Atmos. Meas. Tech. Discuss., 5, 6577–6626, 2012 www.atmos-meas-tech-discuss.net/5/6577/2012/ doi:10.5194/amtd-5-6577-2012 © Author(s) 2012. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Atmospheric Measurement Techniques (AMT). Please refer to the corresponding final paper in AMT if available.

Improving retrieval quality for airborne limb-sounders by horizontal regularisation

J. Ungermann^{1,2}

 ¹Institute of Energy and Climate Research – Stratosphere (IEK-7), Research centre Jülich GmbH, Jülich, Germany
 ²now at: National Center for Atmospheric Research, Boulder, Colorado, USA

Received: 18 June 2012 - Accepted: 31 August 2012 - Published: 13 September 2012

Correspondence to: J. Ungermann (jorn@ucar.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.



Abstract

Modern airborne infrared limb-sounders are capable of measuring profiles so fast that neighbouring profiles are very similar to one another. This can be exploited by retrieving whole 2-D cross-sections instead of simple 1-D profiles. By adding horizontal regulari-

sation in addition to a potentially reduced vertical regularisation, vertical structures can be better retrieved while maintaining or reducing the general noise level.

This paper presents algorithms that are able to perform such a retrieval and efficiently produce typical diagnostic quantities. The characteristics of produced retrieval results for a variety of parametrisations is discussed in a case study that analyses a cross-

section measured by the CRISTA-NF instrument during the RECONCILE campaign between Spitsbergen and Kiruna, Sweden, in March 2010. It is shown that cross-section retrievals can either reduce noise or produce finer vertical structures while maintaining the same noise level. The presented methodology can also be applied in a straightforward way to improve the retrievals for both near-future satellite-borne limb-sounders and current air- and satellite-borne nadir sounder.

1 Introduction

20

Two-dimensional (2-D) cross-sections of trace gas volume mixing ratios derived from limb-sounders measurements taken from aircraft are a valuable tool for examining mixing processes in the upper troposphere/lower stratosphere. This region is of special importance, as it significantly influences radiative forcing (e.g. Forster and Shine, 1997); it also provides the boundary between tropospheric and stratospheric air, so its dynamic properties largely determine the exchange between these layers (e.g. Gettelman et al., 2011). To examine such dynamic structures, one would like to have a highly resolved 3-D image of the atmosphere including information on temperature, pressure, and trace gases of different life times. While there are promising instruments in development that



may deliver such highly resolved volumes (e.g. Friedl-Vallon et al., 2006; Ungermann et al., 2010), current limb imagers can only provide 2-D cross-Sects.

Using high-flying aircraft (such as the Russian M55 Geophysica or the German research plane HALO) as carriers, limb-sounders are typically mounted to be looking

- towards the side. This enables them to have an excellent resolution in flight direction (limited mostly by acquisition time), a good resolution in the vertical direction (limited by their vertical field-of-view; typically in the order of several hundred meters or more) and a much worse resolution perpendicular to the flight path (in the order of several hundred km) (e.g. Ungermann et al., 2012). The latter is obviously a problem, if one tries to rependent direction of several hundred by the seve
- solve atmospheric structures of smaller extent. However, modern (chemical) forecasts are often reliable enough to enable the aircraft to cross long-drawn filamentary structures orthogonally. Under such circumstances, be they achieved by meticulous flight planning or pure chance, the averaging along the line-of-sight becomes less important than the vertical and horizontal resolution.
- There are very thin layers in the atmosphere such as the tropopause inversion layer or tropopause folds with an extent of only 1 km or less (e.g. Birner, 2006; Hegglin et al., 2009). A vertical resolution of 500 m, already quite respectable for an limb-sounder, generates only one to two vertical data points for such structures and, for layers with even less vertical extent, may not reproduce them at all.

This paper describes a retrieval scheme that is able to produce 2-D cross-sections with both better vertical resolution and generally lower noise. Such cross-sections can be more easily analysed due to the consistency between neighbouring profiles and lower smoothing in the vertical directions. By exploiting the high frequency with which modern limb-sounders are able to acquire profiles, it is possible to use the similarity of

neighbouring profiles to increase the vertical resolution at a slightly reduced resolution in the along-flight track direction. This paper introduces the concept of cross-section retrievals, where a whole cross-section consisting of many individual profiles is jointly retrieved in the usual non-linear fashion. This is in principal similar to the concept of tomographic 2-D retrievals of satellite limb sounder data, which were first produced by



Carlotti et al. (2001) and Steck et al. (2005) for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) and by Livesey et al. (2006) for the Microwave Limb Sounder. However, the cross-section retrievals presented here do lack the tomographic component, as consecutively measured radiance profiles do measure different armasses. Also, the intent is different as tomographic satellite retrievals are primarily

⁵ airmasses. Also, the intent is different, as tomographic satellite retrievals are primarily used to increase the along line-of-sight resolution, while the cross-section retrieval presented here shall instead increase the vertical resolution of the retrieval result at the cost of a reduced along flight-track resolution.

The retrieval of trace gases or other quantities from limb-sounder measurements is inherently an ill-posed problem, implying that no or no unique solution may exist and that small measurement errors may cause large differences in retrieval values. One way to obtain a unique and physically reasonable solution is regularisation. Regularisation usually approximates the original ill-posed problem by a similar well-posed problem. This is typically accomplished by adding constraints to the solution. From

- ¹⁵ here on, the term horizontal refers solely to the along-flight track direction. By introducing horizontal regularisation, that is constraints on the horizontal variation of retrieved values, the influence of stochastic error sources becomes reduced, allowing for a reduction of vertical regularisation strength. Further, this paper introduces algorithms to provide an associated set of typical diagnostic information following the linear frame-
- ²⁰ work of Rodgers (2000). Last, the feasibility of the approach is demonstrated upon analysing measurements of the Cryogenic Infrared Spectrometers and Telescope for the Atmosphere–New Frontiers (CRISTA-NF) taken during the RECONCILE campaign (see Ungermann et al., 2012). The effect of horizontal regularisation on the retrieval quality, the impact of noise, and the achievable horizontal resolution is examined and
- discussed. It is demonstrated that horizontal regularisation allows lower vertical regularisation strength in order to reproduce finer features than otherwise possible. Last, this more complicated method is compared against the weighted a posteriori averaging of conventionally retrieved 1-D profiles.



2 Cross-section retrieval

The retrieval is the process of deriving atmospheric quantities such as temperature or trace gas volume mixing ratios from remote sensing measurements. In the case of in-frared limb-sounder measurements, it is an ill-posed inverse problem. This implies that

no map exists that reliably delivers the real atmospheric state given only the measurements. In practise, a discrete representation of the atmospheric state *x* ∈ ℝⁿ is modified under certain constraints until the fit between the real measurements *y* ∈ ℝ^m and simulated measurements *F*(*x*) produced by a forward model *F* : ℝⁿ → ℝ^m is deemed good enough according to some quality criterion. Thereby *F* is a sufficiently accurate forward model capable of simulating the physics of radiative transport and measurement principle. As the problem is ill-posed, the addition of some kind of a priori information is usually required to generate physically reasonable results. This is formalised by introducing a cost function *J* : ℝⁿ → ℝ:

$$J(\mathbf{x}) = (F(\mathbf{x}) - \mathbf{y})^T \mathbf{S}_{\varepsilon}^{-1} (F(\mathbf{x}) - \mathbf{y}) + (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a).$$

- ¹⁵ This specific cost function is a quadratic form, which has the advantage that its minimum can be found rather efficiently. It consists of two distinct terms: the first term penalises differences between the measurements and the radiances provided by the forward model weighted by the measurement error covariance matrix $\mathbf{S}_{e} \in \mathbb{R}^{m \times m}$. This matrix describes the available knowledge about measurement errors, thereby assum-
- ²⁰ ing that they are Gaussian distributed. The second term penalises differences between the derived atmospheric state and so called a priori knowledge, which is usually taken from a climatology, from models, or from ad hoc assumptions about smoothness of atmospheric quantities. Thereby x_a is the expected atmospheric state and $S_a \in \mathbb{R}^{n \times n}$ is a covariance matrix thereof codifying correlations between atmospheric quantities.



(1)

A minimum for the cost function J can then be found using, e.g. an iterative truncated Quasi-Newton method (often within less than 10 iterations) (e.g. Nocedal and Wright, 2006):

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i - \left(\mathbf{S}_a^{-1} + \mathbf{F}'(\boldsymbol{x}_i)^T \mathbf{S}_e^{-1} \mathbf{F}'(\boldsymbol{x}_i) + \lambda_i \mathbf{I}_n \right)^{-1} \cdot \left(\mathbf{S}_a^{-1}(\boldsymbol{x}_i - \boldsymbol{x}_a) + \mathbf{F}'(\boldsymbol{x}_i)^T \mathbf{S}_e^{-1} \left(\boldsymbol{F}(\boldsymbol{x}_i) - \boldsymbol{y} \right) \right)$$
(2)

5

The Levenberg–Marquardt parameter λ_i is chosen in each iteration to decrease the cost function and should go to zero in the limit. The matrix \mathbf{I}_n is the identity matrix in $\mathbb{R}^{n \times n}$ and $\mathbf{F}'(\mathbf{x}_i) \in \mathbb{R}^{m \times n}$ is the Jacobian matrix of F evaluated at \mathbf{x}_i .

¹⁰ This framework describes two popular regularisation techniques. First, optimal estimation (e.g. Rodgers, 2000) that employs an expected state vector x_a from climatological sources and a covariance matrix using variances and covariances from similar sources. Correlations are often chosen in an ad hoc manner or according to model simulations, as reliable true correlations are typically hard to come by. Second, Tikhonov ¹⁵ regularisation (e.g. Tikhonov and Arsenin, 1977), which is basically a superset of optimal estimation. When employed, the matrix S_a^{-1} is often designed to not contain a zeroth moment, implying that only the shape, but not the mean value of the retrieved profile is biased.

This mathematical framework fits well to our approach of cross-section retrievals. The basic difference to conventional profile retrievals is just the exchange of the 1-D representations of the atmospheric state with a 2-D state vector and the addition of horizontal correlation or smoothness conditions to the regularisation.

The remainder of this section discusses technical algorithmic and implementation details necessary to perform cross-section retrievals in practise. First, a specific regu-

²⁵ larisation is presented that is suitable for large-scale retrievals. Second, it is discussed how to practically derive the effects of various error sources on retrieval results for large-scale retrievals. The last subsection presents a novel stochastic approach to



diagnostics very suitable for large-scale retrievals as its computational cost does increase only with the square instead of the third power of the number of unknowns.

2.1 Regularisation

For large scale retrievals, Tikhonov regularisation is very convenient, if not required, as it allows to directly define the needed precision matrix (that is the inverse of the covariance matrix) $\mathbf{S}_{a}^{-1} \in \mathbb{R}^{n \times n}$:

$$\mathbf{S}_{a}^{-1} = (\alpha_{0})^{2} \mathbf{L}_{0}^{T} \mathbf{L}_{0} + (\alpha_{1}^{h})^{2} \mathbf{L}_{1}^{h^{T}} \mathbf{L}_{1}^{h} + (\alpha_{1}^{v})^{2} \mathbf{L}_{1}^{v^{T}} \mathbf{L}_{1}^{v} +$$

The parameters α_0 , α_1^h , and α_1^v allow to weigh the different regularisation terms differently. A good starting point is to set all to 1, even though this paper chooses a smaller α_0 (see Table 2).

Having a sparse precision matrix makes it very inexpensive to set the matrix up and to multiply it with vectors during the solving of equation systems using iterative algorithms such as conjugate gradients. But this poses severe restrictions on which covariance matrices are admissible for the problem, preventing a pure optimal estimation approach, which usually employs arbitrary covariance matrices often derived from

3-D models.

10

As compromise, the Tikhonov regularisation used in this paper is chosen to approximate the precision matrix of an optimal estimation covariance matrix employing the auto-regressive model to fill the covariances Steck and von Clarmann (e.g. 2001). In ²⁰ the parametrisation used in this paper, $L_0 \in \mathbb{R}^{n \times n}$ is a diagonal matrix, with climatological standard deviations on the diagonal. $L_1^h \in \mathbb{R}^{n \times n}$ and $L_1^v \in \mathbb{R}^{n \times n}$ are matrices that by multiplication with the atmospheric state vector calculate the first derivative in horizontal and vertical direction of the continuous atmospheric state by means of finite differences. For convenience, the rows containing only zeros have not been removed ²⁵ from the matrices, implying that the resulting vectors containing the derivatives also

contains "dummy" zero-entries. This has no further effect as these added zeros do not



(3)

change the value of the cost function. In addition, each row of the smoothing matrices is scaled with both the standard deviation of the given quantity and two quantity specific scaling factors c_q^h and c_q^v for \mathbf{L}_1^h and \mathbf{L}_1^v , respectively. The latter scaling factors allow to modify the regularisation strength in the different spatial directions or to turn it

⁵ off altogether; for example, using a correlation length of zero in horizontal direction results in the same retrieval result as would be generated by the simple assembly of the profiles produced by individual 1-D retrievals. The scaling with the standard deviations makes this regularisation similar to optimal estimation, introduces a simple altitudedependent regularisation, and also normalises the regularisation to the extent that the scaling factors c_a^h and c_a^v are of similar magnitude for all involved trace gases.

2.2 Linearised diagnostics

The diagnostics used in this work follow the linearised diagnostics described by Rodgers (2000). The underlying idea is to insert the minimum of the cost function x_f as fix point into Eq. (2) and to replace the vector of measurements y by a first order Taylor approximation around the true atmospheric state $x_T \in \mathbb{R}^n$. This gives a linear relation between the retrieval result, the (unknown) true atmospheric state and the measurement errors $\boldsymbol{\varepsilon} \in \mathbb{R}^m$:

$$\boldsymbol{x}_f = \boldsymbol{A}\boldsymbol{x}_T + (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x}_a + \boldsymbol{G}\boldsymbol{\epsilon}$$

Thereby, the matrix $\mathbf{G} \in \mathbb{R}^{n \times m}$ is the so-called gain matrix mapping differences in the measurements onto differences in the retrieval result. It is defined as

$$\mathbf{G} = \left(\mathbf{S}_{a}^{-1} + \mathbf{F}'(\mathbf{x}_{f})^{T}\mathbf{S}_{c}^{-1}\mathbf{F}'(\mathbf{x}_{f})\right)^{-1}\mathbf{F}'(\mathbf{x}_{f})^{T}\mathbf{S}_{c}^{-1}.$$

The matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ with

 $\mathbf{A} = \mathbf{GF}'(\mathbf{x}_f)$

15

Discussion Paper AMTD 5, 6577-6626, 2012 Improving retrieval quality for airborne limb-sounders **Discussion** Paper J. Ungermann **Title Page** Abstract Introduction Conclusions References **Discussion** Paper Tables Figures Back Close Full Screen / Esc **Discussion** Paper Printer-friendly Version Interactive Discussion

(4)

(5)

(6)

is the averaging kernel matrix that describes the influence of the true atmospheric state on the retrieval result. By changing the regularisation matrix, it is possible to change the impact of the different terms of Eq. (4) on the result.

- One can analyse the averaging kernel matrix to determine how the applied regularisation influences the retrieval result and to quantify the spatial resolution of the result. The gain matrix is also instrumental in estimating the effect of various errors on the retrieval result. With $\mathbf{S} \in \mathbb{R}^{m \times m}$ being the covariance matrix of an error in the measurements, the matrix \mathbf{GSG}^T describes the resulting uncertainty in the retrieval result. Such an error covariance matrix \mathbf{S} can be generated (if only in an ad-hoc manner) for most errors affecting the retrieval, e.g. for the uncertainty of not-retrieved background gases or uncertainties with respect to spectral line strengths (see below). Obviously, many er-
- ror sources are not of Gaussian nature and can thus only be approximated to a certain extent by this approach.
- Providing the diagnostic information for a cross-section retrieval is slightly more com-¹⁵ plicated than for 1-D retrievals, as the involved matrices are of a size that is neither easy to keep in memory without resorting to sparse storage techniques nor straightforward invertible. The inversion is thereby both computationally demanding and, even worse, numerically unstable (scaling the diagonal of the matrix to one and using an LU-decomposition as done by Livesey et al. (2006) works mostly well for the presented ²⁰ type of problem up to a certain size of \approx 40 000 unknowns).

We therefore follow a different approach that avoids all matrix inversions or factorisations. This drastically reduces the amount of required memory at the cost of some additional computation costs. The basic idea is to calculate only a single row of **A** and **G** at a time by solving a linear equation system and to derive all relevant diagnostic quantities for the associated atmospheric quantity from these matrix rows. Doing so for all points in turn is trivially to parallelise on supercomputers.

25

To calculate the *i*-th row of **G** the *i*-th row of $\left(\mathbf{S}_{a}^{-1} + \mathbf{F}'(\mathbf{x}_{f})^{T} \mathbf{S}_{e}^{-1} \mathbf{F}'(\mathbf{x}_{f})\right)^{-1}$ is calculated by solving the equation system $\left(\mathbf{S}_{a}^{-1} + \mathbf{F}'(\mathbf{x}_{f})^{T} \mathbf{S}_{e}^{-1} \mathbf{F}'(\mathbf{x}_{f})\right) \mathbf{s}_{i} = \mathbf{e}_{i}$ with \mathbf{e}_{i} being the *i*-th



unity vector. Naturally, iterative methods are used for this to avoid the calculation of any matrix-matrix product or the generation of dense matrices. The resulting vector s_i is then multiplied with the appropriate matrices to generate the *i*-th row g_i of **G** and the *i*-th row a_i of **A**.

⁵ Using this, one can directly calculate the variance of influence of measurement noise on the retrieval result by evaluating $\boldsymbol{g}_{i}^{T} \boldsymbol{S}_{e} \boldsymbol{g}_{i}$.

For deriving the influence of unknown background gases, one first needs to assemble a covariance matrix S_q describing the state of this gas. This can then be multiplied from both sides with the Jacobian matrix $\mathbf{F}'_{a}(\mathbf{x}_{f})$ of the forward model with respect to this gas to derive a covariance matrix for the influence of this gas on the (simulated) measurements. Assembling a full covariance matrix \mathbf{S}_{q} by, e.g. using the auto-regressive approach (that is $(\mathbf{S}_q)_{i,i} = \sigma_i \sigma_j e^{(v_i - v_j)(c_g^v)^{-1}} e^{(h_i - h_j)(c_g^h)^{-1}}$ with v_i and h_i being vertical and horizontal coordinates of data point i) and some horizontal and vertical correlation lengths is too costly both with respect to memory consumption and computation time. In this case, it is much more efficient to analytically define the (sparse) precision matrix 15 \mathbf{S}_{a}^{-1} instead (e.g. Steck and von Clarmann, 2001; Ungermann, 2011, 134 pp.), which can be done very efficiently. Using the precision matrix instead of the covariance matrix requires the solution of another equation system to generate the desired result. The standard deviation σ_i^g of the influence of the gas on the retrieved quantity i can then be deduced as: 20

$$\left(\sigma_{i}^{g}\right)^{2} = \boldsymbol{x}_{f}^{T} \boldsymbol{\mathsf{F}}_{g}^{\prime}(\boldsymbol{x}_{f})^{T} \boldsymbol{z}, \text{ with } \boldsymbol{\mathsf{S}}_{g}^{-1} \boldsymbol{z} = \boldsymbol{\mathsf{F}}_{g}^{\prime}(\boldsymbol{x}_{f}) \boldsymbol{g}_{i}.$$

This can be done very efficiently using iterative methods as only a couple of diagonals of the precision matrix are non-zero. The effort is therefore typically negligible compared to the effort to calculate the vector g_i itself.

25

The influence of uncertainties of spectral line data of one specific trace gas can be estimated in a similar fashion. Under the assumption that the error is fully correlated, the corresponding covariance matrix $\mathbf{S}_{g}^{\text{spec}}$ is fully determined by its standard deviations 6586



(7)

 $s_g (\mathbf{S}_g^{\text{spec}} = s_g s_g^T)$, which are set to the relative spectral error times the local gas volume mixing ratios (this assumes that the relative spectral error is identical over all used IMWs; should this not be the case, it is straightforward to calculate these individually for each IMW, but more complicated in notation and therefore omitted here). Again, it is too costly to assemble the matrix $\mathbf{S}_g^{\text{spec}}$, especially as the equation is easily simplified. Thus, the standard deviation $\sigma_i^{\text{spec},g}$ induced by uncertainties in spectral characterisation of gas g can be calculated as:

$$\left(\sigma_{i}^{\text{spec},g}\right)^{2} = \boldsymbol{g}_{i}^{T}\boldsymbol{\mathsf{F}}_{g}'(\boldsymbol{x}_{f})^{T}\left(\boldsymbol{s}_{g}\boldsymbol{s}_{g}^{T}\right)\boldsymbol{\mathsf{F}}_{g}'(\boldsymbol{x}_{f})\boldsymbol{g}_{i} = \left(\boldsymbol{s}_{g}^{T}\boldsymbol{\mathsf{F}}_{g}'(\boldsymbol{x}_{f})\boldsymbol{g}_{i}\right)^{2}$$
(8)

In this fashion, all relevant diagnostic quantities can be calculated using only sparse matrices and iterative algorithms. The resulting quantities are identical to the diagnostic quantities that are derived by straightforward matrix inversion, if feasible. The sole disadvantage of this approach is that the effort to calculate the diagnostics for all retrieved quantities quickly surpasses the effort for the retrieval itself. A more efficient way to deduce error bars and to estimate the influence of individual error terms on the full set of retrieved quantities is presented in the next Section.

2.3 Monte Carlo diagnostics

20

Deriving the diagnostics for each individual retrieved quantity requires in the order of $O(n^3)$ operations, as for each of the *n* retrieved elements of the atmospheric state vector, a linear *n*-by-*n* equation system must be solved. This assumes that a fixed number of iterations smaller than *n* produces a sufficiently accurate solution in the iterative solver.

One way of reducing the computational cost is to use stochastic algorithms. Such algorithms can also be used to produce solutions and associated distribution, or to estimate the errors in a non-linear way (e.g. Tarantola, 2004); here, only a simple linear error estimate is desired.



The previous section introduced a number of error effects and associated covariance matrices. Using these Gaussian characterisations of systematic and stochastic errors influencing the measurements, it is easy to generate random errors according to the given distributions by multiplying a root of the covariance matrix with a vector of Gaus-

- sian random variables with standard deviation of one. Such roots are best generated from the sparse precision matrix using the Cholesky-Banachiewicz algorithm, as computing them from the covariance matrix itself is generally too computationally costly (see Ungermann, 2011, p.136). These error terms can then be added to the actual measurements and be fed into a linear retrieval to produce deviating solutions. The
- differences between the actual retrieval result and the outcomes of the linear retrievals with disturbed inputs can the be analysed to derive the error bars. As the involved distributions are assumed to be Gaussian and as the retrieval is linearised for this purpose, the resulting distributions are also Gaussian, simplifying the mathematics for all estimates significantly. Depending on the required accuracy for the error bars, a couple of dozen solved *n*-by-*n* linear equation systems deliver already usable estimates.
- Let $e_i^{mc} \in \mathbb{R}^m$, $1 \le i \le N$, be *N* error vectors generated according to an multivariate Gaussian error distribution, e.g. of the noise error. Usually, one is interested especially in the total error (the sum over all individual errors), for which an error vector according to the joint distribution of all errors may be generated. Given this set of error vectors, a linearised retrieval generates a set of associated solutions x_i^{mc} :

$$\boldsymbol{x}_{j}^{\mathrm{mc}} = \boldsymbol{x}_{f} - \left(\boldsymbol{\mathsf{S}}_{\mathrm{a}}^{-1} + \boldsymbol{\mathsf{F}}'(\boldsymbol{x}_{j})^{T} \boldsymbol{\mathsf{S}}_{\varepsilon}^{-1} \boldsymbol{\mathsf{F}}'(\boldsymbol{x}_{j})\right)^{-1} \cdot \left(\boldsymbol{\mathsf{F}}'(\boldsymbol{x}_{j})^{T} \boldsymbol{\mathsf{S}}_{\varepsilon}^{-1} \boldsymbol{\varepsilon}_{j}^{\mathrm{mc}}\right).$$
(9)

The standard deviation of these Monte Carlo solutions delivers the desired error estimate. According to Kenney and Keeping (1951), the standard deviation can be estimated as

25
$$\sigma^{\rm mc} = \frac{1}{C_4(N)} \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i^{\rm mc} - \bar{x}_i^{\rm mc})^2}.$$



(10)

Thereby, the correction factor

$$c_4(N) = \sqrt{\frac{2}{N-1}} \frac{\Gamma(\frac{N}{2})}{\Gamma(\frac{N-1}{2})}$$

makes this estimator unbiased also for small N (the typically used correction factor of dividing by N - 1 instead of N within the square root of Eq. (10) removes only the bias for the estimate for the variance, which causes noticeable discrepancies between deterministically calculated and stochastically estimated standard deviations of retrieval results for very small values of N). The error of the estimate decreases with the number of samples. Kenney and Keeping (1951) also provide an estimate for the standard deviation of the standard deviation estimate as

10 std(
$$\sigma^{\text{mc}}$$
) = $\sigma^{\text{mc}} \sqrt{\frac{1}{N} \left(N - 1 - \left(\frac{2\Gamma(\frac{N}{2})}{\Gamma(\frac{N-1}{2})}\right)^2\right)}$, (12)

Please note that we distinguish here and in the following between the estimate of the standard deviation of the retrieval result (giving the error of the retrieval result) and the standard deviation of the estimate itself (giving the error of the error estimate).

Equation (12) shows that an exemplary 32 random samples give an estimate of the standard deviation with a one-sigma relative error of about 12.4 %. Thus one will almost certainly encounter outliers if many points are diagnosed. A quadrupling of the number of samples roughly halves the relative error. In our implementation and for the example presented later on, it is possible to calculate roughly 128 linear solutions in the same time that is required for the non-linear retrieval itself, delivering an accuracy of about

6 %. Please note that this is the accuracy of the estimated standard deviation. For example, if the estimated standard deviation points to a 10 % uncertainty for a certain retrieved trace gas (not uncommon for the CRISTA-NF instrument), a relative error of 6 % translates to a one-sigma uncertainty of the retrieval result ranging from 9.4

(11)

to 10.6%. This uncertainty in the error is well within the range of other uncertainties introduced by instrument characterisation or spectral line data.

If one is only interested in a worst-case estimate, for example to determine the relevance of different error sources, it is advantageous to use a very small number of samples to generate a coarse estimate of the standard deviation and add to this one (or multiple) standard deviations of the estimate to derive a reliable upper bound for the error.

Figure 1 shows the converging behaviour for a real diagnosis of the total error for a range of sampling sizes. Shown as yellow circles are the means of the relative error of the standard deviation estimate over all retrieved quantities of one retrieval. Shown as blue circles are the real standard deviation of the estimates generated by comparison of the estimates with deterministically calculated errors. Shown as a dotted horizontal line is the theoretical standard deviation of the standard deviation estimate. As expected for a large number of estimates, the minimum and maximum errors are rather large compared to a single standard deviation. For a small number of samples such as 32 samples, relative errors of the standard deviation estimate of up to 50 % are not uncommon. However, by using more samples, also the maximal errors get reduced

in lockstep with the standard deviation of the estimate.
 This method lends itself well to parallelisation. Each Monte Carlo solution may be
 generated in parallel to the others on an arbitrary number of cores. It is also possible to refine the estimate later on. One may combine the sample mean and sample variance of different sample runs to improve upon the estimate also at a later time, when more processing power may be available or the need for a more accurate estimate has arisen.

As the number of samples for a desired accuracy is independent of the number of unknowns, this algorithm has an effective cost of only $O(n^2)$, albeit with a large constant coefficient. Still, this algorithm is much more efficient than deterministic methods in deriving error bounds for cross-section retrievals and especially for tomographic retrievals (e.g. Carlotti et al., 2001; Ungermann et al., 2010) due to the large amount of



involved unknowns. Thus, this method provides an efficient way to provide diagnostic information for massive scale retrievals.

3 Case study

This section presents a case study that examines the differences between cross-⁵ section retrievals and conventional 1-D retrievals.

The RECONCILE campaign took place from January to March 2011. The high-flying Russian research aircraft M55-Geophysica performed 12 flights in this time frame with a variety of in situ and remote sensing instruments. One of the instruments aboard was CRISTA-NF, an airborne infrared limb-sounder. Its measurements during one flight of
the campaign shall function as an example. The one-dimensional retrieval and validation of this dataset was described by Ungermann et al. (2012). This data set is rather unique in having at the same time a high frequency of taken profiles (one profile every ≈ 15 km) and a high vertical sampling (≈ 250 m). The impact of stochastic errors on the measurements is rather high, first from actual measurement noise introduced by the
detectors and associated components, second from uncertainties in the vertical point-

ing of individual spectra (and even spectral samples within one spectra). The second flight of 2 March 2011 has been selected as example as it offers the most vertical and horizontal structure.

3.1 Instrument

CRISTA-NF is an air-borne infrared passive limb-sounder. It measures emissions of thermally excited atmospheric trace gases in the mid-infrared region (4 to 15 μm). The CRISTA-NF viewing direction is perpendicular to the flight direction of the M55-Geophysica. Spectra are scanned from the flight altitude down to 15 km below this level in vertical steps of ≈ 250 m using a Herschel telescope with a tiltable mirror. A detailed description of the optical system and the cryostat of CRISTA-NF is given by



Kullmann et al. (2004). The optical system of CRISTA-NF has been re-purposed from the CRISTA experiment that was launched by the Space Shuttle on the Shuttle Pallet Satellite SPAS in November 1994 (STS 66) and August 1997 (STS 85) (Offermann et al., 1999; Grossmann et al., 2002). Obtaining a single spectra takes only 1.2 s due

⁵ to the low temperature of the detectors and the whole optical system of the instrument. While the pointing stability of the instrument had been improved upon compared to previous campaigns, its pointing stability is still only in the order of 0.02°, which corresponds to an uncertainty in the tangent point of up to ≈ 100 m vertically. This corresponds roughly to a relative error in measured radiation of ≈ 1 %.

10 3.2 Retrieval processor

15

The Jülich Rapid Spectral Simulation Code Version 2 (JURASSIC2) is used as forward model and retrieval processor. The Python/C++ based JURASSIC2 and its predecessor were used in several experiments and studies (e.g. Hoffmann et al., 2008; Ungermann, 2011). It uses pre-calculated tables of spectrally averaged optical paths to avoid costly line-by-line calculations. Combining the radiances derived from emissivity growth approximation method (EGA; e.g. Weinreb and Neuendorffer, 1973; Gordley and Rus-

- sell, 1981) and the Curtis-Godson Approximation (CGA; Curtis, 1952; Godson, 1953) using a simple regression scheme (Francis et al., 2006), the standard deviation between full line-by-line calculations and the approximate scheme used by JURASSIC2
- ²⁰ is less than 0.22% (Ungermann et al., 2012). This is less than the assumed stochastic error of 1% and much less than assumed uncertainties in spectral line data and parametrisation schemes.

3.3 Retrieval setup

The discussed retrieval derives the primary retrieval targets CCl₄, CFC-11, H₂O, HNO₃,

²⁵ O₃, and ClONO₂ and the secondary retrieval targets aerosol, temperature, PAN, CFC-113, and HCFC-22 from 13 integrated micro windows (see Table 1). The primary



distinction between primary and secondary retrieval targets is that the signal of the primary targets is strong enough to deliver reliable results also in the presence of systematic errors in background gases. The retrieval grid sampling distance is 250 m below 20 km and 1 km above 20 km. All targets are derived between 0 km and 25 km except

- for the trace gases HNO₃, O₃, and CIONO₂ that are subject to an increase in volume mixing ratio above flight altitude. These gases are therefore derived up to an altitude of 65 km. The regularisation is parametrised according to Table 2. A priori information and initial guess are chosen to be identical. For temperature, pressure and water vapour ECMWF ERA-Interim data is used. A zero profile is used for PAN. For the remaining
- ¹⁰ targets appropriate profiles are taken from the climatology assembled by Remedios et al. (2007). All standard deviations used for the regularisation are also taken from this source except for temperature (for which a fixed standard deviation of 1 K was chosen) and PAN (for which the climatology of Glatthor et al., 2007 is used). Additional information about the 1-D retrieval and a validation of the results against measurements by other instruments is given by Ungermann et al. (2012).

3.4 Effect of horizontal regularisation

20

This section quantifies the effect of horizontal regularisation on a cross-section retrieval. Besides the obvious horizontal smoothing of volume mixing rations, it needs to be examined if one can in turn reduce vertical regularisation to improve the vertical resolution at the cost of horizontal resolution and how much the noise can be reduced without degrading the horizontal resolution of the cross-section too much.

The case study necessarily derives multiple gases from the measurements, but to conserve space, mostly only the volume mixing ratios of CFC-11 and $CIONO_2$ will be discussed, as the other trace gases behave quite similar. The choice of CFC-11 is

disadvantageous for cross-section retrievals, as its signal is strong and the noise is low to begin with and therefore the possible gain is small compared to other retrieval gases.
 In contrast, CIONO₂ has a weaker signal and should thus profit more.



The baseline retrieval is a cross-section retrieval that is identical to a conventional series of 1-D retrievals as no horizontal regularisation has been added. The CFC-11 and ClONO₂ volume mixing ratios produced by the baseline retrieval are depicted in Fig. 2. Here we use pseudocolour plots showing each retrieved trace gas volume mixing ratio value as a rectangle. This cross-section exhibits a lot of detail and fine structure. Please note that most structures are horizontal or diagonal and that neighbouring profiles are often very similar. This suggests that consecutive profiles taken at this horizontal sampling are not fully independent from one another. We suppose that many of the inconsistencies between neighbouring profiles stem from random errors in our elevation angle knowledge, which affects lower altitudes more than those close to the instrument.

The CFC-11 cross-section looks noisy in several regions, which is caused by the given noise in the measured signal combined with a rather weak vertical regularisation; but this parametrisation allows a good reproduction of the filament of increased

- ¹⁵ CFC-11 volume mixing ration at 15.5 km on the right, which is situated between two airmasses of smaller volume mixing ratios. The vertical extent seems to be ≈ 500 m, which corresponds to two samples of our employed grid. On the left hand of the figure, two very faint horizontal filaments might be present at 13 km and 14 km. Also, a large artifact is present at 9.5 to 11 km at 10:35 UTC where no measurements were taken due to aircraft movements; this problem is revealed by the decreased information con-
- ²⁰ due to aircraft movements; this problem is revealed by the decreased information content in this area (not depicted).

The $CIONO_2$ cross-section is generally similar in quality to the CFC-11 cross-section. Only at lower altitudes, the $CIONO_2$ volume mixing ratios look much less noisy, mostly, because the measured signal gets rather weak and more a priori knowledge enters the retrieval result, which is reflected in a sharp drop off in measurement contribution (not depicted) below 12 km.

25

A natural starting point for the horizontal regularisation would be the typical scale difference between vertical and horizontal structures in the atmosphere, i.e. ≈ 200 . This strength is referred to as factor-200 horizontal regularisation within this paper.



The result for using thus $c_q^h = 200c_q^v$ for all quantities except aerosol gives Fig. 3 (we found it beneficial to keep $c_{aerosol}^h = 100$ as the vertical strength was very strong to begin with). Starting with the CFC-11 cross-section, the differences are subtle, but it is visually already much more pleasing than the baseline setup. With this regularisation strength, the largest effect is on individual pixels, where closely neighbouring profiles have differing values, i.e. mostly outliers and small-scale oscillations. Some such outliers in the thin vertical filaments have been reduced; especially the region between 10 and 12 km has become quieter. The artifact at 10:35 UTC has become smaller as the missing information starts being filled in from the right side (the left profile is spatially

- even farther away than suggest by the temporal proximity due to a turn of the aircraft and therefore exerts only a minimal influence). Also the CIONO₂ cross-section looks improved. No major features were smoothed over. Several features even look better, for example, the outflow of increased CIONO₂ at 11:45 UTC at 12 km is now consistent over all neighbouring profiles. This seems correct as indicated by retrieved HNO₃ concontrations, which have a strenger signal and show the same behaviour without any
- ¹⁵ centrations, which have a stronger signal and show the same behaviour without any horizontal regularisation (not depicted).

A stronger horizontal regularisation is depicted in Fig. 4, where the horizontal strength has been set to be 2000 times the vertical one (factor-2000). This result is devoid of obvious defects as they have been smoothed over. But strong horizontal gradiente are still reproduced a g the small intrusion of six with law CEC 11 volume

- ²⁰ gradients are still reproduced, e.g. the small intrusion of air with low CFC-11 volume mixing ratios at 11:30 UTC and 14 to 16 km. But the low volume mixing ratios within this filament are slightly increased due to smoothing. Also the diagonal structure at 16 km and 12:00 UTC shows larger CFC-11 concentrations than in the setup without horizon-tal regularisation. However, while certain details have been lost, it certainly reproduces the structure of the setup.
- the major features well. The weaker CIONO₂ volume mixing ratios behave similar. Most features are still present, but increased volume mixing ratios of small filaments have been reduced and vice versa.

For comparison, a setup with obviously too large horizontal regularisation is shown. Using a factor-20000 horizontal regularisation gives a results as shown in Fig. 5.



Previously clear cut horizontal and diagonal structures have been distorted by this setup. Even though the vertical regularisation is unmodified, some structures appear also vertical more smoothed. Depending on the retrieved quantity and the noise level, such a strength might still be useful, as structures with strong horizontal gradients are preserved even here.

3.5 Quantification of noise reduction

5

After presenting the visual differences of the various regularisation strengths, this section focuses on the impact of the horizontal regularisation on the more objective diagnostic noise error estimates.

- As the errors are quite similar from one profile to another, it is sufficient to discuss the errors of a single profile. Figure 6 shows the errors associated with the profile measured at 12:30 UTC for the baseline retrieval setup and factor-2000 horizontal regularisation. The total error is produced by taking the Euclidean norm of a vector consisting of all individual error terms. Only the most significant error terms are depicted in descending
- order. In Fig. 6a, the error according to stochastic noise (which was assumed to be 1 %) is the leading error down to 12 km. Only from there on, the error induced by spectral uncertainties in CFC-11 becomes dominant (this error was set to 5% in an ad hoc manner based on HITRAN11 error quantification). The corresponding plot for factor-2000 regularisation is given in Fig. 6b. It is shown that the noise error was reduced by
- ²⁰ a factor of \approx 3, leaving the total error being dominated by systematic errors implying an improved precision with similar accuracy of the result. The other errors are largely unaffected.

Figure 7 shows the reduction of noise for CFC-11 and CIONO₂ for various horizontal regularisation strength. The effect of noise for CFC-11 is depicted in Fig. 7a. Introducing horizontal regularisation in a natural manner reduced the influence of noise to 70% of the original level. Using a, factor-2000 horizontal regularisation reduces noise further down to 30% the original level. Increasing the horizontal regularisation strength further gives only diminishing returns, especially as the impact of noise is reduced to the same



order as most systematic or quasi-random error sources such as influence of unknown background gases, so that the gain in precision is not notably compared to the given accuracy. For $CIONO_2$, the factor-200 horizontal regularisation already halves the effect of noise. As its signal is weaker, it profits more from additional regularisation.

5 3.6 Achievable horizontal resolution

Adding horizontal regularisation obviously reduces the achievable horizontal resolution. However, it may be that the resulting resolution is still sufficient for the scientific purposes or that the given reduction of noise is seen as more advantageous. It may also be possible to reduce the employed vertical regularisation as the horizontal regularisation exhibits sufficient stabilising influence thereby increasing the achievable vertical resolution.

Focusing first on the worsening of horizontal resolution, Fig. 8 shows the horizontal resolution as derived from the full-width at half-max (FWHM) of the averaging kernel in horizontal direction. The width is determined including linear interpolation in between

- ¹⁵ data points, which causes the step-like structure stemming from the irregular horizontal grid size. Vertical variations are partly caused by differing azimuthal directions of the detector, which causes the lower tangent points of affected neighbouring profiles to be farther apart than upper ones (the profile taken at \approx 10:10 UTC is the most prominent example of this effect). The large value around 10:30 UTC is caused not only by
- the sparse measurement density, but also by a turn of the aircraft implying that the horizontal distance of the two seemingly neighbouring profiles is much larger than suggested by the temporal proximity. A regular horizontal grid would remove several of these artifacts, but on the other hand also introduce many points with purely interpolated information.
- The impact of naturally dimensioned factor-200 horizontal regularisation is very small as shown in Fig. 9. The horizontal resolution for CFC-11 is nearly unchanged to applying no horizontal regularisation. It is typically the maximum of \approx 16 km and the resolution of the baseline setup. This is only slightly larger than the typical minimum distance



between profiles of ≈ 15 km. Several individual points have a noticeably worsened horizontal resolution; these are typically points of profiles where there is a gap in the vertically measured spectra. For these cases, the retrieval prefers to interpolate the mixing ratios from neighbouring profiles rather than doing so from measurements far-

- ther above or below, especially if the neighbouring profile is close by. A second kind of artifact arises in the profile at 10:35 UTC, where the horizontal resolution seemingly was much improved at 9 km altitude. Here, an especially large gap in taken spectra is present and the retrieval extrapolates from the profile to the right, which was taken in close spatial proximity. The averaging kernel matrix indicates that the information is spatially well localised, but centred on the succeeding profile. This dislocation can also
- spatially well localised, but centred on the succeeding profile. This dislocation can also be visualised by plotting the distance between the maximum of the averaging kernel matrix row and the location of the retrieved point; however for the given case study, such a problem arises only for the 10:35 UTC profile.

Figure 10 shows the results of factor-2000 horizontal regularisation strength. Here, the horizontal resolution is obviously worse than for the baseline case, especially in the lower altitudes. But where a sufficient measurement density is given, e.g. around 11:45 UTC, the horizontal resolution is still ≈ 25 km, which is for many long-living trace gases quite sufficient.

3.7 Trade-off of vertical against horizontal resolution

It was established that adding horizontal regularisation reduces the impact of noise at the expense of worse horizontal resolution. One might therefore use this achieved stability to reduce the strength of vertical regularisation to, possibly, achieve less vertical smoothing.

Reducing the vertical regularisation strength by a factor of ten and using a factor-200 ²⁵ horizontal regularisation (related to the original vertical regularisation strength) gives similar noise figures as the baseline retrieval. Therefore, they should give qualitatively comparable results, but with different trade-offs. In Fig. 11, the corresponding CFC-11 and CIONO₂ volume mixing rations are depicted. The figure looks similarly "noisy" as



Fig. 2 but with a different distribution of errors. As neighbouring profiles are slightly more consistent, it is slightly more pleasing to the eye, but objectively with respect to error bars (not depicted) of comparable quality. Especially the lower altitudes around 10 km are quite improved. Some very small structures, like the extension of air with low
 ⁵ CFC-11 volume mixing ratios at 11:15 UTC and 15.5 km altitude seem slightly more

5 CFC-11 volume mixing ratios at 11:15 UTC and 15.5 km altitude seem slightly more consistent.

The horizontal resolution is similar to the setups discussed in Sect. 3.6, which is slightly reduced compared to the baseline case without horizontal regularisation, that is in the order of 15 km (i.e. equal to the profile sampling distance). The difference in vertical resolution is shown in Fig. 12. The vertical resolution of CFC-11 is already quite close to the achievable optimum constrained by the observation geometry, so its vertical resolution cannot improve significantly. But the other major retrieval targets benefit noticeable by up to 30 % at lower altitudes. Also water vapour benefits greatly by more than halving the vertical resolution around 16 km.

10

- ¹⁵ CFC-11 might be a bad example here, as the vertical regularisation was small to begin with and is effectively turned off altogether by the reduction by factor 10. The largest improvement in resolution of a primary target is achieved for ClONO₂. Figure 11b reproduces very well the horizontal filaments at 12 and 14 km altitude, which are difficult to discern in Fig. 2b due to noise and uncertainty in elevation angle. Also the low mix-
- ing ratios in the filament at 15.5 km in the latter part of the flight around 12:30 UTC are more pronounced, just as the area below at 12.5 km. Here is also a more noticeable difference as there is clearly a layer with increased mixing ratios below at 12 km, where previously due to vertical smoothing the volumes of low concentration above and below ran together (correlated mixing ratios in other trace gases suggest this to be real, as
- does examination of the measured radiance values). At only one place, there may to be a negative effect of the horizontal smoothing. At 11:35 UTC in Fig. 2a, there seems to be a vertically continuous mass of low ClONO₂ volume mixing ratios from 8 to 14 km. In Fig. 11b, this faint structure has become less pronounced. It is however unclear,



whether this is an artifact of the regularisation or the proper structure, as only some correlated retrieved trace gases exhibit a similar structures in the baseline retrieval.

Concluding, this technique allows to adjust the characteristics of the retrieved volume mixing ratios to the scientific question and the atmospheric situation. Especially for examining the upper troposphere and lower stratosphere and the tropopause inversion layer, which is of very small vertical extent, this technique might reproduce otherwise invisible details.

3.8 Comparison to smoothing of 1-D retrievals

25

The basic idea behind the cross-section retrieval is to exploit the similarity of neighbouring profiles. This can be exploited also by simple weighted moving averages. This section presents two such less costly alternatives to a full cross-section retrieval, operating solely on the results of a series of individual 1-D retrievals, and compares these to the cross-section retrieval.

For the first weighted moving average, the weights are chosen according to a Lapla-¹⁵ cian distribution (the probability density function is $\sqrt{2}\sigma^{-1}\exp\left(-\sqrt{2}|x-\mu|\sigma^{-1}\right)$ with mean μ and standard deviation σ) with a given FWHM, thereby effectively prescribing a fixed horizontal resolution. The Laplacian distribution was chosen because of its obvious relationship to the auto-regressive model and Tikhonov regularisation. In this way, results with a similar FWHM should be most similar. However, there are many other choices for window functions with different trade-offs, which cannot be fully explored here.

A second, more complex smoothing option is to weigh the profiles not only by a simple scalar but, in addition, with the inverse of their covariance matrix. This corresponds to a maximum likelihood estimator and is mathematically similar to the cross-section retrieval (the major distinction being the retrieval being non-linear).



In this section, x_i refers to the 1-D result of the *i*-th measured profile and S_i is the associated covariance matrix. Then the simple moving average for the *i*-th profile x_i^{avg} is defined as

$$\boldsymbol{x}_{i}^{\text{avg}} = \sum_{j} W_{i,j} \boldsymbol{x}_{j}$$

⁵ The maximum likelihood estimator for the *i*-th profile x_i^{ml} is defined as

$$\boldsymbol{x}_{j}^{\text{ml}} = \left(\sum_{j} w_{i,j} \mathbf{S}_{j}^{-1}\right)^{-1} \sum_{j} w_{i,j} \mathbf{S}_{j}^{-1} \boldsymbol{x}_{j}.$$
 (14)

The weights $w_{i,j}$ are again defined based on a Laplacian probability distribution with a mean of zero and a standard deviation corresponding to a FWHM of 30 km. This correlates roughly with the cross-section retrieval results for a factor-2000 horizontal regularisation. This stronger smoothing was chosen, as the effect of these methods and thereby also their differences become more apparent for stronger smoothing. For the maximum likelihood estimator, the actual width of the weighting functions is naturally also influenced by the involved covariance matrices, should they differ significantly from each other.

¹⁵ The horizontal distance between profiles is taken to be the average distance of the lower 20 km of the profiles. Other values for the FWHM have also been tested and it was asserted that the presented strength delivers the result most similar to the factor-2000 horizontal regularisation. A key difference of cross-section retrievals is that the strength of horizontal smoothing is also influenced by the signal to noise ratio of the measurements and the vertical structure of the stendard deviations of the alimetelary.

measurements and the vertical structure of the standard deviations of the climatology used for a priori, which leads to a stronger smoothing for lower altitudes than for higher altitudes. For either moving average, this effect might be incorporated to some extent by altitude dependent weights, i.e. a diagonal matrix with altitude dependent weights on the diagonal.



(13)

The result in CFC-11 volume mixing ratios for the simple weighted moving average is shown in Fig. 13. The volume mixing ratios look obviously much less noisy than in the baseline retrieval in Fig. 2. In so far this method is an effective means to simply reduce the impact of noise. The result looks overall similar to the corresponding Fig. 4; the only

- ⁵ obvious discrepancies can be found around obvious artefacts of the baseline retrieval. The profile with missing observations at 10:35 UTC still presents too low CFC-11 volume mixing ratios at lower altitudes. Interestingly, also applying stronger smoothing by Laplace distributions with FWHMs up to 200 km does not fully remedy this problem. In addition, there are some individual points with increased CFC-11 volume mixing ra-
- tio around 10:00 UTC, which are likely induced by uncertainties in elevation angle due to aircraft movements during the ascent. Neighbouring profiles fit much better to one another with horizontal regularisation compared to this simple weighted moving average. For this method, diagnostic information can simply be assembled by weighted averaging of the available error covariance matrices of the individual profiles. It is also
 straightforward to generate a 2-D averaging kernel matrix.

The more complicated maximum likelihood estimator delivers very similar results presented in Fig. 14. The only notable difference to Fig. 13 is in the profile at 10:35 UTC, which is slightly better re-constructed. However, some values in the range between 9 and 11 km do not have an increased standard deviation as one might expect, even

- though their value seems obviously biased. As a consequence, the weighted average does not fill these values from the surrounding profiles. The way that this method combines the available profiles is so close to the method of the actual cross-section retrievals that an even closer agreement between this result and the cross-section retrieval results would be expected. However, the cross-section retrieval is a non-linear
- ²⁵ process, while this recombination of profiles is inherently linear. In this way, the crosssection retrieval is more potent.

With respect to providing diagnostic information, the covariance matrix \mathbf{S}_{i}^{ml} of the retrieval result of each profile \mathbf{x}_{i}^{ml} follows from Eq. (14) and can be calculated by the following formula:



$$\mathbf{S}_{i}^{\mathsf{ml}} = \left(\sum_{j} w_{i,j} \mathbf{S}_{j}^{-1}\right)^{-2} \sum_{j} w_{i,j}^{2} \mathbf{S}_{j}^{-1}.$$

It is straightforward to derive further diagnostic quantities like gain and averaging kernel matrix from this. When applying this on the presented numerical example, we found the calculation of $\mathbf{S}_{i}^{\text{ml}}$ to be numerically unstable even after applying typical and more elaborate countermeasures, causing several negative variances to appear in the matrix. This is surprising, as such issues arise typically neither in the diagnostics of conventional 1-D profiles with matrices of similar size nor in the diagnostics of the full cross-section retrieval involving much larger matrices. The inversion and the required two matrix-matrix multiplications seem to be in combination much more sensitive to the condition number of the involved matrices (about $\approx 100\,000$ after symmetrically scaling

the diagonal to 1) of the underlying problem than the single inversion needed for 1-D or cross-section retrievals.

A different questions is, whether such smoothing methods can also be used to reduce the vertical regularisation strength. Initial tests with the discussed test case showed that the increase in oscillations in the 1-D profiles due to vertical decreased regularisation cannot be fully compensated by the a posteriori horizontal smoothing (not depicted). Here, the stabilising influence of horizontal regularisation during the retrieval is more critical as the chosen vertical regularisation strength is insufficient to guarantee numerical stability and well-posedness.

²⁰ Combining individual profiles in a weighted manner can produce some of the benefits of the cross-section retrieval, even though the results are slightly worse as the weighting is only a linear process compared to the non-linear cross-section retrieval. Moving averages have the advantage of being able to provide diagnostic information in a computationally less costly way at the cost of a result of reduced retrieval quality.



(15)

4 Summary and conclusions

5

This paper proposed and demonstrated non-linear 2-D cross-section retrievals of trace gases from airborne infrared limb-sounder measurements. As test case, the 2nd measurement flight on 2 March 2010 during the RECONCILE campaign presented by Ungermann et al. (2012) was used.

Modern limb-sounders are able to measure spectra of a full profile so fast that neighbouring profiles are very similar to one another. This can be exploited by jointly retrieving all profiles of a flight and applying horizontal regularisation.

An algorithm able to retrieve large-scale 2-D cross-sections was presented and it was shown how associated diagnostic information can be produced on common hardware. Stochastic algorithms can provide the error estimates for various error sources for all retrieved quantities with a cost of only $O(n^2)$. These algorithms were then applied to a real-life test case to examine the benefits of the cross-section retrieval.

First, the effect of adding horizontal regularisation of varying strength was exam-

- ¹⁵ ined. Using 200 times the vertical regularisation strength as horizontal strength (factor-200 horizontal regularisation) produced already visibly quieter images without notice-able degradation of features. 200 was chosen as a typical ratio between vertical and horizontal extent of atmospheric structures. Increasing the horizontal regularisation strength to factor-2000 produced a much quieter figure with obvious horizontal smooth-
- ²⁰ ing, but still containing all visible structures. A further increase to factor-20 000 however smoothed away several features and seems therefore undesirable.

After establishing a suitable range for horizontal regularisation strength, its impact on the retrieval error was examined. A detailed analysis shows that for the given case study, most error sources remain unaffected with the notable exception of noise in-

²⁵ duced error. It was demonstrated that adding horizontal regularisation unsurprisingly reduced the impact of noise on the retrieval results. The comparatively weak factor-200 horizontal regularisation reduced noise for the CFC-11 retrieval down to 70% of the original level, while a more aggressive strength can reduce it down to about 30% of



6605

the original level in the given example. For $CIONO_2$, factor-200 horizontal regularisation could reduce noise even down to 50 %.

Adding horizontal regularisation affects obviously the along-flightpath resolution. Factor-200 horizontal regularisation imposes a lower limit of ≈ 17 km for well-resolved

- trace gases to this horizontal resolution, which is just about 2 km worse than the minimal profile distance. The more the horizontal regularisation strength is increased from there on, the worse the horizontal resolution gets. For factor-2000 horizontal regularisation, the imposed minimum mean horizontal resolution is already more than 30 km, more at lower altitudes.
- ¹⁰ Using horizontal regularisation, it is possible to reduce the vertical regularisation strength below the strength necessary for successful 1-D retrieval. In the presented example, the vertical resolution of CIONO₂ could be improved by up to 30 % without increasing the noise level. This enabled thin filaments to become more easily visible and generally reduced the vertical smoothing. Choosing a good combination of vertical ¹⁵ and horizontal regularisation thereby improves the reproduction of boundaries between
- airmasses and thereby the examination of, e.g. mixing processes.

The last section compared the presented approach to the more simple weighted averaging of one-dimensionally retrieved profiles. A simple linear recombination using a Laplace function with a set FWHM and an average that also includes the covariance

²⁰ matrices of the results was examined. Both approached reduce the impact of noise in the same way as the cross-section retrieval. However, they were both less able to reduce artefacts caused by missing observations.

Using cross-section retrievals, it is possible to produce a better representation of the true atmospheric state by exploiting the high measurement density of modern instru-

25 ments and the self-similarity of the atmosphere. By using horizontal regularisation, it is possible to moderately lower the vertical regularisation strength and thereby enable the better reproduction of thin vertical layers important to mixing processes in the upper troposphere/lower stratosphere. This technique might also be used to more reliably derive constant or slowly varying instrument parameters, which cannot be measured pre- or



post flight. Especially for trace gases with weak signature, the technique can reduce the impact of noise significantly without noticeably degrading the horizontal resolution.

This technique can also be used to enhance the retrievals for current limb-imager instruments, which simultaneously take many closely-spaced profiles, such as GLORIA

 Gimballed Limb Observer for Radiance Imaging of the Atmosphere; see Ungermann et al., 2010).

It is straightforward to extend the presented technique also to proposed near-future satellite limb-imagers (Riese et al., 2005; ESA, 2012). Such a rearward-looking instrument measures several swaths in parallel and would probably use 2-D tomographic retrievals to evaluate each swath individually. The proposed technique is capable of evaluating all swaths together in a single 3-D tomographic retrieval. By exploiting the similarity between swaths, the described technique would stabilise the inherently more ill-posed tomographic retrieval problem and reduce the noise level. This might en-

¹⁵ Preusse et al., 2009).

20

25

This method can also be applied to the evaluation of current nadir sounders measurements as these also measure closely spaced profiles. One would obviously use horizontal regularisation in both along-flightpath and across-flightpath direction, but otherwise the method is directly applicable and might be even more beneficial due to the greater ill-posedness of the nadir retrievals compared to limb-sounder retrievals.

hance its scientific capabilities, for example with respect to gravity wave detection (see

Acknowledgements. J. Ungermann is supported by the Deutsche Forschungsgemeinschaft (GZ UN 311/1-1) and the visitor's program of the Atmospheric Chemistry Division at the National Center for Atmospheric Research (NCAR). NCAR is funded by the National Science Foundation. The RECONCILE project is funded under the European Commission Seventh Framework Programme (FP7) under grant number RECONCILE-226365-FP7-ENV-2008-1. The author thanks the RECONCILE coordination and flight planning teams, MDB and Enviscope for the successful implementation of and support during the Arctic field campaign. The

author especially thanks the CRISTA-NF team for the great work before, during, and after the RECONCILE campaign and for the provision of the measurements used in the case study.



The European Centre for Medium-Range Weather Forecasts (ECMWF) is acknowledged for meteorological data support.

The service charges for this open access publication

have been covered by a Research Centre of the 5 Helmholtz Association.

References

25

- Birner, T.: Fine-scale structure of the extratropical tropopause region, J. Geophys. Res., 111, D04104. doi:10.1029/2005JD006301. 2006. 6579
- Carlotti, M., Dinelli, B. M., Raspollini, P., and Ridolfi, M.: Geo-fit approach to the analysis of limbscanning satellite measurements, Appl. Optics, 40, 1872–1885, doi:10.1364/AO.40.001872, 2001. 6580, 6590
 - Curtis, A. R.: Discussion of "A statistical model for water vapour absorption" by R. M. Goody, Q. J. Roy. Meteorol. Soc., 78, 638-640, 1952. 6592
- ESA: Report for Mission Selection: PREMIER, vol. SP-1324/3, ESA Communication Production 15 Office, Noordwijk, The Netherlands, 2012. 6606
 - Forster, P. M. d. F. and Shine, K. P.: Radiative forcing and temperature trends from stratospheric ozone changes, J. Geophys. Res., 102, 10841-10855, doi:10.1029/96JD03510, 1997. 6578 Francis, G. L., Edwards, D. P., Lambert, A., Halvorson, C. M., Lee-Taylor, J. M., and
- Gille, J. C.: Forward modeling and radiative transfer for the NASA EOS-Aura high res-20 olution dynamics limb sounder (HIRDLS) instrument, J. Geophys. Res., 111, D13301, doi:10.1029/2005JD006270.2006.6592
 - Friedl-Vallon, F., Riese, M., Maucher, G., Lengel, A., Hase, F., Preusse, P., and Spang, R.: Instrument concept and preliminary performance analysis of GLORIA, Adv. Space Res., 37, 2287-2291, doi:10.1016/j.asr.2005.07.075, 2006. 6579
 - Gettelman, A., Hoor, P., Pan, L. L., Randell, W. J., Hegglin, M. I., and Birner, T.: The extra tropical upper troposphere and lower stratosphere, Rev. Geophys., 49, RG3003, doi:10.1029/2011RG000355, 2011. 6578

Discussion Pa	AMTD 5, 6577–6626, 2012 Improving retrieval quality for airborne limb-sounders J. Ungermann			
iper Discussion				
n Pap	Title Page			
θŗ	Abstract	Introduction		
	Conclusions	References		
iscussi	Tables	Figures		
on P	14	▶1		
aper		•		
_	Back	Close		
Discus	Full Screen / Esc			
sion	Printer-friendly Version			
Paper	Interactive Discussion			

- Glatthor, N., von Clarmann, T., Fischer, H., Funke, B., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Milz, M., Steck, T., and Stiller, G. P.: Global peroxyacetyl nitrate (PAN) retrieval in the upper troposphere from limb emission spectra of the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), Atmos. Chem. Phys., 7, 2775-2787, doi:10.5194/acp-7-2775-2007, 2007. 6593
- 5 Godson, W. L.: The evaluation of infra-red radiative fluxes due to atmospheric water vapour, Q. J. Roy. Meteorol. Soc., 79, 367–379, 1953. 6592
 - Gordley, L. L. and Russell, J. M.: Rapid inversion of limb radiance data using an emissivity growth approximation, Appl. Optics, 20, 807–813, doi:10.1364/AO.20.000807, 1981. 6592
- Grossmann, K. U., Offermann, D., Gusev, O., Oberheide, J., Riese, M., and Spang, R.: The 10 CRISTA-2 mission, J. Geophys. Res., 107, 8173, doi:10.1029/2001JD000667, 2002. 6592 Hegglin, M. I., Boone, C. D., Manney, G. L., and Walker, K. A.: A global view of the extratropical tropopause transition layer from atmospheric chemistry experiment Fourier transform spectrometer O₃, H₂O, and CO, J. Geophys. Res., 114, D00B11, doi:10.1029/2008JD009984, 2009. 6579

15

- Hoffmann, L., Kaufmann, M., Spang, R., Müller, R., Remedios, J. J., Moore, D. P., Volk, C. M., von Clarmann, T., and Riese, M.: Envisat MIPAS measurements of CFC-11: retrieval, validation, and climatology, Atmos. Chem. Phys., 8, 3671–3688, doi:10.5194/acp-8-3671-2008, 2008. 6592
- Kenney, J. F. and Keeping, E. S.: The Distribution of the Standard Deviation, vol. 2, Van Nos-20 trand, 2nd Edn., 1951. 6588, 6589
 - Kullmann, A., Riese, M., Olschewski, F., Stroh, F., and Grossmann, K. U.: Cryogenic infrared spectrometers and telescopes for the atmosphere - new frontiers, Proc. SPIE, 5570, 423-432, 2004. 6592
- Livesey, N., Van Snyder, W., Read, W., and Wagner, P.: Retrieval algorithms for the 25 EOS Microwave limb sounder (MLS), IEEE T. Geosci. Remote, 44, 1144-1155, doi:10.1109/TGRS.2006.872327, 2006. 6580, 6585
 - Nocedal, J. and Wright, S. J.: Numerical Optimization (Springer Series in Operations Research and Financial Engineering), 2nd Edn., Springer, New York, 2006. 6582
- ³⁰ Offermann, D., Grossmann, K.-U., Barthol, P., Knieling, P., Riese, M., and Trant, R.: Cryogenic infrared spectrometers and telescopes for the atmosphere (CRISTA) experiment and middle atmosphere variability, J. Geophys. Res., 104, 16311-16325, doi:10.1029/1998JD100047, 1999, 6592



- Preusse, P., Schroeder, S., Hoffmann, L., Ern, M., Friedl-Vallon, F., Ungermann, J., Oelhaf, H., Fischer, H., and Riese, M.: New perspectives on gravity wave remote sensing by spaceborne infrared limb imaging, Atmos. Meas. Tech., 2, 299-311, doi:10.5194/amt-2-299-2009, 2009. 6606
- 5 Remedios, J. J., Leigh, R. J., Waterfall, A. M., Moore, D. P., Sembhi, H., Parkes, I., Greenhough, J., Chipperfield, M. P., and Hauglustaine, D.: MIPAS reference atmospheres and comparisons to V4.61/V4.62 MIPAS level 2 geophysical data sets, Atmos. Chem. Phys. Discuss., 7, 9973-10017, doi:10.5194/acpd-7-9973-2007, 2007. 6593

Riese, M., Friedl-Vallon, F., Spang, R., Preusse, P., Schiller, C., Hoffmann, L., Konopka, P.,

Oelhaf, H., von Clarmann, T., and Höpfner, M.: Global limb radiance imager for 10 the atmosphere (GLORIA); scientific objectives. Adv. Space Res., 36, 989–995. doi:10.1016/i.asr.2005.04.115.2005.6606

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, vol. 2 of Series on Atmospheric, Oceanic and Planetary Physics, World Scientific, Singapore, 2000. 6580, 6582, 6584

15

25

Steck, T. and von Clarmann, T.: Constrained profile retrieval applied to the observation mode of the Michelson interferometer for passive atmospheric sounding, Appl. Optics, 40, 3559-3571, doi:10.1364/AO.40.003559, 2001. 6583, 6586

Steck, T., Höpfner, M., von Clarmann, T., and Grabowski, U.: Tomographic retrieval of atmo-

spheric parameters from infrared limb emission observations, Appl. Optics, 44, 3291-3301, 20 doi:10.1364/AO.44.003291, 2005. 6580

Tarantola, A.: Inverse Problem Theory, Society for Industrial and Applied Mathematics, Philadelphia, 2004. 6587

Tikhonov, A. N. and Arsenin, V. Y.: Solutions of ill-posed problems, Winston, Washington DC, USA, 1977. 6582

- Ungermann, J.: Tomographic reconstruction of atmospheric volumes from infrared limb-imager measurements, Forschungszentrum Jülich, Jülich, available at: http://hdl.handle.net/2128/ 4385, last access: September 2012, Ph.D. thesis, Wuppertal University, 2011. 6586, 6588, 6592
- Ungermann, J., Kaufmann, M., Hoffmann, L., Preusse, P., Oelhaf, H., Friedl-Vallon, F., and 30 Riese, M.: Towards a 3-D tomographic retrieval for the air-borne limb-imager GLORIA, Atmos. Meas. Tech., 3, 1647–1665, doi:10.5194/amt-3-1647-2010, 2010. 6579, 6590, 6606



- Ungermann, J., Kalicinsky, C., Olschewski, F., Knieling, P., Hoffmann, L., Blank, J., Woiwode, W., Oelhaf, H., Hösen, E., Volk, C. M., Ulanovsky, A., Ravegnani, F., Weigel, K., Stroh, F., and Riese, M.: CRISTA-NF measurements with unprecedented vertical resolution during the RECONCILE aircraft campaign, Atmos. Meas. Tech., 5, 1173–1191, doi:10.5194/amt-5-1173-2012, 2012. 6579, 6580, 6591, 6592, 6593, 6604
- Weinreb, M. P. and Neuendorffer, A. C.: Method to apply homogeneous-path transmittance models to inhomogeneous atmospheres, J. Atmos. Sci., 30, 662–666, doi:10.1175/1520-0469(1973)030<0662:MTAHPT>2.0.CO;2, 1973. 6592

5

AMTD 5, 6577–6626, 2012			
Improving retrieval quality for airborne limb-sounders J. Ungermann			
Title	Title Page		
Abstract	Introduction		
Conclusions	References		
Tables	Figures		
I4	۶I		
	•		
Back	Close		
Full Scre	Full Screen / Esc		
Printer-frier	Printer-friendly Version		
Interactive Discussion			

Discussion Paper

Discussion Paper

Discussion Paper

Discussion Paper

AN 5, 6577–6	AMTD 5, 6577–6626, 2012		
Improving retrieval quality for airborne limb-sounders			
J. Ungermann			
Title	Title Page		
Abstract	Introduction		
Conclusions	References		
Tables	Figures		
14	۶I		
	•		
Back	Close		
Full Scre	Full Screen / Esc		
Printer-friendly Version			
Interactive Discussion			
CC D			

Discussion Paper

Discussion Paper

Discussion Paper

Discussion Paper

Table 1. A list of employed integrated microwindows and their spectral range.

IMW	range (cm ⁻¹)	IMW	range (cm ⁻¹)
0	777.5–778.5	7	808.0-809.0
1	784.0–785.0	8	810.0–813.0
2	787.0–790.0	9	820.5-821.5
3	791.5–793.0	10	831.0-832.0
4	794.1–795.0	11	846.0-847.0
5	795.5–796.5	12	863.0-866.0
6	796.6–797.5		

Discussion Pa	AMTD 5, 6577–6626, 2012		
per Discussio	Improving retrieval quality for airborne limb-sounders J. Ungermann		
n Pa	Title Page		
oer	Abstract	Introduction	
_	Conclusions	References	
Discus	Tables	Figures	
sion F	I	۶I	
aper			
_	Back	Close	
Disc	Full Screen / Esc		
ussio	Printer-friendly Version		
n Pap	Interactive Discussion		
per	CC O		

Table 2. Parameters employed for the regularisation: Tikhonov strength and correlation lengths.

Parameter	Value	Parameter	Value
α_{0} $\alpha_{1}^{v}, \alpha_{1}^{h}$ $C_{temperature}$ $C_{aerosol}$ C_{CCl4} C_{CFC-11}	$\sqrt{0.1}$ 1 2 km 640 km 1 km 0.3 km 2 km	C_{CIONO_2} C_{H_2O} $C_{HCFC-22}$ C_{HNO_3} C_{O_3} C_{PAN}	4 km 12 km 32 km 4 km 6 km 8 km
CFC-113	<u>د ۱۱۱</u>		



Fig. 1. Accuracy for estimation of errors for varying samples sizes. Please not the varying vertical scale with solid lines indicating identical relative errors. The samples are always powers of two. The spread has been added for better visibility.





Fig. 2. A cross-section of CFC-11 and $CIONO_2$ volume mixing ratios, respectively, for the baseline regularisation (standard vertical regularisation and no horizontal regularisation). The straight black line indicates the altitude of the instrument. Crosses on this line indicate the mean position of the instrument during a profile measurement. The dashed lines indicate the altitude of the highest and lowest tangent point.



Discussion Paper

Discussion Paper

Discussion Paper

Discussion Paper



Fig. 3. A cross-section of CFC-11 and CIONO₂ volume mixing ratios, respectively, with standard vertical regularisation and factor-200 horizontal regularisation.





Fig. 4. A cross-section of CFC-11 and CIONO₂ volume mixing ratios, respectively, with standard vertical regularisation and factor-2000 horizontal regularisation.





Fig. 5. A cross-section of CFC-11 and CIONO₂ volume mixing ratios, respectively, with standard vertical regularisation and factor-20 000 horizontal regularisation.





Fig. 6. Leading errors in CFC-11 volume mixing ratio for a representative profile (12:30 UTC) and the baseline regularisation in panel **(a)** and a factor-2000 horizontal regularisation in panel **(b)**.





Fig. 7. Error induced by noise for CFC-11 and $CIONO_2$, respectively, for one representative profile (12:30 UTC) and varying horizontal regularisation strengths. The given number indicates the multiplier compared to baseline vertical regularisation strength.





Fig. 8. A cross-section of CFC-11 horizontal resolution (FWHM) for the baseline setup.





Fig. 9. A cross-section of CFC-11 horizontal resolution (FWHM) for the factor-200 horizontal regularisation strength.





Fig. 10. A cross-section of CFC-11 horizontal resolution (FWHM) for the factor-2000 horizontal regularisation strength.





Fig. 11. Cross-section of CFC-11 volume mixing ratio for a reduced vertical regularisation strength and factor-200 horizontal regularisation.



Discussion Paper

Discussion Paper









Fig. 13. Cross-section of CFC-11 volume mixing ratios derived by horizontally averaging of 1-D retrieval results using a simple Laplacian distribution with 30 km FWHM to generate the weights.





Fig. 14. Cross-section of CFC-11 volume mixing ratios derived by horizontally averaging of 1-D retrieval results using both a simple Laplacian distribution with 30 km FWHM and the retrieval result covariance matrices to generate the weights.

