Impact of Spaceborne Observations on Tropospheric Composition Analysis and Forecast (ISOTROP)

ESA Contract No: 4000105743

OSSE production and analysis (S5 CO OSSE)

Version 2.0, final
November 2015,
prepared by Jean-Luc Attié,William Lahoz, Rachid Abida, Laaziz El Amraoui, and Philippe Ricaud

ESA Contract No: **4000105743**	SUBJECT: **OSSE production and analysis (S5 CO OSSE)** TITLE**:** **Impact of Spaceborne Observations on Tropospheric Composition Analysis and Forecast (ISOTROP)**	CONTRACTOR: **KNMI**
ESA CR( )No:	No. of Volumes: **1** This is Volume No: **1**	CONTRACTOR'S REFERENCE:

ABSTRACT:
We describe OSSEs with the MOCAGE-PALM and LOTOS-EUROS data assimilation systems with the assimilation activated for target areas and periods, summer 2003 and winter 2003-2004. We assimilate the CO satellite data (LEO), data produced in WP2 and WP3. We assess the performance of the S5 CO OSSEs using standard diagnostics as described in the technical part of the proposal.

The work described in this report was done under ESA Contract. Responsibility for the contents resides in the author or organisation that prepared it.

Names of authors:
Jean-Luc Attié (CNRS-GAME)
Rachid Abida (CNRS-GAME)
Laaziz El Amraoui (CNRS-GAME)
Philippe Ricaud (CNRS GAME)
William Lahoz (NILU)

NAME OF ESA STUDY MANAGER: **Dirk Schuettemeyer** DIV: **EOP SMS** DIRECTORATE:	ESA BUDGET HEADING:

Table of Contents

I Introduction

Over the last decade, the capabilities of satellite instruments for sensing the lower troposphere have strongly improved, and opened the way for monitoring and better understanding atmospheric pollution processes. From their vantage in space, these instruments provide global measurements of many pollutants, their trans-boundary transport, and complement ground-based air quality stations. They have the advantage of providing observations with global coverage and homogeneous quality. However, they also have drawbacks such as their limited spatial and temporal resolution.

Synergistic use of satellite data with ground-based and airborne measurements, assimilated or not into chemistry transport models, has contributed to an improved understanding and forecasts of tropospheric chemistry and dynamics. As part of an integrated observing strategy, satellite measurements provide a novel view on global air quality (AQ). The challenge for future space-borne missions will be to directly assess local scales of processes contributing to atmospheric pollution, and facilitate the use of remote sensing information for improving local-and regional-scale air quality analyses and forecasts.

Copernicus, previously GMES (Global Monitoring for Environment and Security), is the European Programme for the establishment of a European capacity for Earth Observation. The ultimate aim of the Copernicus Atmosphere Service is to provide consistent information on atmospheric variables in support of European policies regarding sustainable development and global governance of the environment. Services cover: AQ; climate change/forcing; stratospheric ozone; and solar radiation. It relies largely on data from satellites observing the Earth.

To ensure the operational provision of Earth-observation data, the Copernicus Space component includes a series of space-borne missions, developed and managed by the European Space Agency (ESA) specifically for Copernicus. Among them, there are three missions, which address atmospheric composition: Sentinel-5 and Sentinel-5 Precursor (S-5, S-5P) from a low earth orbit (LEO), and Sentinel-4 from a geostationary orbit (GEO). The goal of S-4 is to monitor key pollutants and aerosols over Europe at high spatial resolution with very short revisit time (hourly). The goal of S-5 and S-5P is to provide daily measurements at the global scale and high spatial resolution for air pollution, climate related trace gases and aerosols.

There is a need to assess the capabilities of satellite observing satellite systems measuring in the in Ultra-Violet (UV), visible, Near Infrared (NIR), and Short-Wave Infrared (SWIR) at nadir for tropospheric composition monitoring and forecasts, e.g., for upcoming ESA missions. The technique of Observing System Simulation Experiments (OSSEs) provides a way of efficiently demonstrating these capabilities and quantifying the benefit provided by these satellite missions.

Therefore, we conduct a regional-scale Observing System Simulation Experiment (OSSE) over Europe to explore the impact of S-5P carbon monoxide (CO) measurements on lowermost tropospheric air pollution analyses, with a focus on CO boundary layer concentrations. This report provides an important contribution to a forthcoming publication by Abida et al (2015).

We focus on two periods during the year 2003-2004. The first part of this report concerns northern summer of 2003 (June-July-August, JJA), and the second part northern winter 2003-2004 (November-December-January, NDJ).

During summer 2003, Europe experienced a severe heat wave episode associated with extremely hot and dry weather conditions. The long lasting blocking meteorological conditions significantly contributed to the accumulation of pollutants in the planetary boundary layer (PBL) owing to the extended residence time of the air parcels in the PBL (Solberg et al 2008). The spatial distribution of the enhanced levels of carbon monoxide and ozone was much more widespread in that summer than in previous ones (Ordoñez et al 2010, Lee et al. 2006). Furthermore, these exceptional weather conditions resulted in several extreme wildfire episodes over the Iberian Peninsula and the Mediterranean coast (Barbosa et al., 2004). Tressol et al. (2008) point out that between 6 and 10 August 2003 the contribution of biomass burning to the measured CO levels in the lowermost troposphere reached 35%, which is to be compared to the European anthropogenic emissions which contributed 30% to these CO levels. Consequently, the three-month period 1st June - 31st August 2003 encompasses both extreme and normal conditions, which allows us to look at the full range of pollution levels occurring in a summer season over Europe.

In addition, we focus on three other months during northern winter 2003-2004 (November, December 2003 and January 2004) to study differences between summer and winter. Differently to the summer, which experienced few cloudy days, the winter experienced general cloudiness. This winter period allowed us to study the impact of cloudy pixels in the OSSE set-up. For this, we performed two OSSEs, one including pixels with 100% cloudiness, and a second including only pixels with 10% cloudiness; we then compared results from these two OSSEs.

The study period considered ensures sampling of a wide range of meteorological situations, and provides a good compromise between run-time restrictions and obtaining enough statistical information. The OSSE study domain covers most of Europe (5W-35E, 35N-70N), and we perform the OSSE simulations at the spatial resolution of 0.2 degrees. With this resolution, we can track in detail long-range transport plumes of CO.

II S5P Carbon Monoxide

Space-borne instruments on Low-Earth Orbit (LEO), such as MOPITT (Measurements Of Pollution In The Troposphere) and SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric ChartographY), operating respectively in the thermal infrared (TIR) and short-wave infrared (SWIR), have already demonstrated the potential of remote-sensing from space to determine the CO distributions and its main emission sources at the global scale. However, owing to their limited revisit time, and their coarse spatial resolution, a LEO cannot represent regional and local aspects of air quality. Accordingly, to make further substantial contributions to atmospheric chemistry and air quality applications, in particular at regional and local scales, requirements for monitoring CO from space have to be more rigorous. S-5P will respond fully to this challenging need, by providing unprecedented high spatial resolution and improved sensitivity in the PBL, allowing resolution of CO emission sources at finer scales than hitherto.

The S-5P is the ESA pre-operational mission designed to bridge the gap between the end of OMI (Ozone Monitoring Instrument) and SCIAMACHY exploitation, and

the S-5 mission planned for the time period 2020 and beyond (Veefkind et al., 2012). Scheduled launch of S-5P is in late 2015 with a 7 years design lifetime. The S-5P will fly in an early afternoon sun-synchronous LEO with an Equator crossing mean local solar time of 13:30, at which it will pick up a significant pollution signal. This means that S-5P data will contain significantly more of the information needed for air quality forecasts for the next day. In contrast, GOME-2 (Global Ozone Monitoring Experiment 2) data at the local time of 9:30 has a lower predictive value.

S-5P will be a single payload in the TROPOspheric Monitoring Instrument (TROPOMI) jointly developed by The Netherlands and ESA (Veefkind et al., 2012). TROPOMI has heritage from both the OMI and SCIAMACHY missions. The instrument will measure the UV-visible wavelength range (270-500 nm), the near infrared (675-775 nm) and the shortwave infrared (2305-2385 nm). It will deliver a key set of gas and aerosol data products for air quality and climate applications. To enable the sounding of the lower atmosphere at finer scales, TROPOMI has an unprecedented spatial resolution of 7x7 km^2 at nadir. This relatively high spatial resolution is mandatory for AQ applications at local to regional scales. It allows resolution of emission sources accurately and provides an acceptable fraction of spectra without cloud contamination. Furthermore, TROPOMI will feature a wide swath of 2600 km to allow daily global coverage. The S-5P's improved radiometric sensitivity will allow measurements at low albedo, thereby helping to track smaller pollution events and improving the accuracy of air quality assessments and forecasts. The CO product will be available at relatively high spatial and temporal sampling, with a single measurement having an uncertainty of 10% at most. In contrast, one has to average SCIAMACHY data in time (roughly one month) and space (5 degrees by 5 degrees) to obtain realistic CO distributions at comparable accuracy (Gali et al., 2012). The use of S-5P CO measurements with inverse modelling will allow improved quantification of biomass burning emissions and their spatial distribution. Furthermore, simultaneous measurements of CO and, e.g., NO_2 will provide additional information on wildfire episodes.

III Description of the OSSE components

Observing System Simulation Experiments (OSSEs) (Atlas 1997; Masutani et al. 2010a, b) are designed to study the added value of simulated data in a state-of-the-art model using data assimilation (e.g., data from satellite platforms, ground-based networks). In an OSSE, we calculate observations and their associated errors from a representation of reality (the "nature run" or "NR"), and provided these to a data assimilation system to produce estimates of NR states. Afterward, these estimates (analyses or, ideally, forecasts) are compared to the nature run to evaluate the added value of the simulated data compared to the control run, CR (in this case a free model run) and the NR. OSSEs are widely used in the meteorological community for assessing the usefulness of new meteorological satellite data (e.g., Lahoz et al. 2005; Stoffelen et al. 2006); Masutani et al. 2010a, b). However, there are only a few studies concerning OSSEs for AQ related applications (Edwards et al., 2009; Timmermans et al., 2009; Claeyman et al., 2011b; Zoogman et al., 2011a, b; Yumimoto 2013; Zoogman et al., 2013a, 2014). Recently, Timmermans et al. (2015) set out the basis on how to set up an air-quality OSSE using satellite data. To summarize, a series of AQ OSSEs have demonstrated the major benefits that could accrue from a GEO OSSE for AQ monitoring and forecasts. In this report, we present a regional chemical OSSE framework conducted to investigate the impact of S-5P observations on surface CO.

We present the description of the OSSE scheme using the NR, the CR and the Assimilation Run (AR) calculated from the different models in Figure 1.

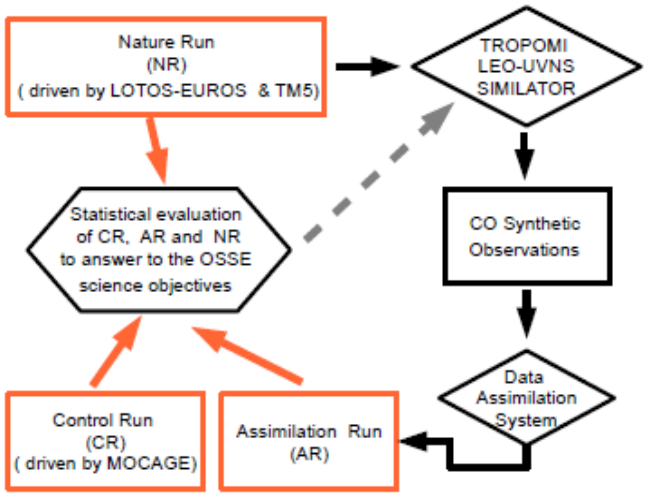

Figure 1: *Diagram of the OSSE components.*

In table 1, we summarize the list of runs which has been done for this OSSE study.

Table 1 : List of runs for MOCAGE in the domain MACC (15W-35E, 35N-70N), domain for the fire episode defined in Fig. 14, model resolution, species included and synthetic observations assimilated.

Run ID	Run	Domain	Resolution	Species	ASSIMILATION	
					Ground	Satellite
RREC	Reference	MACC	0.2°x0.2°	CO	no	no
RRFC		Fire episode				
ORELC100	OSSE,LEO	MACC	0.2°x0.2°	CO	no	LE0/S5P CO
OR-FLC100		Fire episode				
ORELC10	OSSE,LEO CF < 10%	MACC	0.2°x0.2°	CO	no	LEO/S5P CO

III.1 The Nature Run

A key consideration when designing an effective OSSE is to set up the Nature Run (Figure 1), which defines the true atmospheric chemical composition state from which we evaluate the analyses and/or forecasts that use simulated observations. The Nature Run is a long, free-running forecast evolving continuously in a dynamically consistent way (Masutani et al. 2010b). For this study, the Nature Run (NR) consists of two high-resolution free model simulations performed by the regional LOTOS-EUROS air quality model (Schaap et al. 2008) at a resolution of about 7km nested into the global-chemistry transport model TM5 (Huijnen et al. 2010), with a zoom domain over Europe at 1x1 degrees resolution. TM5 has 34 layers with a

model top at 0.1 hPa. The design of the LOTOS-EUROS model allows description of air pollution in the lowermost troposphere. It has four vertical layers following the dynamic mixing layer approach. The first layer is a fixed surface layer of 25 metres thickness, the second layer (boundary layer) follows the mixing layer height, and there are two reservoir layers up to 3.5 km. The LOTOS-EUROS implicitly assumes a well-mixed boundary layer, so constituent concentrations remain constant up to the top of the planetary boundary layer, PBL. The meteorological data used as input for the LOTOS-EUROS model comes from the European Centre for Medium-Range Weather Forecasts (ECMWF). We prescribe surface anthropogenic emissions using the TNO-MACC-II emission database; fire emissions are from the MACC global fire assimilation system (GFAS v1).

The model runs include an appropriate spin-up period of three months. We obtain the Nature Run by combining the LOTOS-EUROS CO profiles from the surface to 3.5 km with the TM5 results from 3.5 km to the top of the atmosphere, and archive data hourly.

The most difficult step in designing any effective OSSE is to demonstrate that the Nature Run exhibits the same statistical behaviour as the real atmosphere in every aspect relevant to the observing system under study (Masutani et al. 2010b). LOTOS-EUROS has been extensively verified with European data and has participated frequently in international model comparisons, mainly for ozone and particulate matter (van Loon et al. 2007; Hass et al. 2003; Cuvelier et al. 2007; Stern et al., 2008).

To evaluate the NR, we compare surface CO data to available ground-based CO measurements over Europe during the northern summer of 2003. We take ground-based stations from the Airbase data set, which, however, have the shortcoming of a lack CO data in rural areas. Furthermore, several ground-based data sets have very coarse resolution, and thus are not directly comparable to NR data. To overcome these shortcomings, we consider all types of ground-based stations. However, we discard stations having less than 75% of hourly data within a single month. As a result, we select 171 ground stations for the comparison. Figure 2 shows the locations of these stations. Note that, most of sites are located in polluted areas, where big emission sources of CO are present.

Figure 2 shows the time series of CO concentrations at the surface, as measured by the ground-based stations and simulated by the Nature Run (from LOTOS-EUROS). We form the measured time series by averaging spatially over all the sites. We form the modelled time-series similarly, but after interpolating the NR surface data to the station locations. We notice generally that the NR captures reasonably well the features of observed CO temporal variability, during the three phases characterizing the summer of 2003: before, during and after the heat wave. The correlation coefficient between these two time series is 0.71. This indicates a good representation of the diurnal cycle of CO in the NR. We also observe that the CO concentration levels in the NR are not as high as for the observed ones. Nonetheless, and most importantly, the simulated CO concentrations and those measured by ground stations are within the same range of values (globally between 200 and 400 μg m-3). The computed relative error of the NR with respect to the observed CO concentrations fluctuates around -10 % on average during normal conditions and reaches -20% within the heat wave period during 31 July-15 August 2003. This means that the NR reproduces the actual surface concentrations with an error range between -10% to -20%, indicating that for the period considered in the OSSE, the NR can be assumed to be reasonably representative of the “true atmosphere” over the European domain.

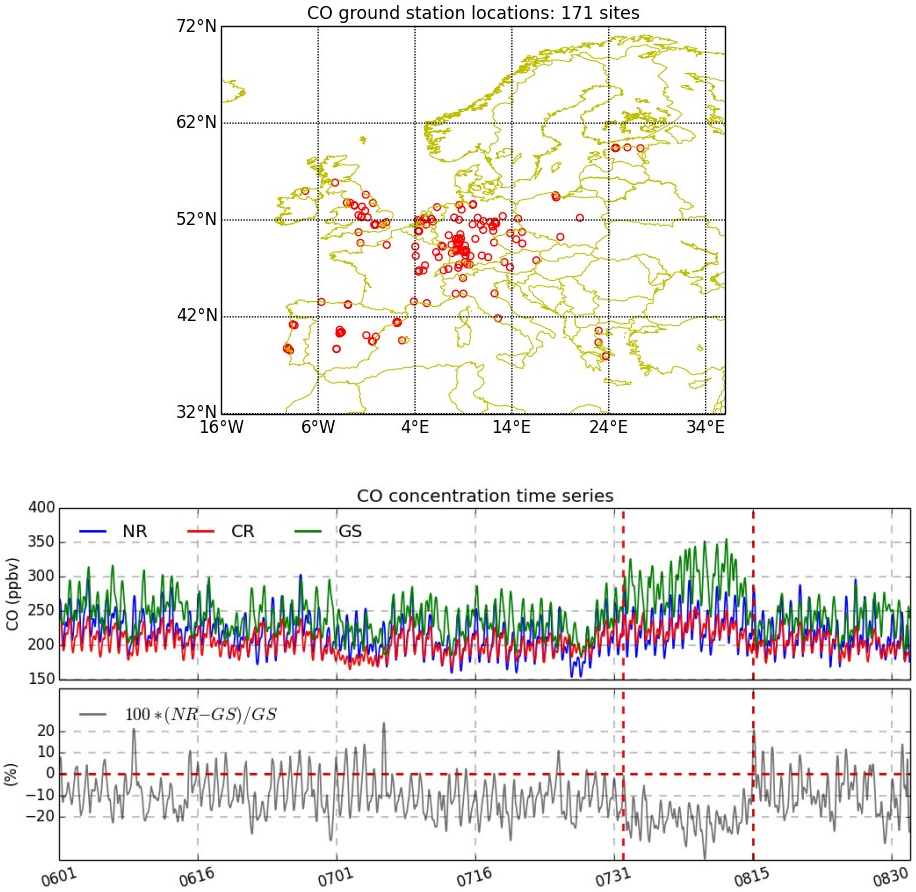

Figure 2: *Top panel: location of selected ground-based stations for CO measurements taken from the Airbase database during summer 2003 (1 June – 31 August). There are 171 sites with locations shown by circles. The labels show longitude, degrees (x-axis) by latitude, degrees (y-axis). Middle panel: simulated and measured time-series of CO concentrations in surface air from nature run (blue line), the control run (red line) and from the selected 171 Airbase sites (green line). We form the CO time-series by averaging the hourly data representative of the 171 sites. The labels show time in MMDD format (x-axis) by CO concentration, parts per billion by volume, ppbv (y-axis). Bottom panel: The gray curve shows the relative error of the nature run with respect to the observations, defined as NR value less ground station value divided by the ground station value and multiplied by 100. The labels show time in MMDD format (x-axis) by relative error, percent (y-axis). The vertical red dashed lines in the middle and bottom panels delineate the 2003 European heat wave period (31 July – 15 August).*

III.2 The Control Run

The second component in the design of an OSSE, is the Control Run (CR) (see Figure 1). We use a state-of-the-art modelling system, which includes all observational data representing current operational observational data. An important requirement for an effective OSSE is generation of the Control Run with a model different to the one used to construct the Nature Run. This avoids the identical twin problem, where we use the same model to produce the Nature Run and perform the assimilation experiments. If the model from which we extract hypothetical observations is the same as the assimilating model, the OSSE results will show unrealistic observation impact and overly optimistic forecast skill (Arnold and Dey 1986; Stoffelen et al. 2006). Consequently, by using two independent models the OSSE will more realistically simulate the assimilation of real observations. For OSSE studies in ISOTROP, we use the MOCAGE model to generate the Control Run. We do not assimilate operationally the CO observations from ground-based stations. Hence, the Control Run in our study is a free model run, i.e., a forecast for the length of the OSSE period.

The MOCAGE model is a three-dimensional CTM developed at Météo France (Peuch et al., 1999) providing the evolution of the atmospheric composition in accordance with dynamical, physical and chemical processes. It has a number of configurations with different domains and grid resolutions, as well as various chemical and physical parametrization packages. We currently use the MOCAGE model for several applications: e.g., Météo-France operational chemical weather forecasts (Dufour et al., 2005), Monitoring Atmospheric Composition and Climate (MACC) services (http://www.gmes-atmosphere.eu), and studies about climate trends of atmospheric composition (Teyssèdre et al., 2007). There is a validation study of the MOCAGE model using a large number of measurements during the Intercontinental Transport of Ozone and Precursors (ICARTT/ITOP) campaign (Bousserez et al., 2007).

This OSSE study uses a two-way nesting configuration to generate the Control Run: a global grid with a horizontal resolution of 2x2 degrees and a regional grid (5W-35E, 35N-70N) with a horizontal resolution of 0.2x0.2 degrees. MOCAGE includes 47 sigma-hybrid vertical levels from the surface up to 5 hPa. The vertical resolution is 40 to 400 m in the boundary layer (7 levels) and about 800 m near the tropopause and in the lower stratosphere. The chemical scheme used is RACMOBUS, which combines the stratospheric scheme REPROBUS (Lefèvre et al., 1994) and the tropospheric scheme RACM (Stockwell et al., 1997). RACMOBUS includes 119 individual species, among which 89 are prognostic variables, and includes 372 chemical reactions.

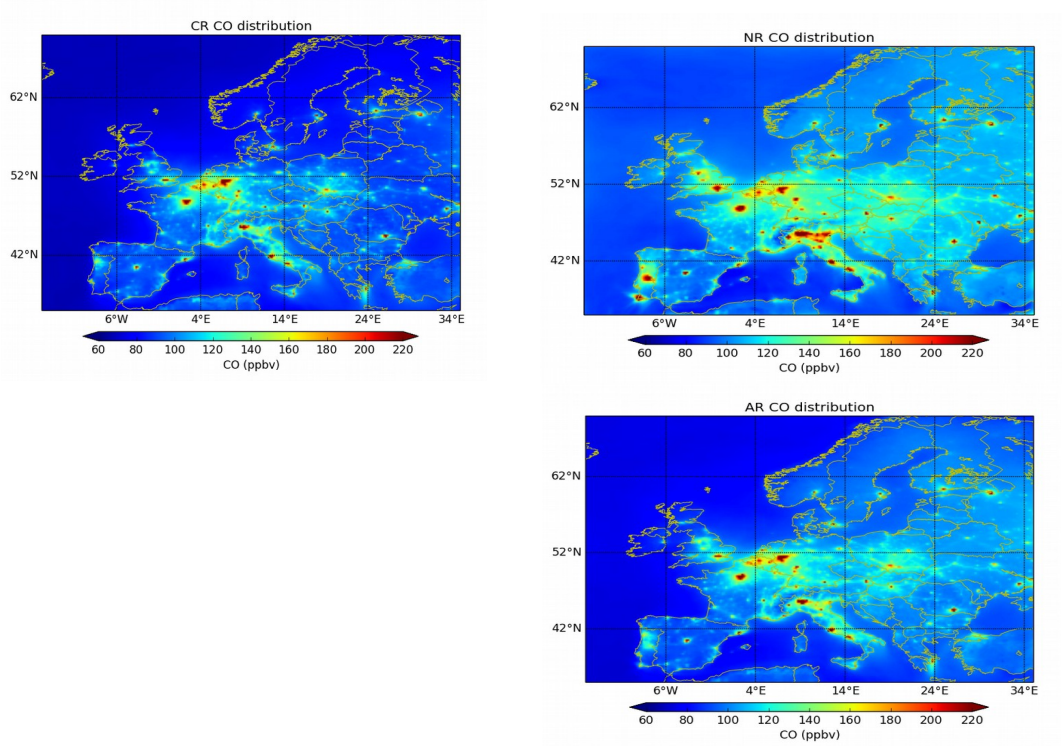

Figure 3: *Distribution of CO surface concentrations, units ppbv, averaged for the period 1 June – 31 August 2003. Top left panel: the control run (CR) from MOCAGE; right top panel: the nature run (NR) from LOTOS-EUROS; bottom panel: the assimilation run (AR) from MOCAGE obtained after assimilating the S-5P CO total column simulated data sampled from the NR. In all panels, the labels show longitude, degrees (x-axis) and latitude, degrees (y-axis). Red/blue colours indicate relatively high/low values of the CO surface concentrations.*

In contrast to the Nature Run, we use the ARPEGE analysis (Courtier et al., 1991) to force the simulation of the Control Run every 3 hours. We prescribe the surface anthropogenic emissions using the MACC-I emission database, and fire emissions from the GFED-v3 inventory. The Control Run includes a spin-up period of three months.

III.3 The Assimilation run

We assimilate simulated S5P CO derived from LOTOS-EUROS NRs into the MOCAGE CTM at the 0.2° scale using the PREV'AIR extended domain.The assimilation system used in this study is MOCAGE-PALM (e.g., El Amraoui et al., 2008a) developed jointly by Météo-France and CERFACS (Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique) in the framework of the ASSET European project (Lahoz et al., 2007b). The assimilation module used in this study is PALM (Projet d'Assimilation par Logiciel Multiméthode), a modular and flexible software, which consists of elementary components that exchange data (Lagarde et al., 2001). It manages the dynamic launching of the coupled components (forecast model, algebra operators and input/output of observational data) and the parallel data exchanges. We use the assimilation system MOCAGE-PALM to assess the quality of satellite ozone measurements (Massart et al., 2007). It is also useful for overcoming possible deficiencies in the model. In this context, we use its assimilation product for many atmospheric studies in relation to ozone loss in the Arctic vortex (El Amraoui et al., 2008a), exchange between the tropics and mid-

latitudes (Bencherif et al., 2007), stratosphere-troposphere exchange (Semane et al., 2007), and exchange between the polar vortex and mid-latitudes (El Amraoui et al., 2008b).

For all the OSSEs (one for summer and two for winter), to speed up the process of assimilation, we used the 3D-Var version of PALM instead of its 3D-FGAT version. Delays in the ISOTROP project outside the control of CNRM, and changes in the operational assimilation system for MOCAGE during this delay, forced us to use the 3D-Var assimilation method. However, the assimilation is done over a domain (Europe) which is covered by the satellite in less than 1 hour. The assimilation time window for both the 3D-FGAT and 3D-VAR is 1 hour, which should theoretically give the same result over our considered domain. However, we may have minor numerical differences if comparing results from the two softwares, but all OSSEs performed use the same assimilation system, and we are interested in the relative performance of the different distributions (CR and AR), the OSSE results are meaningful. From this point of view, the use of either 3D-FGAT or 3D-VAR is not important.

Note that the 3D-Var system we use is univariate, i.e., the background errors of the chemical species assimilated (e.g., CO and ozone) are independent. Nevertheless, assimilation of observations from one species (e.g., CO) on another (e.g., ozone) can still take place through the photochemical equations in the model used in the assimilation system (in this case, MOCAGE), and would depend strongly on the nature of the photochemical scheme. However, this analysis would be time consuming, and would only provide a first order estimate of the impact of CO observations on the ozone distribution. A proper study would require use of a multivariate assimilation system, as used by Zoogman et al. (2014). This is not possible with the current configuration of the MOCAGE-PALM assimilation system.

Finally, factors that could limit the performance of the assimilation system, and thus affect the quality of the results, include the prescription of the background errors (the B-matrix introduced in section III-3.2), and the method used to distribute into a vertical profile the total column information assimilated. In the set-up of the B-Matrix we use standard methods (see below). To distribute the total column information in the vertical we use a method tested and proved satisfactory for the MOCAGE-PALM system (El Amraoui et al., 2014). Based on this, we expect the assimilation system to perform reasonably well. We confirm this with a series of standard diagnostic tests (see Lahoz et al. 2007a for a description) described in section IV-1.1.

In this OSSE study, MOCAGE provides the control run (CR) and by assimilating the simulated CO data from the NR, MOCAGE provides the assimilation run (AR) - see Figure 1. Figure 3 shows the mean fields of surface CO for the CR, the NR and the AR. One can see differences in the CO field over land between the CR and the AR. These come from the contribution of simulated S5P CO data from the NR.

3.1 Pre-processing of S5P-CO

This section discusses how we incorporate S-5P CO column observations into the MOCAGE forward model in order make the different OSSE simulations. S-5P will produce huge amount of data due to its wide swath, and the relatively high spatial resolution of about 7x7 km^2. A pre-processing step is necessary to reduce the data volume for data assimilation. For the scope of this study, we only consider pixels inside the OSSE simulation domain. Note that retrieval pixels in each single cross-track are essentially instantaneous measurements of CO. This has the advantage of

reducing the data volume burden. However, a single cross-track carrying a stronger CO signal over Europe could have more than 80000 valid retrieval pixels. Furthermore, each individual pixel is associated with an averaging kernel vector given at 34 vertical pressure levels, from the surface up to the top of the atmosphere. We show an example of averaging kernels at the surface in Figure 4 as well as the averaging kernels representative of retrievals with various cloud fractions included (less than 10%, greater than 30%, and greater than 80%). In addition, we discard data points with solar zenith angles larger than 80% or errors exceeding 20%. The retrieval over sea is noise-dominated. Thus, we only consider only partial CO columns above cloudy ocean scenes with cloud fraction more than 80% and cloud top heights between the surface and 650hPa. Finally, we apply a spatial weighted mean for binning the measurements into 0.2° by 0.2° grid box, which is the MOCAGE-PALM model resolution.

The weighted mean for pixels falling in the same model grid box is:

$$\tilde{c} = \frac{\sum_i w_i \, c_i}{\sum_i w_i}$$

where $\tilde{c}$ is the weighted average, c_i a single column measurement, and w_i ($=1/\sigma_i^2$) is the inverse of the variance corresponding to measurement c_i, and is the weight assigned to this single measurement. The inverse of the variance associated with the weighted average is

$$\frac{1}{\tilde{\sigma}^2} = \sum_i w_i$$

The spatial binning not only reduces considerably the data volume but also results in an improved spatial representativeness of the CO measurements by reducing the random error of each data pixel.

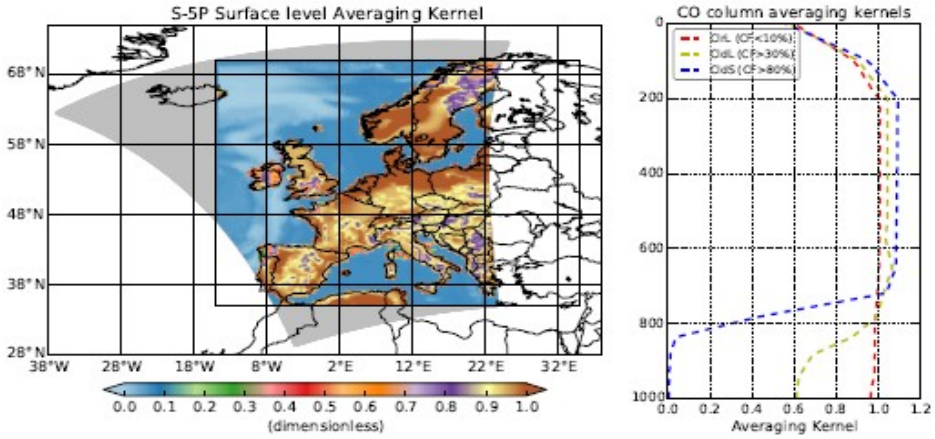

Figure 4: *Left panel: S-5P CO averaging kernel values at the surface. Labels are longitude, degrees (x-axis) by latitude, degrees (y-axis). Right panel: Averaging kernels for land pixels with cloud fraction less than 10% (dashed red lines); for land pixels with cloud fraction greater than 30% (dashed yellow lines); and for sea pixels with cloud fraction greater than 80% (dashed blue lines). The averaging kernels are for an average of the data shown on the swath for 1 June 2003 at 12:34 UTC. Labels are averaging kernel, in units of KK^{-1}, where K is degrees Kelvin (x-axis) by pressure level, hPa (y-axis).*

3.2 Assimilation parameters

One of the matrices to estimate in a common data assimilation system is the background error covariance matrix (B-matrix), which has a large impact on the 3D-Var analysis. This is why it is important to use as realistic as possible a form of B. In MOCAGE-PALM, the B-matrix formulation is based on the diffusion equation approach (Weaver and Courtier, 2001) and can be fully specified by means of a 3D standard deviation field (square root of the diagonal of B, in concentration units or as a percent of the background field) and 3D fields of horizontal (Lx and Ly) and vertical (Lz) local correlation length-scales. We can estimate the B statistics more efficiently using an ensemble method (Bannister 2008). This technique consists on feeding an ensemble of states through the data assimilation system to simulate the most important sources of error. However, this approach is time-consuming. For this study, we have used a simple parametrization for the B-matrix. Lx and Ly are taken as homogeneous and equal to 35 km. Lz is assumed constant and set to one vertical model grid point. As in Emili et al. (2014), the background standard deviation 3D field is parametrized as a vertically varying percentage of the background profile, which vertically decreases from 25%-15% in the troposphere to 10%-5% in the stratosphere. Its shape is the function depicted in Fig. 5. This makes the B-matrix flow-dependent, changing as the background field changes. We base these settings on several 1-day assimilation trials, and they ensure reasonable values of the statistical tests implemented (Fig. 6). The data assimilation procedure weights observations and model 1-h forecasts (from the last analysis point), and updates locations not coincident with the observations through correlation length-scales. Table 2 summarizes the parameters used for the assimilation experiments.

Table 2: Description of the configuration used in the assimilation system

	Description
Assimilation	3D var 1 hour window
Background standard deviation Background correlation zonal Length scale (Lz) Background correlation meridional length scale (Ly) Background correlation vertical length scale (Lx)	in % of the background field (vertically variable) constant 35 km constant 35 km one model grid point
S-5P CO observation errors	from retrieval product

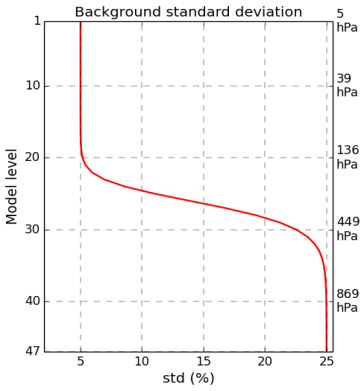

Figure 5: *Background error standard deviation (square root of the diagonal of B) given in % of the background profile, as a function of the model level. A higher uncertainty is set in the troposphere (20 %) than in the stratosphere (5 %).*

IV The summer 2003 period

IV-1 Evaluation of the assimilation run

IV-1.1 Consistency of the assimilation run

We have performed the OSSEs for the summer 2003 period including all pixels in the OSSE domain, regardless of whether they are cloudy or clear-sky. We tested the validity of this approach by performing two OSSEs, one including all pixels, the other just including clear-sky pixels. Comparison of the ARs from these two OSSEs indicated that the impact of including all pixels is small, and justifies the approach taken in this paper. The largest differences between the respective ARs in relation to the NR are 4% in regions over North Europe (North Sea and Scandinavia), with the AR for clear-sky pixels closer to the NR (not shown). We can explain these results by the fact the summer generally has low amounts of cloud.

To evaluate the assimilation run, we calculate the chi-squared (χ^2) diagnostic, and the Observation minus Analysis (OmA) and the Observation minus Forecast (OmF) differences. For the OmA and OmF diagnostics, we interpolate the observations to the model grid and at the time of midnight, where the analyses (A) and the short-term (6 hours) forecasts (F) are given. In addition, we calculate the difference between the observation and the simulation from the model run without assimilation (observation-minus-control run, hereafter OmC). We calculate the OmC diagnostic similarly to the OmA and OmF diagnostics.

Figure 6 shows the χ^2 time series and its associated auto-correlation function calculated over the three-month period. We note that the χ^2 diagnostic starts with a maximum of about 1.56, and takes values down to 0.75 over the whole OSSE three-month period, with a mean of 0.9. The χ^2 series is nearly stable with time since it exhibits relatively small variability, with a standard deviation of about 0.14. Furthermore, the auto-correlation of χ^2 quickly drops to zero, with no correlation after a time lag of 20 days. The calculation of the auto-correlation shows that the χ^2 statistic is uncorrelated after a certain time lag meaning that the mathematical expectation $E(\chi^2)$ is equal to the average of χ^2. We found $E(\chi^2)$= 0.90, which is close to the theoretical value of 1. This result indicates that the a priori error statistics slightly overestimate the actual error statistics. As a result, we infer that the observation and background error covariance matrix, B, and its observational counterpart, R, prescribed in our assimilation system, are reasonably well characterized.

To test whether the observations, forecast and analysis fields, and their associated errors, are consistent with each other, we calculate histograms of OmA , OmF and OmC (normalized by the observation error) over the three-month period. Figure 6 shows that OmA and OmF have Gaussian-like distributions centred at zero. OmF has a mean and standard deviation of 0.10 and 1.73, respectively, whereas OmA has nearly a zero mean and a standard deviation of 1.05. This indicates that an OmA histogram essentially centred at zero, and more peaked than the OmF histogram. We expect this result, since the analysis should be closer to the observations than the forecast. These results verify that the Gaussian error assumption in the observations and the forecast is valid.

Consequently, the self-consistency a posteriori diagnostics involving χ^2, OmA and OmF, indicate that the assimilation system is properly set up (for a review on self-consistency tests for evaluating data assimilation system, see Lahoz and Errera 2007; Lahoz et al., 2007a). Moreover, the shape of OmC histogram, which has a mean and standard deviation of 2.36 and 5.60, respectively, indicates the presence of a large bias between the S5-P observations and the control run (the model run without assimilation). We can reduce this bias in the assimilation experiments, as the analyses are significantly closer to the observations than to the simulations from the control run. This shows that the assimilation of S5-P CO observations has a substantial impact on the forecast and the analysis.

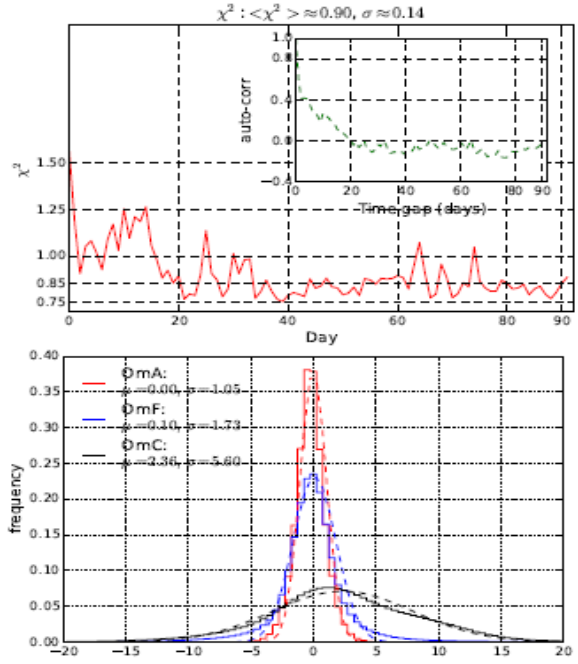

Figure 6: *Self-consistency tests. Top panel: time-series (red line) of* χ^2 *and its associated auto-correlation signal (green line). The labels show time, days (x-axis) and* χ^2 *value (y-axis)* for the χ^2 *plot, and time gap, days (x-axis) and auto-correlation (y-axis) for the auto-correlation plot. Bottom panel: histograms of Observations minus Analysis (OmA -red solid line), Observations minus Forecast (OmF -blue solid line), and Observations minus Control run (OmC -black solid line). We normalize the differences by the observation error. The dashed lines correspond to the Gaussian fits of the different histograms. The labels show the OmA, OmF or OmC differences (x-axis) and the frequency of occurrence of the differences (y-axis). We calculate the diagnostics OmA, OmF, and OmC over the period of 1 June – 31 August 2003.*

IV-1.2 Study of increments

To understand further the impact of the S-5P simulated CO total column at the surface, we calculate the analysis increment (δx) for a single analysis time at 14:00 UTC on 15 June 2003. We calculate this increment as the analysis minus the model first guess (1-h forecast). Notice that the analysis increment is a measure of the error in the 1-h forecast started from the previous analysis. It provides a quantitative diagnostic of the quality of the analysis (Fitzmaurice et al, 2008). **Figure 7 (top)** shows the spatial distribution of δx at the surface (model surface). First, one can see a spread of the impact of the simulated observations across large regions. This is essentially due to the fact S5-P has a wide swath allowing it to sample larger regions, and an enhanced measurement sensitivity in the PBL. The most substantial corrections are over land, where there are sufficient observations to have an impact. Over sea where there are few and relatively inaccurate observations, the analysis increment is negligible. Because of this, there will not be much difference between the model first guess and the analysis. Likewise, this is also true in the regions outside the satellite footprint. Furthermore, the model forecast (from the last analysis point) improves gradually as we introduce new simulated observations into the assimilation system. This explains why we observe also small correlations over some land regions, although the satellite samples them.

To provide further insight into the impact of S5-P CO measurements, we calculate latitude-height and longitude-height cross-sections at 48.8 N 2.6 E, which is near Paris. **Figure 7 (bottom left and right panels)** displays a zoom of the zonal and meridional vertical slices of the analysis increment around 48.8 N 2.6 E. We observe the presence of significant corrections in a deep layer and that they are larger at the surface, and exhibit a second maximum around 650 hPa. This vertical structure is primarily attributable to the forecast error standard deviation (given as a vertically varying fraction of the local CO mixing ratio) which is typically higher in the boundary layer (where the value of the S5P CO averaging kernel is close to 1). The shape of S-5P analysis increments exhibits a second peak around 650 hPa, but is similar to the shape of SCIAMACHY analysis increments, which also extend through a deep layer with maxima at the surface (Tangborn et al., 2009). The fact that these analysis increments stretch out over a deep layer is due to the TROPOMI/S-5P and SCIAMACHY averaging kernels, which are very similar and close to unity over cloud-free land.

Note that the increment fields shown in Fig. 7 show an example of the benefit of the assimilation system, namely redistributing information from the column to the mid troposphere and the surface. Because the averaging kernel shown in Fig. 5 shows uniform sensitivity throughout the troposphere, and there are no emissions for this period impacting the boundary layer, changes at the surface must include information from elsewhere in the tropospheric column. This will include the CO distribution in the mid troposphere, which mainly originates from long-range transport from areas outside the local regions associated with the increments (e.g., Spain and Northern France).

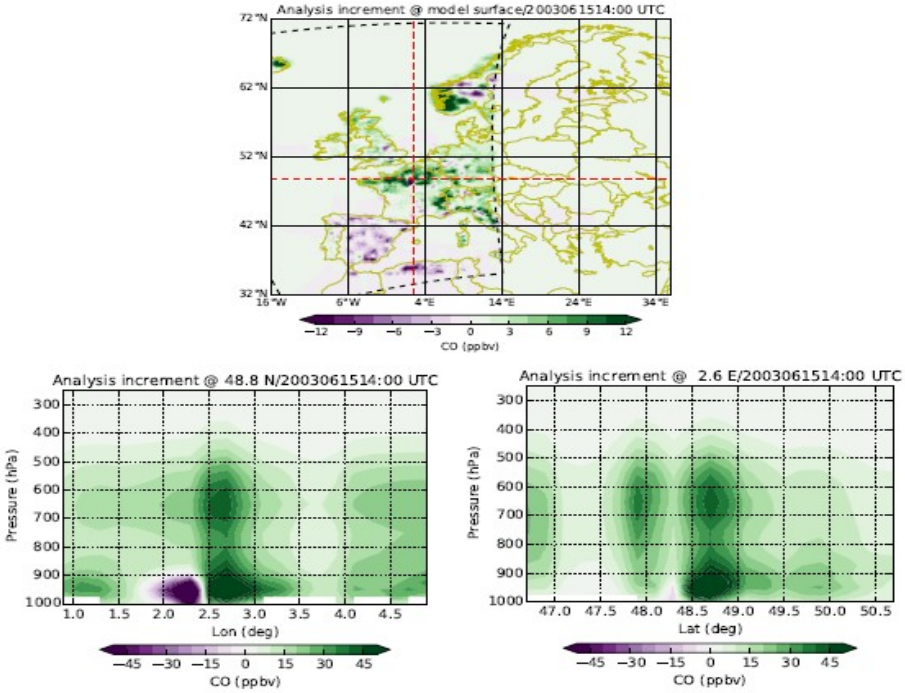

Figure 7 *: S-5P CO analysis increments, units of ppbv, at 14:00 UTC on 15 June 2003: Top panel: geographical distribution at the model surface. Red dashed lines show zonal and meridional vertical slices at 48°8 N, and 2°6 E, respectively. The black dashed line shows the S-5P cross-track at 13:12 UTC, clipped to fit the OSSE simulation domain. Note that we measure the S-5P CO observations at 13:12 UTC. The labels show longitude, degrees (x-axis) and latitude, degrees (y-axis). Left and right bottom panels show, respectively, the longitude-height and latitude-height cross-sections at a location near Paris. The labels for the bottom panels show longitude, degrees (x-axis, left panel), latitude, degrees (x-axis, right panel), and pressure, hPa (y-axis, both panels). Green/purple colours indicate positive/negative values in the increment fields.*

IV-2 Evaluation of the summer OSSE

IV-2.1 Statistical analysis

The main goal of this section is to provide a quantitative assessment of the added value of TROPOMO/S-5P CO total column measurements on the assimilation analysis at the surface. To achieve this, we perform a statistical analysis of the different runs during summer 2003.

We calculate the mean bias (MB, ppbv), its magnitude reduction (MBMR, ppbv), the root mean square error (RMSE, ppbv), its reduction rate (RMSERR, %), along with the mean absolute error (MAE, ppbv), and its reduction rate (MAERR,%). Finally, to measure the linear dependence between two data sets, and the portion of the true variability (variability in NR) reproduced by the CR or AR, we use the correlation coefficient (R). For a single model grid box, we define these statistical indicators with respect to the NR, the truth, as follows:

$$MB(X) = \frac{1}{N}\sum(X - NR) \qquad MBMR = |MB(CR)| - |MB(AR)|$$

$$RMSE(X) = \sqrt{\frac{1}{N}\sum(X - NR)^2} \qquad RMSERR = 100 \times \left(1 - \frac{RMSE(AR)}{RMSE(CR)}\right)$$

$$\mathrm{MAE(X)} = \frac{1}{N}\sum|\mathrm{X} - \mathrm{NR}|; \ \mathrm{MAERR} = 100 \times \left(1 - \frac{\mathrm{MAE(AR)}}{\mathrm{MAE(CR)}}\right),$$

$$R(X) = \frac{\Sigma(X - \overline{X})(NR - \overline{NR})}{\sqrt{\Sigma(X - \overline{X})^2 \, \Sigma(NR - \overline{NR})^2}}$$

where X denotes the CR or the AR; N is the number of data samples; and the overbar symbol represents the arithmetic mean operator. The MB metric gives the average value by which the CR, or the AR differs from the NR over the entire dataset. The RMSE and MAE metrics are more appropriate to measure the overall error distribution in CR or AR.

IV-2.2 Results

Figure 8 presents the zonal and meridional means of the difference between the CR and the AR averaged over the summer 2003 (1 June – 31 August). We also plot the confidence interval representing the areas where the AR is not significantly different to the CR at the 99% confidence limit (highlighted in gray). These two panels in Fig. 8 show that there is benefit from the S-5P CO total column data over the first few levels of the troposphere, i.e., the lowermost troposphere. Between the surface and 800 hPa, a negative peak is present in the zonal difference field (over Scandinavia), and in the meridional difference field (over Eastern Europe). Note that the zonal field shows two areas, one with positive values and the other with negative values representing a CR greater than the AR and a CR smaller than the AR, respectively. The positive peak, at a slightly higher level (i.e., lower pressure) than the negative peak, is representative of the Mediterranean Sea, whereas the negative peak is more representative of the land areas (Scandinavia and Eastern Europe). Figure 8 indicates that the S-5P CO corrects the model in the lower troposphere with a larger impact over land and with a less large impact in the PBL. This is consistent with the behaviour of the analysis increments shown in Fig. 7.

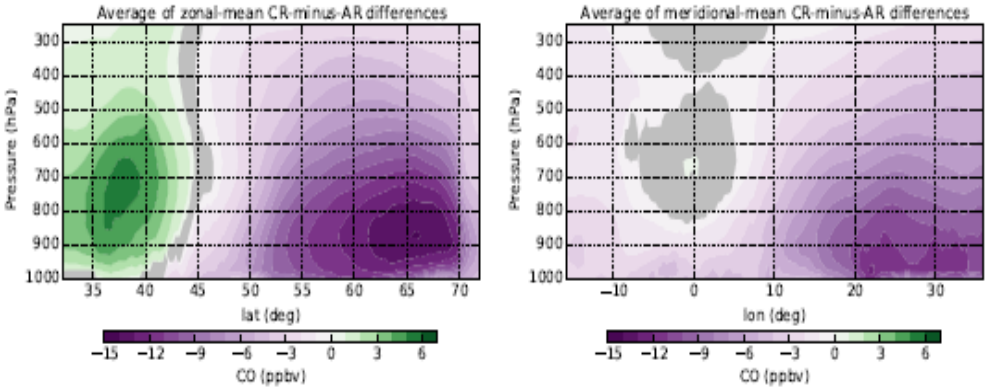

Figure 8: *Zonal (left panel) and meridional (right panel) slices of the difference between the CR and AR CO fields, units of ppbv, averaged over the summer period (1 June – 31 August 2003). The areas highlighted in gray indicate where the AR is not significantly different to the CR at the 99% confidence level. The labels in the left panel are latitude, degrees (x-axis) and pressure, hPa (y-axis). The labels in the right panel are longitude, degrees (x-axis) and pressure, hPa (y-axis). Green/purple colours indicate positive/negative values in the difference fields.*

Figure 9 presents the reduction of the mean biases (MBMR) and the corresponding absolute bias (CR-NR and AR-NR) at the surface averaged over the three-month period of summer. The MBMR indicates the geographical areas where the simulated S5P CO data has the most impact. The MBMR (reddish colour) shows that almost everywhere in the domain, the AR is closer to the NR compared to the difference between the CR and the NR. This indicates that the simulated S5P CO data have a benefit at the surface and above all over the land.

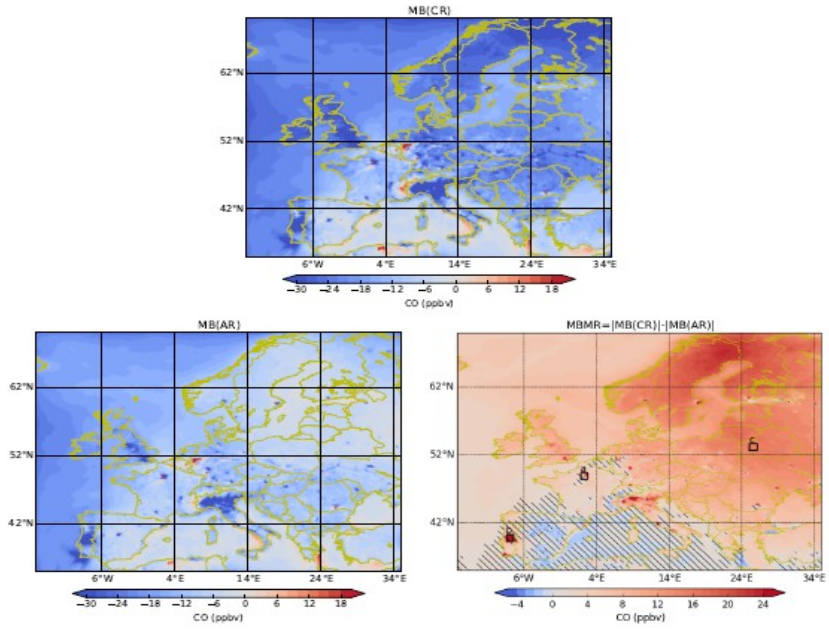

Figure 9: *Mean bias reduction at the surface for CO, units of ppbv: Top panel shows the CR mean bias (CR-NR). Left bottom panel shows the AR mean bias (AR-NR). Right bottom panel shows the mean bias magnitude reduction (absolute value of the mean bias for CR less the absolute value of the mean bias for AR). We average the data over summer 2003 (1 June – 31 August). The labels show longitude, degrees (x-axis) and latitude, degrees (y-*

axis). The hatched area in the right bottom panel shows where the mean bias plotted in this panel (MBMR) is not statistically significant at the 99% confidence level. The three squares in the right bottom panel represent locations for the three time-series shown in Fig. 13. Red/blue colours indicate positive/negative values in the MB/MBMR.

Figure 10 presents the mean absolute error (MAE) between the CR and the NR, and its corresponding reduction rate (MAERR). To quantify the added value of the simulated data, we remove the systematic error either in the MAE or in its corresponding reduction (MAERR). As expected, the added value is smaller when removing the systematic error (East of the domain), but the reduction of the MAE is still about 40% over land and in the regions where the CO sources are coarser and with less intensity (Scandinavia and Russia) for this period, suggesting a significant added value of the satellite data as a means to complement the ground-based network.

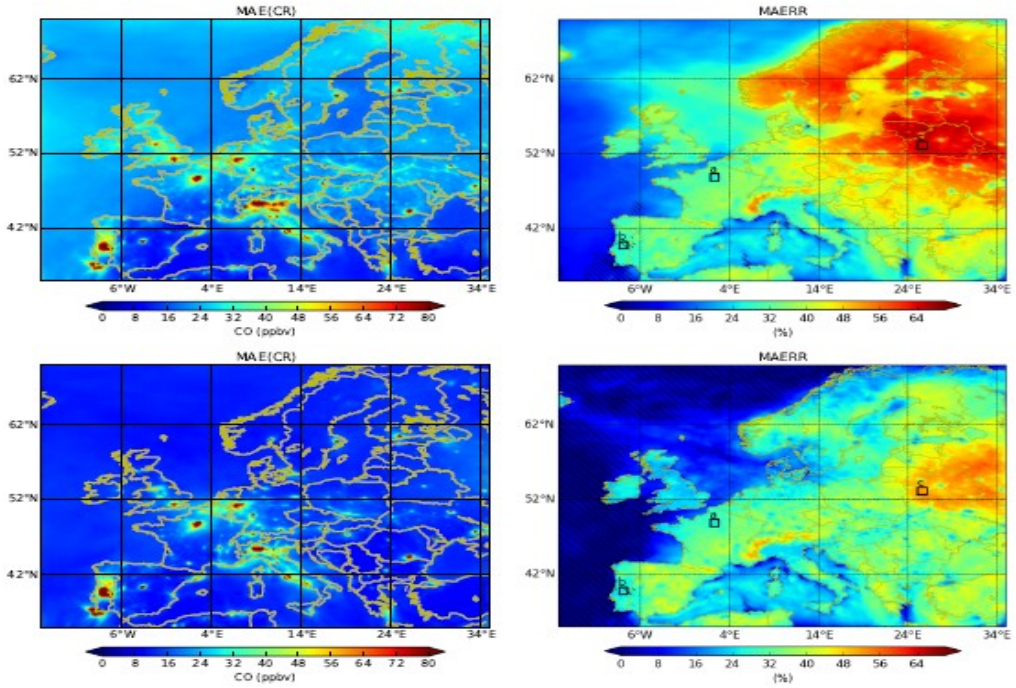

***Figure 10**: Top: Mean absolute error (MAE), units of ppbv, between CR and NR (left panel), and its corresponding reduction rate MAERR, in % (right panel) keeping the systematic error. Bottom: Same as top panel but after removing the systematic error from the MAE. The labels on each panel are longitude, degrees (x-axis) and latitude, degrees (y-axis). The three squares in the two left panels represent locations for the three time-series shown in Fig. 10. Red/blue colours indicate relatively high/low values in the MAE/MAERR.*

In the same way as above, we calculate the RMSE as well as the reduction of the RMSE (**Figure 11),** keeping the systematic error and removing the systematic error (RMSERR). One can see that the larger values of the reduction of RMSE (with and without the systematic error) are located in the same regions as the larger values of MAERR. There is reduction of about 40% in the RMSE in the East domain. By assimilating the S5P CO data, we reduce the variability between the AR and the NR compared to that between the CR and the NR.

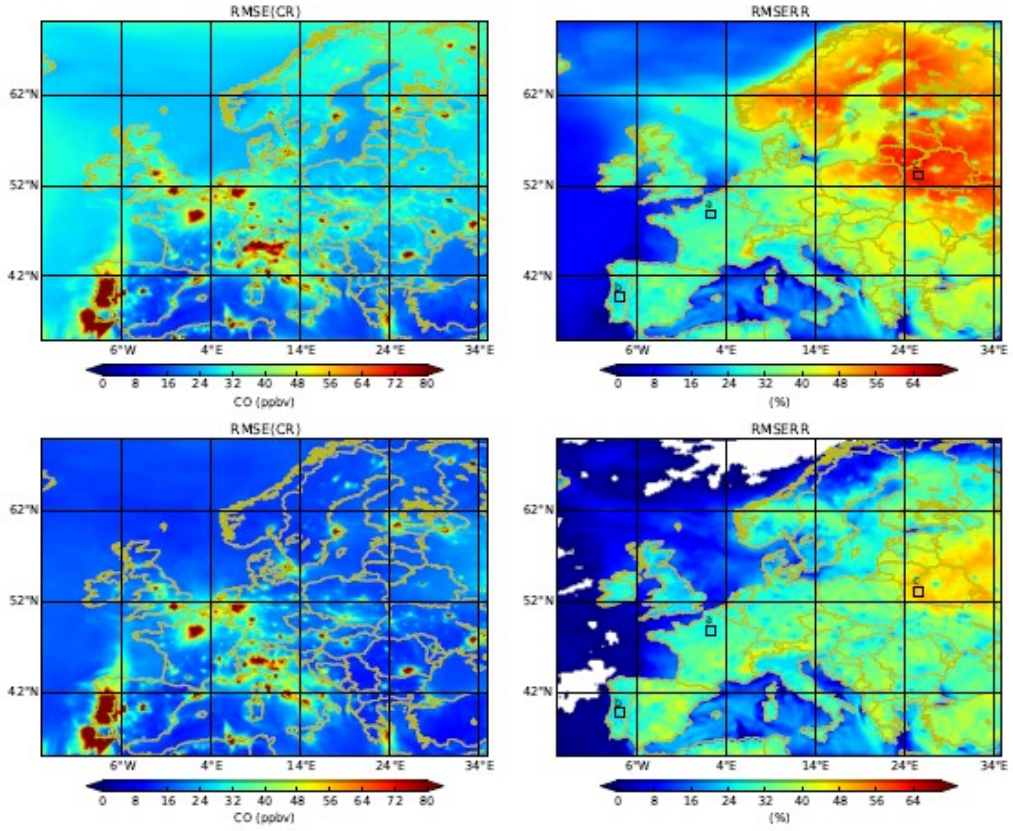

Figure 11: *Top: Root Mean Square Error (RMSE), units of ppbv, between CR and NR (left panel), and its corresponding reduction rate RMSERR, in % (right panel) keeping the systematic error. Bottom: Same as top panel but after removing the systematic error from the RMSE. The labels on each panel are longitude, degrees (x-axis) and latitude, degrees (y-axis). The three squares in the two left panels represent locations for the three time-series shown in Fig. 10. Red/blue colours indicate relatively high/low values in the RMSE/RMSERR.*

In **Figure 12**, we present the correlation between the CR and the NR, and the correlation coefficient between the AR and the NR at the surface for the three-month period. The figure shows that the AR is closer to the NR with correlation values reaching 0.9 over land

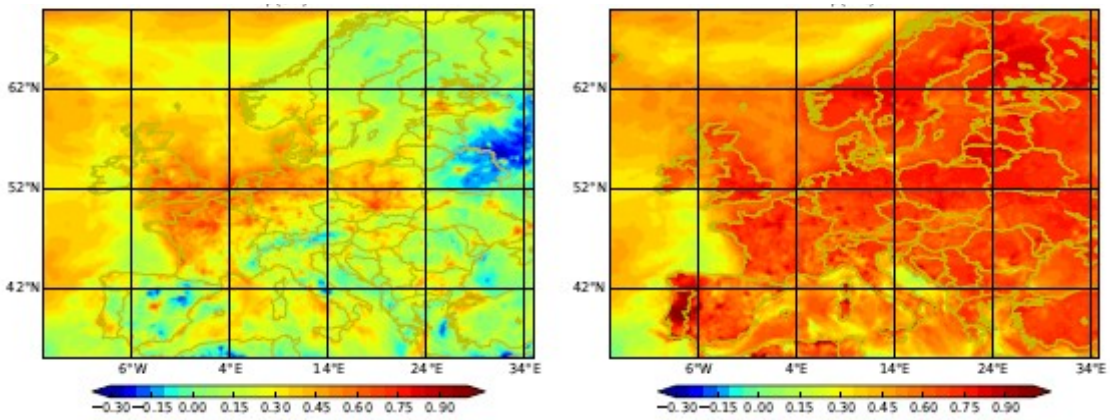

Figure 12: *Correlation coefficient (R) between the CR and the NR (left panel) and the AR and the NR (right panel) at the surface and for the summer period (1 June – 31 August). The labels are longitude, degrees (x-axis) and latitude, degrees (y-axis). Red/blue colours indicate positive/negative values of the correlation coefficient.*

Figure 13 shows an example of time-series from the NR, the CR and the AR over three areas of the study domain represented by the squares shown in Figs. 10 and 11 (right panels). (i) The Paris region (Fig. 13, top panel); (ii) a region over Portugal, where forest fires occurred during the summer (Fig. 13, middle panel); and (iii) an area in the Eastern part of the study domain, where the reduction of RMSE (i.e., RMSERR) was much larger than for other regions (Fig. 13, bottom panel). For all three areas, the AR is generally closer to the NR than the CR, showing the impact of the simulated observations. The biases between the AR and CR vs the NR are (in all cases, the difference is calculated as NR-X, where X is AR or CR): (i) Paris region, CR: 48 ppbv, AR: 38 ppbv; (ii) Portugal, CR: 101 ppbv, AR: 83 ppbv; (iii) Eastern part of domain: CR: 21 ppbv, AR: 5 ppbv.

Over Paris (top panel), the CR is already close to the NR and the impact of the S-5P CO simulated observations is small. Over Portugal (middle panel), the presence of fires is not seen in the CR (e.g., a maximum of CO at the beginning of the heat wave), as the fires were not taken into account in the CR as commonly happens in a real forecasting system. In contrast, over this specific location we see the impact of the fires on the CO concentrations in the AR with, however, much lower values than for the NR. During the fires, the CO concentrations in the AR over Portugal were larger than 500 ppbv, whereas the CR remained relatively unchanged with concentrations less than 200 ppbv. Over the Eastern part of the study area (bottom panel), the temporal variability is not high and magnitude of the bias between the CR and the NR is small, but it is removed in the AR.

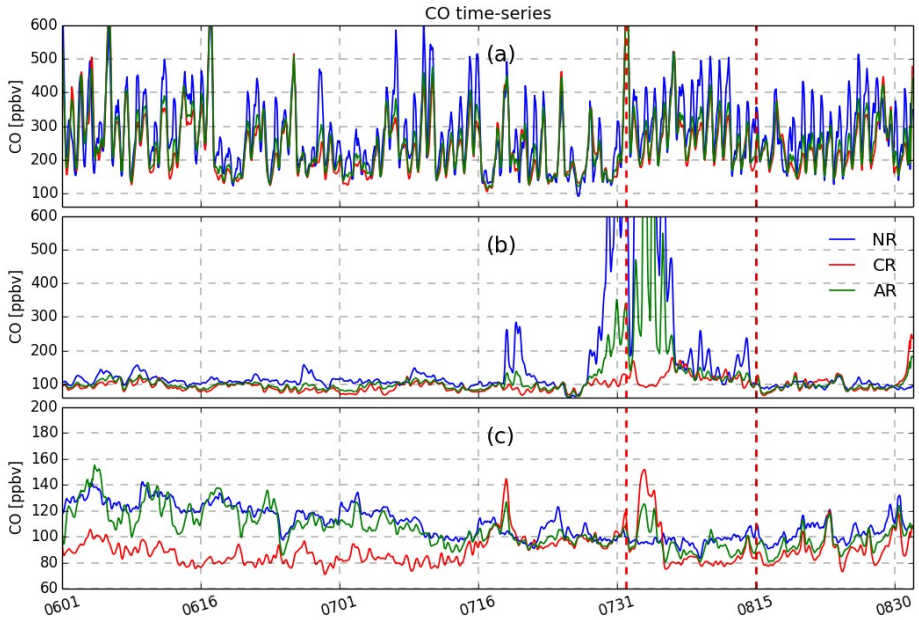

Figure 13: *Time-series for CO surface concentrations (1 June – 31 August) from NR (blue colour), CR (red colour) and AR (green colour) over three different locations represented by squares in Figs. 10-11. Top panel: area near Paris; middle panel: area over Portugal, where forest fires occurred; bottom panel: Eastern part of the study domain. The labels in the three panels are time, in format MMDD (x-axis) and CO concentration, ppbv (y-axis). The plus symbols at the top of each panel indicate availability of observations from the S-5P platform.*

To understand further the performance of the OSSE over the period of the Portugal forest fires, we repeated the OSSE without the default criterion to discard CO column observations with values larger than 75% of the MOCAGE value. This criterion is not appropriate to situations resulting in excessive values in the CO concentrations, as for forest fires. This second OSSE covered the period of the forest fires (25 July – 15 August). For this second OSSE, we compared the total column values and the surface values of the CO fields from the CR and the AR (Figs. 14-15, respectively).

Figure 14 shows the CO total column at 14:15 UTC on 4 August 2003 (during the period of the Portugal forest fires) from the NR (top left panel); the simulated S-5P observations (top right panel); the CR (bottom left panel); and the AR (bottom right panel). We can see that the AR captures the fire event, indicated by relatively high values of the CO total column over Portugal, whereas the CR does not. This confirms the results shown in Fig. 10, which highlight the benefit provided by the S-5P CO total column measurements, in particular regarding the capture of the signature of the Portugal forest fires. Note that the S-5P instrument does not make measurements over the sea (top right panel). This accounts for the sharp edge in the CO total column field seen between the Iberian Peninsula and the Bay of Biscay for the AR (bottom right panel).

Figure 14 shows a further example of the benefit of the assimilation system, namely redistributing information from data-rich areas (the land) to data-poor areas (the sea) through long-range transport. Although the S5-P does not observe CO over the sea (top right panel), the AR (bottom right panel) captures the plume of CO extending

from the north of the Iberian peninsula into the Bay of Biscay, also seen in the NR (top left panel). The CR does not capture this information (bottom left panel).

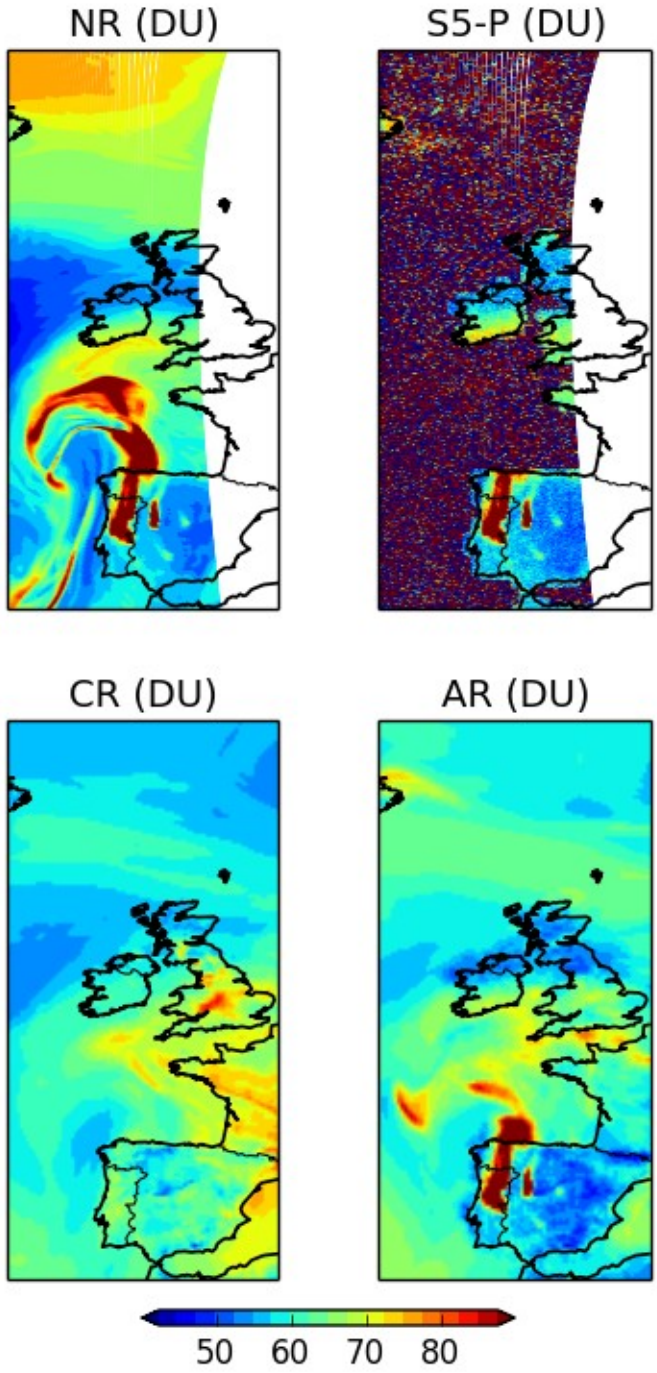

Figure 14: *CO total column at 14:15 UTC on 4 August 2003, Dobson units, DU. Top left panel: NR; top right panel: simulated S-5P observations; bottom left panel: CR; bottom right panel: AR. Red/blue colours indicate relatively high/low values of the CO total column.*

Figure 15 shows the time-series of the surface CO concentrations over the period 25 July – 15 August (that of the Portugal forest fires). In comparison to the original OSSE (see middle panel of Fig. 13), the AR is now closer to the NR, having now peak values of about 900 ppbv, instead of peak values of about 550 ppbv. The CR still has peak values less than 200 ppbv. This indicates that the relatively low values in the AR (in comparison to the NR) for the original OSSE shown in the middle panel

of Fig. 10 result from the application of the default criterion to discard CO column observations that are far away from MOCAGE values. The results from Fig. 15 confirm those shown in Fig. 14, and reinforce the benefit provided by the S-5P CO total column measurements, in particular regarding the capture of the signature of the Portugal forest fires.

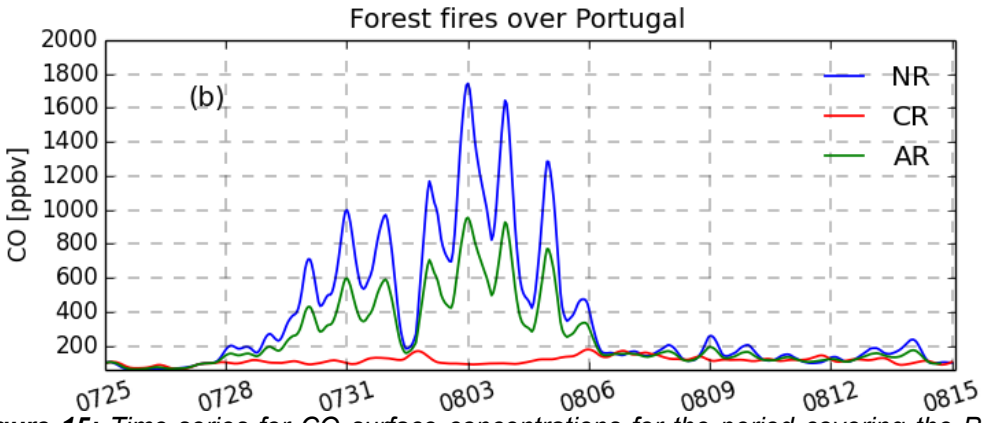

Figure 15: *Time-series for CO surface concentrations for the period covering the Portugal forest fires (25 July – 15 August) from NR (blue colour), CR (red colour) and AR (green colour) over the location associated with the middle panel of Fig.13. These data concern the second OSSE performed and described in the text. The labels are time, in format MMDD (x-axis) and CO concentration, ppbv (y-axis).*

V. The Winter period

We focus on three months during winter 2003 (November, December 2003 and January 2004). The primary goal is to study the differences of S5P CO added value between the summer 2003 and the winter 2003 periods. Different to the summer, which was relatively cloud-free, the winter experienced a heavy cloudiness situation typical of winter over Europe. This period allowed the study of the impact of cloudy pixels in the OSSE. For this, we performed two winter OSSEs, one using all pixels (cloudy pixels and clear pixels), hereafter the 100% case, and a second with less than 10% of cloud cover, hereafter the 10% case. For the 100% case, we assimilate all the land pixels, whereas we assimilate the pixels over sea if the cloud cover is greater than 60%. The 10% case is representative of a clear sky OSSE. Finally, we made the comparison between these two winter OSSEs and the summer OSSE.

V-1 Nature run and control run

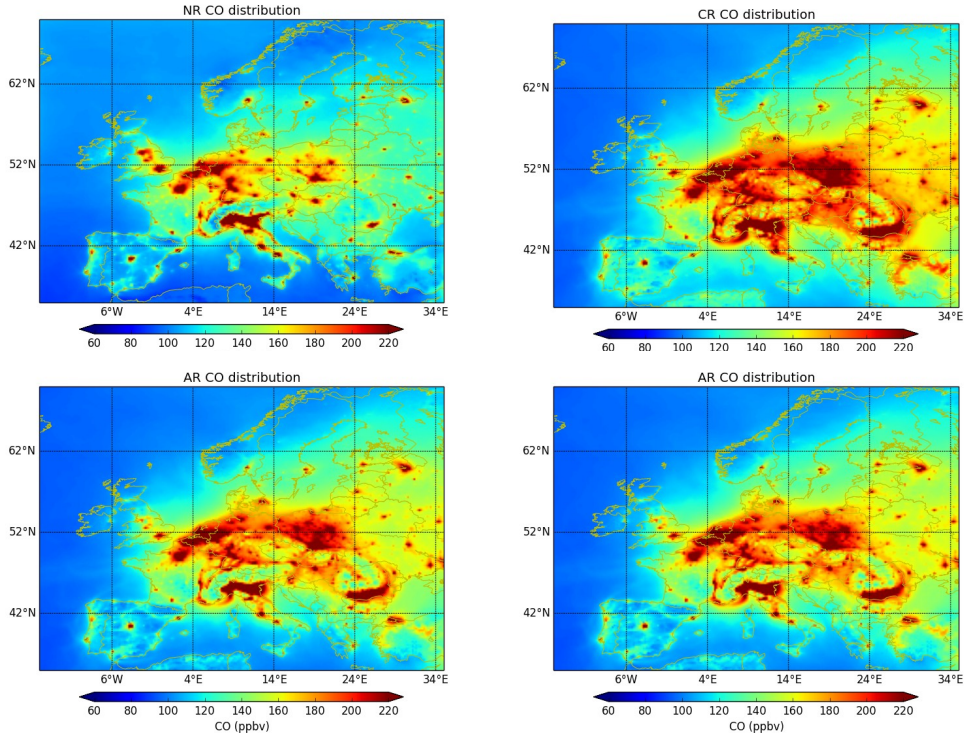

Figure 16: *Top left panel: NR CO averaged for the three months of winter (Nov, Dec 2003 and January 2004); top right panel: CR obtained for the same period. Bottom left and bottom right panels: AR obtained for the 100% case and the 10% case, respectively. All fields are in units of ppbv. The labels on each panel are longitude, degrees (x-axis) and latitude, degrees (y-axis). Red/blue colours indicate relatively high/low values in the NR, CR and AR fields. One can see almost no differences between the two fields.*

V-2 Evaluation of the winter OSSE

V-2.1 Statistical analysis

As for the summer period, the goal of this section is to provide a quantitative assessment of the added value of TROPOMO/S-5P CO total column measurements on the assimilation analysis at the surface and during a period of winter, characterized by significant cloud cover. To assess this added value, we performed a statistical analysis of the different runs during winter 2003-2004.

As for summer 2003, we calculate the mean bias (MB, ppbv), its magnitude reduction (MBMR, ppbv), the root mean square error (RMSE, ppbv), its reduction rate (RMSERR, %), along with the mean absolute error (MAE, ppbv), and its reduction rate (MAERR,%). Finally, to measure the linear dependence between two data sets, and the portion of the true variability (variability in NR) reproduced by the CR or AR, we use the correlation coefficient R, and the coefficient of determination (R^2) given as the squared correlation.

In the results section (V-2.2), we only show the 100% case because the differences between the 100% case and the 10% case are too small to warrant investigation (see Fig. 16). However, we present in section V-3 the differences between the 100 % case and the 10% case.

V-2.2 Results

Contrary to the case for summer 2003, the mean biases calculated for CR (CR-NR) and for AR (AR-NR) (**Figure 17**) show mostly positive values with slightly smaller values for the AR, especially over the Mediterranean Sea. The reduction of the bias (MBMR) is weak but with some significant values over the Alps and Northern Italy (between 20 to 24 ppbv of reduction). Because of the cloud cover during winter, no CO observations are possible at the surface. For this reason, the improvements likely come from the simulated S5P CO concentrations added above the clouds and redistributed by the assimilation system to the lowest layers in the troposphere. One can note that over the Atlantic, the level of significance of the results is less than 99%.

Figure 18 presents the mean absolute error (MAE) and the reduction of MAE for the winter period (MAERR). The MAERR shows that the reduction is larger over the Mediterranean Sea in correlation with smaller values of the mean bias between the AR and the NR. This suggests that although the MAE for the CR has maximum values over land, there is only a small reduction of the MAE overland whereas it is larger over the Mediterranean Sea. Over the Atlantic, the values of MAERR are outside the 99% confidence limits (identified by the shaded areas).

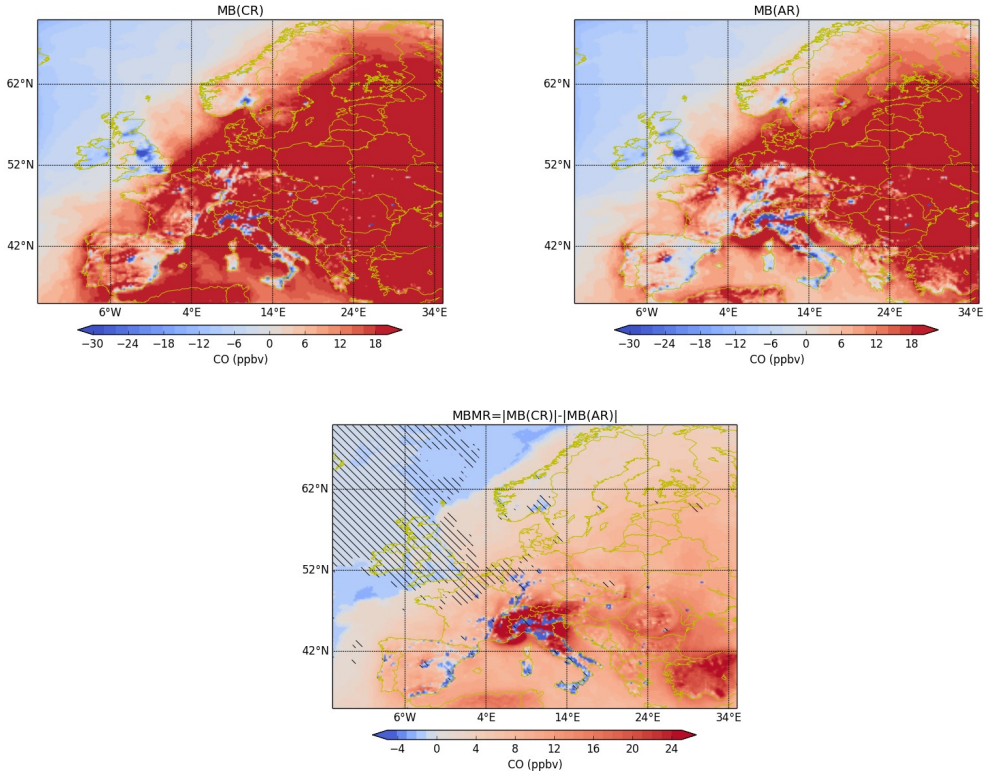

***Figure 17**: Mean bias reduction at the surface for CO, units of ppbv: Left top panel shows the CR mean bias (CR-NR). Right top panel shows the AR mean bias (AR-NR). Bottom panel shows the mean bias magnitude reduction (absolute value of the mean bias for CR less the absolute value of the mean bias for AR). We average the data over the winter 2003-2004 (1 November – 31 January). The labels show longitude, degrees (x-axis) and latitude, degrees (y-axis). The hatched area in the bottom panel shows where the mean bias plotted in this panel (MBMR) is not statistically significant at the 99% confidence level.*

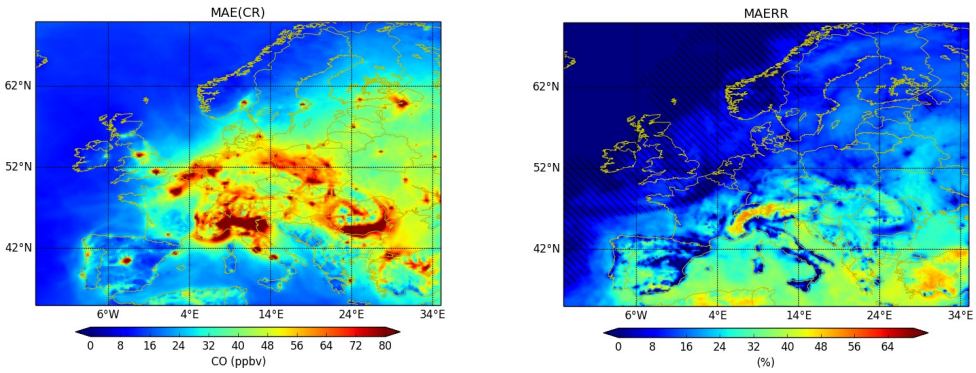

***Figure 18**: Mean absolute error (MAE), units of ppbv, between CR and NR (left panel), and its corresponding reduction rate MAERR, in % (right panel) removing the systematic error. We calculate the fields for the winter 2003-2004 at the surface. The labels on each panel are longitude, degrees (x-axis) and latitude, degrees (y-axis). Red/blue colours indicate relatively high/low values in the MAE/MAERR. The hatched area shows where the reduction rate (MAERR) is not statistically significant at the 99% confidence level.*

Figure 19 shows the RMSE variability of the CO together with its reduction rate. Although the maximum of variability for the CR was over land, the larger reduction is over the Mediterranean Sea and over Portugal with values of about 30-45%. Elsewhere, the values of reduction are below 15% with no reduction over the Atlantic (shown by the white areas).

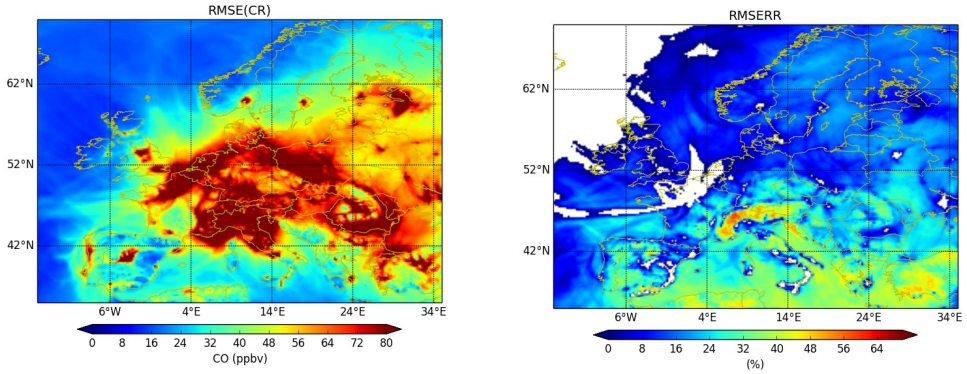

Figure 19:*.Root mean square error (RMSE), units of ppbv, between CR and NR (left panel), and its corresponding reduction rate RMSERR, in % (right panel) removing the systematic error. We calculate the fields for the winter 2003-2004 at the surface. The labels on each panel are longitude, degrees (x-axis) and latitude, degrees (y-axis). Red/blue colours indicate relatively high/low values in the MAE/MAERR.*

Figure 20 (top) presents the correlation coefficients between the CR and the NR and between the AR and the NR. The improvement (value of correlation coefficient closer to 1) is located over the Mediterranean area in accordance with the metrics shown earlier in this section. The values are from 0.4 to 0.8. In addition, we represent in **Figure 20** (bottom) the coefficient of determination R^2 which is a statistic that measures how well the regression line approximates the real data points. An R^2 of 1 indicates that the regression line fits perfectly the data. One can see that there is little difference between R^2(CR) and R^2(AR) in the northern part of the domain. The improvement is again over the Mediterranean Sea, in particular in the South East part of the basin.

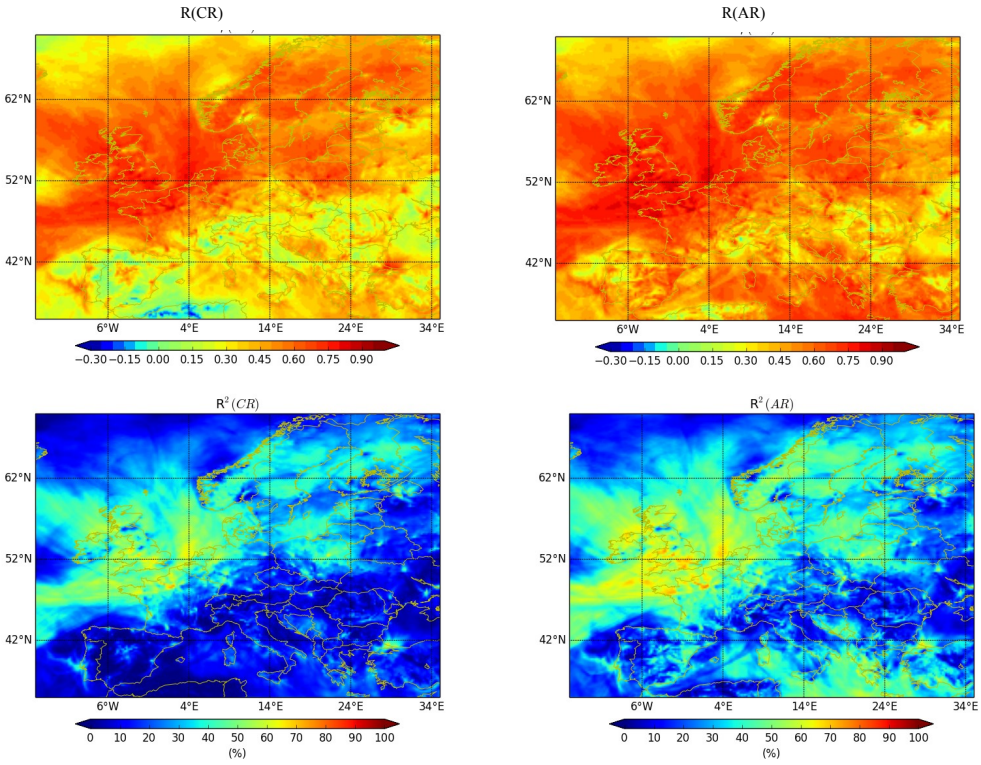

Figure 20 : *Top: Correlation coefficient (R) between CR and NR (left panel), and the AR and the NR (right panel). Bottom: Same as top panel but for the coefficient of determination. We calculate the statistics for the surface and the 2003-2004 winter period. The labels on each panel are longitude, degrees (x-axis) and latitude, degrees (y-axis). Red/blue colours indicate relatively high/low values in the statistic plotted.*

V-3 Difference between the 100% case and the 10% case

The main result of the comparison between the 10% case and the 100% case is that the differences between these two cases are negligible if we look at all the statistical metrics presented. However, to quantify these differences, we calculate **(Figure 21)** the relative difference (in %) between the absolute value of the difference between the assimilation runs (10% case and 100% case) and the nature run. We normalize this relative difference by the bias between the control run and the nature run. The field obtained is only positive with very weak values around 4% with maximum values of about 8 % over the Alps. Therefore, by using all pixels, the AR is closer to the NR than using only pixels with cloud cover less than 10%. This indicates that cloudy pixels do not affect the performance from the AR in comparison to that of the CR.

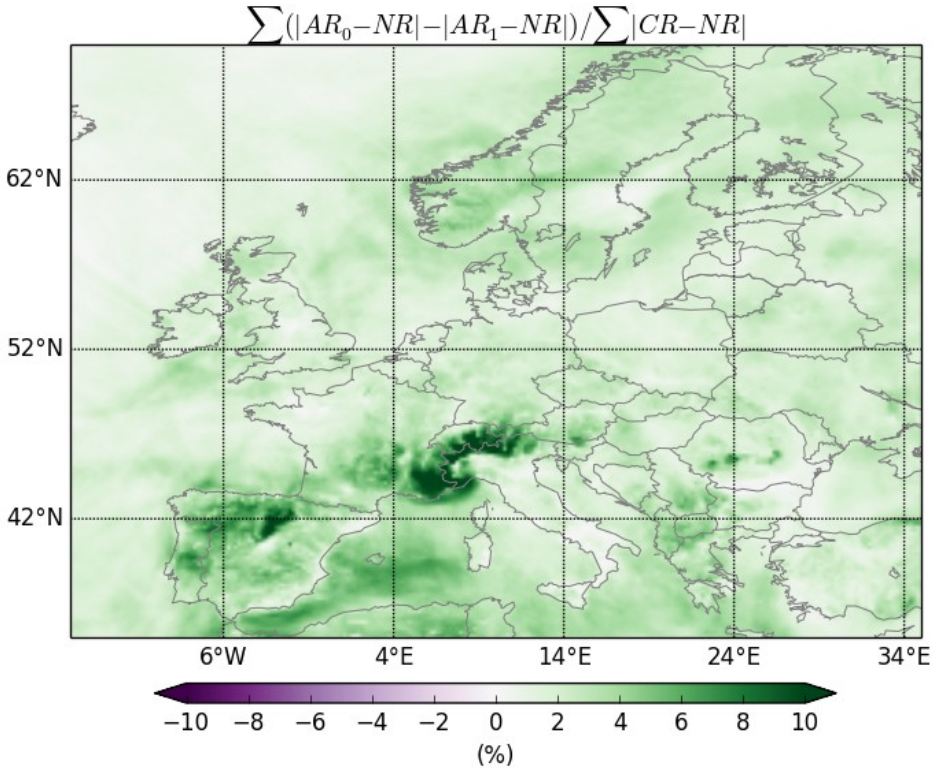

Figure 21: *Relative difference (in %) between the absolute value of the difference between the assimilation runs, (100% case, AR_1 and 10% case, AR_0) and the nature run, NR. We normalize this relative difference by the bias between the control run, CR, and the nature run, NR. See top of the figure for the formula used. We calculate the statistic for the surface and the 2003-2004 winter period. Green/purple colours indicate positive/negative values of the statistic.*

VI. Conclusions

We perform several regional-scale Observing System Simulation Experiments (OSSEs) over Europe to explore the impact of the LEO satellite mission S-5P carbon monoxide (CO) total column measurements on lowermost tropospheric air pollution analyses, with a focus on CO surface concentrations and the Planetary Boundary Layer (PBL). The PBL varies in depth throughout the year, but is contained within the lowermost troposphere (heights 0-3 km), and typically spans the heights 0-1 km. We focus on two periods during 2003-2004. The first part of this report concerns summer of 2003 (JJA) and the second part concerns winter 2003-2004 (NDJ).

This OSSE study provides insight on the impact from LEO S-5P CO measurements on surface CO information. We perform the standard steps of an OSSE for air quality, as identified in Timmermans et al. (2015). (i) Production of a Nature Run, NR. (ii) Test of the realism of the NR. (iii) Different models to produce, on the one hand, the NR, and on the other hand, the OSSE experiments to create the Control Run, CR, and the Assimilation Run, AR. (iv) Calculation of averaging kernels to represent sensitivity of the observations in the vertical. (v) Quantitative evaluation of the OSSE results, including performing statistical significance tests, and self-consistency and chi-squared tests.

Our guiding principle in the set-up of this OSSE study is to avoid overoptimistic results. To achieve this, we address several factors considered likely to contribute to an overoptimistic OSSE. (i) We use different models for the NR and the OSSE experiments. (ii) We check that the differences between the NR and actual measurements of CO are comparable to the CO field differences between the model used for the OSSE and the NR. (iii) We remove the systematic error (calculated as the bias against the NR) in the OSSE outputs (AR and CR) and compare the debiased results to the NR.

The OSSE results for summer 2003 (June, July and August) indicate that simulated S-5P CO total column measurements benefit efforts to monitor surface CO. The largest benefit occurs over land in remote regions (Eastern Europe, including Russia) where CO sources are sparse. Over these land areas, and for the case when we remove the systematic error, we obtain a lower RMSE value (by ~10 ppbv) for the AR than for the CR, in both cases vs the NR. Over sea and Scandinavia, we also obtain a lower RMSE (by ~10%) for the AR than for the CR, in both cases vs the NR. Consistent with this behaviour, we find the AR is generally closer to the NR than the CR to the NR, with a correlation coefficient R, reaching 0.9 over land (NR vs AR). By contrast, the correlation coefficient between the CR and the NR is typically less than 0.5, with very low values over Eastern Europe, where CO sources are sparse. In general, for all the metrics calculated in this paper, there is an overall benefit over land from the S-5P CO total column measurements. Significance tests on the CR and AR results indicate that, generally, the differences in their performance are significant at the 99% confidence level. This indicates that the S-5P CO total column measurements provide a significant benefit to monitor surface CO during northern summer.

We further show that, locally, the AR is capable of reproducing the peak in the CO distribution at the surface due to forest fires (albeit, weaker than the NR signal), even if the CR does not have the signature of the fires in its emission inventory. A second OSSE shows that this relatively weak signal of the forest fires in the AR arises from the use of a criterion to discard CO total column observations too far from model values, a criterion not appropriate to situations resulting in excessive values in the CO concentrations, as for forest fires. This second OSSE shows a much stronger signal in the AR, which is now much closer to the NR than the CR, confirming the benefit of S-5P CO total column measurements.

We now discuss the results for the OSSEs for the three months during winter 2003-2004 (November, December 2003 and January 2004). In contrast to the summer period, the winter experienced a generally cloudy situation. This period allowed us to study the impact of cloudy pixels in the OSSE. For this, we performed two OSSEs, one using pixels with 100% cloudy pixels and a second with only 10% cloudy pixels. We then compare these two OSSEs. The statistics obtained for these two OSSEs indicate that they are very close with differences less than 8%. This indicates that cloudy pixels do not affect the improvement coming from S-5P CO observations. However, even if the CO concentrations during winter are larger than during summer, the impact of the cloud cover is much more significant as it prevents observations at the surface from S-5P. We thus obtain the best improvement with the simulated S-5P observations for summer.

Headline message: Significant improvement from S-5P over summer and winter for CO information in the lowermost troposphere. Cloudy pixels do not have an impact on the improvement, but affect the performance of the S-5P by preventing surface observations of CO.

VII. References

Arnold, C.P., Jr., and C.H. Dey, 1986: Observing-systems simulation experiments: Past, present and future. Bull. Amer. Meteorol. Soc., 67, 687–695.

Atlas, R. 1997: Atmospheric observation and experiments to assess their usefulness in data assimilation. J. Meteor. Soc. Jpn., 75, 111–130.

Baldocchi, D., Hicks B. B., and P. Camara. A canopy stomatal resistance model for gaseous deposition to vegetated surfaces. Atmos. Environ., 21: 91–101, 1988.

Bousserez, N., Attié, J. L., Peuch, V. H., Michou, M., Pfister, G., Edwards, D., Emmons, L., Mari, C., Barret, B., Arnold, S. R., Heckel, A., Richter, A., Schlager, H., Lewis, A., Avery, M., Sachse, G., Browell, E. V., and Hair, J. W.: Evaluation of the MOCAGE chemistry transport model during the ICARTT/ITOP experiment, J. Geophys. Res., 112, D10S42, doi:10.1029/2006JD007595, 2007.

Cathala, M.-L., Pailleux, J., and Peuch, V. -H.: Improving chemical simulations of the upper troposphere – lower stratosphere with sequential assimilation of MOZAIC data, Tellus, 55B, 1–10, 2003.

Claeyman, M., J.-L. Attié, V.-H. Peuch, L. El Amraoui, W.A. Lahoz, B. Josse, P. Ricaud, T. Von Clarmann, M. Höpfner, J. Orphal, J.-M. Flaud, D.P. Edwards, K. Chance, X. Liu, F. Pasternak and R. Cantié, 2011a: A geostationary thermal infrared sensor to monitor the lowermost troposphere: O3 and CO retrieval studies. Atmos. Meas. Tech., 4, 297-317.

Claeyman, M., J.-L. Attié, V.-H. Peuch, L. El Amraoui, W.A. Lahoz, B. Josse, M. Joly, J. Barré, P. Ricaud, S. Massart, A. Piacentini, T. Von Clarmann, M. Höpfner, J. Orphal, J.-M. Flaud and D.P. Edwards, 2011b: A thermal infrared instrument onboard a geostationary platform for CO and O3 measurements in the lowermost troposphere: Observing System Simulation Experiments. Atmos. Meas. Tech., 4, 1637-1661.

Clark, H. L., Cathala, M.-L., Teyssèdre, H., Cammas, J.-P., and Peuch, V.-H.: Cross-tropopause fluxes of ozone unsing assimilation of MOZAIC observations in a global CTM, Tellus, 59B, 39–49, 2007.

Edwards, D.P., A.F. Arellano Jr., and M.N. Deeter, 2009: A satellite observation system simulation experiment for carbon monoxide in the lowermost troposphere. J. Geophys. Res., 114, D14304, doi: 10.1029/2008JD011375.

El Amraoui, L., Peuch, V.-H., Ricaud, P., Massart, S., Semane, N., Teyssèdre, H., Cariolle, D., and Karcher, F.: Ozone loss in the 2002/03 Arctic vortex deduced from the Assimilation of Odin/SMR O3 and N2O measurements: N2O as a dynamical tracer, Q. 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 J.-L. **Attié**, P. Ricaud, W. A. Lahoz, A. Piacentini, V.-H. Peuch, J. X. Warner, R. Abida, J. Barré, and R. Zbinden, 2014: Tropospheric CO vertical profiles deduced from total columns using data assimilation: methodology and validation Atmos. Meas. Tech., 7, 3035-3057

Emberson, L., Ashmore M.R., Cambridge H.M., Simpson D., and J.P. Touvinen. Modelling stomatal ozone flux across europe. Environ. Poll., 109, No. 3:403–414, 2000.

Joly M. and Peuch V-H., 2012 : Objective classification of air quality monitoring sites over Europe. Atm. Env. 47, 111-123. doi 10.1016/j.atmosenv.2011.11.025

Josse, B., Simon, P., and Peuch, V.-H.: Radon global simulation with the multiscale chemistry trasnport model MOCAGE, Tellus, 56, 339–356, 2004.

Lahoz, W.A., Q. Errera, R. Swinbank, and D. Fonteyn, 2007: Data assimilation of stratospheric constituents: a review. Atmos. Chem. Phys., 7, 5745–5773, doi: 10.5194/acp-7-5745-2007.

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

Masutani, M., T.W. Schlatter, R. M. Errico, A. Stoffelen, E. Andersson, W. Lahoz, J.S. Woollen, G.D. Emmitt, L.-P. Riishøjgaard, and S. J. Lord, 2010a: Observing system simulation experiments. Data Assimilation: Making Sense of Observations, W. A. Lahoz, B. Khattatov and R. Ménard, Eds.,Springer, 647-679.

Masutani, M., J.S. Woollen, S.J. Lord, G.D. Emmitt, T.J. Kleespies, S.A. Wood, S. Greco, H. Sun, J. Terry, V. Kapoor, R. Treadon, and K.A. Campana, 2010b: Observing system simulation experiments at the National Centers for Environmental Prediction. J. Geophys. Res., 115, doi: 10.1029/2009JD012528.

Nenes, Pandis S.N., and Pilinis C. Isorropia: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols. Aquat.Geoch., 4:123–152, 1998.

Pradier, S., Attié, J.-L., Chong, M., Escobar, J., Peuch, V.-H., Lamarque, J.-F., Kattatov, 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.

Ricaud, P., Attié, J.-L., Teyssèdre, H., El Amraoui, L., Peuch, V.-H., Matricardi, M., and Schluessel, P.: Equatorial total column of nitrous oxide as measured by IASI on MetOp-A: implications for transport processes, Atmos. Chem. Phys., 9, 3947–3956, 2009a.

Ricaud, P., Pommereau, J.-P., Attié, J.-L., Le Flochmoen, E., El Amraoui, L., Teyssèdre, H., Peuch, V.-H., Feng, W., and Chipperfield, M. P.: Equatorial transport

as diagnosed from nitrous oxide variability, Atmos. Chem. Phys., 9, 8173–8188, 2009b.

Schaap, M., Roemer M., Sauter M., Boersen G., Timmermans R., and Builtjes P.J.H. Lotos- euros documentation. Technical 609 report, TNO Report B&O 2005/297, 2005.

Schaap, M., R. M. A. Timmermans, M. Roemer, G. A. C. Boersen, P. J. H. Builtjes, F. J. Sauter, G. J. M. Velders, and J. P. Beck. The Lotos-Euros model: Description, validation and latest developments. International Journal of Environment and Pollution, 32[2):270–290, 2008.

Simpson, D., Fagerli H., Jonson J.E., Tsyro S., Wind P., and Tuovinen JP. Transboundary acidification, eutrophication and ground level ozone in europe, part 1: unified emep model description. Technical report, EMEP Report 1/2003, Norwegian meteorologyrological Insitute, 2003.

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, Q. J. Roy. Meteor. Soc., 133(S2), 171–178, doi:10.1002/qj. 141, 2007.

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

D. T. Shindell et al., Multimodel simulations of carbon monoxide: Comparison with observations and projected near-future changes", J. Geophys. Res., 111, doi: 10.1029/2006JD007100, 2006.

Wesely. M., Parameterization of surface resistance to gaseous dry deposition in regional scale numerical models. Atmos. Environ, 23:1293Ű1304, 1989.

Zoogman, P., D.J. Jacob, K. Chance, H.M. Worden, D.P. Edwards, and L. Zhang, Improved monitoring of surface ozone by joint assimilation of geostationary satellite observations of ozone and CO. Atmos. Env., 84, 254-261. http://dx.doi.org/10.1016/j.atmosenv.2013.11.048, 2014.