Assimilation of satellite data for the environment

Frédéric Chevallier, Peter Rayner
Laboratoire des Sciences du Climat et de l'Environnement
CEA/CNRS/UVSQ, IPSL, France

Richard Engelen
ECMWF, Reading

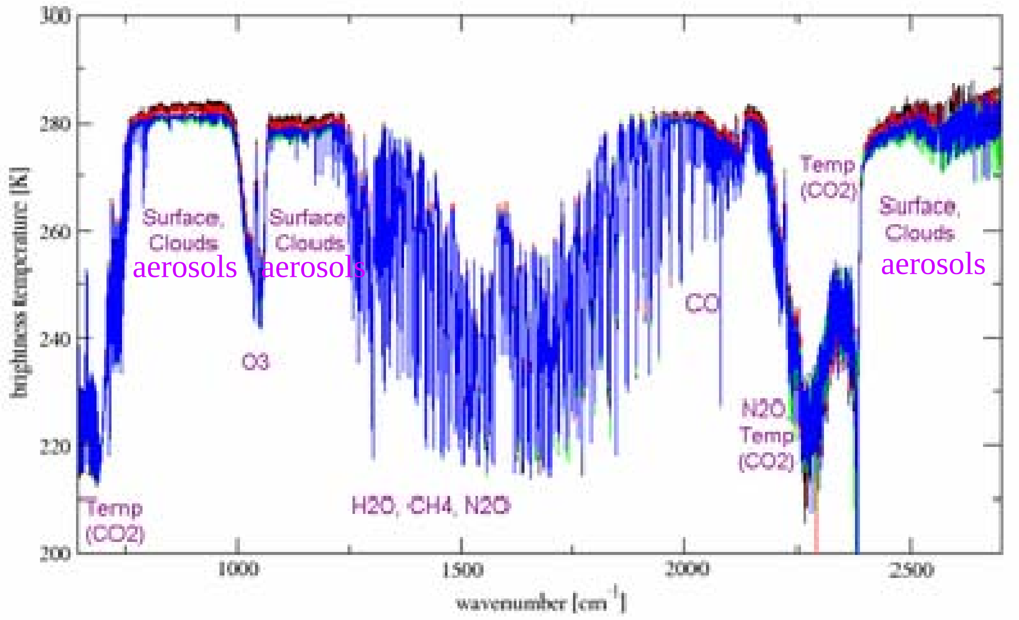
Thanks

- Audrey Fortems, Philippe Peylin and Sophie Szopa
  - LSCE
- The members of the FP6 GEMS consortium
  - Tony Hollingsworth
  - Antje Dethof and Angela Benedetti

Outline

- Introduction: CO_2 from space
- Optimization of atmospheric concentrations
- Optimization of surface fluxes
- Optimization of surface model parameters
- Conclusion

The atmosphere from space



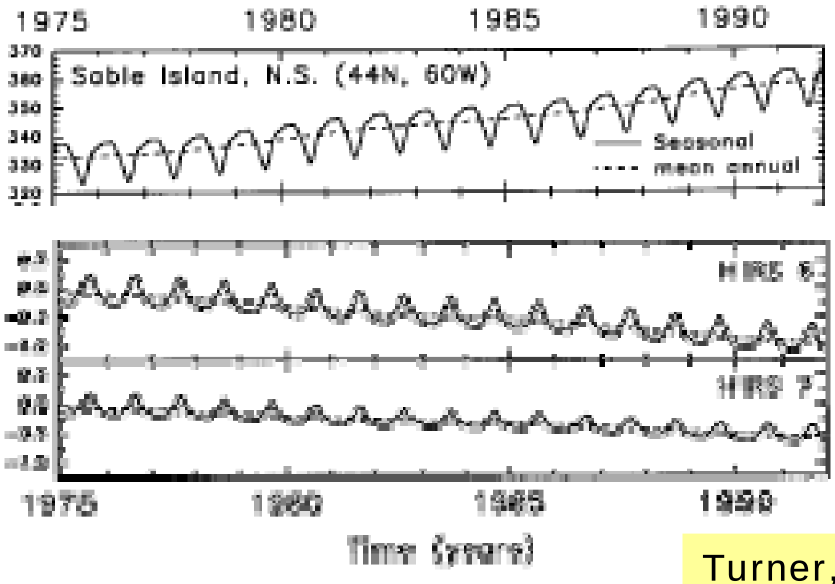
IASI Level 1C Spectra 29/11/2006,
13:42:11 UTC
Source CNES-CNRS ETHER

CO_2 as noise

- HIRS 15-micron channels

CO2 [ppm]

Tb(CO2 variable) –
Tb(CO2@330ppm) [K]



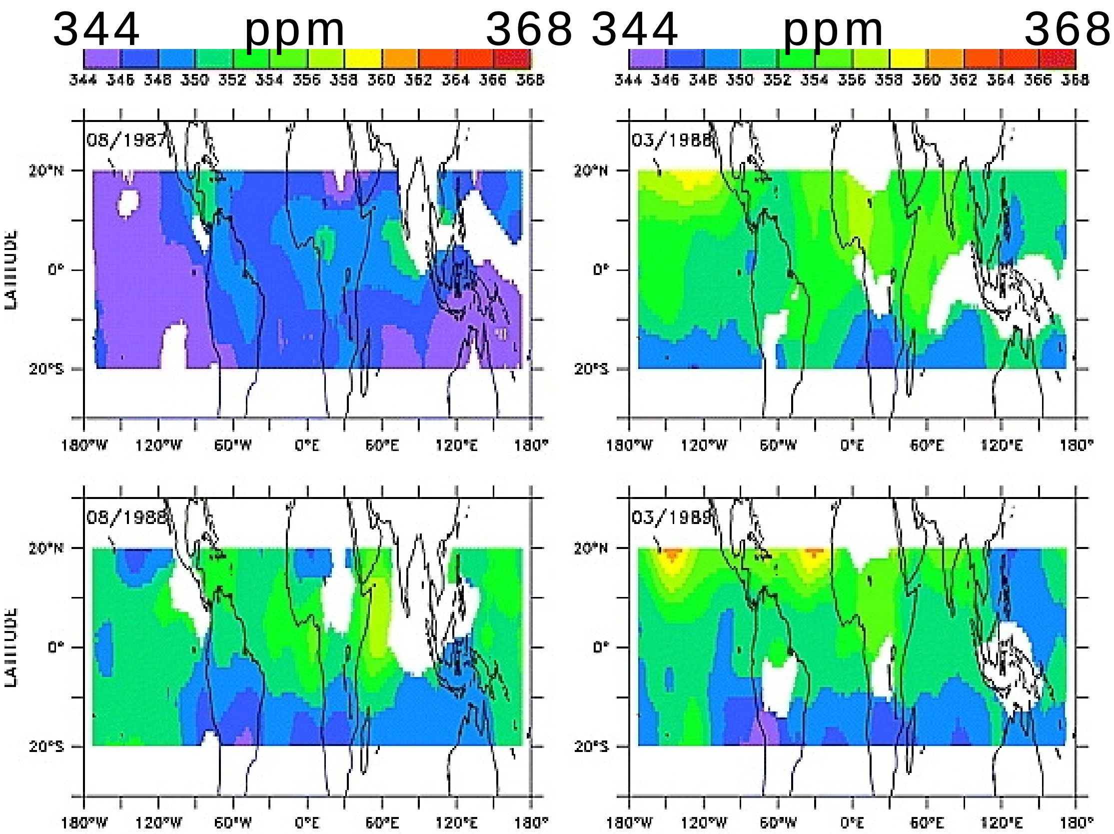
Turner, 1993

- “The brightness temperature differences can be as large as 1 K for a 30-ppmv CO_2 increase and a seasonal variation of a few tenths of a Kelvin may exist” (Turner, 1993).
- Impact on temperature retrievals

CO_2 as signal

- First retrievals of CO_2 concentrations (Chédin et al., 2003).
- Upper troposphere in tropical latitudes
- HIRS+MSU
- 1987-1991
- Neural network

- Model peak to peak twice as less



Chédin et al., 2003

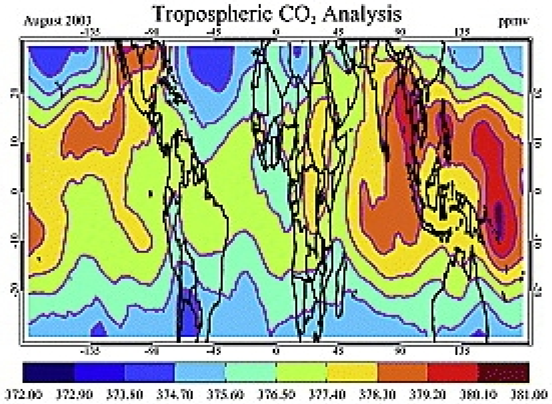
CO_2 as signal (cont'ed)

- Extension to high-spectral resolution measurements
- AIRS
- Upper troposphere in tropical latitudes

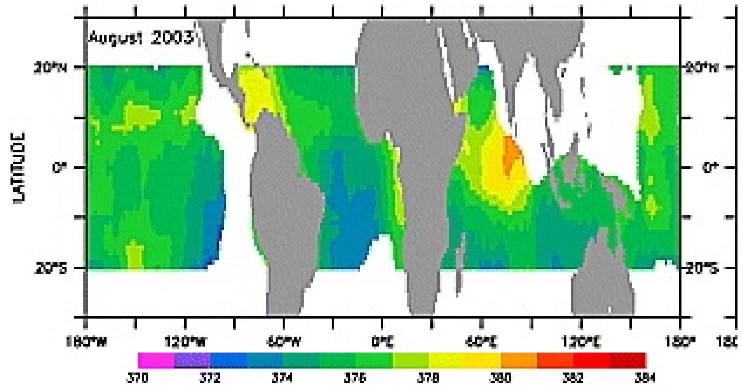
Engelen et al., 2004
1D-Var

August 2003

Crevoisier et al., 2004
Neural network



372 ppm 381

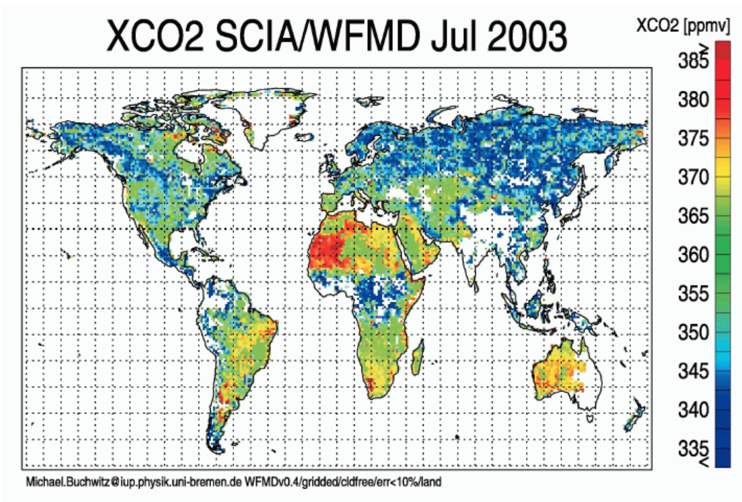


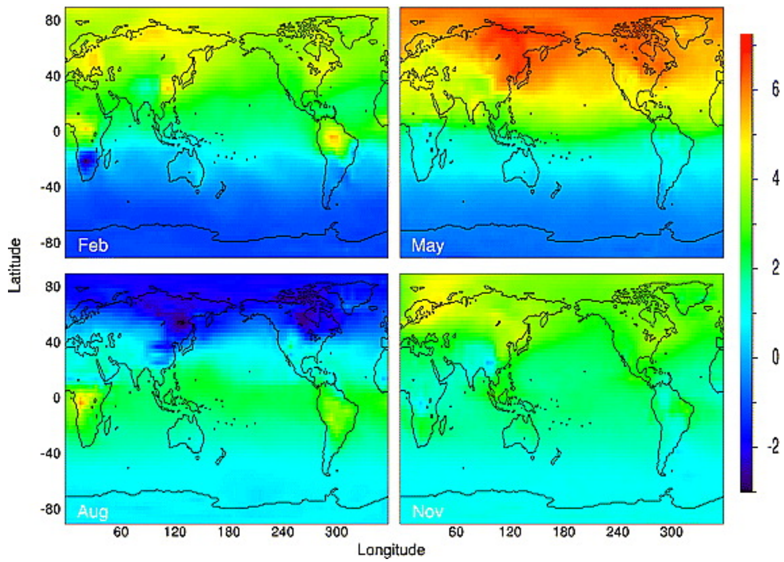
370 ppm 384

CO_2 as signal (cont'ed)

- Extension to high-spectral resolution measurements
- SCIAMACHY
- Total column over lands

Buchwitz et al., 2005
DOAS



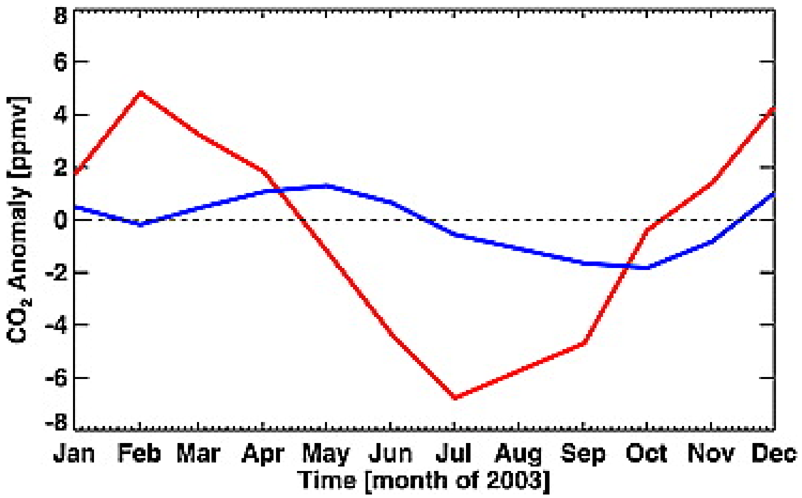


Model simulations of CO_2 column [ppm]
Olsen and Randerson, 2004

The same planet?

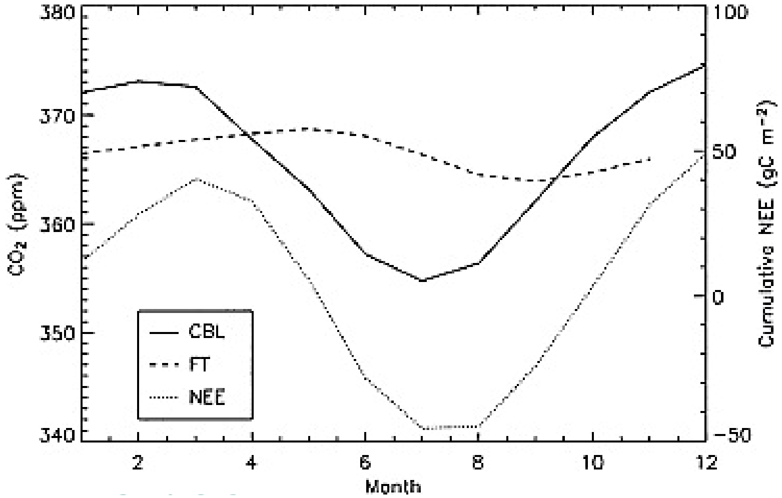
- Total column vs. upper tropospheric column
- Retrieved vs. measured

Barkley et al., 2006
SCIAMACHY (red) and
AIRS (blue) over North America



Chevallier

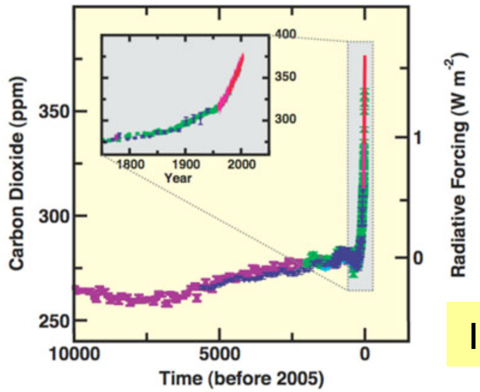
In situ measurements of CO_2 mixing ratios in the free troposphere (FT) and in the continental boundary layer (CBL), and of NEE (Hurwitz et al. 2004).



CO_2 as primary target

- CO_2 concentrations higher than at any time within the last 650,000 years
- OCO (NASA)
  - Launch Dec' 2008
- GOSAT (JAXA, NIES, MoE)
  - Launch Dec' 2008
- More projects
  - A-SCOPE, ACCLAIM, CARBOSAT, ...

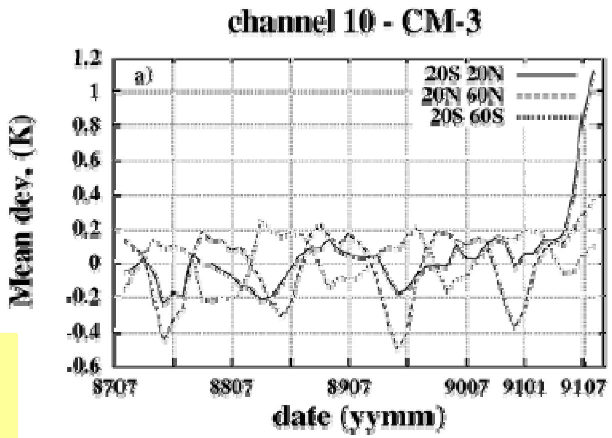
Changes in Greenhouse Gases
from ice-Core and Modern Data



IPCC (2007)

Beyond NWP within NWP systems

- Errors on atmospheric concentrations of gases and aerosols affect NWP systems



NOAA-10 channel 10
Calculated minus obs
Pierangelo et al., 2004

- NWP systems flexible and powerful enough to tackle environmental issues
  - Expertise in data merging
  - Expertise in satellite observations
  - Expertise in atmospheric modelling

- Surface and soil properties

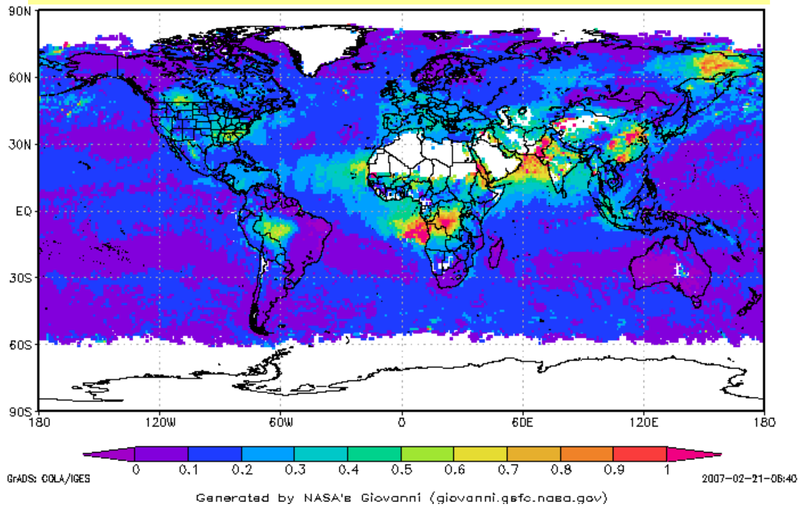
FP6 GEMS project

- Part of the Global Monitoring for Environment and Security (GMES, funded by EC & ESA) Atmosphere theme
- 31 consortium members, 4 years (started in March 2005)
- Coordinated by ECMWF
- Creation of a pre-operational global monitoring system for greenhouse gases, reactive gases, and aerosols in the troposphere and in the stratosphere
- Near-real-time and retrospective global analyses for monitoring atmospheric composition, and short-range forecasts to drive regional air-quality models.

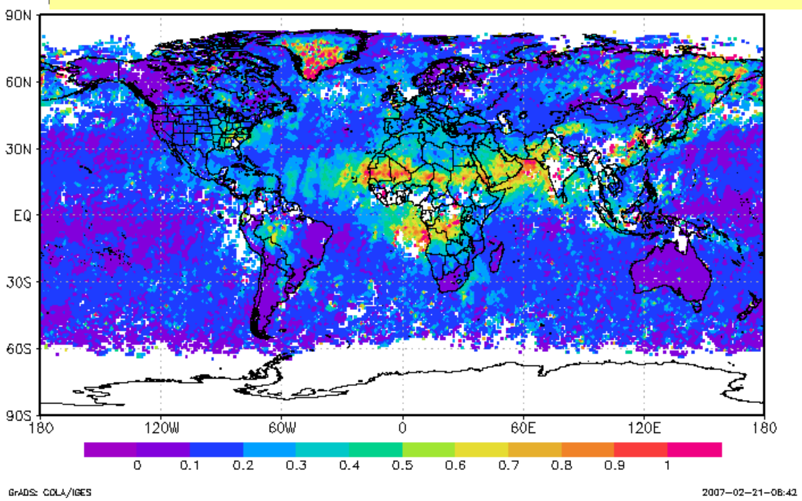
Summary of introduction

- Use of satellite data for the environment = emerging topic
- Increasing interest from the NWP community
- Signal-to-noise may be challenging for some species

MODIS/Terra, August 2003, daytime
Aerosol optical depth @0.55 micron



MISR/Terra, August 2003
Aerosol optical depth @0.55 micron



A. Benedetti

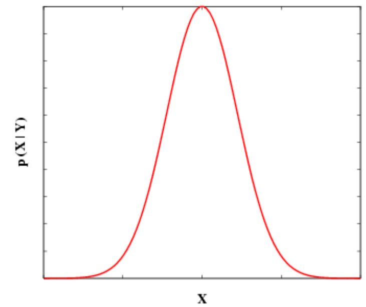
Outline

- Introduction
- Optimization of atmospheric concentrations
- Optimization of surface fluxes
- Optimization of surface model parameters
- Conclusion

Mathematical framework

- Bayes' theorem

$$P(x|y) = \frac{P(y|x).P(x)}{P(y)}$$



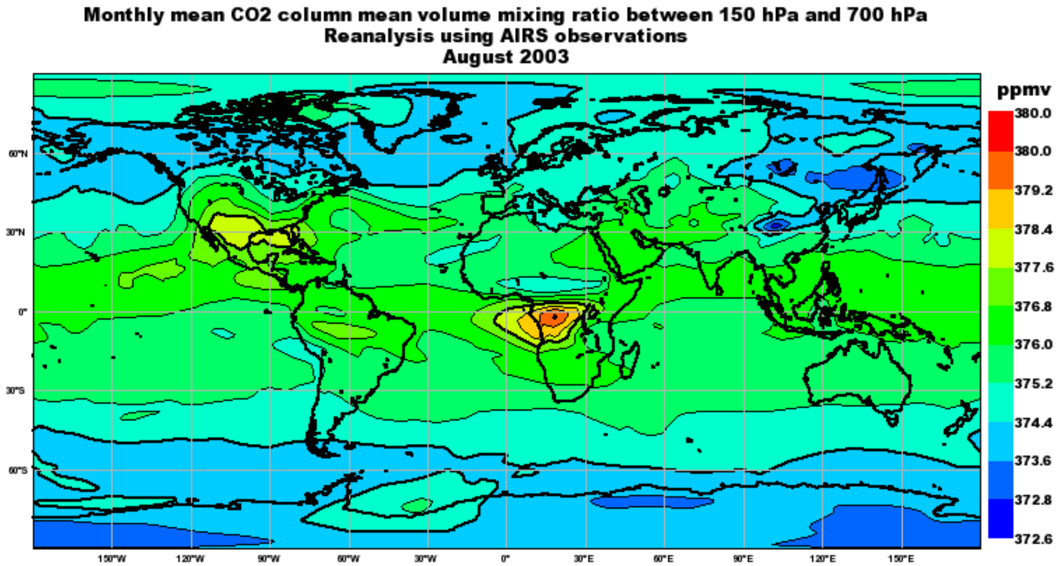
- The optimal solution minimizes the following cost function

$$-2\, lnP(\mathbf{x}|\mathbf{y}) = (\mathbf{x}-\mathbf{x}_b)^{\mathbf{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + (\mathbf{Hx}-\mathbf{y})^{\mathbf{T}}\mathbf{R}^{-1}(\mathbf{Hx}-\mathbf{y})$$

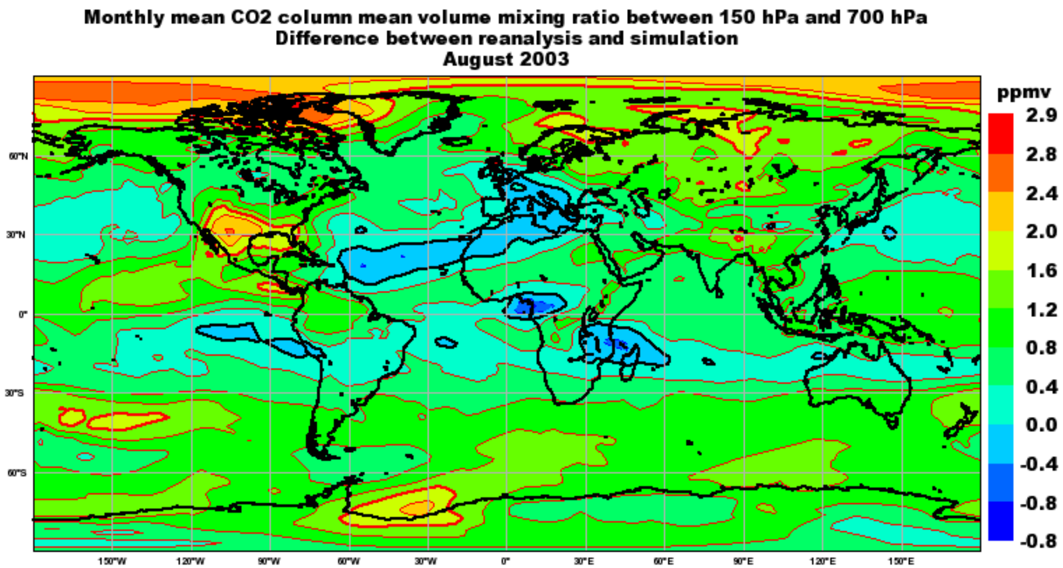
x: state vector (atmospheric concentrations)
y = Hx+ε : observation (atmospheric concentrations)
H: linear observation operator
(short-range chemical transport + interpolation)
B: background error covariance matrix
R: observation error covariance matrix

CO_2 analysis

CO_2 4D-Var analysis using AIRS
August 2003
Started in January 2003



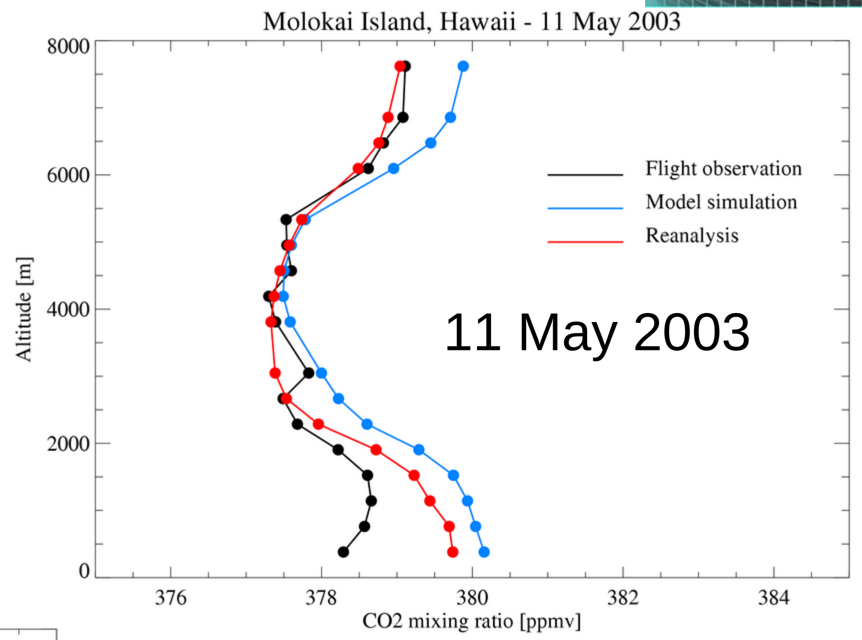
Analysis minus free run
August 2003



R. Engelen

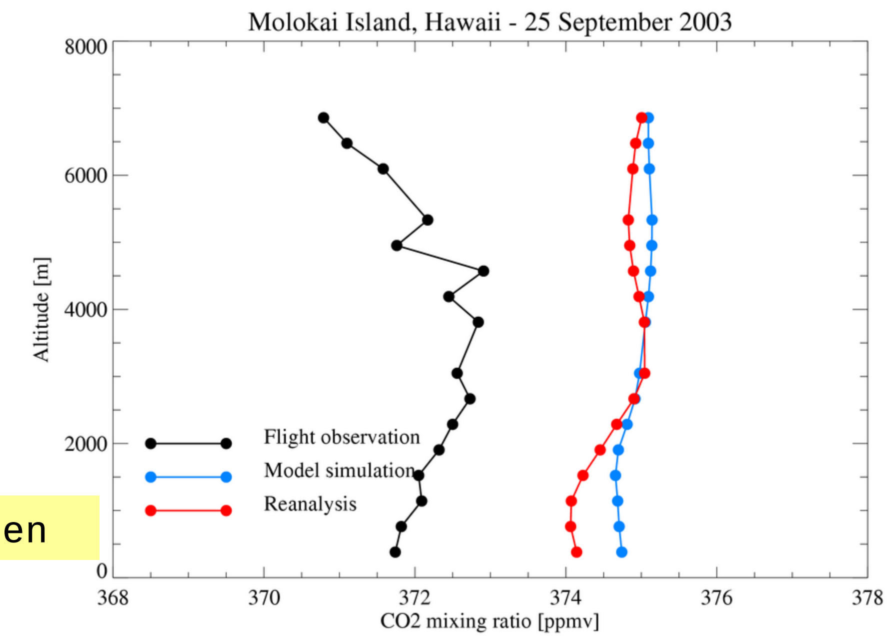
Individual Profiles

Comparisons with flight data over Hawaii (courtesy of Pieter Tans, NOAA/ESRL) shows a clear improvement of the analysis over the free-running model.
But, background error is important.



11 May 2003

10 ppmv

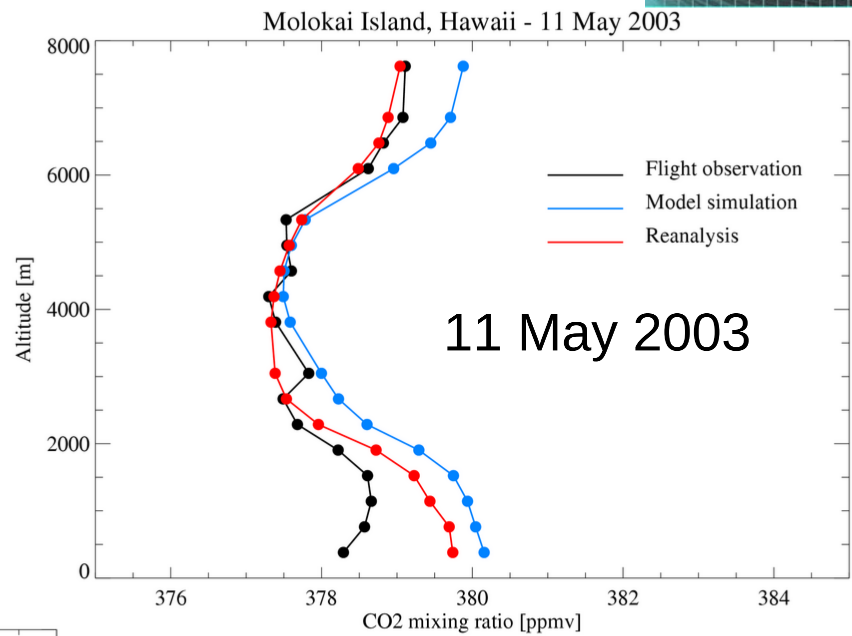


25 September 2003

R. Engelen

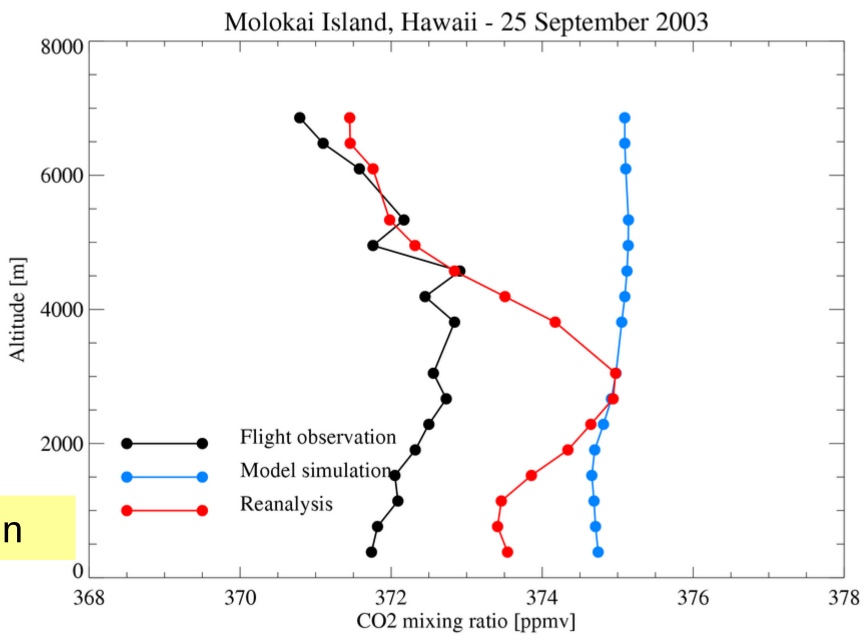
Individual Profiles

Comparisons with flight data over Hawaii (courtesy of Pieter Tans, NOAA/ESRL) shows a clear improvement of the analysis over the free-running model.
But, background error is important.



11 May 2003

10 ppmv

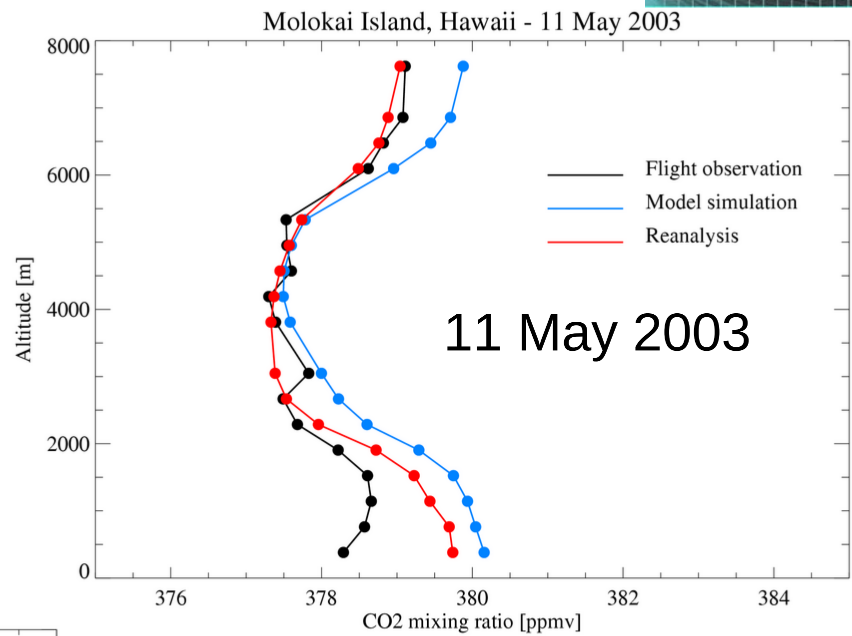


R. Engelen

25 September 2003

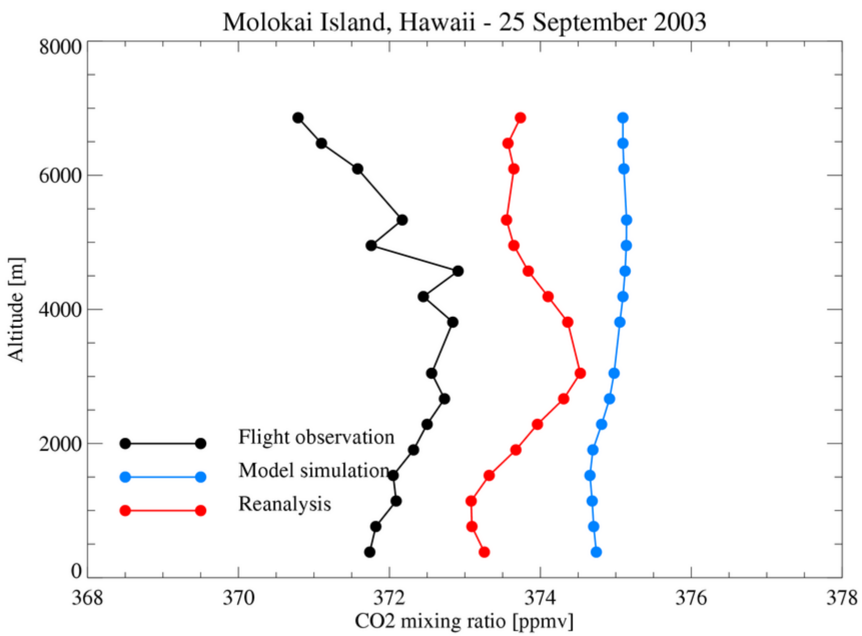
Individual Profiles

Comparisons with flight data over Hawaii (courtesy of Pieter Tans, NOAA/ESRL) shows a clear improvement of the analysis over the free-running model.
But, background error is important.



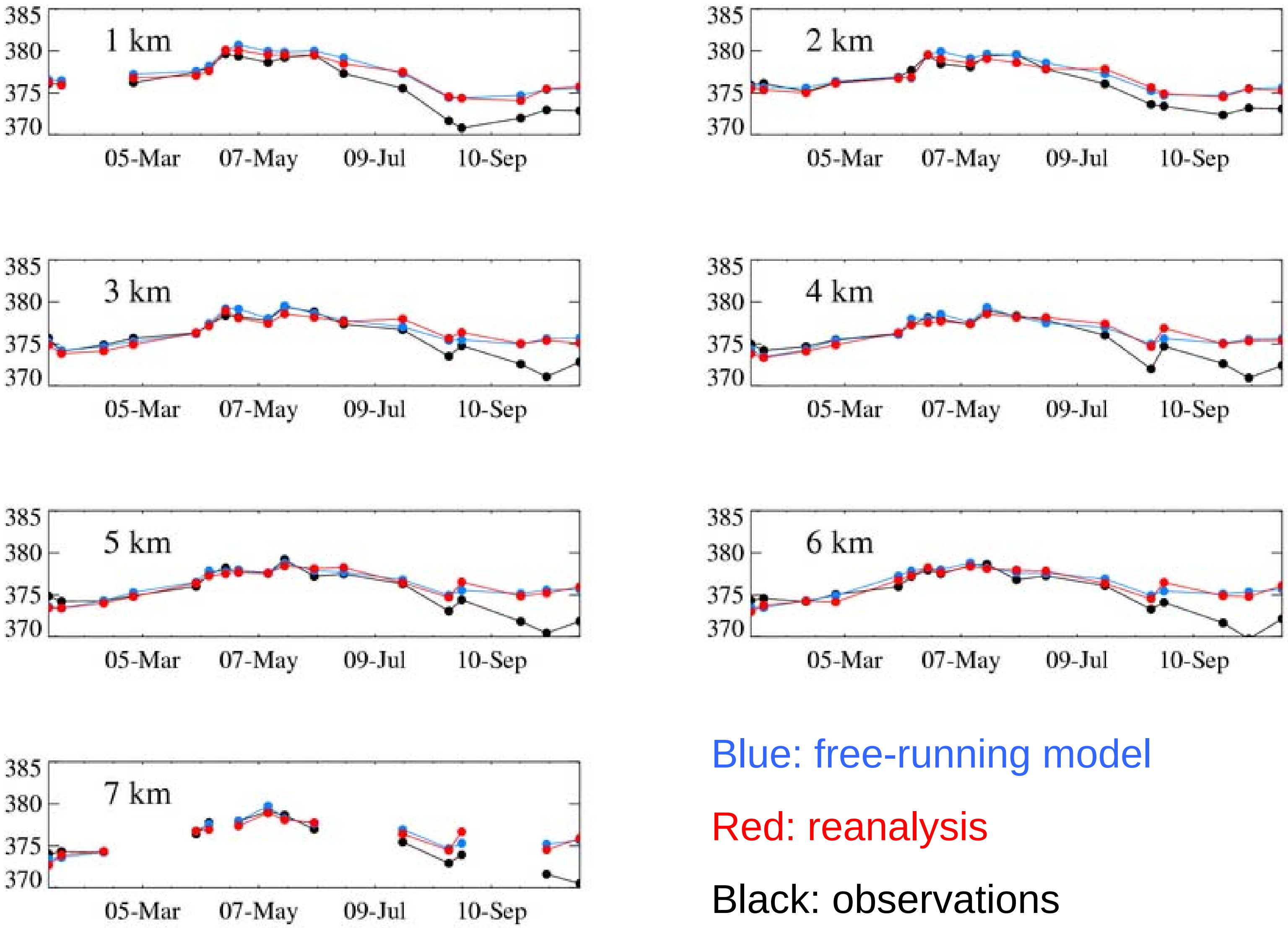
11 May 2003

10 ppmv



25 September 2003

Molokai, Island, Hawaii



Blue: free-running model

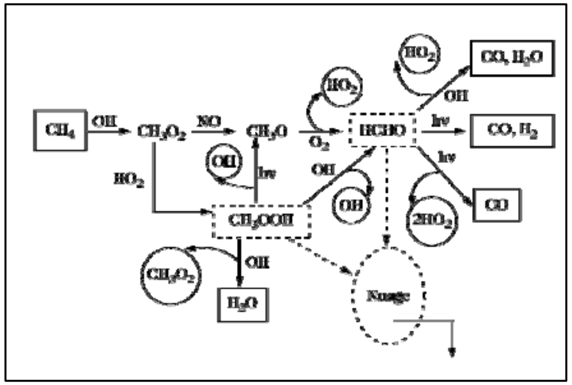
Red: reanalysis

Black: observations

R. Engelen

ECMWF

Analysis of CO

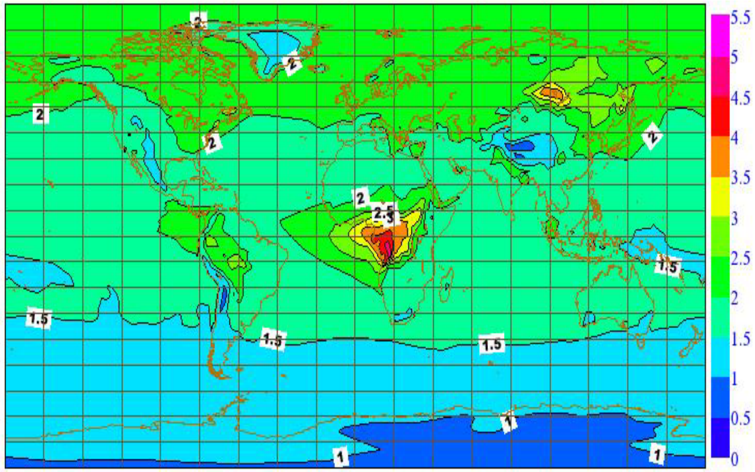


- CO_2 lifetime ~ 100 years
- CO lifetime ~ 2 months
- CO interacts with OH
  - Surface sources (combustion)
  - Chemical production in the atmosphere
  - Chemical loss in the atmosphere
- Observed by MOPITT satellite since 2000

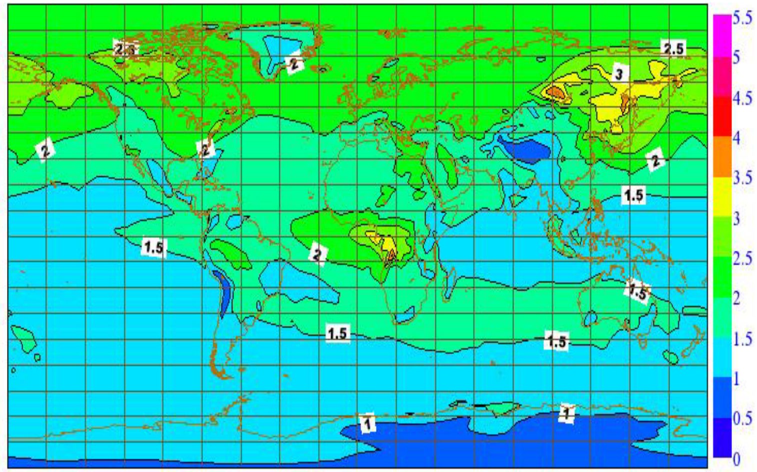
- GEMS analysis system : 2-way coupling between IFS and a chemistry-transport model

Analysis of CO

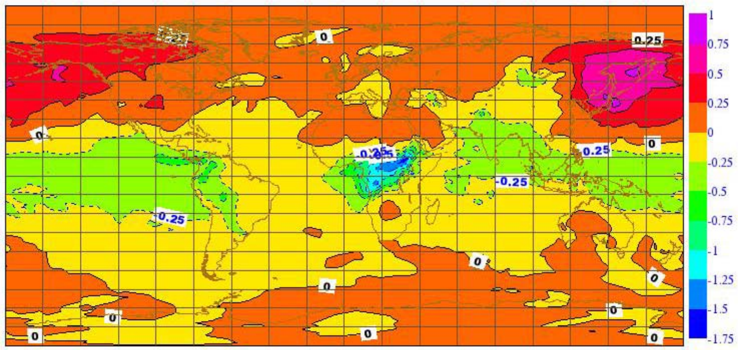
Free running



Assimilation of MOPITT data



Assimilation minus free run

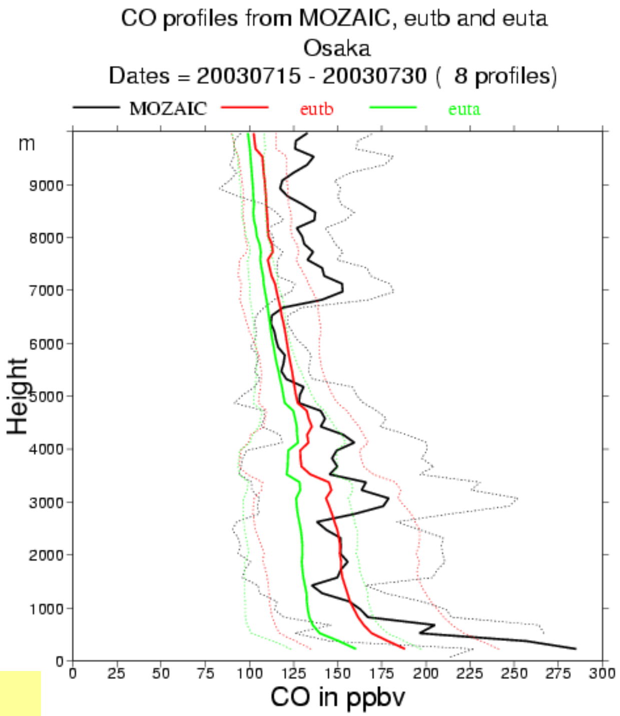


15-30 July 2003
10^{18} modelcules/cm^2

A. Dethof

Analysis of CO (cont’ed)

- Assimilation of MOPITT CO columns leads to improved fit to profile observations from MOZAIC flights



MOZAIC flight

Free-running CO model

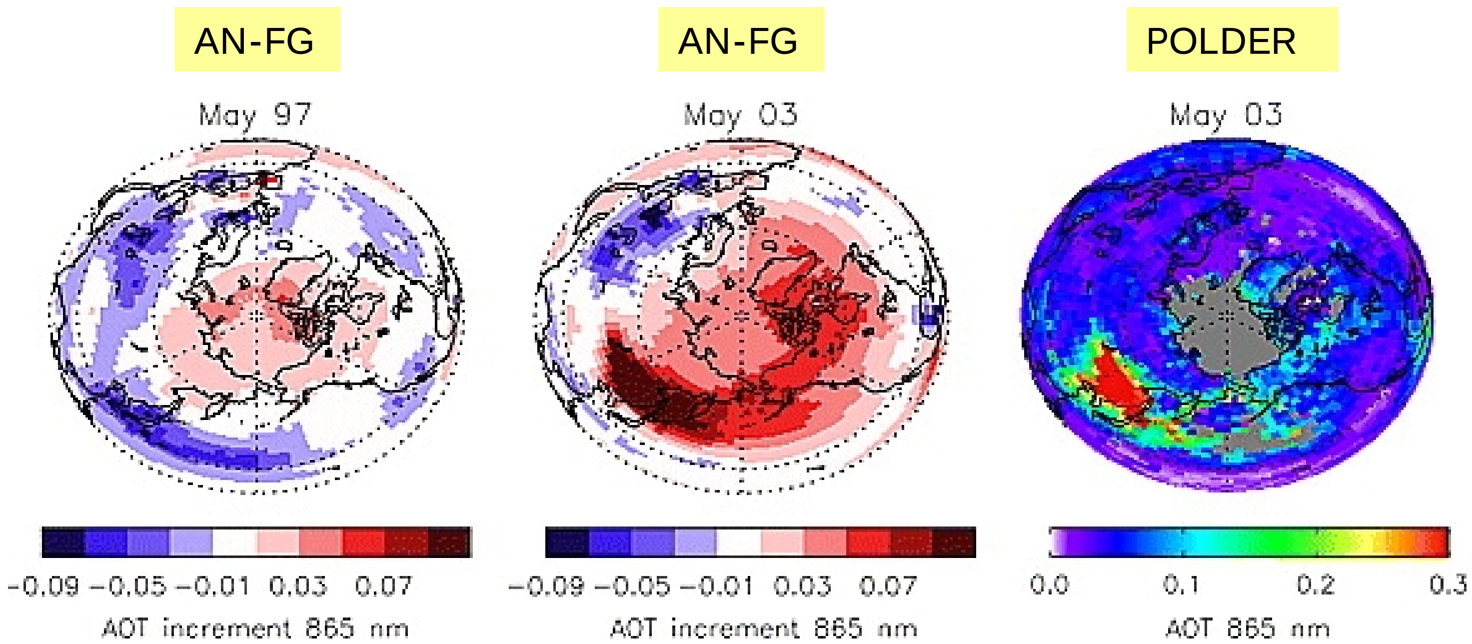
Assimilation of MOPITT data

15-30 July 2003
MOSAIC data over Osaka, JPN

A. Dethof

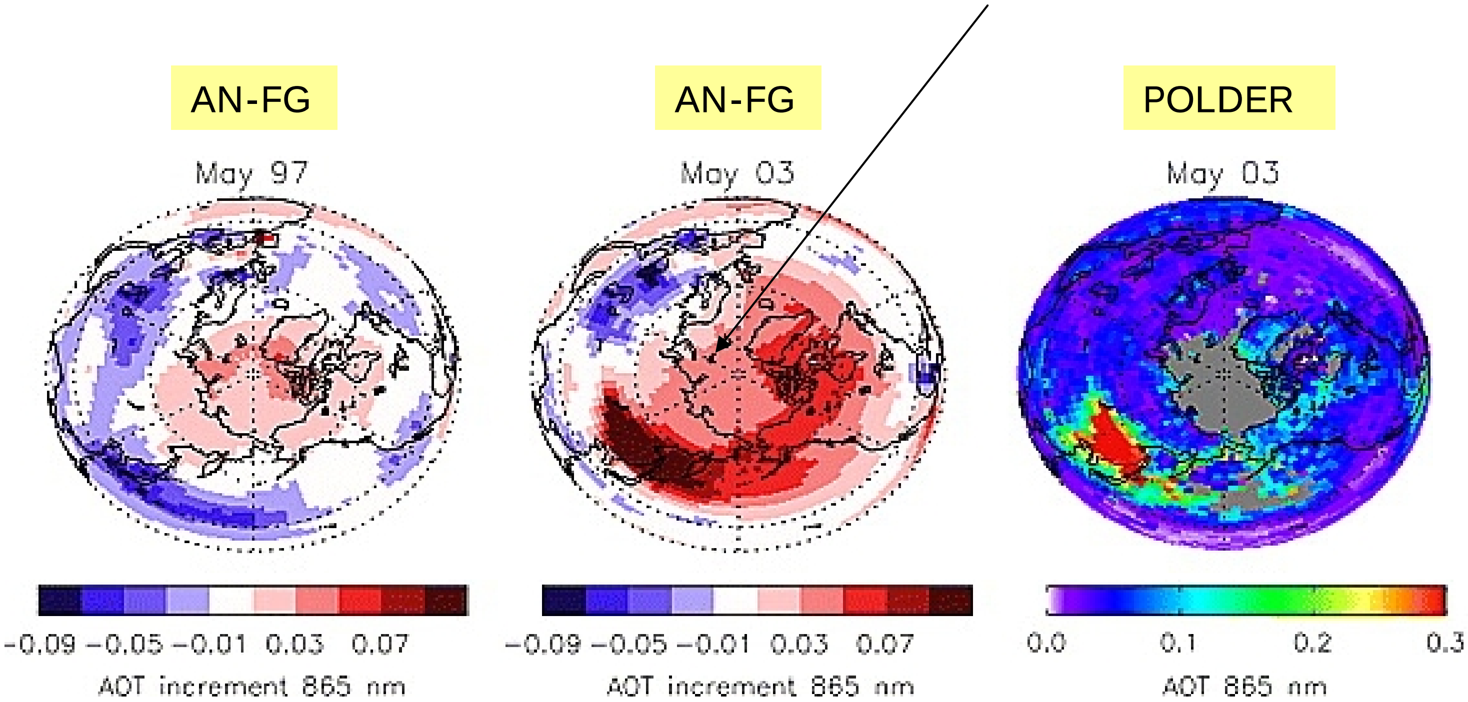
Assimilation of POLDER data within the LMDZ-INCA model

○ Aerosol optical thickness



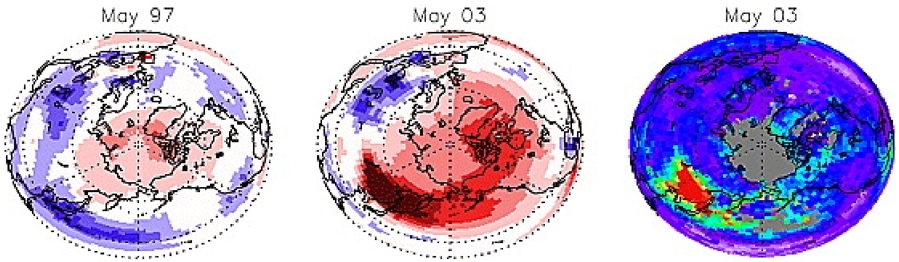
Generoso et al. (2007)

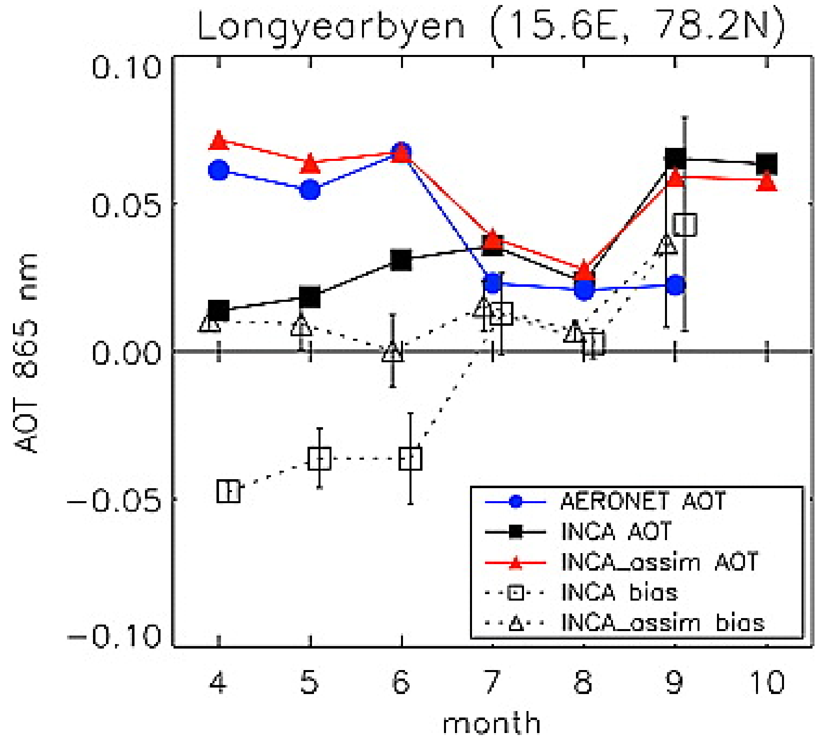
Assimilation of POLDER data within the LMDZ-INCA model



Generoso et al. (2007)

Assimilation of POLDER data within the LMDZ-INCA model





Generoso et al. (2007)

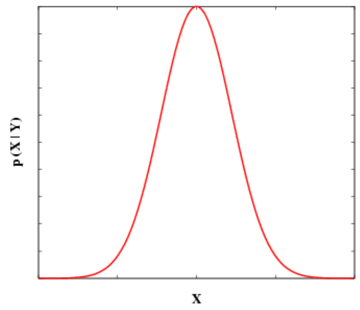
Outline

- Introduction
- Optimization of atmospheric concentrations
- Optimization of surface fluxes
- Optimization of surface model parameters
- Conclusion

Mathematical framework

- Bayes' theorem

$$P(x|y) = \frac{P(y|x).P(x)}{P(y)}$$



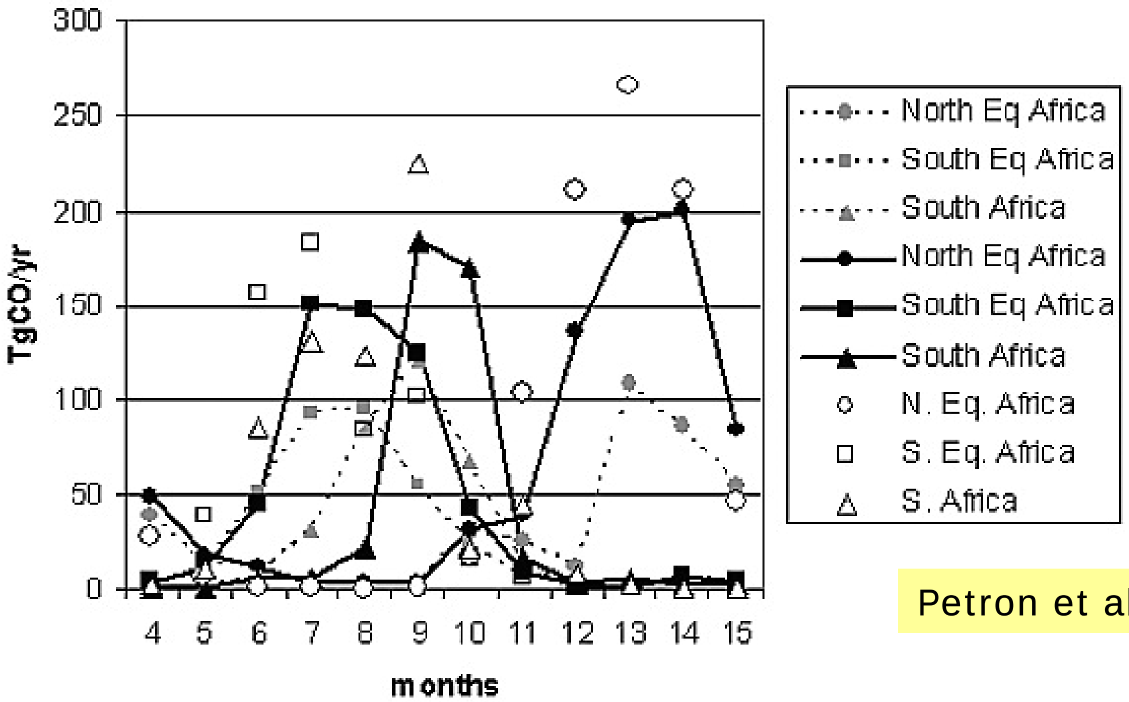
- The optimal solution minimizes the following cost function

$$-2\, lnP(\mathbf{x}|\mathbf{y}) = (\mathbf{x}-\mathbf{x}_b)^{\mathbf{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + (\mathbf{Hx}-\mathbf{y})^{\mathbf{T}}\mathbf{R}^{-1}(\mathbf{Hx}-\mathbf{y})$$

x: state vector (surface fluxes)
y = Hx+ε : observation (atmospheric concentrations)
H: linear observation operator
(long-range chemical transport + interpolation)
B: background error covariance matrix
R: observation error covariance matrix

CO emissions from MOPITT

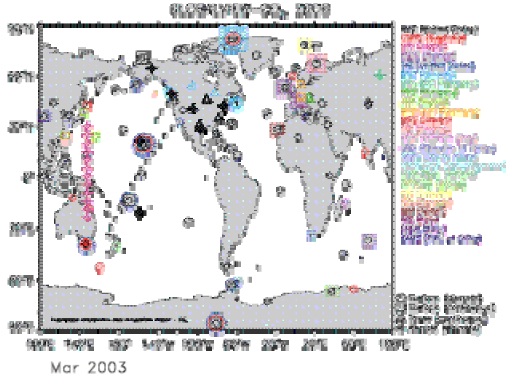
- Comparison of a priori (grey symbols) and a posteriori (black symbols) monthly biomass burning sources in Africa with van der Werf et al. (2004) inventory (white symbols)

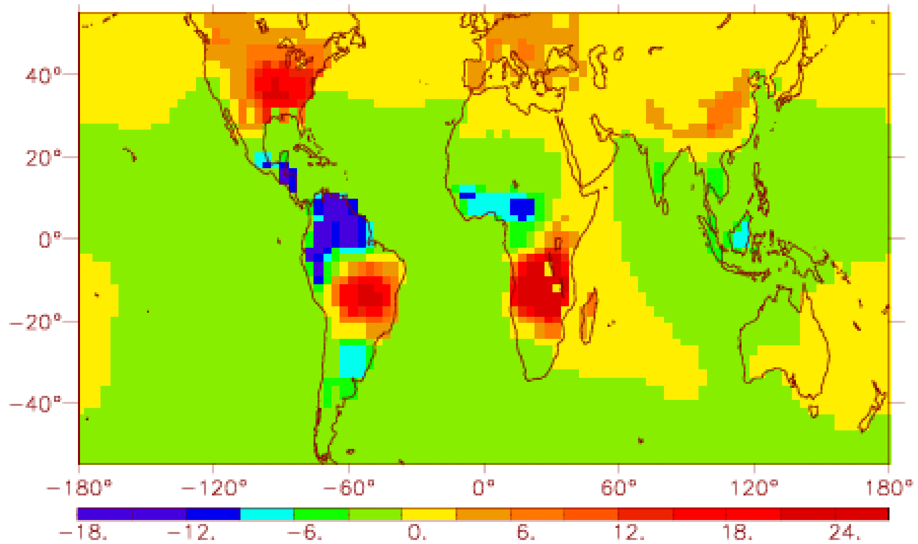


Petron et al. (2004)

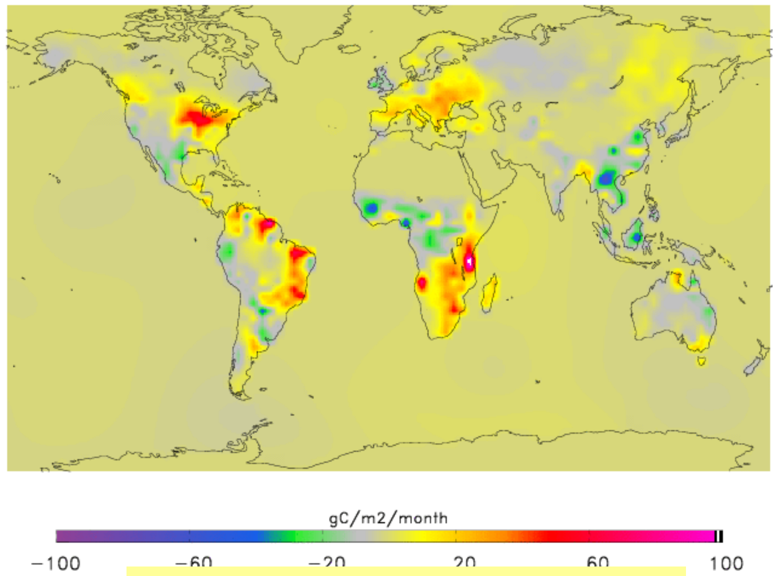
CO_2 fluxes from AIRS

- 11-month inversion
- March 2003, GEMS test re-analysis
- AN-FG, $gC/m^2/month$





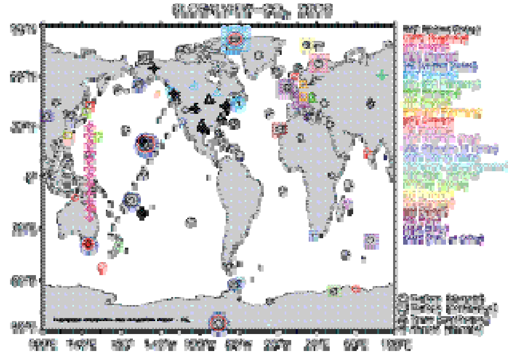
Inversion from satellite data

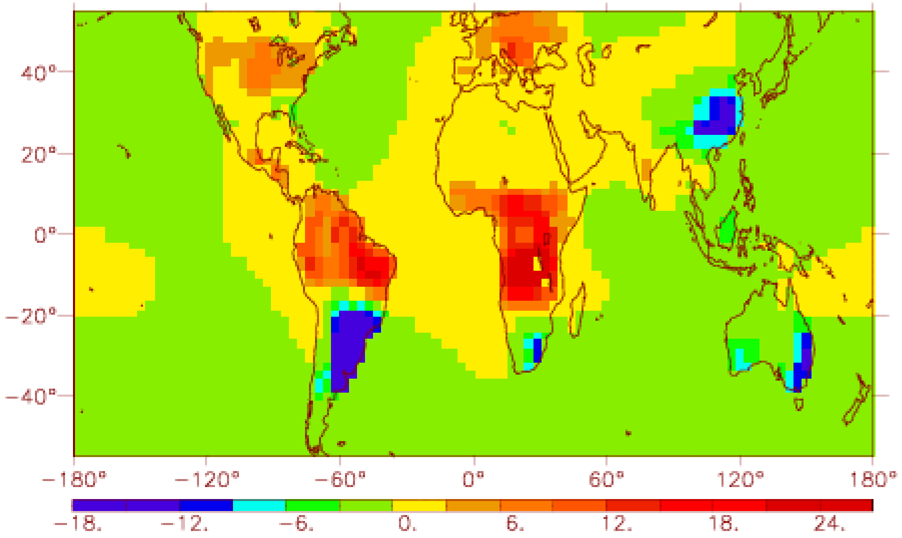


Inversion from in situ data (P. Peylin, pers. comm.)

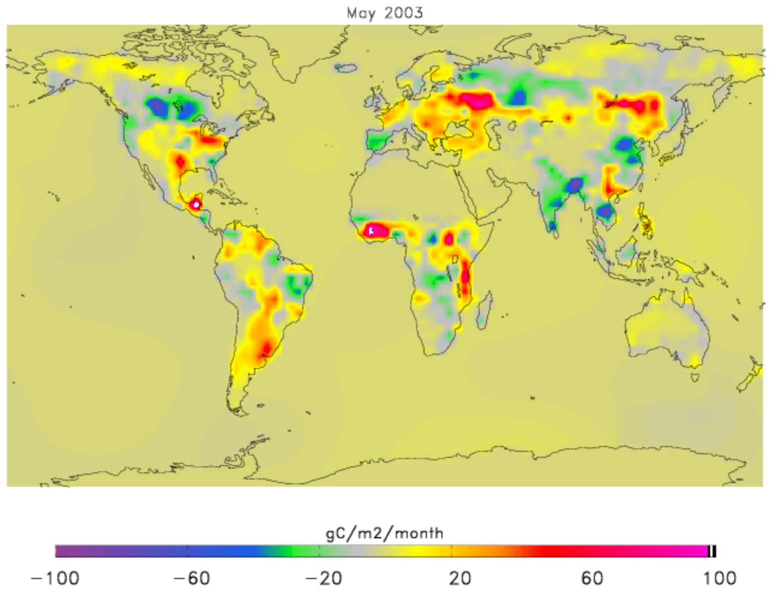
CO_2 fluxes from AIRS

- 11-month inversion
- May 2003 , GEMS test re-analysis
- AN-FG, $gC/m^2/month$





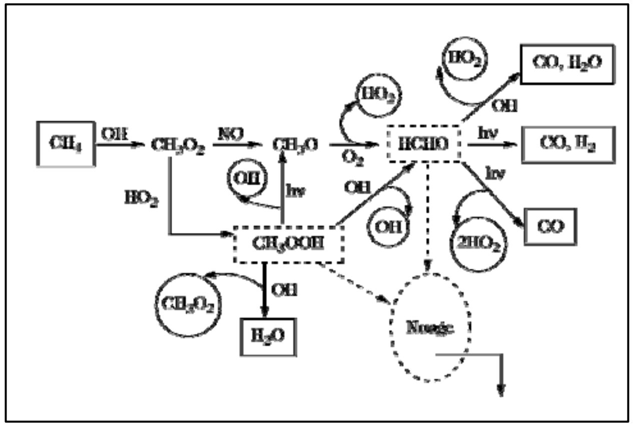
Inversion from satellite data

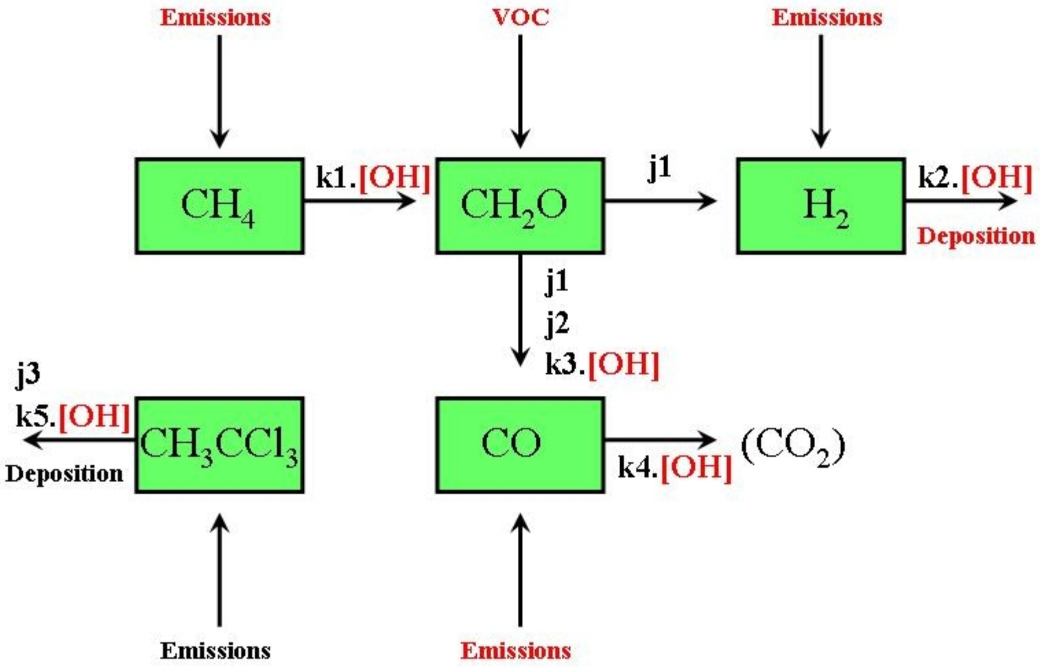


Inversion from in situ data (P. Peylin, pers. comm.)

Multi-tracer inversions

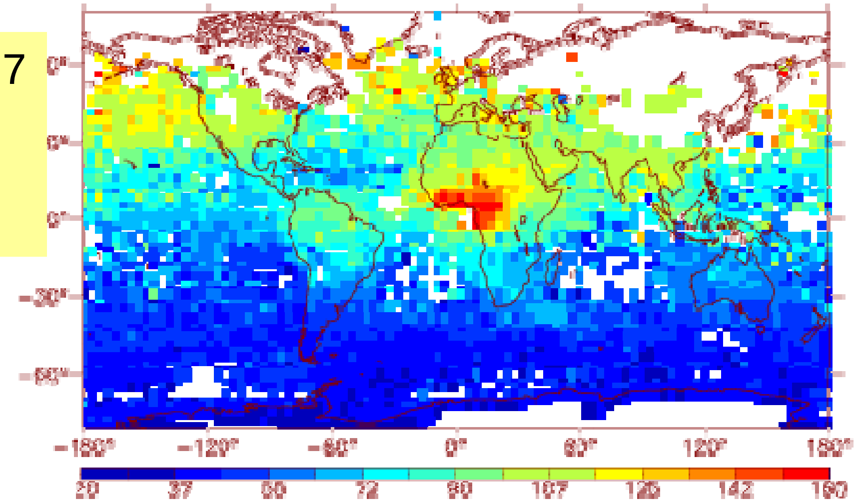
- Simplified atmospheric chemistry
  - Computing time
  - Limited observation information content
- Hydrocarbon oxidation chain

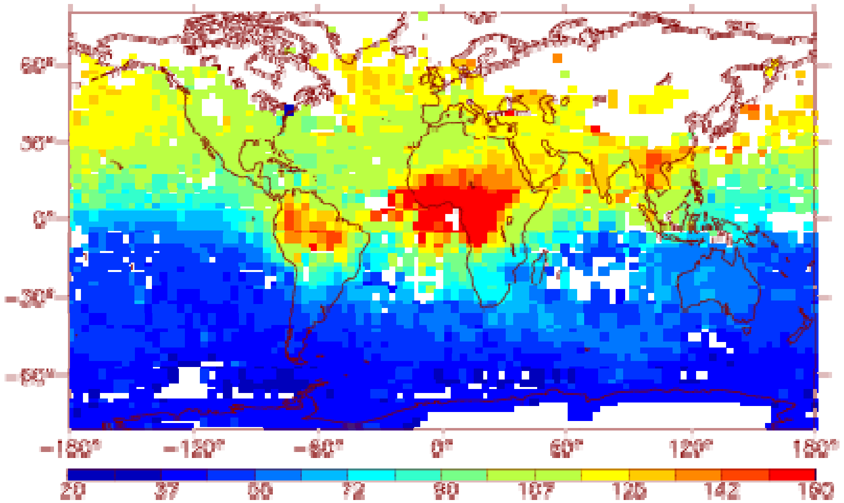




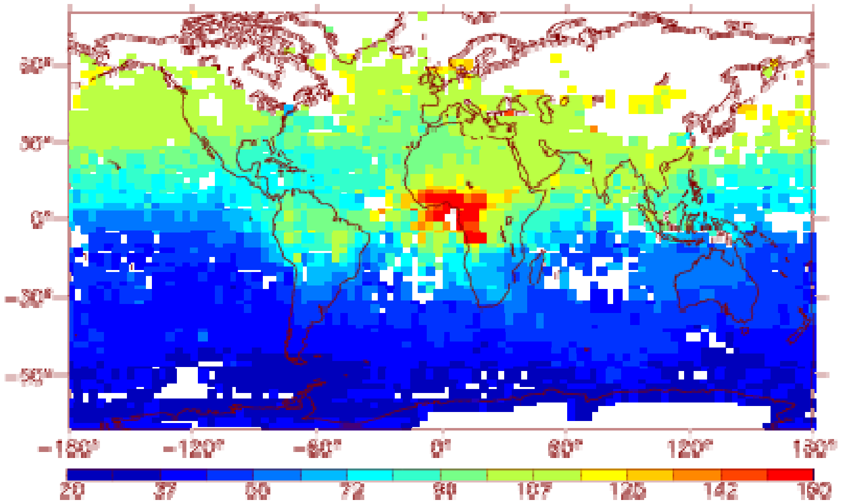
Multi-tracer inversion from MOPITT+ surface MCF

Mean January 2007
MOPITT L2V5.9.4
700 hPa (ppb)
31,400 retrievals



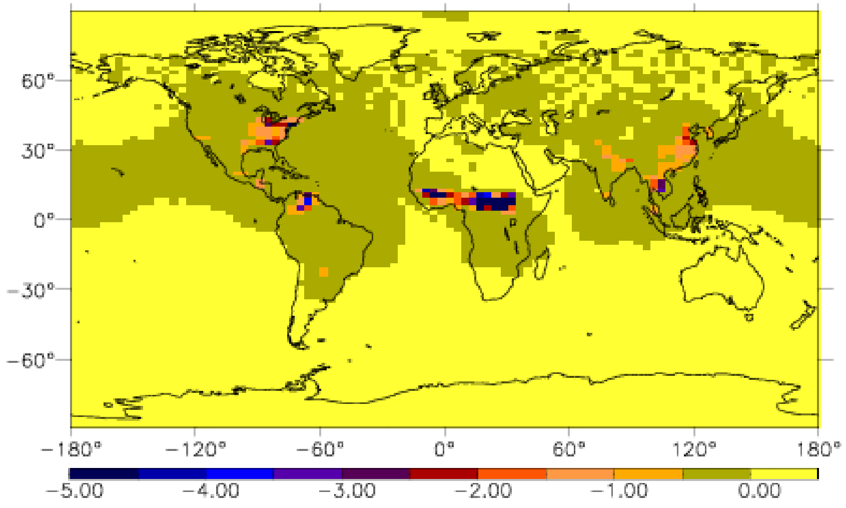


First-Guess equivalent

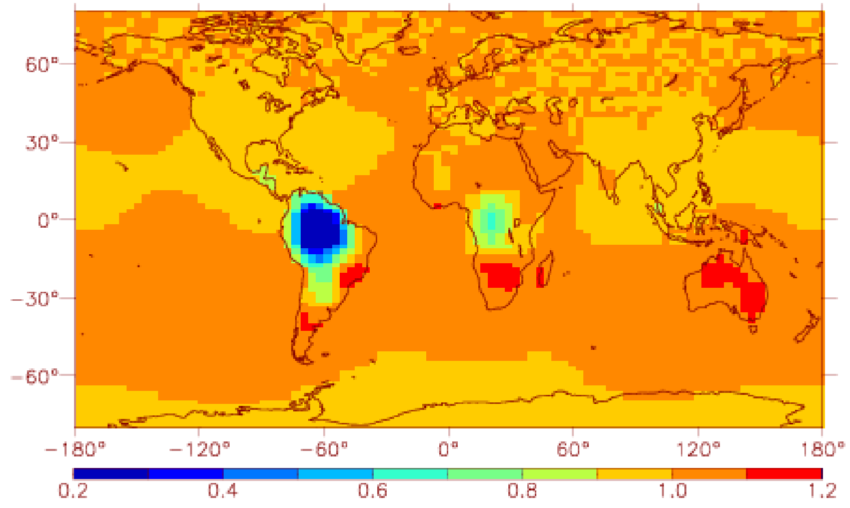


Analysis equivalent

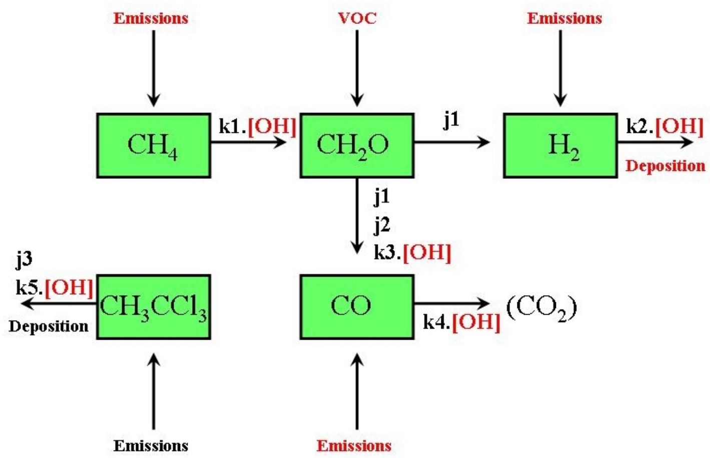
Multi-tracer inversion from MOPITT+ surface MCF



CO flux increments (g/m2 for 1 month)

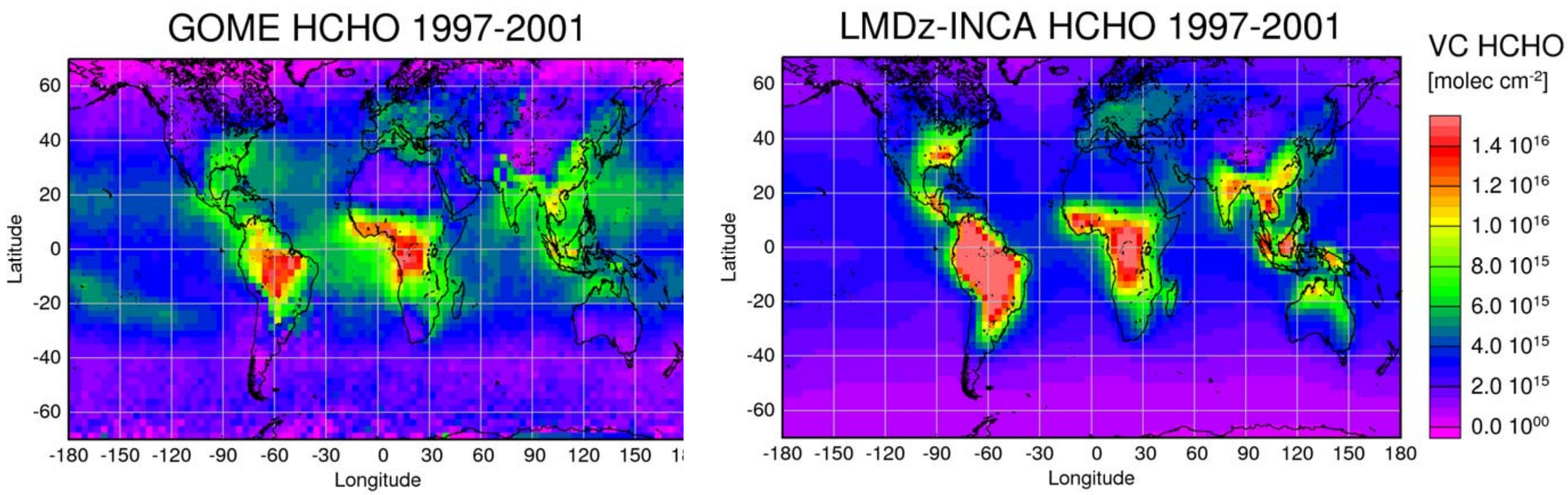


HCHO-production scaling factor (0-1)



Independent HCHO observation

- Too much HCHO in the free model



F. Wittrock, 2006
Bremen University

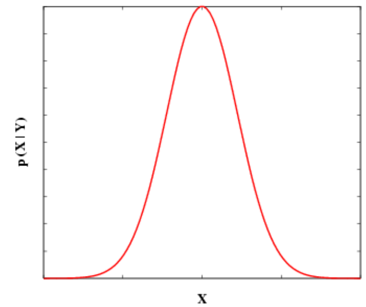
Outline

- Introduction
- Optimization of atmospheric concentrations
- Optimization of surface fluxes
- Optimization of surface model parameters
- Conclusion

Mathematical framework

- Bayes' theorem

$$P(x|y) = \frac{P(y|x).P(x)}{P(y)}$$



- The optimal solution minimizes the following cost function

$$-2\, lnP(\mathbf{x}|\mathbf{y}) = (\mathbf{x}-\mathbf{x}_b)^{\mathbf{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + (\mathbf{H}\mathbf{x}-\mathbf{y})^{\mathbf{T}}\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}-\mathbf{y})$$

x: **state vector (model parameters)**
y = Hx+ε : **observation (atmospheric concentrations + …)**
H: **linear observation operator**
(surface model +long-range chemical transport + interpolation)
B: **background error covariance matrix**
R: **observation error covariance matrix**

Con: involving development

- More complex observation operator

$$\nabla J(\mathbf{x}) = 2\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x_b}) + 2\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{y} - H[\mathbf{x}])$$

- LMDZT transport model includes ~ a few thousands lines of code
- ORCHIDEE model of the terrestrial vegetation includes ~ 40,000 lines of code

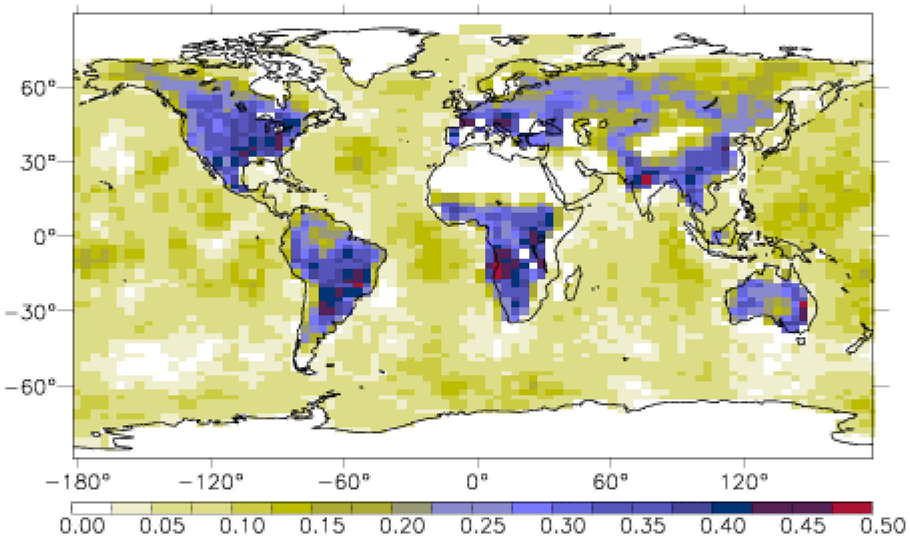
Con: model as hard constraint

- Changes the observation errors as seen by the inversion system
  - [Observation error] = [Measurement error]
    + [representativeness error]
    + [Model error]
- Biases / Variances / Correlations

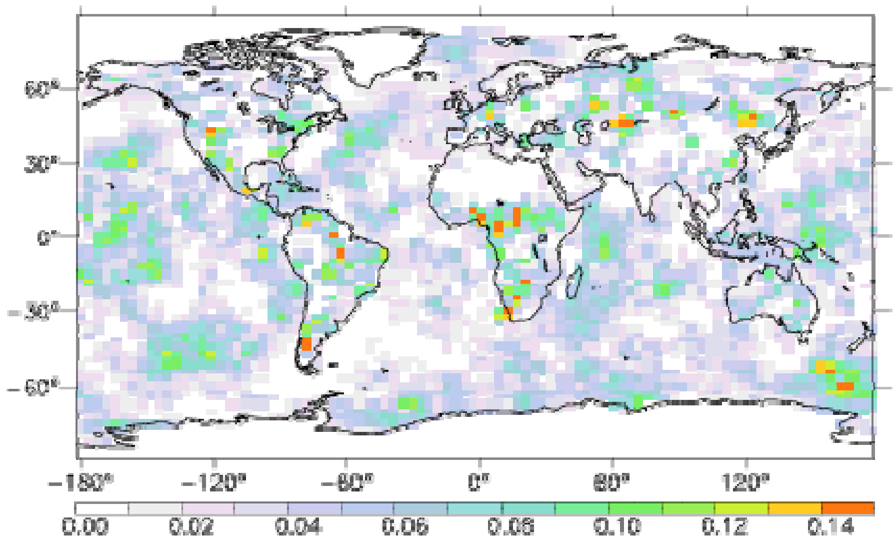
- We may not have enough information from the observations to introduce a weak constraint formulation

Impact of observations error correlations

- Surface fluxes from OCO
- Hypothetical O.5 along-track correlation
- Correlations ignored in the inversion
- Uncertainty reduction 1-sig(post)/sig(prior)



Reference error reduction

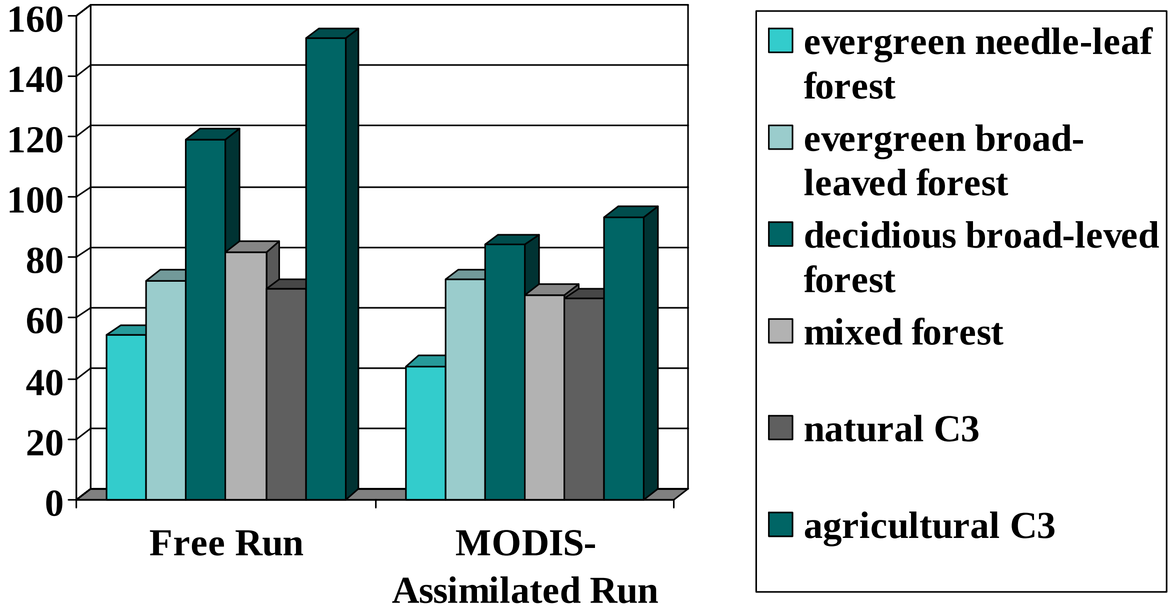


Increment in error reduction

Chevallier (2007)

Pro: assimilate more than atmospheric concentrations

- Assimilation of MODIS LAI within the ORCHIDEE vegetation model
- RMS difference between simulated gross primary production and independent FLUXNET data (40 sites)
- gC/m²/month



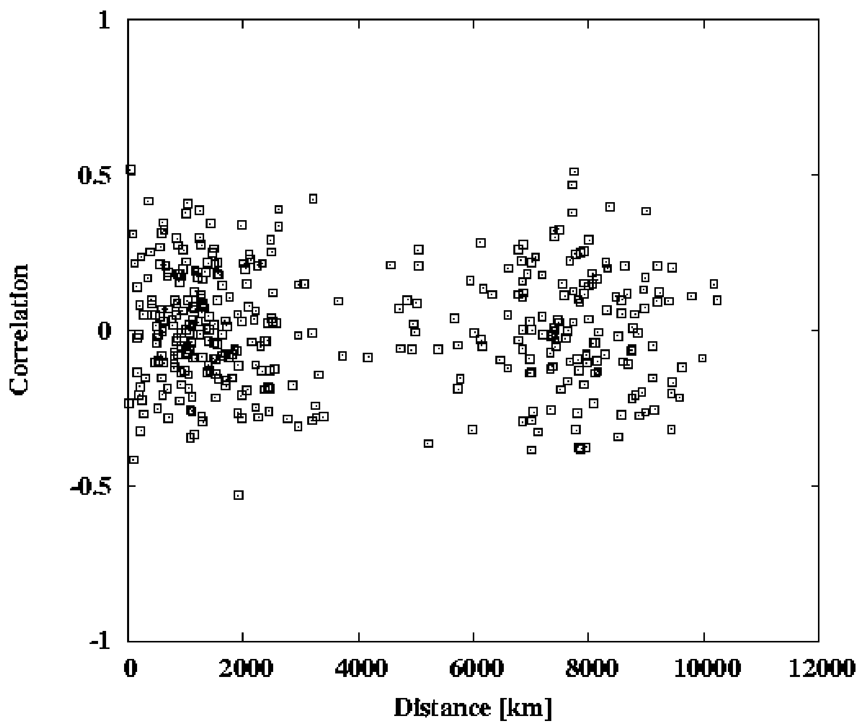
Demarty et al. (2007)

Pro: spread increments

- Background error correlations

$$-2\, lnP(\mathbf{x}|\mathbf{y}) = (\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + (\mathbf{Hx}-\mathbf{y})^T\mathbf{R}^{-1}(\mathbf{Hx}-\mathbf{y})$$

CO_2 fluxes:
Prior error spatial correlations
at a series of sites
Chevallier et al. (2006)



Pro: spread increments (cont'ed)

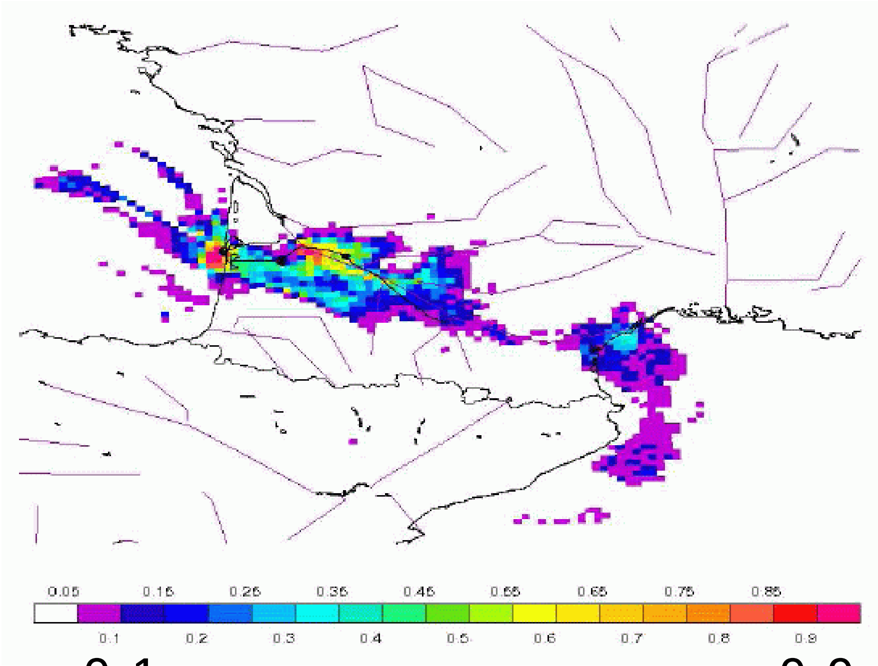
- Optimizing generic parameters may be more efficient than prior errors in spreading the observation information in space and time

Error reduction
for the inversion of CO_2 surface fluxes from CO_2 concentrations at two sites. 4-day period.

No prior error spatial correlations.

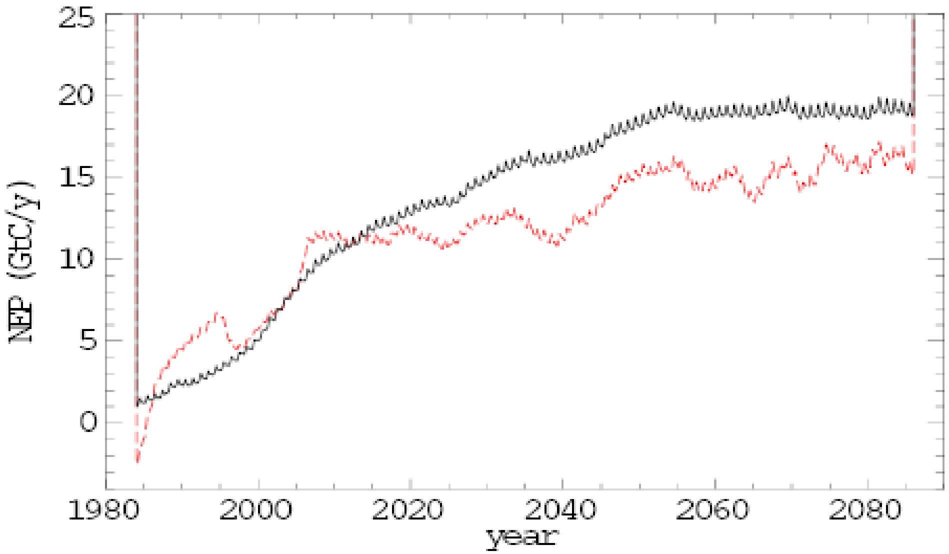
Transport from MesoNH at 8-km resolution.

Lauvaux et al. (2007)



0.1 0.9

Pro: predictive capability



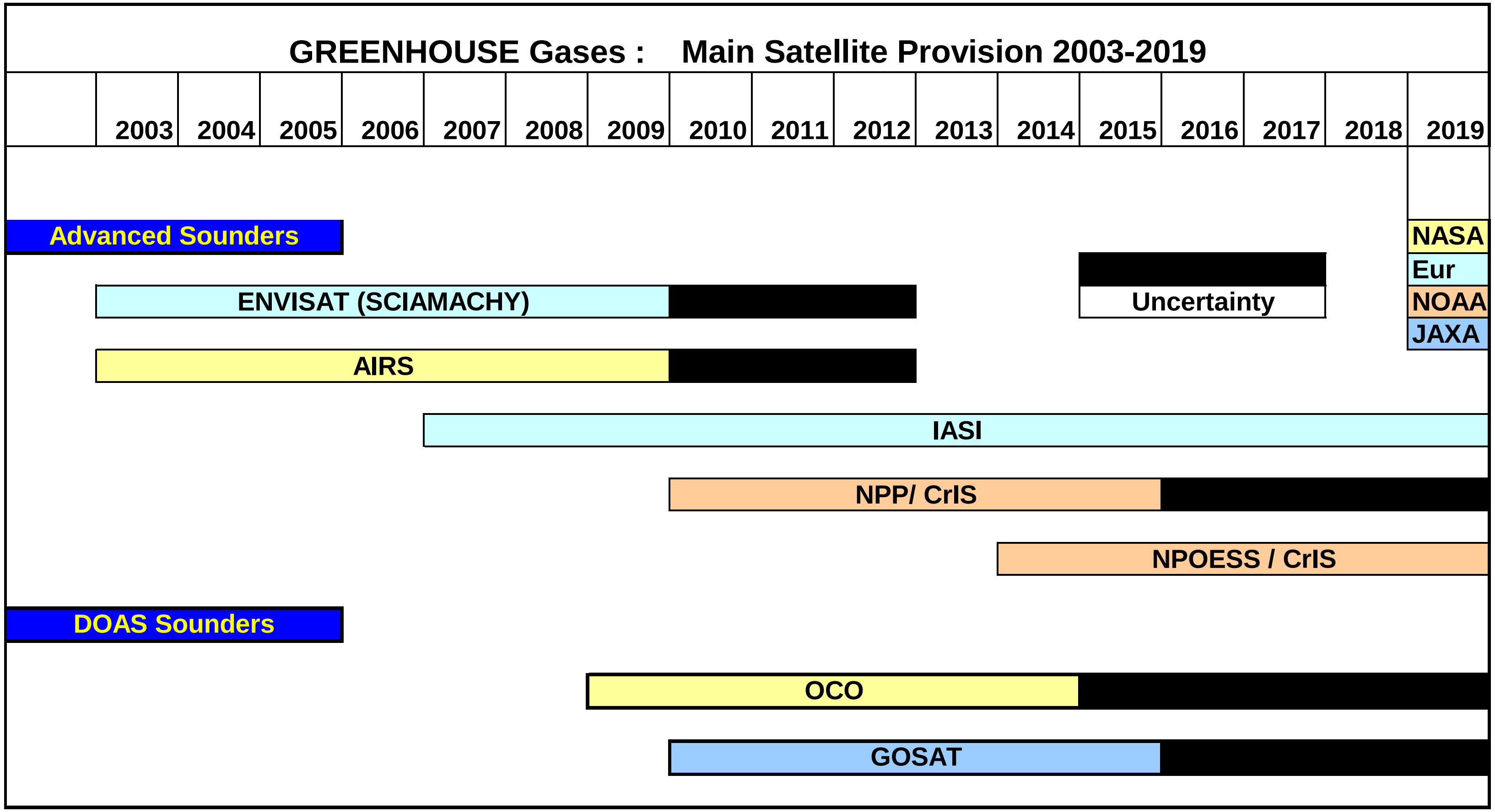
Anomalous terrestrial uptake for the 21st century calculated by the BETHY model forced by output from the IPSL climate model (SREES-A2 scenario run). the red curve uses unoptimized parameters while the black curve uses optimized parameters.

Rayner et al. (2005, 2007)

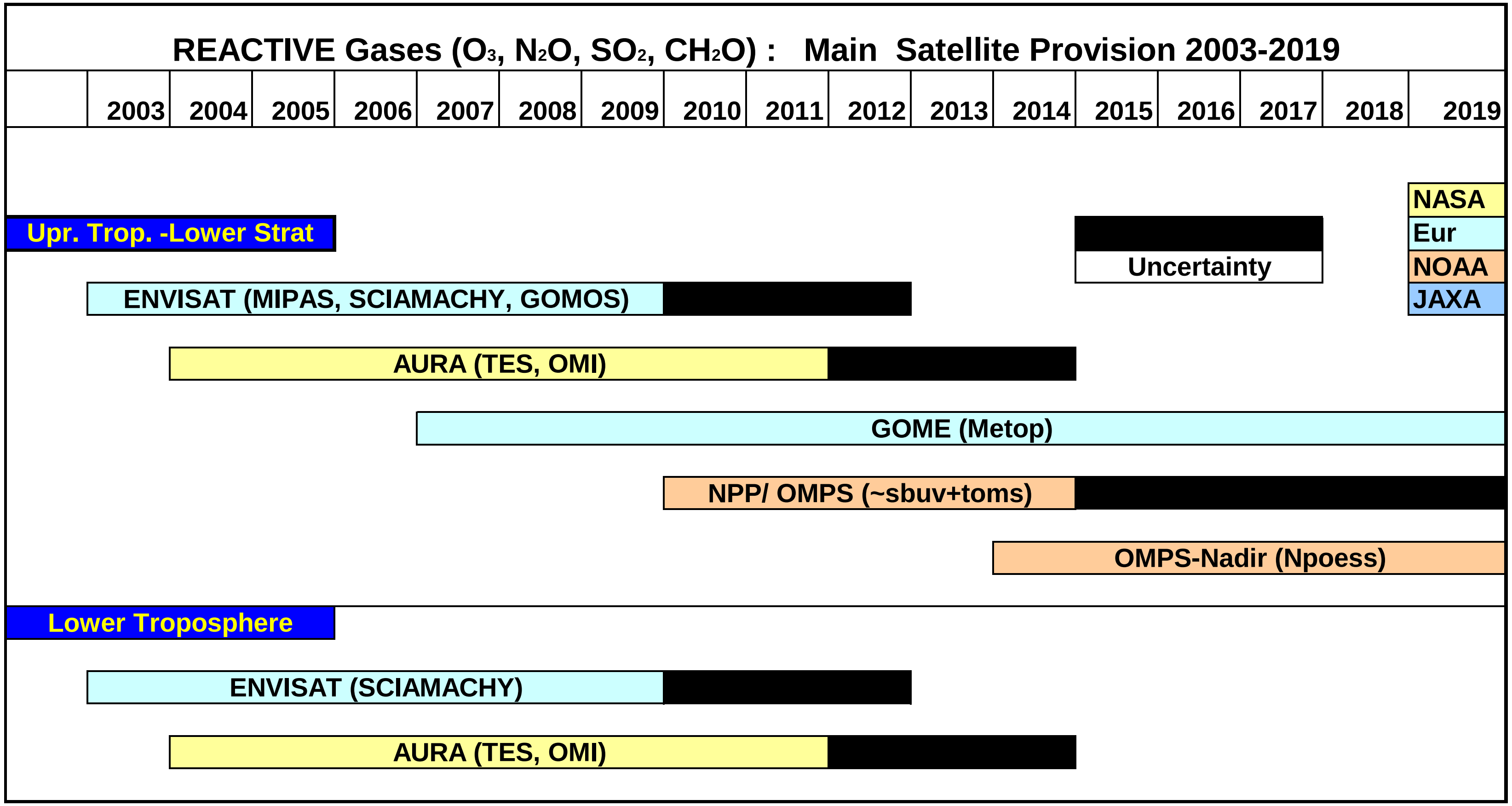
Summary

- From the assimilation of satellite data to the inversion of parameters
- Comprehensive approach
- Increased sophistication
- Large networks of expertise required

Greenhouse gas provision



Reactive gas provision



Aerosol provision