

## Comparing optimized CO emission estimates using MOPITT or NOAA surface network observations

P. B. Hooghiemstra,<sup>1,2</sup> M. C. Krol,<sup>1,2,3</sup> P. Bergamaschi,<sup>4</sup> A. T. J. de Laat,<sup>1,5</sup>  
G. R. van der Werf,<sup>6</sup> P. C. Novelli,<sup>7</sup> M. N. Deeter,<sup>8</sup> I. Aben,<sup>1</sup> and T. Röckmann<sup>2</sup>

Received 20 October 2011; revised 14 January 2012; accepted 30 January 2012; published 27 March 2012.

[1] This paper compares two global inversions to estimate carbon monoxide (CO) emissions for 2004. Either surface flask observations from the National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA/ESRL) Global Monitoring Division (GMD) or CO total columns from the Measurement of Pollution in the Troposphere (MOPITT) instrument are assimilated in a 4D-Var framework. Inferred emission estimates from the two inversions are consistent over the Northern Hemisphere (NH). For example, both inversions increase anthropogenic CO emissions over Europe (from 46 to 94 Tg CO/yr) and Asia (from 222 to 420 Tg CO/yr). In the Southern Hemisphere (SH), three important findings are reported. First, due to their different vertical sensitivity, the stations-only inversion increases SH biomass burning emissions by 108 Tg CO/yr more than the MOPITT-only inversion. Conversely, the MOPITT-only inversion results in SH natural emissions (mainly CO from oxidation of NMVOCs) that are 185 Tg CO/yr higher compared to the stations-only inversion. Second, MOPITT-only derived biomass burning emissions are reduced with respect to the prior which is in contrast to previous (inverse) modeling studies. Finally, MOPITT derived total emissions are significantly higher for South America and Africa compared to the stations-only inversion. This is likely due to a positive bias in the MOPITT V4 product. This bias is also apparent from validation with surface stations and ground-truth FTIR columns. Our results show that a combined inversion is promising in the NH. However, implementation of a satellite bias correction scheme is essential to combine both observational data sets in the SH.

**Citation:** Hooghiemstra, P. B., M. C. Krol, P. Bergamaschi, A. T. J. de Laat, G. R. van der Werf, P. C. Novelli, M. N. Deeter, I. Aben, and T. Röckmann (2012), Comparing optimized CO emission estimates using MOPITT or NOAA surface network observations, *J. Geophys. Res.*, 117, D06309, doi:10.1029/2011JD017043.

### 1. Introduction

[2] Carbon monoxide (CO) is emitted to the atmosphere by the process of incomplete combustion of fossil and bio-fuels and biomass burning. CO is also produced in the atmosphere by oxidation of methane and non-methane volatile organic compounds (NMVOCs). Through its main removal

process, reaction with the radical OH, CO perturbs the oxidation capacity of the atmosphere [Logan *et al.*, 1981] and in particular the methane lifetime. It is also a precursor of tropospheric ozone in high NO<sub>x</sub> conditions, thus contributing to photochemical smog.

[3] The magnitude of CO sources reported in literature shows a large range [e.g., Duncan *et al.*, 2007]. The large uncertainties are caused by several factors, for example increasing emissions from fossil and biofuel combustion in East-Asia for the Northern Hemisphere (NH) as well as interannual variability of CO emissions in the Tropics and boreal NH due to biomass burning [van der Werf *et al.*, 2010]. Also, the amount of CO produced by the oxidation of NMVOCs (mainly isoprene and monoterpenes) is uncertain.

[4] Inverse modeling can be used to optimize CO source estimates. So far, inversion studies used either flask observations from surface stations [e.g., Bergamaschi *et al.*, 2000; Kasibhatla *et al.*, 2002; Pétron *et al.*, 2002; Pison *et al.*, 2009; Hooghiemstra *et al.*, 2011], aircraft observations [Palmer *et al.*, 2003; Heald *et al.*, 2004] or satellite observations [e.g., Arellano *et al.*, 2004; Pétron *et al.*, 2004; Arellano *et al.*, 2006; Stavroukou and Müller, 2006; Chevallier *et al.*, 2009;

<sup>1</sup>SRON Netherlands Institute for Space Research, Utrecht, Netherlands.

<sup>2</sup>Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, Netherlands.

<sup>3</sup>Department of Meteorology and Air Quality, Wageningen University, Wageningen, Netherlands.

<sup>4</sup>Institute for Environment and Sustainability, Joint Research Centre, European Commission, Ispra, Italy.

<sup>5</sup>Royal Netherlands Meteorological Institute, de Bilt, Netherlands.

<sup>6</sup>Faculty of Earth and Life Sciences, Free University, Amsterdam, Netherlands.

<sup>7</sup>Climate Monitoring and Diagnostics Laboratory, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA.

<sup>8</sup>Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder, Colorado, USA.

*Jones et al.*, 2009; *Kopacz et al.*, 2009, 2010; *Fortems-Cheiney et al.*, 2011]. Due to the different spatiotemporal resolution, analytical precision and altitude sensitivity, these different types of measurements potentially yield complementary information about CO in the atmosphere.

[5] Although *Müller and Stavrakou* [2005] used both surface data and observations from Fourier-Transform Infrared Spectrometers (FTIR), to our knowledge, at present no study assimilated both surface and satellite observations jointly in a four-dimensional variational (4D-Var) data assimilation system for CO. For methane, *Bergamaschi et al.* [2007, 2009] and *Meirink et al.* [2008a] performed inversions using both flask measurements from the National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA/ESRL) Global Monitoring Division (GMD) and total columns from the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) instrument. They found that a bias correction scheme was necessary to obtain good agreement of model simulations with both data sets. Hence, before we can actually perform inversions combining surface and satellite observations for CO, it is important to analyze the consistency and possible differences between inversions using either data set.

[6] Therefore, we present in this study two inversions for the year 2004: The first inversion assimilates flask observations from the NOAA surface network. The second inversion uses CO total columns from the Measurement of Pollution in the Troposphere (MOPITT) instrument, version 4 (V4) to constrain the emissions. Since the MOPITT V4 product [*Deeter et al.*, 2010] is modeled using lognormal probability distributions, we will describe in detail how we assimilated these observations in our 4D-Var system. According to *Deeter et al.* [2010], the MOPITT V4 product is improved on retrieval performance, i.e., more retrievals converge leading to more observations, in particular in very clean and highly polluted regions. This was confirmed by *Fortems-Cheiney et al.* [2011], who compared a prior model simulation to MOPITT V3 and MOPITT V4 columns and reported a mean model data bias reduction of 80% for V4 compared to V3. However, the long-term bias drift present in V3 has not been solved and is still present in the V4 product [*Deeter et al.*, 2010]. *Fortems-Cheiney et al.* [2011] reported the first CO inversion results in a variational data assimilation using MOPITT V4 data and the posterior simulation was shown to improve the agreement with independent NOAA surface flasks in the NH and the Tropics. However, the agreement deteriorated significantly in the SH, where posterior modeled CO mixing ratios are much higher compared to the NOAA stations in the remote SH. In the current framework we will explicitly test the consistency in optimized emissions using either NOAA surface flask observations or MOPITT total columns. This is another step in evaluating the effect of the assimilated observations on the inferred emission estimates. To test the validity of our inferred emission estimates, we will validate our results with independent (non-assimilated) observations.

[7] With respect to our previous study [*Hooghiemstra et al.*, 2011], we have optimized CO from NMVOC oxidation on the resolution of the underlying model. This approach was first used by *Stavrakou and Müller* [2006] and more recently in the studies of *Pison et al.* [2009] and *Fortems-Cheiney et al.* [2011]. *Jiang et al.* [2011] recently showed

that aggregating the NMVOC-CO source to a global source has large effects on the inferred emission estimates. Therefore, it is interesting to compare our results to previous (inverse) model studies, because some recent inversion studies [*Jones et al.*, 2009; *Kopacz et al.*, 2010] did not optimize the NMVOC-CO source explicitly on the resolution of the model, bearing the risk that deficiencies in the NMVOC-CO priors used might be projected on either the biomass burning or anthropogenic emissions. For example, *Kopacz et al.* [2010] and *Liu et al.* [2010] found that (apart from underestimated fossil and biofuel combustion emissions in East Asia) in particular biomass burning CO emissions seem to be underestimated in the Global Fire Emissions Database (GFED) v2.

[8] This paper is organized as follows: The 4D-Var system is described in section 2, where we introduce the chemistry transport model TM5 and the prior information used. Furthermore, we describe the observational data sets that are assimilated and those used for validation. In addition, technical details concerning the convergence of the method and data rejection criteria are given. In section 3 the inferred emissions of the two inversions are discussed and compared in detail. Furthermore, our results are compared to recent literature studies and validated with independent data. A series of sensitivity studies is presented in section 4. Conclusions are presented in section 5.

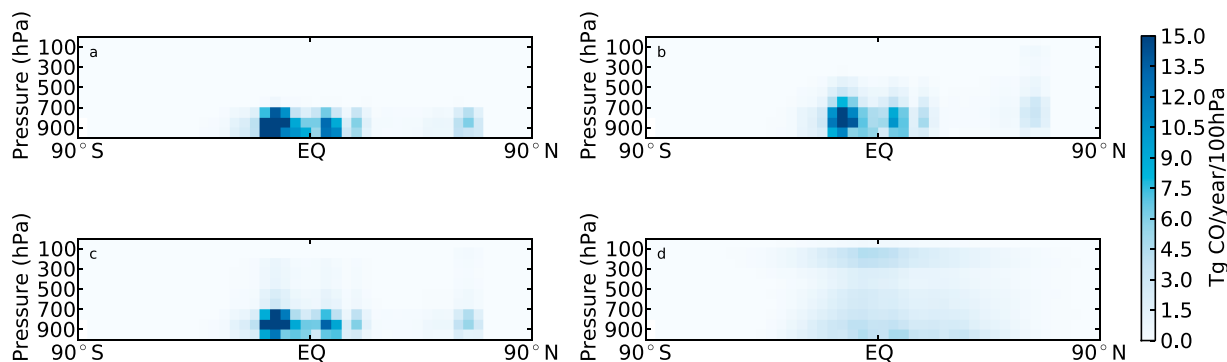
## 2. Method

[9] The 4D-Var system used in this study is based on the system described by *Hooghiemstra et al.* [2011]. In short, 4D-Var inverse modeling optimizes a state vector  $\mathbf{x}$  (containing e.g. emissions) such that modeled CO mixing ratios  $H(\mathbf{x})$  are close to a set of observations  $\mathbf{y}$  weighted with the observational error covariance matrix  $\mathbf{R}$ , while staying close to the prior state  $\mathbf{x}_b$  weighted with the prior error covariance matrix  $\mathbf{B}$ . Mathematically this means that

$$\hat{\mathbf{x}} = \min \arg \mathcal{J},$$

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^\top \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} \sum_{i=1}^M (\mathbf{y}_i - H(\mathbf{x})_i)^\top \mathbf{R}_i^{-1}(\mathbf{y}_i - H(\mathbf{x})_i), \quad (1)$$

where  $\hat{\mathbf{x}}$  is the optimized state vector,  $i$  refers to the time step,  $M$  is the number of time steps with observations and  $T$  is the transpose operator. We use the iterative minimizer CONGRAD [*Fisher and Courtier*, 1995] which is based on the conjugate gradient method, [*Hestenes and Stiefel*, 1952] and the Lanczos algorithm [*Lanczos*, 1950]. After  $N$  iterations CONGRAD returns both  $\hat{\mathbf{x}}$  and the  $N$  leading eigenpairs  $(\lambda_j, \nu_j)$ ,  $j = 1, \dots, N$  of the Hessian of the cost function. These eigenpairs are used to construct an approximation of the posterior error covariance matrix corresponding to  $\hat{\mathbf{x}}$ . We applied a stricter stopping criterion for the iterative minimization method with respect to our previous work. We now require a gradient norm reduction (preduc) of  $10^6$  for convergence (100 in work by *Hooghiemstra et al.* [2011]), since calculations showed that for preduc values of 50, 200 and 1000, the annual posterior emission estimates may still vary regionally. The sensitivity of the inferred emissions



**Figure 1.** Zonally and yearly averaged (a, b, and c) biomass burning emissions and (d) NMVOC-CO emissions. The biomass burning vertical distributions correspond to the base simulation (all emissions released below 2 km, Figure 1a), the FVERT used in this study (Figure 1b) and the FVERT sensitivity simulation by *Hooghiemstra et al.* [2011] (Figure 1c). NMVOC-CO (Figure 1d) is emitted higher up in the troposphere.

with respect to the chosen predic value will be further discussed in section 4. Generally, a predic value of at least 1000 is required for convergence of the emissions in our system. As outlined by *Meirink et al.* [2008b], the system first optimizes large scale patterns to obtain a large cost function reduction. In later iterations fine scale patterns are optimized, which is accompanied by a convergence of the posterior errors. Therefore we use a predic factor of  $10^6$  for our base inversions to obtain the best estimate of the posterior error covariance matrix.

## 2.1. Chemistry Transport Model TM5

[10] The chemistry transport model TM5 is used to simulate CO mixing ratios. TM5 [*Krol et al.*, 2005] uses meteorological fields from the European Centre for Medium-Range Weather Forecasts (ECMWF). These fields drive model transport on a 3-hourly basis (6-hourly for 3-D fields). As in our previous study, we use the TM5-CO only version [*Hooghiemstra et al.*, 2011]. Hence, climatological OH [*Spivakovsky et al.*, 2000], scaled with a factor 0.92 (based on methylchloroform simulations for the years 2000–2006 [*Huijnen et al.*, 2010]), is used to keep the model linear. All simulations are performed on a  $6^\circ \times 4^\circ$  grid resolution with 25 vertical levels.

## 2.2. Prior Information and Error Structure ( $\mathbf{x}_b$ , $\mathbf{B}$ )

[11] The prior state vector  $\mathbf{x}_b$  consists of monthly mean CO emissions for three categories. Anthropogenic (combustion of fossil and biofuels) emissions are taken from the Emissions Database for Global Atmospheric Research (EDGARv4.1, compiled for the year 2004, European Commission and Netherlands Environmental Assessment Agency, <http://edgar.jrc.ec.europa.eu/overview.php?v=41>) and total to 462 Tg CO in 2004. Biomass burning (vegetation fires) emissions from the Global Fire Emissions Database (GFEDv3 [*van der Werf et al.*, 2010]) are used with a total of 334 Tg CO in 2004. The natural source consists of direct emissions from plants and the oceans amounting to 115 Tg CO/yr [*Houweling et al.*, 1998] and the contribution of NMVOC-CO. We add NMVOC-CO to the natural source since the bulk of this source consists of biogenic (isoprene and monoterpenes) emissions. The NMVOC-CO source is based on monthly 3-D

CO production fields (the same as used by *Hooghiemstra et al.* [2011], but following their posterior emission estimates, scaled to an annual total of 400 Tg CO/yr) from a full-chemistry run with the TM4 model [*Myriokefalitakis et al.*, 2008]. These fields are summed over the vertical coordinate and combined with the direct emissions from plants and the oceans to obtain a monthly 2-D emission field, while archiving the corresponding vertical distribution. The resulting 2-D field is added to the state vector and totals to 515 Tg CO/yr. In this way we effectively optimize a volume CO source. A similar approach has been adopted in the studies of *Pison et al.* [2009] and *Fortems-Cheiney et al.* [2011]. However, in their model a separate formaldehyde tracer was added whereas here we emit directly CO. Moreover, *Fortems-Cheiney et al.* [2011] optimized the full 3-D chemical production field of formaldehyde. Here we have chosen to assume the vertical distribution to be known a priori and only optimize 2-D emission fields to reduce the length of the state vector. In contrast to *Hooghiemstra et al.* [2011] we do not optimize CO production from methane oxidation. Instead we use optimized methane mixing ratio fields from a 4D-Var inversion for methane [*Bergamaschi et al.*, 2009; S. Houweling et al., manuscript in preparation, 2012] that are consistent with the NOAA surface network. The production of CO from methane (assuming a CO yield of 1.0) accounts for 865 Tg CO/yr. *Hooghiemstra et al.* [2011] used a constant tropospheric methane mixing ratio of 1800 ppb resulting in 885 Tg CO/yr from methane oxidation. The total prior emissions (including CO from methane oxidation) amount to 2176 Tg CO in 2004. Although the emissions presented here are optimized as 2-D fields, biomass burning and natural emissions are distributed vertically as shown in Figures 1a and 1d, respectively. As shown also by *Val Martin et al.* [2009], biomass burning CO is mostly released within the boundary layer, except when pyrogenic clouds are triggered. In contrast, CO production from the oxidation of methane and NMVOCs occurs at higher altitudes.

[12] The prior error structure used in our inversions is kept the same as in work by *Hooghiemstra et al.* [2011]: Grid-scale monthly errors of 250% of the corresponding grid-scale emissions are chosen for the natural source and the biomass

burning source. For the anthropogenic source the grid-scale error is set to 50% for the Western developed world (North America, Europe and Australia) and to 250% for the rest of the world. Note that these large prior errors implicitly allow for the possibility of negative emissions. This could be avoided by employing a ‘semi-exponential’ description of the emission distribution as is done recently for methane inversions [Bergamaschi *et al.*, 2009, 2010]. However, this would lead to a non-quadratic cost function for which the conjugate gradient method is not suited anymore. This will be discussed in more detail in section 2.3.2. We assign spatial and temporal error correlations to reduce the effective number of variables to be optimized by the inversion. A Gaussian spatial correlation length of 1000 km is used for all emission categories. This is in particular important for the stations-only inversion in which the number of observations is much smaller compared to the number of state vector elements. For a fair comparison, this correlation length was kept 1000 km in the MOPITT-only inversion. An e-folding temporal correlation length of 9.5 months (0.9 month-to-month correlation coefficient) is set for the anthropogenic and natural emissions. For biomass burning emissions an e-folding temporal correlation length of 0.62 months (0.2 month-to-month correlation coefficient) is used.

### 2.3. Observations Assimilated in 4D-Var ( $y$ , $R$ )

[13] In this section we describe the observations that are assimilated in the 4D-Var system. For both NOAA surface flask observations and MOPITT total columns, the observations used are described as well as the assigned uncertainty and the contribution of an observation to the cost function (equation (1)).

#### 2.3.1. NOAA Surface Flask Observations

[14] Currently the NOAA ESRL surface network consists of over 50 surface stations worldwide at which CO mixing ratios are measured weekly with very high analytical precision by using flask samples [Novelli *et al.*, 2003]. However, model simulations on a coarse grid are difficult to compare one-to-one with these flask observations, specifically due to model representativeness errors. For example, in the model the emissions are given per grid box and time step and are instantaneously mixed over the grid volume. In reality, the subgrid-scale variability of the emissions leads to a heterogeneous distribution of CO mixing ratios in that box. Hence, a station located downwind of an emission, would observe higher CO mixing ratios compared to the model. Furthermore, strong gradients in CO mixing ratios due to passing pollution plumes are much sharper in reality than the model can represent.

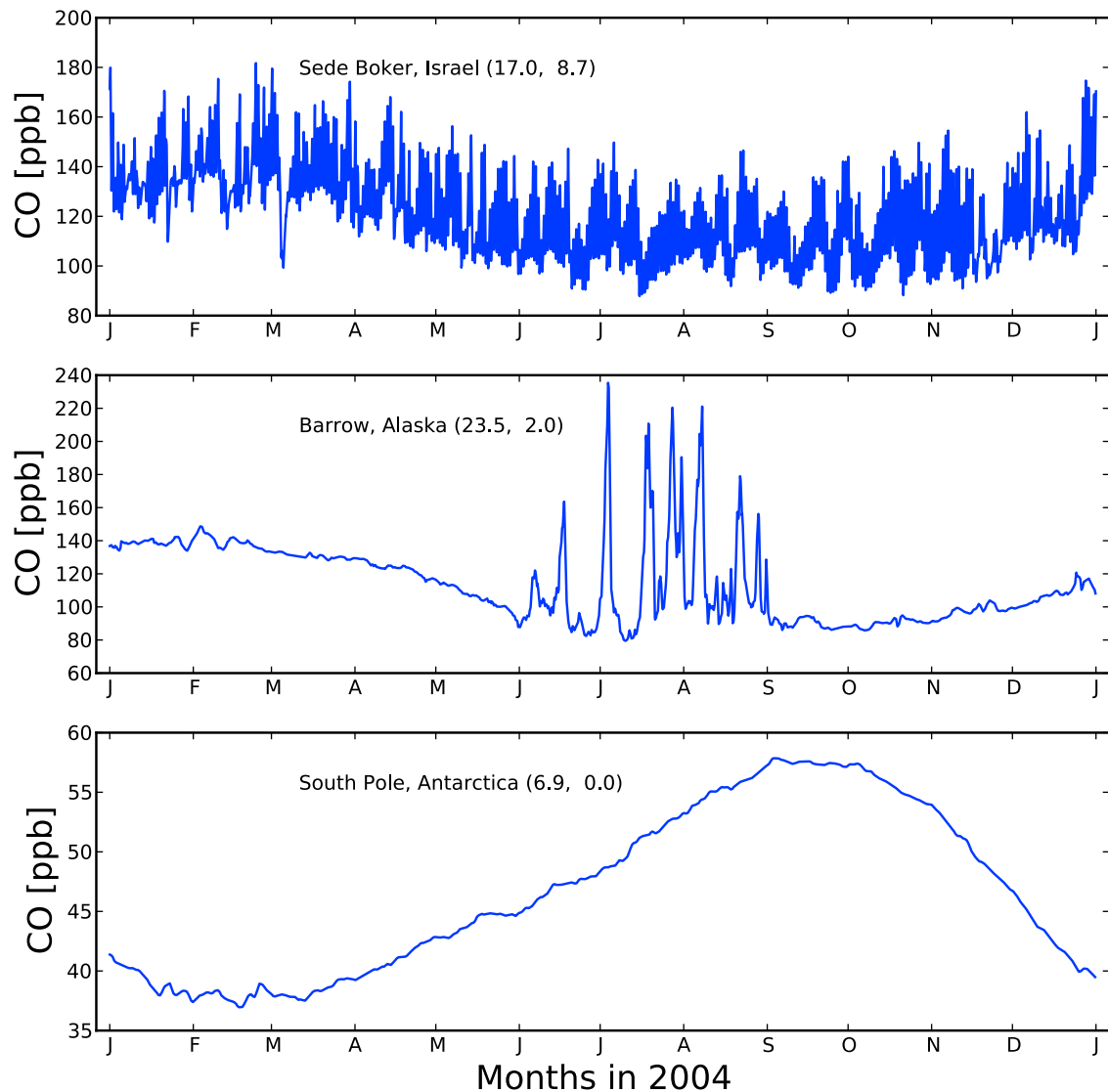
[15] For these reasons, inverse modeling studies deweight or reject some stations before assimilation to prevent biased results. For example, in CarbonTracker Europe (optimization of CO<sub>2</sub> fluxes using Kalman filtering [Peters *et al.*, 2010]) 2 stations are explicitly not assimilated and stations in strong emissions regions are assigned large fixed errors (of 7.5 ppm CO<sub>2</sub>). Bergamaschi *et al.* [2010] (4D-Var optimization of methane fluxes) give an advanced description of the model representativeness error in TM5. The total observational error  $\sigma_{obs}$  is the sum of a measurement error and the model representativeness error, consisting of errors due to local emissions, modeled 3-D gradients and variations in time. It was shown that the observational errors calculated in this

way vary largely from station to station and can vary in time for a certain station throughout the year.

[16] In this study we first apply a quantitative criterion to select the stations that we assimilate in the system and then apply the scheme to estimate the overall observational error of Bergamaschi *et al.* [2010]. The criterion is based on a model simulation with prior sources for the year 2004. The idea is that stations with a large diurnal cycle, most likely due to nearby sources in the model, are excluded whereas background stations and stations influenced by seasonal emissions from for example biomass burning, are maintained. With a model time step of 45 minutes, the model samples each station 32 times per day. From these modeled CO mixing ratio series we compute a daily standard deviation and use the mean daily standard deviation over the whole year as a measure of the diurnal variation. If this measure exceeds a certain threshold (set to 3.5 ppb in this study), the station is not assimilated in the 4D-Var system. As an illustration, the modeled simulation and the mean daily standard deviation for three stations are presented in Figure 2. For comparison also the standard deviation of the complete model time series (annual standard deviation) is given. For station Sede Boker (Figure 2, top), the mean daily standard deviation amounts to 8.7 ppb and the station is not assimilated. In contrast, although station Barrow, Alaska (Figure 2, middle) has an annual standard deviation of 23.5 ppb, the mean daily standard deviation is only 2 ppb. This station is maintained in the assimilation because the model is expected to reproduce the seasonal cycle more accurately than the diurnal cycle. For comparison, station South Pole, Antarctica shows only a very small spread throughout the year (6.9 ppb) and no daily spread as there are no sources of CO nearby. We acknowledge that since the criterion is based on a model simulation, the choice of stations to be assimilated depends strongly on the emissions used in this simulation. Here we used the prior emissions as described in section 2.2 and we believe that we assimilate mainly stations for which the coarse model can reproduce the observations. The location of the 34 stations maintained in the assimilation are shown in Figure 3.

[17] With respect to our previous study, the measurement error has been increased to 3 ppb, because Hooghiemstra *et al.* [2011] found that a measurement error of 1.5 ppb was too conservative in particular on the remote SH. This was likely due to an underestimate of the model error in this region as potential chemistry and transport errors were not included in this error. As a consequence, a large fraction of the observations was not assimilated in the second cycle (see below). With the enhanced observational errors, the total observational error ranges typically from 3–20 ppb. Close to emission regions, the error can even become as large as 100 ppb. In the clean remote SH, the observation error is dominated by the measurement component of 3 ppb. In contrast, in the polluted NH, where most surface CO is released, the model representativeness error is the dominant error term (e.g., see the black bars representing the total observational error in Figure 7).

[18] Each flask observation contributes to the observational part of the cost function. The costs for a mismatch are defined as  $\frac{1}{2} \left( \frac{y_m - y}{\sigma_{obs}} \right)^2$ , where  $y_m$  is the mean modeled CO mixing ratio during a 3 hour period,  $y$  is the observed CO



**Figure 2.** Modeled CO mixing ratio at three NOAA stations for 2004. Numbers between brackets represent the annual standard deviation (ppb) and the mean daily standard deviation (ppb), respectively. The daily standard deviation is computed as the standard deviation of the model CO mixing ratios sampled in each time step during one day. The mean daily standard deviation is the average of these daily standard deviations over the whole year. The annual standard deviation is computed as the standard deviation of the model CO mixing ratios sampled per time step during the year. Our station selection criterion rejects stations with a mean daily standard deviation  $> 3.5$  ppb.

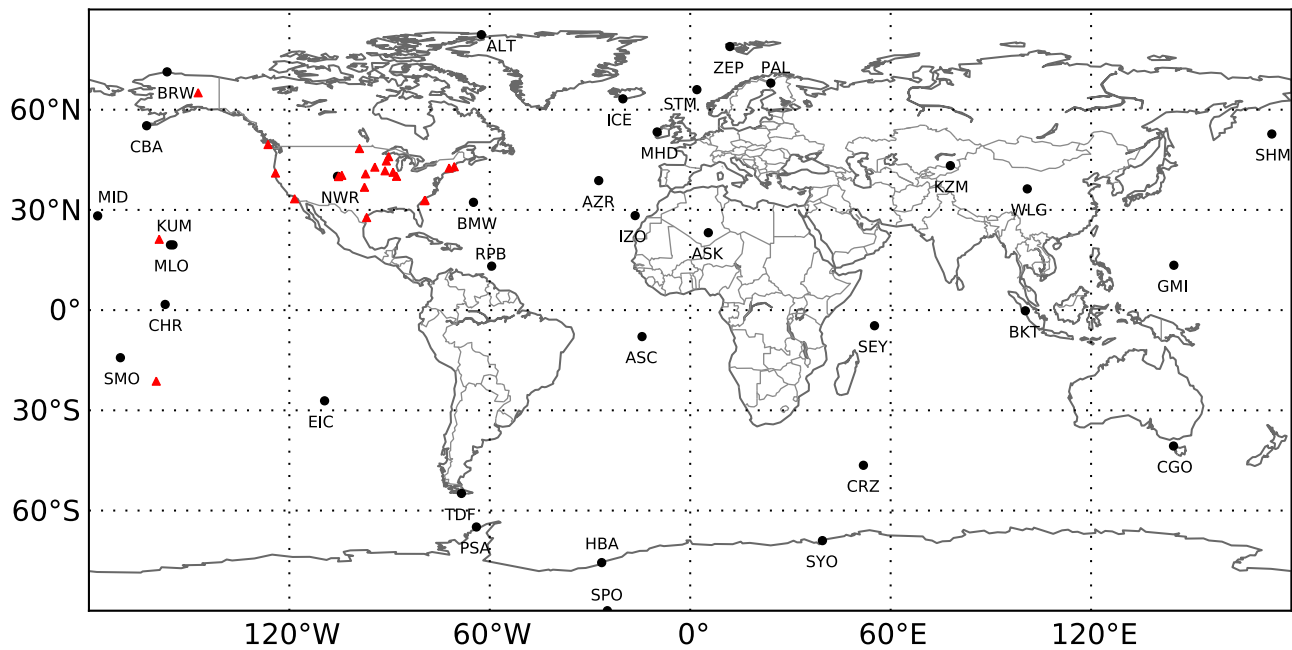
mixing ratio and  $\sigma_{obs}$  is the observation error. We assume the errors to be uncorrelated leading to a diagonal observational error covariance matrix  $\mathbf{R}$ .

[19] We perform the inversion using surface flask observations in 2 cycles following the approach of *Bergamaschi et al.* [2005]. After the first inversion cycle, all observations outside a  $3\sigma$  interval are not used in the second cycle to avoid single outliers to bias the emission estimates. In our previous work, the amount of rejected data points was 15–20% influencing the inferred emissions regionally from cycle 1 to cycle 2. With the larger observation errors used in this study, the number of rejected data points is reduced to around 8% based on a preduc factor of 100 as used by *Hooghiemstra et al.* [2011]. Moreover, for a preduc factor

of  $10^6$ , this number further reduces to 4% since the model fits the observations more accurately. As a result, the difference in inferred emissions from cycle 1 to cycle 2 becomes much smaller. We acknowledge that a rejection of 4% is still a large number given the Gaussian range of  $3\sigma$  that statistically should lead to a rejection of less than 1% of the data. However, this 4% is mainly caused by a few stations that are still difficult to fit, most likely due to transitions from polluted to very clean air masses that the coarse model can not resolve and is difficult to model as a representativeness error.

### 2.3.2. MOPITT V4 CO Total Columns

[20] The MOPITT instrument was launched in December 1999 on board NASA's Terra satellite. Although a cooler failure occurred at one side in May 2001, the instrument is

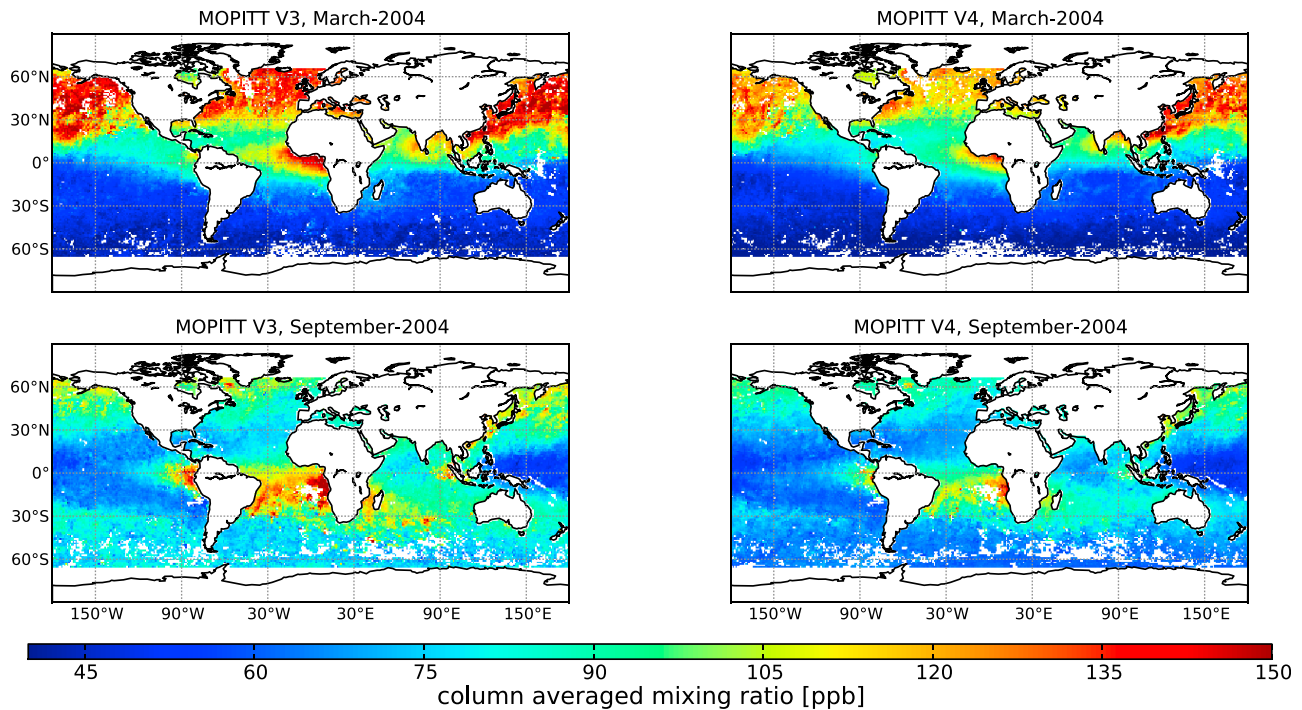


**Figure 3.** Positions of the 34 stations assimilated in the stations-only inversion (black dots) and the NOAA aircraft sites used for validation (red triangles).

already supplying valuable CO observations for 11 years. The MOPITT instrument measures upwelling radiances in a thermal-infrared (TIR) spectral band near  $4.7\mu\text{m}$  and in a short-wave infrared (SWIR) spectral band near  $2.3\mu\text{m}$ . An optimal estimation technique is used to derive CO profiles [Deeter et al., 2003]. A priori information is supplied since the optimization problem is ill-conditioned. In this paper we use MOPITT Version 4, Level 3 data [Deeter et al., 2010], which are based exclusively on TIR observations. This data comes as a daily product, gridded at a  $1^\circ \times 1^\circ$  resolution and the a priori profile, retrieved profile and the corresponding averaging kernel matrix are supplied. As for the extensively validated MOPITT V3 [e.g., Emmons et al., 2009], we use daytime observations between  $65^\circ\text{S}$  and  $65^\circ\text{N}$  only. Except in regions of strong thermal contrast, the MOPITT TIR-based V4 product is mainly sensitive to free tropospheric CO at altitudes from 4–7 km and per profile on average less than 2 independent pieces of information are inferred. Since the total column is generally retrieved more accurately than a single level [Deeter et al., 2003], we only use the CO tropospheric mean mixing ratio expressed in ppb. Figure 4 shows the differences in the CO tropospheric mean mixing ratio between MOPITT V3 and V4 for the months of March and September 2004. In general, MOPITT V4 is significantly lower compared to MOPITT V3. On the NH, differences up to 30 ppb are observed. In the SH, differences are smaller (up to 20 ppb in September) but the relative differences are as large as on the NH due to the North–south gradient in CO mixing ratios.

[21] So far, inversion studies assimilating MOPITT columns always used all MOPITT pixels over both ocean and land surfaces. However, de Laat et al. [2010] and Hooghiemstra et al. [2011] showed that MOPITT columns over deserts are biased high. De Laat et al. [2010] compared observed MOPITT V3 total columns in a latitude band over

the Sahara desert to model columns and SCIAMACHY observed columns (taking into account the averaging kernels for MOPITT and assuming the SCIAMACHY averaging kernels to be unity [de Laat et al., 2010]). They found that while all three were in good agreement over the Atlantic ocean, a sharp increase in MOPITT observed CO total columns at the land-ocean boundary was found, whereas the model and SCIAMACHY data did not show such an increase. Over the Sahara desert MOPITT total columns were on average 25% higher than model and SCIAMACHY columns. Moreover, Deeter et al. [2010] also showed that MOPITT V4 at 700 hPa was 10–30 ppb higher compared to the NOAA station Assekrem, Algeria. Hooghiemstra et al. [2011] conducted a global inversion for the year 2004 using NOAA surface flasks and compared both the prior and the posterior simulation with MOPITT V4 columns. They found differences over the Sahara desert of 15% between MOPITT and the model simulation for both the prior and the posterior simulations. Figure 5 shows the mean modeled and observed total columns between  $15\text{--}26^\circ\text{N}$  for all longitudes and the differences (in black on the right axis) with the columns simulated using the prior emissions. MOPITT columns are up to 20% (and 20 ppb) higher over the Sahara desert and the Arabian Peninsula, located between longitudes  $-15^\circ$  and  $55^\circ\text{E}$ . This discrepancy can not be explained by emissions or by transport. Therefore, we decided not to assimilate MOPITT land pixels in our 4D-Var system in this study. One might expect an unbalanced or even biased system due to this approach as the SH contains more ocean surface compared to the NH and is thus heavier constrained by the observations. However, a sensitivity study using all MOPITT observations (including land pixels) showed only large differences in inferred emission estimates for Africa (see Figure 13). Moreover, this inversion was not able to reduce the prior mismatch over the Sahara desert completely due to



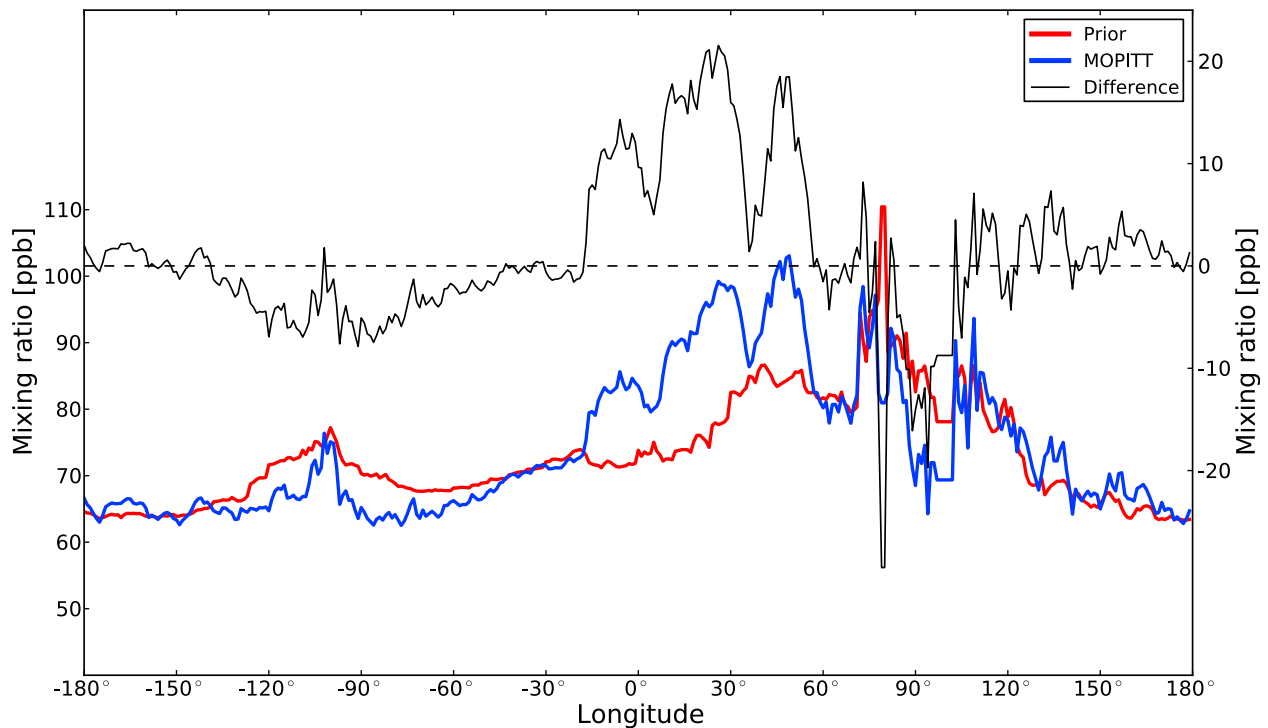
**Figure 4.** Column-averaged CO mixing ratios from MOPITT products (left) V3 and (right) V4 for the months (top) March and (bottom) September 2004.

the lack of emissions in this region in the prior emission inventory. Sensitivity studies will be discussed in detail in section 4.

[22] The contribution to the observational part of the cost function for a MOPITT observation is in principal calculated

in the same way as for the surface stations described above.

Thus, the costs are defined as  $\frac{1}{2} \left( \frac{y_m - y}{\sigma_{obs}} \right)^2$ .  $y_m$  and  $\sigma_{obs}$  are detailed below. In contrast to the MOPITT V3 retrievals resulting in CO profiles in volume mixing ratios (VMR), the



**Figure 5.** MOPITT observed (blue) and prior modeled (red) CO total columns for a latitude band from 15–26°N as a function of longitude. Differences (MOPITT-model) are given in black on the right axis.

MOPITT V4 retrieved profiles are modeled as  $\log(\text{VMR})$  values. The V4 averaging kernels describe the sensitivity of retrieved  $\log(\text{VMR})$  to true atmospheric  $\log(\text{VMR})$  values. According to the MOPITT V4 user guide [Deeter, 2009], an in-situ or model profile should be transformed using the averaging kernel and a priori profile, resulting in a pseudo profile:

$$\ln(\mathbf{x}_{ps}) = \ln(\mathbf{x}_a) + \mathbf{A}(\ln(\mathbf{x}_m) - \ln(\mathbf{x}_a)), \quad (2)$$

where  $\mathbf{x}_{ps}$  is the resulting pseudo profile,  $\mathbf{x}_a$  is the MOPITT V4 a priori profile,  $\mathbf{A}$  is the MOPITT V4 averaging kernel and  $\mathbf{x}_m$  is the modeled CO profile, interpolated to the MOPITT pressure grid. The logarithms in equation (2) however, require the arguments to be positive. Due to the large Gaussian prior errors assigned to the emissions, negative emissions may arise during the iterative 4D-Var process. These negative emissions may occasionally lead to negative model profiles ( $\mathbf{x}_m$ ) and invalid logarithms in equation (2). Another disadvantage of the formulation in equation (2) is that it leads to a non-linear observation operator in the 4D-Var framework and prevents us from using the conjugate gradient method. This method has the important advantage that posterior emission uncertainties can be easily computed [Meirink et al., 2008b]. Therefore, we have chosen to approximate the averaging kernel to first order (derivation in Appendix A), resulting in an averaging kernel  $\tilde{\mathbf{A}}$  that can be used in the following way to construct the pseudo profile:

$$\mathbf{x}_{ps} = \mathbf{x}_a + \tilde{\mathbf{A}}(\mathbf{x}_m - \mathbf{x}_a). \quad (3)$$

[23] From the pseudo profile  $\mathbf{x}_{ps}$ , the scalar tropospheric-mean mixing ratio  $y_m$  is computed by

$$y_m = \frac{1}{p_{\text{surf}}} \sum_{i=1}^{N_{\text{lev}}} (x_{ps})_i (\Delta p)_i, \quad (4)$$

where  $p_{\text{surf}}$  is the surface pressure,  $N_{\text{lev}}$  is the number of levels for the MOPITT profile (10 or less depending on orography) and  $\Delta p$  is the vector of layer thicknesses in pressure units. We analyzed the differences in modeled CO columns using equation (3) compared to equation (2). The global monthly mean differences are typically within 2% (1 ppb). However, larger regional differences up to 10% (15 ppb) may occur as shown in Figure 6. These differences also vary over the year as the linearized approach leads to higher model columns on the SH and the NH Tropics, but slightly lower columns on the NH midlatitudes in March 2004 (Figure 6, top). For September 2004, higher model columns are found over much of the NH. In the SH both larger and smaller model columns are present when using the linearized averaging kernel compared to the formulation of equation (2). Hence, we note that this approach may introduce a small bias and thus slightly biased emission estimates. However, a sensitivity study in which we explicitly corrected for the difference between application of equation (3) and equation (2) by subtracting the difference from the model columns in the prior simulation, led to optimized emissions well within the error bounds of the base inversion for each emission category (see Table 5).

[24] For multiple  $1^\circ \times 1^\circ$  MOPITT profiles in the same  $6^\circ \times 4^\circ$  model grid box (up to 24), we use the same model profile  $\mathbf{x}_m$  to compute  $\mathbf{x}_{ps}$  in equation (3) and  $y_m$  in equation (4). However, since every MOPITT retrieval has its own prior and averaging kernel, the values of  $y_m$  in the same model grid box will differ. Due to the varying orography in a grid box, the surface pressure defined on the  $6^\circ \times 4^\circ$  grid may differ from the retrieved surface pressure for the MOPITT observation that is given on  $1^\circ \times 1^\circ$ . To solve this, a surface pressure filter adopted from Bergamaschi et al. [2009] is used in which only observations are used with a surface pressure that is within 25 hPa of the model surface pressure.

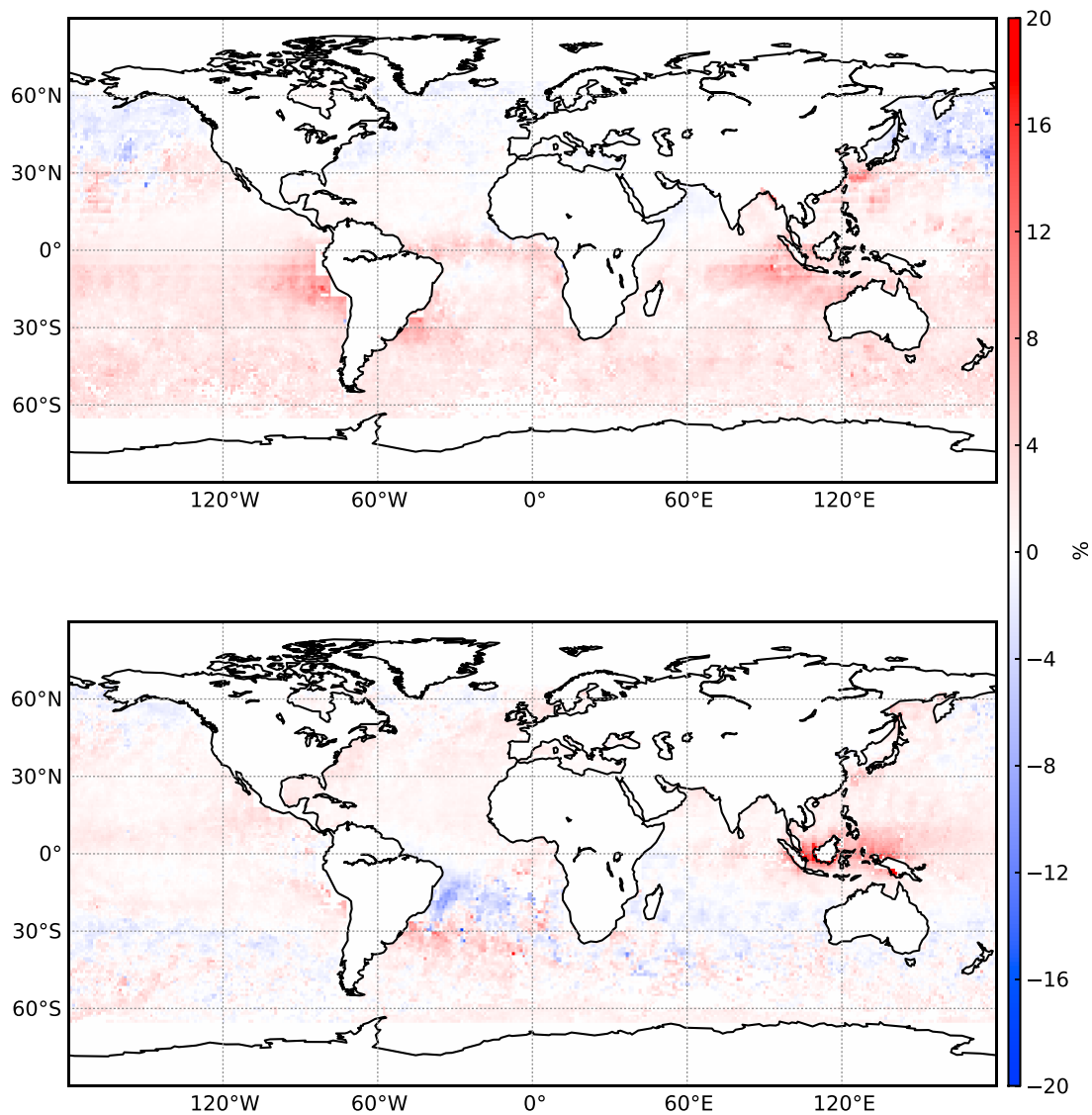
[25] For the MOPITT observations we specify an observation error ( $\sigma_{\text{obs}}$ ) for each ( $1^\circ \times 1^\circ$ ) observation. This error consists of a model error  $\sigma_{\text{mod}}$  and two types of measurement error ( $\sigma_{\text{unc}}$  and  $\sigma_{\text{var}}$ ) such that

$$\sigma_{\text{obs}} = \sqrt{\sigma_{\text{mod}}^2 + \sigma_{\text{unc}}^2 + \sigma_{\text{var}}^2}. \quad (5)$$

The  $\sigma_{\text{var}}$  is given in the MOPITT product and represents the variability of all MOPITT profiles falling in the  $1^\circ \times 1^\circ$  box. The model error  $\sigma_{\text{mod}}$  is non-zero only if multiple MOPITT observations fall in the same  $6^\circ \times 4^\circ$  model grid box and is defined as the standard deviation of the modeled CO total columns  $y_m$  within that grid box. The dominant part of the observation error is  $\sigma_{\text{unc}}$  which represents the uncertainty in the MOPITT observation. The resulting  $\sigma_{\text{obs}}$  is approximately 10% per observation. As for the surface flask observations, we do not include correlations between observations in the observation error covariance matrix. However, correlations are present in both the observations (as roughly the same air mass might be sampled more than once) and in the modeled columns. So far, similar studies have ignored correlations between observations by rebinning the observations on larger spatial scales, e.g., the model resolution. Chevallier [2007] performed an Observing System Simulation Experiment (OSSE) for  $\text{CO}_2$  using simulated OCO measurements binned to the  $3^\circ \times 2^\circ$  model resolution. They investigated the effect of different treatments for the observations on the inferred emissions. It turned out that the best results are obtained by inflation of the observation errors as an approximation to taking all correlations between observations into account. More recently, Mukherjee et al. [2011] introduced the statistical CAR model to take care of observation correlations. The statistical model is described by a few parameters that are jointly optimized with the emissions in the inversion. In addition, this approach was capable to fill in missing observations. Although they showed this approach to be appealing, it was only applied in a so-called big-region approach [Stavrakou and Müller, 2006] in which the length of the state vector remains small. However, Mukherjee et al. [2011] state that this approach is scalable to larger state vectors typically used in 4D-Var systems.

[26] Chevallier [2007] used an arbitrary error inflation factor of 2 in combination with observations that were binned to a  $3^\circ \times 2^\circ$  model resolution. Would Chevallier [2007] have assimilated the observations on a  $1^\circ \times 1^\circ$  resolution, the number of observations would have roughly scaled with a factor 6 and hence it is expected that the





**Figure 6.** Relative model column differences using the linearized averaging kernel equation (3) minus model columns using the lognormal averaging kernels equation (2) for (top) March and (bottom) September 2004.

observational part of the cost function also increases by a factor 6. Therefore, since we assimilate the MOPITT columns on a  $1^\circ \times 1^\circ$  resolution, we should reduce the cost function by inflation of the error by an additional factor of  $\sqrt{6}$ . Therefore, we initially used an inflation factor of  $\sqrt{24}$ , but this led to unrealistic emission estimates as the observations were overfitted (possibly due to the grid-scale emission error of 250%). We ultimately chose a rather large inflation factor of  $\sqrt{50}$ . With this choice we obtained an observational cost function value for the MOPITT data set that was roughly twice the size of the corresponding cost function for the stations-only inversion. The rather large factor of  $\sqrt{50}$  is justified by the fact that there are unknown correlations between the MOPITT observations. Moreover, in a future joint assimilation one needs to balance the observational costs of the individual data sets, otherwise the system may fit mainly the satellite observations and the fit with the

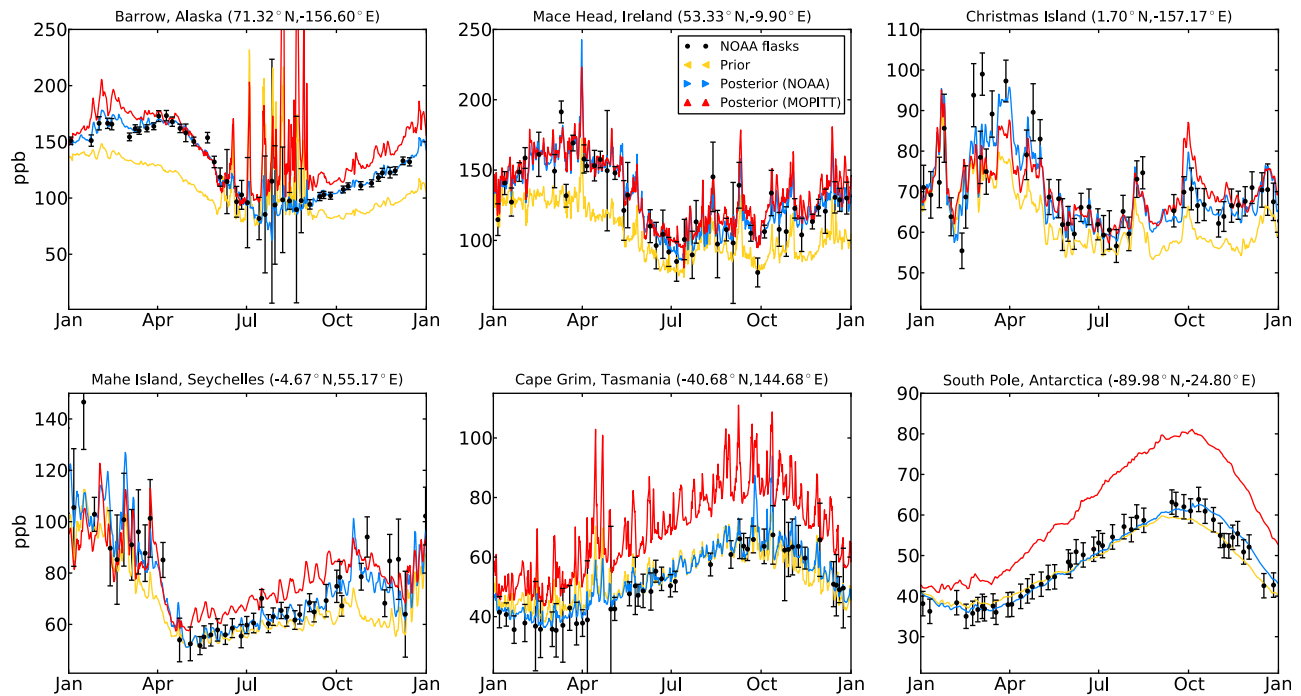
stations might deteriorate significantly in the SH as reported in previous studies [e.g., *Arellano et al.*, 2006; *Kopacz et al.*, 2010; *Fortems-Cheiney et al.*, 2011].

#### 2.4. Observations Used for Validation

[27] The assimilation of CO observations leads to emission changes with respect to the prior emissions. Model simulations using the optimized emission estimates are compared to independent observations for validation. If the inversion yields a more realistic model state, the agreement with non-assimilated observations should improve from the prior to the posterior simulation. Below we describe the observations used for validation in this study.

##### 2.4.1. Aircraft Observations

[28] In addition to surface flasks, NOAA also samples flask data using aircraft. These observations are mainly over North America and below 8 km altitude. In this work we



**Figure 7.** Prior and posterior simulation for 2004 sampled at 6 NOAA stations that were assimilated in the 4D-Var inversion stations-only. The simulation in red used optimized emissions from the MOPITT-only inversion. Black dots represent the NOAA flask observations. Error bars denote the total observation error (including the model representativeness error). The comparison with additional stations is presented in the auxiliary material.

compare modeled CO mixing ratios to flask measurements at altitudes  $>2$  km above the sites shown in Figure 3 as red triangles.

[29] The MOZAIC (Measurement of OZone, water vapour, carbon monoxide and nitrogen oxides by Airbus In-service airCRAFT) program produces in-situ measurements of CO during commercial flights [Nedelec *et al.*, 2003]. These flights are mainly over the NH from Europe to the US, Asia and the Middle East. The SH is poorly covered by these flights. We compare model CO with in-situ measurements at altitudes  $>2$  km to validate our results above the polluted boundary layer. A large fraction of these data is sampled at aircraft cruise altitude (10–12 km). Hence, in the mid and high latitudes of the NH, these flights cross the stratosphere in which the model chemistry and also the vertical transport are less accurate. These measurements are therefore omitted from the comparison.

#### 2.4.2. FTIR Total Column Observations

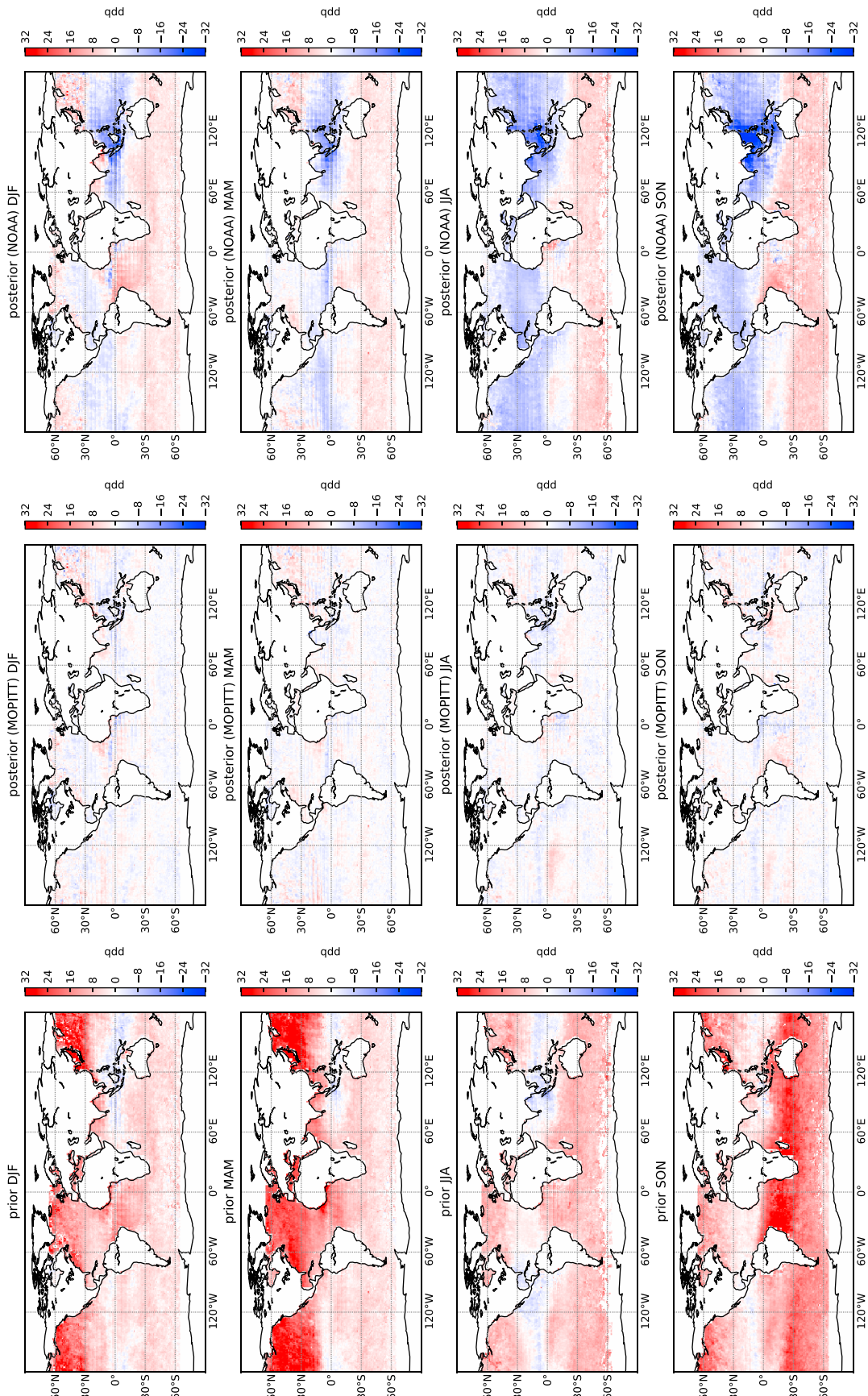
[30] Several Fourier-Transform Infrared Spectrometer (FTIR) stations worldwide measure total columns from the ground. The data used in this paper is publicly available from the Web site of the Network for the Detection of Atmospheric Composition Change (NDACC: <http://www.ndsc.ncep.noaa.gov/>). We compare our modeled CO on the coarse model resolution ( $6^\circ \times 4^\circ$ ) to column data taking into account the averaging kernels if available (and also present the comparison without using the averaging kernels; see Table 3). Due to the small footprint of the FTIR measurements, the model will likely overestimate the observations in mountain regions as the model surface pressure will be larger and hence, the model column will be deeper compared

to the FTIR column. For a fair comparison, the partial model column below the FTIR surface pressure is ignored.

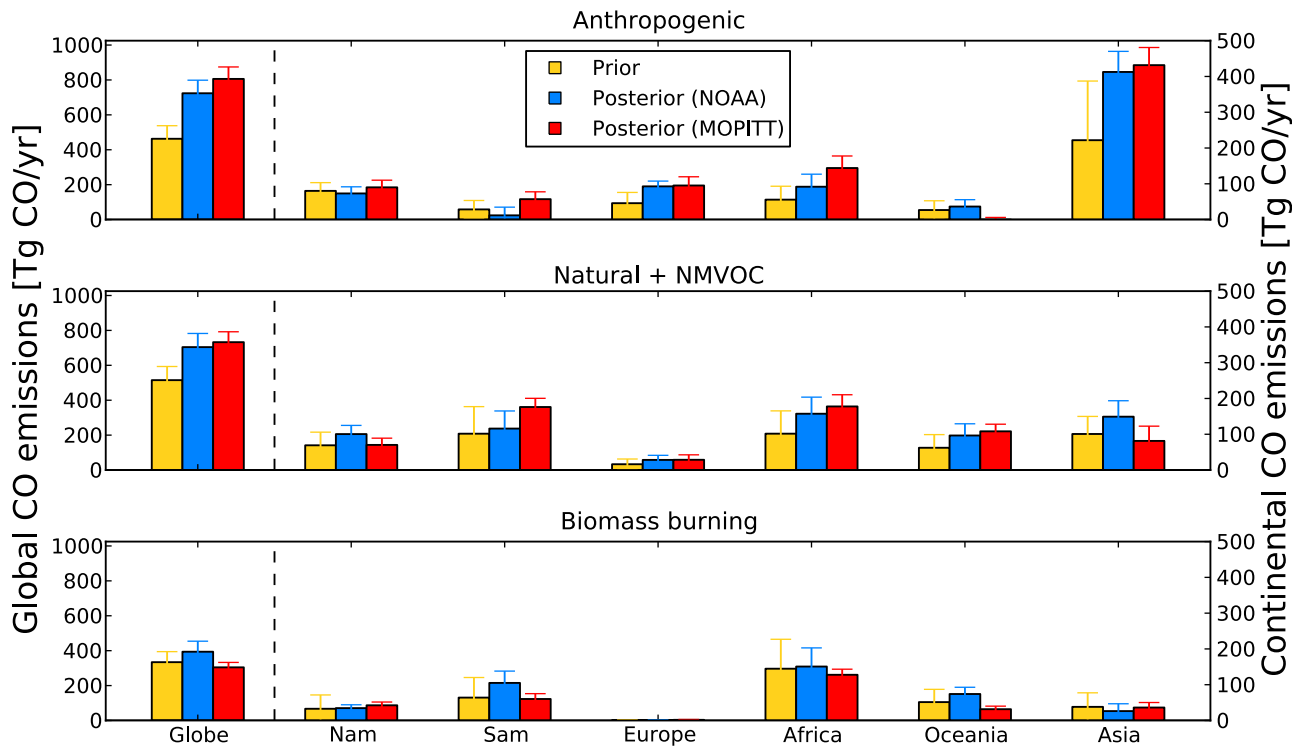
### 3. Results and Discussion

#### 3.1. Emission Increments and the Fit to the Observations

[31] We start this section with a comparison between simulated and observed CO mixing ratios for those observations that have been assimilated in the 4D-Var system. For NOAA surface network observations, Figure 7 shows the prior (yellow line) and posterior simulation for the station inversion (blue line) at 6 stations as well as the flask observations (black dots with computed  $1\sigma$  observation errors). The red line shows the posterior simulation using MOPITT derived emissions and will be discussed in section 3.3. For the NH stations (Figure 7, top), the prior simulation underestimates the observations whereas for the SH (Figure 7, bottom), the prior simulation compares well with the NOAA surface observations. For the MOPITT inversion, the comparison with MOPITT total columns is shown in Figure 8. Three-monthly composites of the difference between the model simulation and the observations are shown for the prior simulation (Figure 8, left) and the posterior simulation (Figure 8, middle). Reddish colors indicate a model underestimate and bluish colors indicate a model overestimate compared to the MOPITT observations. The right most column shows the comparison for the posterior simulation using emissions derived from the stations-only inversion and will be discussed in section 3.3. As for the NOAA stations, the prior simulation (Figure 8, left) underestimates the



**Figure 8.** Column-averaged CO mixing ratio differences MOPITT-TM5 with (left) prior and (middle) posterior emissions. We show 3 month composites per season. For the winter season (December - February) we used January, February and December 2004. (right) MOPITT-TM5 for the posterior simulation with emissions inferred from the stations-only inversion.



**Figure 9.** Prior and posterior emission estimates for the three categories: (top) anthropogenic, (middle) natural + NMVOC and (bottom) biomass burning. Prior emissions are in yellow, posterior emissions assimilating NOAA surface network observations are in blue and posterior emissions assimilating MOPITT total column observations are in red. Note that the emissions on the global scale correspond to the axis on the left whereas the continental emissions correspond to the axis on the right. Error bars show the  $1-\sigma$  uncertainty in the aggregated emissions, based on the approximated inverse Hessian. See text for details.

MOPITT observations in the NH. In contrast to the NOAA stations, however, in the SH the prior simulation also underestimates MOPITT total columns. This inconsistency in the SH for the prior simulation compared to the observations will result in different optimized emissions. Below the inferred emission estimates from the two inversions are discussed.

[32] Figure 9 shows the prior (yellow) and posterior (blue (NOAA) and red (MOPITT)) emissions for the three emission categories for the continents and the globe. The large global increment in anthropogenic emissions is mainly attributed to Asia and to a lesser extent to Europe and Africa. Both inversions yield significantly higher emissions for Asia than the new EDGARv4.1 inventory. Table 1 reports the emissions including the uncertainties as calculated from the posterior error covariance matrix. The anthropogenic source in Asia is increased by 191 Tg CO/yr and 210 Tg CO/yr, respectively for the stations-only and the MOPITT-only inversion. In addition, the uncertainty reduction for both inversions is large (65% and 70%) as the observations constrain the emissions in the region well and because the uncertainty assigned to the prior emissions was large. However, the spatial correlation lengths of 1000 km may lead to some aggregation error and hence a slightly overestimated uncertainty reduction [Meirink *et al.*, 2008b]. The emission increment leads to an improved fit with the MOPITT observations (Figure 8, middle). For Europe, optimized anthropogenic

emissions are nearly a factor 2 higher than the prior estimates, indicating that EDGARv4.1 is also too low for Europe. The 50% uncertainty reduction for the stations-only inversion is larger than for the MOPITT-only inversion (20%), likely due to the three high-latitude European stations (Ocean Station M (STM), Pallas, Finland (PAL) and Ny-Alesund, Spitsbergen (ZEP)) that constrain the emissions well. This is further illustrated in Figure 10 showing the grid-scale annual uncertainty reduction. Note that we only use MOPITT CO columns over the oceans between  $65^{\circ}\text{S}$  and  $65^{\circ}\text{N}$ , which typically lead to less uncertainty reduction for the NH continents. North American emissions remain close to the prior estimates, indicating that the inferred emissions are consistent with EDGARv4.1. In general, the NOAA surface observations and MOPITT total column observations result in comparable emission estimates in the NH.

[33] When we compare the posterior emissions for the SH continents, a first difference between the inversions in the total emission increment is observed: The large prior model underestimate compared to MOPITT total columns in the SH (Figure 8) results in higher posterior emissions compared to the stations-only inversion for South America (+60 Tg CO/yr) and Africa (+50 Tg CO/yr). However, the opposite is true for Oceania, where lower emissions are inferred (−55 Tg CO/yr) when assimilating MOPITT compared to the assimilation with NOAA stations. For the complete SH, the stations-only and MOPITT-only inversion yield 568 and 683 Tg CO/yr,

**Table 1.** Prior and Posterior Emission Estimates per Emission Category, Aggregated to Continental-Scale Regions for the Stations-Only and the MOPITT-Only Inversions in Tg CO/yr

Region	Prior	Posterior	
		Stations-Only	MOPITT-Only
<i>Anthropogenic</i>			
Nam	80 ± 23	73 ± 19	90 ± 20
Europe	46 ± 30	93 ± 15	95 ± 24
Asia	222 ± 165	413 ± 58	432 ± 49
Sam	28 ± 25	12 ± 23	57 ± 21
Africa	56 ± 37	92 ± 35	144 ± 34
Oceania	27 ± 26	36 ± 19	-11 ± 17
NH	401 ± 175	681 ± 66	714 ± 61
SH	62 ± 37	43 ± 32	92 ± 29
Globe	463 ± 180	724 ± 75	806 ± 69
<i>Natural + NMVOC</i>			
Nam	69 ± 37	100 ± 24	70 ± 19
Europe	16 ± 15	28 ± 13	29 ± 14
Asia	101 ± 49	149 ± 45	81 ± 41
Sam	102 ± 76	116 ± 49	176 ± 24
Africa	102 ± 64	157 ± 46	178 ± 33
Oceania	62 ± 37	74 ± 19	108 ± 20
NH	303 ± 87	448 ± 62	299 ± 53
SH	212 ± 82	256 ± 51	434 ± 29
Globe	515 ± 128	704 ± 78	733 ± 60
<i>Biomass Burning</i>			
Nam	33 ± 39	35 ± 9	42 ± 9
Europe	1 ± 1	1 ± 1	2 ± 1
Asia	38 ± 39	26 ± 20	36 ± 14
Sam	64 ± 56	105 ± 33	60 ± 15
Africa	145 ± 82	151 ± 52	128 ± 16
Oceania	51 ± 36	74 ± 19	32 ± 9
NH	144 ± 71	126 ± 42	148 ± 21
SH	191 ± 96	268 ± 43	157 ± 26
Globe	334 ± 119	394 ± 60	304 ± 28

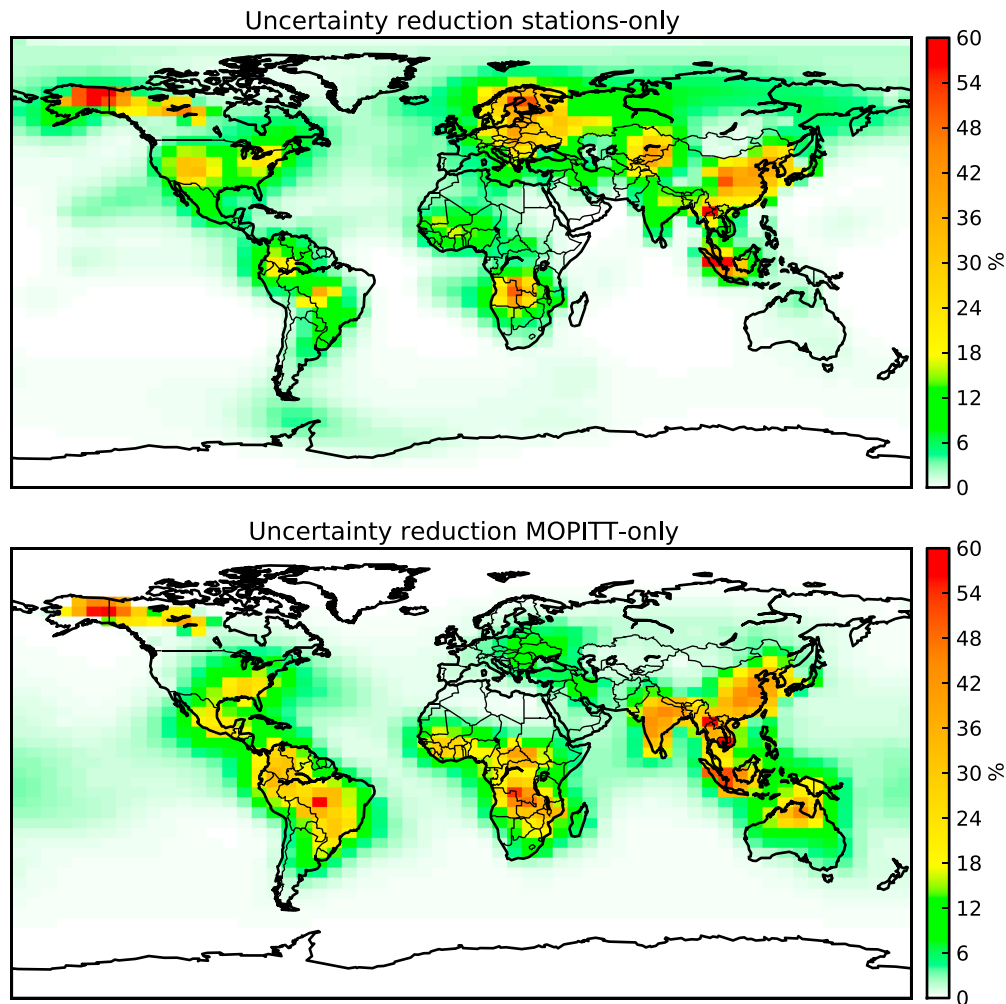
respectively. Due to the 3-day global coverage of MOPITT, the uncertainty reduction for the SH continental regions is much larger compared to the stations-only inversion (Figure 10). For example, African and South American biomass burning emissions show an uncertainty reduction of 80 and 73% respectively for the MOPITT-only inversion. The stations-only inversion results in an uncertainty reduction of 37 and 41% for those two regions, but this is likely due to the large prior errors we assigned to the emissions. In general, the MOPITT-only inversion constrains in particular the emissions in the Tropics, whereas the surface stations constrain the NH.

[34] Furthermore, a remarkable shift from the biomass burning source to the natural source is observed (Table 1 and Figure 9). For example, the stations-only inversion increases GFED3.1 biomass burning emissions for South America (+41 Tg CO/yr), Africa (+6 Tg CO/yr) and Oceania (+23 Tg CO/yr), whereas the MOPITT-only inversion decreases biomass burning emissions for these regions (-4 Tg CO/yr, -17 Tg CO/yr and -19 Tg CO/yr for South America, Africa and Oceania, respectively). However, the MOPITT-only inversion increases the natural source (mainly CO from NMVOC oxidation) significantly. For example, natural emissions are roughly doubled over South America, Africa and Oceania with respect to the prior, whereas the stations-only inversion shows much smaller increments for this

source in these regions. The shift in emission increments from the biomass burning source to the natural source can be explained as follows: Due to the observations (from MOPITT or the NOAA surface network), the emissions over SH continents are required to increase. However, this increment can be added to either of the three sources in a certain region. For the stations-only inversion, since there are almost no surface stations close to the source regions in the SH, it is cheapest in terms of costs in the cost function to increase the biomass burning emissions that stay much closer to the surface compared to the natural emissions and only take place in a specific part of the year (the dry season) and over relative small areas (compared to the natural emissions). These increased biomass burning emissions will be partly diluted and partly chemically removed in the atmosphere and thus only slightly enhance CO concentrations at the stations. The MOPITT instrument is more sensitive to the natural source than the biomass burning source, as the NMVOC-CO in the natural source is released higher up in the troposphere (see Figures 1a and 1d for the vertical distributions of biomass burning and the natural source, respectively). Also, since the natural emissions are specified in the prior throughout the year and over larger geographical areas compared to the biomass burning emissions, increasing this source results in a reduction of the prior mismatch between model and observations with minimal costs in the background part of the cost function. Moreover, increasing biomass burning as in the stations-only inversion does not improve the agreement with MOPITT columns (Figure 8, right) as the model overestimates CO in the Tropics (in particular in Indonesia). We acknowledge that deficiencies in the vertical distribution of the natural emissions, i.e., a too high injection height may cause model data mismatches to be projected on the natural emissions, specifically for the MOPITT-only inversion.

[35] In conclusion, the difference in vertical sensitivity of the two observational data sets (NOAA stations are mainly sensitive to boundary layer CO, whereas MOPITT is mainly sensitive to lofted CO), and the higher spatiotemporal resolution of the MOPITT observations and thus better global coverage leads to a shift in the partitioning of the emissions into different source categories. Furthermore, inconsistencies in the prior mismatch between model and observations lead to different emission increments for the MOPITT-only inversion compared to the stations-only inversion.

[36] It should be noted here that the inverse modeling system can only optimize total CO emissions given a certain mismatch between model and observations. Hence, separation of different sources is only possible if realistic spatial and temporal information from prior inventories is supplied to the inversion system. In the current inversions the system has difficulties to separate the anthropogenic from the natural CO source. Aggregated to continental scales and an annual time scale, the derived posterior correlation coefficient ranges from -0.62 for South America to -0.88 for North America. The system is better capable to separate the biomass burning source from the anthropogenic source due to differences in the spatial patterns of the prior emissions. The posterior correlations are less than -0.2. Although in particular in the Tropics the spatial patterns of the biomass burning and natural CO emissions overlap, the specific



**Figure 10.** Uncertainty reduction for the (top) stations-only and (bottom) MOPITT-only inversions. Uncertainty reduction is defined as  $100\% \left(1 - \frac{\sigma_a}{\sigma_b}\right)$ , where  $\sigma_a$  is the standard deviation of the posterior emission (aggregated over some time window and or spatial region) and  $\sigma_b$  is the standard deviation corresponding to the prior emissions. Here we show the uncertainty reduction for the total emissions (anthropogenic + natural + biomass burning) on the grid-scale, aggregated over the whole year 2004.

timing of biomass burning in the dry season leads to posterior error correlations on yearly time scales of around  $-0.3$ .

### 3.2. Comparison With Recent Studies

[37] Our emission estimates are compared with four recent global inversions, all for the year 2004 (or parts of that year). *Jones et al.* [2009] assimilated MOPITT V3 and TES (Tropospheric Emission Spectrometer) observations separately for November 2004 and we compare our results with their MOPITT-based emissions. *Kopacz et al.* [2010] assimilated observations from three satellite instruments (AIRS, MOPITT V3 and SCIAMACHY). However, the number of AIRS measurements used was three times higher compared to the number of MOPITT observations and even 36 times larger than the number of SCIAMACHY observations. *Fortems-Cheiney et al.* [2011] used MOPITT V4 (as in the current study) and *Hooghiemstra et al.* [2011] assimilated NOAA surface flasks. All satellite derived emissions in the literature studies used observations over land and ocean. The comparison is difficult due to differences in the inversion

setup (e.g., definition of the state vector and error settings), the definition of the aggregation regions, the observations that have been assimilated, and the atmospheric chemistry models used. In Table 2 we report the sum of the anthropogenic and biomass burning emissions. In addition, we give the global source of CO due to oxidation of NMVOCs and methane. Our North American (108 and 132 Tg CO/yr) and European (94 and 97 Tg CO/yr) emission estimates (for the stations-only and the MOPITT-only inversions, respectively) are in the low end of the range reported in the studies in Table 2. For Asia, our emission estimates are slightly lower compared to the other studies. Also in the SH, our emission estimates are somewhat lower compared to the other studies. For example, our South American emission estimates are 117 Tg CO/yr compared to previous emission estimates ranging from 141 to 184 Tg CO/yr. Similarly, for Africa, where our emission estimates of 243 and 272 Tg CO/yr are in particular lower compared to *Kopacz et al.*'s [2010] emission estimate of 343 Tg CO/yr.

**Table 2.** Comparison of Our Derived Total Emissions (Sum of Anthropogenic and Biomass Burning Emissions) Using Either NOAA Stations or MOPITT Observations With Recent Values From Literature<sup>a</sup>

Region	This Study		<i>Jones et al.</i> [2009]	<i>Kopacz et al.</i> [2010]	<i>Fortems-Cheiney et al.</i> [2011]	<i>Hooghiemstra et al.</i> [2011]
	Stations	MOPITT				
Nam	108 ± 21	132 ± 22	146	71	199	166
Europe	94 ± 15	97 ± 24	111	95	137	92
Asia	439 ± 56	468 ± 48	531	660	506	597
Sam	117 ± 37	117 ± 24	141	183	184	156
Africa	243 ± 62	272 ± 38	304	343	283	338
Oceania	110 ± 26	21 ± 19	185	-	117	-
Subtotal	1118 ± 88	1110 ± 71	1417	1350	1441	1390
NMVOC-CO	704 ± 78	733 ± 60	1344	1290	1176	410
CH <sub>4</sub> -CO	865	865				887
Total	2687	2708	2762	2642	2602	2646

<sup>a</sup>All studies shown here performed an inversion for the year 2004 (or parts of that year). The global estimate of CO from oxidation of NMVOCs and methane and the total CO production for 2004 are also given. For studies that only reported an oxidation source of CO (from methane and NMVOCs), we report that number.

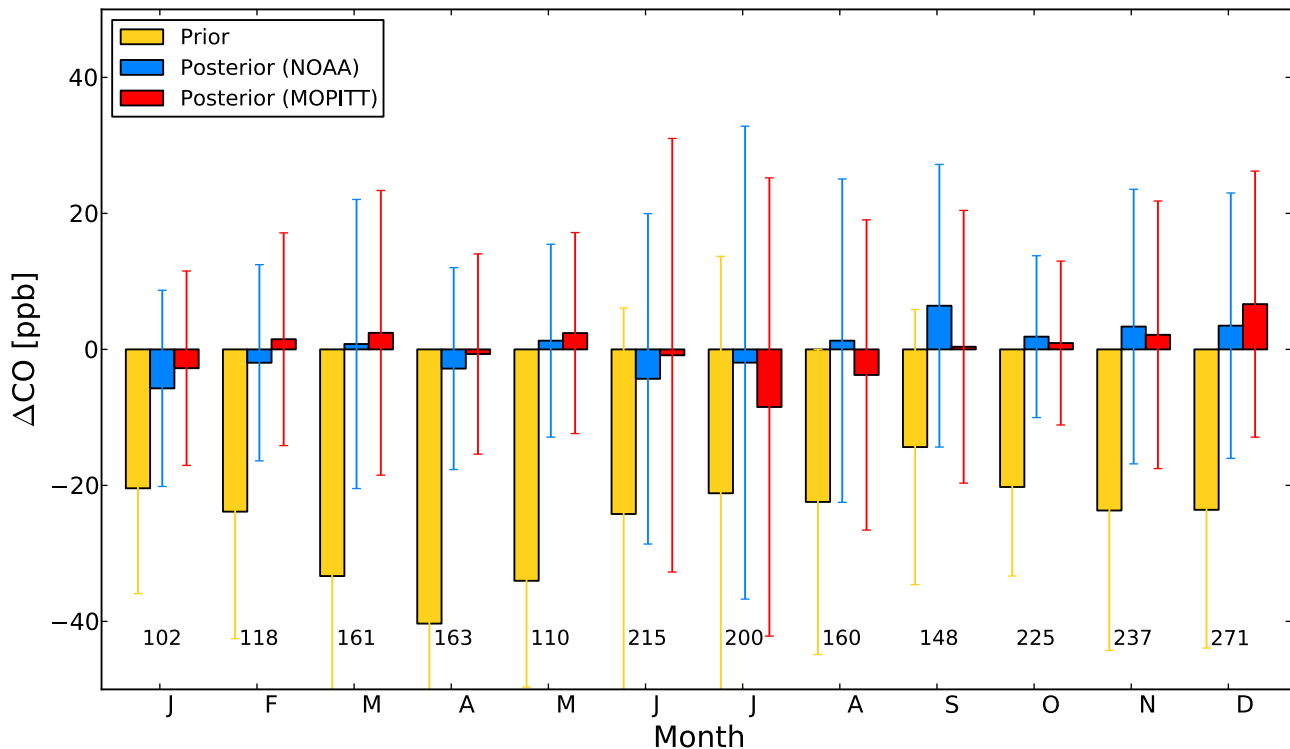
[38] The main reason for these differences is likely our approach to (1) optimize the NMVOC-CO source on the model resolution and (2) using a vertical distribution which releases significant CO in the free troposphere. Some recent studies [*Jones et al.*, 2009; *Kopacz et al.*, 2010] optimized the CO production from oxidation of NMVOCs by a single parameter. But for these studies, posterior production terms for oxidation of methane and NMVOCs remained close to the prior terms, possibly due to too tight error settings on these sources. As a consequence of this approach, any prior mismatch between the model and observations will be projected on either the biomass burning emissions or the anthropogenic emissions. In our previous study [*Hooghiemstra et al.*, 2011] we also used monthly global scaling parameter for NMVOC-CO and CH<sub>4</sub>-CO sources. The relatively large differences in emissions estimates in the current study compared to the results of *Hooghiemstra et al.* [2011] is explained by (1) the aggregation of the NMVOC-CO source in a single parameter and (2) the compensation mechanism which is extensively described by *Hooghiemstra et al.* [2011]. In short, since the observations mainly constrain total emissions, an increase in one emission category may be compensated for by a decrease in another emission category. In the current setup, the natural emissions (see Table 1) increase in both inversions for all regions. This results in global natural emissions of 704 or 733 Tg CO/yr (for the stations-only and the MOPITT-only inversion, respectively). When we add the CH<sub>4</sub>-CO, our total oxidation source of CO amounts to 1569 and 1598 Tg CO/yr (for the stations-only and the MOPITT-only inversions, respectively). This is in contrast with the study of *Fortems-Cheiney et al.* [2011], in which the posterior CO production through oxidation of formaldehyde was reduced compared to the prior (*A. Fortems-Cheiney*, personal communication, 2011) and resulted in a total oxidation source of CO of 1176 Tg CO/yr. The main reason for this difference is likely our large prior grid-scale error of 250%. However, this error choice is justified because also the NMVOC-CO source strength is uncertain and more constraints on this emission category seem needed. Joint assimilation of formaldehyde and CO columns may lead to more accurate emission estimates particularly in the Tropics. *Stavrakou et al.* [2009] for example, used space-based formaldehyde columns to infer isoprene emissions. Although their inversion results were close to the prior emission

estimates on a global scale, large emission increments (up to 55%) were found regionally.

[39] If we compare biomass burning emissions only, our optimized biomass burning emissions from the MOPITT-only inversion are not in agreement with recent studies. For example, *Kopacz et al.* [2010] and *Liu et al.* [2010] found that the GFEDv2 biomass burning inventory was too low by up to a factor 2. Here we started from the more recent GFEDv3 [*van der Werf et al.*, 2010] inventory that is even lower than GFEDv2 by about 70 Tg CO/yr globally and posterior biomass burning emission estimates are another 20 Tg CO/yr reduced (Table 1). *Kopacz et al.* [2010] inverted CO emissions using 3 satellite instruments, including MOPITT. They optimized the total CO surface emissions and attributed large corrections to the total prior emissions in the biomass burning season to deficiencies in the GFEDv2 product. However, *Arellano et al.* [2006] showed that both the anthropogenic source and the biomass burning source increased from the prior to the posterior estimate even during the biomass burning season. This indicates that not all increments in this period should be attributed to biomass burning emissions only. Moreover, the posterior simulation corresponding to the optimized emissions from the joint inversion were not in agreement with MOPITT V3 columns. This was shown to be caused by a not-corrected positive bias in the AIRS observations on the SH (>10%) in combination with the large weight of these observations. *Liu et al.* [2010] performed forward model simulations with two sets of meteorological data (GEOS-4 and GEOS-5) and compared the simulations with TES and MLS data. They showed that simulations with either bottom-up emission estimates or bottom-up emissions scaled using parameters derived by *Kopacz et al.* [2010] were not fully consistent with the observations in the Tropics. They concluded that apart from deficiencies in the emissions, meteorological fields and model transport may dominate model-data mismatches.

### 3.3. Validation of Posterior Emission Estimates With Independent Observations

[40] As a first validation step, we used so-called cross validation of the inferred emissions to validate the two inversions. In this context, cross validation means that the posterior simulation using optimized emissions from either inversion are compared to the other observational data set.



**Figure 11.** Comparison of model with independent aircraft observations  $>2$  km from NOAA, mainly over North America. Per month, the mean difference (model - observation) is given for the prior simulation (yellow), the posterior simulation using optimized emissions from an inversion assimilating NOAA surface network observations (blue) and the posterior simulation using optimized emissions from an inversion assimilating MOPITT total column observations (red). Numbers represent the amount of observations we compared with per month. The error bars are defined as the standard deviation of the differences per month.

[41] Cross validation of the MOPITT-only inversion with the NOAA stations is shown in Figure 7 (red line). Generally good agreement on NH stations is found and the fit with the stations improves even though these observations are not assimilated. In particular for Europe (represented here by station Mace Head) the red line is really close to the posterior simulation corresponding to the stations-only inversion (blue line). For the high latitude NH (represented by station Barrow, Alaska) the agreement with the observations is in particular good in the first part of the year. Biomass burning emissions from the MOPITT-only inversion are significantly higher for this region (compared to the stations-only inversion: 42 Tg CO/yr versus 35 Tg CO/yr) and modeled CO mixing ratios are therefore higher than the observations in summer. However, for the SH, the MOPITT-only simulation largely overestimates the observed CO mixing ratios on the SH stations (Mahe Island, Seychelles (+10 ppb), Cape Grim, Tasmania (+20 ppb) and South Pole station (+20 ppb); see Figure 7 (bottom)), indicating an overestimation of SH emissions, probably due to some bias in the MOPITT observations. The comparison for all assimilated stations is available in the auxiliary material.<sup>1</sup>

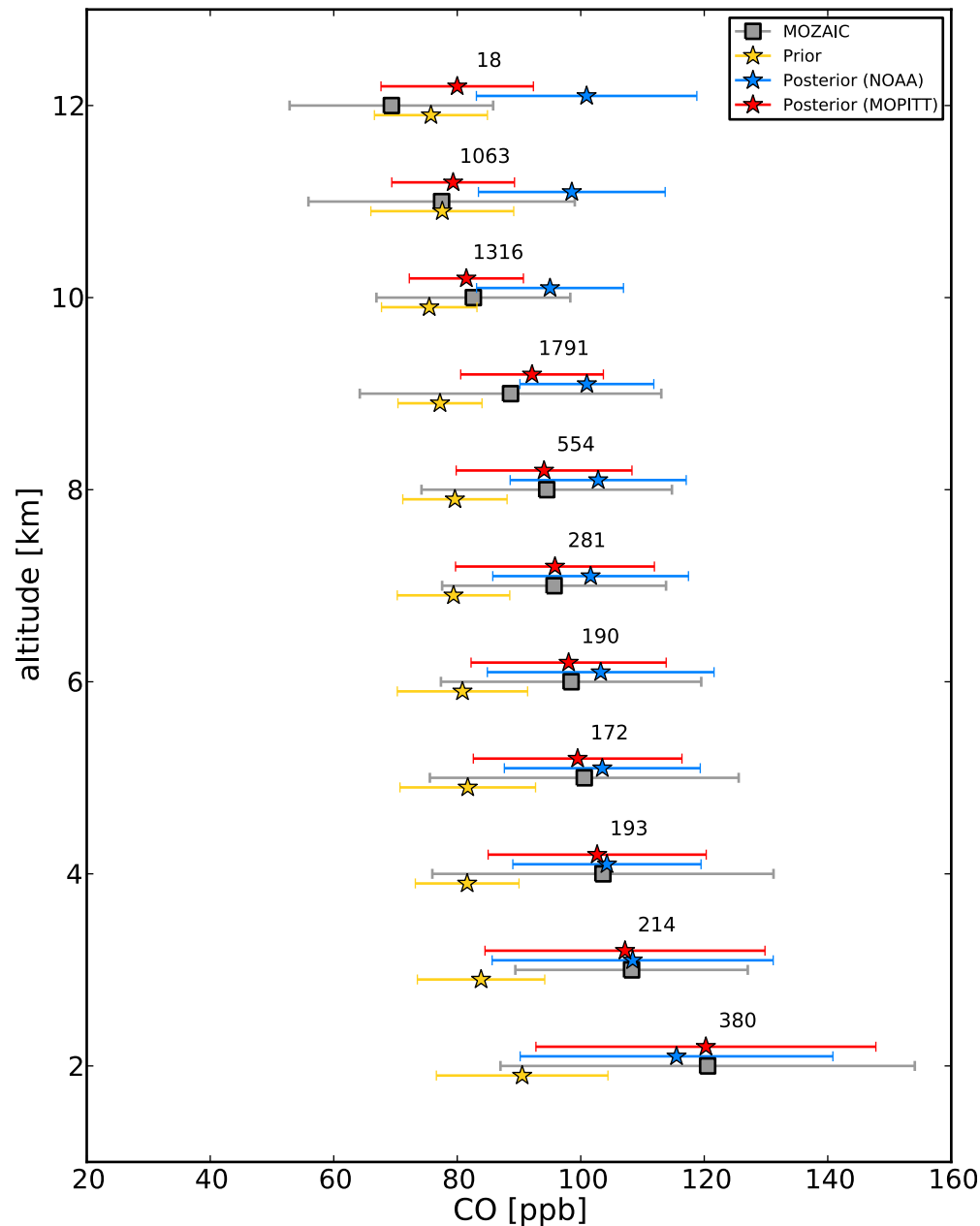
[42] Figure 8 (right) shows the cross validation the other way around. A clear improvement from prior to posterior simulation (for the stations-only inversion) in all seasons is

observed on the NH, mainly due to increased anthropogenic emissions over Asia and Europe. For the SH, the stations-only simulation underestimates the MOPITT total columns south of 30°S in all seasons by about 5–8 ppb and overestimates CO columns in particular in Indonesia (Figure 8, right). From this validation it seems likely that MOPITT CO total columns have some positive bias in the SH south of 30°S.

[43] A second validation is performed using aircraft data. Figure 11 shows monthly mean differences (modeled minus observed CO mixing ratio) for all NOAA aircraft flights above 2 km altitude. The posterior simulation improves the comparison with NOAA aircraft observations compared to the prior simulation for both inversions in very similar ways: The prior underestimate ranging from 10–40 ppb per month is reduced to less than 10 ppb for all months in both inversions. This indicates that the model reproduces CO mixing ratios over North America very well up to 8 km. This agreement was also found by *Deeter et al.* [2010] who found no significant bias when comparing MOPITT V4 data with the NOAA aircraft data. This is in sharp contrast, however, with the validations of MOPITT V3 performed by *Emmons et al.* [2007, 2009]. They reported a positive bias of  $7 \pm 9\%$  with respect to the NOAA aircraft data in 2004, but on the SH, a bias of +20% was reported. Figure 12 shows the comparison for modeled and observed CO mixing ratios from the MOZAIC program. Above 10 km altitude, only measurements south of 40°N have been used to avoid stratospheric

<sup>1</sup>Auxiliary materials are available in the HTML. doi:10.1029/2011JD017043.





**Figure 12.** Comparison of the model with MOZAIC in-situ aircraft data averaged per altitude level. Rectangular symbols represent the mean of all MOZAIC observations in the layer (of width 1 km) above the corresponding altitude. Model simulations are represented by yellow (prior), blue (posterior NOAA) and red (posterior MOPITT) stars. The error bars indicate the standard deviations corresponding to the means. Numbers indicate the amount of observations used in this validation per layer. For altitudes above 10 km, only observations south of 40°N are used to minimize stratospheric sampling.

influence. Since there are almost no flights in the SH, such a cut-off was not necessary there. The majority of the MOZAIC measurements is from flights between Europe and the US or between Europe and East Asia and the middle East. The measurements were averaged into 1 km bins from 2 to 12 km altitude and are shown as grey boxes in Figure 12, the error bars denote the corresponding standard deviation of the observations. The co-sampled model prior simulation is given in yellow and the posterior model simulations corresponding to the stations-only and MOPITT-

only inversions are in blue and red, respectively. Throughout the troposphere, the prior model simulation underestimates observed CO by 10 to 30 ppb. The largest differences are found near the surface reflecting the too low prior emissions on the NH. The posterior simulations show an improved agreement and are typically within a few ppb in the lower troposphere (3–7 km). At higher altitudes (>7 km) the agreement between observations and the MOPITT-only inversion is much better than the agreement with the stations-only inversion due to the higher sensitivity of MOPITT at

**Table 3.** Comparison With Independent FTIR Observations for Seven Sites Around the Globe<sup>a</sup>

Station	Latitude (deg)	Longitude (deg)	Altitude (m.a.s.l.)	Observed	Prior	Posterior (NOAA)	Posterior (MOPITT)
Kiruna, Sweden	67.84	20.41	419	104.8	82.6 (84.1)	107.1 (108.7)	108.0 (109.5)
Zugspitze, Germany	47.42	10.98	2964	82.4	69.0 (69.2)	88.4 (89.1)	84.0 (84.2)
Jungfrauoch, Switzerland	46.55	7.98	3580	86.7	73.6 (70.3)	92.3 (89.1)	88.7 (85.4)
Izaña, Tenerife, Spain	28.30	-16.48	2367	81.7	68.7 (69.1)	85.5 (86.0)	77.8 (78.1)
Reunion Island, France	-20.90	55.50	10	94.5	67.8 (64.4)	83.0 (79.7)	88.3 (84.9)
Lauder, New Zealand	-45.04	169.68	370	52.9	46.0	48.1	56.4
Arrival Heights, Antarctica	-77.82	166.65	250	51.5	43.1	46.4	56.1

<sup>a</sup>Reported values are averaged annual modeled and observed column-averaged CO mixing ratios in ppb. Values between parentheses represent modeled columns when the averaging kernel was not taken into account. Note that for stations Lauder, New Zealand and Arrival Heights, Antarctica, no averaging kernels were available.

this altitude. At aircraft cruise altitudes (10–12 km) this tendency continues. The model prior and MOPITT posterior simulation remain quite close to the observations at these altitudes, but the stations-only simulation largely overestimates the observations. This is attributed to the natural emissions that are approximately 140 Tg CO/yr higher on the NH in the stations-only inversion compared to the MOPITT-only inversion (see Table 1) and are injected higher up in the troposphere. Clearly this is not in agreement with both MOPITT and aircraft observations. Unfortunately, no aircraft profiles are available south of 30S, where the MOPITT bias appears to be most prominent.

[44] Further validation with FTIR total column observations showed overall good agreement on the NH (Table 3): the prior underestimate of 16–20% (13–22 ppb) is significantly reduced for all FTIR stations due to increased emissions over the NH. Although both inversions slightly underestimate the FTIR total columns at Reunion Island, this is different for the other two FTIR sites in the SH (Table 3). For Lauder, New Zealand and Arrival Heights, Antarctica, the MOPITT-only inversion results in too high CO columns for both Lauder (+7% or 3.5 ppb) and Arrival Heights (+9% or 4.6 ppb). For the FTIR measurements made in Lauder, this was already shown by *Yurganov et al.* [2010] who performed a direct comparison between MOPITT and the FTIR observations. For these FTIR sites, the stations-only inversion tends to underestimate the FTIR CO columns by 9% (4.8 ppb) and 10% (4.9 ppb) for Lauder and Arrival Heights, respectively.

#### 4. Sensitivity Studies

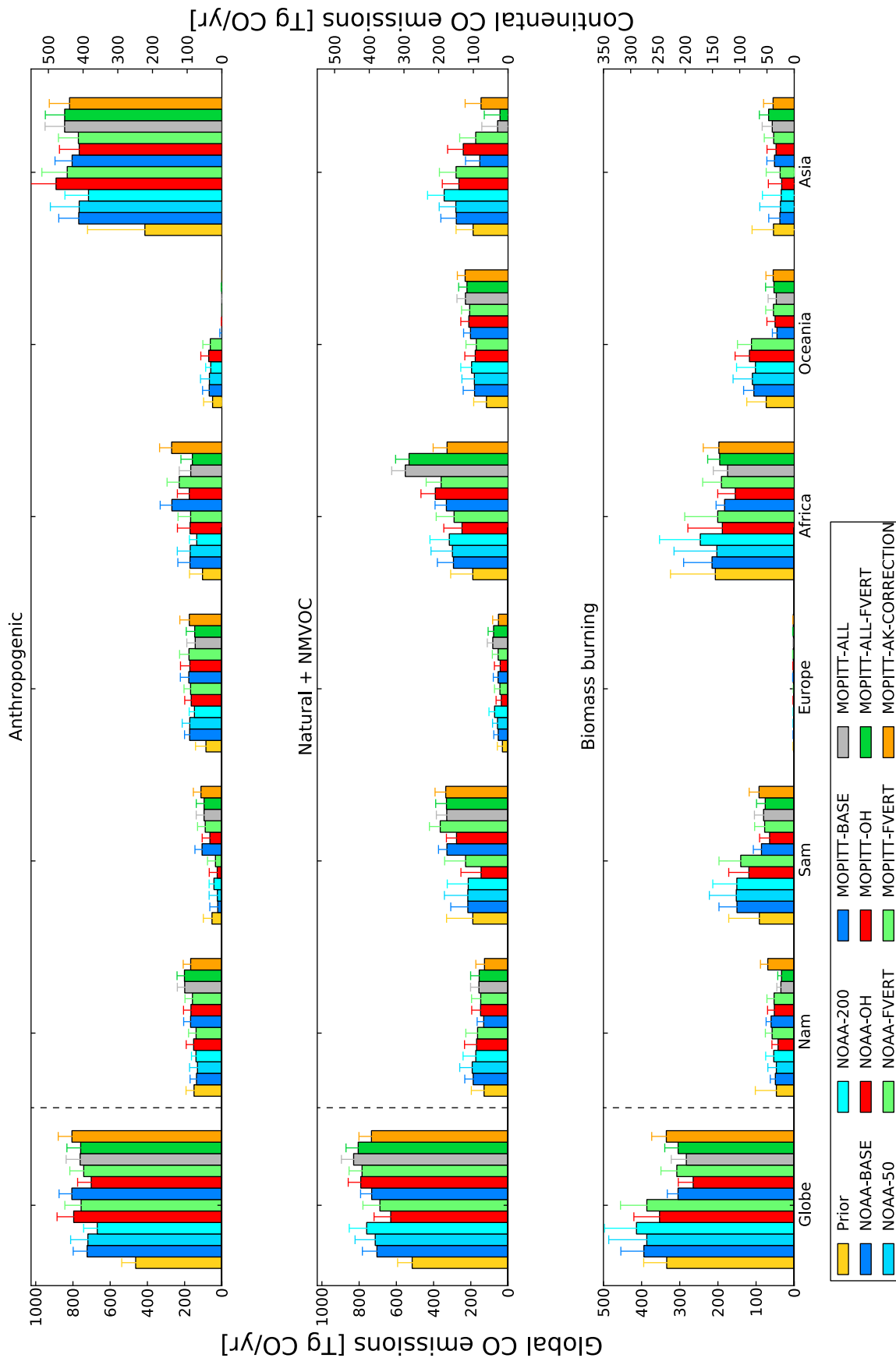
[45] Although it seems that the MOPITT-only inversion results in too high CO in the SH compared to other observations due to a positive bias in the MOPITT V4 product in

the SH, model uncertainties may also play a role. For example, the climatological OH field used and the vertical distribution of biomass burning emissions can have large impacts on the inferred emissions as shown by *Hooghiemstra et al.* [2011]. But also deficiencies in model transport and the effect of linearizing the MOPITT averaging kernels may influence the inferred emission estimates. Therefore, we present a series of sensitivity inversions to address these issues. For the stations-only inversion we performed three sensitivity simulations using (S1) a preduc factor of 50 and (S2) a preduc factor of 200 (instead of a preduc factor of  $10^6$ ) (S3) a different OH field from a full-chemistry simulation with TM5 [*Huijnen et al.*, 2010] and (S4) a different vertical distribution for biomass burning emissions (FVERT). The vertical distribution for this sensitivity inversion is shown in Figure 1b as a function of latitude and model level. This distribution is slightly different compared to the one we used by *Hooghiemstra et al.* [2011] (Figure 1c). The current choice combines the findings in reported literature that boreal forest fires inject CO at higher altitudes [*Val Martin et al.*, 2009], whereas in the Tropics, biomass burning emissions from savannah fires remain largely below 3 km, but tropical forest fires in South East Asia and Indonesia also inject biomass burning CO in the free troposphere. In contrast, the vertical distribution in a sensitivity experiment by *Hooghiemstra et al.* [2011] followed the injection height derived by *Gonzi and Palmer* [2010] that also emitted biomass burning CO in the free troposphere in the Tropics. For the MOPITT-only inversion we performed sensitivity studies (S5) and (S6) similar to studies S3 and S4 described above and in addition we assimilated (S7) all MOPITT observations (S8) all MOPITT observations with a different vertical distribution of biomass burning emissions (Figure 1b) and (S9) corrected for the linearization of the averaging kernel

**Table 4.** Details of the Sensitivity Studies Described in Section 4<sup>a</sup>

Study	Data Assimilated	OH Field	Preduc	FVERT	CORRECTION
S1	stations	<i>Spivakovsky et al.</i> [2000]	50	<2 km	–
S2	stations	<i>Spivakovsky et al.</i> [2000]	200	<2 km	–
S3	stations	<i>Huijnen et al.</i> [2010]	1000	<2 km	–
S4	stations	<i>Spivakovsky et al.</i> [2000]	1000	<2 km	–
S5	MOPITT	<i>Huijnen et al.</i> [2010]	1000	35% > 2 km	No
S6	MOPITT	<i>Spivakovsky et al.</i> [2000]	1000	<2 km	No
S7	MOPITT	<i>Spivakovsky et al.</i> [2000]	1000	35% > 2	No
S8	MOPITT	<i>Spivakovsky et al.</i> [2000]	1000	<2 km	No
S9	MOPITT	<i>Spivakovsky et al.</i> [2000]	1000	35% > 2	Yes

<sup>a</sup>See text for details.



**Figure 13.** Same as Figure 9 but for a series of sensitivity experiments. The bars corresponding to the base inversions (stations-only and MOPITT-only) are also given (in dark blue). Note that the global annual emission estimates are specified on the left axis. Error bars show the  $1-\sigma$  uncertainty in the aggregated emissions, based on the approximated inverse Hessian. See text for details.

**Table 5.** Global Emission Estimates per Emission Category for 2004 for Nine Sensitivities Studies<sup>a</sup>

	Stations-Only					
	Prior	Posterior (Base)	Preduc (50)	Preduc (200)	OH	FVERT
ANT	463 ± 180	724 ± 75	778 ± 107	747 ± 92	796 ± 89	756 ± 87
NAT	515 ± 128	704 ± 78	667 ± 108	689 ± 105	630 ± 90	688 ± 91
BB	334 ± 119	394 ± 60	416 ± 100	391 ± 95	353 ± 67	386 ± 69
Total	1312 ± 251	1822 ± 45	1860 ± 115	1827 ± 73	1779 ± 57	1831 ± 64

	MOPITT-Only						
	Prior	Posterior (Base)	OH	FVERT	All	All-FVERT	Correction
ANT	463 ± 180	806 ± 69	702 ± 73	742 ± 75	761 ± 75	758 ± 73	805 ± 74
NAT	515 ± 128	733 ± 60	791 ± 67	784 ± 70	829 ± 66	804 ± 66	734 ± 67
BB	334 ± 119	304 ± 28	265 ± 39	307 ± 42	282 ± 40	304 ± 35	335 ± 38
Total	1312 ± 251	1843 ± 12	1758 ± 14	1834 ± 16	1873 ± 11	1868 ± 10	1874 ± 16

<sup>a</sup>The results of the base inversions are also included.

(CORRECTION). The correction is applied as follows: For the prior simulation we compute the tropospheric-mean mixing ratio using either equation (2) or equation (3) and archive the differences. In subsequent iterations, these differences are subtracted from the calculated modeled tropospheric-mean mixing ratios. To reduce the computational burden we used a preduc factor of 1000 for the sensitivity studies (except for S1 and S2). The sensitivity studies are detailed in Table 4. The results are summarized below and visualized in Figure 13 and Table 5.

[46] 1. A preduc factor of 50 or 200 (S1 and S2) does influence the posterior emission estimates compared to a fully converged optimization with a preduc factor of  $10^6$ . Although on the global scale the effects are rather small, on a continental scale difference range from 10–40%. However, these differences in optimized emissions are in a similar range as the effect of a different OH field and to a lesser extend, the correction for the linearization of the averaging kernel (see below). Moreover, the comparison with independent observations does not improve from preduc 50 to preduc  $10^6$ . Still, we find a preduc factor of  $10^6$  useful since for smaller preduc values, derived posterior emission uncertainties haven't yet converged and remain larger compared to a fully converged inversion [Meirink *et al.*, 2008b]. In addition, since the small scale emissions are optimized only in these later iterations, a large preduc value is typically necessary to infer CO emissions on a higher spatial resolution (in combination with a spatial correlation length <1000 km).

[47] 2. A different OH field (S3 and S5) has a large effect. As outlined by Hooghiemstra *et al.* [2011], the TM5-based OH field has a larger North-south ratio (1.15) compared to the OH field from Spivakovsky *et al.* [2000] used in the base inversion (North-south ratio of 1.0) and thus results in higher inferred emissions on the NH and lower emissions on the SH. Magnitudes are similar to what was reported by Hooghiemstra *et al.* [2011].

[48] 3. With a different distribution of biomass burning emissions, one would expect more emissions in the stations-only inversion (S4), as these emissions are lofted and not picked up by the stations, and lower emissions for the MOPITT-only inversion (S6), as the emissions are more easily observed by the instrument. However, the sensitivity to this distribution appears to be only small. This is in

contrast with our previous study [Hooghiemstra *et al.*, 2011], where the difference was as large as 70 Tg CO/yr. This is likely due to the new implementation of the NMVOC-CO source, but the lower injection height in the Tropics compared to Hooghiemstra *et al.* [2011] might also play a role. For the MOPITT-only inversion the sensitivity to a different biomass burning injection height is negligible. This is explained by comparing the vertical distribution to the distribution for the NMVOC-CO source (Figures 1b and 1d). Since the NMVOC-CO emissions are injected much higher, the MOPITT-only inversion is mainly sensitive to this source and less sensitive to biomass burning injection heights.

[49] 4. Including MOPITT land observations (S7 and S8) makes a minor difference except for Africa and Asia. Due to the absence of CO sources in the prior over the Sahara desert where the large mismatch is found, natural emissions increase heavily in Africa. As a compensating result, this source decreases for Asia. Since the emissions from the inversion including MOPITT land observations deviate more from the stations-only inferred emissions, this points to a positive bias of MOPITT over desert areas.

[50] 5. Correcting for the linearization of the averaging kernel (S9) results in higher anthropogenic emissions (in Asia and Africa) and somewhat lower natural emissions (including NMVOC-CO). In addition, the biomass burning emissions slightly increase from 311 Tg CO/yr in the base inversion to 335 Tg CO/yr. On both global and continental scales, however, the differences are generally within the error bounds of the base MOPITT inversion.

[51] Despite regional differences, the results of the sensitivity inversions are largely within error bounds of the base inversions on continental and regional scales for the individual emission categories. However, since the system has difficulties separating the different source categories, the inversion finds negative posterior correlations between the categories which lead to much smaller posterior uncertainties for the total CO source compared to the individual categories [Hooghiemstra *et al.*, 2011]. Our reported total CO emissions for the sensitivity studies (Table 5) vary typically more than the calculated posterior error for the base inversion. This posterior error (9 Tg CO/yr for the MOPITT-only base inversion) becomes small because it is the result of a total global source that has to compensate the global total sink

(OH and surface deposition) such that a set of observations is optimally fitted. The sensitivity studies show that the global total emissions are typically estimated within 10%. Hence, the true error is probably larger than the approximation we calculate and depends on the OH sink, the vertical distribution of the emissions and on the set of measurements used in the inversion. Additional model errors may also play a role (e.g., transport). A first attempt to quantify systematic model errors was made by *Jiang et al.* [2011]. They also found that model transport, the OH field and the treatment of the NMVOC-CO source yield differences in inferred emission estimates up to 20%. Nonetheless, we think that the approximation of the posterior uncertainties as presented here are valuable in assessing the information content of the assimilated observations.

## 5. Summary and Conclusions

[52] CO emission estimates have been derived using a 4D-Var framework and utilizing two different observational data sets: surface observations from the NOAA network and total columns from the MOPITT instrument on board NASA's Terra satellite. We have discussed and validated the optimized emission estimates and compared them to values reported in the recent literature. The main conclusions are:

[53] 1. Optimized emissions using either data set show a global increase of CO emissions from the prior to the posterior estimate of 500 Tg CO/yr. Our estimates for total annual CO emissions from the anthropogenic, biomass burning and natural sources (excluding the oxidation of methane) amount to  $1822 \pm 45$  and  $1843 \pm 12$  Tg CO in 2004 for the stations-only and the MOPITT-only inversions, respectively. The regions to which this increment is attributed are East Asia, South America, Europe and Africa. Our results suggest that the EDGARv4.1 bottom up inventories for East Asia and Europe could be too low by up to a factor 2. In South America and Africa, in particular CO production from oxidation of NMVOC-CO is increased when only MOPITT total columns are assimilated, whereas biomass burning emissions are increased when assimilating only NOAA surface observations.

[54] 2. Applying the 4D-Var analysis to CO emission estimates reduces the prior uncertainty for the different source categories significantly. With MOPITT's higher spatiotemporal resolution and better coverage in the Tropics and in the SH compared to the NOAA surface stations, the MOPITT derived emissions show larger uncertainty reductions over the tropical regions. However, due to the high density of NOAA stations on the NH, the high precision of these measurements and the fact that we do not assimilate MOPITT land pixels, uncertainty reduction in the NH mid-latitudes is typically largest for the stations-only inversion.

[55] 3. A detailed comparison of the stations-only and the MOPITT-only inversions shows that in particular the partitioning of the SH sources is different. The difference in optimized biomass burning (or natural) emissions between the stations-only and the MOPITT-only inversions is attributed to the different vertical sensitivity of these observational data sets. Due to the faster vertical mixing in the Tropics, Tropical NOAA stations are quite insensitive to CO

emissions and hardly constrain emissions that are released higher up in the atmosphere. In contrast, MOPITT columns are sensitive to free tropospheric CO and attributes model data mismatches mostly to the natural source. However, also the much higher spatiotemporal resolution of the MOPITT observations, and thus the better global coverage compared to the NOAA observations, plays a role.

[56] 4. We showed that by optimizing the NMVOC CO source on the model resolution, the biomass burning source in the MOPITT-only inversion was no longer increased, as opposed to some recent studies that found that the GFEDv2 inventory underestimated biomass burning emissions. Our approach reduces the risk that all mismatches between the prior model simulation and the observations is projected onto the biomass burning source in the Tropics.

[57] 5. The posterior emission estimates have been validated with aircraft observations from NOAA (observations up to 8 km) and showed large improvement with respect to the prior comparison for both inversions. The comparison with MOZAIC aircraft observations also improved in the lower troposphere. However, higher up in the troposphere (above 8 km), in particular the stations-only posterior simulation diverged from the observations, whereas the MOPITT-only posterior simulation remained close to the observations. Additional comparisons with FTIR total column measurements improved particularly for the NH sites.

[58] 6. Validation in the SH is limited by the amount of independent data available. A cross validation showed that MOPITT-only derived emissions yield too high CO mixing ratios on the SH stations Cape Grim (+20 ppb) and South Pole (+20 ppb). Validation with FTIR total columns at Lauder, New Zealand and Arrival Heights, Antarctica also hint towards a small but significant positive bias in MOPITT. It should be kept in mind however, that model uncertainties such as model transport and the OH climatology used as well as the linearized MOPITT averaging kernels may introduce biases in the optimized emission estimates. However, the inferred emission estimates of sensitivity studies aggregated to continental and global scales are within the error bounds from the base inversions.

[59] With the results presented in this paper it seems an obvious next step to combine surface observations with satellite retrievals to estimate surface sources of CO. In the NH, both data sets seem to be broadly consistent. However, on the high latitude SH and Indonesia, large differences in inferred emission estimates using the two data sets are apparent. Similar to the station - SCIAMACHY inversion performed for methane [*Meirink et al.*, 2008a; *Bergamaschi et al.*, 2009, 2010], a bias correction scheme seems to be necessary to obtain accurate and realistic emission estimates that are in agreement with both data sets and independent observations. Furthermore, extra constraints in the form of formaldehyde columns could serve to better constrain the NMVOC-CO source. Also, higher spatial model resolution should lead to smaller model representativeness errors.

## Appendix A: Derivation of $\tilde{A}$

[60] To properly compare a model profile to a MOPITT retrieved profile, one has to use the averaging kernel as described by *Deeter* [2009]. This averaging kernel is

modeled in terms of a lognormal distribution and hence one typically computes:

$$\ln(\mathbf{x}_i^{ps}) = \ln(\mathbf{x}_i^a) + \sum_{j=1}^n \mathbf{A}_{ij} \left( \ln(\mathbf{x}_j^m) - \ln(\mathbf{x}_j^a) \right), \quad (\text{A1})$$

where  $\mathbf{x}^{ps}$  is the model profile smoothed with the MOPITT averaging kernel  $\mathbf{A}$  and a priori profile  $\mathbf{x}^a$  and  $\mathbf{x}^m$  is the original profile interpolated to the MOPITT pressure grid and  $n$  is the number of pressure levels. Rewriting this to avoid the logarithms yields:

$$\mathbf{x}_i^{ps} = \mathbf{x}_i^a \cdot \prod_{j=1}^n \left( \frac{\mathbf{x}_j^m}{\mathbf{x}_j^a} \right)^{\mathbf{A}_{ij}}. \quad (\text{A2})$$

For a Gaussian distributed averaging kernel ( $\mathbf{K}$ ) equation (A1) would read

$$\mathbf{x}_i^{ps} = \mathbf{x}_i^a + \sum_{j=1}^n \mathbf{K}_{ij} (\mathbf{x}_j^m - \mathbf{x}_j^a), \quad (\text{A3})$$

from which it follows that

$$\mathbf{K}_{ij} = \frac{d\mathbf{x}_i^{ps}}{d\mathbf{x}_j^m} \quad (\text{A4})$$

Now, taking the derivative of equation (A2) to  $\mathbf{x}_j^m$  yields

$$\frac{d\mathbf{x}_i^{ps}}{d\mathbf{x}_j^m} = \frac{\mathbf{A}_{ij}}{\mathbf{x}_j^m} \mathbf{x}_i^{ps}, \quad (\text{A5})$$

hence, if we define

$$\tilde{\mathbf{A}}_{ij} = \frac{\mathbf{A}_{ij}}{\mathbf{x}_j^m} \mathbf{x}_i^{ps} \quad (\text{A6})$$

it is consistent with  $\mathbf{K}_{ij}$  to first order. However, since both  $\mathbf{x}_j^m$  and  $\mathbf{x}_i^{ps}$  depend on the model simulation, they change every iteration when the emissions are perturbed during the iterative optimization. To avoid this and to obtain a formula for the averaging kernel that is constant during the iterative process, we approximate those terms by values given by MOPITT. Naturally,  $\mathbf{x}_j^m$  is approximated by the MOPITT prior  $\mathbf{x}_j^a$  and  $\mathbf{x}_i^{ps}$  is approximated by the MOPITT retrieval  $\mathbf{x}_i^{\text{retr}}$ :

$$\tilde{\mathbf{A}}_{ij} = \frac{\mathbf{A}_{ij}}{\mathbf{x}_j^a} \mathbf{x}_i^{\text{retr}}. \quad (\text{A7})$$

It has been tested that this approximation yields mean model total column CO values that are on average within 2% of the columns using the non-linear formulation. Regionally, larger differences up to 10% are observed.

[61] **Acknowledgments.** This research was supported by the Dutch User Support Programme 2006–2010 under project GO-AO/05. FTIR data used in this publication were obtained as part of the Network for the Detection of Atmospheric Composition Change (NDACC) and are publicly available (see <http://www.ndacc.org>). The Dutch National Computer Facility (NCF) is acknowledged for computer resources. We are most thankful to S. Basu for help during the model development phase. We are thankful to S. Myriokefalitakis for supplying the a priori fields of

NMVO-CO production. We also thank S. Houweling for optimized methane mixing ratio fields. Finally, we thank G. Maenhout for useful discussion during the writing process.

## References

- Arellano, A. F., Jr., P. Kasibhatla, L. Giglio, G. van der Werf, and J. Randerson (2004), Top-down estimates of global CO sources using MOPITT measurements, *Geophys. Res. Lett.*, *31*, L01104, doi:10.1029/2003GL018609.
- Arellano, A. F., Jr., P. Kasibhatla, L. Giglio, G. van der Werf, J. Randerson, and G. Collatz (2006), Time-dependent inversion estimates of global biomass burning CO emissions using Measurement of Pollution in the Troposphere (MOPITT) measurements, *J. Geophys. Res.*, *111*, D09303, doi:10.1029/2005JD006613.
- Bergamaschi, P., R. Hein, M. Heimann, and P. J. Crutzen (2000), Inverse modeling of the global CO cycle: 1. Inversion of CO mixing ratios, *J. Geophys. Res.*, *105*(D2), 1909–1927.
- Bergamaschi, P., M. Krol, F. Dentener, A. Vermeulen, F. Meinhardt, R. Graul, M. Ramonet, W. Peters, and E. J. Dlugokencky (2005), Inverse modelling of national and European CH<sub>4</sub> emissions using the atmospheric zoom model TM5, *Atmos. Chem. Phys.*, *5*, 2431–2460, doi:10.5194/acp-5-2431-2005.
- Bergamaschi, P., et al. (2007), Satellite cartography of atmospheric methane from SCIAMACHY on board ENVISAT: 2. Evaluation based on inverse model simulations, *J. Geophys. Res.*, *112*, D02304, doi:10.1029/2006JD007268.
- Bergamaschi, P., C. Frankenberg, J.-F. Meirink, M. Krol, G. Villani, S. Houweling, F. Dentener, E. J. Dlugokencky, and A. Engel (2009), Inverse modeling of global and regional CH<sub>4</sub> emissions based on recently revised SCIAMACHY retrievals, *J. Geophys. Res.*, *114*, D22301, doi:10.1029/2009JD012287.
- Bergamaschi, P., et al. (2010), Inverse modeling of European CH<sub>4</sub> emissions 2001–2006, *J. Geophys. Res.*, *115*, D22309, doi:10.1029/2010JD014180.
- Chevallier, F. (2007), Impact of correlated observation errors on inverted CO<sub>2</sub> surface fluxes from OCO measurements, *Geophys. Res. Lett.*, *34*, L24804, doi:10.1029/2007GL030463.
- Chevallier, F., A. Fortems, P. Bousquet, I. Pison, S. Szopa, M. Devaux, and D. A. Hauglustaine (2009), African CO emissions between years 2000 and 2006 estimated from MOPITT observations, *Biogeoscience*, *6*, 103–111, doi:10.5194/bg-6-103-2009.
- Deeter, M. N. (2009), MOPITT (Measurements of Pollution in the Troposphere) validated version 4, report, Natl. Cent. for Atmos. Res., Boulder, Colo.
- Deeter, M. N., et al. (2003), Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument, *J. Geophys. Res.*, *108*(D14), 4399, doi:10.1029/2002JD003186.
- Deeter, M. N., et al. (2010), The MOPITT version 4 CO product: Algorithm enhancements, validation, and long-term stability, *J. Geophys. Res.*, *115*, D07306, doi:10.1029/2009JD013005.
- de Laat, A. T. J., A. M. S. Gloudemans, I. Aben, and H. Schrijver (2010), Global evaluation of SCIAMACHY and MOPITT carbon monoxide column differences for 2004–2005, *J. Geophys. Res.*, *115*, D06307, doi:10.1029/2009JD012698.
- Duncan, B. N., J. A. Logan, I. Bey, I. A. Megretskaja, R. M. Yantosca, P. C. Novelli, N. B. Jones, and C. P. Rinsland (2007), Global budget of CO, 1988–1997: Source estimates and validation with a global model, *J. Geophys. Res.*, *112*, D22301, doi:10.1029/2007JD008459.
- Emmons, L. K., et al. (2007), Measurements of Pollution in the Troposphere (MOPITT) validation during the summer 2004 field campaigns over North America, *J. Geophys. Res.*, *112*, D12S02, doi:10.1029/2006JD007833.
- Emmons, L. K., D. P. Edwards, M. N. Deeter, J. C. Gille, T. Campos, P. Nédélec, P. Novelli, and G. Sachse (2009), Measurement of Pollution In The Troposphere (MOPITT) validation through 2006, *Atmos. Chem. Phys.*, *9*, 1795–1803, doi:10.5194/acp-9-1795-2009.
- Fisher, M., and P. Courtier (1995), Estimating the covariance matrices of analysis and forecast error in variational data assimilation, technical report, ECMWF, Reading, U. K.
- Fortems-Cheiney, A., F. Chevallier, I. Pison, P. Bousquet, S. Szopa, M. N. Deeter, and C. Clerbaux (2011), Ten years of CO emissions as seen from Measurements of Pollution in the Troposphere MOPITT, *J. Geophys. Res.*, *116*, D05304, doi:10.1029/2010JD014416.
- Gonzi, S., and P. I. Palmer (2010), Vertical transport of surface fire emissions observed from space, *J. Geophys. Res.*, *115*, D02306, doi:10.1029/2009JD012053.
- Heald, C. L., et al. (2004), Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide, *J. Geophys. Res.*, *109*, D23306, doi:10.1029/2004JD005185.

- Hestenes, M. R., and E. Stiefel (1952), Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Stand.*, *49*, 409–436.
- Hooghiemstra, P. B., M. C. Krol, J. F. Meirink, P. Bergamaschi, G. R. van der Werf, P. C. Novelli, I. Aben, and T. Röckmann (2011), Optimizing global CO emission estimates using a four-dimensional variational data assimilation system and surface network observations, *Atmos. Chem. Phys.*, *11*(10), 4705–4723, doi:10.5194/acp-11-4705-2011.
- Houweling, S., F. Dentener, and J. Lelieveld (1998), The impact of non-methane hydrocarbon compounds on tropospheric photochemistry, *J. Geophys. Res.*, *103*(D9), 10,673–10,696.
- Huijnen, V., et al. (2010), The global chemistry transport model TM5: Description and evaluation of the tropospheric chemistry version 3.0, *Geosci. Model Dev.*, *3*(3), 445–473, doi:10.5194/gmd-3-445-2010.
- Jiang, Z., D. B. A. Jones, M. Kopacz, J. Liu, D. K. Henze, and C. Heald (2011), Quantifying the impact of model errors on top-down estimates of carbon monoxide emissions using satellite observations, *J. Geophys. Res.*, *116*, D15306, doi:10.1029/2010JD015282.
- Jones, D. B. A., K. W. Bowman, J. A. Logan, C. L. Heald, J. Liu, M. Luo, J. Worden, and J. Drummond (2009), The zonal structure of tropical O<sub>3</sub> and CO as observed by the Tropospheric Emission Spectrometer in November 2004. Part 1: Inverse modeling of CO emissions, *Atmos. Chem. Phys.*, *9*(11), 3547–3562, doi:10.5194/acp-9-3547-2009.
- Kasibhatla, P., A. Arellano, J. A. Logan, P. I. Palmer, and P. Novelli (2002), Top-down estimates of a large source of atmospheric carbon monoxide associated with fuel combustion in Asia, *Geophys. Res. Lett.*, *29*(19), 1900, doi:10.1029/2002GL015581.
- Kopacz, M., D. J. Jacob, D. K. Henze, C. L. Heald, D. G. Streets, and Q. Zhang (2009), Comparison of adjoint and analytical Bayesian inversion methods for constraining Asian sources of carbon monoxide using satellite (MOPITT) measurements of CO columns, *J. Geophys. Res.*, *114*, D04305, doi:10.1029/2007JD009264.
- Kopacz, M., et al. (2010), Global estimates of CO sources with high resolution by adjoint inversion of multiple datasets (MOPITT, AIRS, SCIAMACHY, TES), *Atmos. Chem. Phys.*, *10*, 855–876, doi:10.5194/acp-10-855-2010.
- Krol, M. C., S. Houweling, B. Bregman, M. van den Broek, A. Segers, P. van Velthoven, W. Peters, F. Dentener, and P. Bergamaschi (2005), The two-way nested global chemistry-transport zoom model TM5: Algorithm and applications, *Atmos. Chem. Phys.*, *5*, 417–432, doi:10.5194/acp-5-417-2005.
- Lanczos, C. (1950), An iteration method for the solution of eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Stand.*, *45*(4), 255–282.
- Liu, J., J. A. Logan, D. B. A. Jones, N. J. Livesey, I. Megretskaja, C. Carouge, and P. Nedelec (2010), Analysis of CO in the tropical troposphere using Aura satellite data and the GEOS-Chem model: Insights into transport characteristics of the GEOS meteorological products, *Atmos. Chem. Phys.*, *10*(24), 12,207–12,232, doi:10.5194/acp-10-12207-2010.
- Logan, J. A., M. J. Prather, S. C. Wofsy, and M. B. McElroy (1981), Tropospheric chemistry: A global perspective, *J. Geophys. Res.*, *86*(C8), 7210–7254.
- Meirink, J.-F., et al. (2008a), Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: Analysis of SCIAMACHY observations, *J. Geophys. Res.*, *113*, D17301, doi:10.1029/2007JD009740.
- Meirink, J.-F., P. Bergamaschi, and M. C. Krol (2008b), Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: Method and comparison with synthesis inversion, *Atmos. Chem. Phys.*, *8*, 6341–6353, doi:10.5194/acp-8-6341-2008.
- Mukherjee, C., P. S. Kasibhatla, and M. West (2011), Bayesian statistical modeling of spatially correlated error structure in atmospheric tracer inverse analysis, *Atmos. Chem. Phys.*, *11*(11), 5365–5382, doi:10.5194/acp-11-5365-2011.
- Müller, J.-F., and T. Stavrou (2005), Inversion of CO and NO<sub>x</sub> emissions using the adjoint of the IMAGES model, *Atmos. Chem. Phys.*, *5*, 1157–1186, doi:10.5194/acp-5-1157-2005.
- Myriokefalitakis, S., M. Vrekoussis, K. Tsigaridis, F. Wittrock, A. Richter, C. Brühl, R. Volkamer, J. P. Burrows, and M. Kanakidou (2008), The influence of natural and anthropogenic secondary sources on the glyoxal global distribution, *Atmos. Chem. Phys.*, *8*(16), 4965–4981, doi:10.5194/acp-8-4965-2008.
- Nedelec, P., et al. (2003), An improved infrared carbon monoxide analyser for routine measurements aboard commercial Airbus aircraft: Technical validation and first scientific results of the MOZAC III programme, *Atmos. Chem. Phys.*, *3*(5), 1551–1564, doi:10.5194/acp-3-1551-2003.
- Novelli, P. C., K. A. Masarie, P. M. Lang, B. D. Hall, R. C. Myers, and J. W. Elkins (2003), Reanalysis of tropospheric CO trends: Effects of the 1997–1998 wildfires, *J. Geophys. Res.*, *108*(D15), 4464, doi:10.1029/2002JD003031.
- Palmer, P. I., D. J. Jacob, D. B. A. Jones, C. L. Heald, R. M. Yantosca, and J. A. Logan (2003), Inverting for emissions of carbon monoxide from Asia using aircraft observations over the western Pacific, *J. Geophys. Res.*, *108*(D21), 8828, doi:10.1029/2003JD003397.
- Peters, W., et al. (2010), Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, *Global Change Biol.*, *16*, 1317–1337, doi:10.1111/j.1365-2486.2009.02078.x.
- Pétron, G., C. Granier, B. Khattatov, J.-F. Lamarque, V. Yudin, J.-F. Müller, and J. Gille (2002), Inverse modeling of carbon monoxide surface emissions using Climate Monitoring and Diagnostics Laboratory network observations, *J. Geophys. Res.*, *107*(D24), 4761, doi:10.1029/2001JD001305.
- Pétron, G., C. Granier, B. Khattatov, V. Yudin, J.-F. Lamarque, L. Emmons, J. Gille, and D. P. Edwards (2004), Monthly CO surface sources inventory based on the 2000–2001 MOPITT satellite data, *Geophys. Res. Lett.*, *31*, L21107, doi:10.1029/2004GL020560.
- Pison, I., P. Bousquet, F. Chevallier, S. Szopa, and D. Hauglustaine (2009), Multi-species inversion of CH<sub>4</sub>, CO and H<sub>2</sub> emissions from surface measurements, *Atmos. Chem. Phys.*, *9*, 5281–5297, doi:10.5194/acp-9-5281-2009.
- Spivakovsky, C. M., et al. (2000), Three-dimensional climatological distribution of tropospheric OH: Update and evaluation, *J. Geophys. Res.*, *105*, 8931–8980.
- Stavrou, T., and J.-F. Müller (2006), Grid-based versus big region approach for inverting CO emissions using Measurement of Pollution in the Troposphere (MOPITT) data, *J. Geophys. Res.*, *111*, D15304, doi:10.1029/2005JD006896.
- Stavrou, T., J.-F. Müller, I. De Smedt, M. Van Roozendael, G. R. van der Werf, L. Giglio, and A. Guenther (2009), Global emissions of non-methane hydrocarbons deduced from sciamachy formaldehyde columns through 2003–2006, *Atmos. Chem. Phys.*, *9*(11), 3663–3679, doi:10.5194/acp-9-3663-2009.
- Val Martin, M., J. A. Logan, D. Kahn, F. Y. Leung, D. Nelson, and D. Diner (2009), Smoke injection heights from fires in North America: Analysis of 5 years of satellite observations, *Atmos. Chem. Phys. Discuss.*, *9*(5), 20,515–20,566, doi:10.5194/acp-10-1491-2010.
- van der Werf, G. R., et al. (2010), Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), *Atmos. Chem. Phys.*, *10*(23), 11,707–11,735, doi:10.5194/acp-10-11707-2010.
- Yurganov, L., W. McMillan, E. Grechko, and A. Dzholia (2010), Analysis of global and regional CO burdens measured from space between 2000 and 2009 and validated by ground-based solar tracking spectrometers, *Atmos. Chem. Phys.*, *10*(8), 3479–3494, doi:10.5194/acp-10-3479-2010.

I. Aben, SRON Netherlands Institute for Space Research, Sorbonnelaan 2, NL-3584 CA Utrecht, Netherlands.

P. Bergamaschi, Institute for Environment and Sustainability, Joint Research Centre, European Commission, Via E. Fermi 2749, I-21027 Ispra (VA), Italy.

M. N. Deeter, Atmospheric Chemistry Division, National Center for Atmospheric Research, PO Box 3000, Boulder, CO 80307, USA.

A. T. J. de Laat, Royal Netherlands Meteorological Institute, Wilhelminalaan 10, NL-3732 GK de Bilt, Netherlands.

P. B. Hooghiemstra, M. C. Krol, and T. Röckmann, Institute for Marine and Atmospheric Research Utrecht, Princetonplein 5, NL-3584 CC Utrecht, Netherlands. (p.b.hooghiemstra@uu.nl)

P. C. Novelli, Climate Monitoring and Diagnostics Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80303, USA.

G. R. van der Werf, Faculty of Earth and Life Sciences, Free University, De Boelelaan 1085, NL-1081 HV Amsterdam, Netherlands.