WATER VAPOUR RETRIEVAL FROM GOME DATA

S. Noël, M. Buchwitz, H. Bovensmann, R. Hoogen, and J. P. Burrows

Institute of Environmental Physics, University of Bremen, FB 1, P.O. Box 333440, D-28334 Bremen, Germany
phone: +49 421 218 4081, fax: +49 421 218 4555
e-mail: Stefan.Noel@iup.physik.uni-bremen.de

ABSTRACT

A method to derive the atmospheric vertical column of water vapour from the near-infrared measurements of the Global Ozone Monitoring Experiment (GOME) is presented. These columns are determined by a modified DOAS approach, taking into account effects of strong wavelength dependent absorption. The radiative transfer model MODTRAN is used to calculate appropriate reference spectra for different scenarios, i.e. different solar zenith angles (SZAs) and different model atmospheres. This paper shows the feasibility of this approach and gives first estimates on the retrieval precision and the accuracy of a possible new GOME H₂O data product by comparison with selected radiosonde and SSM/I data.

1. INTRODUCTION

The accurate assessment of the impact of present and future anthropogenic pollution and natural phenomena of the atmosphere and the climate-chemistry coupling requires a detailed global knowledge of the temporal and spatial behaviour of atmospheric trace constituents (gases, aerosols, clouds). Water vapour is one of the most abundant atmospheric gases. More than 99% of water vapour is located in the troposphere where it significantly contributes to atmospheric chemistry and of course weather and climate. For example, H₂O plays a major role in the production of OH, the most important oxidising agent. Water vapour is arguably the most variable greenhouse gas. Variations of water vapour are closely related to atmospheric pressure and temperature; thus water vapour may be considered as a tracer for the spatial and temporal variability of the troposphere and particularly of climatic changes. The amount of water vapour is therefore of great importance for global atmospheric models which aim to predict the future climate.

Currently, global water vapour concentrations are mainly estimated from satellite based measurements in the infrared – e.g. with the TIROS-N Operational Vertical Sounder (TOVS) – and in the microwave spectral region – e.g. with the Special Sensor Microwave Imager (SSM/I) at 22 and 37 GHz – in combination with in-situ radiosonde measurements (see Ref. 1, and references therein).

This paper describes a method to derive water vapour column amounts from measurements of the Global Ozone Monitoring Experiment (GOME) which is operating successfully from the ERS-2 satellite since 1995. The GOME instrument is a grating spectrometer which measures solar irradiance and earthshine radiance from the UV to the visible and near-infrared wavelength region (≈ 240-800 nm). The inversion of the ratio of measured irradiances and attenuations provides information about the amount and distribution of atmospheric constituents. The main task of GOME is the determination of global distributions of O₃ and NO₂, but it has been shown that concentrations of much more atmospheric constituents can be derived as well, e.g. BrO, OCIO, SO₂, HCHO (see Ref. 2).

The method described in this paper uses GOME measurements in the near-infrared wavelength region (around 720 nm) to retrieve water vapour total vertical column amounts. In contrast to microwave measurements this approach is not limited to ocean areas. As with IR measurements it is possible to determine H₂O concentrations also over land. However, the determination of total H₂O columns is limited by the comparably low spatial resolution of GOME of about 320×40 km. So far, the retrieval algorithm is applied only to (almost) cloud-free scenarios, but in principle water vapour columns above the cloud top can be derived for cloudy pixels.

H₂O columns from GOME measurements are of special interest for users of the other GOME data products. Moreover, water vapour data from GOME measurements complements well other ERS-2 data products. Combination of GOME water vapour with e.g. cloud detection by ATS&R and liquid water content from the Microwave Sounder will increase our knowledge on the global hydrological cycle in addition to demonstrating the pre-operational potential for meteorological applications.

As the retrieval method described here does not explicitly rely on GOME data, it may be applied to data from other similar sensors, such as the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) (see e.g. Refs. 3, 4). SCIAMACHY is an enhanced version of GOME, which will operate from the ENVISAT-1 platform to be launched in late 2000.

2. RETRIEVAL METHOD

The retrieval of vertical H₂O columns is based on the Differential Optical Absorption Spectroscopy (DOAS) approach which was originally developed for on-ground observations (see e.g. Ref. 5) but has proven to be applicable also for ground-based zenith and space-based nadir measurements (Ref. 6).

2.1. Standard DOAS

DOAS exploits the Beer-Lambert law:

\[ I = I_0 \exp(-\tau) \]

where \( I \) and \( I_0 \) are the measured and the unattenuated intensity, and \( \tau \) is the optical depth along the light path.
The main idea of DOAS is to separate components which vary slowly with wavelength (resulting from Rayleigh and Mie scattering, surface reflection, and broadband absorption) from differential absorption structures. The broadband spectral features are approximated by a low-order polynomial $P$, which leads to

$$\ln \left( \frac{I}{I_0} \right) = P - \tau_A$$

where $\tau_A$ is the slant optical depth due to differential absorption.

For one absorber in the selected wavelength interval, which is the case for the H$_2$O band used in this study, the slant optical depth is given by

$$\tau_A = \int_a \sigma_A n \, ds$$

where $\sigma_A$ is the differential absorption cross section and $n$ is the number density of the atmospheric species; the integration is to be performed along the light path $s$.

If $\sigma_A$ is independent of atmospheric pressure and temperature, the slant optical depth is proportional to the amount of the absorbing species integrated along the light path:

$$\tau_A = \sigma_A C_S$$

where

$$C_S := \int_a n \, ds$$

is the slant column density which is related to the vertical column density $C_V$ by the so-called air mass factor:

$$m := \frac{C_S}{C_V}$$

$C_V$ is defined similar to $C_S$:

$$C_V := \int_0^{\text{TOA}} n \, dz$$

where $n$ is integrated in the vertical direction, $z$, from the ground (altitude 0) to the top of the atmosphere (TOA).

Combining the above yields

$$\ln \left( \frac{I}{I_0} \right) = P - \sigma_A m C_V$$

Using measured values of $I$ and $I_0$, known absorption cross sections $\sigma$, and an air mass factor $m$ (usually derived from radiative transfer calculations), a linear fit provides the polynomial coefficients and the vertical column density $C_V$.

If the differential absorption cross sections depend on pressure and temperature (which is the case for H$_2$O), $\sigma_A$ is not constant, and equation (3) no longer holds. However, it is possible to use a slightly different approach for this case which uses the slant optical depth $\tau_A$ instead of the slant column density $C_S$ in the fitting process (see e.g. Ref. 6).

### 2.2. Modified Approach

The main assumption of “standard” DOAS is that the relative depth of an absorption feature is linearly proportional to the amount of the absorbing species integrated along the light path. This does not hold for molecules like H$_2$O where differential absorption features strongly depend on wavelength and are not resolved by the measuring instrument. In this case, the Beer-Lambert law is not applicable. The differential absorption depth becomes a nonlinear function of the absorber amount, and equation (4) is no longer valid unless at spectral regions where absorption is low, e.g. at the outer edges of a band structure. Unfortunately, this is also the region where absorption cross sections are typically less accurately known.

To make use of the usually much better signal-to-noise ratio of spectral regions with strong absorption, which may enable us to additionally derive H$_2$O vertical profile information out of the GOME data, it is necessary to include the saturation effect in the calculations. This is done by analogy to an approach described by Halthore & al. (Ref. 7) who determined water vapour columns from ground-based occultation measurements using sun photometers operating in the 940-nm band. Although the instrumental setup of Halthore & al. differs from the GOME observational geometry, spectral resolution and wavelength range, it will be shown that the general parametrisation of the saturation effect is applicable to GOME data.

Instead of equation (4) the following relationship between the slant optical depth $\tau_A$ and the slant column density $C_S$ is assumed:

$$\tau_A = a C_S^b$$

The parameters $a$ and $b$ depend on wavelength, observational geometry (mainly solar zenith angle), spectral resolution and atmospheric properties, in particular on the vertical distribution of water vapour, i.e. the shape of the profile. It is a main assumption of the method described here that $a$ and $b$ do not depend on the actual amount of water vapour in the atmosphere, i.e. the total vertical column density. The parameter $a$ contains the differential absorption cross sections (for different temperatures/pressures, averaged along the light path) at instrument resolution, whereas $b$ is directly related to saturation where usually $b \leq 1$. For their broadband photometer measurements Halthore & al. determined single values for $a$ and $b$ which were quite insensitive to atmospheric changes. To use the higher spectral resolution of GOME data $a$ and $b$ are assumed to have a spectral dependence.

If we define

$$c := a m^b$$

the slant optical depth may also be written as a function of the vertical column density:

$$\tau_A = c C_V^b$$

This leads to the following basic equation which replaces (4):

$$\ln \left( \frac{I}{I_0} \right) = P - c C_V^b$$
The parameters \( b \) and \( c \) are obtained from radiative transfer calculations for different model atmospheres and observational geometries (i.e. solar zenith angles). This is explained below. For given values of \( b \) and \( c \), the vertical column amount \( C_V \) may then be determined by a non-linear fit similar to the standard DOAS approach.

2.3. Determination of Parameters \( b \) and \( c \)

The first step in the determination of \( b \) and \( c \) is the derivation of the slant optical depth. For a given scenario, the slant optical depth is calculated by performing two radiative transfer calculations: one for an atmosphere which contains the relevant absorbing species (in our case water vapour), and one for an atmosphere which does not contain this species, i.e. for a zero column density. Using equation (1) the corresponding simulated intensities (i.e. radiiances) \( I_{\text{with}} \) and \( I_{\text{without}} \) are:

\[
\ln \left( \frac{I_{\text{with}}}{I_0} \right) = P - \tau_A
\]

and

\[
\ln \left( \frac{I_{\text{without}}}{I_0} \right) = P
\]

which leads to

\[
\tau_A = \ln \left( \frac{I_{\text{without}}}{I_{\text{with}}} \right)
\]  

(9)

This implicitly assumes that the polynomial \( P \) does not depend on the vertical column amount. This appears to be a good approximation in our case.

Taking the logarithm of equation (7) a linear relation between the logarithms of vertical column density and slant optical depth is obtained:

\[
\ln \tau_A = \ln c + b \ln C_V
\]  

(10)

Varying the vertical column density and computing the corresponding slant optical depths yields \( b \) and \( c \) for each wavelength by a linear fit. To retrieve parameters which are specific for a given model atmosphere, the shape of the water vapour profile is kept constant in this process. This is achieved by scaling the assumed \( \text{H}_2\text{O} \) profile of the model atmosphere by factors from 20\% to 100\% without changing the other atmospheric parameters. It is not useful to apply scaling factors larger than 1 because this may result in a relative humidity larger than 100\% at certain altitudes. On the other hand, taking into account only atmospheric columns which are smaller than the reference column may cause an under-estimation of the retrieved columns. This is (at least partly) compensated by adding more weight to the 100\% atmosphere in the fitting process.

In this study, all radiative transfer calculations are performed with MODTRAN 3.7 (Ref. 8) for the wavelength range between 710 and 740 nm where water vapour is the dominant absorber. However, the method has also been successfully applied to other spectral regions where \( \text{H}_2\text{O} \) absorbs. The simulated spectra are computed at the highest spectral resolution of MODTRAN (1 cm\(^{-1}\), equivalent to about 0.05 nm in this wavelength region) and then convoluted to the GOME spectral resolution (about 0.345 nm in this spectral range) before further processing.

Parameter sets of \( b \) and \( c \) are computed for solar zenith angles between 0\° and 90\° and each of the six MODTRAN standard atmospheres: tropical (TRO), mid-latitude summer (MLS), mid-latitude winter (MLW), sub-arctic summer (SAS), sub-arctic winter (SAW), and 1976 US Standard (STD) atmosphere. MODTRAN is run in multiple scattering mode in nadir viewing geometry assuming no clouds/precipitation, and tropospheric/background stratospheric aerosols. A surface albedo of 0.05 is assumed for ocean scenarios and 0.3 for land.

As an example, Figure 1 shows the slant optical depths computed with MODTRAN assuming a solar zenith angle of 40\°, a surface albedo of 0.05, and a tropical background atmosphere (TRO) with water vapour profiles scaled by the indicated factors. As can be seen from the resulting spectra for \( b \) and \( c \) which are displayed in Figure 2, values of low \( c \), equivalent to low absorption, correspond to values of \( b \) close to 1, equivalent to few saturation. At these regions the current approach is equivalent to standard DOAS. In fact, both methods produce similar results there.

![Figure 1. Slant optical depths computed with MODTRAN for a solar zenith angle of 40°, a surface albedo of 0.05 (equivalent to ocean), and a tropical model atmosphere in which the water vapour profile has been scaled by (from bottom to top) 20\%, 40\%, 60\%, 80\%, and 100\%.](image)

2.4. Selection of Parameter Set

As described above, the parameter sets \( b \) and \( c \) depend on model atmosphere, solar zenith angle, and surface albedo. Thus for each measured spectrum it has to be determined which parameter set will be used for the retrieval.

The surface albedo is chosen with respect to the geographical location of the measurement. Presently, only two different albedo values are used: 0.05 for ocean scenarios and 0.3 as an average value for measurements above land.

Since the solar zenith angle is known, an appropriate parameter set may be selected in one of the following ways:
The third possibility is the most accurate one, but it is very time-consuming because it requires individual radiative transfer calculations for each measured spectrum. The two other possibilities rely on tables of pre-calculated parameter sets and are therefore very useful when applied to a large data set (like the GOME data). Therefore, the present study uses method (2). Since the solar zenith angle dependence of \( b \) and \( c \) has turned out to be rather smooth for SZAs < 80° (higher solar zenith angles are currently not covered) it is sufficient to pre-calculate \( b \) and \( c \) for the solar zenith angles 0°, 20°, 40°, 50°, 60°, 70°, and 80°. For intermediate SZA values interpolation is performed.

The selection of an appropriate model atmosphere is difficult and requires in general a-priori information, e.g. a climatological data base. In our case, this data base is provided by the six MODTRAN atmospheres. Knowing the time and geolocation of a measurement, a parameter set which has been calculated for a suitable scenario (e.g. use the MLS atmosphere for measurements in mid-latitudes during summer) could be selected. However, this method relies on the quality of the climatological data base. This is why a different approach is used here: Retrievals are performed for each of the six atmospheres and then the result which has the smallest residual, i.e. the one which reproduces the measurements best, and for which the retrieved column does not exceed the column of the reference atmosphere is selected. The second criterium avoids extrapolation of column densities which may produce unphysical results. This approach is less dependent on the choice of the model atmosphere. Obviously, performing the retrieval six times is not the most computationally efficient, but it is sufficient at the present time. Other, more sophisticated algorithms will be subject to future investigations.

3. APPLICATION TO SIMULATED DATA

In a first verification the water vapour retrieval algorithm is applied to spectra computed with MODTRAN for the combination of the six MODTRAN atmospheres with the seven different solar zenith angles (see above) each for a surface albedo of 5% and 30%.

Figure 3 shows a comparison between the retrieved water vapour columns and the reference values corresponding to the atmospheric model used in the simulation. The results are very similar for both albedos.

The relative deviation between the retrieved and the reference column density is always below 0.7% and typically around 0.4% (see Figure 4). The retrieved column amounts are always smaller than the reference columns. As explained above, this is characteristic for the approach used here and due to the fact that in the determination of \( b \) and \( c \) no atmospheres with columns larger than the MODTRAN reference can be used.

As can be seen from Figure 5, the relative error of the retrieved water vapour column derived from the non-linear fit is generally less than 0.25% which indicates that this error is a good estimate for the retrieval precision.

There is a small dependence on solar zenith angles visible from Figures 5 and 4 as both estimated retrieval error and deviation from reference columns slightly increase with higher SZAs. However, this effect is very small, which indicates that the algorithm is quite insensitive to different viewing geometries.

It may be noted here, that the atmosphere selected by the retrieval is always the one which has been used in the radiative transfer calculations. This does not only show that the “minimum residual” approach works, but also gives confidence that information on the vertical distribution of water vapour can be retrieved from the data. However, the resulting columns are not too sensitive to the reference atmosphere, as long as the retrieved column is similar to or lower than the reference column, which is ensured by the selection process (see above).

4. APPLICATION TO GOME DATA

This section describes some results applying the water vapour total column retrieval algorithm to radiance/irradiance spectra measured by the GOME instrument.

4.1. Comparison with Radio Sonde Data

As sonde data are commonly used for validation purposes, the retrieval algorithm has been applied to a set of GOME measurements which fulfill the following criteria:

- Radio sonde data and GOME data are available for the same day.
- The distance between the centre of a GOME ground pixel and the location of the radio sonde station is...
Figure 3. Comparison between retrieved water vapour columns (crosses) and reference values (lines) for simulated data assuming a surface albedo of 5% (top) and 30% (bottom). The reference atmospheres for the different data sets are indicated. Within each block of reference atmospheres solar zenith angles increase from left to right (0°, 20°, 40°, 50°, 60°, 70°, and 80°).

- GOME pixels have a cloud fraction smaller than 25%.
- Only GOME centre pixels are used.

Sonde data obtained at several stations at middle and higher latitudes on the northern hemisphere have been taken from the NADIR data base of the Norwegian Institute for Air Research (NILU). The above criteria are met by 47 data sets between July 1996 and June 1997 which cover solar zenith angles between 30° and 80°. Because all sonde measurements are performed over land, the 30% albedo parameter set has been used in the retrieval.

The results are shown in Figure 6. A good correlation between both data sets is obtained, but a systematic offset is visible: low sonde columns being under-estimated by the retrieval. This deviation is most likely caused by surface albedos different from (i.e. in the average lower than) the assumed value of 30%. Further studies will have to use an appropriate albedo data base to minimise this effect. Albedo information may also be obtained from the measured radiances and irradiances outside the H2O band, e.g. at around 750 nm.

The average relative deviation between GOME and radio sonde data is approximately 16% with a standard deviation of 28%. Although this is a significant scatter, almost the standard deviations (up to 20%) are observed when comparing radio sonde data with SSM/I or TOVS results (see Ref. 1). As the estimated retrieval error is generally smaller (average value about 1.3%) the observed scatter most likely arises from differences in the sonde effective pixel and the large GOME ground pixels (about 320 km × 40 km) coupled with the high spatial and temporal variability of atmospheric water vapour.

The relatively small number of appropriate sonde measurements does not allow a more detailed analysis of the relation between radio sonde and GOME water vapour columns. This will be subject to further investigations.

Figure 4. Relative deviation of retrieved water vapour columns from reference values for the simulated data presented in Figure 3. Top: Surface albedo 5%. Bottom: Surface albedo 30%

Figure 5. Estimated relative retrieval error for the simulated data presented in Figure 3. Top: Surface albedo 5%. Bottom: Surface albedo 30%
GOME vs. Sondes

Because of the limited availability of appropriate ground based in-situ measurements, satellite remote sensing results are better suited to assess the quality of the GOME water vapour retrieval algorithm as their spatial and temporal coverage is generally larger. Therefore, GOME total water vapour columns are compared with daily gridded Integrated Water Vapor (IWV) data produced by the Global Hydrology Resource Center (GHRC). These IWV data are based on measurements of the Special Sensor Microwave Imager (SSM/I) from the Defense Meteorological Satellite Program (DMSP) F-14 satellite. Data are available on a global scale but due to limitations of the microwave sensing retrieval method only above ocean. Further information may be found at the GHRC web site (http://ghrc.msfc.nasa.gov).

For the comparison we arbitrarily select two SSM/I water vapour data sets obtained during the descending orbits on two days roughly half a year apart (22 November 1997 and 14 June 1998). For each GOME ground pixel the average of all SSM/I data which cover the same area is computed. To simplify the boundary conditions the data set is further reduced by only taking into account cloud-free GOME pixels (i.e. cloud fraction < 0.1). Furthermore no data sets with an SSM/I column larger than 4.1 g/cm² are taken into account which is the largest water vapour column of the MODTRAN reference atmospheres. All higher columns are currently not covered by an appropriate model atmosphere. This leaves a number of 1030 GOME measurements for which the total water vapour column is retrieved and compared with the corresponding SSM/I data. Since all reference data are valid for ocean areas, computations are performed with the 5% albedo parameter set.

As can be seen from Figure 7, similar results are obtained for both days. Despite the significant scatter a clear correlation between the GOME and SSM/I data sets is visible, but there seems to be a systematic offset as the mean GOME water vapour columns tend to be about 10-20% smaller than the corresponding SSM/I columns. This is indicated by the dotted curve in Figure 7 which shows the result of fitting a straight line with zero bias to the data.

The average deviation between both data sets is about 18% with a standard deviation of 20%. This is the same magnitude as it turned out in the comparison between GOME and radio sonde columns. The scatter is significantly larger than the average retrieval error of 1.6%. Therefore, the variability of atmospheric conditions is considered to be a major source of deviations. This is supported by the statistical error of approximately 14% arising from averaging several SSM/I measurements over one GOME ground pixel. Taking into account that SSM/I measurements show a scatter of up to 20% when compared with radio sonde data (Ref. 1) the SSM/I and GOME results agree considerably well. The retrieval error may be reduced by using more sophisticated retrieval methods which e.g. correct for spectral shifts and use a larger atmospheric data base covering in particular also the higher columns.

In agreement with the results of the comparison between GOME and radio sonde data, the systematic offset is especially visible for small columns. The reason for this discrepancy between SSM/I and GOME is still unclear. In contrast to the radio sonde measurements, it is unlikely that this discrepancy between SSM/I and GOME data may be attributed to an inaccurate albedo value in the determination of b and c. Since the observed deviation is higher than expected from the results for simulated spectra, it is most likely not a problem inherent to method. At the moment, the following are considered as potential sources of error in the GOME H₂O retrieval:

1. Influence of partially cloudy ground pixels.
2. Incorrect air mass factors resulting from conditions which differ from the modelled ones (e.g. aerosols, clouds, albedo).
3. Errors in the absorption cross sections for water vapour.
4. Approximations in the calibration of GOME, espe-

Figure 6. Correlation of GOME and radio sonde total water vapour columns assuming a surface albedo of 30%.

Figure 7. Correlation of GOME and SSM/I total water vapour columns assuming a surface albedo of 5%. The dotted curve shows the result of fitting a straight line with zero bias to the data.
cially w.r.t. polarisation.

5. Water vapour line filling due to Ring effect.

As deviations of the same order of magnitude have been found in comparisons between TOVS and SSM/I water vapour columns at higher latitudes (Ref. 1), potential error sources in the SSM/I will also be investigated.

5. SUMMARY AND CONCLUSIONS

A method to derive total water vapour column densities from GOME data has been presented. This method is based on a modified DOAS approach which considers saturation effects. By application on simulated data self-consistency of the method has been shown, and the retrieval precision has been estimated to be better than 1%.

Comparison of total water vapour columns derived from GOME data with radio sonde measurements and SSM/I Integrated Water Vapor results show a significant scatter in the order of 25–30%. This is slightly larger than typical deviations between SSM/I or TOVS and radio sonde data which may be attributed to the larger GOME ground pixel size in combination with the spatial and temporal atmospheric variability. Between SSM/I and GOME water vapour columns a systematic offset of about 18% is found. It will be a task of further studies to investigate reasons for this offset. Nevertheless, it has been shown that water vapour columns may be derived from GOME measurements and that the results are in general agreement with other data sources.

Currently, several improvements of the retrieval algorithm are under development. This includes using the radiative transfer model GOMETRAN (Ref. 9) instead of MODTRAN. GOMETRAN has been specifically developed for use with GOME-like instruments. Recently, GOMETRAN has been extended and accurately simulates H2O and O3 absorptions (Ref. 10). The introduction of an air mass correction factor derived from measured O3 absorption features will also improve the method such that the resulting water vapour columns will be less influenced by the particular atmospheric conditions. This will also be the first step in the development of a retrieval algorithm which can handle cloudy scenes.

ACKNOWLEDGEMENTS

GOME data are provided by ESA. Radio sonde data are taken from the NILU NADIR data base. SSM/I data have been obtained from the Global Hydrology Resource Center (GHRC). Part of this work was funded by the BMBF under grant 07 UFE 12/8 and the University of Bremen.

REFERENCES


