Atmospheric
Measurement
Techniques

Insights into Tikhonov regularization: application to trace gas column retrieval and the efficient calculation of total column averaging kernels

T. Borsdorff, O. P. Hasekamp, A. Wassmann, and J. Landgraf

SRON – Netherlands Institute for Space Research, Utrecht, the Netherlands

Correspondence to: T. Borsdorff (t.borsdorff@sron.nl)

Abstract. Insights are given into Tikhonov regularization and its application to the retrieval of vertical column densities of atmospheric trace gases from remote sensing measurements. The study builds upon the equivalence of the least-squares profile-scaling approach and Tikhonov regularization method of the first kind with an infinite regularization strength. Here, the vertical profile is expressed relative to a reference profile. On the basis of this, we propose a new algorithm as an extension of the least-squares profile scaling which permits the calculation of total column averaging kernels on arbitrary vertical grids using an analytic expression. Moreover, we discuss the effective null space of the retrieval, which comprises those parts of a vertical trace gas distribution which cannot be inferred from the measurements. Numerically the algorithm can be implemented in a robust and efficient manner. In particular for operational data processing with challenging demands on processing time, the proposed inversion method in combination with highly efficient forward models is an asset. For demonstration purposes, we apply the algorithm to CO column retrieval from simulated measurements in the 2.3 μm spectral region and to O_3 column retrieval from the UV. These represent ideal measurements of a series of spaceborne spectrometers such as SCIAMACHY, TROPOMI, GOME, and GOME-2. For both spectral ranges, we consider clear-sky and cloudy scenes where clouds are modelled as an elevated Lambertian surface. Here, the smoothing error for the clear-sky and cloudy atmosphere is significant and reaches several percent, depending on the reference profile which is used for scaling. This underlines the importance of the column averaging kernel for a proper interpretation of retrieved column densities. Furthermore, we show that the smoothing due to regularization can be underestimated by calculating the column averaging kernel on a too coarse vertical grid. For both retrievals, this effect becomes negligible for a vertical grid with 20–40 equally thick layers between 0 and 50 km.

1 Introduction

Measurements of the vertically integrated column density of atmospheric trace gases are primarily obtained by remote sensing techniques from the ground and space. From space, global information about the vertically integrated column density of atmospheric trace gases was obtained by the following studies. In the ultraviolet, radiance measurements of several spaceborne spectrometers (e.g. the Solar Backscattered Ultra Violet (SBUV) instrument Bhartia et al. (2013), the Global Ozone Monitoring Experiment (GOME) Burrows et al. (1999), and the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) Bovensmann et al. (1999)) have demonstrated the unique ability to measure the vertically integrated column density of several trace gases (e.g. NO_2 Richter and Burrows (2002) and O_3 Lerot et al. (2010)). In the short-wave infrared, nadir-viewing measurements of SCIAMACHY and the Greenhouse Gases Observing Satellite (GOSAT) provide information on the total column of CH_4 and CO_2 (Frankenberg et al., 2006; Butz et al., 2011; Schepers et al., 2012; Frankenberg et al., 2011). In addition, SCIAMACHY

measurements around 2.3 μm are used to retrieve the total column densities of CO (Buchwitz and Burrows, 2004; Gloudemans et al., 2008; Frankenberg et al., 2005). From the ground, long-term measurements of vertically integrated column densities of atmospheric trace gases have been performed in the ultraviolet (UV), visible (VIS), short-wave infrared (SWIR) and mid-infrared (mid-IR) ranges by Dobson and Brewer, differential optical absorption spectroscopy (DOAS), and Fourier transform infrared (FTIR) spectrometers. The atmospheric abundance of several trace gases are retrieved (e.g. O_3 (Schneider et al., 2008; Barret et al., 2002), HNO_3 (Vigouroux et al., 2007; Griesfeller et al., 2005), CO (Yurganov et al., 2005, 2004; Borsdorff and Sussmann, 2009; Rinsland et al., 2002; Pougatchev and Rinsland, 1995), CO_2 (Yang et al., 2002), and CH_4 (Angelbratt et al., 2011; Washenfelder et al., 2003)). For the SWIR and mid-IR range, these measurements are provided by the Network for the Detection of Atmospheric Composition Change (NDACC) Infrared Working Group (IRWG) and the Total Carbon Column Observing Network (TCCON).

A trace gas column retrieval represents a typical inversion problem of atmospheric remote sensing with limited vertical sensitivity. For measurements with a sensitivity limited to the total column abundance of a trace gas, an unregularized profile retrieval would infer a vertical distribution which is dominated by the contribution of measurement noise. Thus, the profile inversion problem is ill-posed and requires regularization to obtain a stable solution. In practice, there are two common ways to regularize the least-squares solution. First, a vertical trace gas profile is retrieved using a Tikhonov regularization approach or optimal estimation to stabilize the inversion. Due to the regularization, the retrieved profile estimates a smoothed version of the true profile, where the smoothing is described by the so-called profile averaging kernel. Subsequently, the retrieved profile and the profile averaging kernel are vertically integrated. By that, an analytical expression is given for the so-called column averaging kernel. It describes the sensitivity of the retrieved column with respect to changes of the true vertical trace gas distribution as a function of altitude and is defined by corresponding derivatives. In an ideal case, the column averaging kernel represents a geometrical integration of the true trace gas profile, whereas in practice it describes a vertically weighted height integration. Here, the weights of the integration are mainly due to atmospheric scattering and the temperature dependence of atmospheric absorption. Hence, for the further usage of the retrieved trace gas columns (e.g. for assimilation or validation) it is essential to account for the total column averaging kernel to prevent misinterpretations. Despite its theoretical advantages, only a few retrieval algorithms derive trace gas columns and the according total column averaging kernels via a profile retrieval.

Second, more frequently used is a regularization approach which is known as profile scaling. Here, for a representative vertical profile, a scaling factor is determined by a standard least-squares fit. For the interpretation of satellite observations, this technique is employed by, for example, Gloudemans et al. (2008) for CO column retrieval from SCIAMACHY measurements in the short-wave infrared at around 2.3 μm and by Lerot et al. (2010) for ozone column retrieval from GOME-2 measurements. Furthermore, this method is the official retrieval approach for the considered species of the TCCON network. The retrieval approach is capable of producing column measurements of trace gases with high precision, which even fulfils the strict requirements of the TCCON network (Wunch et al., 2010). The implementation of this approach is straightforward and its numerical performance is beneficial with respect to computation time. Its main drawback is the lack of the corresponding column averaging kernel. Buchwitz and Burrows (2004) estimated the column averaging kernel for the profile-scaling retrieval by a numerical perturbation of a least-squares fitting. Although valid in general, such an approach can only be applied to a few cases which are assumed to be representative of the overall retrieval product. Furthermore, the accuracy of the numerical perturbation requires careful tuning of the perturbation strength and may result in numerical instability.

The studies of von Clarmann and Grabowski (2007) and Sussmann and Borsdorff (2007) showed that the profile-scaling approach represents a particular form of a Tikhonov profile retrieval for an infinite regularization strength. Despite all the advantages of these types of algorithms including an expression for the averaging kernel, it relies on a full profile retrieval with n layers and for practical applications, in its limit of the regularization strength, the risk of numerical instabilities is involved. Therefore, an operational implementation of this approach would negate the computational advantage of the original profile-scaling method. For the DOAS, Eskes and Boersma (2003) derived a method to determine column averaging kernels which is applicable for optically thin absorbers. A corresponding method for a column retrieval using profile scaling has not yet been reported.

In this study, we present a concept for the retrieval of vertically integrated column densities of atmospheric trace gases from remote sensing measurements with a sensitivity limited to the total column of a trace gas. The approach relies on fitting the total column of a trace gas by a least-squares scaling of a reference profile and provides, in addition, an analytical expression for the column averaging kernel. However, the approach preserves all advantages of a robust numerical implementation of the least-squares scaling approach. It combines the advantages of both regularization strategies, providing an analytical expression for the column averaging kernel with a straightforward numerical implementation. Due to that, it is suited in particular for operational data processing in combination with highly efficient radiative transfer simulations. The paper is structured as followed: Sect. 2.1 provides the retrieval framework of the study, and summarizes Tikhonov regularization of first order, and Sect. 2.2 discusses the generalized singular value decomposition for

the clear representation of the regularized solution. This allows us to discuss the profile-scaling approach in the context of Tikhonov regularization of first order in Sect. 2.3, which leads us to the aspired formulation of the column averaging kernel. The application to CO and O_3 column retrieval is discussed in Sect. 3, which provides further insights into the interpretation of the profile-scaling approach. The same section addresses the representation error of the column averaging kernel. Finally, Sect. 4 gives an outline of our findings about Tikhonov regularization. The new approach to calculate total column averaging kernels for profile-scaling retrievals is summarized and its practical relevance is discussed. For the CO and O_3 total column retrieval, the significance of the smoothing errors is analysed, and we conclude about the representation error of the total column averaging kernel.

2 Inversion problem solved by regularization

2.1 Tikhonov regularization

Tikhonov regularization is a common method to regularize ill-posed inversion problems (Phillips, 1962; Tikhonov, 1963a, b; Twomey, 1963), and in this section we summarize its general concept. To retrieve information about the abundance of an atmospheric trace gas from a measurement y_{meas} of spectral dimension m, we employ a forward model F which describes the measurement within the spectral error e_y, namely

$$y_{meas} = F(x) + e_y. \tag{1}$$

The n dimensional state vector x represents the vertical distribution of a trace gas. A further specification of x or y is not required. After a Taylor expansion of F around a first guess profile x_0, Eq. (1) can be written as

$$y = Kx + e_y \tag{2}$$

with $y = y_{meas} - F(x_0) + Kx_0$ and the Jacobian or kernel matrix $K = \partial F / \partial x(x_0)$. Furthermore, we assume that the measurement noise is described by a non-singular measurement error covariance matrix $S_e \in \mathbb{R}^{m \times m}$. The retrieval of a trace gas abundance involves finding a state vector $\hat{x}$ which reproduces the measurement y within its error via the forward model F. For a remote sensing application with a sensitivity limited to the total column of a trace gas, the inversion of Eq. (2) represents an ill-posed problem, and thus the standard least-squares solution

$$\hat{x} = \min_x \{ \| S_e^{-1/2} (Kx - y) \|_2^2 \} \tag{3}$$

is not unique. To explain this in more detail, we apply a singular value decomposition (SVD) to the Jacobian K, $K = U\Sigma V^T$, and so Eq. (2) can be written as

$$y = \sum_{i=1}^{N} \sigma_i (v_i^T x) u_i + e_y, \tag{4}$$

where u_i and v_i are the column vectors of matrix U and V and the singular values σ_i form the diagonal of matrix Σ. For a discrete ill-posed problem, the singular values can be always ordered such that they decay gradually to zero. This means that from a certain index n the singular values are so small that corresponding terms in Eq. (4) do not significantly contribute to the simulated measurement considering the measurement noise e_y. Consequently, any attempt to determine these contributions infers predominately measurement noise which is amplified in the inversion due to the small singular values. In analogy to the null space of a singular linear problem, Hasekamp and Landgraf (2001) refer to the effective null space, i.e. the part of the state space about which no information can be inferred from the measurement. Thus to reduce the noise propagation on the solution, the least-squares solution has to be regularized. For this purpose, we employ the Tikhonov regularization technique (Phillips, 1962; Tikhonov, 1963a, b; Twomey, 1963), and thus the inverse problem can be formulated as a minimization problem of the following cost function:

$$\hat{x}_{reg} = \min_x \{ \| S_e^{-1/2} (Kx - y) \|_2^2 + \lambda^2 \| L_{n-p} (x - x_a) \|_2^2 \}, \tag{5}$$

where $\| \cdot \|_2$ represents the L_2 norm. The rationale for the side constraint in Eq. (5) is to reduce the noise propagation on the solution and at the same time to extract as much information as possible from the measurement. Here, the difference between the state vector x and an a priori state vector x_a is weighted by a smoothing operator $L_{n-p} \in \mathbb{R}^{p \times n}$, the so-called regularization matrix (Hansen, 1998). The subscript $n - p$ indicates the degree of the corresponding derivative. Generally x_a differs from the linearization point x_0 of the forward model. As an alternative to the side constraint in Eq. (5), one may choose the side constraint $\lambda^2 \| L_{n-p} x \|^2$. Common regularization matrices are the unity matrix $L_0 = I$ or a discrete version of the first derivative:

$$L_1(i, j) = \begin{cases} 1 & \text{if } i = j \\ -1 & \text{if } i = j - 1 \\ 0 & \text{otherwise} \end{cases} . \tag{6}$$

The regularization parameter λ balances the two contributions of the cost function shown in Eq. (5) and thus its value is of crucial importance for the inversion. If λ is chosen too large, the noise contribution to the solution of the measurement is low, but the least-squares residual norm deviates significantly from its minimum. On the other hand, if λ is chosen too small, the measurement is fitted well but the solution norm is high, and therefore the solution is overwhelmed by noise. Thus, λ should be chosen such that the two contributions of the cost function are well balanced. In fact, finding an appropriate value for λ comes along with the definition of the effective null space of the retrieval and frequently the L-curve method (Hansen, 1992, 1993) is used for this purpose.

Formally, the solution of Eq. (5) can be expressed by the gain matrix $\mathbf{G}_{\text{reg}}$,

$$\hat{x}_{\text{reg}} = x_{\text{a}} + \mathbf{G}_{\text{reg}}\, y \tag{7}$$

with

$$\mathbf{G}_{\text{reg}} = \left(\mathbf{K}^T \mathbf{S}_{\text{e}}^{-1} \mathbf{K} + \lambda^2 \cdot \mathbf{L}_{n-p}^T \mathbf{L}_{n-p} \right)^{-1} \mathbf{K}^T \mathbf{S}_{\text{e}}^{-1}. \tag{8}$$

Due to the regularization, the retrieved profile $\hat{x}$ is a smoothed version of the true profile x_{true}. The smoothing can be characterized by the averaging kernel,

$$\mathbf{A} = \frac{\partial \hat{x}}{\partial x_{\text{true}}} = \mathbf{G}_{\text{reg}}\mathbf{K}, \tag{9}$$

and thus the retrieved state vector $\hat{x}_{\text{reg}}$ in Eq. (7) can be written as

$$\hat{x}_{\text{reg}} = \mathbf{A}\,x_{\text{true}} + (\mathbf{I} - \mathbf{A})\,x_{\text{a}} + e_x, \tag{10}$$

where $e_x = \mathbf{G}_{\text{reg}}\,e_y$ represents the error on the retrieved trace gas profile caused by the error e_y on the measurement. The averaging kernel $\mathbf{A}$ smoothes a profile such that $\mathbf{A}\,x_{\text{true}}$ is the part of the solution which is determined from the measurement. However, $(\mathbf{I} - \mathbf{A})$ is the part of the state space to which the measurement, and thus the retrieval, is effectively not sensitive. Thus, the term $(\mathbf{I} - \mathbf{A})\,x_{\text{a}}$ describes the effective null space contribution of the a priori profile x_{a}. It is also known as the smoothing error of the retrieval. As stated before, the definition of the effective null space depends on an appropriate chosen regularization parameter λ.

In the following, we use the side constraint $\lambda^2 \|\mathbf{L}_{n-p}\,x\|^2$. Equation (10) reduces to

$$\hat{x}_{\text{reg}} = \mathbf{A}\,x_{\text{true}} + e_x \tag{11}$$

and hence does not include an effective null space contribution. This representation of the data product is beneficial for its use in data assimilation schemes because it eases the assimilation of observations. When an estimate of the effective null space contribution is needed, it can always be added to the retrieved state vector after the inversion (Rodgers, 2000). For this reason, we ignore this contribution in the following.

For a proper error characterization of the inversion, the retrieval error covariance matrix $\mathbf{S}_x$ is needed. It can be calculated from the measurement covariance matrix $\mathbf{S}_y$ by

$$\mathbf{S}_x = \mathbf{G}_{\text{reg}}\mathbf{S}_y\mathbf{G}_{\text{reg}}^T. \tag{12}$$

The column density $\hat{c}$ of a trace gas is defined by vertical profile integration, namely

$$\hat{c} = \mathbf{C}^T \hat{x}_{\text{reg}}, \tag{13}$$

where $\mathbf{C} = (f_1, \ldots, f_n)$ approximates the corresponding integration. Here, f_k converts the element x_k of the retrieved state vector to the partial column amount of the trace gas in model layer k. The particular form of f_k depends on the units of state vector x and on the chosen vertical grid. Using this formulation, we can characterize the effect of regularization on the column $\hat{c}$ via

$$\hat{c} = \mathbf{A}_{\text{c}}\,x_{\text{true}} + e_{\text{c}}, \tag{14}$$

where $\mathbf{A}_{\text{c}} = \mathbf{C}^T \mathbf{A}$ is the column averaging kernel and e_{c} is the error on the retrieved column. The corresponding effective null space contribution of an a priori profile x_{a} is $(\mathbf{C} - \mathbf{A}_{\text{c}})\,x_{\text{a}}$. The retrieval noise on the retrieved column is given by the standard deviation

$$\sigma_{\text{c}} = \sqrt{\mathbf{C}^T \mathbf{S}_x\,\mathbf{C}}. \tag{15}$$

In this manner, all diagnostic tools suited for the retrieval of the state vector $\hat{x}$ can be transformed to the corresponding diagnostics for the retrieved column $\hat{c}$.

2.2 A general analytic form for the solution

To study the analytic solution $\hat{x}$ in Eq. (7), we first transform the cost function (Eq. 5) into a uniform noise representation by substituting $\tilde{\mathbf{K}} = \mathbf{S}_{\text{e}}^{-1/2}\mathbf{K}$ and $\tilde{y} = \mathbf{S}_{\text{e}}^{-1/2}\,y$, which yields

$$\hat{x} = \min_x \{\|\tilde{\mathbf{K}}x - \tilde{y}\|_2^2 + \lambda^2 \|\mathbf{L}_{n-p}x\|_2^2\} \tag{16}$$

or

$$\hat{x} = \tilde{\mathbf{G}}_{\text{reg}}\tilde{y} \tag{17}$$

with

$$\tilde{\mathbf{G}}_{\text{reg}} = \left(\tilde{\mathbf{K}}^T \tilde{\mathbf{K}} + \lambda \mathbf{L}_{n-p}^T \mathbf{L}_{n-p} \right)^{-1} \tilde{\mathbf{K}}^T. \tag{18}$$

This equation can be simplified using the general single-value decomposition (GSVD) of the matrix pair $(\tilde{\mathbf{K}}, \mathbf{L}_{n-p})$ (e.g. Hansen, 1992). For $m \geq n \geq p$, the GSVD of the matrix pair $(\tilde{\mathbf{K}}, \mathbf{L}_{n-p})$ is

$$\begin{aligned}\tilde{\mathbf{K}} &= \mathbf{U}\mathbf{D}_K\mathbf{W}^{-1} \\ \mathbf{L}_{n-p} &= \mathbf{V}\mathbf{D}_L\mathbf{W}^{-1}\end{aligned} \tag{19}$$

where the non-singular matrix $\mathbf{W}^{-1} \in \mathbb{R}^{n \times n}$ represents a new basis of the state space in which both matrices $\tilde{\mathbf{K}}$ and $\mathbf{L}_{n-p}$ can be represented as the diagonal matrices $\mathbf{D}_K \in \mathbb{R}^{n \times n}$ and $\mathbf{D}_L \in \mathbb{R}^{p \times n}$. $\mathbf{U} \in \mathbb{R}^{m \times n}$ and $\mathbf{V} \in \mathbb{R}^{p \times p}$ are the corresponding back projections with orthonormal columns u_i and v_i, respectively. Therefore, the equations $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, $\mathbf{V}^T\mathbf{V} = \mathbf{I}$, $\tilde{\mathbf{K}}^T\tilde{\mathbf{K}} = \mathbf{W}^{-T}\mathbf{D}_K^2\mathbf{W}^{-1}$, and $\mathbf{L}_{n-p}^T\mathbf{L}_{n-p} = \mathbf{W}^{-T}\mathbf{D}_L^T\mathbf{D}_L\mathbf{W}^{-1}$ are valid. The matrices $\mathbf{D}_K$, $\mathbf{D}_L$ are both diagonal matrices of the form

$$\mathbf{D}_K = \begin{pmatrix} \Sigma & 0 \\ 0 & \mathbf{I}_{n-p} \end{pmatrix}, \quad \mathbf{D}_L = (\mathbf{M},\, 0), \tag{20}$$

where $\Sigma = \text{diag}(\sigma_1, \ldots \sigma_p)$, $\mathbf{M} \in \mathbb{R}^{p \times p}$ hold the corresponding singular values of $(\tilde{\mathbf{K}}, \mathbf{L}_{n-p})$ and $\mathbf{I}_{n-p} \in \mathbb{R}^{(n-p) \times (n-p)}$ is

the unity matrix. As a result, the column vectors $\boldsymbol{w}_i$ of $\mathbf{W}$ for $i = p+1, \ldots n$ span the null space of the regularization matrix $\mathbf{L}_{n-p}$,

$$N\left(\mathbf{L}_{n-p}\right) = \text{span}\left\{\boldsymbol{w}_i, \quad \text{with} \quad i = p+1, \ldots n\right\}. \quad (21)$$

$N\left(\mathbf{L}_{n-p}\right)$ should not be confused with the effective null space of the retrieval as is defined in the previous section.

With this decomposition, we can rewrite the gain matrix in Eq. (18) as

$$\tilde{\mathbf{G}}_{\text{reg}} = \mathbf{W}\left(\Phi\mathbf{D}_K^{-1}\mathbf{U}^T\right) \quad (22)$$

with the diagonal matrix $\Phi = (\mathbf{D}_K^2 + \lambda\mathbf{D}_L^T\mathbf{D}_L)^{-1}\mathbf{D}_K^2$. Equation (18) can also be represented as a linear combination of the column vector $\boldsymbol{w}_j$ of $\mathbf{W}$ since $\Phi_{j,j} = 1$ if $j > p$, namely

$$\tilde{\mathbf{G}}_{\text{reg}} = \sum_{j=1}^{p} \Phi_{j,j} \cdot \frac{\boldsymbol{w}_j\boldsymbol{u}_j^T}{\sigma_j} + \sum_{j=p+1}^{n} \boldsymbol{w}_j\boldsymbol{u}_j^T. \quad (23)$$

The filter factor matrix Φ reveals how the retrieval result is affected by the regularization strength parameter λ. By choosing $\lambda = 0$, the filter factor matrix becomes $\Phi = \mathbf{I}$, and therefore

$$\tilde{\mathbf{G}}_{\text{reg}} = \sum_{j=1}^{p} \frac{\boldsymbol{w}_j\boldsymbol{u}_j^T}{\sigma_j} + \sum_{j=p+1}^{n} \boldsymbol{w}_j\boldsymbol{u}_j^T. \quad (24)$$

This is the gain matrix of the unregularized least-squares fit. By choosing $\lambda \to \infty$, the matrix $\Phi_{j,j} = 0$ if $j \leq p$ and $\Phi_{j,j} = 1$ if $j > p$. Therefore, the gain matrix becomes

$$\tilde{\mathbf{G}}_{\text{reg}} = \sum_{j=p+1}^{n} \boldsymbol{w}_j\boldsymbol{u}_j^T. \quad (25)$$

That means that the solution space of the minimization problem is equal to the null space $N(\mathbf{L}_{n-p})$ of the regularization matrix. In other words, no state vector of the null space $N(\mathbf{L}_{n-p})$ is affected by the smoothing of the averaging kernel, i.e.

$$\boldsymbol{x} = \mathbf{A}\boldsymbol{x} \quad \text{for} \quad \boldsymbol{x} \in N\left(\mathbf{L}_{n-p}\right) \quad \text{and for} \quad \lambda \to \infty \quad (26)$$

and so

$$\text{trace } \mathbf{A} = n - p \quad \text{for} \quad \lambda \to \infty, \quad (27)$$

which is commonly known as the degrees of freedom of the retrieval.

2.3 Profile-scaling retrieval and the total column averaging kernel

Profile-scaling retrievals are widely used as described in Sect. 1. Here, a scaling factor of a reference profile is inferred from a measurement via an unregularized least-squares fit. It

is important to distinguish the reference profile from an a priori profile which is used to fill up the effective null space (see Sect. 2.1). It will be shown later that the reference profile does not have an effective null space contribution. So algebraically, the reference profile cannot be used to fill up the effective null space of the retrieval. The inversion relies on a one-parameter least-squares fit of the scaling factor x_{lsq} with the gain matrix

$$\boldsymbol{g}_{\text{lsq}} = \left(\mathbf{K}_{\text{lsq}}^T\mathbf{S}_{\text{e}}^{-1}\mathbf{K}_{\text{lsq}}\right)^{-1}\mathbf{K}_{\text{lsq}}^T\mathbf{S}_{\text{e}}^{-1} \quad (28)$$

and the Jacobian vector $\mathbf{K}_{\text{lsq}}^T \in \mathbb{R}^m$. This inversion does not provide an analytical approach to calculate the total column averaging kernel. It can be calculated by considering the profile-scaling approach as a particular case of Tikhonov regularization of first order for $\lambda \to \infty$ with the gain matrix

$$\tilde{\mathbf{G}}_{\text{reg}} = \boldsymbol{w}_j\boldsymbol{u}_j^T \quad (29)$$

for $p = n - 1$ in Eq. (25). Here, state vector $\boldsymbol{x}$ is defined as the ratio of the trace gas profile $\boldsymbol{\rho}$ with respect to a reference profile $\boldsymbol{\rho}_{\text{ref}}$, thus

$$\boldsymbol{x} = \boldsymbol{\rho}/\boldsymbol{\rho}_{\text{ref}}. \quad (30)$$

In the following, we show that, although different in dimension, both gain matrices $\boldsymbol{g}_{\text{lsq}}$ and $\tilde{\mathbf{G}}_{\text{reg}}$ represent the same solution of the inverse problem. The solution, given by the gain matrix $\tilde{\mathbf{G}}_{\text{reg}}$ in Eq. (29), lies in the null space of the regularization matrix $\mathbf{L}_1$, which is given by $N(\mathbf{L}_1) = \text{span}\left(\{[1\,1\,1\,\ldots\,1]^T\}\right)$. Thus, the null space of $\mathbf{L}_1$ consists of all state vectors which are constant in altitude, and so the gain matrix $\tilde{\mathbf{G}}_{\text{reg}}$ maps a measurement $\boldsymbol{y}$ to identical element of the state vector. In our case, this is the scaling parameter of the reference profile $\boldsymbol{\rho}_{\text{ref}}$. Consequently, the rows $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_n$ of $\tilde{\mathbf{G}}_{\text{reg}}$ are identical, $\boldsymbol{g}_1 = \boldsymbol{g}_2 = \ldots = \boldsymbol{g}_n$. Furthermore, because the scaling parameter of the regularization scheme is also retrieved by the one-parameter least-squares fit, we obtain

$$\boldsymbol{g}_i = \boldsymbol{g}_{\text{lsq}} \quad \text{for} \quad i = 1, \ldots, N. \quad (31)$$

Through this, we can define an analytic method to calculate the total column averaging kernel of a profile-scaling retrieval in an efficient and numerically stable way. Furthermore, this method is also valid for constrained profile-scaling retrievals since those also fulfil the requirement that any state vector must be part of $N(\mathbf{L}_1)$ and therefore also constant in altitude. The method can be summarized in three steps:

1. The scaling of a reference profile is retrieved using a standard least-squares fit with associated gain matrix $\boldsymbol{g}_{\text{lsq}}$.

2. On any arbitrary height grid with $n = 1, \ldots, N$ model layers, the gain matrix $\tilde{\mathbf{G}} = (\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n)^T$ of Tikhonov regularization of the first kind is given by $\boldsymbol{g}_{\text{lsq}}$ with $\boldsymbol{g}_1 = \boldsymbol{g}_2 = \ldots = \boldsymbol{g}_n = \boldsymbol{g}_{\text{lsq}}$.

3. This defines the averaging kernel in Eq. (9), and so the total column averaging kernel in Eq. (14) is given by

$$\mathbf{A}_{\text{col},j} = \boldsymbol{g}_{\text{lsq}} \boldsymbol{k}_j \sum_{k=1}^{N} f_k. \tag{32}$$

with the column vector $\boldsymbol{k}_j$ of the Jacobian $\mathbf{K} = (\boldsymbol{k}_1, \ldots, \boldsymbol{k}_N)$ and the conversion factors

$$f_k = \Delta z_k \, \boldsymbol{\rho}_{\text{ref},k}; \tag{33}$$

here, Δz_k is the geometrical thickness of model layer k, and the physical units of the reference profile $\boldsymbol{\rho}_{\text{ref}}$ determine the units of the column density $\hat{c}$.

The computational cost of our approach is significantly less than the one of the more established methods which relies on strongly regularized profile retrieval (e.g. Sussmann and Borsdorff, 2007). To compare the numerical cost of both methods, we consider the number of multiplications which have to be performed in one iteration step of the inversion. Faster operations like summations are ignored. For simplicity, we assume that a matrix multiplication of a $n \times m$ matrix with a $m \times m$ matrix requires nm^2 operations and the inversion of a $n \times n$ matrix n^3 operations where $\dim(\boldsymbol{y}) = m$ and $\dim(\boldsymbol{x}) = n$. We ignore the inversion of the matrix $\mathbf{S}$ because this is trivial for uncorrelated measurements noise. Our proposed approach needs a total number of $(4m + mn + 1)$ multiplications to do the profile-scaling retrieval and to calculate the associate total column averaging kernel. In contrast, the strongly regularized profile retrieval requires $(2mn + 3nm^2 + n^3)$ multiplications. We consider two particular cases: (1) $m \gg n$, which means that the number of measurement points clearly exceeds number of parameters n that are actually retrieved, and (2), the case with $m = n$. In case 1 our proposed approach is a factor of $3n$ faster than the regularized profile retrieval, and in case 2 even a factor of $4n$. For a typical representation of a profile on $n = 20$–40 layers, this means a speed-up of the inversion by a factor of up to 160. This numerical advantage of the new inversion scheme becomes particularly important in the context of operational data processing, when the computational burden of the radiative transfer model is of the same order or smaller than that of the inversion. For the CO retrieval from the SWIR spectral region, which is introduced in the following chapter, we found that a profile inversion requires about half of the time spent for the forward calculation and that our proposed method reduces this to less than 1 % of the time required for the non-scattering forward calculation ($n = 40$ vertical layers).

To summarize, the presented approach for calculating the total column averaging kernel relies on the numerical implementation of a one-parameter least-squares fit. It is favourable with respect to both numerical implementation and robustness. Furthermore, it provides a straightforward

manner with which to adapt existing algorithms for profile-scaling retrieval with minor modifications. The new concept also helps to further develop the results of other studies. For example, it eases the calculation of interference errors as presented by Sussmann and Borsdorff (2007) since interference kernels in the profile-scaling case can now be calculated directly with an analytical expression without simulating those retrievals first via a Tikhonov regularization of the first kind. This enables the possibility of including such error estimations in the standard output of operational retrieval algorithms.

3   Applications

In this section, we apply the proposed method for total column retrieval by profile scaling to two specific satellite remote sensing problems in order to illustrate the general features of this kind of retrieval. Moreover, we demonstrate the overall need for the column averaging kernels to correctly interpret the column density retrieved from the measurement. First, we will consider the retrieval of the vertical column density of carbon monoxide (CO) from simulated measurements in the 2.3 μm SWIR spectral region. This spectral range is used to determine atmospheric CO abundances from SCIAMACHY SWIR measurements (Gloudemans et al., 2008; Buchwitz and Burrows, 2004, and references therein) and will also be probed by the TROPOMI instrument as payload of the Sentinel 5 Precursor mission (Veefkind et al., 2012), scheduled for launch in 2015. As a second example, we will consider the retrieval of the vertical column density of ozone (O_3) from simulated ultraviolet (UV) measurements as they are done by several spaceborne spectrometers, like GOME (Burrows et al., 1999), GOME-2 (Callies et al., 2000), SCIAMACHY (Eskes et al., 2005), OMI (Levelt et al., 2006), and TROPOMI (Veefkind et al., 2012). For this purpose, we assume a nadir-viewing geometry of reflected sunlight with a viewing zenith angle of $VZA = 0°$ and a solar zenith angle of $SZA = 45°$. For SWIR we assume a typical surface albedo of 0.05 and for the UV retrieval a surface albedo of 0.1. The measurements are simulated for clear-sky and cloudy conditions, where clouds are described by an elevated Lambertian surface at 7.5 km altitude with a cloud albedo of 0.5 for SWIR and 0.8 for UV. For partially cloudy scenes the independent pixel approach (e.g. Marshak et al., 1995) is employed with a cloud fraction f_{cld}. The model atmosphere is adapted from the US standard atmosphere (NOAA, 1976). Figure 1 shows the simulated measurement under clear-sky conditions for the retrieval windows 2324.5–2338.38 nm for CO and 325.04–334.91 nm for O_3. For the SWIR window, we account for atmospheric absorption by CO, H_2O, HDO, and CH_4. In the UV window, O_3 is considered as the only relevant absorber. For this particular example the forward model of the O_3 retrieval accounts for
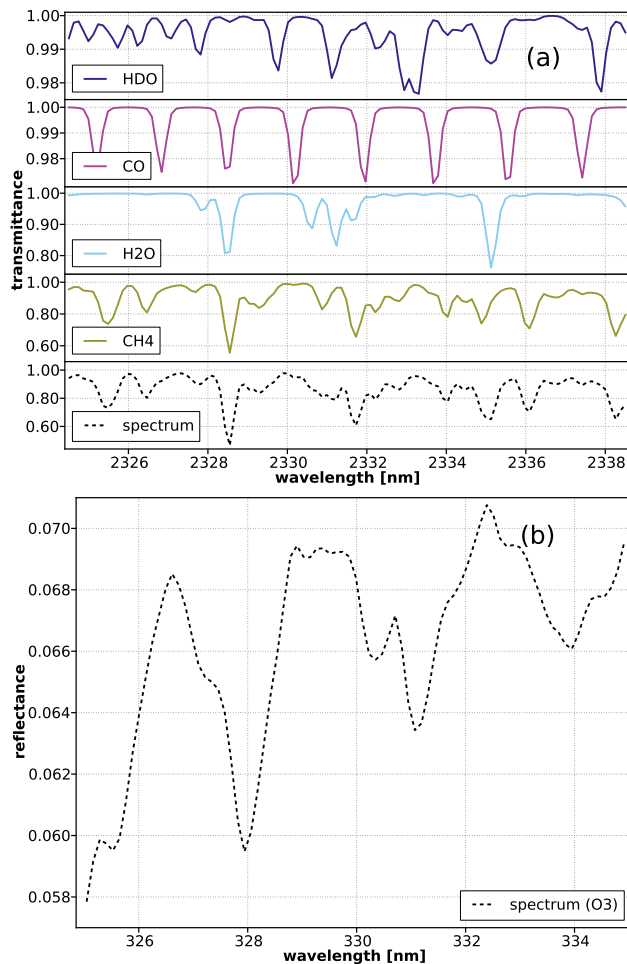
Fig. 1. Simulation of solar absorption spectra. **(a)** CO retrieval window 2324–2339 nm simulated with a spectral resolution of 0.25 nm using HITRAN 2008 spectroscopy (Rothman et al., 2003). The absorptions of the interfering species HDO, H_2O, and CH_4 are separated. **(b)** The corresponding O_3 retrieval window 325–335 nm, which employs the Brion et al. (1993) line list database. All simulations are performed for clear-sky conditions and a solar zenith angle of 45° and a viewing zenith angle of VZA = 0°.

scattering, and the one for the CO retrieval is a non-scattering code.

All species in this study are retrieved using the profile-scaling approach described in the previous section. CO is retrieved together with the interference species H_2O, HDO, and CH_4 to reduce the interference effect caused by the overlapping absorption lines (Sussmann and Borsdorff, 2007; Borsdorff and Sussmann, 2009; Pougatchev and Rinsland, 1995; Rinsland et al., 2002; Rodgers and Connor, 2003). Ozone is inferred without accounting for further atmospheric absorbers. For both retrievals, the measurement noise is assumed to be shot noise with a signal-to-noise ratio of SNR = 100 at the maximum value of the spectrum. Moreover, we also use the profile ρ_{true} that is used to simulate the

measurement spectra as the linearization point for the forward calculation $x_0 = \rho_{true}$. Thus an iterative inversion approach is not needed to account for the non-linearity of the forward model. Throughout this study, the US standard atmosphere profile will serve as the reference profile ρ_{ref} in Eq. (30). Hence, the total column averaging kernels of the retrievals are calculated for the true state of the atmosphere ρ_{true} and reflect the vertical sensitivity of a profile-scaling retrieval which scales the corresponding US standard profile ρ_{ref}.

The resulting total column averaging kernels for the CO and O_3 retrieval are shown in Fig. 2. Here, we assume that the vertical profile contains partial column densities of the individual layers as its components, which implies that the conversion factor in Eq. (33) is $f_k = 1$. This representation eases the interpretation and is commonly used in the literature (e.g. Notholt et al., 2000; Borsdorff and Sussmann, 2009; Rodgers and Connor, 2003). The column averaging kernels differ from the ideal case $\tilde{\mathbf{A}}_c = (1, 1, \ldots, 1)^T$, where $\tilde{\mathbf{A}}_c \mathbf{x}_{true} = c_{true}$. For the clear-sky CO retrieval, $\tilde{\mathbf{A}}_c < 1$ below 5.7 km altitude and $\tilde{\mathbf{A}}_c > 1$ at higher altitudes. The ozone total column averaging kernel shows a more complex shape, with values above and below 1. Only in the range between 21.5 and 29.5 km is it close to its ideal value of unity. For the cloudy case, the retrieved column loses sensitivity to the atmosphere below the cloud but at the same time shows an enhanced retrieval sensitivity above the cloud. This is a typical feature of a profile-scaling approach, and it can be explained most easily for the fully clouded scene with cloud fraction unity. In this case, the total column is determined by the scaling of the reference profile using only the sensitivity above the cloud, and thus a change of the trace gas concentration at these altitudes affects the retrieval twice. First, the profile is adapted above the cloud due to the measurement sensitivity and second also below the cloud, although the measurement is not sensitive to this altitude range. This explains the enhanced value of the column averaging kernel above the cloud and also its dependence on cloud height (not shown). The same rationale is true for the clear-sky and partially cloudy cases, where the altitude ranges of reduced retrieval sensitivity are compensated for by enhanced averaging kernel values at other altitudes. For particular profiles, the differences in the averaging kernel will add up such that the retrieved column is equal to the true column. Due to Eq. (26), this is only the case for profiles that can be expressed as a scaling of the reference profile ρ_{ref}, and so the corresponding state vector $\mathbf{x}$ is element of the null space $N(\mathbf{L}_1)$. Any other profile has an effective null space contribution, which means that the retrieved column is affected by a smoothing error and differs from the true column.

In case that a data product is required that represents an estimate of the true trace gas column, generally one aims to fill up the effective null space using a proper a priori knowledge of the true profile, $(\mathbf{C} - \mathbf{A}_c)\mathbf{x}_a$. One may interpret this term as a correction to the retrieval because the reference
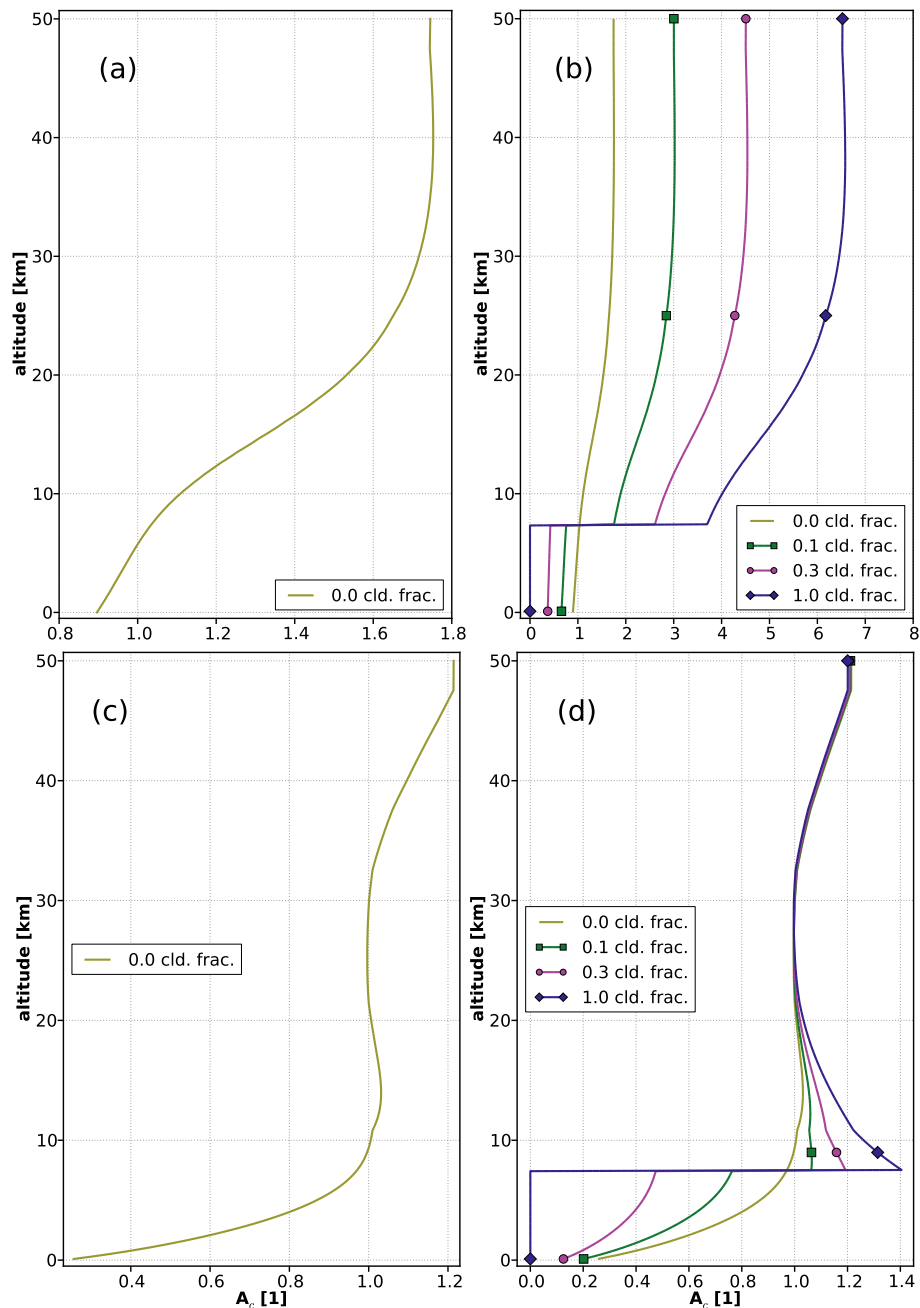
Fig. 2. Total column averaging kernels $\tilde{A}_C$ of the CO **(a, b)** and O_3 **(c, d)** profile-scaling retrieval as function of altitude for different cloud fraction. In the forward calculation a Lambertian surface with albedo 0.5 for SWIR and 0.8 for UV is placed at an altitude of 7.5 km. Measurement geometry and surface albedo are the same as in Fig. 1. The kernels are presented on a vertical grid with 512 equidistant layers.

profile does not represent the true relative distribution of the trace gas. For example, during the processing of an operational retrieval, only a rough estimate of the relative profile is possible. At a later stage, due to sophisticated chemical transport modelling, the estimate can be improved, and so without reprocessing the measurements, the effective null space contribution of the model results can be used to correct the retrieval. In this context, it is interesting to note that

the reference profile cannot be used for this purpose because of its vanishing contribution to the effective null space (see Eq. 26). The same holds for any scaled version of the reference profile.

Thus, the smoothing error $(\mathbf{C} - \mathbf{A}_c)\,\boldsymbol{\rho}_{\text{true}}$ of the scaling approach corresponds to the error when the retrieved column is assumed to be an estimate of the true column. To estimate the relevance of the effective null space contribution,
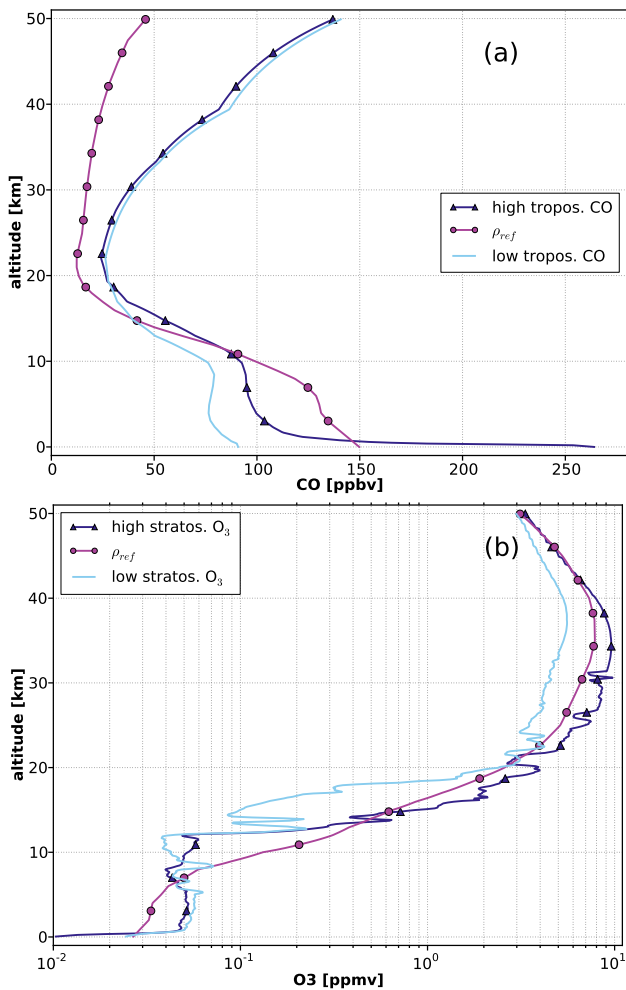
Fig. 3. Ensemble of CO and O_3 volume mixing ratio profiles. **(a)** CO profiles for polluted and unpolluted situations adapted from Levelt et al. (2009) and the CO profile taken from the US standard atmosphere (NOAA, 1976). **(b)** Two O_3 radiosonde measurements at de Bilt, the Netherlands, with high (16 February 2007) and low stratospheric O_3 (19 February 2008). Additionally, the US standard atmosphere ozone profile is depicted.
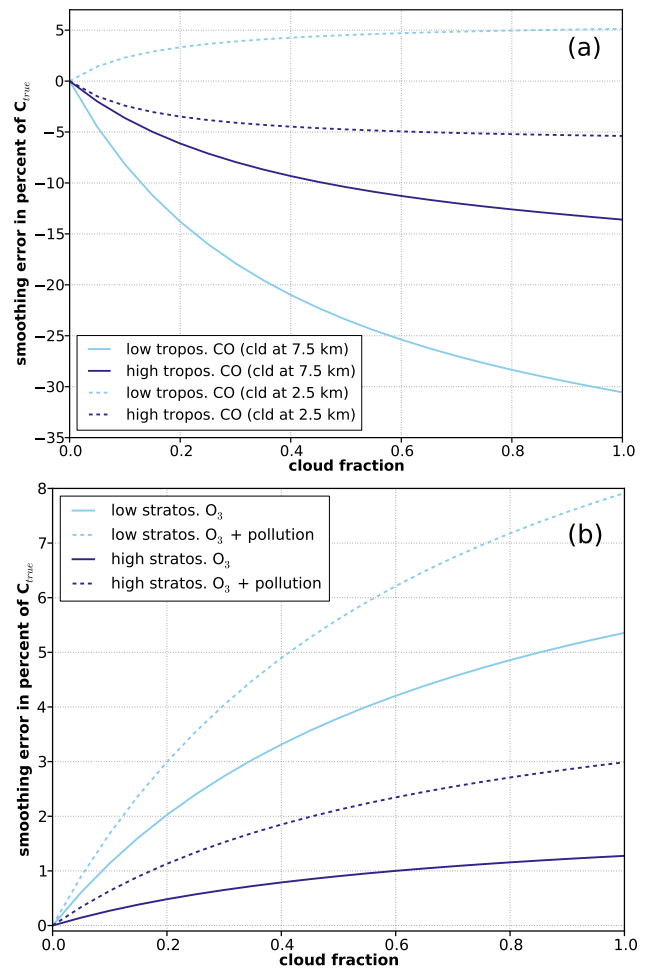
Fig. 4. Smoothing error for cloudy scenes as a function of cloud fraction: **(a)** CO smoothing error for the background and polluted CO profile shown in Fig. 3 for a cloud at 7.5 and 2.5 km, **(b)** O_3 smoothing error for high and low stratospheric ozone profiles shown in Fig. 3, respectively with and without a pollution of 120 ppb in the boundary layer. Values are given in percentage of the known true total column. The difference to the clear-sky case is shown.

we consider measurement simulations for the set of CO and O_3 profiles shown in Fig. 3. The data set comprises a background and a polluted CO profile, two ozone profiles with low and high stratospheric ozone concentrations, and, additionally, both ozone profiles with an enhanced ozone mixing ratio of 120 ppb for all layers below 2.5 km (not shown), which mimics enhanced ozone concentration in the tropospheric boundary layer. Here, all profiles are scaled to the same vertical column density, and as such they differ only in their relative vertical distribution (not shown in Fig. 3). For the retrieval, the US standard profile is employed as the reference profile ρ_{ref} for scaling.

For the different atmospheric profiles, Tables 1 and 2 summarize the smoothing error $(\mathbf{C} - \mathbf{A}_c)\,\rho_{true}$ of the total column

for clear-sky atmospheres and the corresponding contribution of particular partial columns. Here, the partial columns are defined over the maximum altitude range such that the averaging kernel is always > 1 or < 1. The vertical domains indicate the ranges where the retrieval either under- or overestimates the contribution of the true profile to the true total column. When adding up the partial columns to the total column, errors cancel out to a large extent, but overall an error on a percentage level remains for the considered profiles.

The smoothing error increases significantly when one considers partially cloudy scenes. Figure 4 shows the increase of the total column smoothing error as a function of the cloud fraction. For a cloud fraction of $f = 0.6$, the additional smoothing error increases to -25% for the scene of low tropospheric CO concentrations and to -11% for the polluted
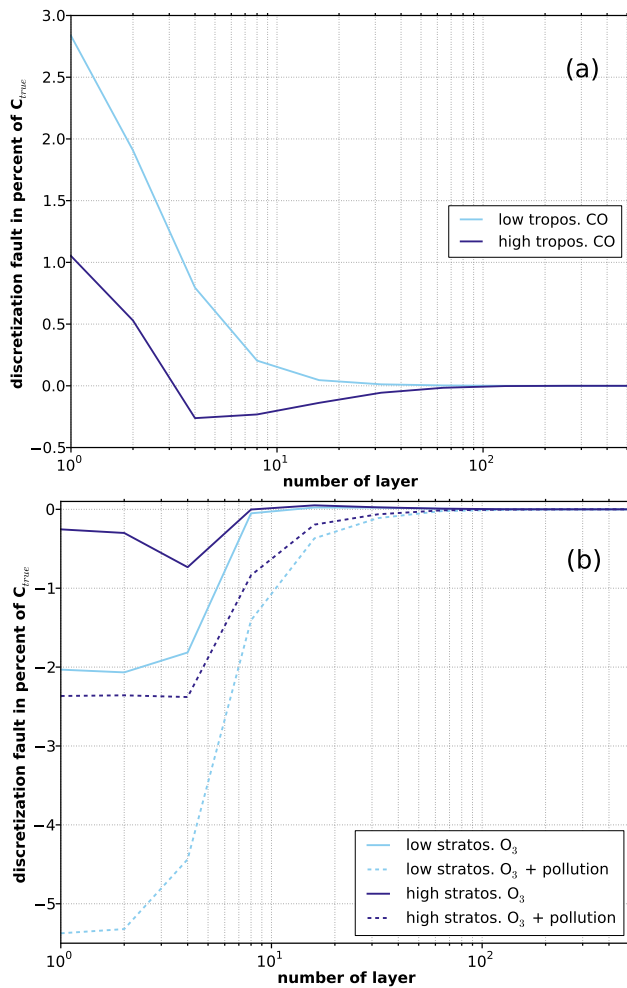
Partial column	High tropos. CO	Low tropos. CO
0–5.7 km	+4.13 %	+3.44 %
5.7–50 km	−5.18 %	−6.27 %
0–50 km	−1.05 %	−2.83 %

in Eq. (32). As a reference, we use a model atmosphere between 0 and 50 km altitude divided into 512 geometrically equidistant layers. Subsequently, we consider the error in the effective column $\mathbf{A}x$ using a vertical grid of N model layers. Figure 5 shows this discretization error as a function of N. Due to the particular form of the column averaging kernel, the discretization error does not always decrease monotonically with an increasing number of model layers. However, a representation of the Jacobian on 20–40 layers is sufficient to reduce the discretization error such that it does not represent a significant error source for the different retrievals. The number of required layers can surely be further reduced by choosing non-equidistant vertical grids which are particularly optimized for a specific application, but the general problem of a discretization error cannot be avoided by choosing different grids.

Fig. 5. Discretization error of the smoothing error for **(a)** CO and **(b)** O$_3$ caused by representing the total column averaging kernel on an equidistant vertical grid with N layers. It is shown how the retrieved column calculated on N layers deviates from the one calculated for 512 layers. Values are given in percentage of the known true total column.

4   Summary and conclusions

In this study, we proposed a concept to retrieve vertical column densities of atmospheric trace gases from remote sensing measurements. The method is based on a least-squares profile-scaling approach, but it allows one to calculate total column averaging kernels via an analytic expression on arbitrary vertical grids. The approach can be implemented in a straightforward manner, and results in a numerically robust and efficient algorithm. In particular, it is suited for operational data processing with high demands on computation time and also provides a straightforward manner to adapt existing algorithms for profile-scaling with minor modifications. For example, we found for the CO total column retrieval from the SWIR spectral region that a profile inversion with $n = 40$ vertical layers needs half of the time required for the non-scattering forward calculation and that our proposed method speeds this up to less than 1 % of the one required for the forward model which also calculates vertically resolved Jacobians.

scene. For the fully clouded scene the additional error increases to -1 and -14 %, respectively. The magnitude of these errors depends on the altitude of the cloud. To demonstrate this, the two cases for CO are recalculated for a cloud placed at 2.5 km and shown in Fig. 4. The smoothing errors of ozone shown in Fig. 4 also significantly increase with higher cloud fractions and can reach up to 8 % for the fully clouded scene. It is interesting to note here that the errors are higher for low stratospheric ozone concentrations. This illustrates the strong dependency of the smoothing error on the shape difference between the reference profile that is used for scaling and the assumed true one.

Finally, we consider the discretization error of the averaging kernel. Here, the vertical gridding of the averaging kernel is given by the vertical discretization of the Jacobian $\mathbf{K}$

We showed that the profile-scaling approach represents a particular form of regularization which is equivalent to a

Table 2. Same as Table 1 but for ozone. Here, the smoothing error is 3to four contributions (0–9.3 km, 9.3–21.5 km, 21.5–29.5 km, and 29.5–50 km), depending on the values of the column averaging kernel. All values are given in percent of the true total column.

Partial column	High O_3	High O_3 + poll.	Low. O_3	Low O_3 + poll.
0–9.3 km	1.61 %	3.67 %	3.27 %	6.53 %
9.3–21.5 km	−0.63 %	−0.61 %	−0.31 %	−0.29 %
21.5–29.5 km	0.09 %	0.09 %	0.09 %	0.09 %
29.5–50 km	−0.82 %	−0.78 %	−1.02 %	−0.96 %
0–50 km	0.25 %	2.37 %	2.03 %	5.37 %

Tikhonov regularization of the first kind with an infinite regularization strength and a vertical profile expressed relative to a reference profile. This equivalence allows us to derive an analytical expression for the total column averaging kernel. Moreover, we showed that such a profile-scaling retrieval does not generally contain an effective null space contribution. This is beneficial for using the data product in data assimilation schemes because it eases the assimilation of observations. Our solution of the inversion problem does not include formal a priori information, which is an advantage because no a priori profiles then have to be provided to the data user. Moreover, we showed algebraically that the reference profile (and any scaled version of it) has a vanishing effective null space contribution and that this cannot be used as a priori profile to fill up the effective null space.

The proposed regularization scheme is of a general nature, and thus it can be applied to many retrieval problems using spaceborne or ground-based remote sensing measurements. For demonstration, we applied it to the CO column retrievals from simulated spectra in the 2.3 µm region and to O_3 column retrievals in the ultraviolet spectral range. This represents the retrieval concept for a series of spaceborne spectrometers like SCIAMACHY, TROPOMI, GOME, and GOME-2. For both retrievals, we considered clear-sky and cloudy scenes where clouds were modelled as an elevated Lambertian surface. We illustrated the dependence of the total column averaging kernel on the cloud coverage of the observed scene. Here, altitudes with a reduced retrieval sensitivity are compensated for by enhanced values at other altitudes, which is a typical characteristic of a profile-scaling retrieval. So the retrieved column may be interpreted as an estimation of the true vertical column density, even when the measurement is not sensitive for the full altitude range. Consequentially, the smoothing error represents the error due to the fact that the scaled reference profile is not the true profile, and therefore the profile scaling is deficient.

By using the US standard model atmosphere to define the reference profile, and by considering both polluted and unpolluted atmospheric abundances and high and low stratospheric ozone for the true profile, we found the smoothing errors for both retrievals in the clear-sky case to be significant, causing errors of up to −2.83 % of the true vertical column density for CO and 5.37 % for O_3. For cloudy cases with a cloud top at 7.5 km, an additional smoothing error occurs which may reach −30 % for CO and 8 % for O_3, depending on cloud coverage. The particular values of the smoothing error depend on cloud altitude and the chosen reference profile. In the ideal case, where the relative distributions of the reference and true profiles are equal, the smoothing error will vanish in all cases. However, in practice the use of the total column averaging kernel is essential for the correct interpretation of retrieved data, in particular for cloudy observations. Here it is recommended to represent the column averaging kernel on a vertical grid with 20–40 equally thick layers that extend between 0 and 50 km to avoid significant discretization errors in the estimate of the smoothing error.

The presented algorithm will be used for the operational data processing of CO columns from TROPOMI measurements. Its functionality will be tested on real SCIAMACHY data for the purpose of CO column retrieval and on real GOME-2 data for the purpose of O_3 column estimates in the near future.

References

Angelbratt, J., Mellqvist, J., Blumenstock, T., Borsdorff, T., Brohede, S., Duchatelet, P., Forster, F., Hase, F., Mahieu, E., Murtagh, D., Petersen, A. K., Schneider, M., Sussmann, R., and Urban, J.: A new method to detect long term trends of methane (CH_4) and nitrous oxide (N_2O) total columns measured within the NDACC ground-based high resolution solar FTIR network, Atmos. Chem. Phys., 11, 6167–6183, doi:10.5194/acp-11-6167-2011, 2011.

Barret, B., de Mazière, M., and Demoulin, P.: Retrieval and characterization of ozone profiles from solar infrared spectra at the Jungfraujoch, J. Geophys. Res., 107, 4788, doi:10.1029/2001JD001298, 2002.

Bhartia, P. K., McPeters, R. D., Flynn, L. E., Taylor, S., Kramarova, N. A., Frith, S., Fisher, B., and DeLand, M.: Solar Backscatter UV (SBUV) total ozone and profile algorithm, Atmos. Meas. Tech., 6, 2533–2548, doi:10.5194/amt-6-2533-2013, 2013.

Borsdorff, T. and Sussmann, R.: On seasonality of stratomeso-spheric CO above midlatitudes: New insight from solar FTIR spectrometry at Zugspitze and Garmisch, Geophys. Res. Lett., 36, L21804, doi:10.1029/2009GL040056, 2009.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goedec, A. P. H.: SCIA-MACHY: Mission Objectives and Measurement Modes, J. Atmos. Sci., 56, 127–150, 1999.

Brion, J., Chakir, A., Daumont, D., Malicet, J., and Parisse, C.: High-resolution laboratory absorption cross section of O_3, Temperature effect, Chem. Phys. Lett., 213, 610–612, doi:10.1016/0009-2614(93)89169-I, 1993.

Buchwitz, M. and Burrows, J. P.: Retrieval of CH_4, CO and CO_2 total column amounts from SCIAMACHY near-infrared nadir spectra: Retrieval algorithm and first results, Remote Sensing of Clouds and the Atmosphere VIII, Proc. SPIE, 5235, 375–388, 2004.

Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstäter-Weißenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K. U., Eisinger, M., and Perner, D.: The global monitoring experiment (GOME): Mission concept and first scientific results, J. Atmos. Sci., 56, 151–175, doi:10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2, 1999.

Butz, A., Guerlet, S., Hasekamp, O., Schepers, A. G., Aben, I., Frankenberg, C., Hartmann, J. M., Tran, H., Kuze, A., Keppel-Aleks, G. T., Wunch, D., Wennberg, P., Deutscher, N., Griffith, D., Macatangay, R., Messerschmidt, J., Notholt, J., and Warneke, T.: Toward accurate CO_2 and CH_4 observations from GOSAT, Geophys. Res. Lett., 38, L14812, doi:10.1029/2011GL047888, 2011.

Callies, J., Corpaccioli, E., Eisinger, M., Hahne, A., and Lefebvre, A.: GOME-2 – Metop's Second Generation Sensor for Operational Ozone Monitoring, ESA Bulletin, 102, ESTEC, Noordwijk, the Netherlands, 2000.

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmos. Chem. Phys., 3, 1285–1291, doi:10.5194/acp-3-1285-2003, 2003.

Eskes, H. J., van der A, R. J., Brinksma, E. J., Veefkind, J. P., de Haan, J. F., and Valks, P. J. M.: Retrieval and validation of ozone columns derived from measurements of SCIA-MACHY on Envisat, Atmos. Chem. Phys. Discuss., 5, 4429–4475, doi:10.5194/acpd-5-4429-2005, 2005.

Frankenberg, C., Platt, U., and Wagner, T.: Retrieval of CO from SCIAMACHY onboard ENVISAT: detection of strongly polluted areas and seasonal patterns in global CO abundances, Atmos. Chem. Phys., 5, 1639–1644, doi:10.5194/acp-5-1639-2005, 2005.

Frankenberg, C., Meirink, J. F., Bergamaschi, P., Goede, A. P. H., Heimann, M., Köner, S., Platt, U., Weele, M. V., and Wagner, T.: Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: analysis of the years 2003 and 2004, J. Geophys. Res., 111, D07303, doi:10.1029/2005JD006235, 2006.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J., Toon, G. C., Butz, A., Jung, M., Kuze, A., and Yokota, T.: New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity, J. Geophys. Res., 38, L17706, doi:10.1029/2011GL048738, 2011.

Gloudemans, A. M. S., Schrijver, H., Hasekamp, O. P., and Aben, I.: Error analysis for CO and CH_4 total column retrievals from SCIAMACHY 2.3 μm spectra, Atmos. Chem. Phys., 8, 3999–4017, 2008,
http://www.atmos-chem-phys.net/8/3999/2008/.

Griesfeller, A., Griesfeller, J., Hase, F., Blumenstock, T., and Naka-jima, H.: Comparison of ILAS-II Data With Ground-based FTIR Measurements of O_3, HNO_3, N_2O, and CH_4 Over Kiruna, Sweden, AGU Fall Meeting Abstracts, p. C950, 2005.

Hansen, C.: Analysis of discrete ill-posed problems by means of the L-curve, Soc. Indust. Appl. Math., 34, 561–580, 1992.

Hansen, C.: The use of the L-curve in the regularization of discrete ill-posed problems, SIAM J. Sci. Comput., 14, 1487–1503, doi:10.1137/0914086, 1993.

Hansen, C.: Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1998.

Hasekamp, O. P. and Landgraf, J.: Ozone profile retrieval from backscattered ultraviolet radiances: The inverse problem solved by regularization, J. Geophys. Res.-Atmos., 106, 8077–8088, doi:10.1029/2000JD900692, 2001.

Lerot, C., van Roozendael, M., Lambert, J. C., Granville, J., van Gent, J., Loyola, D., and Spurr, R.: The GODFIT algorithm: a direct fitting approach to improve the accuracy of total ozone measurements from GOME, Int. J. Remote Sens., 31, 543–550, doi:10.1080/01431160902893576, 2010.

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., and Saari, H.: The ozone monitoring instrument, IEEE T. Geosci. Remote, 44, 1093–1101, doi:10.1109/TGRS.2006.872333, 2006.

Levelt, P. F., Veefkind, J. P., Kerridge, B. J., Siddans, R., de Leeuw, G., Remedios, J., and Coheur, P. F.: Observation Techniques and Mission Concepts for Atmospheric Chemistry (CAMELOT), Report 20533/07NL/HE, European Space Agency, Noordwijk, The Netherlands, 2009.

Marshak, A., Davis, A., Wiscombe, W., and Titov, G.: The Verisimilitude of the Independent Pixel Approximation Used in Cloud Remote Sensing, Remote Sens. Environ., 52, 71–78, 1995.

NOAA: U.S. Standard Atmosphere, 1976, Report NOAA-S/T76-1562, National Oceanic and Atmospheric Administration, US Gov. Printing Office, Washington, DC, 1976.

Notholt, J., Toon, G. C., Rinsland, C. P., Pougatchev, N. S., Jones, N. B., Connor, B. J., Weller, R., Gautrois, M., and Schrems, O.: Latitudinal variations of trace gas concentrations in the free troposphere measured by solar absorption spectroscopy during a ship cruise, J. Geophys. Res., 105, 1337–1350, doi:10.1029/1999JD900940, 2000.

Phillips, B. C.: A technique for the numerical solution of certain integral equations of the first kind, J. Ass. Comput. Mach., 9, 84–97, doi:10.1145/321105.321114, 1962.

Pougatchev, N. S. and Rinsland, C. P.: Spectroscopic study of the seasonal variation of carbon monoxide vertical distribu-

tion above Kitt Peak, J. Geophys. Res., 100, 1409–1416, doi:10.1029/94JD02387, 1995.

Richter, A. and Burrows, J. P.: Tropospheric NO_2 from GOME measurements, Adv. Space Res., 29, 1673–1683, doi:10.1016/S0273-1177(02)00100-X, 2002.

Rinsland, C. P., Jones, N. B., Connor, B. J., Wood, S. W., Goldman, A., Stephen, T. M., Murcray, F. J., Chiou, L. S., Zander, R., and Mahieu, E.: Multiyear infrared solar spectroscopic measurements of HCN, CO, C_2H_6, and C_2H_2 tropospheric columns above Lauder, New Zealand (45° S latitude), J. Geophys. Res., 107, 4185, doi:10.1029/2001JD001150, 2002.

Rodgers, C. D.: Inverse Methods for Atmospheres: Theory and Practice, Vol. 2, World Scientific, Singapore, New Jersey, London, Hong Kong, 2000.

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

Rothman, L. S., Barbe, A., Benner, D. C., Brown, L. R., Camy-Peyret, C., Carleer, M. R., Chance, K., Clerbaux, C., Dana, V., Devi, V. M., Fayt, A., Flaud, J.-M., Gamache, R. R., Goldman, A., Jacquemart, D., Jucks, K. W., Lafferty, W. J., Mandin, J.-Y., Massie, S. T., Nemtchinov, V., Newnham, D. A., Perrin, A., Rinsland, C. P., Schroeder, J., Smith, K. M., Smith, M. A. H., Tang, K., Toth, R. A., Vander Auwera, J., Varanasi, P., and Yoshino, K.: The HITRAN molecular spectroscopic database: edition of 2000 including updates through 2001, J. Quant. Spectrosc. Ra., 82, 5–44, doi:10.1016/S0022-4073(03)00146-8, 2003.

Schepers, D., Guerlet, S., Butz, A., Landgraf, J., Frankenberg, C., Hasekamp, O., Blavier, J. F., Deutscher, N. M., Griffith, D. W. T., Hase, F., Kyro, E., Morino, I., Sherlock, V., Sussmann, R., and Aben, I.: Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms, J. Geophys. Res., 117, D10307, doi:10.1029/2012JD017549, 2012.

Schneider, M., Hase, F., Blumenstock, T., Redondas, A., and Cuevas, E.: Quality assessment of O_3 profiles measured by a state-of-the-art ground-based FTIR observing system, Atmos. Chem. Phys., 8, 5579–5588, doi:10.5194/acp-8-5579-2008, 2008.

Sussmann, R. and Borsdorff, T.: Technical Note: Interference errors in infrared remote sounding of the atmosphere, Atmos. Chem. Phys., 7, 3537–3557, doi:10.5194/acp-7-3537-2007, 2007.

Tikhonov, A.: On the solution of incorrectly stated problems and a method of regularization, Dokl. Acad. Nauk SSSR, 151, 501–504, 1963a.

Tikhonov, A.: On the regularization of ill-posed problems, Dokl. Acad. Nauk SSSR, 153, 49–52, 1963b.

Twomey, S.: On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature, J. Ass. Comput. Mach., 10, 97–101, doi:10.1145/321150.321157, 1963.

Veefkind, J. P., Aben, I., McMullan, K., Forster, H., de Vries, J., Otter, G., Claas, J., Eskes, H. J., de Haan, J. F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P. F.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sens. Environ., 120, 70–83, doi:10.1016/j.rse.2011.09.027, 2012.

Vigouroux, C., De Mazière, M., Errera, Q., Chabrillat, S., Mahieu, E., Duchatelet, P., Wood, S., Smale, D., Mikuteit, S., Blumenstock, T., Hase, F., and Jones, N.: Comparisons between ground-based FTIR and MIPAS N_2O and HNO_3 profiles before and after assimilation in BASCOE, Atmos. Chem. Phys., 7, 377–396, doi:10.5194/acp-7-377-2007, 2007.

von Clarmann, T. and Grabowski, U.: Elimination of hidden a priori information from remotely sensed profile data, Atmos. Chem. Phys., 7, 397–408, doi:10.5194/acp-7-397-2007, 2007.

Washenfelder, R. A., Wennberg, P. O., and Toon, G. C.: Tropospheric methane retrieved from ground based near IR solar absorption spectra, Geophys. Res. Lett., 30, 2226, doi:10.1029/2003GL017969, 2003.

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmos. Meas. Tech., 3, 1351–1362, doi:10.5194/amt-3-1351-2010, 2010.

Yang, Z., Toon, G. C., Margolis, J. S., and Wennberg, P. O.: Atmospheric CO_2 retrieved from ground-based near IR solar spectra, Geophys. Res. Lett., 29, 1339, doi:10.1029/2001GL014537, 2002.

Yurganov, L. N., Blumenstock, T., Grechko, E. I., Hase, F., Hyer, E. J., Kasischke, E. S., Koike, M., Kondo, Y., Kramer, I., Leung, F.-Y., Mahieu, E., Mellqvist, J., Notholt, J., Novelli, P. C., Rinsland, C. P., Scheel, H. E., Schulz, A., Strandberg, A., Sussmann, R., Tanimoto, H., Velazco, V., Zander, R., and Zhao, Y.: A quantitative assessment of the 1998 carbon monoxide emission anomaly in the Northern Hemisphere based on total column and surface concentration measurements, J. Geophys. Res., 109, D15305, doi:10.1029/2004JD004559, 2004.

Yurganov, L. N., Duchatelet, P., Dzhola, A. V., Edwards, D. P., Hase, F., Kramer, I., Mahieu, E., Mellqvist, J., Notholt, J., Novelli, P. C., Rockmann, A., Scheel, H. E., Schneider, M., Schulz, A., Strandberg, A., Sussmann, R., Tanimoto, H., Velazco, V., Drummond, J. R., and Gille, J. C.: Increased Northern Hemispheric carbon monoxide burden in the troposphere in 2002 and 2003 detected from the ground and from space, Atmos. Chem. Phys., 5, 563–573, doi:10.5194/acp-5-563-2005, 2005.