 and 99%. The M-Climate distributions used are pre-computed from re-forecast data, partitioned into 24-hour periods in much the same way as the forecast values are. These M-Climate distributions are also used for other products, such as the EFI (Extreme Forecast Index).

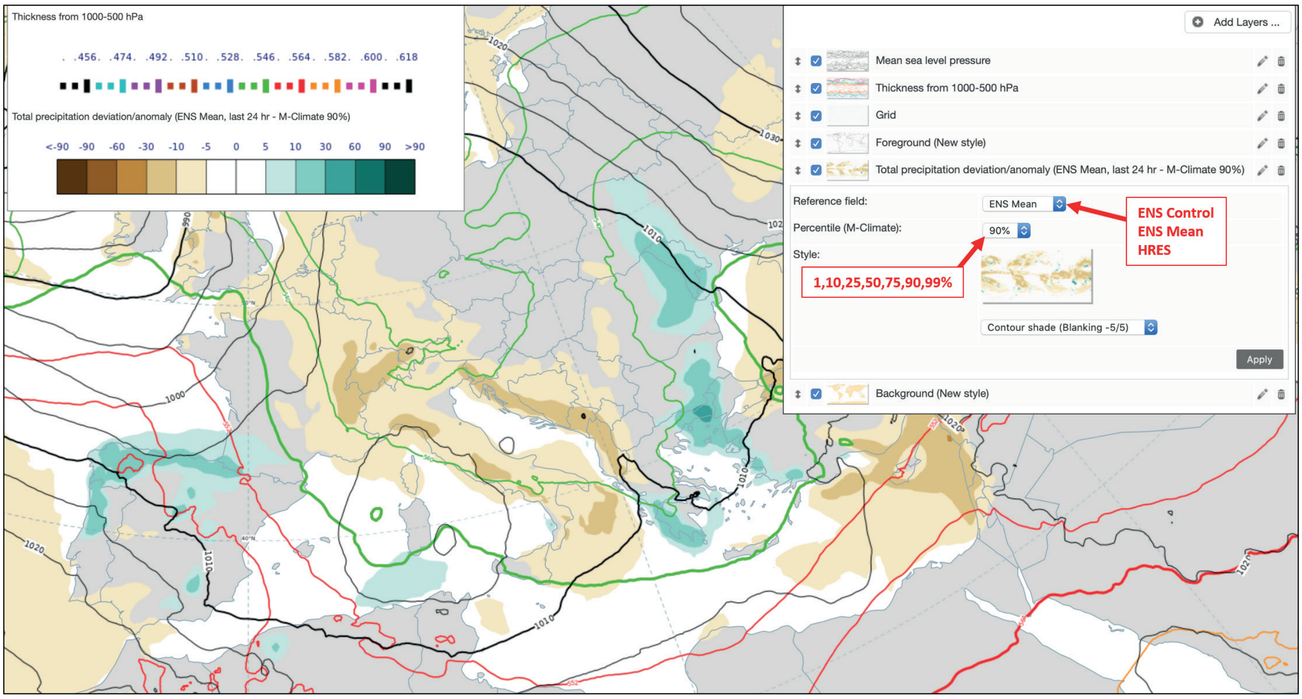
Users can see absolute values of the configured deviation for a site by clicking on the map, or by searching for a city/place.

Extended-range CDFs

To complement the EFI and Shift of Tails (SOT) in the extended range, which was introduced for two parameters with Cycle 46r1 of the Integrated Forecasting System in June 2019, a facility to view extended-range CDFs was also introduced into ecCharts in 2020. Unlike ECMWF's pre-existing CDF-viewing tools, which are used for 24-hour periods at shorter ranges and show absolute values, these CDFs depict anomalies (relative to extended-range M-Climate means). They cover the following parameters:

- 2-metre weekly mean temperature
- 7-day total precipitation

Extended-range CDF diagrams are



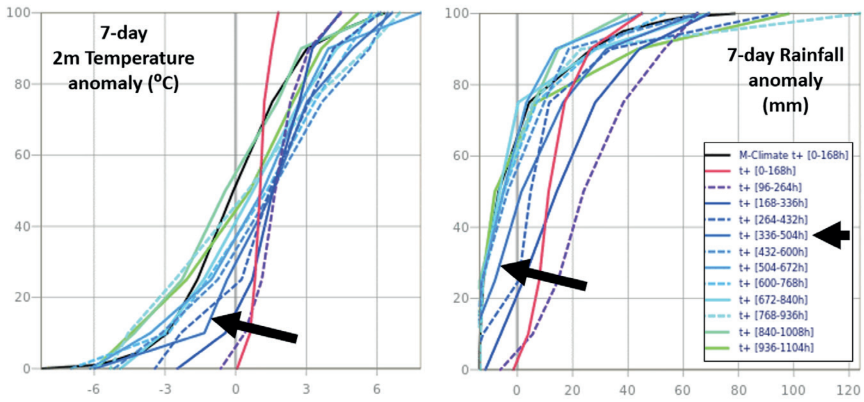
Deviation/Anomaly layer. Rainfall deviation/anomaly (shaded), configured to show the ENS mean forecast versus the 90th percentile of the M-Climate, for 24-hour rainfall ending at 00 UTC on 11 December 2020, 72 hours ahead. Blue shading denotes particularly wet, whilst brown shading denotes not wet (though it may not mean dry). Also shown is an HRES 72-hour forecast, for 00 UTC on the same day, of mean sea level pressure (black) and 1000–500 hPa thickness. Thickness fields suggest that for the particularly wet areas in the eastern Mediterranean/Aegean the cause is largely convection, relating to sea-surface temperature triggering, whilst for Iberia they suggest the cause is primarily fronts and orographic enhancement.

available for all extended-range grid points and for all extended-range time steps. They display all previous and current forecasts validating on the base chart's valid time, which denotes the end of the 7-day period in question. They are available via the menu item 'View > Meteogram window' in the ecCharts interface. They are also accessible from the Dashboard interface menu by choosing 'Add widget > New EFI/CDF meteogram widget'.

A comprehensive description of extended-range CDFs, together with the extended-range EFI/SOT products, can be found in ECMWF's Forecast User Guide: https://confluence.ecmwf.int/display/FUG/Extended+Range+-+CDFs%2C+EFI+and+SOT

Vertical wind shear

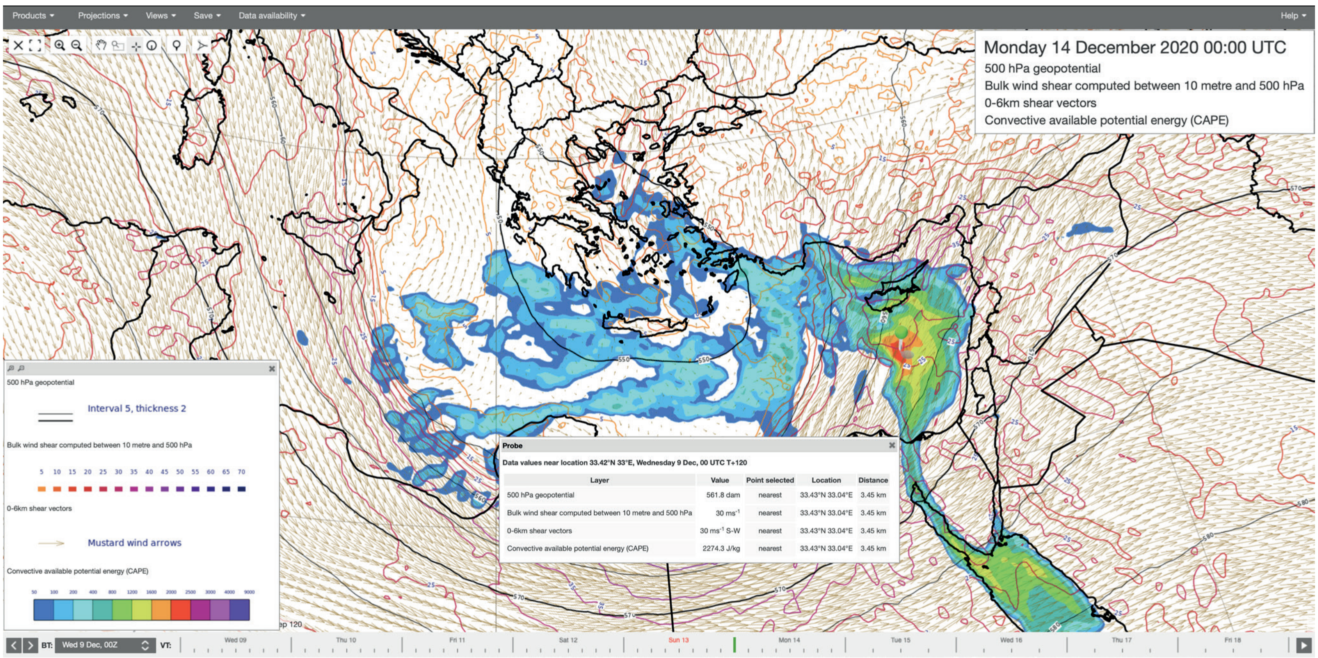
Meanwhile, the 10-metre level has been added as a new level option within the 'vertical wind shear' layer. Whilst previously this could only be plotted as a scalar field (i.e. magnitude), a separate vector layer is now also available, to denote the full wind shear (its magnitude and direction) between 10 metres and 500 hPa (called '0–6 km shear'). These new fields can be combined with CAPE as shown in the third figure, to provide an easy tool to diagnose where CAPE and large deep-layer shear (say over 20 m/s) may favour organised deep, moist convection. The example shows that large CAPE and 0–6 km shear overlap over the eastern Mediterranean in front of the approaching upper trough.



Extended-range CDF diagrams. Extended-range 7-day CDFs for Bologna, Italy, for the week of 7–13 December 2020. The y-axis is the probability to not exceed x-axis values. The x-axis shows parameter values; zeros denote extended-range M-Climate mean values. Those M-Climate distributions are shown as black curves. Colours show forecast distributions from successive extended-range (monthly) forecasts for the same valid week (see legend). Arrows highlight week 3 forecasts (t + 336–504h): there was a weak signal for both wetter and warmer than average, although spread was still large. At shorter leads (dark blue → purple → red) those signals were maintained, and confidence grew (lines became more vertical).

ecCharts sharing

ecCharts sharing functionality has been available to users since the beginning of 2020 as an experimental feature. With this latest update, tools have been improved by taking into account user feedback. The revised tools allow a user to send a product, a projection, a Dashboard widget (a single chart or a meteogram in Dashboard) or a Dashboard tab to other ecCharts users. Once exported and activated by the receiving user, that user will have their own copy. Indeed, an item can be shared with one or many ecCharts users. This is an extremely powerful tool to use in organisations where different users – e.g. shift forecasters – want to all plot data in the same way; it prevents them from all having to go through the same configuration process. Documentation on how to use the sharing feature is available here, under 'Sharing in ecCharts and Dashboard': https://confluence.ecmwf.int/display/ECCHARTS/ecCharts+updates+-+November+2019.



Vertical wind shear layer. Deep-layer 0–6 km shear (colour contours and wind vectors), CAPE (colour shading) and 500 hPa geopotential.

The new interface for ECMWF real-time product configuration

Emma Pidduck, Krzysztof Ściubisz, Cristian Codorean

A new step in ECMWF's product dissemination service has been released to the public to enable users to evaluate and configure their data requirements.

Product dissemination service

ECMWF currently provides real-time data to approximately 220 commercial and research users via its product dissemination service. Each user has their own bespoke data requirements that are delivered by the product dissemination system. The product dissemination system is comprised of four parts, each serving a different stage of the data purchasing process:

- Product Requirements Catalogue (PRC)
- Product Requirements Editor (PREd)
- Product Generation (PGEN)
- Production Data Store (ECPDS)

Three components (PREd, PGEN and ECPDS) saw improvements implemented during 2019 and early 2020, as described in previous articles (Gougeon in Newsletter No. 159 and Zink & Plumridge in Newsletter No. 163). In June 2020, the first step of the product dissemination system, the Product Requirements Catalogue (PRC), was released to the public to enable users to evaluate and configure their data requirements. Previously the product requirements of each user had to be discussed with ECMWF or a Member State before any products could be made available, in many cases requiring back-and-forth revisions to quotes until the final configuration had been reached.

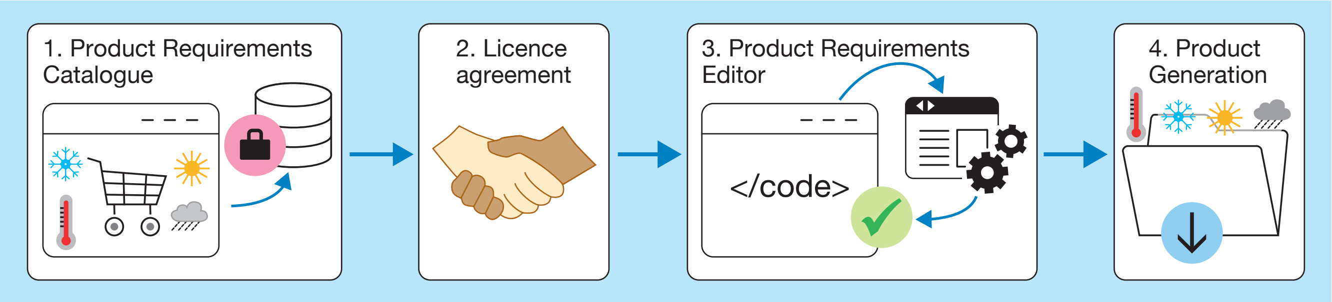
Product Requirements Catalogue

The PRC provides a user-friendly interface to explore and evaluate the real-time products of the ECMWF catalogue, as well as to purchase licences to access the MARS Archive and ecCharts platforms.

The application aims to facilitate a self-service approach and provides many interactive features, such as enabling the user to search for parameters, filter for parameter availability by set and subset, and review the costs and estimated volumes of their bespoke configurations. The connection to the ECMWF parameter database provides parameter descriptions and units from a centralised and maintained database, which also provides relevant links to the Forecast User Guide. A registered user can also configure their own library of pre-set areas. This can be used during product configuration and will intuitively snap to the chosen grid resolution as required, ensuring the user can visualise and understand what will be received in the final delivery.

For ECMWF and the Member and Co-operating States, the PRC also provides a number of additional administrative features. These include provision of the user's final data configuration in the language required for the PREd, folder sorting, and approval or rejection of user requests for licence management. Together, these features enable the user and the licensor to expedite the process between exploring the real-time catalogue, acquiring a licence for real-time or historical data, and subsequent delivery of or access to ECMWF data.

Over time, ECMWF will continue to modify and improve the PRC to further enable users and the licensor to simplify the licensing process. Future features will include the ability to amend licences mid-contract with a clear indication of costs incurred, as well as integration with the ECMWF licence reporting application for the licensor. As the product capabilities expand to support self-service, we envisage that more users will have access to the suite of tools to manage their own requirements.



ECMWF's product dissemination service. Schematic illustrating how users interact with the product dissemination system, including the Product Requirements Catalogue (step 1) and the Product Requirements Editor (step 3). The web-based catalogue reflects the products available for dissemination and provides instant costs and volume estimations for the user's chosen configuration. The agreed data is configured in the Product Requirements Editor. Finally, the Product Generation system uses the product requirements to generate user-tailored data.

ECMWF, ESA and EUMETSAT collaborate in training on atmospheric composition

Chris Stewart, Mark Parrington

An intensive week of engaging lectures by leading atmospheric scientists and hands-on practical sessions in the analysis of satellite data and derived products took place from 16 to 20 November. The event was the second training course on atmospheric composition organised jointly by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), the European Space Agency (ESA) and ECMWF through the Copernicus Atmosphere Monitoring Service (CAMS). It followed a first joint course on atmospheric composition in Cluj, Romania, in November 2019.

Structure and content

The objectives of the training were to present the state of the art in air quality monitoring; to provide an overview of observations, modelling, data assimilation and applications; and to enhance the capacity to access and analyse data. The training course also aimed to foster collaboration amongst participants and encourage the development of individual and team projects using the data, resources and methodologies presented.

Thirty early career scientists and data users, from 20 countries and from both academia and industry, attended the training. A good gender mix was achieved with 12 female and 18 male participants. These were mainly from the member states of the hosting agencies, including associated states, but a few were from other countries to ensure a geographical mix and to encourage the exchange of ideas and networking.

The training took place entirely online, with the mornings dedicated to lectures and discussion sessions, and the afternoons to hands-on practicals. The lectures covered satellite missions and instruments, in-situ networks, trace gas and aerosol retrievals, data assimilation and modelling. These were given by experts from universities, research institutes, space agencies and private industry. ECMWF staff who contributed to the lectures include Vincent-Henri Peuch, Antje Inness, Jérôme Barré and Mark Parrington, with Miha Razinger leading some of the practical sessions. The lectures were publicly streamed on YouTube and accessible to everyone, not only to the 30 selected course participants.

For the practical sessions, participants were divided into smaller groups. The first practical included a primer on the basics of pollution, remote sensing and modelling. On the following two afternoons participants chose between parallel sessions on satellite data processing for atmospheric composition analysis, CAMS data discovery, and exploration of the HARP toolbox. These sessions were carried out using Jupyter Notebooks, with each participant provided with an account on a dedicated JupyterLab environment set up specifically for the training.

On the last two days participants were given the task to develop their own case studies, either individually or in groups. Results were presented at the end and shared on the online wall board, Padlet, to invite comments and stimulate discussion. Some fascinating case studies of participants included analyses of recent events, such as aerosol and trace gas emissions from the California fires, the Raikoko volcano eruption or biomass burning in Australia, using a combination of reanalysis data from CAMS or satellite measurements from instruments such as GOME-2 and TROPOMI. In some cases participants tested the limits of resolution to detect e.g. shipping lanes from atmospheric constituents, or to compare trace gas concentrations over certain European cities before and after the lockdown due to the COVID-19 pandemic. Equally interesting case studies included an analysis of long-range transport of aerosols, a comparison of satellite observations with in-situ measurements, and others.

Feedback

Overall feedback from the course was very positive with participants considering it to be relevant, useful, and with an appropriate balance between theoretical lectures and hands-on practicals. In addressing the fully online nature of the course, participants highlighted the advantages of saving costs in travel and accommodation, and the convenience of tools such as Zoom, Sli.do, Slack and Padlet. However, despite all the efficiencies of online training, participants were almost unanimous in their opinion that they cannot substitute the level of interaction that can be had with face-to-face events.

For more information on the event, please visit https://training.eumetsat.int/course/view.php?id=387.

Participant feedback

"I discovered new commands and the environment of Jupyter that will absolutely enhance my capabilities for data processing. Amazing work!"

"It was great to be given 'recipes' for dealing with the various datasets. This was my expectation of the course and it was met."

"This course has been great, also because it has shown the possibility to have lectures, practical sessions and interactions online, something that was hard to believe just a few months ago!"

Summer of Weather Code in fourth round in 2021

Esperanza Cuartero, Julia Wagemann, Anna Ghelli

ECMWF's Summer of Weather Code 2021 (ESoWC) will launch its fourth edition on 1 February 2021. Innovative weather, climate and atmosphere-related challenges and ideas that will be highly beneficial to ECMWF and Copernicus activities will be proposed on Github, where interested parties can submit a proposal on how to tackle the challenge.

At the core of ESoWC is the provision of innovative and open-source software solutions as well as attracting external expertise and new talents. The application period will end mid-April and the selected teams for the coding phase will be announced at the beginning of May. ESoWC's coding phase between May and end of August is an intense and fruitful collaboration among mentors and developers, leading to a mutual learning experience. The coding phase is the highlight of ESoWC. During this time, the initial proposals are implemented into feasible software developments and innovations.

After the coding period, provided that all milestones and deliverables are fulfilled, the developers will receive a stipend of £5,000. The presentation of the final results will be a virtual event, the Final ESoWC Day, to take place in September 2021.

Innovation and high quality in ESoWC 2020

The last edition of ECMWF's Summer of Weather Code in 2020 included nine open-source projects that once more highlighted the innovative nature of ESoWC. All coding solutions had a high impact; some of them have been earmarked for implementation and further development.

For the first time, ESoWC 2020 was supported by two cloud computing systems, the Copernicus data and information access service WEkEO (https://www.wekeo.eu/) and the European Weather Cloud jointly developed by ECMWF and EUMETSAT (https://www.europeanweather.cloud/). This allowed ESoWC to offer sufficient storage and processing capacities for the compute-intensive projects. The new edition of ESoWC in 2021 will also be supported by both cloud computing systems. ESoWC 2020 projects were at the intersection of machine learning, open-source computing, cloud computing and open data. Four teams explored the use of artificial intelligence in a range of applications:

- DeepGEFF is an example of the **application of deep learning techniques to the science of fire prediction**. The project compared skills between the Global ECMWF Fire Forecasting (GEFF) system and artificial intelligence. The machine learning-based model increased the wildfire prediction accuracy all over the world.
- The Conversational Virtual Assistant (CVA) is a project that **facilitates help desk support to external ECMWF and Copernicus users**. This project developed a chatbot that uses natural language processing and machine learning technologies to assist in finding responses to questions on ECMWF products and services.
- A machine learning project addressed the challenge of monitoring point and contextual **anomalies within the ECMWF production chain system** with the help of artificial intelligence methods and an Anomaly Detection System.
- **A deep-learning approach to improve tropical cyclone prediction** and tracking has created an algorithm able to recognise and classify tropical cyclones based on their intensities, over four classes.

Three teams put a focus on Copernicus open data:

- To predict the dramatic impact of climate extremes, such as floods, wildfires and thawing permafrost, the project UNSEEN-Open developed a **reproducible, open workflow** that allows an assessment of natural hazard events globally.
- **The classification of air quality** project developed a classification scheme to validate and remove outliers from surface air quality observations. This quality control method allowed the comparison of station data with air quality forecasts from the Copernicus Atmosphere Monitoring Service.
- The project DAAQS (Detecting Anomalies in Air Quality Stations) applied machine learning clustering algorithms to **classify similar and dissimilar air quality stations** to measure their reliability and representativeness scores.

Another two teams worked on challenges related to ECMWF's model performance, data storage and archiving:

- The Elefridge project explored the potential of **compressing atmospheric data while preserving real information** to reduce storage and to facilitate data sharing. It provided evidence that the size of climate and weather forecast data archives can be reduced by one to two orders of magnitude without losing valuable information.
- The **HPC performance profiling tool** project aimed to improve the performance analysis of ECMWF's Integrated Forecasting System (IFS). The software application is a graphical, extendable and easy-to-use interface that allows the visualisation of HPC performance data metrics tailored by the user.

More information about ECMWF's Summer of Weather Code can be found at https://esowc.ecmwf.int/ and https://github.com/esowc as well as on Twitter: @esowc_ecmwf.

New way of accessing GRIB data using Julia language

Robert Rosca (EuXFEL), Stephan Siemen, Claudia Vitolo

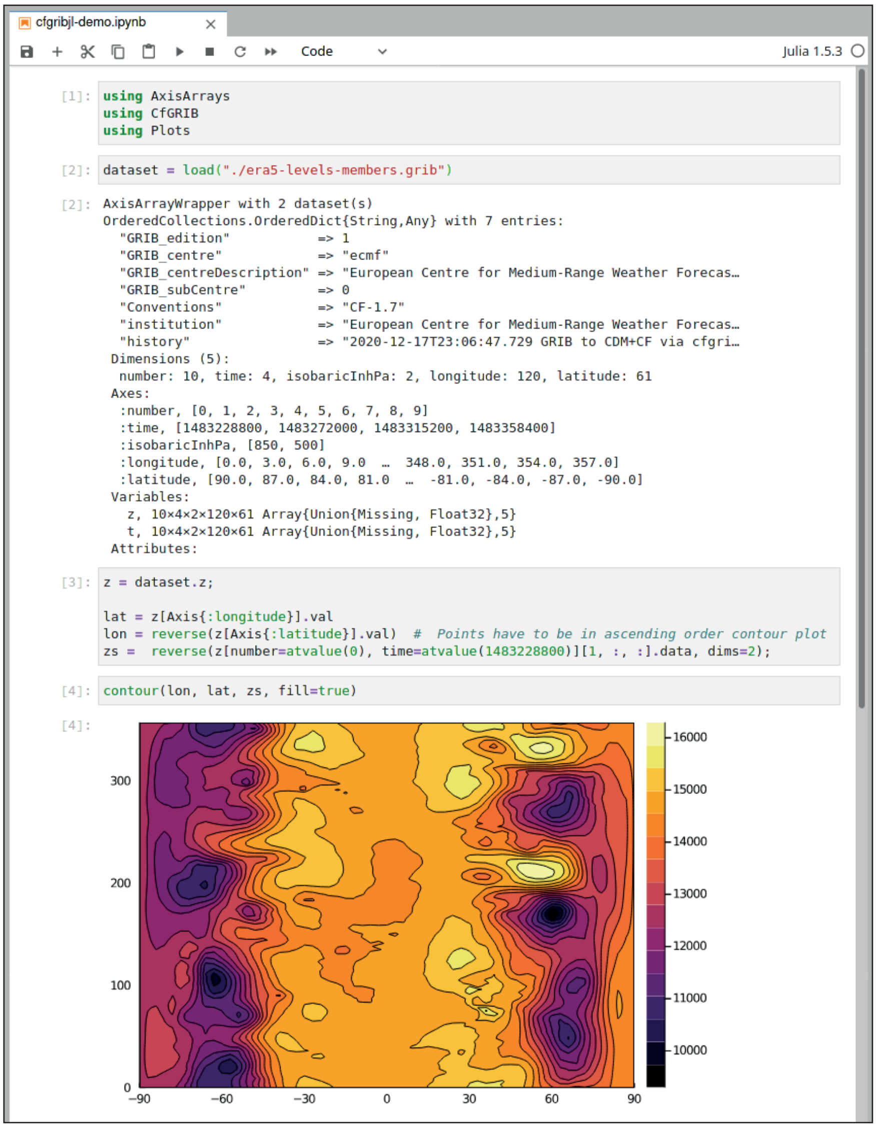
The Julia programming language has become increasingly popular in high-performance computing, including in the climate community. For example, the Climate Modelling Alliance is largely using Julia for its packages. Julia is a language specifically targeted towards scientific computing, with speed, ease of use, and interoperability as its main goals. The question thus presented itself how to make it accessible to the weather community.

Characteristics

The language is still early in its development and adoption, but it has very high ambitions. By using a specialised 'just-ahead-of-time' compilation process, Julia attempts to bring together the best aspects of multiple languages. It builds on top of features that have been proven to work well in the past (straightforward syntax, interactive computer programming environment REPL, mix of dynamic and static typing, multiple dispatch, etc.) and tackles common pain points for scientific computing (speed, parallelism, interoperability, etc.).

Our aim has been to achieve the same high-level interface for Julia as the cfgrib package has been for the Python programming language. The cfgrib Python package reads GRIB data and stores the data in the popular xarray data structure for use in other common scientific packages of the Python ecosystem.

The new cfgrib Julia package is built on the GRIB.jl Julia community package, which already uses ecCodes for the access to GRIB files in a message-by-message style. CfGRIB.jl aims to offer a more user friendly interface to the data – in the same way that the cfgrib Python package exposes the GRIB data in a Python xarray, CfGRIB.jl exposes the



Julia notebook. A short notebook showing an example of loading and plotting data.

data through a number of labelled multi-dimensional array backends.

As Julia is a young language, no clear community-selected 'standard' labelled array package (like xarray for Python) has been adopted yet. That is why CfGRIB.jl uses a flexible array backend system, allowing rapid adoption and integration of up-and-coming Julia packages as the ecosystem continues to develop.

Julia's flexible multiple-dispatch system allows for easy interoperability between a selected array backend and a plotting or mathematical library. For example, you can load data in to the AxisArrays.jl backend, add uncertainties to it with the Measurements.jl package, do some calculations with the numbers, then use Plots.jl to visualise the results; at the end the uncertainties will have been propagated all the way through to your plots. The code and algorithm interoperability between packages enabled by the multiple-dispatch system is one of the biggest strengths of the language.

Outcomes

To implement a Julia interface at the Centre, ECMWF was keen to get developers from the Julia community engaged. They helped to set up the development by first performing a feasibility study, looking at existing climate and labelled array packages in Julia. They then re-implemented many aspects of the Python cfgrib package in Julia while checking feature parity with automated tests.

The CfGRIB.jl package is available on the ECMWF GitHub space. Since this is a new development, users are asked to test it carefully before using it in an operational setting. We hope this development will be of interest to the wider Julia user community and the start of a wider use of GRIB data. We welcome contributions to the code and documentation. An automatic test setup ensures code contributions can be done safely.

A new tool to understand changes in ensemble forecast skill

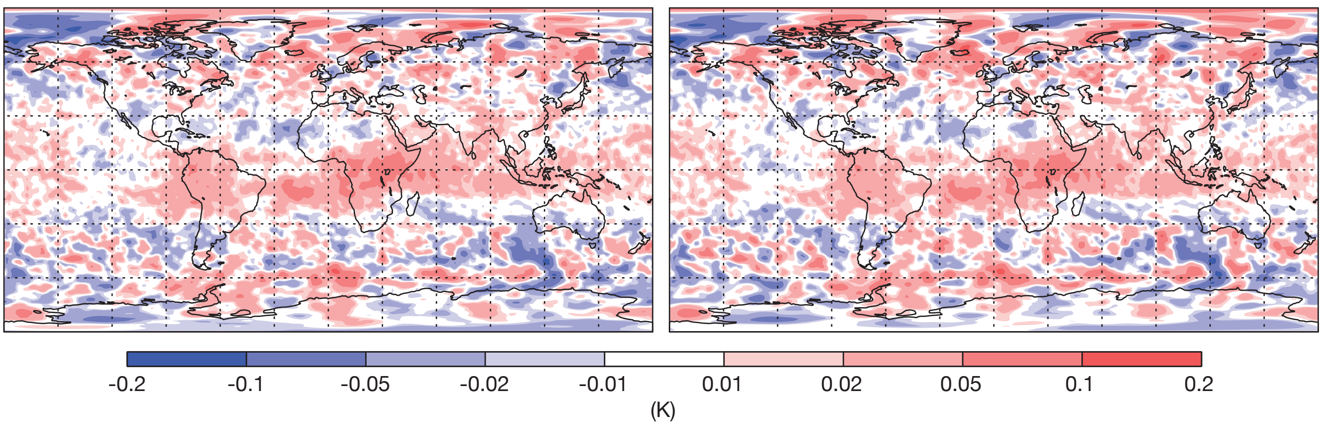
Martin Leutbecher, Thomas Haiden

The continuous ranked probability score (CRPS) is a scoring rule that is popular for assessing the quality of ensemble forecasts. The CRPS can be used to compare different ensemble forecasts and it plays an important role in guiding the development process for forecast systems at ECMWF and beyond. A new tool has been developed that makes it easier to understand the reasons for differences in the CRPS between sets of forecasts.

Outline

The tool consists of two elements. First, the sample of forecast–observation pairs is approximated by a homogeneous Gaussian (hoG) distribution. For each location and lead time, the distribution is characterised by the mean and variance of the error of the ensemble mean together with the ensemble spread. The second element consists of a closed-form expression for the expected CRPS of the hoG distribution. This expression depends on three variables: the ensemble mean error variance, the spread-error ratio and the mean error of the ensemble mean – the bias. Therefore, any difference in the CRPS between forecasts can be linked to changes of these three variables when using this approximation.

In order to examine how well this approximation works, medium-range scores for several upper air variables have been computed and compared with the hoG approximation. Although the actual forecast–observation sample is simplified considerably, the approximation works well. It captures the geographical variations of the CRPS and it gives useful approximations of score changes due



CRPS and its approximation compared. Geographical distribution of CRPS differences (left) and hoG approximation of the CRPS differences (right) between model cycles 47r1 and 46r1 for 5-day ensemble forecasts of temperature at 250 hPa. Each model cycle has been verified against own analysis and the verification period is 16 July – 31 Oct 2019.

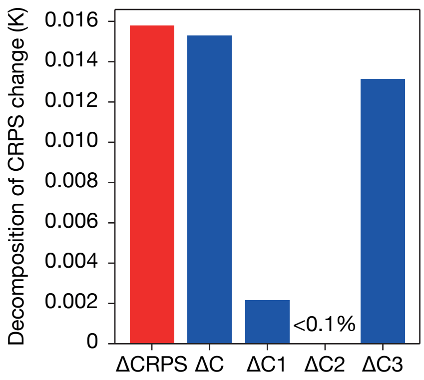
to e.g. changes in bias or ensemble spread. It thus enables users and developers to better understand the reasons for score differences between two ensemble systems. It is planned to add this new diagnostic to existing verification software used routinely at ECMWF for development and monitoring of the forecast system.

Example

A recent example illustrates how the new diagnostic works in practice. Parallel runs with Cycle 47r1 of ECMWF's Integrated Forecasting System showed improvements compared to the previous Cycle 46r1 (see the article comparing the two cycles in Newsletter No. 164). However, in the tropics at 250 hPa, the CRPS for temperature showed a degradation of several percent. The hoG approximation of the CRPS reproduces the geographical distribution of the CRPS changes well, as can be seen in the map plot (first figure). The bar diagram (second figure) shows that for the tropics as a whole, defined here as latitudes between –20° and +20°, the increase in CRPS is well approximated by the hoG model (compare the red and the first blue bar).
The decomposition of the total CRPS change into the three contributions reveals that by far the largest contribution of 86% comes from a change of the bias, 14% is due to an increase in ensemble mean error variance, while changes in spread contribute less than 0.1%.
The absolute change in the bias is small (less than 0.1 K). Yet, it has a considerable impact on the relative CRPS change as the bias is of a similar magnitude as the standard deviation of the error of the ensemble mean.

Deeper implications

In addition to the new diagnostic, this work has potentially deeper implications regarding the criteria used to determine improvements in the representation of uncertainties in ensemble forecasts. At present, the CRPS of raw model output is the main metric guiding the development of medium-range weather forecasts. It will continue to remain a relevant metric as the monitoring and understanding of model biases will always play an important role in model development. However, in the presence of a bias, solely focusing on this metric implies convergence towards a system with spread–error ratios larger than one. In consequence, users who bias-correct their forecasts will end up with an overdispersive probabilistic forecast. To address this issue, it is recommended to also use the CRPS of bias-corrected forecasts for guiding the representation of uncertainties in the medium range. Since extended-range forecasts are already bias corrected, this will have the advantage of making the medium-range and extended-range development more consistent.



CRPS change decomposed. The chart shows the CRPS change for temperature at 250 hPa in the tropics, based on the same data as in the first figure (ΔCRPS, red). It also shows the change in the hoG approximation of the CRPS (ΔC, blue) and its decomposition into contributions from changes in the ensemble mean error variance (ΔC1), changes in the spread–error ratio (ΔC2), and changes in the normalised bias (ΔC3). This decomposition of the approximated CRPS change is exact in the sense that ΔC= ΔC1 + ΔC2 + ΔC3.

Further reading

Further information can be found in an article by Martin Leutbecher and Thomas Haiden, Quarterly Journal of the Meteorological Society, https://doi.org/10.1002/qj.3926.

How to handle errors in satellite data assimilation

Niels Bormann, Bill Bell, Patrick Laloyaux, Karen Clarke

Almost 200 experts from around the world joined a virtual workshop on the treatment of random and systematic errors in satellite data assimilation for numerical weather prediction (NWP), organised by ECMWF and EUMETSAT's NWP SAF (Satellite Application Facility) from 2 to 5 November 2020. Dealing with random and systematic errors in observations and models is at the heart of making optimal use of the wealth of information from satellite data to initialise forecasts or to create reanalyses. Errors and uncertainties arise from many areas, such as in the forecast model, the observations, or the 'observation operators' used to convert model fields to observation equivalents. The challenge is to separate the different errors and to deal with them adequately during the assimilation. The meeting continued the strong tradition of workshops organised jointly with the NWP SAF, fostering further exploitation of satellite data.

One of the aims of the workshop was to connect activities in different communities, spanning data providers, operational forecasting centres and academia, and to identify where NWP can make use of advances made in other fields in the characterisation of observation-related uncertainties. Reanalysis featured as a strong component, where treating biases in observations and models poses its own challenges.

The twenty-three speakers reviewed the strong progress in recent years in several areas. There were lively interactions between participants around the globe. This was in no small part thanks to the dedication shown by the participants, some of whom turned their nights into days to join their European colleagues live, in rich

exchanges through online chats and video-conferencing rooms. Poster sessions, panel discussions and working groups provided further opportunities for discussion.

Characterisation of uncertainties

Several presentations covered the characterisation of observational and model uncertainties which informs our treatment of these in the assimilation. On the topic of instrumental biases, talks covered the latest developments in the on-orbit characterisation of the CrIS hyperspectral infrared instruments and the Aeolus Doppler wind lidar, as well as the ongoing activities of the Global Space-based Inter-Calibration System (GSICS, a World Meteorological Organization/ Coordination Group for Meteorological Satellites initiative). The development of metrology-inspired approaches to the characterisation of observational uncertainties featured prominently in two of the talks, and the topic was revisited frequently in the panel discussion on biases and in working group discussions.

Representation error, resulting from a mismatch between the model representation of the state and observations, was explored in two of the talks. Promising approaches were described involving high-resolution model runs and observations. A talk on the characterisation of uncertainties in historic observations illustrated the value of the long-term perspectives brought by reanalyses.

Correction of observation and model biases

Correcting observational biases is essential for the successful assimilation of many satellite observations, and separating observation and model biases continues to be challenging. Adaptive bias correction methods are now commonly used within the assimilation system to correct for observation biases, and some schemes are developed to address model bias. Several talks discussed the best way to design models of observation bias, with the aim to avoid absorbing model error in the observational bias correction. The concept of scale separation also featured prominently. This applies where model error contains identifiable large-scale structures, which opens a new perspective in the quest to attribute the correct source of biases. Machine learning methods trained on analysis increments or anchor observations also showed some potential to learn the correct structure of model biases.

All the talks illustrated the importance of anchor observations, and more specifically GNSS-RO (Global Navigation Satellite System radio occultation), to estimate the different types of biases. In this context, reanalysis is facing a challenge, with highly variable observational coverage that becomes sparse going further back in time.

Representation of observation errors

The random error characteristics assigned to observations in the assimilation play a key role in determining the weighting of observations in the assimilation. Several talks reviewed the state of the art of modelling observation errors at various centres. Error correlations and situation-dependence of observation errors are increasingly being taken into account in operational systems. The required observation error modelling mostly relies on diagnostics based on differences between observations and model equivalents from short-range forecasts or analyses. Several speakers highlighted the potential, but also pitfalls of these approaches.

To represent situation-dependence, results from the diagnostics are used to determine simple parametrized models that capture main variations in error contributions, for instance from relatively poor surface or cloud modelling contributing to representation error. Concepts for modelling situation-dependent inter-channel error correlations are being developed. A better understanding and treatment of spatially correlated observation errors is also emerging, although algorithmic challenges remain, especially for variational data assimilation systems.

Looking to the future

A growing and increasingly diverse observing system will pose additional challenges for the treatment of systematic and random errors. Hyperspectral sounding data with unprecedented temporal resolution will be available from geostationary satellites, observations from some smaller satellites may have less well-characterised uncertainties, and an increasing number of observations will be driving multiple Earth system components with different model error and levels of maturity. All of these aspects will require continued development of the methods and capabilities used, including for the NWP-independent characterisation of all error sources.

Further information

Recordings of all presentations and panel discussions, all posters and a full workshop report are available from https://events.ecmwf.int/event/170/.

Statistical post-processing of ensemble forecasts at the Belgian met service

Jonathan Demaeyer, Stéphane Vannitsem (both Royal Meteorological Institute of Belgium & EUMETNET), Bert Van Schaeybroeck (Royal Meteorological Institute of Belgium)

The new system for post-processing ECMWF ensemble forecasts at the stations of the Royal Meteorological Institute (RMI) of Belgium was described previously in a short newsletter article (Vannitsem & Demaeyer, 2020). This system has now been operational since the summer of 2020 and we provide a description of its functionality and a preliminary analysis of its added value.

Principles

Statistical post-processing of ensemble weather forecasts has become an essential step in the forecasting chain as it enables the correction of biases and uncertainty estimates of ensembles (Gneiting, 2014). In June 2020, the Royal Meteorological Institute of Belgium, an ECMWF Member State, launched an operational statistical post-processing suite based on ECMWF medium-range ensemble forecasts in 11 reference synoptic stations, with the goal of improving its forecasting chain.

More precisely, the purpose of this new system is to provide forecasts closer to the typical values observed in these stations. Indeed, while ECMWF forecasts have in general good skill in the centre of the country, the temperature forecasts for the seaside region to the north and for the hilly forest region in the south are commonly known by forecasters to display notable biases. In addition, in the southern region, the wind gust variable is also known to be problematic, hampering the assessment of storm intensities and the release of accurate warnings.

The algorithm selected to perform the correction of the weather forecasts for the minimum and maximum temperature and for wind gusts is a linear member-by-member (MBM) Model Output Statistics (MOS) system, post-processing each member of the ECMWF ensemble (Van Schaeybroeck & Vannitsem, 2015). This method consists in correcting the mean and variability of the ensemble members in line with the observed climatology. At the same time, it calibrates the ensemble spread such as to match, on average, the mean square error of the ensemble mean. The MBM method calibrates the ensemble forecasts based on the station observations by minimising the continuous ranked probability score (CRPS). Slightly different configurations of the linear MBM approach are used, depending on the variable considered in order to optimise their reliability.

The forecast suite constitutes a proof-of-concept of research-to-operation implementation resulting from a fruitful collaboration between the different services of the institute.

Operational implementation

To generalise statistical post-processing for a large set of applications, a new Python library has been designed. For a smooth integration in the production chain, this library is then placed inside Docker platforms that are passed to the institute's team in charge of the operational duties.

The RMI post-processing application corrects four variables: temperature (T), minimum temperature (Tmin), maximum temperature (Tmax) and wind gusts. It does so for the 11 'canonical' Belgian synoptic observation stations mentioned above. The MBM MOS post-processing method uses ECMWF re-forecasts over the past 20 years and relates them to the corresponding past synoptic station observations. The ECMWF re-forecast products (Hagedorn, 2008) are issued twice a week (Monday and Thursday at 00 UTC). They consist of ensemble forecasts for the last 20 years at the same calendar date, with the newest version of the model available and with currently 10 ensemble members. A single predictor from the model is used. This means that, in essence, the relationship is obtained as a linear regression over a two-dimensional scatterplot.

The current post-processing applications at RMI are configured based on a window of five weeks of re-forecasts and the corresponding station observations to gather an optimal set of forecasts to evaluate the linear regression parameters. The five weeks cover the full set of available re-forecasts (*https://confluence.ecmwf.int/display/FUG/5.3+Model+Climates*). Twice a week, upon availability, the new re-forecasts are downloaded. Afterwards, the regression parameters are computed, ensuring that the statistical post-processing follows the seasonal variations. Once the relation between the forecasts and the observations is obtained, it is used to correct the ECMWF ensemble forecast

doi: 10.21957/vb78ui92dj

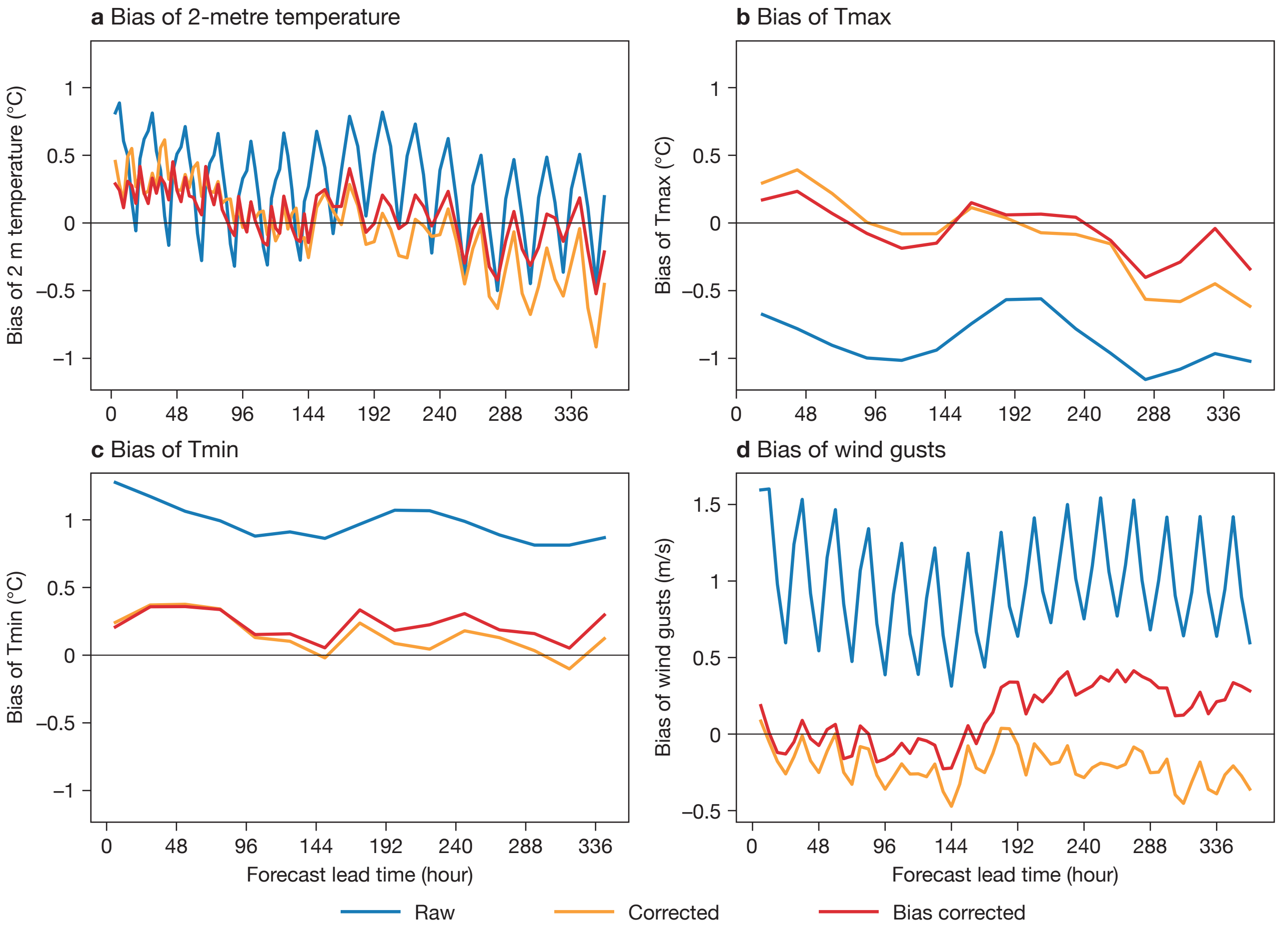


FIGURE 1 Averaged bias over all stations and over the JJAS months for (a) 2 m temperature, (b) Tmax, (c) Tmin and (d) wind gusts, as a function of the lead time. Two corrections are considered: one with a simple statistical correction of the bias of the ensemble ('Bias corrected'), and the full member-by-member correction minimising the CRPS ('Corrected').

issued daily at 00 UTC and transferred to the RMI through the dedicated dissemination channel.

The forecasts are being provided with a lead-time interval of 3 hours for the first 144 hours, and 6 hour-intervals afterwards. The re-forecasts, on the other hand, are provided 6-hourly over the whole time-range. Therefore, the statistical post-processing parameters are interpolated over the missing 3 hour-intervals in the first 144 hours. While experimental and based on the assumption that the post-processing parameters are smooth enough, this interpolation has proved to be skilful, as shown by the scores detailed in the following section.

Post-processing scores for the summer 2020

Statistics have been accumulated for the extended summer of 2020: June, July, August and September (JJAS). Some relevant scores are shown here to highlight the system's performance in providing improved forecasts. The results are obtained by averaging over all stations and all forecast days and months. Figure 1 shows the correction of the systematic bias of the raw ensemble by the MBM adjustment (orange line). The red line is obtained using a simple correction of the systematic bias (no variability or spread correction) and therefore shows the smallest bias (except for the minimum temperature). The reason is that the MBM adjustment seeks to minimise the CRPS, at the expense of a less good correction of the bias.

The CRPS score, which measures the quality of the probabilistic information provided by the ensemble, is shown in Figures 2a,c and 3a,c. A smaller CRPS score corresponds to an improved ensemble forecast. Figures 2b,d and 3b,d show the reliability and resolution components of the CRPS (Hersbach, 2000). The reliability measures the quality of the forecast probabilities with respect to the observed climatological frequencies. The lower the values of the reliability are, the better the calibration of the model's ensemble. The resolution, on the other hand, measures the information content of the ensemble forecast scheme. The higher it is, the more the ensemble forecasts improve upon the climatological forecasts (see Box A).

For all temperature variables (Figures 2 and 3), there is a substantial improvement due to the ensemble calibration

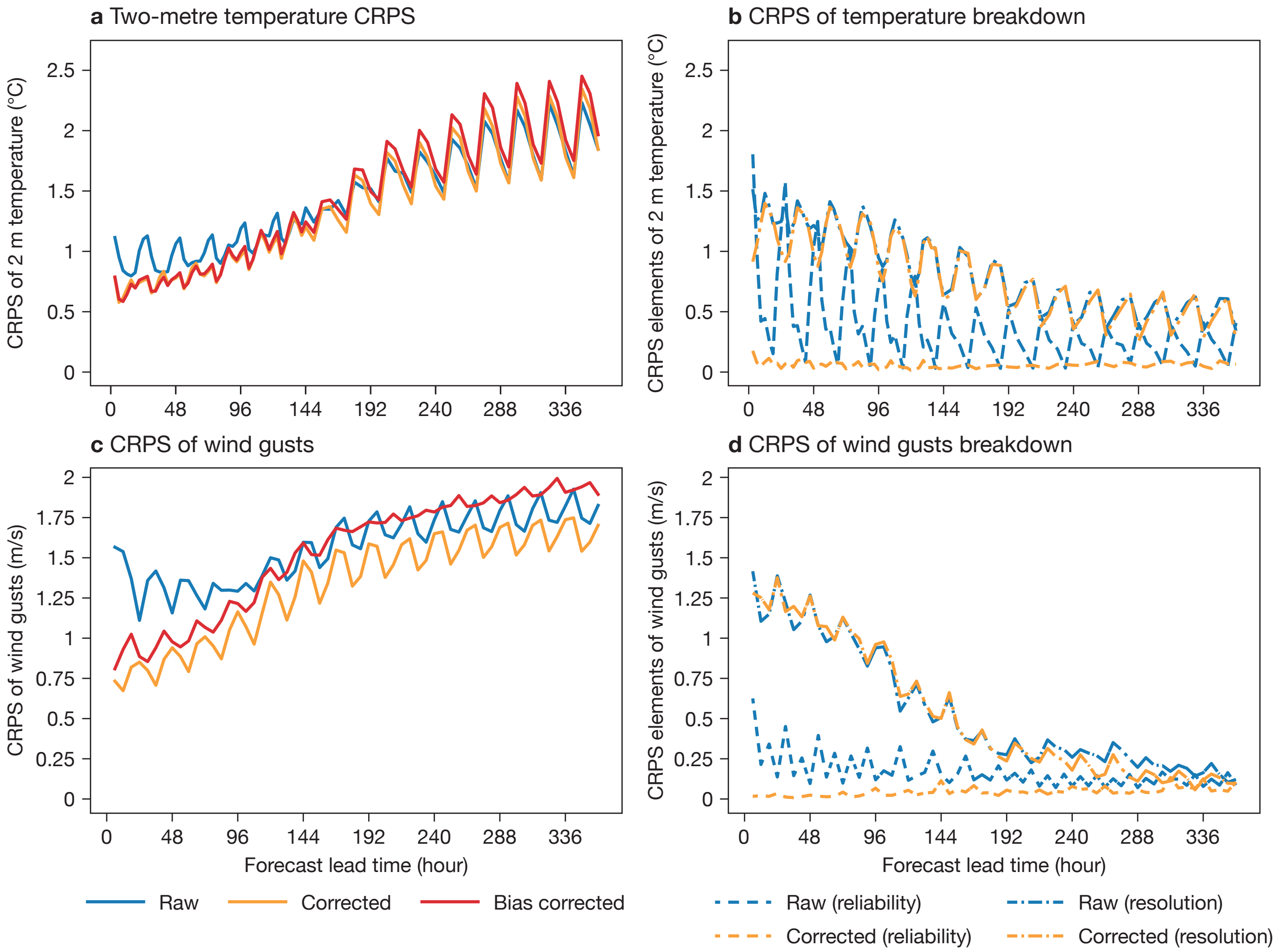


FIGURE 2 Averaged CRPS over all stations and over the JJAS months for (a) 2-metre temperature showing the full member-by-member correction minimising the CRPS ('Corrected') and a simple statistical correction of the bias of the ensemble ('Bias corrected'), with (b) the decomposition of the raw and MBM corrected ensemble CRPS for 2-metre temperature into the reliability and resolution components, (c) the corresponding wind gust variable corrections and (d) the corresponding breakdown of wind gust variable corrections, as a function of the lead time.

during the first half of the forecast period. Skill is gained at all lead times for the wind gust. While a simple bias correction (red line in Figure 2a) suffices to improve the CRPS of the 2-metre temperature to the same level of the MBM method, the variability correction is needed to further reduce the CRPS of the other variables. It indicates that for 2-metre temperature, the bias set aside, the ensemble is already well calibrated, in agreement with the results presented in Vannitsem & Hagedorn (2011).

The right-hand panels of Figures 2 and 3 show that, for each variable, the improvement of the CRPS score is due mainly to a decrease in the reliability contribution. As already indicated above, in the case of 2-metre temperature, the improvement of the reliability is mainly due to the correction of the statistical bias, while it also involves a contribution from the correction of the spread for the other variables. In contrast to the substantial ensemble reliability improvement, the resolution for the temperature and wind gust variables is not substantially modified. The slight decrease in resolution for the minimum and maximum temperature can perhaps be partly circumvented if more predictors

a Breaking down the CRPS score

The CRPS score can be expressed by the reliability (Reli), the resolution (Resol) and the uncertainty (U) as follows:

CRPS = Reli - Resol + U

The forecast is the better the smaller the reliability value is – as close to zero as a possible – and the bigger the resolution. The ensemble system has positive resolution if it performs better than the climatological probabilistic forecast.
The uncertainty is the potential reliability for a forecast system based on the sample climatology.

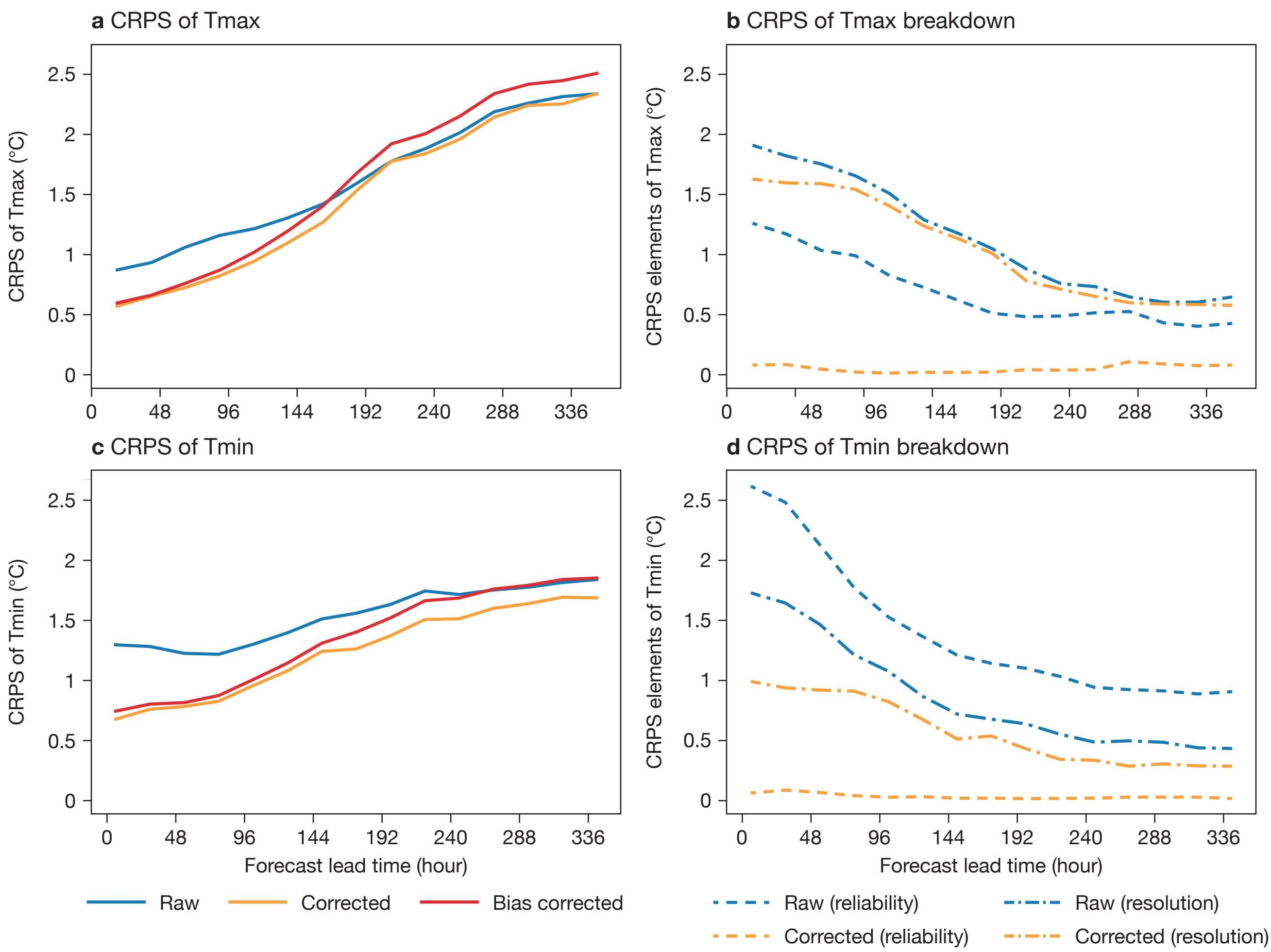


FIGURE 3 Same as Figure 2 but for the maximum and minimum temperatures, as a function of lead time.

are added to the correct variables. This aspect will be explored in the near future.

Finally, we show an example of a forecast at the Elsenborn station in Figure 4, a station located in the Ardennes region featuring large biases with respect to the other stations. The biases result from the complex orography, with an elevation at 570 m of the nearby region. Remarkably, one can see on this graph the correction of the bias, materialised by the shift of the ensemble mean, as well as the correction of the spread, depicted here by the 10–90% quantiles of the ensemble distribution shown with different levels of transparency. Notably, the correction of the bias for the wind gust induces a shift of the ensemble toward smaller values, and hence less extreme ones. We see also that while the observed minimum temperature is out of the raw ensemble distribution at the end of the forecast (Figure 4), the corrected ensemble distribution encompasses the observation, as expected.

The way forward

Following the first steps of the new post-processing programme of the RMI presented here, new developments are expected during the next couple of years:

- Currently only one predictor is used. Further improvements of the skill will be tested by introducing additional predictors.
- As mentioned above, for the moment only the midnight forecasts of ECMWF are post-processed, and not the forecasts issued at noon. Indeed, to correct the latter, one needs re-forecasts issued at noon as well, and these re-forecasts are not currently available at the Centre. One solution would be to post-process the noon forecasts with the parameters of the midnight forecasts. As a consequence, these parameters have to be shifted by 12 hours to match the diurnal cycle, which implies that the parameters are not optimal anymore. This could lead to a less optimal forecast correction and the impact of this shift must be carefully assessed. We nevertheless expect corrections that would justify the post-processing of these noon forecasts.
- Another development avenue is the implementation of a member-by-member post-processing application for gridded probabilistic forecasts. The core of the computation will again be done by the RMI post-processing library placed inside a Docker container. This post-processing application will either use the data

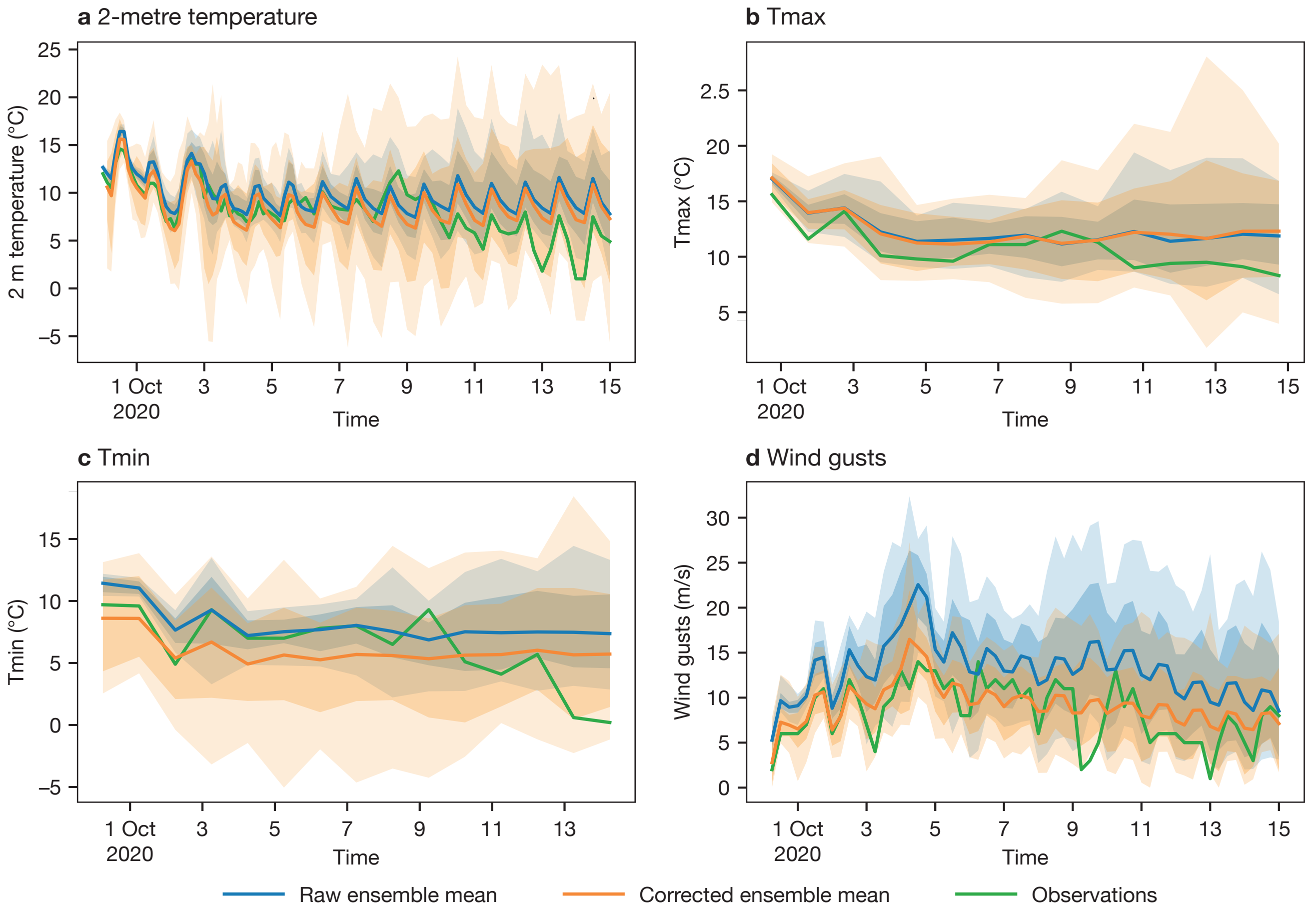


FIGURE 4 Example of a corrected forecast for each of the four variables: (a) 2-metre temperature, (b) maximum temperature (Tmax), (c) minimum temperature (Tmin) and (d) wind gusts, with the corresponding station observations for verification. The station being considered is Elsenborn and the raw ECMWF forecast was issued on 30 September 2020 at 00 UTC. The lower and upper lightly shaded areas represent respectively the 0 to 10% and the 90 to 100% quantiles. The darker shaded areas represent the 10 to 90% quantiles. Solid lines are the ensemble means.

of the RMI INCA system, which contains gridded combined observations, or of the ERA5-land reanalysis. The specific design of this application is still under discussion.

- Finally, longer developments involving more recent and sophisticated techniques are considered: spatial post-processing, machine learning, etc.

Discussions are also ongoing about the link of these activities with the EUMETNET post-processing benchmark, which is in preparation and for which an experimental proof-of-concept will be developed soon on the European Weather Cloud. This could foster the collaboration between national meteorological services on the development of a common platform for post-processing tools and best practices.

Further reading

Gneiting, **T.**, 2014: Calibration of medium-range weather forecasts. *ECMWF Technical Memorandum* **No. 719**.

Hagedorn, **R.**, 2008. Using the ECMWF reforecast dataset to calibrate EPS forecasts. *ECMWF Newsletter* **No. 117**, 8–13.

Hersbach, **H.**, 2000: Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, **15**, 559–570.

Vannitsem, **S.** & **R. Hagedorn**, 2011: Ensemble forecast post-processing over Belgium: comparison of deterministic-like and ensemble regression methods. *Meteorological Applications*, **18**, 94–104.

Vannitsem, **S.** & **J. Demaeyer**, 2020: Statistical post-processing of ECMWF forecasts at the Belgian met service, *ECMWF Newsletter* **No. 164**, 4–5.

Van Schaeybroeck, **B.** & **S. Vannitsem**, 2015: Ensemble post-processing using member-by-member approaches: theoretical aspects. *Quarterly Journal of the Royal Meteorological Society*, **141**, 807–818.

Wilks, **D.S.**, 2011: *Statistical methods in the atmospheric sciences*, **Volume 100**, Academic Press.

Major upgrade of the European Flood Awareness System

Cinzia Mazzetti, Damien Decremer, Christel Prudhomme

On 14 October 2020, the European Flood Awareness System (EFAS) launched a new cycle upgrade, EFAS version 4.0. This was a step-change in EFAS. For the first time, the LISFLOOD hydrological model, the 'engine' of EFAS, was calibrated using sub-daily steps and it is now used with sub-daily steps in all hydrological simulations throughout the system.

The EFAS domain includes 66 countries. For EFAS version 4.0, a total drainage area of 4M km^2 was calibrated so that the hydrological representation of those catchments can be as accurate as possible (Figure 1). This resulted in a marked improvement in the hydrological simulations for most catchments, except in strongly regulated catchments, where the new calibration did not bring much change.

This article introduces the new 6-hourly calibration of the LISFLOOD hydrological model and provides a summary of its performance.

What is EFAS?

EFAS is an operational pan-European flood forecasting system funded by the European Commission through its Copernicus programme. The aim of EFAS is to support preparatory measures before major flood events strike, particularly in large transnational river basins and throughout Europe in general.

EFAS is a component of the Copernicus Emergency Management Service (CEMS). Since the beginning of its operational implementation in 2012, it has been providing flood forecast information to 116 national hydro-meteorological services across Europe, the European Commission's Emergency Response and Coordination Centre (ERCC), and other research institutes. EFAS is managed by the EU Joint Research Centre (JRC) and is delivered by four centres run by different consortia:

- **Computational centre (EFAS-COMP)** It is responsible for producing forecasts and hosting the EFAS-Information System platform. It is operated by ECMWF.
- **Dissemination centre (EFAS-DISS)** It provides a daily analysis of EFAS forecasts and disseminates the information to EFAS partners and the ERCC. It is coordinated by the Swedish Meteorological and Hydrological Institute and also comprises the Dutch Rijkswaterstaat and the Slovak Hydro-Meteorological Institute.

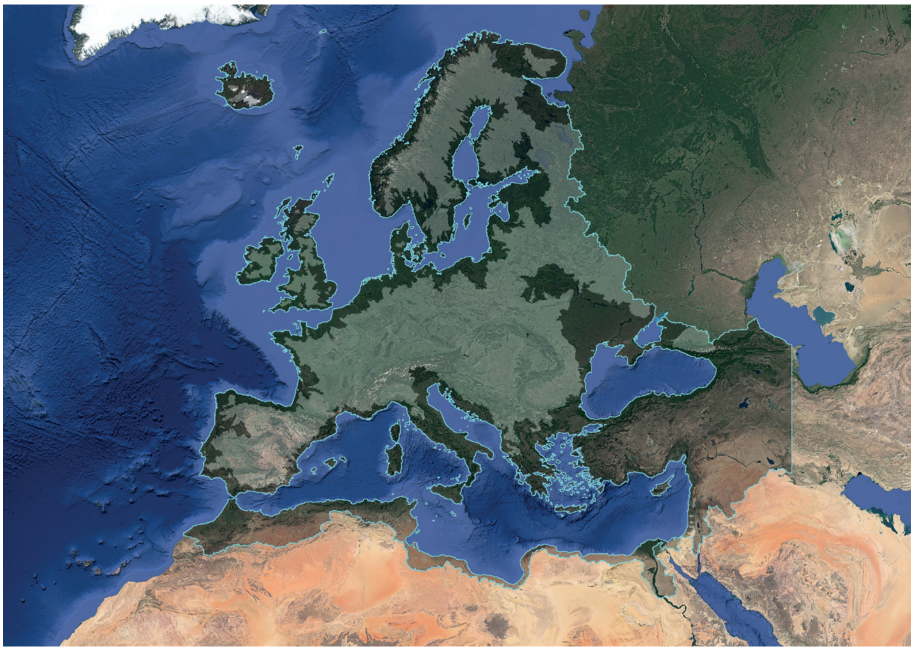


FIGURE 1 Map of the EFAS domain (dark shade) and calibration extent (light shade).

doi: 10.21957/32rgs58mc9

- **Hydrological data collection centre (EFAS-HYDRO)** It collects historical and real-time river discharge and water level data across Europe and makes them available to EFAS-COMP. It is delivered by the Environmental and Water Agency of Andalucía (REDIAM) and Soologic Technological Solutions SL.
- **Meteorological data collection centre (EFAS-METEO)** It collects historical and real-time meteorological data across Europe and provides them in real-time to drive the EFAS modelling chain. It is composed of KISTERS AG and the German national meteorological service (Deutscher Wetterdienst, DWD).

As part of the EFAS computational centre's role, ECMWF is also responsible for developing and integrating into operation any improvement in the forecast model chain, and for developing, managing and running the EFAS web and data services.

Medium-range forecasts in EFAS

EFAS medium-range ensemble flood forecasts are generated by cascading an ensemble of meteorological forecasts (from ECMWF, DWD and COSMO-LEPS from the COSMO consortium), meteorological and hydrological observations, land surface information and model parameters (static maps) through a deterministic hydrological model (LISFLOOD).

The resulting ensemble flood forecasts are post-processed to produce all EFAS products. The products, including flood highlights of different severity levels, are made available to EFAS-DISS and EFAS users as maps and graphs. Three severity levels are highlighted, corresponding to forecasts of floods expected to exceed flood peaks with return periods of 2, 5, and 20 years (a return period indicates the average number of years expected between two floods of the predicted magnitude). Finally, EFAS-DISS duty forecasters analyse the flood summary maps and issue notifications to registered users of the concerned region to inform them of possible upcoming events (Figure 2).

A new 6-hourly calibration

Like other operational forecasting systems, EFAS is always evolving, but the October 2020 upgrade included a step-change in EFAS hydrological modelling for medium-range forecasts. For the first time, the LISFLOOD hydrological model was calibrated at 6-hourly steps over the EFAS pan-European domain, compared to 24-hourly steps previously. At the same time, all hydrological medium-range simulations are produced at sub-daily steps, so that the timing of the start and peak of flood events can be better anticipated.

Upgrade of the hydrological model

LISFLOOD has been developed at the JRC since 2000 and has been used for operational flood forecasting at the pan-European scale since the early days of EFAS. Since 2019 the model is fully open source and the code is developed and maintained through a GitHub repository by the JRC (*https://ec-jrc.github.io/lisflood/*), with support from the EFAS-COMP team at ECMWF.

LISFLOOD is a fully distributed hydrological model, which explicitly considers the spatial distribution of physical properties across catchments and provides estimates of river discharge on the entire geographical domain. Driven by meteorological forcing data (precipitation, temperature, potential evapotranspiration, and evaporation rates for open water and bare soil surfaces), LISFLOOD calculates a complete water balance for every grid cell within the EFAS domain, currently on a 5x5 km grid. Processes simulated include snowmelt, soil freezing, surface runoff, infiltration into the soil, preferential flow, redistribution of soil moisture within the soil profile, drainage of water to the groundwater system, groundwater storage, and groundwater base flow. Runoff produced for every grid cell is then routed through the river network using a kinematic wave approach. The model also includes options to simulate lakes and reservoirs.

For EFAS 4, LISFLOOD was upgraded to run sub-daily

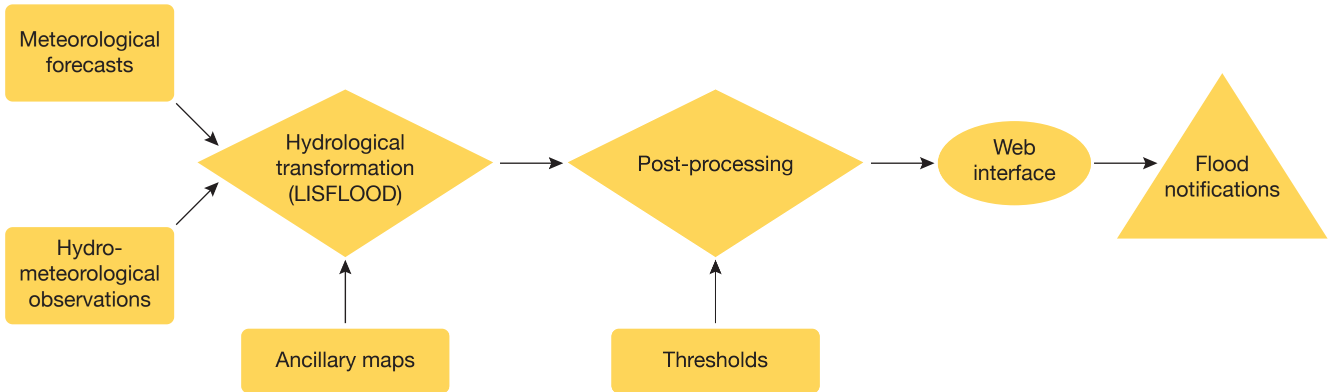


FIGURE 2 EFAS flood forecast and notification process.

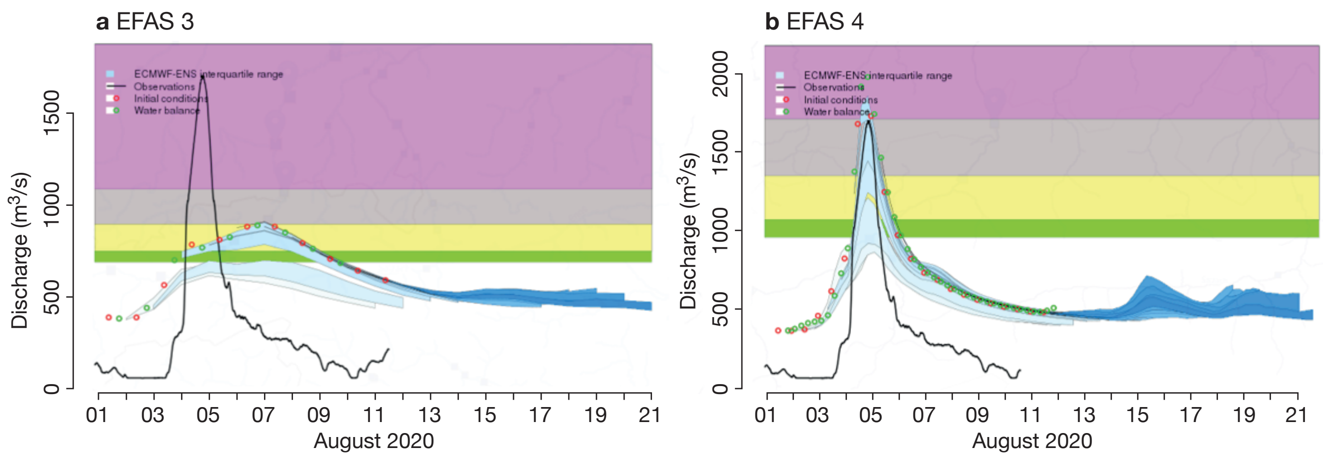


FIGURE 3 Example of the improvement to simulated hydrographs between (a) EFAS 3 and (b) EFAS 4 for the river Inn at Mühldorf (Germany). The observed discharge is represented by the black line, LISFLOOD outputs are represented by the green and red dots.

time steps (6-hourly), routing of flood waves in rivers was improved and the handling of model state files was refined. The upgrades allowed for better representation of hydrology in small to medium size catchments (Figure 3) and for the use of more realistic parameters in the calibration process.

New 6-hourly forcing fields

By upgrading the full EFAS medium-range modelling chain to a 6-hourly timestep, all forcing fields, including observed meteorological data to simulate initial conditions, needed to also be produced at the finer time step. For the LISFLOOD model, this includes gridded maps of precipitation, average air temperature, evaporation rate from free water surface and bare soil surface, and potential evapotranspiration for reference crop surfaces.

In EFAS, the meteorological data collection centre (EFAS-METEO) collects datasets of historical and real-time in-situ meteorological observations on a 24/7 basis from 22 data providers over more than 40k stations and 70k sensors, and it interpolates them to the 5 km hydrological model grid.

For the calibration of EFAS 4, EFAS-METEO produced new datasets with 6-hourly gridded meteorological maps of precipitation and average surface air temperature using point observations for the period 1990–2017 on the model's 5 km spatial resolution. Because of the sparsity of meteorological data such as wind speed, solar radiation or humidity at a 6-hourly time step, estimates of potential evapotranspiration were made using the Penman-Monteith equation with daily data, and then disaggregated using the same evaporation rates for each time step to produce 6-hourly grids.

New daily and sub-daily discharge dataset

The EFAS hydrological data collection centre (EFAS-HYDRO) collects historic and real-time river discharge and water level data across Europe from 44 data partners. Metadata such as name, location and upstream drained area are also collected and maintained. Data from more than 1,800 active stations are collected on a 24/7 basis either as water levels and/or discharge at different temporal resolution, then quality checked and resampled at 1-, 6- and 24-hour time steps. Water level data are transformed to discharge data (the information which is required by LISFLOOD) when rating curves are provided by EFAS partners.

For the calibration of EFAS 4, a dataset containing daily and 6-hourly discharge data at river gauges for the period 1990–2017 was put together at ECMWF, based on data provided by EFAS-HYDRO.

Ancillary maps

LISFLOOD requires a wide range of spatially distributed input parameters and variables such as topography, soil type, land use, channel geometry, and river network. The pan-European setup of LISFLOOD uses a 5 km grid on a Lambert Azimuthal Equal Area projection. EFAS configuration maps were created by the Joint Research Centre (JRC) of the European Commission from various European databases with emphasis on having a homogeneous base all over Europe.

For EFAS 4, the LISFLOOD domain area was slightly extended to include the Jordan catchment. The river network in the Balkans was also improved to better represent physical rivers in the region, with updates to the channels' geometry reflecting the changes done to the drainage network. The LISFLOOD model for EFAS 4 includes 1,423 reservoirs and 210 lakes. Compared to the previous EFAS 3 version, three additional reservoirs were added on the Sava river downstream of Zagreb to better represent the effects of large retention areas.

Calibration stations

LISFLOOD calibration stations were selected from the list of 2,927 river gauging stations with discharge data that was available in the EFAS-HYDRO database in July 2018, when calibration work began. Stations were located on the LISFLOOD 5 km drainage network using a semi-automatic procedure and additional manual checks. Available discharge data were then quality checked to exclude stations where discharge data showed issues with instrumentation, rating curves or water release from reservoirs.

Stations where a minimum of four years of good-quality discharge data were available were selected as calibration points. For the sake of representativeness and to reduce the computation time of the calibration, stations located close to others along the same river (i.e. stations with a difference in drained area smaller than 200 km^2) and with the same data quality were excluded in favour of the station with the longest data period or the largest drainage area.

The selection procedure produced a list of 1,137 calibration stations in 215 different catchments, 406 with 6-hourly and 731 with daily observed discharge time series. In total, 44.5% of the EFAS domain area belongs to a calibrated catchment, corresponding to 4 million km^2 over 9 million km^2, with the catchment area of the stations varying from 468 km^2 (Ishem catchment, Albania) to 807,000 km^2 (Danube catchment, Romania) and a median area of 3,000 km^2.

Compared to the EFAS 3 system, the number of calibration points increased by 426 stations, up from 711 in the previous calibration exercise. However, some of the already existing calibration stations are now providing 6-hourly data and thus different data were used in the EFAS 4 calibration, often on significantly shorter periods.

Calibration methodology

Like most rainfall-runoff models, the equations of the LISFLOOD hydrological model include a range of parameters. Some of these can be determined from physical data, such as reservoir storage-elevation curves, or the drainage area of watersheds. Others vary from one area to another based on changes in climatology and physical factors, e.g. hydraulic soil properties. Some model parameters require calibration, which is generally obtained by tuning parameter values based on a comparison between simulated and observed discharge (Q) at river gauges. This tuning process generally aims to minimise errors in the volume and timing of simulated flow over a multi-year period.

A set of 14 model parameters was selected for calibration following recommendations from previous work on the LISFLOOD model and its application to EFAS. The parameters control snow accumulation and snow melting, overland flow, percolation to the lower groundwater zone, the residence time of the upper and lower groundwater zone, lakes, reservoirs and channel routing. Parameter spaces were defined by physically reasonable lower and upper limits for physically based parameters and largest admissible ranges for empirical parameters.

For EFAS 4 calibration, an Evolutionary Algorithm (EA) was used to generate sets of model parameters and the modified Kling-Gupta efficiency (KGE’) was selected as the objective function (or goodness-of-fit measure) as it provides a way to achieve balanced improvement of simulated mean flow, flow variability, and correlation.

A number of calibration stations were available along the same rivers. This offered information on nested catchments, which is very important for a distributed hydrological model calibration. However, it came with an additional challenge in terms of calibration strategy and run time, as data was generally a mix of 6-hourly and daily discharge observed over different time periods. This was solved by dividing the EFAS domain in 1,137 sub-catchments or inter-catchments and by performing the calibration through a catchment-based parallelisation of the model domain.

Each sub-catchment was calibrated separately but using a multisite cascading calibration (MSCC) approach, where the calibrated discharge from upstream river basins was used as input for downstream ones. The calibration was iteratively performed from upstream to downstream, from the catchment with the smallest area to the largest one, as flow routing calculations must be carried out in a serial manner along the mainstream of a river basin.

To run the parameter optimisation procedure, the LISFLOOD model was run at 6-hourly steps everywhere. However, the objective function (or goodness-of-fit measure) was calculated over a daily-aggregated time series for calibration points with daily observations, to allow a fair comparison between simulated and observed discharge data. This dual calibration strategy (both at 6-hourly and daily time steps) was critical to guarantee the best possible geographical coverage as 6-hourly river discharge data was not available everywhere.

Hydrological model performance

The calibration process resulted in 14 new parameter maps over the pan-European EFAS extent, one per calibrated parameter. For each sub-catchment domain, the parameters identified by the calibration procedure were used in all grid cells, while for areas not covered by the calibration stations, default parameters were used instead. Most of the parameters in LISFLOOD act as multipliers for the ancillary maps describing the geo-physical properties of the catchments, so even if a single value is used for all model pixels in a catchment, spatial variability of the model parameters is preserved through the ancillary maps. Parameter maps were then used to execute a continuous simulation forced with observations for the period Jan 1990 – Dec 2017. The simulated discharge was then compared against observed discharge from the 1,137 calibration stations, excluding the year 1990 so that the impact of the initial conditions did not affect the comparison. For calibration stations with daily data, 6-hourly LISFLOOD time series were aggregated at daily steps before the comparison with the observations.

One important aspect of a hydrological model performance evaluation is to use data which have not been used during the calibration exercise, so that the evaluation is a fair analysis of the model behaviour. This is often achieved using a 'split sample approach', where the hydrometeorological observational record is split into two independent periods, one used to optimise the parameters through calibration, and one used to evaluate the model behaviour. This could not be fully adopted for EFAS 4 for two main reasons. First, although each calibration station was calibrated separately, observational periods were generally different, making it impossible to define the same non-overlapping periods for all calibration points along the same river. Second, some calibration stations had only four years of available data, which is the shortest necessary duration to achieve a robust parameter optimisation, and no data were left for evaluation. This meant that no observational record was available to conduct an independent evaluation for those stations. As a compromise, the calibration was done using a sub-set of the observation record in most of the stations, but the hydrological model performance was evaluated on the full available discharge record (including the data used for the calibration).

Hydrological model performance was measured using the modified Kling-Gupta efficiency (KGE'), the same metric as that used for model calibration. The KGE' is an expression of the distance from the point of ideal model performance in the space described by three components: correlation, variability bias and mean bias. It can vary between any negative number and 1, where 1 indicates perfect agreement between simulations and observations.

All three components of KGE' represent desirable characteristics of the hydrological regime in the context of EFAS: correlation evaluates the flow timing, of paramount importance for EFAS to issue timely warnings. Variability bias measures the statistical

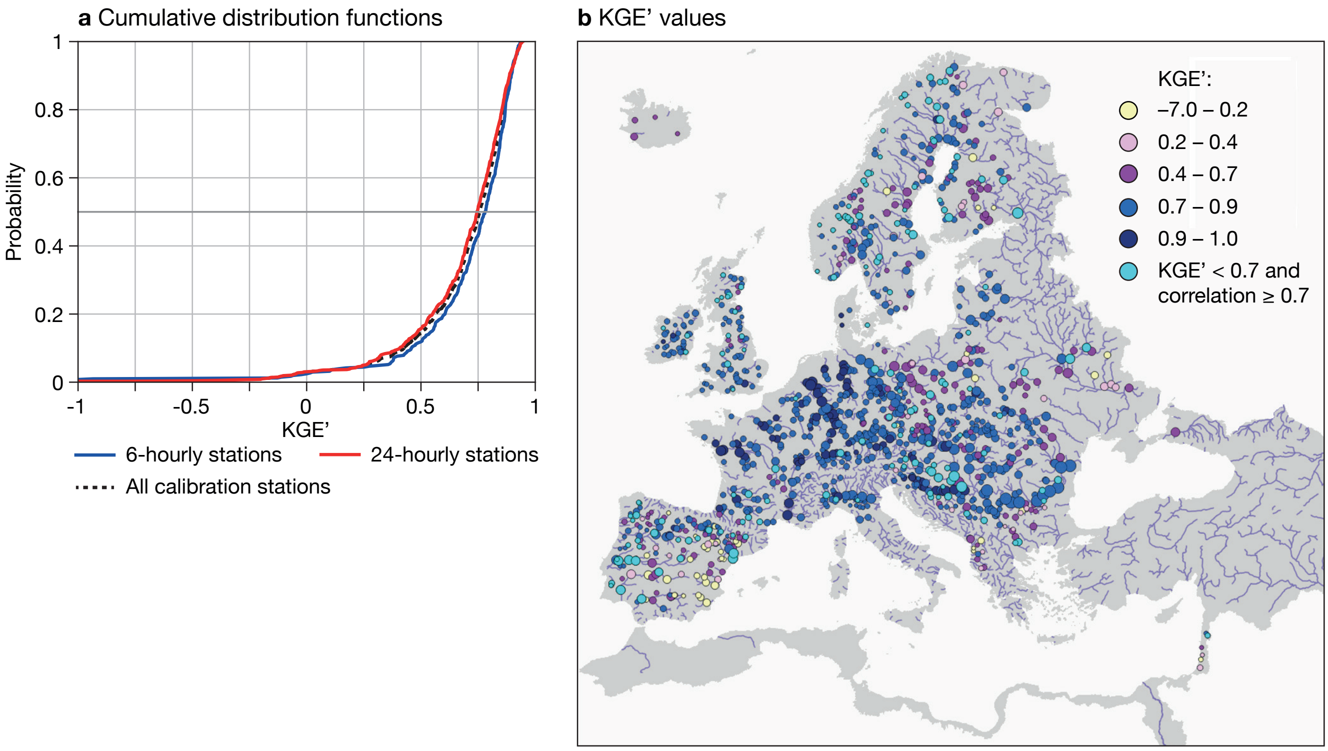


FIGURE 4 Hydrological model performance of EFAS v4.0 as described by the modified Kling-Gupta efficiency score (KGE') calculated over 1991–2017, showing (a) the cumulative distribution function (CDF) for all calibration stations, 6-hourly calibration stations and 24-hourly calibration stations and (b) the KGE' skill score for each calibration station as coloured-coded symbols. Stations with KGE' < 0.7 and correlation ≥ 0.7 are shown separately. The size of the dots is proportional to the drained area of the calibration station.

variability, ensuring that peaks are correctly reproduced by the system. And mean bias describes the long-term water balance, important for any hydrological application. It is important to remember that although the length of the model run is the same for all stations, discharge data availability is uneven, and therefore KGE' is computed using different data periods for each station across the EFAS domain.

The hydrological model performance over 1991–2017 as expressed by the KGE' is shown in Figure 4. Overall, around 50% of the calibration stations achieve a KGE' greater than 0.75, a high value for model performance compared to an optimum of 1. There is very little difference regarding the performance of stations calibrated at a 6-hourly or daily time step, although the former tend to have a slightly better score (Figure 4a). Generally, performance is relatively uniform across the EFAS pan-European domain (Figure 4b). However, areas of higher skill are found in large parts of Central Europe and the main European rivers, whilst lower skill is mostly concentrated in catchments with strongly regulated rivers, like in the Iberian Peninsula. Stations with KGE' < 0.7 and correlation ≥ 0.7 are also highlighted to show stations that might have lower KGE' due to systematic bias, but still have high correlation. Correlation is particularly important for EFAS given that forecasts are compared to model thresholds, so they are bias invariant to a large extent.

Comparing hydrological model performance

For the fairest comparison possible between EFAS versions 3 and 4, the hydrological model performance score KGE' was calculated on simulations using the same meteorological forcing data (but aggregated to daily forcing for EFAS 3) over 1990–2017. For EFAS 4, scores were calculated on river discharge averaged over 24 hours to be comparable to the 24-hourly simulation of EFAS 3 (also matching the calibration time step of EFAS 3). Note that calculating KGE's over daily discharge slightly increases the score of EFAS 4.

Overall, EFAS 4 shows a marked improvement in the hydrological performance compared with EFAS 3, with more stations achieving a higher KGE' score as shown by the cumulative distribution function. About 60% of stations have a KGE' of 0.75 or higher, against only 40% for EFAS 3 (Figure 5a). Improvements are found over most of the EFAS domain, with the exception of some stations in Scandinavia, Spain and central Europe (Figure 5b). Causes for skill score degradation are varied. In the Elbe catchments, EFAS 4 calibration could only be conducted with a much shorter period of data (down to only four years on the main Elbe river, using 6-hourly records) during a period without major flood events, compared with a much longer and hydrologically diverse calibration for EFAS 3. Catchments in Scandinavia and Spain have a large number of

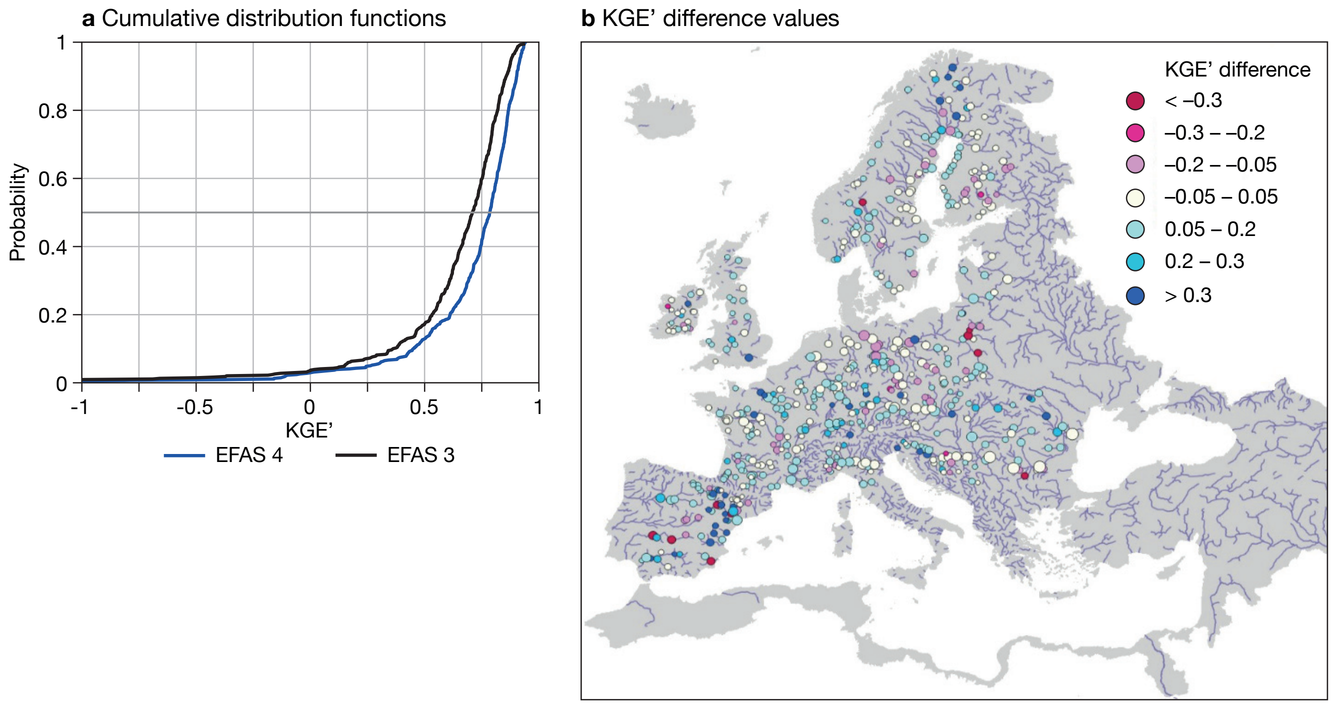


FIGURE 5 Change in the EFAS hydrological model performance between v3 and v4 as described by the modified Kling-Gupta efficiency score (KGE') calculated on daily river discharge over 1991–2017, showing (a) the cumulative distribution function (CDF) for EFAS 3 and EFAS 4 and (b) for each common calibration station, the difference in KGE' values for EFAS 4 and EFAS 3 as colour-coded symbols. Blue (red) shades highlight stations where EFAS 4 performs better (worse) than EFAS 3. The size of the dots is proportional to the drained area of the calibration station.

reservoirs, which can be challenging to model at 6-hourly steps. Finally, the LISFLOOD routing scheme, which does not flatten peaks during flood propagation, is slightly penalised by the 6-hourly time step over large river basins in flat areas. This is because higher and more accurate peak floods from small and medium size upstream catchments are now produced by the model but are not properly propagated downstream. This is the case, for example, in the Danube.

What next for EFAS

The release of EFAS version 4 marks a milestone in EFAS development, delivering a first system calibrated with sub-daily data. The hydrological skill improvement provided by the new calibration and the higher time resolution of the products based on LISFLOOD outputs will allow for a timelier notification of the beginning and peak of predicted flood events, and for EFAS users to better understand and explore flood forecasts.

As with any operational forecasting system, EFAS development never stops. Even though the new EFAS version 4 has only just been released, the EFAS-COMP team at ECMWF has already started working on the next exciting new development: an almost 10-fold increase in the number of model cells, with the spatial resolution going from a 5 km grid to a ~1.8 km (1 arcminute) grid.

Further reading

Arnal, **L.**, **S.-S. Asp**, **C. Baugh**, **A. de Roo**, **J. Disperati**, **F. Dottori et al.**, 2019: EFAS upgrade for the extended model domain – technical documentation. EUR 29323 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-79-92881-9, doi:10.2760/806324, JRC111610.

Gupta, **H.**, **H. Kling**, **K. Yilmaz** & **G. Martinez**, 2009: Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling. *Journal of Hydrology*, **377**, 80–91, doi:10.1016/j.jhydrol.2009.08.003.

Kling, **H.**, **M. Fuchs** & **M. Paulin**, 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, **424–425**, 264–277, doi:10.1016/j.jhydrol.2012.01.011.

Mazzetti, **C.** & **C. Prudhomme**, 2018: Major upgrade for European flood forecasts. *ECMWF Newsletter* **No. 157**, 3.

ecFlow 5 brings benefits to Member States

Avi Bahra, Iain Russell, Sándor Kertész

Managing workflows for large-scale data-intensive computational processes is an ever-growing challenge. These workflows must be repeatable, highly available, monitorable and accurate, while still allowing the flexibility to support changes. At ECMWF this challenge has been met with ecFlow, a workflow package developed in-house to meet the ever-changing requirements of the Centre and its Member and Co-operating States.

ecFlow was designed for general use but has been sculpted by the operational and research needs of weather and climate science. For example, at ECMWF it is used for many purposes including research experiment runs, operational model runs, data post-processing and archiving, and software builds.

ecFlow enables users to run a large number of programs, with dependencies on each other and on time, in a controlled environment. It provides good tolerance for hardware and software failures and allows for controlled restarts. The server, client and graphical user interface (GUI) are highly scalable and can handle workflows with hundreds of thousands of tasks. ecFlow is open source and is written in C++ for optimum performance. It runs on UNIX platforms, with many years of experience on Linux and more recent usage on macOS.

Version 5 of ecFlow brings many modernisations and improvements in terms of features, performance, security and maintainability.

ecFlow's architecture

ecFlow has a client/server architecture (Figure 1). An ecFlow server is responsible for several suites, each a hierarchical collection of tasks. Complex suites can be defined using a Python API that guarantees their syntactic correctness (Figure 2). Simpler suites can be defined through plain text files. The server submits tasks to the machines where they will run, receiving updates as they proceed. Tasks can be defined in any scripting language, for example shell or Python. These scripts can be parametrized, meaning that the same script can be used for many different tasks, with different settings. For example, a script variable 'FORECAST_STEP' could be set to 6 when run in one task and 12 in another. The scripts may also have embedded ecFlow commands that communicate their status back to the server, e.g. to show progress or to trigger another task to start. Sophisticated use of these embedded commands allows tasks to dynamically modify the server's suites, facilitating an adaptive

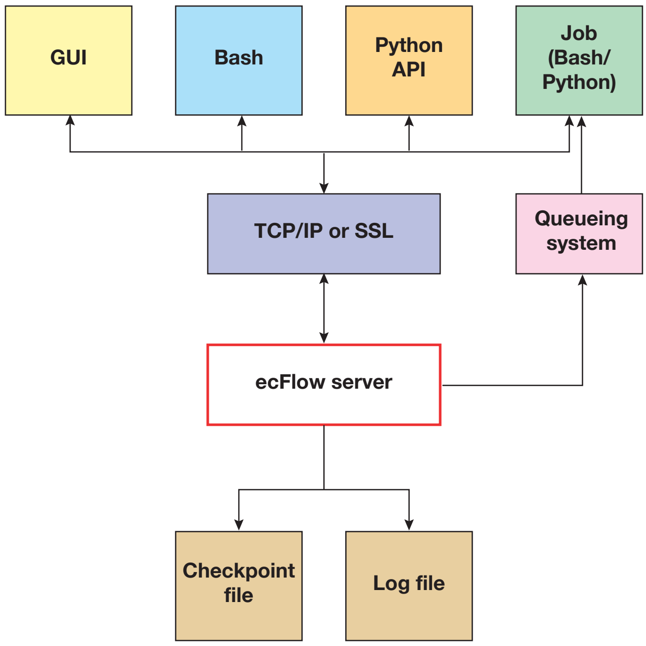


FIGURE 1 Various clients (GUI, Bash, Python API) can communicate bi-directionally with an ecFlow server using standard Transmission Control Protocols/Internet Protocol (TCP/IP) or Secure Sockets Layer (SSL) protocols. The server can run tasks directly or submit them to a queueing system; either way, they can still communicate back to the server. The server keeps track of events in a log file, providing the basis for statistical analyses of past events, such as the average duration of a given task. A checkpoint file is written to disk at regular intervals, providing a backup of the server's internal state at that moment; this mechanism can also be used to provide continuity when upgrading a server to a newer version of ecFlow.

```
import os
from ecflow import Defs,Suite,Family,Task,Edit,Trigger,Client

home = os.path.join(os.getenv("HOME"), "example")
defs = Defs(
        Suite("test",
            Edit(ECF_HOME=home),
            Family("f1",
                Edit(SLEEP=20),
                Task("t1"),
                Task("t2",Trigger("t1 == complete")))))

ci = Client()
ci.load(defs)
```

FIGURE 2 A simple example of a Python script that creates a new suite consisting of a family of two tasks, the second of which will be run as soon as the first has completed. The suite is then loaded onto a server using default settings.

doi: 10.21957/4tba2hz3f7

workflow without requiring manual loading of revised suites into the server. ecFlow is not tied to any particular queueing system that may sit in front of the worker machines, but its tasks can be submitted to any such queueing system through the use of a general submission script. ecFlow client applications include a graphical user interface, ecFlowUI, and a command-line program, ecflow_client, both of which can be used to query and modify the server.

Graphical user interface

ecFlowUI is the graphical user interface to ecFlow (Figure 3). It is written with the C++ Qt library. ecFlowUI supports real-time monitoring of the workflow, allowing jobs to be started, suspended and terminated. Many aspects can be edited on the fly, including the job scripts themselves and their associated variables. Live and historical job output can be viewed with an efficient built-in viewer that can handle output files of arbitrary size. Dependencies between nodes can be visualised in graphical form, and a built-in log analyser can aid in fine-tuning the workflow.

ecFlowUI can monitor several ecFlow servers at once, with facilities to display only those suites or tasks of interest. It can also be used to move a set of tasks from one server to another.

ecFlow version 5

One limiting factor of ecFlow 4 was that its client/server communication was sensitive to changes in the version of the boost library that it links with. This meant that a single client could not necessarily communicate with all the running servers if they had been built with different versions of boost. The technology also limited the ability to make even small changes in communication protocol, which is sometimes necessary in order to allow new features. ecFlow 5 now uses the JSON format for communication, and clients and servers are free to use different versions of boost. This change also allows for new features to be added without breaking compatibility with older servers or clients. With further improvements to the communication, ecFlowUI can now communicate with servers using fewer network requests, meaning less network traffic. An internal improvement is that ecFlow 5 uses features from the C++14 standard, simplifying some code and providing performance benefits.

ecFlow 5 has a number of additional new features requested both by ECMWF users and by Member and

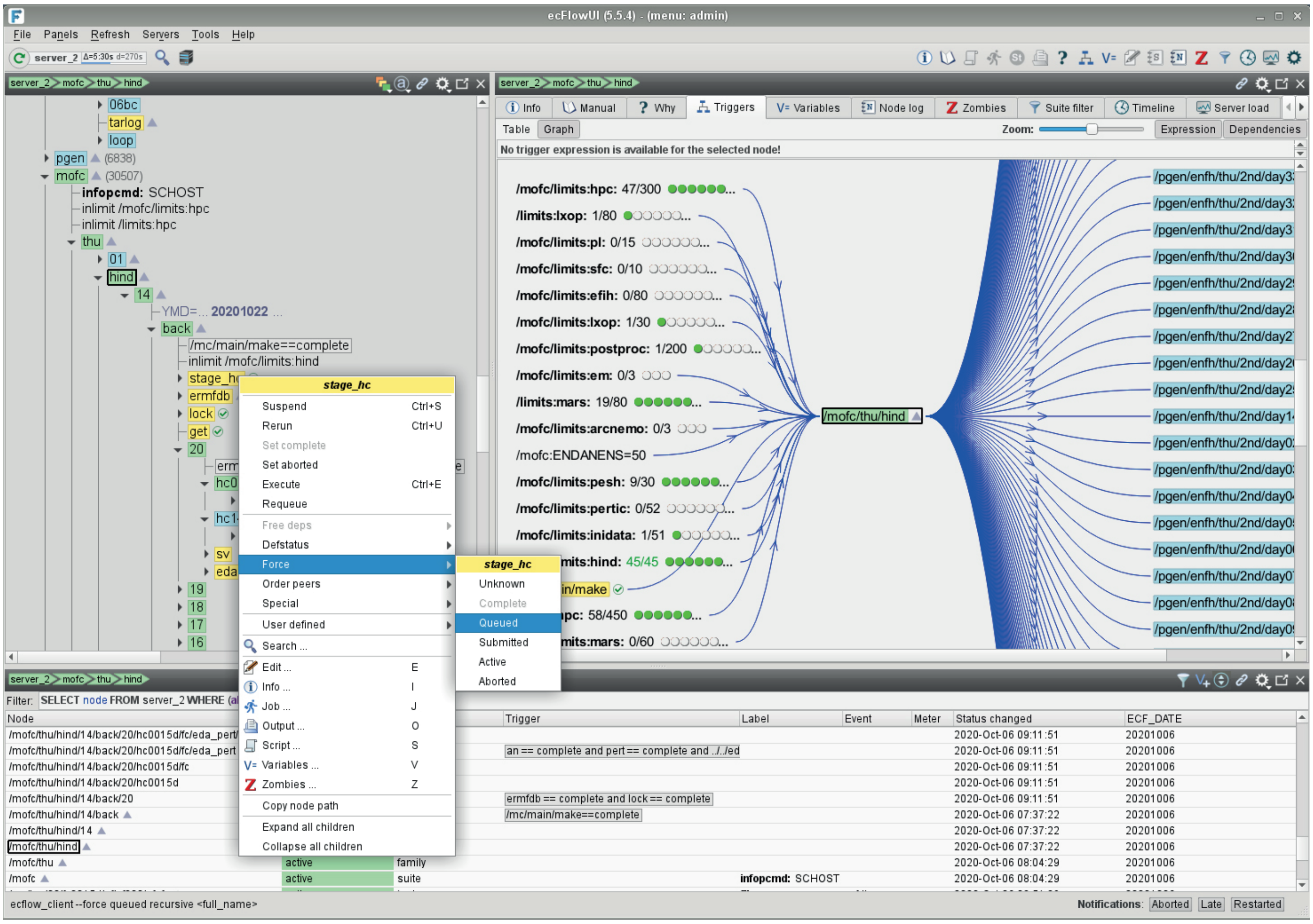


FIGURE 3 ecFlowUI provides a rich environment for viewing and interacting with suites, including a new Trigger Graph view showing dependencies between items in the suite.

Co-operating States. These include:

- Improved security features, such as integrated SSL and password-based access; ecFlowUI can now view both SSL and non-SSL based servers in the same session.
- ecFlowUI now has an interactive trigger graph view to show the interdependency of nodes and attributes.
- Servers now support auto-archive and auto-restore, allowing parts of a suite to be dynamically written to disk when complete and restored later on. This aids the handling of extremely large suites.
- Improved features to help users diagnose problems, for example when a worker machine goes down or a running job becomes detached from the server's records.
- Additional controls to limit the number of submitted or active tasks.
- Various smaller features to help refine the suite definitions.

ecFlow's stability has been validated by daily operational use at ECMWF. In addition, a slew of tests are run every night to ensure that no regressions creep into its releases. With its maturity and proven fitness-for-purpose, future work on ecFlow will emphasise the continuation of this maintenance and stability rather than large new developments.

Migrating to ecFlow 5

Many operational servers at ECMWF have already been migrated to ecFlow version 5. Once ECMWF's computing centre has moved to Bologna, only version 5 will be available. Fortunately, migration from ecFlow 4 to 5 is straightforward and mostly involves stopping the currently running server and then starting it up again using ecFlow 5. The migration page provides more details (*https://confluence.ecmwf.int/display/ECFLOW/Migration+to+ecflow+5*). It is important to note that only an ecFlowUI from version 5 can be used with a version 5 server due to the change in communication protocol. Also noteworthy is that although current versions of ecFlow are built with Python 2 and 3 support, once operational in Bologna only Python 3 will be supported. It is therefore advisable to ensure that any suites are migrated as soon as possible in order to avoid any last-minute problems.

Availability

ecFlow is installed on all of ECMWF's computing platforms, including the Member and Co-operating State server ecgate. If you plan to run an operational ecFlow server at ECMWF, please contact User Services, who will be glad to guide you on the best way to set it up. There are currently a default and a new version of ecFlow 5. To use either one of these, use the commands:

```
module load ecflow/5
```

```
module load ecflow/5new
```

For use external to ECMWF's computing platforms, ecFlow is also available as a binary installation on the conda platform, available through the conda-forge channel with this command:

```
conda install ecflow -c conda-forge
```

The source is also available on github (*https://github.com/ecmwf/ecflow*) or as a tarball from the ecFlow Confluence pages (*https://confluence.ecmwf.int/display/ECFLOW/*).

ECMWF Council and its committees

The following provides some information about the responsibilities of the ECMWF Council and its committees. More details can be found at:
http://www.ecmwf.int/en/about/who-we-are/governance

Council

The Council adopts measures to implement the ECMWF Convention; the responsibilities include admission of new members, authorising the Director-General to negotiate and conclude co-operation agreements, and adopting the annual budget, the scale of financial contributions of the Member States, the Financial Regulations and the Staff Regulations, the long-term strategy and the programme of activities of the Centre.

President Dr Daniel Gellens *(Belgium)*

Vice President Prof. Penny Endersby *(UK)*

Policy Advisory Committee (PAC)

The PAC provides the Council with opinions and recommendations on any matters concerning ECMWF policy submitted to it by the Council, especially those arising out of the four-year programme of activities and the long-term strategy.

Chair Mr Eoin Moran *(Ireland)*

Vice Chair Ms Virginie Schwarz *(France)*

Finance Committee (FC)

The FC provides the Council with opinions and recommendations on all administrative and financial matters submitted to the Council and exercises the financial powers delegated to it by the Council.

Chair Dr Gisela Seuffert *(Germany)*

Vice Chair Mr Lukas Schumacher *(Switzerland)*

Scientific Advisory Committee (SAC)

The SAC provides the Council with opinions and recommendations on the draft programme of activities of the Centre drawn up by the Director-General and on any other matters submitted to it by the Council.
The 12 members of the SAC are appointed in their personal capacity and are selected from among the scientists of the Member States.

Chair Dr Inger-Lise Frogner *(Norway)*

Vice Chair Dr Alain Joly *(France)*

Technical Advisory Committee (TAC)

The TAC provides the Council with advice on the technical and operational aspects of the Centre including the communications network, computer system, operational activities directly affecting Member States, and technical aspects of the four-year programme of activities.

Chair Dr Philippe Steiner *(Switzerland)*

Vice Chair Dr Sarah O'Reilly *(Ireland)*

Advisory Committee for Data Policy (ACDP)

The ACDP provides the Council with opinions and recommendations on matters concerning ECMWF Data Policy and its implementation.

Chair Mr Paolo Capizzi *(Italy)*

Vice Chair Ms Monika Köhler *(Austria)*

Advisory Committee of Co-operating States (ACCS)

The ACCS provides the Council with opinions and recommendations on the programme of activities of the Centre, and on any matter submitted to it by the Council.

Chair Mr Nir Stav *(Israel)*

Vice Chair Dr Kornélia Radics *(Hungary)*

ECMWF publications

(see *www.ecmwf.int/en/research/publications*)

Technical Memoranda

880 **Haiden, T., M. Janousek, F. Vitart, Z. Ben-Bouallegue, L. Ferranti, C. Prates & D. Richardson**: Evaluation of ECMWF forecasts, including the 2020 upgrade. *January 2021*

878 **Düben, P., U. Modigliani, A. Geer, S. Siemen, F. Pappenberger, P. Bauer, A. Brown, M. Palkovič, B. Raoult, N. Wedi & V. Baousis**: Machine learning at ECMWF: A roadmap for the next 10 years. *January 2021*

877 **Polichtchouk, I., P. Bechtold, M. Bonavita, R. Forbes, S. Healy, R. Hogan, P. Laloyaux, M. Rennie, T. Stockdale, N. Wedi, M. Diamantakis, J. Flemming, S. English, L. Isaksen, F. Vána, S. Gisinger & N. Byrne**: Stratospheric modelling and assimilation. *January 2021*

875 **Sandu, I., T. Haiden, G. Balsamo, P. Schmederer, G. Arduini, J. Day, A. Beljaars, Z. Ben-Bouallegue, S. Boussetta, M. Leutbecher, L. Magnusson & P. de Rosnay**: Addressing near-surface forecast biases: outcomes of the ECMWF project 'Understanding uncertainties in surface atmosphere exchange' (USURF). *November 2020*

874 **Bechtold, P., M. Bramberger, A. Dörnbrack, L. Isaksen & M. Leutbecher**: Experimenting with a Clear Air Turbulence (CAT) Index from the IFS. *January 2021*

ESA Contract Reports

Weston, P. & P. de Rosnay: Annual SMOS brightness temperature monitoring report. *January 2021*

Weston, P. & P. de Rosnay: Quarter 4 2020: Operations Service Report. *January 2021*

ECMWF Calendar 2021

Feb 1–4	Training course: Use and interpretation of ECMWF products
Feb 8–10	Virtual workshop: Weather and climate in the cloud
Mar 15–19	Training course: Parametrization of subgrid physical processes
Mar 22–26	Training course: Predictability and ensemble forecast systems
Apr 14	Advisory Committee for Data Policy
Apr 19–23	Training course: Advanced numerical methods for Earth system modelling
Apr 27–28	Finance Committee
Apr 28	Policy Advisory Committee
May 4–7	Training course: EUMETSAT/ECMWF NWP-SAF satellite data assimilation
May 10–14	Training course: Data assimilation
May 17–20	Joint ECMWF/OceanPredict workshop on advances in ocean data assimilation
May 17–21	Online computing training week
May 24–27	Training course: A hands-on introduction to numerical weather prediction models: understanding and experimenting
Jun 1–4	Using ECMWF's Forecasts (UEF2021)
Jun 29–30	Council
Jun 29 – Jul 1	Joint workshop on connecting global to local hydrological modelling and forecasting
Sep 13–17	Annual Seminar 2021
Sep 20–24	19th Workshop on high-performance computing in meteorology
Oct 4–6	Scientific Advisory Committee
Oct 4–7	Training course: Use and interpretation of ECMWF products
Oct 6	Advisory Committee of Co-operating States
Oct 7–8	Technical Advisory Committee
Oct 27–28	Finance Committee
Oct 27	Policy Advisory Committee
Dec 2–3	Council

Contact information

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

Telephone National	0118 949 9000
Telephone International	+44 118 949 9000
Fax	+44 118 986 9450

ECMWF's public website *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).

For any query, issue or feedback, please contact ECMWF's Service Desk at *servicedesk@ecmwf.int*.

Please specify whether your query is related to forecast products, computing and archiving services, the installation of a software package, access to ECMWF data, or any other issue. The more precise you are, the more quickly we will be able to deal with your query.

Newsletter | **No. 166** | Winter 2020/21
European Centre for Medium-Range Weather Forecasts
www.ecmwf.int