OMI tropospheric NO_2 profiles from cloud slicing: constraints on surface emissions, convective transport and lightning NO_x

M. Belmonte Rivas[1-2], P. Veefkind[1,2], H. Eskes[2], P. Levelt[1,2]

[1]{Technical University of Delft, Delft, The Netherlands}

[2]{Royal Netherlands Meteorology Institute, De Bilt, The Netherlands}

Correspondence to: M. Belmonte Rivas (belmonte@knmi.nl)

Abstract

We derive a global annual and seasonal climatology of tropospheric NO_2 profiles from OMI cloudy observations for the year 2006 using the cloud slicing method on six pressure levels centered about 280, 380, 500, 620, 720 and 820 hPa. A comparison between OMI and the TM4 model tropospheric NO_2 profiles reveals striking overall similarities, which confer great confidence to the cloud-slicing approach to provide details that pertain to annual as well as to seasonal means, along with localized discrepancies that seem to probe into particular model processes. Anomalies detected at the lowest levels can be traced to deficiencies in the model surface emission inventory, at mid tropospheric levels to convective transport and horizontal advective diffusion, and at the upper tropospheric levels to model lightning NO_x production and the placement of deeply transported NO_2 plumes such as from the Asian summer monsoon. The vertical information contained in the OMI cloud-sliced NO_2 profiles provides a global observational constraint that can be used to evaluate chemistry transport models (CTMs) and guide the development of key parameterization schemes.

1 Introduction

Global maps of tropospheric NO_2 vertical column densities (VCDs) derived from satellite UV/Vis nadir sounders such as OMI, GOME and SCIAMACHY have contributed to the development of a variety of applications. Clear sky observations of tropospheric NO_2 VCDs, those with cloud fractions typically below 25%, have been used to constrain surface NO_x

emission inventories (Martin et al., 2003) (Mijling and Van der A, 2012) (Miyazaki et al., 2012), detect and monitor point source emission trends (Richter et al., 2005) (Van der A et al., 2008) and constrain surface NO_2 lifetimes (Beirle et al., 2011) to cite a few examples. Still cloudy conditions predominate, which prevent the detection of NO_2 concentrations at the surface. For OMI, more than 70% of the measurements collected in the extratropics is affected by clouds and typically discarded, with the consequent loss of information. The utilization of cloudy data from satellite IR and UV/Vis nadir sounders provides access to a large repository of observations with potential to reveal information about trace gas concentrations at different altitudes and to constrain the parameterizations of a number of cloud related processes.

Clouds are introduced in general circulation models (GCMs) because of their broadband radiative effects and direct relation with the water vapour feedbacks and precipitation (Jakob, 2003). Clouds also affect the redistribution of trace gases via convection and interaction with chemistry, which are essential elements in chemistry transport models (CTMs). Convective transport of polluted plumes (including NO_x, but also HO_x, CO and non-methane hydrocarbons NMHC) from the boundary layer can cause substantial enhancement of upper tropospheric ozone, an important anthropogenic greenhouse gas (Pickering et al., 1992). At high altitudes, enhanced chemical lifetimes and stronger winds are also responsible for the long-range transport of pollutants. Still the exchange between environment and cloud air that determines the way that convective columns evolve (i.e. the entrainment and detrainment rates in mass flux schemes) remains uncertain. The presence of convective clouds not only transports pollutants vertically, it also removes soluble species (like HNO_3) by precipitation, and modulates photolysis rates by altering the actinic fluxes above and below the cloud (Tie et al., JGR, 2003). Associated with the deepest convective clouds, the production of NO_x by lightning is a key component of the NO_2 budget in the upper troposphere, not only because of its relation with O_3 production, but because it affects the general oxidizing capacity of the atmosphere and the lifetimes of tracers destroyed by reactions with OH - like CO, SO_2 and CH_4. Yet the source strength and spatial distribution of lightning NO_x emissions remain uncertain – with a global best estimate of 5±3 Tg a^{-1} (Schumann and Huntrieser, 2007).

In large scale global CTMs, convection and other cloud related processes such as scavenging and lightning NO_x production are represented by sub-grid parameterizations. Most convective parameterizations are tested against temperature and humidity profiles from radiosondes

(Folkins et al., 2006), but chemical tracers provide additional constraints. A number of studies have tried to quantify the effect of different convective schemes on tropospheric CO and O_3 profiles using satellite based climatologies for comparison with model data (Mahowald et al., 1995) (Barret et al., 2010) (Hoyle et al., 2011) finding the largest discrepancies in the tropical middle and upper troposphere. Even though NO_2 may appear unsuitable as a tracer of air motion because of its high reactivity with other NO_y members (such as N_2O_5, HNO_3, PAN, NO_3^- and HNO_4) and the presence of time-varying sources (mainly surface emissions and lightning NO_x, but also aircraft and stratospheric inflows), its short lifetime makes it attractive to study very fast transport mechanisms like convection. A number of studies have demonstrated the capabilities of satellite UV/Vis sounders to estimate the source strength and 3D distribution of lightning NO_x over cloudy scenes [(Boersma et al, 2005), (Beirle et al, 2006), (Martin et al. 2007) and (Miyazaki et al., 2014)]. These studies have found good agreement between modeled and observed lightning NO_2 over the tropical continents – albeit with discrepancies in the geographical and vertical distributions. Other studies have compared the performance of lightning parameterizations against satellite lightning flash densities, like (Tost et al., 2007) and (Murray et al., 2012), to conclude that it is difficult to find a good combination of convective and lightning scheme that accurately reproduces the observed lightning distributions - leaving the problem of the NO_x yield per flash aside. So there is a clear need for measurements with which the development of model parameterizations of convective transport and lightning NO_x schemes can be guided.

In this paper, we use a variation of the cloud slicing technique first developed by (Ziemke et al., 2001) for tropospheric ozone, and later exploited by (Liu et al., 2014) for tropospheric CO and (Choi et al., 2014) for tropospheric NO_2, based on the increments of gas vertical column density above cloud as a function of cloud pressure within a certain longitude/latitude/time cell. Obviously, large cloud fractions and some degree of cloud height diversity within the cell are conditions required for this technique to produce useful results. The cloud slicing approach applied by (Choi et al., 2014) on OMI NO_2 data was able to find signatures of uplifted anthropogenic and lightning NO_2 in their global free-tropospheric NO_2 concentrations, as well as in a number of tropospheric NO_2 profiles over selected regions. In this work, global annual NO_2 VMR profiles are generated at a spatial resolution of 2°x2° on pressure levels centered about 280, 380, 500, 620, 720 and 820 hPa. We give particular consideration to the scattering sensitivity of the OMI measurements above the cloud, as well as to the representativity of the cloud-sliced profiles with regard to a cloudy atmosphere. We

report on results from this methodology as well as its direct applicability as observational constraint using a state-of-the art chemical transport model.

2 Methodology

The methodology to produce observed and modeled annual climatologies of tropospheric NO_2 VMR profiles under cloudy scenes starts with a description of the OMI and TM4 datasets involved. We introduce the pre-processing steps required to estimate NO_2 VCDs above cloud from OMI slant column measurements, followed by the upscaling steps required to bring the spatial resolution of the satellite observations in line with the TM4 model grid for comparison.

OMI NO_2 products

The NO_2 slant columns used in this work are retrieved by the UV/Vis spectrometer OMI [Ozone Monitoring Instrument, (Levelt et al., 2006)] according to the KNMI DOMINO version 2.0 (Boersma et al., 2007) (Boersma et al., 2011). The data files, which include total and stratospheric slant columns, averaging kernel information, cloud fraction, cloud pressure and assimilated trace gas profiles from the TM4 model, are available at http://www.temis.nl/airpollution/no2.html.

Of particular importance to this study are the cloud pressures and fractions retrieved by the OMI O_2-O_2 cloud algorithm (Acarreta et al, 2004). The OMI O_2-O_2 cloud algorithm uses an optically thick lambertian cloud model with a fixed albedo of 0.8; the fraction of this lambertian cloud model covering the pixel is called effective cloud fraction [$c_{eff} = (R_{obs} - R_{clear})/(R_{cloudy} - R_{clear})$, where R_{cloudy} and R_{clear} are modeled clear and cloudy sky reflectances, and R_{obs} is the observed continuum reflectance – i.e. the reflectance with the O_2-O_2 absorption line removed], which is not the same as the geometric cloud fraction but an equivalent amount that yields the same top of atmosphere (TOA) reflectance as observations; the altitude level of the lambertian cloud model is then adjusted so that it results in the same amount of O_2-O_2 absorption as in observations [Stammes et al., 2008]. The OMI O_2-O_2 cloud pressure refers to the optical radiative cloud pressure near the midlevel of the cloud and below the MODIS infrared-based cloud top, which is about 250 hPa higher than OMI for deep convective clouds or about 50-70 hPa higher for extratropical midlevel clouds. The OMI O_2-O_2 cloud pressure has been validated against PARASOL with a mean difference below 50 hPa and a standard deviation below 100 hPa (Stammes et al., 2008). The OMI O_2-O_2 cloud fraction has been validated against MODIS with a mean difference of 0.01 and standard deviation of 0.12 over

cloudy scenes (effective cloud fractions larger than 50% without surface snow or ice) (Sneep et al., 2008). In this paper, we use the cloud radiance fraction defined as CRF = c_{eff} R_{cloudy}/R_{obs} – which represents the weight of the air mass factor of the cloudy part.

TM4 model

The TM4 chemistry transport model has a spatial resolution of 2°x3° with 35 sigma pressure levels up to 0.38 hPa (and approximately 15 levels in the troposphere) driven by temperature and winds from ECMWF reanalyses and assimilated OMI stratospheric NO_2 information from previous orbits. The tropospheric chemistry scheme is based on (Houweling et al., 1998) using the POET emissions (Olivier et al., 2003) database based on the EDGAR inventory for anthropogenic sources, which are typical of years 1990-1995, with biomass emissions of NO_x based on ATSR fire counts over 1997-2003 and released in the lowest model layers. The photolysis rates are calculated as in (Landgraf and Crutzen, 1998) and modified as in (Krol and van Weele, 1997). In the TM4 model, the physical parameterization for convective tracer transport is calculated with a mass flux scheme that accounts for shallow, mid-level and deep convection (Tiedtke, 1989). Large scale advection of tracers is performed by using the slopes scheme of (Russell and Lerner, 1981). The lightning NO_x production is parameterized according to (Meijer et al., 2001) using a linear relationship between lightning intensity and convective precipitation, with marine lightning 10 times less active than continental lightning and scaled to a total annual of 5 TgN/yr (Boersma et al., 2005). The vertical lightning NO_x profile for injection into the model is an approximation of the outflow profile suggested by (Pickering et al., 1998). Including free-tropospheric emissions from air-traffic and lightning, the total NO_x emissions for 1997 amount to 46 TgN/yr. More about this model may be found in (Boersma et al., 2011) and references therein.

2.1 Cloud slicing

A technique initially developed for estimating upper tropospheric ozone using nadir sounders (Ziemke et al., 2001), cloud slicing consists in arranging collections of trace gas VCDs measured above clouds against cloud pressure over a certain area and time period in order to estimate a gas volume mixing ratio (VMR) via the pressure derivative as:

$$VMR = 0.1 \cdot g \cdot M_{air}/N_A \cdot \frac{\partial VCD}{\partial p} \quad (1)$$

where g = 9.8 m/s^2, M_{air} = 28.97 g/mol and N_A=6.022x10^{23} molec/mol with VCD expressed in molec/cm^2 and cloud pressure expressed in hPa. The method determines an average trace gas

volume mixing ratio over a certain area, time period and cloud pressure interval (Choi et al., 2014). In this paper, annual average tropospheric NO_2 VCD lat/lon grids from OMI and TM4 are produced for six tropospheric layers with bottom cloud pressures located within pressure intervals centered at about 330, 450, 570, 670, 770 and 870 hPa. The cloud pressure intervals used for cloud slicing were chosen after several trial runs and are laid out in Table 1 and Fig. 1. An annual climatology of NO_2 VMR profiles is then estimated after differencing the annual tropospheric VCD arrays above cloud with respect to pressure.

Figure 1 shows the latitude-height section of annual zonal mean OMI cloud frequency for the year 2006, showing that cloud slicing does not provide uniform global sampling. Most high clouds (mainly deep cumulus, since cirrus pass generally undetected by OMI) occur along the intertropical convergence zone (ITCZ) near the equator and over tropical continents, but can also be seen in the mid-latitude storm track regions and over mid-latitude continents in the summer; mid-level clouds are prominent in the midlatitue storm tracks, usually guided by the tropospheric westerly jets, and some occur in the ITCZ; low clouds, including shallow cumulus and stratiform clouds, occur essentially over the oceans but are most prevalent over cooler subtropical oceans and in polar regions (Boucher et al, 2013). In summary, cloud sampling proves best at low to mid altitudes in the extratropics and mid to high altitudes in the deep tropics. On the contrary, cloud sampling is typically poor off the west coasts of subtropical (Pacific, Atlantic and Indian) landmasses at high altitudes - which are areas of large-scale subsidence with persistent low stratocumulus, and at low altitudes over the tropical landmasses, particularly the Amazon basin and Central Africa.

2.1.1 NO_2 above cloud

The tropospheric NO_2 vertical column density above the cloud VCD_{above} for an instrument like OMI is defined here as a function of the total slant column SCD as:

$$VCD_{above} = (SCD - SCD_{strat} - SCD_{below})/AMF_{above} \quad (2)$$

Where SCD_{strat} is the stratospheric slant column, SCD_{below} accounts for the slant surface component leaked from below the cloud (i.e. the amount of surface signal that seeps through the cloud for partially cloudy conditions), and AMF_{above} denotes the scattering sensitivity above the cloud. The stratospheric slant column arises from TM4 model stratospheric profiles assimilated to OMI observations over unpolluted areas (Belmonte Rivas et al., 2014). The undercloud leaked component is defined as:

$$SCD_{below} = (1 - CRF) \cdot \sum_{ground}^{CLP} m_{clear}(p) \cdot n(p) \cdot T_{corr}(p) \quad (3)$$

Where *CRF* is the cloud radiance fraction, m_{clear} is the clear sky component of the scattering sensitivity (purely dependent on Rayleigh scattering and surface albedo), *n(p)* is the a priori trace gas profile (i.e. the TM4 model), and T_{corr} is the OMI temperature correction defined below. Note that the summation goes from the ground to the cloud level pressure CLP (see Fig. 2), where the cloud level is given by the OMI O_2-O_2 cloud pressure. The scattering sensitivity above the cloud AMF_{above} is defined as (see appendix A):

$$AMF_{above} = \sum_{CLP}^{tropopause} m(p) \cdot n(p) \cdot T_{corr}(p) \,/\, \sum_{CLP}^{tropopause} n(p) \quad (4)$$

Where *m* is the total scattering sensitivity [usually defined as *(1-CRF)* m_{clear} + *CRF* m_{cloudy} as in (Boersma et al., 2004)]. Note that the summation in this case goes from cloud level to the tropopause (see Fig. 2). The total scattering sensitivity *m* has been derived from the averaging kernel *AK(p)* as:

$$m(p) = AK(p) \cdot AMF / T_{corr}(p) \quad (5)$$

Where *AMF* is the total airmass factor (used to compute the total vertical column VCD = SCD/AMF from the total slant column SCD, and different from the tropospheric airmass factor AMF_{trop} used to compute VCD_{trop} = SCD_{trop}/AMF_{trop}). The temperature correction is defined as in (Boersma et al., 2004) and accounts for the temperature dependence of the NO_2 absorption cross-section and its influence on the retrieved slant column using ECMWF temperatures:

$$T_{corr}(p) = (220 - 11.4)/[T(p) - 11.4] \quad (6)$$

The elements of the averaging kernel contain the height dependent sensitivity of the satellite observation to changes in tracer concentrations and they are calculated with a version of the Doubling Adding KNMI (DAK) radiative transfer model in combination with TM4 simulated tropospheric NO_2 profiles. Of central importance to our cloud slicing approach is that an undercloud leaked component (SCD_{below}) is removed from the tropospheric slant column, and a scattering sensitivity above the cloud (AMF_{above}) is used to estimate the vertical column density above the cloud VCD_{above}. This is in contrast with the methodology applied in (Choi et al., 2014), where undercloud leakages are neglected (making tropospheric estimates more sensitive to surface contamination, particularly at low cloud fractions), and the scattering sensitivity above the cloud assumed equal to the geometric airmass factor.

As far as model quantities are concerned, the NO_2 column above the cloud in TM4 is simply calculated as:

$$VCD_{above} = \sum_{CLP}^{tropopause} n(p) \quad (7)$$

Where *n(p)* is the a priori trace gas profile (i.e. the TM4 model). Note that the a priori gas profiles, originally reported on hybrid sigma pressure grids, have been resampled onto a uniform pressure grid with steps of 23.75 hPa to simplify averaging operations. The cloud level pressure that defines the model above-cloud NO_2 columns in Eq. (7) is the same OMI O_2-O_2 cloud pressure used for cloud slicing. Using OMI's cloud information to sample the TM4 model amounts to assuming that cloud altitudes and fractions in the model are identical to those observed by OMI. We know that differences between instantaneous model and observed cloud fields can be notable, but we also know that current model cloud fields are able to reproduce the average geographical and vertical distribution of observed cloud amounts reasonably well (Boersma et al., 2015), albeit with reports of underestimation of the low cloud fractions in the marine stratocumulus regions, underestimation of the midlevel cloud fractions everywhere, and slight overestimation of the high cloud fraction over the deep tropics (Nam et al., 2014) - errors that are likely related to the microphysical cloud and convection parameterizations. Therefore, using an observed cloud field to probe into model cloud processes, though probably suboptimal in case by case studies, is likely to be fine in an annual average sense.

2.1.2 Spatial averaging

A comparison of OMI observations with a model such as TM4 should also take into account the inhomogeneity of the tropospheric NO_2 field, which is usually large due to the presence of strong point sources and weather-scale variability. The model NO_2 columns should be viewed as areal averages, given that the limit of scales represented in the model is given by its resolution. Thus it is important to aggregate OMI observations to attain the same spatial resolution used by the model. The OMI NO_2 VCD above cloud observations (with a nominal spatial resolution of 13x24 km at the swath center) are aggregated onto daily 1°x1° longitude-latitude bins – later spatially smoothed to 2°x2°– before comparison with the afternoon TM4 model outputs defined on a 2°x3° grid on a daily basis as in Eq. (7). The aggregated OMI product collects all VCDs observed within a specified period (1 day) with solar zenith angle less than 70°, surface albedo less than 30% and CRF larger than 20% at the OMI pixel level

(roughly equivalent to an effective cloud fraction of 10%, which is a minimum condition for cloud fraction and pressure to be properly reported by OMI). No weighting is applied. At this point, populating the grid bins with as many OMI measurements as possible is important in order to avoid spatial representation errors between the two records (a partially filled bin may not be representative of what occurs over the entire cell, which is what the model represents). The aggregated CRF (and all other OMI and model quantities) are then evaluated at grid resolution, and a CRF threshold of 50% at cell level is applied to both observations and model data. The annual mean tropospheric VCD above cloud is then calculated per pressure layer using the CLP thresholds specified in Table 1 on daily gridded OMI and TM4 NO_2 VCD outputs, provided there are at least 30 measurements in a bin.

2.1.3 Error analysis

In the cloud slicing method, the derivation of annual mean VMR profiles from annual layered VCD amounts above cloud follows as:

$$\langle VMR_i \rangle = C \cdot (\langle VCD_{i+1} \rangle - \langle VCD_i \rangle)/(\langle p_{i+1} \rangle - \langle p_i \rangle) \quad (8)$$

where C is defined as $0.1 \times g \times M_{air}/N_A$ as in Eq. (1) and the index *i* refers to the cloud level. We term these objects VMR pseudoprofiles because they are constructed on the conditional provision of cloud presence, and the presence of cloud modifies the underlying NO_2 profile: either directly via chemical or dynamical processes such as lightning NO_x production, advection of (clean/polluted) air from below, suppression of biomass burning or decreased photolysis under the cloud, or more indirectly via selective sampling of seasonal features, such as entangling a wet season column of enhanced lightning at high altitude with a dry season column of enhanced biomass burning at low altitude. One can appreciate that the effect of cloud presence on the profile varies with cloud altitude, which is unfortunate, because we use changes in cloud altitude to sample the underlying profile. This state of affairs introduces a source of systematic error between the cloud-slicing estimate (i.e. the pseudoprofile) and the actual underlying profile, which we term pseudoprofile error. One may evaluate (and further compensate for) the pseudoprofile error associated to conditional cloud sampling by comparing the model VMR profile estimated using the cloud-slicing technique against the underlying "true" mean NO_2 VMR profile from the same model, as described below. Other sources of systematic error may also intervene, including uncertainties in the a priori corrections and errors in the stratospheric column. The effect of uncertainties in the a priori corrections is limited by the impact that a priori corrections have on pseudoprofiles,

which is itself limited (see Supplement). The effect of errors in the stratospheric column are expected to be small, since stratospheric columns only show a small additive bias (Belmonte Rivas et al., 2014) that is bound to cancel via the pressure difference. One could also include temporal representativity errors from mismatched collocations between model and OMI clouds in this category, which (Boersma et al., 2015) estimate to lie around 10%. In this section we provide a brief description of the retrieval error that may be expected from instrumental random noise properties alone, followed by an estimate of pseudoprofile error based on model behavior.

Retrieval error

The retrieval error in the annual mean cloud-slicing profiles is calculated by standard error propagation of Eq. (1). Note that we do not compute VMRs on daily or orbital basis (since one does not achieve the necessary cloud height diversity but in exceptional circumstances), but from the difference of annual mean VCDs. The derivation follows as:

$$\begin{aligned}\delta VMR &= 0.1 \cdot g \cdot \frac{M_{air}}{N_A} \cdot \delta\left(\frac{\partial VCD_{annual}}{\partial p_{annual}}\right) \\ &= 0.1 \cdot g \cdot \frac{M_{air}}{N_A} \cdot \left(\frac{\delta(VCD_1 - VCD_2)}{p_1 - p_2} + \frac{(VCD_1 - VCD_2)}{(p_1 - p_2)^2}\delta(p_1 - p_2)\right) \\ &= 0.1 \cdot g \cdot \frac{M_{air}}{N_A} \cdot \left(2 \cdot \frac{\delta VCD_{annual}}{p_1 - p_2} + \frac{(VCD_1 - VCD_2)}{(p_1 - p_2)^2} \cdot 2 \cdot \delta p_{annual}\right)\end{aligned}$$

Where VCD_1, VCD_2, p_1 and p_2 are all mean annual quantities estimated for contiguous pressure levels. Assuming random Gaussian errors in the determination of single OMI observations with an uncertainty δVCD of 50% in the OMI vertical column density (Boersma, 2004) and an uncertainty δp of 100 hPa in O_2-O_2 cloud pressure (Stammes et al. 2008), the standard error of the mean annual quantity (VCD or pressure) is the standard error of the single retrieval divided by the square root of the number of OMI measurements collected per grid cell N_{grid} in a year:

$$\delta VCD_{annual} = \delta VCD / \sqrt{N_{grid}}$$

$$\delta p_{annual} = \delta p / \sqrt{N_{grid}}$$

So that we obtain:

$$\delta VMR = 0.1 \cdot g \cdot \frac{M_{air}}{N_A} \cdot \left(2\frac{\delta VCD}{\Delta p} + 2\frac{\Delta VCD}{\Delta p} \cdot \frac{\delta p}{\Delta p}\right) \cdot \frac{1}{\sqrt{N_{grid}}} \quad (9)$$

Pseudoprofile (systematic) error

The extent to which cloud-slicing profiles remain physical and accurate representations of an average cloudy atmosphere is limited by the assumptions that underlie the cloud slicing difference, which goes as:

$$VMR(p_{mid}) \propto VCD(p < p_{dn} | p_{cloud} = p_{dn}) - VCD\left(p < p_{up} | p_{cloud} = p_{up}\right) \quad (10)$$

In cloud-slicing, the mean VMR between the pressure levels p_{up} and p_{dn} is given by the difference between the VCD above cloud pressure p_{dn}, provided there is cloud at p_{dn}, and the VCD above cloud pressure p_{up}, provided there is cloud at p_{up} too. The problem is that the presence of cloud modifies the profile. One may think that the column difference in Eq. (10) is an approximation to what happens when clouds are located at p_{mid}, somewhere between p_{up} and p_{dn}. But assuming that the trace gas concentration profile does not change with small changes in cloud altitude (which are otherwise necessary to estimate the VMR slope) entails some error. Ideally, we would like to calculate:

$$VMR_{true}(p_{mid}) \propto VCD(p < p_{dn} | p_{cloud} = p_{mid}) - VCD\left(p < p_{up} | p_{cloud} = p_{mid}\right) \quad (11)$$

Now we have a unique (and physically plausible) cloud condition behind the difference, p_{cloud}=p_{mid}, and a VMR estimate that is representative of gas concentration provided that there are clouds at the p_{mid} level. Yet if we would like to obtain a VMR estimate that is representative of trace gas concentration in a general cloudy atmosphere, then we would calculate:

$$VMR_{ref}(p_{mid}) \propto VCD(p < p_{dn} | \forall p_{cloud}) - VCD\left(p < p_{up} | \forall p_{cloud}\right) \quad (12)$$

That is, VMR_{ref} represents a mean VMR profile provided that there are clouds anywhere in the column, i.e. regardless of cloud altitude. We call the difference between VMR and VMR_{true} sampling error, because the cloud diversity necessary to estimate the trace gas concentration is distorting the underlying profile. We call the difference between VMR_{true} and VMR_{ref} representation error, because a profile measured under high cloud conditions is not representative of a profile under low cloud conditions, nor in general representative of an average cloudy state. The sum of the sampling and representation errors, that is the difference

between the cloud-sliced VMR pseudoprofile and the average profile in a cloudy atmosphere VMR_{ref}, is what we call pseudoprofile error. All VMR, VMR_{true} and VMR_{ref} profiles can be calculated on account of the TM4 CTM, so that a model based estimation of the sampling and representation (pseudoprofile) systematic error becomes available. The general pattern of pseudoprofile errors (see Sect. 3.3) indicates that biases are small in the upper three levels, largely positive (100-200%) over tropical and extratropical outflows in the lower two levels, and negative (up to 100%) over the continents for the lower three levels (particularly over central and South America, Australia, Canada and Siberia). One way to bypass this systematic error is to scale the observed VMR pseudoprofiles by the model profile-to-pseudoprofile ratio as:

$$VMR_{ref,OMI} = VMR_{OMI} \cdot (VMR_{ref,TM4}/VMR_{TM4}) \quad (13)$$

This model-based pseudoprofile correction (applied in Sect. 3.4) remains subject to the accuracy with which the model represents its own profiles, and should be treated with caution.

3 Results and discussion

3.1 NO_2 VCD above cloud

Figure 3a shows the annual mean tropospheric NO_2 VCD aggregates on 1°x1° grids observed by OMI for the year 2006 above clouds with mean pressures centered around 330, 450, 570, 670, 770 and 870 hPa – see Fig. 1 and Table 1. A similar set of annual mean NO_2 VCDs above cloud has been extracted from the TM4 model using identical cloud sampling (i.e. using the cloud fraction and cloud pressure from OMI) for comparison (see Fig. 3b).

Most of the lightning NO_2 emissions are expected above clouds higher than 450 hPa (i.e. the upper two levels in Fig. 3a) although some deep convection may also be present over strong industrial sources (like northeast US, Europe, China, and the Johannesburg area) or biomass burning sources in central Africa, the Amazon basin or northeast India, complicating the problem of process attribution.

The two middle levels in Fig. 3a are expected to carry, along with the NO_2 burden inherited from the upper levels, additional signatures from frontal uplifting into the mid-troposphere by conveyor belts over major industrial sources in northeast US, central Europe and China, as well as convective transport of biomass burning sources over central Africa, South America, Indonesia and northern Australia. The strong convective signatures of surface industrial and

biomass burning sources, along with their low tropospheric outflows, dominate the two lowest levels in Fig. 3a. Note the extensive lack of data over the tropical continents at low altitudes, a region where persistent high cloud precludes penetration into the lowest levels, and over the subtropical subsidence areas.

By differencing the annual average VCD arrays with respect to pressure, we expect to separate the contributions from different altitudes to the total tropospheric VCD column. But before that, let us take a look at the scattering sensitivities above cloud and the effects of correcting for undercloud leakage in these results. Fig. 4 shows the annual mean tropospheric scattering sensitivity above cloud level [AMF_{above} in Eq. (4)] applied to generate the OMI NO_2 VCDs shown in Fig. 3a. Globally, the tropospheric scattering sensitivity above the cloud does not deviate more than a 10% from the geometric airmass factor at most cloud altitudes, except at the lowest levels, where it suffers reductions of up to 30%. This reduction in scattering sensitivity at the lowest cloud levels may come as a surprise, particularly when clouds are known to boost the scattering sensitivity just above the cloud top. However, the pronounced decrease in scattering sensitivity at the lowest cloud levels is related to penetration of substantial amounts of NO_2 (from strong or elevated surface sources) into the cloud mid-level, where extinction acts to reduce the scattering sensitivity. Other than the extinction effect, the variability in scattering sensitivity is governed by changes in the observation geometry (AMF_{above} decreases as the sun angle increases) and the temperature correction introduced in Eq. (6), which is responsible for the subtropical bands and the variability at high southern latitudes.

The corrections for the surface leaked component introduced in Eq. (3) are largest (see Supplement) over polluted regions for the highest clouds (up to 50-66%) and smallest over clean areas like the oceans. In order to verify that the model-based undercloud leak corrections do not appreciably change the OMI NO_2 VCDs arrays, we have performed a separate trial run where the CRF threshold (at grid level) is increased from 50% to 80% (see Supplement) to conclude that none of the prominent VCD signatures seen in Fig. 3a (or none of the VMR features that we will see later) changes appreciably in the restricted CRF>80% case. Results from the CRF>80% trial run include notably diminished cloud frequencies and spatial coverage, seriously thinning the population that produces the annual averages and generally damaging their representativity. This effect is particularly notable in the upper two levels (280 and 380 hPa) and to lesser extent over the large-scale subsidence area in the

lowest level, since deep convective and low marine stratocumulus clouds are not particularly extensive but have a preference for low effective cloud fractions. Excluding the contributions from these cloud types in the CRF>80% case does not change the mid-tropospheric NO_2 patterns relative to the CRF>50% case, but it is biasing the OMI aggregates in the upper troposphere low relative to the modeled average, which is not particularly sensitive to this change. In summary, the CRF>80% trial run does not show any clear signs of a priori information constraining the results, but it provides hints of results being influenced detrimentally by the lower sampling densities afforded by a higher CRF threshold.

3.2 NO_2 VMR pseudoprofiles

The annual mean tropospheric NO_2 VMR pseudoprofiles observed by OMI for the year 2006 are compared against their TM4 model counterparts in Figs. 5a-c. Note that pseudoprofile errors do not affect this comparison, since both observed and modeled pseudoprofiles are observing identical (if somewhat unphysical, because of sampling and representation issues) atmospheric states. After the pressure difference, there remain some instances where negative VMRs are found, but these are mainly associated to poorly populated cells (such as at high latitudes, near the tropics at low altitudes, or around subsidence regions). These instances are identified and dealt with by recourse to information from nearby cells, when available, or otherwise ignored.

Many of the cloud slicing features observed at the upper two levels (280 and 380 hPa) in Fig. 5a can be attributed to actual biomass burning, lightning and deep convection. It may be difficult to separate these components clearly withouth a proper seasonal analysis (deferred to Section 3.6), although one can identify areas of predominant lightning production as those regions that do not seem connected via convection to surface sources underneath and use the OTD-LIS flash rate climatology and the ATSR fire counts (see Fig. 6 below) as interpretation aids for attribution. Positive anomalies (observations larger than modeled amounts) are detected in Fig. 5a over all major industrial areas (eastern US, central Europe and eastern China) both at 280 and 380 hPa levels, suggesting that deep transport of boundary layer NO_2 may be too weak in the model. On the contrary, there are extensive negative anomalies (meaning observations lower than modeled amounts) in background upper tropospheric NO_2 both at 280 and 380 hPa, which is consistent with reports of model overestimation of the amount of NO_2 attributed to lightning over the tropical oceans in (Boersma, 2005).

Negative anomalies in Fig. 5a are particularly large over Siberia, Amazonia and the Bengal Bay. The negative anomaly over eastern Siberia, an area of predominant biomass burning, could be related to excessive fire-induced NO_2 emission over boreal forests in the model (Huijnen et al., 2012). In South America, lightning NO_2 contributions seen by OMI appear confined mostly to the western equatorial coast (Peru, Ecuador and Colombia) on one side, and southern Brasil and off the east coast of Uruguay on the other hand (more in line with the OTD-LIS flash climatology shown in Fig. 6) - in stark contrast with model amounts, which locate the lightning maximum further to the north over the brasilian Matto Grosso, where the maxima in precipitation related to the South American monsoon system usually takes place. It is worth noting that the lightning intensity in the TM4 model is solely driven by convective precipitation, although [Albrecht et al, 2011] report that convective precipitation is not always well correlated with lightning in this area, showing that the most efficient storms in producing lightning per rainfall are located in the south regions of Brazil. The negative anomaly over Amazonia is therefore very likely related to problems with the TM4 lightning scheme. The negative anomaly over the Bengal Bay, an area of maxima in precipitation related to the Indian monsoon, could also be a reflection of excess model lightning linked to convection.

Other notable discrepancies in Fig. 5a include positive anomalies over central Africa and northeast India at 280 hpa. Over central Africa, the pattern of positive anomalies bears only partial resemblance with the pattern of biomass burning emission underneath (see midlevel OMI VMRs in Fig. 5b) – suggesting that upper level positive anomalies in central Africa may be related more to deficiencies in the lighting scheme than to convective transport. Actually, (Barret et al., 2010) report that lightning flash frequencies simulated by TM4 are lower than measured by the LIS climatology over the Southern Sahel, which is consistent with our observations. On the other hand, the large positive anomaly observed over the Tibetan plateau at 280 hPa, which significantly deviates from the OTD-LIS flash rate climatology in the area (confined to the Himalayan foothills only), is likely an effect of deep transport associated with the Asian monsoon. The model does show an enhacement in upper tropospheric NO_2 over India, but not moving far enough north into the Tibetan plateau and failing to reproduce the strong enhancements in upper tropospheric NO_2 over northeast India and Southern China related to the Asian summer monsoon plume – which (Kar et al., 2004) also detected in the MOPITT CO profiles.

The cloud slicing features observed at the mid-tropospheric levels (500 and 620 hPa) in Fig. 5b may be mostly attributed to mid-tropospheric convection of strong surface sources and their associated outflows. We observe a remarkable agreement between model and observations on the localization and intensity of major convective signals over industrial sources (eastern US, central Europe, China and India) as well as over typical biomass burning sources in central Africa, Indonesia and South America. Contrary to what is observed in the upper levels (see prevalent negative anomalies in Fig. 5a), there are extensive positive anomalies (meaning observations larger than modeled amounts) in background middle tropospheric NO_2 both at 500 and 620 hPa in Fig. 5b, particularly over the tropics and subtropics – which is indicative of deficient model mid-tropospheric outflows at these levels. Positive anomalies over the continents are particularly large over China (with an outflow related positive anomaly downwind over the Pacific), central US, and the biomass burning regions in central Africa and South America. While it may be more or less clear that enhanced mid-tropospheric NO_2 concentrations observed over the oceans are related to enhanced convective inflows into this level (without definitely discarding a problem with NO_2 lifetime), the origin of the convective anomalies remains ambiguous. A cursory look at the NO_2 concentrations observed at lower levels might help discriminate whether flux anomalies into the mid-troposphere are related to deficiencies in model prescribed surface emissions or problems with the convective transport scheme, or both.

For example, the pattern of anomalies over China at lowest levels (see Fig. 5c) is prominently positive, but it carries a dipolar positive (China) - negative (Japan) pattern that is no longer observed at higher levels. So, while it is possible that some of the mid-tropospheric convective anomalies are a response to flux anomalies carried from underneath (i.e. a deficiency in the originally prescribed surface emission), as it happens over eastern US and Europe, where negative anomalies are carried upwards (see Fig. 5b), the overall effect does not exclude net deficiencies in model convective transport. As far a biomass burning is concerned, the pattern of anomalies over central Africa and South America in the lowest tropospheric levels (see Fig. 5c) is unfortunately not as evident (given the lack of low cloud detections) as over China but mostly neutral or slightly negative, indicating that mid-tropospheric positive anomalies in this area respond to either a convective transport scheme that is too weak or a model injection height that is too low.

The lower tropospheric levels (720 and 820 hPa) in NO_2 sampled by the cloud slicing technique are shown in Fig. 5c. These levels sustain the highest NO_2 concentrations in the vicinity of major industrial hubs (eastern US, central Europe and China) and the strongest anomalies as well, which in this case can be linked directly to deficiencies in prescribed surface emissions. All major features in the anomaly patterns at these levels can be matched unambiguously to the pattern of OMI to TM4 total tropospheric NO_2 column differences for clear sky-conditions shown later in Fig. 12, characterized by positive anomalies over northeast US, central Europe and Japan, and negative anomalies over China. These low level signatures are consistent with NO_2 increases over China, India and the Middle East, and NO_2 decreases over eastern US and central Europe, which are not reflected in the model emission inventory. Other salient features at these levels include an interesting band of negative anomalies along the ITCZ (perhaps related to rapid convective mixing of relative "clean" air from the boundary layer) and extensive positive anomalies over the oceans (more so at 720 than at 820 hPa) – revealing deficient model outflows at high latitudes and suggesting that poleward transport of NO_2 in the model may not be vigorous enough (a problem likely related with horizontal diffusion in the model).

In summary, there is remarkable agreement between observed and modeled upper/middle/lower tropospheric NO_2 amounts, their main distributions resembling each other at continental scale, with localized differences suggesting that the cloud slicing technique holds promise for testing model features related to anthropogenic emission, convection and uplift, horizontal advection and lightning NO_x production.

3.3 Classification

In the previous section, we studied the geographical distribution of observed and modeled NO_2 amounts on different pressure layers. In this section, we focus on the vertical dimension by looking at NO_2 VMR amounts across pressure layers. In order to simplify the analysis, we have defined a set of geographical classes based on the amount of variance contained in the TM4 model NO_2 profiles. These classes characterize how much of the NO_2 content in the profile can be apportioned to surface sources and how much to outflows – further subdivided into outflows with low, mid or high altitude components. Annual mean NO_2 VMR profiles are plotted for each class, along with reference to pseudoprofile error. A standard empirical orthogonal function (EOF) decomposition of the reference TM4 profiles [VMR_{ref} in Eq. (12)] is employed to characterize the geographical variance of NO_2 vertical profiles under cloudy

conditions and identify major spatial patterns. The first four EOF eigenvectors (out of a total of six) are shown in Fig. 7a. The first EOF represents profiles with higher concentrations near the surface – a profile over a surface source. The second EOF represents profiles with concentrations uniformly distributed across the column – a profile for a generic outflow type. The third and fourth EOF eigenvectors divide the generic outflow type into subtypes with stronger high altitude (EOF3>0), low altitude (EOF3<0) or mid-tropospheric (EOF4>0) components. The classes that result from applying masks based on the conditions defined in Table 2 are shown in Fig. 7b. According to the TM4 model, the classes containing all primary and secondary industrial sources (i.e. strong projections on EOF1) are mainly confined to the US, Europe and China. Other secondary industrial sources relate to India, the Middle East and the Baykal Highway (a major road connecting Moskow to Irkutsk, passing through Chelyabinsk, Omsk and Novosibirsk). Major biomass burning sources include large sectors in Africa and South America, Indonesia, New Guinea, and northern Australia. NO_2 outflows over the tropics (i.e. strong projections on EOF2) are subdivided into generic tropical outflows (with strong upper and mid-tropospheric components, or larger projections on EOF3 and EOF4) and outflows over large-scale subsidence areas (with stronger lower tropospheric components, or negative projections on EOF3). The extratropical outflows differ from the tropical outflows in that the sign of the mid-troposheric projection is reversed, so that extratropical profiles are more C-shaped (according to the model). The boreal outflow differs from the extratropical outflow in that it has an extremely large upper tropospheric component (i.e. a very large projection on EOF3). Finally, we have defined a separate class, labeled clean background, including all those areas without significant projections on either source or outflow eigenvectors.

The average tropospheric NO_2 profiles estimated using the cloud slicing method on OMI and TM4 datasets for all the 11 classes (15 classes when primary and secondary industrial regions are subdivided geographically into China, USA and Europa subclasses) defined in Table 2 and Fig. 7b are shown next in Figs. 8 and 9. These plots compare the OMI and TM4 VMR pseudoprofile estimates calculated in a cloud slicing fashion as in Eq. (10), along with the reference TM4 VMR_{ref} profile calculated as in Eq. (12) for an average cloudy atmosphere. Recall that the difference between the TM4 VMR and VMR_{ref} profiles gives an indication of pseudoprofile error – or the representativity of the cloud-slicing estimate relative to a general cloudy situation. The OMI VMR cloud slicing estimate is bounded by error bars calculated

from standard error propagation as in Eq. (9), and scaling by the square root of the number of profiles collected per grid cell – also shown in right subpanels in Figs. 8 and 9.

The cloud-slicing estimate for the annual tropospheric NO_2 profiles over primary industrial centers in eastern US, Europe and China are shown in the first row in Fig. 8. There is a remarkably good correspondence between observed and modeled tropospheric NO_2 profiles over these strongly emitting areas, particularly over central Europe, attesting to the accuracy and representativity of the cloud-slicing estimates for yearly means. Pseudoprofile errors are small in these areas, so that cloud-slicing estimates remain a good representation of average cloudy conditions. The OMI to TM4 VMR differences at the lowest levels are consistent with known deficiencies in model prescribed surface emissions (OMI smaller than the TM4 over eastern US and central Europe, but larger over China). These low level anomalies are carried upwards to a level of 500-600 hPa, above which the effects of enhanced convective mid-tropospheric and deep transport start to dominate regardless of the signature of the surface difference. The second row in Fig. 8 show the annual tropospheric NO_2 profiles over secondary industrial centers around eastern US, Europe and China. The low level features related to surface emission are identical to those of the primary centers, but the signature of enhanced mid-tropospheric convection is clearer - indicating that vertical transport in the model is too weak or lifetime too short, regardless of the sign of the surface anomaly. The sign of the OMI to TM4 difference is reversed in the upper two levels, in line with the generalized model overestimation of NO_2 in the upper troposphere. The third row in Fig. 8 shows the cloud-slicing estimate for the annual tropospheric NO_2 profiles over secondary industrial pollution centers in India, the Middle East and the Baykal Highway - note that pseudoprofile errors are larger in this case. For India, the differences between OMI and TM4 profiles at low levels point at a large underestimation of model surface emissions, and model overestimation of upper tropospheric NO_2 amounts – this upper level anomaly related to the misplaced Asian summer monsoon signal, which in observations appears located over the Tibetan plateau. For the Middle East, the difference between OMI and TM4 profiles points at large differences at mid-tropospheric level (OMI larger than TM4). The agreement between OMI and TM4 profiles for the Baykal Highway class is reasonably good – allowing for a small underestimation of model surface emissions. After deep transport in China, this is the class with higher upper level NO_2 amounts, most likely related to fire-induced convection from boreal fires. The left panel in the fourth row in Fig. 8 shows the cloud slicing estimate for the annual tropospheric NO_2 profile over tropical biomass burning regions, featuring

positive anomalies at middle levels and negative anomalies at lower and upper levels, again pointing at defective model convective transport into the mid-troposphere (or issues with the pyro-convection heigth). The cloud-slicing estimates for annual tropospheric NO_2 profiles over typical outflow regions are shown in the middle and right panels in the fourth row (tropical and tropical subsidence outflows) and left and middle panels in the fifth row (extratropical and boreal outflows) in Fig. 8. As a salient feature, all of the outflow profiles share a prominent mid-tropospheric plume centered around 620 hPa in the tropics and a little lower in the extratropics, around 720 hPa, with NO_2 amounts much smaller than the model in the upper troposphere and general agreement at the lowest level, producing profiles which are generally S-shaped (instead of C-shaped as in the model). The mid-tropospheric plume is likely related to enhanced convective fluxes of NO_2 over industrial and biomass burning areas (but definitely not discarding issues with NO_2 lifetime or substantial chemical NO_x recycling from HNO_3 and PAN sources at this level). Note also the generalized model overestimation of NO_2 in the upper levels (tropical and extratropical), which is consistent with reports of excess lightning NO_x production over the tropical oceans in (Boersma et al., 2005). The upper level overestimation is particularly large for the boreal outflow class, which we also mentioned could be related to the excess fire-induced convection over Siberia or too large NO_x emission factors. Finally, the cloud-slicing estimate for the annual tropospheric NO_2 profile over the clean Southern Ocean is shown on the right panel of the last row in Fig. 8, with good agreement at the top levels and gradually increasing model underestimation towards the surface, suggesting enhanced lateral contributions at high latitudes from horizontal eddy diffusion.

The left panel in Fig. 9 shows the annual tropospheric NO_2 profile for all the primary surface sources together (eastern US, central Europe and China), indicating that differences at surface level average out globally, leaving the effects of enhanced observed mid-tropospheric convection and deep transport to stand out. The signature of enhanced mid-tropospheric convection becomes even clearer in the mid panel in Fig. 9, which shows the annual tropospheric NO_2 profile for all secondary surface sources together (around primary sources, plus India, the Middle East, the Baykal Highway and the biomass burning areas), where the signature of enhanced deep transport is in this case replaced by model overestimation of upper tropospheric NO_2. The model overestimation of upper level NO_2 appears clearly on the right panel in Fig. 9, which shows the annual tropospheric NO_2 profile for all the outflow classes, along with a prominent model underestimation of mid-tropospheric NO_2 levels. In summary,

and consistent with our comments on Figs. 5a-c, the average profiles that result from applying the cloud slicing technique on observed OMI and modeled TM4 datasets show striking overall similarities, which confer great confidence to the cloud-slicing approach, along with more localized differences that probe into particular model processes and parameterization schemes.

3.4 Cross-sections

We would like to wrap up our results in the form of observed and modeled annual zonal mean and longitudinal NO_2 cross-sections along the tropics (Figs. 10a-b and 11). Note that in order to bypass pseudoprofile errors, the observed NO_2 pseudoprofiles are scaled in this section by the model profile-to-pseudoprofile ratio as in Eq. (13) to form what is called the observation update.

For the annual zonal mean tropospheric NO_2, the left-to-right panel comparison in Fig. 10a shows that although the observation update does not change the strength of major industrial emission over the northern midlatitudes at the lowest levels, the associated convective cloud is reaching higher in altitude. In the tropics and southern latitudes, vertical transport of the combination of biomass burning and industrial emissions is stronger and reaching higher - with a prominent high plume originating from the Johannesburg area. The observation update does bring notably stronger midtropospheric outflows distributed over a broader latitude band and weaker NO_2 signatures at high altitude. The enhanced midtropospheric plume is best appreciated in Fig. 10b, which shows the annual zonal mean tropospheric NO_2 averaged over the Pacific Ocean sector (180W-135W) – the dominant sources of NO_2 over the oceans are thought to include the long-range transport from continental source regions, as well as chemical recycling of HNO_3 and PAN sources [Staudt et al., 2003]. [Schultz et al., 1999] actually shows that the decomposition of PAN originating from biomass burning actually accounts for most of the midtropospheric NO_x in the remote Southern Pacific, suggesting that enhanced convective flux from surface sources may not be the only agent responsible for the enhanced midtropospheric outflows observed by OMI.

Figure 11 shows a picture for the annual longitudinal NO_2 cross-section for tropical latitudes between 10N and 20S, where the observation update raises the convective plumes from major biomass burning areas in South America, central Africa and Indonesia/Northern Australia to higher altitude, between 500 and 600 hPa, with a slight westward tilt and longer downstream

transport of cloud outflow at upper levels caused by the tropical easterly jet, and generally weaker NO_2 signatures at high altitude.

In summary, the OMI cloud-slicing NO_2 profiles seem to suggest that raising the polluted plumes to higher altitudes allows for much longer residence and chemical lifetimes, and longer and more widely distributed horizontal transport of NO_2 (following poleward advection and dispersion by the subtropical jet and by baroclinic waves at lower levels) in the mid-tropophere. These observations are in line with reports in (Williams et al, 2010) showing that the underestimation of upper tropospheric O_3 in TM4 relative to observations over Africa may be linked to a too weak convective uplift using the Tiedtke scheme. The studies of (Tost et al., 2007), (Barret et al., 2010) and (Hoyle et al., 2011) corroborate this finding, indicating that the vertical extent of tropical convection and associated transport of CO and O_3 in the middle and upper troposphere is underestimated in Tiedtke based models. Accurately constraining the convective transport in CTMs should contribute to the determination of the vertical distribution of lighting NO_x, since knowledge of the extent of mixing of air into the cloud as a function of altitude is required to separate the NO_x produced by lightning from that produced by upward transport (Dickerson, 1984).

3.5 Consistency check

Because of their annual and global character, we do not have any direct means to validate the OMI annual tropospheric NO_2 profile climatology estimates in the same way that it has been done, for example, in [Choi et al., 2014]. But we can check their consistency by demanding that the total tropospheric NO_2 column from the cloud-slicing technique does not deviate significantly from the total tropospheric NO_2 column observed in clear sky conditions (see Fig. 12). The total tropospheric NO_2 column $VCD_{slicing}$ from the cloud slicing technique is calculated as the sum of partial vertical column densities obtained from the annual mean pseudoprofile VMR as:

$$VCD_{slicing}(lat, lon) = \sum_{i=1,\dots,6} VMR_i(lat, lon) \cdot (\langle p_{i+1} \rangle - \langle p_i \rangle) / C \quad [14]$$

Where C is the same constant defined in Eq.(8). Note that absent VMR grid values (such as at high altitude over subsidence regions, or at low altitude over the tropical continents) are ignored without provision of a priori information.

We do know that there are some basic differences between NO_2 profiles observed under clear and cloudy conditions though. In the TM4 model, the differences between cloudy (CRF>50%) and clear (CRF<25%) profile climatologies (see left panel in Fig. 13 below), show strong negative anomalies over the biomass burning areas (central Africa, southern America, northern Australia, southern India, but also in the Persian Gulf and Turkestan) most likely related to fire suppression during the wet/cloudy season. Over industrial areas (USA, Europe and China) a more complex pattern of anomalies arises that likely results from the competing effects of suppressed photolysis under clouds (small positive anomaly), venting by passing fronts (large negative anomalies) and accumulation patterns dependent on a predominant synoptic weather type [cyclonic or anticyclonic, (Pope et al., 2014)]. This pattern of differences between cloudy and clear annual NO_2 profile climatologies is well reproduced by OMI observations (see right panel in Fig. 13 below). The sole difference is that OMI sees larger outflows at higher latitudes in the cloudy case – perhaps a deficiency of the model in redistributing its horizontal flows under frontal conditions.

Another more direct way to perform this consistency check is to look at the differences in total NO_2 columns between model (TM4) and observations (OMI) for the clear and cloudy cases separately, as shown in Fig. 14. For the clear sky case (see left panel in Fig. 14) the pattern of anomalies that arises is consistent with existing long-term satellite NO_2 trend studies [(van der A. et al., 2008) and (Richter et al., 2005)] that report significant reductions in NO_2 in Europe and eastern parts of the United States, strong increases in China, along with evidence of decreasing NO_2 in Japan, increasing NO_2 in India, Middle East, and middle Russia - and some spots in central United States and South Africa. The differences between model and clear-sky OMI NO_2 total columns are being used to update the surface emission inventories (Mijling & van der A, 2012) (Ding et al., 2015). What is comforting is that a similar pattern of differences arises in the cloudy case (using the cloud-slicing TM4 and OMI profiles), and with a similar amplitude, verifying that the OMI cloud slicing columns are internally consistent with the clear sky OMI observations in detecting anomalies that can be ultimately related to outdated model emission inventories.

In Figure 14, note that the model total tropospheric NO_2 columns over clean remote areas (i.e. tropical and extratropical outflow regions over the oceans) in the cloudy case do not deviate in general by more than 0.1xE-15 molec/cm2 from observations This is a good result, showing that the estimate of the stratospheric column (by data assimilation) does not produce

significant cloud-cover dependent biases in the tropospheric column. If we recall that the observed cloud-slicing NO_2 profile over clean remote areas is S-shaped, with a much stronger mid-tropospheric component and a much reduced upper tropospheric load than in the model, then we can infer that there has been as much gain in the mid-tropospheric component as there has been loss at high altitude, which is another form of closure.

3.6 Seasonal analysis

The seasonal mean tropospheric NO_2 VMR pseudoprofiles for the DJF, MAM, JJA and SON periods observed by OMI over the year 2006 compared against their TM4 model counterparts are shown next. These plots have been generated using the same cloud slicing gridding and threshold configurations applied for the annual means, with a required minimum of 7 measurements collected per bin during each season (instead of 30 for the annual means). This section is not intended to provide a thorough analysis of seasonal variability in (observed or modeled) tropospheric NO2 profiles, but to reveal the potential of the cloud slicing technique to provide details that pertain to seasonal as well as to annual means.

The largest signatures of seasonal variability expected to arise in these figures will be a) a seasonal cycle in lightning activity in the upper levels (280-380 hPa); b) a seasonal cycle of biomass burning in the mid levels (500-620 hPa); and c) a seasonal cycle over industrial areas at lower levels (720-820 hPa), featuring minimum NO_2 levels in the summer months due to changes in the lifetime of NO_x [van der A. et al., 2008]. The seasonal cycle in lightning NO_x emissions may be verified against the climatology of lightning flashes observed by LIS/OTD (data set available online [ftp://ghrc.nsstc.nasa.gov/pub/lis/climatology], see [Cecil et al., 2014]). The seasonal cycle in biomass burning may be verified against the climatology of ATSR and AVHRR fire counts [(Dwyer et al., 2000) and (Schultz, 2008)].

Africa

Over Africa, persistent lightning activity at upper levels is expected to take place about the Equator (the Congo Basin) all year long, shifting southward towards South Africa in SON and DJF, and northward towards the Gulf of Guinea, the Sahel and Sudan in MAM and JJA, features which are all captured by OMI in Fig. 15 (in reasonable agreement with TM4, though some discrepancies are apparent too). These lightning signatures are not to be confused with traces of NO_2 lifted from biomass burning underneath, which feature a precisely opposite phase. Remarkable biomass burning signatures can be appreciated throughout the entire

tropospheric column in Figures 15-18 as NO_2 enhancements north of Equator (Sahel) in DJF, and south of Equator (Angola and Zambia) in JJA, shifting eastward towards Mozambique and Madagascar in SON (best seen at 500 and 620 hPa in Figs. 17-18). We note that the penetration of seasonal biomass burning signatures into 280-380-500 hPa is stronger in OMI than in TM4. In addition, note the strong enhancement in lightning activity seen by OMI off the Southeast coast of Africa in MAM and JJA at 380 hPa in Fig.16, in connection with the confluence of the warm Agulhas and the cold Antarctic Circumpolar Current, which is virtually missed by TM4.

South America

Over South America, a maximum in lightning activity is expected to occur over central Brazil in SON, as captured by OMI in Figs. 19-20 (in agreement with TM4, though some discrepancies persist relative to the location of the lightning maximum, as we mentioned when describing the annual means), migrating towards the southeast in DJF. Lightning and precipitation are persistent in the northwest (Colombia, Venezuela and Central America) all year-round, intensifying in MAM and JJA, as reasonably captured by OMI in Fig. 15, along with some persistent NO_2 enhancements over la Plata Basin and off into the Brazil-Malvinas Confluence Zone. The lightning signatures at upper levels may be partly overlapped by those from biomass burning lifted from underneath, but their separation is more difficult in this case. For instance, the NO_2 enhacements detected by OMI at 500 hPa over Brazil in SON and DJF in Fig. 21 correlate well with the lightning signatures at 380 hPa, but they also correlate with the biomass burning signals at 620 hPa, indicating that both processes may be occurring at the same time in separate but nearby locations (e.g. combining the start of the wet season in the Amazon Basin in DJF, with the end of the burning season in eastern Brazil). The cycle of biomass burning in South America, which takes place over the dry season, starts in southern Brazil in JJA, to find a maximum in SON eastward towards the coastal states, as OMI captures in Figs. 21-23 (in reasonable agreement with TM4). In DJF, some activity may persist in eastern Brazil and new activity develop over the lower slopes of the Argentinian Andes during the austral summer, generally complicating attribution. Finally, it is interesting to note in Fig. 23 the remarkable decrease in NO_2 levels at 720 hPa over the Amazon Basin during the rainy season in DJF and MAM, as if in connection with an efficient NO_2 removal mechanism.

Southeast Asia and Australia

At upper levels, one should expect to see persistent lightning activity over Indonesia all year round, as qualitatively observed by OMI in Fig. 24 (and in agreement with TM4), migrating northward towards South Asia in MAM and JJA, and southward towards Australia in SON and DJF. These lightning signatures may be mixed to greater or lesser degree with NO_2 lifted from biomass burning and/or anthropogenic sources underneath. Over South Asia, biomass burning is expected to find its maximum in MAM, as OMI captures in Figs. 26-28 particularly over north India and Myanmar. These emissions are likely responsible for a large part of the NO_2 enhancement observed around India at upper levels in MAM. The very strong NO_2 enhancement observed over South Asia at upper levels in JJA (Figs. 24-25) is likely related to deep transport of surface emissions (biomass and industrial) during the monsoon season, which TM4 locates over the Indo-Gangetic area and OMI locates over the Tibetan Plateau. Over Indonesia and Northern Australia, a maximum in biomass burning activity is expected be reached in SON, as OMI captures reasonably well in Figs. 26-28, indicating that the strong NO_2 enhancement seen by OMI over Northern Australia at 280 hPa in SON may well be tainted by deep transport of biomass burning.

Over a major industrial source like China, near-surface concentrations of NO_2 around 720 to 820 hPa are expected to reach minimum/maximum levels in JJA/DJF, just on account of increased/reduced exposure to sunlight (i.e. reduced/increased NO_x lifetime), as shown by both OMI and TM4 in Fig. 28. At mid-tropospheric levels though, other effects such as vertical transport intervene. Note in Figs. 26-27 that the TM4 model registers maximum mid-tropospheric NO_2 levels over China in JJA, and minimum in DJF. However, OMI observes stronger mid-tropospheric NO_2 levels in DJF than in JJA. According to OMI, surface emissions from China (and also from Europe and the US, as well shall see next) are being transported in larger quantities and to higher altitudes than in the model, particularly during the winter months.

Europe and North America

In connection with summer convection, lightning activity at northern midlatitudes is expected to be strongest in JJA. Enhacements in upper tropospheric NO_2 are observed by OMI (and TM4) in Fig. 29 over the eastern Mediterranean in JJA and SON. Enhanced NO_2 levels over Siberia in JJA may also be related to summer convection and increased biomass burning. In the US, lightning activity is expected to reach a maximum in JJA and shift southward towards the Gulf of Mexico in SON and DJF, features which are all registered by OMI in Fig. 29 (in

reasonable agreement with TM4, though some discrepancies are apparent in DJF). Figures 30-31 reveal an interesting discrepancy between the OMI and TM4 pseudoprofiles regarding the intensity and reach of convective penetration at 500 hPa of anthropogenic NO_2 above major industrial areas. As already noted for China, the TM4 model is placing enhancements of mid-tropospheric NO_2 over central Europe and eastern US in MAM and JJA, whereas OMI registers a more uniform distribution of mid-tropospheric signatures across the year, showing maxima in DJF and SON. This disagreement is suggestive of problems with the model convective scheme, possibly related to frontal uplift by conveyor belts in the wintertime. At levels closest to the surface, the variation in NO_2 concentration over major industrial areas (Europe and US, but also China, India and the Middle East) registered by OMI in Figs. 19-20 shows minima in JJA and maxima in DJF, just as expected and in agreement with TM4.

In summary, both the OMI and TM4 model pseudoprofiles show seasonal features that are consistent with the available lightning flash and fire count climatologies, and complementary to the results obtained for the annual means. On a finer scale, we observe some significant differences – on lightning distribution, or the intensity and reach of convective transport over strong biomass or industrial sources - whose detailed examination deserves future work.

4 Summary and conclusions

In this paper, we derive a global annual and seasonal climatology of tropospheric NO_2 profiles from OMI cloudy measurements for the year 2006 using the cloud slicing method on six pressure levels centered at about 280, 380, 500, 620, 720 and 820 hPa. The cloud-slicing profiles have been estimated after differencing annual and seasonal tropospheric NO_2 columns above cloud with respect to pressure, using mean cloud pressures located at about 330, 450, 570, 670, 770 and 870 hPa. We term these objects pseudoprofiles, since the required presence of a probing cloud necessarily draws the cloud-slicing estimate away from the underlying NO_2 profile. The systematic error between the cloud-sliced NO_2 pseudoprofile and the actual average NO_2 profile in a cloudy atmosphere is called pseudoprofile error, which can be evaluated (and possibly corrected) using a CTM model.

The total tropospheric NO_2 content in the cloud slicing profiles is consistent with the OMI clear sky total tropospheric column for the same year, after making allowance for a natural

change in the global NO_2 distribution that occurs in passing from clear to cloudy conditions. This change includes suppression of biomass burning during the wet/cloudy season, suppressed NO_2 photolysis under clouds, venting by weather fronts and accumulation patterns dependent on the predominant (clear or cloudy sky) synoptic weather type. The internal consistency between OMI clear-sky and cloud slicing tropospheric NO_2 columns confirms the capability of cloud slicing profiles to detect CTM model anomalies that can be ultimately related to problems in model emission inventories, but with additional vertical information that allows distinction between surface, mid-tropospheric and upper-tropospheric processes.

The vertical information contained in OMI tropospheric NO_2 profiles derived from the cloud slicing technique provides a wealth of information that can be used to evaluate global chemistry models and provide guidance in the development of sub-grid model parameterizations of convective transport, fire-induced injection, horizontal advective diffusion and lightning NO_x production. Overlapping processes (i.e. the effects of deep convection and lightning NO_x in the upper troposphere, the effects of midtropospheric convection and anomalies in surface emissions in the mid-troposphere) as well as uncertainties in the chemical degradation and NO_x recycling rates currently limit the degree to which discrepancies between observations and simulations can be unambiguously attributed to a single process, although the availability of observational constraints definitely constitutes an improvement.

As an example of such an application, we have performed a comparison between cloud slicing tropospheric NO_2 profiles from OMI and the TM4 model. In the upper troposphere (280 and 380 hPa levels), observed NO_2 concentration anomalies reveal excessive model background NO_2 amounts which are consistent with too strong model lightning emissions over the oceans (and/or too long lifetimes) combined with misplaced lightning NO_2 over central Africa and South America, which is indicative of limitations in the convectively driven model lightning NO_x scheme of (Meijer et al., 2001). Other anomalies suggest observed enhanced deep transport of NO_2 from major industrial centers relative to TM4, including a prominent signal from the Asian summer monsoon plume that the model fails to place accurately, and probable excess model fire-induced convection over Siberia.

In the mid troposphere (500 and 620 hPa levels), observed NO_2 concentration anomalies reveal deficient model background NO_2 amounts suggestive of too small model convective inflows into this level, with deficits particularly large over China, central US, and Europe

during the boreal winter, and the biomass burning regions in central Africa and South America, combined with extensive outflows over the oceans that are stronger and more widely distributed in latitude than in the model. This is consistent with independent reports of underestimation of vertical transport by convective clouds in Tiedtke based models. Raising the NO_2 plumes to higher altitudes allows for much longer residence and chemical lifetimes, and longer and more widely distributed horizontal transport of NO_2 following poleward advection and dispersion by the subtropical jet in the mid-tropophere, all of which end up producing typical outflow profiles over the oceans that are generally S-shaped with a prominent mid-tropospheric plume centered around 620 hPa in the tropics and around 720 hPa in the extratropics. The role that the recycled NO_x component may play in the enhanced mid-tropospheric outflows observed by OMI over remote ocean regions is unclear at this stage, but the cloud slicing technique shows promise to study such effects.

In the lower troposphere (720 and 820 hPa), observed NO_2 concentration anomalies show a pattern that is consistent with deficiencies in model surface emissions related to known NO_2 trends characterized by NO_2 increases over China, India and the Middle East, and NO_2 decreases over eastern US, central Europe and Japan. The lower levels also show extensive positive anomalies over the oceans (particularly at 720 hPa), which are indicative of deficient model outflows at low altitudes (and/or too short model lifetimes) with deficient poleward diffusion of NO_2 at low to mid-tropospheric levels, and an interesting band of negative anomalies along the ITCZ.

Note that support from a CTM model (the TM4 in our case) is required to make provision for the cloud-slicing technique (in order to determine a priori corrections for undercloud leakage), so that a level of trust must initially be placed in the model. When comparing the resulting pseudoprofiles against those of the model, a number of potential model shortcomings are found which work against our initial trust in the model corrections. This conflicting outcome should be understood and justified to the extent that a priori corrections have a limited impact on the resulting cloud slicing profiles.

Handling pseudoprofile error using model-based profile-to-pseudoprofile ratios is another matter. The presence of systematic pseudoprofile errors in cloud-slicing estimates and their general predominance over random instrumental error suggest that the vertical information contained in cloudy pixels may be extracted optimally by an assimilation procedure that improves the atmospheric state (i.e. the model profile shape) at the right time and location

using the averaging kernel of the observation. To date, most data assimilation experiments using OMI NO_2 observations have focused on clear-sky measurements. The current results provide strong motivation to use both clear and cloudy pixels in assimilation experiments, as in e.g. (Miyazaki et al., 2014).

Appendix A: Gas columns above and below cloud

If the tropospheric AMF_{trop} is defined as:

$$AMF_{trop} = CRF \cdot AMF_{cloud} + (1 - CRF) \cdot AMF_{clear} \quad \text{(A1)}$$

Where the clear AMF can be expressed as:

$$\begin{aligned} AMF_{clear} &= \frac{\sum_0^{tropopause} m_{clear}(z) \cdot n(z)}{\sum_0^{tropopause} n(z)} \\ &= \frac{\sum_0^{CLP} m_{clear}(z) \cdot n(z) + \sum_{CLP}^{tropopause} m_{clear}(z) \cdot n(z)}{\sum_0^{tropopause} n(z)} \\ &= \frac{\sum_0^{CLP} m_{clear}(z) \cdot n(z)}{\sum_0^{CLP} n(z)} \cdot \frac{\sum_0^{CLP} n(z)}{\sum_0^{trop} n(z)} + \frac{\sum_{CLP}^{trop} m_{clear}(z) \cdot n(z)}{\sum_{CLP}^{trop} n(z)} \cdot \frac{\sum_{CLP}^{trop} n(z)}{\sum_0^{trop} n(z)} \\ &= AMF_{\substack{clear \\ below}} \cdot \frac{VCD_{below}}{VCD_{trop}} + AMF_{\substack{clear \\ above}} \cdot \frac{VCD_{above}}{VCD_{trop}} \end{aligned} \quad \text{(A2)}$$

Where m_{clear} is the clear-sky scattering sensitivity and n(z) is the model a priori trace gas profile. Similarly, the cloudy AMF can be expressed as:

$$\begin{aligned} AMF_{cloud} &= \frac{\sum_0^{tropopause} m_{cloud}(z) \cdot n(z)}{\sum_0^{tropopause} n(z)} \\ &= \frac{\sum_0^{CLP} m_{cloud}(z) \cdot n(z) + \sum_{CLP}^{tropopause} m_{cloud}(z) \cdot n(z)}{\sum_0^{tropopause} n(z)} \\ &= \frac{\sum_{CLP}^{trop} m_{cloud}(z) \cdot n(z)}{\sum_{CLP}^{trop} n(z)} \cdot \frac{\sum_{CLP}^{trop} n(z)}{\sum_0^{trop} n(z)} = AMF_{\substack{cloud \\ above}} \cdot \frac{VCD_{above}}{VCD_{trop}} \end{aligned} \quad \text{(A3)}$$

Where m_{cloudy} is the cloudy-sky scattering sensitivity. Note that by construction:

$$VCD_{trop} = \sum_0^{tropopause} n(z) = VCD_{above} + VCD_{below} \quad \text{(A4)}$$

Then the tropospheric AMF can be written, after inserting Equations (A2) and (A3) into Eq. (A1), and rearranging terms relating to above and below components separately as:

$$AMF_{trop} = \frac{VCD_{above}}{VCD_{trop}}\left(CRF \cdot AMF_{\substack{cloud \\ above}} + (1 - CRF) \cdot AMF_{\substack{clear \\ above}}\right) + \frac{VCD_{below}}{VCD_{trop}}(1 - CRF)$$

$$\cdot AMF_{\substack{clear \\ below}} = \frac{VCD_{above}}{VCD_{trop}} AMF_{above} + \frac{VCD_{below}}{VCD_{trop}} AMF_{below} \quad \text{(A5)}$$

From this formulation arise definitions for AMF_{above} and AMF_{below}:

$$AMF_{above} \equiv \frac{\sum_{CLP}^{trop}(CRF \cdot m_{cloud}(z) + (1 - CRF) \cdot m_{clear}(z)) \cdot n(z)}{\sum_{CLP}^{trop} n(z)} \quad \text{(A6)}$$

$$AMF_{below} \equiv \frac{\sum_0^{CLP}(1 - CRF) \cdot m_{clear}(z) \cdot n(z)}{\sum_0^{CLP} n(z)} \quad \text{(A7)}$$

Now it is straightforward to write:

$$SCD_{trop} = AMF_{trop} \cdot VCD_{trop}$$

Which after substitution of Eq. (A5) becomes

$$SCD_{trop} = \left(\frac{VCD_{above}}{VCD_{trop}} \cdot AMF_{above} + \frac{VCD_{below}}{VCD_{trop}} \cdot AMF_{below}\right) \cdot VCD_{trop}$$

$$= VCD_{above} \cdot AMF_{above} + VCD_{below} \cdot AMF_{below} = SCD_{above} + SCD_{below} \quad \text{(A8)}$$

Allowing the separation of the slant components above and below the cloud as:

$$VCD_{above} = (SCD_{trop} - SCD_{below})/AMF_{above} \quad \text{(A9)}$$

Now, in [Boersma, ACP, 2005] the above-cloud part of the NO_2 column is retrieved by removing the model predicted ghost column (integrated from the ground to the cloud level pressure, identical to VCD_{below}) that is implicitly added via the tropospheric airmass factor as:

$$VCD_{above} = SCD_{trop}/AMF_{trop} - CRF \cdot VCD_{below} \quad \text{(A10)}$$

However, by virtue of Eq. (A4), formulation in Eq. (A10) in [Boersma, ACP, 2005] should be changed to:

$$VCD_{above} = SCD_{trop}/AMF_{trop} - VCD_{below} \quad \text{(A11)}$$

Which is equivalent to Eq. (A9).

Acknowledgements

This work has been funded by the Netherlands Space Office (NSO) under OMI contract.

References

Acarreta, J.R., De Haan, J.F., Stammes, P.: Cloud pressure retrieval using the O_2-O_2 absorption band at 477 nm, J. Geophys. Res., 109(D5), doi:10.1029/2003JD003915, 2004.

Albrecht, R.I., Naccarato, K.P., Pinto, O., Pinto, I.R.C.A.: Total lightning and precipitation over Brazil: An overview from 12-years of TRMM satellite, AMS Conference, Seattle (WA), January 22-27, abstract No. 185724, 2011.

Arino, O., Casadio, S., Serpe, D.: Global night-time fire season timing and fire count trends using the ATSR instrument series, Remote Sens. Environ., 116, 226-238, 2012.

Barret, B., Williams, J.E., Bouarar, I., Yang, X., Josse, B., Law, K., Pham, M., Flochmoen, E., Liousse, C., Peauch, V.H., Carver, G.D., Pyle, J.A., Sauvage, B., Velthoven, P., Schlager, H., Mari, C., Cammas, J.P.: Impact of West African Mondoon convective transport ind lightning NO_x production upon the upper tropospheric composition: a multimodel study, Atmos. Chem. Phys., 10, 5719-5738, 2010.

Beirle, S., Boersma, F., Platt, U., Lawrence, M.G., Wagner, T.: Megacity emissions and lifetimes of nitrogen oxides probed from space, Science, 333, 1737-1739, 2011.

Beirle, S., Spichtinger, N., Stohl, A., Cummins, K.L., Turner, T., Boccipio, D., Cooper, O.R., Wenig, M., Grzegorski, M., Platt, U., Wagner, T.: Estimating the NO_x produced by lightning from GOME and NLDN data: a case study in the Gulf of Mexico, Atmos. Chem. Phys., 6, 1075-1089, 2006.

Belmonte Rivas, M., Veefkind, P., Boersma, F., Levelt, P., Eskes, H., Gille, J.: Intercomparison of daytime stratospheric NO_2 satellite retrievals and model simulations, Atmos. Meas. Tech., 7, 2203-2225, 2014.

Boersma, F., Eskes, H., Brinskma, E.: Error analysis for tropospheric NO_2 retrieval from space, J. Geophys. Res., 109, doi:10.1029/2003JD003962, 2004.

Boersma, F., Esker, H, Meijer, E. W., Kelder, H.: Estimates of lightning NO_x production from GOME satellite observations, Atmos. Chem. Phys., 5, 2311-2331, 2005.

Boersma, F., Eskes, H., Veefkind, J.P., Brinksma, E.J., van der A, R.J., Sneep, M., van den Oord, G., Levelt, P.F., Stammes, P., Gleason, J.F., Bucsela, E.J.: Near-real time retrieval of tropospheric NO_2 from OMI, Atmos. Chem. Phys., 7, 2103-2118, 2007.

Boersma, F., Eskes, H., Dirksen, R.J., van der A., R.J., Veefkind, J.P., Stammes, P., Huijnen, V., Kleipool, Q.L., Sneep, M., Claas, J., Leitao, J., Richter, A., Zhou, Y., Brunner, D.: An improved tropospheric NO_2 column retrieval algorithm for the ozone monitoring instrument, Atmos. Meas. Tech., 4, 1905-1928, 2011.

Boersma, K. F., Vinken, G. C. M., and Eskes, H. J.: Representativeness errors in comparing chemistry transport and chemistry climate models with satellite UV/Vis tropospheric column retrievals, Geosci. Model Dev. Discuss., in press, gmd-2015-134, 2015.

Boucher, O., D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, P. Rasch, S.K. Satheesh, S. Sherwood, B. Stevens and X.Y. Zhang: Clouds and Aerosols. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Cecil, D.J., Buechler, D.E., Blakeslee, R.J.: Gridded lightning climatology from TRMM-LIS and OTD: Dataset description, Atmos. Res., 135-136, 404-414, 2014.

Choi, S., Joiner, J., Choi, Y., Duncan, B.N., Vasilkov, A., Krotkov, N., Bucsela, E.: First estimates of global free-tropospheric NO_2 abundances derived using a cloud-slicing technique applied to satellite observations from the Aura Ozone Monitoring Instrument (OMI), Atmos. Chem. Phys., 14, 10565-10588, 2014.

Dickerson, R.R.: Measurements of reactive nitrogen compounds in the free troposphere, Atmos. Environ., 18, 12, 2585-2593, 1984.

Ding, J., van der A, R.J., Mijling, B., Levelt, P.F., and Hao, N.: NOx emission estimates during the 2014 Youth Olympic Games in Nanjing, Atmos. Chem. Phys. Discuss., 15, 6337-6372, doi: 10.5194/acpd-15-6337-2015, 2015.

Dwyer, N., Pinnock, S., Gregoire, J.-M., Pereira, J.M.C., Global spatial and temporal distribution of vegetation fire as determined from satellite observations, Int. J. Rem. Sens., 21(6):1289, doi: 10.1080/014311600210182, 2000.

Folkins, I., Bernath, P., Boone, C., Donner, L.J., Eldering, A., Lesins, G., Martin, R.V., Sinnhuber, B.M., Walker, K.: Testing convective parameterizations with tropical

measurements of HNO_3, CO, H_2O and O_3: Implications for the water vapor budget, J. Geophys. Res., 111, D23304, doi:10.1029/2006JD007325, 2006.

Houweling, S., Dentener, F.J., Lelieveld, J.: The impact of non-methane hydrocarbon compounds on tropospheric chemistry, J. Geophys. Res., 103, 10673-10696, 1998.

Hoyle, C.R., Marecal, V., Russo, M.R., Allen, G., Arteta, J., Chemel, C., Chipperfiled, M.P., D'Amato, F., Dessens, O., Feng, W., Hamilton, J.F., Harris, N.R.P., Hosking, J.S, Lewis, A.C., Morgenstern, O., Peter, T., Pyle, J.A., Reddmann, T., Richards, N.A.D., Telford, P.J., Tian, W., Viciani, S., Volz-Thomas, A., Wild, O., Yang, X., Zeng, G.: Representation of tropical deep convecntion in atmospheric models – Part 2: Tracer transport, Atmos. Chem. Phys., 11, 8103-8131, 2011.

Huijnen, V., Flemming, J., Kaiser, J.W., Inness, A., Leitao, J., Heil, A., Eskes, H., Schultz, M., Benedetti, A., Dufour, G., Eremenko, M.: Hindcast experiments of tropospheric composition during the summer 2010 fires over Western Russia, Atmos. Chem. Phys., 12, 4341-4364, 2012.

Jakob, C.: An improved strategy for the evaluation of cloud parameterizations in GCMs, Bull. Am. Meteorol. Soc., doi: 10.1175/BAMS-84-10-1387, 2003.

Kar, J., Bremer, H., Drummond, J.R., Rochon, Y.J., Jones, D.B., Nichitiu, F., Zou, J., Liu, J., Gille, J., Edwards, D.P., Deeter, M., Francis, G., Ziskin, D., Warner, J.: Evidence of vertical transport of carbon monoxide from MOPITT, Geophys. Res. Lett., 31, L23105, doi: 10.1029/2004GL021128, 2004.

Krol, M.C., van Weele, M.: Implications of variations in photodissociation rates for global tropospheric chemistry, Atmos. Environ., 31, 1257-1273, 1997.

Landgraf, J., Crutzen, P.J.: An efficient method for online calculations of photolysis and heating rates, J. Atmos. Sci., 55, 863-878, 1998.

Levelt, P.F., van den Oord, G., Dobber, M.R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J., Saari, H.: The Ozone Monitoring Instrument, IEEE Trans. Geosci. Remote Sens., 44(5), 1093-1101, 2006.

Liu, C., Beirle, S., Butler, T., Hoor, P., Frankenberg, C., Jockel, P., Penning de Vries, M., Platt, U., Pozzer, A., Lawrence, M.G., Lelieveld, J., Tost, H., Wagner, T.: Profile information

on CO from SCIAMACHY observations using cloud slicing and comparison with model simulations, Atm. Chem. Phys., 17, 1717-1732, 2014.

Mahowald, N.M., Rasch, P.J., Prinn, R.G.: Cumulus parameterization in chemical transport models, J. Geophys. Res., 100(D12), 26173-26189, 1995.

Martin, R.V., Jacob, D.J., Chance, K., Kurosu, T.P, Palmer, P.I., Evans, M.J.: Global inventory of nitrogen oxide emissions constrained by space-based observations of NO_2 columns, J. Geophys. Res., 108(D17), 4537, doi:10.1029/2003JD003453, 2003.

Martin, R.V., Sauvage, B., Folkins, I., Sioris, C.E, Boone, C., Bernath, P., Ziemke, J.: Space-based constraints on the production of nitric oxide by lightning, J. Geophys. Res., 112(D09309), doi: 10.1029/2006JD007831, 2007.

Meijer, E.W., Velthoven, P.F.J., Brunner, D.W., Huntrieser, H., Kelder, H.: Improvement and evaluation of the parameterisation of nitrogen oxide production by lightning, Phys. Chem. Earth, 26(8), 557-583, 2001.

Mijling, B., van der A, R.J.: Using daily satellite observations to estimate emissions of short lived air pollutants on a mesoscopic scale, J. Geophys. Res., 117(D17), doi: 10.1029/2012JD017817, 2012.

Miyazaki, K., Eskes, H. J., and Sudo, K.: Global NO_x emission estimates derived from an assimilation of OMI tropospheric NO_2 columns, Atmos. Chem. Phys., 12, 2263-2288, 2012.

Miyazaki, K., Eskes, H. J., Sudo, K., and Zhang, C.: Global lightning NOx production estimated by an assimilation of multiple satellite data sets, Atmos. Chem. Phys., 14, 3277-3305, 2014

Murray, L.T., Jacob, D.J., Logan, J., Hudman, R., Koshak, W.J.: Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS-OTD satellite data, J. Geophys. Res., 117(D20), doi: 10.1029/2012JD017934, 2012.

Nam, C., Quaas, J., Neggers, R., Siegenthaler-Le Drian, C., Isotta, F.: Evaluation of boundary layer cloud parameterizations in the ECHAM5 general circulation model using CALIPSO and CloudSAT satellite data, J. Adv. Model. Earth Syst., 6(2), 300-314, 2014.

Olivier, J., Peters, J., Granier, C., Petron, G., Müller, J. F., and Wallens, S.: Present and future emissions of atmospheric compounds, POET report #2. EU report EV K2-1999-00011, 2003.

Pickering, K.E., Thompson, A.M., Dickerson, R.R., Luke, W.T., McNamara, D.P., Free tropospheric ozone production following entrainment of urban plumes into deep convection, J. Geophys. Res., 97, 17985-18000, 1992.

Pickering, K.E., Wang, Y., Tao, W.K., Price, C., Muller, F.: Vertical distributions of lightning NO_x for use in regional and global chemical transport models, J. Geophys. Res., 103(D23), 31203-31216, 1998.

Pope, R.J., Savage, N.H., Chipperfield, M.P., Arnold, S.R., Osborn, T.J.: The influence of synoptic weather regimes on UK air quality: analysis of satellite column NO_2, Atmos. Sci. Lett., 15(3), 211-217, 2014.

Richter, A., Burrows, J.P., Nuss, H., Granier, C., Niemeier, U.: Increase in tropospheric nitrogen dioxide over Chine observed from space, Nature, 437, 129-132, 2005.

Russell, G.L, Lerner, J.A.: A new finite differencing scheme for the tracer transport equation, J. Appl. Meteor., 20, 1483-1498, 1981.

Schumann, U., Huntrieser, H.: The global lightning-induced nitrogen oxides source, Atmos. Chem. Phys., 7, 3823-3907, 2007.

Schultz, M. G., et al., On the origin of tropospheric ozone and NO_x over the tropical south Pacific, J. Geophys. Res., 104, 5829–5844, 1999.

Schultz, M.G., On the use of ATSR fire count data to estimate the seasonal and interannual variability of vegetation fire emissions, Atmos. Chem. Phys., 2, 387-395, doi: 10.5194/acp-2-387-2002, 2002.

Sneep, M., de Haan, J.F., Stammes, P., Wang, P., Vanbauce, C., Joiner, J., Vasilkov, A.P., Levelt, P.F.: Three way comparison between OMI and PARASOL cloud pressure products, J. Geophys. Res., 113(D15), doi: 10.1029/2007JD008694, 2008.

Stammes, P., Sneep, M., de Haan, J.F., Veefkind, J.P., Wang, P., Levelt, P.F.: Effective cloud fraction from the Ozone Monitoring Instrument: Theoretical framework and validation, J. Geophys. Res., 113(D16), doi: 10.1029/2007JD008820, 2008.

Staudt, A.C., Jacob, D.J., Ravetta, F., Logan, J.A., Bachiochi, D., Krishnamurti, T.N., Sandholm, S., Ridley, B., Singh, H.B., Talbot, B.: Sources and chemistry of nitrogen oxides over the tropical Pacific, J. Geophys. Res., 108(D2), doi: 10.1029/2002JD002139, 2003.

Tiedtke, M.: A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models, Mon. Weather Rev., 117, 1779–1800, 1989.

Tie, X., Madronich, S., Waters, S., Zhang, R., Rasch, P., Collins, W.: Effects of clouds on photolysis and oxidants in the troposphere, J. Geophys. Res., 108(D20), doi: 10.1029/2003JD003659, 2003.

Tost, H., Jockel, P., Lelieveld, J.: Lightning and convection parameterization – uncertainties in global modelling, Atmos. Chem. Phys., 7, 4553-4568, 2007.

Van der A, R.J., Eskes, H., Boesma, K.F., Noije, T.P.C., Roozendael, M., Smedt, I., Peters, D.H.M.U., Meijer, E.W.: Trends, seasonal variability and dominant NO_x sources derived from a ten year records of NO_2 measured from space, J. Geophys. Res., 113(D4), doi: 10.1029/2007JD009021, 2008.

Williams, J.E., Scheele, R., Velthoven, P., Bouarar, I., Law, K., Josse, B., Peuch, V.H., Yang, X., Pyle, J., Thouret, V., Barret, B., Liousse, C., Hourdin, F., Szopa, S., Cozic, A.: Global chemistry simulations in the AMMA (African Monsoon Multidisciplinary Analysis) multimodel intercomparison project, Bull. Am. Meteorol. Soc., 91, 5, 612-624, 2010.

Ziemke, J.R., Chandra, S., Barthia, P.K.: Cloud slicing: A new technique to derive upper tropospheric ozone from satellite measurements, J. Geophys. Res., 106(D9), 9853-9867, 2001.

Table 1. Cloud pressure intervals and mean cloud pressure levels used for cloud slicing (hPa): the VCD pressure interval gives the boundaries of the cloud pressure bin. The VMR pressure interval refers to where the VMR is assumed constant after the pressure difference.

	VCD Pressure Interval	<VCD pressure>	VMR Pressure Interval	<VMR pressure>
Level 1	Tropopause - 380	330	Tropopause - 330	280
Level 2	380 - 500	450	330 - 450	380
Level 3	500 - 620	570	450 - 570	500
Level 4	620 - 720	670	570 - 670	620
Level 5	720 - 820	770	670 - 770	720
Level 6	820 - 1000	870	770 - 870	820

1 Table 2. Model based source and outflow class definitions based on EOF decomposition.

2

Class label	Main condition	Extra condition
Primary industrial	EOF1 > 400 pptv	US, Europe, China
Secondary industrial	100 pptv < EOF1 < 400 pptv	US, Europe, China
Biomass burning	100 pptv < EOF1 < 400 pptv	geographic
Baykal highway	100 pptv < EOF1 < 400 pptv	geographic
Indostan	100 pptv < EOF1 < 400 pptv	geographic
Middle East	100 pptv < EOF1 < 400 pptv	geographic
Tropical outflow	EOF1<50 pptv, EOF2>15 pptv	EOF3>0, EOF4>0
Tropical subsidence	EOF1<50 pptv, EOF2>15 pptv	EOF3<0
Extratropical outflow	EOF1<50 pptv, EOF2>15 pptv	EOF3>0, EOF4<0
Boreal Outflow	EOF1<50 pptv, EOF2>15 pptv	EOF3 >>0
Clean background	EOF1<15 pptv, EOF2<15 pptv	

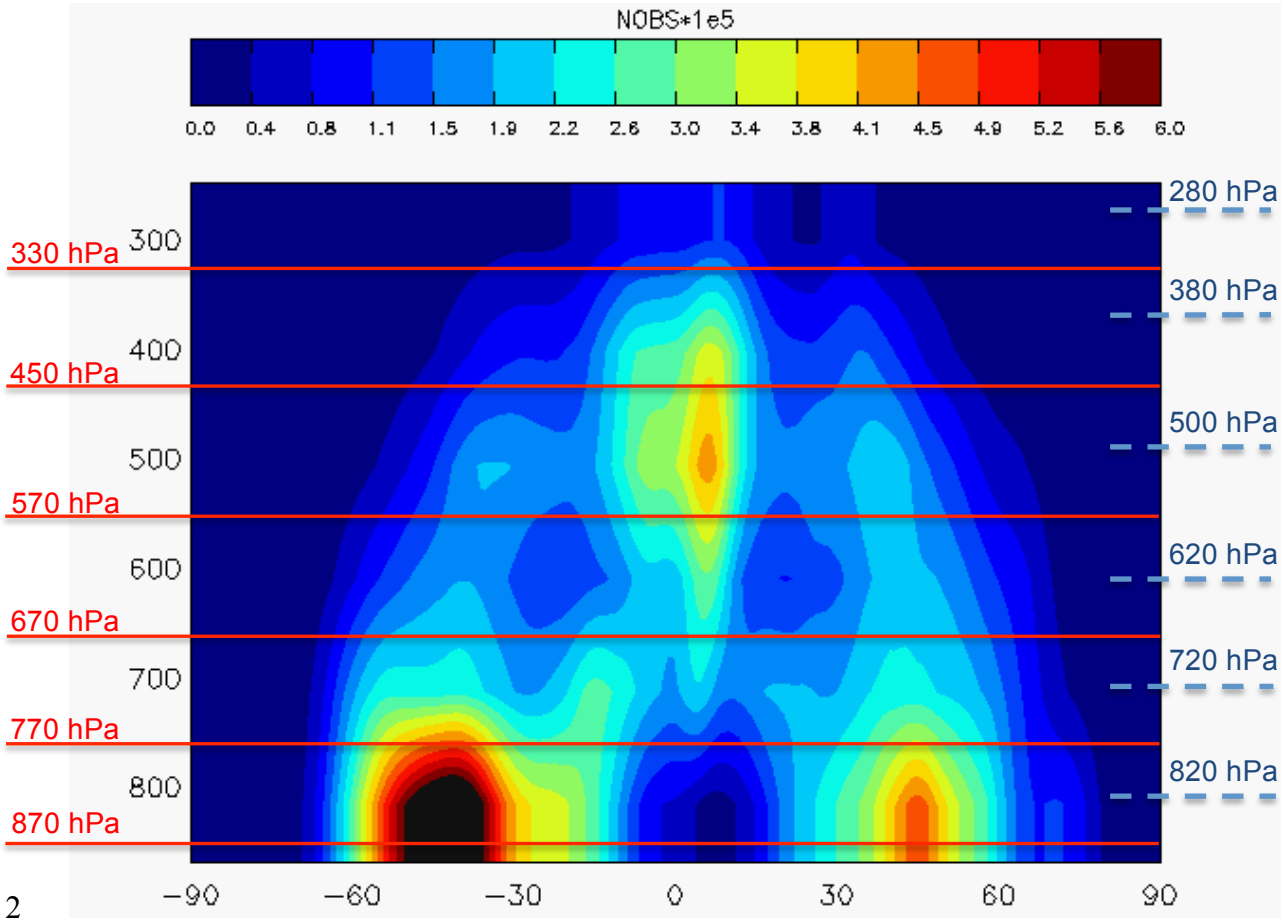


Figure 1. Latitude-height section of annual zonal mean OMI cloud frequencies (CRF>50%) - observed during daytime around 13.45 LST. On the left in red, the bottom pressure boundaries for the calculation of annual mean NO_2 VCDs above cloud (after Table 1). On the right in blue, the approximate pressure for the resulting NO_2 VMR after differenciation of VCDs (also after Table 1).

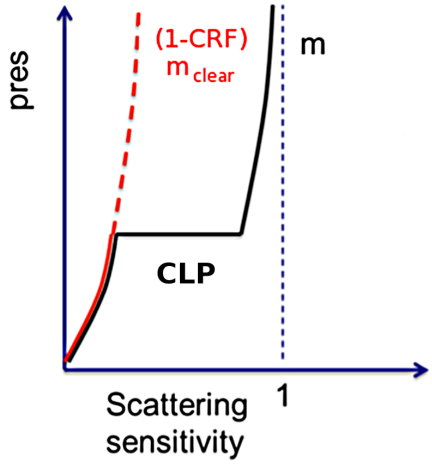


Figure 2. Schematic diagram of the scattering sensitivity above and below the cloud (normalized by the geometric air mass factor): CLP is the cloud level pressure, and *m* is the total scattering sensitivity, usually defined as $(1\text{-}CRF)\ m_{clear} + CRF\ m_{cloudy}$. The red curve illustrates a residual sensitivity to NO_2 contents below the cloud when conditions are partially cloudy.

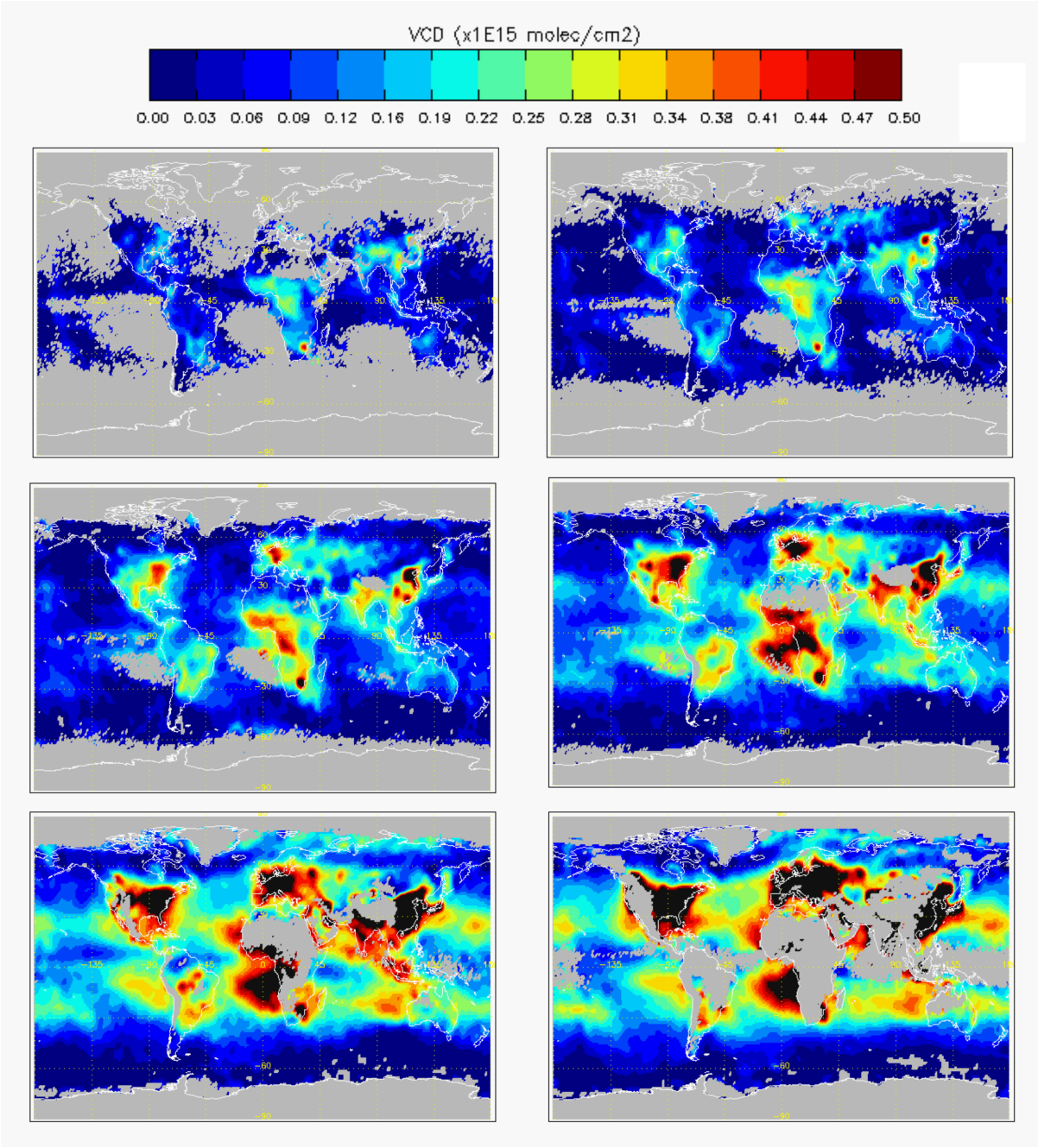


Figure 3a. OMI NO_2 VCDs above cloud - average quantities for the year 2006: for high altitude clouds (top row, 330 and 450 hPa), mid altitude clouds (middle row, 570 and 670 hPa) and low clouds (bottom row, 770 and 870 hPa). Grey means no data available (i.e. insufficient number of cloud detections in the cell).

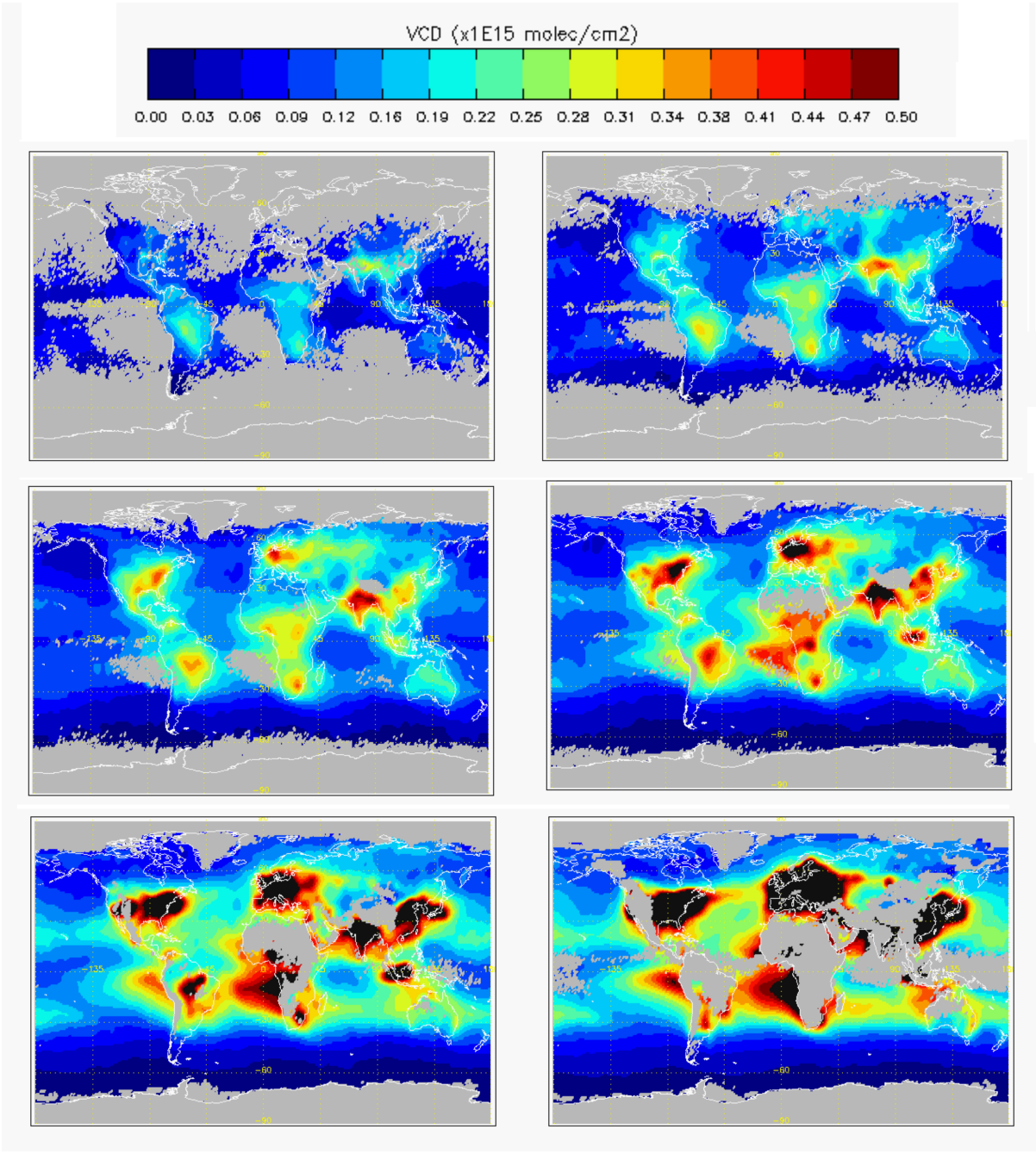


Figure 3b. TM4 NO_2 VCDs above cloud - average quantities for the year 2006: for high altitude clouds (top row, 330 and 450 hPa), mid altitude clouds (middle row, 570 and 670 hPa) and low clouds (bottom row, 770 and 870 hPa). Grey means no data available (i.e. insufficient number of cloud detections in the cell).

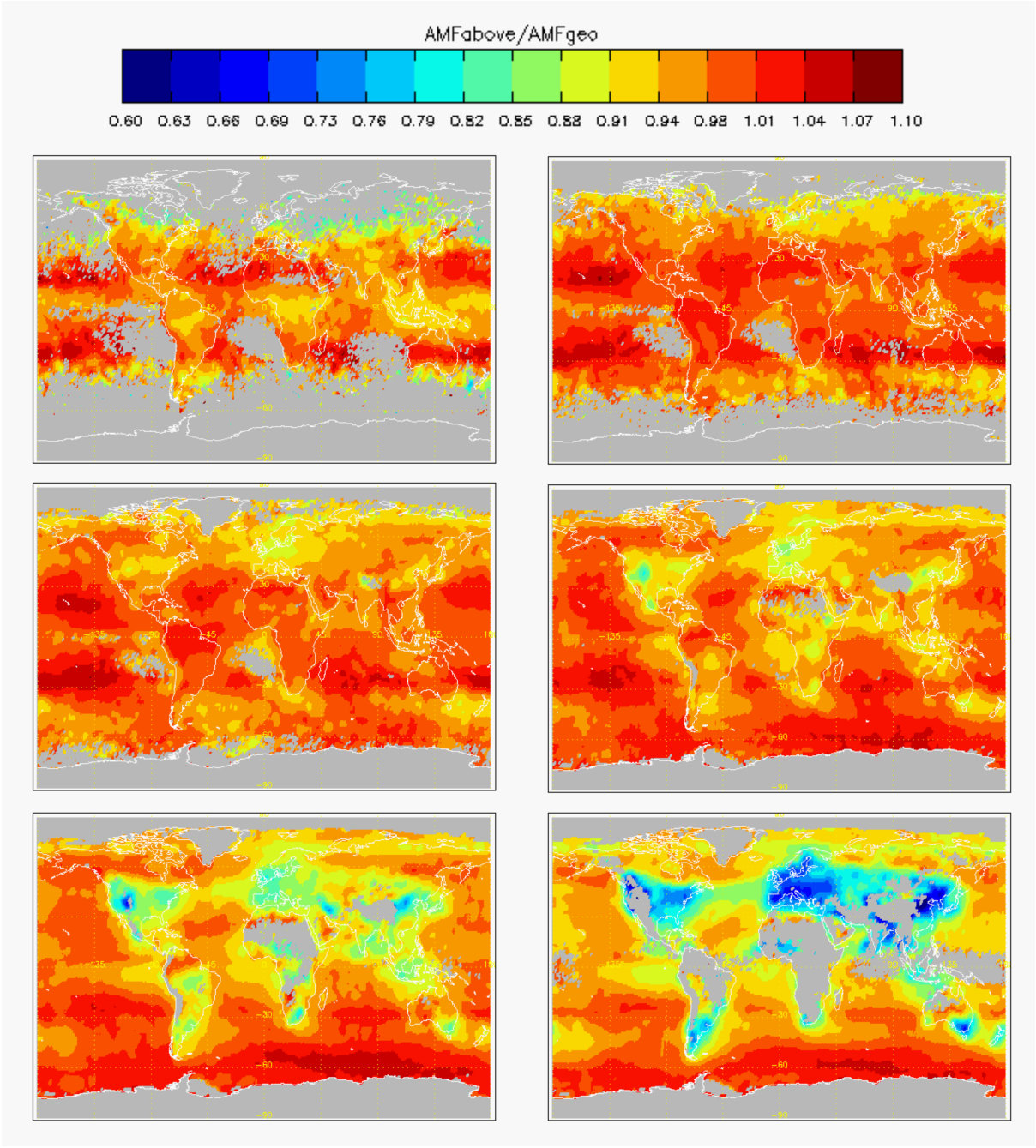


Figure 4. Tropospheric scattering sensitivities above cloud level [AMF_{above}/AMF_{geo} in Eq. (4)]: for high altitude clouds (top row, 330 and 450 hPa), mid altitude clouds (middle row, 570 and 670 hPa) and low clouds (bottom row, 770 and 870 hPa).

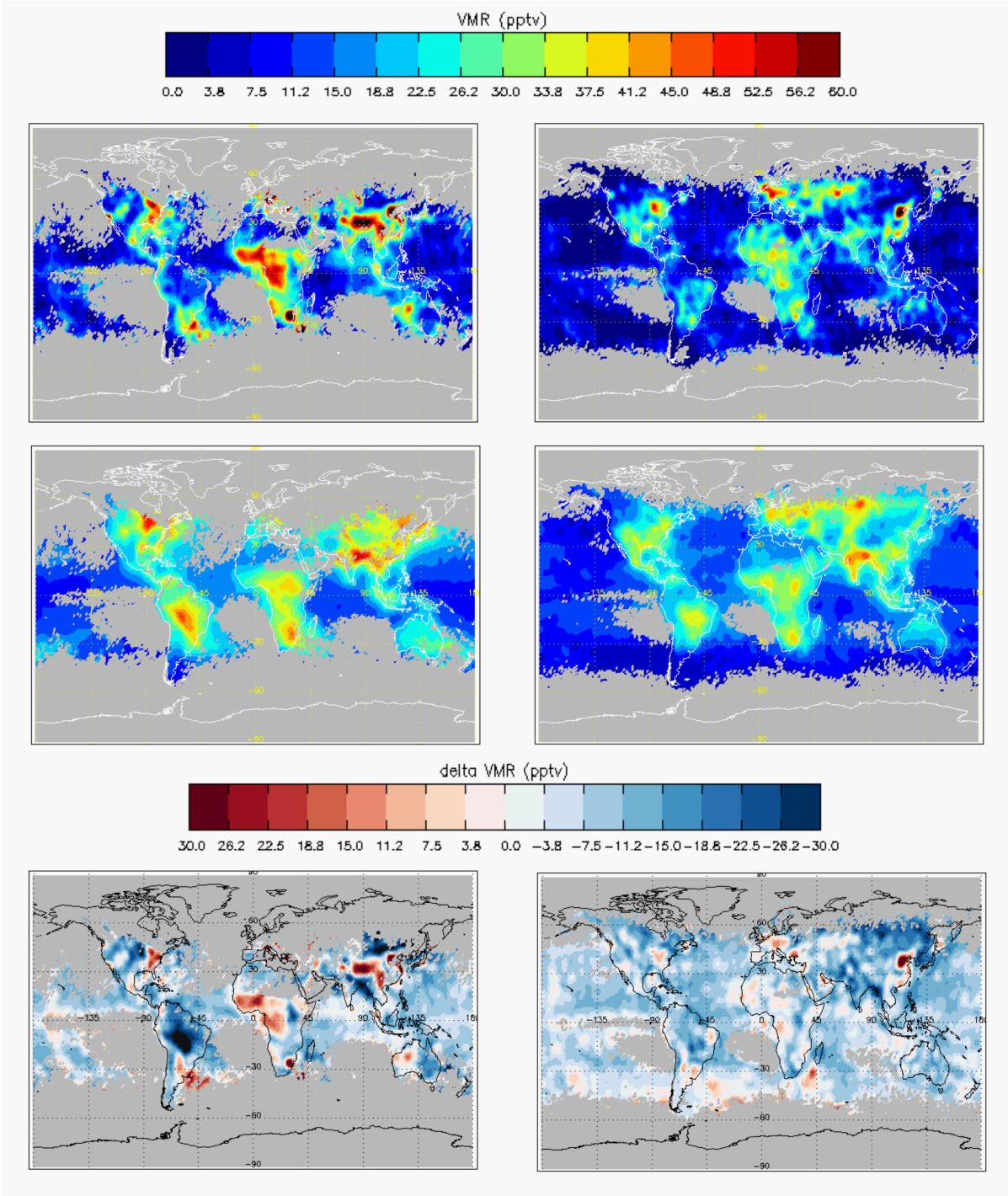


Figure 5a. Upper cloud levels (280 hPa left, 380 hPa right): OMI versus TM4 model NO_2 VMRs (OMI top, TM4 middle, difference bottom) average quantities for the year 2006.

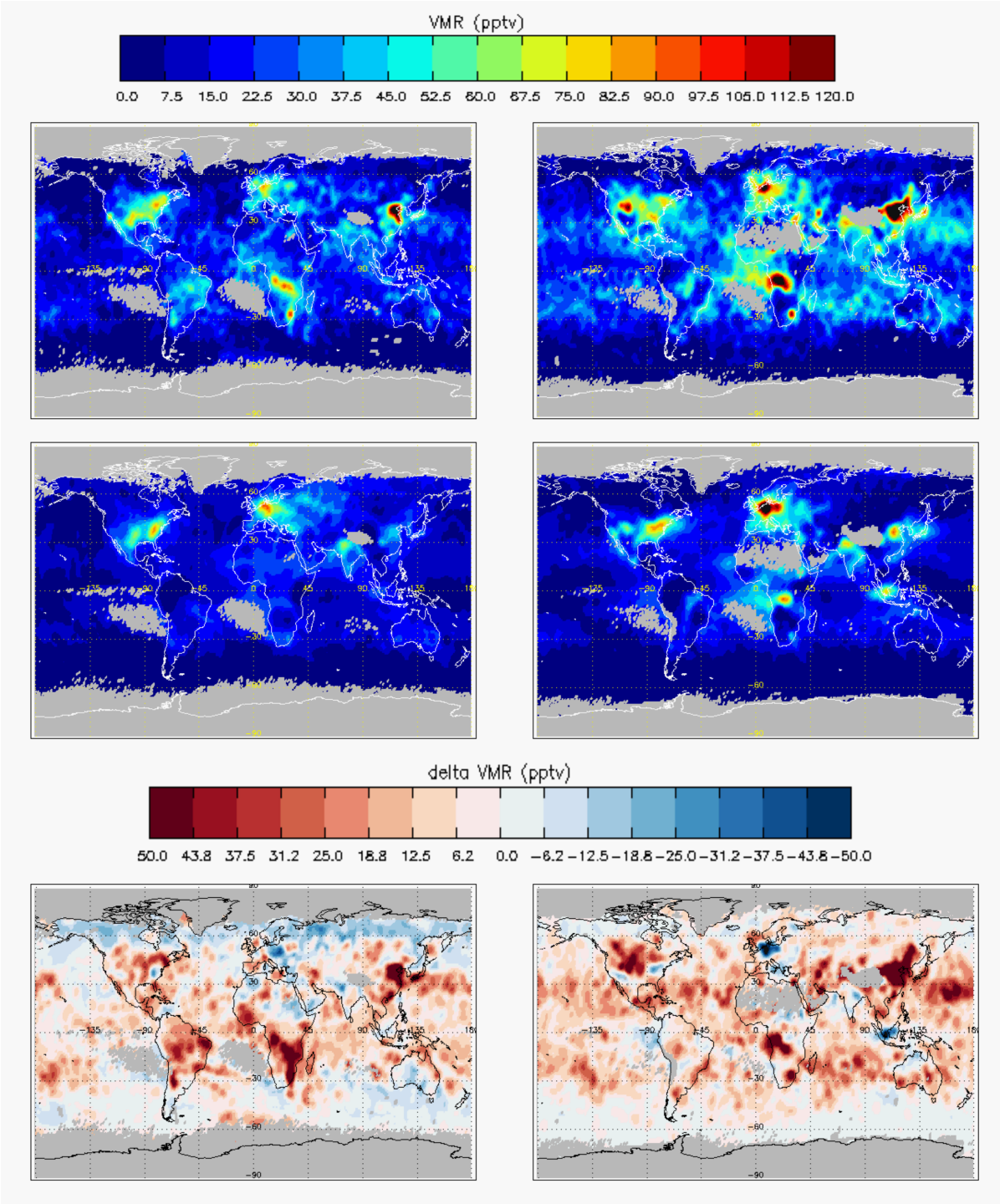


Figure 5b. Middle cloud levels (500 hPa left and 620 hPa right): OMI versus TM4 model NO_2 VMRs (OMI top, TM4 middle, difference bottom) average quantities for the year 2006.

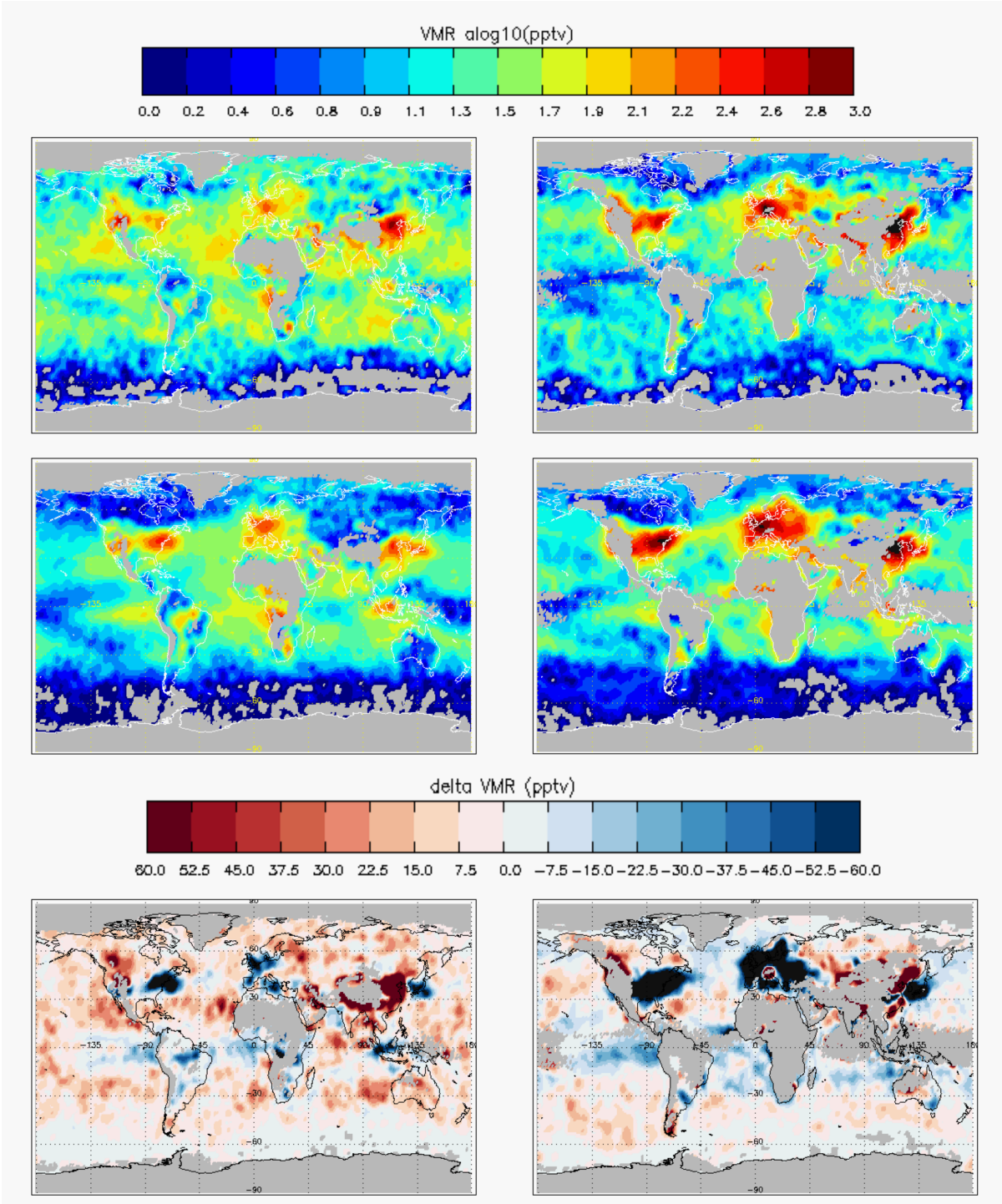


Figure 5c. Lower cloud levels (720 hPa left and 820 hPa right): OMI versus TM4 model NO_2 VMRs (OMI top, TM4 middle, difference bottom) average quantities for the year 2006.

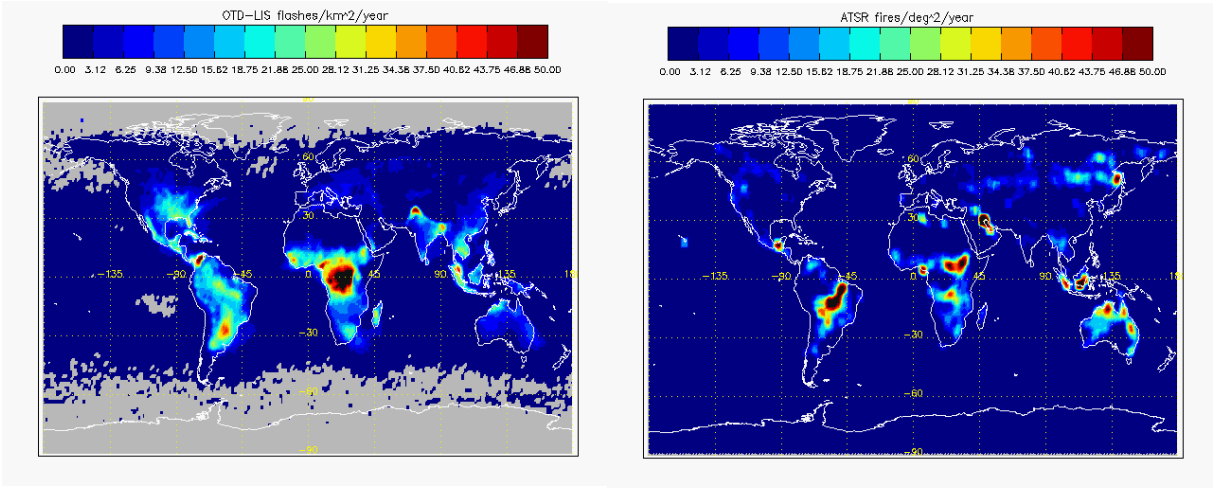


Figure 6. Interpretation aids for process attribution: mean flash rate climatology (1998-2010) from the LIS-OTD sensor (left, [Cecil et al., 2014]) and fire count climatology (1997-2003) from the ATSR sensor (right, [Arino et al., 2012]).

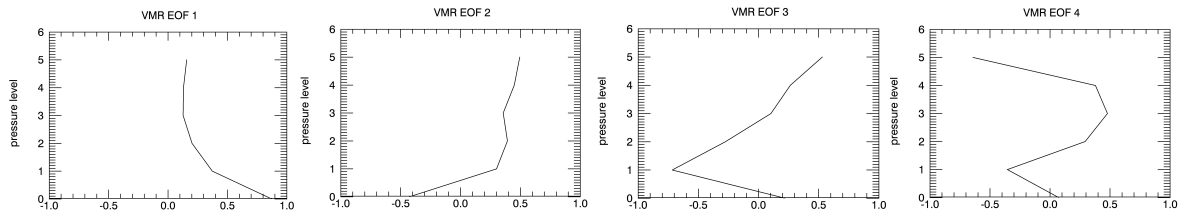


Figure 7a. Classification EOFs: surface source, outflow, high/low outflow, middle outflow.

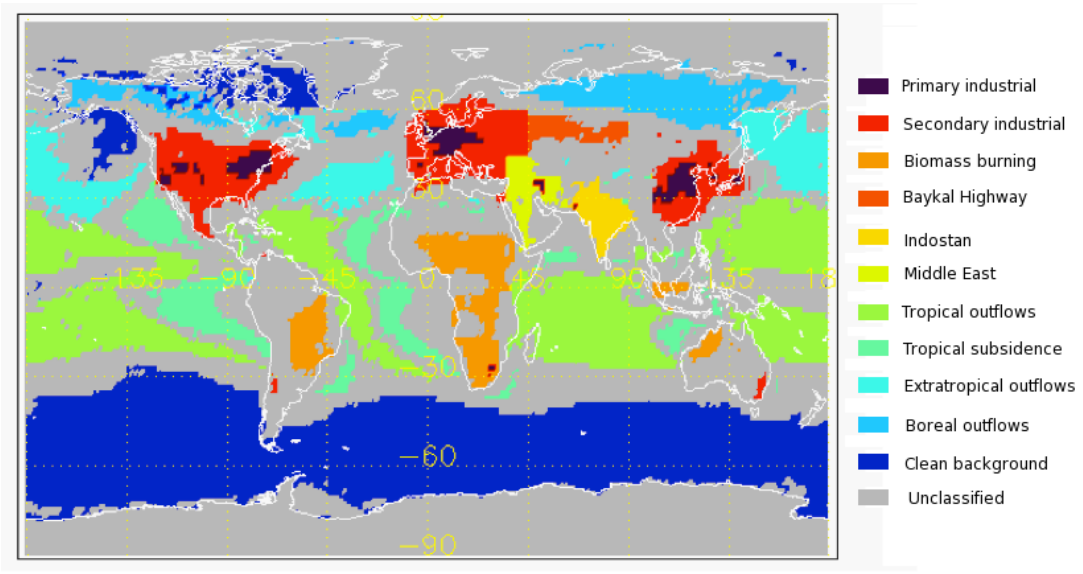


Figure 7b. Model based classes based on EOF decomposition of model NO_2 profiles under cloudy conditions: black (primary industrial), red (secondary industrial), orange (biomass burning), ochre (Baykal Highway), yellow (Indostan), light green (Middle East), green (tropical outflow), turquoise (tropical subsidence), cyan (extratropical outflow), blue (boreal outflow), dark blue (clear background). Gray for unclassified.

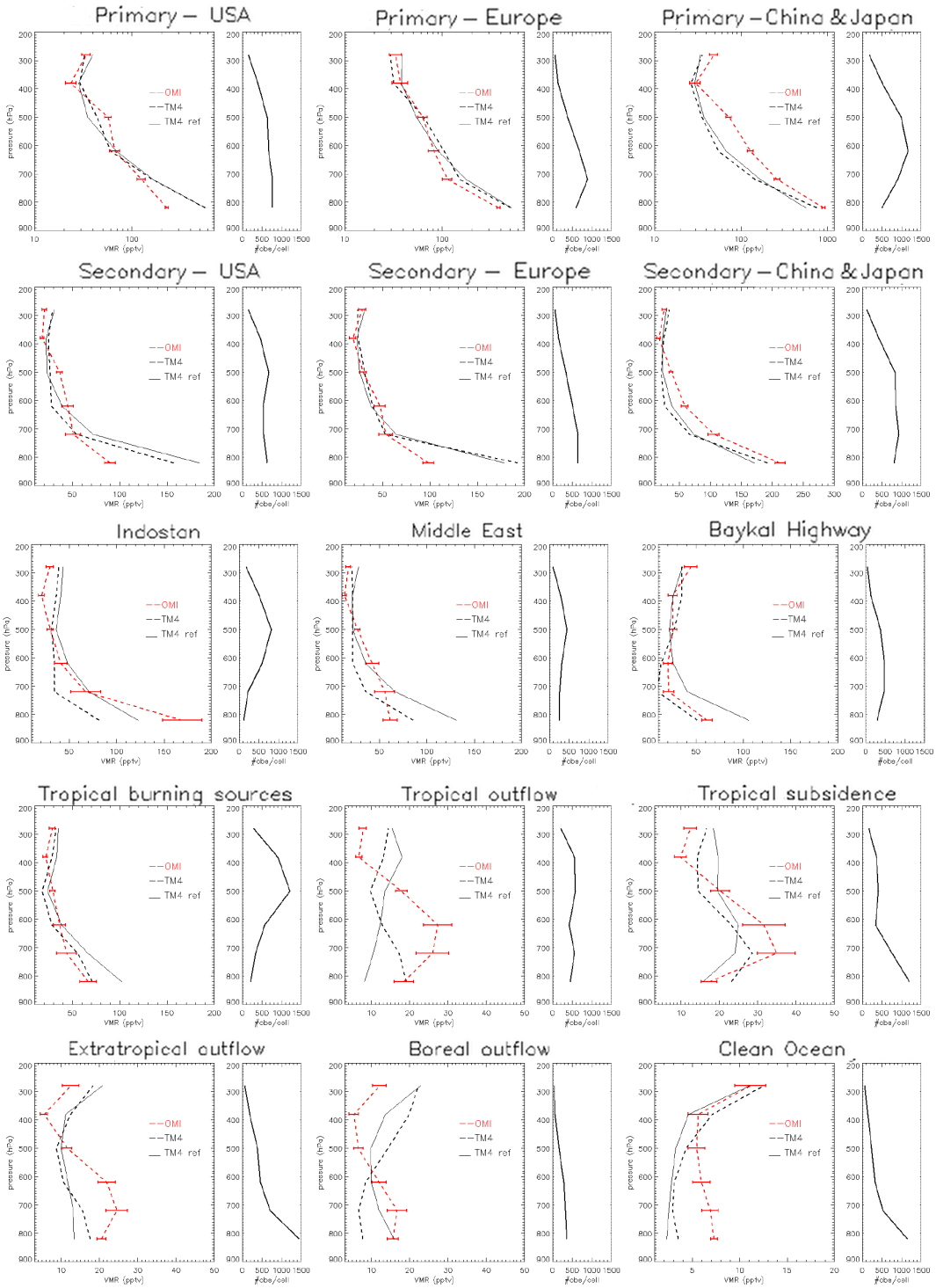


Figure 8. Cloud-slicing NO_2 VMR profiles for the year 2006 by class (OMI pseudoprofile, dashed red line; TM4 pseudoprofile, dashed black line; TM4 profile for cloudy conditions, continuous black line). The error bars show random retrieval errors. The differences between continuous and dashed black lines show systematic pseudoprofile errors. The subpanels on the right show the average number of OMI observations collected per grid cell per year for that class.

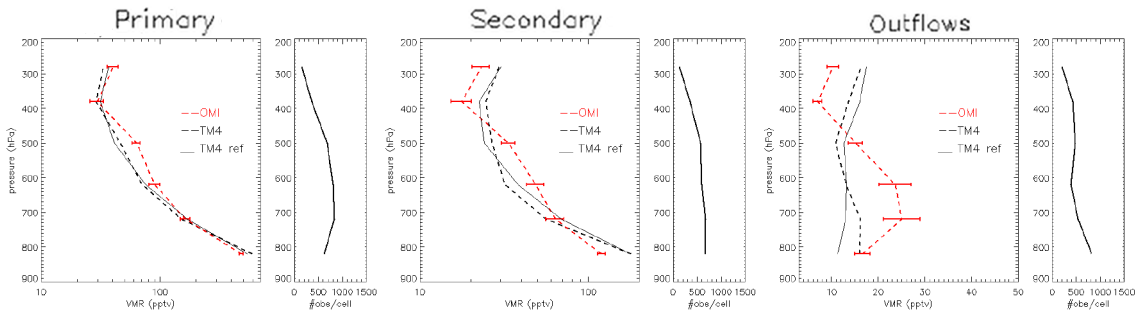
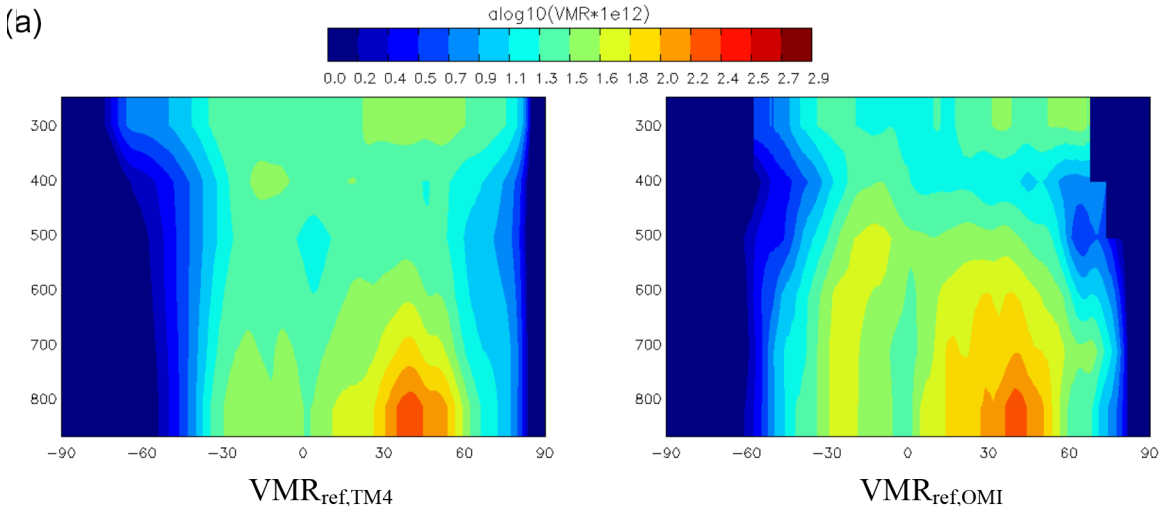


Figure 9. Cloud slicing NO_2 VMR profiles for year 2006 by class: all primary sources (left), all secondary sources (middle) and all outflow classes (right). (OMI pseudoprofile, dashed red line; TM4 pseudoprofile, dashed black line; TM4 profile for cloudy conditions, continuous black line). The error bars show random retrieval errors. The differences between continuous and dashed black lines show systematic pseudoprofile errors.



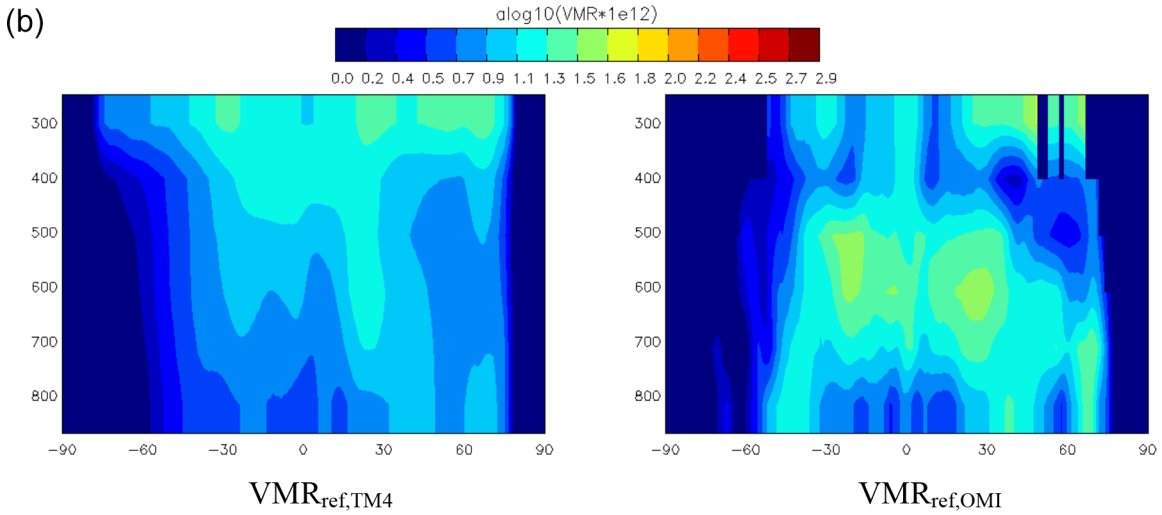


Figure 10a. Latitude-height cross-section of annual zonal mean tropospheric NO_2 from TM4 (left) and OMI (right) with CRF>50%.

Figure 10b. Latitude-height cross-section of annual zonal mean tropospheric NO_2 from TM4 (left) and OMI (right) with CRF>50% over the remote pacific sector (180W-135W).

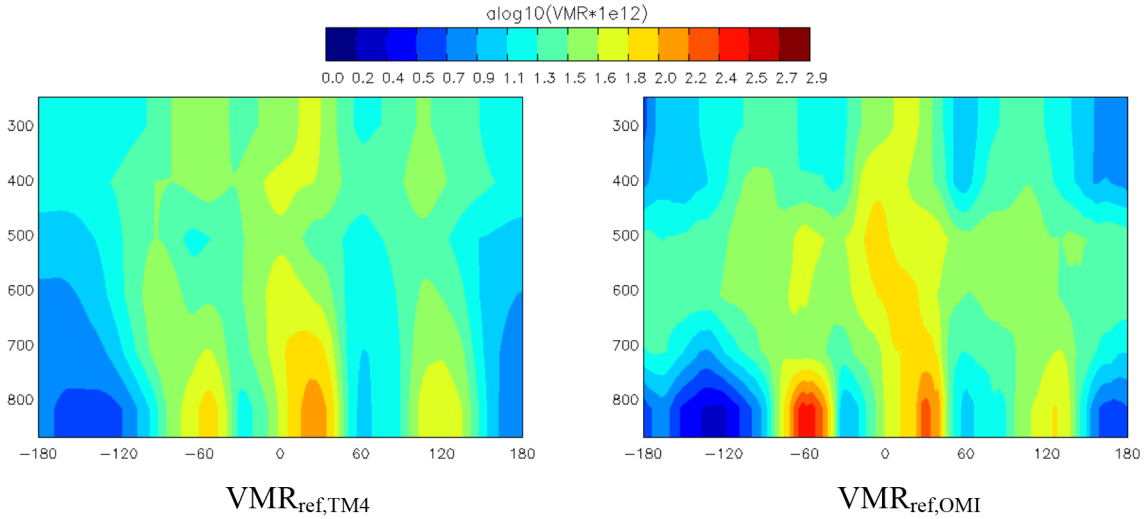


Figure 11. Longitudinal cross-section of annual mean tropospheric NO_2 from TM4 (left) and OMI (right) with CRF>50% over the tropics (10N-20S).

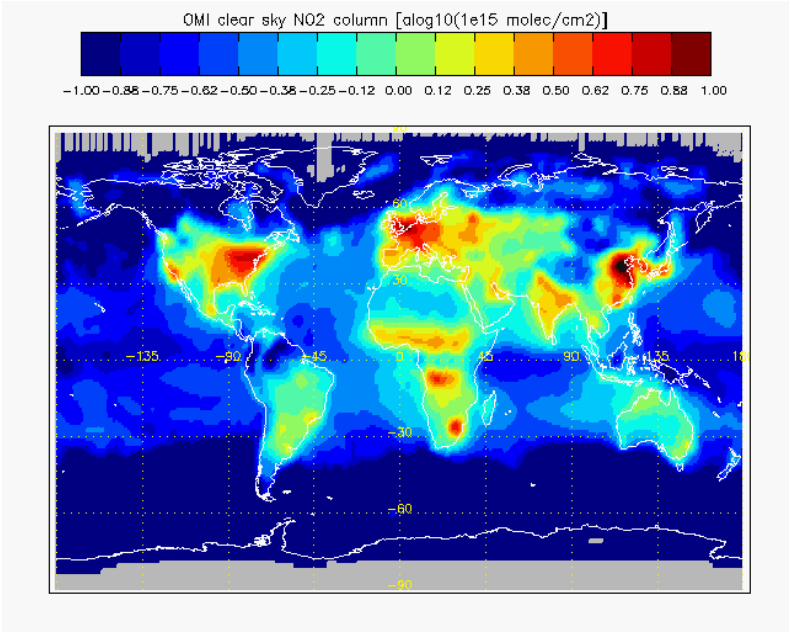


Figure 12. Annual clear sky OMI tropospheric NO_2 total columns for the year 2006.

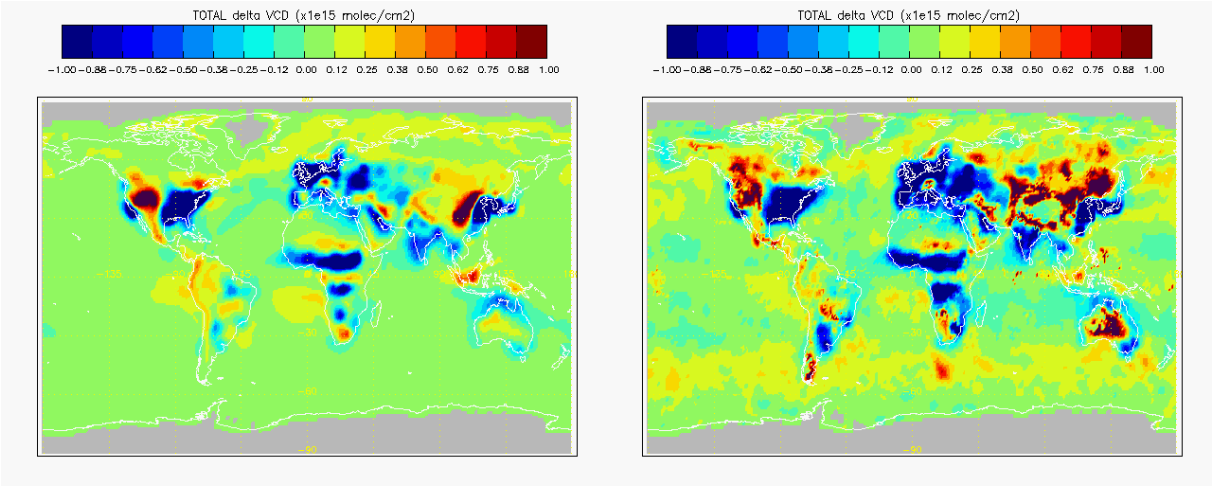


Figure 13. Total tropospheric NO_2 columns differences between cloudy (CRF>50%) and clear (CRF<25%) conditions for TM4 (left) and OMI (right).

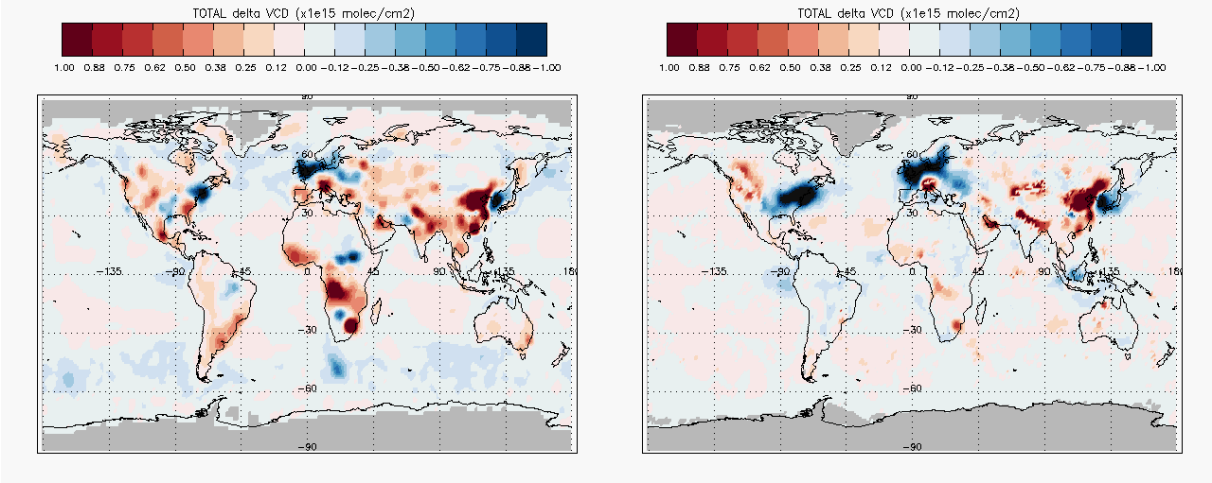


Figure 14. Total tropospheric NO_2 column differences (OMI-TM4) in clear (left) and cloudy (right) conditions for the year 2006.

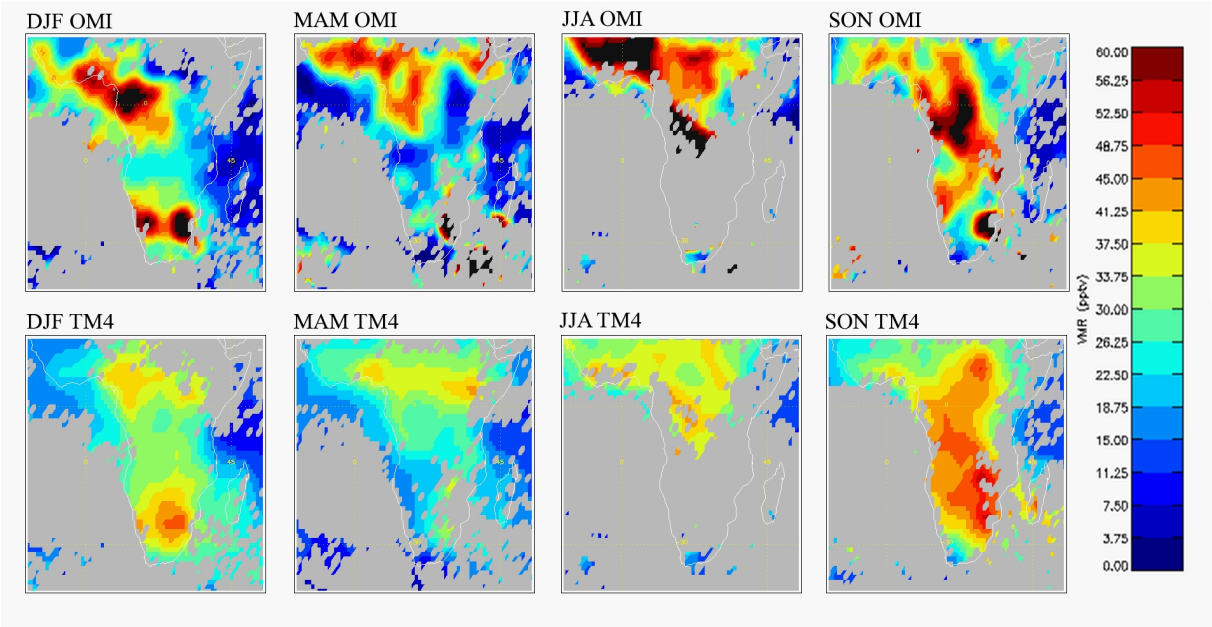


Figure 15. African sector at 280 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

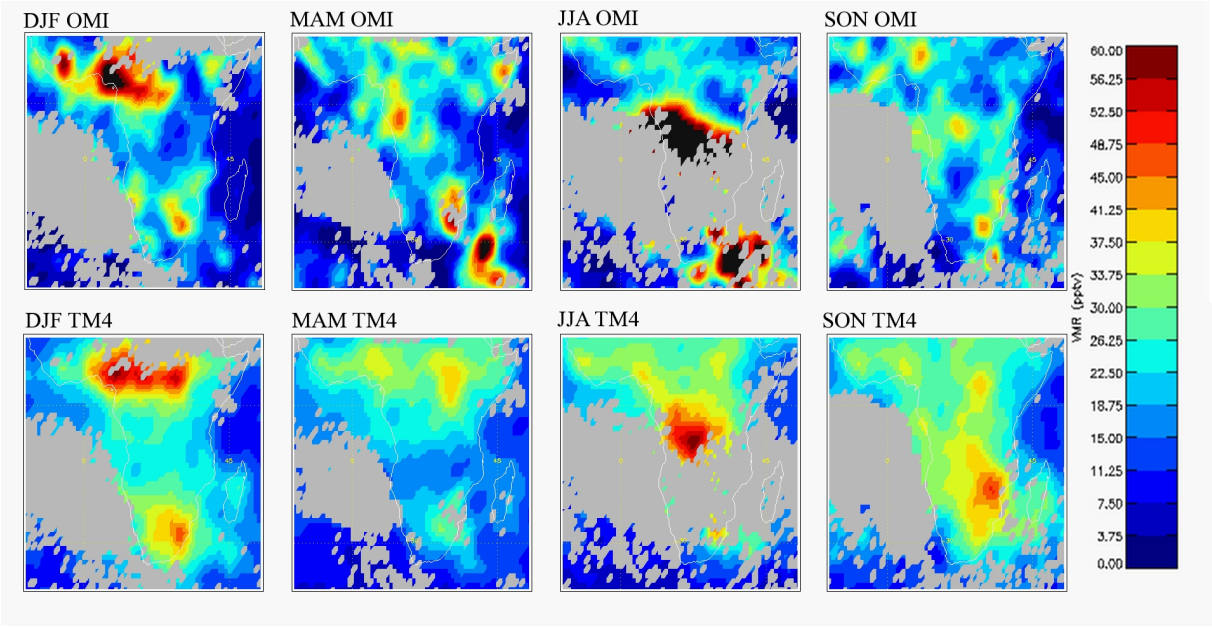


Figure 16. African sector at 380 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

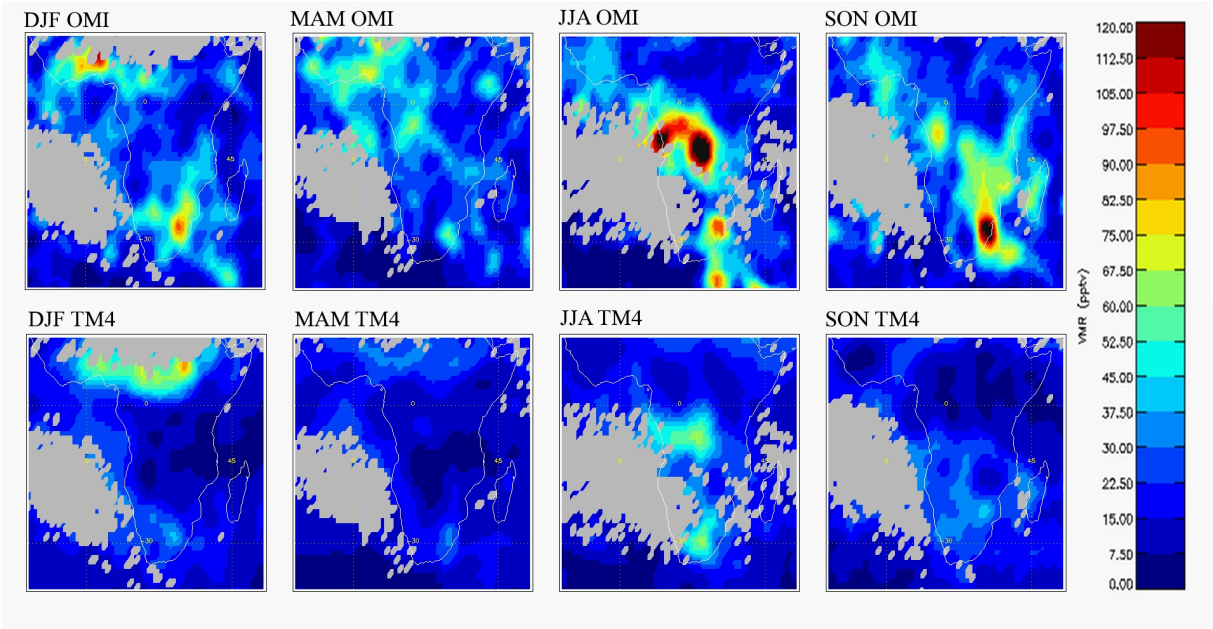


Figure 17. African sector at 500 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

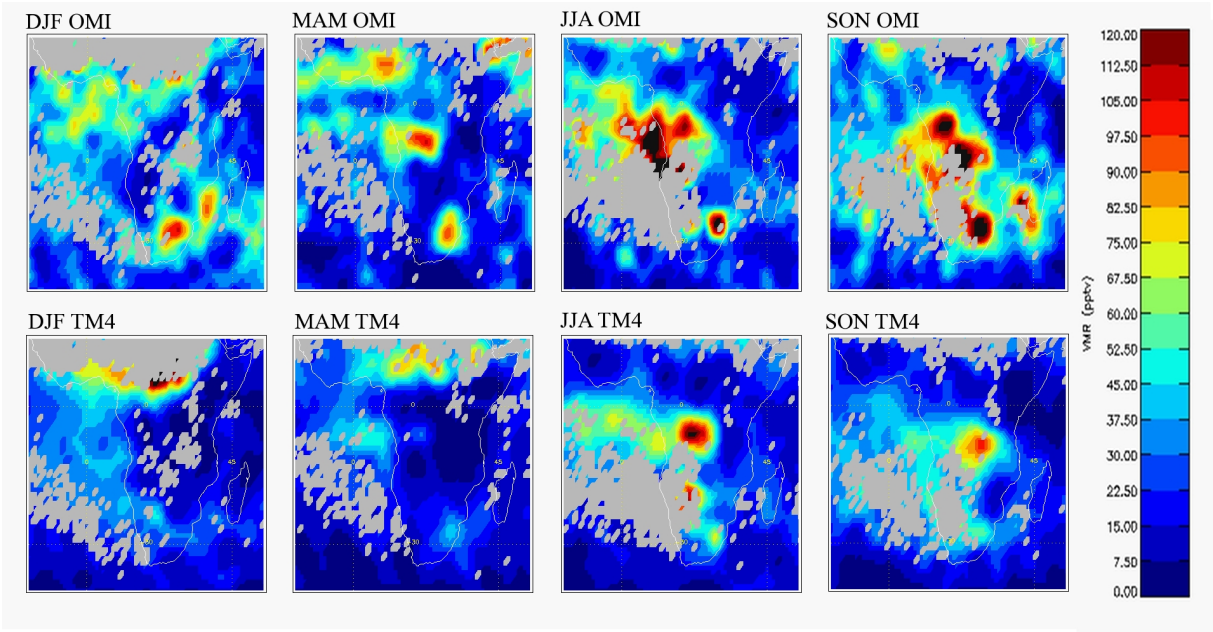


Figure 18. African sector at 620 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

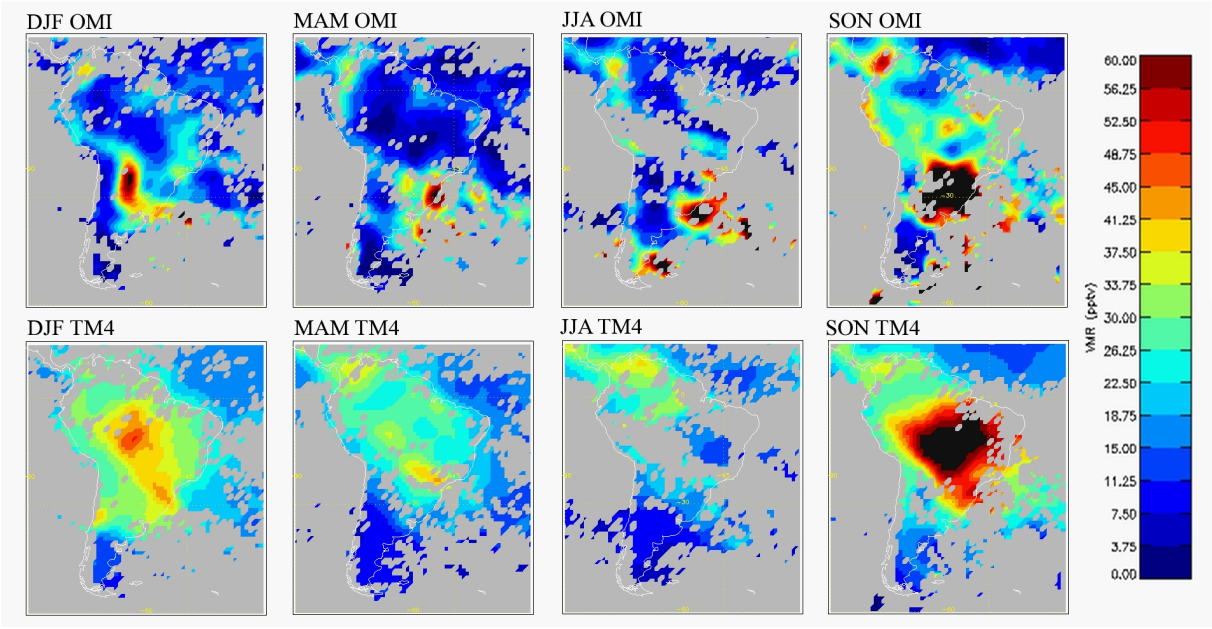


Figure 19. South American sector at 280 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

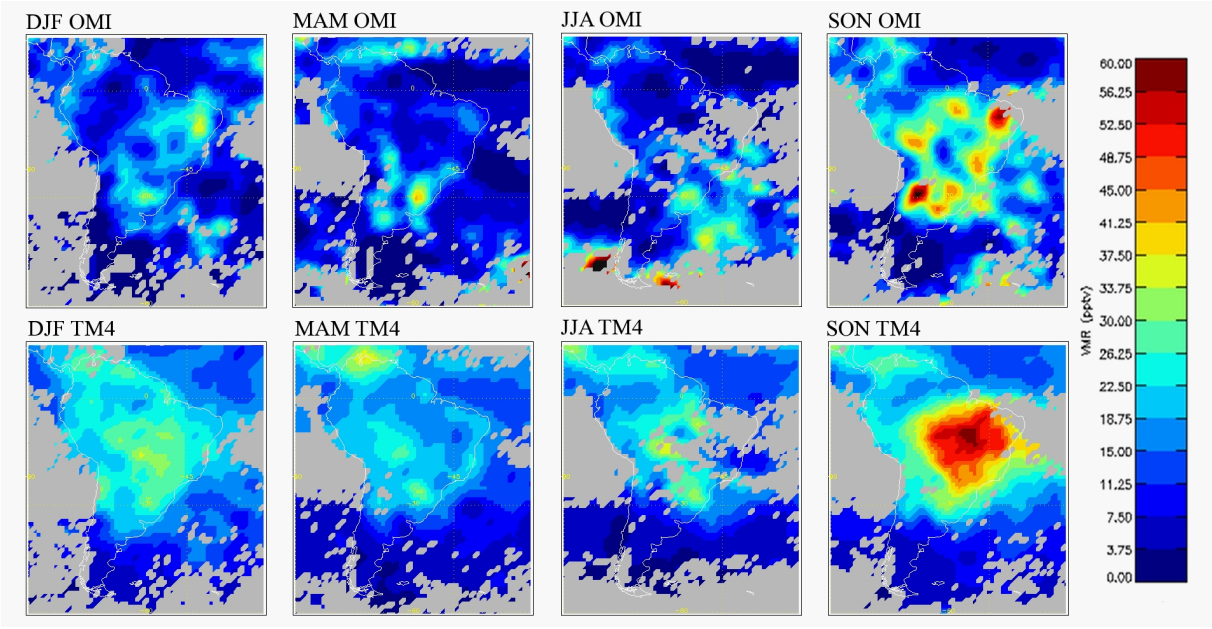


Figure 20. South American sector at 380 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

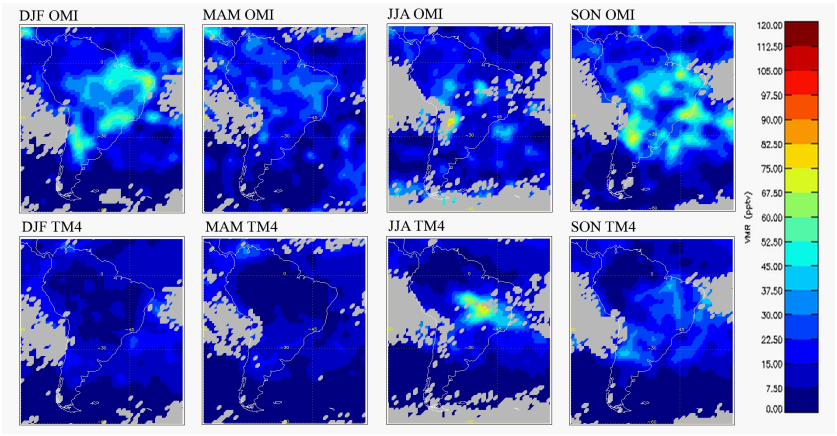


Figure 21. South American sector at 500 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

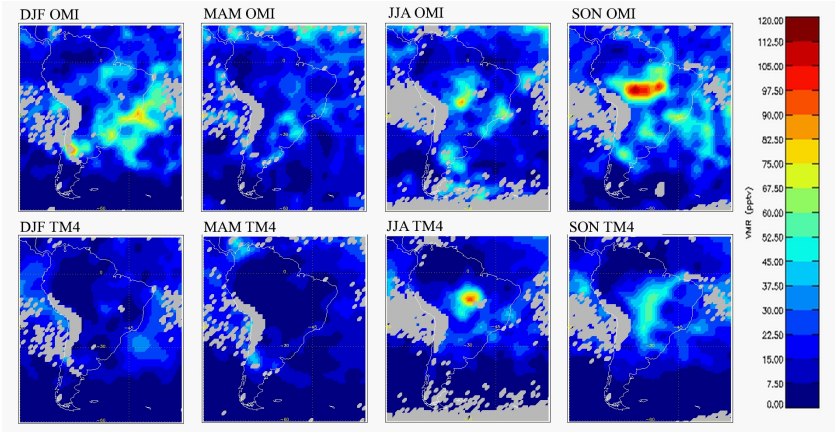


Figure 22. South American sector at 620 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

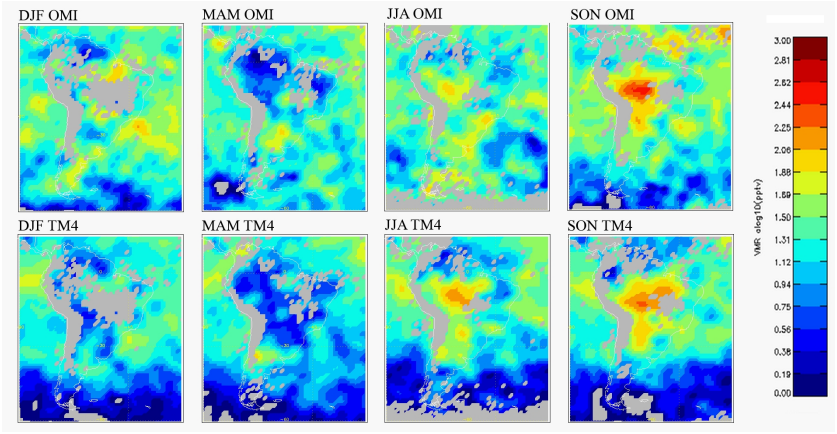


Figure 23. South American sector at 720 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

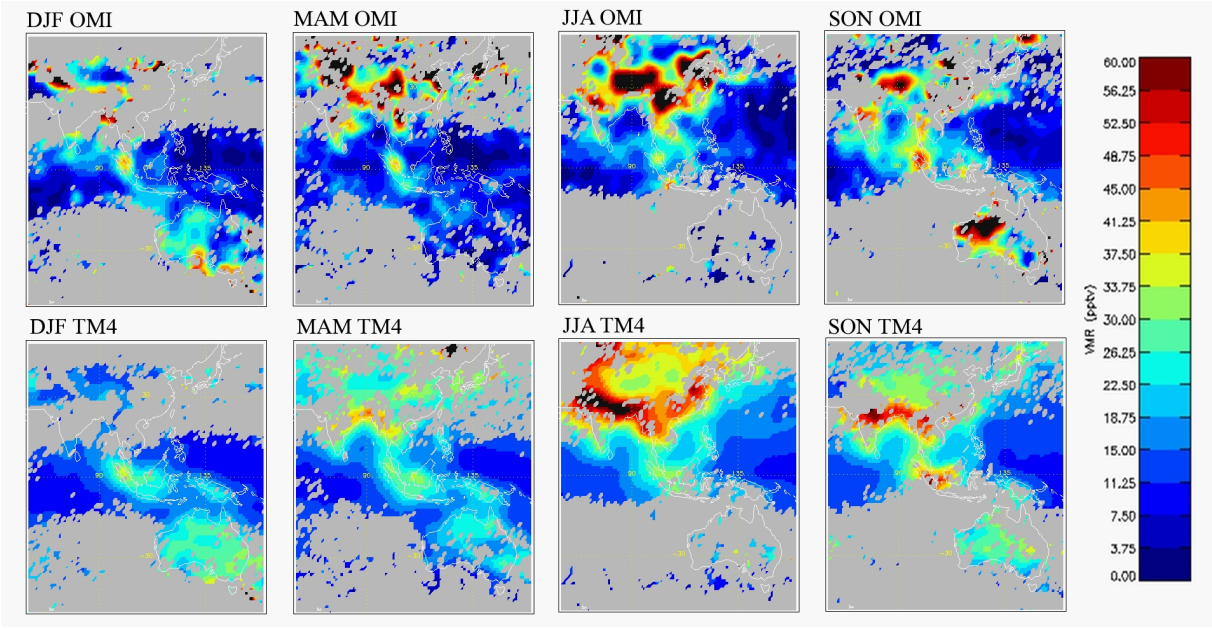


Figure 24. Asian sector at 280 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

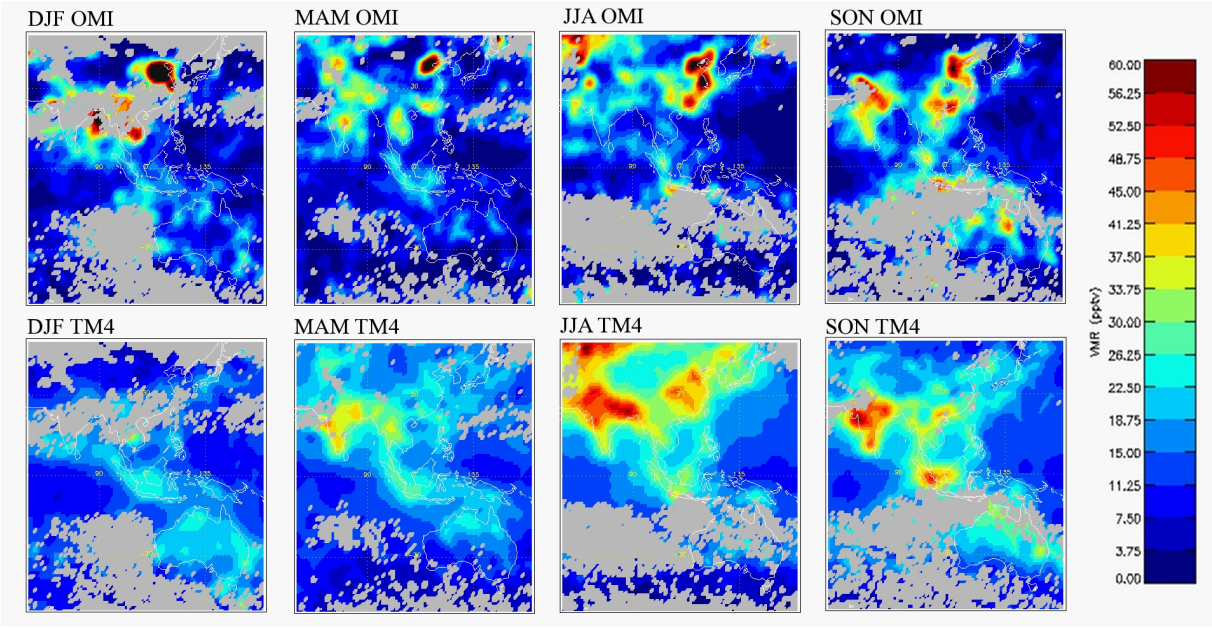


Figure 25. Asian sector at 380 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

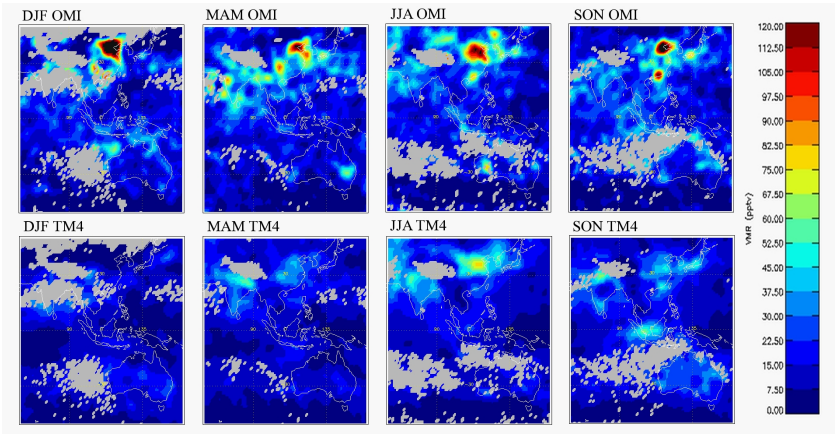


Figure 26. Asian sector at 500 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

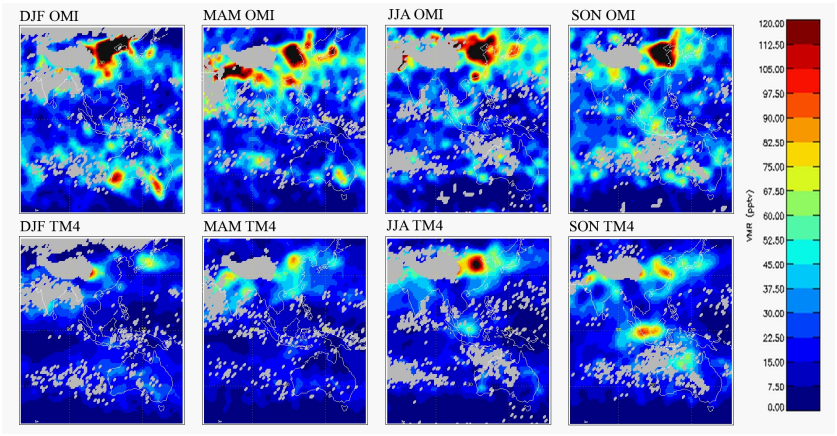


Figure 27. Asian sector at 620 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

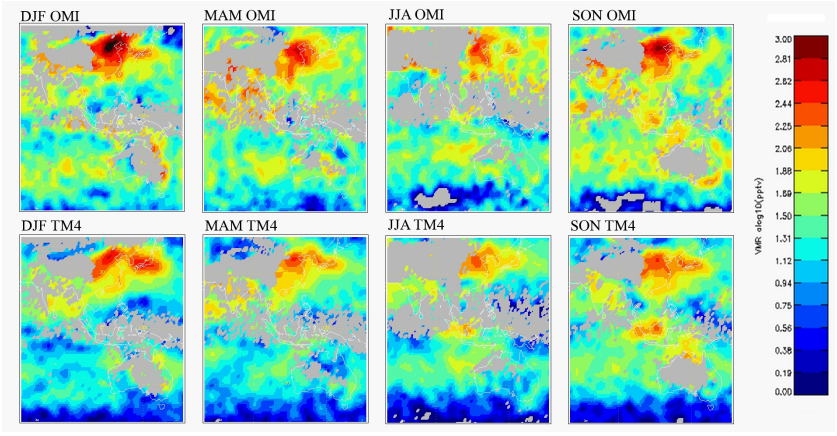


Figure 28. Asian sector at 720 hPa: Seasonal variability in OMI (top row) versus TM4 model (bottom row) average NO_2 VMRs pseudoprofiles for the year 2006.

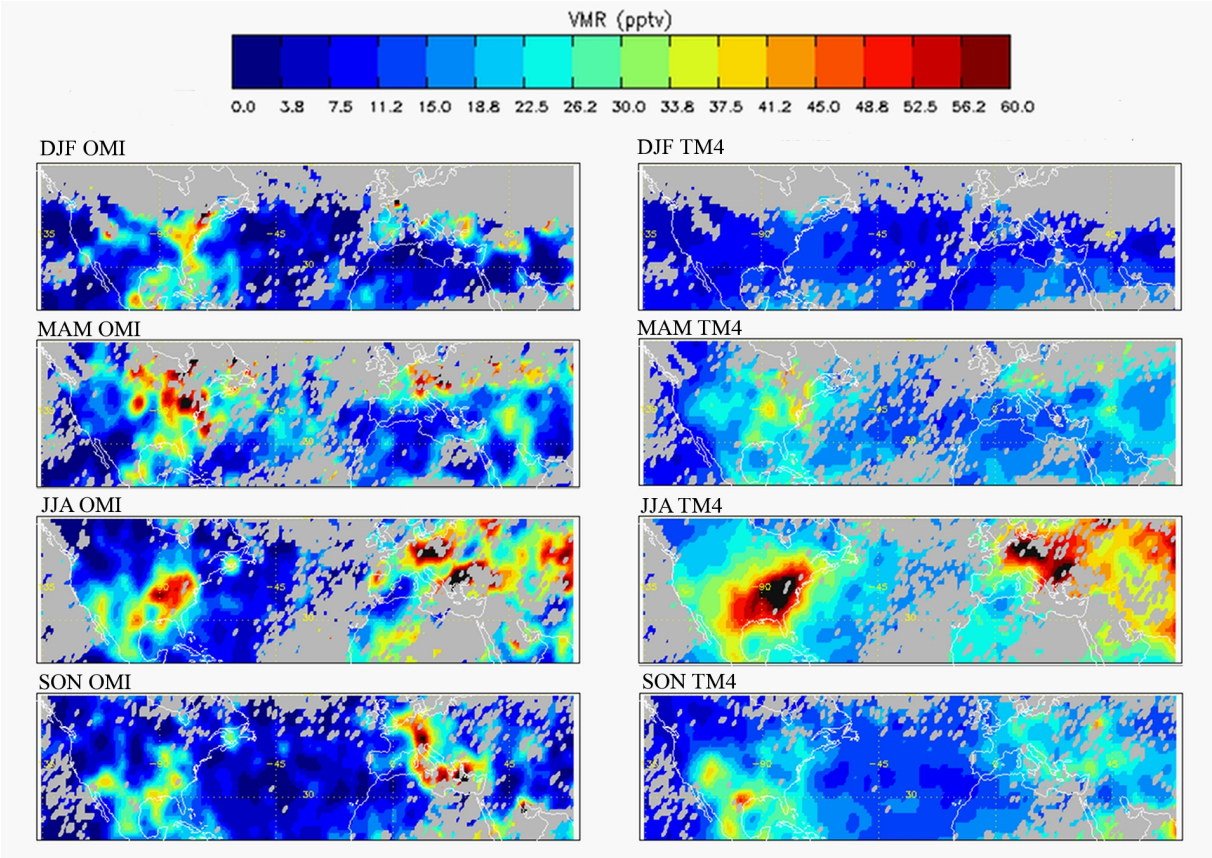


Figure 29. North American and European sector at 380 hPa: Seasonal variability in OMI (left column) versus TM4 model (right column) average NO_2 VMRs pseudoprofiles for 2006.

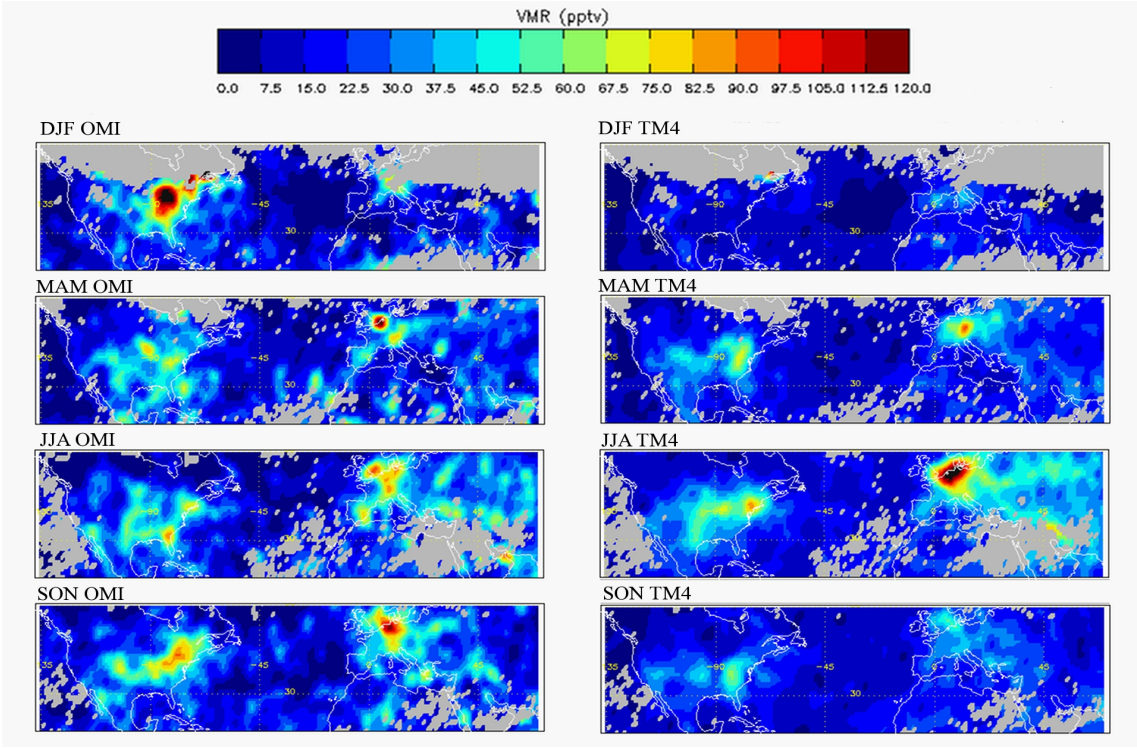


Figure 30. North American and European sector at 500 hPa: Seasonal variability in OMI (left column) versus TM4 model (right column) average NO_2 VMRs pseudoprofiles for 2006.

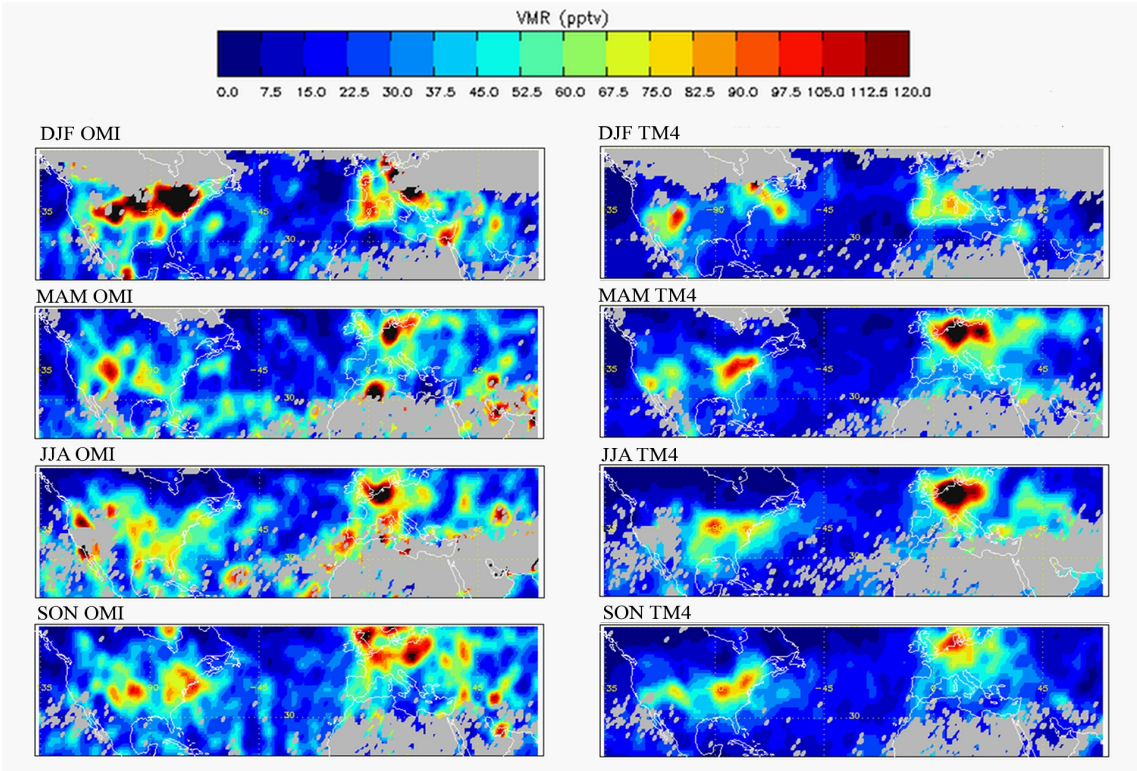


Figure 31. North American and European sector at 620 hPa: Seasonal variability in OMI (left column) versus TM4 model (right column) average NO_2 VMRs pseudoprofiles for 2006.

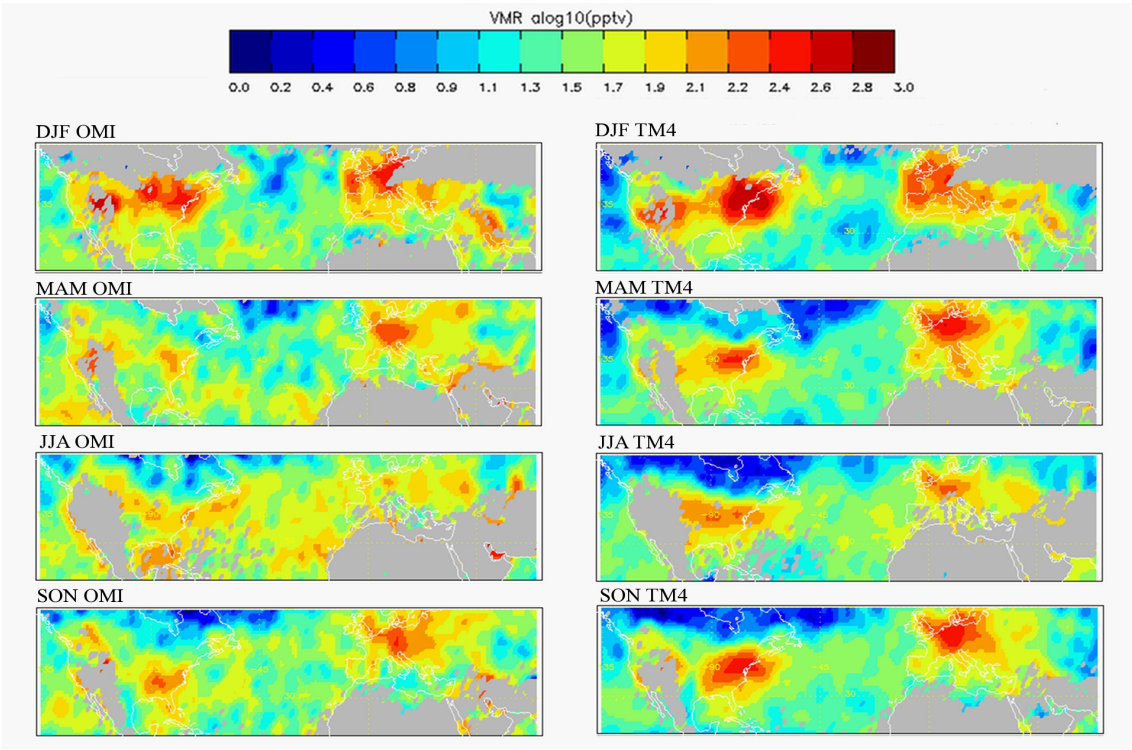


Figure 32. North American and European sector at 720 hPa: Seasonal variability in OMI (left column) versus TM4 model (right column) average NO_2 VMRs pseudoprofiles for 2006.

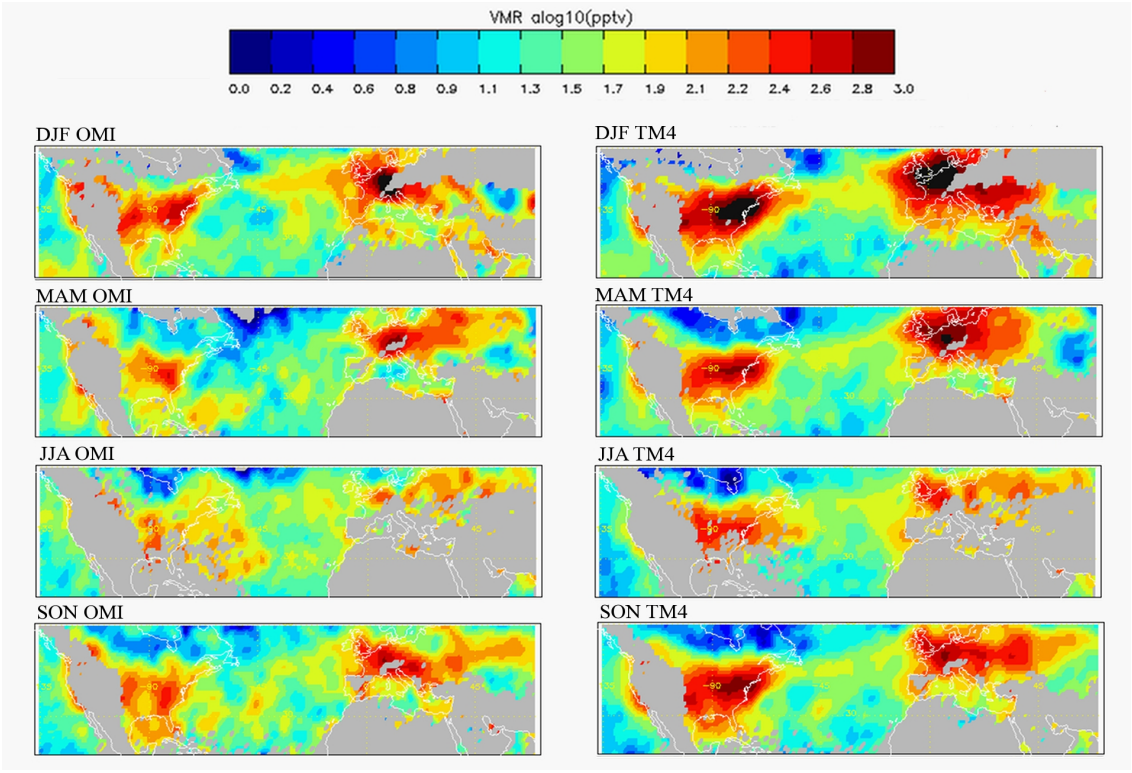


Figure 33. North American and European sector at 820 hPa: Seasonal variability in OMI (left column) versus TM4 model (right column) average NO_2 VMRs pseudoprofiles for 2006.