High Performance Computing for Numerical Weather Prediction

Isabella Weger

European Centre
for Medium-Range Weather Forecasts

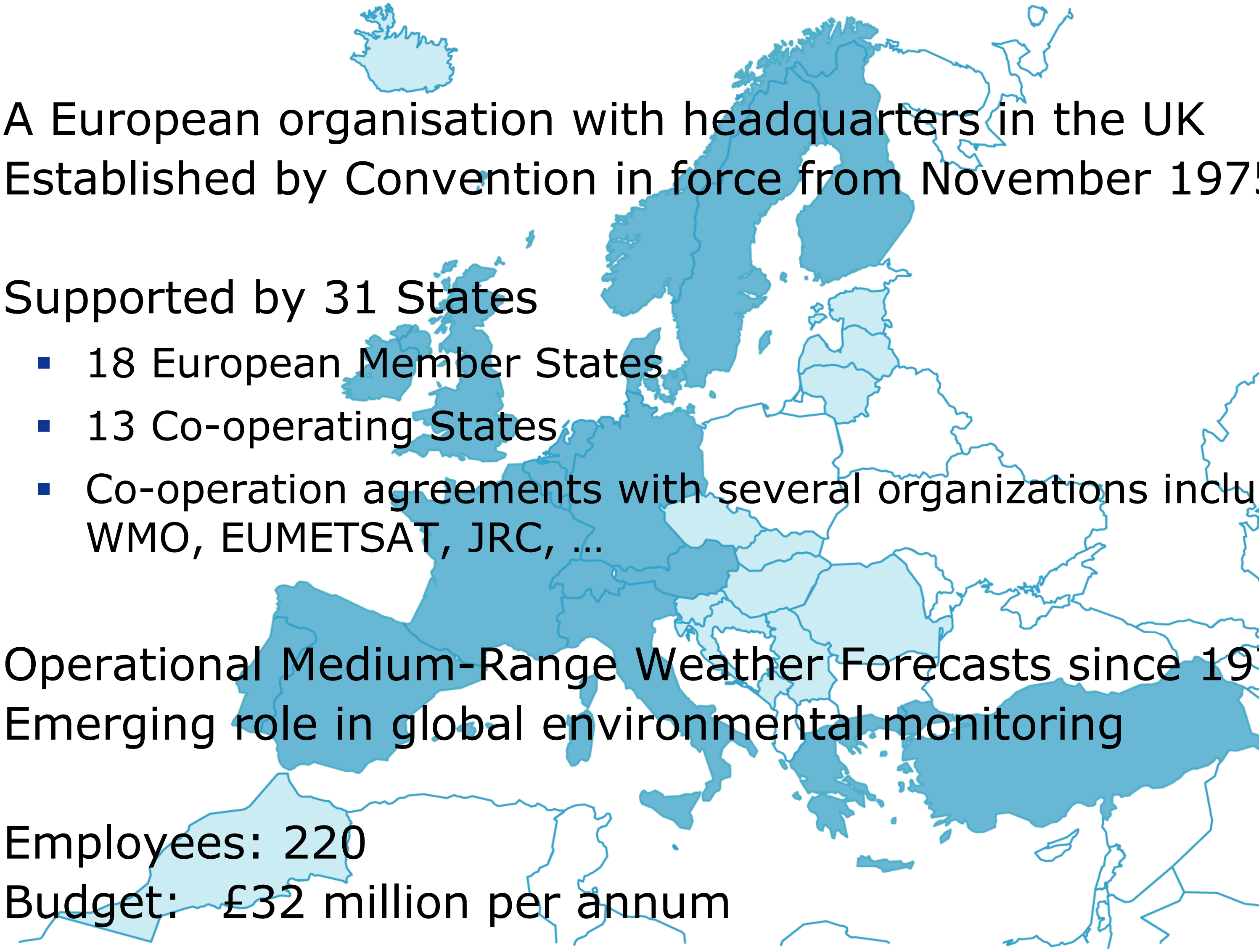
ECMWF

- A European organisation with headquarters in the UK
- Established by Convention in force from November 1975

- Supported by 31 States
  - 18 European Member States
  - 13 Co-operating States
  - Co-operation agreements with several organizations including WMO, EUMETSAT, JRC, …

- Operational Medium-Range Weather Forecasts since 1979
- Emerging role in global environmental monitoring

- Employees: 220
- Budget: £32 million per annum



ECMWF Objectives

- Operational forecasting up to 15 days ahead (including waves)
- Research and development activities in forecast modelling
- Data archiving and related services, collection and storage of appropriate meteorological data
- Operational forecasts for the coming month and season
- Advanced NWP training
- Provision of supercomputer resources
- Provision of archival/retrieval facilities
- Assistance to WMO programmes
- Management of Regional Meteorological Data Communications Network (RMDCN)

ECMWF’s strategy 2006 - 2015

- The principal goal of ECMWF in the coming ten years will be to maintain the current, rapid rate of improvement of its global, medium-range weather forecasting products, with particular effort on early warnings of severe weather events
- Complementary goals are:
  - To improve the quality and scope of monthly and seasonal-to-interannual forecasts
  - To enhance support to Member States national forecasting activities by providing suitable boundary conditions for limited-area models
  - To deliver real-time analyses and forecasts of atmospheric composition
  - To carry out climate monitoring through regular re-analyses of the Earth-system
  - To contribute towards the optimization of the Global Observing System

Types of Atmospheric Prediction

- Short-range weather forecast (0-2 days ahead)
  - Detailed prediction – regional forecasting system
  - Produce forecast an hour or two after observations are made
- Medium-range weather forecast (2 days – 2 weeks ahead)
  - Less detailed prediction – global forecasting system
  - Produce forecast up to 6 hours after observations are made
- Long-range weather forecast (more than 2 weeks ahead)
  - Predict statistics of weather for coming month or season
- Climate prediction
  - Predict what the climate would be like if . . .

ECMWF Forecast Products

Atmosphere global forecasts

- Forecast to ten days from 00 and 12 UTC at 25 km resolution and 91 levels
- 50 ensemble forecasts to fifteen days from 00 and 12 UTC at 50 km resolution

Ocean wave forecasts

- Global forecast to ten days from 00 and 12 UTC at 50 km resolution
- European waters forecast to five days from 00 and 12 UTC at 25 km resolution

Monthly forecasts: Atmosphere-ocean coupled model

- Global forecasts to one month:

  atmosphere: 1.125° resolution, 62 levels

  ocean: horizontally-varying resolution (1/3 ° to 1°), 9 levels

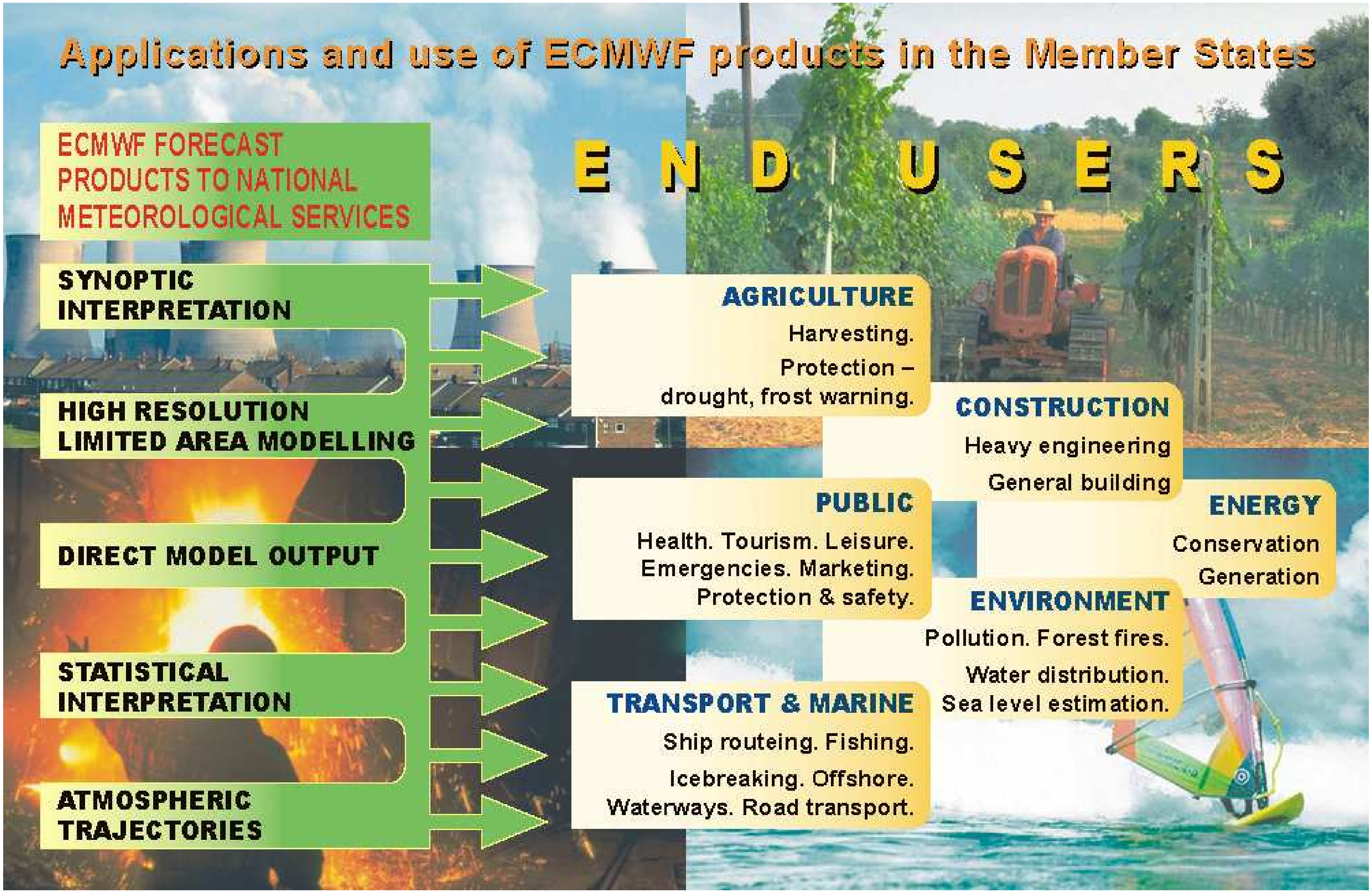
Seasonal forecasts: Atmosphere-ocean coupled model

- Global forecasts to six months:

  atmosphere: 1.8° resolution, 40 levels

  ocean: horizontally-varying resolution (1/3 ° to 1°), 9 levels

Some uses of our forecasts

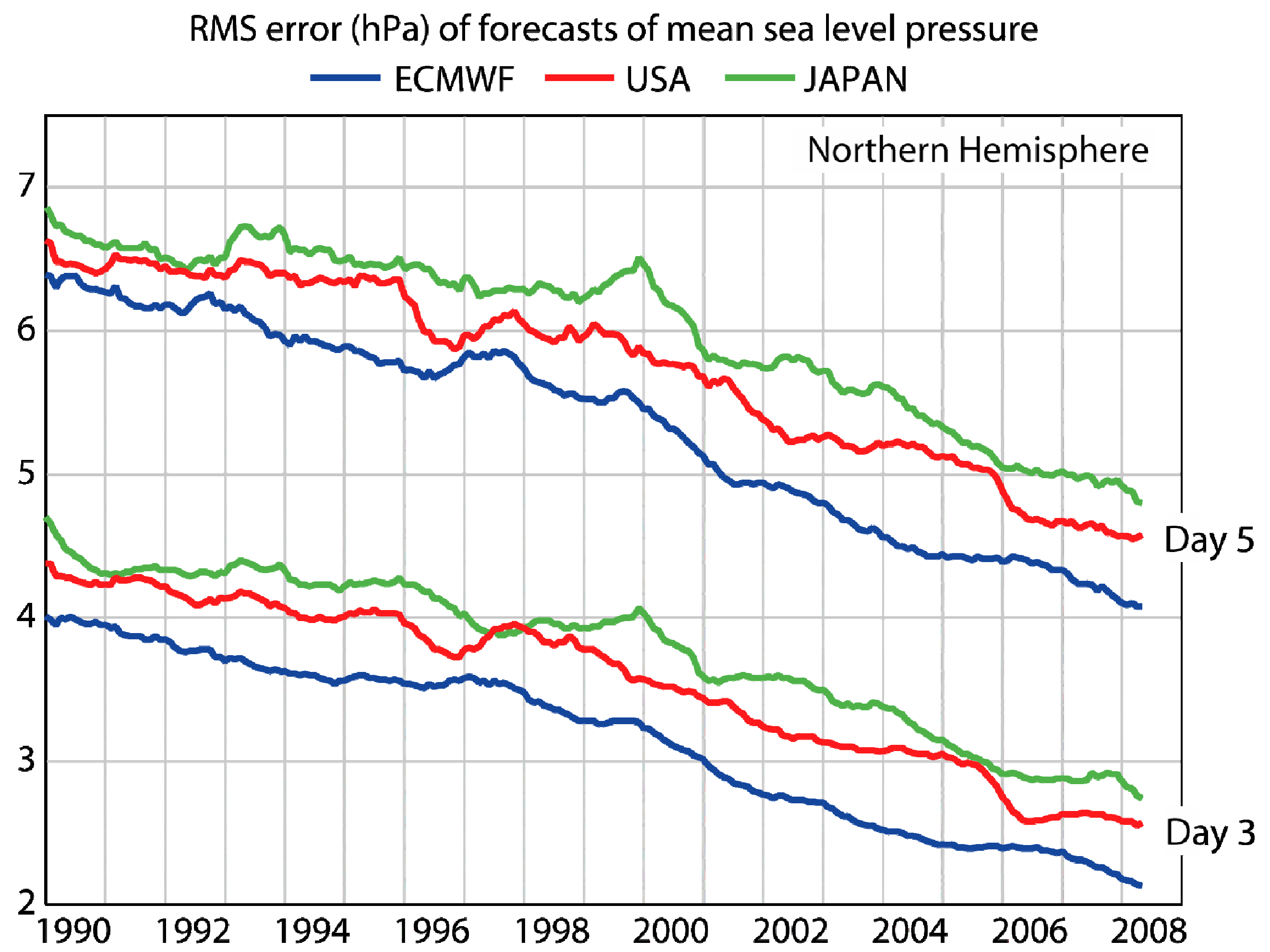


Baseline operational system: 2008

- Four-dimensional variational data assimilation based on 25 / 80 km horizontal resolution and 91-level vertical resolution (4D-Var).
- 25 km 91-level model for single deterministic forecast (twice daily).
- 50 km model (62 levels) to 10 days then 80 km coupled to an ocean model out to 15 days (EPS – run twice daily).
- EPS extended out to 1 month for monthly forecast run once a week.
- 125 km 62-level model coupled to an ocean model (110 km midlat / 33 km tropics) run once a month out to seven months for seasonal forecasts.

(All forecasts include a coupled ocean wave model)

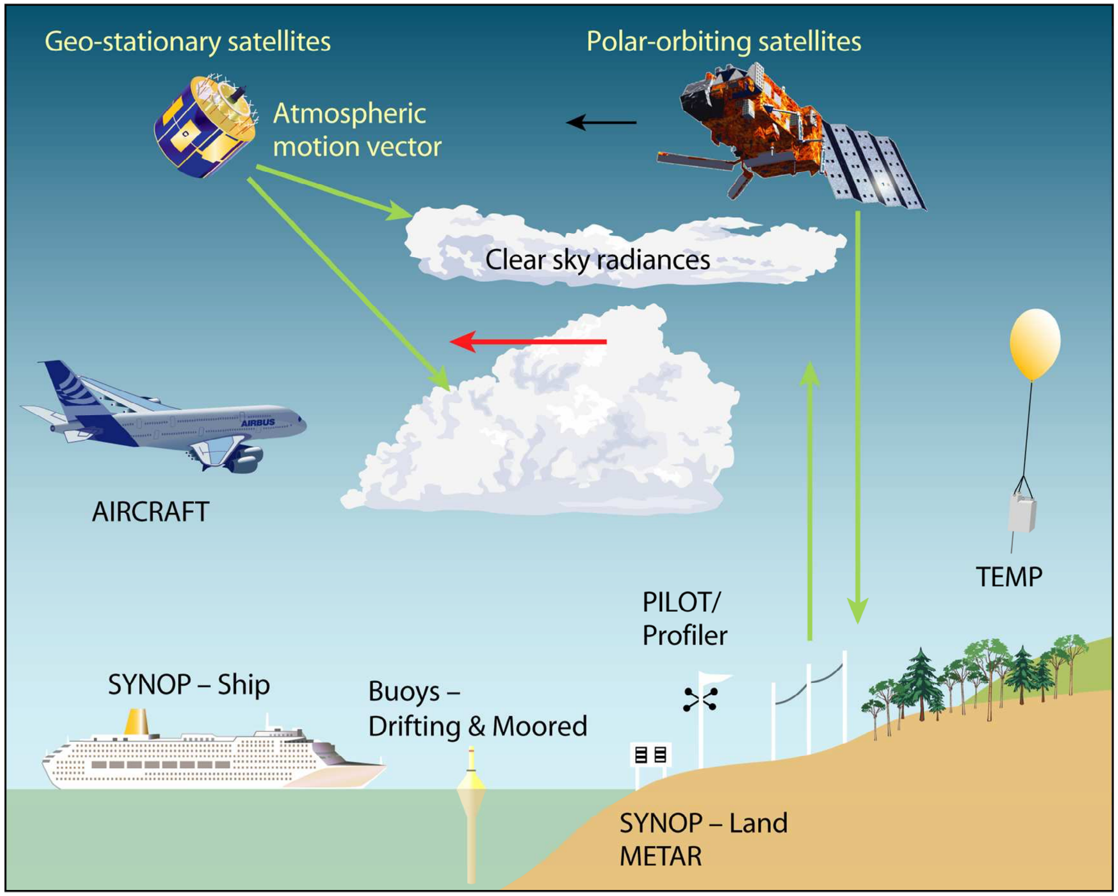
ECMWF scores compared to other major global centres



The medium-range model

Data assimilation: atmospheric fields

- To make accurate forecasts, it is important to know today's weather accurately all over the globe.
- Analyses of the wind, temperature, humidity, ozone and surface pressure of the atmosphere are produced by a four-dimensional variational assimilation system.



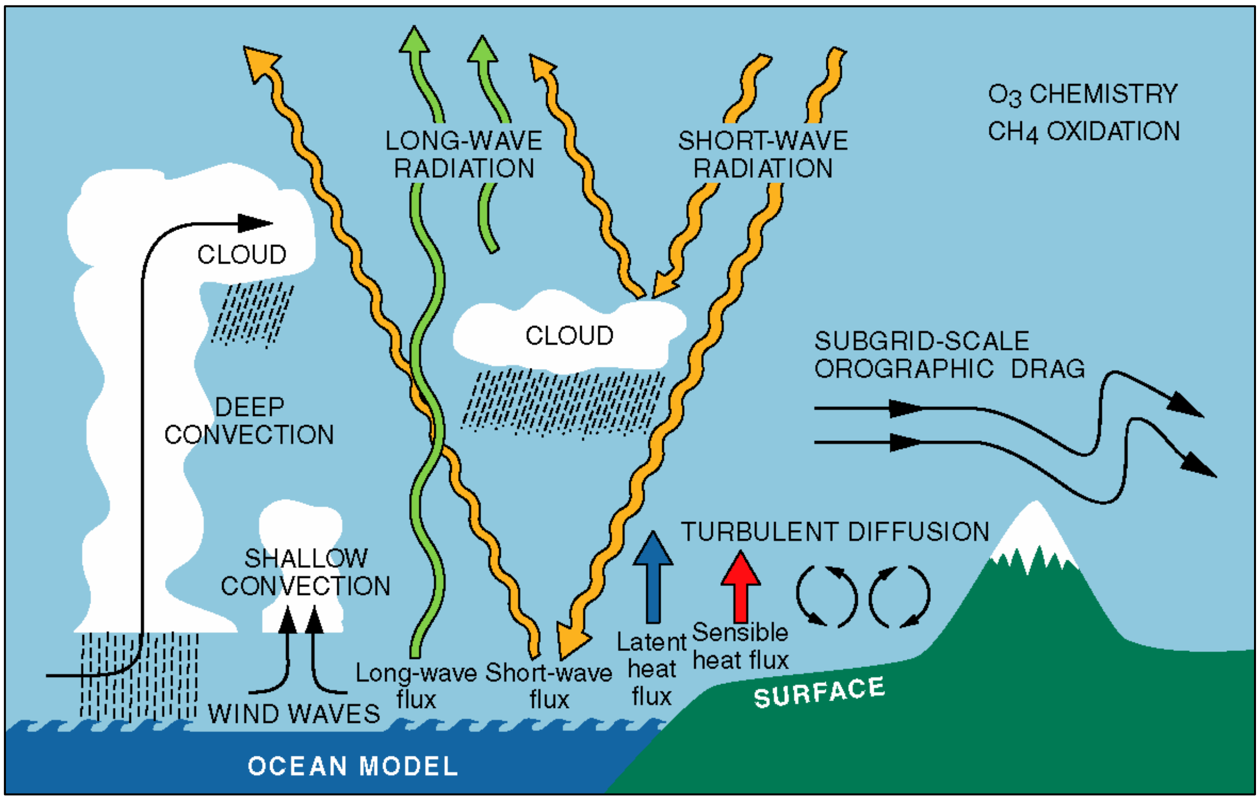
Data sources for the ECMWF Meteorological Operational System (EMOS)

The medium-range model – numerical scheme

- A spectral T_L799L91 (triangular truncation, 799 waves around a great circle on the globe; 91 levels between the earth's surface and 80 km); semi-Lagrangian formulation. The grid points in this model are separated by about 25 km in the horizontal around the globe. Models of lower resolution are used for EPS and seasonal forecasts.
- Variables at each grid point (re-calculated at each time-step): wind (including vertical velocity), temperature, humidity, cloud water and ice and cloud fraction, ozone (also pressure at surface grid-points).
  - Spacing of grid points: ~25 km
  - Time-step: 12 minutes
  - Number of grid points in model: 76,757,590
  - Number of computations required for each ten-day forecast: 1,630,000,000,000,000

The ECMWF numerical weather prediction model

The atmosphere does not evolve in isolation, interactions between the atmosphere and the underlying land and ocean are also important in determining the weather. Ocean ice processes, ocean surface waves, land surface, soil, hydrological and snow processes are all represented at ECMWF in the most advanced operational Earth-system model available anywhere.



These physical processes have smaller scales than the model grid (25 km) and are therefore represented by so-called "Parametrization Schemes" which represent the effect of the small-scale processes on the large-scale flow.

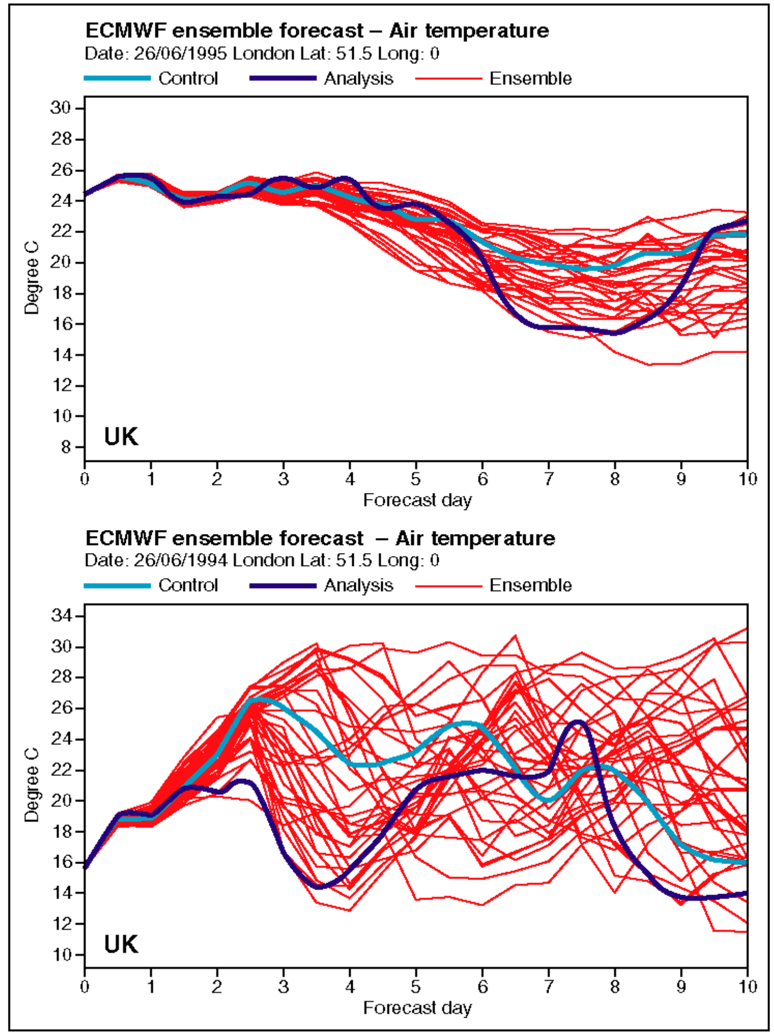
Chaos and weather prediction

The atmosphere is a chaotic system, and as a result, small errors in our estimate of the current state can grow to have a major impact on the subsequent forecast. Because of the limited number of observations available and the uneven spread of these around the globe, there is always some uncertainty in our estimate of the current state of the atmosphere. In practice this limits detailed weather prediction to about a week or so ahead.

Accepting the findings from chaos theory about the sensitivity of the prediction to uncertainties in the initial conditions, it is becoming common now to run in parallel a set, or ensemble, of predictions from different but similar initial conditions. ECMWF currently runs a 51-member ensemble.

EPS forecast for London temperature

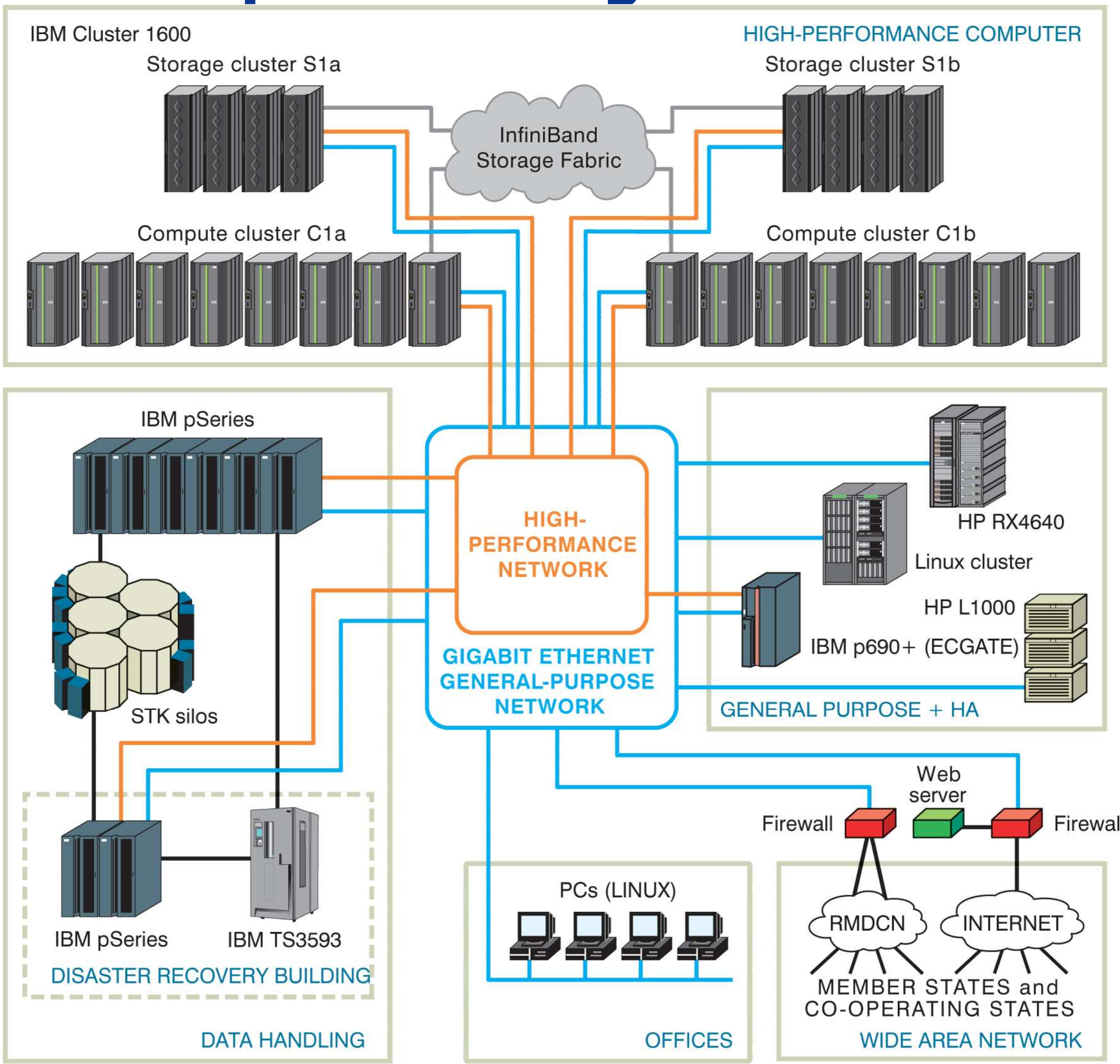
Like the Lorenz system, the predictability of the atmosphere is itself a function of initial state. This figure shows the forecasts for air temperature in London given by an ensemble of 33 forecasts. Top: a relatively predictable situation. All forecasts started on 26 June 1995 give similar forecasts; in this case, small errors in the starting conditions did not seriously affect the forecast. Bottom: an unpredictable situation. Exactly one year earlier, on 26 June 1994, small initial differences had a dramatic effect on the forecasts.



Computing at ECMWF

- ECMWF has a long history of using High Performance Computing in NWP
  - Cray machines from 1977 to 1996 (Cray-1A, XMP, YMP, C90, T3D)
  - Production work moved to distributed memory vector machines in 1996 (Fujitsu VPP700, VPP700E, VPP5000)
  - scalar cluster systems (IBM 1600) have been in production since 2002
- Large meteorological archive for research activities
- Forecasts need to be produced 365 days a year, timeliness is very critical

Computer configuration 2009



Current HPC system – IBM cluster 1600

- IBM HPCF - Phase 4 (2006 – 2009)
  - Two identical IBM AIX Clusters of p5-575 Server
  - Each cluster has 155 compute nodes with 16 Power5+ @ 1.9 GHz SMT processors
  - 32 GB memory per node (2 nodes with 128 GB)
  - 8 nodes per cluster are dedicated to I/O and networking
  - 100 TB of fibrechannel disk storage in total
  - Total sustained performance on ECMWF's codes is around 4 Tflops (based on a typical job mixture)

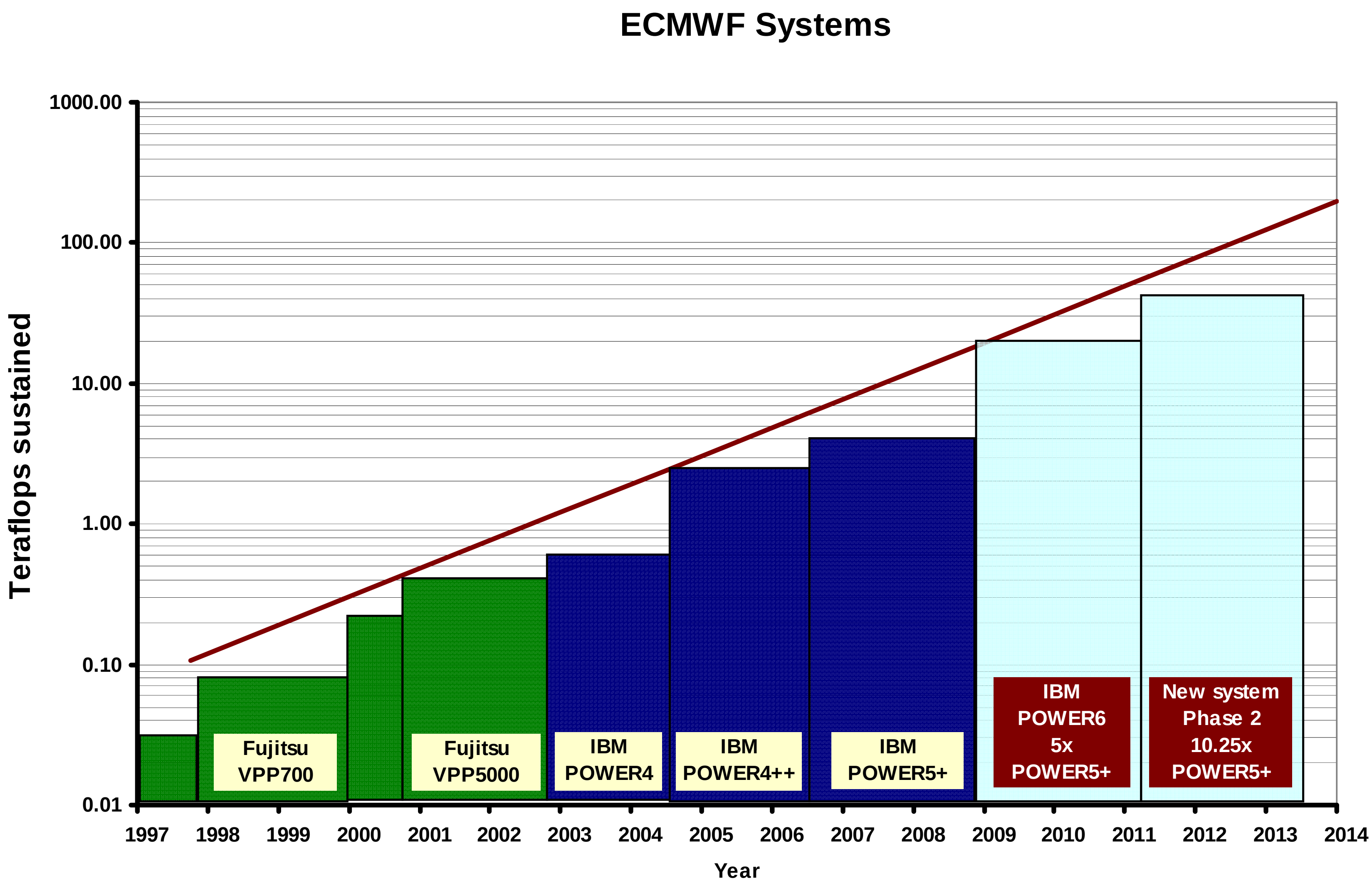
Performance requirements for the next HPCF

- Full operational implementation of the planned research developments in ECMWF’s 10-year strategy would require:
    - at least 20 TFLOPS <u>sustained</u> by April 2009
    - between 150 and 200 TFLOPS <u>sustained</u> by 2015 (N.B. this is measured in the performance sustained on ECMWF’s codes, not peak performance)
- ECMWF’s current HPCF sustains about 4 TFLOPS on a mixture of these application codes, so the above requirement translates to an improvement in performance of:
    - 5 in 2009 (PHASE1)
    - between 10 and 12.5 in 2011 (PHASE2)

The new HPCF solution

- Phase1: (2009 – 2011)
  - 2 compute clusters based on IBM POWER6 (p6-575) servers, with an Infiniband interconnect
  - 2 I/O storage clusters, transferring data to/from directly attached disks from/to the compute clusters over an Infiniband network
  - Performance factor = 5
- Phase2: (2011 – 2013)
  - Similar configuration to Phase1 (2 compute clusters, 2 I/O storage clusters)
  - Based on the follow-on generation of POWER processors
  - Performance factor = 10.5

Growth in HPC performance to 2013



Mflops, Gflops and Tflops at ECMWF

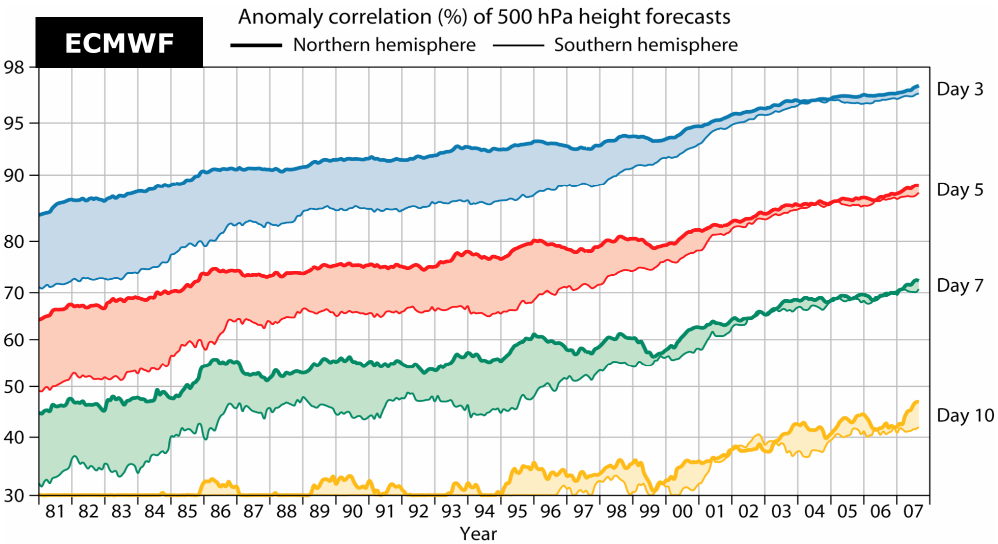
- 1979 – Cray-1:
  - 40 – 50 Mflops sustained on a single CPU
    for a 10-day forecast with 200 km grid resolution
- 1990 – Cray YMP:
  - 1 Gflops sustained on 8 CPUs
    for a 10-day forecast with 125 km grid resolution
- 2006 – IBM p5-575:
  - 2 Tflops sustained on 2112 CPUs (528 x 8 threads)
    for a 10-day forecast with 25 km grid resolution
  - A second, identical system supports other tasks

Increases in computational power have enabled:

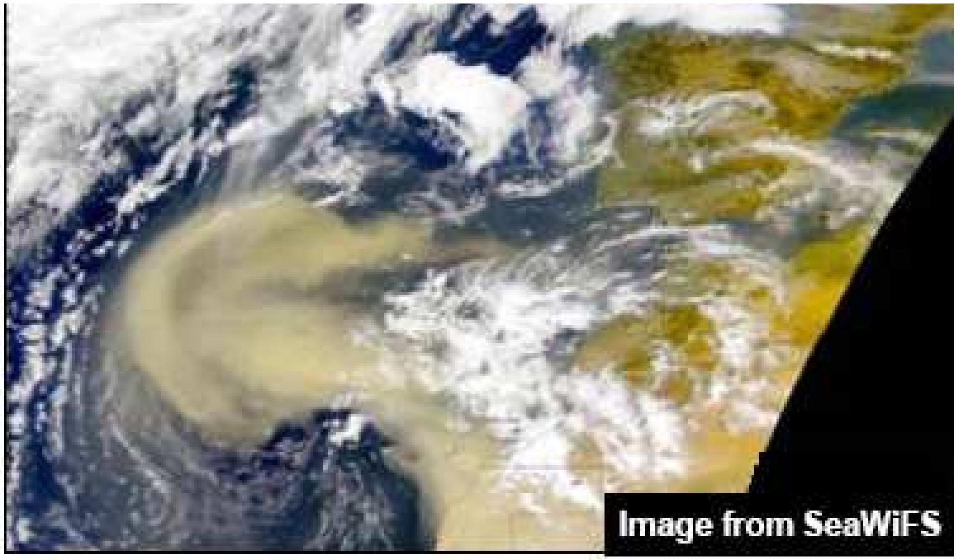
- Increased vertical as well as horizontal resolution
- Very much more sophisticated analyses of observations
- More realistic representation of atmospheric physics and land surfaces, an coupling with ocean-wave models and ocean-circulation models
- Probabilistic weather and climate forecasts based on ensemble methods
- An extended range of activities for forecasting centres:
  - Monthly and seasonal prediction
  - Reanalysis of multi-decadal observations for climate studies

What will Petaflops computing enable?

- A continued increase in the accuracy of weather forecasts

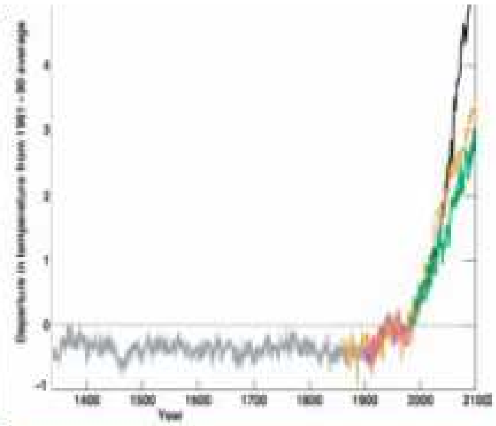


- A wider range of predicted quantities
  - Related, in particular, to air quality



What will Petaflops enable?

- More refined climate models
  - To resolve smaller-scale severe weather systems
  - To refine representation of existing processes
  - To include new processes: carbon cycles, ice-sheet dynamics, …
- More reliable answers to key questions:
  For specified emissions,
  - By how mach will the globe warm?
  - What will happen to the Arctic sea-ice?
  - … to the Greenland ice sheet?
  - … to sea-level and storminess?
  - …



What are the technical challenges?

- Effective utilization of increasing numbers of processor cores
  - Larger problem sizes help, but number of model points per core decreases as resolution increases for given execution time
  - Increases inter-node communication
  - Load balancing becomes more demanding
- Maintaining effectiveness of algorithms
- Maintaining effectiveness of long-lived model and data assimilation codes that represent major investments
  - ECMWF's Integrated Forecasting System has run operationally on
    - CRAY C90 vector shared memory
    - Fujitsu vector distributed memory
    - IBM scalar SMP clusters
  - And will run on ….. ?

ECMWF

Thank you
for your attention

www.ecmwf.int

