# DEVELOPMENT OF THE ECMWF FORECASTING SYSTEM

Jean-Noël Thépaut

European Centre for Medium Range Weather Forecasts
ECMWF

**Acknowledgements: Lars Isaksen, Mike Fisher, Yannick Trémolet, Peter Bauer, Adrian Simmons, Martin Miller, Sakari Uppala, and many other colleagues from the Research and Operational Departments**

# Outline

- Components of the ECMWF forecasting system
- Performance of the NWP system
- Other applications
- Future evolutions and challenges

# Outline

- Components of the ECMWF forecasting system
- Performance of the NWP system
- Other applications
- Future evolutions and challenges

# The operational forecasting system

- **Data assimilation:** twice per day
  12-hour (6-hour) 4D-Var 25 km 91-level; 210/125/80 km minimisations
- **High resolution deterministic forecast:** twice per day
  25 km 91-level, to 10 days ahead
- **Ensemble forecast (EPS):** twice daily
  51 members, 62-level, 50 km to 10 days, then 80 km to 15 days
- **Ocean waves:** twice daily
  Global: 10 days ahead at 40 km; EPS 15 days ahead at 100 km
  European Waters: 5 days ahead at 25 km
- **Monthly forecast:** once a week (coupled to ocean model)
  51-members, 50/80 km 62 levels, to one month ahead
- **Seasonal forecast:** once a month (coupled to ocean model)

  41 members, 125 km 62 levels, to seven months ahead
- **Boundary Conditions:** short cut-off analyses based on 6-hourly 4D-Var initiating a forecast to 3 days, four times per day



ECMWF

# Breakdown of core operational computer usage

| | 1994 | 2008 |
|---|---|---|
| 24h data assimilation | 20% | 37% |
| 10-day deterministic forecast | 40% | 18% |
| Ensemble forecasts | 40% | 45% |

**The issues of computer performance and scalability of the ECMWF NWP system will be addressed by Deborah Salmond and Mats Hamrud**

# Over the last two/three years, forecasting system developments have included

- **T799/L91 higher-resolution forecast system.**
- **Variable-resolution ensemble prediction system (VAREPS) to 15 days.**
- **Significant improvements of model physics.**
- **New satellite data assimilated:**
  - **METOP-A instruments,**
  - **MTSAT AMVs + COSMIC GPS radio occultation,**
  - **More microwave radiances (AMSR-E, TMI and SSMIS),**
  - **More SBUV ozone retrievals and monitoring of OMI (AURA).**
- **New moist linear physics in 4D-Var, and 3rd outer loop: now minimizing at T95 → T159 → T255.**
- **Better treatment of satellite data in the presence of rain and clouds**

# Observation data count for one 12h 4D-Var cycle 0900-2100UTC 3 March 2008

| Screened | | | Assimilated | | |
|---|---|---|---|---|---|
| Synop: | 450,000 | 0.3% | Synop: | 64,000 | 0.7% |
| Aircraft: | 434,000 | 0.3% | Aircraft: | 215,000 | 2.4% |
| Dribu: | 24,000 | 0.02% | Dribu: | 7,000 | 0.1% |
| Temp: | 153,000 | 0.1% | Temp: | 76,000 | 0.8% |
| Pilot: | 86,000 | 0.1% | Pilot: | 39,000 | 0.4% |
| AMV's: | 2,535,000 | 1.6% | AMV's: | 125,000 | 1.4% |
| Radiance data: | 150,663,000 | 96.9% | Radiance data: | 8,207,000 | 91.0% |
| Scat: | 835,000 | 0.5% | Scat: | 149,000 | 1.7% |
| GPS radio occult. | 271,000 | 0.2% | GPS radio occult. | 137,000 | 1.5% |
| TOTAL: | 155,448,000 | 100.00% | TOTAL: | 9,018,000 | 100.00% |

99% of screened data is from satellites

96% of assimilated data is from satellites

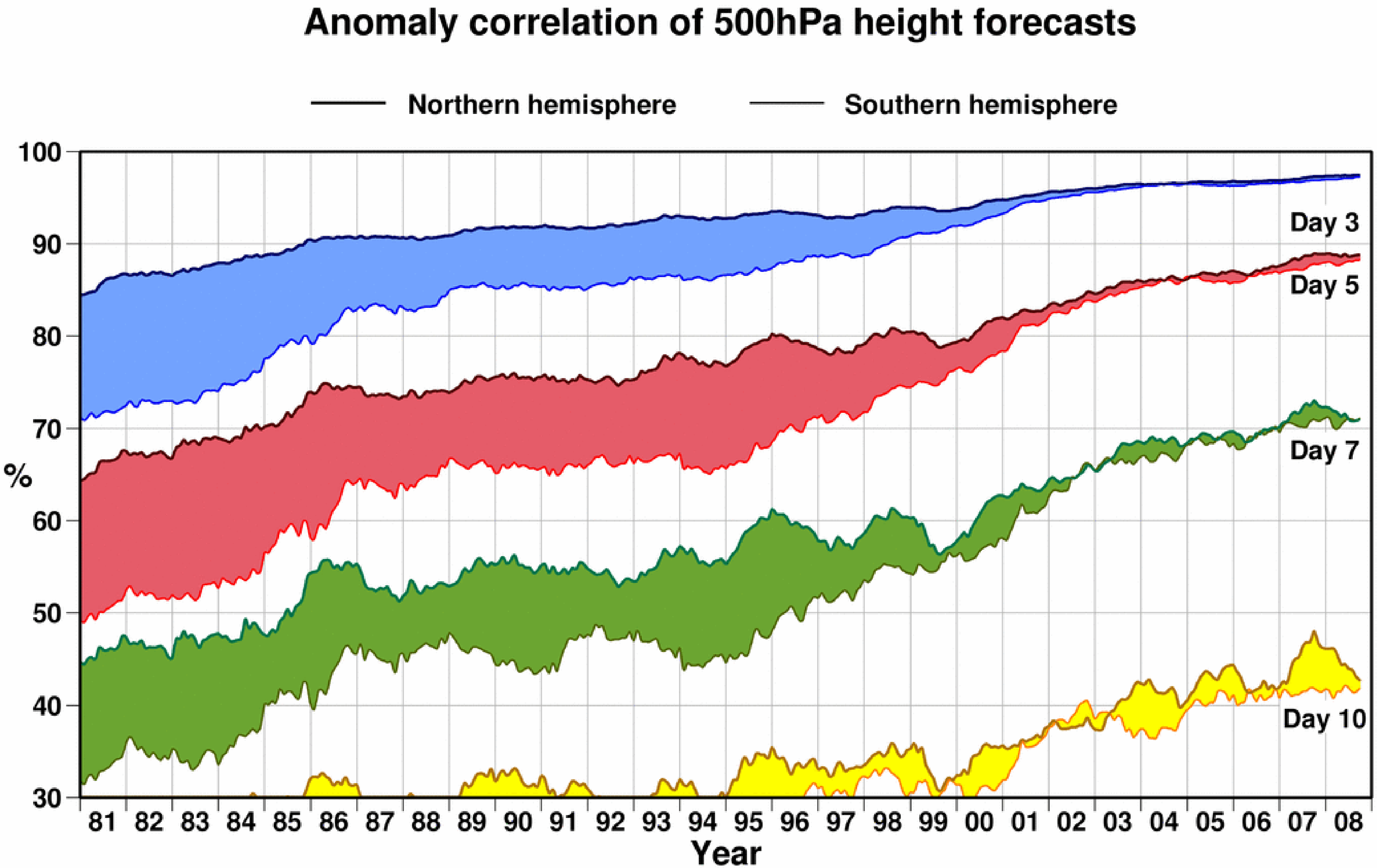


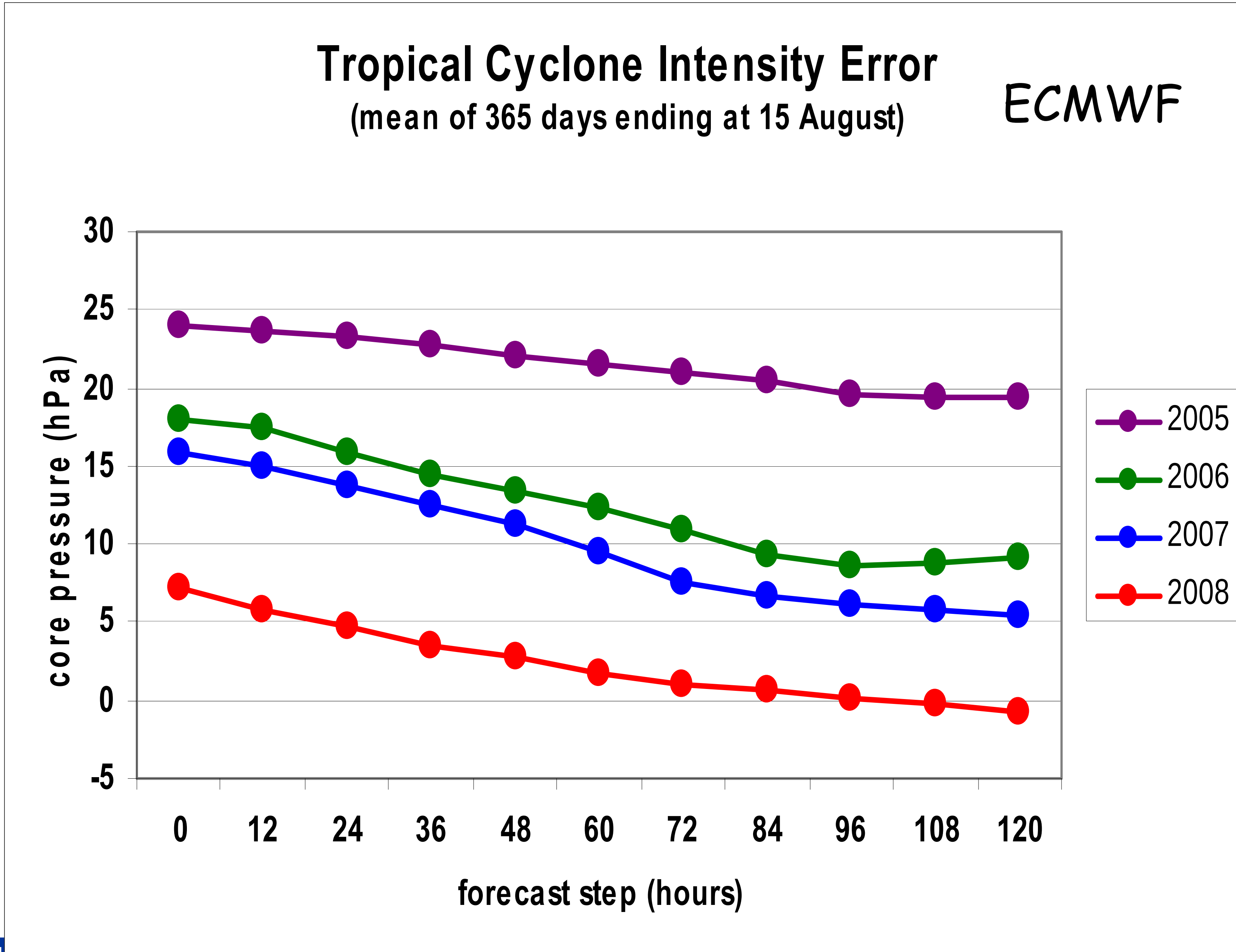
ECMWF

# Outline

- Components of the ECMWF forecasting system
- **Performance of the NWP system**
- Other applications
- Future evolutions and challenges

# Improvement of ECMWF forecasts



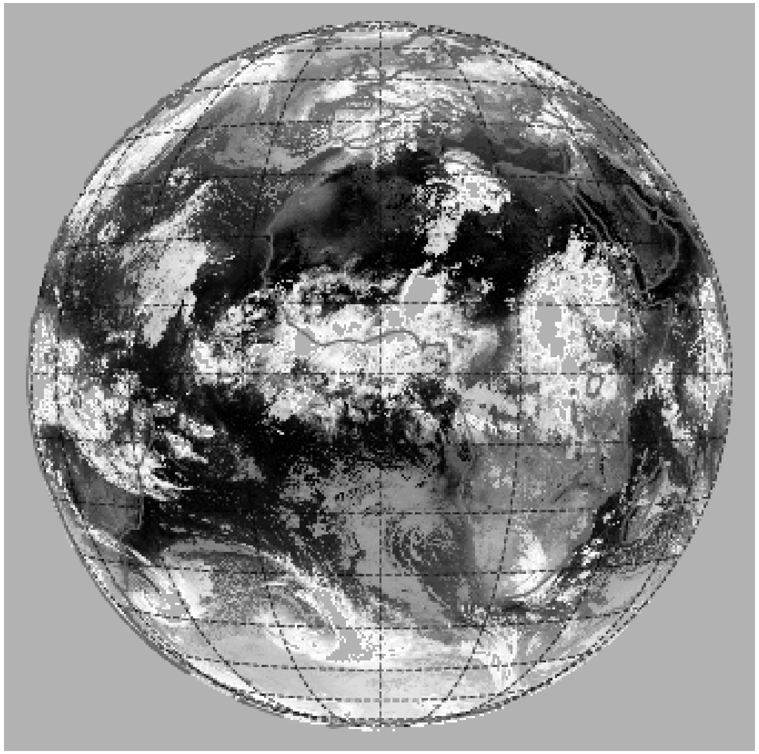

Tropical Cyclone Intensity Error
(mean of 365 days ending at 15 August)
ECMWF
core pressure (hPa)
30
25
20
15
10
5
0
-5
0
12
24
36
48
60
72
84
96
108
120
forecast step (hours)
2005
2006
2007
2008

# Simulated Meteosat imagery

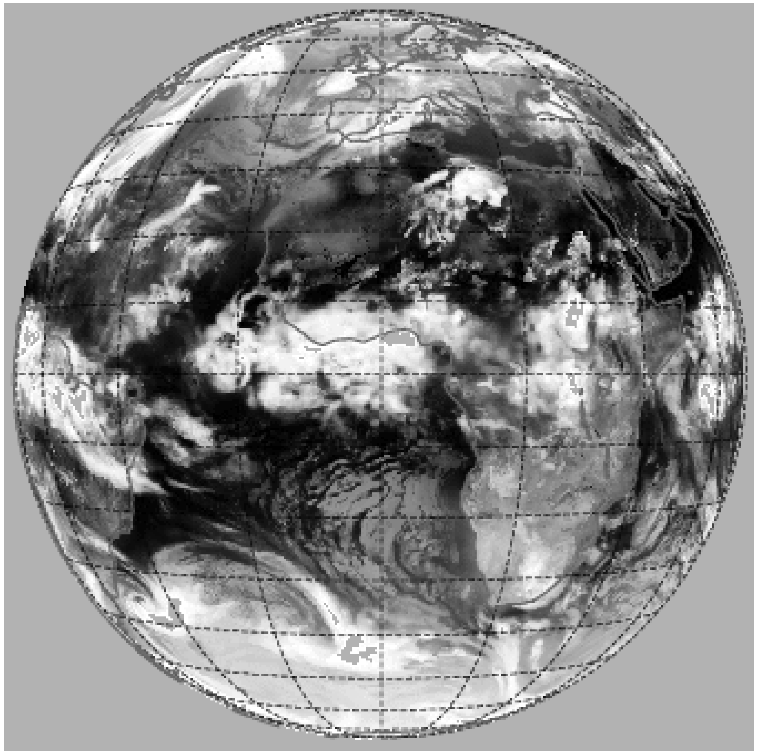
T799 36h forecast from 20080525

(Bechtold 2008)

Meteosat 9 IR10.8 20080525 0 UTC



RTTOV gen. Meteosat 8 IR10.8 ECMWF Fc 20080525 00 UTC:



# Outline

- Components of the ECMWF forecasting system
- Performance of the NWP system
- **Other applications**
- Future evolutions and challenges

# Other applications: reanalyses

To improve the understanding of

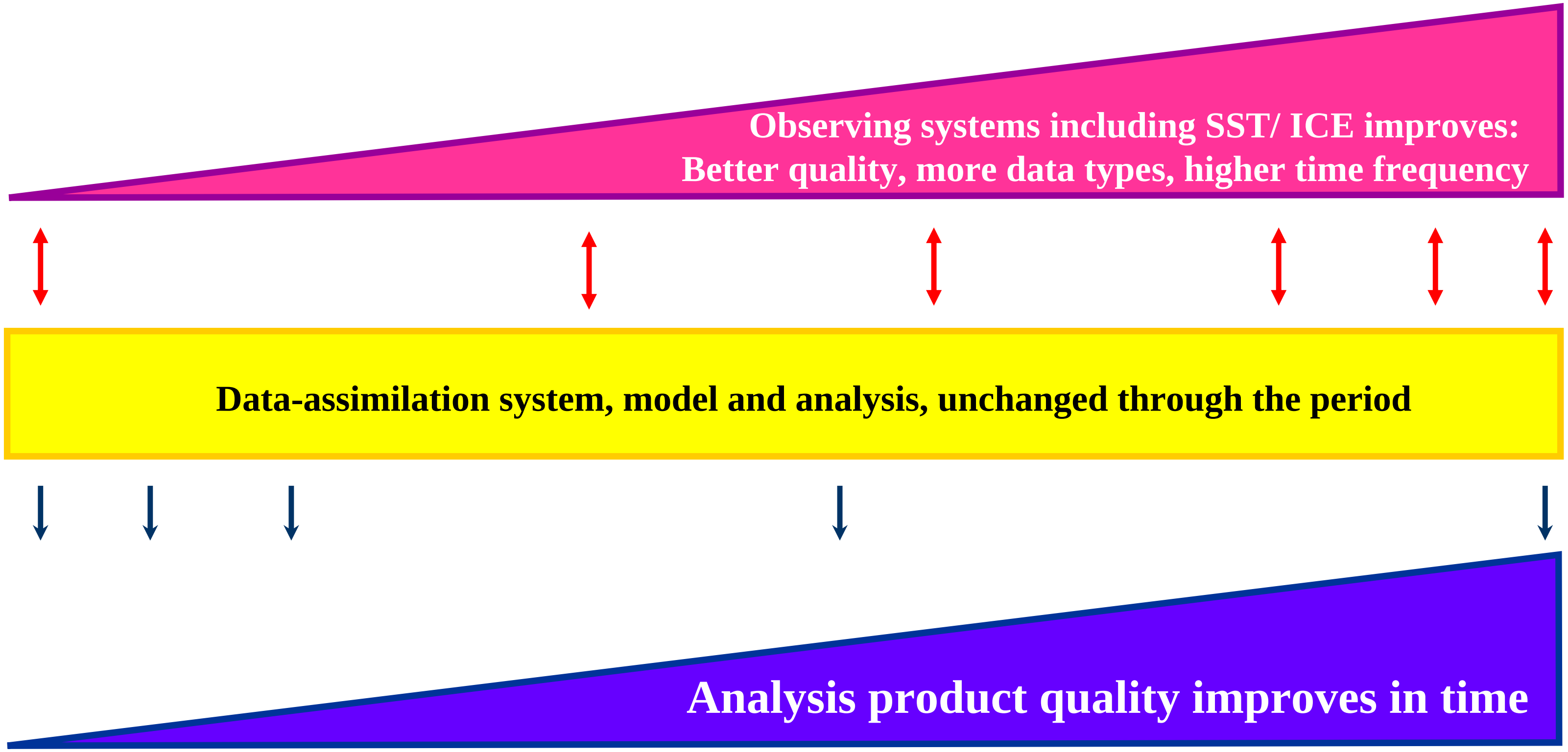
- Weather, climate and general circulation of atmosphere
- Predictability from daily to seasonal, long term variability and climate trends
- Tele-connections
- Atmospheric transport
- Hydrological cycle and surface processes
- Extreme weather, storm tracking, tropical cyclones, …

To provide initial states, external forcing or validation data for

- Climate model integrations
- Ocean models
- Monthly and seasonal forecasting
- Chemical transport models
- …

A substitute for “observed statistics”? An ideal tool to produce and monitor Essential Climate Variables?

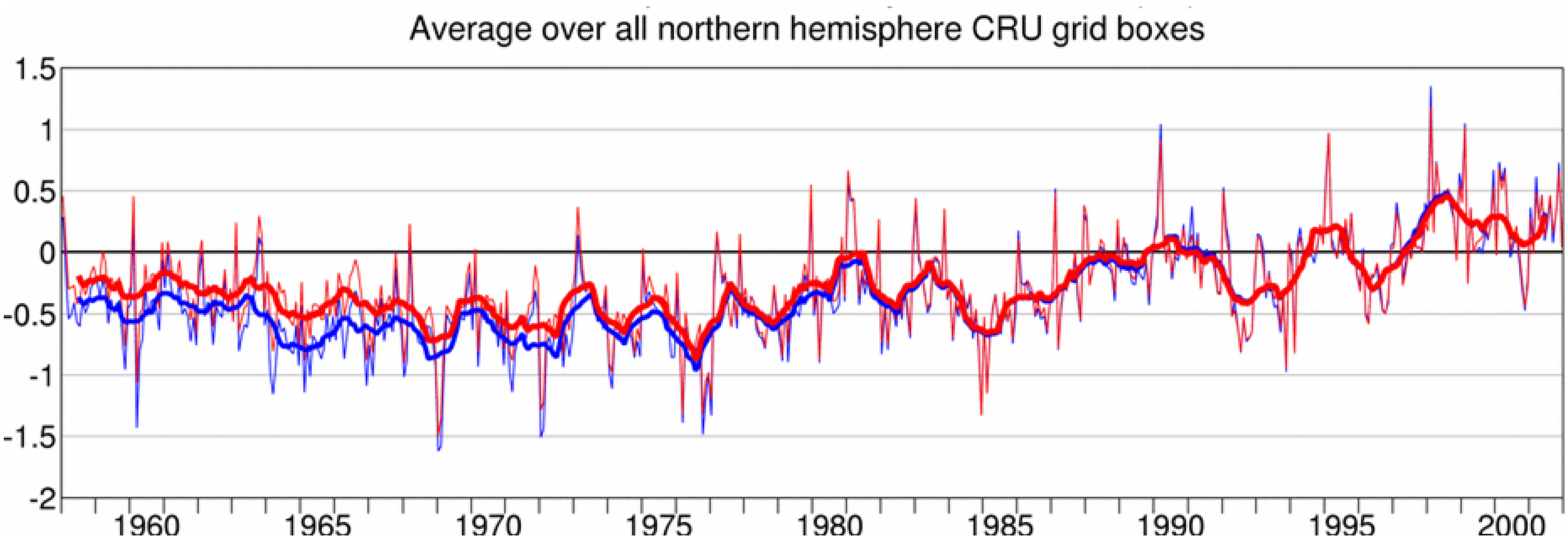
# Reanalysis schematically






ECMWF

# Trend and variability in two-metre temperature




CRUTEM2v (Jones and Moberg, 2003)

ERA-40

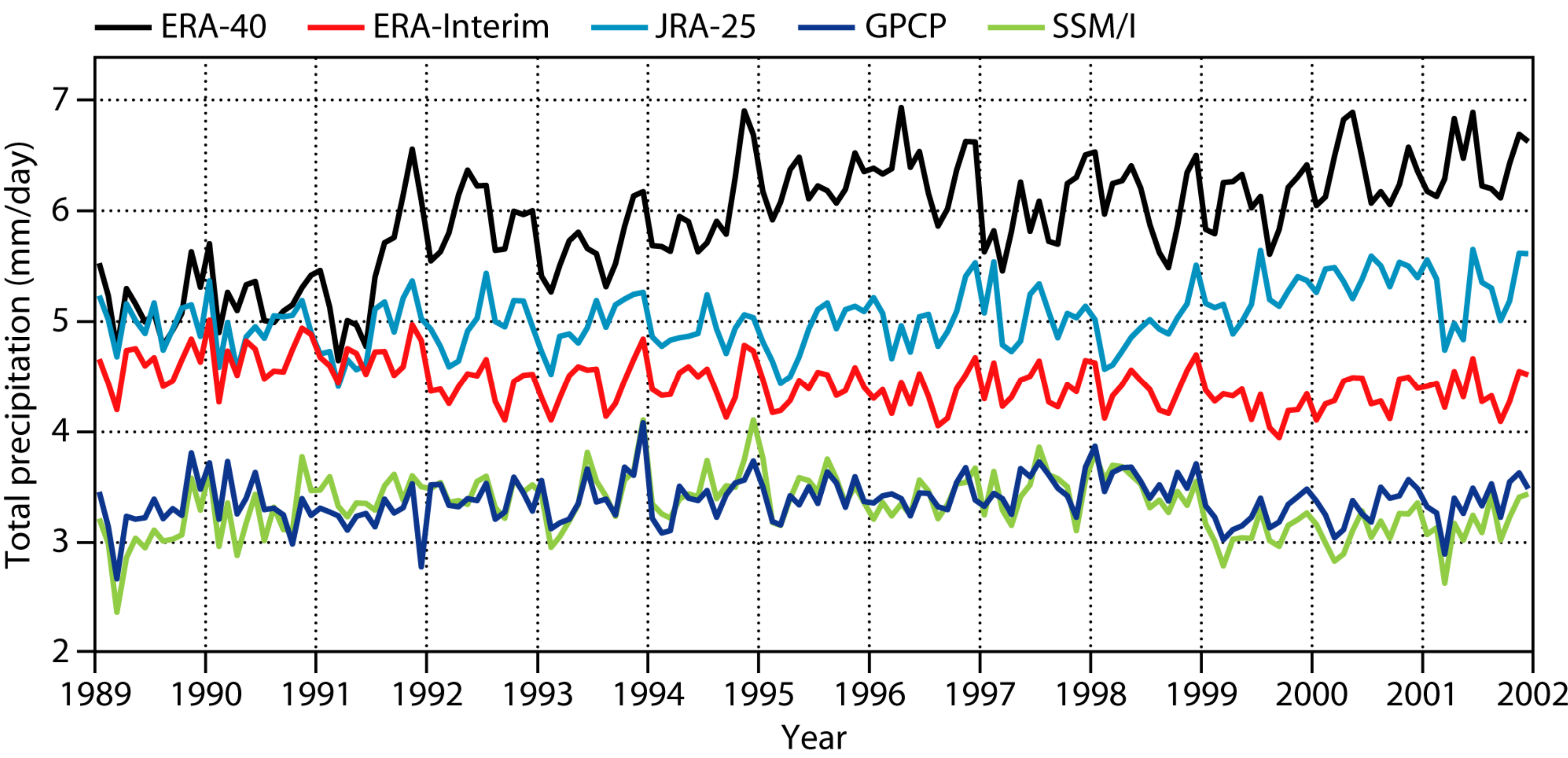
| | | |
|---|---|---|
| **Linear trend (1979-2001):** | **CRUTEM2v** | **0.31$^{O}$C/decade** |
| | **ERA-40** | **0.28$^{O}$C/decade** |
| | **NCEP** | **0.19$^{O}$C/decade** |

## ERA-Interim 1989 → to continue as CDAS →

## ERA-40 1957-2002

- Data-assimilation system
  - T159L60 → T255L60 / 12 hour 4D-Var
  - New humidity analysis and improved model physics
- Satellite level-1c radiances
  - Better RTTOV and improved use of radiances, especially IR and AMSU
  - Assimilation of rain affected radiances through 1D-Var
  - Variational bias correction
- Improved use of radiosondes
  - Bias correction and homogenization based on ERA-40
- Correction of SHIP/ SYNOP surface pressure biases
- Use of reprocessed
  - Meteosat winds
  - GPS-RO data CHAMP / UCAR 2001 →, GRACE and COSMIC
  - GOME O3 profiles 1995 →
- New set of Altimeter wave height data 1991→

# Tropical Ocean areas






ECMWF

## ERA-CLIM?

## ERA-Interim

- Research & Development as a collaborative effort 2009-2011 (under FP7 and with a aimed production starting in 2012)
- 1938 → 2015 and continue as CDAS
- Important components
  - Recovery, organization and homogenization of observations
  - Improved SST & ICE dataset
  - Variational analysis technique aimed for reanalysis
  - Comprehensive adaptive bias handling (including handling of model biases)
  - Research on coupled atmospheric-ocean-land reanalysis?
  - Better historical forcing data (aerosols, greenhouse gases,…)

# Other applications: GEMS
## Global and regional Earth-system Monitoring using Satellite and in-situ data

- An EC FP6 Integrated Project (2005-2009) that is developing:
  - **Global modelling and data assimilation for greenhouse gases, reactive gases and aerosols**
  - **An integrated production system for the above**
  - **Regional forecasting of reactive gases and aerosols**

- ECMWF is providing:
  - **Project coordination**
  - **Modelling and assimilation system for $CO_2$, $CH_4$, $O_3$, CO, $NO_2$, $SO_2$, HCHO and aerosols**
  - **Analyses for ENVISAT/EOS period (2003-2007)**
  - **Support for regional air quality forecasting**

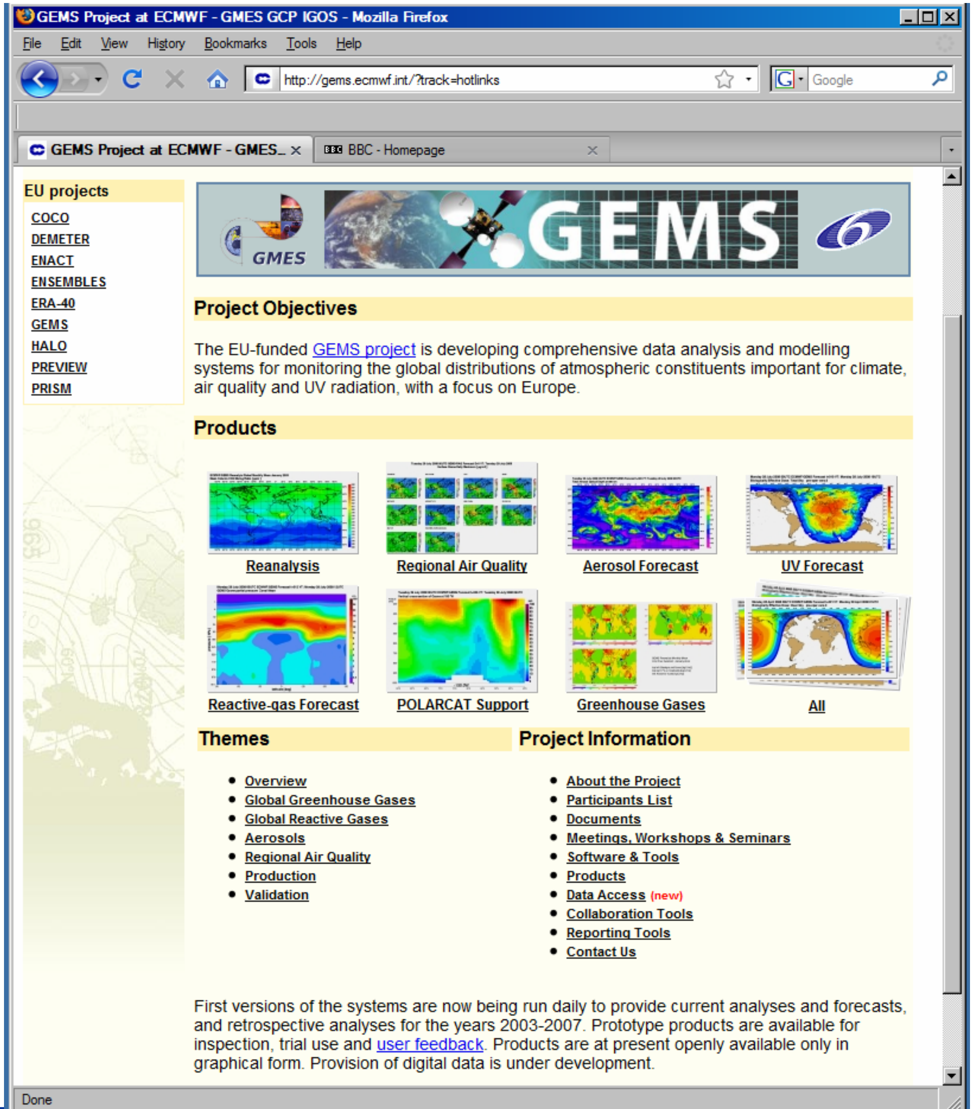
# Status of GEMS

- The system is running a near-real-time global system for reactive gases and aerosols

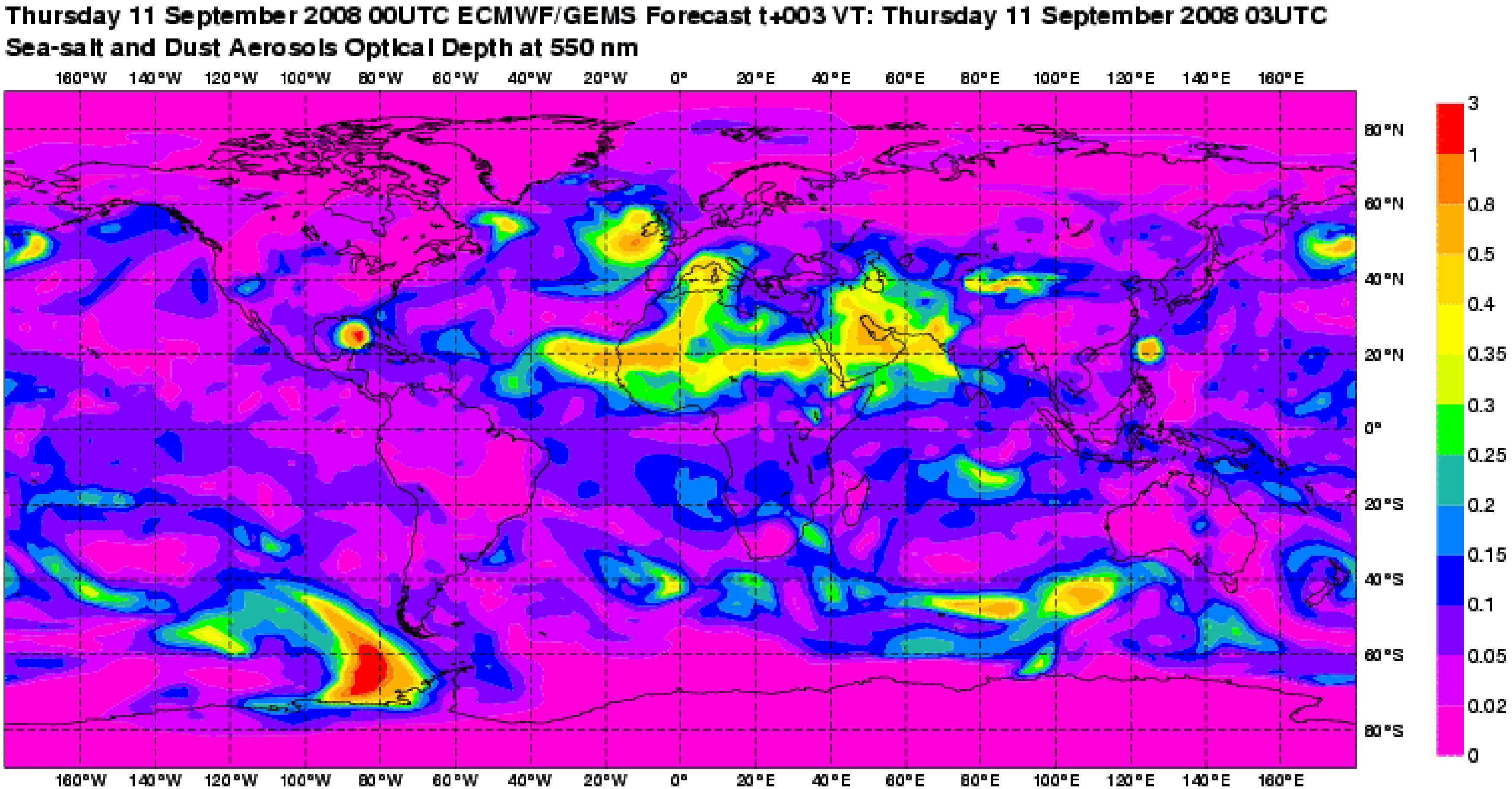
- A combined global reanalysis for 2003-2007 for greenhouse gases, reactive gases and aerosols has reached November 2005

- ECMWF is web-hosting coordinated regional air-quality forecasts from ten systems

- Plans are in place for the follow-on project MACC, with more formalised product delivery and user interaction



# Real-time forecasts (with assimilation of MODIS data)




http://gems.ecmwf.int

# Outline

- Components of the ECMWF forecasting system
- Performance of the NWP system
- Other applications
- **Future evolutions and challenges**



ECMWF

# Future evolutions and challenges

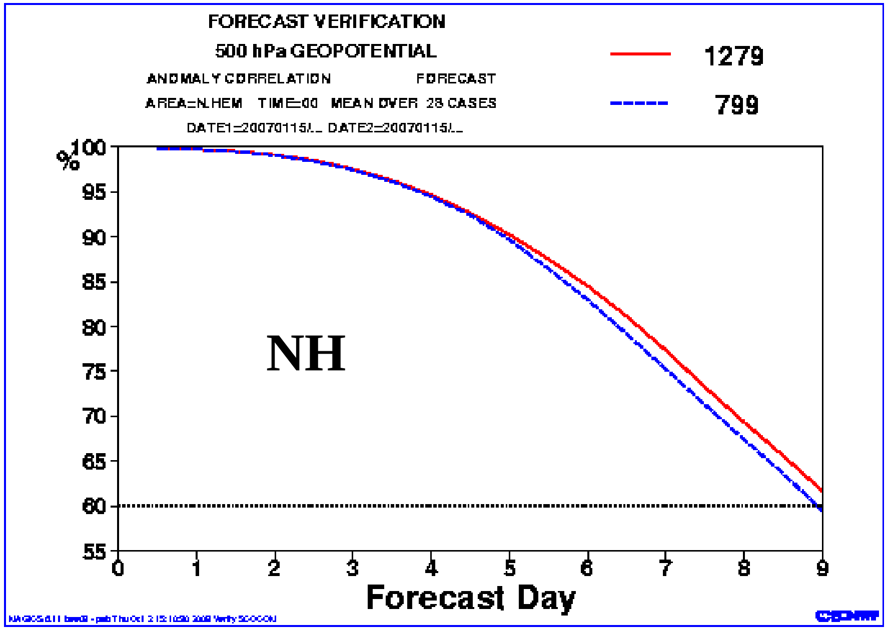
- Model resolution increase
- Increased use of satellite data
- Long window (weak-constraint) 4D-Var
- Ensemble data assimilation
- Modularisation of the IFS
- Non hydrostatic modelling, better physics, etc…

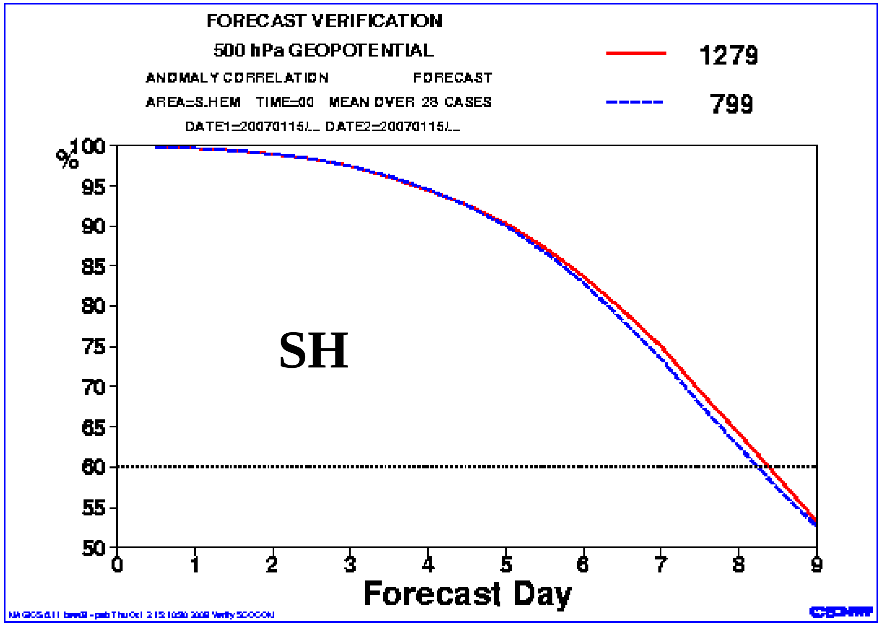
# Model resolution increase

- **The model spectral resolution will be increased from T799 to T1279 in 2009**
    - **The resolution increase of the assimilation and the EPS will be commensurate (T399 and T639 respectively)**
- **The model vertical resolution will be increased from 91 to about 150 levels in 2010**
- **By 2015, the deterministic model resolution could be T2047 (10km)**

# Increased realism via higher resolution (horizontal and vertical):

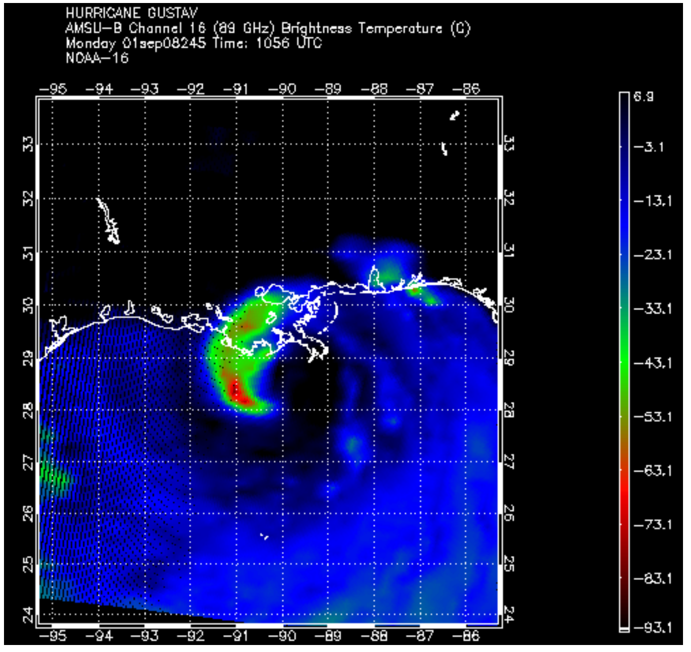
- T1279 (EPS at T639) planned for later next year
- T2047 run as a glimpse of the future



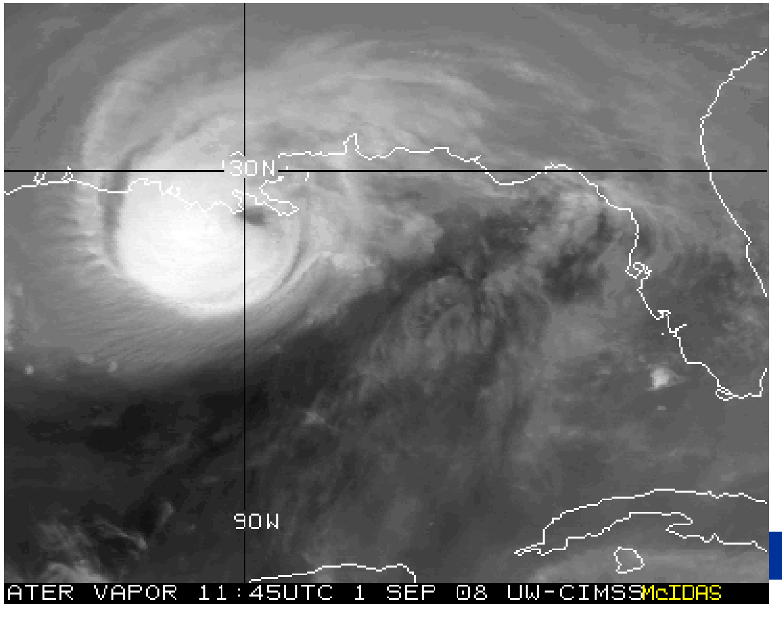




**28 cases (forecast runs only, no assimilation)**

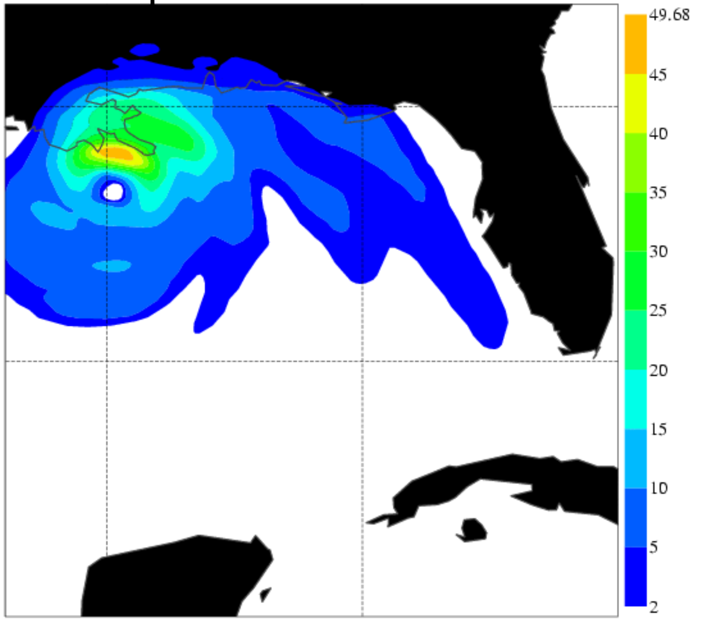
# Hurricane Gustav AMSU-B and 33-36h rainfall



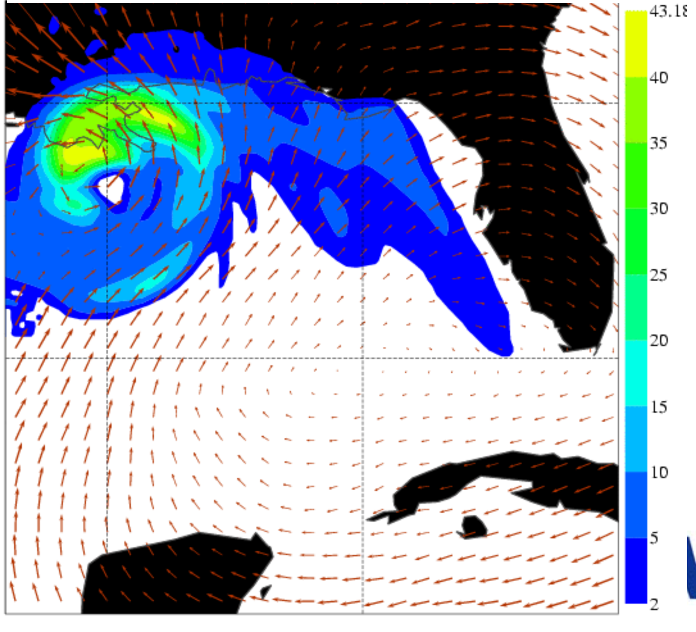

from CIMSS Wisconsin




T799 oper 2008083100 +36h



T1279 with 200hPa wind

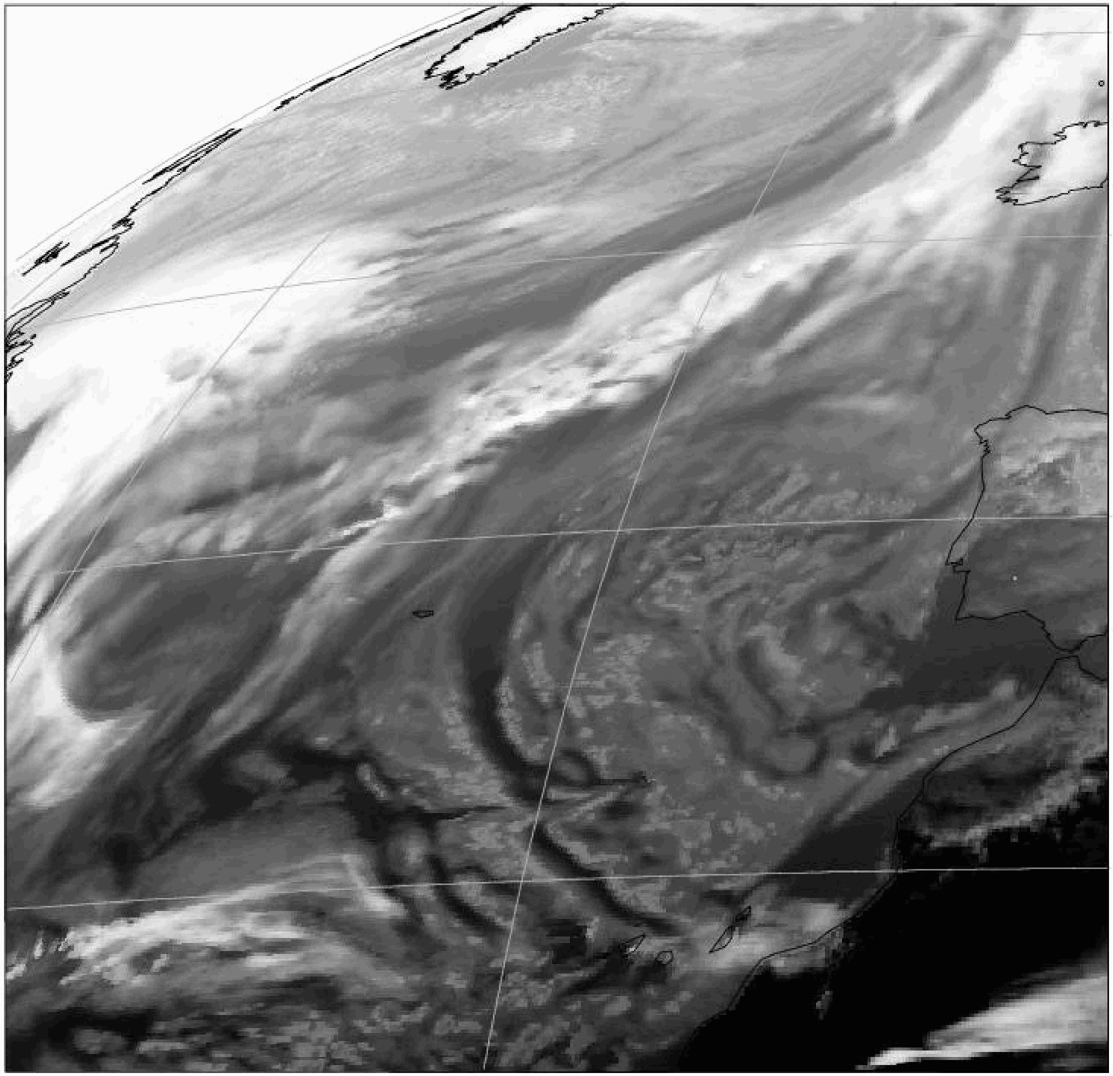


# Simulated infra-red cloud images at T2047 (10kms)

Simulated from a

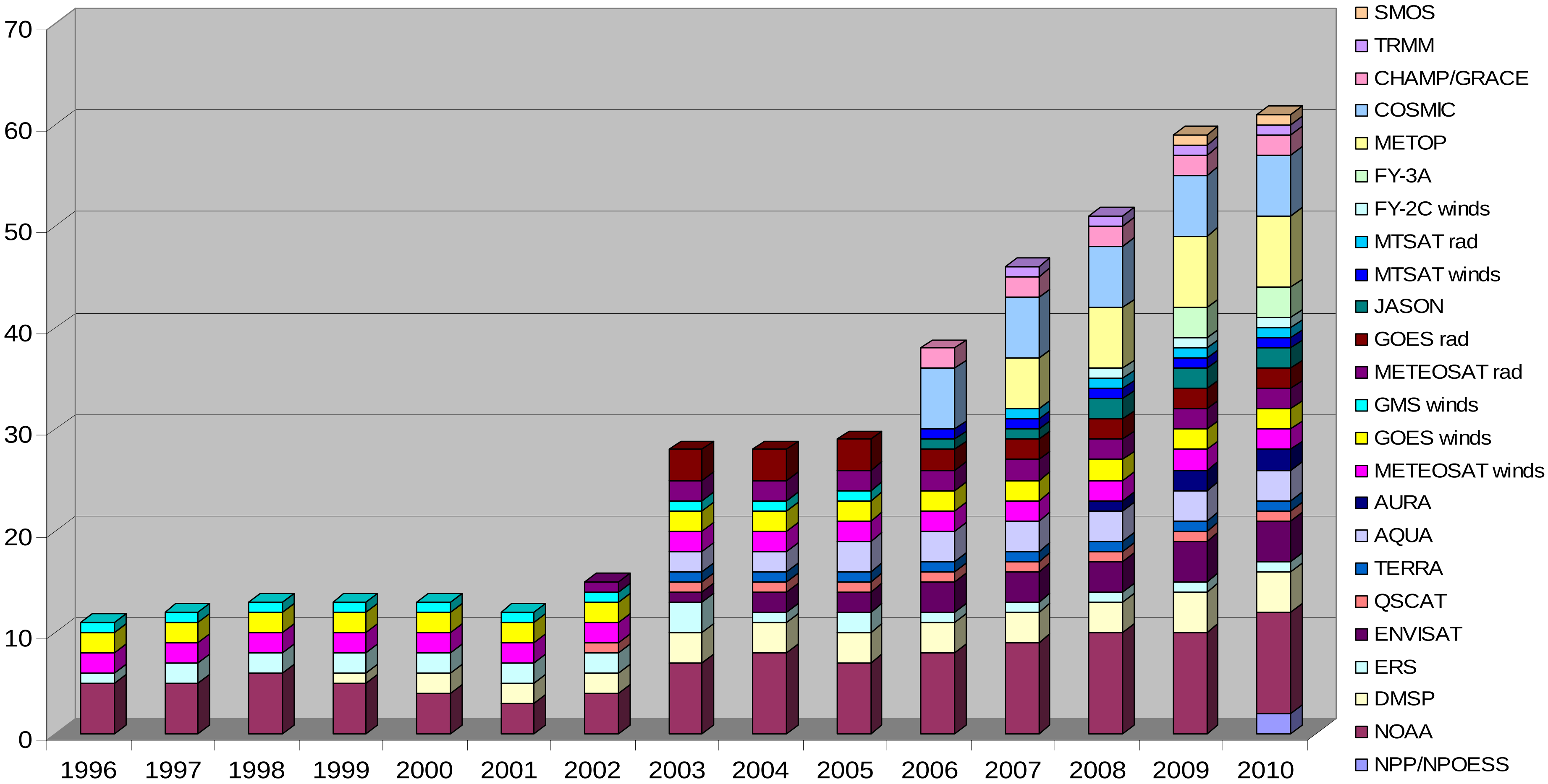
T2047 (~10km) forecast

(15min output)

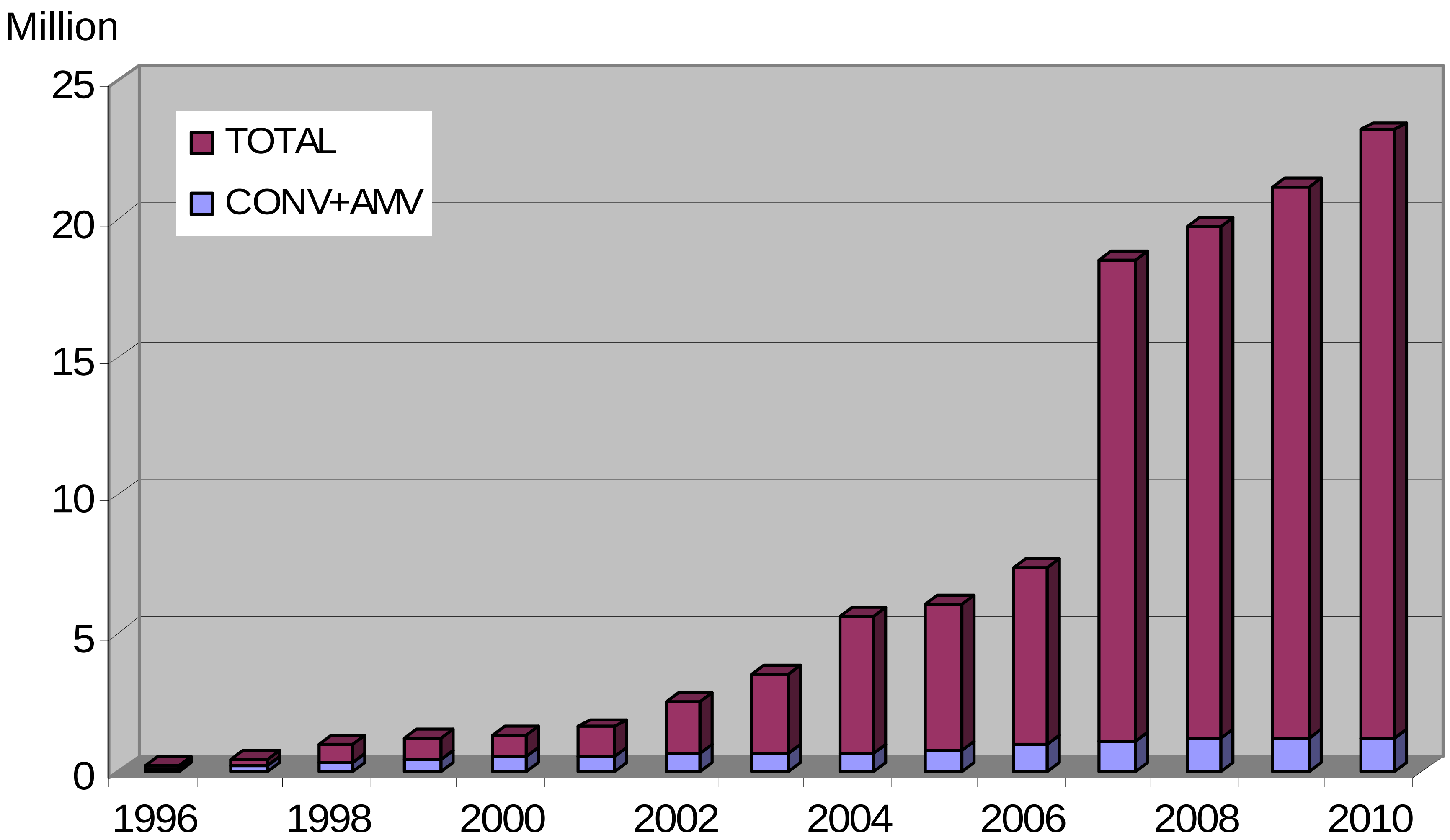


Met-8 IR

# Data sources assimilated at ECMWF



# Data volume assimilated at ECMWF



# Long window 4D-Var (Mike Fisher, Yannick Trémolet)

- **Extending the 4D-Var assimilation window is appealing because:**
  - **True equivalence with the Kalman filter at the end of the window**
  - **Use of all relevant observations to optimally estimate the atmospheric state**
- **Extending the 4D-Var window requires accounting for model error (Weak-constraint 4D-Var)**
- **A formulation, with a 4D-state control variable, has been developed**
  - **Which provides potential for extra-parallelism**

# Weak constraint 4D-Var

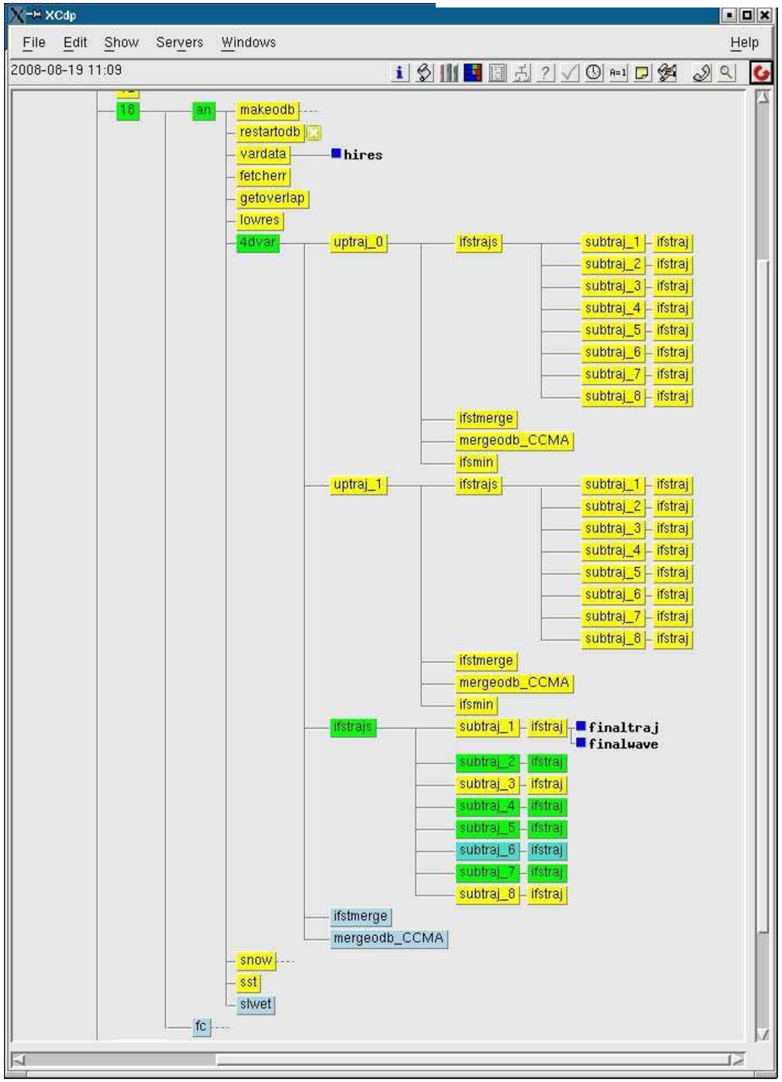
$$J(x) = \frac{1}{2}(x_0 - x_b)^T B^{-1}(x_0 - x_b) + \frac{1}{2}\sum_{i=0}^{n}[\mathcal{H}(x_i) - y_i]^T R_i^{-1}[\mathcal{H}(x_i) - y_i]$$
$$+ \frac{1}{2}\sum_{i=1}^{n}[x_i - \mathcal{M}_i(x_{i-1})]^T Q_i^{-1}[x_i - \mathcal{M}_i(x_{i-1})]$$

- Use $\{x_i\}_{i=0,\ldots,n}$ as the control variable.
- Incremental cost function:

$$J(\delta x) = \frac{1}{2}(\delta x_0 - b)^T B^{-1}(\delta x_0 - b) + \frac{1}{2}\sum_{i=0}^{n}(H\delta x_i - d_i)^T R_i^{-1}(H\delta x_i - d_i)$$
$$+ \frac{1}{2}\sum_{i=1}^{n}(q_i + M_{i-1}\delta x_{i-1} - \delta x_i)^T Q_i^{-1}(q_i + M_{i-1}\delta x_{i-1} - \delta x_i)$$

where $b = x^g - x_b$, $d_i = \mathcal{H}(x_i^g) - y_i$ and $q_i = \mathcal{M}_{i-1}(x_{i-1}^g) - x_i^g$.

- The model does not appear in the $J_o$ term,
- In practice $x_i$ is defined at regular intervals within the assimilation window.




- The outer loop can run in parallel for each sub window

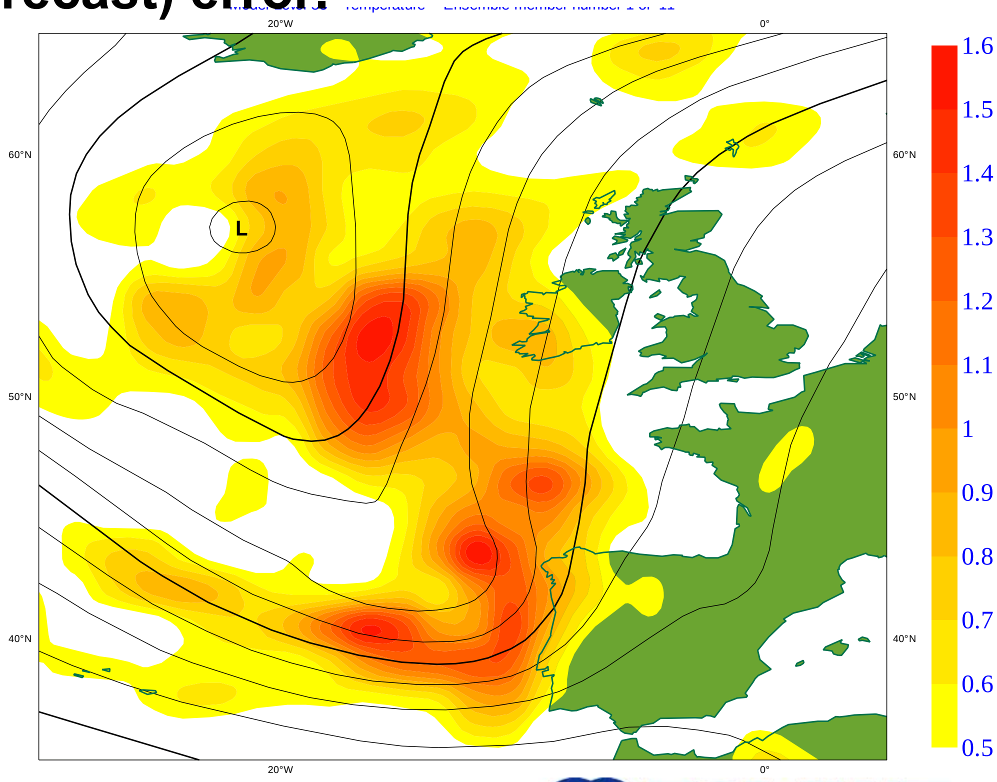
9

# Ensemble data assimilation

- **Run an ensemble (e.g. 10 + 1 control) of analyses with random observation, SST field and model perturbations, and form differences between pairs of analyses (and short-range forecast) fields.**
- **These differences will have the statistical characteristics of analysis (and short-range forecast) error.**

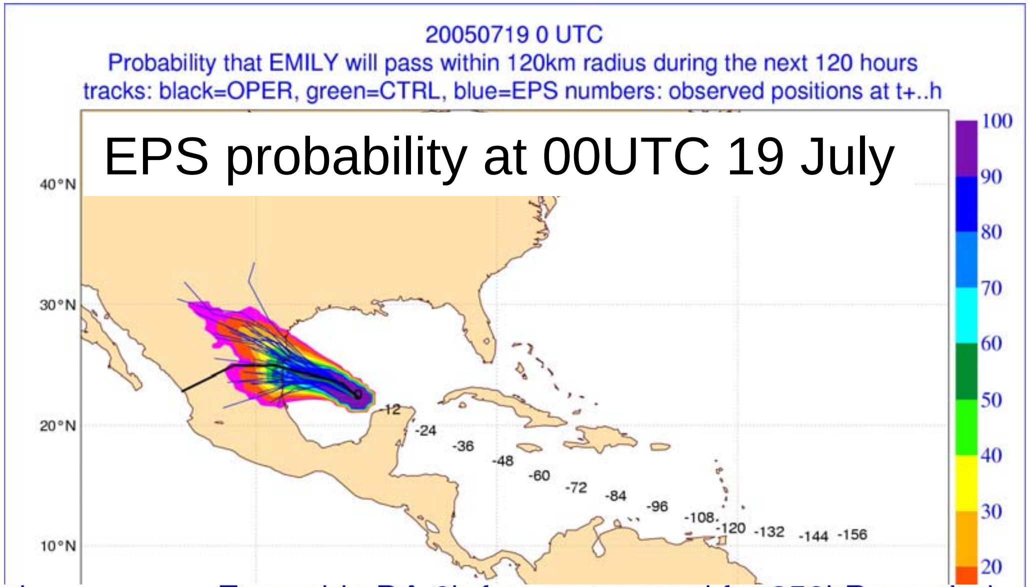
**To be used in specification of background errors = “errors of the day”. To indicate where good data should be trusted in the analysis (yellow shading).**

**This is also used in the initialization of the EPS**



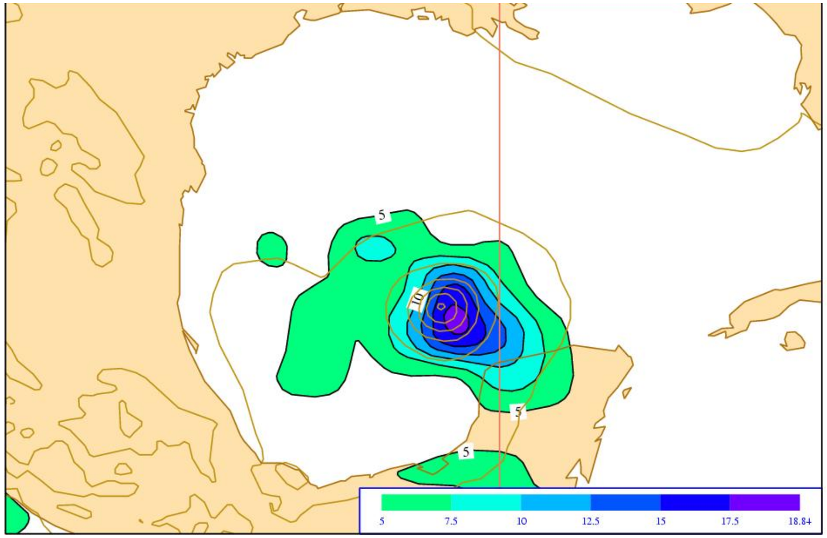
# Hurricane Emily 19-20 July 2005

## Ensemble Data Assimilation spread for zonal wind at 850hPa



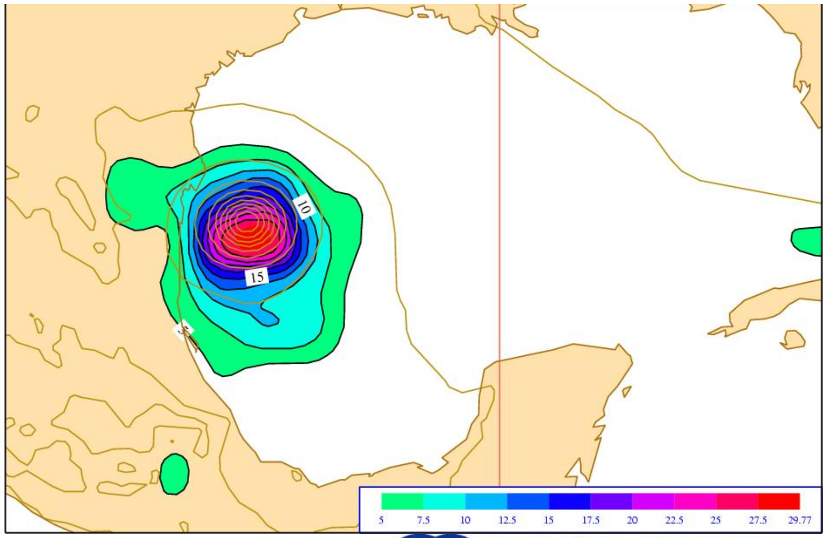

Ensemble DA 6h forecast spread for 850hPa u-wind
T799 10 member ensemble valid 00UTC 19 July 2005

Max. stdev of EnDA spread 19m/s

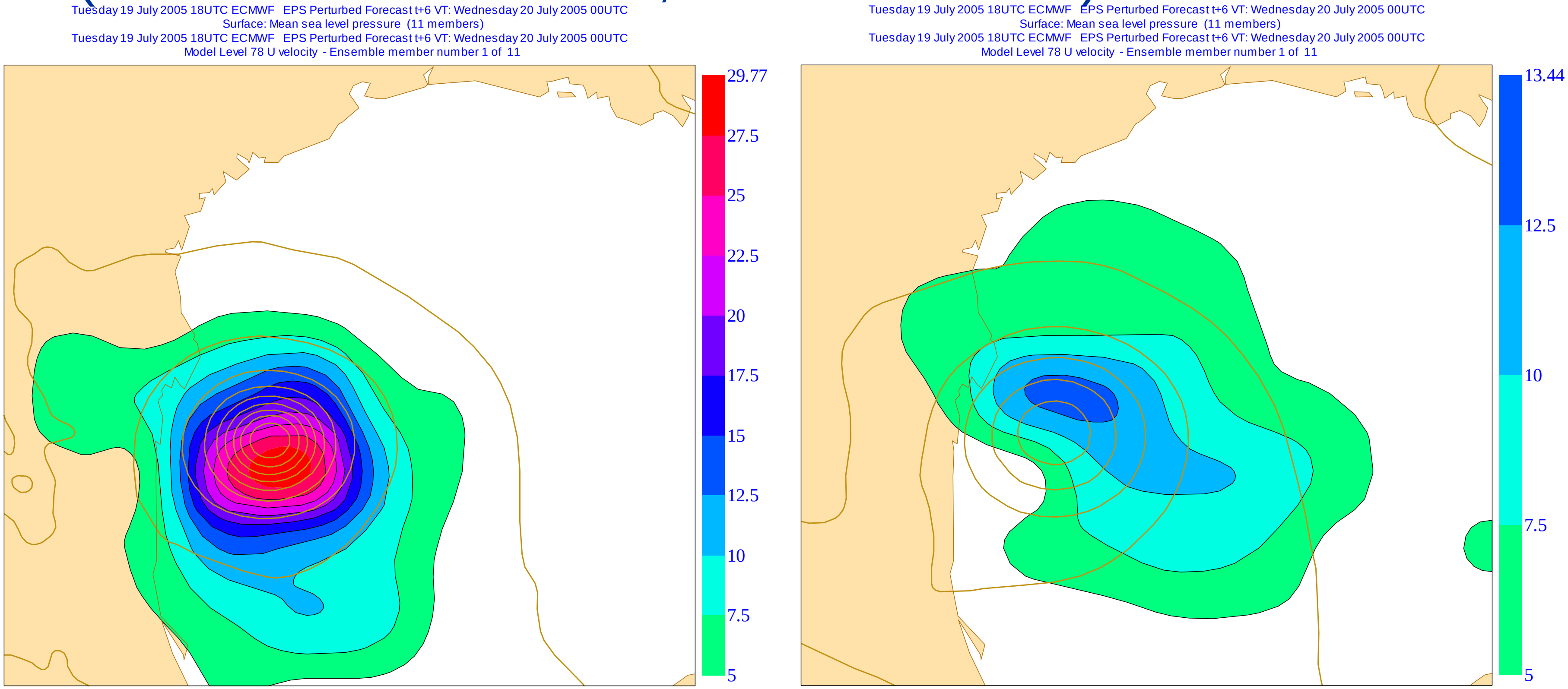



Ensemble DA 6h forecast spread for 850hPa u-wind
T799 10 member ensemble valid 00UTC 20 July 2005

Max. stdev of EnDA spread 30m/s



# Tropical cyclone Emily 00UTC 20/7/2005 (resolution matters, sometimes...)




Standard deviation of zonal wind near 850hPa calculated from two 10-member EnDA ensembles. The contours represent the mean sea level pressure field (5hPa interval).

The right panel is for an ensemble with $T_L399$ outer loop and a single $T_L159$ inner loop.

The left panel is from an ensemble with $T_L799$ outer loop and two minimisations at $T_L95$ and $T_L255$, respectively. Maximum spread values are 13.44ms$^{-1}$ for the lower-resolution ensemble, and 29.77ms$^{-1}$ for the $T_L799$ ensemble..

# Modularisation of the IFS (1) (Yannick Trémolet, Mike Fisher)

- **The IFS code is more than 20 years old. Over this period it has reached a high level of complexity, which is becoming a barrier to future scientific developments, and makes the ramp-up phase for new scientists/visitors unacceptably long.**
- **This makes the case for rethinking the design of the IFS (in particular the data assimilation)**
  - **All data assimilation schemes manipulate a limited number of entities (H, M, R, B, x, y, …)**
  - **To adapt to future scientific developments, these entities should easily be accessible**
- **Information-hiding and abstraction are important: only those parts of the code that need to know about the detailed structure of some entity (e.g. model fields) should be exposed to it.**
- **Object-oriented languages (e.g. Fortran 2003) contain the features required to fully express these ideas.**

# Modularisation of the IFS (2)

- The main idea of abstraction is to separate the algorithm from the detailed implementation of the objects it deals with.
- This will be tried for the entire incremental 4D-Var algorithm.
- The result would be a 4D-Var framework into which we could plug a variety of models.
- Question to the audience:
  - Yannick (yannick.tremolet@ecmwf.int tel: 2110)
  - Mike (mike.fisher@ecmwf.int tel: 2622)
  - Would like to know more about the future F2003 compilers on forthcoming HPCs, and about any OO technologies at large on these future machines