

50 years of ECMWF

ADVANCING WEATHER SCIENCE
THROUGH COLLABORATION

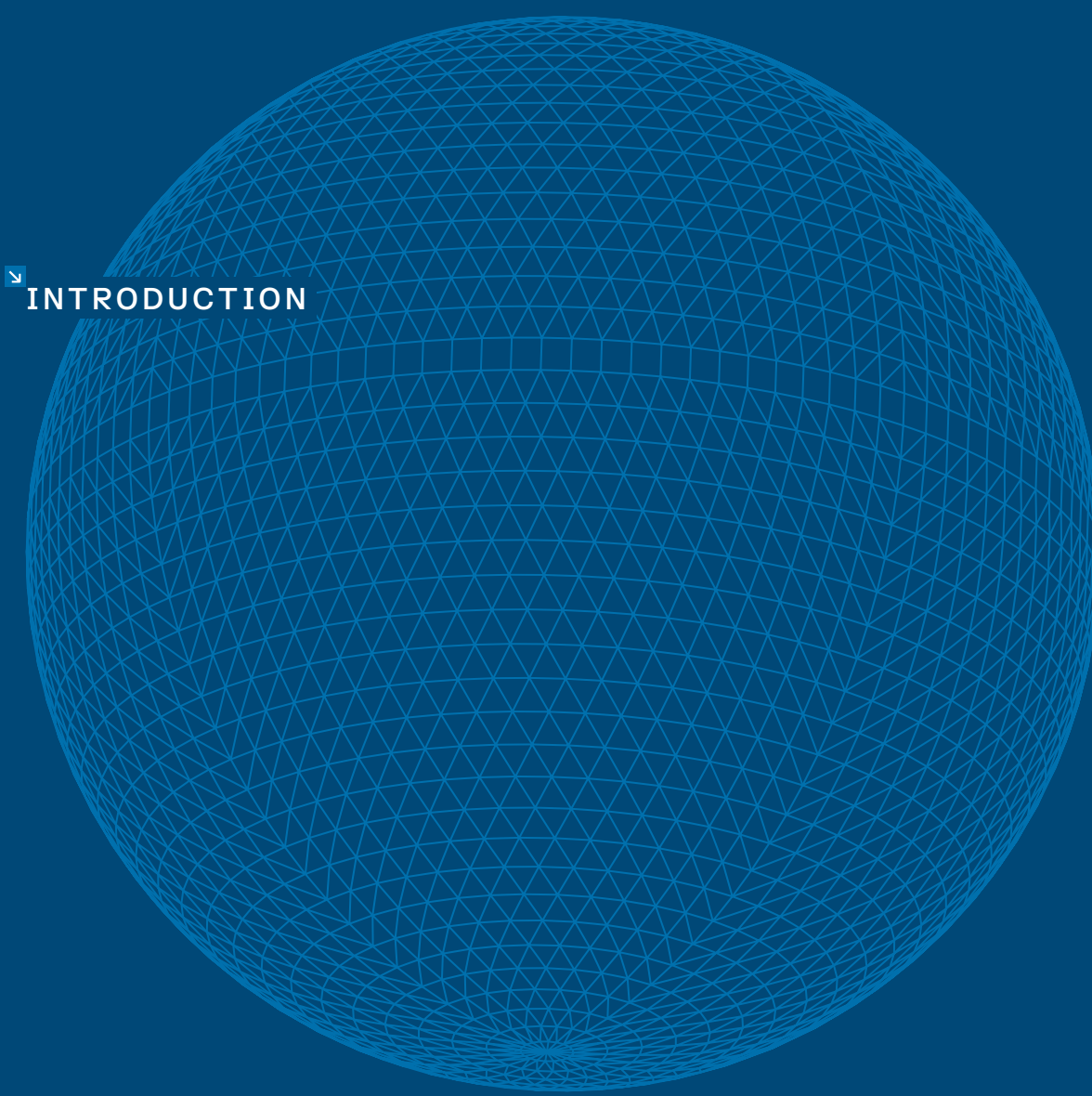
 ECMWF | 50

Concept & coordination Hilda Carr, ECMWF 50th Anniversary Programme Coordinator
Editing & content ECMWF – ecmwf.int
Shinfield Park Reading RG2 9AX United Kingdom
© ECMWF 2025
Design Planning Unit – planningunit.co.uk
Printed Livonia print Jūrkālnes iela 15/25 Rīga LV-1046 Latvia
DOI 10.21957/mpcf-nt20

CONTENTS →

CLICK ON CONTENT HEADERS TO NAVIGATE

INTRODUCTION	4
TIMELINE	16
PAPERS	28
FIFTY YEARS OF EARTH SYSTEM MODELLING AT ECMWF	30
FIFTY YEARS OF DATA ASSIMILATION AT ECMWF	42
TEN YEARS OF COPERNICUS AT ECMWF	62
THE CRITICAL ROLE OF HIGH-PERFORMANCE COMPUTING IN MEDIUM-RANGE WEATHER FORECASTING: HALF A CENTURY OF TECHNOLOGY INNOVATION	84
FIFTY YEARS OF MEETING USERS' NEEDS	108
MACHINE LEARNING FOR NUMERICAL WEATHER PREDICTION	124
PHOTOS	136
REFERENCES	148



INTRODUCTION

The European Centre for Medium-Range Weather Forecasts (ECMWF) was created 50 years ago by a group of European nations, who, convinced that the improvement of medium-range weather forecasts would contribute to the protection and safety of the population, decided to pool their scientific and technical resources to meet this goal.

After years in the making, the Convention creating ECMWF as an independent intergovernmental organisation entered into force on 1 November 1975. Amendments in 2010 opened the possibility for new Members to join and broadened ECMWF's mission to include monitoring of the Earth system.

ECMWF's first building in Reading (UK), still the headquarters to this day, became home to ECMWF's supercomputer and a small staff of experts in atmospheric sciences and supercomputing.

Today, ECMWF has three duty stations, with Italy and Germany joining the UK as host nations. Staff numbers are over 500 across duty stations.

This book presents some of what has made 50 years of collaboration between ECMWF and its Member States, with its broader community of the World Meteorological Organization (WMO), the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), the European Space Agency (ESA), the European Meteorological Network (EUMETNET), and partner organisations around the globe. It is a testament to the talent and expertise of the staff who have worked at ECMWF and to the vision and leadership of the representatives of the Member States who have guided the developments and achievements of these past five decades. It is also, importantly, about the future of our field at this exciting time, when machine learning and artificial intelligence are turning upside down ideas and concepts which have prevailed in the past.

The following pages include words from the directors who have led ECMWF from the early 1980s to today, a timeline of key developments, science and technology progress described by those working in the field today, and a selection of photos of our people and premises.

Above all, this book is an opportunity to express our most sincere thanks to the many who have made ECMWF over the past 50 years.

October 2025.



Florence Rabier

TENURE: JANUARY 2016 TO DECEMBER 2025

LOOKING BACK, MOVING FORWARD: A DECADE OF TRANSFORMATION AT ECMWF →

As I look back on my tenure as Director-General of ECMWF from 2016 to 2025, I am filled with immense pride and deep gratitude. These years have been nothing short of extraordinary – a decade marked by bold transformation, scientific breakthroughs, and collective achievement. None of this would have been possible without the steadfast support of our Member States, the vision and dedication of our staff, and the foresight of and strong foundations laid by my predecessors.

One of the most defining achievements of this era has been our leap forward in representing the Earth’s interconnected systems. We’ve delivered a truly integrated Earth system approach – seamlessly coupling atmosphere, land, ocean, waves, and sea ice in all our forecasts. Our advances in data assimilation, powered by cutting-edge satellite missions like Metop, MTG, Aeolus, and EarthCARE, have been game changing.

In 2023, we fully embraced the ensemble approach, launching a 9 km resolution ensemble forecast system for the medium range. This milestone has significantly enhanced the accuracy and reliability of our forecasts, supporting better decisions across countless sectors.

Our partnership with the European Union has flourished, especially through the ambitious Copernicus programme and the visionary Destination Earth (DestinE) initiative. I am especially proud of Copernicus becoming a trusted global voice on climate and atmospheric composition. Its presence in mainstream media is now routine, and its data – quoted over 200 times in the latest IPCC report – has become a cornerstone of climate policy and action. The socio-economic impact, particularly in renewable energy, is something we can all celebrate.

The official launch of Destination Earth in 2024 was another landmark moment. It showcased our leadership in km-scale global modelling and the creation of Digital Twins for Extreme Weather and Climate – powered by next-generation high-performance computing.

AI IS NOW A CORNERSTONE OF OUR FORECASTING CAPABILITIES →

Perhaps the most transformative force of the past decade has been the rise of artificial intelligence. In 2020, we set out an ambitious strategy to integrate AI and machine learning into our work. By 2025, we had operationalised our first fully AI-based forecasting system – the AIFS – featuring a state-of-the-art ensemble configuration. The pace of innovation has been breathtaking, and AI is now a cornerstone of our forecasting capabilities.

We’ve also championed openness and accessibility. Through open data, new data stores, co-located computing resources, and powerful software tools, we’ve made it easier than ever for users to access and harness our data, fuelling innovation and collaboration across the global weather and climate community, especially with the World Meteorological Organization (WMO).

Our partnerships have been vital to our success. From the Anemoi community framework in machine learning to the National Collaboration Programmes for Copernicus and the community-driven efforts in DestinE, our collaboration with Member States has grown stronger than ever. We’ve also deepened our ties with EUMETSAT and ESA, working together on satellite data assimilation, the European Weather Cloud, and Destination Earth.

Organisationally, we’ve weathered major challenges – from Brexit to the COVID-19 pandemic – and emerged stronger. We’ve grown into a dynamic, multi-site organisation with new sites in Bologna and Bonn, and a new headquarters rising in Reading. We’ve also taken bold steps toward net zero and placed a strong emphasis on diversity, equity, and inclusion, though we know there’s more to do.

THE BEST IS YET TO COME →

In summary, the past decade has been a time of profound evolution. We’ve pushed the boundaries of science and technology, transformed our organisation, and stayed true to our mission: serving our Member States and tackling the global challenges of weather and climate. As I reflect on this incredible journey, I am filled with optimism. The foundation we’ve built is strong, and as I prepare to hand over to Florian Pappenberger, who will lead ECMWF from January 2026, I know from everything he’s already accomplished with us, and from his unshakeable passion and dedication, that the best is yet to come.



Alan Thorpe

TENURE: JULY 2011 TO DECEMBER 2015

THE BIRTH OF THE EU-FUNDED COPERNICUS SERVICES →

In my four and a half years as Director-General, there were many highlights. Probably the highest-profile ECMWF forecast that received global attention was that of Hurricane, and then Superstorm, Sandy in late October 2012. It was by most standards a stunningly accurate prediction of a high-impact event that very unusually turned westwards to strike New York City after tracking northwards from the Caribbean. Nearly all ensemble members exhibited this unusual behaviour. Two days before Sandy formed (9.5 days before landfall in New Jersey), there was already a significant probability of a severe windstorm affecting the north-eastern USA.

Media interest in our prediction peaked in the US because the NOAA prediction did not show until much later this westward turn to strike New York. Not for the first time, ECMWF was in the headlines in the American media, and I remember many interviews I gave trying to answer the question as to how come a European outfit could predict an American weather event so well!

In June 2013 ECMWF signed a new supercomputer contract with Cray, the signing ceremony taking place immediately after our Council of Member States approved the procurement outcome.

As the new year 2014 dawned, ECMWF was deep in complex negotiations with the European Commission regarding its bid to implement two of the new Copernicus Services – which turned into C3S and CAMS – and to provide computational resources for the Emergency Management Service regarding flooding. This was a major development for ECMWF and one that rightly required extensive discussion and scrutiny by the Member States. These services being funded and led by the European Commission and not directly, as is the case for our core activities, by ECMWF Member States represented a significant change. Not least was the new responsibility for ECMWF to itself act as the funding agency for many European organisations for them to deliver many of the specific elements of these services.

Council agreed with our proposal that a new Copernicus services department be created given the size and complexity of implementing these activities. After what was an exhausting period of negotiations, ably led by ECMWF’s Director of Administration, I signed the contract in Brussels on 11 November 2014. Since then, we have seen the hugely successful Copernicus Services at ECMWF develop into

RESEARCH AND COMMUNICATIONS →

the world-leading capability that it has become with data free to all. This capability throws into sharp focus the intimate and complex scientific relationship between climate, atmospheric composition, and weather.

Underpinning all these developments at ECMWF is the crucial in-house research programme that also draws in new science from Member States and from across the world. This research leads to the rhythm of upgrades of ECMWF’s model, reminding us of the great ongoing potential to improve forecast accuracy and reliability.

Scientific interpretation and evaluation of ECMWF’s forecasts is fundamental to future progress. To help facilitate this, a transition was needed in the provision for ECMWF staff and visitors of our forecast information from what was a paper-based presentation room to a modern video-based display facility located at the heart of the organisation. This resulted in the design and fitting-out of the Weather Room, where a weekly weather forecast discussion takes place and where informal meetings occur.

Another development was an expansion of our communications to the world of what we do. I am so impressed with the website and the many other open ways that ECMWF is a source of such a wide range of so much valuable meteorological information and debate.

Finally, a personal highlight came in 2015 when Peter Bauer and I, with Gilbert Brunet, then with Environment Canada, published a review paper in Nature entitled “The Quiet Revolution of Numerical Weather Prediction”. This has become a very highly cited source for the global weather enterprise on the amazingly successful development of weather intelligence over the past fifty years spearheaded by ECMWF and others.



Dominique Marbouty



TENURE: JUNE 2004 TO JUNE 2011



THE PINK COTTAGE →

When appointed Director of ECMWF in 2004, I was looking forward to organise the developments of the Centre’s modelling, computing and services facilities. That was in effect my major occupation, but I also discovered tasks I did not expect.

Since its creation, ECMWF had benefited from the facilities of its neighbour, the Met Office College. It was also seen as a possibility for some extension, should ECMWF continue its growth. With the move of the Met Office to Exeter, which also concerned the College, this room for extension had disappeared.

One month into my contract as Director, I learned that a 5000 m² property denominated Keeper’s Cottage, situated along the northern limit of the Centre’s ground, would be auctioned two months later. There was little buyers’ interest because the access was difficult. The owner was contacted and welcomed the possibility of an offer presented by the Centre, if received before the date of the auction.

The evaluated value was affordable to the Centre but the difficulty was that buying a property is a Council decision and in this pre-videoconference old time, it was impossible to organise a Council meeting at such short notice. I contacted most of the Council representatives, in particular the Finance Committee permanent representatives, and received unanimous support for acquiring the property. So we bought the property. The acquisition was confirmed by the next Council in December 2004 and became designated as “the Pink Cottage”.

NINE YEARS BEFORE BREXIT, A FIRST ATTEMPT AT ECMWF →

At Council 64 in December 2005, an ambitious ECMWF strategy was adopted for the period 2006–2015. From there on, discussions developed during the year 2006 about the required resources for its implementation, in particular in terms of high-performance computing (HPC), for which a doubling of the funding was required.

At Council 65 in July 2006, the UK delegation, worried by the costs implied, requested to conduct an audit of ECMWF. The Council did not agree to this proposal, but proposed that a Program Task Team (PTT) of the Council be set up in order to consider the ECMWF economic value to its Member States, their requirements, and develop an implementation strategy.

THE NEW ECMWF CONVENTION →

The PTT reported very positively to Council 66 in December 2006. On that basis Council adopted an ambitious HPC budget, incidentally outvoting the UK. On 18 April 2007 the President of the ECMWF Council received a letter from the UK delegate announcing, following the decisions of Council 66, that the UK was considering leaving the Centre.

Tense discussions took place at Council 67 in June 2007, which resulted in the UK confirming their ECMWF Member State status. Since then, not only has the UK remained a Member, but it is now developing the new ECMWF headquarters.

ECMWF was created in 1975 through a so-called closed convention that included the list of possible members, and thus excluded eastern European countries. By the end of the 20th century, it became obvious that allowing countries that were now members of the European Union was necessary. At the 51st Council in December 1999, Member States requested that the possibility to enlarge the ECMWF membership be examined.

It took a few years to realise that this meant amending the Convention, and more years to reach a consensus on the text of a new Convention (ECMWF’s role, voting majorities, languages, etc.). This was achieved at an extraordinary Council on 22 April 2005 that unanimously adopted a list of proposed amendments to the Convention and recommended its Members to accept them.

But accepting modifications of the Convention is in fact accepting a new treaty, which is usually done by the parliament. This can take time. For example, in Belgium this means several regional and language parliaments. As one of the amendments was that all national official languages would become official languages of the Centre, for Austria that meant that they needed to provide all official translations of the Convention, which imposed that their adoption had to wait for all the others. By chance Austria was the only one that required this, otherwise it is a catch 22 situation (Germany realised later that they probably should have done similarly!). In the end the Austrian acceptance was announced and the new Convention entered into force on the 6th June 2010, ten years after the start of the process.

And I had the honor to become the first Director-General of ECMWF, as this was one of the many amendments.



David Burridge



TENURE: JANUARY 1991 TO JUNE 2004

INTEGRATED FORECASTING SYSTEM (IFS) →

I joined The Centre in May 1975 (on a three-year contract) and retired from the Directorship in June 2004. The first Director, Aksel Wiin-Nielsen, interviewed me informally for hours on the day I dropped my application off and that was it – “I was in”.

Starting with a clean sheet in 1975, the goal to deliver global forecasts operationally to the Member States in 1979 was met on 1 August 1979 – an outstanding achievement, the more so as they were the best available and they have remained so ever since. Undoubtedly, this early success benefited from the intellectual capital that had been built up elsewhere. However, outstanding original research, major experimentation, and the technical developments (for data management and telecommunications) necessary for operations that were undertaken at the Centre underpinned the development of the first operational system.

After this flying start, the Centre’s forecasts improved throughout the 1980s. But by the late eighties, improvements were becoming small, and some members of the Council and its committees questioned the research and operational programmes. Research being carried out on variational assimilation of observations suggested that it would perform better than the Centre’s original scheme, optimum interpolation. To use this approach, it was decided to rewrite the operational suite, and this became the IFS.

Implemented operationally in 1997 using the IFS, four-dimensional variational data assimilation (4D-Var) revolutionised the assimilation of weather observations to produce initial conditions for the Centre’s forecasting system, which in turn brought major improvements to the Centre’s forecasts.

It would be remiss of me not to mention the huge contribution that Météo-France made to the IFS. Météo-France’s close collaboration with ECMWF in what became the IFS-ARPEGE project proved essential for the scientific and technical development of the IFS and 4D-Var. 4D-Var provided a sound physical and mathematical basis for the use and the analysis of all the available weather observations, particularly the huge volumes of data provided by weather satellites. This in turn facilitated the strong close cooperation that developed between the Centre and EUMETSAT, and ESA.

REANALYSIS →

During my time as Director, the Cray vector computers, much loved by me, were replaced – firstly by Fujitsu VPP systems, which in turn were subsequently replaced by IBM systems. These hardware changes presented significant programming challenges to move from the simple architectures of the early Cray vector computers to distributed-memory and highly-parallel computing. And, of course, from vector to scalar arithmetic. These challenges were overcome brilliantly, and the Centre gained a world-leading reputation in the efficient use of distributed-memory and scalar computers.

It may appear that reanalysis has always been an activity at ECMWF, but in 1975 we had no models, data assimilation system or any other software; the first reanalysis activity was motivated by the First GARP Global Experiment (FGGE 1978/79) data, and a strong team was put together to develop the FGGE Re-Analysis, but it really had to wait until we had a forecasting system.

A reanalysis activity (ERA15) was proposed in the ECMWF long-term strategy being developed during 1989/90. Getting that strategy approved proved impossible and, consequently, funding any new activity was very difficult. At my first ECMWF Council meeting as Director, it was clear that the Council was not going to “roll over” and fund ERA15. And scepticism outside the NWP community made it impossible to get reanalysis included in the EU framework programmes. It looked like reanalysis was a “Cinderella activity”. Fortunately, “angels” were found who provided “mustard seeds”, and these contributions “persuaded” Council to provide new funds for reanalysis.

By 1997, the ERA15 datasets were proving to be very useful, and reanalysis was included in the EU framework programme – the Centre’s ERA-40 proposal was rated second among all the proposals submitted to the environmental part of the programme.

The breadth of the applications of reanalysis is quite staggering and was not really foreseen in the 1980s. Reanalysis matured to become an operational activity at ECMWF, with ECMWF’s reanalyses being the cornerstone of the Copernicus Climate Change Service (C3S).

MORE THAN AN ATMOSPHERIC MODEL →

When I retired in the summer of 2004, the Centre had also developed operational ensemble prediction systems; monthly and seasonal forecasting; and ocean wave prediction. The Centre’s operational systems and other programmes delivered a comprehensive archive of observations, analyses and forecasts which provided a major asset for research on the weather and climate for the world’s research community. Ironically, many of the new activities proposed in the ill-fated 1989/90 draft strategy were now part of the Centre’s programmes.

THE STAFF →

I had the time of my life at the Centre. Being in at the start was more than exciting for me. And, throughout my 29 years at the Centre, I was privileged to work with wonderfully talented colleagues who had the good fortune to be at the forefront of operational weather and climate prediction. Consequently, the Centre was, and still is, in a position to make a valuable contribution to the wellbeing of society.



Lennart Bengtsson



TENURE: JANUARY 1982 TO DECEMBER 1990

BACKGROUND →

I had the great opportunity to be involved with the COST project EMCC (European Meteorological Computing Centre for Research and Operations) in 1969 and from 1974 as member of the planning staff in Bracknell. The period prior to January 1982, when I became Director of ECMWF, was a highly active scientific and technical period including the development of a fully operational system for global medium-range weather forecasts. The work followed broadly the original plan laid out by the COST expert group (Woods, 2006).

SCIENCE ISSUES →

Among the main challenges in 1982 was to assimilate and integrate new types of observations that were obtained by different satellite systems. Data assimilation evolved therefore to become of equal importance to the forecasting model. The importance of data assimilation was originally not fully recognised by all, and it was not even seen as a main task for ECMWF. Indeed, some Member State representatives believed that they could provide the initial data so that ECMWF could focus solely on model development. Because of the rapid evolvement of non-synoptic observing systems, such a limited scope was never seriously considered. Consequently, data assimilation was to become a key task of the forecasting system. Based on the experience gained through the Centre's involvement in producing global datasets during The Global Weather Experiment, 1978–79, it was also agreed that the Centre should include reanalyses of past observations to support research as well as climate monitoring. ECMWF soon became a world-leading institute in data assimilation, with close cooperation with ESA and other non-European space agencies.

The original grid point model was soon replaced by a spectral transform model. Thanks to skillful work by the scientists and the technical experts, the Centre developed efficient algorithms that could make efficient use of the increasingly powerful computers including multi-processing facilities. The modelling work went through increasingly advanced treatment on how to handle radiation processes, clouds, turbulence, land and ocean processes. International intercomparison of forecast prediction skill demonstrated that the forecasts from ECMWF from the very start were superior to those from other main forecasting centres. This made ECMWF an attractive place to visit, with leading scientists visiting from all around the world, including the USA, Canada, Australia, China and Russia. Professor Edward Lorenz developed during such a visit a method to determine the predictability of medium-range forecasts.

TECHNICAL ISSUES →

ECMWF was lucky to obtain a Cray supercomputer in 1978. During my time as Director, we undertook major upgrades of increasingly more powerful Cray computers including multi-processing units. Multi-processing at the time was a new challenge but the technical staff learned to make clever and efficient use of the multi-processing facilities. Seymour Cray and his staff were impressed by the work of ECMWF that implied a fine cooperation with the Cray company. At the time the common procedure was to buy computers, but the Council subsequently decided to support a renting agreement including the ability for easier upgrades when new versions became available. That turned out to be a wise decision.

OTHER ISSUES →

It took a long time and considerable negotiations for Norway to become a member of ECMWF. A delayed membership implied also a considerable cost, as the rules required all previous budget contributions to be paid by a new Member State. A compromise was worked out where Norway did not have to pay any cost related to operational services but only the accumulated costs for research and development. The fact that the Council had to pass a unanimous decision on such a pragmatic solution speaks for the open mind and the fine cooperation between them.

With considerable frustrations and endless efforts, ECMWF had to wait a long time to obtain full membership of the Coordinated Organizations as we faced the veto of NATO. Fortunately, it was possible to bypass the previously strict rule when Lord Carrington became Secretary General of NATO. I was very grateful for his personal and pragmatic support in the process and, consequently, ECMWF could finally enjoy the same full pension and health rights as the staff of NATO, ESA and the OECD.

Reference
Woods. A., 2006: Medium-Range Weather Prediction. The European Approach. Springer, New York, NY. <https://doi.org/10.1007/b138324>

 TIMELINE

This timeline highlights some of the key developments which have marked the past 50 years at ECMWF. It is, of course, far from exhaustive, and the selection of items to include has proven to be a tall order. We hope to have covered as many as possible of the key scientific and technological developments and achievements, as well as the changes which have been made to our governing body, as and when new nations have joined our group of Member and Co-operating States.

1975–1984

- Computing
- Forecasting system
- Membership and governance
- Partnerships

1 AUG 1975

Interim computing facilities: CDC6600 Bracknell Met Office site

The first version of the ECMWF weather forecasting model was developed on a Control Data Corporation (CDC) 6600 computer from 1975 to 1978. The CDC6600 was one of the most powerful systems available at the time, but it still needed 12 days to produce a 10-day forecast. However, this showed that provided a suitably powerful computer could be acquired, useful forecasts could be produced.



15 JUN 1979

Official opening of ECMWF headquarters in Reading, UK

The new headquarters for the Centre were opened by H.R.H. Prince Charles, accompanied by the Lord Lieutenant of the County of Berkshire, in the presence of invited guests, members of the Council and its committees and the staff of the Centre.



1 NOV 1975

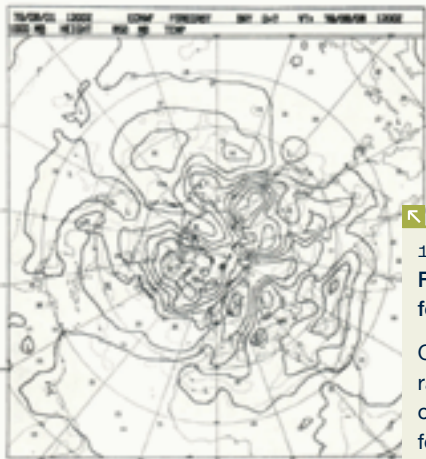
Convention enters into force

The ECMWF Convention was signed in October 1973 and entered into force on 1 November 1975, having been ratified by 13 Member States.



1 JAN 1974–31 DEC 1979

Professor Dr Aksel Wiin-Nielsen, Director



1 AUG 1979

First operational forecasts issued

Operational medium-range forecasting began on 1 August 1979, with forecasts produced to ten days ahead, five days per week.



1978

Cray-1 Series

The first supercomputer owned by ECMWF was installed on 24 October 1978 at the new site in Reading, UK.

The Cray-1A, serial number 9, used about 5 hours of CPU time to produce a 10-day forecast. This was more than 50 times faster than the CDC6600, thereby making the production of 10-day forecasts a feasible undertaking.



Austria



Belgium



Denmark



Spain



Federal Republic of Germany



France



Ireland



Yugoslavia



The Netherlands



Finland



Sweden



Switzerland



The United Kingdom



1 JAN 1976
Portugal



1 MAY 1976
Türkiye



1 SEP 1976
Greece



1 SEP 1977
Italy

Member States
1 NOV 1975



1 DEC 1980
Iceland

Co-operation
Agreement



1 AUG 1980

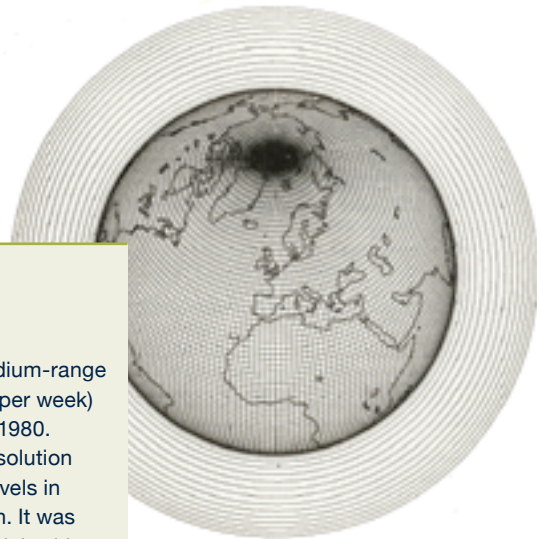
Fully operational forecasting starts

Full operational medium-range forecasting (7 days per week) began on 1 August 1980. The model had a resolution of 210 km with 15 levels in the vertical direction. It was an atmospheric model, with the oceans represented by a sea-surface temperature which was kept constant at each grid-point during the forecast.



1 JAN 1980–31 DEC 1981

Jean Labrousse, Director



1984

ECMWF, Reading



1984

Cray X-MP/22

1981

Telecommunication connections improved

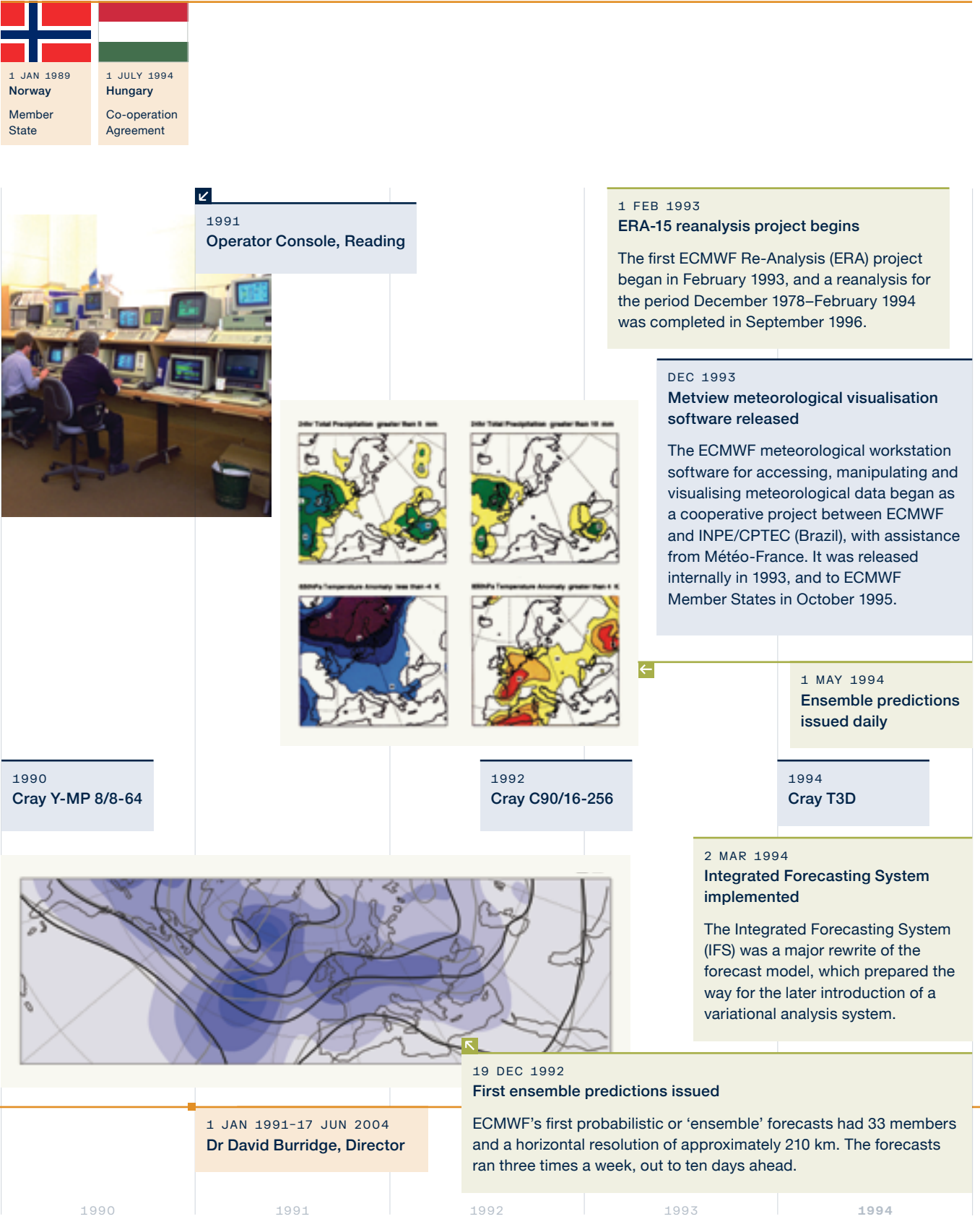
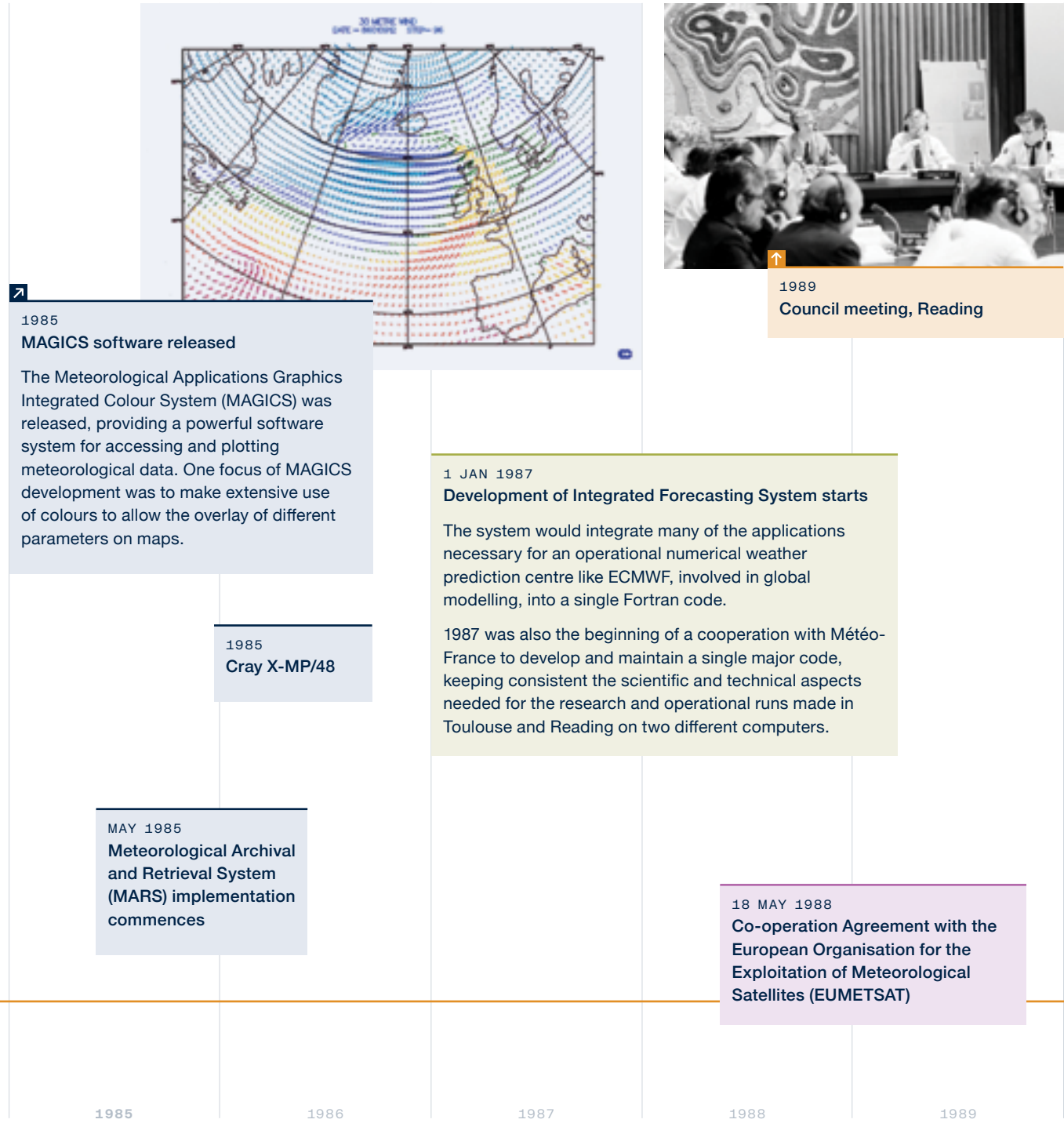
By the end of the year, more than half the Member States had medium-speed connections (2,400 or 4,800 bps). Analysis and forecast products were disseminated to Member States and onto the Global Telecommunications System as they became available during the evening and night.

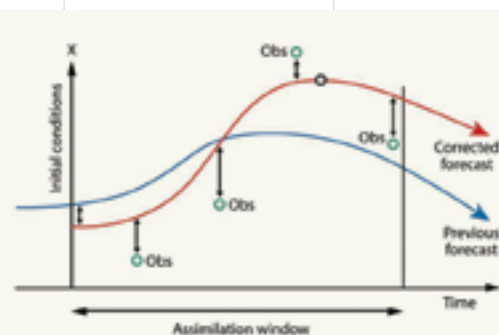
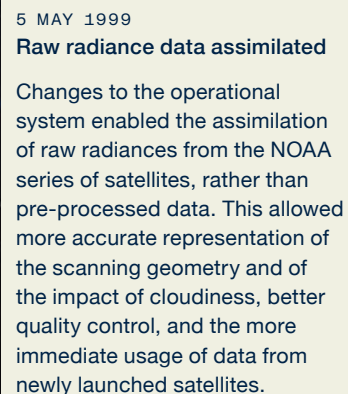
1 JAN 1982–31 DEC 1990

Professor Dr Lennart Bengtsson, Director

1985–1994

- Computing
- Forecasting system
- Membership and governance
- Partnerships

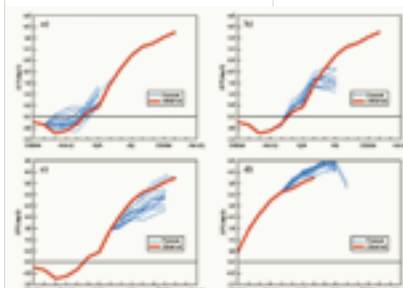




25 NOV 1997

4D-Var data assimilation introduced

Following the use of 1D-Var in satellite retrievals, and one year after the implementation of 3D-Var in global analyses, 4D-Var was made operational. The initial implementation used data in a six-hour time window, centred around the analysis time. It was the first-ever operational application of the 4D-Var technique successfully applied to a high-resolution assimilation and forecast system.



1997

Seasonal forecast system implemented

Research on predictability on the seasonal time scale in the early 1990s led to the implementation of the first ECMWF seasonal forecast system based on a global ocean–atmosphere coupled model in 1997.

29 JUN 1998

Coupled atmosphere–ocean wave forecasts

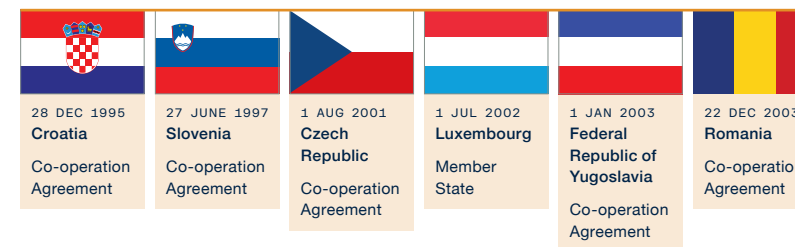
Up to now, ECMWF had run separate atmospheric and ocean wave forecast models. An hourly, two-way coupling of the atmospheric and ocean-wave models was introduced in 1998. Predicted ocean waves now provided information to the atmospheric boundary layer.

25 NOV 1998

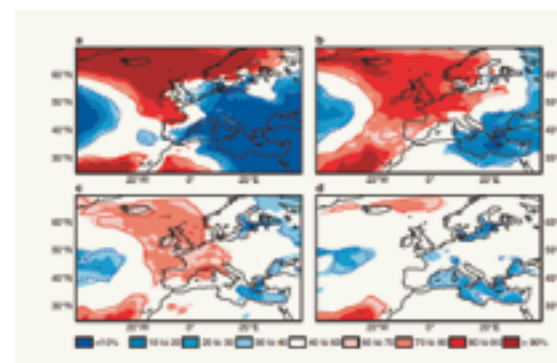
ERA-40 reanalysis production begins

Produced between 1998 and 2003, ERA-40 provided a 45-year reanalysis of the global atmosphere and surface conditions 1957–2002.

1996
Fujitsu VPP700



2000
Fujitsu VPP5000

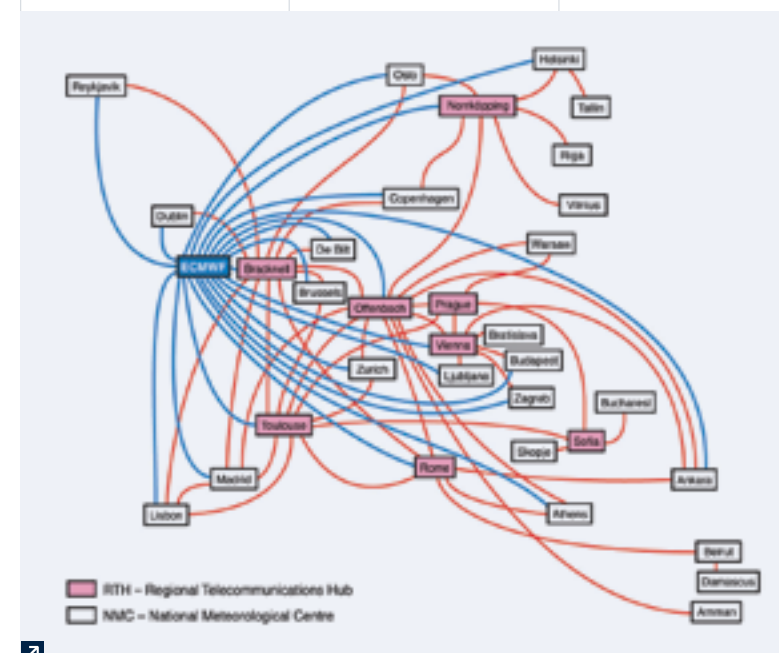


1 OCT 2004
**Sub-seasonal
predictions produced
operationally**

ECMWF had produced
monthly forecasts
(32-day) routinely
since March 2002 and
began producing them
operationally in October
2004, to fill the gap
between medium-range
and seasonal forecasts.

17 OCT 2004
**Meteorological
data archive reaches
1 petabyte**

The MARS archive reached the symbolic size of 1 petabyte of primary data, for a total of 8.6 billion meteorological fields.



MAR 2000
Regional Meteorological Data Communications Network (RMDCN) introduced

The RMDCN introduced a network infrastructure for the connection between ECMWF and its Member and Co-operating States. Before the RMDCN, meteorological data was exchanged using many bilateral links established between pairs of organisations wishing to connect their computer systems.

2003
IBM Cluster 1600
systems



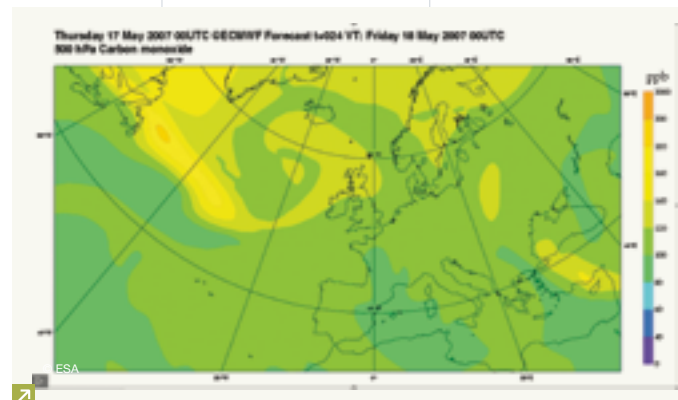
18 JUN 2004-
30 JUN 2011
**Dominique Marbouty,
Director-General**

2005–2014

- Computing
- Forecasting system
- Membership and governance
- Partnerships



2007
Met Ops room, Reading



17 MAY 2007
Atmospheric composition forecasting begins

The first graphs showing daily forecasts of reactive gases such as carbon monoxide and tropospheric ozone were published on the ECMWF website. These experimental daily forecasts were implemented as part of the EU-funded project GEMS, which was the first precursor Research & Innovation project that eventually led to the operational EU Copernicus Atmosphere Monitoring Service (CAMS).

2009
IBM 286 p6-575

31 MAY 2005
Co-operation Agreement with the European Space Agency (ESA)

1 AUG 2006
ERA-Interim reanalysis production begins
Global atmospheric reanalysis from 1979 to 2019.

2006
IBM 155 p5-575+

2005 2006 2007 2008 2009

7 NOV 2005 Estonia Co-operation Agreement	20 NOV 2006 Lithuania Co-operation Agreement	1 DEC 2006 Morocco Co-operation Agreement	5 NOV 2007 Montenegro Co-operation Agreement	1 JAN 2008 Slovakia Co-operation Agreement	30 APR 2008 Latvia Co-operation Agreement	12 JUL 2010 Bulgaria Co-operation Agreement	28 OCT 2010 Israel Co-operation Agreement	9 FEB 2011 North Macedonia ¹ Co-operation Agreement	1 JUN 2011 Iceland Member State

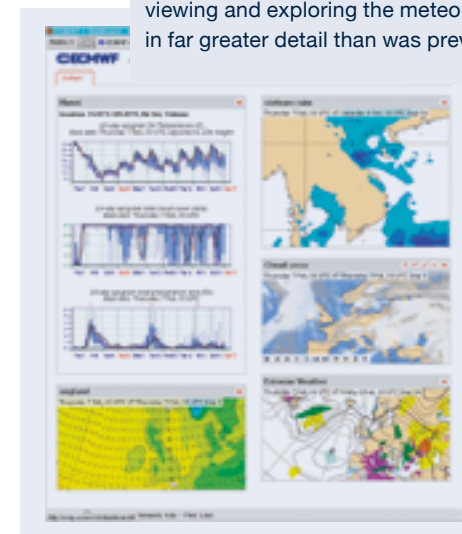
1 DEC 2012 Slovenia Member State

26 JAN 2010
High-resolution forecast resolution increases to 16 km

6 JUNE 2010
Amended Convention enters into force
In 2005, the Convention was amended to allow new Member States to join the organisation, with the amendments coming into force on 6 June 2010.

JUN 2010
Ensemble of Data Assimilations (EDA) implemented
The EDA was introduced to give information about the uncertainty in the analysis of atmospheric observations that provides the starting point for ECMWF's forecasts.

2010 2011 2012



JUN 2012
ecCharts interactive tools released
The new suite of web applications called ecCharts provided fast access to ECMWF's medium-range forecast data as soon as the information was available, offering an interactive set of tools for viewing and exploring the meteorological situation in far greater detail than was previously possible.

2012
IBM 768 POWER7-775

2011
ECMWF designated as the hydrological forecasting computational centre for the Copernicus Emergency Management Service (CEMS)

1 JUL 2011–31 DEC 2015
Professor Alan Thorpe, Director-General

JAN 2014
Scalability Programme commences

2 JUN 2014
ERA-20C reanalysis production begins
ECMWF's first atmospheric reanalysis of the 20th century, from 1900 to 2010.

1 JUL 2014
ECMWF Fellowship programme introduced



2014
Cray XC30



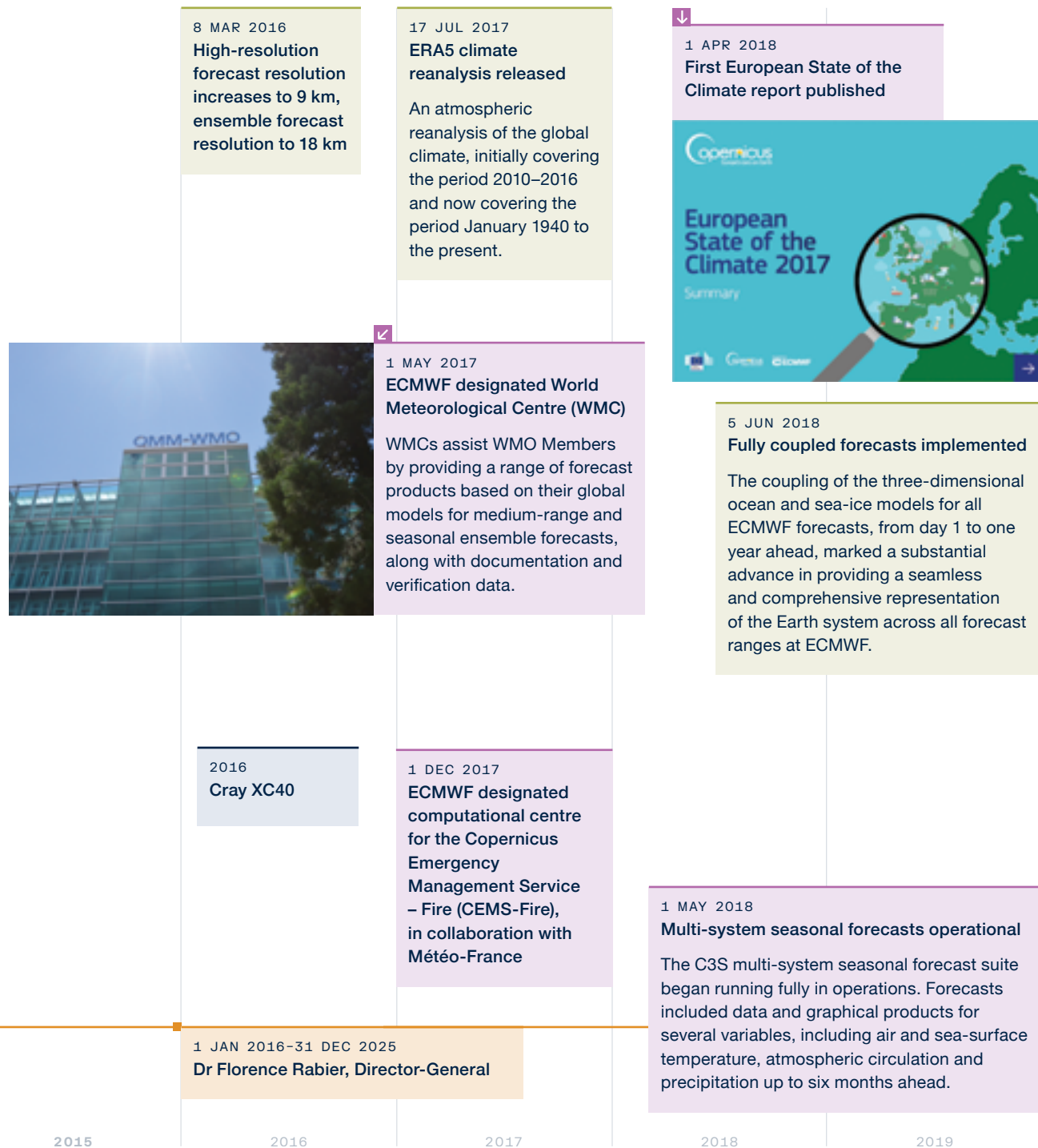
11 NOV 2014
ECMWF to run two Copernicus Services
The European Commission and ECMWF signed an agreement for ECMWF to manage the Copernicus Climate Change Service and Copernicus Atmosphere Monitoring Service.

2013 2014

¹ Formerly known as the former Yugoslav Republic of Macedonia.

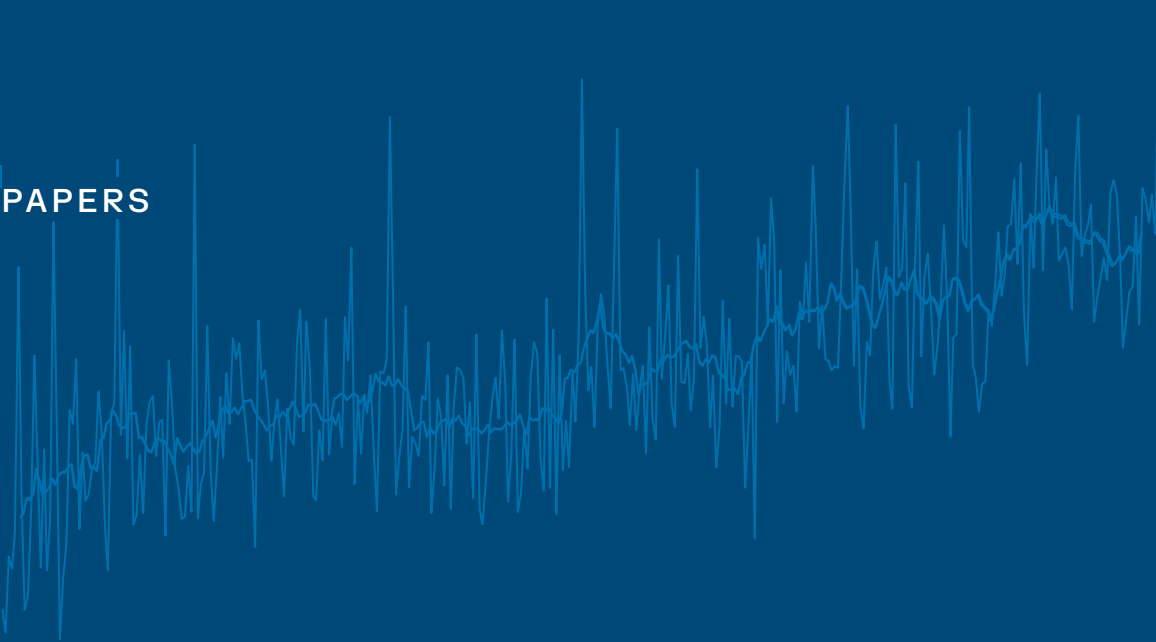
2015–2025

- Computing
- Forecasting system
- Membership and governance
- Partnerships





PAPERS



Some of the key areas of work at ECMWF over the past 50 years are described here, by staff who have led or contributed to their development. The topics covered do not represent the full spectrum of ECMWF’s activities, but they do represent the essence of our work.

We want to thank all the staff at ECMWF who have taken the time to contribute to these papers.

References for all the papers can be found in the ‘References’ section.



Fifty years of Earth system modelling at ECMWF

ABSTRACT



Earth system modelling at ECMWF has been a story of growth and success during the past 50 years. The first operational medium-range forecast was produced on 1 August 1979 and the first ensemble predictions on 24 November 1992. The number of simulated parameters at all grid points spanning the global atmosphere at a given time represented in ECMWF’s Integrated Forecasting System (IFS) simulations went from below 1 million (208 km grid spacing, 19 vertical levels) to 452 billion (9 km, 137 vertical levels, 51 ensemble members), with single forecast predictions in research mode simultaneously involving more than 1,400 billion prediction points (0.7 km grid spacing, 137 vertical levels). The investment in physical modelling has brought world-leading advances in numerical methods and physical parametrizations for radiation, boundary-layer turbulence, convection and cloud processes. These innovations have been developed and implemented in collaboration with many Member and Co-operating States and the wider international science community.

The growing interest in prediction at sub-seasonal and seasonal timescales, as well as the growing availability of satellite data, has fostered a continuous increase in resolution and complexity of ECMWF’s IFS, including extensions to represent the stratosphere, land-surface processes, ocean, sea ice and surface waves.

Targeted developments to describe initial and model process uncertainties have resulted in their trustworthy representation, and a single forecast realisation has expanded into ensembles with 51 simultaneous members, providing information on a range of different weather scenarios at time ranges from a few days to a season ahead. Moreover, the use of a variational data assimilation

algorithm to create the initial conditions for successful forecasts mandated the development of tangent-linear and adjoint model versions.

The addition of a range of aerosols and chemical species as well as greenhouse gases such as CO₂, CH₄, and ozone, together with coupling to relevant chemistry models, facilitated operational atmospheric monitoring forecasts that are part of the EU Copernicus Atmosphere Monitoring Service.

Building on the IFS and other models, the EU Destination Earth initiative goes a step further in building digital twins of the Earth system to harness the power of global kilometre-scale simulations, with a high level of interactivity, and which feed the growing demand of impact sector models.

All these developments are embedded in the exponentially growing supercomputing capacity of chip, storage and networking technologies, fostering both scientific and technological evolutions, resulting in the IFS becoming one of the world’s most efficient, massively parallel Earth system models and one of the leading applications of exascale supercomputing.

The model development at ECMWF is a pertinent example of what can be achieved with a common goal. An incredible amount of help and support from scientists all over the world has influenced the progress and success in both physical and now data-driven modelling. In return, ECMWF has continued to share its knowledge of how to address today’s and future prediction challenges in a changing climate.

INTRODUCTION →

In 1979, the European Centre for Medium-Range Weather Forecasts (ECMWF) started to disseminate weather predictions with the remit to provide skilful and reliable medium-range forecasts to its Member States. Countries had to consider the financial and political investment in a European approach to medium-range forecasting during a time when such forecasts were not yet considered useful (see Woods (2005) for anecdotes).

The 1st ECMWF technical newsletter (1979) stated: “The central premise of the development of these operational forecasts is that the atmosphere may be regarded as a compressible fluid, its behaviour being described by the Navier-Stokes equation and the thermodynamic equations concerned with sources, sinks and the transfer of energy.”

Key for success was careful attention to the initial state of such forecasts by means of data assimilation, the physics governing the non-linear evolution of the Earth system, and the balance between accuracy and cost when solving the discrete form of these equations on the latest available high-performance computing (HPC) architectures. The model development at ECMWF was thus driven by what could be reasonably initialised considering new prognostic variables, the accurate representation of the non-linear evolution of physical processes, and the overall efficiency of execution in a parallel (multi-node) computing environment.

From the beginning this necessitated a suitable infrastructure, later termed the Integrated Forecasting System or IFS, used in both forecasts and data assimilation. In close collaboration with Météo-France, the code infrastructure of the IFS and ARPEGE (Action de Recherche Petites Echelles Grandes Echelles)¹ was born and “many scientific projects, sub-projects, and operational and research options have been built around this initial code since then, covering both data assimilation and forecasting aspects” (Pailleux et al., 2014).

This paper provides an overview of the Earth system model developments at ECMWF during the past 50 years, gives an update on what Earth system modelling entails today, and outlines how Earth system modelling will likely evolve in the future. It will tell the story of how an atmosphere-only model running on a single megaflop compute chip has turned into an Earth system model that is running at kilometre-scale resolution on exascale supercomputers and is used seamlessly for a variety of forecast products from days to seasons, for numerical weather prediction (NWP), climate scenarios and environmental forecasts.

¹ ECMWF and Météo-France share a common global NWP software, termed IFS by ECMWF and ARPEGE by Météo-France.



FIFTY YEARS OF EARTH
SYSTEM MODELLING
AT ECMWF →

A MODEL OF THE
GLOBAL ATMOSPHERE

“ OPERATIONAL FORECASTS BEGAN AT ECMWF IN 1979 AND THE INTEGRATED FORECASTING SYSTEM CODE, DEVELOPED JOINTLY WITH MÉTÉO-FRANCE, WAS INTRODUCED IN MARCH 1994. THE INTEGRATED FORECASTING SYSTEM AND ITS CONTINUING DEVELOPMENT REPRESENTS A HUGE AND HIGHLY SUCCESSFUL COLLABORATIVE EFFORT WITH OUR MEMBER AND CO-OPERATING STATES AND MANY OTHERS.”

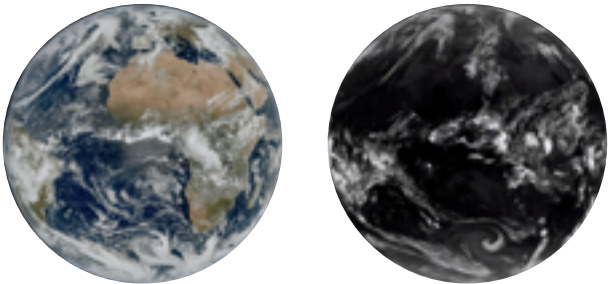
The first ECMWF operational model in 1979 was a finite-difference grid-point model of the global atmosphere with 48 computed latitudes (today we have 2,560) and 15 vertical levels, and with the first physical parametrizations inspired primarily by early collaborations with the Geophysical Fluid Dynamics Laboratory (GFDL). In April 1983, a spectral model was introduced at ECMWF, building on significant developments of the spectral transform method (independently promoted by Eliassen et al. (1970) and Orszag (1970)). In addition, ECMWF developed a fast Fourier transform (Temperton, 1983) that facilitated the dual representation of global prognostic variables in spectral space and grid-point values at specific latitude-longitude locations that satisfied quadrature rules to determine the prognostic variables of temperature, wind, pressure and moisture (Temperton, 1991; Wedi et al., 2013, 2014).

The original spectral model was replaced by the IFS/ARPEGE code, with its first operational use at Météo-France in October 1993, when a stretched-tilted version of ARPEGE became operational, and at ECMWF in March 1994. The efficiency of time-stepping was boosted by a very stable semi-implicit solution procedure (Simmons et al., 1978; Benard, 2003) and an efficient two time-level semi-Lagrangian atmospheric transport scheme (Temperton et al., 2001).

The success and superior efficiency of the semi-Lagrangian, semi-implicit, spectral transform method in NWP, in comparison to alternative methods, has been overwhelming, and many operational forecast centres made the spectral transform their method of choice (Williamson, 2007), although several modelling centres have subsequently adopted alternative approaches. In contrast to the expectations of many modelling experts across the world, ECMWF still successfully

Figure 1: The detail and realism of a cutting edge 2.8 km simulation with the IFS is a striking demonstration of the progress in Earth system modelling at ECMWF over the last 50 years. Digital twinning of a 2.8 km simulation (bottom) compared to the latest in European satellite technology, EUMETSAT's Meteosat Third Generation MTG-I (top), compared in observation space during the day (left) and at night (right), from the inverse observation operator applied to the hydrostatic IFS model fields (images created by Philippe Lopez, ECMWF).

Satellite observations



2.8 km simulation

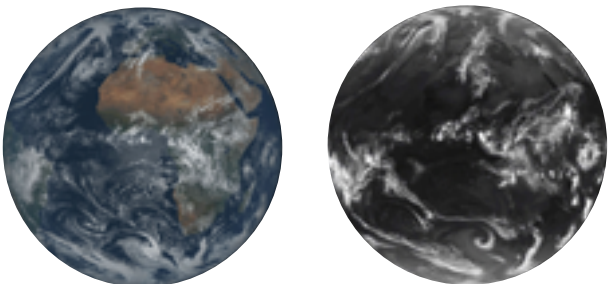
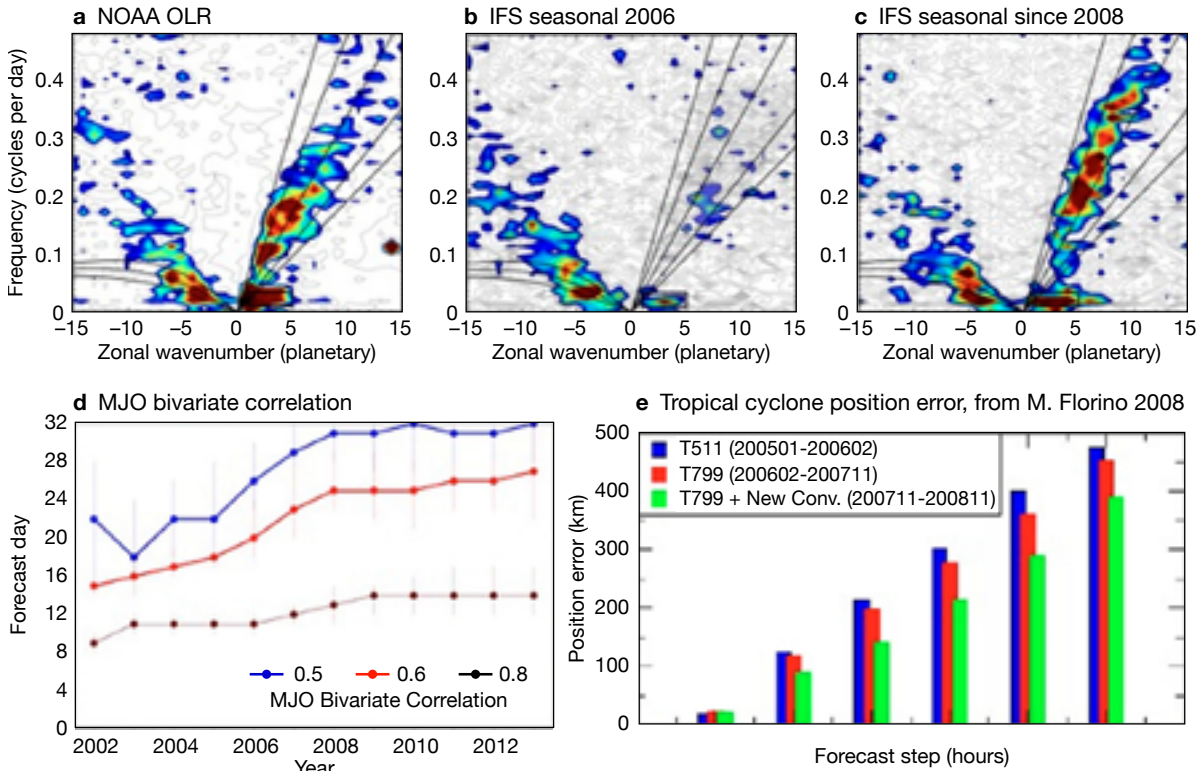


Figure 2: Wavenumber frequency spectra of the outgoing longwave radiation (OLR) from NOAA data (a) and from multi-year integrations with the IFS using the operational cycle in 2006 (b) and with the version that became operational in 2008 (c); the Madden-Julian Oscillation (MJO) spectral band is highlighted by the black rectangle. The bottom row shows (d) skill of the IFS predictions for the MJO between 2002 and 2013 as given by the bivariate correlation with the observed empirical orthogonal functions for wind and outgoing longwave radiation, a value of 0.6 (red line) delimits skilful forecasts and (e) statistics of tropical cyclone positions errors (km) as a function of forecast lead time from the 40 km resolution forecasts in 2005/6 (blue), the 25 km forecasts in 2006/7 (red) and the 25 km forecasts in 2008 (green). Reproduced from Lin et al. (2022).



uses the spectral transform method in operations today, with significantly increased performance and horizontal resolution (Wedi 2014; Malardel et al., 2016). The numerical accuracy and stability were improved with a vertical finite element scheme (Untch and Hortal, 2004), allowing the vertical extension of the model into the stratosphere to accommodate the increasing number of satellite observations in the middle atmosphere and to exploit potential gains in longer term predictability; see Polichtchouk et al. (2021) for a review.

The Navier-Stokes and thermodynamic equations are fundamental to predictions using physical models, but their numerical solution and the non-linear transfer of energy across scales of motion can be impacted by the solution procedures and approximations. One extremely successful (stable and efficient) approximation still made in routine ECMWF forecasts is the hydrostatic approximation, although this is being continuously reviewed, with the ultimate aim of relaxing the approximation in ECMWF's global applications (cf. Wedi and Smolarkiewicz, 2009; Zeman et al., 2021). Non-hydrostatic models are routinely used in limited-area, high-resolution applications across weather services today (Bubnová et al., 1995; Benard et al., 2010) and increasingly in global forecast systems (Wood et al., 2014; Zängl et al., 2015). Research at ECMWF continues to investigate stable and efficient non-hydrostatic modelling frameworks for global kilometre-scale applications on emerging HPC (e.g. Smolarkiewicz et al., 2014; Voitus et al., 2019; Kühnlein et al., 2019; Melvin et al., 2024), in comparison with competitive hydrostatic reference solutions (Wedi et al., 2020; Rackow et al. 2024) and observations (see Figure 1).

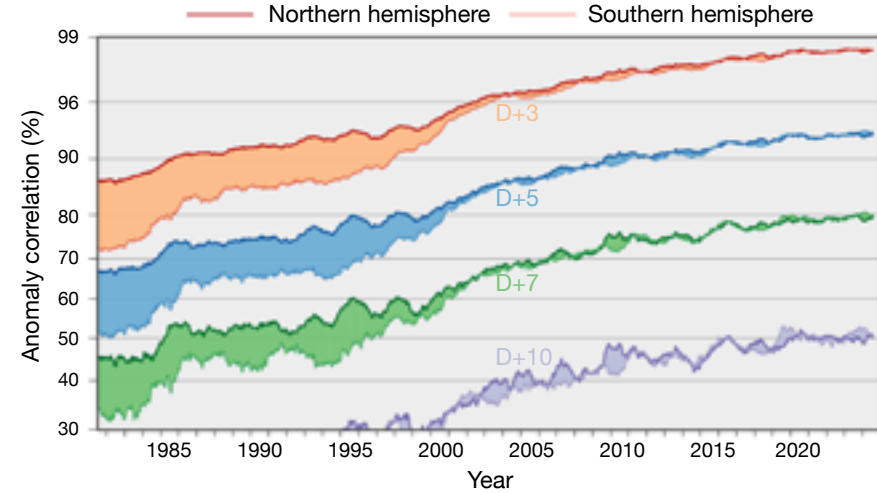
The sources and sinks of energy represented by the physical parametrizations of the IFS have undergone significant evolution over time. There have been many innovative developments in atmospheric radiation (Morcrette et al., 2008, Hogan and Bozzo, 2018), convection parametrization (Tiedtke, 1989; Bechtold et al, 2001; Lin et al., 2022), cloud microphysics and extending the grid-point prognostic variables of clouds and cloud hydrometeors (Tiedtke, 1993; Forbes et al., 2011), and turbulent (Beljaars and Holtslag, 1991) and orographic drag (Lott and Miller, 1996; Beljaars et al., 2004), all contributing significantly to enhanced medium-range forecast skill. A good example of a significant breakthrough in predictive skill is the revised representation of convection during the years 2006 to 2008; see Figure 2.

The developments in physical parametrization have also made possible the growing portfolio of specialised products for a diverse and growing range of applications requiring weather information, such as different forms of hydrometeors, clear air turbulence, Extreme Forecast Indices (EFIs), convective storm triggers, hydrological parameters, and atmospheric composition and air pollution parameters.

The modelling efforts at ECMWF have been regularly scrutinised and ideas exchanged with modelling centres worldwide. This process is facilitated through ECMWF’s long-standing membership of the WMO Working Group on Numerical Experimentation (WGNE²), pioneering intercomparison projects such as AMIP I (Gates et al., 1998), AMIP II (Branković et al., 1999), Transpose-AMIP (Williams et al., 2013), DCMIP2016 (Ullrich et al., 2017), and DIMOSIC (Magnusson et al., 2022), assessing IFS forecast model skill within an international multi-model reference and sharing common systematic model errors and approaches (Zadra et al., 2018; Frassoni et al., 2023).

All these developments have been achieved through collaboration with many Member and Co-operating State weather services and scientists worldwide, as evidenced by the literature references in this article, to drive forward the ambition of extending the predictive skill of ECMWF’s forecasts. Figure 3 shows the evolution in forecast skill of 500 hPa geopotential height and exemplifies well the progress that has been made in medium-range forecasting since ECMWF’s inception.

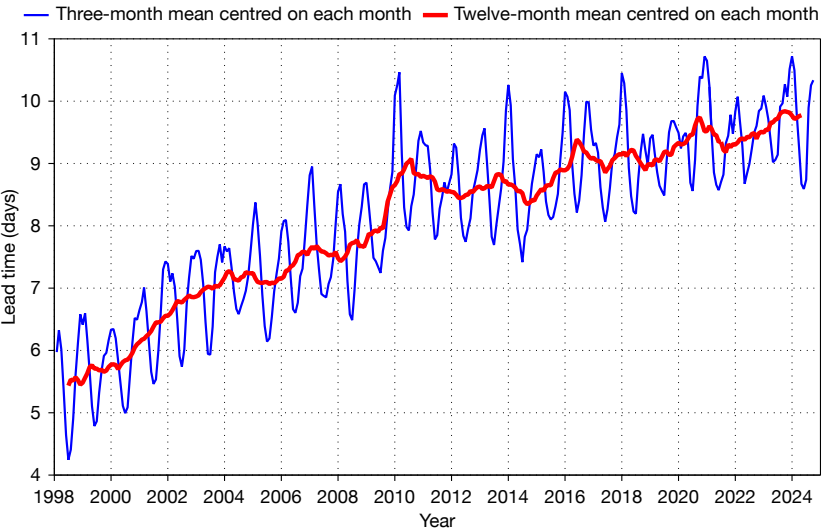
■ **Figure 3:** Anomaly correlation coefficients of 3-, 5-, 7- and 10-day ECMWF 500 hPa geopotential height forecasts for the extratropical northern and southern hemispheres, plotted in the form of annual running means of archived monthly-mean scores for the period from 1 January 1981 to 31 January 2025. Values plotted for a particular month are averages over that month and the 11 preceding months. The shading shows the differences in scores between the two hemispheres at the forecast ranges indicated. Updated version of Simmons and Hollingsworth (2002).



2 <https://www.wcrp-esmo.org/working-groups/wgne>

INCREASING MODEL COMPLEXITY

■ **Figure 4:** The graph shows the number of days into the future that the ensemble system provides useful forecasts of 850 hPa temperature for the northern hemisphere extratropics and how this has evolved over time. Useful is defined by a particular global verification metric (the continuous ranked probability skill score (CRPSS)) reaching a certain level (25%). A CRPSS score of 100% represents a perfect forecast, while 0% represents a system with no skill (compared with using long-term time averages).



3 <https://www.nemo-ocean.eu/doc/>
4 <https://destination-earth.eu/>

Extending the complexity of the IFS model in a seamless prediction approach across timescales opens the door to sub-seasonal and seasonal predictions, while offering a more complete description of the hydrological and carbon cycle required for enhanced Earth system monitoring and prediction.

The IFS system was adapted to describe the coupling between the atmosphere and the ocean through the addition of a three-dimensional ocean and sea-ice model. When introduced in 1997, this approach was only used for seasonal ensembles, before being extended to the sub-seasonal range. In 2018, the same coupled system was introduced across all forecast ranges (Buizza et al., 2017). The ocean component in both the model and the assimilation activities is based on the NEMO³ ocean model and the associated NEMO sea-ice engine (now SI3). Alternatively, the IFS can be coupled to the FESOM-2 ocean and sea-ice model (Koldunov et al., 2019). This is done in Destination Earth⁴ to create alternative warming scenarios and to explore alternative ocean discretisations in view of their scalability and skill in kilometre-scale simulations (Rackow et al., 2024).

The coupling between wind and waves, non-linear surface wave–wave interactions and ocean freak waves (Janssen, 2004, 2013), also has a long history at ECMWF with co-development of the surface wave model WAM (now ecWAM), which began in 1991. In the context of ocean–atmosphere coupling, turbulent effects on the ocean mixed layer have been included and wave–ice interactions are to be added in 2025. Recent improvements in the air–sea interface had significant impact on the accurate representation of extreme events (Majumdar et al., 2023).

In the same spirit, a range of land-surface interactions were added (Balsamo et al., 2009; Agustí-Panareda et al., 2014; Boussetta et al., 2021). This brought to the forefront the accurate and stable coupling across land and ocean boundary layers and the atmosphere (Best et al., 2004; Beljaars et al., 2018 for a review). In recent intercomparisons of latent and sensible heat fluxes at land observation sites, physics-based land-surface models including the IFS schemes are shown to be correct most of the time, but are still outperformed, especially in extremes, by relatively simple, out-of-sample empirical models (Abramowitz et al., 2024). This raises interesting new research questions on the required complexity of boundary layer parametrization, machine learning (ML), and parameter calibration.

“ IN 1992, ECMWF AND THE NATIONAL CENTRES FOR ENVIRONMENTAL PREDICTION (NCEP) WERE THE FIRST TO INTRODUCE OPERATIONAL ENSEMBLE FORECASTS. TODAY, ENSEMBLES ARE AT THE HEART OF ECMWF’S FORECASTS ACROSS MEDIUM, SUB-SEASONAL AND SEASONAL TIMESCALES.”

The step to use variational data assimilation (see ECMWF 50th anniversary paper on data assimilation) mandated the development of corresponding tangent-linear and adjoint models for both the IFS model dynamics and the physics (Mahfouf, 1999; Janiskova et al., 1999). This required highly challenging, bespoke developments, both technical and scientific, and a continuous process is needed to match and ensure compatibility with proposed innovations in the non-linear model. The rigour of this process equally ensured the quality of both the non-linear and tangent-linear models.

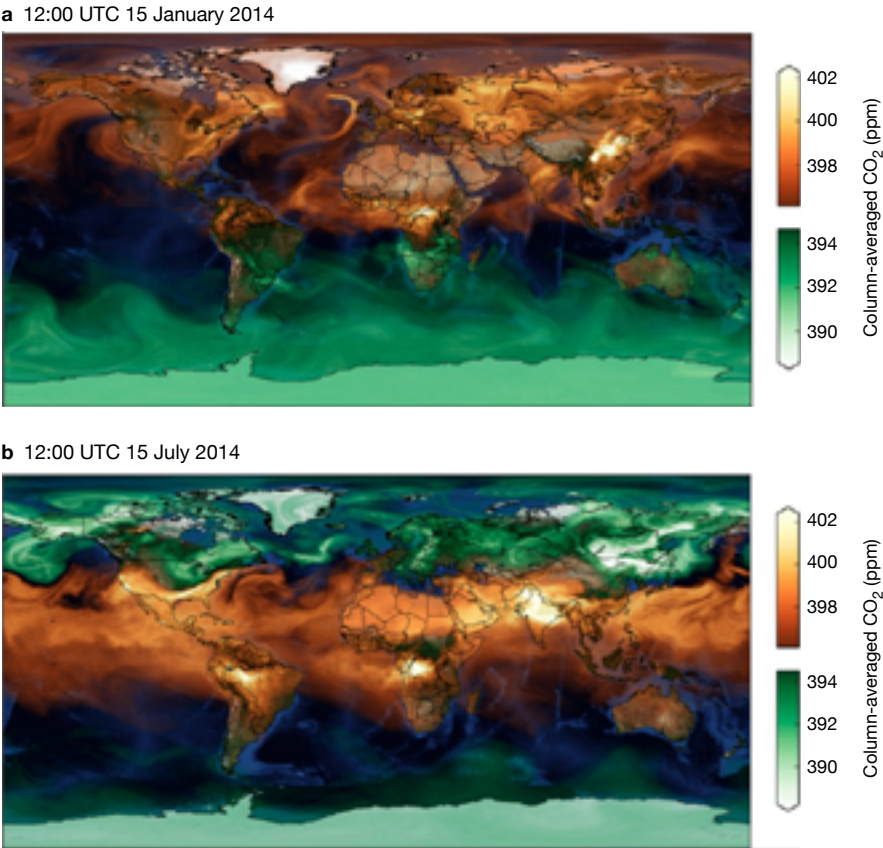
In 1992, ECMWF and the National Centers for Environmental Prediction (NCEP) were the first to introduce operational ensemble forecasts (Buizza, 2019), initiating a distinct shift in thinking about uncertainty in weather forecast products. While the ECMWF high-resolution forecast (HRES) has been a distinct feature of the product spectrum, ECMWF concludes this shift in thinking towards ensembles at the core of medium-range weather prediction, by discontinuing HRES in 2025 and producing instead a data stream with an ensemble of forecasts (ENS) plus a single unperturbed control. Ensembles are at the core of ECMWF’s forecasts across the medium, sub-seasonal and seasonal forecast ranges and they support a range of environmental services. Addressing the fundamental question of predictability using ensemble prediction forecasts, which span the envelope of likely and less likely outcomes, culminates today in forecast statistics derived from an ensemble of 50 perturbed high-resolution (9 km) ensemble members (accounting for a range of initial and model uncertainty) (plus ENS control) instead of a single deterministic forecast (Buizza and Palmer, 1995; Leutbecher et al., 2017; Lang et al., 2021). Figure 4 indicates how the skill of ECMWF ensemble forecasts has increased significantly over the past 33 years of operational use.

Moreover, the extension of ensemble predictions to sub-seasonal (Ferranti et al., 1990; Vitart et al. 2008, 2017) and seasonal (up to 1 year or more ahead, e.g. Johnson et al., 2019) time ranges added to ECMWF’s product portfolio. Providing anomalies in sub-seasonal and seasonal range predictions requires robust background statistics. These statistics can be from a model climate, derived from hindcasts, or from ensemble re-forecasts made using consistent initial conditions of past years, and this takes significant computational effort. Products such as the EFI make use of these forecasts.

While model complexity increased from atmosphere-only to explicitly describe exchanges with the land, hydrosphere (soil, snow, rivers and lakes), ocean, sea ice, and waves, atmospheric composition is another key ingredient in the IFS. The description of all these processes is key to providing today’s Copernicus services⁵ that cover atmospheric composition, marine, land and climate monitoring. The addition of a range of aerosols and chemical species as well as greenhouse gases such as CO₂, CH₄, and ozone, together with coupling to relevant chemistry models, facilitated complementary atmospheric monitoring forecasts (e.g. depicted in Figure 5) that led to the routine offering of atmospheric monitoring services in Copernicus. Beyond monitoring, there are clear two-way interactions between the properties and evolution of the simulated atmospheric circulation and the relevant addition of atmospheric composition components. In particular, simulating the combined hydrological and carbon cycle is challenging modellers to reduce biases in transport, vegetation, land and ocean surface interactions at the highest possible resolutions to also represent anthropogenic and natural emissions (Agustí-Panareda et al., 2014, 2019).

5 <https://www.copernicus.eu/>

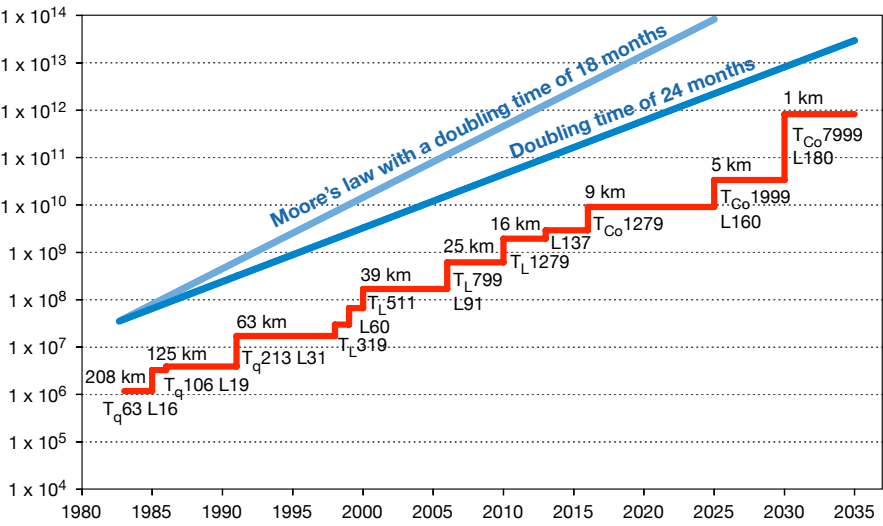
■ **Figure 5:** Snapshots of column-averaged CO₂ (ppm) above the global mean (in red colours) and below the global mean (in green colours) on 15 January 2014 (a) and 15 July 2014 (b) at 12:00UTC from the Copernicus Atmosphere Monitoring Service global CO₂ forecast at high horizontal resolution (~9km). Reproduced from Agustí-Panareda et al. (2019).



A SUPERCOMPUTING APPLICATION →

■ **Figure 6:** Steady increases in horizontal resolution (measured in degrees of freedom given by the number of grid points and vertical levels across the sphere multiplied by the prognostic variables forecast) of the IFS model at ECMWF as a result of increased computing capacity, combined with progress in numerical methods and hardware adaptation efforts. Note that the figure shows the understanding as of 2019, and while flagship simulations and research are already conducted at 1 km, their use in an operational NWP context is uncertain, as is Moore’s law. The ECMWF 50th anniversary paper on high-performance computing provides an overview of the specific HPC evolution at ECMWF. Reproduced after Schulthess et al. (2019).

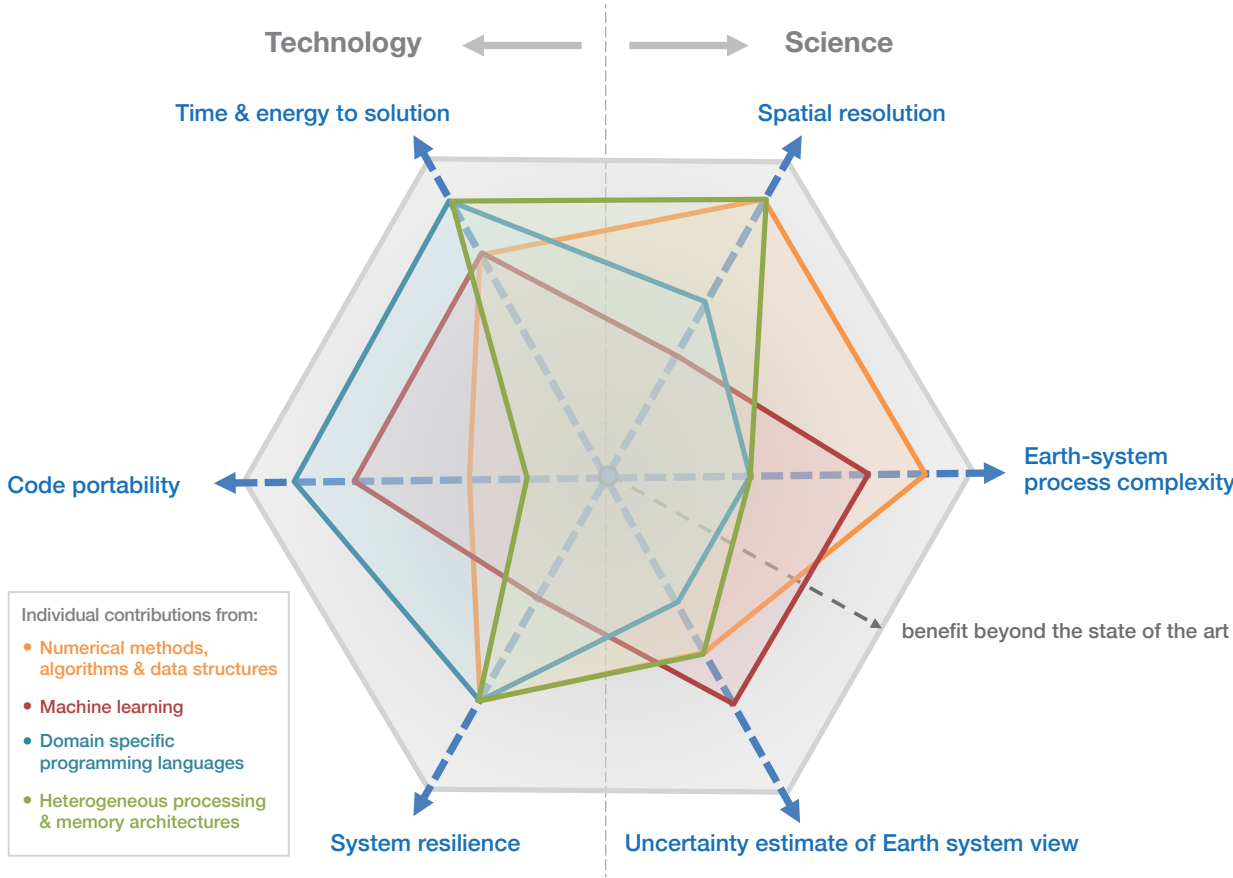
The time-to-solution and energy-to-solution efficiencies of the IFS model have played a significant part in the doubling of horizontal resolution every 8 years (i.e. roughly halving the distance between adjacent model grid points), depicted in Figure 6 (Schulthess et al., 2019), and significantly increasing the number of simultaneous forecast ensemble members.



■ **Figure 7:** The expected contribution of the main system developments necessary to achieve key science and computing technology performance goals. The distance from the centre of the hexagon indicates the magnitude of the individual contributions towards enhanced efficiency for increased spatial resolution, more Earth system complexity and better uncertainty information provided by ensembles as well as resilient, portable and efficient code and workflow execution. Reproduced from Bauer et al. (2021).

With 51 members at 9 km resolution, the ECMWF ensemble is the highest-resolution global ensemble prediction system in the world today. The efficiency of the IFS has been significantly improved, arguably in equal parts by algorithmic changes, continuous refactoring, and supporting parallel computing concepts. The latter include vectorisation across multiple nodes, message passing interfaces (MPI) (Barros et al., 1995), and shared memory programming (e.g. Modzynski et al., 2015) in support of many-core architectures, and more recently accelerator offloading (using OpenMP and OpenACC programming models) in support of accelerator technologies such as graphical processing units (GPUs) (e.g. Mueller et al., 2019). New programming paradigms have been used that are not only based on modularised Fortran code (already mentioned in the 1979 ECMWF technical newsletter on the Cray!) but a mixture of Fortran, Python, C/C++, abstracting data structures such as Atlas (Deconinck et al., 2017) or the FieldAPI, and bespoke source-to-source translation tools such as LOKI⁶. One of the main motivations behind the new methods is an attempt to disentangle the tasks of HPC machine optimisation and scientific progression (known as separation of concerns) (cf. Ubbiali et al., 2025).

Together with its Member States, ECMWF has proactively invested in the Scalability Programme during the last ten years, and is continuing efforts to evolve efficiency, modularisation, coupling aspects and data structures. The announced digital revolution of Earth system science (Bauer et al., 2021) encompasses some of the different foci depicted in Figure 7.



6 <https://github.com/ecmwf-ifs/loki>

ECMWF'S EARTH SYSTEM MODEL TODAY →

After half a century of evolution, the IFS is still operational and efficient and has embraced many developments to serve the increasing envelope of ECMWF's activities. Today, the IFS is an Earth system model that can represent most of the relevant processes in the atmosphere, ocean, sea ice, waves and land surface, representing atmospheric chemistry and increasingly closing the water and carbon cycles.

It is used as the forecast model for NWP running medium-range 50-member ensemble forecasts and a control forecast at 9 km resolution four times a day, as well as sub-seasonal and seasonal ensemble predictions and hindcasts to build the statistics required for reliable forecasts. The IFS is also used within the Copernicus Atmosphere Monitoring and Emergency Management Services. Furthermore, past investment in the IFS has laid important foundations for today's developments in Destination Earth, which readily continues ECMWF's role in supporting the digital preparedness of its Member and Co-operating States (Wedi et al., 2025). The European community Earth System Model (EC-Earth) builds on the OpenIFS⁷ framework for climate projections and Destination Earth uses the IFS in multi-decadal projections.

In Destination Earth, the IFS is combined within an innovative software framework that integrates computational, data, and application layers. This setup allows dynamic and interactive Earth system modelling, designed explicitly to simulate plausible scenarios of future weather and climate and to explore *what-if* questions. Destination Earth implements novel pathways for exploiting the geographically distributed HPC and big data infrastructure provided by EuroHPC⁸. Destination Earth advances Earth system science further by establishing an operational framework for multi-decadal, multi-model climate projections, connected to applications that transform vast climate data into actionable insights for sectors exposed to climate risks in support of both immediate and longer-term climate adaptation strategies.

The IFS is used in trailblazing daily global weather forecasts at kilometre-scale (4.4 km) as well as 5 to 10 km multi-decadal climate scenario projections. These forecasts are efficiently run routinely on several of the largest supercomputers in Europe and worldwide, including the EuroHPC machines Lumi, Leonardo and Marenostrum5. The IFS has been deployed on supercomputers in the US, including Summit and Frontier as part of several successful INCITE projects (Wedi et al., 2020; Polichtchouk et al., 2025), as well as the Fugaku supercomputer in Japan. Most recently the IFS is forming part of the JUPITER Early Access Programme, Europe's first exascale supercomputer, where the ECMWF forecast model received an "exceptionally outstanding recommendation" and has been accepted for early access on the JUPITER Booster system in 2025.

7 <https://www.ecmwf.int/en/research/projects/openifs>

8 <https://eurohpc-ju.europa.eu/>

WHAT WILL THE FUTURE
BRING? →

The recent rise in pure ML models that are competitive with conventional physics-based models, in terms of their deterministic and ensemble skill scores, has generated a significant uncertainty about the future of conventional weather forecasting and Earth system predictions. ECMWF was one of the first to investigate the potential of deep learning for weather forecasting (Dueben and Bauer, 2018). ECMWF is the first operational weather centre to provide daily operational artificial intelligence (AI)-based weather predictions, with ECMWF’s own machine-learned weather forecast model called the Artificial Intelligence Forecasting System (AIFS). The AIFS framework is being further developed to operate in both deterministic and ensemble mode (Lang et al., 2024a; Lang et al., 2024b).

Machine-learned models are not directly based on physical laws but rely on data generated by physics-based models and require significant compute resources during the training phase. However, they are much cheaper to execute than physical models when running forecast simulations in a critical time window and can substantially accelerate forecast production. Therefore, machine-learned approaches and the previously described Earth system modelling have different advantages and disadvantages, and the future will show how the two distinct modelling frameworks (at ECMWF the IFS and the AIFS) will co-exist and develop.

For the foreseeable future, it is unlikely that ML models will entirely replace conventional Earth system models, but physical Earth system models now have a serious competitor for operational NWP and there are many new application areas emerging for ML. The new ECMWF strategy⁹ for 2025 to 2034 envisages ECMWF continuing to innovate at the cutting edge of physical, computational and data science. It also foresees an emphasis on data-driven operational predictions but crucially with an anchoring on physics-based models. Hence the continued development of these models (e.g. improved scale-aware physical processes and coupling of Earth system components) will remain crucial.

Nevertheless, there will likely be a change in the emphasis for physically-based Earth system models, away from models that are primarily targeting operational NWP and forecast scores, towards tools that are as high resolution and physically realistic as possible. For example, tools to be used in the generation of nature-emulating Earth system training datasets for ML models and to be used within data assimilation. ECMWF has a record of sharing so-called *nature runs* with the community that provide detailed, consistent simulation data over a year or more at high resolution (e.g. Hoffmann et al., 2018; Agustí-Panareda et al., 2022). Moreover, the physical understanding of weather regimes and Earth system dynamics, climate projections and adaptation decisions, will likely remain grounded with physics-based Earth system model simulations, albeit combined with other available constraints such as observational records (through techniques including ML) (O’Reilly et al., 2024).

The wide availability of machine-learned tools that can emulate fluid dynamics or even replace selected discretisations for processes as complex as the Earth system is also opening new opportunities for modelling. Combining physics-based models with ML approaches can be less or more invasive, such as post-processing of model outputs, emulation of computationally expensive model components and learning of model error for online bias correction, all the way to emulating entire physical parametrization schemes, or correcting physical models by nudging to the evolving, large-scale dynamics inferred from machine-learned models (Husain et al., 2024).

One of the reasons why ML tools are so successful is the use of software tools such as Pytorch and JAX that make developments for users extremely easy and allow for interactive programming, higher-level abstraction, and rapid development cycles that

9 <https://www.ecmwf.int/sites/default/files/elibrary/2025/81641-ecmwf-strategy-2025-2034.pdf>

are possible when using e.g. Python and Jupyter notebooks, previously used only for visualisation workflows. This further motivates a change in the coding paradigm for physics-based modelling towards higher abstraction and easier-to-use and portable software.

Currently, porting a complex Earth system model to the latest HPC platforms is significantly more difficult for the IFS than for AIFS. To foster collaboration and dialogue with Member States in the area of emerging technologies, such as GPU adaptation, and for data production, access and distributed data proximate compute for NWP, the pilot project on ‘Adaptation to emerging technologies’ led by MeteoSwiss brings together expertise to blueprint solutions for accelerated and distributed workflows.

“ FOR THE FORESEEABLE FUTURE, IT IS UNLIKELY THAT MACHINE-LEARNED MODELS WILL ENTIRELY REPLACE CONVENTIONAL EARTH SYSTEM MODELS, BUT PHYSICAL MODELS NOW HAVE A SERIOUS COMPETITOR FOR OPERATIONAL NUMERICAL WEATHER PREDICTION AND THERE ARE MANY NEW APPLICATION AREAS EMERGING FOR MACHINE LEARNING.”

Moreover, the development of a portable model for multi-scale atmospheric prediction (PMAP) uses the domain-specific Python library GT4Py (GridTools for Python) with IFS (Ubbiali et al., 2025) and numerical modelling concepts of the finite-volume module (Kühnlein et al., 2019), applicable in both global and regional prediction. This has been a first step towards a separation of concerns within a modern coding environment for domain scientists. PMAP is performant but does not run in an operational context yet.

The challenges of data handling and governance, persistence and provenance of FAIR (findable, accessible, interoperable, and reusable) data are similar for both physical modelling and ML approaches. For example, serving ECMWF open data¹⁰ into a geographically distributed computational environment, irrespective of how it is produced on the HPC, is one of the development goals needed to satisfy an increasing demand for streaming weather and climate information. Such information is vital, not only for the immediate task of protecting society from extremes in a changing climate, but also to translate and blend with other distributed data spaces¹¹ in relevant impact sectors such as food, energy, agriculture and health.

CONCLUSION

The model development at ECMWF (also sometimes called the European model in the Americas) is a pertinent example of what can be achieved with a common goal. A huge amount of help and support from scientists all over the world has influenced the progress and success of the forecast model. In return, ECMWF has continued to share its knowledge of how to address today’s and future prediction challenges in a changing climate. Many prediction centres have adopted ideas or open source parts of the model initiated at ECMWF or its Member and Co-operating States.

We hope that this collaborative and successful path will continue long into the future, thus catalysing further shared development efforts in a flexible software infrastructure within an increasingly geographically distributed compute environment.

CONTRIBUTORS

ECMWF would like to thank the contributors to this paper:
Nils Wedi, Peter Dueben and Andy Brown.

10 <https://www.ecmwf.int/en/forecasts/datasets/open-data>

11 <https://digital-strategy.ec.europa.eu/en/policies/data-spaces>



Fifty years of data assimilation at ECMWF

ABSTRACT



Over the past 50 years, data assimilation (DA) has been a cornerstone of ECMWF’s success in numerical weather prediction (NWP), enabling significant advancements in forecast accuracy and extending prediction lead times. Through pioneering research and strong collaborations with its Member States, European meteorological services, and space agencies such as the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) and the European Space Agency (ESA), ECMWF has led the operational adoption of variational DA techniques, transitioning from early methods to the groundbreaking implementation of four-dimensional variational data assimilation (4D-Var). This transformation has allowed the direct assimilation of satellite radiances, unlocking the full potential of spaceborne observations and revolutionising modern data assimilation and forecasting.

Further developments, such as the introduction of the Ensemble of Data Assimilations (EDA), have provided a more robust representation of flow-dependent errors, improving uncertainty quantification in initial conditions. ECMWF continues to drive innovation through the evolution of coupled DA, integrating atmospheric, ocean, and land observations to enhance Earth system modelling. The “all-sky, all-surface” approach has further optimised satellite data assimilation in complex conditions, ensuring the best possible use of European and international investments in space programmes. These advancements are made possible through sustained collaboration with national meteorological services, research institutions, and operational programmes such as the European Union’s Copernicus programme.

Beyond weather forecasting, ECMWF’s world-class DA infrastructure underpins the production of high-impact climate reanalysis datasets, such as ERA5, which have become essential for climate monitoring and research. Looking ahead, artificial intelligence (AI) and machine learning (ML) are set to reshape the DA landscape, offering unprecedented opportunities to enhance observation processing, error correction, and computational efficiency. As ECMWF prepares for future observing systems and AI-driven forecasting, its commitment to scientific excellence, strong partnerships, and collaboration with its Member and Co-operating States ensures that it remains at the forefront of meteorology and climate monitoring science.

INTRODUCTION →

The European Centre for Medium-Range Weather Forecasts (ECMWF) stands as a global leader in numerical weather prediction (NWP), renowned for its pioneering advancements and outstanding forecasting capabilities. Since its establishment 50 years ago, ECMWF has been at the forefront of meteorological research, providing critical weather forecasts that inform decisions for its Member and Co-operating States across a multitude of sectors and applications. Numerical weather prediction, which involves the use of physics-based models to simulate the Earth’s atmosphere, has revolutionised our ability to predict weather, understand atmospheric dynamics and provide quantification of uncertainties about forecast products. At the heart of NWP lies data assimilation – a technique that seamlessly integrates observations from diverse sources into NWP models, ensuring that forecasts are optimally initialised to ensure best possible accuracy.

By incorporating observations from satellites, meteorological stations, buoys, aircraft and other platforms, data assimilation enables models to produce accurate initial conditions – a prerequisite for reliable weather forecasts. Throughout the last 50 years, ECMWF’s innovations in this domain have set the global standard, pioneering significant improvements in forecast skill and extending the lead time of high-confidence predictions. This cornerstone of NWP is an example of the Centre’s success in achieving scientific excellence and operational reliability (Rabier et al., 2000; Bauer et al., 2015).

The success of ECMWF, however, has not been achieved in isolation. Partnerships with national meteorological services, space agencies, and research institutions around the world have played a crucial role in establishing ECMWF as a world leader of weather forecasting. These collaborations have facilitated the sharing of expertise and ideas, resources, and observational data, creating a synergistic environment that has driven innovation. Through initiatives such as the European Union’s Copernicus Earth observation programme and Destination Earth (DestinE), and the World Meteorological Organization’s (WMO’s) collaborative frameworks, ECMWF has greatly benefited from the power of international cooperation to push the boundaries of what is possible in NWP.

In this paper, we describe the fundamental role of data assimilation in ECMWF’s success, its evolution alongside advancements in NWP, and the critical importance of partnerships in shaping ECMWF’s trajectory as a global leader in meteorology.

HISTORICAL OVERVIEW
OF DATA ASSIMILATION
AT ECMWF →

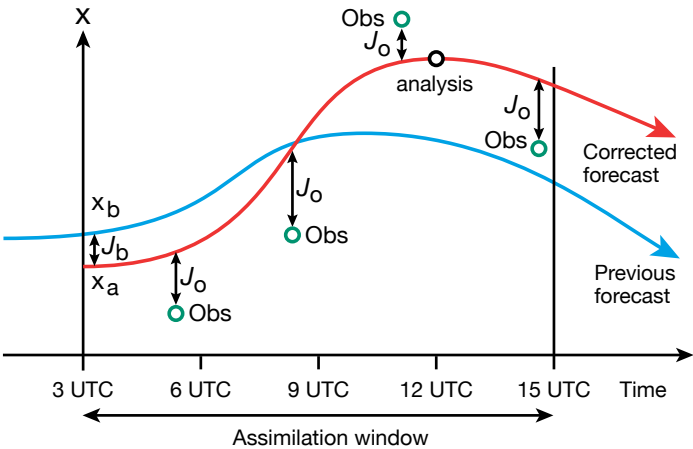
ECMWF was founded in 1975 with the primary mission to produce ten-day weather forecasts using state-of-the-art NWP systems. From the outset, the Centre recognised the importance of data assimilation as the foundation for reliable model initialisation. In its early years, ECMWF implemented a basic three-dimensional optimal interpolation (OI) scheme (Lorenc, 1981), which provided a systematic approach to incorporating observational data into its models. This method, based on statistical interpolation, balanced observational data with prior forecast information (background data), weighting both sources of information according to their relative errors and spatial correlations and minimising the expected error variance of the resulting initial state to improve forecast accuracy. While OI was a significant step forward in data assimilation during its time, it had several shortcomings that limited its effectiveness compared to more advanced techniques like variational data assimilation. Significant limitations of the OI algorithm implemented at the time include the use of static, predefined error covariance matrices, the suboptimal use of observations not linearly related to the analysis variables (e.g. satellite radiances), the lack of model constraints in the analysis procedure, and the local nature of the solver, which can lead to numerical artefacts (e.g. discontinuities) in the resulting analysis fields.

“ DURING THE 1990s, ECMWF PIONEERED THE USE OF VARIATIONAL DATA ASSIMILATION TECHNIQUES, SHIFTING FROM OPTIMAL INTERPOLATION WHICH HAD SIGNIFICANT LIMITATIONS.”

During the 1990s, ECMWF pioneered the use of variational data assimilation (Var) techniques, shifting from OI to more sophisticated approaches, in order to address the limitations described above. This effort was greatly facilitated by a proactive collaboration between ECMWF and Météo-France on what was called the Integrated Forecasting System (IFS)/ARPEGE project, which mobilised significant resources on both sides to address this new revolutionary (at the time) framework (Pailleux et al., 2014). Three-dimensional variational data assimilation (3D-Var) was operationally implemented at ECMWF in 1996 (following an earlier implementation at the US National Centers for Environmental Prediction (NCEP) in 1995). Worth noting is that prior to this implementation, and as described in Eyre et al. (2020), assimilating satellite observations as low-vertical-resolution retrieved profiles had at best a neutral impact in most NWP centres, exhibiting difficulties in specifying appropriate error statistics for the retrievals, contaminated by their climatological background.

An important intermediate step towards direct radiance assimilation was the assimilation of 1D-Var retrievals which used NWP short-range forecasts as background information. This removed large components of the climatological background from the retrieved profiles and was much closer to direct radiance assimilation than the assimilation of retrievals based on climatological information (Eyre et al., 2020). Indeed, the 1D-Var retrieval scheme (Eyre, 1989) used profiles from a short-range forecast as background, whereas other retrieval schemes used statistical background information (Reale et al., 1986). Even with very sophisticated techniques, it is unavoidable that errors in the selected background contribute to the retrieval error. The problem shows up as very systematic air-mass-dependent biases in the retrieved data (Andersson et al., 1991). The errors introduced by the retrieval process are characterised by horizontal correlations that vary with the meteorological conditions and are therefore difficult to accurately account for in the analysis. This problem is fully eliminated by incorporating the retrieval process within the analysis: a combined retrieval/analysis approach enables a more accurate combination of the information contained in the background, in the radiances and in the conventional data (Andersson et al., 1994). All data are analysed simultaneously

■ Figure 1: In the case illustrated here, for a single parameter x the observations are compared with a short-range forecast from a previous analysis over a 12-hour assimilation window. The model state x_b at the initial time is modified to achieve a statistically and dynamically based good compromise x_a by minimising a penalty function. The most important penalty terms are J_b , representing the fit to the previous forecast x_b , and J_o , representing the fit to all the observations within the assimilation window.



Another key improvement in 4D-Var is its ability to better control dynamic imbalances in the analysis. Since it uses the forecast model itself as a constraint, the final analysis is dynamically consistent, reducing unrealistic adjustments to temperature, wind, and pressure fields that could occur in 3D-Var. This leads to a smoother transition between the analysis and forecast phases, improving medium-range prediction accuracy.

Despite these advantages, the transition to 4D-Var also came with challenges. One major drawback was the computational cost. This necessitated advances in high-performance computing (HPC) to make 4D-Var operational. However, this was not enough, and incremental 4D-Var at ECMWF was introduced to address the computational challenges associated with full 4D-Var. Indeed, its original formulation was computationally expensive, requiring multiple integrations of the forecast model and its adjoint, involved in the process of minimising the distance between the model trajectory and the observations over the assimilation time window. To make 4D-Var operationally feasible, ECMWF adopted an incremental approach, first proposed by Courtier et al. (1994). This method allowed for a more efficient optimisation process by splitting the assimilation into multiple lower-resolution linear minimisation steps, known as outer and inner loops. Instead of solving the full nonlinear 4D-Var problem at once, incremental 4D-Var approximates it iteratively, first making a coarse-resolution estimate of how observations should be assimilated and then refining it through a series of linearised adjustments.

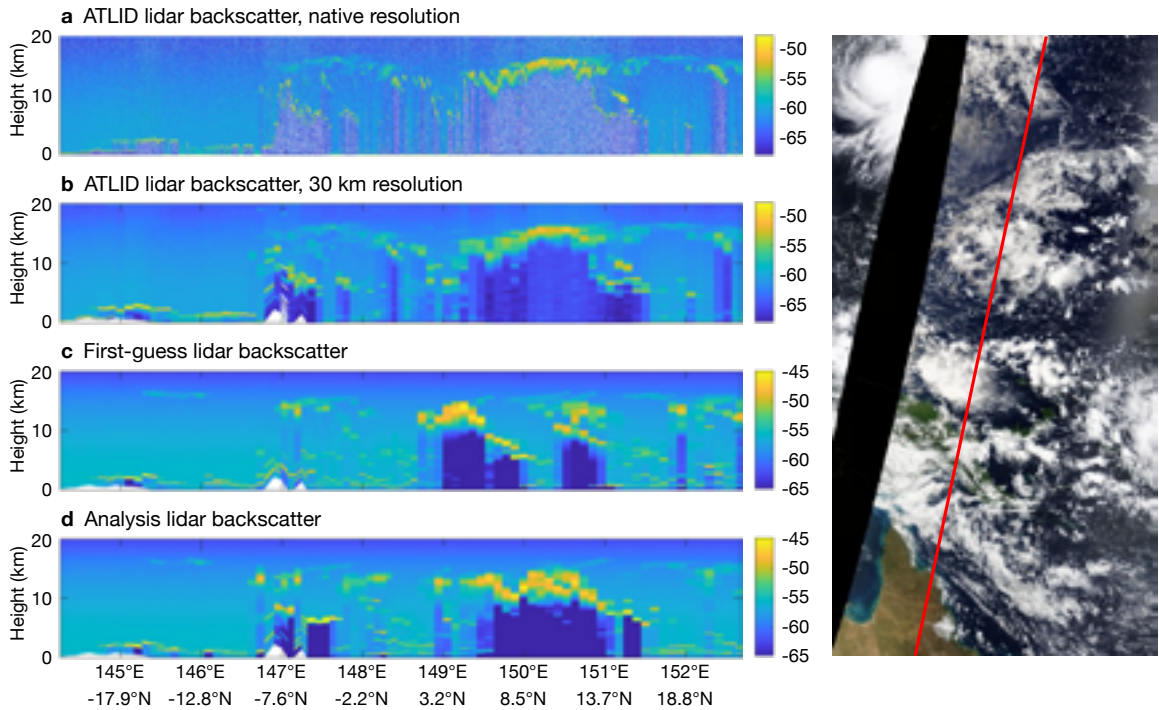
The implementation of incremental 4D-Var significantly reduced computational costs while maintaining the benefits of the full 4D-Var method. The outer loop operates at higher (finer) resolution, using the full nonlinear forecast model to update the control variables. The inner loop, where most of the optimisation occurs, uses a linearised (tangent-linear) version of the model at a reduced resolution to compute corrections more efficiently. This iterative refinement process ensures that the final analysis remains close to the optimal solution while avoiding the prohibitive expense of running a full-resolution nonlinear model at every iteration.

One of the primary advantages of incremental 4D-Var is its ability to make 4D-Var computationally affordable for operational use. Since the inner loop uses a reduced-resolution model, the overall cost is significantly lower compared to that of a full nonlinear 4D-Var system. Additionally, this approach improves numerical stability, as the assimilation increments remain small and are applied gradually, reducing the risk of introducing unrealistic changes to the atmospheric state. By approximating the analysis solution through successive iterations at increasing spatial resolution, incremental 4D-Var retains the ability to capture large-scale atmospheric corrections and, as the assimilation progresses, resolve smaller-scale features more effectively than a single direct minimisation. This makes it particularly useful for global numerical weather prediction at high resolution, which is an inherently multi-scale problem.

A final aspect of incremental 4D-Var that has allowed the algorithm to pass the test of time is its ability to deal efficiently with nonlinearities in the data assimilation system (Bonavita et al., 2018). As the model resolution increases and more observations are ingested that are nonlinearly related to the analysis variables, this capability of incremental 4D-Var has become increasingly important.

To deal with nonlinearities, an important development was related to incorporating increasingly sophisticated linearised physical parametrizations within the inner-loop minimisation process. In the standard formulation of 4D-Var, the inner loop uses a tangent-linear and adjoint model to propagate information about the state and its sensitivities. However, in early implementations, only the dynamical core of

■ **Figure 2:** Example for the assimilation of Atmospheric Lidar (ATLID) total backscatter on 3 August 2024. (a) ATLID total lidar backscatter at native resolution, averaged to the model grid, (b) ATLID total lidar backscatter at 30 km horizontal resolution, (c) first-guess total lidar backscatter, and (d) 4D-Var analysis total lidar backscatter. Backscatter is shown in units of $10 \log_{10} (\text{m}^{-1} \text{sr}^{-1})$. The red line in the satellite image shows the path of the satellite. From Fielding et al., 2025.



the forecast model was linearised, while physical processes such as radiation, convection, and boundary layer interactions were either ignored or represented in a very simplified manner (Mahfouf and Rabier, 2000). This limitation meant that some key atmospheric processes influencing cloud formation, precipitation, and turbulence were not properly accounted for in the assimilation, leading to suboptimal adjustments in the analysis.

To address this, ECMWF introduced linearised physics schemes within the tangent-linear and adjoint models, allowing physical processes to be considered during the minimisation of the 4D-Var cost function (e.g. Janisková et al., 2002). These schemes ensured that physical processes could be consistently represented within the assimilation cycle while maintaining computational efficiency. The introduction of these linearised physics schemes was particularly beneficial for the assimilation of cloud- and precipitation-affected satellite radiances, as well as for improving the representation of boundary layer and convection-related processes.

The development and refinement of these linearised physical parametrizations have continued as ECMWF has increased model resolution and improved satellite data assimilation. In later years, Janisková and Lopez (2013) expanded the use of linearised physics for variational cloud and precipitation assimilation. A recent achievement is the successful assimilation of lidar backscatter observations from the EarthCARE platform (see Figure 2, by Fielding et al., 2025), which would not have been possible without these continual developments.

Incremental 4D-Var remains a cornerstone of ECMWF's data assimilation system, continuously evolving to take advantage of new computational capabilities and improved observational data. By balancing accuracy and efficiency, it has enabled ECMWF to maintain high forecast skill while integrating an ever-growing number of satellite and in-situ observations. The method has proved to be a crucial advancement in numerical weather prediction, allowing for more reliable forecasts and better representation of atmospheric processes.

“ THE INCREMENTAL FORMULATION OF FOUR-DIMENSIONAL VARIATIONAL DATA ASSIMILATION (4D-VAR) HAS PROVED TO BE A CRUCIAL ADVANCEMENT IN NUMERICAL WEATHER PREDICTION, MAKING THE METHOD ACCURATE AND AFFORDABLE, AND ALLOWING FOR MORE RELIABLE FORECASTS AND BETTER REPRESENTATION OF ATMOSPHERIC PROCESSES. ”

Benefiting from this established infrastructure, the Copernicus Atmosphere Monitoring Service (CAMS; described in the ECMWF 50th anniversary paper on Copernicus), which ECMWF operates on behalf of the European Commission, is able to integrate vast amounts of satellite and in-situ observations into its atmospheric composition models. Using 4D-Var, CAMS produces high-quality global analyses of aerosols, greenhouse gases and reactive gases, improving air quality forecasts and environmental monitoring.

It is also worth noting that the 4D-Var framework has enabled the development of the Forecast Sensitivity to Observation Impact (FSOI) methodology to assess the impact of observations on forecast quality. FSOI measures how individual observations influence forecast error reduction. Using the adjoint model, FSOI quantifies the gradient of forecast error with respect to each observation, showing whether a given observation has improved or degraded the forecast. This technique enables convenient and inexpensive real-time assessment of the usefulness of different observing systems, helping optimise data assimilation strategies by prioritising observations that contribute most to forecast improvement. (Cardinali, 2009; Dahoui et al., 2017).

This tool and others are widely used as what we call Observing System Experiments (OSEs) to inform observation providers (e.g. space agencies) about the usefulness of various observing systems, and ECMWF has played a crucial role in shaping the Global Observing System (GOS) through various contributions using its DA infrastructure. These include targeted observation experiments (Buizza et al., 2007) and the Concordiasi project (Rabier et al., 2013).

Since its implementation in 1997, many changes have been made in the 4D-Var system, and some of the advances and challenges are described in the following section.

MAJOR DATA ASSIMILATION ENHANCEMENTS AT ECMWF →

ENSEMBLE OF DATA ASSIMILATIONS

The Ensemble of Data Assimilations (EDA) was introduced at ECMWF with two distinct but connected objectives. One was to provide improved initial conditions for the initialisation of the ECMWF Ensemble Prediction System (Buizza et al., 2008). The other was as a means to better estimate flow-dependent background error covariances within the variational data assimilation system. Before these EDA developments, the original implementation of incremental 4D-Var relied on static background error covariances, which were derived from climatological statistics. While these were carefully tuned, they did not evolve dynamically with the atmospheric flow. This limitation meant that background errors were often misrepresented, particularly in rapidly changing conditions such as during cyclogenesis, tropical cyclone development, or sudden stratospheric warmings (Bonavita et al., 2012).

Recognising the need for a more adaptive approach, ECMWF began developing the EDA in collaboration with Météo-France, which had been conducting pioneering work on ensemble-based estimation of background errors (e.g. Raynaud et al., 2008). Météo-France had explored the concept of using multiple realisations of the data assimilation cycle to diagnose errors dynamically, an approach that showed promise for improving the accuracy of background error covariance estimation. Inspired by these developments, ECMWF integrated the EDA into its operational 4D-Var system, creating an ensemble of perturbed data assimilation cycles to explicitly represent the uncertainties in the background state (Isaksen et al., 2010).

The introduction of the EDA marked a major advancement over the original incremental 4D-Var framework. In its traditional form, incremental 4D-Var minimised a cost function that included a background error covariance matrix (B-matrix), which had been computed from long-term statistics rather than evolving dynamically with the atmosphere. While this approach worked well in many cases, it struggled to correctly weigh observations in regions with high uncertainty, such as areas of active convection, frontal zones, or dynamically unstable regions. By using EDA-generated background errors, ECMWF was able to account for the flow dependency of forecast uncertainty, making the assimilation system much more responsive to the current state of the atmosphere (Bonavita et al., 2016).

The EDA works by running multiple independent 4D-Var analyses, each with stochastically perturbed observations and model states. These perturbations mimic the uncertainties in the observational data and model representation, creating an ensemble of analyses that reflects the possible range of atmospheric states. By computing the spread across the ensemble members, the EDA provides an adaptive estimate of background error covariances, which is then used in the main high-resolution 4D-Var assimilation. This allows the variational system to adjust its weighting of observations dynamically, giving more weight to observations in regions of high uncertainty and less weight where confidence in the background field is stronger (Isaksen et al., 2010).

Figure 3 shows the case of tropical cyclone Aere (north-eastern part of the Philippines on 8/9 May 2011) and is an illustration of how the errors diagnosed by the EDA, here for mean sea-level pressure, are, by design, constructed to estimate the real analysis errors, thus implicitly taking into account the observation network distribution and the model instabilities. In the present case, they act to extrapolate the observational information from the land-based stations into the more uncertain areas to the north-east of the cyclone, thus helping achieve a better positioning of the analysed storm.

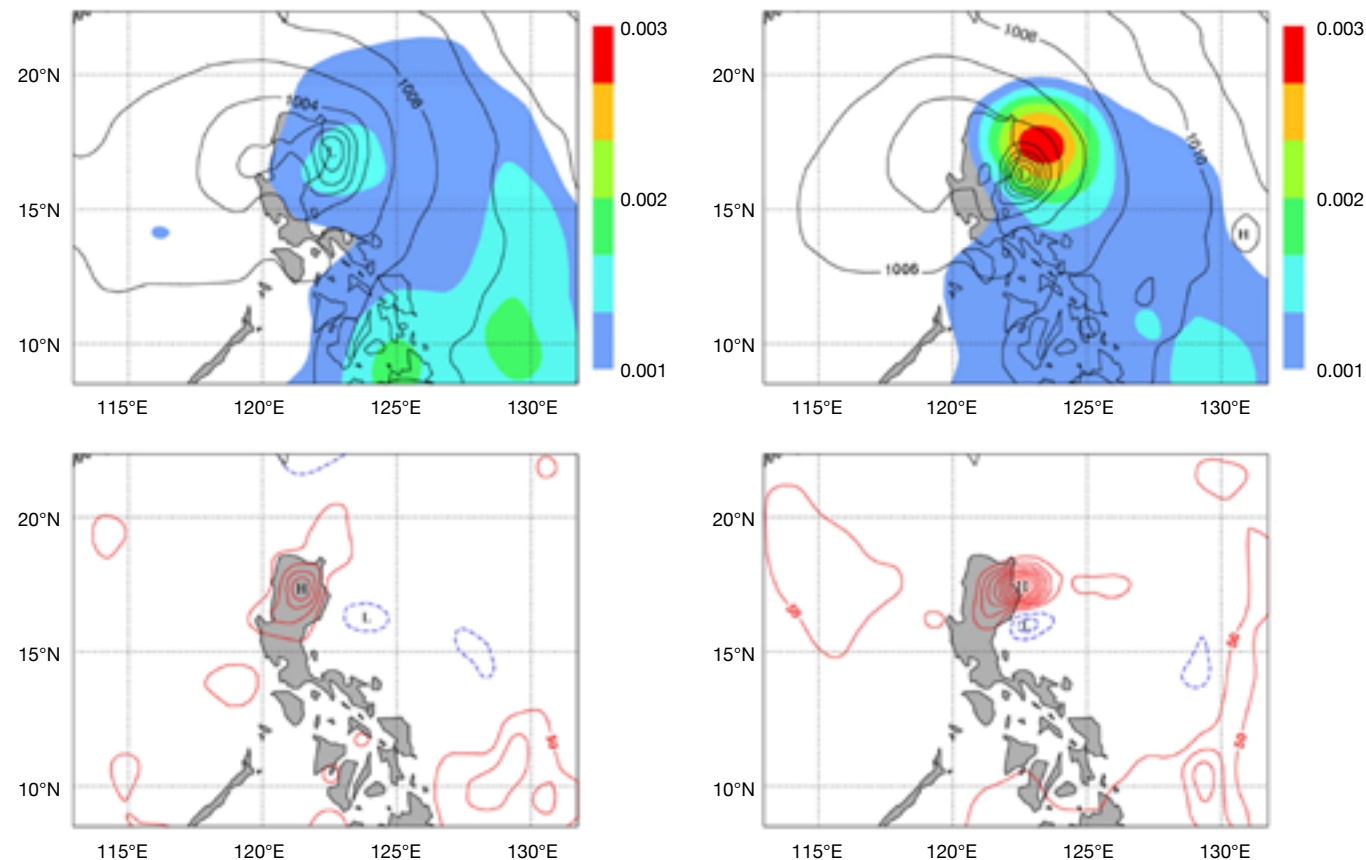


Figure 3: First line: Background mean sea-level pressure forecast valid on 9 May 2011 at 00 UTC (solid line, units: hPa) superimposed on background error estimates for the logarithm of surface pressure (shaded contours). Second line: Surface pressure analysis increments valid on 9 May 2011 at 00 UTC (solid lines indicate positive increments; dashed lines negative increments; isolines of 50 Pa). First column shows fields from the operational ECMWF analysis cycle at the time, with no EDA error estimate, second column from an experiment using EDA error estimates. From Bonavita et al., 2012.

A key benefit of incorporating the EDA into incremental 4D-Var was its impact on forecast sensitivity to observations. In a purely deterministic 4D-Var framework, the system assumes a fixed error distribution, which can lead to overconfidence in certain observations and underuse of others. With the EDA, the system continuously updates its understanding of error growth, leading to more accurate weighting of observational inputs (Bonavita et al., 2012). This proved particularly beneficial for satellite data assimilation, as it allowed ECMWF to assimilate more radiances dynamically, even in areas of high uncertainty, such as cloudy and precipitating regions (Geer et al., 2018).

Beyond its immediate impact on data assimilation, the EDA also played a crucial role in ensemble forecasting at ECMWF. By using EDA-based perturbations to initialise the Ensemble Prediction System (EPS), ECMWF was able to create more realistic ensemble spread, leading to better probabilistic forecasts. This dual application – improving both deterministic analysis and ensemble forecasting – solidified the EDA’s place as a cornerstone of ECMWF’s modern assimilation framework.

As computing power has increased, ECMWF has continued to refine the EDA, increasing the number of ensemble members and improving the perturbation methodologies. This has further strengthened the system’s ability to represent uncertainty and make full use of the ever-expanding volume of satellite and in-situ observations (Lang et al., 2019). However, in recent years, increasing compute power at historical rates has become more challenging, and questions have been asked about the long-term future of investing so many compute resources in the EDA. The main justifications for running the EDA are i) not having to maintain a

dedicated ensemble DA system separate from 4D-Var, ii) the 4D-Var is well tested, and iii) the skill of 4D-Var. The Object-Oriented Prediction System (OOPS, Bonavita et al., 2017) was developed to maintain multiple DA methods easily. It creates the possibility to run Ensemble Kalman Filter (EnKF) or Ensemble 4D-Var (EnVAR) for the EDA, but running a completely independent DA system for the deterministic analysis and the EDA is not desirable. At present, the 4D-Var algorithm has a higher skill level than any other algorithm tested in realistic NWP configurations. But it remains an open question if we can use other methods to replicate the contribution the EDA currently makes at much lower cost, without creating overheads in future support and testing. There is also a broader open question about the future of 4D-Var: whether it is computationally feasible at km-scale. At the time of writing, the answer to this is not clear, but OOPS definitely facilitates the implementation of alternative DA methodologies should these be needed.

“THE ENSEMBLE OF DATA ASSIMILATIONS HAS IMPROVED BOTH ANALYSIS ACCURACY AND ENSEMBLE FORECASTING – SOLIDIFYING ITS PLACE AS A CORNERSTONE OF ECMWF’S MODERN ASSIMILATION FRAMEWORK.”

WEAK-CONSTRAINT 4D-VAR

The implementation of weak-constraint 4D-Var at ECMWF was motivated by the need to address systematic model errors that limited the ability of 4D-Var to more effectively use various types of observations in the stratosphere and, more recently, at the surface. Traditional strong-constraint 4D-Var assumed that the numerical model used in data assimilation was perfect, neglecting the presence of conditional biases usually arising from deficiencies in model physics. However, systematic errors accumulated over time, particularly affecting stratospheric processes, boundary layer dynamics, and fast-evolving atmospheric phenomena. Weak-constraint 4D-Var allows the assimilation system to account for these errors dynamically, potentially leading to improved forecast accuracy and consistency (Trémolet, 2006).

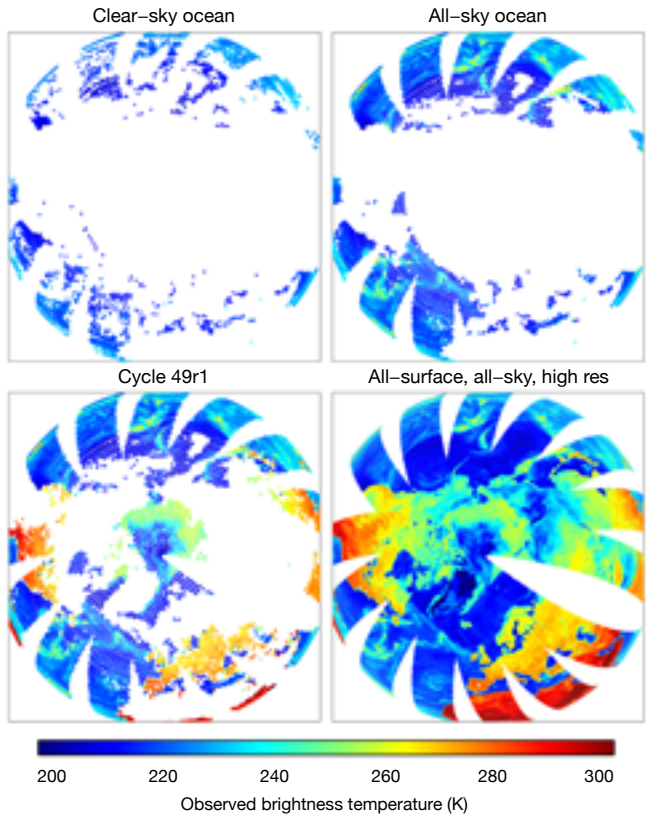
One major improvement resulting from weak-constraint 4D-Var was the reduction in stratospheric temperature biases. Before its implementation, systematic biases in the stratosphere led to persistent temperature drifts, impacting the representation of the jet stream, planetary waves, and stratospheric circulation. Weak-constraint 4D-Var corrected these errors, producing a more realistic depiction of upper-atmospheric dynamics (Laloyaux et al., 2020). More recently, the development of a version of weak-constraint 4D-Var able to estimate time-varying error structures during the assimilation window has allowed its extension to the boundary layer and the surface, with tangible improvements in the use of surface observations (two-metre temperature, surface pressure, scatterometer winds).

ALL-SKY, ALL-SURFACE SATELLITE DATA ASSIMILATION

A particular enhancement of the data assimilation system at ECMWF is related to continual efforts to improve the observation operators (mapping the model into observation space) and the characterisation of observation errors, especially for satellite observations. These developments addressing better surface emissivity models, better representation of microphysics of snow and graupel particles in the microwave, inclusion of observation error correlation, etc. have led to a massive increase in satellite observation usage, in cloudy and rainy conditions, as well as over land, snow and sea-ice surfaces.

An example is shown in Figure 4 from Geer (pers. comm.), representing the progressively increasing usage of microwave radiances (here Advanced Microwave Scanning Radiometer 2, AMSR2) in the DA system, including after the implementation of a new major cycle of the IFS (Cycle 49r1, implemented in

■ **Figure 4:** AMSR2 observed brightness temperatures in the 37 GHz v-polarised channel for the 12-hour DA window around 12 UTC, 17 November 2024, simulating the data coverage at earlier stages of DA development (clear-sky, all-sky, all-sky over sea ice/land after implementation of IFS Cycle 49r1). The bottom-right panel shows all data at the 40 km superobbing scale. In the other panels, the data has been thinned to 1 in every 8 superobs, giving effectively a 100 km spacing between observations. Data from multiple orbits has been allowed to overlap/superimpose.



TOWARDS A COUPLED DATA ASSIMILATION SCHEME

Historically, ECMWF’s data assimilation systems for the atmosphere, ocean, and land operated independently, with the atmospheric 4D-Var system focusing on upper-air data, while ocean and land components were initialised separately.

For the ocean, ECMWF’s ocean data assimilation began with the implementation of the NEMOVAR system, a variational data assimilation software developed collaboratively (CERFACS, ECMWF, Met Office, INRIA/Laboratoire Jean Kuntzmann) to integrate the NEMO ocean model. This system, operationalised in Ocean Analysis System 4 (Ocean-S4, implemented in 2011), used a multivariate three-dimensional variational (3D-Var) First Guess at Appropriate Time (FGAT) approach, assimilating temperature and salinity profiles alongside altimeter-derived sea level anomalies. Building upon this foundation, ECMWF introduced the Ocean ReAnalysis System 5 (ORAS5, introduced in 2017), which incorporated an ensemble generation technique to better represent uncertainties in ocean observations and model physics. ORAS5 provides improved initial conditions for coupled forecasts, thereby enhancing the skill of medium-range weather predictions and seasonal forecasts, the latter being used as an important component of the Copernicus Climate Change Service (C3S) offer. ORAS6 (to be implemented in 2025) further refines ocean reanalysis

capabilities. ORAS6 is based on an ocean ensemble-based variational data assimilation system, offering flow-dependent background error variances and vertical correlation scales.

For the land, initially, the assimilation scheme was a two-dimensional Optimal Interpolation (2D OI) method for analysing screen-level parameters and snow depth, while soil moisture and temperature analyses used a one-dimensional OI (1D OI) approach. This framework, though foundational, had limitations in capturing the complex interactions between land surface variables and atmospheric processes. ECMWF introduced a simplified Extended Kalman Filter (EKF) for soil moisture analysis. This advancement allowed for a more dynamic and responsive assimilation of soil moisture data, improving the representation of land–atmosphere feedbacks (de Rosnay et al., 2013). The EKF approach facilitated the integration of various observational data sources, including satellite-derived soil moisture measurements such as those from the Soil Moisture and Ocean Salinity (SMOS) mission and Advanced Scatterometer (ASCAT) data.

The transition towards a fully coupled DA system at ECMWF involves several methodological advancements. One approach is the development of outer-loop coupling, where the coupled model is introduced at the outer-loop level of the assimilation process. This method allows for the simultaneous adjustment of atmospheric and oceanic states, ensuring consistency across the coupled system. Additionally, efforts are being made to enhance the assimilation of surface-sensitive observations, such as sea-surface temperatures and soil moisture, which are critical for accurately capturing the interactions between different Earth system components (de Rosnay et al., 2022).

Data assimilation for operational NWP is a computationally intensive task that needs to be run daily within strict timeframes on available hardware. This set of requirements poses challenges for DA system developers. Currently, most operational DA systems are run in a hybrid configuration with a high-resolution control analysis based on a global variational solver (either adjoint-based, 4D-Var, or ensemble based, EnVar) and an ensemble DA component for error estimation and cycling (again, either adjoint-based, Ensemble of Data Assimilations (see previous section), or ensemble based, EnKF and its variants). This schematic description already makes it apparent that while DA is conceptually a probabilistic estimation problem, the dimension of the control space for global NWP at current spatial resolutions (O(109)) limits the choice of viable algorithms to those that assume Gaussian errors and only weak nonlinearities in both the observations and the model evolution during the assimilation window (Bonavita et al., 2018).

From a computational perspective, ensemble-based methods (EnVar, EnKF) tend to have better scaling properties than adjoint-based methods, as the analysis sensitivities to observations are directly sampled from the ensemble background forecasts and the solver can be parallelised efficiently. On the other hand, localisation is a known performance limiter for these systems, and the need to sample from the ensemble forecasts requires their storage with fast memory access, which can become impractical for increasing spatial/temporal resolutions and ensemble size.

The adjoint-based methods (4D-Var and its ensemble DA system, EDA) use their ensemble component for background error covariance estimation, but the error evolution in the assimilation window is achieved through running their linearised and adjoint models. This means that for variational methods the main computational

OPERATIONAL CHALLENGES

IMPACT OF ECMWF’S DATA ASSIMILATION ON WEATHER FORECASTING AND CLIMATE MONITORING →

As described in the previous sections, a driver for improved data assimilation at ECMWF has been the goal to make best use of the growing spaceborne observing system. This would not be possible without a very close partnership with space agencies, such as the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) and the European Space Agency (ESA).

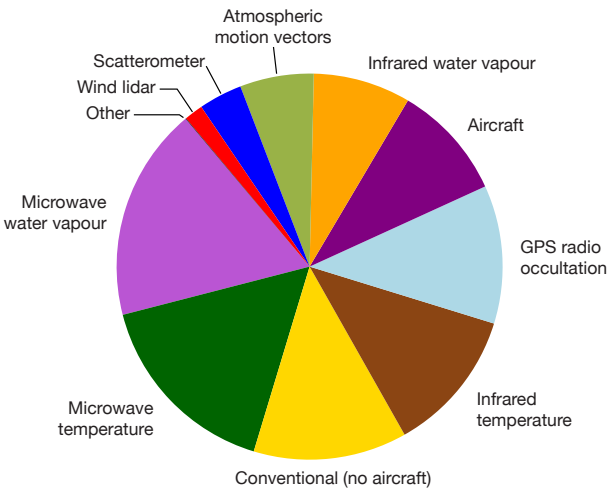
In particular, ECMWF and EUMETSAT have built a very efficient collaborative framework which is critical to enhance the use of satellite data in NWP and environmental monitoring. A key component of this partnership is ECMWF’s active participation in EUMETSAT’s Satellite Application Facilities (SAFs). These SAFs are specialised centres of excellence that focus on processing satellite data for specific applications, such as numerical weather prediction, climate monitoring, radio occultation and atmospheric composition, to name a few. By engaging with these facilities, ECMWF contributes its expertise in NWP to improve the processing and assimilation of satellite observations, thereby enhancing the accuracy of weather forecasts and climate analyses. Another significant aspect of this collaboration is the EUMETSAT Research Fellowship Programme, which places early-career scientists at institutions like ECMWF to develop innovative applications of satellite data. These seconded Fellows work on projects aimed at advancing the assimilation of satellite observations into ECMWF’s forecasting models.

“ A DRIVER FOR IMPROVED DATA ASSIMILATION AT ECMWF HAS BEEN THE GOAL TO MAKE BEST USE OF THE GROWING SPACEBORNE OBSERVING SYSTEM – MADE POSSIBLE THROUGH A VERY CLOSE PARTNERSHIP WITH SPACE AGENCIES, SUCH AS EUMETSAT AND ESA.”

constraint comes from the requirement to run the forecast model and its linearised and adjoint versions efficiently and quickly at ever-increasing resolutions. This problem is compounded by the fact that solvers used in variational DA are intrinsically sequential and there is little scope for domain parallelisation. Ten years ago, this state of affairs led people to question the long-term viability of 4D-Var. However, new ideas have changed the picture in the last few years. One of these is continuous DA (Lean et al., 2019). Continuous DA is based on the incremental implementation of 4D-Var and the concept of letting fresh observations into the assimilation system while 4D-Var is running. In practice, this reduces the time-critical portion of 4D-Var to the duration of the last minimisation update instead of the duration of the whole algorithm (which currently runs with four minimisations). This concept will be further developed in the extending-window DA framework, where the length of the assimilation window itself will vary as a function of observation cutoff time, thus ensuring a more continuous update of the analysis and thus even better ability to describe and forecast fast-evolving weather events.

Another important aspect is that of computational efficiency. In the ECMWF DA system, the EDA is the most computationally demanding component, and efforts have been focused on reducing its cost while maintaining or even improving performance. A recent example of these developments is the soft-centred EDA concept (Hólm et al., 2022). This implementation of the EDA differs from the original one as the perturbed members are simplified, lower-resolution 4D-Var updates and the mean background forecast is re-centred on the unperturbed member background. In addition, the minimisations in the perturbed members start from an initial control vector and preconditioning that is inherited from the output of the first minimisation of the unperturbed member. The resulting EDA is approximately 30% cheaper to run and its performance is superior to that of the original version.

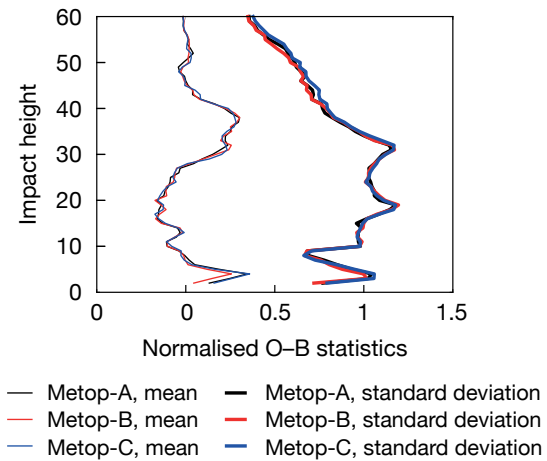
■ **Figure 5:** Relative impact of observing systems on the quality of the operational 24-hour forecast, estimated using their Forecast Sensitivity to Observation Impact (FSOI), and aggregated over the calendar years 2020 to 2024. The impact of microwave and infrared radiance sensors is separated by channel based on primary sensitivity to temperature or water vapour. Ground-based observations are separated into conventional (no aircraft) and aircraft. (Geer, pers. comm.)



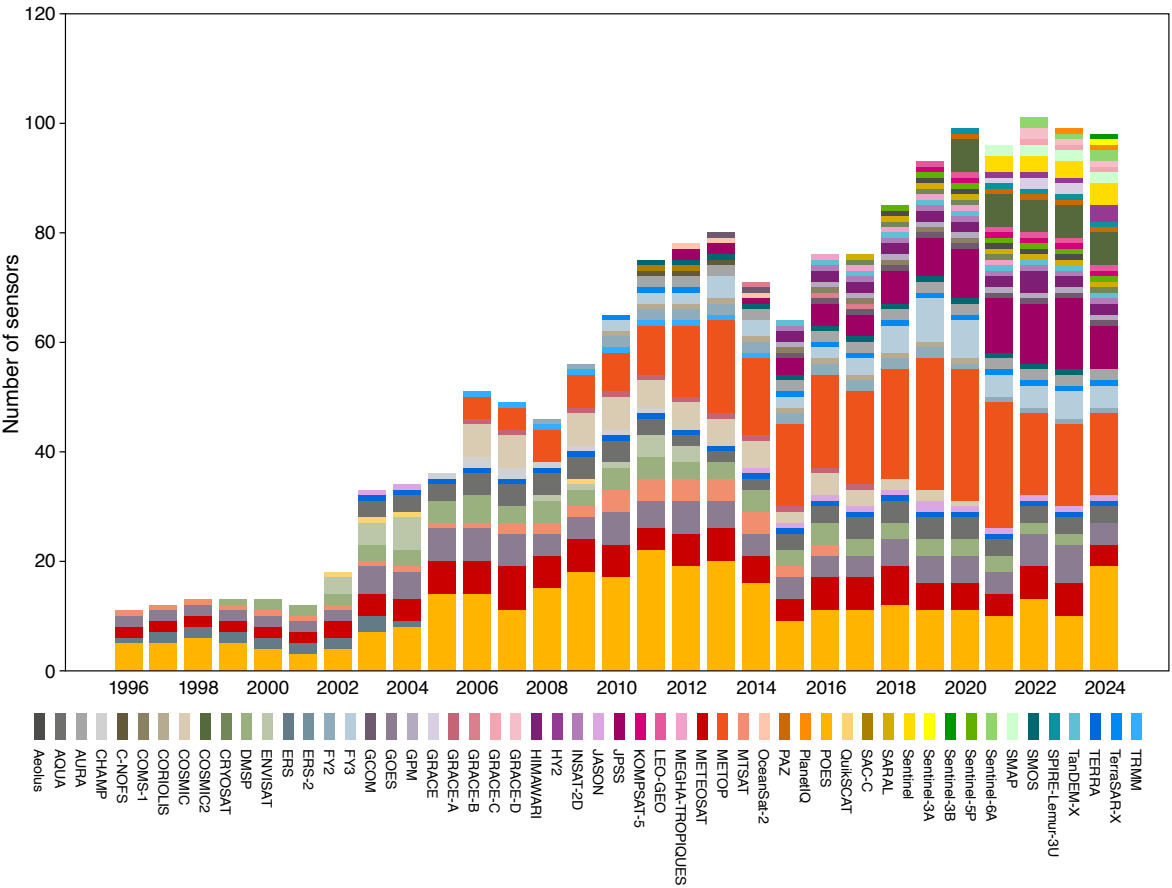
For example, past Fellows have focused on improving the use of atmospheric motion vectors, radiances from geostationary satellites, and microwave radiance data from polar-orbiting satellites – all activities dedicated to maximising the impact of these observations in the ECMWF NWP suite. Furthermore, ECMWF and EUMETSAT jointly regularly conduct various flavours of OSEs (as mentioned in the previous section) to assess and optimise the impact of various satellite data on NWP (see Figure 5).

The insights gained from OSEs inform decisions on future satellite mission designs and data assimilation strategies, ensuring that ECMWF’s models effectively exploit available satellite data and prepare for future missions (Healy et al., 2022). An advantage of this close cooperation with EUMETSAT is also the speed at which provision of feedback on data quality and evaluation of impact of data on the ECMWF system can be done. For example, EUMETSAT’s polar-orbiting MetOp-C satellite was launched on 7 November 2018, and the EUMETSAT radio occultation (RO) team produced high-quality bending angle profiles by 13 November 2018, within only six days of launch, and made them available to the EUMETSAT Radio Occultation Meteorology Satellite Application Facility (ROM SAF) for evaluation. Within days, ECMWF was able to provide quality assessment of these new data in comparison with both the Metop-A and B measurements, by comparing them with NWP information mapped to observation space (see Figure 6, from Healy et al., 2019).

■ **Figure 6:** Observation-minus-background departure statistics. The bending angle observation-minus-background (O-B) departure statistics (standard deviation and mean) as a function of impact height for the three Metop satellites. The departures are normalised by dividing them by the bending angle noise values used when assimilating the data. The statistics are computed for the period 27 November to 2 December 2018.



■ **Figure 7: Increase in satellite sensors monitored at ECMWF from 1996 to 2024.**



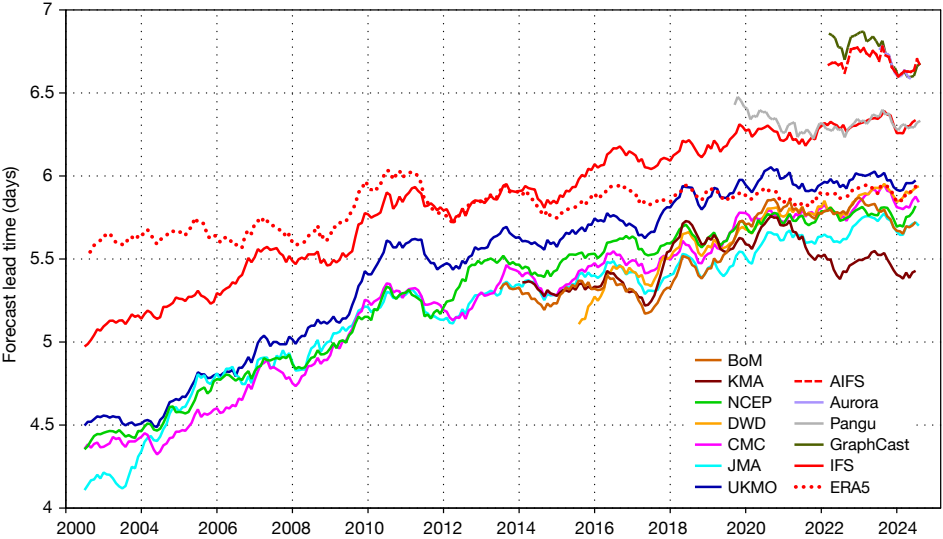
IMPROVED FORECAST ACCURACY

As a result, ECMWF has been a world leader at monitoring and assimilating satellite observations. Figure 7 shows how the data assimilation and model developments over nearly 30 years have enabled the number and diversity of satellite data instruments used to be massively increased.

Forecast skill improvements over the last 45 years have been achieved primarily through improvements to the forecast model, the quality and number of observations and the accuracy of the data assimilation method (Magnusson and Källen, 2013). It is challenging to attribute the contribution of each of these elements, but it is common to compare long-term trends in the performance of forecasts from the reanalysis system with trends in the forecasts from the operational system (see Figure 8). As a first approximation, we can say the reanalysis system shows improvements arising from changes to the observation system, and the trend in the operational system shows improvements from all components, so the difference in trends is an approximation of the combined contribution of model and data assimilation methodology changes.

The lead time at which the anomaly correlation of the 500 hPa geopotential height fell below 85% was 5 days in 2002 and 6.3 days in 2022, so a gain of 0.65 days per decade in this period, a drop from the 1 day per decade improvement reported by Magnusson and Källen (2013). The equivalent change for ERA5 was an increase from 5.5 days in 2002 to 5.9 days in 2022, so an increase of 0.2 days per decade. Therefore, in this 20-year period, we can say, approximately, that a gain of 0.2 days per decade arose from improvements in the Global Observing System, and a gain of 0.45 days per decade arose from improvements in the model and data assimilation. In this context, it is also worth noting ECMWF’s current Artificial Intelligence

■ **Figure 8: Forecast skill changes of various models, including ECMWF’s IFS, AIFS and ERA5. The figure shows the lead time at which the anomaly correlation of 500 hPa geopotential height over the northern hemisphere extratropics falls below 85%.**



“ FORECAST SKILL IMPROVEMENTS OVER THE LAST 45 YEARS HAVE BEEN ACHIEVED PRIMARILY THROUGH IMPROVEMENTS TO THE FORECAST MODEL, THE QUALITY AND NUMBER OF OBSERVATIONS AND ENHANCEMENTS OF THE DATA ASSIMILATION METHOD.”

Forecasting System (AIFS; see ECMWF 50th anniversary paper on machine learning) configuration gains around 0.35 days over the best physics-based models in 2023–24, only marginally less than the model and data assimilation improvements for the last decade.

In considering these changes, the rapid changes in forecast skill of the IFS in 2005–2007, 2015–2017 and 2018–2020 with respect to ERA5 stand out. The main contributor to the forecast skill gain for the latter change was the introduction of continuous DA, which allowed for the ingestion of observations which arrived after the first minimisation in subsequent minimisations in the outer-loop 4D-Var configuration. Therefore, this gain can be attributed mainly to a change in DA methodology, though there were a number of other changes in this period. In 2015–2017, the changes were a mix of model, most notably increased horizontal resolution, and DA changes (and observation changes, but these would also impact ERA5, whose skill also rose during this period). For the older period, it is difficult now to attribute with high confidence, but a major change in background error formulation (Fisher, 2005) was introduced in 2005 and may have contributed to the large improvement seen in this period. Going further back, the transition to variational assimilation and direct radiance assimilation resulted in the largest changes to operational forecast scores at the end of the 1990s (see Figure 3 in the ECMWF 50th anniversary paper on Earth system modelling).

HIGH-IMPACT AND EXTREME WEATHER

In addition to monitoring the impact of data assimilation developments and improved observations on global scores, there have also been attempts to measure progress for high-impact and extreme weather. This is harder to study objectively, because by definition extreme events are rare and, therefore, it is challenging to test changes in a statistically robust way. The assimilation of satellite observations has repeatedly been shown to play a critical role in the accurate forecasting of individual severe weather cases, most notably that of Hurricane Sandy in October 2012 (McNally et al., 2014). Tropical cyclones (TCs) have been studied, most recently by Magnusson et al. (2025). They concluded that near TCs, observations are important

for forecasts mainly up to one day ahead, with the dropsondes particularly helpful to reduce central pressure errors. However, at longer lead times, it is the microwave satellite radiances that are critical to the TC position, and also central pressure up to two days ahead. It was also shown that the development of all-sky microwave assimilation (see section on ‘Major data assimilation enhancements’ above and Geer et al., 2018) is increasing the impact of microwave radiances further, demonstrating that it is not just the observations, but the maturity of the data assimilation method which is important, especially in areas with persistent cloud cover such as TCs. Scatterometer observations were also shown to be of value, with increasing impact as data thinning is reduced. However, other observation types were not shown to have a strong impact on TC forecasts.

ECMWF also engages with partners to examine the impact of observations on forecasts of atmospheric river (AR) events (Lavers et al., 2024). These studies have examined the impact of targeted observations on forecasts of AR events, particularly through the Atmospheric River Reconnaissance (AR Recon) programme, which involves ECMWF and its Member States. In particular, they explore the value of field campaign dropsonde datasets, in the AR Recon seasons 2022/23 and 2023/24. These show where the dropsondes have value, which can be up to four days’ lead time. ECMWF also played a pivotal role in supporting other field campaign experiments, particularly through its involvement in the THORPEX (The Observing System Research and Predictability Experiment) programme. ECMWF’s contributions included providing targeted model runs and assimilating observations from these campaigns to enhance weather prediction accuracy in polar regions. ECMWF participated in the Concordiasi project, with data from Concordiasi being assimilated into ECMWF models, improving weather forecasts and reanalysis efforts in polar regions as well as evaluation of satellite data over difficult surfaces, particularly from the IASI on the MetOp-A satellite (Rabier et al., 2013).

Both ERA-Interim and ERA5 reanalysis datasets, produced by ECMWF, rely on the 4D-Var system to integrate large volumes of observational data into a consistent, long-term dataset. ERA-Interim (Dee et al., 2011), covering the years 1979 to 2019, was based on an earlier version of 4D-Var with a 12-hour assimilation window and a coarser spatial resolution of approximately 79 km. In contrast, ERA5, the production of which was funded under the Copernicus programme, and covering from 1950 to the present, benefits from a more advanced weak-constraint 4D-Var, a higher resolution of approximately 31 km, and hourly output, providing a more detailed and accurate representation of atmospheric, land, and oceanic conditions (Hersbach et al., 2020). ERA5 also assimilates a broader range of satellite observations, including hyperspectral infrared and microwave radiances, with improved bias correction and error representation techniques. These enhancements result in a better depiction of stratospheric processes, and long-term climate trends for screen-level parameters (Simmons et al., 2021). Through the combination of state-of-the-art data assimilation and continuous improvements in observational data usage, 4D-Var in ERA5 continues to enhance the accuracy and reliability of climate reanalysis products, supporting a wide range of scientific, policy and business applications, generating a wide user base, as described in the ECMWF 50th anniversary paper on Copernicus. ERA5 is also crucial for initialising AI-based weather forecasting systems. It provides high-resolution, historical hourly atmospheric data used to train and initialise AI models, including ECMWF’s AIFS. The AIFS leverages Graph Neural Networks (GNNs) trained on ERA5 and operational analyses to learn atmospheric patterns and improve predictions. By using ERA5 as initial conditions, AI models generate accurate forecasts of surface weather and extreme events which compete with forecasts from traditional models.

FUTURE DIRECTIONS AND PROSPECTS →

Last but not least, the CAMS reanalysis (EAC4) also benefits from ECMWF’s advanced data assimilation infrastructure by integrating a vast array of satellite and in-situ observations into a consistent 20-year-long global dataset. This system ensures high-quality atmospheric composition reanalysis, improving accuracy in pollutants, greenhouse gases, and aerosols. Here also, the 4D-Var technique refines temporal consistency. This reanalysis is used for computing climatologies, studying trends, evaluating models, benchmarking other reanalyses, and most importantly, serving as boundary conditions for regional models covering past periods. These applications support policy-making and environmental monitoring efforts.

The development of ECMWF’s data assimilation (DA) system will continue to be driven by the need for accurate initial conditions in Earth system modelling and optimal use of present and future observations to improve forecasts and climate data records. Over the next decade, the DA system will also support the training and initialisation of ECMWF’s Artificial Intelligence Forecasting System (AIFS) and national forecasting efforts via the Anemoi initiative (Dramsche et al., 2024), consolidating ECMWF’s collaborative efforts with its Member and Co-operating States on this critical issue. The focus will expand beyond initial conditions, using the DA system and observations to directly enhance forecast performance.

An important game changer in the next decade will be that AI and ML applications in DA will continue to rapidly evolve. ECMWF is already integrating ML to correct systematic model errors dynamically, extending beyond the capabilities of weak-constraint 4D-Var. Initial studies (Bonavita and Laloyaux, 2020) showed ML-based corrections can improve forecasts significantly. Recent results (Farchi et al., 2025) confirm forecasts based on this hybrid approach can match state-of-the-art data-driven models while retaining physical realism.

“ AN IMPORTANT GAME CHANGER IN THE NEXT DECADE WILL BE THE INCREASING ROLE OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN RESHAPING THE DATA ASSIMILATION LANDSCAPE.”

Machine learning also enhances hybrid observation modelling. As demonstrated by Geer (2024a and 2024b), ML supplements physics-based modelling of complex satellite observations, including radiative properties, hydrometeors, and surface interactions. Additional applications include monitoring observation systems (Dahoui, 2023) and developing latent spaces for variational DA (Melinc and Zaplotnik, 2024). These developments confirm ML’s growing role in enhancing analysis accuracy and forecast skill within mathematically robust DA methodologies.

Within this likely revolution, it remains certainly true that the 4D-Var assimilation system (in a broad sense, and with all its peripheral components) will remain central to NWP and atmospheric composition. Efforts will therefore continue to improve observation and background error covariances while pushing computational resolution limits, leveraging experience from DestinE (Sandu, 2024). Extending assimilation windows will optimise performance and workflow efficiency, with potential benefits for time-critical boundary conditions in regional modelling.

The forthcoming transition to a hybrid HPC system with central processing units (CPUs) and graphical processing units (GPUs) will require code adaptation. Possible avenues could be the enhancement of tangent linear and adjoint calculations using machine-learned emulators to reduce computational costs in 4D-Var. The Ensemble of Data Assimilations, another cost-intensive system, will also benefit from these efficiency gains.

A key priority is improving consistency and efficiency across Earth system components, particularly through interface observations from satellites. The methodology successfully applied to ocean and sea ice observations will be expanded to land surfaces. Beyond initial conditions, DA will be used to learn meteorology-dependent errors for machine-learning-based corrections in medium-range forecasts, optimising model parametrizations, and training data-driven neural network models of the atmosphere.

Copernicus Services (CAMS and C3S) will continue to benefit from DA developments, and efforts to estimate emissions and surface fluxes of greenhouse gases and pollutants will intensify, leveraging satellite data such as Sentinel-5P for CH₄ and NO₂. In the context of the Paris Agreement and the monitoring of CO₂ emissions from space, the operational implementation is being prepared for the Copernicus CO₂ Monitoring (CO2M) mission, set for launch in 2027. The climate (ERA6) and atmospheric composition (EAC5) reanalyses will enter full production within the next few years, integrating scientific advancements with automated quality monitoring, while discussions on future reanalysis activities will begin, emphasising data-driven forecasting applications.

Maximising observational data usage remains a top priority and requires a proactive collaborative approach. Existing satellite observations will be assimilated in more challenging environments, such as complex land/sea ice surfaces and cloudy regions, in collaboration with EUMETSAT, ESA and Member and Co-operating States. Assimilation spatial and temporal resolution will increase, leveraging DestinE experience, and new methodologies will be developed to estimate spatial error correlations crucial for 4D-Var and EDA. The coupled DA framework will extract additional insights from observations at the interfaces between Earth system components.

Infrastructure enhancements will ensure the rapid adoption of new satellite observing systems. Early in the next decade, ECMWF aims for operational use of data from the Meteosat Third Generation Imaging (MTG-I) satellite Flexible Combined Imager (FCI) and Lightning Imager (LI) and from EarthCARE, working closely with EUMETSAT and ESA. The Centre will continue supporting EUMETSAT's mission advisory groups for MTG-S and EPS-SG, while expanding ESA collaborations through the DANTEX initiative (Bormann et al., 2025). Efficient integration of newly launched continuity satellites from the US, China, and Japan and evolving in-situ networks will maintain forecasting system performance.

ECMWF will continue engaging with private sector observation providers, particularly in radio occultation data, while working with EUMETSAT, ESA and the US National Oceanic and Atmospheric Administration (NOAA) to validate and acquire these datasets. Future Observing System Experiments and EDA impact assessments will guide network planning, including optimised conventional observation networks optimised as a result of the Systematic Observations Financing Facility (SOFF) initiative of WMO. This exemplifies the increasingly critical role of partnerships in the data assimilation strategy at ECMWF.



CONCLUSION

ECMWF has established itself as a global leader in data assimilation, consistently pioneering methodologies that have significantly improved numerical weather prediction. The transition from optimal interpolation to 4D-Var has been instrumental in enhancing forecast accuracy and extending lead times. The implementation of the Ensemble of Data Assimilations (EDA) has further strengthened uncertainty representation, refining initial conditions for both deterministic and ensemble forecasts. Advances in satellite data assimilation, particularly the integration of all-sky and all-surface observations, have maximised the use of spaceborne data, improving forecasts for extreme weather events. Continuous developments in weak-constraint 4D-Var have addressed systematic model errors, yielding more reliable analyses, especially in the stratosphere and at the surface. The coupled data assimilation framework is another milestone, promising enhanced Earth system modelling through the simultaneous assimilation of atmospheric, oceanic, and land observations.

Crucially, ECMWF's success is underpinned by the contributions of its Member and Co-operating States and strong partnerships with national meteorological services, space agencies, and research institutions. Collaborations with EUMETSAT, ESA and other agencies have ensured optimal use of satellite observations, while joint initiatives such as Copernicus as well as with the WMO have expanded the impact of ECMWF's advancements.

Looking ahead, ECMWF is at the forefront of integrating AI into data assimilation, exploring ML-based corrections to model biases and advanced observation handling. ECMWF is even pioneering radical research into producing forecasts directly from observations (Alexe et al., 2024 and McNally et al., 2024), essentially incorporating the DA step in a fully end-to-end AI-based forecasting system (called AI-DOP). The next decade will see increasing reliance on hybrid CPU-GPU architectures to optimise computational efficiency, ensuring that advanced DA techniques remain viable at higher resolutions. ECMWF's expertise will continue to shape future reanalysis products such as ERA6, reinforcing its role in climate monitoring and forecasting.

The Centre's commitment to international collaboration, particularly through its Member and Co-operating States and strategic partnerships, remains essential for optimising global observing networks. Additionally, ongoing research into continuous data assimilation and extended-window DA will further refine forecast initialisation, particularly for fast-evolving weather systems. With the impending launch of next-generation satellites and increased observational capabilities, ECMWF is well positioned to harness new data sources for even greater forecast improvements. As numerical weather prediction enters the AI era, ECMWF's data assimilation strategy ensures that both traditional physics-based models and emerging AI-driven approaches benefit from the most accurate initial conditions. By maintaining its focus on accuracy, efficiency, and scientific rigour, ECMWF is well positioned to define the next chapter in data assimilation and Earth system prediction, working hand in hand with its partners to push the boundaries of meteorological science.

CONTRIBUTORS

ECMWF would like to thank the contributors to this paper:
Jean-Noël Thépaut, Massimo Bonavita, Niels Bormann,
Matthew Chantry, Stephen English, Alan Geer and Tony McNally.



Ten years of Copernicus at ECMWF

ABSTRACT

The Copernicus Programme, the European Union’s flagship Earth observation initiative, provides free and open access to high-quality environmental data, supporting decision-making in climate monitoring, air quality, ocean and land monitoring as well as emergency management, for the benefit of public and private economic sectors.

Since its establishment in 2014, part of the Copernicus Programme has been implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) in partnership with European and international institutions, with ECMWF playing a leading role in the Copernicus Atmosphere Monitoring Service (CAMS) and the Copernicus Climate Change Service (C3S). These services deliver essential operational products covering air quality, climate reanalysis, seasonal forecasts, greenhouse gas monitoring, and solar radiation. CAMS and C3S serve thousands to tens of thousands of direct users worldwide (respectively) and reach millions of indirect users. The services support policymakers, businesses, and researchers with authoritative climate and environmental information, including the widely used ERA5¹, ECA4² and EGG4³ reanalyses.

ECMWF is also a key contributor to the Copernicus Emergency Management Service (CEMS), under the leadership of the Joint Research Centre (JRC), supporting global flood and fire forecasting. ECMWF is the Computational Centre of the European Flood Awareness System (EFAS) and Global Flood Awareness System (GloFAS), delivering flood forecasts up to 30 days in advance and a hydrological outlook several months ahead, helping national and international agencies anticipate and mitigate flood risks. In parallel, ECMWF is also responsible for producing daily fire danger assessments. These assessments, which include high-resolution and ensemble forecasts extending up to 15 days, are integrated into the European Forest Fire Information

System (EFFIS) and the Global Wildfire Information System (GWIS) to support national authorities in managing forest fires across Europe and neighbouring regions as well as worldwide. This complements nicely the Global Fire Assimilation System (GFAS) of CAMS, which integrates satellite-based fire radiative power observations to estimate global biomass burning emissions, feeding into atmospheric composition models to assess the impact of wildfires on air quality and climate.

With Copernicus Phase 2 (2021–2028), ECMWF is expanding its contributions, including the Copernicus Anthropogenic CO₂ Emissions Monitoring and Verification Support Capacity (CO2MVS), which will provide independent observation-based assessments of global greenhouse gas emissions. These innovative services will build upon the space component of the Copernicus Programme including the launch of a new generation of satellites between 2025 and 2028 (in particular geostationary Sentinel 4, Sentinel 5, the constellation CO₂M and the Copernicus Contributing missions). New initiatives also include the first global coupled Earth system reanalysis and National Collaboration Programmes tailored to EU Member States’ needs to expand and support Copernicus services and data uptake.

Through its expertise in numerical modelling, data assimilation, and Earth system science, ECMWF continues to ensure that Copernicus services deliver world-class, actionable environmental intelligence for Europe and beyond.

1 Fifth generation ECMWF atmospheric reanalysis
2 Fourth generation ECMWF global reanalysis of atmospheric composition
3 Fourth generation ECMWF global reanalysis of greenhouse gases



INTRODUCTION →

The EU’s Copernicus Programme is a comprehensive Earth observation initiative designed to provide valuable data for monitoring the Earth’s environment using space and in-situ observations. In particular, the Copernicus Sentinel missions are a fleet of Earth observation satellites developed by the European Space Agency (ESA) to support the Copernicus Programme, providing high-resolution, continuous, and freely available environmental data. Each Sentinel satellite series is designed for specific monitoring purposes: Sentinel-1 provides all-weather radar imaging for land and ocean surveillance, Sentinel-2 captures high-resolution optical imagery for land monitoring, Sentinel-3 measures sea and land surface temperatures, ocean colour, and altimetry, Sentinel-4 and Sentinel-5 focus on atmospheric composition monitoring, while Sentinel-6 ensures precise sea level measurements. These satellites, complemented by a wide variety of in-situ observations, work in synergy to provide critical data for climate change monitoring, disaster response, land and ocean management, and air quality assessments, forming the backbone of Europe’s environmental and security services. Harnessed to these observations, Copernicus delivers operational datasets and information services in a range of topical areas (see Figure 1). From these baseline services many other value-added products can be tailored to more specific public, policy and commercial needs.

Figure 1: Conceptual description of the Copernicus Programme. Reproduced from <https://climate.copernicus.eu/about-us>.



The Programme includes six thematic core services: atmosphere monitoring, marine environment monitoring, land monitoring, climate change, emergency management, security. The Copernicus Atmosphere Monitoring Service (CAMS) (Peuch et al., 2022) and Copernicus Climate Change Service (C3S) (Thépaut et al., 2018 and Buontempo et al., 2022) serve as central pillars for monitoring atmospheric composition and climate change, and have been managed by ECMWF since 2014. These services enable governments, industries, researchers, and the public to access critical information for research, decision-making, and policy development. Indeed, CAMS and C3S contribute significantly to global efforts in understanding environmental and climate issues. Their data and forecasting capabilities support a broad range of applications, from air quality monitoring to long-term climate projections, enabling informed decisions across sectors. CAMS provides real-time atmospheric composition analyses and forecasts as well as reanalyses, while C3S delivers authoritative climate data to assess past, present, and future climate

conditions. Both services have been instrumental in shaping European and international policies on air quality, climate mitigation, and adaptation strategies.

As part of ECMWF's broader mission, Copernicus services leverage cutting-edge numerical modelling and data assimilation techniques to provide high-quality and accurate environmental information. ECMWF's expertise in weather forecasting, combined with collaborations across institutions in Europe, ensures that CAMS and C3S products remain at the forefront of Earth system monitoring. This collaboration not only strengthens scientific research and innovation but also enhances operational efficiency, making environmental data more accessible and useful for policymakers and businesses.

CAMS and C3S are deeply embedded in Europe's scientific and operational infrastructure, drawing on expertise from ECMWF and its Member and Co-operating States, which contribute to roughly 70% of the activities. This synergy fosters a collaborative approach, ensuring that data quality, accessibility, and usability continue to evolve in response to emerging challenges.

Through continuous advancements in data assimilation, model improvements, and user engagement, Copernicus services at ECMWF are well positioned to remain a key driver of environmental monitoring and climate resilience in Europe and beyond.

BACKGROUND
AND CONTEXT →

As part of the Copernicus Programme, CAMS and C3S were designed to build upon ECMWF's operational infrastructure and numerical modelling expertise, providing decision-makers, researchers, and businesses with essential environmental intelligence. CAMS focuses on forecasting and monitoring air pollutants, greenhouse gases, and their emissions, while also delivering air quality assessments with products tailored to policy users' needs and solar radiation products. C3S offers authoritative data about past, present and future climate through observations, reanalyses, seasonal forecasts, climate predictions and projections, contributing to both scientific research and evidence-based policymaking across Europe and globally.

CAMS and C3S stand on the shoulders of a series of pioneering research and development projects, such as Global Earth-system Monitoring (GEMS), Monitoring Atmospheric Composition and Climate (MACC) and MACC-II (Hollingsworth et al., 2008; Flemming et al., 2015), and the ERA reanalysis projects such as ERA-20C (Poli et al., 2016). These precursor projects laid much of the foundation for the operational implementation of CAMS and C3S, transforming state-of-the-art research in numerical weather prediction, data assimilation, and Earth system modelling into fully operational and globally recognised services. The investment in these services has been sustained, deliberate, and consistent, ensuring that scientific breakthroughs are not confined to academic circles but are swiftly integrated into operational frameworks that benefit public authorities, industry, and citizens alike. From the outset, Copernicus was conceived with a user-driven approach, engaging stakeholders at every stage to ensure its products and services address real-world needs. This commitment to co-development with meteorological services, space and environmental agencies, research institutions, and private-sector actors has resulted in a dynamic ecosystem where feedback loops continuously refine and enhance service offerings. The participatory nature of Copernicus ensures that its data and tools remain relevant, actionable, and aligned with evolving policy and economic demands.

A defining characteristic of Copernicus is its full, free, and open data policy, which has been instrumental in driving innovation and fostering a new wave of public and private sector applications. By lowering barriers to data access, Copernicus empowers businesses, researchers, and policymakers to develop cutting-edge solutions in areas such as urban planning, agriculture, disaster response, energy management, and climate risk assessment. This democratisation of data amplifies the impact of Copernicus, ensuring that its scientific and technological advancements translate into tangible economic and societal benefits.

Beyond its direct operational value, Copernicus plays a strategic role in positioning the EU as a global soft power in environmental monitoring and climate action. As a provider of authoritative and independent Earth observation data, Copernicus enhances Europe's leadership in global climate governance, supporting international frameworks such as the United Nations Framework Convention on Climate Change (UNFCCC), the Paris Agreement⁴ and its Global Stocktake, and the Global Methane Pledge (GMP)⁵. It also supports the implementation of EU regulations in the air quality, climate and energy fields. Its data and services are used worldwide, reinforcing the EU's ability to shape international environmental policy, support humanitarian efforts, and foster global scientific collaboration.

Through continuous investment, strong research foundations, and an unwavering commitment to operational excellence, Copernicus has evolved from an ambitious vision into a cornerstone of global environmental intelligence. As it progresses into its second phase, it continues to set the benchmark for integrated, policy-relevant, and user-driven Earth observation services, ensuring that Europe remains at the forefront of tackling climate and environmental challenges on a global scale.

METHODOLOGY →

The methodologies employed in CAMS and C3S are built on advanced data collection, assimilation, and modelling techniques. These methodologies ensure that datasets are accurate, comprehensive, and accessible to a broad range of users. CAMS and C3S integrate vast amounts of observational data from satellites, ground-based instruments, and reanalysis products to generate state-of-the-art climate and atmospheric monitoring systems. ECMWF, as the implementing body for both services, plays a critical role in advancing these methodologies, leveraging decades of expertise in numerical weather prediction (NWP) and operational forecasting.

“ ECMWF’S EXPERTISE IN NUMERICAL WEATHER PREDICTION AND OPERATIONAL FORECASTING PROVIDES AN ESSENTIAL FOUNDATION FOR THE COPERNICUS SERVICES THAT IT IMPLEMENTS.”

A cornerstone of both services is data assimilation (see ECMWF 50th anniversary paper on data assimilation), which combines observational data with numerical models to generate reliable and consistent datasets. ECMWF's Integrated Forecasting System (IFS) (see ECMWF 50th anniversary paper on Earth system modelling) is central to this approach, assimilating over 100 satellite data streams for both services and a wide range of climate observations for C3S. The assimilation process, particularly through four-dimensional variational data assimilation (4D-Var), ensures that data used in atmospheric and climate models align with the principles of atmospheric physics and past observations. This system, initially developed for NWP, has seen major advancements since the late 1990s, leading to a step change in the accuracy of atmospheric and climate monitoring.

4 <https://unfccc.int/process-and-meetings/the-paris-agreement>
5 <https://www.globalmethanepledge.org/>

CAMS builds on this capability to generate real-time and reanalysis products for global atmospheric composition. The system produces global analyses and five-day forecasts twice a day, covering greenhouse gases (GHG), reactive gases, and aerosols.

CAMS also incorporates a sophisticated fire emissions monitoring system, the Global Fire Assimilation System (GFAS), which processes satellite data on active fires to estimate emissions from wildfires and biomass burning (in particular carbon and particulate matter emissions). GFAS emission datasets are a key component of the input datasets (emissions inventories, meteorological fields, in-situ measurements and remote sensing observations) that feed the global and regional modelling systems to create high-resolution air quality and atmospheric composition forecasts and analyses. The combination of these datasets allows CAMS to deliver comprehensive and reliable information about atmospheric composition patterns, which are used by policymakers, researchers, and businesses for air quality management, climate impact assessments, and design of emissions mitigation strategies.

C3S focuses on providing reliable, long-term climate data through observational climate records, reanalyses, seasonal forecasts, and climate projections, with its Climate Data Store (CDS) serving as a cloud-based platform that provides seamless access to the vast range of these climate datasets. Designed to be highly accessible, the CDS offers a programming interface (e.g. application programming interfaces (APIs) and Jupyter notebooks) that enables users to analyse data online without requiring extensive computational resources, facilitating data processing, visualisation, and analysis for various sectors such as agriculture, energy, insurance, and water resource management. Over time, the CDS has evolved significantly, enhancing user capabilities to develop custom applications, filter data by region and date, and present information in various formats.

Building on this success, CAMS launched the Atmosphere Data Store (ADS) in June 2020, extending the concept to atmospheric composition data and providing a unified platform for accessing the various global and regional datasets. This evolution reflects a concerted and synergetic effort to streamline data accessibility and usability across Copernicus services, ensuring that scientists, businesses, and policymakers can efficiently leverage climate and atmospheric data for research, decision-making, and environmental policy development (see section on synergies between the two services).

Both CAMS and C3S employ rigorous quality control and evaluation measures to ensure the reliability and scientific integrity of their data products. These quality assurance processes have been at the heart of both services and have progressively been developed in synergy (see section on synergies between CAMS and C3S).

A significant feature of both services is their emphasis on user engagement and application development. CAMS and C3S have actively collaborated with stakeholders, including national meteorological agencies, environmental organisations, and industry partners, to ensure that their data products meet real-world needs. The services offer tailored applications and sector-specific indicators that support decision-making in climate adaptation, risk assessment, and environmental monitoring. For example, C3S produces the European State of the Climate report annually (jointly with WMO since 2023), providing a comprehensive overview of climate trends and anomalies across Europe. Similarly, CAMS delivers policy-relevant assessments, such as the annual European air quality reports that describe air quality status in Europe (since 2013) and its evolution with respect to the implementation of EU regulations, considering health and environmental impact indicators.

MAIN ACHIEVEMENTS →

CAMS MAIN ACHIEVEMENTS

6 A configurable extension of the IFS used for modelling atmospheric composition.

7 <https://eur-lex.europa.eu/eli/dir/2024/2881/oj/eng>.

CAMS and C3S have made significant advancements in delivering high-quality atmospheric and climate information, addressing a wide range of user needs. Their services provide open and free access to data, supporting scientific research, policy development, and operational decision-making across multiple sectors. CAMS has about 40,000 registered users, while C3S has 390,000.

CAMS has established itself as a global leader in air quality and atmospheric composition monitoring. Its data products cover global and European air quality forecasts, analyses and reanalyses, greenhouse gas (GHG) and air pollutant emission inventories, solar irradiance assessments, and climate forcings. These products support a broad range of applications, from real-time monitoring to information for the general public and policy-driven needs.

CAMS integrated global and regional modelling systems offer a unique consistent framework that takes stock of numerous in-situ and satellite observations to provide best estimates of air pollutants and greenhouse gas concentrations in Europe and worldwide.

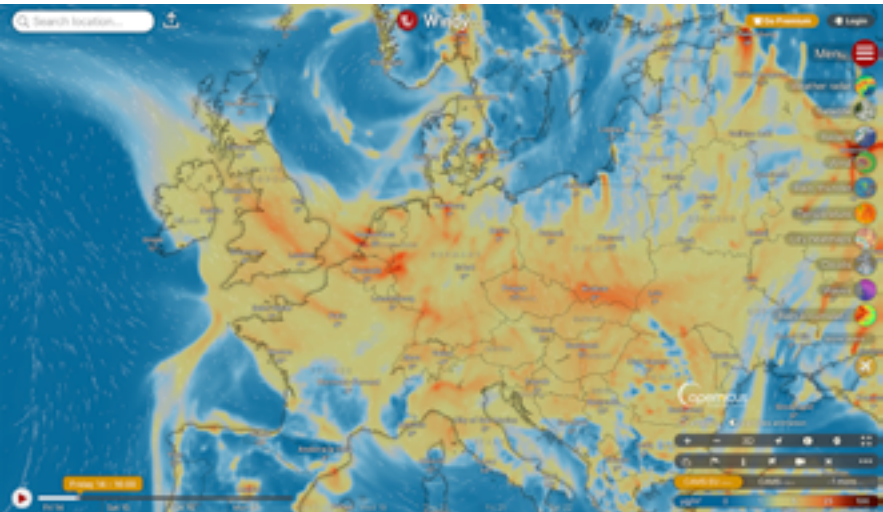
For example, CAMS provides global aerosol information by assimilating satellite observations into the ECMWF IFS-COMPO® model to deliver twice daily forecasts of the distribution of various aerosol types, including desert dust, sea salt, organic matter, black carbon, sulphate, and nitrate, offering insights into phenomena like dust storms, biomass burning, and pollution events. The service’s real-time fire emissions monitoring system, GFAS, based on assimilation of satellite-based fire radiative power data, brings additional key information. Therefore, the CAMS global forecasting system can track long-range air pollution events triggered by natural (wildfires, dust storms, volcanoes) and anthropogenic sources anywhere in the world. This information can be used by national authorities to understand transboundary patterns and to run national systems (as boundary conditions or input data for machine learning downscaling algorithms).

CAMS global analyses and forecasts also provide boundary conditions for implementing the CAMS air quality services over Europe. Managed by Météo-France, these services are built upon 11 European air quality models run in 10 countries with a 10 km resolution, that are combined to derive products that are more robust and are of overall better quality than the individual systems.

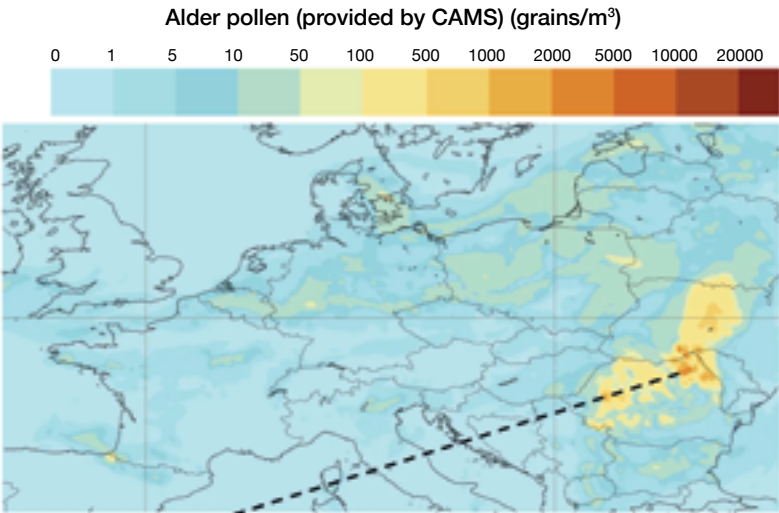
Policy tools are also available that aim to quantify the relative contribution of various sources to air pollution episodes, in forecast mode, to help policymakers assess the efficacy of short-term mitigation measures. This panel of tools is deemed sufficiently relevant and consistent with policy needs to be quoted as a valuable source of information in the revised ambient Air Quality Directive adopted by the European Council in 2024⁷. Predictions of when the weight of natural dust in particulate matter exceeds regulatory limits is one of the most popular applications.

Regarding CAMS’ role in raising public awareness of air pollution, a good example of CAMS data uptake is its integration into the Windy.com platform, a popular weather application. Users can access CAMS forecasts at global and European scales of surface concentrations of nitrogen dioxide, fine particulate matter, and aerosol optical depth (AOD). This integration enables users to track the movement of aerosol plumes from various sources, including dust, biomass burning, air pollutants, sea sprays, and volcanic eruptions. Figure 2 shows an example of a CAMS European forecast of surface nitrogen dioxide (NO₂) visualised on the Windy.com web-based application.

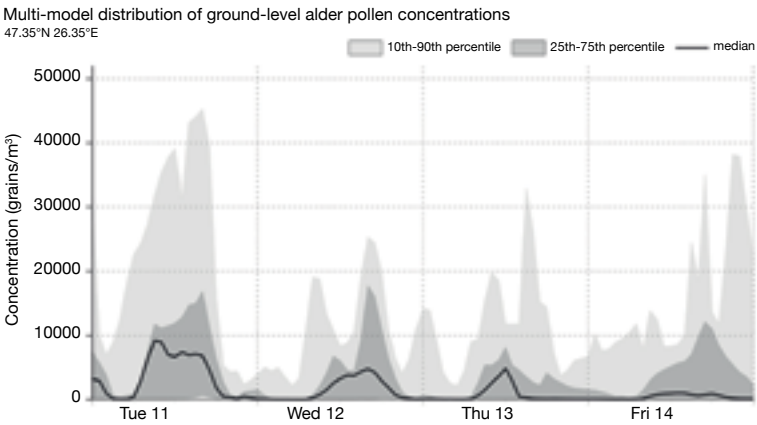
■ **Figure 2:** Sample CAMS European forecast of surface nitrogen dioxide (NO₂) visualised on Windy.com. Four-day forecasts of 8 different CAMS variables (from the global system and, when zooming in over Europe, from the regional ensemble system) can be explored: NO₂, ozone (O₃) and carbon monoxide (CO) surface concentrations; PM2.5 (fine particles with a diameter smaller than 2.5 micrometres) and desert dust particles surface mass concentrations; and vertically-integrated column of aerosol (aerosol optical depth), sulphur dioxide (SO₂) and ozone. Source: Windy.com.



To conclude on the regional near-real-time CAMS services, it should be noted that CAMS also deals with pollen. Six pollen species (alder, birch, grass, mugwort, olive, and ragweed) are monitored and forecasts of their airborne concentrations are produced to help users anticipate exposure issues (see Figure 3).

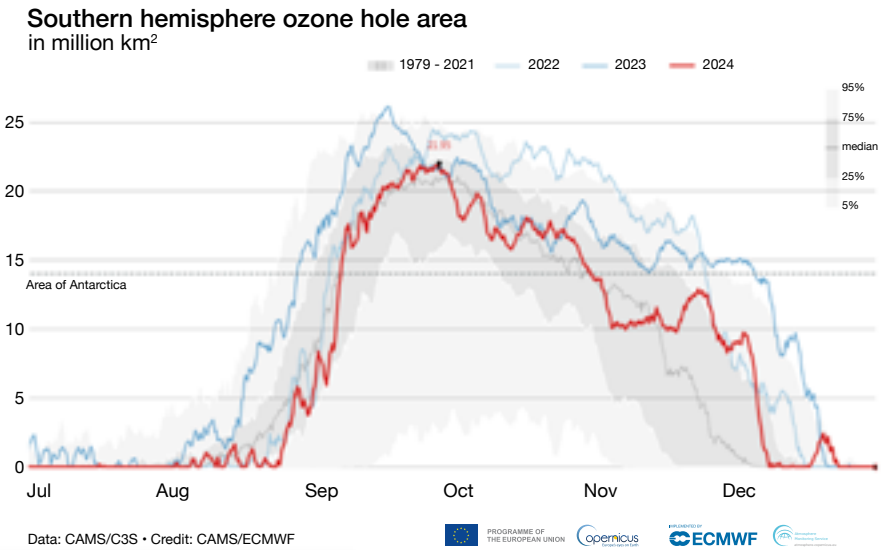


■ **Figure 3:** European-scale alder tree pollen forecast (from 11 March 2025). Alder pollen peaks in the first months of the year, affecting tens of millions in the European population. With appropriate medical support and advice, this information can help allergy sufferers manage their symptoms. This forecast is provided by a large consortium led by Météo-France and Ineris, with specific scientific support from the Finnish Meteorological Institute. Eleven individual regional models are used to consolidate the forecasts and assess uncertainty. See CAMS daily pollen forecasts at: https://atmosphere.copernicus.eu/charts/packages/cams_air_quality/products/europe-air-quality-forecast-pollens



■ **Figure 4:** Evolution of the southern hemisphere ozone hole area in 2024 compared to 2022 and 2023. The shaded areas show the 1979 to 2021 envelope.

Another notable achievement of CAMS is the routine provision of relevant information about the evolution of stratospheric ozone concentrations and the ozone hole. Every year on 16 September, International Day for the preservation of the ozone layer, we launch the Antarctic ozone hole monitoring season and publish regular information about its evolution compared to the previous years (see Figure 4). With these kinds of products, CAMS contributes to major international assessments, such as the Montreal Protocol for the ozone layer, the Intergovernmental Panel on Climate Change (IPCC), the United Nations Framework on Climate Change, and the Global Methane Pledge. One of the most recent achievements to support climate and methane regulation is the new methane hotspot explorer service, developed in collaboration with the Netherlands Institute for Space Research (SRON), to monitor and characterise methane plumes detected by the TROPOMI instrument embedded on the Sentinel-5P satellite.



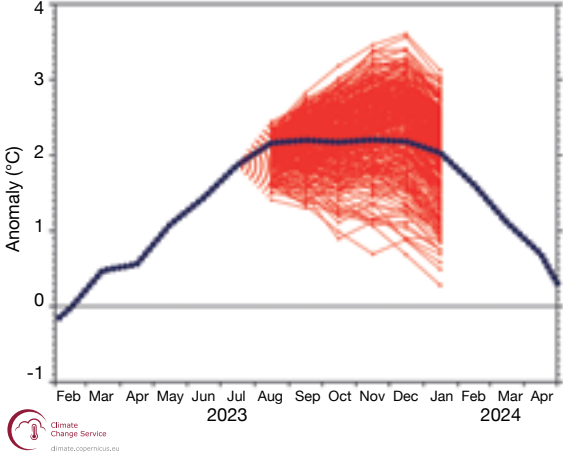
Beyond real-time applications, CAMS is delivering two global reanalyses: one for aerosols and reactive gases (EAC4) (Inness et al., 2019) and another for greenhouse gases (EGG4) (Agustí-Panareda et al., 2023). These datasets provide long-term atmospheric composition records that are widely used in atmospheric sciences, aiding both researchers and policymakers in assessing air pollution trends and climate impacts. By integrating data from over 100 satellite sensors and combining them with numerical models, these reanalyses ensure consistency in historical datasets, supporting research on air quality, emissions regulation, and climate change mitigation. Regional air quality reanalyses over Europe are also available since 2013.

CAMS also provides information on ultraviolet radiation, supporting health-related applications, and offers detailed assessments of climate radiative forcing, quantifying the impact of different atmospheric components on global warming.

A key area where CAMS has gained official recognition is in European policies. It is explicitly mentioned in the EU's Methane Strategy and the Air Quality Directive as a tool that supports Member States in their monitoring and reporting duties. This reflects the service's increasing role in regulatory compliance and environmental governance. The service's ability to provide tailored data for policymakers has led to the development of dedicated policy-support tools, including assessments of emission mitigation strategies and air pollution attribution reports.

C3S MAIN ACHIEVEMENTS

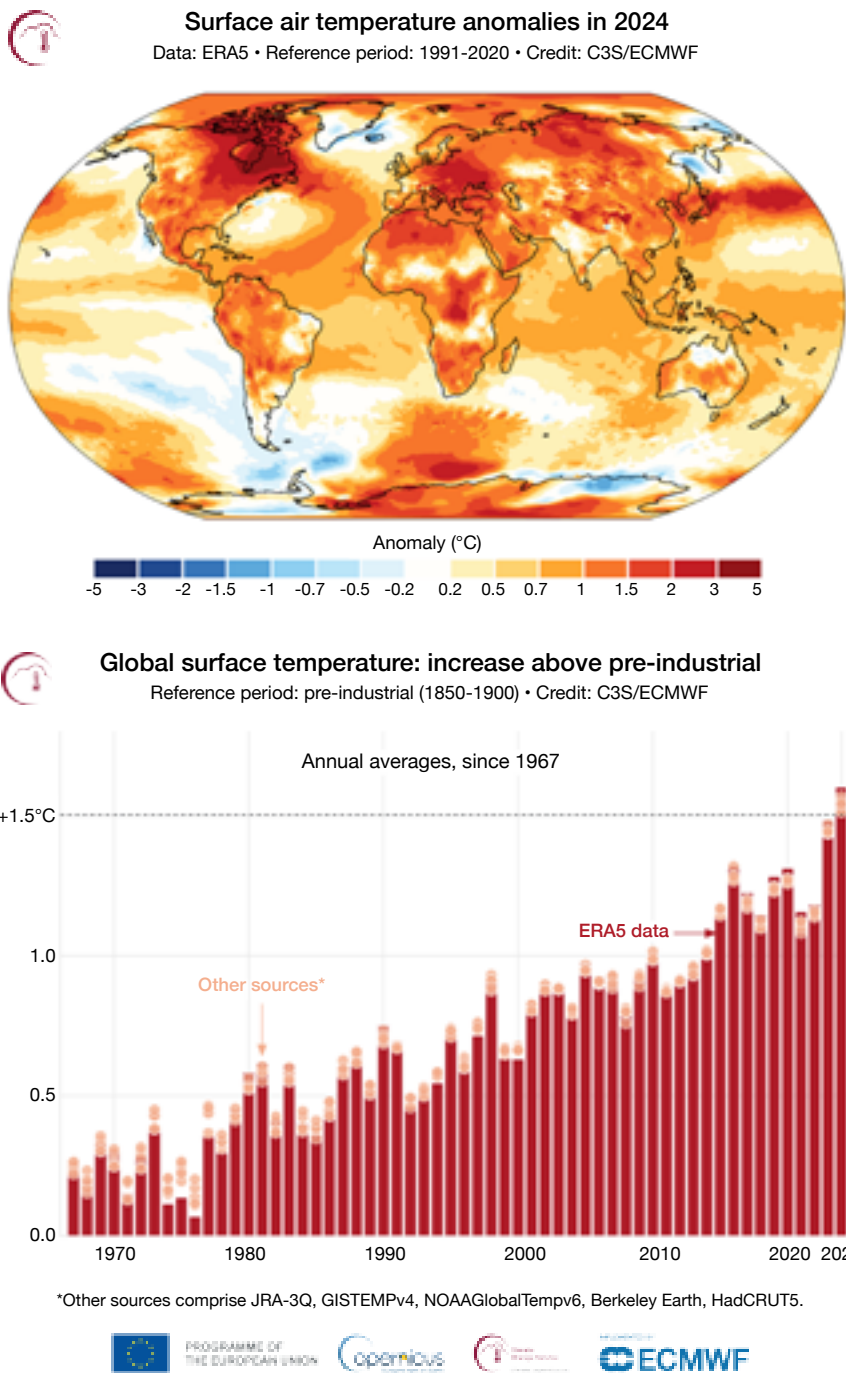
■ **Figure 5:** Niño-3 sea-surface temperature (SST) anomaly forecast from 1 August 2023 based on the C3S multi-system seasonal forecast (from ECMWF, UK Met Office, Météo-France, Euro-Mediterranean Center on Climate Change (CMCC), Deutscher Wetterdienst (DWD), US National Centers for Environmental Protection (NCEP), Japan Meteorological Agency (JMA), Environment and Climate Change Canada (ECCC)). The forecast captures well the peak of El Niño towards the end of 2023. The Niño-3 SST anomaly index is an indicator of eastern tropical Pacific El Niño conditions. It is calculated with SSTs in the box 150°W - 90°W, 5°S - 5°N.



One of C3S’s flagship initiatives is the European State of the Climate report, which provides an annual overview of climate trends and extreme weather events across Europe. Figure 6 shows two example illustrations of the wealth of information about the global and European climate available in these reports.

These reports, along with monthly C3S Climate Bulletins (see Figure 7), have become key references for policymakers and climate scientists. The European State of the Climate reports (jointly published with WMO since 2023) offer detailed analyses of temperature anomalies, precipitation patterns, and other climate indicators, supporting adaptation strategies and informing the public about climate variability and change. In addition, many data from C3S are routinely used in WMO climate assessments and contribute to international climate monitoring efforts.

■ **Figure 6:** Surface air temperature anomaly (°C) for 2024 relative to the average for the 1991–2020 reference period. Data: ERA5 (above). Global surface air temperature (°C) increase above the average for 1850–1900, the designated pre-industrial reference period, based on several global temperature datasets shown as annual averages since 1967 (below).



More broadly, C3S has operationalised the production of Essential Climate Variables (ECVs) by integrating satellite and in situ observations, reanalyses (as mentioned above), and climate models into a single, quality-controlled framework. Through the CDS, C3S provides access to more than 20 ECVs, covering the atmosphere, ocean, cryosphere, and land biosphere, ensuring consistency with Global Climate Observing System (GCOS) standards. The service collaborates with over a hundred entities, including EUMETSAT and its Satellite Application Facilities (SAFs), ESA (and its Climate Change Initiative (CCI)), NASA and NOAA, to secure reliable long-term

Arctic sea ice concentration for February 2025

Data: ERA5. Reference period: 1991-2020. Credit: C3S/ECMWF

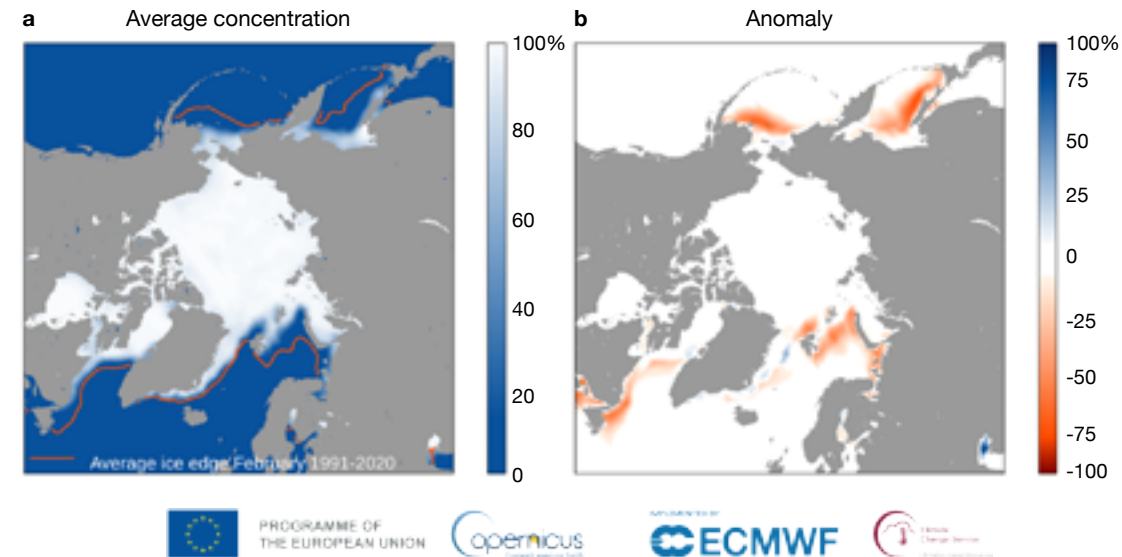


Figure 7: (a) Average Arctic sea ice concentration for February 2025. The thick orange line denotes the climatological sea ice edge for February for the period 1991–2020. (b) Arctic sea ice concentration anomaly for February 2025 relative to the February average for the period 1991–2020. This information is provided as part of the C3S Climate Bulletins published during the first week of each month⁸.

Climate Data Records (CDRs), which are regularly updated. Historical datasets, some extending back to 1755 over land and 1851 over the ocean, are complemented by atmospheric data from different networks. To address data gaps, C3S operates a Data Rescue Service in partnership with the initiative of the WMO (International Data Rescue Initiative (I-DARE)), digitising historical meteorological and early satellite observations to improve reanalyses, particularly for the 1960s and 1970s. High-resolution datasets, such as E-OBS (see e.g. Cornes et al., 2018), enhance European climate records. These efforts support global climate policies, including the Paris Agreement, by ensuring consistent, authoritative climate monitoring, allowing policymakers and scientists to track long-term trends and assess climate risks.

Another major achievement of C3S is its role in supporting the European Union's climate policies. The service provides essential data for climate adaptation and mitigation strategies, including indicators relevant to the Paris Agreement and the Sustainable Development Goals (SDGs). Through partnerships with various EU institutions and international organisations, C3S ensures that its data products align with global climate governance needs.

C3S has also actively engaged in developing capabilities for attributing extreme weather events to climate change. Between November 2019 and June 2021, C3S implemented a prototype service that established protocols for rapid attribution studies, enabling swift assessments of how human-induced climate change influences specific extreme events. The operationalisation of this capability is now under way.

Both CAMS and C3S have seen growing user engagement, with their data being used by businesses, governments, and media outlets worldwide. C3S has been instrumental in supporting climate services across Europe, while CAMS has become a reference for air quality forecasting and emissions monitoring. Their outputs are frequently cited in global media, reinforcing their role as authoritative sources for environmental and climate information.

⁸ <https://climate.copernicus.eu/climate-bulletins>

SYNERGIES BETWEEN CAMS AND C3S →

The synergy between the Copernicus Atmosphere Monitoring Service (CAMS) and the Copernicus Climate Change Service (C3S) is built upon shared infrastructure, scientific methodologies, and complementary objectives, creating an efficient and integrated framework for environmental monitoring. Both services are implemented by ECMWF and leverage its operational capabilities to provide high-quality data for users across various sectors.

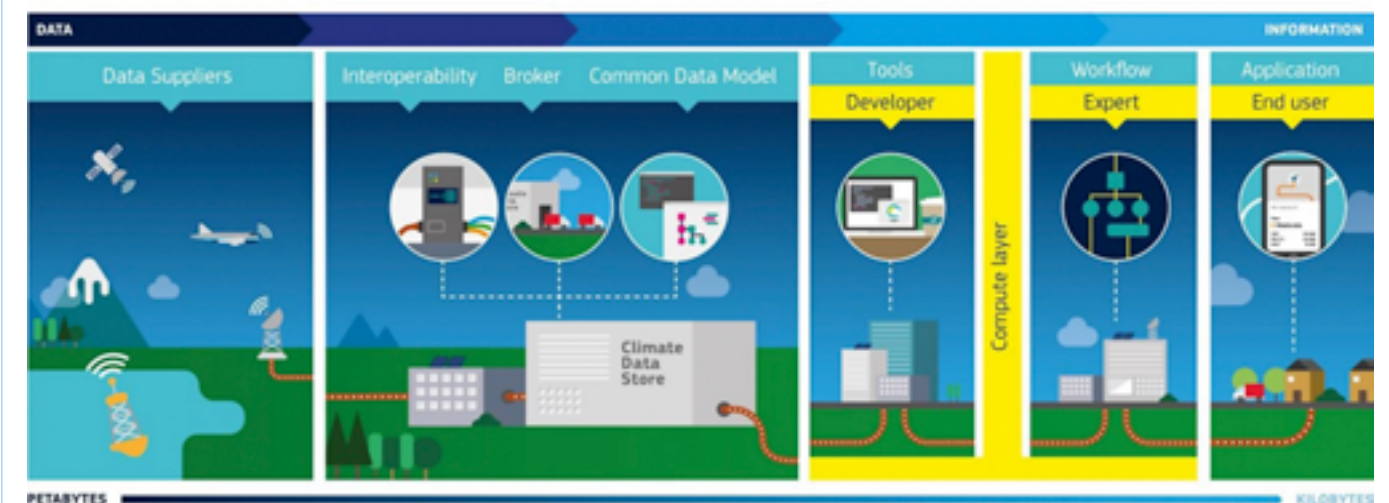
“ THE COPERNICUS ATMOSPHERE MONITORING AND CLIMATE CHANGE SERVICES ARE BUILT ON SHARED INFRASTRUCTURE, METHODOLOGIES AND OBJECTIVES.”

Starting with the Data Stores and as already mentioned above, the Climate Data Store (CDS), which underpins C3S, has transformed access to climate data by providing a cloud-based platform where users can retrieve, process, and visualise datasets. With hundreds of thousands of registered users, the CDS has enabled researchers, businesses, and decision-makers to leverage climate information in an efficient and scalable manner. CAMS launched its own data store following the same principles as the CDS but focusing on atmospheric composition data (the Atmosphere Data Store (ADS)). This joint evolution ensures that both climate and air quality information remain easily accessible, fostering wider adoption of Copernicus products. A graphic of the Data Stores' architecture is provided in Figure 8.

As a result, the CDS and ADS **both** serve as cloud-based portals. An important point is that these user-friendly platforms share a common architecture, ensuring interoperability and allowing users to seamlessly extract, process, and analyse both climate and atmospheric composition data within the ECMWF-managed Common Cloud Infrastructure (CCI), which hosts not only the CDS and ADS but also the European Weather Cloud. This integration enables efficient data retrieval and processing, reducing barriers for users who need to combine datasets from both services for applications such as climate risk assessments and air quality studies, with the option to combine these with core weather products from ECMWF.

Effective decision-making using environmental and climate information requires high-quality data, rigorous evaluation, and strong user support. Evaluation and quality control is paramount to maintain trust and authority. CAMS and C3S have therefore developed comprehensive quality assurance frameworks tailored to their specific objectives. ECMWF ensures strict quality control across all CAMS production systems, providing independent assessments and automated

Figure 8: Visual representation of the information flow through the Climate Data Store.



monitoring of forecast skill scores, statistical performance metrics, and user uptake. CAMS also produces detailed Evaluation and Quality Assurance (EQA) reports quarterly, incorporating expert reviews and long-term observational comparisons. C3S has developed similar quality assurance protocols for climate data, tools, and applications, which are published in the Climate Data Store (CDS). Quality Assurance Reports (QARs) provide detailed insights into data integrity, completeness, uncertainty, and usability, following standardised templates for observational, predictive, and projected datasets. CDS applications undergo rigorous evaluations based on relevance, accessibility, clarity, completeness, and reliability, ensuring compliance with international standards and best practices. Additionally, real-time quality monitoring of the CDS infrastructure is available through interactive dashboards, providing transparency on service performance and data usage. Through these efforts, CAMS and C3S ensure that their data products remain robust, reliable, and fit for policy and research applications.

User engagement and communication are also areas where CAMS and C3S align closely. The services collaborate on training, outreach, and policy support, ensuring that stakeholders, from policymakers to businesses and researchers, can effectively utilise their products. Joint initiatives include co-organised workshops, media collaborations, and tailored user-driven applications that highlight the intersection between atmospheric monitoring and climate change adaptation.

CAMS and C3S prioritise user engagement to ensure their data and services meet evolving user needs. This approach involves collaboration with international organisations (e.g. EUMETSAT, WMO, ESA, UNFCCC), EU institutions, national authorities, industry, and the public to enhance visibility and uptake of their offerings. A User Requirements Database (URDB) systematically collects and analyses user feedback, guiding the services’ evolution. A dedicated user support team maintains resources such as FAQs, knowledge bases, documentation, tutorials, and user forums, complemented by a help desk operating during European business hours. To broaden service adoption, CAMS and C3S support the development of use cases and demonstrators across various sectors, including air quality, health impacts, tourism, agriculture, and insurance.

Communication and outreach efforts target diverse audiences through press activities, content creation, media collaborations, and participation in major scientific conferences like the UNFCCC Conference of the Parties (COP). Their websites receive millions of page views annually, featuring climate monitoring products such as the C3S Climate Bulletins and the European State of the Climate report, which gain extensive media coverage. Additionally, CAMS and C3S sponsor environmental programmes on networks like Euronews and CNN, reaching millions globally. Training initiatives include workshops and online materials. Massive Open Online Courses (MOOCs), sometimes attended by more than 5,000 people, have been organised on topics such as monitoring atmospheric composition, using Jupyter Notebooks to exploit CAMS datasets, renewable energy, and applications of artificial intelligence and machine learning. These events are often organised jointly with partners such as EUMETSAT, ESA, Mercator Ocean International, or the European Environment Agency (EEA).

This strong synergy between the services ultimately enhances the operational efficiency and impact of both services. By aligning their methodologies and infrastructure, CAMS and C3S provide users with a coherent suite of tools and datasets that support informed decision-making in domains ranging from air pollution management to long-term climate adaptation and mitigation strategies.

EXAMPLES OF APPLICATIONS AND DOWNSTREAM SERVICES →

CAMS and C3S have facilitated a vast ecosystem of applications and downstream services that support decision-making across multiple sectors. From regulatory compliance to public health protection and economic risk assessment, these services provide actionable intelligence based on state-of-the-art atmospheric and climate data. CAMS products play an integral role in European Union environmental policies, particularly in air quality management and greenhouse gas (GHG) mitigation, while C3S data serve as a foundation for climate adaptation planning, influencing energy production, agriculture, and disaster resilience strategies.

One of the most impactful applications of CAMS is in public health, where air quality forecasts inform the public and authorities about pollution levels. These forecasts assist national and local air quality monitoring systems, enabling timely interventions to mitigate health risks. Urban planners use CAMS data to develop pollution-reduction strategies, while health agencies rely on its ultraviolet (UV) radiation monitoring capabilities to provide sun safety advisories. Additionally, the service supports smartphone applications and wearable technologies that alert users to high pollution levels, thus empowering individuals to make informed choices about outdoor activities.

As mentioned above, the inclusion of CAMS data in Windy.com enhances public access to real-time air quality information, empowering individuals and organisations to make informed decisions related to health, environmental management, and planning. This collaboration exemplifies how CAMS’ advanced atmospheric composition forecasting capabilities are effectively used in widely accessible platforms, thereby extending their reach and impact.

CAMS also provides critical support for policy-driven air quality management. By delivering detailed pollutant emission datasets, the service enables policymakers to assess the effectiveness of emission control measures and track compliance with international agreements such as the EU Air Quality Directive. Regulatory agencies use CAMS data to pinpoint pollution sources, evaluate long-term air quality trends, and design targeted mitigation policies. The service’s reanalysis products, such as EAC4 for aerosols and reactive gases and EGG4 for greenhouse gases, serve as vital tools for tracking atmospheric changes over time, helping governments and international bodies to fine-tune their environmental policies.

One of the emerging applications of CAMS is in emission monitoring and verification support, particularly for methane emissions. CAMS provides real-time and retrospective data on methane levels, which are essential for tracking emissions from agriculture, fossil fuel extraction, and landfills. These data are crucial in the context of the European Union’s Methane Strategy, which emphasises reducing emissions from key sectors. By integrating satellite-based methane observations with atmospheric modelling, CAMS enables regulators and industries to implement more effective mitigation measures.

On the economic front, C3S provides indispensable climate intelligence for businesses. Energy companies leverage seasonal forecasts to optimise energy production and distribution, ensuring a stable and cost-effective supply of electricity. For instance, hydropower operators use C3S data to anticipate changes in water availability, while wind and solar energy producers utilise its datasets to forecast resource variability. Insurers, on the other hand, rely on C3S climate projections to assess weather-related risks, set premium rates, and design long-term risk mitigation strategies. The latest report⁹ on the economic value of ERA5 underscores its growing importance as a reference dataset for financial risk modelling.

⁹ <https://climate.copernicus.eu/new-study-era5s-socio-economic-benefits>

Another notable example is the European Energy and Climate Data Explorer, which integrates ERA5 reanalysis and regional climate projection data to provide insights into the impact of climate variability on energy demand and supply. This tool supports policymakers, energy companies, and grid operators in designing resilient energy systems that can withstand changing climate conditions. Also, and by leveraging C3S data, companies such as Vortex enhance the precision of their wind energy assessments, contributing to more reliable and efficient renewable energy solutions. Moreover, C3S and the European Network of Transmission System Operators for Electricity (ENTSO-E) have established a collaborative partnership to enhance the integration of climate data into energy sector planning and strategic decision-making for a climate-resilient European electricity grid. These applications demonstrate how C3S's extensive climate datasets empower the energy sector to make data-driven decisions, ultimately promoting the adoption of sustainable energy practices.

A compelling example of C3S's contribution to the climate services landscape is its role in machine-learning driven weather prediction. The ERA5 reanalysis dataset serves as the natural foundation for numerous machine-learning models developed by research institutions and private companies worldwide (see ECMWF 50th anniversary papers on Earth system modelling, data assimilation and machine learning). By training artificial intelligence systems on high-resolution historical climate data, these models can generate more accurate weather forecasts, ultimately benefiting industries ranging from aviation to agriculture.

C3S has also fostered the development of innovative climate monitoring tools. The ERA5 Explorer application, for instance, provides a user-friendly interface for generating local statistics on climate variables based on reanalysis data. Meanwhile, the Global Temperature Trend Monitor helps researchers and policymakers assess long-term warming trends and estimate when the planet might reach the warming threshold of 1.5°C proposed in the Paris Agreement. Such applications exemplify how C3S bridges the gap between raw climate data and real-world decision-making.

The health sector has also benefited from C3S applications. The Monthly Climate Explorer application for COVID-19, for instance, examines possible links between weather patterns, air pollution, and COVID-19 mortality rates, offering valuable insights into the interplay between climate and public health. Such applications highlight the expanding role of climate services in addressing cross-sectoral challenges.

A recurring theme in these applications is the integration of data across multiple domains. The European Climate Data Explorer, hosted on the Climate-ADAPT platform and managed by the European Environment Agency, provides interactive access to numerous climate indices, supporting national and local adaptation strategies across Europe. The CDS and ADS infrastructures, along with application programming interfaces (APIs) and web-based analysis tools, ensure that users from diverse backgrounds can harness the full potential of Copernicus data.

These examples illustrate the transformative impact of CAMS and C3S in fostering a data-driven approach to environmental management, economic planning, and policy development. By enabling a wide range of downstream services, these Copernicus services are not only advancing scientific research but also driving real-world action to address climate change and atmospheric pollution.

ECMWF CONTRIBUTION TO THE COPERNICUS EMERGENCY MANAGEMENT SERVICE →

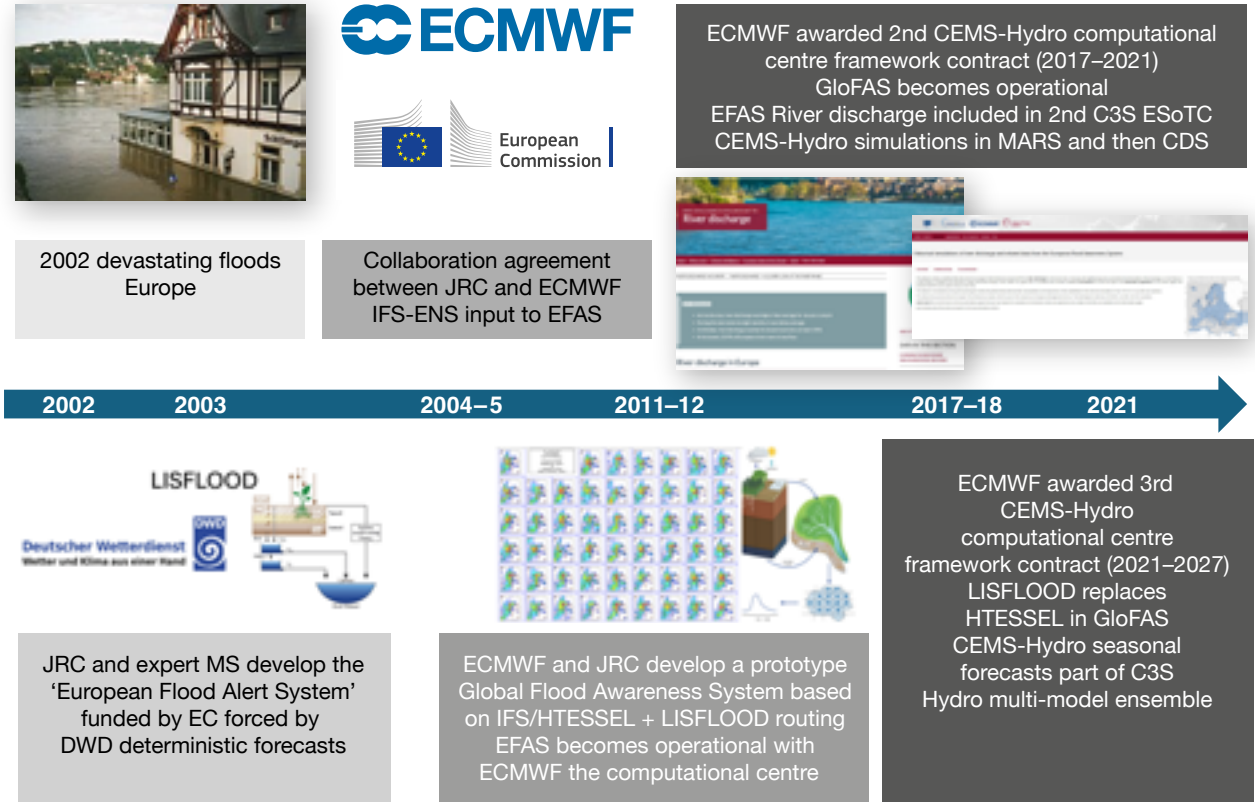
FLOOD AWARENESS SYSTEMS

Beyond the two major services, CAMS and C3S, which ECMWF operates on behalf of the European Union, ECMWF contributes to the Copernicus Emergency Management Service (CEMS), under the leadership of the JRC, in particular the early warning systems for flood and wildfire.

ECMWF has an inspiring and fruitful long-standing collaboration with the JRC, especially to support the development, implementation and operational delivery of CEMS-Flood, the Early Warning for floods component of the Copernicus Emergency Management Service (see Figure 9). CEMS-Flood, managed by the JRC and delivered by a network of centres, currently covers two domains to deliver kilometre-scale hydrological forecasts from short-range (Europe only) to seasonal forecast horizons (global).

In 2004 ECMWF signed a collaboration agreement to provide ensemble forecasts as input to the newly developed experimental 'European Flood Alert System', the precursor of the European Flood Awareness System (EFAS), run as a prototype by the JRC. In 2011, EFAS became an operational service as one of the CEMS Early Warning Systems, and ECMWF was designated as the computational centre for CEMS-Flood. At the same time, ECMWF and the JRC pioneered the set-up of a global flood forecasting system by coupling ECMWF's Integrated Forecasting System with the JRC's hydrological routing component of its distributed hydrological model, OS-LISFLOOD. ECMWF ran the system experimentally until 2017, when it became an official CEMS-Flood service (the Global Flood Awareness System (GloFAS)). The collaboration continued with ECMWF remaining the CEMS hydrological forecast computational centre for the 2nd and 3rd framework contracts (ending in 2027).

Figure 9: Schematic of ECMWF's role in CEMS-Flood.



As the CEMS hydrological forecast computational centre, ECMWF is responsible for the data acquisition from the CEMS-Flood Meteo and Hydrological Data Collection Centres and numerical weather prediction centres such as DWD and the COSMO consortium, for the operational running and maintenance of hydrological forecast chains, and for the delivery of data and associated products to authorised users. It also develops and maintains tools to support the role of the Hydrological Analytics and Dissemination Centre and provides technical and product user support and guidance. In collaboration with the JRC, ECMWF is at the core of technical service development, including building new modelling modules, conducting hydrological model calibration, designing forecast products, optimising modelling chains and product software, and improving all data services. By March 2025, more than 15 separate modelling chains served CEMS-Flood hydrological forecasting operations, with nearly 80 products updated and pushed daily on the EFAS and GloFAS web interface.

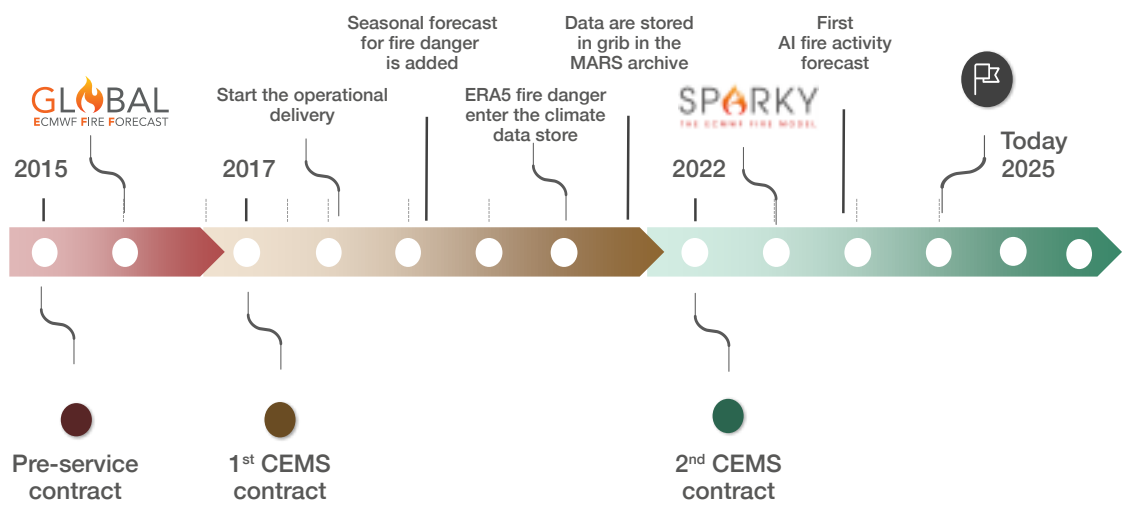
Through its role as the computational centre for CEMS hydrological forecasts and entrusted entity of C3S and CAMS, ECMWF has been instrumental in exploiting synergies across the different services. Since 2017, the daily hydrological simulations of EFAS initial conditions have served as reference data for the C3S Climate Intelligence activities, including the European State of the Climate report and the Climate Bulletins, with CEMS-Flood simulated river discharge first appearing in 2018 as a spotlight variable before becoming a core variable from the 2020 edition onward. ECMWF has also leveraged its in-house expertise in data archiving and discovery, offering to run the CEMS Early Warning Data Store, a tool for data cataloguing and data access built on the same infrastructure as the C3S and CAMS data stores. Between August 2023 and February 2025, over 170 terabytes of CEMS-Flood data have been downloaded through the Data Store to serve more than 620,000 requests from 110 countries.

Until the end of the framework contract as the computational centre for CEMS hydrological forecasting, ECMWF will implement a new web infrastructure to better serve EFAS and GloFAS products, deliver at least one major and several minor service releases for each European and global domain component, and work on new products, continuing the long-standing collaboration with the JRC to advance the service.

Through partnerships and collaborations with the JRC and all CEMS-Flood centres, including DWD, Kister, Ghenova Digital, SMHI, SHMU and Dutch Rijkswaterstaat, ECMWF has taken pride in delivering part of the CEMS-Flood service.

Under CEMS, ECMWF also has the responsibility for producing fire danger forecasts. This complements ECMWF's role under CAMS on various aspects of fires, partly in the context of the Global Fire Assimilation System (GFAS), focusing on detection of active burning fires, their emissions and their impact on atmospheric composition and air quality (see earlier sections). This section covers the CEMS aspects.

ECMWF has been producing fire danger forecasts operationally since 2018 as part of CEMS. A pre-exploitation phase started in 2015, conducted with the support of a grant issued by the JRC, which aimed to explore the use of weather forecasts for the delivery of global fire danger information (Di Giuseppe et al., 2016). Until then, most available systems relied on environmental monitoring through surface weather stations. See Figure 10 for the timeline of ECMWF involvement in support of the fire component of CEMS Early Warning.



■ **Figure 10: Timeline of fire activities and key milestones for ECMWF fire-related work under the Copernicus Emergency Management Service.** During the pre-exploitation phase ECMWF built the Global ECMWF Fire Forecasting (GEFF) model, a model able to calculate fire danger indices from several models developed in the US, Canada and Australia. During the two service contracts ECMWF expanded the availability of data including seasonal prediction and real-time monitoring of fire danger from ECMWF reanalyses. In recent years ECMWF has been exploring the possibility of forecasting fire activity by directly leveraging ML methods and physically-based models of fuel status.

During the pre-exploitation phase, ECMWF developed a global fire danger model, the Global ECMWF Fire Forecasting (GEFF) model, which allows the simultaneous calculation of several fire danger metrics developed in Canada, the US, and Australia. Using GEFF, ECMWF has been delivering forecast products to both the European Forest Fire Information System (EFFIS)¹⁰ and the Global Wildfire Information System (GWIS) platforms. Today, GEFF remains one of the open-source resources widely used by both operational services (e.g. NOAA) and technology companies (e.g. IBM).

After an initial research phase, ECMWF was awarded two consecutive Copernicus service contracts in 2017 and 2022. Thanks to the two service contracts, the range of available products has expanded to include seasonal predictions, and a real-time monitoring dataset based on ERA5, which can be accessed directly through the CDS. ERA5-Fire ranks among the most downloaded datasets in the CDS and is widely used in the fire community for studying changing patterns of landscape flammability worldwide. Through the Copernicus Programme, ECMWF has also introduced a new discipline and data template in the governance of meteorological data managed by the WMO. As a result, fire forecast variables can now be codified similarly to traditional weather variables.

While the first phase focused on adopting existing methods, in recent years, ECMWF has been exploring novel approaches to fire forecasting. By leveraging advances in machine learning and its own weather prediction model, ECMWF is shifting the focus from predicting fire danger (a measure of landscape flammability) to forecasting fire activity.

This progress has been made possible by utilising ECMWF's model infrastructure and initiating the development of a fire model within its physics-based framework, known as SPARKY. The introduction of high-resolution fire activity forecasts (up to 1 km), which incorporate not only weather conditions but also fuel status and human presence, expands our current capability for early detection of fires on a global scale. This advancement could position the Copernicus Programme at the forefront of fire forecasting in the years to come, contributing to building resilience against the escalating threat of extreme fire events in a warming climate.

¹⁰ <https://forest-fire.emergency.copernicus.eu/>



CONCLUSION AND FUTURE WORK

The initial phase of CAMS and C3S has delivered significant value in atmospheric and climate monitoring, firmly establishing both services as globally recognised and authoritative sources of environmental information. Over the past decade, these services have evolved from research initiatives into fully operational programmes that provide critical support for scientific research, policymaking, and commercial applications.

On 22 July, 2021, ECMWF signed a new Contribution Agreement with the European Commission to continue implementing the EU-funded C3S and CAMS for the next seven years. During this Phase 2 (2021–2028), Copernicus is building on this solid foundation by incorporating cutting-edge advancements in Earth system monitoring, high-resolution reanalyses, artificial intelligence (AI) integration, and enhanced user engagement.

“ THE CO₂ MONITORING AND VERIFICATION SUPPORT CAPACITY WITHIN CAMS WILL PROVIDE NEAR-REAL-TIME ASSESSMENTS OF ANTHROPOGENIC CO₂ AND CH₄ EMISSIONS.”

A key component of this evolution is the development of the CO₂ Monitoring and Verification Support (CO2MVS) capacity within CAMS, which will provide near-real-time, observation-based assessments of anthropogenic CO₂ and CH₄ emissions. This capability will leverage data from Sentinel-5P, the upcoming CO₂M mission, and other contributing satellites to monitor emissions at finer spatial and temporal scales, supporting both national greenhouse gas (GHG) inventories and international climate agreements such as the Paris Agreement and the European Green Deal. This marks a transformative step in atmospheric monitoring, shifting from emissions estimates based on statistical inventories to direct, observation-based tracking of GHG emissions. Figure 11 represents the timeline for the implementation of this new service element of CAMS. This initiative fits in and will be an important contributor to a wider global initiative called the Global Greenhouse Gas Watch (G3W) being set up by the WMO.

On the climate side, C3S is preparing the next-generation ERA6 reanalysis, which will provide even greater accuracy, resolution, and coverage. This coupled Earth system reanalysis will integrate data from the atmosphere, ocean, land, and cryosphere to improve long-term climate assessments and inform adaptation strategies. By providing detailed reconstructions of past climate conditions, ERA6 will support risk assessments in critical areas such as water resource management, food security, and infrastructure planning.

Alongside ERA6, the planned EAC5 reanalysis for atmospheric composition will offer enhanced insights into long-term trends in air pollutants and greenhouse gases, further strengthening the link between climate services and policy-relevant air quality monitoring.

To better serve the evolving needs of EU Member States, the Copernicus Programme is expanding its National Collaboration Programme. This initiative allows individual countries to tailor Copernicus data products for national applications, facilitating the integration of Copernicus information with locally available datasets. The National Collaboration Programme enhances the synergy between Copernicus services and national climate adaptation efforts, supporting policy frameworks such as the EU Climate Law and the revised Ambient Air Quality Directives.

Looking beyond 2028, Copernicus services will continue to expand their role in global environmental governance. CAMS will refine its capabilities in air quality and emissions monitoring, improving the accuracy and timeliness of its forecasts while strengthening its contribution to European and international regulations. The service will maximise the impact of satellite data by integrating observations from new missions such as Sentinel-4, Sentinel-5, and CO2M and by promoting the use of low-cost in situ sensor networks in data-sparse regions. C3S, on the other hand, will enhance its support for climate resilience by contributing to providing an operational framework for climate projections. For example, brokering access to climate forcings and scenarios could feed into projections streams produced by, for instance, the Destination Earth (DestinE) initiative which aims to create digital twins of the Earth (Sandu, 2024), enabling interactive, high-resolution climate simulations that will provide decision-makers with dynamic insights into future climate risks.

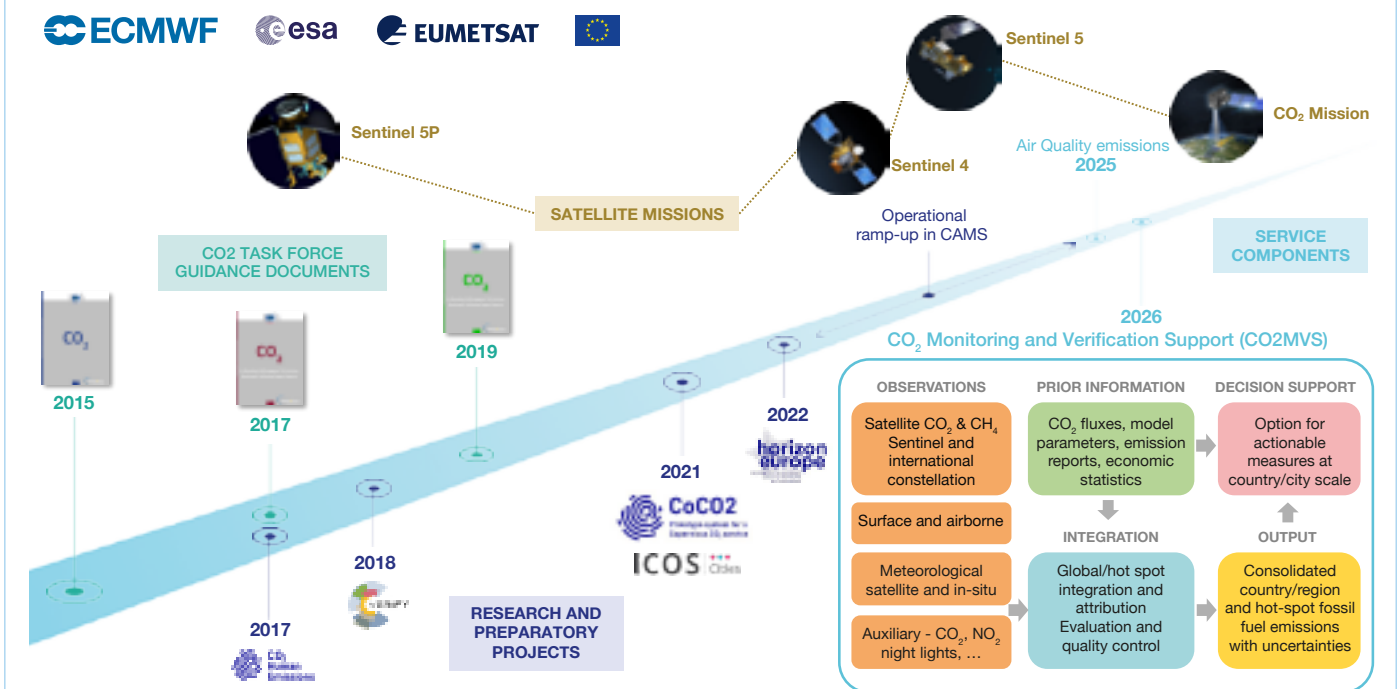


Figure 11: Timeline for the implementation of the Copernicus anthropogenic greenhouse gas emissions monitoring and verification support capacity. The scoping of this new element of CAMS, implemented by ECMWF, started in 2015 with reports from a CO₂ Monitoring Task Force from the European Commission. Supported by dedicated research projects within Horizon 2020 and Horizon Europe, the science and methodologies have been developed and are now being implemented as part of CAMS. In parallel, the space segment is being ramped up by ESA and EUMETSAT through the Sentinel-5P, Sentinel-4, Sentinel-5 and CO2M missions. The overall structure of the envisaged integrated system that combines observations and prior information with modelling capacities at global and local scale is also shown (in the box).

AI and machine learning will play a crucial role in the next phase of Copernicus services. CAMS and C3S are already integrating AI-based techniques to improve data processing, uncertainty quantification, and model downscaling. These innovations will allow Copernicus to provide more accurate, localised forecasts while reducing computational costs and latency. The expansion of AI-driven services will also enhance climate risk assessments and facilitate more precise monitoring of extreme weather events, which are becoming increasingly frequent due to climate change.

Another critical aspect of the future evolution of Copernicus services is their growing economic and societal impact. Climate variability and atmospheric composition are increasingly influencing decision-making in energy production, finance, agriculture, and disaster management. C3S’s seasonal forecasts and ERA5-derived climate risk assessments are already supporting renewable energy companies, including the wind energy sector in optimising production and grid stability. Similarly, insurers are using Copernicus data to assess climate risks and structure innovative financial products that mitigate economic losses from extreme weather events.

The next decade will see CAMS and C3S expand their user-driven portfolios within their operational frameworks, focusing on increased automation, enhanced timeliness, and improved accessibility of their services. The ambition is not only to sustain the current level of service but also to expand capabilities in areas such as air quality monitoring, emissions verification, and climate projection operationalisation. Copernicus 3.0 will aim to integrate near-real-time environmental intelligence, providing rapid-response capabilities for managing climate and air quality crises, such as wildfires, industrial accidents, and extreme weather events.

Equally, ECMWF intends to continue and expand its capabilities in early warning for floods and fire risks in support of CEMS with the JRC.

As climate change becomes an increasingly urgent global challenge, Copernicus services will continue to play a crucial role in climate action, risk mitigation and adaptation strategies. The planned enhancements in reanalysis, forecasting, and emissions tracking will provide policymakers, scientists, and businesses with even more powerful tools to tackle environmental challenges and drive sustainable transformation. Through ongoing innovation and collaboration, Copernicus services will ensure that Europe remains at the forefront of climate and atmospheric science, helping societies worldwide navigate an era of unprecedented environmental change.

CONTRIBUTORS

ECMWF would like to thank the contributors to this paper:

Jean-Noël Thépaut, Carlo Buontempo, Samantha Burgess, Francesca Di Giuseppe, Richard Engelen, Vincent-Henri Peuch, Christel Prudhomme and Laurence Rouil.



The critical role of high-performance computing in medium-range weather forecasting: half a century of technology innovation

ABSTRACT



Over the past five decades, high-performance computing (HPC) and data storage technologies have performed a pivotal role in advancing the operational time-critical capabilities and scientific research at ECMWF and among its Member States. HPC systems have enabled researchers to develop and run increasingly complex numerical weather prediction simulations, improving the model fidelity and increasing forecast lead times. Concurrently, data storage innovations have been vital to support the vast and expanding datasets, preserving half a century of meteorological data essential for weather and climate research. This article highlights the Centre’s infrastructure milestones from the early supercomputers to today’s HPC and storage solutions which underpin real-time prediction and multi-decadal climate analysis. Together, these services have enabled major advances in forecasting precision and scope, allowing the meteorological community to meet the urgent demands of modern weather-related challenges, including forecast reliability, emergency management planning for weather-dependent sectors, and insights into global climate change.

We describe the evolution of ECMWF’s core infrastructure into a highly resilient environment necessary to deliver time-critical services. The article highlights the roles of HPC, networks, data centre design, and storage in delivering ECMWF’s mission to provide reliable, cutting-edge forecasts and support future generations of weather prediction systems. With data volumes and computational demands continuing to grow, ongoing technological innovation remains an essential component to advance medium-range weather forecasting.

Reflecting on the impact of technology over the decades, the continued scientific advancements have been achieved through the growth and diversity of the communities’ skills in exploiting these new capabilities. It has been a pleasure to record the massive progress that has been achieved over the last 50 years. As readers we hope you enjoy the article as much as we did in researching the history of our services, including some of the changes in fashion and designs!

INTRODUCTION →

For decades, high-performance computing (HPC) and storage technologies have been at the heart of weather forecasting services, enabling meteorologists to model complex atmospheric dynamics with ever-increasing accuracy. From the early days of numerical weather prediction (NWP) to today’s exascale generation supercomputers, advancements in computational capabilities and data management have enabled significant improvements in forecast precision and reliability. Accurate medium-range weather forecasting (forecasts ranging from 3 to 15 days ahead) is essential for disaster preparedness, agriculture, transportation, and energy management. The ability to predict weather patterns with precision relies on high-performance computing, which enable meteorologists to process vast amounts of atmospheric data and run sophisticated NWP models. Storage technology advances have allowed the Centre to keep pace with processing higher and higher volumes of data from those sophisticated NWP model runs.

In the 1970s, weather prediction used early mainframe computers, which, whilst state of the art at the time, had limited computational resources, particularly with respect to processing and memory capabilities. Due to computational constraints, the model grid resolutions were relatively coarse and involved simplified physical parametrizations, leading to errors in simulating small-scale atmospheric phenomena such as localised storms. As HPC systems evolved, so did the ability to run sophisticated models that simulate atmospheric behaviour at finer spatial and temporal resolutions. This allowed numerical weather prediction centres to increase model complexity, incorporate more observational data, and extend the range of reliable predictions.

Alongside computational power, storage technologies have played a critical role in handling the exponential growth of meteorological data. The rise of satellite observations and ensemble forecasting has generated vast datasets requiring advanced, affordable high-speed storage, efficient retrieval systems, and scalable data infrastructures. Without these advancements, processing and analysing the massive volumes of atmospheric data necessary for accurate forecasting would not be possible.

25% of ECMWF’s HPC capacity is allocated to Member States, providing national meteorological services with advanced computing and storage access to support research, model development, pre-operational testing, and operational enhancement activities. These resources also support Member States in exploring the use of AI-driven tools alongside traditional physics-based models.

This article describes the transition from ECMWF’s early supercomputers to the latest-generation platforms and it explores the important role of the data centre and network technologies required to host and connect these systems in a highly resilient environment to operate complex time-critical services. It is also and above all a testament to the constant support and collaboration between ECMWF, its Member and Co-operating States, and the broader European Meteorological Infrastructure (EMI).

Figure 1: Control Data Corporation (CDC) 6600 computer hosted by CDC at John Scott House in Bracknell, UK, close to ECMWF's temporary accommodation.

ECMWF Newsletter No 3 (ECMWF, 1975) proudly reported that:

“Usually, 200 jobs a week or more are now being run on the CDC 6600 computer at John Scott House, using more than 20 hours central processor time (of the 40 hours available). On the U.K. Meteorological Office complex, a total of over 200 jobs were run in November (144 on the 360/195, 66 on the 370/158), using over 5 hours of 360/195 central processor time”.

In the 50 years since that was written, the world has seen an almost unimaginable increase in the amount of computer power available; the ECMWF HPC service has increased in computational power by almost 30 million times, and 200 jobs are now run every 20 seconds rather than every week.



ORIGINS OF ECMWF
COMPUTING: FROM CDC
TO CRAY

The beginnings

In December 1975, ECMWF had no dedicated data centre and rented time on a Control Data Corporation (CDC) 6600 computer to run 200 jobs a week (see Figure 1). With 40 hours of central processing unit (CPU) time available each week, the CDC supported initial model development but lacked the power necessary for operational forecasting. Whilst the CDC was itself a significant advancement in computing technology, being a factor of three times faster than the previous fastest system in the world, running a 10-day forecast would have taken 12 days!

1970s: the Cray-1 revolution

The Cray-1 was released in 1976 and was famously built in a C-shaped cabinet surrounded by a ring of benches that hid the power supplies and cooling systems (see Figure 2). The C-shape allowed the length of the myriad of cables connecting the individual module boards to be short and of similar lengths so that electrical signals arrived at the right time. Initially, ECMWF used the very first of these systems, serial number one, that had previously been installed at the Los Alamos National Laboratory in the United States. As this preceded the ECMWF headquarters, it was hosted at the Rutherford Laboratory, located in Didcot not far from Reading, UK. From 1978, ECMWF had its own permanent system, serial number nine, installed at the new headquarters at Shinfield Park in Reading.

Figure 2: The Cray-1A in the ECMWF data centre in Reading (Shinfield Park).

The Cray-1 vector processing machine dramatically improved computational capabilities: the then groundbreaking 160 MFLOPS (million floating-point operations per second) reduced the forecast runtime from a theoretical 12 days to a practical 5 hours. For comparison, the Cray-1A computational capabilities were less than a tenth of the performance of today's smartwatch.

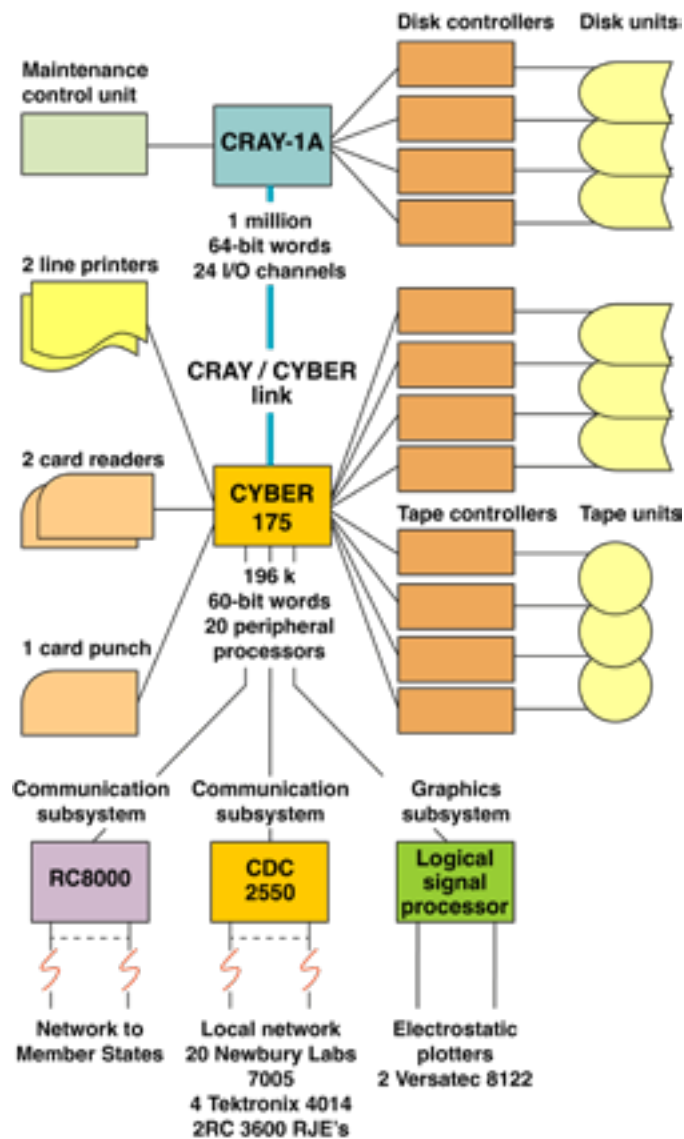


Access to these early supercomputer systems was difficult; the Cray was a dedicated processing unit and required input and output to be fed through a front-end system (Figure 3). Programs had to be prepared and submitted on punch cards (Figure 4). Each punch card encoded a single line of code or data using a series of holes punched in predefined positions, with entire programs involving stacks of these cards. Users would prepare their job decks – often hundreds or thousands of cards long – and physically submit them to the computing centre's operations team. The Cray system relied on batch processing workflows where input from punch cards was read sequentially, compiled, and then executed by the system. This process was time-consuming and inflexible, and required meticulous card handling and error checking, as a single misplaced or incorrectly punched card could crash a job or produce faulty results. Despite these limitations, punch cards remained in use during the early years of high-performance computing, until the arrival of terminals, which gradually started being introduced at the Centre from 1979.

1980s: transition to terminals

Punch cards continued to be used into the early 1980s. However, by 1979 ECMWF had 20 alphanumeric visual display units and 4 graphical units in use; several terminals were placed in staff members' offices. Terminals proved to be very popular, so 34 more were purchased the following year, and by 1982 most staff had a terminal in their office, marking a significant productivity leap. These early systems established ECMWF as a pioneer in supercomputing for numerical weather prediction.

■ **Figure 3:** The Cray-1A and front-end systems, reproduced from ECMWF Technical Newsletter No. 1 (ECMWF, 1979). The main components were the CRAY-1A “number cruncher”, a CDC CYBER 175 “front-end” and a Regnecentralen RC8000 telecommunications system, all linked via high-speed channels.



SCALING UP: CRAY TO FUJITSU AND THE PARALLEL COMPUTING SHIFT

1970s–1990s: the vector era expands

The Cray-1 system achieved its groundbreaking performance by having a small number of processors implementing vector processing where one instruction could perform calculations on a set of data, with hardware dedicated to specific instructions, e.g. addition and multiplication required separate hardware. The vector programming model carried ECMWF supercomputing through the next three decades, scaling its computing capacity with the Cray X-MP, Y-MP, and finally the C90 generations. These systems expanded vector computing, retaining a shared-memory architecture design.

1990s: Fujitsu Vector Parallel Processing (VPP) – a paradigm shift

In 1996, ECMWF transitioned to the Fujitsu VPP700 system – the first distributed-memory vector-parallel machine, providing more than 600 times the performance of the original Cray-1A. The VPP700 had 39 processing elements for computing, 6 for input/output (I/O), and 1 running the batch system and interactive work. Each

■ **Figure 4:** An ECMWF analyst using a punch card preparation system.



processing element had its own 2 gigabyte local memory, requiring applications to explicitly manage data exchange between the processors. The challenge to overcome for many scientific problems is that the calculation being worked on by one processor affects the calculation being performed on other processors, i.e. the weather in one location affects the weather in locations around it. So, there must be a very fast method of communicating information between tasks. This challenge demanded a fundamental redesign of forecasting software and introduced parallelism as a foundation of ECMWF computing models.

Networking challenges

This distributed-memory architecture required significant changes to the forecast system and a fast network to connect the processors. Without shared memory, applications had to send results to all the other parts of the simulation that might need that information. The configuration of the network design became an increasingly major element for software developers and system architects to consider, to manage costs and to optimise performance of the service. Supporting inter-process communication prompted the use of crossbar switches and network topology-aware coding strategies. Future HPC systems would move to more scalable network designs, including hypercubes and tree configurations, before eventually lighter configurations such as the dragonfly topology were introduced.

For the Fujitsu VPP700 system, the networking was a simple “non-blocking crossbar switch” and it was used for operational forecasts from September 1996.

THE MASSIVELY PARALLEL ERA: IBM TO CRAY XC30

Seymour Cray, the founder of Cray Research, was a strong advocate for the vector programming model of high-performance computing and is often quoted as saying: “If you were ploughing a field, which would you rather use: two strong oxen or 1024 chickens?”. A statement of its time, that was soon made obsolete by the evolution of microprocessor technology. This saw the chickens evolve from the simple and specialised integrated circuit chips with a few thousand transistors in the 1970s to the current mass market devices with over 40 billion transistors. The increasing number of transistors on a chip became known as Moore’s law, named after Gordon Moore, one of the co-founders of Intel, who noted in a 1965 article that circuit density was doubling roughly every year and was likely to remain that

way for quite some time (Moore, 1965). The rapid development of commercial microprocessors in the 80s and 90s influenced the next phase of high-performance computing, the massively parallel systems.

2000s: IBM Power Systems

ECMWF’s first massively parallel system arrived in 2003 with IBM Power 4. With around 1,400 processors based on IBM p690 servers, it vastly outperformed its predecessors. Each server had 32 processors logically partitioned into four 8-way nodes, each node with 8 gigabytes (GB) of memory, with a high-memory 12-node subset containing 32 GB per cluster.

The IBM proprietary Colony Switches temporarily provided the high-speed interconnect. However, as this was older technology, it was a short-lived system that was replaced a year later by two IBM p690+ clusters. Each 70-server cluster was connected by a pSeries “Federation” switch, which was four times faster, providing 1,700 MB/s per link bandwidth performance.

High Performance Parallel Interface (HiPPi) networks were introduced to link the HPCs with the storage clusters and with the IRIX-based SGI general-purpose systems as Ethernet could only provide 100 megabits per second (Mbps) links at the time. HiPPi could do 800 Mbps, a major achievement in High Performance Networks (HPN), which did not survive for long as Ethernet advancements quickly took over the market.

The general-purpose systems provided, amongst other things, in-house developed connectivity tools for sharing the HPC resources with Member States, enabling them to access and, almost seamlessly, submit remote jobs from their home site over the 9 Mbps Internet connection. These tools introduced cryptography into ECMWF for the first time. As the systems became more open to remote (full) access users throughout Europe, the need for security increased, access lists on the routers were no longer enough, and the market was in its infancy with regards to dedicated firewall devices. Stateless Private Internet eXchange (PIX) firewalls with very basic packet-filtering were amongst the few commercial devices available. These were not enough to keep the HPC secure, and stateful firewalls were only offered by small high-risk startup companies with very high price tags. ECMWF therefore decided to build its own firewalls based on the open-source Firewall Toolkit (FWTK) which already in the nineties offered stateful application-level security functionality. These initial firewalls were built on existing HP RISC redundant servers running HP-UX in tandem with an at the time leading-edge product that provided automatic failover in case of hardware failures (ServiceGuard). These high availability home-built firewalls were in use until the market matured and ECMWF migrated to dedicated Nokia-Checkpoint devices.

The ECMWF relationship with IBM continued across a rapid evolution through the Power5, Power6 and Power7 generations, scaling to 46,000 processors by 2013 – each processor 30 times faster than the entire original Cray-1, resulting in the Power7 system delivering around 1.4 million times the performance of the original Cray-1.

The idea to create an HPC compute grid by interconnecting the European National Supercomputing Sites materialised in 2002 with the DEISA (Distributed European Infrastructure for Supercomputing Applications) project. This project interconnected ECMWF with sites such as MPI, CINECA, CSC, LRZ, JRZ, BSC, HLRS, EPCC, CNRS, and SARA¹ with features such as cross-mounted shared filesystems (based on a multi-cluster General Parallel File System – GPFS) and a Common Software Stack Environment. DEISA later evolved into PRACE, and PRACE evolved into the

current EuroHPC Joint Undertaking, which provides considerable HPC resources for scientists across Europe.

2010s: Cray XC30 and the dragonfly network

From the late 1990s, the Intel x86 architecture found in desktop computing systems rose to dominance. At their peak in the middle of the 2010s, Intel CPUs accounted for almost 90% of global processor sales, and the economies of scale had made it almost inevitable that the IBM successor systems would be based on Intel technology. In 2014 ECMWF installed two Intel x86-based Cray XC30 systems. Each system had around 3,500 nodes, containing two Intel Ivy Bridge 12 processor chips with 64 gigabytes memory per node, providing a total of 168,000 processors across the entire service.

The number of processors per node sharing a common memory made it important that the first level of parallel programming exploited this feature. However, the large number of nodes put heavy requirements on the high-speed interconnect fabric. Cray developed its own proprietary interconnect, known as Aries. The Aries interconnect used a dragonfly topology to limit the number of connections, particularly the longer cables, as whilst the Cray-1 was contained in one small cabinet, each Cray XC30 cluster was housed in 19 cabinets with the complete system weighing in at 93,000 kilograms. The intra-cabinet and inter-cabinet dragonfly interconnect topology hierarchy added another level of operational complexity, as applications performed better if they could keep communications within a cabinet.

Broadwell upgrade

The modular design of the Cray system allowed a mid-term upgrade in 2016 to the next generation of Intel processors. Cray replaced all the processor boards in the system with Intel Broadwell chip boards. These had 18 processor cores, and consequently took the total processor count to 260,000 without changing the overall physical footprint of the system.

2020s: Atos BullSequana XH2000 in Italy

Fully operational since 2022, ECMWF’s Atos BullSequana XH2000 system is the latest HPC service at ECMWF, the first hosted in the purpose-built Bologna data centre (see the section ‘Designing data centres: from Reading to Bologna – a chronology of key milestones’) and the first to introduce four operational clusters or “complexes” (Figure 5). Hosting over 1 million AMD EPYC Rome processor cores, using NVIDIA-Mellanox InfiniBand interconnect, the system delivers more than 26 petaflops of peak performance – equivalent to 167 million times the CDC 6600. The system weighs more than 200 metric tons, requiring 40 water-cooled cabinets for the main compute nodes, and an additional 40 racks of associated servers and storage. The entire system consumes around 40 million kilowatt hours of electricity per year, equivalent to ca. 15,000 UK homes. The solution is based on AMD Rome chips, each with 64 processors, giving more than a million processor cores in the complete system. Each processor delivers 27 times the performance of the first Cray-1, resulting in an increase of 28 million times the performance of the original Cray-1 (see Table 1).

The system includes two generations of NVIDIA graphics processing unit (GPU) hardware. GPUs were originally developed for manipulating arrays of data for displays. They use massive parallelism to perform relatively simple operations on vast quantities of data as it streams though the system. Exploiting the processing

¹ MPI – Max-Planck-Institute; CINECA – Italian national supercomputing consortium; CSC – IT Centre for Science; LRZ – Leibniz Supercomputing Centre; JRZ – Jülich Supercomputing Centre; BSC – Barcelona Supercomputing Center; HLRS – High-Performance Computing Center Stuttgart; EPCC – Edinburgh Parallel Computing Centre; CNRS – French National Centre for Scientific Research; SARA – Stichting Academisch Rekencentrum Amsterdam.

■ **Figure 5:** Two of the four clusters in the Atos BullSequana XH2000 system in ECMWF's data centre in Bologna, Italy.



■ **Table 1:** Comparison of ECMWF's Cray-1A and Atos BullSequana XH2000 systems.

	Cray-1A	Atos BullSequana XH2000
Year installed	1978	2022
Architecture	Vector processor	Massively parallel cluster
Number of cores	1	1,015,808
Clock speed (MHz)	80	2,250 (2.2 GHz)
Peak perf. per core (MFLOPS)	160	36,000
Peak perf. (MFLOPS)	160	26,687,760,000
Sustained performance (MFLOPS)	50	1,401,003,800
Memory (MiB)	8	2,146,304
Disk space (GB)	2.5	10,084,400,000

Towards heterogeneous HPC

The next phase of high-performance computing at ECMWF will increasingly be driven by data-driven machine learning. Data-driven models have recently surged in skill, but have hardware and software requirements that are very different from traditional HPCs. A hybrid future in which AI workloads are performed alongside the current, traditional HPC jobs will lead to significantly more complex HPC systems, blending CPUs, GPUs, and specialist accelerators, supported by tiered storage and intelligent scheduling systems.

LOCAL- AND WIDE-AREA NETWORKING →

In the late 1970s, external telecommunications for ECMWF's computing infrastructure were basic, as the organisation focused on development of its high-performance computing facility and expertise in numerical weather prediction.

With the growth of modelling capabilities and increasing collaboration with Member States and other meteorological centres, networking infrastructure became a strategic priority early on. Local-area networks within the data centre, and wide-area networks connecting ECMWF to its global partners, have needed to upgrade in turn to match these demands.

When ECMWF produced its first operational medium-range forecast in 1979, it was receiving input observation data in raw GTS form, copied to magnetic tape by the UK Meteorological Office in Bracknell and brought across by car to the Centre. By the late 1980s, connections to facilitate data exchange with national meteorological and hydrological services (NMHSs) were in place using X.25 packet-switching technology, the standard for wide-area networks at the time.

These early networks served two critical functions: collecting observational data from weather stations across Europe and beyond, and distributing forecast products to Member States. Data transmission speeds were extremely modest by modern standards, operating at 9.6 to 64 kilobits per second but, typically for ECMWF, represented cutting-edge technology for the era.

The 1990s marked a period of digital expansion, coinciding with the rise of the Internet and more sophisticated networking standards. ECMWF began to transition to TCP/IP-based networking (Transmission Control Protocol/Internet Protocol), enabling more dynamic and reliable connections. During this period, the Centre developed its Meteorological Archival and Retrieval System (MARS) (Raoult, 1997), which necessitated faster and more scalable data access for research data transfers alongside operational outputs. Collaborations with organisations like the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) also led to increasing volumes of satellite data being acquired and processed within the data centre.

The end of the 1990s witnessed an important development in ECMWF's networking infrastructure with the establishment of the Regional Meteorological Data Communication Network (RMDCN) in collaboration with the World Meteorological Organization (WMO). This service provides a managed, high-availability, private communication infrastructure between ECMWF and the NMHSs of its Member and Co-operating States, and WMO weather centres globally.

Before the RMDCN, the NMHSs exchanged observations using individual dedicated leased lines between each pair of organisations that required it. This resulted in hundreds of international leased lines being deployed at a high cost for the Member States. ECMWF also had at least one international leased line per destination for dissemination purposes. The RMDCN introduced the concept of a single managed network to replace all these individual links, providing a more reliable service and huge financial savings for Member States. A market survey showed that there was no single telecommunications contractor with a presence in all the required countries. The idea then came to build on top of the infrastructure already deployed by civil aviation for interconnecting European airports. This resulted in the contract being awarded to SITA (Société Internationale de Télécommunications Aéronautiques) with the different local public telecommunications operators (PTTs) providing the last mile connectivity between the NMHS home sites and their closest airport.

■ **Figure 6:** The Regional Meteorological Data Communication Network (RMDCN) connections at meteorological organisations around the world. The blue colour indicates ECMWF Member and Co-operating States.



As forecast models grew in complexity and output resolution, the need for faster and more scalable data centre networking became paramount. The HPC facility replacement cycles provided a natural cadence at which to also upgrade, ensuring the facilities remained free of bottlenecks and able to make best use of HPC resources. In the 2010s, ECMWF implemented 10, 40, and later 100 Gbps internal networks within its data centre to support HPC and storage systems, with 400 Gbps now on the horizon.

ECMWF has long been at the forefront of computer networking to support its high-performance computing and forecast delivery. In 2022, ECMWF began transitioning part of its operations to a new data centre in Bologna, Italy, featuring state-of-the-art networking and computing capabilities.

WMO Members now consider the Internet is sufficiently robust and mature to support safety-critical applications – so long as those applications are built in a fault-tolerant way. Many NMHSs around the world already use the Internet to support all or most of their time-critical and safety-related data exchange. ECMWF’s own Internet connections now exceed 100 Gbps in aggregate.

Data exchange has come a long way in the decades since the late 1970s, from using magnetic tape and physical delivery, to today’s time-critical product dissemination and popular Internet-based web services. Challenges for the future are not only in performance and reliability, but increasingly security. With a growing number of publicly accessible web services, ECMWF, like most organisations on the Internet, must face the growing information and cyber security demands of our current age.

**DESIGNING DATA CENTRES:
FROM READING TO
BOLOGNA – A CHRONOLOGY
OF KEY MILESTONES →**

ECMWF has maintained world-leading supercomputing facilities for decades. The evolution of its data centre infrastructure – from the early days at Shinfield Park in Reading to the advanced systems now operational in Bologna – illustrates a continuous trajectory of innovation, resiliency and efficiency in support of meteorological research and operations.

1970s–1980s: foundations at Shinfield Park – early design and infrastructure

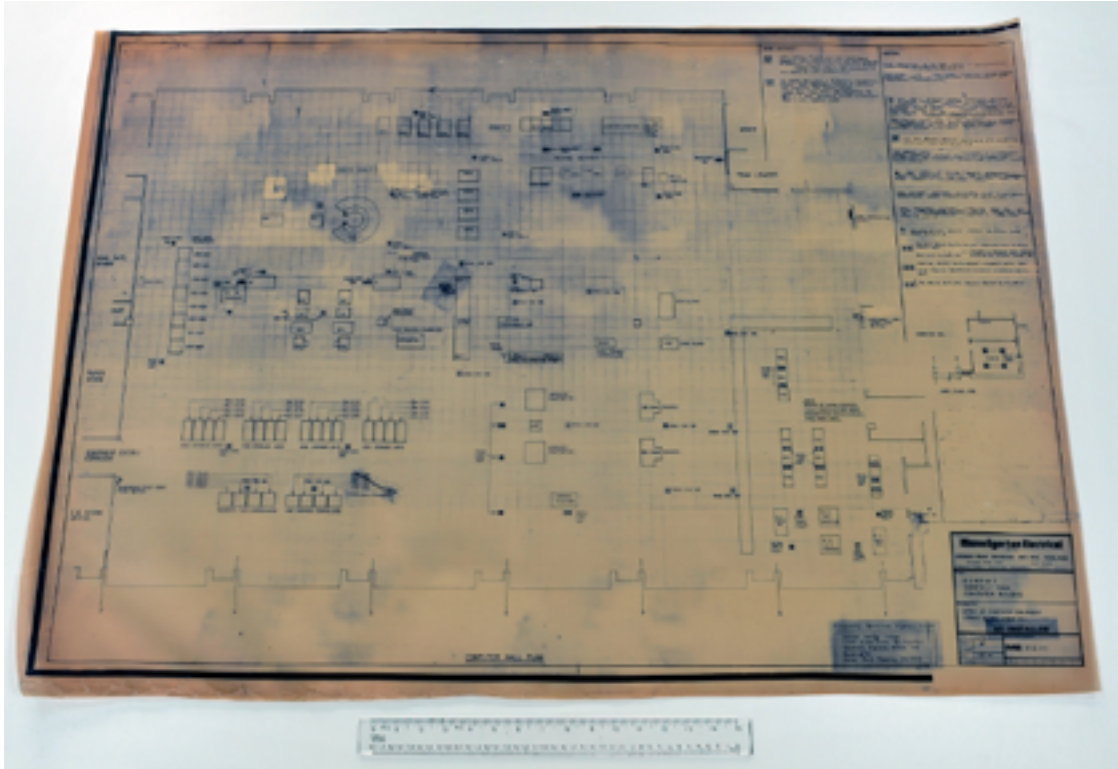
When the initial designs for Shinfield Park were developed, the site was provisioned with a 2000 kVA electrical supply. Original engineering drawings were created manually, using pens on drafting film, reflecting the standards of the time (see Figure 7).

The installation of the Cray-1A supercomputer posed unique challenges, as it was incompatible with the UK’s standard voltage and frequency. To support its operation, an innovative engineering solution implemented a bespoke power delivery system, providing:

- 120 V three-phase at 400 Hz
- 120 V three-phase at 60 Hz and 50 Hz
- Standard UK voltage of 415/240 V for ancillary systems

Although not considered power-hungry by modern standards, the Cray-1A required 288 kW of power. Cooling was provided via three cooling towers at the rear of the site, supplemented by four chillers that also serviced the adjacent conference facilities. Environmental control was maintained by computer room air handling (CRAH) units positioned along the north and south walls which regulated the computer hall’s internal climate.

■ **Figure 7:** The original layout of the computer hall with the Cray-1, dated October 1979.



1992: introduction of the energy centre

To boost power reliability, the energy centre was constructed and two diesel rotary uninterruptible power supply (DRUPS) systems—”KS1” and “KS2”—were installed, each delivering 1,000 kVA. These systems shared a 2000 kVA transformer and provided electrical continuity for mission-critical operations.

1996: transition to Fujitsu

The Centre transitioned from Cray to vector-based Fujitsu systems in 1996, requiring further electrical reconfiguration. These new machines specified different power requirements, necessitating transformer installations to adjust the supply voltage. Additional CRAH units were introduced to manage the increased thermal load in the computer hall, due also to the move away from water cooling to air cooling.

2002: beginning of the IBM era

In 2002, IBM systems were introduced, operating at 415 V/50 Hz, with expanded cooling demands prompting investment by the Centre in additional chiller capacity.

2004: enhanced fire protection systems

- Responding to insurance mandates, multiple layers of fire protection were implemented:
- A water mist (FogTec) system was installed in offices and corridors surrounding the computer building, offering high-pressure mist-based suppression.
 - Within the computer halls, a Very Early Smoke Detection Apparatus (VESDA) was installed. This system used pipes in ceiling voids to draw air samples into a detection chamber, where lasers could identify microscopic smoke particles – providing ultra-sensitive early fire detection.
 - A gas-based fire suppression system using Inergen was installed in both the room and floor voids. On activation, gas was released through nozzles while wall vents opened to reduce oxygen levels, effectively extinguishing combustion. In the basement, diverter valves and 132 bottles of Inergen (stored at 300 bar) ensured system readiness.

Additionally, two more DRUPS units (KS3 and KS4) were installed in the standby house, increasing redundant power capacity to 4,800 kW.

2006–2010: computer hall expansion and upgrades

An extension to the computer hall was completed in 2006, adding 500 m² of data centre floor space. One new cooling loop was installed to support eight additional CRAH units. At the beginning of 2008, the existing high-voltage switchboard was replaced, by adding a temporary board outside and installing the new board in the old one’s place with all the services maintained on temporary cables. New HV transformers and cables were added to increase the site’s mains capacity and in 2010, the legacy DRUPS units KS1 and KS2 were decommissioned and replaced by KS5 and KS6, bringing the total uninterruptible power supply capacity of the site to 5,600 kW.

2008: return to water cooling with Power6 (High Density Computing)

The installation of the first IBM Power6 cluster marked a return to water-cooled computing. Each Power6 rack featured a redundant water-to-water heat exchanger

and dedicated pumps circulating IBM’s proprietary “Blue” cooling water directly to the chips.

To support this infrastructure specification, the following upgrades were performed:

- Three new chillers were installed on the far side of the building.
- As there was no direct access, a 3-metre-square tunnel was bored beneath the existing buildings to link to the new extension.
- Independent piping and pumping systems were installed, creating a completely isolated cooling loop for the Power6 systems.

2013: sustainability – free cooling implementation

The dry air cooler system was commissioned in 2013 to provide energy-efficient “free cooling”. This allowed heated water from the IBM Power7 HPC systems to be cooled by ambient air through large radiator fans, dramatically reducing the reliance on mechanical chillers. As long as the return water temperature was at least 2°C above the ambient air, up to 2 MW of cooling could be achieved without using chillers.

As external temperatures approached the water temperature, fan speeds increased and three-way valves began to divert cooler water from the chiller loop to maintain temperature stability. This system recouped the £1 million construction cost within just 18 months of operation.

During this period, Chiller 4 was also upgraded – its screw compressor was replaced with a high-efficiency turbo core unit, adding an extra 350 kW of cooling capacity.

2014: water quality and cooling standards

In 2014, stricter water quality standards introduced by Cray necessitated a redesign of the cooling systems. Until then, a common water loop had been used for both computing and facility cooling. To meet the new requirements:

- Heat exchangers and dedicated pumps were installed to separate the loops.
- A full dynamic flush of the 70,000-litre system was conducted without interrupting operations, after water samples had revealed high microbial contamination, caused by the existing anti-corrosion chemicals feeding microbial growth. The system was then treated with a new passivation chemical.
- After flushing, biocide treatments were introduced, and a continuous monitoring programme was established to ensure proper water chemistry.

2020s: Bologna data centre

Purpose built to support the Centre’s expanding operational and research workloads, the Bologna data centre combines leading energy efficiency, modular design, and climate-conscious infrastructure. The advanced multi-modal cooling construction strategies included:

- Groundwater wells
- Adiabatic cooling systems
- Backup mechanical chillers

At the heart of its efficiency is the multi-modal cooling system that adapts dynamically to ambient conditions. It leverages groundwater from deep wells and adiabatic cooling – where water is misted into air streams to enhance the cooling without mechanical chillers. This significantly reduces energy consumption and carbon footprint, especially during cooler months. When necessary, high-efficiency backup chillers maintain performance during peak load or adverse climate conditions.

Power reliability is ensured by a 10 MW DRUPS system, providing seamless transitions during outages without relying on traditional battery banks. Fire safety mirrors proven systems at ECMWF’s Shinfield Park site, using inert gas suppression for zero-damage protection of high-value equipment.

In both architecture and operation, the Bologna data centre reflects the Centre’s long-term commitment to environmental responsibility, operational resilience, and scalable computing in support of critical global weather and climate research.

The evolution from manually drawn schematics and air-cooled Cray systems to water-cooled supercomputers and sustainable cooling strategies showcases ECMWF’s dedication to staying at the technological forefront. With the Bologna data centre, ECMWF is well positioned to meet future computational and environmental challenges in the decades ahead.

CURATING 50 YEARS OF DATA: A TIMELINE OF STORAGE SERVICES →

■ **Figure 8:** Growth of data in the ECMWF archive 1978–2025 (logarithmic scale). In 1978, the archive held 50 gigabytes of data. At the end of June 2025, it held 1,300 petabytes, an increase of 27 million times.

1970s–1990s: the early evolution of ECMWF data storage

Numerical weather forecasting requires data, and consequently collecting and storing meteorological data is one of the objectives of ECMWF that was set out in the Convention establishing ECMWF. The Centre has seen transformative growth in data storage capabilities since its inception. Starting with just 50 gigabytes of storage in 1978, the archive had reached 1,300 petabytes by the end of June 2025 – a staggering 27-million-fold increase (see Figure 8).



This exponential growth has paralleled advancements in high-performance computing (HPC), being symbiotically connected: HPC generates massive datasets, and the storage infrastructure preserves selected data for later retrieval by scientists. HPC systems are periodically replaced every five to seven years. Data, on the other hand, are perpetual, so the storing and protection are expensive and complex. Data need to be migrated when a new technology becomes available to improve capacities, performance and reliability and to control costs.

■ **Figure 9:** A Control Data tape drive.

While disk and solid-state storage systems have provided the fast access to data, the bulk storage has always been magnetic tape. The challenge of tape is that while it provides excellent bulk storage capacity, it needs to be loaded on to a drive and accessed by passing the tape over the read and write heads. This process can be fast, but it is strictly linear, starting at the beginning and winding to the end of the used section of tape in order to write new data, or seeking the portion of the tape that contains the data of interest. The Scientific Advisory Committee in 1978 noted: “It is thought that the Centre will be able to cope for 3 or 4 years after the beginning of the operational phase by using high density magnetic tapes (and staging these to disk when necessary to assemble particular required files of data), taking into account the compression methods that will be used and the cuts in the archiving of processed data. However, such an organisation of the Centre’s data bank, based on magnetic tapes in the way suggested, evidently cannot be a long-term policy.” In the event, this model of data staged to disk from tape is exactly what ECMWF has continued to use to manage the growth and complexity.

The Centre initially employed Control Data 669-2 open-reel tape drives (see Figure 9). With 1,600 characters per inch (25 mm), each reel held around 40 MB across 2,400 feet (730 m) of tape, which required manual mounting, severely restricting scalability. By 1982, the system had grown to circa 1,000 gigabytes of data using approximately 15,000 tapes with 10 magnetic tape drives. However, the service provision had become cumbersome and human-intensive to support. Racks of tapes are shown in Figure 10.



■ **Figure 10:** Racked tapes (background) and disk drives (foreground) in the Reading data centre. The blue cases on top of the disk drives were the outside packages for the removable pack, which was put into the drive unit below.



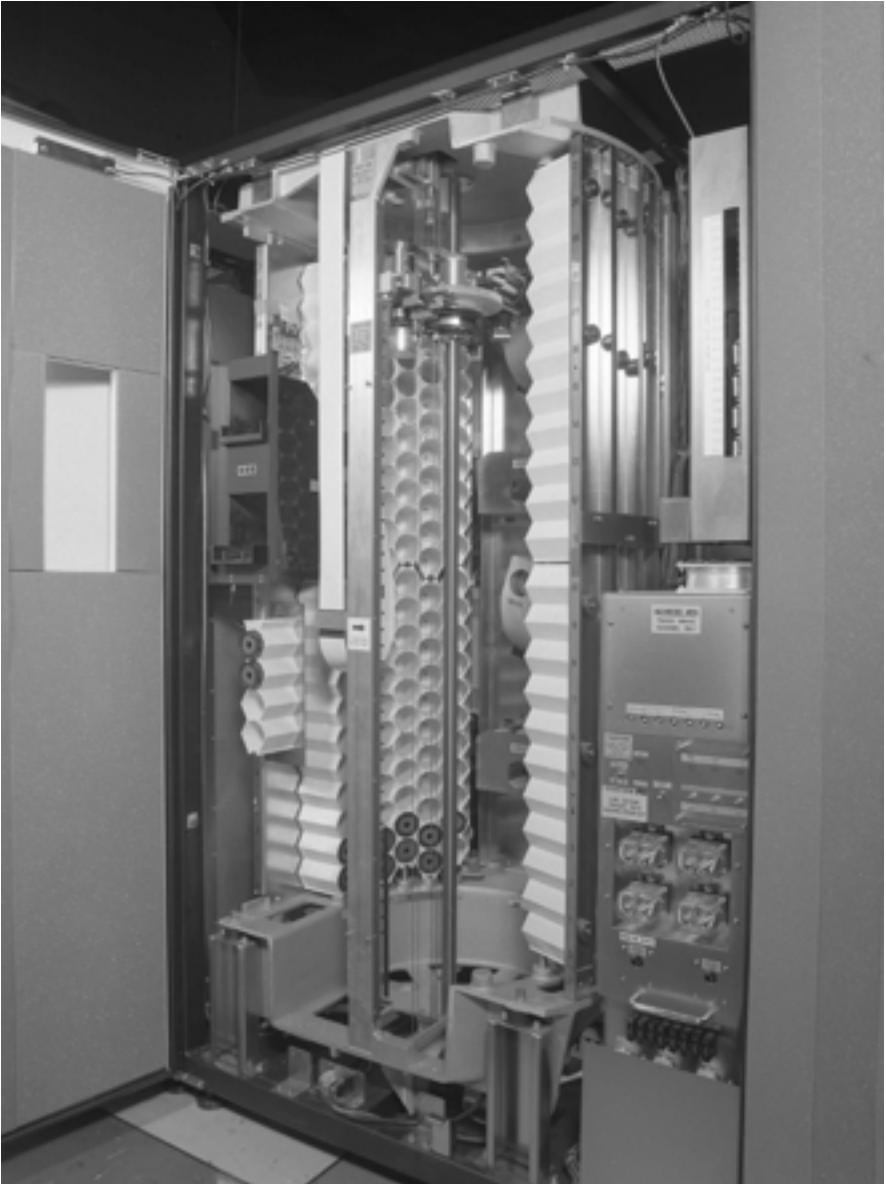
An IBM 3850 Mass Storage System (MSS) was installed at the Centre during 1983 and 1984. The Mass Storage System had a very unusual format. A group of cylindrical small cartridges, 50 mm wide and 100 mm long, each holding a 20 m spool of tape were stored in a hexagonal storage array (see Figure 11). Cartridges were moved by motorized accessor arms (see Figure 12). Each cartridge could store around 50 MB of data. Data was transferred to and from a host through IBM 3350 disk drives. This virtualised the MSS to the host, staging and de-staging data to cartridges in the background. It was the first example of “nearline” storage and pre-dated the Virtualised Tape Libraries (VTLs) that were to follow.

Around 1984, IBM introduced a family of single-reel cartridge tapes. The much smaller data cartridges (102 mm × 127 mm × 25 mm) allowed far more automation. The tape drive now threaded the tape out of the cartridge and over the read/write heads of the drive. ECMWF procured an 8-drive cartridge tape system in 1987, and by 1990 this system had 10 drives and 21,000 cartridge tapes.



■ **Figure 11:** IBM 3850 Cartridge – the first use of helical scan reading and writing technology for magnetic tape.

■ **Figure 12:** IBM 3850 Mass Storage System (MSS). © UKRI Science and Technology Facilities Council, available from <https://www.chilton-computing.org.uk/>



While the cartridge tape significantly increased the capacity of the archive and made handling the tapes much easier, it was not until 1991 that an automated tape library system was procured.

1980s–2020s: creation and development of the Data Handling System

From the early days of the Centre, it was recognised that it was strategically important to address the challenge of the ever-increasing amount of meteorological data. A Data Handling System (DHS) was created in three phases, to be the foundation of the service. A tender in 1982 resulted in the deployment of a Common File System (CFS) developed by Los Alamos National Laboratory (LANL), implemented on an IBM 4341 mainframe supported by IBM disk and tape.

As storage demands outgrew CFS scaling capabilities in the early 1990s, IBM’s ADSM (later Tivoli Storage Manager, TSM) replaced it.

Later, in a similar manner, it became clear that the TSM would not scale well into the future, and in 2002 a competition was run to find a replacement. HPSS (High Performance Storage System, a collaboration between IBM and a number of US Department of Energy laboratories) was chosen as the successor. Migration to HPSS started in earnest in 2003 with MARS and then the ECMWF File Storage (ECFS) using it in production in 2005. It provided a much better platform to scale out, where I/O movers could be added without affecting the performance of other areas of the storage. Some of the original structural decisions made in 2003 still hold today. The Centre will hold a new tender process later in 2025. However, where before it was clear a replacement was necessary, this time HPSS is still a viable option, scaling to the projected capacities into the future.

A key part of managing ECMWF’s storage involves layering the services that users see on top of the automation systems, decoupling the complexity of managing the actual data storage from the use of the data. By using this design, the system is independent from the underlying hardware and software. The data can be physically re-organised without any impact on the system, allowing complex data management to be undertaken.

The Meteorological Archival and Retrieval System (MARS) was created in the early 1980s. It was developed by ECMWF to enable users to retrieve meteorological data via simple “pseudo-meteorological language” queries. It automatically identifies whether data resides on disk or tape, supporting seamless migrations and system upgrades. From managing 70 MB/day in 1985, MARS scaled to over 550 TB/day by 2025.

ECFILE and later ECFS was introduced to provide a file system interface to the users, again abstracting the details of the storage implementation from the user.

In 2022 the DHS was moved en masse to the new data centre in Bologna. Some new components – disks, Storage Area Networks (SAN) tape libraries and servers – were built ahead of time. Most of the existing equipment and tape media were packed up into flight cases and moved by road to Italy. This was done in several stages, but mostly within a concentrated four-week period. During this time, the UK-based disk systems, servers and tape libraries were rebuilt in the two data storage halls at the Bologna data centre. This was done without losing the operational MARS and ECFS services. Both had been primed with data to continue operations using a pure disk-only environment. Within three days HPSS was back and able to cope with the write workloads from the HPC. Over the following month, the remaining tape and disk data from the UK were reintroduced to HPSS to make the entire archive available again.

1990s–2020s: tape evolution, and automation

By 1991, ECMWF had introduced a phased deployment of automated StorageTek 4400 libraries – robotic systems managing up to 6,000 tape cartridges per silo (see Figure 13). The DHS progressively grew to contain five silos. Over time, the tape drives evolved from IBM 3480 (200 MB tapes) through various IBM release versions to the TS1130 drives in 2007, maintaining form factor compatibility but increasing tape density to 1 TB – providing a 5,000-fold increase in capacity in the same physical space.

In 2009, SL8500 libraries from Sun Microsystems (later Oracle) replaced these older StorageTek libraries, scaling capacity to 10,000 tapes. More importantly, the design allowed for up to 64 tape drives in each library, a necessary feature to support the particularly heavy access patterns of ECMWF. Between 2010 and 2015, tape capacity grew from 1 TB to 5 TB to 8.5 TB.

■ **Figure 13:** StorageTek 4400 nearline tape libraries (later rebranded to Powderhorn libraries).



Oracle’s withdrawal from the tape market required ECMWF to pivot to IBM TS4500 and Spectra Logic Tfinity libraries (see Figure 14). ECMWF now operates 14 tape libraries with 740 enterprise-class tape drives and 80 LTO drives. These libraries manage nearly 1.275 exabytes across primary and secondary storage, showcasing both immense scale and complexity.

The Disaster Recovery System (DRS) also evolved, initially using Sony AIT tapes in a Grau AML/J library. This later moved to LTO drives within IBM TS3500 and TS4500 libraries. These systems serve as secondary storage for critical data, adding resilience and redundancy to ECMWF’s storage architecture.

■ **Figure 14:** IBM tape libraries in Bologna.
Photographer: Stefano Marzoli



WHAT WILL THE FUTURE BRING? →

2000s–2020s: disk storage, migrations, and looking ahead

The evolution of disk storage mirrored changes in tape systems. Early Serial Storage Architecture (SSA)-based RAID arrays gave way to Fibre Channel (FC) Storage Area Networks (SANs) attached systems. The first fibre-channel-based systems included IBM FASiT and DS series arrays, which later transitioned to V7000 and XIV systems. These newer models offered better performance and reliability.

From the mid-2010s until 2022, ECMWF deployed DDN SFA7700X and SFA14000 systems. After resolving a few initial implementation challenges these contributed to the delivery of a strong reliable service, only being retired when operations migrated to Bologna. The transition of infrastructure to Bologna enabled consolidation of the SAN resources, reducing complexity and improving manageability.

Migrations between storage systems – from CFS to ADSM/TSM and now to HPSS (High Performance Storage System) – have been monumental tasks. Each migration took over a year, requiring dual system operation and modifications to MARS and ECFS interfaces. A paradigm shift occurred in 2017 when HPSS moved from AIX to Linux, and the workload was distributed across over 100 x86-based Linux servers, compared to just 16 AIX servers previously.

MARS and ECFS have presented consistent interfaces for decades, while the underlying systems have continually evolved to ensure usability, performance, reliability, and scalability. Technologies like Ceph for distributed disk caching are being integrated, and future innovations, especially novel storage media, are being tracked to ensure the long-term future of the data archive.

The design of the next-generation HPC facilities must support an increasingly diverse mix of applications, from the traditional physics-based simulations and data assimilation to machine learning data-driven workflows. At the core of these systems will be heterogeneous compute architectures, combining general-purpose CPUs with high-throughput accelerators such as GPUs or AI-specific processors (e.g. Tensor processing units (TPUs), Graphcore Intelligence Processing Units (IPUs)). Partitions will be optimised for different workloads using flexible resource allocation to support simulation, data assimilation and machine learning inference and training tasks. These technology advancements will enable more accurate forecasts, global-to-local modelling, and faster scientific discovery, ultimately improving the ability to predict extreme weather events and long-term climate change.

By 2027, exascale computing will be widely available, enabling weather model ensembles to run at finer resolutions (4 km to 1 km scales) and supporting large-scale ensemble simulations to quantify uncertainty. AI accelerators will be deeply integrated into research computing platforms, enabling co-execution of machine learning models and traditional numerical solvers. These hybrid modelling systems will be critical for tasks such as emulating direct observation predictions, improving data assimilation cycles for near-real-time forecasting, training of ML models, and accelerating model runtimes.

To accommodate increasing demands for flexibility and collaboration, next-generation HPC systems will support cloud-native services and federated platforms. Containerised applications, workflow orchestration, and elastic scaling will allow ECMWF researchers and Member States to deploy and share models across national HPC centres and research clouds. This will foster a more collaborative and interoperable ecosystem, enhancing Europe’s capacity to respond

to climate and weather-related challenges with agility. These advances will also make it easier for users from diverse backgrounds – including policy, industry, and civil protection – to access and benefit from advanced weather forecasting tools and insights.

CONCLUSION

Over the past 50 years, the transformation of high-performance computing and data storage has revolutionised medium-range weather forecasting, enabling meteorologists to deliver more accurate, timely, and reliable predictions. From the early days of mainframe computing and magnetic tape storage to today’s exascale, modular, cloud-enabled, AI-driven supercomputing environments, technological advancements have continuously pushed the boundaries of weather prediction.

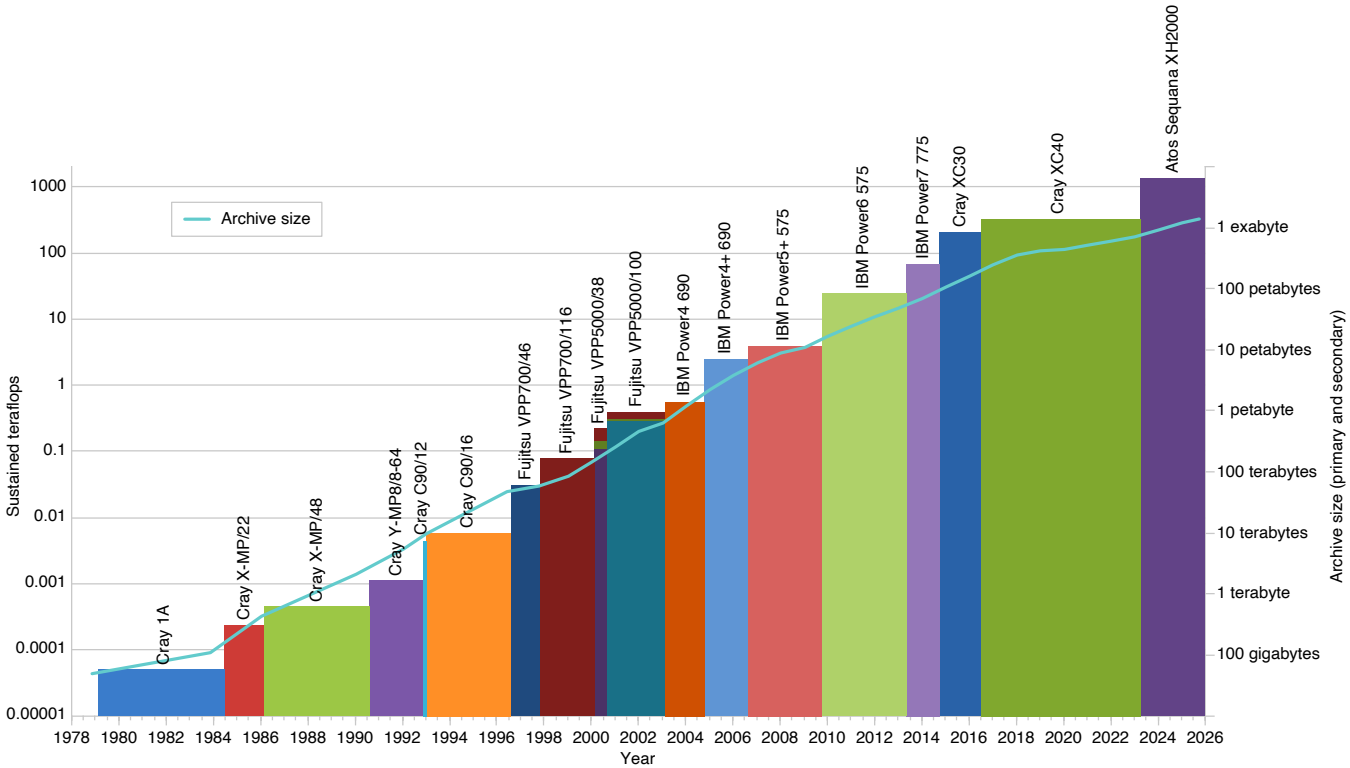


Figure 15: Evolution of ECMWF sustained HPC performance over the years with the data archive trajectory.

The shift from vector-based supercomputers to massively parallel architectures in the 1990s and 2000s allowed for higher-resolution models and the introduction of ensemble forecasting, which significantly improved uncertainty estimation. The ECMWF HPC service has increased in computational power by almost 30 million times from the original Cray-1 service in the 1970s (see Figure 15). Meanwhile, advances in distributed file systems and exascale storage technologies have enabled the handling of vast meteorological datasets, critical for data assimilation and model verification. For example, MARS has scaled from managing 70 MB/day in 1985 to over 550 TB/day in 2025.

These developments required substantial enhancements to the data centre environment to ensure the resilience and capacity to host these increasingly complex services – from mainframes that required specialised cooling through to tackling challenges related to creating carbon-neutral, sustainable data centres using renewable energy and latest technology advancements.

During the last decade, computing has witnessed major disruptive change: the emergence of commercial cloud computing has introduced flexibility and scalability, and graphical processing units have accelerated the growth of artificial intelligence and machine learning platforms. Such changes are transforming weather forecasting. AI-driven models can generate forecasts at lower computational cost, introducing the possibility of hybrid forecasting techniques which integrate traditional physics-based models with data-driven approaches. Service platforms now need to facilitate collaborative research and real-time model updates, provide cost-efficient resource allocation, and support meteorological institutions to adapt to growing data demands.

In conclusion, in just five decades, ECMWF has grown from renting time on a shared machine to operating one of Europe’s powerful, time-critical operations supercomputing facilities. As data-driven science gains prominence, the Centre’s next frontier lies in harmonising traditional physical modelling with machine learning applications, while working in a federated manner across Member and Co-operating States and partner institutions to share resources, expertise, and infrastructure in a truly collaborative European research ecosystem.

CONTRIBUTORS

ECMWF would like to thank the contributors to this paper:

Martin Palkovič, Christine Kitchen, Michael Hawkins, Stephen Richards, Andrew Gundry, Oliver Gorwits, Ricardo Correa.



Fifty years of meeting users’ needs

ABSTRACT



At the forefront of meteorological science and technology, ECMWF is a user-driven organisation.

Over the past 50 years, ECMWF has developed a comprehensive ecosystem of products and services, grounded in its core strengths: research in Earth system modelling and use of observations, cutting-edge technology and infrastructure, advanced data science, and, crucially, always through deep collaboration with its Member and Co-operating States.

With our Member and Co-operating States remaining at the core of our user community, the last 50 years have seen our user base broaden and diversify significantly to cover a range of research, academic and governmental bodies, as well as the private sector. This evolution reflects a growing range of meteorological, climate and environmental services, driven partly in response to the increasing risks from high-impact weather and climate change.

Engaging with our user community, ECMWF has responded over time to meet these needs in several ways: by continuously improving forecast quality, through an evolving range of fit-for-purpose products, and by enhancing access to data and information. Furthermore, supporting our users underpins all that we do, not least through an increasingly comprehensive range of support services and training.

This paper reflects on that journey, a defining feature of which has been, and continues to be, the two-way interaction between ECMWF and its users, with Member States playing a central role in shaping ECMWF’s direction.



INTRODUCTION →

At the forefront of meteorological science and technology, ECMWF is a user-driven organisation. Over the past 50 years, ECMWF has developed a comprehensive ecosystem of products and services, grounded in its core strengths: research in Earth system modelling and use of observations, cutting-edge technology and infrastructure, advanced data science, and deep collaboration with its Member and Co-operating States.

When ECMWF was established in 1975, its mandate was clear: to provide its Member States with high-quality medium-range atmospheric forecasts, to make meteorological data available on an operational basis for their purposes, and to offer advanced training and computing resources for research in numerical weather prediction (NWP). Since then, ECMWF’s portfolio has expanded significantly, keeping pace with scientific and technological progress (which ECMWF and its Member States have themselves helped to drive) and continuously responding to evolving user needs. With the amendment of the Convention in 2010, ECMWF’s mission was expanded to encompass the monitoring of the Earth system, clearly highlighting reanalysis as a key aspect of ECMWF’s offerings.

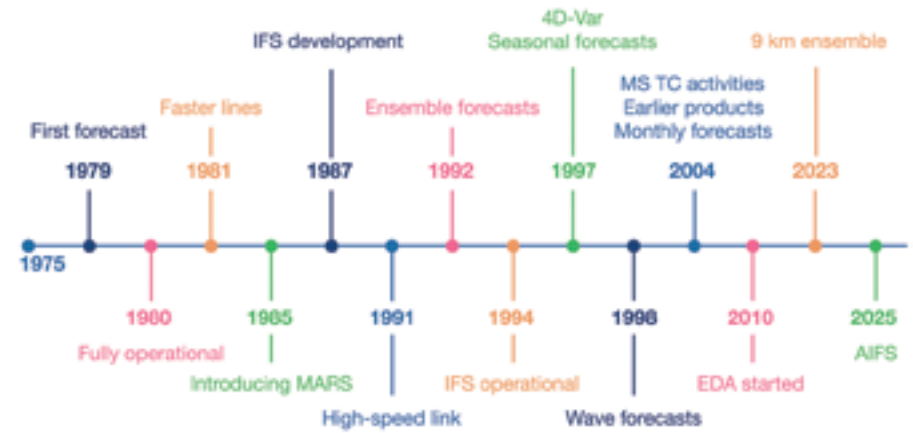
The user community has also transformed. Initially almost exclusively made up of the national meteorological services of Member States, and still very much centred on these, it now spans a wide spectrum: academia and research, governmental organisations in Member States, satellite agencies such as EUMETSAT and ESA, other European and international organisations, other WMO National Meteorological and Hydrological Services (NMHSs), and an increasingly diverse private sector engaged in weather, climate, and environmental services. This growth reflects broader shifts in the accessibility of meteorological data and the rising demand for reliable information in the face of climate change and growing weather-related risks.

ECMWF’s response to this evolution has been multifaceted. First, it has continuously improved its forecasting systems and the quality of the forecasts. Second, it has developed fit-for-purpose products that enable users to make decisions and create value from ECMWF forecasts and analyses. Third, it has ensured easy and reliable access to these products and the underlying data, allowing seamless integration into users’ workflows. Fourth, it has empowered users through dedicated support, training, and services. Finally, it has fostered a culture of dialogue and co-development, with regular workshops, meetings and exchanges, ensuring that ECMWF’s services evolve hand-in-hand with user requirements.

FORECAST AND REANALYSIS PRODUCTS ACROSS TIMESCALES →

At the heart of ECMWF’s ecosystem of forecast products are the ensemble-based configurations across multiple timescales – medium-range, sub-seasonal and seasonal. While the detailed evolution of the forecasting system is discussed elsewhere, it is worth recalling that the introduction of ensemble-based forecasts in 1992 marked a fundamental shift in quantifying uncertainty in weather forecasting. This was complemented by the launch of seasonal forecasting in 1997, which also established a multi-model capability through the EUROSIP initiative. This initiative laid the groundwork for the seasonal multi-system forecast within the Copernicus Climate Change Service (C3S) (see the ECMWF 50th anniversary paper on Copernicus at ECMWF). The sub-seasonal (monthly) forecast system, introduced in 2004, bridged the gap between medium- and long-range prediction (see Figure 1).

■ **Figure 1:** A timeline of key forecasting developments at ECMWF.



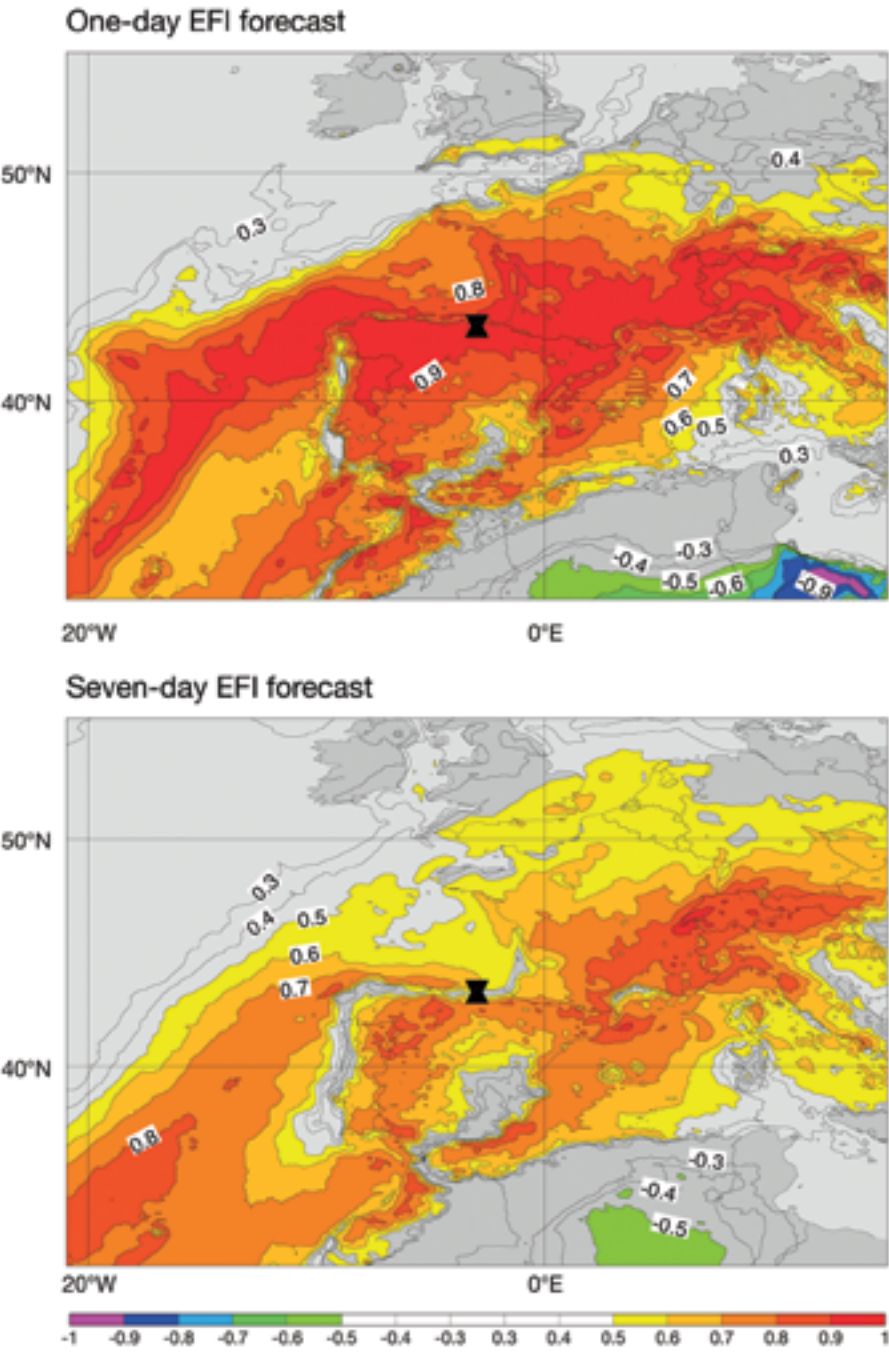
ECMWF products are now routinely used in the operational forecasting chains of Member and Co-operating States, as well as in downstream applications across various sectors. Forecast data serve as inputs for limited-area models (LAMs) and drive a wide range of models, including ocean, wave, atmospheric composition, hydrology, dispersion, and oil-spill forecasting (Hewson and Chevallier, 2024). Such applications also include environmental forecast products provided by ECMWF in the context of Copernicus, such as those offered through the Copernicus Atmosphere Monitoring Service (CAMS) and the Copernicus Emergency Management Service (CEMS) (see the ECMWF 50th anniversary paper on Copernicus at ECMWF). In parallel, the need for calibrated products from the ensemble forecast has led to the delivery of re-forecast (hindcast) sets for each configuration, used to estimate a “model climate” or assess forecast skill.

With the increased focus on supporting Member States in forecasting severe weather events, ECMWF has built derived products to enable forecasters to extract the appropriate information. The overall philosophy was that ECMWF should provide generic products to support forecasters, helping them extract relevant information from the ensemble to inform their decision-making, particularly for early warnings and impact-based forecasting. When ensemble forecasts were first introduced, a limited number of products was available. This included “stamp” maps, displaying all ensemble members’ solutions, probability maps, “plume diagrams” showing the evolution of a small number of surface parameters through the forecast range for selected locations, and cluster products. The product set has now evolved to encompass a wide range of data and products that provide information tailored to meet the needs of different users. Web-based platforms like ecCharts (Lamy-Thépaut et al., 2013) and OpenCharts have further expanded access, allowing forecasters to visualise and interpret ensemble forecast outputs interactively. They include diverse flavours of meteograms, from general surface weather parameters to visibility ranges and precipitation types, leveraging enhanced representation of physical processes in the Integrated Forecasting System (IFS).

A key success story is the Extreme Forecast Index (EFI, see Figure 2) (Lalaurette and van der Grijn, 2003; Zsótér, 2006), introduced in 2003. Developed following the two destructive 1999 storms (named Lothar and Martin), the EFI identifies regions where the ensemble forecasts predict unusually severe weather compared to the model’s own climatology, as assessed using re-forecasts over the past 20 years. It was complemented in 2010 with the Shift of Tails (SOT) index, designed to highlight when the most extreme values in the ensemble forecast shift significantly from the model’s climate extremes. Initially applied to wind, temperature, and

■ **Figure 2:** Extreme Forecast Index (EFI). The EFI for maximum 2-metre temperature on 23 August 2023 in a 1-day and 7-day forecast.

precipitation, EFI and SOT have been extended, based on user requirements, to new variables, such as snowfall, significant wave height, and convective indices, as well as to longer forecast ranges, including sub-seasonal forecasts. EFI verification is performed routinely against station observations. Since its inception, the EFI has been regularly mentioned by forecasters as one of the most popular ensemble products, particularly relevant for impact-based forecasting, as it focuses on anomalies relative to the local climate (as opposed to predefined thresholds).



ENSURING PRODUCT
QUALITY: CONTINUOUS
EVALUATION AND USER
FEEDBACK →

From the early days of its operations, ECMWF has systematically assessed the performance of its forecasting systems to inform both users and internal development. Forecast evaluation serves two essential purposes: providing users with insight into forecast skill and offering model developers diagnostic information to guide improvement. Together, these two functions create a feedback loop that ensures the forecasting system continues to evolve in line with both scientific advances and user needs.

Typically, an ensemble forecast should provide estimates of the probability of events which satisfy the criteria of reliability and sharpness. The forecast is considered reliable if, considering all past situations in which a given probability was forecast for a specific event, the event actually occurred with the same probability. The other desirable property of ensemble forecasts is sharpness, which is characterised by the accuracy of a probability distribution with a minimised width. Another area of interest is any day-to-day jumpiness in the forecasts, as it presents a challenge of interpretation for forecasters.

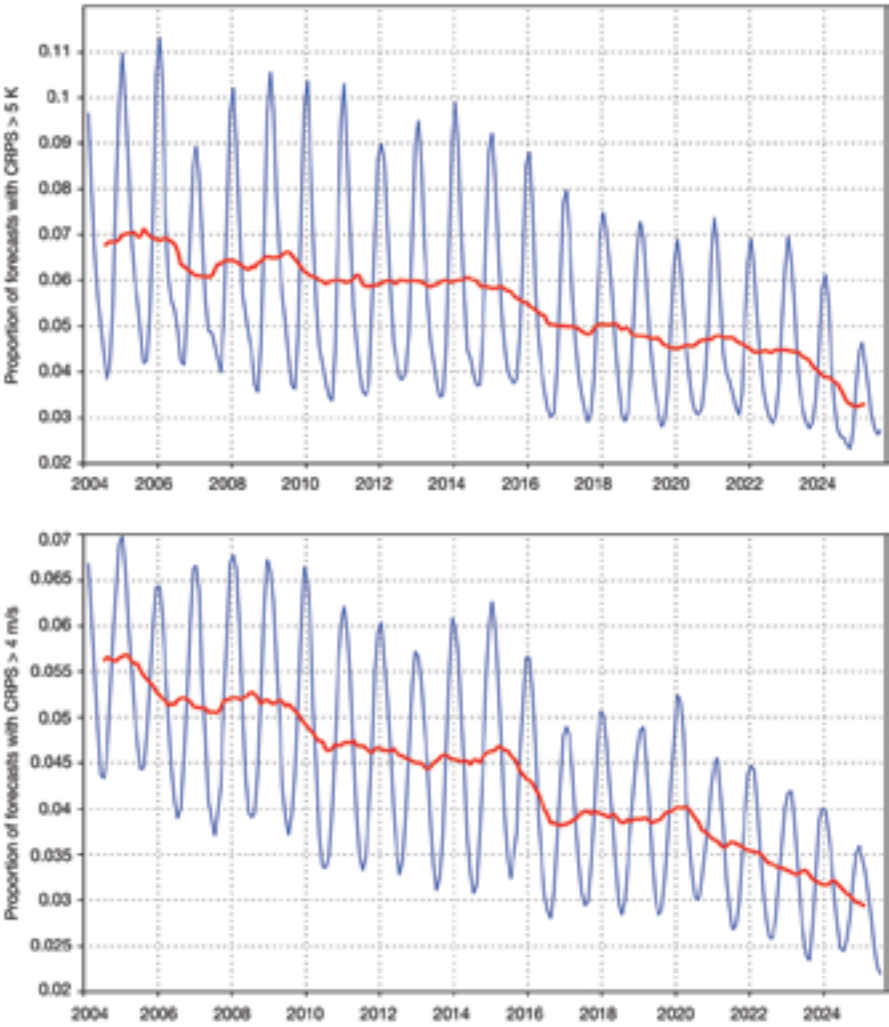
Since the early 1980s, ECMWF has published verification results to demonstrate the benefit of model upgrades and to compare configurations, including in early

Another example of successful product development across timescales is ECMWF’s suite of tropical cyclone products (Magnusson et al., 2021). They are based on the output of the cyclone tracker from Vitart et al. (2012). Tropical cyclone activity maps, including genesis, are produced for both medium-range and sub-seasonal forecasts. Basin-wide statistics of tropical storm frequency and accumulated cyclone energy are produced for sub-seasonal and seasonal forecasts. For medium-range forecasts, a combined product displays strike probability and track plumes, along with intensity along-track, for each active tropical cyclone (as reported at the initial forecast time by a competent designated centre). This product was enhanced in 2024/2025 to also display tracks from ECMWF’s Artificial Intelligence Forecasting System (AIFS) (both AIFS Single and AIFS ENS), as well as experimental data-driven deterministic models. These tropical cyclone products have become an essential part of early warning capabilities in cyclone-prone regions and are provided to all WMO NMHSs as part of ECMWF’s WMO Core data.

Collaboration with Member and Co-operating States has been central to product innovation. Many new products – such as point rainfall forecasts, the extratropical cyclone database, and the expansion of the EFI and meteograms – originated from user feedback in ECMWF Member and Co-operating States or through collaborative development projects. ECMWF regularly engages with its user base through workshops, surveys, and forums to ensure that the product portfolio meets users’ needs. Feedback from Member and Co-operating States on their use of ECMWF products is also reported regularly in the so-called “Green Book” (see for example, Hewson and Chevallier, 2024).

Finally, a valuable by-product of the IFS is the generation of observation-monitoring diagnostics from the data assimilation system. These statistics provide a near-real-time assessment of the quality and impact of ingested satellite and in situ observational data. Over time, this capability has evolved into an advanced alarm system that automatically detects anomalies and data outages, enhanced with machine-learning techniques. In spring 2024, ECMWF operationalised this through a web-based observation dashboard designed to offer both high-level “traffic light” status indicators for non-specialists and interactive drilldowns for experts. All these products are publicly provided to the wider user community.

■ **Figure 3:** Evolution of the fraction of large IFS-ENS 2 m temperature errors (CRPS > 5 K, top) and 10 m wind speed errors (CRPS > 4 m/s, bottom) at forecast day 5 in the extratropics. Verification is against SYNOP observations. 12-month running mean shown in red, 3-month running mean in blue.



Continuous work has been undertaken to enhance forecast verification. First, by defining more user-relevant measures, as exemplified by the introduction of supplementary headline scores for near-surface parameters. Then, by defining measures that are informative about model developments, particularly with increased resolution (double penalty effect – see Lledó et al., 2023). Verification of precipitation forecasts has been a specific focus, for example, stimulating concerted efforts among Member and Co-operating States to collect high-density observational data (Haiden and Duffy, 2016). Finally, with the increasing complexity

1 <https://www.ecmwf.int/en/forecasts/quality-our-forecasts>

DATA SERVICES:
FUTURE-PROOFING DATA
PROVISION THROUGH
FLEXIBLE CLOUD-BASED
SERVICES →

2 <https://confluence.ecmwf.int/display/FCST/Known+IFS+forecasting+issues>
3 <https://confluence.ecmwf.int/display/FCST/Severe+Event+Catalogue>

of the IFS and increased visibility of IFS outputs, it has become necessary to expand ECMWF’s verification capability to other Earth system components through collaborative efforts and leveraging third-party projects.

An annual report summarises recent verification results of ECMWF’s systems (e.g. Haiden et al., 2024), which are also presented in various fora. Member States also provide input and feedback on ECMWF forecast quality. This began in 1980 with an annual “meeting of forecasters”, established by ECMWF’s Technical Advisory Committee, to exchange views on the synoptic quality of the products and produce a summary assessment of forecast quality. Intelligence gained from user feedback is made accessible through the website, for instance, the known IFS forecasting issues.² Since 2014, the Severe Event Catalogue³ has displayed a collection of materials for specific weather events, focusing on meteorological and environmental forecast performance and including feedback and input from Member and Co-operating States. Verification results were also made available on OpenCharts.

Another cornerstone of ECMWF’s quality assurance process is daily in-house monitoring. Forecast performance is reviewed close to real time and reported each weekday, together with regular reports from forecasters in Member and Co-operating States. It is discussed weekly across the entire organisation in the internal Weather Discussion, a forum established in 2014 that supports rapid awareness of emerging issues. Quarterly meetings examine the performance of forecasting systems over the recent season. All these analyses trigger further investigation into identified model features or systematic errors, leading to immediate actions, informing priorities for future model improvements, and stimulating external collaborations. Diagnostic studies have consistently played a crucial role in understanding systematic errors, model sensitivities, and the impact of observation changes.

One of ECMWF’s objectives is to make the results of its research and operations available to Member States “in the most appropriate form”. In the late 2010s, ECMWF projected a sharp rise in data volumes, from approximately 100 terabytes per day to more than 1 petabyte by the mid-2020s, while only a third of the data was being disseminated. At the same time, cloud computing and big data technologies had matured, offering the ability to process data where they reside rather than transferring massive volumes.

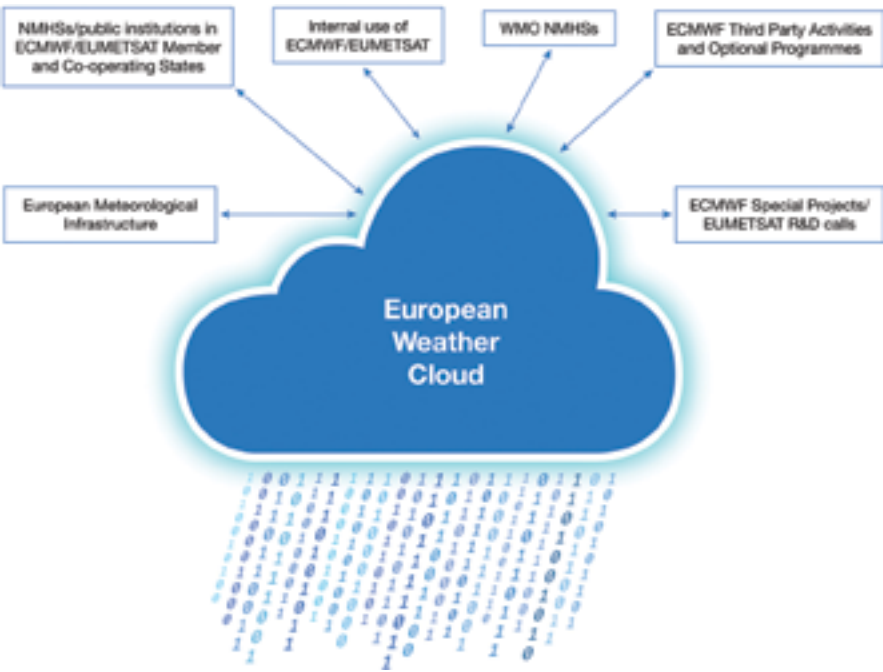
In December 2018, ECMWF’s Council approved a pilot project, in collaboration with EUMETSAT, to establish a federated cloud infrastructure known as the European Weather Cloud (EWC). Its aim was to bring users to the data, allow access to data holdings across the European Meteorological Infrastructure (EMI), and provide flexibility beyond ECMWF’s high-performance computing facility (HPCF) and ecgate services. The initiative began in 2019, and on 26 September 2023, the EWC was declared operational at ECMWF’s Bologna data centre.

The EWC is the cloud-computing-based collaboration platform for meteorological applications in Europe. It consolidates computing resources, data access, and tools in one place to support data exploitation and value-added services. Users can customise and deploy workflows, co-develop new weather and climate applications, and run them close to the data, avoiding costly transfers.

The service is available to eligible users from ECMWF and EUMETSAT Member and Co-operating States, their NMHSs and public institutions, research users through Special Projects and research and development calls, and EUMETNET activities

Figure 4: The European Weather Cloud service is available to different groups of users (Abellan et al., 2024).

relevant to ECMWF’s mission, ECMWF Optional Programmes and certain WMO-aligned activities (see Figure 4).



Because it runs in the same data centre, the EWC benefits from co-location with ECMWF’s broader infrastructure:

- The ECMWF Production Data Store (ECPDS) can deliver tailored data directly to EWC storage for immediate processing.
- The Meteorological Archival and Retrieval System (MARS) provides hundreds of petabytes of data with fast local access.
- The Copernicus Data Stores enable efficient API access to climate and atmospheric composition data.

The EWC complements ECMWF’s high-performance computing facility (HPCF). The HPCF is best suited for large-scale parallel workloads, while the EWC provides customisable environments, such as running public web services. Integrated workflows can link the two: models run on the HPCF, with outputs post-processed and visualised on the EWC.

Over one hundred projects have utilised the EWC, encompassing a range of technical and scientific applications. Examples include:

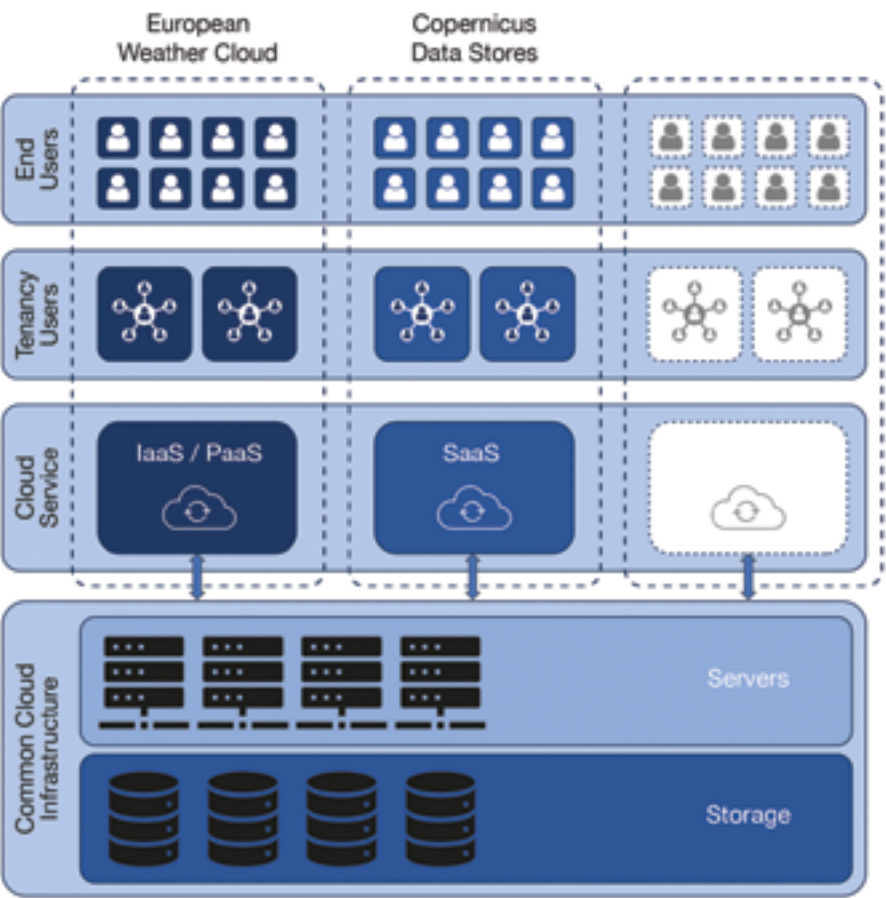
- Machine learning (ML) pilots: Member States and ECMWF collaborate on ML applications for data-driven forecasting, leveraging EWC graphical processing units (GPUs) and collaborative environments.
- International collaborations and system complementarity: a WMO multi-hazard early warning system (the South-East European Multi-hazard Early Warning Advisory System (SEE-MHEWS-A)), hosted at ECMWF. Data from the HPCF flows directly into the EWC, enabling efficient web-based services for participating NMHSs.

- Support for emergency response: in 2020, after an earthquake damaged the headquarters of the Croatian Meteorological and Hydrological Service (DHMZ), a backup system was deployed on the EWC within days, ensuring service continuity.
- System interoperability: whilst use of the EWC is primarily intended for ECMWF and EUMETSAT Member and Co-operating States, it is also available for WMO activities. Météo-France, one of the main contributors to the WMO Information System 2.0 (WIS2.0) (Fucile et al., 2023), used the EWC to host one of the core components of the WMO's new information system, showing how the EWC can interconnect with other cloud-based systems worldwide.

From pilot to operational service, the EWC has demonstrated its value in supporting diverse applications, fostering collaboration, and enabling innovation. Its growing user community benefits from tailored tools, workshops, and an active forum. ECMWF continues to improve the service in line with evolving needs, integrating it with wider initiatives such as the RODEO project⁴ and ML pilot activities. The EWC is now a key element of ECMWF's data strategy: making data accessible, actionable, and collaborative for the European meteorological community and beyond.

The EWC at ECMWF is hosted within the ECMWF Common Cloud Infrastructure (CCI). This is a cloud-computing-based IT infrastructure that hosts and serves multiple projects and services offered at ECMWF. The CCI has been established in ECMWF's data centre in Bologna, Italy, co-located with other ECMWF facilities and services, including the HPCF and the Data Handling System (DHS).

■ **Figure 5:** ECMWF's Common Cloud Infrastructure (CCI) supports the European Weather Cloud, the Copernicus Data Stores and other applications and services. IaaS, PaaS and SaaS stand for 'infrastructure as a service', 'platform as a service', and 'software as a service'.



⁴ <https://rodeo-project.eu/>

HIGH-PERFORMANCE COMPUTING SERVICES: ENABLING MEMBER STATE TIME-CRITICAL BIG DATA PROCESSING →

The CCI is also the home of the Climate Data Store (CDS) for the Copernicus Climate Change Service (C3S) and the Atmosphere Data Store (ADS) for the Copernicus Atmosphere Monitoring Service (CAMS): the main data storage system and backend for the two Copernicus operational services implemented by ECMWF on behalf of the EU. Figure 5 shows the relationship between the CCI, the EWC and the Copernicus Data Stores.

The CCI is divided into two production clouds, designated as CCI1 and CCI2. Each production cloud is hosted in a different computing hall at ECMWF's data centre for redundancy and resilience purposes.

OpenStack, a well-known open-source cloud computing software platform, is the technology used to manage these computing capabilities and expose them to various applications, services, and users. Backing the computing capabilities, approximately 5.5 PiB of usable hard-disk-drive-backed (HDD) storage and around 300 TiB of solid-state-drive-backed (SSD) storage are available on each of the two CCI cloud clusters. Ceph is the storage solution behind them, ensuring scalable and robust access to the data. This storage capacity is used for both classic block storage by the virtual infrastructure deployed in the cloud, as well as object storage using popular APIs, such as Amazon S3 and OpenStack Swift.

A key ECMWF objective, established in its Convention, is to make a share of its supercomputing resources available to Member States: at least 25% of HPC capacity and 10% of data storage. Since ECMWF's creation, these have been used for research and Special Projects (scientific or technical experiments of general interest).

In the early years, access to ECMWF computing was possible only via dedicated leased lines with limited bandwidth, mainly used for disseminating forecasts. The introduction of a dedicated Member State Unix system in the mid-1990s, along with Internet access, opened the way for applications to run directly at ECMWF, with outputs transferred more efficiently.

Some applications required near-real-time access to operational forecasts. To meet this need, ECMWF developed the mechanism of job submission under SMS (Supervisor Monitor Scheduler) control, enabling Member State jobs to run as soon as new data became available. By 2004, about 300 jobs from 60 users ran daily under this system, but monitoring and recovery from failures remained the users' responsibility.

As more complex requests arose (e.g. data for ocean wave forecasts, boundary conditions for deterministic and ensemble-based limited area models (such as COSMO-LEPS), ECMWF recognised the need for a formal framework. In 2004 the Council approved the framework for time-critical applications (Dando and Modigliani, 2025), offering three options:

- **TC-1:** simple time-critical job submission
- **TC-2:** Member State suites monitored by ECMWF
- **TC-3:** Member State suites managed by ECMWF

TC-1 replaced the earlier SMS mechanism. Supported by ECaccess (which gives registered users access to ECMWF's computing and archive facilities), jobs are triggered by "events" linked to ECMWF's operational schedule. Today, about 150 users run 1,800 jobs daily under TC-1, for tasks such as generating bespoke

products from MARS, running trajectory calculations, analysing ensembles, or producing ENS meteograms.

TC-2 supports complex workflows, such as running limited-area models (LAMs) with ECMWF boundary conditions. Suites have high-priority HPC access and are monitored by ECMWF’s 24/7 operations staff.

The first official TC-2 application, COSMO-LEPS (a limited area ensemble prediction system), was launched in July 2005, providing high-resolution ensemble forecasts of high-impact weather. Other TC-2 applications have included:

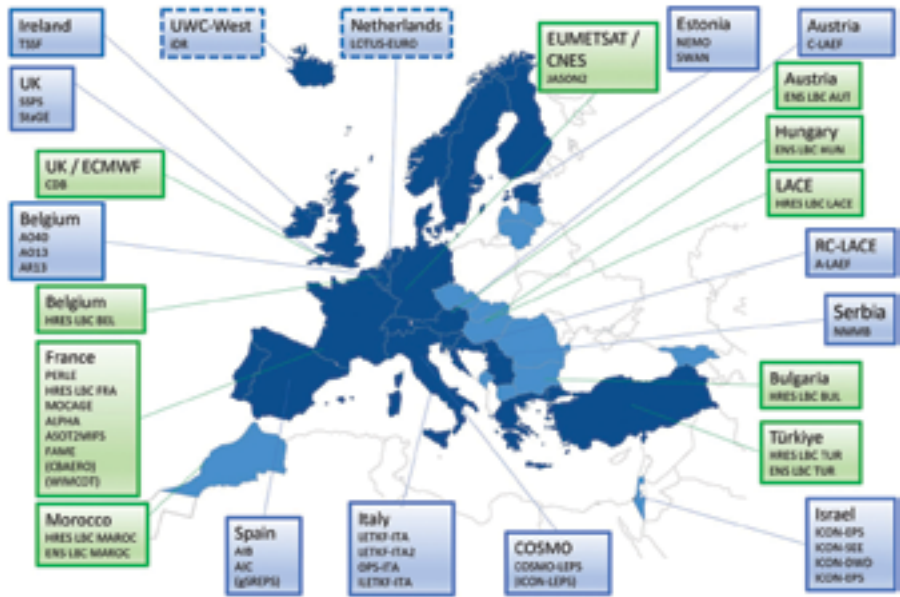
- The UK Met Office’s global ensemble (MOGREPS-15)
- The pan-European GLAMEPS multi-model ensemble
- Austria’s limited-area ensemble forecasting system (LAEF and later convection-permitting C-LAEF)
- The German National Meteorological Service’s Boundary Condition Ensemble Prediction System - Multi Model (BCEPS MuMo)

Several Member States (e.g. Portugal, Germany, Spain and Italy) also use TC-2 to run backups of operational NWP, ensuring continuity in the event of local HPC outages.

TC-3 provides the closest integration with ECMWF operations, with ECMWF managing Member State suites and testing them during IFS upgrades. A long-running TC-3 activity provides boundary conditions for LAMs (AROME, ALADIN, ALARO) and for the MOCAGE chemistry model, with output disseminated via ECPDS. Other TC-3 suites run for Météo-France and partners.

The time-critical framework has twice been critical in emergencies. In 2012, when the Danish Meteorological Institute’s HPC failed, ECMWF provided high-priority access, allowing them to continue NWP production. In 2020, following an earthquake that damaged DHMZ’s headquarters, its NWP system was quickly ported to ECMWF’s HPC, utilising TC framework priority access again.

■ **Figure 6:** TC-2 (in blue boxes) and TC-3 (in green boxes) suites on 1 September 2025, grouped by country, organisation or consortia. Those in boxes outlined with dotted lines or enclosed in round brackets denote new activities expected to be approved within the following 12 months.



Twenty years on, the TC framework remains central to how Member and Co-operating States use ECMWF’s HPC for time-critical activities. Around 2,000 time-critical jobs and suites now run daily (see Figure 6). A recent review considered improvements, including containerised workflows, virtual machines for ecFlow (ECMWF’s workflow package) servers, and synergies with the European Weather Cloud.

The TC framework has enabled Member States to run workflows they could not easily run themselves – such as LAMs, ensemble forecasts, and large-scale post-processing – delivered routinely and on time. It continues to provide significant benefits, underpinning both resilience and innovation, and will remain a cornerstone of ECMWF’s service to its Member States in the years ahead.

**USER SERVICES:
SUPPORTING AND
ENGAGING WITH USERS →**

Since its establishment, ECMWF has realised the importance of having a strong user support service. User support services are vital because they enhance productivity by helping users overcome challenges with products or systems, leading to greater job satisfaction and operational efficiency. They also benefit users by improving the user experience, making them feel supported and encouraging continued engagement.

This has been particularly relevant for ECMWF in assisting and advising users, especially in Member States, on the effective use of its products and services.

A User Support section was established almost immediately, and to facilitate communication with Member States, specific focal points, known as Computing Representatives, were created. Similarly, focal points in the User Support section were established to facilitate more direct contact with Computing Representatives. Regular meetings between ECMWF and these focal points have been organised, aiming to improve information flow in both directions and to facilitate the exchange of experiences among the focal points themselves. This allowed ECMWF to create a tight and effective community, which has enabled, over the last 50 years, access to services, their regular updates and improvements, and the introduction of new activities, such as the framework of time-critical applications or the European Weather Cloud described above. User support services have been a fundamental enabler of ECMWF’s success, allowing its Member States to greatly benefit from its products and services, and will continue to be so in the next 50 years.

Equally important is the two-way engagement with ECMWF’s user community. Forecast evaluation is enriched by user feedback, particularly from forecasters in Member and Co-operating States. Regular meetings with national meteorological representatives, liaison visits to Member and Co-operating States, the regular collection of input on “Use and Verification of ECMWF forecast products in Member and Co-operating States” (“the Green Book”), and the Using ECMWF’s Forecasts (UEF) event (see Figure 7) all contribute to keeping up to date and maintaining an active, ongoing dialogue. The biennial High Performance Computing in Meteorology workshop is also one of our regular engagement activities with the HPC community (see Figure 7). Through this collaboration, users not only share insight into the strengths and limitations of ECMWF forecasts but also shape the evolution and development of new products, ensuring they meet users’ needs.

■ **Figure 7:** Attendees of the 2025 edition of the Using ECMWF's Forecast (UEF) event (above) and the 21st High Performance Computing in Meteorology workshop (below), both held in Bologna 15–19 September 2025, as part of the 50th anniversary celebrations of ECMWF.



TRAINING AND OUTREACH →

Training and capacity building are central to ECMWF’s mission of fostering skilled and informed users. Our training programme spans advanced numerical weather prediction (NWP), ensemble forecasting, seasonal and sub-seasonal techniques, and cutting-edge domains such as machine learning (ML). Courses range from week-long in-person sessions to Massive Open Online Courses (MOOCs) and target diverse audiences: forecasters, scientists, data analysts, and computing specialists. These programmes are aligned with broader initiatives such as Copernicus services and Destination Earth.

ECMWF extends its educational impact through partnerships with the WMO, EUMETSAT, ESA, and national meteorological services, delivering joint training that deepens competence across the European meteorological community. Tailored course offerings are regularly co-organised under the Copernicus Climate Change Service (C3S) and Copernicus Atmosphere Monitoring Service (CAMS) National Collaboration Programmes, embedding training into national capacity-building efforts. These collaborative efforts extend beyond coursework to include workshops, seminars, and webinars – particularly important for onboarding new Member and Co-operating States and strengthening ties with global partners.

Looking ahead, ECMWF plans to further diversify and democratise its educational efforts by scaling-up online learning, launching new MOOCs on climate monitoring, expanding ML course offerings, and establishing new training facilities in Bonn. A deeper emphasis on “training by doing together” – wherein participants actively collaborate on real-world problems – highlights ECMWF’s aspiration to foster an EU-wide community of skilled users capable of harnessing advanced NWP tools in support of societal resilience.



CONCLUSION

Over the past 50 years, ECMWF has developed a comprehensive ecosystem of forecast and analysis products that spans multiple timescales and Earth system components. This ecosystem is underpinned by a versatile and robust operational infrastructure that enhances data accessibility, enabling users to explore, transform, and apply ECMWF data in diverse operational and research contexts. Collaborations with Member and Co-operating States, as well as EU programmes such as Copernicus and Destination Earth, research projects, and international partnerships (e.g. with the WMO) have played an instrumental role, acting as powerful multipliers of value and extending the reach and impact of ECMWF products and services.

Since 1 October 2025, all data in the ECMWF Catalogue of real-time products have been available under a CC-BY-4.0 open licence, with a core subset freely accessible at 0.25 degrees resolution and delivered without additional delays compared to operational availability. This milestone marks the start of a new era of open data at ECMWF, with further expansion planned for 2026 and beyond through a tiered approach that will gradually open higher-resolution products. By partnering with cloud providers and harmonising catalogues and APIs across ECMWF, Copernicus and Destination Earth, ECMWF aims to make data discovery and access more seamless, aligned with European data space initiatives. At the same time, user experience will be enhanced with streamlined documentation, integrated dashboards, and AI-powered support.

The open data policy strengthens ECMWF’s role in providing greater support to WMO national meteorological and hydrological services. This effort is further enabled by stronger relationships with WMO programmes, such as recent initiatives like the WMO’s Systematic Observations Financing Facility (SOFF). This also includes the recent European initiative “Strengthening Early Warning in Africa” under the Africa–EU Space Programme.

Beyond data access and computing services, ECMWF offers specialised software packages tailored to the needs of national meteorological services. With ECMWF’s robust software suite (now known as ECMWF software EnginE (ESEE)), ECMWF Member States have access to a powerful suite of software tools for efficiently retrieving, processing, and analysing the vast meteorological data archive maintained by ECMWF. Some of these software packages are already co-developed with Member States. This empowers Member States to not only leverage ECMWF’s data for their own forecasting models but also seamlessly integrate them with their national observational networks and research datasets.

Today, the user base of ECMWF is broader than ever before. Access to open and free products has brought ECMWF data to new audiences, while traditional users are still at the heart of our user base and continue to rely on ECMWF for critical operational applications. The number of active users across data services, web platforms, and training initiatives illustrates both the scale of the community and the trust placed in ECMWF outputs. This expansion highlights the growing demand for accessible data that can be applied across various sectors to anticipate high-impact weather events and support informed decision-making.

Such growth also brings new responsibilities. The pressure to respond to an increasing number of users with diverse and sometimes competing needs underscores the importance of staying focused while remaining collaborative.

Engagement with Member and Co-operating States, together with mechanisms for user feedback, has been – and will remain – instrumental in ensuring that ECMWF products meet the highest standards of quality, while being adapted to evolving requirements. Co-development, shared tools, and open dialogue are key to sustaining this balance.

Looking ahead, ECMWF is committed to further advancing its open data offering, supporting the joint development of a European meteorological data space, and integrating machine learning both as a source of new capability and as a tool to improve accessibility, interpretation, and delivery. The AI/ML revolution in weather and climate forecasting, described in the ECMWF 50th anniversary paper on machine learning, will certainly be an additional stimulus for the review of products and the means of providing them. It will prompt ECMWF and its Member and Co-operating States to reconsider their workflow and how to best integrate ECMWF forecast products and services. These efforts will ensure that ECMWF data remain not only scientifically excellent but also user-oriented, easily accessible, and actionable.

The story of ECMWF’s first 50 years has been one of doing things together: building world-leading forecast systems through shared infrastructure, co-developed tools, and close engagement with its community of users. The next 50 years will continue this journey, harnessing Earth system science, AI, technology, open data, and collaboration to maximise the societal impact of forecasts – helping more users, enabling and supporting them in the best way, and ensuring that the products developed together deliver lasting value to society.

CONTRIBUTORS

ECMWF would like to thank the contributors to this paper:

Matthieu Chevallier and Umberto Modigliani.

We also acknowledge and thank:

All colleagues across ECMWF and its Member and Co-operating States who, over the past 50 years, have contributed and still contribute to the developments described in this article and many more besides, to ensure ECMWF continues to meet users’ needs.



Machine learning for numerical weather prediction

ABSTRACT



Artificial intelligence (AI), and particularly machine learning (ML), is a mainstay of newspaper headlines, coffee conversations and everyday life across the world. Many fields and disciplines find themselves in the middle of a revolution, where the combination of data, algorithms and compute can provide low-cost solutions for a wide range of tasks. Weather forecasting is no exception to this, with this 50th ECMWF anniversary year seeing the operationalisation of the Artificial Intelligence Forecasting System (AIFS) at ECMWF. We will take a stroll through the history of machine learning at ECMWF, which starts earlier than one may expect, talk about the current state of play and gaze into a crystal ball in discussing the role of machine learning at ECMWF in the years to come.

AWAKENING: 2018–2022 →

Machine learning (ML) is the use of statistical algorithms that can learn from data and generalise to unseen data without explicit instructions. It is a sub-class of artificial intelligence (AI), which is the capability of computer systems to perform tasks typically associated with human intelligence. Many of the statistical methods that have been used for decades in Earth sciences can be counted into the wider class of ML. Examples are multi-dimensional linear regression, or dimensionality reduction via principal components. Even deep learning – the use of neural networks to perform ML – was already applied at ECMWF for the emulation of the radiation scheme more than two decades ago (Chevallier et al., 1998). It may therefore surprise that many claim that we have seen an ML “revolution” during the last couple of years. What happened?

Beyond the domain of Earth system science, AI and ML have seen an enormous rise that was mainly fuelled by:

- A massive increase in computational power with computer hardware customised towards the needs of deep learning, and deep learning being the ideal application for state-of-the-art supercomputers that excel for simple arithmetic and linear algebra.
- The exponential increase of data in many domains – including weather and climate – and the ability of deep learning to learn systems of arbitrary complexity if enough data and compute capacity are available.
- The availability of software libraries such as TensorFlow and PyTorch that allow a user to create complex deep learning architectures with very minimal Python code.
- The massive amount of experience that was collected on how to design efficient deep learning methods with new neural network architectures and training procedures being invented, including convolutional neural networks, recurrent neural networks, generative adversarial networks, attention and transformers, and diffusion networks.

It became more obvious around 2018 that the developments in general machine learning would also impact data assimilation and Earth system modelling. Early success stories across ECMWF’s workflow included the use of neural networks for SMOS soil moisture data assimilation for the land surface (Rodríguez-Fernández et al., 2019) and the use of neural networks for bias correction learned within the 4D-Var data assimilation framework (Bonavita and Laloyaux, 2020). Deep learning has been used successfully for the emulation of the gravity wave drag parametrization schemes (Chantry et al., 2021a), and the deep learning emulators could be used to generate tangent linear and adjoint model code for 4D-Var data assimilation (Hatfield et al., 2021). Furthermore, decision trees have been used for the post-processing of ensemble predictions for precipitation (Hewson and Pilloso, 2020), and there have been plenty of links and similarities between data assimilation and deep learning (Geer, 2021). As ECMWF hosts more than one exabyte of weather and climate data, there were plenty of possible application areas for versatile, scalable tools that allow the extraction of complex information from data – such as deep learning. The potential applications were distributed across all parts of the numerical weather prediction (NWP) pipeline, from observation processing (Dahoui, 2023) to data assimilation, to the forecast model and the post-processing and dissemination of the forecasts. These included *methods to improve our understanding of the Earth system* such as unsupervised learning and causal discovery, uncertainty quantification, and AI powered visualisation; *methods to speed-up conventional models* such as emulators for parametrization

■ **Figure 1:** Objectives for the ML activities at ECMWF as defined in the ML Roadmap in 2021. As described below, ECMWF has been very successful in following the objectives.



schemes including with low numerical precision, the optimisation of the high-performance computing (HPC) and data workflow, and data compression; *methods to improve the models* such as bias correction tools, tools for quality control of observations, feature detection algorithms, and the learning of model components from observations; and *methods that linked different datasets to weather and climate datasets* that have interesting applications for health, energy, transport and pollution applications, as well as for extremes such as wildfires or flooding.

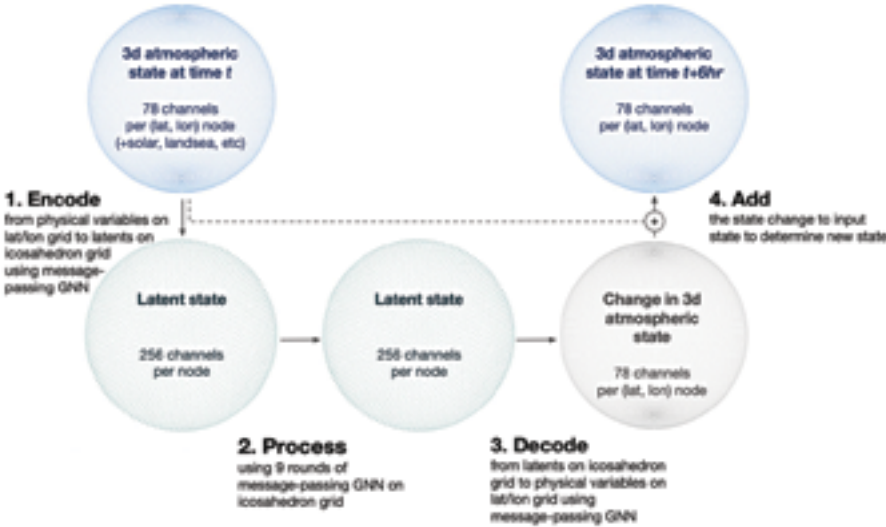
To bring a bit more structure into the multitude of applications that could be explored and to quickly develop the infrastructure and know-how that was needed to move quickly in the developments, ECMWF published a Machine Learning Roadmap for the next ten years in 2021 (Dueben et al., 2021). This roadmap outlined five objectives for the developments in machine learning at ECMWF (see Figure 1).

For the community to fully embrace machine learning required the opportunity for existing domain scientists to learn about the methodologies. Since its inception ECMWF has had a rich history of providing training. In 2022 a first course on machine learning for weather forecasting was run at ECMWF’s Reading site. This has since been repeated most years, with each course being significantly oversubscribed. The topics covered in this course have evolved each year, to cover fresh communities in the field. In 2023 ECMWF introduced a Massive Open Online Course (MOOC) for machine learning in weather and climate, designed to help en-masse the weather forecasting community engage with the topics. The MOOC had 40 hours of content and was a great success attracting more than 9,000 registered participants from 159 countries. It featured contributions from 60 experts around the globe, covering three tiers.

In the first years after the awakening of deep learning for weather and climate, the community mostly focused on “Soft AI” to allow for improvements in computing efficiency via emulation, or “Medium AI” that incorporated machine learning into physics-based models e.g. via learning within data assimilation (Chantry et al., 2021b). However, ECMWF has also written the first-ever paper for “Hard AI” for medium-range NWP. This work aimed to replace the entire forecast model based on a pure deep learning tool trained from reanalysis data (Dueben & Bauer, 2018). ECMWF has also contributed significantly to WeatherBench, the first benchmark dataset to train global weather prediction models from reanalysis data (Rasp et al., 2020). However, at ECMWF and other meteorological centres, “Hard AI” approaches were treated rather as a testbed and scientific playground for new neural network architectures and training mechanisms than as a serious alternative to physical models for operational predictions. This changed in 2022.

DISRUPTION:
2022 AND 2023 →

■ **Figure 2:** Reproduced from Keisler (2022). The figure shows a schematic for building a weather forecasting neural network. 1. The 3D state is encoded into a latent state on a coarser grid. 2. Information is then passed between nodes along edges in successive layers to calculate a latent estimate of the change in the atmospheric data. 3. This information is then decoded onto the state of atmospheric variables. 4. The information is added to the starting time state to give a prediction 6 hours into the future.



Perhaps even more arresting than the forecast skill were the energy costs to make a forecast. Once trained, the above systems could finish a ten-day forecast in a couple of minutes, on a single commercial-grade graphics processing unit (GPU) – to be contrasted with approximately 30 minutes on around 50 nodes for a forecast using the IFS. Roughly, these systems could reduce the energy costs of the forecasting piece of the chain by a factor of 1,000.

ADOPTION: 2023–2025 →

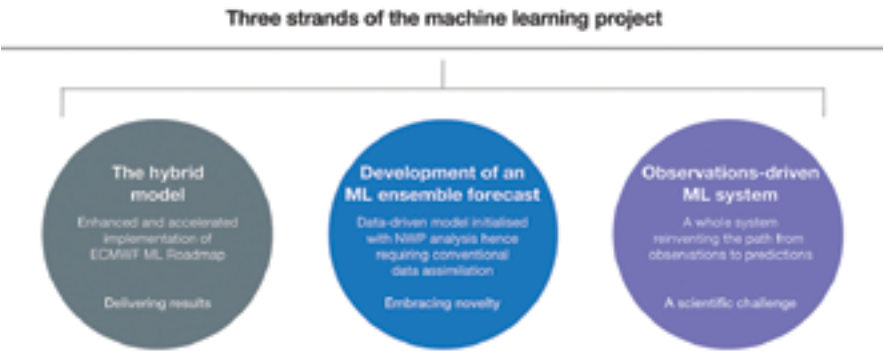
In the summer of 2023, with strong support from the ECMWF Council, ECMWF started the Machine Learning Project, a four-year project to embrace the disruption of machine learning and develop operational systems. This project has three strands (see Figure 3). The first focuses on hybrid combinations of machine learning and physics, as championed in the machine learning roadmap. The hybrid strand, already mature with elements in the operational pipeline, promises value to further improve the IFS. The second focuses on data-driven weather forecasting, trained on reanalysis and analysis data. Here the goal was to build on the scientific publications by first creating a system that matched the skill of these systems, before then aiming to be world-leading in this domain. One of the major targets was an operational ensemble system two years after the start of the project. The third strand is the use of observations to build forecasting systems, including forecasting directly from observations, i.e. building a system inclusive of data-assimilation and forecasting.

Common to all of these works were two aspects which draw sharp contrast with physics-based models and which are part of the explanation for the computational efficiency. One was the relatively minimal prognostic state of the atmosphere required for accurate forecasting. This featured approximately 13 pressure levels up to 50 hPa and only the basic atmospheric variables of humidity, winds, temperature and geopotential. This meant accurate forecasting without an explicit representation of clouds, and with a far coarser representation in the vertical dimension than the 137 levels used by the IFS. The second was the huge timestep, typically 6 hours, made by the models, in contrast to 9 minutes for the IFS. A timestep this large would not be numerically stable for 30 km resolution physics-based models using conventional timestepping schemes.

Due to the preprint nature of contributions, and the work stemming from outside of the meteorological community, there was natural uncertainty to these results. Several features threatened to undercut the validity of the results and warranted closer investigation. Two examples were the spatial smoothing of forecasts induced by optimising root-mean-square errors, and the use of ERA5 as initial conditions, meaning incorporation of fresher observations. Both features are known to artificially inflate skill. To explore this, in early 2023, ECMWF became the first centre to start running these models in real time from the operational ECMWF analysis and it also showcased plots of live forecasts to users on ECMWF’s open charts. The goal was to help the whole community explore and understand these systems. Through in-house scoring, and verification of case studies, it became quickly clear that the results broadly held up to this further scrutiny (Ben Bouallègue et al., 2024). ERA5 initial conditions and forecast smoothing play a part in the skill gains but did not explain away significant improvements in forecasting skill. The open charts were popular on social media, with experienced meteorologists exploring live case studies and generally finding favourable results.

Data-driven models were not a panacea. At that time, current challenges were estimation of small-scale extreme values, e.g. wind speeds in tropical cyclones, or intense small-scale precipitation. These challenges all had roots in the training approach. By minimising the mean-squared-error, models were not rewarded for making bold predictions for harder to predict events. The tool developed at ECMWF to enable this easy running of Pangu-Weather, FourCastNet, GraphCast and more from a single interface, named ai-models, was created as an open-source repository enabling the wider scientific community to run these models more easily and better understand their dynamics.

Figure 3: The three strands of the machine learning project, which started at ECMWF in the summer of 2023.



The first strand, hybridising the IFS with machine learning, explores many approaches to improve the skill and efficiency of the IFS. This includes learning properties of sea-ice state using deliberately minimal learning systems (Geer, 2024); bias correction of 4D-Var data assimilation (Bonavita and Laloyaux, 2020), which will feature in ERA6; and nudging the IFS towards a data-driven model, following Husain et al. (2025), aiming to combine the large-scale improvements of data-driven modelling with the enhanced small-scale details and expanded product set of the IFS.

Prototype work for the Artificial Intelligence Forecasting System (AIFS), began immediately. The design choice of graph neural networks was made for the first implementation. This was chosen due the flexibility of data grid choices and natural encoding of the spherical geometry of the Earth. Building on the previous works by Keiser and GraphCast, rapid progress was made, with a first real-time running system in place by October 2023. After several further experimental model cycles (Lang et al., 2024a), the AIFS Single – named to capture a system designed to produce a single trajectory – was implemented as an operational system in February 2025. AIFS Single 1 outperforms the IFS across the vast majority of scores, with tropical cyclone track accuracy a notable place of significant improvements in forecasting ability (see Figures 4 and 5 for examples). Tropical cyclone intensity estimates by contrast are a current weak point, with AIFS Single significantly underestimating intensity, comparable with other data-driven models. AIFS Single 1 was made fully open source, including both data and model, enabling anyone to easily run the forecasting system themselves. The ability of users to easily run models themselves without HPC systems or extensive HPC knowledge is another advantage of these data-driven systems.

Figure 4: Skill comparison of AIFS Single 1 and the IFS for temperature at 850 hPa in the northern hemisphere extratropics for the spring period MAM 2025.

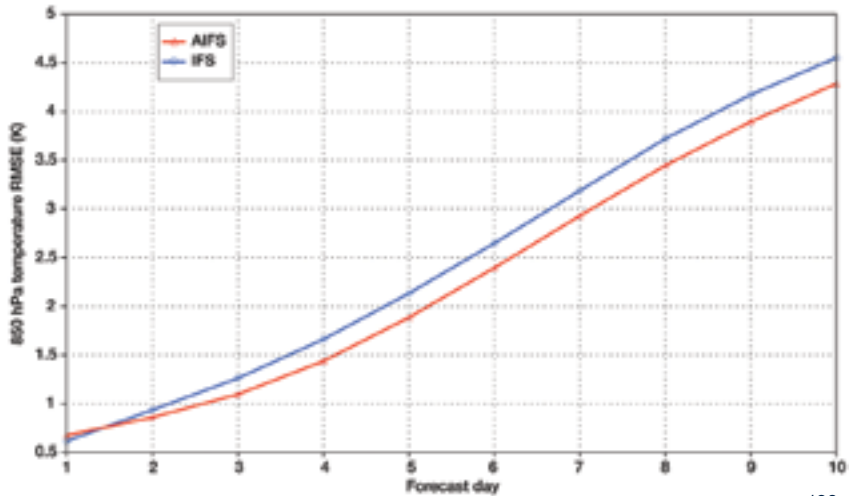
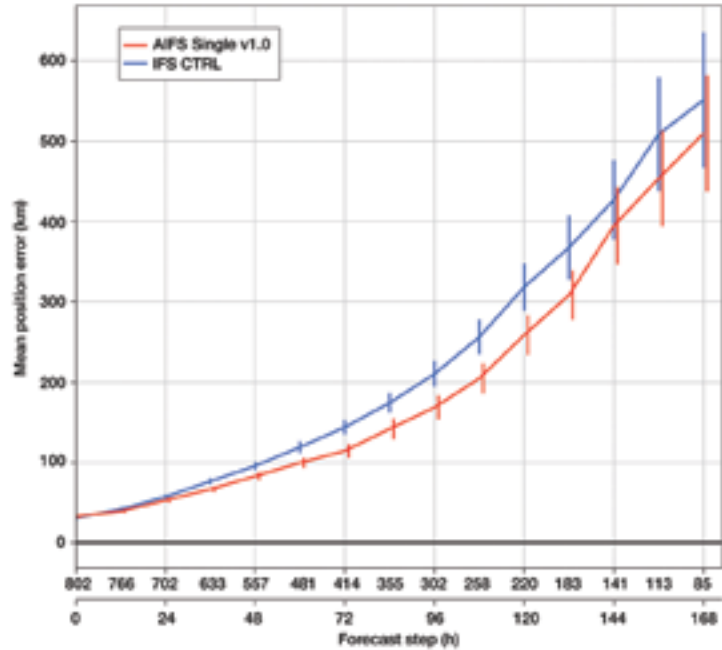


Figure 5: Skill comparison of AIFS Single 1 and the IFS for tropical cyclone track accuracy from July 2024 to June 2025.



Next came the ensemble. Two methodological strands were explored in development towards the first AIFS ensemble system. One used the diffusion training methodology, state-of-the-art in image and video generation ML systems. The second directly targeted optimising the Continuous Ranked Probability Score (CRPS, Lang et al., 2024b). This latter approach proved the more accurate and computationally cheaper. A first ensemble system, AIFS ENS 1, was implemented in July 2025, two years after the start of the ML project.

The CRPS-optimisation versions of the AIFS demonstrate skill not only on the medium-range timescale, but also at sub-seasonal timescales, outperforming the IFS sub-seasonal system across a number of key metrics. To engage the community in this developing field of sub-seasonal forecasting, ECMWF is organising the AI Weather Quest, a competition for sub-seasonal forecasting.¹ At the beginning of the event, 55 models had been entered across 33 teams and 14 countries. Three AIFS variants are being submitted to this competition.

The third strand investigates whether it is possible to encompass the full forecasting system, from observations to predictions, with machine learning. Following the success of machine learning for forecasting from analysis, this is a natural question to ask, one with wide reaching impact if true. A number of groups have engaged with research in this fascinating topic, with different problem framings being tested. Some, like the work of Allen et al. (2025), seek to utilise ERA5 in training, but still produce an end-to-end system without real-time dependencies on ERA5 or similar products. At ECMWF, a novel approach – AI-DOP (Direct Observation Prediction) – was proposed by McNally et al. (2024), seeking to only use observations in building an end-to-end system capable of forecasting future observations from current ones. GraphDOP (Alexe et al., 2024), a prototype of this approach, was created (see Figure 6), building on some of the work for the AIFS. This model showed that accurate forecasting was possible. However, currently the forecast skill of this work lags behind that of the IFS but is continually improving. Whether these works are the equivalent to the work of Keisler, showing promise without yet being state of the art, or whether machine learning fails to surpass physics-based data assimilation, we will learn in the coming years.

¹ <https://aiweatherquest.ecmwf.int/>

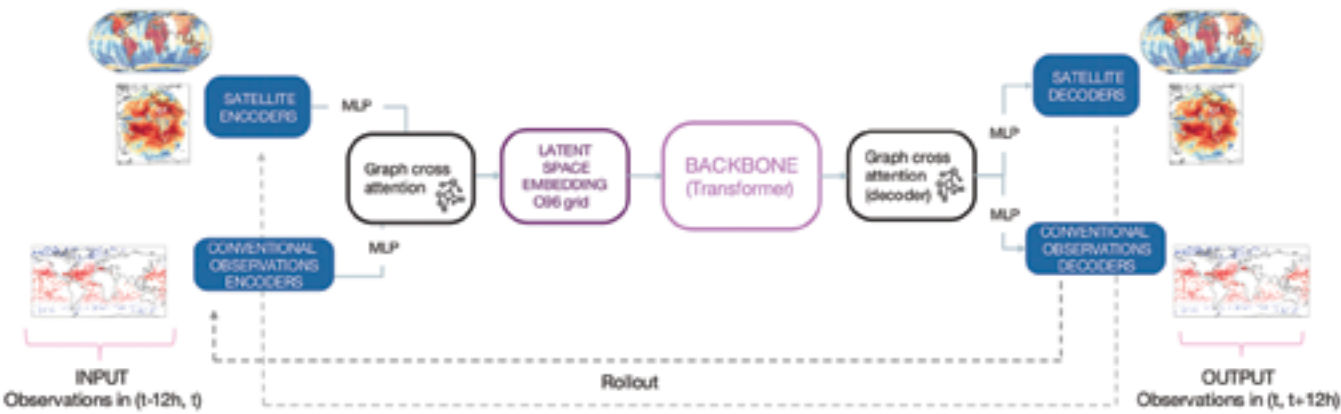


Figure 6: Schematic of the Graph-DOP approach developed in Alexe et al. (2024) which learns to predict future observations from current observations.

At the founding of the ECMWF ML project, opportunity was seen for ECMWF and its Member States to collaborate closely on the topic of data-driven weather forecasting. The ECMWF Member State ML Pilot Project was set up as a vehicle for organising this collaboration, featuring 14 partners across Europe at the time of writing, engaging in five work packages of data-driven modelling. ECMWF refactored the code underlying the AIFS to create a new open-source framework for data-driven modelling, dubbed Anemoi. This identified that the same code underlying a specific data-driven forecast system could also be used to develop global and regional forecasting systems for organisations across Europe, and this code could be co-developed by the European meteorological community, who could make rapid progress together. The first demonstration of this was Nipen et al. (2024), which built the first stretched-grid models, a forecasting system that featured higher spatial resolution over the Nordics but learnt from data around the globe. Anemoi was introduced in 2024 and in 2025 it won the European Meteorological Society (EMS) technology achievement award. The Anemoi community has grown to more than 12 Member States, who use and contribute to Anemoi.

OUTLOOK →

Due to their superiority in deterministic and ensemble forecast scores, the ML models will become the default tool for most applications in NWP. However, it is also unlikely that physical models will disappear from the operational portfolio in the foreseeable future. Physics-based systems currently provide a much wider range of products for users. They can also serve as backup model configurations if unprecedented events are happening (how would an ML forecast model represent the impact of a volcanic eruption on NWP?). Physical models are the prime tool to generate training data when observational datasets are sparse or inconsistent.

While ML currently lags behind physics-based systems for data assimilation, we view that it is likely that there will be an end-to-end ML forecast suite in the future, covering observation ingestion, data assimilation, the forecast model, post-processing and product generation.

One open question is how many different ML models will be trained and used in parallel for operational NWP – one seamless model for all global predictions, or many specialised models for specific predictions, for example for tropical cyclones?

As part of the EU Destination Earth initiative, ECMWF is already developing ML model components for the ocean, ocean waves, sea ice, land surface, and hydrology. ECMWF will, therefore, soon have a full machine-learned Earth system model. There are interesting questions about the coupling of model components

currently being explored – shall we couple ML models for the various components that have each been trained individually, or shall we train all components as a single model? A full machine-learned Earth system model that covers the coupled system will be much more useable for long-term and potentially even climate simulations when compared to the current AIFS NWP model.

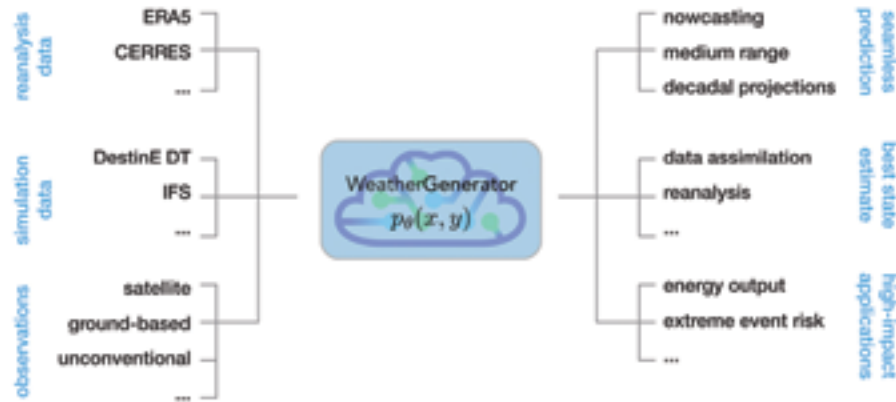
A topic of active discussion is the role for ML models on seasonal to climate timescales, a problem fundamentally about extrapolation into unseen climate regimes. Training on present day alone teaches a model to predict the current climatology, as shown in a recent study by ECMWF and Alfred Wegener Institute (AWI) scientists (Rackow et al., 2024). Work outside of ECMWF has shown that indefinitely stable and conservative data-driven models can be built (Watt-Meyer et al., 2024) and can produce skilful seasonal forecasts (Kent et al., 2025). The ability to create huge ensembles at low cost is an enticing one, particularly on the seasonal timescale. However, building ML models for long-term predictions is challenging due to the finite dataset length that needs to be split into a training and a sufficiently large validation dataset to construct statistically robust results. ECMWF is actively exploring this topic of data-driven seasonal forecasting.

Across the full spectrum of ECMWF activities, close thought is being given to consider the opportunities offered by machine learning. For atmospheric composition, interesting reanalysis datasets and preliminary results from Aurora (Bodnar et al., 2025) have prompted the exploratory development of an AIFS system for atmospheric composition. For reanalysis, the AI-DOP approach is being explored for the construction of reanalysis products.

The ability to make a forecast using a single GPU in just a couple of minutes opens new opportunities for the democratisation of weather forecasting. Within Destination Earth, ECMWF is developing a “forecast-in-a-box” prototype, which packages initial condition retrieval, forecasting, product generation and visualisation into a single portable unit, capable of running locally or on cloud facilities. Alongside MET Norway, ECMWF is working with national forecasters in Malawi to test and further develop this prototype.

Another question about the future of ML for Earth system modelling is whether the domain will follow the developments of large language models (LLMs) towards larger and more generic ML tools that can then be used for multiple application areas – so-called foundation models. LLMs are trained to fill in gaps in huge amounts of text, rather than for a specific task such as the translation from language A to language B. The resulting tools can be used for diverse tasks beyond their training objective, which include translations but also the almost instantaneous creation of a Shakespearean poem about TikTok cat videos. Along this line, it may be possible that a foundation model trained from various Earth system datasets and with a huge latent space with many billions of trainable parameters may perform better in certain tasks when compared with a task-specific model. To explore foundation models for weather and climate applications, ECMWF and a number of Member States have started the WeatherGenerator EU Horizon project that will build such a foundation model and serve as an additional digital twin for Destination Earth (see Figure 7).

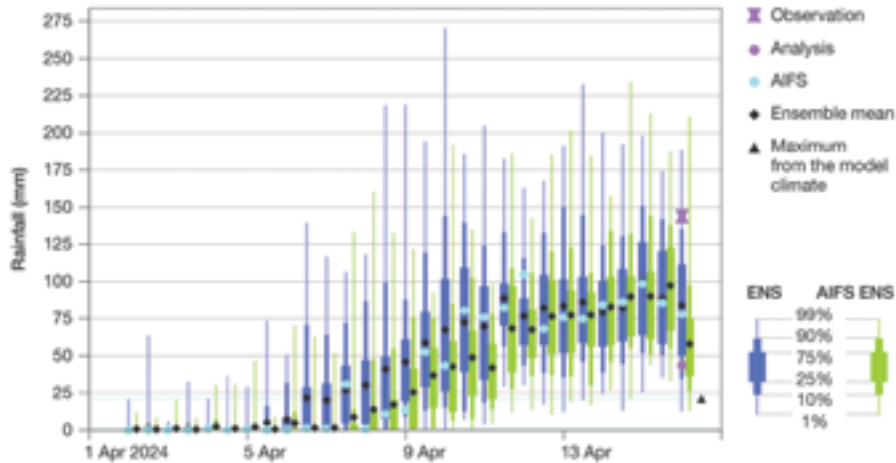
■ **Figure 7:** The WeatherGenerator will serve as foundation model for weather and climate applications. It will be able to digest datasets from various sources, including observations, local and global analysis products, and local and global models, and will be useful for many different application domains in weather and climate science. The input and output of the WeatherGenerator can be local or global, at a resolution between 100 m and 100 km, and can be in the past, present and future.



Data is obviously a key ingredient in accurate forecasting using physics-based approaches. We view that the role of data will grow even larger with the rise of data-driven forecasting. Curating large, calibrated datasets will be vital to feed these models. Direct incorporation of surface observations into model training and inference will increase the value of sharing these datasets. Novel data sources, e.g. cameras measuring visibility, could be directly included in model pipelines.

Representation of extreme, and particularly unprecedented, events remains a somewhat open question. Small-scale extreme events are typically not captured well in reanalysis and analysis datasets, and observation data sources for such events bring their own challenges. Unprecedented events can be categorised from a local or global perspective, i.e. events that have never been seen before in that region versus truly novel events that have never happened anywhere around the globe. Results so far suggest that for local extremes, machine learning models are able to surpass local climatologies significantly by learning to transfer lessons from other parts of the globe. Figure 8 shows an example of this for a case study in the UAE in April 2024. AIFS models confidently predicted record-breaking values well in advance of the event. For globally unprecedented events, researchers at the University of Chicago trained a data-driven forecasting system on ERA5 with all the strongest tropical cyclones removed from the training dataset (Sun et al., 2025). Without these events, the model was unable to produce these unseen strongest events. Further work is required to better understand the value for ML systems in extreme events. Live investigation through case studies will be vital to building trust.

■ **Figure 8:** Showing the evolution of 24-hour rainfall forecasts on 16 April 2024 over the grid box including Dubai. The model climate is about zero precipitation, with a maximum of less than 25 mm based on 1,800 forecasts (marked by the black triangle). Experimental versions of AIFS Single (blue dots) and AIFS Ensemble (green box and whisker) models both predicted precipitation values well-outside the model climatology and values in line with the IFS ensemble (blue box and whisker). All systems underestimated the observed value (purple hourglass).



Without the explicit underpinning of physics in data-driven systems, an increased emphasis is being placed on building trust. Across the wide user-base of ECMWF products, approaches for building trust will differ. For some, extensive verification will be the most important thing, for others case studies. Some users will prioritise physical consistency as an important facet. A holistic view across these dimensions will be important.

What is hopefully clear to all readers is that the world of numerical weather prediction, and ECMWF itself, is amid a revolution. The fundamentals of weather prediction are changing. Machine learning is bringing new opportunities and interesting scientific questions to be answered. Data-driven forecasting offers an opportunity for the meteorological community, particularly in Europe, to work even more closely and benefit from shared tooling without coalescing on a single model. The outside world, including large technology companies, will seize this opportunity, and if ECMWF wants to maintain its position in the community, continued agility and adaptability will be required. The next few years will be key for continued community building and answering the remaining fundamental questions captured above.

CONTRIBUTORS

ECMWF would like to thank the contributors to this paper:

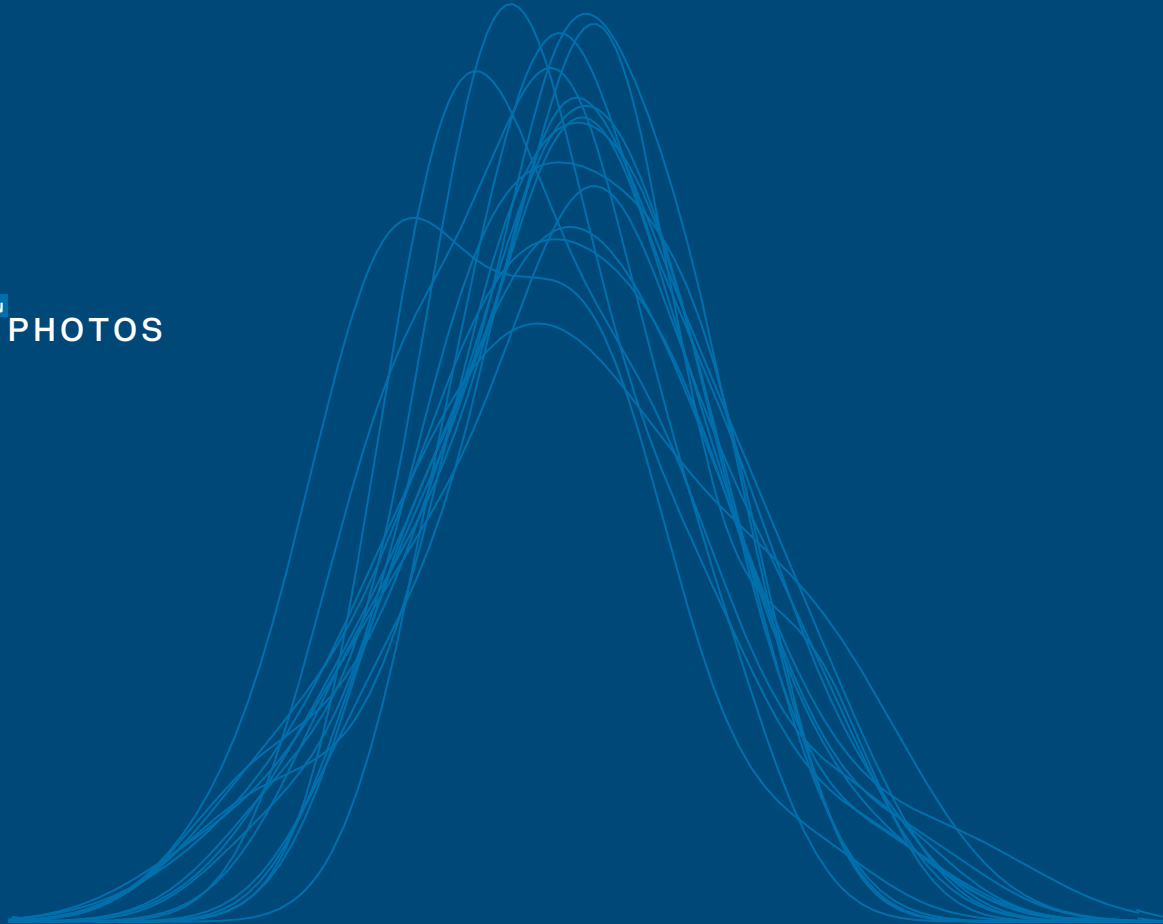
Matthew Chantry and Peter Dueben.

Also the following for providing figures:

Ryan Keisler (Brightband), Michael Meyer Gerber, Zied Ben Bouallègue, Mihai Alexe, Eulalie Boucher, Linus Magnusson and Christian Lessig.



PHOTOS



This section shows photos of our staff and premises, and Council representatives from our Member States.

Bologna



ECMWF staff in Bologna (Italy), February 2025.

Bonn



ECMWF staff in Bonn (Germany), June 2025.



Reading



ECMWF staff in Reading (UK), February 2025.

ARCHITECTS' IMPRESSIONS OF THE ECMWF BUILDINGS IN GERMANY, ITALY AND THE UK

BONN →



BOLOGNA →



READING →



THE THREE SITES IN 2025





MEMBER STATES 2025

Austria	Luxembourg
Belgium	The Netherlands
Croatia	Norway
Denmark	Portugal
Estonia	Serbia
Finland	Slovenia
France	Spain
Germany	Sweden
Greece	Switzerland
Iceland	Türkiye
Ireland	The United Kingdom
Italy	



ECMWF

ECMWF Council delegates, Reading, 2024

REFERENCES



This section contains the references for the papers presented earlier in the book.

FIFTY YEARS OF EARTH SYSTEM MODELLING AT ECMWF

Abramowitz, G., A. Ukkola, S. Hobeichi, J. Cranko Page, M. Lipson, M. G. De Kauwe, S. Green, C. Brenner, J. Frame, G. Nearing, M. Clark, M. Best, P. Anthoni, G. Arduini, S. Boussetta, S. Caldararu, K. Cho, M. Cuntz, D. Fairbairn, C. R. Ferguson, H. Kim, Y. Kim, J. Knauer, D. Lawrence, X. Luo, S. Malyshev, T. Nitta, J. Ogee, K. Oleson, C. Ottl , P. Peylin, P. de Rosnay, H. Rumbold, B. Su, N. Vuichard, A. P. Walker, X. Wang-Faivre, Y. Wang, & Y. Zeng, 2024: On the predictability of turbulent fluxes from land: PLUMBER2 MIP experimental description and preliminary results. *Biogeosciences*, 21, 5517–5538. <https://doi.org/10.5194/bg-21-5517-2024>

Agust -Panareda, A., S. Massart, F. Chevallier, S. Boussetta, G. Balsamo, A. Beljaars, P. Ciais, N. M. Deutscher, R. Engelen, L. Jones, R. Kivi, J.-D. Paris, V.-H. Peuch, V. Sherlock, A. T. Vermeulen, P. O. Wennberg & D. Wunch, 2014: Forecasting global atmospheric CO2. *Atmos. Chem. Phys.*, 14, 11959–11983. <https://doi.org/10.5194/acp-14-11959-2014>

Agust -Panareda, A., M. Diamantakis, S. Massart, F. Chevallier, J. Mu oz-Sabater, J. Barr , R. Curcoll, R. Engelen, B. Langerock, R. M. Law, Z. Loh, J. A. Morgu , M. Parrington, V.-H. Peuch, M. Ramonet, C. Roehl, A. T. Vermeulen, T. Warneke & D. Wunch, 2019: Modelling CO2 weather – why horizontal resolution matters. *Atmos. Chem. Phys.*, 19, 7347–7376. <https://doi.org/10.5194/acp-19-7347-2019>

Agust -Panareda, A., J. McNorton, G. Balsamo, B. C. Baier, N. Bousseres, S. Boussetta, D. Brunner, F. Chevallier, M. Choulga, M. Diamantakis, R. Engelen, J. Flemming, C. Granier, M. Guevara, H. Denier van der Gon, N. Elguindi, J.-M. Haussaire, M. Jung, G. Janssens-Maenhout, R. Kivi, S. Massart, D. Papale, M. Parrington, M. Razinger, C. Sweeney, A. Vermeulen & S. Walther, 2022: Global nature run data with realistic high-resolution carbon weather for the year of the Paris Agreement. *Nature Scientific Data*, volume 9, Article number 160.

Balsamo G., A. Beljaars, K. Scipal, P. Viterbo, B. v. d. Hurk, M. Hirschi & A. K. Betts, 2009: A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System. *J. of Hydrometeorology*, 10, pp. 623–642.

Barros S.R.M, D. Dent, L. Isaksen, G. Robinson, G. Mozdzyński & F. Wollenweber, 1995: The IFS model: A parallel production weather code. *Parallel Computing*, 21, 10, pp.1621–1638.

Bauer P., P. Dueben, T. Hoefler, T. Quintino, T. C. Schulthess & N. P. Wedi, 2021: The digital revolution of Earth-system science. *Nature Computational Science*, 10.1038/s43588-021-00023-0, Vol 1, pp. 104–113.

Bechtold, P., E. Bazile, F. Guichart, P. Mascart & E. Richard, 2001: A mass-flux convection scheme for regional and global models. *Q. J. R. Meteorol. Soc.*, 127, pp. 869–886.

Beljaars, A. & A. Holtslag, 1991: Flux parameterization over land surfaces for atmospheric models. *J. Appl. Meteor.*, 30, 327–341.

Beljaars A. C. M., A. R. Brown & N. Wood, 2004: A new parametrization of turbulent orographic form drag. *Q. J. R. Meteorol. Soc.*, 130, 599, pp. 1327–1347.

Beljaars, A. C. M., G. Balsamo, P. Bechtold, A. Bozzo, R. Forbes, R. J. Hogan, M. K hler, J.-J. Morcrette, A. M. Tompkins, P. Viterbo & N. Wedi, 2018: The Numerics of Physical Parametrization in the ECMWF Model. *Front. Earth Sci.*

Benard, P., 2003: Stability of Semi-Implicit and Iterative Centered-Implicit Time Discretizations for Various Equation Systems Used in NWP. *Monthly Weath. Rev.*, pp. 2479–2491.

Benard, P., J. Vivoda, J. Ma ek, P. Smol kov , K. Yessad, Ch. Smith, R. Bro kov  & J.-F. Geleyn, 2010: Dynamical kernel of the Aladin-NH spectral limited-area model: Revised formulation and sensitivity experiments. *Q. J. R. Meteorol. Soc.*, 136, pp. 155–169.

Best M. J., A. Beljaars, J. Polcher & P. Viterbo, 2004: A Proposed Structure for Coupling Tiled Surfaces with the Planetary Boundary Layer. *Journal of Hydrometeorology*, 5, pp. 1271–1278.

Bode et al., 2025: Towards JUPITER: First System and Application Experiences on JEDI, submitted.

Brankovi , C., C. Jakob, M. Miller, A. Untch & N. Wedi, 1999: Climate diagnostics of the ECMWF AMIP-2 simulations. *ECMWF Technical Memorandum No. 360*.

Bubnov , R, G. Hello, P. B nard & J.F. Geleyn, 1995: Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. *Mon. Weather Rev.*, 123, 515–535.

Buizza, R., 2019: Introduction to the special issue on 25 years of ensemble forecasting. *Q. J. R. Meteorol. Soc.*, Volume 145: Special Supplement on 25 Years of Ensemble Forecasting, pages i-iv, 1–231.

Buizza, R. & T.N. Palmer, 1995: The singular vector structure of the atmospheric general circulation. *Journal of the Atmospheric Sciences*, 52, 1434–1456.

Buizza, R., J.-R. Bidlot, M. Janousek, S. Keeley, K. Mogensen & D. Richardson, 2017: New IFS cycle brings sea-ice coupling and higher ocean resolution. *ECMWF Newsletter No. 150*, 14–17.

Boussetta S., G. Balsamo, G. Arduini, E. Dutra, J. McNorton, M. Choulga, A-A. Panareda, A. Beljaars, N. Wedi, J. Mun z-Sabater, P. de Rosnay, I. Sandu, I. Hadade, G. Carver, C. Mazzetti, C. Prudhomme, D. Yamazaki & E. Zsoter, 2021: ECLand: The ECMWF Land Surface Modelling System. *Atmosphere*, 12, 723, pp. 1–38.

Deconinck W., P. Bauer, M. Diamantakis, M. Hamrud, C. K hnlein, P. Maciel, G. Mengaldo, T. Quintino, B. Raoult, P. K. Smolarkiewicz & N. P. Wedi, 2017: Atlas: A library for numerical weather prediction and climate modelling. *Computer Physics Communications*, 220, pp. 188–204.

Dueben, P. D. & P. Bauer, 2018: Challenges and design choices for global weather and climate models based on machine learning. *Geosci. Model Dev.*, 11, 3999–4009. <https://doi.org/10.5194/gmd-11-3999-2018>

Eliassen, E., B. Machenhauer & E. Rasmussen, 1970: On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. Rep. 2, Institut for Teoretisk Meteorologi, University of Copenhagen, 37 pp.

ECMWF, 1979: ECMWF Technical Newsletter No.1. https://www.ecmwf.int/sites/default/files/elibrary/021979/19020-technical-newsletter-no-1-february-1979_1.pdf

Ferranti L., T. N. Palmer, F. Molteni & E. Klinker, 1990: Tropical-Extratropical Interaction Associated with the 30–60 Day Oscillation and Its Impact on Medium and

Extended Range Prediction. *Journal of the Atmospheric Sciences*, 47, pp. 2177–2199.

Forbes, R. M., A. M. Tompkins & A. Untch, 2011: A new prognostic bulk microphysics scheme for the IFS. *ECMWF Technical Memorandum No. 649*.

Frassoni, A., C. Reynolds, N. Wedi, Z. B. Bouallegue, A. C. Vaz Caltabiano, B. Casati, J. A. Christophersen, C. A. S. Coelho, C. De Falco, J. D. Doyle, L. G. Fernandes, R. Forbes, M. A. Janiga, D. Klocke, L. Magnusson, R. McTaggart-Cowan, M. Pakdaman, S. S. Rushley, A. Verhoef, F. Yang & G. Z ngl, 2023: Systematic Errors in Weather and Climate Models: Challenges and Opportunities in Complex Coupled Modeling Systems. *Bull. Amer. Meteor. Soc.*, E1687-E1693. <https://doi.org/10.1175/BAMS-D-23-0102.1>

Gates, W. L., J. Boyle, C. Covey, C. Dease, C. Doutriaux, R. Drach, M. Fiorino, Gleckler, J. Hnilo, S. Marlais, T. Phillips, G. Potter, B. Santer, K. Sperber, K. Taylor & D. Williams, 1998: An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, 73, 1962–1970.

Hoffman, R. N., S. Malardel; P. Maciel, T. R. Peevey, P. Partain, S. Finley, R. M. Atlas, L. Isaksen, N. Wedi, L. Cucurull, C. D. Kummerow, 2018: The New ECMWF Cubic Octahedral (O1280) Nature Run. <https://www.cira.colostate.edu/imagery-data/ecmwf-nature-run/>

Hogan R. J. & A. Bozzo, 2018: A Flexible and Efficient Radiation Scheme for the ECMWF Model. *Journal of Advances in Modeling Earth Systems*, 10, 8, pp. 1990–2008.

Husain, S. Z., L. Separovic, J.-F. Caron , R. Aider, M. Buehner, S. Chamberland, E. Lapalme, R. McTaggart-Cowan, C. Subich, P. A. Vaillancourt, J. Yanga & A. Zadra, 2024: Leveraging data-driven weather models for improving numerical weather prediction skill through large-scale spectral nudging. [arXiv:2407.06100v2](https://arxiv.org/abs/2407.06100v2) [physics.ao-ph]

Janiskov , M., J.-N. Th paut & J-F. Geleyn, 1999: Simplified and Regular Physical Parameterizations for Incremental Four-Dimensional Variational Assimilation. *Monthly Weather Review*, 127, 26–45.

Janssen, P. A. E. M., 2004: The interaction of ocean waves and wind. Cambridge University Press, 300pp. <https://doi.org/10.1017/CBO9780511525018>

Janssen, P.,  . Breivik, K. Mogensen, F. Vitart, M. Balmaseda, J-R. Bidlot, S. Keeley, M. Leutbecher, L. Magnusson & F. Molteni, 2013: Air-sea interaction and surface waves. *ECMWF Technical Memorandum No. 712*, 34.

Johnson, S. J., T. N. Stockdale, L. Ferranti, M. A. Balmaseda, F. Molteni, L. Magnusson, S. Tietsche, D. Decremer, A. Weisheimer, G. Balsamo, S. P. E. Keeley, K. Mogensen, H. Zuo & B. M. Monge-Sanz, 2019: SEAS5: the new ECMWF seasonal forecast system. *GMD*, 12, 1087–1117.

Koldunov, N. V., V. Aizinger, N. Rakowsky, P. Scholz, D. Sidorenko, S. Danilov & T. Jung, 2019: Scalability and some optimization of the Finite-volume Sea ice–Ocean Model, Version 2.0 (FESOM2). *Geosci. Model Dev.*, 12, 3991–4012. <https://doi.org/10.5194/gmd-12-3991-2019>

K hnlein C., W. Deconinck, R. Klein, S. Malardel, Z. P. Piotrowski, P. K. Smolarkiewicz, J. Szmelter & N. P. Wedi, 2019: FVM 1.0: a nonhydrostatic finite-volume dynamical core for the IFS. *Geosci. Model Dev. Discuss*, 12, pp. 651–676

Lang S, A. Dawson, M. Diamantakis, P. Dueben, S. Hatfield, M. Leutbecher,T. Palmer, F. Prates, C. D. Roberts, I. Sandu & N. Wedi, 2021: More accuracy with less precision. *Q. J. R. Meteorol. Soc.*, 147, 741, pp. 4358-4370.

Lang S., M. Alexe, M. Chantry, J. Dramsch, F. Pinault, B. Raoult, M. C. A. Clare, C. Lessig, M. Maier-Gerber, L. Magnusson, Z. B. Bouall gue, A. P. Nemesio, P. D. Dueben, A. Brown, F. Pappenberger, F. Rabier, 2024a: AIFS - ECMWF’s data-driven forecasting system. *arXiv:2406.01465*

Lang, S., M. Alexe, M. C. A. Clare, C. Roberts, R. Adewoyin, Z. B. Bouall gue, M. Chantry, J. Dramsch, P. D. Dueben, S. Hahner, P. Maciel, A. Prieto-Nemesio, C. O’Brien, F. Pinault, J. Polster, B. Raoult, S. Tietsche & M. Leutbecher 2024b: AIFS-CRPS: Ensemble forecasting using a model trained with a loss function based on the Continuous Ranked Probability Score. *arXiv2412.15832*

Leutbecher, M., S. J. Lock, P. Ollinaho, S. T. K. Lang, G. Balsamo, P. Bechtold, M. Bonavita, H. M. Christensen, M. Diamantakis, E. Dutra, S. English, M. Fisher, R. M. Forbes, J. Goddard, T. Haiden, R. J. Hogan, S. Juricke, H. Lawrence, D. MacLeod, L. Magnusson, S. Malardel, S. Massart, I. Sandu, P.K. Smolarkiewicz, A. Subrama nian, F. Vitart, N. Wedi & A. Weisheimer, 2017: Stochastic

representations of model uncertainties at ECMWF: state of the art and future vision. *Q. J. R. Meteorol. Soc.*, 143, 2315–2339.

Koldunov, N. V., V. Aizinger, N. Rakowsky, P. Scholz, D. Sidorenko, S. Danilov, and T. Jung, 2019: Scalability and some optimization of the Finite-volumE Sea ice–Ocean Model, Version 2.0 (FESOM2). *Geosci. Model Dev.*, 12, 3991–4012. <https://doi.org/10.5194/gmd-12-3991-2019>

Lin, J., T. Qian, P. Bechtold, G. Grell, G. J. Zhang, P. Zhu ... J. Han, 2022: Atmospheric Convection. *Atmosphere-Ocean*, 60(3–4), 422–476. <https://doi.org/10.1080/07055900.2022.2082915>

Lott, F. & M. J. Miller, 1996: A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Q. J. R. Meteorol. Soc.*, 122, 101–127.

Magnusson, L., D. Ackerley, Y. Bouteloup, J.-H. Chen, J. Doyle, P. Earnshaw, Y. C. Kwon, M. K hler, S. T. K Lang, Y.-J. Lim, M. Matsueda, T. Matsunobu & R. McTaggart-Cowan, A. Reinecke, M. Yamaguchi & L. Zhou, 2022: Skill of Medium-Range Forecast Models Using the Same Initial Conditions. *Bull. Amer. Meteor. Soc*, E2050–E2068.

Mahfouf, J.-F.,1999: Influence of physical processes on the tangent-linear approximation. *Tellus*, 51A, 147–166.

Majumdar, S. J., L. Magnusson, P. Bechtold, J.-R. Bidlot & J. D. Doyle, 2023: Advanced Tropical Cyclone Prediction Using the Experimental Global ECMWF and Operational Regional COAMPS-TC Systems. *Mon. Wea. Rev.*, 151, 2029–2048. <https://doi.org/10.1175/MWR-D-22-0236.1>

Malardel S., N. Wedi, W. Deconinck, M. Diamantakis, C. K hnlein, G. Mozdzyński, M. Hamrud & P. Smolarkiewicz, 2016: A new grid for the IFS. *ECMWF Newsletter No.146*.

Melvin, T., B. Shipway, N. Wood, T. Benacchio, T. Bendall, I. Boutle, A. Brown, J. Kent, S. Pring, C. Smith, C. Cotter & J. Thuburn, 2024: A mixed finite-element, finite-volume, semi-implicit discretisation for atmospheric dynamics: Spherical geometry. *Q.J.R. Meteorol. Soc*, 150(764):4252-4269.

Morcrette J.-J., H. W. Barker, J. N. S. Cole, M. J. Iacono & R. Pincus, 2008: Impact of a New Radiation Package, McRad, in the ECMWF Integrated Forecasting System. *Monthly Weath. Rev.*, pp. 4773-4798.

Mozdzynski G., M. Hamrud & N. Wedi, 2015: A Partitioned Global Address Space implementation of the European Centre for Medium Range Weather Forecasts Integrated Forecasting System. *Int. J. High Perform. Comput. Appl.*, 29, 3. <https://doi.org/10.1177/1094342015576773>

Mueller A., W. Deconinck, C. Kühnlein, G. Mengaldo, M. Lange, N. Wedi, P. Bauer, P. K. Smolarkiewicz et al., 2019: The ESCAPE project: Energy-efficient Scalable Algorithms for Weather Prediction at Exascale. *Geosci. Model Dev.*, 12, pp. 4425–4441.

O'Reilly, C., L. Brunner, S. Qasmi, R. Nogherotto, A. P. Ballinger, B. Booth, D. J. Befort, R. Knutti, A. P. Schurer, A. Ribes, A. Weisheimer, E. Coppola & C. McSweeney, 2024: Assessing observational constraints on future European climate in an out-of-sample framework. *npj Climate and Atmospheric Science*, volume 7, Article number: 95

Orszag, S. A., 1970: Transform method for calculation of vector coupled sums: Application to the spectral form of the vorticity equation. *J. Atmos. Sci.*, 27, 890–895.

Pailleux J., J-F. Geleyn, M. Hamrud, P. Courtier, J.-N. Thépaut, F. Rabier, E. Andersson, D. Burridge, A. Simmons, D. Salmond, R. E. Khatib & C. Fischer, 2014: Twenty-five years of IFS/Arpege. *ECMWF Newsletter* No.141.

Polichtchouk, I., P. Bechtold, M. Bonavita, R. Forbes, S. Healy, R. Hogan, P. Laloyaux, M. Rennie, T. Stockdale, N. Wedi, M. Diamantakis, J. Flemming, S. English, L. Isaksen, F. Vána, S. Gisinger & N. Byrne, 2021: Stratospheric modelling and assimilation. *ECMWF Technical Memorandum* No. 877.

Polichtchouk I., K. S. Mogensen, E. R. Sanabia, S. R. Jayne, L. Magnusson, C. R. Densmore, S. Hatfield, I. Hadade, N. Wedi, V. Anantharaj, P. Lopez & A. K. Ekholm, 2025: Effects of atmosphere and ocean horizontal model resolution on tropical cyclone and upper ocean response forecasts in four major hurricanes. MWR-D-24-0104 has been accepted for publication in Monthly Weather Review.

Rackow, T., X. Pedruzo-Bagazgoitia, T. Becker, S. Milinski, I. Sandu, R. Aguridan, P. Bechtold, S. Beyer, J. Bidlot, S. Boussetta, M. Diamantakis, P. Dueben, E. Dutra, R. Forbes, H. F. Goessling, I. Hadade, J. Hegewald, S. Keeley, L. Kluft, N. Koldunov, A. Koldunov, T. Kölling, J. Kousal, K. Mogensen, T. Quintino, I. Polichtchouk,

D. Sármány, D. Sidorenko, J. Streffing, B. Sützl, D. Takasuka, S. Tietsche, M. Valentini, B. Vannière, N. Wedi, L. Zampieri & F. Ziemer, 2024: Multi-year simulations at kilometre scale with the Integrated Forecasting System coupled to FESOM2.5/ NEMOv3.4. *EGUsphere* [preprint]. <https://doi.org/10.5194/egusphere-2024-913>

Schulthess, T., P. Bauer, O. Fuhrer, T. Hoefler, C. Schär & N.P. Wedi, 2019: Reflecting on the goal and baseline for ‘Exascale Computing’: a roadmap based on weather and climate simulations. *Computing in Science and Engineering*, IEEE, Vol. 21, 30–41.

Simmons, A.J. & A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Q.J.R. Meteorol. Soc.*, 128, 647–677

Simmons, A.J., B. Hoskins & D. Burridge, 1978: Stability of the Semi-Implicit Method of Time Integration. *Monthly Weath. Rev.* 106, pp.405–412.

Smolarkiewicz, P., C. Kühnlein & N. P. Wedi, 2014: A consistent framework for discrete integrations of soundproof and compressible PDEs of atmospheric dynamics. *J. Comput. Phys.* 263, pp 185–205

Temperton, C., 1983: Self-sorting mixed-radix fast Fourier trans- forms. *J. Comput. Phys.*, 52, 1–23.

Temperton, C., 1991: On scalar and vector transform methods for global spectral models. *Mon. Wea. Rev.*, 119, 1303–1307.

Temperton C., M. Hortal & A. Simmons, 2001: A two-time-level semi-Lagrangian global spectral model. *Q.J.R. Meteorol. Soc.*, 127, 571, pp. 111–127.

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parametrization in large-scale models. *Mon. Wea. Rev.* 117, 1779–1800.

Tiedtke A., 1993: Representation of Clouds in Large-Scale Models. *Monthly Weath. Rev.* 121, pp. 3040–3061

Ubbiali, S., C. Kühnlein, C. Schär, L. Schlemmer, T.C. Schulthess, M. Staneker, M & H. Wernli, 2025: Exploring a high-level programming model for the NWP domain using ECMWF microphysics schemes. *Geosci. Model Dev.*, 18, 529–546. <https://doi.org/10.5194/gmd-18-529-2025>

Ullrich, P. A., C. Jablonowski, J. Kent, P.H. Lauritzen, R. Nair, K.A. Reed, C.

M. Zarzycki, D. M. Hall, D. Dazlich, R. Heikes, C. Konor, D. Randall, T. Dubos, Y. Meurdesoif, X. Chen, L. Harris, C. Kühnlein, V. Lee, A. Qaddouri, C. Girard, M. Giorgetta, D. Reinert, J. Klemp, S.-H. Park, W. Skamarock, H. Miura, T. Ohno, R. Yoshida, R. Walko, A. Reinecke & K. Viner, 2017: DCMIP2016: a review of non-hydrostatic dynamical core design and intercomparison of participating models. *Geosci. Model Dev.*, 10, 4477–4509. <https://doi.org/10.5194/gmd-10-4477-2017>

Untch, A. & M. Hortal, 2004: A finite-element scheme for the vertical discretization of the semi-Lagrangian version of the ECMWF forecast model. *Q.J.R. Meteorol. Soc.*, 130, pp. 1505–1530.

Vitart, F., R. Buizza, M. Alonso Balmaseda, G. Balsamo, J.-R. Bidlot, A. Bonet, M. Fuentes, A. Hofstadler, F. Molteni & T.N. Palmer, 2008: The new VarEPS-monthly forecasting system: A first step towards seamless prediction. *Q.J.R. Meteorol. Soc.*, 134, 1789–1799. <https://doi.org/10.1002/qj.322>

Vitart, F. et al., 2017: The subseasonal to seasonal (S2S) prediction database. *Bull. Amer. Meteor. Soc.*, 98, 163–173. <https://doi.org/10.1175/BAMS-D-16-0017.1>

Voitus, F., P. Benard, C. Kuehnlein & N.P. Wedi, 2019: Semi-implicit integration of the unified equations in a mass-based coordinate: model formulation and numerical testing. *Q. J. R. Meteorol. Soc.*, 145, pp. 3387–3408.

Wedi N.P. & P.K. Smolarkiewicz, 2009: A framework for testing global non-hydrostatic models. *Q. J. R. Meteorol. Soc.*, 135, 469–484.

Wedi, N. P., 2014: Increasing horizontal resolution in numerical weather prediction and climate simulations: illusion or panacea? *Phil. Trans. R. Soc. A.*, 372, 20130289.

Wedi, N.P., M. Hamrud, & G. Mozdzynski, 2013: A Fast Spherical Harmonics Transform for Global NWP and Climate Models. *Monthly Weath. Rev.*, 141, pp. 3450–3461.

Wedi, N.P., I. Polichtchouk, P. Dueben, V. G. Anantharaj, P. Bauer, S. Boussetta, P. Browne, W. Deconinck, W. Gaudin, I. Hadade, S. Hatfield, O. Iffrig, P. Lopez, P. Maciel, A. Mueller, S.Saarinen, I. Sandu, T. Quintino & F. Vitart, 2020: A baseline for global weather and climate simulations at 1 km resolution. *Journal of Advances in Modeling Earth Systems*, 12, e2020MS002192.

Wedi, N., I. Sandu, P. Bauer, M. Acosta, R. C. Andersen, U. Andrae, L. Auger, G. Balsamo, V. Baousis, V. Bennett, A. Bennett, C. Buontempo, P.-A. Bretonnière, R. Capell, M. Castrillo, M. Chantry, M. Chevallier, R. Correa, P. Davini, L. Denby, F. Doblas-Reyes, P. Dueben, C. Fischer, C. Frauen, I.-L. Frogner, B. Früh, E. Gascón, E. Gérard, O. Gorwits, T. Geenen, K. Grayson, N. Guenova-Rubio, I. Hadade, J. von Hardenberg, U.-U. Haus, J. Hawkes, M. Hirtl, J. Hoffmann, K. Horvath, H. Järvinen, T. Jung, A. Kann, D. Klocke, N. Koldunov, J. Kontkanen, O. Sievi-Korte, J. Kristiansen, E. Kuwertz, J. Mäkelä, I. Maljutenko, P. Manninen, U. S. McKnight, S. Milinski, A. McNally, U. Modigliani, D. Narayanappa, K. P. Nielsen, T. Nipen, H. Nortamo, V.-H. Peuch, S. Polade, T. Quintino, B. Reuter, I. Schicker, S. Smart, M. Sleigh, M. Suttie, P. Termonia, S. Thober, R. Randriamampianina, N. Theeuwes, D. Thiembert, B. Vannière, S. Vannitsem, C. Wittmann, X. Yang, M. Pontaud, B. Stevens & F. Pappenberger, 2025: Implementing Digital Twin technology of the Earth System in Destination Earth. *Journal of the European Meteorological Society*, in review 2025.

Williams, K.D., A. Bodas-Salcedo, M. Déqué, S. Fermepein, B. Medeiros, M. Watanabe, C. Jakob, S. A. Klein, C. A. Senior & D. L. Williamson, 2013: The Transpose-AMIP II Experiment and Its Application to the Understanding of Southern Ocean Cloud Biases in Climate Models, *Journal of Climate*, Vol. 26, pp 3258–3274.

Williamson, D, 2007: The evolution of dynamical cores for global atmospheric models. *J. Meteorology. Soc. Japan*, 85B, pp. 241–269.

Woods, A., 2005 Medium-Range Weather Prediction: The European Approach. Springer. <https://doi.org/10.1007/b138324>

Wood, N., A. Staniforth, A. White, T. Allen, M. Diamantakis, M. Gross, T. Melvin, C. Smith, S. Vosper, M. Zerroukat & J. Thuburn, 2014: An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Q.J.R. Meteorol. Soc.*, 140, 1505–1520.

Zadra, A., K. Williams, A. Frassoni, M. Rixen, A. F. Adames, J. Berner, F. Bouyssel, B. Casati, H. Christensen, M. B. Ek, G. Flato, Y. Huang, F. Judt, H. Lin, E. Maloney, W. Merryfield, A. Van Niekerk, T. Rackow, K. Saito, N. Wedi & P. Yadav, 2018: Systematic errors in weather and climate models: Nature, origins, and ways forward. *Bull. Amer.*

Meteor. Soc., 99, ES67–ES70. <https://doi.org/10.1175/BAMS-D-17-0287.1>

Zängl, G., D. Reinert, P. Rípodas & M. Baldauf, 2015: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Q.J.R. Meteorol. Soc.*, Vol. 141, Issue 687, pp. 563–579.

Zeman, C., N. P. Wedi, P. D. Dueben, N. Ban & C. Schär, 2021: Model intercomparison of COSMO 5.0 and IFS 45r1 at kilometer-scale grid spacing. GMD, Vol. 14, Issue 7, 4617–4639.

FIFTY YEARS OF DATA ASSIMILATION AT ECMWF

Alexe, M., E. Boucher, P. Lean, E. Pinnington, P. Laloyaux, A. McNally, S. Lang, M. Chantry, C. Burrows, M. Chrust, F. Pinault, E. Villeneuve, N. Bormann & S. Healy, 2024: GraphDOP: Towards skilful data-driven medium-range weather forecasts learnt and initialised directly from observations. *arXiv e-prints*, arXiv:2412.15687. <https://doi.org/10.48550/arXiv.2412.15687>

Andersson, E., A. Hollingsworth, G. Kelly, P. Lönnberg, J. Pailleux & Z. Zhang, 1991: Global observing system experiments on operational statistical retrievals of satellite sounding data. *Mon. Wea. Rev.*, 119, 1851–1865. [https://doi.org/10.1175/1520-0493\(1991\)119<1851:GOSEOO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1991)119<1851:GOSEOO>2.0.CO;2)

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. <https://doi.org/10.1002/qj.49712051707>

Bauer, P., A. Thorpe & G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, 525, 47–55. <https://doi.org/10.1038/nature14956>

Bonavita, M., L. Isaksen & E. Hólm, 2012: On the use of EDA background error variances in the ECMWF 4D-Var. *ECMWF Technical Memorandum* No. 664. <https://doi.org/10.21957/3msfrh5zm>

Bonavita, M., E.Hólm, L. Isaksen & M. Fisher, 2016: The evolution of the ECMWF hybrid data assimilation system. *Q.J.R. Meteorol. Soc.*, 142, 287–303. <https://doi.org/10.1002/qj.2652>

Bonavita, M., Y. Trémolet, E. Hólm, S. Lang, M. Chrust, M. Janiskova, P. Lopez, P. Laloyaux, P. de Rosnay, M. Fisher, M. Hamrud & S. English, 2017: A strategy for data assimilation. *ECMWF Technical Memorandum* No. 800. <https://doi.org/10.21957/tx1epjd2p>

Bonavita, M., P. Lean & E. Hólm, 2018: Nonlinear effects in 4D-Var. *Nonlin. Processes Geophys.*, 25, 713–729. <https://doi.org/10.5194/npg-25-713-2018>

Bonavita, M. & P. Laloyaux, 2020: Machine learning for model error inference and correction. *Journal of Advances in Modeling Earth Systems*, 12, e2020MS002232. <https://doi.org/10.1029/2020MS002232>

Bormann, N., P. de Rosnay, S. Healy, H. Zuo, S. E. English & F. Catapan, 2025: ECMWF and ESA start project to better exploit Earth system satellite data. *ECMWF Newsletter* No. 182, 10–12. <https://doi.org/10.21957/m32jp51ad6>

Buizza, R., M. Leutbecher & L. Isaksen, 2008: Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. *Q.J.R. Meteorol. Soc.*, 134, 2051–2066. <https://doi.org/10.1002/qj.346>

Cardinali, C., 2009: Monitoring the observation impact of the short-range forecast. *Q.J.R. Meteorol. Soc.*, 135, 239–250. <https://doi.org/10.1002/qj.366>

Courtier, P., J.-N. Thépaut & A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Q.J.R. Meteorol. Soc.*, 120, 1367–1387. <https://doi.org/10.1002/qj.49712051912>

Dahoui, M., L. Isaksen & G. Radnoti, 2017: Assessing the impact of observations using observation-minus-forecast residuals. *ECMWF Newsletter* No. 152, 27–31. <https://doi.org/10.21957/51j3sa>

Dahoui, M., 2023: Use of machine learning for the detection and classification of observation anomalies. *ECMWF Newsletter* No. 174, 23–27. <https://doi.org/10.21957/n64md0xa5d>

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kállberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J.

Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavalato, J.-N. Thépaut & F. Vitart, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.*, **137**, 553–597. <https://doi.org/10.1002/qj.828>

de Rosnay, P., M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel & L. Isaksen: 2013: A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Q.J.R. Meteorol. Soc.*, **139**, 1199–1213. <https://doi.org/10.1002/qj.2023>

de Rosnay et al., 2022: Coupled data assimilation at ECMWF: current status, challenges and future developments. *Q.J.R. Meteorol. Soc.*, **148**, 2672–2702. <https://doi.org/10.1002/qj.4330>

Dramsch, J., B. Raoult, M. Chantry, T. García, L. Denby, F. Prill, N. Sokka, A. Vocino, J. Wijnands, T. Nipen, C. Osuna, S. Akodad, M. Van Genderachter, D. Van den Bleeken, 2024: Introducing Anemoi: a new collaborative framework for ML weather forecasting. *ECMWF Newsletter No. 181*, 6–8. <https://www.ecmwf.int/en/newsletter/181/news/introducing-anemoi-new-collaborative-framework-ml-weather-forecasting>

Eyre, J. R., 1989: Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. I: Theory and simulation for TOVS. *Q.J.R. Meteorol. Soc.*, **115**, 1001–1026. <https://doi.org/10.1002/qj.49711548902>

Eyre, J.R., S.J. English & M. Forsythe, 2020: Assimilation of satellite data in numerical weather prediction. Part I: The early years. *Q.J.R. Meteorol. Soc.*, **146**, 49–68. <https://doi.org/10.1002/qj.3654>

Farchi, A., M. Chrust, M. Bocquet & M. Bonavita, 2025: Development of an offline and online hybrid model for the Integrated Forecasting System. *Q.J.R. Meteorol. Soc.*, e4934. <https://doi.org/10.1002/qj.4934>

Fielding, M., M. Janisková, S. Mason, R. Hogan, W. McLean, A. Benedetti & R. Forbes, 2025: EarthCARE data begin to make an impact. *ECMWF Newsletter No. 183*, 21–27. <https://doi.org/10.21957/65ds8j72mf>

Fisher, M., 2005: “Wavelet” Jb – A new way to model the statistics of background errors. *ECMWF Newsletter No. 106*, 23–28. <https://doi.org/10.21957/ef5601ak43>

Geer, A. J., K. Lonitz, P. Weston, M. Kazumori, K. Okamoto, Y. Zhu, E. H. Liu, A. Collard, W. Bell, S. Migliorini, P. Chambon, N. Fourrié, M.-J. Kim, C. Köpken-Watts & C. Schraff, 2018:

All-sky satellite data assimilation at operational weather forecasting centres. *Q.J.R. Meteorol. Soc.*, **144**, 1191–1217. <https://doi.org/10.1002/qj.3202>

Geer, A. J., S. Migliorini & M. Matricardi, 2019: All-sky assimilation of infrared radiances sensitive to mid- and upper-tropospheric moisture and cloud. *Atmos. Meas. Tech.*, **12**, 4903–4929. <https://doi.org/10.5194/amt-12-4903-2019>

Geer, A. J., 2024a: Combining machine learning and data assimilation to estimate sea ice concentration. *ECMWF Newsletter No. 177*, 14–21. <https://doi.org/10.21957/agh93vs26>

Geer, A. J., 2024b: Simultaneous inference of sea ice state and surface emissivity model using machine learning and data assimilation. *Journal of Advances in Modeling Earth Systems*, **16**, e2023MS004080. <https://doi.org/10.1029/2023MS004080>

Healy, S., N. Bormann & K. Lean, 2019: Operational assimilation of Metop-C data. *ECMWF Newsletter No. 159*, 3-5. <https://www.ecmwf.int/en/newsletter/159/news/operational-assimilation-metop-c-data>

Healy, S., N. Bormann, A. Geer, E. Hólm, B. Ingleby, K. Lean, K. Lonitz & C. Lupu, 2022: Methods for assessing the impact of current and future components of the global observing system. *ECMWF Technical Memorandum No. 916*. <https://doi.org/10.21957/2f240fe55f>

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. D. Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. V. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume & J.-N. Thépaut, 2020: The ERA5 global reanalysis. *Q.R.J. Meteorol. Soc.*, **146**, 1999–2049. <https://doi.org/10.1002/qj.3803>

Hólm, E., M. Bonavita & S. Lang, 2022: Soft re-centring Ensemble of Data Assimilations. *ECMWF Newsletter No. 171*, 6–8. <https://www.ecmwf.int/en/newsletter/171/news/soft-re-centring-ensemble-data-assimilations>

Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher & L. Raynaud, 2010: Ensemble of Data Assimilations at ECMWF. *ECMWF*

Technical Memorandum No. 636. <https://doi.org/10.21957/obke4k60>

Janisková, M., J.-F. Mahfouf, J.-J. Morcrette & F. Chevallier, 2002: Linearized radiation and cloud schemes in the ECMWF model: Development and evaluation. *Q.J.R. Meteorol. Soc.*, **128**, 1505-1527. <https://doi.org/10.1002/qj.200212858306>

Janisková, M. & P. Lopez, 2013: Linearized physics for data assimilation at ECMWF. In: Park, S., L. Xu (eds) Data Assimilation for Atmospheric, Ocean and Hydrological Applications (Vol. II). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-35088-7_11

Laloyaux, P., M. Bonavita, M. Dahoui & E. Hólm, 2020: Towards an unbiased stratospheric analysis. I, **146**, 2392–2409. <https://doi.org/10.1002/qj.3798>

Lang, S., E. Hólm, M. Bonavita & Y. Trémolet, 2019. A 50-member Ensemble of Data Assimilations. *ECMWF Newsletter No. 158*, 27–29. <https://www.ecmwf.int/en/newsletter/158/meteorology/50-member-ensemble-data-assimilations>

Lavers, D. A. et al., 2024: Advancing Atmospheric River Science and Inspiring Future Development of the Atmospheric River Reconnaissance Program. *Bull. Amer. Meteor. Soc.*, **105**, E75–E83. <https://doi.org/10.1175/BAMS-D-23-0278.1>

Lean, P., M. Bonavita, E. Hólm et al., 2019: Continuous data assimilation for the IFS. *ECMWF Newsletter No. 158*, 21–26. <https://doi.org/10.21957/9pl5fc37it>

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721. [https://doi.org/10.1175/1520-0493\(1981\)109<0701:AGTDMS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1981)109<0701:AGTDMS>2.0.CO;2)

Magnusson, L. & E. Källén, 2013: Factors Influencing Skill Improvements in the ECMWF Forecasting System. *Mon. Wea. Rev.*, **141**, 3142–3153, <https://doi.org/10.1175/MWR-D-12-00318.1>

Magnusson, L., S.J. Majumdar, M.L. Dahoui, N. Bormann, M. Bonavita, P.A. Browne et al., 2025: The role of observations in ECMWF tropical cyclone initialisation and forecasting. *Q.J.R. Meteorol. Soc.*, 1–23. <https://doi.org/10.1002/qj.4924>

Mahfouf, J.-F. & F. Rabier, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics.

Q.J.R. Meteorol. Soc., **126**(564), 1171–1190. <https://doi.org/10.1002/qj.49712656416>

McNally, T., M. Bonavita & J. Thépaut, 2014: The Role of Satellite Data in the Forecasting of Hurricane Sandy. *Mon. Wea. Rev.*, **142**, 634–646. <https://doi.org/10.1175/MWR-D-13-00170.1>

McNally, T., C. Lessig, P. Lean, M. Chantry, M. Alexe & S. Lang, 2024: Red sky at night... producing weather forecasts directly from observations. *ECMWF Newsletter No. 178*, 30–34. <https://doi.org/10.21957/tmc81jo4c7>

Melinc, B. & Z. Zaplotnik, 2024: 3D-Var data assimilation using a variational autoencoder. *Q.J.R. Meteorol. Soc.*, **150**(761), 2273–2295. <https://doi.org/10.1002/qj.4708>

Pailleux, J., J.-F. Geleyn, M. Hamrud, P. Courtier, J.-N. Thépaut, F. Rabier, E. Andersson, D. M. Burridge, A. J. Simmons, D. Salmond, R. El Khatib & C. Fischer, 2014: Twenty-five years of IFS/ARPEGE. *ECMWF Newsletter No. 141*, 22–30. <https://doi.org/10.21957/ftu6mfvy>

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf & A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q.J.R. Meteorol. Soc.*, **126**, 1143–1170. <https://doi.org/10.1002/qj.49712656415>

Rabier, F., A. Bouchard, E. Brun, A. Doerenbecher, S. Guedj, V. Guidard, F. Karbou, V.-H. Peuch, L. E. Amraoui, D. Puech, C. Genthon, G. Picard, M. Town, A. Hertzog, F. Vial, P. Cocquerez, S. Cohn, T. Hock, H. Cole, J. Fox, D. Parsons, J. Powers, K. Romberg, J. VanAndel, T. Deshler, J. Mercer, J. Haase, L. Avallone, L. Kalnajs, C. R. Mechoso, A. Tangborn, A. Pellegrini, Y. Frenot, A. McNally, J.-N. Thépaut, G. Balsamo and P. Steinle, 2013: The Concordiasi infield Experiment over Antarctica: First Results from Innovative Atmospheric Measurements. *Bull. Amer. Meteor. Soc.*, **94**, ES17–ES20. <https://doi.org/10.1175/BAMS-D-12-00005.1>

Raynaud, L., L. Berre & G. Desroziers, 2008: Spatial averaging of ensemble-based background-error variances. *Q.J.R. Meteorol. Soc.*, **134**, 1003–1014. <https://doi.org/10.1002/qj.245>

Reale, A., D. Gray, M. Chalfant, A. Swaroop & A. Nappi, 1986: Higher resolution operational satellite retrievals. In Conference on Satellite Meteorology/Remote Sensing and Applications, 2nd, Williamsburg, VA (pp. 16-19).

Sandu, I., 2024: Destination Earth’s digital twins and Digital Twin Engine – state of play. *ECMWF Newsletter No. 180*, 22–29. <https://doi.org/10.21957/is1fc736jx>

Simmons, A. J., H. Hersbach, J. Muñoz-Sabater, J. Nicolas, F. Vamborg, P. Berrisford, P. de Rosnay, K. Willett & J. Woollen, 2021: Low frequency variability and trends in surface air temperature and humidity from ERA5 and other datasets. *ECMWF Technical Memorandum No. 881*. <https://doi.org/10.21957/ly5vbtbfid>

Temperton, C. & D. L. Williamson, 1981: Normal Mode Initialization for a Multilevel Grid-Point Model. Part I: Linear Aspects. *Mon. Wea. Rev.*, **109**, 729–743. [https://doi.org/10.1175/1520-0493\(1981\)109<0729:NMIFAM>2.0.CO;2](https://doi.org/10.1175/1520-0493(1981)109<0729:NMIFAM>2.0.CO;2)

Thépaut, J.-N., P. Courtier, G. Belaud & G. Lemaître, 1996: Dynamical structure functions in a four-dimensional variational assimilation: A case study. *Q.J.R. Meteorol. Soc.*, **122**, 535–561. <https://doi.org/10.1002/qj.49712253012>

Trémolet, Y., 2006: Accounting for an imperfect model in 4D-Var. *Q.J.R. Meteorol. Soc.*, **132**, 2483–2504. <https://doi.org/10.1256/qj.05.224>

TEN YEARS OF COPERNICUS AT ECMWF

Agustí-Panareda, A., J. Barré, S. Massart, A. Inness, I. Aben, M. Ades, B. C. Baier, G. Balsamo, T. Borsdorff, N. Bousseriez, S. Boussetta, M. Buchwitz, L. Cantarello, C. Crevoisier, R. Engelen, H. Eskes, J. Flemming, S. Garrigues, O. Hasekamp, V. Huijnen, L. Jones, Z. Kipling, B. Langerock, J. McNorton, N. Meilhac, S. Noël, M. Parrington, V.-H. Peuch, M. Ramonet, M. Razinger, M. Reuter, R. Ribas, M. Suttie, C. Sweeney, J. Tarniewicz & L. Wu, 2023: Technical note: The CAMS greenhouse gas reanalysis from 2003 to 2020. *Atmospheric Chemistry and Physics*, **23**(6), 3829–3859. <https://doi.org/10.5194/acp-23-3829-2023>

Buontempo, C., S. N. Burgess, D. Dee, B. Pinty, J.-N. Thépaut, M. Rixen, S. Almond, D. Armstrong, A. Brookshaw, A. Lopez Alos, B. Bell, C. Bergeron, C. Cagnazzo, E. Comyn-Platt, E. Damasio-Da-Costa, A. Guillory, H. Hersbach, A. Horányi, J. Nicolas, A. Obregon, E. Penabad Ramos, B. Raoult, J. Muñoz-Sabater, A. Simmons, C. Soci, M. Suttie, F. Vamborg, J. Varndell, S. Vermoote, X. Yang & J. Garcés de Marcilla,

2022: The Copernicus Climate Change Service: Climate Science in Action. *Bull. Amer. Meteor. Soc.*, **103**(12), E2669–E2687. <https://doi.org/10.1175/BAMS-D-21-0315.1>

Cornes, R.C., G. van der Schrier, E.J.M. van den Besselaar & P.D. Jones, 2018: An Ensemble Version of the E-OBS Temperature and Precipitation Datasets. *Journal of Geophysical Research: Atmospheres*, **123**(17), 9391–9409. <https://doi.org/10.1029/2017JD028200>

Di Giuseppe, F., F. Pappenberger, F. Wetterhall, B. Krzeminski, A. Camia, G. Libertá & J. San Miguel, 2016: The potential predictability of fire danger provided by numerical weather prediction. *Journal of Applied Meteorology and Climatology*, **55**(11), 2469–2491. <https://doi.org/10.1175/JAMC-D-15-0297.1>

Flemming, J., H. Flentje, M.G. Schultz, O. Stein, A. Inness, L. Jones, V.-H. Peuch, A.-M. Blechschmidt, A. Richter, W. Spangl, R. Weller & C. Zerefos, 2015: Evaluation of the MACC operational forecast system: Potential and challenges of global near-real-time modelling with respect to reactive gases in the troposphere. *Atmospheric Chemistry and Physics*, **15**(23), 14005–14030. <https://doi.org/10.5194/acp-15-14005-2015>

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G.D. Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume & J.-N. Thépaut, 2020: The ERA5 global reanalysis. *Q.J.R. Meteorol. Soc.*, **146**(730), 1999–2049. <https://doi.org/10.1002/qj.3803>

Hollingsworth, A., R. J. Engelen, C. Textor, A. Benedetti, O. Boucher, F. Chevallier, A. Dethof, H. Elbern, H. Eskes, J. Flemming, C. Granier, J. W. Kaiser, J.-J. Morcrette, P. Rayner, V.-H. Peuch, L. Rouil, M. G. Schultz, A. J. Simmons & The Gems Consortium, 2008: Toward a monitoring and forecasting system for atmospheric composition: The GEMS project. *Bull. Amer. Meteor. Soc.*, **89**(8), 1147-1164. <https://doi.org/10.1175/2008BAMS2355.1>

Inness, A., M. Ades, A. Agustí-Panareda, J. Barré, A. Benedictow, A.-M. Blechschmidt, J.J. Dominguez, R. Engelen, H. Eskes, J.

Flemming, V. Huijnen, L. Jones, Z. Kipling, S. Massart, M. Parrington, V.-H. Peuch, M. Razinger, S. Remy, M. Schulz & M. Suttie, 2019: The CAMS reanalysis of atmospheric composition. *Atmospheric Chemistry and Physics*, **19**(6), 3515–3556. <https://doi.org/10.5194/acp-19-3515-2019>

Matthews, G., C. Baugh, C. Barnard, C. Carton De Wiart, J. Colonese, D. Decremers, S. Grimaldi, E. Hansford, C. Mazzetti, K. O'Regan, F. Pappenberger, A. Ramos, P. Salamon, D. Tasev & C. Prudhomme, 2024: On the operational implementation of the European Flood Awareness System (EFAS). In Flood forecasting: *A global perspective* (2nd ed., pp. 251–274). Elsevier. <https://doi.org/10.1016/B978-0-443-14009-9.00005-5>

Matthews, G., C. Baugh, C. Barnard, C. Carton De Wiart, J. Colonese, S. Grimaldi, D. Ham, E. Hansford, S. Harrigan, S. Heiselberg, H. Hooker, S. Hossain, C. Mazzetti, L. Milano, F. Moschini, K. O'Regan, F. Pappenberger, D. Pfister, R.M. Rajbhandari, P. Salamon & C. Prudhomme, 2024: On the operational implementation of the Global Flood Awareness System (GloFAS). In *Flood Forecasting: A Global Perspective* (2nd ed., pp. 299–350). Elsevier. <https://doi.org/10.1016/B978-0-443-14009-9.00006-7>

Muñoz-Sabater, J., E. Dutra, A. Agustí-Panareda, C. Albergel, G. Arduini, G. Balsamo, S. Boussetta, M. Choulga, S. Harrigan, H. Hersbach, B. Martens, D. G. Miralles, M. Piles, N. J. Rodríguez-Fernández, E. Zsoter, C. Buontempo & J.-N. Thépaut, 2021: ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, **13**(9), 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>

Peuch, V.-H., R. Engelen, M. Rixen, D. Dee, J. Flemming, M. Suttie, M. Ades, A. Agustí-Panareda, C. Ananasso, E. Andersson, D. Armstrong, J. Barré, N. Bousserez, J. J. Dominguez, S. Garrigues, A. Inness, L. Jones, Z. Kipling, J. Letertre-Danczak, M. Parrington, M. Razinger, R. Ribas, S. Vermoote, X. Yang, A. Simmons, J. Garcés de Marcilla & J.-N. Thépaut, 2022: The Copernicus Atmosphere Monitoring Service: From Research to Operations. *Bull. Amer. Meteor. Soc.*, **103**(12), E2650–E2668. <https://doi.org/10.1175/BAMS-D-21-0314.1>

Poli, P., H. Hersbach, D. P. Dee, P. Berrisford, A. J. Simmons, F. Vitart, P. Laloyaux, D. G. H. Tan, C. Peubey, J.-N. Thépaut, Y. Trémolet, E. V. Hólm, M. Bonavita, L. Isaksen & M. Fisher, 2016: ERA-20C: An atmospheric reanalysis of

the twentieth century. *Journal of Climate*, **29**(11), 4083–4097. <https://doi.org/10.1175/JCLI-D-15-0556.1>

Sandu, I., 2024: Destination Earth’s digital twins and Digital Twin Engine – state of play. *ECMWF Newsletter*, No. 180, 22–29. <https://doi.org/10.21957/is1fc736jx>

Thépaut, J.-N., D. Dee, R. Engelen & B. Pinty, 2018: The Copernicus Programme and its climate change service. *IEEE Int. Geoscience and Remote Sensing Symp., Valencia, Spain*, 1591–1593. <https://doi.org/10.1109/IGARSS.2018.8518067>

THE CRITICAL ROLE OF HIGH-PERFORMANCE COMPUTING IN MEDIUM-RANGE WEATHER FORECASTING: HALF A CENTURY OF TECHNOLOGY INNOVATION

ECMWF, 1975: *Newsletter* No. 3.

ECMWF, 1979: *Technical Newsletter* No. 1. <https://www.ecmwf.int/en/elibrary/81039-technical-newsletter-no-1-february-1979>

Moore, G. E., 1965: Cramming more components onto integrated circuits. *Electronics*, Volume 38, Number 8.

Raoult, B., 1997: Architecture of the New MARS Server, *Sixth Workshop on Meteorological Operational Systems*, 17–21 November 1997. <https://www.ecmwf.int/en/elibrary/76107-architecture-new-mars-server>

FIFTY YEARS OF MEETING USERS’ NEEDS

Abellan, X., V. Baousis, R. Correa, R. Cuccu, C. Duma, C. Kominos, S. Langlois & U. Modigliani, 2024: The European Weather Cloud is now operational. *ECMWF Newsletter* No. 180, 37–43. <https://doi.org/10.21957/xawe573uI9>

Dando, P. & U. Modigliani, 2025: Twenty years of the Framework for Member State time-critical applications. *ECMWF Newsletter* No. 182, 24–32. <https://doi.org/10.21957/p31dmu80sd>

Fucile, E., J. Tandy, T. Kralidis & R. Giraud, 2023: WIS 2.0: WMO data sharing in the 21st century. *ECMWF Newsletter* No. 176, 35–39. <https://doi.org/10.21957/rg3f1ky704>

Haiden, T. & S. Duffy, 2016. Use of high-density observations in precipitation verification. *ECMWF Newsletter* No. 147, 20–25. <https://doi.org/10.21957/hsacrdem>

Haiden, T., M. Janousek, F. Vitart, M. Tanguy, F. Prates & M. Chevalier, 2024: Evaluation of ECMWF forecasts. *ECMWF Technical Memorandum* No. 918. <https://doi.org/10.21957/52f2f31351>

Hewson, T. & M. Chevallier, 2024: Use and Verification of ECMWF products. *ECMWF Technical Memorandum* No. 919. <https://doi.org/10.21957/734d5d4d39>

Lalaurette, F. & G. van der Grijn, 2003: Ensemble forecasts: Can they provide useful early warnings? *ECMWF Newsletter* No. 96, 10–18. <https://www.ecmwf.int/en/elibrary/78236-newsletter-no-96-winter-200203>

Lamy-Thépaut, S., C. Sahin & B. Raoult, 2013: ecCharts service. *ECMWF Newsletter* No. 134, 7–9. <https://www.ecmwf.int/en/elibrary/78196-newsletter-no-134-winter-201213>

Lledó L., T. Haiden, J. Schröttle & R. Forbes, 2023: Scale-dependent verification of precipitation and cloudiness at ECMWF. *ECMWF Newsletter* No. 174, 18–22. <https://doi.org/10.21957/c92Ioli749>

Magnusson, L., S. Majumdar, R. Emerton, D. Richardson, M. Alonso Balmaseda, C. Baugh et al., 2021: Tropical cyclone activities at ECMWF. *ECMWF Technical Memorandum* No. 888. <https://doi.org/10.21957/zzxzzygwv>

Vitart, F., F. Prates, A. Bonet & C. Sahin, 2012: New tropical cyclone products on the web. *ECMWF Newsletter* No. 130, 17–23. <https://doi.org/10.21957/ti1191e2>

Zsótér, E., 2006: Recent developments in extreme weather forecasting. *ECMWF Newsletter* No. 107, 8–17. <https://doi.org/10.21957/kl9821hnc7>

MACHINE LEARNING FOR NUMERICAL WEATHER PREDICTION

Alexe, M., E. Boucher, P. Lean, E. Pinnington, P. Laloyaux, A. McNally et al., 2024: GraphDOP: Towards skilful data-driven medium-range weather forecasts learnt and initialised directly from observations. <https://arxiv.org/abs/2412.15687>

Allen, A., S. Markou, W. Tebbutt, J. Requeima, W. P. Bruinsma, T. R. Andersson et al., 2025. End-to-end data-driven weather prediction. *Nature*, **641**(8065), pp.1172–1179. <https://doi.org/10.1038/s41586-025-08897-0>

Ben Bouallègue, Z., M. C. Clare, L. Magnusson, E. Gascon, M. Maier-Gerber, M. Janoušek et al., 2024: The rise of data-driven weather forecasting: A first statistical assessment of machine learning–based weather forecasts in an operational-like context. *Bulletin of the American Meteorological Society*, **105**(6), pp.E864–E883.

Bi, K., L. Xie, H. Zhang et al., 2023: Accurate medium-range global weather forecasting with 3D neural networks. *Nature* **619**, 533–538. <https://doi.org/10.1038/s41586-023-06185-3>

Bonavita, M. & P. Laloyaux, 2020: Machine learning for model error inference and correction. *Journal of Advances in Modeling Earth Systems*, **12**, e2020MS002232. <https://doi.org/10.1029/2020MS002232>

Bodnar, C., W. P. Bruinsma, A. Lucic et al., 2025: A foundation model for the Earth system. *Nature*, **641**, 1180–1187. <https://doi.org/10.1038/s41586-025-09005-y>

Chantry, M., S. Hatfield, P. Dueben, I. Polichtchouk & T. Palmer, 2021a: Machine learning emulation of gravity wave drag in numerical weather forecasting. *Journal of Advances in Modeling Earth Systems*, **13**(7), p.e2021MS002477.

Chantry, M., H. Christensen, P. Dueben & T. Palmer, 2021b: Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI. *Philosophical Transactions of the Royal Society A* **379**, no. 2194: 20200083.

Chevallier, F., F. Chéruy, N. A. Scott & A. Chédin, 1998: A Neural Network Approach for a Fast and Accurate Computation of a Longwave Radiative Budget. *J. Appl. Meteorol.*, **37**, 1385–1397. [https://doi.org/10.1175/1520-0450\(1998\)037<1385:ANN_AFA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1998)037<1385:ANN_AFA>2.0.CO;2)

Dahoui, M., 2023: Use of machine learning for the detection and classification of observation anomalies. *ECMWF Newsletter* No. 174, 23–27. <https://doi.org/10.21957/n64md0xa5d>

Dueben, P. D. & P. Bauer, 2018: Challenges and design choices for global weather and climate models based on machine learning.

Geosci. Model Dev., **11**, 3999–4009. <https://doi.org/10.5194/gmd-11-3999-2018>

Dueben, P., U. Modigliani, A. Geer, S. Siemen, F. Pappenberger, P. Bauer et al., 2021: Machine learning at ECMWF: A roadmap for the next 10 years. *ECMWF Technical Memorandum* No 878. <https://doi.org/10.21957/ge7ckgmn>

Geer, A. J., 2021: Learning earth system models from observations: machine learning or data assimilation? *Philosophical Transactions of the Royal Society A* **379**, no. 2194: 20200089. <https://doi.org/10.1098/rsta.2020.0089>

Geer, A. J., 2024: Simultaneous inference of sea ice state and surface emissivity model using machine learning and data assimilation. *Journal of Advances in Modeling Earth Systems*, **16**(7), e2023MS004080. <https://doi.org/10.1029/2023MS004080>

Hatfield, S., M. Chantry, P. Dueben, P. Lopez, A. Geer & T. Palmer, 2021: Building tangent-linear and adjoint models for data assimilation with neural networks. *Journal of Advances in Modeling Earth Systems*, **13**, e2021MS002521. <https://doi.org/10.1029/2021MS002521>

Hewson, T. & F. Pilloso, 2020: A new low-cost technique improves weather forecasts across the world. *arXiv*:2003.14397v1

Husain, S. Z., L. Separovic, J. F. Caron, R. Aider, M. Buehner, S. Chamberland et al., 2025: Leveraging data-driven weather models for improving numerical weather prediction skill through large-scale spectral nudging. *Weather and Forecasting*, **40**(9), pp.1749–1771.

Keisler, R., 2022: Forecasting global weather with graph neural networks. *arXiv preprint arXiv:2202.07575*

Lam, R. et al., 2023: Learning skillful medium-range global weather forecasting. *Science* **382**, 1416–1421. <https://doi.org/10.1126/science.adi2336>

Lang, S., M. Alexe, M. Chantry, J. Dramschi, F. Pinault, Raoult, B. et al., 2024a: AIFS--ECMWF’s data-driven forecasting system. *arXiv preprint arXiv:2406.01465*

Lang, S., M. Alexe, M. C. Clare, C. Roberts, R. Adewoyin, Z. B. Bouallègue et al., 2024b: AIFS-CRPS: ensemble forecasting using a model trained with a loss function based on the continuous ranked probability score. *arXiv preprint arXiv:2412.15832*

McNally, A., C. Lessig, P. Lean, E. Boucher, M. Alexe, E. Pinnington et al., 2024: Data driven weather forecasts trained and initialised directly from observations. *arXiv preprint arXiv:2407.15586*

Nipen, T.N., H. H. Haugen, M. S. Ingstad, E. M. Nordhagen, A. F. S. Salihi, P. Tedesco et al., 2024: Regional data-driven weather modeling with a global stretched-grid. *arXiv preprint arXiv:2409.02891*

Kent, C., A. A. Scaife, N. J. Dunstone, D. Smith, S. C. Hardiman, T. Dunstan & O. Watt-Meyer, 2025: Skilful global seasonal predictions from a machine learning weather model trained on reanalysis data. *npj Climate and Atmospheric Science*, **8**(1), p.314.

Rackow, T., N. Koldunov, C. Lessig, I. Sandu, M. Alexe, M. Chantry & T. Jung, 2024: Robustness of AI-based weather forecasts in a changing climate. *arXiv preprint arXiv:2409.18529*

Rasp, S., P. D. Dueben, S. Scher, J. A. Weyn, S. Mouatadid & N. Thuerey, 2020: WeatherBench: A benchmark data set for data-driven weather forecasting. *Journal of Advances in Modeling Earth Systems*, **12**, e2020MS002203. <https://doi.org/10.1029/2020MS002203>

Rodríguez-Fernández, N., P. de Rosnay, C. Albergel, P. Richaume, F. Aires, C. Prigent & Y. Kerr, 2019: SMOS Neural Network Soil Moisture Data Assimilation in a Land Surface Model and Atmospheric Impact. *Remote Sens*, **11**, 1334.

Sun, Y.Q., P. Hassanzadeh, M. Zand, A. Chattopadhyay, J. Weare & D. S. Abbot, 2025: Can AI weather models predict out-of-distribution gray swan tropical cyclones? *Proceedings of the National Academy of Sciences*, **122**(21), e2420914122. <https://doi.org/10.1073/pnas.2420914122>

Watt-Meyer, O., B. Henn, J. McGibbon, S. K. Clark, A. Kwa, W. A. Perkins et al., 2025: ACE2: accurately learning subseasonal to decadal atmospheric variability and forced responses. *npj Climate and Atmospheric Science*, **8**(1), 205.







ECMWF is the European Centre for Medium-Range Weather Forecasts.

We are both a research institute and a 24/7 operational service, producing global numerical weather predictions and other data for our Member and Co-operating States and the broader community. We operate a world-class supercomputer facility for weather forecasting and hold one of the largest meteorological data archives. Other strategic activities include delivering advanced training and assisting the WMO in implementing its programmes.

We are a key player in Copernicus, the Earth Observation component of the European Union's Space programme, offering quality-assured information on climate change (Copernicus Climate Change Service), atmospheric composition (Copernicus Atmosphere Monitoring Service), flooding and fire danger (Copernicus Emergency Management Service), and through the EU's Destination Earth initiative, we are developing prototype digital twins of the Earth.