Radiophysics and Quantum Electronics, Vol. 46, Nos. 8–9, 2003

BAYESIAN APPROACH TO RETRIEVING A VERTICAL OZONE PROFILE FROM RADIOMETRIC MEASUREMENT DATA

Ya. I. Molkov, D. N. Mukhin, E. V. Suvorov, and A. M. Feigin

UDC 551.510.534

We propose a technique for retrieving a vertical profile of the atmospheric ozone number density from ground-based radiometry data. The technique is based on the Bayesian approach to solving inverse problems and permits one, by allowing for measurement noise and using certain a priori *information on the retrieved profile, to construct the probability distribution of the ozone number density in the entire altitude range being monitored. Using the proposed technique, we compare the retrieval results obtained for various (both well-known and suggested for the first time in this paper) methods of approximation and regularization of retrieved profiles. Model examples demonstrate that the proposed technique is capable of retrieving ozone-profile disturbances which accompany the formation of ozone holes.*

1. INTRODUCTION

Determination of the number densities of various chemical species in the atmosphere by the methods of passive ground-based remote sensing has been the subject of a large number of papers (see, e.g., [1–4]). Such methods allow one to continuously monitor the atmospheric composition irrespectively of the visibility of bright radio sources on the sky and at the altitude ranges hardly accessible for balloons. In addition, these methods are topical since the underlying experiments are considerably cheaper than balloon-borne (see, e.g., [5]) or satellite-born (see, e.g., [6]) observations.

One of the main problems immanent to ground-based techniques is related to retrieval of the altitude distributions (profiles) of the number densities of atmospheric trace species using sensing data. Experimental data used for such a retrieval include spectra of the atmospheric emission in the absorption bands of studied species, which are obtained on the ground. Hence the emission from the entire column of the atmosphere contributes to a received signal. This means that the desired profiles are related to the experimentally measured quantities by integral formulas, so that retrieval of a vertical profile is an inverse problem which, in general, is ill-posed (ill-conditioned).

Most popular methods for solving ill-posed problems, e.g., the Tikhonov method [7], the optimal estimation method [8], etc., are based on an approach in which *a priori* information on a retrieved profile is used. This approach yields solutions within a certain given class specified by the form of *a priori* information. This is achieved by regularization which consists in imposing a few certain additional conditions specifying the required class of solutions. Hardness of the constraints imposed by these condition is controlled by a special parameter.

For example, assumptions concerning smoothness of a desired distribution are explicitly imposed in the often used versions of the Tikhonov method. Namely, the inverse problem is solved under conditions constraining the values of derivatives of the retrieved distribution over the corresponding coordinate.

Institute of Applied Physics of the Russian Academy of Sciences, Nizhny Novgorod, Russia. Translated from Izvestiya Vysshikh Uchebnykh Zavedenii, Radiofizika, Vol. 46, Nos. 8–9, pp. 742–659, August–September, 2003. Original article submitted June 22, 2003.

In another method called the method of optimal estimation [8], a solution closest to a specific profile is sought under *a priori* assumptions on the altitude distribution of variations in a desired quantity.

The value of the regularization parameter in such methods is determined from the condition that the variance of the measurement noise is equal to the rms discrepancy between the data calculated using the retrieved profile and the experimental data.

Similar regularization techniques make an inverse problem well-conditioned, but, at the same time, introduce systematic retrieval errors which increase with strengthening of *a priori* requirements to the solution. As a result, the estimate of the retrieval accuracy can be biased, while the value of such a bias, which is not checked anyhow in the process of retrieval, is a function of the regularization parameter. As a result, the uncertainty of the retrieved profile decreases with increasing regularization strength, but the biases of the estimates become larger. On the contrary, if the regularization strength decreases, then the problem becomes less conditioned and, therefore, the retrieval accuracy becomes worse. Hence, the problem of choosing the optimal regularization is immanent to the above-described techniques. Solving this problem requires, as was mentioned above, physically reasonable *a priori* information on the characteristics of the retrieved distributions. However, a researcher often misses such information.

Another important aspect of the methods mentioned above is the proper choice of the cost function being optimized in the course of retrieval. In what follows we show that such a function should be determined with allowance for the assumptions concerning noise in the experimental data, and choosing an incorrect cost function leads to systematic errors.

In this paper, we propose to solve the problem of profile retrieval by applying the Bayesian approach for solution of inverse problems [9], in which *a priori* information on noise in the data is explicitly taken into account. In this approach, as well as in other methods, a model is used which includes, firstly, an integral functional assumed exactly known from the "first principles." Such a functional should solve the direct problem, i.e., allow one to calculate the received radiation spectrum given a certain altitude distribution of the desired quantity. Secondly, the model comprises a function approximating the altitude distribution and parameterizing the problem. Hence, a retrieved profile is determined by a set (vector) of parameters of this function (model parameters). In addition, a brand new element incorporating the considered problem into the framework of the Bayesian approach is introduced in the model. Namely, it is assumed that the unavoidable "randomness" of the received signal (the presence of a noise component) makes it possible to apply the statistical formalism to a calculation of the model parameters. By this, the problem of retrieval is reduced to a search for probability distributions of the parameters, determining, in turn, the probability distributions of the desired quantity. The latter distributions make it possible to estimate the confidence levels of retrieval for any altitude. Within the framework of such an approach, regularization consists in seeking solutions in a certain ensemble with an *a priori* given probability distribution (statistical regularization, see [10]), and the parameters of this distribution (for example, the variance in the case of a normal distribution) determine the "stiffness" of the regularization. In contrast to the standard retrieval techniques, in which these parameters are calculated from the discrepancy condition, this approach allows one to determine these parameters only on the basis of *a priori* assumptions on the characteristics of a retrieved profile.

The approach outlined above is considered below in view of its use for retrieval of the vertical distribution of the atmospheric ozone number density from data of radiometric (passive) sensing of the atmosphere in the millimeter wavelength range. The experimental data include the spectra of the atmospheric optical depth in the rotational lines of an ozone molecule.

We use model data to demonstrate the capabilities of the proposed technique. At first, we choose a certain model profile, substitute it into the known formulas, and calculate the spectrum of the ozone optical depth. Next, we add noise to the obtained "emission line." Then the noisy data are used to solve the inverse problem of retrieval.

To solve this problem, we used two different models. In the first model, a piecewise-uniform approximation of the profile was used, which was applied by many authors within the framework of the Tikhonov method and the optimal-estimation method. In the second model, a function in the form of an artificial neural network, which was the superposition of a certain number of strongly nonlinear function, was used as the approximator. We perform a comparative analysis of ozone profiles retrieved on the basis of these models and analyze how the retrieval depends on the way by which *a priori* information on the profile is input into the model.

2. DESCRIPTION OF THE BAYESIAN APPROACH TO SOLVING THE PROBLEM OF RETRIEVAL OF A VERTICAL OZONE DISTRIBUTION

The relationship between the ozone number density and the ozone optical depth τ is given by the integral formula [1]

$$\tau(f) = \int_0^\infty N(z)W(z,f)\,\mathrm{d}z, \tag{1}$$

which is linear with respect to the number-density profile $N(z)$. Here, $W(z,f)$ is the absorption cross section of an ozone molecule [11], f is the frequency, and z is the altitude. The altitude dependence of the kernel W of this integral is determined by the altitude distributions of the atmospheric pressure and temperature, which are assumed known. The experimental data include the values τ_i of the ozone optical depth corresponding to the frequencies f_i within a certain absorption (emission) line of ozone.[1]

Let noise immanent to measurement data due to instrumental errors be described by a random quantity ξ which is included additively in our model:

$$\tau_i = \int_0^\infty N(z)W(z,f_i)\,\mathrm{d}z + \xi_i. \tag{2}$$

The form of the joint probability distribution function $P_\xi(\xi_1,\,\dots\,,\xi_n)$ of this random quantity is determined by the assumptions concerning the noise, while the parameters of this function (such as, e.g., the covariance matrix) can be found using information on the measurement techniques or estimated by direct analysis of measurement data. Hence, this function is assumed *a priori* known.

The statistical formulation of the problem of ozone-profile retrieval consists in finding the probability distribution of various profiles $N(z)$ for given experimental data $\boldsymbol{\tau}, \mathbf{f}$. In other words, the statistical approach should yield the *a posteriori* probability density function $P(N(z)\,|\,\boldsymbol{\tau},\mathbf{f})$. It is evident that this can only be done if the probability distribution is mathematically meaningful or, in other words, the problem is parameterized. Such a procedure consists in approximating the profile $N(z)$ by a certain functional dependence $N(z,\boldsymbol{\mu})$ with the vector $\boldsymbol{\mu}$ specifying the parameters of the model. As a result, the desired distribution takes the form $P(\boldsymbol{\mu}\,|\,\boldsymbol{\tau},\mathbf{f})$, i.e., becomes the distribution of the model parameters specifying various profiles. Such a distribution function, also called the likelihood functional of a model [9], is constructed using the known Bayes theorem which relates the *a priori* and *a posteriori* conditional probabilities of random quantities:

$$P(\boldsymbol{\mu}\,|\,\boldsymbol{\tau},\mathbf{f}) \propto P(\boldsymbol{\tau},\mathbf{f}\,|\,\boldsymbol{\mu})P(\boldsymbol{\mu}). \tag{3}$$

According to this theorem, the desired likelihood functional is, firstly, proportional to the probability density function $P(\boldsymbol{\tau},\mathbf{f}\,|\,\boldsymbol{\mu})$ of the event that a model with a given set of parameters generates an observed spectrum. It is not difficult to see that the form of this function is determined by the assumptions on how noise is included in the model and by *a priori* information on the noise probability distribution:

$$P(\boldsymbol{\tau},\mathbf{f}\,|\,\boldsymbol{\mu}) = P_\xi\big(\boldsymbol{\tau} - \textstyle\int_0^\infty N(z,\boldsymbol{\mu})W(z,\mathbf{f})\,\mathrm{d}z\big)\,. \tag{4}$$

Secondly, the desired functional is proportional to the function $P(\boldsymbol{\mu})$ of the *a priori* probability density of

[1] In fact, the radiation temperature is initially found from the experimental data and then it is reduced to the optical depth [3].

the model parameters, which specifies a class of sought solutions and includes the actually available (and not dictated by the adopted method of regularization) *a priori* information on a retrieved profile. If no such information is available, then the *a priori* distribution of the model parameters becomes uniform, i.e., this function is a constant nondependent of these parameters. In this case, likelihood functional (3) is entirely determined by function (4). It is important that since the problem is ill-posed, the maximum value of the likelihood functional can correspond to an infinite set of the parameter vectors and, thus, the profiles $N(z, \boldsymbol{\mu})$. In the general case, this makes the desired distribution nonintegrable. As a result, it is impossible to find the most "optimal" profile and to estimate the confidence intervals of retrieval. Hence, in many cases, including an *a priori* distribution function of the parameters into a model plays crucial role for the correct regularization of a problem, and the parameters of this function (such as, e.g., the variance of the distribution) determine the "stiffness" of regularization.[2] However, we should note that if the noise variance tends to zero, then distribution function (3) of the parameters tends to a δ-function, so that the parameter set $\boldsymbol{\mu}$ corresponding to an exact profile yields the only solution whose probability is nonzero. Hence, the solution converges in the statistical sense.

In contrast to the Tikhonov method, in which a discrepancy condition imposes a stiff relationship between measurement data and regularization parameters, these parameters in our approach are *a priori* determined from physical considerations concerning the studied system. The regularization methods considered in the next section are based on using only very general *a priori* assumptions concerning the possible solutions under conditions when information on specific profiles (for example, the statistic of earlier measurements) is unavailable.

According to the method proposed in this paper, two additional procedures should be carried out after finding a likelihood functional. The first one is constructing (by the Monte-Carlo method) a statistical ensemble of model parameters, which has a probability distribution corresponding to the obtained likelihood functional. In the second procedure, confidence intervals of retrieval, corresponding to the selected confidence level, are calculated for the corresponding statistical ensemble.

3. TYPES OF MODELS

In this paper, we use two types of models in which different methods of approximation of the altitude distribution of the ozone density are used.

3.1. Piecewise-uniform model

Within the framework of this model, the entire atmosphere is divided into layers of certain thickness, and the density distribution is assumed constant in each layer. If the profile is approximated in such a way, then Eq. (2) is written as

$$\tau_i = \sum_{j=1}^{n} N_j W_j(f_i) + \xi_i,$$

or, in the matrix/vector form, as

$$\boldsymbol{\tau} = \mathbf{W}\mathbf{N} + \boldsymbol{\xi}, \tag{5}$$

where

$$W_j(f_i) = \mathbf{W}_{ij} = \int_{z_{j-1}}^{z_j} W(z, f_i)\,\mathrm{d}z. \tag{6}$$

The model is parameterized by the quantities N_j having the meaning of the ozone densities in the corresponding layers, and by the total number n of the layers.

[2] Note that, in addition to the said above, regularization can be the result of choosing the approximating function $N(z, \boldsymbol{\mu})$ with certain properties (see, e.g., the case of an artificial neural network discussed in the next section).

According to the procedure described in the previous section, we construct *a posteriori* probability density function (3) for the model parameters. At the first step, we should adopt certain assumptions concerning the statistical parameters of noise $\boldsymbol{\xi}$. Let noise ξ_i in different frequency channels be uncorrelated and have a normal distribution with the same variance σ_ξ^2, i.e., the joint probability density function of noise has the form

$$P(\boldsymbol{\xi}) \propto \prod_i \exp\left(-\frac{\xi_i^2}{2\sigma_\xi^2}\right) = \exp\left(-\frac{\|\boldsymbol{\xi}\|^2}{2\sigma_\xi^2}\right). \tag{7}$$

This yields (see Eq. (4)) the expression

$$P(\boldsymbol{\tau}, \mathbf{f} \,|\, \mathbf{N}) \propto \exp\left(-\frac{\|\boldsymbol{\tau} - \mathbf{WN}\|^2}{2\sigma_\xi^2}\right) \tag{8}$$

for the probability density of observation of the data $\boldsymbol{\tau}, \mathbf{f}$ for a given profile $\mathbf{N}$. Note that the expression in the exponential index in Eq. (8), taken with the opposite sign, corresponds to the cost function constructed when solving an optimization problem by the linear least-square method [12]. Evidently, an optimization problem with a sufficiently large number of layers, n, is ill-conditioned. As a result, the maximum of function (8) corresponds to linear hypersurfaces extending to infinity in the space of model parameters $\mathbf{N}$. Therefore, the set of "optimal" profiles turns out to be "spread" in the infinite limits with respect to the ozone number density. Hence, the problem should be regularized by constructing an *a priori* distribution of problem parameters, which specified a certain class of solutions. Consider two cases of such a distribution.

In the first case, the distribution of the density difference in the neighboring layers (in other words, the distribution of the derivatives of the profile $N(z)$ in the finite-difference representation) is *a priori* specified. It is assumed that this distribution is Gaussian with variance σ_N^2 being the parameter which determines the degree of regularization of the problem:

$$P(\mathbf{N}) \propto \exp\left(-\frac{1}{2\sigma_N^2}\sum_{j=1}^{n}(N_j - N_{j-1})^2\right) = \exp\left(-\frac{||\mathbf{CN}||^2}{2\sigma_{\Delta N/\Delta z}^2}\right). \tag{9}$$

Here, $\mathbf{C}$ is a matrix which, being multiplied by $\mathbf{N}$, yields the vector of altitude derivatives of the profile. *A priori* distribution of this kind imposes a constraint on the first derivatives of the profile, and thereby specifies the smoothness class of solutions. The fact that the maximum probability density in the statistical ensemble relevant to Eq. (9) corresponds to profiles with zero derivative over the altitude is not surprising since, within the framework of the outline statistical approach, information is contained in the moments of distribution (9) and not in the parameters of individual elements of the ensemble.

With allowance for Eq. (9), the *a posteriori* probability density given by Eq. (3) takes the form

$$P(\mathbf{N}|\boldsymbol{\tau}, \mathbf{f}) \propto \exp\left(-\frac{\|\boldsymbol{\tau} - \mathbf{WN}\|^2}{2\sigma_\xi^2} - \frac{||\mathbf{CN}||^2}{2\sigma_{\Delta N/\Delta z}^2}\right). \tag{10}$$

It is seen that the exponential index of this function comprises an expression whose form is similar to the functional which is regularized within the framework of the widely used Tikhonov method. We recall that the "Tikhonov" optimization is aimed at obtaining a solution satisfying a certain smoothness condition. Let us point out a significant difference of the proposed approach. In the Tikhonov method, the regularization parameter is determined in a unique and uncontrollable way from the condition of the maximum admissible mean square deviation of an experimentally measured line from the line calculated using a retrieved profile (the discrepancy condition). In the method proposed in this paper, this parameter is specified *a priori*, and its choice can be influenced both by the conventional views concerning the characteristics of the atmosphere in the studied region, which directly affect the smoothness of the altitude distribution being retrieved (such as, e.g., diffusion and the velocity of convective transport), and by the purposes of a specific study (for example, finding disturbances of the vertical distribution of ozone, which have certain spatial scales).

We also propose the second form of an *a priori* distribution of model parameters, which follows from the obvious physical assumption that the ozone number density is nonnegative at any altitude and has the form of a uniform distribution in a bounded area of the positive half-space of model parameters. This distribution if free of any *a priori* assumptions concerning the characteristics of the altitude distribution of ozone. In other words, minimum *a priori* information is used. The desired *a posteriori* distribution takes the form

$$P(\mathbf{N}\,|\,\boldsymbol{\tau},\mathbf{f}) \propto \exp\left(-\frac{\|\boldsymbol{\tau}-\mathbf{WN}\|^2}{2\sigma_\xi^2}\right)\Pi(0<N_j<N_{\max}). \tag{11}$$

The function Π is equal to unity if the condition in its argument is valid and is zero otherwise. The quantity $N_{\max}$ corresponds to an *a priori* specified constraint on the value of the number density. In this case, the only regularization parameter is the number n of atmospheric layers used for retrieval (the number of model parameters). We should note that such regularization was also used within the framework of the Tikhonov method [13].

3.2. Model based on an artificial neural network

In this model, we propose to approximate the profile by a function in the form of a simplest artificial neural network which is the superposition of a certain number of strongly nonlinear functions ("sigmoids") and plays the role of a "universal approximator" [14]:

$$N(z,\boldsymbol{\mu}) = \sum_{i=1}^{m} \alpha_i \,\mathrm{th}(\omega_i z + \gamma_i). \tag{12}$$

The neural-network coefficients α_i, ω_i, and γ_i specify the vector of model parameters: $\boldsymbol{\mu} = \{\boldsymbol{\alpha},\boldsymbol{\omega},\boldsymbol{\gamma}\}$. In this case, *a posteriori* probability density function (3) has the following form:

$$P(\boldsymbol{\mu}\,|\,\boldsymbol{\tau},\mathbf{f}) \propto \exp\left(-\frac{\|\boldsymbol{\tau}-\int N(z,\boldsymbol{\mu})W(z,\mathbf{f})\,\mathrm{d}z\|^2}{2\sigma_\xi^2}\right)\exp\left(-\frac{\|\boldsymbol{\omega}\|^2}{2\sigma_\omega^2}-\frac{\|\boldsymbol{\gamma}\|^2}{2\sigma_\gamma^2}\right). \tag{13}$$

Here, we used the Gaussian *a priori* distribution

$$P(\boldsymbol{\mu}) \propto \exp\left(-\frac{\|\boldsymbol{\omega}\|^2}{2\sigma_\omega^2}-\frac{\|\boldsymbol{\gamma}\|^2}{2\sigma_\gamma^2}\right)$$

for the parameters $\boldsymbol{\omega}$ and $\boldsymbol{\gamma}$. The *a priori* distribution of the parameters $\boldsymbol{\omega}$ constrains variations in the first derivative of the profile $N(z)$ and, thus, specifies the degree of smoothness of the profile. The variance of this distribution is determined on the basis of *a priori* assumptions concerning the scales of profile disturbances (see Sec. 2). The distribution of the parameters $\boldsymbol{\gamma}$ contains *a priori* information on the altitude localization of the profile, so that this distribution specifies the domain of definition of function (12). The interval of retrieval altitudes determines the corresponding variance. It is assumed that these parameters can be estimated independently of the experiment on the basis of *a priori* data on the parameters of the atmosphere and specifications of the measuring device being used. The necessity of introducing such *a priori* distributions is also stipulated by the fact that the function $\exp[-\|\boldsymbol{\tau}-\int N(z,\boldsymbol{\mu})W(z,\mathbf{f}\,\mathrm{d}z\|^2/(2\sigma_\xi^2)]$ in Eq. (13) is nonintegrable with respect to the parameters $\boldsymbol{\omega}$ and $\boldsymbol{\gamma}$ since neural network (12) has constant nonzero asymptotes in the region where these parameters are large in the absolute value. The absence of *a priori* information about these parameters would make desired distribution (13) nonintegrable and, therefore, the statistical analysis of this distribution, impossible. Such problems are absent for the parameters $\boldsymbol{\alpha}$, and the *a posteriori* probability density corresponding to model (12) is integrable over these parameters. Owing to this fact, no additional constraints were imposed on these parameters, i.e., their *a priori* distribution was assumed uniform and was not included in Eq. (13).

Another regularization parameter of the considered model is the number of neurons, m, which determines the number of intervals in which the solution is monotonic. Thus, model (12) makes it possible to separate the regularization parameters specifying the number of monotonicity intervals of a profile from the parameters determining the degree of its smoothness. By this, we ensure a more complete and optimal use of *a priori* information on an ozone distribution profile. It will be shown in the next section that, concerning retrieval of local disturbances of a profile, such a model is superior in comparison with the piecewise-uniform one.

4. DEMONSTRATION OF THE CAPABILITIES OF THE APPROACH USING MODEL DATA

To demonstrate the capabilities of the outlined approach, we simulated a spectrum of the optical depth of atmospheric ozone in the frequency range corresponding to the ozone emission line at a frequency of 142 GHz. With allowance for state-of-the-art instrumentation parameters, the modeling procedure employed a fairly modest ozonometer sensitivity and analyzed spectrum bandwidth and consisted in the following. First of all, a model altitude distribution of the ozone number density was constructed. Next, this profile was used for calculating the spectrum of the ozone optical depth in accordance with Eq. (1). Then, white Gaussian noise was added to the resulting spectrum. The resulting noisy line was used for retrieval of the profile of the ozone number density. The mean square deviation of the noise component was chosen equal to 10^{-3} Np, which corresponds to typical conditions of an observation with averaging over $10-15$ readouts and amounts to about 50% of the signal level in spectral channels at frequencies about $110-130$ MHz apart from the line center.

As an example, we present in Fig. 1*a* a model profile of ozone, which includes both the intervals in which the ozone number density is a smooth function of altitude and a sharp "trough" in the altitude range from about 15 to 30 km. Firstly, such an altitude distribution of ozone allows us to study the capabilities of the proposed approach. Secondly, this distribution simulates an actual case realized during the formation of the ozone hole in the winter–spring polar stratosphere [15]. The adopted profile was used to calculate the spectrum of the ozone optical depth. The line calculated according to Eq. (1) and the noisy line are shown in Fig. 1*b*.

The likelihood functionals of the models described in the previous section (see Eqs. (10), (11), and (13)) were used for the retrieval. According to the Monte-Carlo technique, we generated statistical ensembles of model parameters, which have probability distributions specified by these functionals. To obtain an ensemble of quantities distributed in accordance with a given multi-dimensional nonlinear function, we used a technique described in detail in [16, 17]. Generating an ensemble of 10^4 elements required about 20 min of calculations on a PC with a CPU clock rate of 2.5 GHz. Next, the obtained ensembles of parameters were used to calculate the ensembles of vertical profiles of the ozone number density. This permitted calculating the probability distributions of the ozone number density at any altitude. These distributions were then used to calculate the confidence intervals for the number density, which determine the retrieval accuracy.

Figure 2 shows the results of retrieval of the ozone number-density distribution using piecewise-uniform model (5) with the *a priori* parameter distribution given by Eq. (9). The results correspond to three different values of the regularization parameter $\sigma_{\Delta N/\Delta z}$ specifying the *a priori* mean square deviation of the first derivative of the profile $\mathbf{N}$. It is seen that the profile is "oversmoothed" due to a too strong regularization for $\sigma_{\Delta N/\Delta z} = 5$ (Fig. 2*a*). In this case, the conditionality of the problem is better than for $\sigma_{\Delta N/\Delta z} = 10$ and $\sigma_{\Delta N/\Delta z} = 20$ (the confidence intervals have the smallest widths), but, at the same time, the strongest biases of the estimates are observed for $\sigma_{\Delta N/\Delta z} = 5$. The "actual" profile at some altitudes is not covered even by the 95% confidence interval. On the contrary, in the case of a too weak regularization, for $\sigma_{\Delta N/\Delta z} = 20$ (Fig. 2*c*), the problem becomes "more ill-posed," and the resulting confidence intervals are

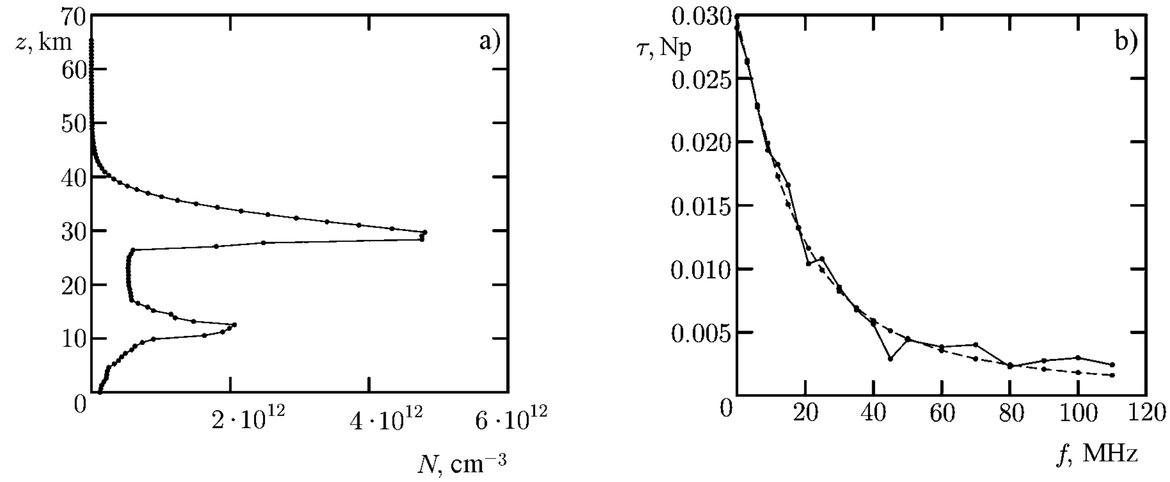
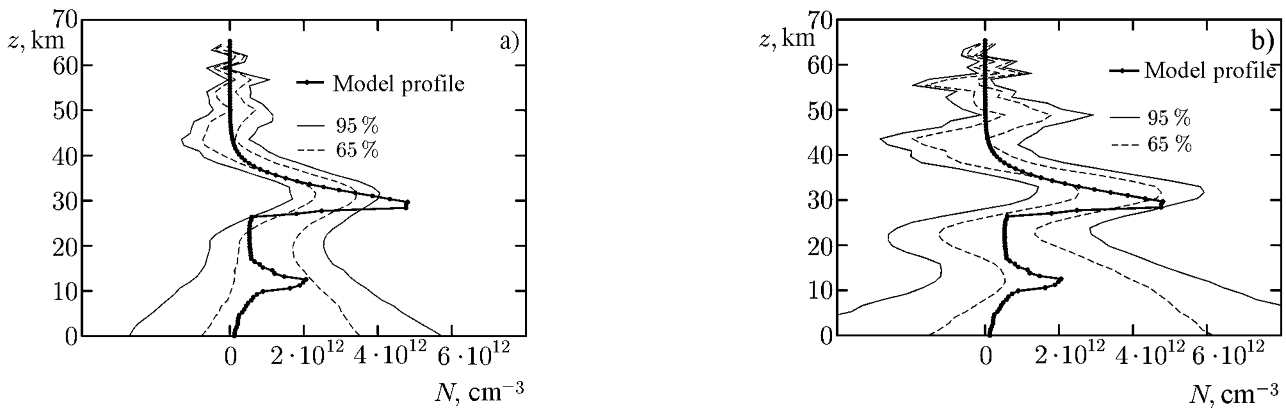


Fig. 1. Model altitude dependence of the ozone number density (*a*) and the spectrum of ozone optical depth (*b*) calculated according to Eq. (1) from the model profile. The spectrum without noise is shown by a dashed line, and the spectrum containing noise with $\sigma = 10^{-3}$ Np, by a solid line.



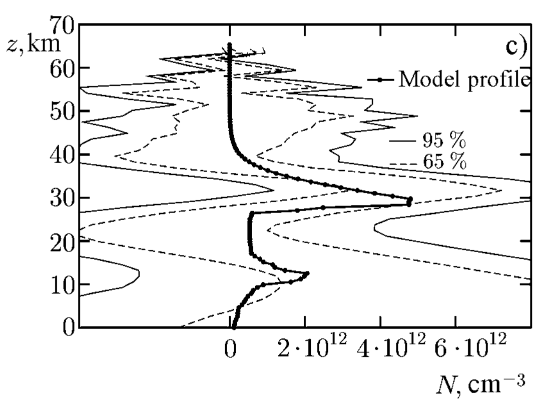


Fig. 2. Results of retrieval of the ozone number-density profile using functional (10) for $\sigma_{\Delta N/\Delta z} = 5$ (*a*), $\sigma_{\Delta N/\Delta z} = 10$ (*b*) and $\sigma_{\Delta N/\Delta z} = 20$ (*c*). The solid and dashed lines, respectively, border the 95% and 65% confidence intervals of the retrieved number density of ozone. The model ozone profile is shown by dots.

wide, thus making the retrieval accuracy inadmissible. The value $\sigma_{\Delta N/\Delta z} = 10$ (Fig. 2*b*) corresponds to the "optimal" regularization in the considered case. Hence, the Bayesian approach for retrieval of the vertical ozone profile using the piecewise-nonuniform model demonstrated high sensitivity to the parameters specifying the class of solution. This complicates the problem of choosing the optimal values of these parameters under conditions of limited *a priori* information on the number-density profile.

Figure 3 shows the result of retrieval of the ozone profile within the framework of the same model (5), but with another type of regularization which does not allow for the negative values of the number density.

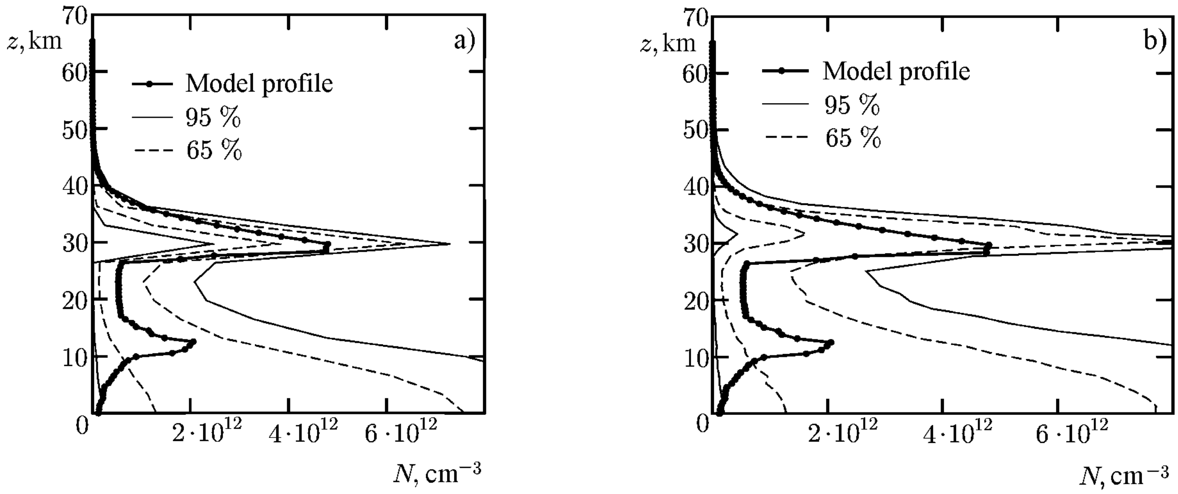
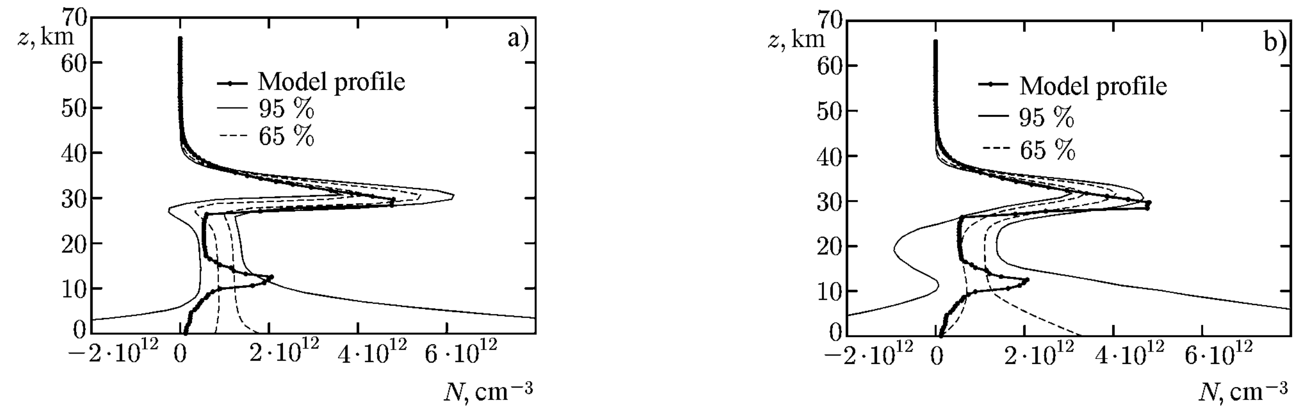


Fig. 3. The same as in Fig. 2, but using likelihood functional (11) for $n = 20$ (*a*) and $n = 50$ (*b*).



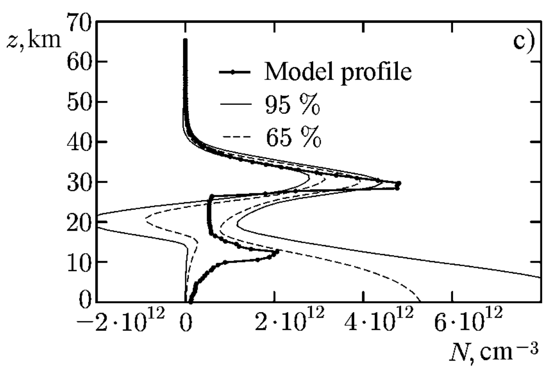


Fig. 4. Results of retrieval using model (12) (probability functional (13)) for $\sigma_\omega = 100$ (*a*), $\sigma_\omega = 25$ (*b*), and $\sigma_\omega = 10$ (*c*). The nomenclature is identical to Figs. 2 and 3.

The likelihood functional for such method of retrieval is described by Eq. (11). Figures 3*a* and 3*b* show the retrieval results for the models including 20 and 50 atmospheric layers, respectively. It is seen that an increase in the number of layers n worsens the retrieval accuracy due to "weakening" of regularization. Nevertheless, even such a minimal regularization permits retrieval local disturbances of a profile, although the retrieval accuracy is low.

The results of retrieval using the model based on artificial neural network (12) with a likelihood functional given by Eq. (13) are shown in Fig. 4. The neuron number m was put equal to five, which corresponds to five plausible intervals of monotonic behavior of the retrieved profiles. The dependence of the retrieval on the mean square deviation σ_ω of *a priori* distribution of the parameters $\boldsymbol{\omega}$ specifying the degree of smoothness of the profile was analyzed. As this parameter increases from 10 (Fig. 4*c*) to 100 (Fig. 4*a*), the requirements to smoothness of a solution become less stiff, and the sharp variation in the number density at altitudes $27-30$ km is retrieved with good accuracy for $\sigma_\omega = 100$. However, this parameter does not affect the "irregularity" of the profile since the latter feature is controlled only by the number of neurons m (as was mentioned above, the corresponding regularization is already taken into account in the model itself). Thus, an *a priori* distribution of the parameter in this model constrains only the values of local disturbances of the profile. This is different from, e.g., the piecewise-uniform model in which the regularization determines the "irregularity" of a solution in the entire altitude range of retrieval. It is seen in Fig. 4 that such use of *a priori* information, which is more optimal in our opinion, makes it possible to retrieve local disturbances of a number-density profile of ozone with better accuracy than within the framework of the piecewise-uniform model.

5. CONCLUDING REMARKS

In this paper, we propose a method for retrieving vertical number-density profiles of chemical species in the atmosphere using data of passive radiometric measurements. The proposed method is based on the Bayesian approach for solving ill-posed inverse problems. It allows one to construct probability distributions of the desired quantity in the entire altitude interval of sounding by using information on noise unavoidably present in the measurement data and using *a priori* known characteristics of the profile.

The capabilities of the method are demonstrated by applying it to the problem of retrieval of the vertical number-density profile of atmospheric ozone from a model spectrum of the optical depth in one emission line of ozone. The models based on two different approximations of the profile, the piecewise-uniform approximation and the approximation by a neural-network function, are used in the process of retrieval.

In the case of the model using a piecewise-uniform approximation, we consider two methods of regularization of the problem, i.e., taking into account *a priori* information on the profile, within the framework of the Bayesian approach. The first method allows for obtaining solutions satisfying a given smoothness criterion. It is shown that the results of retrieval obtained by this method of regularization are highly sensitive to the value of the regularization parameter. If the requirements to smoothness of a solution weaken (i.e., more indented profiles are allowed), then the retrieval uncertainty increases. Otherwise, the systematic biases of the estimates become larger. The second method of retrieval, proposed by the authors of this paper and using obvious physical assumptions on the nonnegative number density, is more rough with respect to the values of the regularization parameters, but ensures the smaller error of retrieval. A combination of the first and second regularization methods should probably be more efficient. Another model proposed in this paper assumes that the desired profile is approximated by a function in the form of a neural network. Retrieval of a profile by using such a model ensures the more complete allowance for *a priori* information on the profile, namely, makes it possible to separate the regularization parameters specifying the number of monotonicity intervals from the parameters responsible for the profile smoothness. In comparison with the piecewise-uniform model, this provides for the better accuracy of retrieval of local disturbances of a profile.

The authors are grateful to I. V. Kuznetsov, V. G. Ryskin, and A. A. Shvetsov for fruitful discussions and to K. P. Gaikovich for a helpful discussion.

This work was supported by the Russian Foundation for Basic Research (project No. 02–02–16540 and 03–05–64311) and the Program of Fundamental Studies "Physics of the Atmosphere: Electric Processes and Radiophysical Methods of Study" by the Division of Physical Sciences of the Russian Academy of Sciences (projects 2.2 and 2.3).

REFERENCES

1. T. L. Erukhimova and E. V. Suvorov, *Radiophys. Quantum Electron.,* **44**, Nos. 1–2, 129 (2001).
2. K. P. Gaikovich, E. P. Kropotkina, and S. V. Solomonov, *Izvestiya, Atmos. Oceanic Phys.,* **35**, No. 1, 78 (1999).
3. O. S. Mocheneva, T. L. Erukhimova, and E. V. Suvorov, *Radiophys. Quantum Electron.,* **38**, No. 8, 499 (1995).
4. A. K. Randegger, *Pure Appl. Geophys.,* **118**, 1052 (1980).
5. D. J. Hofmann, S. J. Oltmans, J. M. Harris, et al., *J. Geophys. Res. D,* **102**, No. 7, 8931 (1997).
6. J. W. Waters, *Microsc. Opt. Tech. Lett.,* **4**, 2 (1991).
7. A. N. Tikhonov, *Sov. Math. Doklady,* **4**, 1035 (1963).
8. C. D. Rodgers, *J. Geophys. Res. D,* **95**, No. 5, 5587 (1990).
9. J. Tamminen and E. Kyrola, *J. Geophys. Res. D,* **106**, No. 13, 14377 (2001).
10. V. F. Turchin, V. P. Kozlov, and M. S. Malkevich, *Sov. Phys. — Uspekhi,* **13**, 681 (1971).
11. C. H. Townes and A. L. Schawlow, *Microwave Spectroscopy,* McGraw-Hill, New York (1955).
12. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, in: *Numerical Recipes,* Cambridge Univ. Press (1989), p. 499.
13. K. P. Gaikovich, *Digest of IGARSS'94, Pasadena, USA, August 8–12, 1994. Vol. 4,* p. 1901.
14. K. Hornik, M. Stinchcombe, and H. White, *Neural Networks,* **2**, 359 (1989).
15. B. Nardi, W. Bellon, L. D. Oolman, and T. Deshler, *Geophys. Res. Lett.,* **26**, No. 6, 723 (1999).
16. W. R. Gilks and P. Wild, *Appl. Statistics,* **41**, No. 2, 337 (1992).
17. W. K. Hasting, *Biometrica,* **57**, 97 (1970).