



1 **Impact of Spaceborne Carbon Monoxide**  
2 **Observations from the S-5P platform on**  
3 **Tropospheric Composition Analyses and Forecasts**

4

5 R. Abida<sup>1</sup>, J.-L. Attié<sup>1,2</sup>, L. El Amraoui<sup>1</sup>, P. Ricaud<sup>1</sup>, W. Lahoz<sup>3</sup>, H. Eskes<sup>4</sup>, A. Segers<sup>5</sup>, L. Curier<sup>5</sup>, J.  
6 de Haan<sup>4</sup>, J. Kujanpää<sup>6</sup>, A. O. Nijhuis<sup>4</sup>, D. Schuettmeyer<sup>7</sup>, J. Tamminen<sup>6</sup>, R. Timmermans<sup>5</sup>, P.  
7 Veefkind<sup>4</sup>, and B. Veihelmann<sup>7</sup>

8

9 <sup>1</sup> CNRM-GAME, Météo-France/CNRS UMR 3589, Toulouse, France

10 <sup>2</sup> Université de Toulouse, Laboratoire d'Aérodynamique, CNRS UMR 5560, Toulouse, France

11 <sup>3</sup> NILU – Norwegian Institute for Air Research, P.O. Box 100, 2027 Kjeller, Norway

12 <sup>4</sup> Royal Netherlands Meteorological Institute (KNMI), P.O. Box 201, 3730 AE De Bilt, The  
13 Netherlands

14 <sup>5</sup> TNO, Business unit Environment, Health and Safety, P.O. Box 80015, 3508 TA Utrecht, The  
15 Netherlands

16 <sup>6</sup> Finnish Meteorological Institute, Earth Observation Unit, P.O. Box 503, 00101 Helsinki, Finland

17 <sup>7</sup> ESA/ESTEC, Earth Observation Programmes, Noordwijk, The Netherlands

18

19 Submitted to Atmospheric Chemistry and Physics Discussions, 13 November 2015



## 20 **Abstract**

21 We use the technique of Observing System Simulation Experiments (OSSEs) to quantify the impact of  
22 spaceborne carbon monoxide (CO) total column observations from the Sentinel-5 Precursor (S-5P) platform  
23 on tropospheric analyses and forecasts. We focus on Europe for the period of northern summer 2003, when  
24 there was a severe heat wave episode associated with extremely hot and dry weather conditions. We describe  
25 different elements of the OSSE: (i) the Nature Run (NR), i.e., the “Truth”; (ii) the CO synthetic observations;  
26 (iii) the assimilation run (AR), where we assimilate the observations of interest; (iv) the control run (CR), in  
27 this study a free model run without assimilation; and (v) efforts to establish the fidelity of the OSSE results.  
28 Comparison of the results from AR and the CR, against the NR, shows that CO total column observations  
29 from S-5P provide a significant benefit (at the 99% confidence level) at the surface, with the largest benefit  
30 occurring over land in remote regions. Furthermore, the S-5P CO total column observations are able to  
31 capture phenomena such as the forest fires that occurred in Portugal during summer 2003. These results  
32 provide evidence of the benefit of S-5P observations for monitoring processes contributing to atmospheric  
33 pollution.  
34



## 35 **1. Introduction**

36 Over the last decade, the capabilities of satellite instruments for sensing the lower troposphere have  
37 improved, and opened the way for monitoring and better understanding of atmospheric pollution processes,  
38 e.g., tropospheric chemistry (Jacob, 2000), long-range transport (HTAP, 2007), and emissions (e.g. Streets  
39 D., 2013 and references therein). Satellite instruments provide global measurements of many pollutants (e.g.,  
40 ozone; carbon monoxide, CO; nitrogen dioxide, NO<sub>2</sub>; and aerosols), including information on their trans-  
41 boundary transport, and complement in situ measurements from ground-based stations (e.g., the EMEP,  
42 <http://www.nilu.no/projects/ccc/emepdata.html>, and Airbase, <http://acm.eionet.europa.eu/databases/airbase/>,  
43 networks). Low Earth Orbit (LEO) satellite platforms have the advantage of providing observations with  
44 global coverage, but at a relatively low temporal resolution. Geostationary Earth Orbit (GEO) satellite  
45 platforms provide observations at a continental scale, i.e., not global, but at a much higher temporal  
46 resolution.


47

48 Satellite data, either in synergy with ground-based and airborne measurements and/or assimilated into  
49 models such as chemistry transport models (CTMs), contribute to an improved understanding of tropospheric  
50 chemistry and dynamics and improved forecasts of atmospheric pollutant fields (see, e.g., Elbern et al.,  
51 2010). As part of an integrated observing strategy, satellite measurements provide a global view on air  
52 quality (AQ). The challenge for future space-borne missions will be to assess directly the local scales of  
53 transport and/or chemistry for tropospheric pollutants (1 hour or less, 10 km or less) and to facilitate the use  
54 of remote sensing information for improving local- and regional-scale (from country-wide to continental  
55 scales) AQ analyses and forecasts. Building on this effort, various LEO satellite platforms and/or  
56 constellations of GEO satellite platforms will help extend AQ information from continental scales to global  
57 scales (e.g., Lahoz et al., 2012, and references therein for LEO/GEO platforms; Barré et al., 2015, for GEO  
58 platforms).

59

60 An atmospheric species of interest for monitoring AQ is CO, owing to its relatively long time-scale in the  
61 troposphere; its distribution provides information on the transport pathways of atmospheric pollutants.  
62 Spaceborne instruments on LEO satellite platforms demonstrate the potential of remote sensing from space



63 to determine the CO distribution and its main emission sources at the global scale (Edwards et al., 2004,  
64  Buchwitz et al., 2006; Worden et al., 2013 and references therein). These LEO satellite platforms  
65 include MOPITT (Measurements Of Pollution In The Troposphere), IASI (Infrared Atmospheric Sounding  
66 Interferometer), AIRS (Atmospheric InfraRed Sounder) and TES (Tropospheric Emission Spectrometer)  
67 operating in the thermal infrared (TIR) and SCIAMACHY (SCanning Imaging Absorption spectromETER for  
68 Atmospheric ChartographY) operating in the short-wave infrared (SWIR), respectively. By contrast, to our  
69 knowledge, there are no GEO satellite platforms measuring the CO distribution. However, despite their  
70 potential, owing to limited revisit time, and relatively coarse spatial resolution, LEO instruments are not  
71 optimal for monitoring regional and local aspects of air quality.

72  
73 Copernicus is the current European Programme for the establishment of a European capability for Earth  
74 Observation (<http://www.copernicus.eu/pages-principales/services/atmosphere-monitoring>). The main  
75 objective of the Copernicus Atmospheric Services is to provide information on atmospheric variables (e.g.,  
76 the essential climate variables, ECVs; [https://www.wmo.int/pages/prog/gcos/index.php?name=](https://www.wmo.int/pages/prog/gcos/index.php?name=EssentialClimateVariables)  
77 EssentialClimateVariables) in support of European policies regarding sustainable development and global  
78 governance of the environment. The Copernicus Atmospheric Services cover: AQ, climate change/forcing,  
79 stratospheric ozone and solar radiation. The services rely mainly on data from Earth Observation satellites.

80  
81 To ensure operational provision of Earth Observation data, the space component of the Copernicus  
82 programme includes a series of spaceborne missions developed and managed by the European Space Agency  
83 (ESA) and EUMETSAT. Among them, three missions address atmospheric composition. These are the  
84 Sentinel-5 (S-5) and Sentinel-5 Precursor (S-5P) from a LEO satellite platform, and the Sentinel-4 (S-4)  
85 from a GEO satellite platform. The goal of the S-4 is to monitor key atmospheric pollutants (e.g., ozone;  
86 NO<sub>2</sub>; sulphur dioxide, SO<sub>2</sub>; bromine monoxide, BrO; and formaldehyde) and aerosols at relatively high  
87 spatio-temporal resolution over Europe and North Africa (8 km; 1 hour). The goal of the S-5 and S-5P  
88 platforms is to provide global daily measurements of atmospheric pollutants (e.g., CO, ozone, NO<sub>2</sub>, SO<sub>2</sub>,  
89 BrO, and formaldehyde), climate related trace gases (e.g., methane, CH<sub>4</sub>) and aerosols, at relatively high  
90 spatial resolution (from below 8 km to below 50 km, depending on wavelength).

91



92 The S-5P is the ESA pre-operational mission required to bridge the gap between the end of the OMI (Ozone  
93 Monitoring Instrument) and the SCIAMACHY missions and the start of the S-5 mission planned for 2020  
94 onwards. The S-5P scheduled launch is in 2016 with a 7 years design lifetime. The S-5P will fly in an early  
95 afternoon sun-synchronous LEO geometry with an Equator crossing mean local solar time of 13:30, chosen  
96 to allow the instrument to measure the strong pollution signal present in the afternoon. In contrast, the  
97 GOME-2 (Global Ozone Monitoring Experiment - 2) platform collects data at a local solar time of 09:30  
98 (when the pollution signal is relatively weak) and thus has a lower predictive value (Veeffkind et al., 2012,  
99 and references therein). The S-5P LEO platform will address the challenge of limited revisit time from LEOs  
100 by providing unprecedented high spatial resolution of 7x7 km, and improved sensitivity in the Planetary  
101 Boundary Layer (PBL), allowing resolution of, e.g., derived CO emission sources at finer scales than  
102 hitherto. The PBL varies in depth throughout the year, but is contained within the lowermost troposphere  
103 (heights 0-3 km), and typically spans the heights 0-1 km.

104  
105 A method to objectively determine the added value of future satellite observations such as S-4, S-5 and S-5P,  
106 and to investigate the impact of different instrument designs, is that of Observing System Simulation  
107 Experiments (OSSEs) commonly based on data assimilation (e.g., Lahoz and Schneider, 2014). The OSSEs  
108 have been extensively used and shown to be useful in the meteorological community to test the impact of  
109 future meteorological observations on the quality of weather forecasts (Nitta, 1975; Atlas, 1997; Lord et al.,  
110 1997; Atlas et al., 2003). In a recent paper, Timmermans et al. (2015) review the application of OSSEs to  
111 assess future missions to monitor AQ. The OSSEs are increasingly being used by the space agencies to assess  
112 the added value of future instruments to be deployed as part of the Global Observing System (e.g., work on  
113 the ESA Earth Explorer ADM-Aeolus; Tan et al., 2007).

114  
115 Although the usefulness of OSSEs is well established, they have limitations, discussed in Masutani et al.  
116 (2010a, b). A frequent criticism of OSSEs is that they are overoptimistic, largely owing to the difficulties of  
117 representing the real Earth System (e.g., the atmosphere), even with state-of-the-art numerical models.  
118 Nevertheless, even if overoptimistic, OSSEs provide bounds on the impact of new observing systems. For  
119 example, if additional instruments provide no significant impact within an OSSE, they are unlikely to do so  
120 in reality.



121

122 In this paper, we describe a regional-scale OSSE over Europe for northern summer 2003 (1 June – 31  
123 August) to explore the impact of S-5P CO total column measurements on lowermost tropospheric air  
124 pollution analyses, with a focus on CO PBL concentrations. The severe heat wave experienced in Europe  
125 during northern summer 2003, and the concomitant atmospheric pollution and fire episodes, had a strongly  
126 negative societal impact, being responsible for the deaths of over 14,000 people in France (Vautard et al.,  
127 2005). This period had extremely hot and dry weather conditions and the long lasting atmospheric blocking  
128 conditions significantly contributed to the accumulation of pollutants in the PBL owing to extended  
129 residence time of the air parcels (Solberg et al., 2008). The spatial distribution of the enhanced levels of CO  
130 and ozone was much more widespread over Europe during that summer than in previous ones (Lee et al.,  
131 2006; Ordoñez et al., 2010). These exceptional weather conditions also resulted in several extreme wildfire  
132 episodes over the Iberian Peninsula and the Mediterranean coast (Barbosa et al., 2004). Tressol et al. (2008)  
133 point out that between 6 and 10 August 2003 the contribution of biomass burning to measured CO levels in  
134 the lowermost troposphere reached 35% of the total CO field at these levels, a value comparable to typical

6-1  
135 European anthropogenic emissions which represent 30% of this total CO field. Thus, the three-month period  
136 1 June - 31 August 2003 includes both extreme and normal conditions, and provides an opportunity to study  
137 the full range of pollution levels that occur in a summer season over Europe.

138

139 The OSSE study domain covers the larger part of Europe (5W-35E, 35N-70N), and we perform the OSSE  
140 simulations at the spatial resolution of 0.2 degrees (latitude and longitude). This corresponds to a spatial  
141 resolution of ~20 km (meridionally) and ~15 km (zonally, at 45N). With this spatial resolution, we can track  
142 long-range transport plumes of CO. The length of the study period ensures we can sample different  
143 meteorological situations typical for summertime, and provides an acceptable compromise between run-time  
144 restrictions and provision of sufficient information for statistically significant results.

145

146 The structure of the paper is as follows. In Sect. 2 we describe the various components of the OSSE; in Sect.  
147 3 we present the results from the OSSE for S-5P during summer 2003 over Europe. Finally, Sect. 4 provides  
148 conclusions and identifies further work. A guiding principle in the OSSE set-up in this paper is to avoid  
149 overoptimistic results.



150

## 151 **2. The OSSE set-up**

152 The OSSE concept consists of simulating observations and their associated errors from a representation of  
153 reality (the “Nature Run” or NR) and providing this information to a data assimilation system to produce  
154 estimates of the NR states. Thereafter, one compares these estimates of the NR states from an assimilation  
155 run, AR (where the observation of interest has been assimilated), and from a control run, CR (in this case a

156 free model run), against the NR. The performance of the AR and the CR against the NR quantifies the benefit  
157 of the observation of interest.

158

159 The OSSEs are widely used in the meteorological community for assessing the usefulness of new  
160 meteorological satellite data. Recent examples (not exhaustive) include the work of Lahoz et al. (2005),  
161 Stoffelen et al. (2006), and Tan et al. (2007); Masutani et al. (2010a) reviews the OSSE methodology and  
162 provides a comprehensive list of references of OSSEs for meteorological applications. By contrast, there are  
163 relatively few studies concerning OSSEs for AQ applications (Edwards et al., 2009; Timmermans et al.,  
164 2009a, b; Claeys et al., 2011; Zoogman et al., 2011; 2014a, b; Yumimoto, 2013). In a recent review,  
165 Timmermans et al. (2015) comment that documented AQ OSSEs have demonstrated the benefits that could  
166 accrue from proposed and planned satellite platforms for AQ monitoring and forecasting. In the study  
167 described in this paper, the set-ups for the NR, and the CR and AR, use different models, thereby avoiding  
168 the identical twin problem typically associated with overly optimistic OSSE results (see, e.g., Masutani et al.,  
169 2010a). In Sects. 2.1-2.5 we describe the various elements of the OSSE study described in this paper. Figure

170 1 in Timmermans et al. (2015) provides a schematic showing the relationships between the various elements  
171 in an OSSE.

172

### 173 **2.1 The Nature Run**

174 A key element of an OSSE is the NR that defines the true state used to evaluate analyses and/or forecasts  
175 using simulated observations. The NR commonly consists of a long, free-running forecast evolving  
176 continuously in a dynamically consistent way (Masutani et al. 2010a, b). For this study, the basis of the NR  
177 consists of two high-resolution free model simulations performed with: (i) the regional LOTOS-EUROS air



178 quality model (Schaap et al., 2008), and (ii) the global chemistry transport model TM5 (Huijnen et al., 2010).  
179 We obtain the NR by combining the LOTOS-EUROS CO profiles from the surface to 3.5 km with the TM5  
180 CO profiles from 3.5 km to the top of the atmosphere (identified by the TM5 model top at 0.1 hPa). We use  
181 spatial interpolation to merge the values near the boundary between the two models at a height of 3.5 km.  
182 The model simulations used to construct the NR have a spin-up period of three months. We archive the NR  
183 output data on an hourly basis.

184

185 To construct the NR, we run the LOTOS-EUROS model at a horizontal resolution of about 7 km nested into  
186 the TM5 model, the latter run with a zoom domain over Europe at 1x1 degrees resolution. The TM5 model  
187 has 34 layers with a model top at 0.1 hPa. The LOTOS-EUROS model describes air pollution in the  
188 lowermost troposphere. It has four vertical layers following the dynamic mixing layer approach. The first  
189 layer is a fixed surface layer of 25 metres thickness, the second layer (boundary layer) follows the mixing  
190 layer height, and there are two reservoir layers spanning the rest of the atmosphere up to 3.5 km. The implicit  
191 assumption of the LOTOS-EUROS model is the presence of a well-mixed boundary layer, so constituent  
192 concentrations are constant up to the top of the Planetary Boundary Layer. The meteorological data used as  
193 input for the LOTOS-EUROS model come from the European Centre for Medium-Range Weather Forecasts  
194 (ECMWF). Prescription of surface anthropogenic emission is from the TNO-MACC-II emission database  
195 (Kuenen et al., 2014), and fire emissions are from the MACC global fire assimilation system (GFAS v1;  
196 Kaiser et al., 2012).

197

198 In the design of an OSSE, it is important to demonstrate that the NR exhibits the same statistical behaviour  
199 as the real atmosphere in every aspect relevant to the observing system under study (Masutani et al., 2010a,  
200 b). For the LOTOS-EUROS model used to build the lowermost levels of the NR, there is extensive  
201 verification by comparison with European data and by frequent participation in international model  
202 comparisons. This is the case for ozone and particulate matter (see Hass et al., 2003; Cuvelier et al., 2007;  
203 van Loon et al., 2007; Stern et al., 2008; Manders et al., 2009; Curier et al., 2012; Márecal et al., 2015). To  
204 evaluate the NR, we compare the surface CO data to available ground-based CO measurements over Europe  
205 during northern summer 2003 (1 June – 31 August). For this comparison, we use the ground-based stations  
206 from the Airbase database. We consider all types of ground-based stations from this database because of the





207 limited number of available measurements, but we discard stations with less than 75% of hourly data within  
208 a month. This provides 171 ground-based stations for the comparison against the NR (note this approach  
209 results in a paucity of stations over France).

210

211 Figure 1 shows the location of the selected Airbase ground-based stations measuring CO over Europe during  
212 northern summer 2003 (top panel), and the time-series of CO concentrations during 1 June – 31 August  
213 2003, measured by the selected Airbase ground-based stations and simulated by the NR and the CR (bottom  
214 panel). Note that most ground-based stations selected are located in polluted areas, where big emission  
215 sources of CO are present. We form the time-series from the ground-based stations by averaging spatially  
216 over all the sites. We form the NR time-series similarly, but interpolate the NR surface data to the station  
217 location. We do not add random observation errors to the NR time-series.

218

219 From Fig. 1, we see that, generally, the NR captures reasonably well the features of observed CO temporal  
220 variability during the three phases characterizing the summer of 2003: before, during and after the heat wave  
221 (the heat wave occurred on 31 July – 15 August). The correlation coefficient,  $\rho$ , between the ground-based  
222 data and NR time-series shown in the middle panel is 0.71. From this, we conclude that the NR has a realistic  
223 representation of the CO diurnal cycle. Note that CO concentration levels in the NR are slightly lower than  
224 observed ones. The bias of the NR with respect to observed CO concentrations fluctuates around -10 % on  
225 average during normal conditions and reaches -20% within the heat wave period. This means that the NR  
226 reproduces the surface concentrations with a negative bias (NR lower than ground-based stations) between  
227 10 and 20%. Nonetheless, the simulated CO concentrations and those measured by the ground-based stations

228 generally fall within the same range of values (between 200 and 400  $\mu\text{gm}^{-3}$ ). Thus, for the OSSE period  
229 considered, we conclude that the NR is representative of the variability of actual observations over the  
230 European domain, albeit with a negative bias.

231

232 Additionally, from Fig. 1 the behaviour of the CO time-series from the CR compared to the NR, is similar to  
233 the behaviour of the NR CO time-series compared to the Airbase data. This suggests that the configuration of  
234 our model is reasonably realistic, and reduces the likelihood that the OSSE produces overoptimistic results.

235

9-1

9-2

9-3



## 236 **2.2 The S-5P CO simulated measurements**

237 The S-5P will deploy the TROPOspheric Monitoring Instrument (TROPOMI) jointly developed by The  
238 Netherlands and ESA (Veeffkind et al. 2012). The TROPOMI instrument has heritage from both the OMI and  
239 the SCIAMACHY missions. The TROPOMI instrument will make measurements in the UV-visible  
240 wavelength range (270-500 nm), the near infrared, NIR (675-775 nm) and the shortwave infrared, SWIR  
241 (2305-2385 nm). It will deliver a key set of gas and aerosol data products for air quality and climate  
242 applications, including ozone, NO<sub>2</sub>, formaldehyde, SO<sub>2</sub>, methane and CO.

243

244 To enable sounding of the lower atmosphere at finer scales, TROPOMI has an unprecedented spatial  
245 resolution of 7x7 km<sup>2</sup> at nadir. This relatively high spatial resolution is necessary for air quality applications  
246 at local to regional scales. It will **resolve emission sources with relatively high accuracy**, and will obtain an  
247 acceptable fraction of cloud-free spectra. In contrast to the advantages provided by the relatively high spatial  
248 resolution of S-5P and design improvements, the SCIAMACHY CO data needs averaging in time (roughly  
249 one month) and space (5x5 degrees) to obtain realistic CO distributions at comparable uncertainty (Galli et  
250 al., 2012). Furthermore, TROPOMI will have a wide swath of 2600 km to allow for daily global coverage.

251 The relatively high radiometric sensitivity of S-5P will allow measurements at **low** albedo, thus helping track  
252 smaller pollution events and improving the accuracy of air quality assessments and forecasts. The use of S-  
253 5P CO total column measurements with inverse modelling techniques will also help quantify **more accurately**  
254 biomass burning emissions and map their spatial distribution. The simultaneous measurements of CO and,  
255 e.g. NO<sub>2</sub>, will provide additional information on wildfire and other pollution episodes (Veeffkind et al., 2012).

256

257 The NR results were used to generate a set of synthetic S-5P observations. This involves several steps: 1)  
258 Generating realistic S-5P orbits and geolocation and viewing/solar geometries for the appropriate overpass  
259 time; 2) Using the ECMWF modelled cloud distributions to generate effective cloud fractions; 3) Generation  
260 of lookup tables for the averaging kernels and observation errors; 4) Collocation and application of the NR to  
261 derive a set of synthetic observations for 3 Summer months and 3 Winter months. These steps are discussed  
262 in the sub-sections below.

263

10-1

10-2

10-3



264 **2.2.1 Orbit simulator**

265 The System Tool Kit (STK, available from AGI, <http://www.agi.com/products/>) is used to generate the S-5P  
266 orbit geometry and the geolocation of the edges of the swath as a function of time. Based on these  
267 characteristics, the location of the individual observations with a spatial distance of 7 km are generated. Time  
268 and longitude shifts are applied to the STK orbits to generate the orbits for the three Summer and three  
269 Winter months. Subsequently, the solar and viewing geometries are computed. Finally, segments of the orbits  
270 are maintained that have an overlap with the modelling domain.

271

272 **2.2.2 Cloud properties**

273 Cloud fields are obtained from the high-resolution operational weather forecast archive of the ECMWF.  
274 Meteorological fields of liquid water content, ice water content, specific humidity and cloud fraction are  
275 retrieved at a resolution of 0.25 x 0.25 degree for June-August 2003 and November 2003 - January 2004.  
276 These quantities are converted to cloud optical properties. The optical properties determine the reflectance,  
277 and are used to estimate effective cloud fractions and effective cloud top heights as would be retrieved from  
278 the satellite observations (Acarreta et al., 2004). The distribution of effective cloud fractions was compared  
279 with the distribution of effective cloud fractions **obtained from CMI observations**, and a reasonable  
280 agreement was found for Summer and Winter months.

281

282 These effective cloud fractions (and corresponding cloud radiance fractions) are used to provide weights to  
283 the cloud-free and cloud-covered fractions of the surface scene. The cloud altitude is used for the  
284 computation of the averaging kernel.

285

286 **2.2.3 Averaging kernel and measurement uncertainty lookup tables**

287 Because of the large number of observations that will become available from the S-5P instrument, full  
288 radiative transfer calculations for each observation separately is not feasible. We have chosen to build look-  
289 up tables for a set of geometries based on a radiative transfer code that employs the adding-doubling method  
290 in combination with optimal estimation (radiative transfer toolbox DISAMAR; de Haan, 2012). Look-up  
291 tables are set up for the averaging kernels (1D vectors as a function of altitude) and the measurement

11-1



292 uncertainty. Results are stored for a number of surface albedos, cloud/surface pressures, solar zenith angles,  
293 viewing zenith angles and relative azimuth angles. The look-up table details are provided in Table 1. Kernels  
294 are provided on 21 pressure levels between 1050.0 and 0.1 hPa. Uncertainties are specified for clear-sky and  
295 cloudy-sky separately.

296

297 Each simulation with DISAMAR consist of a forward calculation of the satellite-observed spectrum,  
298 followed by a retrieval step based on the optimal estimation method (Rodgers, 2000). Instrument noise, listed  
299 in Table 1, is converted into uncertainties for the retrieved CO column. A-priori trace gas profiles are taken  
300 from the CAMELOT study (Levelt et al., 2009). We assume that both the cloud and the surface are  
301 Lambertian reflectors. Kujanpää et al. (2015) provide further details of this procedure.

302

303 In particular the albedo is of major influence for the uncertainty, because it directly determines the signal  
304 observed by the instrument. This dependence is shown in Fig.2. Over land, albedo values are typically of the  
305 order of 0.1-0.2, with typical column errors of the order of 2 DU, or about  $10^{17}$  molecules  $\text{cm}^{-2}$ . Because  
306 typical CO columns over Europe are  $2 \cdot 10^{18}$  molecules  $\text{cm}^{-2}$ , this is a relatively small error of the order of 5%.

307 These numbers are in good agreement with the results presented in the CO ATBD of TROPOMI (document  
308 expected to be publicly released by the end of 2015). Over Sea, the albedo is very low, and the noise  
309 dominates the signal. In order to simulate this behaviour in a realistic way we have added the albedo values  
310 0.005, 0.01 and 0.02 to the albedo list.

311

312 We note that the uncertainties reported here are substantially lower than those reported for SCIAMACHY  
313 (e.g. Gloudemans et al., 2008). This reflects a difference in specifications of the instruments, and the fact that  
314 SCIAMACHY observations were hindered by ice build-up on the detectors. Real TROPOMI observations  
315 will show if the relatively small errors are realistic.

316

#### 317 **2.2.4 Synthetic observations generation**

318 The generation of the synthetic observations consists of the following steps:

- 319 • Collocation of the Nature run vertical profiles of CO to the locations of the observations.



- 320 • Computation of the effective cloud fraction, cloud radiance fraction, and cloud pressure from the  
321 ECMWF cloud fields collocated to the observations.
- 322 • Collocation of the NIR albedo map (Surface albedo at 2300 nm is interpolated from a climatology  
323 provided by SRON and based on SCIAMACHY observations (P. Tol, personal communication)) to  
324 the locations of the observations.
- 325 • Extract interpolated values for the observation kernel and uncertainties from the look-up table.
- 326 • Compute the synthetic observation from the inner product of the kernel with the nature run CO  
327 profile. This is done for both a clear sky and fully ~~clouded~~ situation, using the cloud pressure.
- 328 • Add a random noise amount to each observation, by drawing numbers from a Gaussian distribution  
329 with a width determined from the uncertainty estimate.

13-1

13-2

- 330 • Compute the partially clouded synthetic observation by weighting the clear and cloudy results with  
331 the cloud radiance fraction.

332

333 Over land, and in clear sky cases, the averaging kernel is close to 1, showing that the S-5P instrument is  
334 observing the vertical column to a good approximation (see Fig. 3). In cloud-covered cases the kernel equals  
335 0 for layers below the cloud pressure (yellow line in Fig. 3). For low-albedo cases (over ocean), Rayleigh  
336 scattering becomes non-negligible, and the kernel is decreasing towards the surface, but the noise is  
337 dominant in this case.

338

339 The results of this process is shown in Fig. 4. The figure demonstrates the high resolution of the NR (about 7  
340 km) and the corresponding simulated amount of detail. The bottom panel shows the corresponding CO  
341 observations. Over land the NR features are clearly present due to the relatively low uncertainty. Over the  
342 ocean and Mediterranean the signal is dominated by the noise. An improved information content is observed  
343 near Iceland, related to a thick cloud cover, where the higher signal reduces the relative noise.

344

### 345 **2.3 Pre-processing of S-5P CO total column observations**

346 This section describes the pre-processing of S-5P CO total column observations prior to assimilation into the  
347 MOCAGE model (Peuch et al., 1999) for the OSSE simulations. Using the MOCAGE model for the AR and

13-3



348 CR simulations avoids the identical twin problem associated with using the same model for both the NR and  
349 the OSSE simulations, which typically produces overoptimistic results (Arnold and Dey, 1986; Stoffelen et  
350 al., 2006). Section 2.4 provides further details of the MOCAGE model.

351

352 The S-5P will produce large amounts of data owing to its wide swath and relatively high spatial resolution of  
353 about  $7 \times 7 \text{ km}^2$ . Thus, a pre-processing step is necessary to reduce the data volume for the data assimilation  
354 experiments. For this study, we consider only pixels inside the OSSE simulation domain (Note that retrieval  
355 pixels in each single cross-track are essentially instantaneous measurements of CO<sub>2</sub>). This has the advantage  
356 of alleviating the data volume burden. However, a single cross-track over Europe could have more than  
357 80,000 valid retrieval pixels. Furthermore, each individual pixel is associated with an averaging kernel vector  
358 given at 34 vertical pressure levels, from the surface up to the top of the atmosphere (identified as 0.1 hPa).

359

360 Figure 3 shows an example of averaging kernels at the surface, as well as the averaging kernels  
361 representative of retrievals including pixels with different cloud fractions (less than 10%, greater than 30%,  
362 and greater than 80%). In addition, we discard data points with solar zenith angles larger than 80° or errors  
363 exceeding 20%. The retrieval over sea is noise-dominated. Because of this, we only consider CO partial  
364 columns above cloudy sea scenes with cloud fraction more than 80% and cloud top heights between the  
365 surface and 650 hPa. Finally, we apply a spatially weighted mean to bin the measurements into  $0.2^\circ \times 0.2^\circ$   
366 grid boxes ( $\sim 20 \times 15 \text{ km}$  at 45N), the assimilation model resolution; this is the set-up used for the OSSE  
367 assimilation experiments (CR and AR), and is described in El Amraoui et al. (2008a). It combines the  
368 MOCAGE model and the PALM (Projet d'Assimilation par Logiciel Multiméthode) data assimilation  
369 module. Sections 2.4-2.5 provide further details of the CR and AR set-ups.

370

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

372

$$\bar{c} = \frac{\sum_l w_l c_l}{\sum_l w_l}$$

373

374



375 where  $\bar{c}$  is the weighted average,  $c_i$  a single column measurement, and  $w_i (=1/\sigma_i^2)$  is the inverse of the  
376 variance corresponding to measurement  $c_i$ , and is the weight assigned to this single measurement. The  
377 inverse of the variance associated with the weighted average is  
378

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

379  
380

381 The spatial binning not only reduces considerably the data volume but also results in an improved spatial  
382 representativeness of the CO measurements by reducing the random error of each data pixel.  
383

#### 384 **2.4 The Control Run**

385 To generate the CR, it is important to use a state-of-the-art modelling system, which simulates the  
386 observational data representing, for example, a current operational observational system. An important  
387 requirement for an effective OSSE is to generate the CR with a model different to the one used to construct  
388 the NR to avoid the identical twin problem (see Sect. 2.3). If the model from which we extract hypothetical  
389 observations is the same as the assimilating model, the OSSE results tend to show unrealistic observation  
390 impact and overly optimistic forecast skill (Arnold and Dey, 1986; Stoffelen et al., 2006). Consequently, by  
391 using two independent models the OSSE will more realistically simulate the assimilation of real  
392 observations. This follows our guiding principle to design an OSSE that is not too overoptimistic.

393

394 As mentioned in Sect. 2.3, we use the MOCAGE model to generate the CR. In this OSSE study, the CR is a  
395 free model run. The MOCAGE model is a three-dimensional CTM developed at Météo France (Peuch et al.,  
396 1999) providing the evolution of the atmospheric composition in accordance with dynamical, physical and  
397 chemical processes. It provides a number of configurations with different domains and grid resolutions, as  
398 well as various chemical and physical parameterization packages. Current use of MOCAGE includes several  
399 applications: e.g., the Météo-France operational chemical weather forecasts (Dufour et al., 2004); the  
400 Monitoring Atmospheric Composition and Climate (MACC) services (<http://www.gmes-atmosphere.eu>;  
401 Mérécal et al., 2015); and studies of climate trends of atmospheric composition (Teyssède et al., 2007).



402 Validation of MOCAGE simulations against a large number of measurements took place during the  
403 Intercontinental Transport of Ozone and Precursors (ICARTT/ITOP) campaign (Bousserez et al., 2007).

404

405 In this study, we use a two-way nesting configuration to generate the CR and the AR (we describe the AR  
406 set-up in Sect. 2.5): a global grid with a horizontal resolution of 2x2 degrees and a regional grid (5W-35E,  
407 35N-70N) with a horizontal resolution of 0.2x0.2 degrees. The MOCAGE model includes 47 sigma-hybrid  
408 vertical levels from the surface up to 5 hPa. The vertical resolution is 40 to 400 m in the boundary layer (7  
409 levels) and about 800 m in the neighbourhood of the tropopause and in the lower stratosphere. The chemical  
410 scheme used is RACMOBUS, which combines the stratospheric scheme REPROBUS (Lefèvre et al., 1994)  
411 and the tropospheric scheme RACM (Stockwell et al., 1997). The RACMOBUS scheme includes 119  
412 individual species, of which 89 are prognostic variables, and considers 372 chemical reactions.

413

414 We force the CR (and the AR) every 3 hours with the ARPEGE analysis (Courtier et al., 1991). We prescribe  
415 the surface anthropogenic emission using the MACC-I emission database ([https://gmes-  
416 atmosphere.eu/about/project\\_structure/input\\_data/d\\_emis/](https://gmes-atmosphere.eu/about/project_structure/input_data/d_emis/)). We do not include the fire emissions in the  
417 CR and AR experiments described in this paper, **as they are a priori not known**. This means that any  
418 signature of fire emissions in the AR (see Sect. 2.5) can only come from assimilation of the CO  
419 measurements. Note that for the NR, the surface anthropogenic emissions come from the MACC-II  
420 inventory, helping to differentiate the CR from the NR. As for the NR, the CR has a spin-up period of three  
421 months.

422

## 423 **2.5 The Assimilation run**

424 We assimilate simulated S-5P total column CO observations derived from the LOTOS-EUROS NR into the  
425 MOCAGE CTM at a 0.2° spatial resolution using the MACC extended domain (5W-35E, 35N-70N). The  
426 assimilation system used in this study is MOCAGE-PALM (e.g., El Amraoui et al., 2008a) developed jointly  
427 by Météo-France and CERFACS (Centre Européen de Recherche et de Formation Avancée en Calcul  
428 Scientifique) in the framework of the ASSET European project (Lahoz et al., 2007b). The assimilation  
429 module used in this study is PALM, a modular and flexible software, which consists of elementary  
430 components that exchange data (Lagarde et al., 2001). It manages the dynamic launching of the coupled





17-1  
431 components (forecast model, algebra operators and input/output of observational data) and the parallel data  
432 exchanges. Massart et al. (2009) used the assimilation system MOCAGE-PALM to assess the quality of  
433 satellite ozone measurements. The MOCAGE-PALM assimilation system also helps identify and overcome  
434 model deficiencies. In this context, its assimilation product has been used in many atmospheric studies in  
435 relation to ozone loss in the Arctic vortex (El Amraoui et al., 2008a); tropics/mid-latitudes exchange  
436 (Bencherif et al., 2007); stratosphere-troposphere exchange (Semane et al., 2007); and exchange between the  
437 polar vortex and mid-latitudes (El Amraoui et al., 2008b). For this OSSE, to speed up the assimilation  
438 process we use the 3D-Var version of PALM. In the OSSE, the MOCAGE model provides the CR and by  
439 assimilating the simulated CO data from the NR, the MOCAGE model provides the AR. Thus, we produce  
440 the CR and AR outputs with a model different from that used to produce the NR (see Sect. 2.1).

441

442 A key element of the data assimilation system is the background error covariance matrix (the **B**-matrix)  
443 (Bannister, 2008). It has a large impact on the 3D-Var analysis used in this study and, thus, it is important to  
444 use a form of **B** that is as realistic as possible. In MOCAGE-PALM, we base the **B**-matrix formulation on the  
445 diffusion equation approach (Weaver and Courtier, 2001). It can be fully specified by means of the 3-D  
446 standard deviation field (square root of the diagonal elements of **B**, in concentration units or as a percentage  
447 of the background field) and 3-D fields of the horizontal ( $L_x$  and  $L_y$ ) and vertical ( $L_z$ ) local correlation  
448 length-scales. We can estimate the **B**-matrix elements more efficiently using an ensemble method (Bannister,  
449 2008). This technique consists of feeding an ensemble of states through the data assimilation system to  
450 simulate the important sources of error. However, this approach is time-consuming and, therefore, not used in  
451 this study.

452

453 For this study, we use a simple parameterization for the **B**-matrix:  $L_x$  and  $L_y$  are assumed homogeneous and  
454 equal to 35 km (about two model grid lengths);  $L_z$  is constant and set to one vertical model layer. As in Emili  
455 et al. (2014), the background standard deviation 3-D field is parameterized as a vertically varying percentage  
456 of the background profile, which decreases from values of 25% at the surface to values of 15% in the upper  
457 troposphere, and decreases further throughout the stratosphere to values of 5% in the upper stratosphere (not  
458 shown). We base these settings on several 1-day assimilation trials; they ensure reasonable values of standard  
459 self-consistency tests, e.g., providing chi-squared ( $\chi^2$ ) values close to 1 (see Fig. 3 in Sect. 3.1). Furthermore,



460 a value of  $L_x$  and  $L_y$  of 35 km corresponds to more than one grid length of the model, allowing the model to  
461 resolve these features. The data assimilation procedure will weight both the observations and the model 1-  
462 hour forecasts (from the last analysis point), and will update locations not coincident with the observations  
463 through the correlation length-scales. Table 2 summarizes the parameters used for the assimilation  
464 experiments.

### 465 **3. Results**

#### 466 **3.1 Evaluation of the assimilation run**

467 In this section, we evaluate the impact of the assimilation of the S-5P CO total column. First, we evaluate the  
468 consistency of the assimilation run by separating the clear-sky pixels from their cloudy counterparts (Sect.  
469 3.1.1). Second, to further understand the impact on the surface CO field of the simulated S-5P CO total  
470 column measurements, we investigate the analysis increment ( $\delta x$ ) to provide a quantitative diagnostic of the  
471 quality of the analysis for a selected date, 15 June 2003 (Sect. 3.1.2).

472

##### 473 **3.1.1 Consistency of the assimilation run**

474 We have performed two OSSEs. The first one includes all pixels in the OSSE domain, regardless of whether  
475 they are cloudy or clear-sky and the second only includes clear-sky pixels. A pixel is considered as clear  
476 when the cloud fraction is less than 10%. Comparison of the ARs from these two OSSEs indicated that the  
477 impact of including all pixels is small. The largest differences between the respective ARs in relation to the  
478 NR are 4% in regions over North Europe (North Sea and Scandinavia), with the AR for clear-sky pixels  
479 closer to the NR (not shown). We can explain these results by the fact the summer generally has low amounts  
480 of cloud. Consequently, we only present the results from the OSSE with all pixels.

481 To evaluate the AR, we calculate the  $\chi^2$  diagnostic associated with the Observation minus Forecast (OmF)  
482 differences (see, e.g., Lahoz et al., 2007a). Here, we normalize the OmF differences by the background error.  
483 We also calculate histograms of the Observation minus Analysis (OmA) differences, the observation and the  
484 simulation from the CR (observation-minus-control run, hereafter OmC) differences, and the OmF



485 differences. We use the observational error to normalize the differences building the histograms of OmA,  
486 OmC and OmF.

487

488 Figure 5 (top panel) shows the chi-squared time-series for OmF and its associated auto-correlation function  
489 calculated over the three-month period of the OSSE experiments, computed as daily averages. The chi-  
490 squared diagnostic starts with a maximum of about 1.56, and takes values down to 0.75, with a mean of 0.9  
491 over the OSSE three-month period. The chi-squared time-series is nearly stable since it exhibits relatively  
492 small variability (a standard deviation of about 0.14). Furthermore, the auto-correlation of the chi-squared  
493 statistic drops to zero, with no correlation after a time delay of 20 days. The calculation of the auto-  
494 correlation shows that the chi-squared statistic is uncorrelated after a time lag of 20 days; this means that  
495 after this time the mathematical expectation  $E(\chi^2)$  is equal to the average of the chi-squared statistics. We  
496 find  $E(\chi^2) = 0.90$ , which is close to the theoretical value of 1 (see Lahoz et al., 2007a). This result indicates  
497 that the a priori error statistics as represented in the **B**-matrix slightly overestimate the actual error statistics  
498 from the OmF differences.

499

500 To test whether the observations, forecast and analysis fields, and their associated errors, are consistent with  
501 each other, we calculate the histograms of OmA, OmF and OmC only over land (normalized by the  
502 observation error) over the three-month period (Fig. 5, bottom panel). For a properly set up assimilation  
503 system, the OmF and OmA normalized histograms should be close to a Gaussian distribution with mean zero  
504 and standard deviation one. Figure 3 (bottom panel) shows that the OmA and OmF differences are close to a  
505 Gaussian distribution centred near to or at zero. The OmF has a mean and standard deviation of 0.10 and  
506 1.73, respectively, whereas the OmA has nearly a zero mean and a standard deviation of 1.05. This indicates  
507 that the centre of the OmA histogram is closer to zero and more peaked than the histogram of OmF. We  
508 expect this, since the analyses should be closer to the observations than the forecasts. Furthermore, the  
509 histogram for OmA indicates that the errors in the **R**-matrix, the observational counterpart of the **B**-matrix,  
510 are a good representation of the analysis error.

511

512 Based on the above results, we conclude that the background error covariance matrix, **B**, and its  
513 observational counterpart, **R**, prescribed in our assimilation system are reasonably well characterized (see,



514 e.g., Lahoz et al., 2007a, for a discussion of the specification of errors in a data assimilation system).  
515 Furthermore, the above results are consistent with the assumption that the errors in the observations and the  
516 forecasts are Gaussian.

517

518 The shape of the OmC normalized histogram, which has a mean and standard deviation of 2.36 and 5.60,  
519 respectively, indicates the presence of a relatively large bias between the S-5P observations and the CR. The  
520 assimilation reduces this bias, as shown by the analyses being significantly closer to the observations than  
521 the simulation from the CR. This shows that the assimilation of simulated S-5P CO total column  
522 observations has a significant impact on the CO forecasts and analyses.

523

### 524 **3.1.2 Study of increments**

525 To understand further the impact on the surface CO field of the simulated S-5P CO total column  
526 measurements, we calculate the analysis increment ( $\delta x$ ) for a single analysis time at 14:00 UTC on 15 June  
527 2003. We calculate this increment as the analysis minus the model first guess (1-hour forecast). The analysis  
528 increment provides a quantitative diagnostic of the quality of the analysis (see, e.g., Fitzmaurice and Bras,  
529 2008).

530

531 Figure 6 (top panel) shows the spatial distribution of  $\delta x$  at the model surface. One can see the spread of the  
532 impact of the simulated observations across large regions. This is owing to S-5P having a wide swath  
533 allowing it to sample larger regions. The most substantial corrections are over land, where there are sufficient  
534 observations to have an impact. Over sea, the increments tend to be negligible, as any observations found  
535 there have relatively large errors. Thus, there will not be much difference between the model first guess and  
536 the analysis. Likewise, this is also true in the regions outside the satellite footprint.

537

538 To provide further insight into the impact of S-5P CO measurements, we calculate latitude-height and  
539 longitude-height cross-sections at 48.8N, 2.6E, near Paris, for 15 June 2003. Figure 4 (bottom left and  
540 bottom right panels) shows a zoom of the zonal and meridional vertical slices of the analysis increment. We  
541 see significant corrections to the model first guess (identified by large increments) confined to a deep layer.  
542 These corrections are larger at the surface, and exhibit a second maximum around 650 hPa. This vertical



21-1 543 structure is mainly attributable to the forecast error standard deviation (given as a vertically varying fraction  
544 of the local CO mixing ratio), the square root of the diagonal entry of the **B**-matrix, and which is higher in  
545 the boundary layer (where the value of the S-5P CO averaging kernel is close to 1). **The shape of the S-5P**  
546 **analysis increments also exhibits a second peak around 650 hPa.** The increments for this particular day thus  
547 show a clear impact from the S-5P CO measurements in the PBL and the free troposphere.

548  
549 The shape of the S-5P increments is similar to that of typical SCIAMACHY analysis increments, which also  
550 extend through a deep layer and have a maximum at the surface (Tangborn et al., 2009). The fact that both  
551 these analysis increments stretch out over a deep layer is owing to similarities in the S-5P and SCIAMACHY  
552 averaging kernels - both are close to unity over cloud-free land (see Fig. 4). Note that the situation shown in  
553 Fig. 6 is a snapshot and depends on the particular conditions for this time. An average of the increments over  
554 the summer period would tend to show a uniform distribution in height.

555

## 556 **3.2 Evaluation of the summer OSSE**

### 557 **3.2.1 Summer averages**

558 Figure 7 shows the fields of surface CO from the CR, and the NR and the AR, averaged over the northern  
559 summer period. One can see the general change of CO over land between the CR (top left panel) and the AR  
560 (bottom panel). We can ascribe this to the contribution of simulated S-5P total column CO data sampled from  
561 the NR. This figure shows several differences between the CR and AR fields that indicate the superior

21-4 562 behaviour of the AR in capturing features in the NR. For example, over Eastern Europe and Russia, the AR

563 **CO concentration values are closer to those in the NR;** in particular, the CR shows generally lower values  
564 than in the NR. Nevertheless, over Portugal, where the NR shows the forest fires that occurred over the  
565 summer, the AR captures them only slightly better than the CR. We expect the relatively poor performance of

21-2 566 the CR regarding fires, as the fires are not included in the CR set-up (see Sect. 2.4). Although the AR

21-3 567 captures the forest fires slightly better than the CR (through assimilation of CO measurements), **the relatively**

568 **poor temporal resolution of the S-5P ultimately limits its performance.** A geostationary satellite, **given its**

569 **relatively high temporal resolution,** should be able to capture better the temporal variability of these forest  
570 fires (Edwards et al., 2009).

571



572 **3.2.2 Statistical metrics**

573 In this section, we provide a quantitative assessment of the benefit from S-5P CO total column measurements  
574 on the CO surface analysis. For this, we perform a statistical analysis of the different OSSE experiments for  
575 northern summer 2003.

576

577 We calculate the mean bias (MB, in parts per billion by volume, ppbv), its magnitude reduction (MBMR,  
578 ppbv), and the root mean square error (RMSE, ppbv), and its reduction rate (RMSERR, %). Note that  
579 although recent papers have raised concerns over the use of the RMSE metric (Willmott and Matsuura, 2005;  
580 Willmott et al., 2009), Chai and Draxler (2014) discuss circumstances where the RMSE is more beneficial.  
581 We use the correlation coefficient,  $\rho$  to measure the linear dependence between two datasets, and the fraction  
582 of the true variability (i.e., variability represented by the NR) reproduced by the CR or AR.

583

584 For a single model grid box, we define the statistical metrics (MB, RMSE,  $\rho$ ) with respect to the NR as:

585

586 
$$MB(X) = \frac{1}{N} \sum (X - NR)$$

587

588 
$$MBMR = |MB(CR)| - |MB(AR)|$$

589

590 
$$RMSE(X) = \sqrt{\frac{1}{N} \sum (X - NR)^2}$$

591

592 
$$RMSERR = 100 \times \left( 1 - \frac{RMSE(AR)}{RMSE(CR)} \right)$$

593

594 
$$\rho(X) = \frac{\sum (X - \bar{X})(NR - \bar{NR})}{\sqrt{\sum (X - \bar{X})^2 \sum (NR - \bar{NR})^2}}$$

595



596 where X denotes the CR or the AR; N is the number of data samples; the vertical bars denote the absolute  
597 value operator; and the overbar symbol represents the arithmetic mean operator. The MB metric gives the  
598 average value by which the CR or the AR differs from the NR over the entire dataset.

599

### 600 **3.2.3 Results of the statistical tests**

601 Figure 8 presents the zonal and meridional means of the difference between the CR and the AR averaged  
602 over the northern summer 2003 (1 June – 31 August). We also plot the confidence interval representing the  
603 areas where the AR is not significantly different to the CR at the 99% confidence limit (highlighted in the  
604 grey colour). These two figures show that there is benefit from the S-5P CO total column data over the first  
605 few bottom levels of the troposphere, i.e., the lowermost troposphere. Between the surface and 800 hPa, a  
606 negative peak is present in the zonal difference field (over Scandinavia), and in the meridional difference  
607 field (over Eastern Europe). Note that the zonal field shows two areas, one with positive values and the other  
608 with negative values representing a CR greater than the AR and a CR smaller than the AR, respectively. The  
609 positive peak, at a slightly higher level (i.e., lower pressure) than the negative peak, is representative of the  
610 Mediterranean Sea, whereas the negative peak is more representative of the land areas (Scandinavia and  
611 Eastern Europe). Figure 8 indicates that the S-5P CO corrects the model in the lower troposphere with a  
612 larger impact over land and with a less large impact in the PBL. This is consistent with the behaviour of the  
613 analysis increments shown in Fig. 6.

614

615 Figure 9 shows the performance of the biases between the CR and the NR, and the AR and the NR at the  
616 surface, and averaged over the northern summer of 2003 (1 June – 31 August). The MBMR, which compares  
617 the magnitude of the CR vs NR and AR vs NR biases, indicates the geographical areas where the simulated  
618 S5P CO total column data have the most impact. The MBMR shows that the AR is closer to the NR than the  
619 CR, almost everywhere in the domain (reflected by the prevalence of the red colours in the bottom left  
620 panel). This indicates that the simulated S-5P CO total column data generally provide a benefit at the surface,

621 **and especially over land areas where the CO sources are sparse.**

622

623 We also calculate the RMSE as well as the reduction rate of the RMSE, RMSERR (Figure 10), both keeping  
624 the systematic error (Fig. 10 top), and removing the systematic error (Fig. 10 bottom). We calculate the



24-2

24-1

625 **systematic error** in the AR and CR by subtracting the NR field from each of them, producing a ~~debiased~~ AR  
626 and CR. For the case where we remove the systematic error, we perform the statistics on the debiased AR  
627 and CR. If we examine the RMSE statistics, Fig. 8 shows that the CR gets closer to the NR over the Atlantic  
628 Ocean and over the Eastern domain including Russia and Scandinavia, when we remove the systematic error.  
629 For example, over these areas we obtain ~30 ppbv and ~10 ppbv for the RMSE keeping and removing the  
630 systematic error, respectively. For the reduction of the RMSE, RMSERR, the behaviour for the CR is similar  
631 overall, showing a reduction rate of 60% and 30-45% keeping and removing the systematic error,  
632 respectively. Note that over Scandinavia the reduction rate goes down from 60% to about 10% after  
633 removing the systematic error.

634

635 These results indicate that S-5P CO data show more benefit when keeping the systematic error in the  
636 calculation of the RMSE. Following our guiding principle of avoiding an overoptimistic OSSE, we consider  
637 only the values of RMSE obtained when we remove the systematic error. For this case, the average reduction  
638 rate for the AR is around 20-25% over land (except Scandinavia) and close to 10% over sea and over  
639 Scandinavia.

640

641 In Figure 11, we show the correlation between the CR and the NR, and the correlation between the AR and  
642 the NR, at the surface for the three northern summer months (1 June – 31 August). Figure 11 shows that the  
643 AR is closer than the CR to the NR with the correlation coefficient reaching 0.9 over land. By contrast, the  
644 correlation coefficient between the CR and the NR is typically less than 0.5, with very low values over  
645 Eastern Europe, where CO sources are sparse.

646

#### 647 **3.2.4 Time-series at selected locations**

648 Figure 12 shows time-series from the NR, the CR and the AR over the three areas of the study domain  
649 represented by the squares shown in Figs. 9 (bottom panel) and 10 (right panels). (i) The Paris region (Fig.  
650 12, top panel); (ii) a region over Portugal, where forest fires occurred during the northern summer (Fig. 12,  
651 middle panel); and (iii) an area in the Eastern part of the study domain, where the reduction of RMSE (i.e.,  
652 RMSERR) was much larger than for other regions (Fig. 12, bottom panel). For all three areas, the AR is  
653 generally closer to the NR than the CR, showing the impact of the simulated observations. We calculate the





654 biases between the AR and CR vs the NR by computing the difference  $NR-X$ , where  $X$  is AR or CR, and  
655 normalizing by the number of observations over the northern summer period. The biases are: (i) Paris region,  
656 CR: 48 ppbv, AR: 38 ppbv; (ii) Portugal, CR: 101 ppbv, AR: 83 ppbv; (iii) Eastern part of domain: CR: 21  
657 ppbv, AR: 5 ppbv.

658

25-1 659 Over Paris (top panel), the CR is already close to the NR and the impact of the S-5P CO simulated

25-2 660 observations is small. Over Portugal (middle panel), the presence of fires is not seen in the CR (e.g., a

661 maximum of CO at the beginning of the heat wave), as the fires were not taken into account in the CR as

662 they are not known a priori (see Sect. 2.4). In contrast, over this specific location we see the impact of the

663 fires on the CO concentrations in the AR with, however, much lower values than for the NR. During the

664 fires, the CO concentrations in the AR over Portugal were larger than 500 ppbv, whereas the CR remained

25-3 665 relatively unchanged with concentrations less than 200 ppbv. Over the Eastern part of the study area (bottom

666 panel), the temporal variability is not high and the magnitude of the bias between the CR and the NR is

667 small, but it is removed in the AR.

668

669 **3.2.5 Sensitivity tests for fire episode**

25-4 670 The assimilation system we use has a default criterion to discard CO column observations with values larger

671 than 75% of the MOCAGE value. This criterion is not appropriate to situations resulting in excessive values

672 in the CO concentrations, as is the case for forest fires. To understand further the performance of the OSSE

673 over the period of the Portugal forest fires we perform a second OSSE without this default criterion. This

674 second OSSE covers the period of the forest fires (25 July – 15 August). For this second OSSE, we compare

675 the total column values and the surface values of the CO fields from the CR and the AR (Figs. 13-15,

676 respectively).

677

678 Figure 13 shows the CO total column at 14:15 UTC on 4 August 2003 (during the period of the Portugal

679 forest fires) from the NR (top left panel); the simulated S-5P observations (top right panel); the CR (bottom

680 left panel); and the AR (bottom right panel). We can see that the AR captures the fire event, indicated by

681 relatively high values of the CO total column over Portugal, whereas the CR does not. This confirms the

682 results shown in Fig. 12, which highlight the benefit provided by the S-5P CO total column measurements, in



683 particular regarding the capture of the signature of the Portugal forest fires. Note that the S-5P measurement  
684 is noise-dominated over the sea (top right panel). This accounts for the sharp edge in the CO total column  
685 field seen between the Iberian Peninsula and the Bay of Biscay for the AR (bottom right panel).

686

687 Figure 14 shows the time-series of the surface CO concentrations over the period 25 July – 15 August (that  
688 of the Portugal forest fires). In comparison to the original OSSE (see middle panel of Fig. 12), the AR is now  
689 closer to the NR, having now peak values of about 900 ppbv, instead of peak values of about 550 ppbv. The  
690 CR still has peak values less than 200 ppbv. This indicates that the relatively low values in the AR (in  
691 comparison to the NR) for the original OSSE shown in the middle panel of Fig. 12 result from the  
692 application of the default criterion to discard CO column observations that are far away from MOCAGE  
693 values. The results from Fig. 14 confirm those shown in Fig. 13, and reinforce the benefit provided by the S-  
694 5P CO total column measurements, in particular regarding the capture of the signature of the Portugal forest  
695 fires.

696

#### 697 **4. Conclusions**

698 We perform a regional-scale Observing System Simulation Experiment (OSSE) over Europe to explore the  
699 impact of the LEO satellite mission S-5P carbon monoxide (CO) total column measurements on lowermost  
700 tropospheric air pollution analyses, with a focus on CO surface concentrations and the Planetary Boundary  
701 Layer (PBL). The PBL varies in depth throughout the year, but is contained within the lowermost  
702 troposphere (heights 0-3 km), and typically spans the heights 0-1 km. We focus on northern summer 2003,  
703 which experienced a severe heat wave with severe societal impact.

704

705 This OSSE study provides insight on the impact from LEO S-5P CO measurements on surface CO  
706 information. We perform the standard steps of an OSSE for air quality. (i) Production of a Nature Run, NR.  
707 (ii) Test of the realism of the NR. (iii) Different models to produce, on the one hand, the NR, and on the other  
708 hand, the OSSE experiments to create the Control Run, CR, and the Assimilation Run, AR. (iv) Calculation  
709 of synthetic observations, observation uncertainty, and averaging kernels to represent sensitivity of the  
710 observations in the vertical. (v) Quantitative evaluation of the OSSE results, including performing statistical

26-1



711 significance tests, and self-consistency and chi-squared tests. Based on the specifications of the TROPOMI  
712 instrument, relatively low CO column uncertainties of around 5% are anticipated over the European  
713 continent.

714

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

721

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

735

736 We further show that, locally, the AR is capable of reproducing the peak in the CO distribution at the surface  
737 due to forest fires (albeit, weaker than the NR signal), even if the CR does not have the signature of the fires  
738 in its emission inventory. A second OSSE shows that this relatively weak signal of the forest fires in the AR  
739 arises from the use of a default criterion to discard CO total column observations too far from model values,



740 a criterion not appropriate to situations resulting in excessive values in the CO concentrations, as is the case  
741 for forest fires. This second OSSE shows a much stronger signal in the AR, which is now much closer to the  
742 NR than the CR, confirming the benefit of S-5P CO total column measurements.

743

744 Further work will involve extending the OSSE approach to other S-5P measurements, such as ozone total  
745 column, and NO<sub>2</sub> and formaldehyde tropospheric columns. These studies will complement similar studies on  
746 the benefit from Sentinel-4 and -5 measurements. Collectively, these OSSE studies will provide insight into  
747 the relative benefits from the Sentinel-4, -5 and -5P platforms for monitoring atmospheric pollution  
748 processes.

749

## 750 **5. Acknowledgments**

751 Support for this work came partly from the ESA funded project “Impact of Spaceborne Observations on  
752 Tropospheric Composition Analysis and Forecast” (ISOTROP –ESA contract number 4000105743). WAL  
753 acknowledges support from an internal project from NILU. RA, JLA, PR, LE and WAL acknowledge  
754 support from the RTRA/STAE. JK and JT acknowledge support from the Academy of Finland (Project no.  
755 267442).

756



## 757 **6. References**

- 758 Acarreta, J. R., J. F. De Haan, and P. Stammes (2004), Cloud pressure retrieval using the O<sub>2</sub>-O<sub>2</sub> absorption  
759 band at 477 nm, *J. Geophys. Res.*, 109, D05204, doi:10.1029/2003JD003915.
- 760 Arnold, C.P., Jr., and C.H. Dey, 1986: Observing-systems simulation experiments: Past, present and future.  
761 *Bull. Amer. Meteorol. Soc.*, 67, 687–695.
- 762 Atlas, R. 1997: Atmospheric observation and experiments to assess their usefulness in data assimilation. *J.*  
763 *Meteor. Soc. Jpn.*, 75, 111–130.
- 764 Atlas, R., G.D. Emmitt, Terry, E. Brin, J. Ardizzone, J.C. Jusem, et al., 2003: Recent observing system  
765 simulation experiments at the NASA DAO, in Preprints, 7<sup>th</sup> Symposium on Integrated Observing  
766 Systems (Long Beach, CA: American Meteorological Society).
- 767 Bannister, R.N., 2008: A review of forecast error covariance statistics in atmospheric variational data  
768 assimilation. I: characteristics and measurements of forecast error covariances. *Q. J. R. Meteorol. Soc.*,  
769 134, 1951–1970. doi: 10.1002/qj.339.
- 770 Barbosa, P., J. San-Miguel-Ayanz, A. Camia, M. Gimeno, G. Liberta, and G. Schmuck, 2004: Assessment of  
771 fire damages in the EU Mediterranean Countries during the 2003 Forest Fire Campaign. Official  
772 Publication of the European Commission, S.P.I.04.64, Joint Research Center, Ispra, 2004.
- 773 Barré, J., D. Edwards, H. Worden, A. Da Silva, and W. Lahoz, 2015: On the feasibility of monitoring air  
774 quality in the lower troposphere from a constellation of northern hemisphere geostationary satellites  
775 (Part 1). *Atmos. Env.*, 113, 63-77, doi: 10.1016/j.atmosenv.2015.04.069.
- 776 Bencherif, H., L. El Amraoui, N. Semane, S. Massart, D.C. Vidyaranya, A. Hauchecorne, and V.-H. Peuch,  
777 2007: Examination of the 2002 major warming in the southern hemisphere using ground-based and  
778 Odin/SMR assimilated data: stratospheric ozone distributions and tropic/mid-latitude exchange. *Can. J.*  
779 *Phys.*, 85, 1287–1300.
- 780 Bousserrez, N., J.L. Attié, V.-H. Peuch, M. Michou, G. Pfister, D. Edwards, L. Emmons, C. Mari, B.  
781 Barret, S.R. Arnold, A. Heckel, A. Richter, H. Schlager, A. Lewis, M. Avery, G. Sachse, E.V.  
782 Browell, and J.W. Hair, 2007: Evaluation of the MOCAGE chemistry transport model during  
783 the ICARTT/ITOP experiment, *J. Geophys. Res.*, 112, D10S42, doi: 10.1029/2006JD007595.
- 784 Buchwitz, M., R. de Beek, S. Noël, J.P. Burrows, H. Bovensmann, O. Schneising, I. Khlystova, M. Bruns, H



- 785 Bremer, P. Bergamaschi, S. Körner and M. Heimann Atmospheric carbon gases retrieved from  
786 SCIAMACHY by WFM-DOAS: version 0.5 CO and CH<sub>4</sub> and impact of calibration improvements on  
787 CO<sub>2</sub> retrieval Atmospheric Chemistry and Physics 6, 2727–2751, 2006
- 788 Chai, T., and R.R. Draxler, 2014: Root mean square error (RMSE) or mean absolute error (MAE)?  
789 Arguments against avoiding RMSE in the literature. Geosci. Model Dev., 7, 1247-1250.
- 790 Claeysman, M., J.-L. Attié, V.-H. Peuch, L. El Amraoui, W.A. Lahoz, B. Josse, M. Joly, J. Barré, P. Ricaud, S.  
791 Massart, A. Piacentini, T. Von Clarmann, M. Höpfner, J. Orphal, J.-M. Flaud and D.P. Edwards, 2011: A  
792 thermal infrared instrument onboard a geostationary platform for CO and O<sub>3</sub> measurements in the  
793 lowermost troposphere: Observing System Simulation Experiments. Atmos. Meas. Tech., 4, 1637-1661.
- 794 Courtier, P., C. Freydier, J. Geleyn, F. Rabier, and M. Rochas, 1991: The ARPEGE project at Météo France,  
795 in: Atmospheric Models, vol.2, pp. 193–231, Workshop on Numerical Methods, Reading, UK, 1991.
- 796 Curier, R.L., R. Timmermans, S. Calabretta-Jongen, H. Eskes, A. Segers, D. Swart, and M. Schaap, 2012:  
797 Improving ozone forecasts over Europe by synergistic use of the LOTOS-EUROS chemical transport  
798 model and in-situ measurements. Atmos. Env., 60, 217-226, doi:10.1016/j.atmosenv.2012.06.017.
- 799 Cuvelier, C., P. Thunis, R. Vautard, M. Amann, B. Bessagnet M. Bedogni, R. Berkowicz, J. Brandt, F.  
800 Brocheton, P. Builtjes, A. Coppalle, B. Denby, G. Douros, A. Graf, O. Hellmuth, C. Honoré, A. Hodzic,  
801 J. Jonson, A. Kerschbaumer, F. de Leeuw, E. Minguzzi, N. Moussiopoulos, C. Pertot, G. Pirovano, L.  
802 Rouil, M. Schaap, R. Stern, L. Tarrason, E. Vignati, M. Volta, L. White, P. Wind and A. Zuber, 2007:  
803 CityDelta: A model intercomparison study to explore the impact of emission reductions in European  
804 cities in 2010, Atmos. Env., 41, 189-207, doi: 10.1016/j.atmosenv.2006.07.036.
- 805 Dufour, A., M. Amodei, G. Ancellet, and V. H. Peuch, 2004: Observed and modeled “chemical weather”  
806 during ESCOMPTE. Atmos. Res., 74, 161–189.
- 807 Edwards, D. P., L. K. Emmons, J. C. Gille, A. Chu, J.-L. Attié, L. Giglio, S. W. Wood, J. Haywood, M. N.  
808 Deeter, S. T. Massie, D. C. Ziskin, and J. R. Drummond (2006), Satellite Observed Pollution From  
809 Southern Hemisphere Biomass Burning, J. Geophys. Res., 111, D 14312, doi:10.1029/2005JD006655.
- 810 Edwards, D. P., L. K. Emmons, D. A. Hauglustaine, A. Chu, J. C. Gille, Y. J. Kaufman, G. Pétron, L. N.  
811 Yurganov, L. Giglio, M. N. Deeter, V. Yudin, D. C. Ziskin, J. Warner, J.-F. Lamarque, G. L. Francis, S. P.  
812 Ho, D. Mao, J. Chan, and J. R. Drummond (2004), Observations of Carbon Monoxide and Aerosol



- 813 From the Terra Satellite: Northern Hemisphere Variability, *J. Geophys. Res.*, 109, D24202,  
814 doi:10.1029/2004JD0047272004
- 815 Edwards, D.P., A.F. Arellano Jr., and M.N. Deeter, 2009: A satellite observation system simulation  
816 experiment for carbon monoxide in the lowermost troposphere. *J. Geophys. Res.*, 114, D14304, doi:  
817 10.1029/2008JD011375.
- 818 El Amraoui, L., V.-H. Peuch, P. Ricaud, S. Massart, N. Semane, H. Teyssèdre, D. Cariolle, and F. Karcher,  
819 2008a: Ozone loss in the 2002/03 Arctic vortex deduced from the Assimilation of Odin/SMR O<sub>3</sub> and  
820 N<sub>2</sub>O measurements: N<sub>2</sub>O as a dynamical tracer. *Q. J. R. Meteorol. Soc.*, 134, 217–228.
- 821 El Amraoui, L., N. Semane, V.-H. Peuch, and M.L. Santee, 2008b: Investigation of dynamical processes in  
822 the polar stratospheric vortex during the unusually cold winter 2004/2005. *Geophys. Res. Lett.*, 35,  
823 L03803, doi: 10.1029/2007GL031251.
- 824 Elbern, H., A. Strunk, and L. Nieradzki, 2010: “Inverse modelling and combined state-source estimation for  
825 chemical weather,” in *Data Assimilation: Making Sense of Observations*, eds W.A. Lahoz, B. Khattatov,  
826 and R. Ménard (Berlin: Springer), 491–513.
- 827 Emili, E., B. Barret, S. Massart, E. Le Flochmoen, A. Piacentini, L. El Amraoui, O. Pannekoucke, and D.  
828 Cariolle, 2014: Combined assimilation of IASI and MLS observations to constrain tropospheric and  
829 stratospheric ozone in a global chemical transport model. *Atmos. Chem. Phys.*, 14, 177–198, doi:  
830 0.5194/acp-14-177-2014.
- 831 Fitzmaurice, J., and R.L. Bras, 2008: Comparing Reanalyses Using Analysis Increment Statistics. *J.*  
832 *Hydrometeor.*, 9, 1535–1545.
- 833 Galli, A., A. Butz, R.A. Scheepmaker, O. Hasekamp, J. Landgraf, P. Tol, D. Wunch, N. M. Deutscher, G.C.  
834 Toon, P.O. Wennberg, D.W.T. Griffith, and I. Aben, 2012: CH<sub>4</sub>, CO, and H<sub>2</sub>O spectroscopy for the  
835 Sentinel-5 Precursor mission: an assessment with the Total Carbon Column Observing Network  
836 measurements. *Atmos. Meas. Tech.*, 5, 1387-1398.
- 837 Gloudemans, A. M. S., Schrijver, H., Hasekamp, O. P., and Aben, I.: Error analysis for CO and CH<sub>4</sub> total  
838 column retrievals from SCIAMACHY 2.3 μm spectra, *Atmos. Chem. Phys.*, 8, 3999-4017,  
839 doi:10.5194/acp-8-3999-2008, 2008.
- 840 de Haan, J.F., DISAMAR Algorithms and background, RP-TROPOMI-KNMI-066, KNMI, January 2012.



- 841 Hass, H., M. van Loon, C. Kessler, R. Stern, J. Matthijssen, F. Sauter, Z. Zlatev, J. Langner, V. Foltescu and  
842 M. Schaap, 2003: Aerosol modelling: Results and Intercomparison from European Regional scale  
843 modelling systems, Special Rep. EUROTRAC-2 ISS, Munich, 2003.
- 844 HTAP, 2007: Hemispheric Transport of Air Pollution 2007, Air Pollution Studies No. 16. UN Publication,  
845 ECE/EB.AIR/94, Geneva.
- 846 Huijnen, V., H.J. Eskes, A. Poupkou, H. Elbern, K.F. Boersma, G. Foret, M. Sofiev, M., et al., 2010:  
847 Comparison of OMI NO<sub>2</sub> tropospheric columns with an ensemble of global and European regional air  
848 quality models. Atmos. Chem. Phys., 10, 3273–3296, doi: 10.5194/acp-10-3273-2010.
- 849 Jacob, D.J., 2000: Heterogeneous chemistry and tropospheric ozone. Atmos. Env., 34, 2131–2159.
- 850 Kaiser, J.W., A. Heil, M.O. Andreae, A. Benedetti, N. Chubarova, L. Jones, J.-J. Morcrette, M. Razinger,  
851 M.G. Schultz, M. Suttie, and G.R. van der Werf, G. R., 2012: Biomass burning emissions estimated with  
852 a global fire assimilation system based on observed fire radiative power. Biogeosciences, 9, 527-554,  
853 doi: 10.5194/bg-9-527-2012.
- 854 Kuenen, J.J.P., A.J.H. Visschedijk, M. Jozwicka, and H.A.C. Denier van der Gon, 2014: TNO-MACC\_II  
855 emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory  
856 for air quality modelling. Atmos. Chem. Phys., 14, 10963-10976, doi:10.5194/acp-14-10963-2014,  
857 2014.
- 858 Kujanpää, Jukka, Albert Oude Nijhuis, Henk Eskes, Johan de Haan, Pepijn Veeffkind, Johanna Tamminen,  
859 Synthetic Observation Product Specification (SOPS), Report of the ESA project "Impact of Spaceborne  
860 Observations on Tropospheric Composition Analysis and Forecast" (ISOTROP), 12 August 2015.
- 861 Lagarde, T., A. Piacentini, and O. Thual, 2001: A new representation of data assimilation methods: the  
862 PALM flow charting approach. Q. J. R. Meteorol. Soc., 127, 189–207.
- 863 Lahoz, W.A., R. Brugge, D.R. Jackson, S. Migliorini, R. Swinbank, D. Lary, et al., 2005: An observing  
864 system simulation experiment to evaluate the scientific merit of wind and ozone measurements from the  
865 future SWIFT instrument. Q. J. R. Meteorol. Soc., 131, 503–523. doi:10.1256/qj.03.109.
- 866 Lahoz, W.A., Q. Errera, R. Swinbank, and D. Fonteyn, 2007a: Data assimilation of stratospheric  
867 constituents: a review. Atmos. Chem. Phys., 7, 5745–5773, doi: 10.5194/acp-7-5745-2007.
- 868 Lahoz, W.A., A.J. Geer, S. Bekki, N. Bormann, S. Ceccherini, H. Elbern, Q. Errera, H.J. Eskes, D. Fonteyn,  
869 D.R. Jackson, B. Khattatov, M. Marchand, S. Massart, V.-H. Peuch, S. Rharmili, M. Ridolfi, A. Segers,





- 870 O. Talagrand, H.E. Thornton, A.F. Vik, and T. von Clarmann, 2007b: The Assimilation of Envisat data  
871 (ASSET) project. *Atmos. Chem. Phys.*, 7, 1773–1796.
- 872 Lahoz, W.A., V.-H. Peuch, J. Orphal, J.-L. Attié, K. Chance, X. Liu, et al., 2012: Monitoring air quality from  
873 space: the case for the geostationary platform. *Bull. Am. Meteorol. Soc.*, 93, 221–233. doi:  
874 10.1175/BAMS-D-11-00045.1.
- 875 Lahoz, W.A., and P. Schneider, 2014: Data assimilation: making sense of earth observation. *Front. Environ.*  
876 *Sci.*, 2, 16. <http://dx.doi.org/10.3389/fenvs.2014.00016>.
- 877 Lee, J.D., A.C. Lewis, P.S. Monks, M. Jacob, J.F. Hamilton, J.R. Hopkins, N.M. Watson, J.E. Saxton, C.  
878 Ennis, L.J. Carpenter, N. Carslaw, Z. Fleming, B.J. Bandy, D.E. Oram, S.A. Penkett, J. Slemr, E.  
879 Norton, A.R. Rickard, L.K. Whalley, D.E. Heard, W.J. Bloss, T. Gravesstock, S.C. Smit, J. Stanton, M.J.  
880 Pilling, and M.E. Jenkin, 2006: Ozone photochemistry and elevated isoprene during the UK heatwave  
881 of August 2003. *Atmos. Env.*, 40, 7598–7613.
- 882 Lefèvre, F., G.P. Brasseur, I. Folkins, A.K. Smith, and P. Simon, 1994: Chemistry of the 1991–1992  
883 stratospheric winter: three dimensional model simulations. *J. Geophys. Res.*, 99, 8183-8195.
- 884 Levelt, P., et al., 2009: Observation Techniques and Mission Concepts for Atmospheric Chemistry  
885 (CAMELOT), ESA Study, Contract no. 20533/07/NL/HE.
- 886 Lord, S.J., E. Kalnay, R. Daley, G.D. Emmitt, and R. Atlas, 1997: “Using OSSEs in the design of the future  
887 generation of integrated observing systems, 1st Symposium on Integrated Observing Systems (Long  
888 Beach, CA: American Meteorological Society).
- 889 Manders, A.M.M., M. Schaap, and R. Hoogerbruggem, 2009: Testing the capability of the chemistry  
890 transport model LOTOS-EUROS to forecast PM10 levels in The Netherlands. *Atmos. Env.*, 4050-459  
891 doi:10.1016/j.atmosenv.2009.05.006.
- 892 Marécal, V., V.-H. Peuch, C. Andersson, S. Andersson, J. Arteta, M. Beekmann, A. Benedictow, R.  
893 Bergström, B. Bessagnet, A. Cansado, F. Chéroux, A. Colette, A. Coman, R.L. Curier, H.A.C. Denier  
894 van der Gon, A. Drouin, H. Elbern, E. Emili, R.J. Engelen, H.J. Eskes, G. Foret, E. Friese, M. Gauss, C.  
895 Giannaros, J. Guth, M. Joly, E. Jaumouillé, B. Josse, N. Kadygrov, J.W. Kaiser, K. Krajsek, J. Kuenen,  
896 U. Kumar, N. Liora, E. Lopez, L. Malherbe, I. Martinez, D. Melas, F. Meleux, L. Menut, P. Moinat, T.  
897 Morales, J. Parmentier, A. Piacentini, M. Plu, A. Poupkou, S. Queguiner, L. Robertson, L. Rouïl, M.  
898 Schaap, A. Segers, M. Sofiev, M. Thomas, R. Timmermans, Á. Valdebenito, P. van Velthoven, R. van



- 899 Versendaal, J. Vira, and A. Ung, 2015: A regional air quality forecasting system over Europe: the  
900 MACC-II daily ensemble production, *Geosci. Model Dev.*, 8, 2777-2813, doi:10.5194/gmd-8-2777-  
901 2015
- 902 Massart, S., C. Clerbaux, D. Cariolle, A. Piacentini, S. Turquety, and J. Hadji-Lazaro, 2009: First steps  
903 towards the assimilation of IASI ozone data into the MOCAGE-PALM system. *Atmos. Chem. Phys.*, 9,  
904 5073–5091. doi: 10.5194/acp-9-5073-2009.
- 905 Masutani, M., T.W. Schlatter, R. M. Errico, A. Stoffelen, E. Andersson, W. Lahoz, J.S. Woollen, G.D.  
906 Emmitt, L.-P. Riishøjgaard, and S. J. Lord, 2010a: Observing system simulation experiments. *Data*  
907 *Assimilation: Making Sense of Observations*, W. A. Lahoz, B. Khattatov and R. Ménard, Eds., Springer,  
908 647-679.
- 909 Masutani, M., J.S. Woollen, S.J. Lord, G.D. Emmitt, T.J. Kleespies, S.A. Wood, S. Greco, H. Sun, J. Terry, V.  
910 Kapoor, R. Treadon, and K.A. Campana, 2010b: Observing system simulation experiments at the  
911 National Centers for Environmental Prediction. *J. Geophys. Res.*, 115, doi: 10.1029/2009JD012528.
- 912 Nitta, T., 1975: Some analyses of observing systems simulation experiments in relation to First GARP  
913 Global Experiment, in GARP Working Group on Numerical Experimentation, Report No. 10, US GARP  
914 Plan (Washington, DC), 1–35.
- 915 Ordoñez, C., N. Elguindi, O. Stein, V. Huijnen, J. Flemming, A. Inness, H. Flentje, E. Katragkou, P. Moinat,  
916 V.-H. Peuch, A. Segers, V. Thouret, G. Athier, M. van Weele, C. S. Zerefos, J.-P. Cammas, and M.G.  
917 Schultz, 2010: Global model simulations of air pollution during the 2003 European heat wave. *Atmos.*  
918 *Chem. Phys.*, 10, 789-815.
- 919 Peuch, V.-H., M. Amodei, T. Barthet, M.L. Cathala, M. Michou, and P. Simon, 1999: MOCAGE, MOdèle de  
920 Chimie Atmosphérique à Grande Echelle, in: *Proceedings of Météo France: Workshop on atmospheric*  
921 *modelling*, pp. 33–36, Toulouse, France, 1999.
- 922 Rodgers, C. D., *Inverse methods for atmospheric sounding: Theory and Practice*, Series on Atmospheric,  
923 *Oceanic and Planetary Physics–Vol. 2.*, Singapore, World Scientific, 2000.
- 924 Schaap, M., R.M.A. Timmermans, M. Roemer, G.A.C. Boersen, P.J.H. Builtjes, F.J. Sauter, G.J.M. Velders,  
925 and J.P. Beck, 2008: The Lotos-Euros model: Description, validation and latest developments.  
926 *International Journal of Environment and Pollution*, 32, 270–290.



- 927 Semane, N., V.-H. Peuch, L. El Amraoui, H. Bencherif, S. Massart, D. Cariolle, J.-L. Attié, and R. Abida,  
928 2007: An observed and analysed stratospheric ozone intrusion over the high Canadian Arctic UTLS  
929 region during the summer of 2003, *Q. J. R. Meteorol. Soc.*, 133, 171–178, doi: 10.1002/qj.141.
- 930 Solberg, S., Ø Hov, A. Søvde, I.S.A. Isaksen, P. Coddeville, H. De Backer, C. Forster, Y. Orsolini, and K.  
931 Uhse, 2008: European surface ozone in the extreme summer 2003, *J. Geophys. Res.*, 113, D07307, doi:  
932 10.1029/2007JD009098.
- 933 Stockwell, W.R., F. Kirchner, M. Kuhn, and S. Seefeld, 1997: A new mechanism for regional atmospheric  
934 chemistry modeling. *J. Geophys. Res.*, 102, 25847-25879, doi: 10.1029/97JD00849.
- 935 Stoffelen, A., G.J. Marseille, F. Bouttier, D. Vasiljevic, S. DeHaan, and C. Cardinali, 2006: ADM-Aeolus  
936 Doppler wind lidar observing system simulation experiment. *Q. J. R. Meteorol. Soc.*, 619, 1927–1948.  
937 doi:10.1256/qj.05.83.
- 938 Stern, R., P. Builjtes, M. Schaap, R. Timmermans, R. Vautard, A. Hodzic, M. Memmesheimer, H. Feldmann,  
939 E. Renner, R. Wolke, and A. Kerschbaumer, 2008: A model inter-comparison study focussing on  
940 episodes with elevated PM10 concentrations. *Atmos. Env.*, 42, 4567-4588. [http://dx.doi.org/](http://dx.doi.org/10.1016/j.atmosenv.2008.01.068)  
941 10.1016/j.atmosenv.2008.01.068.
- 942 Streets D. G., 2013: Emissions estimation from satellite retrievals: A review of current capability  
943 *Atmospheric Environment*, 77 1011-1042
- 944 Tan, D.G.H., E. Andersson, M. Fisher, and L. Isaksen, 2007: Observing system impact assessment using a  
945 data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission. *Q. J. R.*  
946 *Meteorol. Soc.*, 133, 381–390. doi: 10.1002/qj.43.
- 947 Tangborn, A., I. Štajner, M. Buchwitz, I. Khlystova, S. Pawson, R. Hudman, et al., 2009: Assimilation of  
948 SCIAMACHY total column CO observations: global and regional analysis of data impact. *J. Geophys.*  
949 *Res.*, 114. doi: 10.1029/2008JD010781.
- 950 Teyssède, H., M. Michou, H.L. Clark, B. Josse, F. Karcher, D. Olivié, V.-H. Peuch, D. Saint-Martin, D.  
951 Cariolle, J.-L. Attié, P. Nédélec, P. Ricaud, V. Thouret, A.R.J. van der A., A. Volz-Thomas, and F.  
952 Chéroux, F., 2007: A new tropospheric and stratospheric Chemistry and Transport Model MOCAGE-  
953 Climat for multi-year studies: evaluation of the present-day climatology and sensitivity to surface  
954 processes. *Atmos. Chem. Phys.*, 7, 5815–5860, <http://www.atmos-chem-phys.net/7/5815/2007/>.



- 955 Timmermans, R.M.A., M. Schaap, H. Elbern, R. Siddans, S. Tjemkes, R. Vautard, et al., 2009a: An  
956 Observing System Simulation Experiment (OSSE) for Aerosol Optical Depth from Satellites. *J. Atmos.*  
957 *Ocean Tech.*, 26, 2673–2682. doi: 10.1175/2009JTECHA1263.1.
- 958 Timmermans, R.M.A., A.J. Segers, P.J.H. Builtjes, R. Vautard, R. Siddans, H. Elbern, et al., 2009b: The  
959 added value of a proposed satellite imager for ground level particulate matter analyses and forecasts.  
960 *IEEE J. Sel. Top. Appl. 2*, 271–283. doi:10.1109/JSTARS.2009.2034613.
- 961 Timmermans, R., W.A. Lahoz, J.-L. Attié, V.-H. Peuch, L. Curier, D. Edwards, H. Eskes, and P. Builtjes,  
962 2015: Observing System Simulation Experiments for Air Quality. *Atmos. Env.*, 115, 199-213,  
963 doi:10.1016/j.atmosenv.2015.05.032.
- 964 Tressol, M., C. Ordoñez, R. Zbinden, J. Brioude, V. Thouret, C. Mari, P. Nedélec, J.-P. Cammas, H. Smit, H.-  
965 W. Patz, and A. Volz-Thomas, 2008: Air pollution during the 2003 European heat wave as seen by  
966 MOZAIC airliners. *Atmos. Chem. Phys.*, 8, 2133-2150.
- 967 Vautard, R., C. Honoré, M. Beekmann, and L. Rouïl, 2005: Simulation of ozone during the August 2003 heat  
968 wave and emission control scenarios. *Atmos. Env.*, 39, 2957–2967.
- 969 van Loon, M., R. Vautard, M. Schaap, R. Bergstrom, B. Bessagnet, J. Brandt, P.J.H. Builtjes, J.H.  
970 Christensen, C. Cuvelier, A. Graff, J.E. Jonson, M. Krol, J. Langner, P. Roberts, L. Rouïl, R. Stern, L.  
971 Tarrason, P. Thunis, E. Vignati, and L. White, 2007: Evaluation of long-term ozone simulations from  
972 seven regional air quality models and their ensemble, *Atmos. Env.*, 41, 2083–2097. doi:  
973 10.1016/j.atmosenv.2006.10.073.
- 974 Veefkind, J.P., I. Aben, K. McMullan, H. Förster, J. de Vries, G. Otter, J. Claas, H.J. Eskes, J.F. de Haan, Q.  
975 Kleipool, M. van Weele, O. Hasekamp, R. Hoogeveen, J. Landgraf, R. Snel, P. Tol, P. Ingmann, R.  
976 Voors, B. Kruizinga, R. Vink, H. Visser, and P.F. Levelt, 2012: TROPOMI on the ESA Sentinel-5  
977 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air  
978 quality and ozone layer applications. *Remote Sens. Env.*, 120, 70-83.
- 979 Weaver, A., and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion  
980 equation. *Q. J. R. Meteorol. Soc.*, 127, 1815–1846.
- 981 Willmott, C., and K. Matsuura, 2005: Advantages of the Mean Absolute Error (MAE) over the Root Mean  
982 Square Error (RMSE) in assessing average model performance. *Clim. Res.*, 30, 79-82.



983 Willmott, C., K. Matsuura, and S.M. Robeson, 2009: Ambiguities inherent in sums-of-squares-based error  
984 statistics. *Atmos. Env.*, 43, 749-752.

985 Worden, H. M., M. N. Deeter, Christian Frankenberg, Maya George, Florian Nichitiu, John Worden,  
986 Ilse Aben et al. "Decadal record of satellite carbon monoxide observations." *Atmospheric Chemistry*  
987 *and Physics* 13, no. 2 (2013): 837-850.

988 Yumimoto, K., 2013: Impacts of geostationary satellite measurements on CO forecasting: an observing  
989 system simulation experiment with GEOS- Chem/LETKF data assimilation system. *Atmos. Env.*, 74,  
990 123–133. doi: 10.1016/j.atmosenv.2013.03.032.

991 Zoogman, P., D.J. Jacob, K. Chance, L. Zhang, P. Le Sager, A.M. Fiore, A. Eldering, X. Liu, V. Natraj, and  
992 S.S. Kulawik, 2011. Ozone air quality measurement requirements for a geostationary satellite mission.  
993 *Atmos. Env.*, 45, 7143-7150.

994 Zoogman, P., D.J. Jacob, K. Chance, X. Liu, M. Lin, A.M. Fiore, and K. Travis, 2014a. Monitoring high-  
995 ozone events in the US Intermountain West using TEMPO geostationary satellite observations. *Atmos.*  
996 *Chem. Phys.*, 14, 6261-6271. <http://dx.doi.org/10.5194/acp-14-6261-2014>.

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

1000

1001

1002

1003

1004

1005

1006

1007



1008 **Tables**

1009

1010 **Table 1:** Spectral and radiometric settings for DISAMAR, and the look-up table nodes.

<b>Spectral and radiometric settings</b>	
Spectral range [nm]	2330-2345
Spectral resolution (FWHM) [nm]	0.25
Spectral sampling [nm]	0.1
SNR Earth radiance	120
SNR Solar irradiance	5000
Additional calibration error (%)	1.0, correlation length 100 nm
<b>Node points</b>	
cos(SZA)	0.1 - 1.0, step 0.1
cos(VZA)	0.3 - 1.0, step 0.1
Relative azimuth [degree]	0.0, 180.0
Cloud/surface pressure	1100 - 200, step -100
Cloud/surface albedo	0.0, 0.005, 0.01, 0.02, 0.04, 0.06, 0.1, 0.2, 0.3, 0.4, 0.8, 0.9
Pressure layers	1100, 1000, 900, 800, 700, 600, 500, 400, 300, 200, 137.50, 68.75, 34.38, 17.19, 8.59, 4.30, 2.15, 1.07, 0.54, 0.27, 0.13, 0.07

1011

1012

1013

1014



1015

1016 **Table 2:** Description of the configuration used in the assimilation system

	<b>Description</b>
Assimilation	3D-var, 1 hour window
Background standard deviation	in % of the background field (vertically variable)
Background correlation zonal Length scale ( $L_x$ )	constant 35 km
Background correlation meridional length scale ( $L_y$ )	constant 35 km
Background correlation vertical length scale ( $L_z$ )	one vertical model layer
S-5P total column CO observation errors	from retrieval product and weighting to account for the total column

1017

1018

1019

1020

1021

1022

1023

1024

1025

1026

1027

1028



1029 **Figures**

1030

1031

1032

1033

1034

1035

1036

1037

1038

1039

1040

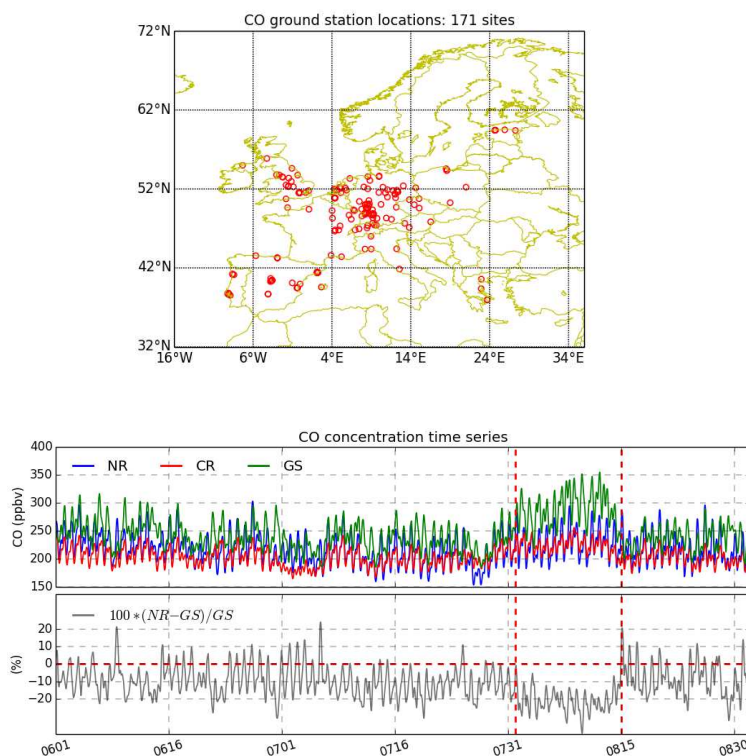
1041

1042

1043

1044 **Figure 1:** Top panel: location of selected ground-based stations for CO measurements taken from the Airbase  
1045 database during northern summer 2003 (1 June – 31 August). There are 171 sites with locations shown by  
1046 circles. The labels show longitude, degrees (x-axis) by latitude, degrees (y-axis). Middle panel: simulated  
1047 and measured time-series of CO concentrations in surface air from nature run (blue line), the control run (red  
1048 line) and from the selected 171 Airbase sites (green line). We form the CO time-series for the ground-based  
1049 stations by averaging the hourly data representative of the 171 sites. The labels show time in MMDD format  
1050 (x-axis) by CO concentration, parts per billion by volume, ppbv (y-axis). Bottom panel: The gray curve  
1051 shows the relative error of the nature run with respect to the observations, defined as NR value minus ground  
1052 station value divided by the ground station value and multiplied by 100. The labels show time in MMDD  
1053 format (x-axis) by relative error, percent (y-axis). The vertical red dashed lines in the middle and bottom  
1054 panels delineate the 2003 European heat wave period (31 July – 15 August).

1055



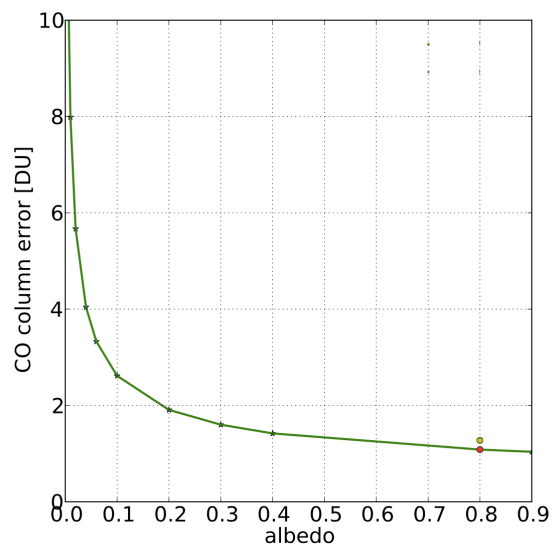




1056

1057

1058



**Figure 2:** Dependence of the CO column uncertainty on the surface albedo. Simulation settings are: solar zenith angle 53 degrees, viewing zenith angle 26 degrees, relative azimuth angle 0 degree, cloud/surface pressure 1100 hPa.

1059

1060

1061

1062

1063

1064

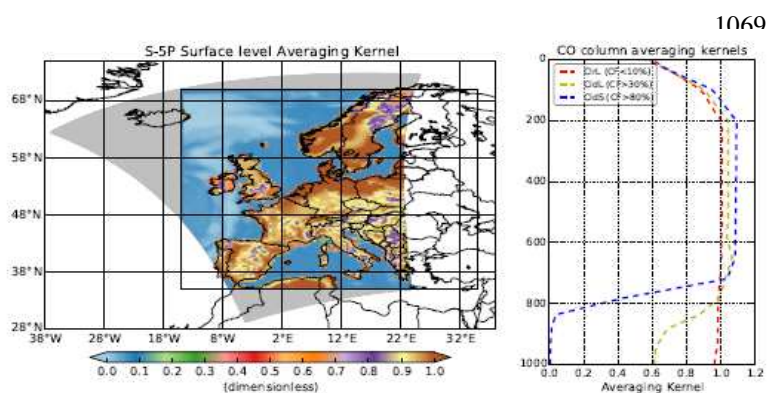
1065

1066



1067

1068



1069

1077

1078 **Figure 3:** Left panel: S-5P CO averaging kernel values at the surface. Labels are longitude, degrees (x-axis)  
1079 by latitude, degrees (y-axis). Right panel: Averaging kernels for land pixels with cloud fraction less than 10%  
1080 (dashed red lines); for land pixels with cloud fraction greater than 30% (dashed yellow lines); and for sea  
1081 pixels with cloud fraction greater than 80% (dashed blue lines). The averaging kernels are for an average of  
1082 the data shown on the swath for 1 June 2003 at 12:34 UTC. Labels are averaging kernel, normalized (x-axis)  
1083 by pressure level, hPa (y-axis).

1084

1085

1086

1087

1088

1089

1090

1091

1092

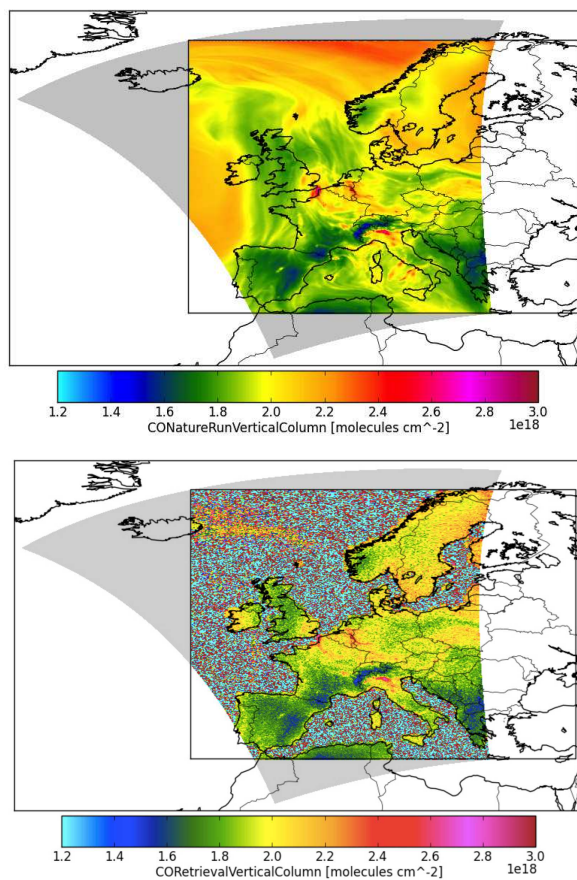
1093

1094



1095

1096



**Figure 4:** Top: Nature run collocated to the synthetic S-5P observations for the 12:34 orbit on 1 June 2003.  
Bottom: corresponding synthetic observations.

1097

1098

1099

1100

1101

1102

1103



1104

1105

1106

1107

1108

1109

1110

1111

1112

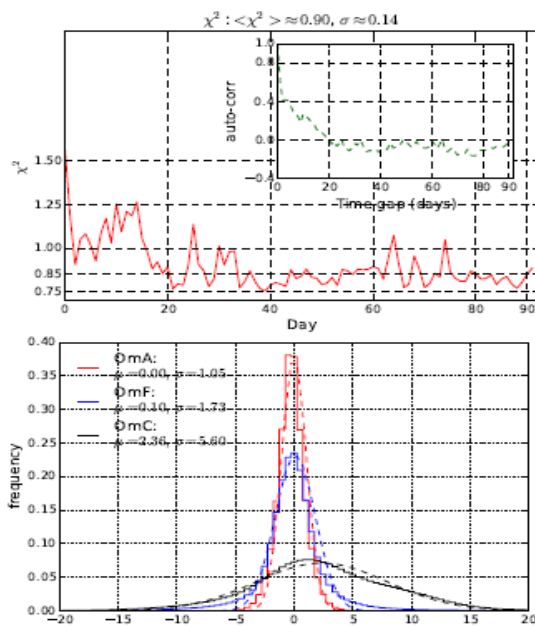
1113

1114

1115

1116

1117



1118 **Figure 5:** Self-consistency tests. Top panel: time-series (red line) of  $\chi^2$  for OmF and its associated auto-  
 1119 correlation signal (green line). For the  $\chi^2$  diagnostic we normalize the OmF differences by the background  
 1120 error. The labels show time, days (x-axis) and  $\chi^2$  value (y-axis) for the  $\chi^2$  plot, and time gap, days (x-axis) and  
 1121 auto-correlation (y-axis) for the auto-correlation plot. Bottom panel: histograms of Observations minus  
 1122 Analysis (OmA -red solid line), Observations minus Forecast (OmF -blue solid line), and Observations  
 1123 minus Control run (OmC -black solid line). We normalize these differences by the observation error. The  
 1124 dashed lines correspond to the Gaussian fits of the different histograms. The labels show the OmA, OmF or  
 1125 OmC differences (x-axis) and the frequency of occurrence of the differences (y-axis). We calculate the  
 1126 diagnostics OmA, OmF, and OmC over the period of 1 June – 31 August 2003.

1127

1128

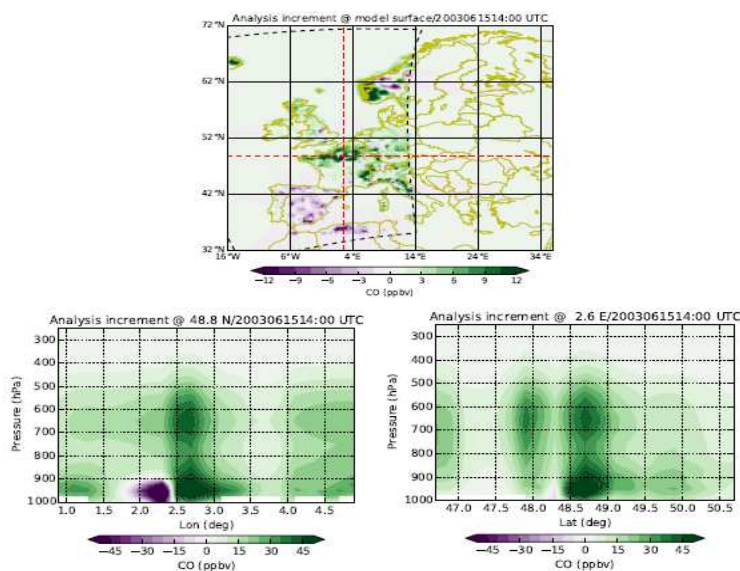
1129

1130



1131

1132



1144 **Figure 6:** S-5P CO analysis increments, units of ppbv, at 14:00 UTC on 15 June 2003: Top panel:  
1145 geographical distribution at the model surface. Red dashed lines show zonal and meridional vertical slices at  
1146 48°8 N, and 2°6 E, respectively. The black dashed line shows the S-5P cross-track at 13:12 UTC, clipped to  
1147 fit the OSSE simulation domain. Note that we measure the S-5P CO observations at 13:12 UTC. The labels  
1148 show longitude, degrees (x-axis) and latitude, degrees (y-axis). Left and right bottom panels show,  
1149 respectively, the longitude-height and latitude-height cross-sections at a location near Paris. The labels for  
1150 the bottom panels show longitude, degrees (x-axis, left panel), latitude, degrees (x-axis, right panel), and  
1151 pressure, hPa (y-axis, both panels). Green/purple colours indicate positive/negative values in the increment  
1152 fields.

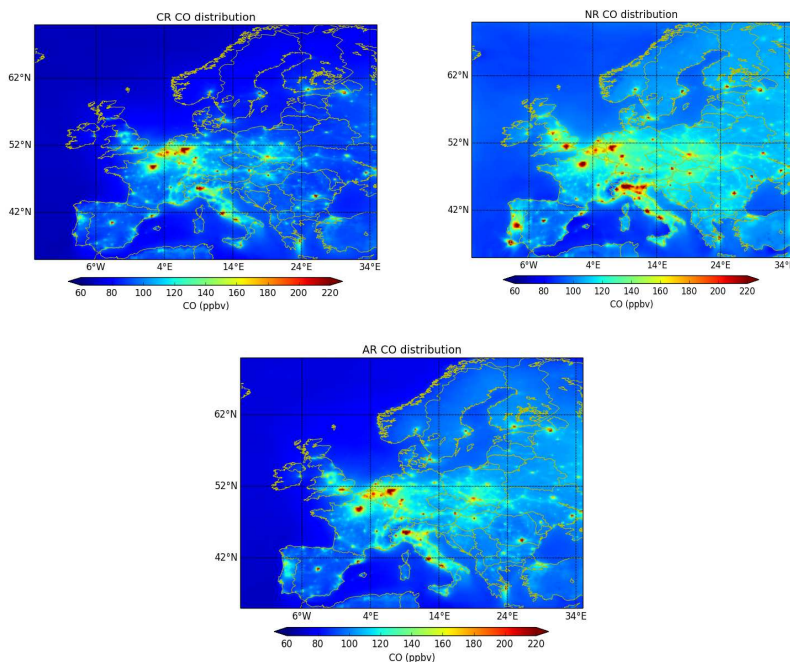
1153

1154

1155

1156

1157



1158

1159

1160 **Figure 7:** Distribution of CO surface concentrations, units ppbv, averaged for the period 1 June – 31 August  
1161 2003. Top left panel: the control run (CR) from MOCAGE; right top panel: the nature run (NR) from  
1162 LOTOS-EUROS; bottom panel: the assimilation run (AR) from MOCAGE obtained after assimilating the S-  
1163 5P CO total column simulated data sampled from the NR. In all panels, the labels show longitude, degrees  
1164 ( $x$ -axis) and latitude, degrees ( $y$ -axis). Red/blue colours indicate relatively high/low values of the CO surface  
1165 concentrations.

1166

1167

1168

1169

1170

1171

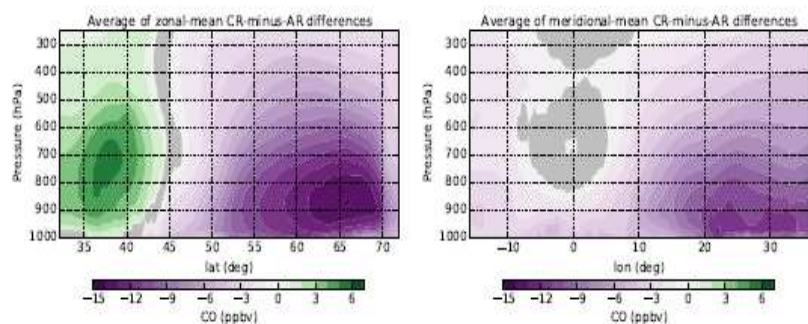
1172

1173

1174



1175



1176

1183

1184 **Figure 8:** Zonal (left panel) and meridional (right panel) slices of the difference between the CR and AR CO  
1185 fields, units of ppbv, averaged over the summer period (1 June – 31 August 2003). The areas highlighted in  
1186 grey colour indicate where the AR is not significantly different to the CR at the 99% confidence level. The  
1187 labels in the left panel are latitude, degrees (x-axis) and pressure, hPa (y-axis). The labels in the right panel  
1188 are longitude, degrees (x-axis) and pressure, hPa (y-axis). Green/purple colours indicate positive/negative  
1189 values in the difference fields.

1190

1191

1192

1193

1194

1195

1196

1197

1198

1199

1200

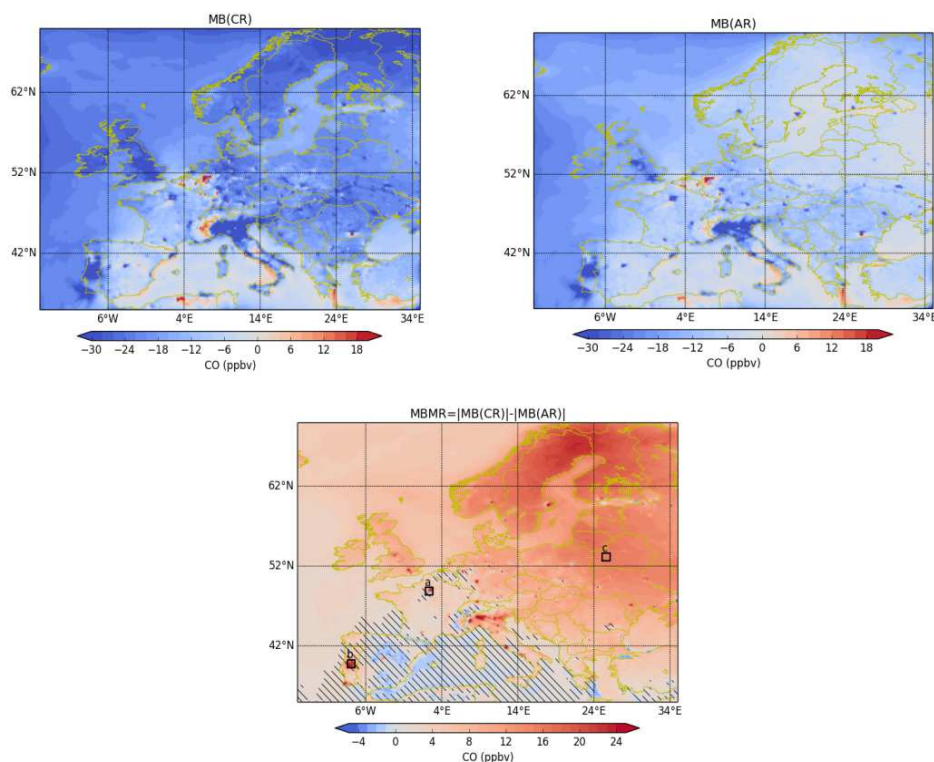
1201

1202



1203

1204



1205

1206 **Figure 9:** Mean bias reduction at the surface for CO, units of ppbv: Left top panel shows the CR mean bias  
1207 with respect to the NR (CR-NR). Right top panel shows the AR mean bias with respect to the NR (AR-NR).  
1208 Bottom panel shows the mean bias magnitude reduction (absolute value of the mean bias for CR minus the  
1209 absolute value of the mean bias for AR). We average the data over northern summer 2003 (1 June – 31  
1210 August). The labels show longitude, degrees (x-axis) and latitude, degrees (y-axis). The hatched area in the  
1211 bottom panel shows where the mean bias plotted in this panel (MBMR) is not statistically significant at the  
1212 99% confidence level. The three squares in the bottom panel represent locations for the three time-series  
1213 shown in Fig. 12. Red/blue colours indicate positive/negative values in the MB/MBMR.

1214

1215

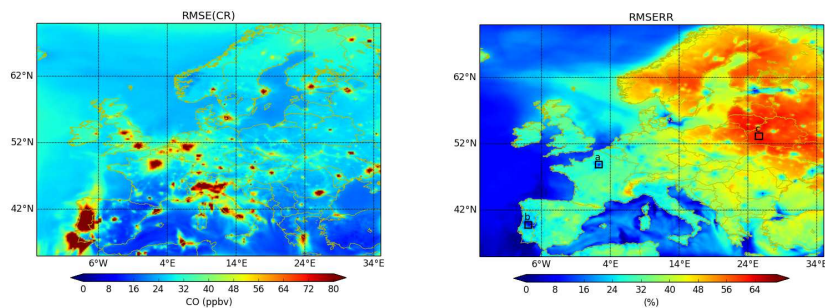
1216

1217

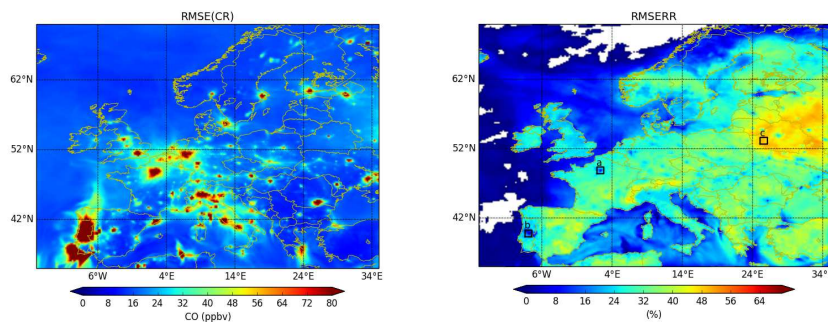




1218



1219



1220

1221

1222 **Figure 10:** Top: Root Mean Square Error (RMSE), units of ppbv, between CR and NR (left panel), and its  
1223 corresponding reduction rate RMSERR, in % (right panel) keeping the systematic error. Bottom: Same as top  
1224 panel but calculating the RMSE after removing the systematic error. The labels on each panel are longitude,  
1225 degrees (x-axis) and latitude, degrees (y-axis). The three squares in the two right panels represent locations  
1226 for the three time-series shown in Fig. 12. Red/blue colours indicate relatively high/low values in the  
1227 RMSE/RMSERR.

1228

1229

1230

1231

1232

1233

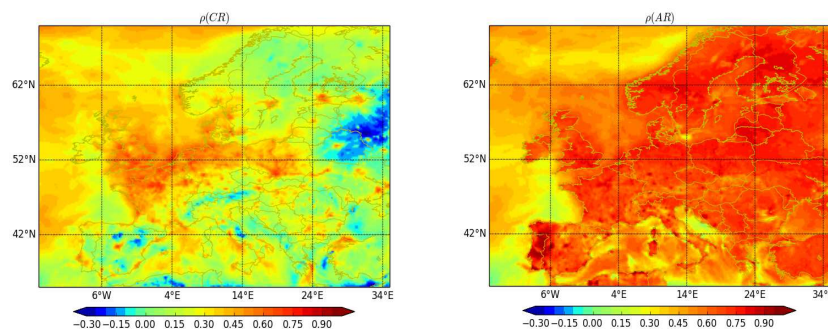
1234

1235



1236

1237



1238

1239 **Figure 11:** Correlation coefficient between the CR and the NR (left panel) and the AR and the NR (right  
1240 panel) at the surface and for the northern summer period (1 June – 31 August). The labels are longitude,  
1241 degrees (x-axis) and latitude, degrees (y-axis). Red/blue colours indicate positive/negative values of the  
1242 correlation coefficient.

1243

1244

1245

1246

1247

1248

1249

1250

1251

1252

1253

1254

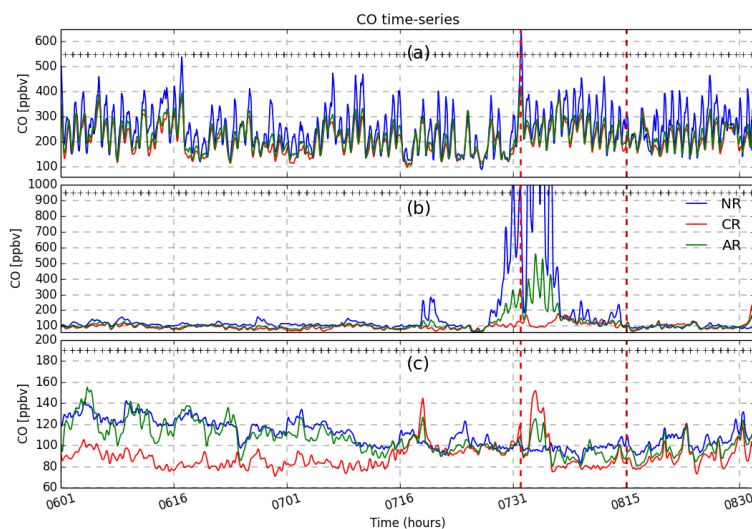
1255

1256

1257



1258



1259

1260 **Figure 12:** Time-series for CO surface concentrations (1 June – 31 August) from NR (blue colour), CR (red  
1261 colour) and AR (green colour) over three different locations represented by squares in Figs. 9 and 10. Top  
1262 panel: area near Paris; middle panel: area over Portugal, where forest fires occurred; bottom panel: Eastern  
1263 part of the study domain. The labels in the three panels are time, in format MMDD (x-axis) and CO  
1264 concentration, ppbv (y-axis). The plus symbols at the top of each panel indicate availability of observations  
1265 from the S-5P platform.

1266

1267

1268

1269

1270

1271

1272

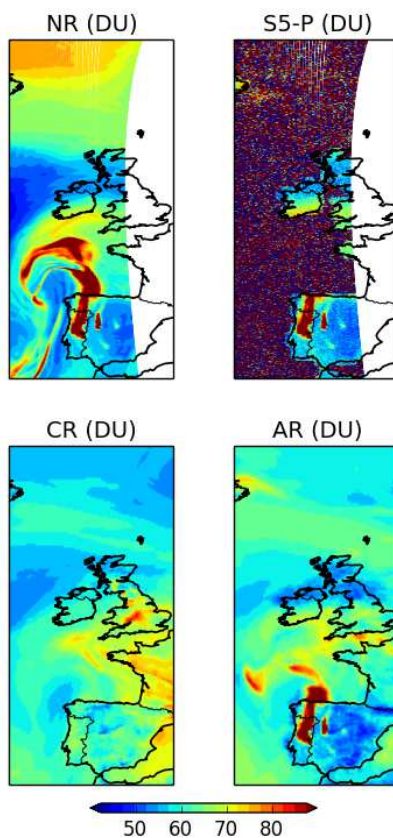
1273

1274

1275



1276



1277

1278 **Figure 13:** CO total column at 14:15 UTC on 4 August 2003, Dobson units, DU. Top left panel: NR; top  
1279 right panel: simulated S-5P observations; bottom left panel: CR; bottom right panel: AR. Red/blue colours  
1280 indicate relatively high/low values of the CO total column.

1281

1282

1283

1284

1285

1286

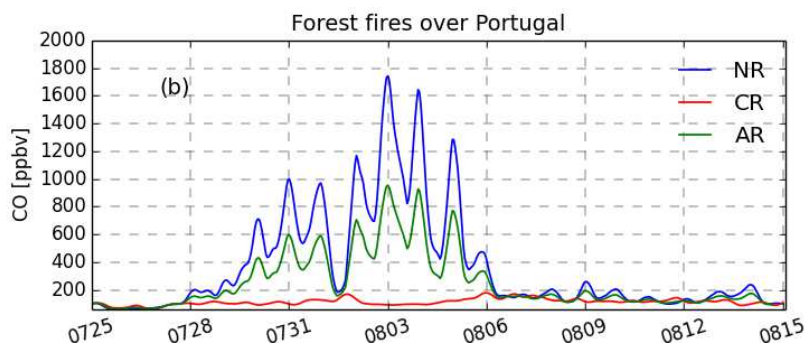
1287

1288



1289

1290



1291

1292

1293 **Figure 14:** Time-series for CO surface concentrations for the period covering the Portugal forest fires (25  
1294 July – 15 August) from NR (blue colour), CR (red colour) and AR (green colour) over the location associated  
1295 with the middle panel of Fig.12. These data concern the second OSSE we perform to understand the  
1296 behaviour of the original OSSE over the period of the forest fires (see text for more details). The labels are  
1297 time, in format MMDD (x-axis) and CO concentration, ppbv (y-axis).

1298

1299

1300

1301

1302

1303

1304

1305

- 4-1 May 9, 2016, 6:40 PM, Ref  
Need to include CrIS.
- 4-2 May 9, 2016, 6:40 PM, Ref  
What about planned?
- 4-3 May 9, 2016, 6:40 PM, Ref  
Why is methane defined twice?
- 4-4 May 9, 2016, 6:40 PM, Ref  
How is this point relevant if you're not discussing geostationary options?
- 5-1 May 9, 2016, 6:40 PM, Ref  
Relative to what, MOPITT? Not clear how since they both have NIR channels.
- 5-2 May 9, 2016, 6:40 PM, Ref  
Poorly formed sentence. And not quite true. The resolution of an inverse emission estimate is controlled by the data and transport/diffusion. Not clear which is the limiting factor without analysis.
- 5-3 May 9, 2016, 6:40 PM, Ref  
True, but how is relevant?
- 5-4 May 9, 2016, 6:40 PM, Ref  
So, TROPOMI needs to be assessed relative to the existing satellites.
- 6-1 May 9, 2016, 6:40 PM, Ref  
So, if one knew CO concentrations or emissions better, what societal or scientific benefit would have been achieved? Improved forecasts? Better attribution?
- 7-1 May 9, 2016, 6:40 PM, Ref  
Shouldn't the control run include assimilation of the existing observing system, e.g., MOPITT, AIRS, CrIS, rather than a free run? This would be true for your 2003 test case when MOPITT and AIRS was available.
- 7-2 May 9, 2016, 6:40 PM, Ref  
I shouldn't have to look up the figure. How does it relate to the specific OSSE elements listed?
- 9-1 May 9, 2016, 6:40 PM, Ref  
Did I miss something? Where is the CR described?
- 9-2 May 9, 2016, 6:40 PM, Ref  
There is also some high frequency component that is missed. What is that frequency? is that the night time values? Needs to be discussed.
- 9-3 May 9, 2016, 6:40 PM, Ref  
That statement needs to be limited to the ultimate performance of the OSSE. The assimilation can't saying anything better than 10-20% in accuracy.
- 10-1 May 9, 2016, 6:40 PM, Ref  
At what accuracy?
- 10-2 May 9, 2016, 6:40 PM, Ref  
How low?

- 10-3** May 9, 2016, 6:40 PM, Ref  
More accurately compared to what?
- 11-1** May 10, 2016, 9:54 AM, Ref  
But OMI is quite a bit larger in footprint than TROPOMI. MODIS would be a better choice. Describe how the cloud fraction is related to the 7km footprint. How different are cloud ODS between the UV and the NIR? It seems like you are assuming they are the same.
- 12-1** May 9, 2016, 6:40 PM, Ref  
provides
- 12-2** May 9, 2016, 6:40 PM, Ref  
Is it released now?
- 12-3** May 9, 2016, 6:40 PM, Ref  
sea
- 13-1** May 10, 2016, 9:54 AM, Ref  
Cloudy
- 13-2** May 10, 2016, 9:54 AM, Ref  
It looks like you are assuming that the retrievals will work under partially cloudy scenes. MOPITT NIR only works under clear sky. Provide a reference on NIR CO retrievals under partially cloudy conditions and justify why a weighted approach works. Also I haven't heard any discussion of aerosols. These will be important for emissions like biomass burning and industry.
- 13-3** May 10, 2016, 9:54 AM, Ref  
This is the first time MOCAGE is mentioned even though it is implicitly referenced in the CR run. Needs to be mentioned earlier.
- 16-1** May 10, 2016, 9:54 AM, Ref  
That is not exactly true. The existence of burning and the burnt area can be obtained from optical measurements. That is very important a priori information. Is that being ignored?
- 17-1** May 10, 2016, 9:54 AM, Ref  
In the introduction, the authors argued for the value of TROPOMI to resolve emissions. Here the focus has shifted to concentration estimation. What is the scientific rationale? What are the limitations of this OSSE, then, to make statements about resolving sources?
- 21-1** May 10, 2016, 9:54 AM, Ref  
The correction in the free trop points to the role of boundary conditions in the assimilation, which would be an important consideration for a GEO. How much change is occurring at the boundary of the nested grid?
- 21-2** May 10, 2016, 9:54 AM, Ref  
That is not obvious. The biggest weakness is using 3D-var and having poor prior statistics. Given that significant fires generally last longer than a day, a proper inversion system would pick those up. That does not diminish the value of GEO sounders.
- 21-3** May 10, 2016, 9:54 AM, Ref  
Please elaborate as to whether it is the diurnal sampling or the effective sampling density that is more important for a geo.

- 
- 21-4** May 10, 2016, 9:54 AM, Ref  
That's too qualitative. Could please provide some simple metrics, e.g., means, to quantify these statements?
- 23-1** May 10, 2016, 9:54 AM, Ref  
Why?
- 24-1** May 10, 2016, 9:54 AM, Ref  
unbiased
- 24-2** May 10, 2016, 9:54 AM, Ref  
Why is this called a systematic error? It looks like you are merely removing the bias term in the error, which is simply a statement that the assimilation is not an unbiased estimator.
- 25-1** May 10, 2016, 9:54 AM, Ref  
However, both the CR and AR miss the high frequency min/max. Why?
- 25-2** May 10, 2016, 9:54 AM, Ref  
This is a weakness of the OSSE design, not the measurements.
- 25-3** May 10, 2016, 9:54 AM, Ref  
Please explain why the variability is high in Paris but not in E. Europe.
- 25-4** May 10, 2016, 9:54 AM, Ref  
This seems like a discovery for the authors post-facto. I recommend using this version of the OSSE only rather than devoting a whole section to it. It's what should have been done originally.
- 26-1** May 10, 2016, 9:54 AM, Ref  
The first several paragraphs are repetitious of the introduction. I recommend removing. In fact, the OSSE has shown that S5P will have a similar or bigger impact on the free trop rather than merely the surface.