ECMWF Newsletter

Number 140 – Summer 2014

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

www.ecmwf.int/publications/newsletter

CONTENTS

PUBLICATION POLICY

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 The *ECMWF Newsletter* is not peer-reviewed.

Editor: Bob Riddaway

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 Bob.Riddaway@ecmwf.int

Guidance about submitting an article is available at www.ecmwf.int/publications/newsletter/guidance.pdf

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

Weather science

Weather science and forecasting have developed hugely over recent decades and expectations are high for even greater advances in the years to come. We can look forward to new sources of atmospheric observations, faster supercomputers and further advances in the science as well as the transition to a fully Earth-system approach. The "weather enterprise" is emerging as a diverse partnership between public, private and academic-sector meteorologists, forecasters, computing experts, social scientists and application developers from many countries. The importance of harnessing the talents and enthusiasm of the next generation is crucial.

It is with this backdrop that the first ever World Weather Open Science Conference (WWOSC), with the theme "The weather: what's the outlook?", took place in Montreal, Canada, from 16 to 21 August 2014. More than 1,000 participants from over 50 countries discussed the development of weather science and its many applications, making WWOSC a resounding success.

Participants (including many from ECMWF) investigated opportunities for achieving major breakthroughs in weather science at the same pace as in the last 20 to 30 years, if not faster. Frequently the focus was on the prediction of extreme weather hazards. The goal of achieving seamless predictions – integrated modelling systems for all time scales, from a few minutes to weeks, months and years ahead – received strong support. The integration of weather and climate science was confirmed as an important objective for the next decade, as was the international cooperation between nations and between scientists, practitioners and users.

The Conference explored applications and uses of weather forecasts, focussing on social sciences and economic impacts of weather in many sectors such as energy, health and insurance. A large group attending the conference were early-career scientists who played a major and active part in the meeting. They discussed the possibility of forming an association for early-career weather, climate and environment scientists dedicated to working together to develop weather science and its applications.

The WMO World Weather Research Programme is a key vehicle for internationally-collaborative research projects that helps address the issues raised at the Conference. ECMWF staff, along with many colleagues around the world, will play a major part in three new programmes – polar prediction, subseasonal-to-seasonal and high-impact weather.

What is clear today is that no one individual, group or nation can go it alone – meteorology is a huge global enterprise that demands co-operation and collaboration if society is to have advance warning of extreme weather so that people can get out of harm's way.

Alan Thorpe

End of Conference Statement

"It now seems apparent that over the next 20 years, forecasters are going to need Earth-system modelling. Today's weather forecasts and climate predictions are likely to evolve towards seamless weather-climate-impacts forecasting. In addition to the atmosphere and oceans, they will integrate increasingly accurate information on topography, land-use change, vegetation, rivers, lakes, clouds and socio-economic trends to provide user-specific decision-support services that will touch almost every part of our lives."

Michel Beland and Alan Thorpe, Conference Co-Chairs

Second OpenIFS user meeting at Stockholm University

ABDEL HANNACHI (Stockholm University), **GLENN CARVER** (ECMWF)

The OpenIFS project at ECMWF is in its third year since it was initiated in December 2011. The aim of the project is to provide a long-term, sustainable and easy-to-use version of the operational Integrated Forecasting System (IFS) of ECMWF for education and research at universities and research centres. The OpenIFS model is currently based on IFS cycle 38r1 and includes all the forecast capability but without any data assimilation. The single column model (SCM) version of IFS is also made available by the OpenIFS project and is an ideal tool for teaching and research in physical parametrizations. Information about OpenIFS can be found at http://www.ecmwf.int/en/research/projects/openifs.

The first use of OpenIFS outside ECMWF was at the Meteorological Institute of Stockholm University (MISU) in November 2012 by masters students as part of their numerical weather prediction (NWP) MSc course. This is described in more detail in *Hannachi et al*. (2012, *ECMWF Newsletter No 134*, 12–15). The idea then arose of organising a meeting to bring together OpenIFS users to learn and share experience.

The first OpenIFS user meeting was organised by Victoria Sinclair and Heikki Järvinen and took place at the University of Helsinki, Finland, in June 2013 (see *Carver et al*., 2013, *ECMWF Newsletter No 136*, 4). The second meeting took place in Stockholm, hosted by MISU from 10 to 12 June 2014. Research at MISU includes atmospheric dynamics, atmospheric physics, chemical meteorology and physical oceanography. The department was founded by Carl-Gustaf Rossby when he returned to his native country Sweden from the USA in 1947. Nowadays, it is an internationally respected research centre with 35 scientists and over 30 PhD students. It is part of the Bolin Centre for Climate Research at Stockholm University.

The programme of the meeting at MISU was composed of lectures and practicals. The first day was mostly dedicated to lectures and the next two days were devoted to practicals using Metview and the SCM.

Eleven lectures were presented during the first day, led by keynote talks by Erland Källén on '*Progress and prospects for the ECMWF forecasting system*' and Peter Bechtold on '*It is spring/summer not only in the IFS: from the large waves to the daily showers*'. These were followed by presentations covering various ways in which OpenIFS and IFS have been used in research and teaching – see the OpenIFS website for more information about the meeting.

Participants in the OpenIFS user meeting. There were 36 participants from European universities in Helsinki, Oxford, Stockholm, Uppsala, Minsk, Sofia, Dublin and Munich, the Swedish Meteorological and Hydrological Institute (SMHI) and ECMWF.

The second day was devoted to practicals on convection based on two case studies presented by Peter Bechtold. The first case study was based on a tornado that hit Little Rock, Arkansas, 27 April 2014 at 7.00 pm local time. Two 30-hour forecasts at T255L137 were prepared: a 'control' run and a 'perturbed' run in which deep convection was turned off. The objective of the exercise was to identify areas of threat, understand the synoptic situation and understand the role of the large-scale and convection adjustment. The role of deep convection was particularly noticeable in identifying the areas of threat. The second case study focussed on the role of the convective diurnal cycle correction recently introduced into Cy40r1 of the IFS (*Bechtold et al*., 2013, *ECMWF Newsletter No. 136*, 15–22) to forecast convection processes over Africa. Metview was used to change areas of focus and compare with Amazonia to identify the role played by moisture availability.

The third day was mostly dedicated to practicals using the SCM led by Filip Váňa. The two cases were repeated by extracting initial profiles and forcing tendencies from operational IFS forecasts. Particular focus was on the effect of the convective diurnal cycle, convection time-scale adjustment and sensitivity to the entrainment rate. The diurnal cycle was found, as expected, to be one of the most important parameters for convection over Central Africa.

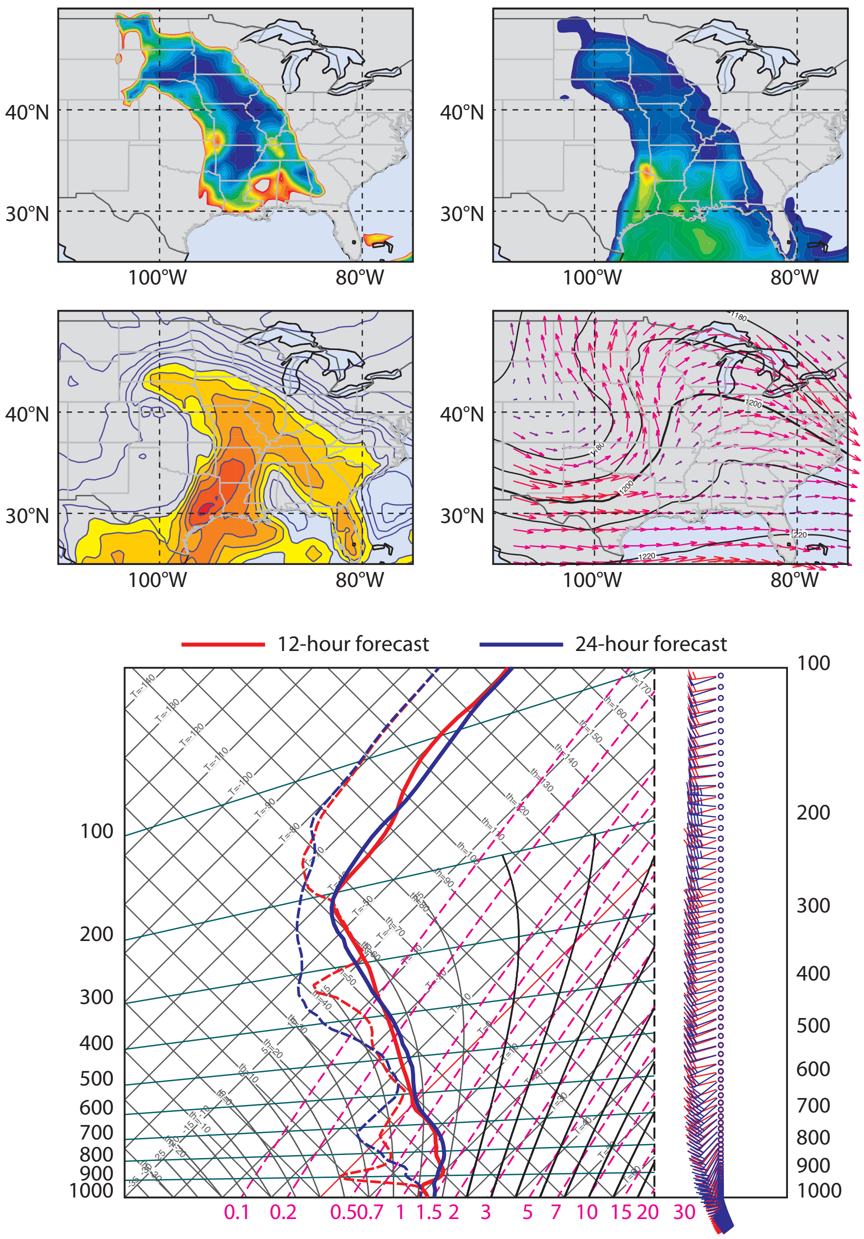
A key feature of the meeting was the use of a Linux virtual machine. This is an effective way of delivering remote workshop practicals involving real models, data and analysis tools. It was successfully first used at the University of Helsinki OpenIFS user workshop in 2013. The virtual machine allowed for the packaging of the exercises, binary executables and supporting files at ECMWF before transferring to Stockholm by ftp. A virtual machine 'player' was installed on each classroom computer which the participants used to access Metview, the SCM and accompanying data to complete the exercises.

Towards the end of the last day an extended discussion was held between the participants and ECMWF staff about OpenIFS and future perspectives. This was well received by the participants. They enjoyed the experience of using Metview and running the SCM and, above all, discussing the results of the exercises with ECMWF staff. There was an eagerness to continue these types of meetings with perhaps an extension of the meeting duration, with a wider scope, to allow more time for exercises and more idealised experiments. The issue of data assimilation was also raised.

The success of the workshop was in part due to the support provided at the University of Stockholm by Michael Tjernström, Abdel Hannachi and other members of MISU staff. From ECMWF the preparation of the practical exercises and support of Peter Bechtold, Glenn Carver, Erland Källén, Sándor Kertész and Filip Váňa ensured the workshop was scientifically interesting and informative.

Feedback from the workshop participants was very positive and we look forward to the next user workshop in 2015 which will be organised by the OpenIFS users at the University of Oxford, UK, to be held at ECMWF.

Practical exercises. The exercises were a key component of the workshop. Prepared case studies were used by the participants to consider a number of questions on the role of convection in the formation of a tornado in the USA and the convective diurnal cycle over Africa.



Case study of a tornado that hit Little Rock, Arkansas, 27 April 2014. Top four panels: The Metview plot used to identify the area of threat for the tornado case using 21-hour forecasts from 00 UTC on 27 April for the T255L137 control (top left: Convective Inhibition (CIN); top right: CAPE; bottom left: equivalent potential temperature at 850 hPa; bottom right: wind and geopotential at 200 hPa). High values are in red. Bottom panel: Tephigrams valid at two different forecast times at a location in the area of deep convection overlaid in Metview. The case study was analysed using Metview macros prepared by Sándor Kertész.

Recognition of the contributions to meteorology by Florence Rabier and Tim Palmer

BOB RIDDAWAY

Florence Rabier

On 9 July an event was held at ECMWF to celebrate Florence Rabier becoming a Chevalier de la Légion d'Honneur – France's highest honour.

The award was presented to Florence by Jean-Claude André from the Academie des Sciences in recognition of her outstanding contributions to meteorology.

As well as carrying out fundamental research into four-dimensional variational data assimilation (4D-Var) and use of satellite data, Florence has led the way in bringing research through to operations. Also she has been involved in a wide range of international activities and is committed to the development of young scientists and communicating meteorological science to the general public.

After taking up a post as a meteorological engineer at Météo-France, Florence carried out research into 4D-Var and showed that this technique was a critical step forward for improving meteorological forecasts. This was followed by six years at ECMWF studying the sensitivity of forecasts to initial errors using adjoint methods and the implementation of 4D-Var in the forecasting system.

Florence returned to Météo-France in 1998 and joined GMAP (Groupe de Modélisation et d'Assimilation pour la Prévision), where she took a leading role in introducing new satellite data into the NWP models. In addition, she co-chaired the international working group for observational strategy and data assimilation for the WMO THORPEX programme. In 2010 Florence became deputy head of GMAP with responsibilities which included transferring research results into operations and adapting the assimilation scheme to take advantage of new satellite data.

In 2013 Florence was appointed as Director of Forecasts at ECMWF with responsibility for the Forecast Department. This has a strong user-focus, with meteorological input to forecast production, forecast evaluation and diagnostics, forecast products and applications, software development, and catalogue and data services.

Tim Palmer

Tim Palmer has been awarded the Dirac Medal by the Institute of Physics for the development of probabilistic weather and climate prediction systems. The award is made for outstanding contributions to theoretical (including mathematical and computational) physics.

For many years Tim was Head of the Probability Forecasting and Diagnostics Division at ECMWF where he led the development of the ECMWF ensemble forecasts and research into extended-range forecasting. In 2012 he became the Royal Society Research Professor in Climate Physics at the University of Oxford, but has maintained close collaboration with ECMWF. Recently Tim was appointed as an ECMWF Fellow because he is carrying out pioneering work in areas relevant to the strategic goals of ECMWF.

Tim and his team at ECMWF have led a revolution in the fields of weather and climate by establishing a physical basis for understanding nonlinear error growth in prediction models and for developing practical ways of estimating flow-dependent predictability. They have challenged old ideas and have changed the way that weather and climate are viewed both by the public, associates in the same field, and scientists in other disciplines.

Tim's work in weather and climate predictions is a blend of theoretical insight and practicality. Based on his insights into chaotic behaviour of fluids, he and his team have created a system that gives an indication of the probability of drought, flood, tropical cyclones or hydro-meteorological hazards in general. For example, ECMWF's probabilistic predictions have been used in Bangladesh where, for the first time, societies can anticipate slow-rise, long-lived floods.

Normally, the societal benefits of ideas take a long time to permeate to the practical level. But Tim, having practical applications in mind, has formed an almost immediate link between his theoretical insights and practical applications.

Forecasting the severe flooding in the Balkans

LINUS MAGNUSSON,
FREDRIK WETTERHALL,
FLORIAN PAPPENBERGER,
IVAN TSONEVSKY

Weeks of wet conditions followed by exceptionally intense rain from 13 to 16 May 2014 led to disastrous and wide-spread flooding in the Balkans, in particular in Bosnia-Herzegovina and Serbia. Critical flooding was also reported in other regions including southern Poland, Slovakia, and the Czech Republic. The event in Bosnia-Herzegovina and Serbia was reported to be the worst for more than 100 years, with more than 80 fatalities reported from the flood-affected regions and a significant loss of livestock. More than a million inhabitants were estimated to have been affected by this flood event and both Bosnia-Herzegovina and Serbia activated the EU Community Civil Protection Mechanism for help.

During the month leading up to the main precipitation event, precipitation had already been double the average. The main precipitation event itself lasted about three days, beginning on 13 May. On that day there was an eastward-moving cold front. Thereafter an upper-level, cut-off low formed, together with a deep surface low over south-eastern Europe. Later the low moved north-westwards, bringing further heavy rain over Serbia as easterly winds on its northern flank hit the upslope of a mountainous region between Serbia and Bosnia-Herzegovina. Finally, on the 16th the low pressure system started to fill. At the end of the episode heavy rain was also observed in Slovakia and Austria.

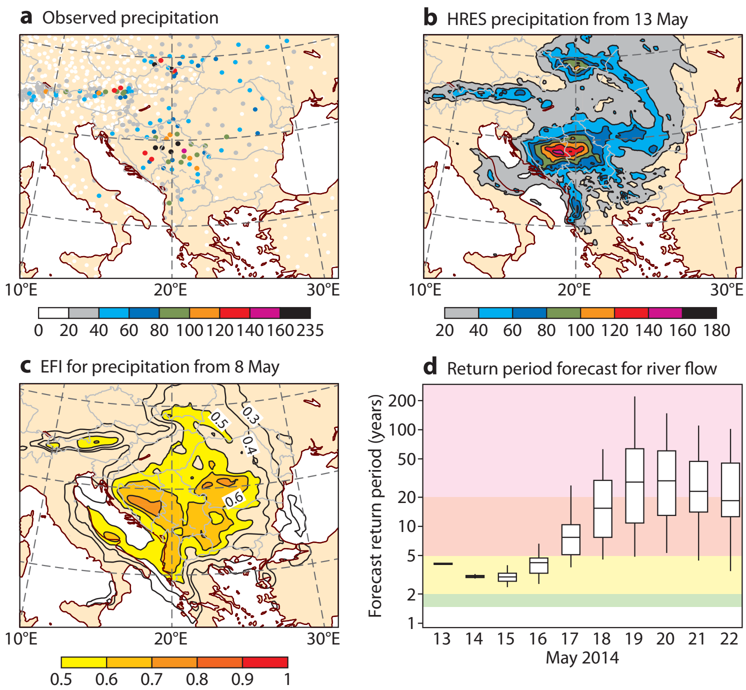
Panel a of the figure shows the observed three-day accumulation of precipitation starting at 06 UTC on 13 May. The highest values were in a band from Tuzla in the west to Belgrade in the east, with more than 160 mm over three days. The four most extreme reports were Tuzla (264 mm), Loznica (213 mm), Valjevo (190 mm) and Belgrade (174 mm).

Panel b shows the three-day accumulation of precipitation from the high-resolution forecast (HRES) initialised at 00 UTC on 13 May. The location of the precipitation is good. The values are underestimated in the worst-affected area, although forecast totals did reach 140 mm. An underestimation of the largest precipitation totals by HRES was also seen for the flooding in central Europe at the beginning of June 2013 (see *ECMWF Newsletter No. 136*).

For severe weather events, having accurate predictions well in advance can be very valuable. The Extreme Forecast Index (EFI) compares the ensemble forecast (ENS) distribution with the model climate distribution derived from a set of re-forecast data. Panel c shows the EFI for three-day precipitation accumulation starting at 00 UTC on 13 May from the forecast issued on 00 UTC on 8 May.

Already at this very early stage, the ENS suggested there was likely to be extreme precipitation in the Balkans. This early signal of extreme precipitation was connected to a relatively confident forecast of the aforementioned cut-off low.

The European Flood Awareness System (EFAS) uses ECMWF precipitation forecasts to drive its flood forecasting model. EFAS produces medium-range probabilistic flood forecasts for Europe and the computations are executed at ECMWF as part of the Copernicus services. The forecast from the 9th (which was the earliest forecast that covered the day of peak flow) already had ensemble members indicating a risk of a very severe event (not shown). Panel d shows the forecast from the 13th for the return period of the stream flow for the river Sava for a point close to Belgrade. This forecast had more than 50% probability for an event with



Observations and forecasts covering the period of 13 to 16 May 2014. (a) Observed precipitation between 06 UTC on 13 May and 00 UTC on 16 May and (b) HRES for precipitation accumulated for the same three-day period from 00 UTC on 13 May. (c) Extreme Forecast Index for three-day precipitation accumulation valid for 00 UTC on 13 May to 00 UTC on 16 May issued on 00 UTC on 8 May. (d) Return period forecast for the flow in river Sava for a point close to Belgrade from the forecast initialised on 13 May.

a return period of greater than 20 years. You can also see in the figure that even on the 13th the river flow was already higher than normal, due to the wet conditions in the region during previous weeks.

To summarize, the EFI and the EFAS flood forecast system gave an early warning of the Balkan flooding. This was based on a well-predicted cut-off low that gave persistent precipitation over the region for three days coupled with already-saturated soil from precipitation during the previous weeks. Precipitation amounts from HRES were generally too low for the areas where the largest totals were measured (although values above 140 mm were predicted) and the precipitation amounts in the ENS were in general less than in HRES.

The problem of extreme precipitation amounts being under-forecast was also identified for the flooding in central Europe in 2013 and is presented in *ECMWF Tech. Memo. No. 723*. In the investigation of that case, a number of experiments were conducted with different resolutions and changes to the model physics. The conclusion is that both improved physics and increased resolution could lead to better representation of precipitation extremes. In the next model cycle (to be introduced operationally later in 2014) the cloud physics will change and this should provide a general improvement in the precipitation extremes. The next increase in horizontal resolution is planned for 2015.

10th Anniversary of HEPEX

FREDRIK WETTERHALL, FLORIAN PAPPENBERGER, ROBERTO BUIZZA

In 2004, Philippe Bougeault, Tony Hollingsworth and Roberto Buizza from ECMWF and John Schaake from NOAA initiated the Hydrological Ensemble Prediction Experiment (HEPEX). Ten years on, it has developed into a broad network of enthusiastic and devoted scientists, forecasting groups, and end-users helping to find unique and novel solutions to challenges in hydrological forecasting.

ECMWF continues to take a leading role in HEPEX: Florian Pappenberger was co-chair between 2011 and 2014, a position that Fredrik Wetterhall has now taken on for the next three years. Ever since the start, ECMWF has been active in the development and application of hydrological ensemble predictions. Good examples of ECMWF's involvement are the now operational European Flood Awareness System (EFAS) and the pre-operational global system GloFAS. Both systems have relied on ECMWF's ensemble forecasts from the start.

Hosted by the National Centers for Environmental Prediction (NCEP/NOAA) at their impressive new facilities in Maryland, USA, the 10th anniversary workshop opened with an address by Roberto Buizza, first co-chair of HEPEX from 2004 to 2007. This was followed by presentations about the state-of-the-art techniques and uses of ensemble predictions for floods, droughts and water resource management. During the workshop scientists and end-users had many opportunities to share their views and experiences.

The workshop reviewed what has been achieved in the decade of hydrological probabilistic forecasting. Participants considered how we have learned to cope with the wealth of meteorological and hydrological data and which gaps still exist. A further question addressed was how to produce and communicate probabilistic information in a way that effectively helps decision-makers to take the best decision.

More information about the mission and activities of HEPEX is available at http://www.hepex.org/.

Participants at the HEPEX 10th anniversary workshop. The workshop brought together more than 70 participants from around the world, with a further 30 following the meeting online via a webcast and Twitter. The workshop was held at the NCEP/NOAA Center for Weather and Climate Prediction in Maryland, USA.

ECMWF revisits the meteorology of the D-Day period

ADRIAN SIMMONS, HANS HERSBACH, PAUL POLI, JEAN BIDLOT, DICK DEE

ECMWF and several partners are building records of atmospheric conditions over the 20th century by applying the techniques of modern numerical weather analysis and forecasting to historical surface and upper-air observations. This technique is known as reanalysis.

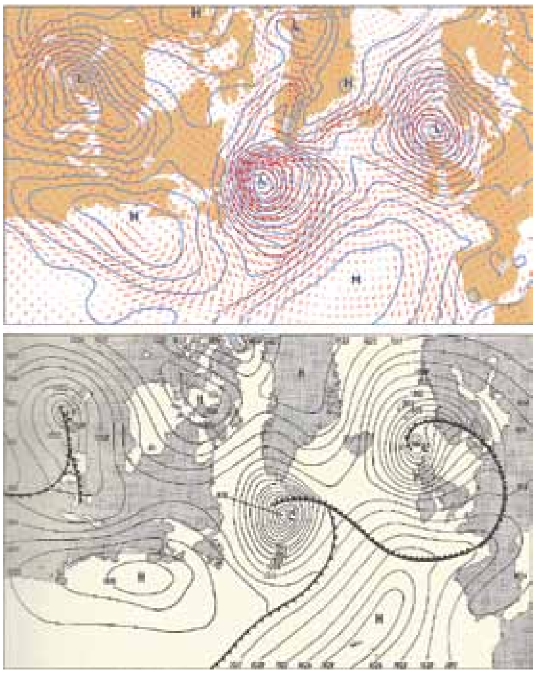
Weather forecasts critical to the success of the D-Day landings of 6 June 1944 were made on the nights of 3/4 and 4/5 June 1944.

- Forecasts for conditions on 5 June made on the evening of 3 June and confirmed early in the morning of 4 June.
- Forecasts for conditions on 6 June made on the evening of 4 June and confirmed early in the morning of 5 June.

The first of these two forecasts, presented to the Supreme Commander of the Allied Forces in Europe, General Eisenhower, by his meteorological advisor, Group Captain J.M. Stagg, led to the postponement of the invasion planned for 5 June; the second enabled Eisenhower to make the decision to go ahead the following day.

Ten years ago, ECMWF had completed ERA-40, a reanalysis that provided a three-hourly record of conditions back to 1957, when the atmospheric observing system was enhanced in preparation for the International Geophysical Year. To mark the retirement of the then ECMWF Director, David Burridge, who was born in June 1944, and recognising the challenge and interest in reanalysing the meteorological conditions of the D-Day period as its sixtieth anniversary approached, we stepped back a further few years, and applied a slightly updated version of the ERA-40 data assimilation system to reconstruct the weather of June 1944. This was not a comprehensive study; instead we simply made use of existing tools and a quite limited number of readily available observations. Although far from perfect, the results agreed in the broadest terms with the historical record.

Ten years on, a new opportunity to re-examine the meteorology and sea-state conditions for June 1944 presents itself. ECMWF has recently completed ERA-20C, a reanalysis for the period from 1900 to 2010 assimilating surface pressure and marine surface wind observations. This follows the first reanalysis of this type, the NOAA-CIRES 20th Century Reanalysis. We are also currently producing a second reanalysis that builds on ERA-20C from 1939 onwards by including upper-air observations. This second reanalysis provides the basis for the discussion of the analyses and forecasts for June 1944 that is provided at: http://www.ecmwf.int/en/research/projects/era-clim/d-day-analyses.



Surface pressure charts for 5 June 1944. Top: New ERA analysis of surface pressure and ten-metre wind for 12 UTC. Bottom: surface-pressure analysis reproduced from Stagg's book *Forecast for Overlord* (W.W. Norton & Co., Inc., 1971) for 13 UTC. The contour interval for surface pressure is 3 hPa.

From our perspective, the expectations we had at the outset of our second reanalysis of the D-Day period, covering not only the landings themselves but also a persistent major storm two weeks later, have been met or exceeded. We are now able to produce analyses that in general terms match those reported by Stagg, and that are a matter for discussion where disagreement can be seen. Our forecasts, even when penalised for starting from analyses that included some observations taken several hours later, do depict the essentials of the evolution of the critical weather systems in June 1944 for 48 hours or longer ahead, despite the limitations of the observations used. The detail cannot be perfect, however, and variables such as cloud that remain today a challenge in some types of synoptic situation are yet more of a problem with the observations available from seventy years ago. And what we can do today in the quiet of our offices or even connected through the internet from the comfort of home should in no way detract from the achievements of Stagg and his colleagues under a very different set of circumstances.

Beyond the intrinsic interest in seeing how our results compare with what was achieved at the time by the teams of meteorologists supporting the D-Day landings on the Normandy coast, studying cases of weather extremes such as occurred that month is part of the general evaluation of our climate records. These studies also provide pointers to where additional observational data can be found, and identify problems in existing data holdings that can be corrected for future reanalyses.

Use of GPS-RO in operational NWP and reanalysis applications

SEAN HEALY

ECMWF hosted a workshop on the applications of GPS radio occultation (GPS-RO) measurements from 16 to 18 June 2014.

The workshop, jointly organised by ECMWF and the EUMETSAT Radio Occultation Meteorology Satellite Application Facility (ROM-SAF), reviewed the use of the GPS-RO data at the major NWP centres, and explored how the assimilation of the measurements can be improved using more sophisticated observation operators. The value of GPS-RO as 'anchor measurements' for estimating the bias correction of satellite radiances was emphasised. Also the workshop explored the potential new applications of monthly-mean and seasonal GPS-RO climatologies in NWP/climate model development and testing, and reviewed the latest work on climate monitoring with GPS-RO.

The NWP working group endorsed the ECMWF strategy of moving towards assimilation with a two-dimensional observation operator. They also recommended the use of a more sophisticated GPS-RO observation error covariance matrix in the 4D-Var.

The climate working group stressed the need for increased co-operation between GPS-RO and broader climate communities. Although it seems clear that GPS-RO will become increasingly useful for climate applications in the coming years, it was emphasised that new products would have to be accompanied by a robust estimate of the uncertainty.

The value and continued importance of GPS-RO in both operational NWP and reanalysis applications was emphasised at the plenary session.

Overall, this was a useful and enjoyable workshop, which benefited greatly from the diversity of the participants' backgrounds and research interests.

Participants at the GPS-RO workshop. Around 30 international scientists, with expertise in both GPS radio occultation and climate applications, attended the workshop.

ECMWF training courses are changing!

The 2015 training programme will see the introduction of a new module on the use and interpretation of ECMWF products. This additional module will run over three days and will be offered in the autumn of 2015. It will complement the two traditional '*Use and interpretation of ECMWF products*' modules that are scheduled for end of January and beginning of February 2015.

The new three-day module will focus on ensemble forecast products and their use and interpretation, with particular emphasis on practical activities based on case studies of weather events. The participants will be asked to complete activities before the start of the course, which will require 3 to 4 hours of self-study. More information on the course content and the exact dates can be found on the ECMWF website at:

http://www.ecmwf.int/en/learning/training.

ECMWF Annual Report 2013

The ECMWF Annual Report 2013 is now available. It is structured around the eight key objectives set out in the Centre's four-year programme of activities.

Weather prediction is critical for people and society, and in reviewing ECMWF over 2013 we reflect on our achievements for our Member States and society at large.

Looking back on a year of technical and scientific innovation, the report demonstrates how European co-operation and investment has helped citizens prepare for severe weather and anticipate the impacts of the weather on their lives.

The content of the report includes the following.

- 2013 at a glance
- Improving our forecasts
- Advancing weather science
- Improving users' experience
- More power
- Our people
- Developing European infrastructure
- European investment in ECMWF
- Working together
- Looking to the future
- How we work

The report is available from

http://www.ecmwf.int/sites/default/files/Annual_Report_2013.pdf

Launch of a new fellowship programme

ERLAND KÄLLÉN

ECMWF has a long history of working with respected institutions around the world through both formal and informal collaborations. The new ECMWF fellowship programme fosters collaboration with renowned international scientists by formalising links with individuals, rather than institutions, who are carrying out pioneering work in areas relevant to the strategic goals of ECMWF. Key to the programme's success will be close working relationships between individuals, with each Fellow having a main scientific and technical contact point at ECMWF.

July 2014 marked the start of a three-year tenure for the first ECMWF Fellows.

- *Tilmann Gneiting*: Group Leader for Computational Statistics at the Heidelberg Institute for Theoretical Studies and Professor of Computational Statistics at Karlsruhe Institute of Technology.
- *Rupert Klein*: Professor of Scientific Computing in the Mathematics Department at the Free University of Berlin.
- *Tim Palmer*: Royal Society Professor of Climate Physics at University of Oxford.

More information about the first three Fellows can be found at: http://www.ecmwf.int/en/about/media-centre/news/2014/ecmwf-launches-new-fellowship-programme

The ECMWF Directorate will appoint new Fellows every year, based on their expertise in areas of scientific and technical research and development relevant to the Centre. The total number of Fellows is expected to become around ten when the programme is fully operational. Fellowships will be for three years, with the possibility of being renewed once for another three years. Fellows are encouraged to provide regular, short reports of the key results of their work, which will be published on ECMWF's website and in other relevant publications. To support their research, the Centre will provide access to ECMWF computing facilities and databases, and a limited amount of annual funding for visits to the Centre to share the results of their research work.

ECMWF's fellowship programme is something we have been working on for a while, and we believe it will be very important for the future guidance and inspiration of our research. Scientific collaboration is the essence of progress and it is essential for us to be able to share and exchange with like-minded researchers around the world.

Use of radiosonde and surface observations provided in BUFR format

BRUCE INGLEBY, ENRICO FUCILE, TOMAS KRAL, DRASKO VASILJEVIC, LARS ISAKSEN

Alphanumeric codes (including TEMP and SYNOP used for radiosonde and surface reports respectively) have been used for more than 50 years and are restrictive in some respects. For over 15 years the World Meteorological Organization (WMO) has been promoting and coordinating a transition to the binary BUFR format. BUFR surface reports contain extra metadata and facilitate higher frequency reporting but the basic content is much the same (except that some contain radiation measurements that are not available in SYNOPs). BUFR radiosonde reports allow much higher vertical resolution (removing the need to select 'significant' levels) and the reporting of time/position at each level. The transition process has been slow because in each country it involves experts in observations, codes, telecommunications, databases, and finally numerical weather prediction (NWP). From November 2014 circulation of some reports in alphanumeric format may cease.

ECMWF has been working on decoding and storing the observations provided in BUFR format (BUFR reports) and on their use in the NWP system. Validating the reports is important: various errors have been found and reported to the data producers, either directly or via EUCOS (EUMETNET Composite Observing System) and WMO, and some of the errors have been resolved. Information exchange with other NWP centres is ongoing and helps to ensure that all available bulletins reach ECMWF.

Current land surface coverage

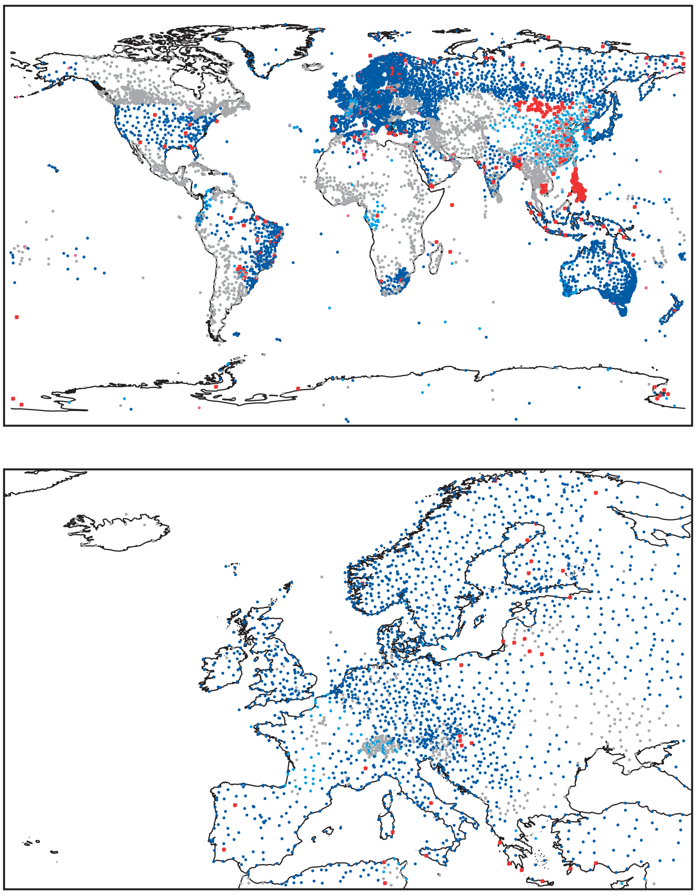
The top panel of the first figure shows current global coverage for land surface stations, about 65% of them issue some BUFR reports but some (shown in light blue) make significantly fewer BUFR reports than SYNOP reports.

As shown in the bottom panel, the best coverage is over Europe but even here there are gaps. For land stations BUFR reports include position information. TEMP/SYNOP reports rely on the separate provision of positions in *WMO Publication 9, Volume A*. Up to 3% of SYNOP reports are unusable because the position information is not available (*Ingleby*, 2014, *Q. J. R. Meteorol. Soc.*, doi: 10.1002/qj.2372). The change to BUFR format should largely avoid that problem but some positions in BUFR reports are erroneous (often due to wrong conversion from degrees/minutes to decimals). Some of the position errors have been corrected but some remain (shown in red).

Changes to radiosonde reporting

About 45% of global radiosonde stations provided some BUFR reports (see the second figure). Comparing data between TEMP and BUFR formats is complicated because the ascents are split up in different ways and TEMPs often give the nominal time whereas BUFR reports contains launch time. Full radiosonde ascents take about two hours. To improve timeliness data is sent when the ascent reaches 100 hPa (TEMP parts A and B contain standard and significant levels respectively) and after balloon burst (TEMP parts C and D).

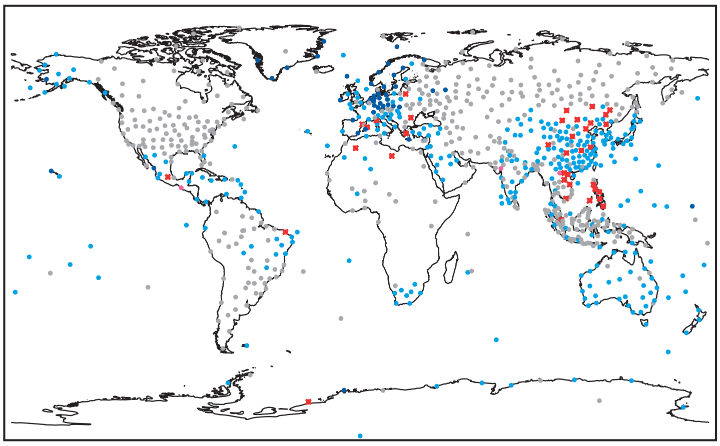
In BUFR format there should be one report when the radiosonde reaches 100 hPa, and one report containing the



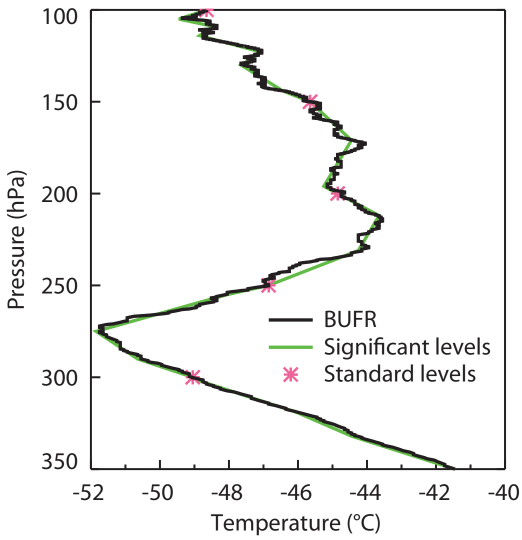
1: Land surface data available at ECMWF, 1–5 July 2014. Top: About 65% of global land stations provide some BUFR reports. Bottom: Almost 85% of European land stations in the area shown provide some BUFR reports. The light/dark blue dots are where the number of BUFR reports is below/above 60% of the number of SYNOP reports. Grey: no BUFR reports; purple: no alphanumeric reports; red *: position error.

whole ascent. There is also a move to reporting of high-resolution data (e.g. every two seconds during the ascent) which can result in 5,000 or more levels compared to 100 or so for TEMP code (all parts combined). The third figure shows an example of the much higher vertical resolution in a BUFR report which should result in minor improvements to the representation of radiosonde ascents in the ECMWF analysis system. *Ingleby & Edwards* (2014, *Atmos. Sci. Lett.*, doi: 10.1002/asl2.518) provide more details of the reports, document a small offset in the temperatures (due to rounding in the TEMP coding/decoding) and discuss processing options including vertical averaging and the treatment of balloon drift. Treatment of radiosonde drift is planned for the ECMWF NWP system, but the immediate priority is making sure that we can correctly use the BUFR reports before they become the only source of radiosonde data.

High-resolution radiosonde data is currently available from about 6% of stations: from Germany, Scandinavia, Greenland and a few stations in Spain, Russia and elsewhere (dark blue points in the second figure; in some cases the report up to 100 hPa is at low resolution). Unfortunately most of the other BUFR reports are currently reformatted from TEMP code (still as separate parts A, B, C and D which breaches the BUFR coding rules and complicates processing). There is a trend for GPS radiosondes not to include a pressure sensor, and one BUFR template includes geopotential height but not pressure (not processed at ECMWF yet).



2: Radiosondes available and decoded at ECMWF for 1–5 July 2014. This shows that about 45% of global radiosonde stations report some BUFR. Dark blue: reports with at least 400 levels; light blue: generally reformatted TEMP; grey: no BUFR reports; purple: no TEMP reports; red *: position error.



3: Part of a temperature profile from a radiosonde ascent. The BUFR profile contains more detailed information than is available from the significant levels and standard levels. Initially the ECMWF assimilation system will subsample the BUFR profile which is at much higher vertical resolution than current forecast models. Shown are the high-resolution BUFR profile, profile based on significant levels and data at standard levels. The ascent is from the autosonde station at Castor Bay, Northern Ireland, UK, at 00 UTC on 22 August 2012.

Observation processing changes at ECMWF

In June SAPP (Scalable Acquisition and Pre-Processing system for observations, see separate article starting on page 37) took over the supply of observations to the NWP system. (NB For years ECMWF has used BUFR format internally even for observations arriving in alphanumeric formats; in general there are differences in BUFR template from the new data streams – and the arriving land surface data uses four different BUFR templates.) From late 2014 SAPP will pass the data to COPE (Continuous Observation Processing Environment), another new component system. Eventually COPE will run quasi-continuously, transforming variables as required and performing as much data selection and checking as possible before the start of each assimilation cycle.

In the near future ECMWF will process TEMP/SYNOP and BUFR reports in parallel and monitor both sets of reports (BUFR ship and buoy data will be added to the system next year). We will start assimilating subsets of BUFR data (in place of alphanumeric data) when the quality and reliability are acceptable. We will continue to feedback any problems to data producers and share information with other NWP centres, both directly and via https://software.ecmwf.int/wiki/display/TCBUF/.

Separate developments/plans include the trialling of aircraft humidities and short-haul (TAMDAR) data, improved radiosonde processing and use of surface temperature, humidity and wind in the atmospheric analysis.

Working together to address weather forecasting challenges

ANNA GHELLI,
FLORENCE RABIER

The annual meeting on 'Using ECMWF's Forecasts' took place at ECMWF from 4 to 6 June 2014. This meeting provides a forum for exchanging ideas and experiences on the use of ECMWF data and products. It is also an opportunity for participants to provide feedback to ECMWF on forecast performance and the range of products and services available, and to be updated on the recent developments affecting the Integrated Forecasting System (IFS).

The title *'Working together to address weather forecasting challenges'* highlighted two important elements: the recognition that weather forecasting is changing and that forecasting centres and international organisations need to work together to address the challenges.

Four thematic areas served as a focal point for the discussion: integration of NWP data into impact models, severe weather, seamless forecasting and uncertainty, and NWP challenges.

The meeting offered a number of activities aimed at encouraging, in a relaxed atmosphere, dialogue and generation of ideas. The participants could also enjoy the newly-opened weather room.

Participants in annual meeting on 'Using ECMWF's Forecasts'. The meeting gathered an international audience of about 80 people from National Meteorological Services (NMSs), commercial users and international organisations.

The weather room. The light and vibrant colours in the weather room created the perfect setting to browse through the weather information displayed on the large screens and engage in discussions with other enthusiasts, meteorologists or weather amateurs.

The highlights from the four thematic areas are summarized here.

- The World Meteorological Organization has launched Severe Weather Forecast Demonstration Projects aimed at providing some NMHSs in developing countries with access to NWP products to improve their severe weather forecasts. Unfortunately there are still a large number of NMHSs in developing countries that are not able to take full advantage of these projects, usually due to lack of capacity, connectivity and/or financial resources. Moreover, cascading systems that allow impact forecasting are not widely available. Financing models and opportunities to support and further develop these systems need to be explored. A skeleton of such a cascading system is implemented in the European Flood Awareness System (EFAS) which links a weather forecasting model to a hydrological model to provide risk maps to its members.
- More businesses are now becoming aware of the impact that weather may have on their activities. However, they do not have the threshold knowledge required to translate forecasts into the multi-variable optimized solutions relevant to them. Intermediaries can help by speaking with the business owner in their language and delivering a bespoke forecast which shows the potential impact on profit/loss and/or quality.
- Forecasters value ECMWF's ensemble forecast (ENS) products as an essential support for severe weather forecasting. The assessment of ECMWF model performance in the medium to extended range for specific severe weather events, now available on the new Forecaster User portal, is an important addition to understanding the strengths and weaknesses of ECMWF's forecasts. Moreover, information of issues affecting forecasts is available on the portal: https://software.ecmwf.int/wiki/display/FCST/Forecast+User+Home.
- NMSs recognize the importance of uncertainty in the forecasting process and they are embracing the idea that ensemble forecasts should be used for all forecasting activities. It was pointed out that the full spectrum of probabilities should be complemented with weather scenarios with low and high probabilities.
- Extended- and long-range forecasts have the challenging task of finding sources of predictability in the observational record and identifying their dynamical mechanisms which can be captured by models. International projects, such as the S2S (Subseasonal to Seasonal Prediction Project), are the way forward in identifying dynamical mechanisms. To help achieve the goal of S2S, an extensive database of sub-seasonal to seasonal forecasts

and re-forecasts will be established at ECMWF.

- The presentation of the information across time ranges is still evolving. Probabilities such as upper and lower terciles are difficult to interpret and commercial customers are increasingly interested in receiving more specific products (e.g. information about heating degrees, droughts and heat waves) which are directly relevant to their businesses and have useful skill. At the same time, attempts to show information in a seamless way across scales are being developed using regime analysis.
- Moving towards higher horizontal resolution to improve accuracy will bring a variety of challenges for the forecasting system. In the autumn of 2014, along with other model improvements, the IFS will include a lake model which will not only allow a better description of the inland water bodies, but will improve the forecast for coastal locations where the current land-sea mask is a crude approximation of the real land/sea position.
- Future increases in model resolution will also challenge the current computing set-up and ECMWF is working on a scalability project aimed at preparing the whole forecasting chain, including the IFS, for highly parallel high-performance computing architectures.

User voice corner and helpdesk activities

Participants were offered the opportunity to provide feedback to ECMWF during a specially designed session whereby in small groups led by one member of ECMWF, comments and suggestions were encouraged on three areas.

- Performance of ECMWF forecasts in the last 12 months.
- Improvements or addition of ECMWF Forecast products.
- ecCharts and the new ECMWF website.

During a plenary session, all the groups presented their findings and a list of requests to ECMWF was collated.

The helpdesk activities that ran on the last day aimed at providing support for technical and meteorological aspects. The new ECPDS dissemination interface was demonstrated to commercial customers who get their data directly from ECMWF.

Panel discussion

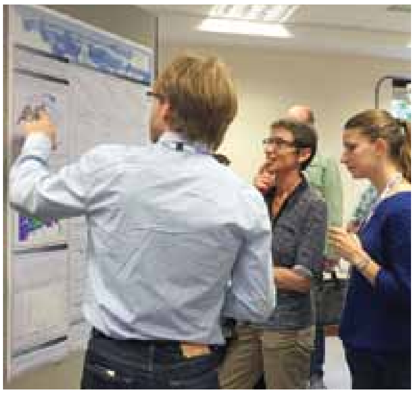
This year a lively panel discussion was included on the last day of the meeting. The five panelists, representing NMSs, private sector companies and international organisations, gave their views on the following question: 'Over the next ten years, in the light of how you expect your business to develop, what is it in general terms that you will need from an NWP producing centre like ECMWF?'.

The panelists highlighted the new role for forecasters dictated by an ever closer link between NMSs and end users. Forecasters are no longer data producers – they are model experts and advisors for customers. In this way forecasters add value to the raw model output. In the same way, private enterprises that provide weather information add value by tailoring products for the end users. The challenge of the development of cascading systems that forecast impacts was recognized as an important priority as such systems would provide scientific information to the decision-makers.

The panelists agreed that models at higher resolution than at present are likely to give improved skills when forecasting severe weather, but higher resolution at the global scale will bring a number of challenges, both scientific and technical. In these two areas ECMWF can provide help by understanding the science and improving the current models, and by preparing the IFS for highly parallel high-performance computing architectures with the Scalability Project.

One important role of ECMWF has been to provide boundary conditions for limited-area models. The panelists stressed that this needs to continue in the future. In addition, users would like to have more frequent analyses/forecasts (e.g. ensembles four times a day and hourly output) to drive their models.

During the panel discussion, training was mentioned as a valuable service offered by ECMWF. The challenges for the future will be to make sure that such training is part of a capacity building framework.



Group activities and poster session. These events provided opportunities for giving feedback and demonstrating how to take full advantage of what ECMWF can provide.

What is next?

Participants expressed their appreciation for this year's meeting and the opportunity of giving feedback to ECMWF. We would like to conclude with what we think are the three points to take away from the meeting.

- Ensemble forecasts are widely used, their accuracy has improved, and forecasters rely on ENS products for their severe weather forecasts. This trend has fed into NMSs strategies which are suggesting the use of ensemble forecasting at all time ranges.
- There are some gaps, such as the integration of impact models into ensemble weather forecast models and the handling of weather information at all forecast ranges in a seamless manner. These gaps will need to be addressed to be able to deliver relevant weather information to decision-makers.
- There are still challenges to communicate and use probabilistic information.

This workshop demonstrated that there is the will for users to engage with ECMWF to discuss forecast performance and the usefulness of products, and to create innovative solutions for solving future challenges. We are looking forward to seeing you next year to continue our engagement with users.

Wave experts meet at ECMWF

PETER JANSSEN, JEAN BIDLOT

ECMWF hosted the 2014 meeting of the WISE (Waves In Shallow Water Environment) group that took place from 8 to 12 June. The group's aim is to improve the knowledge of physical processes involved in wave development and evolution from basin-scale to coastal environments.

This year's programme included talks by participants and ECMWF staff, keynote addresses on specific topics, and side meetings. ECMWF's Adrian Simmons and Peter Janssen both addressed delegates: Adrian gave the introductory presentation on a reanalysis of the weather and ocean waves during the D-Day landing (see the news item on page 7), whilst Peter presented developments in coupled ocean wave and ocean circulation modelling at ECMWF.

Wind waves represent a basic parameter of interest for marine applications. In the open ocean, they dictate the design and operational conditions of offshore structures and vessels. Close to the coast, the design and planning of any activity depends on the waves and on the local environment they create. The impact of waves on coastline flooding and defence is also important.

On a wider perspective, wind waves regulate to a considerable extent all transfers taking place at a variety of scales at the air-sea interface (e.g. momentum, energy, humidity, heat and aerosol). These transfers play a key role for atmosphere, ocean and sea ice dynamics. Modelling these exchanges is essential in the development a fully coupled Earth-system.

As well as considering marine applications and transfers at the air-sea interface, the meeting also provided a forum to present and discuss novel techniques aimed at improving the numerics and reliability of wave models.


2014 edition of the Waves In Shallow water Environment meeting took place at ECMWF from 8 to 12 June

10th Eumetcal International Workshop in 2015 to be held jointly with ECMWF

The 10th Eumetcal International Workshop will be on "advancing numerical weather prediction training and education". It will be held from 15 to 18 June 2015 in Reading, UK, jointly organized with and hosted by ECMWF.

NWP products and models are continually developing in response to societal needs. This workshop aspires to bring together the academic, forecasting and training communities to discuss future training needs and solutions, as well as contribute to improving the use and understanding of NWP in hydro-meteorological services and universities.

A second announcement in fourth quarter 2014 will provide further information on the workshop themes and practicalities.

More information will become available at the workshop webpage:
http://bit.ly/EumetcalWS2015

Potential to use seasonal climate forecasts to plan malaria intervention strategies in Africa

FRANCESCA DI GIUSEPPE, ADRIAN M TOMPKINS

Despite a reduction in the number of cases of malaria over the past century, it remains a disease with a heavy health burden. The life cycles of the malaria parasite and its mosquito vector are affected by climate, principally temperature and rainfall. Thus, in addition to other factors such as land cover, human migration, malaria interventions and socio-economic conditions which can alter the local disease prevalence significantly, year-to-year fluctuations in climate can lead directly to variability in the intensity of malaria transmission. The period of the malaria parasite development in both host and vector results in a temperature-dependent lag of approximately one to two months between the onset of suitably wet conditions for vector proliferation and the appearance of malaria symptoms. Thus accurate real-time forecasts of temperature and rainfall conditions could provide useful information concerning malaria transmission if employed in Malaria Early Warning Systems (MEWS). Researchers across 13 European and African research institutes have worked together to integrate data from climate modelling and disease forecasting systems to predict the likelihood of a malaria epidemic.

In this article we analyse the potential of an integrated malaria prediction system by identifying regions and months for which climate variability directly translates into significant variability in malaria transmission. Local temperature anomalies are predictable from one to two months ahead, while reliable rainfall forecasts are only available in eastern and southern Africa one month ahead. Nevertheless, the inherent lag between the rainy seasons and the onset of malaria transmission results in there being potential predictability of malaria cases three to four months in advance. This would extend the early warning available from environmental monitoring by one to two months.

Predicting malaria transmission: why?

Weather forecasts could extend the advance warning of outbreaks provided by climate monitoring and supply them on a regional or even continental scale. One of the first attempts to show the potential for this use of climate forecasts dates back to 2006. A collaboration involving colleagues from ECMWF showed the potential for using seasonal forecasts from DEMETER (Development of a Multi-model Ensemble System for Seasonal to Interannual Climate Prediction) to drive a simple statistical model for malaria. This provided a prediction of the total number of malaria cases for Botswana (see *Thomson et al.*, 2005 for further details). The forecasts appeared to have skill up to six months ahead, perhaps facilitated by the higher predictability in this location compared with others due to strong teleconnections to ENSO (El Niño–Southern Oscillation). At that time no attempt was made to assess the potential predictability of an operational system at a pan-African scale. Note that the derivation of a statistical malaria model to predict cases is only possible if long, reliable and readily available health data records exist; this is not the case for many countries in Africa.

During the three years (2010–2013) of QWeCI (Quantifying Weather and Climate Impacts on health in developing countries), a project funded by the European Commission's Seventh Framework Research Programme, we have attempted to build a MEWS (*Thompkins & Guiseppe*, 2014). Given the noticeable advance in the quality of ECMWF's seasonal predictions, this time the long-range forecasts have been used to drive a dynamical malaria model (*Tompkins & Ermert*, 2013). This model attempts to represent the physics of the malaria transmission cycle explicitly, and allows it to be applied in countries where health data records are short or not available.

The first month of the forecast uses temperature and rainfall provided by the monthly extension of ECMWF's ensemble forecast. These are then substituted by data from the lower-resolution and longer-range seasonal forecasts from System 4 for months two to four (*Di Giuseppe et al.*, 2013). The integration of the VECTRI malaria model produces an ensemble of forecasts of a range of epidemiological and entomological parameters; the most useful parameters for operational applications being:

- *Parasite ratio* (PR): proportion of people infected.
- *Entomological inoculation rate* (EIR): the number of infective bites per person per time.

The malaria model is initialized from realistic initial conditions provided by a malaria analysis system obtained using forcing from the ERA-Interim reanalysis – see Box A.

Using this system we will try to answer two key questions: where in Africa would a MEWS providing regularly updated forecasts be most useful and what is the prediction skill of the malaria forecasts in these locations?

Predicting malaria transmission: where and when?

Previous efforts to provide early warnings have pinpointed a selection of locations subject to sporadic epidemics or irregular short transmission seasons. In these locations adult immunity is lower and the intermittent transmission implies health facilities may not be adequately prepared

AFFILIATIONS
Francesca Di Giuseppe: ECMWF, Reading, UK
Adrian M Tompkins: ICTP, Trieste, Italy

for significant outbreaks. The first step is therefore to locate areas and calendar months with high variability in malaria prevalence, using the daily malaria analysis system.

The standard deviation of the parasite ratio for each calendar month provides an insight to the modes of variability of malaria transmission, with Figure 1 showing four example months. In each of these months the standard deviation of PR shows three distinct modes.

- *Malaria free*. There is a mode at zero parasite ratio which simply identifies malaria-free zones.
- *Low variability*. The second mode identifies regions where the year-to-year standard deviation is non-zero, but less than around 10%, with the upper bound changing slightly from month to month. This mode is associated with endemic zones, where transmission regularly occurs in those months each year.
- *High variability*. The third mode encompasses the remaining higher values of standard deviation exceeding 10%. The highest standard deviations belong to locations where transmission is very intermittent and does not occur every year but instead in occasional epidemic outbreaks.

Events in the high variability mode are termed as 'true' or 'classic' epidemic. However, this mode also includes locations where transmission is regular each year, but the transmission season is irregular in the particular month in question (e.g. occurring later or earlier according to variability in the onset of monsoon rains). Thus this mode also encompasses situations of 'unusual seasonal transmission' and may include mesoendemic and hyperendemic zones – see Box B for the definitions of mesoendemic and hyperendemic as well as hypoendemic and holoendemic.

Based on this analysis of the variability of malaria prevalence rates, we identified all locations where the standard deviation of PR for a particular month exceeds 10% of the population to identify the epidemic mode. In order to relate prevalence more closely to expected cases,

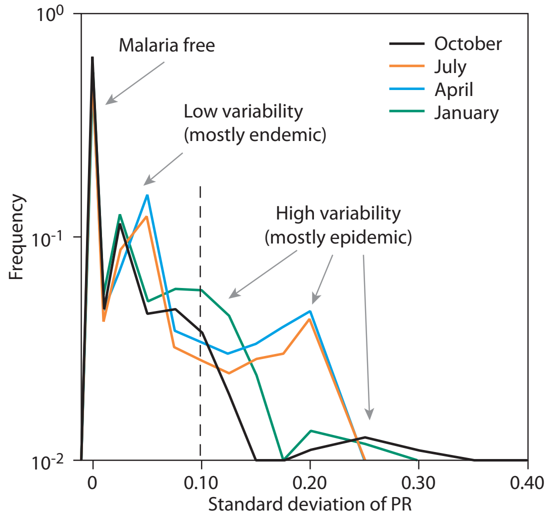


Figure 1 Probability density function of the standard deviation of the PR for four representative months of the year. The vertical dashed line represents a standard deviation of 10% used as a threshold in the analysis to identify high variability regions.

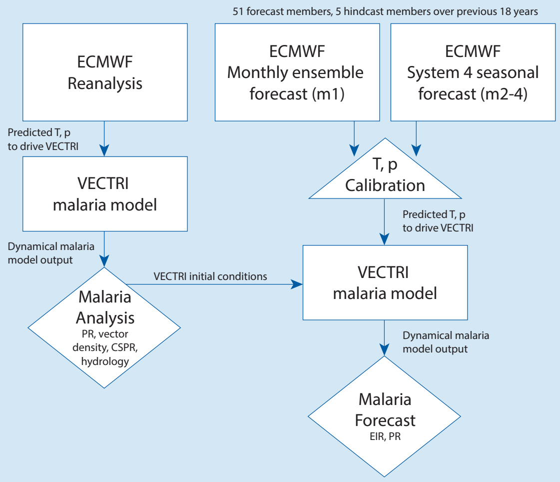
an additional filter is applied to exclude months in which transmission intensity is minimal and no new cases are expected. This is accomplished by excluding months where the EIR falls below 0.01 per month. The locations are then subdivided into two categories.

- The first category includes the hypoendemic and mesoendemic zones where immunity is likely to be low and the malaria hazard impacts the entire population.
- The second category concerns the hyperendemic regions, where children under five are most at risk (most holoendemic locations are excluded by the PR variability threshold since the interannual variation is low for all months).

This subdivision is made since malaria interventions and preparations are likely to differ in the two transmission environments. The separation of hypoendemic/

Schematic of the forecast system

A



Schematic of the forecasting system set up, with boxes representing models, triangles for processes, and diamonds for products. The operational numerical weather prediction reanalysis of temperature (T) and precipitation (p) is used to drive the malaria model to provide a malaria analysis of epidemiological and entomological parameters (PR=parasite ratio, CSPR=Circumsporozoite Protein Rate), which are used as initial conditions for the forecast.

The malaria forecast uses climate information from the ensemble climate forecasts in the first month (m1), which is seamlessly combined with information from the seasonal forecasts in months two to four (m2-4). Both temperature and precipitation are calibrated before application to the malaria model, which then provides forecasts of PR and entomological inoculation rate (EIR).

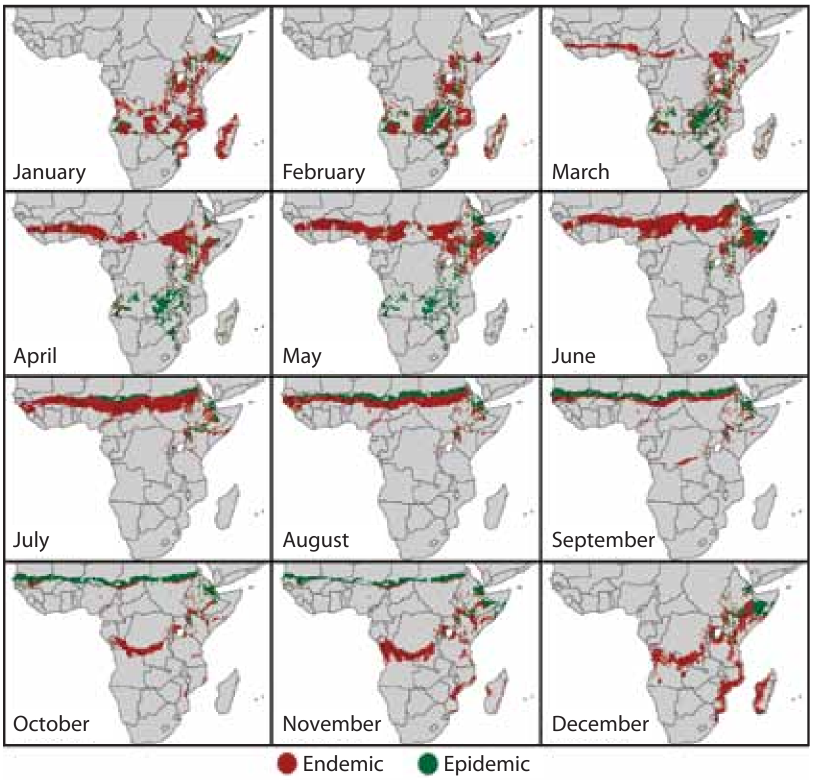


Figure 2 Locations of high variability in parasite prevalence are shown for each calendar month of the year, subdivided into endemic (hyperendemic and holoendemic, red) and epidemic (hypoendemic and mesoendemic, green) transmission zones.

mesoendemic from hyperendemic regions is made using the common threshold of 0.5 for annual mean PR. This threshold can be applied to the whole population since the model neglects immunity, which increases parasite clearance rates in adults relative to malaria-naive children. However, it also implies that the model is likely to overestimate the hyperendemic region relative to the lower transmission classes.

The map in Figure 2 identifies where and when a reliable malaria forecast would have a relatively high potential value to the decision maker. The map identifies epidemic regions such as the Sahel fringe, the East African highlands and the southern-most transmission regions skirting through Botswana, southern Africa and Mozambique. In each region the malaria model's representation of the vector and parasite life cycles results in a peak malaria transmission that is lagged with respect to the rainy season by one to two months. It is in these zones that the indication of an outbreak by a reliable forecast could trigger the whole host of intervention strategies, depending on the status of the national malaria control programme in question (control, elimination or post-elimination surveillance).

It is emphasized that the assessment of 'where and when' is only for the theoretical climate-related variability of malaria and may not reflect the situation in reality. If control measures have led to local eradication, for example, then indication of heightened hazard due to climate anomalies may serve to tighten surveillance and response measures to imported cases in the region in question to prevent explosive growth of secondary cases. Likewise, this assessment obviously neglects zones prone to epidemics due to non-climate factors such as population movements and breakdowns of health services due to conflict or other socio-economic factors.

In addition to these epidemic zones, the map also reveals areas with annual seasonal transmission but where high variability is prevalent in certain months of the year, for example, in a band spanning the northern half of the Ivory Coast, Ghana, Togo and Benin in West Africa during April and May. Malaria transmission variability is highest

Terms used to describe regional malaria risk

B

The following terms have been used to empirically gauge regional risk.

- *Hypoendemic*. Little transmission, and the effects of malaria on the community are unimportant.
- *Mesoendemic*. Variable transmission that fluctuates with changes in one or many localconditions (e.g. weather or disturbance to the environment).
- *Hyperendemic*. Seasonally intense malaria transmission with disease in all age groups.
- *Holoendemic*. Perennial intense transmission with protective clinical immunity among adults.

in these regions during the rain season onset (also in April/May at these latitudes) and is associated with the interannual fluctuations of the monsoon cycle which are highly variable from year to year. In these hyperendemic zones, health services are likely to be geared up to dealing with regular malaria transmission, unless funding shortages or breakdown of health structures preclude this. However, a skilful forecast would still be useful for ensuring timely mobilization of indoor residual spraying (IRS) teams using insecticide to kill mosquitoes that spread malaria.

Our subsequent analysis uses these maps to exclude locations for each calendar month where climate is deemed less relevant for interannual variability of malaria transmission.

Predicting malaria transmission: how?

We now consider the potential malaria predictability in the locations that have been identified in Figure 2. Here the word 'potential' is emphasized as we use the malaria analysis system to quantify the malaria prediction skill. This is necessary as no continent-wide health data set exists. Thus the actual skill will be necessarily less than reported in this analysis due to the neglect of errors in the malaria modelling system. Nethertheless, the analysis should provide useful indicators to where skill deriving from climate prediction exists, and thus in which countries further national-level investigation of the system should be focussed.

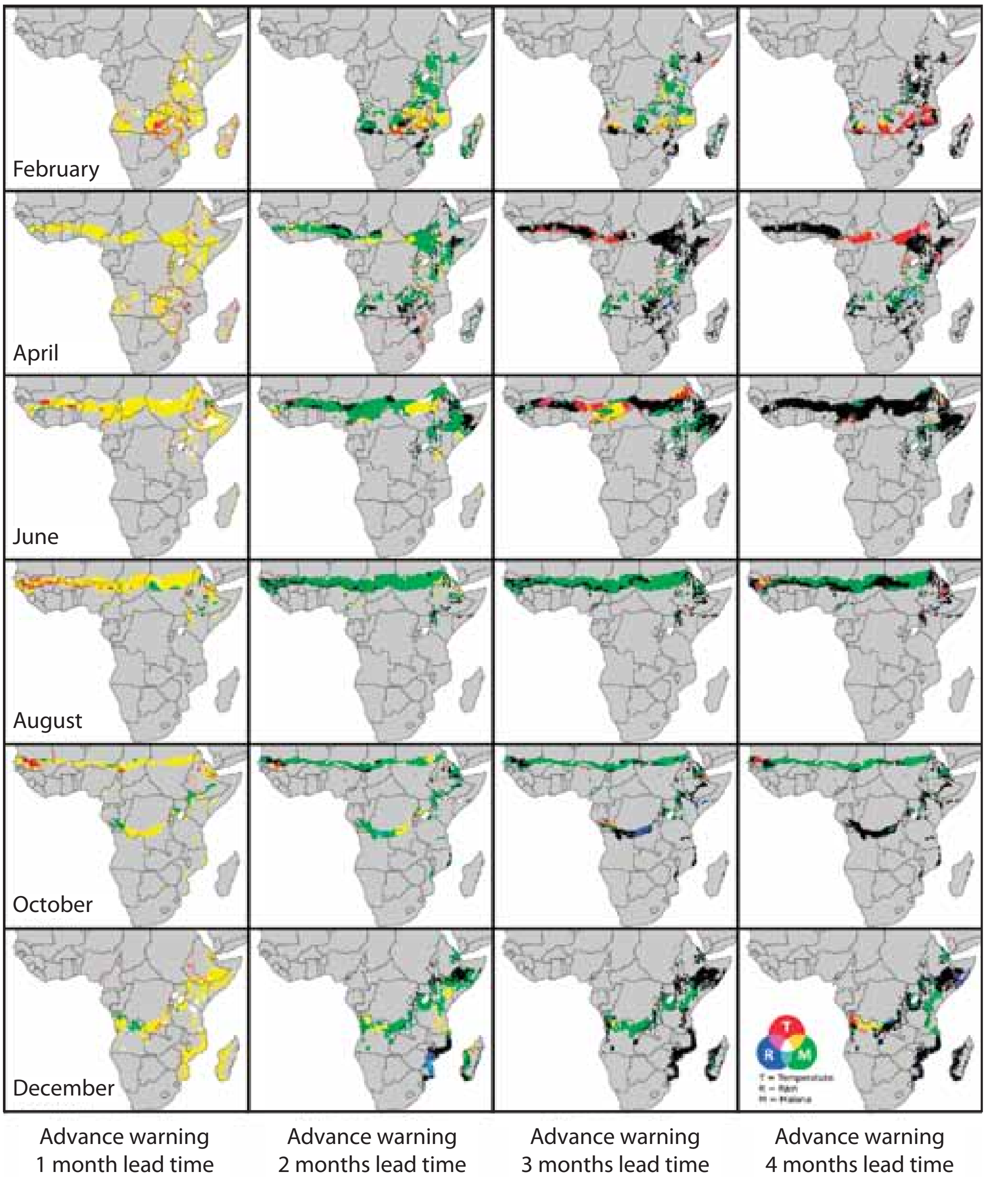


Figure 3 Composite plot of temperature (red), rainfall (blue) and malaria (green) forecast anomaly correlation coefficients that are statistically skilful at the 95% confidence level for issuing warnings for 1 through 4 months in advance (lead time) for alternate months of the year. White points mark cells with skill in all three variables, black points mark cells without any skill. See legend for colour definitions of intersecting categories.

The skill of the forecast is examined for both the climate and malaria forecasts using the operational forecasting system output for 2012. We analyze the forecasts of 2012, and the 18 years of hindcasts, giving an evaluation period of 1994 to 2012. Although ideally one would evaluate an ensemble system using probabilistic skill scores, the small ensemble size of the hindcast (five members) precludes this and thus an assessment is made for the skill of the ensemble mean anomaly correlation. Temperature and rainfall are validated against the ERA-Interim reanalyses. We identify the locations where skill in predicting malaria transmission is statistically significant one to four months in advance (referred to generically as 'malaria prediction skill') for each month. In addition to malaria, we show the skill in predicting anomalies of rainfall and temperature to identify which of these variables generate any identified malaria prediction skill.

Examining first the predictions one month in advance (Figure 3, first column), encouragingly, there is model skill in malaria predictions in the target prediction zones throughout the year. In some regions the predictability derives from correctly forecasting variations in temperature. However, in southern Africa, in a band stretching from Botswana through to Malawi and also across Eastern Africa, there are wide areas in which malaria predictive skill derives from both rainfall and temperature; for these the analysis does not show which variable contributes most to the skilful malaria prediction. In these regions rainfall predictability tends to be higher owing to stronger teleconnections with the El Niño phenomenon. Outside of these regions, skill in rainfall prediction appears limited in the areas of interest for malaria forecasting, in broad agreement with studies using the predecessor of the seasonal forecast (System 3).

In some locations the malaria forecasts are not significantly skilful, marked by a limited number of black points where predictions of all variables fail, or by blue, purple or red points which indicate skill in climate but not malaria prediction. In the northern-most Sahel belt spanning Senegal, Mali and Niger in July and August, wide areas display skill in temperature only (red colours) while in some places rainfall is also correctly predicted (purple colours) but no malaria prediction skill ensues. In this northern-most zone of the Sahel, rainfall variability and the northern extent of the monsoon limit malaria transmission. Thus, where rainfall predictions (despite bias correction) are inaccurate, a frequent shortcoming in atmospheric models, malaria predictions will also fail. Where both rainfall and temperature are skilfully predicted, the failure to translate this into accurate malaria prediction could relate to the non-linear relationship between transmission and rainfall, where intense rain events flush early-stage larvae breeding sites and monsoon breaks lead to puddle desiccation. This non-linearity is fully sampled by the high day-to-day variability of rainfall in the tropics, thus significant skill in predicting seasonal rainfall anomalies may not be sufficient if sub-seasonal rainfall variability is poorly represented.

Analyzing the malaria prediction skill for longer lead times of two to four months (Figure 3, columns 2 to 4), it is seen that the climate prediction system exhibits a sharp drop in skill at predicting rainfall and temperature two months in advance compared to one month. Despite this, there are wide areas for which the pilot MEWS still has significant skill for malaria prediction in months 2 and 3, and in smaller regions even four months ahead. This is due to the inherent lags between the rainfall anomalies and the resulting malaria transmission season. The skill in predicting malaria transmission in the second and third months derives from the climate information contained in the forecast initial conditions and the first more skilful month from the monthly forecast. This highlights the crucial role that the malaria analysis system has in correctly initializing the malaria modelling system. In areas where rainfall and temperatures are predictable beyond one month, such as in Eastern Africa, the malaria prediction advanced warning is extended beyond the three-month range.

The analysis indicates that, by driving the malaria model with long-range forecasts, useful information regarding the future transmission season in epidemic and seasonally variable endemic regions can be deliverable at least one, and in limited regions, two to three months earlier than would otherwise be the case using climate observations, which themselves provide more advance warning than the direct monitoring of symptomatic malaria cases.

From potential predictability to decision support: the next steps

The results we obtained have demonstrated the potential for skilful malaria predictions up to four months in advance over wide areas that were identified as having highly variable transmission for specific months. This is an important result; it demonstrates for the first time that climate forecasts may usefully extend the early warning available from environmental monitoring on a continental scale and reaffirms the potential importance of accurate climate information on this continent. However, it is important to emphasize that this study is just the first step because it is limited to identifying the potential skill in such a system.

The next phase is underway, in collaboration with two health ministries in Africa; this involves conducting a detailed evaluation of the system at the health district level. Such endeavours are complicated by the relatively short and often incomplete records of malaria cases available in most countries, which are infrequently confirmed by laboratory tests; this highlights the need for improved clinical datasets for the evaluation of MEWS required for their uptake. This interactive process will likely identify areas of the malaria modelling system in need of further development.

In addition, the next phase will tackle the issues associated with integrating such forecast information into the present decision-making process. This will provide guidance on forecast products most useful at a district and national level, and how to best communicate forecast uncertainty to decision makers to effectively complement existing planning strategy.

As cited earlier, there is wide evidence that climate information could increase advance warnings and the cost-effectiveness of malaria interventions. However, demonstrations of MEWS based on climate monitoring have not been widely adopted in an operational environment to assist planning in African health ministries. To a large extent these ministries base decisions concerning drug distribution and interventions on long-term maps of mean malaria prevalence. In the past, health ministries have aimed to increase the efficiency of disease monitoring systems, improve the reaction time to the onset of epidemics and ensure coping strategies are sufficient with minimum cost. They have nevertheless been unfamiliar with the operational paradigm of using climate or other information sources to predict outbreaks in advance. Using climate information to predict outbreaks implies the incorporation of uncertainty into the decision-making process and the risk of costs associated with a prediction failure (termed a 'forecast miss'). The potential benefits of such information may be deemed to be outweighed by the perceived risk of an incorrect action.

The above considerations serve to emphasize that integrating climate forecast information into the decision-making process will require extensive country-level evaluation of the system's past performance, including cost–loss analysis of potential intervention actions taken on the basis of the information. To carry out such an analysis adequately, improvements in the representation of model uncertainty and increased ensemble sizes will be necessary, since the present system uses a single weather ensemble forecast to drive a single malaria model. Additional climate forecasting systems and malaria models should improve the representation of model uncertainty.

A further complication is that appropriate actions based on forecast information will also depend on a country's malaria intervention phase (malaria control, progress towards elimination or post-elimination surveillance). While forecasts may have such use in guiding timely IRS team mobilization in hyperendemic zones where earlier than usual malaria transmission is predicted, here the emphasis has been on epidemic regions. With this approach one can envisage forecast information aiding a wide range of district-level management decisions in countries still in the phase of malaria control (e.g. ensuring adequate drug supply to clinics and reassigning vector-control intervention actions to at-risk districts). District offices could also use forecasts to enhance information campaigns to vulnerable populations to increase bite-avoidance behaviour such as ensuring use of nets supplied in previous mass distribution campaigns. This may involve the combination of the malaria hazard forecasts with high spatial resolution modelling of population vulnerability which can change dramatically over small spatial scales.

This research was supported by the European Commission's Seventh Framework Research Programme under QWeCI (Quantifying Weather and Climate Impacts on Health in Developing Countries, grant agreement 243964) and HEALTHY FUTURES (266327).

FURTHER READING

Thomson, **M.C.**, **F.J. Doblas-Reyes**, **S.J. Mason**, **R. Hagedorn**, **S.J. Connor**, **T. Phindela**, **A.P. Morse** & **T.N. Palmer**, 2006: Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature*, **439**, 576–579.

Tompkins, **A.M.** & **V. Ermert**, 2013: A regional-scale, high resolution dynamical malaria model that accounts for population density, climate and surface hydrology. *Malaria J.*, **12**, doi:10.1186/1475-2875-12-65.

Di Giuseppe, **F.**, **F. Molteni** & **A.M. Tompkins**, 2013: A rainfall calibration methodology for impacts modelling based on spatial mapping. *Q. J. R. Meteorol. Soc.*, **139**, 1389–1401.

Tompkins, **A.M.** & **F. Di Giuseppe**, 2014: The potential predictability of malaria using monthly and seasonal climate forecasts in Africa. Submitted to *J. Appl. Meteorol. Climatol.*

As a part of the QWeCI activity, in April 2013 the International Centre for Theoretical Physics (ICTP) in Trieste, Italy, hosted a school related to climate and health. Lectures given by a large number of experts and leaders in the field from across the globe were combined with many hands-on exercises for the participants to practice the techniques they had learned. The school was funded by ICTP, with additional generous support from the World Meteorological Organization.

Automatic checking of observations at ECMWF

MOHAMED DAHOUI, NIELS BORMANN, LARS ISAKSEN

Observations are essential for numerical weather prediction (NWP) systems. They are used by the data assimilation system to produce the best estimate of the initial conditions (the analysis). The quality of the analysis and the subsequent forecasts depend, amongst other factors, on the quality and availability of observations. A reliable monitoring system is therefore needed for early detection of observational data issues that can potentially degrade the quality of the analysis. The steady increase in data volumes makes it difficult to rely solely on manual checking procedures.

In 2008, ECMWF started to use an automatic data checking system to trigger warnings if there are sudden changes in the quality or availability of the data actively assimilated by the ECMWF four-dimensional variational data assimilation system (4DVAR). The checking system was initially limited to satellite observations, but last year it was extended to cover in-situ measurements. More recently, the system has been complemented with an automatic checking of persistent quality improvements of in-situ data that are blacklisted (not used because of previously poor or unknown data quality) in the Integrated Forecasting System (IFS), for potential future use in the analysis.

In this article a brief description of the automated checking process is first provided. It is followed by a description of the system configuration for satellite and in-situ observations. Thoughts about the potential of the system to improve the data assimilation and model diagnostics are then discussed. The concluding section presents the status of the system validation and the on-going work to improve it.

Automatic checking process

For each set of observations, selected statistical quantities (e.g. number of observations, bias correction, mean bias-corrected innovations (i.e. observation minus background) and analysis departures) are checked against an expected range. An appropriate alert message is generated if statistics are outside the specified limits. A severity level ('slightly', 'considerably', 'severely' or 'severely persistent') is then assigned to each message depending on how far statistics are from the expected values. The automatic checking uses two kinds of ranges: soft and hard limits (see Figure 1 and Box A).

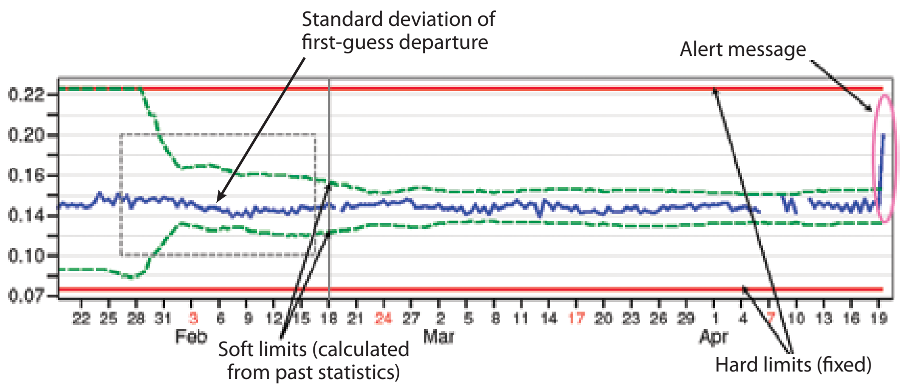


Figure 1 Example of an alert for ATMS Channel 9 produced at 12 UTC on 19 April 2013.

Automated checking system limits

A

Soft limits

Soft limits are updated automatically using statistics from the last twenty days (the last two days and extremes are excluded during this process). The soft limits are calculated as:

upper soft limit=mean+A x stdev

lower soft limit=mean-A x stdev

where *stdev* is the standard deviation. In the current configuration A is set to 5.

Hard limits

The hard limits are estimated during the process of adding new data to the alarm system. They are adjusted occasionally when a drift is noticed or during IFS upgrades. The hard limits are updated manually when needed.

Severity levels

The classification into *'slightly'*, *'considerably'*, *'severely'* and *'severely persistent'* depends on the expected variability and is as follows.

- *Slightly*: outside ±5 *stdev* from the mean.
- *Considerably*: outside ±7.5 *stdev* from the mean.
- *Severely*: outside±10 *stdev* from the mean.
- *Severely persistent*: Severe problems occurring frequently during the past 10 days.

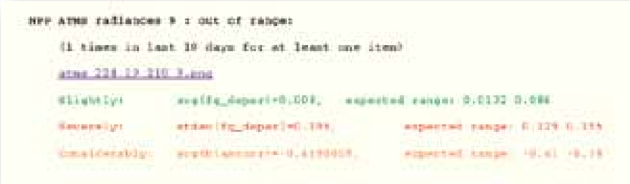


Figure 2 Example of an automatic alert message for ATMS Channel 9 triggered at 12 UTC on 19 April 2013.

Soft limits are designed to detect sudden changes in statistics, whilst hard limits are used to detect slow drifts in statistics. The automatic alarm system uses an 'ignore' facility to filter out warnings for data with known problems (e.g. expected outages or orbital manoeuvres). For each alert message (apart for data missing events) the alarm system generates a relatively long time series providing a comprehensive view of the time evolution of the various statistical quantities (e.g. innovations, counts and bias correction). When a problem is detected, the system includes the number of times the same issue has been detected during the past ten days (see Figure 2).

Alert messages are sent by email to subscribed users (mainly from within ECMWF with the exception of few external users) according to their registered preferences (i.e. data types of interest and levels of severity needed). The warnings are also published on the ECMWF monitoring web pages, which have public access (*http://www.ecmwf.int/forecasts/see-our-forecasts/medium-range-forecast-charts/monitoring-observing-system*). To help with follow-up investigations, warnings are archived in a relational events database allowing conditional extraction of alerts (e.g. to find possible links with forecast busts or weather events). The events database is currently not publically available.

Data Types	Statistical quantities
Radiances	• Data counts • Average of innovations • Standard deviation of innovations • Bias correction
AMVs	• Data counts • Average of innovations • Standard deviation of innovations • Average pressure (pressure of the assigned height) • Standard deviation of pressure
Scatterometer (wind speed)	• Data counts • Average of innovations for the best ambiguity • Standard deviation of innovations for the best ambiguity
Ozone	• Data counts • Average of innovations • Standard deviation of innovations • Bias correction
GPS Radio Occultation	• Data counts • Average of normalized innovations • Standard deviation of normalized innovations

Table 1 Current setup for the automatic checking of satellite data.

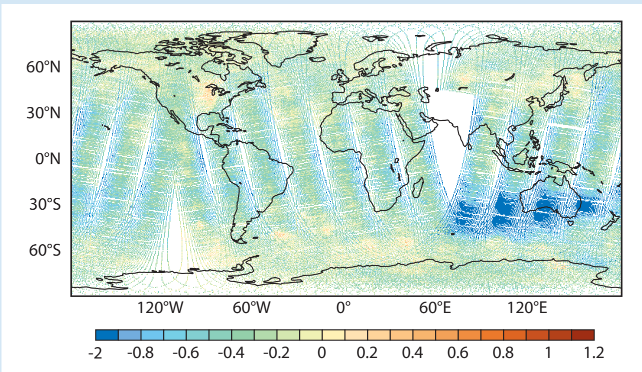
The automated checking system is designed for ECMWF internal use. Results are based on feedback information from the ECMWF data assimilation system and therefore they reflect ECMWF's data usage. Alerts published on the ECMWF website are provided for information on an as-is basis. This information can potentially be used by other NWP centres as additional information that can be compared to their own monitoring.

Lunar intrusion event

B

On 19 April 2013, the automatic checking system triggered severe warnings for several NPP/ATMS channels. The alerts were caused by a sudden sharp increase in the innovations standard deviation (see the plot below for ATMS channel 9 innovations covering 09 to 21 UTC on 19 April 2013) due to the moon intrusion into the ATMS field of view used for the calibration of the instrument.

To filter out contaminated data, ECMWF was relying on the Quality Control flags provided with the data. The algorithm used to detect and flag moon intrusion events was evolving at the time. It has now reached a level of maturity allowing the detection and flagging of most moon intrusion events.



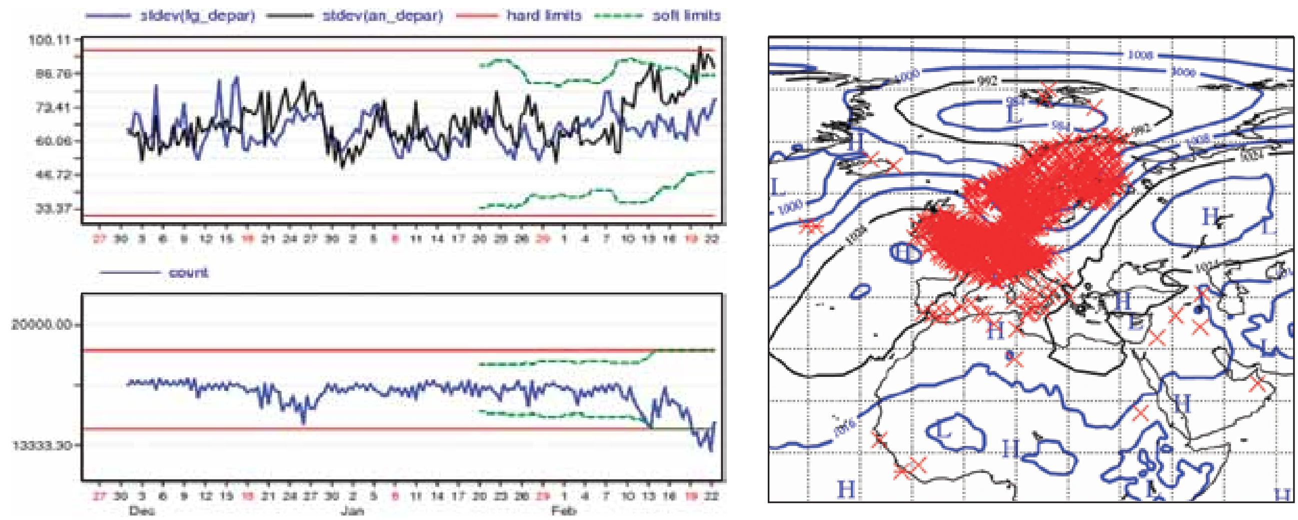


Figure 3 Left: Statistics for SYNOP surface pressure over Europe: standard deviations of first-guess departure [stdev(fg_depar)] and analysis departure [stdev(an_depar)] along with hard and soft limits and observation count. Right: surface pressure observations with low variational quality control weight for a 12-hour 4DVAR 'delayed cut-off data assimilation' cycle in February 2012.

Alarm system for satellite data

For satellite data, potential quality and availability issues are mainly related to the instrument, ground segment or telecommunication problems. As the same satellite instrument covers a large area, any quality issues affect the assimilation on a large, often global scale which makes it especially important to detect such issues as quickly as possible. For satellite data, we check separately global statistics from each satellite, instrument, channel, parameter and statistical quantities. The only exceptions to this are the Atmospheric Motion Vectors (AMVs), which are averaged over the total vertical column. The system offers the flexibility to add new statistical quantities. The current setup is summarized in Table 1.

An example of an alert is given in Figures 1 and 2, showing a sudden increase in the standard deviations of innovations for a channel on ATMS that is used to retrieve profiles of atmospheric temperature and moisture. Further investigation showed that this was caused by temporary calibration problems: for certain parts of some orbits, the moon entered the space view used for the calibration of the instrument, leading to incorrectly calibrated observations (see Box B). Flagging of such data was subsequently introduced in the ECMWF system, and the data providers also updated their processing to mitigate such problems.

Alarm system for in-situ measurements

Unlike satellite observations, in-situ measurement issues are typically specific to each individual station. However, on some occasions widespread issues might be caused by data routing problems, significant weather events, model errors or data assimilation issues. To cover both aspects, the automatic data checking system has been extended to perform two kinds of automatic check.

Individual stations

For each assimilation cycle, the system checks the quality of available stations based on the estimated Probability of Gross Error (PGE), the mean and root-mean-square of innovations and bias correction. The parameters checked are the temperature, pressure, specific humidity,

Widespread intermittent rejections of surface pressure over Europe

C

Since the IFS resolution upgrade in 2010 and throughout 2011, the data assimilation system was exhibiting intermittent widespread rejections of surface pressure observations over Europe. After investigations it was found that the root-mean-square analysis fit to SYNOP surface pressure observations over Europe was poor. In particular, it was worse than that of the background, indicating that 4DVAR was failing to draw to these observations. The root cause of the problem was related to differences in the propagation of gravity waves between the inner and outer loops of the 4DVAR. These differences are a result of the different time steps used in the inner and outer loops. The investigations led to a successful implementation of a solution that suppresses the generation of gravity waves with a strong surface pressure component.

The automatic data checking of in-situ data was not operational at the time. However, as a result of this case we added (for all in-situ measurements) a new test comparing the standard deviation of innovations against the standard deviation of analysis departures. An alert message is triggered when the standard deviation of innovations is smaller than the one of analysis departures.

relative humidity and wind vector difference. PGE values are provided directly by the 4DVAR for all 'used' data. The automatic system calculates the percentage of in-situ reports with PGE above a pre-defined threshold (currently 0.75). It then triggers a warning if this percentage is high or if there are significant changes in the other statistical quantities. The system performs the same check over a period of ten days. Any such issues are flagged as 'severely persistent'.

Main data types over a number of pre-defined geographical areas.
This component of the system detects widespread issues. It follows the same method as applied for satellite data but with an additional test comparing the standard deviation of innovations and analysis departures. Such a test is important to highlight areas and situations where the model or the analysis is not performing well. One analysis-related example is the intermittent widespread rejection of surface pressure observations over Europe in 2012 (see Figure 3 and Box C). The system currently monitors data from:

- Nine predefined geographical areas: North America, Europe, North Atlantic, Asia, South Pacific, South America, Africa and Indian Ocean. It offers the flexibility to add new geographical areas.
- Ten data types: SYNOP, METAR, SHIP, AMDAR, ACARS, AIREP, TEMP, PILOT, DRIBU and PROFILERS.

Automatic detection of improvements

Currently in-situ measurements are added to a blacklist when affected by frequent quality issues. Removing improved measurements from the blacklist requires an assessment of their behavior (e.g. using monthly values of the root-mean-square of innovations). These statistics are checked against fixed values ignoring the dependency of departures on geographical areas and significant weather events.

To make the procedure automatic and efficient, the data checking system has been extended to detect persistent improvement in the quality using a PGE-like quantity. For each flagged station, the system produces a time series plot that will help in making the final decision. The plot includes the estimated PGE, mean and root-mean-square of innovations, data counts and spread of the Ensemble of Data Assimilations (EDA).

Potential for model performance monitoring

The automatic data checking system was primarily designed to detect data-related issues. However, since the quality assessment is mainly based on innovations, there is a potential to detect situations where the forecast model or the data assimilation itself has weaknesses. This can be the case in the following situations.

- The system triggers alerts for independent satellites/instruments that provide similar observations of the atmosphere. Large departures associated with independent data suggest that the forecast model is not able to fully capture the phenomena being observed. With the increase in satellite data sources, there is a potential to improve the diagnostics of the model and data assimilation algorithms. As an example, on many occasions the system triggered alerts associated with microwave and infrared satellite upper-stratospheric channels. These alerts were caused by sudden stratospheric warming episodes where the initial rapid circulation changes were not well captured by the model, leading to larger than usual differences between the model first-guess field and the observations.
- Widespread warnings affecting in-situ measurements (apart from availability issues) are likely to be related to the forecast model or the data assimilation (see example in Figure 3).
- Individual warnings occurring in the vicinity of dynamic atmospheric situations need to be investigated. They can potentially point to model or data assimilation issues (e.g. a weather system moving too slowly in the model).

Alarm system validation and future developments

For the last two years, the automatic alarm system has been running operationally at ECMWF. Although the results are not being verified objectively, the alarm system has reached a good level of maturity allowing the detection of almost all severe data quality issues (according to the daily data monitoring and other external notifications). However, the system still produces many 'slight' warnings that are related to small changes in statistics. Such warnings are still being distributed to internal users to keep track of all data-related changes. The system is performing well by keeping the rate of 'severe' false alarms to a very low level.

The system is fully automatic in detecting issues but at the moment it does not act automatically on any other component of the data assimilation. The results from the system are used to manually update the blacklist, but in future some kind of semi-automatic updating of the blacklist might be possible.

Work is on-going to improve the reliability of the system and explore further its potential for improving the daily monitoring of the operational data assimilation and forecasting system at ECMWF.

All-sky assimilation of microwave humidity sounders

ALAN GEER, FABRIZIO BAORDO,
NIELS BORMANN, STEPHEN ENGLISH

Satellite radiance observations give dense, frequent, global coverage of the atmosphere but they are difficult to use in cloudy areas. In the last decade ECMWF has pioneered the assimilation of cloud and precipitation-affected satellite observations in operational numerical weather forecasting. The implementation of Cycle 40r3 of the Integrated Forecasting System (IFS) will mark a milestone in this development: for the first time all the main humidity-sensitive microwave instruments will go through the 'all-sky' assimilation path, which allows us to exploit water vapour information in cloudy areas and to directly assimilate cloud and precipitation (see Box A).

All-sky assimilation of microwave imager observations, sensitive mainly to the lower troposphere, has been operational since 2009 (*Bauer et al.*, 2010). The new addition is the microwave humidity sounding channels at 183 GHz which are a core capability of the European and US polar-orbiting operational meteorological satellites. These channels were previously assimilated using the 'clear-sky' method: in other words, all cloud-affected observations had to be discarded. All-sky assimilation allows a more complete use of the data, particularly in the winter storm tracks.

The direct information content of microwave humidity sounders is humidity, cloud and precipitation, but when they are used inside any four-dimensional variational data assimilation system (hereafter referred to as 4D-Var)

Clear-sky and all-sky assimilation

A

When doing clear-sky assimilation there is no attempt to model the effect of cloud or precipitation. At the observation location, an assimilation system can only improve the fit to clear-sky observations by modifying the water vapour and temperature fields (though as explained in the main text, a 4D-Var system will have the possibility of doing this by adjusting the surrounding dynamical fields). Cloud-affected observations are discarded because cloud would otherwise be aliased into spurious water vapour or temperature increments, leading to degradation of the forecasts. Drawbacks of this technique are:

- Cloud detection may not remove all cloud-affected observations and the forecasts may be degraded anyway.
- The quality control is asymmetric: if cloud-affected observations are removed, the remaining observations are representative only of cloud-free areas. In the case of humidity observations this will lead to a dry bias.
- Observations have to be discarded in the most meteorologically interesting areas, such as frontal systems and mid-latitude storm tracks.

All-sky assimilation uses an observation operator capable of simulating the effect of cloud and precipitation. At the location of the observation, the assimilation system can adjust water vapour, temperature, cloud or precipitation to fit the observation better. No cloud screening is applied and all observations, whether clear, cloudy or precipitating, are assimilated in the same way using the same observation operator.

Three key components in an all-sky approach are:

- In a 4D-Var system, there must be a tangent linear and adjoint representation of the moist physical processes in order to make use of cloud and precipitation information coming from the all-sky radiances. These key pieces of the system were put in place at ECMWF by the Physical Aspects Section in the middle of the last decade (see *Janiskova & Lopez*, 2012).
- There must be an observation operator capable of representing cloud and precipitation effects on the observed radiances. To use cloudy microwave observations we have had to develop radiative transfer models that can represent the scattering of radiation from rain and frozen particles. At the same time, these models need to be fast enough to be used in an operational NWP system. A key recent development leading to the use of microwave humidity channels in the 183 GHz range has been better modelling of the scattering from frozen particles using non-spherical particle models, for example snowflakes (*Geer & Baordo*, 2014).
- The quality of first-guess cloud and precipitation is still much worse than humidity and temperature. In fact, it could be argued that at the smaller scales, and particularly when it comes to convection, cloud and precipitation are not even representable by current models – in other words, the model cannot be expected to supply a precise location, timing and intensity for cloud and precipitation. In all-sky assimilation this is addressed as a 'representivity error' by applying smaller observation errors in clear-sky areas and larger observation errors in cloud and precipitation (*Geer & Bauer*, 2011). For relatively high-spatial-resolution instruments like microwave imagers (resolutions down to 10 km) the observations are also averaged together (superobbed) onto an 80 km spatial scale to help remove small-scale cloud and precipitation variability. Microwave humidity sounders have relatively large spatial scales (40 to 60 km) and are not superobbed.

they also provide information on wind and temperature. Particularly in the mid-latitude upper troposphere, strong humidity and cloud gradients offer a chance to infer wind and mass information from the moisture fields. Indeed, the latest upgrade has its main benefit in the wind and geopotential forecasts, which are significantly improved out to at least day 4. We can see the microwave humidity observations as an important part of the observing system, improving the large-scale analysis of wind and mass and benefitting forecast scores into the medium range.

Though it is good that forecasts are being improved, it is important to know why it is happening. It could be that we now have better geographical coverage of upper-tropospheric water vapour in dynamically interesting cloudy regions. The assimilation of cloud and precipitation could itself be benefitting the forecasts. More crucially, we need to know if we are directly inferring winds from the humidity, cloud and precipitation fields or whether the benefit comes through some more subtle aspect of the data assimilation system. All-sky assimilation may be beginning to derive information that has historically come from cloud motion vectors. The more we understand the all-sky assimilation, the more we will know how best to use satellite observations in the future – whether as atmospheric motion vectors, clear-sky or all-sky radiances. In this article we make a start on investigating these questions.

Microwave humidity sounding around 183 GHz

Figure 1 shows the observations from the microwave humidity sounder (MHS) on the European Metop-B satellite. These are the observations available in the 00 UTC assimilation window on 15 August 2013, during the southern hemisphere winter where synoptic activity in the southern storm track is at its height.

Figure 1a shows MHS channel 3 (at 183±1 GHz), sensitive to upper-tropospheric humidity, cloud ice and frozen precipitation. In the absence of cloud or precipitation, low relative humidity corresponds to a high observed brightness temperature and vice-versa. The climatological subtropical high pressure regions are marked by high brightness temperatures indicating a very dry free-troposphere. In the mid-latitudes and in the inter-tropical convergence zone (ITCZ) there is much higher relative humidity and the brightness temperatures are relatively low. Also, in the mid-latitudes there is great variability and a mixture of dry and moist air masses. The great hope of humidity sounding is that the movement, position and water amounts of these air masses can be used to infer information on the wind fields and ultimately on the large-scale synoptic structures of the atmosphere (e.g. *Andersson et al.*, 1994, *Allen et al.*, 2014). Where the wind is aligned with a gradient in humidity, the model's humidity advection term should allow a 4D-Var assimilation system to infer information on the wind fields. Additionally, the effect of the continuity equation on the water vapour amounts can provide information on the convergence and divergence of the winds.

The all-sky brightness temperatures in Figure 1a include cloudy and precipitating scenes. These are hard to

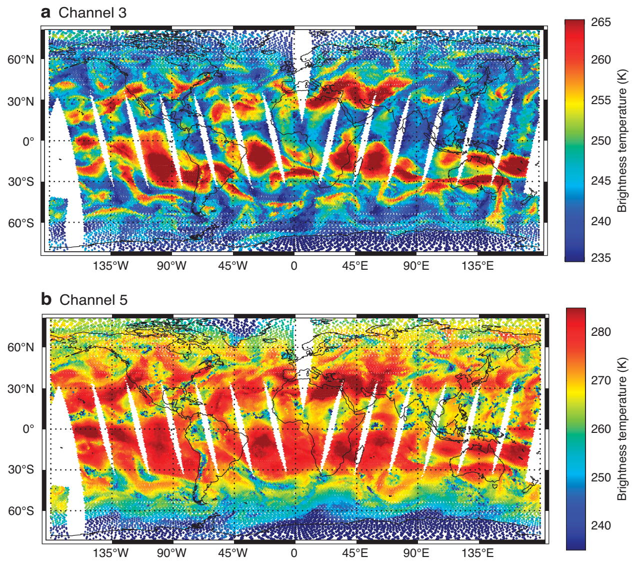


Figure 1 Observed brightness temperatures from Metop-B MHS in the 00 UTC assimilation window on 15 August 2013: (a) Channel 3, at 183±1 GHz, sensitive to upper-tropospheric humidity, cloud ice and frozen precipitation; (b) Channel 5, at 190 GHz, sensitive to mid-tropospheric humidity, cloud ice and frozen precipitation. High brightness temperatures correspond to low humidity; low brightness temperatures correspond to high humidity, cloud ice or frozen precipitation (typically deep convection).

distinguish, at least in a global view. At the microwave frequencies used by MHS, cloud ice and frozen precipitation cause scattering, which reduces brightness temperatures. High relative humidity also reduces brightness temperatures: without additional information it is hard to distinguish cloudy or deep-convective areas from those which just have high relative humidity.

Figure 1b shows MHS channel 5 (at 190 GHz), which peaks lower in the atmosphere, sometimes low enough to observe the surface. Again humidity is the main factor controlling the observed brightness temperature, but cloud and precipitation are more easily distinguished against the warmer clear-sky background: spots of localised low brightness temperature (the dark blue colours) indicate deep-convective systems or broader areas of thick ice cloud, for example in the ITCZ and in mid-latitude frontal zones.

Single cycle example

To illustrate the effect of assimilating microwave humidity channels in all-sky conditions, we can perform a single-cycle assimilation experiment and compare the increments generated by the all-sky observations on their own to those from the full observing system (including all-sky humidity observations). In the all-sky-only experiment we will assimilate humidity data from five microwave humidity sounders on various satellites (i.e. MHS on Metop-A, Metop-B, NOAA-18, NOAA-19 and SSMIS on DMSP-F17). The experiment is performed for the 00 UTC assimilation window on 15 August 2013, the same date as the observations we examined in the last section.

Figure 2 shows the normalised all-sky first-guess departures (first-guess departure divided by observation error) corresponding to the observations from MHS channels 3 and 5 in Figure 1. This is a fundamental quantity in the data assimilation system because, when squared, it gives the contribution of that observation to the 4D-Var cost function. In other words, the normalised first-guess departure is a guide to the influence of that observation in the data assimilation, though its ultimate effect in the analysis is also controlled by the background error term. The largest normalised departures are in the mid-latitude storm tracks and in the ITCZ: hence these are the areas where we might expect to have greatest influence on the analysis.

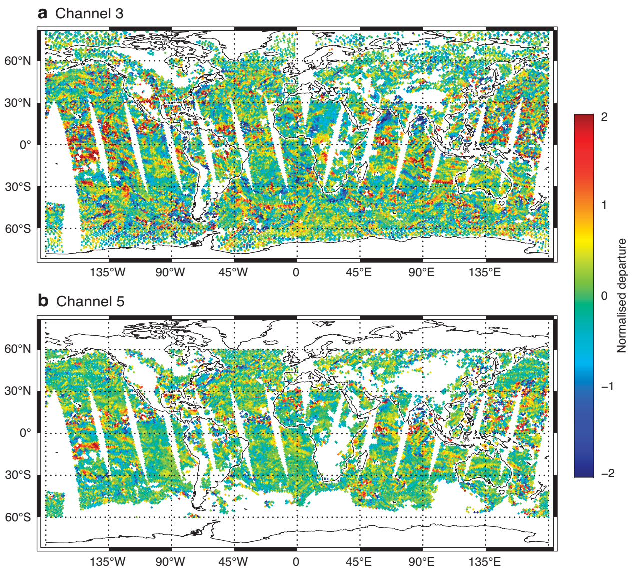


Figure 2 Normalised, bias corrected first-guess departures from Metop-B MHS in the 00 UTC assimilation window on 15 August 2013: (a) Channel 3, sensitive to the upper troposphere; (b) Channel 5, sensitive to the mid or lower troposphere. Normalised first-guess departures are computed as the observation minus the first guess divided by the observation error. Squared, these would give the observation's weight in the 4D-Var cost function; as they are, they retain information on the direction in which the analysis should move: either moistening or drying the model.

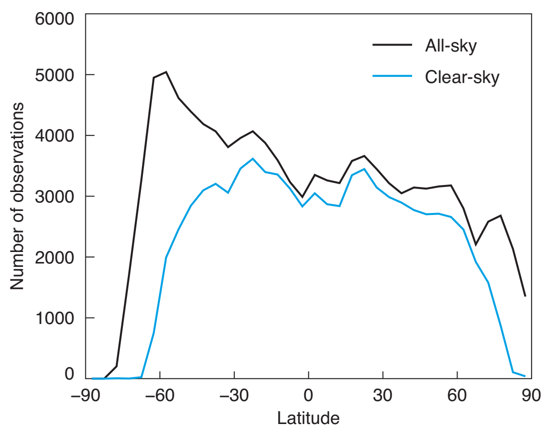


Figure 3 Number of MHS channel 3 observations actively assimilated, per 5° latitude bin, in 00 UTC assimilation window on 15 August 2013 for the all-sky approach and if clear-sky quality control is applied to remove cloudy scenes.

Only active observations are shown, so it is clear that in the lower-peaking channels the quality control removes a lot of data. We cannot yet assimilate these observations over high orography, sea-ice or at high latitudes. Here, there is too much influence from surface emission, which is difficult to model over land, snow and ice surfaces. We also have to remove the 'cold air outbreak' regions where a systematic lack of supercooled cloud liquid water in the forecast model causes a difficult-to-correct bias between model and observations. Overall, we might expect the upper-tropospheric channel to have the greatest impact on the data assimilation because it has a much greater geographical coverage.

Looking closely at the departures in the tropics, there are many positive and negative regions corresponding to displacements in the position of large convective systems in the ITCZ. It is still very hard for the model to simulate these features in the right place and at the right time, even in the 12-hour first-guess forecast. In the mid-latitudes, we can see elongated regions of negative and positive departures with a width of the order 100 to 300 km and lengths up to 1000 km. These correspond to displacements and intensity variations in mid- and upper-tropospheric humidity and in the cloud fields. This is the information we hope will lead back to the wind fields and ultimately to improvements in the analysis of the large-scale synoptic situation.

One of the key advantages of the all-sky approach is simply its enhanced coverage in meteorologically interesting areas.

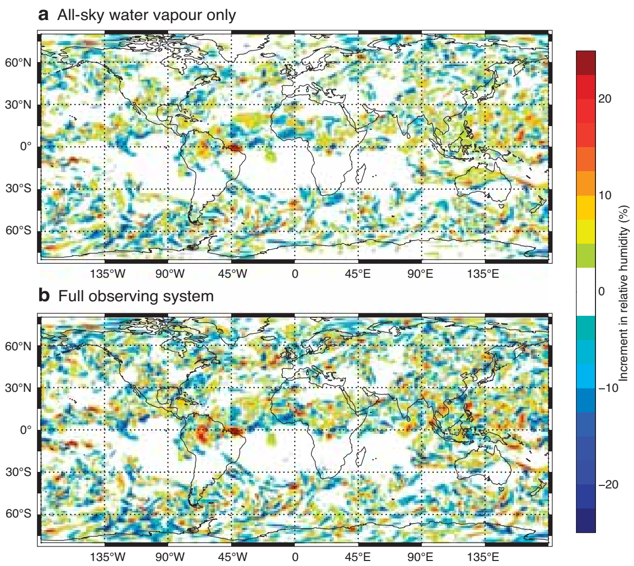


Figure 4 Increments in relative humidity on model level 95 (the closest level to 500 hPa) at 06 UTC on 15 August 2013: (a) assimilating only the all-sky microwave humidity observations; (b) assimilating the full observing system, including all-sky humidity channels. The first guess is identical in both cases and is created using the full observing system. Correlation between panels (a) and (b) is 72%.

Figure 3 shows the number of MHS channel 3 observations that can be assimilated in this analysis using either a clear-sky or an all-sky approach. In either case, coverage is lowest in the northern hemisphere, mainly because high altitude land surfaces need to be discarded. But polewards of 50°S, across the southern hemisphere storm tracks, the all-sky technique provides at least double the number of observations. In the clear-sky technique, cloud screening removes a majority of observations in the winter high latitudes.

On its own, the all-sky humidity sounder assimilation can generate mid-tropospheric increments that replicate a substantial portion of those from the full observing system. Figures 4 and 5 show the humidity and wind increments 9 hours into the assimilation window. The humidity increments resemble the normalised departures from Figure 2 in terms of their size and morphology, though they cannot be compared quantitatively because the satellite observations are valid at different times through the assimilation window and are sensitive to a broad layer of humidity, not just a single model level. However, we do know that the analysis fits the all-sky observations much better than the first guess and that this fit is very similar whether the full observing system is used or just the all-sky water vapour data (not shown). So these humidity increments are being generated specifically to fit the all-sky observations. The morphology of the wind increments is qualitatively consistent too, on similar 100–400 km scales, and in similar areas of the globe: these wind increments will be associated with displacements, stretching and compression of the humidity field to make it fit the observations.

It is surprising how well the all-sky-only increments resemble those from the full observing system. There are plenty of other areas – stratosphere and lower troposphere included – where the all-sky-only assimilation does not replicate the full system. But all-sky assimilation on its own can generate mid- and upper-tropospheric wind increments that are reasonably consistent with those from the full observing system.

Forecast scores and observation fits

Assimilating the all-sky observations on their own is an unrealistic scenario. What really matters is whether the new technique can improve forecasts in the context of the full observing system: we need to know the incremental benefit of adding all-sky assimilation of microwave humidity sounders.

Adding clear-sky or all-sky assimilation to the otherwise full observing system improves standard 500 hPa geopotential height forecast scores out to at least day 4 (Figure 6). Beyond day 4, even though the error bars suggest there is impact, we should be cautious on these scores given the relatively short duration of these experiments (a combined total of around four months). At shorter ranges, the all-sky assimilation significantly outperforms the clear-sky assimilation and brings the total improvement from water vapour sounding channels to around 2% in forecast scores. This may sound small, but the improvement in ECMWF forecast scores over the last 10 or 20 years amounts to only around a percent annually. In the full observing system, there is a substantial amount of often overlapping

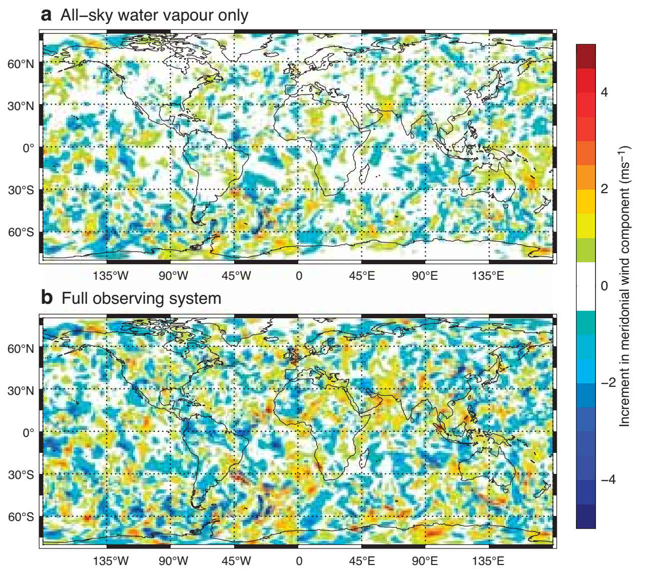


Figure 5 Increments in meridional wind component on model level 95 (the closest level to 500 hPa) at 06 UTC on 15 August 2013: (a) assimilating only the all-sky microwave humidity observations; (b) assimilating the full observing system, including all-sky humidity channels. The first guess is identical in both cases and is created using the full observing system. Correlation between panels (a) and (b) is 58%. Increments in the zonal wind component show qualitatively similar patterns.

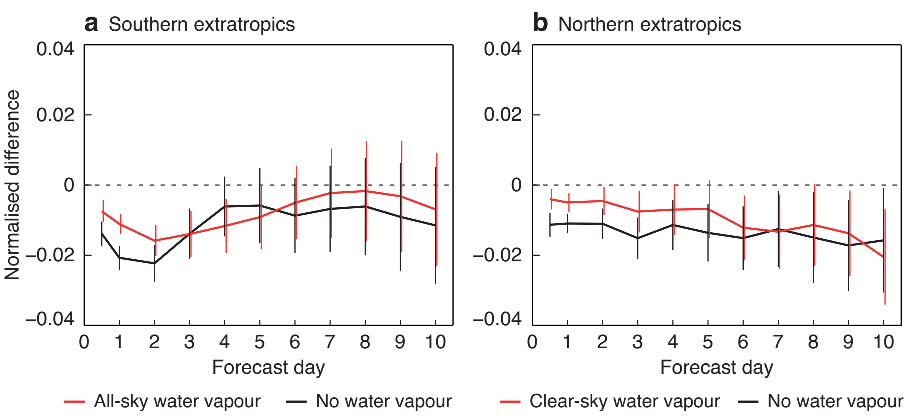


Figure 6 Normalised change in root-mean-square geopotential error at 500 hPa between the assimilation of clear-sky or all-sky microwave water vapour sounding observations and the otherwise full observing system in the (a) southern and (b) northern extratropics, based on a maximum of 236 forecasts from summer (August/ September 2013) and winter (January/ February 2014) experiments. Error bars indicate the 95% confidence range; verification is against own analysis.

information. New techniques and new instruments typically fill small gaps or add just a little more precision to what is already a very good analysis.

For the 48-hour forecast there is over 10% reduction in the wind errors in the upper troposphere in the southern mid-latitudes when the all-sky observations are added to the otherwise full observing system (Figure 7). Clear-sky assimilation of the same data gives impact in similar places, but in general it is less (not shown).

First-guess fits to assimilated observations confirm the short-range forecast scores (Figure 8). Clear-sky assimilation of microwave water vapour channels is beneficial to temperatures and winds, represented by AMSU-A observations and conventional wind data (Figures 8a and 8b respectively). Compared to clear-sky, all-sky improves the quality of wind fits to conventional observations between 50 hPa and 250 hPa. Temperature-sounding AMSU-A data is better fitted across channels spanning the troposphere (channels 5 to 9) and to a lesser extent the stratosphere (channels 10 to 14). Interestingly it is the relative humidity fits that see least benefit from the move to all-sky assimilation. Channels 11 and 12 of HIRS (Figure 8c) are infrared channels that are sensitive to mid- and upper-tropospheric humidity in a very similar way to the microwave humidity sounding channels at 183 GHz. Clear-sky assimilation already brings substantial improvements in upper-tropospheric humidity; all-sky improves the situation by a smaller margin. However, it is likely the wind and temperature improvements that are most important in improving forecast quality. In other words, the all-sky assimilation seems to have particular benefits to the dynamical fields that cannot be so easily obtained from the clear-sky approach.

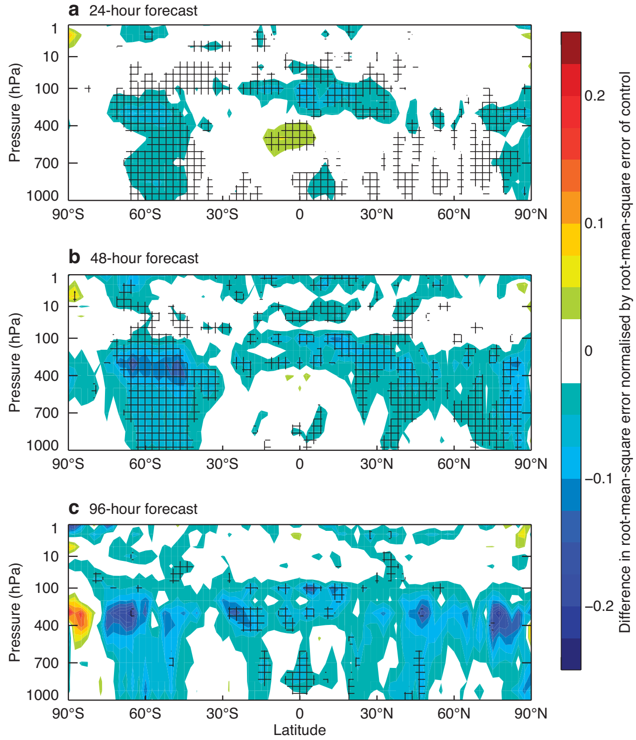


Figure 7 Normalised change in root-mean-square vector wind error for (a) 24-hour, (b) 48-hour and (c) 96-hour forecasts when all-sky microwave water vapour assimilation is added to an otherwise complete global observing system, based on a maximum of 236 forecasts from summer (August/September 2013) and winter (January/February 2014) experiments. Cross-hatching indicates statistical significance at 95%; verification is against own analysis.

Impact in the absence of other observations

To further understand how the all-sky assimilation benefits the forecasts, the single-cycle, single-observing system experiments were repeated twice-daily over a period of 16 days. This is a framework in which the first guess is always of the highest possible quality, coming from the full observing system, and the background errors are correctly specified. These experiments answer the question "what would

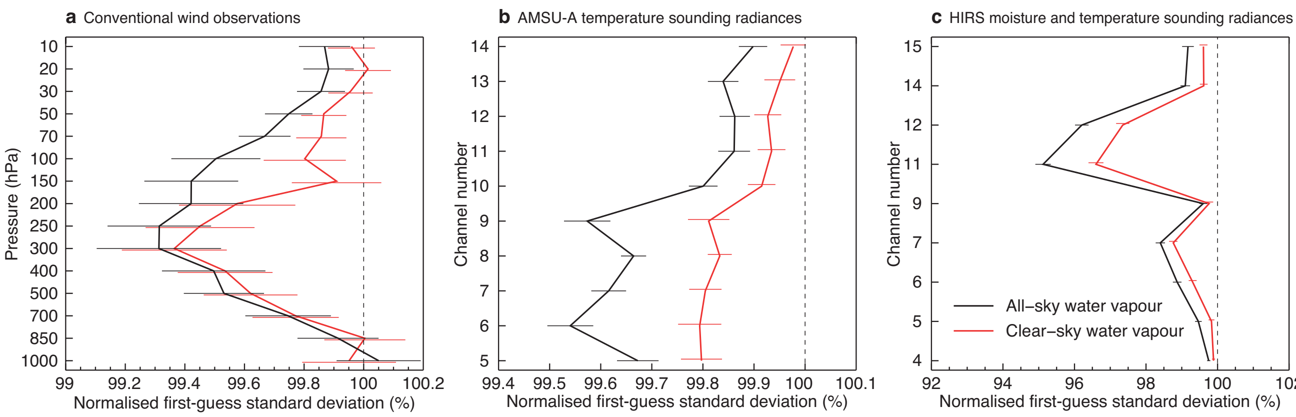


Figure 8 Normalised standard deviation of first-guess departures for: (a) conventional meridional and zonal winds from aircraft, profilers, pilot balloons and radiosondes; (b) temperature-sensitive AMSU-A observations (combining instruments on seven different European and US polar orbiting satellites); (c) HIRS infrared observations, with channels 11 and 12 sensitive to upper-tropospheric water vapour (on the Metop-A satellite). The departures are normalised by the results of the full observing system excluding microwave humidity observations, so that 100% corresponds to the full observing system experiment. The area is global and the period is a combination of August/September 2013 and January/February 2014.

happen if for 12 hours we lost all observations except...". In this set of experiments we have two reference points: (a) the quality of forecasts when no observations are assimilated and (b) the quality of forecasts when all observations are assimilated. We can create a metric that puts these points at 0% and 100% impact respectively. If forecast quality were degraded compared to the no data case, this 'impact' would be a negative number.

Figure 9 shows the impact of assimilating only clear-sky microwave humidity channels. Mid- and upper-tropospheric humidity errors are reduced by 50% to 80% of that generated by the full observing system. As we saw in the fits to observations (Figure 8) clear-sky assimilation benefits the wind and temperature fields too. In Figure 9, the reduction in wind errors is largest at 400 hPa, roughly consistent with the fit to conventional wind observations in Figure 8. In the absence of other observations, clear-sky water vapour assimilation can replicate about 40% of the impact of the full observing system in these upper-tropospheric winds.

The impact of all-sky humidity assimilation in the full observing system is shown in Figure 10. In much of the tropics and northern hemisphere, impact is similar to clear-sky, and in the relative humidity, it is even a little less. Only in the mid-latitude southern hemisphere is the impact substantially larger. Here, all-sky humidity observations can reproduce up to around 60% of the impact of the full observing system on winds in the upper troposphere and 40% in temperature. The impact is greater also around the tropopause, for example in tropical winds at 100 hPa and in extending the zone of impact in the mid-latitudes up a little – perhaps from 300 hPa to 200 hPa. In terms of latitude, the region where all-sky produces most benefit is from 30°S to 70°S – exactly the zone where all-sky assimilation is able to add substantially more observations in these examples (Figure 3). These experiments are for the southern hemisphere winter, but in the northern hemisphere winter, more impact occurs in the northern storm tracks (not shown) and this is also reflected in the score improvements in northern latitudes in Figure 7, which combines results from August/September 2013 and January/February 2014.

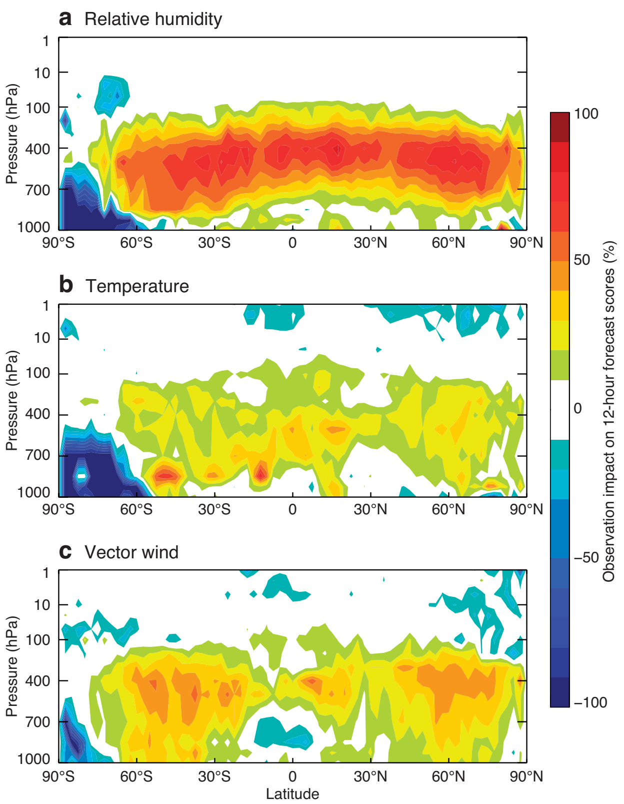


Figure 9 Percentage reduction of 12-hour forecast error in (a) relative humidity, (b) temperature and (c) wind when only clear-sky microwave water vapour sounding channels are assimilated, based on the period 15–31 August 2013. The assimilation is not cycled; the first guess comes from a full-observing system experiment. The quality of a forecast from the first guess (i.e. without any data assimilation) defines 0%; the quality of a forecast based on an analysis with the full observing system defines the 100% level. Clear-sky microwave assimilation can on its own replicate about 70% of the impact of the full observing system on mid- and upper-tropospheric relative humidity.

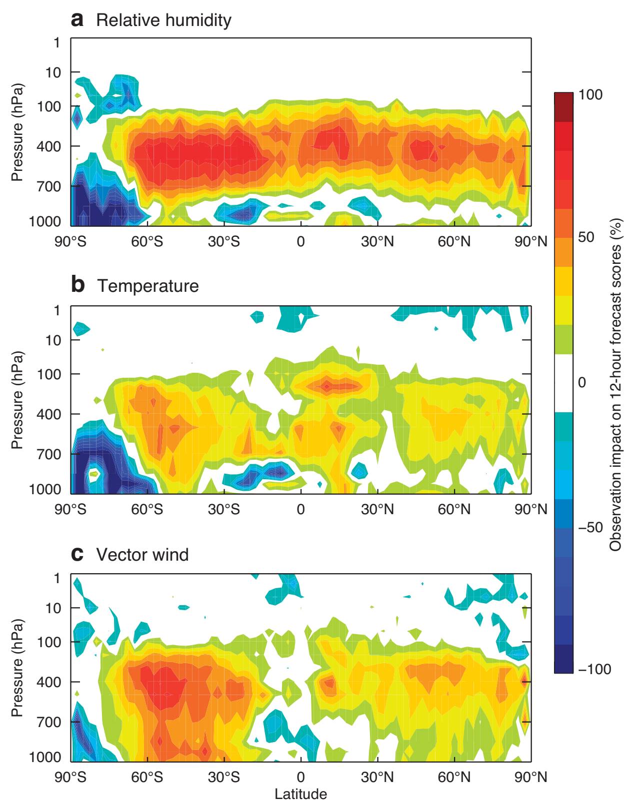


Figure 10 As Figure 9 but for all-sky assimilation of the microwave water vapour sounding observations. All-sky is comparable to clear-sky in the tropics and northern hemisphere but it is substantially better in the southern hemisphere, particularly in reducing temperature and wind errors.

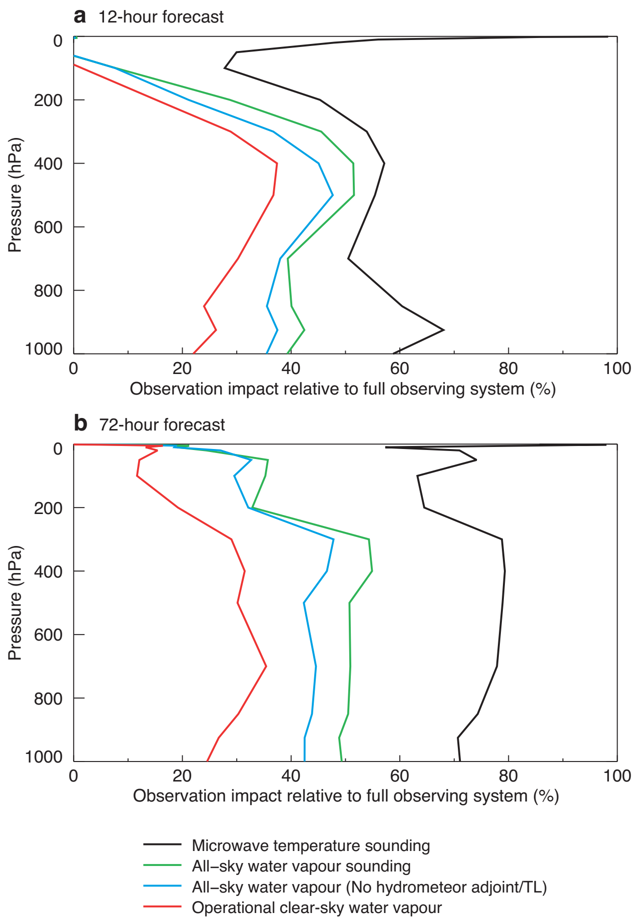


Figure 11 Southern hemisphere extratropical error reduction in vector wind error for (a) 12-hour and (b) 72-hour forecasts for different observing systems as a percentage of the impact from the full observing system. Results are given for microwave temperature sounders (AMSU-A, ATMS), all-sky water vapour sounders (MHS, SSMIS), all-sky water vapour sounders with adjoint and tangent linear (TL) sensitivity to hydrometeors (cloud and precipitation) turned off, and operational clear-sky water vapour sounders (just MHS).

A minor issue is the degradation of forecast quality in the Antarctic at low levels in the microwave humidity assimilation. Partly, the issue is that this is a non-existent part of the atmosphere; ground levels are as high as 3,000 m and surface pressure can be at around 600 hPa to 700 hPa. Even at levels where this may represent a true degradation in forecasts, it may simply be that the fact that we do not assimilate microwave humidity observations in these areas and the unconstrained assimilation system is making spurious increments. In the full observing system, other observations, such as the microwave temperature sounders, can constrain these areas.

Looking at the southern hemisphere extratropics where the all-sky observations have greatest impact in this experiment, Figure 11 shows the impact on the 12-hour and 72-hour forecasts. As a reference point, this figure includes the impact of all the microwave temperature sounders, which are probably the single most important group of observations in our system. These can replicate around 50% to 60% of the full observing system's impact on temperature and wind fields, with benefits not just in the mid and upper troposphere but in the lower troposphere and stratosphere too. All-sky assimilation maintains its advantage over the clear-sky assimilation at least out to 72 hours, and brings the impact towards that of the microwave temperature sounding observations.

An important question is whether the benefits of all-sky assimilation come through constraining the water vapour fields in the presence of clouds or through the cloud fields themselves. A crude way to test this is to turn off the adjoint and tangent linear sensitivity to cloud and precipitation in the observation operator used in 4D-Var. Having done this, the assimilation system can only directly improve the fit to all-sky observations by changing the moisture fields at the observation location. The assimilation system is prevented from directly adjusting clouds and precipitation to fit the observations (though this may be a natural secondary result of improving the humidity). Figure 11 shows that turning off the cloud and precipitation sensitivity in the minimisation has relatively little impact. It looks like the majority of impact in going from clear-sky to all-sky assimilation is coming from a better constraint on relative humidity in the presence of cloud, rather than directly through the sensitivity to cloud and precipitation. Nevertheless, the cloud and precipitation sensitivity is clearly beneficial.

Summary

Clear-sky assimilation of microwave humidity sounders improves forecasts, but all-sky assimilation is better, particularly in the extratropics. Mid-latitude wind fields are improved in the analysis and in shorter-range forecasts. It is important to understand how this is achieved, whether through the 4D-Var tracing effect or in combination with other mechanisms. Perhaps improved humidity fields lead to improved modelling of the radiative or latent heating of the atmosphere, or perhaps dynamical information could be inferred through the background error correlations between humidity and temperature (ECMWF does not model correlations between humidity and wind). We hope to perform some more detailed tests on the system to better understand these and other possibilities, but there is evidence in the current experiments to suggest the tracing effect is most important.

All-sky humidity first-guess departures highlight regions where humidity and cloud need to be increased or decreased. These regions are prevalent in the winter storm tracks, where there are strong humidity gradients in the mid and upper troposphere and relatively large errors in the position and intensity of humidity features. In the absence of other observations, all-sky assimilation can create wind increments in these regions that substantially resemble those from the full observing system. Given that the forecast model in 4D-Var maintains the meteorological consistency of wind, temperature and moisture, wind increments must be contributing to the adjustment of the relative humidity and clouds through advection, convergence and divergence. The wind increments are also bound through the model to be consistent with adjustments to the large-scale synoptic structures, which is potentially the mechanism that can benefit longer-range forecasts.

It is the winter storm tracks where all-sky assimilation has the greatest ability to fill the gaps left by clear-sky assimilation. These are also the regions where the humidity tracing effect is likely to be largest, due to the strong humidity gradients, and where the humidity observations may be most beneficial to dynamical forecasts. In the context of the full observing system, the greatest impact on short-range wind forecasts is in the expected areas: in the upper troposphere in the storm-track regions. These characteristics are all consistent with a 4D-Var assimilation system that is adapting dynamical fields to better fit humidity, cloud and precipitation features in the observations.

FURTHER READING

Allen, **D.R.**, **K.W. Hoppel** & **D.D. Kuhl**, 2014: Wind extraction potential from 4D-Var assimilation of stratospheric O_3, N_2O, and H_2O using a global shallow water model. *Atmos. Chem. Phys.*, **14**, 3347–3360, doi:10.5194/acp-14-3347-2014, 2014.

Andersson, **E.**, **J. Pailleux**, **J.-N. Thépaut**, **J.R. Eyre**, **A.P. McNally**, **G.A. Kelly** & **P. Courtier**, 1994: Use of cloud-cleared radiances in three/four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.*, **120**, 627–653, doi: 10.1002/qj.49712051707.

Bauer, **P.**, **A.J. Geer**, **P. Lopez** & **D. Salmond**, 2010: Direct 4D-Var assimilation of all-sky radiances. Part I: Implementation. *Q. J. R. Meteorol. Soc.*, **136**, 1868–1885, doi: 10.1002/qj.659.

Geer, **A.J.** & **F. Baordo**, 2014: Improved scattering radiative transfer for frozen hydrometeors at microwave frequencies. *Atmos. Meas. Tech.*, **7**, 1839–1860, doi:10.5194/amt-7-1839-2014

Geer, **A.J.** & **P. Bauer**, 2011: Observation errors in all-sky data assimilation. *Q. J. R. Meteorol. Soc.*, **137**, 2024–2037, doi: 10.1002/qj.830.

Janiskova, **M.** & **P. Lopez**, 2012: Linearized physics for data assimilation at ECMWF. *ECMWF Tech. Memo. No. 666*.

Predictability of cold drops based on ECMWF's forecasts over Europe

NIKOLETT GAÁL, ISTVÁN IHÁSZ

Cold drops are closed cyclonic eddies in the middle and upper troposphere that are separated from the main western stream. Being elliptically shaped with a diameter of hundreds of kilometres, they resemble miniature cyclones on satellite images. Within a cold drop the air mass is isolated from the cooler air of higher latitudes and it carries air substantially colder than its surroundings to the warm regions of lower latitudes. Cold drops are sometimes called upper-level lows (ULLs).

It is important to study cold drops because they can sometimes be responsible for severe weather affecting a region for a couple of days. They can occur at any time of the year and often bring high amounts of heavy rainfall. The unstable nature of cold drops provides perfect conditions for the formation of massive thunderstorms in summer (these can lead to flooding, or in rare cases, even tornadoes), and intense snowfall during winter. Synoptic studies of the cold drops have received less attention during the past decades than their importance would suggest.

We have studied the formation, development conditions, and synoptic and dynamical backgrounds of cold drops using ECMWF's ERA-Interim reanalysis. The first step was to develop an algorithm to recognize the cold drops. Then a statistical investigation determined the typical horizontal and vertical structures of the cold drops. For the 280 selected cases, predictability of the intensity and position of the cold core was investigated using ECMWF's high-resolution forecast (HRES). Also, based on ECMWF's ensemble forecast (ENS), a new type of ensemble plume containing four meteorological variables was developed to help forecasters predict the onset of a cold drop.

Life cycle of a cold drop

A classic cold drop has four development phases, starting with the ULL becoming isolated from the main airstream and ending with it dissolving or fusing with another stream (*Nieto et al.*, 2005). There are four phases: upper-level low, tear-off, cut-off and final phases.

- *Upper-level low phase.* For a cold drop to form, there have to be unstable waves inside the main airstream, where the temperature wave lags the geopotential wave

AFFILIATIONS

Nikolett Gaál: Eötvös Loránd University, Budapest, Hungary

István Ihász: Hungarian Meteorological Service, Budapest, Hungary

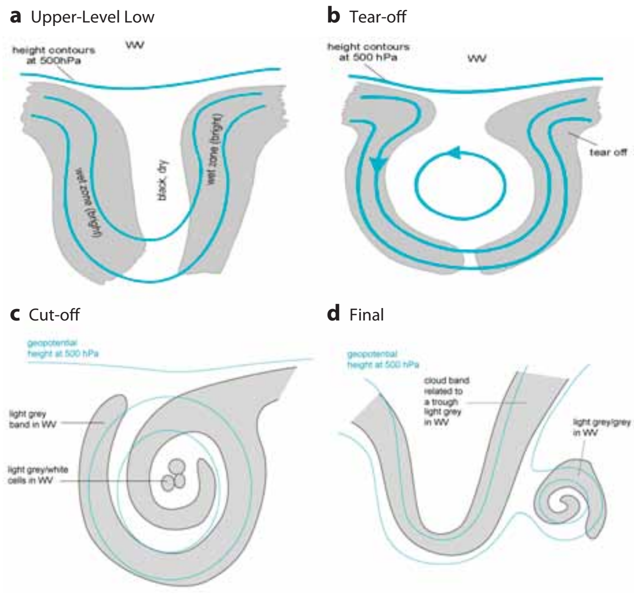


Figure 1 Phases of the cold drop: (a) Upper-level low, (b) tear-off, (c) cut-off and (d) final Phases (ZAMG, 2007).

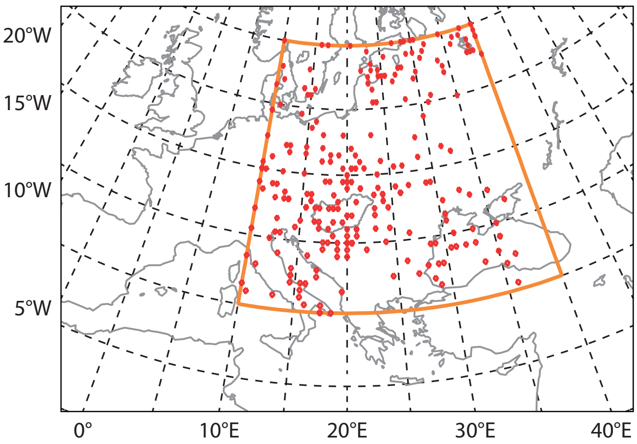


Figure 2 Centres (local minima) of 280 cold drops between 2002 and 2011.

(Figure 1a). This is the phase where the ULL is still behind the frontal cloud mass, so it shows as a clearly visible cloud trail on satellite images.

- *Tear-off phase*. The main meteorological process of this phase is the trough tearing off from the main airstream (Figure 1b). The wave amplitude gets higher ('the wave gets deeper'), followed by cold air detaching from the airstream in its southern regions. The bottom part of the ULL slowly isolates from the main stream, leading to a closed circulation in the upper troposphere.
- *Cut-off phase*. The isolation is completely finished in this phase, and the ULL is fully developed. The wind field shows the most advanced closed circulation at 500 hPa (Figure 1c).
- *Final phase*. Convection begins to develop in the ULL (except for its coldest parts). The upper air mass gets warmer by the convection and the ULL slowly decays (Figure 1d). In most cases, the ULL reattaches to the main westerly stream before dissolving completely (*ZAMG*, 2007).

The development of the cold drop usually takes 3–10 days. There are two kinds of cold drops, based on size and lifetime: 'small eddies' which last 2–4 days and 'big eddies' which can last up to 14 days. Larger eddies are usually more common than smaller ones.

The horizontal and vertical structure of cold drops

Because cold drops rarely occur at any specific geographical location, it is not easy to get a large sample of these events in order to examine their characteristics. Consequently, we decided to collect cold drops by developing an objective method for recognizing them. We used a two-step process.

- Several occurrences were collected when cold drops were identified by forecasters in the last ten years.
- The general characteristics of cold drops were determined by applying an objective method.

At first we gathered 70 cases from those submitted by synoptic meteorologists as cold drop situations between 2002 and 2011. These cases were studied daily at 6-hourly intervals, so this led to us having 280 different states. We used ERA-Interim to extract a few meteorological parameters such as temperature, geopotential, relative humidity and winds at standard pressure levels (850, 700, 500 and 400 hPa) to study the three-dimensional structure of the atmosphere. Since cold drops usually have small horizontal extensions and they rarely occur at any specific location, we chose a large area for our study. Figure 2 shows the study area and the core positions of the 280 cold drops studied between 2002 and 2011.

For the period between 1979 and 2008, we calculated the monthly average temperatures at 500 hPa from ERA-Interim. The core temperature of the cold drops was compared to the 30-year average along with the minimum and maximum monthly-mean temperatures for each month (Figure 3). It can be clearly seen that core temperature is always located below the monthly mean value. We can also see that cold drops do not have an absolute temperature threshold.

To have enough samples for a statistical analysis, we developed an algorithm for recognizing cold drops. The version used in this study was based on 500 hPa temperature, potential temperature at 2 PVU, and isentropic potential vorticity at 315 K.

Having found the core of the cold drops we explored their typical horizontal and vertical structures using a three-step process.

- Local minima of the temperature fields are determined at 400, 500, 700 and 850 hPa.
- Horizontal gradients (in °C/100 km) at distances of 100, 250, 500 and 700 km from the local minimum are calculated at each pressure level.

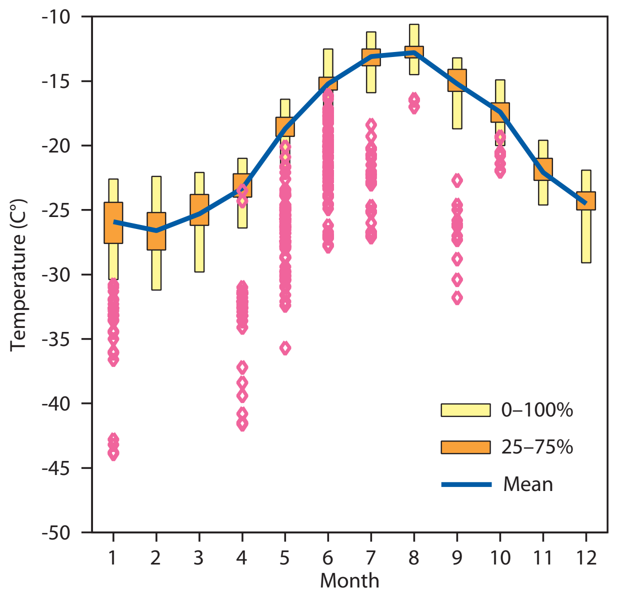


Figure 3 Monthly mean of the 500 hPa temperature (brown line) along with the corresponding coldest monthly mean (blue lines), and the warmest monthly mean (red lines) based on data from ERA-Interim for 1979 to 2008. Magenta dots show the minima of 280 cold drops.

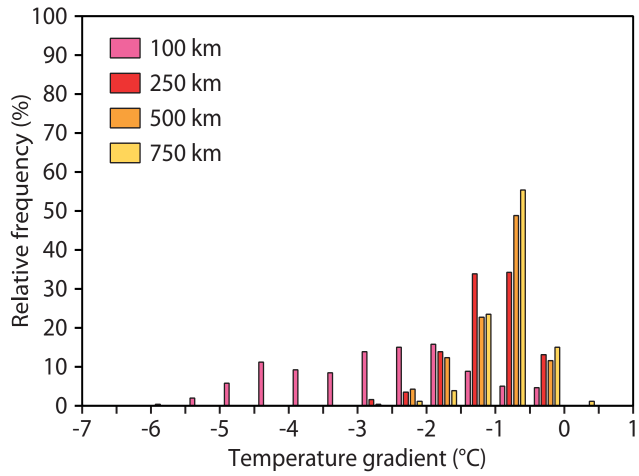


Figure 4 Histograms of the average horizontal temperature gradient (°C/100 km) around the centre of cold drops at 100, 250, 500 and 750 km at 400 hPa based on data from ERA-Interim for 1979 to 2008.

- Histograms are produced for each selected distance and pressure for the 280 cold drops.

An example of these histograms is shown in Figure 4. As expected the results indicate that the thermal gradient gets lower as we move further away from the core. This unique thermal structure of cold drops gives us a good way of distinguishing them from the cyclones in temperate zones, which are characterized by their large extent and thermal asymmetry (except in their cut-off phase). Another unique feature of cold drops is that their inner core is only apparent in the upper troposphere, while in the lower troposphere it is barely visible (as opposed to what occurs in temperate cyclones). This information can be used by the recognition algorithm to separate cold drops from cyclones.

Relationship among meteorological fields in cold drops

In this part of the work we extended the number of meteorological variables used for studying cold drops (*Gaál & Ihász*, 2014). The additional variables are 300 hPa wind speed, potential temperature at 2 PVU, isentropic potential vorticity at several levels including 315 K and 320 K, horizontal temperature advection at 500 hPa, and wind shear between 850 and 300 hPa. Here only two of the new meteorological variables are discussed: 300 hPa wind speed and potential temperature at 2 PVU.

Studying the connection between cold drops and jet streams, we found there is a very strong relationship between position of the cold core and structure of the jet stream. A typical lifecycle can be seen in Figure 5. As expected from thermal wind considerations, the jet stream is always located around the cold drop and we never see the wind maximum near the core. Before the tear-off phase the most intensive part of the jet stream quickly moves to the southern part of the wave. Then this maximum wind moves towards the northeast; this change causes a tearing of the drop. In most cases, due to further rotation of the cold drop, two intensive parts of the wind shear are visible: one to the left of the core and the other one to the right. For large cold vortices, three or four separated intensive parts of the jet stream can often be found. In the final phase of the lifecycle, when the cold drop joins to the main flow, an intensive part of the jet is always found towards the east. However, when the cold drop does not join to the main stream, it becomes stationary, typically with increased warming at the core and a decrease in the intensity of the jet stream.

As well as studying cold drops on standard pressure levels, it could be beneficial to apply 'pv-thinking'. Studying the relationship between the local minimum of the potential temperature at 2 PVU and position of the cold drops at 500 hPa, we found these two minima are typically very close to each other (Figure 6). If there is more than one cold core there are also more minima in the potential temperature at 2 PVU. We found that in the tearing phase, the pattern of the potential temperature at 2 PVU field is slightly behind that of the 500 hPa temperature with a delay of 12–18 hours.

Comprehensive study of several meteorological fields based on ERA-Interim, such as that just described, has two benefits. Firstly, good guidance could be provided for the successful prediction of cold drops using high-resolution and ensemble forecasts as we would have a better understanding of the typical relationships between various fields when cold drops occur. Secondly, it would be possible to recognise all cold drop cases in the European region during several decades for the whole reanalysis period.

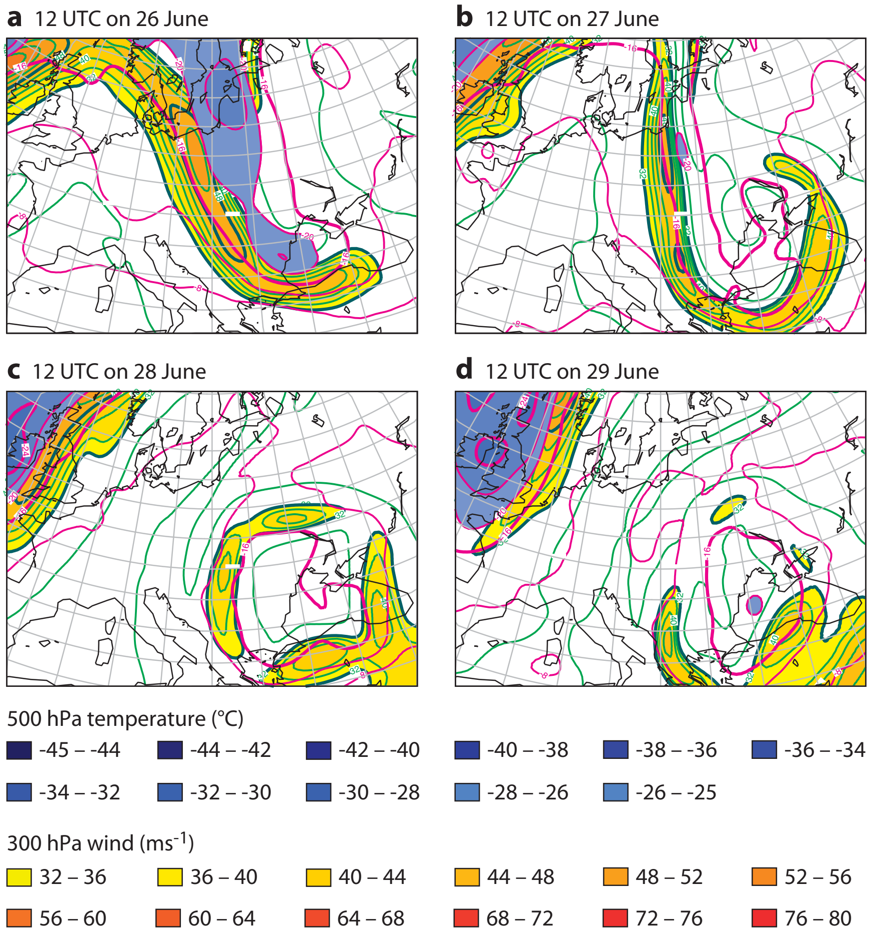


Figure 5 300 hPa wind and 500 hPa temperature at 12 UTC on (a) 26 June, (b) 27 June (c) 28 June and (d) 29 June 2011 using ECMWF's high-resolution forecast.

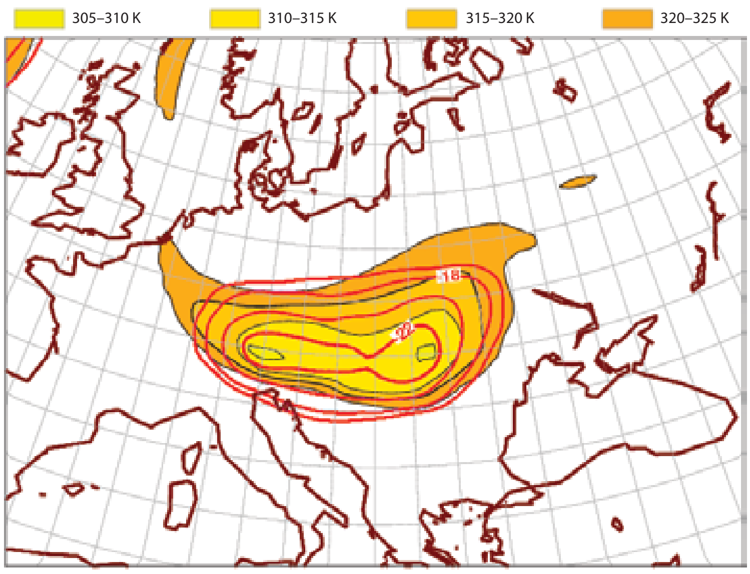


Figure 6 Potential temperature at 2 PVU (shading and black isopleths) and 500 hPa temperature (red isopleths) at 12 UTC on 2 July 2011 using ECMWF's high-resolution forecast.

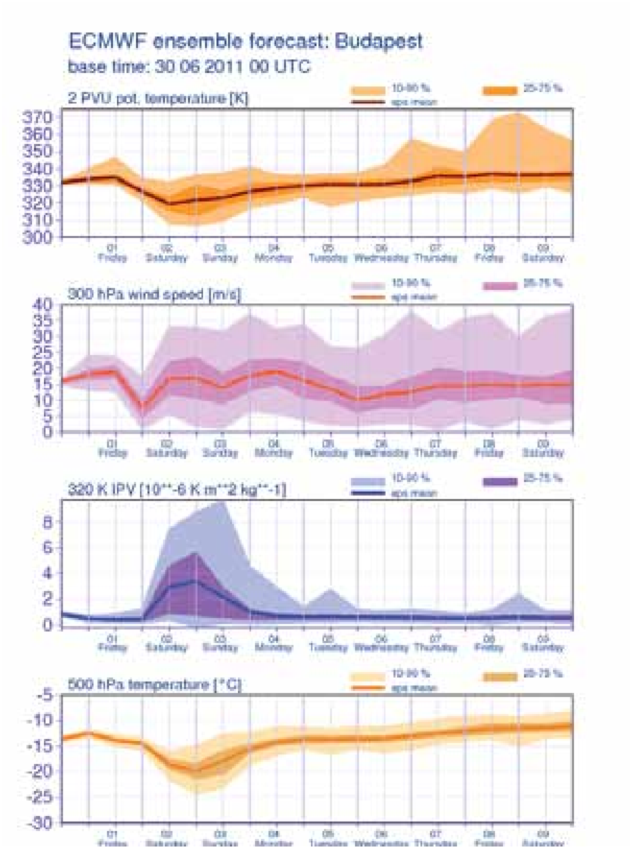


Figure 7 Ensemble plume, containing potential temperature at 2 PVU, 300 hPa wind speed, isentropic potential vorticity at 320 K and 500 hPa temperature for Budapest, based on ECMWF's ensemble forecast starting at 00 UTC on 30 June 2011.

New ensemble plumes for supporting recognition of the cold drops

Predicting cold drops using numerical weather prediction models is not an easy task. So it is worthwhile using ECMWF's ensemble forecast alongside the high-resolution forecast. At the Hungarian Meteorological Service a wide range of the ensemble-based graphical products are used in operations, among them plumes, meteograms, spaghetti and probability maps. As a result of our investigation of the cold drops we developed new tools using a variety of meteorological fields to support the better recognition and forecast of the cold drops. We developed two new types of ensemble plume diagram. One contains three variables: 500 hPa temperature, isentropic potential vorticity at 320 K and potential temperature at 2 PVU. The other one contains four variables: in addition to the previous three variables, the 300 hPa wind speed is also included.

The usefulness of these new tools can be demonstrated using a case study. On 2 July 2011 a cold drop travelled across Hungary – Figure 7 shows the ensemble plume based on four variables. We can see the strong U-shape on the second forecast day of the 500 hPa temperature (bottom panel), showing a high probability of about an 8 °C drop in 24 hours and a similar rise after the passing of the cold drop. A corresponding change can be seen in the potential temperature at 2 PVU (top panel). The change in isentropic potential vorticity of 320 K (lower middle panel) is the opposite of that found for the other two variables, so these three can support the recognition of cold drops. As already mentioned, the behavior of the 300 hPa wind speed (upper middle panel) is not usually synchronized in space with that of the other three variables but it can provide some useful information.

Summary and outlook

Because cold drops rarely occur at any specific geographical location, it is not an easy task to acquire a large sample of cases to study and summarize the typical characteristics of cold drops. Consequently, we collected information about cold drops from a large area and developed an objective method for their recognition.

As well as determining the general characteristics of the horizontal and vertical structure of cold drops, several new methods were developed for studying them with a view to providing guidance about the forecasting of cold drops. The two most important aspects are usage of the ensemble and high-resolution forecasts on isentropic levels. For operational practice, use of the ensemble forecast alongside the high-resolution forecast can provide very valuable additional information because the intensity and position of the cold drops are quite often uncertain. In addition to applying standard pressure level fields, potential temperature at 2 PVU and isentropic potential vorticity fields are also useful as they provide information about features of the cold drops that are not apparent on pressure levels. An objective process for detecting cold drops is required that is able to reliably distinguish them from cyclones. Our results so far make this a very real and achievable goal.

One of our plans is to identify the areas that have a potential for the formation of cold drops based on ensemble forecasts. In the future, we would like to run further tests with our detection algorithm for cold drops to study the last 30 years of cold drops, and we would also like to carry out further experimentation about the forecasting of cold drops.

FURTHER READING

Gaál, N. & I. Ihász, 2014: Evaluation of the cold drops based on ERA Interim and ECMWF's ensemble model over Europe. *Időjárás* (accepted for publication)

Nieto, R., L. Gimeno, L. Torpe, P. Ribera, P. Gallego, R. Garcia-Herrera, J.A. Garcia, M. Nunez, A. Redano & J. Lorente, 2005: Climatological features of cutoff low systems in the Northern Hemisphere. *J. Climate*. 3085–3012.

ZAMG, 2007: Manual of synoptical and satellite meteorology: Conceptual models and cases studies. *http://www.zamg.ac.at/docu/Manual/*

SAPP: a new scalable acquisition and pre-processing system at ECMWF

ENRICO FUCILE, CRISTIANO ZANNA,
IOANNIS MALLAS, MARIJANA CREPULJA,
MARTIN SUTTIE, DRASKO VASILJEVIC

A new scalable acquisition and pre-processing system (SAPP) was introduced into operations at ECMWF on 3 June 2014, replacing the previous acquisition and pre-processing suite (HAPP) after 20 years of continuous work. The longevity of the previous processing system is both a testament to its good design and a reflection of its underlying complexity. Such a system requires great care in making any significant changes to the framework of the processing chain and, consequently, this acts as a deterrent to major upgrades or re-engineering of the system. Only a change in computing architecture and a sharp increase in satellite data in the recent years have triggered the need for re-engineering to provide a shift of the processing model from a highly available application (i.e. one that ensures a certain degree of operational continuity) to a scalable system (i.e. one that can be enlarged to handle a growing amount of work).

The function of this fundamental module of the processing chain is to:

- Acquire observations from a multitude of sources (in the order of 200).
- Decode the various formats (e.g. BUFR, GRIB, HDF, netCDF, text).
- Apply initial quality control.
- Convert to a consolidated format before delivering to the data assimilation processing.

The main design requirements for SAPP are scalability, improved monitoring and administration, continuous observation processing, and use of Commercial Off-the-Shelf (COTS) Linux boxes.

Scalability is required to cope with the increasing volume of satellite data, while improved monitoring and administration is needed for the growing variety of data coming from a multitude of remote sensing and in-situ platforms.

The requirement of building a system with COTS Linux boxes is driven by several factors:

- The need to choose a relatively cheap solution.
- The aim of minimising the diversity of systems to be maintained in-house, as this solution is suitable for many other projects.
- The portability of a system on such a platform compared with that of a system depending on special high-availability mechanisms which are tightly linked to a specific platform.

There are also drawbacks in this choice which are mostly connected with the reliability and stability of the system. To overcome these problems it has been decided that there should be enough redundancy in the system to cope with any system failures. Indeed, the operational configuration consists of two Linux clusters, one of which is operational and the other is used as a backup and for testing and other activities (e.g. the re-processing of past observations which has become more important with the growing reanalysis needs).

A project to build a Continuous Observation Processing Environment (COPE) was initiated at ECMWF with the aim of reviewing the entire observation processing chain to (a) implement continuous processing and complex monitoring and (b) build a unified framework for filtering observations. The continuous processing requirement coming from COPE was to some extent already implemented in the existing system. However, SAPP has improved this aspect by building the system around the concept of continuity rather than running quasi-

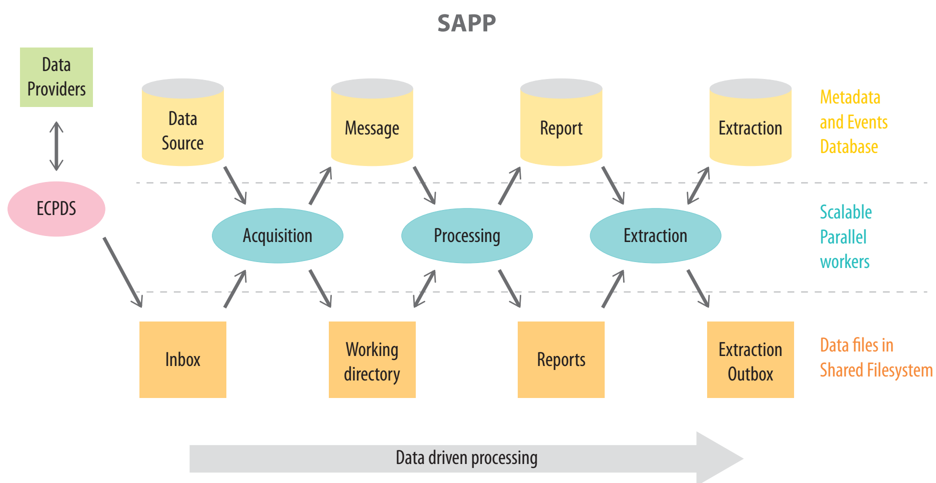


Figure 1 SAPP system design diagram.

continuous tasks from an external scheduling system. The immediate advantage is a 'natural' load balancing because most of the observations come more or less regularly throughout the day. This means that processing them as soon as they enter the system is a very efficient way of spreading the load in time as compared to processing at regular intervals which can cause resources saturation.

The aim of this article is to illustrate the design concepts of SAPP and show some of the improvements in comparison with the previous system.

Design

The overall system design is illustrated in Figure 1. The data is downloaded from more than 200 sources by ECPDS, which is the system providing the file moving service from and to external and internal sources or destinations. ECPDS has been developed at ECMWF and was extended so that it could handle data sources to support SAPP in the acquisition task. ECPDS is not part of the SAPP system and its development is not in the scope of this article.

The incoming data reaching the inbox directory is immediately processed by SAPP. The system design is based on a data-driven processing model in which the data is continuously pulled by parallel tasks in successive stages: acquisition, processing and extraction. The system is designed to work on a Linux cluster with multiple workers running on several nodes, spawned by a master process acting as load balancer and scheduler. A special case of workers are the processors performing the decoding, quality control and conversion of the incoming data in a standard BUFR format (*WMO*, 2013) that is then delivered to the data assimilation system through the extraction.

The process dispatcher, which is spawning the processors, can be configured to create a maximum number of processors per node. Also it assigns to each processor a maximum number of incoming messages to be processed. The dispatcher prepares the work for each processor, tagging the data in the metadata database as 'being processed' and writing on the same database the start time of the processing. At the end of the processing the result status (i.e. as 'success' or 'failure') and the time when processing finished are written on the database. This allows monitoring of the load of the system and recovery of anomalous situations by the dispatcher.

A timeout attribute is set for each processing type and the scheduler re-spawns a processor with the same data if the original processing task runs for longer than the timeout period. This allows the system to put back in the processing queue the data involved in a task which crashed or was taking longer than expected. All the timeout events are logged in the system so that anomalous behaviour of processors can be detected.

We have used a relational database to store metadata and events and a shared filesystem to make the data available to the processing nodes. The decision to keep the data on a filesystem was taken to minimise the work on the relational database engine. In addition, this ensures that the database was more manageable in terms of replication and backup, which are the main resilience mechanisms used to recover situations of database corruption. This design choice also leads to better performance and more flexibility of the system.

As an example of the benefit of the chosen design, consider the acquisition process. It is very efficient because it analyses the incoming files to insert into the database only the location and size of each message within the file. In this step, the original file remains untouched and the process dispatcher prepares the work for each processor by making a list of messages in terms of file names, position in the file and size of the message. The processor is able to access the input file which is on the shared filesystem and then a random access reading of the message that has to be processed is performed. This approach is very efficient as the preparation of the work for a single process is reduced to making a list of file name, file offsets and message sizes.

In the approach taken there is no real moving of data or preparation of an input file for the processor. This is a key aspect in the design of the system which leads to a significant improvement in performance compared with the previous HAPP system. It also has the important advantage of providing flexibility in the development phase of each processor as a processor can be developed outside of SAPP because it is working on files. The new processor can then be plugged into the system at a later stage after consolidation and when it is considered to be ready for a continuous test phase preceding the operational implementation.

SAPP versus HAPP

A summary of the comparison between the new scalable and the previous system is reported in Table 1. The improvements are mainly due to the design choices already described and satisfying the new requirements.

HAPP was not able to scale in the context of a problem that is intrinsically scalable (i.e. the processing of two different data that are not interdependent and can be performed simultaneously). The old system was able to run only one processor per type of data on one single node whereas SAPP is designed to run simultaneously several observation processors on several nodes, thereby providing the necessary scaling capability which is in the main requirements.

The old system was built around the concept of high availability of the application which is relocated to another node if there is any problem on the node on which it was running. This concept is not conveniently applied to a scalable model such as the one we were trying to design for SAPP. Therefore we decided to achieve a resilient system through an intrinsic and seamless fault tolerance. In SAPP, if a node has a problem and the processing of that node is not coming back to the process dispatcher as expected, new processes are spawned on the available nodes and the system continues to work without the need for any intervention or the use of any mechanism for relocation or recovery. The fault tolerance in this case is built into

HAPP	SAPP	SAPP improvements
One process per data stream on a single node	Several processes per data stream per node on several nodes	Scale out by adding more nodes to the cluster; more efficient use of resources
Highly available by relocation to another node in case of failure	Fault tolerant. Unprocessed items are made available for processing again	Seamless recovery of node or process failure
SMS suite	SAPP dispatcher	Load balancing
Configuration on files	Configuration on database managed by WEB interface	Configuration unified under unique user friendly interface
Data and metadata on database	Data on file system; metadata on database	Smaller database better performance; database backup/replication
Processors get/put data from/to database	Processors get/put data from/to files	Processors can be developed and tested as standalone and plugged into the system when ready; no need for database calls from FORTRAN
Extraction is involving retrieval of BLOB from database	Extraction relies on metadata on database and data in files	Extraction is more flexible due to the richer metadata content and faster because data are on file system, not on database
No events log, only some error files	Processing status is logged in database	Tracking data each processing step. Monitoring and live data statistics possible
No BUFR, GTS, GRIB tools used for routing	ecCodes used for routing and report metadata	Simpler use of decoding tools, more flexibility, easier maintenance
Developed with a mixture of C and FORTRAN	Developed with Python	Quicker prototyping, reduced amount of code, easier maintenance

Table 1 Comparison between HAPP and SAPP design features.

the system. Actually the system needs a highly available database server to provide enough resilience for the operational service. This is achieved with well-established replication and backup techniques which are common in the database administration domain.

Another design feature of the old system that prevented effective load balancing of the processing was the use of an SMS (Supervisor Monitor Scheduler) suite for all the activities which was designed for single process tasks. This has been addressed by designing the new system on an internal dispatcher which is used for load balancing.

One of the main requirements was the improvement of the administration of the system because the old system was based on a large set of files spread in several directories and several SMS definition files making the configuration of the system unnecessarily difficult. This requirement was implemented in SAPP by making a unique web administration interface based on an open source web framework: django.

The basic design of HAPP relied on the database not only for the metadata, but also as storage for the data, with a big overhead on the database engine both in insertion and extraction. Moreover, the processors were designed to query the database to get the data to be processed and to insert the output directly into the database. This results in a tight link between the processors and the system, and the impossibility of developing them out-of-the-box as is now the case in SAPP.

The processors are mostly written in FORTRAN and the insert and select on the database had to be performed in HAPP through a software layer, strongly dependent on the specific database engine which makes the portability to another relational database management system (RDBMS) more difficult. In SAPP, the original FORTRAN processors are used with minor modifications to read and write on files without accessing the database, while the SAPP system components are written using the Python programming language and all the accesses to the database are performed in Python, resulting in more flexibility in changing the RDBMS.

In HAPP there was only a very limited log of the activities which mainly involved error files, while SAPP logs the status and time of processing for each step as indicated in Figure 1. This gives an advantage in tracking the data through the processing steps and in building monitoring views and alarms which are extensively used in the management of daily operations.

A significant difference between SAPP and HAPP related to improving the maintainability of the system is in the use of a single component for the decoding of GTS headers (*WMO*, 2009), BUFR and GRIB (*WMO*, 2013) for the purpose of routing the messages to the appropriate processors and gathering the metadata from the produced reports. Indeed, we have developed a new decoding library (ecCodes) which is able to decode GTS headers, BUFR and GRIB with the same function calls and tools and these can be considered an important building block of SAPP. This new decoding library is described in some detail in the next section.

Routing rules and ecCodes

A data acquisition and pre-processing system like SAPP has to be able to decode the incoming data in each of the processing stages, mostly because the metadata coded inside the messages is used to route the data to different processing chains; it is also used for monitoring purposes. To provide decoding capabilities to SAPP, a new decoding library has been developed based on the successful design used for grib_api *(http://software.ecmwf.int/wiki/display/GRIB/Home)*, which is the GRIB encoding/decoding software recently developed at ECMWF. This new library (ecCodes) is able to decode GTS headers, BUFR, GRIB and can be extended to decode other types of messages.

The basic principle of ecCodes (*Fucile*, 2014) is to provide access to the information items of a message through key/values pairs. Some simple functions are provided to get and set the value associated with a key, which in some cases can also be a list of values. The approach is very simple to learn and the library provides a convenient Python interface which made the implementation of the routing rules for incoming messages very simple. Since the SAPP core system is written in Python, it was particularly useful to have a decoding library available in the same language.

ecCodes is currently able to decode all the data processed in SAPP and is used to get metadata such as latitude and longitude of the station, time of observation and many other parameters from the reports produced by the processors. The metadata enables the system to provide space and time statistics in real time while the data is flowing through the processing stages.

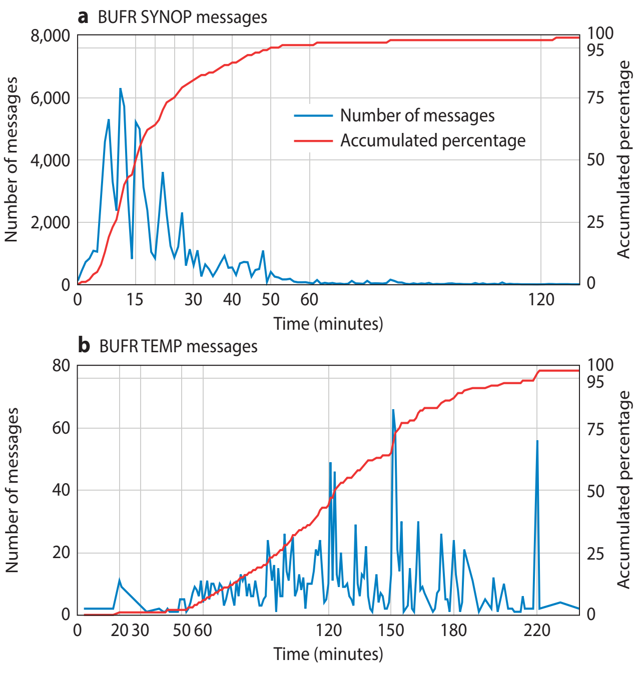


Figure 2 Number of (a) BUFR SYNOP and (b) BUFR TEMP messages as a function of the timeliness, computed as the difference between the observation and the acquisition time. The accumulated percentage of the total number of observation is also reported.

Re-processing

The activity of re-processing entire observation data streams is becoming more important with the increased need to prepare reanalyses associated with the Copernicus programme. A problem with the old HAPP system was the amount of time needed to re-process multiple years of data. Typically, for 20 years of data from a single satellite instrument, the re-processing could take more than three weeks. We estimate that the old system was processing one year of data in one day. This can appear to be a good speed, but it results in limiting the number of re-processing operations that can be completed per month to no more than one because of the limited resources available on the old system. Also, some human expertise was occasionally needed to monitor that the activity was progressing as expected.

SAPP is currently capable of processing the same data stream at a rate of one year per hour and, hence, of completing the processing of 20 years of data in two days. This will permit a significant enhancement in the re-processing throughput and satisfy the increasing need for reanalyses.

Monitoring

One of the advantages of basing the design of SAPP on a database containing metadata and events is the possibility of having several instantaneous and statistical views of the activity of the system and the data flow. A set of monitoring database queries has been developed to provide real-time alarms about events affecting the regular data processing. As an example, an alarm is sent to the console operators when a data source is not receiving data within a given period which is automatically computed by making an average of the time of inactivity for each source. Other alarms are targeted on the quality and format of the incoming data. These alarms are connected with the percentage of unsuccessful processing or rejected data per data stream. When the percentage of success goes under a fixed threshold, an alarm is sent to the console operators and the intervention of an analyst is required.

A simple use of the metadata stored in the database is shown in Figure 2 where the timeliness for BUFR SYNOP and BUFR TEMP (*WMO*, 2013) data is reported. The timeliness is computed as the difference between the observation time and the acquisition time (i.e. the time when the observation has reached SAPP). The number of messages is reported as a function of the timeliness. Also an accumulated percentage of messages received is reported which makes it easier to see when the data has reached significant thresholds in terms of percentage. In the comparison between the two plots it is clear that 95% of BUFR SYNOP data reached the processing chain 50 minutes after the observation time, while 95% of BUFR TEMP data was acquired 220 minutes after the observation time.

Figure 3 Geographic coverage of BUFR SYNOP observations over Italy and Greece as extracted in real time from SAPP and plotted on a google map.

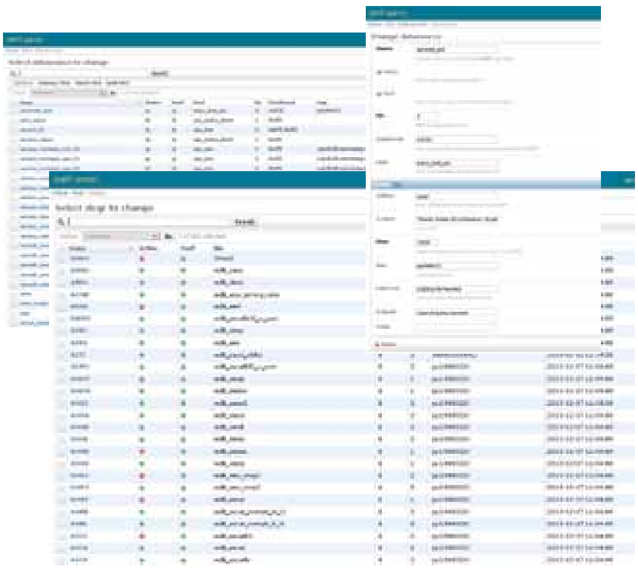


Figure 4 View of SAPP administration web interface.

SAPP can provide statistics in space where map coverage can be produced in real time from the system by submitting queries based on the latitude and longitude metadata of the observation processed. This is used to plot maps of observations coverage and has been particularly useful in the development of a wiki on the data migration from Traditional Alphanumeric Codes (TAC) to BUFR (TAC to BUFR wiki) – see *http://software.ecmwf.int/wiki/display/TCBUF*. Most of the content of this wiki is produced automatically by SAPP and is, in some cases, providing a complex and precise view of the migration status. In Figure 3 we have a coverage plot on google maps of the BUFR SYNOP data produced in WMO Region VI from Italy and Greece. Other plots or tables reporting the stations producing TAC, but not yet producing BUFR, can be consulted on the wiki and are produced with relatively simple queries on the SAPP metadata database.

Administration

A web administration interface has been developed to conveniently configure the system which is processing hundreds of data sources and using hundreds of different data processors. The web administration interface (see Figure 4) is a quick and simple way of configuring the system by defining new processors with their specific properties, adding new data sources or defining new routes between existing data sources and processors. This speeds up work that is made complex by the growing number of different observations and the complexity of the data received on the GTS. Indeed, the system has the capability to define a specific routing based on the GTS header of a message. This is particularly useful when new data is introduced on the GTS or some data generated by a particular provider is affected by some problem and needs to be removed from a routing table.

Summary and outlook

The new SAPP system has improved several aspects of pre-processing observation data by providing the scalability required by future needs in data volumes and diveristy of data. It also provides improved monitoring, seamless failover recovery and a better framework to develop further observation pre-processing. This is achieved in the context of a wider project to develop a Continuous Observation Processing Environment (COPE) aimed at building a more scalable, efficient and effective system for the successive processing stages involving filtering observations and complex quality control procedures.

FURTHER READING

WMO, 2013: Volume on Codes. *Publication N. 306*. http://www.wmo.int/pages/prog/www/WMOCodes/WMO306_vI2/Publications/2011editionUP2013/WMO306_vI2_2011UP2013.pdf.

WMO, 2009: Manual on the Global Telecommunication System. *Publication N. 386*, http://library.wmo.int/pmb_ged/wmo_386-v1-2009_en.pdf.

Fucile, **E.**, 2014: Semantic approach to WMO Codes. *SENSORNETS 2014 Proceedings*, 409–414.

Metview's new user interface

SÁNDOR KERTÉSZ, IAIN RUSSELL

Metview is ECMWF's meteorological workstation software for accessing, manipulating and visualising meteorological data, incorporating both an interactive and a batch mode. By bringing together various software technologies developed at ECMWF, Metview provides access to the MARS archive, allows the examination and manipulation of data formats such as GRIB, BUFR and ODB, and can overlay data to generate high quality meteorological plots.

Metview's user interface was originally developed using the Motif toolkit in the early 1990s. By the year 2000, Metview had a new user interface, but was still based on the same toolkit. Today, Motif is far behind the cutting edge of user interface development and is becoming harder to install on some new systems. Also, it is laborious adding new features to code developed using Motif; this means that Metview's user interface has barely evolved in the last decade.

When the Qt application framework became available under the Open Source LGPL licence in 2009, it was assessed to be the best option for future Metview developments. Because of Metview's modular nature, some self-contained parts of the user interface, such as the Grib Examiner, the Macro Editor and the interactive Display Window were already re-written in Qt by the end of 2010, bringing new features and ease of use. However, the main user interface was not tackled until more recently due to the size of the task. Now, with the latest available versions of Metview, the new user interface is ready for use.

The new user interface

Design of the interface

Metview's new interface was designed to be familiar to existing Metview users whilst being intuitive to first-time users who are used to modern user interfaces. All the functionality of the previous interface is available in one way or another, but many additional features make it easier and quicker to work with Metview. The Metview *desktop* is still the main interface, essentially being a browser for the icons contained in a particular folder (Figure 1). As well as allowing different desktops to be open in different windows, the new interface allows multiple tabs in each window, thus reducing screen clutter.

Layout

As shown in Figure 2, Metview now has three icon layouts: *Classic* and *Simple* emulate the views previously available, whilst the *Detailed* layout displays the icons in a table format with sortable columns such as file size and modification time. A slider in the status bar at the bottom of the desktop allows more control over the size of the icons.

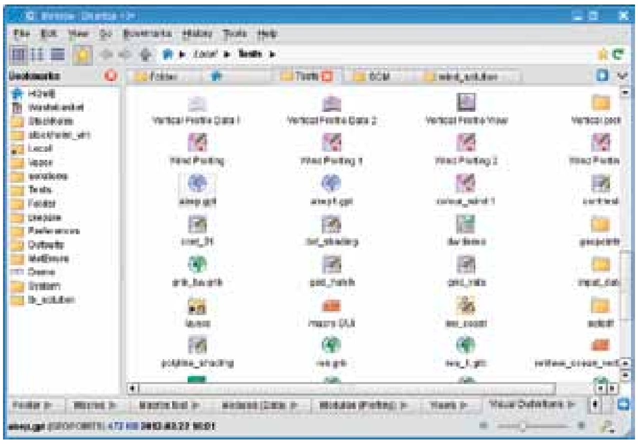


Figure 1 The main window of Metview's new user interface. The central part contains the folder views organised into a set of tabs. To the left there is the bookmarks sidebar, while the folder navigation actions together with clickable folder location bar are located in the toolbar at the top. The bottom part of the desktop, just like in the previous user interface, is occupied by the icon drawers. The status bar, at the very bottom of the desktop, displays useful information about the icon that the mouse cursor currently hovers over.

Figure 2 Users can switch between three icon view modes: there are Classic (left), Simple (middle) and Detailed (right) views available.

Navigation

The top portion of the interface contains tools for navigating through the file system (see Figure 1). The Back, Forward, Up and Home icons, combined with the clickable folder location bar, allow for fast navigation between folders. Folders can be bookmarked, as can a whole set of tabs.

Icons

The Create New Icon dialogue (see Figure 3) has been made faster to find icons and now includes a quick filter mode, a grouped icon mode and a list of recently-used icons. The icon drawers at the bottom of the desktop are no longer updated by the system and are intended solely for user icons. When the mouse cursor hovers over an icon, useful

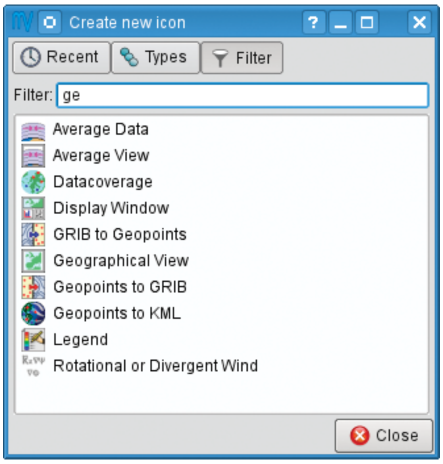


Figure 3 The Create New Icon dialogue offers various ways to quickly look up the particular icon to be created.

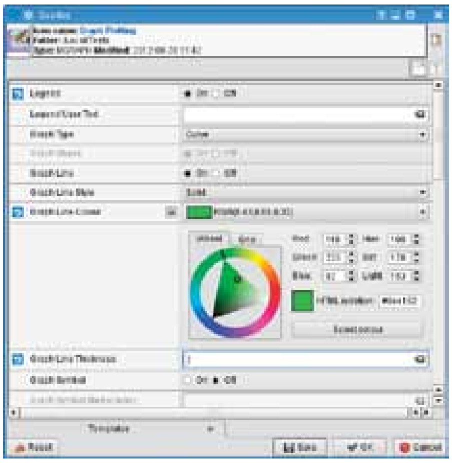


Figure 4 Although icon editors retain their original functionality, they have been completely revamped and equipped with a new set of parameter editors such as the new colour editor.

information is displayed in the status bar. Metview now offers the ability to compress and archive icons, and to restore them later. This facility is also used when e-mailing icons to another user.

Icon editors retain their previous functionality, but with some enhanced 'helper' tools such as the colour-selection panel shown in Figure 4.

Search

Metview users often work with folders containing hundreds (even thousands) of icons and there has always been a need for an icon search facility. The new user interface now offers the *Icon filter* toolbar enabling users to search for specific icons in cluttered desktops (Figure 5). As the filter term is being typed in, all the non-matching icons are greyed out leaving only the searched-for icons highlighted.

The *Advanced search* dialogue has also been added to find icons in subfolders according to name, type, contents and other search criteria (Figure 6). This powerful new feature can easily expose the contents of the Metview folder hierarchy to users.

Moving to the new user interface

The new interface can read settings, such as icon positions, from the previous interface but not vice-versa. This means that on starting the new interface for the first time, everything will be as expected. However, icons which have been moved within the new interface will not be seen in their new position in the previous interface; that said, we expect that users will not want to revert to the old interface!

It is planned to make the new interface the default during 2014. Until then, it can be invoked by starting Metview with the command-line option `–desktop`.

More information about Metview's new user interface is available at: *http://software.ecmwf.int/metview*

Figure 5 The Icon filter bar attached to the bottom of the folder view allows icons to be found quickly by greying out the all non-matching icons in the desktop.

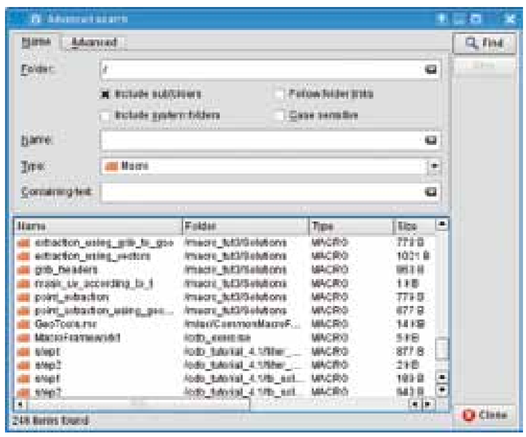


Figure 6 The Advanced search dialogue is a powerful new feature for finding icons in subfolders according to various search criteria.

ECMWF Calendar 2015

Jan 26–30	Training Course: Use and Interpretation of ECMWF Products
Feb 2–6	Training Course: Use and Interpretation of ECMWF Products
Feb 9–13	Computer User Training Course: ecFlow or HPC
Feb 23–27	Computer User Training Course: GRIB
Mar 2–6	Computer User Training Course: INTRO/MARS
Mar 9–13	Computer User Training Course: Magics/Metview
Mar 16-20	NWP Training Course: Data Assimilation
Mar 23-27	EUMETSAT/ECMWF NWP SAF Training Course: Assimilation of Satellite Data
Mar 31	EUMETSAT Licensing Agents Workshop and Data Policy Group
Apr 1	Advisory Committee for Data Policy
Apr 2	ECOMET Working Group
Apr 13–17	Computer User Training Course: ecFlow or HPC
Apr 20–29	NWP Training Course: Predictability
Apr 21	Policy Advisory Committee
Apr 22–23	Finance Committee
May 11–21	NWP Training Course: Parametrization
May 18–22	Security and Computing Representatives' Meetings
Jun 1–5	NWP Training Course: Advanced Numerical Methods
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 3	ERA-CLIM2 Workshop on Observations for Earth System Reanalysis
Aug 31–Sep 4	Annual Seminar
Sep 28–30	Visualisation Week: Workshop on Meteorological Operational Systems
Sep 29–Oct 1	Visualisation Week: European Working Group on Operational Meteorological Workstations (EGOWS)
Oct 1 (pm)	Visualisation Week: RMetS Seminar on Visualisation in Meteorology
Oct 2 (am)	Visualisation Week: OGC MetOcean Interoperability Session
Oct 5–7	Training Course: Use and Interpretation of ECMWF Products
Oct 12–14	Scientific Advisory Committee
Oct 15–16	Technical Advisory Committee
Oct 21	Policy Advisory Committee
Oct 22–23	Finance Committee
Nov 2-6*	Workshop on Extended Range Predictability
Dec 2 (pm)*	Royal Meteorology Society Meeting
Dec 8-9	Council

* To be confirmed

ECMWF publications

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

Technical Memoranda

727 **Migliorini, S.**: Optimal flow-dependent selection of channels from advanced sounders in the presence of cloud. *July 2014*

726 **Cardinali, C., N. Žagar, G. Radnoti & R. Buizza**: Representing model error in ensemble data assimilation. *June 2014*

725 **Bechtold, P., I. Sandu, D. Klocke, N. Semane, M. Ahlgrimm, A. Beljaars, R. Forbes & M. Rodwell**: The role of shallow convection in ECMWF's Integrated Forecasting System. *July 2014*

724 **Kaiser, J.W., N. Andela, J. Atherton, M. de Jong, A. Heil, R. Paugam, S. Remy, M.G. Schultz, G.R. van der Werf, T.T. van Leeuwen & M.J. Wooster**: Recommended fire emission service enhancements. *May 2014*

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

721 **Mozdzynski, G. & J.-J. Morcrette**: Reorganization of the radiation transfer calculations in the ECMWF IFS. *April 2014*

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

Newsletter | Number 140 – Summer 2014
European Centre for Medium-Range Weather Forecasts
www.ecmwf.int