

A thermal infrared instrument onboard a geostationary platform for CO and O₃ measurements in the lowermost troposphere: Observing System Simulation Experiments

M. Claeysman^{1,2}, J.-L. Attié^{1,2}, V.-H. Peuch², L. El Amraoui², W. A. Lahoz^{2,3}, B. Josse², M. Joly², J. Barré², P. Ricaud¹, S. Massart⁴, A. Piacentini⁴, T. von Clarmann⁵, M. Höpfner⁵, J. Orphal⁵, J.-M. Flaud⁶, and D. P. Edwards⁷

¹Laboratoire d'Aérodynamique, Université de Toulouse, CNRS/INSU, Toulouse, France

²CNRM-GAME, Météo-France and CNRS URA 1357, Toulouse, France

³NILU, N-2027 Kjeller, Norway

⁴CERFACS, Toulouse, France

⁵Forschungszentrum Karlsruhe, Inst Meteorol & Klimaforsch, Karlsruhe, Germany

⁶Laboratoire Interuniversitaire des Systèmes Atmosphériques, CNRS UMR 7583, Université de Paris-Est, Créteil, France

⁷National Center for Atmospheric Research, Boulder, Colorado, USA

Correspondence to: M. Claeysman, Laboratoire d'aérodynamique 14 avenue Edouard Belin, 31400 Toulouse, France.
(marine.claeysman@aero.obs-mip.fr)

Abstract. This paper presents observing system simulation experiments (OSSEs) to compare the relative capabilities of two geostationary thermal infrared (TIR) instruments to monitor ozone (O₃) and carbon monoxide (CO) for air quality (AQ) purposes over Europe. The originality of this study is to use OSSEs to assess the impact of various key parameters (emissions, meteorology, initial condition and the 3 parameters together) on these infrared measurements. The first instrument (GEO-TIR) has a configuration optimized to monitor O₃ and CO in the lowermost troposphere (LmT; defined to be the atmosphere between the surface and 3 km), and the second instrument (GEO-TIR2) is designed to monitor temperature and humidity. Both instruments measure radiances in the same spectral TIR band. Results show that GEO-TIR could have a significant impact (GEO-TIR is closer to the reference atmosphere than GEO-TIR2) on the analyses of O₃ and CO LmT column. The value of the measurements for both instruments is mainly over the Mediterranean Basin and some impact can be found over the Atlantic Ocean and Northern Europe. The impact of GEO-TIR is mainly above 1 km for O₃ and CO but can also improve the surface analyses for CO. The analyses of GEO-TIR2 show low impact for O₃ LmT column but a significant impact (but lower than for GEO-TIR) for CO above 1 km. The results of this study indicate the beneficial impact from an infrared instrument (GEO-TIR) dedicated to monitoring O₃ and CO concentrations in the LmT, and the value of this information for constraining AQ models.

1 Introduction

The atmospheric composition of pollutants in the lowermost troposphere (LmT; defined to be the atmosphere between the surface and 3 km) is a societal issue because it is associated with air quality (AQ). Poor AQ can lead to negative health effects such as respiratory problems, heart disease and lung cancer. Monitoring and forecasting AQ is becoming routine (e.g., Prev'air in France, Honoré et al. (2008)). This concerns both gaseous and particle species and includes ground-level ozone (O_3), nitrogen oxides (NO_x) and suspended particulate matter (PM), all of which are identified as potential health hazards (Brunekreef and Holgate, 2002).

O_3 is a key trace gas in the troposphere that plays a significant role in atmospheric chemistry, air quality and radiative forcing (Jacob, 2000). It is a secondary pollutant produced by the photochemical oxidation of hydrocarbons and carbon monoxide (CO) in the presence of nitrogen oxides (NO_x). It is a precursor to the formation of the hydroxyl radical which impacts the oxidizing capacity of the atmosphere. It is also an irritant gas which can affect severely the respiratory tract and cause damage to vegetation (Seinfeld and Pandis, 1997). In Europe, tropospheric O_3 levels increased rapidly between 1970 and 1990 as a result of increases in precursor emissions (e.g., Lamarque et al., 2005); but this increase has slowed down or declined since 1990 (e.g., Oltmans et al., 2006). CO is a reactive gas which also plays an important role in tropospheric chemistry (Jacob, 2000). It is an O_3 precursor and a tracer of pollution (Turquety et al., 2009). In addition to atmospheric chemical sources, CO is also a primary pollutant, emitted during incomplete combustion processes, which makes CO a good tracer for urban/industrial fossil fuel burning (e.g., Branis, 2009), wildfires (e.g., Cristofanelli et al., 2009) and tropical biomass burning (e.g., Edwards et al., 2006).

In Europe, despite the definition and the implementation of regulations and laws regarding pollutants, AQ is still a real concern for the public and the authorities. Reduction of the AQ impact on health may be achieved both with long- and short-term actions (Menut and Bessagnet, 2010). Long-term actions concern global improvement of AQ by reducing anthropogenic emissions. Short-term actions consist in anticipating pollution events, a few days before they happen, to warn the public in advance in order to reduce exposure and help authorities take effective control measures. AQ monitoring and forecasting is required to achieve these actions.

Current monitoring and forecasting systems mostly rely on three-dimensional models (e.g., Vaughan et al., 2004; McKeen et al., 2005; Honoré et al., 2008; Hollingsworth et al., 2008). Traditionally, AQ monitoring has been done using measurements from ground-based stations. Ground-based in situ observations have the disadvantage of an inhomogeneous spatial coverage, and can present inconsistencies in their spatial representativeness, their measurement methods and correction factors (Ignaccolo et al., 2008). The main advantage of satellite observations is the large spatial coverage. Ground-based observations and satellite observations of pollutants complement each other; the former sample the surface, the latter sample in the vertical, typically as a column. For AQ purposes, satellite observations have to measure tropospheric composition at adequate spatial ($\sim 10 \times 10 \text{ km}^2$)

55 and temporal (~ 1 hour) resolution (Fishman et al., 2008; Martin, 2008). To complement in situ information (e.g., AQ networks, sondes, aircraft measurements), denser observations at continental scales in the lowermost troposphere (LmT; defined to be the atmosphere between the surface and 3 km) are needed for AQ relevant species (e.g., O_3 and CO). These observations can only be provided by a Geostationary Earth Orbit (GEO) platform (Bovensmann and Orphal, 2005; Edwards, 2006).

60 Several GEO missions have been proposed to monitor AQ. In the USA, the GEO-CAPE concept (National Research Council, 2007) dedicated to the measurement of tropospheric trace gases is planned toward the end of the decade. In Japan, a similar mission has been proposed by the Japan Society of Atmospheric Chemistry to monitor O_3 from GEO (Akimoto et al., 2008) and has been recently endorsed by the Japanese Space Agency (JAXA). In Europe, several GEO missions have
65 been proposed to monitor tropospheric constituents at high temporal and spatial resolution such as GeoTrobe (Burrows et al., 2004) and GeoFIS (Flaud et al., 2004; Orphal et al., 2005). The Meteosat Third Generation - Thermal Infrared Sounder (MTG-IRS) is a planned mission to be launched in 2017. MTG-IRS will be able to provide information on horizontally, vertically, and temporally resolved water vapour and temperature structures of the atmosphere. It will also provide O_3 and CO
70 measurements in the troposphere within the long-wave infrared and the mid-wave infrared bands, respectively.

The sentinel 4 UVN (ultraviolet-visible-near infrared) payload is also a planned mission and will be deployed on the two MTG - Sounder (MTG-S) satellites in GEO orbit over Europe; there are planned for launch in 2017 and 2024 and UVN is expected to provide measurements of O_3 and
75 nitrogen dioxide column, and aerosol optical depth. To complement Sentinel 4 UVN, the mission Monitoring the Atmosphere from Geostationary orbit for European Air Quality (MAGEAQ) has been proposed as a candidate for the Earth Explorer Opportunity Mission EE-8 call of the European Space Agency (Peuch et al., 2009, 2010). MAGEAQ is a multispectral instrument (thermal infrared and visible) designed to provide height-resolved measurements of O_3 and CO in the LmT.

80 A method to determine the beneficial impact of future instruments is the Observing System Simulation Experiment (OSSE) (Atlas, 1997). This method is widely used in the meteorological community for assessing the usefulness of new meteorological satellite data (e.g., Lahoz et al., 2005; Stoffelen et al., 2006; Masutani et al., 2010b). However, there are actually few studies concerning OSSEs on chemical species. Two recent OSSE studies have been conducted concerning a GEO plat-
85 form for AQ purposes. The first one consists of an OSSE for CO in the LmT using a multispectral (near-infrared and thermal infrared) instrument (Edwards et al., 2009). The second one concerns a satellite imager to monitor the aerosol optical depth to improve ground level particulate analyses and forecasts (Timmermans et al., 2009).

The aim of this paper is to present a new OSSE for a GEO instrument in the thermal infrared
90 band (called GEO-TIR) with instrument characteristics optimized to monitor O_3 and CO in the LmT. GEO-TIR presents instrument characteristics (signal to noise ratio: SNR and spectral sampling

interval: SSI) similar to the thermal infrared instrument proposed in the MAGEAQ mission and described in Claeysman et al. (2010b). In order to accurately assess the impact of GEO-TIR O_3 and CO observations in an AQ model, we perform several OSSEs to evaluate the sensitivity of the analyses to various key parameters: emissions, meteorology and initial conditions, and for all these parameters simultaneously. We also perform OSSEs for another GEO thermal infrared instrument but with instrument characteristics optimized for temperature and humidity (GEO-TIR2) to evaluate the relative added value of GEO-TIR with respect to GEO-TIR2. GEO-TIR2 has SNR and SSI similar to those of MTG-IRS (Clerbaux et al., 2008b). We first evaluate the added value over Europe of GEO-TIR in the LmT column considering several statistical measures (correlation, bias, standard deviation) and then, the vertical impact of GEO-TIR, considering several AQ statistical measures (e.g., good detection, false alarms, missing events).

This paper is organized as follow. In section 2, we describe the OSSE method, the chemistry transport model (CTM), the assimilation scheme used, the synthetic observations, the different experiments, and the statistical measures. In section 3, we discuss the added value of GEO-TIR in an AQ model in the LmT, by comparison with GEO-TIR2. Summary and conclusions are presented in section 4.

2 Methodology and experiment setup

2.1 The Observing System Simulation Experiment

Observing System Simulation Experiments (e.g., Atlas, 1997) are used to assess the impact of future observing systems. To simulate a future observing system, existing observations are generally replaced by synthetic observations, generated by sampling a nature run, according to the instrument characteristics (observational geometry, spatial and temporal resolution, errors). In some cases, a subset of the future observations can be represented by current observations, but the observing platform of interest is always simulated (see Masutani et al. (2010a) for further discussion). In this study, the nature run simulates the true state of the atmosphere and the synthetic observations are simulated through the nature run; no current observations are used. Synthetic observations are then assimilated in the control run of the OSSE. The OSSE discussed is composed of the following elements:

1. A nature run produced using a state-of-the-art model which represents the true atmosphere;
2. Synthetic observations which are sampled through the nature run corresponding to the instruments considered;
3. A control run, which yields an alternative representation of the atmosphere, different from the nature run. In this study, the control run is a free model run and includes no assimilated observations. The differences between the control run and the nature run should ideally be similar to the differences between a state-of-the-art model and the real atmosphere;

4. An assimilation run using synthetic observations from the instruments of interest generated from the nature run and the same model setup configuration as for the control run;
5. Assessment of the added value of the instruments of interest by statistical comparison between nature run, control run and assimilation run. In fact, the assessment is based on the differences between the nature run and control run, and between the nature run and assimilation run. If the difference between the assimilation and the nature run is significantly smaller than the difference between the control run and the nature run, we conclude that the instrument of interest has added value.

Note that in the OSSE described in this paper, the future observing system only comprises two GEO observing platforms and no other observations (e.g., ground stations). We think this is justified because at this stage we are only interested in providing a reasonably accurate first order estimate of the added value of the proposed observing platform. Furthermore, because of model uncertainties, we focus in a comparison of the relative performance of two instruments and not predicting the absolute performance of the two instruments. In a later work, we will extend this study to include a more complete representation of the future observing system, including the ground-based network, and refine our estimate of the added value of the proposed observing platform.

The different elements of the OSSE are described in more detail below.

2.2 The reference atmosphere

The MOCAGE (MODèle de Chimie Atmosphérique à Grande Echelle) model is used to simulate the nature run. MOCAGE is a three-dimensional CTM for the troposphere and stratosphere (Peuch et al., 1999) which simulates the interactions between the dynamical, physical and chemical processes. It uses a semi-Lagrangian advection scheme (Josse et al., 2004) to transport the chemical species. Its vertical resolution is 47 hybrid levels from the surface up to 5 hPa with a resolution of about 150 m in the lower troposphere increasing to 800 m in the higher troposphere. Turbulent diffusion is calculated with the scheme of Louis (1979) and convective processes with the scheme of Bechtold et al. (2001). The chemical scheme used in this study is RACMOBUS. It is a combination of the stratospheric scheme REPROBUS (Lefèvre et al., 1994) and the tropospheric scheme RACM (Stockwell et al., 1997). It includes 119 individual species with 89 prognostic variables and 372 chemical reactions.

MOCAGE has the flexibility to be used for stratospheric studies (El Amraoui et al., 2008a) and tropospheric studies (Dufour et al., 2004). It is used in the operational AQ monitoring system in France: Prev'air (Rouil et al., 2008) and in the pre-operational GMES (Global Monitoring for Environment and Security) atmospheric core service (Hollingsworth et al., 2008).

The model uses 2 nested domains, at 2° over the globe and at 0.5° over Europe, from 32°N to 72°N and from 16°W to 36°E. The nature run simulation covers the period from July 1st, 2009 to September 1st, 2009. The simulated field for July 1st, 2009 has been obtained from a free run

with RACMOBUS started from a June climatological initial field. The meteorological analyses of Météo-France, ARPEGE (Courtier et al., 1991) were used to force the dynamics of the model every 6 hours. The emission inventory used in the nature run is the inventory provided by TNO (Netherlands Organization for Applied Scientific Research) (Visschedijk and Denier van der Gon, 165 2005), for the Global and regional Earth-system Monitoring using Satellite and in-situ data (GEMS) project (Hollingsworth et al., 2008); hereinafter noted GEMS-TNO. This inventory has a high spatial resolution of $10 \times 10 \text{ km}^2$, and a temporal resolution of 1 hour. It is representative of the year 2003.

2.3 The synthetic observations

In this study, we generate synthetic observations for two nadir infrared GEO platforms. The first 170 one (GEO-TIR) has a SSI (0.05 cm^{-1}) and a Noise Equivalent Spectral Radiance (NESR: $1.00 \text{ nW}/(\text{cm}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$ and $6.04 \text{ nW}/(\text{cm}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$ for the CO and O₃ spectral windows, respectively) equivalent to the infrared instrument proposed in the MAGEAQ mission to EE-8, dedicated to monitoring CO and O₃ the LmT (Claeyman et al., 2010b). The second one (GEO-TIR2) has a SSI (0.625 cm^{-1}) and a NESR ($6.12 \text{ nW}/(\text{cm}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$ and $24.5 \text{ nW}/(\text{cm}^2 \cdot \text{sr} \cdot \text{cm}^{-1})$ for the CO and O₃ spectral 175 windows, respectively) similar to the MTG-IRS (Stuhlmann et al., 2005). Considering the high computing cost associated with generating OSSEs, we define the pixel size at $0.5^\circ \times 0.5^\circ$, corresponding to the model spatial resolution and the revisit time at 1 hour for both instruments. A resolution of $0.5^\circ \times 0.5^\circ$ for AQ monitoring over Europe is commonly used in operational systems (e.g., Prev'air in France).

180 To represent the synthetic observations in the OSSE, we need temperature and water vapour fields and their uncertainty. Following the MTG-IRS retrieval study of Clerbaux et al. (2008a), we assign uncertainties at each vertical level of 0.5 K for temperature and 10% for water vapour. The number of pixels at $0.5^\circ \times 0.5^\circ$ of an instrument onboard a geostationary platform is very important. In our case, we have to consider about 100 000 profiles per instrument per species per day over the defined 185 domain (Europe). Then, to study 2 months of synthetic observations for the 2 instruments, we set up a method much faster than using traditional radiative transfer and retrievals models. In the following we define the method and its validation.

Retrievals of LmT O₃ and CO in the infrared strongly depend on the thermal contrast between the surface and the air immediately above it (see e.g., Deeter et al., 2007; Eremenko et al., 2008; Clerbaux et al., 2009). Several parameters (e.g., measurement and temperature error) have to be taken 190 into account to assess the sensitivity of such retrievals. However, among these errors, the smoothing error is the main contributor to the shape of the averaging kernels, which represent the sensitivity of the retrieval to the true atmosphere at different altitudes. From these averaging kernels, one can deduce for example the surface sensitivity of the retrieval. Because of the strong dependence of 195 the averaging kernels on the thermal contrast, we construct a look-up table containing the specific values of the thermal contrast and their corresponding averaging kernels. In addition, to refine the

method, we include in the look-up table other errors such as the measurement error and the temperature error, assuming a linear regime. This look-up table is built using the forward model KOPRA (Karlsruhe Optimized and Precise Radiative transfer Algorithm) (Stiller et al., 2002). The retrieval system KOPRA-fit (Höpfner et al., 1998), based on the Tikhonov-Phillips regularization is also employed (Tikhonov, 1963; Phillips, 1962). We generate the averaging kernels and the corresponding covariance matrix error for several thermal contrast values between -20 K and 20 K with a step of 0.2 K representing a total number of 201 values for each instrument configuration. The range of thermal contrast values has been established using statistics on the thermal contrast found in the temperature analyses of the current version of the ARPEGE global model. This method allows us to provide quickly (with a speed up factor of more than 70 in terms of calculation speed) the required parameters (errors and averaging kernels) that correspond to any thermal contrast. From these parameters we reconstruct the different trace gas profiles using the quantity (Rodgers, 2000):

$$\mathbf{x}_{sim} = \mathbf{A}\mathbf{x}_t + (\mathbf{I} - \mathbf{A})\mathbf{x}_a + \varepsilon \quad (1)$$

with \mathbf{x}_{sim} the simulated retrieved profile, \mathbf{x}_t the true profile corresponding to the calculated profile (nature run) from MOCAGE CTM, \mathbf{x}_a the a priori profile - a climatology over Europe calculated from the MOCAGE model and \mathbf{A} the averaging kernel matrix. ε is defined as a random Gaussian error with a standard deviation corresponding to the square root of the diagonal elements of the error covariance matrix. Note that these quantities are defined in terms of $\ln(\text{VMR})$, where VMR stands for the volume mixing ratio. For further details on the averaging kernel shapes and covariance matrix errors of GEO-TIR and GEO-TIR2 the reader should refer to Claeys et al. (2010b).

A similar method was used in Edwards et al. (2009) to simulate CO infrared observations using 3 different averaging kernel sets. We validate the method by comparing the values from the look-up table and the results calculated with the comprehensive KOPRA-fit method. The details of the statistics obtained for the validation exercise for both GEO-TIR and GEO-TIR2 and for observations at altitudes between the surface and 10 km are shown in Table 1. The statistics show a very good agreement between the values provided by the look-up table and the KOPRA-fit method. All the correlation coefficients are greater than 0.9 for both O_3 and CO, and for the two instrument configurations. In addition, standard deviations (between 1.7% and 4.8%) and biases (between -0.4% and 1.3 %) are small. Moreover, the histograms of the relative difference between the look-up table and KOPRA-fit (not shown) show a Gaussian-like shape around the value 0 confirming the good match of the two methods.

This validation highlights the good correspondence between both methods. We then use the look-up table to generate observations for two instrument configurations (GEO-TIR and GEO-TIR2) over the two months of the study. To account for cloudy scenes, cloud estimates from the ARPEGE model are used to assign a cloud fraction to the observation pixels. Pixels with a cloud fraction greater than 0.5 are filtered out. The vertical grid is provided by the retrieval, with a step of 1 km from the surface to 39 km. Since we are interested in the relative added value, we use for both instruments the same

approximations to generate the observations. This makes the problem tractable, and is not expected
235 to change the results.

2.4 The assimilation scheme

The assimilation system used in this study is MOCAGE-PALM (Massart et al., 2005). The assimila-
tion module is implemented within the PALM framework (Buis et al., 2006). The used assimilation
technique is the 3D-FGAT (First Guess at Appropriate Time, Fisher and Andersson (2001)). This
240 technique is a compromise between the 3D-Var (3d-variational) and the 4D-Var (4d-variational)
methods. It has been validated during the assimilation of ENVISAT data project (ASSET, Lahoz
et al. (2007)) and has produced good quality results compared to independent data and other assimi-
lation systems (Geer et al., 2006). Further details on the assimilation system can be found in Massart
et al. (2009), El Amraoui et al. (2008b) and Claeysman et al. (2010a).

2.5 The experiments

To study the sensitivity of the OSSEs to various key parameters, we perform several experiments
summarized in Table 2. For these simulations, we also used MOCAGE but with different configu-
rations in order to have an alternative representation of the atmospheric composition, different from
the nature run. For all experiments (except the nature run), we perform 3 simulations: the control run
250 without data assimilation, the assimilation run with assimilation of GEO-TIR and the assimilation
run with the assimilation of GEO-TIR2.

The first sensitivity test concerns the meteorology. In the nature run we use the ARPEGE analysis
every 6 hours whereas in the control run and assimilation run we use the 48 hour forecasts every 6
hours. It is denoted hereafter EXP1.

255 In a second sensitivity test, we change the emission inventory. Instead of the detailed GEMS-TNO
inventory used in the nature run, we use a global inventory where emissions are given as a monthly
mean for biomass burning and a yearly mean for other sources (Dentener et al., 2006) representing
the year 2000 (EXP2). Both inventories use different daily and monthly emission factors. Figure
1 shows the emission map of CO and NO_x (NO+NO₂, an O₃ precursor), emitted over Europe on
260 July 6th, 2009 according to both inventories. In the GEMS-TNO inventory, emissions show a higher
variability than in the global inventory. For example, over Paris or over Madrid the maximum values
are higher in the GEMS-TNO inventory, whereas in Northern Europe or in Spain over rural areas,
CO and NO_x emissions are lower in the GEMS-TNO inventory. However, both inventories show
the same emissions of NO_x from ships. In Figure 2, the emission diurnal cycle is shown for CO and
265 NO_x and emissions are accumulated over Europe for each hour of July 6th, 2009. Generally, more
CO and NO_x are emitted by the global inventory than by the GEMS-TNO inventory but locally over
large European cities the opposite is the case. Three peaks are observed in the global inventory at
6h, 12h and 18h UTC for both CO and NO_x emissions whereas only 2 are observed at 8h and 17h

UTC for CO, and at 8h and 18h UTC for NO_x in the GEMS-TNO inventory.

270 In the third sensitivity test, the initial conditions are modified (EXP3). In the nature run, the initial condition from July 1st, 2009 is provided from a previous free run. For the control run and the assimilation runs, we change the initial condition every week by taking the field from the nature run one week before (e.g., the initial field from July 1st, 2009 in the control run and assimilation runs is provided by the field from May 25th, 2009 from the nature run). We repeat this change every week
275 to keep a significant difference between the nature run and the control run (see section 3); after one week the influence of the initial condition is very low in the LmT on O₃ and CO concentrations (not shown). This modification introduces discontinuity in the O₃ and CO time-series, and this effect is considered in the next section.

The last experiment (EXP4) involves all of the 3 sensitivity tests (meteorology, emissions and
280 initial condition). This experiment contains the main errors encountered in an AQ model (e.g., Menut and Bessagnet, 2010), except the chemical scheme and the transport scheme which are kept the same for all experiments presented here. Although this may impact the results of the study, we consider that for this OSSE, the nature run and the control run, and the nature run and the assimilation runs have enough realistic differences to make the experiments meaningful (see section 3).

285 Table 3 presents the correlation, the bias and the RMS between the 4 control runs (EXP1a, EXP2a, EXP3a and EXP4a) and the nature run averaged over 2 months over Europe. The 4 sensitivity tests generate different errors: EXP1a is characterized by high RMS (~10 % for O₃ and ~7 % for CO) and low bias (0.19 % for O₃ and -1.02 % for CO); EXP2a by high bias (~8 % for O₃ and CO), high correlation (>0.9) and low RMS (~5%); EXP3a by a low correlation (<0.7), high RMS (~13 % for
290 O₃ and ~8 % for CO) and low bias (~2%). EXP4a mixes up all these effects and is characterized by high bias (~7 %), high RMS (~17 % for O₃ and ~11 % for CO) and low correlation (~0.5). The 4 experiments have different characteristics and can bring information on the capabilities of GEO-TIR and GEO-TIR2 to constrain several parameters in the LmT.

2.6 Statistical Analysis

295 The impact of the observations (GEO-TIR and GEO-TIR2) is evaluated by comparing the results from the control run and the assimilation runs with the “truth” represented by the nature run. To provide a degree of robustness to our OSSEs, we perform significance tests to check at the 0.95 and 0.99 confidence limit if differences between the control run and the nature run and differences between the assimilation runs and the nature run are significant, as was done in Lahoz et al. (2005).

300 The null hypothesis is that the means of the differences between the control run and the nature run and the differences between the assimilation runs and the nature run are the same. The datasets have sufficient data to assume a normal distribution.

We used the two-sample hypothesis z -test defined as:

$$Z = \frac{|\overline{CR - NR}| - |\overline{AR - NR}|}{\sqrt{\frac{\sigma_{CR-NR}^2}{N} + \frac{\sigma_{AR-NR}^2}{N}}} \quad (2)$$

where NR is the nature run dataset, CR is the control run dataset, AR is the assimilation run dataset, σ is the root-mean square (RMS) and N is the number of grid points. Vertical lines indicate absolute value.

Furthermore, in order to quantify the GEO-TIR and GEO-TIR2 added values, we compute indicators commonly used in AQ modelling evaluation: absolute difference, RMS difference and temporal correlation. For the protection of public health, the WHO (World Health Organization, 2005; Krzyzanowski and Cohen, 2008) has established a threshold at $100 \mu g.m^{-3}$ of O_3 concentrations for the maximum of a 8-hour running average. We use this threshold to calculate 3 contingency tables: the percentage of good detections (GD), the percentage of correct analyses above threshold (GD+) and the percentage of false alarms (FA) calculated as follows:

$$GD = 100 \times \frac{(NR1_AR1 + NR0_AR0)}{N} \quad (3)$$

$$GD+ = 100 \times \frac{NR1_AR1}{NR1} \quad (4)$$

$$FA = 100 \times \frac{NR0_AR1}{AR1} \quad (5)$$

305 where NR1_AR1 represents the number of grid points where the nature run is greater than $100 \mu g.m^{-3}$ and the assimilation run (or control run) is above $100 \mu g.m^{-3}$; NR0_AR0 represents the number of grid points where the nature run is less than $100 \mu g.m^{-3}$ and the assimilation run (or control run) is less than $100 \mu g.m^{-3}$; N is the number of all grid points; NR1 represents the number of grid points where the nature run is greater than $100 \mu g.m^{-3}$; NR0_AR1 represents the number of
 310 grid points where the nature run is less than $100 \mu g.m^{-3}$ and the assimilation run (or control run) is greater than $100 \mu g.m^{-3}$; and AR1 represents the number of grid points where the assimilation run (or control run) is greater than $100 \mu g.m^{-3}$.

3 Results

3.1 Evaluation of the nature run

315 We compare the nature run provided by the MOCAGE model to O_3 and CO ground-based station observations over France from July 1st, 2009 to August 31th, 2009, to verify that the nature run is representative of the “true atmosphere”.

Figure 3 shows the time-series of CO (panels a and b) and O_3 (panels e and f) simulated by MOCAGE (nature run) and observed by ground stations over France in July and August 2009. CO

320 from the nature run is generally higher than CO from ground stations. Some maxima are well represented (e.g., 28th and 29th July, 2009), some maxima are overestimated (e.g., 10th August, 2009) and some other are underestimated (e.g., 19th August, 2009). However, most importantly, the CO concentrations simulated in the nature run are in the same range of values (globally between 50 and 500 $\mu\text{g.m}^{-3}$) as those observed by ground stations, and show similar temporal variability.

325 O₃ concentrations simulated in the nature run are also globally overestimated compared to ground measurements. However, the diurnal cycle of production and destruction of O₃ is well represented in the nature run. The minima of O₃ in the nature run are generally overestimated, except over particular periods, where the nature run and the observations show a good agreement (e.g., from July 28th to August 1st; from August 5th to August 6th or from August 16th to August 20th).

330 Table 4 shows the correlation, the bias and the RMS between the nature run and the ground stations over France on a hourly mean basis for O₃ and CO. The correlation coefficients are 0.76 and 0.63 for O₃ and CO, respectively. For both O₃ and CO a positive bias is observed (12 $\mu\text{g.m}^{-3}$ ($\sim 18\%$) and 19.9 $\mu\text{g.m}^{-3}$ ($\sim 17\%$), respectively). The RMS is larger for CO (59.9 $\mu\text{g.m}^{-3}$ $\sim 52\%$) than for O₃ (18.2 $\mu\text{g.m}^{-3}$ $\sim 26\%$) likely because CO concentrations have a great variability and can be

335 locally very high at the surface ($> 1000 \mu\text{g.m}^{-3}$).

Despite the fact that the simulations are performed at $0.5^\circ \times 0.5^\circ$, the results concerning the comparison between ozone surface observations and the nature run over France are comparable to those commonly observed in AQ forecasting. For example, Pagowski et al. (2006) computed bias, RMS and correlation of hourly concentrations forecast over the Eastern USA and Southern Canada for

340 July and August 2004 for seven AQ models compared to hourly surface ozone measurements over 350 sites. The bias ranges between 10.6 and 62.2 $\mu\text{g.m}^{-3}$, the RMS between 33.0 and 74.9 $\mu\text{g.m}^{-3}$ and the correlation between 0.55 and 0.72. In another study on the French AQ forecasting system Prev'air (Honoré et al., 2008) the bias for the ozone hourly forecasts was 12.3 $\mu\text{g.m}^{-3}$, the RMS 28.2 $\mu\text{g.m}^{-3}$ and the correlation 0.67. The scores of the nature run are in the same range of values

345 than Pagowski et al. (2006) which indicates that the nature run is fairly close to the ground-based observations and can be assumed to be representative of the “true atmosphere” over the European domain.

3.2 Spatial distribution of the impact of geostationary infrared measurements in the lowermost troposphere

350 Figure 4 presents the area of Europe where differences between various experiments are significant at the 0.95 and 0.99 confidence limit for the O₃ and CO LmT columns using the two-sample hypothesis z-test (section 2.6). This test assesses whether the control run and the GEO-TIR assimilation run; and the GEO-TIR2 assimilation run and the GEO-TIR assimilation run, are significantly different (with a confidence limit of 95 and 99%). Figure 4 shows that EXP2, EXP3 and EXP4 have large areas

355 of significance at the 0.99 confidence limit (red areas). Areas which are not significant at the 0.99

confidence limit nor at the 0.95 confidence limit are generally over sea, which is less important for AQ purposes as we are interested in highly populated areas. However, EXP1 shows less significant areas at the 0.95 confidence limit than other experiments - the implications of this are discussed later. All the statistics presented hereafter are for a period of 2 months (July and August 2009).

Our objective is to have a statistically robust evaluation of the added value of GEO-TIR synthetic observations for air quality hindcasts. However, it will be difficult to substantiate the reasons for the spatial distribution of the OSSE increments averaged over two months; indeed, over such a period, there is a combination of different conflicting effects explaining variations of the strength of the constraint brought by GEO-TIR and GEO-TIR2 synthetic observations. These can only be understood by studying cases on a day-by-day basis, which is outside the scope of this paper.

3.2.1 Sensitivity study on meteorology: experiment 1

We performed a sensitivity study using different meteorology for the control run (EXP1a) and assimilation runs (EXP1b and EXP1c) compared to the ones used for the nature run, to determine the capability of GEO-TIR to reduce differences generated by the meteorology used in our analyses.

Figure 5 shows the correlation, the bias and the RMS for the O₃ LmT column between the nature run and the control run and the improvement added by the assimilation of GEO-TIR compared to the control run and to the assimilation run for GEO-TIR2.

The correlation between the nature run and the control run for O₃ ranges between 0.5 and 0.9. The added value of GEO-TIR (red colours) is mainly over Spain, North Africa and the Atlantic Ocean where the results are significant at the 0.95 confidence limit. The assimilation of GEO-TIR increases the correlation from ~0.7 in the control run to ~0.8 in the GEO-TIR assimilation run, mainly over the Atlantic ocean and over Spain. Similar results are observed concerning the added value of the GEO-TIR assimilation run compared to the GEO-TIR2 assimilation run: GEO-TIR is closer to the nature run. The bias between the nature run and control run for O₃ is low (between -8% and 8%) and mainly negative over the Mediterranean Basin and positive over Northern Europe. The GEO-TIR assimilation run reduces the bias over the Mediterranean Basin and over the Nordic countries, which are regions with significance at the 0.99 confidence limit compared to the control run and to the GEO-TIR2 assimilation run. The RMS between the nature run and the control run is between 4 and 25% for O₃. The GEO-TIR assimilation run reduces globally the RMS to 5% over sea and land areas.

Figure 6 shows the same diagnostics but for the CO LmT column. The correlation between the nature run and the control run for CO ranges also between 0.5 and 0.9. The positive impact of the GEO-TIR assimilation run on the control run is bigger than for O₃ with a significant improvement of the correlation (e.g., from 0.7 between the nature run and the control run to 0.85 between the nature run and the GEO-TIR assimilation run over Spain and France, or from 0.85 between the nature run and the control run to 0.95 between the nature run and the GEO-TIR assimilation run

over Turkey). The assimilation of GEO-TIR2 also improves the correlation between the nature run and GEO-TIR2 assimilation run compared to the correlation between the nature run and the control run (e.g., over the Atlantic ocean or over Turkey) but the impact of GEO-TIR is higher. The bias
 395 between the control run and the nature run for CO is low and mainly negative ($\sim -3\%$) except over the Po valley where the bias is high and positive (15%). This large difference between the control run and the nature run over the Po Valley can be explained by differences in the winds since the meteorology in the nature run is significantly different to that in the control and assimilation runs. In the control run, pollutants are trapped in the Po Valley which is surrounded by the Alps whereas
 400 in the nature run, pollutants are transported by the winds. For this particular event, the GEO-TIR assimilation run reduces considerably the bias observed compared to the control run and to the GEO-TIR2 assimilation run, and does this to a lesser extent over France and Eastern Europe. The RMS between the control run and the nature run for CO is $\sim 7\%$ but can reach 25% over the Po valley. The GEO-TIR assimilation run reduces globally the RMS observed in the control run and in the GEO-
 405 TIR2 assimilation run ($\sim 2\%$), with a particular emphasis on the Po valley where the RMS added value is $\sim 11\%$ compared to the control run and $\sim 7\%$ compared to the GEO-TIR2 assimilation run. Note that results observed over the Po valley for CO are significant at the 0.99 confidence limit.

In this experiment, we have analysed the capabilities of both instruments to correct errors in the meteorology. The resulting control run generally shows low biases for both CO and O₃ but impacts
 410 the correlation and the RMS. For this particular experiment, the GEO-TIR assimilation run improves considerably the RMS and locally the bias and the correlation.

3.2.2 Sensitivity study on emissions: experiment 2

In this experiment (EXP2), we use another emission inventory in the control run (EXP2a) and assimilation runs with a coarser spatio-temporal resolution than the one used in the nature run (see section
 415 2.5). Figure 7 shows also the correlation, the bias and the RMS for the O₃ LmT column between the nature run and the control run and the improvement added by the assimilation of GEO-TIR (EXP2c) compared to the control run and to the assimilation run for GEO-TIR2 (EXP2b).

The correlation between the nature run and control run is very high for O₃ (> 0.95), especially over sea where both inventories use the same emissions. The impact of the GEO-TIR assimilation
 420 on the correlation coefficient is relatively small compared to the control run and the GEO-TIR2 assimilation run and is located over the Eastern Mediterranean Basin where the correlation between the nature run and the control run is lower (~ 0.7). However, the bias between the nature run and control run is positive and high (up to 20%) because emissions of NO_x and CO are higher in the inventory used in the control run and assimilation runs (Figure 2). The impact of the GEO-TIR
 425 assimilation run is very high and can reduce by a factor of 2 the bias over the Mediterranean Basin both for the control run and the GEO-TIR2 assimilation run. The RMS between the nature run and control run for O₃ is very low over sea (less than 4%), but over land it can reach 15% (e.g., Spain,

South West of France, Northern Africa). The GEO-TIR assimilation run reduces by $\sim 1\%$ the RMS compared to the control run and to the GEO-TIR2 assimilation run over Southern Europe (except
 430 over the Atlantic ocean) but locally over specific areas (e.g., over Spain), GEO-TIR can bring an improvement of 5%. Note that the significance is at the 0.99 confidence limit almost everywhere for O_3 for this experiment (except over a small region over the Atlantic ocean, see Figure 4).

Figure 8 shows similar diagnostics to Figure 7 but for CO. As for O_3 , the correlation coefficient between the control run and the nature run is very high which leads to a very low impact of GEO-
 435 TIR compared to the control run and to the GEO-TIR2 assimilation run. This impact can locally be slightly negative (e.g., over the Atlantic ocean). This negative impact may come from the observation errors, which are discussed in detail in Claeyman et al. (2010b) for an instrument similar to GEO-TIR. As for O_3 , the bias between the control run and the nature run is very high and can reach 20% as the inventory used in the control run and assimilation runs emitted more CO, but only locally. Over
 440 large cities (e.g., Paris, Turin, Amsterdam, Saint Petersburg, consistent with the emission map in Figure 1), the results for CO in the LmT reflect differences between the global and the GEMS-TNO emissions inventories. The GEO-TIR assimilation run reduces the overall bias to 15% and 10% over the Mediterranean Basin compared to the control run and GEO-TIR2 assimilation run, respectively, but brings little improvement over these large cities where CO concentrations in the control run and
 445 GEO-TIR2 assimilation run are low. The RMS between the nature run and the control run is $\sim 7\%$ over land and very low over the Atlantic ocean, but can locally reach 20% (e.g., South Italy, Greece). GEO-TIR improves also the RMS compared to the control run and especially over land and over the Mediterranean Basin compared to the GEO-TIR2 assimilation run. The RMS of GEO-TIR degrades over the Atlantic ocean (where significance is not at the 0.95 confidence limit) but also in South East
 450 Europe compared to GEO-TIR2 assimilation run where the RMS between the control run and the nature run is low. This can also be explained by the GEO-TIR observation errors.

In this experiment, we analyse the capability of the 2 observing systems to correct errors in the emissions. This experiment shows that GEO-TIR is able to considerably reduce the global bias observed in the control run in the LmT for both O_3 and CO and can also bring significant skill
 455 compared to GEO-TIR2.

3.2.3 Sensitivity study on the initial condition: experiment 3

In this experiment (EXP3), we change the initial condition every week (see section 2.5) in the control run (EXP3a) and in the assimilation runs to quantify the capability of GEO-TIR (EXP3c) and GEO-TIR2 (EXP3b) to correct for these differences. Figure 9 shows that the correlation for the O_3 LmT
 460 column between the nature run and the control run ranges between 0.3 (e.g., over Atlantic Ocean or Turkey) and 0.9 (e.g., over Italy). The correlation coefficient for O_3 is lower than in previous experiments (EXP1 and EXP2) since the modification of the initial condition every week artificially brings down the correlation. The GEO-TIR assimilation run improves the correlation compared to

the control run and to the GEO-TIR2 assimilation run, both over land and sea. This positive impact
 465 of GEO-TIR can improve the correlation (e.g., from 0.3 between the nature run and the control run
 and 0.5 between the nature run and GEO-TIR2 assimilation run up to 0.8 between the nature run and
 the GEO-TIR assimilation run over Turkey). The bias between the control run and the nature run
 for O_3 is low in the Southern part of Europe and is mainly positive over the Atlantic Ocean and over
 Russia. The added value of GEO-TIR compared to the control run and the GEO-TIR2 assimilation
 470 run is overall low but positive ($\sim 1\%$) and is higher over Russia where the significance is at the
 0.99 confidence limit (but can reach 6% and 4% compared to the control run and the GEO-TIR2
 assimilation run, respectively). The RMS between the nature run and the control run is higher in the
 Northern part of Europe ($\sim 20\%$) than in the Southern part ($\sim 7\%$). The assimilation of GEO-TIR
 reduces the RMS by $\sim 2\%$, particularly where the RMS difference between the nature run and the
 475 control run is high $\sim 5\%$ (e.g., Northern Atlantic ocean).

The correlation between the nature run and the control run for the CO LmT column ranges be-
 tween 0.3 (e.g., over Aegean Sea) and 0.9 (e.g., over France and Germany). The assimilation of
 GEO-TIR improves considerably the correlation compared to the control run (from ~ 0.7 between
 the nature run and the control run to ~ 0.9 between the nature run and the GEO-TIR assimilation run)
 480 over the Mediterranean Basin, where the significance is at the 0.99 confidence limit. The GEO-TIR
 assimilation run also improves the correlation compared to the GEO-TIR2 assimilation run espe-
 cially over the Aegean Sea, Spain and North Africa. The bias and the RMS between the nature
 run and the control run for CO are low: $\sim 2\%$ for the bias and between 4 and 12% for the RMS.
 The impact of GEO-TIR assimilation run on the bias is then positive but very low compared to the
 485 control run and the GEO-TIR2 assimilation run; and the impact on the RMS is locally high, 7% and
 6% compared to the control run and the GEO-TIR2 assimilation run, respectively over Turkey and
 over Spain, and is positive but low elsewhere.

The modification of the initial condition mainly impacts the correlation for both CO and O_3 .
 This experiment shows that the assimilation of GEO-TIR can improve considerably the correlation
 490 coefficient over land and sea for the CO and O_3 LmT column.

3.2.4 Sensitivity study on the emissions, meteorology and initial condition: experiment 4

We perform a final sensitivity test by simultaneously changing the emissions, the meteorology and
 the initial condition (Figure 11). The control run (EXP4a) for the O_3 LmT column is characterized
 by low correlation (between 0 and 0.7), high bias ($\sim 15\%$ on average), and high RMS ($\sim 17\%$ on
 495 average) compared to the nature run. By construction, we expect this experiment to provide results
 that differ the most from the nature run. The impact of the assimilation of GEO-TIR (EXP4c) is
 high compared to the control run and the GEO-TIR2 assimilation run (EXP4b). The added value
 of GEO-TIR for the correlation coefficient is positive over Europe and increases significantly the
 correlation coefficient (e.g., over Turkey, Germany, Atlantic Ocean). The GEO-TIR assimilation run

500 reduces the bias by 3% and 2% in average but locally the impact is $\sim 5\%$ and $\sim 6\%$, compared to the control run and the GEO-TIR2 assimilation run, respectively. The RMS is considerably reduced all over all Europe up to 12% and 10% compared to the control run and the GEO-TIR2 assimilation run, respectively.

The differences between the nature run and the control run for the CO LmT column (Figure 12) are 505 similar to those for O_3 : low correlation coefficient (between 0 and 0.8), high bias ($\sim 11\%$ on average) and high RMS ($\sim 11\%$ on average). As for O_3 , this CO experiment provides results that differ the most from the nature run, as expected. The impact of the assimilation of GEO-TIR is positive over all the Europe, where the significance is at the 0.99 confidence limit: it increases the correlation (from 0.4 between the nature run and the control run and 0.6 between the nature run and the GEO-TIR2 510 assimilation run up to 0.8 between the GEO-TIR assimilation run over Turkey), reduces the bias (up to 20% and 15% over the Po valley compared to the control run and the GEO-TIR2 assimilation run, respectively); and reduces the RMS (up to 14% and 9% over Turkey compared to the control run and the GEO-TIR2 assimilation run, respectively).

We have presented a statistical analysis over 2 months to characterize the added value of the two 515 instrument configurations. The results of the 4 experiments show that the assimilation of GEO-TIR improves significantly the O_3 and CO LmT columns compared to the control run and the assimilation of GEO-TIR2. The assimilation of GEO-TIR is able to effectively constrain the O_3 and CO fields perturbed by different sources of error in air quality prognoses: meteorology, emission, initial state.

The added value of GEO-TIR is high over land and over sea. Concerning results over land, nadir 520 infrared measurements are well known to be sensitive to the LmT with high thermal contrast and high surface temperature (namely over land during day) (e.g., Deeter et al., 2007; Eremenko et al., 2008; Clerbaux et al., 2009). Concerning results over sea, they suggest that via direct assimilation and/or transport of successive increments by the model, the added value of GEO-TIR also impacts the sea (e.g., vertical and horizontal transport, Foret et al. (2009)).

525 The largest effects are mainly located over the Mediterranean Basin, where the cloud cover is weaker and surface temperatures and thermal contrasts are high over country surrounding coastal areas. In contrast, the added value of GEO-TIR is rather limited over the North Western part of the domain (Atlantic Ocean). Due to predominant winds blowing from the West in the area, air masses are largely influenced by incoming fluxes situated outside the field of view of our simulated geostationary platforms, and the effects of assimilation are strongly mitigated. Also, the spatial distribution 530 of the efficiency of GEO-TIR simulated observations to bring the assimilation run statistically close to the nature run are governed to a large extent by the spatial distribution of the differences between the nature run and the different control runs: GEO-TIR can in fact better constrain fields where the nature and control runs differ most, while where nature and control runs agree, little effect from the 535 assimilation is expected, as seen in practice.

3.3 Vertical distribution of the impact of geostationary infrared measurements in the lower-most troposphere

In section 3.2.1, we have quantified the added value of the assimilation of GEO-TIR for four sensitivity studies on the CO and O₃ LmT column over Europe. In this section, we concentrate on the vertically resolved added value of GEO-TIR in the lower troposphere (0-5 km) compared to the control run and the GEO-TIR2 assimilation run. Figure 13 show the correlation, the absolute difference and the RMS between the control run and the nature run, the GEO-TIR2 assimilation run and the nature run, and the GEO-TIR assimilation run and the nature run, for the four sensitivity studies (EXP1, EXP2, EXP3 and EXP4) averaged over Europe for 2 months (July and August 2009) as a function of altitude (surface up to 5 km) for O₃. For the O₃ correlation, the impact of the assimilation of GEO-TIR improves considerably it for EXP3 and EXP4, slightly for EXP1 but is not significant for EXP2. The vertical improvement of the correlation by the assimilation of GEO-TIR is very low at the surface, slight at 1 km, but high from 2 to 5 km, whereas the impact of GEO-TIR2 is very low for all levels between the surface and 5 km for O₃. For the bias and the RMS, similar conclusions can be made: the impact of GEO-TIR is highly dependent on the experiment and the altitude, and reduces the bias and the RMS mainly for altitudes above ~1 km whereas the impact of GEO-TIR2 is very low for O₃.

The results are highly dependent on the experiments, but the impact of the assimilation of GEO-TIR improves considerably the O₃ analyses compared to the nature run above 1 km. Note that Honoré et al. (2008) showed that the mean model bias of daily ozone maxima was mostly under 5 $\mu\text{g}/\text{m}^{-3}$ (~7%), RMS was generally less than 20 $\mu\text{g}/\text{m}^{-3}$ (~30%) and temporal correlation was more than 0.8 on average over Western Europe compared to O₃ surface observations, which indicates that the correlation and the bias observed between the nature run and the control runs are realistic. The RMS in the control run is underestimated which may be because in the study from Honoré et al. (2008) the average is made over land and over Western Europe, whereas in this study the average is made over Europe (including the sea where O₃ concentrations show less variability at the surface).

Figure 14 shows similar results as Figure 13 but for CO. The assimilation of GEO-TIR improves considerably the CO correlation for EXP1, EXP3 and EXP4 but has little impact on EXP2, which has already a high correlation coefficient. The positive impact of GEO-TIR is mainly situated above 1 km except for EXP4, which has a lower correlation (~0.7); the assimilation of GEO-TIR improves the correlation at the surface. The assimilation of GEO-TIR2 CO also improves the correlation (but not at the surface) but the GEO-TIR assimilation run is closer to the nature run. The assimilation of GEO-TIR and GEO-TIR2 also reduces the bias and the RMS, especially for EXP2 and EXP4 which show high biases, but the GEO-TIR assimilation run is closer to the nature run than the GEO-TIR2 assimilation, particularly at the surface.

3.4 Ozone evaluation at the surface

As for AQ purposes we are mainly interested by pollutant surface concentrations, we focus on the added values of both geostationary instruments on ozone surface concentrations.

We compute the percentage of good detection (GD), the percentage of correct detection above threshold (GD+) and the percentage of false alarms (FA) (see section 2.6) for the control run, the GEO-TIR2 assimilation run and the GEO-TIR assimilation run for the four experiments at the surface over land for the European domain (Table 5). The observations are simulated throughout the nature run. We select the threshold at $100 \mu\text{g.m}^{-3}$ for the maximum of the 8-hour running average, established by the WHO (World Health Organization, 2005) for the protection of public health. We do not compute the same scores for CO since the threshold for the protection of public health for the maximum of the 8-hour running average is $10000 \mu\text{g.m}^{-3}$ which is seldom observed in a non-confined atmosphere. Furthermore, CO is interesting for AQ because it is a proxy for pollutant sources and transport processes and not because of its direct impact on human health.

We have already shown that in general the added value of GEO-TIR and GEO-TIR2 for O_3 at the surface is low. However, for particular cases (high concentrations of O_3 above the threshold) the results presented in Table 5 indicate that the assimilation of geostationary instruments can help better in detecting high concentration events. In all cases, except EXP1 for GEO-TIR2, good detection and false alarm scores are enhanced both for GEO-TIR and GEO-TIR2. Concerning threshold–overshoot detections, results are more contrasting. The GEO-TIR assimilation is better than the control run for 2 experiments (EXP1 and EXP3) and GEO-TIR2 only for 1 (EXP1).

Comparing the GEO-TIR and GEO-TIR2 assimilation runs, in only 2 out of 12 cases, GEO-TIR2 is better than GEO-TIR (Table 5). Finally, GEO-TIR gives better scores than both the GEO-TIR2 assimilation run and the control run in 9 out of 12 cases.

4 Summary and Conclusions

In this paper, we perform an OSSE for geostationary infrared instruments to determine their relative added values for O_3 and CO concentrations in the lowermost troposphere (LmT; defined to be the atmosphere between the surface and 3 km) in an AQ model over Europe. The originality of this study is to use an AQ model in an OSSE to assess the impact of various key parameters (emissions, meteorology, initial condition and the 3 parameters together) on analyses derived using two infrared instruments. The first one (GEO-TIR) has an instrument configuration (SNR and SSI) dedicated to monitoring O_3 and CO in the LmT, similar to the MAGEAQ infrared instrument (Peuch et al., 2010); the second one (GEO-TIR2) has an instrument configuration (SNR and SSI) mainly dedicated to measure temperature and humidity and is similar to the MTG-IRS instrument (Clerbaux et al., 2008b). For both instruments we use a pixel size of $0.5^\circ \times 0.5^\circ$ and a revisit time of one hour.

We first concentrate on the capability of GEO-TIR and GEO-TIR2 to simulate the distributions of

the O₃ and CO LmT column over Europe, using statistical diagnostics averaged over 2 months (July and August 2009). The GEO-TIR assimilation runs are closer to the nature run than the GEO-TIR2 assimilation runs for almost all experiments. The positive impact of GEO-TIR is highly dependent on the experiment and similar behaviour is observed for the O₃ and CO LmT columns. For experiments involving changes in emissions GEO-TIR is able to significantly reduce the systematic bias produced by excessive emissions. For experiments involving changing the initial conditions or the meteorology, GEO-TIR is also able to considerably increase the correlation coefficient with respect to the nature run and reduce the RMS in comparison to the control run. The added value of GEO-TIR impacts both over land and sea areas, but is mainly situated near the Mediterranean Basin. The different experiments also show that when the bias and the RMS are very low or the correlation very high, the GEO-TIR assimilation run has little impact and can even slightly degrade the analyses at particular locations if the control run error is very small and the observation error is bigger. We show that the added value of the two instruments is experiment dependent and is mainly governed by the spatial distribution of the differences between the nature run and the different control runs. Even if nadir infrared instruments are well-known to be sensitive in the LmT for high thermal contrast (mainly over land during daytime), the assimilation and the successive transport of increments by the model during 2 months bring added value of GEO-TIR and GEO-TIR2 also over the sea in the LmT.

We also quantify the vertically resolved impact of both GEO-TIR and GEO-TIR2 from the surface to 5 km over Europe during 2 months (July and August 2009). For O₃, the impact of GEO-TIR is significant (the GEO-TIR assimilation run is closer to the nature run) from 1 to 5 km whereas at the surface the impact of GEO-TIR is low. In general, the impact of the assimilation of GEO-TIR2 is very low for O₃ (GEO-TIR2 assimilation runs are very close to the control runs for all experiments). For CO, the GEO-TIR assimilation runs are mainly closer to the nature run, but the assimilation of GEO-TIR2 also has a positive impact above 1 km. However at the surface, the assimilation of GEO-TIR brings bigger improvement than the assimilation of GEO-TIR2.

We also analyse the impact of the assimilation of GEO-TIR on O₃ AQ scores at the surface. The assimilation of GEO-TIR reduces the percentage of false alarms and increases the percentage of good detections for all experiments although improvement can be slight.

Finally, the results shown in this paper using OSSEs suggest that the assimilation of GEO-TIR into an AQ model can considerably improve the information on O₃ and CO fields in the LmT. However, the OSSE used in this study is based only on the assimilation of profiles and can certainly be improved by assimilating radiances and a much bigger observing system including ground-based stations, sondes, balloons, aircraft, low earth orbit satellites, and other observations. Such a wider study would be very expensive and require a lot of resources but would give a more accurate assessment of the added value of GEO-TIR. Another perspective for the GEO-TIR instrument would be to add channels in the visible (Chappuis bands) as for the MAGEAQ instrument, and to perform an

OSSE for O₃ combining this new instrument with ground-based measurements. It would be very useful to perform further OSSEs to characterize how this combination of satellite and ground-based data could improve AQ monitoring and forecasting.

Acknowledgements. This work was funded by the Centre National de Recherches Scientifiques (CNRS), Astrium-EADS, the Centre National de Recherches Météorologiques (CNRM) of Météo-France and the RTRA/STAE (POGEQA project). The authors acknowledge ETHER, the French atmospheric composition database (CNES and CNRS-INSU) and the Région Midi-Pyrénées (INFOAIR project).

650 References

- Akimoto, H., Irie, H., Kasai, Y., Kanaya, Y., Kita, K., Koike, M., Kondo, Y., Nakazawa, T., and Hayashida, S.: Planning a geostationary atmospheric observation satellite, Commission on the Atmospheric Observation Satellite of the Japan Society of Atmospheric Chemistry, 2008.
- Atlas, R. M.: Observing system simulation experiments: Methodology, examples and limitations, in:
655 CGC/WMO Workshop, Geneva 7-9 April, WMO TD No. 868, W.M.O., Geneva, Switzerland, 1997.
- Bechtold, P., Bazile, E., Guichard, F., Mascart, P., and Richard, E.: A mass-flux convection scheme for regional and global models, *Quart. J. Roy. Meteor. Soc.*, 127, 869–886, 2001.
- Bovensmann, H. and Orphal, J.: The geostationary component of an operational atmospheric chemistry monitoring system: Specification and expected performance, ESA Study: Operational Atmospheric Chemistry
660 Monitoring Missions 17237/03/NL/GS (CAPACITY), 2005.
- Branis, M.: Air quality of Prague: traffic as a main pollution source, *Environ. Monit. Assess.*, 156, 377–390, 2009.
- Brunekreef, B. and Holgate, S. T.: Air pollution and health, *The Lancet*, 360, 1233–1242, 2002.
- Buis, S., Piacentini, A., and Déclat, D.: PALM: a computational framework for assembling high-performance
665 computing applications, *Concurrency Computat.: Pract. Exper.*, 18, 231–245, 2006.
- Burrows, J. P., Bovensmann, H., Bergametti, G., Flaud, J. M., Orphal, J., Noël, S., Monks, P. S., Corlett, G. K., Goede, A. P., von Clarmann, T., Steck, T., Fischer, H., and Friedl-Vallon, F.: The geostationary tropospheric pollution explorer (GeoTROPE) mission: objectives, requirements and mission concept, *Adv. Space Res.*, 34, 682–687, 2004.
- 670 Claeysman, M., Attié, J.-L., El Amraoui, L., Cariolle, D., Peuch, V.-H., Teyssédre, H., Josse, B., Ricaud, P., Massart, S., Piacentini, A., Cammas, J.-P., Livesey, N. J., Pumphrey, H. C., and Edwards, D. P.: A linear CO chemistry parameterization in a chemistry-transport model: evaluation and application to data assimilation, *Atmos. Chem. Phys.*, 10, 6097–6115, 2010a.
- Claeysman, M., Attié, J.-L., Peuch, V.-H., El Amraoui, L., Lahoz, W. A., Josse, B., Ricaud, P., von Clarmann,
675 T., Höpfner, M., Orphal, J., Flaud, J.-M., Edwards, D. P., Chance, K., Liu, X., Pasternak, F., and Cantié, R.: A geostationary thermal infrared sensor to monitor the lowermost troposphere: O₃ and CO retrieval studies, *Atmos. Meas. Tech. Discuss.*, 3, 3489–3534, 2010b.
- Clerbaux, C., Coheur, P.-F., Scharf, O., Hurtmans, D., and Boynard, A.: The potential of MTG-IRS and S4-TIR to detect high pollution events at urban and regional scales, *Eumetsat EUM/CO/07/4600000447/SAT*, 2008a.
- 680 Clerbaux, C., Edwards, D. P., Deeter, M., Emmons, L., Lamarque, J.-F., Tie, X. X., Massie, S. T., and Gille, J.: Carbon monoxide pollution from cities and urban areas observed by the Terra/MOPITT mission, *Geophys. Res. Lett.*, 35, L03817, doi:10.1029/2007GL032300, 2008b.
- Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Turquety, A. R. S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal
685 infrared IASI/MetOp sounder, *Atmos. Chem. Phys.*, 9, 6041–6054, 2009.
- Courtier, P., Freydl, C., Geleyn, J. F., Rabier, F., and Rochas, M.: The ARPEGE project at Météo France, in: *Atmospheric Models*, vol.2, pp. 193–231, Workshop on Numerical Methods, Reading, U.K., 1991.
- Cristofanelli, P., Marinoni, A., Arduini, J., Bonafé, U., Calzolari, F., Colombo, T., Decesari, S., Duchi, R., Facchini, M. C., Fierli, F., Finessi, E., Maione, M., Chiari, M., Calzolari, G., Messina, P., Orlandi, E., Roccato,

- 690 F., and Bonasoni, P.: Significant variations of trace gas composition and aerosol properties at Mt. Cimone during air mass transport from North Africa: contributions from wildfire emissions and mineral dust, *Atmos. Chem. Phys.*, 9, 4603–4619, 2009.
- Deeter, M. N., Edwards, D. P., Gille, J. C., and Drummond, J. R.: Sensitivity of MOPITT observations to carbon monoxide in the lower troposphere, *J. Geophys. Res.*, 112, D24306, doi:10.1029/2007JD008929, 2007.
- 695 Dentener, F., Kinne, S., Bond, T., Boucher, O., Cofala, J., Generoso, S., Ginoux, P., Gong, S., Hoelzemann, J. J., Ito, A., Marelli, L., Penner, J. E., Putaud, J.-P., Textor, C., Schulz, M., van der Werf, G. R., and Wilson, J.: Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom, *Atmos. Chem. Phys.*, 6, 4321–4344, 2006.
- Dufour, A., Amodei, M., Ancellet, G., and Peuch, V.-H.: Observed and modelled "chemical weather" during
700 ESCOMPTE, *Atmos. Res.*, 74, 161–189, 2004.
- Edwards, D. P.: Air Quality Remote Sensing From Space, *Eos Trans. AGU*, 87(33), 327, doi:10.1029/2006EO330005, 2006.
- Edwards, D. P., Petron, G., Novelli, P. C., Emmons, L. K., Gille, J. C., and Drummond, J. R.: Southern Hemisphere carbon monoxide interannual variability observed by Terra/Measurement of Pollution in the
705 Troposphere (MOPITT), *J. Geophys. Res.*, 111, D16303, doi:10.1029/2006JD007079, 2006.
- Edwards, D. P., Arellano Jr., A. F., and Deeter, M. N.: A satellite observation system simulation experiment for carbon monoxide in the lowermost troposphere, *J. Geophys. Res.*, 114, D14304, doi:10.1029/2008JD011375, 2009.
- El Amraoui, L., Peuch, V.-H., Ricaud, P., Massart, S., Semane, N., Teyssèdre, H., Cariolle, D., and Karcher,
710 F.: Ozone loss in the 2002–2003 Arctic vortex deduced from the assimilation of Odin/SMR O₃ and N₂O measurements: N₂O as a dynamical tracer, *Quart. J. Roy. Meteor. Soc.*, 134, 217–228, 2008a.
- El Amraoui, L., Semane, N., Peuch, V.-H., and Santee, M. L.: Investigation of dynamical processes in the polar stratospheric vortex during the unusually cold winter 2004/2005, *Geophys. Res. Lett.*, 35, L03803, doi:10.1029/2007GL031251, 2008b.
- 715 Eremenko, M., Dufour, G., Foret, G., Keim, C., Orphal, J., Beekmann, M., Bergametti, G., and Flaud, J.-M.: Tropospheric ozone distributions over Europe during the heat wave in July 2007 observed from infrared nadir spectra recorded by IASI, *Geophys. Res. Lett.*, 35, L18805, doi:10.1029/2008GL03803, 2008.
- Fisher, M. and Andersson, E.: Developments in 4D-Var and Kalman filtering, in: *Technical Memorandum Research Department*, vol. 347, ECMWF, Reading, UK, 2001.
- 720 Fishman, J., Al-Saadi, J. A., Creilson, J. K., Bowman, K. W., Burrows, J. P., Richter, A., Chance, K. V., Edwards, D. P., Martin, R. V., Morris, G. A., Pierce, R. B., Ziemke, J. R., Schaack, T. K., and Thompson, A. M.: Remote Sensing of Tropospheric Pollution from Space, *Bull. Amer. Meteor. Soc.*, 89, 6, 805–821, 2008.
- Flaud, J. M., Orphal, J., Bergametti, G., Deniel, C., von Clarmann, T., Friedl-Vallon, F., Steck, T., Fischer,
725 H., Bovensmann, H., Burrows, J. P., Carlotti, M., Ridolfi, M., and Palchetti, L.: The Geostationary Fourier Imaging Spectrometer (GeoFIS) as part of the Geostationary Tropospheric Pollution Explorer (GeoTroPE) mission: objectives and capabilities, *Adv. Space Res.*, 34, 688–693, 2004.
- Foret, G., Hamaoui, L., Schmechtig, C., Eremenko, M., Keim, C., Dufour, G., Boynard, A., Coman, A., Ung, A., and Beekmann, M.: Evaluating the potential of IASI ozone observations to constrain simulated surface

- 730 ozone concentrations, *Atmos. Chem. Phys.*, 9, 8479–8491, 2009.
- Geer, A. J., Lahoz, W. A., Bekki, S., Bormann, N., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Juckes, M. N., Massart, S., Peuch, V.-H., Rharmili, S., and Segers, A.: The ASSET intercomparison of ozone analyses: Method and first results, *Atmos. Chem. Phys.*, 6, 5445–5474, 2006.
- Hollingsworth, A., Engelen, R., Textor, C., Boucher, O., Chevallier, F., Dethof, A., Elbern, H., Eskes, H., 735 Flemming, J., Granier, C., Kaiser, J. W., Morcrette, J. J., Rayner, P., Peuch, V.-H., Rouil, L., and the GEMS Consortium: Toward a monitoring and forecasting system for atmospheric composition: The GEMS Project, *Bull. Amer. Meteor. Soc.*, 89, 8, 1147–1164, 2008.
- Honoré, C., Rouil, L., Vautard, R., Beekmann, M., Bessagne, B., Dufour, A., Elichegaray, C., Flaud, J.-M., Malherbe, L., Meleux, F., Menut, L., Martin, D., Peuch, A., Peuch, V.-H., and Poisson, N.: Predictability of 740 European air quality: Assessment of 3 years of operational forecasts and analyses by the PREV’AIR system, *J. Geophys. Res.*, 113, D04301, doi:10.1029/2007JD008761, 2008.
- Höpfner, M., Stiller, G. P., Kuntz, M., von Clarmann, T., Echle, G., Funke, B., Glatthor, N., Hase, F., Kemnitzer, H., and Zorn, S.: The Karlsruhe optimized and precise radiative transfer algorithm. Part II: Interface to retrieval applications, in: *Optical Remote Sensing of the Atmosphere and Clouds*, Beijing, China, 15–17 745 September 1998, edited by Wang, J., Wu, B., Ogawa, T., and Guan, Z., vol. 3501, pp. 186–195, 1998.
- Ignaccolo, R., Ghigo, S., and Giovenali, E.: Analysis of air quality monitoring networks by functional clustering, *Environmetrics*, 19, 672–686, 2008.
- Jacob, D. J.: Heterogeneous chemistry and tropospheric ozone, *Atmos.-Environ.*, 34, 2131–2159, 2000.
- Josse, B., Simon, P., and Peuch, V.-H.: Radon global simulation with the multiscale chemistry transport model 750 MOCAGE, *Tellus*, 56, 339–356, 2004.
- Krzyzanowski, M. and Cohen, A.: Update of WHO air quality guidelines, *Air. Qual. Atmos. Health*, 1, 7–13, doi:10.1007/s11869-008-0008-9, 2008.
- Lahoz, W. A., Brugge, R., Jackson, D. R., Migliorini, S., Swinbank, R., Lary, D., and Lee, A.: An observing system simulation experiment to evaluate the scientific merit of wind and ozone measurements from the 755 future SWIFT instrument, *Quart. J. Roy. Meteor. Soc.*, 131, 503–523, 2005.
- Lahoz, W. A., Errera, Q., Swinbank, R., and Fonteyn, D.: Data assimilation of stratospheric constituents: a review, *Atmos. Chem. Phys.*, 7, 5745–5773, 2007.
- Lamarque, J.-F., Hess, P., Emmons, L., Buja, L., Washington, W., and Granier, C.: Tropospheric ozone evolution between 1890 and 1990, *J. Geophys. Res.*, 110, D08304, doi:10.1029/2004JD005537, 2005.
- 760 Lefèvre, F., Brasseur, G. P., Folkins, I., Smith, A. K., and Simon, P.: Chemistry of the 1991–1992 stratospheric winter: three dimensional model simulations, *J. Geophys. Res.*, 99, 8183–8195, 1994.
- Louis, J.-F.: Parametric model of vertical eddy fluxes in the atmosphere, *Boundary Layer Meteorol.*, 17, 187–202, 1979.
- Martin, R. V.: Satellite remote sensing of surface air quality, *Atmos.-Environ.*, 42, 7823–7843, 2008.
- 765 Massart, S., Cariolle, D., and Peuch, V.-H.: Towards an improvement of the atmospheric ozone distribution and variability by assimilation of satellite data, *C. R. Geosci.*, 337, 1305–1310, 2005.
- Massart, S., Clerbaux, C., Cariolle, D., Piacentini, A., Turquety, S., and Hadji-Lazaro, J.: First steps towards the assimilation of IASI ozone data into the MOCAGE-PALM system, *Atmos. Chem. Phys.*, 9, 5073–5091, 2009.

- 770 Masutani, M., Schlatter, T. W., Errico, R. M., Stoffelen, A., Andersson, E., Lahoz, W. A., Woollen, J. S.,
David Emmit, G., Riishøjgaard, L.-P., and Lord, S. J.: Observing System Simulation Experiments, in *Data
Assimilation: Making Sense of Observations*, W.A. Lahoz, B. Khatattov, R. Ménard, (Eds.), Springer, July,
2010a.
- Masutani, M., Woollen, J. S., Lord, S. J., Emmitt, G. D., Kleespies, T. J., Wood, S. A., Greco, S., Sun, H., Terry,
775 J., Kapoor, V., Treadon, R., and Campana, K. A.: Observing system simulation experiments at the National
Centers for Environmental Prediction, *J. Geophys. Res.*, 115, D07101, doi:10.1029/2009JD012528, 2010b.
- McKeen, S., Wilczak, J., Grell, G., Djalalova, I., Peckham, S., Hsie, E.-Y., Gong, W., Bouchet, V., Menard, S.,
Moffet, R., McHenry, J., McQueen, J., Tang, Y., Carmichael, G. R., Pagowski, M., Chan, A., Dye, T., Lee,
G. F. P., and Mathur, R.: Assessment of an ensemble of seven real-time ozone forecasts over eastern North
780 America during the summer of 2004, *J. Geophys. Res.*, 110, D21307, doi:10.1029/2005JD005858, 2005.
- Menut, L. and Bessagnet, B.: Atmospheric composition forecasting in Europe, *Atmos.-Environ.*, 28, 61–74,
2010.
- National Research Council: *Earth Science and Applications from Space: National Imperatives for the Next
Decade and Beyond*, in: the National Academy of Sciences, Washington, D. C, USA, 2007.
- 785 Oltmans, S. J., Lefohn, A. S., Harris, J. M., Galbally, I., Scheel, H. E., Bodeker, G., Brunke, E., Claude, H.,
Tarasick, D., Johnson, B. J., Simmonds, P., Shadwick, D., Anlauf, K., Hayden, K., Schmidlin, F., Fujimoto,
T., Akagi, K., Meyer, C., Nichol, S., Davies, J., Redondas, A., and Cuevas, E.: Long-term changes in
tropospheric ozone, *Atmos.-Environ.*, 40, 3156–3173, 2006.
- Orphal, J., Bergametti, G., Beghin, B., Hébert, J.-P., Steck, T., and Flaud, J.-M.: Monitoring tropospheric pol-
790 lution using infrared spectroscopy from geostationary orbit, *Comptes Rendus Physique*, 6, 888–896, 2005.
- Pagowski, M., Grell, G. A., Devenyi, D., Peckham, S. E., McKeen, S. A., Gong, W., Delle Monache, L.,
McHenry, J. N., McQueen, J., and Lee, P.: Application of dynamic linear regression to improve the skill of
ensemble-based deterministic ozone forecasts, *Atmos.-Environ.*, 40, 3240–3250, 2006.
- Peuch, V.-H., Amodei, M., Barthet, T., Cathala, M.-L., Michou, M., and Simon, P.: MOCAGE, MODéle de
795 Chimie Atmosphérique à Grande Echelle, in: *Proceedings of Météo France: Workshop on atmospheric
modelling*, pp. 33–36, Toulouse, France, 1999.
- Peuch, V.-H., Attié, J.-L., Claeysman, M., Amraoui, L. E., Ricaud, P., Semane, N., Massart, S., Piacentini, A.,
Cariolle, D., Flaud, J.-M., Bergametti, G., Cantie, R., Pasternak, F., Lehors, L., von Clarmann, T., Höpfner,
M., and Orphal, J.: Data assimilation experiments within the POGEQA project, in: *American Geophysical
800 Union, Fall Meeting*, San Francisco, California, USA, December, 2009.
- Peuch, V.-H., Orphal, J., and the MAGEAQ consortium: MAGEAQ: Monitoring the Atmosphere from Geosta-
tionary orbit for European Air Quality, Proposal submitted to ESA for Earth Explorer Opportunity Mission
EE-8, 2010.
- Phillips, D.: A technique for the numerical solution of certain integral equations of first kind, *J. Ass. Comput.*
805 *Mat.*, 9, 34–97, 1962.
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, Series on Atmospheric,
Oceanic and Planetary Physics, vol. 2, World Scientific, 2000.
- Rouïl, L., Honoré, C., Vautard, R., Beekmann, M., Bessagnet, B., Malherbe, L., Meleux, F., Dufour, A.,
Elichegaray, C., Flaud, J.-M., Menut, L., Martin, D., Peuch, A., Peuch, V.-H., and Poisson, N.: An Opera-

- 810 tional Forecasting and Mapping System for Air Quality in Europe, *Bull. Amer. Meteor. Soc.*, 90, 1, 73–83, 2008.
- Seinfeld, J. H. and Pandis, S. N.: *Atmospheric Chemistry and Physics, from Air Pollution to Climate Change*, John Wiley & Sons, 1997.
- 815 Stiller, G. P., von Clarmann, T., Funke, B., Glatthor, N., Hase, F., Höpfner, M., and Linden, A.: Sensitivity of trace gas abundances retrievals from infrared limb emission spectra to simplifying approximations in radiative transfer modelling, *J. Quant. Radiat. Transfer*, 72 (3), 249–280, 2002.
- Stockwell, W. R., Kirchner, F., Khun, M., and Seefeld, S.: A new mechanism for regional atmospheric chemistry modelling, *J. Geophys. Res.*, 102, 25 847–25 879, 1997.
- 820 Stoffelen, A., Marseille, G. J., Bouttier, F., Vasiljevic, D., Haan, S. D., and Cardinali, C.: ADM–Aeolus Doppler wind lidar observing system simulation experiment, *Quart. J. Roy. Meteor. Soc.*, 619, 1927–1948, 2006.
- Stuhlmann, R., Rodriguez, A., Tjemkes, S., Grandell, J., Arriaga, A., Bézy, J.-L., Aminou, D., and Bensi, P.: Plans for EUMETSAT’s Third Generation Meteosat (MTG) Geostationary Satellite Program., *Adv. Space Res.*, 36, 975–981, 2005.
- 825 Tikhonov, A.: On the solution of incorrectly stated problems and a method of regularization, *Dokl. Acad. Nauk SSSR*, 151, 501–504, 1963.
- Timmermans, R. M. A., Segers, A. J., Builtjes, P. J. H., Vautard, R., Siddans, R., Elbern, H., Tjemkes, S. A. T., and Schaap, M.: The added value of a proposed satellite imager for ground level particulate matter analyses and forecasts, *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, 2, 4, 271–283, 2009.
- 830 Turquety, S., Hurtmans, D., Hadji-Lazaro, J., Coheur, P.-F., Clerbaux, C., Josset, D., and Tsamalis, C.: Tracking the emission and transport of pollution from wildfires using the IASI CO retrievals: analysis of the summer 2007 Greek fires, *Atmos. Chem. Phys.*, 9, 4897–4913, 2009.
- Vaughan, J., Lamb, B., Frei, C., Wilson, R., Bowman, C., Figueroa-Kaminsky, C., Otterson, S., Boyer, M., Mass, C., Albright, M., Koenig, J., Collingwood, A., Gilroy, M., and Maykut, N.: A Numerical Daily Air Quality Forecast System for The Pacific Northwest, *Bull. Amer. Meteor. Soc.*, 85, 4, 549–561, 2004.
- 835 Visschedijk, A. H. J. and Denier van der Gon, H. A.: Gridded European anthropogenic emission data for NO_x , SO_2 , NMVOC, NH_3 , CO, PM_{10} , $\text{PM}_{2.5}$ and CH_4 for the year 2000, TNO B& O-A Rapport 2005/106, 2005.
- World Health Organization: *Air Quality Guidelines: Global Update 2005*, Tech. rep., 2005.

Table 1. Correlation coefficient, standard deviation (%) and bias (%) between observations generated with the look-up tables and observations generated with KOPRA-fit, calculated with respect to observations generated with KOPRA-fit for O₃ (1st and 2nd columns) and CO (3rd and 4th columns) and for the configurations of GEO-TIR (2nd and 4th columns) and GEO-TIR2 (1st and 3rd columns). These statistics have been calculated for data with altitudes between the surface and 10 km.

	GEO-TIR O ₃	GEO-TIR2 O ₃	GEO-TIR CO	GEO-TIR2 CO
Corr. Coeff.	0.99	0.99	0.99	0.93
Stdev (%)	4.4	2.3	1.7	4.8
Bias (%)	-0.8	1.3	0.1	-0.5

Table 2. Description of the different experiments. The NR experiment represents the true atmosphere or the nature run. The experiments with a change in the meteorology, emissions and initial condition are referred as EXP1, EXP2 and EXP3, respectively. The experiments with a change in the meteorology and emission and initial condition are referred as EXP4. The a, b and c extensions represent the experiments with no assimilation, with assimilation of GEO-TIR2 and with assimilation of GEO-TIR, respectively. See text for further details.

Experiment	Meteorology	emissions	Initial condition	Assim
NR	ARPEGE analysis	GEMS	free run	No
EXP1a	ARPEGE forecast 48h	GEMS	free run	No
EXP1b	ARPEGE forecast 48h	GEMS	free run	GEO-TIR2
EXP1c	ARPEGE forecast 48h	GEMS	free run	GEO-TIR
EXP2a	ARPEGE analysis	GLOBAL	free run	No
EXP2b	ARPEGE analysis	GLOBAL	free run	GEO-TIR2
EXP2c	ARPEGE analysis	GLOBAL	free run	GEO-TIR
EXP3a	ARPEGE analysis	GEMS	changed every week	No
EXP3b	ARPEGE analysis	GEMS	changed every week	GEO-TIR2
EXP3c	ARPEGE analysis	GEMS	changed every week	GEO-TIR
EXP4a	ARPEGE forecast 48h	GLOBAL	changed every week	No
EXP4b	ARPEGE forecast 48h	GLOBAL	changed every week	GEO-TIR2
EXP4c	ARPEGE forecast 48h	GLOBAL	changed every week	GEO-TIR

Table 3. Correlation, bias and RMS in % calculated for ozone and CO LmT column between the nature run and the control run for the 4 experiments averaged over 2 months.

	Ozone			CO		
Experiment	Corr.	Bias (%)	RMS (%)	Corr.	Bias (%)	RMS (%)
EXP1	0.793	0.19	10.42	0.780	-1.02	6.78
EXP2	0.935	8.60	5.31	0.919	8.46	5.22
EXP3	0.693	2.07	12.98	0.693	1.73	8.13
EXP4	0.528	7.78	17.27	0.545	7.11	11.41

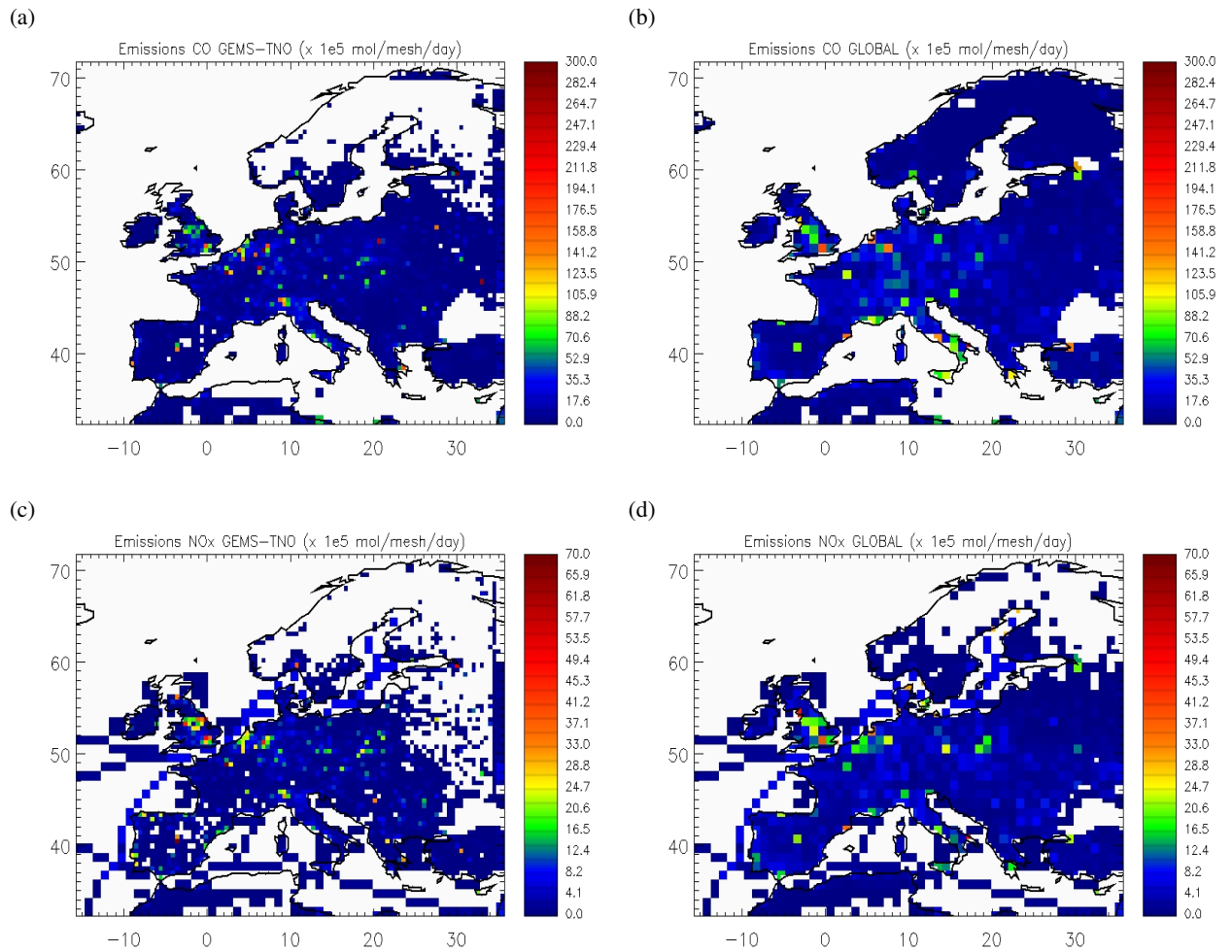


Fig. 1. Total CO emitted by day and by model mesh (0.5 °) over Europe on July 6th, 2009 with GEMS-TNO emission inventory (a) and global emission inventory (b). (c) and (d) are as (a) and (b) but for NO_x=NO+NO₂.

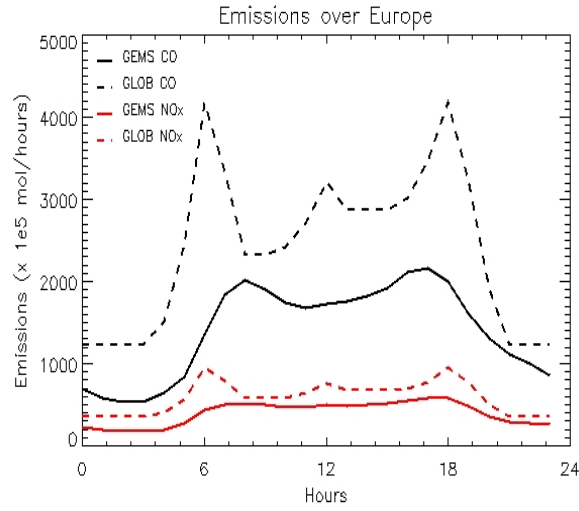


Fig. 2. Diurnal cycle of CO (black line) and ozone (red line) total emissions over Europe for the GEMS-TNO emission inventory (solid line) and for the global emission inventory (dashed line) for the July 6th, 2009.

Table 4. Correlation coefficients (2nd column), biases (3rd column) and root-mean square (4th column) in $\mu g.m^{-3}$ and in % (in brackets) between ground based station observations and MOCAGE nature run for France from July 1st, 2009 to August 31th, 2009 computed on an hourly mean basis.

Species	Corr.	Bias	RMS
Ozone	0.76	12.0 (~18 %)	18.2 (~26 %)
CO	0.63	19.9 (~17 %)	59.9 (~52 %)

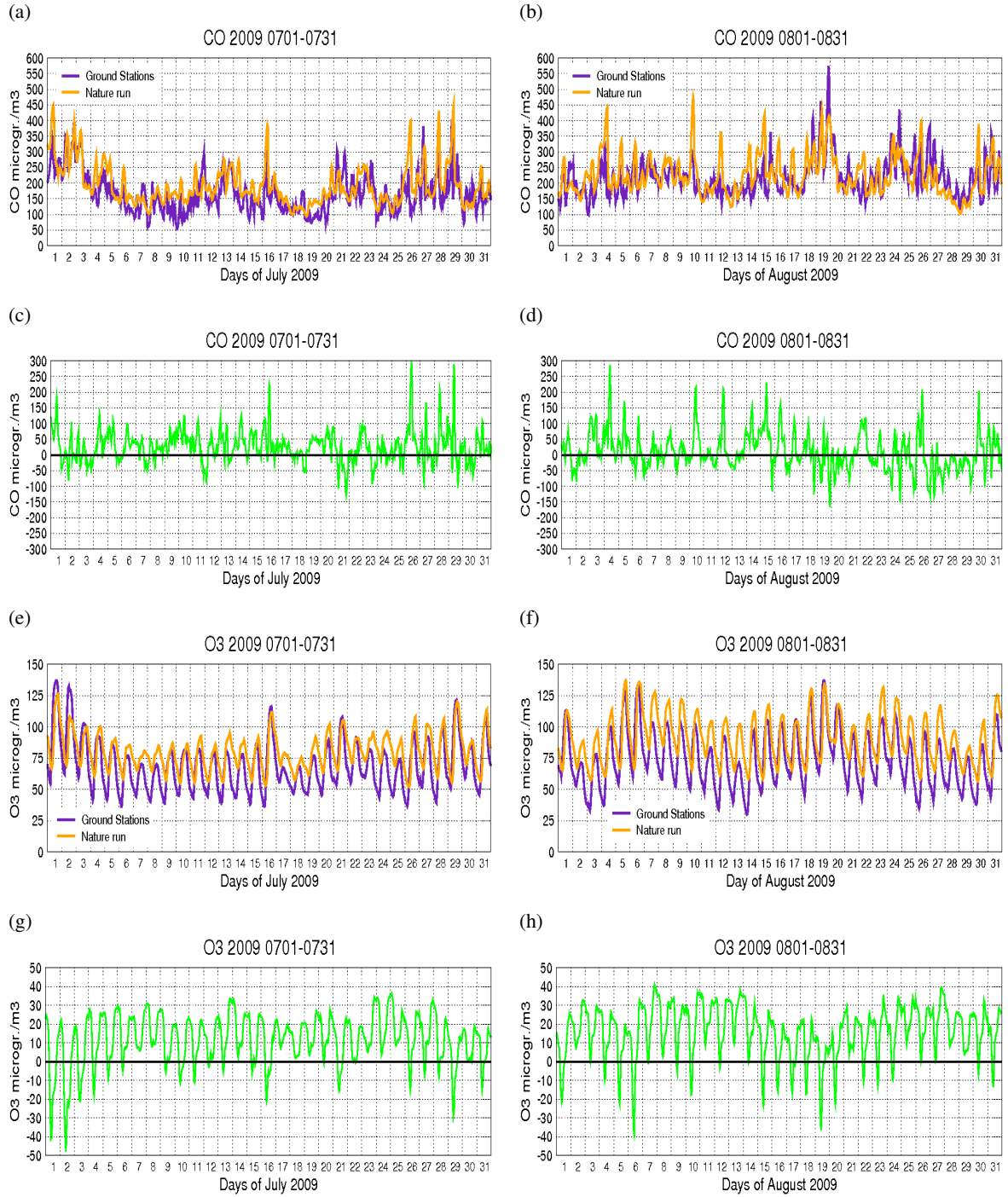


Fig. 3. Evolution of the CO concentration from the nature run (orange) and measured by ground based stations (purple), averaged each hour over France in July 2009 (a) and August 2009 (b) and respective differences between the nature run and the surface observations (c and d). e, f, g and h are as a, b, c and d but for O₃. For CO, all types of ground based stations are considered because of their limited numbers, whereas for O₃ only rural ground stations are considered in order to be closer to the model resolution of $0.5^{\circ} \times 0.5^{\circ}$.

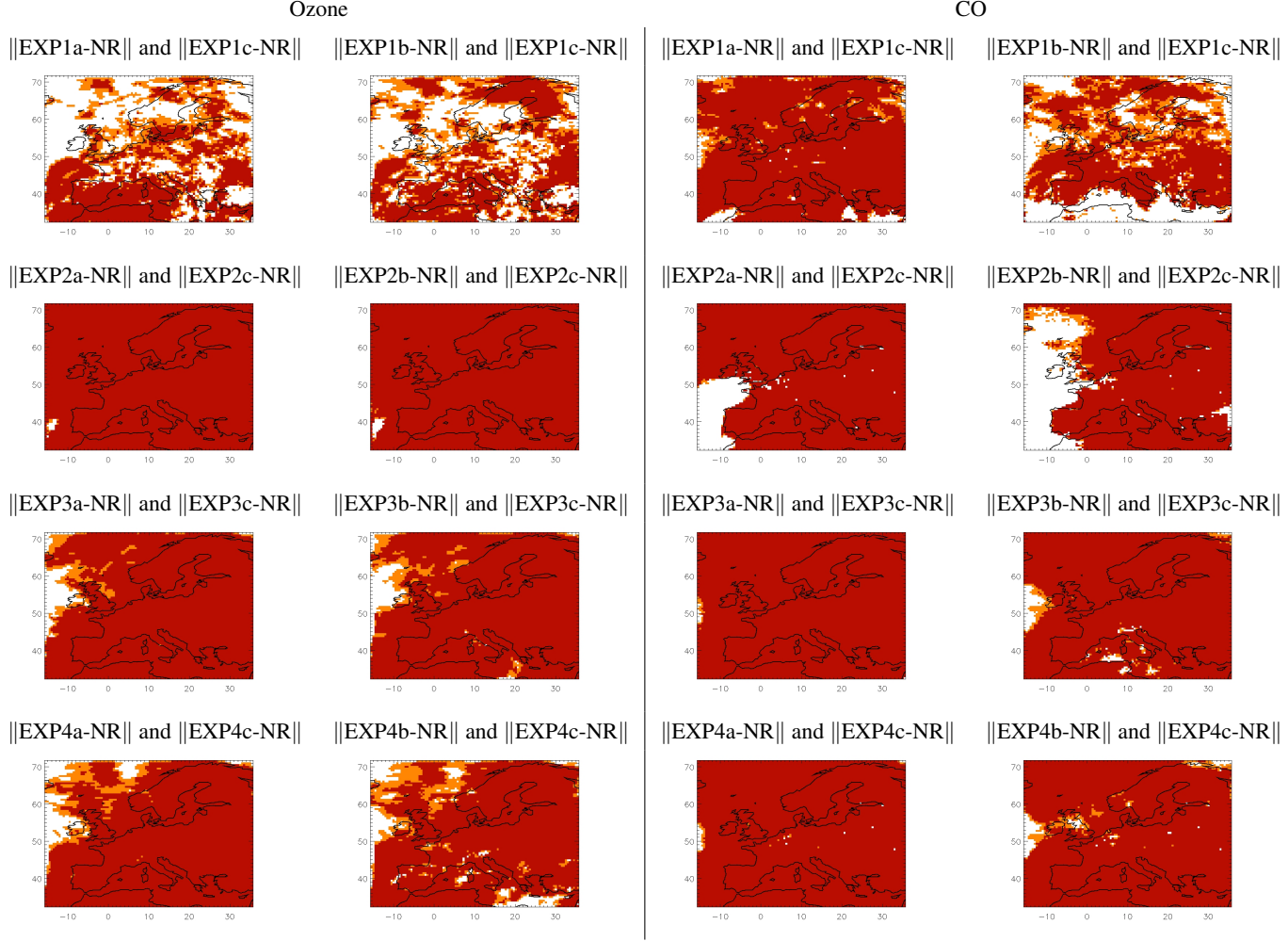


Fig. 4. Z-test where $\|EXP*a-NR\|$ and $\|EXP*c-NR\|$ (1^{st} and 3^{rd} columns); and $\|EXP*b-NR\|$ and $\|EXP*c-NR\|$ (2^{st} and 4^{th} columns) are different at the 0.95 confidence level (orange and red) and 0.99 confidence level (red). Double vertical bars indicate absolute value and * indicates experiment 1, 2, 3 or 4. The 1^{st} and 2^{nd} columns concern the O_3 LmT column and the 3^{rd} and 4^{th} columns concern the O_3 LmT column. The 1^{st} row is for EXP1 (change in the meteorology), the 2^{nd} row is for EXP2 (change in the emissions), the 3^{rd} row is for EXP3 (change in the initial condition) and the 4^{th} row is for EXP4 (change in the meteorology, in the emissions and in the initial condition). See text for further details.

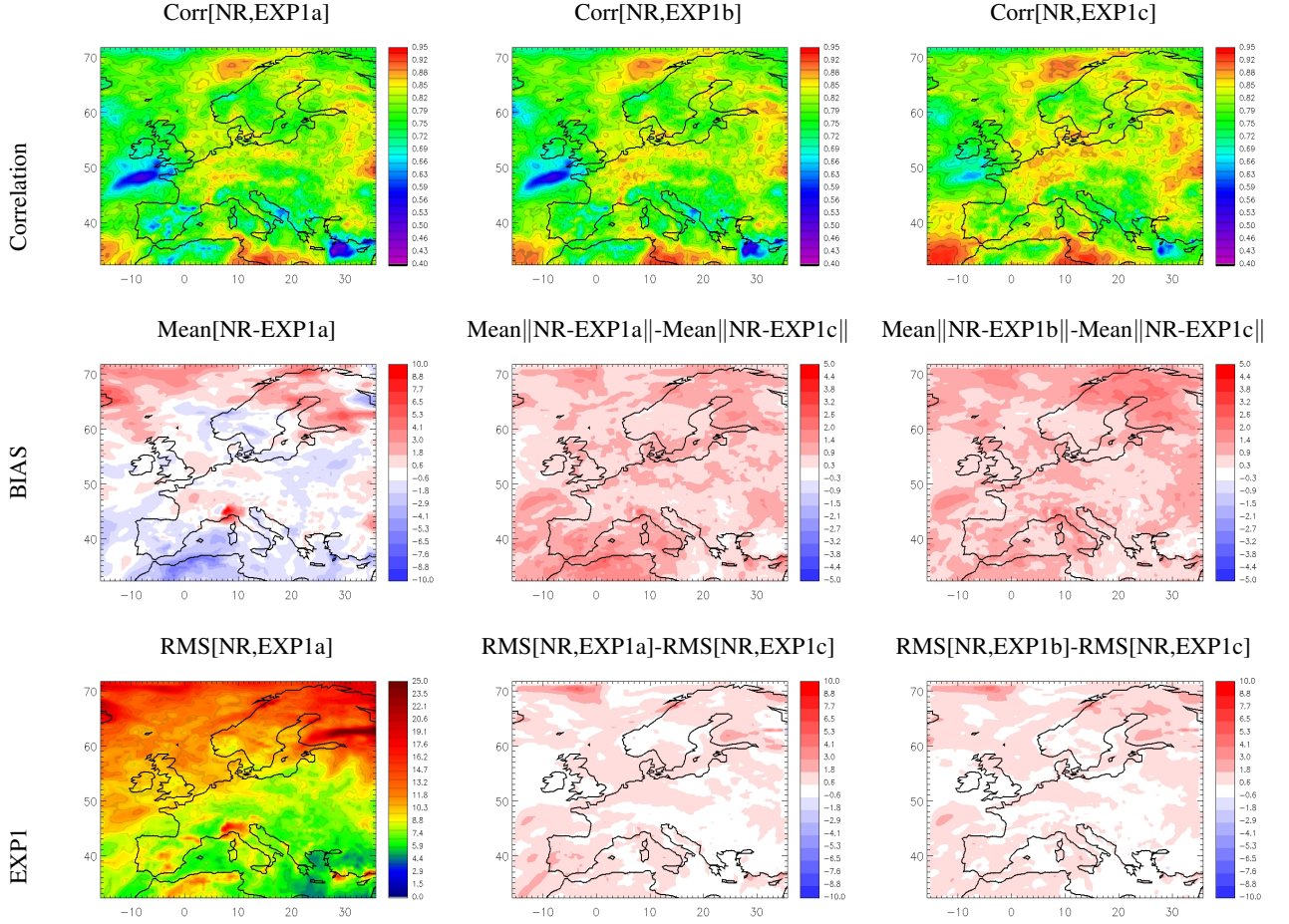


Fig. 5. Correlations (upper panel), absolute difference (middle panel) and root-mean square (lower panel) between the nature run (NR) and the control run (EXP1a) for the O₃ LmT column (1st column) for the experiment with a change in the meteorology (EXP1). For the 2nd and 3rd rows: (i) the 2nd column shows the difference between the nature run and the assimilation of GEO-TIR (EXP1c); and between the nature run and EXP1a; (ii) the 3rd column shows the difference between the nature run and EXP1c; and between the nature run and the assimilation of GEO-TIR2 (EXP1b). Red colours indicate that the assimilation of GEO-TIR improves the correlation (1st column) and reduces the absolute difference (2nd row) or the RMS (3rd row) whereas blue colours indicate a deterioration from using GEO-TIR.

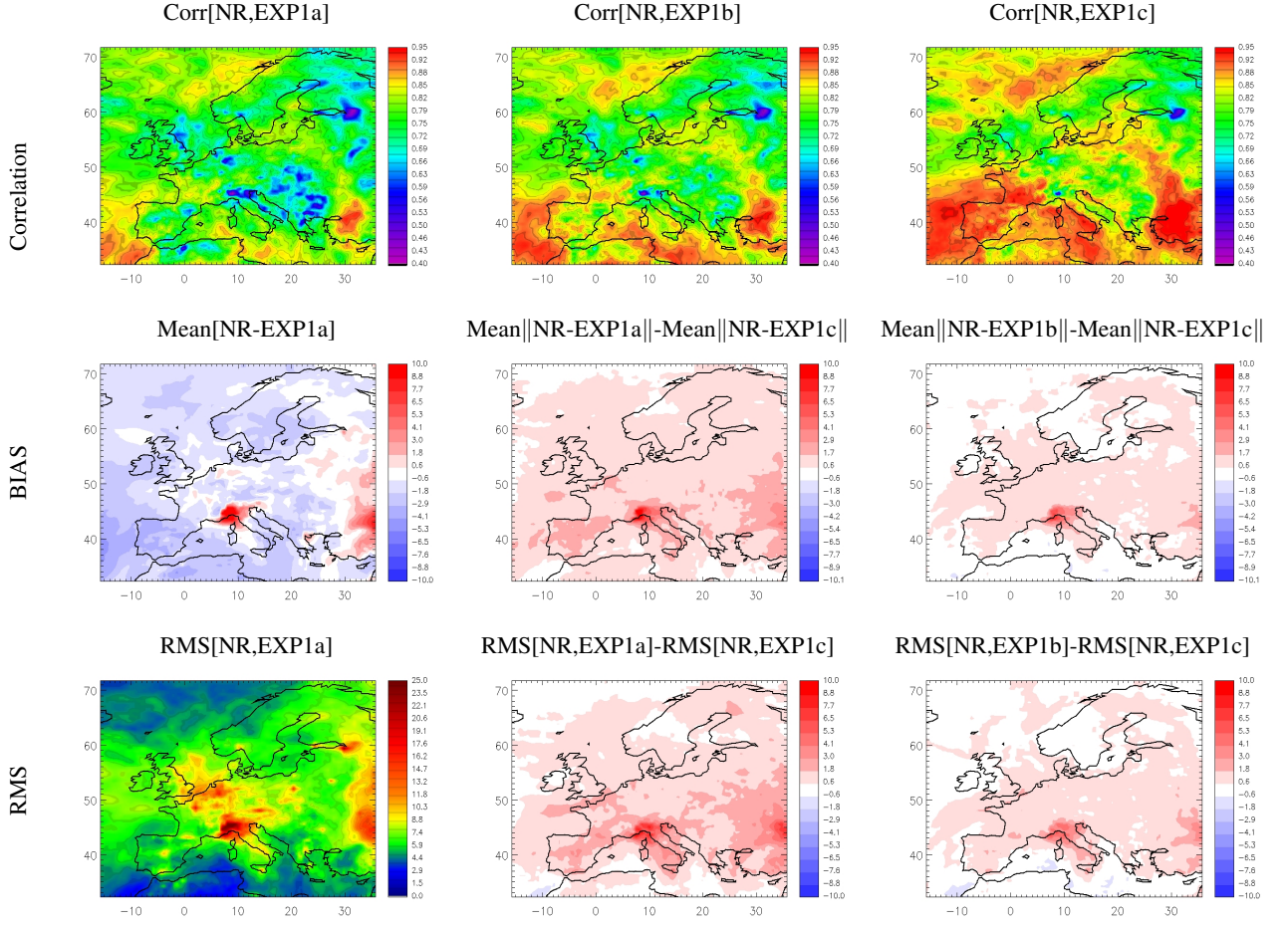


Fig. 6. Same as Figure 5 (EXP1) but for CO.

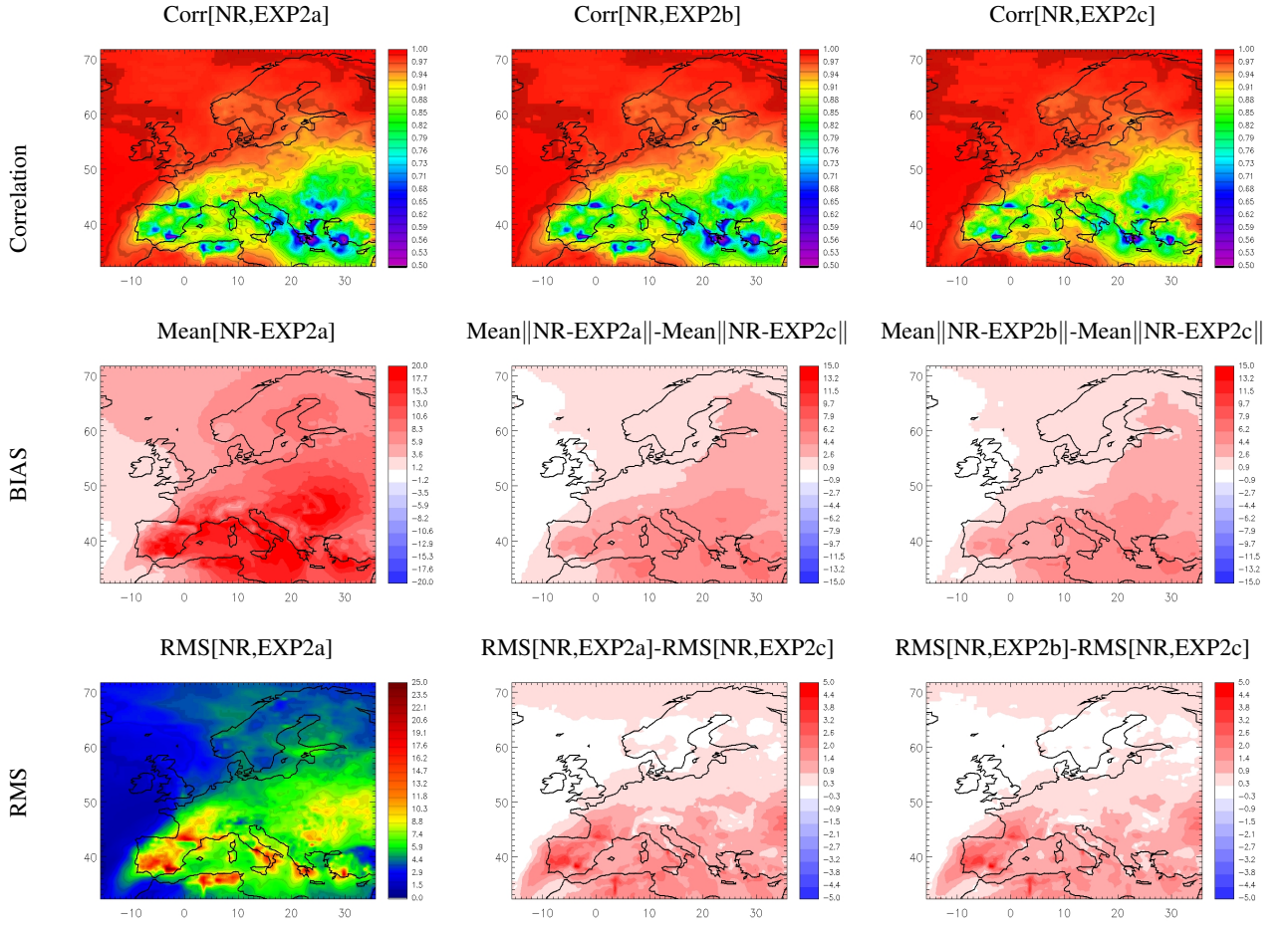


Fig. 7. Same as Figure 5 but for the experiment with a change in the emissions (EXP2).

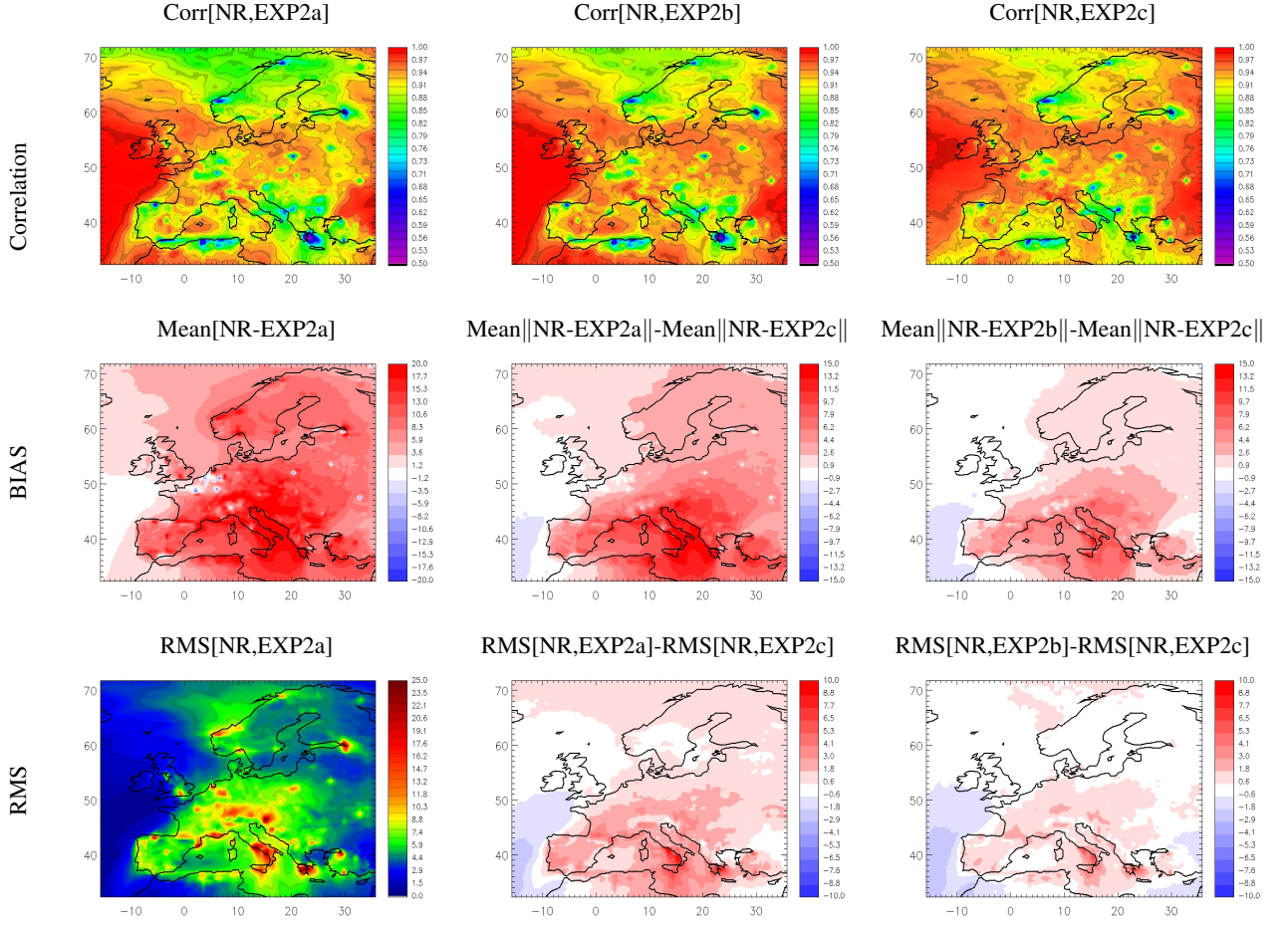


Fig. 8. Same as Figure 7 (EXP2) but for CO.

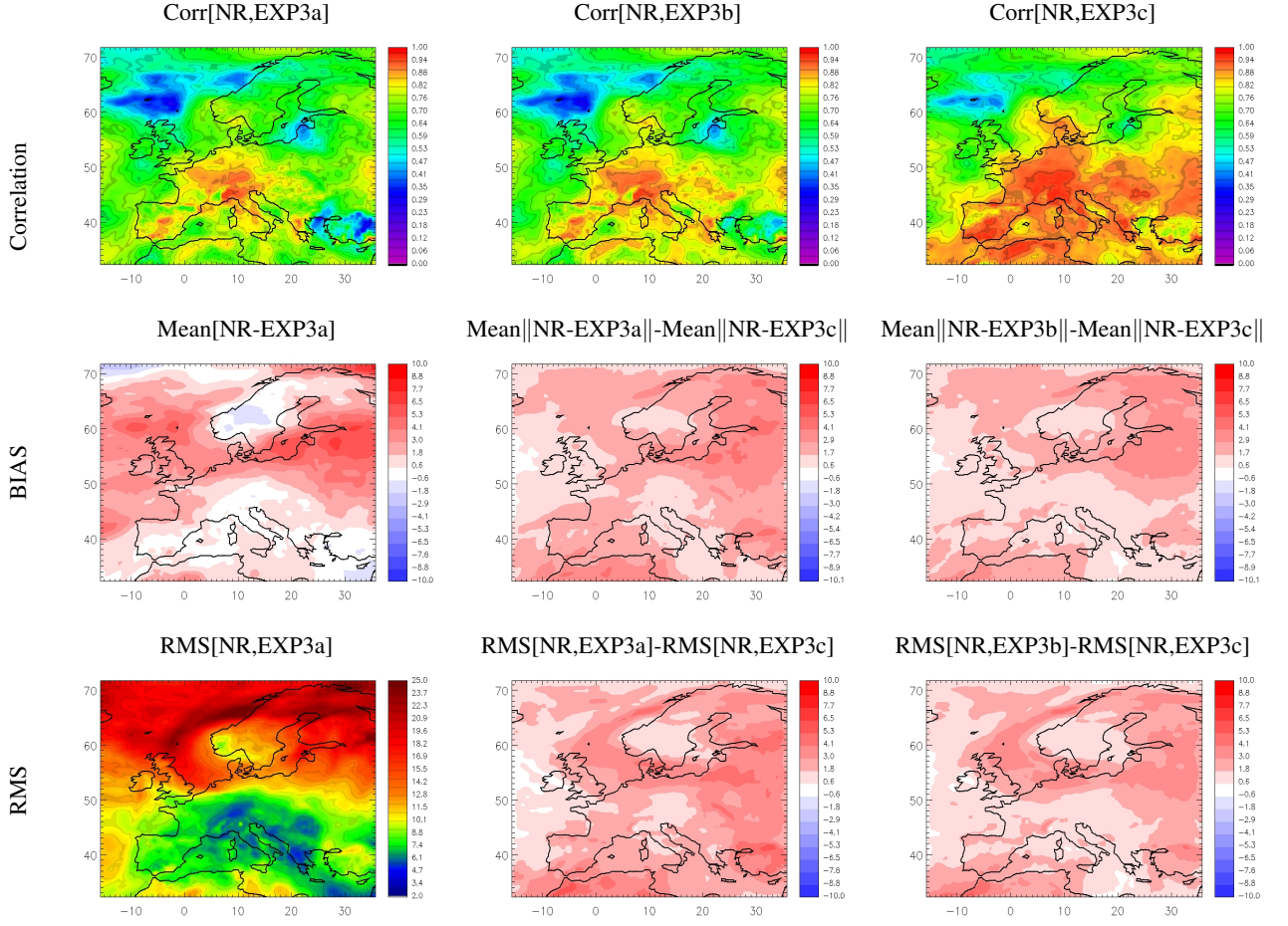


Fig. 9. Same as Figure 5 but for the experiment with a change in the initial condition (EXP3).

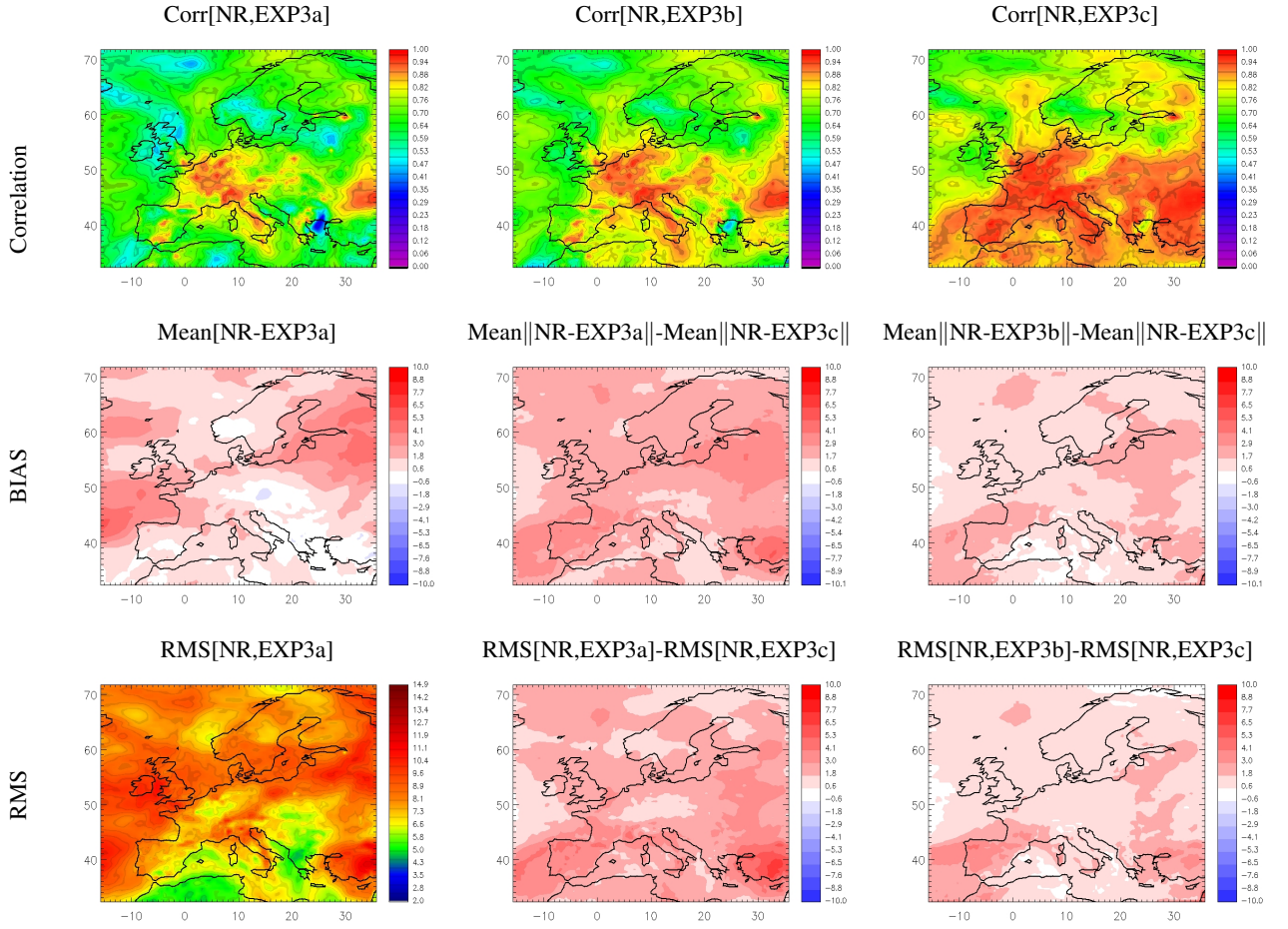


Fig. 10. Same as Figure 9 (EXP3) but for CO.

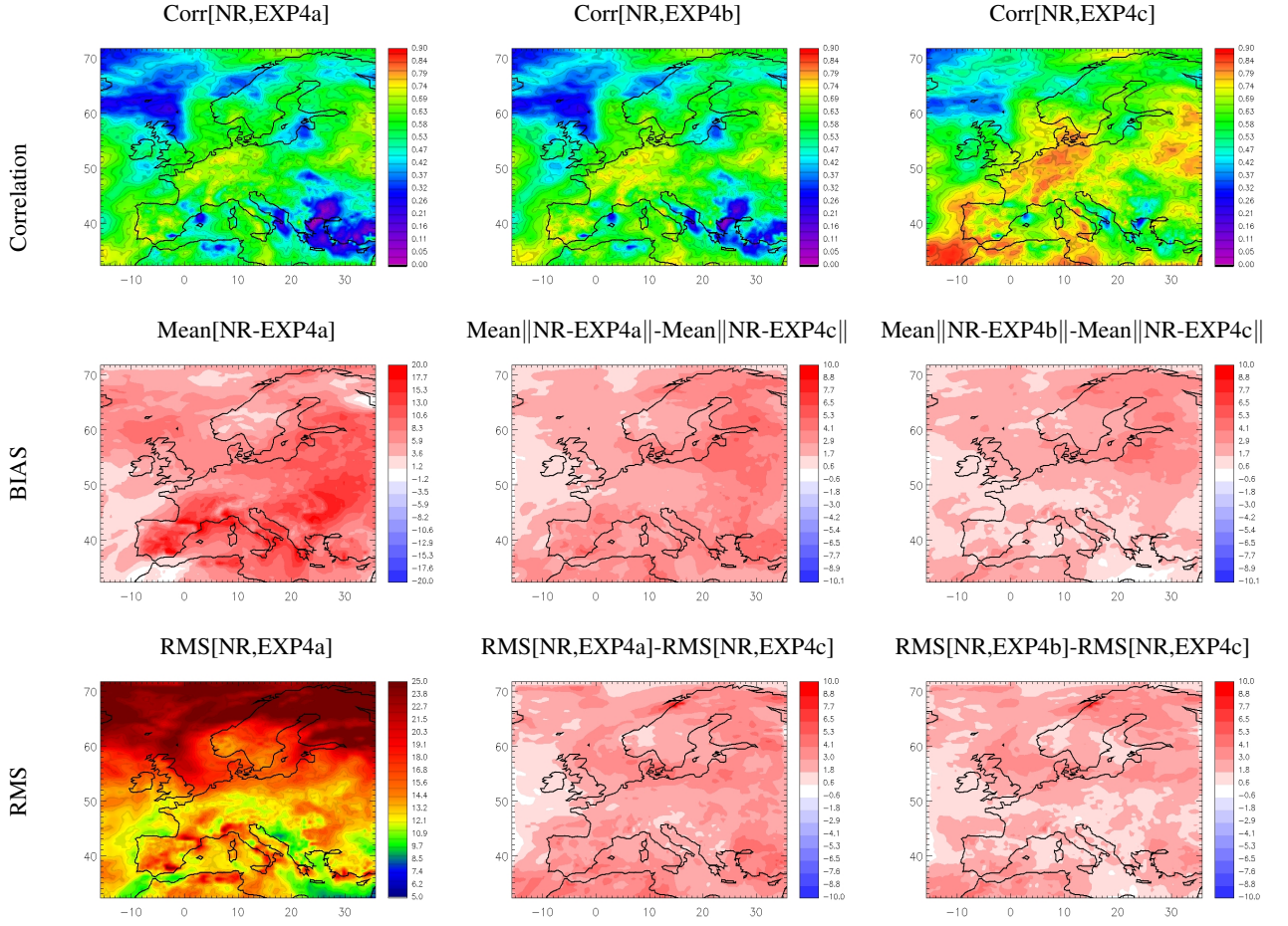


Fig. 11. Same as Figure 5 but for the experiment with a change in the emissions, meteorology and initial condition (EXP4).

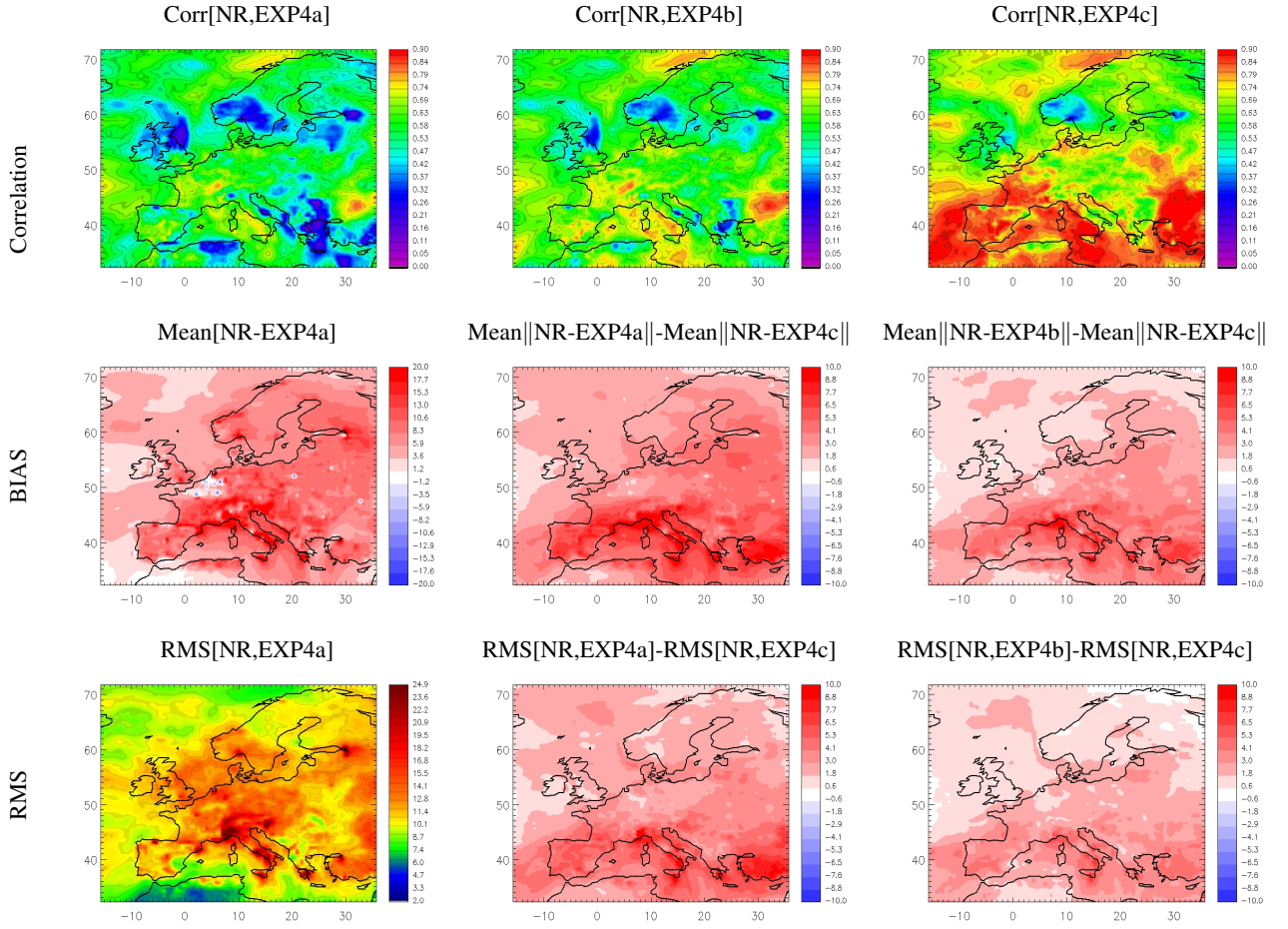


Fig. 12. Same as Figure 11 (EXP4) but for CO.

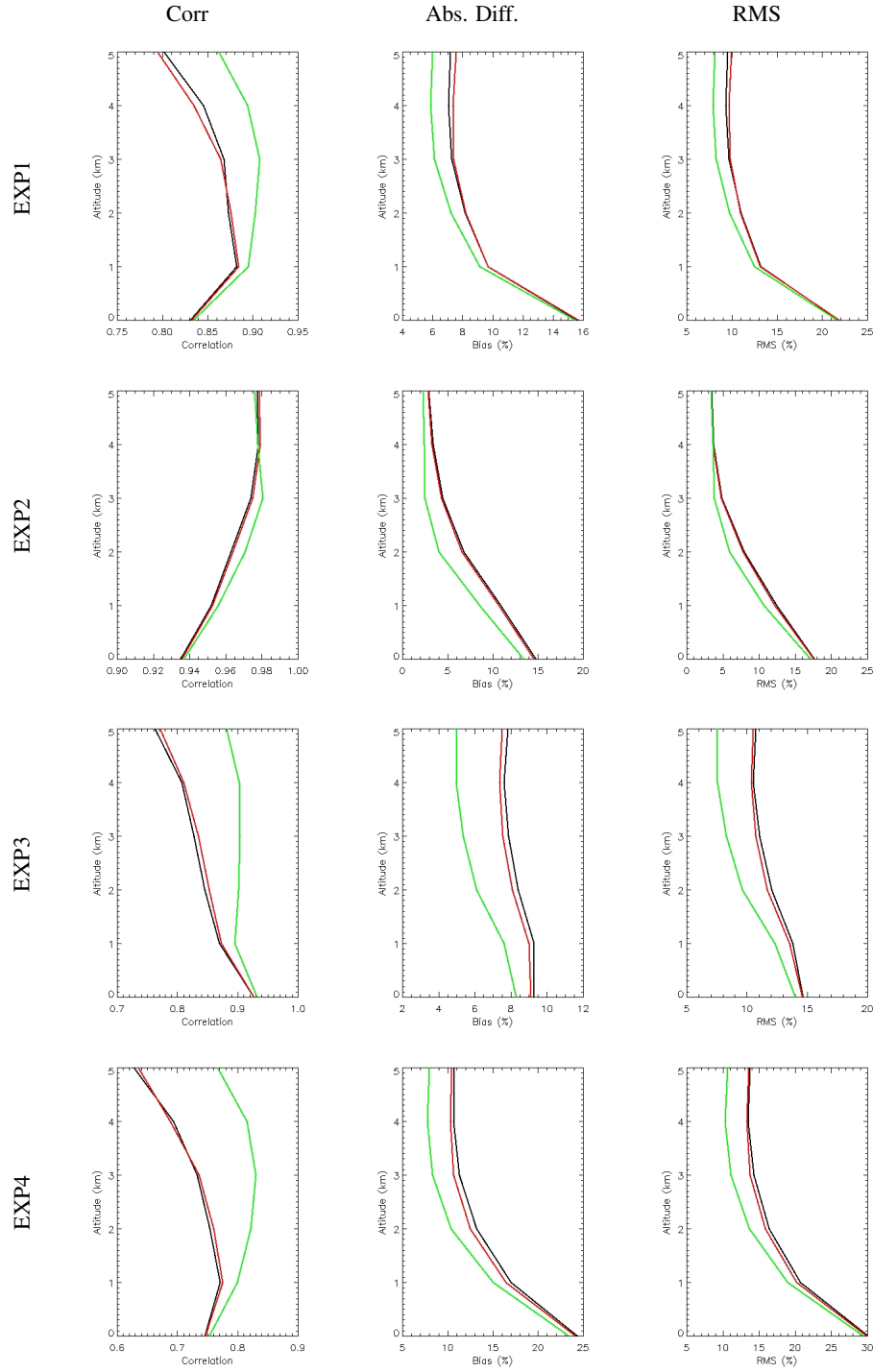


Fig. 13. Correlation (left), absolute difference in % (middle) and RMS difference in % (right) between the nature run (NR) and the control run (black); between the nature run and the assimilation run of GEO-TIR2 (red) and between the nature run and the assimilation run of GEO-TIR (green). Percentages are with respect to the nature run. The 1st row is for EXP1 (change in the meteorology), the 2nd row is for EXP2 (change in the emissions), the 3rd row is for EXP3 (change in the initial condition) and the 4rd row is for EXP4 (change in the meteorology, in the emissions and in the initial condition).

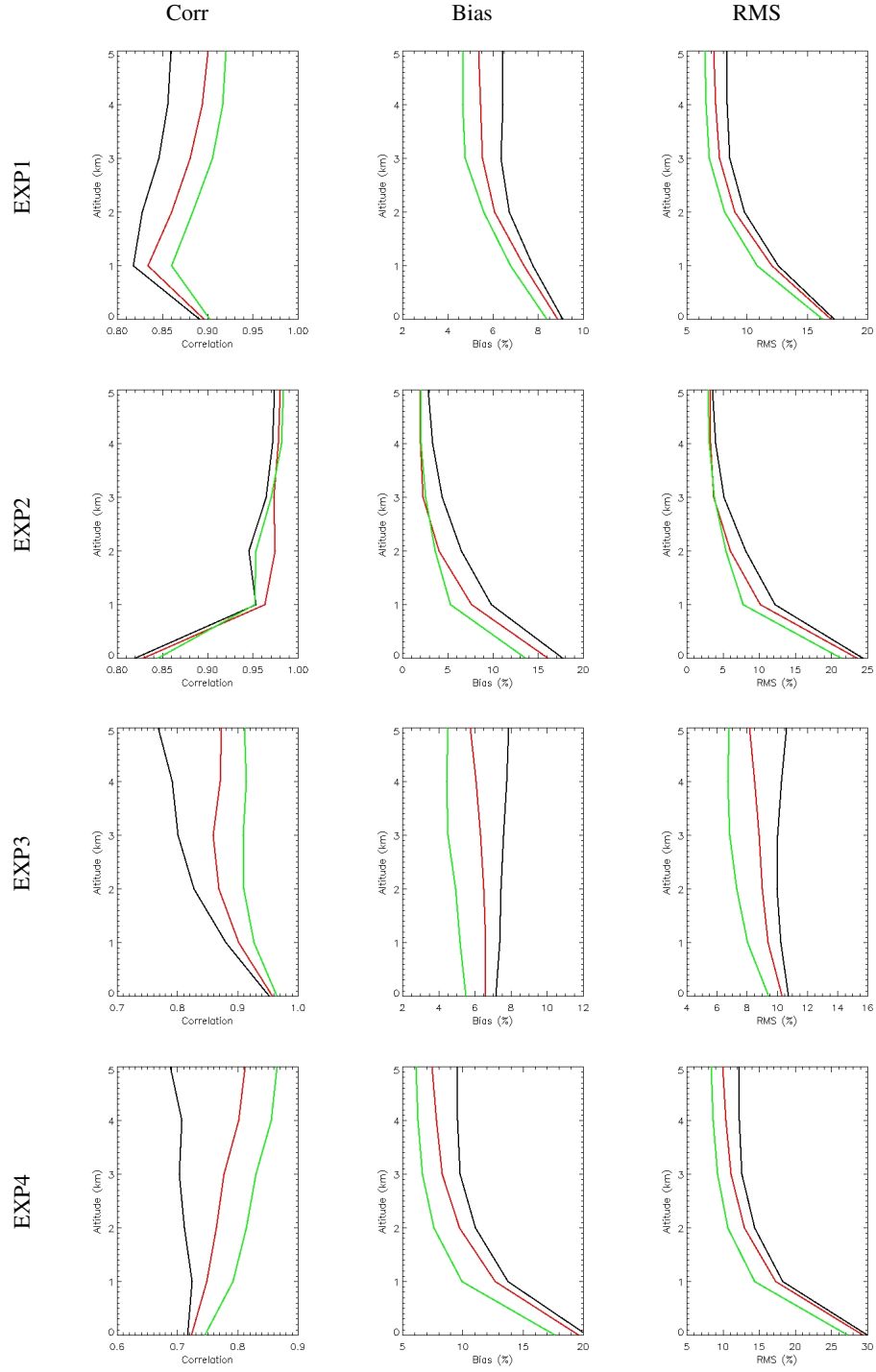


Fig. 14. Same as Figure 13 but for CO.

Table 5. Scores for O₃ (percentage of good detection (GD), percentage of correct forecast above threshold (GD+), percentage of false alarm (FA)) obtained over Europe during July and August 2009 by comparing the control run to the nature run (2nd column), the assimilation run with GEO-TIR2 to the nature run (3rd column) and the assimilation run with GEO-TIR to the nature run (3rd column). Bold scores indicates that the assimilation run is better than the control run, underlined scores indicates that one of the assimilation runs (GEO-TIR or GEO-TIR2) is better than the other one. See text for details about the different scores.

Experiment	No Assim			GEO-TIR2			GEO-TIR		
	GD	GD+	FA	GD	GD+	FA	<u>GD</u>	GD+	<u>FA</u>
EXP 1	86.9	75.2	25.4	86.7	<u>77.9</u>	27.3	<u>87.3</u>	77.0	<u>25.3</u>
EXP 2	75.6	99.2	48.4	75.9	<u>99.2</u>	48.2	<u>77.5</u>	99.1	<u>46.5</u>
EXP 3	92.8	89.9	16.4	93.2	89.4	14.9	<u>94.6</u>	<u>91.6</u>	<u>12.0</u>
EXP 4	72.8	90.1	51.4	73.0	89.9	51.1	<u>74.4</u>	<u>90.0</u>	<u>49.6</u>