The Retrieval Algorithm of the SCIAMACHY Limb Processor Version 3.0

Adrian Doicu[(1)], **Siegfried Hilgers**[(1)], **Albrecht von Bargen**[(1)] **and Bernd Aberle**[(1)]

German Aerospace Center, Oberpfaffenhofen, 82234 Wessling, Germany, Email: adrian@doicu.dlr.de

1. INTRODUCTION

Several satellite instruments measure radiance emitted from the Earth's limb to monitor vertical profiles of gases. A sequence of observations (limb scans) corresponding to different tangent altitudes are used for atmospheric retrieval. A recently launched instrument that employ this technique is the SCanning Imaging Absorption SpectroMeter for Atmospheric Chartography (SCIAMACHY) aboard the ENVISAT satellite [1]. The SCIAMACHY instrument measures scattered solar radiance and provides informations about the Earth's atmospheric composition with respect to O_3 and other minor constituents such as NO_2 and BrO.

Inverse problems arising in atmosphric remote sensing are nonlinear and ill–posed. Frequently, we are dealing with multi–parameter problems, i.e., several atmospheric profiles are retrived together with a set of auxiliary parameters. Specific constraints have to be imposed on the state vector in order to obtain realistic solutions. This kind of inverse problems can be solved by using a multi–parameter regularization method with simple bounds on the variables.

In this paper we present the main features of an inversion algorithm for atmospheric retrieval. The algorithm incorporates two regularization methods: Tikhonov regularization [2] and the iteratively regularized Gauss–Newton method [3] and has been employed for SCIAMACHY limb scatter measurements [4,5].

2. INVERSION ALGORITHM

The main steps of the retrieval algorithm are: the derivation of the discrete data model, the formulation and the solution of the bound–constrained minimization problem, the selection of the regularization parameter, the derivation of the global regularization matrix for multi–parameter problems and the error analysis depending on the setting in which the problem is treated.

2.1 Discrete data model

The discretization of the radiative transfer equation leads to the data model:

$$\mathbf{y}^\delta = \mathbf{F}\left(\mathbf{x}\right) + \delta, \tag{1}$$

where the mapping $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ represents the forward model, $\mathbf{x} \in \mathbb{R}^n$ is the state vector, $\mathbf{y}^\delta \in \mathbb{R}^m$ is the noisy data vector and $\delta \in \mathbb{R}^m$ is the measurement error. Simple bounds on the variables are imposed in order to obtain a solution with physical meaning, i.e.,

$$\mathbf{l} \le \mathbf{x} \le \mathbf{u}. \tag{2}$$

For a multi–parameter problem, the state vector $\mathbf{x}$ has several components $\mathbf{x}_i$,

$$\mathbf{x} = \left[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_P\right]^T, \tag{3}$$

where $\mathbf{x}_i$ stands for an atmospheric profile or the set of auxiliary parameters and P is the number of components. In our analysis we consider a semi–stochastic data model in the sense that the exact solution $\hat{\mathbf{x}}$ is deterministic but the measurement error δ is stochastic with zero mean and the covariance matrix $\mathbf{C}_\delta = \mathcal{E}\left\{\delta \cdot \delta^T\right\} = \sigma^2 \mathbf{I}_m$, where $\mathcal{E}$ is the expected value operator and $\mathbf{I}_m$ is the identity matrix (of rank m).

2.2 Bound–constrained minimization problem

The ill–posed problem is solved in the least–squares sense by means of regularization methods. The corresponding bound–constrained minimization problem is of the form

$$\left\{ \begin{array}{c} \min_{\mathbf{x}\in\mathbb{R}^n} \;\; \mathcal{F}(\mathbf{x}) = \frac{1}{2}\left\{ \left\| F(\mathbf{x}) - \mathbf{y}^\delta \right\|^2 + \lambda^2 \Lambda\left(\mathbf{x}\right) \right\} \\ \text{subject to the simple bounds } \mathbf{l} \le \mathbf{x} \le \mathbf{u} \end{array} \right. , \tag{4}$$

where λ is the regularization parameter, Λ is the global regularization term and the symbol $\|\cdot\|$ denotes the Euclidian norm. The regularization term adds additional knowledge about the solution to the information comming from the measurement.

Due to the multi–parameter character of the problem, specific constraints have to be imposed on each component $\mathbf{x}_i$. The global regularization term is expressed as a linear combination of regularization terms corresponding to each component, i.e.,

$$\Lambda(\mathbf{x}) = \sum_{i=1}^{P} \omega_i \left\| \mathbf{L}_i \left(\mathbf{x}_i - \mathbf{x}_{\mathtt{a}i} \right) \right\|^2 = \left\| \mathbf{H} \left(\mathbf{x} - \mathbf{x}_{\mathtt{a}} \right) \right\|^2, \tag{5}$$

where $\mathbf{L}_i$ is the i–th regularization matrix corresponding to $\mathbf{x}_i$, $\mathbf{x}_{\mathtt{a}}$ is the a priori state vector, the best beforehand estimator of $\hat{\mathbf{x}}$, and $\mathbf{H}$ is the global regularization matrix

$$\mathbf{H} = \begin{bmatrix} \sqrt{\omega_1}\mathbf{L}_1 & \dots & \mathbf{0} \\ \dots & \dots & \dots \\ \mathbf{0} & \dots & \sqrt{\omega_P}\mathbf{L}_P \end{bmatrix}. \tag{6}$$

The regularization matrices $\mathbf{L}_i$ are chosen in accordance with the pecularities of the solutions $\mathbf{x}_i$. For atmospheric profiles, regularization matrices accounting for the smoothness of the solution (discrete approximations of the first and second derivatives or a priori covariance matrices) are used, while for auxiliary parameters, diagonal matrices controlling the magnitude of the solution are appropriate. The weighting factors ω_i satisfy the normalization condition $\sum_i \omega_i = 1$ and give the contribution of each regularization matrix $\mathbf{L}_i$ to the global regularization matrix $\mathbf{H}$.

The bound–constrained minimization problem can be formulated as a least squares problem and iterative methods for sums of squares can then be employed to solve the minimization problem. A trust–region method with a local active–set strategy to select the step is implemented in our code.

2.3 Regularization parameter selection

At each iteration $\mathbf{x}_{k\lambda}^{\delta}$, the iterative methods consider the linearization of $\mathcal{F}$ arround $\mathbf{x}_{k\lambda}^{\delta}$. The regularized function for the corresponding linear subproblem can be written in standard form as:

$$\mathcal{F}_k^{\mathtt{linear}}(\mathbf{u}) = \left\| \mathbf{K}_k \mathbf{u} - \mathbf{w}_k^{\delta} \right\|^2 + \lambda^2 \left\| \mathbf{H}\mathbf{u} \right\|^2, \tag{7}$$

where $\mathbf{u} = \mathbf{x} - \mathbf{x}_{\mathtt{a}}$, $\mathbf{K}_k = \mathbf{K}\left(\mathbf{x}_{k\lambda}^{\delta}\right)$ is the Jacobian matrix and $\mathbf{w}_k^{\delta} = \mathbf{K}_k\left(\mathbf{x}_{k\lambda}^{\delta} - \mathbf{x}_{\mathtt{a}}\right) - \mathbf{r}_{k\lambda}^{\delta}$. $\mathbf{r}_{k\lambda}^{\delta}$ is the residual vector at the iteration step k and is given by $\mathbf{r}_{k\lambda}^{\delta} = \mathbf{F}\left(\mathbf{x}_{k\lambda}^{\delta}\right) - \mathbf{y}^{\delta}$. The parameter–choice methods which are implemented in our code are: the L–curve method, the generalized cross-validation method, the unbiased predictive risk estimator method, the minimum bound method and the noise error criterium.

The L–curve method consists in the analysis of the graph of the curve obtained by plotting the norm of the constraint vector versus the norm of the residual in log–log scale. This curve exhibits a typical "L" shape, and the optimal value of the regularization parameter is at the corner of the "L". The solution corresponding to the corner of the L–curve balances the constraint and residual norms and also the smoothing and noise errors. The corner of the L–curve is defined as the point on the graph with maximal curvature, and since the residual and constraint norms can be computed analytically, function minimization routines can be used to calculate the maximum of the curvature.

The noise error is a quantity which can be computed accurately and a simple parameter–choice method can be formulated. The idea is simply to choose the regularization parameter λ such that the noise error is equal to an a priori upper bound, i.e.,

$$\sqrt{\mathcal{E}\left\{ \left\| \epsilon_{k\lambda,\mathtt{noise}}^{\mathtt{linear},\delta} \right\|^2 \right\}} = \Delta \left\| \mathbf{x}_{k\lambda}^{\delta} \right\|, \tag{8}$$

where typically $\Delta = 0.05, ..., 0.1$.

For nearly linear problems, the regularization parameter computed at the first iteration can be used for the subsequent iterations. For strongly nonlinear problems we have two options: the regularization parameter can be keep constant or it can be variable during the iterative process. In the first case, the regularization method is the Tikhonov regularization [2], while in the second case, the regularization method is the iteratively regularized Gauss–Newton method [3]. The iteratively regularized Gauss–Newton method can be regarded as a Tikhonov regularization with a variable regularization parameter, i.e., (λ_k) is a monotonically decreasing sequence. The following recursive procedure is used to construct the sequence of iterates [4]:

$$\lambda_k = \beta_k \lambda_{\mathtt{min}} + (1 - \beta_k)\, \lambda_{k-1}, \tag{9}$$

where $\beta_k = \left\| \mathbf{r}_{k\lambda}^{\delta} \right\| / \left\| \mathbf{r}_{k-1,\lambda}^{\delta} \right\|$. The upper and lower bounds of the sequence of regularization parameters $\lambda_{\mathtt{max}}$ and $\lambda_{\mathtt{min}}$ have to be chosen in advanced and the decay rate is given by the ratio of the residuals at two consecutive steps. For

$\beta_k \approx 0$ (large reduction of the residual), the new value of the regularization parameter is close to the previous value, while for $\beta_k \approx 1$ (small reduction of the residual), the regularization parameter is considerably decreased towards $\lambda_{\min}$. This parameter–choice method allows enough regularization to be applied at the beginning of iterations and then to be gradually decreased.

2.4 Global regularization matrix

The weighting factors ω_i are computed by using an a priori parameter choice rule which consists in the minimization of the error $\left\|\mathbf{x}_{\lambda\omega}^{\delta} - \hat{\mathbf{x}}\right\|^2$ with respect to λ and $\omega = [\omega_1, \omega_2, ..., \omega_P]$ for a set of true solutions $\hat{\mathbf{x}}$, i.e., $\left(\lambda_{\mathtt{opt}}, \omega_{\mathtt{opt}}\right) = \arg\min_{\alpha,\omega} \left\|\mathbf{x}_{\lambda\omega}^{\delta} - \hat{\mathbf{x}}\right\|^2$. Solving the minimization problem for different noise variances yields the a priori selection criterion of the form $\lambda_{\mathtt{opt}} = \lambda_{\mathtt{opt}}\left(\sigma^2\right)$ and $\omega_{\mathtt{opt}} = \omega_{\mathtt{opt}}\left(\sigma^2\right)$.

As an alternative, we may consider a set of P one–parameter regularized functions and compute the weighting factors is

$$\omega_i = \frac{\left(\lambda_{\mathtt{opt}i}\right)^2}{\sum_{j=1}^{P}\left(\lambda_{\mathtt{opt}j}\right)^2},\ i = 1, 2, ..., P, \tag{10}$$

where $\lambda_{\mathtt{opt}i}$ is the optimal regularization parameters for the i one–parameter regularized function. In practice, $\lambda_{\mathtt{opt}i}$ can be computed by using the generalized cross–validation method or the L–curve method.

2.5 Error analysis

The accuracy of a retrieval method can be characterized by the discrepancy between the approximate solution $\mathbf{x}_{\lambda}^{\delta}$ and the exact solution $\hat{\mathbf{x}}$. The approximate solution $\mathbf{x}_{\lambda}^{\delta}$ is defined as the limit of the sequence of iterates whether the regularization parameter is keep constant or varies during the iterative process. Considering the linearization arround $\mathbf{x}_{\lambda}^{\delta}$ and assuming that the linearity relation holds at $\mathbf{x} = \hat{\mathbf{x}}$ yields

$$\epsilon_{\lambda,\mathtt{total}}^{\delta} = \epsilon_{\lambda,\mathtt{smooth}} + \epsilon_{\lambda,\mathtt{noise}}^{\delta}, \tag{11}$$

where $\epsilon_{\lambda,\mathtt{total}}^{\delta} = \mathbf{x}_{\lambda}^{\delta} - \hat{\mathbf{x}}$ is the total error, $\epsilon_{\lambda,\mathtt{smooth}} = \left(\mathbf{A}_{\lambda} - \mathbf{I}_n\right)\left(\hat{\mathbf{x}} - \mathbf{x}_{\mathtt{a}}\right)$ is the smoothing error and $\epsilon_{\lambda,\mathtt{noise}}^{\delta} = \mathbf{K}_{\lambda}^{\dagger}\delta$ is the noise error. Since the smoothing error is a deterministic quantity, $\mathbf{x}_{\lambda}^{\delta}$ is always a biased estimator of $\hat{\mathbf{x}}$. An estimator for the smoothing error vector is $\epsilon_{\lambda,\mathtt{smooth}}^{\delta} = \mathbf{K}_{\lambda}^{\dagger}\mathbf{r}_{\lambda}^{\delta}$, while the covariance matrix of the noise error is $\mathbf{C}\left(\epsilon_{\lambda,\mathtt{noise}}^{\delta}\right) = \sigma^2\mathbf{K}_{\lambda}^{\dagger}\mathbf{K}_{\lambda}^{\dagger T}$. For data with significant noise level, $\epsilon_{\lambda,\mathtt{smooth}}^{\delta} = \left(\mathbf{A}_{\lambda} - \mathbf{I}_n\right)\left(\mathbf{x}_{\lambda}^{\delta} - \mathbf{x}_{\mathtt{a}}\right)$ can be used as an alternative estimator of the smoothing error vector.

3. NUMERICAL SIMULATIONS

In this section we consider a numerical experiment concerning the retrieval of O_3 profile from SCIAMACHY limb scatter measurements. The model does not include the boundary condition at the Earth's surface and the rotational Raman scattering. The Ring effect is taken into account by adding a Ring reference spectrum to the simulated radiance. Under these circumstances, the surface albedo and the Ring parameters (multiplicative factors of the Ring reference spectra) are auxiliary parameters which have to be retrieved together with the atmospheric profile O_3.

For our simulations, an altitude retrieval grid with 20 grid points between 14 and 100 km is considered. The number of limb scans is 14 and the corresponding tangent altitudes vary between 14 and 60 km. The spectral domain of analysis is a window of 323.4–333.4 nm in Channel 2 of the SCIAMACHY instrument. The a priori and initial gas profiles were assumed to be identical and were chosen from the U.S. standard atmosphere. The initial value of the surface albedo is 0.3, while the initial ring parameters are chosen such that the average ring spectrum (over each scan) is 15% from the simulated radiance. The deviation of the atmospheric profile with respect to the a priori profile is assumed to be smaller than 30%, the surface albedo is considered to be bounded by the reference values 0.2 and 0.8, and we impose that the ring spectrum will not exceed 30% from the simulated radiance.

The regularization matrix for ozone is the Choleski factor of an a priori covariance matrix, while the regularization matrix for the auxiliary parameters is a diagonal matrix. The multi–parameter regularization algorithm gives $\omega_{\mathrm{O}_3} = 0.98$ and $\omega_{\mathtt{auxpar}} = 0.02$.

For the numerical simulations we choose the L–curve method and the noise error criterium as parameter–choice methods and consider three schemes for regularization parameter selection. In the first scheme, the regularization parameter is constant during the iterative process and is given by $\lambda = \lambda_{\mathtt{LCR},0} = 8.04 \cdot 10^{-1}$. In the second and third schemes, the regularization parameter is variable and is computed accordingly to Eq. (9). We choose, $\lambda_{\mathtt{max}} = \lambda_{\mathtt{LCR},0}$. For the second scheme we set $\lambda_{\mathtt{min}} = \lambda_{\mathtt{noise},0}(0.05) = 5.52 \cdot 10^{-2}$, while for the third scheme we choose $\lambda_{\mathtt{min}} = \lambda_{\mathtt{noise},0}(0.1) = 6.02 \cdot 10^{-6}$.

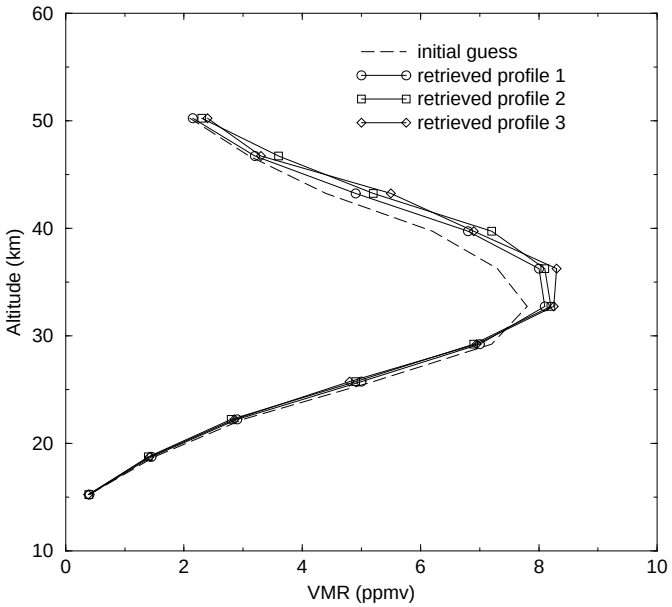

Figure 1: Retrieved profiles for the regularization methods with constant regularization parameter, $\lambda = \lambda_{\mathtt{LCR},0}$ (curve 1), variable regularization parameter, $\lambda_{\mathtt{min}} = \lambda_{\mathtt{noise},0}(0.05)$ (curve 2) and variable regularization parameter, $\lambda_{\mathtt{min}} = \lambda_{\mathtt{noise},0}(0.1)$ (curve 3).

The retrieved profiles are plotted in Figure 1. Clearly, the solution given by the first scheme is more close to the a priori than the other two solutions. The best trade–off between the smoothing and noise errors appears to be given by the second scheme. In the first case, the solution is oversmoothed and is dominated by smoothing errors, while in the third case, the solution is underregularized and is dominated by noise errors. It should be observed that the solution obtained by the third scheme is smooth and no oscillations are visible. This peculiarity of the solution is tipically for the iteratively regularized Gauss–Newton method. We avoid oscillating solutions by decreasing the regularization parameter in a safe way.

4. CONCLUSIONS

An inversion algorithm for atmospheric remote sensing has been presented. The algorithm combines the Tikhonov regularization and the iteratively regularized Gauss–Newton method into a common shell and is devoted to the solution of multi–parameter and bound–constrained inversion problems. The bound constrained minimization problem is solved by using a Gauss–Newton method with an active–set algorithm and a trust–region method with a local active–set strategy to select the step. Several parameter–choice methods as for instance the L–curve method and the noise error criterium are implemented in the code. These methods guarantee an optimal choice of the regularization parameter. The error analysis is performed in a semi–stochastic setting and consists in the computation of the smoothing and noise errors at the (approximate) solution. All these features lead to an efficient and robust inversion algorithm for atmospheric remote sensing.

References

1. Bovensmann H., Burrows J. P., Buchwitz M., Frerick J., Noel S., Rozanov V.V., Chance K. V. and Goede A. H. P., SCIAMACHY–Mission objectives and measurements models, J. Atmos. Sci. 56, 127–150, 1999.

2. Engl H. W., Hanke M. and Neubauer A., *Regularization of Inverse Problems*, Dordrecht, Kluver, 1996.

3. Bakushinskii A. B., The problem of the convergence of the iteratively regularized Gauss–Newton method, Comput. Math. Math. Phys. 32, 1353–1359, 1992.

4. Doicu A., Schreier F. and Hess M., Iteratively regularized Gauss–Newton method for atmospheric remote sensing,

Comput. Phys. Commun. 148, 214–226, 2002

5. Doicu A., Schreier F. and Hess M., Iteratively regularized Gauss–Newton method for bound–constraint problems in atmosphere remote sensing, Comput. Phys. Commun. 153, 59–65, 2003.