



Improvements of the NO₂ Total Column Retrieval for GOME-2

O3M SAF Visiting Scientist Activity

Final Report

17/11/2006

Hendrik Nüß¹, Andreas Richter¹,
Pieter Valks², and John P. Burrows¹

¹Institute for Environmental Physics

University of Bremen

Otto-Hahn-Allee 1

D-28359 Bremen

²German Aerospace Center (DLR),
Remote Sensing Technology Institute (IMF),
Oberpfaffenhofen, D-82234, Wessling, Germany

Contents

1	Temperature Dependence of the NO ₂ cross-section	4
1.1	Introduction	4
1.2	Simple Scaling Approach.....	5
1.3	Slant Column Fit on Synthetic Data.....	6
1.4	Results and discussion.....	8
2	Estimation of the Stratospheric Contribution.....	9
2.1	Introduction	9
2.2	Reference Sector Method	9
2.3	Using SLIMCAT for the Stratosphere	10
2.4	Using Masking and Smoothing	11
2.5	Summary and Recommendations.....	13
3	Selection of Tropospheric a priori Profile.....	14
3.1	Introduction	14
3.2	Problems.....	14
3.2.1	Identification of polluted scenes	14
3.2.2	Misinterpretation of stratospheric variability.....	15
3.2.3	Introduction of a priori Information.....	15
3.2.4	Selection of typical profiles.....	17
3.2.5	Spatial resolution.....	18
3.2.6	Treatment of clouds.....	19
3.2.7	Use of Monthly Profiles	19
3.3	Tropospheric NO ₂ “climatology”	20
3.4	Summary and Recommendations.....	21
4	Vertical Sensitivity.....	22
4.1	Introduction	22
4.2	Effect of Aerosols.....	22
4.3	Effect of Surface Albedo.....	24
4.4	Effect of Surface Altitude	25
4.5	Summary and Recommendations.....	26
5	Error Study	27
5.1	Introduction	27
5.2	Comparison of different cloud treatments.....	27
5.3	Comparison of different Airmass Factors	30
5.3.1	Change in model version.....	32
5.3.2	Change from daily to monthly AMF.....	34

5.3.3	Change from monthly AMF to monthly profiles	34
5.4	Comparison of different tropospheric NO ₂ products	35
5.5	Summary and Recommendations	41
6	Acknowledgements	43
7	References	44

1 Temperature Dependence of the NO₂ cross-section

1.1 Introduction

The NO₂ absorption cross-section has a marked temperature dependence, both in the absolute value and in the differential structures. For optical remote sensing measurements of the NO₂ amount in the atmosphere, this has to be taken into account to improve the accuracy of the retrieved values.

Ground-based zenith-sky measurements of NO₂ columns are usually performed using a NO₂ cross-section at a “stratospheric” temperature in the analysis, although earlier work used room temperature cross-sections as only these were available at that time. For compatibility reasons, some networks (SAOZ, NDSC) still provide at least part of their data based on analysis with room temperature cross-sections. However, it has been shown that the quality of the fit improves significantly if a cross-section at appropriate temperature is used [Sanders, 1996], and in particular for measurement with a large contribution from tropospheric NO₂ a further improvement can be obtained by including two NO₂ cross-sections at different temperatures (possibly orthogonalised to each other) in the fit. In cases of strong pollution, the temperature dependence of the cross-section can even be used to differentiate between “warm” tropospheric NO₂ and “cold” stratospheric NO₂ [Richter, 1997].

In addition to the impact on the quality of the spectral retrieval, the absolute column is also affected as at least in the “standard” fitting window 425 - 450 nm, the temperature dependence is in first order a scaling of the differential structures. This is illustrated in Fig. 1a, where differential cross-sections of NO₂ as measured with the GOME-FM [Burrows et al., 1998] are plotted for different temperatures. The change with temperature is substantial (25% between 221K and 293K) and can not be ignored in the atmospheric temperature range. By simply scaling the cross-sections with appropriate factors, they can be brought into reasonable agreement as shown in Fig. 1b. This quasi-linear dependence can be used in a correction scheme developed by Boersma et al. (2004) to adjust the column value according to temperature after the fit. The basic idea is to derive a correction factor as a function of temperature and then apply it to the slant column using a temperature profile, a profile of NO₂ and the altitude dependent airmass factors:

$$c_l = \frac{221 - 11.4}{T_l - 11.4} \quad (1)$$

where c_l is the correction factor for layer l and T_l is the temperature of layer l . In the airmass factor calculation, this factor is applied to each layer:

$$AMF = \frac{\sum_l m_l(\mathbf{b}) x_{a,l} c_l}{\sum_l x_{a,l}} \quad (2)$$

where the m_l are the airmass factors for the individual layers l based on the model parameters \mathbf{b} and $x_{a,l}$ is the a priori NO₂ concentration in layer l .

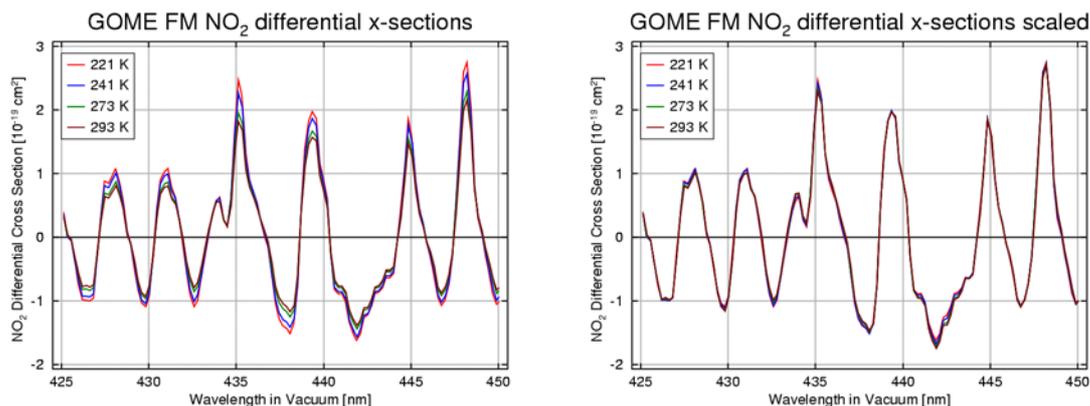


Fig. 1: Temperature dependence of the differential NO₂ absorption cross-section. Left: differential cross-sections, right: scaled differential cross-sections.

As the correction is largest for large temperatures, it is more relevant for tropospheric NO₂ than for stratospheric NO₂. Unfortunately, the vertical profile of NO₂ can not be inferred from nadir UV-visible measurements, and therefore has to be taken from other sources. This introduces rather large uncertainties in the correction, for example if the a priori assumes that a large part of the NO₂ is located close to the surface at high temperatures while in reality it is higher up in the atmosphere at much lower temperature. The temperature itself is also derived from external information, but uncertainties here are much lower if meteorological analysis fields are used.

In principle, it would be preferable to determine the correct temperature to be used for the NO₂ cross-sections from the fit itself as this would be independent of external information. In the case of O₃, which also exhibits strong temperature dependent differential structures, this is routinely done in ground-based measurements and also in the GDP-4 processor. The approach taken is to supply the fit with two ozone cross-sections; one taken at a typical stratospheric temperature and the other one being the difference between the cross-sections measured at two temperatures. If the temperature dependence is linear and the structures are large enough and not too much correlated with other absorption features, the fit will “select” the right temperature by combining the two cross-sections and at the same time retrieve the correct slant column. As shown below, the same approach is feasible for NO₂ in synthetic data, but experience with GOME and SCIAMACHY data shows, that in real measurements the SNR is usually not large enough to successfully apply this method.

1.2 Simple Scaling Approach

As the simplest approach to determine the temperature dependence of the NO₂ cross-sections, the differential cross-sections at different temperatures have been scaled to the differential cross-section at 220K using a linear least-squares fit. The resulting scaling factors are given in Table 3 for the GOME-FM cross-sections, the SCIAMACHY-FM cross-sections [Bogumil et al., 2003] and for comparison also for the high resolution FTS measurements of VanDaele et al. (1998). The results show a high degree of consistency indicating that i) the temperature dependence of the NO₂ absorption cross-section is well characterised and ii) that the effect depends only weakly on spectral resolution, at least in the range of values studied here. It is to be expected that the GOME-2 NO₂ cross-sections will have very similar temperature dependence, but this will have to be verified.

As shown in Fig. 2, the temperature dependence can in very good approximation be described by a linear regression which can then be used to determine the value at any specific temperature.

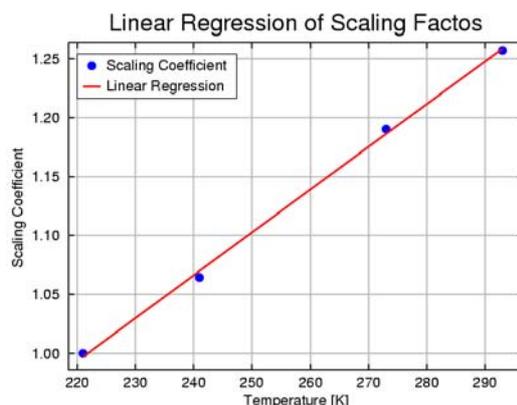


Fig. 2: Scaling coefficients for the temperature dependence of the NO₂ cross-sections and a linear regression through the data

1.3 Slant Column Fit on Synthetic Data

A different approach to evaluating the effect of temperature on the NO₂ DOAS analysis is to do an end-to-end simulation on modelled data. To this end, radiance spectra have been simulated with the radiative transfer model SCIATRAN [Rozanov et al., 2001] for a number of solar zenith angles using NO₂ cross-sections at different but constant temperatures (see Table 1 for settings used). The resulting spectra have then be analysed using the DOAS settings foreseen for the GOME-2 operational processor (see Table 2), and the resulting slant columns compared to the “true” slant columns which were derived from the simulation using the same NO₂ temperature as the fit (221K). Again, a set of scaling factors was derived which is given in Table 3. As the analysis was performed on a series of solar zenith angles (SZA), one can use this as a check of the stability of the factors found. As shown in Fig. 3 for the 293K case, the variation with SZA is small and can be neglected.

parameter	value
viewing geometry	satellite nadir
surface albedo	0.05
wavelengths	422 - 453, 0.2 nm steps
irradiance	Kurucz convoluted to GOME resolution
cross-sections	GOME-FM NO ₂ (221K or 241K or 273K or 241K) GOME-FM O ₃ 221 K Greenblatt O ₄
SZA	10, 50, 70, 80, 85, 90

Table 1: Settings used for the simulated radiances used for the determination of the effect of temperature changes on the NO₂ analysis

parameter	value
wavelength window	425 - 450 nm
polynomial	cubic
cross-sections	GOME-FM NO ₂ at 221 K GOME-FM O ₃ at 221 K Greenblatt O ₄

Table 2: Settings used for the DOAS analysis of the simulated radiances

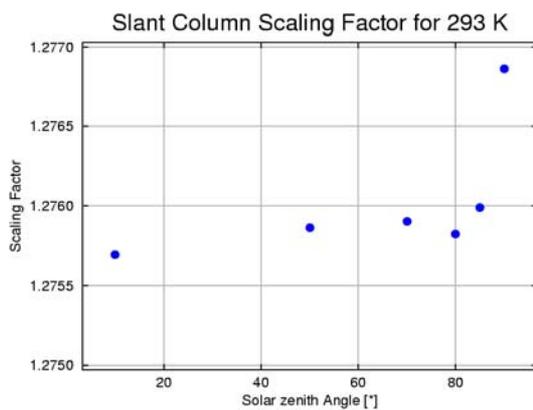


Fig. 3: SZA dependence of the scaling factor derived from the synthetic data fit for 293 K

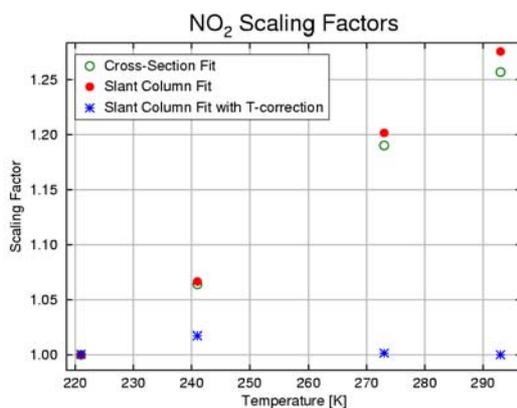


Fig. 4: Scaling factors for different NO₂ temperatures derived with the cross-section fit (open circles), the slant column fit on simulated data (filled circles) and the residual factors when including the temperature dependence in the fit of synthetic data (stars).

Method	221 K	241 K	273 K	293 K
scaling GOME-FM	1.000	1.064	1.190	1.257
scaling SCIAMACHY-FM	1.000*	1.056*	1.175*	1.266*
scaling VanDaele FTS	1.000**	---	---	1.292**
scaling VanDaele FTS convoluted to GOME-FM	1.000**	---	---	1.294**
SC fit GOME-FM	1.000	1.067	1.202	1.276
scaling GOME-FM in OMI window (405 - 465 nm)	1.000	1.059	1.177	1.242
Boersma et al., 2004	1.000	1.095	1.248	1.344

Table 3: Temperature effect of the NO₂ cross-section as derived by different methods. The reference was 221K. The numbers are the factors that have to be applied to a slant column retrieved using the NO₂ cross-section at 221 K if the real temperature was 241/273/293 K. The values from Boersma et al. (2004) are inverted to account for the difference in definition. **vanDaele Temperatures are 220K and 294K, *SCIAMACHY-FM temperatures are 223K, 243K, 273K and 293K.

As mentioned above, the temperature dependence of the NO₂ cross-section can be corrected in the DOAS fit by introducing an additional cross-section which is computed from the difference of NO₂ cross-sections taken at two temperatures. With this correction included, the slant column derived from the fit is much closer to the “true” value (the slant column derived using the correct 221K cross-section without temperature correction). This is illustrated in Fig. 4 where the scaling factors are shown with and without temperature correction. The small deviation (about 1%) for 241 K is related to the non-linearity of the temperature dependence, which could either be a spectroscopic effect or the result of measurement errors in the GOME-FM cross-sections. This approach would make any corrections on the AMF level unnecessary. Unfortunately, application of this method is limited by the signal to noise ratio of the spectra and can usually not be applied to real satellite measurements

1.4 Results and discussion

As shown in Table 3, the two methods (scaling of cross-section and slant column fit on synthetic data) do not agree fully in their results. The largest difference occurs at the largest temperature difference (293K) where the two approaches yield results with roughly 2% difference. The reason for this difference is not obvious, but the most probable explanation is that some of the other contributions to the radiative transfer simulation (absorption by O₃ and O₄, scattering) have some correlation with the temperature signature in the NO₂ cross-section. In real data, there are many more parameters that could interfere with these structures (surface albedo, Ring signature, other minor absorbers, instrumental effects) and even larger differences are to be expected. This can not be avoided and limits the accuracy of the approach to correct the effect not on the fitting but on the airmass factor level even if the input parameters are perfectly known (which they are not). With this limitation in mind, it is not necessarily true that the more sophisticated computation based on slant column simulations is actually providing higher accuracy.

For comparison, Table 3 also lists the values given by Boersma et al (2004) for the temperature dependence of the NO₂ slant column. The details of the derivation are not given in that paper and also not in the reference cited [Chance et al., 2002] and at this point it is not clear why the difference is that large although the same quantity is derived.

2 Estimation of the Stratospheric Contribution

2.1 Introduction

One of the aims of the operational GOME-2 NO₂ product is to derive an accurate total column. Unfortunately, the two main contributions to the column, the stratospheric and the lower tropospheric NO₂ amounts, are of varying but comparable magnitude and contribute to the signal with different weights. From the measurements alone, there is no way to separate the two components which is necessary to correct for the difference in measurement sensitivity at the two altitude levels. The usual approach to this problem is to use measurements over clean areas for the stratospheric NO₂ amounts and to extrapolate them with some assumptions (or by using a model) to polluted areas. There, the estimated stratospheric contribution is subtracted, and the tropospheric column determined. Finally, the two values are added to yield the total column. While there is no alternative to this approach if no other information is available (such as the limb profiles in SCIAMACHY measurements), it must be kept in mind (and communicated to the data user) that the retrieval can only produce accurate total columns where one of the two components is small compared to the other. In fact, only one of the parts is determined from the measurement over that pixel while the other is based on external information or assumptions e.g. from smoothing over latitude.

In the following, different methods used in the literature for quantification of the stratospheric contribution are briefly discussed and compared in sample applications to GOME data.

2.2 Reference Sector Method

As a first approximation, it is often assumed that the Pacific area is clean, and that stratospheric NO₂ depends only on latitude, not on longitude (e.g. *Richter and Burrows, 2002, Martin et al., 2002, Beirle et al., 2003*). However, in particular in winter in mid- and high latitudes, this assumption does not hold and introduces significant errors. This is illustrated in Fig. 5 using SLIMCAT NO₂ columns (*Chipperfield, 1999*).

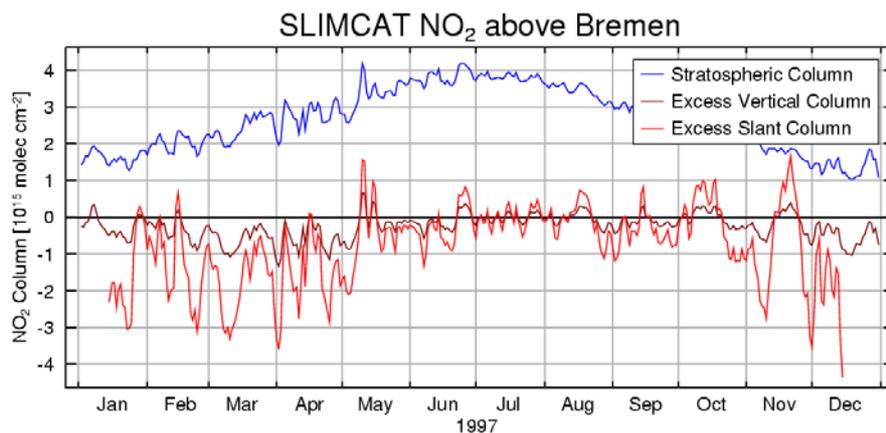


Fig. 5: SLIMCAT NO₂ columns above Bremen (53°N) for 1997. The upper line is the stratospheric column with the typical mid-latitude seasonality: low winter values and high summer columns. The brown curve is the vertical excess column with respect to the Pacific area (180° - 210° longitude). In particular in winter, the stratospheric NO₂ column over Bremen varies strongly and is on average lower than over the Pacific. The red curve results when the excess vertical column is multiplied with the stratospheric airmass factor yielding the excess slant column.

The stratospheric column over Bremen has a maximum in summer and a minimum in winter. In addition to this smooth variation which results mainly from the change in sunlight hours with season, there are short term fluctuations with an amplitude of about 10^{15} molec cm^{-2} . When subtracting the Pacific area as is done in the RSM (reference sector method), significant, mostly negative vertical columns result over Bremen instead of the zero values expected if the assumption on zonal homogeneity of stratospheric NO_2 would hold. After multiplication with the stratospheric airmass factor, an excess slant column of up to -4×10^{15} molec cm^{-2} results. Considering that the tropospheric airmass factor is of the order of 1, this already is approximately the error made in the tropospheric columns.

2.3 Using SLIMCAT for the Stratosphere

If the SLIMCAT model results are available, they can be used to subtract the stratospheric component of the measurements. This is illustrated in Fig. 6, where GOME measurements and SLIMCAT model data are compared between 40° and 50°N on May 2-4, 1999. On these days, there is a marked trough of lower GOME NO_2 columns over the Pacific, which would result in enhanced tropospheric columns at all other longitudes when using the reference sector method. By using the SLIMCAT model column, which also shows this behaviour, the stratosphere can be corrected much more accurately. Unfortunately, the absolute values of the SLIMCAT stratospheric column and the GOME measurements over clean areas do not always agree well, and an additional scaling step has to be performed to link the two datasets and to apply this type of correction [Richter et al., 2005]

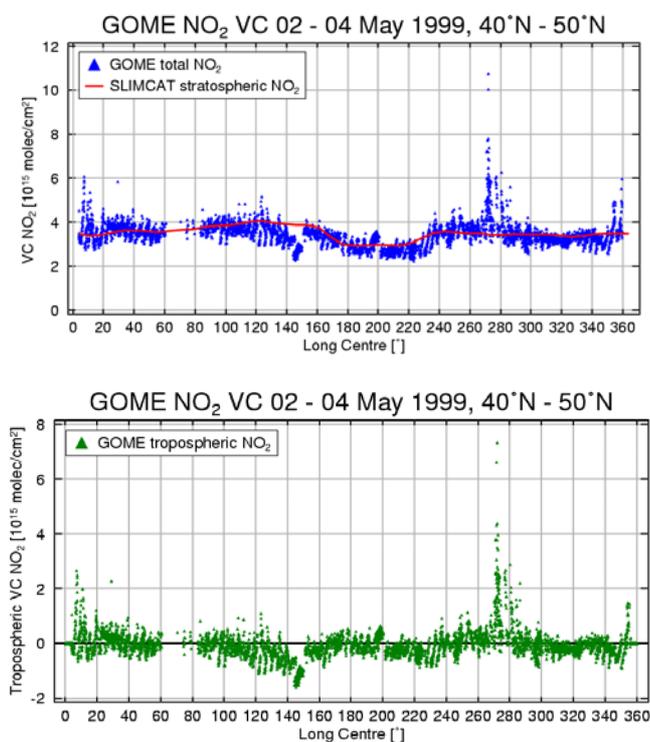


Fig. 6: GOME NO_2 measurements between 40° and 50° north on May 2-4, 1999 (upper panel), the corresponding SLIMCAT columns (red line in upper panel) and the tropospheric column resulting from the difference (lower panel). The airmass factors used for the lower figure are stratospheric and not appropriate for tropospheric columns.

An in many ways similar approach is taken by *Boersma et al.* (2004) for the correction of the stratosphere by assimilating GOME NO_2 columns from clean regions into a model. This

forces the model to the satellite data in more than just one area, and provides an “intelligent interpolation” of the sparse GOME field.

2.4 Using Masking and Smoothing

If no external information is available, the reference sector method can be improved by using data not only from the Pacific, but also from other areas of the world which are known to be relatively clean. Using these masked areas, a smoothed and interpolated stratospheric field can be constructed (*Leue et al.*, 2001, *Boersma et al.*, 2002). The advantage is, that zonal variations of the stratospheric NO₂ column can be corrected using only data from the instrument, and that the correction is self-consistent and does not need ad hoc scaling corrections as discussed in the section on the SLIMCAT correction. The disadvantage is that large scale pollution events or transport of pollution can bias the “stratospheric” field, leading to a systematic underestimation of tropospheric columns. Also, some polluted areas such as Europe or the US are so large that substantial interpolation is needed which reduces the accuracy of the method.

Several options are available for the details of the masking and smoothing. For the masking, model data can be used as for OMI (*Boersma et al.*, 2002) or again the measured data themselves. The latter approach risks to misinterpret stratospheric features (see section 3.2.2) while using model data might bias the results towards the a priori. To illustrate this problem consider a large biomass burning area in Africa which is for some reasons missing in the model. It will therefore not be masked and the high values will be attributed to the stratospheric field. As a result, the tropospheric NO₂ columns in this area will not be enhanced as much as they should. Although this is an unlikely scenario, there is some possibility for systematic errors.

The smoothing can be performed in different ways, and basically the choice is between allowing more degrees of freedom which will improve the fit to real stratospheric variability but potentially also removes some tropospheric enhancements and using less degrees of freedom which results in poorer fits and problems at high stratospheric gradients but reduces the risk of removing tropospheric NO₂. Sensitivity studies for the OMI instrument have considered both w2 and w4 fits and found better results with w2 fits in model data (*Buscela et al.*, 2006). Tests with the w2 and w4 fits in comparison to running 30 degrees boxcar averaging as foreseen for GOME-2 showed that the latter is more similar to the w4 fits and often performs better than w2. In contrast to the OMI tests these studies were performed on real GOME data which might influence the results.

It has also been suggested that measurements over clouds can be used to estimate the stratospheric columns, but at least in mid-latitudes, substantial NO₂ amounts are often present above and within clouds over polluted regions (*Wang et al.*, 2005) and lightning events can have a similar effect. Therefore, this method is not recommended.

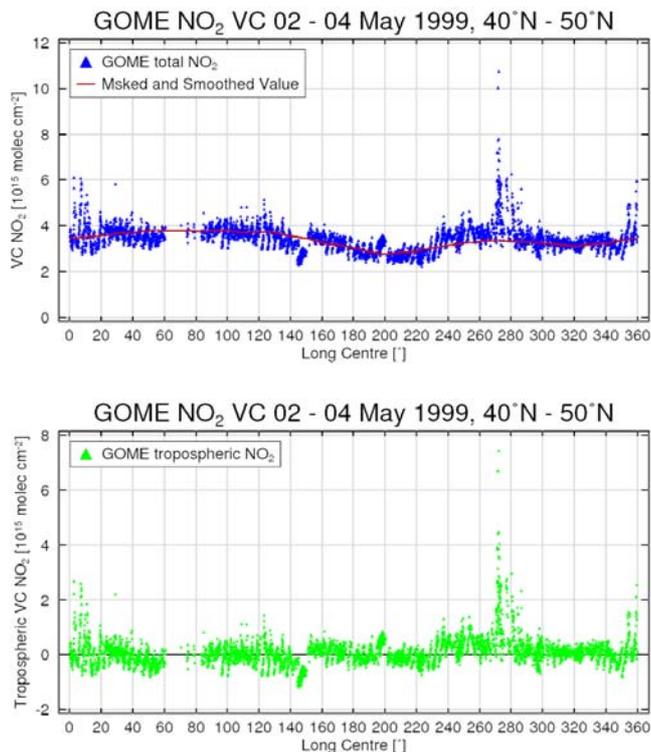


Fig. 7: As Fig. 6 but for the masking and smoothing method. GOME NO₂ measurements between 40° and 50° north on May 2-4, 1999 (upper panel), the corresponding masking and smoothing columns (brown line in upper panel) and the tropospheric column resulting from the difference (lower panel). The airmass factors used for the lower figure are stratospheric and not appropriate for tropospheric columns.

An example for the results obtained using the masking and smoothing method is shown in Fig. 7 for the same days as used in Fig. 6. As can be seen, the results are very similar to those obtained using the SLIMCAT data. In some areas, the fit seems to be a bit closer to the GOME total columns which is to be expected given the fact that the SLIMCAT data are model results on a coarse spatial grid. In this example, 3 days of GOME data are used which leads to obvious discontinuities e.g. at 140° and also is not feasible in an operational environment intended to produce near-realtime data. However, with the improved spatial coverage of GOME-2 similar results should be obtainable on a daily basis.

A slight complication of the masking and smoothing algorithm is introduced by the photochemical evolution of stratospheric NO₂ over the day. As result of the large swath of GOME-2, in mid and high latitudes several measurements per day will be made over the same location albeit at different local times. Therefore, the NO₂ columns used in the smoothing procedure are from different points on the diurnal evolution of stratospheric NO₂ and can not simply be combined to one stratospheric NO₂ field. Also, the individual measurements need a different correction in particular at high latitudes. This problem can be solved by using photochemical corrections for each measurement in the creation of the stratospheric background as well as in the correction for each point.

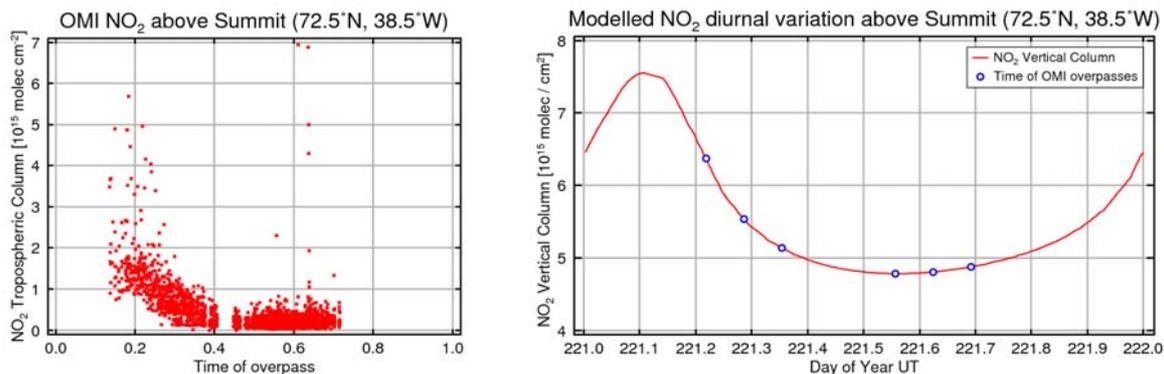


Fig. 8: Effect of diurnal variation of stratospheric NO₂ on (apparent) OMI tropospheric NO₂ column over the clean air site Summit, Greenland. OMI data source: EOS Aura OMI OMNO2-ECS2, Generated: 22-Jul-2006 by <http://avdc.gsfc.nasa.gov>

2.5 Summary and Recommendations

It is clear from previous studies and the examples shown here that zonal variability of stratospheric NO₂ is too large to ignore in the retrieval of tropospheric NO₂ columns. The model based approaches (using either SLIMCAT data or assimilated NO₂ fields) have proven to be efficient but are not compatible with the operational environment foreseen for GOME-2. Therefore, another approach has to be taken.

The masking and smoothing approach used for OMI data is promising and tests using running boxcar averaging and different degrees of masking performed well on many selected GOME data sets. They therefore are the method of choice for the GOME-2 NO₂ retrieval. However, the exact degree of smoothing and the areas masked will have to be optimised on real GOME-2 data based on results from validation with independent measurements and models.

For northern latitudes, the effect of diurnal variation of stratospheric NO₂ on the smoothed background NO₂ has to be evaluated to avoid artefacts as observed in the current OMI product at high latitudes.

3 Selection of Tropospheric a priori Profile

3.1 Introduction

As discussed in the last section, the signal measured from a nadir viewing satellite comprises both the tropospheric and the stratospheric components, albeit with different sensitivity. In particular in the troposphere, the measurement sensitivity is a strong function of altitude and depends on a variety of parameters including surface albedo, aerosol loading, viewing geometry and last but not least cloud coverage. The approach usually taken is to calculate the measurement sensitivity as a function of the above parameters for each altitude level and to create a look-up table. For an individual measurement, the corresponding vertical sensitivity is retrieved from the data base and multiplied with the assumed vertical profile of the absorber. The latter comes from another data base which can have varying degrees of complexity from a small number of typical profiles to an online 3d-CTM calculation. Some considerations for the selection of these profiles are discussed in the following sections.

3.2 Problems

3.2.1 Identification of polluted scenes

As the measurements themselves do not contain information on the vertical distribution of the NO₂, the absolute column can be used to identify polluted situations for which a tropospheric column is to be calculated. This can conveniently be done by defining a threshold value on the residual column obtained after subtracting the stratospheric column from the measured column, and using only those measurements which are above that threshold. This approach, which is used in the initial OMI NO₂ processor (*Boersma et al.*, 2002), ensures positive tropospheric columns and avoids interpretation of errors from noise or incomplete removal of the stratospheric component as tropospheric NO₂.

However, such a threshold introduces a systematic high bias when calculating monthly averaged tropospheric NO₂ distributions, as only the (clearly) positive part of the error distribution contributes to the signal, and all smaller and negative parts are ignored. Using such a technique on a stochastic NO₂ field will on average result in positive tropospheric NO₂ column where zero should be obtained. Therefore, no selection for high columns should be performed at all to avoid this bias. Clearly, users of the tropospheric product must be advised on how to interpret the data, as negative values are to be expected within the error range of the method.

3.2.2 Misinterpretation of stratospheric variability

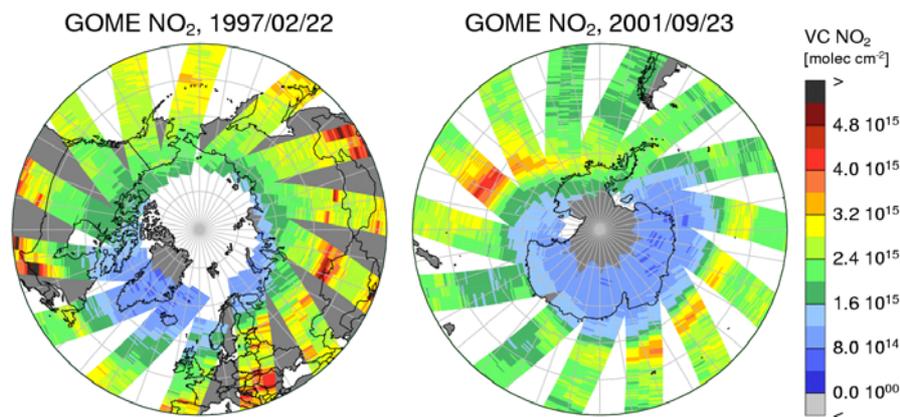


Fig. 9: GOME NO₂ columns for two days illustrating situations where the stratospheric NO₂ distribution deviates clearly from the zonal homogeneity which is often assumed in the retrieval. In the left figure, very low NO₂ columns extend over Greenland, Iceland and parts of the North Atlantic. In the right figure, the area with low NO₂ is also displaced, and in addition two areas of strongly enhanced stratospheric NO₂ columns are observed outside the polar vortex area, a situation typical for the breaking up of the polar vortex in the Southern hemisphere.

If polluted scenes are to be identified from the measurements themselves and not from a priori knowledge or model results, the obvious approach is to use an absolute threshold above background (probably as function of latitude and season) above which an airmass is considered to be polluted and an appropriate vertical profile is used in the airmass factor calculation. While this will probably work well in most situations, in high latitudes in spring stratospheric NO₂ occasionally varies strongly and can lead to structures which are difficult to distinguish from tropospheric pollution. An example is given in Fig. 9 where asymmetric stratospheric NO₂ distributions are shown for both hemispheres and ring like high NO₂ region in the SH stratosphere that is of similar magnitude and spatial structure as the pollution signatures in the NH. They therefore will be flagged as such and reported in a tropospheric NO₂ product based on a threshold selection technique.

Such stratospheric structures will only affect a limited number of measurements but their occurrence has to be noted in a disclaimer of the product.

3.2.3 Introduction of a priori Information

A general problem of the vertical change in measurement sensitivity is the dependence of the results on the assumptions made for the airmass factor calculation.

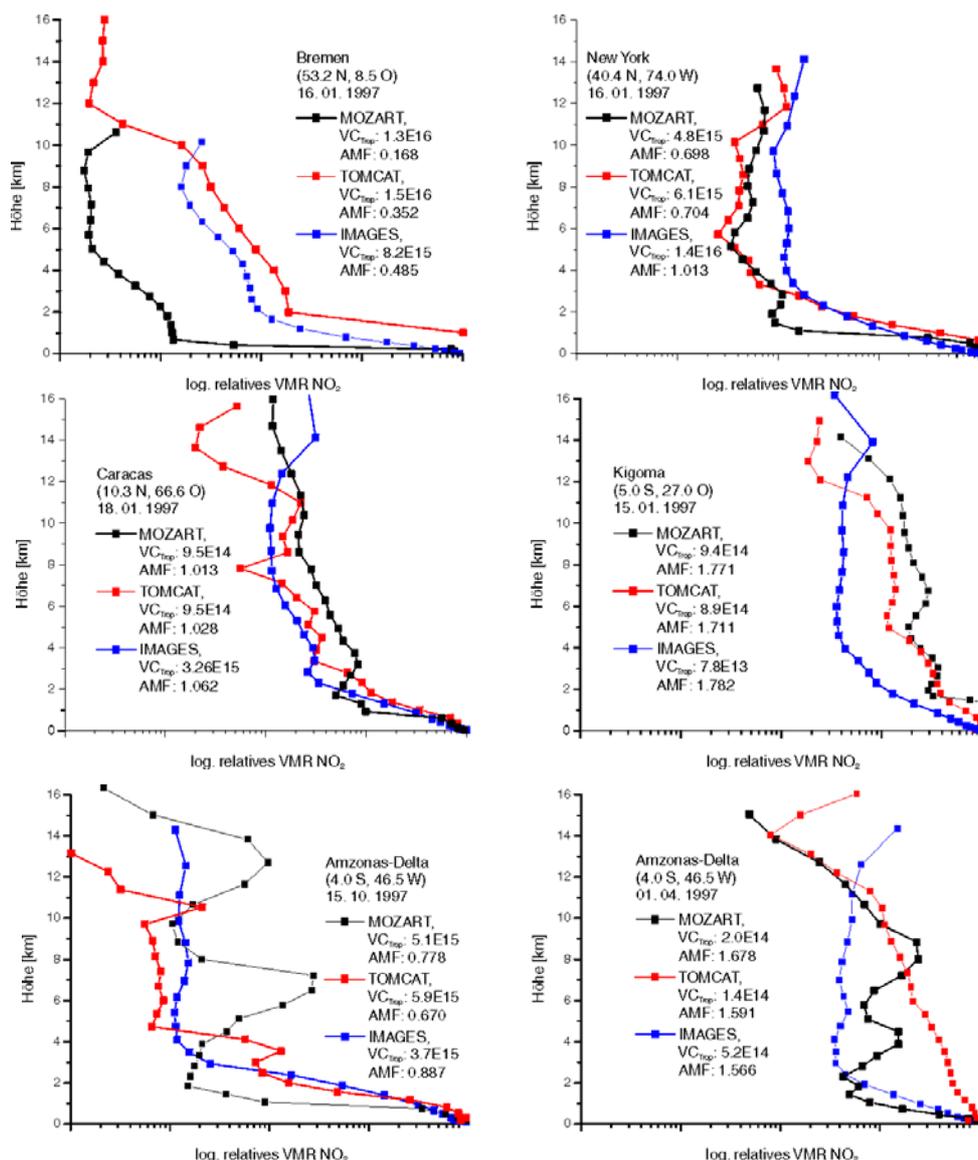


Fig. 10: NO₂ profiles for different locations on a given day as predicted by the 3d-CTM models MOZART, TOMCAT, and IMAGES. The profiles are normalised to the lowest value and shown on a logarithmic scale.

One interesting question is how good the agreement between the vertical profiles predicted by different state-of-the-art models is. The degree of variability observed is illustrated in Fig. 10, where profiles over 6 locations are shown for individual days. Results from three models (MOZART, TOMCAT and IMAGES) are shown, normalized to the lowest value.

As is evident from the figure, the agreement is good over some areas, but in particular over biomass burning and over urban areas, the profiles differ significantly, both in shape and in total column. Also shown in the figures are the air mass factors computed from the individual profiles. In particular over “Bremen” and “New-York”, the air mass factors differ by a factor of two with MOZART profiles giving the lowest values, IMAGES the largest.

These examples show that there is currently no consensus in the models on what the vertical profile of NO₂ over a given area is. Some of the differences can be explained by low model resolution (IMAGES) or differences in the wind fields used, and for many parameters it can be decided, which treatment is more accurate. However, as also was shown in a recent IGAC initiated model comparison (*van Noije, 2005*), the spread of modelled vertical NO₂ profiles is

large and the uncertainty introduced in the air mass factors (and thereby the satellite retrieved NO₂ columns) can be up to 100% over polluted areas.

3.2.4 Selection of typical profiles

One of the most critical inputs are the vertical NO₂ profiles needed for the air mass factor calculation. Different possible approaches have been used in the past to obtain these data:

- a single profile for all seasons and locations (e.g. *Velders et al.*, 2001, *Richter and Burrows*, 2002, *Leue et al.* 2001, *Beirle et al.*, 2004)
- a set of several typical profiles (*Boersma et al.*, 2002)
- a monthly climatology of air mass factors based on runs from the MOZART-2 model (e.g. *Richter et al.*, 2005)
- on-line CTM calculations for the time and location of measurement (e.g. *Martin et al.*, 2002, *Boersma et al.*, 2004)

The advantages of disadvantages of using input from online models have been discussed above. If the vertical profiles for the air mass factor are selected from a limited set of profiles, the question arises, which one to select for an individual measurement.

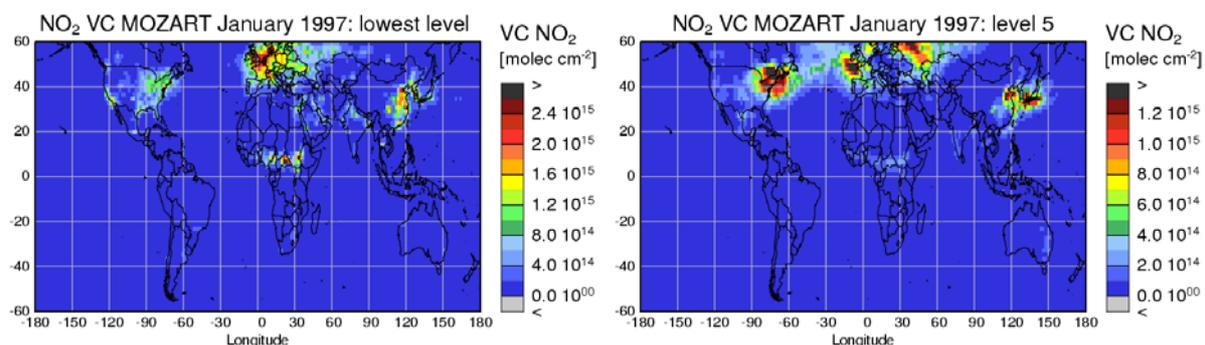


Fig. 11: Partial NO₂ column as modelled by the MOZART-2 model for January 1997 on sigma levels 1 (surface) and 5 (1.2 - 1.5 km over the ocean). A clear separation between source areas and areas influenced by transport in or into the free troposphere can be observed. Please note the different colour scales.

In principle, model results can be used to identify areas with high emissions (profile dominated by lowest layers), areas which are only affected by transport (NO₂ in the free troposphere) and regions with both emissions and convection or transport. This is illustrated in Fig. 11 where the partial columns for the surface layer and the sigma layer between approximately 1.2 and 1.5 km are shown. While surface NO₂ is large over the source regions, the effect of transport and convection is apparent over the Atlantic (outflow from the US), West of Europe and over Eastern Europe (outflow from Western Europe) and over China and Japan. There also is some NO₂ in the higher layer over areas with biomass burning in Africa and Australia. Using emission maps, fire counts, lightning measurements, and meteorological data, typical profiles can be assigned to individual regions on a seasonal basis. However, inspection of the daily profiles shows that in most areas, variability is large.

Another possible choice is to use a large number of model simulations (e.g. several years of model runs) and to compute an averaged profile for a given location in a given time period. This approach is reasonable for areas with relatively constant situations, e.g. over large NO_x source regions or areas with seasonal biomass burning. However, in many regions, changing

atmospheric situations can lead to strongly varying vertical distributions. This is true for areas that are affected by transport of polluted air masses or by sporadic lightning events. Under these conditions, the average profile can be quite distinct from the profile observed during a period with enhanced tropospheric columns.

One possibility to make the average profile more representative for situations with enhanced tropospheric columns is to use a minimum threshold for the tropospheric column when averaging the profiles. Depending on the threshold, this will lead to gaps in the data base in areas where the threshold is never reached in the model which have to be filled by interpolation in space or time or by using a default profile under such conditions. If the model gives a good representation of the real situation, these areas will rarely be affected by large tropospheric columns in the measurements and such an approximation will only affect few values. However, the choice of threshold value is to some degree arbitrary and will have an impact on the results.

3.2.5 Spatial resolution

Today, the spatial resolution of the models typically employed to compute the vertical profiles used in the air mass factor calculations is much lower than that of GOME-2 measurements, in particular if operated in high resolution mode. On the other hand, the spatial scales of both NO₂ emissions and meteorological processes involved in vertical and horizontal transport are often comparable or even smaller than those of the satellite measurements. As a result, the real vertical profiles of NO₂ often change from measurement to measurement while the model profile remains the same over a relatively large grid box. This is illustrated in Fig. 12, where a 30 x 30 km² measurement of SCIAMACHY over Spain and Portugal is compared with typical model resolution (2.5°x 2.5°) and high model resolution (1°x1°). Clearly, even the high resolution model will not resolve the NO₂ distribution even over a large metropolitan area such as Madrid. To account for horizontal gradients and the irregular grid of the satellite measurements, probably one order of magnitude higher resolution is needed which is not feasible today for global models. It is safe to assume that the large variations in NO₂ column are linked to large changes in vertical distribution, mainly depending on the presence or absence of sources and the amount of vertical and horizontal transport.

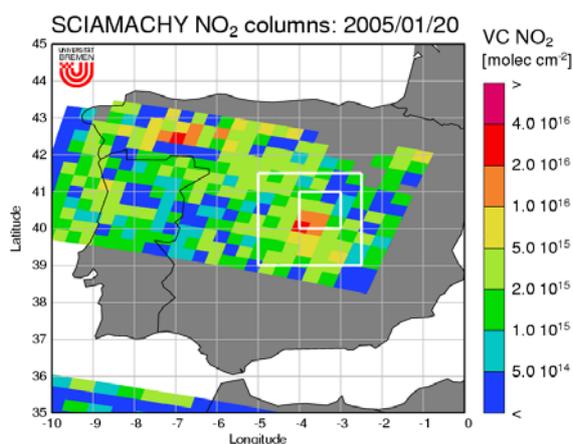


Fig. 12: Spatial variability of tropospheric NO₂ compared to typical (2.5°x2.5°) and high (1°x1°) spatial resolution of global models used to determine the vertical profiles for the air mass factor calculations. The air mass factors used in the analysis shown are based on a 1.875° x 1.875° model run.

The mismatch in spatial resolution has several effects:

- As result of the non-linearity in the NO_x chemistry, the coarse resolution model results will not necessarily be a good representation of the average situation in the grid box.
- The profile computed will neither be appropriate for the urban area, nor for the rural area around the city, and NO₂ columns in the city will be underestimated while NO₂ outside the city will be overestimated.
- Artificial gradients can be introduced in the NO₂ fields at the limits of model grid boxes. This can be avoided by using interpolated fields, but this will further smear out the spatial variability in the air mass factors.
- The effect of clouds can be misinterpreted (see next section).

There is no simple solution to the spatial resolution problem which currently can only be treated by adding its effect to the error budget. In the absence of high resolution model results even the error budget is difficult to access.

One approach that could be taken is to use emission inventories with high spatial resolution and trajectory models such as FLEXPART (*Stohl et al., 1998*) with little or no chemistry to produce the high resolution fields needed. As long as the NO₂ fields are dominated by emissions and transport, these models will produce reasonable NO_x distributions. However, at the current time, this solution is not yet feasible in an operational environment. Also, many of the objections discussed above still hold.

3.2.6 Treatment of clouds

One of the most important error sources in the retrieval of tropospheric NO₂ are clouds. Clouds shield the lower atmosphere from view, and at the same time enhance the signal of absorption above them as result of the albedo effect. Two different approaches can be taken in the treatment of clouds:

1. cloud screening without further correction (e.g. *Richter and Burrows, 2002*). This approach does not rely on any external information on the vertical NO₂ profile, but will result in a significant underestimation of tropospheric NO₂ columns in regions with strong sources and broken cloud fields. However, to some degree a cancellation of the two cloud effects will reduce the error of this simple assumption.
2. simulation of the effect of clouds on the retrieval and application of an appropriate correction factor (*Martin et al., 2002, Boersma et al., 2004*). This approach will yield more accurate results if the model simulation is a good approximation of the real situation, but has the problem of strongly introducing model assumptions into the final columns. While this is also true for clear pixels, the situation is worse for cloudy pixels as part of the column is shielded from view, and therefore no information is in the measurements. Also, the model resolution is much coarser than typical cloud fields, and therefore the assumption that the model profile is valid in an area with broken clouds is questionable.

For the GOME-2 processor, the second approach is foreseen and two different cloud algorithms have been tested. This is described in section 5.2.

3.2.7 Use of Monthly Profiles

If no online 3d-CTM calculations are available, the NO₂ vertical profiles have to be retrieved from a “climatology”. One way of implementing this is to use monthly averages of profiles from a 3d-CTM run. However, through this averaging, the daily variations which are

dominated by transport and clouds will be lost and an error is introduced in the retrieved columns. To assess the magnitude of this error, airmass factors have been calculated from daily NO₂ fields modelled by MOZART for 1997 and the mean, minimum, maximum, and standard deviations be analysed for each month. As all other inputs (reflectivity, aerosols, emissions) vary on a monthly basis, the variations observed are all due to meteorology.

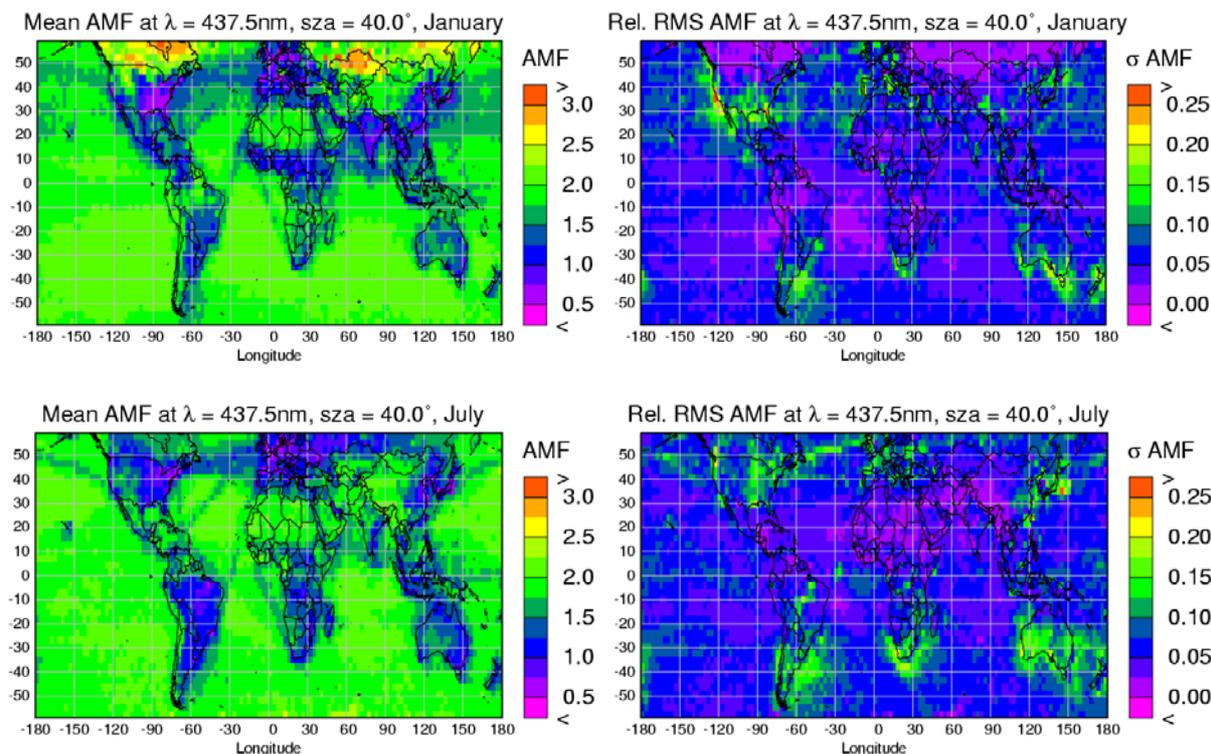


Fig. 13: Monthly average and relative RMS of the NO₂ AMF at 40° solar zenith angle for January and July. Tropospheric AMF are of the order of 1.5 and smaller in source regions and about 2 over clean regions at this SZA. The relative RMS is usually between 5% and 10% but can be up to 20% locally.

Some examples are given in Fig. 13. The relative RMS is lower than 10% for most areas, but can be up to 20% in particular in regions affected by transport and along coastal regions. For individual locations, extreme values can vary up to a factor of 3 within one month. All in all, the uncertainty appears acceptable given the size of other error in the analysis.

In real data, the SZA changes of course with season and latitude, and for a monthly average the satellite sampling and the relation between absolute NO₂ columns and airmass factors also plays a role. For example, small AMF are often associated to large columns while small columns have larger AMF which is not accounted for when using the averaged profiles. As a result, differences are levelled out.

3.3 Tropospheric NO₂ “climatology”

In order to be able to use appropriate profiles for airmass factors without having access to an online 3d-CTM, a “climatology” of tropospheric NO₂ profiles was produced based on a long-term run of the MOZART model at T63 resolution (1.875°). The data were provided by MPI Hamburg as part of the European project RETRO and use the settings (emissions, chemical scheme) agreed upon within the frame of that project. As the profiles are based on model data only, they should not be called a climatology. However, for lack of measurements, averages

over a long time series of simulations driven by re-analysed meteorological wind fields are the best data set available. Clearly, the problems discussed in sections 3.2.3, 3.2.5 and 3.2.7 also apply to this data set and in the future, better model runs at higher spatial resolution will become available that should then be used for the airmass factor calculations.

The data used here are from the period November 1994 to October 1998 and are averaged over months. Two versions of the climatology were created, one with all the data included and one with selected profiles which correspond to polluted situations only (column $> 5 \cdot 10^{14}$ molec cm^{-2}). The latter is more consistent with the threshold approach foreseen for the GOME-2 data analysis but the difference between the two data sets is small. The files are in NET-CDF format and were provided to DLR.

3.4 Summary and Recommendations

Vertical profiles of NO_2 in the troposphere are needed as input for the airmass factor calculations. They should be as representative as possible and at the same time introduce as little bias as possible into the product. Different approaches can be used, e.g. a small set of "typical" profiles, a "climatology" of model profiles or online 3d-CTM calculations.

Using typical profiles requires an algorithm to select them according to some criteria, and each approach (model results, magnitude of measured NO_2 column, data bases on emissions, fires etc.) will fail under certain conditions.

For the GOME-2 project, the second (model based) option was selected mainly for practical reasons and monthly averaged profiles were created in this project from a 4 year MOZART T63 run.

When using such a data set, a number of problems have to be kept in mind, namely that

- current global models do not have the spatial resolution needed to resolve the horizontal variability of NO_2 on the scale of satellite pixels,
- current models do not agree on the absolute amount and vertical distribution of NO_2 in the troposphere,
- daily variability of NO_2 and its vertical distribution is large where transport plays a role, limiting the usefulness of climatological profiles,

A solution of these problems will only be possible with much improved models and their validation in coming years.

4 Vertical Sensitivity

4.1 Introduction

The vertical sensitivity depends on a number of parameters, including measurement geometry and solar position, surface albedo, aerosol loading and surface altitude. The impact of these parameters on the retrieval of tropospheric NO₂ has been discussed in previous publications (*Richter and Burrows, 2002, Martin et al., 2002, Boersma et al., 2004*). Still, it is worthwhile to re-visit the sensitivity as a function of aerosols, surface albedo and surface altitude.

4.2 Effect of Aerosols

Not all retrieval approaches for tropospheric NO₂ include the effect of aerosols (e.g. *Boersma et al., 2004*), and those who do, come to different conclusions on how large the effect is (*Ni β , 2005*). Here, we show the results from a large number of sensitivity runs with the SCIATRAN RTM model (*Ro \acute{z} anov et al., 1997, Ro \acute{z} anov et al., 2001*) using the LOWTRAN (*Shettle and Fenn, 1976*) parameterisation scheme. Four different aerosol types are considered (none, rural, maritime, urban) and for the latter three different visibilities.

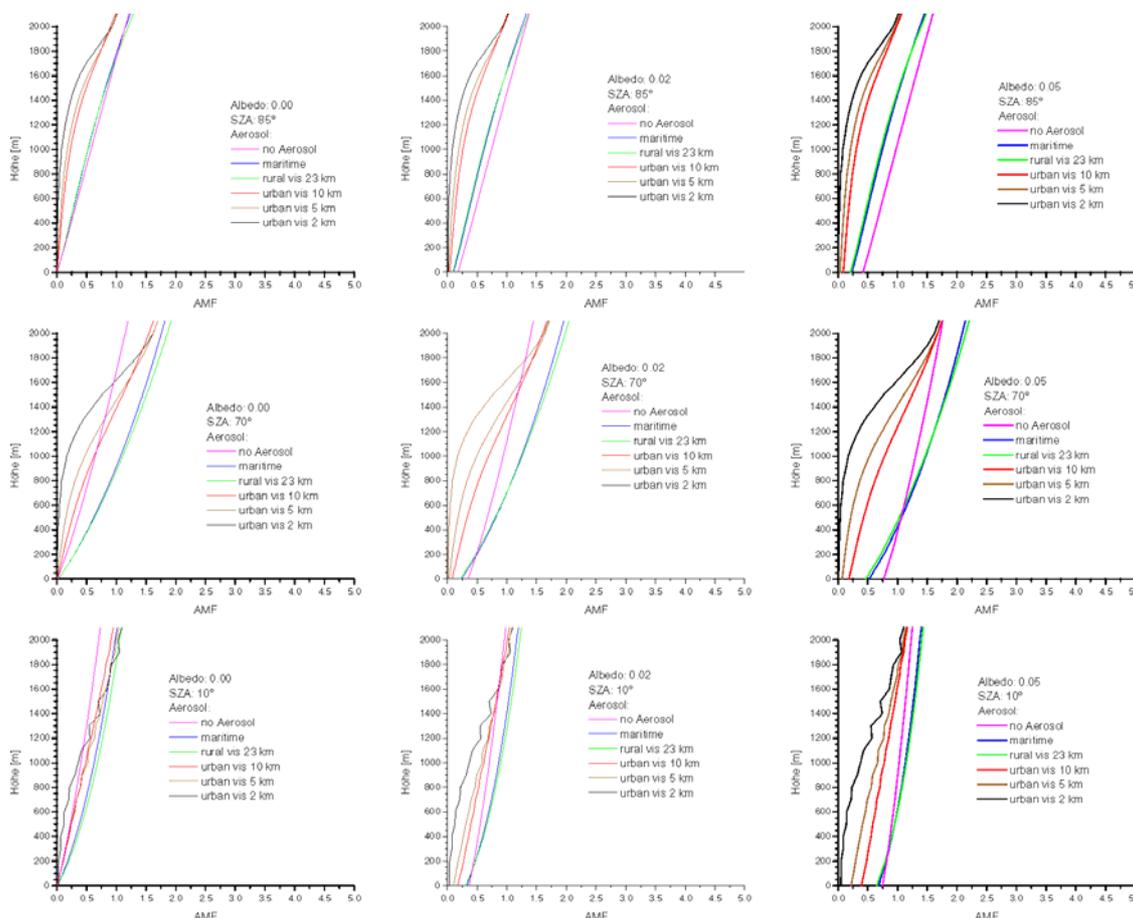


Fig. 14. The effect of aerosols on NO₂ airmass factors in the lowest 2 km. The SZA decreases from top to bottom (85°, 70°, 10°) while surface albedo increases from left to right (0, 2%, 6%). Colours indicate the results for different aerosol scenarios, all taken from the LOWTRAN parameterisation in SCIATRAN. The oscillations in the urban aerosol scenario at high sun result from the phase function approximation for this aerosol type and are not expected in the real atmosphere.

The results are given in Figure Fig. 14 and can be summarized as follows.

- In all cases, the sensitivity decreases significantly towards surface. In contrast to the situation for stratospheric absorbers, the sensitivity decreases towards larger solar zenith angles as result of the larger scattering probability and extinction.
- An increase in surface albedo by a few percent leads to a strong increase in sensitivity in the lowest few hundred meters, in particular for the cases without aerosol or with maritime or rural aerosol. The effect of surface albedo is discussed in more detail in the next section.
- A thick aerosol “hides” lower layers just as a thin cloud would do, and where large NO₂ concentrations close to the ground are combined with high aerosol loading as e.g. in the industrialized parts of China, the information content of the measurements is low and the results depend largely on the a priori assumptions.
- Maritime and rural aerosols which have reflective properties decrease the measurement sensitivity within and below the aerosol layer but actually increase it above. This is similar to the effect of thin clouds. Absorbing aerosols on the other hand result in reduced sensitivity throughout the atmosphere. This is of particular relevance as little is known on the absorption and scattering properties of aerosols, and changes in e.g. power plant technology can change the typical aerosol composition.

In summary, the effect of aerosols on the retrieval can be very large, in particular at low sun, over dark surfaces and for urban aerosol. These results are in contrast to those published by *Martin et al.*, (2003) who found only a moderate change (20%) in their tropospheric NO₂ columns when switching from no aerosol to full aerosol. The reasons for this discrepancy are not yet understood but are one topic in an ongoing comparison exercise between the GOME NO₂ products produced at the University of Bremen, the KNMI, the University of Heidelberg and the SAO/Dalhousie University.

4.3 Effect of Surface Albedo

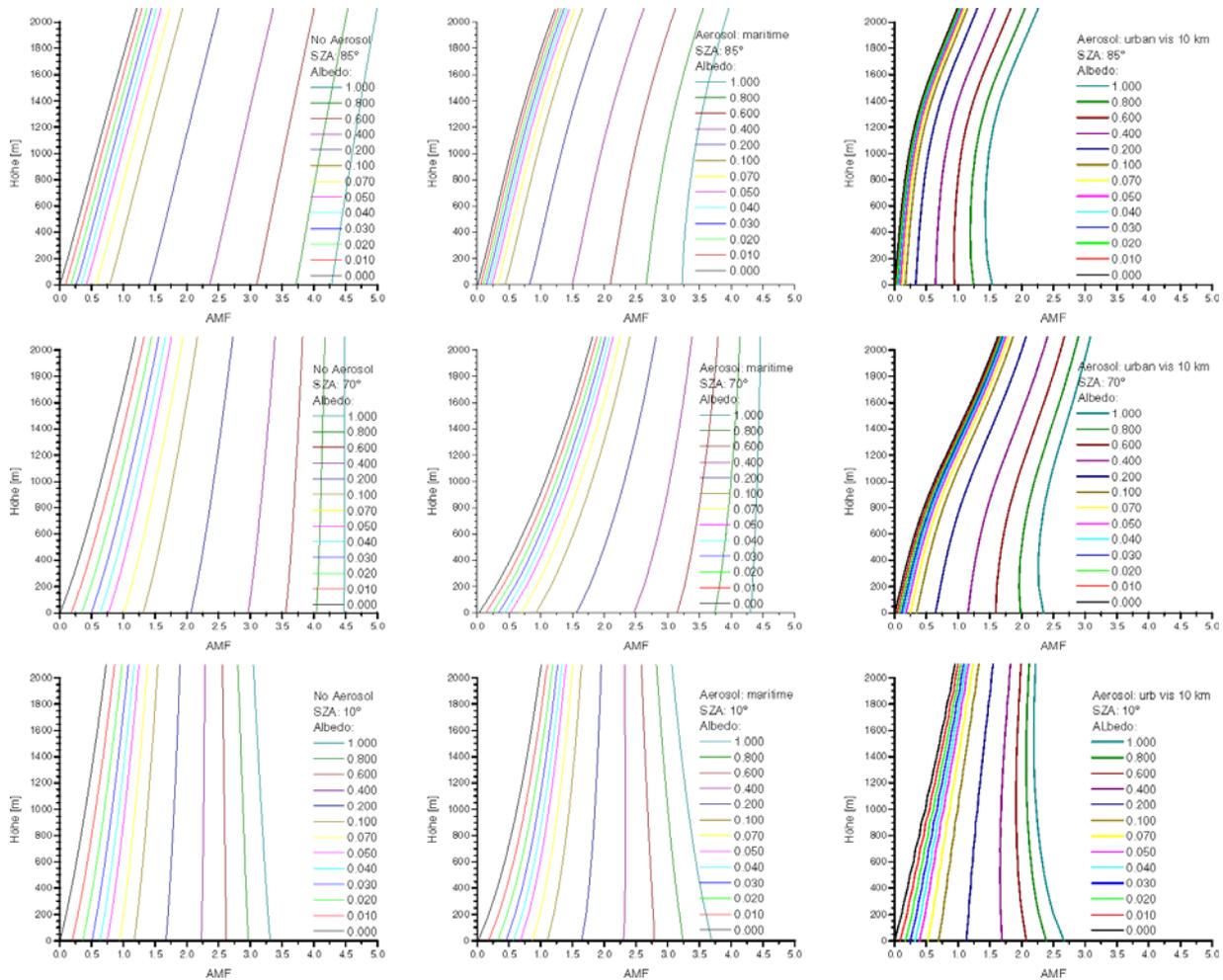


Fig. 15: Dependence of air mass factor on surface reflectivity. The solar zenith angle decreases from top to bottom (85° , 70° , 10°). From left to right, the aerosol scenarios change from no aerosol to maritime aerosol to urban aerosol with 10 km visibility. Colours indicate different surface albedo (0..1)

As already pointed out in the last section, surface albedo is a critical parameter in air mass factor calculations. In single scattering approximation, high albedo can double the number of photons close to the surface, and multiple scattering can further amplify the effective absorption. The effect on the air mass factor is however larger than a factor of two as can be understood when considering the extreme case of the lowest layer over a completely dark surface. Even if attenuation in the atmosphere is small and the incoming intensity is large, the magnitude of the absorption signal from the lowest layer observed from space is determined by the absorption in the layer and the relative contribution to the total intensity. If the surface is dark, only those photons will be able to reach the satellite which are scattered below the absorber which is a small number for the lowest layer. If in contrast the surface has a large albedo, most photons are reflected back and the absorption in the surface layer contributes with similar weight to the total signal as absorption in other altitudes.

The effects of albedo are shown in more detail in Fig. 15 and can be summarized as follows.

- In all cases, an increase in albedo results in an increase in sensitivity to the lowest 2 km for the reasons discussed in the last paragraph.

- At high sun, the sensitivity curve can actually be inverted with altitude with larger sensitivity close to the surface than at higher altitudes. This is the result of multiple scattering which is largest where density is largest.
- The effect is less pronounced at larger optical depth (high aerosol loading) as result of the increased extinction.

These results highlight the importance of the albedo input for the airmass factor calculations, and given the fact that these are at best monthly climatologies at moderate spatial resolution identify albedo as one of the main error sources wherever surface albedo is varying in space or time.

For measurements over clouds, the same arguments apply as for measurements over bright surfaces, and sensitivity can be much enhanced. This is particularly relevant for low clouds or fog with a higher probability of NO₂ above them but also for reflecting aerosols which behave similar to thin clouds. In practice, it is not always possible to separate persistent aerosol from enhanced surface albedo in the data bases as they rely on finding minimum values in long time series of atmospheric reflectance and correction of the Rayleigh component only. In that sense, errors in the surface albedo input might actually compensate errors from aerosols.

4.4 Effect of Surface Altitude

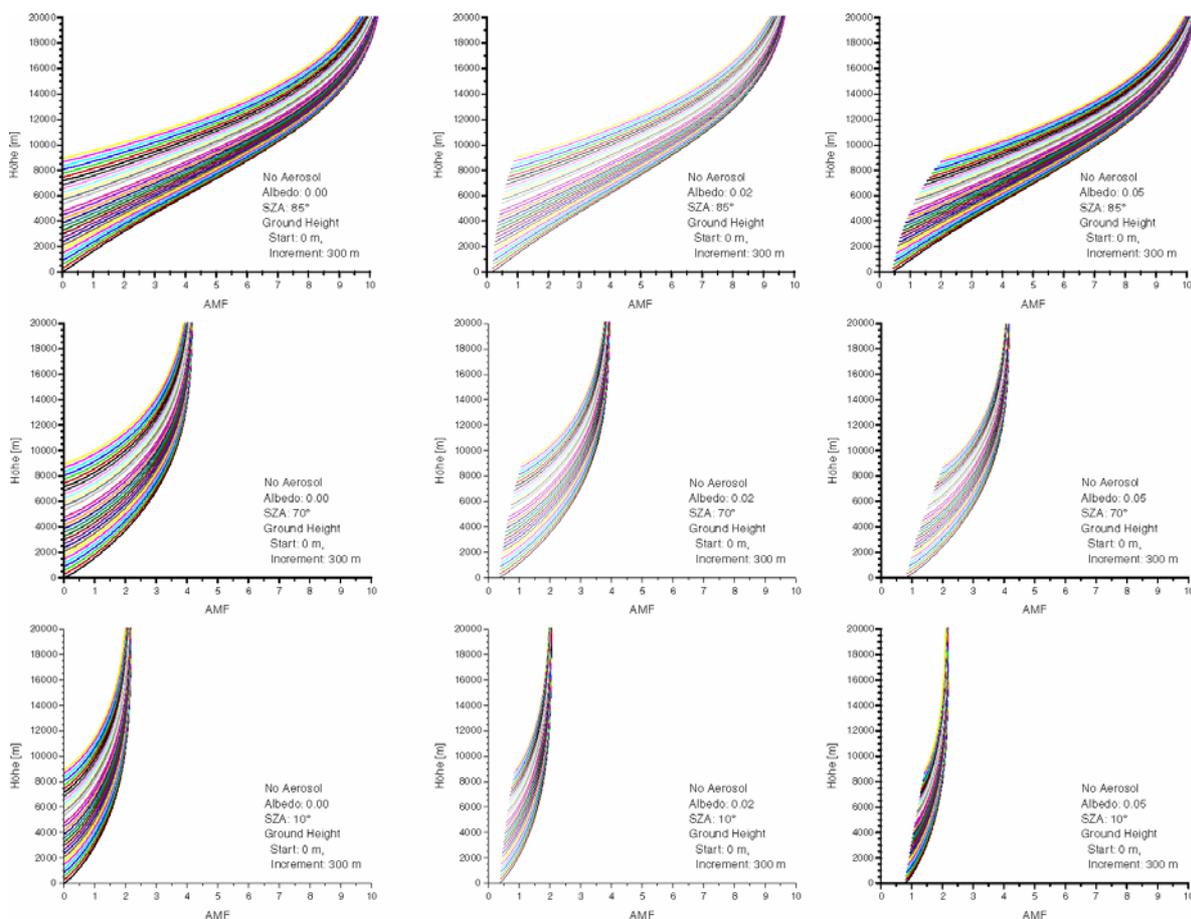


Fig. 16: Dependence of airmass factor on surface altitude. From top to bottom, the solar zenith angle decreases (85°, 70°, 10°), from left to right the surface albedo increases (0, 2%, 5%). Colours indicate different surface altitudes (0 ..9 km). No aerosols were included in these simulations.

Another parameter which has a significant impact on airmass factors is surface altitude. The main effect is that at the same altitude in the atmosphere, less photons will be scattered from below over elevated surfaces and therefore the sensitivity of the measurement will be decreased. This is illustrated in Fig. 16. As can readily be seen, the measurement sensitivity over high altitude areas is in general reduced. The effect is less pronounced over bright surfaces as in these cases scattering plays a smaller role. Also, at high sun the importance of surface altitude is smaller.

In contrast to other parameters, surface altitude is known with high accuracy and can be implemented at high spatial resolution and with good accuracy. However, problems arise from two sides: First of all, the a priori profiles are available only at low spatial resolution, and often topography changes within one model grid box. In the model, this is represented by an average pixel surface altitude and as a result, the a priori profile will often not extend to the lowest surface altitude actually present in the area. Numerically, this can be fixed by simple linear extrapolation or extension of the lowest profile value, but in reality, at least the lower part of the profile is probably scaling with altitude above ground rather than absolute altitude. A similar problem arises from the fact that also the satellite pixel often covers an area with varying surface altitude, and using the average altitude not necessarily provides the right answer in the airmass factor calculation. Ideally, weighted results from subpixels at the resolution of the topographical data base should be used, but in view of the resulting error (which is comparatively small) and the price in computational time this probably should not have high priority.

4.5 Summary and Recommendations

In summary, the effects of aerosols and surface albedo on the airmass factor are large and can easily amount to 50% errors in individual measurements.

For aerosols, this is in contrast to previous findings and from our sensitivity studies, it is highly recommended to include aerosols in the airmass factor calculations. As the effect of aerosols also depends critically on type, vertical profile and optical depth a good climatology or on-line model simulations would be needed for detailed correction. If on-line calculations are not available, measurements of aerosol OD and speciation derived from measurements from the same satellite would provide the necessary input for aerosol distribution and in combination with vertical distributions from climatology could greatly improve the treatment of aerosols.

With respect to surface albedo, the existing data bases are too inaccurate, have too low spatial resolution and also do not capture variability over a month and between years. While on average the values are probably reasonable, individual measurements can be off significantly as result of local or short term changes. As in the case of aerosols, independent measurements of surface albedo from the same satellite would much improve the situation as would retrieval from the GOME-2 measurements themselves. This has so far not been demonstrated for tropospheric species but is part of e.g. the WFDOAS retrieval scheme of ozone columns which also retrieves an effective surface reflectivity. This of course relies on accurate radiometric calibration of the spectra.

An additional problem with surface albedo might arise from the large line of sight viewing angles used by GOME-2. The existing albedo data bases are limited to nadir measurements, and depending on surface type and solar position, the effective albedo might be quite different at the edges of the scan (e.g. for trees, cities, crop or other 3-d structures). Therefore, in principle the BRDF of the surfaces also needs to be known.

5 Error Study

5.1 Introduction

In order to facilitate a realistic assessment of the uncertainties in the final GOME-2 total column product, the effect of different approaches for the treatment of clouds and the a priori NO₂ profiles in the troposphere has been analysed. In addition, an end-to-end comparison of different tropospheric NO₂ products derived from GOME measurements has been performed.

Different treatments of clouds can be used from a simple cloud screening to using cloud fraction information from the PMDs and cloud top from O₂ absorption (OCRA/ROCINN) to using cloud fraction and cloud top height from the O₂ band (FRESCO). Those options considered for NO₂ retrieval from GOME-2 data have been compared.

For the airmass factor calculation, monthly averaged airmass factors for a realistic case i.e. the average airmass factor that results for a month of real GOME data have been compared to daily airmass factors and airmass factors based on monthly averages of NO₂ profiles. These values not only depend on the tabulated airmass factors, but also on the measurement pattern, the cloud fields and the treatment of clouds in the retrieval. Two months have been selected, January and June 1997 as they represent summer and winter situations in the two hemispheres.

As the last point, an end-to-end comparison of three different tropospheric NO₂ products provided by KNMI / BIRA on the TEMIS web site, the IUP Bremen product and the result of the DLR analysis using the settings foreseen for the GOME-2 processor was undertaken. This comparison has been performed on GOME data from 2000, but again only results for January and June are presented here.

5.2 Comparison of different cloud treatments

In the DLR and TEMIS retrieval schemes, the tropospheric air mass factor for a partly cloudy scene is determined with the independent pixel approximation (IPA). The tropospheric air mass factor AMF_{trop} is written as a linear combination of a cloudy and a clear-sky air mass factor:

$$AMF_{trop} = w \cdot AMF_{cloudy} + (1 - w) AMF_{clear} \quad (3)$$

where AMF_{cloudy} is the air mass factor for a completely cloud-covered scene, and AMF_{clear} is the air mass factor for a cloud-free scene. The radiance-weighted cloud fraction w is the fraction of the photons that originates from the cloudy part of the scene ($w = f \cdot I_{cloudy} / I$), and depends on the cloud fraction f , the surface- and cloud-albedo, and the viewing geometry [Martin *et al.*, 2002]. The air mass factor for a completely cloud-covered scene is obtained from Eq. (2), with $m_l = 0$ below the cloud-top. For the DLR and TEMIS products (see also section 5.4), the cloud fraction and cloud top height have been obtained from FRESCO [Koelemeijer *et al.*, 2001]. For the calculation of the monthly mean NO₂ fields, only scenes with a cloud radiance fraction < 50% are used (corresponding to a cloud fraction of about 0.2).

With the cloud screening method, as used in the Bremen retrieval scheme, only nearly cloud-free pixels are used with a FRESCO cloud fraction less than 0.2, similar to the scheme described above. However, with the cloud screening method, the cloud is neglected for fractions less than 0.2, while with the IPA method, the influence of small cloud fractions on the radiative transfer is explicitly taken into account via Eq. (3).

As described by *Boersma et al.* [2004], the uncertainty in the cloud fraction is one of the most important error sources in the retrieval of the tropospheric NO₂ column. For polluted situations, *Boersma et al.* [2004] found uncertainties up to 30% in the tropospheric NO₂ column as a consequence of the uncertainty in the cloud fraction. This analysis was based on a reported uncertainty in the FRESCO cloud fraction of 0.05 [*Koелеmeijer et al.*, 2001]. However, comparison between different cloud products derived from GOME measurements [*Tuinder et al.*, 2004] shows much larger differences, and an uncertainty of 0.10 seems more realistic at least at high cloud fractions leading to even larger impact of the cloud treatment on the overall uncertainty of the NO₂ columns.

Here, the effect of the uncertainty in the cloud fraction on the retrieved tropospheric NO₂ columns for polluted conditions has been studied by using two different GOME cloud products. First, tropospheric NO₂ columns have been calculated with the DLR retrieval scheme using FRESCO cloud fractions and cloud-top heights. Then, the NO₂ calculations have been redone with the same retrieval scheme, but now using cloud fractions and cloud-top heights from the OCRA/ROCINN algorithm [*Loyola and Ruppert*, 1998; *Loyola*, 2004]. In both cases, only scenes with a cloud radiance fraction < 50% have been used. The main difference between the two cloud algorithms is that in OCRA, the cloud fraction is determined from GOME-PMD measurements, while the FRESCO algorithm uses GOME radiances in the O₂ A-band to retrieve the cloud fraction. It should be noted that differences between the FRESCO and ROCINN cloud-top heights also effect the tropospheric NO₂ retrieval. However, the sensitivity of the air mass factor to the cloud height is generally very small over polluted areas. This is due to the fact that for most scenes, the cloud tops are well above the pollution layers [*Boersma et al.*, 2004]. Another difference between FRESCO and ROCINN is that the latter retrieves cloud albedo while in FRESCO a fixed value of 0.8 is used which is appropriate for optically thick clouds. While this is a good approximation for the retrieval of stratospheric ozone, it is less good for the retrieval of tropospheric NO₂ where low and thin clouds are more relevant which often have quite different albedo.

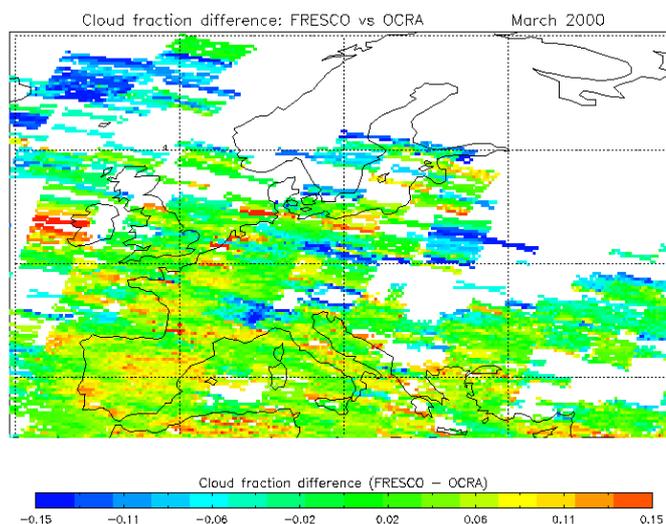


Fig. 17: Difference between FRESCO and OCRA cloud fractions for March 2000. Only cloud fraction data from GOME measurements used in the tropospheric NO₂ column calculations have been taken into account, i.e. only for scenes with a cloud radiance fraction < 50%

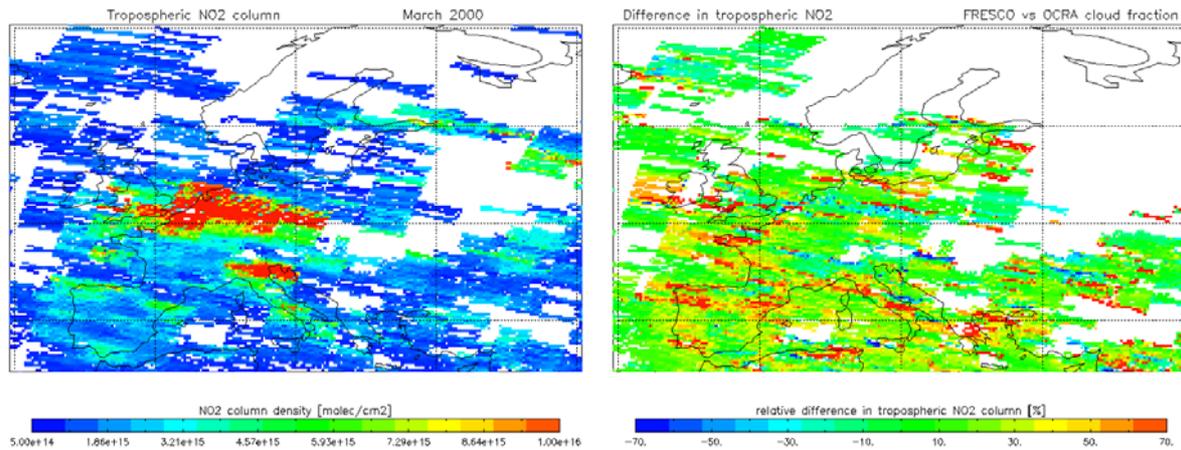


Fig. 18: Tropospheric NO₂ column retrieved with the DLR scheme using the FRESKO cloud fractions for March 2000 (left), and the relative difference between the tropospheric NO₂ columns retrieved with the FRESKO and OCRA cloud fractions using the same retrieval scheme (right).

Fig. 17 shows the difference between the FRESKO and OCRA cloud fractions for the European area for March 2000. It is important to note that in this figure, only cloud fraction data from the GOME measurements that are used in the tropospheric NO₂ column calculations, have been taken into account, i.e. only scenes with a cloud radiance fraction < 50%. As shown in Fig. 17, absolute cloud fraction differences of 0.05 and larger are often observed over Europe. The FRESKO cloud fractions are mostly higher than the OCRA cloud fractions, but negative differences are also observed for some areas. We do not attempt to quantify the accuracy of the FRESKO or OCRA cloud fractions from these comparisons, but this analysis indicates that for small cloud fractions, an assumed uncertainty in the cloud fraction of 0.05 should be regarded as a minimum value.

Fig. 18a shows the tropospheric NO₂ column retrieved with the DLR scheme using the FRESKO cloud fractions for the same area and month. Fig. 18b shows the difference between the tropospheric NO₂ columns retrieved with the FRESKO and OCRA cloud fractions, using the same retrieval scheme. Fig. 17 and Fig. 18b demonstrate the strong correlation between the cloud fraction differences and the tropospheric NO₂ column differences: larger cloud fractions usually result in larger tropospheric NO₂ columns. This is to be expected, since for polluted conditions, the cloudy air mass factor AMF_{cloudy} is generally smaller than the clear-sky air mass factor AMF_{clear} , resulting in smaller tropospheric air mass factors for larger cloud fractions (Eq. (3)). As a result, the tropospheric NO₂ columns retrieved with the FRESKO cloud fraction are mostly larger than the ones retrieved with the OCRA cloud fraction. Fig. 18b illustrates that the differences in the tropospheric NO₂ column can be larger than 30%.

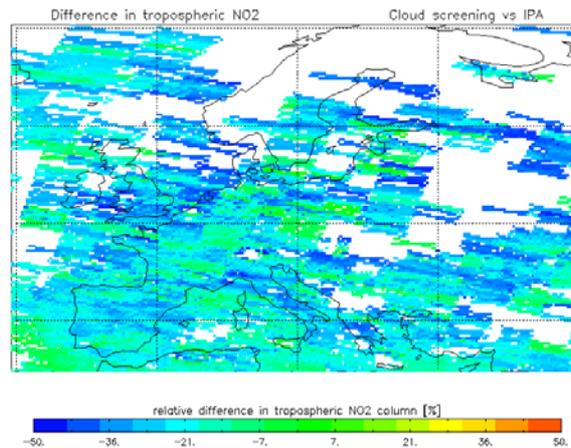


Fig. 19: Difference between the tropospheric NO₂ columns retrieved with the cloud screening and the IPA method for March 2000. In both retrievals, FRESCO cloud fraction have been used and only scenes with a cloud radiance fraction < 50%.

The method for calculating the tropospheric air mass factor (the cloud screening method and the IPA method), has an important effect on the retrieved tropospheric NO₂ column as well. Fig. 19 shows the difference between the tropospheric NO₂ columns retrieved with the cloud screening and the IPA method for the European area for March 2000. Note that for both retrievals, FRESCO cloud fractions have been used, and only scenes with a cloud radiance fraction < 50%. As can be seen in Fig. 19, the tropospheric NO₂ columns retrieved with the cloud screening method are systematically smaller than the ones retrieved with the IPA method by about 15-40%. With the cloud screening method, the cloud is neglected in the air mass factor calculation for small cloud fractions (as explained above), and only a clear-sky air mass factor AMF_{clear} is used. For polluted conditions, this generally results in larger tropospheric air mass factors and smaller tropospheric NO₂ columns.

The main conclusion of these analyses is that the treatment of clouds in the air mass factor calculations and the accuracy of the cloud fraction have a large impact on the retrieved tropospheric NO₂ column. To reduce the uncertainty in the tropospheric NO₂ column, it is important to further improve the cloud retrieval schemes and to continue the validation of the cloud parameters, especially for low cloud fractions.

5.3 Comparison of different Airmass Factors

For the analysis of the sensitivity of the airmass factors to different approaches to the vertical NO₂ profile used, five different airmass factor sets were available. For all the cases, only the model profiles used and the way they were averaged have been varied, but not the other input parameters such as aerosol fields, surface albedo or cloud treatment. These settings were all kept fixed as described in *Richter et al.*, 2005 and in more detail in *Nüß*, 2005, the main points being:

- Albedo: GOME climatology from (*Koelemeijer et al.*, 2003)
- Aerosol: 3 types (maritime, rural, urban at two visibilities based on sea / land type and EDGAR-2 CO₂ emissions (*Nüß*, 2005))
- Block-AMF calculated using the radiative transfer model SCIATRAN (*Rozanov et al.*, 2001)
- Topography: TerrainBase (*Row et al.*, 1994)
- Clouds: FRESCO (*Koelemeijer et al.*, 2001), screening for cc ≤ 20% only

- Stratosphere: SLIMCAT (*Chipperfield*, 1999) correction scaled on the Pacific sector
- all inputs were first gridded to the lowest spatial resolution (that of the model profiles) although the other inputs are available at higher spatial resolution. This point is discussed in more detail in section 3.2.5.

Using these settings, the following airmass factor sets were created:

1. The **IUP Bremen standard airmass factors** which are based on an old MOZART-2 run for 1997 at T43 resolution where the airmass factors were computed from daily profiles for each grid cell but then averaged over months to create „typical“ values. This data set is used in the IUP Bremen tropospheric NO₂ product which has been used in several publications and is available on the web at <http://www.iup.physik.uni-bremen.de/doas>. It serves as a comparison standard to link the new data sets with previous work.
2. **Daily airmass factors based on the new MOZART-2 T63** daily profiles. This is the run used for the profiles climatology for the GOME-2 tropospheric NO₂ airmass factors. For each location and each day, an individual profile and thus an individual airmass factor are available. This is considered to be the “best” possible set of airmass factors and serves as the reference case.
3. **Monthly airmass factors using MOZART-2 T63** profiles. This data set corresponds to case 1., only that the new model profiles are used.
4. **MOZART-2 T63 airmass factors based on monthly profiles.** In this data set, the monthly averaging was performed not on the airmass factors, but already on the profiles. This is not necessarily equivalent and conceptually, it is not clear if one should prefer to use an average airmass factor (as in cases 1. and 3.) or an airmass factor based on the average profile. However, from the point of view of implementation in the operational GOME-2 processor, this approach has large advantages.
5. **MOZART-2 T63 airmass factors based on climatological monthly profiles.** This data set is equivalent to case 4., only that the profiles have not only been averaged over months, but over all respective months of the MOZART run. This makes the profiles closer to a climatological mean profile for e.g. January. See section 3.3 for details.

In order to compare the different sets of airmass factors in a realistic way, the following approach has been taken: One month of real GOME NO₂ data have been analysed using the respective AMF set with cloud screening ($\leq 20\%$ FRESCO cloud cover) and binned on a 0.5 x 0.5 degree grid. However, instead of computing the average NO₂ vertical column, the average airmass factor has been determined for each pixel. Ratios between different data sets have then been taken on the gridded results with gaps where there were no data e.g. as result of persistent clouds or snow and ice (which are also flagged as cloudy). This approach differs from that taken in section 3.2.7, where AMF at a fixed SZA were compared. The values shown here are more representative of what happens in the real data evaluation, but at the same time depend on the exact measurement pattern and the assumed cloud fractions. In another month, with another satellite instrument or with another cloud detection algorithm, the results will be slightly different.

It should also be noted that as result of the non-linear nature of the monthly average of vertical columns, the monthly averaged vertical column is not the ratio of monthly averaged slant column and monthly averaged AMF:

$$\langle SC_i / AMF_i \rangle \neq \langle SC_i \rangle / \langle AMF_i \rangle$$

As high NO₂ columns and small airmass factors are strongly correlated (at least if the model provides a realistic NO₂ field), the results can deviate quite strongly in areas with variable NO₂. This is closely related to the problem of limited sampling by the satellite measurement and the error introduced by monthly averaged airmass factors (see section 3.2.7 and below).

5.3.1 Change in model version

The first test is concerned with the effect of a change in model version used for the vertical NO₂ profiles used in the airmass factor calculations. This comparison is motivated by the fact, that the NO₂ profile climatology produced in this project is based on a different MOZART-2 model run than the one used for the airmass factors applied in the IUP Bremen NO₂ product. Also, a recent comparison of a large number of state of the art CTM models of the troposphere showed large differences between individual models although they were using the same emissions and meteorology (*van Noije, 2005*).

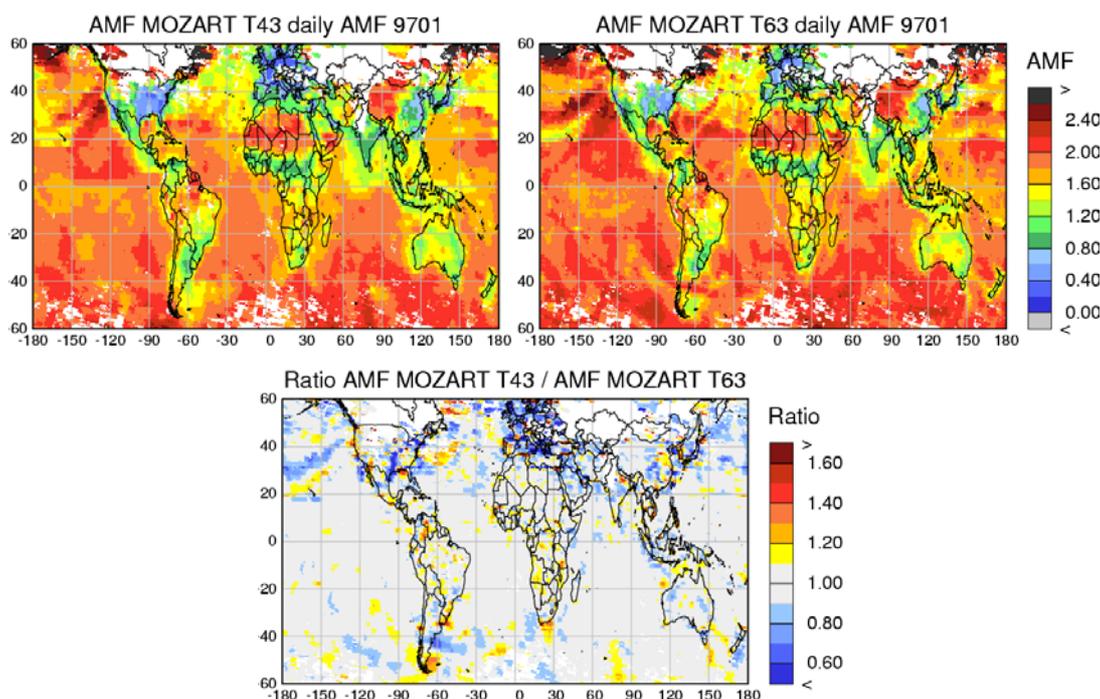


Fig. 20: Comparison of airmass factors calculated for GOME measurements January 1997 with the standard T43 MOZART run (left) and the new T63 run (right). The lower panel shows the ratio of the two upper ones.

Here, two model runs of the same CTM (MOZART-2) were compared that differ in many ways, most notably the spatial resolution (T63 vs. T43), the emission scenarios used (POET vs. RETRO) and also aspects of the model parameterisations. The results are shown in Fig. 20 and Fig. 21 for January and July, respectively. Overall, the airmass factors look very similar from both model runs, but in the ratio, differences of up to a factor of 2 can occur. Closer inspection of the figures reveal, that three different effects contribute to these differences:

1. the change in resolution which leads to changes in the representation of the meteorological patterns (e.g. in the Southern Hemisphere) which result in different airmass factors
2. the change in resolution which provides more details in T63 and therefore in areas of large gradients results in large differences
3. a less confined boundary layer which leads to less accumulation in polluted areas in particular in winter (Fig. 21) and therefore to larger airmass factors.

Many of the relatively large deviations are seen over the ocean or in remote regions, where the absolute effect on the tropospheric NO₂ is small, and also one can expect the effects to cancel to a large degree over time. However, the change in boundary layer treatment is systematic and leads to a considerable change in the seasonality of the computed columns, the new MOZART-2 profiles yielding a smaller seasonality. This is in agreement with the findings of *van Noije et al. (2005)* who reported systematic differences in the seasonality over urban areas between different GOME NO₂ products, the IUP Bremen data having the largest. The MOZART-2 data used in that study are similar to the T63 run used here.

Considering that it is not obvious which of the model runs is giving the more accurate representation of the real atmosphere, the large changes observed can be interpreted as indication of the degree of uncertainty that results from this error source alone even if state of the art model profiles are used. The magnitude of the uncertainties seen in this study are at least locally clearly larger than the model profile uncertainty of 10% derived by *Boersma et al., (2004)* who relied on an analysis of one single TM5 model run.

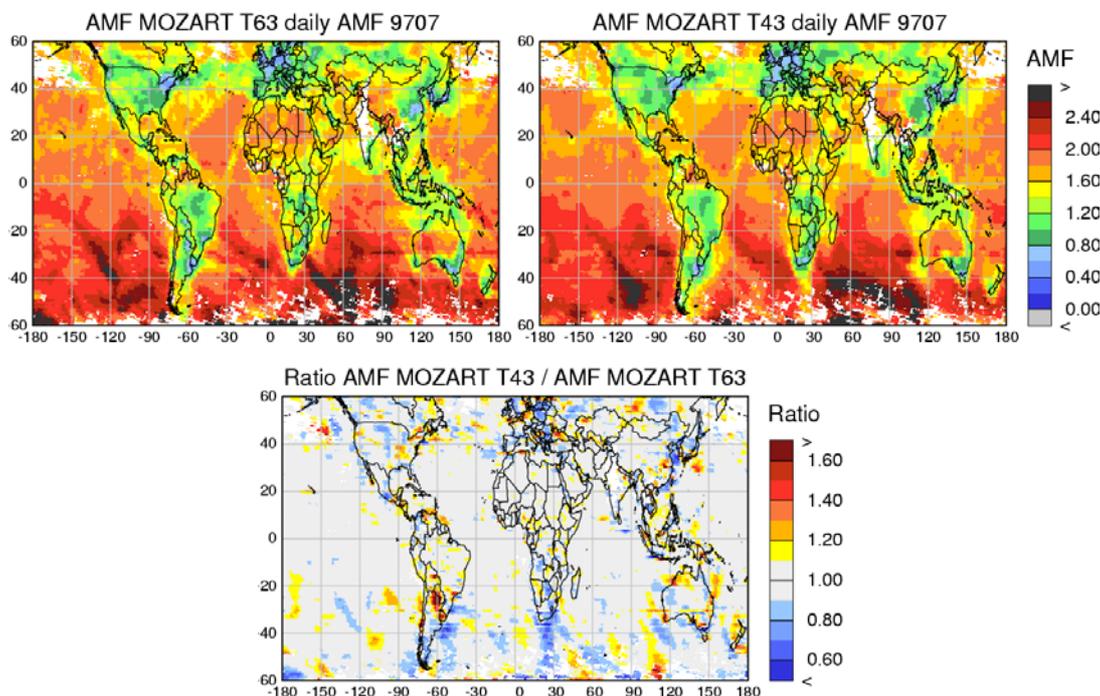


Fig. 21: Comparison of airmass factors calculated for GOME measurements July 1997 with the standard T43 MOZART run (left) and the new T63 run (right). The lower panel shows the ratio of the two upper ones.

5.3.2 Change from daily to monthly AMF

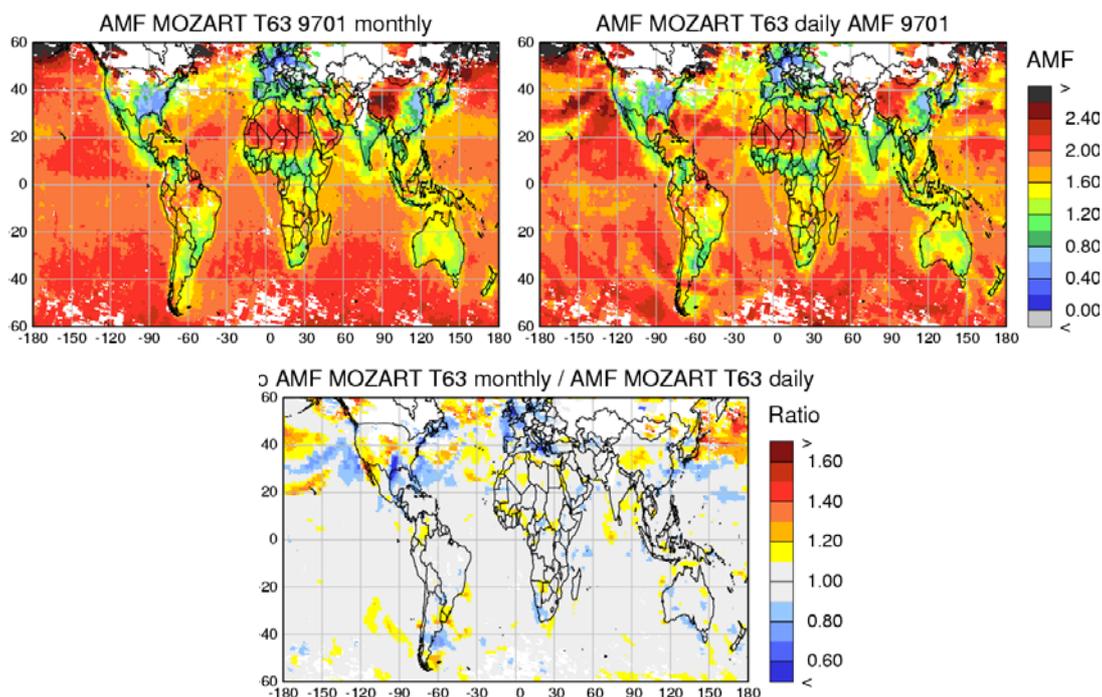


Fig. 22: Comparison of airmass factors calculated for GOME measurements January 1997 with AMF averaged over one month (left) and daily AMF (right). The lower panel shows the ratio of the two upper ones.

As already briefly discussed in section 3.2.7, the use of monthly averaged airmass factors leads to some uncertainties in the retrieved NO_2 columns. As shown in Fig. 22 for January 1997, the averaged airmass factors are much more homogeneous and more clearly show the underlying emission patterns, e.g. in the case of shipping emissions. This is to be expected and in fact the idea behind using averaged fields, as a typical value instead of a specific one is used.

The differences shown in Fig. 22 can locally be as large as 30%, but usually are much smaller in polluted areas. The reason for the differences are mainly sampling effects, as with a global coverage within 3 days for each pixel only a subset of all daily model profiles is applied in the “daily” case, whereas all profiles contribute to the “monthly” case.

5.3.3 Change from monthly AMF to monthly profiles

In the IUP-Bremen NO_2 retrieval, monthly averaged airmass factors have calculated from a MOZART-2 run for 1997 and tabulated for use with GOME and SCIAMACHY data. This method has been chosen to be independent from online model results and also to be independent from any trends or drifts that might be present in model data. The method uses averages of airmass factors instead of airmass factors based on averaged profiles as this provides a better approximation of the real average value over a month in the data analysis. However, if clouds are to be incorporated in the radiative transfer calculations, the vertical profile is needed for each satellite pixel, and the approach taken in Bremen can not be applied. Therefore, monthly averages of atmospheric profiles are first computed and the resulting AMF compared with those discussed in the last section.

As shown in Fig. 23, the differences between the two methods are generally small, with the exception of areas along the coasts which are affected by occasional outflow of pollution from

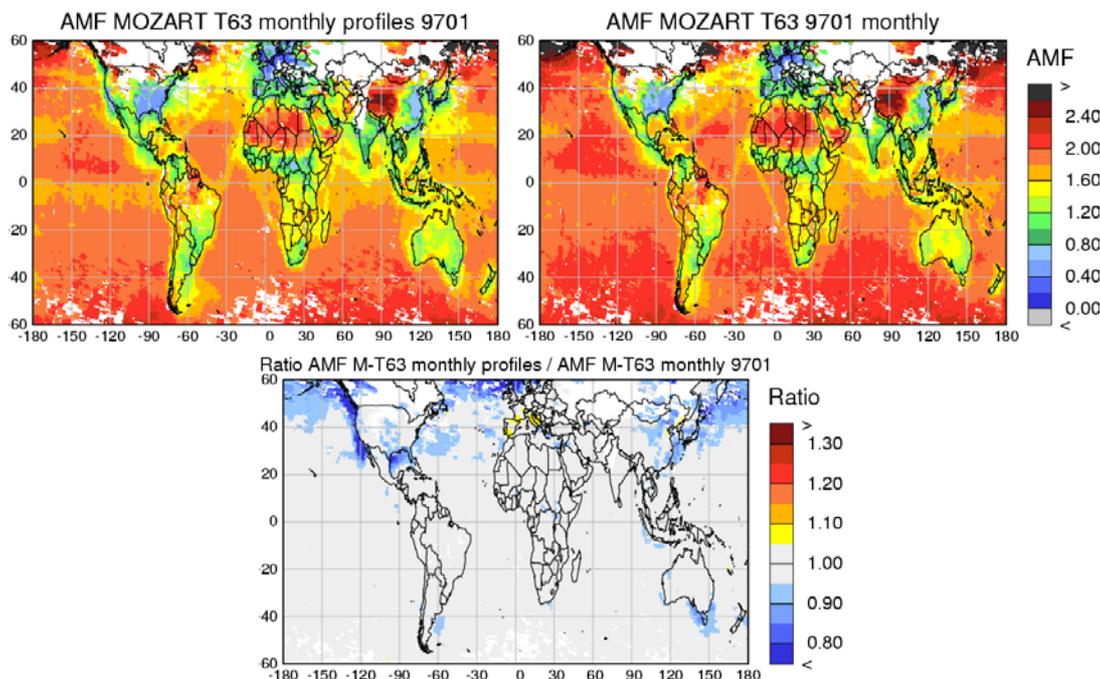


Fig. 23: Comparison of airmass factors calculated for GOME measurements January 1997 with profiles averaged over one month (left) and airmass factors averaged over one month (right). The lower panel shows the ratio of the two upper ones. Please note the change in colour scale.

the continents. Also, at high latitudes differences of up to 15% are observed, probably as result of profile changes linked to stratospheric variability in the winter hemisphere. In nearly all cases, the averaged profiles lead to smaller airmass factors.

The differences seen can be understood by the non-linear effect of the vertical profile on the airmass factor: Assuming that in a background region the air is usually clean, but on one day a transport event is bringing strongly polluted air to the location. The airmass factor will be large on the clean days, but small on the polluted day. On average, the clean days dominate with little (1/30) impact of the polluted (small) airmass factor. If, on the other hand, profiles are averaged, then the very large total amount of NO₂ in the boundary layer on the polluted day will have a large effect on the averaged profile even though it is only observed on one day. While the airmass factor does not depend on the total NO₂ column, the averaged profile does. Thus, averaging profiles will lead to airmass factors weighted towards high NO₂ situations.

5.4 Comparison of different tropospheric NO₂ products

In this section, different GOME NO₂ products are compared on an end-to-end level to assess the consistency of the results and to analyse the effect of some specific algorithm settings. An overview over the settings used in the different retrieval schemes is given in Table 4. To make the IUP Bremen and DLR products more comparable, the IUP Bremen data set has been re-computed using the MOZART-2 T63 run which also is the basis for the climatology used in the DLR product. A detailed comparison of different models and data products for tropospheric NO₂ has recently been performed by *van Noije et al.* (2005), the results of which are not repeated here. However, some specific points are treated in more detail in addition to a comparison to the results obtained using the settings foreseen for the GOME-2 operational processor.

	TEMIS	IUP	DLR-OCRA	DLR-FRESCO
Slant Columns	IASB	Bremen	GDP-4	GDP-4
T-correction	YES	NO	YES	YES
Stratosphere	assimilation	SLIMCAT + scaling to Pacific	masking + smoothing	masking + smoothing
Clouds	FRESCO + correction	FRESCO, only screening	OCRA / ROCINN + correction	FRESCO + correction
RTM model	DAK	SCIATRAN	LIDORT	LIDORT
Aerosols	none	3 types	none	none
Albedo	GOME/TOMS LER	GOME LER	GOME/TOMS LER	GOME LER*
NO₂ profile	TM5	MOZART T63 1997 monthly averages	MOZART T63 climatology	MOZART T63 climatology
Pollution Threshold	no	no	10 ¹⁴ molec cm ⁻²	10 ¹⁴ molec cm ⁻²

Table 4: Overview on the settings used for the different GOME NO₂ products compared in this section. * GOME/TOMS LER is intended as baseline but only data with GOME LER were available at time of writing for technical reasons. This also improves comparability between the Bremen and DLR products.

The first comparison is for January 2000 and shown in Fig. 24. A number of observations can be made from the plots:

1. the absolute columns of the IUP Bremen and TEMIS products agree on average, although locally significant differences exist. The DLR product has much lower values overall
2. the IUP Bremen columns are larger over anthropogenic emission regions but smaller over areas with biomass burning
3. the DLR product is lower overall, but larger over background regions and in the area of the Southern Atlantic Anomaly.

As previous comparisons have shown that the slant columns of IASB, IUP Bremen and GDP-4 are in good agreement, the differences observed are mainly the result of different airmass factors. In part, the differences observed can be understood:

1. the IUP Bremen products assumes high aerosol loading over anthropogenic pollution, and therefore uses small airmass factors which lead to large retrieved columns.
2. the DLR product used an threshold of 10¹⁴ molec/cm² for the pollution correction, and unpolluted measurements were not included in the averaged tropospheric fields. As discussed in section 3.2.1, this will lead to a high bias, in particular in remote regions. As aerosols are not included in the RTM calculations, lower values are expected over source regions. For the GOME-2 operational product, it is foreseen to use a threshold of 0 molec/cm² as in the OMI operational data.
3. the IUP Bremen product uses cloud screening only and no correction for partially cloudy pixels which will often lead to an underestimation of NO₂ in polluted regions.

The difference in the TEMIS and IUP-Bremen products over areas with biomass burning has also been noted in *van Noije et al. (2005)*, and still is under examination.

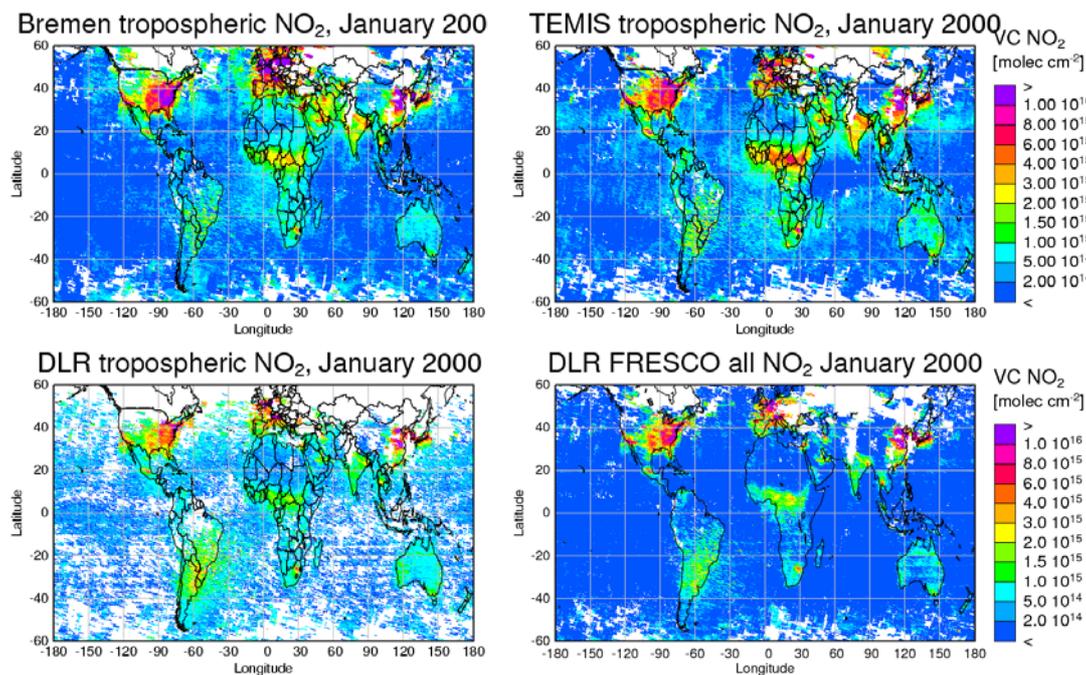


Fig. 24: Comparison of tropospheric NO₂ columns derived from GOME measurements in January 2000. Top left: Bremen analysis, top right KNMI/BIRA analysis from the TEMIS web site. Bottom left: DLR standard analysis (OCRA/ROCINN with pollution threshold), Bottom Right: DLR product with FRESCO cloud fraction and all values included).

More relevant in the context of this study is the question, why the DLR product is so much lower than the IUP Bremen product although the airmass factors are based on similar vertical profiles from MOZART-2. In principle, the opposite should be expected as the DLR product accounts for the effect of clouds in partially cloudy pixels which increases the retrieved columns. However, this effect can be offset by neglecting the effect of aerosols. It can also not be excluded that the LIDORT airmass factors are generally larger than those computed with SCIATRAN.

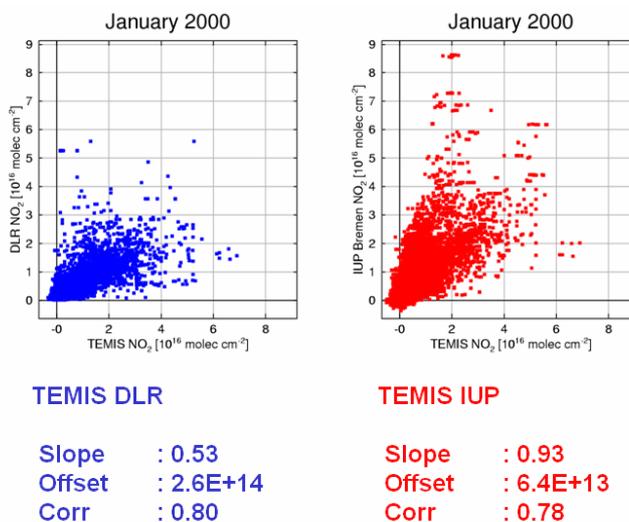


Fig. 25. Correlation plots between different versions of the tropospheric NO₂ for January 2000

Another view of the same comparison is presented in Fig. 25. All three data sets show correlation coefficients of the order of 0.8, but while TEMIS and IUP Bremen have a similar absolute value, the DLR product is about a factor of two lower. Also apparent are the cut-off at 10^{14} molec cm^{-2} in the DLR product and some residual high values which are from the Southern Atlantic Anomaly area (SAA). The scatter is large in both comparisons, and in the case of the TEMIS / IUP Bremen comparison can be traced to different behaviour of different areas - good agreement over Asia, a shift in pattern over the US and generally lower TEMIS values over Europe.

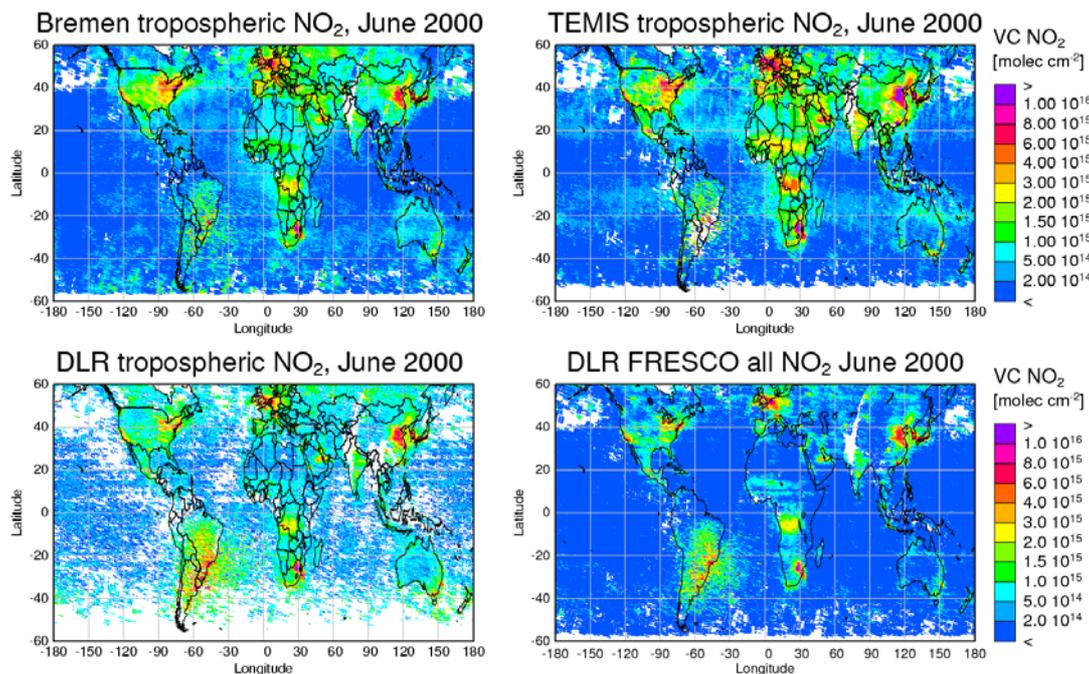


Fig. 26: As Fig. 24, but for June 2000

In Fig. 26, the three NO_2 products are shown for June 2000. For this month, the IUP Bremen product is clearly lower than the TEMIS data set, but still somewhat larger than the DLR product. The difference between TEMIS and IUP Bremen data products could have several reasons:

1. the temperature correction which has not been applied to the IUP Bremen data is more important in summer in the NH where the bulk of the NO_2 pollution is located, leading to an underestimation of the IUP Bremen product. However, as discussed in section 1.4, this effect does not exceed 25% for extreme cases.
2. the correction for partial cloud cover should also increase the TEMIS product
3. the effect of aerosols which are corrected for in the IUP Bremen data set is much smaller in summer, resulting in less cancellation of effects

There is however no clear reason why the DLR product should again be lower than the IUP Bremen data set; as pointed out above, the opposite should be the case considering the difference in cloud treatment and temperature correction. The only remaining explanation (apart from RTM model differences) is the neglecting of aerosols in the DLR product or a systematic difference in cloud treatment.

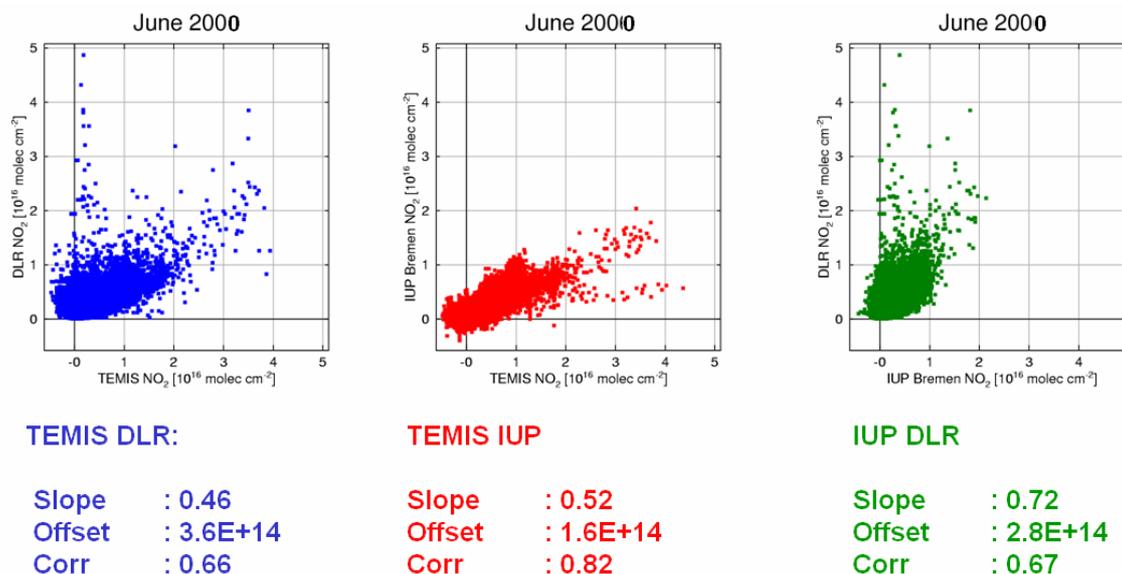


Fig. 27: As Fig. 25, but for June 2000

In Fig. 27, correlation plots are shown for the different data sets for June. The situation differs from that in January in that the TEMIS / IUP Bremen correlation is high and has little scatter but TEMIS data are nearly a factor of two larger. TEMIS data can be expected to be larger as result of the cloud treatment but a factor of two is surprising and must be related to differences in the vertical profiles used for the AMF.

The agreement between IUP Bremen and DLR product is better in absolute terms, the DLR columns being larger than the IUP columns. The scatter however is much larger than for the comparison with TEMIS which is reflected in the lower correlation coefficient. The slope computed for the comparison is affected by the application of a pollution threshold as will be discussed in more detail below with Fig. 29.

In order to investigate the impact of different cloud treatment on the NO₂ columns, three different approaches have been compared (see Fig. 28):

1. the first scenario, where only “polluted” pixels were included and OCRA / ROCINN cloud information was used
2. a scenario including both polluted and unpolluted pixels and using OCRA / ROCINN cloud information. For pixels flagged as unpolluted, a 0 value was used in the averaging.
3. a scenario including both polluted and unpolluted pixels but FRESCO cloud fractions were applied

As can be seen from the figure, the scatter is reduced when the pollution threshold is set to 0 (please note that these are monthly averages, and reducing the threshold will decrease many individual DLR values but no sharp selection is obvious as would be the case for an individual day). When using the FRESCO cloud algorithm, the agreement is much improved and many outliers disappear leading to a correlation coefficient of 0.77. This is in line with the results already discussed in section 3.2.6 where substantial differences were found when using the two cloud retrievals. As the TEMIS product is also based on FRESCO cloud values, the overall consistency is best for this set-up. It should however be pointed out that this does not imply that OCRA/ROCINN is worse than FRESCO as no absolute standard is available.

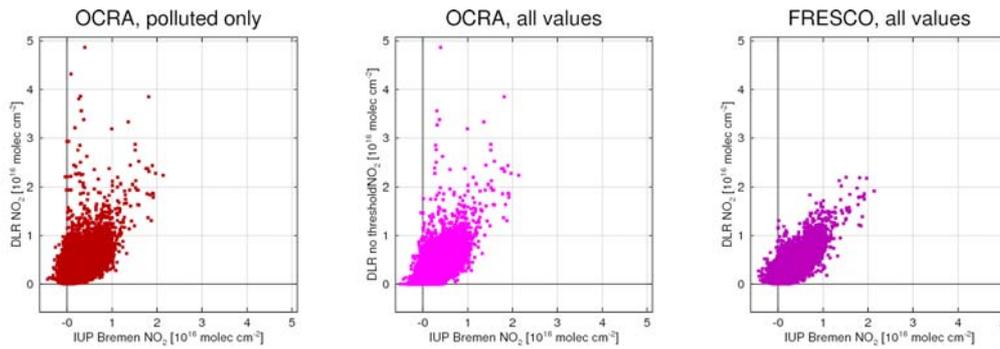


Fig. 28: Correlation plots between the IUP Bremen data (x-axis) and different versions of the DLR product for June 2000. From left to right: baseline (OCRA, threshold 10^{14} molec cm^{-2}), OCRA (threshold 0 molec cm^{-2}), FRESCO (threshold 0 molec cm^{-2})

A quantitative comparison of the two retrievals is complicated by the effect of the threshold application in the DLR product. While visual inspection indicates good agreement between DLR FRESCO and IUP Bremen, and the correlation coefficient is about 0.77, the slope of a linear regression is smaller than 0.7. The reason for this unexpected result is illustrated in Fig. 29, where averaged data is shown on top of all data points. As result of the offset application, a “hockey stick” shape is created, which forces a linear regression line to the wrong side. While for the purpose of this comparison the main result (good agreement between DLR FRESCO and IUP Bremen data set in summer) can be deduced from the figure, this problem will have to be kept in mind when in the future GOME-2 operational products are compared to other satellite data.

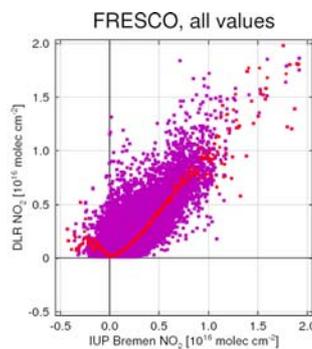


Fig. 29: Zoom in of Fig. 28c with the averaged columns overplotted. While the overall ratio is close to 1, the curvature close to the origin and the turn over for negative values has a large impact on a linear regression and reduces the apparent slope to 0.7. This is a consequence of the threshold applied in the averaging of the DLR columns.

A complementary view of the results already discussed is the frequency distribution of the absolute differences between different products. This is shown in Fig. 30.

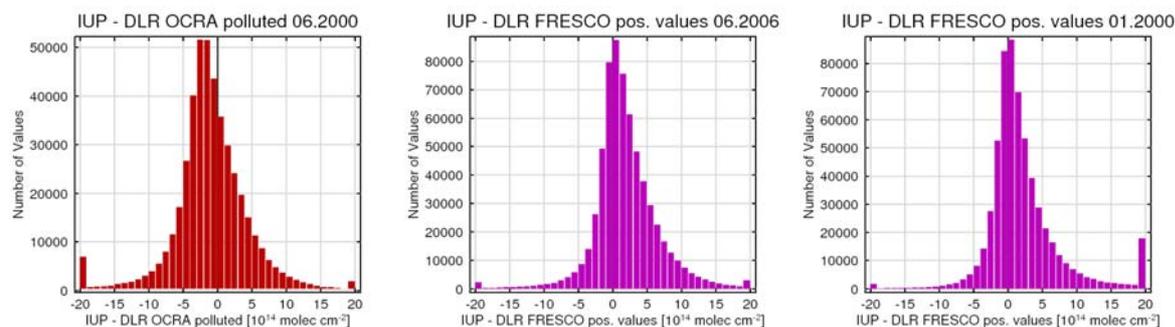


Fig. 30: Frequency distribution of differences between the IUP Bremen and DLR NO₂ columns for June 2000 (left and middle) and January 2000 (right). For June, both the results for the baseline product and the FRESCO product with threshold 0 molec cm⁻² are shown.

In all cases, the differences have well centred distributions with a FWHM of the order of 3 – 6 10^{14} molec cm⁻² which is somewhat larger than the differences of about 1×10^{14} molec cm⁻² cited for different slant columns by *Boersma et al.* (2004). Considering the large additional uncertainties introduced by airmass factors and cloud treatment, these results are quite encouraging. Also, the centre of all the distributions differs from 0 only by several 10^{14} molec cm⁻² which again is a good result. However, as is evident from the large number of extreme values shown at the edges of the figures (all values have been accumulated in the last bin which are outside of the range of values shown) and the skewed form of the distribution this is only part of the story. The narrow part of the distribution is dominated by the low values over clean regions but over polluted regions, systematic differences exist in particular in January (the Bremen retrieval being higher). As this is a multiplicative effect from the airmass factors, it doesn't show up very clearly in the absolute differences shown in Fig. 30.

Comparison of the OCRA and FRESCO figures again highlights the large impact the choice of cloud treatment has on the product, mainly visible in a general offset and a broader distribution for the OCRA/ROCINN product but also in a number of outliers. It is very probable that a large part of the remaining difference in the comparison between DLR-FRESCO and IUP data product are also related to the different cloud treatment (IUP Bremen uses FRESCO data for screening only but does not apply any correction for partially cloudy scenes).

5.5 Summary and Recommendations

In summary, relatively large differences were found between the different GOME tropospheric NO₂ products, in agreement with the results of *van Noije et al.* (2005). Both absolute values and the spatial distribution are different, in particular in winter. This is the result of the many different assumptions that went into the individual retrievals, which can change the final columns by up to a factor of 2. This is the case for the comparison of the DLR baseline product with TEMIS in both seasons and with the IUP Bremen product in winter. By using more similar assumptions, e.g. by replacing the OCRA cloud product with the FRESCO product, the agreement between DLR and IUP Bremen products can be improved as expected.

Not unexpectedly, the following parameters turned out to be of large importance for the absolute value of the tropospheric columns:

- Using monthly instead of daily airmass factors has a significant impact on individual days but on average leads to comparable results. Similarly, the use of airmass factors

based on averaged profiles instead of averages of airmass factors has an impact, but seems acceptable in view of the large advantages for implementation.

- The choice of cloud treatment has a large impact on both the absolute value and the distribution of the tropospheric NO₂ columns as demonstrated by comparing the DLR products with OCRA/ROCINN and FRESKO, respectively. Using a cloud screening only also has a significant effect and probably leads to underestimation of the columns.
- The choice of GOME or GOME/TOMS LER makes a large difference in the airmass factors and therefore retrievals, which is in line with the theoretical studies discussed in section 4.3.
- The use of aerosols in the airmass factor calculations also has a large impact in particular over polluted regions in winter when the sun is low (compare Fig. 14). This explains the large difference between DLR and IUP product in January. The University of Bremen team believes that ignoring aerosols leads to large errors in the airmass factors in particular for absorbing aerosols. Therefore, the availability of accurate aerosol data from GOME-2 (or from other MetOp instruments) is of great importance.
- The choice of a threshold in the DLR product has a significant impact on correlations with other products, and has to be considered very carefully. In the opinion of the University of Bremen team, no such threshold should be applied (threshold of negative infinity) to avoid artificial offsets and differences to other products.

For the most relevant input parameters, no clear recommendation can be given based on the results of this study. A more robust estimate of tropospheric NO₂ from satellite measurements will only be possible if the input parameters used (surface reflectance, aerosol scenarios, cloud fractions and altitudes, vertical distributions of NO₂) are validated separately with external data sources so that a decision can be made on which of the different inputs available best to use. A mere validation of the end product (tropospheric NO₂ columns) can contribute to this end, but will always be of limited value as cancellation of errors often occurs.

6 Acknowledgements

The authors would like to thank M. Chipperfield for providing the stratospheric NO₂ columns from SLIMCAT, ESA / DLR for providing GOME lv1 data, and the OMI team for discussions and reports which were very useful for this study. The MOZART model runs used for the NO₂-"climatology" and the airmass factors were provided by Claire Granier and Ulrike Niemayer from MPI Hamburg. Some of the software and results used in this were developed within the European project POET and through funding of the University of Bremen.

7 References

- Beirle, S., Platt, U., Wenig, M. and Wagner, T., Weekly cycle of NO₂ by GOME measurements: a signature of anthropogenic sources, *Atmospheric Chemistry and Physics*, Vol. 3, pp 2225-2232, 2003
- Beirle, S., 2004. Estimating source strengths and lifetime of Nitrogen Oxides from satellite data. *PhD Thesis, University of Heidelberg, 2004.* http://archiv.ub.uni-heidelberg.de/volltextserver/volltexte/2005/5225/pdf/diss_komplett.pdf
- Boersma et al., 2002: Boersma, K. F., E. J. Bucsela, E. J. Brinksma, and J. F. Gleason, OMI Algorithm Theoretical Basis Document, K. Chance (Ed.), *Volume IV OMI Trace Gas Algorithms - Chapter 2, NO₂, ATBD-OMI-04*, Version 2, August, 2002.
- Boersma, K.F., H.J. Eskes and E.J. Brinksma, Error Analysis for Tropospheric NO₂ Retrieval from Space, *J. Geophys. Res.*, **109** D04311, doi:10.1029/2003JD003962, 2004
- Bogumil, K., J. Orphal, T. Homann, S. Voigt, P. Spietz, O.C. Fleischmann, A. Vogel, M. Hartmann, H. Bovensmann, J. Frerik and J.P. Burrows, Measurements of Molecular Absorption Spectra with the SCIAMACHY Pre-Flight Model: Instrument Characterization and Reference Data for Atmospheric Remote-Sensing in the 230-2380 nm Region, *J. Photochem. Photobiol. A.*, **157**, 167-184, 2003
- Bucsela, E.J., E.A. Celarier, M.O. Wenig, J.F. Gleason, J.P. Veefkind, K.F. Boersma and E.J. Brinksma, *Algorithm for NO₂ Vertical Column Retrieval From the Ozone Monitoring Instrument*, IEEE Trans. Geo. Rem. Sens., 2006, Vol. **44**, No. 5, 1245-1258, doi:10.1109/TGRS.2005.863715.
- Burrows, J. P., Dehn, A., Deters, B., Himmelmann, S., Richter, A., Voigt, S., and J. Orphal, Atmospheric remote-sensing reference data from GOME: Part 1. Temperature-dependent absorption cross-sections of NO₂ in the 231--794 nm range, *J. Quant. Spectrosc. Rad. Transfer*, 60:1025-1031, 1998
- Chance et al., OMI Algorithm Theoretical Basis Document Volume IV OMI Trace Gas Algorithms, 2002
- Chipperfield, M. P.: Multiannual Simulations with a Three-Dimensional Chemical Transport Model, *J. Geophys. Res.*, **104**, 1781–1805, 1999.
- Greenblatt, G. D., Orlando, J. J., Burkholder, J. B., Ravishankara, A. R., Absorption measurements of oxygen between 330 and 1140 nm, *J. Geophys. Res.*, **95**(D11):18577-18582, 1990
- Koelemeijer, R. B. A., P. Stammes, J. W. Hovenier, and J. F. de Haan, A fast method for retrieval of cloud parameters using oxygen A band measurements from the Global Ozone Monitoring Experiment, *J. Geophys. Res.*, **106**, 3475-3490, 2001.
- Koelemeijer, R. B. A., de Haan, J. F., & Stammes, P., A database of spectral surface reflectivity in the range 335 – 772 nm derived from 5.5 years of GOME observations, *J. Geophys. Res.*, **108**(D2), 4070, doi:10.1029/2002JD002429, 2003.
- Leue, C., M. Wenig, T. Wagner, U. Platt, and B. Jähne, Quantitative analysis of NO_x emissions from GOME satellite image sequences, *J. Geophys. Res.*, **106**, 5493-5505, 2001.
- Loyola, D., and T. Ruppert, A new PMD cloud-recognition algorithm for GOME, *ESA Earth Observation Quarterly*, **58**, 45-47, 1998.
- Loyola, D., Automatic Cloud Analysis from Polar-Orbiting Satellites using Neural Network and Data Fusion Techniques, IEEE International Geoscience and Remote Sensing Symposium, 4, 2530-2534, Alaska, 2004.
- Martin, R. V., K. Chance, D. J. Jacob, T. P. Kurosu, R. J. D. Spurr, E. Bucsela, J. F. Gleason, P. I. Palmer, I. Bey, A. M. Fiore, Q. Li, R. M. Yantosca and R. B. Koelemeijer, An improved retrieval of tropospheric nitrogen dioxide from GOME, *J. Geophys. Res.*, **107**, 4437-4456, 2002.
- Martin, R. V., D. J. Jacob, K. Chance, T. P. Kurosu, P. I. Palmer and M. J. Evans, Global inventory of nitrogen oxide emissions constrained by space-based observations of NO₂ columns, *J. Geophys. Res.*, **108**, 4537-4548, 2003.
- J. H. Nüß: An improved tropospheric NO₂ retrieval for GOME and SCIAMACHY, PhD Thesis, University of Bremen, 2005 (in German)
- Richter, A.: Absorptionsspektroskopische Messungen stratosphärischer Spurengase über Bremen, 53° N, *PhD-Thesis, University of Bremen*, June 1997 (in German)
- Richter, A. and J. P. Burrows, Retrieval of Tropospheric NO₂ from GOME measurements, *Adv. Space Res.*, **29**, 16673-1683, 2002.

- Richter, A., Burrows, J. P., Nüß, H., Granier, C., Niemeier, U., Increase in tropospheric nitrogen dioxide over China observed from space, *Nature*, **437**, 129-132, doi: 10.1038/nature04092, 2005
- Row; Hastings and Dunbar, *Terrainbase worldwide digital terrain data. Release 1.0*, National Geophysical Data Center, Boulder, 1994.
- Rozanov, V., Diebel, D., Spurr, R. J. D. & Burrows, J. P., GOMETRAN: A radiative transfer model for the satellite project GOME - the plane parallel version, *J. Geophys. Res.*, **102**, 16683-16695, (1997).
- Rozanov, A., Rozanov, V., and Burrows, J. P.: A numerical radiative transfer model for a spherical planetary atmosphere: Combined differential-integral approach involving the picard iterative approximation. *J. Quant. Spectrosc. Radiat. Transfer*, **69**, 491-512, 2001.
- Sanders, R. W., Improved analysis of atmospheric absorption spectra by including the temperature dependence of NO₂, *J. Geophys. Res.*, **101**(D15):20945-20952, 1996
- Shettle and Fenn, Models of the atmospheric aerosols and their optical properties, *AGARD Conference Proceedings on optical propagation in the atmosphere*, 1976.
- Stohl, A., Hittenberger, M., and Wotawa, G.: Validation of the Lagrangian particle dispersion model FLEXPART against large scale tracer experiment data, *Atmos. Environ.*, **32**, 4245-4264, 1998.
- Tuinder, O. N. E., deWinter-Sorkina, R., Builtjes, P. J. H., Retrieval methods of effective cloud cover from the GOME instrument: an intercomparison, *Atmos. Chem. Phys.*, **4**, 255-273, 2004
- Vandaele A.C., C. Hermans, P.C. Simon, M. Carleer, R. Colin, S. Fally, M.F. Merienne, A. Jenouvrier, and B. Coquart, Measurements of the NO₂ absorption cross-section from 42000 cm⁻¹ to 10000 cm⁻¹ (238-1000 nm) at 220 K and 294 K, *J.Q.S.R.T.*, **59**, 171-184, 1998
- van Noije, T. P. C., Multi-model ensemble simulations of tropospheric NO₂ compared with GOME retrievals for the year 2000, *submitted to Atmos. Chem. Phys Discussins.*, 2005
- Velders, G. J. M., C. Granier, R. W. Portmann, K. Pfeilsticker, M. Wenig, T. Wagner, U. Platt, A. Richter and J. P. Burrows, Global Tropospheric NO₂ Column distribution: Comparing three-dimensional model calculations with GOME measurements, *J. Geophys. Res.*, 106, 12643-12660, 2001.
- Wang, P., A. Richter, M. Bruns, V. V. Rozanov, J. P. Burrows, K.-P. Heue, T. Wagner, I. Pundt, U. Platt, Measurements of tropospheric NO₂ with an airborne multi-axis DOAS instrument, *Atmos. Chem. Phys.*, **5**, 337-343, 2005