Assimilation of Radio Occultation Data at DWD

Andreas Rhodin

andreas.rhodin@dwd.de

Deutscher Wetterdienst, Offenbach, Germany

2nd GRAS SAF User Workshop

Helsingør

12th June 2003

Contents

Status of the new DWD 3D-Var system.

Observation Space Assimilation System OSAS (PSAS)

Solution of the equations: Preconditioning, nonlinear terms

Implementation of a GPS ray-tracer operator.

DWD 3D-Var

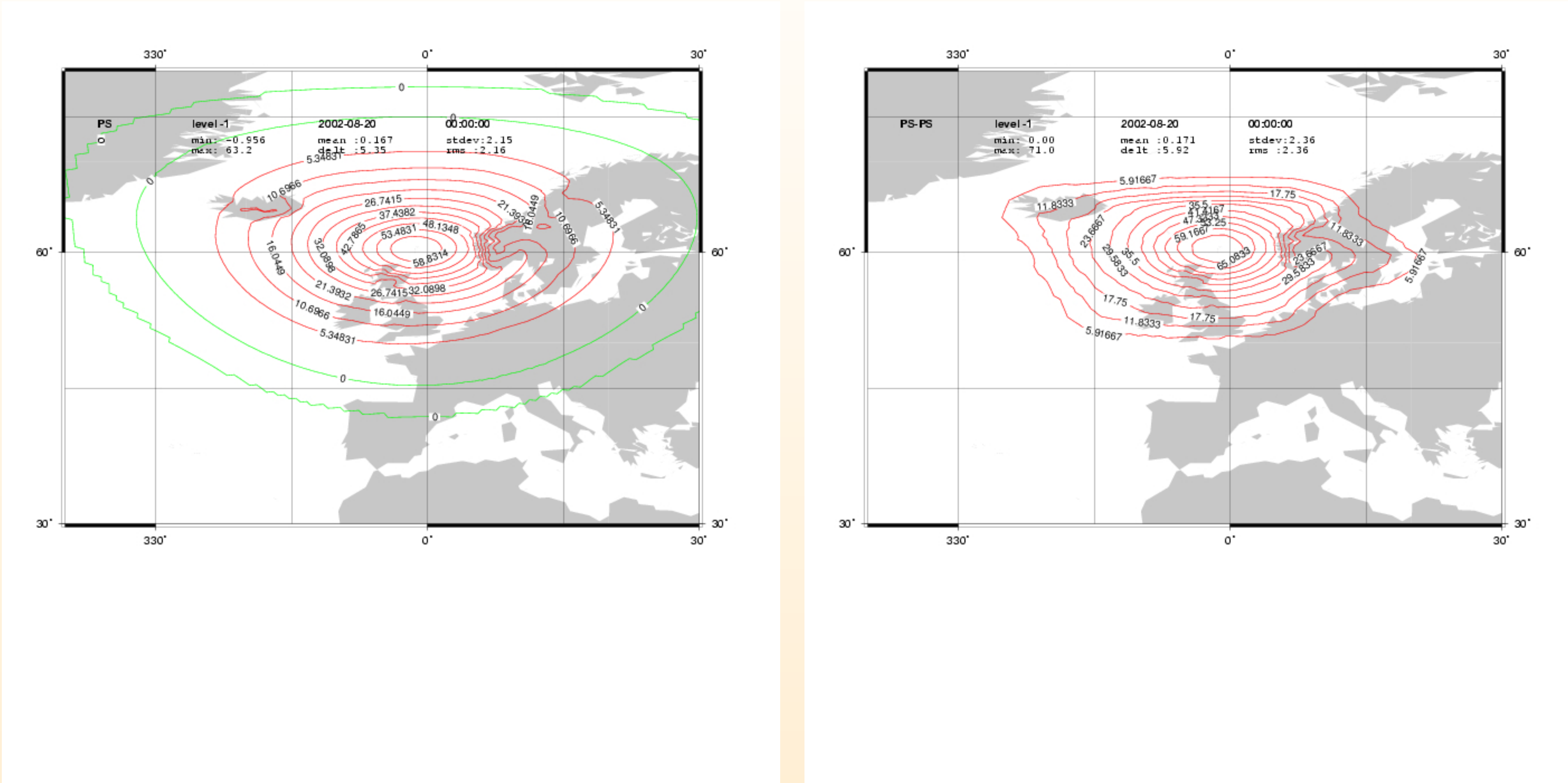
- A 3D-Var system is build up in order to replace the operational OI. The new system is more appropriate for assimilation of remote sensing data.
- The control variables are still chosen in observation space (OSAS/PSAS). This approach provides full flexibility for the specification of the background error covariance matrices.
- The development builds up on activities started at MPIfM, including the implementation of a GPS Ray-tracing Operator (Michael Gorbunov, Luis Kornblueh).

  The Ray-tracing Operator will remain in the DWD 3D-Var System as a prototype for a limb-sounding operator.

Status

- Up to now the background error model of the OI are used.
- In situ observation operators are implemented.
- (Re)implementation of the GPS ray-tracer operator is under way.
- Some of the deficiencies of the OI (due to local solution of the equations in observation boxes) are not present in the 3D-Var.

Response to a single pressure observation



Surface pressure analysis increment for a single observation in the 3D-Var (left) and OI (right).

3D-Var Formulation

$$J(\mathbf{x}) = \frac{1}{2}\big((\mathbf{x}-\mathbf{x}_b)\mathbf{P}_b^{-1}(\mathbf{x}-\mathbf{x}_b) + (H(\mathbf{x})-\mathbf{o})\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{o})\big)$$

$\mathbf{R}$: Observation error covariance matrix:
Sparse, Only data within one observation are correlated.

$\mathbf{P}_b$: Forecast Error Covariance Matrix
Dense within an influence radius of some 1000 km, zero outside

- size of $\mathbf{x}$: $n \approx 10^7 \ldots 10^8$
- size of $\mathbf{P}_b$: $n^2 \approx 10^{15}$

ca. 10% nonzero elements.

OSAS - Observation Space Assimilation System

Cost function:

$$J(\mathbf{x}) = \frac{1}{2}\big((\mathbf{x}-\mathbf{x}_b)\mathbf{P}_b^{-1}(\mathbf{x}-\mathbf{x}_b) + (H(\mathbf{x})-\mathbf{o})\mathbf{R}^{-1}(H(\mathbf{x})-\mathbf{o})\big)$$

Gradient of the cost function:

$$\frac{d}{d\mathbf{x}}J = \mathbf{H}^T\mathbf{R}^{-1}[H(\mathbf{x})-\mathbf{o}] + \mathbf{P}_b^{-1}[\mathbf{x}-\mathbf{x}_b]$$

$\mathbf{H}$ is the Jacoby matrix of the operator H.
In the minimum of J its gradient is zero and the analyzed state $\mathbf{x}_a$ is:

$$\mathbf{x}_a - \mathbf{x}_b = [\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{P}_b^{-1}]^{-1}\mathbf{H}^T\mathbf{R}^{-1}[\mathbf{o} - H(\mathbf{x}_b)]$$

Algebraic manipulation yields:

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{P}_b\mathbf{H}^T[\mathbf{H}\mathbf{P}_b\mathbf{H}^T + \mathbf{R}]^{-1}[\mathbf{o} - H(\mathbf{x}_b)]$$

For in situ observations of model variables $\mathbf{H}\mathbf{P}_b\mathbf{H}^T$ is the background correlation matrix at observation points (denoted $\mathbf{B}$ in the following).

Solution of the OSAS equation

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{P}_b\mathbf{H}\,(\mathbf{B}+\mathbf{R})^{-1}(\mathbf{o} - H(\mathbf{x}_b))$$

with $\mathbf{B} = \mathbf{H}^T\mathbf{P}_b\mathbf{H}$.
$\mathbf{B}$ is the forecast error covariance matrix for the observed quantities.
$(\mathbf{B}+\mathbf{R})$ is a symmetric and positive definite Matrix.

The set of equations is solved in two steps:

1. Solve the system of linear equations:

$$\mathbf{z} = (\mathbf{B}+\mathbf{R})^{-1}(\mathbf{o} - H(\mathbf{x}_b))$$

2. Perform the matrix vector multiplication:

$$\mathbf{x}_a - \mathbf{x}_b = (\mathbf{P}_b\mathbf{H})\,\mathbf{z}$$

CG algorithm - Preconditioning - Parallelization

In order to solve the system of linear equations

$$\mathbf{z} = (\mathbf{B} + \mathbf{R})^{-1}(\mathbf{o} - H(\mathbf{x}_b))$$

a preconditioned Conjugate Gradient (CG) algorithm is used.

For preconditioning, an approximation to $(\mathbf{B} + \mathbf{R})^{-1}$ is required:

- The matrix $\mathbf{B}$ is decomposed into blocks with size $\approx 500 \times 500$. Matrix elements are sorted, so that strongly correlated elements (nearby observations) are located within the same block.
- The block-diagonal matrix $\tilde{\mathbf{B}}$ is used for the preconditioner $(\tilde{\mathbf{B}} + \mathbf{R})^{-1}$. Block diagonals are inverted by Cholesky decomposition.
- The block structure is the basis of parallelization. Matrix blocks are distributed over different processor elements.
- The block structure can also be used for a single observation mode (similar to 1D-Var) by taking only one observation per box.

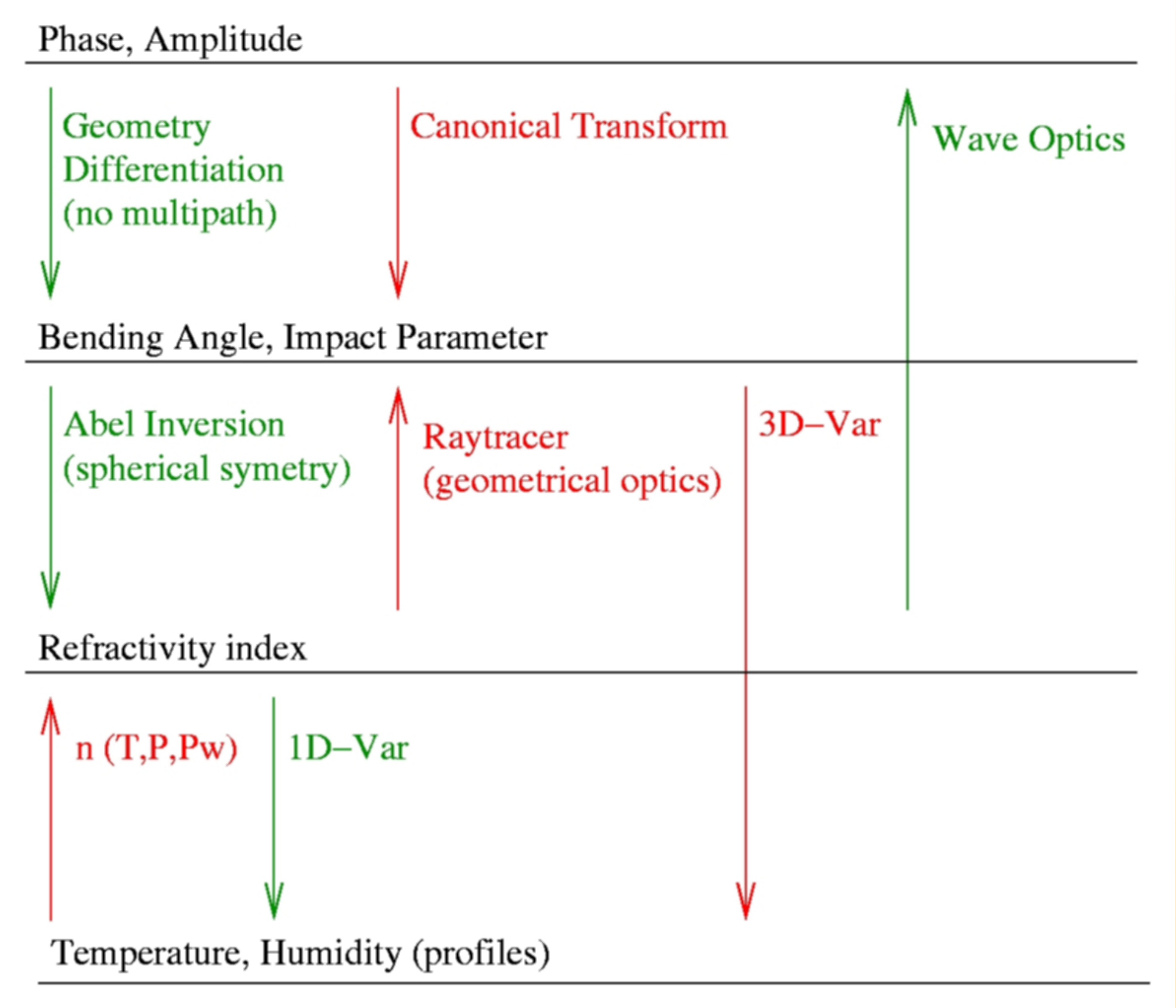
Nonlinear Observation Operators

$$\mathbf{z} = (\mathbf{B} + \mathbf{R})^{-1}(\mathbf{o} - H(\mathbf{x}_b))$$

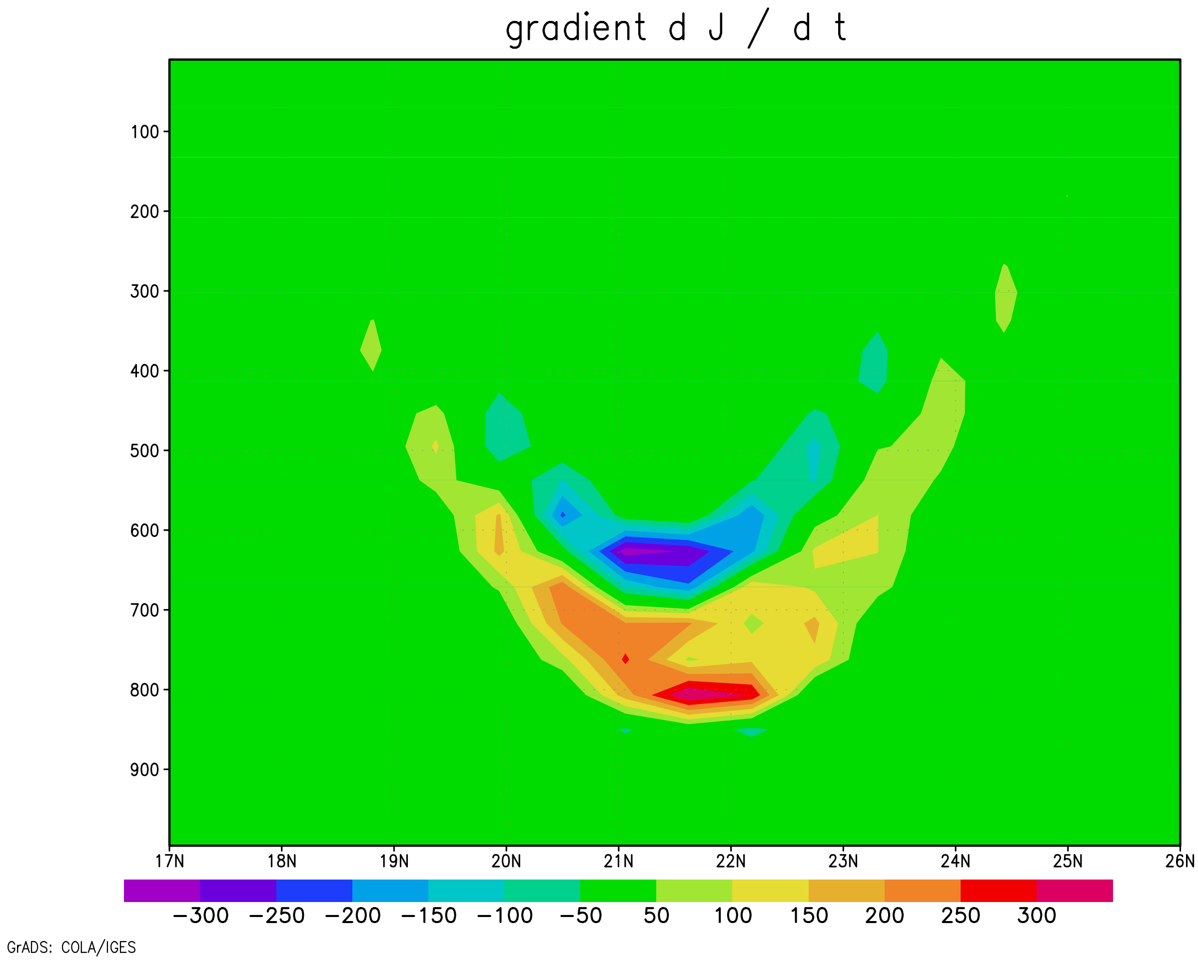
- For nonlinear observation operators H and non-Gaussian error distributions (variational quality control) the linear OSAS equation only holds approximately.
- The CG algorithm serves as a Newton step within an outer loop.
- Due to nonlinearities in H and $\mathbf{R}$, the matrices $\mathbf{R}$ and $\mathbf{B} = \mathbf{H}\mathbf{P}_b\mathbf{H}^T$ must be updated in each outer iteration.

  Evaluation of the nonlinear terms in general is expensive.

GPS Radio Occultations - Observables

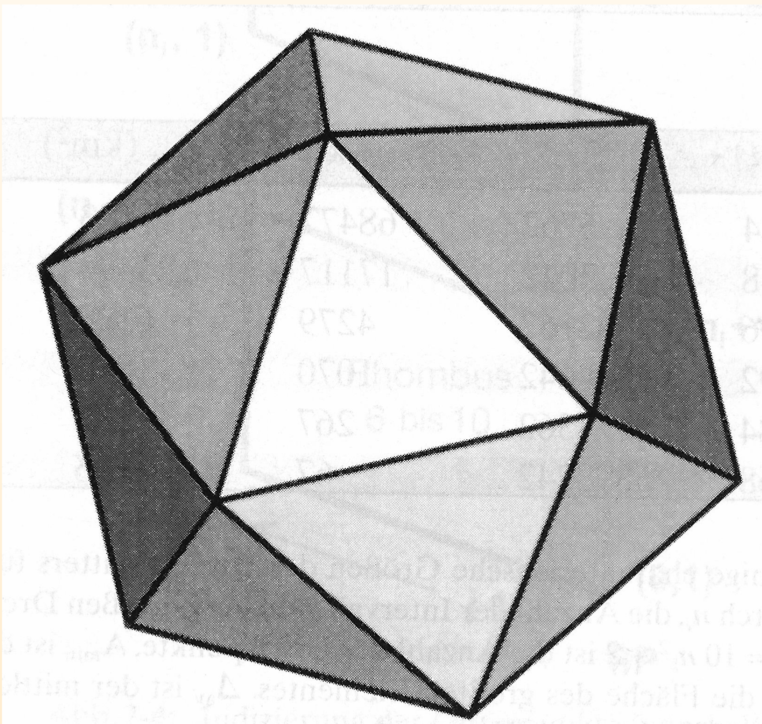


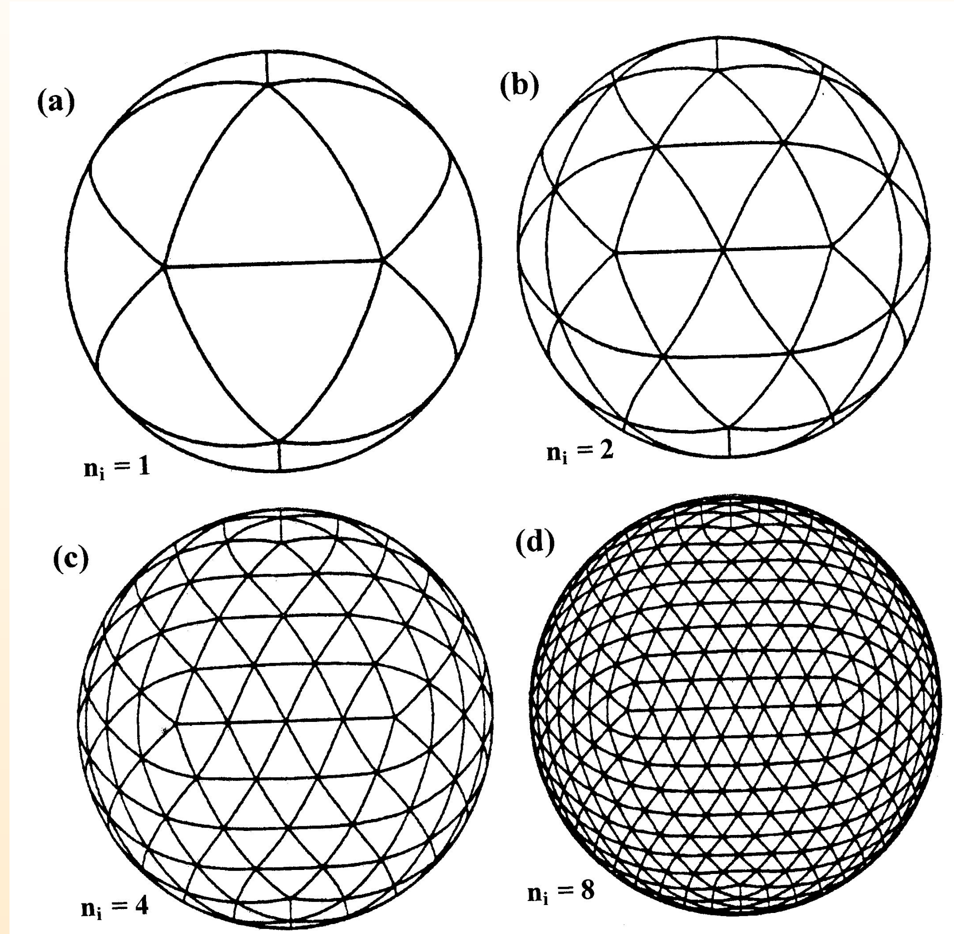
GPS Raytracer - Sensitivity



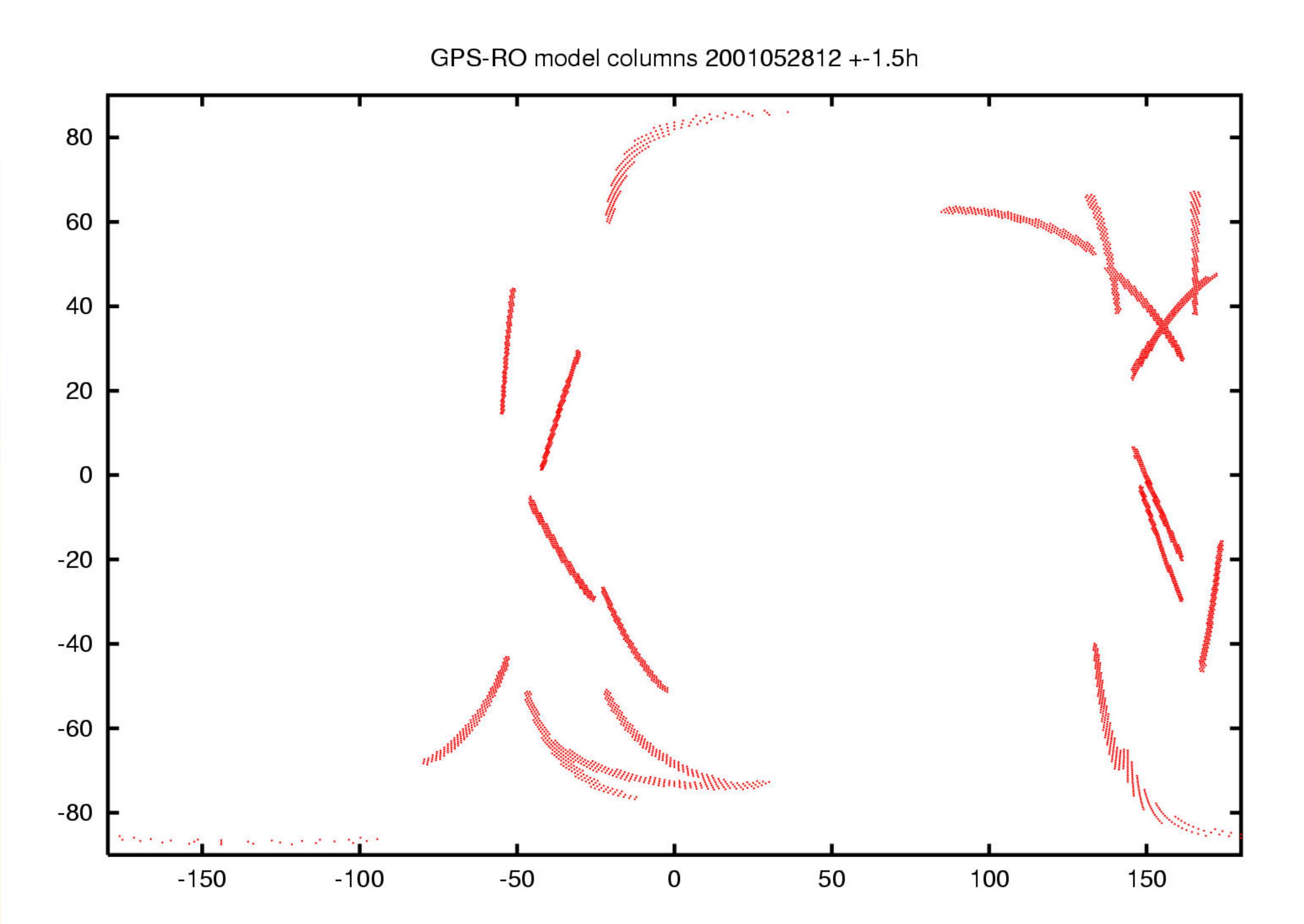
Gradient of the cost function with respect to the atmospheric temperature for a ray-tracer observation operator

Global Model GME - Icosahedral Grid





Selected model columns



Model columns (3277 out of 166410) required to run the GPS-ray-tracer operator for 19 occultations observed by CHAMP within an assimilation window of 3 hours.

Ray-tracing operator - Implementation

Preprocessor:

Derive $\epsilon(p)$ by the CT method

Observation Operator:

$$H(\mathbf{x}) = \epsilon(\mathbf{x}, p)$$

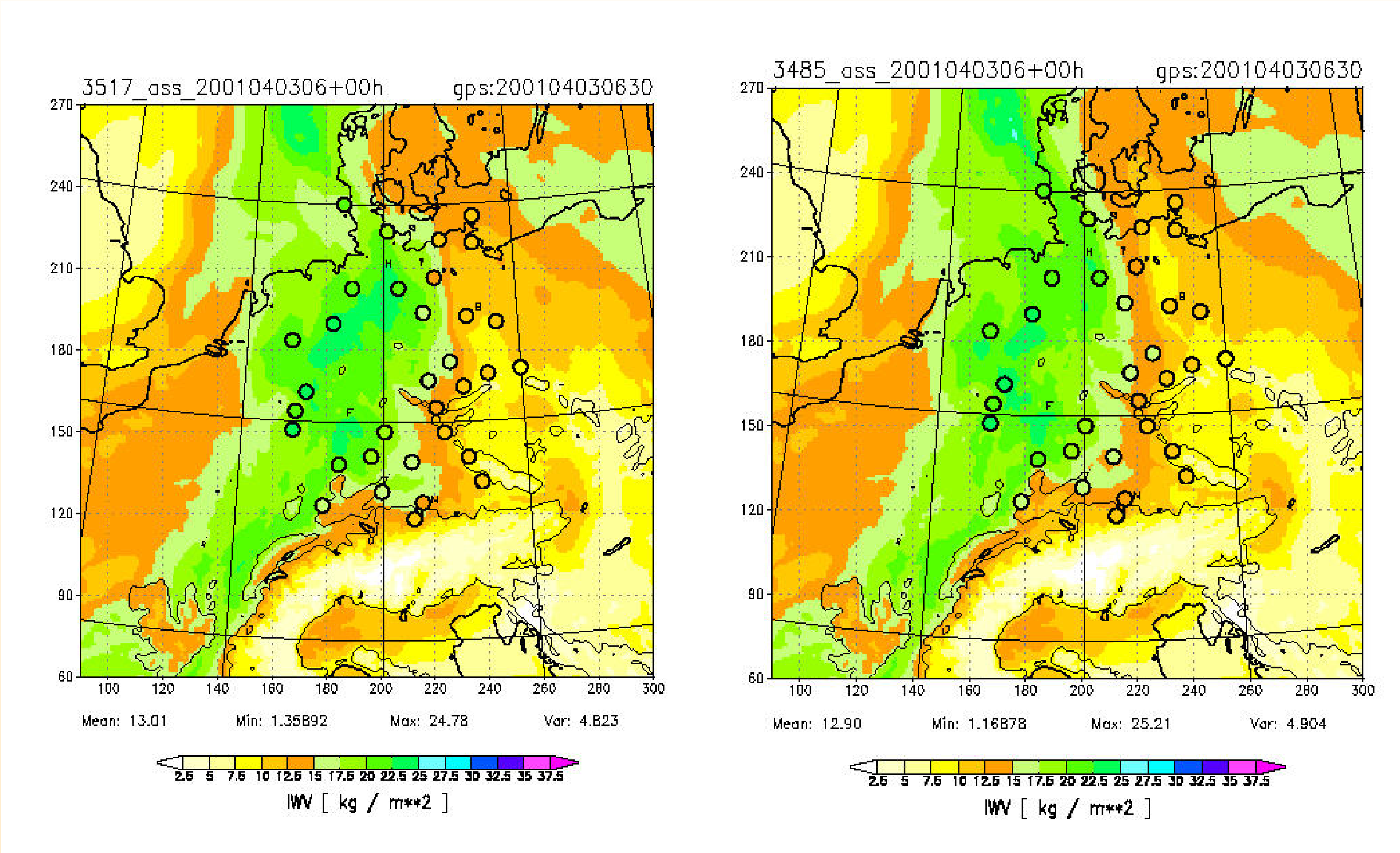
Implementation:

1. Identify model columns the observation operator H depends on.
2. Gather model columns required by H on one processor element.
3. Run the observation operator and its adjoint.
4. Solve the linearized problem $\mathbf{z} = (\mathbf{B} + \mathbf{R})^{-1}(\mathbf{o} - H(\mathbf{x}_b))$.
5. Post multiplication: project guess to the model columns required.
   goto 3 .
6. Post multiplication: project guess to the whole model grid.

Additional Staff for DWD 3D-Var

1 scientist	5 years	Assimilation of GPS data: integrated water vapour radio occultations
3 scientists	3 years	3D-Var: observation operators numerical solvers forecast error estimation

Impact of additional GPS integrated water vapour data



Integrated water vapour in the local model LM without (left) and with (right) assimilation of GPS data

Expectations to the GRAS SAF

- Test of less sophisticated observation operators.
  - ▷ *Is running a ray-tracing observation operator worth the effort?*
- What kind of observational errors shall be assigned to the bending angle observations?
  What are the errors of the bending angles provided by the CT method?
- Quality control: Can outliers be identified a priori?
- Data in real time.
  - ▷ *The time window for observations currently is $2\frac{1}{2}$ hours at 0 and 12 UTC.*