Error estimation for profiles retrieved from ground-based solar absorption FTIR measurements: PROFFIT Ver 9.3

By F. Hase

Motivation: Value of NDSC FTIR measurements relies heavily on specification of the quality of data products (trends, validation, …).

Standard error estimation by current profile retrieval codes (SFIT2, PROFFIT9) is poor. Activities to improve the situation underway (S. Wood, error estimation working group), still neither a separate error calculation tool nor an extended error estimation option in the codes available. I try to sound the latter option: PROFFIT93

Central strategic decision: Error estimation option in the retrieval code EEO vs. separate tool ST

- ST does not require any extensions in retrieval code
- ST can include any error sources difficult to include into retrieval code architecture (e.g. pressure broadening, pressure shift, …)

- ST can not start from the standard radiative transfer output + retrieval code output alone.
- Powerful ST is considerably more complex than the retrieval tool itself.
- ST requires own input, much overlap to original input. Interface??
- ST updating along with the retrieval code (shared input, line mixing…).

Decision: Go for EEO, accept limited capability …

General outline of error estimation:

OE- Retrieval equation:

$$\vec{x} = \vec{x}_a + \left(K^T S_\varepsilon^{-1} K + S_a^{-1}\right)^{-1} K^T S_\varepsilon^{-1} \left(\vec{y} - K\vec{x}_a\right) \quad \text{general}$$

$$\vec{x} = \vec{x}_a + \left(K^T K + \sigma^2 S_a^{-1}\right)^{-1} K^T \left(\vec{y} - K\vec{x}_a\right) \quad \text{uncorrelated noise}$$

Any error modifies the spectrum, thereby disturbs the retrieved quantity

$$\partial\vec{x} = \left(K^T S_\varepsilon^{-1} K + S_a^{-1}\right)^{-1} K^T S_\varepsilon^{-1} \partial\vec{y} = G_y \partial\vec{y}$$

Gain matrix G: sensitivity of retrieval wrt measurement error
(examples: noise, SZA, ILS, …)

Associated Covariance $S_x = E(\partial\vec{x}^T \partial\vec{x}) = G_y^T E(\partial\vec{y}^T \partial\vec{y}) G_y = G_y^T S_y G_y$

Smoothing error: A chimera?

Retrieval does not preserve details of true profile.

$$\vec{x}_{obs} = \vec{x}_a + \left(K^T S_\varepsilon^{-1} K + S_a^{-1}\right)^{-1} K^T S_\varepsilon^{-1} K\left(\vec{x}_{true} - \vec{x}_a\right) = \vec{x}_a + A\left(\vec{x}_{true} - \vec{x}_a\right)$$

$$\delta\vec{x}_{smooth} = \left(\vec{x}_{obs} - \vec{x}_a\right) - \left(\vec{x}_{true} - \vec{x}_a\right) \qquad S_{smooth} = (1-A)S_a(1-A)^T$$

BUT: True state and observed state (best estimate of the true state) reside in different spaces! Any true state is projected into an observed state, but no way back…

Smoothing error abets a dangerous misconception: observed state directly comparable with true state. Consequences:

- … "smoothing error is large, no need to take further errors into account"
- … direct comparison of "different" profiles (FTIR-in-situ, gb-satellite…)
- … mixing incompatible assumptions on S_a

Decision: Smoothing error not included! (Note: x_a and A essential)

Error sources taken into account by PROFFIT v 93:

- Offset in spectrum zero baseline
- SZA
- ILS (modulation + phase)
- Tilt (for each MW)
- Channeling (4 frequencies, cos + sin each)
- Spectroscopic errors / interfering species (intensity error)
- Temperature (concept of T-patterns)

  Note: retrieval + reporting in altitude domain! – Hydrostatic equilibrium needs readjustment. Analytical handling difficult, therefore retrieval is repeated for each T-pattern. Expensive, recalculation of absorption coeffs, but no T derivs required.

User interface:

PROFFIT switches to extended error mode if additional input file “errcalc.inp” is found in path. Required additional input given therein.

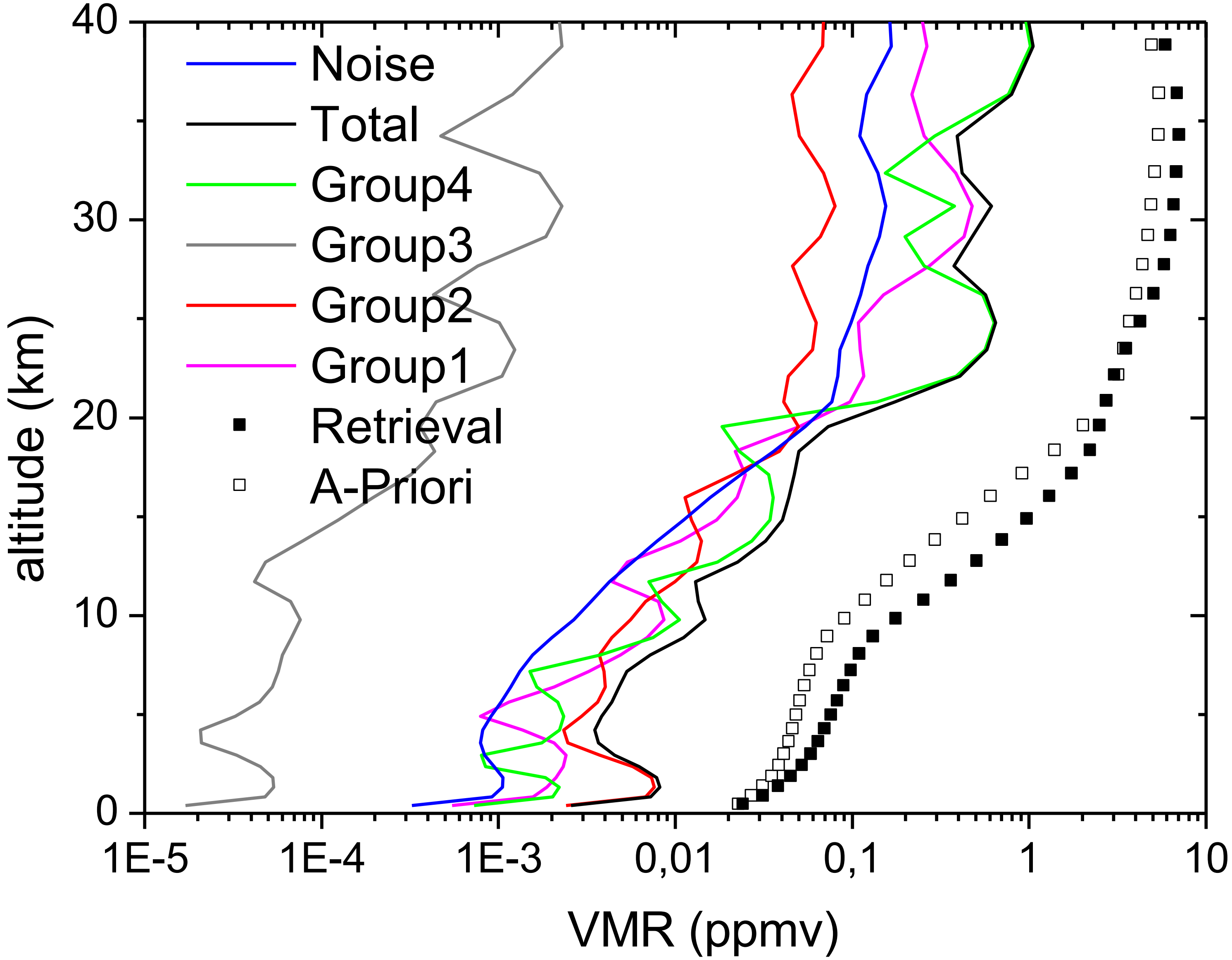
********** Target species for error calculation : #3
********** Channeling frequencies (4 frequencies, cm-1): 0.005 0.2 1.0 3.0
********** Assumed sds for: Offset (nW/(cm**2 sterad cm-1)   LOS (rad)   ILSmod (-)   ILSphas (rad): 1500.0 0.0015 0.04 0.01
********** Assumed sds for Tilt per MW (cm): 0.005 0.005 0.005 0.001
********** Assumed sds channeling (cos + sin): 0.002 0.002 0.002 0.002 0.002 0.002 0.002 0.002
********** Assumed sds interfering species (-): 0.02 0.02 0.0 0.01 0.01 0.01 0.01 0.01
********** Number of T-patterns: 3
********** T-patterns (nlevels x nTpatterns)
… … …

************ Number of partial columns: 2
************ for each partial column: level contribution (nlevels x nPartialColumns)
… …

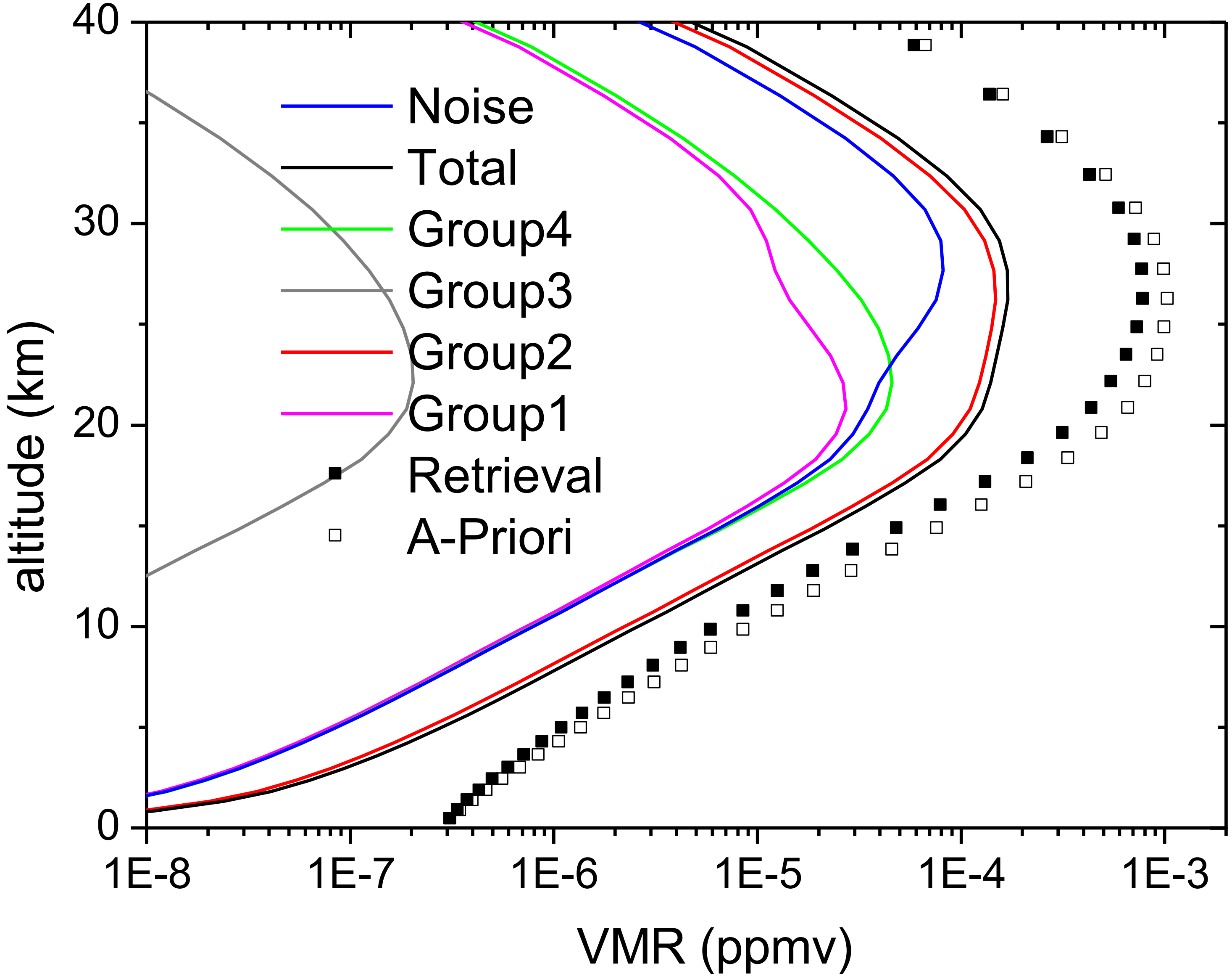
PROFFIT93 Output:

- Jakobians for error parms (offset, LOS, ILS, channel, tilt, spectr.)
- Individual error patterns for each parm (LOS: single pattern,
  channeling: 2 x NMWs x NChans, spectroscopy / interfering
  species NMWs x NSpezies, temperature NTpatt, …)
- Full covariance due to all error parms (incl. noise covariance)
- Error pattern groups: Gaussian summation of all patterns of the same kind (+ Grouping / Gaussian summation of different kinds of errors)
  Final groups:
  1: Zero baseline offset, LOS, ILS
  2: Channeling + tilt
  3: Spectroscopy / interfering species
  4: All T-patterns
  5: Noise
  6: Sqrt(Diagonal) of full covariance
- Full covariance for partial columns + sensitivities

Results: O_3 UFTIR, main isotope



Results: $ClONO_2$



Estimation of errors due to interfering species

Problem: Error due to interfering species strongly underestimated!

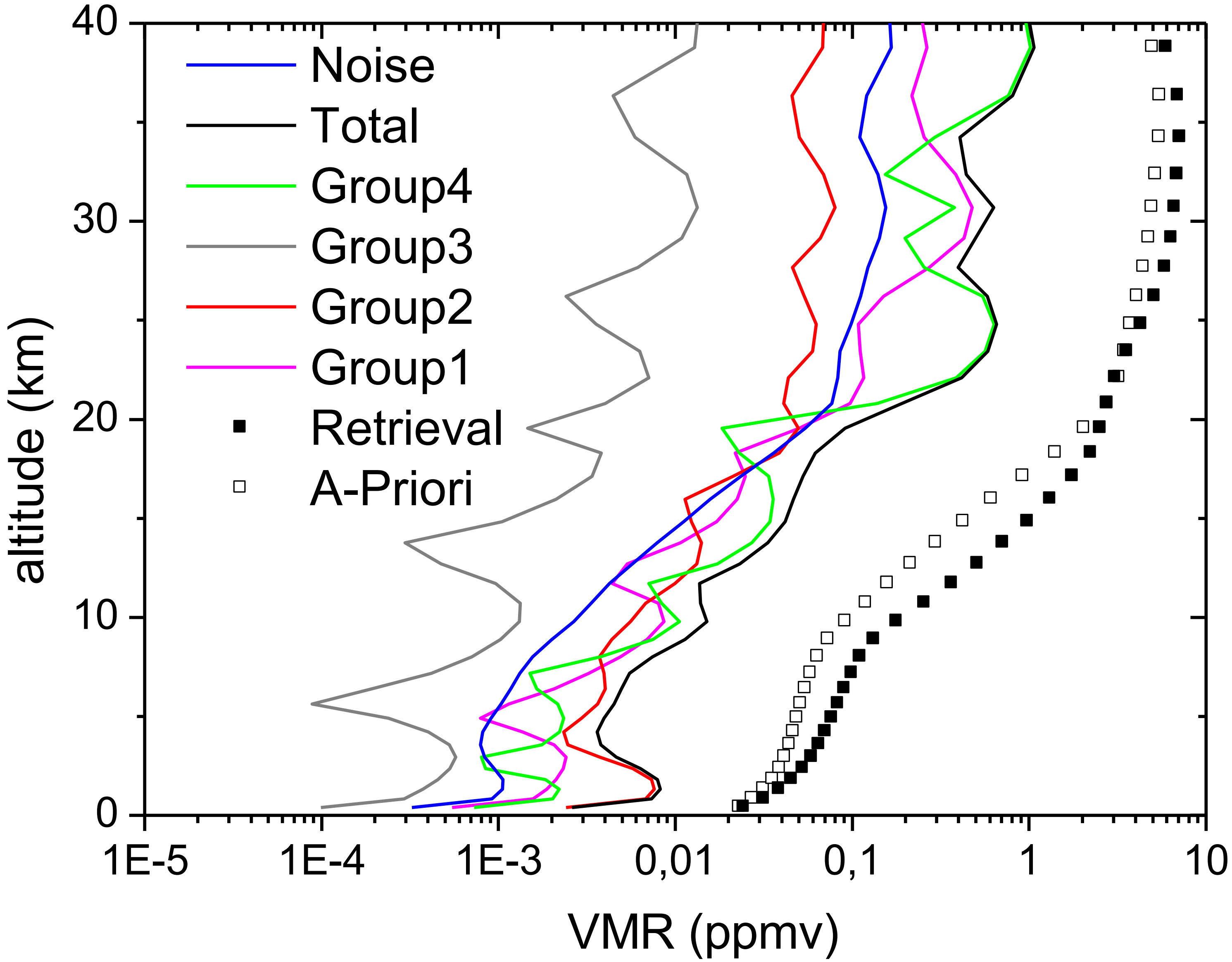
Reason: Retrieval strategy minimises impact of interfering species on target, gain matrix maps spectral errors mainly on interfering species

But we aim at something else …
…residual spectral features due to interfering species affecting target

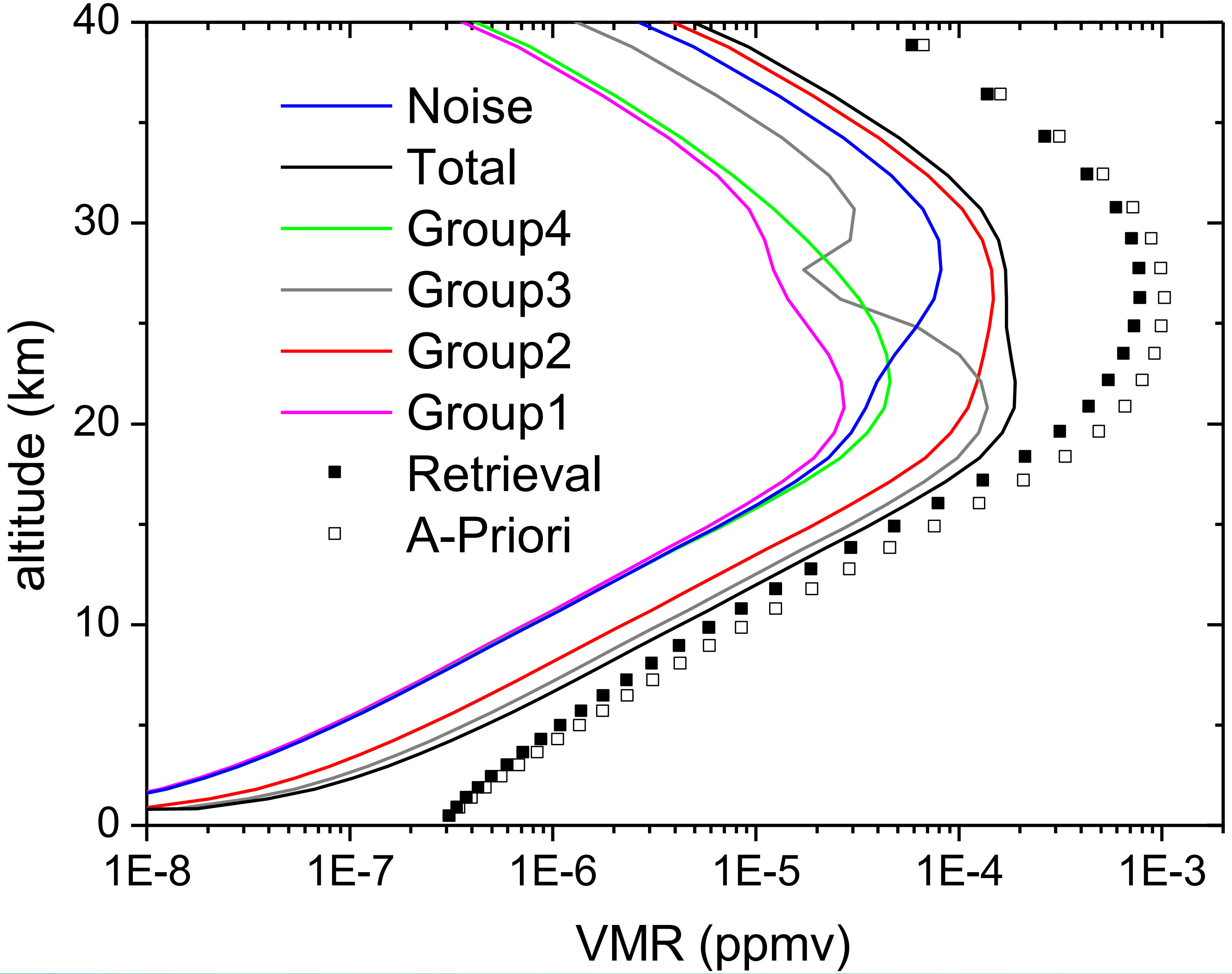
Solution: use modified gain matrix here, allow only fit of target species and aux parms (scale, tilt, shift, …)

$$\partial\vec{x} = \left(K^T S_\varepsilon^{-1} K + S_a^{-1} + C \cdot diag\left\{K^T K_{\text{species}\neq\text{target}}\right\}\right)^{-1} K^T S_\varepsilon^{-1} \partial\vec{y} = \widetilde{G}_y \partial\vec{y}$$

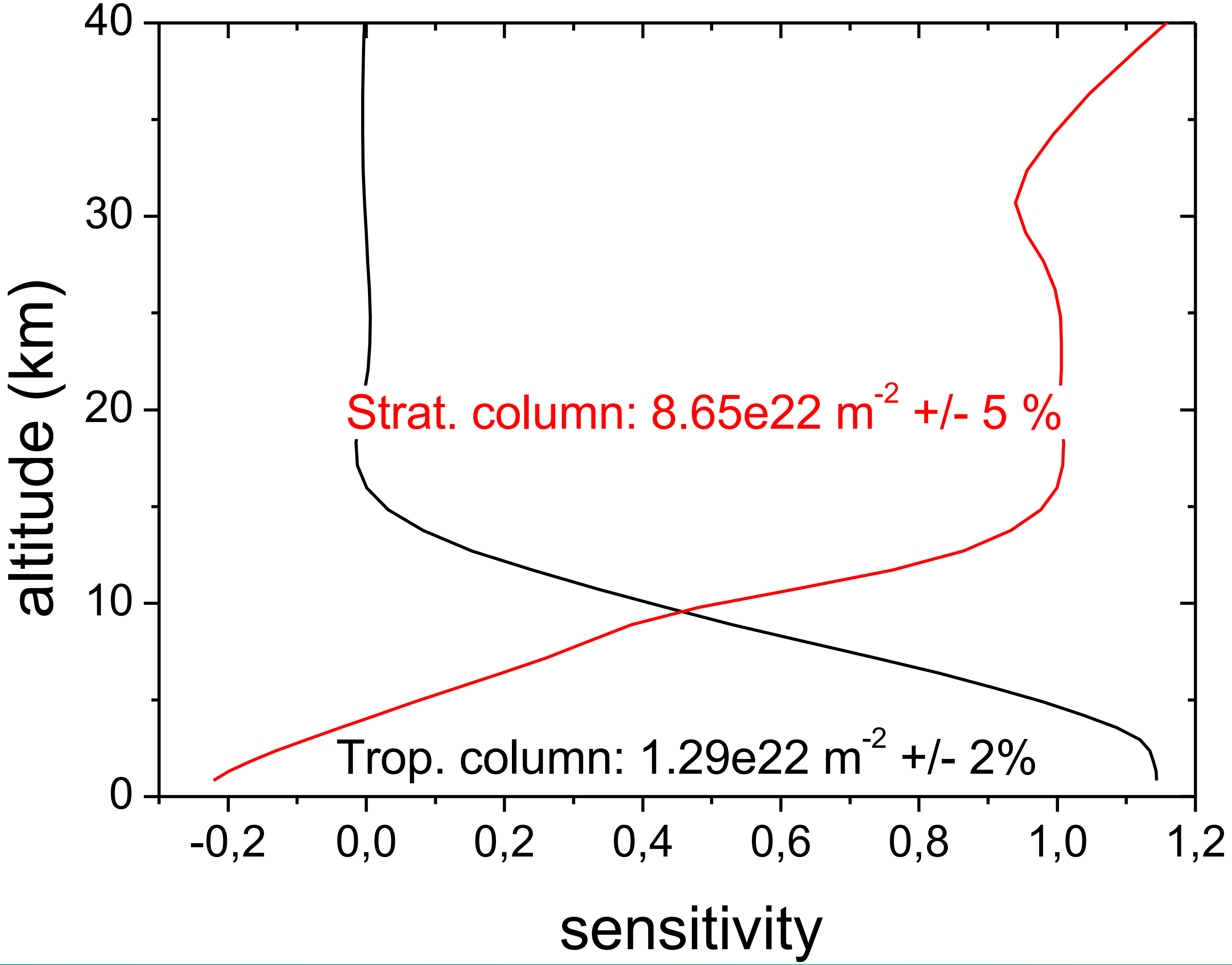
Results: O_3 UFTIR, main isotope



Results: $ClONO_2$



Results: O_3 UFTIR, main isotope



Summary:

- Proffit Ver. 93 offers extended error calculation capabilities
- Smoothing error not included (dangerous concept!)
- Noise error typically not the leading error source
- Offset, LOS, ILS, tilt, channeling: straightforward treatment
- T error more complicated: perturbation method, pattern concept
- Error due to interfering species: use of modified gain matrix
- Rigorous determination of spectroscopic / interfering species errors impossible in EEO approach

Note on Computational errors

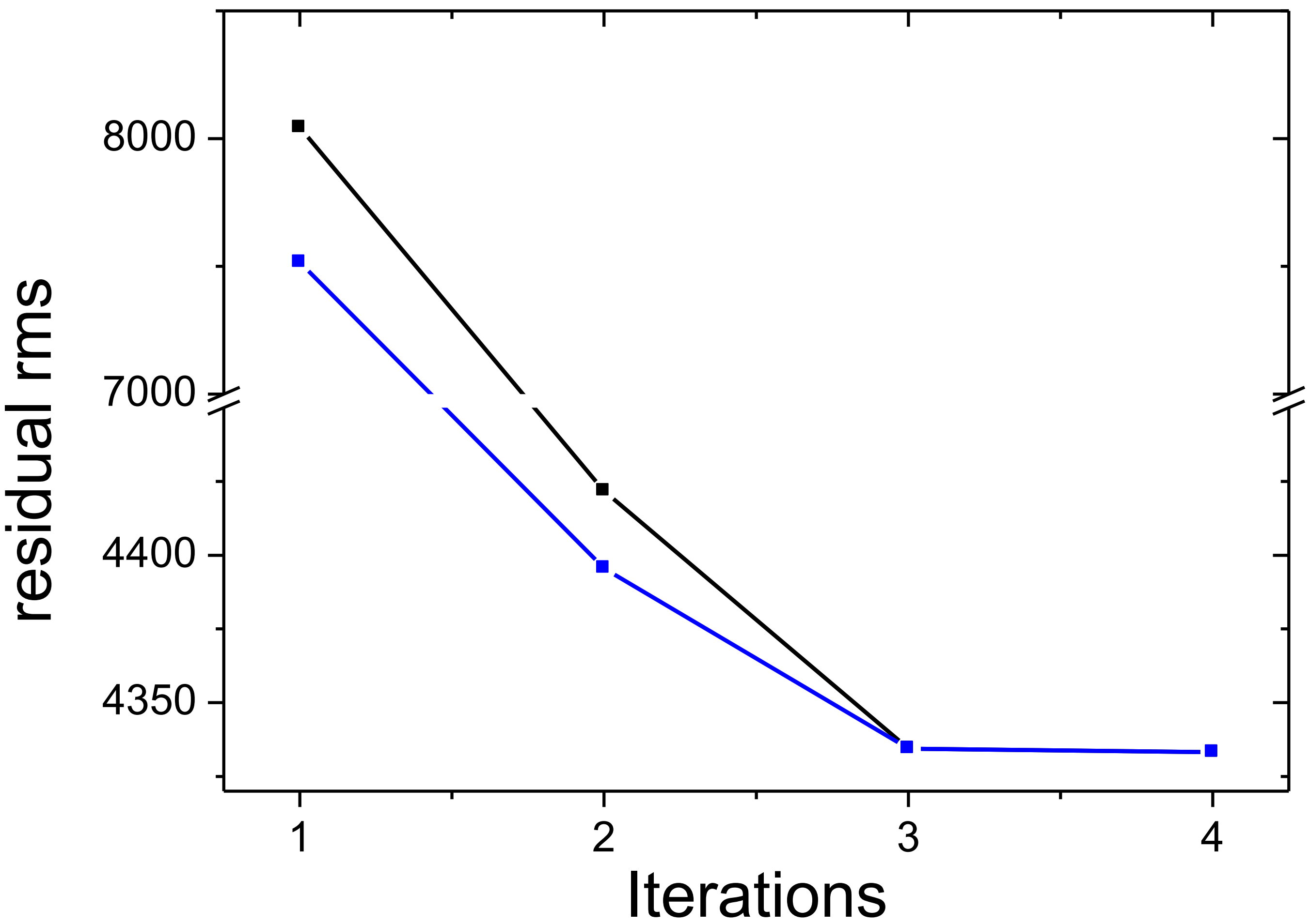
- Truncation error: can always be reduced to insignifcance …

- Limited accuracy of matrix inversion (*pointed out by A. Wiacek, UoT*)

$$\left(K^T K + \sigma^2 S_a^{-1}\right)^{-1}$$

Ratio of largest to smallest eigenvalue can become quite large!

Truncation



Computational error

$$M_{true} = \left(K^T K + \sigma^2 S_a^{-1}\right)^{-1}$$

$$M_{comp} = \left(K^T K + \sigma^2 (S_a^{-1} + E)\right)^{-1}$$

Computed matrix equivalent to additional (unphysical) constraints.
Results in coupling between independent components in state vector.
Computed solution + averaging kernels refer to total constraint.

Problem might remain undetected if stable solution is still reached!

Simple test of computational accuracy

Test original total cost function! Does reduction of constraint still improve solution?

Most critical run found by Aldona in case of PROFFIT: UFTIR O_3 retrieval (148 elements in state vector!).

Simple test of computational accuracy

Result: PROFFIT resides well in stable region for UFTIR O_3 retrieval
(constraint could be reduced by another factor 2^10…)