



# ECMWF Newsletter



.....

Number 143 – Spring 2015

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

Skill of cloudiness forecasts  
.....

Atmospheric composition in the IFS  
.....

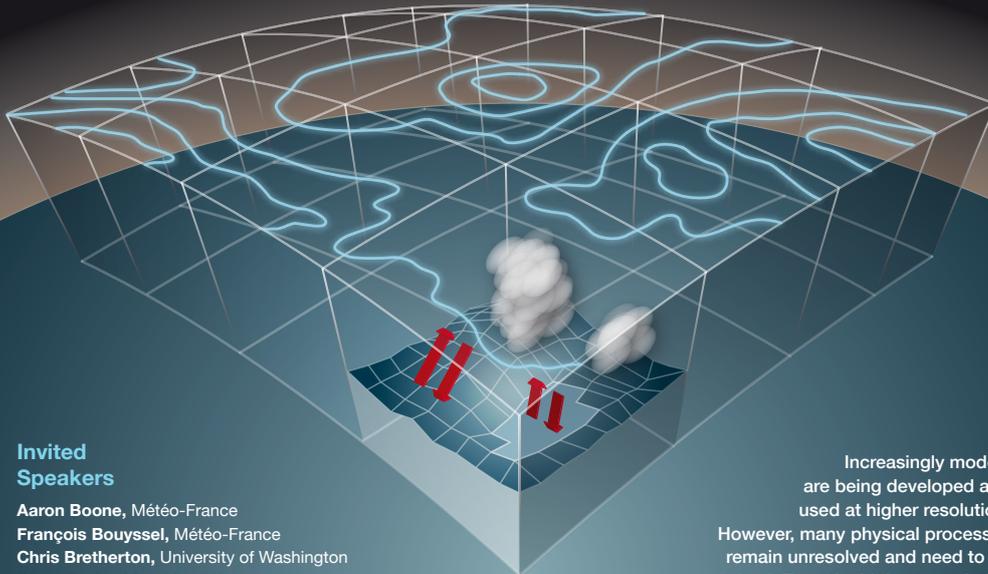
Snow data assimilation  
.....

Supercomputing at ECMWF  
.....

# ANNUAL SEMINAR 2015

## Physical processes in present and future large-scale models

1–4 September



### Invited Speakers

Aaron Boone, Météo-France  
François Bouyssel, Météo-France  
Chris Bretherton, University of Washington  
Steve Derbyshire, Met Office  
Paul Dirmeyer, George Mason University  
Andreas Dörnbrack, DLR  
Saulo Freitas, CPTEC  
Andrew Gettelman, NCAR  
Marat Khairoutdinov, Stony Brook University  
Martin Köhler, DWD  
Jason Milbrandt, Environment Canada  
Louise Nuijens, MPI  
Peter Preusse, Forschungszentrum Jülich GmbH  
Sonia Seneviratne, ETH Zürich  
Keith Shine, University of Reading  
Gunilla Svensson, Stockholm University  
Martin Wild, ETH Zürich

### ECMWF Speakers

Peter Bechtold  
Anton Beljaars  
Richard Forbes  
Alan Geer  
Thomas Haiden  
Robin Hogan  
Marta Janisková  
Martin Leutbecher  
Philippe Lopez  
Irina Sandu

Increasingly models are being developed and used at higher resolution.

However, many physical processes remain unresolved and need to be parametrized. Even the highest-resolution limited area models still need a parametrized representation of shallow convection, turbulence, microphysics, radiation and land surface processes. Schemes for deep convection and sub-grid orography will still be needed in the foreseeable future for global NWP and climate models.

The seminar will give an overview of the relevant issues. Emphasis will be on the understanding, observation and representation of processes across scales. This is particularly relevant at very high resolution where part of the process, e.g. convection and small-scale orography, is resolved.



<http://www.ecmwf.int/en/annual-seminar-2015>

email: [events@ecmwf.int](mailto:events@ecmwf.int)

© Copyright 2015

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

Literary and scientific copyright belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director-General. Appropriate non-commercial use will normally be granted under condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

**CONTENTS****EDITORIAL**

Scalability challenge. .... 1

**NEWS**

Work on Copernicus Climate Change Service under way. .... 2

El Niño set to strengthen but longer-term trend uncertain... 3

Upbeat mood as MACC project draws to a close..... 4

Forecasts for US east coast snow storm in January 2015 ..... 6

New training module for Metview software..... 7

Benefits of statistical post-processing ..... 8

Modelling the Quasi-Biennial Oscillation..... 8

Warm conditions continue from 2014 into 2015 ..... 9

The role of hindcast length in assessing seasonal climate predictability ..... 11

Stochastic workshop explores simulation of forecast model uncertainty..... 12

Piotr Smolarkiewicz granted Poland's top academic title .... 13

Annual Seminar proceedings published ..... 13

**METEOROLOGY**

The skill of ECMWF cloudiness forecasts ..... 14

Atmospheric composition in ECMWF's Integrated Forecasting System..... 20

Snow data assimilation at ECMWF ..... 26

**COMPUTING**

Supercomputing at ECMWF ..... 32

**GENERAL**

ECMWF publications..... 38

ECMWF Calendar 2015 ..... 39

Contact information ..... 39

Index of newsletter articles ..... 40

**PUBLICATION POLICY**

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

Editor: Georg Lentze

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

Any queries about the content or distribution of the *ECMWF Newsletter* should be sent to [Georg.Lentze@ecmwf.int](mailto:Georg.Lentze@ecmwf.int)  
Guidance about submitting an article is available at [www.ecmwf.int/en/about/news-centre/media-resources](http://www.ecmwf.int/en/about/news-centre/media-resources)

**CONTACTING ECMWF**

Shinfield Park, Reading, Berkshire RG2 9AX, UK

Fax: +44 118 986 9450

Telephone: National 0118 949 9000

International +44 118 949 9000

ECMWF website [www.ecmwf.int](http://www.ecmwf.int)

## Scalability challenge

Why is weather prediction so difficult? It's amazing that it can be done at all! It needs a huge volume of observations to describe the current state of the weather globally and an excellent scientific understanding of the laws of physics that govern the atmosphere (and oceans etc.). It needs numerical methods to solve the mathematical equations that describe the laws of physics and a supercomputer able to solve these equations in time for the forecast to be useful. It also needs data networks to disseminate the forecast data to users.

But arguably the part that is the most mysterious is the computer because it does something that is way beyond the capacity of the human brain. The computer is involved in all stages of the process, from the observational data handling at one end to the use of the forecasts by society at the other. Perhaps it is not surprising that to do this computers are complex to construct and consume significant amounts of energy.

How the computer works, how we present information to it, and how we formulate the equations for it to solve all make a real difference to the forecasts that are produced. To describe the computer as a tool is to diminish the significance of its role at the core of weather prediction. For example, we can be in the strange position of trying to inject an estimate of uncertainty into the equations the computer solves at the same time as requiring the computation to be done to an amazingly high level of precision. The future requires even more prodigious amounts of computing power to be at our disposal for forecasts to be at the level of detail, skill and confidence that users require.

We are reaching a computing crossroads where a new generation of computing systems with exascale capabilities promise much greater energy efficiency – but this will rely on parallel processing at levels to which current numerical weather prediction (NWP) codes are not adapted. Changes are needed throughout the entire NWP processing chain if we are to exploit these new opportunities for energy efficiency. The aim has to be to use the computer as effectively and efficiently as possible and in ways consistent with the kind of forecasts we want to create. In a nutshell, this is what the ECMWF Scalability Programme, launched in 2013, is aiming to do and it is a game-changer by any stretch of the imagination.

This programme is collaborative in so many ways – across member states, with computer vendors, with computer scientists, linking weather and climate modelling, etc. It brings together experts from around the world for a coordinated approach to hardware and software development. It is likely to result in a revolution – albeit a quiet one – in the way numerical weather prediction is done.

The key is to address the scalability challenge by using novel mathematical solutions and computing techniques. The programme's objectives are: an integrated forecasting system combining a flexible framework to apply the latest science with maximum achievable parallelism; portable code structures ensuring efficiency and code readability, and exploiting a range of expected future technologies; and metrics and a framework for code testing, allowing a quantitative assessment of scalability. The sky's the limit!

**Alan Thorpe**

## Work on Copernicus Climate Change Service under way

### GEORG LENTZE

ECMWF has begun to host a series of international workshops to prepare for the implementation of the EU-funded Copernicus Climate Change Service (C3S).

The goal of C3S is to provide authoritative, quality-assured information about the past, current and future states of the climate in Europe and worldwide. The service is being implemented by ECMWF as part of the Copernicus Earth observation programme coordinated and managed by the European Commission.

A web-based Climate Data Store (CDS) will give users of the service access to climate information. More than 70 participants from European institutions, national meteorological services, research institutes and companies in 19 countries came together at ECMWF from 3 to 6 March to discuss the best way forward on the CDS.

Speaking at the end of the workshop, Head of Copernicus Climate Change Service Jean-Noël Thépaut said: *“This first event is a milestone for C3S as it is the very first workshop engaging the community since the signing of the delegation agreement with the European Commission. This workshop has been a great opportunity for us and the participants to better define the requirements of a wide variety of users, to hear about existing solutions and their strengths and weaknesses,*



**More than 70 participants** from European institutions, national meteorological services, research institutes and companies in 19 countries came together at ECMWF.

*to identify gaps, and to discuss ways to address them.”*

He added: *“We saw a number of very impressive web portals dealing with climate information, showing that C3S can build upon existing capabilities and provide an enhanced service at European level. Another important element is the need to tailor the CDS and associated portal functionalities to the wide potential user base of C3S, from European Directorates-General to climate scientists and small and medium-sized enterprises. The working groups have been extremely helpful in advising on key aspects for the development of the CDS, such as co-design, toolboxes, standards and required interoperability.”*

#### **“Make the data relevant”**

Ned Gardiner from the US National Oceanic and Atmospheric Administration (NOAA), who gave a presentation on the US Climate

Resilience Toolkit, emphasized the need to connect with users. While it is “never a mistake to just provide the data”, the real challenge is to enable users to “gain meaning” from the material, said Ned, who has long worked in the field of climate data visualization. It is important to “make the data relevant to people,” he feels, and one way of doing this is to “connect with people’s values”.

Another participant, MET Norway Assistant Research Director Trond Iversen, agreed that “the key word is relevance”. “Access to quality-assured data which are relevant for each user’s needs is very important for C3S to be successful,” he said. “Any user who needs to take into account potentially adverse impacts of weather, sea-state, sea level in their planning for the coming decades should ideally have as much relevant data as possible and nothing more,” according to Trond.

A second workshop on climate projections took place in Reading on 20 and 21 April. It will be followed by workshops on climate communication to take place in mid-June in Brussels, and on climate observation requirements from 29 June to 2 July. These activities are part of the two-year C3S proof-of-concept stage. It is expected that some elements of the service will become available in this period.

ECMWF is also implementing the Copernicus Atmosphere Monitoring Service (CAMS). C3S and CAMS are two of the six services which make up the EU’s Copernicus Earth observation programme.



**Jean-Noël Thépaut** is the head of the Copernicus Climate Change Service at ECMWF.

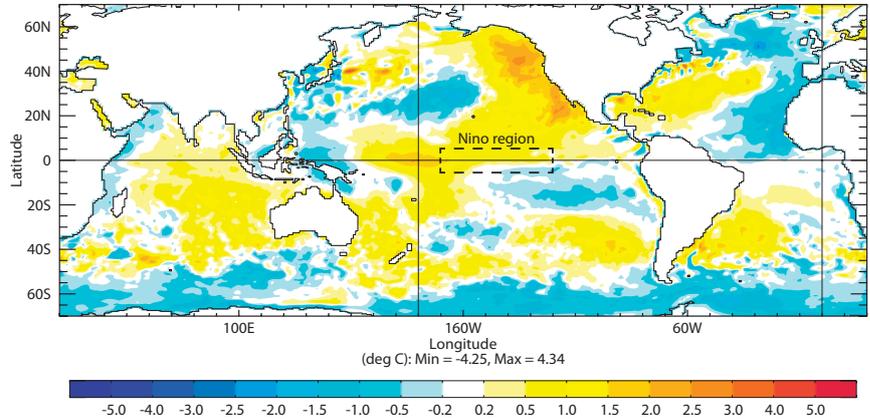
# El Niño set to strengthen but longer-term trend uncertain

**TIM STOCKDALE**

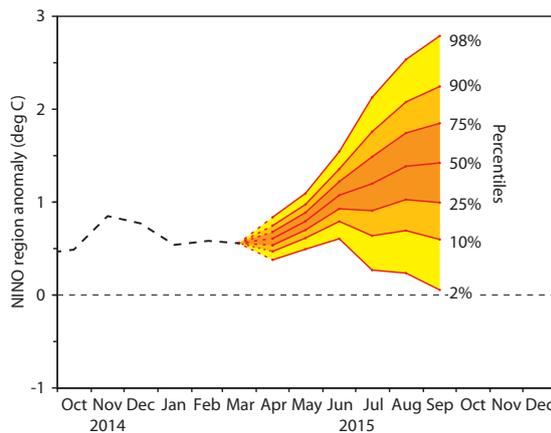
Weak El Niño conditions observed since October 2014 are expected to strengthen over the next few months and may turn into a substantial event, but there is a great deal of uncertainty beyond June 2015, current forecasts suggest.

An El Niño event is a prolonged period of abnormally high sea-surface temperatures in the tropical Pacific Ocean. It goes hand in hand with changes in atmospheric conditions and can have strong repercussions on global weather patterns. El Niño can also significantly affect the global average temperature and hence influence the global warming signal.

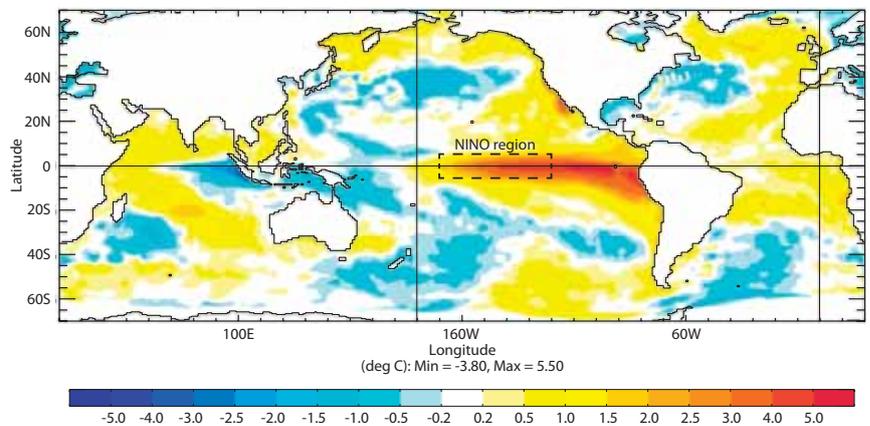
The latest ECMWF seasonal forecast from 1 April 2015 shows that in the coming months the equatorial Pacific surface temperature is expected to rise further, giving a moderate strengthening of El Niño conditions. After June, the uncertainty becomes large: in some ensemble members El Niño conditions start to dissipate, in others the event remains large or in some cases amplifies further. The very warmest ensemble members produce temperature anomalies which are considered unrealistic, due to limitations in how the model responds to extreme El Niño events. Within the usual range of El Niño variability, though, the forecast products correctly represent the possible amplitudes of the expected anomaly. As always in El Niño forecasting, the eventual outcome will depend very much on how strongly the winds respond to the warming sea-surface temperatures. Ensemble forecasts account for the uncertainties inherent in the prediction of weather and ocean parameters by producing a set of possible outcomes. However, different forecast models and different analyses of the initial state will give slightly different predictions for the range of possible outcomes. Experience has shown that, at seasonal timescales, the most reliable forecasts are obtained by combining and calibrating the output from a number of independent forecasting systems.



**Average sea-surface temperature anomalies in March 2015.** The chart shows sea-surface temperature anomalies compared to the 1981–2009 average.



**EUROSIP probability spread of sea-surface temperature anomalies.** The chart shows the probability spread of sea-surface temperature anomalies over the Niño3.4 region in the central-eastern equatorial Pacific, produced on 1 April 2015 from the ensembles provided by four different forecasting centres: ECMWF, the UK Met Office, Météo-France and the US National Centers for Environmental Prediction.



**Average sea-surface temperature anomalies in November 1997.** The chart shows sea-surface temperature anomalies, compared to the 1981–2009 average, at the height of the 1997–1998 El Niño event.

At ECMWF, output from four major forecasting centres (ECMWF, the UK Met Office, Météo-France and the US National Centers for Environmental Prediction) is combined to make the EUROSIP multi-model products. The spread of the probability distribution for El Niño temperature anomalies calculated by EUROSIP supports the ECMWF model in showing a large uncertainty after June. It suggests that there is approximately a 75% chance that sea-surface temperature anomalies will exceed 1 degree Celsius by August. Conversely, there is a 25% chance that they will not.

The uncertainty in the development of the current El Niño conditions has prompted the World Meteorological Organization (WMO) to be cautious. “While current forecasts imply that

a careful watch must be kept on the tropical Pacific Ocean temperatures, it is too early to assess the strength of any potential event,” the WMO said in its El Niño update published on 16 March.

#### An “exceptional” El Niño in 1997

Present ocean conditions have both similarities and differences to those that preceded an event in 1997, when a combination of a large ocean signal and strong westerly winds led to an ‘El Niño of the century’. The event of 1997 coincided with the start of experimental real-time seasonal forecasting at ECMWF. The success in predicting both the El Niño and its global impacts helped raise the profile of seasonal forecasting at ECMWF.

In December 1997, at the height of the El Niño event, ECMWF began

publishing seasonal predictions on its website. The text accompanying the products explained that, “taking into account the exceptional El Niño event of 1997, and following overarching WMO requirements, the ECMWF Council has decided to make a range of products from the experimental programme of seasonal prediction available on this site.”

Today, ECMWF regularly publishes a broad range of long-range or seasonal forecasts on its website, including El Niño ensemble forecasts, expected anomalies and probability information on distributions of variables such as surface air temperature, precipitation and mean sea level pressure, and probabilistic outlooks for tropical storm numbers and any shifts in the distribution of their tracks.

## Upbeat mood as MACC project draws to a close

**RICHARD ENGELEN,  
GEORG LENTZE**

MACC-III, the third phase of the EU-funded Monitoring Atmospheric Composition and Climate project, held its General Assembly at ECMWF in Reading from 19 to 21 January.

The meeting was the final one in a long series that started in July 2005 in Hamburg, Germany, with the launch of the GEMS project (Global and regional Earth-system Monitoring using Satellite and in situ data). Since then, the GEMS, PROMOTE, MACC, MACC-II, and MACC-III projects have been building the pre-operational Copernicus Atmosphere Monitoring Service (CAMS), which is being implemented by ECMWF and will enter its operational phase later this year.

#### “Visionary”

Several speakers referred to the historic nature of the moment and highlighted the impressive amount of progress that has been made since the early days. Speaking on the sidelines of the event, Brian Kerridge, Head of the Remote Sensing Group at the Rutherford Appleton Laboratory in the UK, said the original concept behind MACC had been “visionary”.

Those working on the project “have pioneered a state-of-the-art system for monitoring and forecasting atmospheric



*The General Assembly* brought together more than 100 participants from a dozen countries.

composition, and they’ve done that in a unique way by adopting an integrated approach,” Brian said (see Box).

Paolo Laj, the director of the Laboratoire de Glaciologie et Géophysique de l’Environnement at the University of Grenoble, France, said MACC’s biggest achievement was to have held the community of experts together for such a long time. “There is now a demand from society to apply what we have been doing over the last 20 years,” he added.

#### “Blueprint” for other regions

David Edwards, the interim director of the NCAR Earth System Laboratory in Boulder, Colorado, said there was a need to look beyond Europe since MACC could be “a blueprint for how this integrated system could be achieved in other countries, in developing countries.”

Guy Brasseur, a scientist at the Max Planck Institute for Meteorology,

Hamburg, agreed. Pointing out that there is nothing like MACC in the USA, Asia or South America, he said that “in fact MACC is not a European project, it’s a world-wide project supported by Europe.”

Participants in the General Assembly discussed progress on various topics, and a one-day plenary meeting presented the current state of MACC-III to the European Commission’s project officer and to project reviewers. Concluding meetings were held between the MACC-III Management Board and the Advisory Board and User Advisory Board.

The discussions and presentations showed that the pre-operational system is ready to be made operational, while at the same time there is no lack of ideas on how to further improve the services provided. Everyone agreed that CAMS will need the right balance between robust operational services and further development based on the extensive expertise available in Europe.

## Four views on MACC's achievements and the way forward



**Brian Kerridge**, Head of the Remote Sensing Group at the Rutherford Appleton Laboratory, Oxfordshire, UK – member of the MACC Advisory Board

*“The original concept behind MACC was visionary. Those working on the project have pioneered a state-of-the-art system for monitoring and forecasting atmospheric composition, and they’ve done that in a unique way by adopting an integrated approach.*

*“Through this excellent initiative, we’re at the point of being able to use satellite data to inform our understanding of air quality and its impact on human health. In the future, we’re going to go one step further and unravel what’s happening right at the surface, where ozone affects ecosystem health and agriculture.*

*“Going forward into the Copernicus Atmosphere Monitoring Service, we expect to see further innovation, further development, and to retain this cutting-edge capability in the European system.”*



**David Edwards**, Interim Director NCAR Earth System Laboratory, Boulder, Colorado, USA – member of the MACC Advisory Board

*“I think the great achievement of the MACC project in general is to have brought together observationalists and modellers in a way which we often talk about but don’t often manage to achieve.*

*“One thing the MACC project has done as well is to provide a blueprint for how this integrated system could be achieved in other countries, in developing countries, and it’s showing a way forward for how the system could be applied not only to the European domain but to other regions as well, such as China.*

*“We are looking to a future of having these coupled models where, instead of talking about weather forecasting, we’re really going to be talking about Earth-system prediction and being able to say what is happening to the Earth system as a system.”*



**Paolo Laj**, Director of the Laboratoire de Glaciologie et Géophysique de l'Environnement, CNRS, University of Grenoble, France – member of the MACC Advisory Board

*“There are a lot of achievements, the biggest of which is that the community has held together for more than 10 years.*

*“There is now a demand from society to apply what we have been doing over the last 20 years. There are still gaps, but they can be addressed in parallel with the first applications.*

*“There could be a greater focus on countries which face big air quality issues, such as countries in Africa and South America. In that area of the world there may well be an expectation that Europe will accept a share of responsibility in trying to improve the environmental situation there.”*



**Guy Brasseur**, Max Planck Institute for Meteorology, Hamburg, Germany – former member of the MACC Advisory Board

*“The project has established for the first time the capability of bringing together satellite, surface and aircraft observations and also modelling in order to provide a way of predicting air quality at the global scale – that’s what’s really unique.*

*“The biggest achievement is to have created a unique team of specialists from all over Europe who can deal with different aspects: emissions from industry, emissions from fire, the way the meteorology disperses pollutants, also people looking at satellite data, people doing measurements in the field, people working on forecasts or on assimilating data into the model.*

*“Essentially this serves the whole world: there is nothing like this in the US, in Asia or South America, and they all need the data. I really believe that in fact MACC is not a European project, it’s a world-wide project supported by Europe.”*

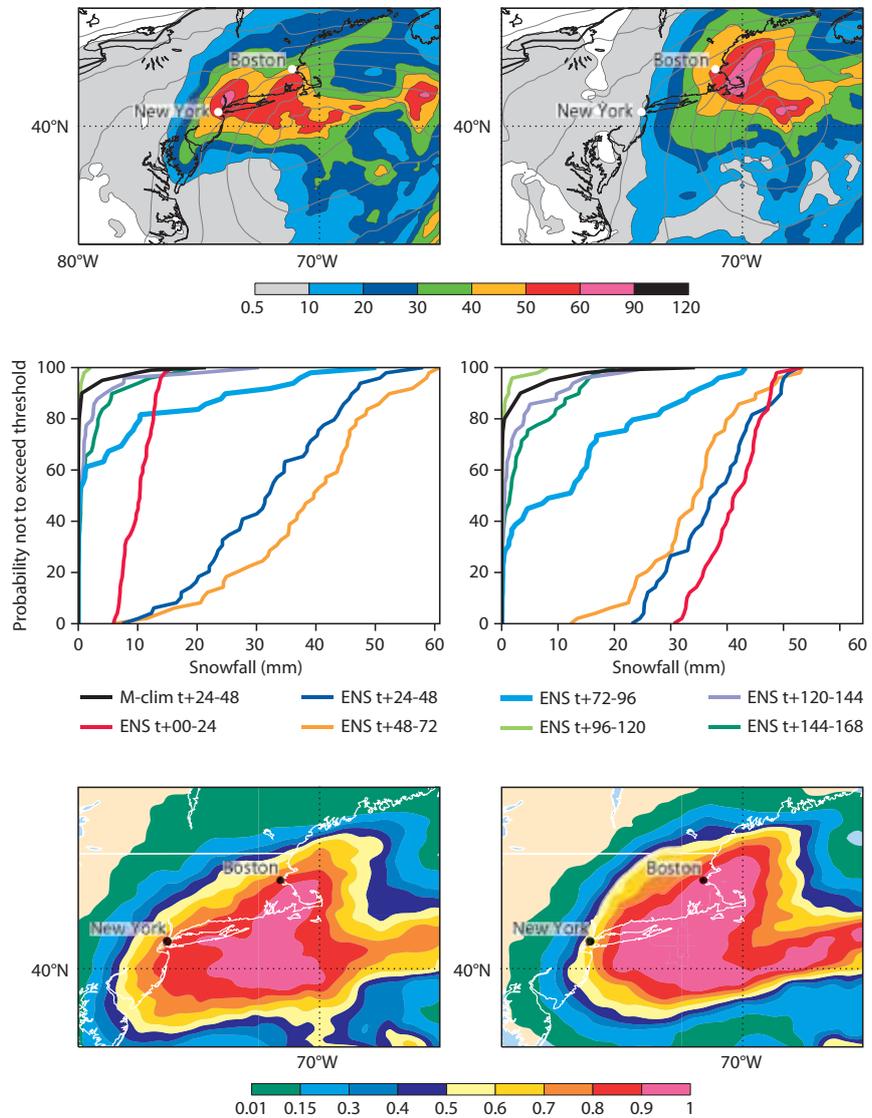
# Forecasts for US east coast snow storm in January 2015

DAVID RICHARDSON,  
ERVIN ZSOTER,  
LINUS MAGNUSSON,  
TIMOTHY HEWSON,  
FLORENCE RABIER

A severe snow storm caused major disruption to parts of the north-east USA on 27 January. ECMWF forecasts consistently predicted the development of a major snow storm, but small changes in the predicted track led to large differences in snowfall forecasts for New York City.

On 27 January 2015, a blizzard hit north-eastern parts of the USA. The storm was expected to affect an area extending from the state of New Jersey north-eastwards, which includes New York City (NYC) and Boston, and preventive actions were taken (closing subways, motorways etc.). However, New Jersey and NYC only received relatively little snow, with the worst affected areas being in a band from eastern Long Island towards Boston and further north. This was the beginning of one of the snowiest periods on record for Massachusetts.

The snowfall was caused by a cyclone that formed close to Florida early on 26 January and then rapidly developed over the Atlantic late that day as it moved northward close to the US east coast. Both the actual and the predicted areas of heavy snowfall exhibited sharp gradients on the western side that were close to NYC, as seen in the top panels of the figure. Around NYC itself, 10–25 cm of snow was observed (e.g. Newark airport – 16 cm, Central Park – 24 cm, La Guardia airport – 28 cm), though Islip on Long Island, 40 km to the east, recorded over 50 cm. Further northeast many stations reported more than 75 cm of snowfall with the worst affected areas being in Massachusetts. It is not at all unusual to see large gradients in total precipitation in the area to the left of a cyclone's track; this can occur throughout the extra-tropical regions of the northern hemisphere. The first indication in ECMWF forecasts of a potential major snow storm affecting the north-east USA appeared over a week ahead, although



**Top panels** High-resolution forecast (HRES) of 24-hour precipitation (mm of rainfall equivalent) valid 27 January (shade) and mean sea level pressure valid at 12 UTC on 27 January (contour) from 00 UTC on 25 January (top-left) and 00 UTC on 27 January (top-right).

**Middle panels** Ensemble forecast (ENS) cumulative distribution functions (CDFs) for snowfall (defined as mm of rainfall equivalent, diagnosed as snow) from forecasts valid 27 January for NYC (middle-left) and Boston (middle-right).

**Bottom panels** Probability for more than 30 mm of precipitation on 27 January in ENS forecasts from 00 UTC on 25 January (bottom-left) and 00 UTC on 26 January (bottom-right).

the next few forecasts generally kept any developing systems out in the Atlantic, with limited impact on coastal areas. From Saturday 24 January onwards, successive ECMWF forecasts consistently predicted the development of a major snow storm that would affect the north-east coast of the USA on 27 January. However, whilst there was high confidence in the occurrence

of a severe storm, its precise track remained less certain.

The lines in the middle panels of the figure represent ECMWF ensemble forecasts (ENS) from different initial dates for snowfall on 27 January for NYC (left) and Boston (right). Early forecasts (dark green, lilac, light green lines) suggested a high probability of at most small amounts of snowfall.

Later forecasts (light blue and yellow lines) indicated a greater probability of larger amounts of snowfall. For Boston this assessment remained stable in the very latest forecasts (dark blue and red lines), but for NYC the latest forecasts moved back to the left in the chart, suggesting once again a much smaller probability of large snowfall amounts.

On 25 January, most ENS members were taking a slightly more westerly track than the forecasts from the previous day. A change of 100–150 km in position (only 3–5 grid-lengths of the model) had a very large impact on the snow forecasts for NYC in particular: the probability of ‘substantial snow’ (more than 30 mm of precipitation, say, equivalent roughly to 30 cm or 1 foot of fresh snow) was over 80% for both NYC and Boston. This outcome was reinforced by the HRES (high resolution) forecast (top-left panel), which was predicting over 40 mm precipitation widely in north-eastern coastal regions, with over 50 mm in NYC. Although these totals were towards the upper end of the range predicted by the ENS, the overall signal was that a major event was likely to occur in both NYC and Boston. The ENS probabilities (bottom-left

panel in the figure) show a sharp gradient over the New York area. Those ENS members that did maintain a track further to the east showed that significantly less snow could occur over NYC even though that was a less likely outcome.

Later forecasts from 25 and 26 January consolidated the likely outcome for Boston: the probability of an exceptional snowfall event increased with each new forecast, as shown by a rightward shift and a steepening of the lines in the right-hand chart of the middle panel. In contrast, the situation for NYC remained uncertain. Indeed the likelihood of ‘substantial snow’ decreased from 80% to 60% and then 40% as more ensemble members predicted the storm to pass slightly further to the east (bottom-right panel).

NYC was consistently predicted to be near the edge of the area forecast to be affected by the storm. Because of the very sharp gradients in precipitation, a relatively small change of less than 100 km in the storm’s track could have a very large impact on the snow forecast for NYC. The ENS has a grid spacing of 32 km, and an uncertainty of perhaps one or two grid lengths has to be expected on some occasions.

This was one such occasion where the uncertainty persisted until the last minute: only in the forecast from 00 UTC on 27 January (top-right panel) was the uncertainty in the ECMWF ensemble finally resolved and the chance for severe snowfall in NYC eliminated, which was too late for operational use.

This case highlights the importance of using the full range of forecast information available, and the challenges in communicating this information to users. The ECMWF forecasts predicted potential extreme snowfall for both NYC and Boston. For both cities the forecast 2–3 days ahead indicated that this was highly probable, though not certain.

The ensemble approach is essential to provide forecast users with information about the range of future weather scenarios and the likelihood of high-impact weather events. The ensemble aims to account for all sources of uncertainty in the forecasting system, and ECMWF is continuing to improve its forecasting system, including by increasing its horizontal resolution. The aim is to both reduce and reliably quantify the uncertainty.

---

## New training module for Metview software

-----  
**IAIN RUSSELL, SÁNDOR KERTÉSZ,  
 FERNANDO II**

ECMWF’s meteorological workstation application, Metview, was the focus of a new training module at ECMWF from 9 to 13 March. The redesigned ‘Data Analysis and Visualisation using Metview’ module introduces users to Metview’s tools for accessing, processing and visualising meteorological data. It is part of ECMWF’s ‘Use of Computing Facilities’ training course.

Participants from across Europe with a range of scientific and technical backgrounds learned how to use Metview’s high-level concepts and interactive analysis tools. They also learned how to run their Metview

macros remotely in batch on ecgate, a Linux cluster available exclusively to registered users from Member and Co-operating States. The training module follows the workflow from the prototyping of data and graphical products to routinely generating them in a scheduling system. Participants studied many practical examples from the day-to-day work of analysts and researchers at ECMWF and saw first-hand how the analyst on duty works. This practical approach enabled many participants to develop solutions for their work in their home institutions. Preparations for the course resulted in a large amount of new documentation and tutorials for Metview becoming available online.

Metview brings together various software technologies developed at

ECMWF to provide a complete working environment for both the operational and research meteorologist. Through a single, unified interface it provides access to the MARS archive, allows users to examine and manipulate data formats such as GRIB, BUFR and ODB, and can overlay data to produce high-quality meteorological plots.

ECMWF’s extensive education and training programme supports Member and Co-operating States in training scientists in numerical weather forecasting and in using ECMWF forecast products and computer facilities. ECMWF training courses are freely available to Member and Co-operating States.

For more information on Metview, see: <https://software.ecmwf.int/metview>

## Benefits of statistical post-processing

**CHANTAL DUNIKOWSKI,  
JOANNE JEPPESEN,  
GEORG LENTZE**

ECMWF Fellow Prof Tilmann Gneiting visited the Centre from 11 to 13 February to discuss recent research into statistical post-processing (SPP) for numerical weather forecasts. Prof Gneiting's well-attended seminar on 11 February reported on joint work by the Heidelberg Institute for Theoretical Studies (HITS) and ECMWF in this area.

### Increase in skill

On the sidelines of the seminar, Prof Gneiting explained that SPP can increase the skill of forecasts by a considerable margin. *"With post-processing, forecasts six days ahead might have as much skill as four days ahead without post-processing, that's the order of magnitude in improvement we are dealing with when we think about forecasts for specific locations,"* he said.

He added, however, that SPP worked better for some parameters than for others. While there were considerable benefits for variables such as temperature, wind speed and cloud cover, the gains for precipitation were less pronounced. *"This is something we discussed during this visit and which is still a bit of a mystery to us,"* he said.

### "Exciting news"

Prof Gneiting explained that SPP will continue to be useful even as "raw forecasts" improve. He said: *"The raw forecasts being produced are getting*

*better every year, and SPP makes them even better. The natural hypothesis is that maybe in 20 years' time we will no longer need SPP. Actually, we have found that the benefits of SPP stay about the same even as the raw forecasts improve."*

He added: *"That's exciting news, because we can reap benefits twice: we can continue to improve the model and the ensemble and continue to benefit from SPP. This is great news for the quality of weather forecasts and will be the basis for further collaboration."*

Looking to the future, Prof Gneiting said a key challenge was to extend the application of SPP beyond a single variable and location. *"For the past 10 years, we have mostly focussed on one location, one variable, and one look-ahead time. The ensembles generated at ECMWF honour the physical relationships, but if we then do statistical modelling for each variable individually, we destroy those relationships, and that's highly undesirable,"* he explained.

*"My group and others elsewhere have been looking at a type of technique called the 'empirical copula' approach, which is computationally cheap and easy to implement. These are techniques that we are only beginning to investigate closely now, and that we will need to develop further over the next 5 to 10 years."*

The full interview with Prof Gneiting can be found at:



**Tilmann Gneiting leads the Computational Statistics Group at the Heidelberg Institute for Theoretical Studies** (HITS) and is Professor of Computational Statistics at the Karlsruhe Institute of Technology (KIT) in Germany. A mathematical statistician by training, his interest in statistical post-processing for ensembles and probabilistic forecasting started in the early 2000s at the University of Washington in Seattle, USA.

Prof Gneiting on his collaboration with ECMWF:

Prof Gneiting on his collaboration with ECMWF:

*"Being a Fellow is a great honour for me and has certainly intensified our collaboration. It's also a great opportunity for my group, several of whom have become involved in joint work. The Fellowship Programme enables us to demonstrate our research developments in real-world applications. ECMWF is able to take advantage of the latest research results and can help us to identify areas where future research would bring valuable additional benefits."*

[www.ecmwf.int/en/about/media-centre/news/2015/interview-ecmwf-fellow-professor-tilmann-gneiting](http://www.ecmwf.int/en/about/media-centre/news/2015/interview-ecmwf-fellow-professor-tilmann-gneiting)

## Modelling the Quasi-Biennial Oscillation

**TIM STOCKDALE**

A workshop on the Quasi-Biennial Oscillation (QBO) held in Canada in March agreed to cooperate on a set of experiments and analyses to improve the QBO in models. ECMWF is playing a major role in helping to design the 'initialized forecast' experiments, where we have much experience, and other groups are contributing expertise on longer-term simulations and diagnostics. The Quasi-Biennial Oscillation, which

is relevant to seasonal prediction and climate studies, is one of the most remarkable phenomena in the Earth's atmosphere. High above the equator, in the stratosphere, strong zonal winds blow in a continuous circuit around the Earth. At a given altitude, the winds might start as westerlies, but over time they weaken and eventually reverse, becoming strong easterlies.

### Detailed mechanisms still unclear

Looking at different heights, the peak amplitude of the westerly winds

migrates slowly downwards, with the zone of easterly winds coming behind it also migrating downwards. When the winds approach the base of the stratosphere, at around 80 hPa, they dissipate. New bands of zonal winds appear in the mid-stratosphere, at around 5 hPa, to replace the bands which migrate downwards. At any one time, there is one region of easterlies and one region of westerlies. The whole cycle progresses at a fairly (but not entirely) uniform rate, taking on average

26 months to return to the starting state. Although the basic physics of the QBO is well known, the quantitative details and balances of the different processes are still rather unclear. Worse, many of the models used for numerical weather prediction (NWP) or climate modelling are unable to produce a QBO, or they produce a QBO which looks very different from observations. For example, only four of more than 30 models used for the last IPCC report have any sort of QBO.

At ECMWF, the IFS does have sufficient vertical resolution and physics to allow a reasonable simulation of the QBO, for example as seen in ECMWF's ERA-20CM, a set of extended model runs covering the 20th century. Our seasonal forecasts also have skill in predicting the future evolution of the QBO signal. However, we would like to improve the accuracy and skill of the IFS, so we are working with other

members of the international scientific community to better understand and model the processes driving the QBO.

### Applications

Why is the QBO important? It is certainly relevant for seasonal prediction, where the state of stratospheric winds affects interactions between the tropics and the mid-latitudes, and may also affect the tropical troposphere directly and possibly how the solar cycle interacts with the atmosphere. For groups working on climate, the QBO has a role in modulating transport out of the tropical stratosphere to higher altitudes, thus influencing the concentration of gases in the stratosphere, which in turn might lead to further climate feedbacks.

The poor representation of the QBO in climate change models means that no-one knows what will happen to the QBO in the decades ahead – will it remain largely unchanged, will its period lengthen, or will it change more



**The Quasi-Biennial Oscillation** was the subject of a workshop in the Canadian city of Victoria from 16 to 18 March. Tim Stockdale represented ECMWF at the meeting.

radically? There is also increasing interest in stratospheric wind forecasts in their own right: Google has a programme, dubbed Project Loon, to provide global internet coverage using actively controlled stratospheric balloons. This both requires and contributes to data on stratospheric winds, with forecasts being needed on various timescales to control the system.

## Warm conditions continue from 2014 into 2015

### ADRIAN SIMMONS, PAUL POLI

Some data producers have declared 2014 the warmest year on record, but a comparison of four different datasets paints a more nuanced picture. Meanwhile, 2015 has also started warm, with exceptionally high 12-month means up to February or March in all datasets.

#### Was 2014 a record year?

Two of the four datasets examined are produced using reanalysis: ECMWF's ERA Interim and the Japan Meteorological Agency's JRA-55. The other two are datasets that are conventionally used to characterize the long-term warming of the atmosphere since the 19th century: HadCRUT4, produced by the UK Met Office in collaboration with the Climatic Research Unit of the University of East Anglia, and MLOST, produced by the US National Climatic Data Center (NCDC). Each of these datasets uses a different sea-surface temperature (SST) analysis. NASA's GISTEMP dataset, which is often studied alongside HadCRUT4 and MLOST, uses the same SST analysis as MLOST.

Figure 1 shows time series of 12-month running averages of the four estimates

of global temperature. It is based on the latest available monthly data, running from November 1978 to February 2015 for HadCRUT4 and MLOST, and to March 2015 for the two reanalyses. Running averages rather than calendar-year averages are shown because the warmest or coldest 12-month periods within a warm or cold spell generally do not coincide with a calendar year. HadCRUT4 is an ensemble of 100 realisations; its median and individual elements are plotted.

Each of the datasets provides a similar overall picture: warming since the 1970s is not in doubt, nor is the occurrence of warmer and colder spells linked with El Niño events, volcanic eruptions and other sources of variability. The datasets nevertheless differ in their estimates of the strengths of individual warm spells. ERA-Interim shows the largest peaks, exceeding the 2014 calendar-year average for 12-month periods within 2005–2006 and 2009–2010, including the calendar years of 2005 and 2010. In JRA-55, 2014 is the warmest calendar year by a narrow margin, but there was a warmer 12-month spell in 2009–2010. MLOST shows the lowest pre-2014 maxima; 2014 clearly is the

warmest year for this particular dataset, even though the temperature anomaly for the year is not much different from those of the other datasets.

The range of uncertainty provided by the HadCRUT4 ensemble more-or-less encompasses the spread between the other datasets shown in Figure 1, but the maxima and minima of these other datasets tend to lie close to the extreme ensemble members. This suggests that some of the differences among the datasets are not due to random observational uncertainties, but rather may be understood, and then perhaps reduced in new versions of the datasets.

One difference that has already been studied is a fundamental one that tends to make the reanalyses warmer than the other datasets in recent years. Model-assisted use of a much broader range of observational data, including information on sea-ice cover, enables the reanalyses to fully resolve warm anomalies in the Arctic, where HadCRUT4 and MLOST have limited data coverage. Other differences that merit further study and may be more reconcilable are weaker overall warming trends in the SSTs used by JRA-55 and MLOST, occasional

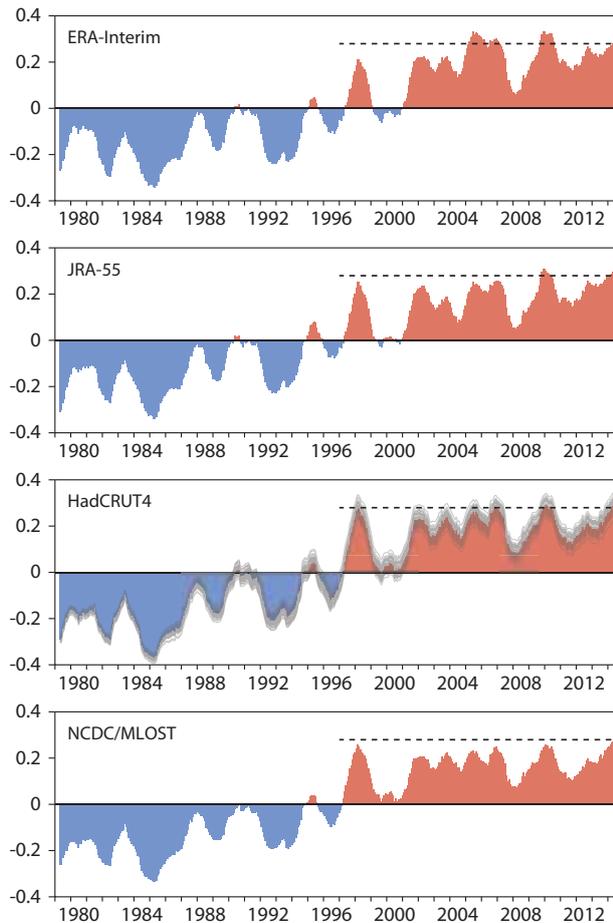
discrepancies between the reanalyses over the Antarctic, where ERA-Interim is distinctly warmer than JRA-55 in 2005, and more warming in recent years over land for MLOST than for the three other datasets.

**Warm start to 2015**

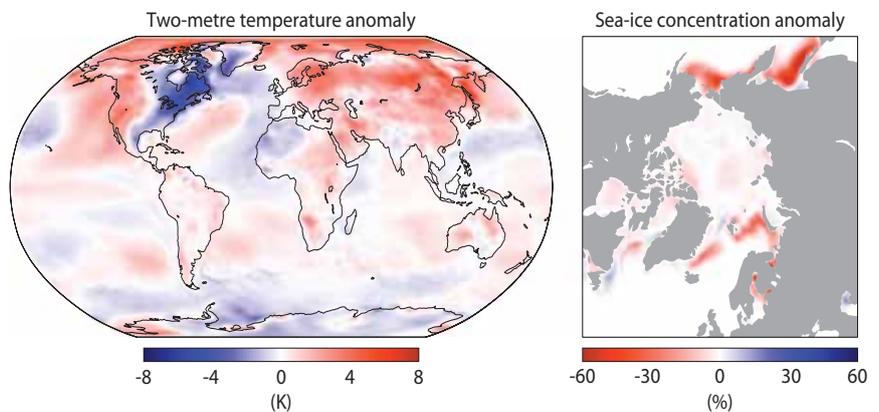
Conditions have remained warm, overall, for the start of 2015. Figure 2 shows maps from ERA-Interim of anomalies in average surface air temperature and sea-ice concentration for January to March. This period was anomalously warm over much of Europe, Asia, the eastern North Pacific, western North America and the Arctic Ocean. The US National Snow and Ice Data Center reported that its lowest recorded winter-maximum Arctic sea-ice extent occurred in late February. Below-average sea ice and associated above-average temperatures for the season were prominent over the Bering and Okhotsk Seas. In contrast, eastern North America was anomalously cold, though the warmth of the adjacent Atlantic Ocean likely contributed, through moisture supply, to the large eastern-seaboard snowfalls that caught media attention this winter.

Further south, it was relatively warm over much of the tropical Pacific and Indian Oceans and southern subtropics. Colder than usual conditions around Antarctica appear to have continued into 2015; here winter sea ice reached record extent and surrounding open-ocean temperatures were below average in 2014.

What is to come? The warmth of the opening months of 2015 has already been sufficient for 12-month means that end in February or March to be higher than for any preceding 12-month mean in the cases of JRA-55, the HadCRUT4 median and MLOST. Only ERA-Interim has peaks in 2005 and 2009/10 that reach above its 12-month mean to March 2015. With the prospect of a strengthening El Niño but with uncertainty for the second half of the year, as also discussed in this Newsletter, it seems quite likely but not certain that global temperatures as recorded in ERA-Interim, too, will reach their highest values yet during 2015.



**Figure 1** Twelve-month running-mean anomalies relative to 1981–2010 in global-mean surface temperature from the ERA-Interim and JRA-55 reanalyses and from the HadCRUT4 and NCDC/MLOST datasets. Values are based on surface air temperature analyses over land and sea ice, and sea-surface temperature analyses elsewhere. They are true global averages in the case of the reanalyses and area-weighted averages over all the 5°x5° grid boxes for which values are provided in the case of HadCRUT4 and MLOST. The median of the 100 HadCRUT4 estimates is plotted in red and blue colours; individual estimates are plotted with semi-transparent grey lines that appear as darker red or blue lines where anomalies are smaller than the median. The dashed line in each panel shows the average of the four analyses for the calendar year 2014, taking the median in the case of HadCRUT4. ERA-Interim SSTs are reduced by 0.1K prior to 2002 to compensate for a shift in values associated with a change in source SST analyses.



**Figure 2** Anomalies in two-metre temperature (K) and sea-ice concentration (%) for January–March 2015 relative to the 1981–2010 mean for the quarter, from ERA-Interim.

# The role of hindcast length in assessing seasonal climate predictability

ANTJE WEISHEIMER

It has recently been argued that single-model seasonal forecast ensembles are overdispersive, implying that the real world is more predictable than indicated by estimates of so-called perfect model predictability, particularly over the North Atlantic. However, such estimates are based on relatively short forecast datasets comprising just 20 years of seasonal predictions.

By analysing a consistent set of seasonal forecast ensembles from eight individual models of the DEMETER and ENSEMBLES projects over various subsets of a longer 42-year period, a recent paper by *Shi et al.* (2015) found that sampling uncertainty due to the length of the hindcast periods is large. It was shown that the skill of forecasting the North Atlantic Oscillation during winter varies within the datasets such that some of the models have high levels of skill for 20-year sub-periods, whereas none of the models have significant skill over all sub-periods (see Table 1 for the three DEMETER models).

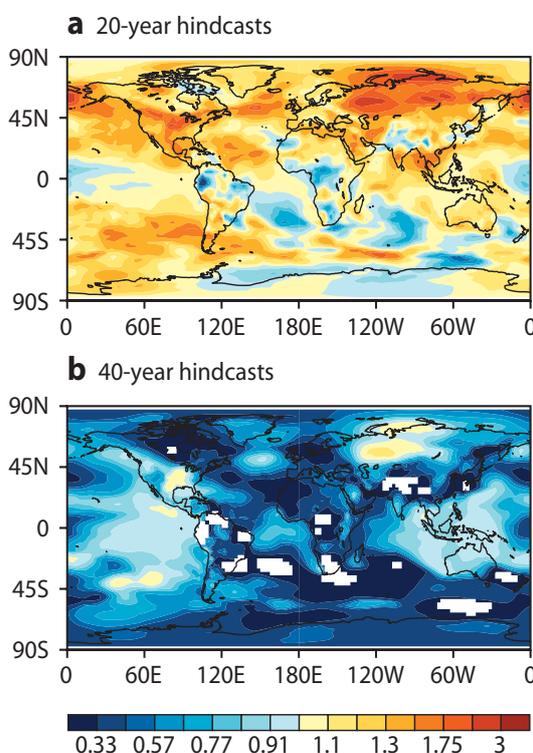
In order to systematically study the effect of the length of the hindcast period on assessments of seasonal predictability, and in particular on the question of whether the forecasts over the North Atlantic are overdispersive and thus underconfident, a large pool of hindcast years was generated by randomly sampling from the huge number of all possible combinations of 5, 10, 15, 20, ... 40 years out of the total 42-year period. For example, there exist 861 possible combinations of randomly sampled 40 years and more than 500 billion combinations of randomly sampled 20 years.

Based on an analysis of the predictable components in the observations and model hindcasts, *Shi et al.* (2015) demonstrate that while individual model ensembles can indeed appear overdispersive over 20-year periods, estimates of seasonal predictability over 40-year periods are more stable and show no

evidence of overdispersion. Instead, the predominant feature on these longer timescales is underdispersion, particularly in the tropics, see Figure 1. The seemingly paradoxical suggestion that the real world is more predictable than indicated by perfect model estimates is thus fully consistent with the sampling uncertainties due to the short hindcast lengths of the seasonal predictions studied. As a consequence, it remains crucially important to develop reliable methods to represent model uncertainty.

Model \ Hindcast period	1960–1979	1980–2001	1960–2001
Météo-France	0.26	<b>0.59</b>	<b>0.38</b>
ECMWF	-0.42	<b>0.45</b>	-0.12
UK Met Office	-0.05	0.21	-0.15

**Table 1** Correlation of the North Atlantic Oscillation index of the ensemble mean with reanalysis data for three DEMETER models for different hindcast periods. Correlations where a t test suggests significance at the 95% level are marked in bold.



**Figure 1** Maximum ratio of predictable component of the observations to predictable component of the model for seasonal forecasts of mean sea level pressure anomalies during boreal winter December to February from retrospective forecasts started on the 1st November each year based on a large sample of a) 20-year hindcasts and b) 40-year hindcasts. A ratio above 1 suggests overdispersion, a ratio below 1 indicates underdispersion. The results shown here are for the Météo-France model used in DEMETER but all other 7 models are broadly similar. For details see *Shi et al.* (2015).

## Further reading

**Shi, W., N. Schaller, D. MacLeod, T. N. Palmer & A. Weisheimer, 2015:** Impact of hindcast length on estimates of seasonal climate predictability, *Geophys. Res. Lett.*, **42**, 1554–1559, doi: 10.1002/2014GL062829.

# Stochastic workshop explores simulation of forecast model uncertainty

**ROBERTO BUIZZA,  
ANTJE WEISHEIMER**

ECMWF hosted a workshop on ‘Stochastic Parametrization’ on 9 and 10 March, which reviewed progress in the simulation of uncertainties in forecasting models and examined the scope for further improvements.

The event was organised by Prof Tim Palmer from Oxford University, and Dr Antje Weisheimer, who works for both ECMWF and Oxford University. It brought together about 40 participants, mainly from ECMWF, Oxford University and the UK Met Office.

Prof Palmer, one of the first three ECMWF Fellows appointed in 2014, said stochastic parametrization was becoming “a more and more important technique, not only for weather prediction, but also for climate prediction”.

*“In order for ECMWF’s forecasts to be useful for quantitative decision-making, it is essential that they come with reliable estimates of forecast uncertainty. The computational models used to make the forecasts are themselves only approximations of reality, and hence are a source of forecast uncertainty. Stochastic parametrization is an important technique to represent model uncertainty quantitatively in an ensemble forecast system,”* he said.

## Stochastic schemes at ECMWF

An ensemble forecast system predicts a set of possible scenarios, taking into account uncertainties arising from our imperfect knowledge of the current state of the atmosphere and from limitations in the forecasting model.

Stochastic schemes are now used routinely in the atmosphere component of ECMWF’s ensemble forecasts and are being developed and tested in the land, ocean and sea-ice components. The schemes simulate model uncertainty by introducing stochastic perturbations into the equations of the forecasting model.

The term ‘stochastic’ means that we cannot precisely determine these perturbations and instead define them in statistical terms. In the ensemble forecast context, the perturbations also have spatial and temporal structures selected to simulate as accurately and reliably as possible the effect of model uncertainty on forecast quality.

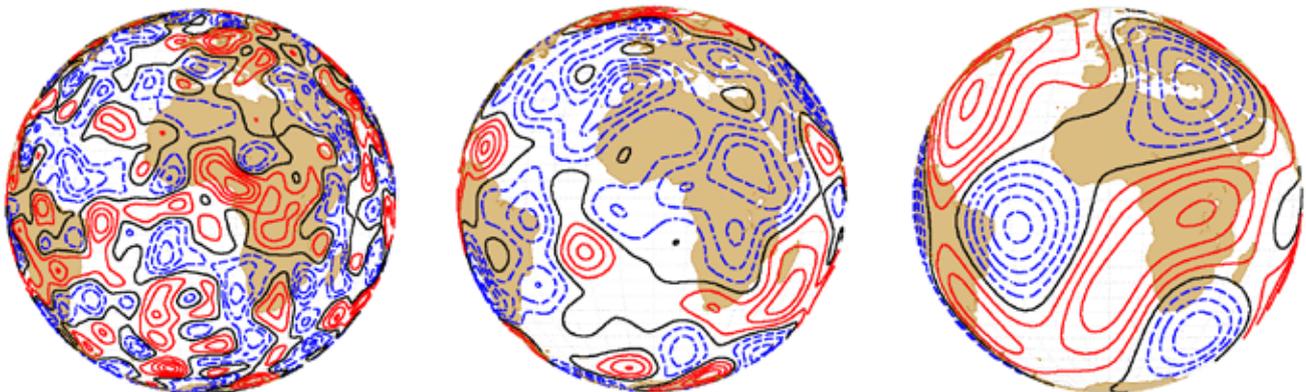
The contours in the figure show examples of perturbation patterns in one of the two schemes used in ECMWF ensembles, the Stochastically Perturbed Parametrization Tendency (SPPT) scheme. SPPT introduces stochastic perturbation at three different spatial and temporal scales: a fast one with a spatial scale of about 500 km and a temporal scale of six hours (left panel), and two slower ones with a larger spatial scale and a slower temporal scale (middle and right panels).

## Search for ‘best approach’

Stochastic parametrization is an area where ECMWF did some pioneering work in 1998–99, after Prof Palmer started to work on the problem in 1996. Today, model uncertainty simulation using stochastic schemes is one of the main areas of collaboration between ECMWF, Prof Palmer and his group at Oxford University, and the Met Office.

This workshop provided an opportunity for scientists working in this area to present ongoing work and to discuss preliminary results and new ideas. It was the third in a series started two years ago to promote collaboration and the exchange of information between the three groups working in this area at Oxford University, ECMWF and the Met Office. The first workshop was held at Oxford University in 2013, followed by a workshop at the Met Office last year.

Many speakers at this third workshop showed how aspects of model behaviour can be improved by using stochastic schemes. They also indicated that the way models react depends on the models themselves. Thus, although it is clear that stochastic schemes improve probabilistic forecasts, it is difficult to draw general conclusions on the relative benefits of the schemes and on which scheme performs best. The next few years should provide more clarity on the approaches that can be followed to further advance the simulation of uncertainties in forecasting models.



*Examples of perturbation patterns* at three spatial scales used in ECMWF medium-range/monthly (ENS) and seasonal (S4) ensembles.

## Piotr Smolarkiewicz granted Poland's top academic title

JOANNE JEPPESEN

Dr Piotr Smolarkiewicz, a member of ECMWF's Research Department, has been made Professor of Physical Sciences by the President of Poland. Piotr received the country's highest academic title at a ceremony held at the Presidential Palace in Warsaw on 18 February.

The University of Warsaw, where Piotr studied physics and obtained his PhD, nominated him for the title. A two-year review process involving a panel of referees culminated in a recommendation to President Bronisław Komorowski. "I feel honoured to have received this title. I regard it as an expression of appreciation from my colleagues at the University of Warsaw," Piotr said.

After studying physics, Piotr took up a post-doctoral fellowship at the National Center for Atmospheric Research

(NCAR), Boulder, USA, where he remained for over 30 years. During that time his work brought him into contact with colleagues at ECMWF, where he spent a sabbatical year in 1999/2000 and became a regular visitor. He joined ECMWF in March 2013 as the principal investigator of the PantaRhei project. Funded as an Advanced Grant by the European Research Council, PantaRhei is developing an interdisciplinary forecasting system capable of simulating multi-scale fluid flows.



**Piotr Smolarkiewicz** joined ECMWF in March 2013



**Polish President Bronisław Komorowski**

handed over the certificate during a ceremony in Warsaw.

Photo credit: Eliza Radzikowska-Białobrzewska/  
Chancellery of the President of the Republic of Poland

## Annual Seminar proceedings published

STEPHEN ENGLISH

The presentations given at ECMWF's Annual Seminar 2014 on The Use of Satellite Observations in NWP have been published on the web. The seminar, attended by over 100 scientists from around the world, was held from 8 to 12 September at ECMWF's headquarters in Reading, UK. It is an important part of ECMWF's educational activities.

Presentations covered satellite data assimilation in a range of areas: weather, atmospheric chemistry,

marine, land surface and climate monitoring. Relatively new and forthcoming techniques, such as spaceborne lidar and radar, were presented alongside an evaluation of the state of the art for well-established observations. The seminar confirmed that hyperspectral infrared radiances and radio occultation bending angles are now a core part of the global observing system, alongside conventional observations, microwave radiances and scatterometer and atmospheric motion vector wind observations.

The seminar concluded with an examination of techniques to estimate the overall value of the satellite global observing system and its individual components, alongside a gap analysis, with a view to ensuring that the total global observing system meets requirements. Although satellite observations are now very important, the continued importance of conventional observations was also emphasised.

All presentations are available at: <http://old.ecmwf.int/publications/library/do/references/list/3032015>

# The skill of ECMWF cloudiness forecasts

THOMAS HAIDEN, RICHARD FORBES,  
MAIKE AHLGRIMM, ALESSIO BOZZO

Correctly predicting cloudiness is an important part of a successful weather forecast. Cloud cover is not just of interest in its own right but also has a major impact on other parameters, such as temperature and solar radiation. It is, however, often highly variable in terms of time and location and can therefore be difficult to forecast. The skill of ECMWF cloudiness forecasts improves if time-averaged values and ensemble forecasts are used, but some regions of the world pose particular challenges.

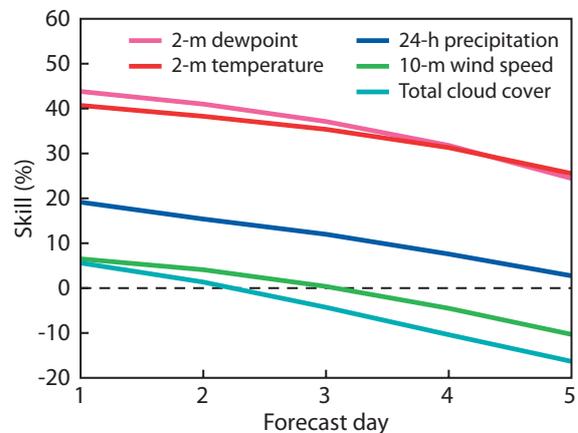
We investigate cloud forecast skill in the ECMWF model by verifying total cloud cover and solar radiation forecasts against surface observations and satellite data, by analysing the scale-dependence of skill, and by evaluating both high-resolution and ensemble forecasts of cloudiness and solar radiation.

## Skill in predicting cloudiness

The difference in downward solar radiation at the surface between an overcast and a clear sky can reach several  $100 \text{ Wm}^{-2}$  during daytime and thereby strongly impact the weather experienced at the surface as well as the development of the atmospheric boundary layer. At night, the reduced net outgoing longwave radiation in the presence of clouds, especially of low clouds, may alter the structure of the stable boundary layer and strongly affect minimum temperatures. Hence the skill in predicting cloudiness has implications for the skill in predicting other parameters, too.

However, cloudiness is one of the more difficult parameters to forecast. This is illustrated in Figure 1, which shows the skill relative to climatology of the high-resolution forecast (HRES) in ECMWF's Integrated Forecasting System (IFS) in predicting various surface parameters in Europe. As a verification measure, we use the standard deviation of the forecast error, which is independent of any constant biases in either the forecast or the verifying observations. For 2-metre temperature and 24-hour precipitation there is positive skill beyond forecast day 5, whereas for 10-metre wind speed and total cloud cover it becomes negative as early as on days 3 and 2, respectively. If skill is measured relative to persistence, as was done by Köhler (2005) in a similar diagram, then the ranking of precipitation and temperature is reversed but total cloud cover still ranks lowest.

Deterministic forecast skill for total cloud cover lags behind other parameters and has improved relatively little over the last ten years, despite some recent progress (Haiden et al., 2015). One explanation could be that, unlike 24-hour precipitation, total cloud cover observed at surface stations (SYNOP observations) is a near-instantaneous quantity. A substantial part of the diurnal variability in cloud cover is due to sub-grid scale variations in vertical motion and humidity fields, which have low predictability.



**Figure 1** Forecast skill relative to climatology of various surface parameters for Europe in 2014. Skill is defined as the standard deviation of the forecast error divided by the standard deviation of the climatology error (average over the period 1980–2009), where the error is calculated using SYNOP observations. This verification measure is subtracted from 1 and multiplied by 100 to give a skill score that is 100% for a perfect forecast and 0% when equivalent to climatology.

## Time and space scale-dependence of cloud forecast skill

### Total cloud cover

Cyclonic activity and convective processes create cloud systems with a high degree of spatial and temporal variability. Thus part of the low skill of cloudiness forecasts verified at a certain time at a given location may be due to the limited representativeness of the observations. If we verify not just instantaneous values of total cloud cover against SYNOP observations (at 12 UTC, say) but also averages over several reporting times (e.g. 06, 12, 18 UTC), the loss of predictability associated with intra-diurnal variations in cloud cover can be estimated.

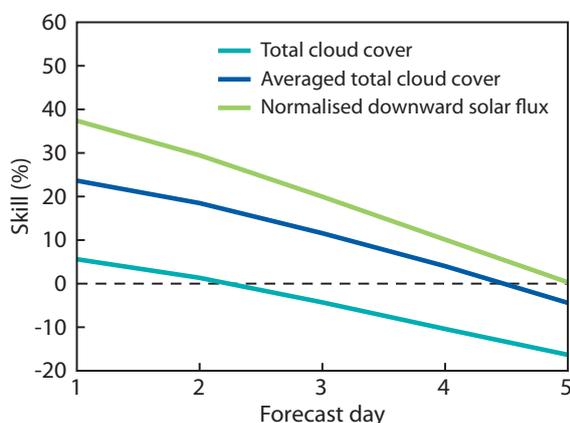
There is also a spatial representativeness mismatch between forecasts and SYNOP observations of total cloud cover, whether they are made visually or by an instrument. The area covered by visual observation typically varies between 10 and 100 km around a station, depending on visibility and topography (Mittermaier, 2012). Automated observations are derived from ceilometers, which measure cloud cover overhead. Depending on the wind speed in the cloud layer, the scanned area may or may not be representative of the model grid-scale.

Verifying cloud cover averaged over a certain time period reduces this spatial mismatch to a certain extent, depending on how strongly local cloud cover changes are governed by advection. It does not work in situations such as convective clouds over mountainous terrain in the presence of weak flow, where quasi-stationary sub-grid scale variations may persist during the day.

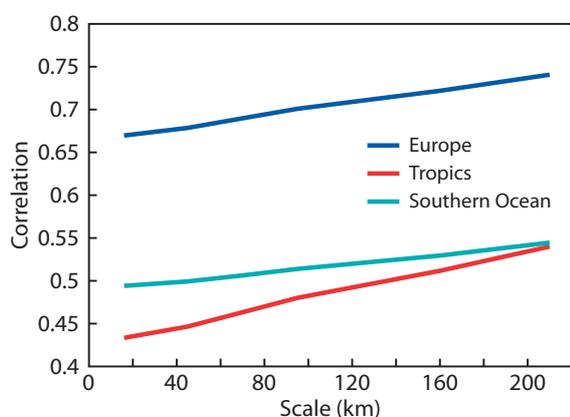
A comparison of the skill, relative to climatology, in predicting instantaneous and averaged cloudiness in Europe is given in Figure 2. The difference in performance for the two quantities is equivalent to about two forecast days and is only weakly dependent on lead time. Comparison with Figure 1 shows that forecasts of averaged cloudiness have a skill comparable to that of 24-hour precipitation.

### Solar radiation

Daily averages of downward solar radiation can be used as a proxy for average cloud cover. To assess the skill of such radiation forecasts, we use daily averages of downward solar radiation at the surface provided by the Climate Monitoring Satellite Application Facility (CM SAF). The data is based on observations from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard the geostationary Meteosat satellites and has been verified extensively against Baseline Surface Radiation Network (BSRN) data.



**Figure 2** Forecast skill for Europe of instantaneous (00, 12 UTC) and averaged [(06+12+18 UTC)/3] total cloud cover in 2014 from verification against SYNOP, as well as forecasts of daily averages of the normalized downward solar flux at the surface, verified against satellite-derived values.



**Figure 3** Correlation between forecasts and satellite-derived observations of daily averages of the normalized downward solar flux at the surface on forecast day 3 in 2014 as a function of horizontal scale for Europe (35°–70°N, 12°W–42°E), tropics (20°N–20°S) and Southern Ocean (50°–70°S).

Since the main interest here is to use the radiation totals as a proxy for cloudiness, the amount of incoming solar radiation at the surface is normalized by comparing it with typical clear-sky values for a given latitude and season. Day-to-day variations in the normalized values are primarily due to variations in cloudiness, but in some desert areas they also represent the effect of varying dust aerosol content.

Figure 2 shows how the skill of the daily averaged solar radiation forecast compares to that of the averaged cloud cover forecast. Note that the radiation verification has been made for all grid points within the European domain using the daily mean derived from satellite observations every 15 minutes, whereas the SYNOP cloudiness verification is only for about 700 station locations averaged over just 3 validity times. The resulting medium-range skill is slightly higher than for the averaged cloud skill, but both skill measures show the forecast outperforming climatology up to day 4.

The dependence of forecast skill on spatial scale is analysed by averaging the forecast and satellite-derived radiation fields from the original grid at 16 km resolution onto increasingly coarser grids. As a measure of correspondence between time-series of forecasts and satellite-derived values at each grid point, we use the correlation coefficient. Figure 3 shows that, in Europe, the correlation increases by about 0.07 when the scale increases from 20 to 200 km. In the tropics, where the correlation is much lower to begin with, the increase is larger, about 0.11. Over the Southern Ocean, the level of correlation is comparable to the tropics, but with an increase of only 0.05. This suggests that the low radiation forecast skill over the Southern Ocean is more strongly associated with processes on scales resolved by the model.

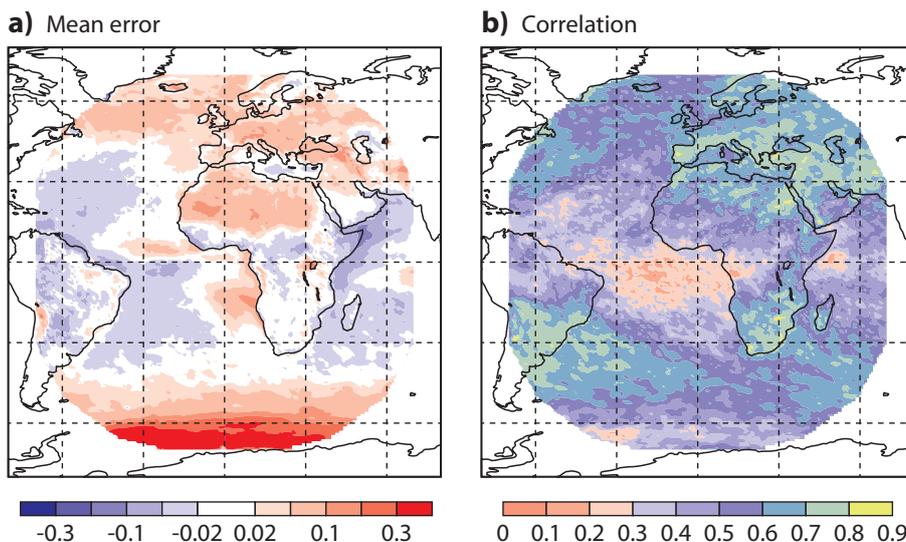
### Geographical variations in skill

Overall, the bias in ECMWF's solar radiation forecast is relatively small compared to the magnitude of non-systematic errors, even at short lead times. However, there are certain regions where a specific cloud type dominates. Here, systematic model deficiencies have a more marked effect. The region off the south-western African coast, where the forecast underestimates cloud forcing (Figure 4a), is dominated by marine stratocumulus. Forecasts predict too much solar radiation at the surface over the Southern Ocean, as well as in the mid-latitudes of the northern hemisphere. In terms of absolute values, these correspond to biases of 10–20 Wm<sup>-2</sup>. There is also a positive bias over North Africa and a small negative bias over the western part of the Atlantic and Indian Oceans.

Note that an overestimation of solar radiation at the surface can be due to a negative bias in cloud fraction, an underestimation of cloud optical thickness, or other reasons, such as an underestimate of aerosol. Some of these systematic errors are discussed in more detail below.

### The Southern Ocean and supercooled liquid water

The high latitudes of the Southern Ocean are an area where overestimation of solar radiation at the surface is a common problem in both numerical weather prediction and climate models. There may be a number of reasons for



**Figure 4** Verification of normalized downward solar radiation at the surface on forecast day 3 in terms of (a) the mean error and (b) correlation of daily averages for the year 2014.

systematic radiation errors in this region, but *Forbes & Ahlgrimm* (2014) highlight the large areas of stratiform low-level cloud that are prevalent over the Southern Ocean as a cloud regime where cloud forcing can be underestimated. Clouds found in this region are often mixed-phase, containing both liquid and ice at sub-zero temperatures, and the correct cloud phase as well as cloud cover and amount of condensate is required to represent the correct impact on radiation.

In 2010, the cloud scheme in the IFS was changed from a diagnostic temperature-dependent liquid/ice phase split to include separate prognostic variables for cloud liquid and ice and associated sources and sinks. This enabled the model to better represent the vertical structure and variable proportions of liquid and ice found in these stratiform clouds (*Forbes & Ahlgrimm*, 2014). This was not only important for the Southern Ocean but also reduced radiation errors for winter stratiform boundary layer cloud over Northern Hemisphere land, with a direct impact on 2-metre temperatures. A significant contribution to the remaining radiation errors over the Southern Ocean and over the North Atlantic mid-to-high latitudes in Figure 4 is due to the lack of supercooled liquid water in convective clouds, which will be addressed in a future IFS cycle.

**Boundary layer cloud and cloud liquid water**

The cloud types dominating the subtropical oceanic regions where surface downward solar radiation is overestimated are stratocumulus (including transitions to cumulus) and stratus. Marine stratocumulus are typically only a few hundred metres thick, and cloud-top radiative cooling plays an important part in their generation and maintenance. The boundary layer associated with stratocumulus is usually well-mixed and topped by an inversion. Thus there are similarities with continental stratus in central Europe, suggesting that model changes

aimed at improving the prediction of the former may also improve forecasts of the latter (*Haiden & Trentmann*, 2015).

Work is under way to better integrate the boundary layer, shallow convection and cloud schemes with the aim of improving the stratocumulus-to-cumulus transition and stratocumulus cloud cover and radiative forcing over land and ocean, as well as improving the overly strong radiative forcing in the trade cumulus regions in the western half of the ocean basins.

As mentioned above, model biases in cloud radiative forcing may be due to errors in cloud fraction or optical properties, which for warm-phase cloud are dependent on the assumed effective radius of cloud particles and the liquid water path. Precipitation processes (formation and evaporation) modify the amount of liquid condensate in the cloud as well as the humidity in the sub-cloud layer, which in turn can affect the cloud cover. The next IFS cycle (41r1), to be operational this spring, includes changes to the parametrization of rain formation (autoconversion and accretion) and evaporation (*Ahlgrimm & Forbes*, 2014). This increases the liquid water path, particularly for stratiform cloud over land, which will help to reduce the bias in shortwave cloud forcing.

**Africa and dust aerosol**

A large negative bias can be seen over the Arabian Sea, and this is more prominent during the summer months. It is linked to the monthly mean climatological aerosol distribution used in the operational model. The distribution of mineral dust in particular dominates the total aerosol optical depth (AOD) in the region, and the negative bias in the solar radiation at the surface suggests an overestimation of the dust AOD. Also related to the climatological dust AOD is the positive bias over the Sahara, although part of it appears to be the result of an underestimation of the surface albedo in the CM SAF radiation product in the area. Tests with a new aerosol climatology based on prognostic aerosol

species implemented in the MACC (Monitoring Atmospheric Composition and Climate) system at ECMWF show that the solar radiation at the surface will benefit from an improved AOD distribution and revised optical properties of the aerosol species.

**Variation in skill**

The skill in forecasting total cloud cover decreases substantially from the mid-latitudes towards the subtropics and tropics. At lower latitudes, a larger portion of the vertical motion field is due to convective processes, which have lower predictability. Figure 4b shows the temporal correlation between forecasts and satellite-derived values. It measures the ability of the model to forecast day-to-day variations of total cloud cover.

In Europe, correlation values of around 0.8 are found, implying that the forecast is able to capture about two thirds of the day-to-day variability. Within Europe, the highest skill is found in Mediterranean land areas. This is due to the less frequent occurrence of low stratus in these areas compared with central and northern Europe. As analysed by *Haiden & Trentmann* (2015), low stratus is underestimated in large parts of continental Europe. It is one of the primary causes of large 2-metre temperature forecast errors in the short to medium range during autumn and winter in Europe.

Compared to the mid-latitudes, skill is rather low in tropical and subtropical areas, with correlations dropping below 0.3, which means that less than 10% of the day-to-day variability is captured by the model. The large area of low skill in the southern tropical Atlantic extends from the Intertropical Convergence Zone (ITCZ) of frequent deep convection into areas more dominated by cumulus and stratocumulus.

**Use of ensemble forecasts of cloudiness**

We investigate the usefulness of ensemble forecasts (ENS) of cloudiness by first verifying the mean and the median of the ensemble members' total cloud cover forecast against SYNOP observations. We then verify the mean and median of the ensemble members' radiation forecast against satellite data. The error-spread relationship is also considered.

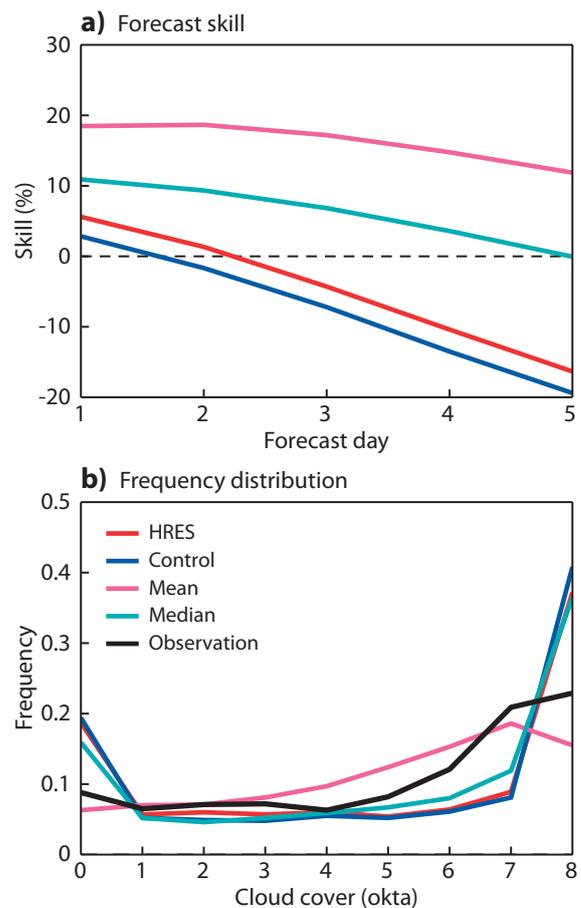
**Verification against SYNOP observations**

Figure 5a compares the skill of instantaneous total cloud cover relative to climatology of the HRES, the ensemble control (CTRL), ensemble mean (MEAN), and ensemble median (MEDIAN). While the CTRL performs slightly worse than the HRES, both MEDIAN and MEAN show considerably higher skill, and a smaller loss of skill with lead time. The fact that MEDIAN and MEAN outperform the single forecasts by such a large margin, particularly in the short range, is perhaps surprising and we need to ask to what extent it is due to the fact that cloud cover is a parameter that is physically limited to a finite interval, causing error measures to favour forecasts that avoid extremes.

Figure 5b compares the frequency distribution of forecasts

at day 3 with observations. In both the HRES and the CTRL there is a much higher occurrence of clear-sky and overcast conditions than in the observations, while in the range from 1–7 okta the frequency is underestimated. The MEDIAN behaves similarly but its distribution is slightly smoothed, which brings it closer to the observations.

The MEAN is much smoother still, with the typical U-shape of the distribution almost lost by the averaging over ensemble members. Although the MEAN underestimates the occurrence of 0 and 8 oktas, it is closest to the observed distribution overall and therefore has highest skill. However, it is important to remember that there is a spatial representativeness mismatch between the modelled and observed cloud cover and uncertainty in the comparison. For example, in SYNOP observations of cloud cover a value of 0 okta is given only if the sky is completely cloudless, and 8 oktas are given only if there is no gap visible through the cloud layer(s). To take account of these uncertainties, if the frequencies in Figure 5b for 7 and 8 okta are taken together, then HRES, CTRL and MEDIAN all match the observed combined frequency for the high cloud cover regime quite well, and better



**Figure 5** (a) Forecast skill of total cloud cover (00,12 UTC) in 2014 for Europe and (b) frequency distribution for forecast day 3 from the high-resolution forecast (HRES), the ensemble control (CTRL), ensemble mean (MEAN) and ensemble median (MEDIAN) and for observed (SYNOP) cloud cover (Observation).

than MEAN. This is not the case though for the combined frequency of 0 and 1 okta, which is too high in all forecasts except the MEAN.

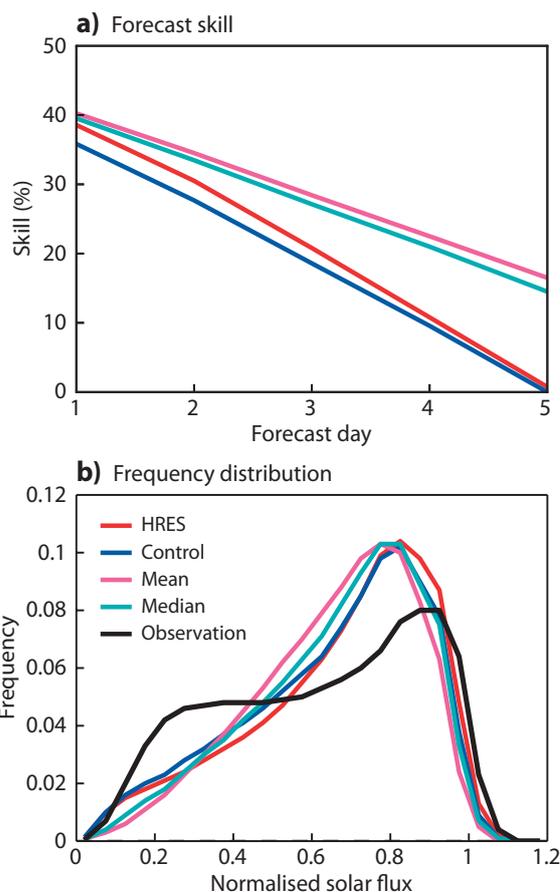
**Verification against satellite data**

To see whether the above results carry over to daily averaged radiation forecasts, we perform a similar analysis based on satellite-derived normalized downward solar radiation. Figure 6a shows that the skill of the HRES, CTRL, MEAN and MEDIAN follows the same ranking as in the verification against SYNOP (Figure 5a). However, the benefit of using ensemble information is smaller, especially at short lead times.

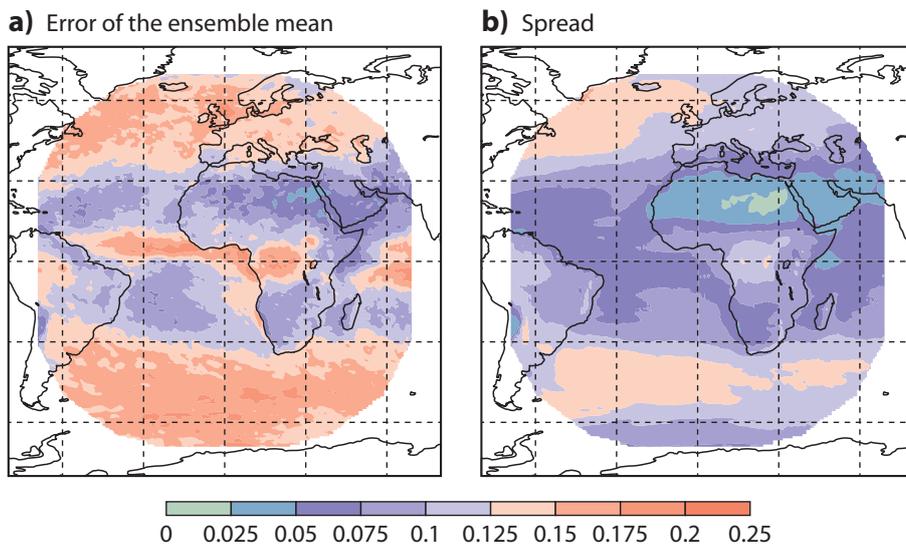
The frequency distribution of the normalized radiation values (Figure 6b) is rather different from that of instantaneous cloud cover. In the satellite-derived data it peaks near 0.9 (corresponding to 90% of the typical clear-sky flux at the given latitude and season) and has an almost bimodal shape with a plateau near 0.4. The forecast distributions, which are relatively similar to each other, put the peak closer to 0.8 and they appear to underestimate the frequency of cases in the part of the distribution that represents cloudy to overcast conditions. This is consistent with the results from an evaluation of radiation bias for different cloudiness regimes over central North America (Ahlgrimm & Forbes, 2012), where the IFS underestimates the cloud forcing of more overcast cloud (corresponding to normalised solar fluxes of 0.1 to 0.4) due to too little liquid water path and overestimates the cloud forcing in the broken cloud cover regime (corresponding to normalised solar fluxes closer to 1).

**Error-spread relationship**

An important property of ensemble forecasts is the agreement between error and spread. While this has been



**Figure 6** (a) Forecast skill of the normalised downward solar flux at the surface in 2014 for Europe and (b) frequency distribution for forecast day 3 from the high-resolution forecast (HRES), the ensemble control (CTRL), ensemble mean (MEAN) and ensemble median (MEDIAN) and for satellite-derived fluxes (Observation).



**Figure 7** (a) Error of the ensemble mean (i.e. standard deviation of the ensemble mean), and (b) the spread (i.e. ensemble standard deviation around the mean) for daily averages of the normalised downward solar radiation at the surface in 2014 on forecast day 3.

monitored for upper-air fields and surface parameters for some time, the solar radiation data makes it possible to obtain a more complete picture with respect to cloudiness. Figure 7a shows the standard deviation of the error of the ensemble mean. There are broad areas of increased error due to cyclonic activity in the mid-latitudes, and due to deep convection associated with the ITCZ. The geographical distribution of ensemble spread (Figure 7b) captures the increase of error from the sub-tropics to the mid-latitudes reasonably well. In the tropics, a comparable match between error and spread is found only over land, whereas over the tropical Atlantic and Indian Ocean the ensemble appears to be strongly underdispersive.

### Summary and outlook

While positive forecast skill for instantaneous total cloud cover in the mid-latitudes extends only out to days 2–3, forecasts of daily averaged quantities, such as daily means of solar radiation at the surface, are found to be skilful out to day 5. By using ensemble-derived quantities such as ensemble mean or median total cloud cover, the skilful range can be extended further by 1–2 days. The decrease of forecast skill from the mid-latitudes towards the sub-tropics and tropics is considerable. For a forecast at day 3, the correlation between forecast and satellite-derived downward solar radiation decreases from 0.6–0.8 to 0.2–0.4. Similarly, in the mid-latitudes ensemble error and spread match reasonably well, whereas the ensemble forecast is more strongly underdispersive in the area of the ITCZ. Cloud systems generated by deep convection are inherently less predictable than those associated with extra-tropical

cyclones, and the ensemble forecast is not yet able to fully capture the uncertainty in initial conditions and model physics for such systems.

Starting with the introduction of the 5-species cloud microphysics scheme in 2010, changes to microphysical parametrization have led to an increase in cloud forecast skill in recent years (Haiden *et al.*, 2015). At the time of writing, the new model cycle 41r1 has just been evaluated in the operational parallel suite, showing a reduction in total cloud cover root-mean-square error (RMSE) of about 1–2% compared to the operational forecast. With regard to evaluation, there are plans to include satellite-derived bulk cloud properties, such as liquid water path and cloud top temperature, in order to stratify cloud forecast errors and facilitate their attribution to specific cloud processes.

A number of further improvements to cloudiness forecasts, and reduced systematic errors in radiation forecasts, are on the horizon. In particular, work is under way on improved integration of the boundary layer, shallow convection and cloud schemes to improve the stratocumulus and stratocumulus-to-cumulus transition, and on an improved representation of turbulence driven by cloud top cooling for stable boundary layers, which will help to reduce deficiencies in night-time continental stratus. A better representation of mixed-phase cloud, particularly for convection in cold-air outbreaks, will reduce the high-latitude, oceanic, solar radiation bias, and an improved aerosol climatology will reduce errors over North Africa and the Arabian Sea.

---

### FURTHER READING

**Ahlgrimm, M. & R. Forbes**, 2012: The impact of low clouds on surface shortwave radiation in the ECMWF model. *Mon. Wea. Rev.*, **140**, 3783–3794.

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

**Forbes, R.M. & M. Ahlgrimm**, 2014: On the representation of high-latitude boundary layer mixed-phase cloud in the ECMWF global model. *Mon. Wea. Rev.*, **142**, 3425–3445.

**Haiden, T., M. Janousek & D. Richardson**, 2015: Forecast performance 2014. *ECMWF Newsletter No. 142*, **4**.

**Haiden, T. & J. Trentmann**, 2015: Verification of cloudiness and radiation forecasts in the greater Alpine region. *Meteorol. Z.*, **24**, (in press).

**Köhler, M.**, 2005: Improved prediction of boundary-layer clouds. *ECMWF Newsletter No. 104*, 18–22.

**Mittermaier, M.**, 2012: A critical assessment of surface cloud observations and their use for verifying cloud forecasts. *Q. J. R. Meteorol. Soc.*, **138**, 1794–1807.

# Atmospheric composition in ECMWF's Integrated Forecasting System

VINCENT-HENRI PEUCH, RICHARD ENGELEN,  
JOHANNES FLEMMING,  
VINCENT HUIJNEN (KNMI), ANTJE INNESS

Weather conditions can have a strong influence on atmospheric composition, and composition can in turn have a significant impact on weather-related atmospheric processes, such as radiation and microphysics. Integrating atmospheric composition into numerical weather prediction models can thus not only help to produce reliable forecasts of composition but can also improve the skill of weather forecasts. The benefits of considering the interaction between composition and meteorology have been documented in several individual cases (*Baklanov et al., 2014*).

As part of the EU-funded Monitoring Atmospheric Composition and Climate research and development projects (MACC), ECMWF has devoted significant efforts to integrating a detailed representation of atmospheric composition and associated processes into its Integrated Forecasting System (IFS). The resulting system, C-IFS (Composition-IFS), is presented in this article. The system has been extensively validated both in model-only and data assimilation configurations and replaced the previous system in September 2014, in anticipation of transitioning to the operational phase of the Copernicus Atmosphere Monitoring Service (CAMS). C-IFS is effectively one of the first examples of a new class of global forecasting systems capable of delivering combined weather and environmental services.

Just like the IFS, C-IFS will undergo continuous development, including a resolution upgrade expected for later this year.

## From IFS-MOZART to C-IFS

Modules for the simulation of atmospheric chemistry have been implemented in the IFS since 2009 as part of the MACC series of projects. Forecasts and data assimilation of reactive gases were initially produced with a coupled system, comprising the IFS and a chemical transport model, MOZART, which simulates atmospheric chemistry. IFS-MOZART was used for MACC reanalysis and for daily near-real-time forecasts from May 2007 to September 2014. But the coupled approach had limitations, such as the need for interpolation between IFS and MOZART model grids and the duplicate simulation of transport processes.

The drive towards closer integration led to the development of a new model configuration, called Composition-IFS (C-IFS). In C-IFS, the chemistry scheme complements the previously integrated modules for aerosol and greenhouse gases. C-IFS is computationally much more efficient than IFS-MOZART because the chemistry scheme is fully

integrated into the IFS. The first version of C-IFS also uses a different chemical mechanism (CB05), which is simpler and only accounts for tropospheric chemical processes.

C-IFS is expected to provide reanalyses, analyses and forecasts at the global scale, covering atmospheric composition comprehensively (aerosols, reactive gases and greenhouse gases). These datasets constitute the 'global' service line of CAMS. In the following, we present a brief overview of evaluation results for certain key reactive gases. More detailed validation results can be found on the MACC/CAMS website in the form of quarterly reports ([www.copernicus-atmosphere.eu](http://www.copernicus-atmosphere.eu)). For more details on C-IFS, see Box A.

## Evaluation of C-IFS

An evaluation exercise covering the year 2008 has been carried out to compare C-IFS and IFS-MOZART regarding their skill in representing tropospheric composition (*Flemming et al., 2015*). The MACC reanalysis (*Inness et al., 2013*), based on IFS-MOZART but including the assimilation of chemical satellite retrievals of CO, O<sub>3</sub> and NO<sub>2</sub>, has been used as a benchmark. The evaluation showed that C-IFS performed as well as or better than IFS-MOZART and broadly as well as the MACC reanalysis against independent observations.

Compared to IFS-MOZART, C-IFS had smaller biases for:

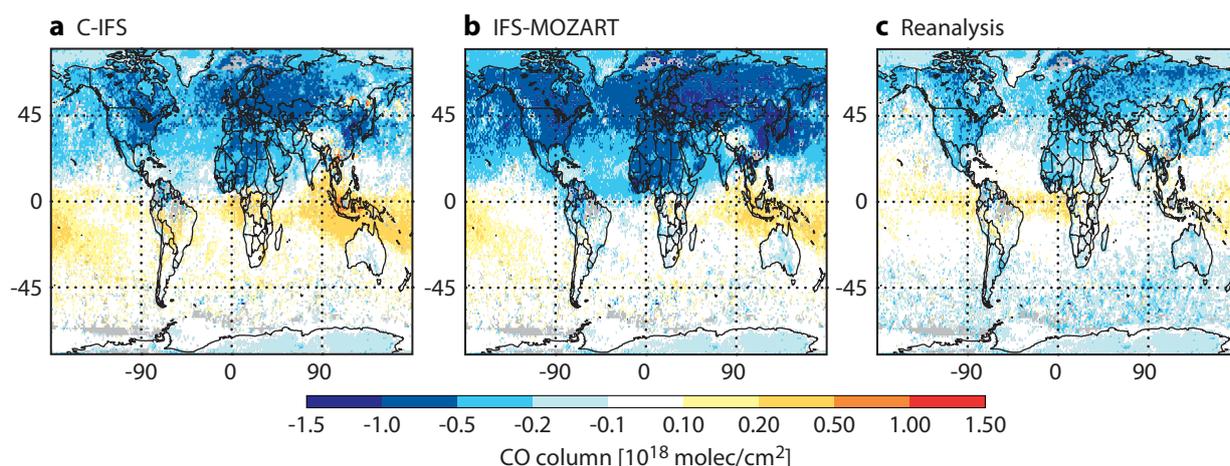
- CO in the northern hemisphere,
- O<sub>3</sub> in the upper troposphere, and
- winter-time SO<sub>2</sub> at the surface in Europe.

## The main features of C-IFS

A

C-IFS applies the CB05 Carbon Bond chemical mechanism of the chemical transport model TM5. The implementation of CB05 was carried out in close co-operation with the Dutch meteorological service, KNMI. CB05 describes tropospheric chemistry with 55 species and 126 reactions.

C-IFS benefits from the detailed cloud and precipitation physics of the IFS for the calculation of wet deposition and NO emissions from lightning. Wet deposition modelling accounts for the sub-grid scale distribution of clouds and precipitation. Dry deposition is modelled using pre-calculated monthly-mean dry deposition velocities with a superimposed diurnal cycle. Surface emissions and dry deposition fluxes are applied as surface boundary conditions in the diffusion scheme. Lightning emissions of NO can be calculated either by cloud height or by convective precipitation. The latter parametrization is used in MACC applications. Anthropogenic emissions are taken from the MACCcity inventory and biomass burning emissions from the Global Fire Assimilation System (GFAS).



**Figure 1** Bias of carbon monoxide (CO) total column with respect to retrieval of MOPITT V6 satellite data for April 2008 of (a) C-IFS, (b) IFS-MOZART and (c) Reanalysis.

As an illustration, Figure 1 shows the bias against MOPITT satellite instrument data for total column CO for April 2008, the month with the highest CO concentrations, of C-IFS, IFS-MOZART and MACC reanalysis. (See Box B for a list of satellite instruments referred to in this article.) Interestingly, the diurnal cycle of surface  $O_3$  also showed greater realism in the C-IFS simulation against independent, spatially representative European air quality observations (not shown). As all configurations used the same emission data, such improvements can partly be explained by differences in the chemical mechanism and the simulation of wet and dry deposition. Further, improvements in  $SO_2$  and the diurnal cycle of  $O_3$  found in C-IFS are most probably caused by the more consistent interplay of diffusion and sink and source processes in the integrated C-IFS compared to the coupled IFS-MOZART system, which by design cannot achieve perfect consistency.

There is still room for improvement of C-IFS. It underestimated surface  $O_3$  over Europe and North America in the spring of 2008 and overestimated it in late summer and autumn. CO was underestimated by C-IFS, in particular in Europe and North America throughout the year and especially in spring and winter, and in the biomass burning season in Southern Africa. Winter-time tropospheric  $NO_2$  over China as retrieved from the GOME-2 instrument was twice as high as the fields modelled by C-IFS, IFS-MOZART and the MACC reanalysis.

### Extending the model

Although only one chemical mechanism is used for operational forecasts, C-IFS can apply multiple chemistry schemes. The implementation of the chemistry schemes of the MOCAGE and MOZART chemical transport models (CTMs) has technically been completed, but further optimisation and evaluation is required. Both schemes offer a description of stratospheric chemistry, which is not included in the scheme CB05 currently used in C-IFS. In parallel, work is under way to combine the CB05 mechanism with the stratospheric chemistry mechanism

of the BASCOE CTM. In the end, C-IFS is expected to offer several options to represent chemical processes in the troposphere and the stratosphere.

There are plans to further improve the link between the physics and chemistry packages in the IFS. For example, detailed information from the IFS surface scheme will be used for the calculation of dry deposition and biogenic emissions during the simulation. A first important step will be to replace the pre-calculated climatological dry deposition velocities with values calculated based on actual weather and surface parameters in the IFS. Once the stratospheric chemistry is fully implemented, the impact of the simulated  $O_3$  fields on the IFS radiation scheme and the corresponding feedback on the temperature fields will be investigated. Another ongoing development is to create closer links between the C-IFS greenhouse gas, aerosol and gas-phase chemistry modules.

### Data assimilation

C-IFS has been used since September 2014 to provide daily analyses and 5-day forecasts for MACC and CAMS. To improve the quality of the atmospheric composition forecasts, several of the chemistry variables ( $O_3$ , CO,  $NO_2$ ,  $SO_2$ , HCHO) have been included as control variables in the data assimilation part of the IFS. This means that the initial conditions for these fields can be modified by assimilating observations of atmospheric composition.

In this article, we discuss the assimilation of CO,  $O_3$  and  $NO_2$  retrievals from various sensors. This is described here by comparing  $O_3$ , CO, and  $NO_2$  fields from an assimilation run for the year 2008 (CIFS-AN) with fields from a control run without data assimilation (CIFS-CTRL) and against independent observations. More details about the setup of these C-IFS data assimilation experiments and the evaluation results can be found in *Inness et al. (2015)*, where we also show that C-IFS in data assimilation mode performs similarly well or better than the coupled system used in the MACC reanalysis for CO,  $O_3$  and  $NO_2$ , especially in the lower troposphere and at the surface.

MOPITT data for total column CO are assimilated in CIFS-AN. This leads to an improved total column CO field compared to CIFS-CTRL, and also to some improvements in the vertical distribution of CO and CO concentrations in the lower troposphere. The largest impact is found in the northern hemisphere winter, when the CO lifetime is longest (Figure 2a). In the tropics, there is also some improvement in CIFS-AN, compared to CIFS-CTRL, in surface and lower tropospheric CO, particularly during the South African biomass burning season (Figure 2b).

The simple stratospheric ozone photochemical parametrization used by the standalone C-IFS (CB05) system to model the stratospheric ozone field is designed to be used in a data assimilation context and leads to a biased stratospheric and total column ozone field in CIFS-CTRL. In CIFS-AN, a combination of ozone total column and stratospheric profile retrievals are assimilated, which greatly improves the total column, the stratospheric ozone field and the upper tropospheric ozone field compared to CIFS-CTRL.

Because no tropospheric O<sub>3</sub> data as such are assimilated, the differences between CIFS-CTRL and CIFS-AN in tropospheric O<sub>3</sub> come from the residual of total column O<sub>3</sub> and the stratospheric profile data and are smaller in the mid- and lower troposphere than in the upper troposphere, as characteristics of the chemistry scheme become more important. For example, a large positive bias in lower tropospheric ozone over East Asia is not reduced by the analysis, and there is little difference in lower tropospheric ozone over Europe and North America during the summer (Figure 3). Nevertheless, there is some improvement in CIFS-AN in the troposphere compared to CIFS-CTRL: the positive ozone bias seen in CIFS-CTRL over Europe and North America during winter and spring in the lower troposphere is reduced.

Testing is under way for the assimilation of O<sub>3</sub> profile retrievals from the GOME-2 and IASI instruments. This will hopefully provide a better constraint on both the stratospheric ozone profiles and the tropospheric ozone burden.

Tropospheric NO<sub>2</sub> column data from the OMI instrument

**Satellite instruments mentioned**

**B**

MOPITT (Measurements of Pollution in The Troposphere) – an instrument carried on the Terra satellite of NASA’s Earth Observing System, measuring tropospheric carbon monoxide.

GOME-2 (Global Ozone Monitoring Experiment-2) and IASI (Infrared Atmospheric Sounding Interferometer) – two instruments carried on the MetOp meteorological satellite operated by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT).

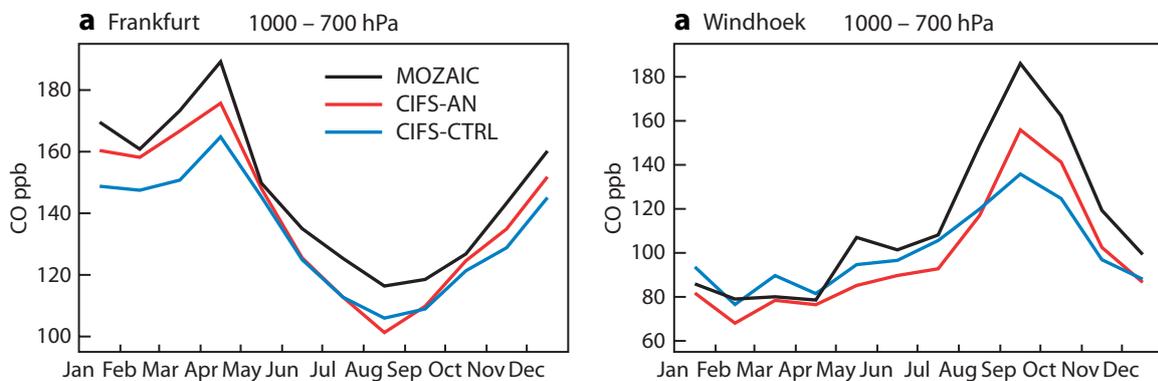
OMI (Ozone Monitoring Instrument) – instrument carried on the Aura satellite of NASA’s Earth Observing System.

are assimilated in CIFS-AN. As shown in Figure 4, their impact is small because of the short lifetime of NO<sub>2</sub>. Compared to NO<sub>2</sub> retrievals from the GOME-2 instrument, C-IFS with or without assimilation of OMI NO<sub>2</sub> data severely underestimates wintertime NO<sub>2</sub> over East Asia (Figure 4b) and overestimates NO<sub>2</sub> over Southern Africa during the biomass burning season. At other times and over Europe, the agreement is better (Figure 4a).

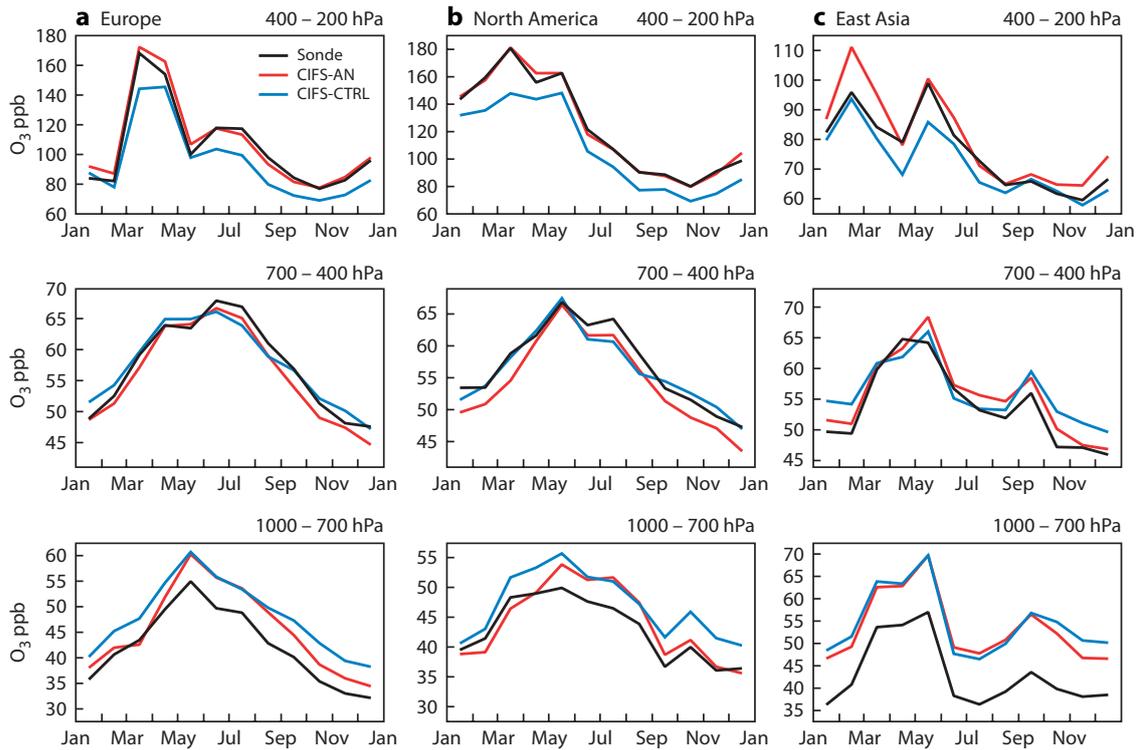
Even though the assimilation leads to large NO<sub>2</sub> analysis increments, this information is not retained by the model, and most of the impact of the data assimilation is lost from one analysis cycle to the next. It might be possible to improve this slightly by using a shorter assimilation window, e.g. 6-hour 4D-Var, and by using NO<sub>2</sub> retrievals from more than one satellite with different overpass times, but ideally the NO<sub>2</sub> data should be used to adjust the emissions in addition to the initial conditions.

**Towards a higher resolution**

C-IFS currently runs at a horizontal resolution of approximately 80 km at 60 vertical levels (T255L60). While this is a significantly higher resolution than is typically used in current global chemical transport models, it is far less than is used for numerical weather prediction (NWP), which



**Figure 2** Time series of monthly mean tropospheric carbon monoxide (CO) in ppb over (a) Frankfurt (50°N, 8.6°E, 837 profiles) and (b) Windhoek (22.5°S, 17.5°E, 323 profiles) averaged between 1000–700 hPa from MOZAIC aircraft data, CIFS-AN and CIFS-CTRL in 2008.

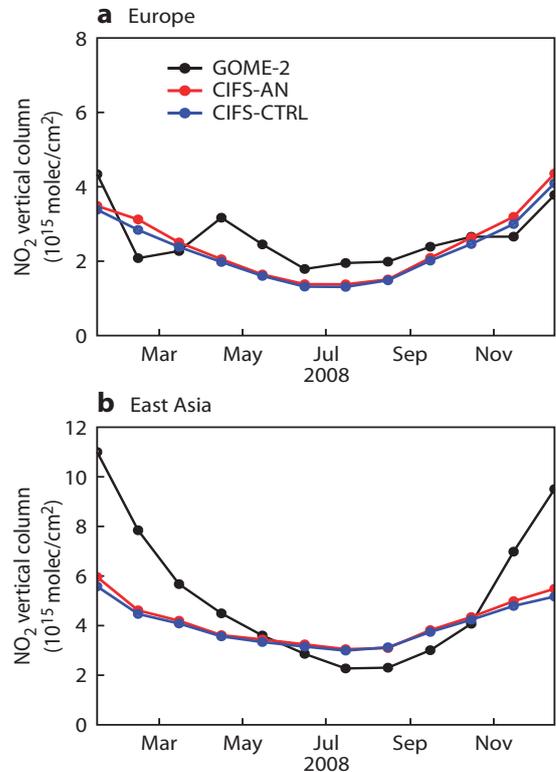


**Figure 3** Time series of monthly mean tropospheric ozone ( $O_3$ ) in ppb over (a) 11 sites in Europe, (b) 11 sites in North America and (c) 4 sites in East Asia averaged in the pressure bands 1000–700 hPa (bottom), 700–400 hPa (middle) and 400–200 hPa (top) from ozone sondes, CIFS-AN and CIFS-CTRL in 2008.

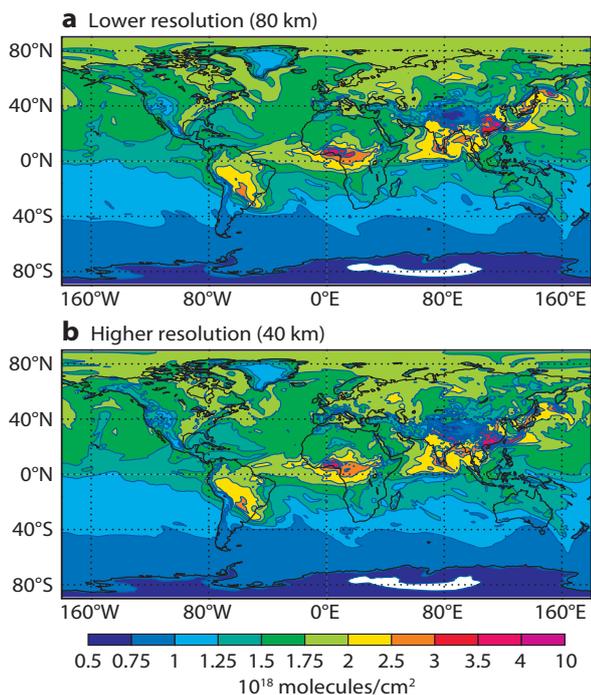
currently runs at approximately 16 km spacing and 137 levels (T1279L137). Increasing the resolution (in particular in the horizontal) of the MACC system is one of the key user requirements. With the move to CAMS, ECMWF will be in a position to increase the resolution of C-IFS in 2015. We have carried out tests at T511L60 (40 km at 60 levels) with the current C-IFS (CB05) system in data assimilation mode.

Two short data assimilation runs were performed at 40 km resolution (T511L60) from 1 to 5 December 2013 and 1 to 5 June 2014. The results were compared with a reference run at 80 km resolution (T255L60). Figure 5 shows 5-day global forecasts of total column carbon monoxide starting from the analysis of 5 December 2013. The standard-resolution experiment is shown in Figure 5a and the increased-resolution experiment is shown in Figure 5b. At first sight, differences are fairly small, because global CO concentrations are mostly determined by large-scale sources and transport.

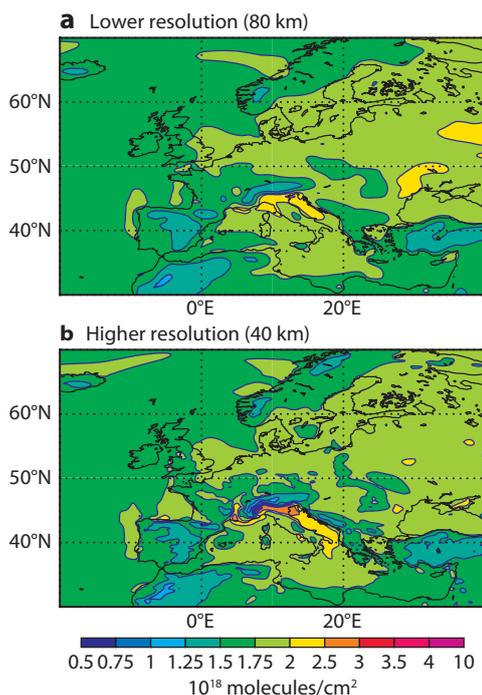
However, a closer look at, for instance, the European domain makes the impact of an increased resolution more visible, as shown in Figure 6. In this particular case, there is a significant difference in the area around the Alps and the Rhône and Po valleys. At the 40 km resolution, the mountains and valleys are much better resolved, resulting in lower total column CO values over the Alps (the column is smaller because the mountains are higher), and higher values in the river valleys, possibly also because of stronger blocking by the mountains for the Po valley and stronger confinement to the river valley for the Rhône.



**Figure 4** Time series of area-averaged tropospheric nitrogen dioxide ( $NO_2$ ) columns in  $10^{15}$  molecules/ $cm^2$  from GOME-2 retrievals, CIFS-AN and CIFS-CTRL for (a) Europe and (b) East Asia. (Figure provided by A.-M. Blechschmidt, University of Bremen.)



**Figure 5** Five-day forecast of total column carbon monoxide (CO) in  $10^{18}$  molecules/cm<sup>2</sup> starting from 5 December 2013 at (a) T255L60 and (b) T511L60 resolutions.



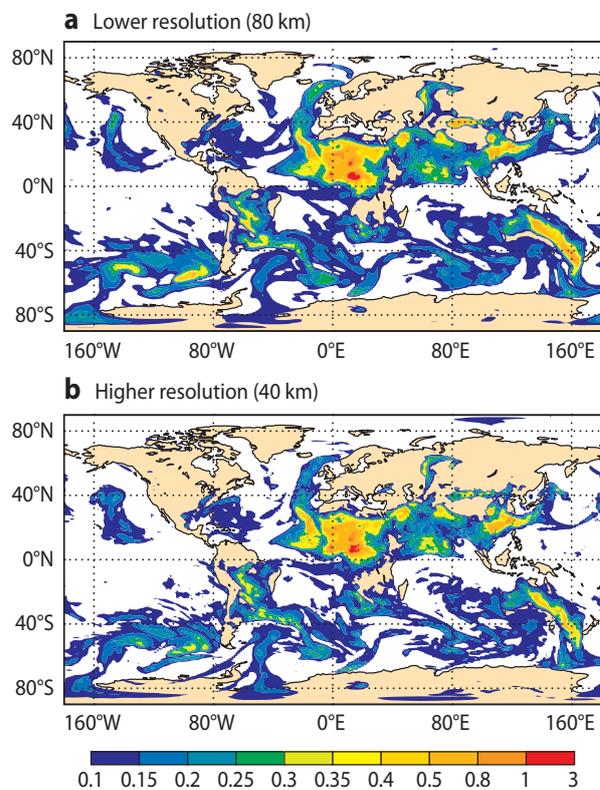
**Figure 6** Five-day forecast of total column carbon monoxide (CO) in  $10^{18}$  molecules/cm<sup>2</sup> for the European domain starting from 5 December 2013 at (a) T255L60 and (b) T511L60 resolutions.

For aerosol the picture is similar. However, as shown in Figure 7, there are larger differences in the global aerosol optical depth (AOD) than in total column CO. This is due to the fact that the amount of sea salt and dust aerosol in the atmosphere is a function of surface wind speeds, which are resolution-dependent, and also because aerosol lifetime is typically of the order of a week (instead of a month for CO), which leaves less time for mixing processes to smooth out high-frequency variability due to emission and transport processes.

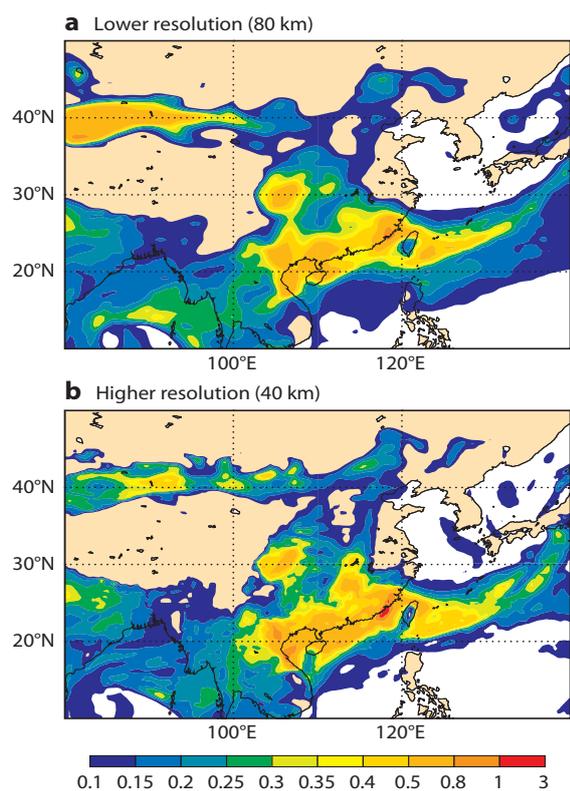
This becomes even clearer when we zoom in, this time on South East Asia. Figure 8 shows the AOD distribution for the two runs. The higher-resolution experiment shows more spatial variability, as expected, but also shows much lower values for desert dust over the Gobi desert north of Tibet. The low-resolution run has a smooth field with AOD values between 0.5 and 0.8, while the higher-resolution run has AOD values up to 0.5 with more spatial variability. Validation of longer runs will have to be used to assess which run matches the observed concentrations better.

While the comparisons for CO and AOD are mostly seeing differences due to transport and the meteorological impact on sources, a change in resolution may also be able to improve the modelling of atmospheric chemistry. For example, tropospheric ozone concentrations are determined by highly non-linear chemistry, and a change in resolution will potentially change the extent to which such chemical processes can be captured.

C-IFS has thus been tested successfully in data assimilation mode at a resolution of T511L60 (40 km at 60 levels). While



**Figure 7** Five-day forecast of aerosol optical depth (AOD) starting from 5 December 2013 at (a) T255L60 and (b) T511L60 resolutions.



**Figure 8** Five-day forecast of aerosol optical depth (AOD) for an Asian domain starting from 5 December 2013 at (a) T255L60 and (b) T511L60 resolutions.

the global distribution of the various species does not change significantly, there are interesting differences at the regional scale. Preliminary results suggest that various factors play a role, such as better-resolved emissions, better-resolved orography, better-resolved transport, the impact of different meteorological conditions on natural surface fluxes, and non-linear chemistry. A thorough investigation is under way to verify these results against independent observations. This should shed further light on the benefits of increasing the resolution of the C-IFS run operationally for CAMS.

### Next steps

With MACC-III finishing in mid-2015, CAMS will take over the delivery of services and will progressively transition to operational status. While most CAMS service elements are currently the subjects of Invitations to Tender, global CAMS products will be delivered by ECMWF using C-IFS. Activities are under way between ECMWF's Research, Forecast and Computing Departments in order to organise the transition to operational status. This transition aims to minimise the impact on current users of MACC-III, of whom there are about three thousand now.

In addition to supporting the delivery of CAMS services, C-IFS offers an excellent framework for studying and testing feedbacks between composition and meteorology at the global scale. The coupling between ozone and radiation in the critical region of the upper troposphere and the lower stratosphere is for instance expected to be improved by the

explicit treatment of tropospheric ozone. This is expected to improve numerical weather prediction beyond what could be achieved with the linear scheme for stratospheric ozone currently used in the IFS. Aerosol is also a key area for progress. The accurate treatment of highly variable emissions (dust, biomass burning, sea salt, etc.) is expected to capture variability especially in the tropical band. Finally, consideration of atmospheric composition in 4D-Var data assimilation could also bring benefits, in particular because the representation of radiative properties of the atmosphere can be made more realistic for radiance assimilation and because certain species have tracer-like properties that can provide information on atmospheric motions and winds.

Currently, MACC production is run separately from the weather forecasting operation. Research is under way to investigate the merging of production chains by enabling the handling of different simultaneous data grids in the IFS. In addition to providing operational synergies, this will pave the way for next-generation, fully integrated weather and environmental monitoring and forecasting applications.

### FURTHER READING

- Baklanov, A., K. Schlünzen, P. Suppan, J. Baldasano, D. Brunner, S. Aksoyoglu, G. Carmichael, J. Douros, J. Flemming, R. Forkel, S. Galmarini, M. Gauss, G. Grell, M. Hirtl, S. Joffre, O. Jorba, E. Kaas, M. Kaasik, G. Kallos, X. Kong, U. Korsholm, A. Kurganskiy, J. Kushta, U. Lohmann, A. Mahura, A. Manders-Groot, A. Maurizi, N. Moussiopoulos, S.T. Rao, N. Savage, C. Seigneur, R.S. Sokhi, E. Solazzo, S. Solomos, B. Sørensen, G. Tsegas, E. Vignati, B. Vogel, & Y. Zhang, 2014: Online coupled regional meteorology chemistry models in Europe: current status and prospects, *Atmos. Chem. Phys.*, **14**, 317–398, doi:10.5194/acp-14-317-2014.
- Flemming, J., V. Huijnen, J. Arteta, P. Bechtold, A. Beljaars, A.-M. Blechschmidt, B. Josse, M. Diamantakis, R.J. Engelen, A. Gaudel, A. Inness, L. Jones, E. Katragkou, V. Marecal, V.-H. Peuch, A. Richter, M. G. Schultz, O. Stein and A. Tsikerdekis, 2015: Tropospheric chemistry in the Integrated Forecasting System of ECMWF, *Geosci. Model Dev.*, **8**, 975–1003, doi:10.5194/gmdd-8-975-2015.
- Inness, A., A.-M. Blechschmidt, I. Bouarar, S. Chabrillat, M. Crepulja, R.J. Engelen, H. Eskes, J. Flemming, A. Gaudel, F. Hendrick, V. Huijnen, L. Jones, J. Kapsomenakis, E. Katragkou, A. Keppens, B. Langerock, M. de Mazière, D. Melas, M. Parrington, V.-H. Peuch, M. Razinger, A. Richter, M.G. Schultz, M. Suttie, V. Thouret, M. Vrekoussis, A. Wagner and C. Zerefos, 2015: Data assimilation of satellite retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS, *Atmos. Chem. Phys. Discuss.*, **15**, 4265–4331, doi:10.5194/acpd-15-4265-2015.
- Inness, A., F. Baier, A. Benedetti, I. Bouarar, S. Chabrillat, H. Clark, C. Clerbaux, P. Coheur, R.J. Engelen, Q. Errera, J. Flemming, M. George, C. Granier, J. Hadji-Lazaro, V. Huijnen, D. Hurtmans, L. Jones, J.W. Kaiser, J. Kapsomenakis, K. Lefever, J. Leitão, M. Razinger, A. Richter, M.G. Schultz, A.J. Simmons, M. Suttie, O. Stein, J.-N. Thépaut, V. Thouret, M. Vrekoussis, C. Zerefos, & the MACC team, 2013: The MACC reanalysis: an 8 yr data set of atmospheric composition, *Atmos. Chem. Phys.*, **13**, 4073–4109, doi:10.5194/acp-13-4073-2013.

# Snow data assimilation at ECMWF

PATRICIA DE ROSNAY, LARS ISAKSEN,  
MOHAMED DAHOU

Over the last few years, significant changes to snow analysis have been implemented in the operational ECMWF land data assimilation system. The Optimal Interpolation interface to observations was revised, allowing the implementation of land surface observation monitoring capabilities for snow and screen-level observations in Cycle 40r1 of ECMWF's Integrated Forecasting System (IFS). Developments also include improvements in the combined use of snow depth and snow cover data. Cycle 40r1 has been operational since November 2013.

An evaluation against independent observations shows that the revised snow data assimilation performs better than the earlier version on a range of scores. The revision has brought clear improvements both in snow depth and near-surface weather parameter forecasts. Continuous observation monitoring provides evidence of the high quality of short-range ECMWF snow depth forecasts, but it also shows a slight continuing overestimation of snow depth in the background field and the analysis compared with independent observations.

## History of ECMWF snow data assimilation

The first ECMWF snow data assimilation system was implemented more than 25 years ago, based on the Cressman interpolation technique. It was limited to the use of in situ snow depth observations (SYNOP observations). ECMWF's snow analysis was revised in 2004 to introduce 24-km Interactive Multi-Sensor Snow and Ice Mapping System (IMS) snow cover information, in addition to the SYNOP snow depth measurements. This led to a more realistic representation of the extent of snow cover in the operational analysis (*Drusch et al., 2004*). However, the persistent large amounts of snow in the northern hemisphere during the cold winter of 2009/2010 highlighted several problems in the operational snow analysis and motivated further work to improve the snow data assimilation method and the processing of snow observations.

In November 2010 (IFS Cycle 36r4), the snow analysis code was revised to use a two-dimensional (2D) Optimal Interpolation (OI) method instead of the Cressman interpolation. The difference between the Cressman and the OI analyses mainly results from differences in the structure functions that determine how an observation influences the analysis in the surrounding region. Compared to the Cressman interpolation, the 2D OI produces smoother and more realistic snow analysis patterns that are in better agreement with in situ observations (*de Rosnay et al., 2014*). In the same IFS cycle, the use of IMS was revised in several ways to improve the data and model collocation. The snow cover IMS product itself, which is provided by the US National Environment

Satellite Data and Information Service (NESDIS), was upgraded to a higher resolution (4 km), providing useful spatial detail particularly in coastal areas.

These revisions in the use of IMS snow cover data had a positive impact on atmospheric forecast quality, both for screen-level parameters and for the northern hemisphere geopotential height at 500 hPa, which significantly improved up to the 4-day forecast range. Further improvements in IMS data acquisition and assimilation in 2012 led to a reduction in IMS snow cover data latency by 12 hours. The latest improvements, implemented in November 2013 (IFS Cycle 40r1), include the revision of the surface analysis interface to observations to enable the monitoring of conventional land surface and IMS snow cover observations. This has been achieved by using the Observation Data Base (ODB) format for both types of observations. Along with these modifications, the snow data assimilation algorithm was revised to assimilate IMS non-zero snow cover observations together with in situ snow depth observations in the OI analysis.

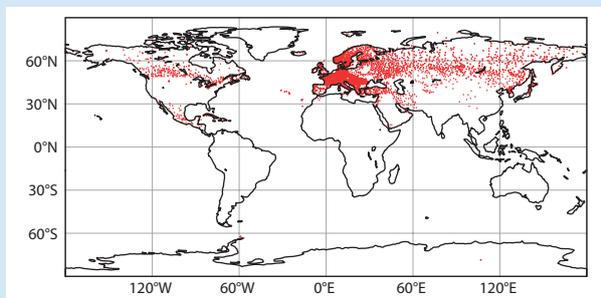
## Snow observations

In situ snow depth observations constitute a very important and very reliable source of information for the snow data assimilation system. They are provided by the SYNOP station network and are made available in near real-time on the Global Telecommunication System (GTS) for numerical weather prediction (NWP) applications. In addition to SYNOP observations, National Meteorological Services (NMSs) maintain national snow depth measurement networks. For example the US SNOTEL (SNOWpack TELEmetry) network provides snow depth measurements that are used for snow data assimilation at the National Oceanic and Atmospheric Administration (NOAA). The Cooperative Observer Program (COOP) also constitutes a very dense in situ snow depth measurement network in the US. However, these additional in situ snow depth observations are currently not available on the GTS for operational NWP applications.

In Europe, an initiative launched in 2011 by ECMWF and NMSs in ECMWF's Member States has begun to improve the availability of in situ snow depth observations. NMSs were encouraged to use a dedicated BUFR format developed by ECMWF to report their additional national snow depth observations on the GTS. This has led to a significant increase in the availability of in situ observations for operational NWP. To date, seven ECMWF Member States (Sweden, the Netherlands, Denmark, Romania, Hungary, Norway and Switzerland) report snow depth daily from more than 600 additional stations in their networks. These additional in situ observations are used alongside previously available SYNOP reports in the operational snow depth analysis (Box A). In 2013, this European initiative was extended to the World Meteorological Organization (WMO) in the context of the Snow Watch project of the

### Snow observations used at ECMWF

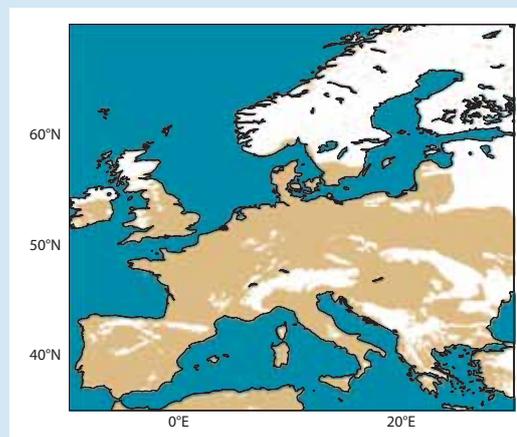
The snow analysis relies on SYNOP and national networks of snow depth ground observations available on the GTS as well as on the IMS snow cover information available for the northern hemisphere.



The map shows the spatial distribution of in situ station reports available on the GTS on 20 January 2015. On this day, 16,112 snow depth observations were reported from 3,810 stations. A total of 2,844 stations reported snow depth using the Traditional Alphanumeric Code (TAC), 663 stations reported additional snow data using the dedicated snow BUFR template, 303 new SYNOP stations reported in BUFR, and 1,573 stations reported both in BUFR and TAC. In IFS Cycle 40r1, the first two types of snow depth reports are used.

A

The IMS product provides cover maps for the northern hemisphere. It combines microwave and visible sensors to provide binary snow cover information in all weather conditions. An IMS observation of snow indicates that at least 50% of the grid cell is snow-covered.



IMS snow cover data over Europe on 20 January 2015

Global Cryosphere Watch programme in order to make data from synoptic and climate networks more widely available over the GTS (Brun *et al.*, 2013). Improving access to snow observations is also a key objective of the current European Cooperation Science and Technology (COST) Action on snow (ES1404).

Although these developments have improved the availability of in situ snow depth observations, snow depth observations are still unavailable to NWP for many regions of the world. Satellite observations provide useful information, especially in remote locations with sparse in situ station coverage. The IMS snow cover product is widely used by the NWP community to analyse snow depth. It is a multi-sensor product that combines satellite data primarily from visible sensors with microwave data and weather station reports to provide snow cover information with good accuracy in all weather conditions. Several validation and comparative studies have demonstrated the accuracy of the IMS product against in situ and other products, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) snow cover product (Brubaker *et al.*, 2009, Helfrich *et al.*, 2007). At ECMWF, the IMS northern hemisphere daily product has been used operationally since 2004, initially at a resolution of 24 km and since 2010 at a resolution of 4 km. The 4-km IMS product includes sea ice extent as well as the snow cover information over land currently used by the land data assimilation system.

### Snow data assimilation method

ECMWF snow analysis uses a 2D OI method that is performed every 6 hours. A short-range IFS forecast

provides the snow depth background field, computed from the prognostic snow water equivalent and snow density variables. The observations used are in situ snow depth measurements and the IMS northern hemisphere snow cover product. The IMS product only provides information on the presence of snow for each 4-km grid cell. A snow cover IMS observation indicates that at least 50% of the grid cell is covered by snow, but it does not indicate the snow depth. To enable the snow cover product to be used quantitatively in the snow analysis, the IMS information is converted into quantitative snow depth information using the relation between snow cover and snow depth as detailed in Box B. In the snow data assimilation, the quality control includes a redundancy check and a first guess departure check. More details can be found in the IFS Cycle 40r1 documentation available on ECMWF's website. The revised snow OI algorithm implemented in IFS Cycle 40r1 enables observations to be assimilated jointly from in situ reports and IMS snow cover in the OI. It replaces the two-step algorithm that was used from 2004 to 2013, in which, prior to the analysis, the snow depth background field was replaced with 0.1 m of snow at locations where the background field had no snow while the IMS indicated the presence of snow (Box B).

### Data assimilation experiments and results

To evaluate the impact of the revised snow analysis on snow depth and forecasts of near-surface weather parameters, global data assimilation experiments were performed. They were conducted at T511 (40-km grid) with IFS Cycle 40r1 from 1 October 2012 to 30 April 2013. The

two experiments differ only in their snow analysis, with one experiment set up to use the IFS Cycle 38r2 snow analysis ('old') and the other experiment using the revised IFS Cycle 40r1 snow analysis ('revised').

**Evaluation using independent snow depth observations**

To evaluate the impact of the revised snow analysis, it is important to use independent observations. For the purpose of this study, a fixed subset of 251 in situ stations (around 10% of available stations) reporting near-real-time snow depth observations were randomly selected to provide validation data. Their snow depth reports were excluded from the assimilation (for both experiments) during the entire period. Figure 1 shows the location of the independent validation snow depth stations reporting on 10 December 2012. They are mostly located in Europe and Siberia. A few validation stations are also located in Mongolia, Iran, Canada and Central America.

The presence of snow on the ground in the data assimilation experiments was evaluated for every day from 1 October 2012 to 30 April 2013. The evaluation presented here is based on a 2x2 snow/no-snow contingency table:

	Snow observed	No snow observed
Snow in analysis	a Hits	b False alarm
No snow in analysis	c Misses	d Correct no snow

The following scores are used for the evaluation:

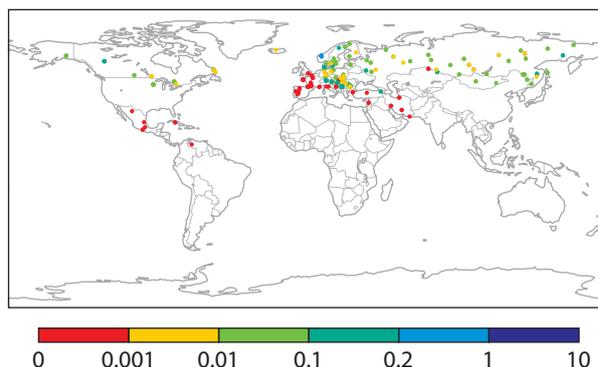
- Accuracy =  $a + d / (a + b + c + d)$
- Hit rate (probability of detection) =  $a / (a + c)$
- False alarm ratio =  $b / (a + b)$
- Threat score =  $a / (a + b + c)$
- Frequency bias index =  $(a + b) / (a + c)$

Table 1 summarises the scores obtained for the two data assimilation experiments conducted with the old and revised snow analyses. It shows that the revised scheme better captures the snow/no-snow occurrences with all the scores consistently improved against independent observations. The overall accuracy is increased from 0.92 to 0.94. The impact of the snow analysis revision on the snow occurrence accuracy is also shown in Figure 2a. Most

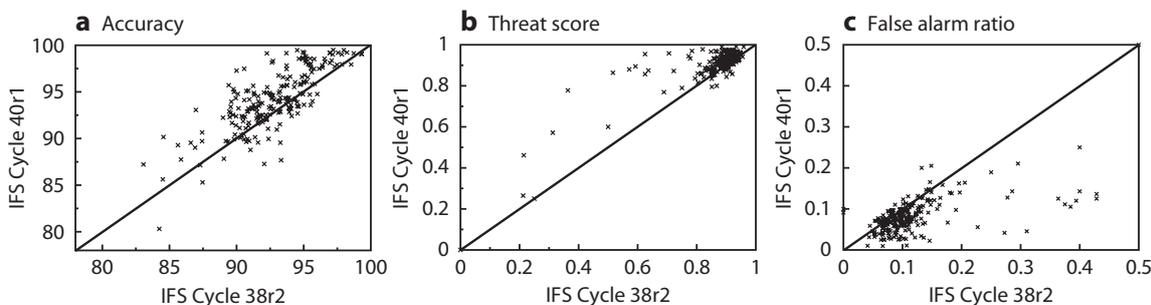
of the points are above the diagonal, which is indicative of improvements in snow and no-snow detection.

This score is strongly influenced by the two most common categories (correct no snow in October and April and correct snow in December and January). The hit rate score is higher for the revised analysis (0.98) than for the old analysis (0.96), which shows that the revised analysis captures the occurrence of snow on the ground better than the old analysis. The false alarm ratio has been reduced from 0.12 to 0.09 and the threat score increased from 0.86 to 0.9. These two scores are also shown in the form of scatter plots in Figures 2b–c. Here again we see an overall improvement in the threat score and the false alarm ratio at the daily time scale. The frequency bias index is reduced from 1.10 to 1.07 in the revised snow analysis, which suggests that the new system reduces the overestimation of snow on the ground (compared to independent in situ observations).

Figure 3 shows the temporal evolution of the revised snow analysis scores against the independent in situ observations of snow depth. The four elements of the contingency table are represented as a percentage of the total number of events. Since almost all the in situ



**Figure 1** In situ stations used for research experiment evaluation. The colour scale illustrates snow depth in metres on 10 December 2012. These stations are the subset of 251 stations used in the evaluation, which were randomly selected from the SYNOP network and national networks. They are not used in the data assimilation experiments and therefore they constitute an independent validation dataset.



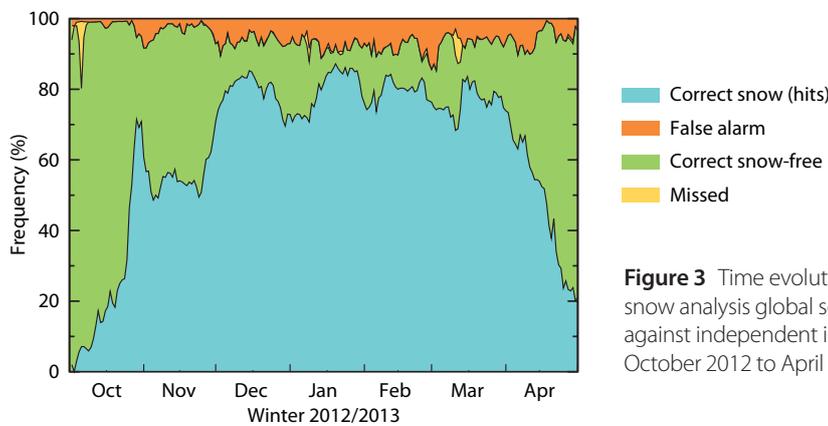
**Figure 2** Snow analysis scores for the revised IFS 40r1 snow analysis versus the IFS 38r2 analysis for (a) accuracy, (b) threat score, and (c) false alarm ratio in the period October 2012 to April 2013. Each cross represents the scores computed against 251 independent in situ snow depth observations for a given date. The scatter plots show the results for each of the 212 days from 1 October 2012 to 30 April 2013. The black line represents the one-to-one line.

stations are located in the northern hemisphere, annual cycles of correct snow and correct no-snow events follow the northern hemisphere winter season. In October, most stations report snow-free conditions, which are well captured by the analysis. From October to December, the proportion of snow events correctly detected increases to about 80%. The proportion of the sum of hits and correct

no-snow events represents the accuracy, which remains above 90% throughout the season, with a mean value of 0.94, as indicated in Table 1. Figure 3 shows that misses are very infrequent compared to false alarms. This is consistent with the slight overestimation of snow occurrence in the IFS shown in Table 1, with frequency bias index values larger than one.

	Accuracy	Hit rate	False alarm ratio	Threat score	Frequency bias index
IFS 38r2 snow depth analysis	0.92	0.96	0.12	0.86	1.10
IFS 40r1 snow depth analysis	0.94	0.98	0.09	0.90	1.07

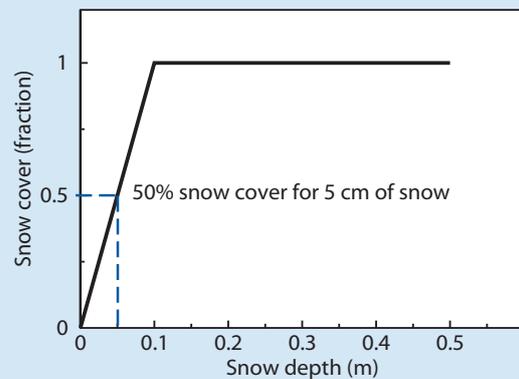
**Table 1** Global evaluation of snow depth analyses using IFS Cycles 38r2 and 40r1 for the period from 1 October 2012 to 30 April 2013. Scores are computed using independent snow depth observations.



**Figure 3** Time evolution of the revised 40r1 snow analysis global scores (in percent) against independent in situ stations for October 2012 to April 2013.

### IMS snow cover data assimilation in the IFS

IMS provides binary information on the presence of snow for each grid cell but it does not give any quantitative information on snow depth. In IFS Cycle 40r1, the model relation between snow cover and snow depth (Dutra et al., 2010), as shown in the figure on the right, is used to convert snow cover IMS information into a snow depth estimate relevant for data assimilation. The assumption is that 0.05 m of snow corresponds to 50% snow coverage, which is the IMS binary threshold. For cells with binary snow-free conditions, IMS is converted into 0 cm of snow depth. So, snow-free IMS observations enter the analysis with an assumed snow depth of 0 cm whatever the model background. Snow cover observations enter the analysis with an assumed snow depth of 5 cm if the model background indicates snow-free conditions. Snow cover IMS observations do not enter the analysis if the model background already indicates snow cover (see table). In this way, the model background, in situ and IMS observations are optimally combined in the OI analysis. Standard deviation values of background, in situ and IMS snow depth error are set at 0.03 m, 0.04 m and 0.08 m, respectively. This single-step algorithm replaces the previous two-step algorithm used until IFS Cycle 38r2: In a first step, a Direct Insertion (DI) approach was used to update the model background with 0.1 m in the case of snow-covered conditions in IMS and snow-free conditions in the model. In a second step, IMS snow-free observations were used in the analysis using the same observation error as with in situ data.



Model relation between snow cover and snow depth used to convert IMS binary information of snow cover into quantitative snow depth information in the event of a snow-free background.

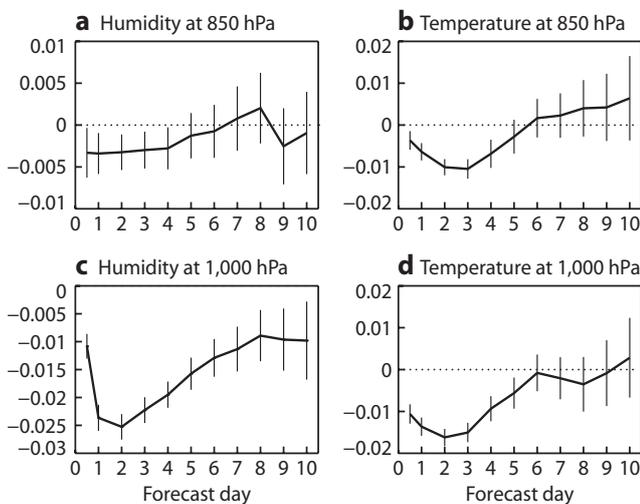
IMS \ Model	Snow	Snow-free
Snow	not used	Data assimilation 0.05 m
Snow-free	Data assimilation 0 m	Data assimilation 0 m

IMS equivalent snow depth observation that enters the snow analysis depending on the IMS and background snow status.

B

**Forecasts scores**

Figure 4 shows the mean impact for October 2012 to April 2013 of the revised snow analysis on temperature and humidity forecasts at 1000 hPa and at 850 hPa. It shows that the revised analysis improves both temperature and humidity forecasts. At 1000 hPa, the root mean square error is significantly reduced at all forecast ranges for humidity and until day 5 for temperature. At 850 hPa, the error is reduced until forecast day 4 for both humidity and temperature. Figure 5 shows a map of the root mean square error difference between the revised and old snow analyses for the 72-hour temperature and humidity forecasts. It shows a clear error reduction in continental areas of the northern hemisphere, particularly in parts of North America and Northeast Asia.



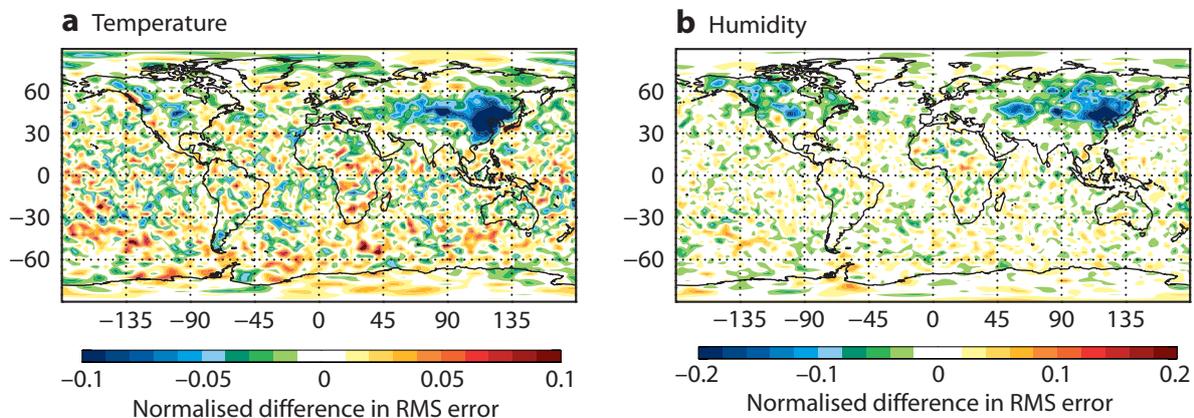
**Figure 4** Impact of the revised snow analysis on the normalised root mean square error difference between IFS Cycles 40r1 and 38r2 (40r1 minus 38r2) for (a) humidity forecasts at 850 hPa; (b) temperature forecasts at 850 hPa; (c) humidity forecasts at 1,000 hPa and (d) temperature forecasts at 1,000 hPa in the extratropical northern hemisphere (20° to 90° latitude). Scores are computed against own analysis from 01 October 2012 to 30 April 2013. Error bars indicate 95% confidence range.

**Observation monitoring**

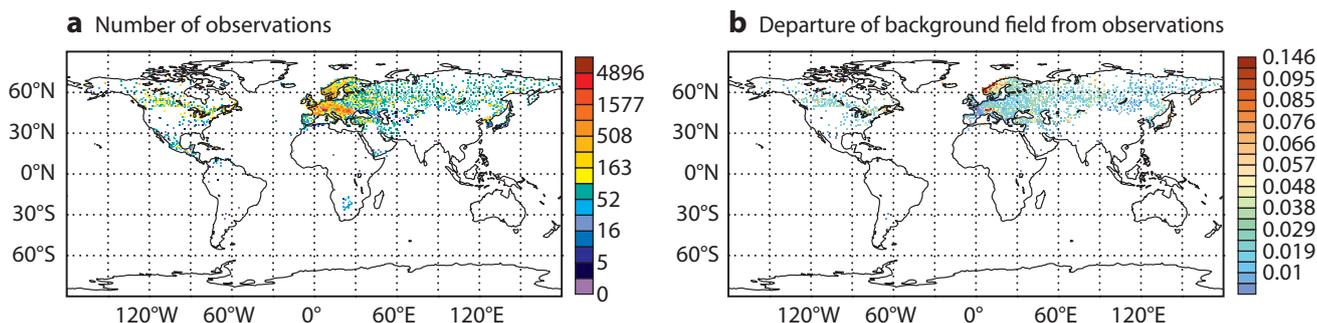
IFS Cycle 40r1 has been used for operational forecasts since 19 November 2013. As part of the implementation of the revised snow analysis, the interface to observations was also revised to allow the monitoring of observations of conventional screen-level parameters, snow depth and IMS snow cover. Operational monitoring is an important part of the ECMWF data assimilation system. It ensures the continuous evaluation of observation counts and departures of background fields and analyses from observations both spatially and in time. It helps to identify and subsequently blacklist suspect observations. When relevant, feedback is communicated to data providers. Observation monitoring is also important for the evaluation of IFS model and data assimilation system experiments.

Figure 6a shows a monitoring map of the mean number of observations used in the surface analysis for December 2014 to February 2015. It illustrates the uneven distribution of snow depth reports, with very dense snow depth observations available over Europe and sparse observation reporting over the USA and China and in the southern hemisphere. ECMWF’s operational monitoring of the number of snow depth observations provides a clear and continuous record of current near-real-time availability of in situ snow depth reports on the GTS. It feeds into WMO snow watch initiatives aiming to improve snow depth report availability.

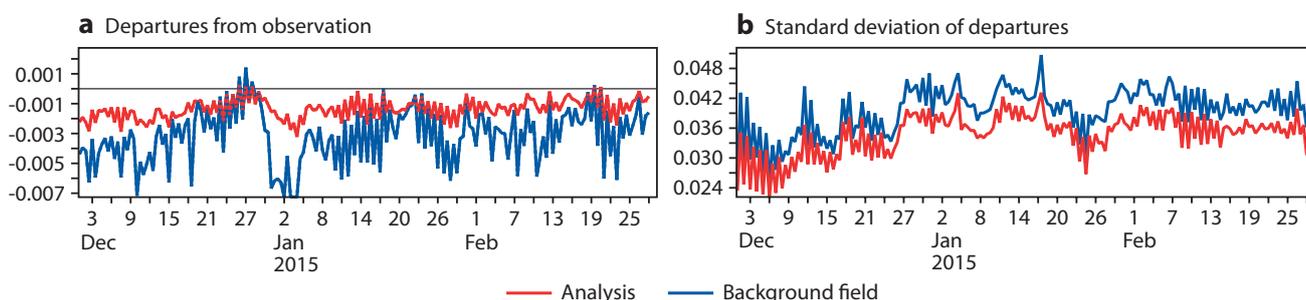
Figure 6b shows the standard deviation of the ECMWF innovations (observation minus background) for December 2014 to February 2015. It shows a mean standard deviation of around 0.029 m, which is indicative of the high quality of short-range ECMWF snow depth forecasts evaluated against snow depth reports. The time evolution of these statistics is also monitored operationally, as shown in Figure 7 for December 2014 to February 2015. Figure 7a shows the evolution of the mean departure of the background and of the analysis from observations. The negative values shown are consistent with a slight overestimation of snow depth, as discussed in the previous



**Figure 5** Normalised impact of the revised snow analysis compared to the old snow analysis on the root mean square error of 72-hour forecasts of (a) temperature and (b) humidity at 1,000 hPa. Scores are computed against own analysis for October 2012 to April 2013. Blue shades indicate a smaller root mean square error in the revised snow analysis, yellow and red shades indicate a larger root mean square error in the revised snow analysis.



**Figure 6** Snow monitoring statistics for December 2014 to February 2015 showing (a) the number of observations per 1-degree grid cell and (b) the standard deviation of the departure of background field values from observations, in metres.



**Figure 7** Monitoring time series from December 2014 to February 2015 of the ECMWF operational IFS Cycle 40r1 suite for conventional snow depth showing (a) mean departures of background field and analysis from observations, in metres (b) standard deviation of background field and analysis departures from observations, in metres.

section, when compared against independent observations. In the operational snow analysis, the mean background and analysis departures over the whole period are -0.003 m and -0.001 m, respectively. The standard deviation, shown in Figure 7b, has mean values of 0.029 m and 0.025 m for background and analysis departures, respectively. Both statistics show good stability and consistent performance of the analysis through the winter season.

**Summary**

ECMWF’s snow analysis has been revised substantially over the last few years. For the last two winters, IFS Cycle 40r1 has been used in operations. It has benefited from improvements in IMS snow cover data assimilation and new operational monitoring capabilities. Research experiments show a clear improvement of both snow depth and near-surface weather parameter forecasts directly related to the snow analysis revisions. An evaluation of snow depth analysis against independent snow depth reports shows that ECMWF’s analysis performs very well in estimating snow depth and snow cover. The operational monitoring of snow depth observations helps to assess the quality of both observations and ECMWF’s snow depth analysis and background field. This highlights the need for further improvements in the availability of near-real-time snow depth reports on the GTS for NWP applications.

**FURTHER READING**

**Brubaker, K., R. Pinker & E. Deviatova**, 2009: Evaluation and comparison of MODIS and IMS snow-cover estimates for the continental United States using station data. *J Hydrometeorology*, **6**, 1002–1017

**Brun, E., J. Lawrimore, P. de Rosnay & J. Friddell**, 2013: A Global Cryosphere Watch Initiative for improving in-situ snow observations and their access and for rescuing/collecting historical in-situ snow data, WMO Global Cryosphere Watch Snow Watch document, October 2013

**de Rosnay, P., G. Balsamo, C. Albergel, J. Muñoz-Sabater & L. Isaksen**, 2014: Initialisation of land surface variables for Numerical Weather Prediction, *Surveys in Geophysics*, **35(3)**, 607–621, doi: 10.1007/s10712-012-9207-x

**Drusch, M., D. Vasiljevic, P. Viterbo**, 2004: ECMWF’s global snow analysis: assessment and revision based on satellite observations. *J Appl Meteorol*, **43**, 1282–1294

**Dutra, E., G. Balsamo, P. Viterbo, P. Miranda, A. Beljaars, C. Schär & K. Elder**, 2010: An improved snow scheme for the ECMWF land surface model: description and offline validation. *J Hydrometeorol*, **11**, 899–916, doi:10.1175/2010JHM1249.1

**Helfrich S.R., D. McNamara, B. Ramsay, T. Baldwin, T. Kasheta**, 2007: Enhancements to, and forthcoming developments in the interactive multisensor snow and ice mapping system, (IMS). *Hydrol Process*, **21**, 1576–1586, doi:10.1002/hyp.6720

# Supercomputing at ECMWF

MIKE HAWKINS, ISABELLA WEGER

ECMWF's latest High-Performance Computing Facility (HPCF), two Cray XC-30s with over 160,000 processor cores in a resilient configuration, is at the heart of ECMWF's production of weather forecasts and cutting-edge research in numerical weather prediction (NWP). The two clusters, which are among the most powerful supercomputers in Europe, have been producing ECMWF's operational forecasts since September 2014. In addition to ECMWF's operational and research activities, ECMWF's Member States also have access to the HPCF and scientists may apply to use it for Special Projects. Figure 1 shows 'Ventus', one of the two Cray clusters installed in ECMWF's data centre.

Numerical weather prediction has always relied on state-of-the-art supercomputers to run a complex numerical model of the atmosphere in the shortest possible period of time. ECMWF's current Integrated Forecasting System (IFS) is based on a numerical model with 293 million grid points, ingests 40 million observations per day and takes 2 hours and 10 minutes to produce a 10-day high-resolution global forecast.

ECMWF's first operational forecast in 1979 was produced on a single-processor Cray-1A. The fastest supercomputer at its time, it had a peak performance of 160 million floating-point operations per second, around a tenth of the computing power of a modern smartphone. The peak performance of the Cray XC-30 system is 21 million times greater, equivalent to a stack of smartphones more than 15 kilometres tall.

The demand for more accurate and reliable forecasts and for better early warnings of severe weather events, such as windstorms, tropical cyclones, floods and heat waves, requires continual improvements of ECMWF's numerical models and data assimilation systems. Finer model grid

resolutions, a more realistic representation of physical processes in the atmosphere, and the assimilation of more observations are the main drivers for better computational performance. ECMWF's forecasting system has also developed towards a more comprehensive Earth-system model: the atmospheric model is coupled to ocean, wave, sea ice and land-surface models and now includes the composition of the atmosphere (e.g. aerosols and greenhouse gases).

The growing computational requirements resulting from ECMWF's scientific and operational strategy require ECMWF to replace and upgrade its HPCF on a regular basis. Competitive procurements are run every four to five years and contracts have built-in upgrade cycles of about two years to take advantage of improvements in technology and to better match the system to operational and research needs.

Figure 2 shows the evolution of sustained HPC performance at ECMWF, with the Cray XC-30 as the most recent system. Sustained performance is measured by ECMWF benchmark codes that represent the current operational version of the IFS.

## Looking back to the beginnings of supercomputing

From 1976 to 1978, ECMWF leased access to a Control Data Corporation (CDC) 6600 computer, hosted by CDC at John Scott House in Bracknell. The CDC6600 was one of the most powerful systems available at the time and considered the first successful supercomputer. It allowed the development of the first version of ECMWF's weather forecasting model, but it still needed 12 days to produce a 10-day forecast.

The CDC6600 experience showed that, provided a suitably powerful computer could be acquired, useful forecasts could be produced. Figure 3 shows an example architecture envisaged at the time. The first supercomputer owned by ECMWF was installed on 24 October 1978 at the new



Figure 1 'Ventus' is one of the two Cray XC-30 clusters installed in ECMWF's data centre.

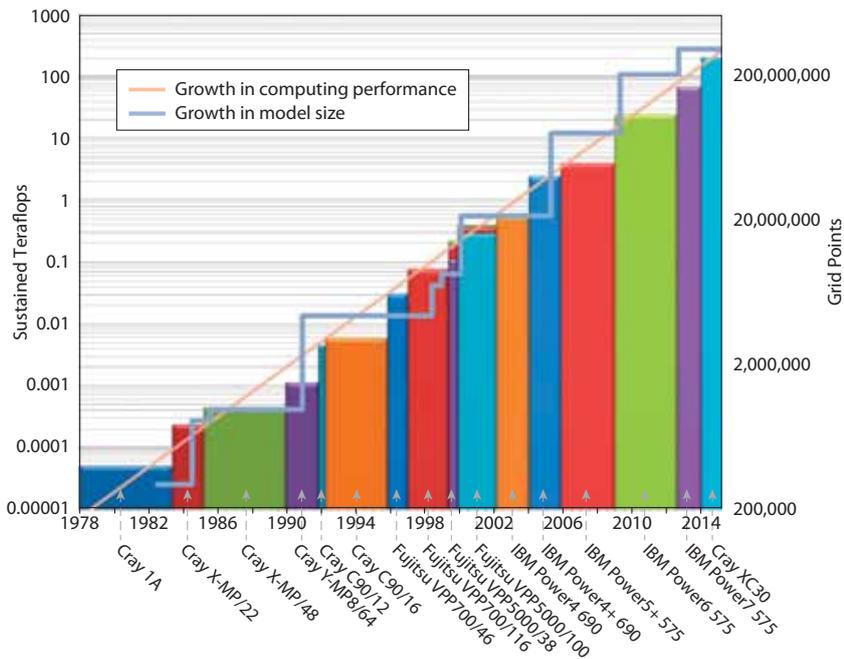


Figure 2 HPC growth versus size of ECMWF's high-resolution forecast model.

Shinfield Park site. This system was a CRAY-1A, serial number 9, manufactured by Cray Research. Before then, the Centre's scientists also had access to an IBM 360/195 at the British Meteorological Office and later to 'Serial 1', hosted at the Rutherford Laboratory, the first production model of the CRAY 1 series, in order to test out all the programs required to produce a real operational forecast.

**From the 70s to the 90s – the first Cray era**

The Cray-1A was a single-processor computer with a memory of eight megabytes and a 2.4 gigabyte storage system. The processor could produce two results per cycle, with a cycle time of 12.5 nanoseconds, giving a theoretical peak performance of 160 megaflops (160 million arithmetic calculations per second), about one tenth the performance of a modern smartphone. Running the operational weather model, the machine was capable of a sustained performance of 50 megaflops, allowing an operational 10-day forecast to be produced in five hours.

The era of Cray systems at ECMWF lasted 18 years, until 1996. In that time, a succession of systems advanced the total sustained performance the model could achieve from 50 megaflops to 6,000 megaflops. Despite the relatively long time period, most of the systems were quite similar in design. They all had a small number of fast processors, 16 in the last system, and each of these processors had access to all the memory on the machine. This 'shared memory' configuration is the basic building block of the large systems we use today.

The other important feature of the Cray systems was the use of vector instructions, single instructions that could work on single-dimensional arrays of data. With a vector instruction, to add 10 numbers together, the set of 10 numbers is loaded and then added up in one go. This parallelism gives much better performance than doing the additions one after the

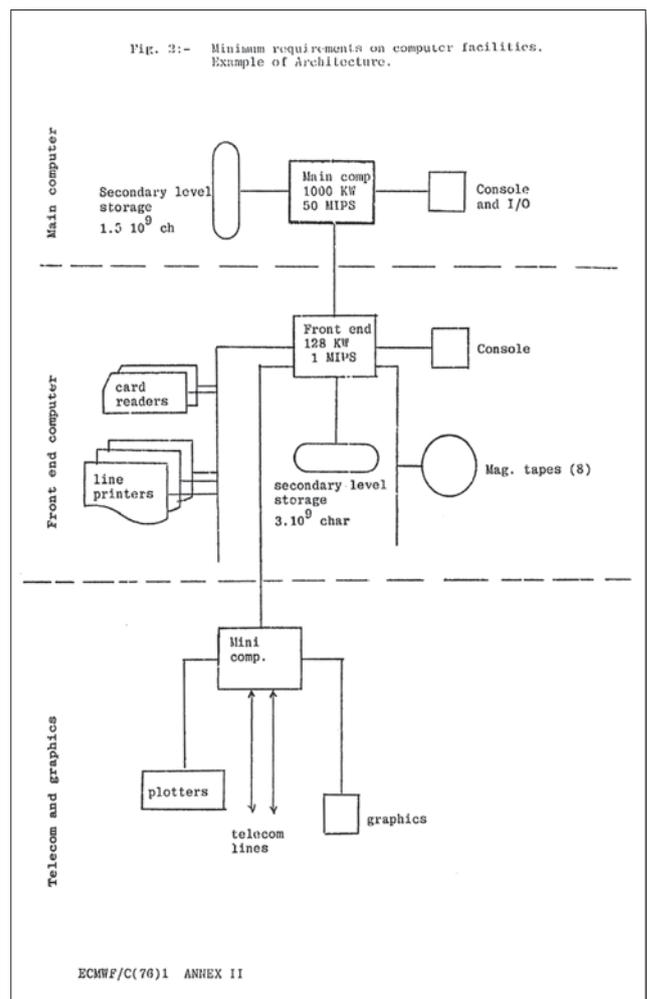


Figure 3 Example architecture to meet the minimum requirements of the computer facility – from a paper to the ECMWF Council, April 1976.

other in a loop, as would be done on a 'scalar' system. Vector instructions are also used in modern processors. They were the key building block of the next era of supercomputing at ECMWF, the Fujitsu vector systems.

The last of the Cray machines was quite different from the others. Delivered in 1994, the Cray T3D was a massively parallel system. Rather than the 16 specially designed and built processors of its predecessor, the system had 128 Alpha processors, produced by Digital Equipment Corporation (DEC). Each processor had 128 megabytes of memory. The name of the system came from the network, which was arranged in the shape of a torus. With the memory distributed between the processors rather than shared into one big pool, substantial changes to the weather forecasting system were required to operate efficiently on this type of architecture.

**From the 90s to 2002 – the Fujitsu era**

In 1996, a small Fujitsu test system, a VPP300/16, was installed. This was followed by a larger VPP700/46 system, which started producing operational forecasts from 18 September 1996. The VPP700 had 39 processing elements for computing, six for input/output (I/O), and one running the batch system and interactive work. Unlike the Cray systems, the VPP systems had a distributed-memory, parallel architecture. Each of the vector processing elements only had direct access to their own two gigabytes of memory. A built-in, fully non-blocking crossbar switch acted as a high-speed network that allowed the processing elements to communicate with each other. The model ran on 18 processing elements and achieved a sustained performance of around 30 gigaflops, a fivefold increase over the last Cray. The system as a whole had a peak performance equivalent to around 60 modern smartphones.

The distributed memory of the Fujitsu necessitated a rewrite of the forecast model. Cray-specific code features had to be removed and the standard Message Passing Interface (MPI) adopted so that the code would work efficiently on the new system. The rewrite made the code fully portable to different

architectures, an important feature retained to this day. Figure 4 shows the structure of ECMWF's computer systems towards the end of the Fujitsu era. The Fujitsu systems continued successfully for six years, increasing the sustained performance 13-fold to 400 gigaflops.

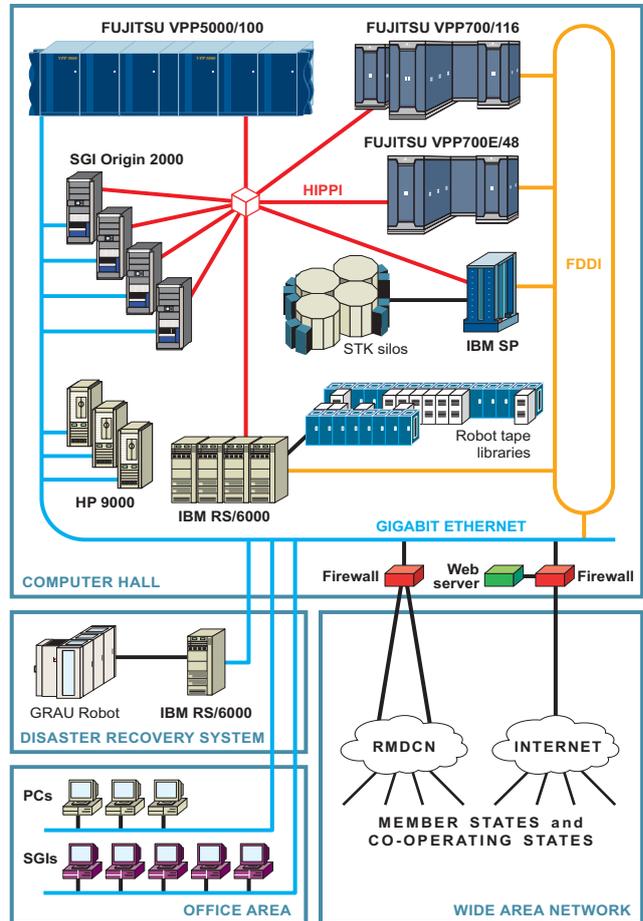


Figure 4 ECMWF computer systems in 2000.

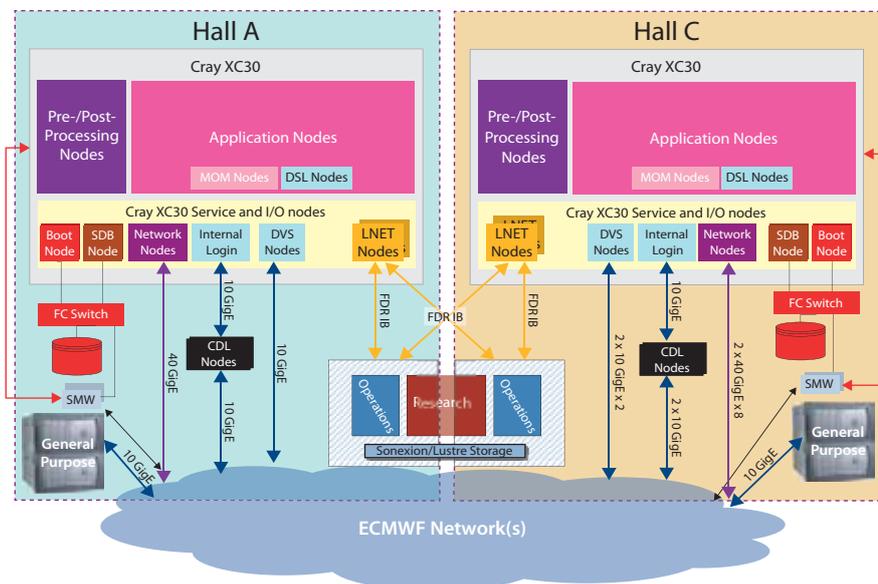


Figure 5 High-level diagram of the Cray HPCF showing major components.

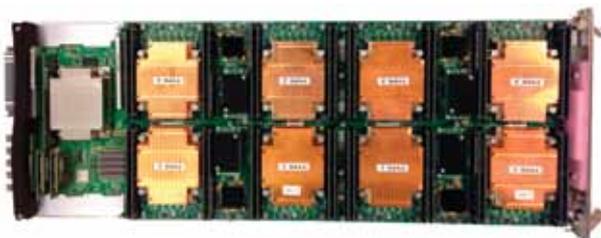
	Previous system	New system
Compute clusters	2	2
Peak performance (teraflops)	1,508	3,593
Sustained performance on ECMWF codes (teraflops)	70	200
EACH COMPUTE CLUSTER		
Compute nodes	768	3,505
Compute cores	23,424	84,120
Operating system	AIX 7.1	Cray CLE 5.2
High-performance interconnect	IBM HFI	Cray Aries
High Performance Parallel Storage (petabytes)	3.14	6.0
General-purpose storage (terabytes)	Not applicable	38
EACH COMPUTE NODE		
Memory in compute node (gibibytes)	64 (20 nodes with 256)	64 (60 x 128, 4 x 256)
Processor type	IBM POWER7	Intel E5-2697 v2 'Ivy Bridge'
CPU chips per node	4	2
Cores per CPU chip	8	12
Clock frequency (gigahertz)	3.8	2.7

**Table 1** Comparison of the current system's specification with that of the previous HPCF.

**From 2002 to 2014 – the IBM era**

In 2002, following a competitive tender exercise started in 2001, new machines from IBM replaced the Fujitsu systems. Two IBM Cluster 1600 systems, consisting of 30 p690 servers connected by an SP2 switch, produced their first operational forecasts on 4 March 2003. These machines differed from the Fujitsu systems in two important ways. First, they had no vector-processing capability, and second, they were high-volume production, standard computers linked by a special high-performance interconnect.

IBM systems provided ECMWF's computing service until the current Cray system replaced them in 2014. They took the sustained performance from the gigaflop range into the terascale, achieving 70 teraflops of sustained performance on the POWER7 system in 2012.



**Figure 6** An XC30 compute blade. On the main part of the blade you can see the heat sinks for the eight CPU chips of the four nodes. At the back of the blade is the Aries router.

**The current Cray HPCF**

The current Cray system is the result of a competitive procurement carried out in 2012 and 2013. This resulted in ECMWF awarding a two-phase service contract to Cray UK Ltd to supply and support this HPCF until mid-2018. The contract was signed on 24 June 2013.

The first-phase system started producing operational forecasts on 17 September 2014. The layout comprising two identical Cray XC30 systems continues ECMWF's successful design of having two self-sufficient clusters with their own storage, but with equal access to the high-performance working storage of the other cluster. This cross-connection of storage provides most of the benefits of having one very large system, but dual clusters add significantly to the resilience of the system. They enable flexibility in performing maintenance and upgrades and, when combined with separate resilient power and cooling systems, they provide protection against a wide range of possible failures. Figure 5 shows a high-level diagram of the system, the parts of which are described more fully below.

Each compute cluster weighs almost 45 metric tonnes and provides three times the sustained performance on ECMWF codes of the previous system. Table 1 compares the current system's specification with that of the previous HPCF.

**System description**

The bulk of the system consists of compute nodes, which each have two 12-core Intel processors. As shown in Figure 6, up to four compute nodes sit on a blade. Sixteen blades sit in a

chassis, and there are three chassis in a cabinet. This gives a maximum of 192 nodes or 4,608 processor cores per cabinet. The number of compute nodes in a cabinet will sometimes be less than the maximum since, as well as compute nodes, each cluster has a number of 'service nodes'. These have space for a PCI-Express card and are twice the size of a compute node so that only two fit on a blade. There are 19 cabinets in each of ECMWF's two clusters.

The Intel Xeon EP E5-2697 v2 'Ivy Bridge' used in the system was released in September 2013. It is an update of the original Xeon E5 'Sandy Bridge' processor following Intel's 'tick-tock' development strategy. In Intel terms, the Sandy Bridge processor was a 'tock', an introduction of a new microarchitecture and new features. Ivy Bridge is a 'tick' as it takes the architecture from the 'tock' and builds it using a new manufacturing technology, in this case a shrink to a 22-nanometre process that gives a greater transistor density on the chip. This allows more to be packed onto a chip while retaining the same overall energy consumption. For comparison, a 22-nanometre transistor is so small that about 4,000 can fit across the average width of a human hair.

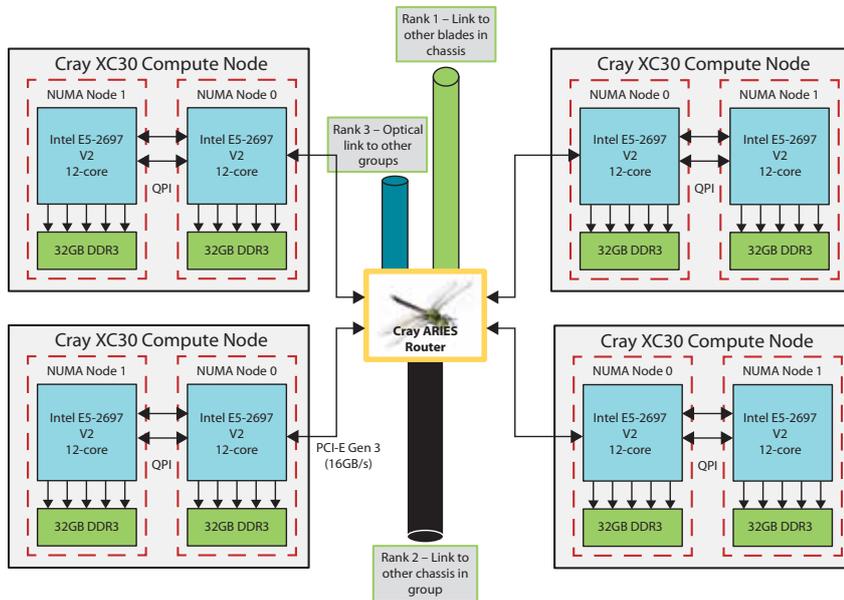
The peak performance of one processor core is around 22 gigaflops per second. This is more than the peak

performance of the Cray C90/16 system ECMWF had in 1996. There are more than 84,000 such cores in each of the XC30 clusters.

**High-performance interconnect**

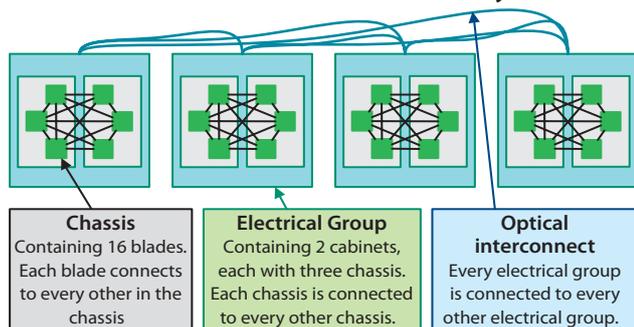
Connecting all of the processing power together is the Aries™ interconnect developed by Cray. This interconnect uses a 'dragonfly' topology, shown in Figure 7. The name stems from the shape of the dragonfly's body and wings, which represent local electrical connections on the one hand and longer-distance optical connections on the other.

Each blade in the system has a single Aries chip, and all the nodes on the blade connect to it via PCI-Express Gen3 links capable of a transfer rate of 16 gigabytes per second in each direction. Each Aries chip then connects via the chassis backplane to every other blade in the chassis. A chip has five other electrical connections, one to each chassis in a group of two cabinets. Cray describe this as an 'electrical group'. As shown in Figure 8, a further network level uses optical links to connect every electrical group to every other electrical group in the system. Electrical connections are cheaper than optical ones but are limited in length to a few metres. The Aries chip design also removes the need for external routers.



**Figure 7** A diagram of an XC30 compute blade. Each blade has four dual-socket nodes and an Aries router chip in a 'dragonfly' arrangement.

**Aries interconnect for an 8 cabinet Cray XC30**



**Figure 8** The Cray Aries interconnect has a large number of local electrical connections and a small number of longer-distance optical connections.

**Operating system**

The nodes of the Cray system are optimised for their particular function. The bulk of the nodes run in 'Extreme Scalability Mode'. In this mode, each node runs a stripped-down version of the Linux operating system. Reducing the number of operating system tasks running on a node to the minimum is a key element of providing a highly scalable environment for applications to run in. Any time spent not running the user's application is a waste. If the application is a tightly coupled parallel one, where results need to be exchanged with processes running on other nodes in order to progress, then delays caused by operating system interruptions on one node can cause other nodes to go idle waiting for input, increasing the runtime of the application.

The other two types of nodes in a Cray system are 'Service' nodes and 'Multiple Applications Multiple User (MAMU)' nodes.

MAMU or Pre-/Post-processing nodes (PPN) for ECMWF run full versions of the Linux operating system and allow more than one batch application to be run on a node. This mode is important: approximately four-fifths of the jobs run on the ECMWF systems require less than one full node to run on. These jobs are the preparation and clean-up for the main parallel jobs. While there are a huge number of these jobs, they account for less than 1% of the processing time offered by the system.

Service nodes are generally not visible to users. They perform a number of functions, such as connecting the compute system to the storage and the ECMWF networks, running the batch scheduler and monitoring and controlling the system as a whole.

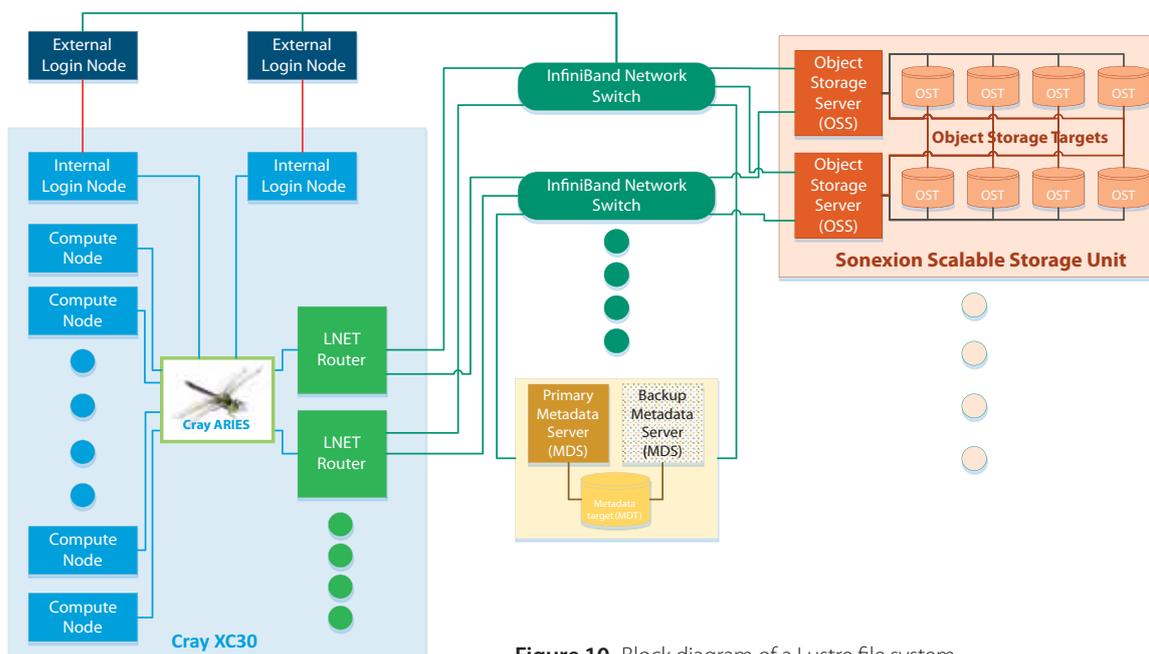
**High-performance storage**

High-performance working storage for the compute clusters is provided by Lustre file systems from integrated



**Figure 9** A Cray Sonexion storage appliance. The rack contains a metadata server and six storage building blocks.

Cray Sonexion appliances, shown in Figure 9. Each cluster has two main pools of storage, one for time-critical operational work, the other for research work. Segregating time-critical from research storage helps avoid conflicts between workloads and thus limits the variability of run times for time-critical work. While each cluster has its own high-performance working storage and is self-sufficient, it also has access, at equal performance, to the storage resources of the other cluster. This cross-mounting allows work to be flexibly run on either cluster, in effect making it in many regards a single system. There is a risk that an issue on one storage system can affect both compute clusters but, if necessary, the cross-mounts can be dropped to limit the impact of the instability to just one compute cluster. Each of our XC30 systems has about 6 petabytes of storage and offers up to 250 gigabytes per second of I/O bandwidth.



**Figure 10** Block diagram of a Lustre file system.

**Lustre file system**

The Lustre architecture has been developed in response to the requirement for a scalable file system for large supercomputers.

A Lustre file system, shown schematically in Figure 10, has several components. A metadata server (MDS) supports the directory hierarchy and information about individual files, such as who owns them and who can access them. The MDS stores its data on a metadata target (MDT), a small RAID array connected to the primary and backup server for resilience. Object Storage Servers (OSS) handle the actual data. Each OSS has a number of Object Storage Targets (OSTs) where the data is actually stored. When a Lustre client wants to do something like write a file, it contacts the MDS. This checks the authorisation of the user and that they have permission to access the file location. If successful, it sends back a list of OSTs from which the file can be read or to which it can be written. The client can then deal with the OSSs that host the OSTs. Since the OSTs are independent of each other, if the client has been given more than one, then it is possible for it to use them in parallel for higher performance. How many OSTs are given out for each file is a configurable parameter called 'stripe count', which can be set on a file, a directory or the entire file system. When more than one OST is used,

data is striped across all of them in chunks controlled by the stripe size parameter.

**General-purpose storage**

The second type of storage in the HPCF is the general-purpose storage provided by a NetApp Network File System. This storage provides space for home file systems and for storing applications. At 38 terabytes, its capacity is relatively small compared to the Lustre file systems, but the general-purpose storage is very reliable and offers a number of advanced features, such as file system snapshots and replication, which Lustre currently does not implement.

**Outlook**

Improving ECMWF's forecasts will require further advances in modelling the Earth's physical processes, the use of more observational data and finer model grid resolutions. This means that ECMWF will have to continue to develop its computing capability.

A key requirement for the future will be the scalability of the forecasting system to prepare it for the next generation of HPCFs: exascale facilities performing a billion billion calculations per second. Space is also an issue. ECMWF is looking for new premises to accommodate the kind of supercomputer centre it will need to maintain its leading position in global numerical weather prediction.

---

## ECMWF publications

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

**Technical Memoranda**

- 747 **Inness, A., A. Benedetti, J. Flemming, M. Parrington, J.W. Kaiser & S. Remy:** The ENSO signal in atmospheric composition fields: Emission driven versus dynamically induced changes. *March 2015*
- 746 **R.J. Hogan & A. Bozzo:** Mitigating surface temperature errors using approximate radiation updates. *February 2015*
- 745 **A. Inness, A.-M. Blechschmidt, I. Bouarar et al.:** Data assimilation of satellite retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS. *December 2014*
- 740 **S. Boussetta, G. Balsamo, E. Dutra, A. Beljaars & C. Albergel:** Analysis of surface albedo and Leaf Area Index from satellite observations and their impact on numerical weather prediction. *November 2014*
- 736 **Zuo, H., M.A. Balmaseda & K. Mogensen:** The ECMWF-MyOcean2 eddy-permitting ocean and sea-ice reanalysis ORAP5. Part 1: Implementation. *February 2015*

**ERA Report Series**

- 17 **de Boissésion, E., P. Laloyaux & M. Balmaseda:** Capturing tropical instability waves in the ECMWF Coupled Reanalysis System. *February 2015*

**EUMETSAT/ECMWF Fellowship Programme Research Report**

- 35 **Lonitz, K. & A. Geer:** New screening of cold-air outbreak regions used in 4D-Var all-sky assimilation. *February 2015*

## ECMWF Calendar 2015

Jun 1–5	NWP Training Course: Advanced Numerical Methods for Earth-System Modelling
Jun 8–10	Using ECMWF's Forecasts (UEF2015)
Jun 10–12	OpenIFS Workshop
Jun 15–19	Workshop on Advancing Training and Teaching in Numerical Weather Prediction
Jun 25–26	Council
Jun 29–Jul 2	Copernicus Climate Observation Requirements Workshop
Aug 31–Sep 4	Annual Seminar
Sep 28–Oct 2	Visualisation Week: <ul style="list-style-type: none"> <li>• Workshop on Meteorological Operational Systems</li> <li>• European Working Group on Operational Meteorological Workstations (EGOWS)</li> <li>• RMetS Seminar on Visualisation in Meteorology</li> <li>• OGC MetOcean Interoperability Session</li> </ul>

Oct 5–7	Training Course: Use and Interpretation of ECMWF Products
Oct 12–14	Scientific Advisory Committee
Oct 15–16	Technical Advisory Committee
Oct 21	Policy Advisory Committee
Oct 22–23	Finance Committee
Oct 26	Advisory Committee for Co-operating States (ACCS)
Nov 2–5	Workshop on Sub-Seasonal Predictability
Dec 8–9	Council

## Contact information

ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, UK  
 Telephone National 0118 949 9000  
 Telephone International +44 118 949 9000  
 Fax +44 118 986 9450

ECMWF's public website <http://www.ecmwf.int/>  
 E-mail: The e-mail address of an individual at the Centre is firstinitial.lastname@ecmwf.int. For double-barrelled names use a hyphen (e.g. j-n.name-name@ecmwf.int).

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

## Index of newsletter articles

This is a selection of articles published in the *ECMWF Newsletter* series during recent years.

Articles are arranged in date order within each subject category.

Articles can be accessed on the ECMWF public website – <http://www.ecmwf.int/en/research/publications>

	No.	Date	Page		No.	Date	Page
<b>NEWS</b>							
Work on Copernicus Climate Change Service under way	143	Spring 2015	2	Exploring the potential of using satellite data assimilation in hydrological forecasting	141	Autumn 2014	10
El Niño set to strengthen but longer-term trend uncertain	143	Spring 2015	3	New Cray High-Performance Computing Facility	141	Autumn 2014	11
Upbeat mood as MACC project draws to a close	143	Spring 2015	4	Anton Beljaars elected as an AMS Fellow	141	Autumn 2014	12
Forecasts for US east coast snow storm in January 2015	143	Spring 2015	6	Second OpenIFS user meeting at Stockholm University	140	Summer 2014	2
New training module for Metview software	143	Spring 2015	7	Recognition of the contributions to meteorology by Florence Rabier and Tim Palmer	140	Summer 2014	4
Benefits of statistical post-processing	143	Spring 2015	8	Forecasting the severe flooding in the Balkans	140	Summer 2014	5
Modelling the Quasi-Biennial Oscillation	143	Spring 2015	8	10th Anniversary of HEPEX	140	Summer 2014	6
Warm conditions continue from 2014 into 2015	143	Spring 2015	9	ECMWF revisits the meteorology of the D-Day period	140	Summer 2014	7
The role of hindcast length in assessing seasonal climate predictability	143	Spring 2015	11	Use of GPS-RO in operational NWP and reanalysis applications	140	Summer 2014	8
Stochastic workshop explores simulation of forecast model uncertainty	143	Spring 2015	12	Launch of a new fellowship programme	140	Summer 2014	9
Piotr Smolarkiewicz granted Poland's top academic title	143	Spring 2015	13	Use of BUFR radiosonde and surface observations	140	Summer 2014	10
Annual Seminar proceedings published	143	Spring 2015	13	Working together to address weather forecasting challenges	140	Summer 2014	11
ECMWF Copernicus Services – Open for Business	142	Winter 2014/15	2	Wave experts meet at ECMWF	140	Summer 2014	12
Additional clustering time-periods available for dissemination and in MARS	142	Winter 2014/15	3	Interview with a departing graduate trainee	139	Spring 2014	2
Forecast performance 2014	142	Winter 2014/15	4	Enhancing the biomass-burning emissions database: release of a new version of GFAS	139	Spring 2014	3
Membership of the Scientific Advisory Committee	142	Winter 2014/15	5	Presentation of maps for the new website	139	Spring 2014	5
Serbia becomes ECMWF's 21st Member State	142	Winter 2014/15	6	Start of the ERA-CLIM2 project	139	Spring 2014	6
Flow-dependent background error covariances in 4DVAR	142	Winter 2014/15	7	TIGGE-LAM improves regional ensemble forecasts	139	Spring 2014	7
Forecasts for a fatal blizzard in Nepal in October 2014	142	Winter 2014/15	8	Scalability programme at ECMWF	139	Spring 2014	8
New blog for software developers	142	Winter 2014/15	9	Metview's interface to 3D interactive graphics	139	Spring 2014	9
Recognition of ECMWF's role in THORPEX	142	Winter 2014/15	9	Migrating the RMDCN	139	Spring 2014	10
Update on migration to BUFR for radiosonde, surface and aircraft observations at ECMWF	142	Winter 2014/15	10	Top class training	139	Spring 2014	12
Sharing knowledge about climate data	142	Winter 2014/15	11	Global partnership for enhanced resilience to flood risk	139	Spring 2014	12
Copernicus Climate Change and Atmosphere Monitoring Services	141	Autumn 2014	2	Copernicus Climate Change Service Workshop	139	Spring 2014	13
Recent cases of severe convective storms in Europe	141	Autumn 2014	3	Metview's 20 <sup>th</sup> anniversary	138	Winter 2013/14	2
Licensing ECMWF products	141	Autumn 2014	5	New model cycle 40r1	138	Winter 2013/14	3
Closing the GRIB/NetCDF gap	141	Autumn 2014	6	ECMWF's contribution to GEO	138	Winter 2013/14	11
Peter Janssen awarded the EGU Fridtjof Nansen Medal for 2015	141	Autumn 2014	7	Using Earth System science at ECMWF	137	Autumn 2013	6
MACC-III forecasts the impact of Bardarbunga volcanic SO <sub>2</sub>	141	Autumn 2014	8	Floods in Central Europe in June 2013	136	Summer 2013	9
ERA-20C goes public for 1900–2010	141	Autumn 2014	9				
Use of high-performance computing in meteorology	141	Autumn 2014	10	<b>VIEWPOINT</b>			
				Decisions, decisions...!	141	Autumn 2014	12
				Using ECMWF's Forecasts: a forum to discuss the use of ECMWF data and products	136	Summer 2013	12
				Describing ECMWF's forecasts and forecasting system	133	Autumn 2012	11
				<b>COMPUTING</b>			
				Supercomputing at ECMWF	143	Spring 2015	32

	No.	Date	Page		No.	Date	Page
SAPP: a new scalable acquisition and pre-processing system at ECMWF	140	Summer 2014	37	Breakthrough in forecasting equilibrium and non-equilibrium convection	136	Summer 2013	15
Metview's new user interface	140	Summer 2014	42	Convection and waves on small planets and the real Earth	135	Spring 2013	14
GPU based interactive 3D visualization of ECMWF ensemble forecasts	138	Winter 2013/14	34	Global, non-hydrostatic, convection-permitting, medium-range forecasts: progress and challenges	133	Autumn 2012	17
RMDCN – Next Generation	134	Winter 2012/13	38	Development of cloud condensate background errors	129	Autumn 2011	13
A new trajectory interface in Metview 4	131	Spring 2012	31	<b>PROBABILISTIC FORECASTING &amp; MARINE ASPECTS</b>			
A new framework to handle ODB in Metview 4	130	Winter 2011/12	31	Have ECMWF monthly forecasts been improving?	138	Winter 2013/14	18
Managing work flows with ecFlow	129	Autumn 2011	30	Closer together: coupling the wave and ocean models	135	Spring 2013	6
<b>METEOROLOGY</b>				20 years of ensemble prediction at ECMWF	134	Winter 2012/13	16
<b>OBSERVATIONS &amp; ASSIMILATION</b>				Representing model uncertainty: stochastic parametrizations at ECMWF	129	Autumn 2011	19
Snow data assimilation at ECMWF	143	Spring 2015	26	<b>METEOROLOGICAL APPLICATIONS &amp; STUDIES</b>			
Assimilation of cloud radar and lidar observations towards EarthCARE	142	Winter 2014/15	17	The skill of ECMWF cloudiness forecasts	143	Spring 2015	14
The direct assimilation of principal components of IASI spectra	142	Winter 2014/15	23	Calibration of ECMWF forecasts	142	Winter 2014/15	12
Automatic checking of observations at ECMWF	140	Summer 2014	21	Twenty-five years of IFS/ARPEGE	141	Autumn 2014	22
All-sky assimilation of microwave humidity sounders	140	Summer 2014	25	Potential to use seasonal climate forecasts to plan malaria intervention strategies in Africa	140	Summer 2014	15
Climate reanalysis	139	Spring 2014	15	Predictability of the cold drops based on ECMWF's forecasts over Europe	140	Summer 2014	32
Ten years of ENVISAT data at ECMWF	138	Winter 2013/14	13	Windstorms in northwest Europe in late 2013	139	Spring 2014	22
Impact of the Metop satellites in the ECMWF system	137	Autumn 2013	9	Statistical evaluation of ECMWF extreme wind forecasts	139	Spring 2014	29
Ocean Reanalyses Intercomparison Project (ORA-IP)	137	Autumn 2013	11	Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector	139	Spring 2014	34
The expected NWP impact of Aeolus wind observations	137	Autumn 2013	23	iCOLT – Seasonal forecasts of crop irrigation needs at ARPA-SIMC	138	Winter 2013/14	30
Winds of change in the use of Atmospheric Motion Vectors in the ECMWF system	136	Summer 2013	23	Forecast performance 2013	137	Autumn 2013	13
New microwave and infrared data from the S-NPP satellite	136	Summer 2013	28	An evaluation of recent performance of ECMWF's forecasts	137	Autumn 2013	15
Scaling of GNSS radio occultation impact with observation number using an ensemble of data assimilations	135	Spring 2013	20	Cold spell prediction beyond a week: extreme snowfall events in February 2012 in Italy	136	Summer 2013	31
ECMWF soil moisture validation activities	133	Autumn 2012	23	The new MACC-II CO2 forecast	135	Spring 2013	8
Forecast sensitivity to observation error variance	133	Autumn 2012	30	Forecast performance 2012	134	Winter 2012/13	11
Use of EDA-based background error variances in 4D-Var	130	Winter 2011/12	24	Teaching with OpenIFS at Stockholm University: leading the learning experience	134	Winter 2012/13	12
Observation errors and their correlations for satellite radiances	128	Summer 2011	17	Uncertainty in tropical winds	134	Winter 2012/13	33
Development of cloud condensate background errors	128	Summer 2011	23	Monitoring and forecasting the 2010-11 drought in the Horn of Africa	131	Spring 2012	9
Use of SMOS data at ECMWF	127	Spring 2011	23	Characteristics of occasional poor medium-range forecasts for Europe	131	Spring 2012	11
<b>FORECAST MODEL</b>				A case study of occasional poor medium-range forecasts for Europe	131	Spring 2012	16
Atmospheric composition in ECMWF's Integrated Forecasting System	143	Spring 2015	20	The European Flood Awareness System (EFAS) at ECMWF: towards operational implementation	131	Spring 2012	25
Towards predicting high-impact freezing rain events	141	Autumn 2014	15	New tropical cyclone products on the web	130	Winter 2011/12	17
Improving ECMWF forecasts of sudden stratospheric warmings	141	Autumn 2014	30	Increasing trust in medium-range weather forecasts	129	Autumn 2011	8
Improving the representation of stable boundary layers	138	Winter 2013/14	24	Use of ECMWF's ensemble vertical profiles at the Hungarian Meteorological Service	129	Autumn 2011	25
Interactive lakes in the Integrated Forecasting System	137	Autumn 2013	30	Developments in precipitation verification	128	Summer 2011	12
Effective spectral resolution of ECMWF atmospheric forecast models	137	Autumn 2013	19				



Newsletter | Number 143 – Spring 2015  
European Centre for Medium-Range Weather Forecasts  
[www.ecmwf.int](http://www.ecmwf.int)