Supplement of

Improved spectral fitting of nitrogen dioxide from OMI in the 405–465 nm window

J. H. G. M. van Geffen et al.

Correspondence to: J. H. G. M. van Geffen (geffen@knmi.nl)
S1 DOAS retrieval of NO$_2$ slant column densities

The DOAS procedure minimises the difference between the measured spectrum $R_{\text{meas}}(\lambda)$ and a modelled spectrum $R_{\text{mod}}(\lambda)$ within a given wavelength window, in the form of minimisation of a chi-squared merit function:

$$\chi^2 = \sum_{i=1}^{N_{\lambda}} \left( \frac{R_{\text{meas}}(\lambda_i) - R_{\text{mod}}(\lambda_i)}{\Delta R_{\text{meas}}(\lambda_i)} \right)^2 ,$$

(S1)

with $N_{\lambda}$ the number of wavelengths in the fit window and $\Delta R_{\text{meas}}(\lambda_i)$ the precision of the measurements.

The measured reflectance $R_{\text{meas}}(\lambda)$ is given by:

$$R_{\text{meas}}(\lambda) = \frac{\pi I(\lambda)}{\mu_0 I_0(\lambda)} ,$$

(S2)

with $I(\lambda)$ the radiance at top-of-atmosphere, $I_0(\lambda)$ the extraterrestrial solar irradiance spectrum, and $\mu_0 = \cos(\theta_0)$ the cosine of the solar zenith angle (the viewing geometry dependence of $I$ is omitted for brevity). Here $I(\lambda)$ and $I_0(\lambda)$ share the same wavelength grid, i.e. an appropriate wavelength calibration has been applied prior to the DOAS fit (see Sect. 4.2). The $I_0(\lambda)$ in use for the DOAS retrieval of OMNO2A is an average of the solar spectra measured by OMI using the Quartz Volume Diffusor in 2005, which is used on a daily basis.

In the OMI slant column retrieval the modelled reflectance is expressed in terms of intensities, which leads to a non-linear fit problem. This approach allows to describe the effects of inelastic scattering after a scattering event has occurred:

$$R_{\text{mod}}(\lambda) = P(\lambda) \cdot \exp \left[ - \sum_{k=1}^{N_k} \sigma_k(\lambda) \cdot N_{s,k} \right] \cdot \left( 1 + C_{\text{ring}} \frac{I_{\text{ring}}(\lambda)}{I_0(\lambda)} \right) ,$$

(S3)

with $P(\lambda)$ a polynomial of degree $N_p$, $\sigma_k(\lambda)$ the cross section and $N_{s,k}$ the slant column amount of molecule $k$ taken into account in the fit (NO$_2$, O$_3$, etc.), $C_{\text{ring}}$ the Ring fitting coefficient and $I_{\text{ring}}(\lambda)/I_0(\lambda)$ the sun-normalised synthetic Ring spectrum. The Ring spectrum describes the differential spectral signatures arising from inelastic Raman scattering of incoming sunlight by NO$_2$ and O$_3$ molecules. The last term between brackets in Eq. (S3) describes both the contribution of the direct differential absorption (i.e. the 1), and the modification of these differential structures by inelastic scattering (the $+C_{\text{ring}} I_{\text{ring}}(\lambda)/I_0(\lambda)$ term) to the reflectance spectrum.

An alternative approach used by most DOAS applications – including the QDOAS software (Danckaert et al., 2012) used by BIRA-IASB for the processing of data from GOME-2, SCIAMACHY and SAOZ – applies a fitting in terms of the optical density, i.e. the logarithm of the reflectance. This allows for a linearisation of the problem, by writing the modelled reflectance as follows:

$$\ln \left[ R_{\text{mod}}(\lambda) \right] = P^*(\lambda) - \sum_{k=1}^{N_k} \sigma_k(\lambda) \cdot N_{s,k} - \sigma_{\text{ring}}(\lambda) \cdot C^*_{\text{ring}} ,$$

(S4)

with $\sigma_{\text{ring}}(\lambda)$ the differential (pseudo-absorption) spectrum of the Ring effect and $C^*_{\text{ring}}$ its fitting coefficient, where $\sigma_{\text{ring}}(\lambda)$ is constructed from the Ring radiance spectrum $I_{\text{ring}}(\lambda)$, a reference solar spectrum $I_{\text{ref}}(\lambda)$ (which is different from the measured solar spectrum $I_0(\lambda)$ in Eq. (S3)) and a low order polynomial. DOAS applications using Eq. (S4) often also fit a non-linear offset parameter to account of atmospheric and/or instrumental stray light or residual dark current signals; such an offset parameter is not needed in the non-linear approach of Eq. (S3), because in that case the offsets are captured by the polynomial. Note that the use for the linear DOAS approach of Eq. (S4) the treatment of the measurement errors in the spectrum, and therewith the details of the NO$_2$ slant column error, is different from the non-linear approach of Eq. (S3).

A measure of the goodness of the fit is the so-called root-mean-square (RMS) error, which is defined as follows in case of the intensity fitting approach:

$$\text{RMS} = \sqrt{\frac{1}{N_{\lambda}} \sum_{i=1}^{N_{\lambda}} (R_{\text{meas}}(\lambda_i) - R_{\text{mod}}(\lambda_i))^2} ,$$

(S5)

while in the linear fitting mode of QDOAS the definition is:

$$\text{RMS}^* = \sqrt{\frac{1}{N_{\lambda}} \sum_{i=1}^{N_{\lambda}} (\ln[R_{\text{meas}}(\lambda_i)] - \ln[R_{\text{mod}}(\lambda_i)])^2} ,$$

(S6)

which is always larger than the RMS of Eq. (S5) for the spectra investigated here. The difference between the measured and modelled reflectances is usually referred to as the residual of the slant column fit:

$$R_{\text{resid}}(\lambda) = R_{\text{meas}}(\lambda) - R_{\text{mod}}(\lambda) .$$

(S7)

S2 The OMI slit function & convolution of the reference spectra

The spectral resolution of the instrument covers a finite wavelength interval, which effectively averages the incident (ir)radiance – which varies on a much finer wavelength scale – over that interval: the incoming light is convolved by the so-called instrument transfer function (ITF) or slit function, with the full-width at half-maximum (FWHM) of the slit function determining the spectral resolution of the instrument. Since the reference spectra are generally measured with a much higher spectral resolution than the OMI resolution, they have to be convolved with the instrument slit function as well for usage in e.g. a DOAS retrieval. Some reference spectra are
spectrally smooth (cf. Fig. 3), in which case a convolution will not make much difference. For consistency, however, all reference spectra for OMNO2A are convolved as described below.

S2.1 The OMI slit function

During on-ground calibrations prior to launch, the OMI slit function was measured and described by a parametrised broadened Gaussian function with a FWHM value of about 0.63 nm (Dirksen et al., 2006). The OMI slit function, designated $S$ here, is a function of wavelength and depends on the viewing angle, i.e. it differs from detector row to detector row: $S_n(\lambda_i, \lambda)$, with $n = 0, \ldots, 59$ the row number and $\lambda_i$ the wavelength of detector pixel $i$, which is counted along the flight track from south to north. For each $\lambda_i$ the slit function covers the wavelength range $\lambda = [\lambda_i - 1.5 : \lambda_i + 1.5]$ nm, where the 1.5 nm represents approximately three times the FWHM of OMI. Within this range, the slit function is given in steps of 0.01 nm.

The operational OMNO2A processor is programmed such that it is not feasible to take the row dependency in the cross sections, i.e. in the slit function, explicitly into account. To circumvent this issue an average slit function $S_a$ is defined for convolving the reference spectra, by taking an average over the slit functions of the middle 36 rows:

$$ S_a(\lambda_i, \lambda) = \frac{1}{36} \sum_{n=12}^{47} S_n(\lambda_i, \lambda), \quad (S8) $$

which is representative for most detector rows (see below).

The main plot of Fig. S1 shows the difference between the average slit function and the slit functions of selected rows for the wavelength $\lambda_i = 435$ nm: the slit functions of rows 12–47 differ by at most 1%; for the outer rows the difference w.r.t. the average is up to 4%. The top inset of Fig. S1 shows the average slit function $S_a(\lambda_i, \lambda)$ for $\lambda_i = 435$ nm. The slit function is slightly asymmetric w.r.t. the central wavelength, and this asymmetry differs from row to row, which is why the difference curves in the main plot are not symmetric.

A measure for the difference between the average slit function $S_a$ and the slit function $S_n$ of an individual row is the area of the difference curves shown in Fig. S1:

$$ A(S_n(\lambda_i)) = \int_{\lambda_i - 1.5}^{\lambda_i + 1.5} |S_a(\lambda_i, \lambda) - S_n(\lambda_i, \lambda)| d\lambda, \quad (S9) $$

where the integral is evaluated by way of a simple summation with steps of $\Delta\lambda = 0.01$ nm. The value of this area for $\lambda_i = 435$ nm is shown in the bottom inset of Fig. S1, with the differences for rows 12 and 47 marked by a filled circle; $A(S_{12}(\lambda_i))$ and $A(S_{47}(\lambda_i))$ are about 0.004. From this graph it is clear that $S_a$ is representative for most of the detector rows. For the outer 5 rows on either side $S_a$ is less representative; the slit function for rows 3 and 55, for example, differs by up to 2.5% from the average, with $A(S_3(\lambda_i))$ and $A(S_{55}(\lambda_i))$ about 0.011.

The wavelength dependency of the slit function is illustrated in Fig. S2, which shows the difference of the average slit function for selected wavelengths with the average slit function at $\lambda_i = 435$ nm (which is shown in the top inset of Fig. S1); these differences are between $-1.5\%$ for 405 nm and $+0.7\%$ for 465 nm. The curves of the differences are
nearly symmetric w.r.t. the central wavelength, which means that the asymmetry of the slit function does not vary much with wavelength.

### S2.2 Convolution of reference spectra

Convolution of a reference spectrum $\sigma_k^h(\lambda)$ for trace gas $k$ can be written as follows:

$$\sigma_k(\lambda_i) = \frac{1}{S_0(\lambda_i)} \int_{\lambda_i-1.5}^{\lambda_i+1.5} \sigma_k^h(\lambda) S_0(\lambda_i, \lambda) d\lambda,$$

(S10)

where the superscript $h$ indicates it concerns a high-resolution spectrum, and $S_0(\lambda_i)$ is the normalisation factor:

$$S_0(\lambda_i) = \int_{\lambda_i-1.5}^{\lambda_i+1.5} S_0(\lambda_i, \lambda) d\lambda.$$

(S11)

The limited spectral resolution of typical DOAS instruments leads to an interference between the absorption cross sections of the trace gases and the Fraunhofer structures in the solar spectrum $I_0(\lambda)$, and the division in Eq. (S2) does not fully remove the Fraunhofer structures, because the division and the convolution cannot be exchanged (see e.g. Platt et al. (1997)). This so-called "I$_0$-effect" can be corrected for by including a high-resolution solar reference spectrum $I_{0r}(\lambda)$ in the convolution. In the case of weak absorbers, like those relevant in the NO$_2$ fit window, Eq. (S10) is then written as follows:

$$\sigma_k(\lambda_i) = \frac{1}{S_0(\lambda_i)} \int_{\lambda_i-1.5}^{\lambda_i+1.5} \sigma_k^h(\lambda) I_{0r}(\lambda) S_0(\lambda_i, \lambda) d\lambda,$$

(S12)

with $S_0(\lambda_i)$ the normalisation factor:

$$S_0'(\lambda_i) = \int_{\lambda_i-1.5}^{\lambda_i+1.5} I_{0r}(\lambda) S_0(\lambda_i, \lambda) d\lambda.$$

(S13)

For most atmospheric absorbers the I$_0$-effect is weak and can often be neglected, but usually it is corrected for in the reference spectra used in the NO$_2$ retrieval. The $I_{0r}(\lambda)$ spectrum is taken from Dobber et al. (2008); see Sect. 4.1. And $I_{0r}(\lambda)$, computed following Eq. (S10), represents the convolved reference solar spectrum.

### S3 OMNO2A processor version numbering

The OMI NO$_2$ slant column retrieval processor OMNO2A provides the data in HDF-EOS files. These files are input for the processors that provide NO$_2$ vertical stratospheric and tropospheric column data, notably the DOMINO datasets from KNMI and NASA’s "Standart Product" (SP). The OMNO2A data files are given a version number consisting of three or four digits, where the 4th digit refers to minor bug fixes only and is ignored here.

Table S1 lists the changes in the OMNO2A versions relevant in view of the updates described in this paper. The first column gives the main version number, the second the date this version was introduced. The data range covered by the version is given in the third column, but only for the versions of which the NO$_2$ SCD data is in use for subsequent processing, which is listed in the last column of the Table. The version of the reference spectra (Sect. 4.1) is given in the fourth column and the version of the wavelength calibration (Sect. 4.2) in the fifth column.

The current OMNO2A processor runs in two branches, each with its own version number: 1.1.1 for the
DOMINO v2.0 NO₂ data products and 1.2.3 for the NASA SP v2.1 NO₂ data products. The differences between these two lies mainly in the update of the surface albedo and cloud cover data transferred to the NO₂ SCD data files as well as a number of additional HDF-EOS attributes for tracking data pixel quality; these differences do not affect the NO₂ SCD data itself. The version number “v1” used in Sect. 5, refers to these NO₂ SCD data, i.e. to the data of OMNO2A versions 1.0.5 – 1.2.3, while the version number “v2” refers to the forthcoming data with OMNO2A version 2.x.y.

For the forthcoming TROPOMI NO₂ processing it is planned to provide more information regarding the versions of the different processing elements in the meta data of the final NO₂ data product.

S4 Water vapour reference spectrum

Absorption by water vapour (H₂O vap) takes place in the form of a multitude of spectrally fine absorption lines, rather than as a smooth function of wavelength, so that Eq. (S12) cannot be simply applied to create a reference spectrum suitable for the DOAS retrieval. Instead, an effective reference spectrum for H₂O vap absorption suitable for use in the DOAS fit is determined following Sneep et al. (2013):

\[ R(\lambda) = R_0(\lambda) \left( 1 - N_{\text{H}_2\text{O}_{\text{vap}}} \cdot \sigma_{\text{H}_2\text{O}_{\text{vap}}}(\lambda) \right), \]

where \( R(\lambda) \) is a reflectance spectrum with a water vapour concentration with slant column value \( N_{\text{H}_2\text{O}_{\text{vap}}} \) and \( R_0(\lambda) \) a reflectance spectrum without water vapour.

Both \( R(\lambda) \) and \( R_0(\lambda) \) are simulated with the radiative transfer code DISAMAR (de Haan, 2011) in a line-by-line forward calculation on the basis of the high-resolution solar spectrum \( I_{\text{ref}}(\lambda) \) mentioned in Sect. 4.1 and the absorption lines of all six isotopes of H₂O vap from the HITRAN 2012 database (Rothman et al., 2013), without any other atmospheric trace gases present (obviously the main atmospheric gases O₂ and N₂ are present). The simulations use US standard atmosphere profiles (Anderson et al., 1986) for pressure, temperature and H₂O vap, where the latter has a total column amount of 4.76 \( \times \) 10²² molec/cm². In the simulations of \( R(\lambda) \) and \( R_0(\lambda) \), convolution with the OMI slit function and the \( I_{\text{ref}} \)-correction (cf. Sect. S2.2) are automatically applied.

After simulating \( R(\lambda) \) and \( R_0(\lambda) \), and using that \( N_{\text{H}_2\text{O}_{\text{vap}}} = 4.76 \times 10^{22} \) molec/cm², Eq. (S14) provides the absorption reference spectrum \( \sigma_{\text{H}_2\text{O}_{\text{vap}}}(\lambda) \). Since this spectrum results from narrow line absorption features, there is no need to subtract a low-order polynomial, as is custom to generate differential absorption spectra: \( \sigma_{\text{H}_2\text{O}_{\text{vap}}}(\lambda) \) is the absorption reference spectrum suitable for use in a DOAS retrieval.

S5 Comparison between OMNO2A and QDOAS

The NO₂ spectral fits for SCIAMACHY and GOME-2 presented in this study have been performed with the QDOAS software in fit windows and with a fitting method different from OMNO2A (cf. Table 1). A full study of the effect of the fit window limits, the degree of the low-order polynomial, and the fitting method is beyond the scope of this study. It is worthwhile, however, to obtain an estimate of the sensitivity of the NO₂ SCD to the spectral fitting approach. Such estimates are important for satellite intercomparisons and the generation of long-term seamless multi-sensor data records.

The QDOAS package (version 2.105, May 2013) is used for this study, as it is more flexible in usage than the OMNO2A processor and offers the linear fit approach. All fits are performed with QDOAS using the same v2014 (i.e. the updated) reference spectra on the OMI Pacific Ocean test orbit.

Fig. S3 shows that the difference in NO₂ SCD between the OMNO2A and QDOAS fit results in the OMI fit window 405–465 nm, using a 5th-degree polynomial (red line with filled circles) varies between \(-0.2 \text{ and } +0.1 \times 10^{15} \text{ molec/cm}^2\). The agreement between these two is therefore quite good, considering there are several differences between the processors: the fitting method differs, the Ring effect is included differently (compare Eqs. (S4) and (S3)), and the wavelength calibration of QDOAS differs from the OMNO2A wavelength calibration.
QDOAS has the option to apply a non-linear intensity fitting method instead of the linear optical density fitting method Eq. (S4), similar to the OMNO2A non-linear fitting method Eq. (S3), but with the Ring effect treated as a pseudo-absorber:

\[
R_{\text{mod}}(\lambda) = P(\lambda) \cdot \exp \left[ - \sum_{k=1}^{N_s} \sigma_k(\lambda) \cdot N_{i,k} \right]
- \sigma_{\text{ring}}(\lambda) \cdot C_{\text{ring}}^* .
\]

(S15)

The red line with open circles in Fig. S3 shows the difference between the results of this approach and the OMNO2A results, which appears to be larger than the difference with the linear fitting method of QDOAS: about \(-0.3 \times 10^{15}\) molec/cm², almost independent of latitude.

The SCIAMACHY and GOME-2 NO₂ data is retrieved in the fit window 425–450 nm, using a 3rd-degree polynomial. The difference between the OMI orbit processed with QDOAS in this manner and the OMNO2A data is shown by the blue line with squares in Fig. S3. At \(+0.2 - 0.6 \times 10^{15}\) molec/cm², the difference is clearly larger than for the OMNO2A fit window.

In their study to improve the GOME-2 NO₂ retrieval, Richter et al. (2011) apply the extended fit window 425–497 nm (they include absorption by soil in the fit, but such absorption signatures are absent over the Pacific). The measurements by the SAOZ instrument, discussed in Sect. 2.3, are analysed in a similar fitting window (Table 1). The black line with triangles in Fig. S3 shows that OMNO2A is higher by \(0.4 - 0.9 \times 10^{15}\) molec/cm² than applying a linear fit in this extended fit window.

In summary, the selection of the fit window (and with that the degree of the polynomial) and the fitting method determines the NO₂ fit results, i.e. there is no “true” NO₂ SCD, but at most a fit window and fit method specific slant column value. Judging from the curves in Fig. S3, the variability in the fit window and fit method selection introduces differences in the retrieved NO₂ SCD between \(-0.3\) and \(+0.6 \times 10^{15}\) molec/cm² (i.e. up to \(0.2 \times 10^{15}\) molec/cm² in terms of the NO₂ VCD). To know more about the "true" NO₂ SCD, independent measurements that do not depend on the DOAS technique are needed.

Something that stands out in Fig. S3 is that the differences in the areas around latitudes 20°S and 20°N differ markedly from the differences at other latitudes for those curves where QDOAS was used in the linear fitting mode, while for QDOAS’s non-linear fitting mode the differences with OMNO2A are nearly independent of latitude. The areas around latitudes 20°S and 20°N are the areas where absorption in liquid water plays a role (cf. Sect. 5.3). This may indicate that the linear fitting method deals differently with the polynomial-like signature of H₂O liq and/or O₃ and/or O₂–O₂ absorption (cf. Fig. 3) than the non-linear fitting method, possibly due to interference of the reference spectra with the DOAS polynomial.

A detailed investigation of this issue is beyond the scope of this paper, but a clue for it is visible in Fig. S4, which shows the differences of the SCDs of NO₂, O₃ and O₂–O₂ and the fit coefficient of H₂O liq between the new v2 OMNO2A results and results obtained with QDOAS in the linear (filled symbols) and nonlinear (open symbols) fitting approach for the standard OMI fit window 405–465 nm. The differences are given in arbitrary units which are constructed as follows. From each difference “OMNO2A – QDOAS” the average value is subtracted to ensure the difference lies around zero. The differences for the QDOAS linear fitting approach (filled symbols) are then scaled such that the maximum values near latitude 20°S are around one (but not exactly one, to make sure the curves do not overlap fully). The same scaling is also applied to the curves for the QDOAS nonlinear approach (open symbols), so that the differences for the two fitting approaches can be compared in magnitude.

The structure around latitudes ±20° is clearly visible in all four SCDs for the QDOAS linear approach (filled symbols). For NO₂ (red circles) the structure is fully gone at 20°N and smaller and negative at 20°S when using the nonlinear fit approach, as is also visible in Fig. S3. For both O₃ (blue squares) and H₂O liq (black diamonds) the structure is fully absent when using the nonlinear fit approach, while for O₂–O₂ (magenta triangles) the magnitude of the structure is about halved. For higher latitudes, both south and north, the differences are much smaller for the nonlinear fitting approach in QDOAS than for the linear approach, a further indication that the two fitting approaches behave differently.
The curves for the differences in the fitting coefficients for $\text{H}_2\text{O}_{\text{vap}}$ and the Ring effect (not shown) do not show clear latitudinal structures. For the Ring coefficient the two fitting approaches of QDOAS, Eqs. (S4) and (S15), give the same difference with OMNO2A (Eq. S3), showing that the $C^{*}\text{ring}$ in both approaches are essentially the same, while the difference between QDOAS and OMNO2A is related to the different implementation of the Ring effect in the fitting method. For the $\text{H}_2\text{O}_{\text{vap}}$ the difference with OMNO2A is somewhat closer to zero for the nonlinear than for the linear fitting approach.

References


