TECHNICAL MEMORANDUM

671

The MACC reanalysis: An 8-year data set of atmospheric composition

A. Inness[1] , F. Baier[2], A. Benedetti[1], I. Bouarar[3], S. Chabrillat[4], H. Clark[15,16], C. Clerbaux[3], P. Coheur[14], R. J. Engelen[1], Q. Errera[4], J. Flemming[1], M. George[3], C. Granier[3,6,7], J. Hadji-Lazaro[3], V. Huijnen[8], D. Hurtmans[14], L. Jones[1], J. W. Kaiser[1,17,18], J. Kapsomenakis[12], K. Lefever[4], J. Leitão[9], M. Razinger[1], A. Richter[10], M. G. Schultz[11], A. J. Simmons[1], M. Suttie[1], O. Stein[11], J.-N. Thépaut[1], V. Thouret[5], M. Vrekoussis[12,13], C. Zerefos[12] and the MACC team

Research Department

October 2012

Submitted to ACP

[1]ECMWF, Reading, UK, [2]DLR, Oberpfaffenhofen, Germany
[3]UPMC Univ. Paris 06; Université Versailles St-Quentin; CNRS/INSU, LATMOS-IPSL, Paris, France, [4]BIRA, Brussels, Belgium, [5]CNRS, Paris, France
[6]NOAA /ESRL and CIRES, University of Colorado, Boulder, CO, USA
[7]Max Planck Institute for Meteorology, Hamburg, Germany
[8]KNMI, De Bilt, The Netherlands, [9]European Commission, DG JRC, Ispra, Italy
[10]IUP-Bremen, Germany, [10]FZ Jülich, Germany, [12]Academy of Athens, Greece
[13]EEWRC, The Cyprus Institute, Cyprus
[14] Spectroscopie de l'Atmosphère, Service de Chimie Quantique et Photophysique, Université Libre de Bruxelles (ULB), Brussels, Belgium
[15] Université de Toulouse, UPS, LA (Laboratoire d'Aérologie), Toulouse, France
[16] CNRS, LA (Laboratoire d'Aérologie), UMR 5560, 31400 Toulouse, France
[17]King's College London, London, UK, [18]Max-Planck-Institute for Chemistry, Mainz, Germany

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

Abstract

An eight-year long reanalysis of atmospheric composition data covering the period 2003-2010 was constructed as part of the FP7 funded Monitoring Atmospheric Composition and Climate project by assimilating satellite data into a global model and data assimilation system. This reanalysis provides fields of chemically reactive gases, namely carbon monoxide, ozone, nitrogen oxides, and formaldehyde, as well as aerosols and greenhouse gases globally at a resolution of about 80 km for both the troposphere and the stratosphere. This paper describes the assimilation system for the reactive gases and presents validation results for the reactive gases analysis fields to document the dataset and to give a first indication of its quality.

Tropospheric CO values from the MACC reanalysis are on average 10-20% lower than routine observations from commercial aircrafts over airports through most of the troposphere, and have larger negative biases in the boundary layer at urban sites affected by air pollution, possibly due to an underestimation of CO or precursor emissions

Stratospheric ozone fields from the MACC reanalysis agree with ozone sondes and ACE data to within ±10 % in most situations. In the troposphere the reanalysis shows biases of -5% to +10% with respect to ozone sondes and aircraft data in the extratropics, but has larger negative biases in the tropics. Area averaged total column ozone agrees with ozone fields from a multi sensor reanalysis data set to within a few percent.

NO_2 fields from the reanalysis show the right seasonality over polluted urban areas of the NH and over tropical biomass burning areas, but underestimate wintertime NO_2 maxima over anthropogenic pollution regions and overestimate NO_2 in Northern and Southern Africa during the tropical biomass burning seasons.

Tropospheric HCHO is well simulated in the MACC reanalysis even though no satellite data are assimilated. It shows good agreement with independent SCIAMACHY retrievals over regions dominated by biogenic emissions with some anthropogenic input, such as the Eastern US and China, and also over African regions influenced by biogenic sources and biomass burning.

1 Introduction

MACC (Monitoring Atmospheric Composition and Climate) is a research project with the aim of establishing the core global and regional atmospheric environmental services for the European GMES (Global Monitoring for Environment and Security) initiative. The project was funded from 1 June 2009 to 31 December 2011 under the Seventh Framework Programme of the European Union. MACC built on the predecessor projects GEMS (Global and regional Earth-system Monitoring using Satellite and in-situ data; Hollingsworth et al. 2008) and PROMOTE (http://www.gse-promote.org/). The project combined state-of-the-art atmospheric modelling with Earth observation data to provide information services covering European Air Quality, Global Atmospheric Composition, Climate, and UV and Solar Energy. The global model and data assimilation system used in MACC was based on ECMWF's Integrated Forecast System (IFS). In GEMS, IFS had been extended to include chemically reactive gases (Flemming et al. 2009; Inness et al. 2009), aerosols (Morcrette et al. 2009; Benedetti et al. 2008) and greenhouse gases (Engelen et al. 2009), so that ECMWF's 4D-VAR data assimilation system could be used to assimilate satellite observations of atmospheric composition at global scale. Chemical transport models (CTMs) were coupled to the IFS using the OASIS-4 coupler (Valcke and Redler 2006) to provide initial fields and chemical production and loss rates for the reactive gases (Flemming et al. 2009).

MACC generated data records of atmospheric composition for recent years, data for monitoring present conditions, and forecasts of the distribution of key constituents for a few days ahead. As part

of MACC an eight-year long reanalysis over the period 2003-2010 of atmospheric composition data was constructed. The MACC reanalysis built on the experience gained by producing a reanalysis of atmospheric composition as part of the GEMS project. MACC used a newer model than the one used in GEMS, and benefited from the assimilation of more and reprocessed satellite data and from having a higher horizontal resolution (80 km instead of 125 km as in GEMS). The period 2003-2010 was chosen based on consideration of the available satellite data of atmospheric composition. Reactive gases were calculated with a system configuration where the CTM MOZART-3 (Kinnison et al. 2007; Stein et al. 2009) was coupled to the IFS (Stein et al. 2012).

Assimilation of satellite data on atmospheric composition with focus on stratospheric ozone has been carried out for over a decade (Hólm et al. 1999; Khattatov et al., 2000; Dethof and Hólm 2004; Geer et al. 2006; Lahoz et al. 2007; Arellano et al. 2007; Dragani 2010 and 2011), and global ozone forecasts are now produced routinely by several meteorological centres. ECMWF, for example, produces daily ozone analyses and forecasts, KNMI uses the TM3-DAM system to produce operational ozone forecasts and analyses (Eskes et al. 2002), and NCEP assimilates several ozone products into its operational Global Forecast System (http://www.cpc.ncep.noaa.gov/products/stratosphere/strat_a_f/). The Belgian Institute for Space Aeronomy has the Belgian Assimilation System for Chemical Observations (BASCOE, Errera et al. 2008), while the German Aerospace Centre applies the SACADA model (Elbern et al. 2009). Both 4D-Var systems are dedicated to the assimilation of stratospheric chemical observations making use of explicit chemistry. Data assimilation is now also increasingly being used for other chemical trace gases in both global and regional model systems (see overview articles by Sandu and Chai 2011, Zhang et al. 2011, Baklanov et al. 2008), and data assimilation code has been implemented in several chemical transport models (e.g. GEOS-Chem, Bey et al. 2001; WRF-Chem, Grell et al. 2005 and http://ruc.fsl.noaa.gov/wrf/WG11/). Assimilation of tropospheric constituents, however, is still in its infancy.

While several centres have produced meteorological reanalyses, for example NCEP (Kalnay et al. 1996), ECMWF (Gibson et al., 1997; Uppala et al., 2005, Dee et al., 2011), JMA (Onogi et al., 2007) and NASA/DAO (Schubert et al., 1993) there has been less activity with respect to reanalyses of atmospheric composition. ECMWF included the assimilation of ozone data in several of its reanalysis projects, and reanalyzed ozone fields are available from ERA-40 (Dethof and Hólm 2004) and ERA-Interim (Dragani 2010 and 2011). At KNMI a 30-year long ozone dataset was produced from a multi sensor reanalysis (Van der A et al. 2010).

MACC was in a position to combine a wealth of atmospheric composition data with a numerical model and data assimilation system to produce a reanalysis of atmospheric composition. This paper describes the setup of the reactive gases data assimilation system used in the MACC reanalysis of atmospheric composition. The reactive gases that were included as IFS model variables in the MACC reanalysis were ozone (O_3), carbon monoxide (CO), nitrogen oxides (NO_x = NO + NO_2), and formaldehyde (HCHO). These four gases were chosen because they play a key role in the chemistry of the atmosphere and have been measured by space-borne instruments with sufficient density and continuity to deliver strongly constrained analyses.

Carbon monoxide has natural and anthropogenic sources (Seinfeld and Pandis, 2006; Kanakidou and Crutzen, 1999). It is emitted from the soil, plants and the ocean, but its main sources are incomplete fossil fuel and biomass burning, which leads to enhanced surface concentrations. Another important

source of CO is the oxidation of anthropogenic and biogenic volatile organic compounds (VOCs). In areas with large biogenic emissions (e.g. tropical rain forests), oxidation of biogenic VOCs contributes strongly to the production of CO (Griffin et al. 2007). Hudman et al. (2008) found that over the Eastern US during summer the biogenic sources of CO were higher than the anthropogenic ones due to decreasing anthropogenic emissions. The highest CO concentrations are found over the industrial regions of Europe, Asia and North America. Surface concentrations are higher during the winter than during the summer months because of the shorter lifetime in the summer due to higher OH concentrations and more intense mixing processes. Tropical biomass burning is most intense during the dry season (December-April in the Northern Hemisphere (NH) tropics, July-October in the Southern Hemisphere (SH) tropics). CO has a lifetime of several weeks and can serve as a tracer for regional and inter-continental transport of polluted air. The main loss process is the reaction with the hydroxyl radical (OH) radical.

Ozone is an important species for chemistry of the troposphere. Tropospheric ozone is a regional scale pollutant and, at high concentrations near the surface, it is harmful to humans and vegetation. Photolysis of ozone, followed by reaction with water vapour, provides the primary source of the hydroxyl radical. Ozone is also a significant greenhouse gas, particularly in the upper troposphere (Hansen et al. 1997). The majority of tropospheric ozone formation occurs when NO_x, CO, and VOCs react in the atmosphere in the presence of sunlight. In urban areas in the NH high ozone levels usually occur during spring and summer. About 90% of the total ozone amount resides in the stratosphere, a result of oxygen photolysis as first discovered by Chapman. This ozone layer absorbs a large part of the sun's harmful UV radiation. Anthropogenic chlorofluorocarbons led to a global decrease of the ozone total column, with potentially catastrophic consequences avoided thanks to the Montreal Protocol (Newman *et al.*, 2009). Over Antarctica ozone destruction during Austral spring still leads to strong and rapid depletion of the ozone layer ("ozone hole"). Stratospheric ozone destruction happens also on a smaller scale over the Arctic in boreal spring (Manney *et al.*, 2011).

Nitrogen oxides play a key role in tropospheric chemistry and are the main ingredient in the formation of ground level ozone. Their sources are anthropogenic emissions, biomass burning, soil emissions and, at altitude, lightning and aviation. NO_x has a lifetime of a few days in the free troposphere and less in the boundary layer, so that concentrations are larger over land than over the cleaner oceans. The largest concentrations are found over industrial and urban regions of the Eastern US, California, Europe, China and Japan. Loss processes for NO_x are the formation reactions of with OH to HNO_3, with O_3 to NO_3 at night and formation of PAN as well as dry deposition.

Formaldehyde is one of the most abundant hydrocarbons in the atmosphere. Even though its primary emission sources are industrial activities, fossil fuel burning, and biomass burning, the largest contribution to the HCHO budget is its secondary source from the oxidation of VOCs, in particular isoprene (Atkinson 1994; Palmer et al. 2003, 2006; Abbot et al. 2003; Millet et al. 2008). The main source of HCHO in the background troposphere is the oxidation of methane, which accounts for more than half of the global HCHO production (Stavrakou et al. 2009). In the continental boundary layer the oxidation of non-methane VOCs dominates. The main sinks of HCHO are photolysis and oxidation by OH. HCHO has a short lifetime of a few hours, making it a good indicator of hydrocarbon emission areas. While the data quality of the individual satellite retrievals of HCHO was not sufficient to allow

active assimilation in the MACC reanalysis, passive monitoring was performed by the assimilation system.

In this paper we describe results for the fields of CO, O_3, NO_x and HCHO. The paper is structured in the following way: Section 2 describes the coupled IFS GRG system and the data assimilation setup for the reactive gases. This includes information about the data assimilation system, aspects of the coupling between the IFS and the CTM, and information about satellite and emission data that were used in the reanalysis. Section 3 shows results from the reanalysis and comparisons with independent observations, and Section 4 presents the conclusions.

2 Description of the MACC chemical data assimilation system

2.1 Model system

The MACC data assimilation system for chemically reactive gases was constructed by extending ECMWF's integrated forecast system (IFS) to include fields for O_3, CO, NO_x, and HCHO. Source and sink terms for these gases are supplied by a CTM that is coupled to the IFS using the OASIS4 coupler (Valcke and Redler 2006) and which holds a detailed representation of the atmospheric chemical system together with its sources and sinks. For the MACC reanalysis the MOZART-3 CTM (115 species) which features a full description of stratospheric and tropospheric chemical processes was coupled to the IFS. A description of the MOZART-3 CTM as implemented in the MACC system can be found in Stein (2009) and Stein et al. (2012). In the coupled setup the IFS and the CTM run in parallel, exchanging fields through the OASIS4 coupler every hour (Flemming et al. 2009). This means the IFS supplies the meteorological data and updated mixing ratios for the MACC global reactive gases (GRG) species O_3, CO, NO_x, and HCHO to the CTM, and the CTM provides IFS with initial conditions for the four GRG species and with chemical tendency fields every hour. These are tendencies due to chemistry, wet deposition and atmospheric emissions, and tendencies due to surface fluxes (emission, dry deposition). The tendencies for the individual species are combined before the exchange and one total tendency per species is given from the CTM to the IFS.

In the MACC reanalysis the IFS has a spectral truncation of T255 and the grid point space is represented in a reduced Gaussian grid (Hortal and Simmons 1991) of about 80 km horizontal resolution. The vertical coordinate system is given by 60 hybrid sigma-pressure levels, with a model top at 0.1 hPa. In order to avoid difficulties in the vertical interpolation by the OASIS4 coupler, the CTMs use the same 60 vertical levels. The coupler only has to perform horizontal interpolations for which the bi-linear mode is applied. The MOZART-3 resolution is 1.125° x 1.125°, lower than the IFS resolution, because of the high computational cost of the CTM. The IFS is run on a higher horizontal resolution than the CTM as this improves the quality of the meteorological forecasts and because a lower resolution would limit the acceptance of high resolution observations within the data assimilation. More details of the CTMs and the coupling setup are given in Flemming et al. (2009). A modification of MOZART-3 as described in Kinnison et al. (2007) was used in the MACC reanalysis from 2003-2008 (Version 3.1). From 1 January 2009 onwards MOZART Version 3.5 was used in the MACC reanalysis, and this implementation is described in Stein et al. (2012). The updated version has a better representation of ozone depletion inside the Antarctic vortex (Flemming et al. 2011).

Ozone had already been included in the IFS as an additional model variable, and ozone data have been assimilated at ECMWF since 1999 (Hólm et al. 1999, Dethof and Hólm 2004). However, the ECMWF approach differs from the MACC approach because it uses a built-in chemistry routine with a parameterization of photochemical sources and sinks based on Cariolle and Teyssèdre (2007) instead of a coupled CTM to provide the chemical tendencies. Moreover, this relatively simple chemical scheme is only suited for the description of stratospheric ozone, while the MOZART-3 CTM represents the whole tropospheric and stratospheric chemistry.

2.2 Data assimilation

ECMWF's IFS has used an incremental formulation of 4-dimensional variational data assimilation (4D-Var) since 1997. In 4D-Var a cost function is minimized to combine the model background and the observations to obtain the best possible forecast through the length of the assimilation window by adjusting the initial conditions. The GRG species are integrated into the ECMWF variational analysis as additional model variables. They are minimized together with the other ECMWF fields, which means they can, in principle, influence the analysis of wind and other meteorological variables in 4D-Var. However, given the uncertainty of the GRG observations and the lack of observational constraints of variables such as wind or temperature in the stratosphere and mesosphere, a possible influence of the GRG observations on the meteorological fields was suppressed in the reanalysis. Nevertheless, this might be a worthwile interaction to study in the future (Semane et al. 2009).

2.2.1 Observation operators for reactive gases

Observation operators are needed to calculate the model equivalent of the assimilated observations, i.e. of the satellite retrievals of atmospheric composition. The observations used in the IFS are total or partial column data, i.e. integrated layers bounded by a top and a bottom pressure. The model's background column value is calculated as a simple vertical integral between the top and the bottom pressure given by the partial or total column, at the time and location of the observation.

It is also possible to use averaging kernel information in the observation operator. This removes the impact of the retrieval a-priori profile in the assimilation (Eskes and Boersma, 2003). Equation (1) shows how the retrieved quantity $\hat{x}_r$ can be described as a linear combination of the a-priori profile x_a and the true profile x_t by using the averaging kernel matrix A (Deeter et al. 2009). The averaging kernels indicate the sensitivity of the retrieved profile to the true profile, with the remainder of the information coming from the a-priori profile (Rodgers 2000; Emmons et al. 2004; Emmons et al. 2007). Where the averaging kernel matrix elements are small, the retrieved profile shows little sensitivity to changes in the true profile and the retrieved quantity is largely determined by the a-priori.

$$\hat{x}_r \approx x_a + A\left(x_t - x_a\right) \tag{1}$$

Equation (1) can be used in the observation operator to transform the model field to have the same vertical resolution and a-priori dependence as the satellite retrievals. Thus, the difference between the retrieved quantity $\hat{x}_r$ and its model equivalent $\hat{x}_m$ can be given as

$$\hat{x}_r - \hat{x}_m = A\left(x_t - x_m\right) \tag{2}$$

This means that the model field and the retrieved quantity can be compared in a way that is not affected by the a-priori profile dependence or by different vertical resolutions. Averaging kernels were used in the MACC reanalysis if they were provided by the data producers (more details will be given in Section 2.3).

NO2 observation operator

The fast diurnal NO_2 - NO interconversion caused by solar radiation can not be handled by the coupled model with an exchange frequency for the chemical tendencies of one hour, and the absence of a full chemistry module in the IFS prevents direct assimilation of short-lived chemical species. Therefore, NO_x is used as the IFS model variable instead of NO_2. Its longer chemical lifetime allows both a better simulation by the coupled forward model and a correct assimilation in the adjoint model. The use of NO_x also reduces spatial variability everywhere which is of advantage for the data assimilation. Since the satellite observations assimilated in the MACC system are NO_2 data, a diagnostic NO_2/NO_x interconversion operator was developed, including its tangent linear and adjoint. This operator is based on a simple photochemical equilibrium between the NO_2 photolysis rate JNO_2 and the ozone mixing ratio:

$$\frac{[NO_2]}{[NOx]} \approx \frac{k[O_{3eff}]}{JNO_2 + k[O_{3eff}]} \tag{3}$$

Here k is the rate coefficient of the reaction $O_3 + NO \rightarrow NO_2 + O_2$ and depends on temperature, while JNO_2 depends on surface albedo, solar zenith angle, overhead ozone column, cloud optical properties and temperature. A parameterized approach for the calculation of JNO_2 was used based on the band scheme by Landgraf and Crutzen (1998) in combination with actinic fluxes parameterized following Krol and Van Weele (1997). In the stratosphere O_{3eff} is equal to O_3, but in the troposphere it is equal to

$$[O_{3eff}] = [O_3] + [XO_2]\cos\alpha \tag{4}$$

to account for the influence of per-oxy-radicals (XO_2=HO_2 + RO_2). In an ad-hoc approach a per-oxy-radical concentration of 80 ppt in the troposphere was assumed (Kleinman et al. 1995), which was scaled by the cosine of the solar zenith angle α to account for the diurnal cycle of the per-oxy-radical concentration. This improved the match of the NO_2/NO_x ratios from the operator and the MOZART-3 fields (Flemming et al., 2011a).

2.2.2 Observation errors for the reactive gases

The observation error and background error covariance matrices determine the relative weight given to the observation and the background in the analysis. For the reactive gases, observation errors given by the data providers were used. If these values were below 5%, a minimum value of 5% was taken. The observation error was assumed to include any observation operator error, and a representativeness error that could arise because of differences in resolution of observation and the model, and that accounts for scales unresolved by the model. The satellite data were thinned in the data pre-processing to ensure a minimum distance between two observations from the same platform. This was done to reduce the data volume and helped to avoid redundant observations that did not contain any independent information. It also avoided the introduction of spatial observation error correlation that was not accounted for in the data assimilation algorithm. In the MACC reanalysis, the reactive gases

satellite retrievals were thinned to a horizontal resolution of 1° x 1° by randomly selecting an observation in the grid box.

Variational quality control (Andersson and Järvinen, 1999) and background quality checks were applied to the reactive gases observations. In the background quality check, the square of the normalized background departure was considered as suspect when it exceeded its expected variance by more than a predefined multiple. In this case the observation was not used in the analysis.

2.2.3 Background errors for the reactive gases

In the ECMWF data assimilation system the background error covariance matrix is given in a wavelet formulation (Fisher 2004, 2006). This allows both spatial and spectral variations of the horizontal and vertical background error covariances. The background error standard deviations determine the relative weight of the background in the analysis, while the correlations determine how the analysis increments are spread in the horizontal and in the vertical. This is particularly important for vertically integrated observations, such as total column trace gas retrievals. In this case the vertical structure of the increments is determined by the vertical correlations of the background errors since the observations do not give information about this distribution.

The background error correlations used in the operational ECMWF data assimilation system were derived from an ensemble of forecast differences, using a method proposed by Fisher and Andersson (2001). This ensemble consisted of ten members, all run for one month. For the MACC ozone field the same background error statistics used in ECMWF's operational ozone assimilation were used. A different method had to be chosen to determine background error statistics for the other GRG fields because they had not been included in the ensemble of forecast runs. The National Meteorological Center (NMC) method (Parrish and Derber 1992) was used to derive initial background error statistics for the reactive gases. For this, 150 days of 2-day forecasts were run with the coupled system initialized from fields produced by the free running MOZART-3 CTM, and the differences between 24-h and 48-h forecasts valid at the same time were used as a proxy for the background errors. These differences were then used to construct a wavelet background error covariance matrix according to the method described by Fisher (2004, 2006). This background error covariance matrix contains the statistics for the reactive gases as well as the original statistics for the other meteorological fields. Background errors determined with the NMC method usually have longer horizontal and vertical correlations than those calculated with the analysis ensemble method (Fisher and Andersson 2001).

For the assimilation of NO_x data it was found that the analysis based on mixing ratio was prone to large extrapolation errors, due to the large range of NO_x concentrations which make it difficult to model the background error covariances. Therefore a logarithmic control variable was developed for NO_x.

The GRG background errors are univariate in order to minimize the feedback effects of the GRG fields on the other variables. Examples of the GRG background error standard deviation profiles and correlations used in the MACC reanalysis can be found in Inness et al. (2009).

2.3 Satellite data

2.3.1 Satellite data used in the reanalysis

Table 1 lists the data sets that are actively assimilated in the MACC reanalysis to constrain the reactive gases. These contain profile (PROF), total column (TC), partial column (PC) and tropospheric column (TRC) data. The usage criteria for the datasets are also given in the table. In addition to the listed criteria, data were not used if quality flags given by the data producers mark the data as bad quality. HCHO data were not assimilated in the MACC reanalysis because the data quality of individual satellite retrievals was not sufficient, but analysis fields are available. Monthly mean HCHO observations generally have a total error of 20-40%, but individual observations can have large errors (greater than 50% for individual SCIAMACHY or GOME pixels, De Smedt et al. 2008), which did not favour assimilating individual observations. Hence, the HCHO reanalysis fields were entirely determined by the MOZART-3 chemistry, the MACCity and biomass burning emissions (see section 2.4 below), and the atmospheric transport.

Averaging kernels were used in the observation operators if they were provided by the data producers. This was the case for MOPITT CO, IASI CO and SCIAMACHY NO2. Figure 1 shows averaging kernels from MOPITT CO and IASI CO day time total column retrievals averaged over a 5° x 25° box over Europe in July 2009, and global mean averaging kernels for SCIAMACHY NO2 averaged over the period June 2009 to May 2010. MOPITT and IASI have the main sensitivity to CO in the mid troposphere, i.e. between 300-700 hPa. SCIAMACHY has broad averaging kernels for NO2, indicating some sensitivity to all tropospheric levels. The global mean profile peaks at 300 hPa. Even though the main sensitivity is to NO2 above 700 hPa, as the vertical profile of NO2 in polluted regions is dominated by the lowest layers, the retrieved column (which is the product of NO2 profile times averaging kernel) is mainly determined by NO2 in the boundary layer, at least for cloud free situations. The sensitivity at higher altitudes is larger, but as result of the low NO2 concentrations usually found there, this has only small effect on the NO2 columns retrieved from satellite data.

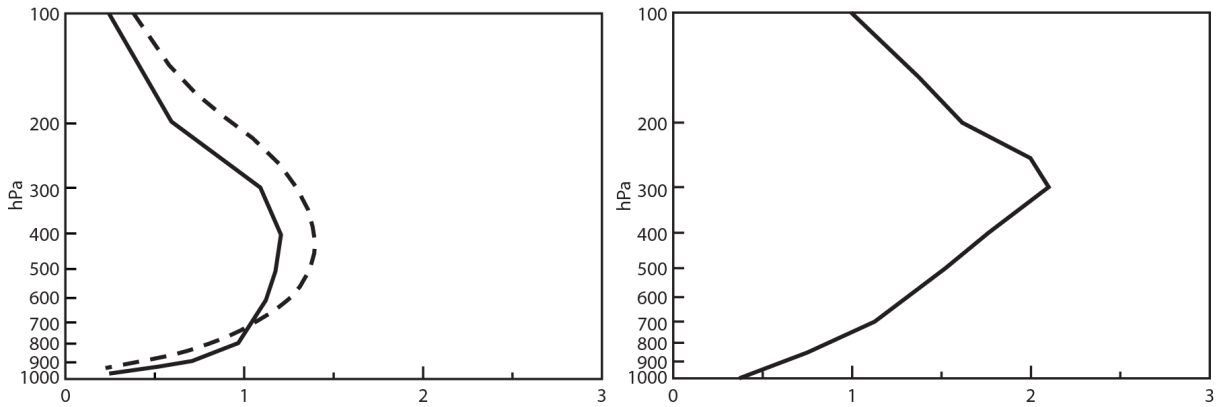

Figure 1: Left panel: Mean averaging kernels for MOPITT (solid) and IASI (dashed) day time total column CO retrievals for July 2009 averaged over a 5° x 25° box over Europe (46-51°N, 3-28°E). Right panel: Global mean averaging kernels for tropospheric column NO_2 from SCIAMACHY averaged over the period June 2009 to May 2010.

Sensor	Satellite	Provider	Version	Period	Type	Data usage criteria	Reference
GOME	ERS-2	RAL		20030101-20030531	O_3 PROF	Used if SOE>15° and 80°S<lat<80°N	Siddans et al. 2002
MIPAS	ENVISAT	ESA		20030127-20040326	O_3 PROF	All data used	Carli et al. 2004
MLS	AURA	NASA	V02	20040808-20090315, NRT data from 20090316	O_3 PROF	All data used	Waters et al. 2006
OMI	AURA	NASA	V003	From 20041001, NRT data 20070321-20071231	O_3 TC	Used if SOE >10°	Bhartia et al. 2002; Levelt et al. 2006
SBUV/2	NOAA-16	NOAA	V8	From 20040101	O_3 PC	Used if SOE>6°	Bhartia et al. 1996
SBUV/2	NOAA-17	NOAA	V8	From 20030101	O_3 PC	Used if SOE>6°	Bhartia et al. 1996
SBUV/2	NOAA-18	NOAA	V8	From 20050604	O_3 PC	Used if SOE>6°	Bhartia et al. 1996
SCIAMACHY	ENVISAT	KNMI		From 20030101	O_3 TC	Used if SOE>6°	Eskes et al. 2005
IASI	METOP-A	LATMOS/ULB		From 20080401	CO TC	Used if 70°S<lat<70°N	George et al. 2009, Clerbaux et al. 2009
MOPITT	TERRA	NCAR	V4	From 20030101, NRT data after 20100323	CO TC	Used if 65°S<lat<65°N	Deeter et al. 2010
SCIAMACHY	ENVISAT	KNMI	V1.04	20030101-20070630	NO_2 TRC	Used if SOE>6° and 60°S<lat<60°N	Boersma et al., 2004
SCIAMACHY	ENVISAT	KNMI	V1.1	From 20070911	NO_2 TRC	Used if SOE>6° and 60°S<lat<60°N	http://www.temis.nl, Wang et al. 2008

Table 1: Satellite retrievals of reactive gases that were actively assimilated in the MACC reanalysis. PROF denotes profile data, TC total columns, TRC tropospheric columns, PC partial columns, and SOE solar elevation. PC SBUV/2 data consist of 6 layers between the surface and 0.1 hPa. NRT (near-real time) data are available within a few hours after the observation was made, and are being used in operational forecast systems. For periods towards the end of the MACC reanalysis period, NRT data were used for some of the species when no offline products were available.

2.3.2 Bias correction and quality control for the reactive gases satellite data

To improve the assimilation of atmospheric composition data, the variational bias correction scheme (VarBC) developed at ECMWF for radiance data (Dee 2004, McNally et al. 2006, Auligné et al. 2007, Dee and Uppala 2009) was extended to data of atmospheric composition. Without applying a bias correction scheme to the data, the assimilation would either have to be limited to one retrieval product for a reactive gas, or data would be used that are inconsistent with each other or with the reactive gases forecast. Biases in the variational scheme are estimated during the analysis by including bias parameters in the control vector. The bias corrections are continuously adjusted to optimize the consistency with all information used in the analysis. Extending VarBC to atmospheric composition data has the advantage that it can be easily applied to a variety of sensors and species.

For the reactive gases in the MACC reanalysis a start from zero bias was used at the beginning of the experiment. The reanalysis was started on 1 December 2002, so that a month-long spin-up of the bias correction and the fields was possible. Solar elevation and a global constant were used as bias predictors, and SBUV/2 data (from various NOAA platforms) and MLS ozone profiles from 1 January 2008 onwards were used as anchor for the bias correction, i.e. no bias correction was applied to these data. Using anchors for the bias correction should help to avoid drifts in the system. The reason for choosing SBUV/2 data as anchor was that the data were available for the whole reanalysis period, they had been reprocessed and inter-calibrated and should hence make a good anchor for the other ozone data sets. While experience at ECMWF had shown this to work well when total column ozone data were assimilated, it was found in the MACC reanalysis that SBUV/2 ozone data (whose lowest layer is between 16 hPa and the surface) could not stop the bias correction drifting for individual MLS layers that have a finer vertical resolution. This drift did not affect the total ozone column much, but it became noticeable in the troposphere and above 15 hPa after a few years (see section 3.2 below). Once the cause of the apparent ozone trend had been identified, it was decided to stop bias correcting MLS data, and from 1 January 2008 onwards, both MLS and SBUV/2 data were used as anchors, i.e. assimilated without bias correction. For the assimilation of CO retrievals a globally constant predictor was used for IASI data while MOPITT CO retrievals were used as an anchor. The latter was used as the anchor simply because these data were available from the beginning of the reanalysis, not because their data quality was assumed to be better than that of the IASI data. No bias correction was applied to NO_2 data, because retrievals from only one instrument (SCIAMACHY) were assimilated.

2.4 Emissions

For the MACC reanalysis an updated dataset of anthropogenic emissions (MACCity) was produced (Granier et al., 2011a). The emissions are injected into the surface layer in the CTM, which is about 10m deep, and quickly distributed within the boundary layer. The anthropogenic emissions for MACCity were developed as an extension of the historical ACCMIP emissions dataset (Lamarque et al., 2010) developed for the IPCC AR5 assessment. The ACCMIP dataset provides decadal emissions up to the year 2000. The 2000-2011 MACCity emissions were obtained by using the 2005 and 2010 emissions from the future scenarios called Representative Concentration Pathways (RCPs; Moss et al. 2010). For the MACCity emissions, scenario RCP 8.5 was chosen, since it includes information on regional emissions after 2000 (Van Vuuren et al., 2011; Riahi et al., 2011): a linear interpolation was then applied to obtain the yearly MACCity emissions. Ship emissions are based on Eyring et al. (2010) and a source-specific seasonality developed for the RETRO project (http://retro.enes.org/) was applied

to the emissions. Monthly average emissions were derived using the seasonal patterns developed within the RETRO project.

Biomass burning emission for the MACC reanalysis for the years 2003-2008 were generated from a preliminary version (v3.0) of the Global Fire Emissions Database (GFED, van der Werf et al. 2010) and Fire Radiative Power (FRP) observations by the MODIS satellite instruments (Justice et al. 2002). The Global Fire Assimilation System (GFAS) of MACC was used to create daily gridded FRP maps, with which the monthly GFED emission were redistributed amongst the days of each month (Kaiser et al. 2011). Thus the monthly budgets of GFEDv3.0 were maintained. In Central and South America, the carbon combustion rate is about 8% larger in GFEDv3.0 than in the final, published GFEDv3.1 dataset. In all other regions, it is smaller. The global average is smaller by 9%. The carbon combustion rate in GFEDv3.1 is in turn smaller than in GFEDv2 (van der Werf et al. 2006) in all regions but Boreal North America, Australia and the Middle East; the global average is reduced by 13%.

For the years 2009 and 2010, daily biomass burning emissions from MACC's GFAS, Version 1.0 (Kaiser et al. 2012), were used. These are based on MODIS FRP observations, have a horizontal resolution of 0.5x0.5 degrees and are now available from 2003 onwards, but were not available when the reanalysis was started. The global average carbon combustion rate is 8% larger than in GFEDv3.1 and thus about 18% larger than during the first 5 years of the reanalysis. (Due to slightly modified emission factors and burning patterns, the carbon monoxide emissions are only about 6% larger than in GFEDv3.1, though.) Overall, the GFAS Version 1.0 emissions are approximately 20% larger than the data used prior to 2009.

Biogenic emissions in the MOZART-3 CTM came from a recent update (Barkley, 2010) of MEGAN2 (Guenther et al. 2006, http://acd.ucar.edu/~guenther/MEGAN/MEGAN.htm) and were read into MOZART-3 as monthly surface flux fields without interannual variation. These data were for 2003 and have no interannual variability. The biogenic sources are in particular sensitive to temperature, and there might be inconsistencies between those of the MACC model and of the Goddard Earth Observing System used to generate the biogenic emissions. However, this effect is expected to be small. MOZART-3 also included several other natural emissions like NO_x from soils and oceanic emissions from various sources, including the POET inventory (Granier et al., 2005, available from the ECCAD database, http://eccad.sedoo.fr/) and accounted for in-situ production of NO_x by lightning.

After the MACC reanalysis had been started it became apparent that using the MACCity emission led to an underestimation of CO concentrations in the Northern Hemisphere compared to independent observations (Granier et al. 2011b, see also Section 3.1 below). This could be due to an underestimation of surface emissions, to an underestimation of the chemical production of CO from the oxidation of VOCs, or to some missing reaction in the CTM. It should be noted that low CO values are found by most of all CTMs regardless of the emission inventory used (e.g. Shindell et al. 2006; Kopacz et al. 2010; Fortems-Cheiney et al. 2011), and that the MACCity anthropogenic emissions are in the same range as the emissions provided by the few other emission inventories available for the post-2000 period (Granier et al., 2011a). MACCity CO emissions are much higher than the most recent inventory developed for 2005, i.e. EDGAR-v4. In the MACC project offline simulations with the MOZART-3 CTM were conducted to test the model sensitivity to different emission inventories. It was found that the model results improved when increasing anthropogenic CO and VOC emissions.

Further studies are being carried out to establish exactly why the MACCity emissions and other emissions inventories lead to low CO values in the coupled IFS-MOZART system.

2.5 Control run

It would have been computationally too expensive to produce a control analysis experiment that was identical to the MACC reanalysis, but did not actively assimilate observations of reactive gases. Instead, a MOZART-3 stand-alone run was carried out that applied the same settings (model code, resolution, emissions) as MOZART in the MACC reanalysis. The meteorological data for the stand alone run were taken from the reanalysis, but the control run had free running chemistry. The results from this control run can be used to detect the impact of the assimilation of GRG observations in the MACC-reanalysis. Since the meteorological input data were derived from interpolation of archived 6-hourly output from the MACC reanalysis, and not through hourly exchange as in the reanalysis, the standalone run was not a completely clean control run. However, these differences would be small.

3 Results

This section presents fields from the MACC reanalysis and compares them with observations. It provides a basic first assessment of the quality of the reactive gases reanalysis fields, and it highlights problems in the reanalysis that users should be aware of. Shown are time series and seasonal climatologies of the MACC analysis fields, as well as comparisons with independent observations where possible. Also shown are some results from the control run to highlight the impact of assimilating atmospheric composition data on the reanalysis fields. First, CO, O_3 and NO_x analysis fields are assessed for which observations were assimilated in the MACC reanalysis (see Table 1). Subsequently, results are presented for HCHO fields where no observations were assimilated in the MACC reanalysis.

3.1 CO analysis

MOPITT Version 4 CO retrievals were assimilated in the MACC reanalysis (see Table 1) from 2003 to 2010. IASI CO retrievals were assimilated from April 2008 onwards. Both instruments measure in the thermal infrared part of the spectrum and provide data during day and night. In the MACC reanalysis both day and night time data were assimilated. Because averaging kernels were used in the observation operators, differences in sensitivity to CO in the lower troposphere between day and night time observations (Deeter et al. 2003) were accounted for in the analysis.

Figure 2 shows time series of zonal mean total column CO data from the MACC reanalysis, the control run, MOPITT and IASI data for the period 2003-2010. The CO field shows a pronounced seasonal cycle in both hemispheres, as well as differences between the hemispheres. Larger anthropogenic emissions in the NH lead to larger CO values compared to the SH. The NH emissions peak in late winter/early spring because of increased fossil fuel burning for heating and increased power requirements (Edwards et al., 2004). This together with the seasonal cycle of OH which accounts for 90% of CO loss (Thompson, 1992) leads to maximum CO values in the NH in March and April. In the SH, the seasonal cycle is determined by a large contribution from biomass burning and some contribution from the oxidation of biogenic VOCs that lead to high CO values between the equator and 40°S with maximum values during September and October, the months of peak fire activity in the SH (Torres et al. 2010, Kaiser et al. 2012). The interannual variability is dominated by

changing emissions from biomass burning in both hemispheres (Kaiser et al. 2012) which is largely related to precipitation rates (Torres et al. 2010). In the SH, CO maxima are lower in 2003, 2008 and 2009 compared to other years. In October 2006 a CO maximum is visible just south of the equator. This is a result of the 2006 wildfires in Indonesia which led to the highest CO emissions over Indonesia in over a decade (Fortems-Cheiney et al. 2011; Kaiser et al. 2012). The lower values seen in the reanalysis in the NH from 2008 onwards are the result of assimilating IASI data in addition to MOPITT data.

The reanalysis agrees well with the assimilated MOPITT and IASI data. However, the control run with the free running MOZART-3 CTM underestimates CO concentrations. It starts from realistic initial conditions, but drifts towards lower CO concentrations within the first 6 months, though the seasonal cycle is well represented. The low bias of the control points to either a problem with the MACCity emissions or a short coming in the MOZART CTM, and is being investigated in other studies.

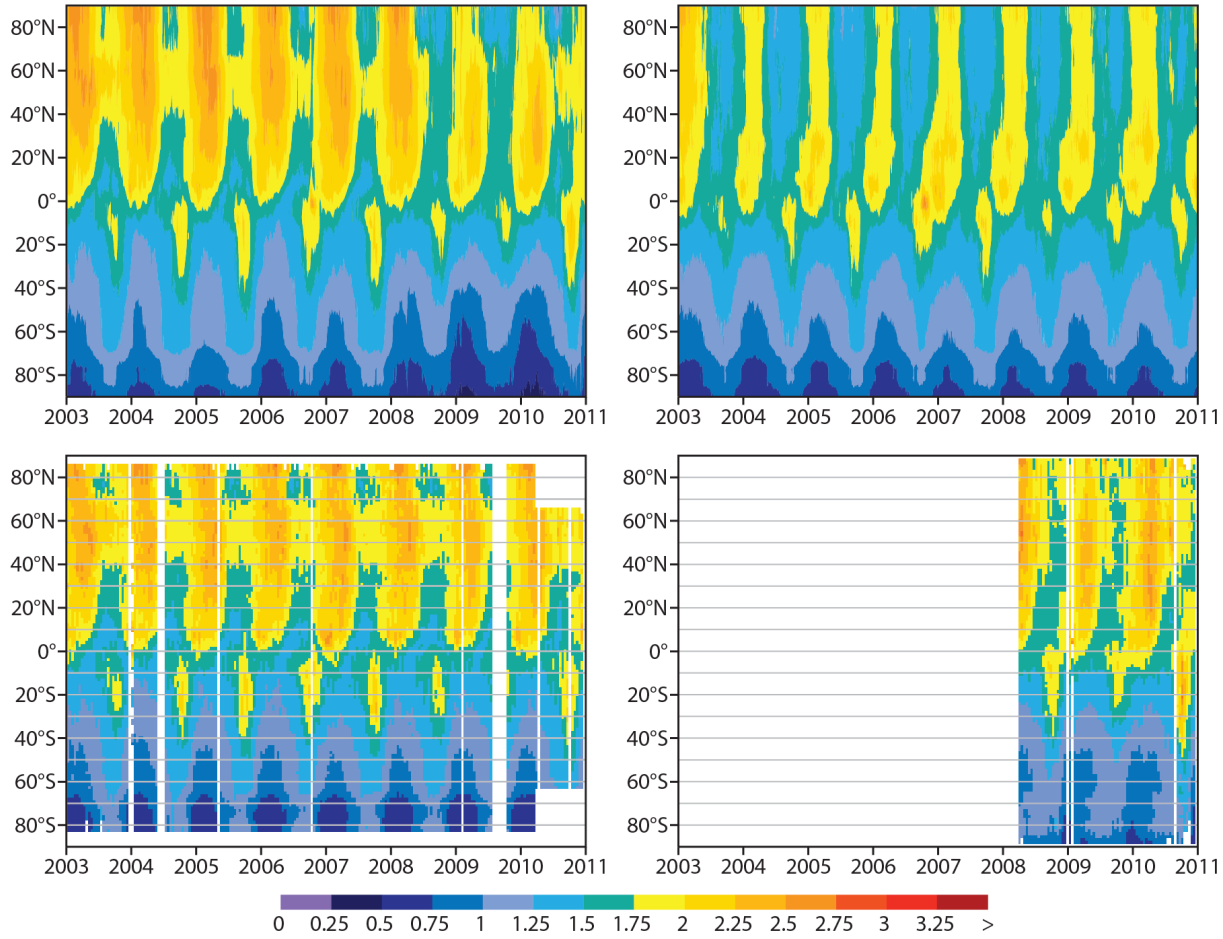

Figure 2: Time series of zonal mean total column CO field in 10^{18} molec/cm^2 for the period 2003 to 2010 from the MACC reanalysis (top left), the control run (top right), MOPITT (bottom left) and IASI (bottom right). The MOPITT time series shows the change to NRT data in 2010, for which no data are available polewards of 65°.

Figure 3 shows the reanalysis seasonal mean CO columns for the years 2003-2010. The figure again highlights the interhemispheric differences, with CO values generally larger in the NH than in the SH, except in the biomass burning regions in the tropics. This reflects the greater anthropogenic emissions in the NH (e.g. Fortems-Cheiney 2011). The largest values are found over South-East Asia in DJF and MAM, and there is transport of CO rich air from South-East Asia out into the Pacific. Values in the NH have a minimum in JJA and are still low in SON.

CO from biomass burning in the tropics shows a different seasonality. In Africa, maximum CO columns are seen north of the equator in DJF, when biomass burning takes place in the Sahel region and equatorial West Africa during the local dry season. In MAM the fire signal over Africa is much weaker, and by JJA the affected area moves south of the equator. In SON the signal is weaker than in JJA but extends further to the south and east. This agrees well with studies by Torres et al. (2010) who found that biomass burning in Central Africa normally starts in June at about 10°S and then moves southwards. It generally peaks in August, but the season extends to November. In all seasons, there is indication of CO rich air being transported from Africa downwind into the Atlantic. Convective activity over Africa (and also over South America) transports the smoke into the free troposphere from where it is then transported by the prevailing winds (e.g. Jonquières et al. 1998; Mari et al. 2008; Real et al. 2010).

In South America the strongest biomass burning signal is seen in SON. Deforestation fires and agricultural fires occur south of 10°S during August-October with peak in September. In Indonesia and North-West Australia the strongest biomass burning signal is also seen in SON, but the climatological biomass burning signal here is weaker than over Africa and South America. The large values over Indonesia in SON come from the strong fires in 2006 (see also Figure 2).

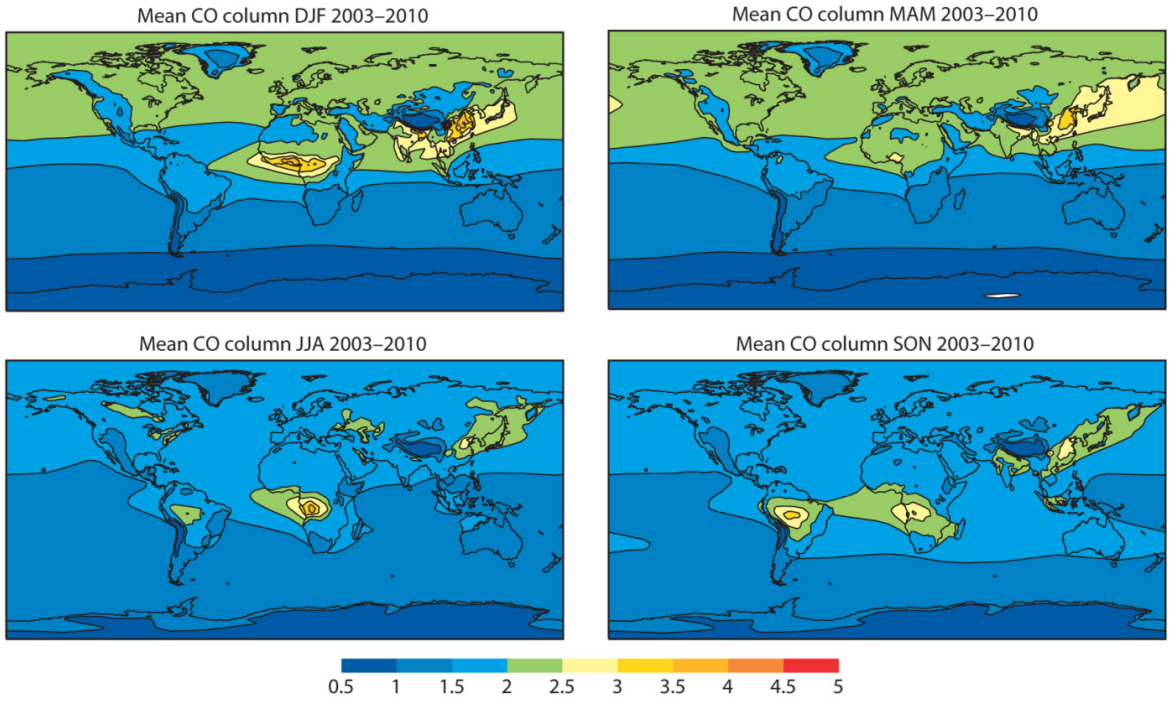

Figure 3: Seasonal mean CO total columns from the MACC reanalysis in 10^{18}molec/cm^2 for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010.

The vertical structure of the MACC CO field in the troposphere can be seen in the seasonal mean cross sections in Figure 4. In the NH the highest values are found in DJF and MAM when anthropogenic emissions are largest and the photochemical lifetime of CO is longest. There is some indication of large scale convective transport lifting CO into the upper troposphere in the tropics. Deep convection is known to carry biomass burning products into the upper troposphere (Pickering et al. 1996; Thompson et al. 1996; Kar et al. 2004). In the SH, high CO from biomass burning can be seen between 0-20° in JJA and SON, and again there is evidence of large scale transport lifting CO into the free troposphere.

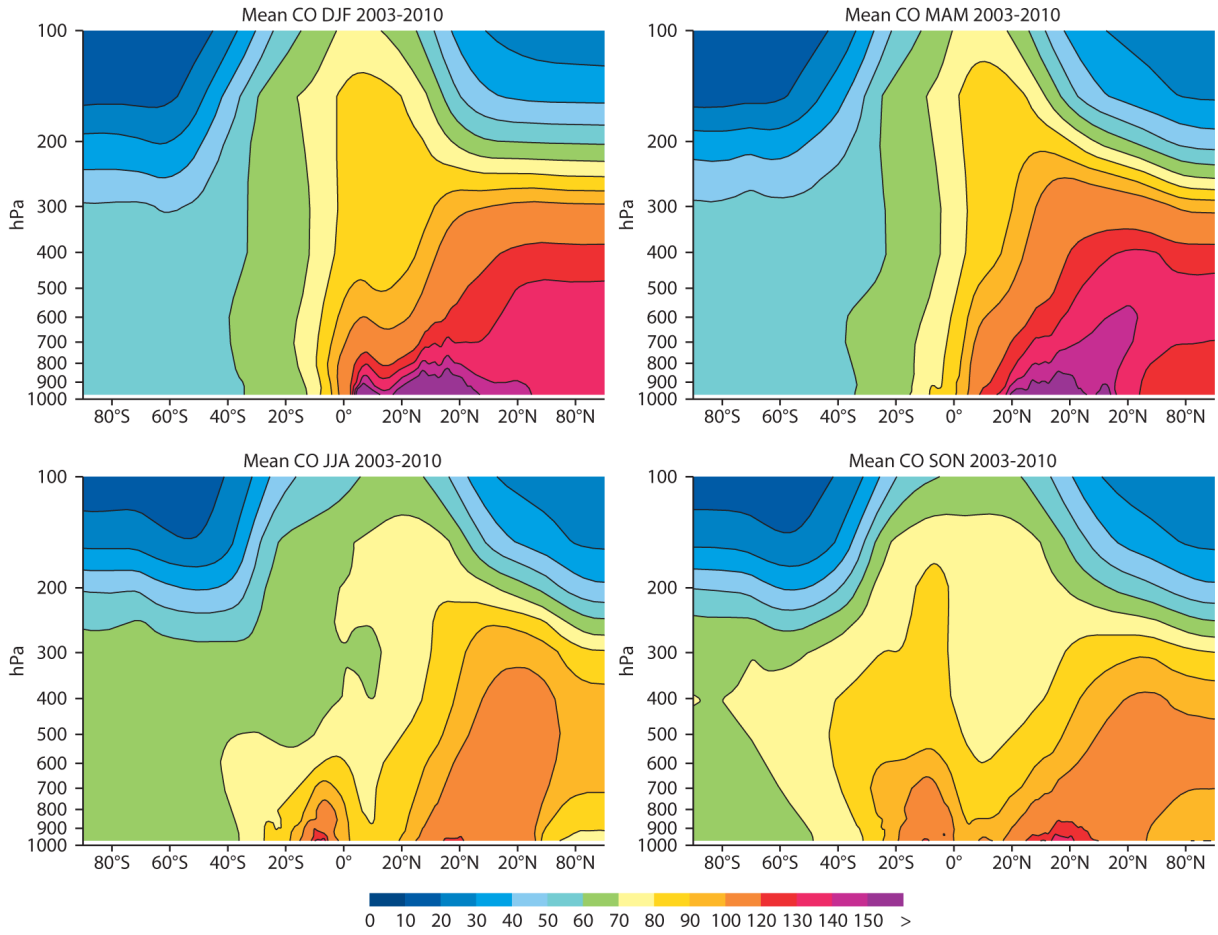

Figure 4: Seasonal mean zonal average CO altitude cross sections from the MACC reanalysis in ppbv for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010.

To validate CO from the reanalysis, fields are compared with independent observations, i.e. data that were not used in the assimilation. Figure 5 shows a comparison of CO concentrations from the reanalysis Tenerife, Key Biscayne and South Pole stations. The uncertainty of the NOAA/GMD CO observations is around 1-3 ppmv (Novelli et al., 2003). The results show that the magnitude and seasonal variability of surface CO is generally well captured by the reanalysis over most stations and improved compared to the control run, apart at South Pole from 2003-2007. This indicates that, despite low sensitivity of satellite measurements near the surface, assimilation of CO data from such products can lead to a good representation of surface CO concentrations. Assimilation of satellite

measurements can therefore overcome to some extent the underestimation of surface CO concentrations in the Northern Hemisphere seen in most global models (Isaksen et al. 2009; Holloway et al. 2000; Shindel et al. 2006) which could be related to uncertainties in surface emissions as mentioned in Section 2.4. Interestingly, the reanalysis agrees better with the observations over South Pole station after April 2008, i.e. after the assimilation of IASI CO data was introduced. On the other hand, over some stations at high northern latitudes (Alert and Barrow, not shown), the reanalysis tends to underestimate CO during the same period. These lower values after April 2008 can also be seen in Figure 2. This could be related to long-range transport of air masses from regions where satellite retrievals are assimilated.

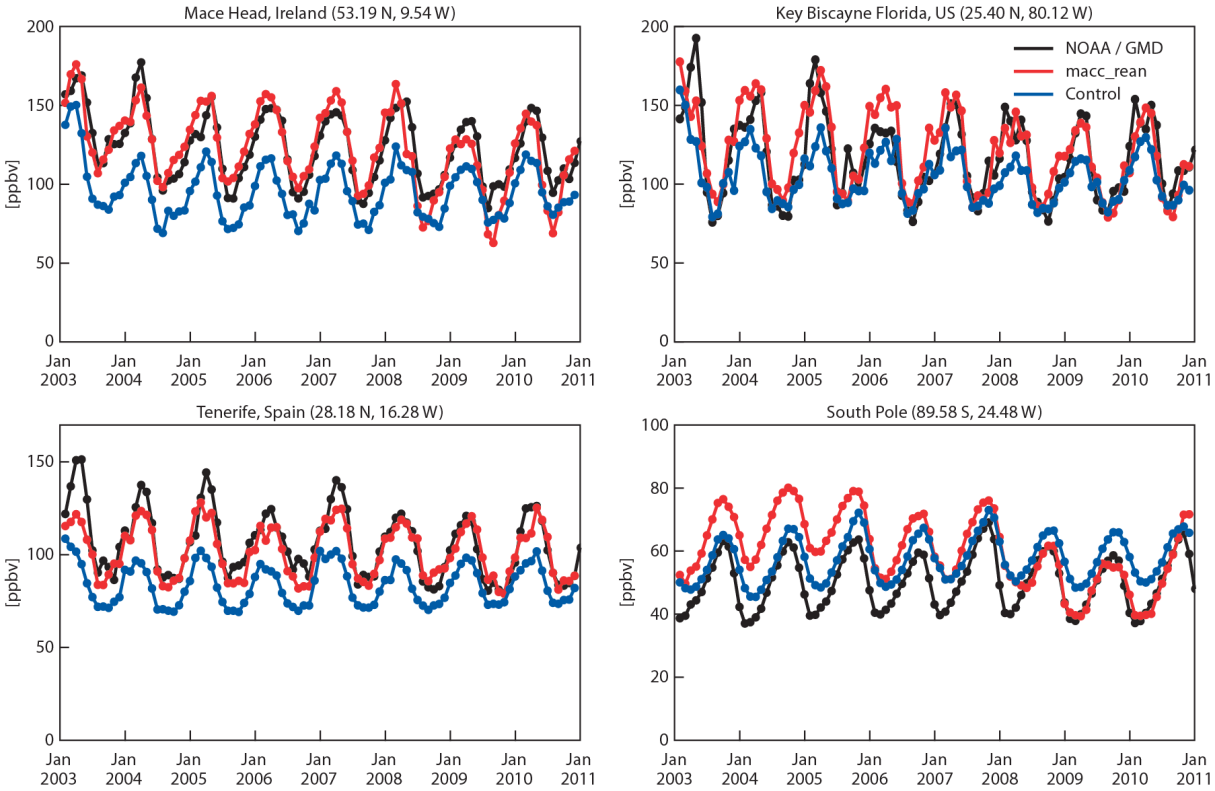

Figure 5: Time series (2003-2010) of monthly mean CO concentrations (ppbv) from the MACC reanalysis (red), the control run (blue), and from NOAA/GMD ground-based measurements (black) over Mace-Head (top left), Key Biscayne (top right), Ascension Island (bottom left), and South Pole (bottom right) stations.

Next, CO from the reanalysis is compared with data from the MOZAIC (Measurement of Ozone, Water Vapour, Carbon Monoxide and Nitrogen Oxides by Airbus in-service Aircraft) programme (Marenco et al. 1998; Nédélec et al. 2003). This programme provides profiles of various parameters, including CO and O_3 taken during aircrafts' ascents and descents at various airports. MOZAIC data were available from January 2003 till December 2010, but data availability is patchy in 2010.

MOZAIC CO data have an accuracy of ± 5 ppbv, a precision of ± 5%, and a detection limit of 10 ppbv (Nédélec et al. 2003).

Figure 6 shows the time mean vertically averaged relative CO bias between 300 and 1000 hPa from the MACC reanalysis. The reanalysis has a negative bias in the troposphere with the exception of a few airports. The averaged biases are usually less than 15 %, but larger at some tropical airports. The control run has a bias that is more than twice as large (not shown).

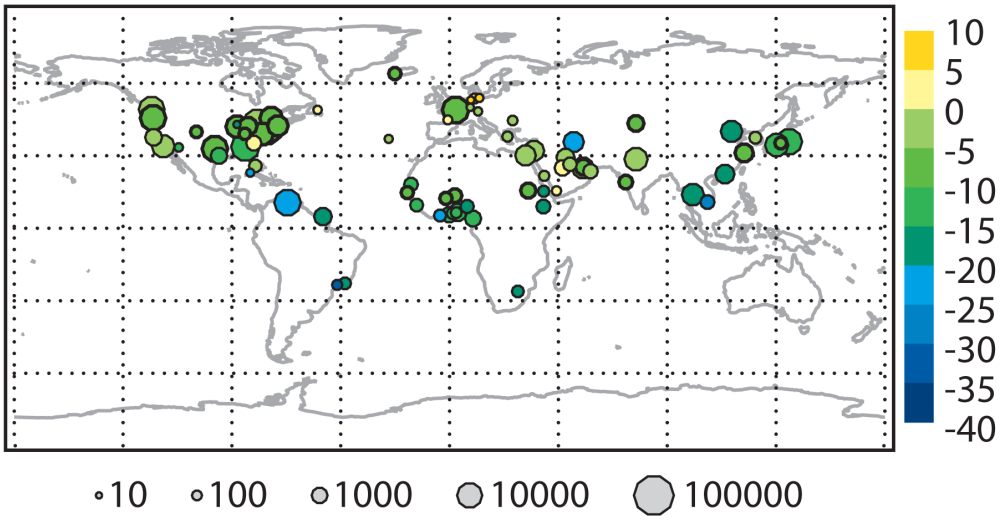

Figure 6: Vertically (300-1000 hPa) averaged CO bias in % of MACC reanalysis minus MOZAIC ascent/descent data averaged over the period January 2003 to December 2010. The diameter of the circles indicates the number of profiles over the respective airports.

Figure 7 shows time mean profiles from the reanalysis (black) and the MOZAIC data (red) averaged over NH extratropical airports (solid lines) and tropical airports (dashed lines). Also shown are the mean differences between the reanalysis and MOZAIC data for both hemispheres. The highest CO concentrations are found near the surface and values get smaller in the free troposphere. The biases between 700-300 hPa are less than -5 % in the NH and around -10 % in the tropics, but larger in the lower troposphere. By assimilating CO satellite data the biases are greatly reduced in the reanalysis compared to the control run without data assimilation. This was also seen by Elguindi et al. (2010). The assimilated satellite data have only little sensitivity to CO in the lower troposphere (see Figure 1). Here the influence of the emissions is important and accurate emissions are crucial to reproducing the high CO values seen in the MOZAIC data. The low bias of the MACC reanalysis suggests an underestimation of CO or precursor emissions (Elguindi et al. 2010). The importance of the emissions for correct CO values in the lower troposphere in the MACC system was also found in a study of the 2010 Russian wildfires (Huijnen et al., 2012). Furthermore, some of the bias in the lower troposphere is likely to be a representativeness error, because the model with a horizontal resolution of T255 (corresponding to a reduced Gaussian grid of about 80 km x 80 km) is not able to reproduce the high values observed by MOZAIC over polluted airports. Larger positive biases are seen in the NH above 300 hPa.

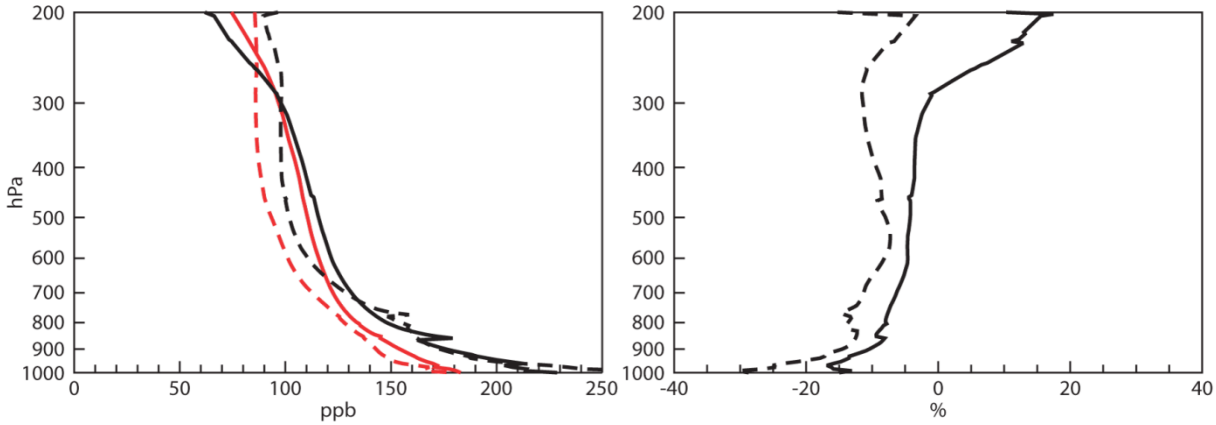

Figure 7: Left panel: Mean CO profiles in ppbv from the MACC reanalysis (red) and MOZAIC data (black). The solid lines show the means for NH airports (north of 30°N), the dashed lines the means for tropical airports (30°S-30°N). Right panel: CO bias in % of MACC reanalysis minus MOZAIC for NH airports (solid) and tropical airports (dashed). Data are averaged over the period January 2003 to December 2010.

Figure 8 shows a time series of monthly mean CO bias profiles at Frankfurt airport. Frankfurt is the most frequented airport in the MOZAIC database and the most reliable in terms of data availability. A total of 7182 MOZAIC profiles were available over Frankfurt between January 2003 and December 2010. Concentrations are larger during the winter months and larger concentration extend higher up in the troposphere. The plot confirms that CO is underestimated in the surface layer as seen in Figure 7. This was also found for other MOZAIC urban sites affected by air pollution such as Beijing, Tokyo and Cairo (not shown). Whilst the bias is generally less than 10 % in the free troposphere (850 hPa to the upper troposphere), large positive biases are found in the upper troposphere and lower stratosphere, with the largest values during winter and spring, possibly due to too much upward transport of CO.

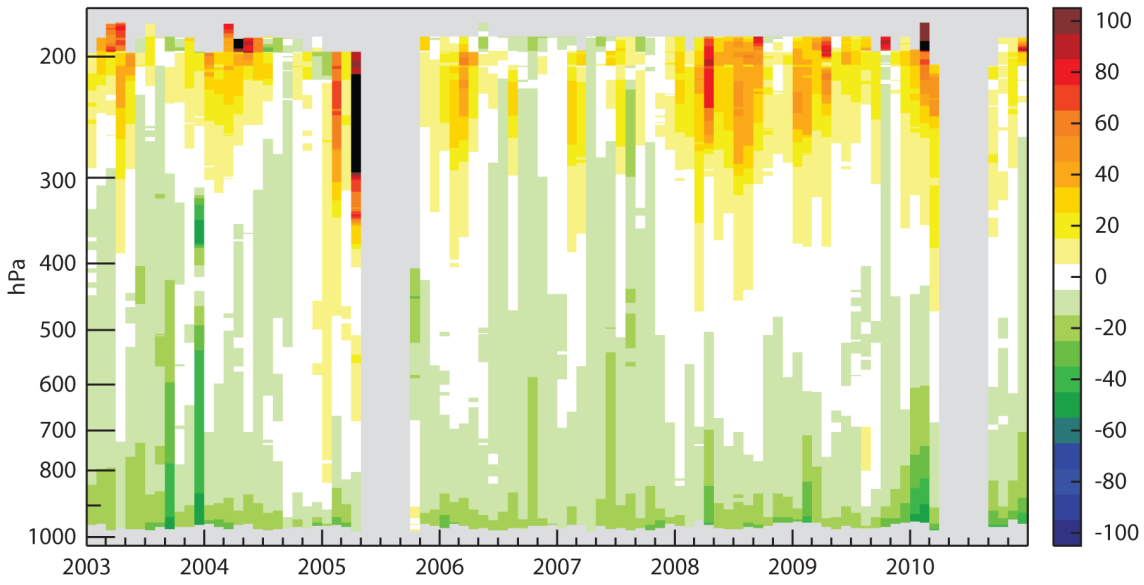

Figure 8: Time series of monthly mean CO biases (MACC reanalysis minus MOZAIC) in % at Frankfurt airport (50.0°N, 8.6°E) for the period January 2003 to December 2010.

3.2 O_3 analysis

The ozone retrievals assimilated in the MACC reanalysis are listed in Table 1. MIPAS and MLS measure in the mid-infrared and microwave part of the spectrum, respectively, and are the only ozone data used in the MACC reanalysis that are available independent of illumination condition, including during the polar night. Together with the GOME O_3 profile retrievals they are also the only assimilated ozone profile data with higher vertical resolution, which had been shown in the past to be crucial for obtaining a realistic vertical ozone distribution in MACC and ECMWF analyses (Flemming et al., 2011, Dethof 2003). By assimilating GOME, MIPAS or MLS data with their vertically resolved information in the stratosphere together with total column ozone data of OMI and SCIAMACHY, tropospheric ozone can also be constrained.

3.2.1 Total column and stratospheric ozone

Figure 9 shows time series of zonal mean total column ozone from the MACC reanalysis, the control run, SCIAMACHY and OMI data. The reanalysis shows a realistic seasonal cycle in both hemispheres. In the NH extratropics, ozone values are highest during boreal winter and spring. This is a result of poleward and downwards transport of ozone by the large scale Brewer-Dobson circulation (Weber et al. 2011; Brewer 1949; Dobson 1956). In the tropics, where there is slow large scale ascent, the ozone columns are lower. In the SH, the reanalysis shows the very low values of the Antarctic ozone hole, and also the higher values seen in a belt around the Antarctic. The assimilation of MIPAS and MLS data gives information during the polar night when the UV instruments GOME, SBUV/2, SCIAMACHY and OMI can not observe the ozone field.

The time series of total column ozone from the control run illustrates that the free running model generally overestimates the ozone column. The control run has a positive total column bias compared to SCIAMACHY and OMI data everywhere. This bias points to problems with the stratospheric ozone distribution in the MOZART-3 CTM, which dominates the ozone column. Similar results were found by Kinnison et al. (2007) when they used ECMWF meteorological fields to drive the MOZART-3 model. Also, the Antarctic ozone hole is not deep enough in the control run, a known problem of MOZART-3 in the MACC configuration (Flemming et al. 2011). The model update to Version 3.5 on 1 January 2009 slightly improves the representation of the ozone hole, but values are still not as low as observed.

Figure 10 shows the mean relative bias between the MACC reanalysis and KNMI's multi sensor reanalysis (MSR, van der A et al., 2010) for the years 2003-2008 which is based on SBUV/2, GOME, TOMS, SCIAMACHY and OMI observations, and the SCIAMACHY data (Eskes et al., 2003) for the years 2003-2010, in order to evaluate the data after 2008. These data sets are not fully independent because SBUV/2, OMI and SCIAMACHY data were also assimilated in the MACC reanalysis. Nevertheless, it provides a good insight into the consistency of the MACC total column ozone fields. There is good agreement between the datasets at various latitude bands, showing a clear improvement in total columns compared to the control run. Biases of the MACC reanalysis with respect to the MSR and the SCIAMACHY assimilation system are generally of similar magnitude, which constrains the evaluation for 2009-2010, for which no MSR data is available. Zonal-average, monthly mean biases of the reanalysis compared to the MSR over the extratropical northern hemisphere are less than 3%. The reanalysis shows a general positive bias except for the NH winter season. During this season observed

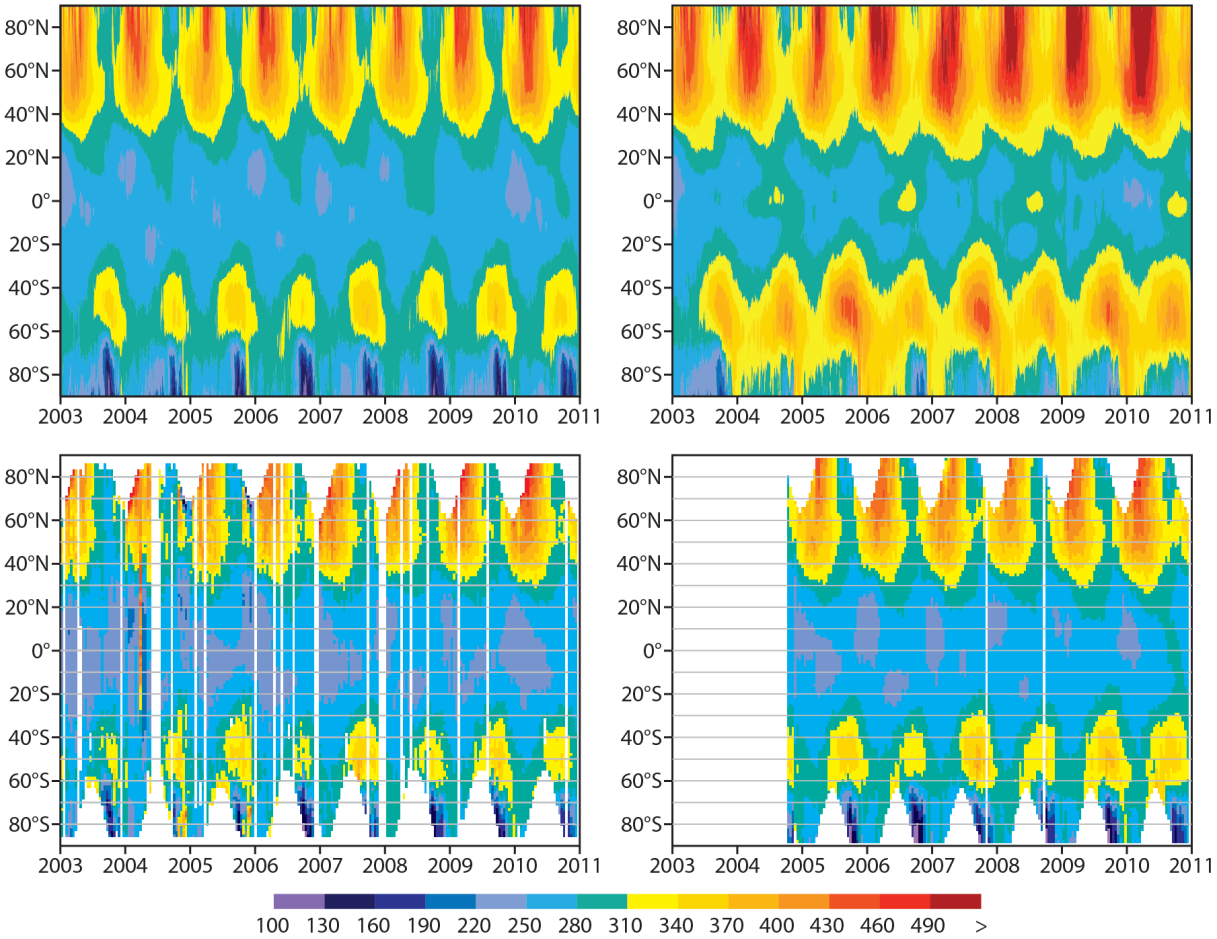

Figure 9: Time series of zonal mean total column O_3 field in Dobson Units (DU) for the period 2003 to 2010 from the MACC reanalysis (top left), the control run (top right), SCIAMACHY (bottom left) and OMI (bottom right).

Figure 10: Time series of the mean difference of the MACC reanalysis minus total O_3 columns from the multi sensor reanalysis (solid) and SCIAMACHY fields (dashed) in % averaged over (a) the NH extratropics (30°N-90°N),(b) tropics(30°S-30°N) and (c) SH extratropics (90°S-30°S) for the years 2003 to 2010.

average O_3 total columns increase. The negative biases suggest a slightly too slow response of the reanalysis to this increase, compared to the MSR. Biases in the tropics are very small (<2%). Only a period of ~1.5 year, mainly covering 2008, a slightly larger of ~3% was found. Over the extratropical southern hemisphere the biases are less than 5% during 2003-2004, and below 3% during later years when MLS and OMI data were assimilated. The seasonal oscillations seen in the NH and SH are the result of a seasonally varying model bias in the MOZART-3 CTM and the fact that MLS and MIPAS data are assimilated in the reanalysis and give information about ozone in the polar night that is not included in the MSR.

Seasonal mean climatologies of total column ozone from the reanalysis for the years 2003-2010 are shown in Figure 11. The figure shows the largest ozone columns over the NH extratropics in DJF and MAM and lowest values in SON, as already seen in Figure 9. In the Tropics the lowest values are seen in DJF, when the Brewer Dobson circulation is strongest. In the SH, the Antarctic ozone hole is visible in SON with seasonal mean total column values lower than 200 DU. Ozone columns in the circum Antarctic belt are at their highest in SON, when large scale descent brings down ozone rich air.

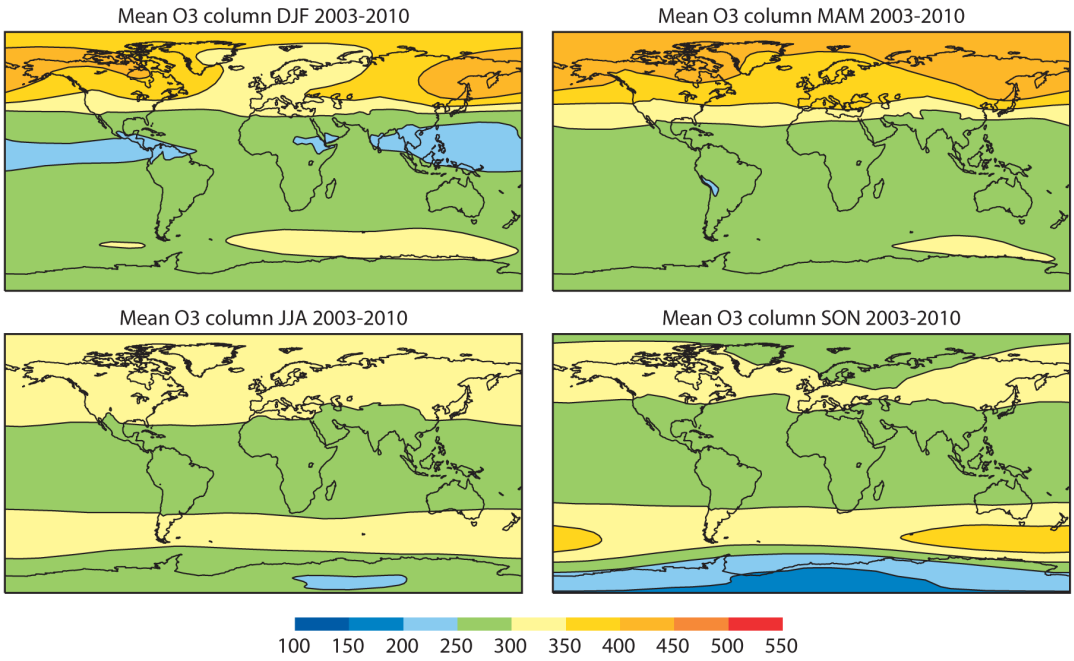

Figure 11: Seasonal mean total column ozone field from the MACC reanalysis in DU for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010.

The vertical structure of the MACC ozone field can be seen in the seasonal mean cross sections in Figure 12. Ozone concentrations in the stratosphere are the result of the balance of ozone production, ozone loss and transport. The figure clearly depicts the ozone layer which is at higher altitude in the tropics than in the extratropics as a result of the large scale Brewer-Dobson circulation. Even though the tropics are an area of net ozone production, the lowest zonal mean ozone values are found in the tropical upper troposphere from where ozone is transported upwards into the stratosphere and then

poleward and downward in the extratropics. The concentrations in the tropics are lowest in DJF when the upwelling branch of the Brewer-Dobson circulation is strongest. At the same time, the ozone layer in the NH extratropics is strongest because descent brings ozone rich air down. In SON, the impact of the chemical ozone destruction over Antarctica is clearly visible. Here the ozone layer is very thin and values around the ozone maximum are less than 9 mPa in the seasonal mean.

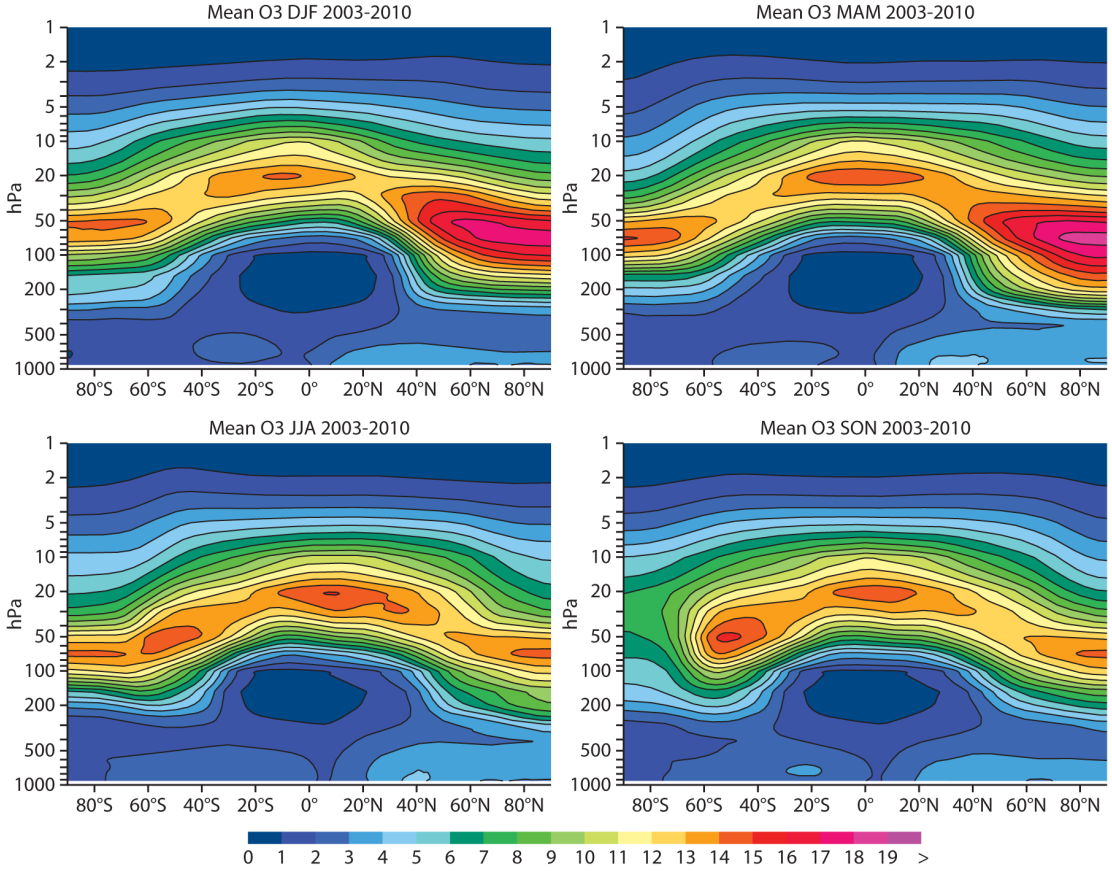

Figure 12: Seasonal mean zonal average ozone cross sections from the MACC reanalysis in mPa for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010.

Next, MACC ozone is validated against independent data that were not assimilated in the reanalysis. First, stratospheric ozone fields from the reanalysis are compared with data from the Atmospheric Chemistry Experiment Fourier Transform Spectrometer (ACE-FTS, Dupuy et al. 2009). ACE-FTS data show good agreement with correlative measurements, with a small positive bias with mean relative differences of about +5% between 15-45 km, and a larger positive bias above 42-45 km (Dupuy et al. 2009). ACE-FTS data were available from January 2004 to September 2010, with a gap from December 2009 to May 2010. Figure 13 shows time series of monthly mean relative biases between the reanalysis and ACE-FTS data at 10, 46 and 100 hPa for extra-polar and polar regions. The reanalysis generally has a small positive bias. At 100 hPa the biases are less than 10% in most regions. Noticeable are larger biases between April and August 2004 and after March 2009, a result of changes in the assimilated data. Between April and August 2004 no ozone profile data were assimilated in the

reanalysis because MIPAS data were not available anymore and MLS data were not available yet. The larger biases after March 2009 are the result of assimilating NRT MLS data (see Table 1). For the NRT MLS data the data producers recommend not to use the ozone values below 64 hPa (bottom three levels) because those values are largely influenced by the a-priori of the retrieval. Therefore the NRT MLS data can not constrain ozone in the lower stratosphere and upper troposphere. The largest negative biases at 100 hPa are seen over the South Pole in September and November

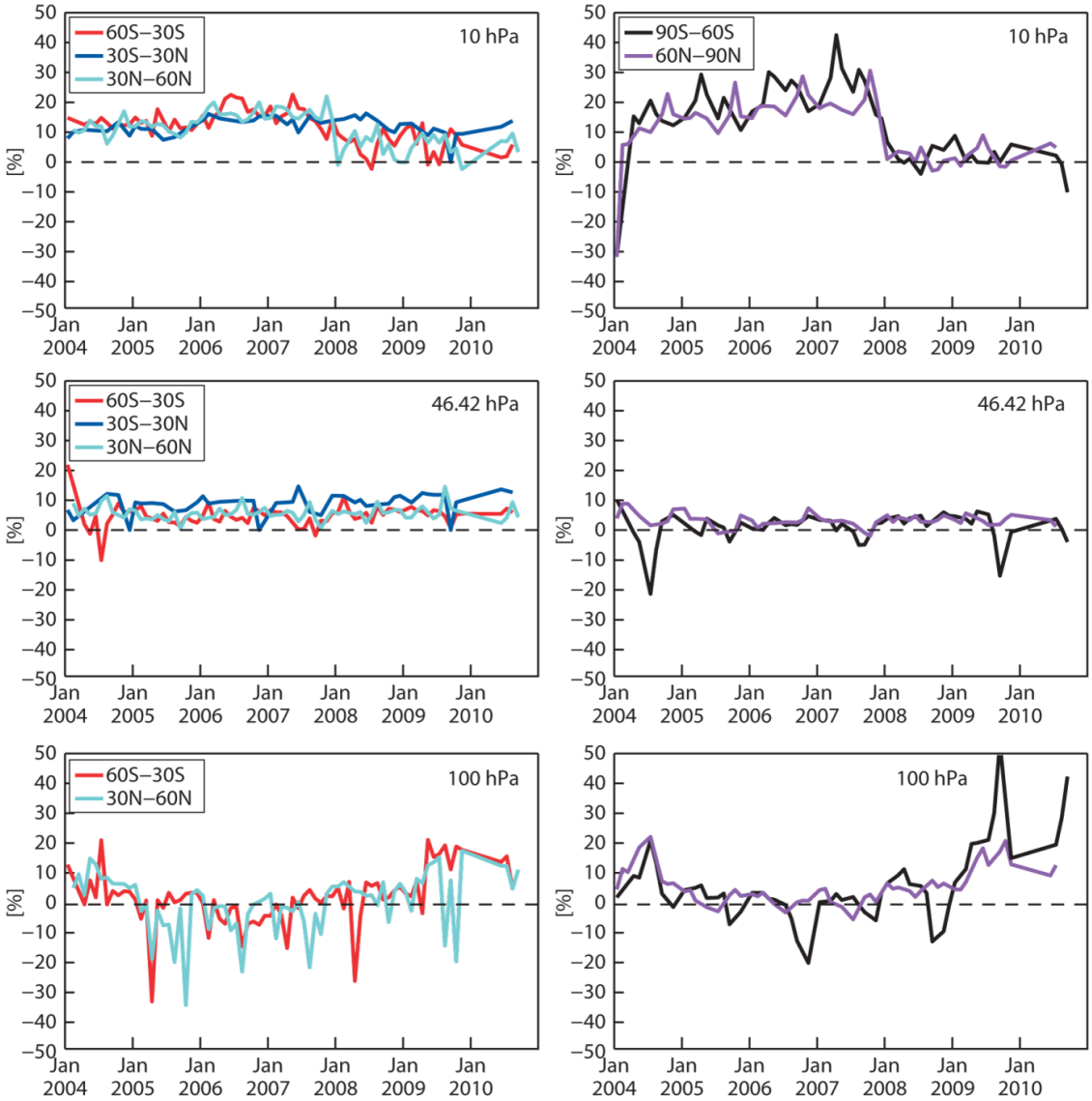

Figure 13: Time series of monthly mean relative ozone biases between the MACC reanalysis and ACE-FTS data (MACC minus ACE-FTS) in % for the period January 2004 to September 2009 at 10 hPa (top), 46 hPa (middle), and 100 hPa (bottom). The left panels show time series for extra-polar regions, the right panels time series for polar regions. There are not enough good quality ACE-FTS data at 100 hPa between 30°S and 30°N for a meaningful validation, hence this curve is omitted from the bottom left panel.

The change to NRT MLS data does not have a noticeable impact at higher levels (46 and 10 hPa). At 46 hPa, the biases in the extra-polar regions are less than 10% and less than 5% in the polar regions. At 10 hPa we see an impact of the changed bias correction in January 2008 (section 2.3.2), particularly in the polar regions, and biases are reduced after January 2008. At 46 and 100 hPa the change to the variational bias correction does not have a noticeable impact. In this altitude range, where the bulk of the ozone column is located, the ozone bias correction is successfully anchored by SBUV/2 data alone. The monthly mean standard deviations of the differences between ACE data and the reanalysis are usually around 10% at 10 and 46 hPa, but up to 20% at 100 hPa, and even larger over the South Pole during the ozone hole season (not shown).

Figure 14 shows profiles of seasonal mean relative biases of MACC ozone with respect to ACE-FTS data for extra-tropical and tropical areas. The biases are negative above 3 hPa and below 100 hPa (where ACE_FTS errors are large) and mainly positive between 3 and 100 hPa. In the stratosphere, they are smallest between 20 and 50 hPa where they are less than 5% for most areas. The largest positive bias of up to 15% is found around 10 hPa, i.e. at the level of maximum ozone mixing ratio. Below 100 hPa the ACE-FTS data have large uncertainties.

Next, MACC ozone is compared with ozone sonde data. The ozone sondes are available for the whole 8 years of the reanalysis and come from a variety of data sources: WOUDC, SHADOZ, NDACC, MATCH campaigns, and from the ECWMF Meteorological Archive and Retrieval System (MARS). The precision of Electrochemical Concentration Cell ozone sondes is of the order of ±5% in the range between 200 and 10 hPa, between -14% and +6% above 10hPa, and between -7% and +17% below 200 hPa (Komhyr et al. 1995). Larger errors are found in the presence of steep gradients and where the ozone amount is low. The same order of precision was found by Steinbrecht et al. (1996) for Brewer-Mast sondes. We did not include Carbon Iodine sondes in the validation because they have larger biases.

Figure 15 shows the time mean vertically averaged relative biases between 5 and 100 hPa from the MACC reanalysis minus ozone sondes. The mean stratospheric biases are less than ±10% for most stations and in many cases even less than ±5%. Larger biases are found over South East Asia. The control run has considerably larger positive biases in the stratosphere than the reanalysis. Biases of up to 40% can be seen over Antarctica which is in agreement with the large total column biases seen in Figure 9.

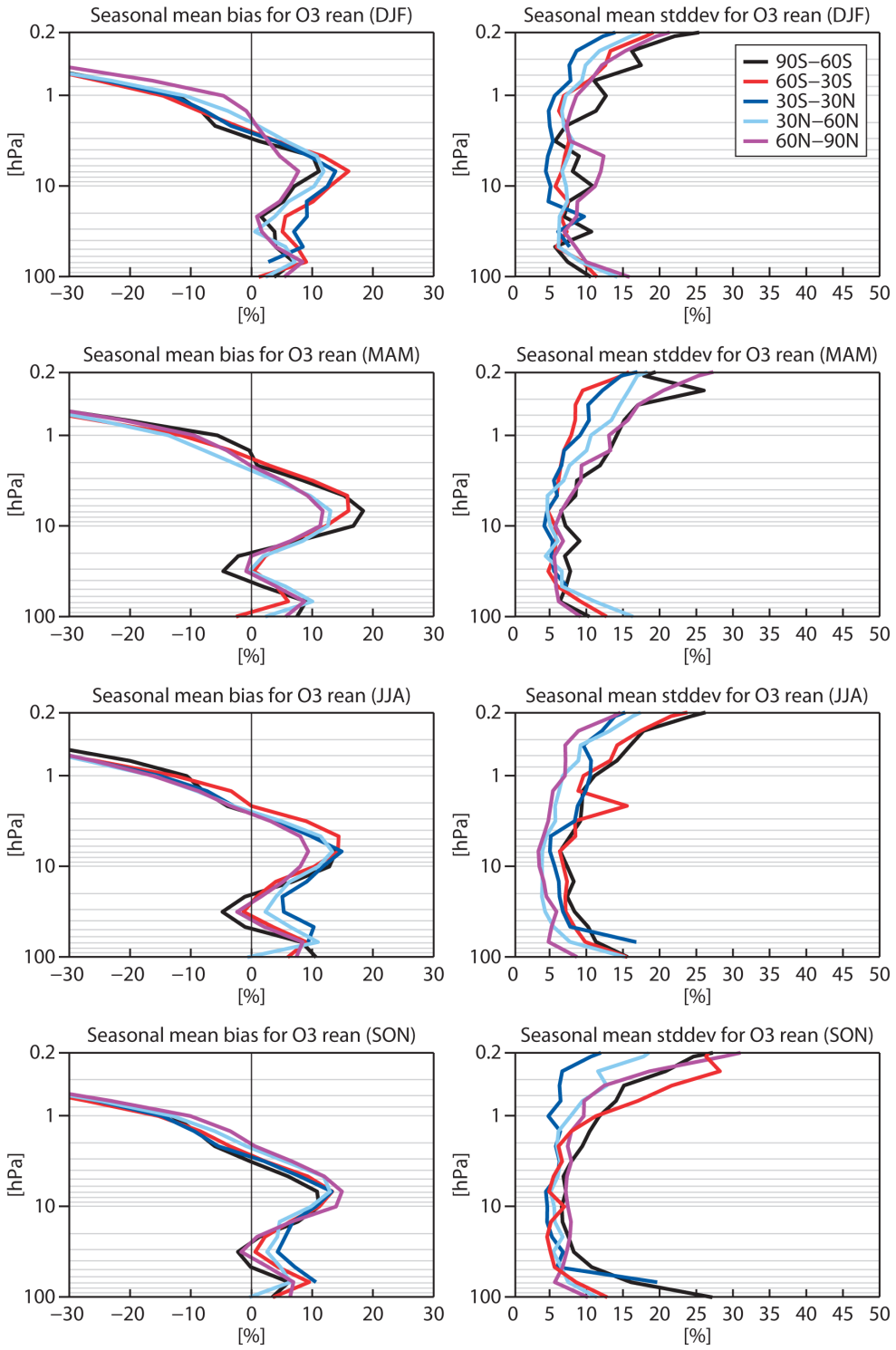

Figure 14: Seasonally averaged relative ozone bias profiles (left) of MACC reanalysis minus ACE-FTS data, and seasonally averaged standard deviation profiles of the differences (right) in % for polar and extra-polar areas (as defined in the figures).

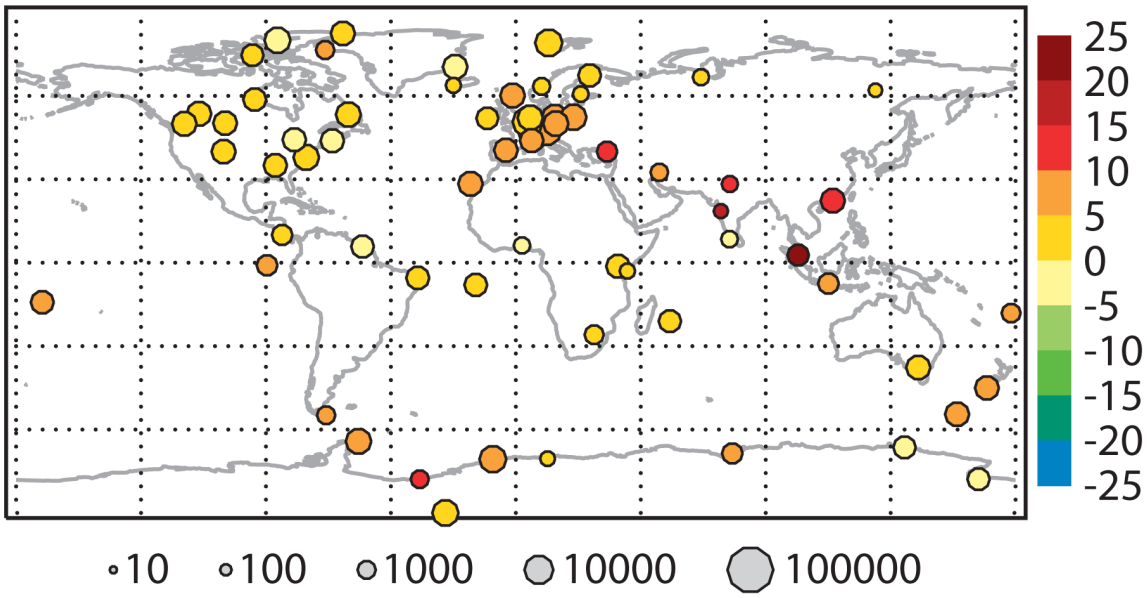

Figure 15: Mean stratospheric O_3 bias in % of MACC reanalysis minus ozone sonde data averaged between 5-100 hPa for the period January 2003 to December 2010. The diameter of the circles indicates the number of profiles over the respective stations.

To assess the vertical structure of these biases in more detail Figure 16 shows time averaged ozone profiles and bias profiles from the reanalysis and ozone sondes for the NH extratropics, the Tropics, and the SH extratropics. The figure demonstrates that the reanalysis agrees to within ±5% in the NH and to within -5% to +10% in the SH, where the largest biases are seen near the surface. In the tropics the reanalysis and sondes agree to within ±10% above 70 hPa, but have larger negative relative differences below 100 hPa. The tropospheric bias is discussed further in the next subsection.

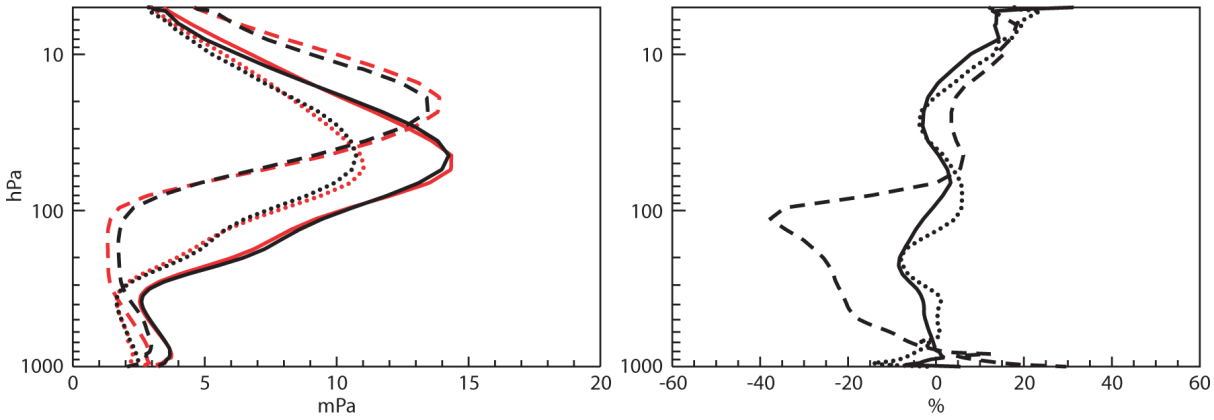

Figure 16: Time mean ozone profiles (left) in mPa from the MACC reanalysis (red) and ozone sondes (black), and averaged ozone bias in % of MACC reanalysis minus ozone sondes (right) averaged over the period January 2003 to December 2010. Solid lines show means for the NH extratropical stations (i.e. north of 30°N), dashed lines for tropical stations (30°S-30°N), and dotted lines for SH extratropical stations (south of 30°S).

3.2.2 Tropospheric ozone

Figure 16 shows larger tropospheric than stratospheric relative biases between MACC ozone and ozone sondes, particularly in the tropics. To investigate this bias in more detail the reanalysis ozone is compared with ozonesondes and MOZAIC data in the altitude range between 200 and 1000 hPa. The MOZAIC ozone data have a detection limit of 2ppbv and a precision of ± (2ppbv +2%) (Marenco et al. 1998).

Figure 17 shows the mean relative ozone bias of MACC ozone with respect to ozone sondes and MOZAIC data averaged between 200 and 1000 hPa averaged over the period from January 2003 to December 2010. With respect to ozone sondes, the reanalysis biases are within ± (5-10) % in the NH and over the Antarctic, but larger negative biases are found in the Tropics. With respect to MOZAIC data, the reanalysis has mainly positive biases of less than 10% over Europe, North and South America and negative biases of up to -10% over Africa. Larger positive biases with respect to MOZAIC are found over East Asia. Larger biases over East Asia were also seen for CO data (Figure 6) suggesting that either the horizontal resolution is not high enough to reproduce the high values seen over polluted airports, or that the differences are due to the chemical coupling O_3 and other fields. At the very high NO_x concentrations over Asia, especially in polluted regions near cities (where the airports are), O_3 will be destroyed. The model has too little NO_x here (see section 3.3 below), so it might not destroy enough O_3, which could explain the positive bias seen in these regions.

The differences seen in the biases with respect to ozone sondes and MOZAIC data are consistent with findings by Tilmes et al. (2011) who showed that ozone sondes measure higher concentrations than MOZAIC data in the free troposphere over Western Europe, North and South America, but to agree to within the given error bars. Saunois et al. (2011) investigated uncertainties as a result of different sampling frequencies and found these to be of the order 7-14% in the free troposphere and larger above and below. These sampling frequency uncertainties have to be considered when comparing ozone sonde and MOZAIC data which have a very different measurement frequency, and they are large enough to explain the differences between the ozone sonde and MOZAIC biases seen in Figure 17.

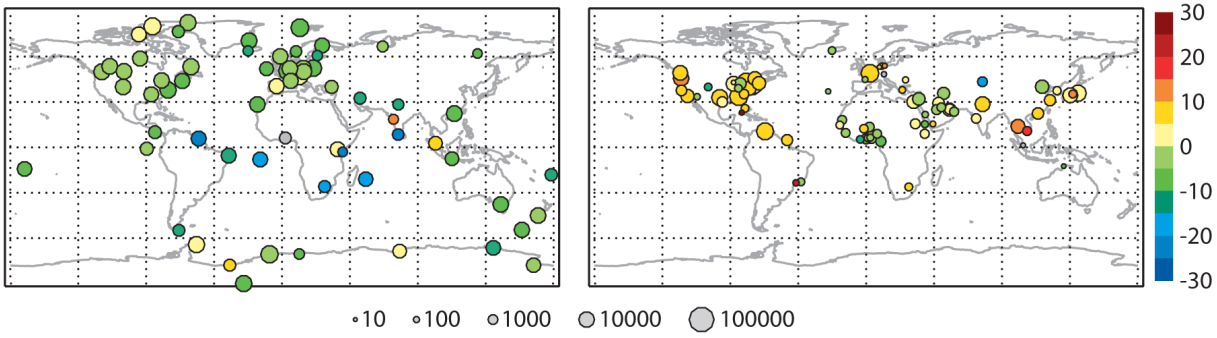

Figure 17: Mean tropospheric ozone bias in % of MACC reanalysis minus ozone sondes (left) and minus MOZAIC data (right) for the period January 2003 to December 2010 averaged between 200-1000 hPa. The diameter of the circles indicates the number of profiles over the respective stations. Grey circles depict biases of greater than -30%.

Figure 18 shows the time mean relative tropospheric bias profiles of the analysis minus ozone sondes and minus MOZAIC data. In the tropical troposphere, the reanalysis shows a large negative bias with respect to sondes above 650 hPa, and the opposite below. Elsewhere the biases are small which is in agreement with Figure 17. Compared to MOZAIC data the reanalysis shows the largest positive bias below 800 hPa. The bias of the reanalysis is small and positive in the free troposphere in the NH extratropics, and small and negative in the tropics.

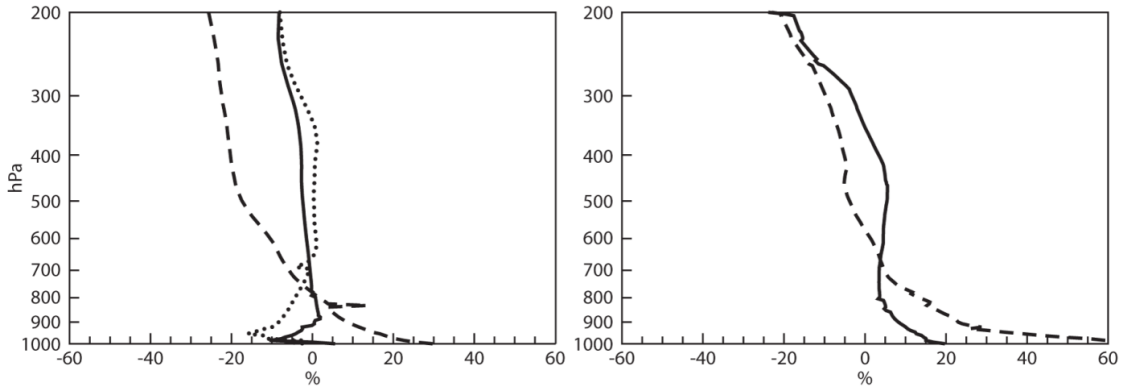

Figure 18: Time mean ozone biases in % from the MACC reanalysis minus sondes (left) and MACC minus MOZAIC data (right) averaged over the period January 2003 to December 2010. Solid lines show means for the NH extratropical stations (i.e. north of 30°N), dashed lines for tropical stations (30°S-30°N), and dotted lines for SH extratropical stations (south of 30°S). Note that there are no MOZAIC flights in the SH in our database.

A time series of the monthly mean biases with respect to ozone sondes at Hohenpeissenberg and MOZAIC profiles at Frankfurt (Figure 19) shows good agreement between the two data sets. Large negative biases are seen at the beginning of the reanalysis, when the system was still adjusting and fewer ozone retrievals were available. From the end of 2004 until the end of 2007, biases in the troposphere increase to up to 50% in the free troposphere and more near the surface. These drifts were traced back to the problem with the variational bias correction for ozone (see section 2.3.2) which was resolved on 1 January 2008. Figure 19 shows that the biases in the free troposphere return to below 20% after January 2008. The large negative biases above 300 hPa seen after March 2009 are a result of using NRT MLS data instead of the offline product, as already discussed above in Section 3.2.1.

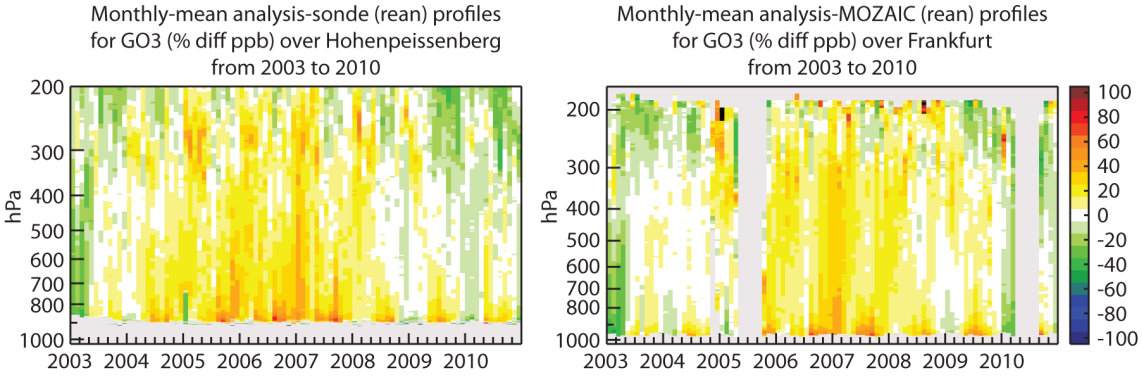

Figure 19: Time series of monthly mean ozone biases (MACC reanalysis minus observations) with respect to ozone sondes at Hohenpeissenberg(left, 47.5°N, 11E°) and MOZAIC profiles at Frankfurt airport (right, 50.0°N, 8.6°E) for the period January 2003 to December 2010 in %.

3.2.3 Surface ozone

Surface ozone from the reanalysis is validated against data from the European Monitoring and Evaluation Programme (EMEP, http://www.emep.int/). Ozone measurements within EMEP are conducted with commercial UV monitors. An overview on the calibration/maintenance and data quality can be found on the EMEP web site. A typical accuracy for a commercial UV instrument is about 1.5 ppbv. The precision is also close to 1.5 ppbv for a 10s measurement Hourly mean EMEP observations of O_3 from all available (close to surface level) EMEP stations (altitude<600m) are used for the validation of the reanalysis simulations for the years 2003-2008. Only stations meeting the 75% availability threshold per day and per month are taken into account. Ground level three-hourly averages from the reanalysis are used to produce daily ozone averages, and the data are interpolated horizontally to the location of the EMEP stations. The EMEP surface ozone values and the interpolated surface reanalysis values are compared on a seasonal basis for the latitude bands of 30N-40N (Southern Europe), 40N-50N (Central Europe) and 50N-70N (Northern Europe) and the results are shown in Figure 20. Over Northern Europe, the reanalysis underestimates O_3 levels during the first half of the year and overestimates O_3 during the second half. This results in a negative bias (model minus observations) during winter and spring and a positive one during summer and autumn. Over Central and Southern Europe, the seasonal variability of ozone from the reanalysis agrees well with the observations, but a large negative bias is observed over Central Europe during winter and spring. Over Southern Europe the reanalysis systematically over-predicts ozone mixing ratios with the highest discrepancies observed during summer and autumn. It should be noted that the reported absolute biases are significantly higher than the typical accuracy (1ppbv) of the commercial UV monitoring instruments.

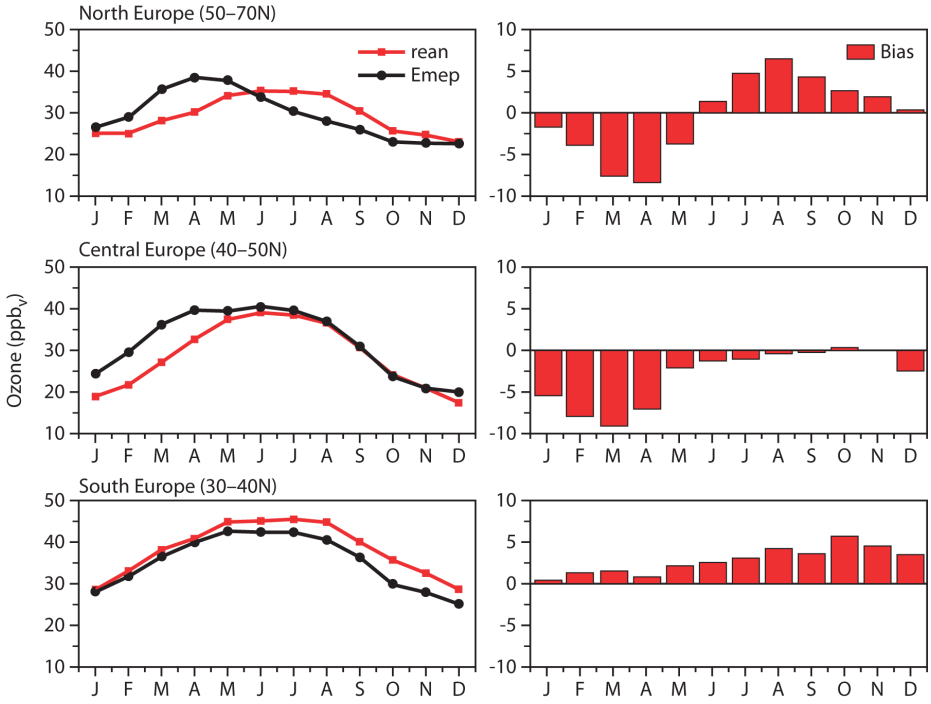

Figure 20: Mean monthly annual variability during the period 2003-2008 of the MACC reanalysis (black) and observations (blue), and the absolute bias (red bars) over Northern Europe (top panels, based on 72 monitoring stations), Central Europe (middle, based on 27 stations) and Southern Europe (bottom, based on 5 stations) in ppbv.

To better understand the above mentioned seasonal variability of the biases, the model simulations and the EMEP observations have been separated into daytime (12:00-15:00h) and nighttime (0:00-3:00) datasets (Figure 21). Over South Europe, the overestimated reanalysis ozone (Figure 20) coincides with positive daytime biases, which are largest between May and October. Over Central Europe and despite the overall good agreement between the observations and the reanalysis during the warm period (Mar-Oct), large amplitudes were computed between the nighttime and daytime biases. Specifically, during daytime the reanalysis strongly overpredicts ozone levels. During nighttime on the other hand, the reanalysis has large negative biases. Lastly, over North Europe, three well defined temporal periods describe the biases. During the first period (Jan-Apr) the reanalysis systematically underestimates the surface ozone mixing ratios. The second period (Apr-Aug) is characterized by a progressive shift of the biases from negative to positive values whilst for the third period (Aug-Dec) the day and night biases are positive. The differences of the daytime and nighttime biases are larger

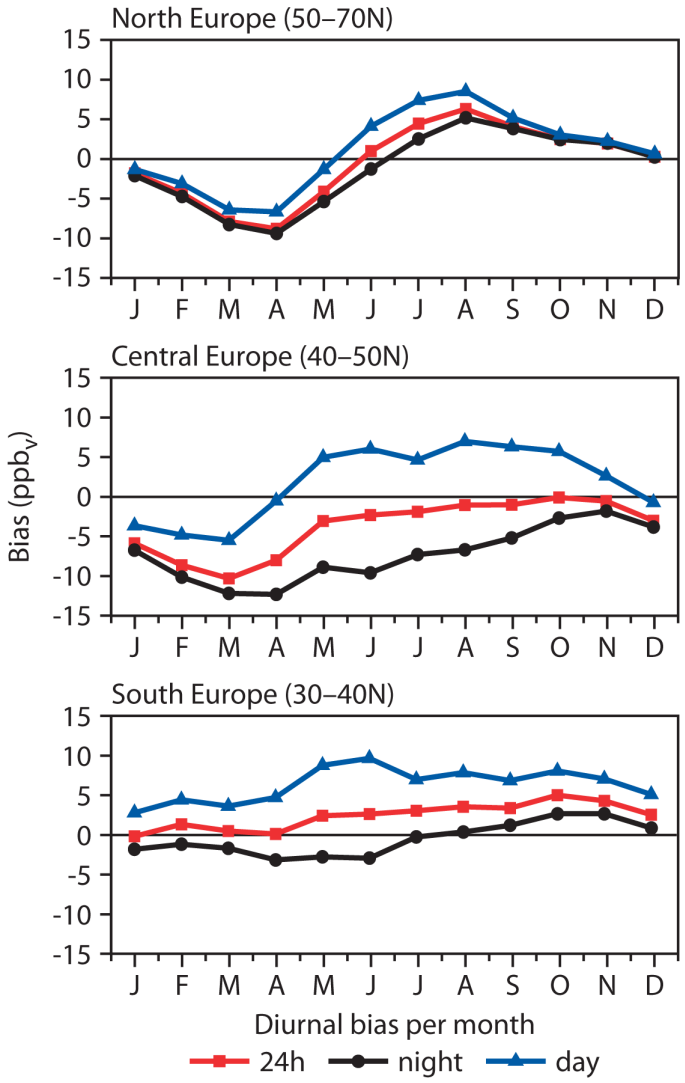

Figure 21: Monthly mean variability (2003-2008) of the biases (MACC reanalysis minus EMEP) during daytime (blue line), nighttime (black line) and during the 24h time period (red line) over Northern Europe (top panel), Central Europe (middle) and Southern Europe (bottom).

during the warmer periods when the highest amplitudes are observed. The reason for the diurnal variations of the biases is not understood at present and is being investigated further. It could be a result of using NO_x emissions that do not have a diurnal cycle in the MOZART-3 CTM. During the day this would result in too low emissions and hence too little O_3 destruction, during the night emissions and hence O_3 would be too large.

3.2.4 Discussion of ozone analysis

To put the magnitude of the ozone biases seen in the MACC reanalysis into perspective, MACC ozone is compared with ozone data from the ERA Interim reanalysis (Dee et al. 2011). This dataset is ECMWF's latest reanalysis product, and its ozone analysis fields have better quality than the ERA-40 ozone fields (Dragani 2010; Dragani 2011). Figure 22 shows time series of monthly mean ozone bias profiles with respect to sondes from the reanalysis and ERA Interim for a station at high northern latitudes (Alert), a tropical station (Natal), and an Antarctic station (South Pole). The MACC bias plots confirm what was already seen in Figures 13 and 19. At all 3 stations the biases are larger during 2004 when no profile data were available, and after March 2009 when NRT MLS data were used. At Alert and South Pole the issues with the bias correction lead to larger biases in the troposphere and above 15 hPa between August 2004 and December 2007. At all three stations tropospheric biases are reduced after 1 January 2008 when the problem with the variational bias correction for ozone was resolved. Figure 22 also shows that at South Pole the largest positive stratospheric biases are seen in austral spring during the ozone hole season. Here, the control run has a large bias and the ozone analysis can not completely remove this bias.

Figure 22 illustrates that the biases of ERA Interim ozone are larger than those of the MACC reanalysis almost everywhere. This is particularly noticeable in the tropical troposphere, where biases greater than 100% are found near the surface in ERA Interim. At Alert and South Pole, there are large changes in ERA Interim between negative biases during winter/ spring and positive biases during summer/ autumn. In ERA Interim ozone data are assimilated without bias correction and it is possible that a bias in one of the UV sensors can lead to a seasonally varying ozone bias with respect to sondes, depending on when data were available in the analysis. In ERA Interim NRT MLS data have been used since November 2008, and like in the MACC reanalysis they lead to larger departures in the upper troposphere and lower stratosphere where the profile information is missing after this date. Also, the representation of the ozone hole is worse when NRT MLS data are used. The comparison shown in Figure 22 illustrates that ozone from the MACC reanalysis has smaller errors than other available reanalysis products.

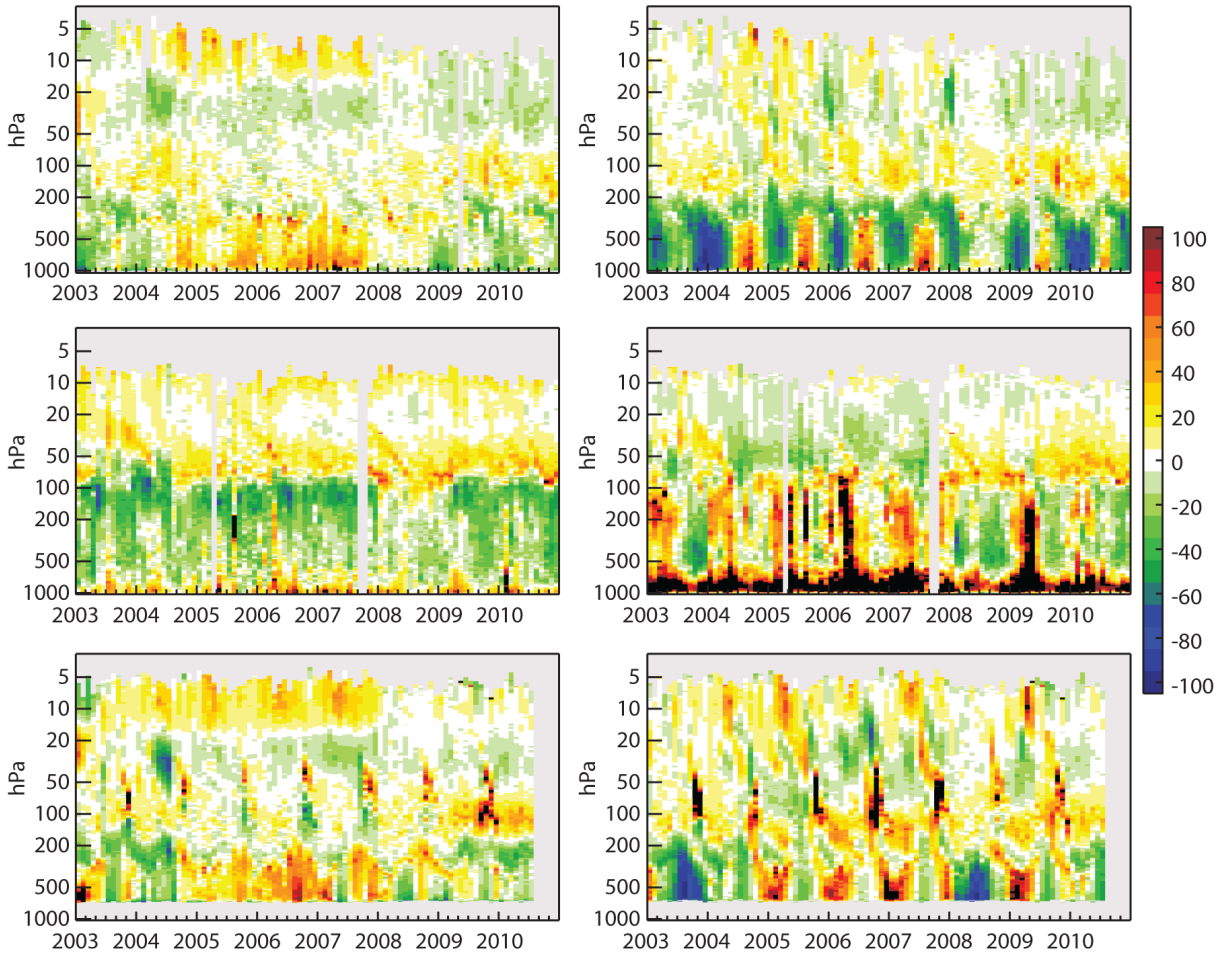

Figure 22: Time series of monthly mean ozone bias profiles from the MACC reanalysis (left panels) and ERA Interim (right panels) with respect to ozone sondes at Alert (82.5°N, 62.3°W, top), Natal (5.5°S, 35.3°W, middle) and South Pole (bottom) for the period 2003 to 2010 in %.

3.3 NO_2 analysis

Tropospheric NO_2 column retrievals from SCIAMACHY produced in the TEMIS project (Boersma et al. 2004) were assimilated in the MACC reanalysis. The largest uncertainties in the retrieval of tropospheric NO_2 are due to errors in the description of clouds, albedo, a-priori profile shape of the trace gas and aerosol description, and assumptions made for aerosol conditions. While some of the uncertainties as those from the a-priori profile shape are canceled in the assimilation by using the averaging kernels of the product, others have to be taken into account. Boersma et al. (2004) found that retrieval results were generally best for regions with strong NO_2 sources or high surface albedo. Here, the errors of the retrieved tropospheric NO_2 columns were of the order 35-60 %. In clean areas with small background concentrations the signal to noise ratio was very small and relative errors were very large. In the MACC reanalysis SCIAMACHY tropospheric NO_2 columns V1.04 data were used until the end of June 2007, and V1.1 data after 11 September 2007 (see Table 1). Version 1.1 used an improved cloud algorithm in the retrieval of tropospheric trace gases and as a result V1.1 tropospheric NO_2 columns were lower than those from V1.04, in particular over heavily polluted areas with low

clouds (Wang et al. 2008). This data change did however not have a large impact on the NO_2 reanalysis fields (see Figure 26 below).

SCIAMACHY has a local overpass time of 10:00. A dedicated observation operator was used to convert the model NO_x to NO_2 in the reanalysis at the time and location of the observations (see also section 2.2.1). This is important, because NO_2 has an atmospheric lifetime in the boundary layer of the order of an hour in summer and about one day in winter and shows large spatial and temporal variability around source areas. Figure 23 shows seasonal mean tropospheric NO_2 columns from the MACC reanalysis. The largest tropospheric NO_2 columns are found in the industrial areas of the NH, with maxima during DJF over Europe, China and the Eastern US. This is the result of seasonal variations in the photochemical lifetime of NO_2 and increased anthropogenic emissions during this season. In the SH there are some hot spots with high NO_2 columns from fossil fuel emissions over the urban areas and coal fired power plants of South Africa, and over cities in South-East Australia. In the tropics, the largest NO_2 values come from biomass burning events during the local dry season. This is during DJF in northern Africa and during JJA and SON in Africa south of the equator. Over South America the strongest NO_2 signal is seen in SON when biomass burning activity is strongest here. Values in this region are lower than over Africa as result of the lower NO_x / CO emission ratio for tropical forest fires compared to Savannah fires. Over Indonesia and North-West Australia the biomass burning signal is strongest during SON. The seasonality of the NO_2 signal from biomass burning is in good agreement with the CO signal (see Figure 3).

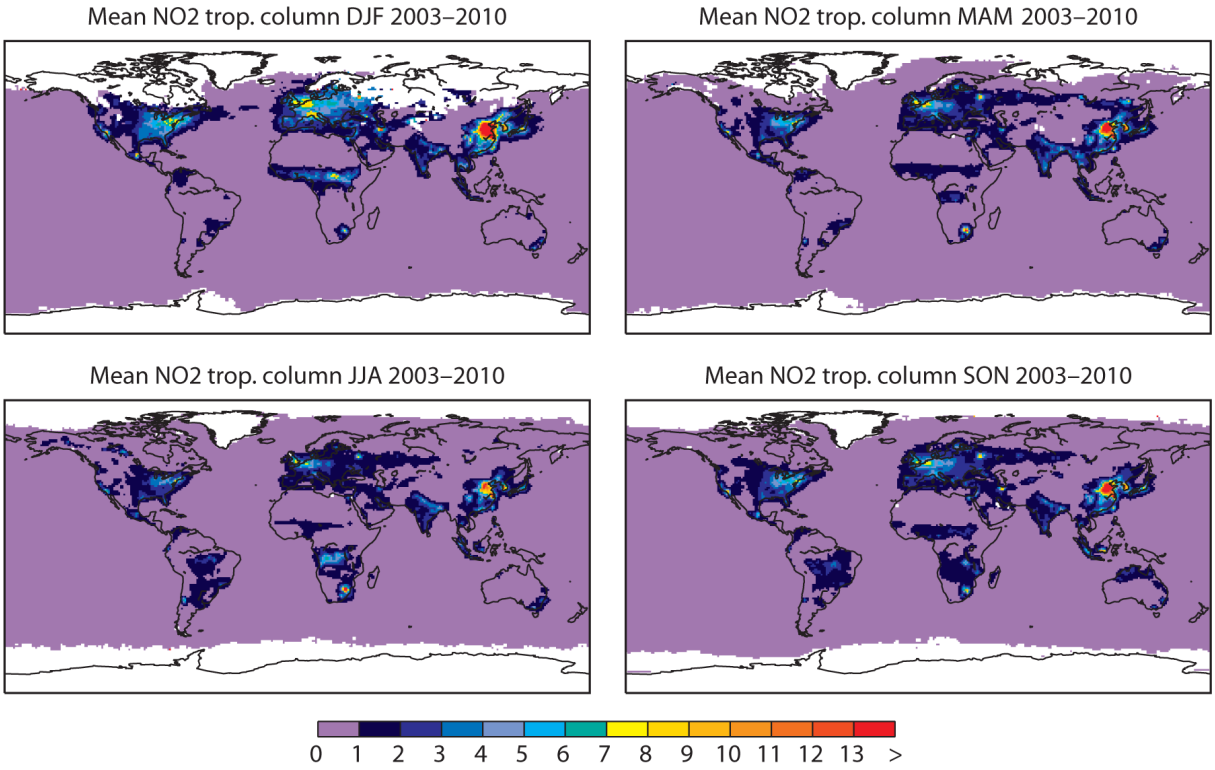

Figure 23: Seasonal mean tropospheric NO_2 columns from the MACC reanalysis in 10^{15}molec/cm^2 for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010. The analysis data were sampled to match the coverage and overpass time of the SCIAMACHY data.

To validate the tropospheric NO2 fields from the MACC reanalysis a comparison was made with SCIAMACHY tropospheric column data retrieved by IUP-Bremen. This dataset is different to the TEMIS SCIAMACHY data that are assimilated in the reanalysis (see Table 1). Even though the two SCIAMACHY retrievals are based on the same level 1 spectral irradiance data, the retrieval products are completely independent, from the spectral fit to the assumptions made on the a priori used for the air mass factor calculations. While this is not ideal in terms of having fully independent validation data, it can provide a critical evaluation of the model performance on a global scale. The IUP retrieval of tropospheric NO2 columns from SCIAMACHY measurements is performed in several steps, starting with the retrieval of NO2 slant column (SC) with the DOAS (Differential Optical Absorption Spectroscopy) technique in the 425-450 nm wavelength windows. The tropospheric SC is calculated by subtracting the stratospheric contribution from the total columns retrieved, assuming that the region over the Pacific (180°E – 220°E) is clean in the lower atmosphere. Airmass factors determined from radiative transfer calculations are used to convert the tropospheric slant columns to vertical columns. As a rough estimate, systematic uncertainties in polluted regions are of the order 30-50%. Further details regarding the retrieval can be found, for example, in Richter and Burrows (2002) and Richter et al. (2005). NO2 tropospheric columns are only determined for clear sky pixels, i.e. for cloud fractions smaller than 20% according to the cloud cover data from the FRESCO database (Koelemeijer et al., 2001; Wang et al. 2008). For the comparison with the IUP-satellite retrievals, the model was sampled at the time and location of the SCIAMACHY overpass and the satellite data been averaged to the model spatial resolution.

Figure 24 shows the seasonal mean differences between tropospheric NO_2 columns from the MACC reanalysis and IUP-SCIAMACHY. As the satellite data are relative to the clean Pacific reference sector, the same correction has been applied to the model data. These plots show that the largest differences are found over the urban areas of the NH, where the MACC reanalysis underestimates NO_2 with respect to the SCIAMACHY data. The largest differences in the NH are found in DJF over East Asia and the Eastern US. The best agreement for the polluted areas of the NH is seen over the US in JJA. Negative differences are also seen for NO_2 hotspots in South Africa. Several reasons might explain the low bias of the reanalysis NO_2 with respect to the IUP-SCIAMACHY data over the urban areas of the NH. Firstly, the anthropogenic NO_x emissions from the MACCity dataset could be too low, but considering the observed downward trend in NO_x emissions (van der A et al., 2008) this is unlikely. Secondly, there is no daily cycle in emissions applied in MOZART-3, which may result in underestimation of emissions at 10:00, i.e. during day time and shortly after the rush hour. Thirdly, because of the short lifetime the information brought into the system by assimilating SCIAMACHY data is lost quickly, and the impact of the NO_2 assimilation is actually small. Finally, with the relatively low resolution of MOZART-3, the maximum emissions in urban areas might not be modelled correctly. Over Northern Europe the reanalysis is higher than the IUP-SCIAMACHY data in MAM and SON. This might be related to a problem in life time leading to too much transport of pollution into these areas.

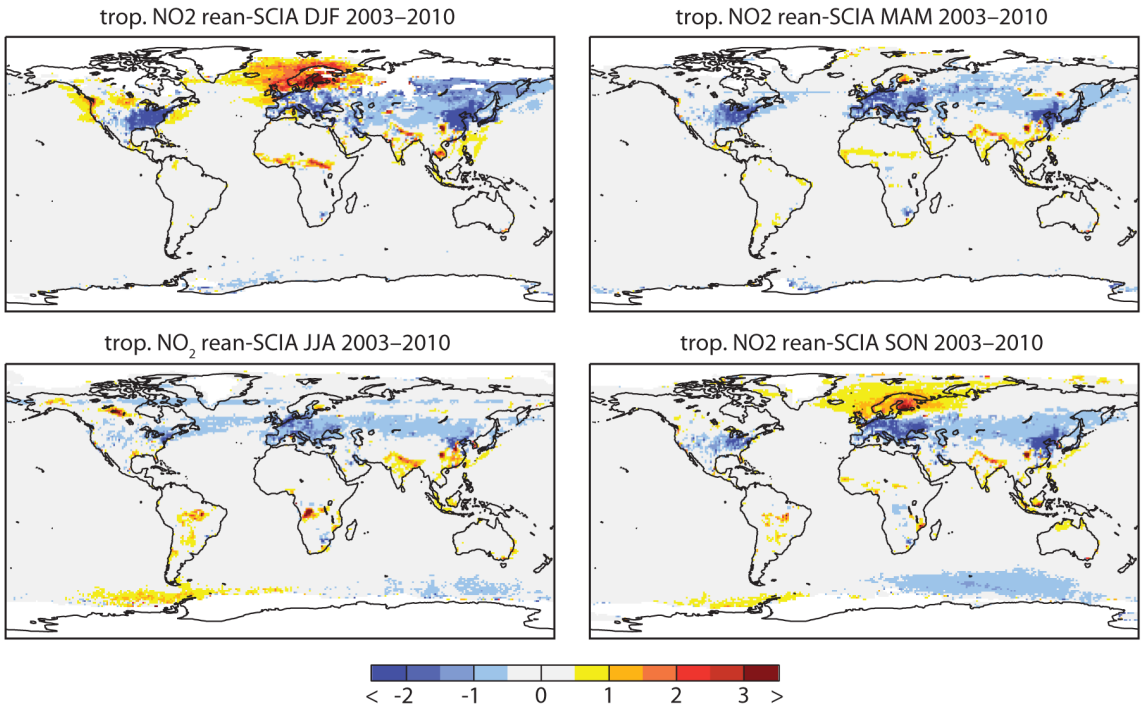

Figure 24: Seasonal mean differences between tropospheric NO_2 vertical columns from the MACC reanalysis and IUP-SCIAMACHY data in 10^{15}molec/cm^2 for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010. For proper comparison with the measurements, the reference sector was also subtracted from the reanalysis data.

Over tropical biomass burning regions the reanalysis NO_2 columns are higher than the IUP-SCIAMACHY data, for example in northern Africa during DJF, in southern Africa in JJA, and in South America in JJA and SON. Over Indonesia the reanalysis generally overestimates the NO_2 columns relative to the data, with the largest deviations in SON. Over north-west Australia the differences are also positive in SON.

Other interesting aspects are clear spots of positive differences for the boreal fires (e.g. Asia in MAM, North America in JJA) that are captured in the reanalysis, but not seen by the satellite. This could be related to a too large NO_x / CO emission ratio used for fires in boreal forests. The reanalysis background concentrations over clean areas agree to within ± 0.5x10^{15}molec/cm^2 with IUP-SCIAMACHY data, which is close to the detection limit of the instrument.

Next, time series of area-averaged monthly mean tropospheric NO_2 from the reanalysis and from IUP-SCIAMACHY data are compared. The areas used for the NO_2 comparison are shown in blue in Figure 25. Only land points are used to calculate the monthly area averages. The MACCity emissions have a positive trend for NO_x emissions over China, and negative trends over Western Europe and the US (Granier et al. 2011a).

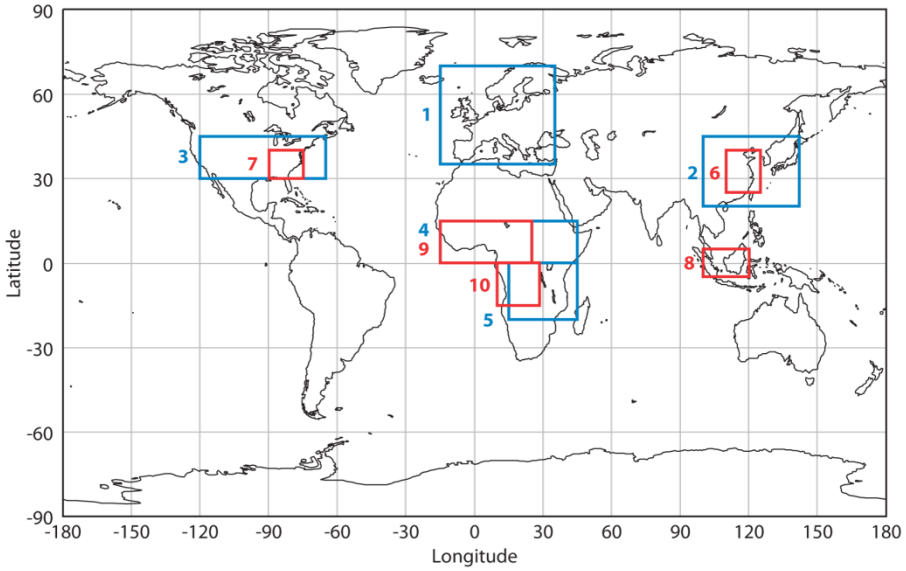

Figure 25: Regions used for validation of MACC NO_2 data (blue) and MACC HCHO data (red) against IUP- SCIAMACHY data. NO_2 regions: (1) Europe, (2) East Asia, (3) US, (4) Northern Africa, and (5) Southern Africa. HCHO regions: (6) China, (7) Eastern US, (8) Indonesia, (9) Northern Africa, and (10) Southern Africa.

Figure 26 shows the time series for the polluted regions of the NH for the reanalysis, the control run and the IUP-SCIAMACHY data. The plots show the impact of the NO_2 assimilation is small, because reanalysis and control run are very similar.. This can have several reasons. First, the NO_2 data have larger observation errors than the CO or O_3 data, and are therefore given less weight in the analysis. Secondly, because of the short lifetime of NO_2 and the fact that NO2 data are only assimilated once per day the impact of the analysis is lost again quickly. A larger impact might be obtained by assimilating data from an additional instrument with different overpass time (e.g. NO_2 from OMI which has a local overpass time of 13:40). Figure 26 shows that the simulation of tropospheric NO_2 in the reanalysis has a realistic seasonal cycle with maxima during the winter months and minima during the summer. There is good agreement during the summer, but winter values are too low. Particularly, over East Asia there is an underestimation of tropospheric NO_2 by about a factor of 2 during winter in the reanalysis, as already seen in Figure 24. For Europe and the US, the agreement between the IUP-SCIAMACHY data and the reanalysis is better. Considering the low resolution of the MOZART-3 CTM, maximum NO_x emissions in urban areas are not modelled adequately and at least some of the differences seen in Figure 26 are likely to be representativeness errors.

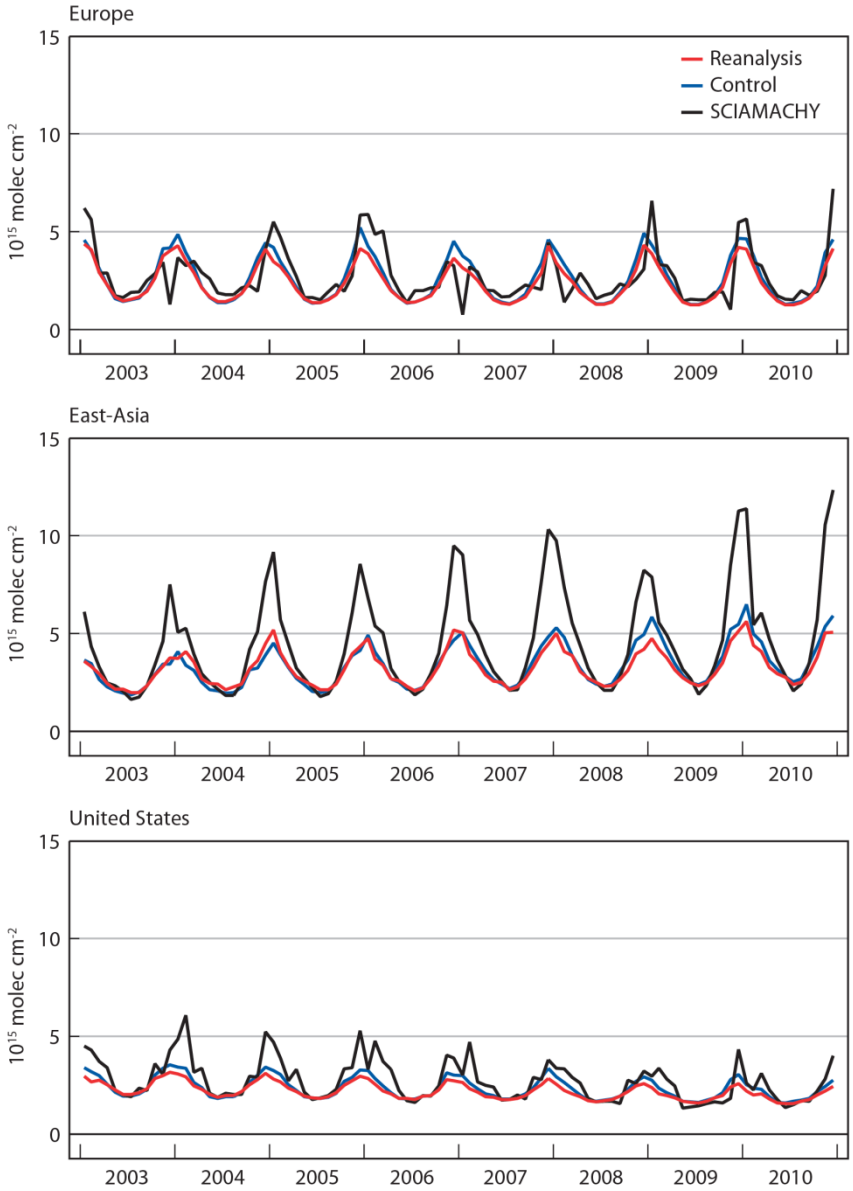

Figure 26: Time series of monthly mean area averaged tropospheric NO_2 columns in 10^{15}molec/cm^2 from the MACC reanalysis (red), the control run (blue) and IUP-SCIAMACHY data (black) for the period 2003 to 2010 for Europe (top), East Asia (middle) and the US (bottom). Only land points were used in these calculations.

Figure 27 shows time series of monthly mean tropospheric NO_2 columns from the reanalysis and IUP-SCIAMACHY for the biomass burning areas of Northern and Southern Africa. Again the differences between the reanalysis and the control run are negligible. The figure shows that the reanalysis has the right seasonality with maximum values in Northern Africa during DJF and in Southern Africa during June to September. However, the reanalysis overestimates the tropospheric NO_2 columns during the biomass burning seasons, as already seen in Figure 24. The overestimation during the biomass burning season could again be related to a too large NO_x emission factors used for fires.

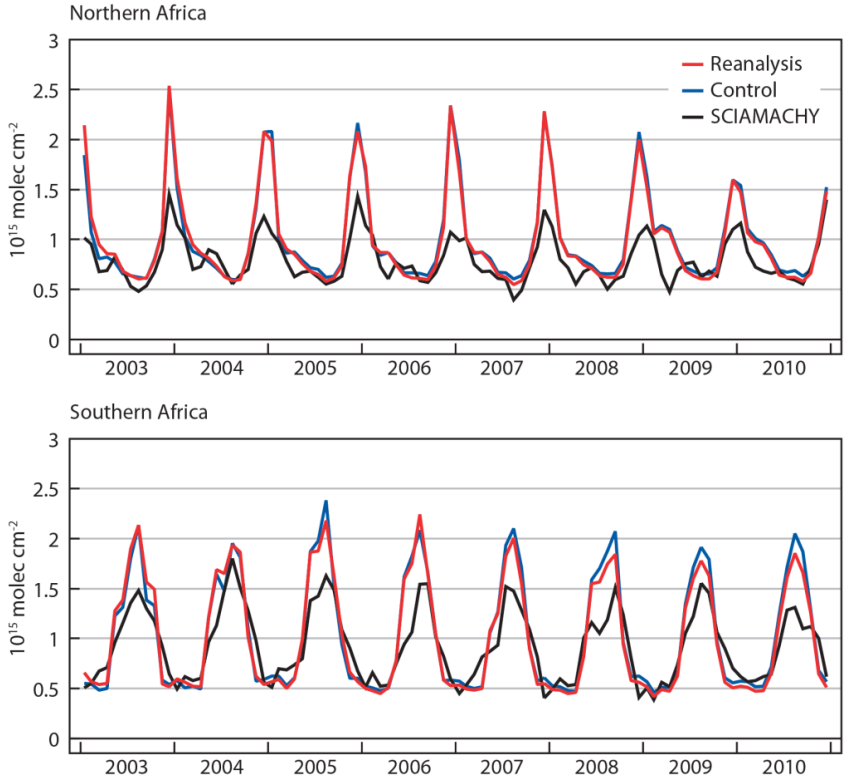

Figure 27: Time series of monthly mean area averaged tropospheric NO_2 columns in 10^{15}molec/cm^2 from the MACC reanalysis (red), the control run (blue) and from IUP-SCIAMACHY data (black) for the period 2003 to 2010 for Northern Africa (top) and Southern Africa (bottom). Only land points were used in these calculations.

The correlation between the global and seasonally averaged reanalysis NO_2 and the IUP-SCIAMACHY data (over all surfaces) averaged over the years 2003 to 2010 is 0.83 in DJF, 0.87 in MAM, 0.82 in JJA and 0.84 in SON, indicating a good spatial agreement between the reanalysis and the satellite retrievals. Table 2 shows correlations between the seasonal trends of reanalysis and IUP-SCIAMACHY data for the five NO_2 regions (data considered only from land grid boxes). The high correlation over East Asia shows that the reanalysis captures the seasonal NO_2 trend well, despite the winter time biases.

Region	World	Europe	East-Asia	US	Northern Africa	Southern Africa
Correlation	0.89	0.72	0.96	0.85	0.86	0.91

Table 2: Correlation between the seasonal trends of the monthly averaged MACC reanalysis and IUP-SCIAMACHY NO_2 data over land.

Table 3 summarizes the seasonal mean biases between IUP-SCIAMACHY NO_2 data and the reanalysis fields for the five regions discussed above. For this table daily differences between the reanalysis and SCIAMACHY were calculated over land and then averaged over the years 2003-2010 for the corresponding months and regions. The table confirms that the largest negative biases can be seen over the industrial areas of the NH, with maxima during the winter months. The biases in the African biomass burning areas are smaller. Over Northern Africa the bias is positive apart from JJA

when it is small and negative. In Southern Africa biases are largest and positive in JJA and small and negative in the other seasons.

	DJF		MAM		JJA		SON	
	bias	RMS	bias	RMS	bias	RMS	Bias	RMS
East-Asia	-4.10	6.91	-1.12	2.60	-0.01	1.34	-1.80	3.58
Europe	-0.41	3.09	-0.62	1.34	-0.34	0.88	-0.05	1.73
US	-1.12	2.70	-0.61	1.22	-0.05	0.66	-0.22	0.90
Northern Africa	0.53	0.67	0.20	0.40	-0.01	0.22	0.16	0.33
Southern Africa	-0.02	0.11	-0.02	0.21	0.53	0.86	-0.01	0.47

Table 3: Seasonal mean biases and rms errors in 10^{15}molec/cm^2 of MACC reanalysis and IUP SCIAMACHY NO_2 averaged over the years 2003 to 2010 for the 5 regions illustrated above. Only land points were used in the calculations.

Figure 28 shows seasonal mean zonal average NO_x altitude cross sections from the MACC reanalysis to illustrate the vertical structure of the NO_x field. NO_x concentrations are largest near the surface and fall off rapidly with height. This illustrates the dominating importance of the emissions for the NO_x field.

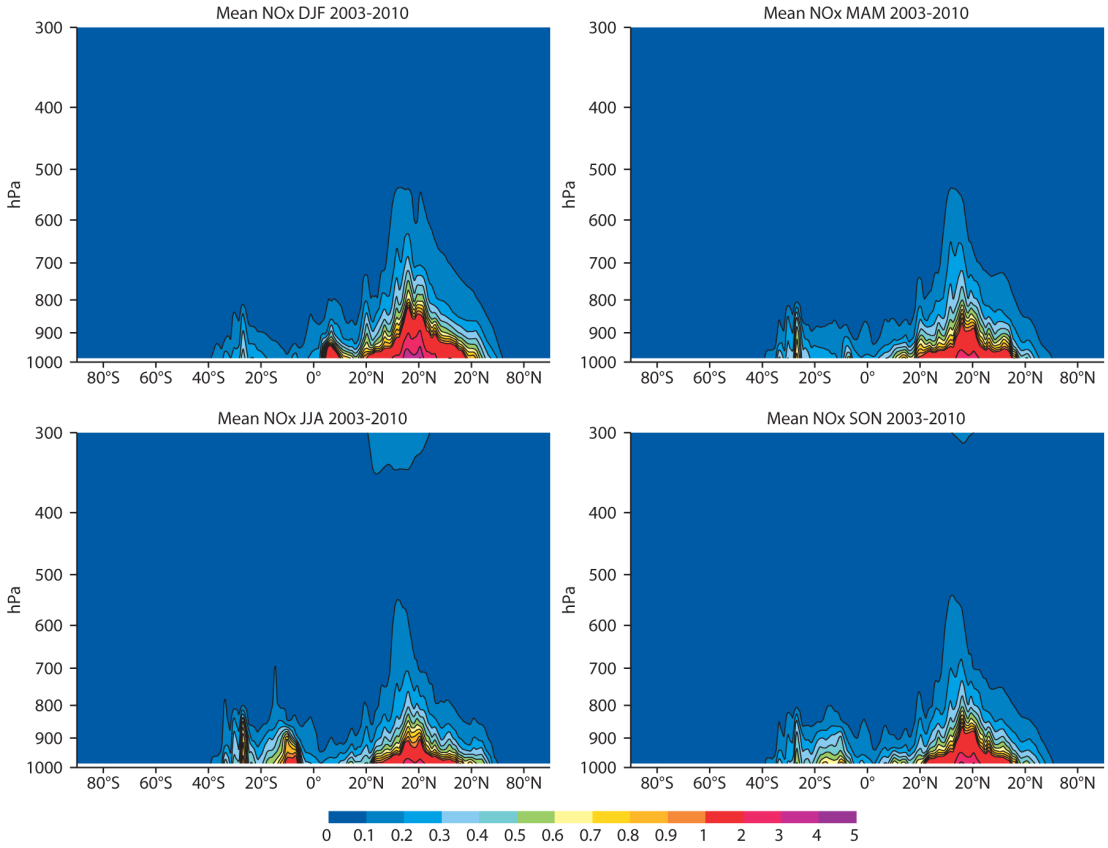

Figure 28: Seasonal mean zonal average NO_x cross sections from the MACC reanalysis in ppb for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010.

3.4 HCHO analysis

HCHO data are not assimilated in the MACC reanalysis because the data quality of individual satellite retrievals is not sufficient. Monthly mean observations generally have a total error of 20-40%, but individual observations can have much large errors (De Smedt et al. 2008). Over Europe, for example, the mean HCHO column is smaller than the random error of SCIAMACHY observations which does not favour assimilating individual observations. Hence, the HCHO reanalysis fields are entirely determined by the MOZART-3 chemistry, the MACCity emissions and the atmospheric transport.

Figure 29 shows seasonal mean tropospheric HCHO columns from the MACC reanalysis. The largest HCHO columns are found in the tropics and reflect the regions of high biogenic VOC emissions and biomass burning. High values are also found over the South-Eastern US in JJA and are indicative of the oxidation of isoprene emitted during the growing season in the summer. Values over Europe are much lower but also peak in JJA. Figure 29 agrees well with the global maps shown in De Smedt et al. (2008). Over South America, the maximum HCHO values are seen during the fire season (August to November). In Africa, north of the equator, maximum values are found during the main fire season in DJF. South of the equator, the maximum values are observed in JJA. Over South-East Asia the largest HCHO concentrations occur during MAM and JJA, and are likely to be associated with biogenic VOC emissions. For the region of Indonesia the HCHO concentrations are always high, with a minimum observed during DJF. Over northern Australia HCHO concentrations are largest during the main growing season SON and DJF.

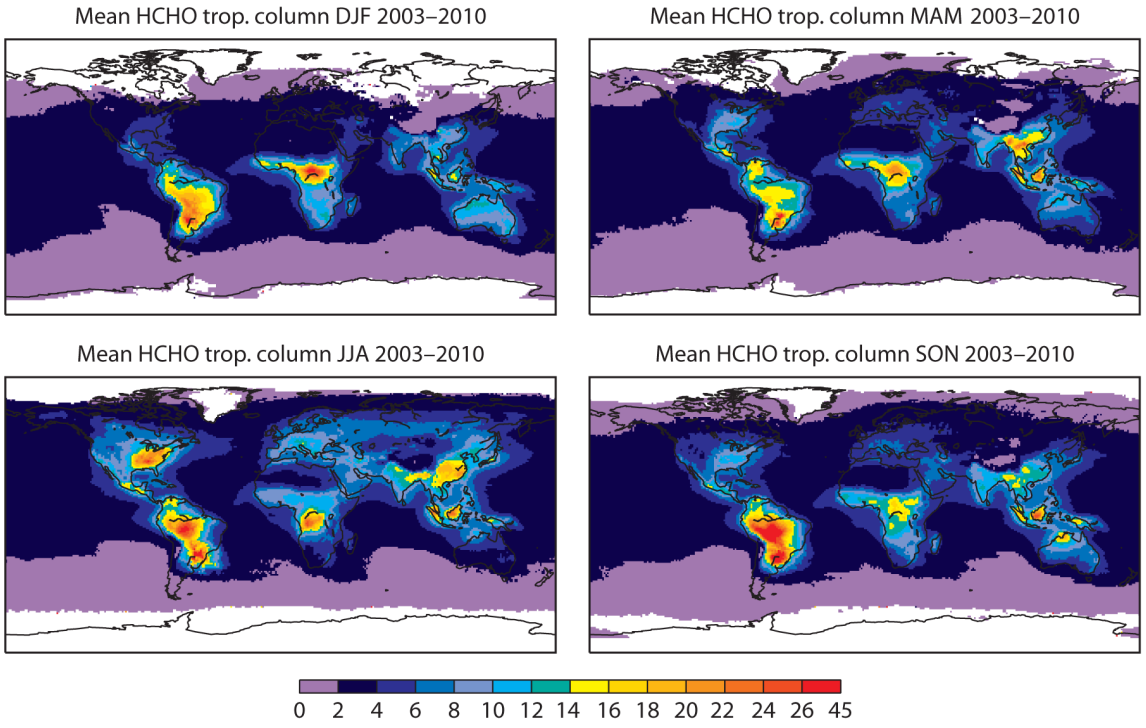

Figure 29: Seasonal mean tropospheric HCHO columns from the MACC reanalysis in 10^{15}molec/cm^2 for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010. The analysis data were sampled to match the coverage and overpass time of the SCIAMACHY data.

To validate the tropospheric HCHO columns from the reanalysis they are compared with SCIAMACHY HCHO data retrieved by IUP-Bremen. The IUP retrieval of HCHO columns follows a similar approach as the NO_2 method described in the previous section, using also the approach of reference sector to correct for instrumental drift and appropriate airmass factors to convert the slant to vertical columns. To account for the tropospheric HCHO amount present over the region between 180–200°E used for normalization, a mean value of 3.5×10^{15} mol cm^{-2} is added. Further details on the retrieval can be found in Wittrock (2006) and Wittrock et al. (2006). The total error on the monthly and regionally averaged data is between 20-40%, and the detection limit is 2×10^{15}molec/cm^2. The uncertainty in the mean of the observations is estimated to be of the order 10^{16} molec/cm^2 (Wittrock 2006). However, for HCHO hotspots, both the absolute values and the seasonality can be retrieved with confidence. Figure 30 shows the seasonal mean differences between IUP HCHO tropospheric columns and the reanalysis fields. The figure shows that there are limitations with the satellite retrievals at low solar elevations, that lead to large differences and large scatter in the NH during DJF and in the SH during JJA. The difference plots also show scatter over the area of the South Atlantic anomaly in JJA. This localized discrepancy is due to an artifact in the observations, because here the SCIAMACHY instrument is exposed to high energy solar particles, leading to a reduced signal to noise ratio and a large scatter in the data. Figure 30 shows that the reanalysis overestimates the HCHO tropospheric columns with respect to SCIAMACHY in regions with high biogenic emissions and biomass burning. This is the case in the Eastern US, Europe and China during JJA, and in northern Africa, South America and Indonesia throughout the year. In southern Africa the seasonal mean differences between the IUP-SCIAMACHY data and the reanalysis are small.

The plots indicate that the reanalysis underestimates HCHO for background concentrations over the oceans, where oxidation of methane is the main source of HCHO. This could point to a problem with the retrieval over sea, but the HCHO values here are close to the detection limit of the instrument and the differences hence not very meaningful.

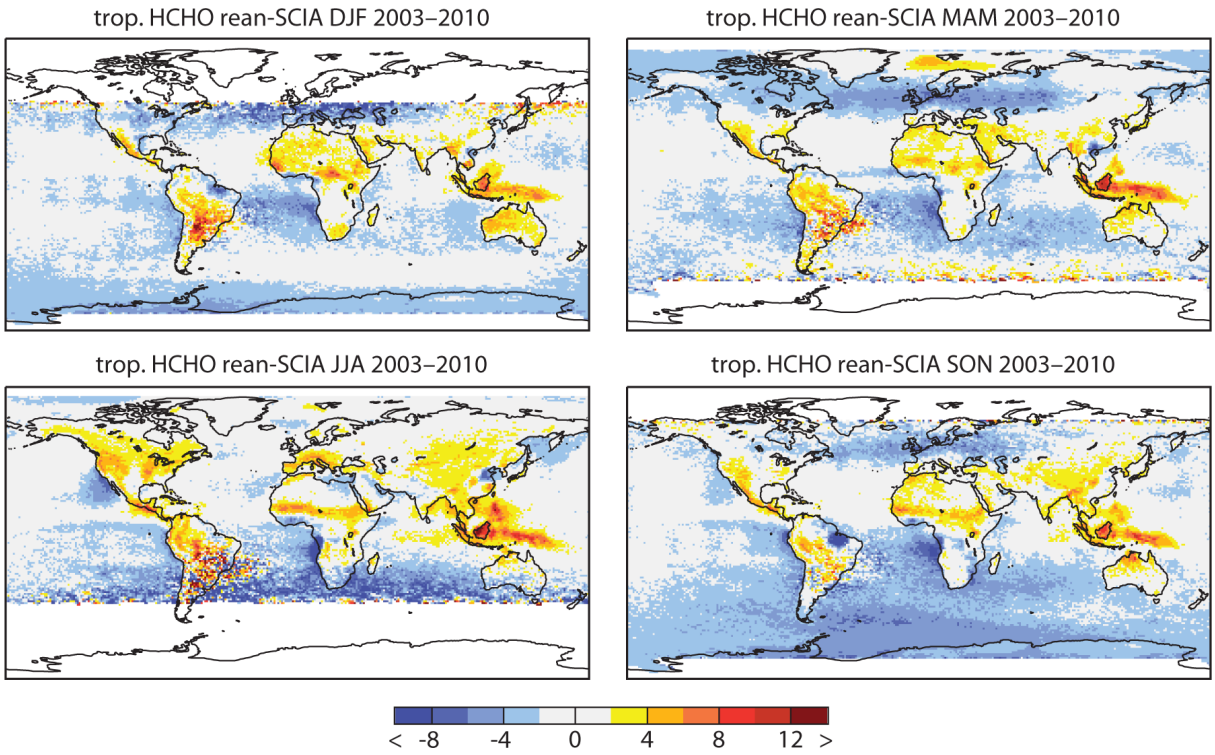

Figure 30: Seasonal mean differences between tropospheric HCHO columns from the MACC reanalysis and IUP-SCIAMACHY data in 10^{15}molec/cm2 for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010.

Next, time series of monthly mean tropospheric HCHO from the reanalysis and IUP-SCIAMACHY data are compared. The regions used for the HCHO comparison are shown in red in Figure 25 and focus on areas with a strong HCHO signal. Only land points are used to calculate the monthly area averages.The differences between the reanalysis and the control are very small which was to be expected because no HCHO data were assimilated in the MACC reanalysis. For this reason the control run is not shown in the timeseries plots below. Figure 31 shows time series over China and the Eastern US, regions dominated by biogenic emissions with some anthropogenic input. Figure 32 shows time series for Northern Africa, Southern Africa, and Indonesia, i.e. regions with biogenic sources and biomass burning. The seasonality and magnitude is well captured for China and the Eastern US. In the biomass burning areas of Northern Africa, the reanalysis overestimates HCHO during the main fire season, but the agreement between reanalysis and data in Southern Africa is good. The largest differences are seen over Indonesia where the reanalysis is almost constantly higher than the satellite data by at least $5x10^{15}$ molec/cm^2. However, during the Indonesian fires in October 2006 the reanalysis and the satellite data show good agreement, capturing the very high values of HCHO registered for this month.

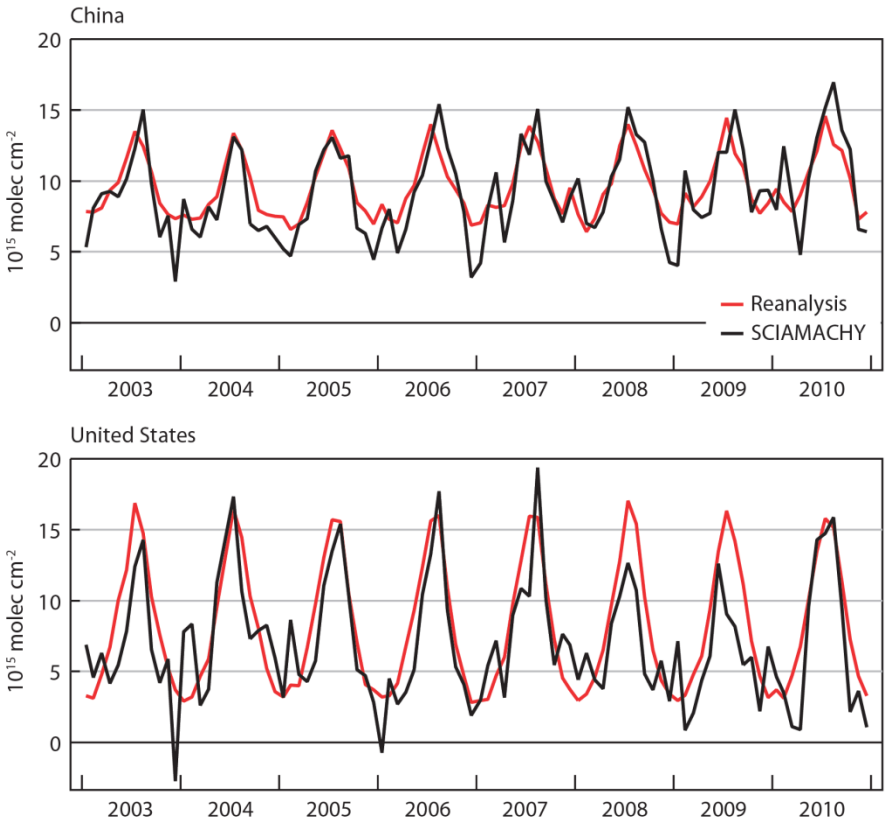

Figure 31: Time series of monthly mean area averaged tropospheric HCHO columns in 10^{15}molec/cm^2 from the MACC reanalysis (red), the control run (blue) and IUP-SCIAMACHY (black) for China (top) and the Eastern US (bottom). Only land points were used in these calculations.

The regional correlations between the monthly mean reanalysis time series and the IUP-SCIAMACHY data are shown in Table 4 (data considered only from land grid boxes). They are low for the world, confirming that a meaningful statement can only be made for regions with large HCHO concentrations. In the five regions discussed above the correlations are lowest over Indonesia and Northern Africa, confirming what was seen in Figure 32.

Region	World	China	US	Indonesia	Northern Africa	Southern Africa
Correlation	0.24	0.85	0.81	0.69	0.54	0.84

Table 4: Correlation between the seasonal cycles of monthly averaged MACC reanalysis and IUP-SCIAMACHY HCHO data over land.

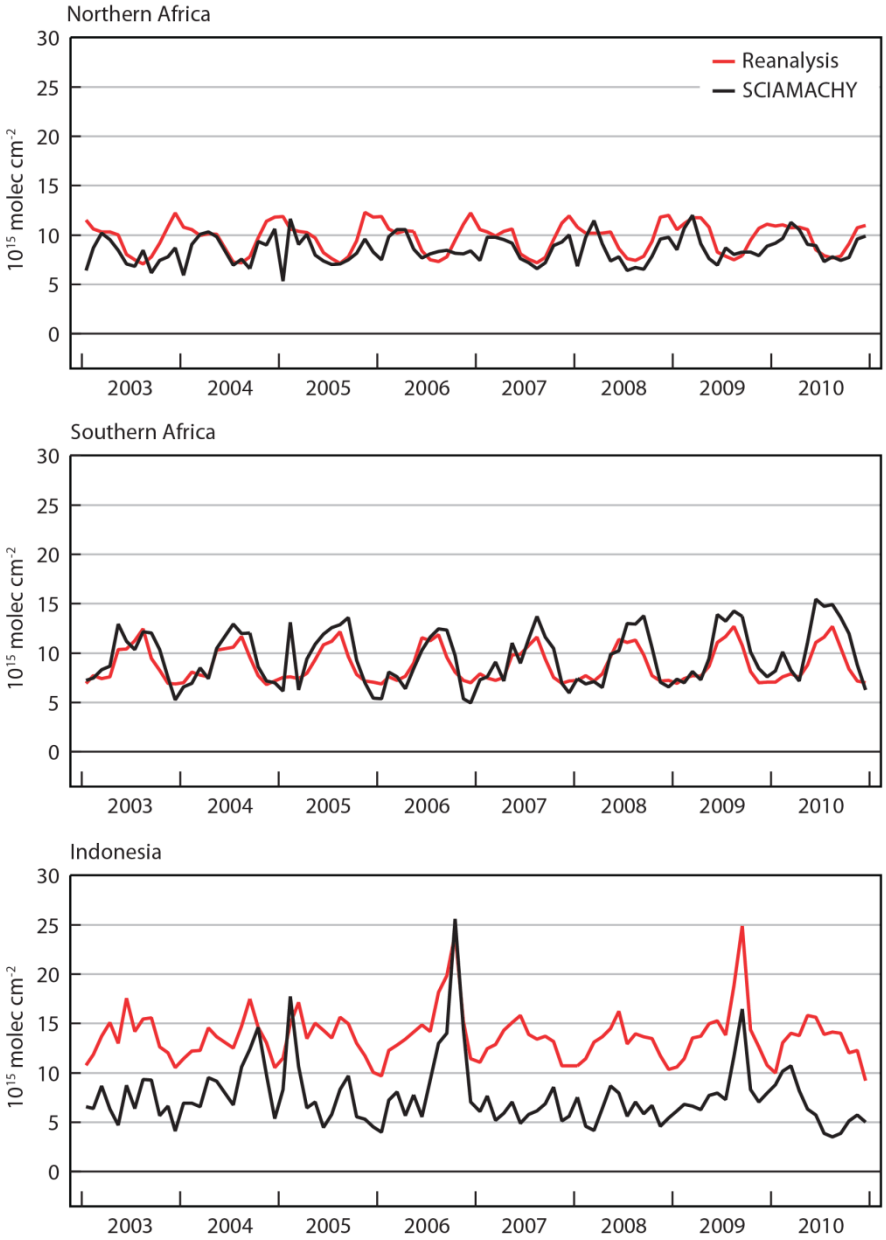

Figure 32: Time series of monthly mean area averaged tropospheric HCHO columns in 10^{15}molec/cm^2 from the MACC reanalysis (red), the control run (blue) and IUP-SCIAMACHY (black) for Northern Africa (top), Southern Africa (middle) and Indonesia (bottom). Only land points were used in these calculations.

Figure 33 shows seasonal mean zonal average cross sections of HCHO from the reanalysis. It shows that the largest concentrations are confined to the boundary layer and decrease with height. There is some sign of transport of HCHO into the upper troposphere by deep convection in the tropics. In DJF the zonal mean maximum is found around 10°N, pointing to high HCHO values in Africa north of the equator. In JJA the highest values are seen around 40°N (North American signal) and around 10°S (combination of South America, Africa and Indonesia).

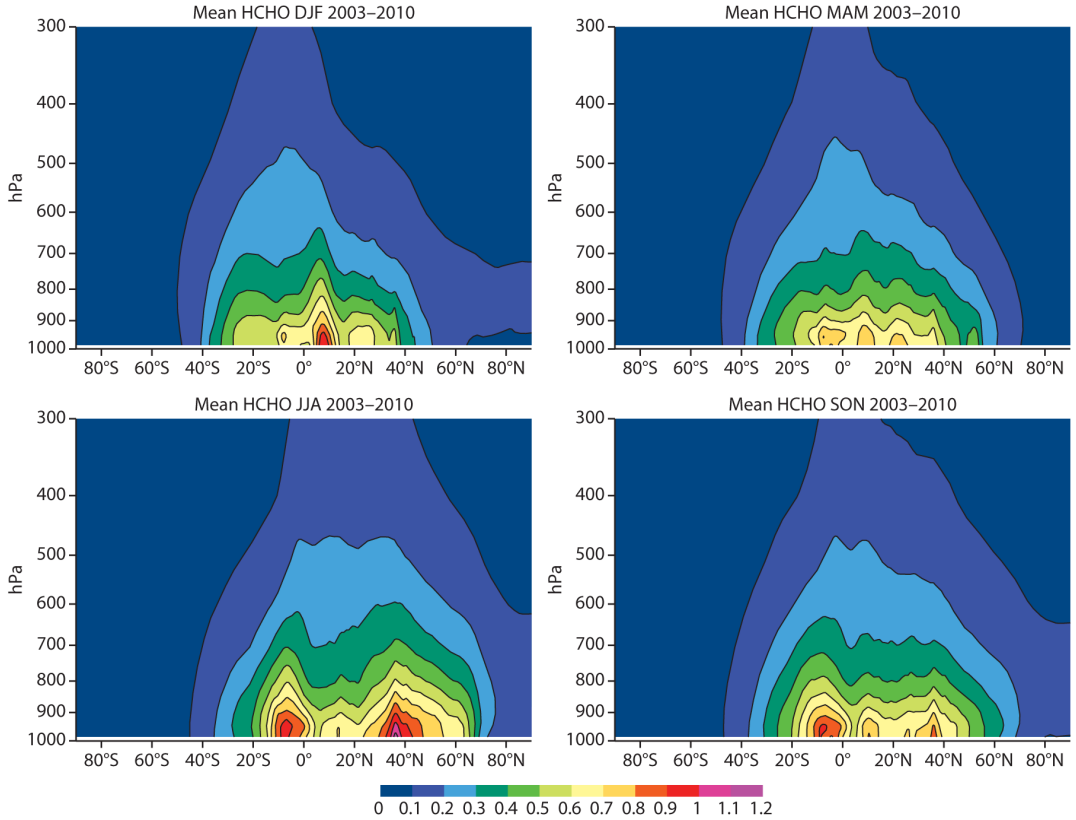

Figure 33: Seasonal mean zonal average HCHO cross sections from the MACC reanalysis in ppbm for DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right) averaged over the years 2003 to 2010.

HCHO	DJF		MAM		JJA		SON	
	Bias	RMS	bias	RMS	bias	RMS	bias	RMS
China	0.50	1.31	1.02	1.27	-1.23	3.74	0.03	1.64
East-US	-1.90	1.48	1.43	1.25	2.25	1.70	1.55	1.09
Indonesia	5.01	2.34	7.83	2.55	8.57	2.81	6.29	2.05
Northern Africa	2.56	3.15	1.00	2.54	0.23	2.94	1.40	2.73
Southern Africa	-0.42	1.22	-0.43	2.26	-1.24	4.12	-1.80	2.38

Table 5: Seasonal mean biases and rms error in 10^{15}molec/cm^2 of MACC reanalysis and IUP-SCIAMACHY HCHO averaged over the years 2003 to 2010 for the 5 regions discussed above. Only land points were used in the calculation.

3.5 Summary of issues with the reactive gases reanalysis fields

There are some issues with the reactive gases analysis fields that a user should be aware of, because they cause discontinuities in the data set:

- The biomass burning emissions were changed on 20090101 from a preliminary version of GFED3 to GFAS Version 1.0. The GFAS Version 1.0 emissions budget is about 18 % higher than those used during the first 6 years.
- Using MACCity emissions to provide lower boundary conditions for the MOZART-3 CTM lead to too low CO analysis values, especially in the boundary layer
- After 20100323 NRT MOPITT CO data were used in the reanalysis instead of the offline product. This change did not have a noticeable impact on the reanalysis fields.
- Assimilation of IASI CO after 20080401 led to some changes in the CO field.
- Using variational bias correction for MLS ozone profiles led to increased tropospheric ozone and changes to ozone above 15 hPa. However, it did not affect the total column ozone field. These drifts stopped on 1 January 2008 when the bias correction was switched off for MLS and afterwards agreement with independent ozone sondes and MOZAIC data was improved.
- Using NRT MLS data instead of the offline product after 20090316 resulted in larger departures in the upper troposphere and lower stratosphere, because the lowest 3 layers (68-100, 100-146, and 146-215 hPa) of the MLS data could not be used.
- NRT OMI ozone columns were assimilated instead of the off-line product between 20070321-20071231. This did not have a noticeable impact on the ozone analysis.
- A model upgrade to MOZART 3.5 was implemented on 20090101, which slightly improved the representation of the ozone hole in the control run but did not affect the other model fields
- NO_2 SCIAMACHY retrievals V1.04 were assimilated until 20070630, SCIAMACHY V1.1 data were assimilated after 20070911, but this only has a minor impact on the analysis fields.

4 Conclusions

A data assimilation system for global reactive gases, aerosols and greenhouse gases was developed and consolidated as part of the EU funded GEMS and MACC projects. This system was used in the MACC project to produce an 8-year long reanalysis of atmospheric composition data for the period 2003-2010, by assimilating satellite data to constrain O_3, CO, NO_2, CO_2, CH_4, and aerosol optical depth. The reanalysis data are constrained in a consistent way by observations and the model simulation. This paper describes the assimilation system for the reactive gases used in the MACC reanalysis and presents some validation results of the reanalysis fields for CO, O_3, NO_2 and HCHO. Total column values are generally in very good agreement with independent observations, but profiles can show some problems in the boundary layer where concentrations are dominated by emissions. There are some discontinuities in the dataset related to instrument changes and issues with the bias correction of ozone data.

Assimilating MOPITT and IASI CO retrievals in the MACC reanalysis leads to an improved CO field compared to a MOZART-3 standalone run carried out with the MACC configuration. The reanalysis CO field has a realistic seasonal cycle and inter-hemispheric differences. Total column values in 2008 to 2010 are low compared to the satellite retrievals at high northern latitudes, but in other areas the agreement with MOPITT and IASI is good. Using MACCity emissions to provide boundary conditions for the MOZART-3 CTM leads to tropospheric CO values that are 10-20% too low compared to MOZAIC data through most of the troposphere. In the boundary layer at urban sites affected by air pollution the negative biases are larger, suggesting an underestimation of CO or precursor emissions. Surface CO from the MACC reanalysis agrees well with NOAA/GMD observations which indicates that, despite uncertainties of satellite measurements in the lower troposphere, assimilation of CO data from such products can lead to a good representation of surface CO concentrations for unpolluted regions.

Comparison with independent data have shown ozone from the MACC reanalysis to be considerably better than a free running MOZART-3 CTM. Stratospheric ozone fields from the MACC reanalysis agree with ozone sondes and ACE data to within ±10 % in most situations. In the troposphere the reanalysis shows biases of -5% to +10% with respect to ozone sondes and MOZAIC aircraft observations in the extratropics, but has larger negative biases in the tropics (up to -40% around 100 hPa). These biases are partly due to biases in the underlying MOZART-3 CTM, but a time varying bias in the troposphere is the result of using the variational bias correction scheme without MLS as an anchor before 2008 (see also section 2.3.2). Area averaged total column ozone agrees with data from KNMI's multi sensor reanalysis product to within a few percent. Surface ozone from the reanalysis agrees with EMEP surface observations over Europe to within ± (5-10) ppbv. However, there are some diurnal variations in the surface ozone biases that need to be investigated further. The biases of the MACC reanalysis with respect to ozone sondes are smaller than biases of ERA Interim ozone fields.

Assimilating NO_2 retrievals from SCIAMACHY in the MACC reanalysis has only little impact, and the NO_2 fields from the reanalysis and the control run are very similar. A possible reason for this is the short lifetime of NO2, so that the impact of the data have in the analysis is lost again quickly. NO2 fields from the reanalysis show the right seasonality over polluted urban areas of the NH and over tropical biomass burning areas, but underestimate wintertime NO_2 maxima over anthropogenic regions and overestimate NO_2 in Northern and Southern Africa during the tropical biomass burning seasons.

Tropospheric HCHO is quite well simulated in the MACC reanalysis even though no satellite data are assimilated. It agrees well with independent IUP-SCIAMACHY observations over regions dominated by biogenic emissions with some anthropogenic input, such as the Eastern US and China, and also over African regions influenced by biogenic sources and biomass burning. Over Indonesia, however, the reanalysis has a large positive bias that is not seen in the control run, but the high HCHO values observed during the Indonesian fires of 2006 are very well captured by the reanalysis.

The MACC reanalysis is a valuable 8-year long atmospheric composition data set that can, for example, be used as boundary conditions for regional models, climatological studies or for model evaluation. In addition to the four GRG IFS fields analyzed in this paper more chemical species are available from the MOZART-3 CTM output. The MACC model and assimilation system is also run in NRT to produce daily analyses and 5-day forecasts of reactive gases and aerosols. Data from the

MACC reanalysis and the NRT analysis are available from the MACC data server http://www.gmes-atmosphere.eu. Further validation results from the MACC reanalysis can be found on the MACC verification web page http://www.gmes-atmosphere.eu/services/gac/global_verification/.

Since November 2011, the MACC-II project has begun as a successor to MACC. This project will continue to deliver the daily analyses and forecasts of atmospheric composition. There are no plans for a new reanalysis in MACC-II, but the MACC reanalysis will be extended to more recent years and selected periods will be rerun to test changes to the assimilation system and new input datasets.

Acknowledgements

MACC was funded by the European Commission under the EU Seventh Research Framework Programme, contract number 218793. The CO products from IASI are available from the Ether French atmospheric database (http://ether.ipsl.jussieu.fr). We thank Doug Kinnison at NCAR, Boulder for support with the MOZART model. We acknowledge the free use of tropospheric NO_2 column data from the SCIAMACHY sensor from www.temis.nl. MOPITT data are from NASA. Mihalis Vrekoussis acknowledges European Union research fellowship (FP7-PEOPLE-2009-RG Grant, PERG-GA-2009-256391).

References

Abbot, D. S., Palmer, P. I., Martin, R. V., Chance, K. V., Jacob, D. J., and Guenther, A.: Seasonal and interannual variability of North American isoprene emissions as determined by formaldehyde column measurements from space, *Geophys. Res. Lett.*, **30**(17), 1886, doi:10.1029/2003GL017336, 2003.

Andersson, E. and Järvinen, H.: Variational quality control. *Quart. J. R. Meteorol. Soc.*, **125**, 697-722, 1999.

Arellano, A. F., K. Raeder, J. Anderson, P. Hess, L. Emmons, D. Edwards, G. Pfister, T. Campos, and Sachse, G.: Evaluating model performance of an ensemble-based chemical data assimilation system during INTEX-B field mission, *Atmos. Chem. Phys.*, **7**, 5695-5710, 2007.

Atkinson, R.: Gas-phase tropospheric chemistry of organic compounds. *J. Phys. Chem. Ref. Data Monogr.*, **2**, 13-46, 1994.

Auligné, T., McNally, A. P., and Dee, D. P.: Adaptive bias correction for satellite data in a numerical weather prediction system, *Quart. J. R. Meteorol. Soc.*, **133**, 631–642, 2007.

Baier, F., Lefever, K. , Errera, E. and van der A, R.: Validation of G-RG stratospheric trace gas services, subtask 1.5, MACC validation report document. Available from http://www.gmes-atmosphere.eu/documents/deliverables/g-rg , 2011.

Baklanov, A., Korsholm, U., Mahura, A., Petersen, C., and Gross, A.: ENVIRO-HIRLAM: on-line coupled modelling of urban meteorology and air pollution, *Advances in Science and Research*, **2**, 41–46, 2008.

Barkley, M.: Description of MEGAN biogenic VOC emissions in GEOS-Chem, http://acmg.seas.harvard.edu/geos/wiki_docs/emissions/megan.pdf , 2010.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R. J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J. W., Kinne, S., Mangold, A., Razinger, M., Simmons, A. J., Suttie, M., and the GEMS-AER team: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Data Assimilation. *J. Geophys. Res*., D13205, **114**, doi:10.1020/2008JD011115, 2008.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, *J. Geophys. Res*., **106**, 23,073– 23,096, 2001

Bhartia, P. K., McPeters, R. D., Mateer, C. L., Flynn, L. E., and Wellemeyer, C., Algorithm for the estimation of vertical ozone profiles from the backscattered ultraviolet technique, *J. Geophys. Res*., **101**(D13), 18,793– 18,806, 1996.

Bhartia, P. K. and Wellemeyer, C.: TOMS-V8 Total O3 algorithm, in: *OMI ozone product ATBD Volume II, NASA Goddard Space Flight Center*, Greenbelt, MD, USA, 2002.

Boersma, K.F., Eskes, H.J., and Brinksma, E.J.: Error analysis for tropospheric NO2 retrieval from space. *J. Geophys. Res*., **109**, D04311, doi:10.1029/2003JD003962, 2004.

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van den Oord, G. H. J. , Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: Near-real time retrieval of tropospheric NO2 from OMI, *Atmos. Chem. Phys*., **7**, 2103–2118, 2007.

Boersma, K. F., Jacob, D.J., Eskes, H. J., Pinder, R. W., Wang, J., and van der A, R. J.: Intercomparison of SCIAMACHY and OMI tropospheric NO2 columns: Observing the diurnal evolution of chemistry and emissions from space, *J. Geophys. Res*., **113**, D16S26, doi:10.1029/2007JD008816, 2008.

K. F. Boersma, Eskes, H. J., 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., and Brunner, D.: An improved tropospheric NO2 column retrieval algorithm for the Ozone Monitoring Instrument. *Atmos. Meas. Tech*., **4**, 1905–1928, 2011.

Brewer, A.W.: Evidence for a world circulation provided my measurements of helium and water vapour distribution in the stratosphere. *Quart. J. R. Meteorol. Soc.*, **75**, 351-363, 1949.

Cariolle, D. and Teyssèdre, H.: A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations. *Atmos. Chem. Phys*., **7**, 2183-2196, 2007.

Carli, B., Alpaslan, D., Carlotti, M., Castelli, E., Ceccherini, S., Dinelli, B. M.,Dudhia, A., Flaud, J. M., H¨opfner, M., Jay, V., Magnani, L., Oelhaf, H., Payne, V., Piccolo, C., Prosperi, M., Raspollini, P., Ridolfi, M., Remedios, J. and Spang, R.: First results from MIPAS/ENVISAT with operational Level 2 code. *Adv. Space Res*., **33**, 7, 1012-1019, doi:10.1016/S0273-1177(03)00584-2, 2004.

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, *Atmos. Chem. Phys*., **9**, 6041-6054, 2009.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. R. Meteorol. Soc.*, **120**, 1367-1388, 1994.

Dee, D. P.: Variational bias correction of radiance data in the ECMWF system, *Proceedings of the ECMWF workshop on assimilation of high spectral resolution sounders in NWP*, Reading, UK, 28 June–1 July 2004, 97–112, 2004.

Dee, D. P. and Uppala, S.: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quart. J. R. Meteorol. Soc.,* **135**, 1830–1841, 2009.

Dee, D P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M. , van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitarta, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553-597, 2011.

Deeter: Calculation and application of MOPITT averaging kernels. Available from http://www.acd.ucar.edu/mopitt/data-interpretation.shtml , 2002.

Deeter, M. N., Emmons, L. K., Francis, G. L., et al.: Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument, *J. Geophys. Res*., **108**, 4399, doi:10.1029/ 2002JD003186, 2003.

Deeter, M. N., Edwards, D. P., Gille, J. C., and Drummond, J. R.: CO retrievals based on MOPITT near-infrared observations, *J. Geophys. Res.*, **114**, D04303, doi:10.1029/2008JD010872, 2009.

Deeter, M. N., et al.: The MOPITT version 4 CO product: Algorithm enhancements, validation, and long-term stability, *J. Geophys. Res*., **115**, D07306, doi:10.1029/2009JD013005, 2010.

Dethof, A.: Assimilation of ozone retrievals from the MIPAS instrument on board ENVISAT. *ECMWF Tech. Memorandum* 428, available from http://www.ecmwf.int/publications/library/do/references/list/14 , 2003.

Dethof, A. and Hólm, E.V.: Ozone assimilation in the ERA-40 reanalysis project. Quart . J. Roy. Met. Soc., 130, 2851-2872, 2004.

De Smedt, I., Müller, J.F., Stavrakou, T., van der A, R., Eskes, H., and Van Roozendael, M.: Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors. *Atmos. Chem. Phys*., **8**, 4947-4963, 2008.

Dobson, G.M.: Origin and distribution of the polyatomic molecules in the atmosphere, *Proc. Roy. Soc. London*, **A236**, 187-193, 1956.

Dragani, R.: On the quality of the ERA-Interim ozone reanalyses: comparisons with in situ data. *ECMWF ERA Report Series*, 2, avaiable from http://www.ecmwf.int/publications/library/do/references/list/782009 , 2010.

Dragani, R.: On the quality of the ERA-Interim ozone reanalyses: comparisons with satellite data. *Quart. J. Roy. Meteor. Soc.,* **137**: 1312–1326. doi: 10.1002/qj.821, 2011.

Dupuy, E. et. al.: Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), ACP, 9, 287-343, 2009.

Edwards, D. P., Emmons, L. K., Hauglustaine, D. A., Chu, A., Gille, J. C. , Kaufman, Y. J. , Pétron, G., Yurganov, L. N., Giglio, L., Deeter, M. N., Yudin, V., Ziskin, D. C., Warner, J., Lamarque, J.-F., Francis, G. L., Ho, S. P., Mao, D., Chan, J., and Drummond, J. R.: Observations of Carbon Monoxide and Aerosol From the Terra Satellite: Northern Hemisphere Variability, *J. Geophys. Res*., **109**, D24202, doi:10.1029/2004JD0047272004, 2004.

Elbern, H., Schwinger, J. and Botchorishvili, R.: Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: System configuration, *J. Geophys. Res*., **115**, D06302, doi:10.1029/2009JD011953, 2010.

Elguindi N.; Clark H.; Ordonez C.; Thouret, V., Flemming, J., Stein, O., Huijnen, V., Moinat, P., Inness, A., Peuch, VH., Stohl, A., Turquety, S., Athier, G., Cammas, J.P., Schultz, M.: Current status of the ability of the GEMS/MACC models to reproduce the tropospheric CO vertical distribution as measured by MOZAIC, *Geoscientific Model Development*, 3, 501-518 DOI: 10.5194/gmd-3-501-2010, 2010.

Emmons, L. K., Deeter, M. N., Gille, J. C., et al.: Validation of Measurements of Pollution in the Troposphere (MOPITT) CO retrievals with aircraft in situ profiles, *J. Geophys. Res*., **109**, D03309, doi:10.1029/2003JD004101, 2004.

Emmons, L. K., Pfister, G. G., Edwards, D. P., Gille, J. C., Sachse, G., Blake, D., Wofsy, S., Gerbig, C., Matross, D., and N´ed´elec, P.: Measurements of Pollution in the Troposphere (MOPITT) validation exercises during summer 2004 field campaigns over North America, *J. Geophys. Res*., **112**, D12S02, doi:0.1029/2006JD007833, 2007.

Engelen R. J., Serrar, S., Chevallier, F.: Four-dimensional data assimilation of atmospheric CO2 using AIRS observations, *J. Geophys. Res*., **114**, D03303, doi:10.1029/2008JD010739, 2009.

Errera, Q., Daerden, F., Chabrillat, S., Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D.: 4D-Var Assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses, *Atmos. Chem. Phys*., **8**, 6169-6187, 2008.

Eskes, H.J., Van Velthoven, P.F.J. and Kelder, H.M.: Global ozone forecasting based on ERS-2 GOME observations. *Atmos. Chem. Phys.* **2,** 271-278, 2002.

Eskes, H.J. and Boersma. K. F.: Averaging kernels for DOAS total-column satellite retrievals. *Atmos. Chem. Phys*., **3**, 1285-1291, 2003.

H. Eskes, van Velthoven, P., Valks, p., and Kelder, hH.: Assimilation of GOME total ozone satellite observations in a three-dimensional tracer transport model, *Quart. J. Roy. Meteor. Soc*. **129**, 1663, 2003.

H. J. Eskes, van der A, R. J., Brinksma, E. J. , Veefkind, J. P. de Haan, J. F., and Valks, P. J. M.: Retrieval and validation of ozone columns derived from measurements of SCIAMACHY on Envisat, *Atmos. Chem. Phys. Discuss*, 5, 4429-4475, 2005.

Fisher, M. and Andersson, E.: Developments in 4D-Var and Kalman Filtering. ECMWF Technical Memorandum 347, 2001. http://www.ecmwf.int/publications/library/do/references/list/14

Fisher, M.: Generalized frames on the sphere with application to background error covariance modelling. Seminar on resent developments in numerical methods for atmospheric and ocean modelling, 6-10 September 2004. Proceedings, ECMWF, pp. 87-101. http://www.ecmwf.int/publications/library/do/references/list/17334

Fisher, M.: Wavelet Jb - A new way to model the statistics of background errors. *ECMWF Newsletter*, No.106, 23-28, 2006. http://www.ecmwf.int/publications/newsletters/archive_chronological_list.htm

Flemming, J., Inness, A., Jones, L., Eskes, H. J., Huijnen, V., Schultz, M. G., Stein, O., Cariolle, D., Kinnison, D., and Brasseur, G.: Forecasts and assimilation experiments of the Antarctic ozone hole 2008, *Atmos. Chem. Phys*., **11**, 1961-1977, doi:10.5194/acp-11-1961-2011, 2011.

Flemming, J., Dethof, A., Moinat, P. , Ordóñez, C., Peuch, V.-H. , Segers, A., Schultz, M., Stein, O., van Weele, M.: Coupling Global Atmospheric Chemistry Transport Models to ECMWF Integrated Forecasts System for Forecast and Data Assimilation within GEMS, in: *Integrated Systems of Meso-Meteorological and Chemical Transport Models*, Baklanov, Alexander; Mahura, Alexander; Sokhi, Ranjeet (Eds.), Spinger, 2011a.

J. Flemming, Inness, A., Flentje, H., Huijen, V., Moinat, P., Schultz, M.G. and Stein O.: Coupling global chemistry transport models to ECMWF's integrated forecast system. *Geosci. Model Dev*., **2**, 253-265. www.geosci-model-dev.net/2/253/2009/ , 2009.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Szopa, S., Deeter, M. N., and Clerbaux, C.: Ten years of CO emissions as seen from Measurements of Pollution in the Troposphere (MOPITT), *J. Geophys. Res*., **116**, D05304, doi:10.1029/2010JD014416, 2011.

Geer, A. J., Peubey, C., Bannister, R. N., Brugge, R., Jackson, D. R., Lahoz, W. A., Migliorini, S., O'Neill, A. and Swinbank, R.: Assimilation of stratospheric ozone from MIPAS into a global general-circulation model: The September 2002 vortex split. *Quart. J. R. Meteorol. Soc.*, **132**: 231–257. doi: 10.1256/qj.04.181, 2006.

George, M., Clerbaux, C., Hurtmans, D., Turquety, S., Coheur, P.-F., Pommier, M. , Hadji-Lazaro, J., Edwards, D. P., Worden, H., Luo, M., Rinsland, C., and McMillan, W.: Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors. *Atmos. Chem. Phys*., **9**, 8317–8330, 2009.

Gibson, J. K., Kållberg, P., Uppala, S., Nomura, A., Hernandez, A., and Serrano, E.: ERA Description. *ERA-15 Project Report Series* 1, ECMWF. Available from www.ecmwf.int/publications, 1997.

Granier, C., Bessagnet, B., Bond, T., D'Angiola, A., Denier van der Gon, H., Frost, G. J., Heil, A., Kaiser, J. W., Kinne, S., Klimont, Z., Kloster,S., Lamarque, J.-F., Liousse, C., Masui, T., Meleux, F., Mieville, A., Ohara, R., Raut, J.-C., Riahi, K., Schultz, M. G., Smith, S. G., Thompson, A., van Aardenne, J., van der Werf, G. R., and van Vuuren, D. P.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980-2010 period. *Climatic Change*, **109**, 163-190. DOI: 10.1007/s 10584-011-0154-1, 2011a.

Granier et al.: Updated inventory of Global and European regional emissions based on the RETRO system: Evaluation of the anthropogenic emissions used in the MACC project. *MACC deliverable available from MACC website*, 2011b.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C. , and Eder, B.: Fully coupled ‘‘online’’ chemistry within the WRF model, *Atmos. Environm.*, **39**, 6957– 6975, 2005.

Griffin, R.J., Chen, J., Carmody, K. and Vutukuru, S.: Contribution of gas phase oxidation of volatile organic compounds to atmospheric carbon monoxide levels in two areas of the united States. *J. Geophys. Res.*, **11**, D10S17, doi:10.1029/2006JD007602, 2007.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P.I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), Atmos. Chem. Phys., 6, 3181-3210, 2006.

Hansen, J., Sato, M., and Ruedy, R.: Radiative forcing and climate response, *J. Geophys. Res*., **102**, 6831–6864, 1997.

Hollingsworth, A., Engelen, R. J., Textor, C., Benedetti, A., Boucher, O., Chevallier, F., Dethof, A., Elbern, H., Eskes, H., Flemming, J., Granier, C., Kaiser, J. W., Morcrette, J.-J., Rayner, R., Peuch, V.-H., Rouil, L., Schultz, M. G., Simmons, A. J. and The GEMS Consortium: Toward a Monitoring and Forecasting System For Atmospheric Composition: The GEMS Project. *Bull. Amer. Meteor. Soc*., **89**, 1147–1164, doi:10.1175/2008BAMS2355.1, 2008.

Holloway, J. S., Jakoubek, R. O. , Parrish, D. D., Gerbig, C., Volz-Thomas, A., Schmitgen, S., Fried, A., Wert, B. ,Henry, B., and Drummond, J. R.: Airborne intercomparison of vacuum ultraviolet fluorescence and tunable diode laser absorption measurements of tropospheric carbon monoxide, *J. Geophys. Res*., **105**(D19), 24,251–24,261, doi:10.1029/2000JD900237, 2000.

Hólm, E. V., Untch, A., Simmons, A., Saunders, R., Bouttier, F. and Andersson, E.: Multivariate ozone assimilation in four-dimensional data assimilation. Pp. 89-94 in *Proceedings of the Soda Workshop on Chemical Data Assimilation*, 9-10 December 1998, KNMI, De Bilt, The Netherlands, 1999.

Hortal, M. and Simmons, A.J.: Use of reduced Gaussian grids in spectral models. *Mon. Wea. Rev*., **119**, 1057-1074, 1991.

Hudman, R.C., Murray, L.T., Jacob, D.J., Millet, D.B., Turquety, S., Wu, S., Blake, D.R., Goldstein, A.H., Holloway, J., Sachse, G.W.: Biogenic versus anthropogenic sources of CO over the United States. *Geophysical Research Letters* **35**, L04801, 2008.

Huijnen, V., Flemming, J., Kaiser, J. W., Inness, A., Leitão, J., Heil, A., Eskes, H. J., Schultz, M. G., Benedetti, A., Hadji-Lazaro, J., Dufour, G., and Eremenko, M.: Hindcast experiments of tropospheric composition during the summer 2010 fires over western Russia, *Atmos. Chem. Phys*., **12**, 4341-4364, doi:10.5194/acp-12-4341-2012, 2012.

Inness, A., Flemming, J., Suttie, M. and Jones, L.: GEMS data assimilation system for chemically reactive gases. *ECMWF Tech. Memorandum* No.587. Available from http://www.ecmwf.int. , 2009.

Isaksen, I.S.A., Granier, C., Myhre, G., Berntsen, T. K., Dalsøren, S. B., Gauss, M., Klimont, Z., Benestad, R., Bousquet, P., Collins, W., Cox, T., Eyring, V., Fowler, D., Fuzzi, S., Jöckel, P., Laj, P., Lohmann, U., Maione, M., Monks, P., Prevot, A. S. H., Raes, F., Richter, A., Rognerud, B., Schulz, M., Shindell, D., Stevenson, D. S., Storelvmo, T., Wang, W.-C. , van Weele, M., Wild, M., and Wuebbles, D.: Atmospheric composition change: Climate-chemistry interactions. *Atmos. Environ.*, **43**, 5138-5192, doi:10.1016/j.atmosenv.2009.08.003, 2009.

Jonquières, I., Marenco, A., Maalej, A. and Rohrer, F.: Study of ozone formation and transatlantic transport from biomass burning emissions over West Africa during the airborne Tropospheric Ozone Campaigns TROPOZ I and TROPOZ II, *J. Geophys. Res*., **103**(D15), 19,059–19,073, doi:10.1029/98JD00819, 1998.

Justice, C. O., Giglio, L., Korontzi, S., Owens, J., Morrisette, J. T., Roy, D., Descloitres, J., Alleaume, S., Petitcolin, F., and Kaufman, Y.: The modis fire products. *Remote Sensing of Environment*, **83**:244–262, 2002.

Kaiser, J. W., Heil, A., Schultz, M. G., Stein, O., van der Werf, G. R., Wooster, M. J., and Xu, W.: Final report on implementation and quality of the D-FIRE assimilation system. *MACC deliverable D D-FIRE 7*, ECMWF, 2011.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, **9**:527–554, 2012.

Kalnay et al.: The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteor. Soc*., **77**, 437-470, 19, 1996.

Kanakidou, M., and Crutzen, P. J.: The photochemical source of carbon monoxide: Importance, uncertainties, and feedbacks, *Chemosphere Global Change Sci.*, **1**, 91–109, 1999.

Kar J., Bremer, H., Drummond, J. R., Rochon, Y. J., Jones, D. B. A., Nichitiu, F., Zou, J., Liu, J., Gille, J. C., Edwards, D. P., Deeter, M. N., Francis, G., Ziskin, D., and Warner, J.: Evidence of vertical transport of carbon monoxide from Measurements of Pollution in the Troposphere (MOPITT), *Geophys. Res. Lett.*, **31**(23), L23105, doi:10.1029/2004GL021128, 2004.

Khattatov, B.V., Lamarque, J.-F., Lyjak, L. V., Menard, R., Levelt, P., Tie, X., Brasseur, G. and Gille, J. C.: Assimilation of sattelite observations of long-lived chemical species in global chemistry transport models. *J. Geophys. Res.*, **105**, D23, 29135-29144, 2000.

Kinnison, D. E., Brasseur, G. P., Walters, S., Gracia, R. R., Marsh, D. R., Sassi, F., Harvey, V. L., Randall, C.E., Emmons, L., Lamarque, J. F., Hess, P., Orlando, J. J., Tie, X. X., Randel, W., Pan, L. L., Gettelman, A., Granier, C., Diehl, T., Niemeier, U. and Simmons, A. J.: Sensitivity of chemical tracers to meteorological parameters in the MOZART-3 chemical transport model. *J. Geophys. Res*, **112**, D20302, doi:10.1029/2006JD007879, 2007.

Kleinman, L., Y.-N. Lee, S. R. Springston, J. H. Lee, L. Nunnermacker, J. Weinstein-Lloyd, X. Zhou, and L. Newman: Peroxy radical concentration and ozone formation rate at a rural site in the southeastern United States, *J. Geophys. Res.*, **100**(D4), 7263-7274, 1995.

Koelemeijer, R., Stammes, P., Hovenier, J.W. and De Haan, J.F.: A fast method for retrieval of cloud parameters using oxygen a band measurements from the Global Ozone Monitoring Experiment. *J. Geophys. Res.* **106**(D4): doi: 10.1029/2000JD900657. issn: 0148-0227, 2001.

Koelemeijer, R., Stammes, P., Hovenier, J. W. and De Haan, J. F.: Global distributions of effective cloud fraction and cloud top pressure derived from oxygen A band spectra measured by the Global Ozone Monitoring Experiment: Comparison to ISCCP data. *J. Geophys. Res.* **107**(D12): doi: 10.1029/2001JD000840. issn: 0148-0227, 2002.

Kopacz, M., Jacob, D. J., Fisher, J. A., Logan, J. A., Zhang, L., Megretskaia, I. A., Yantosca, R. M., Singh, K., Henze, D. K., Burrows, J. P., Buchwitz, M., Khlystova, I., McMillan, W. W., Gille, J. C., Edwards, D. P., Eldering, A., Thouret, V., and Nedelec, P.: Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES), *Atmos. Chem. Phys*., **10**, 855-876, doi:10.5194/acp-10-855-2010, 2010.

Krol, M.C.,and van Weele, M.: Implications of variations in photodissociation rates for global tropospheric chemistry, *Atmospheric Environment*, **31**, 1257-1273, 1997.

Lefever, K., Chabrillat, S., Errera, Q., Viscardy, S., Christophe, Y., Wagner, A., Flentje, H., Thomas, W., Flemming, J., Inness, A., and O. Stein: Case study validation report for stratospheric ozone, subtasks 2.5 and 2.8, *MACC validation report document*. Available from http://www.gmes-atmosphere.eu/documents/ deliverables/g-rg, 2011.

Lahoz, W. A., Geer, A.J. , Bekki, S., Bormann, N., Ceccherini, S., Elbern, H., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Khattatov, B., Massart, S., Peuch, V. H., Rharmili, S., Ridolfi, M., Segers, A., Talagrand, O., Thornton, H. E., Vik A. F., and Von Clarman, T.: The Assimilation of Envisat data (ASSET) project, *Atmos. Chem. Phys*., **7**, 1773-1796, 2007.

Lamarque, J.-F., Bond, T. C. , Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D. , Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850-2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmos. Chem. Phys*., **10**, doi:10.5194/acp-10-7017-2010, 7017-7039, 2010.

Landgraf, J. and 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. H. J., Dobber, M. R., M¨alkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., Saari, H.: The Ozone Monitoring Instrument, *IEEE T. Geosci. Remote Sensing*, **44**(5), 1093–1101, 2006

Manney, G. L. et al.: Unprecedented Arctic ozone loss in 2011. *Nature*, doi: 10.1038/nature10556, 2011.

Marenco, A., V. Thouret, P. Nédélec, H.G. Smit, M. Helten, D. Kley, F. Karcher, P. Simon, K. Law, J. Pyle, G. Poschmann, R. Von Wrede, C. Hume, and T. Cook, Measurement of ozone and water vapour by Airbus in-service aircraft : The MOZAIC airborne programme, an overview, *J. Geophys. Res.*, **103**, D19, 25,631-25,642, 1998.

Mari, C. H., Cailley, G., Corre, L., Saunois, M., Attie, J. L., Thouret, V., and Stohl, A.: Tracing biomass burning plumes from the Southern Hemisphere during the AMMA 2006 wet season experiment, *Atmos. Chem. Phys*., **8**, 3951–3961, 2008.

McNally, A.P., Watts, P. D., Smith, J. A., Engelen, R. J., Kelly, G. A., Thépaut, J.-N., and Matricardi, M.: The assimilation of AIRS radiance data at ECMWF. *Quart. J. Roy. Meteor. Soc*., **132**, 935-957, 2006.

Millet, D. B., Jacob, D. J. , Boersma, K. F. , Fu, T.-M. , Kurosu, T. P. , Chance, K. , Heald, C. L., and Guenther, A.: Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, *J. Geophys. Res*., **113**, D02307, doi:10.1029/2007JD008950, 2008.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment. *Nature,* **463**(7282):747-56, 2010.

Morcrette, J.J. et al.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling, *J. Geophys. Res*., **114**, D06206, doi:10.1029/2008JD011235, 2009.

Nedelec, P., Cammas, J.-P., Thouret, V., Athier, G., Cousin, J.-M., Legrand, C., Abonnel, C., Lecoeur, F., Cayez, G., and Marizy, C.: An improved infrared carbon monoxide analyser for routine measurements aboard commercial aircraft: technical validation and first scientific results for the MOZAIC III programme, *Atmos. Chem. Phys.*, **3**, 1551-1564, 2003.

Newman, P. A., Oman, L. D., Douglass, A. R., Fleming, E. L., Frith, S. M., Hurwitz, M. M., Kawa, S. R., Jackman, C. H., Krotkov, N. A., Nash, E. R., Nielsen, J. E., Pawson, S., Stolarski, R. S., and Velders, G. J. M.: What would have happened to the ozone layer if chlorofluorocarbons (CFCs) had not been regulated?, *Atmos. Chem. Phys*., **9**, 2113-2128, doi:10.5194/acp-9-2113-2009, 2009.

Novelli, P. C., Masarie, K. A. , Lang, P. M. , Hall, B. D., Myers, R. C. , and Elkins, J. W.: Reanalysis of tropospheric CO trends: Effects of the 1997–1998 wildfires, *J. Geophys. Res*., **108**, 4464, doi:10.1029/2002JD003031, 2003.

Novelli, P.C. and Masarie, K.A. : Atmospheric Carbon Monoxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1988-2009, Version: 2011-10-14, Path: ftp://ftp.cmdl.noaa.gov/ccg/co/flask/event/ , 2010.

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S.,Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., and Taira, R.: The JRA-25 Reanalysis. *J. Meteor. Soc. Japan*, **85**, 369\u2013432, 2007.

Palmer, P. I., Jacob, D. J. , Fiore, A. M. , Martin, R. V. , Chance, K., and Kurosu, T. P.: Mapping isoprene emissions over North America using formaldehyde column observations from space, *J. Geophys. Res*., **108**(D6), 4180, doi:10.1029/2002JD002153, 2003.

Palmer, P. I., Abbot, D. S., Fu, T. M., Jacob, D. J., Chance, K., Kuruso, T. P., Guenther, A., et al.: Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of formaldehyde column. *J. Geophys. Res*., **111**: D12315, doi: 10.1029/2005JD006689, 2006.

Parrish, D.F. and Derber, J. C.: The National Meteorological Center's spectral statistical-interpolation analysis scheme. *Mon. Weather Rev*., **120**, 1747-1763, 1992.

Pickering, K. E., et al.: Convective transport of biomass burning emissions over Brazil during TRACE A, *J. Geophys. Res*., **101**, 23,993– 24,012, 1996.

E. Real, Orlandi, E., Law, K. S., Fierli, F., Josset, D., Cairo, F., Schlager, H., Borrmann, S., Kunkel, D., Volk, C. M. , McQuaid, J. B., Stewart, D. J. , Lee, J., Lewis, A. C., Hopkins, J. R., Ravegnani, F., Ulanovski, A., and Liousse, C.: Cross-hemispheric transport of central African biomass burning pollutants: Implications for downwind ozone production. *Atmos. Chem. Phys*., **10**, 3027-3046, 2010.

Riahi K, Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., Rafaj., P: RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change*. **109**:33-57. DOI 10.1007/s10584-011-0149-y, 2011.

Richter, A. and Burrows, J. P.: Retrieval of Tropospheric NO2 from GOME Measurements, *Adv. Space Res*., **29**(11) ,1673-1683, 2002.

Richter, A., Burrows, J. P., Nuß, H., Granier, C., and Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, *Nature*, **437**, 129–132, doi:10.1038/nature04092, 2005.

Rodgers, C. D.: *Inverse methods for atmospheric sounding – theory and practice*, Series on Atmospheric, Oceanic and Planetary Physics, World Scientific Publishing, Singapore, 2000.

Sandu, A. and Chai, T.: Chemical data assimilation – An overview. *Atmosphere*, **2**, 426-463; doi:10.3390/atmos2030426, 2011.

Schubert, S. D., Pfaendtner, J., and Rood, R.: An assimilated data set for Earth Science applications. *Bull. Am. Meteor. Soc*., **74**, 2331â2342, 1993.

Seinfeld, J. H., and Pandis, S. N.: *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, John Wiley, Hoboken, N. J., 2006.

Semane, N., Peuch, V.-H. , Pradier, S., Desroziers, G., El Amraoui, L., Brousseau, P., Massart, S., Chapnik, B., and Peuch, A.: On the extraction of wind information from the assimilation of ozone profiles in Météo–France 4-D-Var operational NWP suite. *Atmos. Chem. Phys*., **9**, 4855-4867, 2009.

Siddans, R., Reburn, W. J., Kerridge, B. J., and Munro, R.: Height resolved ozone information in the troposphere and lower stratosphere from GOME. *Technical report, British Atmospheric Data Centre* (BADC). Available from http://badc.nerc.ac.uk/home/index.html , 2002.

Shindell, D.T., Faluvegi, G., Stevenson, D. S., Krol, M. C., Emmons, L. K., Lamarque, J.-F., Pétron, G., Dentener, F. J., Ellingsen, K., Schultz, M. G., Wild, O., Amann, M., Atherton, C. S., Bergmann, D. J., Bey, I., Butler, R., Cofala, J., Collins, W. J., Derwent, R. G., Doherty, R. M., Drevet, J., Eskes, H. J., Fiore, A. M., Gauss, M., Hauglustaine, D. A., Horowitz, L. W., Isaksen, I. S. A., Lawrence, M. G., Montanaro,V., Müller, J.-F., Pitari, G., Prather, M. J., Pyle, J. A., Rast, S., Rodriguez, J. M., Sanderson, M. G., Savage, N. H., Strahan, S. E., Sudo, K., Szopa, S., Unger, N., van Noije, T. P. C., and Zeng, G.: Multi-model simulations of carbon monoxide: Comparison with observations and projected near-future changes. *J. Geophys. Res*., **111**, D19306, doi:10.1029/2006JD007100, 2006.

Stavrakou, T. , Müller, J.-F., De Smedt, I. , Van Roozendael1, M., van der Werf, G. R., Giglio, L., and A. Guenther: Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns. *Atmos. Chem. Phys*., **9**, 1037–1060, 2009.

Stein, O.: Model documentation of the MOZART CTM as implemented in the GEMS system. Available from: http://gems.ecmwf.int/do/get/PublicDocuments/1531/1172 , 2009.

Stein, O., Flemming, J., Inness, A., Kaiser, J. W., Schultz, M. G.: Global reactive gases and reanalysis in the MACC project *J. Integrative Environm. Science*, 2012, s, DOI:10.1080/1943815X.2012.696545, 2012.

Thompson, A.M.: The oxidizing capacity of the earth's atmosphere: Probable past and future changes. *Science* **256**:1157-1165, 1992.

Thompson, A. M., Pickering, K., McNamara, D., Schoeberl, M., Hudson, R., Kim, J., Browell, E., Kirchhoff, V., and Nganga, D..: Where did tropospheric ozone over southern Africa and the tropical Atlantic come from in October 1992? Insights from TOMS, GTE TRACE A, and SAFARI 1992, *J. Geophys. Res.*, **101**, D19, doi:10.1029/96JD01463, 1996.

Torres, O., Chen, Z., Jethva, H., Ahn, C., Freitas, S.R. and Bhartia, P.K.: OMI and MODIS observations of the anomalous 2008-2009 Southern Hemisphere biomass burning. *Atmos. Chem. Phys*., **10**, 3505-3513, 2010.

Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, 2005.

Valcke, S. and Redler, R.: OASIS4 User Guide (OASIS4_0_2). PRISM–Support Initiative, Technical Report No 4. Available from http://www.prism.enes.org/Publications/index.php , 2006.

Van der A, R.J., Allaart, M.A.F., and Eskes, H.J. Multi sensor reanalysis of total ozone. *Atmos. Chem.Phys*. **10**, 11277–11294, 2010.

van der A, R. J., Eskes, H. J., Boersma, K. F., van Noije, T. P. C., Van Roozendael, M., De Smedt, I., Peters, D. H. M. U., Meijer, E. W.: Trends, seasonal variability and dominant NOx source derived from a ten year record of NO2 measured from space. *J. Geophys. Res*., **113**, D0430210.1029/2007JD009021, 2008.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasib- hatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmos. Chem. Phys*., **10**:11707–11735, 2010.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., and Kasibhatla, P. S.: Interannual variability in global biomass burning emissions from 1997 to 2004. *Atmos. Chem. Phys*., **6**(11):3423–3441, 2006.

Van Vuuren, D.P., Edmonds, J., Thomson, A., Riahi, K., Kainuma, M., Matsui, T., Hurtt, G.C., Lamarque, J.-F., Meinshausen, M., Smith, S., Granier, C., Rose, S.K., Hibbard, K.A., Representative Concentration Pathways: An overview, *Climatic Change*, 2010.

P.Wang, Stammes, P., van der A, R., Pinardi, G., and van Roozendael, M.: FRESCO+: an improved O2 A-band cloud retrieval algorithm for tropospheric trace gas retrievals. *Atmos. Chem. Phys*., **8**, 6565–6576, 2008.

Waters, J. W., et al.: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite, *IEEE Trans. Geosci. Remote Sens*., **44**(5), 1075– 1092, 2006.

Weber, M., Dikty, S., Burrows, J.P., Garny, H., Dameris, M., Kubin, A., Abalichin, J., and Langematz, U.: The Brewer-Dobson circulation and total ozone from seasonal to decadal time scales. *Atmos. Chem. Phys*., **11**, doi:10.5194/acp-11-11211-2011, 11221-11235, 2011.

Wittrock, F.: *The retrieval of oxygenated volatile organic compounds by remote sensing techniques*, Ph.D. thesis, Univ. of Bremen, Germany, 192 pp., 2006.

Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., Volkamer, R., Beirle, S., Platt, U., and Wagner, T.: Simultaneous global observations of glyoxal and formaldehyde from space, *Geophys. Res. Lett.*, **33**, L16804, doi:10.1029/2006GL026310, 2006.

Zhang, L., Li, Q. B., Jin, J., Liu, H., Livesey, N., Jiang, J. H., Mao,Y., Chen, D., Luo, M., and Chen, Y.: Impacts of 2006 Indonesian fires and dynamics on tropical upper tropospheric carbon monoxide and ozone *Atmos. Chem. Phys.*, **11**, 10929-10946, 2011.