# Fundamentals of Data Assimilation

JCSDA Summer Colloquium on DA, Stevenson, Washington.

Andrew Lorenc, July 2009.

# Fundamentals of DA (1)

1. ***Motivation & Historical Background:***
2. ***Objective Analysis & Data Assimilation:***
3. ***Longer-term challenges:***

# *1. Motivation & Historical Background:* DA for NWP

- Much clever mathematics is available.
- Have to simplify real-world problems before we can apply it in practice.
- Which simplifications to make, and what it is important to get right, depend on insight into the application.
  **This, for me, is the most interesting bit of DA R&D, rather than the mathematics *per se*.**
- This lecture is focussing on DA for NWP. I assume that:
  **The best data assimilation scheme is that which leads to the best forecast.**

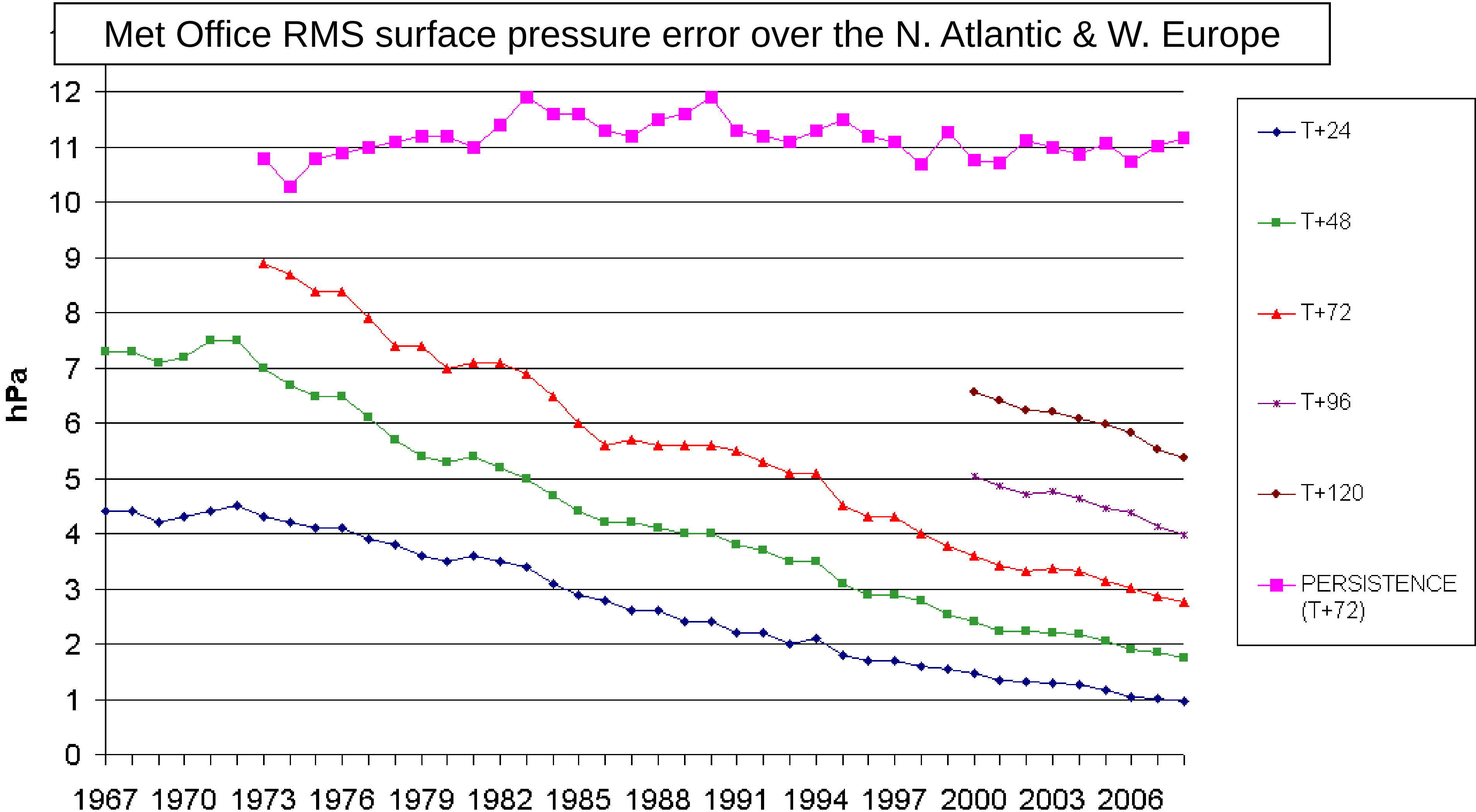
# *Historical Background:*
What has been important for getting the best NWP forecast?

NWP systems are improving by 1 day of predictive skill per decade. This has been due to:

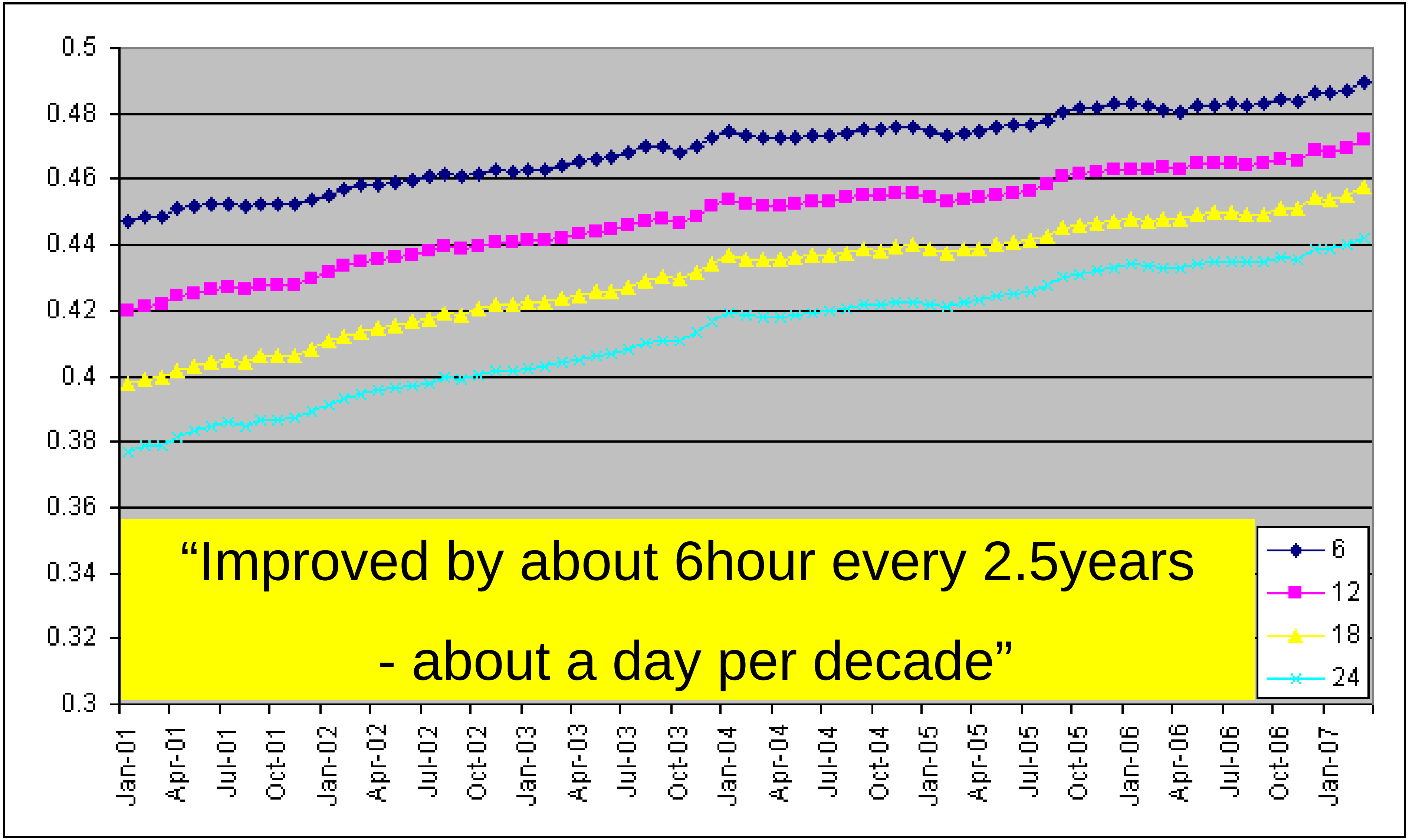
1. *Model improvements, especially resolution.*
2. *Careful use of forecast & observations, allowing for their information content and errors. Achieved by variational assimilation e.g. of satellite radiances.*
3. *4D-Var.*
4. *Better observations.*

# Performance Improvements

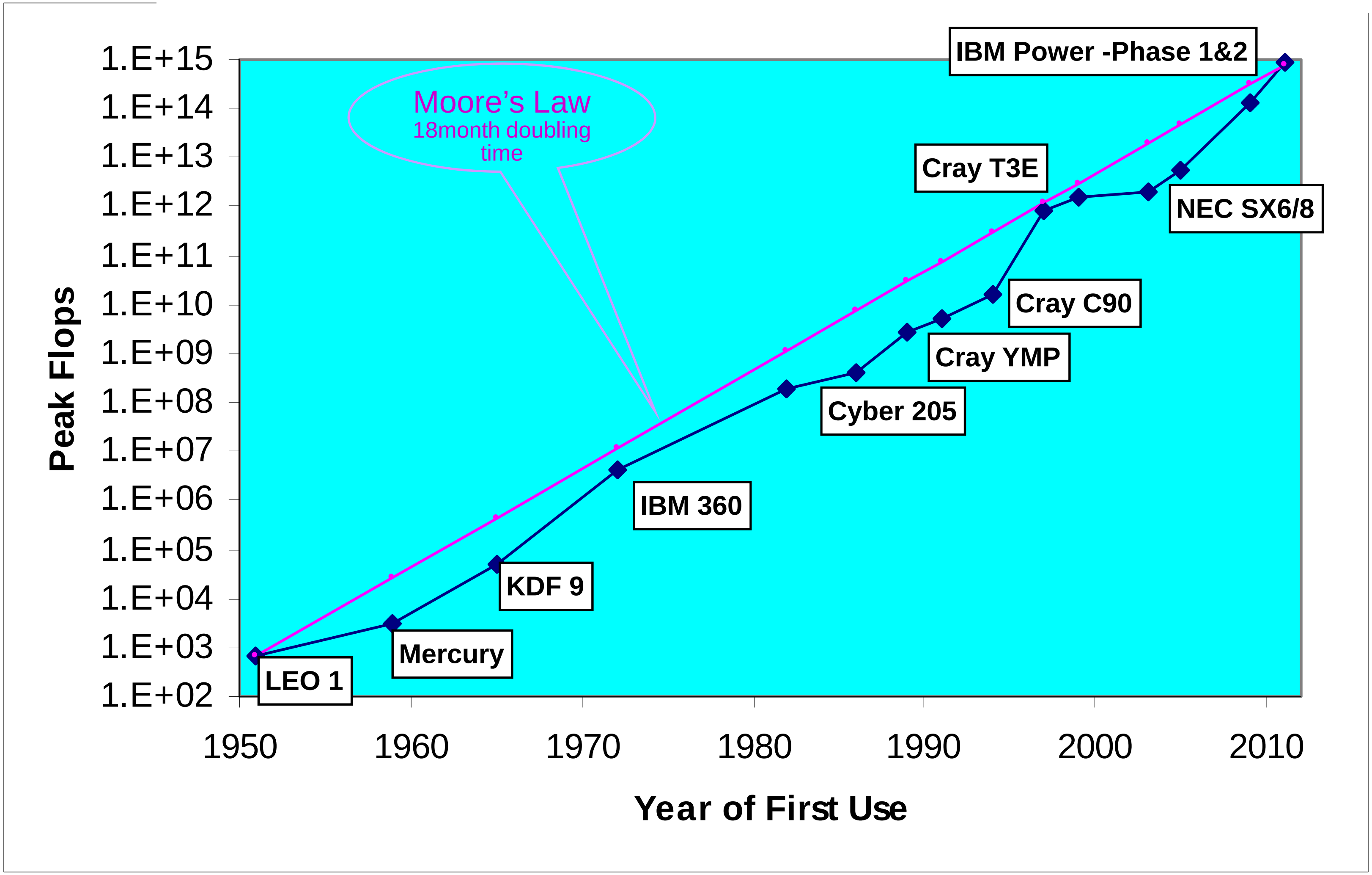
## “Improved by about a day per decade”



# UK Index Improvement:
## skill scores vs UK SYNOPS for T wind ppn cloud visibility



# 60 Years of Met Office Computers



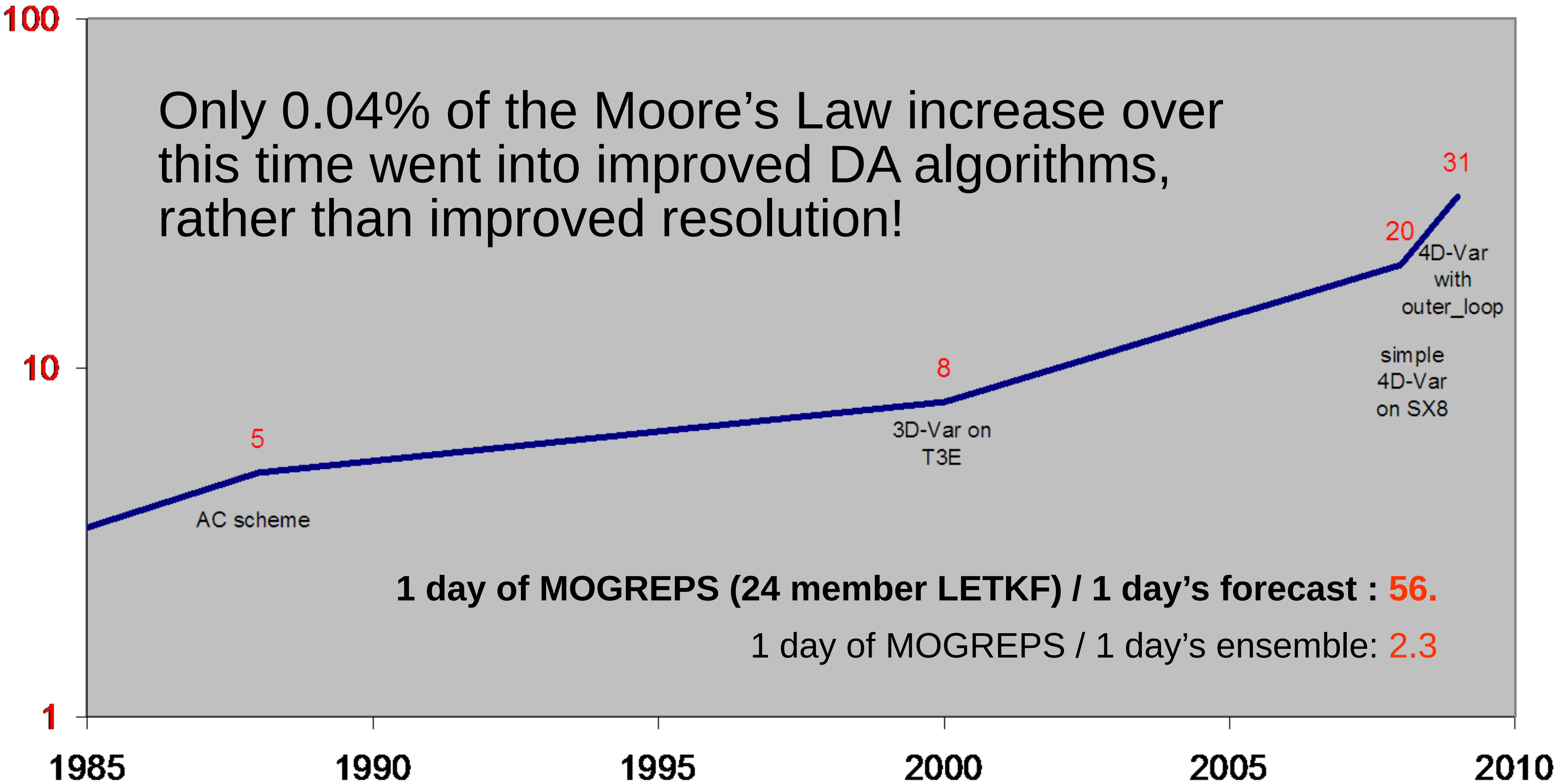
# Importance of forecast model

- A large part of the increase in ***assimilation*** accuracy comes from improvements to the model
- A large part of the increase in model accuracy comes from improvements in resolution
- The resolution has been limited by computer power, so the increase in skill is related to Moore's Law.
- Still true today – a larger part of this year's increases in computer power at the Met Office will be spent on increased resolution than on improved algorithms.

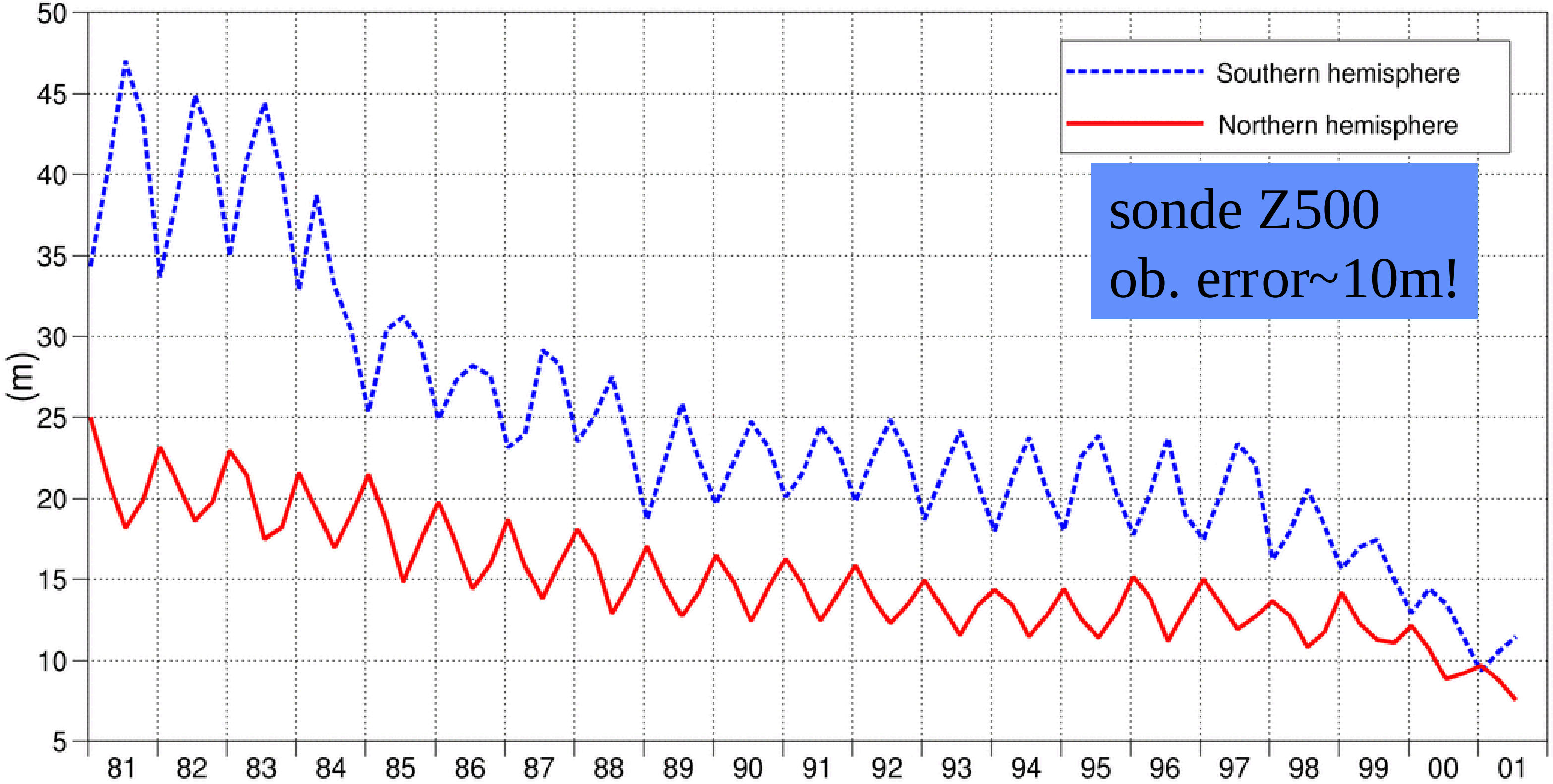
*NWP is an extreme example here.*
*Other applications of DA place less emphasis on the model and more on use of data.*

# Ratio of global computer costs:

## 1 day's DA (total incl. FC) / 1 day's forecast.



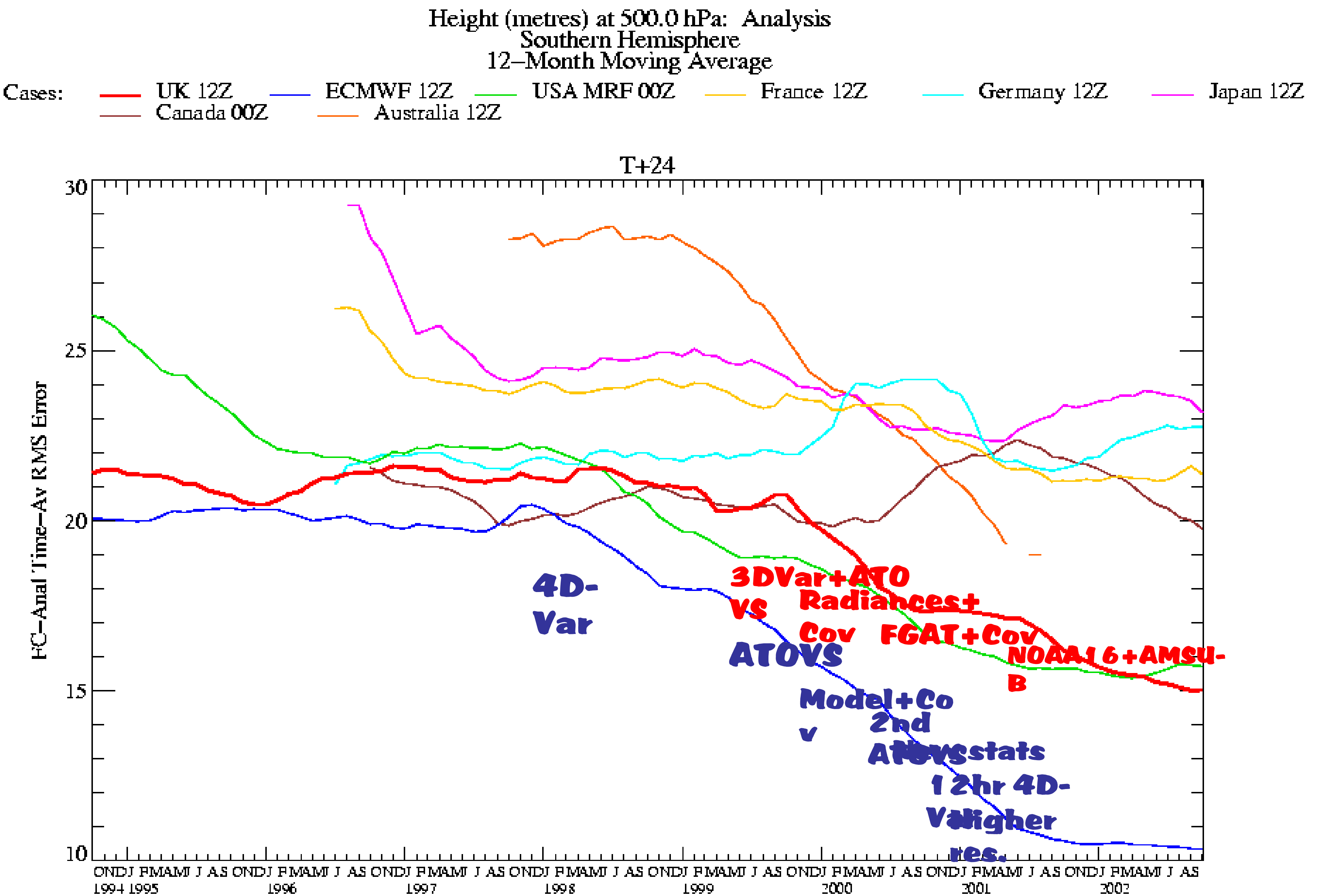
# *Evolution of the r.m.s day-one 500hPa height forecast error 1981-2001*



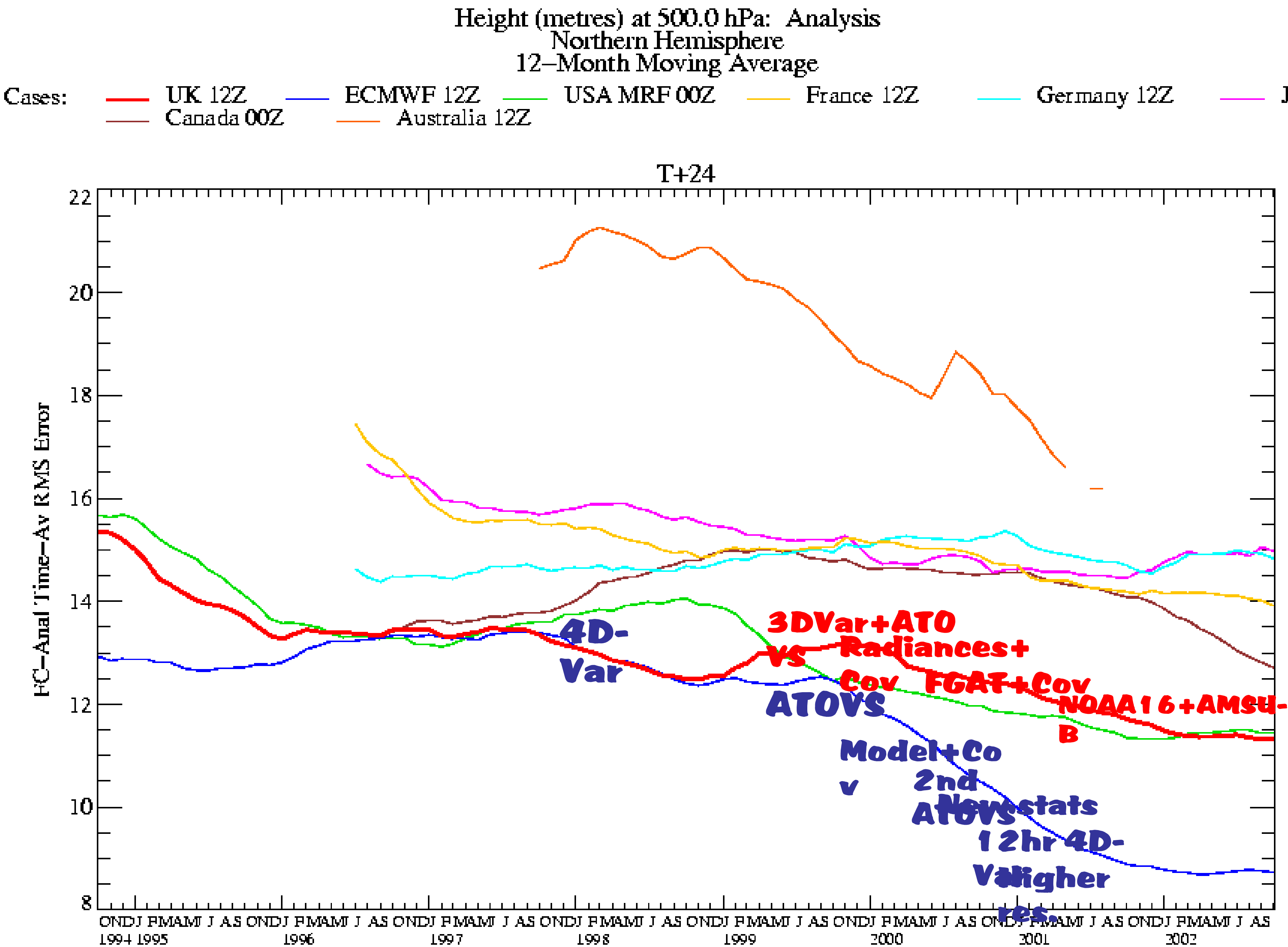

ECMWF

*Simmons & Hollingsworth, 2002*

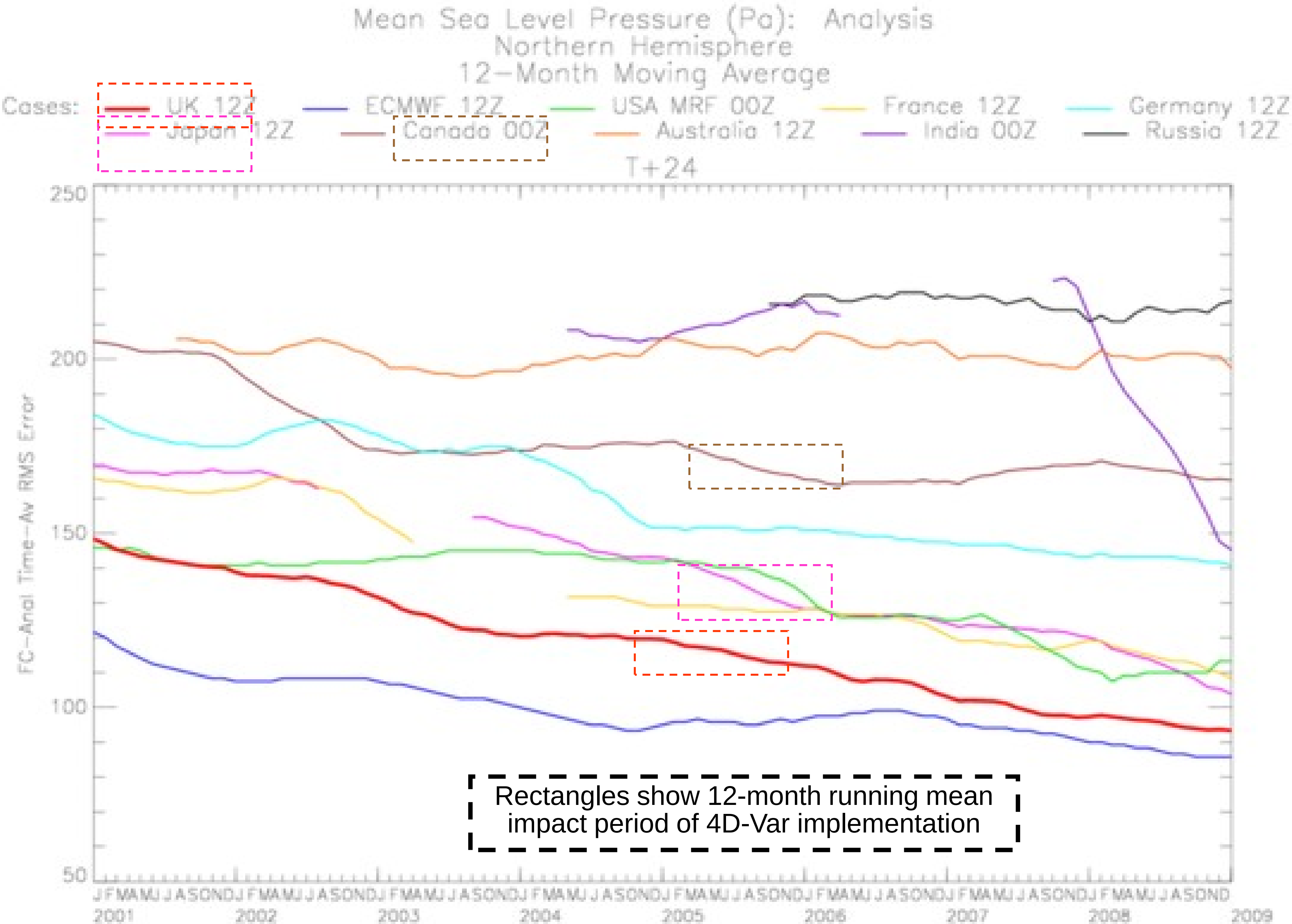
 



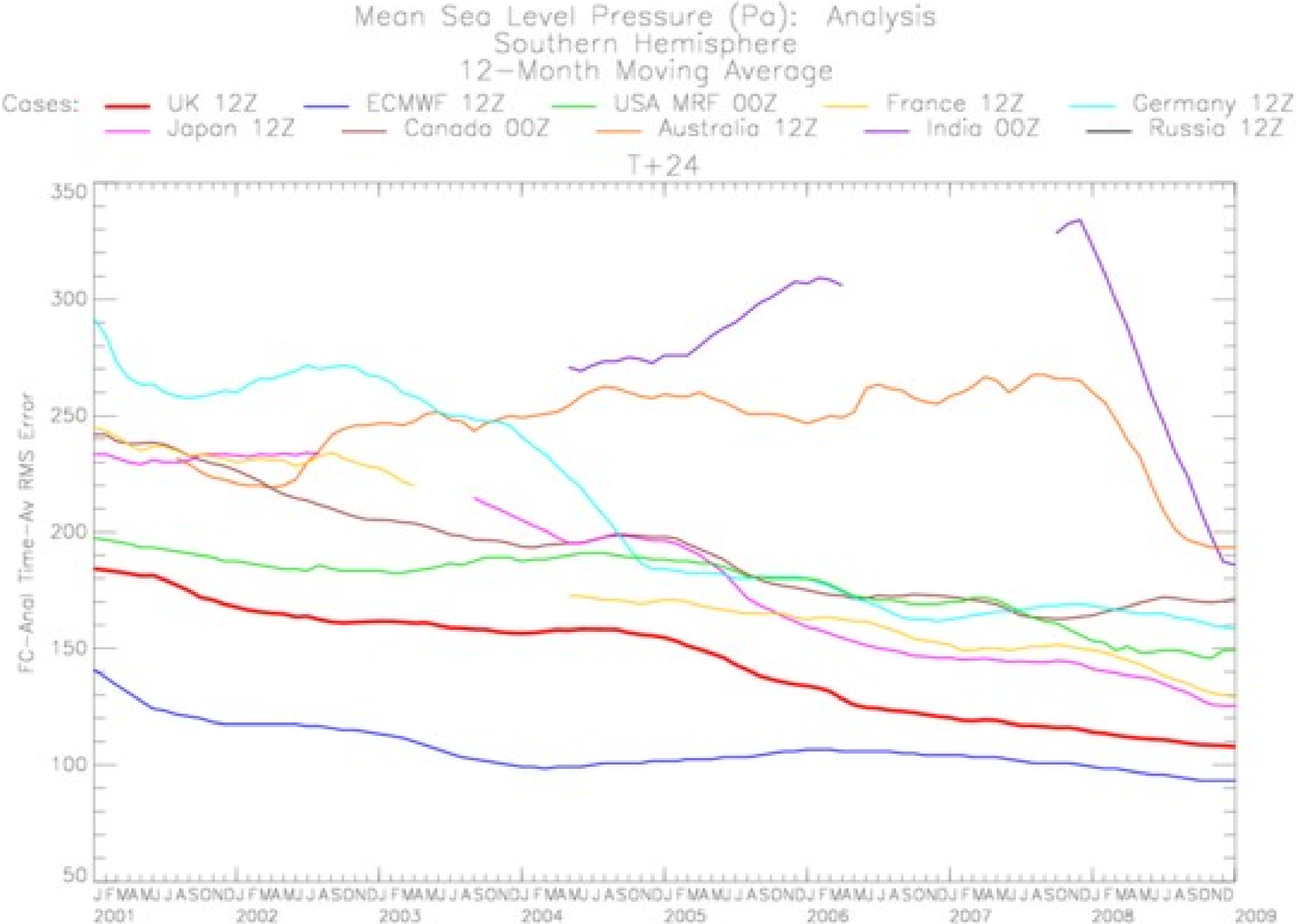
Height (metres) at 500.0 hPa: Analysis
Southern Hemisphere
12–Month Moving Average
Cases:
UK 12Z
ECMWF 12Z
USA MRF 00Z
France 12Z
Germany 12Z
Japan 12Z
Canada 00Z
Australia 12Z
T+24
FC–Anal Time–Av RMS Error
30
25
20
15
10
ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS
1994 1995
1996
1997
1998
1999
2000
2001
2002
4D-Var
3DVar+ATOVS
Radiances+Cov
FGAT+Cov
NOAA16+AMSU-B
ATOVS
Model+Cov
2nd ATOVS
New stats
12hr 4D-Var
Higher res.



Height (metres) at 500.0 hPa: Analysis
Northern Hemisphere
12–Month Moving Average
Cases:
UK 12Z
ECMWF 12Z
USA MRF 00Z
France 12Z
Germany 12Z
Japan 12Z
Canada 00Z
Australia 12Z
T+24
FC–Anal Time–Av RMS Error
22
20
18
16
14
12
10
8
ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS ONDJ FMAMJ J AS
1994 1995
1996
1997
1998
1999
2000
2001
2002
4D-Var
3DVar+ATOVS
Radiances+Cov
FGAT+Cov
NOAA16+AMSU-B
ATOVS
Model+Cov
2nd ATOVS
New stats
12hr 4D-Var
Higher res.



Mean Sea Level Pressure (Pa): Analysis
Northern Hemisphere
12-Month Moving Average
Cases: UK 12Z ECMWF 12Z USA MRF 00Z France 12Z Germany 12Z
Japan 12Z Canada 00Z Australia 12Z India 00Z Russia 12Z
T+24
FC-Anal Time-Av RMS Error
250
200
150
100
50
2001
2002
2003
2004
2005
2006
2007
2008
2009
Rectangles show 12-month running mean
impact period of 4D-Var implementation


Mean Sea Level Pressure (Pa): Analysis
Southern Hemisphere
12-Month Moving Average
Cases:
UK 12Z
ECMWF 12Z
USA MRF 00Z
France 12Z
Germany 12Z
Japan 12Z
Canada 00Z
Australia 12Z
India 00Z
Russia 12Z
T+24
FC-Anal Time-Av RMS Error
350
300
250
200
150
100
50
J F M A M J J A S O N D
2001
2002
2003
2004
2005
2006
2007
2008
2009

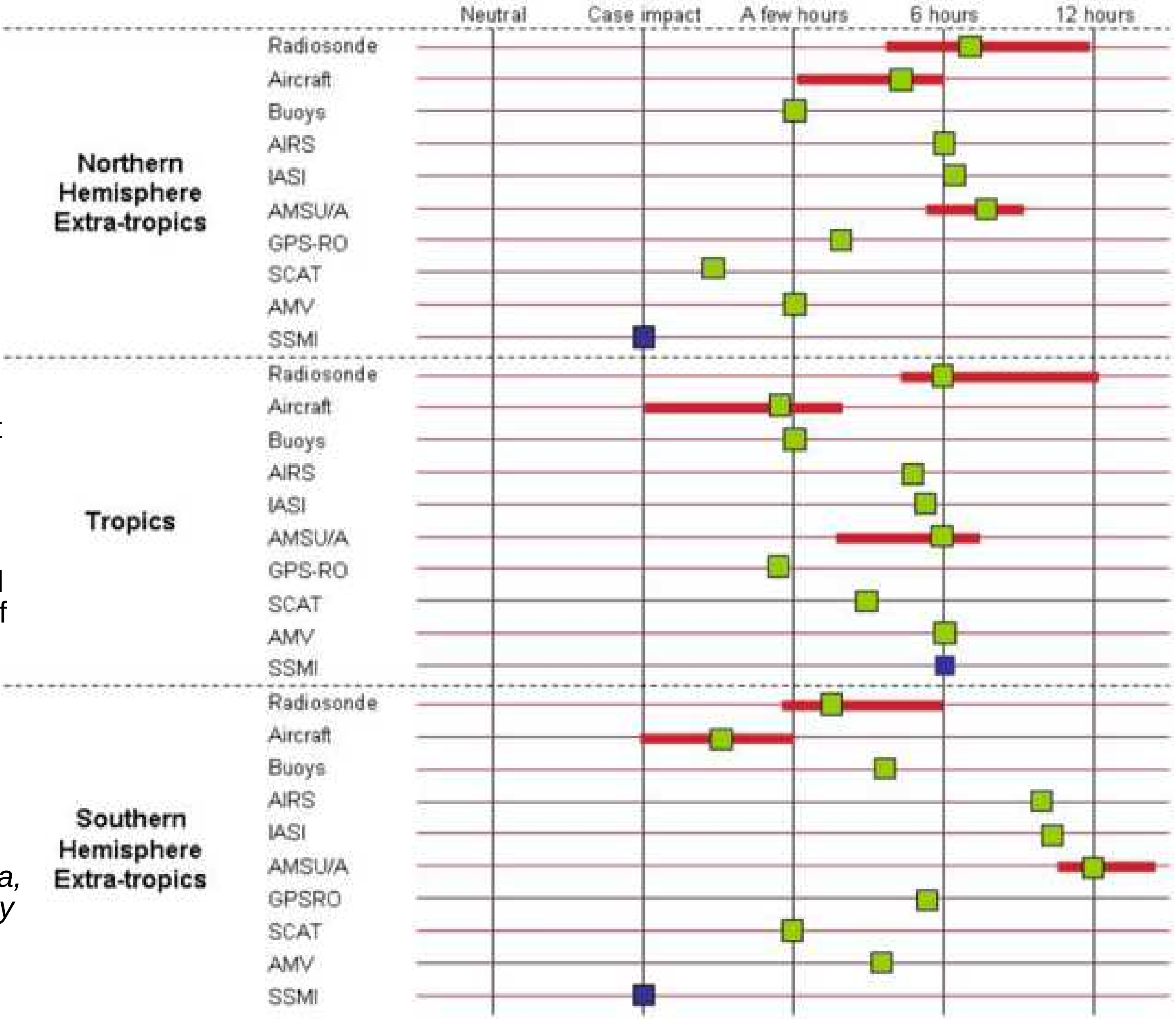
# Recent improvements are not simply due to better observations

- Whole observing systems give up to 6 hours improvement in skill (*Fourth WMO Workshop on the Impact of Various Observing Systems on NWP*). This is only equivalent to 2~3 years improvement.

# Impact of different observing systems.

Current contributions of parts of the existing observing system to the large-scale forecast skill at short and medium-range. The green colour means the impact is mainly on the mass and wind field. The blue colour means the impact is mainly on humidity field. The contribution is primarily measured on large-scale upper-air fields. The red horizontal bars give an indication of the spread of results among the different impact studies so far available.

*Fourth WMO Workshop on the Impact of Various Observing Systems on NWP. Geneva, Switzerland, 19-21 May 2008*

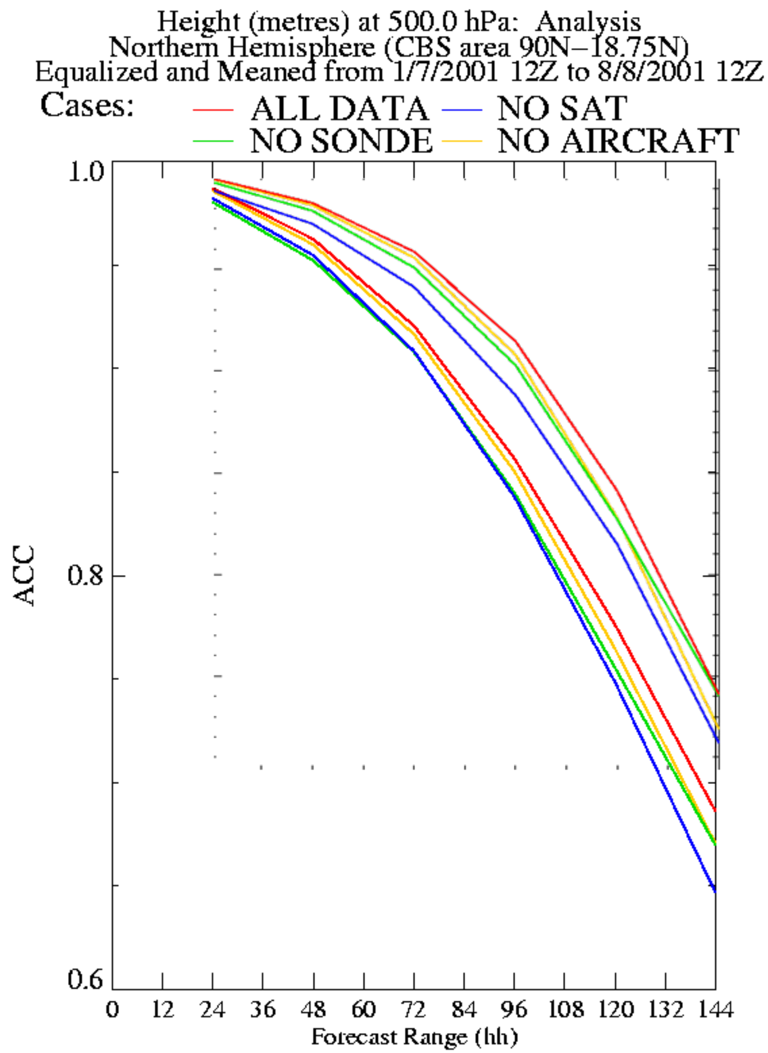


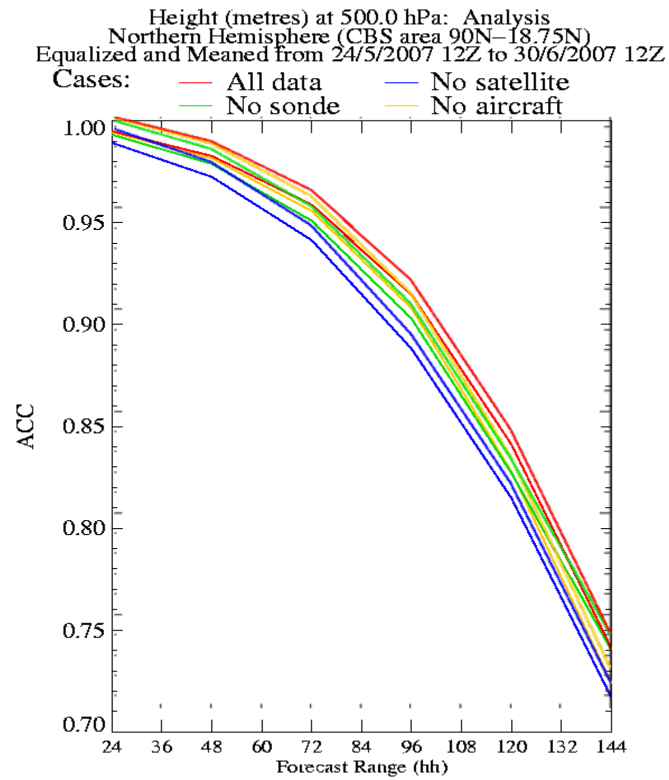
# Recent improvements are not simply due to better observations

- Whole observing systems give up to 6 hours improvement in skill (*Fourth WMO Workshop on the Impact of Various Observing Systems on NWP*). This is only equivalent to 2~3 years improvement.

- “**No satellite**” OSEs now give **better** forecasts than “**All Obs**” OSEs did 6 years ago (*Richard Dumelow*).

# Change in OSE results 2001-2007. N-hem 500hPa height ACC.




Not the same period,
so only make qualitative comparisons!

*Richard Dumelow*

# 2. From Objective Analysis to Data Assimilation

# Objective Analysis

- Bergthorsson and Doos 1955: "Numerical weather map analysis". *Tellus*
- Cressman, G.P. 1959: "An operational objective analysis scheme". *Mon. Wea. Rev.*
- Gandin, L.S. 1963: "Objective analysis of meteorological fields."

*The name "analysis" for the best estimate model state used to initialise an NWP forecast is very well established, but perhaps best estimate state, or synthesis, would be more appropriate.*

# **Analysis**: separation of a whole into its component parts

- For Bergeron, weather map analysis was rather a fine art than applied science. The analysis was, for a Bergen School connoisseur, not only a method to determine the "initial state", but also a process whereby the forecasters could familiarise himself with the weather, create an inner picture of the synoptic situation. (Persson, 2004)
- The machine computed analyses were not as readily accepted as the NWP. Quite a few 'old-timers' were disgusted by the thought that the artistic, intuitive and mystical qualities professionals were so proud of could be simulated by a computer. (Bushby, 1986)

# J. Charney, M. Halem, and R. Jastrow (1969)

*J. Atmos. Sci.* **26**, 1160-1163.

# Use of incomplete historical data to infer the present state of the atmosphere

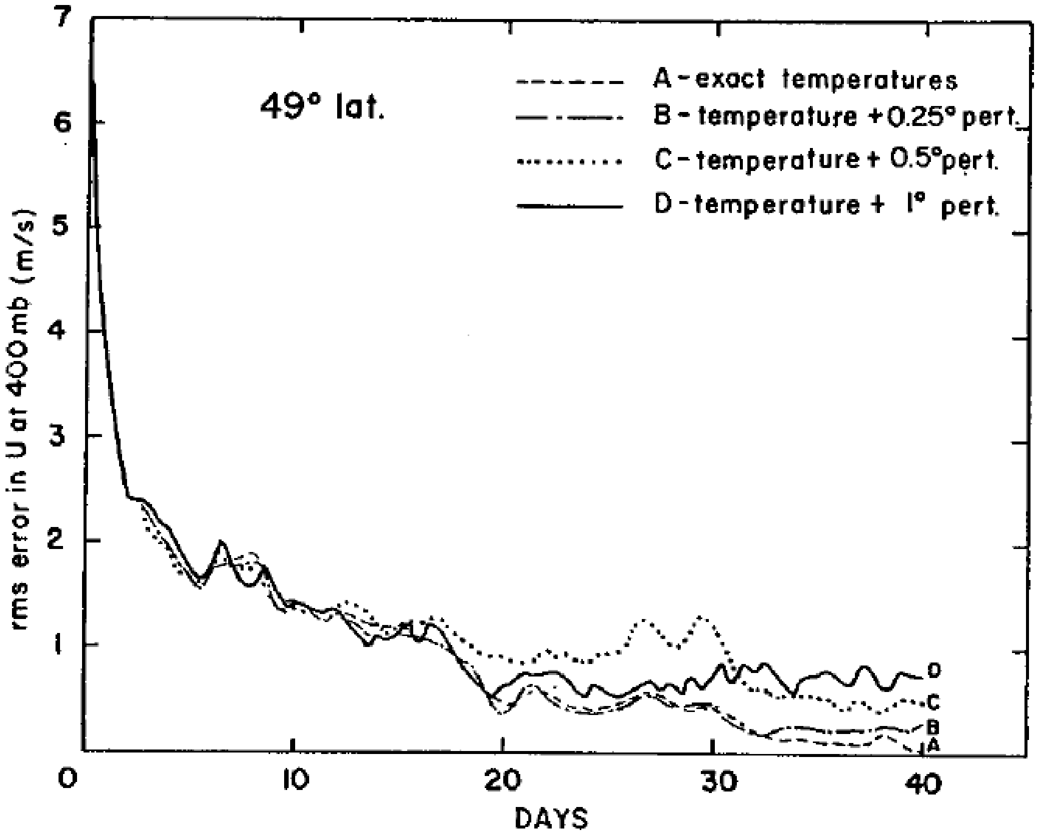



FIG. 1. The rms error in zonal wind (m sec$^{-1}$) at 400 mb at 49° latitude, in cases where temperatures with random error perturbations of 0, 0.25, 0.5 and 1C are inserted every 12 hr at all grid points.

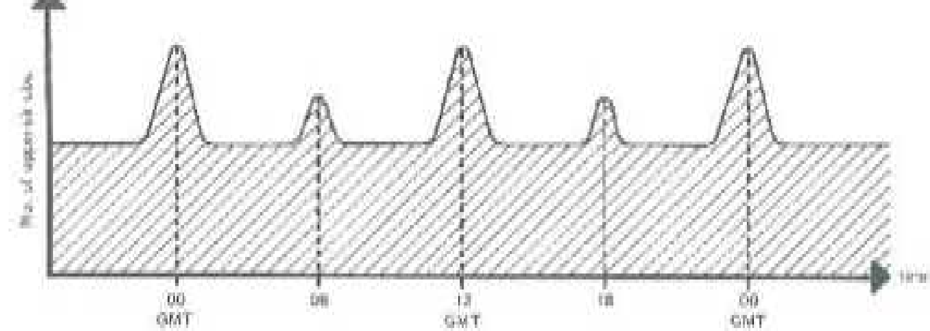
**OSSE using Mintz-Arakawa model :**
**9° × 7° × 2 levels.**

✓ ***Satellite sounders could become a major part of the global OS.***

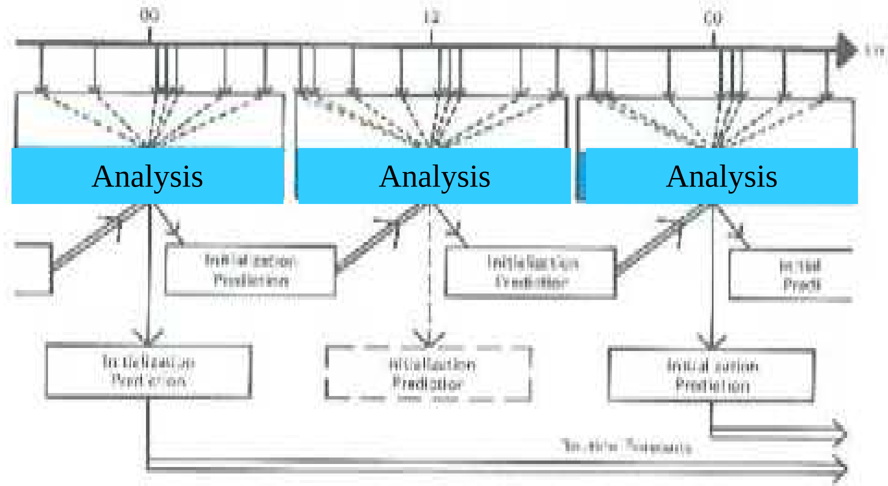
✗ ***Direct insertion of satellite temperature***

# The Four Dimensional Assimilation Problem

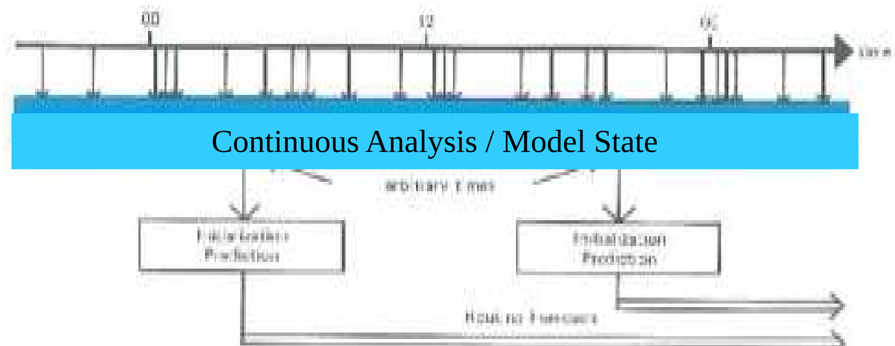
from the First GARP Global Experiment: Objectives and Plans (1973)




The number of upper-air observations as a function of time




The intermittent data assimilation scheme



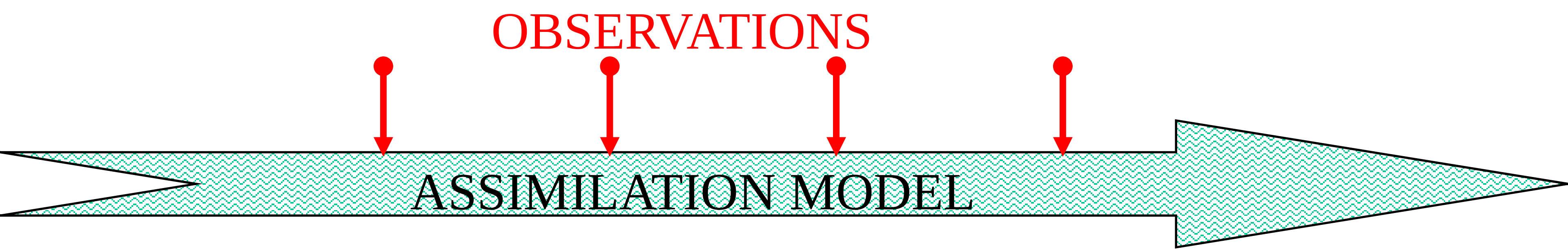

The continuous data assimilation scheme

***Data Assimilation is the process of absorbing and incorporating observed information into a prognostic model.***

OED "assimilate, v. t. … II: to absorb and incorporate."

This is normally done by integrating the model forward in time, adding observations.

- The model state summarises in an organised way the information from earlier observations.
- It is modified to incorporate new observations, by combining new & old information in a statistically optimal way.




- *At any time, the model state usually contains more information than the current observations.*
- *Only parameters well represented by the model can be assimilated in this way.*

# Data Assimilation is fitting models to data

... finding the model state most consistent with the observations

Good example of scientific method:

- The model, observation operators and assumed bias and error statistics embody our hypothesis
- Data assimilation provides ob-analysis misfit statistics to test this hypothesis
- This leads to improvements:
  - Observing system cal-val
  - Model improvements.

Naturally done in the process of properly assimilating new observation types.

Can also be done without actually assimilating, by comparing assimilation with independent obs.

This has become a major part of the process of developing and improving atmospheric models – even climate models not meant for initial value forecasting applications.

# 3. Continuing Challenges:

## *Cup is Half empty*

- Even making linear Gaussian assumptions, NWP systems are so large that error distributions are unknowable, cannot be represented even if known, and cannot be used in practical computation.
- We can only afford linear DA algorithms – the atmosphere is nonlinear and chaotic.
- The atmosphere exhibits significant features with a wide range of scales. Because of the nonlinearities these interact giving an enormous, coupled problem.

## *Cup is Half full*

- We have an excellent physical understanding; we can proceed by a physically based approach for sub-processes.
- Nonlinearities give us climate, attractors, "balance" and recognisable features (fronts, cyclones).
- Increasing computer power allows us to make progress. The selection of appropriate approximations is what makes the work interesting!

# Information content of imagery sequences

- Humans can make reasonable forecasts based on imagery alone (satellite or radar): information scarcely used in NWP.
- Time-sequences aid the interpretation of images.
- Some important information is multi-scale; details at high-resolution are used to recognise patterns whose larger-scale movements are significant.

FORSYTHE, M.: ATMOSPHERIC MOTION VECTORS: PAST, PRESENT AND FUTURE

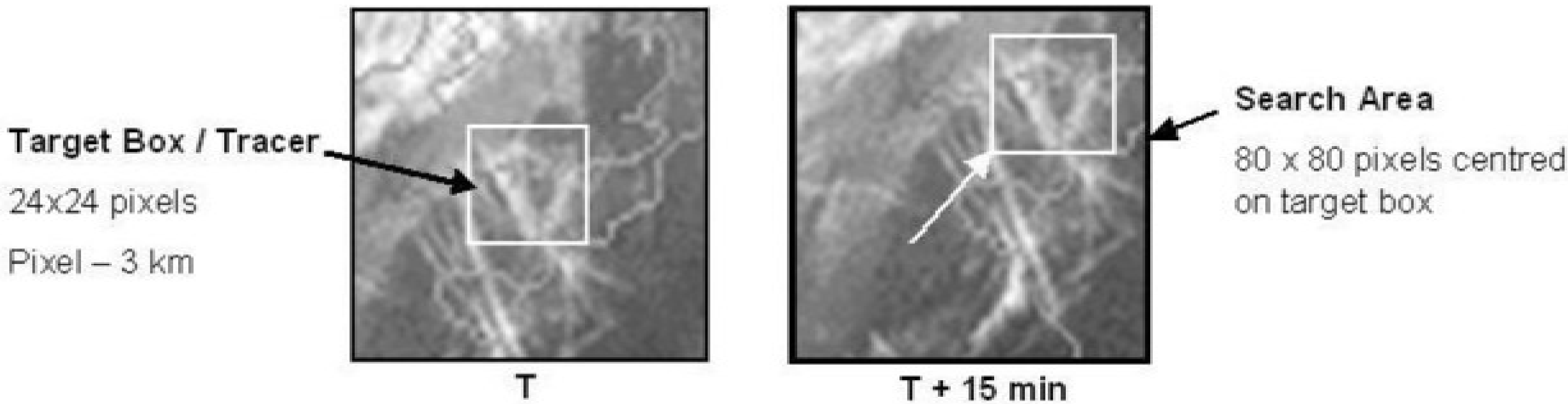
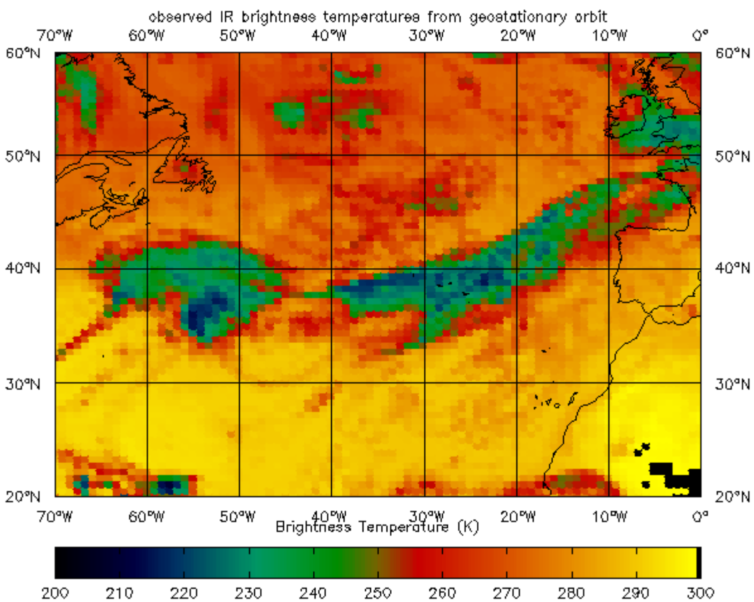


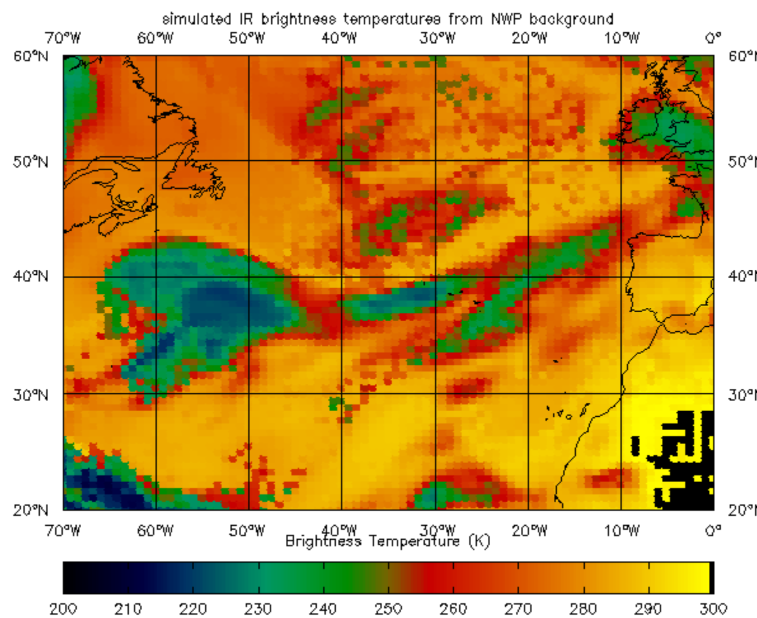

***Figure 1:*** *An illustration of the AMV tracking step for Meteosat-9 IR AMVs. The location of the target in the later image is determined by best match of the individual pixel counts of the target with all possible locations of the target in the search area using cross-correlation in the Fourier domain. The wind vector is taken as the displacement between the locations of the target boxes in the two images.*

- I am not suggesting we could replace AMVs by 4DDA in the near future!
- However they provide an example of demonstrated useful information from imagery sequences, which a method should in principle be able to extract.
- 4DDA methods could, in theory, improve on current AMV techniques in allowing for development and dynamical coupling of features.

# Comparison of observed and modelled cloud

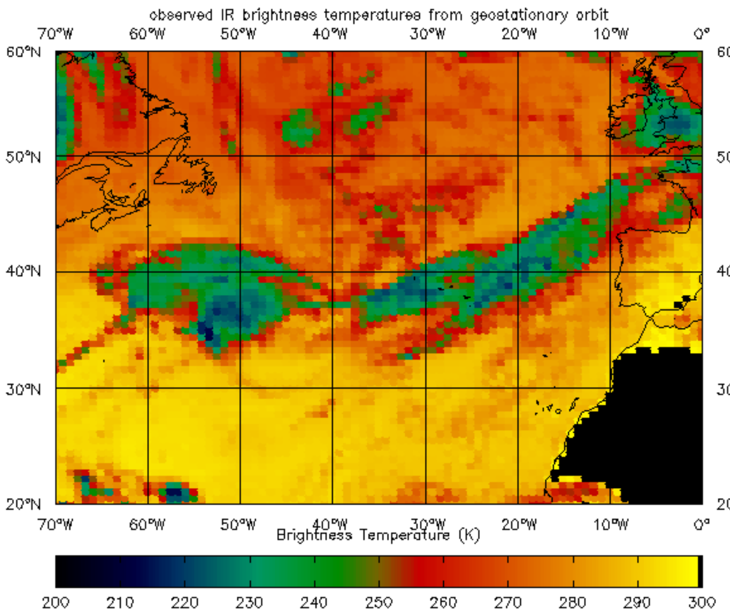
## 9Z 13-10-2002

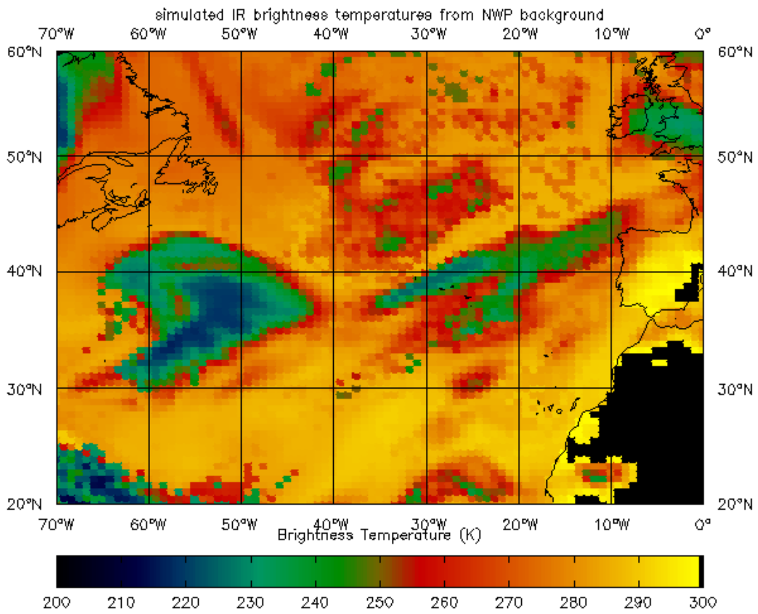


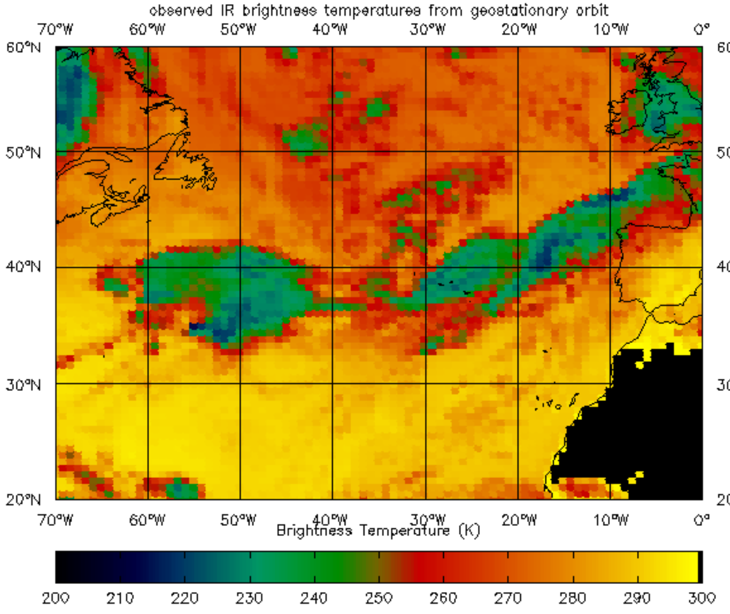

Samatha Pullen

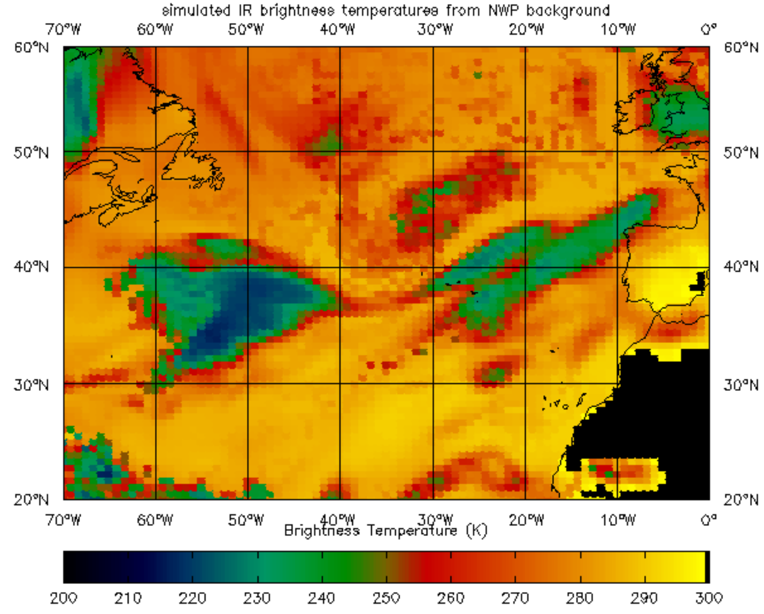
# 12Z 13-10-2002



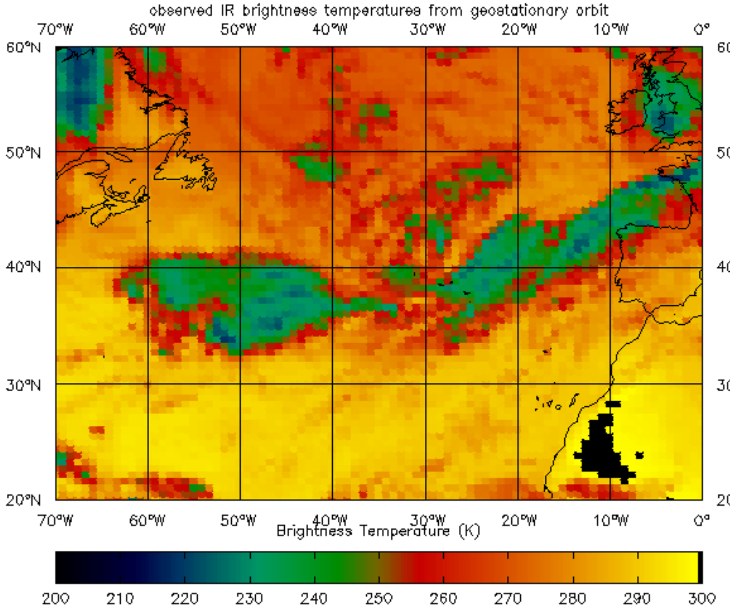



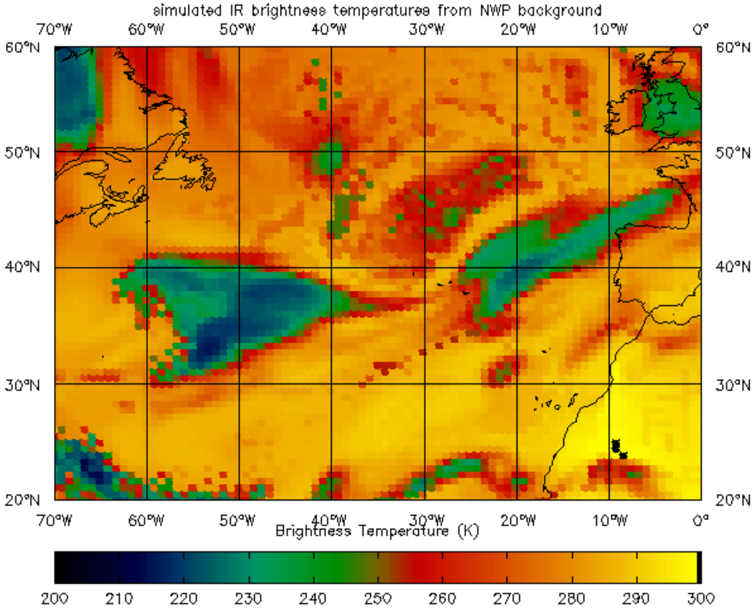
# 152 13-10-2002



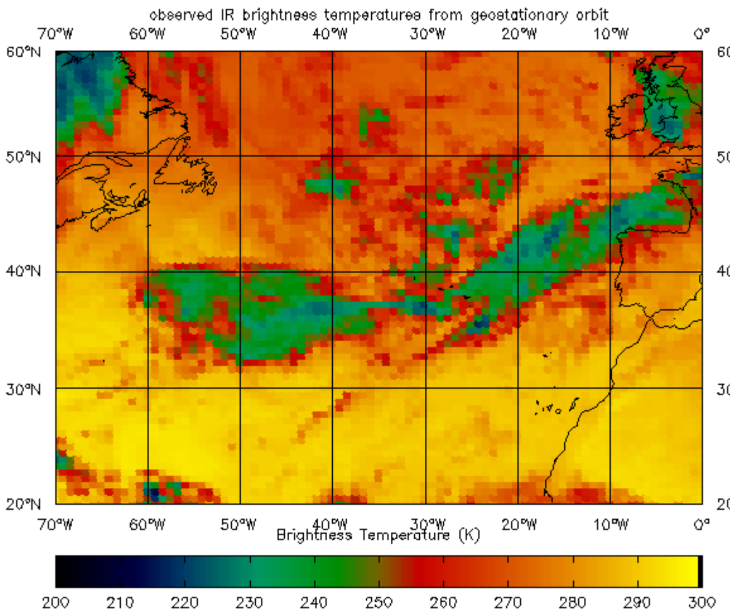



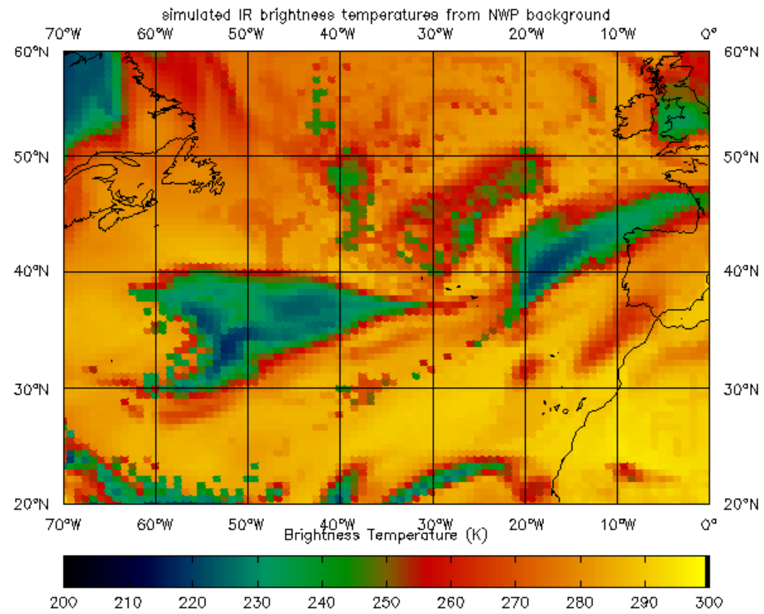
# 182 13-10-2002



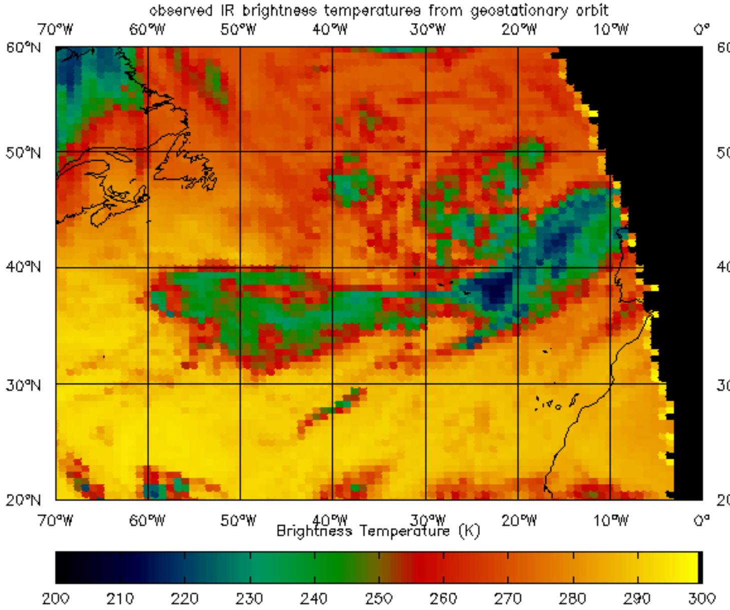



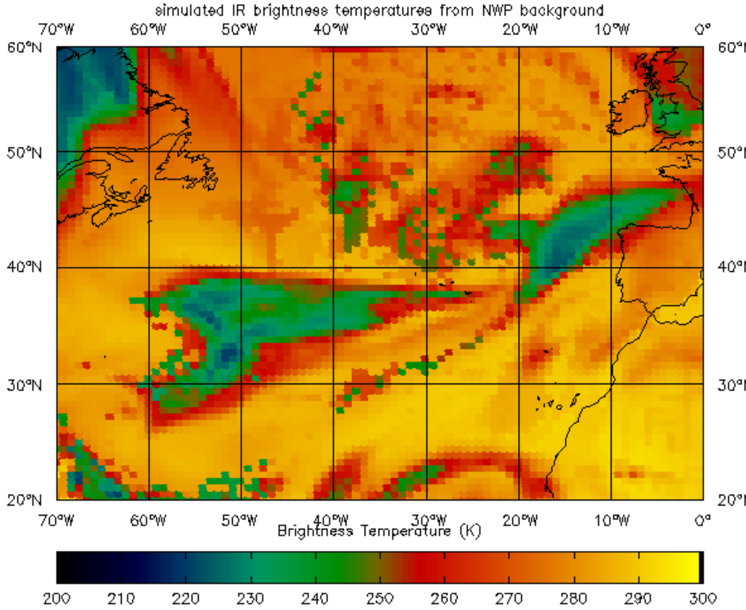
# 212 13-10-2002



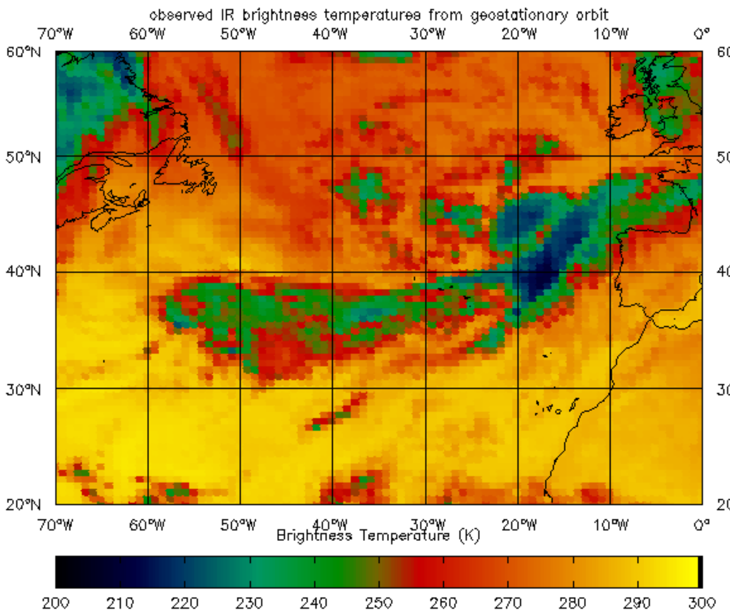



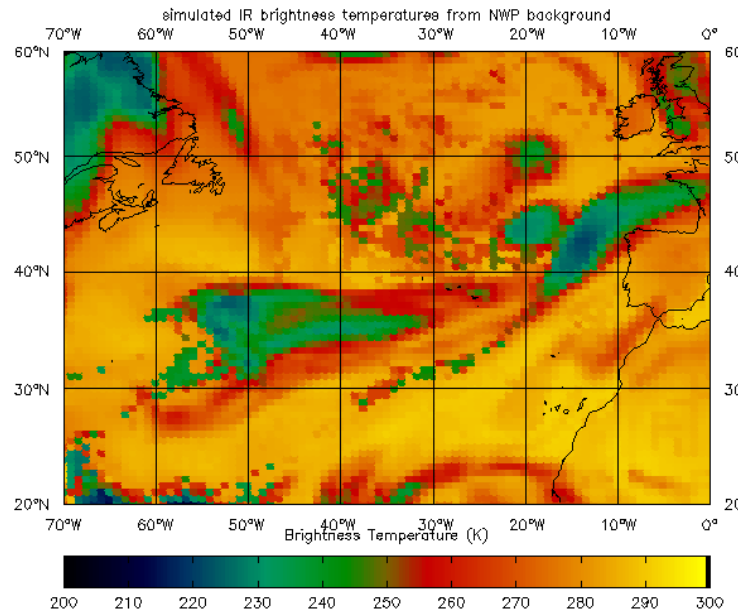
# 0Z 14-10-2002


observed IR brightness temperatures from geostationary orbit
70°W 60°W 50°W 40°W 30°W 20°W 10°W 0°
60°N 50°N 40°N 30°N 20°N
Brightness Temperature (K)
200 210 220 230 240 250 260 270 280 290 300

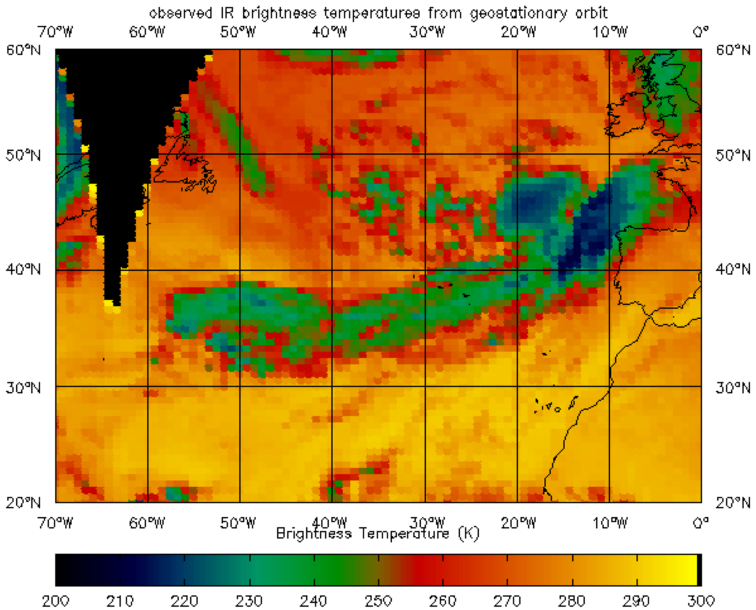

simulated IR brightness temperatures from NWP background
70°W 60°W 50°W 40°W 30°W 20°W 10°W 0°
60°N 50°N 40°N 30°N 20°N
Brightness Temperature (K)
200 210 220 230 240 250 260 270 280 290 300

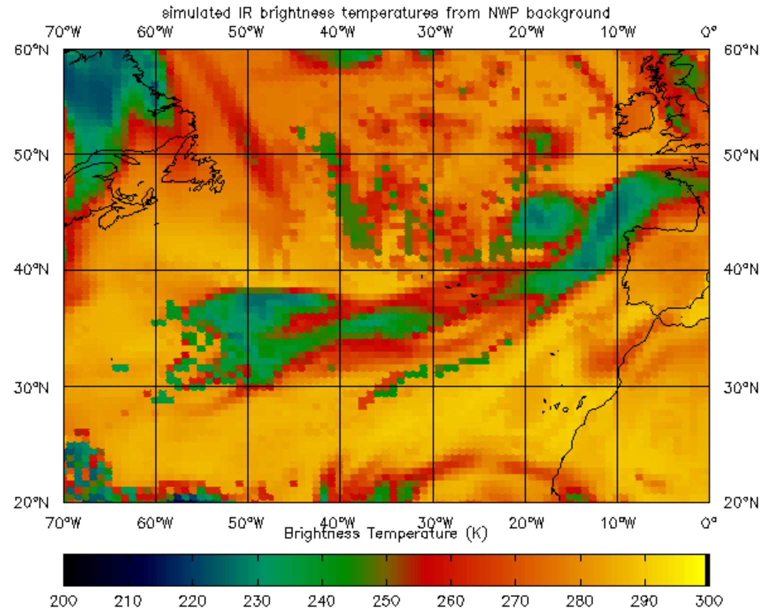
# 32 14-10-2002


observed IR brightness temperatures from geostationary orbit
70°W 60°W 50°W 40°W 30°W 20°W 10°W 0°
60°N 50°N 40°N 30°N 20°N
Brightness Temperature (K)
200 210 220 230 240 250 260 270 280 290 300


simulated IR brightness temperatures from NWP background
70°W 60°W 50°W 40°W 30°W 20°W 10°W 0°
60°N 50°N 40°N 30°N 20°N
Brightness Temperature (K)
200 210 220 230 240 250 260 270 280 290 300

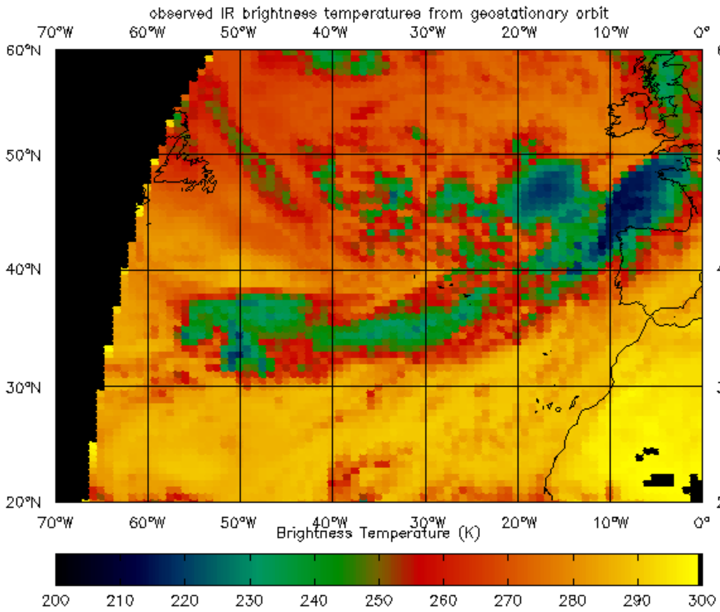
# 62 14-10-2002

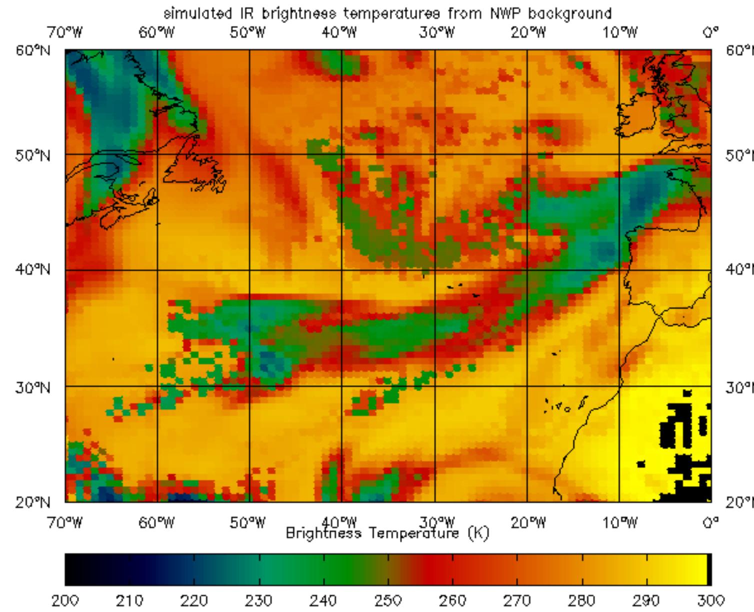

observed IR brightness temperatures from geostationary orbit
70°W
60°W
50°W
40°W
30°W
20°W
10°W
0°
60°N
50°N
40°N
30°N
20°N
Brightness Temperature (K)
200
210
220
230
240
250
260
270
280
290
300


simulated IR brightness temperatures from NWP background
70°W
60°W
50°W
40°W
30°W
20°W
10°W
0°
60°N
50°N
40°N
30°N
20°N
Brightness Temperature (K)
200
210
220
230
240
250
260
270
280
290
300

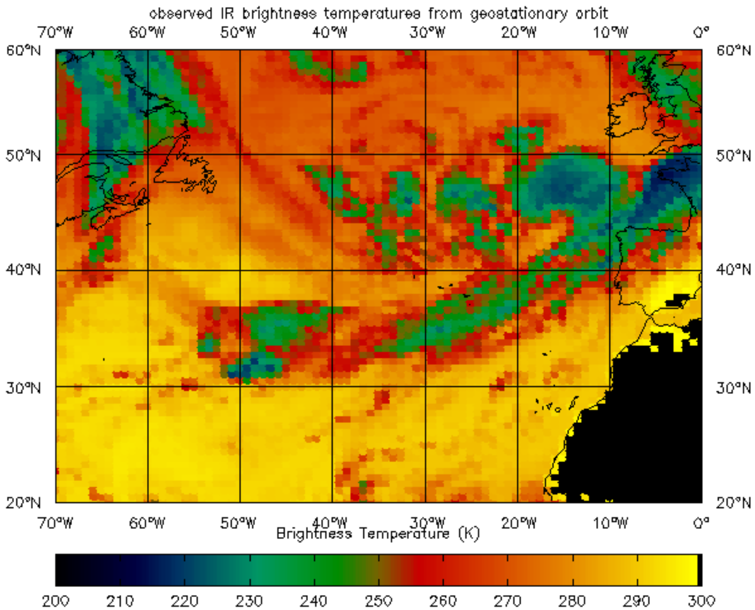
Met Office

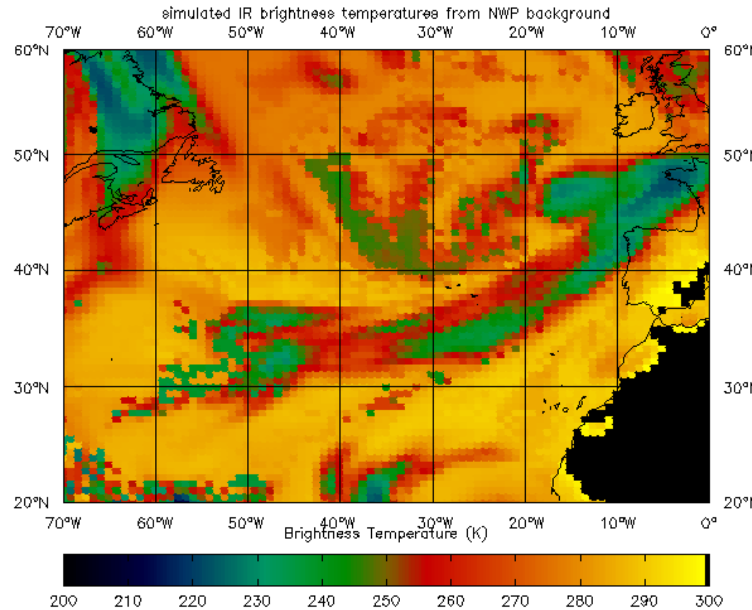
# 92 14-10-2002



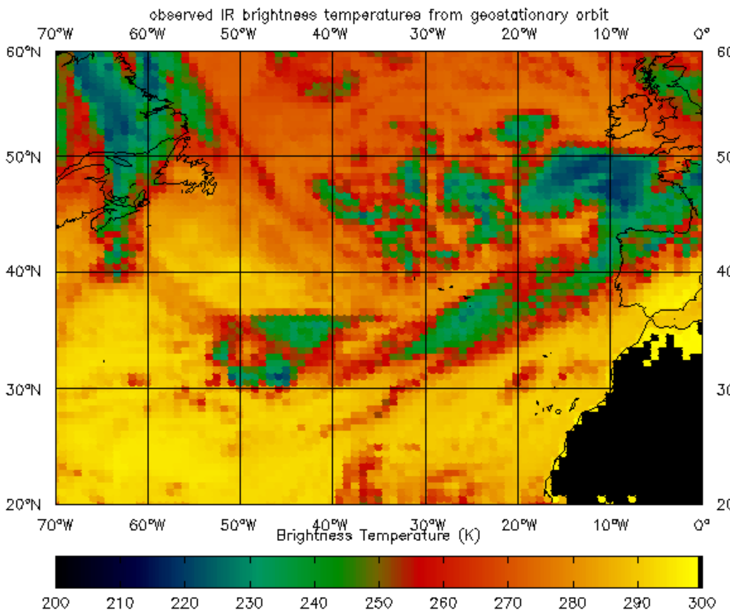



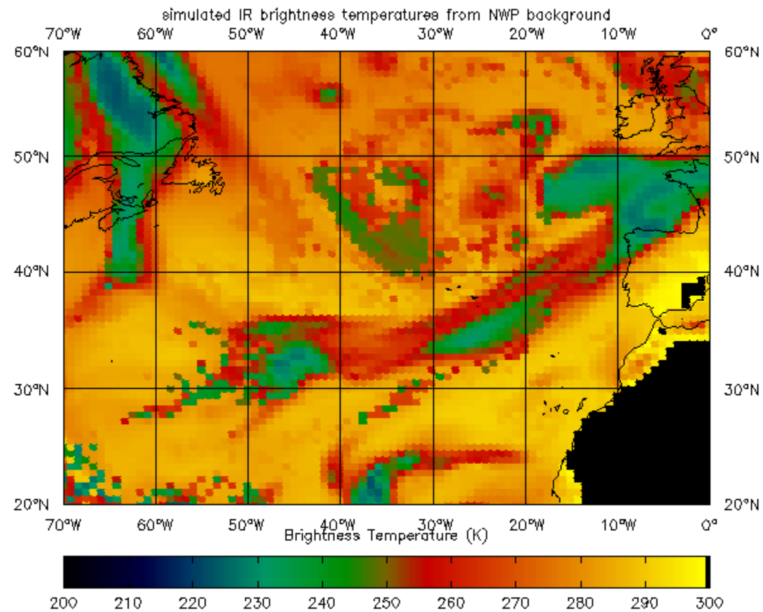
# 12Z 14-10-2002



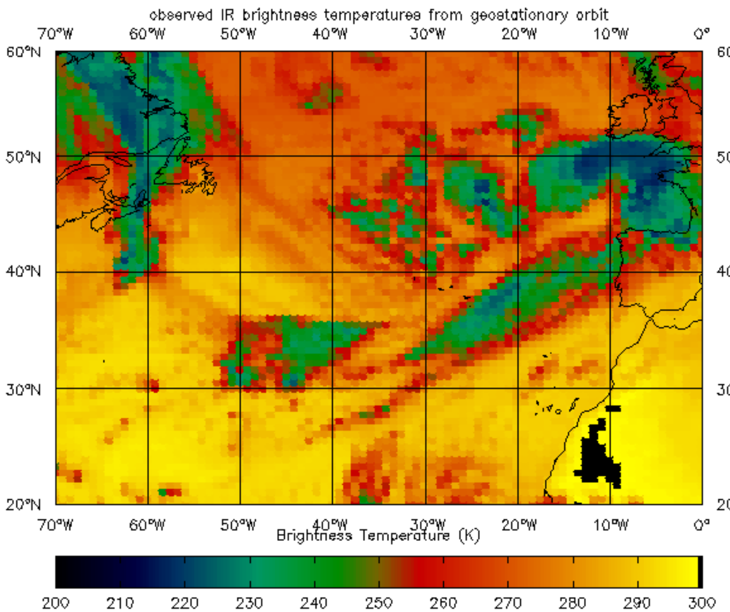



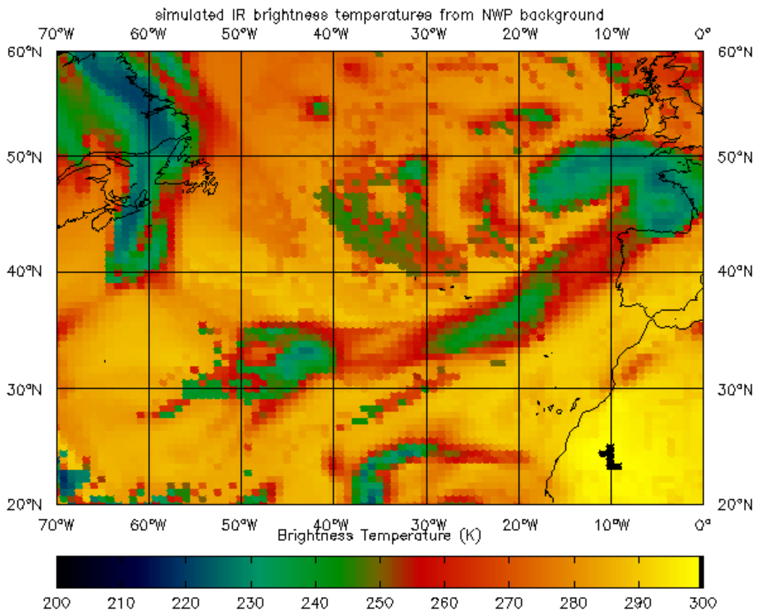
# 152 14-10-2002



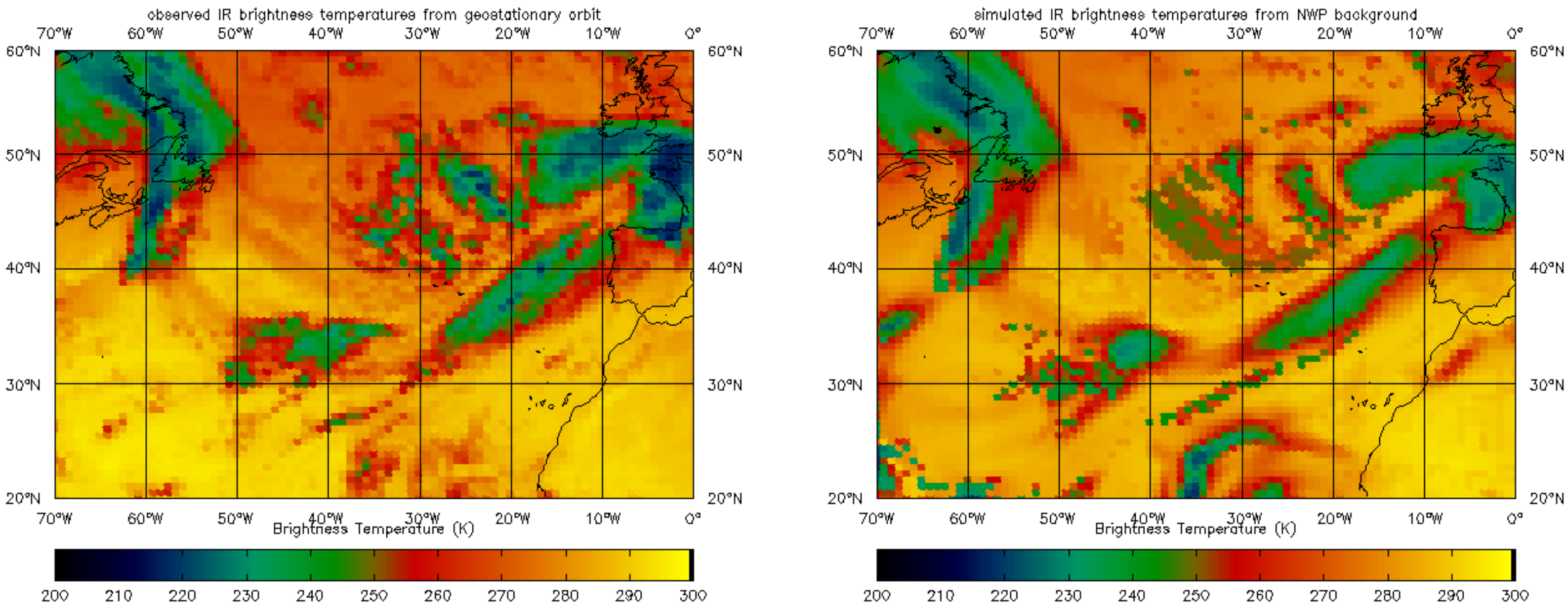



# 182 14-10-2002



# 21Z 14-10-2002



# Nonlinear 4D-Var

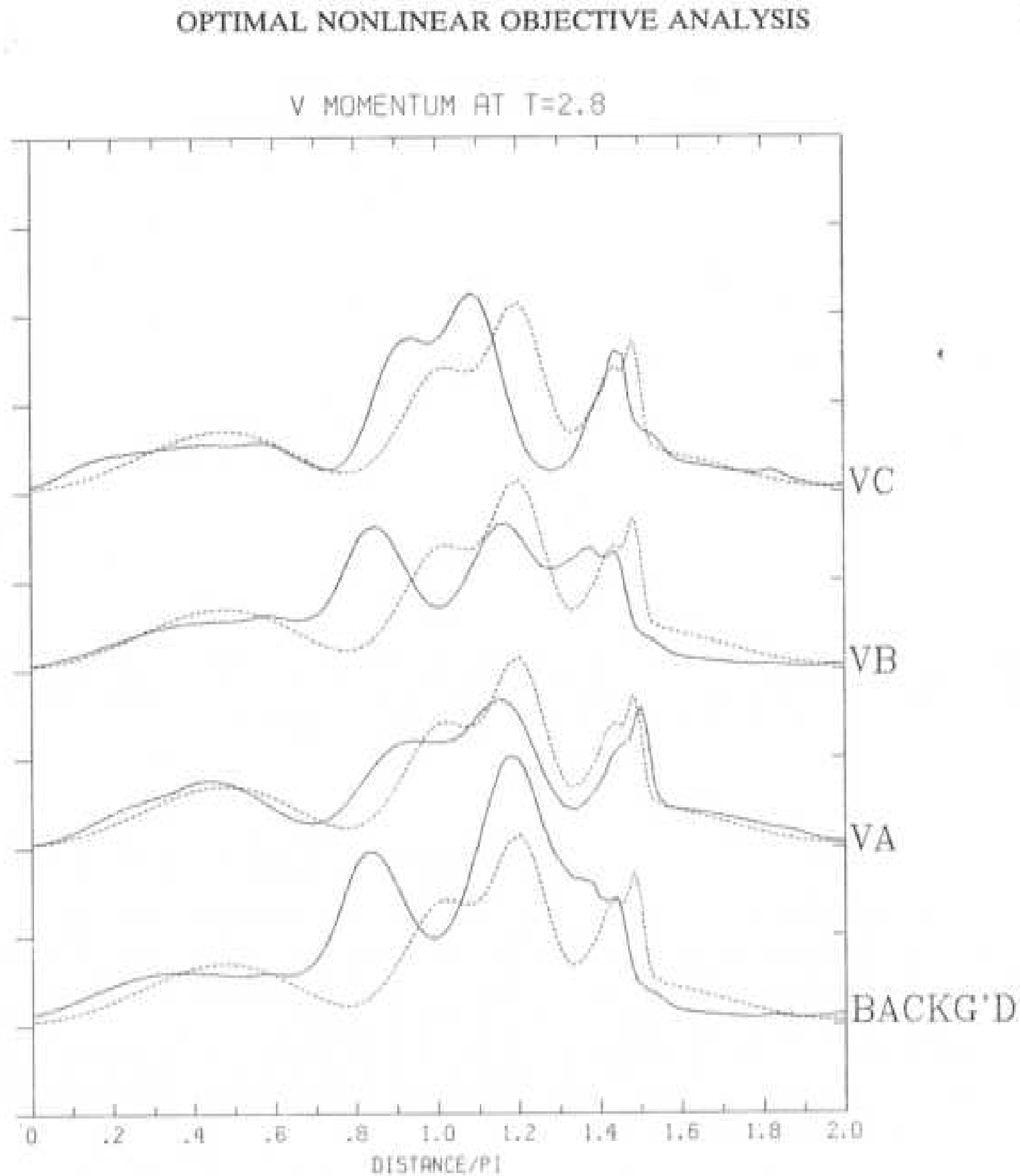
Lorenc 1988 showed that nonlinear 4D-Var of tracer obs at two times in a shallow water model improved forecast.

Cycled 3D-Var of tracer at two times

3D-Var of tracer at one time

4D-Var of tracer at two times

Forecast from background



# 4D-Var “retrieved” winds

## T42L19, 24hr, adiabatic, not incremental, no $J_b$

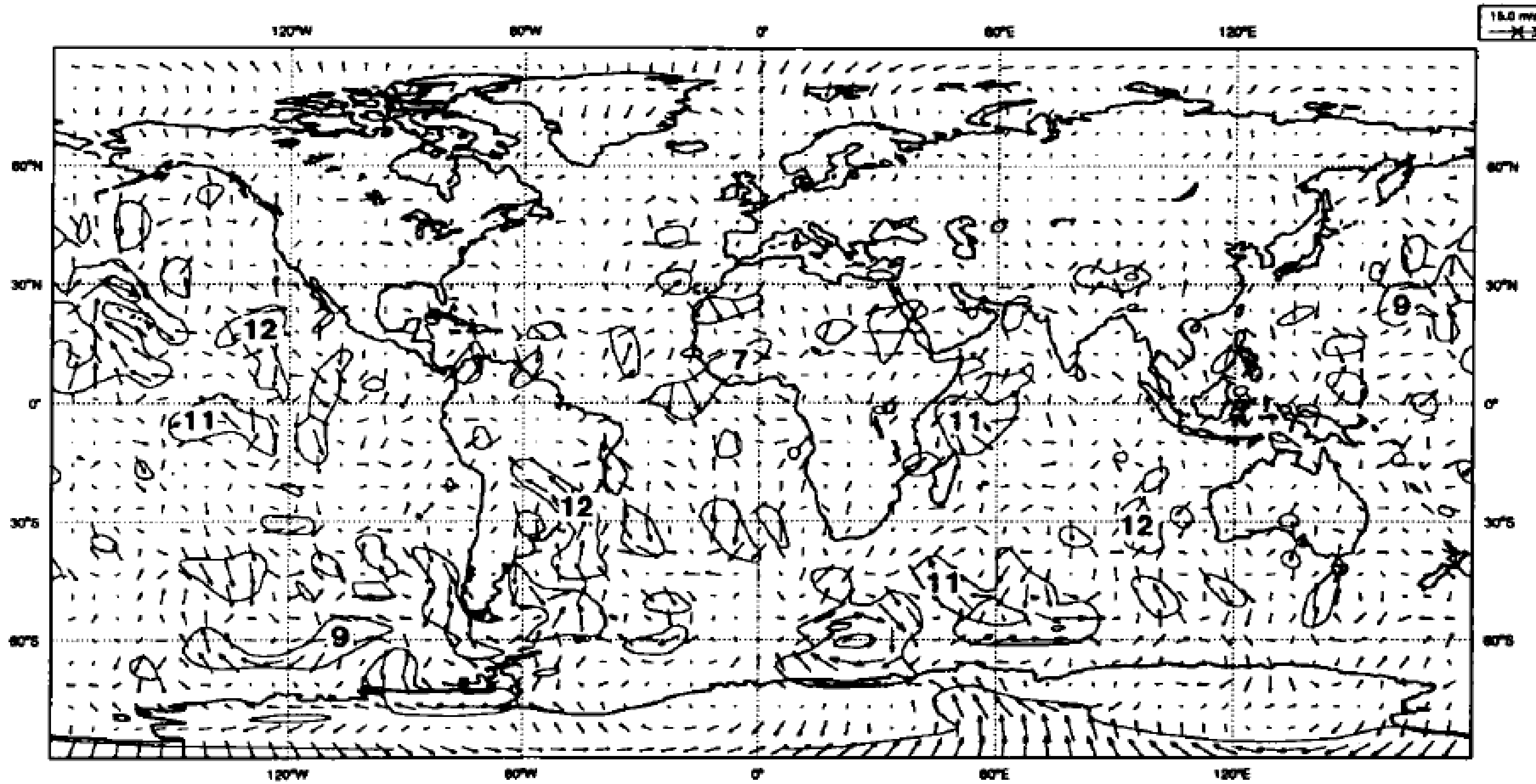



Figure 16. As Fig. 15, but for the wind field. Wind speed is contoured with an interval of 5 m $s^{-1}$.

# 4D-Var “retrieved” winds

 QJ 1994

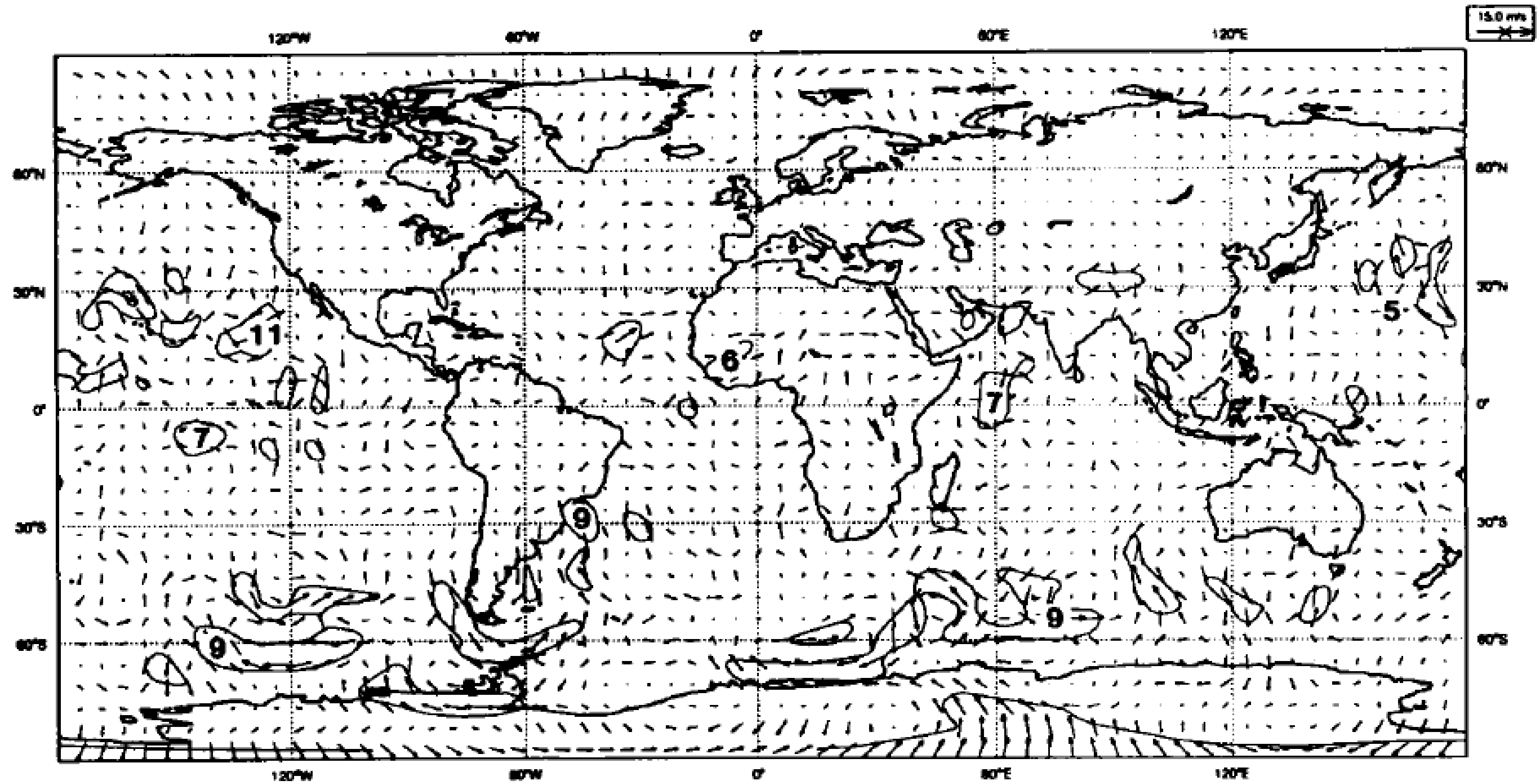



**Figure 17.** As Fig. 16 but humidity TOVS channels HIRS-11 and HIRS-12 have been excluded from the assimilation using TOVS radiances.

# Linearized Extended Kalman Filter

Daley (1995, 1996) studied linearized equations in EKF.

*Wind field can be recovered provided:*

- *sufficient structure in the constituent field,*
- *observations are frequent and accurate,*
- *data voids are small.*

i.e. filter estimated field must stay close enough to the truth for gradients to be accurate.

# Equations for tracer advection

$$\frac{Dm}{Dt} = S$$

$$\frac{\partial m}{\partial t} + \nabla\left(\mathbf{u}\ m\right) = S$$

Determining $\mathbf{u}$ & $m$ simultaneously is a nonlinear problem.

$$\frac{\partial m'}{\partial t} + \mathbf{u} \bullet \nabla m' + \mathbf{u}' \bullet \nabla m = S'$$

In the linearised equations,

changes to the wind depend on the gradient of the linearisation state $m$,

biases in observations or model $S'$ can change the wind.

# Will linear incremental 4D-Var work? *Not very well!*

- Wind increments are calculated using gradients of the guess.
- In a long window (several ob times):
  - Cannot alter both the initial $m$ (to fit early obs) and the wind $\mathbf{u}$ which advects it (to fit late obs).
  - the guess is less likely to be accurate.
- In a short-window cycle (mimicking EKF):
  - $\mathbf{u}'$ will be derived from the movement of background $m$ to fit observations.
  - But 4D-Var does not know in which areas background $m$ is unreliable (due to past data voids) and may derive unreliable $\mathbf{u}'$ .

# Multi-scale DA

- If displacement (between obs) $\geq$ size of features (or if features have sharp edged, e.g. cloud/no cloud):
  - Multiple maxima in fit to obs are possible;
  - Linearisation fails if obs increments fall in regions with zero gradient; $$\frac{\partial m'}{\partial t} + \mathbf{u} \bullet \nabla m' + \mathbf{u}' \bullet \nabla m = S'$$
  - So we need a good guess at the displacement.
- Might obtain this from a preliminary iteration at reduced resolution (such that features are smoothed).
- This fits well with multiple outer-loop 4D-Var.