→ ADVANCED ATMOSPHERIC TRAINING COURSE 2014

# [Global] Data Assimilation

Vincent-Henri Peuch

ECMWF
EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

**Acknowledgement:** Angela Benedetti, Hans Hersbach, Antje Inness, Sébastien Massart (ECMWF)

« La composition chimique de l'air, la nature des poussières qu'il charrie nous offrent aussi leurs renseignements particuliers. On ne saurait trop s'entourer de lumière dans la solution d'un problème aussi difficile et aussi grave que la prévision du temps [...] »

*The chemical composition of air and the nature of the dusts it carries provide us also with their particular information. One has never too many pieces of information for solving such a difficult and serious problem as weather forecasting*

**H. Marié-Davy**

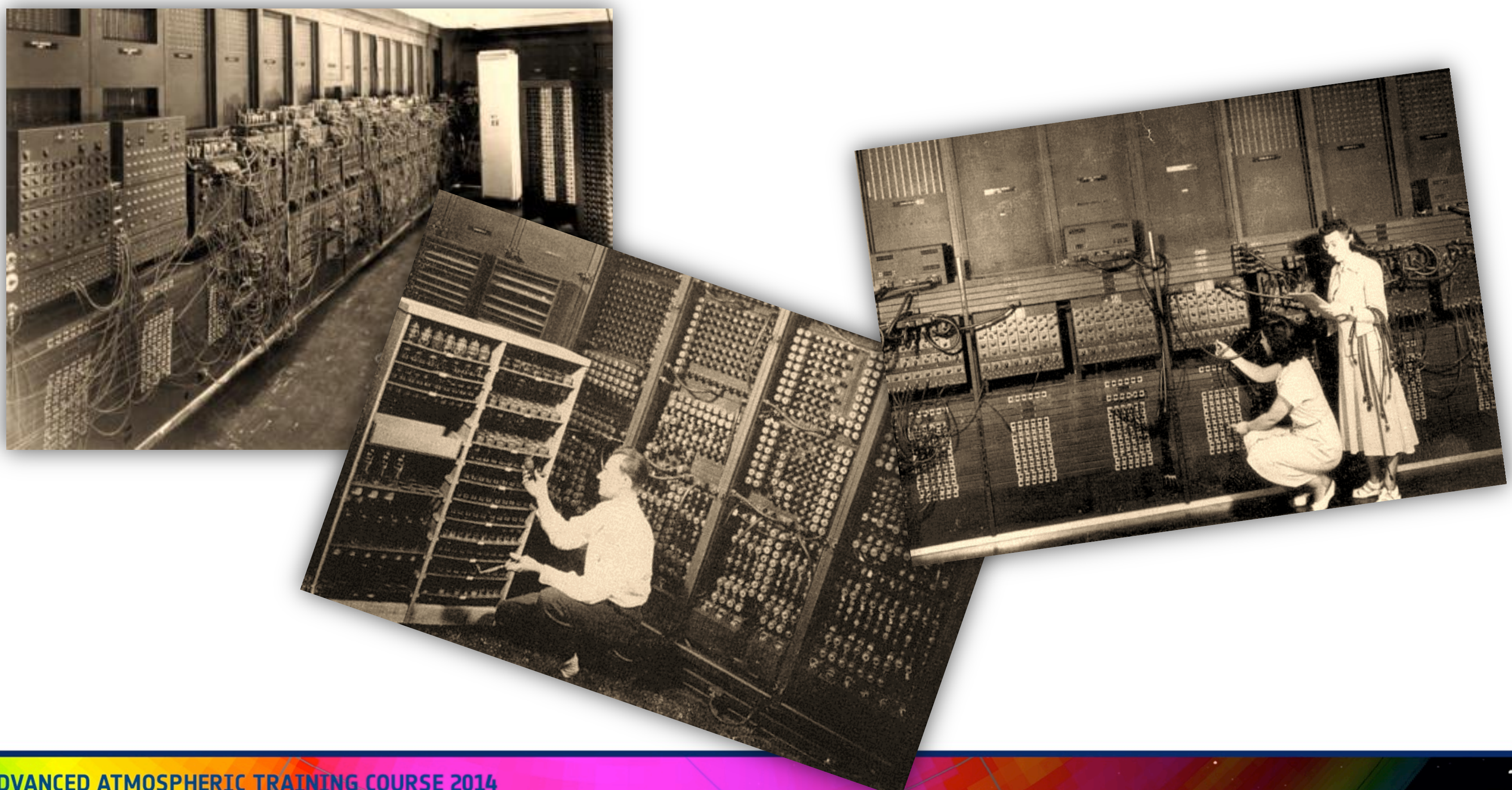
"Les mouvements de l'atmosphère
et les variations du temps"
Masson, Paris, **1877**

# L.F. Richardson's Forecast Factory

Weather Prediction by Numerical Process,
Cambridge Univ. Press, 1922, ch. 11/2, 219-220.

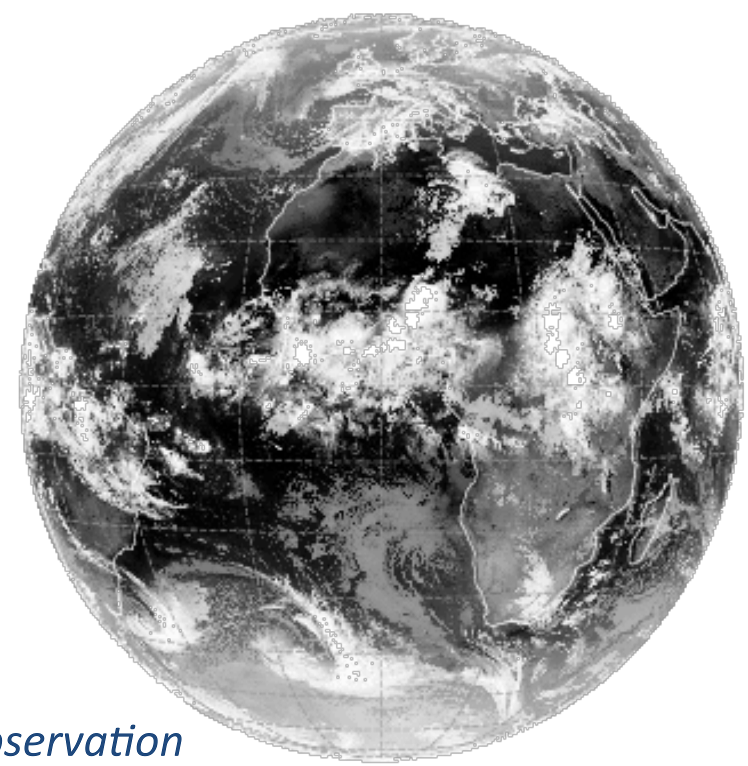
"Computers" solving an initial value problem using the latest observations made globally and the laws of physics to forecast the weather

# It took three decades for Richardson's vision to start being explored really using (machine) "computers".

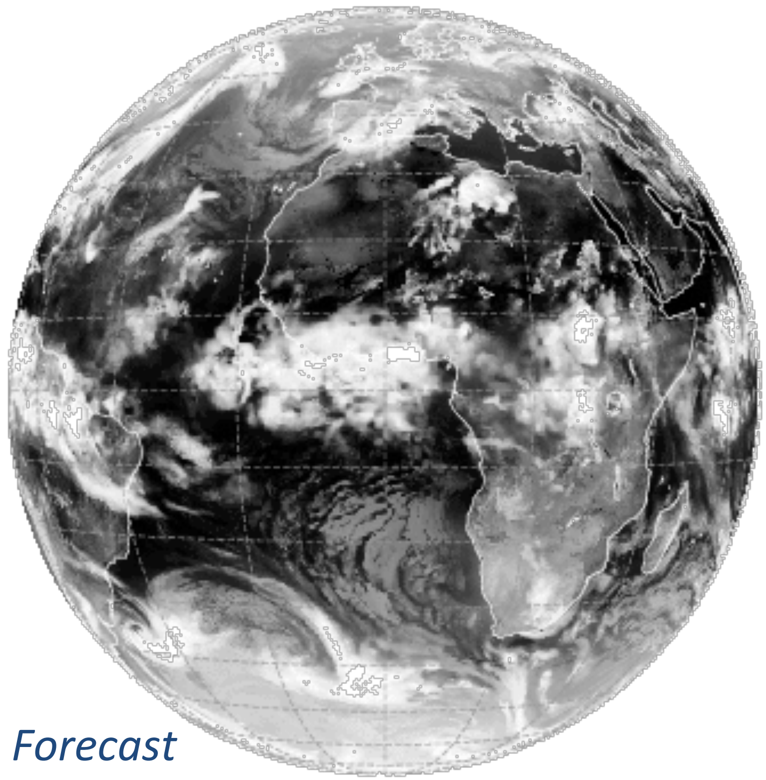


# And another three decade to establish reliable Numerical Weather Predictions

# And yet another three decades to reach today's increasingly accurate forecasts, including in the medium-range

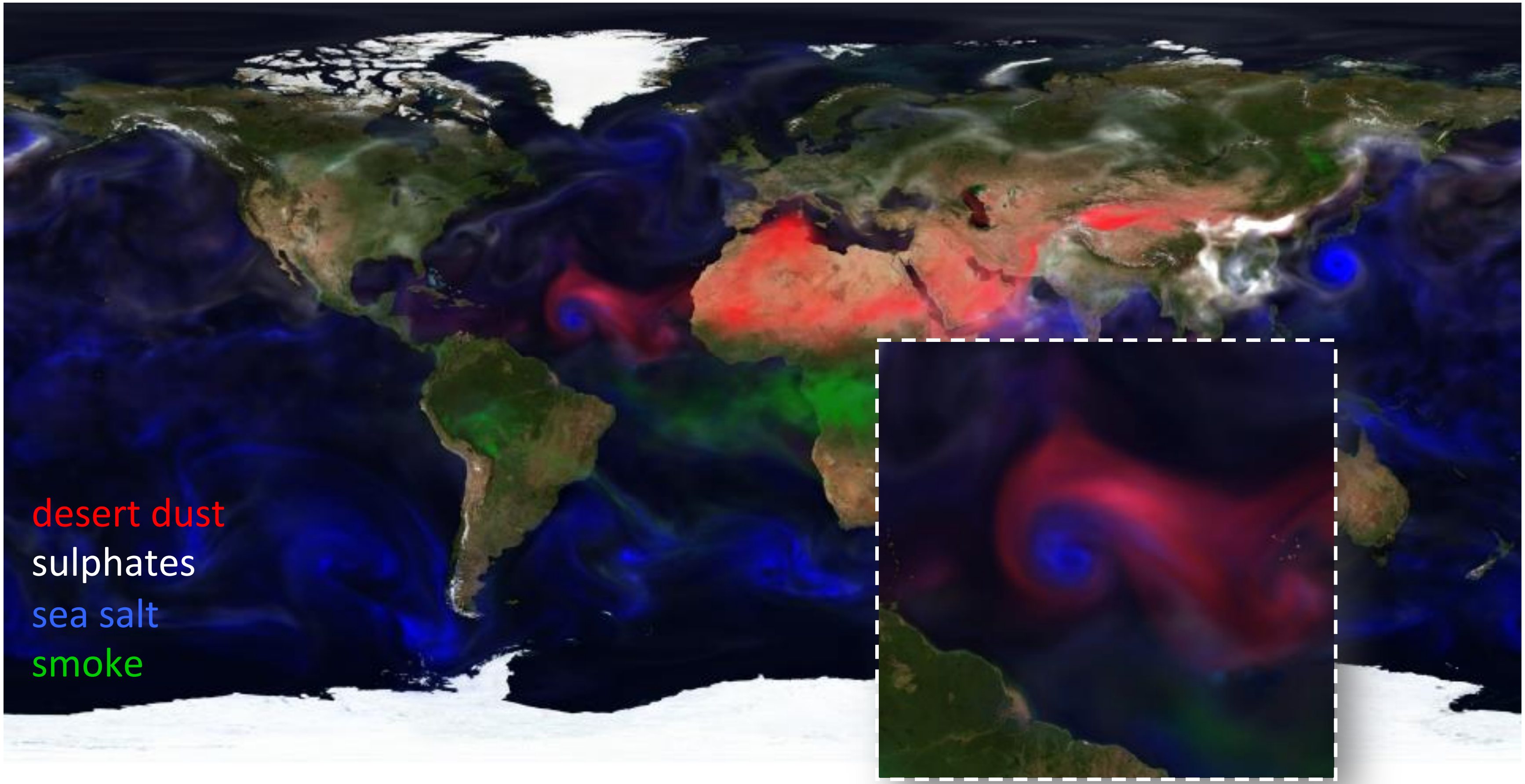



*Observation*

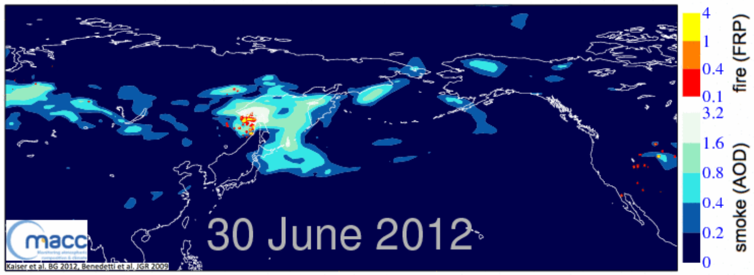


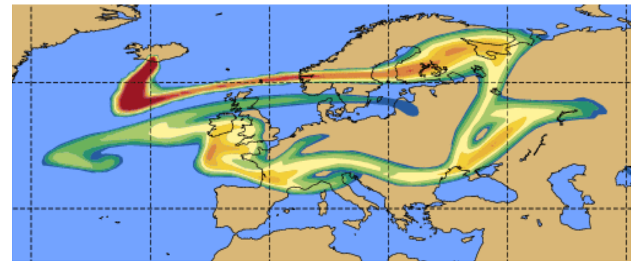

*Forecast*

# Today: Monitoring atmospheric aerosol in an integrated weather-composition global model



# Tomorrow : "super-seamless" environmental prediction



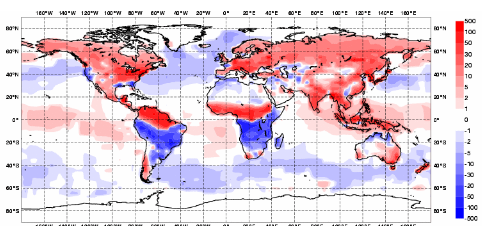



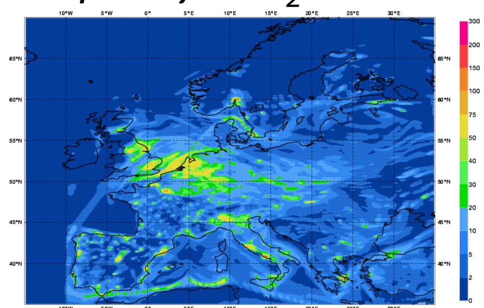
*Volcanic ash*

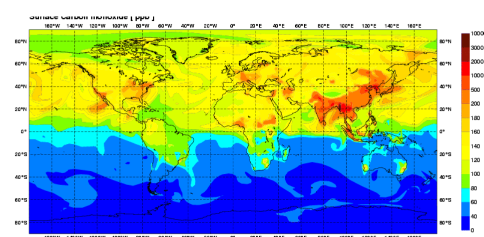
*Fire and smoke*

*Air quality: $NO_2$*

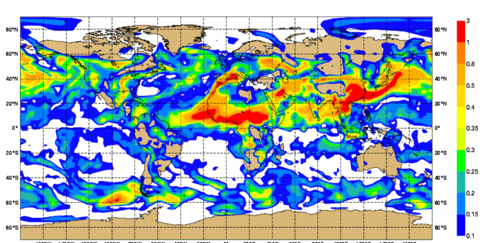
*Global pollution: CO*





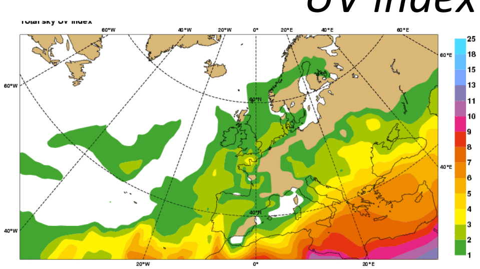


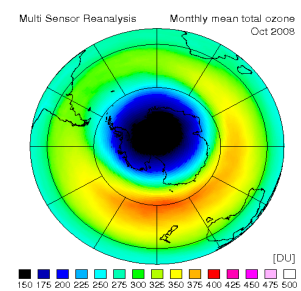
*Flux inversions: $CO_2$*



*Aerosol*

*UV index*





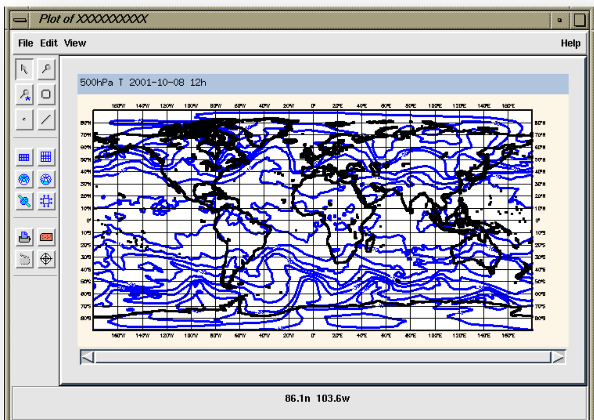

*Ozone layer*

# DATA ASSIMILATION

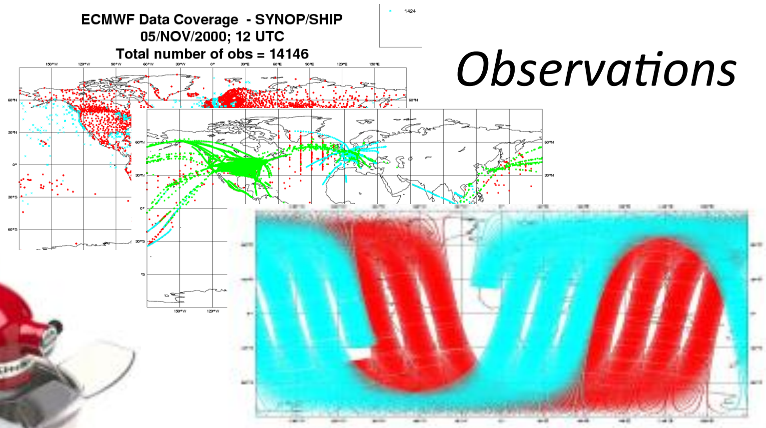
# What?



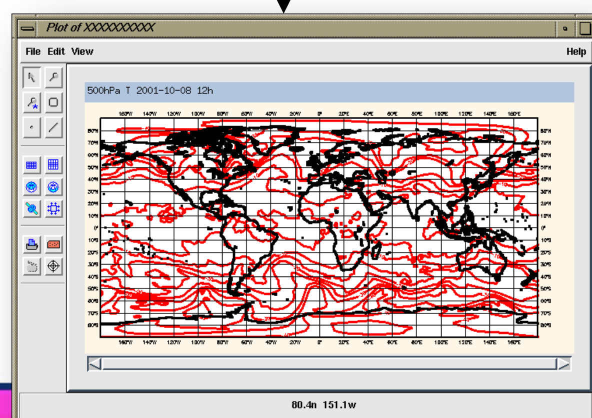
## Mixing information from different sources




*Background information*

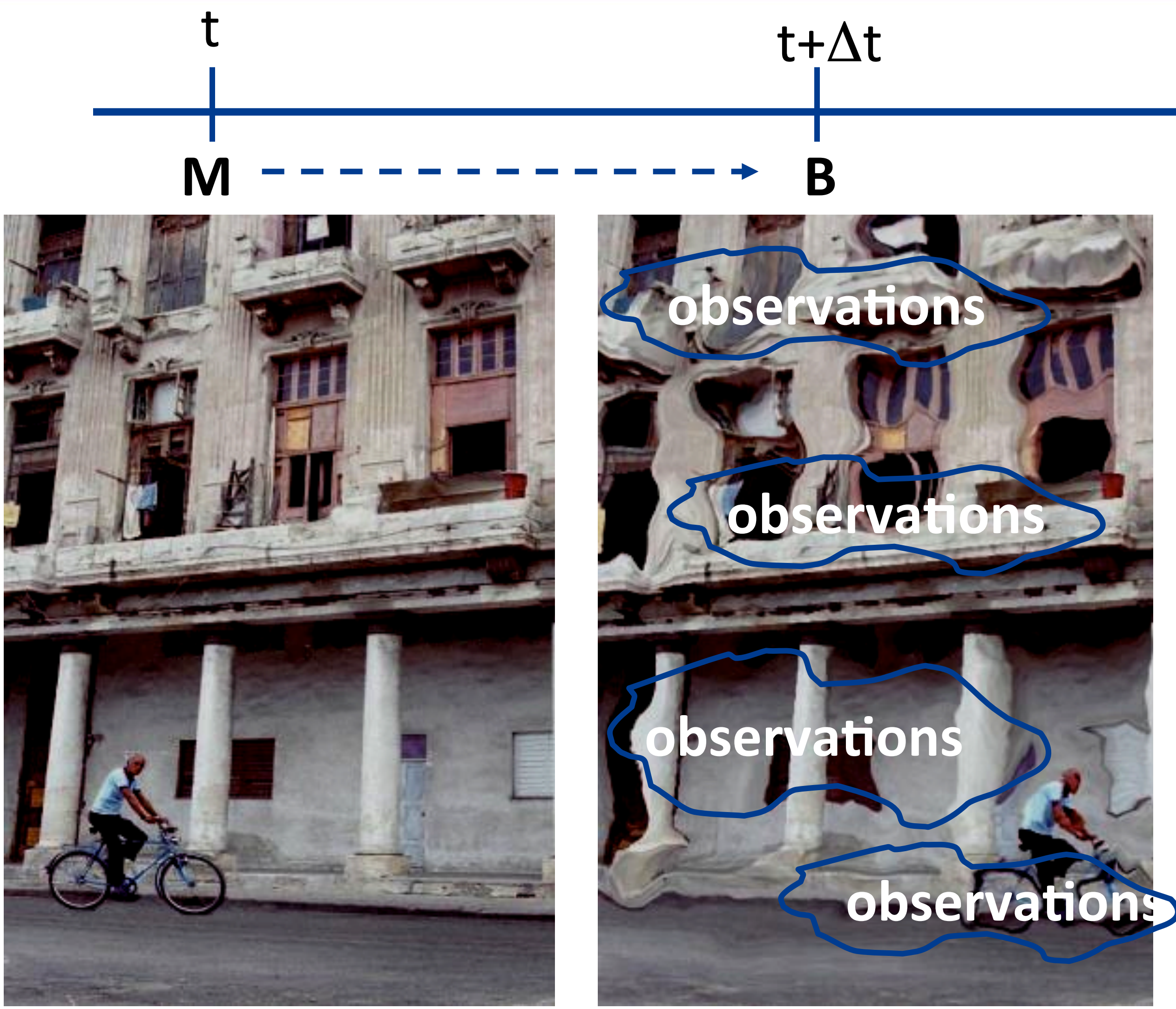



*Observations*



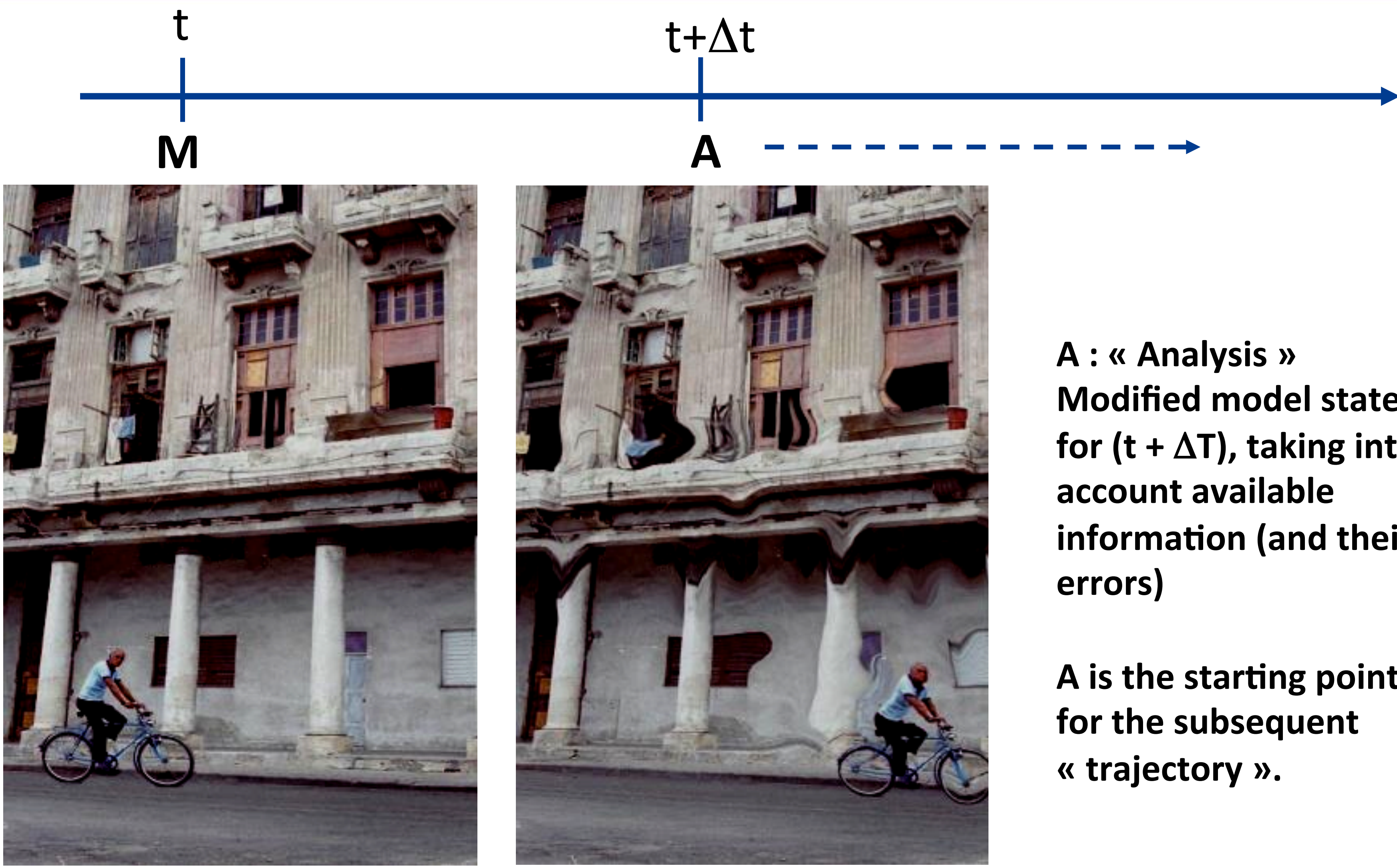

*Analysis*

*Initial conditions for next forecast*



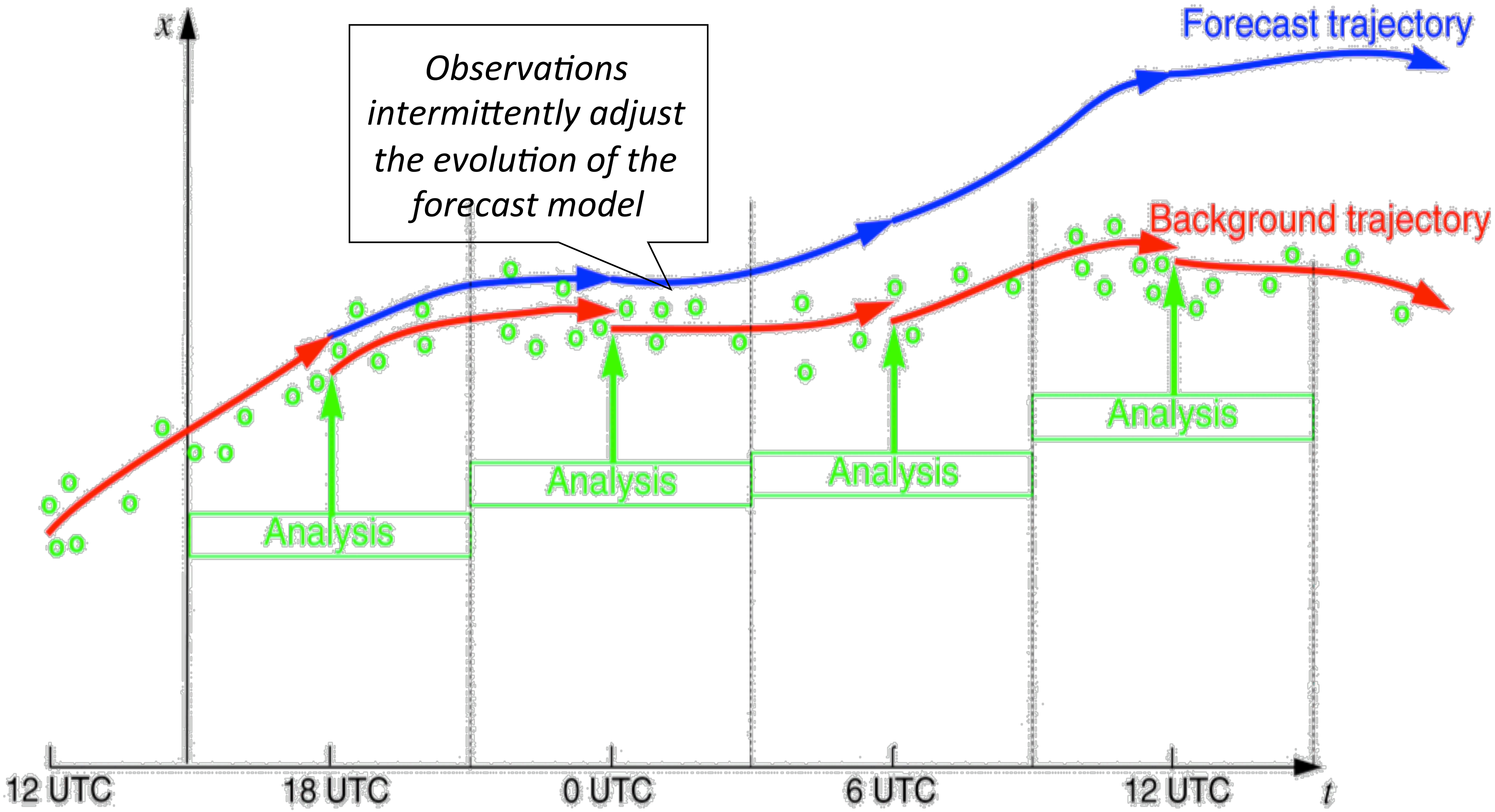

**M :**
**known model state (t)**

**B :**
**« Background »**
**(or « First Guess »)**



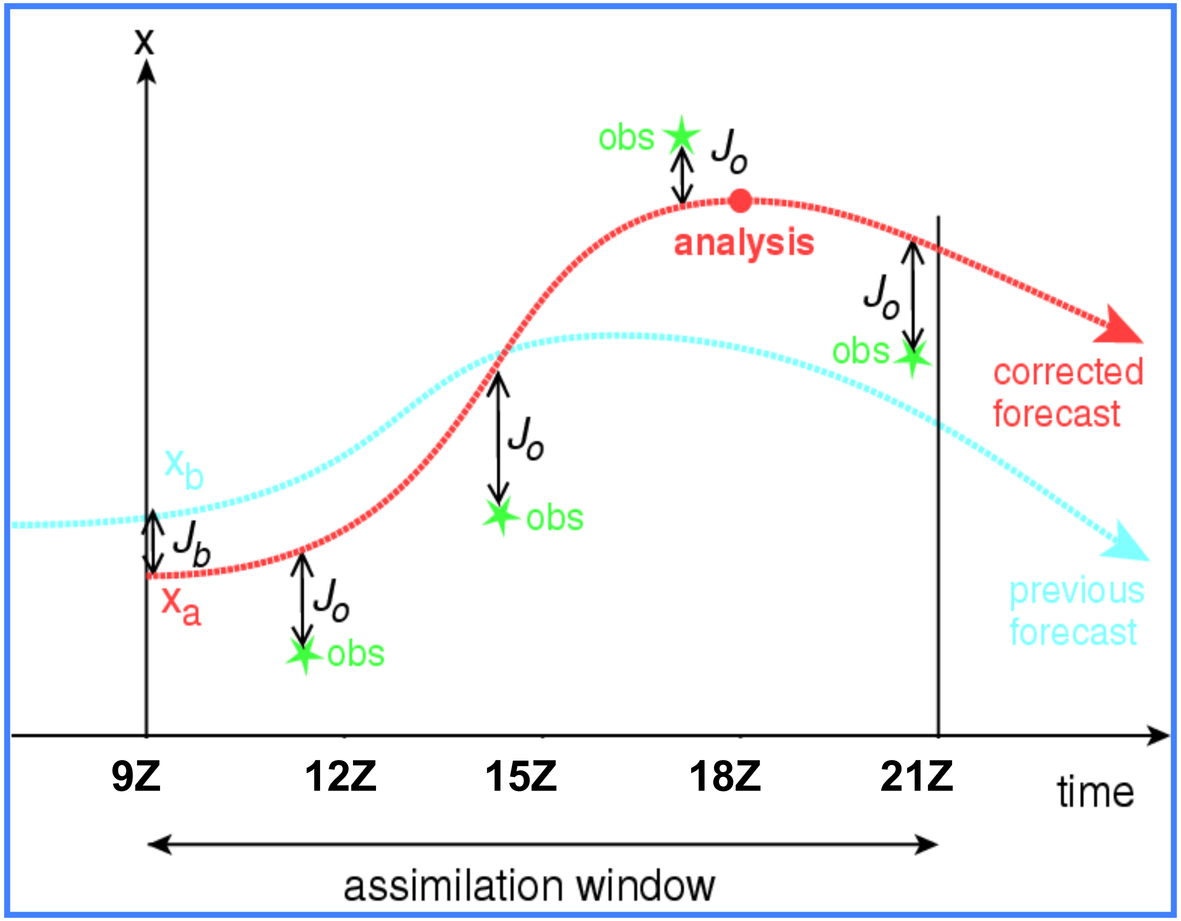

**A : « Analysis »
Modified model state for (t + ΔT), taking into account available information (and their errors)**

**A is the starting point for the subsequent « trajectory ».**


Observations intermittently adjust the evolution of the forecast model
Forecast trajectory
Background trajectory
Analysis
Analysis
Analysis
Analysis
x
t
12 UTC
18 UTC
0 UTC
6 UTC
12 UTC

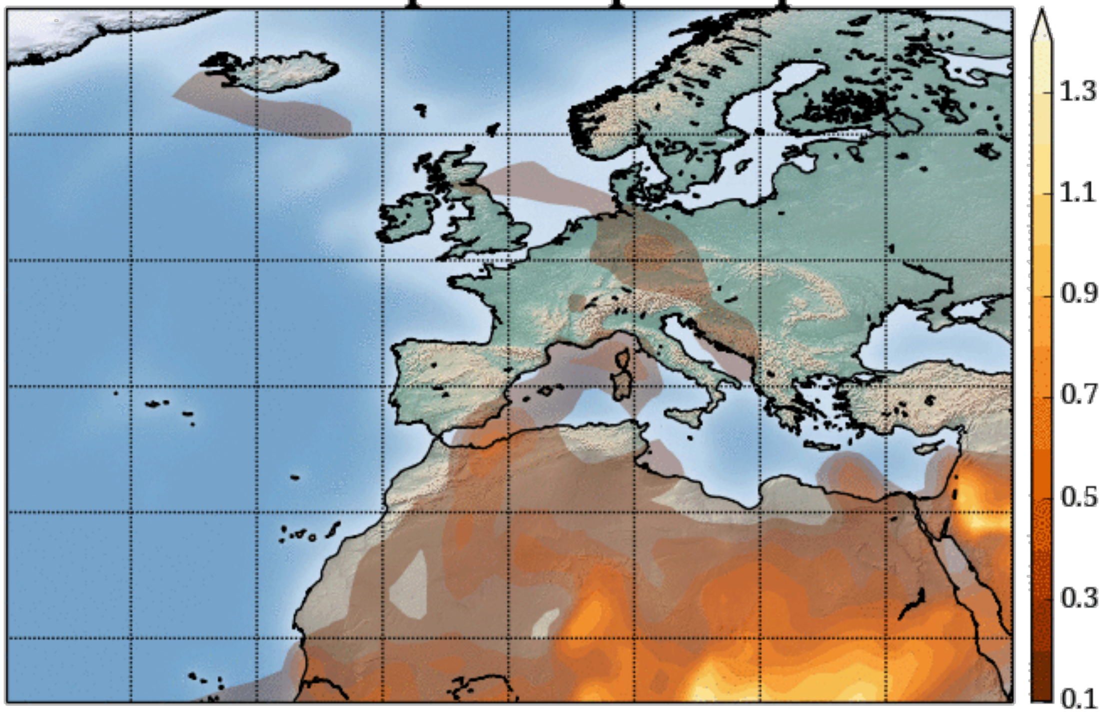
# A few Characteristics about the ECMWF NWP 4D-Var

- "Observation – model" values are computed at the observation time at high resolution: 16 km
- 4D-Var finds the 12-hour forecast that take account of the observations in a dynamically consistent way.
- This is based on a tangent linear and adjoint forecast models, used in the iterative minimization process.
- 80,000,000 model variables (surface pressure, temperature, wind, specific humidity and ozone) are adjusted



# Why?

MACC-II dust aerosol optical depth 2 April 2014 01 UTC

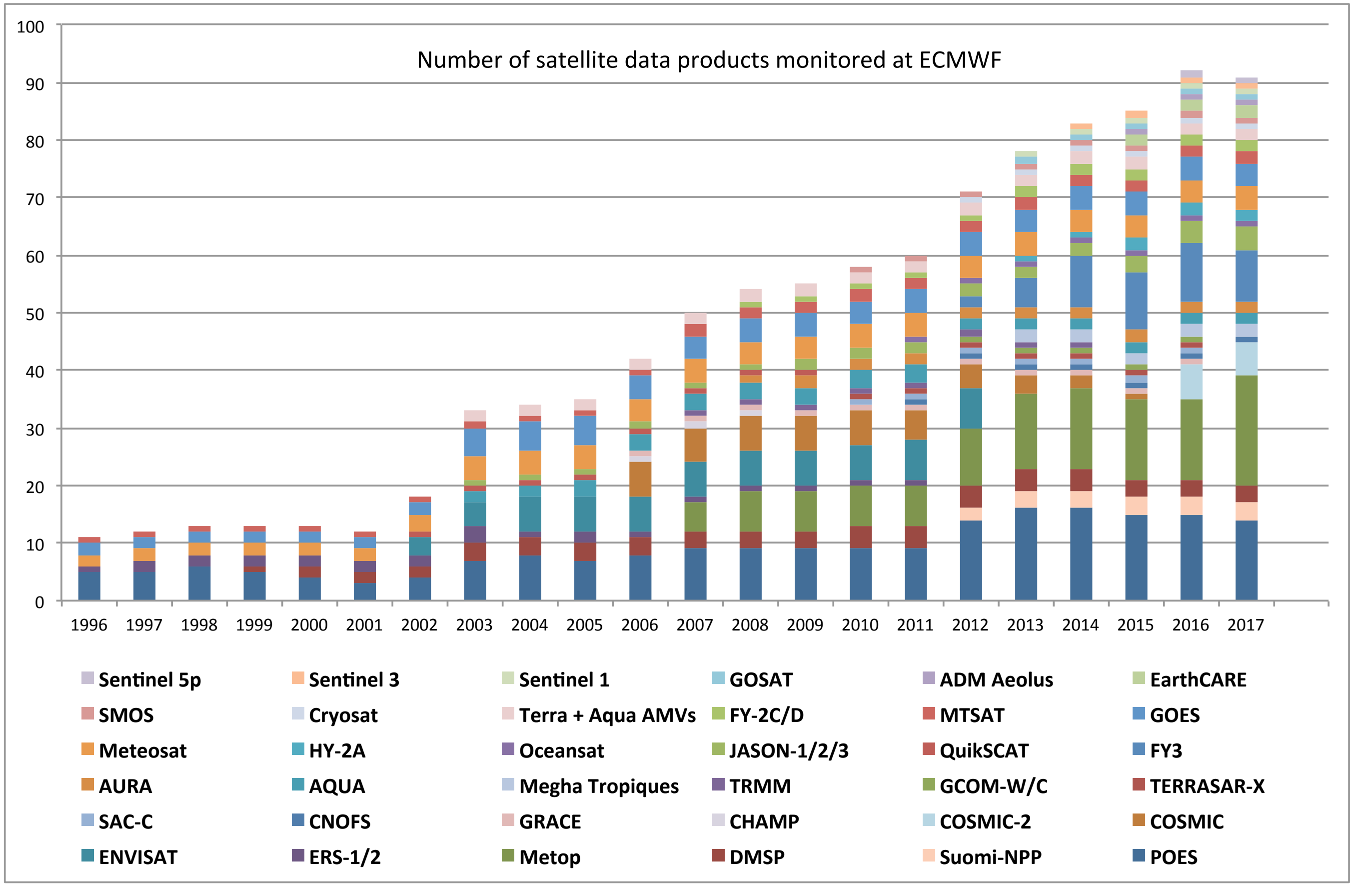



- To base forecasts on “observed” state
- To provide maps with no gaps
- To combine information from different sources, with a result consistent with specified errors
- To use observations at their specific time and place (advanced DA techniques)
- ...

A good read:

*[W. Lahoz et al., Data assimilation of stratospheric constituents: a review, ACP, 7, 5745–5773, 2007.]*

# Some challenges: quantity

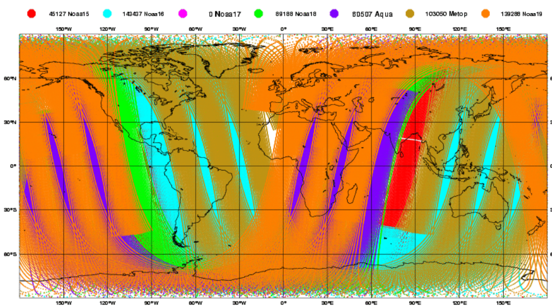



70+ instruments provide ~15mil observations in each 12-hour window

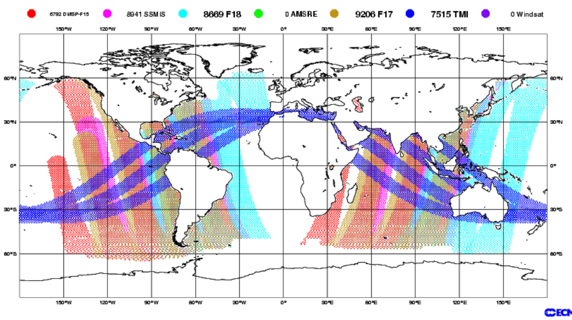
# Some challenges: variety

*Example of 6-hourly satellite data coverage*

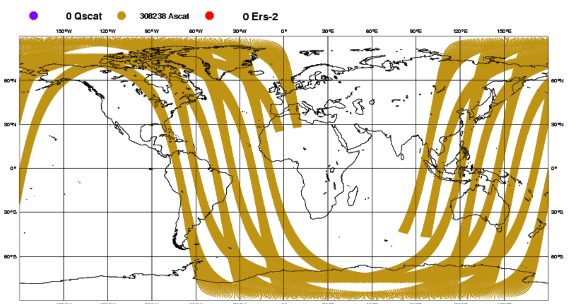
LEO Sounders



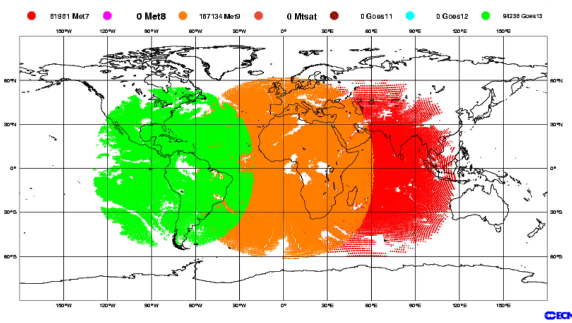
LEO Imagers



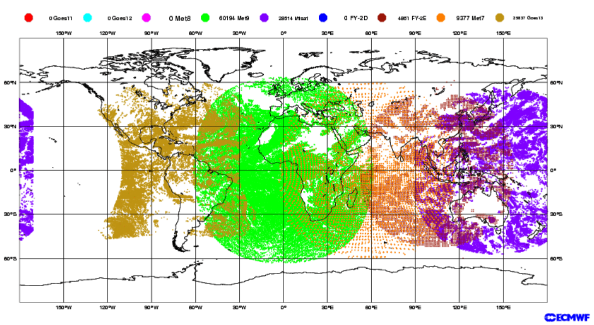
Scatterometers



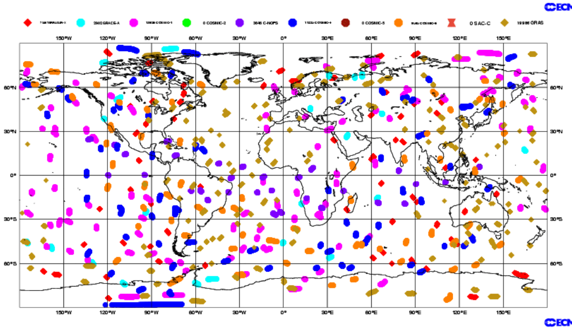

GEO imagers

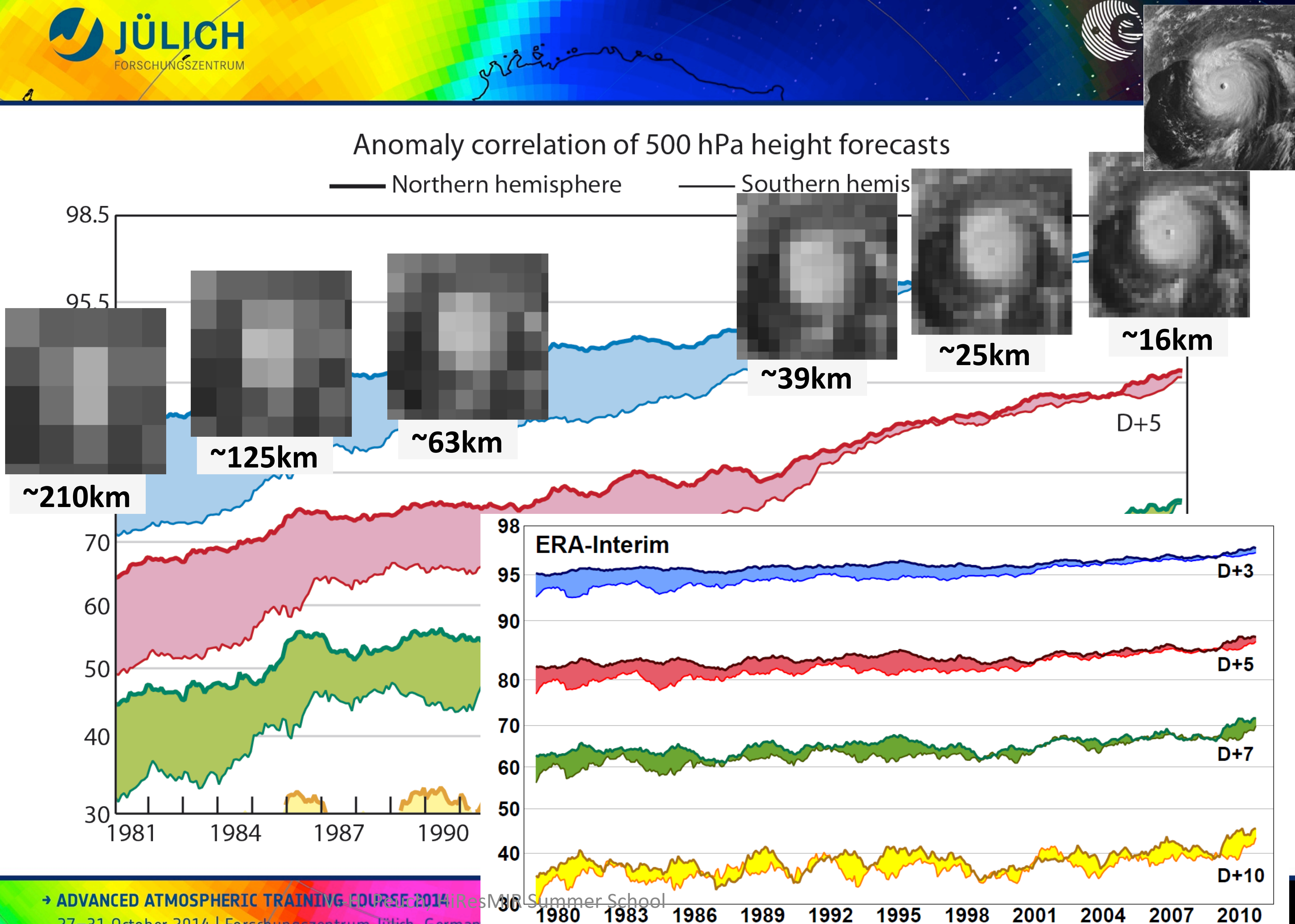


Satellite Winds (AMVs)

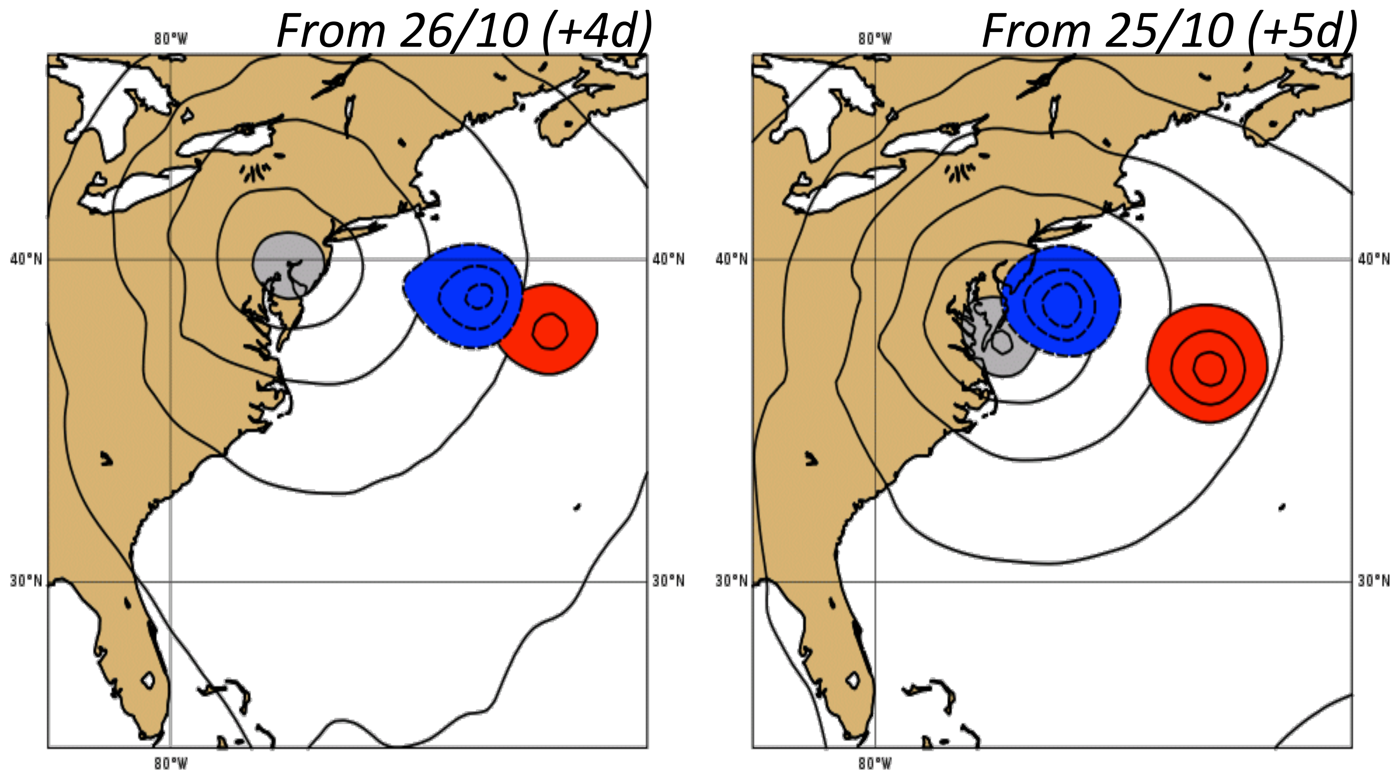


GPS Radio Occultation




Anomaly correlation of 500 hPa height forecasts
Northern hemisphere
Southern hemis
98.5
95.5
70
60
50
40
30
1981
1984
1987
1990
~210km
~125km
~63km
~39km
~25km
~16km
D+5
ERA-Interim
98
95
90
80
70
60
50
40
30
D+3
D+5
D+7
D+10
1980
1983
1986
1989
1992
1995
1998
2001
2004
2007
2010

# Observations and how they are assimilated make a difference




*Shading below 970hPa*

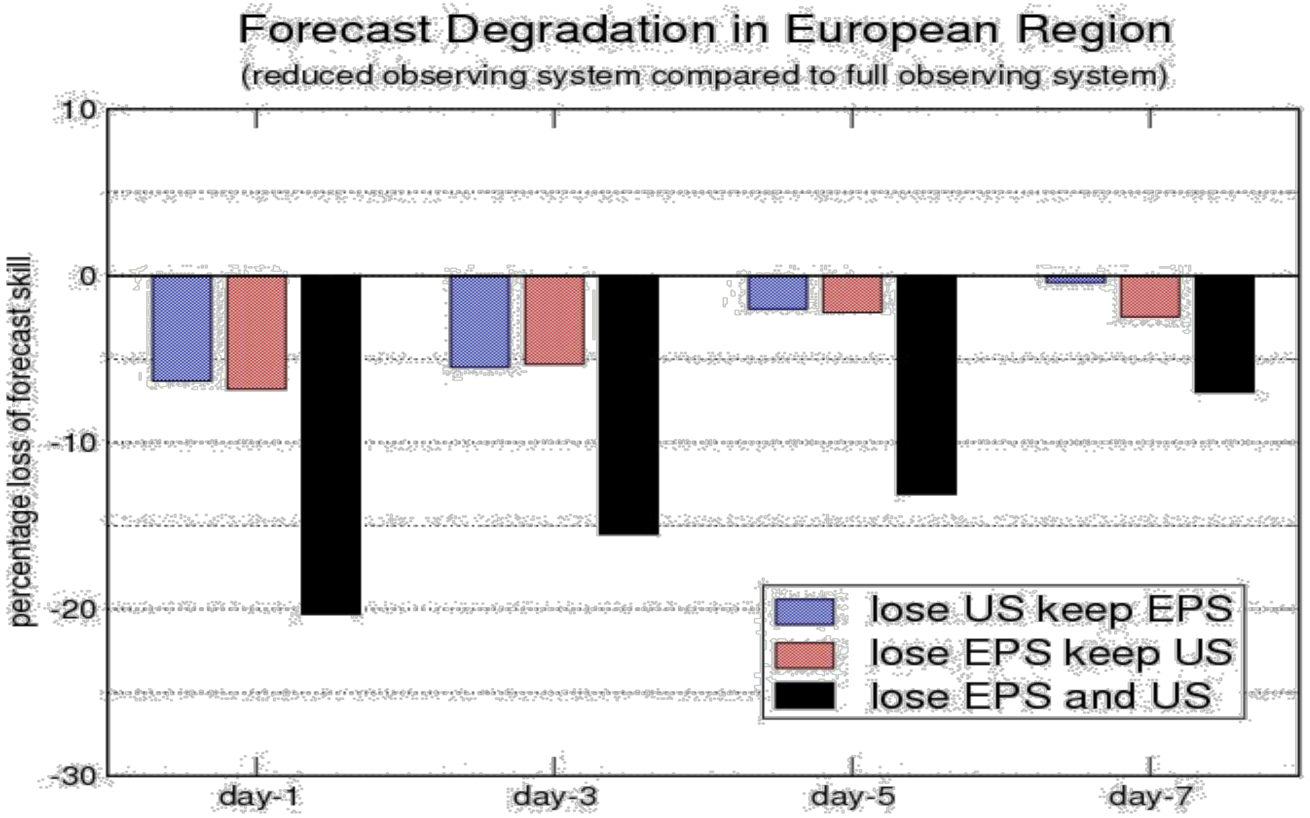
**Full system**

**No polar, brute force**

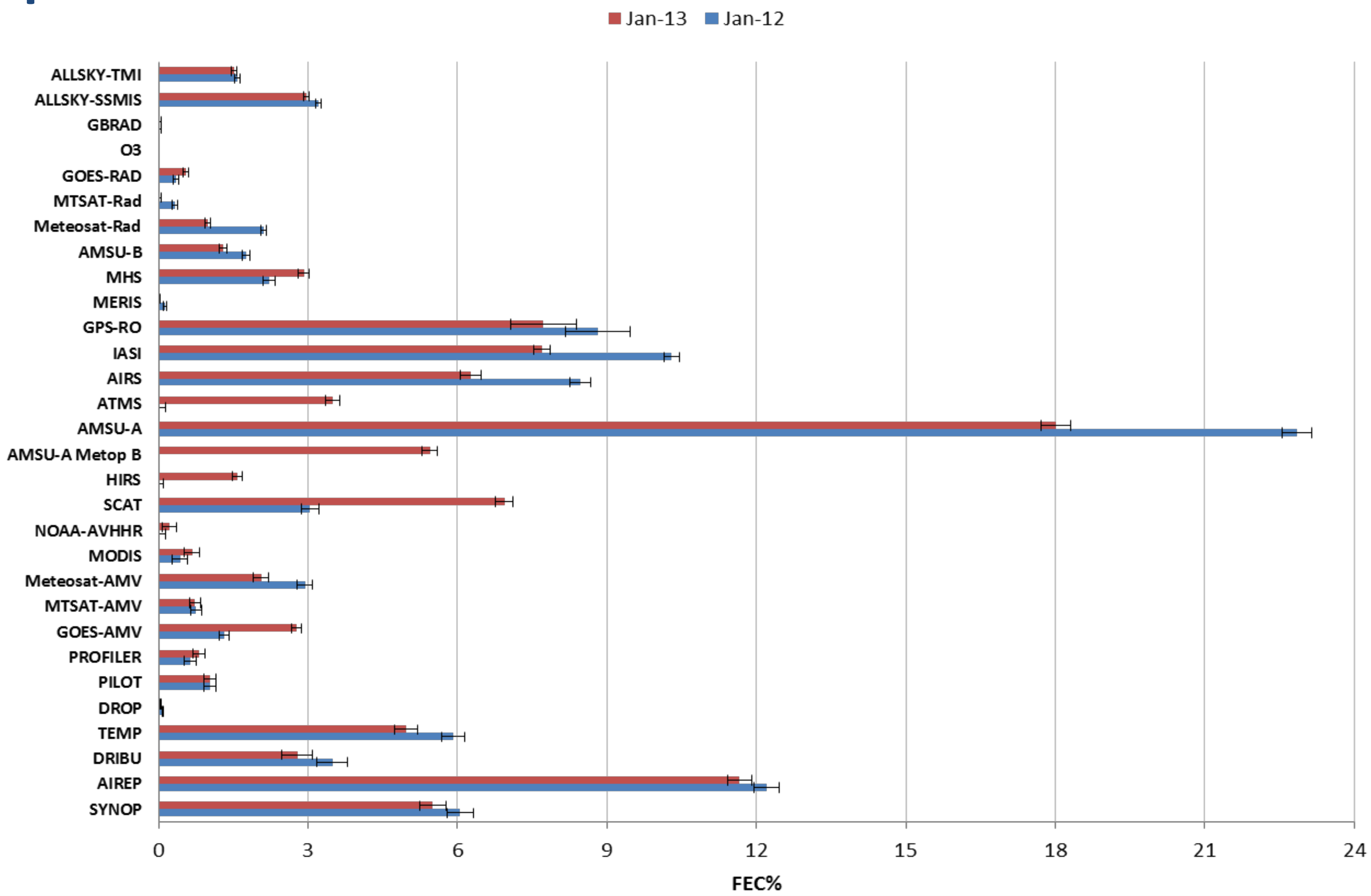
**No polar, EDA informed covariances**

*(T. McNally, M. Bonavita)*

# Impact of polar orbiting operational satellite



# Impact of Assimilated observations on Forecast Error Reduction



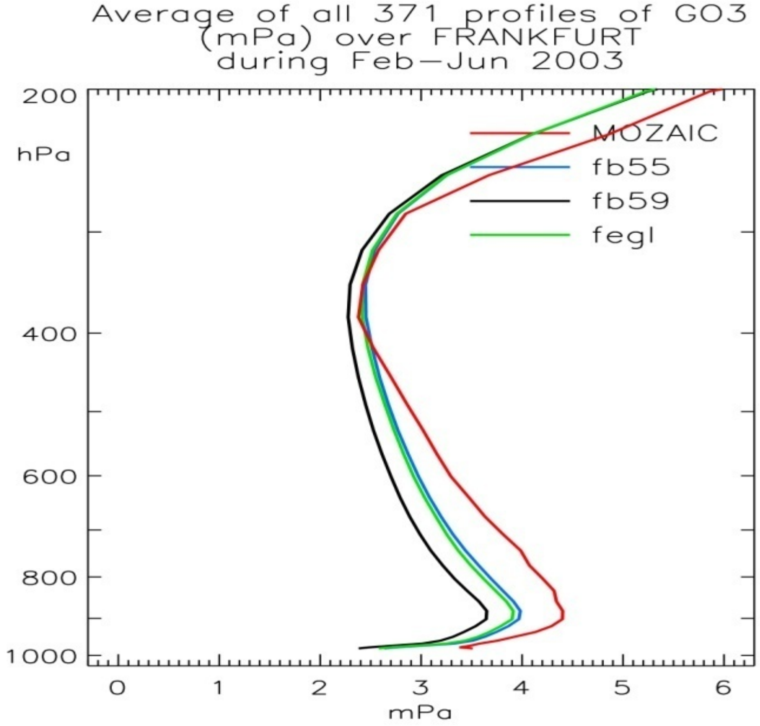
# CHEMICAL DATA ASSIMILATION

# Specific challenges for Atmospheric Composition DA

- Quality of NWP depends predominantly on initial state
- AC modelling depends on initial state (lifetime) and emission & surface fluxes
- CTMs have larger biases than NWP models
- Most processes take place in boundary layer, which is not well observed from space
- Only a few species (out of 100+ modeled) can be observed
- Data availability
- More complex and expensive, e.g. atmospheric chemistry ("stiff" equation systems), aerosol physics

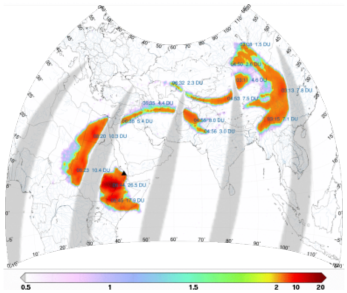
# But... Benefits of chemical coupling

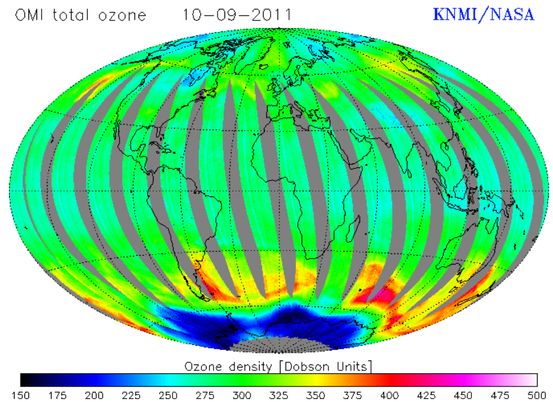
- Background NOx levels determine $O_3$ prod/loss
- Assimilation of $NO_2$ has an impact on ozone field (through chemical feedbacks in the CTM)
- Assimilation of $NO_2$ can improve $O_3$ field

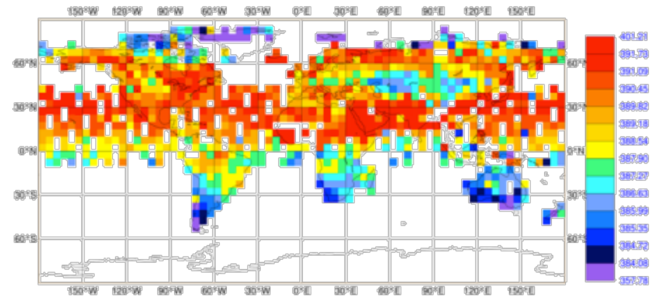


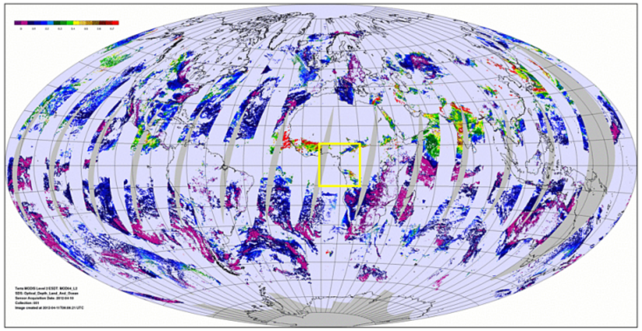

*Validation with MOZAIC/IAGOS ozone data*

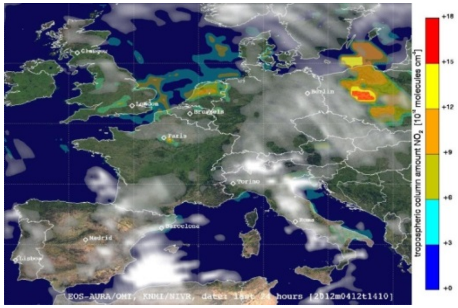
- Control (no CO or $NO_2$ assim, only $O_3$ assim) — black
- MOZAIC observation — red
- CO & $NO_2$ assim — blue
- NO2 assim — green









## Satellite observations

Atmospheric composition observations traditionally come from UV/VIS measurements. This limits the coverage to day-time only. Infrared/microwave are now adding more and more to this spectrum of observations (MOPITT, AIRS, IASI, MLS, MIPAS …)

# Data used in MACC-II/-III global NRT analysis @ECWMF

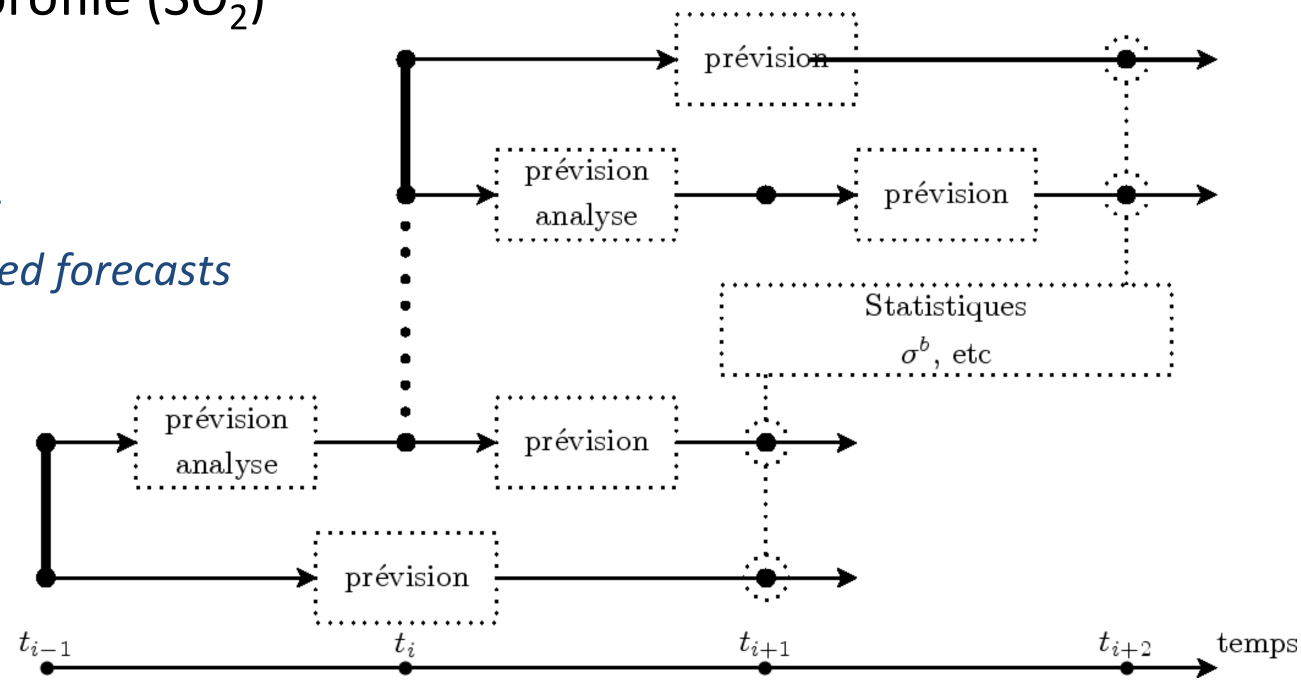
| Instrument | Satellite | Satellite operator | Data provider | Species | Status |
|---|---|---|---|---|---|
| **MODIS** | Terra | NASA | NASA/NOAA | Aerosol, fires | **Active** |
| **MODIS** | Aqua | NASA | NASA/NOAA | Aerosol, fires | **Active** |
| **SEVIRI** | Meteosat-9 | EUMETSAT | IM | Fires | **Active** |
| **Imager** | GOES-11, 12 | NOAA | NOAA | Fires | **Passive** |
| **Imager** | MTSAT-2 | JMA | JMA | Fires | Planned |
| **MLS** | Aura | NASA | NASA | $O_3$ | **Active** |
| **OMI** | Aura | NASA | NASA | $O_3$ | **Active** |
| **SBUV-2** | NOAA-17,18,19 | NOAA | NOAA | $O_3$ | **Active** |
| **SCIAMACHY** | Envisat | ESA | KNMI | $O_3$ | **Died** |
| **GOME-2** | Metop-A | EUMETSAT | DLR | $O_3$ | **Passive** |
| **SEVIRI** | Meteosat-9 | EUMETSAT | EUMETSAT | $O_3$ | **Passive** |
| **IASI** | Metop-A | EUMETSAT | LATMOS | CO | **Active** |
| **MOPITT** | Terra | NASA | NCAR | CO | **Active** |
| **GOME-2** | Metop-A | EUMETSAT | DLR | $NO_2$ | **Passive** |
| **OMI** | Aura | NASA | KNMI | $NO_2$ | **Active** |
| **OMI** | Aura | NASA | NASA | $SO_2$ | **Active** |
| **GOME-2** | Metop-A | EUMETSAT | DLR | $SO_2$ | **Passive** |
| **GOME-2** | Metop-A | EUMETSAT | DLR | HCHO | **Passive** |

# REACTIVE GASES DATA ASSIMILATION

# Practical issues and pragmatic approaches (1)

- How to specify Background errors?
  - NMC method (CO, NOx, HCHO)
  - Analysis ensemble method ($O_3$)
  - Prescribed profile ($SO_2$)

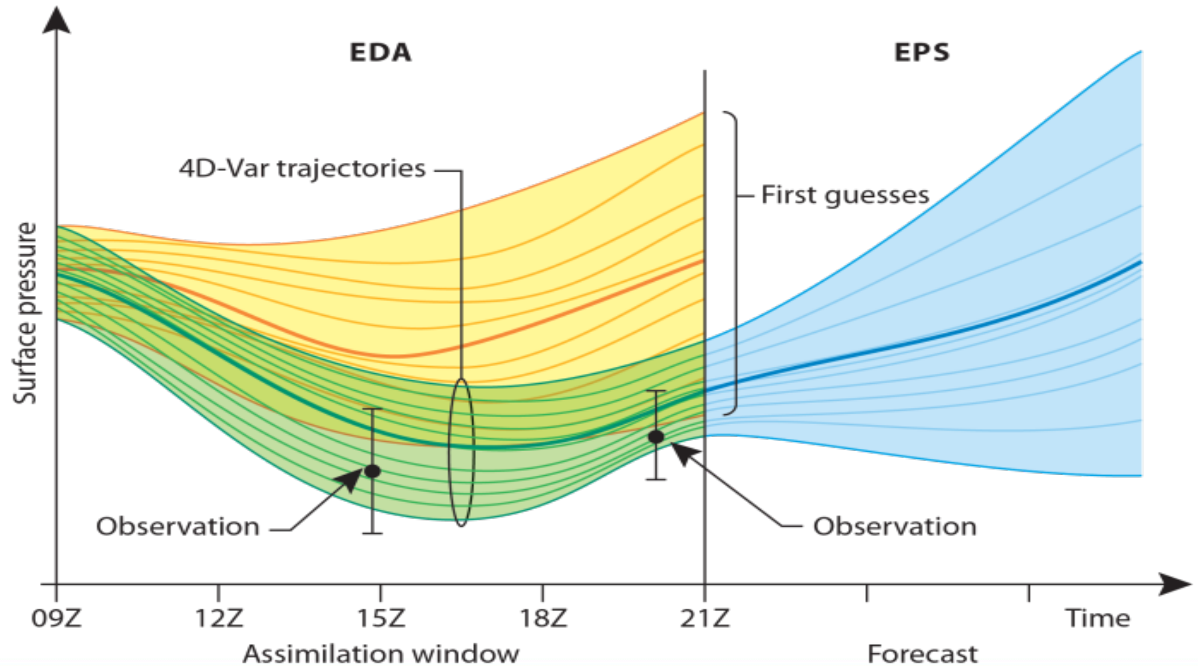
*NMC*
*Lagged forecasts*



# Practical issues and pragmatic approaches (1)

- How to specify Background errors?
  - NMC method (CO, NOx, HCHO)
  - Analysis ensemble method ($O_3$)
  - Prescribed profile ($SO_2$)

*EDA*
*Ensemble of Data Assimilation*



# Practical issues and pragmatic approaches (3)

- Variational bias correction used for reactive gases

The original problem:

$J_b$: background constraint

$$J(\mathbf{x}) = (\mathbf{x}_b - \mathbf{x})^T \mathbf{B}^{-1}(\mathbf{x}_b - \mathbf{x}) + [\mathbf{y} - \mathbf{h}(\mathbf{x})]^T \mathbf{R}^{-1}[\mathbf{y} - \mathbf{h}(\mathbf{x})]$$

$J_o$: observation constraint

The modified problem:

$J_b$: background constraint for x

$J_\beta$: background constraint for $\beta$

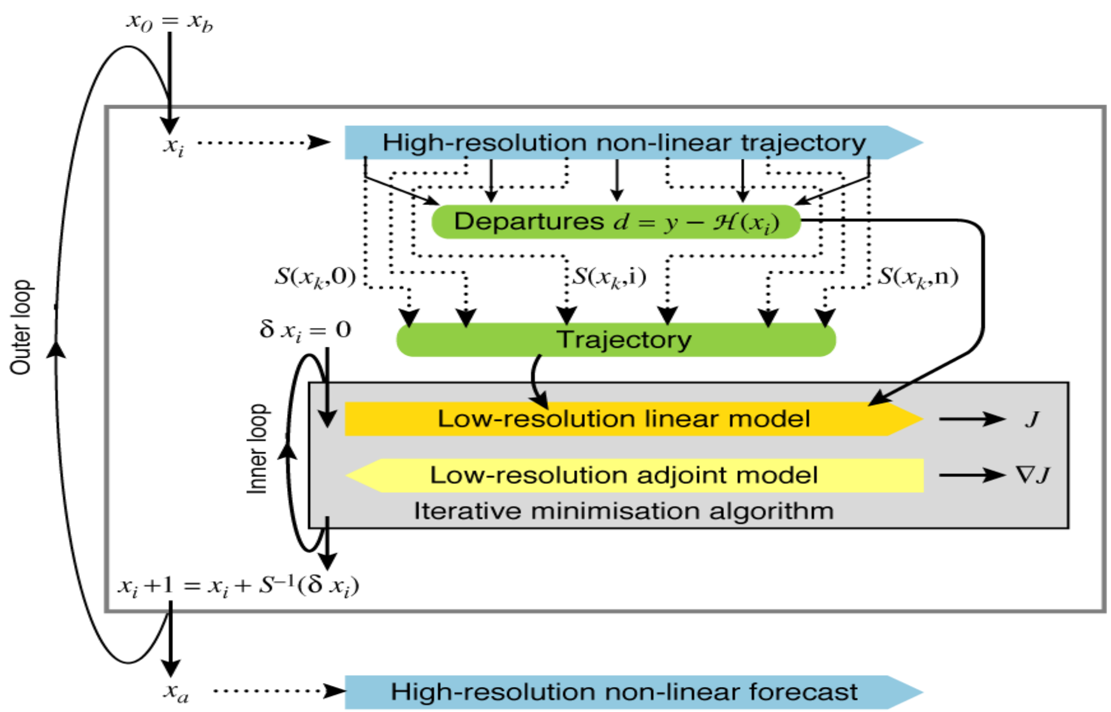
$$\mathbf{J}(\mathbf{x}, \beta) = (\mathbf{x}_b - \mathbf{x})^T \mathbf{B}_x^{-1}(\mathbf{x}_b - \mathbf{x}) + (\beta_b - \beta)^T \mathbf{B}_\beta^{-1}(\beta_b - \beta) + [\mathbf{y} - \mathbf{b}_o(\mathbf{x}, \beta) - \mathbf{h}(\mathbf{x})]^T \mathbf{R}^{-1}[\mathbf{y} - \mathbf{b}_o(\mathbf{x}, \beta) - \mathbf{h}(\mathbf{x})]$$

Parameter estimates from previous analysis

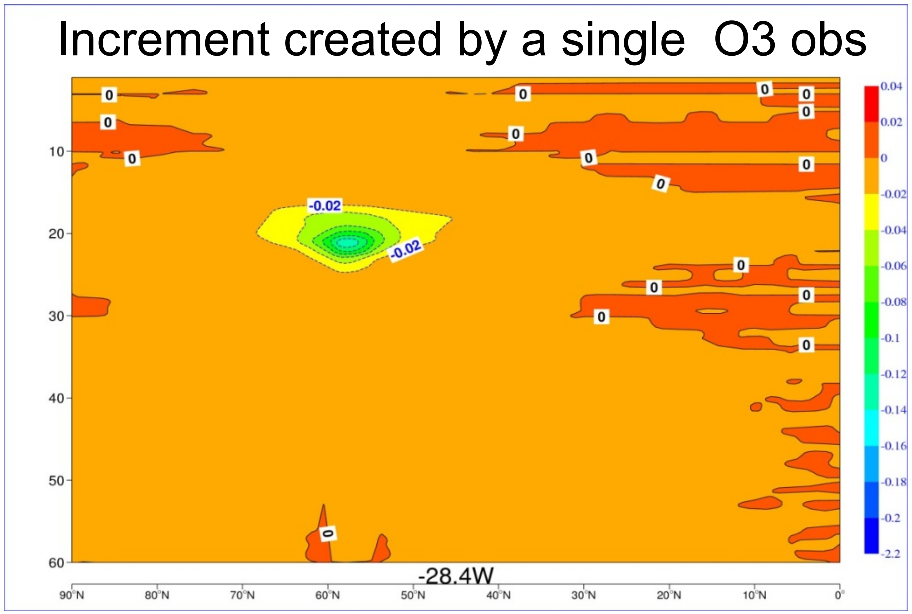
$J_o$: bias-corrected observation constraint

# Practical issues and pragmatic approaches (4)

- Difficulties assimilating species with short lifetimes (e.g. $NO_2$): NOx as control variable and $NO_2$-NOx interconversion operator
- Chemistry included in outer loop not in minimisation; adjoint of transport only

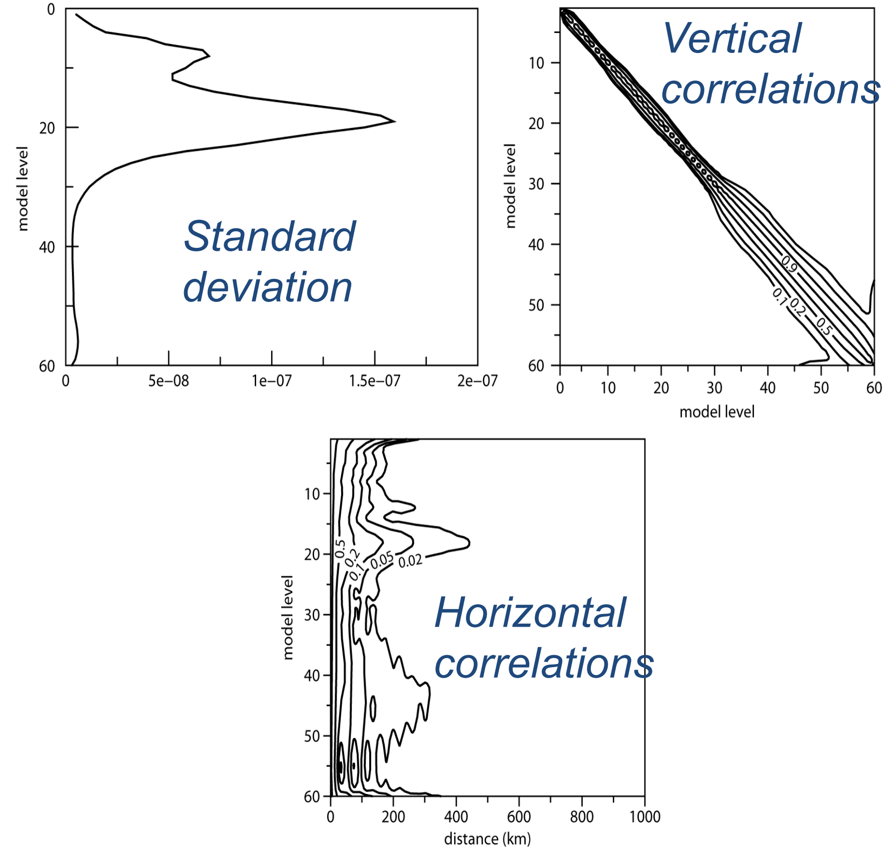


# Increment from a single TCO3 observation

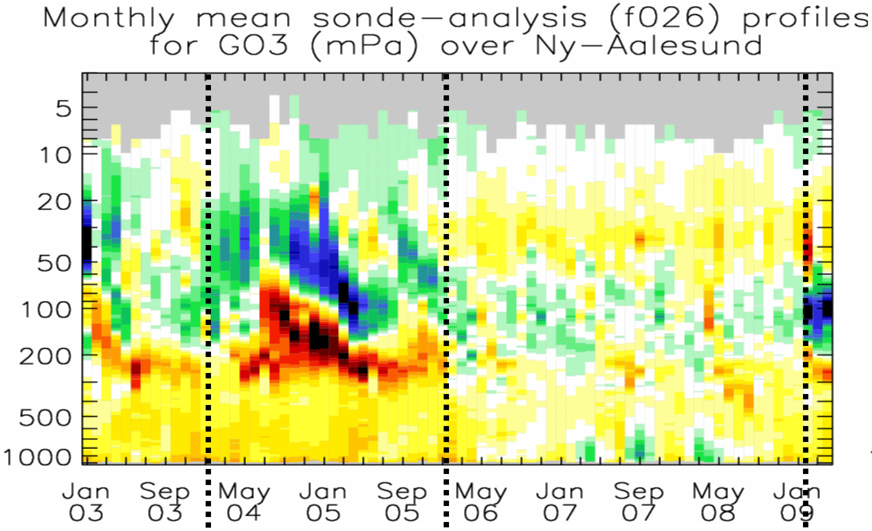


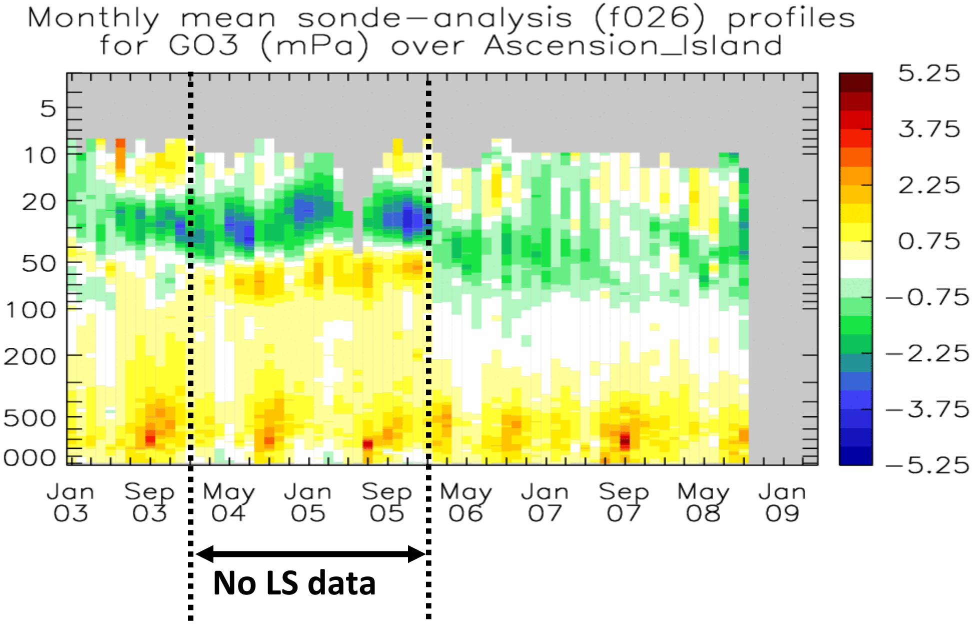

Ozone observation of 247 DU, 66 DU lower than background

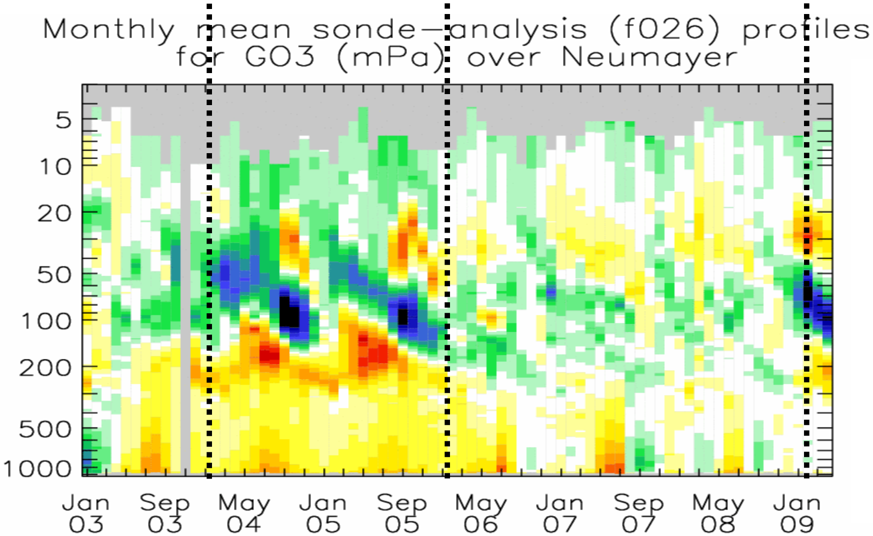
Ozone background errors




Profile data are important to obtain a good vertical analysis profiles



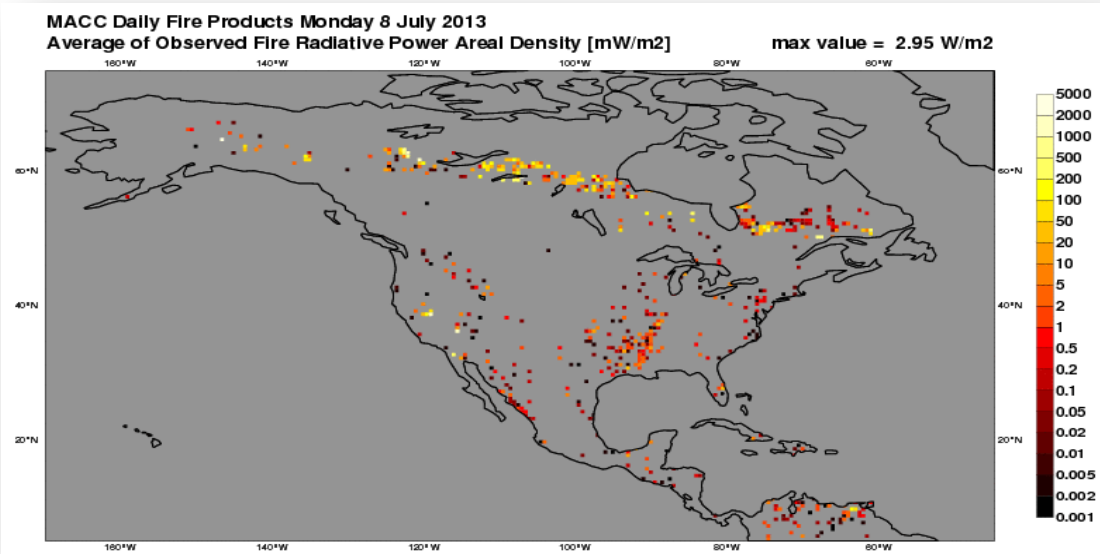




Limb-sounding ozone data assimilated in 2003 (MIPAS) and 2006-2008 (MLS)

These data are clearly beneficial

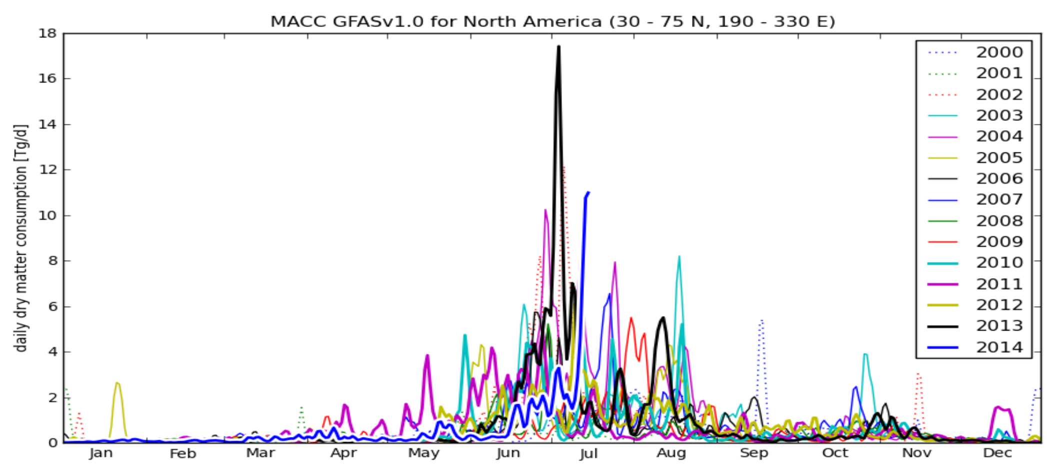
OMI data are used from July 2007

# The Global Fire assimilation System

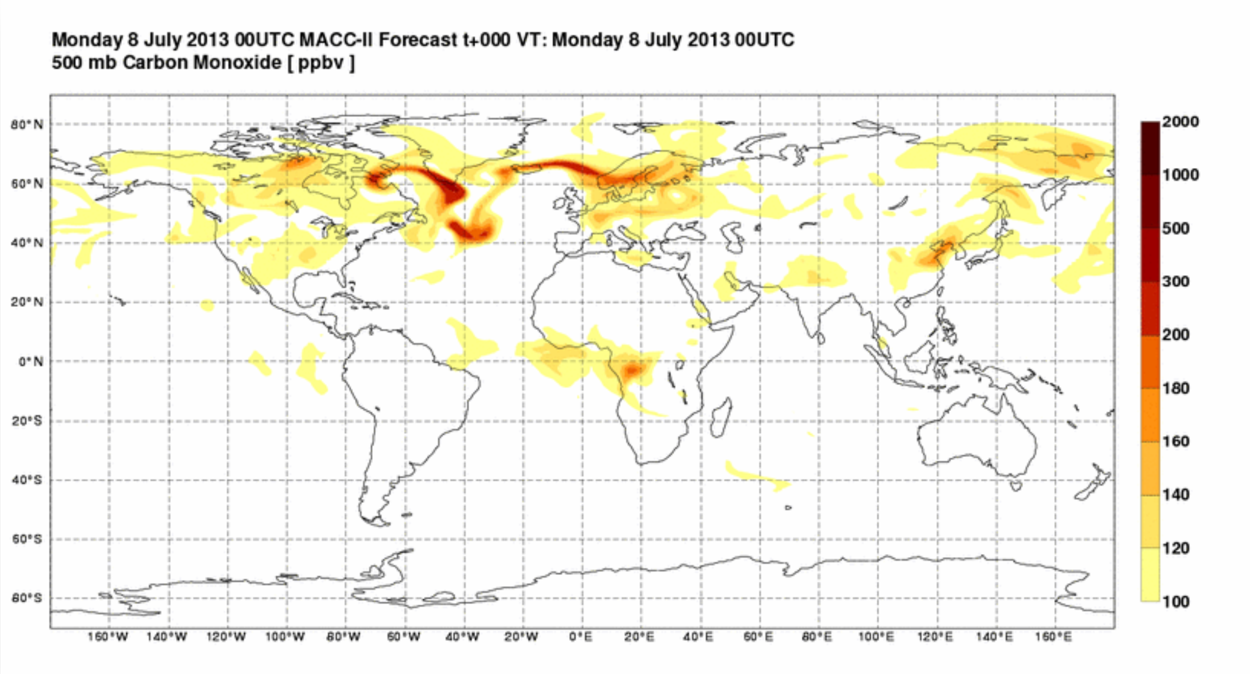


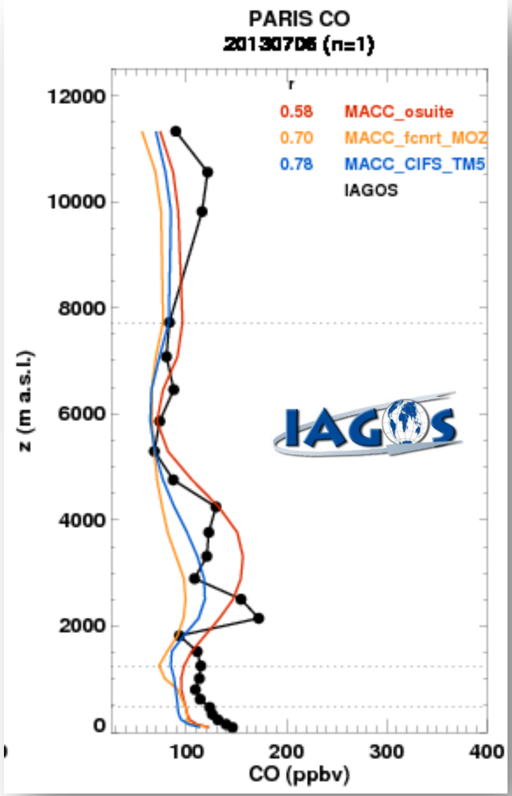

Fire Radiative Power from satellite measurements (MODIS and GEOs)...

...Converted to emissions (with estimation of plume height)



# Canadian fires (July 2013)



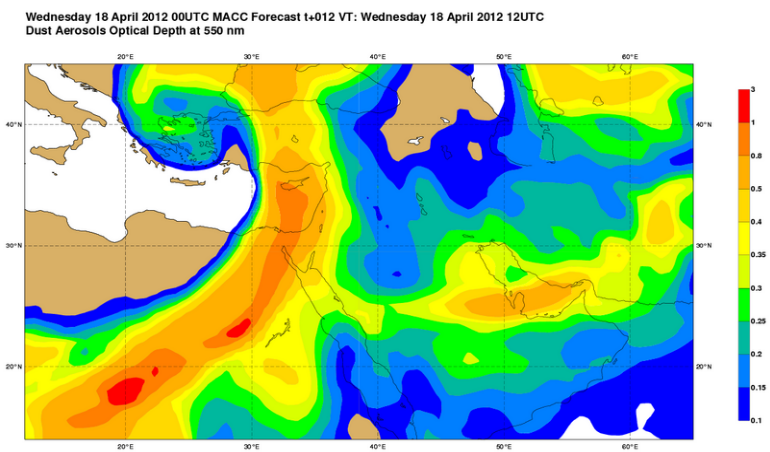




Satellite observation of meteorology and composition (in the case of CO from MOPITT and IASI) are used for assimilation while the pollutant plumes travels in the atmosphere (one daily update)

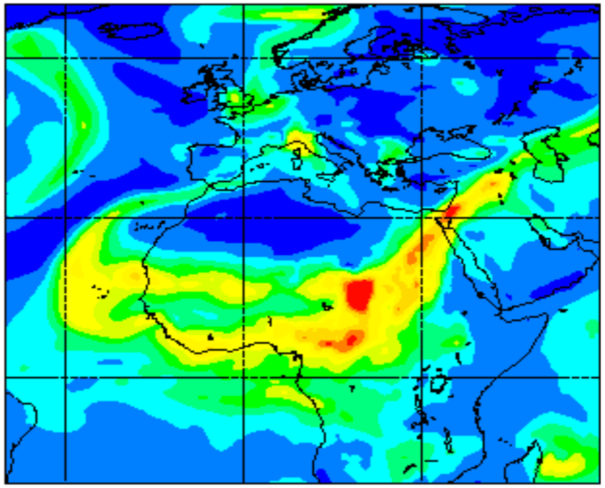
# AEROSOL DATA ASSIMILATION

Aerosol assimilation is difficult because:

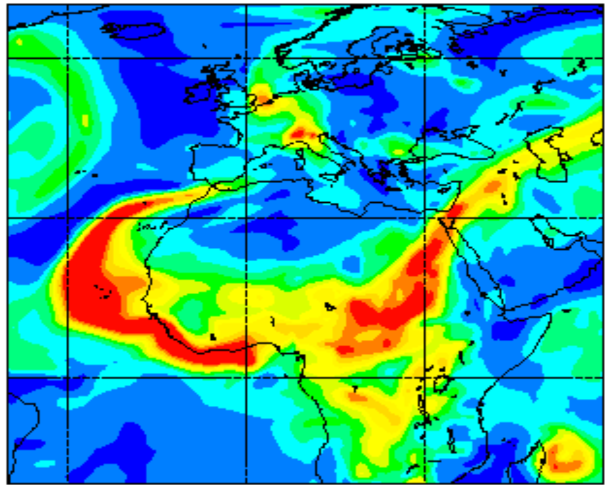
- There are numerous unknowns (depending on the aerosol model) and very little observations to constrain them
- The concentrations vary hugely with for instance strong plumes of desert dust in areas with very little background aerosol, which makes it difficult to estimate the background error covariance matrix



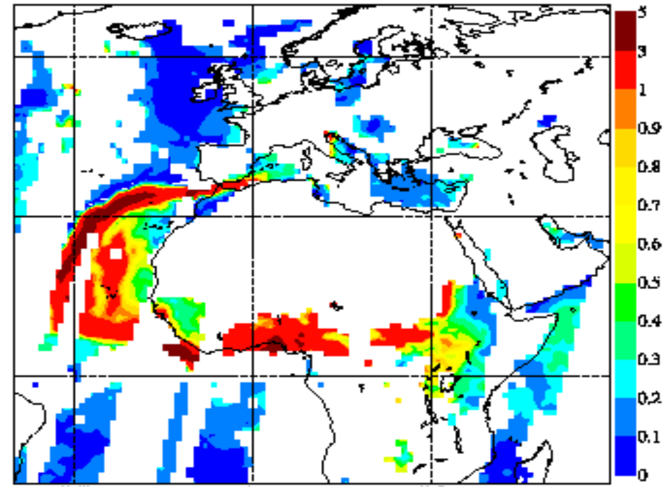

Model simulation



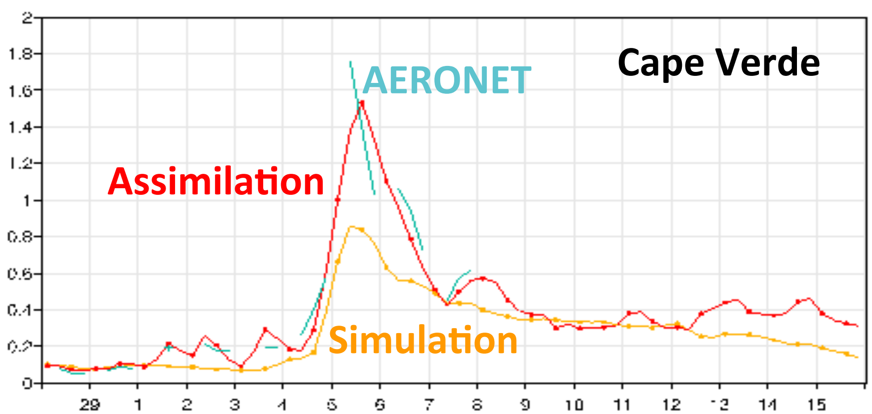
Assimilation

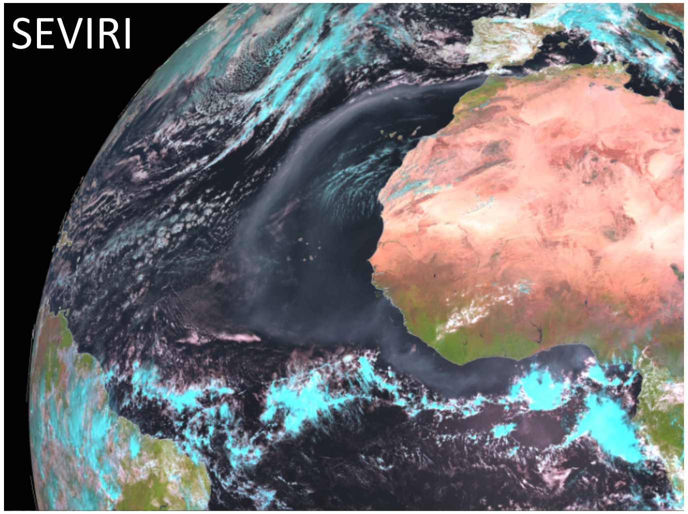


MODIS

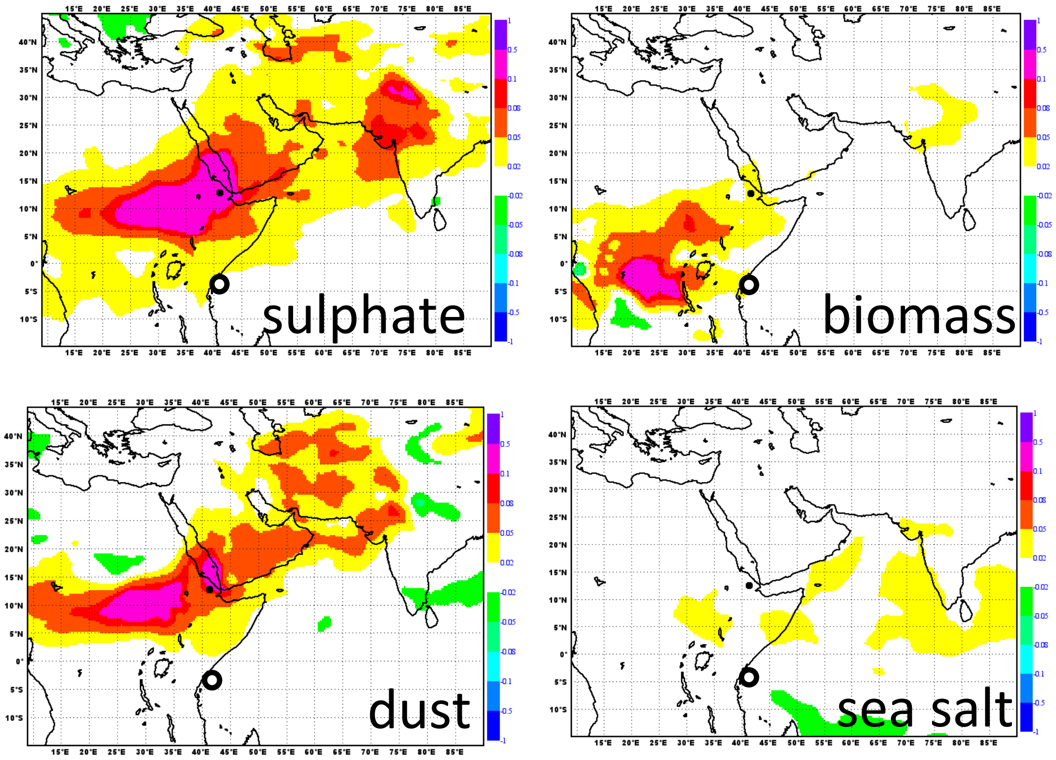



Aerosol optical depth at 550nm (upper)

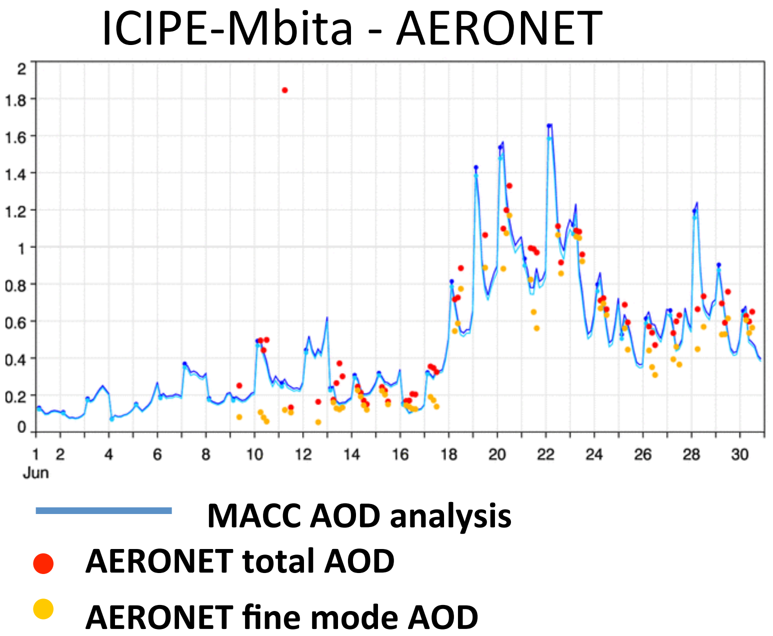


# Right for wrong reasons (aliasing)

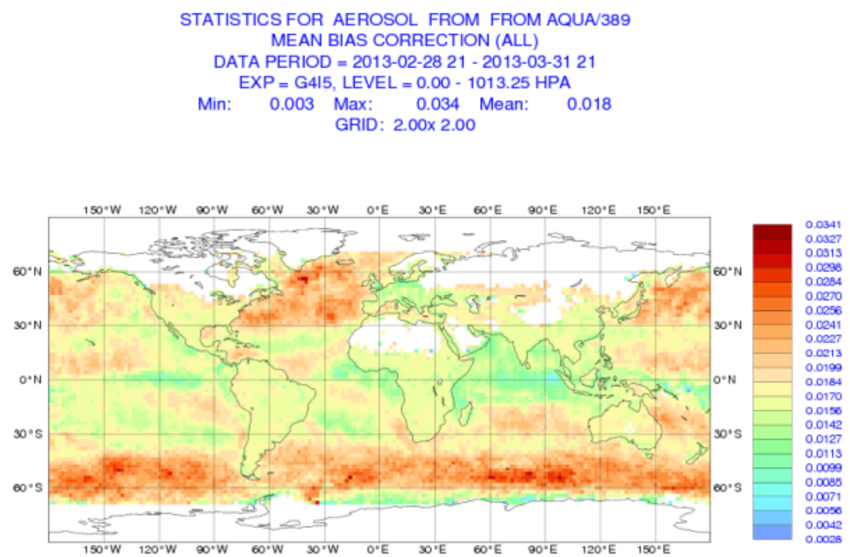


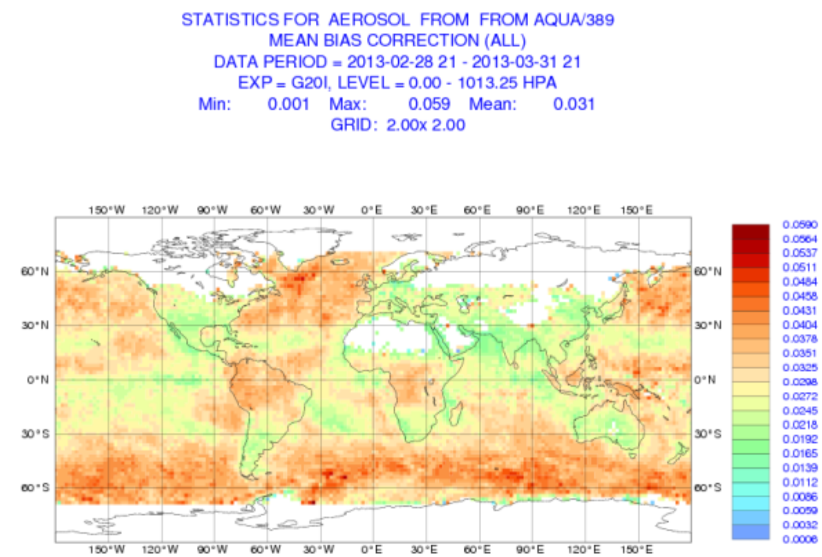

Eruption of the Nabro volcano in 2011 put a lot of fine ash into the stratosphere.
This was observed by AERONET stations and the MODIS instrument.




*The MACC aerosol model did not contain stratospheric aerosol, so the observed AOD was attributed to the available aerosol types.*

# AOD assimilation & issues

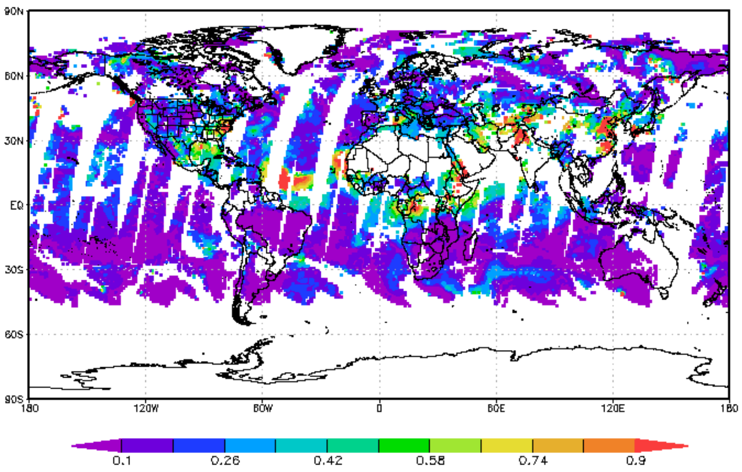


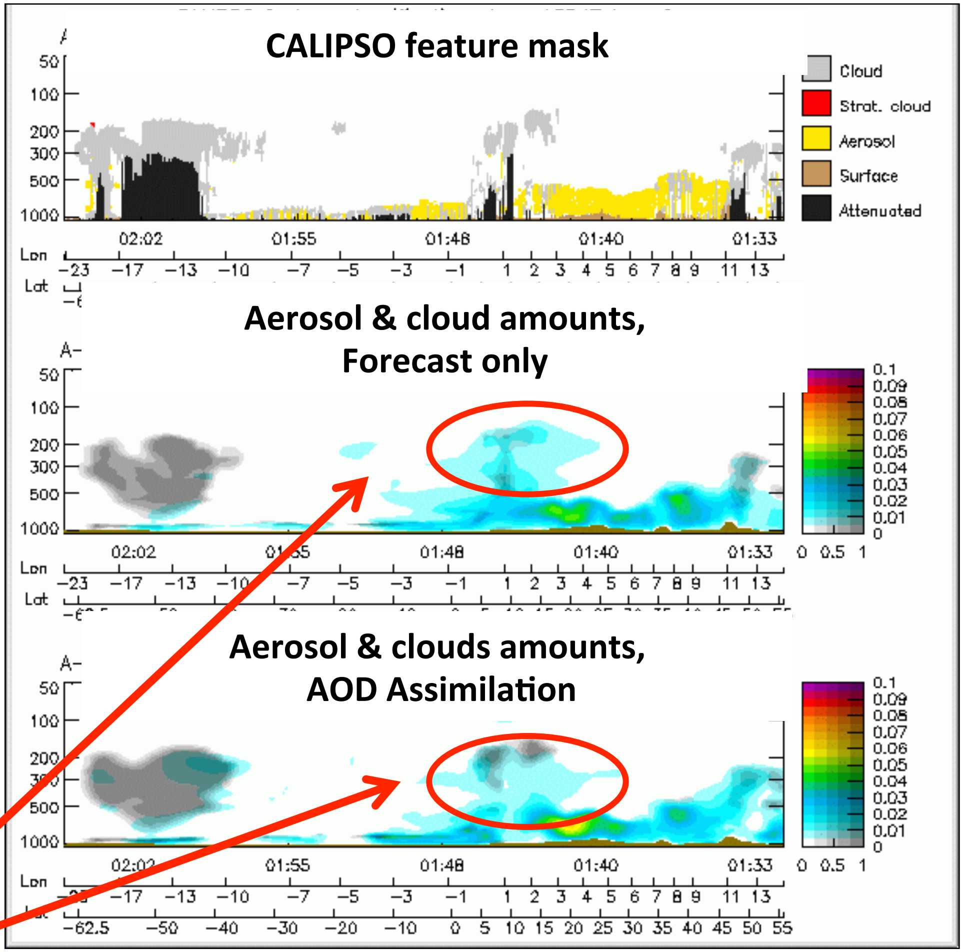




Variational bias correction for March 2013 for MODIS AOD at 550 nm for global constant and wind speed over ocean formulation (old, left) and global constant, wind speed over ocean, and cloud cover formulation (new, right).

(A. Benedetti)

## MODIS Aerosol Optical Depth

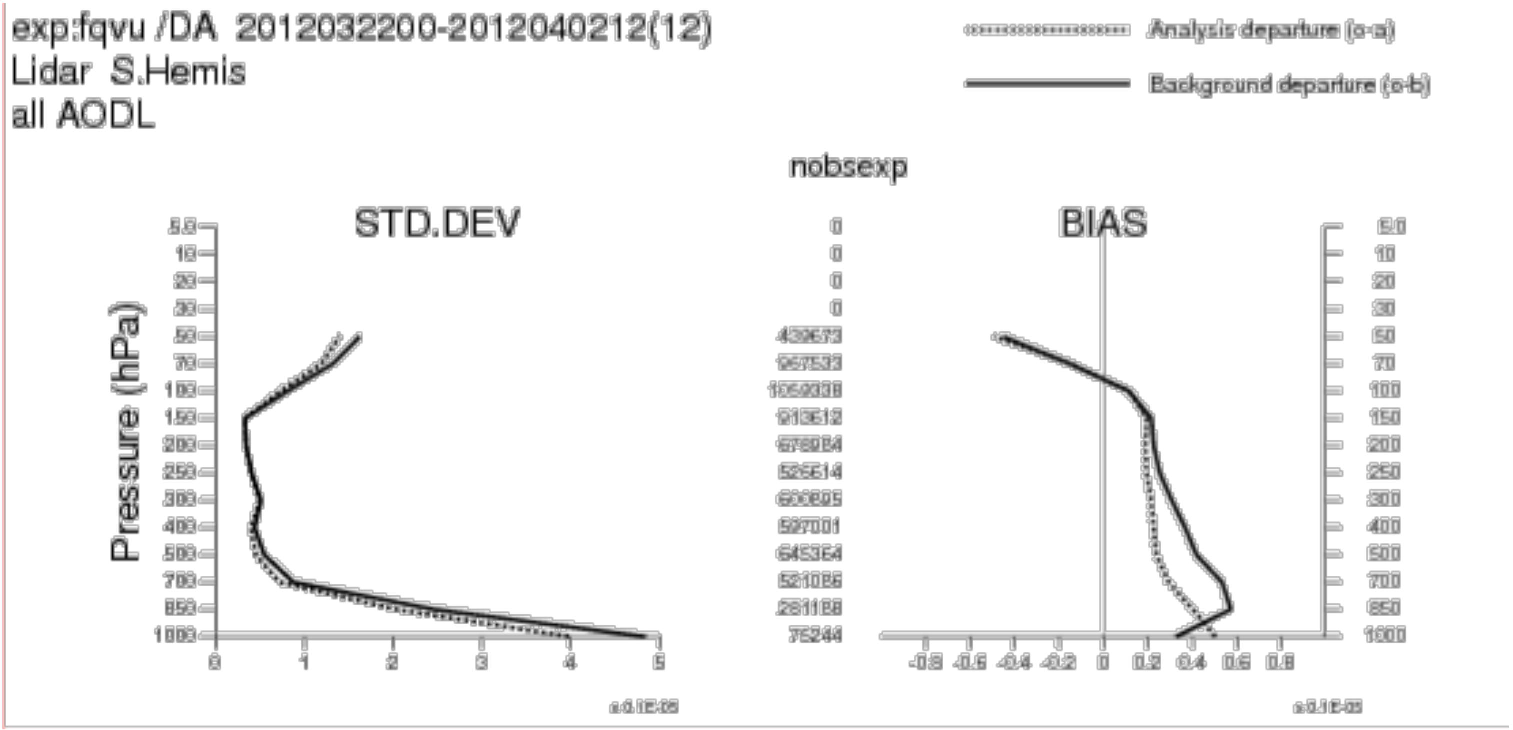



- AOD is a column-integrated quantity
- Assimilation of AOD does not modify the vertical profile
- Profile data are needed (lidar)

# Assimilation of CALIOP data (L1.5)

Period: March 21- April 2, 2012
Data: all operational data plus MODIS AOD and CALIOP Level 1.5 backscatter



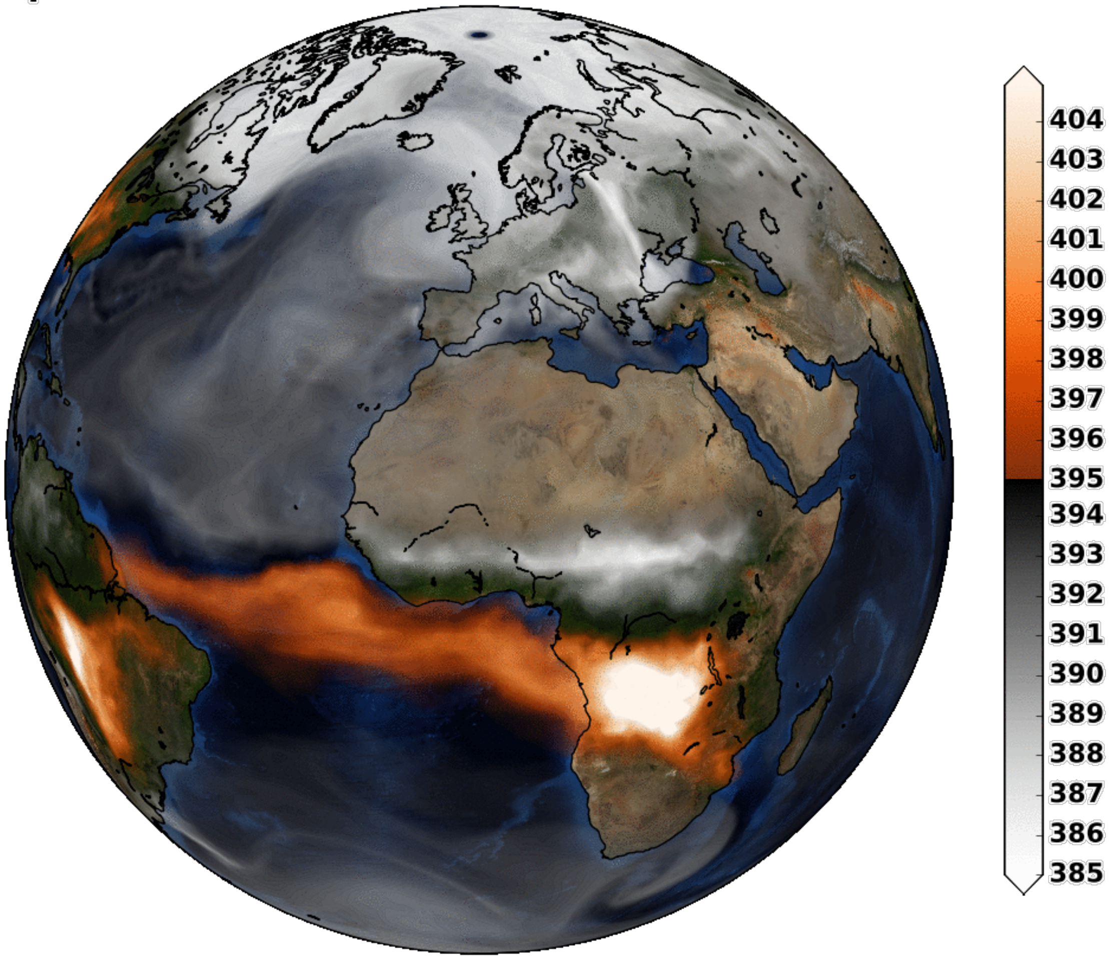

Lidar backscatter x 1e7 (sr m)$^{-1}$

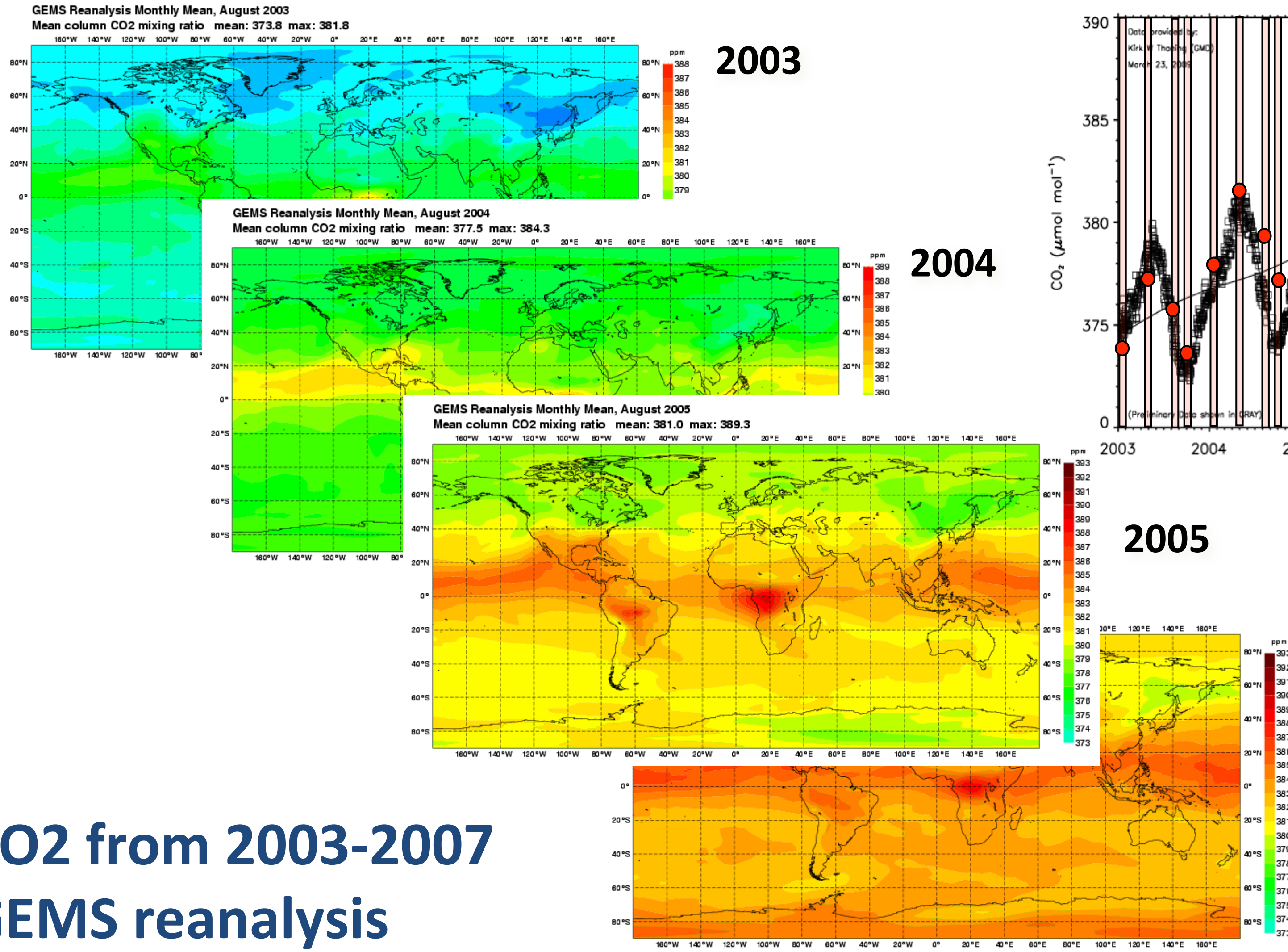


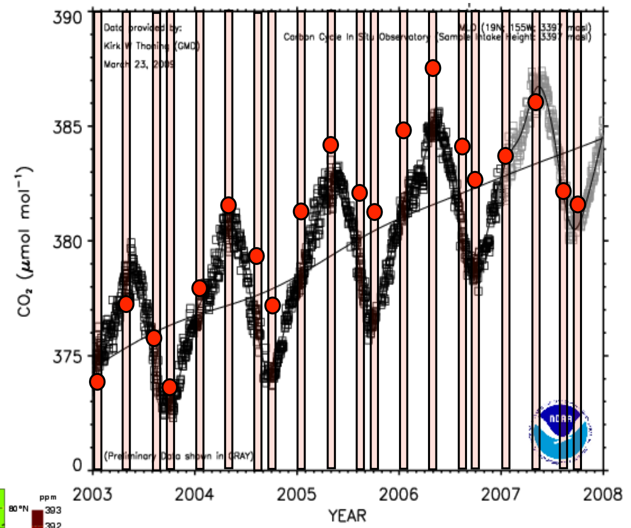
(A. B

# GREENHOUSE GAS ASSIMILATION

MACC column-averaged dry-air mole fraction of $CO_2$ [ppm]
September 2013

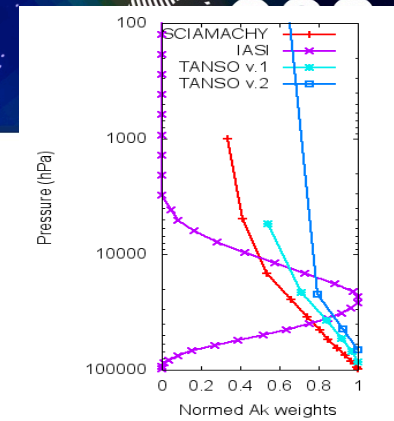




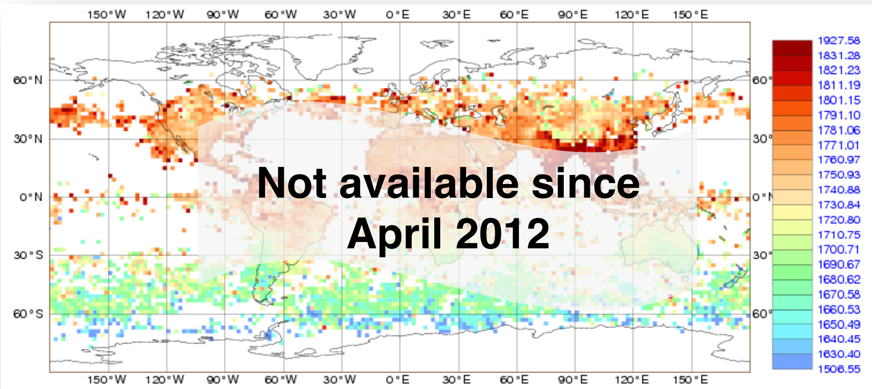

# CO2 from 2003-2007 GEMS reanalysis

# Assimilated GHG data @ECMWF



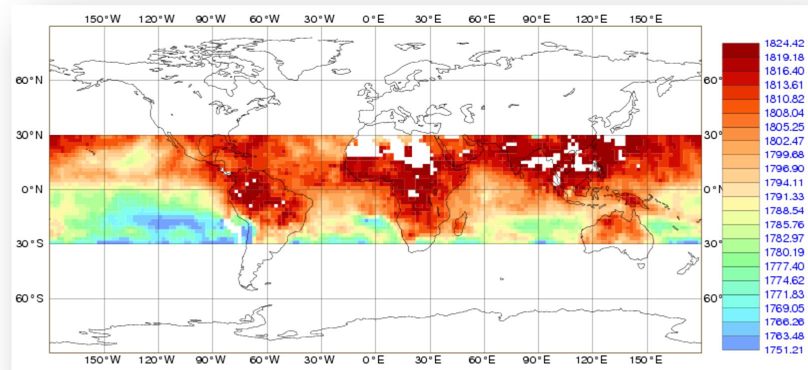

*ENVISAT/SCIAMACHY*

$CH_4$ and $CO_2$ – Lower tropo.



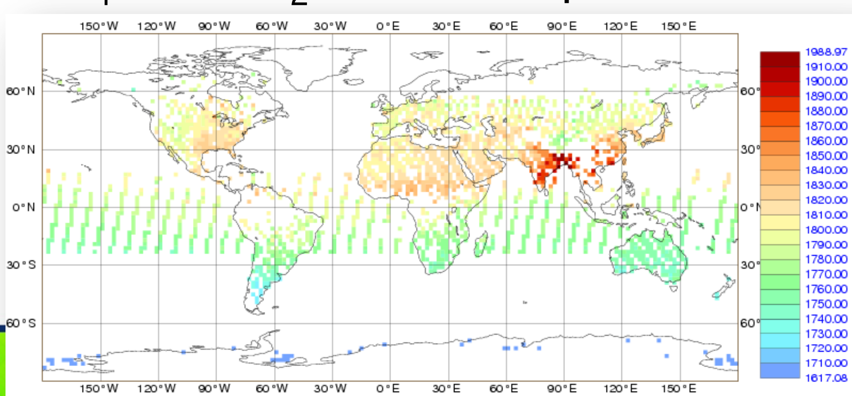

*METOP-A/IASI*

$CH_4$ and $CO_2$ – Middle tropo.




*GOSAT/TANSO*

$CH_4$ and $CO_2$ – Lower tropo.




Column-averaged dry-air mole fractions of CO2 and CH4 provided by:

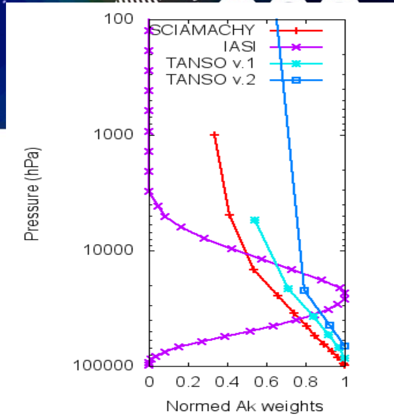
SRON Netherlands Institute for Space Research

LMD

Universität Bremen

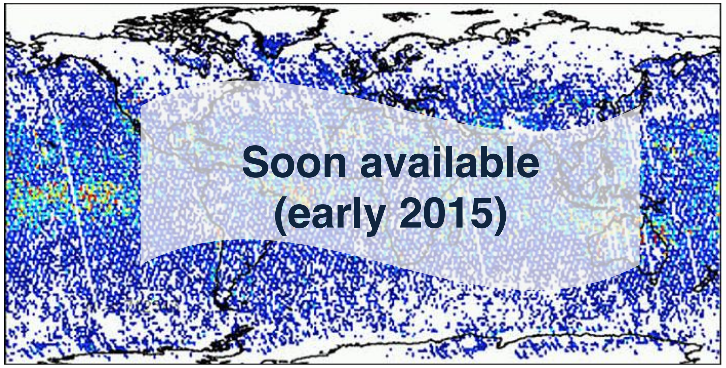
JÜLICH FORSCHUNGSZENTRUM

# Assimilated GHG data @ECMWF



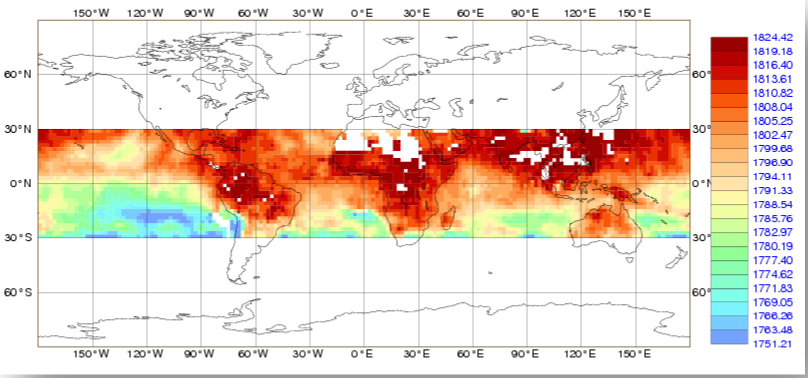

*OCO-2*

$CO_2$ – Lower tropo.



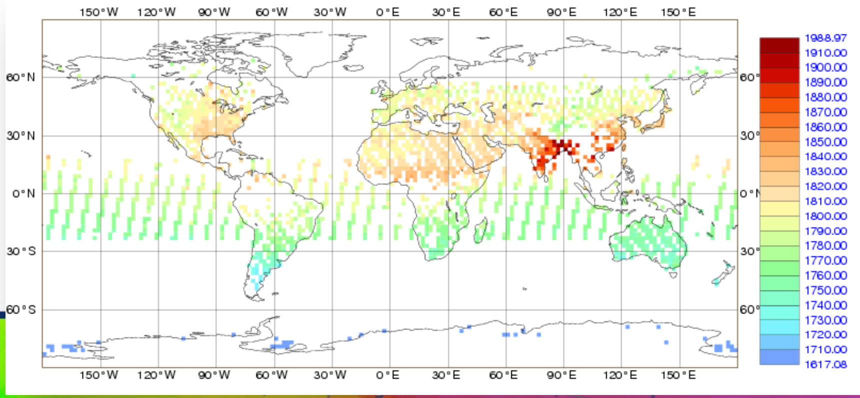

*METOP-A/IASI*

$CH_4$ and $CO_2$ – Middle tropo.



*GOSAT/TANSO*

$CH_4$ and $CO_2$ – Lower tropo.



Column-averaged dry-air mole fractions of CO2 and CH4 provided by:

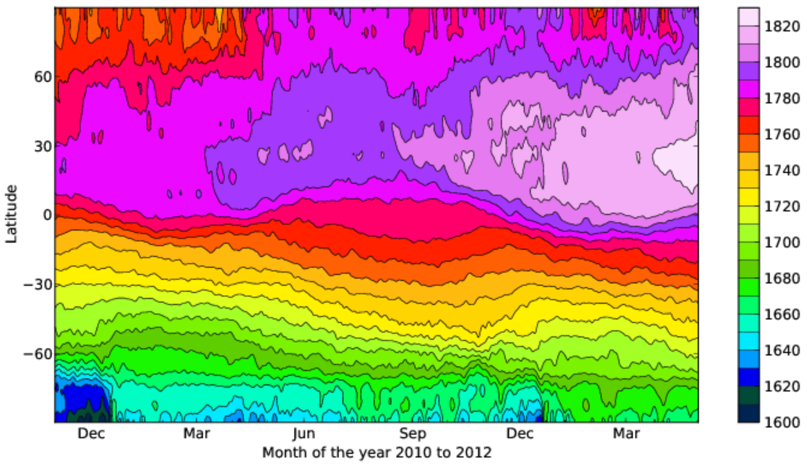
SRON Netherlands Institute for Space Research

LMD

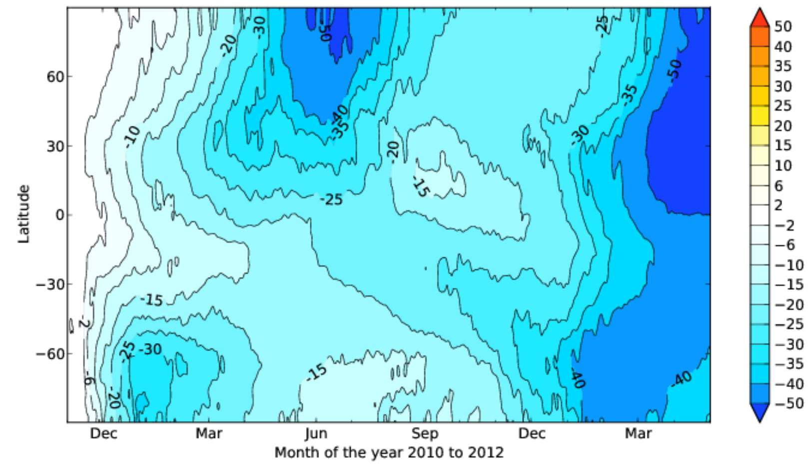
Universität Bremen

# Impact of each instrument

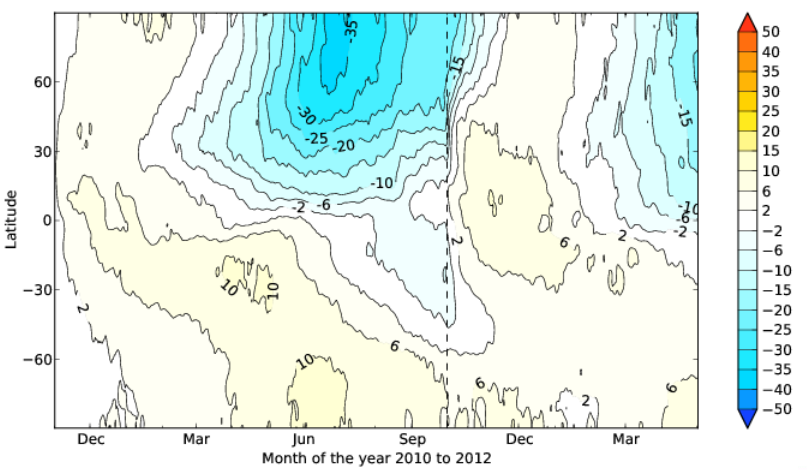
*Forecast model (FC)*



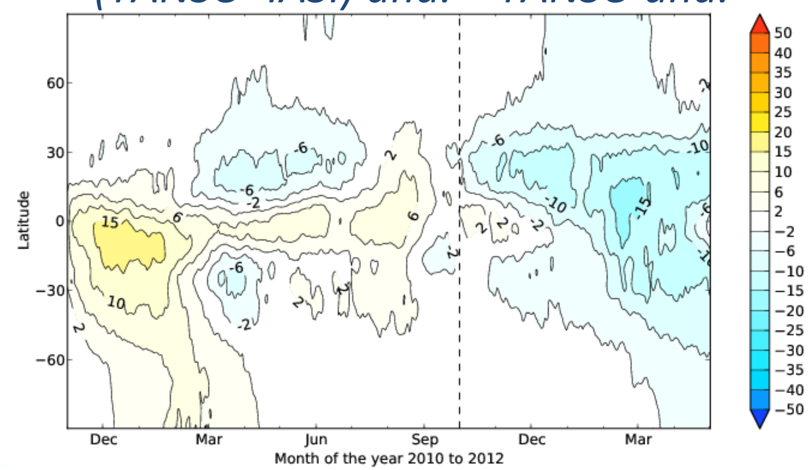

*SCIAMACHY ana. – FC*




*TANSO ana. – FC*

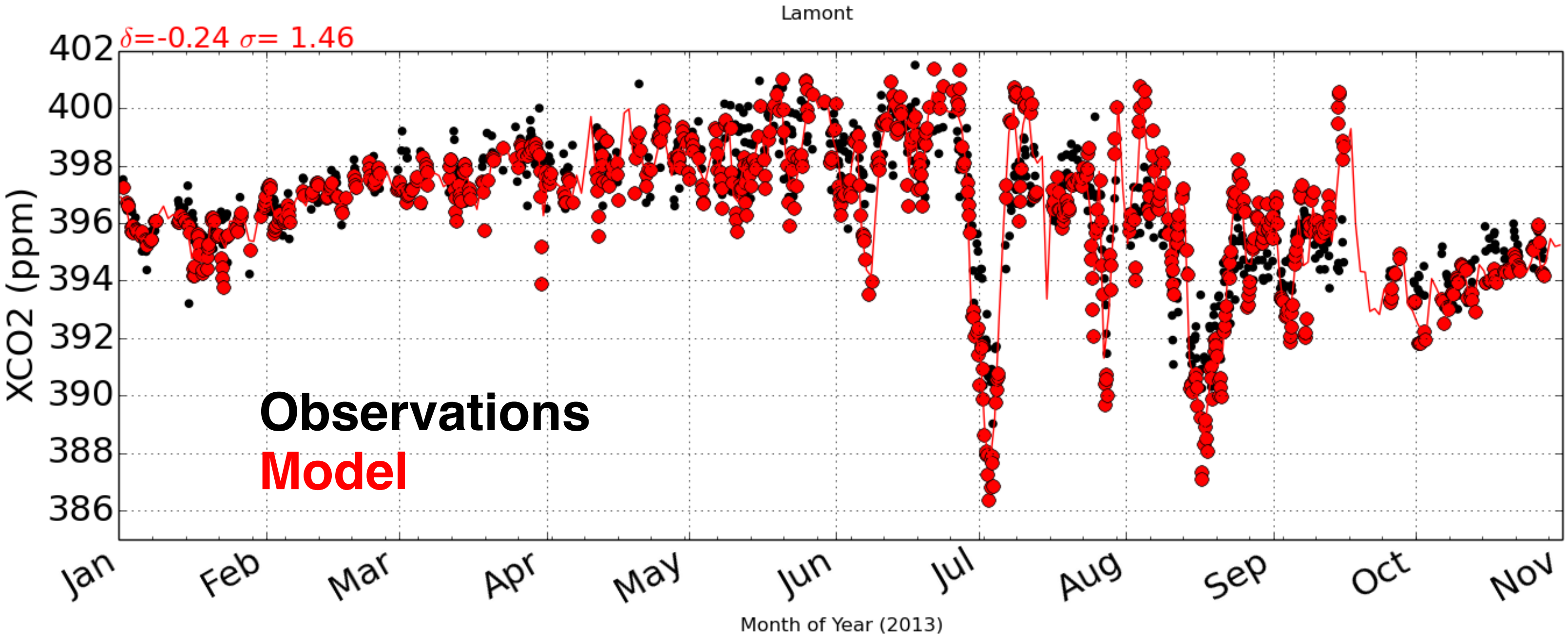



*(TANSO+IASI) ana. – TANSO ana.*



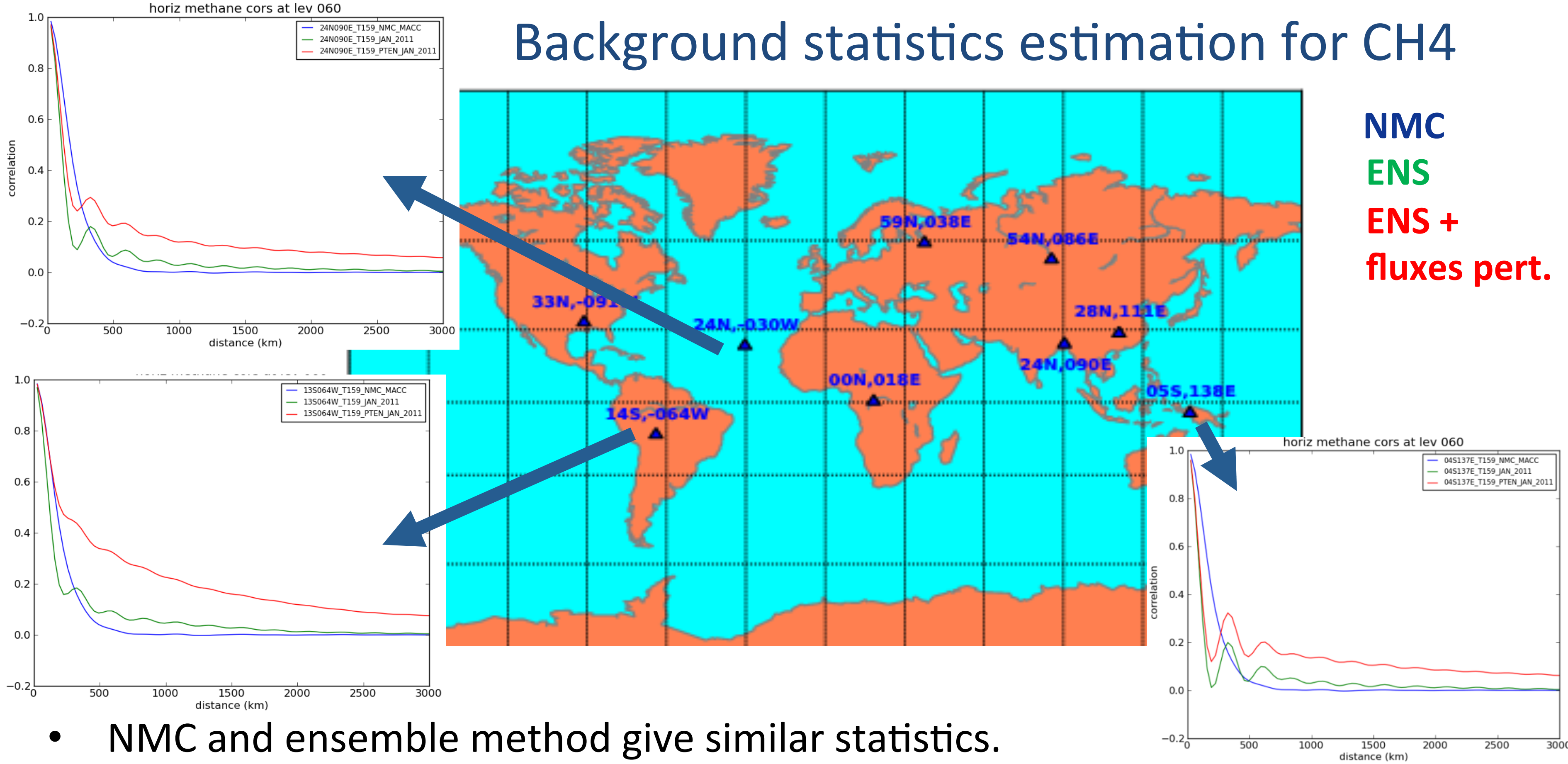
# Validation of $XCO_2$ at Lamont

**2013**




TCCON

# Background statistics estimation for CH4




- NMC and ensemble method give similar statistics.
- Ensemble method + perturbed fluxes give different statistics and geographical differences
- Work in progress to test the new background statistics

# Concluding remarks

- Chemical data assimilation is catching up with latest NWP data assimilation developments: reactive gases, greenhouse gases, aerosols.
- Specific challenges for DA of atmospheric composition compared to NWP (but also extra benefits through chemical coupling and potential impact).
- Still work in progress (issue of model biases, aerosol lidar assimilation, estimation of background errors…).
- We are providing **near-real time analyses and forecasts** of atmospheric composition as well as **reanalyses (2003-2012)** for reactive gases, greenhouse gases, and aerosols as part of the MACC-III project (www.copernicus-atmosphere.eu).

http://old.ecmwf.int/newsevents/training/meteorological_presentations/2013/DA2013/index.html

Vincent-Henri.Peuch@ecmwf.int

ECMWF
EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

"Miss Peterson, may I go home? I can't assimilate any more data today."