Atmos. Chem. Phys. Discuss., 6, 6723–6751, 2006
www.atmos-chem-phys-discuss.net/6/6723/2006/
© Author(s) 2006. This work is licensed
under a Creative Commons License.

Atmospheric Chemistry and Physics Discussions

Elimination of hidden a priori information from remotely sensed profile data

T. von Clarmann and U. Grabowski

Forschungszentrum Karlsruhe, Institut für Meteorologie und Klimaforschung, Karlsruhe, Germany

Received: 18 April 2006 – Accepted: 3 July 2006 – Published: 18 July 2006

Correspondence to: T. von Clarmann (thomas.clarmann@imk.fzk.de)

Abstract

Profiles of atmospheric state parameters retrieved from remote measurements often contain a priori information which causes complication in the use of data for validation, comparison with models, or data assimilation. For such applications it often is desirable to remove the a priori information from the data product. If the retrieval involves an ill-posed inversion problem, formal removal of the a priori information requires resampling of the data on a coarser grid, which, however, is a prior constraint in itself. The fact that the trace of the averaging kernel matrix of a retrieval is equivalent to the number of degrees of freedom of the retrieval is used to define an appropriate information-centered representation of the data where each data point represents one degree of freedom. Since regridding implies further degradation of the data and thus causes additional loss of information, a re-regularization scheme has been developed which allows resampling without additional loss of information. For a typical $ClONO_2$ profile retrieved from spectra as measured by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), the constrained retrieval has 9.7 degrees of freedom. After application of the proposed transformation to a coarser information-centered altitude grid, there are exactly 9 degrees of freedom left, and the averaging kernel on the coarse grid is unity. Pure resampling on the information-centered grid without re-regularization would reduce the degrees of freedom to 7.1.

1 Introduction

While in early applications of remote sensing to atmospheric research retrieval approaches often were based on ad hoc methods or unconstrained maximum likelihood estimation, recent methods mostly rely on information theory (Rodgers, 2000). It is now common pracise to represent the state parameter to be retrieved on an altitude grid which is finer than the altitude resolution of the instrument (see, e.g. von Clarmann et al., 2003b). Since inference of vertical profiles of atmospheric state parameters on such a fine grid otherwise would lead to an ill-posed inversion problem, stochastical (Rodgers, 2000) or smoothness (Tikhonov, 1963; Twomey, 1963; Steck and von Clarmann, 2001) constraints are applied which make the retrieval stable, i.e. regularize the inversion problem. The disadvantage of regularization is that the elements of the solution of the retrieval problem do no longer carry independent information, and that the solution does not merely depend on the measurement but includes a certain portion of a priori information. This adds severe complication to many fields of application of remotely sensed data, including validation, comparison to model results, or data assimilation (Rodgers and Connor, 2003; Calisesi et al., 2005). For such applications a representation of the retrieved state parameters which is free of formal a priori information is desirable and often advantageous. Singular value decomposition-based approaches have been developed for similar purposes: Ceccherini et al. (2003) have proposed such an approach for validation purposes targeted at maximum likelihood estimates of the atmospheric state. Joiner and da Silva (1998) proposed two methods, namely null-space filtering of retrievals and partial eigen-decomposition retrievals. All these approaches, however, transform the estimated atmospheric state parameters into a space without obvious physical meaning, where most data users do not really feel comfortable. This paper is targeted at a physically obvious representation of retrieved data, where the data user does not need to consider sophisticated diagnostic metadata but can directly use the data as they are without running risk of major misinterpretation.

While considerations outlined above are applicable to a manifold of remote sensing instruments, in this paper they are discussed on the basis of the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) (Fischer and Oelhaf, 1996; Endemann and Fischer, 1993; SAG, 1999) which is an Earth observation instrument onboard the Envisat research satellite, which records high-resolution limb emission spectra of the Earth's atmosphere. From these spectra, vertical profiles of atmospheric constituents' abundances and temperature are retrieved. Operational near-real time data analysis (Ridolfi et al., 2000; Nett et al., 1999), under responsibility of the European Space Agency (ESA), avoids problems with ill-posedness of the inversion by using a grid for representation of the vertical profiles which is defined by the tangent altitudes of the measurements, and by restriction of the altitude coverage of the retrieved profiles to heights were the measurements contain a clear signal of the target quantity, and where no problems due to saturation effects in the measured spectra occur. However, there also exist further MIPAS data processors (von Clarmann et al., 2003a), which aim at an extended MIPAS data product for scientific use. One of these processors is the one operated at Forschungszentrum Karlsruhe, Institut for Meteorology and Climate Research (von Clarmann et al., 2003b). This processor uses constrained retrieval approaches, with all implications mentioned above, in particular mapping of a priori information onto the retrieved profiles and, as a consequence, interdependences of retrieved data. While all diagnostic information is available for correct quantitative use of retrieved data, the experience was made that a majority of data users refuses to deal with this diagnostic information and thus runs risk to misinterpret the data. Therefore it was found preferable to distribute the data in an easy-to-handle representation. This paper proposes and discusses a method how such a representation can be set up.

2 Theoretical background

With very few exceptions explicitely mentioned, we use the theoretical retrieval concept by Rodgers (2000). While all MIPAS data processors the authors are aware of perform the retrieval in the framework of Newtonean iteration, for the discussion of the method presented in this paper the linear formulation of the retrieval problem is used. The estimate of the n-dimensional atmospheric state vector $\hat{\boldsymbol{x}}$ is calculated from the m-dimensional vector containing the measured spectral radiances, $\boldsymbol{y}$, as

$$\hat{\boldsymbol{x}} = \boldsymbol{x}_a + (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{R})^{-1} \mathbf{K}^T \mathbf{S}_\epsilon^{-1} (\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_a)) = \boldsymbol{x}_a + \mathbf{G}(\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_a)), \quad (1)$$

where $\boldsymbol{x}_a$ contains the a priori information on the atmospheric state, $\boldsymbol{F}(\boldsymbol{x})$ is the radiative transfer forward model which provides the spectral radiances as a function of $\boldsymbol{x}$,

$\mathbf{K}$ is the $m \times n$ Jacobian matrix, whose elements are $\frac{\partial y_i}{\partial x_j}$; superscript ^{T} denotes transposed matrices; $\mathbf{S}_\epsilon$ is the measurement covariance matrix. $\mathbf{R}$ is a regularization term, which, in maximum a posteriori (formerly called "optimal estimation") applications as suggested by Rodgers (2000), is the inverse of the a priori covariance matrix, while, in our applications, it typically is a Tikhonov first order smoothing operator (Tikhonov, 1963); see Steck and von Clarmann (2001); Steck (2002) for implementation details. Setting $\mathbf{R}$ zero leads to the unconstrained least squares solution, which is also called maximum likelihood solution. $\mathbf{G}$ is called the gain function. All complications arising from iterative processing to solve problems caused by non-linearity of $F(\boldsymbol{x}_\mathrm{a})$ is omitted here since it does not contribute to the aspect of the problem discussed. The retrieval covariance matrix $\mathbf{S}_\mathrm{x}$ of $\hat{\boldsymbol{x}}$ is

$$\mathbf{S}_\mathrm{x} = (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{R})^{-1}, \quad (2)$$

while the noise error covariance matrix $\mathbf{S}_\mathrm{m}$ of $\hat{\boldsymbol{x}}$ is

$$\mathbf{S}_\mathrm{m} = \mathbf{G} \mathbf{S}_\epsilon \mathbf{G}^T. \quad (3)$$

Removal of the effect of the a priori information without rerunning the retrieval, i.e. the transformation of the maximum a posteriori solution $\hat{\boldsymbol{x}}_\mathrm{MAP}$, or any other regularized solution $\hat{\boldsymbol{x}}_\mathrm{REG}$ to the related maximum likelihood solution $\hat{\boldsymbol{x}}_\mathrm{ML}$ according to

$$\hat{\boldsymbol{x}}_\mathrm{ML} = (\mathbf{S}_\mathrm{x}^{-1} - \mathbf{R})^{-1} \left[\mathbf{S}_\mathrm{x}^{-1} \hat{\boldsymbol{x}}_\mathrm{REG} - \mathbf{R} \boldsymbol{x}_\mathrm{a} \right] \quad (4)$$

usually first needs resampling on a coarser grid to avoid singularity of $\mathbf{S}_\mathrm{x}^{-1} - \mathbf{R} = \mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K}$. It is worthwhile mentioning that the maximum likelihood solution is not really unconstrained but implicitly constrained by the coarse grid on which it is represented and the related interpolation convention.

The $n \times n$ averaging kernel matrix

$$\mathbf{A} = (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{R})^{-1} \mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} = \mathbf{G} \mathbf{K} \quad (5)$$

is a helpful diagnostic tool, which is useful to rewrite Eq. (1) such that it becomes obvious which part of the solution is controlled by the measurement and which by the a priori information:

$$\hat{\boldsymbol{x}}_{\mathrm{REG}} = \boldsymbol{x}_a + \mathbf{A}(\boldsymbol{x} - \boldsymbol{x}_a) + \mathbf{G}\epsilon, \quad (6)$$

where $\boldsymbol{x}$ represents the true atmospheric state and ϵ is the measurement noise characterized by $\mathbf{S}_\epsilon$. In the following we assume $\hat{\boldsymbol{x}}=\hat{\boldsymbol{x}}_{\mathrm{REG}}$.

The number of degrees of freedom of the signal was shown to be

$$dgf_{\mathrm{signal}} = \mathrm{tr}(\mathbf{A}) \quad (7)$$

for optimal estimation applications where $\mathbf{R}=\mathbf{S}_a^{-1}$ (Rodgers, 2000). However, this concept also holds for other regularization methods such as smoothing in the sense of Tikhonov (1963) or Twomey (1963) (see Appendix A). To avoid confusion, we follow the suggestion of Steck (2002) and use the term "number of degrees of freedom of the retrieval ($dgf_{\mathrm{retrieval}}$)" here rather than "number of degrees of freedom of the signal (dgf_{signal})" which is reserved for maximum a posteriori retrievals in a Bayesian sense.

3 Appropriate representation

In the following we address the question which representation of the retrieved profile is appropriate. The grid on which the profile is represented shall be coarse enough for the inversion of $\mathbf{K}^T\mathbf{S}_\epsilon^{-1}\mathbf{K}$ to be stable, but no coarser, in order to avoid unnecessary loss of information. We assume that the strength of the regularization in Eq. (1) has been chosen appropriately (in a Bayesian sense or any other optimality criterion, Steck, 2002) and that the number of degrees of freedom of the regularized solution is the maximum number of degrees of freedom reasonably obtainable from the given measurement.

A linear transformation of the retrieved profile $\hat{\boldsymbol{x}}$ to a coarser grid uses the transformation matrix

$$\mathbf{W}^* = (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T \quad (8)$$

ACPD
6, 6723–6751, 2006

T. von Clarmann and U. Grabowski

where $\mathbf{W}$ is the interpolation matrix which samples the coarse grid profile on the fine grid. The profile on the coarse grid then is

$$\tilde{\hat{x}} = \mathbf{W}^{*}\hat{x} \tag{9}$$

and the coarse-grid Jacobian is

$$\tilde{\mathbf{K}} = \mathbf{K}\mathbf{W} \tag{10}$$

The averaging kernel matrix on the coarse grid, $\tilde{\mathbf{A}} \in I\!R^{k\times n}$ is

$$\tilde{\mathbf{A}} = \tilde{\mathbf{G}}\mathbf{K} = \mathbf{W}^{*}\mathbf{G}\mathbf{K} = \mathbf{W}^{*}\mathbf{A} \tag{11}$$

The $\tilde{\mathbf{A}}\in I\!R^{k\times n}$ averaging kernel matrix is understood to represent the response of a retrieval in the k-dimensional grid to a delta perturbation in the finer n-dimensional grid. With $\mathbf{W}^{*}\mathbf{W}=\tilde{\tilde{\mathbf{I}}}$ where $\tilde{\tilde{\mathbf{I}}}$ is $k\times k$ unity, the averaging kernel $\tilde{\tilde{\mathbf{A}}}\in I\!R^{k\times k}$ related to a retrieval in the coarser k-dimensional grid is

$$\begin{aligned}
\tilde{\tilde{\mathbf{A}}} &= \tilde{\mathbf{G}}\tilde{\mathbf{K}} && (12)\\
&= \mathbf{W}^{*}\mathbf{G}\mathbf{K}\mathbf{W} && (13)\\
&= \mathbf{W}^{*}\mathbf{A}\mathbf{W} && (14)\\
&= \left(\mathbf{W}^{T}\mathbf{K}^{T}\mathbf{S}_{\epsilon}^{-1}\mathbf{K}\mathbf{W} + \mathbf{W}^{T}\mathbf{R}\mathbf{W}\right)^{-1}\mathbf{W}^{T}\mathbf{K}^{T}\mathbf{S}_{\epsilon}^{-1}\mathbf{K}\mathbf{W} && (15)
\end{aligned}$$

Averaging kernels of Eqs. (11) and (12) are not directly comparable, because they have delta functions of different widths as a reference: The delta funktion in the k-dimensional grid, which is the reference to Eq. (12), is wider than the delta function in the n-dimensional grid, which is the reference for Eq. (11).

The $k\times k$ covariance matrix of the retrieval represented in the coarse grid is

$$\mathbf{S}_{\tilde{x}} = \mathbf{W}^{*}\mathbf{S}_{x}\mathbf{W}^{*T} \tag{16}$$

An appropriate representation of the retrieval in a sense as discussed above fulfills the conditions

$$\tilde{\tilde{\mathbf{A}}} = \tilde{\tilde{\mathbf{I}}}, \tag{17}$$

and

$$k = \text{int}(dgf_{\text{retrieval}}) \tag{18}$$

where $dgf_{\text{retrieval}}$ here is the number of degrees of freedom of the retrieval in the fine grid, which is not typically an integer. As follows from Eq. (15), Eq. (17) can be satisfied with

$$\mathbf{W}^T\mathbf{R}\mathbf{W} = \mathbf{0} \tag{19}$$

This means that the transformed retrieval shall conserve as much of the degree of freedom as possible while the state parameters in the k-dimensional grid shall not be constrained to each other. This kind of representation of a vertical profile where each data point represents one degree of freedom we call "information-centered".

The number of useful gridpoints can be obtained by singular value decomposition, but not the placement of the altitude gridpoints. Therefore, we propose an approach which is based on the number of degrees of freedom of the retrieval. If the number of degrees of freedom, i.e. the number of independent pieces of information in the measurement, is equal to the dimension of the retrieval vector, an averaging kernel can be inverted, i.e. the a prior information can be removed from the retrieval. In the following we discuss two approaches to remove the a priori information from the retrieval and to represent the retrieval on a coarser altitude grid.

3.1 Staircase representation

First, we determine, how many degrees of freedom dgf_c shall be represented by each component of the k-dimensional state vector in the coarse grid. The ideal value would

ACPD

6, 6723–6751, 2006

be one, but since dgf generally is not an integer, dgf_c of each altitude step j is calculated as

$$\forall j \ : \ dgf_c = \frac{dgf}{\text{int}(dgf)} \tag{20}$$

Certainly, the excess information related to $k \times (dgf_c - dgf)$ is lost through slight under-sampling on the coarse k-dimensional grid. This, however, is the price to pay for an easy representation.

The new gridpoints $\tilde{h}_j$ are distributed over altitude such that

$$\sum_{l=l_1(j)}^{l_2(j)} a_{l,l} \approx 1 \quad \text{where} \quad \mathbf{A} = (a_{i_1 i_2}) \tag{21}$$

where l_1 and l_2 are the lowermost and uppermost gridpoints in the fine grid to be represented by one new coarse-gridpoint $\tilde{h}_j$. The new gridpoint then has to be placed somewhere between the altitudes $h(l_1)$ and $h(l_2)$. One can even go a step further and determine the placement of the new gridpoint $\tilde{h}_j$ by linear interpolation of the sub-trace $\sum_{l=l_1(j)}^{l_2(j)} a_{l(z),l(z)}$ in z. Alternatively, the closest altitude in the n-dimensional grid can be chosen, which offers some operational advantages. The level of sophistication of finding the exact information-centered altitudes is not really necessary, because no two subsequent limb measurements will be exactly identical but may be desired to be represented on the same grid, which requires approximations anyway.

In summary, we work bottom up along the diagonal of the averaging kernel matrix adding up diagonal elements until the trace of the block of the matrix considered by now exceeds dgf_c and assign an altitude gridpoint to this submatrix of the averaging kernel. Since this grid is approximately information-centered, it is considered a kind of natural representation of the retrieved profile $\hat{\boldsymbol{x}}$.

Resampling of the constrained oversampled fine-grid retrieval, however, degrades the profile and further reduces the information below $dgf_{\text{retrieval}}$. This follows from

Eq. (15), because Eq. (19) is not usually satisfied for arbitrarily chosen **W** and **R** matrices. This leads to $k \times k$ averaging kernels unequal unity, which is equivalent to less than k degrees of freedom. To compensate for this additional loss of information, it seems necessary to first remove the a priori information in the fine-grid retrieval. This, however, is not possible in most cases, since the unconstrained n-grid solution suffers from ill-posedness. Instead, we search for a new constraint $\mathbf{R}'$, which in effect is equivalent to the resampling of the retrieval to the appropriate coarse grid, i.e. which fulfills Eq.(19). In order to substitute the original constraint $\mathbf{R}$ by the new constraint $\mathbf{R}'$, we make use of Eq. 10.48 of Rodgers (2000)

$$\hat{\boldsymbol{x}}' = (\mathbf{S}_{\mathrm{x}}^{-1} - \mathbf{R} + \mathbf{R}')^{-1} \left[\mathbf{S}_{\mathrm{x}}^{-1} \hat{\boldsymbol{x}} - \mathbf{R}\boldsymbol{x}_{\mathrm{a}} + \mathbf{R}'\boldsymbol{x}'_{\mathrm{a}} \right], \tag{22}$$

where $\hat{\boldsymbol{x}}'$ is the transformed profile in the n-dimensional grid, $\mathbf{S}_{\mathrm{x}}$ is the estimated $n \times n$ covariance matrix of $\hat{\boldsymbol{x}}$, $\boldsymbol{x}_{\mathrm{a}}$ is the original a priori profile, and $\boldsymbol{x}'_{\mathrm{a}}$ is the new a priori profile, which is set to zero at all altitude gridpoints for the application outlined below. The new averaging kernel matrix after this substitution is

$$\mathbf{A}' = \mathbf{A}(\mathbf{R}') = (\mathbf{K}^T \mathbf{S}_{\epsilon}^{-1} \mathbf{K} + \mathbf{R}')^{-1} \mathbf{K}^T \mathbf{S}_{\epsilon}^{-1} \mathbf{K} \tag{23}$$

The new constraint term $\mathbf{R}'$ and the particular resampling matrix $\mathbf{W}'$ are chosen to fulfill the following conditions

$$\mathrm{tr}(\mathbf{A}(\mathbf{R})) \approx \mathrm{tr}(\mathbf{A}(\mathbf{R}')), \tag{24}$$

which is not required to be satisfied exactly because of $dgf \neq \mathrm{int}(dgf)$, and

$$\mathrm{tr}(\tilde{\tilde{\mathbf{A}}}') = \mathrm{tr}(\mathbf{W}'^{*} \mathbf{A}(\mathbf{R}') \mathbf{W}') \Leftrightarrow \tilde{\tilde{\mathbf{A}}}' = \tilde{\tilde{\mathbf{I}}} \Leftrightarrow \mathbf{W}'^T \mathbf{R}' \mathbf{W}' = \mathbf{0}, \tag{25}$$

where $\mathrm{tr}(\tilde{\tilde{\mathbf{A}}}')$ is the number of degrees of freedom of the resampled profile. This is achieved by a constraint $\mathbf{R}'$ in the fine grid which is equivalent to a resampling onto the coarse grid. The formal resampling of an n-grid profile subject to such a constraint onto the k-grid then conserves all information.

In the following we present a pair of $\mathbf{R}'$ and $\mathbf{W}'$ matrices which fulfill these two conditions. For convenience, we assume that the coarse grid is a subset of the fine grid. The new constraint $\mathbf{R}'$ is supposed not to constrain values at the gridpoints which also are members of the coarse grid, while it shall produce values between the coarse gridpoints which do not carry additional information but are completely determined by the values at the coarse gridpoints and a chosen interpolation scheme, e.g. a staircase function.

A constraint $\mathbf{R}'$ compliant to this requirement is a block-diagonal $n \times n$ matrix, of the form

$$\mathbf{R}'_{\Box} = \begin{bmatrix} \mathbf{R}'_{\Box\text{sub};1}, & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}'_{\Box\text{sub};2}, & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \mathbf{0}, & \mathbf{0} & \cdots & \mathbf{R}'_{\Box\text{sub};k} \end{bmatrix}, \tag{26}$$

where each of the disjoint diagonal blocks $\mathbf{R}'_{\Box\text{sub};j}$ represents one gridpoint in the coarse grid, representing one degree of freedom. The coarse gridpoints shall be placed where

$$\sum_{l=l_1(j)}^{l_2(j)} a_{l,l} = \frac{dgf_\text{c}}{2} + (j-1)dgf_\text{c}, \tag{27}$$

while the block boundaries are placed where

$$\sum_{l=l_1(j)}^{l_2(j)} a_{l,l} = j\ dgf_\text{c}. \tag{28}$$

Blocks of the size $i(j) \times i(j)$ are set zero when $i(j)=1$, while blocks equal or larger than 2×2 are set up as first order Tikhonov-type regularization matrices (Tikhonov, 1963) of the type

$$\mathbf{R}'_{\Box\text{sub}} = \gamma \mathbf{L}_1 \mathbf{L}_1^T \tag{29}$$

where $\mathbf{L}_1$ is a local-gridwidth weighted $i(j)\times(i(j)-1)$ first order difference matrix of the type

$$\mathbf{L}_1^T(j) = \begin{pmatrix} (h_l - h_{l+1})^{-1} & 0 & \cdots & 0 \\ 0 & (h_{l+1} - h_{l+2})^{-1} & \cdots & 0 \\ \vdots & \cdots & & \vdots \\ 0 & 0 & \cdots & (h_{i_j-1} - h_{i_j})^{-1} \end{pmatrix} \times$$

$$\times \begin{pmatrix} -1 & 1 & 0 \cdots & 0 & 0 \\ 0 & -1 & 1 \cdots & 0 & 0 \\ \vdots & & \cdots 0 & \vdots & \\ 0 & 0 & 0 \cdots & -1 & 1 \end{pmatrix} \tag{30}$$

for regular fine grid, or a similar matrix which each non-zero matrix element weighted by $h_{l+1}-h_l$ where j indicated the row of $\mathbf{L}^T$ gridwidth-compensated second order difference operators. Such a regularization does not constrain the $\hat{x}$-values which correspond to coarse-grid-points but, when sufficient regularization strength for each matrix-block is chosen such that no altitude resolution is allowed within each block, values at the additional gridpoints are forced to follow the staircase function between the coarse-grid points. Such large γ values are necessary to avoid singularity of the first term in Eq. (22). Transformation of arbitrary fine gridded profiles to the coarse grid then uses

$$\mathbf{W}_{\square}'^{*} = \begin{pmatrix} \frac{1}{i_1} \cdots \frac{1}{i_1} & 0 \cdots 0 & \cdots 0 \cdots 0 \\ 0 \cdots 0 & \frac{1}{i_2} \cdots \frac{1}{i_2} & \cdots 0 \cdots 0 \\ & \cdots & \\ 0 \cdots 0 & 0 \cdots 0 & \cdots \frac{1}{i_k} \cdots \frac{1}{i_k} \end{pmatrix}, \tag{31}$$

where i_j is the number of fine-grid points to be represented by the jth coarse-grid point. The number of non-zero entries in each row of $\mathbf{W}'^{*}$ is i_j. For staircase profiles, however, which have been generated with $\mathbf{R}'_{\square}$ and $\gamma\rightarrow\infty$, also a simpler transformation matrix of the type

$$\mathbf{W}'^{*}_{\Box} = \begin{pmatrix} 1\ 0 \cdots 0\ 0\ 0 \cdots 0 & 0\ 0 \cdots 0 \\ 0\ 0 \cdots 0\ 1\ 0 \cdots 0 & 0\ 0 \cdots 0 \\ & \ddots \quad 0 \\ 0\ 0 \cdots 0\ 0\ 0 \cdots 0 & 1\ 0 \cdots 0 \end{pmatrix} \tag{32}$$

will transform the matrix to the coarse grid without loss of information, because due to the regularization applied all values within a block are the same. Reformation of the staircase-profile from the coarse-grid profile uses

$$\mathbf{W}'_{\Box} = \begin{pmatrix} 1\ 0 \cdots 0 \\ \vdots \quad\quad \vdots \\ 1\ 0 \cdots 0 \\ 0\ 1 \cdots 0 \\ \vdots \quad\quad \vdots \\ 0\ 1 \cdots 0 \\ \cdots \\ 0\ 0 \cdots 1 \\ \vdots \quad\quad \vdots \\ 0\ 0 \cdots 1 \end{pmatrix}, \tag{33}$$

where the number of non-zero entries in the ith column is n_i. It should be noted that also $\mathbf{W}'^{*}_{\Box}$ satisfies the condition

$$\mathbf{W}'^{*}_{\Box}\mathbf{W}' = \tilde{\tilde{\mathbf{I}}} \tag{34}$$

The validity of Eq. (19) can easily be verified for regularization and transformation matrices as defined in Eqs. (26–30) and (33). The re-regularized and resampled profile thus contains k=int(dgf) degrees of freedom.

ACPD

6, 6723–6751, 2006

3.2 Triangular representation

Neither is the staircase representation considered a realistic representation of the atmosphere, nor do most forward models support this representation. A representation where atmospheric state parameters are interpolated (e.g. linearly) between the altitude gridpoints is preferred in most applications. As with the staircase representation, we require the information content of the resampled retrieval to be the integer value of the degrees of freedom of the fine-grid retrieval (Eq. 20) and define this to be the number of coarse-grid points. In order not to lose information at the uppermost and lowermost end of the retrieved profile, we define

$$\tilde{h}_1 = h_1 \quad (35)$$

and

$$\tilde{h}_k = h_n \quad (36)$$

where $\tilde{\boldsymbol{h}}=(\tilde{h}_1,\cdots,\tilde{h}_k)^T$ is the grid of the coarse representation, while $\boldsymbol{h}=(h_1,\cdots,h_n)^T$ is the original fine altitude grid. In order to satisfy Eq. 20), the other coarse-grid points $\tilde{h}_j$ are placed such that

$$\sum_{l=l_1(j)}^{l_2(j)} a_{l,l} \approx \frac{k}{k-1}. \quad (37)$$

In order to emulate this coarse grid retrieval in the fine grid retrieval, the following regularization matrix can be constructed:

$$\mathbf{R}'_\Delta = \begin{bmatrix} \mathbf{R}'_{\Delta\text{sub};1} & 0\cdots0 \\ 0 & 0\cdots0 \\ \vdots & \vdots \\ 0 & 0\cdots0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0\cdots0 \\ 0 & \mathbf{R}'_{\Delta\text{sub};2} & 0\cdots0 \\ 0 & 0 & 0\cdots0 \\ \vdots & & \vdots \\ 0 & 0 & 0\cdots0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \mathbf{R}'_{\Delta\text{sub};k-1} \end{bmatrix} \quad (38)$$

Contrary to $\mathbf{R}'_{\square}$, the diagonal blocks of $\mathbf{R}'_{\Delta}$ are not disjoint but each two adjacent blocks share a common diagonal element. Each block j is of the size

$$i(j) \times i(j) = (l_2(j) - l_1(j) + 1) \times (l_2(j) - l_1(j) + 1).$$

Blocks $\mathbf{R}'_{\Delta\text{sub};j}$ of the size 2×2 are set zero, while blocks equal or larger than 3×3 are set up as second order Tikhonov-type regularization matrices (Tikhonov, 1963) of the type

$$\mathbf{R}'_{\Delta\text{sub}} = \gamma \mathbf{L}_2 \mathbf{L}_2^T \tag{39}$$

where $\mathbf{L_2}^T$ is a local-gridwidth weighted $(i(j)-2) \times i(j)$ second order difference matrix of tridiagonal form with sub-diagonal, diagonal, and super-diagonal elements $\lambda_i \Delta_i$, $-\lambda_i(\Delta_{i-1}+\Delta_i)$, $\lambda_i \Delta_{i-1}$, respectively, where

$$\Delta_l = h_{l+1} - h_l \tag{40}$$

and

$$\lambda_i = \frac{2}{\Delta_{i-1}\Delta_i(\Delta_{i-1} + \Delta_i)}. \tag{41}$$

In this application, transformation matrix $\mathbf{W}'_{\Delta}$ transforms the coarse-grid profile to the fine grid by linear interpolation, with

$$\mathbf{W}'_{\Delta} = \begin{pmatrix} \frac{\tilde{h}_2-h_1}{\tilde{h}_2-\tilde{h}_1} & \frac{h_1-\tilde{h}_1}{\tilde{h}_2-\tilde{h}_1} & 0 & & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \frac{\tilde{h}_2-h_{i_1-1}}{\tilde{h}_2-\tilde{h}_1} & \frac{h_{i_1-1}-\tilde{h}_1}{\tilde{h}_2-\tilde{h}_1} & 0 & & 0 & 0 \\ 0 & \frac{\tilde{h}_3-h_{i_1}}{\tilde{h}_3-\tilde{h}_2} & \frac{h_{i_1}-\tilde{h}_2}{\tilde{h}_3-\tilde{h}_2} & & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \frac{\tilde{h}_3-h_{i_1+i_2-1}}{\tilde{h}_3-\tilde{h}_2} & \frac{h_{i_1+i_2-1}-\tilde{h}_2}{\tilde{h}_3-\tilde{h}_2} & & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & & \frac{\tilde{h}_k-h_{n-i_k}}{\tilde{h}_k-\tilde{h}_{k-1}} & \frac{h_{n-i_k}-\tilde{h}_{k-1}}{\tilde{h}_k-\tilde{h}_{k-1}} \\ \vdots & \vdots & \vdots & & \vdots & \\ 0 & 0 & 0 & & \frac{\tilde{h}_k-h_{n-1}}{\tilde{h}_k-\tilde{h}_{k-1}} & \frac{h_{n-1}-\tilde{h}_{k-1}}{\tilde{h}_k-\tilde{h}_{k-1}} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}. \quad (42)$$

The generalized formulation of matrix $\mathbf{W}'^{*}_{\Delta}$, which controls the transformation from the fine to the coarse grid for an arbitrary profile leads to a lengthy expression. Multiplication of any fine-grid profile $\boldsymbol{x}$ by $\mathbf{W}'^{*}_{\Delta}$ returns the footpoints of the contiguous sequence of segments of which each is a regression line constrained such that the regression lines intersect at the pre-defined coarse-grid altitudes. If need be (e.g. for validation purposes as discussed in Sect. 4.1), $\mathbf{W}'^{*}_{\Delta}$ can easily be computed numerically for the actual interpolation scheme $\mathbf{W}'_{\Delta}$ according to Eq. (8). For the special case discussed here, where the fine-grid profile has been re-regularized such that all profile values on the fine-grid falling between the coarse-grid points can be generated by linear interpolation and thus carry no independent information, the following simpler matrix can be used for the transformation:

$$\mathbf{W}'^{*}_{\Delta} = (w^{*}_{j,l}); \quad w^{*}_{j,l} = \begin{cases} 1 & : \quad h_l \in \{\tilde{h}_1 \cdots \tilde{h}_k\} \\ 0 & : \quad h_l \notin \{\tilde{h}_1 \cdots \tilde{h}_k\} \end{cases} \tag{43}$$

This interpolation matrix just picks out the independent data points of the interpolated profile. Again, for $\gamma \rightarrow \infty$ the re-regularized retrieval based on the constraint $\mathbf{R}'_{\Delta}$ is transformed onto the coarse grid by transformation matrix $\mathbf{W}^{*}_{\Delta}$ without additional loss of information, and we have

$$\mathbf{W}'^{*}_{\Delta}\mathbf{W}' = \tilde{\tilde{\mathbf{I}}}. \tag{44}$$

4 Application areas

4.1 Validation

Validation of remotely sensed profiles of atmospheric state parameters recently has been described as a χ^2-test of the difference profile (Rodgers and Connor, 2003). After transformation of the profiles to the same altitude grid and *a priori* profile, these authors evaluate the estimated covariance matrix $\mathbf{S}_{\text{diff}}$ of the difference of both profiles $\hat{\boldsymbol{x}}_1$ and $\hat{\boldsymbol{x}}_2$ and evaluate the χ^2

$$\chi^2 = (\hat{\boldsymbol{x}}_1 - \hat{\boldsymbol{x}}_2)^T \mathbf{S}^{-1}_{\text{diff}} (\hat{\boldsymbol{x}}_1 - \hat{\boldsymbol{x}}_2), \tag{45}$$

in order to judge if the profile differences are significant. Besides various sources of retrieval errors, an important contributor to $\mathbf{S}_{\text{diff}}$ is the smoothing error of the difference, $\mathbf{S}_{\text{smooth.,diff}}$, which originates from different altitude resolutions of both profiles and which is calculated as

$$\mathbf{S}_{\text{smooth.,diff}} = (\mathbf{A}_2 - \mathbf{A}_1)\mathbf{S}_{\text{a}}(\mathbf{A}_2 - \mathbf{A}_1)^T. \tag{46}$$

ACPD
6, 6723–6751, 2006

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀◀ | ▶▶
◀ | ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

EGU

This covariance matrix can only be evaluated when the a priori covariance matrix $\mathbf{S}_a$ is known, which often is not the case. Unfortunately, there are situations when assumptions on $\mathbf{S}_a$ are dominating the calculation of χ^2.

For an assumed a priori covariance matrix $\tilde{\mathbf{S}}_a$, the deviation of the estimated smoothing error of the difference of the two profiles from the true smoothing error is

$$(\mathbf{A}_2 - \mathbf{A}_1)(\tilde{\mathbf{S}}_a - \mathbf{S}_a)(\mathbf{A}_2 - \mathbf{A}_1)^T. \quad (47)$$

It cannot easily be associated with certain altitude scales of atmospheric variability.

In the following, we shortly discuss various approximations aiming at making profiles comparable such that the smoothing error of the difference of the compared profiles vanishes. Since these approximations are not perfect, there remains a residual smoothing error, which we characterize below.

An approximation widely used is to consider the better resolved one of the two profiles, $\hat{\boldsymbol{x}}_2$ to be compared as an ideal measurement and to smooth it by application of the averaging kernel of the coarser resolved measurement before comparison. This approximation does not rely on an estimate of the a priori covariance. The residual smoothing error is

$$\mathbf{S}_{\text{smooth.,diff}} = \mathbf{A}_1(\mathbf{I} - \mathbf{A}_2)\mathbf{S}_a(\mathbf{I} - \mathbf{A}_2)^T \mathbf{A}_1^T, \quad (48)$$

where $\mathbf{I}$ is $n \times n$ identity. $\mathbf{S}_{\text{smooth.,diff}}$ is zero in the ideal case when $\mathbf{A}_2$=$\mathbf{I}$.

When comparing retrievals whose altitude resolution is too different to be ignored but too similar that the approach mentioned above would be justified, one may wish to crosswise smooth the retrievals in order to achieve cancellation of smoothing errors. This, however, is not appropriate since the application of averaging kernels to profiles is not commutative. The residual smoothing error is

$$\mathbf{S}_{\text{smooth.,diff}} = (\mathbf{A}_1\mathbf{A}_2 - \mathbf{A}_2\mathbf{A}_1)\mathbf{S}_a(\mathbf{A}_1\mathbf{A}_2 - \mathbf{A}_2\mathbf{A}_1)^T. \quad (49)$$

Instead, we suggest to first re-regularize and resample the finer resolved one of the two profiles, $\hat{\boldsymbol{x}}_2$, following the procedure described in Sect. 3. Then the resulting profile

$\tilde{\tilde{x}}'_2$ is smoothed by the averaging kernel $\tilde{\tilde{\mathbf{A}}}_1$ of the other profile regridded to the k-grid. Using the decomposition

$$\mathbf{S}_a = \mathbf{S}_a^{\text{fine}} + \mathbf{S}_a^{\text{coarse}}, \quad (50)$$

where $\mathbf{S}_a^{\text{coarse}}$ is the part of the variability of the state variable which could completely be represented on the coarse grid, and a component $\mathbf{S}_a^{\text{fine}}$ which is the part of the variability which can only be represented on the fine grid, the residual smoothing error can be written in the k-grid as

$$\begin{aligned}\mathbf{S}_{\text{smooth.,diff}} &= \mathbf{W}'^{*}(\mathbf{A}_1\mathbf{A}'_2 - \mathbf{A}'_2\mathbf{A}_1)\mathbf{S}_a^{\text{fine}}(\mathbf{A}_1\mathbf{A}'_2 - \mathbf{A}'_2\mathbf{A}_1)^T\mathbf{W}'^{*T} + \\ &\quad (\tilde{\tilde{\mathbf{A}}}_1\tilde{\tilde{\mathbf{A}}}'_2 - \tilde{\tilde{\mathbf{A}}}'_2\tilde{\tilde{\mathbf{A}}}_1)\mathbf{W}^{*\prime}\mathbf{S}_a^{\text{coarse}}\mathbf{W}^{*\prime T}(\tilde{\tilde{\mathbf{A}}}_1\tilde{\tilde{\mathbf{A}}}'_2 - \tilde{\tilde{\mathbf{A}}}'_2\tilde{\tilde{\mathbf{A}}}_1)^T \\ &= \mathbf{W}'^{*}(\mathbf{A}_1\mathbf{A}'_2 - \mathbf{A}'_2\mathbf{A}_1)\mathbf{S}_a^{\text{fine}}(\mathbf{A}_1\mathbf{A}'_2 - \mathbf{A}'_2\mathbf{A}_1)^T\mathbf{W}'^{*T}.\end{aligned} \quad (51)$$

The second term disappears because of $\tilde{\tilde{\mathbf{A}}}'_2=\mathbf{I}$ and all residual smoothing error is associated with subscale variability $\mathbf{S}_a^{\text{fine}}$. This approximation is exact, i.e. the residual smoothing error disappears, in the case that the variability of the true state is sufficiently well characterized by $\mathbf{S}_a^{\text{coarse}}$, i.e. if the variability can be represented completely in the coarse grid established in Sect. 3. In real applications, however, there often is no reliable information on the variability of the true state available, and it may be helpful to get rid of smoothing error components related to large scale variability and to be able to relate the ignored part of the smoothing error to a smaller scale.

4.2 Correlation analysis

For many scientific applications, correlations between mixing ratios of various trace species are analyzed, e.g. Kondo et al. (1999); Rex et al. (1999); Plumb et al. (2000); Wetzel et al. (2002); Ray et al. (2002); Glatthor et al. (2005); Mengistu Tsidu et al. (2005); Muscari et al. (2003); Esler and Waugh (2002); Müller et al. (1999). Different or varying altitude resolution of different species can cause artefacts in the correlations.

To avoid these, it is recommended to first resample and re-regularize the data on a common grid with matching altitude resolution. If the correlations are to be compared to reference correlations, these have to be resampled and degraded for consistency. If sub-scale effects are important, the same caveat as in Sect. 4.1 applies.

4.3 Trace gas families

In many applications, trace gas families such as $[ClO_x] = [Cl] + [ClO] + 2[Cl_2O_2] + [OClO]$, $[ClO_y] = [ClO_x] + [ClONO_2] + [HCl] + [HOCl]$, $[NO_x] = [NO] + [NO_2]$ or $[NO_y] = [NO_x] + [HNO_3] + [ClONO_2] + 2[N_2O_5] + [HNO_4]$ are analyzed. Summation of mixing ratios, however, is only meaningful if related profiles of different species are of the same altitude resolution. Also here the proposed regridding and re-regularization approach helps, however, on the grid appropriate to optimally represent the coarser resolved retrieval.

4.4 Data assimilation

When remotely sensed data of atmospheric state parameters are used in data assimilation, the data should represent the measurements with no significant loss of information, and should not add any apparent information (a priori) that does not appear in the measurements. Further, the amount of data transferred to the assimilator should be as small as convenient, consistent with retaining its full information content (Rodgers, 2000). The approach suggested in this paper satisfies these conditions quite well, as discussed in the following.

The loss of information is limited to $dgf - \text{int}(dgf)$, which is less than unity by definition. Only if a common altitude grid for a series of remotely sensed vertical profiles is chosen for convenience, further loss of information will occur. This, however, can simply be avoided in variational data analysis by evaluation of the objective function characterizing the residual between model run and observations on the individual grids.

Besides the smoothness of the profiles implied by the coarse grid, there is no further

ACPD

6, 6723–6751, 2006

Elimination of a priori

T. von Clarmann and U. Grabowski

Title Page | Abstract | Introduction | Conclusions | References | Tables | Figures | ◀◀ | ▶▶ | ◀ | ▶ | Back | Close | Full Screen / Esc | Printer-friendly Version | Interactive Discussion

EGU

a priori information in the observations, which might lead to overweighting the observations or may cause some conflict with the assumption of statistical independence between the observations and the model forecast. While the observations after the transformation discussed in this paper carry independent information, the retrieval errors typically are still correlated in altitude. In limb sounding, the off-diagonal elements of the retrieval covariance matrix are zero only if one deals with slant path column amounts rather than profiles.

Also the amount of data to be transferred from the data provider to the assimilator is reduced by a factor of approximately

$$\frac{n+n^2+\frac{(n+1)n}{2}}{2k+\frac{(k+1)k}{2}},$$

where n, n^2, $\frac{(n+1)n}{2}$ are the numbers of variables needed to store the profile, the averaging kernel matrix and the retrieval covariance matrix in the fine grid, while $2k$ and $\frac{(k+1)k}{2}$ are the numbers of variables needed to store the profile plus the altitude grid definition and the retrieval covariance matrix in the coarse grid. Since the averaging kernel is unity in the coarse grid, it needs not to be provided. Certainly, in the fine grid also the a priori profile of size n, the grid definition of size n and the a priori covariance matrix of size $\frac{(n+1)n}{2}$ will be needed by the assimilator, but these quantities are assumed not to vary from observation to observation and thus do not contribute significantly when large numbers of observations are provided. Still data reduction is considerable, mainly because the covariance matrices are smaller and the averaging kernel matrices are not needed.

5 Case study: MIPAS

Here we illustrate the behaviour of the proposed re-regularization/resampling method by application to MIPAS-type limb emission simulated measurements. First, a vertical profile of $ClONO_2$ volume mixing ratios is retrieved, using the regular processing setup as described in Höpfner et al. (2004). For radiative transfer modeling, the Karlsruhe Optimized and Precise Radiative Transfer Algorithm (KOPRA) (Stiller, 2000) was used. Retrievals were performed with the MIPAS data processor as described in von Clarmann et al. (2003b). Volume mixing ratios are sampled on a 1-km grid from 4–44 km, a 2-km grid from 44–70 km, and at 70, 75, 80, 90, 100 and 120 km. Since this grid is finer than the altitude resolution provided by the measurement geometry, i.e. the vertical distance of adjacent tangent altitudes, and since there is not enough spectral information in the measurement to extract altitude-resolved information on this fine altitude grid, the retrieval needs regularization to be stable. In this case, a constraint has been chosen which minimizes the first order differences of mixing ratios at adjacent altitude gridpoints as discussed in (Steck and von Clarmann, 2001, and references therein). While represented at many more altitude gridpoints, the retrieved profile has only 9.7 degrees of freedom. Our proposed method selects the following 9 altitude gridpoints to represent the profile: 8.0, 11.0, 15.0, 18.0, 22.0, 26.0, 30.0, 36.0, and 44.0 km. Resampling of the profile on this coarse grid further reduces the degrees of freedom to 7.1 (see discussion in the paragraph above Eq. 22). Contrary to this, our proposed method conserves exactly 9 degrees of freedom, and the averaging kernel matrix is exactly 9×9 unity. The original retrieval and the re-regularized/resampled profile are shown in Fig. 1. Certainly re-regularized/resampled profile is not free of a priori information because the coarse grid is a priori information in itself, but this kind of a priori information is obvious also to data users who are not familiar with the averaging kernel formalism. In consequence, the risk of misinterpretation of data is largely reduced.

6 Conclusions

We have proposed a re-regularization/resampling scheme which allows to represent a retrieval on an appropriate altitude grid such that its averaging kernel becomes unity. This, of course, does not mean that the profile is now free of a priori information but that all a priori is inherent in the grid definition and interpolation scheme, and thus is more obvious to the non-expert data user who neither has the tools available to work with averaging kernel matrices nor wants to work in spaces without obvious physical meaning such as those spanned by the singular vectors of the solution. In a scientific community which is segmented to an extent into data providers and data users who do not interact directly but communicate their results via databases, as favored by, e.g., the responsible officials in the Global Monitoring for Environment and Security (GMES), the importance of easy-to-use data representation will increase in order to avoid misinterpretation of data. Further, the proposed data representation is particularly useful for validation purposes and data assimilation.

Appendix A

Here we show that the concept to calculate the number of degrees of freedom of the retrieval as the trace of the averaging kernel is applicable to smoothing constraints, too. This is not self-evident since the inference of this in Rodgers (2000), Eqs. 2.48–2.56, involves the non-inverted a priori covariance matrix $\mathbf{S}_a$. The smoothing constraint $\mathbf{R}$, which replaces the $\mathbf{S}_a$ term in a Tikhonov-type retrieval is singular due to its rank $n-1$. Smoothing constraints can formally be understood as Bayesian constraints in cases when there is a priori knowledge only on the vertical gradient of the profile but not on the values themselves.

We bypass this problem of singularity of the regularization matrix by decomposing the Tikhonov-type retrieval in an unconstrained part of which the degrees of freedom are trivial to estimate and an optimal estimation/maximum a posteriori-type part, where the arguments of Rodgers (2000) (Eqs. 2.48–2.56) hold.

The retrieval of $\boldsymbol{x}$ can be rewritten as a retrieval of an integrated quantity, e.g. vertical column of a trace species, altitude-averaged temperature etc., plus the retrieval of $n-1$ first order differences of values at adjacent altitude gridpoints. This transformation of the retrieval vector is done by multiplying the vector with matrix **C** of the type

$$\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & & \ddots & & \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

This transformation of the co-ordinate system does not change the degrees of freedom. The Tikhonov-type first order regularization matrix

$$\mathbf{R} = \begin{bmatrix} 1 & -1 & 0 & & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ & & & & \ddots & & & \\ 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \cdots & & -1 & 1 \end{bmatrix}$$

is transformed to

$$\mathbf{C}^{-1^T}\mathbf{R}\mathbf{C}^{-1} = \mathbf{R}^* = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ & & & & \ddots & \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

This proves that the retrieval of the integrated quantity is not constrained by the first order difference smoothing approach. Therefore, by definition, this quantity adds exactly one degree of freedom to the retrieval, and the related diagonal element of the averaging kernel is unity while all off-diagonal elements are zero. The $n-1$ differences now are constrained by a diagonal matrix (lower right $(n-1)\times(n-1)$ block in $\mathbf{R}^*$), which

ACPD

6, 6723–6751, 2006

can be understood as an inverse diagonal covariance matrix in the sense of Bayesian maximum a posteriori retrieval, assuming that there is no knowledge on higher order differences (which would cause off-diaginal elements in the relevant block of $\mathbf{C}^*$. Therefore, the number of degrees of freedom of the retrieval of differences is the trace of the relevant $(n-1)\times(n-1)$ averaging kernel matrix $\mathbf{A}^*$. Thus the number of degrees of freedom of the retrieval is

$$dgf_{\text{retrieval}} = 1 + \text{tr}(\mathbf{A}^*)$$

We recombine the retrieval of the integrated quantity and the differences retrieval and write the related averaging kernel matrix $\mathbf{A}^{*\prime} \in IR^{n\times n}$ as

$$\mathbf{A}^{*\prime} = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{A}^* \end{bmatrix}$$

and find that

$$\text{tr}(\mathbf{A}^{*\prime}) = dgf_{\text{retrieval}}.$$

Since the number of degrees of freedom are conserved when reversibly transforming back to the original coordinate system, we obtain

$$dgf_{\text{retrieval}} = \text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^{*\prime}).$$

Following this approach the equivalence can also be shown for higher order difference operators as regularization constraint, or linear combinations of these.

Acknowledgements. This work was partly funded by BmBF under contract 50 EE 0512. The authors would like to thank F. Hase and G. Kopp for helpfull comments.

References

Calisesi, Y., Soebijanta, V. T., and van Oss, R.: Regridding of remote soundings: Formulation and application to ozone profile comparison, J. Geophys. Res., 110, D23306, doi:10.1029/2005JD006122, 2005. 6725

Ceccherini, S., Carli, B., Pascale, E., Prosperi, M., Raspollini, P., and Dinelli, B. M.: Comparison of measurements made with two different instruments of the same atmospheric vertical profile, Appl. Opt., 42, 6465–6473, 2003. 6725

Endemann, M. and Fischer, H.: Envisat's High-Resolution Limb Sounder: MIPAS, ESA bulletin, 76, 47–52, 1993. 6725

Esler, J. G. and Waugh, D. W.: A method for estimating the extent of denitrification of Arctic polar vortex air from tracer–racer scatter plots, J. Geophys. Res., 107, 4169, doi:10.1029/2001JD001071, 2002. 6741

Fischer, H. and Oelhaf, H.: Remote sensing of vertical profiles of atmospheric trace constituents with MIPAS limb-emission spectrometers, Appl. Opt., 35, 2787–2796, 1996. 6725

Glatthor, N., von Clarmann, T., Fischer, H., Funke, B., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Milz, M., Steck, T., Stiller, G. P., Mengistu Tsidu, G., and Wang, D. Y.: Mixing processes during the Antarctic vortex split in September/October 2002 as inferred from source gas and ozone distributions from ENVISAT-MIPAS, J. Atmos. Sci., 62, 787–800, 2005. 6741

Höpfner, M., von Clarmann, T., Fischer, H., Glatthor, N., Grabowski, U., Kellmann, S., Kiefer, M., Linden, A., Mengistu Tsidu, G., Milz, M., Steck, T., Stiller, G. P., Wang, D.-Y., and Funke, B.: First spaceborne observations of Antarctic stratospheric $ClONO_2$ recovery: Austral spring 2002, J. Geophys. Res., 109, D11308, doi:10.1029/2004JD005322, 2004. 6744

Joiner, J. and da Silva, A. M.: Efficient methods to assimilate remotely sensed data based on information content, Q. J. R. Meteorol. Soc., 124, 1669–1694, 1998. 6725

Kondo, Y., Koike, M., Engel, A., Schmidt, U., Mueller, M., Sugita, T., Kanzawa, H., Nakazawa, T., Aoki, S., Irie, H., Toriyama, N., Suzuki, T., and Sasano, Y.: NO_y–N_2O correlation observed inside the Arctic vortex in February 1997: Dynamical and chemical effects, J. Geophys. Res., 104, 8215–8224, 1999. 6741

Mengistu Tsidu, G., Stiller, G. P., von Clarmann, T., Funke, B., Höpfner, M., Fischer, H., Glatthor, N., Grabowski, U., Kellmann, S., Kiefer, M., Linden, A., López-Puertas, M., Milz, M., Steck, T., and Wang, D. Y.: NO_y from Michelson Interferometer for Passive Atmospheric Sounding on Environmental Satellite during the Southern Hemisphere polar vortex split in September/October 2002, J. Geophys. Res., 110, D11301, doi:10.1029/2004JD005322, 2005. 6741

Müller, R., Groß, J.-U., McKenna, D. S., Crutzen, P. J., Brühl, C., Russell III, J. M., Gordley, L. L., Burrows, J. P., and Tuck, A. F.: Chemical ozone loss in the Arctic vortex in the winter 1995–96: HALOE measurements in conjunction with other observations, Ann. Geophys., 17, 101–114, 1999. 6741

Muscari, G., de Zafra, R. L., and Smyshlyaev, S.: Evolution of the NO_y–N_2O correlation in the Antarctic stratosphere during 1993 and 1995, J. Geophys. Res., 108, 4428, doi:10.1029/2002JD002871, 2003. 6741

Nett, H., Carli, B., Carlotti, M., Dudhia, A., Fischer, H., Flaud, J.-M., Perron, G., Raspollini, P., and Ridolfi, M.: MIPAS Ground Processor and Data Products, in Proc. IEEE 1999 International Geoscience and Remote Sensing Symposium, 28 June–2 July 1999, Hamburg, Germany, 1692–1696, 1999. 6725

Plumb, R. A., Waugh, D. W., and Chipperfield, M. P.: The effects of mixing on tracer relationships in the polar vortices, J. Geophys. Res., 105, 10 047–10 062, 2000. 6741

Ray, E. A., Moore, F. L., Elkins, J. W., Hurst, D. F., Romashkin, P. A., S.Dutton, G., and Fahey, D. W.: Descent and mixing in the 1999–2000 northern polar vortex inferred from in situ tracer measurements, J. Geophys. Res., 107, 8285, doi:10.1029/2001JD000961, 2002. 6741

Rex, M., von der Gathen, P., Braathen, G. O., Harris, N. R. P., Reimer, E., Beck, A., Alfier, R., Krüger-Carstensen, R., Chipperfield, M., de Backer, H., Balis, D., O'Connor, F., Dier, H., Dorokhov, V., Fast, H., Gamma, A., Gil, M., Kyrö, E., Litynska, Z., Mikkelsen, I. S., Molyneux, M., Murphy, G., Reid, S. J., Rummukainen, M., and Zerefos, C.: Chemical Ozone Loss in the Arctic Winter 1994/95 as determined by the Match Technique, Atmos. Environ., 32, 35–59, 1999. 6741

Ridolfi, M., Carli, B., Carlotti, M., von Clarmann, T., Dinelli, B., Dudhia, A., Flaud, J.-M., Höpfner, M., Morris, P. E., Raspollini, P., Stiller, G., and Wells, R. J.: Optimized Forward and Retrieval Scheme for MIPAS Near-Real-Time Data Processing, Appl. Opt., 39, 1323–1340, 2000. 6725

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, vol. 2 of Series on Atmospheric, Oceanic and Planetary Physics, edited by: Taylor, F. W., World Scientific, 2000. 6724, 6725, 6726, 6727, 6728, 6732, 6742, 6745

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res., 108, 4116, doi:10.1029/2002JD002299, 2003. 6725, 6739

SAG: MIPAS, Michelson Interferometer for Passive Atmospheric Sounding, An ENVISAT Instrument for Atmospheric Chemistry and Climate Research, Scientific Objectives, Mission Concept and Feasibility, Instrument Design and Data Products, European Space Agency, 1999. 6725

Steck, T.: Methods for determining regularization for atmospheric retrieval problems, Appl. Opt., 41, 1788–1797, 2002. 6727, 6728

Steck, T. and von Clarmann, T.: Constrained profile retrieval applied to the observation mode of the Michelson Interferometer for Passive Atmospheric Sounding, Appl. Opt., 40, 3559–3571, 2001. 6725, 6727, 6744

Stiller, G. P. (Ed.): The Karlsruhe Optimized and Precise Radiative Transfer Algorithm (KOPRA), vol. FZKA 6487 of Wissenschaftliche Berichte, Forschungszentrum Karlsruhe, 2000. 6744

Tikhonov, A.: On the solution of incorrectly stated problems and method of regularization, Dokl. Akad. Nauk. SSSR, 151, 501–504, 1963. 6725, 6727, 6728, 6733, 6737

Twomey, S.: On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature, Journal of the ACM, 10, 97–101, 1963. 6725, 6728

von Clarmann, T., Ceccherini, S., Doicu, A., Dudhia, A., Funke, B., Grabowski, U., Hilgers, S., Jay, V., Linden, A., López-Puertas, M., Martín-Torres, F.-J., Payne, V., Reburn, J., Ridolfi, M., Schreier, F., Schwarz, G., Siddans, R., and Steck, T.: A blind test retrieval experiment for infrared limb emission spectrometry, J. Geophys. Res., 108, 4746, doi:10.1029/2003JD003835, 2003a. 6726

von Clarmann, T., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Mengistu Tsidu, G., Milz, M., Steck, T., Stiller, G. P., Wang, D. Y., Fischer, H., Funke, B., Gil-López, S., and López-Puertas, M.: Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), J. Geophys. Res., 108, 4736, doi:10.1029/2003JD003602, 2003b. 6724, 6726, 6744

Wetzel, G., Oelhaf, H., Ruhnke, R., Friedl-Vallon, F., Kleinert, A., Kouker, W., Maucher, G., Reddmann, T., Seefeldner, M., Stowasser, M., Trieschmann, O., von Clarmann, T., and Fischer, H.: NO_y partitioning and budget and its correlation with N_2O in the Arctic vortex and in summer midlatitudes in 1997, J. Geophys. Res., 107, 4280, doi:10.1029/2001JD000916, 2002. 6741

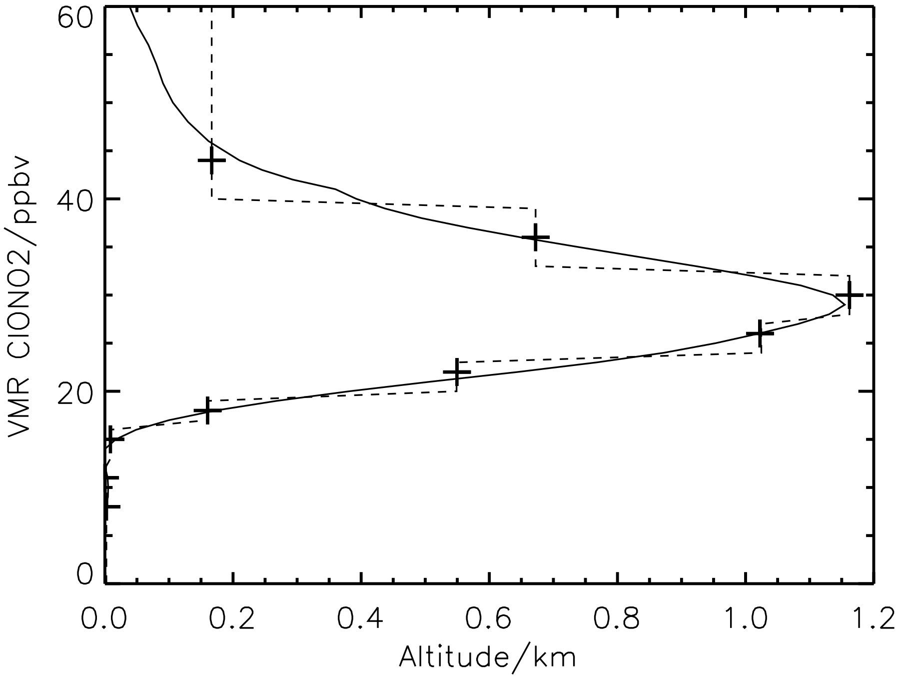

Fig. 1. A retrieval of MIPAS $ClONO_2$ on a fine altitude grid (solid line) and the related re-regularized resampled profile which contains nearly the same number of degrees of freedom and is, except for the coarse grid itself, free of a priori information (dashed line). + signs indicate the selected altitude gridpoints of the coarse grid.

ACPD
6, 6723–6751, 2006