

Referee # 1

We thank the referee for his/her helpful comments. Our response to referee #1 is below. Items in bold and italics are the referee comments.

L30: “. . .with the largest benefit occurring over land in remote regions”, explain remote from what? Sources?

We mean regions far away from important CO sources. We rephrased this in the introduction. See line 29

L60-68: several sensors are listed but all references are from the MOPITT team. Please add references for the relevant sensors.

We added additional references in particular for AIRS and IASI CO in the revised version. See line 63

L81-103 paragraph: In the discussion of S-5, S-4, and S-5P, etc., please list the timeframe of these missions. Since OMI and SCIAMACHY are discussed, as well as S-5P, this should be a good place to introduce TROPOMI. Among the sensors/missions discussed in this paragraph, which ones have CO, since it's the topic of study here?

We mention now the time frame for S-4. S-5 and S-5P time frames are already mentioned (see lines 88 and 93-95). We also mention TROPOMI in this paragraph but note we describe the instrument characteristics and show how the the S-5P simulated measurements are generated in section 2.2. We replace GOME-2 by IASI which measures CO and has the same revisit local time (line 98-99).

L99-101: “The S-5P LEO platform will address the challenge of limited revisit time from LEOs by providing unprecedented high spatial resolution of 7x7 km, and improved sensitivity in the Planetary Boundary Layer (PBL), allowing resolution of, e.g., derived CO emission sources at finer scales than hitherto.” How?

We clarify this point in the revised version. We remove the vague sentence concerning the sources (lines 102-104)

L223: “. . .the NR has a realistic representation of the CO diurnal cycle.” Does CO have diurnal cycle? Also, describe the ground measurement methods. Is it in situ or radiometric?

At air surface usually carbon monoxide exhibits a diurnal variation, generally with two peaks, one in the morning and the second in the evening. Ground station measurements are real in situ observations taken from AirBase data set. We mention this in the revised version. See line 213.

L232-233: “. . .the behaviour of the CO time-series from the CR compared to the NR, is similar to the behaviour of the NR CO time-series compared to the Airbase data.” Not clear, are the differences similar? You might want to add the difference $100*(NRCR)/NR$.

We likely are in the configuration where the NR is between the GS data and the CR with similar behaviour that means in our assimilation we may need a similar correction for the CR to obtain the NR or for the NR to obtain the GS. We rephrase this in the revised version. See lines 246-248

L237-242: TROPOMI should have been introduced in the Introduction.

See previous answer.

L308: “Over sea, . . .” should probably be “Over the ocean, . . .”

corrected.

L387: “. . .different to . . .” to “. . .different from . . .”

Corrected (line 408 in the revised version)

L391: “. . . the OSSE will more realistically simulate. . .” to “. . . the OSSE will simulate more realistically . . .”

Corrected (line 412)

L392: “This follows our guiding principle to . . .” to “This allows us to . . .”

Corrected (line 413)

L394-395: “As mentioned in Sect. 2.3, we use the MOCAGE model to generate the CR. In this OSSE study, the CR is a free model run.” to “In this OSSE study, the CR is a free model run using MOCAGE.”

Corrected (line 415)

L409: “. . . and about 800 m in the neighbourhood of the tropopause . . .” to “. . .and approximately 800 m near the tropopause . . .”

Corrected

L417: “. . . as they are a priori not known.” To “. . . as their a priori is not known.”

We corrected by (... as their a priori is unknown). See line 439

L420: “. . . , helping to differentiate the CR from the NR.” To “. . . , which helps to differentiate the CR from the NR.”

Corrected (line 441)

L420: “As for the NR, . . .” to “Similar to the NR, . . .”

Corrected (line 442)

L435-436: add “the”

Corrected. See line 451

L453-454: “. . . for the B-matrix: Lx and Ly are . . . ; Lz is constant and . . .” to “. . .for the B-matrix, where Lx and Ly are . . . ; and Lz is constant and . . .”

Corrected. See line 474-475

L459: “. . . (see Fig. 3 in Sect. 3.1).” should be Fig. 5

Corrected. All the figure are renumbered in the revised version

L504: Should be Figure 5, not Figure 3

Corrected

L564-570: The reasons the AR are not performing well over fire emissions are not explained correctly in this paragraph, as suggested by L670-676. Should move part of L670-676 to this part of the paper to properly explain why AR did not work well over fires.

We took into account this suggestion.

L642: "Figure 11 shows that the AR . . ." would be better to "The AR . . ."

We correct this in the revised version (line 668)

L651: ". . . (iii) an area in the Easter part of the study domain, . . ." should name the location/country/regions.

We will correct this in the revised version.

L662-665: same issue explaining MOCAGE assimilation over fires.

We will rephrase this paragraph and give more details.

Fig. 5. Legends and labels are not legible when overlap with dashed lines. Should redo with care and use larger fonts.

We enlarge former Fig. 5 (now Fig. 6) but we could provide all the figures in their original format to have the best quality.

Referee # 2

We thank the referee for his/her helpful comments. Our response to referee #2 is below. Items in bold and italics are the referee comments.

While they establish a bias and variance between the NR and independent observations, their fundamental threshold is that the difference appears "reasonable".

What they have not done is to relate the errors between the NR and independent observations to the interpretation of the performance of the AR. So, if the accuracy and precision of the NR is twice as bad, what should one infer about the performance of the AR?

We wrote a paragraph which clarified this point. See lines 232-237.

I would argue that a more important implementation of their guiding principle is to assimilate real observations, e.g., MOPITT and/or SCIAMACHY, into their system and compare the analysis fields to independent observations. Then, they could do an OSSE for the same observing system and assess the statistical difference between the AR and NR sampled at independent observations versus AR(real) against independent observations. That would provide a better sense of what the OSSE limitations actually are.

We think that the suggestion of the referee is valuable, and suitable for a study in its own right. However, it would only provide direct information on the error of the OSSE for MOPITT and/or SCIAMACHY CO and not directly for S5P. The S5P instrument has different characteristics to MOPITT and SCIAMACHY. Instead, and to keep the study tractable, we focus on the evaluation of the uncertainty in the NR, and compare it to published studies using the MOCAGE DA system. This in line with standard practice in AQ OSSEs as discussed in Timmermans et al (2015).

As it stands, I'm still suspicious of the overall performance. Furthermore, we don't know how well S5p will perform given other sensors, e.g., MOPITT, CrIS, are already taking CO data with comparable performance.

In this AQ OSSE, we follow standard practice by comparing the performance of S5P against a free model run. If the proposed satellite data are to have added value, they must perform better than a model. This is the first step one must take to evaluate the added value of a proposed satellite instrument. We mentioned this in the conclusion, see lines 759-769

Otherwise, the overall work is reasonable and the authors have performed some nice statistical analysis of the results. Of course, in principle, this OSSE should have been performed *before* S5p was funded to assess its potential. But, practice is still catching up with theory. I've attached comments of the manuscript in the accompanying pdf.

We thank the referee for this comment. We address below the other points from the referee.

4-1
Need to include CrIS.

We added the information on CrIS and the reference Fu et al., 2016, see lines 63 and 66

4-2
What about planned?

We will check if there are plans for GEO missions to measure CO. To our knowledge no GEO missions to measure specifically CO are planned yet.

4-3
Why is methane defined twice?

We identify that the formula for methane is CH₄.

4-4
How is this point relevant if you're not discussing geostationary options?

We think it helps the reader to contrast the characteristics of GEO and LEO satellite platforms with respect to atmospheric composition.

5-1
Relative to what, MOPITT? Not clear how since they both have NIR channels.

We mean that the S5P with its SWIR band will do better than our model in the PBL. Furthermore, compared to TIR instruments such as IASI, we expect S5P to do better in the PBL (Veefkind et al., 2012). See lines 101-104

5-2
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.

We removed the sentence

5-3
True, but how is relevant?

We think it is helpful to remind the reader where the PBL is located.

5-4
So, TROPOMI needs to be assessed relative to the existing satellites.

See response to the general comments from this referee.

6-1

So, if one knew CO concentrations or emissions better, what societal or scientific benefit would have been achieved? Improved forecasts? Better attribution?

We provide an example of the benefits of improved knowledge of the CO distribution. See lines 140-143

7-1

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.

See the response to the general comments of this referee.

7-2

I shouldn't have to look up the figure. How does it relate to the specific OSSE elements listed?

We provide this figure and relate it to the OSSE elements listed (this is Fig. 1 in the revised paper). We renumbered all the other figures.

9-1

Did I miss something? Where is the CR described?

We refer to section 2.4, which describes the CR. See lines 178-180

9-2

There is also some high frequency component that is missed. What is that frequency? is that the night time values? Needs to be discussed.

We discuss this high frequency component. See lines 231-236

9-3

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.

We rephrase this sentence following the referee's suggestion. See lines 247-248

10-1

At what accuracy?

We quote Veeffkind et al (2012) on this point: 15% (accuracy) and 10% (precision). See lines 261-262 in the revised version.

10-2

How low?

We quote Veeffkind et al., 2012 on this point (the value is 2%). See line 267

10-3

More accurately compared to what?

Original sentence reads:

“The use of S-5P CO total column measurements with inverse modelling techniques will also help quantify more accurately biomass burning emissions and map their spatial distribution.”

We remove “more accurately”

The statement refers to the current observing system, consisting of, e.g., IASI, MOPITT, OMI, GOME-2, including measurements of the species CO and NO₂. S5P will provide global coverage, enhanced sensitivity for CO at the surface (compared to, e.g., IASI) and 3.5 to 7 km high spatial resolution observations.

11-1

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.

The cloud fractions were derived at the resolution of the ECMWF 0.25 x 0.25 degree grid. This is ca. 30 x 30 km² at the equator and decreases as a function of latitude. The ground pixel of OMI UV-2 and VIS channels is 13 x 24 km² at nadir increasing to 13 x 128 km² at edges of the swath. We consider that the ECMWF grid cells and OMI pixels are of comparable size for the purpose of comparing the cloud fraction distributions (ca. 0.5 million pixels or cells in each distribution). We model clouds with a simple Lambertian reflectors and ignore any wavelength dependency of cloud fraction.

We include this information in the revised version. See lines 296-301.

12.1

Provides

We think provide is appropriate

12-2

Is it released now?

Yes, it is available from the ESA Sentinel-5P TROPOMI document library:

<https://sentinel.esa.int/web/sentinel/user-guides/sentinel-5p-tropomi/document-library>.

We provide this information in the revised paper. See line 329

12.3

sea

Corrected

13.1

Cloudy

We understand clouded and cloudy are both appropriate here. If insisted upon, we will change this.

13-2

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.

TROPOMI NIR CO retrievals in partially cloudy conditions are discussed in Landgraf et al.: "Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements", Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2016-114. They are also discussed in Vidot et al.: "Carbon monoxide from shortwave infrared reflectance measurements: A new retrieval approach for clear-sky and partially cloudy atmospheres", Remote Sens. Environ., doi:10.1016/j.rse.2011.09.032. We did not include aerosol effects in our study. We add these references to the paper. See line 354.

13-3

This is the first time MOCAGE is mentioned even though it is implicitly referenced in the CR run. Needs to be mentioned earlier.

At the start of section 2, when we first mention the CR, we will introduce MOCAGE and provide appropriate references. See lines 178-180.

16-1

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?

The visible information on burnt area and burning does not provide direct knowledge on CO concentrations. It is not used in this study.

17-1

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?

We reword the text in the introduction to state that we focus on CO concentrations. We will clarify the scientific rationale of the paper and indicate the limitations of the OSSE. See lines 118-123 and 150-152.

21-1

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?

First, note that the OSSE concerns S5-P, which is a LEO. Nevertheless, we focus on the surface level and we assume that the effect of the free troposphere on the boundary layer is secondary. Due to the revisit time of S5-P, we expect the impact at the boundaries to be small. Second, for efficiency reasons and storage limitations, we set up our DA system to only store the data over the regional domain. This means that without rerunning the OSSE, we cannot quantify the response to the reviewer's comment.

21-2

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.

We clarify that. We are now talking about concentrations and not emission inversions. See lines 589 and 594.

21-3

Please elaborate as to whether it is the diurnal sampling or the effective sampling density that is more important for a geo.

The comment from the referee is not clear to us. We think that diurnal sampling (high temporal resolution) will be the determining factor, as the relatively coarse model resolution would compromise the high spatial sampling.

21-4

That's too qualitative. Could please provide some simple metrics, e.g., means, to quantify these statements?

We quantify this difference in the revised paper. See lines 584-589

23-1

Why?

Owing to the relatively small variability of CO over remote land regions, the S5-P data can provide a larger benefit compared to regions where the variability is relatively high. We make this point in the revised version. See lines 651-653

24.124
Unbiased

Corrected

24-2

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.

When calculating the RMSE we remove the bias between the AR and the NR and between the CR and the NR. We then make the common equivalence between systematic error and bias. We clarify this point in the revised version. We call it now bias. See lines 651-653.

25-1

However, both the CR and AR miss the high frequency min/max. Why?

The AR and the CR capture the variability but not the values of the peaks. However, the LEO only samples at most twice a day over Paris and may not capture the peaks. We indicate the S5-P revisit time by the plus signs at the top of the panel and when you zoom in one sees that the peaks do not coincide with the time of the S5 P measurements. Thus, S5P cannot capture the value at the peak. Another factor could be that the emission inventory used in the AR has lower values than the one used in the NR. We clarify this in the revised version. See lines 683-687.

25-2

This is a weakness of the OSSE design, not the measurements.

Maybe we have mis-understood the referee's comment, but our view is that because we do not know the fires a priori, we cannot include them in the CR and the AR. In our view, this result shows the benefits from the measurements regarding the identification and quantification of fire emissions.

25-3

Please explain why the variability is high in Paris but not in E. Europe.

The variability is higher over Paris than over E. Europe, because there are higher emissions over Paris than over E. Europe (as shown in Fig. 7 – old paper submission). We clarify this in the revised version. See lines 695-696.

25-4

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.

We agree with the reviewer. However, we think it is relevant to present the results in this way because it shows the limitations of using standard operational criteria as we did in the first experiment of the OSSE. We make this point in the revised version. See lines 697-698 and 726

26-1

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.

We will edit the conclusions following the reviewer's comments. The focus of this study is the surface; however, a study of the increments (see figs. 6 and 8 in the old paper submission) indicates that S5-P has benefits in the free troposphere. We mention this in the revised version. See the conclusions.

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

Mis en forme : List Paragraph

Mis en forme : Début de section :
Nouvelle page, Numérotation :
Recommencer à chaque section

4 | R. Abida¹, J.-L. Attié^{1,2}, L. El Amraoui¹, P. Ricaud¹, W. Lahoz³, H. Eskes⁴, A. Segers⁵, L. Curier⁵, J.
5 | de Haan⁴, J. Kujanpää⁶, A. O. Nijhuis⁴, D. Schuetttemeyer⁷, J. Tamminen⁶, R. Timmermans⁵, P.
6 | Veefkind⁴, and B. Veihelmann⁷

7 |
8 | ¹ CNRM-GAME, Météo-France/CNRS UMR 3589, Toulouse, France

9 | ² Université de Toulouse, Laboratoire d'Aérodynamique, CNRS UMR 5560, Toulouse, France

10 | ³ NILU – Norwegian Institute for Air Research, P.O. Box 100, 2027 Kjeller, Norway

11 | ⁴ Royal Netherlands Meteorological Institute (KNMI), P.O. Box 201, 3730 AE De Bilt, The
12 | Netherlands

13 | ⁵ TNO, Business unit Environment, Health and Safety, P.O. Box 80015, 3508 TA Utrecht, The
14 | Netherlands

15 | ⁶ Finnish Meteorological Institute, Earth Observation Unit, P.O. Box 503, 00101 Helsinki, Finland

16 | ⁷ ESA/ESTEC, Earth Observation Programmes, Noordwijk, The Netherlands
17 |

18 | Submitted to Atmospheric Chemistry and Physics Discussions, 13 November 2015

19 **Abstract**

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

33

34 1. Introduction

35 Over the last decade, the capabilities of satellite instruments for sensing the lower troposphere have
36 improved, and opened the way for monitoring and better understanding of atmospheric pollution processes,
37 e.g., tropospheric chemistry (Jacob, 2000), long-range transport (HTAP, 2007), and emissions (e.g. Streets
38 ~~D.~~, 2013 and references therein). Satellite instruments provide global measurements of many pollutants (e.g.,
39 ozone; carbon monoxide, CO; nitrogen dioxide, NO₂; and aerosols), including information on their trans-
40 boundary transport, and complement in situ measurements from ground-based stations (e.g., European
41 [Monitoring and Evaluation Programme \(EMEP\)](http://www.nilu.no/projects/ccc/emepdata.html)), <http://www.nilu.no/projects/ccc/emepdata.html>, and
42 Airbase, <http://acm.eionet.europa.eu/databases/airbase/>, networks). Low Earth Orbit (LEO) satellite
43 platforms have the advantage of providing observations with global coverage, but at a relatively low
44 temporal resolution. Geostationary Earth Orbit (GEO) satellite platforms provide observations at a
45 continental scale, i.e., not global, but at a much higher temporal resolution.

46
47 Satellite data, either in synergy with ground-based and airborne measurements and/or assimilated into
48 models such as chemistry transport models (CTMs), contribute to an improved understanding of tropospheric
49 chemistry and dynamics and improved forecasts of atmospheric pollutant fields (see, e.g., Elbern et al.,
50 2010). As part of an integrated observing strategy, satellite measurements provide a global view on air
51 quality (AQ). The challenge for future space-borne missions will be to assess directly the local scales of
52 transport and/or chemistry for tropospheric pollutants (1 hour or less, 10 km or less) and to facilitate the use
53 of remote sensing information for improving local- and regional-scale (from country-wide to continental
54 scales) AQ analyses and forecasts. Building on this effort, various LEO satellite platforms and/or
55 constellations of GEO satellite platforms will help extend AQ information from continental scales to global
56 scales (e.g., Lahoz et al., 2012, and references therein for LEO/GEO platforms; Barré et al., 2015, for GEO
57 platforms).

58
59 An atmospheric species of interest for monitoring AQ is CO, owing to its relatively long time-scale in the
60 troposphere; its distribution provides information on the transport pathways of atmospheric pollutants.
61 Spaceborne instruments on LEO satellite platforms demonstrate the potential of remote sensing from space

62 to determine the CO distribution and its main emission sources at the global scale (Edwards et al., 2004,
63 2006; Buchwitz et al., 2006; Worden et al., 2013; [Fu et al., 2016](#), [Warner et al., 2013](#), [George et al., 2015](#))
64 and references therein). These LEO satellite platforms include MOPITT (Measurements Of Pollution In The
65 Troposphere), IASI (Infrared Atmospheric Sounding Interferometer), AIRS (Atmospheric InfraRed
66 Sounder), TES (Tropospheric Emission Spectrometer) [and CrIS \(Cross-track Infrared Sounder\)](#) operating in
67 the thermal infrared (TIR) and SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric
68 ChartographY) operating in the short-wave infrared (SWIR), respectively. By contrast, to our knowledge,
69 there are no GEO satellite platforms measuring the CO distribution. However, despite their potential, owing
70 to limited revisit time, and relatively coarse spatial resolution, LEO instruments are not optimal for
71 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 [the European Organisation for the Exploitation of Meteorological Satellites \(EUMETSAT\)](#).
84 Among them, three missions address atmospheric composition. These are the Sentinel-5 (S-5) and Sentinel-5
85 Precursor (S-5P) from a LEO satellite platform, and the Sentinel-4 (S-4) from a GEO satellite platform. The
86 goal of the S-4 is to monitor key atmospheric pollutants (e.g., ozone; NO₂; sulphur dioxide, SO₂; bromine
87 monoxide, BrO; and formaldehyde) and aerosols at relatively high spatio-temporal resolution over Europe
88 and North Africa (8 km; 1 hour). [We expect launch of this mission in 2021 with a lifetime of 8.5 years.](#) The
89 goal of the S-5 and S-5P platforms is to provide global daily measurements of atmospheric pollutants (e.g.,

90 CO, ozone, NO₂, SO₂, BrO, and formaldehyde), climate related trace gases (e.g., methane, CH₄) and
91 aerosols, at relatively high spatial resolution (from below 8 km to below 50 km, depending on wavelength).

92

93 The S-5P is the ESA pre-operational mission required to bridge the gap between the end of the OMI (Ozone
94 Monitoring Instrument) and the SCIAMACHY missions and the start of the S-5 mission planned for 2020
95 onwards. The S-5P scheduled launch is in 2016 with a 7 years design lifetime. The S-5P will fly in an early
96 afternoon sun-synchronous LEO geometry with an Equator crossing mean local solar time of 13:30, chosen
97 to allow the instrument to measure the strong pollution signal present in the afternoon. [_We describe the](#)
98 [instrument TROPOMI \(TROPOspheric Monitoring Instrument\) onboard S-5P in section 2.2.](#) In contrast, ~~the~~
99 ~~GOME 2 (Global Ozone Monitoring Experiment — 2) platform~~ [IASI \(Infrared Atmospheric Sounding](#)
100 [Interferometer\) onboard Metop platform](#) collects data at a local solar time of 09:30 (when the pollution
101 signal is relatively weak) and thus has a lower predictive value (Veefkind et al., 2012, and references
102 therein). The S-5P LEO platform will address the challenge of limited revisit time from LEOs by providing
103 unprecedented high spatial resolution of 7x7 km, and [with its SWIR band,](#) improved sensitivity in the
104 Planetary Boundary Layer (PBL) [compared to a TIR instrument such as IASI.](#) The PBL varies in depth
105 throughout the year, but is contained within the lowermost troposphere (heights 0-3 km), and typically spans
106 the heights 0-1 km.

107

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

117

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

124
125 In this paper, we describe a regional-scale OSSE over Europe for northern summer 2003 (1 June – 31
126 August) to explore the impact of S-5P CO total column measurements on lowermost tropospheric air
127 pollution analyses, with a focus on CO PBL concentrations. The severe heat wave experienced in Europe
128 during northern summer 2003, and the concomitant atmospheric pollution and fire episodes, had a strongly
129 negative societal impact, being responsible for the deaths of over 14,000 people in France (Vautard et al.,
130 2005). This period had extremely hot and dry weather conditions and the long lasting atmospheric blocking
131 conditions significantly contributed to the accumulation of pollutants in the PBL owing to extended
132 residence time of the air parcels (Solberg et al., 2008). The spatial distribution of the enhanced levels of CO
133 and ozone was much more widespread over Europe during that summer than in previous ones (Lee et al.,
134 2006; Ordoñez et al., 2010). These exceptional weather conditions also resulted in several extreme wildfire
135 episodes over the Iberian Peninsula and the Mediterranean coast (Barbosa et al., 2004). Tressol et al. (2008)
136 point out that between 6 and 10 August 2003 the contribution of biomass burning to measured CO levels in
137 the lowermost troposphere reached 35% of the total CO field at these levels, a value comparable to typical
138 European anthropogenic emissions which represent 30% of this total CO field. Thus, the three-month period
139 1 June - 31 August 2003 includes both extreme and normal conditions, and provides an opportunity to study
140 the full range of pollution levels that occur in a summer season over Europe. [The improved knowledge of
141 CO distribution will improve its forecast and allows a better knowledge of the long range transport of
142 pollution plumes. In addition, CO being one of the ozone precursors, it is likely to use its measurement to
143 improve the ozone distribution calculated by the model.](#)

144
145 The OSSE study domain covers the larger part of Europe (5W-35E, 35N-70N), and we perform the OSSE
146 simulations at the spatial resolution of 0.2 degrees (latitude and longitude). This corresponds to a spatial

147 resolution of ~20 km (meridionally) and ~15 km (zonally, at 45N). With this spatial resolution, we can track
148 long-range transport plumes of CO. The length of the study period ensures we can sample different
149 meteorological situations typical for summertime, and provides an acceptable compromise between run-time
150 restrictions and provision of sufficient information for statistically significant results. [The focus of this OSSE](#)
151 [is CO concentrations and the goal is to evaluate the benefit of S-5P CO columns after assimilation in a](#)
152 [chemistry transport model, in particular CO concentrations at the surface.](#)
153

154 The structure of the paper is as follows. In Sect. 2 we describe the various components of the OSSE; in Sect.
155 3 we present the results from the OSSE for S-5P during summer 2003 over Europe. Finally, Sect. 4 provides
156 conclusions and identifies further work. A guiding principle in the OSSE set-up in this paper is to avoid
157 overoptimistic results.
158

159 **2. The OSSE set-up**

160 The OSSE concept consists of simulating observations and their associated errors from a representation of
161 reality (the "Nature Run" or NR) and providing this information to a data assimilation system to produce
162 estimates of the NR states. Thereafter, one compares these estimates of the NR states from an assimilation
163 run, AR (where the observation of interest has been assimilated), and from a control run, CR (in this case a
164 free model run), against the NR. The performance of the AR and the CR against the NR quantifies the benefit
165 of the observation of interest.
166

167 The OSSEs are widely used in the meteorological community for assessing the usefulness of new
168 meteorological satellite data. Recent examples (not exhaustive) include the work of Lahoz et al. (2005),
169 Stoffelen et al. (2006), and Tan et al. (2007); Masutani et al. (2010a) reviews the OSSE methodology and
170 provides a comprehensive list of references of OSSEs for meteorological applications. By contrast, there are
171 relatively few studies concerning OSSEs for AQ applications (Edwards et al., 2009; Timmermans et al.,
172 2009a, b; Claeys et al., 2011; Zoogman et al., 2011; 2014a, b; Yumimoto, 2013). In a recent review,
173 Timmermans et al. (2015) comment that documented AQ OSSEs have demonstrated the benefits that could
174 accrue from proposed and planned satellite platforms for AQ monitoring and forecasting. In the study

175 described in this paper, the set-ups for the NR, and the CR and AR, use different models, thereby avoiding
176 the identical twin problem typically associated with overly optimistic OSSE results (see, e.g., Masutani et al.,
177 2010a). In Sects. 2.1-2.5 we describe the various elements of the OSSE study described in this paper. Figure
178 1 provides a schematic showing the relationships between the various elements in an OSSE. [In this study, we](#)
179 [used the LOTOS-EUROS model as the NR and the MOCAGE \(Modèle de Chimie Atmosphérique de](#)
180 [Grande Echelle\) Chemistry Transport Model as the CR \(for details, see Sects. 2.1 and 2.4, respectively\)](#)

181

182 **2.1 The Nature Run**

183 A key element of an OSSE is the NR that defines the true state used to evaluate analyses and/or forecasts
184 using simulated observations. The NR commonly consists of a long, free-running forecast evolving
185 continuously in a dynamically consistent way (Masutani et al. 2010a, b). For this study, the basis of the NR
186 consists of two high-resolution free model simulations performed with: (i) the regional LOTOS-EUROS air
187 quality model (Schaap et al., 2008), and (ii) the global chemistry transport model TM5 (Huijnen et al., 2010).
188 We obtain the NR by combining the LOTOS-EUROS CO profiles from the surface to 3.5 km with the TM5
189 CO profiles from 3.5 km to the top of the atmosphere (identified by the TM5 model top at 0.1 hPa). We use
190 spatial interpolation to merge the values near the boundary between the two models at a height of 3.5 km.
191 The model simulations used to construct the NR have a spin-up period of three months. We archive the NR
192 output data on an hourly basis.

193

194 To construct the NR, we run the LOTOS-EUROS model at a horizontal resolution of about 7 km nested into
195 the TM5 model, the latter run with a zoom domain over Europe at 1x1 degrees resolution. The TM5 model
196 has 34 layers with a model top at 0.1 hPa. The LOTOS-EUROS model describes air pollution in the
197 lowermost troposphere. It has four vertical layers following the dynamic mixing layer approach. The first
198 layer is a fixed surface layer of 25 metres thickness, the second layer (boundary layer) follows the mixing
199 layer height, and there are two reservoir layers spanning the rest of the atmosphere up to 3.5 km. The implicit
200 assumption of the LOTOS-EUROS model is the presence of a well-mixed boundary layer, so constituent
201 concentrations are constant up to the top of the Planetary Boundary Layer. The meteorological data used as
202 input for the LOTOS-EUROS model come from the European Centre for Medium-Range Weather Forecasts
203 (ECMWF). Prescription of surface anthropogenic emission is from the TNO-MACC-II emission database

204 (Kuenen et al., 2014), and fire emissions are from the MACC global fire assimilation system (GFAS v1;
205 Kaiser et al., 2012).

206

207 In the design of an OSSE, it is important to demonstrate that the NR exhibits the same statistical behaviour
208 as the real atmosphere in every aspect relevant to the observing system under study (Masutani et al., 2010a,
209 b). For the LOTOS-EUROS model used to build the lowermost levels of the NR, there is extensive
210 verification by comparison with European data and by frequent participation in international model
211 comparisons. This is the case for ozone and particulate matter (see Hass et al., 2003; Cuvelier et al., 2007;
212 van Loon et al., 2007; Stern et al., 2008; Manders et al., 2009; Curier et al., 2012; [Marécal](#) et al., 2015). To
213 evaluate the NR, we compare the surface CO data to available [in situ](#) ground-based CO measurements over
214 Europe during northern summer 2003 (1 June – 31 August). For this comparison, we use the ground-based
215 stations from the Airbase database. We consider all types of ground-based stations from this database
216 because of the limited number of available measurements, but we discard stations with less than 75% of
217 hourly data within a month. This provides 171 ground-based stations for the comparison against the NR
218 (note this approach results in a paucity of stations over France).

219

220 Figure [2+](#) shows the location of the selected Airbase ground-based stations measuring CO over Europe
221 during northern summer 2003 (top panel), and the time-series of CO concentrations during 1 June – 31
222 August 2003, measured by the selected Airbase ground-based stations and simulated by the NR and the CR
223 (bottom panel [and see Sect. 2.4 for the definition of the CR](#)). Note that most ground-based stations selected
224 are located in polluted areas, where big emission sources of CO are present. We form the time-series from
225 the ground-based stations by averaging spatially over all the sites. We form the NR time-series similarly, but
226 interpolate the NR surface data to the station location. We do not add random observation errors to the NR
227 time-series.

228

229 From Fig. [2+](#), we see that, generally, the NR captures reasonably well the features of observed CO temporal
230 variability during the three phases characterizing the summer of 2003: before, during and after the heat wave
231 (the heat wave occurred on 31 July – 15 August). [One can notice that the observed and simulated CO time-](#)
232 [series exhibit some high frequency component due principally to the fact that the 171 sites representing these](#)

233 [time series are mostly located in emission source areas \(there are only 5 background rural sites among the](#)
234 [171 sites sample\)](#) . The CO diurnal pattern over a background rural site during the course of a summer day
235 [shows a peak between 7h and 8h in the morning. However, in polluted regions, the CO diurnal pattern shows](#)
236 [more variability](#). The correlation coefficient, ρ , between the ground-based data and NR time-series shown in
237 the middle panel is 0.71. From this, we conclude that the NR has a realistic representation of the CO diurnal
238 cycle. Note that CO concentration levels in the NR are slightly lower than observed ones. The bias of the NR
239 with respect to observed CO concentrations fluctuates around -10 % on average during normal conditions
240 and reaches -20% within the heat wave period. This means that the NR reproduces the surface concentrations
241 with a negative bias (NR lower than ground-based stations) between 10 and 20% . Nonetheless, the simulated
242 CO concentrations and those measured by the ground-based stations generally fall within the same range of
243 values (between 200 and 400 μgm^{-3}). Thus, for the OSSE period considered, we conclude that the NR is
244 representative of the variability of actual observations over the European domain, albeit with a negative bias.

245
246 Additionally, from Fig. 2+ the behaviour of the CO time-series from the CR compared to the NR, is similar
247 to the behaviour of the NR CO time-series compared to the Airbase data. [This suggests that the NR from](#)
248 [LOTOS-EUROS model from which we sample the S-5P simulated observations is reasonably realistic. This](#)
249 reduces the likelihood that the OSSE produces overoptimistic results.

250

251 **2.2 The S-5P CO simulated measurements**

252 The S-5P will deploy the TROPOspheric Monitoring Instrument (TROPOMI) jointly developed by The
253 Netherlands and ESA (Veefkind et al. 2012). The TROPOMI instrument has heritage from both the OMI and
254 the SCIAMACHY missions. The TROPOMI instrument will make measurements in the UV-visible
255 wavelength range (270-500 nm), the near infrared, NIR (675-775 nm) and the shortwave infrared, SWIR
256 (2305-2385 nm). It will deliver a key set of gas and aerosol data products for air quality and climate
257 applications, including ozone, NO_2 , formaldehyde, SO_2 , methane and CO.

258

259 To enable sounding of the lower atmosphere at finer scales, TROPOMI has an unprecedented spatial
260 resolution of $7 \times 7 \text{ km}^2$ at nadir. This relatively high spatial resolution is necessary for air quality applications
261 at local to regional scales. It will resolve emission sources with [15% of accuracy and 10% precision](#)

262 | [\(Veefkind et al., 2012\)](#), and will obtain an acceptable fraction of cloud-free spectra. In contrast to the
263 | advantages provided by the relatively high spatial resolution of S-5P and design improvements, the
264 | SCIAMACHY CO data needs averaging in time (roughly one month) and space (5x5 degrees) to obtain
265 | realistic CO distributions at comparable uncertainty (Galli et al., 2012). Furthermore, TROPOMI will have a
266 | wide swath of 2600 km to allow for daily global coverage. The relatively high radiometric sensitivity of S-5P
267 | will allow measurements at low albedo [\(order of 2%: Veefkind et al., 2012\)](#), thus helping track smaller
268 | pollution events and improving the accuracy of air quality assessments and forecasts. The use of S-5P CO
269 | total column measurements with inverse modelling techniques will also help quantify ~~more accurately~~
270 | biomass burning emissions and map their spatial distribution. The simultaneous measurements of CO and,
271 | e.g. NO₂, will provide additional information on wildfire and other pollution episodes (Veefkind et al., 2012).

272

273 | [We used](#) the NR results to generate a set of synthetic S-5P observations. This involves several steps: 1)
274 | Generating realistic S-5P orbits and geolocation and viewing/solar geometries for the appropriate overpass
275 | time, 2) Using the ECMWF modelled cloud distributions to generate effective cloud fractions, 3) [Generating](#)
276 | lookup tables for the averaging kernels and observation errors. 4) Collocation and application of the NR to
277 | derive a set of synthetic observations for [three](#) summer months and [three](#) winter months. [We discuss](#) these
278 | steps are discussed in the sub-sections below.

279

280 | **2.2.1 Orbit simulator**

281 | [We](#) use the System Tool Kit (STK, available from AGI, <http://www.agi.com/products/>) to generate the S-5P
282 | orbit geometry and the geolocation of the edges of the swath as a function of time. Based on these
283 | characteristics, [we generate](#) the location of the individual observations with a spatial distance of 7 km. [We](#)
284 | [apply](#) time and longitude shifts to the STK orbits to generate the orbits for the three [summer](#) and three [winter](#)
285 | months. Subsequently, [we compute](#) the solar and viewing geometries. Finally, [we maintain](#) segments of the
286 | orbits that have an overlap with the modelling domain.

287

288 | **2.2.2 Cloud properties**

289 [We obtain](#) cloud fields from the high-resolution operational weather forecast archive of the ECMWF. [We](#)
290 [retrieve](#) meteorological fields of liquid water content, ice water content, specific humidity and cloud fraction
291 at a resolution of 0.25 x 0.25 degree for June-August 2003 and November 2003 - January 2004. [We convert](#)
292 [these](#) quantities to cloud optical properties. The optical properties determine the reflectance, and [we use](#)
293 [them](#) to estimate effective cloud fractions and effective cloud top heights as retrieved from the satellite
294 observations (Acarreta et al., 2004). The distribution of effective cloud fractions was compared with the
295 distribution of effective cloud fractions obtained from OMI observations, and a reasonable agreement was
296 found for [summer](#) and [winter](#) months. [We derive the cloud fractions at the resolution of the ECMWF 0.25 x](#)
297 [0.25 degree grid. This is ca. 30 x 30 km² at the Equator and decreases as a function of latitude. The ground](#)
298 [pixel of OMI UV-2 and VIS channels is 13 x 24 km² at nadir increasing to 13 x 128 km² at edges of the](#)
299 [swath. We consider that the ECMWF grid cells and OMI pixels are of comparable size for comparing the](#)
300 [cloud fraction distributions \(ca. 0.5 million pixels or cells in each distribution\). We model clouds with a](#)
301 [simple Lambertian reflectors and ignore any wavelength dependency of cloud fraction.](#)

Mis en forme : Exposant

Mis en forme : Exposant

Mis en forme : Exposant

302
303 [We use](#) these effective cloud fractions (and corresponding cloud radiance fractions to provide weights to the
304 cloud-free and cloud-covered fractions of the surface scene. [We use](#) the cloud altitude for the computation of
305 the averaging kernel.

307 *2.2.3 Averaging kernel and measurement uncertainty lookup tables*

308 Because of the large number of observations that will become available from the S-5P instrument, full
309 radiative transfer calculations for each observation separately [are](#) not feasible. We have chosen to build look-
310 up tables for a set of geometries based on a radiative transfer code that employs the adding-doubling method
311 in combination with optimal estimation (radiative transfer toolbox DISAMAR; de Haan, 2012). Look-up
312 tables are set up for the averaging kernels (1D vectors as a function of altitude) and the measurement
313 uncertainty. Results are stored for a number of surface albedos, cloud/surface pressures, solar zenith angles,
314 viewing zenith angles and relative azimuth angles. [We provide](#) the look-up table details in Table 1. [We](#)
315 [provide](#) kernels on 21 pressure levels between 1050.0 and 0.1 hPa. [We specify uncertainties](#) for clear-sky
316 and cloudy-sky separately.

317

318 Each simulation with DISAMAR consists of a forward calculation of the satellite-observed spectrum,
319 followed by a retrieval step based on the optimal estimation method (Rodgers, 2000). [We convert](#) instrument
320 noise, listed in Table 1, into uncertainties for the retrieved CO column. [We take a-priori](#) trace gas profiles
321 from the CAMELOT study (Levelt et al., 2009). We assume that both the cloud and the surface are
322 Lambertian reflectors. Kujanpää et al. (2015) provide further details of this procedure.

323

324 In particular, the albedo is of major influence for the uncertainty, because it directly determines the signal
325 observed by the instrument. [We show](#) this dependence in Fig. [32](#). Over land, albedo values are typically of
326 the order of 0.1-0.2, with typical column errors of the order of 2 DU, or about 10^{17} molecules cm^{-2} . Because
327 typical CO columns over Europe are 2×10^{18} molecules cm^{-2} , this is a relatively small error of the order of
328 5%. These numbers are in good agreement with the results presented in the CO ATBD of TROPOMI
329 (document [available from https://sentinel.esa.int/web/sentinel/user-guides/sentinel-5p-tropomi/document-](https://sentinel.esa.int/web/sentinel/user-guides/sentinel-5p-tropomi/document-library)
330 [library](#)). Over [the ocean](#), the albedo is very low, and the noise dominates the signal. [To](#) simulate this
331 behaviour in a realistic way we have added the albedo values 0.005, 0.01 and 0.02 to the albedo list.

332

333 We note that the uncertainties reported here are substantially lower than reported for SCIAMACHY (e.g.
334 Gloudemans et al., 2008). This reflects a difference in specifications of the instruments, and the fact ice
335 build-up on the detectors [affected the SCIAMACHY observations](#). Real TROPOMI observations will show if
336 the relatively small errors are realistic.

337

338 **2.2.4 Synthetic observations generation**

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

- 340 • Collocation of the Nature run vertical profiles of CO to the locations of the observations.
- 341 • Computation of the effective cloud fraction, cloud radiance fraction, and cloud pressure from the
342 ECMWF cloud fields collocated to the observations.
- 343 • Collocation of the NIR albedo map (surface albedo at 2300 nm is interpolated from a climatology
344 provided by SRON and based on SCIAMACHY observations; P. Tol, personal communication) to
345 the locations of the observations.

- 346
- Extract interpolated values for the observation kernel and uncertainties from the look-up table.
- 347
- Compute the synthetic observation from the inner product of the kernel with the nature run CO
- 348 | profile. [We do this](#) for both a clear sky and fully clouded situation, using the cloud pressure.
- 349
- Add random noise amount to each observation, by drawing numbers from a Gaussian distribution
- 350 | with a width determined from the uncertainty estimate.
- 351
- Compute the partially clouded synthetic observation by weighting the clear and cloudy results with
- 352 | the cloud radiance fraction ([Landgraf et al., 2016](#); [Vidot et al., 2011](#)).
- 353

354 Over land, and in clear sky cases, the averaging kernel is close to 1, showing that the S-5P instrument is

355 | observing the vertical column to a good approximation (see Fig. [43](#)). In cloud-covered cases the kernel

356 | equals 0 for layers below the cloud pressure (yellow line in Fig. [43](#)). For low-albedo cases (over ocean),

357 | Rayleigh scattering becomes non-negligible, and the kernel ~~is~~ [decreases](#) towards the surface, but the noise

358 | is dominant in this case.

359

360 | [We show](#) the results of this process in Fig. [54](#). The figure demonstrates the high resolution of the NR (about

361 | 7 km) and the corresponding simulated amount of detail. The bottom panel shows the corresponding CO

362 | observations. Over land the NR features are clearly present due to the relatively low uncertainty. Over the

363 | ocean and Mediterranean, noise [dominates the signal](#). [We observe an](#) improved information content near

364 | Iceland, related to ~~a~~ thick cloud cover, where the higher signal reduces the relative noise.

365

366 **2.3 Pre-processing of S-5P CO total column observations**

367 This section describes the pre-processing of S-5P CO total column observations prior to assimilation into the

368 MOCAGE model (Peuch et al., 1999) for the OSSE simulations. Using the MOCAGE model for the AR and

369 CR simulations avoids the identical twin problem associated with using the same model for both the NR and

370 the OSSE simulations, which typically produces overoptimistic results (Arnold and Dey, 1986; Stoffelen et

371 | al., 2006). Section 2.4 provides further details of the MOCAGE model.

372

373 The S-5P will produce large amounts of data owing to its wide swath and relatively high spatial resolution of
374 about 7x7 km². Thus, a pre-processing step is necessary to reduce the data volume for the data assimilation
375 experiments. For this study, we consider only pixels inside the OSSE simulation domain (Note that retrieval
376 pixels in each single cross-track are essentially instantaneous measurements of CO₂). This has the advantage
377 of alleviating the data volume burden. However, a single cross-track over Europe could have more than
378 80,000 valid retrieval pixels. Furthermore, each individual pixel is associated with an averaging kernel vector
379 given at 34 vertical pressure levels, from the surface up to the top of the atmosphere (identified as 0.1 hPa).

380

381 Figure 43 shows an example of averaging kernels at the surface, as well as the averaging kernels
382 representative of retrievals including pixels with different cloud fractions (less than 10%, greater than 30%,
383 and greater than 80%). In addition, we discard data points with solar zenith angles larger than 80° or errors
384 exceeding 20%. The retrieval over sea is noise-dominated. Because of this, we only consider CO₂ partial
385 columns above cloudy sea scenes with cloud fraction more than 80% and cloud top heights between the
386 surface and 650 hPa. Finally, we apply a spatially weighted mean to bin the measurements into 0.2° x 0.2°
387 grid boxes (~20 x 15 km at 45N), the assimilation model resolution; this is the set-up used for the OSSE
388 assimilation experiments (CR and AR), and is described in El Amraoui et al. (2008a). It combines the
389 MOCAGE model and the PALM (Projet d'Assimilation par Logiciel Multiméthode) data assimilation
390 module. Sections 2.4-2.5 provide further details of the CR and AR set-ups.

391

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

393

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

394

395

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

399

$$\frac{1}{\bar{\sigma}^2} = \sum_t w_t$$

400

401

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

404

405 **2.4 The Control Run**

406 To generate the CR, it is important to use a state-of-the-art modelling system, which simulates the
407 observational data representing, for example, a current operational observational system. An important
408 requirement for an effective OSSE is to generate the CR with a model different [from](#) the one used to
409 construct the NR to avoid the identical twin problem (see Sect. 2.3). If the model from which we extract
410 hypothetical observations is the same as the assimilating model, the OSSE results tend to show unrealistic
411 observation impact and overly optimistic forecast skill (Arnold and Dey, 1986; Stoffelen et al., 2006).
412 Consequently, by using two independent models the OSSE will [simulate more realistically](#) the assimilation
413 of real observations. This [allows us](#) to design an OSSE that is not too overoptimistic.

414

415 In this OSSE study, the CR is a free model run [using MOCAGE](#). The MOCAGE model is a three-
416 dimensional CTM developed at Météo France (Peuch et al., 1999) providing the evolution of the atmospheric
417 composition in accordance with dynamical, physical and chemical processes. It provides a number of
418 configurations with different domains and grid resolutions, as well as various chemical and physical
419 parameterization packages. Current use of MOCAGE includes several applications: e.g., the Météo-France
420 operational chemical weather forecasts (Dufour et al., 2004); the Monitoring Atmospheric Composition and
421 Climate (MACC) services (<http://www.gmes-atmosphere.eu>; [Marécal](#) et al., 2015); and studies of climate
422 trends of atmospheric composition (Teyssède et al., 2007). Validation of MOCAGE simulations against a
423 large number of measurements took place during the Intercontinental Transport of Ozone and Precursors
424 (ICARTT/ITOP) campaign (Bousserez et al., 2007).

425

426 In this study, we use a two-way nesting configuration to generate the CR and the AR (we describe the AR
427 set-up in Sect. 2.5): a global grid with a horizontal resolution of 2x2 degrees and a regional grid (5W-35E,

428 35N-70N) with a horizontal resolution of 0.2x0.2 degrees. The MOCAGE model includes 47 sigma-hybrid
429 vertical levels from the surface up to 5 hPa. The vertical resolution is 40 to 400 m in the boundary layer (7
430 levels) and [approximately](#) 800 m [near](#) the tropopause and in the lower stratosphere. The chemical scheme
431 used is RACMOBUS, which combines the stratospheric scheme REPROBUS ([REactive Processes Ruling](#)
432 [the Ozone BUdget in the Stratosphere](#); Lefèvre et al., 1994) and the tropospheric scheme RACM ([Regional](#)
433 [Atmospheric Chemistry Mechanism](#); Stockwell et al., 1997). The RACMOBUS scheme includes 119
434 individual species, of which 89 are prognostic variables, and considers 372 chemical reactions.

435

436 We force the CR (and the AR) every 3 hours with the ARPEGE analysis (Courtier et al., 1991). We prescribe
437 the surface anthropogenic emission using the MACC-I emission database ([https://gmes-](https://gmes-atmosphere.eu/about/project_structure/input_data/d_emis/)
438 [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
439 CR and AR experiments described in this paper, as their [a priori](#) [is un](#)known. This means that any signature
440 of fire emissions in the AR (see Sect. 2.5) can only come from assimilation of the CO measurements. Note
441 that for the NR, the surface anthropogenic emissions come from the MACC-II inventory, [which helps](#) to
442 differentiate the CR from the NR. [Similar to](#) the NR, the CR has a spin-up period of three months.

443

444 **2.5 The Assimilation run**

445 We assimilate simulated S-5P total column CO observations derived from the LOTOS-EUROS NR into the
446 MOCAGE CTM at a 0.2° spatial resolution using the MACC extended domain (5W-35E, 35N-70N). The
447 assimilation system used in this study is MOCAGE-PALM (e.g., El Amraoui et al., 2008a) developed jointly
448 by Météo-France and CERFACS (Centre Européen de Recherche et de Formation Avancée en Calcul
449 Scientifique) in the framework of the ASSET European project (Lahoz et al., 2007b). The assimilation
450 module used in this study is PALM, a modular and flexible software, which consists of elementary
451 components that exchange [the](#) data (Lagarde et al., 2001). It manages the dynamic launching of the coupled
452 components (forecast model, algebra operators and input/output of observational data) and the parallel data
453 exchanges. Massart et al. (2009) used the assimilation system MOCAGE-PALM to assess the quality of
454 satellite ozone measurements. The MOCAGE-PALM assimilation system also helps identify and overcome
455 model deficiencies. In this context, its assimilation product has been used in many atmospheric studies in
456 relation to ozone loss in the Arctic vortex (El Amraoui et al., 2008a); tropics/mid-latitudes exchange

457 (Bencherif et al., 2007); stratosphere-troposphere exchange (Semane et al., 2007); and exchange between the
458 polar vortex and mid-latitudes (El Amraoui et al., 2008b). For this OSSE, to speed up the assimilation
459 process we use the 3D-Var version of PALM. In the OSSE, the MOCAGE model provides the CR and by
460 assimilating the simulated CO data from the NR, the MOCAGE model provides the AR. Thus, we produce
461 the CR and AR outputs with a model different from that used to produce the NR (see Sect. 2.1).

462

463 A key element of the data assimilation system is the background error covariance matrix (the **B**-matrix)
464 (Bannister, 2008). It has a large impact on the 3D-Var analysis used in this study and, thus, it is important to
465 use a form of **B** that is as realistic as possible. In MOCAGE-PALM, we base the **B**-matrix formulation on
466 the diffusion equation approach (Weaver and Courtier, 2001). It can be fully specified by means of the 3-D
467 standard deviation field (square root of the diagonal elements of **B**, in concentration units or as a percentage
468 of the background field) and 3-D fields of the horizontal (L_x and L_y) and vertical (L_z) local correlation
469 length-scales. We can estimate the **B**-matrix elements more efficiently using an ensemble method (Bannister,
470 2008). This technique consists of feeding an ensemble of states through the data assimilation system to
471 simulate the important sources of error. However, this approach is time-consuming and, therefore, not used in
472 this study.

473

474 For this study, we use a simple parameterization for the **B**-matrix, where L_x and L_y are assumed
475 homogeneous and equal to 35 km (about two model grid lengths); and L_z is constant and set to one vertical
476 model layer. As in Emili et al. (2014), the background standard deviation 3-D field is parameterized as a
477 vertically varying percentage of the background profile, which decreases from values of 25% at the surface
478 to values of 15% in the upper troposphere, and decreases further throughout the stratosphere to values of 5%
479 in the upper stratosphere (not shown). We base these settings on several 1-day assimilation trials; they ensure
480 reasonable values of standard self-consistency tests, e.g., providing chi-squared (χ^2) values close to 1 (see
481 Fig. 63 in Sect. 3.1). Furthermore, a value of L_x and L_y of 35 km corresponds to more than one grid length of
482 the model, allowing the model to resolve these features. The data assimilation procedure will weight both the
483 observations and the model 1-hour forecasts (from the last analysis point), and will update locations not
484 coincident with the observations through the correlation length-scales. Table 2 summarizes the parameters
485 used for the assimilation experiments.

486 **3. Results**

487 **3.1 Evaluation of the assimilation run**

488 In this section, we evaluate the impact of the assimilation of the S-5P CO total column. First, we evaluate the
489 consistency of the assimilation run by separating the clear-sky pixels from their cloudy counterparts (Sect.
490 3.1.1). Second, to further understand the impact on the surface CO field of the simulated S-5P CO total
491 column measurements, we investigate the analysis increment (δx) to provide a quantitative diagnostic of the
492 quality of the analysis for a selected date, 15 June 2003 (Sect. 3.1.2).

493

494 **3.1.1 Consistency of the assimilation run**

495 We have performed two OSSEs. The first one includes all pixels in the OSSE domain, regardless of whether
496 they are cloudy or clear-sky and the second only includes clear-sky pixels. [We consider a pixel](#) as clear when
497 the cloud fraction is less than 10%. Comparison of the ARs from these two OSSEs indicates [that the impact](#)
498 of including all pixels is small. The largest differences between the respective ARs in relation to the NR are
499 4% in regions over North Europe (North Sea and Scandinavia), with the AR for clear-sky pixels closer to the
500 NR (not shown). We can explain these results by the fact the summer generally has low amounts of cloud.
501 Consequently, we only present the results from the OSSE with all pixels.

502 To evaluate the AR, we calculate the χ^2 diagnostic associated with the Observation minus Forecast (OmF)
503 differences (see, e.g., Lahoz et al., 2007a). Here, we normalize the OmF differences by the background error.
504 We also calculate histograms of the Observation minus Analysis (OmA) differences, the observation and the
505 simulation from the CR (observation-minus-control run, hereafter OmC) differences, and the OmF
506 differences. We use the observational error to normalize the differences building the histograms of OmA,
507 OmC and OmF.

508

509 Figure [65](#) (top panel) shows the chi-squared time-series for OmF and its associated auto-correlation function
510 calculated over the three-month period of the OSSE experiments, computed as daily averages. The chi-
511 squared diagnostic starts with a maximum of about 1.56, and takes values down to 0.75, with a mean of 0.9
512 over the OSSE three-month period. The chi-squared time-series is nearly stable since it exhibits relatively

513 small variability (a standard deviation of about 0.14). Furthermore, the auto-correlation of the chi-squared
514 statistic drops to zero, with no correlation after a time delay of 20 days. The calculation of the auto-
515 correlation shows that the chi-squared statistic is uncorrelated after a time lag of 20 days; this means that
516 after this time the mathematical expectation $E(\chi^2)$ is equal to the average of the chi-squared statistics. We
517 find $E(\chi^2) = 0.90$, which is close to the theoretical value of 1 (see Lahoz et al., 2007a). This result indicates
518 that the a priori error statistics as represented in the **B**-matrix slightly overestimate the actual error statistics
519 from the OmF differences.

520

521 To test whether the observations, forecast and analysis fields, and their associated errors, are consistent with
522 each other, we calculate the histograms of OmA, OmF and OmC only over land (normalized by the
523 observation error) over the three-month period (Fig. 65, bottom panel). For a properly set up assimilation
524 system, the OmF and OmA normalized histograms should be close to a Gaussian distribution with mean zero
525 and standard deviation one. Figure 63 (bottom panel) shows that the OmA and OmF differences are close to
526 a Gaussian distribution centred near to or at zero. The OmF has a mean and standard deviation of 0.10 and
527 1.73, respectively, whereas the OmA has nearly a zero mean and a standard deviation of 1.05. This indicates
528 that the centre of the OmA histogram is closer to zero and more peaked than the histogram of OmF. We
529 expect this, since the analyses should be closer to the observations than the forecasts. Furthermore, the
530 histogram for OmA indicates that the errors in the **R**-matrix, the observational counterpart of the **B**-matrix,
531 are a good representation of the analysis error.

532

533 Based on the above results, we conclude that the background error covariance matrix, **B**, and its
534 observational counterpart, **R**, prescribed in our assimilation system are reasonably well characterized (see,
535 e.g., Lahoz et al., 2007a, for a discussion of the specification of errors in a data assimilation system).
536 Furthermore, the above results are consistent with the assumption that the errors in the observations and the
537 forecasts are Gaussian.

538

539 The shape of the OmC normalized histogram, which has a mean and standard deviation of 2.36 and 5.60,
540 respectively, indicates the presence of a relatively large bias between the S-5P observations and the CR. The
541 assimilation reduces this bias, as shown by the analyses being significantly closer to the observations than

542 the simulation from the CR. This shows that the assimilation of simulated S-5P CO total column
543 observations has a significant impact on the CO forecasts and analyses.

544

545 *3.1.2 Study of increments*

546 To understand further the impact on the surface CO field of the simulated S-5P CO total column
547 measurements, we calculate the analysis increment (δx) for a single analysis time at 14:00 UTC on 15 June
548 2003. We calculate this increment as the analysis minus the model first guess (1-hour forecast). The analysis
549 increment provides a quantitative diagnostic of the quality of the analysis (see, e.g., Fitzmaurice and Bras,
550 2008).

551

552 | Figure 76 (top panel) shows the spatial distribution of δx at the model surface. One can see the spread of the
553 | impact of the simulated observations across large regions. This is owing to S-5P having a wide swath
554 | allowing it to sample larger regions. The most substantial corrections are over land, where there are sufficient
555 | observations to have an impact. Over sea, the increments tend to be negligible, as any observations found
556 | there have relatively large errors. Thus, there will not be much difference between the model first guess and
557 | the analysis. Likewise, this is also true in the regions outside the satellite footprint.

558

559 To provide further insight into the impact of S-5P CO measurements, we calculate latitude-height and
560 | longitude-height cross-sections at 48.8N, 2.6E, near Paris, for 15 June 2003. Figure 54 (bottom left and
561 | bottom right panels) shows a zoom of the zonal and meridional vertical slices of the analysis increment. We
562 | see significant corrections to the model first guess (identified by large increments) confined to a deep layer.
563 | These corrections are larger at the surface, and exhibit a second maximum around 650 hPa. This vertical
564 | structure is mainly attributable to the forecast error standard deviation (given as a vertically varying fraction
565 | of the local CO mixing ratio), the square root of the diagonal entry of the **B**-matrix, and which is higher in
566 | the boundary layer (where the value of the S-5P CO averaging kernel is close to 1). The shape of the S-5P
567 | analysis increments also exhibits a second peak around 650 hPa. The increments for this particular day thus
568 | show a clear impact from the S-5P CO measurements in the PBL and the free troposphere.

569

570 The shape of the S-5P increments is similar to that of typical SCIAMACHY analysis increments, which also
571 extend through a deep layer and have a maximum at the surface (Tangborn et al., 2009). The fact that both
572 these analysis increments stretch out over a deep layer is owing to similarities in the S-5P and SCIAMACHY
573 averaging kernels - both are close to unity over cloud-free land (see Fig. 54). Note that the situation shown in
574 Fig. 76 is a snapshot and depends on the particular conditions for this time. An average of the increments
575 over the summer period would tend to show a uniform distribution in height.

576

577 **3.2 Evaluation of the summer OSSE**

578 **3.2.1 Summer averages**

579 Figure 87 shows the fields of surface CO from the CR, and the NR and the AR, averaged over the northern
580 summer period. One can see the general change of CO over land between the CR (top left panel) and the AR
581 (bottom panel). We can ascribe this to the contribution of simulated S-5P total column CO data sampled from
582 the NR. This figure shows several differences between the CR and AR fields that indicate the superior
583 behaviour of the AR in capturing features in the NR. For example, over Eastern Europe and Russia, the AR
584 CO concentration values are closer to those in the NR (with a mean bias between -1.5 and +1.5 ppbv); in
585 particular, the CR shows generally lower values than in the NR (mean bias around -6 ppbv). Nevertheless,
586 over Portugal, where the NR shows the forest fires that occurred over the summer, the AR captures them
587 only slightly better than the CR. We expect the relatively poor performance of the CR regarding fires, as the
588 fires are not included in the CR set-up (see Sect. 2.4). Although the AR, in the operational set-up, captures
589 the CO concentrations emitted by forest fires slightly better than the CR (through assimilation of CO
590 measurements), the relatively poor temporal resolution of the S-5P ultimately limits its performance.
591 However, the most important deficiency is due to the criterion used in the operational set-up in which a data
592 screening test is activated to discard observations far away from the model (see section 3.2.5). A
593 geostationary satellite, given its relatively high temporal resolution, should be able to capture better the
594 temporal variability of CO from these forest fires (Edwards et al., 2009).

595

596 **3.2.2 Statistical metrics**

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

600
601 We calculate the mean bias (MB, in parts per billion by volume, ppbv), its magnitude reduction (MBMR,
602 ppbv), and the root mean square error (RMSE, ppbv), and its reduction rate (RMSERR, %). Note that
603 although recent papers have raised concerns over the use of the RMSE metric (Willmott and Matsuura, 2005;
604 Willmott et al., 2009), Chai and Draxler (2014) discuss circumstances where the RMSE is more beneficial.
605 We use the correlation coefficient, ρ to measure the linear dependence between two datasets, and the fraction
606 of the true variability (i.e., variability represented by the NR) reproduced by the CR or AR.

607
608 For a single model grid box, we define the statistical metrics (MB, RMSE, ρ) with respect to the NR as:

609

$$610 \quad MB(X) = \frac{1}{N} \sum (X - NR)$$

611

$$612 \quad MBMR = |MB(CR)| - |MB(AR)|$$

613

$$614 \quad RMSE(X) = \sqrt{\frac{1}{N} \sum (X - NR)^2}$$

615

$$616 \quad RMSERR = 100 \times \left(1 - \frac{RMSE(AR)}{RMSE(CR)} \right)$$

617

$$618 \quad \rho(X) = \frac{\sum (X - \bar{X})(NR - \overline{NR})}{\sqrt{\sum (X - \bar{X})^2 \sum (NR - \overline{NR})^2}}$$

619

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

623

624 3.2.3 Results of the statistical tests

625 Figure 98 presents the zonal and meridional means of the difference between the CR and the AR averaged
626 over the northern summer 2003 (1 June – 31 August). We also plot the confidence interval representing the
627 areas where the AR is not significantly different to the CR at the 99% confidence limit (highlighted in the
628 grey colour). These two figures show that there is benefit from the S-5P CO total column data over the first
629 few bottom levels of the troposphere, i.e., the lowermost troposphere. Between the surface and 800 hPa, a
630 negative peak is present in the zonal difference field (over Scandinavia), and in the meridional difference
631 field (over Eastern Europe). Note that the zonal field shows two areas, one with positive values and the other
632 with negative values representing a CR greater than the AR and a CR smaller than the AR, respectively. The
633 positive peak, at a slightly higher level (i.e., lower pressure) than the negative peak, is representative of the
634 Mediterranean Sea, whereas the negative peak is more representative of the land areas (Scandinavia and
635 Eastern Europe). Figure 98 indicates that the S-5P CO corrects the model in the lower troposphere with a
636 larger impact over land and with a [smaller](#) impact in the PBL. This is consistent with the behaviour of the
637 analysis increments shown in Fig. 76.

638

639 Figure 109 shows the performance of the biases between the CR and the NR, and the AR and the NR at the
640 surface, and averaged over the northern summer of 2003 (1 June – 31 August). The MBMR, which compares
641 the magnitude of the CR vs NR and AR vs NR biases, indicates the geographical areas where the simulated
642 S-5P CO total column data have the most impact. The MBMR shows that the AR is closer to the NR than the
643 CR, almost everywhere in the domain (reflected by the prevalence of the red colours in the bottom left
644 panel). This indicates that the simulated S-5P CO total column data generally provide a benefit at the surface,
645 and especially over land areas where the CO sources are sparse. [This suggests that owing to the relatively
646 small variability of CO over remote land regions, the S-5P data can provide a larger benefit compared to
647 regions where the variability is relatively high.](#)

648

649 We also calculate the RMSE as well as the reduction rate of the RMSE, RMSERR (Figure 1140), both
650 keeping the systematic error (Fig. 1140 top), and removing the systematic error (Fig. 1140 bottom). We
651 calculate the ~~systematic error~~ [bias](#) in the AR and CR by subtracting the NR field from each of them,

652 producing a ~~unbiased~~~~debiased~~ AR and CR. For the case where we remove the systematic error, we perform
653 the statistics on the ~~unbiased~~~~debiased~~ AR and CR. If we examine the RMSE statistics, Fig. 98 shows that the
654 CR gets closer to the NR over the Atlantic Ocean and over the Eastern domain including Russia and
655 Scandinavia, when we remove the systematic error. For example, over these areas we obtain ~30 ppbv and
656 ~10 ppbv for the RMSE keeping and removing the systematic error, respectively. For the reduction of the
657 RMSE, RMSERR, the behaviour for the CR is similar overall, showing a reduction rate of 60% and 30-45%
658 keeping and removing the systematic error, respectively. Note that over Scandinavia the reduction rate goes
659 down from 60% to about 10% after removing the systematic error.

660

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

666

667 In Figure 124, we show the correlation between the CR and the NR, and the correlation between the AR and
668 the NR, at the surface for the three northern summer months (1 June – 31 August). Figure 11 shows that the
669 AR is closer than the CR to the NR with the correlation coefficient reaching 0.9 over land. By contrast, the
670 correlation coefficient between the CR and the NR is typically less than 0.5, with very low values over
671 Eastern Europe, where CO sources are sparse.

672

673 3.2.4 Time-series at selected locations

674 Figure 134 shows time-series from the NR, the CR and the AR over the three areas of the study domain
675 represented by the squares shown in Figs. 109 (bottom panel) and 114 (right panels). (i) The Paris region
676 (Fig. 134, top panel). (ii) A region over Portugal (5°W-40°N), where forest fires occurred during the
677 northern summer (Fig. 134, middle panel). (iii) An area in the Eastern part of the study domain (25°E-
678 53°N), where the reduction of RMSE (i.e., RMSERR) was much larger than for other regions (Fig. 134,
679 bottom panel). For all three areas, the AR is generally closer to the NR than the CR, showing the impact of
680 the simulated observations. We calculate the biases between the AR and CR vs the NR by computing the

681 difference NR-X, where X is AR or CR, and normalizing by the number of observations over the northern
682 summer period. The biases are: (i) Paris region, CR: 48 ppbv, AR: 38 ppbv; (ii) Portugal, CR: 101 ppbv, AR:
683 83 ppbv; (iii) Eastern part of domain: CR: 21 ppbv, AR: 5 ppbv. [Note that the AR and the CR capture the](#)
684 [variability but not the values of the peaks. However, the LEO only samples at most twice a day over Paris](#)
685 [and may not capture the peaks. In figure 13, we indicate the S-5P revisit time by the plus signs at the top of](#)
686 [the panel and one can see that the peaks do not coincide with the time of the S-5P measurements. Another](#)
687 [factor could also be that the emission inventory used in the AR has lower values than the one used in the NR.](#)
688

689 Over Paris (top panel), the CR is already close to the NR and the impact of the S-5P CO simulated
690 observations is small. Over Portugal (middle panel), the presence of fires is not seen in the CR (e.g., a
691 maximum of CO at the beginning of the heat wave), as the fire [emissions](#) were not taken into account in the
692 CR as they are [unknown a priori](#) (see Sect. 2.4). In contrast, over this specific location we see the impact of
693 the fires on the CO concentrations in the AR with, however, much lower values than for the NR. During the
694 fires, the CO concentrations in the AR over Portugal were larger than 500 ppbv, whereas the CR remained
695 relatively unchanged with concentrations less than 200 ppbv. Over the Eastern part, [where there are lower](#)
696 [emissions compared for instance to Paris](#), (bottom panel), the temporal variability is not high and the
697 magnitude of the bias between the CR and the NR is small, but it is removed in the AR. [Moreover, note that](#)
698 [the operational screening test was still in force \(see section below\).](#)
699

700 *3.2.5 Sensitivity tests for fire episode*

701 The assimilation system we use has a default criterion to discard CO column observations with values larger
702 than 75% of the MOCAGE value. This criterion is not appropriate to situations resulting in excessive values
703 in the CO concentrations, as is the case for forest fires. To understand further the performance of the OSSE
704 over the period of the Portugal forest fires we perform a second OSSE without this default criterion. This
705 second OSSE covers the period of the forest fires (25 July – 15 August). For this second OSSE, we compare
706 the total column values and the surface values of the CO fields from the CR and the AR (Figs. [14-16](#)~~13-15~~,
707 respectively).
708

709 | Figure [14+3](#) shows the CO total column at 14:15 UTC on 4 August 2003 (during the period of the Portugal
710 | forest fires) from the NR (top left panel); the simulated S-5P observations (top right panel); the CR (bottom
711 | left panel); and the AR (bottom right panel). We can see that the AR captures the fire event, indicated by
712 | relatively high values of the CO total column over Portugal, whereas the CR does not. This confirms the
713 | results shown in Fig. [13+2](#), which highlight the benefit provided by the S-5P CO total column measurements,
714 | in particular regarding the capture of the signature of the Portugal forest fires. Note that the S-5P
715 | measurement is noise-dominated over the sea (top right panel). This accounts for the sharp edge in the CO
716 | total column field seen between the Iberian Peninsula and the Bay of Biscay for the AR (bottom right panel).
717

718 | Figure [15+4](#) shows the time-series of the surface CO concentrations over the period 25 July – 15 August (that
719 | of the Portugal forest fires). In comparison to the original OSSE (see middle panel of Fig. [13+2](#)), the AR is
720 | now closer to the NR, having now peak values of about 900 ppbv, instead of peak values of about 550 ppbv.
721 | The CR still has peak values less than 200 ppbv. This indicates that the relatively low values in the AR (in
722 | comparison to the NR) for the original OSSE shown in the middle panel of Fig. [13+2](#) result from the
723 | application of the default criterion to discard CO column observations that are far away from MOCAGE
724 | values. The results from Fig. [15+4](#) confirm those shown in Fig. [14+3](#), and reinforce the benefit provided by
725 | the S-5P CO total column measurements, in particular regarding the capture of the signature of the Portugal
726 | forest fires. [This sensitivity test also shows the limitations of using standard operational criteria.](#)
727

728 | **4. Conclusions**

729 | We perform a regional-scale Observing System Simulation Experiment (OSSE) over Europe to explore the
730 | impact of the LEO satellite mission S-5P carbon monoxide (CO) total column measurements on lowermost
731 | tropospheric air pollution analyses, with a focus on CO surface concentrations and the Planetary Boundary
732 | Layer (PBL). The PBL varies in depth throughout the year, but is contained within the lowermost
733 | troposphere (heights 0-3 km), and typically spans the heights 0-1 km. We focus on northern summer 2003,
734 | which experienced a severe heat wave with severe societal impact [over Europe](#).
735 |
736

737 ~~This OSSE study provides insight on the impact from LEO S-5P CO measurements on surface CO~~
738 ~~information. We perform the standard steps of an OSSE for air quality. (i) Production of a Nature Run, NR.~~
739 ~~(ii) Test of the realism of the NR. (iii) Different models to produce, on the one hand, the NR, and on the other~~
740 ~~hand, the OSSE experiments to create the Control Run, CR, and the Assimilation Run, AR. (iv) Calculation~~
741 ~~of synthetic observations, observation uncertainty, and averaging kernels to represent sensitivity of the~~
742 ~~observations in the vertical. (v) Quantitative evaluation of the OSSE results, including performing statistical~~
743 ~~significance tests, and self-consistency and chi-squared tests. Based on the specifications of the TROPOMI~~
744 ~~instrument, relatively low CO column uncertainties of around 5% are anticipated over the European~~
745 ~~continent.~~

746
747 Our guiding principle in the set-up of this OSSE study is to avoid overoptimistic results. To achieve this, we
748 address several factors considered likely to contribute to an overoptimistic OSSE. (i) We use different
749 models for the NR and the OSSE experiments. (ii) We check that the differences between the NR and actual
750 measurements of CO are comparable to the CO field differences between the model used for the OSSE and
751 the NR. (iii) We remove the systematic error (calculated as the bias against the NR) in the OSSE outputs (AR
752 and CR) and compare the ~~unbiased~~~~debiased~~ results to the NR. (iv) We perform a quantitative evaluation of
753 the OSSE results, including performing statistical significance tests, and self-consistency and chi-squared
754 tests. Based on the specifications of the TROPOMI instrument, we anticipate relatively low CO column
755 uncertainties of around 5% over the European continent.
756 Also, our approach was to study the performance of S-5P alone without taking into account the other existing
757 or future missions (i.e. MOPITT, CrIS or IASI).

758
759 The OSSE results indicate that simulated S-5P CO total column measurements during northern summer 2003
760 benefit efforts to monitor surface CO. The largest benefit occurs over land in remote regions (Eastern
761 Europe, including Russia) where CO sources are sparse. Over these land areas, and for the case when we
762 remove the systematic error, we obtain a lower RMSE value (by ~10 ppbv) for the AR than for the CR, in
763 both cases vs the NR. Over sea and Scandinavia, we also obtain a lower RMSE (by ~10%) for the AR than
764 for the CR, in both cases vs the NR. Consistent with this behaviour, we find the AR is generally closer to the
765 NR than the CR to the NR, with a correlation coefficient reaching 0.9 over land (NR vs AR). By contrast, the

766 correlation coefficient between the CR and the NR is typically less than 0.5, with very low values over
767 Eastern Europe, where CO sources are sparse. In general, for all the metrics calculated in this paper, there is
768 an overall benefit over land from the S-5P CO total column measurements [in the free troposphere, but also at](#)
769 [the surface](#). Significance tests on the CR and AR results indicate that, generally, the differences in their
770 performance are significant at the 99% confidence level. This indicates that the S-5P CO total column
771 measurements provide a significant benefit to monitor surface CO.

772
773 We further show that, locally, the AR is capable of reproducing the peak in the CO distribution at the surface
774 due to forest fires (albeit, weaker than the NR signal), even if the CR does not have the signature of the fires
775 in its emission inventory. A second OSSE shows that this relatively weak signal of the forest fires in the AR
776 arises from the use of a default criterion to discard CO total column observations too far from model values,
777 a criterion not appropriate to situations resulting in excessive values in the CO concentrations, as is the case
778 for forest fires. This second OSSE shows a much stronger signal in the AR, which is now much closer to the
779 NR than the CR, confirming the benefit of S-5P CO total column measurements [and the limitations of using](#)
780 [standard operational criteria in this case.](#)

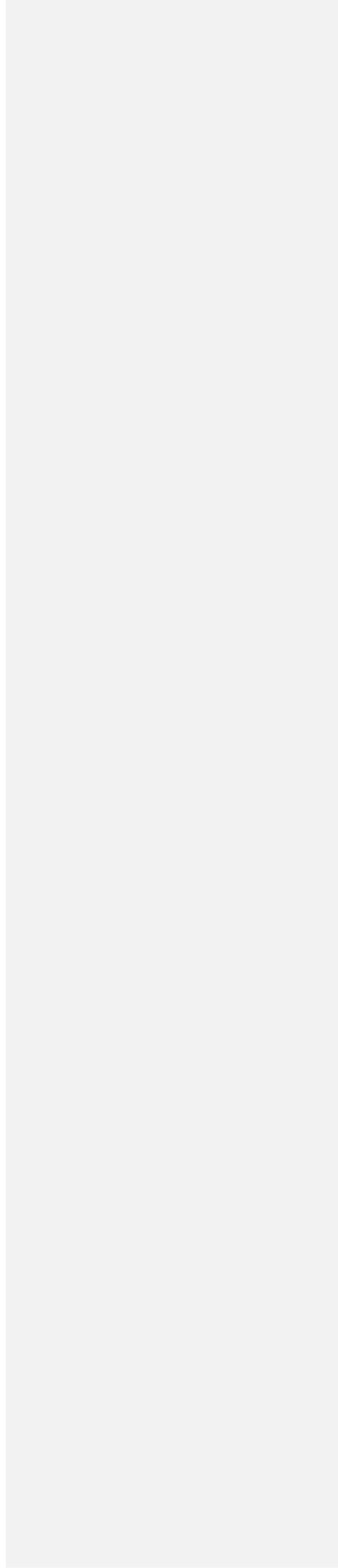
781
782 Further work will involve extending the OSSE approach to other S-5P measurements, such as ozone total
783 column, and NO₂ and formaldehyde tropospheric columns. These studies will complement similar studies on
784 the benefit from Sentinel-4 and -5 measurements. Collectively, these OSSE studies will provide insight into
785 the relative benefits from the Sentinel-4, -5 and -5P platforms for monitoring atmospheric pollution
786 processes.

787

788 **5. Acknowledgments**

789 Support for this work came partly from the ESA funded project “Impact of Spaceborne Observations on
790 Tropospheric Composition Analysis and Forecast” (ISOTROP –ESA contract number 4000105743). WAL
791 acknowledges support from an internal project from NILU. RA, JLA, PR, LE and WAL acknowledge
792 support from the RTRA/STAE. JK and JT acknowledge support from the Academy of Finland (Project no.
793 267442).

794



796 **6. References**

- 797 Acarreta, J. R., J. F. De Haan, and P. Stammes (2004), Cloud pressure retrieval using the O₂-O₂ absorption
798 band at 477 nm, *J. Geophys. Res.*, 109, D05204, doi:10.1029/2003JD003915.
- 799 Arnold, C.P., Jr., and C.H. Dey, 1986: Observing-systems simulation experiments: Past, present and future.
800 *Bull. Amer. Meteorol. Soc.*, 67, 687–695.
- 801 Atlas, R. 1997: Atmospheric observation and experiments to assess their usefulness in data assimilation. *J.*
802 *Meteor. Soc. Jpn.*, 75, 111–130.
- 803 Atlas, R., G.D. Emmitt, Terry, E. Brin, J. Ardizzone, J.C. Jusem, et al., 2003: Recent observing system
804 simulation experiments at the NASA DAO, in Preprints, 7th Symposium on Integrated Observing
805 Systems (Long Beach, CA: American Meteorological Society).
- 806 Bannister, R.N., 2008: A review of forecast error covariance statistics in atmospheric variational data
807 assimilation. I: characteristics and measurements of forecast error covariances. *Q. J. R. Meteorol. Soc.*,
808 134, 1951–1970. doi: 10.1002/qj.339.
- 809 Barbosa, P., J. San-Miguel-Ayanz, A. Camia, M. Gimeno, G. Liberta, and G. Schmuck, 2004: Assessment of
810 fire damages in the EU Mediterranean Countries during the 2003 Forest Fire Campaign. Official
811 Publication of the European Commission, S.P.I.04.64, Joint Research Center, Ispra, 2004.
- 812 Barré, J., D. Edwards, H. Worden, A. Da Silva, and W. Lahoz, 2015: On the feasibility of monitoring air
813 quality in the lower troposphere from a constellation of northern hemisphere geostationary satellites
814 (Part 1). *Atmos. Env.*, 113, 63–77, doi: 10.1016/j.atmosenv.2015.04.069.
- 815 Bencherif, H., L. El Amraoui, N. Semane, S. Massart, D.C. Vidyaranya, A. Hauchecorne, and V.-H. Peuch,
816 2007: Examination of the 2002 major warming in the southern hemisphere using ground-based and
817 Odin/SMR assimilated data: stratospheric ozone distributions and tropic/mid-latitude exchange. *Can. J.*
818 *Phys.*, 85, 1287–1300.
- 819 Bousserez, N., J.L. Attié, V.-H. Peuch, M. Michou, G. Pfister, D. Edwards, L. Emmons, C. Mari, B.
820 Barret, S.R. Arnold, A. Heckel, A. Richter, H. Schlager, A. Lewis, M. Avery, G. Sachse, E.V.
821 Browell, and J.W. Hair, 2007: Evaluation of the MOCAGE chemistry transport model during
822 the ICARTT/ITOP experiment, *J. Geophys. Res.*, 112, D10S42, doi: 10.1029/2006JD007595.
- 823 Buchwitz, M., R. de Beek, S. Noël, J.P. Burrows, H. Bovensmann, O. Schneising, I. Khlystova, M. Bruns, H

824 Bremer, P. Bergamaschi, S. Körner and M. Heimann Atmospheric carbon gases retrieved from
825 SCIAMACHY by WFM-DOAS: version 0.5 CO and CH₄ and impact of calibration improvements on
826 CO₂ retrieval Atmospheric Chemistry and Physics 6, 2727–2751, 2006

827 Chai, T., and R.R. Draxler, 2014: Root mean square error (RMSE) or mean absolute error (MAE)?
828 Arguments against avoiding RMSE in the literature. Geosci. Model Dev., 7, 1247-1250.

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

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

835 Curier, R.L., R. Timmermans, S. Calabretta-Jongen, H. Eskes, A. Segers, D. Swart, and M. Schaap, 2012:
836 Improving ozone forecasts over Europe by synergistic use of the LOTOS-EUROS chemical transport
837 model and in-situ measurements. Atmos. Env., 60, 217-226, doi:10.1016/j.atmosenv.2012.06.017.

838 Cuvelier, C., P. Thunis, R. Vautard, M. Amann, B. Bessagnet M. Bedogni, R. Berkowicz, J. Brandt, F.
839 Brocheton, P. Builtjes, A. Coppalle, B. Denby, G. Douros, A. Graf, O. Hellmuth, C. Honoré, A. Hodzic,
840 J. Jonson, A. Kerschbaumer, F. de Leeuw, E. Minguzzi, N. Moussiopoulos, C. Pertot, G. Pirovano, L.
841 Rouil, M. Schaap, R. Stern, L. Tarrason, E. Vignati, M. Volta, L. White, P. Wind and A. Zuber, 2007:
842 CityDelta: A model intercomparison study to explore the impact of emission reductions in European
843 cities in 2010, Atmos. Env., 41, 189-207, doi: 10.1016/j.atmosenv.2006.07.036.

844 Dufour, A., M. Amodei, G. Ancellet, and V. H. Peuch, 2004: Observed and modeled “chemical weather”
845 during ESCOMPTE. Atmos. Res., 74, 161–189.

846 Edwards, D. P., L. K. Emmons, J. C. Gille, A. Chu, J.-L. Attié, L. Giglio, S. W. Wood, J. Haywood, M. N.
847 Deeter, S. T. Massie, D. C. Ziskin, and J. R. Drummond (2006), Satellite Observed Pollution From
848 Southern Hemisphere Biomass Burning, J. Geophys. Res., 111, D 14312, doi:10.1029/2005JD006655.

849 Edwards, D. P., L. K. Emmons, D. A. Hauglustaine, A. Chu, J. C. Gille, Y. J. Kaufman, G. Pétron, L. N.
850 Yurganov, L. Giglio, M. N. Deeter, V. Yudin, D. C. Ziskin, J. Warner, J.-F. Lamarque, G. L. Francis, S. P.
851 Ho, D. Mao, J. Chan, and J. R. Drummond (2004), Observations of Carbon Monoxide and Aerosol

852 From the Terra Satellite: Northern Hemisphere Variability, *J. Geophys. Res.*, 109, D24202,
853 doi:10.1029/2004JD0047272004

854 Edwards, D.P., A.F. Arellano Jr., and M.N. Deeter, 2009: A satellite observation system simulation
855 experiment for carbon monoxide in the lowermost troposphere. *J. Geophys. Res.*, 114, D14304, doi:
856 10.1029/2008JD011375.

857 El Amraoui, L., V.-H. Peuch, P. Ricaud, S. Massart, N. Semane, H. Teysse re, D. Cariolle, and F. Karcher,
858 2008a: Ozone loss in the 2002/03 Arctic vortex deduced from the Assimilation of Odin/SMR O₃ and
859 N₂O measurements: N₂O as a dynamical tracer. *Q. J. R. Meteorol. Soc.*, 134, 217–228.

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

863 Elbern, H., A. Strunk, and L. Nieradzki, 2010: “Inverse modelling and combined state-source estimation for
864 chemical weather,” in *Data Assimilation: Making Sense of Observations*, eds W.A. Lahoz, B. Khattatov,
865 and R. M nard (Berlin: Springer), 491–513.

866 Emili, E., B. Barret, S. Massart, E. Le Flochmoen, A. Piacentini, L. El Amraoui, O. Pannekoucke, and D.
867 Cariolle, 2014: Combined assimilation of IASI and MLS observations to constrain tropospheric and
868 stratospheric ozone in a global chemical transport model. *Atmos. Chem. Phys.*, 14, 177–198, doi:
869 0.5194/acp-14-177-2014.

870 Fitzmaurice, J., and R.L. Bras, 2008: Comparing Reanalyses Using Analysis Increment Statistics. *J.*
871 *Hydrometeor.*, 9, 1535–1545.

872 [Fu, D., Bowman, K. W., Worden, H. M., Natraj, V., Worden, J. R., Yu, S., Veefkind, P., Aben, I., Landgraf,](#)
873 [L., Strow, L., and Han, Y., 2016: High-resolution tropospheric carbon monoxide profiles retrieved from](#)
874 [CrIS and TROPOMI, *Atmos. Meas. Tech.*, 9, 2567-2579, 2016, doi:10.5194/amt-9-2567-2016](#)

875 Galli, A., A. Butz, R.A. Scheepmaker, O. Hasekamp, J. Landgraf, P. Tol, D. Wunch, N. M. Deutscher, G.C.
876 Toon, P.O. Wennberg, D.W.T. Griffith, and I. Aben, 2012: CH₄, CO, and H₂O spectroscopy for the
877 Sentinel-5 Precursor mission: an assessment with the Total Carbon Column Observing Network
878 measurements. *Atmos. Meas. Tech.*, 5, 1387-1398.

879 [George M., Clerbaux C., Bouarar I., Coheur P.-F., Deeter M. N., Edwards D. P., Francis G., Gille J. C., Hadji-](#)
880 [Lazaro J., Hurtmans D., Inness A. et al. 2015: An examination of the long-term CO records from](#)
881 [MOPITT and IASI: comparison of retrieval methodology. Atmos. Meas. Tech., 8 \(10\), 4313-4328.](#)

882 Gloudemans, A. M. S., Schrijver, H., Hasekamp, O. P., and Aben, I.: Error analysis for CO and CH₄ total
883 column retrievals from SCIAMACHY 2.3 μm spectra, Atmos. Chem. Phys., 8, 3999-4017,
884 doi:10.5194/acp-8-3999-2008, 2008.

885 de Haan, J.F., DISAMAR Algorithms and background, RP-TROPOMI-KNMI-066, KNMI, January 2012.

886 Hass, H., M. van Loon, C. Kessler, R. Stern, J. Matthijsen, F. Sauter, Z. Zlatev, J. Langner, V. Foltescu and
887 M. Schaap, 2003: Aerosol modelling: Results and Intercomparison from European Regional scale
888 modelling systems, Special Rep. EUROTRAC-2 ISS, Munich, 2003.

889 HTAP, 2007: Hemispheric Transport of Air Pollution 2007, Air Pollution Studies No. 16. UN Publication,
890 ECE/EB.AIR/94, Geneva.

891 Huijnen, V., H.J. Eskes, A. Poupkou, H. Elbern, K.F. Boersma, G. Foret, M. Sofiev, M., et al., 2010:
892 Comparison of OMI NO₂ tropospheric columns with an ensemble of global and European regional air
893 quality models. Atmos. Chem. Phys., 10, 3273–3296, doi: 10.5194/acp-10-3273-2010.

894 Jacob, D.J., 2000: Heterogeneous chemistry and tropospheric ozone. Atmos. Env., 34, 2131–2159.

895 Kaiser, J.W., A. Heil, M.O. Andreae, A. Benedetti, N. Chubarova, L. Jones, J.-J. Morcrette, M. Razinger,
896 M.G. Schultz, M. Suttie, and G.R. van der Werf, G. R., 2012: Biomass burning emissions estimated with
897 a global fire assimilation system based on observed fire radiative power. Biogeosciences, 9, 527-554,
898 doi: 10.5194/bg-9-527-2012.

899 Kuenen, J.J.P., A.J.H. Visschedijk, M. Jozwicka, and H.A.C. Denier van der Gon, 2014: TNO-MACC_II
900 emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory
901 for air quality modelling. Atmos. Chem. Phys., 14, 10963-10976, doi:10.5194/acp-14-10963-2014,
902 2014.

903 Kujanpää, Jukka, Albert Oude Nijhuis, Henk Eskes, Johan de Haan, Pepijn Veeffkind, Johanna Tamminen,
904 Synthetic Observation Product Specification (SOPS), Report of the ESA project "Impact of Spaceborne
905 Observations on Tropospheric Composition Analysis and Forecast" (ISOTROP), 12 August 2015.

906 Lagarde, T., A. Piacentini, and O. Thual, 2001: A new representation of data assimilation methods: the
907 PALM flow charting approach. Q. J. R. Meteorol. Soc., 127, 189–207.

908 Lahoz, W.A., R. Brugge, D.R. Jackson, S. Migliorini, R. Swinbank, D. Lary, et al., 2005: An observing
909 system simulation experiment to evaluate the scientific merit of wind and ozone measurements from the
910 future SWIFT instrument. *Q. J. R. Meteorol. Soc.*, 131, 503–523. doi:10.1256/qj.03.109.

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

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

917 Lahoz, W.A., V.-H. Peuch, J. Orphal, J.-L. Attié, K. Chance, X. Liu, et al., 2012: Monitoring air quality from
918 space: the case for the geostationary platform. *Bull. Am. Meteorol. Soc.*, 93, 221–233. doi:
919 10.1175/BAMS-D-11- 00045.1.

920 Lahoz, W.A., and P. Schneider, 2014: Data assimilation: making sense of earth observation. *Front. Environ.*
921 *Sci.*, 2, 16. <http://dx.doi.org/10.3389/fenvs.2014.00016>.

922 [Landgraf et al.: “Carbon monoxide total column retrievals from TROPOMI shortwave infrared](#)
923 [measurements”, *Atmos. Meas. Tech. Discuss.*, doi:10.5194/amt-2016-114](#).

924 Lee, J.D., A.C. Lewis, P.S. Monks, M. Jacob, J.F. Hamilton, J.R. Hopkins, N.M. Watson, J.E. Saxton, C.
925 Ennis, L.J. Carpenter, N. Carslaw, Z. Fleming, B.J. Bandy, D.E. Oram, S.A. Penkett, J. Slemr, E.
926 Norton, A.R. Rickard, L.K. Whalley, D.E. Heard, W.J. Bloss, T. Gravestock, S.C. Smit, J. Stanton, M.J.
927 Pilling, and M.E. Jenkin, 2006: Ozone photochemistry and elevated isoprene during the UK heatwave
928 of August 2003. *Atmos. Env.*, 40, 7598–7613.

929 Lefèvre, F., G.P. Brasseur, I. Folkins, A.K. Smith, and P. Simon, 1994: Chemistry of the 1991–1992
930 stratospheric winter: three dimensional model simulations. *J. Geophys. Res.*, 99, 8183-8195.

931 Levelt, P., et al., 2009: Observation Techniques and Mission Concepts for Atmospheric Chemistry
932 (CAMELOT), ESA Study, Contract no. 20533/07/NL/HE.

933 Lord, S.J., E. Kalnay, R. Daley, G.D. Emmitt, and R. Atlas, 1997: “Using OSSEs in the design of the future
934 generation of integrated observing systems, 1st Symposium on Integrated Observing Systems (Long
935 Beach, CA: American Meteorological Society).

Mis en forme : Anglais (Royaume-Uni)

Mis en forme : Anglais (Royaume-Uni)

Mis en forme : Anglais (Royaume-Uni)

936 Manders, A.M.M., M. Schaap, and R. Hoogerbrugge, 2009: Testing the capability of the chemistry
937 transport model LOTOS-EUROS to forecast PM10 levels in The Netherlands. *Atmos. Env.*, 4050-459
938 doi:10.1016/j.atmosenv.2009.05.006.

939 Marécal, V., V.-H. Peuch, C. Andersson, S. Andersson, J. Arteta, M. Beekmann, A. Benedictow, R.
940 Bergström, B. Bessagnet, A. Cansado, F. Chéroux, A. Colette, A. Coman, R.L. Curier, H.A.C. Denier
941 van der Gon, A. Drouin, H. Elbern, E. Emili, R.J. Engelen, H.J. Eskes, G. Foret, E. Friese, M. Gauss, C.
942 Giannaros, J. Guth, M. Joly, E. Jaumouillé, B. Josse, N. Kadygrov, J.W. Kaiser, K. Krajsek, J. Kuenen,
943 U. Kumar, N. Liora, E. Lopez, L. Malherbe, I. Martinez, D. Melas, F. Meleux, L. Menut, P. Moinat, T.
944 Morales, J. Parmentier, A. Piacentini, M. Plu, A. Poupkou, S. Queguiner, L. Robertson, L. Rouïl, M.
945 Schaap, A. Segers, M. Sofiev, M. Thomas, R. Timmermans, Á. Valdebenito, P. van Velthoven, R. van
946 Versendaal, J. Vira, and A. Ung, 2015: A regional air quality forecasting system over Europe: the
947 MACC-II daily ensemble production, *Geosci. Model Dev.*, 8, 2777-2813, doi:10.5194/gmd-8-2777-
948 2015

949 Massart, S., C. Clerbaux, D. Cariolle, A. Piacentini, S. Turquety, and J. Hadji-Lazaro, 2009: First steps
950 towards the assimilation of IASI ozone data into the MOCAGE-PALM system. *Atmos. Chem. Phys.*, 9,
951 5073–5091. doi: 10.5194/acp-9-5073-2009.

952 Masutani, M., T.W. Schlatter, R. M. Errico, A. Stoffelen, E. Andersson, W. Lahoz, J.S. Woollen, G.D.
953 Emmitt, L.-P. Riishøjgaard, and S. J. Lord, 2010a: Observing system simulation experiments. *Data*
954 *Assimilation: Making Sense of Observations*, W. A. Lahoz, B. Khattatov and R. Ménard, Eds., Springer,
955 647-679.

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

959 Nitta, T., 1975: Some analyses of observing systems simulation experiments in relation to First GARP
960 Global Experiment, in GARP Working Group on Numerical Experimentation, Report No. 10, US GARP
961 Plan (Washington, DC), 1–35.

962 Ordoñez, C., N. Elguindi, O. Stein, V. Huijnen, J. Flemming, A. Inness, H. Flentje, E. Katragkou, P. Moinat,
963 V.-H. Peuch, A. Segers, V. Thouret, G. Athier, M. van Weele, C. S. Zerefos, J.-P. Cammas, and M.G.

964 Schultz, 2010: Global model simulations of air pollution during the 2003 European heat wave. *Atmos.*
965 *Chem. Phys.*, 10, 789-815.

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

969 Rodgers, C. D., *Inverse methods for atmospheric sounding: Theory and Practice, Series on Atmospheric,*
970 *Oceanic and Planetary Physics–Vol. 2.*, Singapore, World Scientific, 2000.

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

974 Semane, N., V.-H. Peuch, L. El Amraoui, H. Bencherif, S. Massart, D. Cariolle, J.-L. Attié, and R. Abida,
975 2007: An observed and analysed stratospheric ozone intrusion over the high Canadian Arctic UTLS
976 region during the summer of 2003, *Q. J. R. Meteorol. Soc.*, 133, 171–178, doi: 10.1002/qj.141.

977 Solberg, S., Ø Hov, A. Søvde, I.S.A. Isaksen, P. Coddeville, H. De Backer, C. Forster, Y. Orsolini, and K.
978 Uhse, 2008: European surface ozone in the extreme summer 2003, *J. Geophys. Res.*, 113, D07307, doi:
979 10.1029/2007JD009098.

980 Stockwell, W.R., F. Kirchner, M. Kuhn, and S. Seefeld, 1997: A new mechanism for regional atmospheric
981 chemistry modeling. *J. Geophys. Res.*, 102, 25847-25879, doi: 10.1029/97JD00849.

982 Stoffelen, A., G.J. Marseille, F. Bouttier, D. Vasiljevic, S. DeHaan, and C. Cardinali, 2006: ADM-Aeolus
983 Doppler wind lidar observing system simulation experiment. *Q. J. R. Meteorol. Soc.*, 619, 1927–1948.
984 doi:10.1256/qj.05.83.

985 Stern, R., P. Builtjes, M. Schaap, R. Timmermans, R. Vautard, A. Hodzic, M. Memmesheimer, H. Feldmann,
986 E. Renner, R. Wolke, and A. Kerschbaumer, 2008: A model inter-comparison study focussing on
987 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)
988 10.1016/j.atmosenv.2008.01.068.

989 Streets D. G., 2013: Emissions estimation from satellite retrievals: A review of current capability
990 *Atmospheric Environment*, 77 1011-1042

991 Tan, D.G.H., E. Andersson, M. Fisher, and L. Isaksen, 2007: Observing system impact assessment using a
992 data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission. *Q. J. R.*
993 *Meteorol. Soc.*, 133, 381–390. doi: 10.1002/qj.43.

994 Tangborn, A., I. Štajner, M. Buchwitz, I. Khlystova, S. Pawson, R. Hudman, et al., 2009: Assimilation of
995 SCIAMACHY total column CO observations: global and regional analysis of data impact. *J. Geophys.*
996 *Res.*, 114. doi: 10.1029/2008JD010781.

997 Teyssèdre, H., M. Michou, H.L. Clark, B. Josse, F. Karcher, D. Olivié, V.-H. Peuch, D. Saint-Martin, D.
998 Cariolle, J.-L. Attié, P. Nédélec, P. Ricaud, V. Thouret, A.R.J. van der A., A. Volz-Thomas, and F.
999 Chéroux, F., 2007: A new tropospheric and stratospheric Chemistry and Transport Model MOCAGE-
1000 Climat for multi-year studies: evaluation of the present-day climatology and sensitivity to surface
1001 processes. *Atmos. Chem. Phys.*, 7, 5815–5860, <http://www.atmos-chem-phys.net/7/5815/2007/>.

1002 Timmermans, R.M.A., M. Schaap, H. Elbern, R. Siddans, S. Tjemkes, R. Vautard, et al., 2009a: An
1003 Observing System Simulation Experiment (OSSE) for Aerosol Optical Depth from Satellites. *J. Atmos.*
1004 *Ocean Tech.*, 26, 2673–2682. doi: 10.1175/2009JTECHA1263.1.

1005 Timmermans, R.M.A., A.J. Segers, P.J.H. Builtjes, R. Vautard, R. Siddans, H. Elbern, et al., 2009b: The
1006 added value of a proposed satellite imager for ground level particulate matter analyses and forecasts.
1007 *IEEE J. Sel. Top. Appl. 2*, 271–283. doi:10.1109/JSTARS.2009.2034613.

1008 Timmermans, R., W.A. Lahoz, J.-L. Attié, V.-H. Peuch, L. Curier, D. Edwards, H. Eskes, and P. Builtjes,
1009 2015: Observing System Simulation Experiments for Air Quality. *Atmos. Env.*, 115, 199-213,
1010 doi:10.1016/j.atmosenv.2015.05.032.

1011 Tressol, M., C. Ordoñez, R. Zbinden, J. Brioude, V. Thouret, C. Mari, P. Nédélec, J.-P. Cammas, H. Smit, H.-
1012 W. Patz, and A. Volz-Thomas, 2008: Air pollution during the 2003 European heat wave as seen by
1013 MOZAIC airliners. *Atmos. Chem. Phys.*, 8, 2133-2150.

1014 Vautard, R., C. Honoré, M. Beekmann, and L. Rouil, 2005: Simulation of ozone during the August 2003 heat
1015 wave and emission control scenarios. *Atmos. Env.*, 39, 2957–2967.

1016 van Loon, M., R. Vautard, M. Schaap, R. Bergstrom, B. Bessagnet, J. Brandt, P.J.H. Builtjes, J.H.
1017 Christensen, C. Cuvelier, A. Graff, J.E. Jonson, M. Krol, J. Langner, P. Roberts, L. Rouil, R. Stern, L.
1018 Tarrason, P. Thunis, E. Vignati, and L. White, 2007: Evaluation of long-term ozone simulations from

1019 seven regional air quality models and their ensemble, *Atmos. Env.*, 41, 2083–2097. doi:
1020 10.1016/j.atmosenv.2006.10.073.

1021 Veefkind, J.P., I. Aben, K. McMullan, H. Förster, J. de Vries, G. Otter, J. Claas, H.J. Eskes, J.F. de Haan, Q.
1022 Kleipool, M. van Weele, O. Hasekamp, R. Hoogeveen, J. Landgraf, R. Snel, P. Tol, P. Ingmann, R.
1023 Voors, B. Kruizinga, R. Vink, H. Visser, and P.F. Levelt, 2012: TROPOMI on the ESA Sentinel-5
1024 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air
1025 quality and ozone layer applications. *Remote Sens. Environ.*, 120, 70-83.

1026 [Vidot et al., 2011: Carbon monoxide from shortwave infrared reflectance measurements: A new retrieval
1027 approach for clear-sky and partially cloudy atmospheres, *Remote Sens. Environ.*,
1028 doi:10.1016/j.rse.2011.09.032.](#)

1029 [Warner J., F. Carminati, Z. Wei, W. Lahoz and J.-L. Attié, 2013: Tropospheric carbon monoxide variability
1030 from AIRS under clear and cloudy conditions. *Atmos. Chem. Phys.*, 13, 12469-12479. doi:10.5194/acp-
1031 13-12469-2013.](#)

1032 Weaver, A., and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion
1033 equation. *Q. J. R. Meteorol. Soc.*, 127, 1815–1846.

1034 Willmott, C., and K. Matsuura, 2005: Advantages of the Mean Absolute Error (MAE) over the Root Mean
1035 Square Error (RMSE) in assessing average model performance. *Clim. Res.*, 30, 79-82.

1036 Willmott, C., K. Matsuura, and S.M. Robeson, 2009: Ambiguities inherent in sums-of-squares-based error
1037 statistics. *Atmos. Env.*, 43, 749-752.

1038 Worden, H. M., M. N. Deeter, Christian Frankenberg, Maya George, Florian Nichitiu, John Worden,
1039 Ilse Aben et al. "Decadal record of satellite carbon monoxide observations." *Atmospheric Chemistry
1040 and Physics* 13, no. 2 (2013): 837-850.

1041 Yumimoto, K., 2013: Impacts of geostationary satellite measurements on CO forecasting: an observing
1042 system simulation experiment with GEOS- Chem/LETKF data assimilation system. *Atmos. Env.*, 74,
1043 123–133. doi: 10.1016/j.atmosenv.2013.03.032.

1044 Zoogman, P., D.J. Jacob, K. Chance, L. Zhang, P. Le Sager, A.M. Fiore, A. Eldering, X. Liu, V. Natraj, and
1045 S.S. Kulawik, 2011. Ozone air quality measurement requirements for a geostationary satellite mission.
1046 *Atmos. Env.*, 45, 7143-7150.

Mis en forme : Anglais (Royaume-Uni)

1047 Zoogman, P., D.J. Jacob, K. Chance, X. Liu, M. Lin, A.M. Fiore, and K. Travis, 2014a. Monitoring high-
1048 ozone events in the US Intermountain West using TEMPO geostationary satellite observations. Atmos.
1049 Chem. Phys., 14, 6261-6271. <http://dx.doi.org/10.5194/acp-14-6261-2014>.

1050 Zoogman, P., D.J. Jacob, K. Chance, H.M. Worden, D.P. Edwards, and L. Zhang, 2014b: Improved
1051 monitoring of surface ozone by joint assimilation of geostationary satellite observations of ozone and
1052 CO. Atmos. Env., 84, 254-261. <http://dx.doi.org/10.1016/j.atmosenv.2013.11.048>.

1053

1054

1055

1056

1057

1058

1059

1060

1061

1062 **Tables**

1063

1064 **Table 1:** Spectral and radiometric settings for DISAMAR, and the look-up table node [points](#).

Spectral and radiometric settings	
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
Node points	
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

1065

1066

1067

1068

1069

1070

1071 **Table 2:** Description of the configuration used in the assimilation system

	Description
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

1072

1073

1074

1075

1076

1077

1078

1079

1080

1081

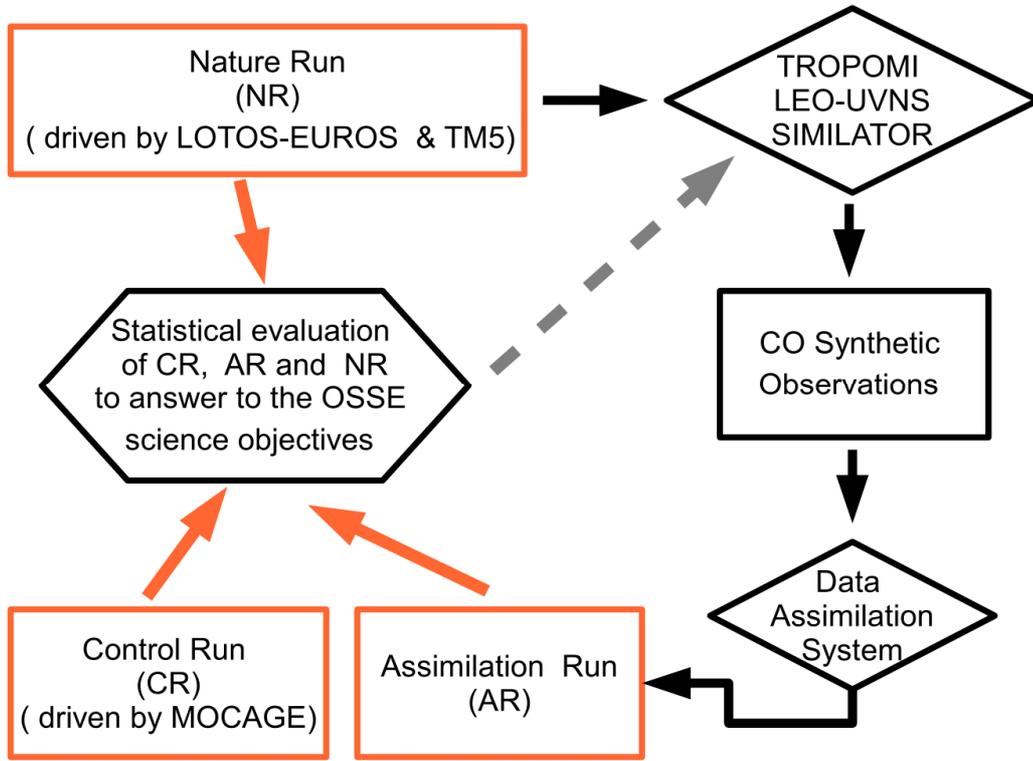
1082

1083

1084

1085 | [Figures](#)

1086



1087

1088 | [Figure 1: Diagram of the Observing System Simulation Experiments \(OSSE\) components](#)

1089

1090 **Figures**

1091

1092

1093

1094

1095

1096

1097

1098

1099

1100

1101

1102

1103

1104

1105

1106

1107

1108

1109

1110

1111

1112

1113

1114

1115

1116

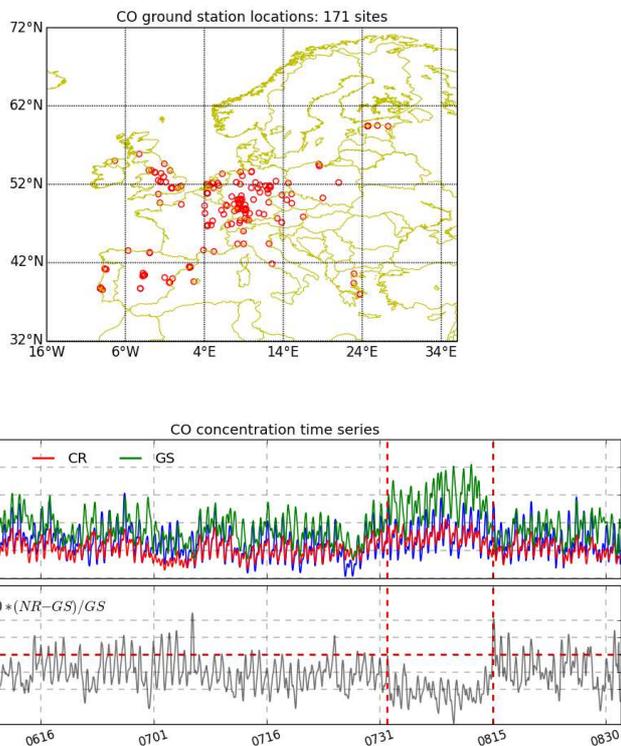


Figure 24: Top panel: location of selected ground-based stations for CO measurements taken from the Airbase database during northern summer 2003 (1 June – 31 August). There are 171 sites with locations shown by circles. The labels show longitude, degrees (x-axis) by latitude, degrees (y-axis). Middle panel: simulated and measured time-series of CO concentrations in surface air from nature run (blue line), the control run (red line) and from the selected 171 Airbase sites (green line). We form the CO time-series for the ground-based stations by averaging the hourly data representative of the 171 sites. The labels show time in MMDD format (x-axis) by CO concentration, parts per billion by volume, ppbv (y-axis). Bottom panel: The grey curve shows the relative error of the nature run with respect to the observations, defined as NR value minus ground station value divided by the ground station value and multiplied by 100. The labels show time in MMDD format (x-axis) by relative error, percent (y-axis). The vertical red dashed lines in the middle and bottom panels delineate the 2003 European heat wave period (31 July – 15 August).

1117

1118

1119

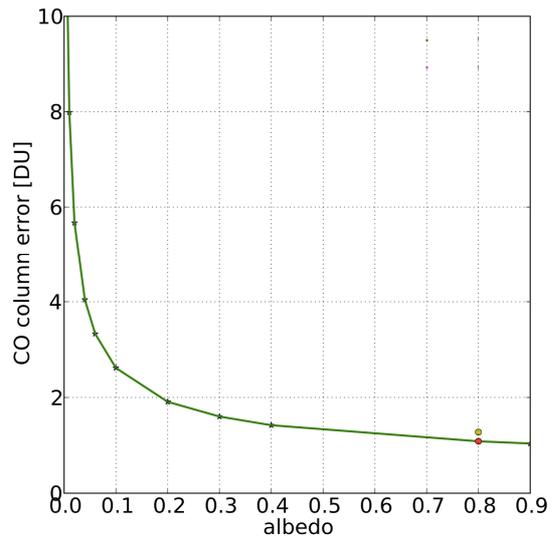


Figure 32: Dependence of the CO column uncertainty (Dobson Unit) on the surface albedo. Simulation settings are: solar zenith angle 53 degrees, viewing zenith angle 26 degrees, relative azimuth angle 0 degree, cloud/surface pressure 1100 hPa.

1120

1121

1122

1123

1124

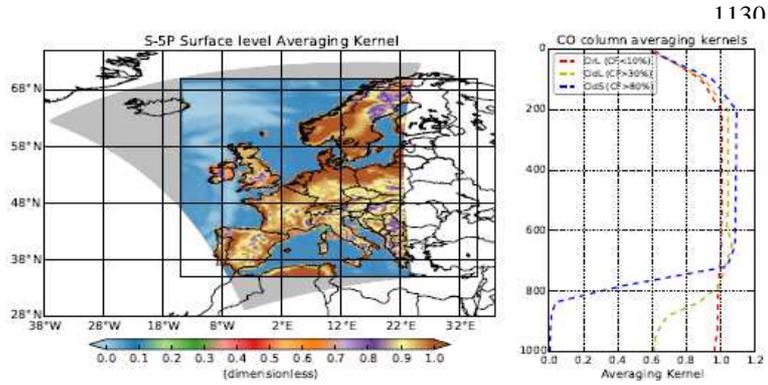
1125

1126

1127

1128

1129



1138

1139 **Figure 43:** Left panel: S-5P CO averaging kernel values at the surface. Labels are longitude, degrees (x-axis)
1140 by latitude, degrees (y-axis). Right panel: Averaging kernels for land pixels with cloud fraction less than 10%
1141 (dashed red lines); for land pixels with cloud fraction greater than 30% (dashed yellow lines); and for sea
1142 pixels with cloud fraction greater than 80% (dashed blue lines). The averaging kernels are for an average of
1143 the data shown on the swath for 1 June 2003 at 12:34 UTC. Labels are averaging kernel, normalized (x-axis)
1144 by pressure level, hPa (y-axis).

1145

1146

1147

1148

1149

1150

1151

1152

1153

1154

1155

1156

1157

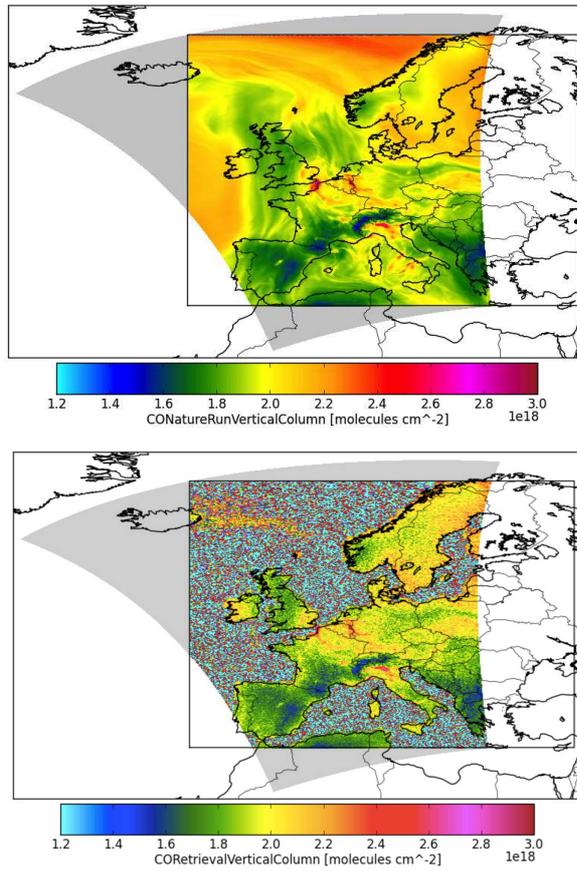


Figure 54: Top: Nature run collocated to the synthetic S-5P observations for the 12:34 orbit on 1 June 2003.

Bottom: corresponding synthetic observations.

1158

1159

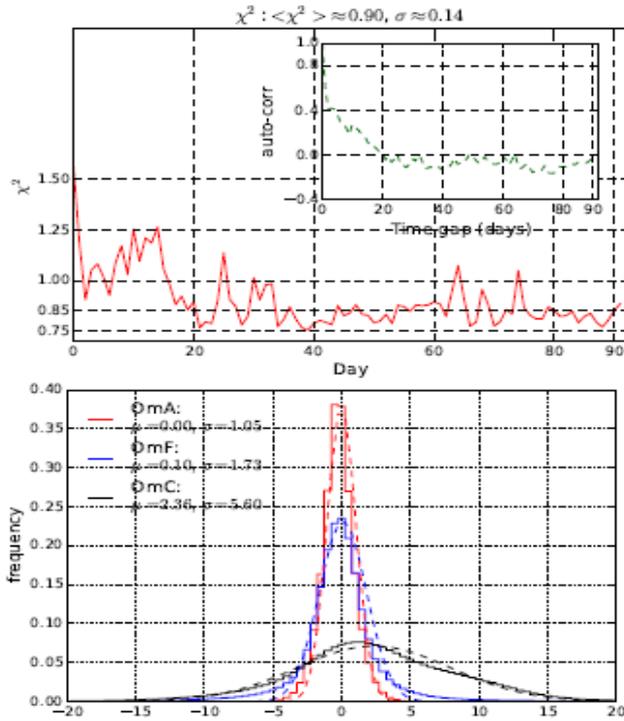
1160

1161

1162

1163

1164



1179 **Figure 6S:** Self-consistency tests. Top panel: time-series (red line) of χ^2 for OmF and its associated auto-
 1180 correlation signal (green line). For the χ^2 diagnostic we normalize the OmF differences by the background
 1181 error. The labels show time, days (x-axis) and χ^2 value (y-axis) for the χ^2 plot, and time gap, days (x-axis) and
 1182 auto-correlation (y-axis) for the auto-correlation plot. Bottom panel: histograms of Observations minus
 1183 Analysis (OmA -red solid line), Observations minus Forecast (OmF -blue solid line), and Observations
 1184 minus Control run (OmC -black solid line). We normalize these differences by the observation error. The
 1185 dashed lines correspond to the Gaussian fits of the different histograms. The labels show the OmA, OmF or
 1186 OmC differences (x-axis) and the frequency of occurrence of the differences (y-axis). We calculate the
 1187 diagnostics OmA, OmF, and OmC over the period of 1 June – 31 August 2003.

1188

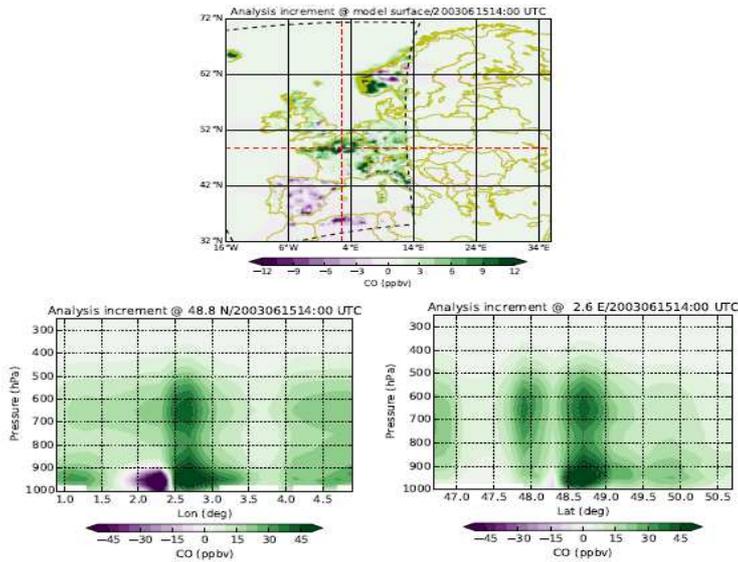
1189

1190

1191

1192

1193



1205 | **Figure 76:** S-5P CO analysis increments, units of ppbv, at 14:00 UTC on 15 June 2003: Top panel:
1206 | geographical distribution at the model surface. Red dashed lines show zonal and meridional vertical slices at
1207 | 48°8 N, and 2°6 E, respectively. The black dashed line shows the S-5P cross-track at 13:12 UTC, clipped to
1208 | fit the OSSE simulation domain. Note that we measure the S-5P CO observations at 13:12 UTC. The labels
1209 | show longitude, degrees (x-axis) and latitude, degrees (y-axis). Left and right bottom panels show,
1210 | respectively, the longitude-height and latitude-height cross-sections at a location near Paris. The labels for
1211 | the bottom panels show longitude, degrees (x-axis, left panel), latitude, degrees (x-axis, right panel), and
1212 | pressure, hPa (y-axis, both panels). Green/purple colours indicate positive/negative values in the increment
1213 | fields.

1214

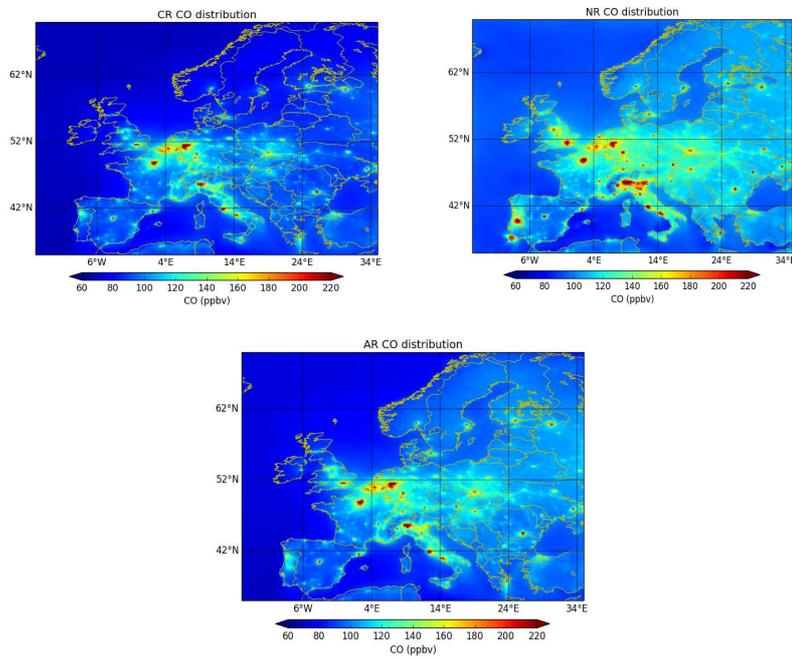
1215

1216

1217

1218

1219



1220

1221 **Figure 87:** Distribution of CO surface concentrations, units ppbv, averaged for the period 1 June – 31 August
1222 2003. Top left panel: the control run (CR) from MOCAGE; right top panel: the nature run (NR) from
1223 LOTOS-EUROS; bottom panel: the assimilation run (AR) from MOCAGE obtained after assimilating the S-
1224 5P CO total column simulated data sampled from the NR. In all panels, the labels show longitude, degrees
1225 (x-axis) and latitude, degrees (y-axis). Red/blue colours indicate relatively high/low values of the CO surface
1226 concentrations.

1227

1228

1229

1230

1231

1232

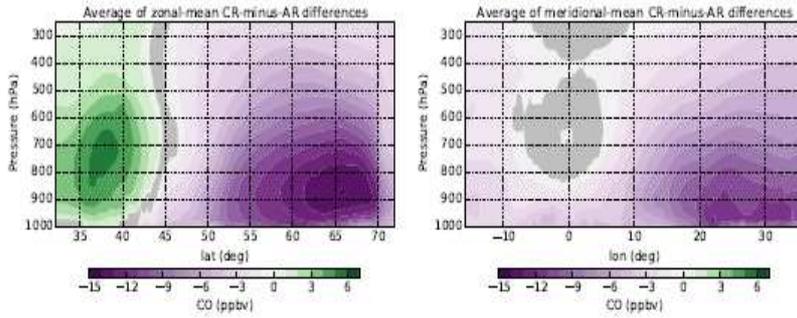
1233

1234

1235

1236

1237



1244

1245 **Figure 98:** Zonal (left panel) and meridional (right panel) slices of the difference between the CR and AR
1246 CO fields, units of ppbv, averaged over the summer period (1 June – 31 August 2003). The areas highlighted
1247 in grey colour indicate where the AR is not significantly different to the CR at the 99% confidence level. The
1248 labels in the left panel are latitude, degrees (x-axis) and pressure, hPa (y-axis). The labels in the right panel
1249 are longitude, degrees (x-axis) and pressure, hPa (y-axis). Green/purple colours indicate positive/negative
1250 values in the difference fields.

1251

1252

1253

1254

1255

1256

1257

1258

1259

1260

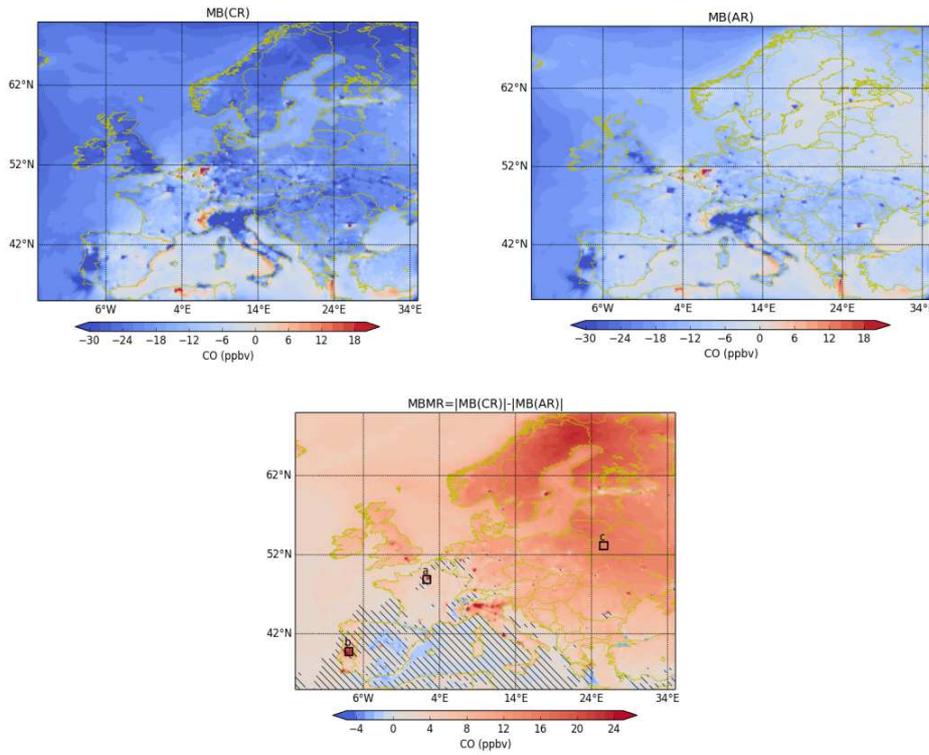
1261

1262

1263

1264

1265



1266

1267 **Figure 109:** Mean bias reduction at the surface for CO, units of ppbv: Left top panel shows the CR mean
1268 bias with respect to the NR (CR-NR). Right top panel shows the AR mean bias with respect to the NR (AR-
1269 NR). Bottom panel shows the mean bias magnitude reduction (absolute value of the mean bias for CR minus
1270 the absolute value of the mean bias for AR). We average the data over northern summer 2003 (1 June – 31
1271 August). The labels show longitude, degrees (x-axis) and latitude, degrees (y-axis). The hatched area in the
1272 bottom panel shows where the mean bias plotted in this panel (MBMR) is not statistically significant at the
1273 99% confidence level. The three squares in the bottom panel represent locations for the three time-series
1274 shown in Fig. 1342. Red/blue colours indicate positive/negative values in the MB/MBMR.

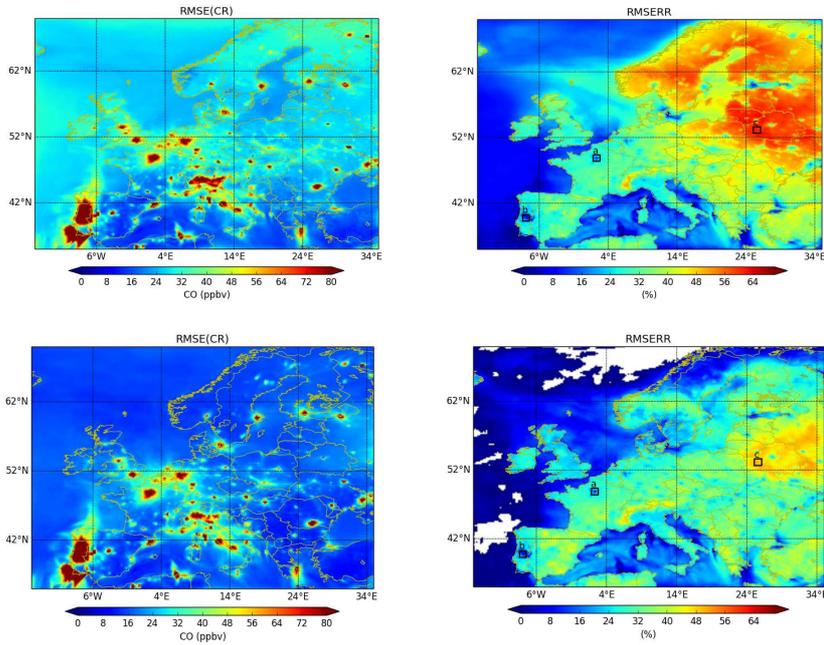
1275

1276

1277

1278

1279



1280

1281

1282

1283 **Figure 1140:** Top: Root Mean Square Error (RMSE), units of ppbv, between CR and NR (left panel), and its
1284 corresponding reduction rate RMSERR, in % (right panel) keeping the systematic error. Bottom: Same as top
1285 panel but calculating the RMSE after removing the systematic error. The labels on each panel are longitude,
1286 degrees (x-axis) and latitude, degrees (y-axis). The three squares in the two right panels represent locations
1287 for the three time-series shown in Fig. 1342. Red/blue colours indicate relatively high/low values in the
1288 RMSE/RMSERR.

1289

1290

1291

1292

1293

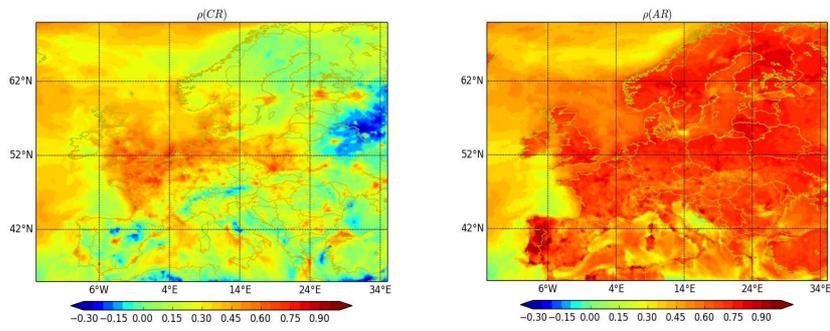
1294

1295

1296

1297

1298



1299

1300 **Figure 12.11:** Correlation coefficient between the CR and the NR (left panel) and the AR and the NR (right
1301 panel) at the surface and for the northern summer period (1 June – 31 August). The labels are longitude,
1302 degrees (x-axis) and latitude, degrees (y-axis). Red/blue colours indicate positive/negative values of the
1303 correlation coefficient.

1304

1305

1306

1307

1308

1309

1310

1311

1312

1313

1314

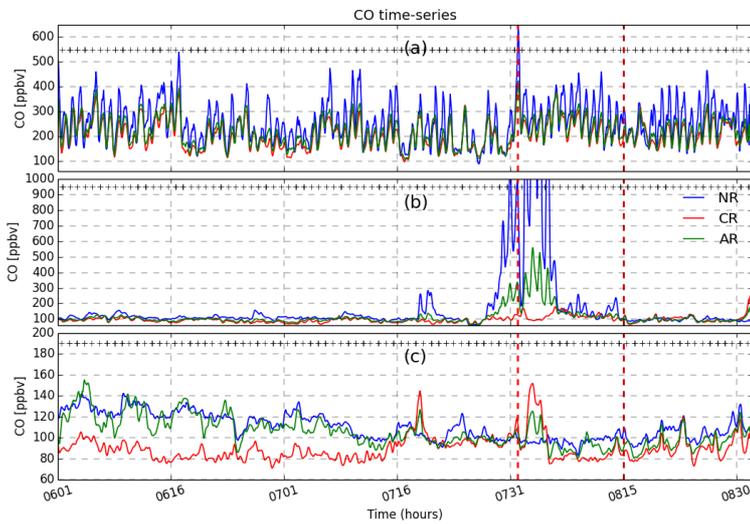
1315

1316

1317

1318

1319



1320

1321 **Figure 1312:** Time-series for CO surface concentrations (1 June – 31 August) from NR (blue colour), CR
1322 (red colour) and AR (green colour) over three different locations represented by squares in Figs. 109 and
1323 1140. Top panel: area near Paris; middle panel: area over Portugal, where forest fires occurred; bottom panel:
1324 Eastern part of the study domain. The labels in the three panels are time, in format MMDD (x-axis) and CO
1325 concentration, ppbv (y-axis). The plus symbols at the top of each panel indicate availability of observations
1326 from the S-5P platform.

1327

1328

1329

1330

1331

1332

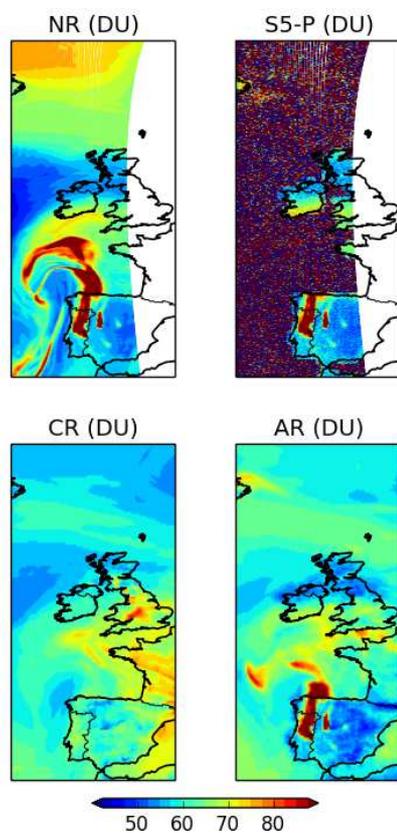
1333

1334

1335

1336

1337



1338

1339 **Figure 1413:** CO total column at 14:15 UTC on 4 August 2003, Dobson units, DU. Top left panel: NR; top
1340 right panel: simulated S-5P observations; bottom left panel: CR; bottom right panel: AR. Red/blue colours
1341 indicate relatively high/low values of the CO total column.

1342

1343

1344

1345

1346

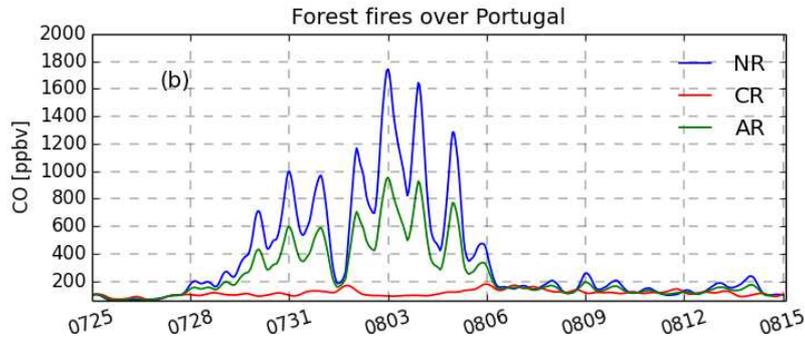
1347

1348

1349

1350

1351



1352

1353

1354 **Figure 1514:** Time-series for CO surface concentrations for the period covering the Portugal forest fires (25
1355 July – 15 August) from NR (blue colour), CR (red colour) and AR (green colour) over the location associated
1356 with the middle panel of Fig.1312. These data concern the second OSSE we perform to understand the
1357 behaviour of the original OSSE over the period of the forest fires (see text for more details). The labels are
1358 time, in format MMDD (x-axis) and CO concentration, ppbv (y-axis).

1359

1360

1361

1362

1363

1364

1365

1366