Cross validation of MIPAS/HALOE trace gas observations by means of four dimensional variational assimilation

J. Schwinger, H. Elbern

Rhenish Institute for Environmental Research at the

University of Cologne

Outline:

- Why using data assimilation techniques for cross validation?
- Description of the SACADA assimilation system
- Consistency checks and tuning of error variances
- Cross validation results MIPAS/HALOE ozone profiles

Why using data assimilation techniques for cross validation?

The conventional approaches:

“Direct matching”:

- **Finding pairs of collocated profiles (within limited miss-distance and miss-time)**

“Trajectory mapping”:

Use a model to bridge time and space between measurements of two different sensors (e.g. Morris et al., 2000)

Recent studies using the direct matching approach:

Instruments	Miss distance	Miss time	Source
SMR / MLS	+/- 5° Lon +/- 1.5 Lat	+/- 6 h	Barret et al. (2006)
MIPAS / SMR MIPAS / HALOE	+/-10° Lon +/- 5° Lat (up to 750 km)	+/- 6 h +/- 12 h	Wang et al. (2005)
ASUR / SCIA ASUR / MIPAS ASUR / SMR ASUR / OSIRIS	1000 km	+/- 6 h	Kuttippurath et al. (2007)
MIPAS / Ozone sondes	400 km	+/- 6h	Migliorini et al. (2004)

The Data Assimilation approach:

1) Assimilate observations of instrument A to obtain the best estimate of the atmospheric state

2) Compare observations of instrument B to the analysis

Prerequisite:

Coverage of instrument A must be dense enough

Advantages:

-> Cross validation is more or less a by-product of the assimilation procedure

-> As "coincidence" is not an issue, neither in space nor in time, all available information of sensor B can be used

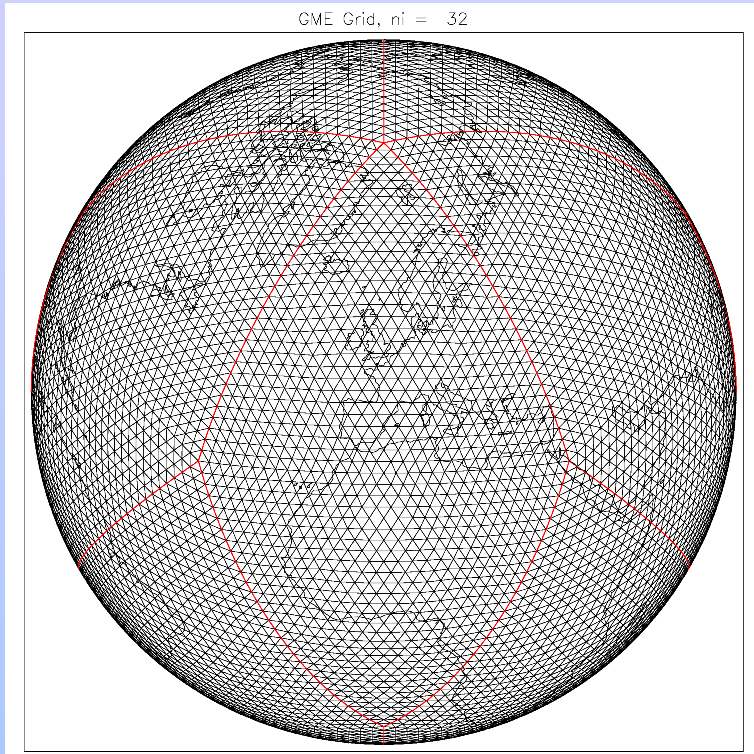
Issues that must be addressed include: Statistical assumptions, specification of background and observation error covariances, model biases…

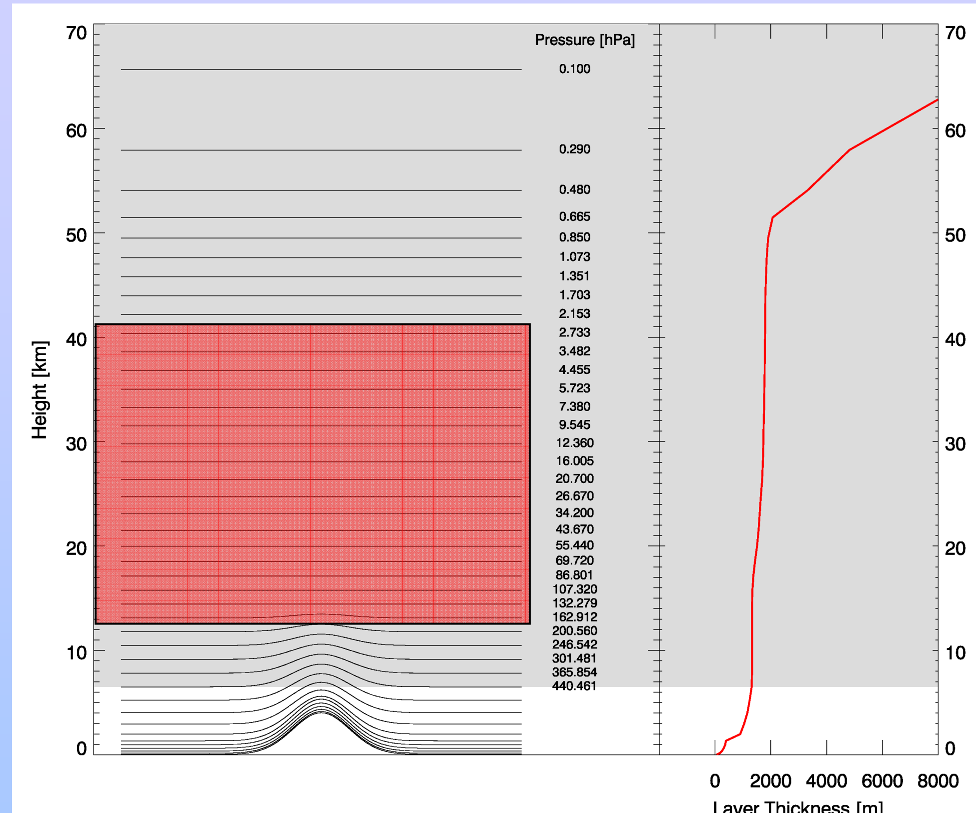
SACADA: Synoptic Analyses for Chemical Constituents with Advanced Data Assimilation

Principal features of the SACADA assimilation System:

- 4D-var data assimilation approach
- German Weather Services global forecast model (GME) serves as an online meteorological driver
- Icosahedral grid and semi Lagrange transport scheme have been adopted from GME
- Horizontal grid spacing ~240 km
- 42 level ranging from the surface to 0.1 hPa

SACADA system - Model grid:





horizontal separation of grid points 220-260 km, nearly homogeneous over the globe

hybrid sigma / pressure coordinate system

Assimilation for this study on 19 levels from 2.4 – 168 hPa (red shaded region)

Features of the SACADA Assimilation System (cont.)

- State of the art chemistry module
  - -Accounts for 167 gas phase and 10 heterogeneous reactions on aerosol and PSC surface
  - -Reaction rates taken from JPL-Recommendations (2006)
  - -KPP-generated chemistry solver (2^{nd} order Rosenbrock) no family assumptions
- Adjoint modules have been build for advection, gas phase chemistry and heterogeneous chemistry
- Flow dependent parameterisation of the Background Error Covariance Matrix (BECM) using a diffusion approach (Weaver and Courtier, 2001)

Issues that must be addressed include: Statistical assumptions, specification of background and observation error covariances, model biases…

- Specification of error variances/covariances?
  - -> Tuning of error variances, consistency tests
- Bias of the model and/or observations?
  - -> Model bias against observations can be detected (O-B differences)
- Analysis error?
  - -> Estimate in observation space is available

Tuning of error covariances (Desroziers et al. 2005)

$$E\left\{\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}}\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}T}\right\} = \mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{T}$$

$$E\left\{\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}}\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}T}\right\} = \tilde{\mathbf{R}}$$

$$\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}} := H(\boldsymbol{x}^{\mathrm{a}}) - H(\boldsymbol{x}^{\mathrm{b}})$$

$$\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}} := \boldsymbol{y} - H(\boldsymbol{x}^{\mathrm{a}})$$

$$\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}} := \boldsymbol{y} - H(\boldsymbol{x}^{\mathrm{b}})$$

If $\mathbf{B}$ and $\mathbf{R}$ are correctly specified, then $\mathbf{B} = \tilde{\mathbf{B}}$ and $\mathbf{R} = \tilde{\mathbf{R}}$ and

$$E\left\{\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}}\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}T}\right\} = \mathbf{H}\mathbf{A}\mathbf{H}^{T}$$

In practise: Iterative approach

Iterative approach:

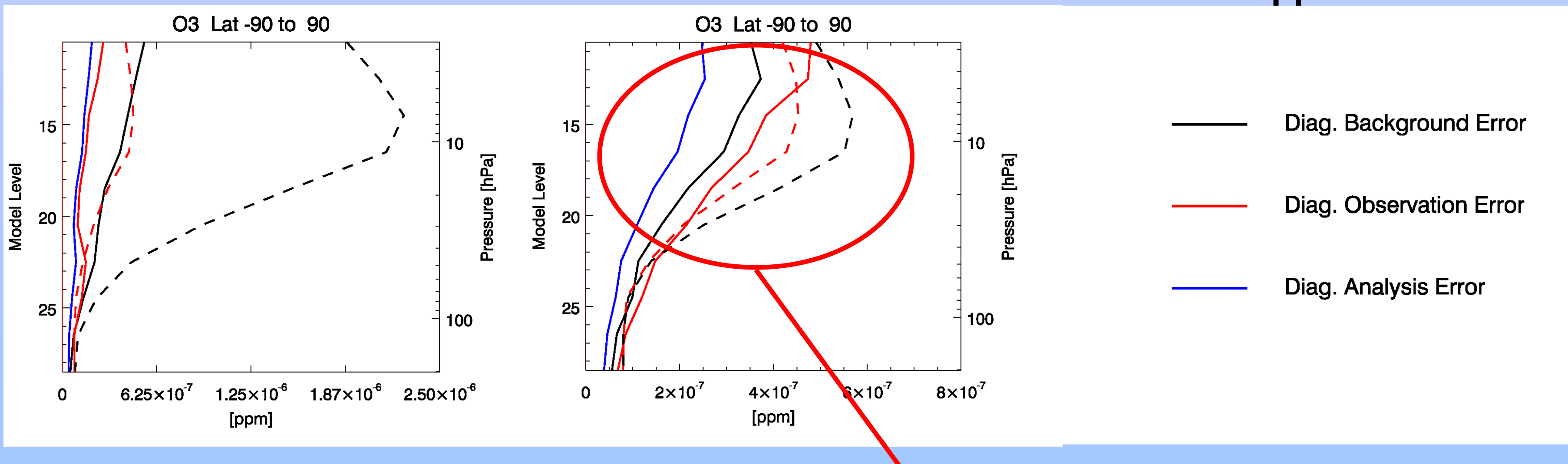
Adapt two parameter model for background and observation errors until satisfying coincidence of specified and diagnosed errors (only diagonal elements of B and R considered)

First guess specification:

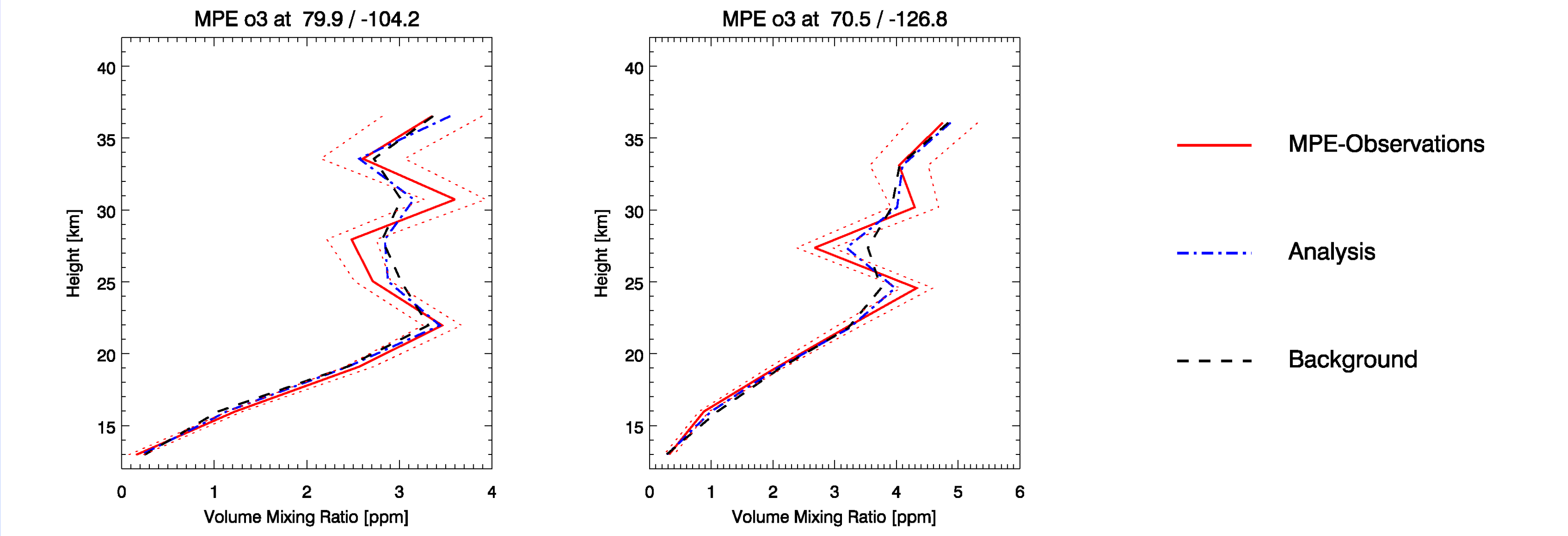
Background error: 30%
Min. observation error: 10%

Final tuning result:

Background error: 8% / 0.08 ppm
Min. observation error: 6% / 0.08 ppm



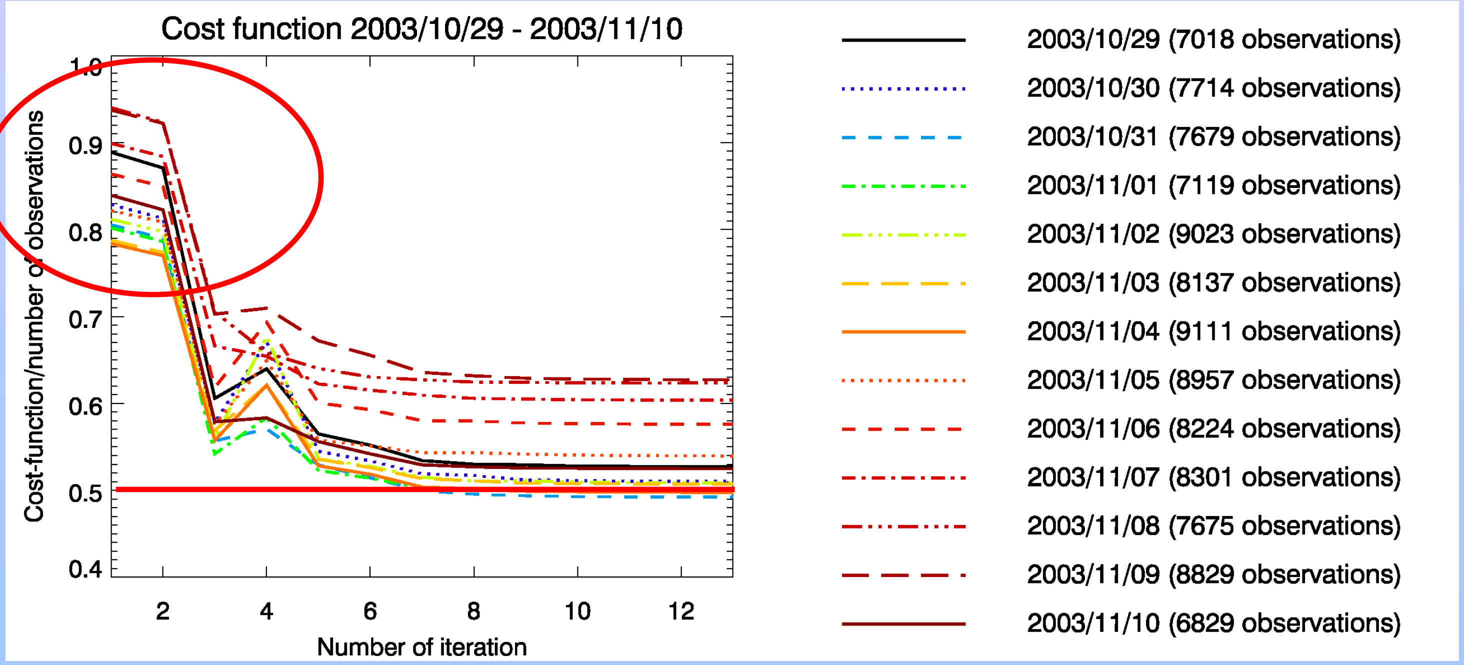
Really?



About 10% of all MIPAS-ESA ozone profiles show these kind of oscillations (artefacts from the retrieval algorithm?)

Another consistency test for error covariances:

$$E\{J(\boldsymbol{x}^{\mathrm{a}})/p\} = 1/2 \quad \text{where} \quad p := \text{nb of observations}$$

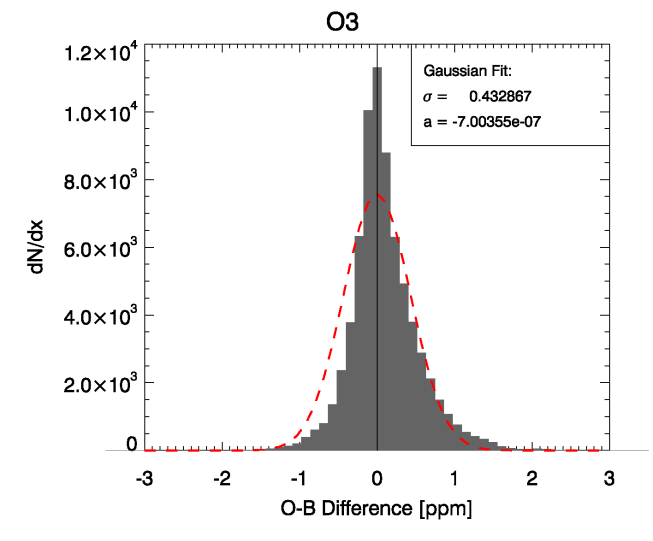


Background field is already within error bars of observations

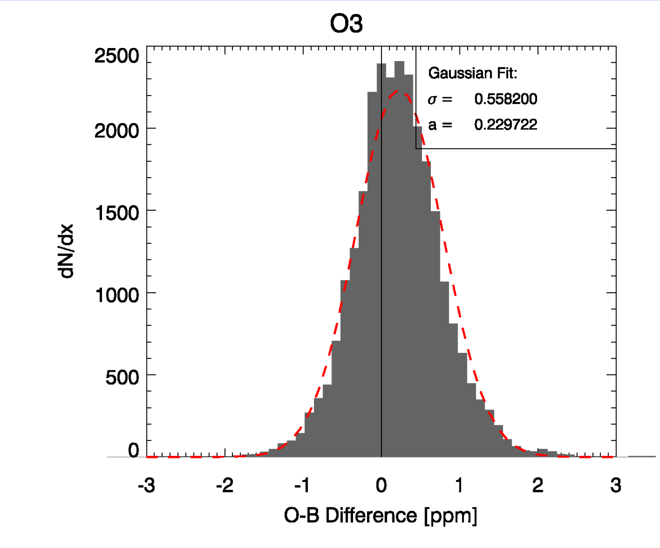
Inspection of O-B differences:

- Differences are approximately gaussian distributed
- Slight bias is found in the upper model region

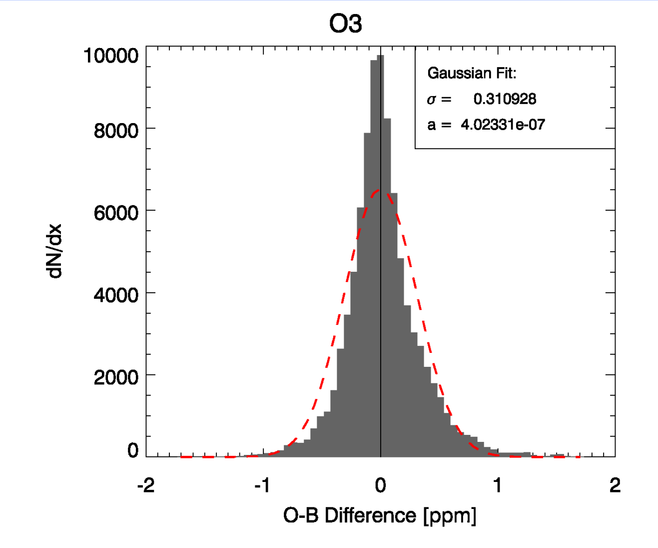
2.4 – 168 hPa



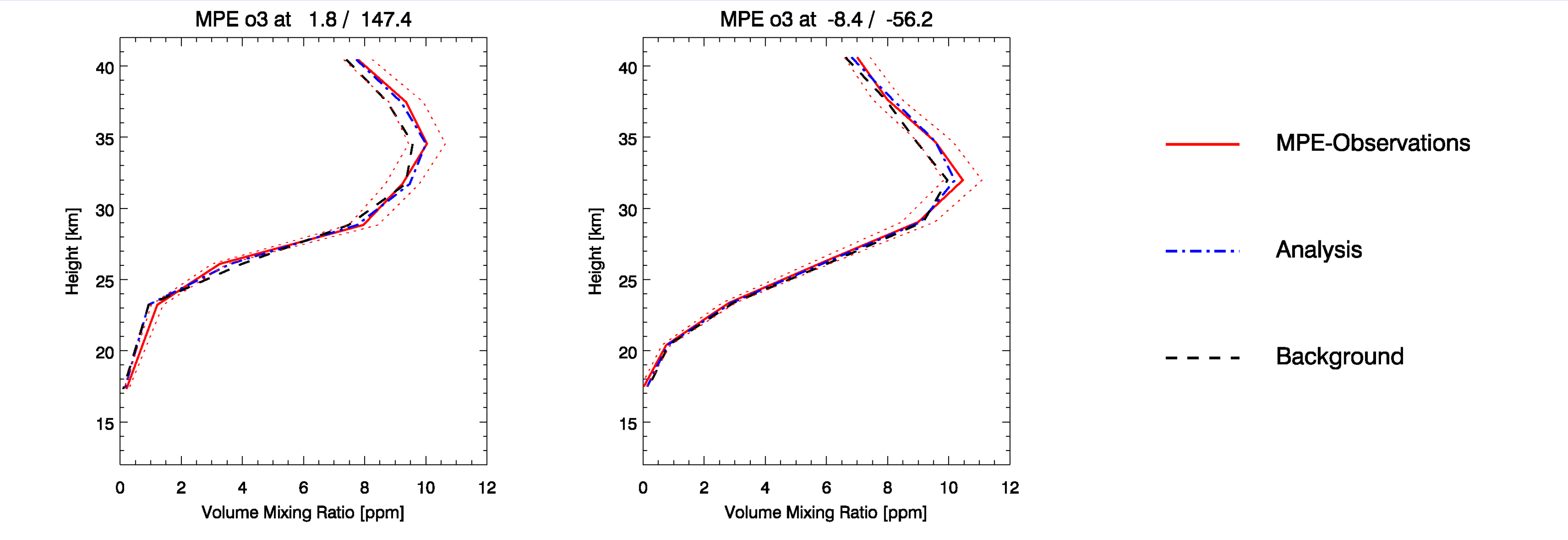
2.4 – 10.7 hPa



10.7 – 168 hPa



Bias of background w.r.t. MIPAS is found at tropical latitudes (where data is sparsest) in the region of vmr-maximum -> Inconsistency of data with the model photochemistry

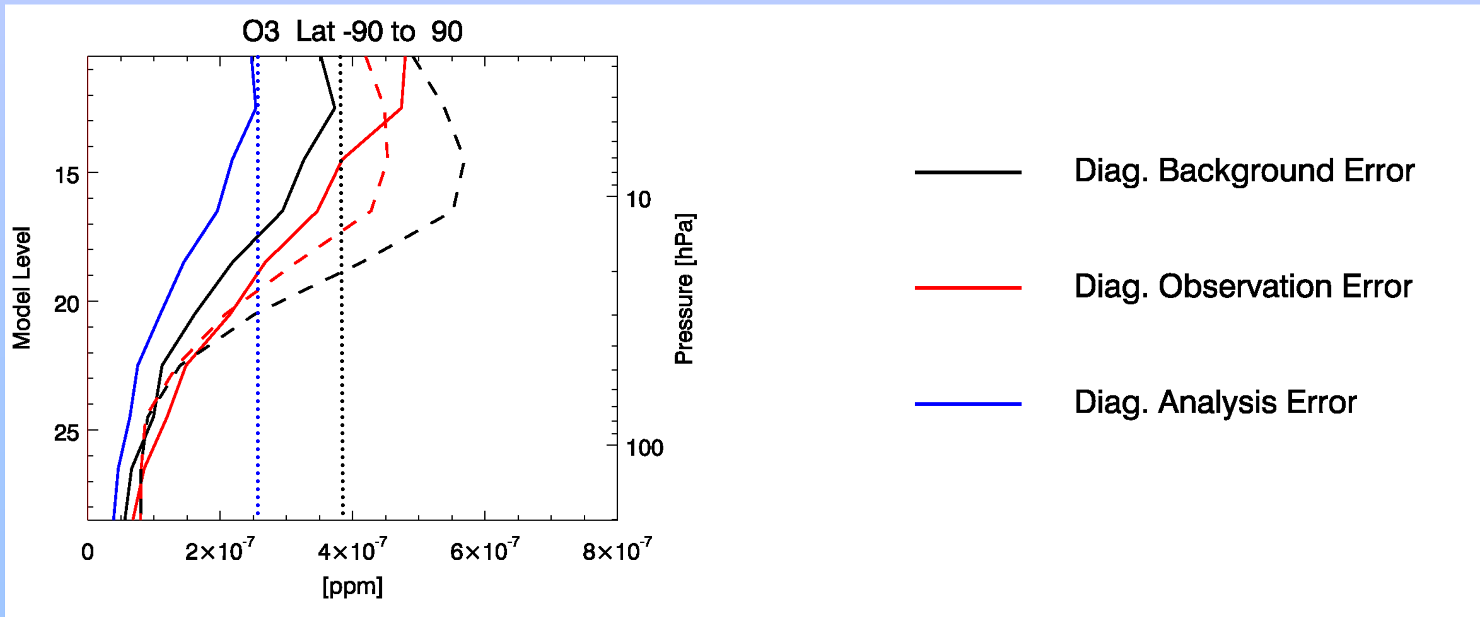


Estimate of analysis error in observation space

- In observation space: Estimate according to Desroziers
- Elsewhere: Will not be larger than Background error

$0.25 < \sigma^a < 0.4$ ppm at ~ 2 hPa decreasing to

$0.05 < \sigma^a < 0.1$ ppm at ~100 hPa



Assimilation system set-up:

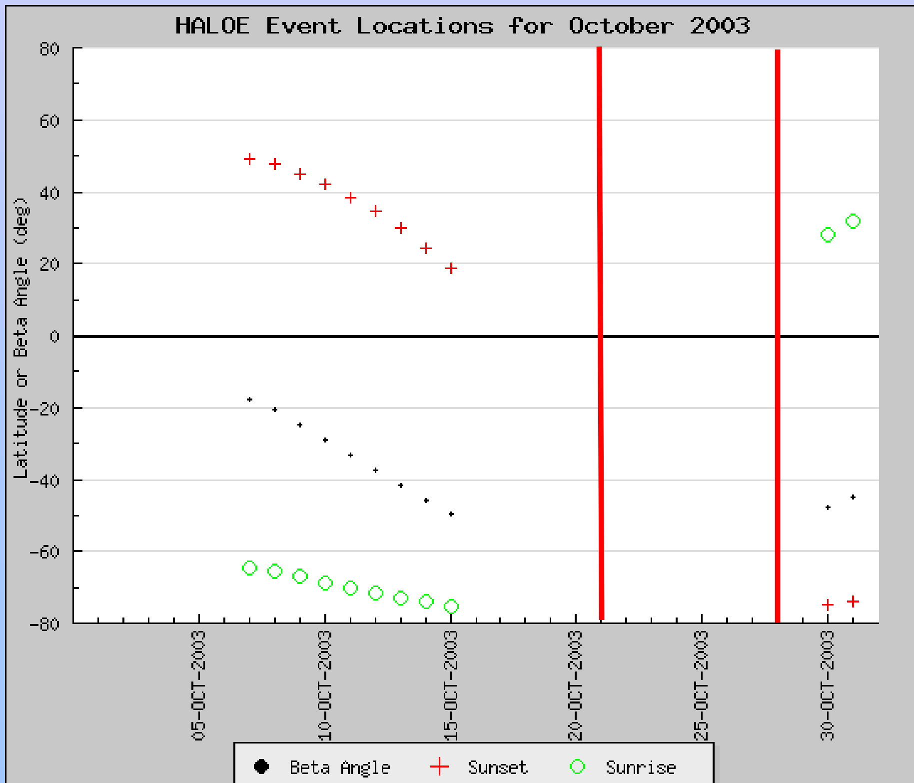
- Case study period Oct 21, 2003 – Nov 10, 2003
- Assimilation of MIPAS profiles (data version 4.61)
- Spin up assimilation Oct 21 – Oct 29 starting from SOCRATES zonal mean fields
- Successive adjustment of background error variances:

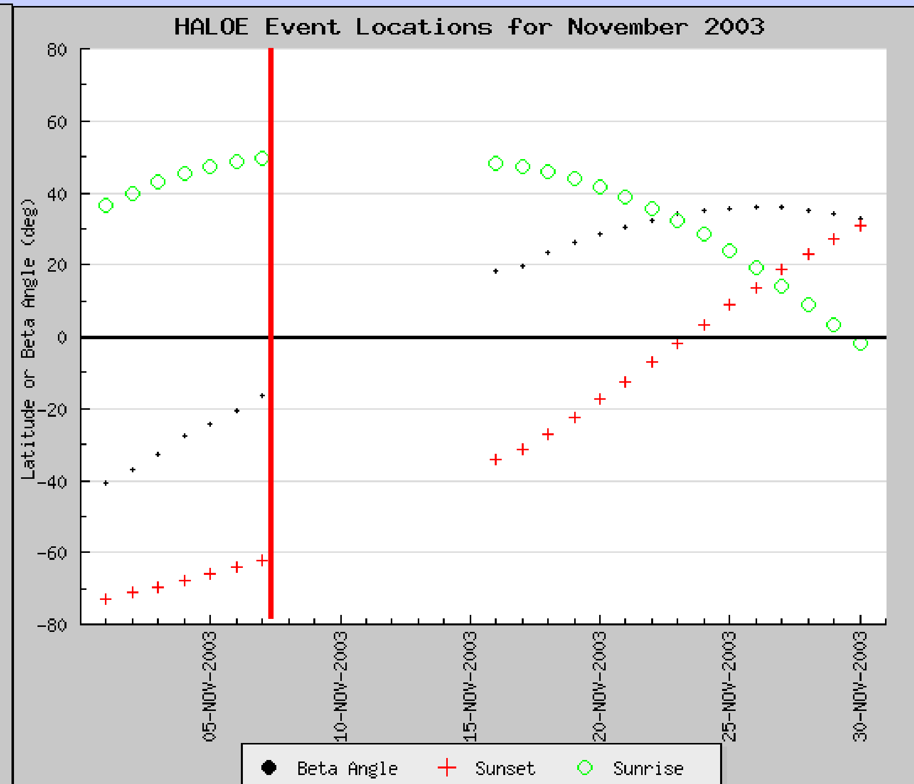
Date	Background Error	Correlation Length h/v
10/21 – 10/26	30% / 0.08 ppb	400 km / 1.5 km
10/27 – 10/28	10% / 0.08 ppb	400 km / 1.5 km
10/29 – 11/10	8% / 0.08 ppb	400 km / 1.5 km

Spin-up assimilation
8 days (10/21 – 10/28)

Cross validation
9 days (10/30 – 11/07)

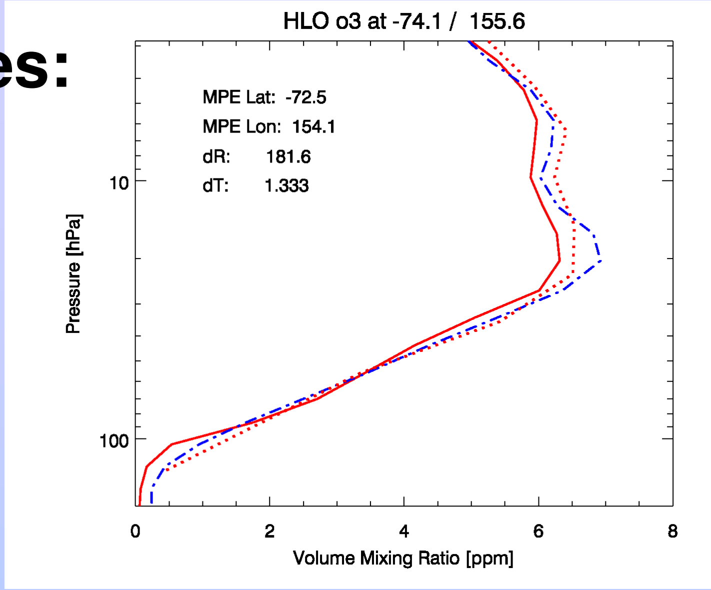
Total of 241 HALOE Profiles

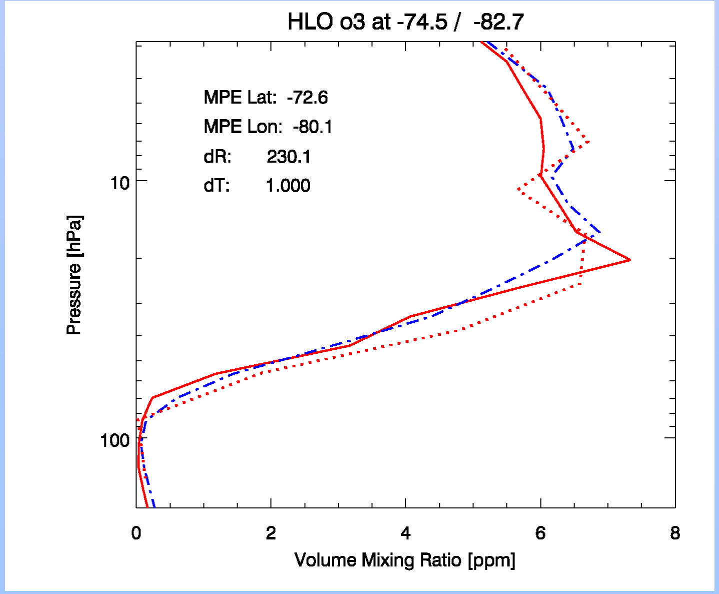


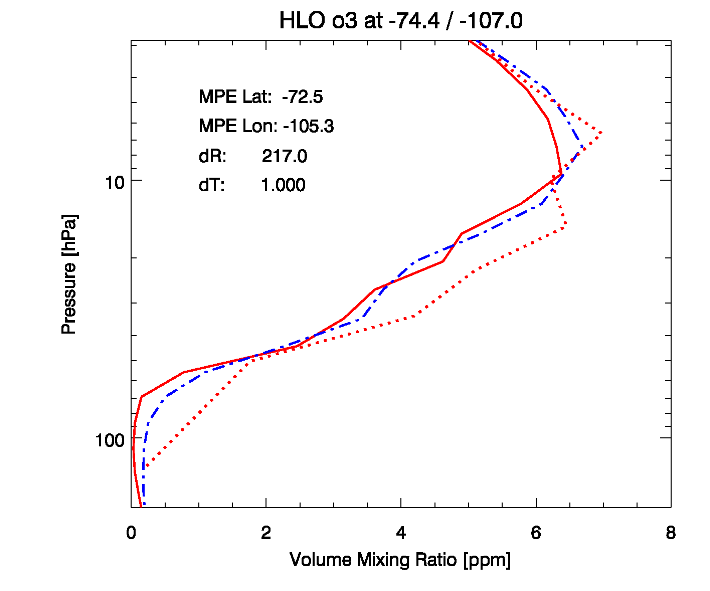


Comparison of 3 individual profiles:

- **HALOE O_3 profiles (not assimilated)**
- **Analysis obtained from MIPAS O_3 profiles**
- **Collocated MIPAS profile (assimilated)**

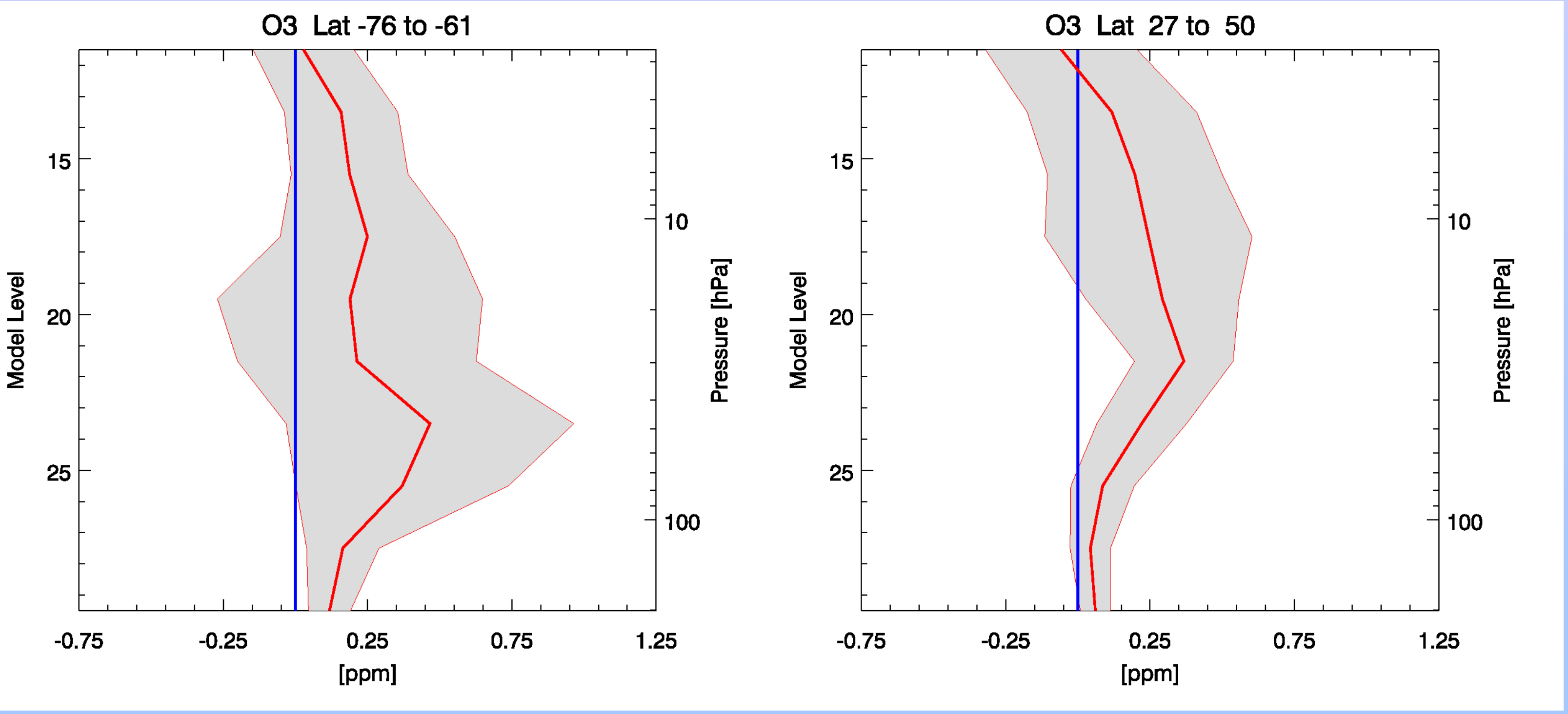






Mean MIPAS – HALOE differences

10/30/2003 – 11/07/2003



Total of 121 HALOE Profiles

Total of 120 HALOE Profiles

Summary, conclusions and outlook:

- Using assimilated fields is the most “natural” way of comparing profiles from different sensors (no arbitrary collocation criteria)
- The SACADA system proves to be useful for satellite data intercomparison
- Prerequisite:
  - At least one sensor must provide dense coverage in order to make the analysis error small
  - Consistency between data and model
- Statistics of comparison is much ‘better’, especially if data from the second sensor is sparse
- In principle, even the comparison of reactive species is possible

Acknowledgements

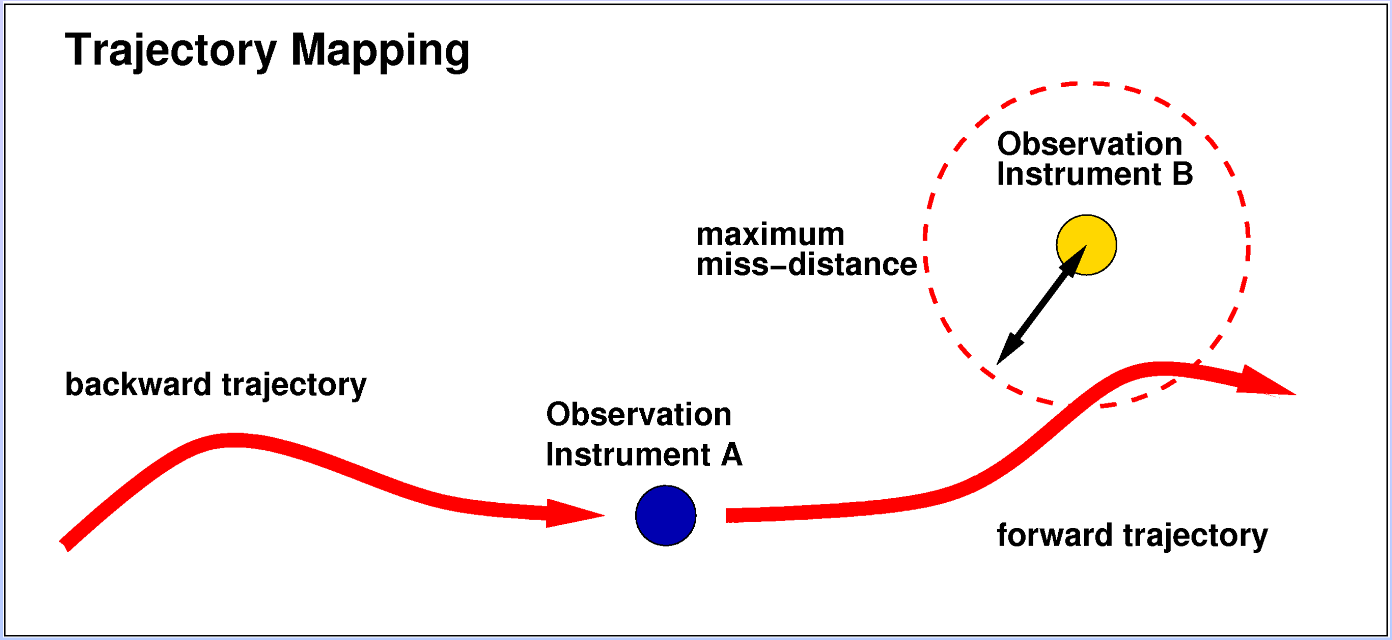
We are especially grateful for support by the following groups, persons and institutions:

German Weather Service for provision of the GME software

ESA for providing the MIPAS data

HALOE data has been supplied by Hampton University and NASA Langley Research Center, Hampton, Virginia

The SACADA project was funded by the German Federal Ministry of Education and Research (BMBF) in the framework of AFO-2000

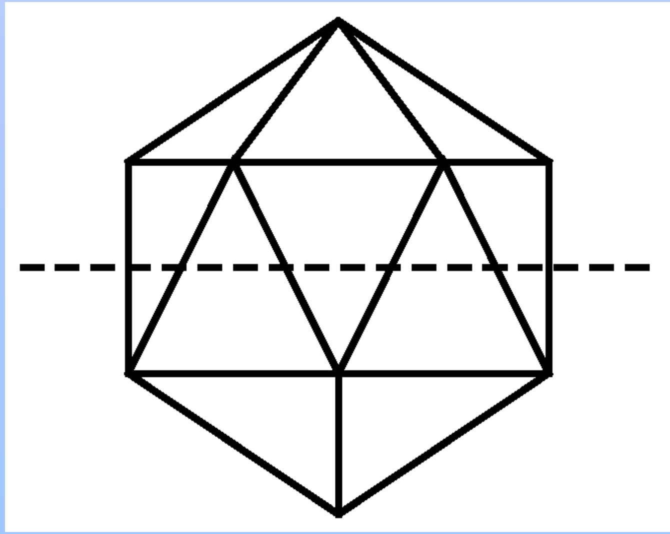


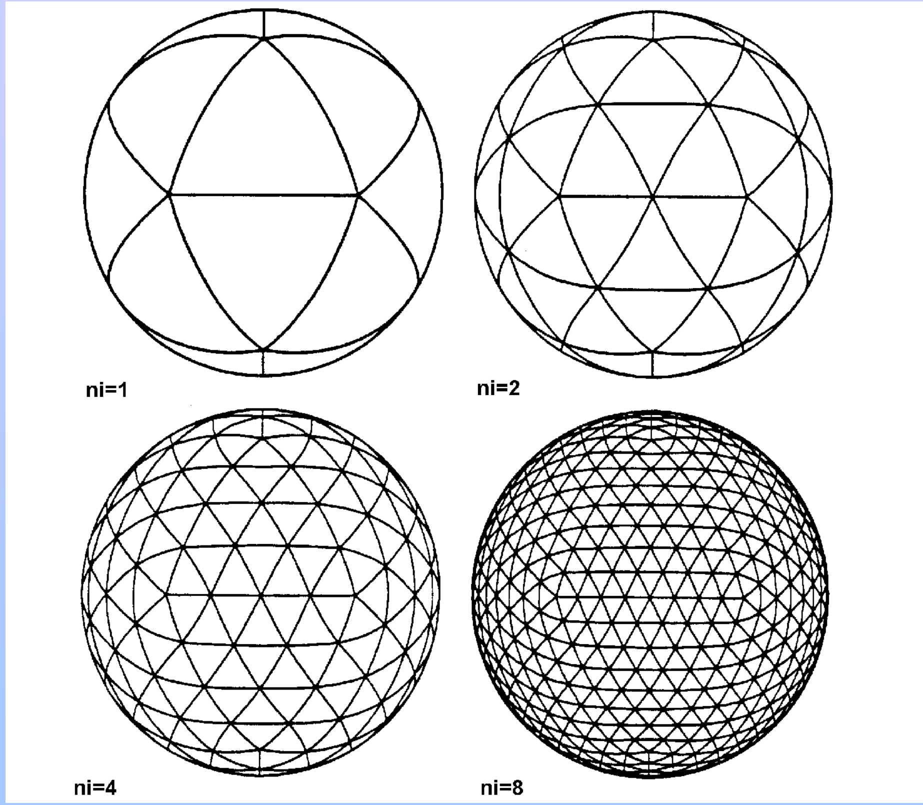
"Trajectory mapping,…, allows a stricter definition of 'coincidence' to be used than in the traditional technique, increases the number of coincident data pairs available for comparison, and at the same time enables a decrease in statistical uncertainty to be achieved."

From "Assessment of trends in the Vertical Distribution of Ozone",
SPARC Report No 1, 1998

Construction of Icosahedral Grid

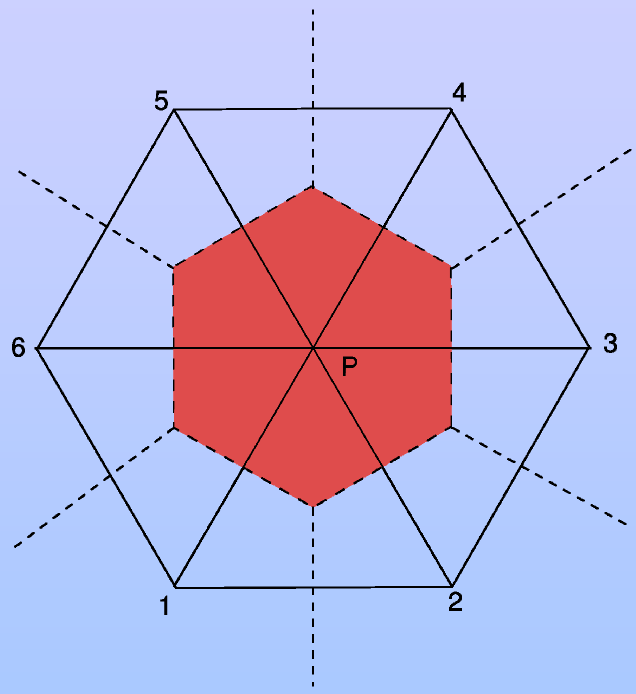
An Icosahedron (20 equilateral triangles) is placed in a sphere. Vertices are connected by great circles. Great circle arcs are subdivided into n_i intervals to form a regular grid.

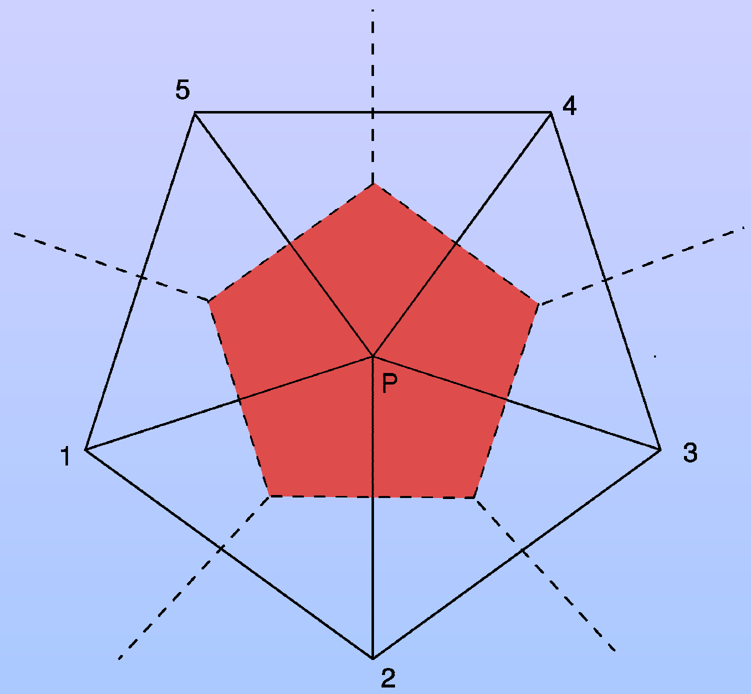




Source of Figures: DWD

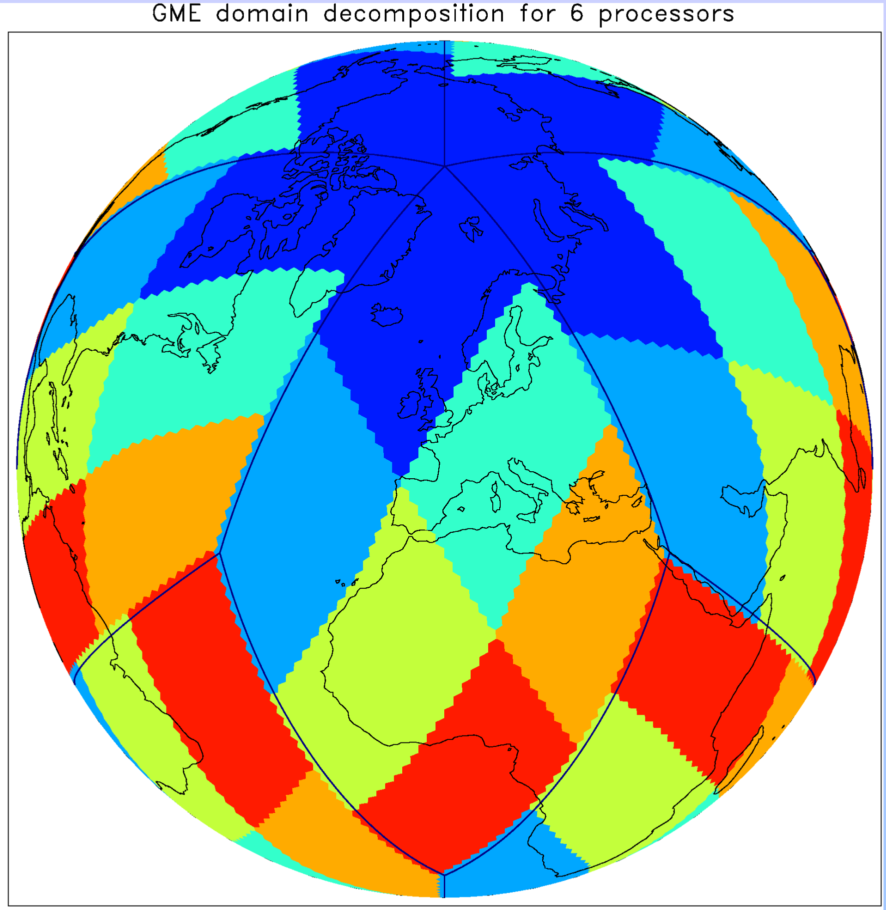
Grid Cells of Icosahedral Grid





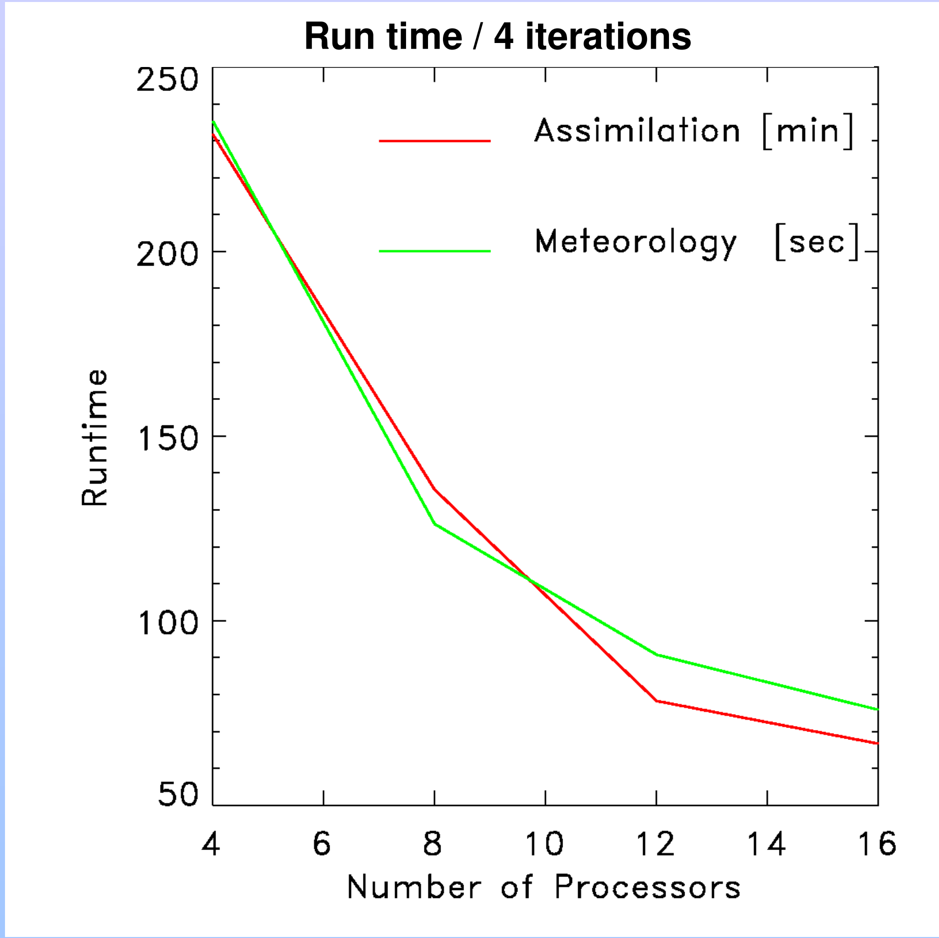
The area of representativeness belonging to one grid cell is a hexagon or (at the twelve special points) a pentagon

Computational Aspects, Parallelisation Strategy



Domain decomposition for 6 processors à simple but effective strategy for load balancing

Computational Aspects



With an assimilation window of 24h and 12 using iterations, the wall clock run time is about 3 hours using 16 processors

Incremental formulation of the cost function

Define new control variable

$$\boldsymbol{v} = \mathbf{B}^{-1/2}\left[\boldsymbol{x}_0 - \boldsymbol{x}^{\mathrm{b}}\right]$$

Cost function

$$J(\boldsymbol{v}) = J^{\mathrm{b}} + J^{\mathrm{o}} = \frac{1}{2}\boldsymbol{v}^T\boldsymbol{v} + J^{\mathrm{o}}$$

Gradient of cost function

$$\nabla_{\boldsymbol{v}} J = \boldsymbol{v} + \mathbf{B}^{T/2}\nabla_{\boldsymbol{x}_0} J^{\mathrm{o}}$$

Minimasation

New initial values

$$\boldsymbol{x}_0 = \mathbf{B}^{1/2}\boldsymbol{v} + \boldsymbol{x}^{\mathrm{b}}$$

=> Inverse of **B** is never needed. $\mathbf{B}^{1/2}$ and $\mathbf{B}^{T/2}$ can be modelled using a diffusion operator

Formulation of the background error covariance matrix:

Diffusion approach (Weaver and Courtier, 2001)

$\mathbf{B}^{1/2}$ and $\mathbf{B}^{T/2}$ encoding quasi Gaussian correlations can be modelled using a diffusion operator:

$$\mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{T/2} = \left(\mathbf{\Sigma\Lambda L}^{1/2}\mathbf{W}^{-1/2}\right)\ \left(\mathbf{W}^{-1/2}\mathbf{L}^{T/2}\mathbf{\Lambda\Sigma}\right)$$

$\mathbf{\Sigma}$: Matrix of background error variances (diagonal)

$\mathbf{\Lambda}$: Matrix of normalisation factors (diagonal)

$\mathbf{L}$: Diffusion Operator

$\mathbf{W}$: Matrix of grid cell area elements (diagonal)

SACADA Assimilation System - Setup