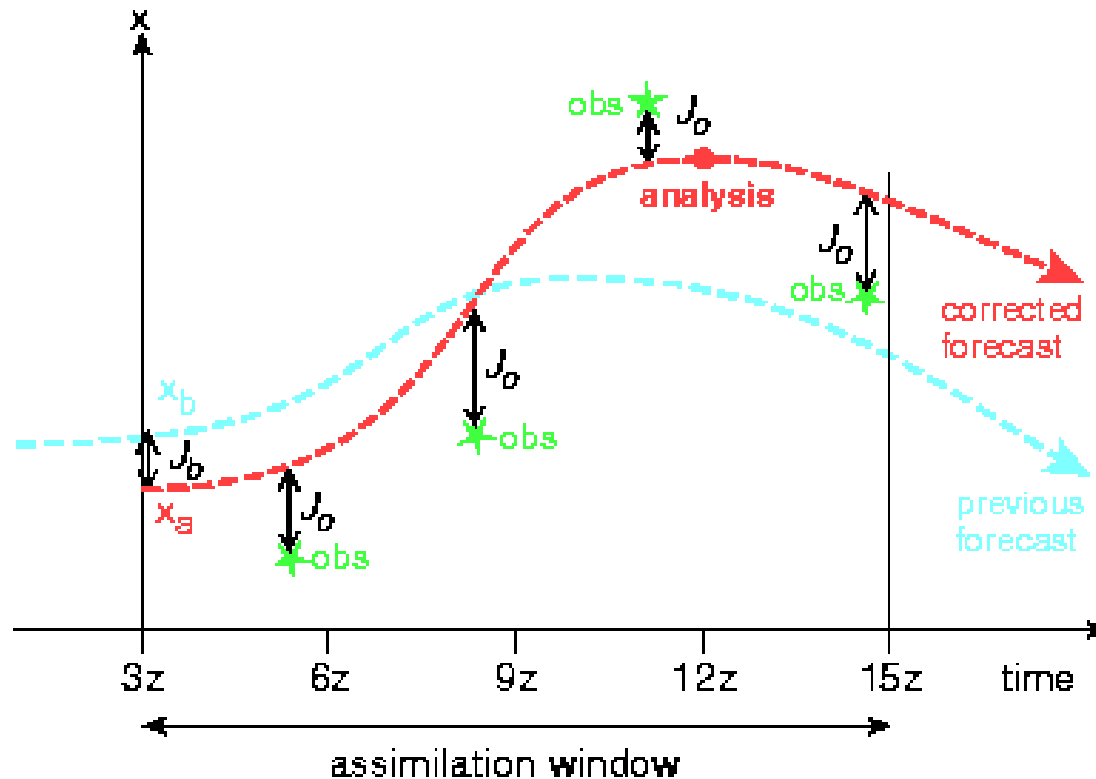


Status of four-dimensional variational data assimilation

Erik Andersson
and the ECMWF DA-Section



Contents – Status of 4D-Var (45 minutes)

As requested, but not necessarily in this order

- **Operational aspects**
- **Recent developments**
- **Work underway**
- **A clear list of assumptions and approximations - with characterisation and motivation for each**

Our duty is to provide the tools to maximize the value (benefit) of often expensive observations

- **Creating coherent analyses of the state of the atmosphere, ocean and land surface**
- **Improving the accuracy of weather forecasting**
- **Monitoring atmospheric constituents and pollution**
- **Creating climate re-analyses - documenting climate change**
- **Providing estimates of uncertainty (analysis error) – and initialize EPS**

It is not likely that one single methodology will satisfy all the demands of all these application areas

Variational, ensemble and other approximations of the Kalman Filter will all remain in our communal toolbox for many years to come

All(?) of the main NWP centres now run variational data assimilation schemes operationally

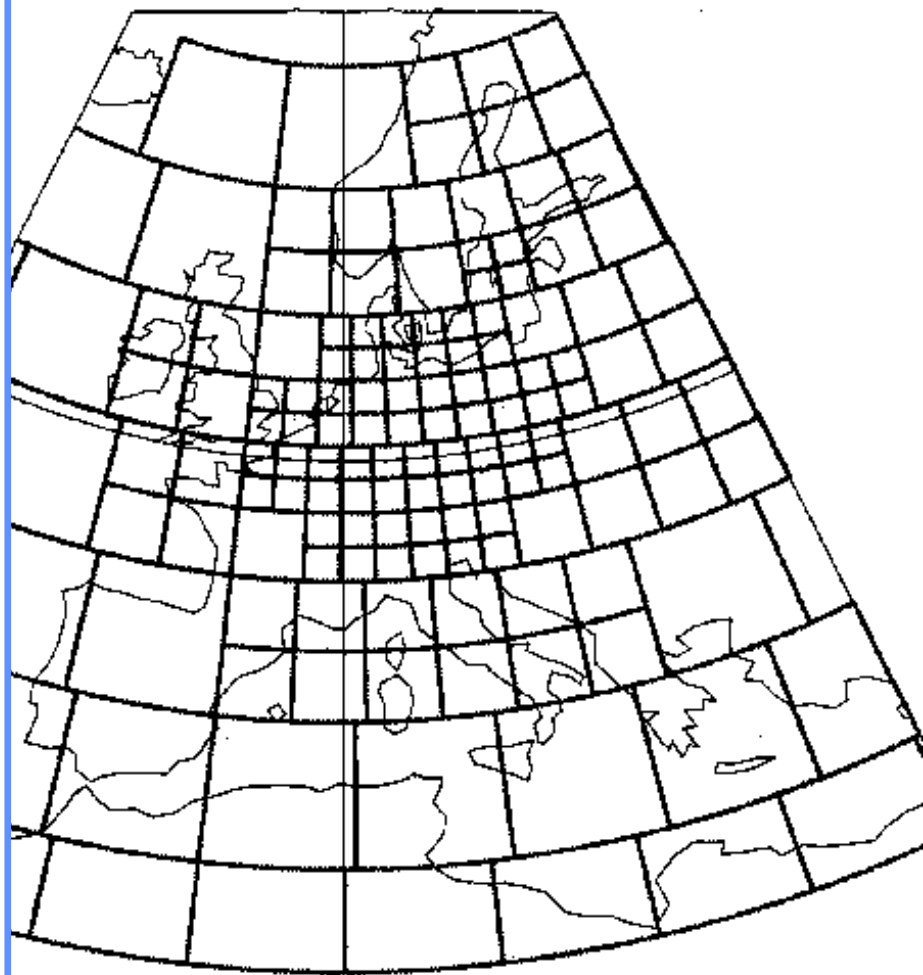
- **ECMWF**
- **France**
- **Canada**
- **Japan**
- **United Kingdom**
- **USA (NCEP, NRL, WRF)**
- **Germany**
- **HIRLAM countries**
- **ALADIN countries**
- **Taiwan**
- **Korea**
- **Australia**
- **China**

At the WMO Geneva workshop on “Impact of various observing systems”, May 2008, an impressive range of positive results were presented – clearly demonstrating the ability of variational systems to benefit from the varying components of the global observing system.

1986 issues in data assimilation (ECMWF's OI)

Initial data is of critical importance for NWP success in the medium range

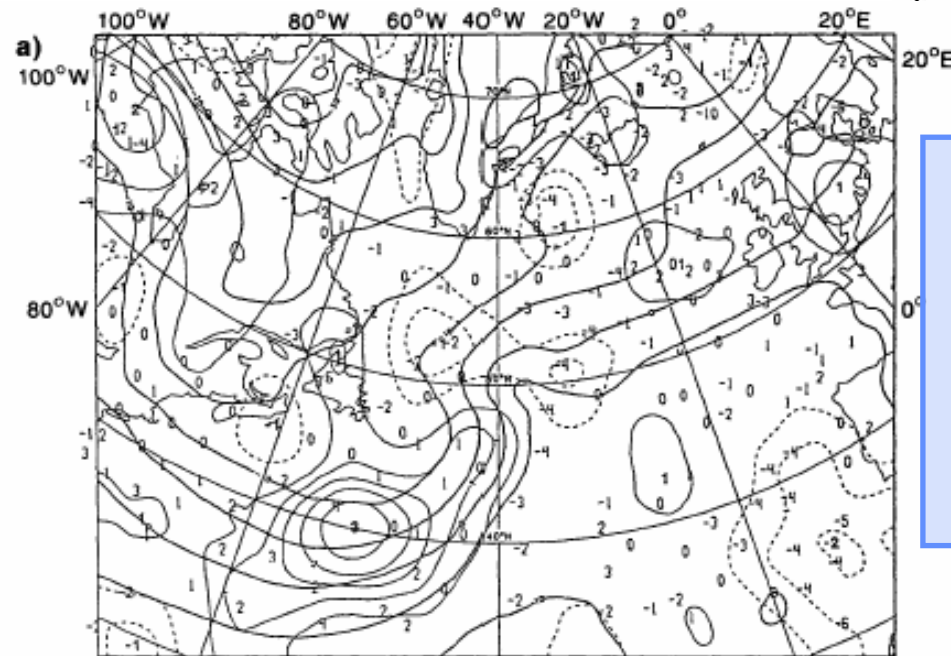
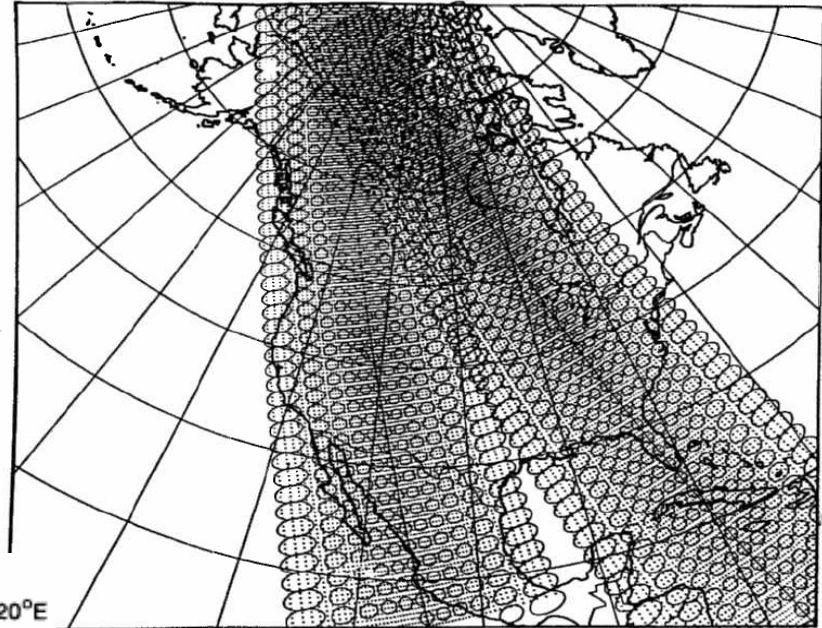
- **Need to utilize effectively all available observations - without data selection**
- **Need to analyse large-scale background errors, extending beyond the local domains (the 'OI boxes')**
- **Need to involve model dynamics: use of TL/AD models will be a major part of predictability research + 4D-Var**



The idea of direct radiance assimilation

There was frustration of seeing so little impact in OI from use of TOVS retrievals.

- Creates noisy analyses at tropopause level.
- Contains unwanted 'climate' information → correlated errors



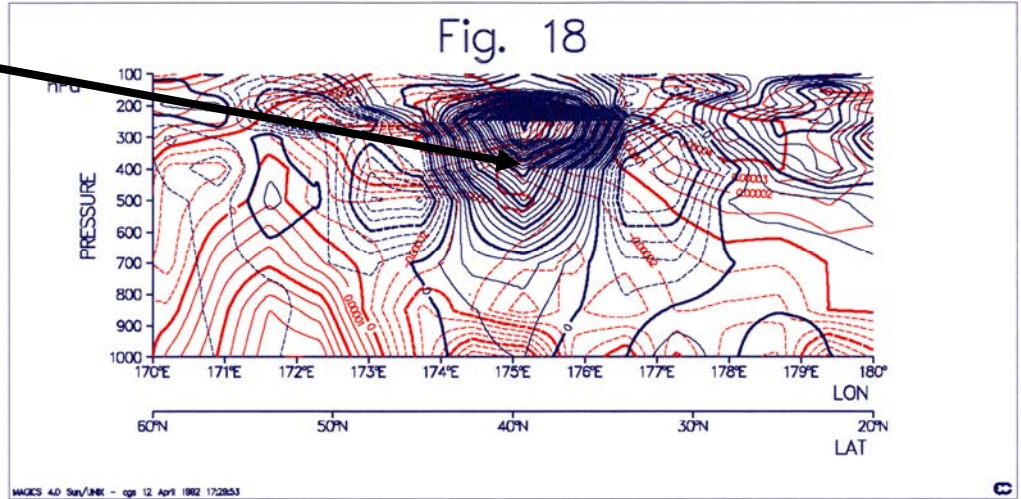
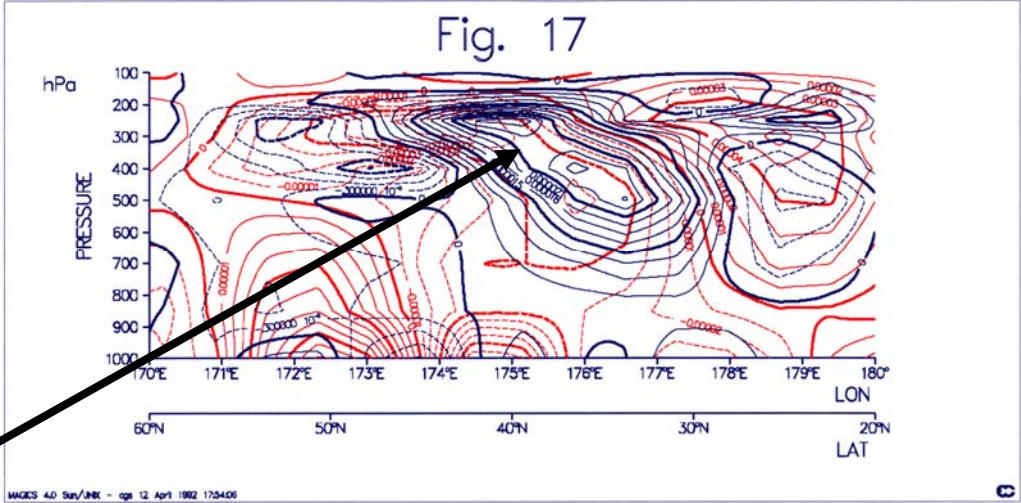
Direct use of radiances in a 3D-Var through a direct RT model and its adjoint "is a natural and attractive solution" (1991)

Sloping increments in a baroclinic region, 1993

24h 4D-Var

Impact of removing 200 hPa aircraft observations from the assimilation

OI cycled over 24h



The basic elements needed for 4D-Var

- **A forecast model and its tangent-linear and adjoint counterparts (or a perturbation model)**
- **The observation operators, code to compute J_o and its gradient**
- **The first-guess operator, code to compute J_b and its gradient**
- **Mass/wind balance operators**
- **General minimization algorithm**

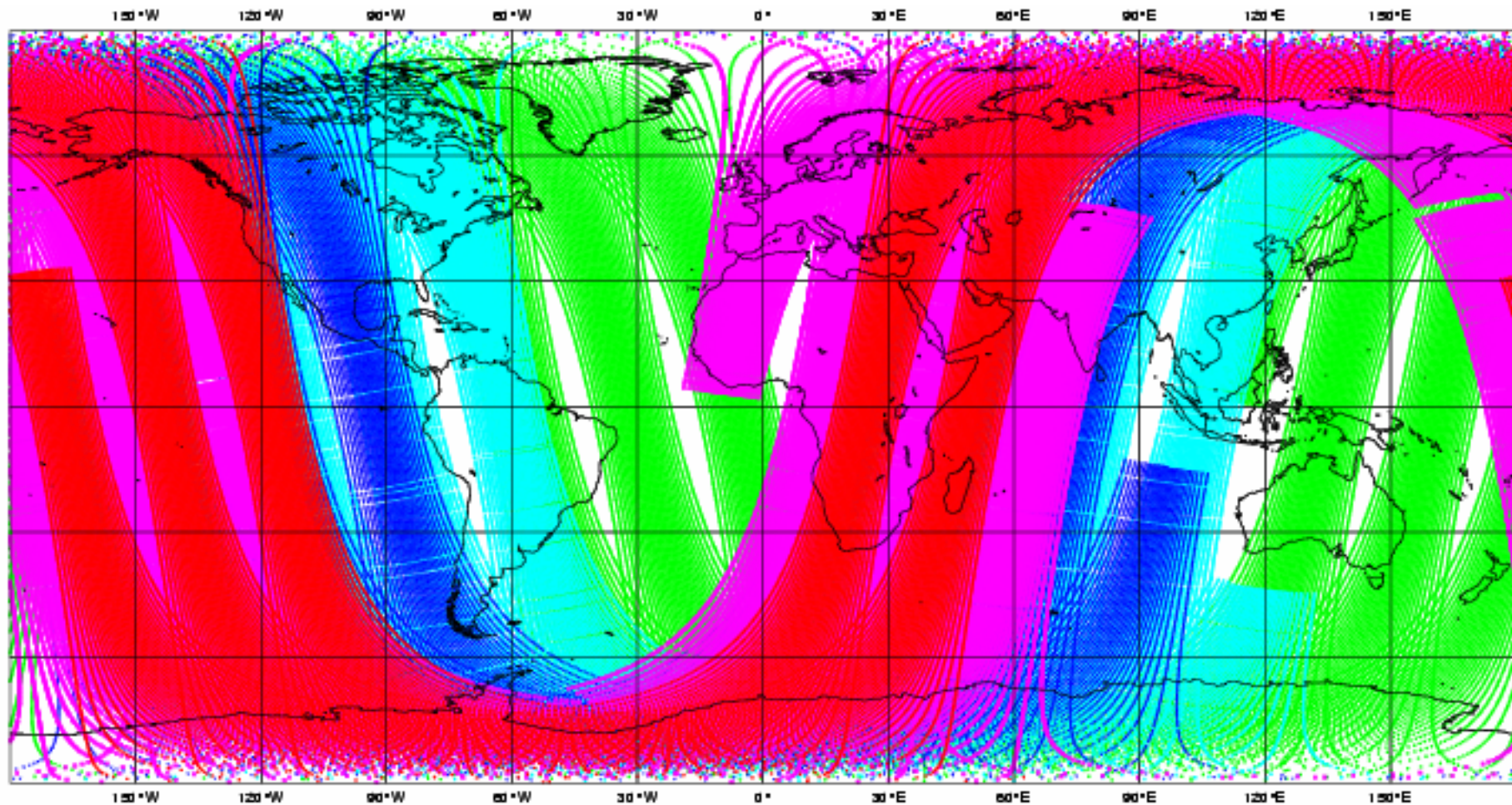
A forecast model and its TL/AD counterparts (or perturbation model)

- **Tremendous effort has gone into coding accurate TL and adjoints**
- **TL of semi-lagrangian time stepping**
- **Perturbation FC model developed by UK**
- **Focus on moist physical processes i.e. convection to facilitate assimilation of cloud and precipitation data**

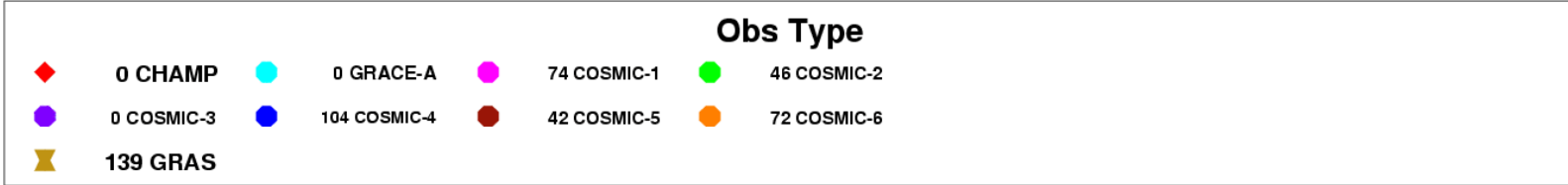
The observation operators, code to compute Jo and its gradient

- **Masses of new observations have been added over the past 10 years.**
- **Observation operators have been developed for all the main observing systems**
- **Software development is often coordinated, most notably RTTOV**
- **Observation operator routines can often be shared between NWP centres.**
- **Most recent successes include AMSU-A, AIRS, IASI, SSMIS, GPS-RO, ...**
- **Observation bias correction schemes have been developed**

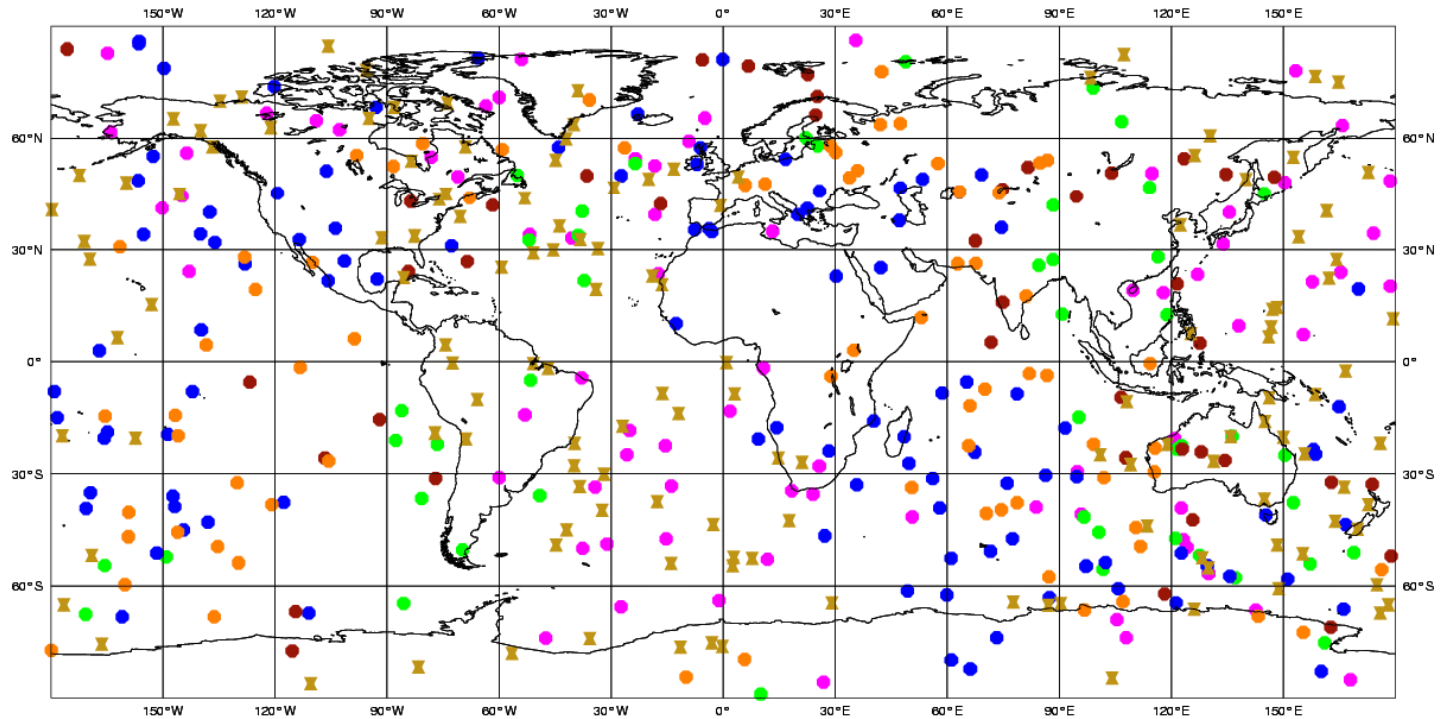
Current data coverage for AMSU-A instruments



GPS radio-occultation. Current 6-hour data coverage.



ECMWF Data Coverage (All obs DA) - GPSRO 15/MAY/2008; 00 UTC Total number of obs = 477



ECMWF

The first-guess operator, code to compute Jb and its gradient

- **Spectral-Jb**
- **Grid-point Jb**
- **Wavelet-Jb**
- **Humidity, Ozone and Jb for tracer constituents such as CO₂ and aerosols**
- **Codes tend not to be exchanged**
- **The current trend is towards formulations that can accommodate some regional variation in the specified covariance statistics**
- **Incorporation of flow-dependent variances ('errors of the day')**

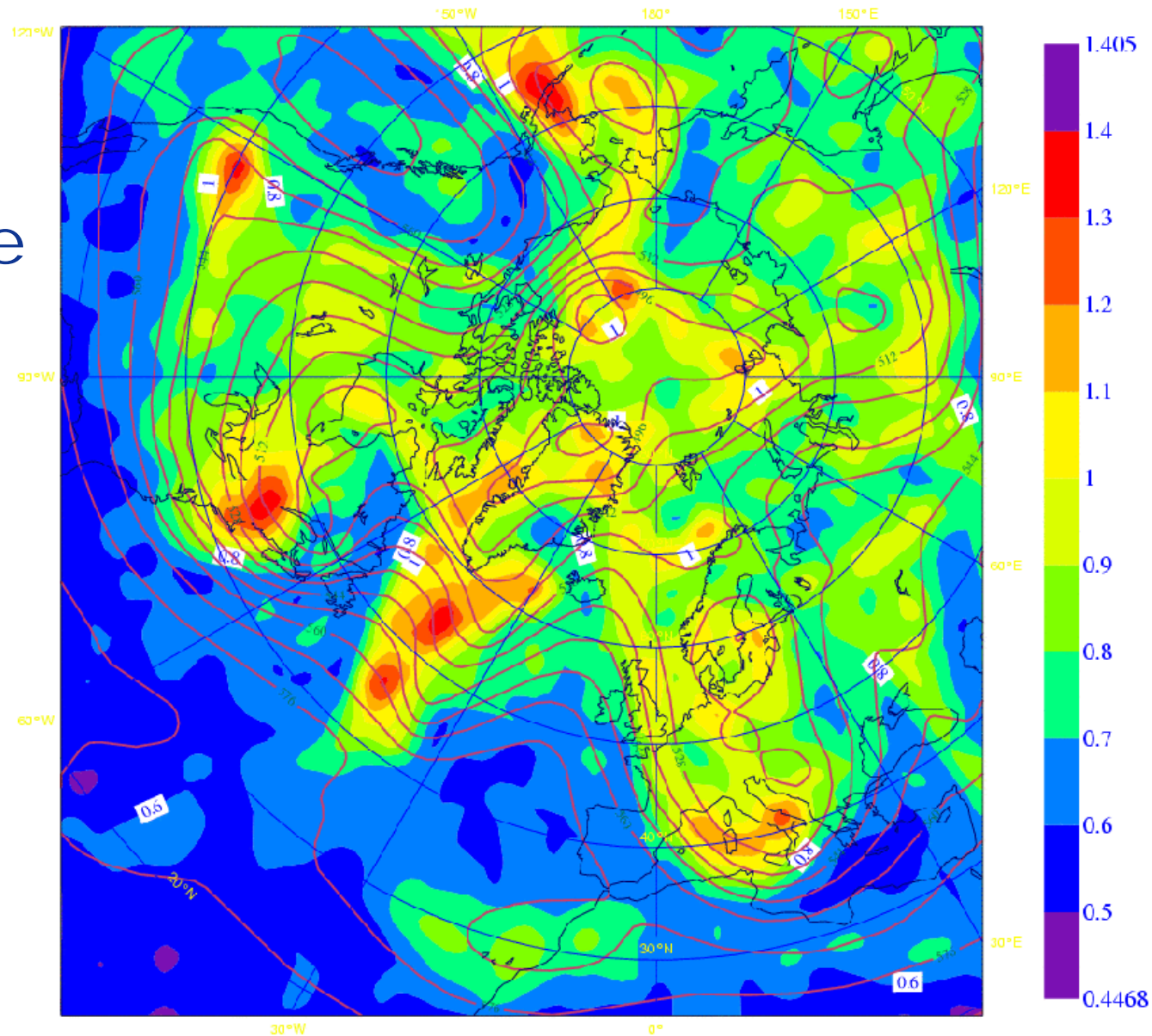
Mass/wind (moisture) balance operators

- **Linear balance**
- **(linearized) non-linear balance**
- **Omega equation**
- **Diabatic omega equation**
- **Boundary layer balances**
- **Humidity/temperature physical relationships implemented as 'balances'**

Temperature HBHT (K)

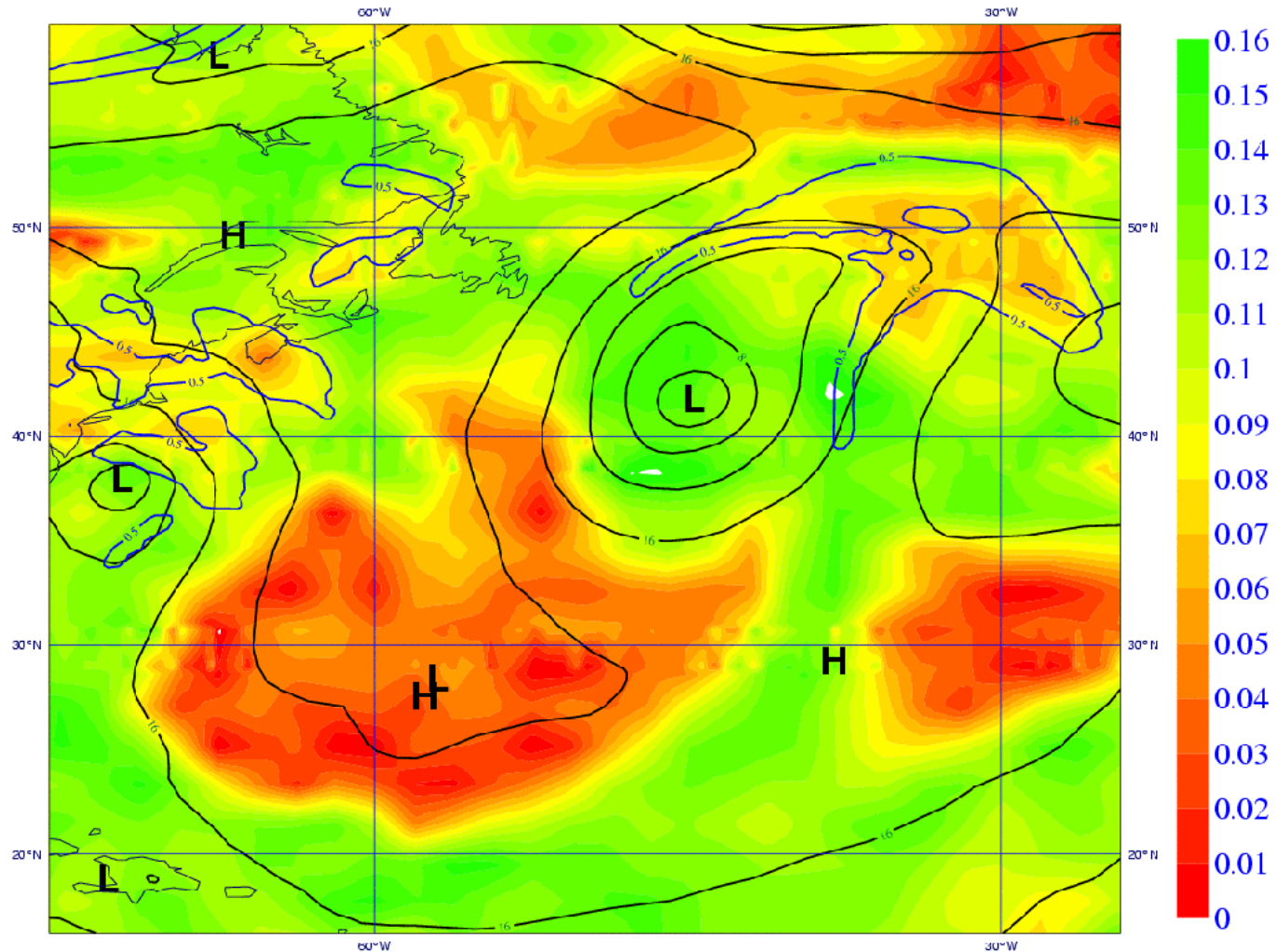
T lev39 HBHT
(shaded),

Z 500 hPa
(contoured)



Relative Humidity HBH^T

H is $H(T, q, p)$



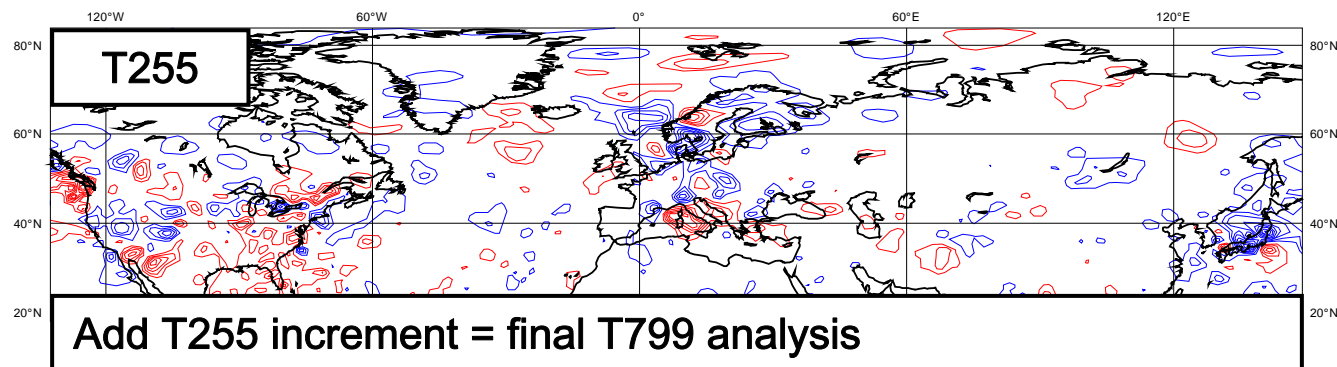
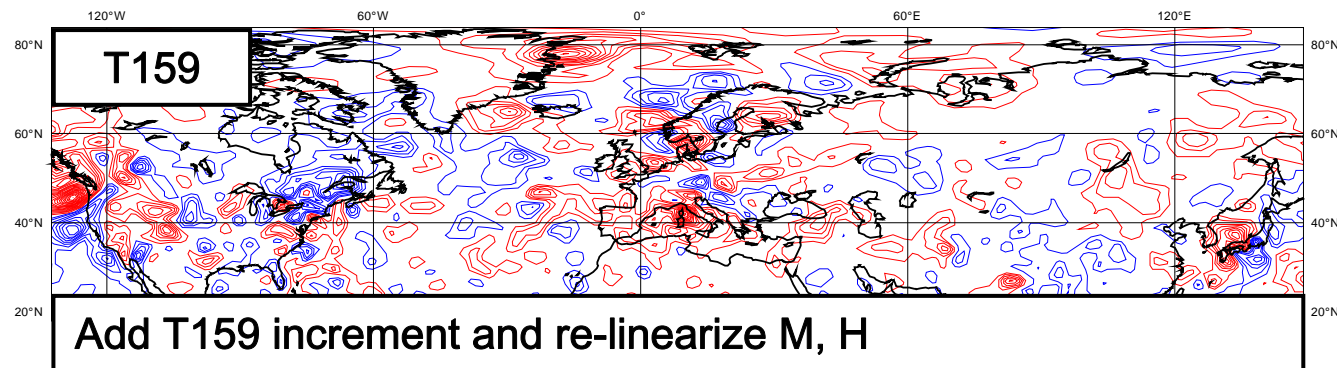
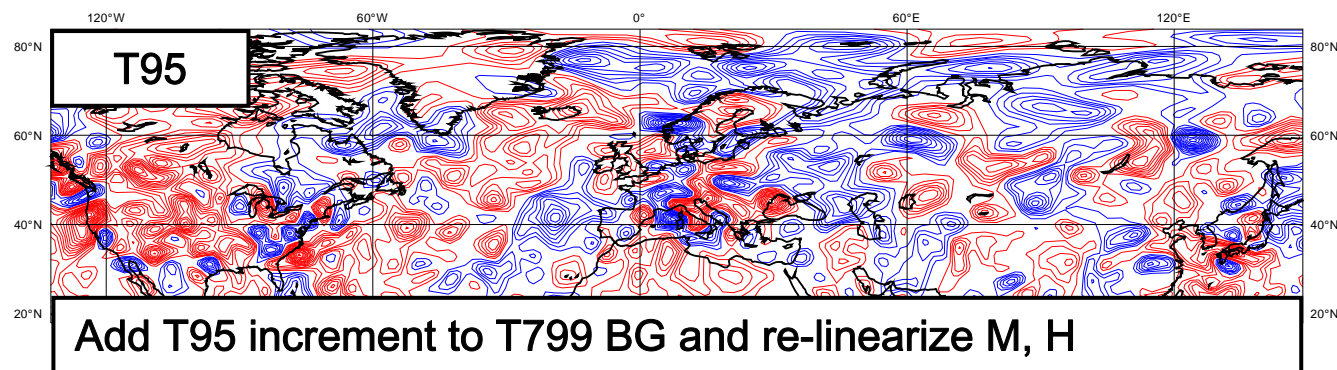
The humidity analysis formulation of E.Holm

General minimization algorithm

- **M1QN3 (non-quadratic)**
- **Conjugate gradient (quadratic)**
- **Incremental 4D-Var with linear inner-loop iterations, nested within non-linear outer-loop updates**
- **A triple loop structure adopted by UK**
- **Scalable to > 1000 processors**
- **The trend towards higher-resolution analysis requires efficient solution algorithms for parallel computing environments**

4D-Var is efficient, accurate and allows non-linearity

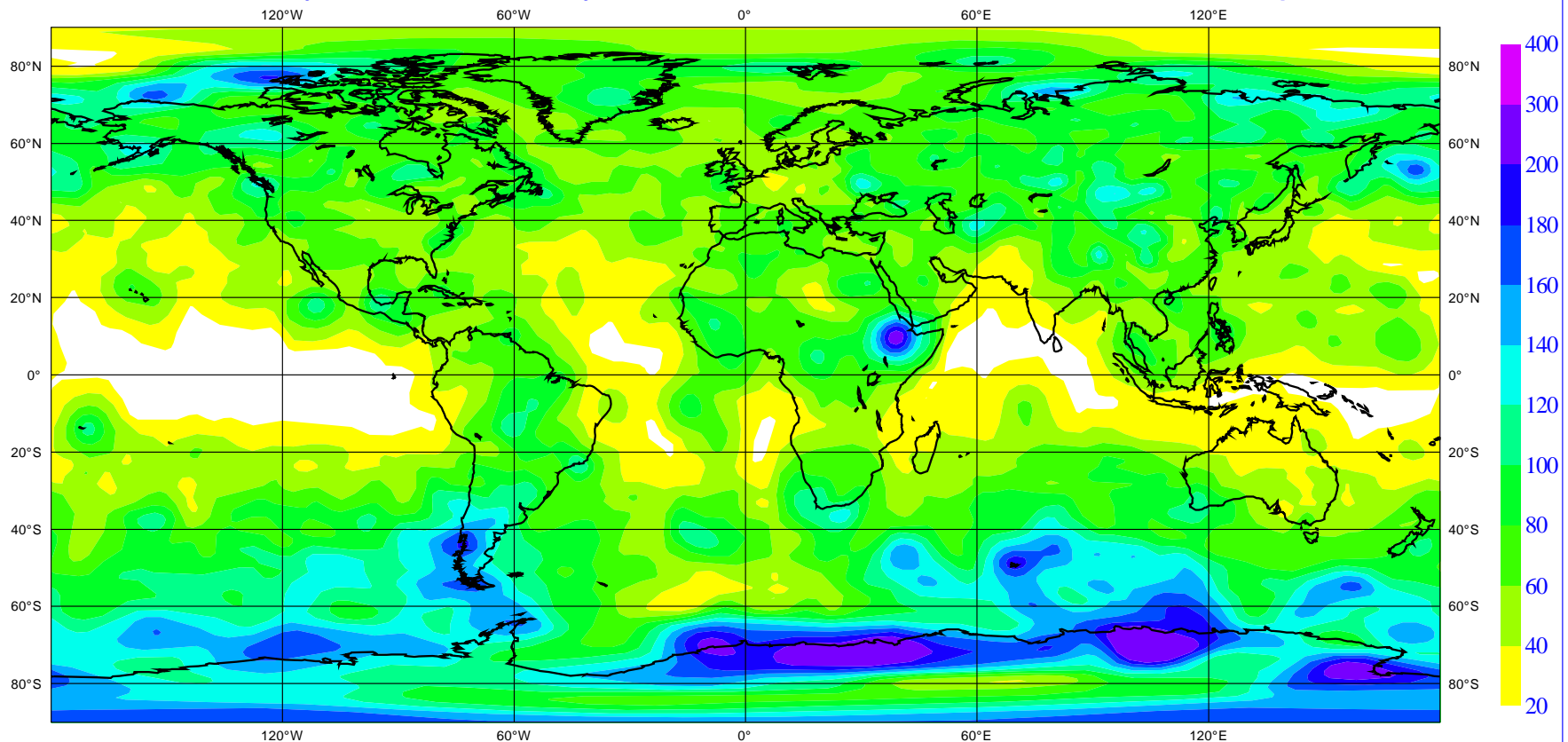
- ML=80 (900 hPa) temp. analysis increments for each of the three minimizations.
- Decreasing amplitudes $T95 > T159 > T255$.
- Small corrections added at T255 where data density is highest.
- Model and obs. operators are re-linearized twice.



Analysis increments 1994 (OI)

OI prior to the 1D-Var retune. Much 'work' done by isolated coastal radiosonde stations

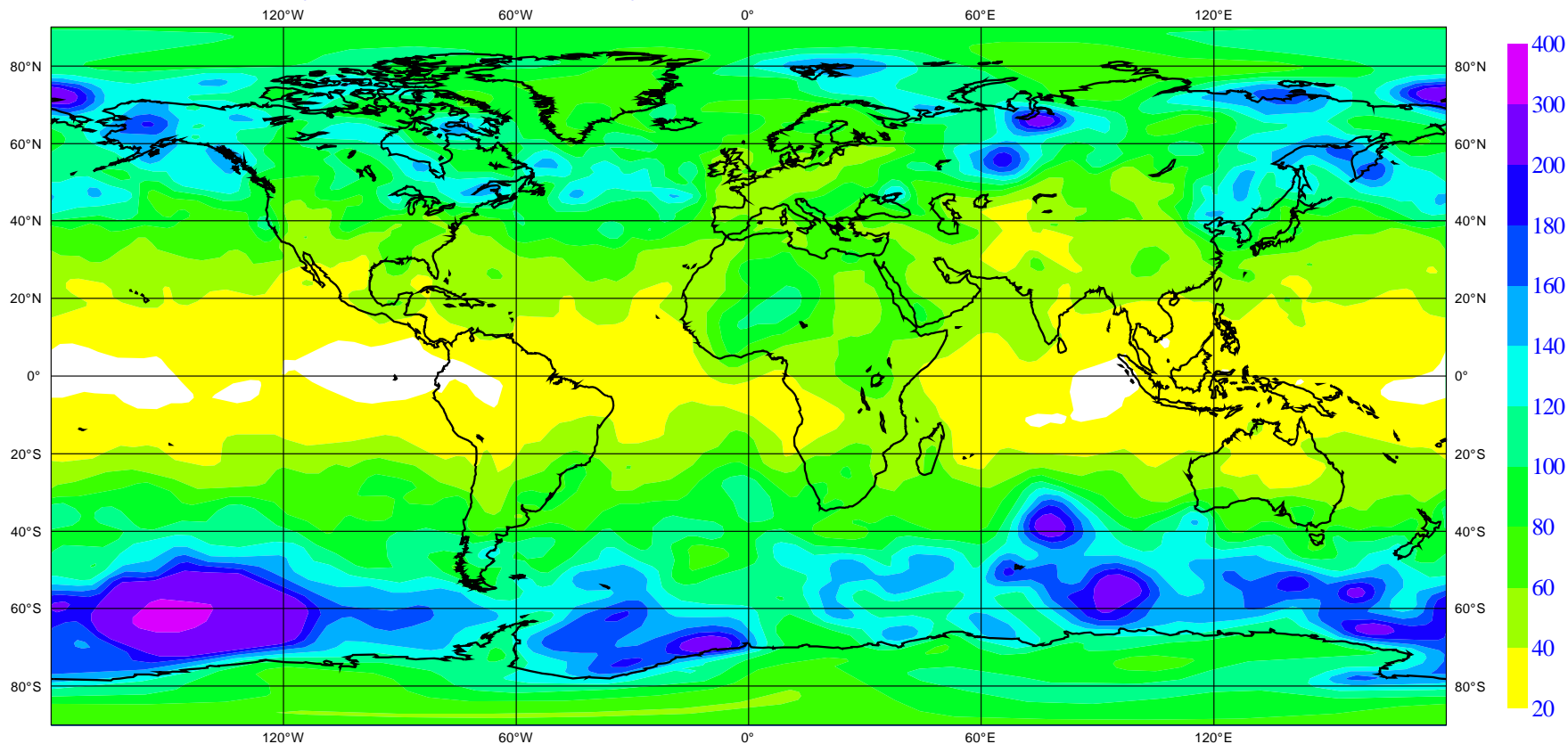
ECMWF Analysis VT:Saturday 1 October 1994 12UTC 500hPa **Geopotential



Analysis increments 1997 (3D-Var)

3D-Var, filled in the oceans. Some obvious data problems can be seen

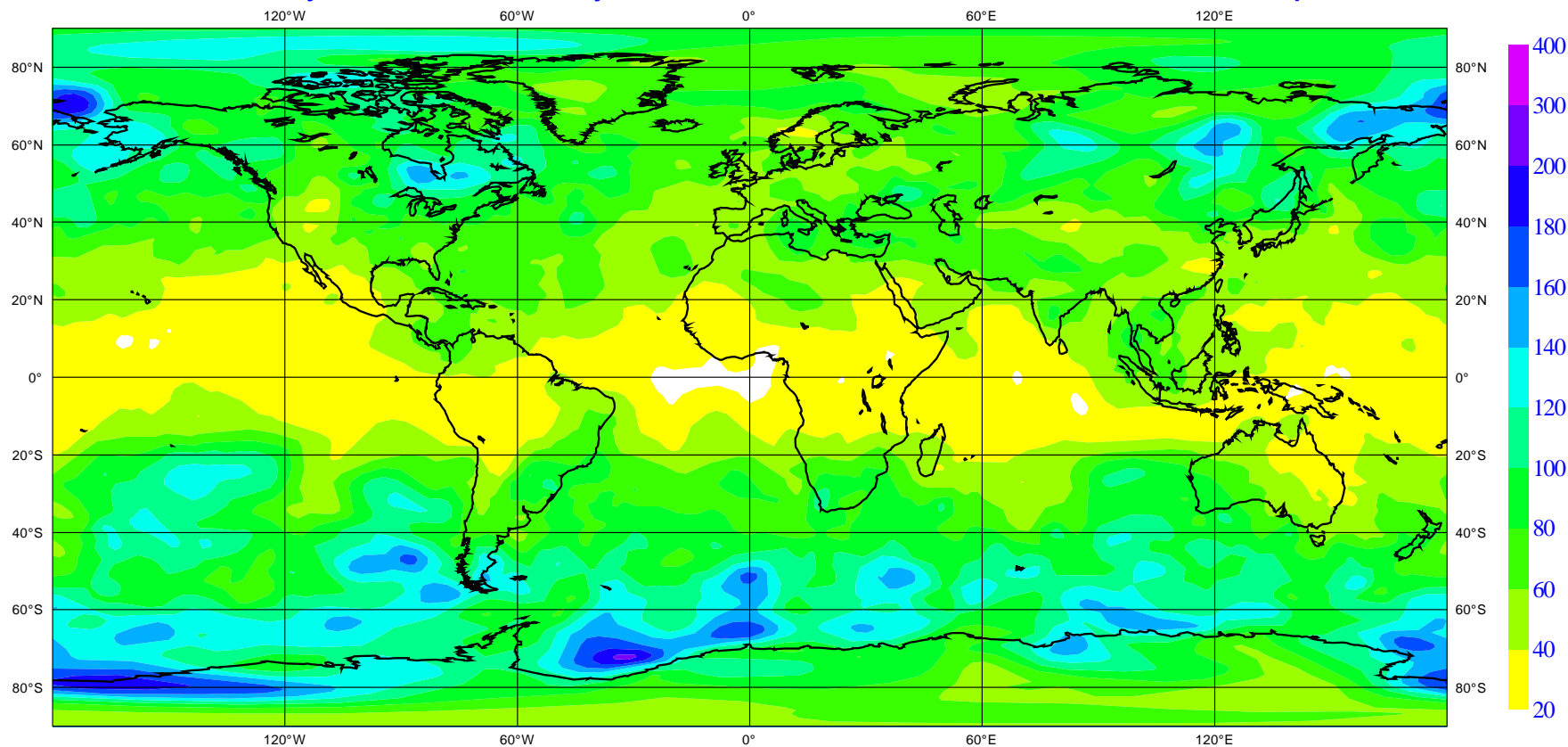
ECMWF Analysis VT:Wednesday 1 October 1997 12UTC 500hPa **Geopotential



Analysis increments 1998 (4D-Var)

4D-Var improved the accuracy significantly

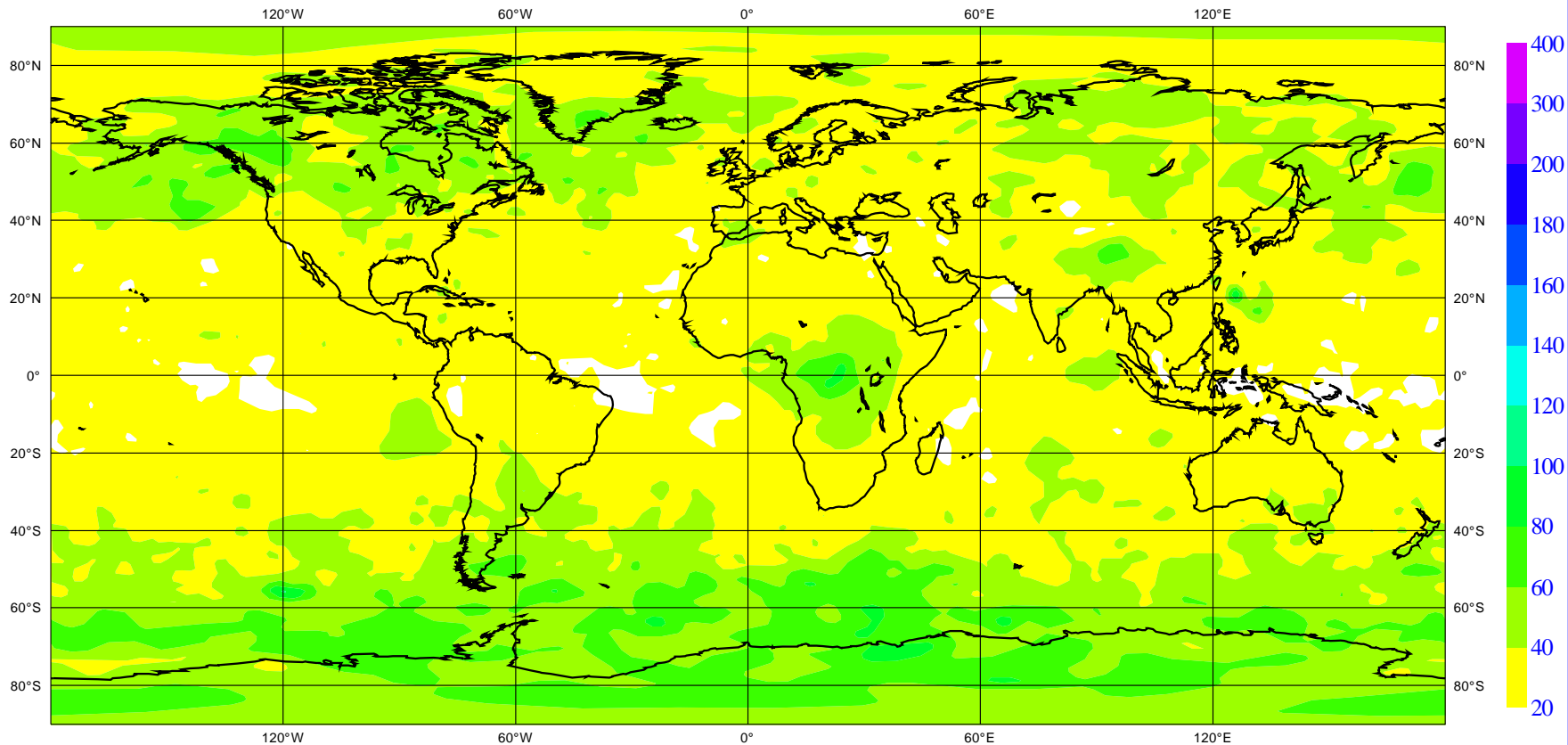
ECMWF Analysis VT:Thursday 1 October 1998 12UTC 500hPa **Geopotential



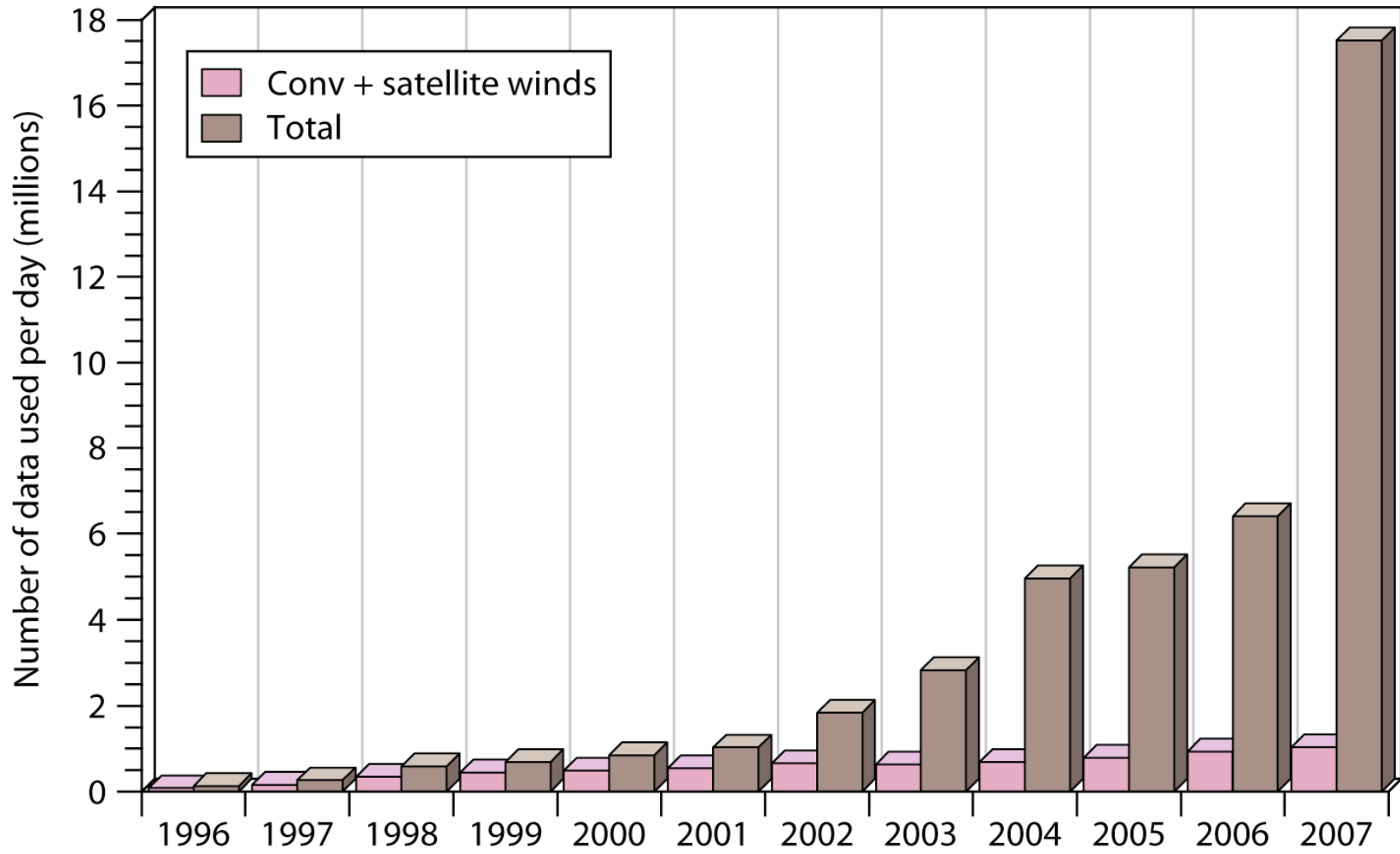
Analysis increments 2007 (Now!)

Small increments globally, due to high data density and a very accurate forecast model

ECMWF Analysis VT:Monday 1 October 2007 12UTC 500hPa **Geopotential

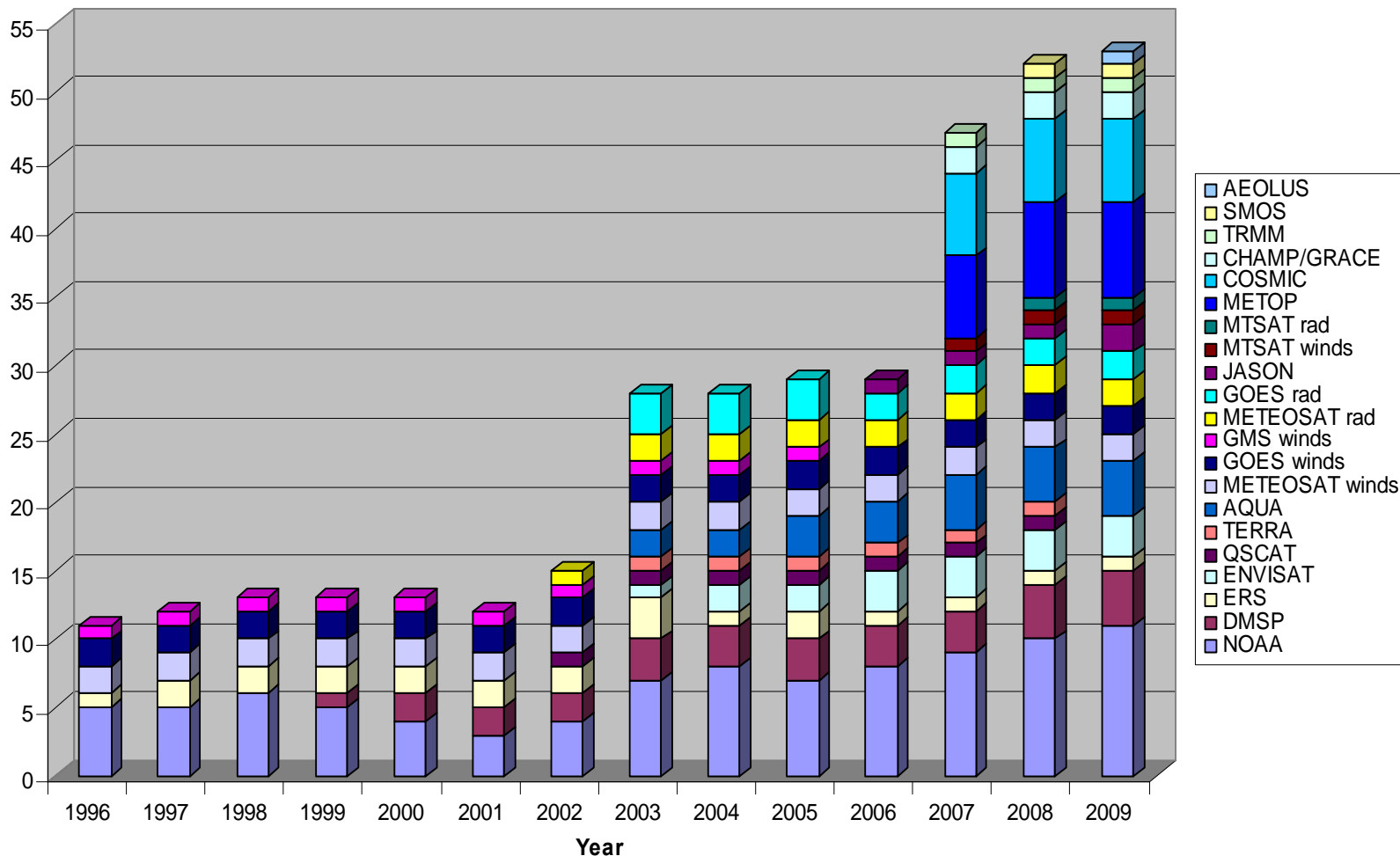


Increasing number of assimilated data



Number of satellite/sensor types to 2009

Number of satellite sources used at ECMWF



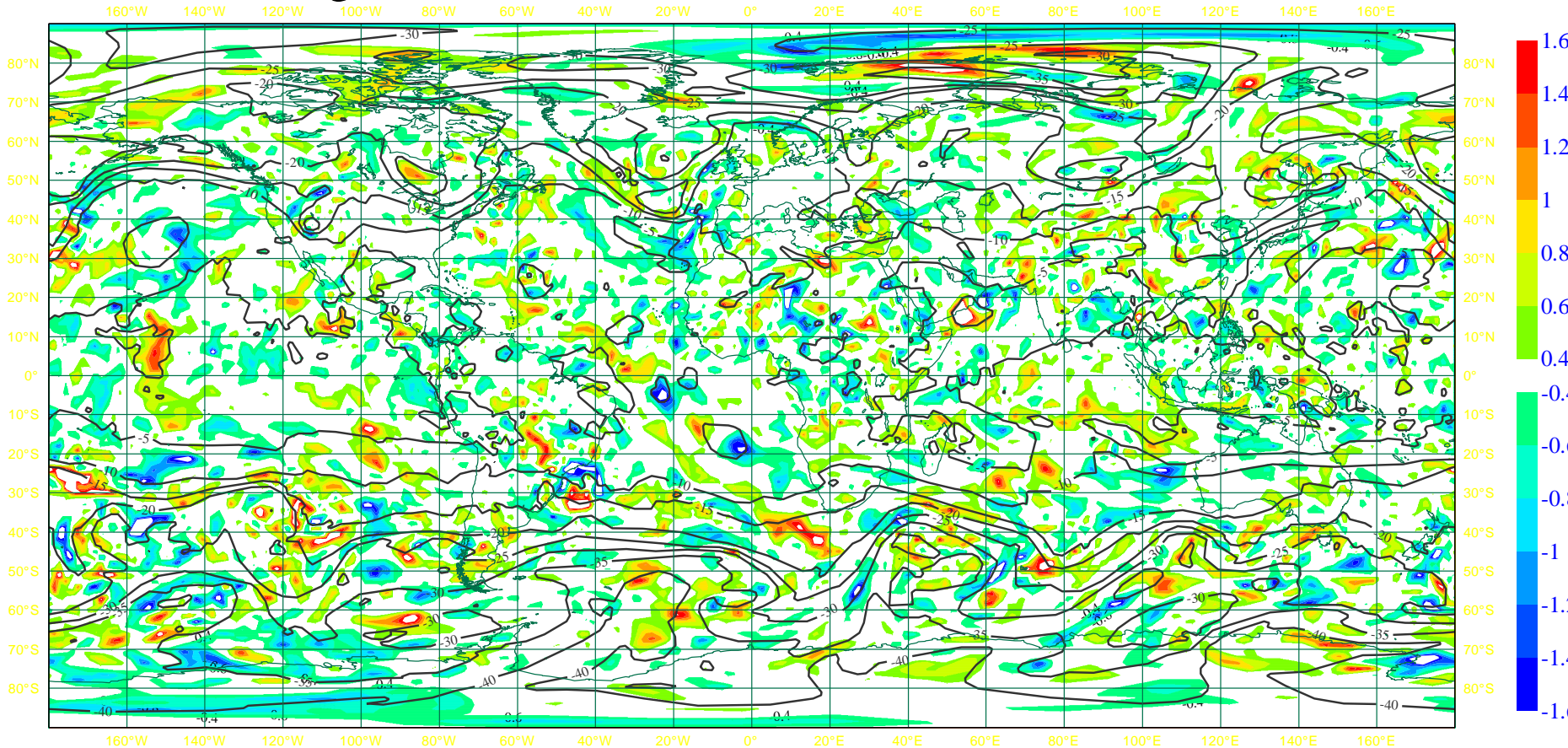
Now that we have so many observations, the balance of priorities is shifting:

- **Initial condition errors are smaller**
- **Analysis increments are smaller**
- **The initial errors are smaller-scale, quasi-random, less structured**
- **The accuracy of observation operators is becoming even more important**
- **Bias correction becomes more important**
- **Importance of Jb diminishes**
- **Need for balance constraints diminishes**
- **A larger part of the short-range fc error is due to moist-physical processes, and less to baroclinic error growth**
- **Timeliness demands increase**

One realisation of a typical 3h forecast difference, from 4D-Var ensemble

Contours: 500hPa T field member 1

Shading: 500hPa T difference, member 2 minus member 1



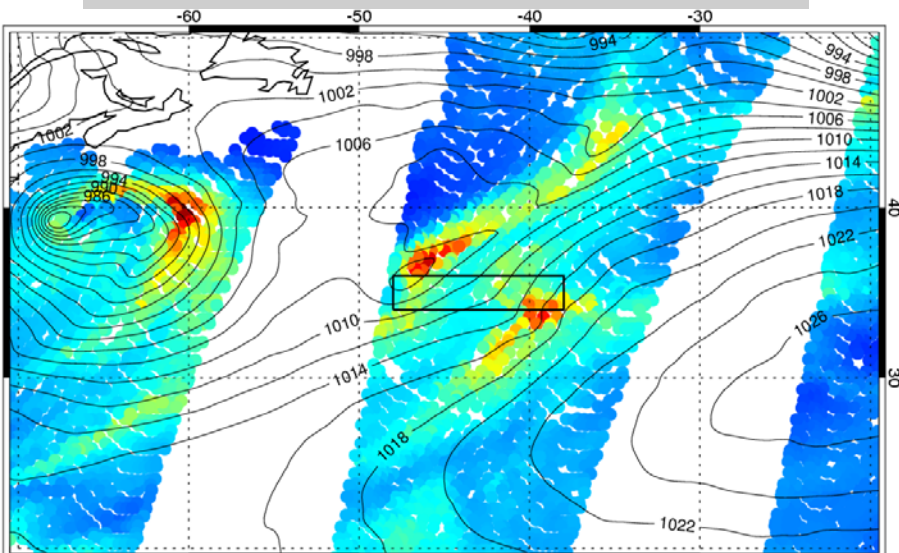
Ongoing 4D-Var research topics

- **Longer assimilation window, with the ultimate goal to emulate the Kalman Smoother, and to realize vanishing sensitivity to initial conditions**
- **Weak constraint 4D-Var**
- **Cloud and rain analysis**
- **Accounting for correlated observation error**
- **Improved Jb balance constraints (which provide improved flow-dependence)**
- **Ensembles of 4D-Var, and their coupling to ensemble prediction**
- **...**

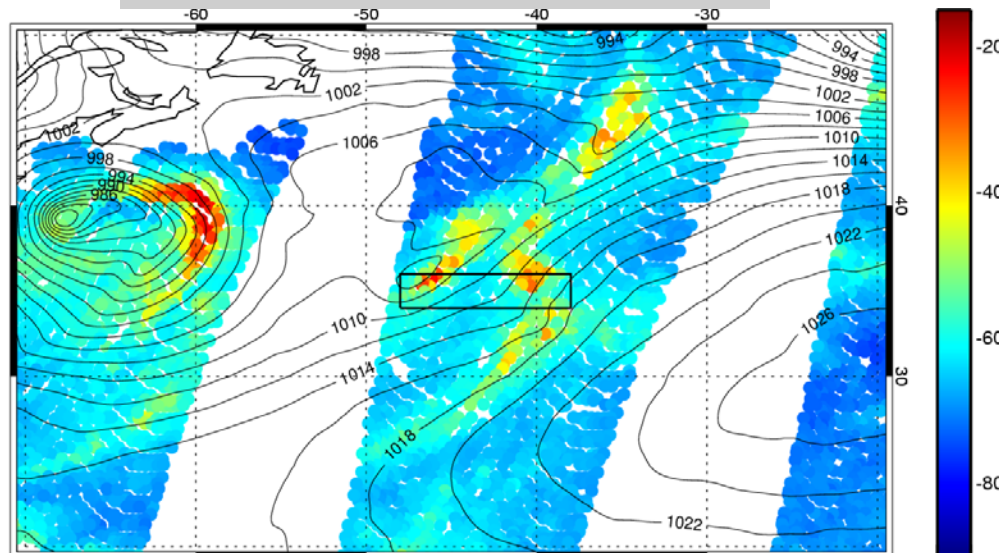
Assimilation of rain-affected microwave radiances

- Assimilation of rain-affected SSM/I radiances in 1D+4D-Var active since June 2005.
- Main difficulties: inaccurate moist physics parameterizations (location/intensity), formulation of observation errors, bias correction, linearity.
- Major improvements accomplished in 2007 and SSMIS, TMI, AMSR-E data included.
- Direct 4D-Var radiance assimilation envisaged for early 2009.

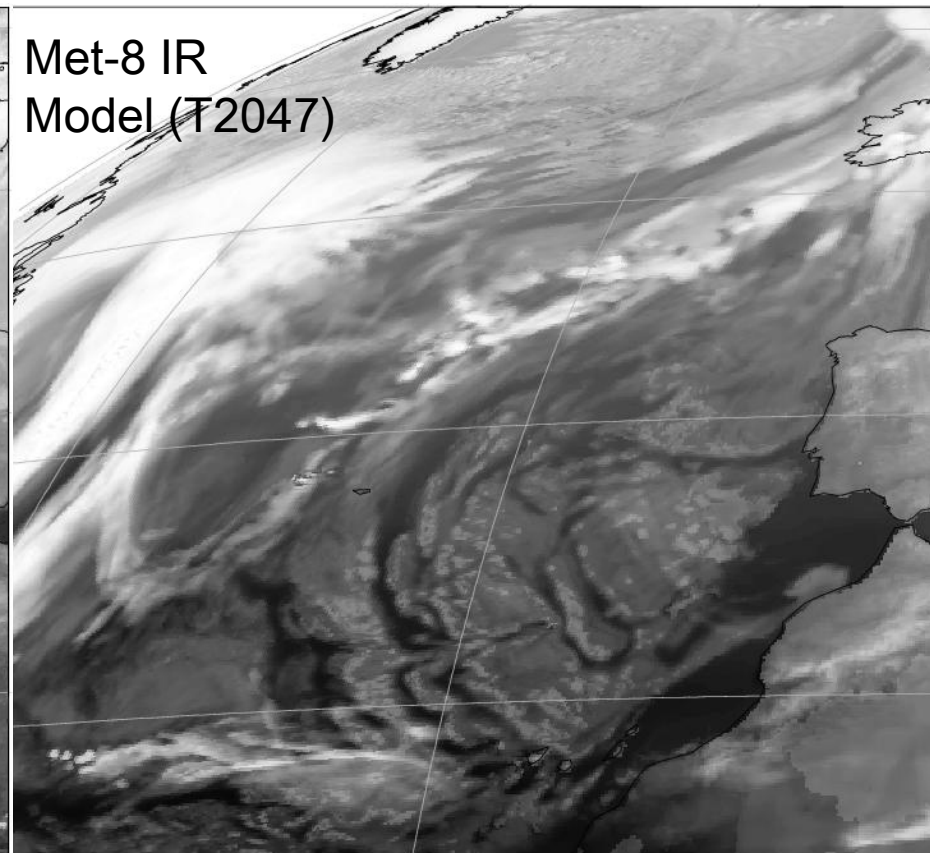
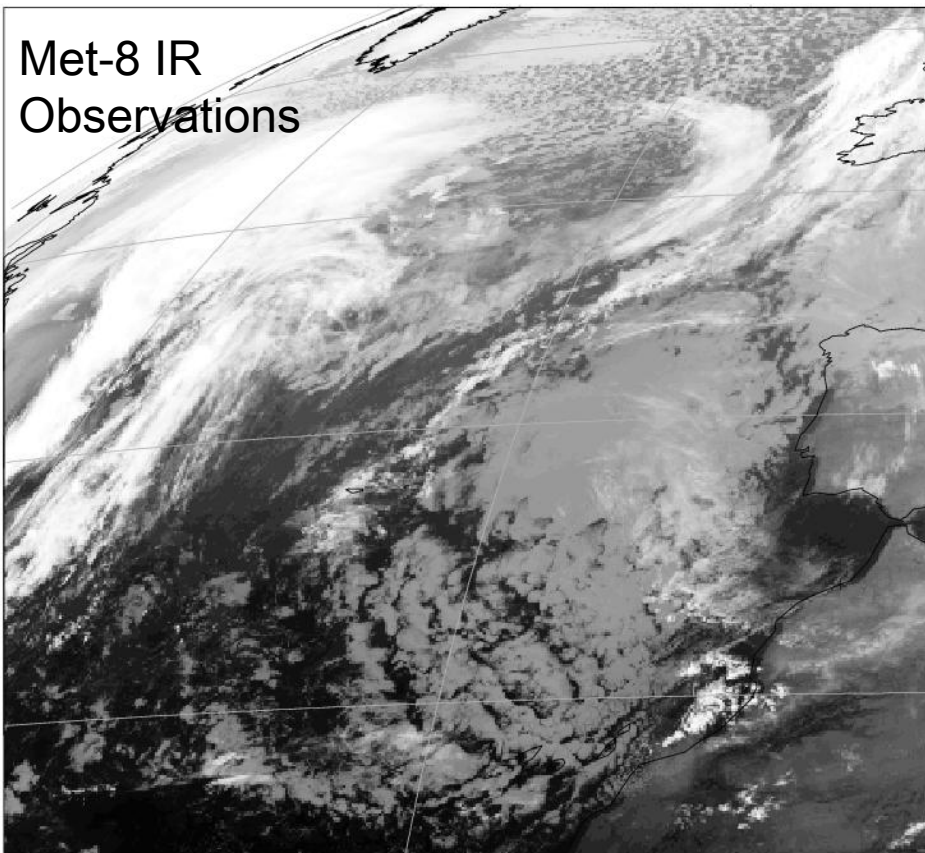
4D-Var first guess SSM/I ΔT_b 19v-19h [K]



SSM/I observational ΔT_b 19v-19h [K]



Assimilation of clouds and precipitations: The model has to look like the observations!



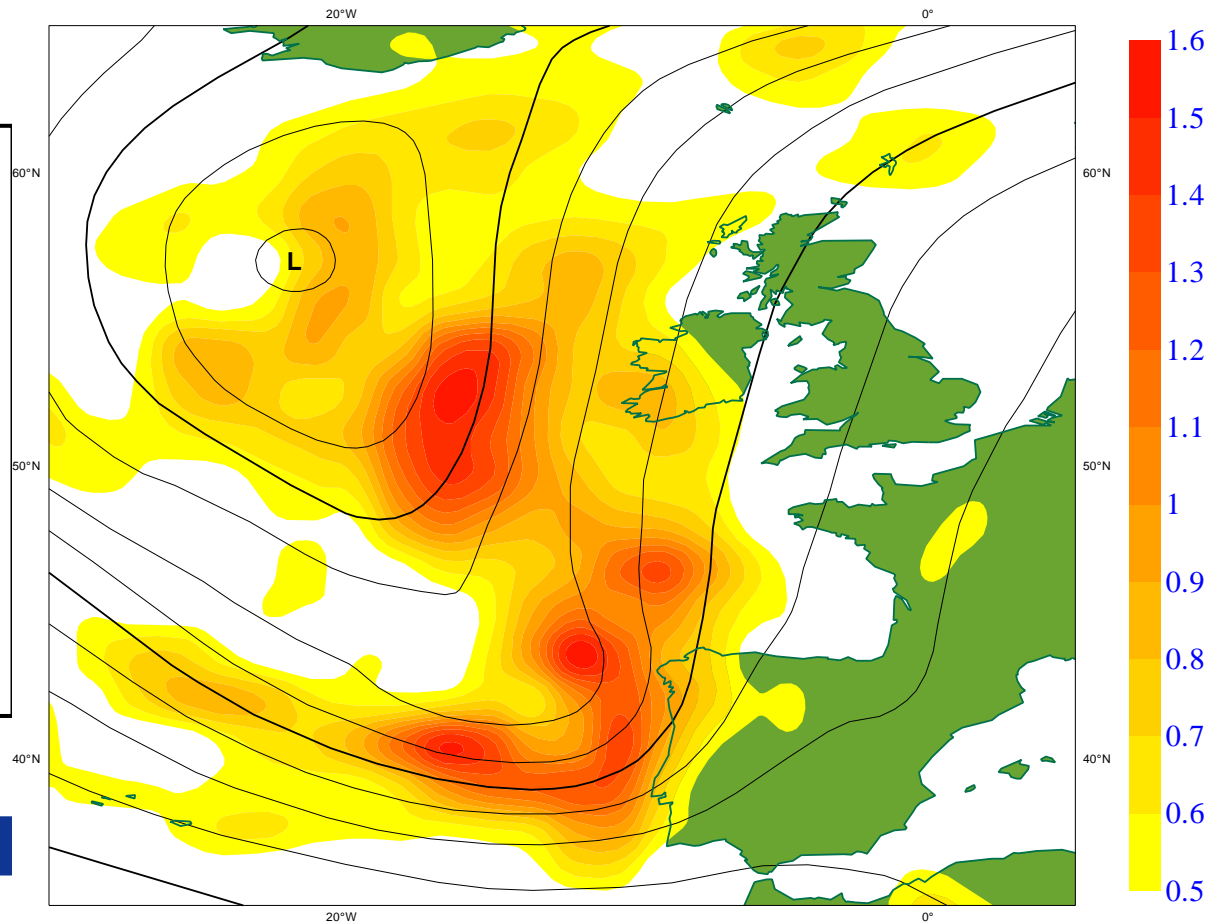
This requires progress in model physics and spatial resolution

Ensembles of data assimilations

- Run an ensemble of 4D-Var assimilations with random observation and SST perturbations, and form differences between pairs of analyses (and forecast) fields.
- These differences will have the statistical characteristics of analysis (and forecast) error (or *uncertainty*).

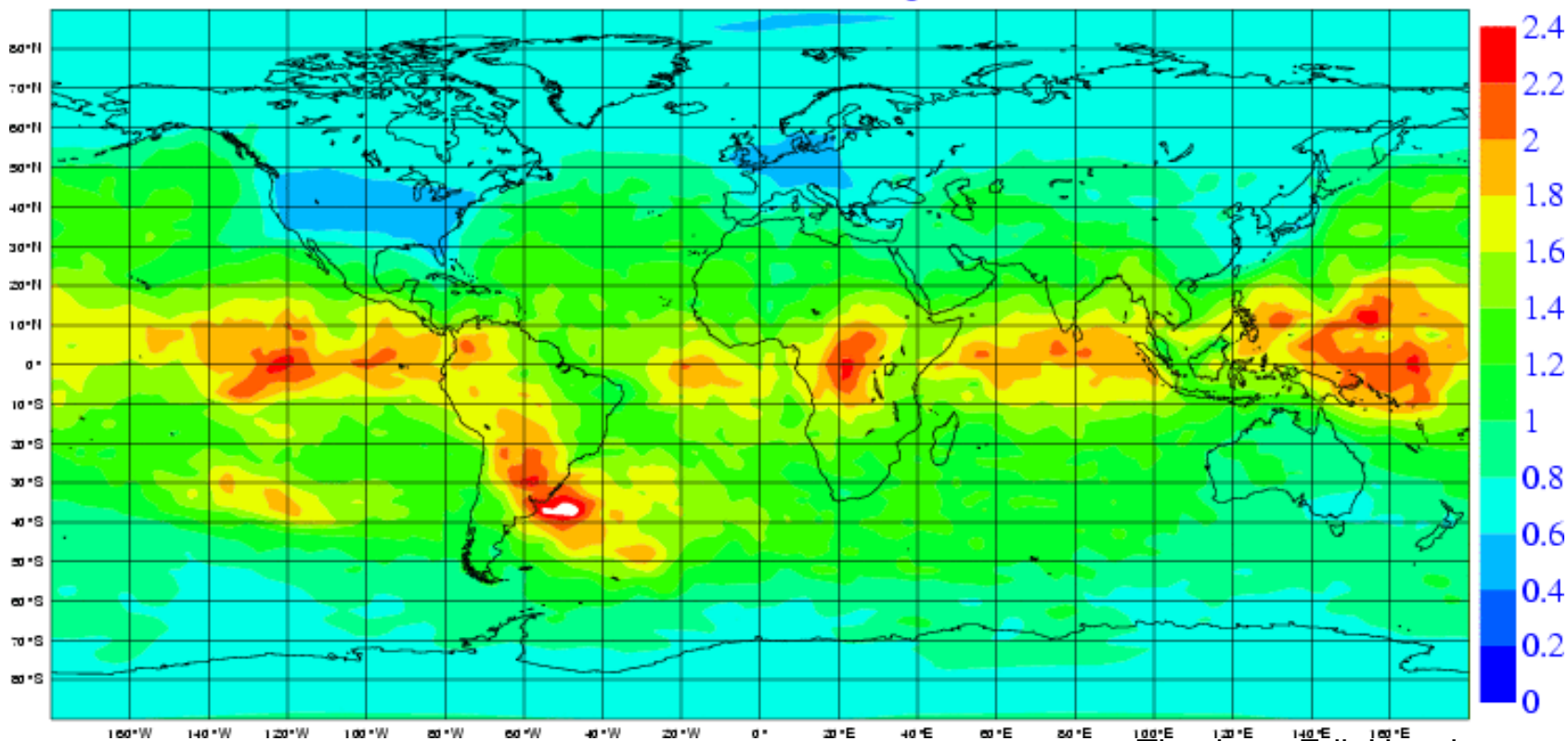
To be used to indicate where good data should be trusted in the analysis (yellow shading).

Also for initialization of the EPS.



Static estimation of analysis error for 200hPa u-wind from the standard deviation of 40 days of 10 member ensemble DA with SPBS looks reasonable. Scaling factor of 1.5-2 required due to underestimation of model errors in ensemble.

Estimation of analysis error (EXP: evh4)
parameter: 200 hPa u (m/s)
period: 22/09/2006 - 31/10/2006
min: 0.40, max: 2.53, average: 1.05



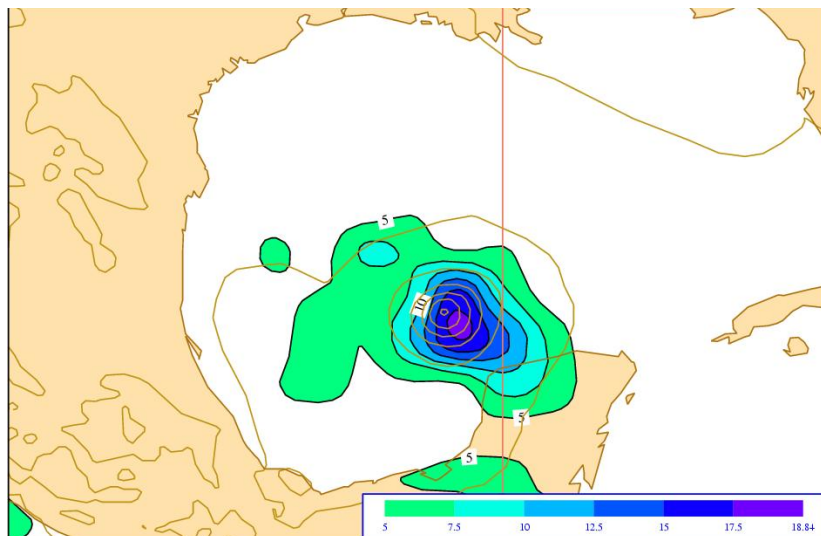
Thanks to Edit Hagel

Hurricane Emily 19-20 July 2005

Ensemble Data Assimilation spread for zonal wind at 850hPa

Ensemble DA 6h forecast spread for 850hPa u-wind
T799 10 member ensemble valid 00UTC 19 July 2005

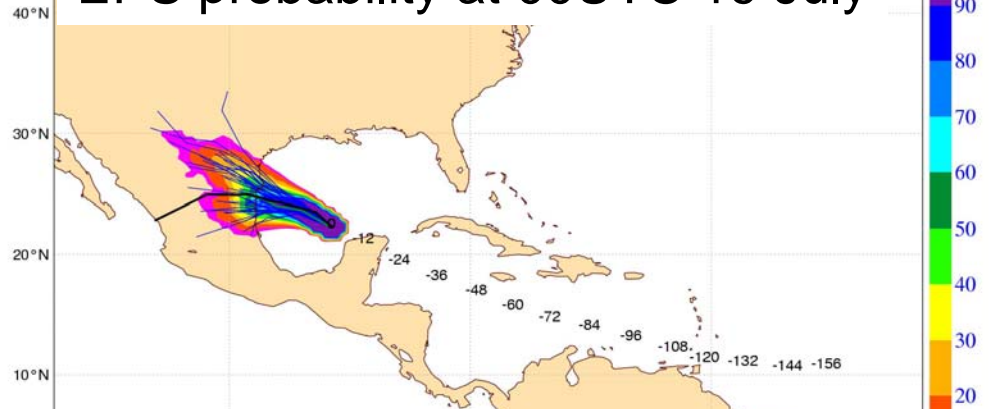
Max. stdev of EnDA spread 19m/s



Status of 4D-Var

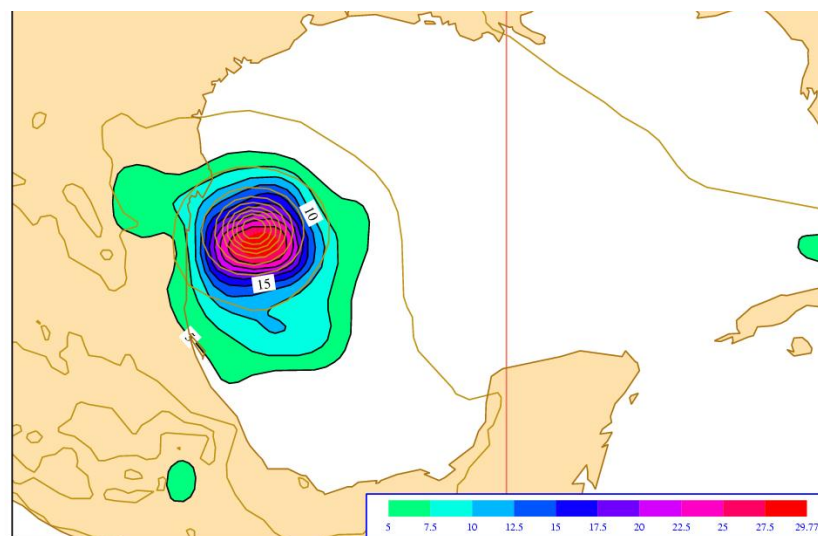
20050719 0 UTC
Probability that EMILY will pass within 120km radius during the next 120 hours
tracks: black=OPER, green=CTRL, blue=EPS numbers: observed positions at t+..h

EPS probability at 00UTC 19 July



Ensemble DA 6h forecast spread for 850hPa u-wind
T799 10 member ensemble valid 00UTC 20 July 2005

Max. stdev of EnDA spread 30m/s



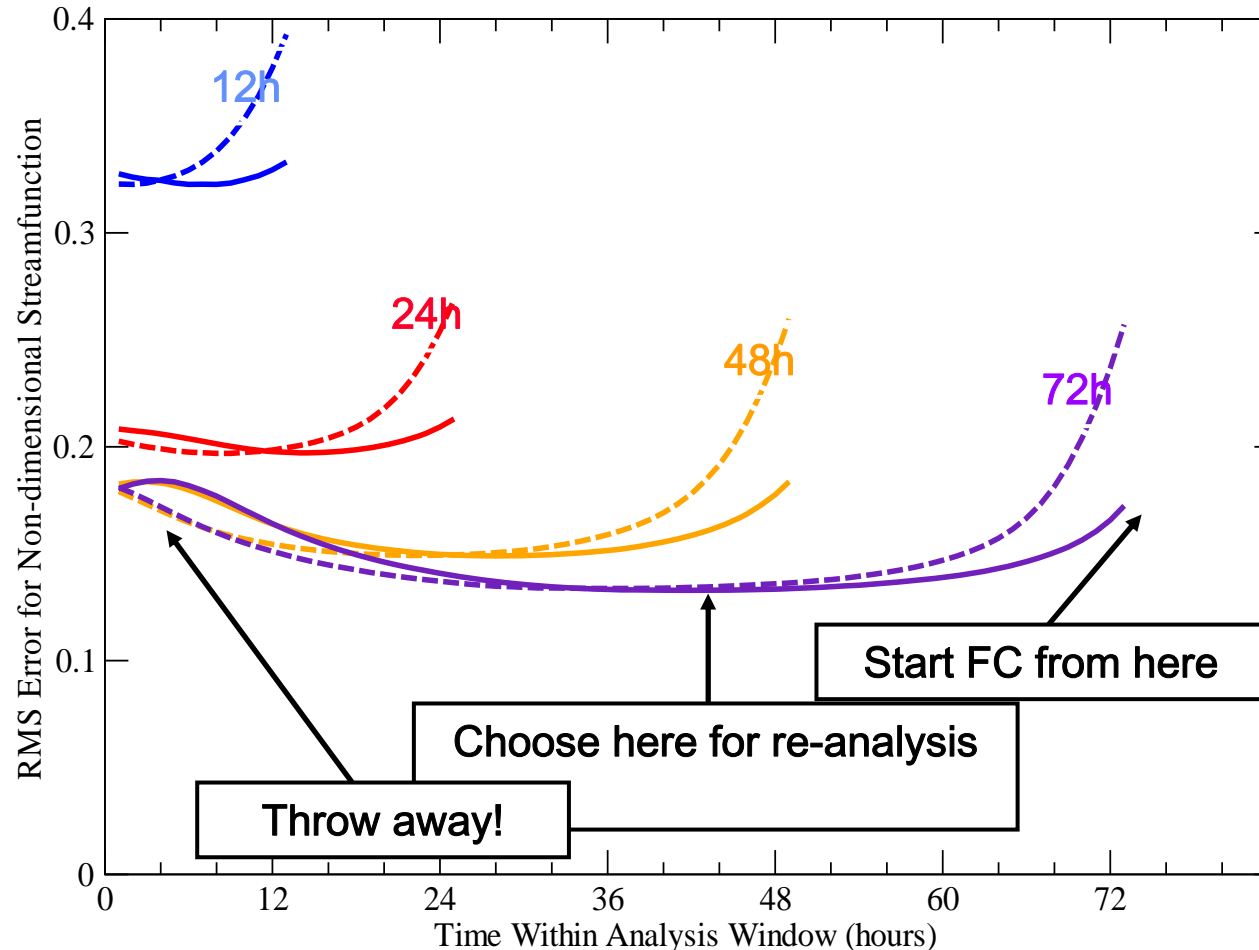
Slide 33

The ultimate goal – a life without Jb?

When the assimilation window is long enough, the addition of the newly arrived data will no-longer change the analysis at the start of the window (Mike Fisher)

Long-Window 4D-Var in a Two-Level QG Model

Mean Analysis and First Guess Error for Different Window Lengths



- Ignore the i.c. analysis increments
- Jb irrelevant!
- Instead: estimates model error
- Requires specification of model error
- This is TRUE 4D VAR!

Known assumptions and approximations

- **Only ‘weakly’ non-linear**
- **Model is perfect (in strong-constraint 4D-Var)**
- **Near-Gaussian statistics**
- **No multiple minima near the first-guess state**
- **That the outer-loop iterations converge**
- **That suitably transformed variables can be defined that display globally nearly-homogeneous covariance structure**
- **That perturbations evolve ‘smoothly’.**
- **That observation errors are uncorrelated (up to now)**
- **...**

Summary: Main 4D-Var characteristics

- **Has enabled NWP centres around the world to utilize a wide variety of observations, contributing to improved NWP performance**
- **Require TL/AD models, and these are useful also for ‘adjoint sensitivity’ and singular vectors**
- **Generates suitably balanced analyses**
- **Copes well with millions of observations**
- **Runs well on 1000s of processors**
- **Accounts for non-linearities**
- **Proven successful for climate re-analyses and environmental monitoring**
- **Allows serial correlations (in both R and Q)**

The main current challenges

- **How to extend parallelism to >10,000 processors**
- **Computational cost, especially for LAM applications**
- **Linearisation of small-scale moist physical processes**
- **Improved coupling between moisture and dynamic variables**
- **Improved analysis of the boundary layer**
- **Improved wind analysis in the tropics (in the lack of direct wind observations)**

Finally

- **This concludes my talk**
- **Thank you for years of stimulating exchanges!**
- **This concludes my contribution to data assimilation, ...**
- **I've moved to a new job where I now manage operational systems (still at ECMWF)**
- **Now I can look forward to a few relaxing days here in Buenos Aires**
- **Then, hope to see you all again in some other (related?) context!**

Going beyond the 'basic' set-up

- 1. Efficient and accurate solution algorithms (multi-resolution) with inner/outer loops to T255**
- 2. Parallel, large-volume observation processing**
- 3. Adaptation to various computer architectures, up to >1000 processors**
- 4. Masses of new observations added!**
- 5. Wavelet Jb, Grid-point Jb**
- 6. Various 'balance' constraints in Jb**
- 7. Humidity, ozone analyses**
- 8. TL moist physics**
- 9. Observation bias correction schemes**
- 10. Weak-constraint, accounting for model error**
- 11. Adjoint-sensitivity diagnostics of obs impact**
- 12. Ensembles of 4D-Var**
- 13. Climate re-analyses**
- 14. Aerosol, CO₂, NH₄, SO₂, ...**
- 15. ...**

The innovation covariance

The innovation covariance can be written

$$\langle \mathbf{d}, \mathbf{d}^T \rangle = \hat{\mathbf{H}} \hat{\mathbf{P}}^f \hat{\mathbf{H}}^T + \hat{\mathbf{O}} + \hat{\mathbf{F}} - (\hat{\mathbf{H}} \hat{\mathbf{X}}^T + \hat{\mathbf{X}} \hat{\mathbf{H}}^T)$$

with

(Joiner and Dee, *QJ* 2000)

$$\hat{\mathbf{P}}^f = \hat{\mathbf{M}} \hat{\mathbf{B}} \hat{\mathbf{M}}^T + \hat{\mathbf{Q}}$$

Confusion surrounding ‘*Model error*’:

- **Q = Model error, due to imperfections in M**
- **MBM^T = Predictability error, due to evolution of errors in the initial conditions**
- **P^f = MBM^T + Q = Forecast error**
- **B = Bg-error = Initial condition error**

4D-Var approximations

In our 4D-Var the 'true' co-variances are approximated:

$$\hat{\mathbf{O}} + \hat{\mathbf{F}} \approx \mathbf{R}$$

R diagonal

$$\hat{\mathbf{X}} \approx 0$$

No cross co-variances

$$\hat{\mathbf{Q}} \approx 0$$

Perfect model assumption

$$\hat{\mathbf{H}}\hat{\mathbf{P}}^f\hat{\mathbf{H}}^T \approx \mathbf{H}\mathbf{M}\mathbf{B}\mathbf{M}^T\mathbf{H}^T$$

Tangent linear obs. operators

Tangent linear forecast model

4D-Var: $\mathbf{H}\mathbf{M}\mathbf{B}\mathbf{M}^T\mathbf{H}^T + \mathbf{R}$

TL dynamics?

TL physics?

3D-Fgat: $\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}$

Truncation?

Ol: $\mathbf{B}_o + \mathbf{R}$

B given through \mathbf{J}_b -modelling

Validation

From samples of innovations we can compute

$$\langle \mathbf{d}, \mathbf{d}^T \rangle$$

If we knew how to diagnose

$$\mathbf{HMBM}^T \mathbf{H}^T + \mathbf{R}$$

in the full 4D-Var system,

then the two could be compared, and some shortcomings due to the modelling assumptions might become apparent.

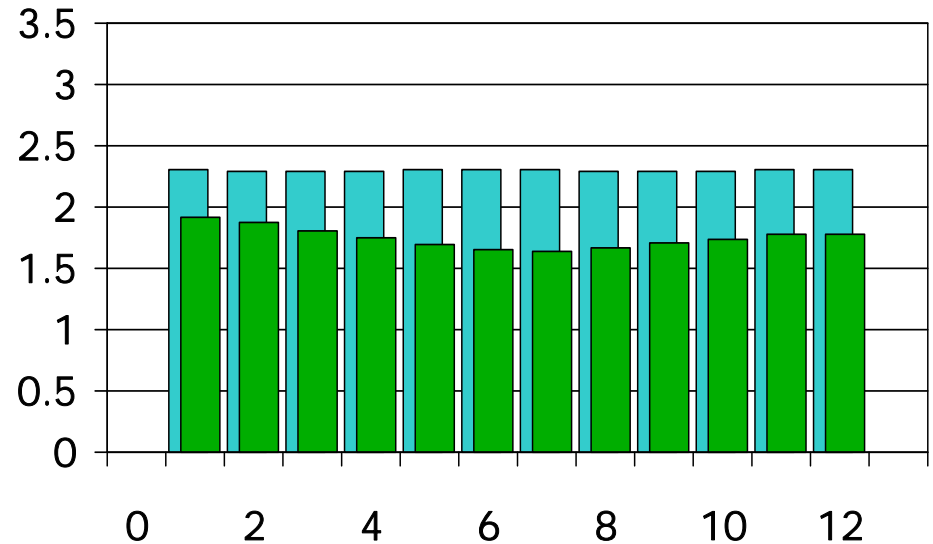
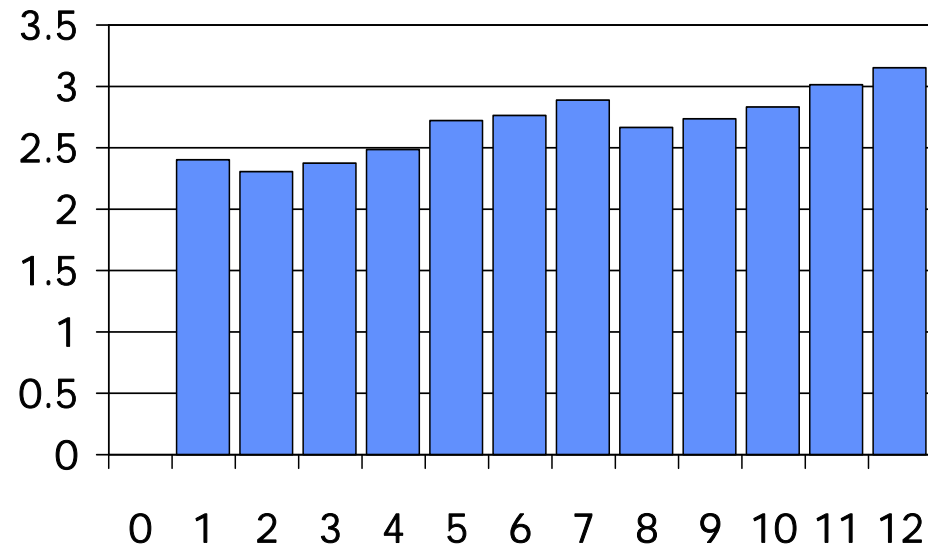
Discrepancies could be due to H, M, B, R or Q !!!

In a 'well-tuned' system: $\langle \mathbf{d}, \mathbf{d}^T \rangle \approx \mathbf{HMBM}^T \mathbf{H}^T + \mathbf{R} (+\mathbf{Q})$

American wind profilers, U-component (m/s), 300-200 hPa

■ Obs-Bg

■ R ■ HMBMH



20030205-12 to 20030211-12,

About 12,000 data per bin