Vital improvements to the retrieval of tropospheric NO\textsubscript{2} columns from the Ozone Monitoring Instrument

J.D. Maasakkers - 0655330

R-1822-A

July 3, 2013

Supervisors : Dr. K.F. Boersma (Royal Netherlands Meteorological Institute and Eindhoven University of Technology)
Prof. Dr. Ir. G.J.F. van Heijst (Eindhoven University of Technology)
Abstract

Nitrogen oxides (NO$_x$ = NO + NO$_2$) are important trace gases in the atmosphere. They catalyze ozone production and contribute to aerosol formation, with important effects on climate and human health. Satellite observations provide an important tool to understand the effects of NO$_x$ by providing daily global measurements of tropospheric NO$_2$ columns. We present several improvements for the tropospheric NO$_2$ column retrieval algorithm for the Ozone Monitoring Instrument (OMI), the Dutch OMI NO$_2$ retrieval (DOMINO v3). First, we couple the retrieval to the Tracer Model 5 (TM5) chemistry transport model. The retrieval with TM5 results in similar patterns of tropospheric NO$_2$ but decreases overall concentrations over Europe (-10%), the US (-15%), and China (-11%) because of more realistic anthropogenic NO$_x$ emissions in the model. Secondly, the transition to TM5 allows retrieval at a higher model resolution of 1° × 1°. The higher resolution increases retrieved tropospheric NO$_2$ columns over hot spots related to higher a priori surface NO$_2$. Tropospheric columns over cities increase by up to 22% while columns over outflow areas decrease related to lower modeled surface NO$_2$. As a third innovation, the effective surface pressure calculation is improved by replacing the terrain height at the center of the OMI pixel with an average terrain height representative for the complete pixel. The new terrain height has a local effect of up to $0.5 \times 10^{15}$ molec./cm$^2$. Fourthly, we correct a temperature-dependent bias in the cloud retrieval, which reduces the retrieved cloud pressure for regions where the atmospheric temperature is below that of the mid-latitude summer climatology. The reduction in cloud pressure leads to an increase in retrieved tropospheric NO$_2$ of up to 11% over London in November because of more pronounced screening. As a last improvement, a correction of the relative azimuth angle dependence of the air mass factor gives local daily changes of up to $1.0 \times 10^{15}$ molec./cm$^2$ but averages out over time. All changes combined lead to lower global retrieved tropospheric NO$_2$ columns in this study compared to DOMINO v2 by 10% because of lower columns over outflow areas and lower overall emissions. Retrieved columns over cities generally increase, by up to 23%. This increase is mainly caused by the higher model resolution. We present an improved error parametrization for the tropospheric NO$_2$ product, reducing the average global error by 13% and errors for high tropospheric columns by up to 41%, reflecting better knowledge of albedo and cloud errors and mixing in TM. We characterize the mean error per pixel for November 2004 as $0.45 \times 10^{15}$ molec./cm$^2$ plus 42% of the tropospheric NO$_2$ column. Furthermore, we analyze the assimilation of the stratosphere by looking at the model forecast, model analysis, and satellite observation. We find that the assimilation performs well but also identify possible air mass factor calculation errors as a function of the viewing angle and solar zenith angle. Finally, we improve the representation of stratospheric NO$_3$ by implementing nudging of Odin HNO$_3$:O$_3$ ratios. Though the Odin nudging produces more realistic stratospheric HNO$_3$ concentrations, more research is recommended into the stratospheric NO$_2$ of the OMI assimilated TM.
Contents

1 Introduction 1

2 Retrieval theory 3
  2.1 Remote sensing 3
  2.2 Chemistry Transport Models 4
  2.3 Data assimilation 5

3 Retrieval framework 7
  3.1 The Ozone Monitoring Instrument 7
  3.2 Tracer Model 5 9
    3.2.1 Horizontal advection and vertical transport 10
    3.2.2 Sources and sinks 10
    3.2.3 Chemistry 11
    3.2.4 Stratospheric NO₂ in TM5 11
    3.2.5 Model performance 14
  3.3 The Dutch OMI Tropospheric NO₂ retrieval 15
    3.3.1 The NO₂ slant column 15
    3.3.2 Computation of the AMF 16
    3.3.3 Calculation of the tropospheric vertical NO₂ column 17
    3.3.4 Assimilation of OMI measurements in TM5 18
    3.3.5 Error calculation in DOMINO 19
  3.4 The OMI cloud retrieval 21
    3.4.1 Computation of the cloud retrieval look up table 22
  3.5 ODIN HNO₃ 23

4 Retrieval development results 24
  4.1 OMI NO₂ retrieval with TM5 24
  4.2 The impact of higher a priori NO₂ profiles on the retrieved tropospheric NO₂ columns 26
  4.3 Implementing terrain height averaging in DOMINO 29
  4.4 Correcting the OMI cloud retrieval temperature dependence 31
  4.5 Correcting the relative azimuth angle dependence of the AMF 35
  4.6 Updated error calculation for DOMINO 36
  4.7 Combined impact of the retrieval improvements 40

5 Evaluation of the stratospheric TM assimilation 42
  5.1 Using O-F and A-F statistics to diagnose assimilation in DOMINO 42
  5.2 Using Odin HNO₃ forcing in the TM5 stratosphere 50

6 Conclusion 55

7 Outlook 56

A NO₃ chemistry in TM5 57

B Temperature dependence of the reflectance for the O₂-O₂ cloud retrieval 59
1 Introduction

Nitrogen oxides (NO\textsubscript{x} = NO + NO\textsubscript{2}) play an important role in tropospheric chemistry. They catalyze the production of ozone (O\textsubscript{3}) and contribute to aerosol formation. NO\textsubscript{x} is linked to the oxidizing efficiency of the atmosphere since O\textsubscript{3} plays an important role in the formation of OH. OH in turn lowers the lifetime of methane, dampening the greenhouse effect. Although NO\textsubscript{2} is a minor greenhouse gas itself, elevated NO\textsubscript{x} concentrations cause a substantial net cooling of the atmosphere [Shindell et al., 2009]. The cooling effect is mainly related to the removal of methane via OH and the contribution to aerosol formation. In the stratosphere, NO\textsubscript{x} presents an effective O\textsubscript{3} loss which was one of the reasons for the phase-out of the NO\textsubscript{2} emitting Concorde fleet.

NO\textsubscript{x} itself has a negative impact on the ecosystem (e.g. the depletion of oxygen in water) and human health. Recently, the World Health Organization stated that it is reasonable to infer that NO\textsubscript{2} has direct short-term health effects. For long-term exposure, it is hard to distinguish the contribution of NO\textsubscript{2} from other pollutants. However, studies have suggested that long-term NO\textsubscript{2} exposure is associated with respiratory and cardiovascular mortality and children’s respiratory symptoms [World Health Organization, 2013].

Sources of nitrogen oxides are both natural and anthropogenic. Global NO\textsubscript{x} emissions are dominated by anthropogenic combustion processes (about 33 Tg N yr\textsuperscript{-1}, over 60% of the global NO\textsubscript{x} emissions [Schultz et al., 2007]). Major emissions are associated with oceangoing ships, power plants, and automobiles. The contribution of aircraft is under 4% of the anthropogenic emissions. Natural sources of NO\textsubscript{x} include lightning, biomass burning, and soil emissions. The natural sources are responsible for about 20 Tg N yr\textsuperscript{-1}. Over unpolluted areas, almost all NO\textsubscript{2} is located in the stratosphere. Over polluted regions the tropospheric contribution to the total column can reach a value of 90%.

Understanding the effect of NO\textsubscript{x} can be achieved from space by using satellite based spectrometer instruments. Until 1995, the remote sensing of NO\textsubscript{2} was focused on the stratosphere, studying the ozone chemistry. In 1995, ESA’s Global Ozone Monitoring Experiment (GOME) was launched [Burrows et al., 1999]. This instrument measured the light scattered back from the Earth. Using the absorption signature of NO\textsubscript{2} in the measured spectrum, the tropospheric NO\textsubscript{2} column could be retrieved, giving global information about the spatial distribution and intensity of pollution and emissions. The same principle is used by the Ozone Monitoring Instrument (OMI) [Levelt et al., 2006] which was launched aboard EOS-Aura in 2004 and provides daily global coverage. Over the last decades, tropospheric NO\textsubscript{2} has also been measured by SCIAMACHY [Bovensmann et al., 1999] and the two GOME-2 instruments [Munro et al., 2006]. In 2014, the successor of OMI, the TROPOspheric Monitoring Instrument (TROPOMI) will be launched. Although satellite observations of NO\textsubscript{2} are important for monitoring and studying concentrations of nitrogen oxides, considerable uncertainties on the accuracy and robustness of the retrievals, and their fitness for model evaluation still persist. In this study, we present a number of relevant improvements to the tropospheric NO\textsubscript{2} column retrieval algorithm from OMI, the DOMINO algorithm [Boersma et al., 2007 and Boersma et al., 2011] and diagnose the assimilation of NO\textsubscript{2} in the stratosphere.

The outline of this report is as follows. Chapter 2 gives a brief general introduction of the three cornerstones of the retrieval-assimilation-modeling approach: remote sensing, chemistry transport modeling, and data assimilation. Chapter 3 introduces the retrieval framework. The OMI satellite instrument and TM5 chemistry transport model will be introduced, followed by a detailed discussion of the DOMINO retrieval. Subsequently, the OMI cloud retrieval which provides cloud parameters for DOMINO and the Odin SMR instrument that is used for its HNO\textsubscript{3} climatology are examined. Chapter 4 contains several improvements to the DOMINO retrieval. The retrieval will be linked with the TM5 chemistry transport model.
model, the resolution of TM will be increased, the surface pressure calculation will be improved, the
temperature dependence of the cloud retrieval will be corrected, a bug will be corrected in the computa-
tion of the AMF, and the computation of errors will be revisited. The chapter ends with an overview
of the impact of all improvements. In Chapter 5, the OMI assimilated TM stratosphere will be topic of
discussion. The chapter starts with an analysis of the assimilation of NO$_2$ concentrations in TM with
OMI measurements. The chapter ends with a study of nudging the Odin HNO$_3$:O$_3$ ratios, which are the
main driver of stratospheric NO$_2$ in TM. The conclusions and outlook are given in respectively Chapter
6 and 7.
2 Retrieval theory

This chapter gives a short and general introduction into the three cornerstones of the retrieval-assimilation-modeling approach used to retrieve tropospheric NO$_2$ columns. The first section discusses remote sensing. The second section introduces the concept of chemistry transport models and the final third part of the chapter concerns data assimilation.

2.1 Remote sensing

Gas molecules in the atmosphere absorb radiation of a given wavelength if the energy can be used to transition the molecule in a higher electronic state. The electronic states of the molecule are quantized by electronic, vibrational, and rotational states. Electronic transitions are only caused by UV radiation (the electronic transitions of O$_3$ and O$_2$ are very efficient in the stratosphere) and visible radiation; rotational and vibrational transitions are associated with IR radiation. NO$_2$ has its clearest absorption features in the 405-465 nm visible window, where the interference from other trace gases is minimal.

Using a spectrometer, the radiation flux per wavelength can be measured for different wavelengths, constituting a spectrum. The spectrum of solar radiation measured outside the atmosphere is close to the radiation of a black body at 5800K [Jacob, 1999]. Radiation that is backscattered from the Earth atmosphere and surface has been influenced by the absorption of trace gases. Examining the ratio of the Earth radiance and Solar irradiance (the reflectance spectrum) gives information about the abundance of trace gases in the atmosphere. This principle is used by satellites to retrieve the concentrations of trace gases. Nadir instruments monitor the backscattered radiance from the Earth and retrieve the column of a trace gas above a certain surface. They offer a high horizontal resolution but contain limited vertical information. Limb instruments monitor the edge of the atmosphere, allowing the retrieval of (stratospheric) vertical profiles but with a limited horizontal resolution. OMI, the satellite instrument used in this study, is a nadir UV/Vis backscatter instrument and will be introduced in section 3.1.

Knowledge about the average photon path is required to transform the measured reflectance spectrum into the atmospheric state (the column) of the trace gas. The observed spectrum ($y$) can be expressed as:

$$ y = F(x, b) + \varepsilon + \Delta F $$(1)

Where $F$ is a forward model that describes how the measured spectrum depends on the vertical distribution and quantity of the gas ($x$) and other forward model parameters describing the state of the atmosphere ($b$). Examples of the forward model parameters are the viewing geometry (commonly used angles to define the viewing geometry are shown in Figure 1), meteorological parameters, cloud information, surface properties, and information about other absorbing gases. Finally, $\varepsilon$ gives the measurement error and $\Delta F$ the error in the forward model.

For a weakly absorbing gas like NO$_2$ there is only enough information in the measured spectrum to retrieve the vertical column density $x$. The vertical column density can be retrieved from the measured spectrum by using a retrieval procedure $R$. The retrieval also requires the forward model parameters and the assumed a priori vertical distribution distribution of the trace gas, $x_a$:

$$ x = R(y, b, x_a) $$(2)

For NO$_2$, the radiative transfer can be assumed to be linear, which allows a two step retrieval (Equation 3). In the first step, the abundance of the tracer along the average photon path (the slant column, $S$) is
retrieved from the measured spectrum. The second step converts this slant column into a vertical column via the concept of the air mass factor. The air mass factor \( M \) relates the absorption along the average photon path between sun and satellite to the absorption along a photon path for an overhead Sun and ground-based detector. The air mass factor depends on the forward model parameters and the assumed a priori profile [Boersma, 2005]. This is the principle of the DOMINO retrieval of tropospheric NO\(_2\) columns which will be addressed in Section 3.3.

\[
x = R(y, b, x_a) = \frac{S(y)}{M(x_a, b)}
\]  

(3)

Figure 1: Illustration of often used angles in (nadir) remote sensing. All angles are defined relative to the measured pixel, which is located in the origin. Here \( \theta \) is the viewing zenith angle, \( \theta_0 \) is the solar zenith angle, \( \phi \) is the viewing azimuth angle, \( \phi_0 \) is the solar azimuth angle, and \( \phi_{rel} \) is the relative azimuth angle which is by convention defined as \(|\phi - \phi_0|\).

### 2.2 Chemistry Transport Models

Global chemistry transport models (CTMs) simulate the concentrations of trace gases by giving a mathematical description of processes driving atmospheric composition [Jacob, 2007b]. CTMs do not explicitly solve the basic equations of motion but use prescribed meteorological fields\(^4\). As analyzed meteorological fields can be used for the past, CTMs can closely resemble the actual historic state of the atmosphere. Besides, CTMs can use forecast meteo to predict future concentrations.

Starting point of the CTM is the chemical continuity equation for tracer \( i \) which describes the mass conservation of the tracer:

\[
\frac{\partial c_i(x,t)}{\partial t} = -u(x,t) \cdot \nabla c_i(x,t) + R_i(c(x,t),x,t) + E_i(x,t) - S_i(x,t)
\]  

(4)

Here, \( c_i \) is the concentration of a gas at location \( x \) and time \( t \). The change of concentration in time depends on transport of the tracer (the first term on the right side gives the flux divergence, where \( u \) is

\(^4\)Global circulation models with chemistry do explicitly solve the equations of motion but are highly complex and require large computational resources
the wind vector), net chemical production of the tracer \((R_i)\), emissions \((E_i)\), and sinks \((S_i\), for example deposition and scavenging).

In a CTM, this equation is solved numerically, for discrete times (separated by model time step \(\Delta t\)) and discrete locations, the individual grid cells of the model. Most CTMs use a Cartesian horizontal grid with a resolution between 5°×4° and 1°×1°. The vertical domain is usually defined by pressure levels that variate based on the surface pressure of the grid cell. Boundary conditions are periodic in the horizontal direction. Vertical boundary conditions are usually a flux of zero at the top of the atmosphere, and a calculated emission and deposition flux at the surface.

Discretization of Equation 4 gives Equation 5 which can be solved for \(\bar{c}_i\), the average concentration of a tracer for a discrete time and location (grid cell).

\[
\frac{\partial \bar{c}_i}{\partial t} = -\bar{u} \cdot \nabla \bar{c}_i + \left( \frac{\partial \bar{c}_i}{\partial t} \right)_{\text{diff}} + \left( \frac{\partial \bar{c}_i}{\partial t} \right)_{\text{conv}} + R_i(\bar{c}) + E_i - S_i
\]

Here, \(\bar{u}\) is the average wind speed for the grid cell. The additional transport terms (diffusion and convection) have to be parameterized because they are not resolved by the ”coarse” discrete grid and time step. The last three terms give the net chemical production, emission, and sink for the tracer in the grid cell during the time step of the model. The dependence on discrete position and time is left out for the purpose of readability. The simplest way to solve the discretized continuity equations is the backward Euler method. Most global CTMs use more advanced higher-order implicit schemes. Using the principle of operator splitting, different processes are calculated after each other, based on the assumption that the coupling between different processes can be neglected within the model time step. This simplifies the solving of the equation since it reduces its dimensionality.

For this study, the modeling of nitrogen dioxide \((NO_2)\) is most important. Because of the quick cycle between nitrogen oxide \((NO)\) and NO\(_2\), they are often considered together as nitrogen oxides or NO\(_x\). NO\(_x\) is a subset of NO\(_y\), which also consists of nitrous oxide \((N_2O)\), nitrate \((NO_3^-)\), nitric acid \((HNO_3)\), and peroxyacyl nitrate \((PAN)\). The CTM used here, the Tracer Model 5 (TM5) [Krol et al., 2005 and Huijnen et al., 2010b], will be introduced in Chapter 3.2.

### 2.3 Data assimilation

The central idea of data assimilation is to modify a model based on measurements. In the case of the retrieval of NO\(_2\), a modeled NO\(_2\) slant column is modified to better match an observed slant column. While measurements are only performed at specific times and locations, when combined with a model, they can provide a continuous 3-D field of concentrations representative for the measurements. This way, data assimilation can be used to generate a representative forecast of the state of the atmosphere at the time of the measurement. After the measurement, the a priori information (model forecast) can be corrected towards the observation. After the model information has been transformed to how the satellite would have observed the model (through the observation operator), the model can be adjusted to the observation, leading to the model analysis.

Via assimilation, it is possible to correct model biases in a durable way. For stratospheric NO\(_2\), this can be done by adjusting the concentrations of NO\(_2\) because the lifetime of NO\(_2\) in the stratosphere is long enough for the gas to be stable until the next measurement. For the assimilation of the troposphere, one needs to adjust the emissions in the model based on observations. In this procedure, the model is used as the link between emissions and measured concentrations. This specific form of estimating emissions
using data assimilation is called inverse modeling and is not used in the DOMINO algorithm.

Here, we use the Kalman filter technique which is an extension to the optimal interpolation approach. In the optimal interpolation approach, the weights of the model and measurement in the analysis are based on the error covariance estimates of both [Jeuk, 2000]. The assumption for this method to work is that the difference between forecast and observation has a Gaussian distribution. The Kalman filter technique is an extension to the optimal interpolation assimilation scheme, that also takes into account the time evolution of the forecast error [Kalman and Bucy, 1961]. The scheme used in the DOMINO retrieval is a fast version of the Kalman filter because it parameterizes the correlations between forecast errors [Eskes et al., 2003]. The assimilation performed in DOMINO will be the subject of section 3.3.4. A comprehensive theoretical description of the applied Kalman-filtering technique is beyond the scope of this thesis but can be found in Eskes et al. (2003).
3 Retrieval framework

This chapter discusses the existing framework used to retrieve tropospheric NO\textsubscript{2} columns from the Dutch-Finnish Ozone Monitoring Instrument (OMI). The chapter starts with a discussion of the OMI instrument used to determine the reflectance spectrum, the ratio of Earth radiance and solar irradiance. The reflectance spectrum is the starting point of the retrieval. From this reflectance spectrum, a slant column of NO\textsubscript{2} can be determined. However, this slant column has a contribution from the troposphere and stratosphere and gives the abundance of NO\textsubscript{2} along the average slant path of photons reaching the instrument. For practical applications, the vertical tropospheric NO\textsubscript{2} column is desired. For the conversion of the total slant column to the tropospheric vertical column, a chemistry transport model is required. The TM5 chemistry transport model that is used for the estimation of the stratospheric contribution to the slant column and the a priori vertical distribution of NO\textsubscript{2}, is discussed in the second section. Furthermore, the Dutch OMI NO\textsubscript{2} (DOMINO) retrieval is examined and the chapter finishes with a description of the OMI cloud retrieval which provides the cloud parameters used in DOMINO and the Odin instrument, which is used for its HNO\textsubscript{3} climatology.

3.1 The Ozone Monitoring Instrument

The Ozone Monitoring Instrument (OMI) [Levelt et al., 2006] is a visible/ultraviolet solar backscatter spectrometer equipped to the nadir side of the EOS-Aura (Latin for breeze) satellite. Aura was launched in July 2004 and is part of the National Aeronautics and Space Administration’s (NASA) Earth Observing System (EOS) mission. The satellite has a Sun-synchronous ascending polar orbit at 705 km altitude and 98.2° inclination. The spacecraft completes 14 to 15 orbits per day with a local overpass time of 13:45 h, repeating its ground track every 16 days.

![Figure 2: Measuring principle of OMI showing the swath of 13 km $\times$ 2600 km consisting of 60 pixels with different widths being processed by a 2-D CCD detector. The figure is based on Levelt et al. (2006).](image)

OMI is a collaboration of the Dutch Agency for Aerospace Programs (NSO) and Finnish Meteorological Institute (FMI). The instrument measures direct sunlight and the sunlight backscattered by the atmosphere and surface between 270 and 500 nm. By measuring both UV and visible radiation, OMI is able to measure concentrations of important trace gases such as O\textsubscript{3}, SO\textsubscript{2}, HCHO, BrO, and NO\textsubscript{2}; aerosols characteristics; and effective cloud pressures and fractions. OMI provides daily global coverage with a spatial resolution of up to 13 $\times$ 24 km$^2$ in the regular 'Global observation mode'. This coverage is achieved by using swath registration perpendicular to the flight direction of the satellite with a field of
Light entering OMI is pseudo-depolarized using a scrambler to make the measurement polarization independent. Subsequently, the light is split into three beams, each proceeding to one of the three channels: UV-1 (270-310 nm), UV-2 (310-365 nm), or VIS (365-500 nm, containing strong NO$_2$ spectral lines between 405 nm and 465 nm). Measurements of the Earth radiance (in W/m$^2$/nm/sr) and solar irradiance (in W/m$^2$/nm) are taken by three 2-D charge-coupled device (CCD) detectors, one for each channel. One dimension of the CCD is used to obtain spectral information (wavelength dimension), the other for spatial data (swath dimension). Five measurements of the Earth radiance are taken per pixel, each with an exposure time of 0.4s leading to nominal measurement time of 2s. The CCD measurements result in an Earth radiance spectrum of light scattered and absorbed by the Earth atmosphere and surface. Once per day, OMI directly measures the Sun, the solar irradiance, serving as a reference for the Earth radiance measurements. Spectral calibration is achieved by a cross-correlation of Fraunhofer lines in theoretical and observed in-flight irradiance spectra. A detailed description of the calibration of OMI can be found in Dobber et al. (2006). The signal-to-noise ratio of the OMI reflectance lies between 1000 and 1500 for the visible domain.

The reflectance spectrum ($R$) is calculated from the ratio between the measured Earth radiance ($I$) and the solar irradiance ($E$) as a function of wavelength ($\lambda$) and is calculated using Equation 6. Here, $\mu_0$ is the cosine of the solar zenith angle. The reflectance spectrum is the starting point of the DOMINO retrieval of tropospheric NO$_2$ columns discussed in Section 3.3.

$$R(\lambda) = \frac{\pi I(\lambda)}{\mu_0 E(\lambda)}$$

Examples of the solar irradiance spectrum, Earth radiance spectrum, and resulting reflectance spectrum measured by OMI are shown in Figure 3. Scientific data from OMI is available from October 2004 up to today (July 2013). The optical degradation of the VIS channel used of NO$_2$ was under 2% between 2004 and 2010. However, OMI does suffer from across-track variability (or stripes) related to calibration differences in the solar irradiance spectrum and radiance measured for different viewing angles [Kleipool et al., 2008]. The bias translates into an apparently viewing angle dependent error in the order of 10% of the slant column. In DOMINO v2, a stripe-correction is in place. Furthermore, since 2007, row anomalies have occurred for several viewing angles, leading to substantial signal suppression for that viewing angle for either part of or an entire orbit. More than half of the rows are (partially) affected. The cause of the anomaly is currently unknown and measurements influenced by the anomaly should not be used [Boersma et al., 2011].
3.2 Tracer Model 5

The global atmospheric chemistry transport model (CTM) that is used for the a priori profiles and assimilation of the stratosphere is Tracer Model 5 (TM5) [Krol et al., 2005 and Huijnen et al., 2010b]. TM5 is the successor of the three-dimensional global atmospheric TM4 CTM [Dentener et al., 2003 and Williams et al., 2009] that was used for DOMINO v2 retrievals. Here we use the TM5-chem-v3.0 benchmark version. The model simulates the concentrations of 42 tracers and uses a regular global grid with a resolution of up to $1^\circ \times 1^\circ$. In the vertical dimension, 34 hybrid pressure levels are used, with the highest layer at 0.1 hPa. Compared to TM4, the TM5 model has more layers in the troposphere and fewer layers in the stratosphere to get a better representation of the free troposphere. The vertical levels of TM4 and TM5 are shown in Figure 4. The model is run with meteorology from the ERA-Interim reanalysis, a list of all meteorological variables used in TM5 can be found in Huijnen et al. (2010b).

Within a time step (the time step will be discussed in detail in section 3.2.1), TM5 uses operator splitting. Each operation is performed twice in a symmetric order. The time step starts with horizontal advection (X,Y), followed by vertical advection (Z) and vertical mixing (V). After the transport, sources and sinks (S) are calculated and at end the first half of the time step, the chemistry (C) is applied. In the second half of the time step, the processes are performed in opposite order. A schematic overview of a TM5 time step is given in Eq. 7.

\[(XYZVSC)(CSVZYX)\]  

\[ (7) \]

In the subsequent subsections, the different operations will be discussed. In section 3.2.4 special attention will be given to stratospheric NO\textsubscript{x} chemistry in TM5. Finally, in section 3.2.5, the performance of the model will be examined.
3.2.1 Horizontal advection and vertical transport

TM5 uses the mass-conserving advection scheme as introduced in Berkvens et al. (1999) which is based on Russel and Lerner (1981). The horizontal advection determines the required time step of the model. The time step of the model is reduced till the Courant-Friedrichs-Lewy (CFL) stability criterion is fulfilled (Eq. 8, [Bregman et al., 2003]), which means that the time step ($\Delta t$) is smaller than the dimension of a grid cell ($\delta x$) divided by the wind speed ($v$) within the grid cell.

$$\Delta t < \frac{\delta x}{v}$$

This criterion gives a time step of on average 30 minutes for $3^\circ \times 2^\circ$ and 15 minutes for $1^\circ \times 1^\circ$ simulations. Only 27 of the 42 tracers are transported, 15 short-lived chemical species are non-transported. NO$_2$, the relevant gas for this research, is such a short-lived tracer and is not transported individually. However, together with NO, NO$_3$, N$_2$O$_5$, and HNO$_4$ this gas is transported as lumped aggregate in the (transported) NO$_x$ tracer$^\dagger$. Vertical transport also considers convection [Tiedtke, 1989], diffusion based on Louis (1979) for the free troposphere and Holtslag and Boville (1993) within the planet boundary layer, and scavenging due to precipitation. The stratosphere-troposphere exchange is based on the overturning circulation from the ERA-Interim meteorology.

3.2.2 Sources and sinks

Emissions are available with a monthly time resolution. Anthropogenic emissions are based on the 2000 emissions from the RETRO project [Schultz et al., 2007]. For east Asia, emissions are replaced by the REAS inventory for 2006 by Ohara et al. (2007). This provides an important difference with TM4, in which emissions were taken from the 1997 POET inventory [Granier et al., 2005]. Emissions over Europe and the US are lower in the RETRO inventory compared to the POET inventory. To simulate the weekend effect (NO$_x$ emissions are lower over the weekend), emissions over Europe and the US were scaled by

$^\dagger$NO$_x$ is defined as NO + NO$_2$ but the tracer here is used to represent NO$_3$, N$_2$O$_5$, and HNO$_4$ as well.
1.11 during the week, 0.85 on Saturday, and 0.6 on Sunday in TM4. Such a first order weekend-effect is not implemented in TM5.

Biogenic emissions are given by a 12-year average of the ORCHIDEE model [Lathièra et al., 2006]. Shipping emissions are from the AMVER inventory [Endresen et al., 2003], scaled up with a factor 1.95 to give more realistic emission totals. Lightning NO\(_x\) is calculated by using the linear relation between observed lightning flashes and convective precipitation [Meijer et al., 2001]. Biomass burning emissions are taken from the 8-day Global Fire Emissions Database v2 (GFEDv2) emission inventory [van der Werf et al., 2006]. Finally, aircraft emissions are based on Schmitt and Brunner (1997). An overview of NO\(_x\) (in N) emission totals is given in Table 1, all sources are emitted as NO in the model. NO emissions are applied within the chemistry step because of the quick cycling to NO\(_2\).

<table>
<thead>
<tr>
<th>Emissions</th>
<th>Annual emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anthropogenic</td>
<td>32.9 Tg N yr(^{-1})</td>
</tr>
<tr>
<td>Of which: Shipping</td>
<td>6.3 Tg N yr(^{-1})</td>
</tr>
<tr>
<td>Of which: Aircraft</td>
<td>0.7 Tg N yr(^{-1})</td>
</tr>
<tr>
<td>Soil</td>
<td>9.3 Tg N yr(^{-1})</td>
</tr>
<tr>
<td>Lightning</td>
<td>5.9 Tg N yr(^{-1})</td>
</tr>
<tr>
<td>Biomass burning</td>
<td>5.1 Tg N yr(^{-1})</td>
</tr>
<tr>
<td>Total</td>
<td>53.2 Tg N yr(^{-1})</td>
</tr>
</tbody>
</table>

Surface emissions are injected in the lowest model layer, except when the thickness of this layer is less than 3.5 hPa. If this is the case, emissions are divided over the first two layers. This operator step also includes the application of stratospheric boundary conditions which will be discusses in section 3.2.4. As sinks, wet deposition (in-cloud and below-cloud scavenging) and dry deposition following the resistance approach by Wesely (1989) are included.

3.2.3 Chemistry

The chemistry step consists of the regular chemistry scheme, the photolysis, dry deposition, and NO\(_x\) emissions. The chemical mechanism is the Carbon Bon Mechanism 4 (CBM4), which was also used in TM4 and described in Houweling et al. (1998). Reactions rates are based on Sander et al. (2006), Atkinson et al. (2006), and Yarwood et al. (2005). TM5 uses the Euler Backward Iterative (EBI) solver as chemical solver [Hertel et al., 1993]. Photo dissociation rates are calculated using a look up table of temperature dependent absorption coefficients and quantum yields, and an offline parameterization for actinic fluxes based on Landgraf and Crutzen (1998). Furthermore, the model includes heterogeneous phase reactions and aerosol chemistry. To prevent the artificial loss of nitrogen, a mass balance step is applied to the NO\(_x\) family after chemistry. A complete overview of the TM5 NO\(_x\) CBM4 and photolysis reactions considered in TM5 is given in Appendix A. The TM5 chemistry is similar to the TM4 chemistry used for DOMINO v2.0.

3.2.4 Stratospheric NO\(_2\) in TM5

Modeled stratospheric NO\(_x\) is important for the retrieval of tropospheric NO\(_2\) because the OMI assimilated stratospheric NO\(_2\) column is subtracted from the measured NO\(_2\) slant column, resulting in the tropospheric slant column. The natural source of NO\(_x\) in the stratosphere is nitrous oxide (N\(_2\)O). N\(_2\)O is
is a product of nitrification and denitrification in the biosphere and since it is inert in the troposphere, it will eventually reach the stratosphere. In the stratosphere, 95% of N₂O is converted to N₂ by photolysis and oxidation by O. However, 5% is oxidized by the oxygen atom to form NO (R0). The other (smaller) source of NO₃ in the stratosphere is the emission of NOₓ by stratospheric aircraft. This source is negligible after the phase-out of the Concorde fleet.

Table 2 shows an overview of the important stratospheric reactions involving NOₓ. Stratospheric NO₂ is photolyzed to NO (R1, in Table 2), which reacts with O₃ to produce NO₂ (R2). R1-R2 constitutes a null cycle between NO and NO₂ which has a time scale of about one minute. Because the oxygen atom from R2 and oxygen from R1 will react to reproduce O₃, this cycle has no net effect on ozone. A fraction of NO₂ produced by R1 reacts with oxygen atoms to produce NO (R3). Since this reaction prevents O₃ from reproducing ozone, it does give an effective O₃ loss in combination with R1.

Table 2: Important stratospheric chemistry reactions regarding the NO₂ concentration, a full overview of the TM5 chemistry is given in Appendix A.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Rate</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>N₂O + O → 2NO</td>
<td>R0</td>
<td></td>
</tr>
<tr>
<td>NO + O₃ → NO₂ + O₂</td>
<td>R1</td>
<td></td>
</tr>
<tr>
<td>NO₂ + hv → NO + O</td>
<td>R2</td>
<td></td>
</tr>
<tr>
<td>NO₂ + O → NO + O₂</td>
<td>R3</td>
<td></td>
</tr>
<tr>
<td>NO₂ + OH → HNO₃</td>
<td>R4</td>
<td></td>
</tr>
<tr>
<td>NO₂ + O₃ → NO₃ + O₂</td>
<td>R5</td>
<td></td>
</tr>
<tr>
<td>NO₃ + hv → NO₂ + O</td>
<td>R6</td>
<td></td>
</tr>
<tr>
<td>NO₃ + NO₂ → N₂O₅</td>
<td>R7</td>
<td></td>
</tr>
<tr>
<td>HNO₃ + hv → NO₂ + OH</td>
<td>R8</td>
<td></td>
</tr>
<tr>
<td>HNO₃ + OH → NO₃ + H₂O</td>
<td>R9</td>
<td></td>
</tr>
<tr>
<td>N₂O₅ + hv → NO₃ + NO₂</td>
<td>R10</td>
<td></td>
</tr>
<tr>
<td>N₂O₅ + H₂O → 2HNO₃</td>
<td>R11</td>
<td></td>
</tr>
</tbody>
</table>

Furthermore, NO₂ can react with the OH radical to produce HNO₃ (R4, only occurs during the day) and be oxidized to NO₃ by O₃ (R5). During the day, the resulting NO₃ radical is photolyzed back to NO₂ (R6) but at night, the resulting NO₃ radical is subsequently converted to N₂O₅ (R7). Both HNO₃ and N₂O₅ are eventually converted back to NOₓ but have relative long lifetimes (R8-R10). NO₃ will be converted back to NO₂ by R6. The main sink of stratospheric NOₓ is transport to the troposphere where HNO₃ is removed by deposition. A schematic overview of stratospheric NOₓ chemistry is shown in Figure 5 [Jacob, 1999].

However, TM5 does not simulate N₂O concentrations and the chemistry scheme does not include halogenated species or photolytic destruction below 202 nm which are relevant for the NO₃ chemistry [Huijnen et al., 2010b]. Without some external forcing, this makes it difficult for TM5 to autonomously simulate stratospheric NOₓ and O₃ concentrations. To solve this, ozone and nitric acid are constrained in the stratosphere using climatologies. O₃ is nudged to the monthly specific and yearly varying assimilated multi-sensor reanalysis (MSR) dataset which is based on measurements of a combination of space-borne ozone-measuring instruments [van der A et al., 2010]. TM5 volume mixing ratios of O₃ are nudged towards the MSR climatology above the tropopause with a nudging time of 2.4-4 days depending on the latitude. HNO₃ concentrations are prescribed at 10hPa by multiplying the MSR O₃ concentration with a HNO₃:O₃ ratio based on measurements by the Microwave Limb Sounder (MLS) instrument aboard the Upper Atmosphere Research Satellite (UARS) satellite (Eq. 9). The UARS climatology is monthly
specific. The nudging makes the HNO\textsubscript{3} climatology the effective source of NO\textsubscript{y} and therefore NO\textsubscript{2} in the stratosphere. This implies that the nudged HNO\textsubscript{3} profile has a strong influence on the vertical profile of NO\textsubscript{2} in the stratosphere. As will be discussed in section 3.3, stratospheric NO\textsubscript{2} will also be adjusted (assimilated) using OMI measurements.

\[
[HNO_3]_{\text{Clim}} = \left( \frac{[HNO_3]}{[O_3]} \right)_{\text{UARS}} \times [O_3]_{\text{MSR}} \tag{9}
\]

In DOMINO v2, the same UARS ratio was used with a nudging time of 2 months (\(\delta t_{\text{nudge}}\)) [Dirksen et al., 2011]. To exclude this difference from the initial comparison between DOMINO v2 and the prototype DOMINO v3, we will initially use the same nudging time. The HNO\textsubscript{3} concentration ([HNO\textsubscript{3}]\textsubscript{Clim}, Eq. 9) that is calculated using the climatologies is used to nudge the TM concentration from the previous time step ([HNO\textsubscript{3}]\textsubscript{TM,\(\cdot\delta t\)}) to the new model concentration ([HNO\textsubscript{3}]\textsubscript{TM}, Eq. 10). This way, the TM HNO\textsubscript{3} concentration is not directly prescribed but the result of a combination of the current model state and the climatologies. The resulting stratospheric TM NO\textsubscript{2} concentration is driven by the HNO\textsubscript{3} and the TM NO\textsubscript{2} chemistry that provides the link between HNO\textsubscript{3} and NO\textsubscript{2}. Another difference between TM5 and TM4 is that in TM4, the stratospheric ozone concentrations were nudged to the Fortuin-Kelder climatology [Fortuin and Kelder, 1998]. This difference will not be compensated.

\[
[HNO_3]_{TM} = \frac{[HNO_3]_{TM,\cdot\delta t} + [HNO_3]_{\text{Clim}} \times \frac{\delta t}{\delta t_{\text{nudge}}}}{1 + \frac{\delta t}{\delta t_{\text{nudge}}}} \tag{10}
\]
3.2.5 Model performance

Tropospheric NO$_2$ columns simulated by TM5 have been shown to have a global spatial correlation coefficient of 0.89 with DOMINO v1.0.2 observations and show good seasonal correlation. Absolute tropospheric columns of TM5 are $0.5-1.0 \times 10^{15}$ molec./cm$^2$ lower than DOMINO v1.0.2 values [Huijnen et al., 2010b]. This difference can be attributed to the high value DOMINO v1.0.2 columns [Boersma et al., 2008] and a possible underestimation of emissions in TM5. Furthermore, the NO$_2$ profiles of TM5 show good agreement with aircraft measurements from the INTEX-B campaign over Mexico in March 2006 [Singh et al., 2009] with differences up to 0.1 ppbv [Huijnen et al., 2010b]. Finally it should be noted that there is ongoing discussion about the model lifetime of NO$_2$ [Stavrakou et al., 2013]. Especially the uptake coefficient on aerosol of N$_2$O$_5$ and production rate of HNO$_3$ from NO and NO$_2$ have large uncertainties.
3.3 The Dutch OMI Tropospheric NO\textsubscript{2} retrieval

The Dutch OMI Tropospheric NO\textsubscript{2} retrieval (DOMINO) retrieves tropospheric vertical NO\textsubscript{2} columns from the measurements by the OMI instrument. The DOMINO v2 retrieval is widely used in the scientific community, for instance recently by McLinden et al. (2012) to study trends in oil sands emissions over Canada. The DOMINO product has been validated by Irie et al. (2012). Tropospheric NO\textsubscript{2} DOMINO v2.0 columns are publicly available within three hours via the Tropospheric Emission Monitoring Internet Service (TEMIS), \url{http://www.temis.nl/} in the orbital HDF-EOS5 format [Boersma et al., 2009]. The retrieval is based on the retrieval-assimilation-modeling (RAM) approach which has previously been used for GOME, SCIA, and GOME-2. DOMINO uses the TM model to calculate the stratospheric contribution to the slant column and predict the vertical NO\textsubscript{2} profile that is needed to calculate the tropospheric air mass factor (AMF). The principle of the DOMINO retrieval is schematically presented in Figure 6.

![Figure 6: A conceptual view of the 3-step DOMINO retrieval. In the first step, a DOAS fit of the reflectance spectrum gives the NO\textsubscript{2} slant column, the abundance of NO\textsubscript{2} along the average photon path through the atmosphere to the instrument. In the second step, the TM stratospheric NO\textsubscript{2} concentrations are used to estimate and subtract the stratospheric contribution to the slant column. In the final step, the tropospheric slant column is converted to the tropospheric vertical column by through dividing by the tropospheric AMF.]

3.3.1 The NO\textsubscript{2} slant column

The starting point of the DOMINO retrieval is the reflectance spectrum as measured by OMI (\(R(\lambda)\), Section 3.1). The NO\textsubscript{2} slant column (the abundance of NO\textsubscript{2} along the average photon path) is retrieved using the Differential Optical Absorption Spectroscopy (DOAS) method [Platt and Stutz, 2008] in the 405-465 nm window. This fitting window is wider than the fitting window used by similar UV-Vis backscatter instruments to account for the relatively low signal-to-noise ratio of OMI. The DOAS method consists of minimizing the chi-squared difference between a modeled spectrum and the observed Earth reflectance spectrum. The modeled spectrum is optimized by fitting the correct concentrations (slant columns) of absorbing tracers.

For the modeled spectrum, the absorption cross section (\(\sigma\)) of NO\textsubscript{2} is taken from the 220 K spectrum from Vandaele et al. (1998). The temperature dependence of the cross section will be corrected in the
AMF calculation (Section 3.3.2). Furthermore, cross sections for O$_3$ and H$_2$O are fitted. Now, the simulated reflectance spectrum ($R_{\text{model}}$) is related to the slant columns ($S_i$, the abundance of the species along the photon path) and cross sections ($\sigma_i$) of the tracers, as described in Equation 11 where the summation is over the tracers. The background reflectance is modeled as a fifth-order polynomial ($P_5$, given by Eq. 12 where $a_m$ are the polynomial coefficients). The polynomial accounts for spectrally smooth structures resulting from molecular multiple scattering and absorption, aerosol scattering and absorption, and the surface reflectance [Boersma et al., 2002]. The last term of Equation 11 represents the inelastic Raman scattering by N$_2$ and O$_2$ molecules. The so-called Ring effect is modeled by $C_{\text{ring}}$, the Ring fitting coefficient and $I_{\text{ring}}(\lambda)/E(\lambda)$, the Sun-normalized synthetic Ring spectrum. Modeling the ring effect similar to the absorption by tracers would assume the effect to occur along the entire light path from Sun to satellite. However, the ring effect does not occur along the light path from the Sun until the light is scattered by a molecule for the first time. Therefore, the effect is modeled as 1 + the Ring fitting coefficient to take into account that it only occurs for the second half of the light path.

$$R_{\text{model}}(\lambda) = \exp\left(\sum_i -\sigma_i S_i\right) \times P_5(\lambda) \times \left(1 + C_{\text{ring}} \frac{I_{\text{ring}}(\lambda)}{E(\lambda)}\right)$$ (11)

$$P_5(\lambda) = \sum_{m=0}^5 a_m \lambda^m$$ (12)

The slant column densities and polynomial coefficients are now obtained through a linear least-square fitting (minimizing $\chi^2$) of the modeled ($R_{\text{model}}$) to the measured ($R_{\text{meas}}$) reflectance as described in Equation 13. Here, the difference for each wavelength is divided by the precision of the measurement $\Delta R_{\text{meas}}$, which depends on the precision of the radiance and irradiance measurements.

$$\chi^2 = \sum_{\lambda_i=405}^{465} \frac{R_{\text{meas}}(\lambda_i) - R_{\text{model}}(\lambda_i)}{\Delta R_{\text{meas}}(\lambda_i)}$$ (13)

By minimizing $\chi^2$, the slant column densities ($S_i$) for different tracers are found that explain the measured reflectance spectrum. However, it is the vertical column of NO$_2$ above a certain surface area that one is interested in. Therefore, the concept of the air mass factor is introduced in the next section.

### 3.3.2 Computation of the AMF

The retrieved slant column is not a geophysical parameter, it mainly depends on the viewing and solar angles of the measurement. The air mass factor (AMF) gives the relation between the slant column density along the optical path from Sun to satellite (measured by the satellite) and the vertical column density above the surface the satellite is viewing. The altitude-dependent air mass factor gives this relation for a partial column, related to one vertical layer in the TM model. In DOMINO, the altitude-dependent AMFs are calculated offline with the Doubling Adding KNMI (DAK) radiative transfer model [de Haan et al., 1987 and Stammes et al. 2001]. The AMFs are calculated for 439 nm (the center of the retrieval window) in DAK by simulating 24 atmospheric layers and representing clouds and the surface as opaque Lambertian surfaces. DAK takes into account polarization and Rayleigh scattering. Results of the radiative transfer modeling are stored as a function of forward model parameters in the 6-D AMF look up table (LUT). The forward model parameters are: solar zenith angle ($\theta_0$), viewing zenith angle ($\theta$), relative azimuth angle ($\phi_{rel}$), surface albedo ($A_s$, based on Kleipool et al. (2008)), surface pressure ($p_s$), and the atmospheric pressure of the TM layer ($p_l$).
In the DOMINO retrieval, values for the AMF are computed by a 6-D linear interpolation of the LUT. Subsequently, we use the independent pixel approximation, assuming that a pixel can be divided in sub-pixels that are either completely clouded or clear. The reflectance for the pixel is the sum of the reflectance from the sub-pixels [Chambers et al., 1997]. The AMF for the cloudy sub-pixels is calculated using the same LUT but with the cloud pressure \(c_p\), taken from the OMI cloud product (Section 3.4) instead of the \(p_s\) and cloud albedo \(A_c\), assumed to be 0.8, Section 3.4) instead of \(A_s\). The altitude-dependent AMF \(m_l\), Eq. 14 can now be expressed as a linear combination of the clear-sky AMF \(m_{clr}\) and cloudy AMF \(m_{cld}\). The weighing is based on the radiance weighted cloud fraction \(w\), which is given by Eq. 15. In Eq. 15, \(R_{clr}\) is the radiance from the clear part of the pixel, \(R_{cld}\) the radiance from the cloudy part of the pixel, and \(c\) the effective cloud fraction from the OMI cloud retrieval.

\[
m_l = w m_{cld} + (1 - w) m_{clr} \\
w = \frac{c R_{cld}}{c R_{cld} + (1 - c) R_{clr}}
\]

3.3.3 Calculation of the tropospheric vertical NO\(_2\) column

The slant column measured by OMI has a significant stratospheric contribution. The subtraction of the stratospheric contribution of the slant column, to get the tropospheric component, is based on the stratosphere simulated by TM. Stratospheric NO\(_2\) in TM is assimilated with OMI measurements as will be described in Section 3.3.4. Now that the altitude-dependent AMFs \(m_l\) are known, the TM partial columns (NO\(_2\) column per vertical layer, the a priori profile, set of \(n_l\)) can be converted to slant columns as they would have been observed by OMI. Furthermore, a temperature correction \(c_l\) is applied to take into account the difference between the cross-section temperature \(T_\sigma = 220K\) [Vandaele et al., 1998]) and the actual temperature of NO\(_2\) in the atmosphere which depends on the NO\(_2\) profile, temperature profile and AMFs. The temperature profile is taken from TM5 \(T_{TM5}\) and is based on the ERA-INTERIM reanalysis. The temperature correction is given in Equation 16 where \(T_x\) (11.39 K) is the result of a least-square fit of the fitted slant column versus the cross section temperature [Boersma et al., 2004].

\[
c_l = \frac{T_\sigma - T_x}{T_{TM5} - T_x}
\]

Now, the tropospheric slant column \(S_{tr}\) can be calculated by subtracting all stratospheric TM5 vertical levels (down to the WMO tropopause) multiplied by the temperature correction and altitude-dependent AMF (the forecast stratospheric slant column) from the measured slant column \(S\), Eq. 17. Since the stratosphere in TM5 is mainly determined by the assimilation of OMI observations and the stratospheric AMF calculation is straightforward, the stratospheric slant column has a relatively small error.

\[
S_{tr} = S - \sum_{l=tropopause}^{lmax} m_l n_l c_l
\]

In the final step of the retrieval, the tropospheric slant column is converted to the tropospheric vertical column \(V_{tr}\). This is done by dividing by the tropospheric AMF \(M_{tr}\), which depends on the a priori profile \((x_{a, tr})\) and forward model parameters \((b)\). To calculate the AMF, the slant column forecast from TM is calculated by multiplying the altitude-dependent AMFs with the partial TM NO\(_2\) columns below the tropopause. Subsequently, this slant column forecast is divided by the model vertical column forecast which is the sum of all the partial NO\(_2\) columns in TM below the tropopause:

\[
M_{tr}(x_{a, tr}, b) = \frac{\sum_{l=tropopause}^{lmax} m_l n_l c_l}{\sum_{l=tropopause}^{lmax} n_l}
\]
The computation of the tropospheric NO\textsubscript{2} vertical column is shown in Equation 19, where the tropospheric slant column (Eq. 17) is divided by the tropospheric AMF.

\[ V_{tr} = \frac{S_{tr}}{M_{tr}(x_{a, tr}, b)} \] (19)

The geometric AMF is a simplification of the AMF which does not depend on the a priori profile. A basic representation of the geometric AMF (\(M_{geo}\)) only depends on the cosines of the viewing zenith angle \(\theta\) and solar zenith angle \(\theta_0\):

\[ M_{geo} = \frac{1}{\cos(\theta_0)} + \frac{1}{\cos(\theta)} \] (20)

A more precise formulation of the geometric AMF is used in DOMINO. This \(M_{geo}\) takes into account the curvature of the Earth as given in Equation 21 [Leue, 1999]. Here, \(\delta\) is defined as the ratio between the height of the atmosphere and the radius of the Earth. A similar correction for the \(\theta\) term can be neglected.

\[ M_{geo} = \frac{\sqrt{\cos^2(\theta_0) + \delta^2 + 2\delta - \cos(\theta_0)}}{\delta} + \frac{1}{\cos(\theta)} \] (21)

It is known that the geometrical tropospheric NO\textsubscript{2} column has a too large value because it fails to represent the reduced sensitivity to NO\textsubscript{2} in the boundary layer. Thus it gives too small tropospheric NO\textsubscript{2} columns over polluted regions. For the stratospheric and total AMF, the geometric approximation is representative when the tropospheric contribution to the slant column is relatively small.

### 3.3.4 Assimilation of OMI measurements in TM5

After a tropospheric column is retrieved, the TM5 model is assimilated with the OMI measurement to improve the agreement between the stratospheric slant column measured by OMI and the forecasted slant column by the model. The forecasted slant column is used to estimate the stratospheric contribution of the OMI slant column (Eq. 17). The assimilation is based on the Kalman filter technique with a prescribed horizontal correlation. It updates the forecasted model field (\(x_f\)) to the analyzed model field (\(x_a\)) according to Equation 22.

\[ x_a = x_f + PH^T(HPH^T + R)^{-1}(y - y_m) \] (22)

Here, \(P\) is the forecast error covariance matrix which accounts for model imperfections; \(H\) is the observation operator (the average altitude-dependent AMF for the super observation); and \(R\) is the combined observation and representativeness error covariance which effectively filters out OMI observations with a high tropospheric contribution [Eskes and Boersma, 2003]. The super observation (\(y\)) is the average of all OMI observations that are located in the considered TM grid cell and \(y_m\) gives the model forecast for that grid cell. The model forecast is equal to \(Hx_f\), best explained as how OMI would have observed the modeled profile \(x_f\).

\(PH^T(HPH^T + R)^{-1}\) determines the best model adjustment based on the difference between observation and forecast (\(y - y_m\)). The size of the adjustment is based on the ratio between the uncertainty in the model (\(P\)) and observation (\(R\)). For observations of polluted regions, with a high tropospheric contribution to the slant column, the error in the measurement will be bigger because of the strong dependence of
the tropospheric contribution on the assumed state of the atmosphere (i.e. a priori profile, cloud parameters, albedo). The error in the observation ($σ_0$) used in matrix $R$ is based on the modeled profile shape and given in Equation 23.

$$σ_0 = \frac{AS_{\text{trop}} + BS_{\text{strat}}}{S}$$  \hspace{1cm} (23)

Here $S$ is the slant column, $S_{\text{trop}}$ is the tropospheric contribution, and $S_{\text{strat}}$ is the stratospheric contribution. Constants $A$ and $B$ have been chosen as 4.0 and 0.25 respectively. This ensures that the assimilation is most effective for slant columns with a small tropospheric contribution and mainly influences the concentration of NO$_2$ in the stratosphere. In covariance matrix $P$, a second-order autoregressive Thiebaux function (Eq. 24) with a correlation length of 600 km ($r_{\text{cor}}$) is used to describe the correlation ($ρ$) between errors of super observations at distance $r$. The correlation function filters out structures smaller than 600 km in the difference between observation and forecast, reducing the influence of small-scale pollution on the analysis. To prevent physical strong gradients in stratospheric NO$_2$ (like the Noxon cliff [Noxon, 1979]) to be filtered out, the correlation is set to zero for large stratospheric concentration differences.

$$σ(r) = \left(1.0 + \frac{2.0r}{r_{\text{cor}}} \right) \exp \left(-\frac{2.0r}{r_{\text{cor}}} \right)$$  \hspace{1cm} (24)

Because of the fact that NO$_y$ is a well-conserved quantity, not only NO$_2$ but also the fully correlated tracers NO, NO$_3$, N$_2$O$_5$, HNO$_4$, and lump tracer NO$_x$ are updated with the same ratio as NO$_2$. The assimilation can be diagnosed by examination of the difference between the slant column retrieved from OMI ($y$) and the slant column forecasted by the model ($y_{\text{m}}$), referred to as O-F. The difference between the TM analysis slant column after assimilation and the forecast (A-F), can serve as a tool to analyze the effectiveness of the model forcing. Both diagnostics will be used in Chapter 5.1 to diagnose the assimilation in the prototype DOMINO v3. A more detailed discussion of assimilation of the stratosphere in DOMINO can be found in Dirksen et al. (2011).

In the regular DOMINO retrieval, the assimilation is performed immediately after the processing of a slant column (at the same model time). In the near real-time (NRT) retrieval [Boersma et al., 2007] the assimilation is only done once a day. This means that the forecasted columns are based on a model run that has not been assimilated with OMI for a maximum time of 24 h. Furthermore, in the NRT retrieval, the forecast meteo from ECMWF is used instead of the analyzed fields. These modifications allow the publication of retrieved columns within 3 hours after the start of an orbit.

### 3.3.5 Error calculation in DOMINO

The DOMINO retrieval calculates errors for the total, tropospheric, and stratospheric NO$_2$ column. Furthermore, errors are calculated for the total and tropospheric averaging kernel. Here, we will focus on the computation of the error in the main product, the tropospheric vertical NO$_2$ column. The error in the tropospheric vertical column ($σ_{V_{\text{t}}}$, Eq. 25), consists of contributions from the error in the slant column ($σ_{S}$), the subtraction of the stratospheric contribution ($σ_{S_{\text{strat}}}$), and the tropospheric AMF ($σ_{M_{\text{tr}}}$). By using the relations of Eq. 19 and 17, the error is given by Equation 26.

$$σ_{V_{\text{t}}} = \sqrt{\left(\frac{∂V_{\text{t}}}{∂S}σ_S\right)^2 + \left(\frac{∂V_{\text{t}}}{∂S_{\text{strat}}}σ_{S_{\text{strat}}}\right)^2 + \left(\frac{∂V_{\text{t}}}{∂M_{\text{tr}}}σ_{M_{\text{tr}}}\right)^2}$$  \hspace{1cm} (25)

$$σ_{V_{\text{t}}} = \sqrt{\left(\frac{σ_S}{M_{\text{tr}}}\right)^2 + \left(\frac{σ_{S_{\text{strat}}}}{M_{\text{tr}}}\right)^2 + \left(\frac{(S - S_{\text{strat}})}{M_{\text{tr}}}σ_{M_{\text{tr}}}\right)^2}$$  \hspace{1cm} (26)
The overall error in the slant column (σs) mainly reflects the uncertainty in the calibration of the reflectance spectrum and the dominating error: the measurement noise. Contributions from errors in the NO2 cross-section spectrum and the temperature dependence of the cross section are negligible. A value of 0.55×10^{15}molec./cm² was found by Boersma et al. (2007) for v.1.02 retrievals based on Collection 2 lvl1b data [Dobber et al., 2008] by subtracting the across-track variability (striping) from the total variance of the NO2 vertical column (slant columns divided by $M_{geo}$) in a small box over the Pacific Ocean. Since all vertical columns are expected to be similar over the Pacific, the variance gives an indication for the error in the slant column. The same procedure will be applied in Section 4.6 to estimate a value for the error in the NO2 slant column for v3 retrievals based on Collection 3 lvl1b data. A value of 0.4×10^{15}molec./cm² was used for DOMINO v2.

The error in the subtraction of the stratospheric background contribution (σbg) was previously estimated to be 0.25×10^{15}molec./cm² based on the average difference between the model slant column forecast and analysis after assimilation with OMI. The value should be relatively small since stratospheric NO2 in TM is assimilated with OMI measurements to optimize correspondence.

The error in the AMF (Eq. 27) consists of contributions of errors in the cloud fraction, cloud pressure, surface albedo, a priori profile ($x_a$), and mixing in TM (mix). Based on Boersma et al. (2004) a term representing the error correlation ($\rho_{A_s,c}$) between the error in the cloud fraction and surface albedo is taken into account. The AMF also depends on the viewing geometry, however, the measurement geometry is known with such high accuracy that uncertainties do usually not contribute to the error in the AMF. For the pixels at the edge of the swath, errors in the sphericity correction are expected to be smaller than 2% [van Geffen et al., 2013].

$$\sigma_{M_tr} = \left( \sigma^2_{M_tr}\sim c + \sigma^2_{M_tr}\sim p_c + \sigma^2_{M_tr}\sim A_s + \sigma^2_{M_tr}\sim x_a + \sigma^2_{M_tr}\sim mix + 2\rho_{A_s,c} \sigma_c \sigma_{A_s} \right)^{1/2}$$  \hspace{1cm} (27)

The AMF errors due to errors in to $c$, $p_c$, and $A_s$ are calculated by multiplying the assumed error in the quantity with the derivative of $M_{tr}$ to the quantity at the value of $M_{tr}$ (Equation 28). The assumed error in the cloud pressure (σp_c) is 50 hPa and the error in the cloud fraction (σc) 0.05, both based on Accareta et al. (2004). The AMF is most sensitive to a change in the cloud pressure when the cloud is located at a height near the peak in the NO2 vertical profile, which is close to the surface in regions with high emissions. The sensitivity to the cloud fraction is highest for low cloud fractions, especially over polluted regions. The assumed error in the albedo (σAs) is 0.02, the sensitivity of σM_tr to the albedo is highest for albedos in the 0.0-0.2 range [Boersma et al., 2004].

$$\sigma_{M_tr} = \left( \left( \frac{\partial M_{tr}}{\partial c} \sigma_c \right)^2 + \left( \frac{\partial M_{tr}}{\partial p_c} \sigma_{p_c} \right)^2 + \left( \frac{\partial M_{tr}}{\partial A_s} \sigma_{A_s} \right)^2 + \sigma^2_{M_tr}\sim x_a + \sigma^2_{M_tr}\sim mix + 2\rho_{A_s,c} \sigma_c \sigma_{A_s} \right)^{1/2}$$  \hspace{1cm} (28)

The error contribution of the a priori profile ($\sigma_{M_tr}\sim x_a$) is estimated to be 10% of the tropospheric AMF based on the work of Hains et al. (2010). They found that replacing a priori TM4 profiles in DOMINO v2 with measured profiles from the INTEX-B campaign changed the retrieved tropospheric NO2 column by up to 15%. This error contribution also takes into account the error in the emissions in TM since they directly influence the a priori profile. The mixing error estimates the error in the AMF related to errors in the boundary layer mixing in the TM model. The error is calculated by considering the difference in AMF between having all NO2 of the lowest 5 model layers concentrated in the lowest layer (no mixing) or divided equally over the five layers (total mixing). Especially for high emission areas, this contribution is large, as will be the topic of discussion in Section 4.6.
An explicit error for the presence and influence of aerosols in the atmosphere is not calculated. Because of the nature of the cloud retrieval, an implicit aerosol correction is applied by the use of the effective cloud fraction and pressure. Boersma et al. (2004) concluded that an explicit aerosol correction (and error estimation) is complicated because of the entanglement of the cloud parameters and aerosol concentrations. Furthermore, Boersma et al. (2011) demonstrated the entanglement of coinciding MODIS scattering aerosol and OMI effective cloud retrievals over the South-Eastern US. This implies a disentanglement should start within the cloud retrieval.

For unpolluted pixels over the ocean, the error will be dominated by the error in the observation (slant column) and stratospheric contribution. However, over polluted regions, the error in the AMF dominates. This means that for pixels over populated regions, the error in the tropospheric vertical column is dominated by the knowledge about the state of the atmosphere.

3.4 The OMI cloud retrieval

The cloud retrieval for OMI is based on measuring the continuum reflectance and depth of the absorption by the collision-pair of oxygen (O$_2$-O$_2$) [Acarreta et al., 2004]. The absorption is the result of a short time interaction between two colliding O$_2$ molecules and depends on the square of the O$_2$ pressure. The strongest collision complex absorption can be found at 477 nm, the band used for the cloud retrieval. Since clouds effectively shield O$_2$-O$_2$ complexes below the cloud, the measured O$_2$-O$_2$ slant column can be used to retrieve the height of clouds. The result is an effective cloud pressure near the middle of the cloud for clouds with a vertical extent from 1 - 5 km. This principle is visualized in Figure 7. The effective cloud fraction is found by matching a modeled top-of-atmosphere reflectance with the observed continuum reflectance [Acarreta and de Haan, 2002].

The retrieval of cloud pressure and cloud fraction follows a two step approach. In the first step, the O$_2$-O$_2$ slant column ($N_s$) is retrieved using a DOAS approach similar to the one used for NO$_2$ in DOMINO but now between 460 and 490 nm. This also yields the continuum reflectance ($R_c$, the reflectance in the absence of absorption). In the second step, the retrieved $N_s$ and $R_c$ are the input for the look up tables (LUT) together with geometrical parameters ($\theta, \theta_0, \phi_{rel}$), terrain height and, albedo that describe the knowledge of the atmospheric state. A linear interpolation of the LUTs (computed by offline radiative transfer modeling discussed in section 3.4.1) gives an effective cloud pressure and cloud fraction.

The theoretical error in the cloud fraction is estimated to be smaller than 0.05 and the error in the cloud pressure is smaller than 100 hPa. For pixels with an effective cloud fraction above 0.2 and cloud pressure below 300 hPa, the error is smaller than 50 hPa. However, since scenes with a cloud fraction under 0.2 are most important for the DOMINO retrieval, an error of 100 hPa is a conservative threshold [Acarreta et al., 2004].
Figure 7: Measurement principle of the cloud pressure. OMI measures the slant column of O$_2$-O$_2$ down to the effective cloud pressure. The O$_2$-O$_2$ under the cloud is shielded from OMI making the O$_2$-O$_2$ slant column (S) an effective measure for the effective cloud pressure ($p_c$) which is located at the center of the cloud. The cloud retrieval also gives an effective cloud fraction, which is not visualized here.

3.4.1 Computation of the cloud retrieval look up table

The DAK radiative transfer model is used to compute a relation between the input parameters ($N_s$ and $R_c$) and the cloud height and cloud fraction. The radiative transfer model takes into account polarization effects and multiple Rayleigh scattering. DAK also uses the independent pixel approximation and assumes the mid-latitude summer temperature profile [Anderson et al., 1986]. Both the cloud and surface are assumed to be opaque Lambertian surfaces with fixed albedos. The albedo of the cloud is assumed to be 0.8 since it is impossible to retrieve both cloud fraction and albedo independently from one reflectance spectrum. The same assumption is made in the retrieval of NO$_2$.

Radiative transfer simulations are done for all possible atmospheric states. In essence, this means that the reflectance is modeled as a function of all the parameters in the 7-D LUT\(^{vi}\), resulting in values for $N_s$ and $R_c$. The results are placed in inverted look up tables so that for every measurement of $R_c$ and $N_s$, the corresponding cloud fraction and cloud pressure can be found in the retrieval.

\(^{vi}\)The related cloud height is subsequently converted into an effect cloud pressure by using the mid-latitude summer pressure profile.
3.5 **ODIN HNO$_3$**

As described in Section 3.2.4, stratospheric HNO$_3$ concentrations in TM5 are nudged towards a climatology to compensate for the incomplete stratospheric chemistry. In Chapter 5, we will introduce nudging of stratospheric HNO$_3$ based on the ratio between the Odin climatologies of HNO$_3$ and O$_3$ which are described here.

Odin is an international aeronomy and astronomy mini satellite mission led by Sweden which accommodates two instruments, the Optical Spectrograph and Infra-Red Imaging System (OSIRIS) and the Sub-Millimetre Radiometer (SMR) [Frisk et al., 2003]. Here we use the climatologies of HNO$_3$ [Urban et al., 2009] and O$_3$\textsuperscript{vii} based on measurements by SMR. SMR measures the vertical profile of trace gases using the thermal emissions of the gases between 480 and 580 GHz. This measuring principle, unlike OMI’s, does not require sunlight. Using 14-15 orbits and 60 limb scans per day, global volume mixing ratio profiles are measured every three days. The instrument is able to measure concentrations above 15 km with a horizontal resolution of around 300 km and a vertical resolution of 3 km. The resulting climatology of HNO$_3$ has been compared to the UARS/MLS climatology from the 1990s and good agreement was found. As an advantage, Odin provides a better resolution and more recent data than UARS/MLS [Urban et al., 2009].

The zonal average Odin climatologies are publicly available for every month between August 2001 and April 2009 at 10° latitudinal resolution on [http://odin.rss.chalmers.se/climatologies](http://odin.rss.chalmers.se/climatologies). A more comprehensive discussion of the HNO$_3$ climatology can be found in Urban et al. (2009).

\textsuperscript{vii}SMR measurements of O$_3$ have been validated by Jégou et al. (2008).
4 Retrieval development results

Starting point for this project was the Dutch OMI NO$_2$ retrieval (DOMINO) v2.0 as described in Boersma et al. (2011) which uses the Tracer Model 4 (TM4) [Dentener et al., 2003]. The TM4 model is scientifically frozen, it is no longer maintained or developed. Here we present a number of improvements to the DOMINO retrieval. We start with a better representation of relevant atmospheric physics and chemistry by using the successor of TM4, TM5 (Section 4.1), for a priori profiles and the estimation of the stratospheric column. Subsequently, we improve the a priori TM5 NO$_2$ profiles by increasing the model resolution (Section 4.2). Furthermore, we implement a better calculation of the effective surface pressure (Section 4.3) and correct a temperature-dependent bias in the cloud retrieval (Section 4.4). Finally, we solve an issue with the interpretation of the AMF look-up table (Section 4.5) and revisit the computation of errors for the tropospheric NO$_2$ product in Section 4.6. Section 4.7 gives an overview of the combined impact of the retrieval improvements in the prototype DOMINO v3.

4.1 OMI NO$_2$ retrieval with TM5

We update DOMINO from TM4 [Dentener et al., 2003 and Williams et al., 2009] to TM5$^{viii}$ [Krol et al., 2005 and Huijnen et al., 2010]. TM5 is a supported model with the potential to be used at a higher resolution than $3^\circ \times 2^\circ$ (v2). For the retrieval of NO$_2$, the most important change is the update of anthropogenic NO$_x$ emissions. The emissions over Europe and the US are lower in the RETRO inventory (TM5) compared to the POET inventory (TM4) consistent with observed reductions in NO$_x$ emissions between 1997 and 2000 [Richter et al. 2005, Frost et al. 2006, and others]. Emissions over Europe have decreased from 4.7 Tg N yr$^{-1}$ (POET) to 4.1 Tg N yr$^{-1}$ (RETRO). This reduction is expected to decrease NO$_2$ concentrations near the surface in TM5 and subsequently increase the calculated AMF. The increased AMF will lead to reduced retrieved tropospheric NO$_2$ columns.

![OMI tropospheric NO$_2$ columns for 20-30 October 2004](image)

Figure 8: OMI tropospheric NO$_2$ columns for 20-30 October 2004. The left panel shows the currently available DOMINO v2 product, which uses TM4. The right panel shows the DOMINO v3 prototype which uses TM5. For all shown NO$_2$ tropospheric vertical column maps, pixels are only taken into account if the cloud radiance fraction is smaller than 0.5 and the surface albedo is smaller than 0.2. For difference plots, pixels from the compared retrievals both have to obey the criteria to be taken into account, this way, only directly comparable pixels are considered. Negative columns are shown as white pixels, grey pixels indicate the lack of clear-sky measurements.

$^{viii}$The DOMINO retrieval is now run as a project within the TM5 framework, simplifying future updates and allowing the retrieval to be run using the TM5 compilation structure.
Figure 8 shows the DOMINO v2 tropospheric NO\textsubscript{2} columns (based on TM4 at 3°×2°) on the left and the updated TM5-based (3°×2°) DOMINO (v3 prototype) columns on the right, averaged over 20-30 October 2004. Both retrievals are initialized with the same SCIAMACHY based restart file at October 1st 2004. The figures show a very similar spatial distribution of NO\textsubscript{2}. By going from TM4 to TM5, the number of negative columns decreases with 27\% for this time period. The reduction of the number of negatives is most likely caused by the different vertical layers in TM5 compared to TM4 and the different O\textsubscript{3} climatologies used in the models. The different layers facilitate a better representation of the boundary between the stratospheric and tropospheric part of the slant column. The O\textsubscript{3} climatology used for the stratosphere (Section 3.2.4) is important for the stratospheric NO\textsubscript{2} simulation because it drives HNO\textsubscript{3} concentrations in TM through the assumed climatological O\textsubscript{3}:HNO\textsubscript{3} ratio from UARS. HNO\textsubscript{3} in turn influences the modeled stratospheric NO\textsubscript{2}. Finally, reduced columns are visible over regions with anthropogenic emissions or biomass burning. This change is mostly related to the updated emission inventories.

Figure 9: Difference in DOMINO tropospheric NO\textsubscript{2} columns retrieved with the TM4 and TM5 model using the retrieval-assimilation-modeling approach. The main changes are a reduction of the number of negative columns over the oceans and lower columns related to updated anthropogenic emissions.

Figure 9 presents the difference between the new TM5-based retrieval and the old TM4-based retrieval. The global average tropospheric NO\textsubscript{2} column does not show a large change for 20-30 October 2004 by going from TM4 to TM5 (the global average is constant at 0.50×10\textsuperscript{15} molec./cm\textsuperscript{2}). However, the European ([-20°,45°;35°,65°]) average decreases by 10\% from 1.83×10\textsuperscript{15} molec./cm\textsuperscript{2} to 1.66×10\textsuperscript{15} molec./cm\textsuperscript{2}, reflecting the lower emissions in especially eastern Europe from RETRO compared to POET. Similarly, concentrations over the eastern US ([-90°,-68°;35°,45°]) decrease by 15\% and columns over China
([85°,125°;30°,40°]) decrease by 9%. The decrease over China is caused by a decrease of emissions close to Beijing. Meanwhile, emissions over mainland China have increased related to economic development between 1997 and 2006. While total emissions decrease over continents, they are more concentrated in hot spot areas in RETRO compared to POET, causing an increase of 7% for tropospheric columns over the Netherlands ([4°,7°;51°,53.5°]). Furthermore, a reduction in the number of negative columns over high and low latitude oceans can be seen as an increase in tropospheric NO2 in Figure 9. The average column over the Pacific (here defined as [-180°,-140°;50°,25°]) increases by 37% to 0.22 × 10^15 molec./cm^2. As can be seen in Figure 9, the posteriori stripe-correction [Boersma et al., 2011] has not yet been applied to the DOMINO v3 prototype columns, leading to a stripe pattern in the difference figure. For the final release of DOMINO v3, the stripe-correction will be applied. We conclude that the transition from TM4 to TM5 does not cause dramatic changes in retrieved tropospheric NO2. While patterns of enhanced NO2 are very similar, reductions over polluted regions are primarily related to the improved a priori NOx emissions in TM5.

4.2 The impact of higher a priori NO2 profiles on the retrieved tropospheric NO2 columns

DOMINO v2 used TM4 a priori tropospheric NO2 profiles at 3° × 2° for the calculation of the AMF. As suggested previously by Boersma et al. (2007) and demonstrated by Heckel et al. (2011), using better resolved (higher resolution) a priori profile shapes leads to a better understanding of spatial gradients in NO2. While DOMINO v2 has shown good consistency with ground-based MAX-DOAS measurements (e.g. Mendolia et al. (2013) and Irie et al. (2012)), it underestimates ground-based measurements of tropospheric NO2 columns over hot spot regions like Beijing [Ma et al., 2013] and Delhi [Shaiganfar et al., 2011]. This underestimation has been largely attributed to the gradient smoothing effect of low resolution a priori profiles. This gradient smoothing effect has been estimated to be up to 20% by Ma et al. (2013). The transition to TM5 now allows an increase of the used model resolution from 3° × 2° to 1° × 1°, presumably reducing gradient smoothing effects over hot-spot regions. To illustrate the overall effect of the resolution change, Figure 10 shows tropospheric NO2 columns over Europe as simulated by the (OMI assimilated) TM5 model for 20-30 October 2004 at 3° × 2° (left) and 1° × 1° (right).

Figure 10: Tropospheric NO2 columns over Europe simulated by (the OMI assimilated) TM5 at 3° × 2° (left) and 1° × 1° (right) for 20-30 October 2004 at 13:30 h local time, the overpass time of OMI. The image illustrates the importance of using higher resolution a priori profiles to correctly represent patterns in tropospheric NO2. The white box indicates the region discussed in the Barcelona analysis.
The importance of the increased resolution is especially visible close to big cities. Where previously, emissions from the city were smeared out over the large grid cell, they are now mainly concentrated in one of the six $1^\circ \times 1^\circ$ cells. This can be for example clearly seen for Barcelona in Spain. Averaged globally, tropospheric columns are very similar in both simulations ($0.31 \times 10^{15}$ molec./cm$^2$ for both resolutions). Over Europe, there is a small increase (from $1.76 \times 10^{15}$ molec./cm$^2$ to $1.82 \times 10^{15}$ molec./cm$^2$), probably indicative of too rapid numerical diffusion in the $3^\circ \times 2^\circ$ simulation [Huijnen et al., 2010a]. The resolution upgrade also influences the a priori TM4 profile shape which is important for the AMF calculation. Hot spots areas (like megacities) will have more NO$_2$ in the lowest layers because of the improved spatial allocation of emission sources. Even more significantly, outflow regions will have lower NO$_2$ concentrations near the surface compared to the $3^\circ \times 2^\circ$ simulation. This is illustrated in Figure 11 which gives the vertical NO$_2$ profiles for the white box given in Figure 10, containing Barcelona. The figure shows the vertical profile of the $3^\circ \times 2^\circ$ simulation, the average of the six $1^\circ \times 1^\circ$ cells that make up the $3^\circ \times 2^\circ$ cell in the high resolution simulation, and the six individual $1^\circ \times 1^\circ$ cells including one for Barcelona. The figure shows considerable variability between the profiles of the different $1^\circ \times 1^\circ$ cells. Most notably, the cell that now contains all the emissions from Barcelona, given by the blue line, shows a surface concentration that is three times as large as in the $3^\circ \times 2^\circ$ simulation. On top of that, the concentration of NO$_2$ in the free troposphere is lower over Barcelona at $1^\circ \times 1^\circ$ compared to $3^\circ \times 2^\circ$. Since now more NO$_2$ is located near the surface, where the sensitivity of OMI is low, the AMF will be lower, resulting in a higher retrieved tropospheric NO$_2$ column over Barcelona. The comparison of the averages shows that small deviations occur for the free troposphere, where the upward transport from the boundary layer in the coarse resolution simulations was facilitated by horizontal numerical diffusion and the large size of the grid cell.

Figure 11: Vertical NO$_2$ mixing ratios for the white box shown in Figure 10 containing Barcelona. The orange line shows the profile for the $3^\circ \times 2^\circ$ grid cell, the black line shows the average of the six $1^\circ \times 1^\circ$ grid cells contained within the boundaries of the $3^\circ \times 2^\circ$ cell. Furthermore, the dashed lines show the six individual $1^\circ \times 1^\circ$ cells, the cell containing Barcelona is shown in blue.
Figure 12 shows the difference in retrieved tropospheric NO$_2$ columns over Europe by improving the resolution of TM5. In general, all hot spot areas like cities and industrial regions show elevated concentrations relative to the retrieval with the coarse a priori profiles. This reflects the increase in the simulated NO$_2$ concentration near the surface and lower free troposphere concentrations in the 1$^\circ$ × 1$^\circ$ retrieval. Since relatively more NO$_2$ of the total column is now attributed to the lower troposphere, where the sensitivity of the instrument is lower, the tropospheric AMFs decrease. A decreased AMF leads to an increased tropospheric vertical column, as observed over cities in Figure 12. This is expected to reduce the (smoothing gradient) bias between OMI and ground-based MAX-DOAS measurements in megacities. The average tropospheric NO$_2$ column over Europe decreases by 5% (from 1.66 × 10$^{15}$ molec./cm$^2$ to 1.57 × 10$^{15}$ molec./cm$^2$) while retrieved columns over the Netherlands increase by about 7%. The tropospheric columns over the localized hot spot Barcelona ([2$^\circ$,3$^\circ$;42$^\circ$,43$^\circ$]) increase by 22% related to the higher a priori NO$_2$ concentration near the surface as discussed above. The overall decrease over Europe is caused by the fact that for the calculation of the AMF, the reduced surface NO$_2$ over outflow areas has a relatively larger impact than the increased surface NO$_2$ over hot spot areas. This result is in good agreement with Figure 13 in Huijnen et al. (2010), who used DOMINO averaging kernels on the high resolution EURAD-IM model to demonstrate the potential improvement when using very high resolution (0.4$^\circ$ × 0.4$^\circ$) a priori profile shapes from that model instead of from the coarse TM4 CTM.

Figure 12: Difference in retrieved DOMINO tropospheric NO$_2$ columns over Europe between assimilation with TM5 at 1$^\circ$ × 1$^\circ$ and 3$^\circ$ × 2$^\circ$. The resolution increase leads to higher a priori surface NO$_2$ over cities, decreasing the AMF and increasing the retrieved vertical column. Over outflow areas, a priori surface NO$_2$ decreases, increasing the tropospheric AMF and subsequently decreasing the retrieved vertical NO$_2$ columns.

---

One has to note that those values cannot directly be compared to the tropospheric columns simulated by TM5 because not the whole European domain is covered by observations during the 20-30 October time period.
4.3 Implementing terrain height averaging in DOMINO

The AMF calculation requires a representative value of the surface pressure for the spatial extent of the OMI pixel. However, the only available surface pressure in the retrieval code is based on the coarse resolution TM5 surface pressure \( \rho_{TM5} \) for the grid cells that coincide with the OMI pixel. Therefore, DOMINO v2 corrected an effective pixel surface pressure by taking into account the difference between the smoothed TM terrain height and the terrain height from a high resolution climatology (DEM_3KM).

In this procedure, the TM5 surface pressure is converted to a surface pressure associated with the OMI pixel \( \rho_{eff} \) by following the hypsometric equation and assuming linearly changing temperature with height (Equation 29) [Wallace and Hobbs, 1977].

\[
\rho_{eff} = \rho_{TM5} \times \left( \frac{T_{surf}}{T_{surf} + \Gamma(h_{TM5} - h_{eff})} \right)^{-\frac{g}{R}} \tag{29}
\]

With \( T_{surf} \) the surface temperature, lapse rate \( \Gamma \) (6.5 K km\(^{-1}\)), gas constant for dry air \( R \) (287 J kg\(^{-1}\) K\(^{-1}\)), and gravitation constant \( g \) (9.8 ms\(^{-2}\)). Finally, \( h_{TM5} \) and \( h_{eff} \) are respectively the TM5 grid cell surface height and OMI pixel height from the Global 3 km Digital Elevation Model data (DEM_3KM) provided with the OMI slant columns. As described in Zhou et al. (2009) and Boersma et al. (2011), this correction has a significant impact on the calculated AMF and thus on retrieved NO\(_2\) over regions with a high level of orographic variability. The sensitivity of the AMF to the surface pressure depends on the cloud pressure and fraction, and the a priori NO\(_2\) profile shape [Zhou et al., 2009]. For cloud free pixels, a higher surface pressure (lower terrain height) means that more NO\(_2\) can be expected low in the atmosphere which leads to a lower tropospheric AMF and thus a higher retrieved vertical tropospheric NO\(_2\) column. For partly cloudy pixels, the AMF is more sensitive to a change in terrain height because the overall sensitivity below the cloud is very low and the estimated fraction of a priori NO\(_2\) underneath the cloud can make a dramatic impact. The impact of this correction (Equation 29) is largest in winter when, because of the shallow boundary layer, most NO\(_2\) is located close to the surface. Columns over Beijing were found to be higher by up to 14% in January because of a net decrease in terrain height of 450 m.

Here, we make an additional improvement to the surface pressure correction. As illustrated in Figure 13, in the original DOMINO v2.0 retrieval, \( h_{eff} \) is based on the sampling of the high resolution DEM_3KM database at the center of the OMI pixel only. This leads to a discrepancy between the terrain height reported with the OMI slant columns and the average terrain height representative for the pixel. To better represent the true pixel terrain height, we sample the DEM_3KM database over the complete spatial extent of an OMI pixel as illustrated on the right side of Figure 13. An averaging grid with a distance between sampling points of 1 km is laid over the OMI pixel. This means that for larger pixels at the edges of the OMI swath, more samples are taken. The resolution of the averaging grid is higher than

![Figure 13: Schematic overview of the old DOMINO v2.0 (left) and new prototype DOMINO v3.0 (right) procedures to determine the effective terrain height of a pixel. In the old situation, the DEM_3KM database was sampled at the center of the pixel. In the new procedure, the database is sampled on a 1 km grid laid within the pixel, leading to an average terrain height for the pixel.](image-url)
the resolution of the DEM database, leading to an effective weighing of the degree of overlap between a DEM_3KM cell and the OMI pixel.

Figure 14 shows the impact of the updated terrain height. The left plot shows the difference between the pixel-averaged terrain height and the terrain height at the center of the pixel (DOMINO v2). As expected, large differences occur over mountainous areas, resulting in local changes in $h_{eff}$ of over 500m. The global average terrain height increases with about 0.7% for valid pixels because pixels with a decrease in terrain height (mountains) are often covered by snow and are filtered out. As shown in the right map, the associated change in $p_{eff}$ leads to considerable changes in NO$_2$ over areas with a lot of NO$_2$ near the surface. Especially eastern China shows substantial local changes in retrieved NO$_2$ in the order of $0.5 \times 10^{15}$ molec./cm$^2$. The average tropospheric NO$_2$ over China decreases by less than 1% related to an increase in the average terrain height of less than 1%.

Figure 14: The left image shows the difference in retrieved terrain height between using the pixel averaging described above and using the terrain height at the center of the pixel. The right image shows the impact of the new pixel averaging on the retrieved tropospheric NO$_2$. Especially areas with large terrain height differences over small scales and high NO$_2$ concentrations (for example China), show up as affected.

Furthermore, the new terrain height avoids extreme surface pressures resulting from the surface pressure correction [Lin et al., 2013]. The paper by Lin et al. (2013) showed that the center sampling of the OMI pixel occasionally resulted in large differences between the TM5 and OMI terrain height so that the corrected surface pressure would occasionally reach values over 1050 hPa. By averaging the terrain height over the pixel, such extreme surface pressure values no longer occur. This is illustrated in Figure 15, which shows different height values for 100 subsequent pixels on 20 October 2004. The left graph shows $h_{TM5}$, the old $h_{eff}$, and the new $h_{eff}$. The right graph shows the resulting $p_{eff}$ using both the old and new $h_{eff}$. Both graphs illustrate that the smoothing of the terrain height leads to a smaller range of surface pressures. In conclusion, the improved surface pressure calculation leads to more realistic surface pressures and only significantly affects retrieved NO$_2$ over hot spot regions with large terrain height differences like China.
4.4 Correcting the OMI cloud retrieval temperature dependence

The AMF calculation requires accurate information on effective cloud fraction and cloud pressure. The DOMINO NO\textsubscript{2} retrieval uses the cloud parameters from the OMCLDO2 OMI cloud product. The OMCLDO2 retrieval uses the absorption of the O\textsubscript{2}-O\textsubscript{2} collision complex at 477 nm to determine the cloud fraction ($c$) and effective cloud pressure ($p_c$) [Acarreta et al., 2004]. The radiative transfer calculations in the retrieval assume a constant mid-latitude summer (mls) temperature profile climatology ($T_{mls}$) [Anderson et al., 1986] for every measurement. Since air density itself depends on temperature ($T$), the O\textsubscript{2}-O\textsubscript{2} density will depend on $T^2$ [de Haan, 2010]. Previous cloud retrievals used the O\textsubscript{2} column for their calculations. When the O\textsubscript{2} column is used, the temperature dependence will drop out of the equations, rendering the O\textsubscript{2} column temperature independent. However, the O\textsubscript{2} absorption is not within the wavelength range measured by OMI so cannot be used. A derivation of the temperature dependence of the O\textsubscript{2}-O\textsubscript{2} column is given in Appendix B.

The use of the mls temperature profile introduces an error in the correct interpretation of the radiative transfer modeling results whenever the true temperature profile of the atmosphere deviates from the climatology [de Haan, 2010]. Errors are introduced in the cloud fraction and most significantly in the cloud pressure. The bias is highest for measurements with small cloud fractions which are most important for the NO\textsubscript{2} retrieval. Figure 16 shows the constant $T_{mls}$ profile and (true) $T_{TM5}$ profile for 20-30 October 2004 over London. The figure shows a substantial difference between the mls climatology and the ERA-INTERIM temperature profile in TM5.

The OMI cloud retrieval uses a look up table (LUT, Section 3.4) to match the retrieved O\textsubscript{2}-O\textsubscript{2} slant column ($N_{\text{slant,meas}}$) and observed continuum reflectance to a simulated slant column and reflectance with an associated cloud pressure and cloud fraction. The retrieved slant column is correct but the interpretation of the slant column by the LUT leads to errors because the slant columns in the LUT have been calculated using the mls atmosphere with an invariant temperature profile. The temperature-dependent bias could be corrected in the OMI cloud retrieval, but the use of daily temperature profiles would require far-reaching changes to the cloud retrieval. A correction within the DOMINO framework is easier to implement because of the availability of ERA-INTERIM temperature profiles and the calculation of altitude dependent AMFs and reflectances. Here, we calculate a pseudo slant column (‘corrected’ OMI column)
Figure 16: The graph shows the mid-latitude summer temperature profile which is used in the original OMI cloud retrieval and the temperature profile as provided by the ERA-interim reanalysis averaged over London [-1°.1°;51°.52°] for 20-30 October 2004.

that would have been measured if the atmosphere would have had the mls temperature profile. The use of the pseudo column in the LUTs of the cloud retrieval gives the correct cloud parameters. We apply a correction to the measured O$_2$-O$_2$ slant column ($N_{s, \text{meas}}$), using the actual temperature profile as provided by the ERA-interim reanalysis used in TM5 ($T_{\text{T M5}}$) to represent the state of the atmosphere during the measurement. We follow the correction proposed by de Haan (2010) given in Eq. 30, with $N_{s, \text{mls}}$ the 'corrected' OMI O$_2$-O$_2$ column matched to the LUT, $N_{s, \text{meas}}$ the originally retrieved O$_2$-O$_2$ column, $p$ the atmospheric pressure, $p_{\text{TOA}}$ the pressure at the top of the atmosphere, and $m$ the NO$_2$ altitude-dependent air mass factors as calculated by the DOMINO retrieval. Theoretically, one should use the O$_2$-O$_2$ AMFs for this correction factor. However, the O$_2$-O$_2$ AMFs are not available and the NO$_2$ AMFs (439 nm) calculated in the DOMINO product strongly resemble the O$_2$-O$_2$ AMFs (477 nm) and are appropriate to be used here.

$$N_{s, \text{mls}} = N_{s, \text{meas}} \times \frac{\int_{p_{\text{TOA}}}^{p_{\text{TOA}}} m(p) \frac{p}{p_{\text{TOA}}(p)} dp}{\int_{p_{\text{TOA}}}^{p_{\text{TOA}}} m(p) \frac{p}{p_{\text{TOA}}(p)} dp}$$ \hspace{1cm} (30)

Equation 30 only holds for scenes with a cloud cover of 100%. By following the independent pixel approximation, Eq. 30 can be rewritten to Equation 31 by distinguishing the clear and cloudy part of the pixel, giving a generic description for every cloud fraction. In this equation, $R_{\text{clr}}$ is the reflectance from the clear part of the pixel, $R_{\text{cld}}$ the reflectance from the cloudy part of the pixel, and $p_s$ the surface pressure.

$$N_{s, \text{mls}} = N_{s, \text{meas}} \times \frac{(1 - c)R_{\text{clr}} \int_{p_{\text{TOA}}}^{p_{\text{TOA}}} m(p) \frac{p}{p_{\text{TOA}}(p)} dp + cR_{\text{cld}} \int_{p_{\text{TOA}}}^{p_{\text{TOA}}} m(p) \frac{p}{p_{\text{TOA}}(p)} dp}{(1 - c)R_{\text{clr}} \int_{p_{\text{TOA}}}^{p_{\text{TOA}}} m(p) \frac{p}{p_{\text{TOA}}(p)} dp + cR_{\text{cld}} \int_{p_{\text{TOA}}}^{p_{\text{TOA}}} m(p) \frac{p}{p_{\text{TOA}}(p)} dp}$$ \hspace{1cm} (31)

The correction factor is applied to the O$_2$-O$_2$ slant column by integrating over the TM5 pressure levels. The altitude dependent AMFs, reflectances, and cloud parameters that are required for the calculation of the correction factor, are initially calculated with the original $p_c$ and $c$ from the cloud product. The
'corrected' OMI slant column is then matched to the lookup tables of the cloud retrieval to calculate the corrected cloud fraction and cloud pressure.

The relation between the surface temperature in TM5 and the correction in the cloud pressure is given in Figure 17. The largest changes occur for temperatures far below the mid-latitude summer surface temperature of 294 K. For temperatures higher than the MLS temperature profile, the cloud pressure increases. This is in good accordance with the results of Boersma et al. (2011) who found too low cloud pressures over the southeastern US during the summer when temperatures were higher than the MLS climatology.

![Figure 17](image_url)

Figure 17: The graph shows the relation between surface temperature and the correction to the cloud pressure for 30 October 2004. The largest changes occur for temperatures far below the mid-latitude summer surface temperature of 294 K. The largest changes in cloud pressure occur over the polar regions, because of the presence of snow and ice, the NO_2 measurements for those locations will not be taken into account. For surface temperatures larger than 294 K, cloud pressures are increased, in good accordance with Boersma et al. (2011) who found too low OMI cloud pressures during the 2005 US heat wave. The blue line shows the fraction of pixels per surface temperature.

Figure 18 shows distribution of the impact of the temperature correction. The top graph shows the change in cloud pressure for 20-30 October 2004. For high and low latitudes, the lower than mid-latitude summer surface temperature (294 K) leads to a smaller O_2-O_2 column. Generally, smaller O_2-O_2 columns imply higher clouds, so the cloud pressure decreases. The increase of the height of the cloud increases the screening (thus reduces the sensitivity to NO_2) which reduces the tropospheric AMF. Consequently, the correction leads to higher tropospheric NO_2 columns for high and low latitudes. The change in tropospheric NO_2 is shown in the bottom graph of Figure 18. The largest differences occur for regions where a large abundance of NO_2 is present near the (corrected) height of the cloud. As an example, a substantial decrease in tropospheric NO_2 is visible over the UK and the Netherlands. Here, in polluted regions (high NO_2 concentrations related to large anthropogenic emissions) with lower tropospheric clouds and temperature below the mid-latitude summer conditions (Figure 16), the temperature correction leads to an increase of the tropospheric NO_2 column. Differences of up to 11% occur over London ([-1°,1°;51°,52°]) related to the temperature difference shown in Figure 16. Over China, the average tropospheric NO_2 column increases by 5% because of the correction of the cloud pressure.

Because the surface albedo at 477 nm is not known with the same precision as in the original cloud retrieval, a small bias is introduced in the resulting cloud fraction and cloud pressure compared to the original retrieval. However, this bias is negligible compared to the temperature-dependent correction.
Figure 18: The top image shows the impact of the temperature correction on the retrieved cloud pressure. For high and low latitudes, the real temperature is lower than the temperature used in the retrieval, this reduces the cloud pressure (thus increasing the height of the cloud). As shown in the bottom figure, a decrease in cloud pressure leads to an increase in retrieved tropospheric NO$_2$. The correction only has a big impact on locations with large NO$_2$ concentrations near the height of the cloud.
4.5 Correcting the relative azimuth angle dependence of the AMF

The AMF depends on the relative azimuth angle ($\phi_{rel}$, Eq. 32) which is conventionally defined as the absolute difference between the viewing azimuth angle ($\phi$) and the solar azimuth angle ($\phi_0$).

$$\phi_{rel} = |\phi - \phi_0|$$ (32)

In DOMINO v2.0, the relative azimuth angle has been erroneously defined as $|\phi_0 - \phi|$, as a result of the ambiguous output of an early version of the DAK model (Version 3.0, 18 August 2003, J. de Haan). The different relative azimuth angle is not taken into account in the interaction between the DOMINO retrieval and the LUT, causing an error in the AMF computation which grows for larger values of $\phi_{rel}$.

To correct the dependence, we compute a separate value of the relative azimuth angle to be used for the linear interpolation of the LUT, $\phi_{LUT}$:

$$\phi_{LUT} = |\pi - \phi_{rel}|$$ (33)

The left graph of Figure 19 shows the relative change in the tropospheric AMF after the correction for 15 November 2004. Differences range from -15% to +15%. The right graph of the figure shows the impact on the resulting vertical tropospheric NO$_2$ columns for the same day. Differences in both directions of the order of $1.0 \times 10^{15}$ molec./cm$^2$ occur in regions with high tropospheric NO$_2$ columns like the eastern US. Since the correction has an altering sign related to the relative azimuth angle of the pixel, the influence of the bug fix will largely cancel out for longtime averages. Furthermore, the global average tropospheric NO$_2$ decreases with 0.5% for 15 November 2004 while the number of negative tropospheric NO$_2$ columns decreases with 2%. The fact that the number of negatives decreases provides additional confidence that the LUT is now interacted with in the right way.

Figure 19: Impact of the bug fix for the interaction between DOMINO and the AMF LUT considering the relative azimuth angle. The figures show the relative impact on the tropospheric AMF (left, ranging from -15% to +15%) and the absolute impact on the tropospheric vertical NO$_2$ column (right, differences up to $10^{15}$ molec./cm$^2$ over hot spot areas) for 15 November 2004. Over longer time periods, the impact is less severe because of the differing sign of the impact.
4.6 Updated error calculation for DOMINO

We present an improved error calculation for the DOMINO algorithm. Changing the error estimates of the cloud fraction, surface albedo, slant column density, and stratospheric slant column will affect errors calculated for the total, tropospheric and stratospheric vertical columns; and total and tropospheric averaging kernels. Here we focus on the impact on the error in the main product, the tropospheric vertical \( \sigma_{\text{NO}_2} \).

The cloud fraction error currently used in DOMINO is based on Acarreta et al. (2004). Stammes et al. (2008) found a high correlation between MODIS [King et al., 1998] and OMI cloud fractions of 0.92 with an average bias of 0.01 and a standard deviation of 0.12 (mainly contributed to the model used to convert MODIS cloud optical thickness to cloud fractions). Here we estimate the error in the cloud fractions for OMI measurements of \( \text{NO}_2 \) that satisfy the filtering criteria, therefore we focus on measurements with a solar zenith angle smaller than 65\(^\circ\). To estimate the error we evaluate the probability distribution of unclipped cloud fractions (cloud fractions used in the retrieval are set to (clipped) 0 when retrieved as smaller than 0 and set to 1 when retrieved as larger than 1 before being used) as shown in Figure 20 by the solid line. The tail of negative cloud fractions is mirrored in the y-axis and the standard deviation of this distribution (the dashed line in Figure 20) serves as a simple estimate for the uncertainty in the cloud fraction. The standard deviation in this distribution is found to be 0.017. As a conservative estimate, we assume the error in the cloud fraction to be 0.025, half the value previously assumed for DOMINO v2 (0.05).

Figure 20: Probability distribution of unclipped (not cut off at 0) cloud fractions for measurements with a solar zenith angle smaller than 65\(^\circ\) for 1 October 2004. The tail of negative cloud fractions is mirrored in the x-axis and the resulting probability distribution is used to calculate the standard deviation from 0 as a measure for the error in the cloud fraction. Cloud fractions used in the retrieval are clipped at 0 and 1 before being used.

Secondly, we decrease the assumed error in the surface albedo (Lambertian equivalent reflectance, LER) from 0.02 to 0.015 based on Kleipool et al. (2008) who state the accuracy of the their LER product as: ”0.01 to 0.02 for the longer wavelengths and increasing toward the shorter ultraviolet wavelengths”. Furthermore, for DOMINO v2, the error in the cloud pressure was estimated to be 50 hPa [Boersma et al.,
2011]. Based on Sneep et al. (2008), this error estimation might have been too optimistic. However, since we presented an update to the cloud pressure resolving the temperature bias of on average 50 hPa in Section 4.4; we argue that the error estimate for the cloud pressure given for DOMINO v2 (50hPa) is now a reasonable estimate for the error in the cloud pressure for the DOMINO v3 prototype presented here.

As a third change, the mixing error no longer contributes to the error in the AMF. The mixing error was previously used to represent the potential error in the boundary layer mixing in TM. In DOMINO v2, the mixing error was calculated as half the difference in AMF between having all the NO\(_2\) of the lowest 5 layers concentrated in the first layer ("no mixing") and distributed equally over the 5 layers ("well mixed"). This error strongly overstated the actual error. TM5 NO\(_2\) profiles haven been proven to compare well with profiles measured by the aircraft INTEX-B campaign (Section 3.2.5, [Huijnen et al., 2010b]). While the contribution of the mixing error to the global average error was in the order of 1%, the error would reach unrealistic large values for regions with a lot of NO\(_2\) in the surface layer, becoming responsible for more than 75% of the AMF error. We now hypothesize that the error in the a priori profile is already sufficiently represented by the profile error which is formulated as 10% of the AMF (Section 3.3.5).

Based on Dirksen et al. (2011) we reduce the error in the error in the stratospheric slant column from \(0.25 \times 10^{15}\) molec./cm\(^2\) to \(0.2 \times 10^{15}\) molec./cm\(^2\). The value given by Dirksen et al. (2011) is based on an examination of the average difference between the model stratosphere forecast and model analysis after assimilation, and a comparison with measurements from ground stations at high altitude.

Finally, we revisit the error assumptions for the slant column density which is \(0.4 \times 10^{15}\) molec./cm\(^2\) in DOMINO v2. The error in the retrieved slant column is estimated by evaluation of the variance of the slant column (\(\sigma_{\text{var}}\)) in 25 \(2^\circ \times 2^\circ\) boxes over the Pacific [-180°W, -178°W ; -60°S, -10°S] for 20-30 Oct 2004. Over the Pacific, all vertical column measurements can expected to be similar. For each slant column, we calculate its difference with the box average. To calculate a representative average, each slant column is first divided by its own \(M_{\text{geo}}\) (Eq. 21) and the resulting average vertical column is subsequently multiplied with the average \(M_{\text{geo}}\). To find the variance of slant columns, a Gaussian function is fitted through the distribution. The standard deviation of the fitted Gauss curve is found to be: \(\sigma_{\text{var}} = 0.65 \times 10^{15}\) molec./cm\(^2\). The variance in the slant column (Eq. 34) is a combination of the error in the slant column (\(\sigma_{\text{meas}}\)) and the effect of the striping (\(\sigma_{\text{stripe}}\)). For the start of the mission, the striping has been estimated to contribute 0.36 \(\times 10^{15}\) molec./cm\(^2\) by Boersma et al. (2011).

\[
\sigma_{\text{var}}^2 = \sigma_{\text{meas}}^2 + \sigma_{\text{stripe}}^2
\]

Using Eq. 34 we infer an error in the measured slant column of \(0.54 \times 10^{15}\) molec./cm\(^2\). This is a higher value than used for DOMINO v2 (0.4 \(\times 10^{15}\) molec./cm\(^2\)). The discrepancy is caused by the fact that the DOMINO v2 error parametrization estimated a higher influence of striping, which is valid for later years of the mission. For our error parametrization, we will use a slant column error of \(0.55 \times 10^{15}\) molec./cm\(^2\).

The top graph of Figure 21 shows the distribution of errors calculated using the new assumptions for the error in the slant column, stratospheric column, surface albedo, and cloud fraction and discarding the mixing error. The figure on the bottom shows the difference with the old error parametrization for the same DOMINO v3 retrieval. The largest differences occur for areas with large tropospheric NO\(_2\) columns where the error in the AMF dominates the overall error budget. The impact for several key regions over the world is shown in Table 3.
Figure 21: The top graph shows the average error in the vertical tropospheric NO\textsubscript{2} column calculated using the new error parametrization for 20-30 October 2004 in 10\textsuperscript{15} molec./cm\textsuperscript{2}. The figure shows the increasing error with the tropospheric NO\textsubscript{2} column. The larger the tropospheric column, the bigger the error in the tropospheric AMF used to calculate the vertical column. The bottom graph shows the difference between the old and new error parametrization for the same DOMINO v3 prototype retrieval for 20-30 October 2004. The largest differences occur for regions with large abundances of NO\textsubscript{2} in the boundary layer where the discarding of the mixing error has the largest impact.
Table 3: The effect of the improved DOMINO error parameterization. The first column states the region defined by the boundaries given in column two. Column three give the average tropospheric NO\textsubscript{2} column for that region for 20-30 October 2004 in 10\textsuperscript{15} molec./cm\textsuperscript{2}, column four gives the old error in 10\textsuperscript{15} molec./cm\textsuperscript{2}, and column five gives the new error in 10\textsuperscript{15} molec./cm\textsuperscript{2}. Column six gives the relative change in the average error: \(100 \times \frac{\sigma_{V_{tr,\text{new}}} - \sigma_{V_{tr,\text{old}}}}{\sigma_{V_{tr,\text{old}}}}\).

<table>
<thead>
<tr>
<th>Region</th>
<th>Coordinates</th>
<th>(V_{tr}) (*)</th>
<th>(\sigma_{V_{tr},\text{old}}) (*)</th>
<th>(\sigma_{V_{tr},\text{new}}) (*)</th>
<th>Relative change</th>
</tr>
</thead>
<tbody>
<tr>
<td>World</td>
<td>-180°, -180° - 60°, 80°</td>
<td>0.52</td>
<td>0.67</td>
<td>0.58</td>
<td>-13%</td>
</tr>
<tr>
<td>Eastern US</td>
<td>-90°, -60° - 35°, 45°</td>
<td>4.57</td>
<td>4.10</td>
<td>2.69</td>
<td>-34%</td>
</tr>
<tr>
<td>China</td>
<td>85°, 125° - 30°, 40°</td>
<td>3.45</td>
<td>2.31</td>
<td>1.77</td>
<td>-23%</td>
</tr>
<tr>
<td>Europe</td>
<td>-20°, 45° - 35°, 65°</td>
<td>2.07</td>
<td>2.07</td>
<td>1.41</td>
<td>-32%</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>4°, 7° - 51°, 53.5°</td>
<td>7.53</td>
<td>5.47</td>
<td>3.84</td>
<td>-30%</td>
</tr>
<tr>
<td>London</td>
<td>-1°, 1° - 51°, 52°</td>
<td>8.87</td>
<td>7.22</td>
<td>4.32</td>
<td>-41%</td>
</tr>
<tr>
<td>Barcelona</td>
<td>2°, 3° - 41°, 42°</td>
<td>6.40</td>
<td>3.06</td>
<td>2.09</td>
<td>-32%</td>
</tr>
</tbody>
</table>

* In 10\textsuperscript{15} molec./cm\textsuperscript{2}

To give a simple representation of the error, we fit a linear slope for the error in the tropospheric NO\textsubscript{2} column as a function of the tropospheric NO\textsubscript{2} column. The data (points and bold line) and fit (dashed line) for the old (black) and new (blue) error parameterization are shown in Figure 22. The relation found for the old parametrization is: \(\sigma_{V_{tr}} = 0.66V_{tr} + 0.45 \times 10^{15}\) molec./cm\textsuperscript{2}. While the new error parameterization is characterized by: \(\sigma_{V_{tr}} = 0.42V_{tr} + 0.45 \times 10^{15}\) molec./cm\textsuperscript{2}. This error quantification has a larger slope but smaller constant part than the relation reported by Boersma et al. (2011) for DOMINO v2: \(\sigma_{V_{tr}} = 0.25V_{tr} + 1.00 \times 10^{15}\) molec./cm\textsuperscript{2} shown by the orange curve in Figure 22. The intercept of the error curve with the y-axis is determined by errors in the slant column, the slope is mainly related the AMF error. Even though our relation is only based on 10 days of data, it provides a relevant representation of errors in DOMINO to compare to Boersma et al. (2011) since our method is based on propagated errors and their associated tropospheric NO\textsubscript{2} columns.

Figure 22: Errors in the vertical tropospheric NO\textsubscript{2} column as a function of the vertical tropospheric NO\textsubscript{2} column. The black points and line represents the error calculation with the mixing error and old error assumptions for the albedo and cloud fraction. Values calculated using the new parametrization are shown in blue. Linear fits of the old and new errors are shown as dashed lines. The error for DOMINO v2 as given by Boersma et al. (2011) is shown by the orange line.
Table 4: The change in tropospheric NO\textsubscript{2} per improvement for different regions. The first column states the region as defined in Table 3. Additionally, the Pacific is defined as [-180\degree,-140\degree;-50\degree,25\degree]. In the second column, the average tropospheric NO\textsubscript{2} column over the region as retrieved with DOMINO v2 is given in 10\textsuperscript{15} molec./cm\textsuperscript{2}. The third column gives the relative difference in the average 20-30 October 2004 tropospheric NO\textsubscript{2} column related to the transition from TM4 to TM5. The fourth-seventh columns give the difference related to: improving the spatial resolution to 1\degree\times1\degree, implementing terrain height averaging to calculate the effective surface pressure, the temperature-dependent correction of the cloud pressure, and the relative azimuth angle bug fix. All individual differences are calculated between a retrieval with the update and a retrieval without. The last column gives the average tropospheric NO\textsubscript{2} column as retrieved with the prototype DOMINO v3 in 10\textsuperscript{15} molec./cm\textsuperscript{2} and the overall difference with v2. The prototype DOMINO v3 prototype includes all improvements from this chapter.

<table>
<thead>
<tr>
<th>Region</th>
<th>v2*</th>
<th>TM5</th>
<th>1\degree\times1\degree</th>
<th>p_{eff}</th>
<th>p_{c}</th>
<th>\Phi_{rel}</th>
<th>v3*</th>
</tr>
</thead>
<tbody>
<tr>
<td>World</td>
<td>0.50</td>
<td>0%</td>
<td>-8%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0.45  (-10%)</td>
</tr>
<tr>
<td>Pacific</td>
<td>0.16</td>
<td>+38%</td>
<td>-14%</td>
<td>0%</td>
<td>0%</td>
<td>-1%</td>
<td>0.19  (+19%)</td>
</tr>
<tr>
<td>Eastern US</td>
<td>5.50</td>
<td>-15%</td>
<td>-4%</td>
<td>0%</td>
<td>+1%</td>
<td>0%</td>
<td>4.54  (-17%)</td>
</tr>
<tr>
<td>China</td>
<td>3.19</td>
<td>-9%</td>
<td>-11%</td>
<td>-1%</td>
<td>+5%</td>
<td>0%</td>
<td>2.73  (-14%)</td>
</tr>
<tr>
<td>Europe</td>
<td>1.83</td>
<td>-10%</td>
<td>-5%</td>
<td>0%</td>
<td>+2%</td>
<td>0%</td>
<td>1.58  (-14%)</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>9.68</td>
<td>+8%</td>
<td>+7%</td>
<td>0%</td>
<td>+5%</td>
<td>+2%</td>
<td>11.88 (+23%)</td>
</tr>
<tr>
<td>London</td>
<td>8.13</td>
<td>+4%</td>
<td>0%</td>
<td>+11%</td>
<td>-1%</td>
<td>9.29</td>
<td>6.03  (+23%)</td>
</tr>
<tr>
<td>Barcelona</td>
<td>4.89</td>
<td>+3%</td>
<td>+22%</td>
<td>0%</td>
<td>+2%</td>
<td>0%</td>
<td>6.03  (+23%)</td>
</tr>
</tbody>
</table>

* In 10\textsuperscript{15} molec./cm\textsuperscript{2}

4.7 Combined impact of the retrieval improvements

We now discuss the combined impact of the individual retrieval improvements. The impact of the five major retrieval developments for several important regions is shown in Table 4.

The third column shows the impact of the transition from TM4 to TM5. In general, concentrations over Europe and the US decrease because of lower overall anthropogenic emissions in TM5. However, since the emissions in RETRO are more concentrated in hot spots than in POET, tropospheric NO\textsubscript{2} over cities (London and Barcelona) increases. The fourth column gives the relative change in tropospheric NO\textsubscript{2} related to the increase in model resolution from 3\degree\times2\degree to 1\degree\times1\degree. As concentrations over outflow areas decrease because of lower a priori NO\textsubscript{2} near the surface, average concentrations over the continents (Eastern US, China, and Europe) are lower. On the other hand, modeled surface concentrations for emission hot spots are higher, leading to lower tropospheric AMFs and an increase in the tropospheric column of over 20\% in cities like Barcelona. The impact of the OMI terrain height averaging is shown in the fifth column. For the considered areas, the average impact is always smaller than 2\%. The sixth column shows the difference caused by the cloud correction. Over most shown regions, the average temperature for October is lower than in the mid-latitude summer climatology. This lower temperature leads to higher clouds, reducing the sensitivity of the instrument to NO\textsubscript{2} and in turn, in higher tropospheric columns by 1\% up to 11\%. The last shown impact is the impact of the relative azimuth angle bug fix. As the impact generally averages out of the course of multiple days, only minor differences are seen for the shown 11 day averages. The last column gives the tropospheric NO\textsubscript{2} columns as retrieved with the prototype DOMINO v3 and overall differences between v3 and v2. Because the different modifications have been tested individually (usually at 3\degree\times2\degree), the final difference between v3 and v2 may deviate from the sum of all changes reported in Table 4. Overall, concentrations over larger regions decrease because of lower columns over outflow areas and lower overall emissions. Retrieved columns over cities generally increase because of the higher model resolution.
Figure 23 shows the tropospheric NO$_2$ columns as retrieved with the prototype DOMINO v3 for 20-30 October 2004 in the top graph and the difference with the DOMINO v2 columns for the same time period in the bottom graph. The difference figure shows the combined impact of the TM update, the increased resolution, the surface pressure improvement, the cloud correction, and the relative azimuth angle bug fix. In general, concentrations over outflow regions decrease while concentrations over hot spot regions and the oceans (used to be negative) increase. The number of negative pixels is decreased with 15% in comparison with DOMINO v2. The resulting errors are not shown here but are highly similar to the ones presented in Figure 21.

Figure 23: The top figure shows the tropospheric NO$_2$ columns for 20-30 October 2004 retrieved with the prototype DOMINO v3. The retrieval is done with TM5 at 1° $\times$ 1°, including the cloud correction, relative azimuth angle bug fix, and terrain height averaging. Grey pixels indicate the lack of clear-sky measurements, for this figure, negative pixels are shown as dark blue instead of the previously used white. The bottom figure shows the difference between tropospheric NO$_2$ columns retrieved with the prototype DOMINO v3 and DOMINO v2 (20-30 October 2004). Decreases are seen over the continents related to lower emissions in TM5 compared to TM4. Furthermore, retrieved columns over hot spots increase while columns over outflow areas decrease because of the transition to 1° $\times$ 1°. Lower retrieved columns are partly compensated for higher latitudes by the cloud correction which leads to higher columns. Finally, tropospheric columns over oceans increase compared to DOMINO v2.
5 Evaluation of the stratospheric TM assimilation

In this section, the assimilation of TM5 stratospheric NO$_2$ with OMI observations is examined. In the first section (5.1) the assimilation is analyzed by examination of the adjustment of the model through the assimilation and the difference between the observation and forecast. In the second section, the nudging of stratospheric HNO$_3$ to the Odin climatology is discussed. The nudging of HNO$_3$ is - in combination with the chemistry - the main driver of stratospheric NO$_2$ in the TM5 model.

5.1 Using O-F and A-F statistics to diagnose assimilation in DOMINO

The difference between the observation (O) and the model forecast (F) of the NO$_2$ slant column (O-F) and the difference between the analysis (A) and the forecast of the slant column (A-F) can be used to diagnose the performance of the assimilation. Figure 24 shows the difference between O and F for 10 November (top figure) and the average for 5-15 November (bottom figure). Figure 25 shows the difference between analysis and forecast for 10 November in the top graph and the bottom figure shows the average between 5 and 15 November. In the assimilation, the TM forecast (and analysis) is converted into a slant column forecast by applying the altitude-dependent AMFs. Here we use the geometric approximation of the AMF to convert the measured slant column and slant column forecast/analysis into total vertical NO$_2$ columns.

The O-F graph for 10 November in Figure 24 clearly shows the striping related to calibration differences for different viewing angles. Since no such pattern is present in the model forecast, the striping shows up in the O-F figure. Furthermore, large differences between observation and forecast are visible over hot spot areas. Over those areas the tropospheric contribution to the total column is large which makes the forcing of the model towards the observation less stringent because of the large error associated with the tropospheric contribution (Equation 23 in Section 3.3.4). Comparing the O-F for 10 November with the A-F for the same day (top graph of Figure 25) gives a good indication of how the assimilation works. The striping is filtered out in the A-F plot for 10 November and only large scale differences in O-F that are not related to tropospheric pollution have adjusted the modeled stratosphere and show up in the differences between the analysis and forecast.

Looking at the 11 day averages, the difference between the observation and forecast is generally within $0.5\times10^{15}$molec./cm$^2$, except for high emission areas where the assimilation is set to be less stringent and differences are over $1.0\times10^{15}$molec./cm$^2$. The differences between the analysis and the forecast are generally smaller (of the order $0.1\times10^{15}$molec./cm$^2$) showing that the model is close to equilibrium with the assimilation. Comparing 11 day averages from Figures 25 and 24 again shows that the assimilation (A-F) is hardly influenced by tropospheric pollution and small-scale concentration patterns. This is the result of the higher error associated with high tropospheric contribution measurements (Eq. 23) and the spatial correlation (Eq. 24) which has a smoothing effect on the model adjustment.
Figure 24: The OMI observation minus TM5 forecast of the total vertical column for 10 November (top figure) and 5-15 November (bottom figure) 2004. Both observation and forecast of the slant column were divided by the same geometrical AMF (Eq. 21) to get the forecasted and measured total vertical columns. The DOMINO v3 TM5-based retrieval with the cloud correction and surface pressure averaging is used at $3^\circ \times 2^\circ$.
Figure 25: The TM5 analysis minus TM5 forecast of the total vertical column for 10 November (top figure) and 5-15 November (bottom figure) 2004. Both analysis and forecast were divided by the same geometrical AMF (Eq. 21). The same DOMINO version used for Figure 24 is used. The difference between analysis and forecast gives the model adjustment towards the satellite observation as described in Section 3.3.4.
To study the assimilation of TM5 in more detail, we focus on the O-F and A-F diagnostics above the Pacific Ocean. By choosing this region, we minimize the influence of tropospheric pollution and small-scale concentration differences. As shown in Figure 26, both the distribution of O-F and A-F show near Gaussian behavior with respective standard deviations of 0.33 and 0.08. The Gaussian distribution of O-F and O-A is required for the Kalman filter and is in good accordance with Eskes et al. (2003) who found similar behavior for the assimilation of GOME total ozone satellite observations.

Examination of patterns in O-F as a function of retrieval parameters may help identify systematic errors in any of the retrieval steps. Figure 27 shows the difference between the observation and forecast (black line) and the analysis and forecast (green line) relative to the model forecast (F) as a function of a number of parameters characterizing the retrieval. The figures show the dependence on: the vertical column observation (top left), solar zenith angle (top right), relative azimuth angle (bottom left), and pixel number (bottom right). The top left figure shows that there are large differences between the observation and forecast, we attribute these to the effects of striping. Very small observations (smaller than $2 \times 10^{15}$ molec./cm$^2$) are related to negative striping and will be smaller than the forecast. On the other edge of the graph, large observations are related to positive striping and will exceed the model forecast. That the A-F differences are close to zero percent on average shows that the stripes have little overall effect on the assimilation. The top right figure shows there is a small bias in both O-F and A-F as a function of the solar zenith angle which ranges from +3% to -3% in the A-F. This bias suggests small errors in the calculation of the AMF used to convert the vertical model forecast to the slant column forecast. The figure on the bottom right shows the average O-F and A-F as a function of the pixel number. It again shows that the stripe bias does not affect the assimilation, the striping is filtered out in the assimilation of the model (A-F). The across-track slope that was previously seen in the O-F and A-F figures for 10 November is now apparent in both the relative azimuth angle (bottom left) and pixel number dependence of A-F and O-F. The forecast is generally too high for pixels at the left side of the swath (large relative azimuth angle) and too small on the right side of the track (small relative azimuth angle). This bias suggests another error in the AMF computation. The cause of the possible bias in the AMF is currently under investigation and might be related to the adjustment of the AMF to the curvature of the Earth.
Figure 27: Visualization of O-F and A-F diagnostics for total vertical columns over the Pacific [-180°, -100°; -60°, 10°] for 5-15 November 2004. Clockwise from the top left corner: O-F and A-F relative to the forecast as a function of the vertical column observation, solar zenith angle, pixel number, and the relative azimuth angle.

Figure 28 shows the spatial correlation of O-F and A-F over the Pacific Ocean. The orange line gives the theoretical second order Thiebaux correlation function (Eq. 24) with a correlation length of 600 km, as used in the assimilation. The correlation coefficients in O-F are rather low because of the striping. Because of the occurrence of striping, there are large differences between the observations for neighboring pixels. To study the correlation between pixels without being influenced by the striping, the correlation is calculated between pixels with the same pixel number (same viewing zenith angle). The correlation in the O-F for pixels of the same viewing zenith angle is significantly higher, as shown by the green curve in Figure 28.

As the adjustment of the model is partly determined by the spatial correlation between pixels (Section 3.3.4) the A-F correlation is expected to show good agreement with the second-order Thiebaux correlation function that is used in the assimilation. Figure 28 shows that this is indeed the case. Considering the
spatial correlation while leaving out the self-correlation\textsuperscript{xii}, the correlation length for O-F is found to be (61 ± 4) km between all pixels (strongly striping related) and (470 ± 20) between pixels with the same pixel number. The significant higher along-track correlation is caused by exclusion of the striping. The correlation length of A-F is found to be (683 ± 5) km, close to the correlation length of the used Thiebaux correlation function.

Figure 28: Correlation of O-F (black line) and A-F (blue line) as a function of the distance between two pixels. The correlation in the O-F is strongly limited by the striping, therefore, the green line shows the correlation in O-F for pixels with the same viewing zenith angle (pixel number). The orange line shows the theoretical Thiebaux correlation (Eq. 24) with a correlation length of 600 km which is used to determine the model adjustment.

\textsuperscript{xii}The correlation lengths are found by fitting a scaled Thiebaux correlation function
As a next step, we consider the resulting tropospheric NO$_2$ columns for the same region and time. The top left figure of 29 shows the probability distribution of the tropospheric NO$_2$ columns over the Pacific for 5-15 November 2004. Like the O-F and A-F diagnostics, this probability function has a Gaussian distribution. The shown Gaussian fit is centered around 0.20 \(10^{15}\) molec./cm$^2$ (a realistic background value for tropospheric NO$_2$) and has a standard deviation of 0.45 \(10^{15}\) molec./cm$^2$.

Figure 29: Tropospheric NO$_2$ columns over the Pacific [-180°,-100°,-60°,10°] for 5-15 November 2004. The top left figure shows the probability distribution of the tropospheric column. The blue line gives a Gaussian fit which is centered around 0.20 \(10^{15}\) molec./cm$^2$ and has a standard deviation of 0.45 \(10^{15}\) molec./cm$^2$. The other graphs show the dependence of the tropospheric column on: the solar zenith angle, relative azimuth angle, and pixel number. Tropospheric columns with a cloud radiance fraction larger than 50% or albedo larger than 0.2 are filtered out.

Additionally, Figure 29 shows the tropospheric NO$_2$ column as a function of the solar zenith angle (top right), relative azimuth angle (bottom left), and pixel number (bottom right). The tropospheric column as function of the solar zenith angle shows that pixels with negative values in the Pacific box tend to have a large solar zenith angle. The strong dependence of the tropospheric product on the solar zenith angle is again suggestive of systematic errors in the AMF calculation that should be investigated before the release of DOMINO v3. Furthermore, the bottom figures show that striping effect is propagated into the tropospheric column. This can be expected since the tropospheric contribution to the slant column is
calculated by subtracting the forecast for the stratosphere (no striping) from the observation which does contain striping. As mentioned before, a posteriori stripe-correction is in place for DOMINO v2 and will be implemented for v3.

As a conclusion, the examination of DOMINO using O-F, A-F and the tropospheric column as diagnostics has given confidence in the functioning of the assimilation. The striping does not perturb the assimilation and in an average sense all three diagnostics show a Gaussian distribution. Furthermore, the average difference between analysis and forecast for 5-15 November 2004 is generally within $0.1 \times 10^{15}$ molec./cm$^2$. However, the examination has also revealed possible errors in the AMF computation: negative tropospheric columns occur often for large values of the solar zenith angle and the AMF seems to introduce an across-track bias in the forecast.
5.2 Using Odin HNO$_3$ forcing in the TM5 stratosphere

In TM5, HNO$_3$ is nudged towards a climatology in the stratosphere. Since the actual source of stratospheric NO$_x$ (N$_2$O) is missing, the HNO$_3$ nudging is the effective source of NO$_x$ and therefore NO$_2$ in the stratosphere. For all previous presented results, stratospheric NO$_x$ at 10 hPa is nudged towards the NO$_3$ concentration at 10 hPa and the retrieval is initialized with the same restart file as DOMINO v2 on 1 October 2004. The restart file is based on SCIAMACHY assimilation. Here, we compare three different set ups for the nudging of HNO$_3$ in TM5. The first one is the standard approach mentioned above which is equal to the set up used for DOMINO v2. However, the retrieval does include all improvements from Chapter 4. As a second approach, we nudge HNO$_3$ in the stratosphere with the standard TM5 setting that directly prescribes the HNO$_3$ concentration based on the climatology. NO$_x$ concentrations above 10 hPa are not nudged and the assimilation is initialized with a free model spin up until October 1st from when OMI measurements are assimilated. This will be referred to as the ’prescribing UARS HNO$_3$’ set up.

For the third set up, ’nudging Odin HNO$_3$’, we introduce the nudging of stratospheric HNO$_3$ towards the Odin climatology discussed in Section 3.5. While the UARS nudging is only applied at 10 hPa, we now also nudge the 5.5 hPa and 28 hPa levels as described by Equation 35.

\[
[HNO_3]_{TMS}^{\text{TMS},-\delta t} + [HNO_3]_{\text{Clim}} \times \frac{\delta t}{\delta t_{nudge}}
\]

Here, \([HNO_3]_{TMS}^{\text{TMS},-\delta t}\) is the nudged HNO$_3$ mixing ratio, \([HNO_3]_{TMS}^{\text{TMS},-\delta t}\) is the HNO$_3$ concentration before nudging, \(\delta t\) is the model time step, and \(\delta t_{nudge}\) is the nudging time scale. The used \(\delta t_{nudge}\) depends on the pressure level and is 5 days for 5.5 hPa, 20 days for 10 hPa, and 60 days for 28 hPa. The HNO$_3$ concentration based on the climatology (\([HNO_3]_{\text{Clim}}\)) is given in Equation 36.

\[
[HNO_3]_{\text{Clim}} = \left(\frac{[HNO_3]}{[O_3]}\right)_{\text{Odin}} \times [O_3]_{TMS} \times \frac{[HNO_3]_{TMS}}{[HNO_3]_{TMS}^{\text{TMS},-\delta t}}
\]

The climatological HNO$_3$ concentration is calculated by multiplying the ratio of the zonal average Odin HNO$_3$ and O$_3$ climatologies with the zonal averaged modeled O$_3$ concentration: \([O_3]_{TMS}\). The Odin HNO$_3$ and O$_3$ climatologies are available with a resolution of 10° and are linearly interpolated to a resolution of 1° [J.E. Williams, personal communication]. The last term of Equation 36 gives the ratio between the TM5 HNO$_3$ mixing ratio (\([HNO_3]_{TMS}^{\text{TMS},-\delta t}\)) of the grid cell before nudging and the longitudinal average TM5 HNO$_3$ mixing ratio for that latitude (\([HNO_3]_{TMS}^{\text{TMS},-\delta t}\)). Adding this term allows longitudinal variation of stratospheric HNO$_3$ in the model [H. J. Eskes, personal communication]. Following Dirksen et al. (2011), for pressure levels above 5 hPa, the mixing concentration of NO$_x$ is nudged towards the mixing ratio of NO$_3$ at 5 hPa with a nudging time of 20 days. Similar to the UARS set up, the Odin retrieval has been initialized with a free model run, now with Odin nudging in the stratosphere. An overview of the three set-ups is shown in Table 5.
Table 5: Three different set ups for the nudging of HNO$_3$ and the model initialization that is used for the set-up. The table shows the name of the set-ups, the HNO$_3$:O$_3$ ratio climatology that is used, the model layers at which the ratios are nudged, the nudging time, and the restart file that is used to initialize the model.

<table>
<thead>
<tr>
<th>Set-up</th>
<th>HNO$_3$</th>
<th>Nudging altitude</th>
<th>Nudging time</th>
<th>Restart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard approach (v2)</td>
<td>UARS</td>
<td>10 hPa</td>
<td>2 months</td>
<td>SCIAMACHY based</td>
</tr>
<tr>
<td>Prescribing UARS HNO$_3$</td>
<td>UARS</td>
<td>10 hPa</td>
<td>Direct forcing</td>
<td>Free model spin up</td>
</tr>
<tr>
<td>Nudging Odin HNO$_3$</td>
<td>Odin</td>
<td>5.5, 10, 28 hPa</td>
<td>5, 20, 60 days</td>
<td>Free model spin up</td>
</tr>
</tbody>
</table>

Figure 30 shows the three resulting TM5 stratospheric HNO$_3$ fields for 5-15 November 2004 at 13:30 h local time. Although HNO$_3$ is not directly assimilated with OMI measurements, there is an influence from the OMI assimilation (starting 1 October) because NO$_3$ species that are assimilated react to form HNO$_3$. By comparison with independent measurements [Wespes et al., 2007], it is clear that with the free model run initialization, the UARS prescribing at 10 hPa (bottom left figure) is not able to provide realistic stratospheric HNO$_3$ columns. The global distribution of HNO$_3$ found using the Odin HNO$_3$ ratios much better resembles the distribution found by Wespes et al. (2007, Figure 7) based on measurements by the Interferometric Monitor for Greenhouse Gases onboard ADEOS. The standard approach and Odin nudging provide similar stratospheric HNO$_3$ concentrations. The Odin retrieval does so without the help of an externally generated model initialization as is the case for the standard approach that uses the SCIAMACHY restart file.

Figure 30: Modeled (OMI assimilated TM5) stratospheric HNO$_3$ columns for 5-15 November 2004 at 13:30 h local time. The top figure shows the standard approach that is initialized with the DOMINO v2 restart file and uses UARS HNO$_3$ ratios with a nudging time of 2 months. The bottom left figure shows the prescribing UARS approach and the bottom right figure shows the Odin 3-layer nudging HNO$_3$ set up. Both bottom retrievals were initialized with a free model run until October 1st.
Figure 31 shows the partial columns of HNO$_3$ (left) and NO$_2$ (right) simulated with the three TM5 set-ups for 5-15 November 2004 at 13:30 h local time (the OMI overpass time) between 30°N and 60°N. Besides the partial columns from the three different model set-ups, the plots also show partial columns based on the Odin climatologies for HNO$_3$ and NO$_2$ averaged to the TM5 vertical grid. As expected, the HNO$_3$ Odin climatology agrees well with the HNO$_3$ profile produced by the Odin nudging. This proofs that the nudging of Odin HNO$_3$:O$_3$ ratios as described above, performs well. Furthermore, the Odin nudging gives lower HNO$_3$ concentrations under 100 hPa compared to the standard approach and UARS retrievals.

The NO$_2$ Odin climatology is based on measurements by the OSIRIS instrument aboard Odin\textsuperscript{xxii}. The climatology is averaged to the TM5 vertical grid and provides an independent validation\textsuperscript{xxiii}. Comparison of all NO$_2$ profiles with the Odin climatology shows too high stratospheric columns in the OMI assimilated TM5 model. This probably reflects a recently found high bias in the OMI slant columns of on average 30% [Hendrick et al., 2012, Van Geffen et al., 2013b, and Belmonte Rivas et al., 2013]. Because of the assimilation of OMI measurements in (mainly) the stratosphere, this will cause too large concentrations of stratospheric NO$_2$. On top of that, it is an open question whether the limited TM5 stratospheric chemistry is able to produce realistic NO$_2$ profiles even when the HNO$_3$ concentrations are perfect. Furthermore, the peak of NO$_2$ is located lower in the stratosphere for the Odin free model run, more NO$_2$ than in the UARS free model run was already located in the stratosphere.

\textsuperscript{xxii}It is impossible to use the Odin NO$_2$ climatology to nudge stratospheric NO$_2$ because it is based on measurements by OSIRIS and has large gaps related to the position of the Sun. For example, for January, no data is available for the northern hemisphere. On top of that, nudging of NO$_2$ ratios would make the retrieved OMI tropospheric column directly related to the NO$_2$ measurements of OSIRIS.

\textsuperscript{xxiii}The NO$_2$ climatology at 55°N is not available because of a lack of measurements, the Odin satellite value is based on 35° & 45°N.
because of the multi-layer nudging. This means that the assimilation will not have to scale the partial columns as much to be in equilibrium with OMI observations, leaving the free troposphere concentrations uninfluenced.

Figure 32 shows the differences in the retrieved NO\textsubscript{2} tropospheric and stratospheric columns between the UARS nudging set up and the standard approach (left figures), and the differences between the Odin nudging set-up and the standard approach (right figures). The UARS retrieval with prescribed HNO\textsubscript{3} shows lower stratospheric concentrations because of the low NO\textsubscript{2} concentrations in the initial model field (on the 1st of October) that lead to too high tropospheric concentrations when the assimilation starts. The low stratospheric concentrations lead to unrealistically high retrieved tropospheric columns. The set-up with the nudging of Odin HNO\textsubscript{3} shows a small reduction of NO\textsubscript{2} in the troposphere for the lower latitudes compared to the standard approach. This decrease of NO\textsubscript{2} is associated with a larger number of negative columns. Besides this change, the nudging of Odin HNO\textsubscript{3} provides results reasonably similar to the standard approach.

Figure 32: Differences in retrieved OMI tropospheric NO\textsubscript{2} and stratospheric NO\textsubscript{2} in \(10^{15}\) molec/cm\(^2\) for the free model initialized retrievals compared to the standard approach. The standard approach is initialized with the SCIAMACHY restart file and uses UARS HNO\textsubscript{3} nudging at 10 hPa with a nudging time of 2 months. The left figures show the differences between the standard approach and the set-up with the UARS HNO\textsubscript{3} prescription at 10 hPa. The figures on the right show the difference between the standard approach and the Odin HNO\textsubscript{3} nudging at 3 levels.
As another diagnostic, we evaluate the average O-F and the difference between the observation and the model analysis (O-A). Table 6 shows the global average O-F and O-A and their standard deviations for 10 November 2004. After assimilation with OMI for over 40 days, there is still a large average O-F for the retrieval with the prescribed UARS HNO$_3$ and free model spin up. This difference between model and observation is also seen after assimilation (O-A). The values found for the standard approach and Odin retrievals are significantly lower and similar, giving confidence that Odin is able to provide NO$_2$ concentrations in the stratosphere that are in equilibrium with OMI. The main advantage of the Odin approach is that it does not require the SCIAMACHY based restart file to be in equilibrium with OMI.

Table 6: Results of the analysis of the average O-F and O-A and standard deviations ($\sigma$) for 10 November 2004 using different ways to nudge HNO$_3$ in the stratosphere and different spin ups to start the retrieval.

<table>
<thead>
<tr>
<th>Set-up</th>
<th>$&lt;\text{O-F}&gt;$</th>
<th>$\sigma$</th>
<th>$&lt;\text{O-A}&gt;$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard approach ($\nu$2)</td>
<td>-0.01</td>
<td>0.22</td>
<td>0.01</td>
<td>0.14</td>
</tr>
<tr>
<td>Prescribing UARS HNO$_3$</td>
<td>0.04</td>
<td>0.23</td>
<td>0.03</td>
<td>0.16</td>
</tr>
<tr>
<td>Nudging Odin HNO$_3$</td>
<td>-0.01</td>
<td>0.22</td>
<td>0.00</td>
<td>0.14</td>
</tr>
</tbody>
</table>

* In $10^{15}$ molec./cm$^2$

We conclude that the initial vertical distribution of NO$_2$ in the stratosphere is highly relevant for the retrieved tropospheric NO$_2$ columns. While the Odin nudging of HNO$_3$ in the stratosphere provides an improvement over using the prescription of UARS HNO$_3$, the increased number of negatives compared to the standard approach and modeled NO$_2$ profile that does not match observations are still worrisome. Since the vertical distribution of stratospheric NO$_2$ is mainly driven by the forcing of HNO$_3$ and the chemistry that provides the link between HNO$_3$ and NO$_2$, we recommend to investigate this issue further. Especially in view of the high bias in the OMI NO$_2$ slant columns.
6 Conclusion

We presented an improved tropospheric nitrogen dioxide column retrieval algorithm (prototype DOMINO v3) for OMI and diagnosed the assimilation of stratospheric NO2. As a first improvement, the DOMINO retrieval was coupled to the TM5 chemistry transport model instead of the TM4 model which was used in earlier versions. The TM5-based retrieval did not cause dramatic changes in the patterns of retrieved tropospheric NO2. However, the transition from TM4 to TM5 did reduce the number of negatives and decreased average retrieved NO2 concentrations over Europe, the US, and China because of updated NOx emission estimates in TM5 compared to TM4. As a second improvement, TM5 allowed an increase in resolution from 3°×2° to 1°×1°. The higher resolution for the simulation of NO2 vertical distributions led to a better understanding of the spatial distribution of NO2. Over pollution hot spots, the 1°×1° simulation better captured NO2 gradients leading to local increases in the retrieved tropospheric NO2 column of up to 22%. On the other hand, NO2 columns over outflow areas were lower because of reduced surface concentrations in the higher resolution model forecast.

As a third innovation, the representation of the surface pressure was improved. We implemented terrain height averaging over the OMI pixel replacing sampling at the center of the pixel. This change had a minor impact on retrieved columns on a continent scale but had a local effect of up to 0.5×1015 molec/cm² over mountainous areas. Fourthly, the temperature-dependent bias in the cloud retrieval was corrected using the TM5 temperature profile. This correction led to lower cloud pressures for regions with a temperature below the mid-latitude summer climatology. The decrease in cloud pressure resulted in an increase of retrieved tropospheric NO2 of up to 11% over London during November because of a stronger screening effect.

We corrected the relative azimuth angle dependence of the air mass factor calculation. This led to differences up to 1015 molec/cm² within one orbit. However, the impact averaged out for longer time periods because of the bipolar character of the correction. All changes combined led to lower global retrieved tropospheric NO2 columns by 10% because of lower columns over outflow areas and lower overall emissions. Retrieved columns over cities generally increased by up to 23%, mainly related to the higher model resolution. Subsequently, a more realistic error estimation was introduced that reduced the global average error in the tropospheric column by 13%. The main reductions occurred over hot spot areas (up to 41%) where the unrealistically high contribution of the mixing error was now discarded. We found that the absolute error in the tropospheric column for 5-15 November 2004 is best described by 0.45×1015molec/cm² plus 42% of the tropospheric NO2 column.

In an analysis of the assimilation of the stratosphere we concluded that striping does not affect the assimilation. However, a viewing geometry-dependent bias was found for the analysis minus forecast diagnostic, suggesting an error in the AMF calculation. Furthermore, a large number of negative tropospheric columns was found to be associated with large solar zenith angles, suggesting another systematic error in the AMF computation. Stratospheric HNO3 in TM5 was improved by implementing nudging of the HNO3:O3 ratio at three model layers based on the Odin climatology instead of using the UARS climatology at one layer only. However, the current analysis of the resulting retrieved tropospheric columns is complicated by the fact that the assimilated OMI columns are probably too high. Further research on the stratospheric chemistry and constraint of TM5 was recommended.
7 Outlook

Directly arising from this work are two subjects for follow-up research. The first is the bias in the difference between the observation and TM5 forecast of the total vertical NO$_2$ column as a function of the solar zenith angle and viewing zenith angle. These biases hint at an error in the computation of the altitude-dependent AMFs that are used to translate the modeled vertical column into the slant column. As a second issue, the posteriori stripe correction introduced for DOMINO v2 by Boersma et al. (2011) should be re-implemented in DOMINO v3 for the retrieved tropospheric NO$_2$ columns. The O-F and A-F diagnostics could serve as a foundation of an online stripe correction. Further down the horizon, two elements of the retrieval that could be improved will be discussed here: the slant column calculation and the a priori TM5 profiles.

Recent comparisons of co-located stratospheric NO$_2$ columns from various satellite instruments (SCIAMACHY, OMI, and GOME-2) reveal that OMI retrieves higher stratospheric NO$_2$ columns than other satellite instruments and ground-based instruments by on average 30% [Hendrick et al., 2012, Van Geffen et al., 2013b, and Belmonte Rivas et al., 2013]. The magnitude of the differences in NO$_2$ cannot be solely explained by the difference in measurement time (which is in the order of $0.5 \times 10^{15}$ molec./cm$^2$). Van Geffen et al. (2013b) found that using a wavelength calibration at the middle of the fit window (425.5-443 nm) instead of at the start of the window (408-423 nm) as done previously, reduces both the retrieved slant column and root-mean-square error of the fit. The slant column density decreases with on average 13% and the root-mean-square error with on average 20%. On top of that, it was found that reducing the fit window from 405-465 nm to 415-465, reduces the retrieved slant column as well, with about 11% (the root-mean-square error decreases with another 6%). The changes are largest over open ocean waters. Similar changes do not occur when the width of the fitting window is further decreased, indicative of a physical effect in the 405-415 nm window that is not yet properly taken into account in the DOAS fit. Current causes of the phenomena under consideration are the absorption by liquid water and/or the vibrational Raman scattering over open ocean waters [van Geffen et al., 2013].

On the other side of the retrieval, the a priori profiles from TM5 could be improved. The most at hand improvements are in the emissions of NO$_x$ from ships and soil. Instead of using the currently used emission inventory, emissions of soil NO$_x$ could be calculated online using the parametrization by Hudman et al. (2012) which was recently updated with a canopy reduction factor for implementation in GEOS-Chem by Maasakkers (2012). The large advantage of online calculation of soil emissions is their highly meteo-dependent timing. Large emissions of soil NO$_x$ occur in strong pulses with the first rain after a period of drought. Furthermore, shipping NO$_x$ emissions could be improved based upon Vinken et al. (2011) who account for the non-linear in-plume chemistry. Using their high resolution plume-in-grid approach to model ship emissions during the first hours after emission could reduce modeled NO$_x$ concentrations by up to 50% [Vinken et al., 2011]. As a last possible TM improvement, a full stratospheric chemistry (including N$_2$O, halogenated species, and photolytic destruction below 202 nm) would allow TM to generate realistic a priori profiles without the nudging of HNO$_3$ or O$_3$ boundary conditions in the stratosphere. This would also improve the spatial distribution of NO$_2$ in the stratosphere, no longer being influenced by a zonal average constraint. If the stratospheric nudging is maintained, more research is required to improve the simulation of stratospheric NO$_2$ via (Odin) HNO$_3$ forcing.

Finally, in 2014, the successor of OMI, the TROPOspheric Monitoring Instrument (TROPOMI) will be launched. For TROPOMI, the same DOMINO algorithm used for OMI (discussed here) will be applied [van Geffen et al., 2013]. However, with a spatial resolution of $7 \times 7$ km$^2$, the detail retrieved by TROPOMI will far surpass OMI measurements and represent a new step in our understanding of air pollution.
A \ NO_y\ chemistry in TM5

This appendix gives an overview of the NO_y chemistry in TM5. The information is based on Huijnen et al. (2010b). Table 7 gives the NO_y reactions in TM5 with their rate constants. Reaction rates are based on Sander et al. (2006), Atkinson et al. (2006)\textsuperscript{xiv}, and Yarwood et al. (2005)\textsuperscript{xv}. Table 8 gives the photolysis reactions taken into account in the TM5 NO_y chemistry.

Table 7: Overview of the NO_y chemistry in TM5 with rate constants.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Rate constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO + O_3 \rightarrow NO_2</td>
<td>3.0 \times 10^{-12} e^{-1390/T}</td>
</tr>
<tr>
<td>NO + HO_2 \rightarrow NO_2 + OH</td>
<td>3.5 \times 10^{-12} e^{250/T}</td>
</tr>
<tr>
<td>NO + CH_3O_2 \rightarrow NO_2 + CH_2O + HO_2</td>
<td>2.8 \times 10^{-12} e^{300/T}</td>
</tr>
<tr>
<td>NO_2 + OH (+M) \rightarrow HNO_3</td>
<td>K_2 = 1.8 \times 10^{-30} (300/T)^3</td>
</tr>
<tr>
<td>HNO_3 + OH \rightarrow NO_3</td>
<td>K_w = 2.8 \times 10^{-11}</td>
</tr>
<tr>
<td>NO_2 + O_3 \rightarrow NO_3</td>
<td>K_0 = 0.1 \times 10^{-11} (300/T)^{0.9}</td>
</tr>
<tr>
<td>NO + NO_3 \rightarrow 2 NO_2</td>
<td>K_2 = 2.6 \times 10^{-11} (2199/T)^{0.7}</td>
</tr>
<tr>
<td>NO_2 + NO_3 \rightarrow N_2O_5</td>
<td>1.2 \times 10^{-13} e^{-2450/T}</td>
</tr>
<tr>
<td>N_2O_5 \rightarrow NO_2 + NO_3</td>
<td>1.5 \times 10^{-11} e^{170/T}</td>
</tr>
<tr>
<td>HNO_4 + OH \rightarrow NO_2</td>
<td>K_0 = 2.0 \times 10^{-30} (300/T)^4.4</td>
</tr>
<tr>
<td>NO_2 + HO_2 \rightarrow HNO_4</td>
<td>K_w = 1.4 \times 10^{-12} (300/T)</td>
</tr>
<tr>
<td>HNO_4 (+ M) \rightarrow NO_2 + HO_2</td>
<td>2.7 \times 10^{-27} e^{1000/T}</td>
</tr>
<tr>
<td>NO_2 + CH_3O_2 \rightarrow NO_2 + CH_2O + XO_2 + CH_2O</td>
<td>1.3 \times 10^{-12} e^{380/T}</td>
</tr>
<tr>
<td>NO_2 + C_2O_3 \rightarrow PAN</td>
<td>K_0 = 2.0 \times 10^{-31} (300/T)^{0.3}</td>
</tr>
<tr>
<td>PAN \rightarrow NO_2 + C_2O_3</td>
<td>K_w = 2.9 \times 10^{-12} (300/T)^{1.1}</td>
</tr>
<tr>
<td>NO_3 + OLE \rightarrow 0.91 XO_2 + CH_2O + 0.09 XO_2N + NO_2</td>
<td>2.1 \times 10^{-27} e^{1000/T}</td>
</tr>
<tr>
<td>+ ALD2 + RXPAR</td>
<td>5.8 \times 10^{-16}</td>
</tr>
<tr>
<td>NO_3 + ISOP \rightarrow 0.9 HO_2 + 0.9 ORGNTR + 0.45 OLE</td>
<td>\frac{1}{2} (6.5 \times 10^{-15} + 1.4 \times 10^{-12} e^{-1880/T}) \textsuperscript{xiv}</td>
</tr>
<tr>
<td>+ 0.12 ALD2 + 0.08 CH_3COCHO + 0.1 NO_2 + 0.03 CH_2O</td>
<td>8.1 \times 10^{-12} e^{270/T}</td>
</tr>
<tr>
<td>NO + XO_2 \rightarrow NO_2</td>
<td>K_0 = 2.7 \times 10^{-28} (300/T)^{7.1} \textsuperscript{xiv}</td>
</tr>
<tr>
<td>OLE + NO \rightarrow NO_2 + XO_2 + CH_2O</td>
<td>K_w = 1.2 \times 10^{-11} (300/T)^{0.9} \textsuperscript{xiv}</td>
</tr>
<tr>
<td>ORGNTR + OH \rightarrow NO_2 + XO_2</td>
<td>K_0 = 4.9 \times 10^{-3} e^{-12100/T} \textsuperscript{xiv}</td>
</tr>
<tr>
<td>NO + NH_2 \rightarrow (sink)</td>
<td>K_w = 5.4 \times 10^{-18} e^{-13830/T} \textsuperscript{xiv}</td>
</tr>
<tr>
<td>NO_2 + NH_2 \rightarrow (sink)</td>
<td>\frac{1}{2} (4.1 \times 10^{-15} + 4.0 \times 10^{-14} e^{-400/T}) \textsuperscript{xiv}</td>
</tr>
<tr>
<td>NO_2 + HO_2 \rightarrow OLE</td>
<td>3.15 \times 10^{-12} e^{-450/T} \textsuperscript{xiv}</td>
</tr>
<tr>
<td>NO_2 + HO_2 \rightarrow ORGNTR</td>
<td>2.6 \times 10^{-12} e^{365/T} \textsuperscript{xxv}</td>
</tr>
<tr>
<td>ORGNTR + OH \rightarrow NO_2 + XO_2</td>
<td>2.6 \times 10^{-12} e^{365/T} \textsuperscript{xxv}</td>
</tr>
<tr>
<td>NO + NH_2 \rightarrow (sink)</td>
<td>5.9 \times 10^{-12} e^{-360/T} \textsuperscript{xxv}</td>
</tr>
<tr>
<td>NO_2 + NH_2 \rightarrow (sink)</td>
<td>4.0 \times 10^{-12} e^{450/T}</td>
</tr>
<tr>
<td>NO_2 + HO_2 \rightarrow ORGNTR</td>
<td>2.1 \times 10^{-12} e^{650/T}</td>
</tr>
</tbody>
</table>

\textsuperscript{xiv} Reaction rate based on Atkinson et al. (2006)
\textsuperscript{xv} Reaction rate based on Yarwood et al. (2005)
Table 8: Overview of photolysis reactions in TM5, reaction products are calculated explicitly by the chemical solver [Huijnen et al., 2010b]

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO$_2$ + hv $\rightarrow$ NO + O$_3$</td>
<td></td>
</tr>
<tr>
<td>HNO$_3$ + hv $\rightarrow$ NO$_2$ + OH</td>
<td></td>
</tr>
<tr>
<td>HNO$_4$ + hv $\rightarrow$ NO$_2$ + HO$_2$</td>
<td></td>
</tr>
<tr>
<td>N$_2$O$_5$ + hv $\rightarrow$ NO$_2$ + NO$_3$</td>
<td></td>
</tr>
<tr>
<td>NO$_3$ + hv $\rightarrow$ NO$_2$ + O$_3$</td>
<td></td>
</tr>
<tr>
<td>NO$_3$ + hv $\rightarrow$ NO</td>
<td>Temperature independent</td>
</tr>
<tr>
<td>PAN + hv $\rightarrow$ NO$_2$ + C$_2$O$_3$</td>
<td>Uses C4 N absorption cross section values</td>
</tr>
<tr>
<td>ORGNTR + hv $\rightarrow$ NO$_2$ + HO$_2$</td>
<td></td>
</tr>
</tbody>
</table>
B Temperature dependence of the reflectance for the O$_2$-O$_2$ cloud retrieval

This appendix gives an illustration of the influence of the atmospheric temperature profile on the depth of the O$_2$-O$_2$ absorption near 477 nm. The derivation is based on de Haan (2010). The vertical optical thickness of O$_2$-O$_2$ in the atmosphere ($\tau_{O_2-O_2, vert}$), a function of wavelength ($\lambda$, the dependence is usually weak), is given by Equation 37.

$$\tau_{O_2-O_2, vert}(\lambda) = \int_{z_{\text{bottom}}}^{z_{\text{TOA}}} \sigma_{O_2-O_2}(\lambda) n_{O_2}(z) n_{O_2}(z) \, dz$$  \hspace{1cm} (37)

Here, $z_{\text{TOA}}$ is the top of the atmosphere, $z_{\text{bottom}}$ is the surface height, $\sigma_{O_2-O_2}(\lambda)$ is the wavelength dependent absorption cross section of O$_2$-O$_2$, and $n_{O_2}(z)$ is the altitude ($z$) dependent oxygen number density. The Earth radiance measured by OMI ($I$) in case of 100% cloud cover at cloud height $z_c$ and the absence of other absorbers or scatterers, is given in Equations 38 and 39.

$$I(\lambda) = I_0(\lambda) \exp \left( - \tau_{O_2-O_2}(\lambda) \right)$$ \hspace{1cm} (38)

$$I(\lambda) = I_0(\lambda) \exp \left( - \int_{z_c}^{z_{\text{TOA}}} m(z, \lambda) n^2_{O_2}(z) \sigma_{O_2-O_2}(\lambda) \, dz \right)$$ \hspace{1cm} (39)

Here, $I_0$ is the (theoretical) Earth radiance uninfluenced by O$_2$-O$_2$ (the second term gives the absorption by the collision complex) and $m$ is the altitude resolved AMF which is a function of height and wavelength. The AMF gives the relation between the vertical atmosphere and the average light path of a measured photon. It converts the optical thickness for OMI to the vertical optical thickness (Eq. 37). The cloud effectively shields the absorption of O$_2$-O$_2$ underneath the cloud, making the depth of the O$_2$-O$_2$ signature in the reflectance a measure of the cloud height. We now introduce scale height $H$ (Eq. 40) which depends on the gas constant ($R_g$, 8.31 J K$^{-1}$ mole$^{-1}$), altitude dependent temperature ($T$), gravitational acceleration ($g$) and molecular mass of dry air ($M$, 28.96 $\times$ 10$^{-3}$). The altitude and latitude dependence of the gravitational acceleration are insignificant here.

$$H(z) = \frac{R_g T(z)}{g M}$$ \hspace{1cm} (40)

By assuming hydrostatic equilibrium we can relate $dp$ to $dz$ (Eq. 41) and use this relation to rewrite Equation 39 to 42.

$$\frac{dp}{p} = - \frac{dz}{H(z)}$$ \hspace{1cm} (41)

$$I(\lambda) = I_0(\lambda) \exp \left( - \int_{\ln(p_c)}^{\ln(p_{\text{TOA}})} H(p) m(p, \lambda) n^2_{O_2}(p) \sigma_{O_2-O_2}(\lambda) \, dp \right)$$ \hspace{1cm} (42)

As a next step, we express $n_{O_2}$ as a function of pressure ($p$), the volume mixing ratio of oxygen ([O$_2$] = 0.21), the Boltzmann’s constant $k_B$, and $T(p)$:

$$n_{O_2}(p) = \frac{[O_2] \, p}{k_B T(p)} = \frac{0.21 \, p}{k_B T(p)}$$ \hspace{1cm} (43)

Using the relations of Eq. 40 and Eq. 43 for the scale height and altitude dependent number density of oxygen in Equation 42, leads to Equation 44.

$$I(\lambda) = I_0(\lambda) \exp \left( - \int_{\ln(p_c)}^{\ln(p_{\text{TOA}})} \frac{R_g T(p)}{g M} m(p, \lambda) \left( \frac{0.21 \, p}{k_B T(p)} \right)^2 \sigma_{O_2-O_2}(\lambda) \, dp \right)$$ \hspace{1cm} (44)
Now, by simplifying the equation and the use of $\frac{d\ln(p)}{dp} = \frac{1}{p}$, one gets Equation 45:

$$I(\lambda) = I_0(\lambda) \exp \left( \frac{-0.21 \frac{R_g}{k_B M} \sigma_{O_2-O_2}(\lambda) \int_{p_{TOA}}^{p_{c}} \frac{m(p, \lambda) p}{T(p)} dp}{p} \right)$$

(Equation 45)

Equation 45 shows the $\frac{1}{T(p)}$ temperature dependence of the depth of the $O_2-O_2$ absorption and thus of the $O_2-O_2$ slant column. If instead of the collision complex $O_2-O_2$, the cloud product would have been based on $O_2$, Equation 42 would have been Equation 46. This equation can be rewritten to Equation 47, showing no temperature dependence for the $O_2$ cloud retrieval since the temperature dependence of $n_{O_2}$ cancels with the temperature dependence of $H(p)$.

$$I(\lambda) = I_0(\lambda) \exp \left( -\int_{\ln(p_{TOA})}^{\ln(p_{c})} H(p) m(p, \lambda) n_{O_2}(p) \sigma_{O_2}(\lambda) d\ln(p) \right)$$

(Equation 46)

$$I(\lambda) = I_0(\lambda) \exp \left( \frac{-0.21 R_g}{k_B M} \sigma_{O_2}(\lambda) \int_{p_{TOA}}^{p_{c}} \frac{m(p, \lambda) dp}{T(p)} \right)$$

(Equation 47)
References


