

Discretization Independent Retrieval of Atmospheric Ozone Profile

Barbara D'Ambrogi¹, Sari Mäenpää¹ and Markku Markkanen²

¹ University of Oulu, Department of Mathematical Sciences, FIN-90401 Oulu, Finland

² Sodankylä Geophysical Observatory, FIN-99600 Sodankylä, Finland

(Received: November 1998; Accepted: March 1999)

Abstract

A simulation based study on the remote sensing problem of retrieving the atmospheric ozone profile is performed so that the effects of the discretization error can be analyzed.

The retrieved profile is recovered by using the stochastic inverse theory in a finite dimensional case. Provided the a priori information is given, the method gives as a solution a probability density function having the most probable ozone profile as mean value.

The goal of this paper is to show that it is possible to choose the a priori probability distribution so that the retrieved profile does not depend on the density of the discretization.

Numerical simulations show that the estimated ozone profile converges pointwisely when the discretization is densified i.e. the retrieved profile remains almost constant when increasing the number of the points in the discretization set.

Key words: Microwave sounding of ozone, convergence results, stochastic processes, a priori information

1. Introduction

During the past decade many ground based microwave radiometres have been constructed to monitor the atmospheric ozone profile. These radiometres measure the atmospheric brightness temperature (ABT) from which the ozone profile is retrieved. There already exist many papers about the retrieval theory and many different retrieval methods have been applied to these kind of problems (*Rodgers, 1976* and *Twomey, 1965*).

It is known (*Randegger, 1980*) that the ABT equation leads to a Fredholm integral equation of the first kind for the ozone profile. Moreover it is known that the discretized inverse problem is ill-posed i.e. the solution is not stable with respect to the measurement error.

There exist many methods for solving ill-posed inverse problems, in this paper we use a non iterative method, the statistical inverse method. This method has been already applied to this problem with good results by *Rodgers (Rodgers, 1976)*. The aim of this paper is to show that with this method it is possible to get a solution that does not

depend on the density of the discretization. We claim that by choosing the *a priori* distribution in the appropriate way we do not need to densify the discretization in order to obtain a more precise solution. This is of course important in terms of computing times.

The use of the *a priori* information is required since the ill-posedness of the problem. The *a priori* information is usually based on various quantitative and qualitative features of the true profile. *A priori* statements are made for values of the unknown profile at a specified set of points called a discretization set. This discretization set is relevant since the profile is estimated only at these points.

Since *a priori* information comes from observations of the true profile, it could be stated for the whole profile rather than a part of it like values at some specified discretization set. Moreover, the meaning of the *a priori* assumptions should become clearer when stated for the whole profile. We formulate the *a priori* information in such a way that it applies to any equally spaced discretization set of the unknown profile. Hence, with one measurement vector we are able to obtain solutions with various discretizations that can be compared.

In the beginning of this paper, the direct theory of the atmospheric radiation transfer is presented for a better understanding of those who are not so familiar with this problem. A brief explanation of the inversion method is also presented and a particular attention is obviously given to the construction of the *a priori* covariance matrix. Numerical simulations are presented to show that the solution does not depend on the density of the discretization. The claim is also supported by the comparison of the variances obtained by using different amounts of point in the discretization set.

2. Direct theory

In this section we describe the integral equation relating the measurements and the density of ozone. We assume that a microwave radiometer measures the intensity of downwelling radiations in the direction of the zenith. Moreover, this radiometer operates at the center frequency of 110.836 GHz with a 1200 MHz bandwidth and a 20 MHz resolution in addition to a 50 MHz bandwidth and a 85 kHz resolution.

At the frequency of 110.836 GHz there is a rotational transition in the spectrum of an ozone molecule. The spectrum broadens as a function of the pressure in the atmosphere due to collisions between molecules. Below the altitude of 75 km this broadening can be modelled by the Van Vleck-Weisskopf spectral line shape function (*Van Vleck and Weisskopf, 1946*)

$$F(\nu, z) = \frac{1}{\pi} \frac{\nu}{\nu_0} \left(\frac{\mathcal{L}(z)}{(\nu - \nu_0)^2 + \gamma_L^2(z)} + \frac{\mathcal{L}(z)}{(\nu + \nu_0)^2 + \gamma_L^2(z)} \right), \quad (1)$$

where z is the altitude, ν is the frequency, ν_0 is the center frequency 110.836 GHz and the Lorentz half-width parameter in 1/cm is

$$\gamma_{\mathcal{L}}(z) = \gamma_{air}(p(z)) \left(\frac{T_s}{T(z)} \right)^n, \quad (2)$$

where p is the pressure, T is the temperature and $T_s=300$ K.

The line shape function multiplied by the transition intensity $S(z)$ defines the length absorption coefficient (*Kroto*, 1975 and *Rosenkranz*, 1993)

$$K(\nu, z) = F(\nu, z) S(\nu, z). \quad (3)$$

Here the line strength function, calculated by using a reference value $S_{T_{ref}}$, is

$$S(\nu, z) = S_{T_{ref}} \frac{Q_r(T_{ref}) Q_v(T_{ref})}{Q_r(T(z)) Q_v(T(z))} \left(\frac{e^{-E_l/kT(z)}}{e^{-E_l/kT_{ref}}} \right) \frac{1 - e^{-h\nu/kT(z)}}{1 - e^{-h\nu/kT_{ref}}}, \quad (4)$$

where k is the Boltzmann constant, $T_{ref}=300$ K is the reference temperature,

$$Q_r(z) = \left(\frac{5.34 \cdot 10^6}{\Sigma} \right) \sqrt{\frac{T(z)^3}{ABC}} \quad (5)$$

is the rotational partition function and

$$Q_v(z) = \frac{1}{(1 - e^{-h\omega_1/kT(z)})(1 - e^{-h\omega_2/kT(z)})(1 - e^{-h\omega_3/kT(z)})} \quad (6)$$

is the vibrational partition function, where h is Planck's constant. The spectrometric parameters are given in Table 1.

Table 1. Molecular quantities of $^{16}O_3$.

Molecular parameter	Value	Reference
Figure of symmetry Σ	2	(Gora, 1959)
Rotational constant A	106 536.24 MHz	(Depannemaecker et al., 1977)
Rotational constant B	13 349.26 MHz	(Depannemaecker et al., 1977)
Rotational constant C	11 834.36 MHz	(Depannemaecker et al., 1977)
1. vibration frequency ω_1	716 cm^{-1}	(Rachavachari et al., 1989)
2. vibration frequency ω_2	1.1089 cm^{-1}	(Rachavachari et al., 1989)
3. vibration frequency ω_3	1 135 cm^{-1}	(Rachavachari et al., 1989)
Lower energy level E_l	17.5973 cm^{-1}	(Rothman, 1992)
Line intensity (300 K) $S_{T_{ref}}$	1.188E-23 $\text{cm}^{-1}/(\text{molec cm}^{-2})$	(Rothman et al. 1998)
Air-broadened halfwidth γ_{air}	0.0812 $\text{cm}^{-1}/\text{atm}$ at 300 K	(Rothman et al. 1998)
Coeff. of temp. dependence n	0.76	(Rothman et al. 1998)

The intensity of radiation as seen on the ground obeys the radiative transfer equation (*Randegger, 1980*)

$$T_B(\nu) = \left(\frac{h\nu}{k}\right) \exp[-\beta(\nu)] \int_0^H \frac{\exp[-\tau(0,z)]}{(1 - \exp[h\nu/kT(z)])} K(\nu,z) \rho(z) dz + T_{atm} \quad (7)$$

where the intensity T_B is expressed by the Rayleigh-Jeans law in terms of the equivalent black body brightness temperature. In equation (7) ρ is the density of ozone as a function of the altitude z , H is the maximum altitude, $\exp[-\tau(0,z)]$ corresponds to the ozone absorption below z , $\exp[-\beta(\nu)]$ represents the attenuation in troposphere due to water vapour and T_{atm} is the intensity of radiation emitted by sources other than ozone. The local thermodynamic equilibrium approximation is assumed to be valid.

In the term $\exp[-\tau(0,z)]$ the unknown ozone density is replaced with a standard ozone profile from AFGL atmospheric models (*Anderson et al., 1986*). The linearization is plausible since the term differs only slightly from unity (*Brillet, 1989*).

After estimating the tropospheric term $\exp[-\beta(\nu)]$ and the linear term T_{atm} (*Brillet, 1989*) the equation for the brightness temperature reduces to

$$T_B^c(\nu) = \int_0^H A(\nu,z) \rho(z) dz, \quad (8)$$

where the integral kernel is

$$A(\nu,z) = \left(\frac{h\nu}{k}\right) \frac{\exp[-\tau(0,z)]}{(1 - \exp[h\nu/kT(z)])} K(\nu,z). \quad (9)$$

In equation (8) the brightness temperature T_B is corrected for tropospheric attenuation and emission due to other sources.

3. Discretization independent retrieval

By discretizing the integral equation (8) we obtain a matrix equation $Y = AX + \varepsilon_d$ that relates the M -dimensional measurement vector Y with the N -dimensional ozone density vector X and discretization error ε_d . When the experimental noise ε_m is included, we obtain a matrix equation $Y_m = AX + \varepsilon_m + \varepsilon_d$.

The number N is taken so large that the discretization error ε_d is much smaller than the measurement error. Therefore, densifying the discretization set would not have any noticeable effect on the overall noise $\varepsilon = \varepsilon_d + \varepsilon_m$.

We estimate the density of the ozone X from the measured values of Y_m by applying the stochastic inverse theory (*Lehtinen, 1988*). It states that all available characteristics of the vector X on the basis of one measured value y of Y_m and a *a priori* probability density function (p.d.f) $D_X^{pr}(x)$ are contained in a *a posteriori* p.d.f.

$$D_X^{post}(x) = \frac{D_{Y_m|x}(y|x)D_X^{pr}(x)}{D_{Y_m}(y)}. \quad (10)$$

From the *a posteriori* p.d.f. suitable estimates and errorbounds for X can be extracted.

When the noise ε is gaussian with zero mean and a covariance matrix Σ_ε and the *a priori* p.d.f. is gaussian with zero mean and a covariance matrix Σ_{pr} , then the *a posteriori* p.d.f. is gaussian with a covariance matrix

$$\Sigma_{post} = \left(A^T \Sigma_\varepsilon^{-1} A + \Sigma_{pr}^{-1} \right)^{-1} \quad (11)$$

and mean

$$x = \Sigma_{post} A^T \Sigma_\varepsilon^{-1} y. \quad (12)$$

This method is already used in atmospheric retrieval problems (Rogers, 1976). We will apply the method with a special choice of the *a priori* p.d.f.

The method is stable with respect to the noise. Considering this, it would seem unlikely that the estimated profile should greatly change when the discretization is densified. On the other hand, the size of the unknown vector X would grow and bring more points to the estimated profile. Our conjecture (which we investigate later numerically) is that with right *a priori* information the estimated profile converges in common points when the discretization set is densified. Moreover, we claim that even the *a posteriori* p.d.f. restricted to the common points do not change when the discretization is densified. This is later numerically verified by studying both covariance matrices and mean values.

The crucial point is the choice of *a priori*. The *a priori* should be extendable to the whole atmosphere and it should hold necessary and trueful characteristics of the ozone density.

3.1 Construction of the *a priori* p.d.f.

From the long term behaviour of ozone density (Anderson et al., 1986) we deduce that our solution has to satisfy next conditions:

The value of the ozone density on the ground is near zero.

Between 0 and $t_0=40$ km the solution is smooth.

Above t_0 the smoothness grows.

The solution is zero at the altitude of $T=120$ km.

The solution decreases exponentially above t_0 .

Denote with $Z^{(n)} = (Z_0^{(n)}, \dots, Z_n^{(n)})$ the *a priori* random vector which corresponds to the *a priori* ozone density on a given discretization set ($0 = s_0, \dots, s_n = 120$ km). We denote the separation $s_j - s_{j-1}$ with h_n .

Condition 1. fulfills when $Z_0^{(n)}$ is a Gaussian random variable with zero mean and variance 1 (when the ozone density is given in 10^{18} molec/ m^3).

For Condition 2. we form an n -dimensional Gaussian random vector $\eta^{(n)}$ with zero mean and covariance matrix $a^2 \cdot h_n^{-1} \cdot \delta_{ij}$. The idea is that $\eta^{(n)}$ has very irregular dependencies between its components. The irregularity is scaled by the factor a which will be later related to the classical regularization parameter.

The irregular dependencies are smoothed with integrations of various orders. For Condition 2. we integrate once. Hence we obtain a random vector $X^{(n)}$ which is a discrete integral of the random vector $\eta^{(n)}$

$$X_j^{(n)} = \sum_{i=1}^j h_n \eta_i^{(n)}.$$

This vector has as a covariance matrix

$$C_X^{(n)}_{ij} = E[X_i^{(n)} X_j^{(n)}] = h_n \cdot a^2 \cdot \min(i, j).$$

The random vector $X^{(n)}$ has $X_0^{(n)} = 0$ contradicting Condition 1. We overcome the difficulty by choosing such l that $a^2 \cdot h_n \cdot l \approx 1$ and repeating above calculations with an $n + l$ dimensional vector η . Then the vector $U_j^{(n)} = X_{l+j}^{(n)}$ has as its first component a Gaussian random variable with zero mean and variance approximately 1. The vector $U_j^{(n)}$, where $jh_n \leq 40$ km, represents the ozone profile from ground to the altitude of 40 km.

For Condition 3. we integrate twice an n -dimensional vector ξ independent of η with zero mean and covariance matrix $b^2 \cdot h_n^{-1} \cdot \delta_{ij}$. We obtain a random vector $Y^{(n)}$ which is

$$Y_j^{(n)} = h_n \sum_{k=1}^j h_n \sum_{i=1}^k \xi_i = h_n^2 \sum_{i=1}^j (j-i+1) \xi_i.$$

This vector has as a covariance matrix

$$C_Y^{(n)}_{ij} = b^2 \cdot h_n^3 \sum_{k=1}^{\min(i, j)} (1+(j-k))(1+(i-k)).$$

To fulfill Condition 4. we extract a suitable term

$$V_j^{(n)} = Y_j^{(n)} - \frac{j}{n_0} Y_{n_0}^{(n)},$$

where $h_n \cdot n_0 = T - t_0$. For Condition 5. we modulate the process $Y^{(n)}$ by multiplying the discrete noise $\xi^{(n)}$ with the function f

$$L_j^{(n)} = h_n^2 \sum_{i=1}^j (j-i+1) f(t_i) \xi_i.$$

where $f(t) = \exp(-t/s)$. The parameter s is chosen appropriately. This has covariance matrix

$$C_L^{(n)}{}_{ij} = b^2 \cdot h_n^3 \sum_{k=1}^{\min(i,j)} (1+j-k)(1+i-k) f(t_k)^2.$$

Now we define the modulated random vector

$$K_j^{(n)} = L_j^{(n)} - \frac{j}{n_0} L_{n_0}^{(n)},$$

which has as a covariance matrix

$$C_K^{(n)}{}_{ij} = C_L^{(n)}{}_{ij} - \frac{j}{n_0} C_L^{(n)}{}_{in_0} - \frac{i}{n_0} C_L^{(n)}{}_{n_0j} + \frac{i \cdot j}{n_0^2} C_L^{(n)}{}_{n_0n_0}. \quad (13)$$

Now we can form *a priori* random vector

$$Z_j^{(n)} = \begin{cases} U_j^{(n)}, & j < k \\ \frac{n-j}{n-k} U_k^{(n)} + K_{j-k}, & j \geq k, \end{cases}$$

where $h_n \cdot k = t_0$. This has covariance matrix

$$C_Z^{(n)}{}_{ij} = \begin{cases} a^2 \cdot h_n \cdot \min(i, j), & \text{when } i, j \leq k \\ a^2 \cdot \frac{n-j}{n-k} h_n \cdot i, & \text{when } j > k, i \leq k \\ a^2 \cdot \frac{n-i}{n-k} h_n \cdot j, & \text{when } j \leq k, i > k \\ a^2 \cdot \frac{(n-i)(n-j)}{(n-k)^2} \cdot h_n \cdot \min(i, j) + E[K_{j-k} K_{i-k}], & \text{when } i, j > k \end{cases}$$

Let us briefly explain what this *a priori* means for the whole ozone density. By studying covariance matrices $C_Z^{(n)}$ of vectors $Z^{(n)}$ one can show that, when n grows, the components of $Z^{(n)}$ at fixed grid points converge to random variables which are the values of the stochastic process

$$Z_t = \begin{cases} a \cdot B_{t+1}, & t < t_0 \\ \frac{T-t}{T-t_0} a \cdot B_{t_0+1} + K_{t-t_0}, & t \geq t_0, \end{cases} \quad (14)$$

$$\|A\bar{x} - y\|^2 + \alpha \|B\bar{x}\|^2 = \inf_x (\|Ax - y\|^2 + \alpha \|Bx\|^2)$$

Hence our *a priori* gives the Twomey-Tikhonov method in this special case and even an interpretation for the determination of the regularization parameter.

We have two undetermined scaling factors a and b that, once chosen, apply to any discretization set. For example, when studying values below t_0 , we notice that in the Twomey-Tikhonov method the regularization parameter is $\frac{\alpha}{a^2 h_n}$, where α is the measurement error, h_n is the discretization step and a is the parameter influencing the magnitude of the *a priori* irregularity.

In the following, we demonstrate numerically that *a priori* covariance matrix $C_Z^{(n)}$ has a correct dependence from h_n enabling the discretization independent retrieval.

4. Numerical results

In the following, we present results from four simulations based on the theory presented above. As simulated measurements we calculate the ABT by using the AFGL sub-artic ozone profile in the summer (*Anderson et al.*, 1986). In the simulation we add a random noise that has standard deviation equal to 2 percent of the maximum of the profile.

Each of these four simulations is performed using a grid doubly as dense as the previous one. In Figure 1 it is shown that already with the first three different grids we recover the same solution.

Because the solution vectors have different sizes, the comparison of the results is done at the common points of the three grids. In Table 2 the differences of each two successive solutions is shown. To calculate the comparison scalar we use the following:

$$\sum_{j=1}^{2^{(i-1)} \cdot 46 + 1} \frac{x_j^{i+1} - x_j^i}{2^{(i-1)} \cdot 46 + 1} \quad (17)$$

where x_j^i is the I : th retrieval profile calculated in the points that are in common with x_j^{i-1} . This shows that the solution does not change almost at all when the grid is made denser.

Table 2. Differences between different solutions divided by the number of common points.

Between 47 points and 93 points	1.070E-04
Between 93 points and 185 points	1.090e-05
Between 185 points and 369 points	2.412e-06

Fig. 1. Inverse profile of the ozone in sub-arctic summer with three different grids.

In Figure 2 it is shown that also the differences between the standard deviations of each two successive grids depict this feature. Moreover we have compared the covariance matrices of successive grids; also in that case the difference between any two successive grids become smaller as the grid is made denser. Simulations made with much denser grids show that the solution is convergent but that the profile is already so near to the limit profile with the used grid that it does not justify the use of any denser grid.

5. *Conclusions*

Inverse problems related to natural phenomena, like the one considered in this paper, are usually ill-posed. Stable solutions are recovered by applying some kind of regularization. Simulations show that taking the regularization errors in Twomey-Tikhonov regularization as functions of the distance between neighboring points (h_n) leads to a solution that does not depend on the density of the discretization set. This is a significant result when we think in terms of computing times and becomes really important when the model is described by large matrices.

Fig. 2. Differences between the standard deviations of each two successive grids.

The method we used is called stochastic inversion and it gives the most probable ozone profile with error estimates. The *a priori* information is chosen so that also the error estimates remain constant when the discretization changes. We have numerically verified that with the right *a priori* information the mean value and the covariance matrix of the *a posteriori* p.d.f. converge when the discretization is densified. The discretization has clearly to be at least dense enough to give a solution that is near to the limit solution. On the other hand, our simulations have shown that one has no reason to use a denser grid since the profile remains approximatively the same.

We used the stochastic inverse theory in a finite dimensional case to obtain the estimated solutions but made our *a priori* assumptions extendable to a continuous stochastic process. On the contrary, the commonly used *a priori* p.d.f. with diagonal covariance matrix does not extend well to the whole atmosphere. The extended version should have only uncorrelated points, a behavior similar to white noise which is not exactly a stochastic process but a generalized stochastic process.

The true ozone density is intended to fall in the class consisting of sample paths of the stochastic process. If the ozone density should behave differently, the *a priori* stochastic process should be accordingly modified. This concerns mainly Condition 5. of the exponential decay.

When we used only modified Brownian motion as *a priori* we ended up with an oscillating tail. We concluded that the Brownian motion does not restrict the estimated solution enough, hence arriving at our more constrained *a priori*.

Finally, the results are a bit more general than what is presented here. Based on our preliminary analysis there seems to be no objection in using the same technique in two dimensional case.

References

- Anderson, G.P., S.A. Clough, F.X. Kneizys, J.H. Chetwynd and E.P. Shettle, 1986. AFGL Atmospheric Constituent Profiles (0-120 km). Air Force Geophysics Laboratory, Hanscom AFB, MA 01731, USA.
- Brillet, J., 1989. A theoretical study of ozone measurements made with ground-based microwave sensors. *J. Geophys. Res.* **94**, 12 833-12 850.
- Depannemaecker, M.J.C., B. Duterage and M.J. Bellet, 1977. Systematic calculations of rotational spectra of normal and substituted (^{18}O in place of ^{16}O) ozone molecules. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **17**, 519-530.
- Gora, E.K., 1959. The rotational spectrum of ozone. *J. Molec. Spectroscopy*, **3**, 78-99.
- Kroto, H.W., 1975. Molecular rotation spectra. John Wiley & Sons, London - New York - Sidney - Toronto.
- Lehtinen, M.S., 1988. On Statistical Inversion Theory. *Theory and applications of inverse problems*. In: Proc. Symp., Helsinki/Finland, Pitman Res. Notes Math. Ser. 167, 46-57.
- Raghavachari, K., G.W. Trucks, J.A. Pople and E. Replogle, 1989. Highly correlated systems: Structure, binding energy and harmonic vibrational frequencies of ozone. *Chemical Physics Letters*, **158**, 4835-4856.
- Randegger, A.K., 1980. On the determination of the atmospheric ozone profile for ground based microwave measurements. *Pure Appl. Geophys.*, **118**, 1052-1065.
- Rodgers, C.D., 1976. Retrieval of Atmospheric Temperature and Composition From Remote Measurements of Thermal Radiation. *Reviews of Geophysics and Space Physics*, **14**, 609-624.
- Rosenkranz, P.W., 1993. Absorption of microwaves by atmospheric gases. In: M.A. Janssen (Ed.), Atmospheric remote sensing by microwave radiometry.
- Rothman, L.S., C.P. Rinsland, A. Goldman, S.T. Massie, D.P. Edwards, J.-M. Flaud, A. Perin, C. Camy-Peyret, V. Dana, J.-Y. Mandin, J. Schroeder, A. McCann, R.R. Ganache, R.B. Wattson, K. Yoshino, K. Chance, K. Jucks, L.R. Brown, V. Nemtchinov and P. Varanasi, 1998. The 1996 HITRAN molecular spectroscopic database and HAWKS (HITRAN atmospheric workstation). *Journal of Quantitative Spectroscopy and Radiative Transfer*, **60**, 665-710.

- Twomey, S., 1975. Comparison of constrained linear inversion and an iterative nonlinear algorithm applied to the indirect estimation of particle size distributions. *Journal of Computational Physics*, **18**, 188-200.
- Twomey, S., 1965. The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements. *Journal of the Franklin institute*, **279**, 95-109.
- Van Vleck, J.H. and V. F. Weisskopf, 1946. On the shape of Collision-Broadened Lines. *Reviews of Modern Physics*, **17**, 227-236.