ELSEVIER

Journal of Quantitative Spectroscopy & Radiative Transfer 86 (2004) 115–131

Journal of Quantitative Spectroscopy & Radiative Transfer

www.elsevier.com/locate/jqsrt

The retrieval of profile and chemical information from ground-based UV-visible spectroscopic measurements

R. Schofield[a,b,*], B.J. Connor[a], K. Kreher[a], P.V. Johnston[a], C.D. Rodgers[a,c]

[a] *National Institute of Water and Atmospheric Research, Lauder, New Zealand*
[b] *School of Geography and Environmental Science, Auckland University, Auckland, New Zealand*
[c] *Atmospheric, Oceanic and Planetary Physics, Oxford University, Oxford, UK*

Received 15 April 2003; received in revised form 12 June 2003; accepted 9 July 2003

Abstract

An algorithm has been developed to retrieve altitude information at different diurnal stages for trace gas species by combining direct-sun and zenith-sky UV-visible differential slant column density (DSCD) measurements. DSCDs are derived here using differential optical absorption spectroscopy. Combining the complementary zenith-sky measurements (sensitive to the stratosphere) with direct-sun measurements (sensitive to the troposphere) allows this vertical distinction. Trace gas species such as BrO and NO_2 have vertical profiles with strong diurnal dependence. Information about the diurnal variation is simultaneously retrieved with the altitude distribution of the trace gas. The retrieval is a formal optimal estimation profile retrieval, allowing a complete assessment of information content and errors.

Keywords: Ozone; DOAS; Air mass factor; Inverse; Radiative transfer

1. Introduction

Retrieval of the altitude distribution of trace gases from differential slant column density, $\Delta\Phi$ (DSCD) measurements using differential optical absorption spectroscopy (DOAS) was first attempted by Brewer et al. in 1973 [1] for NO_2. Brewer et al. and later Noxon [2] performed simple analyses to infer stratospheric and tropospheric NO_2 concentrations from measurements of the slant column density (SCD). McKenzie et al. [3] performed a formal retrieval of O_3 and NO_2 profiles

* Corresponding author. National Institute of Water and Atmospheric Research, Lauder, New Zealand. Fax: +64-3-447-3348.
E-mail address: r.schofield@niwa.co.nz (R. Schofield).

0022-4073/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0022-4073(03)00278-4

using zenith-sky SCDs with no error treatment. Preston et al. [4] implemented Rodgers [5] optimal estimation technique to retrieve NO_x profiles (and in turn NO_2 profiles using a chemical box model) from zenith-sky NO_2 DSCD measurements with a complete error and sensitivity analysis. The concept of a combined DOAS-optimal estimation profile retrieval has also been applied to the limb sounding satellite experiment OSIRIS to retrieve ozone, NO_2, BrO and OClO stratospheric profiles [6].

Profile retrievals involving ozone and NO_x (a quantity that does not have a significant diurnal variation) avoid the complications that arise due to the strong diurnal variation typical for radical or radical-like trace gas species (NO, NO_2, IO, OClO, BrO, etc.) involved in photochemical cycles. This work explores the complications and errors arising from diurnal variation when incorporated in the retrieval of altitude information. We find that retrieving several profiles describing the diurnal variation maximises the amount of information that can be retrieved from the measurements. It is also found that the errors arising from treating the diurnal variation as a forward model parameter (i.e. supplying the diurnal variation a priori (initial knowledge)) are too large to be ignored.

The variation with time of profiles of diurnally varying species causes complication in the profile retrieval from measurements. The highest sensitivity to altitude is obtained using twilight spectra for the geometries considered here. This is also when the photochemistry is changing most rapidly. Livesey et al. [7] detail the use of the optimal estimation technique in retrieving temperature and composition profiles along the line of sight of limb sounding satellite measurements. The concept of retrieving the 'horizontal gradient' from limb sounding satellite measurements has been developed further by Kemnitzer et al. [8]. A similar concept is applied here to retrieve a set of profiles at different diurnal stages concurrently from the observed DSCDs.

The concept of combining complementary measurements in a formal retrieval was developed for ozone using the Huggins bands by Jiang et al. [9] in a theoretical study. Jiang et al. found combining direct and diffuse measurements had the potential to be an effective technique to retrieve the tropospheric ozone column amount. In this study, the advantages of combining the two spectroscopic viewing geometries of direct-sun and zenith-sky, to give tropospheric and stratospheric sensitivity, is investigated. This retrieval method has the advantage that it may be extended to include many measurement platforms.

2. Spectroscopy

Spectroscopic radiance measurements can be made from a number of viewing platforms, e.g. direct-sun as ground-based or a balloon platform [10], or observing scattered light as zenith viewing [11], and off-axis [12]. UV/visible spectroscopy is used in satellite experiments, for example, GOME [13] and SCIAMACHY [14]. There are a vast number of trace gas species detectable using spectroscopy in the UV-visible wavelength region e.g. HONO, OH, NO_3, BrO, ClO, OClO, NO_2, NO, NH_3, IO, O_3, SO_2, CS_2, HCHO, O_4, H_2O.

Direct-sun spectroscopy samples a single light path through the atmosphere. The path through the troposphere is up to 20 times longer than the vertical path when the sun is close to the horizon.

Zenith-sky spectroscopy samples light that has traversed a number of paths before being scattered into the detector. The density of air and aerosols, and the wavelength region determine the altitude from which most of the detected light has been scattered from. Very generally for high

sun measurements at $\lambda = 350$ nm this maximum scattering altitude is ~5–10 km. Below this altitude the light traverses a vertical path into the detector. For low sun this maximum scattering altitude is higher ~15–30 km. The light path enhancement for the stratosphere can be up to 20 times that of the vertical path. However, the light path enhancement through the troposphere into the detector is close to one.

The atmospheric sensitivity of the direct-sun and zenith-sky viewing geometries is quite different, because of the different altitudes of path enhancement. The zenith-sky measurements are more sensitive to the stratosphere, while the direct-sun measurements have a higher tropospheric sensitivity. The twilight periods are important for obtaining enhanced pathlengths for both viewing geometries.

2.1. Differential optical absorption spectroscopy (DOAS)

The well-known DOAS fitting technique is used to determine DSCDs (for DOAS reviews see Noxon [2], and Platt [15]). DOAS targets the detection of species with distinct absorption features (~5 nm in width or less) on a background that contains broadband absorbers, Mie, and Rayleigh scattering features.

A ratio of radiance spectra with respect to a reference spectrum is taken. The reference spectrum is taken when there is no or only a small amount of absorber present. Complicating features common to both spectra are removed e.g., the Fraunhofer lines when a solar light source is used. Low-order polynomials are fitted to approximate the Rayleigh, Mie and broadband absorption features that vary slowly with wavelength. The cross sections of each absorber expected to be present are fitted in a non-linear least-squares fitting procedure to determine their DSCDs [16,17].

Subsequently, a retrieval is performed using the DSCDs as the measurement vector $\mathbf{y}$. It is this latter retrieval that is explored by this work to illustrate how different measurement platforms can be combined to provide altitude distribution on a temporal scale for trace gas concentrations. This retrieval method can be adapted to use any spectroscopically derived quantity (SCD or DSCD) or the radiance measurements directly. An accurate error assessment of the quantity being used in the retrieval's measurement vector $\mathbf{y}$ is required.

The measurement error covariance matrix ($\mathbf{S}_\varepsilon$) incorporates all the errors that are present in determining the DSCDs. Errors in the absorption cross-sections, their temperature dependencies, Ring, polarization, dark current, instrument function, Mie and Rayleigh effects have to be considered in the construction of $\mathbf{S}_\varepsilon$.

3. Retrieval algorithm

Remote sensing techniques often have the disadvantage that the quantity being measured has a complex relationship to that which is desired to be known. Optimal estimation is a retrieval method that allows the inversion of the 'measurements' $\mathbf{y}, \Delta\Phi$ (DSCDs), into the desired state $\mathbf{x}$. The state is in this instance a set of profiles at different diurnal stages. For a given state $\mathbf{x}$ a forward model must describe the expected DSCDs $\hat{\boldsymbol{y}}$. This is the radiative transfer part of the algorithm. For a complete retrieval description refer to Rodgers [5]. Fig. 1 displays the basic structure of the retrieval algorithm illustrating how complementary sets of measurements are combined. The following sections give a summary of the expressions and terms used in the algorithm.

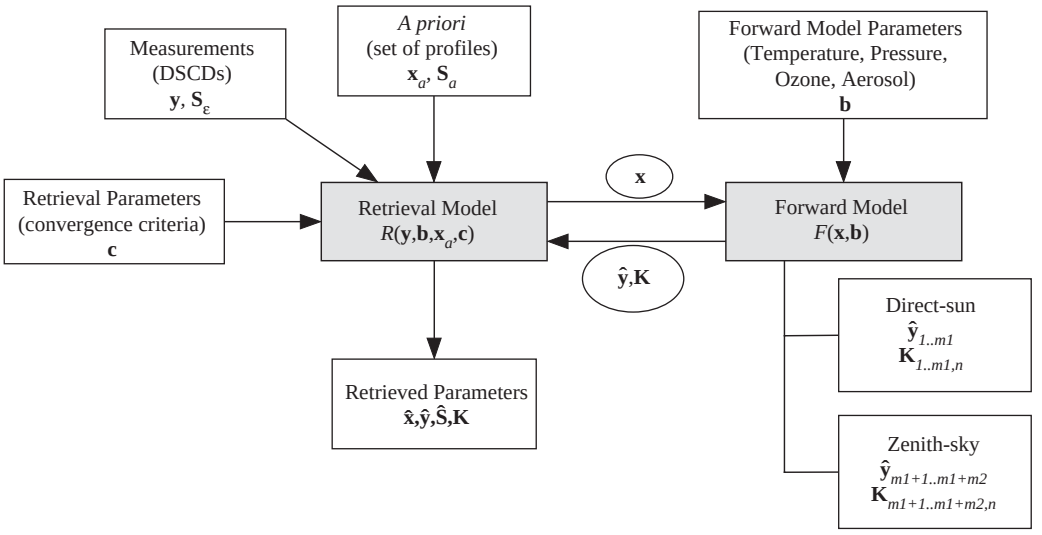

Fig. 1. A schematic diagram that displays the basic structure of the retrieval algorithm.

3.1. Forward model: radiative transfer

The forward model is an approximation to the complete description of the physics of the measurement and is expressed as

$$\mathbf{y} = F(\mathbf{x}, \mathbf{b}) + \varepsilon, \tag{1}$$

where $\mathbf{y}$ is the measurement vector, F is the forward model, the radiative transfer calculation of $\hat{\mathbf{y}}$ (the modelled DSCDs) and ε is the error in the measurement. The forward model requires a description of the state ($\mathbf{x}$), the quantity that is being retrieved. Also required are the forward model parameters ($\mathbf{b}$), which affect the calculated DSCDs, but are not retrieved.

The radiative transfer descriptions for the direct-sun and zenith-sky geometries contain essentially the same core equations. A spherical curved earth divided into discrete atmospheric shells is used in the model. Radiative descriptions for the two different viewing geometries include the effects of refraction, Rayleigh scattering, Mie scattering and molecular absorption.

Refraction is included in the radiative transfer using refractive indices from Bucholtz [18]. The path through each discrete shell, and the earth centric angle are calculated using polar coordinates to avoid the singularity problems when refracting through a tangent point [5,19].

Rayleigh cross-sections and Rayleigh phase functions from Bucholtz [18] are used to account for the Rayleigh scattering out of the calculated light path and the single scattering into the detector for the zenith-sky geometry. Aerosol scattering out of the light path is determined from aerosol extinction profiles from LIDAR and backscatter data [J.B. Liley, Personal Communication, 2003]. Mie scattering into the detector is approximated using a one parameter Henyey–Greenstein-based phase function [20].

Absorption from the optically thick species ozone is included, with profiles for Lauder from ozonesonde data [21]. This is also the source of temperature and pressure profiles as forward model

parameters. Absorption from NO_2 is also included when calculating DSCDs for weaker absorbers. NO_2 profiles for Lauder are obtained from SAGE II [22].

DOAS fitting provides the DSCDs. SCDs Φ, and DSCDs $\Delta\Phi$, are modelled as the number of molecules per unit area. Where j is any given solar zenith angle (SZA) and the subscript R refers to the reference solar zenith angle [23]:

$$\Phi_j = -\frac{1}{\sigma(\lambda)} \ln\left(\frac{I_j(\lambda)}{I'_j(\lambda)}\right) \quad (\text{molecules cm}^{-2}), \tag{2}$$

$$\Delta\Phi_j = \mathbf{y}_j = \Phi_j - \Phi_R \quad (\text{molecules cm}^{-2}) \tag{3}$$

I_j is the intensity that reaches the detector at a given wavelength (λ) and SZA (j) in the model atmosphere with the absorber of interest present and I'_j is the intensity without the absorber of interest. $\sigma(\lambda)$ is the cross-section for the absorber of interest.

The direct-sun viewing geometry requires only one path, as the effects of multiple scattering are negligible. For the zenith-sky viewing geometry a single scattering approximation is made. The influence of multiple scattering on the DSCDs for BrO is found to be negligible up until about 92° [24,25].

The weighting function matrix **K** describes the sensitivity of the measurements to changes in the state and is determined by perturbation of each state vector quantity. This is the most computationally expensive part of the retrieval

$$\mathbf{K} = \frac{\partial F}{\partial \mathbf{x}}. \tag{4}$$

The measurement vector **y** in the retrieval algorithm is constructed using the direct-sun measurements to make up the first m_1 values, then the zenith-sky measurements make up the subsequent m_2 values in the vector. As **K** describes how each measurement responds to a change in the state at level n, it is constructed similarly with the direct-sun weighting functions preceding the zenith-sky weighting functions. Further measurement platforms could be added to the basic structure displayed in Fig. 1. To add further platforms an appropriate forward model and means of calculating the weighting functions is required.

3.2. Retrieval model and characterisation

An optimal estimation type of retrieval is used to obtain the best solution from a set of all possible solutions for the case where the problem is ill-posed. The retrieval problem here is formally ill-posed, as there are more elements in the state vector than there are measurements. There is an infinite set of solutions and the a priori constraint allows a single solution to be determined.

A retrieval is characterised by the averaging kernel matrix, **A**, which describes the sensitivity of the retrieved state ($\hat{\mathbf{x}}$) to the true state ($\mathbf{x}$). It is calculated from the weighting function and covariance matrices by

$$\mathbf{A} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} = (\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\varepsilon}^{-1}\mathbf{K} + \mathbf{S}_{\mathrm{a}}^{-1})^{-1}\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\varepsilon}^{-1}\mathbf{K}. \tag{5}$$

The rows of $\mathbf{A}$ are the averaging kernels, which express the relationship between the true state and the retrieved state. An averaging kernel for a retrieval point or quantity describes how the true atmospheric state is smoothed in the retrieval of that quantity. In the ideal case the averaging kernel matrix would be the identity matrix.

The covariance matrix of the uncertainties in the measurements, $\mathbf{S}_\varepsilon$, is constructed by placing the variance arising from one standard deviation of the error from the DOAS fitting along the diagonal of the $m \times m$ matrix. This covariance matrix includes the systematic errors that arise in the determination of the DSCDs.

The construction of $\mathbf{S}_a$ is not so straightforward. If the equation is to be used formally as an optimal estimator, then $\mathbf{S}_a$ describes whatever is known about the state (a priori knowledge), however little. In principle the best that can be carried out for a case taken at random from some ensemble of profiles is to use the ensemble covariance. If no ensemble is available then care is needed. Too tight a constraint leads to an $\mathbf{S}_a$ matrix which implies low error in the a priori state ($\mathbf{x}_a$). This will result in a retrieval biased towards the a priori with the information from the measurements being given less weight. Conversely too loose a constraint, assuming a high error in the a priori state, results in the measurement being over-fitted, and noise being interpreted as data.

In this work, a percentage of the peak for each of the temporal profiles is used as a conservative estimate of the a priori variance. The value of this percentage is determined by an l-curve method. The percentage is varied from a very low value (tightly constrained) to a very high value (loosely constrained). A plot of the root mean square of the retrieval fit to the measurements versus this percentage is constructed. The percentage chosen is the value where increasing the a priori error no longer results in a marked improvement of the measurement fit.

The area of an averaging kernel is the sum of its elements, giving a qualitative indication of the influence of the a priori on the retrieved state. When the area is approximately unity then most of the information is being supplied by the measurements rather than the a priori. When the area is close to zero then the retrieved value has been derived from the a priori information and the measurements have not added any information. Where the area of the averaging kernel is negative, or much larger than one, the retrieved values may not be a good representation of the true atmosphere.

The 'degrees of freedom for signal' describe the number of useful independent quantities that can be determined from a set of measurements. It can be evaluated by taking the trace of the averaging kernel matrix $\mathbf{A}$. Each value along the diagonal of $\mathbf{A}$ can be viewed as the degrees of freedom per retrieval point.

The information content of the measurements can be thought of as a measure of the factor by which knowledge is improved by making the measurement. It can be thought of as a multivariate generalization of the scalar concept of signal-to-noise ratio for the retrieval. Shannon's definition of information content is the reduction of entropy (or uncertainty) in the state that is achieved by making the measurement:

$$H = \frac{1}{2}\log_2\left(\frac{|\mathbf{S}_a|}{|\hat{\mathbf{S}}|}\right) \quad \text{or} \quad H = \frac{1}{2}\log_2|\mathbf{I} - \mathbf{A}|, \tag{6}$$

where the units of H are bits. Two bits of information means that the entropy has been reduced by a factor of four in making the measurement. $\mathbf{I}$ is the identity matrix and $\hat{\mathbf{S}}$ is retrieval covariance matrix (refer to Rodgers [5] for the derivation of all covariance matrices).

The covariance due to each forward model parameter ($\mathbf{S}_f$) is a useful error term to evaluate. This covariance provides a quantitative evaluation of the error that a forward model parameter produces in the final retrieved quantity. If this error is too large, then consideration should be given to retrieving the forward model parameter.

The spatial resolution of a retrieval is an important diagnostic, but in this case it is complicated by the two-dimensional nature of the state vector. Rather than considering horizontal and vertical resolution separately as some 'width' of the averaging kernel, we use the simpler measure of the reciprocal of the diagonal values of $\mathbf{A}$. The diagonal values of $\mathbf{A}$ give the number of degrees of freedom per state vector element. The reciprocal gives the number of state vector elements required to describe a degree of freedom.

Inversion of the measurements can be performed using the following formula for a linear problem:

$$\hat{\mathbf{x}} = \mathbf{x}_a + \mathbf{S}_a\mathbf{K}^T(\mathbf{K}\mathbf{S}_a\mathbf{K}^T + \mathbf{S}_\varepsilon)^{-1}(\mathbf{y} - \mathbf{K}\mathbf{x}_a). \tag{7}$$

The measurements considered in this work are in optically thin spectral regions and have a linear dependence on the profile of the trace gas. The optical density of a species depends upon its absorption coefficient and the measurement geometry. If this general retrieval technique is applied to retrieve profile information for a species that is optically thick, then a non-linear retrieval will be necessary.

4. Diurnally varying species

Spectroscopy in the UV-visible region is advantageous in the detection of highly reactive species, many of which play critical roles in ozone chemistry. These species are either radical in nature or their concentrations depend directly on radical chemistry. In general their resultant vertical profiles have strong diurnal variations due to photochemistry. The case study discussed below focuses on the retrieval of altitude information for the radical species of bromine monoxide (BrO). BrO is present in the atmosphere at low concentrations, has a strong diurnal variation and is involved in catalytic ozone destruction cycles in the stratosphere. Recent studies have suggested that BrO is present in the free troposphere at all latitudes [26,27]. The stratospheric diurnal variation of BrO is displayed in Fig. 2.

Diurnal variation of the trace gas profile adds complication to the retrieval problem. There are two approaches that allow the retrieval of profiles of diurnally varying trace gases. The first approach is to assume that the chemistry is known a priori from chemical models and that a single profile varies in a predetermined way, for example as a scalar multiple or a more complex change of profile shape. The diurnal variation is thus being treated as a forward model parameter. The retrieval errors associated with this approach are explored below. However, the main concern with this approach is that while existing chemical models explain diurnal variations for the stratosphere (SLIMCAT [25,28], etc), tropospheric diurnal variations are essentially unknown and an area of active research. Assuming the diurnal variation from stratospheric models alone causes discontinuity in the weighting functions at the troposphere–stratosphere boundary. Consequently this produces errors in the final retrieved state.

The second approach, used here to account for the diurnal variation of the species, is to retrieve it. Sets of number density profiles defined on a time (or SZA) grid, to describe the diurnal variation of the species, are retrieved. This leads to an increased number of retrieved parameters but avoids

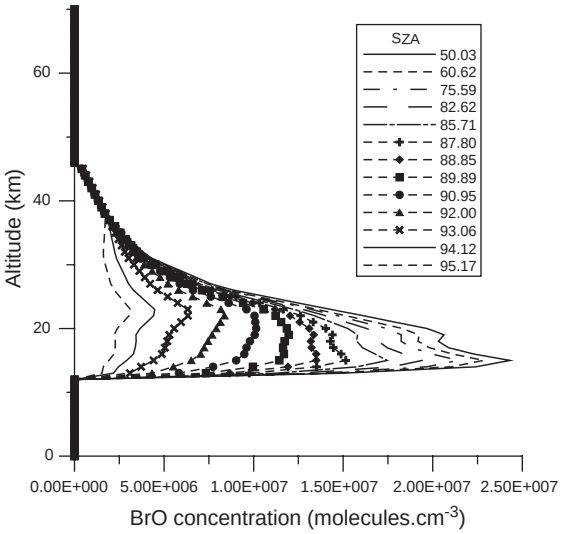

Fig. 2. The diurnal variation of the BrO profile in the stratosphere at a range of SZAs calculated for the sunset on day 254, 2001 at Lauder, New Zealand using a stationary Lagrangian box model.

the problems associated with assuming the diurnal variation a priori and does allow the chemical assumptions to be tested.

5. Case study: BrO

The retrieval method outlined above can be implemented to retrieve profile information by combining any number of measurement platforms targeting any species that is measured by a spectroscopic technique. The technique of retrieving a set of profiles at different diurnal stages allows the retrieval of difficult diurnally varying species. To demonstrate how this method works, a retrieval has been performed for BrO utilizing direct-sun and zenith-sky measurements made for the sunset on day 254, 2001 at Lauder (45°S, 170°E).

Direct-sun measurements spanning an SZA range of 63°–87° are combined with zenith-sky viewing measurements spanning 62°–92° SZA to construct the measurement vector $\mathbf{y}$. The covariance matrix of the measurements $\mathbf{S}_\varepsilon$ is constructed from the DOAS fitting error. All errors in the determination of the DSCDs using the DOAS fitting of the radiance measurements are included in the construction of the $\mathbf{S}_\varepsilon$ matrix. These include errors from the Ring effect, polarization, Rayleigh scattering, and Mie scattering. Errors due to the ozone, NO_2 and the BrO absorption cross-sections and their temperature dependencies are also included in the construction of $\mathbf{S}_\varepsilon$. $\mathbf{S}_\varepsilon$ is assumed to be diagonal in this work.

The a priori of the state vector was constructed using the model diurnal variation displayed in Fig. 2 for day 254, 2001 over Lauder. Only six number density profiles were chosen to find a balance between adequately describing the diurnal variation and minimizing the number of state components. As the model diurnal variation only describes the stratosphere, a tropospheric component

is added as suggested from mid-latitude balloon flight observations [29]. A constant number density of 5×10^6 molecules cm^{-3} is assumed for the troposphere for the 75° profile. This constant tropospheric number density is reduced to 3×10^6 and to 2×10^6 molecules cm^{-3} for the 84° and 87° profiles, respectively.

The a priori covariance matrix was constructed by taking 50% of the profile peak as the error at all levels of each a priori profile. This error was found to fit the measurements well without over-fitting them. As $\mathbf{S}_a$ is being used as a tuning parameter the retrieval is not strictly 'optimal'. To be optimal, the true atmospheric variability, or just the prior uncertainty (reflecting the lack of knowledge) would need to be derived from an ensemble of state profiles.

Ozonesonde data for Lauder closest to day 254, 2001 was used to provide the forward model parameters of temperature, pressure and ozone number density [21]. The NO_2 profile was provided from SAGE II for Lauder [22]. The aerosol extinction profile for Lauder was provided by LIDAR and backscatter data [J.B. Liley, Personal Communication, 2003].

The weighting functions are calculated by perturbing each level of each of the six a priori BrO profiles independently. The resultant change in the calculated DSCDs is divided by the initial perturbation to give a column of **K**. The weighting functions for perturbing the state vector components for the 75°, 84° and 87° SZA profiles are displayed in Fig. 3. The resultant weighting functions for the direct-sun and zenith-sky measurements are provided separately. Note that the two measurement modes have very different sensitivities to the three diurnal BrO profiles.

Perturbing the 75° SZA state vector profile leads to a negative sensitivity for high SZA measurements in both the direct-sun and zenith-sky cases. This is due to the fact that the measurements are differential with respect to the lowest measurement of each mode. Reference SZAs of 62° and 63° are used for the zenith-sky and direct-sun cases, respectively, in this case study. For example, the resultant change in the 86.13° zenith-sky SCD to a perturbation at any altitude of the 75° state profile is less than the resultant change in the reference SCD for 62.13°.

The zenith-sky DSCDs display the highest sensitivities to perturbations in the stratosphere for all three state profiles. In contrast, the direct-sun weighting functions display a tropospheric sensitivity. The interesting shape of the direct-sun weighting functions for the 84° profile perturbations nicely illustrates where along a given measurement path the perturbation has the most influence. It is clear that the measurement sensitivities are very different for each of the three state vector profiles. To retrieve only one of these with an a priori assumed diurnal variation leads to complex and unrealistic features in the weighting functions. These features propagate as errors into the retrieved quantities.

The averaging kernels produced by this retrieval method are not only a function of altitude but also of profile or time (SZA). Each element of the state vector has an averaging kernel that is two dimensional. So while it is possible to view the averaging kernel in a traditional altitude sense, their variation in the space of profiles (time sense) should also be considered, thus they are best seen as two dimensional. Contour plots are used to view the averaging kernels and derived quantities such as resolution and area. While contours give the impression of continuity, the values are discrete in both space and time.

The retrieval is conducted with a state vector that is defined on a 1 km grid. The information content and degrees of freedom for each of the state vector profiles is displayed in Table 1. The measurements produce two pieces of independent information in each of the 75°,84° and 87° BrO profiles. In each of these profiles the factor by which the entropy has been reduced due to making

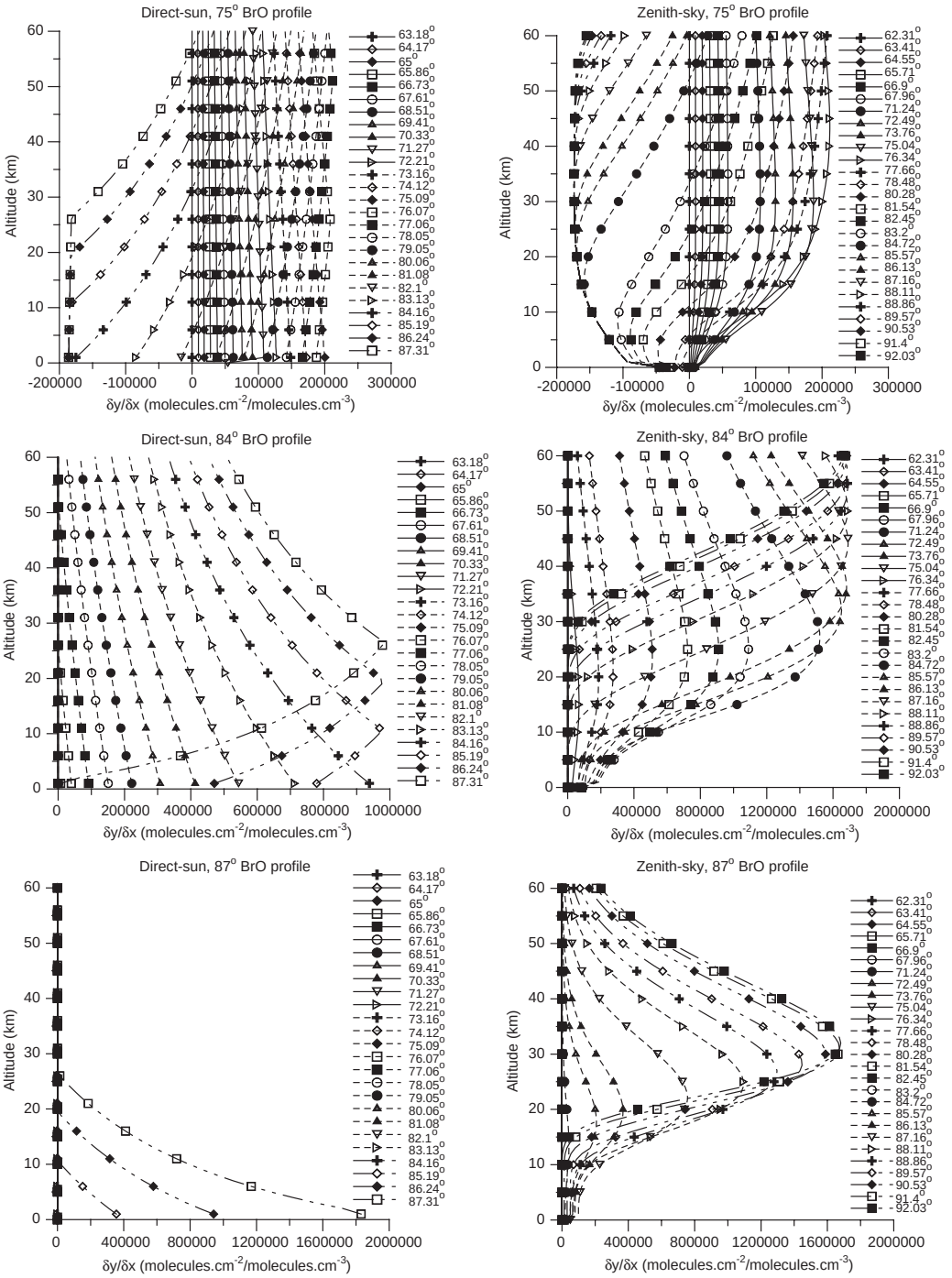

Fig. 3. Weighting functions for direct-sun (left panels) and zenith-sky (right panels) resulting from perturbations in the 75° SZA state vector profile (top). The measurement response to perturbations of the 84° SZA state vector profile is displayed in the middle and 87° SZA state vector profile at the bottom. It can be clearly seen that the measurements, both comparing between the measurement modes and within one measurement mode, have very different sensitivities for the three diurnal BrO profiles.

Table 1
Information content and degrees of freedom in each of the BrO profiles and the total of these quantities in the retrieval. With two degrees of freedom (independent pieces of information from the measurements) in each of the 75°, 84° and 87° BrO profiles, only two quantities should be derived from the retrieval for each profile

BrO profile	0°	75°	84°	87°	92°	95°	Total
Information content (bits)	0.03	2.9	4.0	2.3	0.02	0.0	13.5
Degrees of freedom	0.05	2.1	2.6	2.3	0.04	0.0	7.1

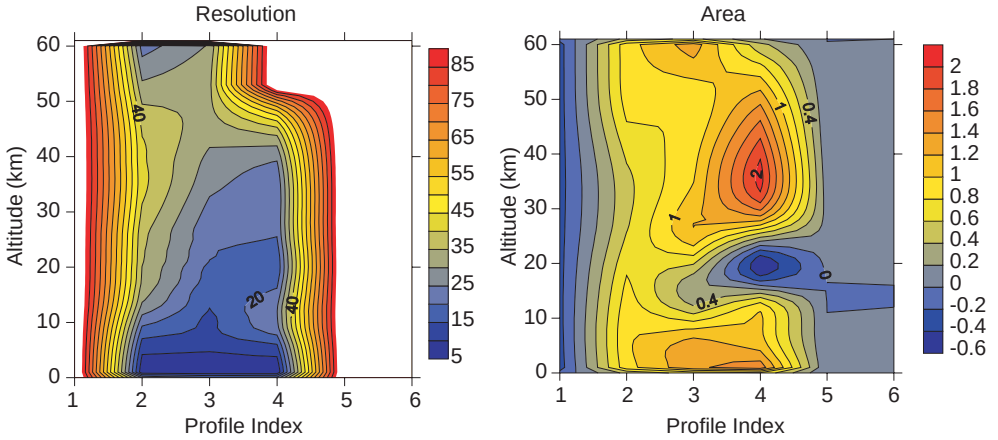

Fig. 4. Resolution and area of the averaging kernels in altitude space (y-axis in km) and profile space (x-axis) for the BrO profiles. The profile index refers to the solar zenith angle corresponding to the profile: 1 = 0°, 2 = 75°, 3 = 84°, 4 = 87°, 5 = 92° and 6 = 95°. The left panel displays the resolution as the number of state vector elements required per degree of freedom. In the troposphere for the state vector profiles 75°, 84° and 87° the resolution is 5–15 km and in the stratosphere it is 40 km for the 75° profile and 20 km for the 84° and 87° profiles. The right panel displays the area of the averaging kernels, and is a qualitative measure of how much information on the broad structure is being provided from the measurements rather than the a priori. For the state vector profiles 75° and 84° the retrieved quantities have come from the measurements. The 87° profile illustrates a heavy reliance on the a priori data at 20 km due to the fact that the zenith-sky measurement sensitivity is in the higher stratosphere (above 25 km) and the direct-sun sensitivity is close to the surface.

the measurement is between five and 16. Information content and the degrees of freedom are relative to the a priori covariance that has been chosen for the retrieval.

The resolution and area of the averaging kernels of the retrieval are displayed as two-dimensional contour plots in altitude and profile space (Fig. 4). Where the area of the averaging kernel is negative, and where it is much larger than one, the retrieved values are no longer a good representation of the true atmosphere. This occurs in the 87° profile case at ~20 km due to the fact that the zenith-sky measurement sensitivity is higher in the stratosphere. The direct-sun measurements are most sensitive near the ground when sampling the 87° profile.

The area, resolution, information content and degrees of freedom collectively indicate that the usefully resolved retrieval product is the tropospheric and stratospheric column amount for the 75°,

84° and 87° state vector profiles. A matrix $\mathbf{g}$ is defined to evaluate the stratospheric and tropospheric columns from the retrieved state ($\hat{\mathbf{x}}$). $\mathbf{g}$ sums over the relevant altitudes in the state vector to give tropospheric and stratospheric columns for each of the six state profiles ($\hat{\mathbf{x}}_c$):

$$\hat{\mathbf{x}}_c = \mathbf{g}\hat{\mathbf{x}}. \tag{8}$$

The averaging kernels for the tropospheric and stratospheric columns for each of the six state profiles are evaluated using the following equation:

$$\mathbf{A}_c = \mathbf{g}\mathbf{A}. \tag{9}$$

The averaging kernels for the tropospheric and stratospheric columns for the 75°, 84° and 87° BrO profiles are displayed in Fig. 5. These display how the retrieved column quantities clearly separate the troposphere and stratosphere for the BrO profiles of 75°, 84° and 87°.

The averaging kernels demonstrate that good separation between the tropospheric and stratospheric columns is possible for the state vector profiles of 75°, 84° and 87°. These averaging kernels show how the retrieved column quantities are related to the true atmospheric state. The averaging kernel for the stratospheric column quantity of the 87° BrO profile shows this quantity to be the poorest representation of the true atmosphere. Thus, the retrieval error for the 87° stratospheric column is expected to be higher than the other retrieved stratospheric columns, as displayed in Table 2.

The covariance matrices for the retrieved column quantities are evaluated in a similar fashion to the averaging kernels i.e.:

$$\mathbf{S}_c = \mathbf{g}\mathbf{S}\mathbf{g}^T. \tag{10}$$

Table 2 displays the values for the a priori and retrieved tropospheric and stratospheric columns along with the error breakdown for the retrieved quantities of 75°, 84° and 87°.

The error contribution from the forward model parameters to the final retrieval error is important in the assessment of which parameters should be retrieved rather than assumed to be known to a high enough precision a priori. The contribution of errors arising from the variability of temperature, ozone and aerosol profiles are negligible in the total retrieval error. The covariance matrices for the forward model parameters used in the $\mathbf{S}_f$ calculation are derived from the variability of these measured quantities over Lauder.

The smoothing error is the largest and least well-estimated source of error in the retrieved quantities. It arises from the difference between the idealised column integral and the averaging kernel, applied to fine structure in the true profile as represented statistically by the a priori covariance $\mathbf{S}_a$. The use of $\mathbf{S}_a$ as a tuning parameter means that the smoothing error is not necessarily meaningful, and is likely to be an overestimate with the formulation used here, because the diagonal covariance describes relatively large amplitude of fine structure. The real state will probably have longer scale correlations, and less fine structure.

The error reduction from the a priori to the retrieved value demonstrates how much the entropy has been reduced by making the measurement. Qualitatively this is the information content of the measurement.

The fit of the modelled DSCDs to the measured DSCDs is essential for the retrieved quantities to represent the measurements. Fig. 6 displays the modelled fit of the measured DSCDs with the residuals. A good fit is achieved with fairly low residual.

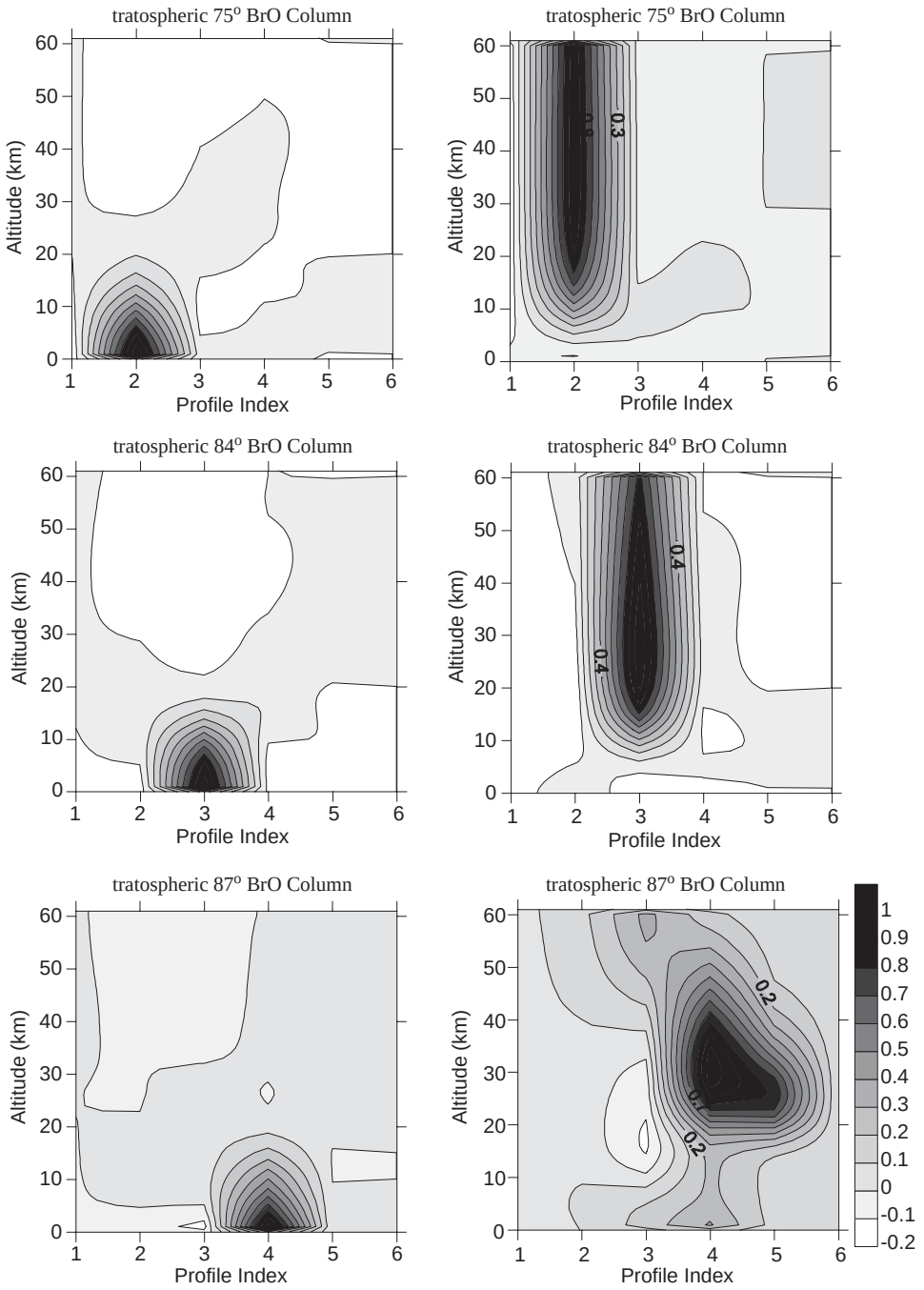

Fig. 5. Averaging kernels in altitude space (y-axis in km) and profile space (x-axis) for the BrO profiles. The profile index refers to the solar zenith angle corresponding to the profile: $1=0^\circ$, $2=75^\circ$, $3=84^\circ$, $4=87^\circ$, $5=92^\circ$ and $6=95^\circ$. The left panels display the averaging kernels for the tropospheric columns and the right panels the averaging kernels for the stratospheric columns. The top panels display the averaging kernels for the 75° profile, the middle panels the 84° profile and the lower panels the 87° profile. Averaging kernels show how the retrieved column quantities are related to the true atmospheric state.

Table 2
The retrieved and a priori columns for the troposphere and stratosphere are displayed along with the retrieval error. The error is then broken down into the error contribution from retrieval noise, smoothing and forward model parameter errors. The forward model parameters investigated are temperature, pressure, ozone and aerosol

Profile SZA(deg)	75	84	87
Tropospheric column a priori value ($\times 10^{12}$ molecules cm^{-2})	5.3 ± 3.0	3.8 ± 2.7	2.4 ± 2.2
Tropospheric column retrieved value ($\times 10^{12}$ molecules cm^{-2})	**2.7** ± 1.5	**1.5** ± 1.1	**1.7** ± 1.1
Total retrieval error	1.5	1.1	1.1
Noise error	0.9	0.6	0.7
Smoothing error	1.2	0.9	0.9
Temperature error	0.3	0.03	0.02
Pressure error	0.03	0.03	0.02
Ozone error	0.03	0.008	0.001
Aerosol error	0.02	0.005	0.003
Stratospheric column a priori value ($\times 10^{12}$ molecules cm^{-2})	2.25 ± 0.73	2.02 ± 0.64	1.72 ± 0.52
Stratospheric column retrieved value ($\times 10^{12}$ molecules cm^{-2})	**2.38** ± 0.18	**2.20** ± 0.16	**1.76** ± 0.37
Total retrieval error	0.18	0.16	0.37
Noise error	0.07	0.08	0.09
Smoothing error	0.17	0.14	0.36
Temperature error	0.004	0.01	0.008
Pressure error	0.003	0.005	0.005
Ozone error	0.001	0.002	0.001
Aerosol error	0.003	0.002	0.001

The weighting functions and averaging kernels display how the direct-sun and zenith-sky geometries complement each other with different tropospheric and stratospheric sensitivity. The information content and degrees of freedom for the signal along with the area of the averaging kernels and resolution characterise the retrieval. These quantities demonstrate that while it is possible to retrieve a large number of elements, the final retrieved product must take these into consideration. Only six independent variables are presented as the final retrieval product. The averaging kernels of these final retrieved quantities show there is good tropospheric and stratospheric separation for the three BrO profiles of 75°, 84° and 87°.

6. Conclusions

This paper presents the use of ground-based spectroscopic measurements to provide the altitude distribution of time-varying atmospheric absorbers. A case study of BrO demonstrates the details and the potential of this retrieval technique.

The use of two different sampling techniques provides information about a state that would not be accessed by only considering each technique independently. Combining complementary

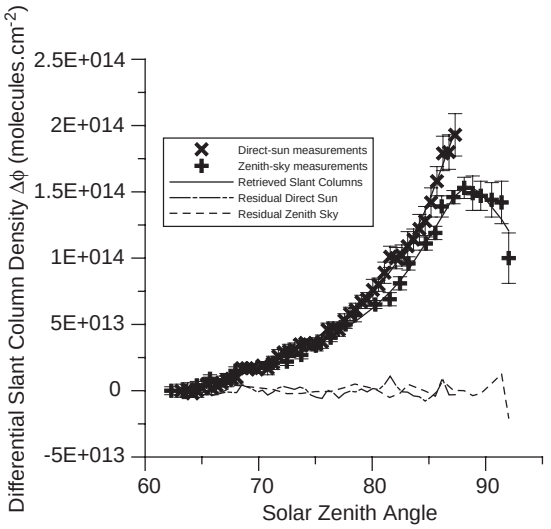

Fig. 6. The measured and retrieved (modelled) DSCDs and the residuals for the direct-sun and zenith-sky sunset measurements made at Lauder (45°S, 170°E) on day 254, 2001.

measurements facilitates the tropospheric and stratospheric column separation in this retrieval. The direct-sun viewing geometry is simple and provides good total column and tropospheric sensitivity. The zenith-sky viewing geometry, while measuring diffuse light, leading to slightly more complex modelling, contains good stratospheric information. With a number of possible platforms for spectroscopic measurements, there is potential for combining a large ensemble of complementary measurements.

A two-dimensional retrieval results in the characterisation and error analysis becoming more complex, but also more interesting. Sets of profiles defined on a time (or SZA) grid, to describe the diurnal variation of the species, are retrieved. Treatment of the diurnal profiles as retrieval parameters eliminates the complications that would arise if the diurnal variation would be treated as a forward model parameter. If the diurnal variation were treated as a forward model parameter discontinuities in the weighting functions would propagate as errors into the final retrieved state. Retrieving a set of profiles that describe the diurnal variation is advantageous as the measurements can provide information on the diurnal variation of the atmosphere.

The general algorithm for retrieving profile and chemical information by combining measurement modes is applied in a case study of BrO. Direct-sun and zenith-sky DSCDs measured at Lauder (45°S, 170°E) on day 254, 2001 are investigated for their profile and diurnal information. Characterisation of the retrieval is performed using averaging kernels, resolution, information content and degrees of freedom for signal. The retrieval characterisation demonstrates that tropospheric and stratospheric separation is possible for BrO profiles at 75°, 84° and 87° SZA. The stratospheric column is seen to decrease from 2.38 to 2.20 to 1.76×10^{13} molecules cm^{-2}, respectively, as expected. The tropospheric column is seen to be $\sim 2.0 \times 10^{12}$ molecules cm^{-2}, which is smaller than previous estimates. The analysis of a more comprehensive measurement set is the focus of future work. The error analysis showed that the forward model parameters of temperature, pressure, ozone and aerosol all produce negligible error effects in the final retrieved column amounts.

Acknowledgements

The authors thank J.B. Liley for kindly supplying the aerosol data used in the radiative transfer calculations. RS would like to thank the New Zealand Foundation for Research in Science and Technology Top Achiever Doctoral fellowship scheme for funding this research. This work was also funded in part by the Visiting Scientist programme of the National Institute of Water and Atmospheric Research, and in part by the Foundation for Research in Science and Technology, programme C01X0035. One of us (CDR) would like to thank the staff of NIWA, Lauder, for their hospitality during the extended visits when this work was carried out.

References

[1] Brewer AW, McElroy CT, Kerr JB. Nitrogen dioxide concentration in the atmosphere. Nature 1973;246:129–33.
[2] Noxon JF. Nitrogen dioxide in the stratosphere and troposphere measured by ground-based absorption spectroscopy. Science 1975;189:547–9.
[3] McKenzie R, Johnston PV, McElroy CT, Kerr JB, Solomon S. Altitude distributions of stratospheric constituents from ground-based measurements at twilight. J Geophys Res 1991;96:15499–511.
[4] Preston KE, Jones RL, Roscoe HK. Retrieval of NO_2 vertical profiles from ground-based UV-Visible measurements—method and validation. J Geophys Res 1997;102:19089–97.
[5] Rodgers CD. Inverse methods for atmospheric sounding, theory and practice. In: Taylor FW, editors. Series on Atmospheric, Oceanic and Planetary Physics, Singapore: World Scientific; 2000.
[6] Strong K, Joseph BM, Dosanjh R, McDade IC, McLinden CA, McConnell JC, Stegman J, Murtagh DP, Llewellyn EJ. Retrieval of vertical concentration profiles from OSIRIS UV-visible limb spectra. Can J Phys 2002;80:409–34.
[7] Livesey JJ, Read WG. Direct retrieval of line-of-sight atmospheric structure from limb sounding observations. Geophys Res Lett 2000;27:891–4.
[8] Kemnitzer H, Hilgers S, Schwarz G, Steck T, Clarmann TV, Hopfner M, Ressel K. Trace gas retrieval including horizontal gradients. Adv Space Res 2002;29:1631–6.
[9] Jiang Y, Yung YL, Sander SP. Detection of tropospheric ozone by remote sensing from the ground. JQSRT 1997;57:811–8.
[10] Ferlemann F, Camypeyret C, Fitzenberger R, Harder H, Hawat T, Osterkamp H, Schneider M, Perner D, Platt U, Vradelis P, Pfeilsticker K. Stratospheric BrO profiles measured at different latitudes and seasons—Instrument description, spectral analysis and profile retrieval. Geophys Res Lett 1998;25:3847–50.
[11] Solomon S, Schmeltekopf AL, Sanders RW. On the interpretation of zenith sky absorption measurements. J Geophys Res 1987;92:8311–9.
[12] Sanders RW, Solomon S, Smith JP, Perliski L, Miller HL, Mount GH, Keys JG, Schmeltekopf AL. Visible and near-ultraviolet spectroscopy at McMurdo station, Antarctica 9. Observations of OClO from April to October 1991. J Geophys Res 1993;98:7219–28.
[13] Burrows JP, Weber M, Buchwitz M, Rozanov V, Ladstätter-Weienmayer A, Richter A, DeBeek R, Hoogen R, Bramstedt K, Eichmann K-U, Eisinger M. The Global Ozone Monitoring Experiment (GOME): mission concept and first scientific results. J Atmos Sci 1999;56:151–75.
[14] Noel S, Bovensmann H, Burrows JP, Frerick J, Chance KV, Goede AHP. Global atmospheric monitoring with SCIAMACHY. Phys Chem Earth Pt C 1999;24:427–34.
[15] Platt U. Differential Optical Absorbtion Spectroscopy (DOAS). In: Sigrist MW, editor. Air monitoring by spectroscopic techniques. New York: Wiley; 1994. p. 27–76.
[16] Aliwell SR, Roozendael MV, Johnston PV, Richter A, Wagner T, Arlander DW, Burrows JP, Fish DJ, Jones RL, Tornkvist KK, Lambert J-C, Pfeilsticker K, Pundt I. Analysis for BrO in zenith-sky spectra: an intercomparison exercise for analysis improvement. J Geophys Res 2002;107:ACH10-1–10-20.
[17] Hofmann D, et al. Intercomparison of UV Visible spectrometers for measurements of stratospheric NO_2 for the Network for the Detection of Stratospheric Change. J Geophys Res 1995;100:16765–91.

[18] Bucholtz A. Rayleigh-scattering calculations for the terrestrial atmosphere. Appl Opt 1995;34:2765–73.
[19] Kneizys FX, Chetwynd JHJ, Clough SA, Shettle EP, Abreu LW, Fenn RW, Gallery WO, Selby JEA. Atmospheric transmittance/radiance: computer code Lowtran 6. Massachusetts. Air force Geophysics Laboratory; 1983. p. 11–16.
[20] Toublanc C. Henyey–Greenstein and Mie phase functions in Monte Carlo radiative transfer computations. Appl Opt 1996;35:3270–4.
[21] Bodeker GE, Boyd IS, Matthews WA. Trends and variability in vertical ozone and temperature profiles measured by ozonesondes at Lauder, New Zealand: 1986–1996. J Geophys Res 1998;103:28661–81.
[22] (NASA-LaRC), Sage II NO_2 profiles were supplied by the NASA Langley Research Center (NASA-LaRC) and the NASA Langley Radiation and Aerosols Branch.
[23] Sarkissian A, Roscoe NK, Fish DJ. Ozone measurements by zenith-sky spectrometers—an evaluation of errors in air-mass factors calculated by radiative transfer models. JQSRT 1995;54:471–80.
[24] Perliski LM, Solomon S. On the evaluation of air mass factors for atmospheric near-ultraviolet and visible absorption spectroscopy. J Geophys Res 1993;98:10363–74.
[25] Sinnhuber BM et al. Comparison of measurements and model calculations of stratospheric bromine monoxide. J Geophys Res 2002;107:ACH11-1–11-18.
[26] Van Roozendael M, Wagner T, Richter A, Pundt I, Arlander DW, Burrows JP, Chipperfield M, Fayt C, Johnston PV, Lambert J-C, Kreher K, Pfeilsticker K, Platt U, Pommereau J-P, Sinnhuber BM, Tornkvist KK, Wittrock F. Intercomparison of BrO measurements from ERS-2 GOME, ground-based and balloon platforms. Adv Space Res 2002;29:1661–6.
[27] Richter A, Wittrock F, Ladstätter-Weienmayer A, Burrows JP. GOME measurements of stratospheric and tropospheric BrO. Adv Space Res 2002;29:1667–72.
[28] Chipperfield MP. Multiannual simulations with a three-dimensional chemical transport model. J Geophys Res 1999;104:1781–805.
[29] Fitzenberger R, Bosch H, Camy-Peyret C, Chipperfield MP, Harder H, Platt U, Sinnhuber BM, Wagner T, Pfeilsticker K. First profile measurements of tropospheric BrO. Geophys Res Lett 2000;27:2921–4.