Manuscript prepared for Atmos. Meas. Tech.
with version 3.0 of the LaTeX class copernicus.cls.
Date: 5 December 2013

Tropospheric CO vertical profiles deduced from total columns using data assimilation: methodology and validation

L. El Amraoui[1], J.-L. Attié[1,2], P. Ricaud[1], W. A. Lahoz[1,4], A. Piacentini[5], V.-H. Peuch[6], J. X. Warner[7], R. Abida[1], J. Barré[8], and R. Zbinden[1]

[1]CNRM-GAME, Météo-France and CNRS, UMR3589, Toulouse, France
[2]Laboratoire d'Aérologie, Université de Toulouse, CNRS/INSU, Toulouse, France
[3]EUMETSAT, Darmstadt, Germany
[4]NILU, 2027 Kjeller, Norway
[5]CERFACS, Toulouse, France
[6]ECMWF, Reading, UK
[7]Department of Atmospheric & Oceanic Science, University of Maryland, Baltimore County, USA
[8]NCAR, Boulder, USA

Correspondence to: L. El Amraoui (laaziz.elamraoui@meteo.fr)

Abstract. **This paper presents a validation of a method to derive the vertical profile of carbon monoxide (CO) from its total column using data assimilation. We choose version 3 of MOPITT CO total columns to validate the proposed method. MOPITT products have the advantage of providing both the vertical profiles and the total columns of CO. Furthermore, this version has been extensively validated by comparison with many independent datasets, and has been used in many scientific studies.**

The first step of the paper consists in the specification of the observation errors based on the Chi-square (χ^2) test. The observations have been binned according to 3 types: over land during daytime, over land during nighttime, and over sea. Their respective errors using the χ^2 metric have been found to be 8%, 11% and 7%.

In a second step, the CO total columns, with their specified errors, are used within the assimilation system to estimate the vertical profiles. These are compared to the retrieved profiles of MOPITT V3 at global and regional scales. Generally, the two datasets show similar patterns and good agreement at both scales. Nevertheless, total column analyses slightly overestimate CO concentrations compared to MOPITT observations. The mean bias between both datasets is 15% and 12% at 700 hPa and 250 hPa, respectively.

In a third step, the assimilation of total column has been compared to the assimilation of MOPITT vertical profiles. Both analyses show very good agreement. In terms of lon-lat maps, the mean bias between the two datasets is 6 % and 8 % at the pressure levels 700 and 200 hPa, respectively. In terms of zonal means, the CO distribution is similar for both analyses with a

mean bias which does not exceed 12 %.

Finally, the two analyses have been validated using the aircraft MOZAIC independent observations in terms of vertical profiles over 8 airports. Over most airports, both analyses agree well with aircraft profiles. For more than 50% of recorded measurements, the difference between the analyses and MOZAIC does not exceed 5 ppbv (parts per billion volume).

1 Introduction

Carbon monoxide (CO) is an important atmospheric species as it influences tropospheric chemistry and climate (Crutzen and Andreae, 1990). The main sources of CO emissions are biomass burning, fossil fuel and the oxidation of methane and non-methane hydrocarbons (Granier et al., 2000). For this reason and because of larger anthropogenic emissions in the Northern Hemisphere (NH) than in the Southern Hemisphere (SH), tropospheric CO background values are much higher in the NH than in the SH. The major global sink of CO in the troposphere is the chemical reaction with the hydroxyl radical (OH). Therefore, CO concentrations are higher in winter than summer owing to the seasonal variations of OH abundances. Since OH is the only significant tropospheric sink for CO and many other atmospheric trace gases emitted into the troposphere, CO has the potential to indirectly control the oxidation capacity of the troposphere. Therefore, an increase in CO emissions could reduce OH concentrations and, consequently, the oxidation capacity of the troposphere and its ability to remove pollutants (Mahieu et al., 1997).

Most of the CO in the troposphere is found in the lower troposphere or boundary layer. Compared to its inter-hemispheric mixing time of several years, CO is not well mixed in the free troposphere where it has a relatively long lifetime of several weeks to a few months. This makes CO a useful tracer of air pollution, and allows studies of long-range transport of pollutants in the troposphere.

For more than 10 yr, global observations of tropospheric CO have been performed from several satellite instruments which provide many opportunities to study tropospheric CO on a global scale. The Infrared Atmospheric Sounding Interferometer (IASI) onboard the MetOP-A (Meteorological Operational Program) satellite launched in October 2006 provides augmented horizontal resolution of the CO total column. Monitoring of this species will continue with the METOP-B satellite which carries a suite of sophisticated instruments. These two satellites are polar orbiters and provide global observations. The data they collect on the atmosphere and the environment are complementary and allow the monitoring of the atmospheric composition and its evolution in near real-time.

Most tropospheric sensors operate with a nadir-viewing geometry and typically provide vertically integrated information, implying limited vertical resolution. This could present a limitation for some process studies such as long-range transport of pollutants because of missing information on vertical levels. Furthermore, most chemistry and transport models (CTMs) are subject to large uncertainties concerning the distribution of CO concentrations. This is because CO sources are not well known since their estimates are generally derived from inventory-based, bottom-up techniques which are highly uncertain (e.g., Jones et al., 2003). Another issue concerns the CO emissions from biomass burning which have unexpected sources in terms of time, location and magnitude and thus are subject to large uncertainties (Bian et al., 2007).

Chemical data assimilation consists of combining in an optimal way observations provided by instruments with a priori knowledge about a physical system such as model output. It allows constraints to be put on models using observations, and thus can be used to overcome model deficiencies. It also provides a four-dimensional (time and space) description of the dynamical and chemical state of the atmosphere. Typically, data assimilation systems produce observation minus forecast (OMF) statistics that are used for monitoring biases between the observations and the models (e.g., El Amraoui et al., 2010). The specific objective of chemical data assimilation is to produce a self-consistent picture of the atmosphere taking into account both the available observations and our theoretical understanding of the atmospheric system.

Assimilation of CO satellite observations in the troposphere has been performed using different sensors. These include MAPS (Lamarque et al., 1999), IMG (Clerbaux et al., 2001), MOPITT (e.g. Pradier et al., 2006; Claeyman et al., 2010; El Amraoui et al., 2010) and SCIAMACHY (e.g. Tangborn et al., 2009). Most of the CO analyses in these studies have revealed improvements of the CO distribution in comparison to the free model run. However, no assessment of the impact of the assimilation of the total column on the CO vertical profile has been done hitherto.

The main goal of this study is to assess the benefit of the CO total column assimilation on the CO vertical distribution at global and regional scales. **The philosophy of this study is the following : The CO total column is generally produced from the profiles using a simple integration over the vertical levels. The question we pose is : can we use an assimilation system to derive the CO vertical profile from its total column using the adjoint of the integration operator ?**

We choose Version 3 of MOPITT CO measurements to validate the proposed method. The motivation for this choice is presented in Sect. 2.1. The proposed method has the advantage of allowing fast computation of the vertical profiles and the analyses of CO. It will be particularly useful in the future when there will be many missions providing large volumes of data for which level 2 retrievals with their corresponding characteristics (covariance matrices and averaging kernels) will be very expensive in terms of computer resources (i.e., IASI onboard METOP-A and METOP-B or future geostationary missions). Furthermore, the assimilation of such data in CTMs taking into account all these characteristics will likely be very costly in terms of time computation and memory. This will be a significant shortcoming regarding the operational use of these data. Thus, the validation of the method proposed in this paper could be an alternative way to produce CO fields at global scale with relatively modest resources.

First, we describe the approach which consists of deducing the vertical distribution of CO in the troposphere from the assimilation of total MOPITT column measurements (hereinafter

noted TOTCOL_ANALYSES). Second, we validate the CO vertical profiles deduced from TOTCOL_ANALYSES with the MOPITT retrieved vertical profiles. In a third step, we compare TOTCOL_ANALYSES against the assimilation of MOPITT CO vertical profiles taking into account the corresponding error covariance matrices and averaging kernels (hereinafter noted PROFILE_ANALYSES). Finally, both analyses, from total column and from profiles, have been validated using independent in-situ MOZAIC observations.

The paper outline is as follows: Sect. 2 presents the MOPITT CO measurements as well as the corresponding total columns, the data assimilation system used in this study, and the data used for the evaluation of the vertical profiles deduced from the assimilation of CO MOPITT total column: the official vertical profiles of MOPITT V3 measurements. The method used for the assimilation of MOPITT CO total columns, the specification of the errors as well as the a posteriori diagnostics are presented in Sect. 3. The comparisons of CO vertical profiles deduced from TOTCOL_ANALYSES to those of MOPITT observations are presented in Sect. 4. Section 5 presents a validation of the CO vertical profiles calculated from TOTCOL_ANALYSES against PROFILE_ANALYSES. Conclusions are presented in Sect. 6.

2 Data and analysis

2.1 Terra/MOPITT carbon monoxide observations

The MOPITT instrument (Drummond and Mand, 1996) is onboard the Terra platform and has been monitoring global tropospheric CO from March 2000 to date. The pixel size is $22\,\mathrm{km} \times 22\,\mathrm{km}$ and the vertical profiles for MOPITT version 3 are retrieved on 7 pressure levels (surface, 850, 700, 500, 350, 250 and $150\,\mathrm{hPa}$) with the maximum likelihood method (Rodgers, 2000). The retrieved profiles are characterized by their error covariance matrices and their averaging kernels, providing information on the vertical sensitivity of the measurements. In particular, the Degrees of Freedom for Signal (DFS), the trace of the averaging kernel matrix, indicates the number of independent pieces of information in the measurements. It depends, via the surface temperature, on latitude and time of day. The MOPITT V3 CO level 2 product consists of retrieved values and estimated uncertainties of the CO total column and CO profile (see: http://www.acd.ucar.edu/mopitt/retrievals.shtml). The retrieved CO total column is obtained as a byproduct of the retrieved profile by integrating the retrieved profile from the surface to the top of the atmosphere (see: www.acd.ucar.edu/mopitt/avg_krnls_app.pdf).

The main motivation for using the MOPITT V3 is because these data have been extensively validated against many independent datasets (e.g., Emmons et al., 2004; Emmons et al., 2007; Emmons et al., 2009; Deeter et al., 2007; Yurganov et al., 2008). The temporal and spatial behavior of MOPITT V3 data is well understood (e.g., Emmons et al., 2009).

For MOPITT V3 the DFS is low for vertical profiles as well as for the total columns. Figure 1

shows an example of the spatial variation of the DFS of MOPITT V3 profiles in terms of longitude-latitude averaged over the month of August as well as its frequency distribution over the same period. The typical value of DFS for MOPITT V3 profiles is around 1.5 and is located primarily on the sea above the tropics.

We will demonstrate in this paper that, for this kind of data, the presented method consisting of deducing the profiles from the total columns, remains valid when only using the adjoint of the integration operator. For other types of data for which the DFS is greater than that of MOPITT V3, the present method has to be tested and evaluated against independent observations. This will be the subject of an ongoing work for which we will evaluate the validity conditions of the adjoint operator with respect to different values of the DFS.

2.2 MOCAGE CTM and data assimilation system

The assimilation system used in this study is MOCAGE-VALENTINA (e.g., Emili et al., 2013 : accepted in Atmos. Chem. Phys.) which is an extension of the MOCAGE-PALM (e.g., El Amraoui et al., 2008a,b) system initially developed in the framework of the ASSET (ASSimilation of Envisat daTa) European project. It is developed jointly by Météo-France and CERFACS (Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique). MOCAGE (MOdèle de Chimie Atmosphérique à Grande Echelle) (Peuch et al., 1999) is a 3-D-CTM which covers the planetary boundary layer, the free troposphere, and the stratosphere. It provides a number of optional configurations with varying domain geometries and resolutions, as well as chemical and physical parameterization packages. It has the flexibility to use several chemical schemes for stratospheric and tropospheric studies. In this study, MOCAGE is forced dynamically by wind and temperature fields from the ARPEGE model analyses, the global operational weather prediction model of Météo-France (Courtier et al., 1991). The MOCAGE horizontal resolution used for this study is 2° both in latitude and longitude and the model uses a semi-Lagrangian transport scheme. It includes 47 hybrid (σ, P) levels from the surface up to 5 hPa, where $\sigma = P/P_\mathrm{s}$; P and P_s are the pressure and the surface pressure, respectively. MOCAGE has a vertical resolution of about 800 m in the vicinity of the tropopause and in the lower stratosphere, whereas in the boundary layer MOCAGE has 7 levels with a vertical resolution between 40 and 400 m. In the free troposphere, MOCAGE has a vertical resolution which varies from 400 to 800 m.

The technique implemented within VALENTINA and used for the assimilation of MOPITT CO observations, is the 3-D-FGAT (First Guess at Appropriate Time) method. This method is a compromise between the 3-D-Var and 4-D-Var techniques (Fisher and Andersson, 2001). It compares the observation and background fields at the correct time and assumes that the increment to be added to the background state is constant over the entire assimilation window. The choice of this assimilation technique limits the size of the assimilation window, since it has to be short enough compared to chemistry and transport timescales. This technique has already produced good-quality results compared to independent data especially for O_3 and CO (e.g., Semane et al., 2007; El Amraoui et al., 2010; Claeyman et al., 2011; Rabier et al., 2010; Bencherif et al., 2011; Lahoz et al., 2012).

3 Assimilation of MOPITT CO total column: methodology and error specification

3.1 Assimilation methodology

For variational systems, the assimilation method is based on the minimization of the cost function, J. They exist in a variety of formulations. We use the notation of Ide et al. (1997):

$$J(\mathbf{x}) = \frac{1}{2}\left[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)\right]^{\mathrm{T}} \mathbf{B}^{-1} \left[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)\right] + \frac{1}{2}\sum_{i=0}^{N} \left[\mathbf{y}^o(t_i) - H_i\left(\mathbf{x}(t_i)\right)\right]^{\mathrm{T}} \mathbf{R}_i^{-1} \left[\mathbf{y}^o(t_i) - H_i(\mathbf{x}(t_i))\right] \tag{1}$$

The first term on the right-hand side of Eq. (1) is the misfit to the background state and the second term represents the misfit to the observations. $\mathrm{x}^{\mathrm{b}}(t_0)$ and $\mathbf{y}(t_i)$ are the background state at the initial time and the observation at time t_i, respectively. B and R are the background and the observation error covariance matrices, respectively. $\mathrm{x}(t_i)$ is the model state at the observation time t_i and represents the propagation of the initial state $\mathrm{x}(t_0)$ by the model operator M:

$$\mathbf{x}(t_i) = M_i\mathbf{x}(t_0) \tag{2}$$

H_i is the observation operator, generally non-linear, which maps the model state $\mathrm{x}(t_i)$ to the measurement space where $\mathrm{y}^o(t_i)$ is located. The subscript i refers to time and N is the number of time steps in the assimilation window $[t_0, t_N]$.

For the incremental variational 3-D-FGAT method, the cost function, J in Eq. (1) can be expressed as:

$$J\left[\delta\mathbf{x}(t_0)\right] = \frac{1}{2}\delta\mathbf{x}(t_0)^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x}(t_0) + \frac{1}{2}\sum_{i=0}^{N}\left[\mathbf{d}(t_i) - \mathbf{H}_i(\delta\mathbf{x}(t_i))\right]^{\mathrm{T}} \mathbf{R}_i^{-1}\left[\mathbf{d}(t_i) - \mathbf{H}_i(\delta\mathbf{x}(t_i))\right] \tag{3}$$

$\left(\delta\mathbf{x}(t_0) = \mathbf{x}(t_0) - \mathbf{x}^{\mathrm{b}}(t_0)\right)$ is the increment vector which represents the difference between the model state x and the background state x^{b} at time t_0, and $\delta\mathbf{x}(t_i) = \mathbf{M}(t_i, t_0)\delta\mathbf{x}(t_0)$, where $\mathbf{M}$ is the tangent linear of M.

The first term on the right-hand side of eq.3 is the background cost function (J_{b}), and the second term represents the observation cost function (J_{o}). $\mathbf{d}(t_i) = \mathbf{y}^o(t_i) - \mathbf{H}_i\left[\mathrm{x}^{\mathrm{b}}(t_i)\right]$ is the departure, at the observation time t_i, between the observation vector $\mathbf{y}^o(t_i)$ and its model equivalent in the observation space $\mathbf{H}_i\left[\mathrm{x}^{\mathrm{b}}(t_i)\right]$. The $\mathbf{H}$ operator represents the tangent-linear of the H operator.

For the assimilation of MOPITT total columns, the observation vector y^{o} contains the 2D field of CO total columns, while the model state x and, consequently, the background state x^{b} is the 3D field of CO vertical profiles updated by the model during the forecast step. The observation operator H which maps the model state to the observation space is then a vertical integration over all model levels taking into account the vertical profile of both the pressure and the density of air.

Note that, in this study, although we assimilate the CO total column the control variable is the 3D CO field. The assimilation process seeks for the optimal 3D increment $\delta\mathrm{x}$ of the CO vertical profiles and thus the observation component of the cost function acts just as one constraint. Another constraint, regularizing the solution by keeping it in the proximity of the background information, is the background cost function (J_{b})

in which we use the background error covariance matrix B. The assimilation increment is therefore a 3D field and its vertical structure depends on the H operator through its adjoint H^{T}, mapping back a variation in the 2D total column space, toward the model 3D space, and the vertical correlation coefficients in B.

More explicitly, the variational 3D-FGAT method consists of minimizing the cost function of equation 3. Since the observation operator is linear, the analysis state can be expressed as:

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{K}\left(\mathbf{y} - \mathbf{H}\mathbf{x}^{\mathrm{b}}\right) \quad \text{where} \quad \mathbf{K} = \mathbf{B}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1} \tag{4}$$

The update of x^{a} after the minimization of the cost function is done by using:

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \delta\mathbf{x}^{\mathrm{a}} \ ; \ \delta\mathbf{x}^{\mathrm{a}} = \mathbf{B}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1} \cdot \mathbf{d} \tag{5}$$

d is the innovation vector. The correction $\delta\mathrm{x}^{\mathrm{a}}$ to be added to x^{b} to obtain x^{a} is normalized by $\mathrm{B}\mathrm{H}^{\mathrm{T}}\left(\mathrm{H}\mathrm{B}\mathrm{H}^{\mathrm{T}} + \mathrm{R}\right)^{-1}$, after it is introduced into the model space (here the CO vertical profile) via H^{T} and is finally multiplied by B.

3.2 Background error covariance matrix

The background error covariance matrix is a key component in data assimilation. It contributes first to filter and propagate spatially the observed information, and second to define the correlations between the control variables of the models during the assimilation process. For the MOCAGE-Valentina assimilation system, the background covariance matrix B is split into a diagonal matrix of the forecast error variances of the assimilated species in each grid point of the model Σ and a positive definite symmetric correlation matrix C :

$$\mathbf{B} = \Sigma\mathbf{C}\Sigma^{\mathrm{T}} \tag{6}$$

The correlation matrix C contains both horizontal and vertical operators. The horizontal correlation is modeled using a two dimensional diffusion equation (Weaver and Courtier, 2001) with a homogenous length-scale both in latitude and longitude. The vertical correlation is

modeled using a Gaussian function in terms of the logarithm of the pressure. Thus the vertical correlation $C^v_{i,j}$ between two pressure levels p_i and p_j is as follows:

$$C^v_{i,j} = \exp\left[-k \cdot \log^2\left(\frac{p_i}{p_j}\right)\right] \tag{7}$$

The dimensionless parameter k is determined from many validation experiments of the MOPITT V3 vertical profiles assimilation in comparison to other independent data such as AIRS and MOZAIC (e.g., El Amraoui et al., 2010). It was found that $k = 100$ gives better analyses compared to the independent data and consequently better characterizes the vertical correlation of the B matrix in the troposphere.

The horizontal correlation $C^h_{k,l}$ between two points k and l separated by a distance $\delta_{k,l}$ is:

$$C^h_{k,l} = \exp\left[\frac{-\delta^2_{k,l}}{2\left(L_x^2 + L_y^2\right)}\right] \tag{8}$$

L_x and L_y are the longitude and latitude length scales in km, respectively.

$$L_x = 2\mathrm{R_e} \cdot \sin\left(\frac{\alpha_x \pi}{360}\right) \ \text{ and } \ L_y = 2\mathrm{R_e} \cdot \sin\left(\frac{\alpha_y \pi}{360}\right) \tag{9}$$

$\mathrm{R_e}$ is the Earth radius (6371.22 km), α_x and α_y are the longitude and latitude length scales, respectively in $^\circ$. In this study, both α_x and α_y are constant and are fixed to 2° which corresponds to a length scale of about 220 km.

3.3 Error specification

The first step of the proposed method consists in specifying the observation error covariance matrices. The assimilation process needs, at least, specification of the error covariance matrices ($\mathbf{R}$ and $\mathbf{B}$ matrices in Eqs. 1 and 2).

To validate the method, we assume in this study that the CO total column from MOPITT has neither error covariance matrix nor averaging kernel information. We specify the corresponding errors of the CO total columns based on the χ^2 test (e.g., El Amraoui et al., 2010): observation errors of the MOPITT CO total columns are estimated using this test. Different values of the observation error have been selected and several assimilation tests with these values have been conducted over a one-month study, August 2008. The appropriate value of the observation error is that for which the χ^2 test is the closest to 1. A value of χ^2 close to 1 indicates consistency between both error-covariance matrices ($\mathbf{R}$ and $\mathbf{B}$), whereas a value of χ^2 lower (greater) than 1 implies an overestimation (underestimation) of the observation and/or background errors (e.g., Lahoz et al., 2007).

Since the sensitivity of MOPITT measurements in the thermal infrared (TIR) wavelength depends, via the surface temperature and thermal contrast, on daytime and nighttime periods, specification of the measurement error is made by binning the observations according to day, night, land and sea. The specification of the errors will be done for three types of measurements: over land during daytime (LAND_DAY), over land during nighttime (LAND_NIGHT) and over sea during daytime and nighttime (SEA). For each type of measurements we assume that all observations have the same percentage error and that errors are uncorrelated.

Figure 2 shows an example of the time evolution of the χ^2 test over the period of study, August 2008 (left hand side) and the corresponding Gaussian fit of the normalized χ^2 test (right hand side) for different observation error values (diagonal of $\mathbf{R}$) corresponding to the measurement type SEA. We note that the χ^2 test is very sensitive to the observation error value. For low values of $\mathbf{R}$, the χ^2 test gives high values and vice-versa. The optimal observation error value (diagonal of $\mathbf{R}$) for which χ^2 is the closest to 1 for SEA measurements is $\sim 7\,\%$.

Table 1 summarizes the χ^2 results for all type of measurements. The optimal values of the observation error (diagonal of $\mathbf{R}$) are indicated in boldface. They are $8\,\%$, $11\,\%$ and $7\,\%$ for LAND_DAY, LAND_NIGHT and SEA measurements, respectively. These values will be used as the observation error values for each corresponding type of measurements in the assimilation of MOPITT CO total column measurements.

3.4 A posteriori diagnostics

Each of the three types of MOPITT measurements (LAND_DAY, LAND_NIGHT and SEA) has been assimilated, in terms of total column, using the corresponding observation error selected according to the χ^2 test discussed in Sect. 3.3. Figure 3 shows the OMF and the OMA (observation minus analysis) diagnostics for TOTCOL_ANALYSES for the whole assimilation period (August 2008). Figure 3-left shows the OMF distributions normalized by the observation errors for the three types of measurements. The OMF histograms are fitted by a Gaussian function. The comparison between the OMF histograms for all types of measurements and the corresponding fitted Gaussian function is very good. This good agreement supports the assumption that the specified observations and their corresponding forecasts have Gaussian errors. We note that the mean of all the normalized OMF values is positive but close to zero (lying between 0.4 and 0.7), which suggests that the bias between the model and the observations is very small for all the three types of measurements.

Figure 3-right shows the OMA and OMF histograms for all MOPITT CO total columns during the whole assimilation period. For the three types of measurements, the OMA histogram is narrower than that for OMF and the bias is reduced. Furthermore the standard deviation of OMA is smaller than that of OMF: $\sigma_{OMA} = 4.9$, 4.8 and 4.0 DU (Dobson Unit) for LAND_DAY, LAND_NIGHT and SEA measurements, respectively. The corresponding values for σ_{OMF} are: 14.1, 13.7 and 7.1 DU, respectively. This indicates that, as expected, the analyses for the different types of measurements are closer to the observations than the forecasts.

4 Comparison of CO derived from total column assimilation to MOPITT V3 observations

4.1 Comparison in terms of horizontal maps

In this section we validate the vertical profiles calculated from TOTCOL_ANALYSES in comparison to the MOPITT CO observations in terms of vertical profiles at global and regional scales. Figure 4 presents a comparison between both datasets in terms of longitude-latitude maps at 700 and 250 hPa for the three types of observations. Since the sensitivity of MOPITT measurements through the averaging kernels is not vertically uniform, TOTCOL_ANALYSES in terms of vertical profiles have been smoothed by the MOPITT averaging kernels to take into account the vertical resolution as well as the a priori information used in the retrieval process of MOPITT vertical profiles. This is performed through the transformation of the vertical profile issued from TOTCOL_ANALYSES ($\mathbf{x}_{\mathrm{assim}}$) using the averaging kernels of the MOPITT CO vertical profiles ($\mathbf{A}$) and the a priori CO profile ($\mathbf{x}_{\mathrm{apriori}}$) to create an analysed vertical profile ($\mathbf{x}_{\mathrm{comp}}$) appropriate for a quantitative comparison to the MOPITT CO retrievals:

$$\mathbf{x}_{\mathrm{comp}} = \mathbf{A}\mathbf{x}_{\mathrm{assim}} + (\mathbf{I} - \mathbf{A})\,\mathbf{x}_{\mathrm{apriori}} \tag{10}$$

Note that the two quantities: MOPITT observations and $\mathbf{x}_{\mathrm{comp}}$ have been averaged in boxes of $2^\circ \times 2^\circ$ (corresponding to the grid mesh of the MOCAGE model) over the month of comparison, August 2008. Figure 4 shows that the general features of both datasets are consistent over the globe at 700 and 250 hPa and that the CO concentrations in the two fields have the same patterns particularly over the emission regions over central Africa, South-Eastern Asia and southern America. Generally, the fields of TOTCOL_ANALYSES slightly overestimate CO concentrations, especially at 700 hPa. The maximum differences between both datasets for this type of measurements range from -10 % to 40 % for the 700 hPa pressure level with a mean bias of 15 %. Whereas, at the pressure level of 250 hPa, the differences range from -10 % to 15 % with a mean bias of 12 %. The mean differences between both datasets are slightly higher at 700 hPa than at 250 hPa. **This could be explained by the way the assimilation system redistributes the increment after the minimization of the cost function. The information in terms of CO content given to the system is important in the lower levels compared to the higher levels.**

4.2 Comparison in terms of zonal means

In this section, we compare the CO vertical profiles calculated from TOTCOL_ANALYSES, in terms of zonal means, to the MOPITT observations. Figure 5 shows, for the three types of measurements, CO monthly zonal means of MOPITT observations and their corresponding collocated TOTCOL_ANALYSES in terms of vertical profiles for August 2008 from the surface up to 150 hPa: the upper level of the MOPITT V3 observations. For the three types of measurements, the two zonal mean distributions (observation and TOTCOL_ANALYSES) show similar patterns. They both show the regions of CO emissions, particularly the biomass burning region in the latitude range between 0° and 20° S as well as the CO emissions in the NH.

For LAND_DAY and LAND_NIGHT measurement types, the CO vertical distribution is similar in both fields: over the emission regions, the maximum of CO extends up to 220 hPa over Africa and up to 150 hPa in the subtropical regions of the NH. These features of upper troposphere CO outflow reflect surface CO emissions lifted by convection. Nevertheless, MOPITT total column analyses slightly overestimate CO concentrations in the NH and in the tropical regions (up to: $+30$ % for LAND_NIGHT and $+20$ % for LAND_DAY). For the SEA measurement type, both zonal means have generally the same distributions. In the NH, both fields show high CO concentrations corresponding to the anthropogenic emissions over North America, Europe and Asia. However, in the SH the maximum difference between the two zonal means for the SEA type ranges between -10 % and $+20$ %. Generally, in the SH both fields show very moderate CO concentrations reflecting very low CO emissions over this region.

4.3 Comparison in terms of vertical profiles at regional scales

In this section, we compare the vertical profiles calculated from TOTCOL_ANALYSES at different regional scales in to the MOPITT observations. Figure 6 shows the main regional domains for which the evaluation of MOPITT total column analyses is done by comparison to MOPITT observations. These domains are considered as the regions having the bulk of the CO sources which are significantly different. The choice of these domains is consistent with the results of Liu et al. (2006) who state that the most important sources of CO variability in the troposphere are synoptic disturbances which have spatial scales of hundreds to thousands of kilometers. Consequently, it is important to have a statistical assessment of the variability of the two fields (MOPITT observations and TOTCOL_ANALYSES) over these regional areas. This will allow us to examine their respective behavior with respect to different types of emissions at the different regional scales.

Figure 7 presents a comparison between MOPITT CO profiles and their co-located profiles derived from the CO total column analyses over the six regional domains. Both datasets are averaged over each domain for the 7 MOPITT levels. Over the six domains, the different mean profiles are matching very well. Note also that the CO concentrations over sea are generally lower than those over land, especially at lower levels. The vertical profiles from the two datasets are very similar **and agree within their standard deviations**. Note also that the most significant variabilities of both datasets over all domains, especially domains 5, 6 and 3, are located at the lowermost levels (between the surface and 700 hPa). This reflects the variability of CO sources near the surface in Africa, South America and East Asia.

The mean bias as well as the corresponding RMS (Root Mean Square) between both datasets over the six domains of comparison for the three types of measurements are presented in Fig. 8.

The absolute mean bias does not exceed $14\,\%$, and is generally higher at lower levels (from the surface up to 700 hPa). For LAND_NIGHT and SEA types, the mean bias is generally positive for all domains at all pressure levels, reflecting an overestimation of the vertical profile deduced from TOTCOL_ANALYSES in comparison with MOPITT observations. The LAND_DAY type is generally characterized by a large positive bias with a corresponding RMS higher than that of other types particularly at the lowermost levels. This reflects a higher variability of TOTCOL_ANALYSES for LAND_DAY compared to the other types of measurements.

For all types of measurements over all domains, both the bias and the RMS are large between the surface and 700 hPa. This is in agreement with the results of Fig. 7 showing high variability in this altitude range. From 500 hPa up to 150 hPa, both quantities have generally small values. The vertical profile of the correlation coefficient between both datasets over the six domains of comparison is presented in Fig. 9. The correlation coefficient ranges from $\sim$0.6 to 0.95. The correlation is generally good in the mid-troposphere (500 hPa). This means that the added value to the model from MOPITT total column is more pronounced in the mid-troposphere compared to the lower levels. This is due to the redistribution of the CO column information by the assimilation system which is important in the lower levels compared to the high levels.

The results show that the vertical profiles deduced from TOTCOL_ANALYSES and those of MOPITT observations are generally in good agreement from the surface up to 150 hPa, particularly in the mid-troposphere.

5 Comparison of CO deduced from total column assimilation and CO deduced from vertical profile assimilation

In this section, we compare the vertical profiles calculated from MOPITT CO TOTCOL_ANALYSES for which the observation errors have been specified using the method presented in this paper (see: Sect. 3.3) and the vertical profiles issued from PROFILE_ANALYSES. The objective is to evaluate the differences between both analyses.

5.1 Comparison in terms of horizontal maps

In this section, we compare the vertical profiles derived from both analyses in terms of horizontal maps at different pressure levels.

Figure 10 presents a comparison, at 700 hPa, between the vertical profiles calculated from TOTCOL_ANALYSES with those from PROFILE_ANALYSES. These later are considered as the reference since they are assimilated with all their retrieval characteristics. Consequently, they should present the most realistic state of the atmosphere. Both fields are presented at global scale and averaged over the month of August 2008. The CO total column analyses and vertical profile analyses are very similar at 700 hPa. The mean bias between both quantities over the globe is very low ($\sim 6\,\%$ in average). This mean bias is still in the range of the mean specified observation errors ($\sim$7–8%) except over some local areas where the maximum difference ranges between ~ -12 % and $\sim +14$ %. However, CO fields are different from the fields of the model free run highlighting the added value of the assimilation results (Fig. 10). For example, over the regions of South-America, central Africa and Asia, the free run results differ from the analyses at 700 hPa. Figure 11 presents the same comparison as for Fig. 10 but at 200 hPa. The same conclusion as for 700 hPa can be deduced: the profiles deduced from TOTCOL_ANALYSES are very close to those issued from PROFILE_ANALYSES with exactly the same patterns especially over the emission regions : Africa and south of Asia. The maximum mean bias between both fields is ranging between -3 and $+10$ %. However, the comparison between the model free run field and the vertical profile analyses shows a bias which exceeds 60 % even if the general patterns between both fields are almost the same. These results confirm again that the CO fields deduced from PROFILE_ANALYSES and obtained from TOTCOL_ANALYSES are almost the same with very small differences. The relative mean bias between the two datasets is very small and is generally within the specified errors.

5.2 Comparison in terms of zonal means

In this section, we evaluate the differences between the two analyses in terms of zonal means. In this way, we present in Fig. 12 a comparison of CO zonal mean fields between the PROFILE_ANALYSES, TOTCOL_ANALYSES and the MOCAGE free run model. The CO distribution is similar for both analyses (total column and vertical profiles). Over the SH in the extratropics both fields show moderate values of CO from the surface up to the mid-troposphere ($\sim$400 hPa). CO concentrations from TOTCOL_ANALYSES are slightly overestimated compared to those from PROFILE_ANALYSES. The mean bias between both analyses (Fig. 12-Middle) is positive and does not exceed 12% over the vertical. In the tropics, both fields show strong CO emissions over Africa that can reach $\sim$200 hPa. Over this region, the differences between the two fields are very small ranging from -5 % to $+9$ %. In the NH, the two fields show very high CO concentrations in the mid-troposphere. These high CO concentrations correspond to anthropogenic emissions from North America, Europe and East Asia. The mean bias between both analyses ranges between -12 and $+12$ % which shows that both fields are very similar over the altitude range from the surface up to 150 hPa. However, the comparison between the zonal means deduced from PROFILE_ANALYSES against those of the MOCAGE model free run (Fig. 12-Bottom) shows a bias ranging between -35 % and 45 %, particularly in the mid-troposphere of the tropical regions and the lower troposphere of the extra-tropics ($<$ 40 %). These results show that the information derived from the total columns using data assimilation is capable of modifying the vertical structure of the CO distribution over the whole troposphere, showing features very similar to those obtained from the assimilation of MOPITT CO profiles.

5.3 Comparison in terms of vertical profiles at regional scales

In this section, we evaluate the differences between the two analyses in terms of vertical profiles at regional scales. We compare, at the same regional scales shown in Fig. 6, the CO vertical profiles deduced from TOTCOL_ANALYSES, and those obtained from PROFILE_ANALYSES. The vertical profiles calculated from both analyses are averaged over different domains for August 2008. Figure 13 presents the vertical profiles with their associated standard deviations. This later represents the variability of the CO concentration over each domain for the month of August 2008. The profiles calculated from both analyses as well as their associated standard deviations are similar for all domains.

Both analyses show the same behavior for the CO fields in terms of vertical structure at the regional scales, and have similar variability. The maximum standard deviation is generally found at pressure levels between the surface and 700 hPa for both analyses, especially for domains 5 (Africa), 6 (South America), and 3 (East Asia). This is in good agreement with the results of Fig. 8 which illustrates again the variability of CO sources over Africa, South America and East Asia. Figure 13 confirms that TOTCOL_ANALYSES and PROFILE_ANALYSES provide almost the same vertical structure over regional scales. This shows again that the assimilation of total column impacts all the vertical levels of the profile in the same way as the assimilation of the vertical profiles.

Figure 14 presents the vertical profiles of the mean bias and the corresponding RMS between the two assimilation set-ups (TOTCOL_ANALYSES and PROFILE_ANALYSES) averaged over each domain for the month of August 2008. For all regional domains, the mean bias has low values at all pressure levels and is generally less than 10% except for domains 2 (Europe) and 3 (East Asia) at 150 hPa, where it reaches $\sim 13\,\%$. The values of the RMS range between $+10$ % and $+15$ % for most domains. All these values are smaller or in the range of the expected errors of the assimilation results, and are generally smaller than the observation error values used in the assimilation process. The only exception concerns domain 6 (South America), for which the corresponding RMS is about $20\,\%$ in the altitude range between the surface and 400 hPa. This could be attributed to the large variability of the CO field in this domain (see Fig. 12). The vertical profiles of the correlation coefficient between the two analyses over the six domains of comparison are presented in Fig. 15. For the different domains, the correlation coefficients range between 0.75 and 0.99, with most of the values close to 0.9. This shows a very good agreement between the two assimilation results.

The results shown in this section concern the following statistics calculated between both datasets: bias, RMS and correlation coefficient. They show that the comparison between the vertical profiles deduced from TOTCOL_ANALYSES and those obtained from PROFILE_ANALYSES is consistent. Both analyses are in very good agreement at the global and regional scales.

5.4 Validation of the analyses with MOZAIC independent data

To further evaluate both analyses, we compare theme to MOZAIC measurements. The MOZAIC programme was launched in January 1993. The measurements started in August 1994, with the installation of ozone and water vapor sensors aboard 5 commercial aircraft. In 2001, the instrumentation was upgraded by installing carbon monoxide sensors on all aircraft and a total odd nitrogen instrument ($\mathrm{NO_y}$) aboard one aircraft. Ozone is measured by UV absorption (Thermo Instruments, Model 49–103). The instruments are calibrated before and after each period of deployment ($\sim$every 12 months) and in-flight quality control is achieved, both for bias and calibration factor, with a built-in ozone generator. A comparison of the first 2 years of MOZAIC observations with data of the ozonesonde network showed good agreement (Thouret et al., 1998). For CO measurements, the infra–red (IR) gas filter correlation technique is employed (Thermo Environmental Instruments, Model 48CTL). This IR instrument provides excellent stability, which is important for continuous operation without frequent maintenance. The sensitivity of the instrument was improved by several modifications (Nédélec et al., 2003), achieving a precision of ± 5 ppbv (parts per billion volume) or $\pm 5\%$ for a 30 s response time.

The comparison was conducted with collocated vertical profiles for the three datasets (MOZAIC, TOTCOL_ANALYSES and PROFILE_ANALYSES) over eight MOZAIC airports visited over the assimilation period (Atlanta, Caracas, Dallas, Frankfurt, Hyderabad, London, Philadelphia and Windhoek). These airports are located in the domain lat:[51.6°N–22.6°S], lon:[96.8°W–78.4°E]. For the three datasets, collocated observations are selected in 2° radius area over each of the eight airports. The comparisons of TOTCOL_ANALYSES and PROFILE_ANALYSES to MOZAIC observations at all visited airports are presented in Fig. 16. The two analyses behave similarly over all airports. The agreement of both analyses compared to MOZAIC is very good. We note that at only one level (850 hPa) over Caracas, the difference between MOZAIC and both analyses exceed 40 ppbv. Note also that for only 6% of measurements, the difference between the analyses and MOZAIC is in the range of 20–25 ppbv. This difference does not exceed 5 ppbv for more than 50% of measurements. These results show a very good agreement between the two analyses compared to the in–situ MOZAIC independent data.

6 Conclusions

The aim of this paper is to describe a method to derive the vertical profile of CO from its total column with no associated error covariances and averaging kernels using data assimilation. We have chosen version 3 of MOPITT CO total columns to validate the proposed method since it has the advantage of providing both the vertical profiles and the total columns of CO.

The method is based on the estimation of the observation error covariance matrices (diagonal of the R matrix) using the χ^2 test to obtain consistency between model and observa-

tion errors. This specification has been done by discriminating the observations according to day, night, land and sea. The appropriate observation errors are 8% and 11 % for measurements performed over land during daytime (LAND_DAY) and over land during nighttime (LAND_NIGHT), respectively. For measurements performed over sea during daytime and nighttime (SEA), the observation error is 7 %. The a posteriori diagnostics concerning the analyses for all specified total column observations confirm that the specified errors, for different types, using the proposed method as well as the corresponding forecasts error, have a Gaussian structure.

In the first comparison, CO profiles from MOPITT total column analyses and MOPITT observations show similar patterns in terms of longitude–latitude maps at 700 hPa and 250 hPa. The mean bias at 700 hPa between the two datasets is 15 $\%$, 18 $\%$ and 12 $\%$ for LAND_DAY, LAND_NIGHT and SEA types, respectively. At 250 hPa, these respective mean biases are $+12\,\%$, $+8\,\%$ and $+7\,\%$ for LAND_DAY at 250 hPa. The comparison of the zonal means shows that the CO vertical distribution is homogeneous in both fields from the surface up to 150 hPa. At regional scales, the comparison of the two datasets in terms of vertical profiles shows very good agreement. The mean bias is generally large at low levels but does not exceed $+10$ %.

In the second comparison, the analyses issued from the MOPITT vertical profiles and the CO total column analyses show a very good agreement over the globe. The general aspect is consistent. The CO fields present the same features particularly over the emission regions in central Africa, South-Eastern Asia and Southern America. The mean bias between both datasets is 6 $\%$ and 8 $\%$ at 700 and 200 hPa, respectively. In terms of zonal means, the CO distribution is similar for the two analyses with very low differences. The total column analyses tend to slightly overestimate the CO concentrations. The maximum mean bias does not exceed 15 $\%$ over all levels. Over regional scales, the vertical profiles calculated from both analyses are in very good agreement. The mean bias is very small and generally does not exceed $+10$ %, whereas the vertical profile of the correlation coefficient ranges from 0.75 to 0.99. These results concerning the CO distributions, vertical profiles, mean bias, RMS and correlation coefficient, confirm that the analyses of the CO total column assimilation are in very good agreement with the analyses calculated from the assimilation of the MOPITT CO profiles. This agreement is confirmed at global and regional scales.

Both analyses have also been validated using in-situ MOZAIC independent data. The comparison was conducted with collocated vertical profiles for the three datasets over eight airports visited over the assimilation period located in the domain lat:[51.6°N–22.6°S], lon:[96.8°W–78.4°E]. The comparisons of the two analyses to MOZAIC data over all the visited airports show a very good agreement. Only at one level (850 hPa) over Caracas, the difference between MOZAIC and the two analyses exceeds 40 ppbv. However, the difference between the analyses and MOZAIC is in the range of 20–25 ppbv for Only 6% of measurements. This

difference does not exceed 5 ppbv for more than 50% of measurements.

Note finally that the DFS of MOPITT V3 is relatively low for vertical profiles ($\sim$1.5) as well as for the total columns($\sim$1). In this paper we have demonstrated that for this kind of data, the present method consisting of deducing the profiles from the total columns, remains valid when only using the adjoint of the integration operator. Note that for other types of data for which the DFS is greater than that of MOPITT V3, the present method has to be tested and evaluated against independent observations.

Acknowledgements. This work is funded in France by the Centre National de Recherches Météorologiques (CNRM) of Météo-France and the Centre National de Recherches Scientifiques (CNRS). The authors acknowledge the Région Midi-Pyrénées (INFOAIR project) and thank the RTRA/STAE fondation for their support. The authors acknowledge also for the strong support of the European Commission, Airbus, and the Airlines (Lufthansa, Austrian, Air France) who carry free of charge the MOZAIC equipment and perform the maintenance since 1994. MOZAIC is presently funded by INSU-CNRS (France), Météo-France, and Forschungszentrum (FZJ, Julich, Germany). The MOZAIC data based is supported by ETHER, the French atmospheric composition database (CNES and CNRS-INSU).

References

Bencherif, H., El Amraoui, L., Kirgis, G., Bellevue, J. L. D., Hauchecorne, A., Hauchecorne, A., Mz, N., Portafaix, T., Pazmino, A., and Goutail, F.: Analysis of a rapid increase of stratospheric ozone during late austral summer 2008 over Kerguelen (49.4°S, 70.3°E), Atmos. Chem. Phys., 11, 363–373, doi:10.5194/acp-11-363-2011, www.atmos-chem-phys.net/11/363/2011/, 2011.

Bian, H., Chin, M., Kawa, S. R., Duncan, B., Arellano, A., and Kasibhatla, P.: Sensitivity of global CO simulations to uncertainties in biomass burning sources, J. Geophys. Res., 112, D23308, doi:10.1029/2006JD008376, 2007.

Claeyman, M., Attié, J.-L., El Amraoui, L., Cariolle, D., Peuch, V.-H., Teyssèdre, H., Josse, B., Ricaud, P., Massart, S., Piacentini, A., Cammas, J.-P., Livesey, N. J., Pumphrey, H. C., and Edwards, D. P.: A linear CO chemistry parameterization in a chemistry-transport model: evaluation and application to data assimilation, Atmos. Chem. Phys., 10, 6097–6115, doi:10.5194/acp-10-6097-2010, http://www.atmos-chem-phys.net/10/6097/2010/, 2010.

Claeyman, M., Attie, J.-L., Peuch, V.-H., El Amraoui, L., Lahoz, W. A., Josse, B., Ricaud, P., von Clarmann, T., Hopfner, M., Orphal, J., Flaud, J.-M., Edwards, D. P., Chance, K., Liu, X., Pasternak, F., and Cantie, R.: A geostationary thermal infrared sensor to monitor the lowermost troposphere: O_3 and CO retrieval studies, Atmos. Meas. Tech., 4, 297–317, doi:10.5194/amt-4-297-2011, www.atmos-meas-tech.net/4/297/2011/, 2011.

Clerbaux, C., Hadji-Lazaro, J., Hauglustaine, D., Megie, G., Khattatov, B., and Lamarque, J.-F.: Assimilation of carbon monoxide measured from satellite in a three-dimensional chemistry-transport model, J. Geophys. Res., 106, 15 385–15 394, 2001.

Courtier, P., Freydier, C., Geleyn, J., Rabier, F., and Rochas, M.: The ARPEGE project at Météo France, in: Atmospheric Models, vol.2, pp. 193–231, Workshop on Numerical Methods, Reading, U.K., 1991.

Crutzen, P. J. and Andreae, M. O.: Biomass Burning in the Tropics: Impact on Atmospheric Chemistry and Biogeochemical Cycles, Science, 250, doi:10.1126/science.250.4988.1669, 1990.

Deeter, M. N., Edwards, D. P., Gille, J. C., and Drummond, J. R.: Sensitivity of MOPITT observations to carbon monoxide in the lower troposphere, Journal of Geophysical Research (Atmospheres), 112, D24306, doi:10.1029/2007JD008929, 2007.

Drummond, J. and Mand, G.: The measurements of pollution in the troposphere (MOPITT) instrument: Overall performance and calibration requirements, J. Atmos. Oceanic Technol., 13, 314–320, 1996.

El Amraoui, L., Peuch, V.-H., Ricaud, P., Massart, S., Semane, N., Teyssèdre, H., Cariolle, D., and Karcher, F.: Ozone loss in the 2002–2003 Arctic vortex deduced from the assimilation of Odin/SMR O_3 and N_2O measurements: N_2O as a dynamical tracer, Quart. J. Roy. Meteor. Soc., 134, 217–228, 2008a.

El Amraoui, L., Semane, N., Peuch, V.-H., and Santee, M. L.: Investigation of dynamical processes in the polar stratospheric vortex during the unusually cold winter 2004/2005, Geophys. Res. Lett., 35, L03803, doi:10.1029/2007GL031251, 2008b.

El Amraoui, L., Attié, J.-L., Semane, N., Claeyman, M., Peuch, V.-H., Warner, J., Ricaud, P., Cammas, J.-P., Piacentini, A., Josse, B., Cariolle, D., Massart, S., and Bencherif, H.: Midlatitude stratosphere–troposphere exchange as diagnosed by MLS O_3 and MOPITT CO assimilated fields, Atmos. Chem. Phys., 10, 2175–2194, doi:10.5194/acp-10-2175-2010, http://www.atmos-chem-phys.net/10/2175/2010/, 2010.

Emili, E., Barret, B., Massart, S., Le Flochmoen, E., Piacentini, A., El Amraoui, L., Pannekoucke, O., and Cariolle, D.: Combined assimilation of IASI and MLS observations to constrain tropospheric and stratospheric ozone in a global chemical transport model, Atmos. Chem. Phys. Discuss., 13, 21 455–21 505, doi: 10.5194/acpd-13-21455-2013, http://www.atmos-chem-phys-discuss.net/13/21455/2013/, 2013 : accepted in Atmos. Chem. Phys.).

Emmons, L., Deeter, M., Gille, J., Edwards, D., Attié, J.-L., Warner, J., Ziskin, D., Francis, G., Khattatov, B., Lamarque, V. Y. J.-F., Ho, S., Mao, D., Chen, J., Drummond, J., Novelli, P., Sachse, G., Coffey, M., Hannigan, J., Gerbig, C., Kawakami, S., Kondo, Y., Takegawa, N., Schlager, H., Baehr, J., and Ziereis, H.: Validation of Measurements of Pollution in the Troposphere (MOPITT) CO retrievals with aircraft in situ profiles, J. Geophys. Res., 109, doi:10.1029/2003JD004101, 2004.

Emmons, L., Edwards, D., Deeter, M., Gille, J., Campos, T., Nédélec, P., Novelli, P., and Sachse, G.: Measurements of Pollution In The Troposphere (MOPITT) validation through 2006, Atmos. Chem. Phys., 9, 1795–1803, 2009.

Emmons, L. K., Pfister, G. G., Edwards, D. P., Gille, J. C., Sachse, G., Blake, D., Wofsy, S., Gerbig, C., Matross, D., and Nédélec, P.: Measurements of Pollution in the Troposphere (MOPITT) validation exercises during summer 2004 field campaigns over North America, Journal of Geophysical Research (Atmospheres), 112, D12S02, doi:10.1029/2006JD007833, 2007.

Fisher, M. and Andersson, E.: Developments in 4D-Var and Kalman filtering, in: Technical Memorandum No. 347, p. 36, Research Department, ECMWF, Reading, UK, 2001.

Granier, C., Petron, G., Muller, J. F., and Brasseur, G.: The impact of natural and anthropogenic hydrocarbons on the tropospheric budget of carbon monoxide, Atmos. Environ., 34, 5255–5270, 2000.

Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C.: Unified Notations for Data Assimilation: Operational, Sequential and Variational, J. Meteor. Soc. Japan, 75, 181–189, 1997.

Jones, D., Bowman, K. W., Palmer, P., Worden, J., Jacob, D., Hoffman, R., Bey, I., and Yantosca, R.: Potential of observations from the Troposphere Emission Spectrometer to constrain continental sources of carbon monoxide, J. Geophys. Res., 108(D24), 4789, doi:10.1029/2003JD003702, 2003.

Lahoz, W. A., Errera, Q., Swinbank, R., and Fonteyn, D.: Data assimilation of stratospheric constituents: a review, Atmos. Chem. Phys., 7, 5745–5773, 2007.

Lahoz, W. A., Peuch, V.-H., Orphal, J., Attie, J.-L., Chance, K., Liu, X., Edwards, D., Elbern, H., Flaud, J.-M., Claeyman, M., and El Amraoui, L.: Monitoring Air Quality from space: The case for the geostationary platform, Bull. Amer. Meteor. Soc., pp. 221–233, doi:10.1175/BAMS-D-11-00045.1, 2012.

Lamarque, J.-F., Khattatov, B., Gille, J., and Brasseur, G.: Assimilation of MAPS CO in a global three-dimensional model, J. Geophys. Res., 104, D21, doi:10.1029/1999JD900807, 1999.

Liu, J., Drummond, J. R., Jones, D. B. A., Cao, Z., Bremer, H., Kar, J., Zou, J., Nichitiu, F., and Gille, J. C.: Large horizontal gradients in atmospheric CO at the synoptic scale as seen by spaceborne Measurements of Pollution in the Troposphere, J. Geophys. Res., 111, doi:doi:10.1029/2005JD006076, 2006.

Mahieu, E., Zander, R., Delbouille, L., Demoulin, P., Roland, G., and Servais, C.: Observed trends in total vertical column abundances of atmospheric gases from IR solar spectra recorded at the Jungfraujoch, Journal of Atmospheric Chemistry, 28, 227–243, doi:10.1023/A:1005854926740, 1997.

Nédélec, P., Cammas, J.-P., Thouret, V., G., G. A., Cousin, J.-M., Legrand, C., Abonnel, C., Lecoeur, F., Cayez, G., and Marizy, C.: An improved infra-red carbon monoxide analyser for routine measurements aboard commercial Airbus aircraft: Technical validation and first scientific results of the MOZAIC III Program, Atmos. Chem. Phys., 3, 1551–1564, 2003.

Peuch, V.-H., Amodei, M., Barthet, T., Cathala, M. L., Michou, M., and Simon, P.: MOCAGE, MOdéle de Chimie Atmosphérique à Grande Echelle, in: Proceedings of Météo France: Workshop on atmospheric modelling, pp. 33–36, Toulouse, FRANCE, 1999.

Pradier, S., Attié, J.-L., Chong, M., Escobar, J., Peuch, V.-H., Lamarque, J.-F., Khattatov, B., and Edwards, D.: Evaluation of 2001 springtime CO transport over West Africa using MOPITT CO measurements assimilated in a global chemistry transport model, Tellus, 58B, 163–176, 2006.

Rabier, F., Bouchard, A., Brun, E., Doerenbecher, A., Guedj, S., Guidard, V., Karbou, F., Peuch, V., El Amraoui, L., Puech, D., Genthon, C., Picard, G., Town, M., Hertzog, A., Vial, F., Cocquerez, P., Cohn, S., Hock, T., Fox, J., Cole, H., Parsons, D., Powers, J., Romberg, K., VanAndel, J., Deshler, T., Mercer, J., Haase, J., Avallone, L., Kalnajs, L., Mechoso, C., Tangborn, A., Pellegrini, A., Frenot, Y., Thpaut, J., McNally, A., Balsamo, G., and Steinle, P.: The Concordiasi Project in Antarctica, Bull. Amer. Meteor. Soc., 91, 69–86, doi:10.1175/2009BAMS2764.1, 2010.

Rodgers, C.: Inverse Methods for Atmospheric Sounding: Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics, Vol. 2, 238 pp., World Scientific, Hackensack, New Jersey, 2000.

Semane, N., Peuch, V.-H., El Amraoui, L., Bencherif, H., Massart, S., Cariolle, D., Attié, J.-L., and Abida, R.: An observed and analysed stratospheric ozone intrusion over the high Canadian Arctic UTLS region during the summer of 2003, Quart. J. Roy. Meteor. Soc., 133, 171–178, 2007.

Tangborn, A., Stajner, I., Buchwitz, M., Khlystova, I., Pawson, S., Burrows, J., Hudman, R., and Nedelec, P.: Assimilation of SCIAMACHY total column CO observations: Global and regional analysis of data impact, J. Geophys. Res., 114, doi:10.1029/2008JD010781, 2009.

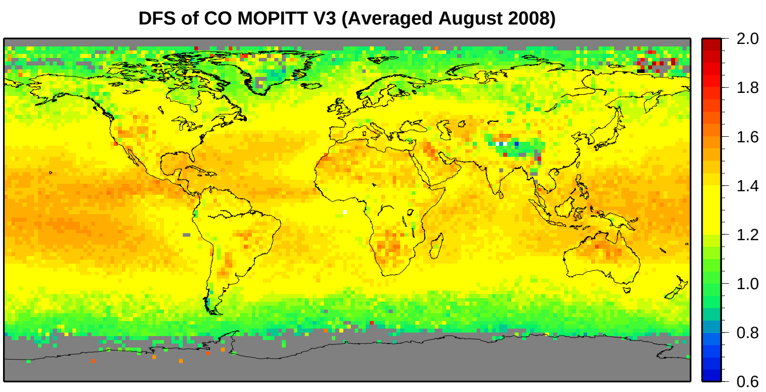
Thouret, V., Marenco, A., Logan, J., Nédélec, P., and Grouhel, C.: Comparisons of ozone measurements from the MOZAIC airborne program and the ozone sounding network at eight locations, J. Geophys. Res., 103, 25 695–25 720, 1998.

Weaver, A. and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, Quart. J. Roy. Meteor. Soc., 127, 1815–1846, doi:10.1002/qj.49712757518, http://dx.doi.org/10.1002/qj.49712757518, 2001.

Yurganov, L. N., McMillan, W. W., Dzhola, A. V., Grechko, E. I., Jones, N. B., and van der Werf, G. R.: Global AIRS and MOPITT CO measurements: Validation, comparison, and links to biomass burning variations and carbon cycle, Journal of Geophysical Research (Atmospheres), 113, D09301, doi:10.1029/2007JD009229, 2008.

Table 1. Mean and median values of χ^2 test for different error values of MOPITT V3 total column observations. The error values of the observations for which the χ^2 test is the closest to 1 are indicated in boldface. They are 8 % for LAND_DAY; 11 % for LAND_NIGHT and 7 % for SEA. These error values are fixed within the assimilation system for all experiments concerning MOPITT V3 total columns.

	χ^2							
	LAND_DAY			LAND_NIGHT			SEA	
R (%)	Mean value	Median value	R (%)	Mean value	Median value	R (%)	Mean value	Median value
4	4.37	3.98	8	2.15	2.04	4	2.59	2.54
5	2.66	2.36	9	1.60	1.51	5	1.70	1.67
6	1.78	1.59	10	1.15	1.16	6	1.26	1.24
7	1.27	1.18	**11**	**1.05**	**1.01**	**7**	**0.96**	**0.94**
8	**0.97**	**0.92**	12	0.85	0.79	8	0.75	0.74
9	0.78	0.72	13	0.72	0.68	9	0.51	0.50



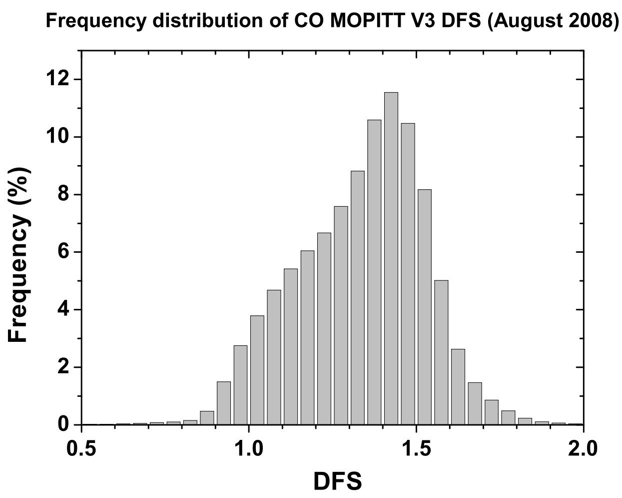


Fig. 1. (Top):Lon-lat map of the averaged DFS over August 2008 corresponding to the vertical profiles of MOPITT 3. (Bottom): the frequency distribution of the DFS corresponding to all vertical profiles measured during the same period.

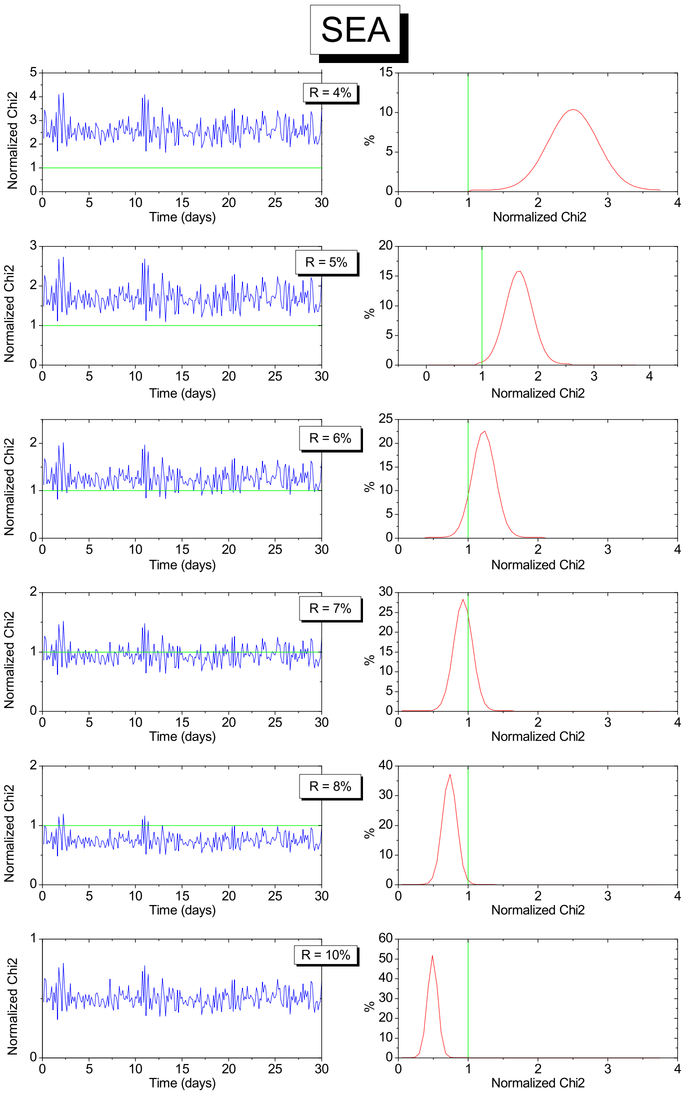


Fig. 2. Example of χ^2 diagnostic used to estimate the error of the total column for observations made over sea. (Left): time evolution of the normalized χ^2 value over the month of August 2008 for different values of observation error (Right): the Gaussian fit of the corresponding normalized histogram. The mean as well as the median values of χ^2 for each observation error value are reported in Table 1.

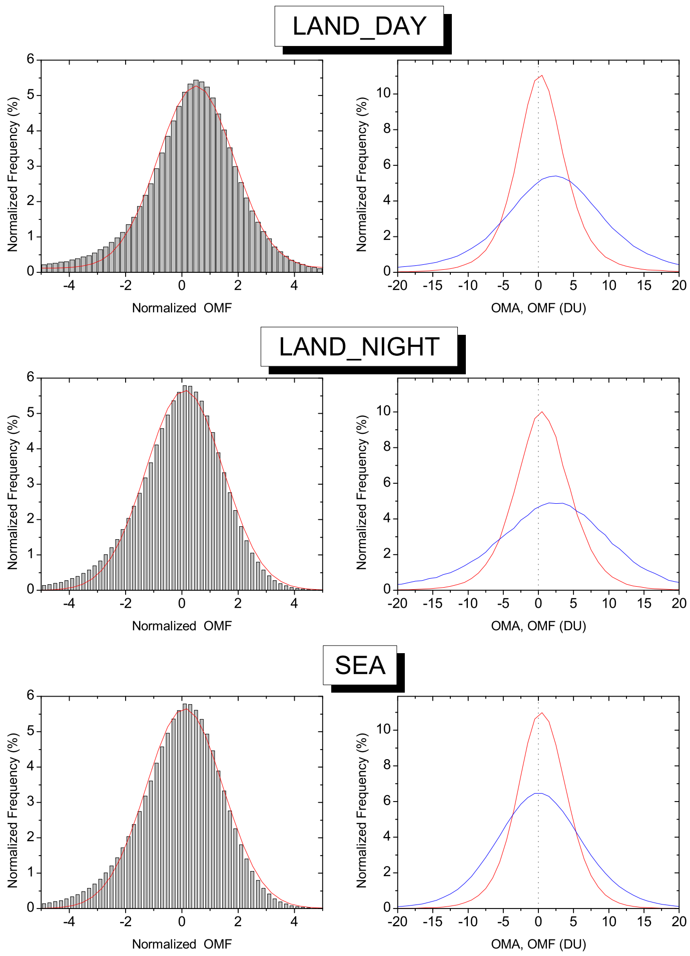


Fig. 3. A posteriori verification of the observation error specification for the analyses issued from the MOPITT CO total column for which the observation errors are estimated using the proposed method. (Left): histograms of OMF (Observations Minus Forecast) differences normalized by the specified observation errors. The red line is a Gaussian fit to the histogram. The good agreement between the histogram and the fit function supports the assumption of Gaussian errors in the observations and the forecast. (Right): histograms of Observations Minus Analysis (OMA: red lines) and OMF (blue lines).

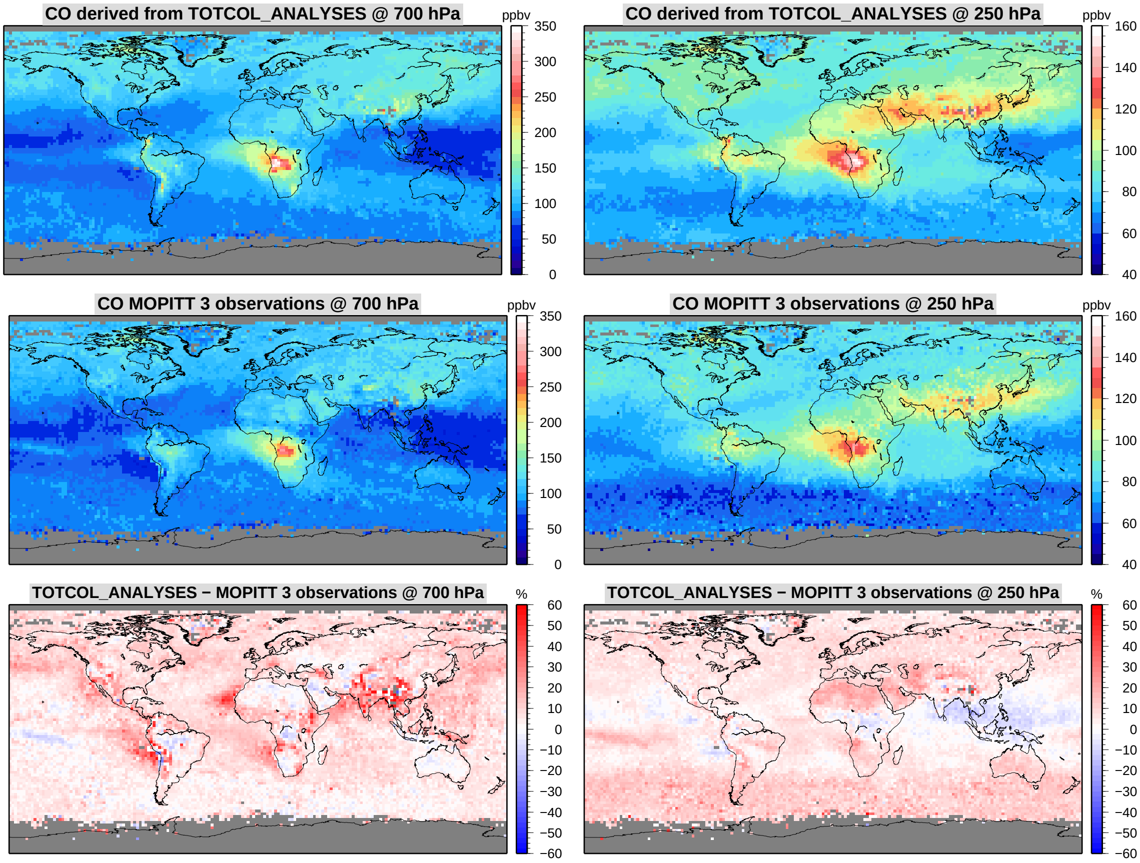


Fig. 4. Comparison of CO analyses obtained by the assimilation of MOPITT V3 CO total column observations with the optimal error estimated by the χ^2 test (top) to the operational MOPITT V3 CO retrieved profiles (middle) at 700 hPa (left panels) and 250 hPa (right panels). The corresponding relative differences between both datasets (TOTCOL_ANALYSES – observations) are indicated in the bottom panels for both pressure levels. Blue and red colors indicate negative and positive differences, respectively. Note that this figure corresponds to an average over August 2008 for all observations carried out over land during daytime.

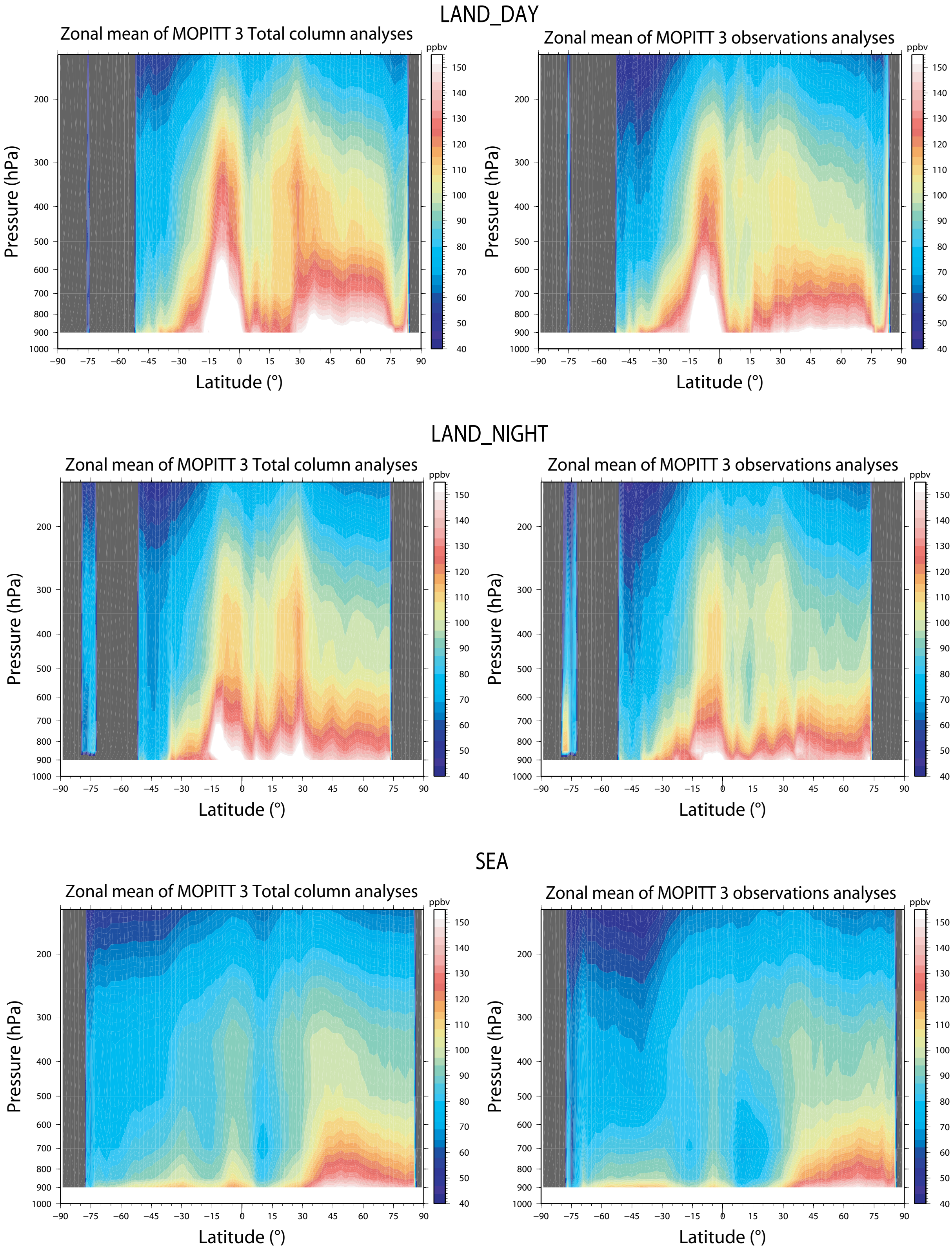


Fig. 5. Zonal mean of MOPITT CO TOTCOL_ANALYSES (left panels) compared to the zonal mean of the MOPITT CO observations (right panels) for August 2008. The comparison is done for observations made: over land during daytime (upper panel), over land during nighttime (middle panel) and over sea during daytime and nighttime (bottom panel).

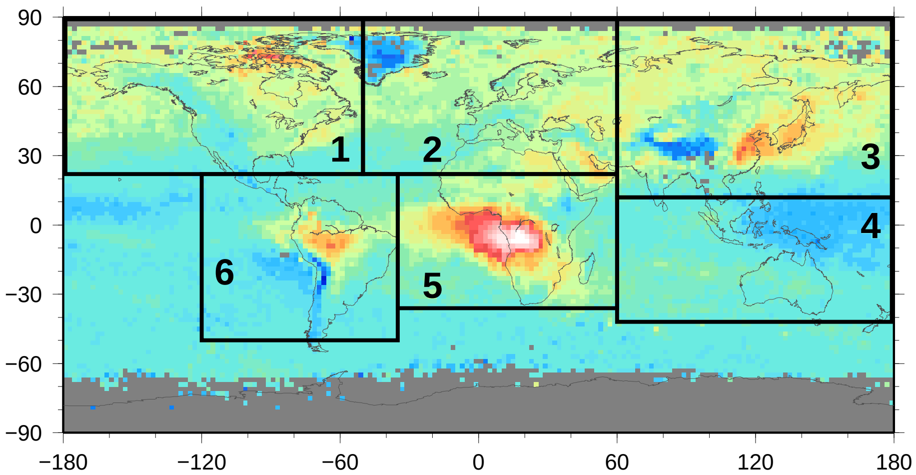


Fig. 6. Main domains of CO emissions considered for the regional validation of the proposed method dealing with the validation of CO TOTCOL_ANALYSES.

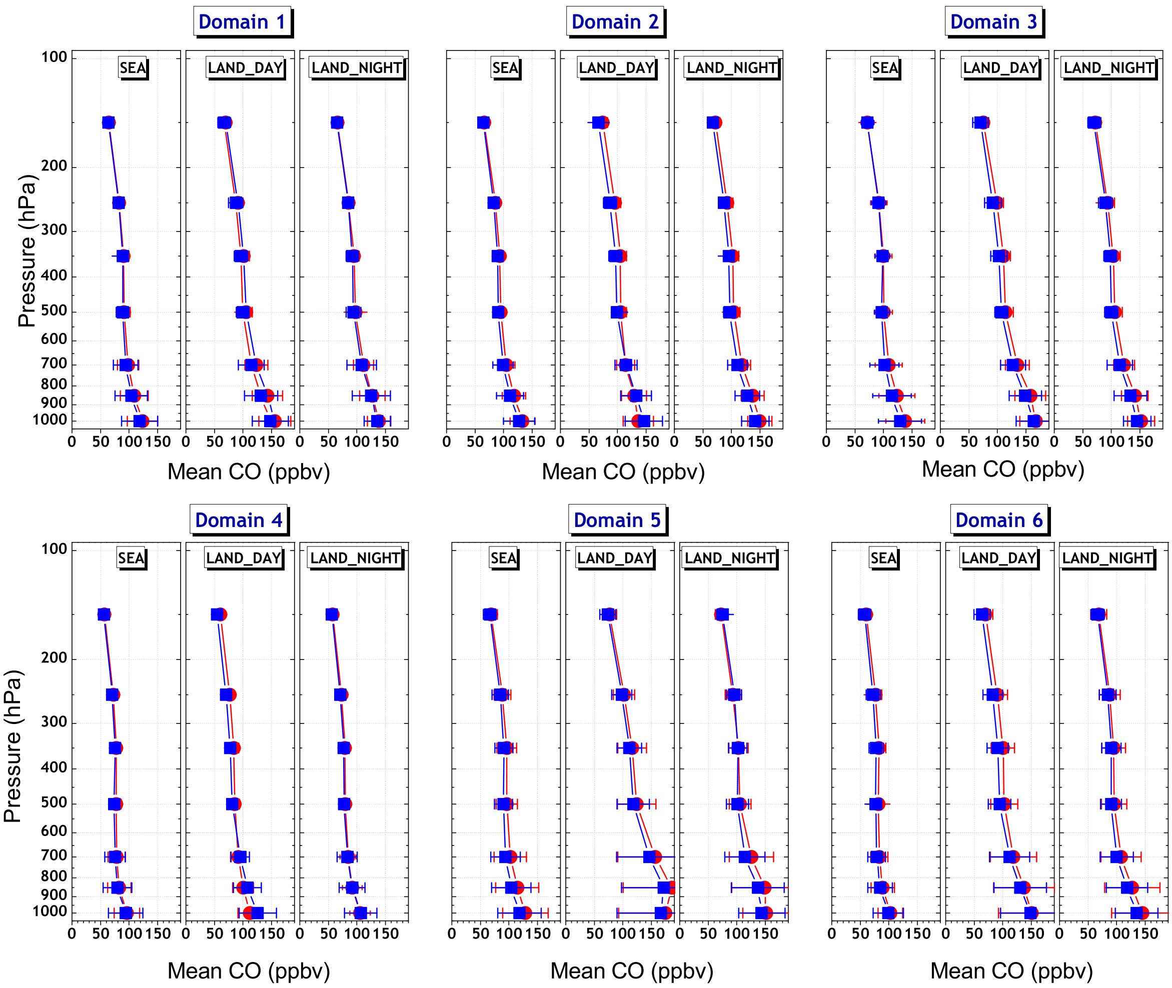


Fig. 7. The mean CO vertical profiles in parts per billion by volume (ppbv) deduced from MOPITT V3 CO TOTCOL_ANALYSES (blue) compared to the operational MOPITT V3 observations (red). Both datasets are averaged over August 2008, over all the domains defined in Fig. 5 and are associated with their corresponding standard deviations.

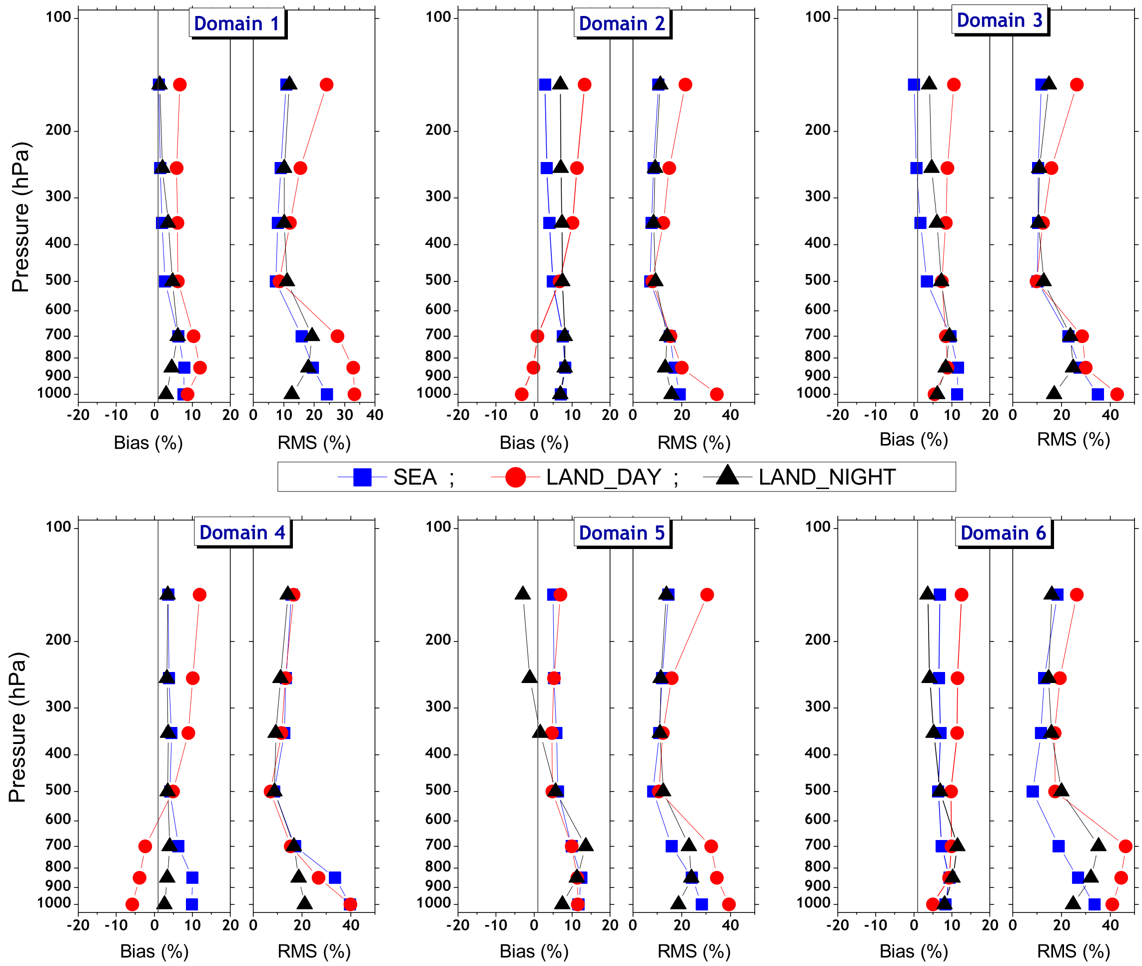


Fig. 8. The mean bias and the corresponding RMS (Root Mean Square) between CO vertical profiles deduced from the MOPITT V3 CO TOTCOL_ANALYSES and the MOPITT V3 observations. The comparison is made for observations carried out over land during daytime (red), those carried out over land during nighttime (black) and those carried out over sea (blue).

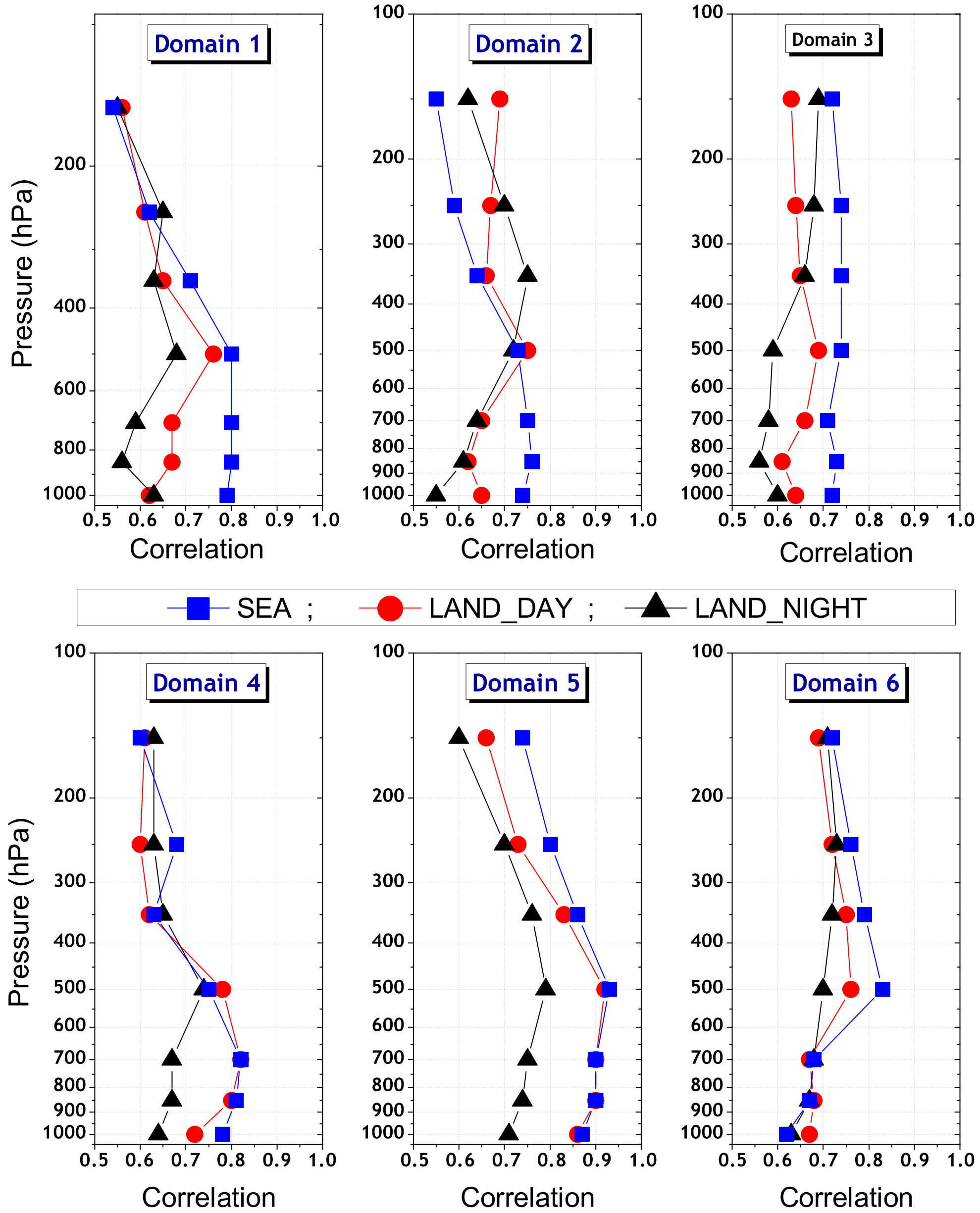


Fig. 9. Same as Fig. 8 but for the correlation coefficient between CO vertical profiles deduced MOPITT CO TOTCOL_ANALYSES and the MOPITT V3 observations. The comparison is made for each level of the MOPITT V3 retrievals.

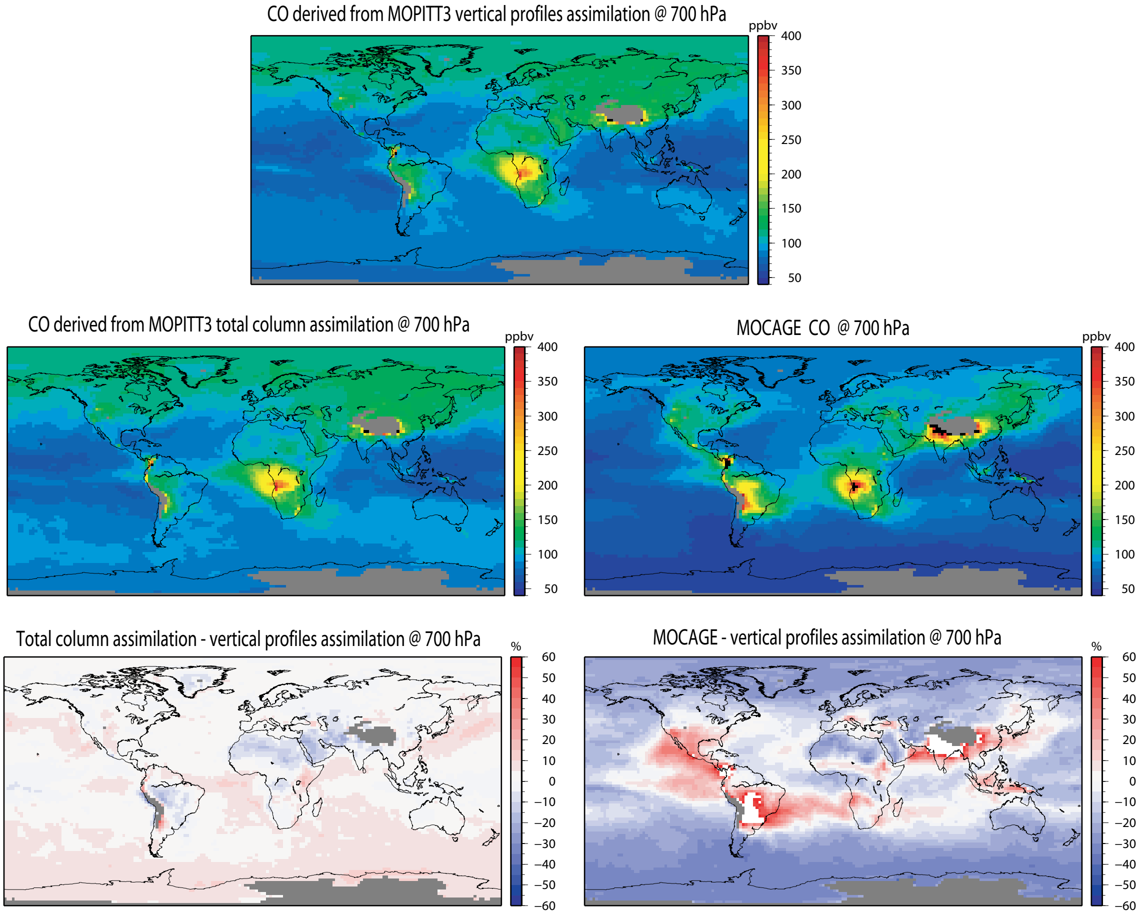


Fig. 10. Maps of CO field at 700 hPa for: (top) MOPITT PROFILE_ANALYSES taking into account averaging kernels and observation error covariance matrices; (middle-left) MOPITT TOTCOL_ANALYSES, and (middle-right) the MOCAGE free-run field. The Figures in the bottom present the difference in % between TOTCOL_ANALYSES and PROFILE_ANALYSES (left), and the difference between the model and PROFILE_ANALYSES (right).

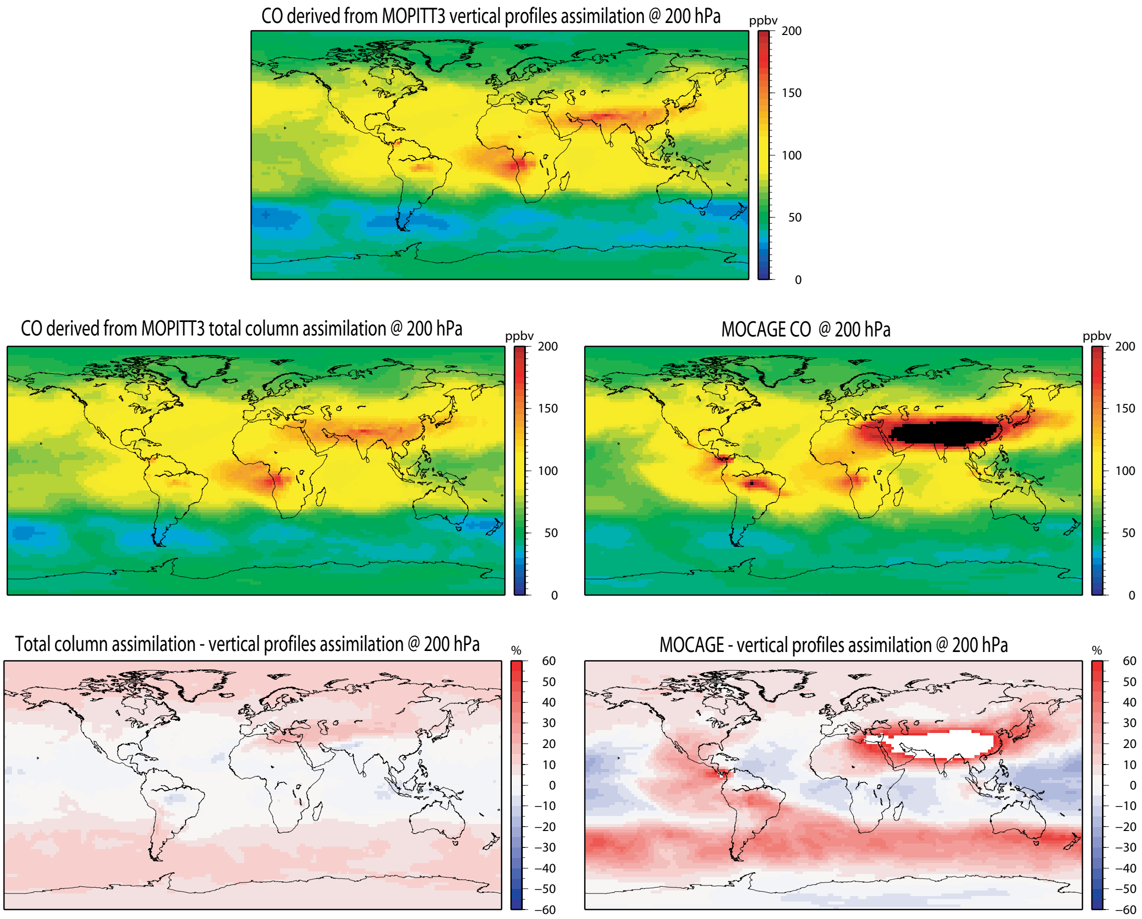


Fig. 11. Same as Fig. 10 but for the pressure level 200 hPa.

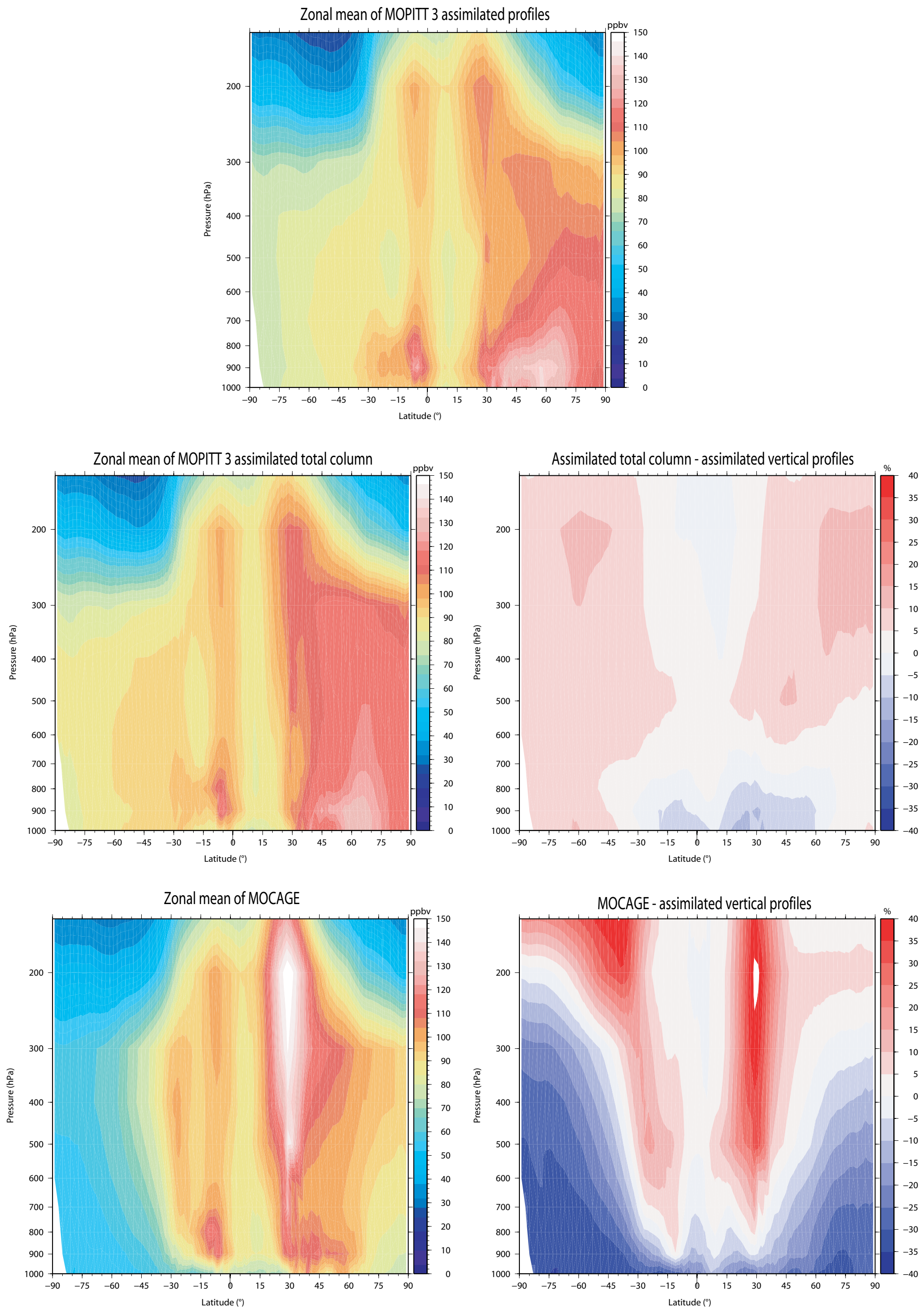


Fig. 12. Zonal means of CO field for the month of August 2008 as obtained by: (top) : MOPITT PROFILE_ANALYSES taking into account averaging kernels and observation error covariance matrices ; (middle-left) : MOPITT TOTCOL_ANALYSES ; and (bottom-left) : the MOCAGE free-run model. The figures in the right present the difference in % between TOTCOL_ANALYSES and PROFILE_ANALYSES (middle); and the difference between the free-run model and PROFILE_ANALYSES (bottom).

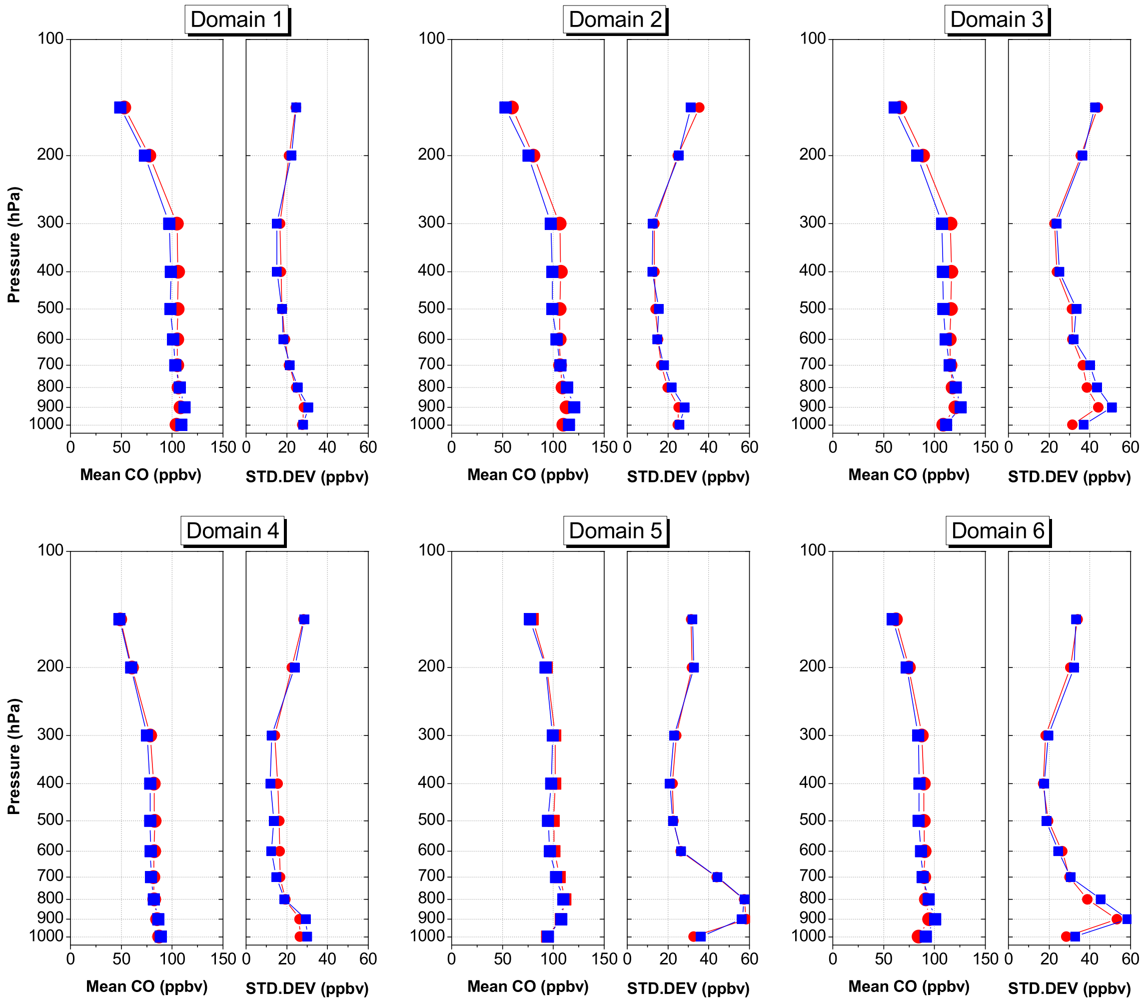


Fig. 13. Mean CO vertical profiles and their associated standard deviations in parts per billion by volume (ppbv) deduced from MOPITT CO TOTCOL_ANALYSES (red) compared MOPITT PROFILE_ANALYSES taking into account the averaging kernels as well as the error covariance matrices (blue). Both datasets are averaged over the month of August 2008 and over all the regional domains defined in Fig. 6.

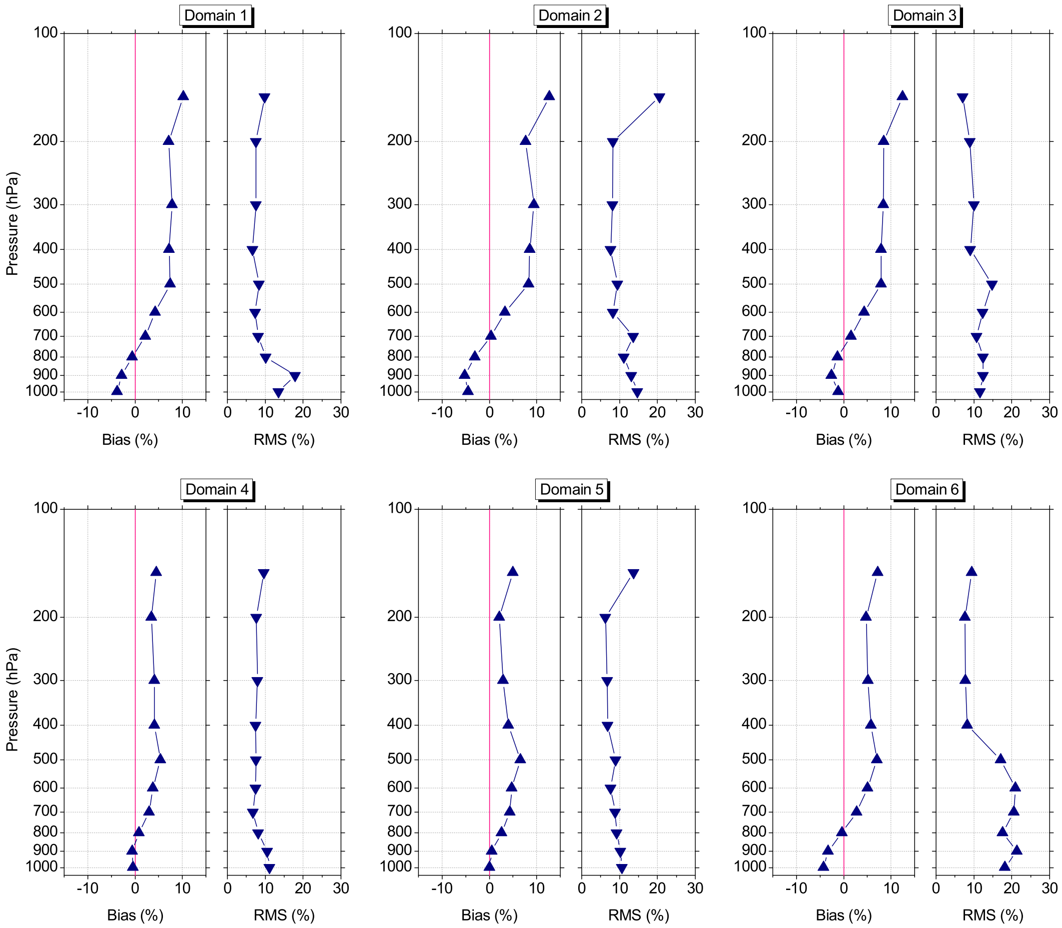


Fig. 14. Same as Fig. 13 but for the mean bias and the corresponding RMS (Root Mean Square).

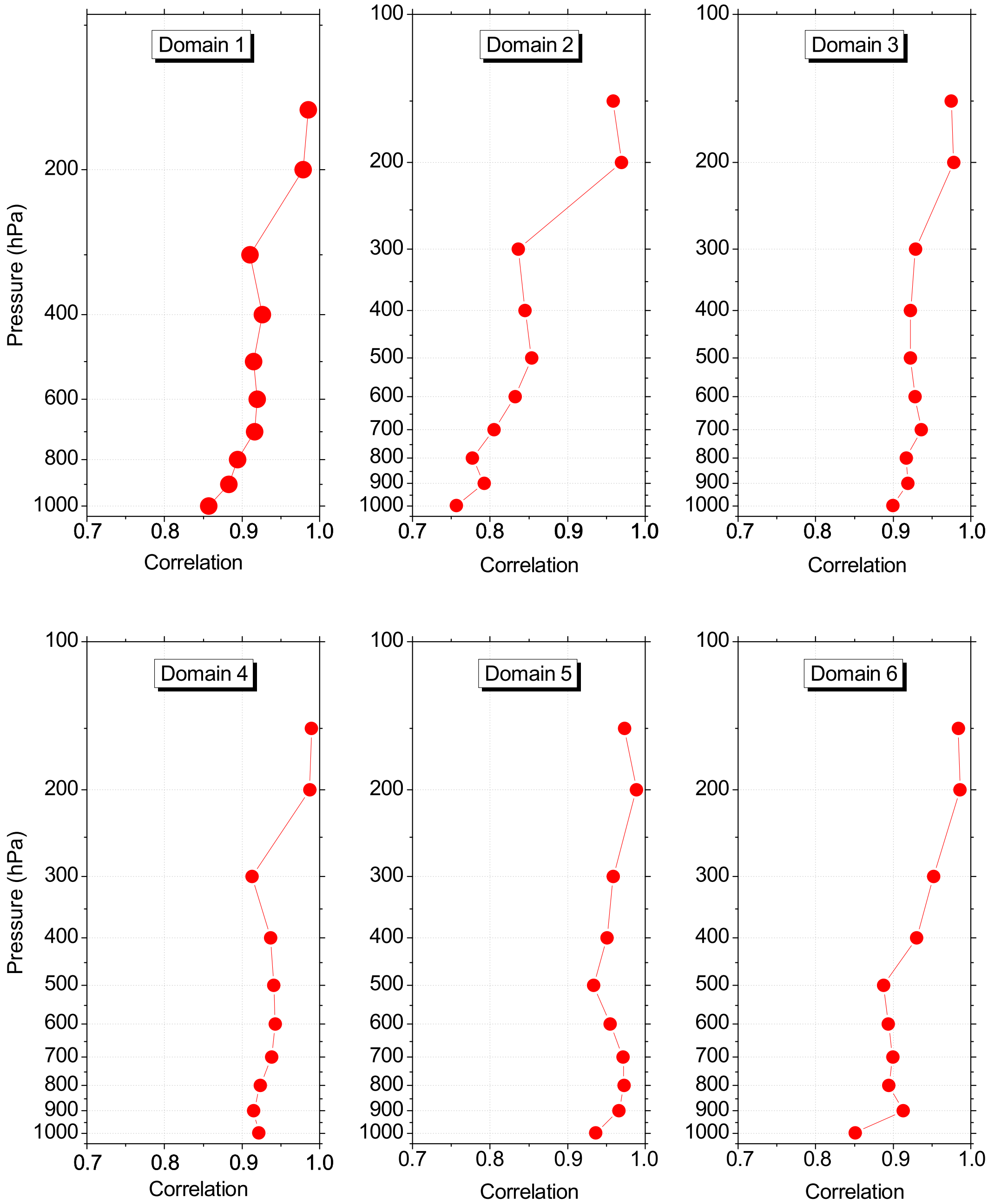


Fig. 15. Same as Fig. 13 but for the correlation coefficient between both analyses (TOTCOL_ANALYSES and PROFILE_ANALYSES).

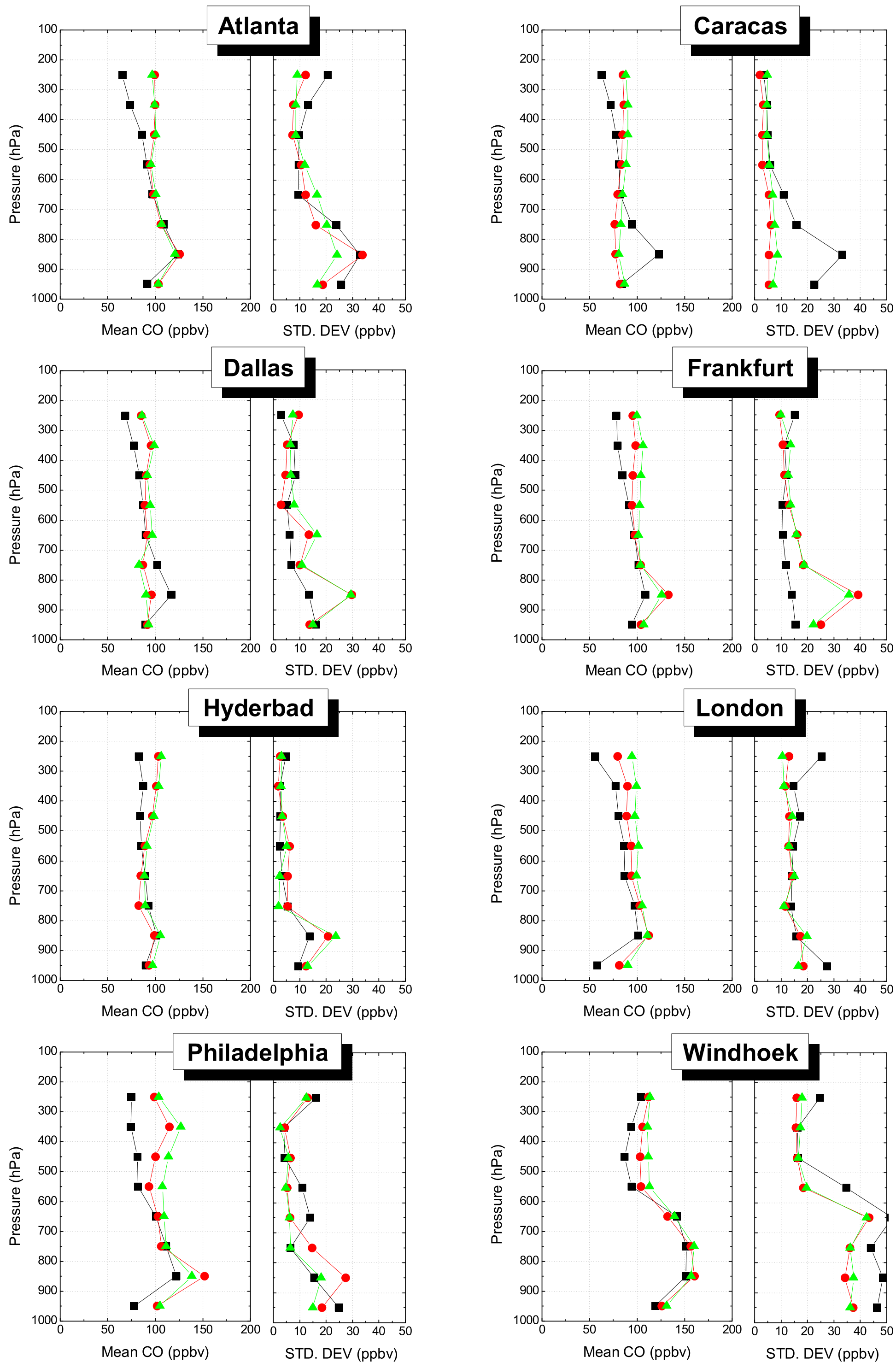


Fig. 16. Comparaison of both analyses with MOZAIC in terms of vertical profiles. **Black**: MOZAIC ; **Red**: analyses of MOPITT V3 profiles (taking into account averaging kernels, covariance error matrices and a priori profile) ; **Green**: analyses of MOPITT V3 total columns for which the error specification was done following the χ^2 test (see sec. 3.)