



Validation report of the CAMS near-real time global atmospheric composition service

Status up to 1 September 2016

Issued by: KNMI / Henk Eskes

Date: 28/11/2016

Ref: CAMS84_2015SC1_D84.1.5_201611_v1.0

This document has been produced in the context of the Copernicus Atmosphere Monitoring Service (CAMS). The activities leading to these results have been contracted by the European Centre for Medium-Range Weather Forecasts, operator of CAMS on behalf of the European Union (Delegation Agreement signed on 11/11/2014). All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission and the European Centre for Medium-Range Weather Forecasts has no liability in respect of this document, which is merely representing the authors view.



Validation report of the CAMS near-real-time global atmospheric composition service. Status up to 1 September 2016

EDITORS:

H.J. ESKES (KNMI), A. WAGNER (MPG), M. SCHULZ (METNO),
Y. CHRISTOPHE (BIRA-IASB), M. RAMONET (LSCE)

AUTHORS:

S. BASART (BSC), A. BENEDICTOW (METNO), A.-M. BLECHSCHMIDT (IUP-UB),
S. CHABRILLAT (BIRA-IASB), H. CLARK (CNRS-LA), E. CUEVAS (AEMET),
H. FLENTJE (DWD), K. M. HANSEN (AU), U. IM (AU),
J. KAPSOMENAKIS (AA), B. LANGEROCK (BIRA-IASB), A. RICHTER (IUP-UB),
N. SUDARCHIKOVA (MPG), V. THOURET (CNRS-LA), T. WARNEKE (UBC),
C. ZEREFOS (AA)

REPORT OF THE COPERNICUS ATMOSPHERE MONITORING SERVICE, VALIDATION SUBPROJECT.

AVAILABLE AT:

[HTTP://ATMOSPHERE.COPERNICUS.EU/QUARTERLY_VALIDATION_REPORTS](http://atmosphere.copernicus.eu/quarterly_validation_reports)

CITATION:

VALIDATION REPORT OF THE CAMS NEAR-REAL-TIME GLOBAL ATMOSPHERIC COMPOSITION SERVICE. SYSTEM EVOLUTION AND PERFORMANCE STATISTICS; STATUS UP TO 1 SEPTEMBER 2016.

H.J. ESKES, A. WAGNER, M. SCHULZ, Y. CHRISTOPHE, M. RAMONET, S. BASART, A. BENEDICTOW, A.-M. BLECHSCHMIDT, S. CHABRILLAT, H. CLARK, E. CUEVAS, H. FLENTJE, K.M. HANSEN, U. IM, J. KAPSOMENAKIS, B. LANGEROCK, A. RICHTER, N. SUDARCHIKOVA, V. THOURET, T. WARNEKE, C. ZEREFOS,
COPERNICUS ATMOSPHERE MONITORING SERVICE (CAMS) REPORT,
CAMS84_2015SC1_D84.1.5_201611_v1.pdf, NOVEMBER 2016.

STATUS:

VERSION 1, FINAL

DATE:

28/11/2016



1 Executive Summary

The Copernicus Atmosphere Monitoring Service (<http://atmosphere.copernicus.eu>, CAMS) is a component of the European Earth Observation programme Copernicus. The CAMS global near-real time (NRT) service provides daily analyses and forecasts of reactive trace gases, greenhouse gases and aerosol concentrations. The CAMS system was developed by a series of MACC research projects (MACC I-II-III) until July 2015. This document presents the validation statistics and system evolution of the CAMS NRT service for the period until 1 September 2016. Updates of this document appear every 3 months.

This summary is split according to areas of interest to users: Climate forcing, regional air quality, and stratospheric ozone. Specific attention is given to the ability of the CAMS system to capture recent events. We focus on the 'o-suite' composition fields, operationally produced by the C-IFS (Composition-IFS) modelling system at ECMWF. The o-suite generates daily analyses and forecasts, using the available meteorological and atmospheric composition observations which are ingested in the ECMWF 4D-Var assimilation system. For analyses and forecasts of trace gases the CB05 tropospheric chemistry is used, while for aerosol this is the CAMS prognostic aerosol module. We furthermore assess the impact of the composition observations by comparing the validation results from the 'o-suite' to a 'control' configuration without assimilation. Also the pre-operational high-resolution forecasts of CO₂ and CH₄ are assessed in this report.

The o-suite data availability for the period June-August 2016 was very good, with 100% of the forecasts delivered before the target time, 22:00 UTC.

Climate forcing

Tropospheric ozone (O₃)

Model ozone is validated with respect to surface and free tropospheric ozone observations from the GAW and ESRL networks, IAGOS airborne data and ozone sondes. For free tropospheric ozone against sondes the o-suite modified normalized mean biases (MNMBs) are on average smaller $\pm 10\%$ over the Northern Hemisphere (NH), and between $\pm 20\%$ for stations in the Tropics (Fig. S1). This is an improvement compared to the control experiment without the assimilation of composition observations. For June to August 2016 good agreement is found over the NH mid latitudes in the free troposphere, which is confirmed with IAGOS evaluations over Paris, Amsterdam and Frankfurt.

The o-suite shows an overestimation of surface ozone for Europe during June and August 2016 with MNMBs of around 10% on average. For USA the o-suite shows an overestimation of surface ozone of around 25% on average. For Asia, the o-suite shows an overestimation of surface ozone MNMBs of around 50% on average. For the tropics, the surface ozone is overestimated around 30%. For Antarctic and Arctic stations, the o-suite shows a good correspondence with observed surface ozone mixing ratios (MNMBs $< \pm 10\%$). The data assimilation corrects the negative offset visible for the control run.

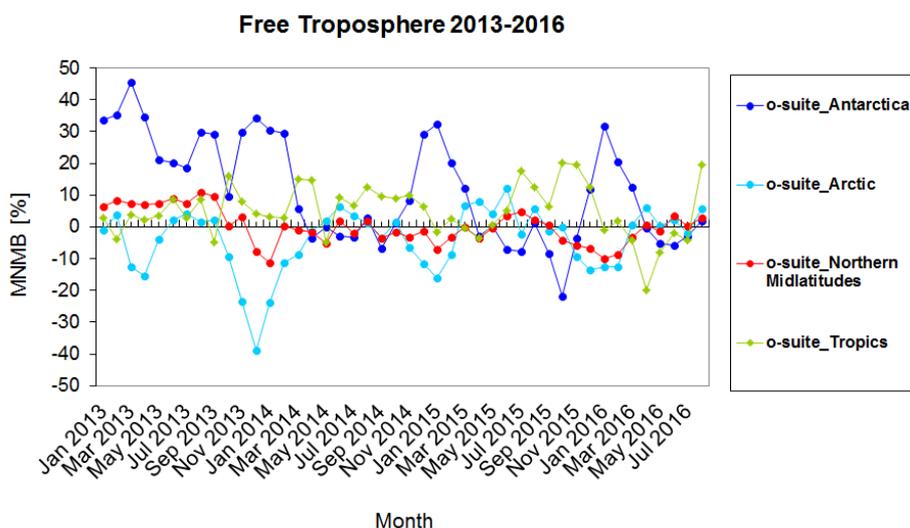


Figure S1: Time series of MNMB of ozone in the o-suite, compared against ozone sondes, averaged over different latitude bands.

Tropospheric Nitrogen dioxide (NO₂)

Model validation, with respect to SCIAMACHY/Envisat NO₂ data before April 2012 and GOME-2/MetOp-A NO₂ data afterwards, shows that tropospheric NO₂ columns are well reproduced by the NRT model runs, indicating that emission patterns and NO_x photochemistry are generally well represented, although modelled shipping signals are larger than the satellite retrievals. Compared to satellite data, all model runs underestimate tropospheric background values over Africa, Europe and the US. Since December 2014, the agreement between satellite retrievals and model results for time series over East-Asia and Europe is better than for previous years (Fig. S2), and observed columns of NO₂ decreased recently, likely associated with reduced emissions. Spring and summertime values over East-Asia are overestimated by the o-suite in 2015, a feature which does not occur for previous years. Compared to satellite data, tropospheric background values over Africa, South America and Australia are currently underestimated by the models, while local maxima over Central Africa are overestimated, likely due to overestimation of fire emissions for Central Africa. Evaluation against MAX-DOAS observations illustrates the positive impact of data assimilation for urban sites, leading to an increase in NO₂.

Tropospheric Carbon Monoxide (CO)

Model validation with respect to GAW network surface observations, IAGOS airborne data, FTIR observations (NDACC and TCCON) and MOPITT and IASI satellite retrievals reveals that the seasonality of CO can be reproduced well by both model versions. A small, consistent negative bias of -5% against MOPITT appears in the o-suite throughout the year over Europe and the US, but for the latest period (JJA) it is further reduced. Also compared to IAGOS aircraft observations over Europe and Asia, modelled free tropospheric CO mixing ratios show an underestimation compared to the measurements, whereas the control run partly overestimates the observations. This is confirmed with comparison against GAW surface observations for Asia (MNMBs between 6% and -18%). During the fire season over Alaska and Siberia negative biases are within 5%. The two northern hemisphere TCCON sites, however, show a slight overestimation of CO in the o-suite.

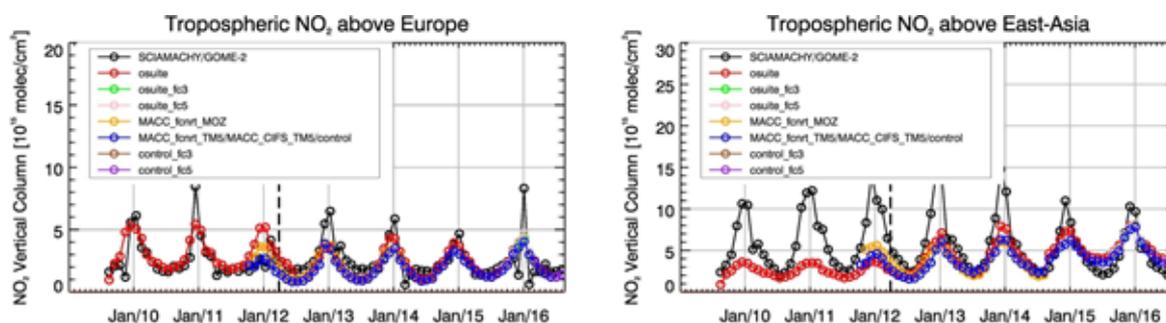


Figure S2: Time series of tropospheric NO₂ columns from SCIAMACHY (up to March 2012), GOME-2 (from April 2012 onwards) compared to model results for Europe and East-Asia. The o-suite is in red, control is in blue (before Sept 2014 blue and yellow represent older model configurations).

Especially the control run shows an overestimation of CO total columns in the tropics, SH and Asia. This overestimation is reduced by the data assimilation for the o-suite. The positive impact of the assimilation of satellite CO on model results shows especially over East and South Asia and North and South Africa, and Réunion Island, whereas for Europe and the US, the control run corresponds better to satellite and surface CO observations. The forecasts (D+1, D+4) are mostly identical to analysis (within 1% difference).

Formaldehyde

Model validation, with respect to SCIAMACHY/Envisat HCHO data before April 2012 and GOME-2/MetOp-A HCHO data afterwards, shows that modelled monthly HCHO columns represent well the magnitude of oceanic and continental background values and the overall spatial distribution in comparison with mean satellite HCHO columns. Compared to GOME-2 satellite retrievals, there is a strong overestimation of values for Northern Australia and Central Africa. As for tropospheric NO₂, the latter may be due to an overestimation of HCHO emissions from fires for Central Africa. For time series over East-Asia and the Eastern US, both regions where HCHO columns are probably dominated by biogenic emissions, models and retrievals agree rather well. However, the yearly cycle over East-Asia is underestimated by the models.

The validation of model profiles with ground-based UV-VIS DOAS measurements over Xianghe, near Beijing, shows that background column values are underestimated by around 30%, in agreement with satellite observations for this region. Also local pollution events are not captured correctly, in part due to the relatively coarse horizontal resolution of the global models, and in part associated with uncertainties in HCHO and precursor emissions. Note that no formaldehyde observations are assimilated in the system.

Aerosol

We estimate that the o-suite aerosol optical depth showed an average positive bias in the latest three months of +30%, measured as modified normalized mean bias against daily Aeronet sun photometer data. The +3 day forecasted aerosol distributions, since July 2012, show 10-30% less aerosol optical depth (AOD) than those from the initial day, as shown all in Figure S3a. The correlation, shown in figure S3b, shows month-to-month variation ranging from 0.65 to 0.85, indicating the simulation reproduces approximately 50% of the day to day AOD variability across all Aeronet stations, with higher correlations in the last 6 months (at least compared to the year

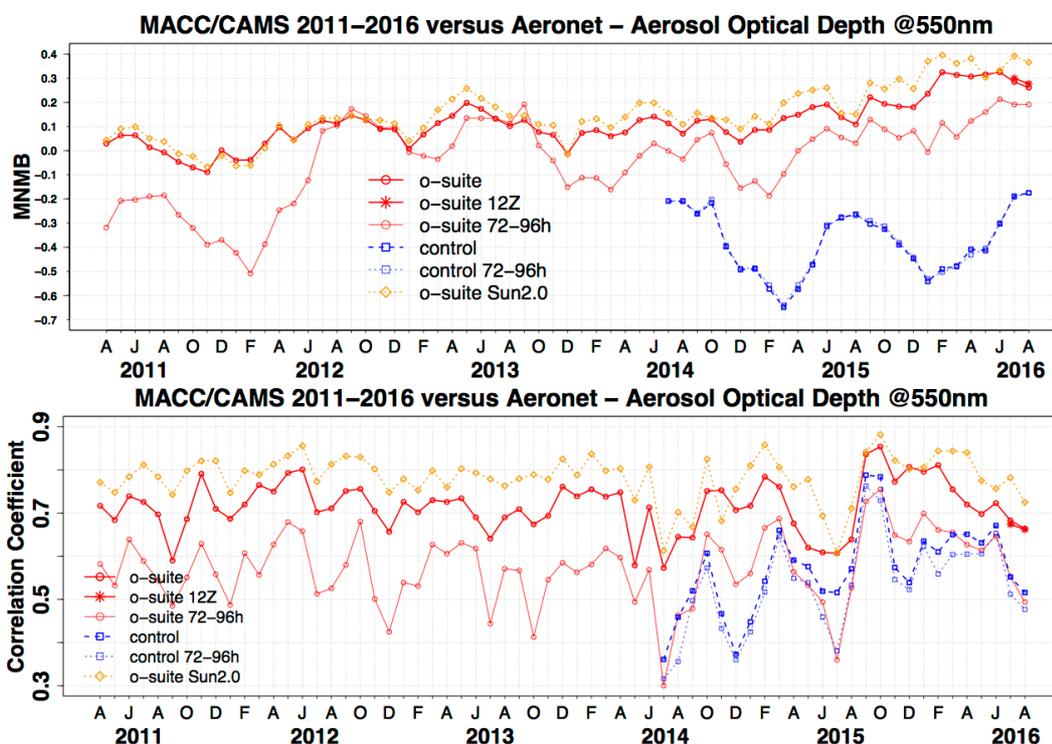


Figure S3. Aerosol optical depth at 550nm in IFS model simulations for April 2011 - August 2016 against daily matching Aeronet NRT level 1.5 and level 2.0 data a) Modified normalized mean bias (MNMB); o-suite (thick red curve); o-suite at last forecast day (light red curve); Control (blue dashed); Control at last forecast day (light blue dashed); o-suite but evaluated against quality assured Aeronet level 2.0 data (orange dashed); b) Corresponding correlation coefficient. Note that quality assured level 2.0 data amount decreases from ca 2800 daily data points per month (mean in 2014) to ca. 200 data points in the last two month of the time series. The results of the 12Z forecast run have been compared with the 00Z run.

before). The latter indicates that assimilating the MODIS deep blue product since September 2015 improves aerosol AOD simulation. The o-suite forecast at +3 days shows slightly lower correlation, as a consequence of imperfect forecasted meteorology and fading impact of the initial assimilation of MODIS AOD and MODIS fire info on model performance. The second o-suite running each day at 12UTC showed almost identical performance as the o-suite starting at 00UTC.

The regional AOD performance of the o-suite with respect to the AERONET data exhibits a seasonal cycle depending on region. A lower correlation in autumn and winter in North America can be noted. The smallest bias is shown in East Asia, and last months show a higher positive bias in North America (+50%).

The aerosol Ångström exponent contains information about the size distribution of the aerosol, and implicitly composition. The o-suite continues to show a positive global bias against Aeronet data of +20%, indicating too fine particles in the model, possibly dominated by sulphate, which represents ca 45% of global mean AOD. Correlation is lower in autumn and winter.

For this report PM10 data were used directly as defined by the IFS system, and PM10 concentrations decreased by 50% compared to earlier validation reports, eg those in 2015. An evaluation of these PM10 surface concentrations against a climatological average (2000-2009) at

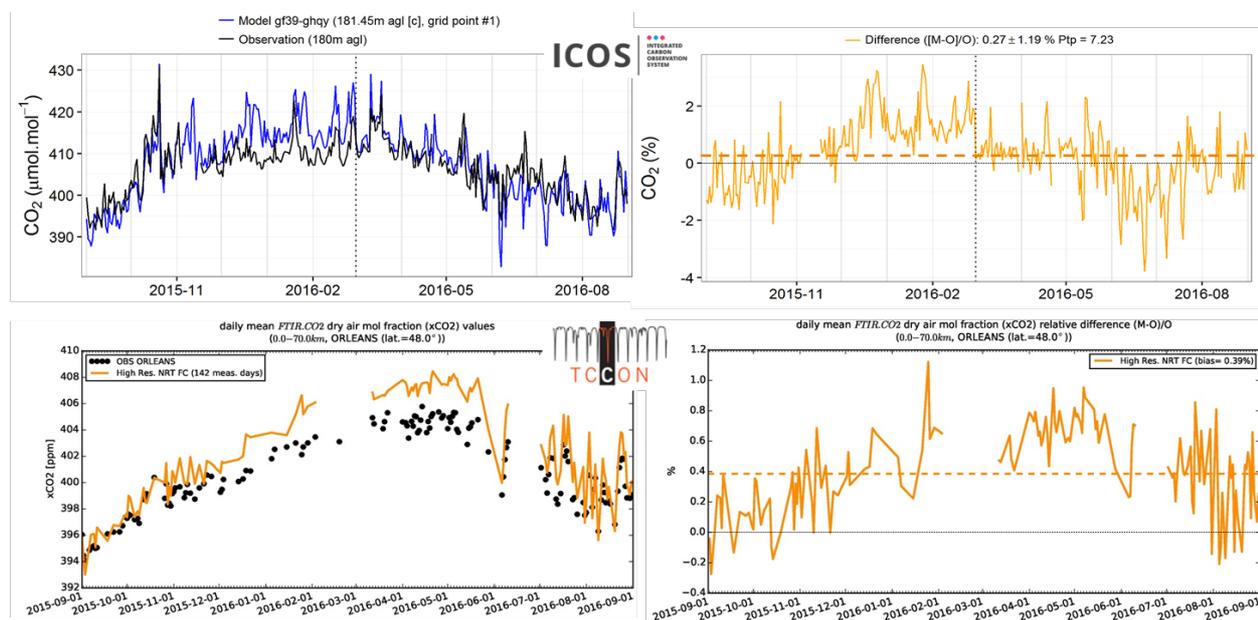


Figure S4. CO₂ mole fractions observed at Trainou, 100km south of Paris, from the ICOS tall tower (180 m agl) and the TCCON FTS instrument (total column) for September 2015 - August 2016. Left figures show the comparison of simulated and observed mole fractions, and differences are shown on the right panel for ICOS (above) and TCCON (below).

150 background sites in North America and Europe indicates overestimations at some sites closer to the coast, possibly due to high simulated sea salt concentrations. However, PM₁₀ concentrations more inland exhibit an underestimation with MNMB bias of -30% both in Europe and North America.

During summer, CAMS o-suite can simulate the main areas of dust activity in comparison with MODIS aerosol product and the observed variability in the AERONET sites with correlation values of 0.56 in the Middle East and 0.41 in the Sahara. Although CAMS o-suite tends to overestimate the dust aerosol optical depth (DOD) over desert dust regions, during summer the model present underestimations in the Sahara associated to strong and fast dust outbreaks associated with a mesoscale convective system. Maximum dust activity over long-range transport regions is observed in Central and Western Mediterranean and subtropical North Atlantic. The Mediterranean is the region where the highest correlations are achieved ($r > 0.70$).

A preliminary evaluation of vertical profiles of aerosol backscatter coefficient derived from the German ceilometer network indicates that during dust events model profiles confirm the suspected presence of dust in the observations, and vice versa. Small-scale structures in dust plumes are not resolved, most likely due to model resolution. Profiles during elevated sea salt periods show more disagreement with observations and sea salt seems to be overestimated inland during storm events, confirming PM₁₀ bias findings above.

Greenhouse gases

Pre-operational high-resolution forecasts of CO₂ and CH₄ have been compared to ICOS surface (15 sites) and TCCON total column (3 sites) measurements, for a one year period from (Sept. 2015 to Sep. 2016). Most of the stations are located in Europe (9 ICOS and 2 TCCON sites) providing a better



representativeness over this continent. All stations show an over estimation of the CO₂ seasonal cycle in Europe, which is illustrated on figure S4 at the Trainou tall tower where are combined surface (ICOS) and total column (TCCON) measurements. For both observations the model simulates too high CO₂ values in wintertime up to 2% for the tall tower and 1% for the total column. The surface measurements also indicate an over estimation up to 2% of the CO₂ drawdown during the growing season (June-July 2016). This disagreement is not seen by the total column observations which agree pretty well to the simulations during summertime. The Trainou site also demonstrates the importance of the tall tower to characterize CO₂ variations at continental sites, since the bias decrease from -5 ppm at 5 m agl down to -0.5 ppm at 180m agl.

The evaluation of the model performances are more difficult to evaluate in tropics and Southern hemisphere due to the limited number of observations. La Reunion Island where are installed ICOS and TCCON instruments is a challenging site for the model due to the hotspot of emissions close to the stations and the topography of the Island. The CO₂ variability of the model is too large for both dataset. The CH₄ mole fractions are clearly better simulated at all southern sites with the new experiment (ghqy) started on March 2016.

System performance in the Arctic

The CAMS model runs are validated using surface ozone measurements from the ESRL-GMD and the IASOA networks (5 sites) and ozone concentrations in the free troposphere are evaluated using balloon sonde measurement data.

For the period from December 2014 to May 2016 the simulations of the surface ozone concentrations are on average in good agreement with the observations apart from ozone depletion events in spring (March to June), which are not captured by the model simulations. These events are related to halogen chemistry reactions that are not represented in the C-IFS model.

During the period June 2016 – August 2016 the o-suite slightly underestimates observed O₃ mixing ratios (between -5 and -16%) for three of the four stations. MNMBs for the control run are larger (between -10 and -36%). For the Arctic free troposphere, the o-suite shows low MNMBs (between -2 and 6%), whereas the control has a negative offset.

System performance in the Mediterranean

The model is compared to surface O₃ observations from the AirBase network. Our analysis shows a considerable contrast in the model MNMBs between the Mediterranean shore of Spain (MNMBs around ±20%) and Eastern Mediterranean (MNMBs≈0). Temporal correlation coefficients between simulated and observed surface ozone mixing ratios are higher for the control run (on average around 0.5).

Regional air quality

Ozone, CO and aerosol boundary conditions

Free tropospheric ozone concentrations in the o-suite in the northern midlatitudes are generally in good correspondence with ozone sondes, MNMBs in the range of ±10% (for the last six months ±3%). The o-suite shows a positive bias in surface ozone concentrations in Europe, with MNMBs for GAW and ESRL stations ranging between 5% and 20% for the period June to August 2016, and also

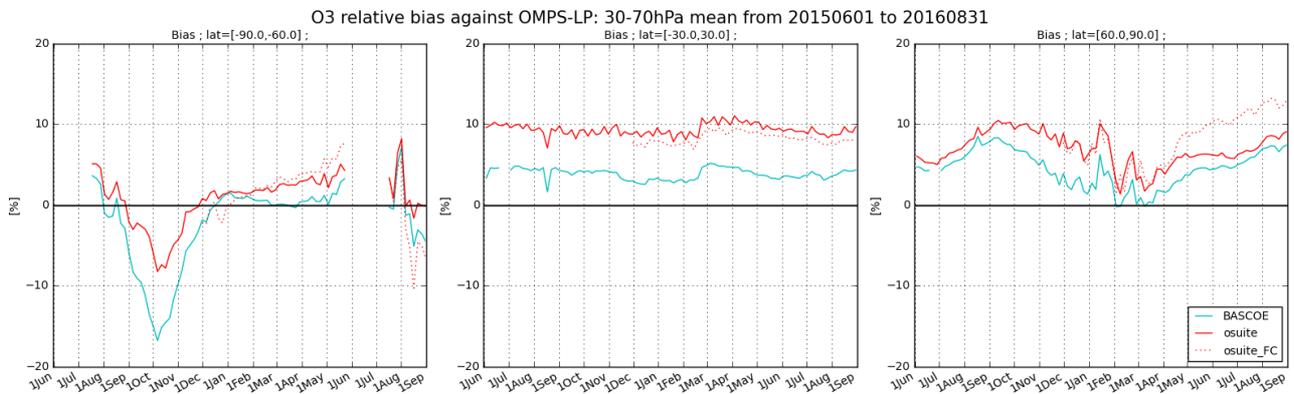


Figure S4: Time series of the normalized mean bias (%) between ozone from o-suite analyses (red, solid) or 4th day forecasts (red, dotted), or BASCOE analyses (cyan) and OMPS-LP satellite observations, in the middle stratosphere (30-70hPa averages).

positive biases over North American stations (between 20 and 30%). The o-suite mostly underestimates surface CO concentration in Europe and Asia with MNMBs with respect to GAW of around -12%.

Ozone layer

Ozone partial columns and vertical profiles

Ozone columns and profiles have been compared with the following observations: vertical profiles from balloon-borne ozonesondes; ground-based remote-sensing observations from the NDACC (Network for the Detection of Atmospheric Composition Change); and satellite observations by the limb-scanning instrument OMPS-LP. Furthermore, the o-suite analyses are compared with those delivered by two independent assimilation systems: BASCOE, and TM3DAM.

Compared to ozone sondes the model O₃ partial pressures are mostly slightly overestimated in all latitude bands (MNMB between 0 and +10%).

Comparisons with the NDACC network include microwave observations for Ny Alesund (78.9°N) and Bern (47°N) and LIDAR observations at Hohenpeissenberg (47.8°N) and Lauder (45°S). Among these stations the o-suite performs best at Bern with stratospheric columns evolving since September 2015 with seasonally averaged relative biases smaller than 5%, which is smaller than the reported measurement uncertainties. At Ny Alesund, the seasonally averaged bias of the stratospheric column almost vanishes during summer months, while during the rest of the year the o-suite overestimates (>10%) the ozone abundance between 25km and 35km. Compared with the LIDAR at Lauder and Hohenpeissenberg, the o-suite does not show significant biases with the observed ozone between 20km and 35km.

The comparison with OMPS-LP delivers a good agreement in the middle stratosphere and confirms the overestimation by the o-suite in the lower stratosphere. This overestimation reaches 10% in the Tropics (70 hPa) and 20% in the mid-latitudes and the Arctic (100 hPa). The time evolution of the normalized mean bias in the lower middle stratosphere (Fig. S4) shows a systematic overestimation by the o-suite (5-10%) in the Tropics and the Northern Hemisphere. Also, the 4th day forecasts



exhibits an increased bias in the Northern Hemisphere since April 2016, and an underestimation in the Antarctic in August 2016.

Other stratospheric trace gases

Due to the lack of stratospheric chemistry in the C-IFS-CB05 scheme, the only useful product in the stratosphere is ozone. Other species, like NO₂, have also been evaluated but the results are only indicative.

Events

A fire event took place in the central part of South America during the 20th of August 2016. Both runs could capture the location of the plum, showing CO values that are in good agreement with the satellite data over the region of the fire case. The transportation pathway of the plum could be reproduced by both runs. However, over the oceans the control run simulates larger values compared to the satellite data.

In mid-June 2016, winds lofted thick plumes of dust from northern Africa's deserts high into the air. On June 19, winds had already swept a plume of dust westward over the Atlantic Ocean and the archipelago of Cabo Verde (Cape Verde) that reached the Canary Islands by June 22. Otherwise, a thinner plume of Saharan dust also spread north toward Europe starting on June 16. This dust outbreak from the Western Sahara was coincident with a plume of dust from Africa's northeast, carried eastward over the Red Sea. CAMS o-suite can timely reproduce the spatial distribution of the different dust events as observed by MODIS over Northern Africa and the Middle East. Also, the model tracks fairly well the changes in the shape and size of the dust layer throughout the period of analysis. The model shows a clear underestimation of the dust event in the Red Sea. Finally, the comparison of the modelled surface concentration shows how the model fits fairly well the reduction of the visibility over the dust sources.



Table of Contents

1. Introduction	14
2. System summary and model background information	17
2.1 System based on the ECMWF IFS model	17
2.1.1 o-suite	17
2.1.2 Control	19
2.1.3 High-resolution CO ₂ and CH ₄ forecasts	20
2.2 Evolution of the IFS-based system	20
2.3 Other systems	22
2.3.1 BASCOE	22
2.3.2 TM3DAM and the multi-sensor reanalysis	22
2.3.3 SDS-WAS multimodel ensemble	23
2.4 CAMS products	23
2.5 Availability and timing of CAMS products	23
3. Validation results for reactive gases and aerosol	25
3.1 Tropospheric Ozone	25
3.1.1 Validation with sonde data in the free troposphere	25
3.1.2 Ozone validation with IAGOS data	26
3.1.3 Validation with GAW and ESRL-GMD surface observations	36
3.1.4 Validation with AirBase observations in Mediterranean	40
3.1.5 Validation with IASOA surface observations	43
3.2 Tropospheric nitrogen dioxide	44
3.2.1 Evaluation against GOME-2 retrievals	44
3.2.2 Evaluation against ground-based DOAS observations	48
3.3 Carbon monoxide	49
3.3.1 Validation with Global Atmosphere Watch (GAW) Surface Observations	49
3.3.2 Validation with IAGOS Data	52
3.3.3 Validation against FTIR observations from the NDACC network	55
3.3.4 Evaluation with MOPITT and IASI data	57
3.3.5 Evaluation against TCCON CO	62
3.3.6 Validation with AirBase observations in Mediterranean	65
3.4 Formaldehyde	67
3.4.1 Validation against satellite data	67
3.4.2 Validation against UVVIS DOAS observations from the NDACC network	71
3.5 Aerosol	73
3.5.1 Global comparisons with Aeronet and PM	73
3.5.2 Dust forecast model intercomparison: Validation of DOD against AERONET, and comparisons with Multimodel Median from SDS-WAS	77
3.5.3 Aerosol validation over the Mediterranean	83
3.5.4 Backscatter profiles	87
3.6 Stratospheric ozone	95
3.6.1 Validation against ozone sondes	95
3.6.2 Validation against observations from the NDACC network (MWR, LIDAR)	96



3.7 Stratospheric NO₂	101
4. Validation results for greenhouse gases	103
4.1 CH₄ and CO₂ validation against ICOS observations	103
4.2 CH₄ and CO₂ validation against TCCON observations	109
4.2.1 Evaluation against TCCON CO ₂	109
4.2.2 Evaluation against TCCON CH ₄	109
5. Events	112
5.1 Fire case in the central part of South America in late August 2016	112
5.2 A dusty period over North Africa, Middle East and Europe: mid-June 2016	113
6. References	118
Annex 1: Acknowledgements	123



1. Introduction

The Copernicus Atmosphere Monitoring Service (CAMS, <http://atmosphere.copernicus.eu/>) is a component of the European Earth Observation programme Copernicus. The CAMS global near-real time (NRT) service provides daily analyses and forecasts of trace gas and aerosol concentrations. The CAMS system was developed by a series of MACC research projects (MACC I-II-III). The CAMS near-real time services consist of daily analysis and forecasts with the Composition-IFS system with data assimilation of trace gas concentrations and aerosol properties. This document presents the system evolution and the validation statistics of the CAMS NRT global atmospheric composition analyses and forecasts. The validation methodology and measurement datasets are discussed in Eskes et al. (2015).

In this report the performance of the system is assessed in two ways: both the longer-term mean performance (seasonality) as well as its ability to capture recent events are documented. Table 1.1 provides an overview of the trace gas species and aerosol aspects discussed in this CAMS near-real time validation report. This document is updated every 3 months to report the latest status of the near-real time service.

This report covers results for a period of at least one year to document the seasonality of the biases. Sometimes reference is made to other model versions or the reanalysis to highlight aspects of the near-real time products.

Key CAMS NRT products and their users are: Boundary conditions for regional air quality models (e.g. AQMEII, air quality models not participating in CAMS); Long range transport of air pollution (e.g. LRTAP); Stratospheric ozone column and UV (e.g. WMO, DWD); 3D ozone fields (e.g. SPARC).

As outlined in the MACC-II Atmospheric Service Validation Protocol (2013) and MACC O-INT document (2011), relevant user requirements are quick looks of validation scores, and quality flags and uncertainty information along with the actual data. This is further stimulated by QA4EO (Quality Assurance Framework for Earth Observation, <http://www.qa4eo.org>) who write that “all earth observation data and derived products is associated with it a documented and fully traceable quality indicator (QI)”. It is our long-term aim to provide such background information. The user is seen as the driver for any specific quality requirements and should assess if any supplied information, as characterised by its associated QI, are “fit for purpose” (QA4EO task team, 2010).

CAMS data are made available to users as data products (grib or netcdf files) and graphical products from ECMWF, <http://atmosphere.copernicus.eu/global-near-real-time-data-access>. The stratospheric ozone service is provided by BIRA-IASB at <http://copernicus-stratosphere.eu>.

A summary of the system and its recent changes is given in section 2. Section 3 gives an overview of the performance of the system from a seasonal (climatological) perspective, for various species. Section 4 describes the performance of the system during recent events. Extended validation can be found online via regularly updated verification pages, <http://atmosphere.copernicus.eu/user-support/validation/verification-global-services>. Table 1.2 lists all specific validation websites that can also be found through this link.



Table 1.1: Overview of the trace gas species and aerosol aspects discussed in this CAMS near-real time validation report. Shown are the datasets assimilated in the CAMS analysis (second column) and the datasets used for validation, as shown in this report (third column). Green colors indicate that substantial data is available to either constrain the species in the analysis, or substantial data is available to assess the quality of the analysis. Yellow boxes indicate that measurements are available, but that the impact on the analysis is not very strong or indirect (second column), or that only certain aspects are validated (third column).

Species, vertical range	Assimilation	Validation
Aerosol, optical properties	MODIS Aqua/Terra AOD	AOD, Ångström: AERONET, GAW, Skynet, MISR, OMI, lidar, ceilometer
Aerosol mass (PM10, PM2.5)	-	European AirBase stations
O ₃ , stratosphere	MLS, GOME-2A, GOME-2B, OMI, SBUV-2	Sonde, lidar, MWR, FTIR, OMPS, BASCOE and MSR analyses
O ₃ , UT/LS	Indirectly constrained by limb and nadir sounders	IAGOS, ozone sonde
O ₃ , free troposphere	Indirectly constrained by limb and nadir sounders	IAGOS, ozone sonde
O ₃ , PBL / surface	-	Surface ozone: WMO/GAW, NOAA/ESRL-GMD, AIRBASE
CO, UT/LS	-	IAGOS
CO, free troposphere	IASI, MOPITT	IAGOS, MOPITT, IASI, TCCON
CO, PBL / surface	Indirectly constrained by satellite IR sounders	Surface CO: WMO/GAW, NOAA/ESRL
NO ₂ , troposphere	OMI, partially constrained due to short lifetime	SCIAMACHY, GOME-2, MAX-DOAS
HCHO	-	GOME-2, MAX-DOAS
SO ₂	GOME-2A, GOME-2B (Volcanic eruptions)	-
Stratosphere, other than O ₃	-	NO ₂ column only: SCIAMACHY, GOME-2
CO ₂ , surface, PBL		ICOS
CO ₂ , column		TCCON
CH ₄ , surface, PBL		ICOS
CH ₄ , column		TCCON



Table 1.2: Overview of quick-look validation websites of the CAMS system.

Reactive gases – Troposphere
GAW surface ozone and carbon monoxide: http://macc.copernicus-atmosphere.eu/d/services/gac/verif/grg/gaw/gaw_station_ts/
IAGOS tropospheric ozone and carbon monoxide: http://www.iagos.fr/cams/
Surface ozone from EMEP (Europe) and NOAA-ESRL (USA): http://www.academyofathens.gr/cams
Tropospheric nitrogen dioxide and formaldehyde columns against satellite retrievals: http://www.doas-bremen.de/macc/macc_veri_iup_home.html
Tropospheric CO columns against satellite retrievals: http://cams.mpimet.mpg.de
Reactive gases - Stratosphere
Stratospheric composition: http://www.copernicus-stratosphere.eu
NDACC evaluation in stratosphere and troposphere (the NORS server) http://nors-server.aeronomie.be
Aerosol
Evaluation against selection of Aeronet stations: http://www.copernicus-atmosphere.eu/d/services/gac/verif/aer/nrt/
Aerocom evaluation: http://aerocom.met.no/cgi-bin/aerocom/surfobs_annualrs.pl?PROJECT=MACC&MODELLIST=MACC-VALreports&
WMO Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS) model intercomparison and evaluation: http://sds-was.aemet.es/forecast-products/models
Satellite data monitoring
Monitoring of satellite data usage in the Reanalysis and Near-Real-Time production: http://copernicus-atmosphere.eu/d/services/gac/monitor/

This validation report is accompanied by the "Observations characterization and validation methods" report, Eskes et al. (2016), which describes the observations used in the comparisons, and the validation methodology. This report can also be found on the global validation page, <http://atmosphere.copernicus.eu/user-support/validation/verification-global-services>.



2. System summary and model background information

The specifics of the different CAMS model versions are given (section 2.1) with a focus on the model changes (section 2.2). An overview of products derived from this system is given in section 2.3. Several external products used for validation and intercomparison are listed in section 2.4. Timeliness and availability of the CAMS products is given in section 2.5.

2.1 System based on the ECMWF IFS model

Key model information is given on the CAMS data-assimilation and forecast run o-suite and its control experiment, used to assess the sensitivity to assimilation. The forecast products are listed in Table 2.1. Table 2.2 provides information on the satellite data used in the o-suite. Further details on the different model runs and their data usage can be found at <http://atmosphere.copernicus.eu/documentation-global-systems>.

Information on older experiment types, including MACC_fcnrt_MOZ and MACC_CIFS_TM5 can be found in older Validation reports available from http://www.gmes-atmosphere.eu/services/aqac/global_verification/validation_reports/.

2.1.1 o-suite

The o-suite consists of the C-IFS-CB05 chemistry combined with the CAMS bulk aerosol model. The chemistry is described in Flemming et al. (2015), aerosol is described by Morcrette et al. (2009). The forecast length is 120 h. The o-suite data is stored under expver '0001' of class 'MC'. On 3 September 2015 the meteorological model has been updated significantly, moving from cy40r2 to cy41r1. On 21 June 2016 the model resolution has seen an upgrade from T255 to T511, and forecasts are produced twice per day. Here a summary of the main specifications of this version of the o-suite is given.

- The meteorological model is based on IFS version cy41r1, see also <http://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model/cy41r1-summary-changes>; the model resolution is T255L60.
- The modified CB05 tropospheric chemistry is used (Williams et al., 2013), originally taken from the TM5 chemistry transport model (Huijnen et al., 2010)
- Stratospheric ozone during the forecast is computed from the Cariolle scheme (Cariolle and Teyssèdre, 2007) as already available in IFS, while stratospheric NO_x is constrained through a climatological ratio of HNO₃/O₃ at 10 hPa.
- Monthly mean dry deposition velocities are based on the SUMO model provided by the MOCAGE team.
- Data assimilation is described in Inness et al. (2015) and Benedetti et al. (2009) for chemical trace gases and aerosol, respectively. Satellite data assimilated is listed in Table 2.2 and Fig. 2.1.
- Anthropogenic and biogenic emissions are based on MACCity (Granier et al., 2011) and a climatology of the MEGAN-MACC emission inventories (Sindelarova et al., 2014)
- NRT fire emissions are taken from GFASv1.2 (Kaiser et al. 2012).



Table 2.1: Overview of model runs assessed in this validation report.

Forecast system	Exp. ID	Brief description	Status
o-suite	0001	Operational CAMS DA/FC run	20160621- present (0067) 20150903-20160620 (g9rr) 20140918-20150902 (g4e2)
Control	gjjh geuh g4o2	control FC run without DA	20160621-present (gjjh) 20150901-20160620 (geuh) 20140701-20150902 (g4o2)
GHG run	gf39 ghqy	High resolution, NRT CO ₂ and CH ₄ runs without DA	20150101-20160229 (gf39) 20160301-present (ghqy)

Table 2.2: Satellite retrievals of reactive gases and aerosol optical depth that are actively assimilated in the o-suite.

Instrument	Satellite	Provider	Version	Type	Status
MLS	AURA	NASA	V3.4	O3 Profiles	20130107 -
OMI	AURA	NASA	V883	O3 Total column	20090901 -
GOME-2A	Metop-A	Eumetsat	GDP 4.7	O3 Total column	20131007 -
GOME-2B	Metop-B	Eumetsat	GDP 4.7	O3 Total column	20140512 -
SBUV-2	NOAA	NOAA	V8	O3 21 layer profiles	20121007 -
IASI	MetOp-A	LATMOS/ULB	-	CO Total column	20090901 -
IASI	MetOp-B	LATMOS/ULB	-	CO Total column	20140918 -
MOPITT	TERRA	NCAR	V5-TIR	CO Total column	20130129-
OMI	AURA	KNMI	DOMINO V2.0	NO2 Tropospheric column	20120705 -
OMI	AURA	NASA	v003	SO2 Tropospheric column	20120705-20150901
GOME-2A/2B	METOP A/B	Eumetsat	GDP 4.7	SO2 Tropospheric column	20150902-
MODIS	AQUA / TERRA	NASA	Col. 5 Deep Blue	Aerosol total optical depth	20090901 - 20150902 -



Figure 2.1: Satellite observation usage in the real-time analysis, from Oct. 2014 onwards. CO: Top three rows; O₃ columns and profiles: rows 4-8; Aerosol Optical Depth: rows 9-10.

The aerosol model includes 12 prognostic variables, which are 3 bins for sea salt and desert dust, hydrophobic and hydrophilic organic matter and black carbon, sulphate aerosols and its precursor trace gas SO₂ (Morcrette et al., 2009). Aerosol total mass is constrained by the assimilation of MODIS AOD (Benedetti et al. 2009). A variational bias correction for the MODIS AOD is in place based on the approach used also elsewhere in the IFS (Dee and Uppala, 2009).

A brief history of updates of the o-suite is given in Table 2.4, and is documented in earlier MACC-VAL reports:

http://www.gmes-atmosphere.eu/services/aqac/global_verification/validation_reports/

2.1.2 Control

The control run (expver=gjjh/geuh/g4o2) applies the same settings as the respective o-suites, based on the coupled C-IFS-CB05 system with CAMS aerosol for cy41r1/cy40r2, except that data assimilation is not switched on. The only two exceptions with regard to this setup are:

- at the start of every forecast the ECMWF operational system is used to initialise *stratospheric ozone*, considering that stratospheric ozone, as well as other stratospheric species are not a useful product of this run. As a consequence, the behavior of this control run will not be discussed in the stratospheric contribution of this report. The reason for doing so is that this ensures reasonable stratospheric ozone as boundary conditions necessary for the tropospheric chemistry.
- The full meteorology in the control run is also initialized from the ECMWF operational NWP analyses. Note that this is different from the o-suite, which uses its own data assimilation setup for meteorology. This can cause slight differences in meteorological fields between o-suite and control, e.g. as seen in evaluations of upper stratospheric temperatures.



2.1.3 High-resolution CO₂ and CH₄ forecasts

The pre-operational forecasts of CO₂ and CH₄ use an independent setup of the IFS as the osuite, at a resolution of TL1279, i.e. ~16 km horizontal, and with 137 levels. This system runs in NRT, and does not apply data assimilation for the greenhouse gases.

The land vegetation fluxes for CO₂ are modelled on-line by the CTESSEL carbon module (Boussetta et al., 2013). A biogenic flux adjustment scheme is used in order to reduce large-scale biases in the net ecosystem fluxes (Agusti-Panareda, 2015). The anthropogenic fluxes are based on the annual mean EDGARv4.2 inventory using the most recent year available (i.e. 2008) with estimated and climatological trends to extrapolate to the current year. The fire fluxes are from GFAS (Kaiser et al., 2012).

Methane fluxes are prescribed in the IFS using inventory and climatological data sets, consistent with those used as prior information in the CH₄ flux inversions from Bergamaschi et al. (2009). The anthropogenic fluxes are from the EDGAR 4.2 database (Janssens-Maenhout et al, 2012) valid for the year 2008. The biomass burning emissions are from GFAS v1.2 (Kaiser et al., 2012).

The high resolution forecast experiments from March 2015 to May 2016 analyzed in this report correspond to two experiments:

- "gf39" from Jan 2015 to Feb 2016. This run was set up to replace run gcbt, which had a bug in the code resulting in spikes in concentration fields.
- "ghqy" from March 2016 to present. The initial conditions used in ghqy on 1st of March 2016 are from the GHG analysis (experiment gg5m). Furthermore, the meteorological analysis used to initialize the ghqy forecast changed resolution and model grid in March 2016.

The high-resolution model run also include a linear CO scheme (Massart et al., 2015), which is also briefly assessed in this report.

2.2 Evolution of the IFS-based system

A list with o-suite system changes from September 2014 until March 2016 are given in Table 2.3. A full list with all changes concerning the assimilation system can be found at <http://atmosphere.copernicus.eu/user-support/operational-info/global-system-changes>. The CAMS o-suite system is upgraded regularly, following updates to the ECMWF meteorological model as well as CAMS-specific updates such as changes in chemical data assimilation. These changes are documented in e-suite validation reports, as can be found from the link above. Essential model upgrades are also documented in Table 2.4.



Table 2.3: Recent changes in the CAMS o-suite setup.

Date	Change
2015.03.23- 2014.04.14	Temporarily no assimilation of MOPITT CO
2015.04.15	Only allow OMI - SO ₂ assimilation for rows 1-20.
2015.09.03	Update of o-suite to CY41R1 C-IFS-CB05 with experiment id g9rr
2016.02.18- 2016.04.21	Terra satellite went into safe mode, implying no data available for MODIS (until 2016.04.11) and MOPITT (until 2016.04.21).
2016.02.26- 2016.03.01	Problem with GFAS fire emissions due to TERRA MODIS coming back on with inaccurate data, mostly pronounced on CO and aerosol over western United States.
2016.05.30- 2016.06.16	Missing NO ₂ and O ₃ data from OMI, due to temporary problems with OMI instrument.

Table 2.4: Long-term o-suite system updates.

Date	o-suite update
2009.08.01	Start of first NRT experiment f7kn with coupled MOZART chemistry, without aerosol. Also without data assimilation.
2009.09.01	Start of first MACC NRT experiment f93i, based on meteo cy36r1, MOZART v3.0 chemistry, MACC aerosol model, RETRO/REAS and GFEDv2 climatological emissions, T159L60 (IFS) and 1.875°×1.875° (MOZART) resolution.
2012.07.05	Update to experiment fnyp: based on meteo cy37r3, MOZART v3.5 chemistry, where changes mostly affect the stratosphere, MACCity (gas-phase), GFASv1 emissions (gas phase and aerosol), T255L60 (IFS) and 1.125°×1.125° (MOZART) resolution. Rebalancing aerosol model, affecting dust.
2013.10.07	Update of experiment fnyp from e-suite experiment fwu0: based on meteo cy38r2, no changes to chemistry, but significant rebalancing aerosol model. Assimilation of 21 layer SBUV/2 ozone product
2014.02.24	Update of experiment fnyp from e-suite experiment fzpr: based on meteo cy40r1. No significant changes to chemistry and aerosol models.
2014.09.18	Update to experiment g4e2: based on meteo cy40r2. In this model version C-IFS-CB05 is introduced to model atmospheric chemistry.
2015.09.03	Update to experiment g9rr: based on meteo cy41r1.
2016.06.21	Update to experiment 0067: based on meteo cy41r1, but a resolution increase from T255 to T511, and two production runs per day



2.3 Other systems

2.3.1 BASCOE

The NRT analyses and forecasts of ozone and related species for the stratosphere, as delivered by the Belgian Assimilation System for Chemical Observations (BASCOE) of BIRA-IASB (Lefever et al., 2014; Errera et al., 2008), are used as an independent model evaluation of the CAMS products. The NRT BASCOE product is the ozone analysis of Aura/MLS-SCI level 2 standard products, run in the following configuration (version 05.07):

- The following species are assimilated: O₃, H₂O, HNO₃, HCl, HOCl, N₂O and ClO.
- It lags by typically 4 days, due to latency time of 4 days for arrival of non-ozone data from Aura/MLS-SCI (i.e. the scientific offline Aura/MLS dataset).
- Global horizontal grid with a 3.75° longitude by 2.5° latitude resolution.
- Vertical grid is hybrid-pressure and consists in 86 levels extending from 0.01 hPa to the surface.
- Winds, temperature and surface pressure are interpolated in the ECMWF operational 6-hourly analyses.
- Time steps of 20 minutes, output every 3 hours

See the stratospheric ozone service at <http://www.copernicus-stratosphere.eu/>.

It delivers graphical products dedicated to stratospheric composition and allows easy comparison between the results of o-suite, BASCOE and TM3DAM. The BASCOE data products (HDF4 files) are also distributed from this webpage. Other details and bibliographic references on BASCOE can be found at <http://bascoe.oma.be/>. A detailed change log for BASCOE can be found at http://www.copernicus-stratosphere.eu/4_NRT_products/3_Models_changelogs/BASCOE.php.

2.3.2 TM3DAM and the multi-sensor reanalysis

One of the MACC products was a 30-year reanalysis, near-real time analysis and 10-day forecast of ozone column amounts performed with the KNMI TM3DAM data assimilation system, the Multi-Sensor Reanalysis (MSR) system (van der A et al., 2010, 2013), http://www.temis.nl/macc/index.php?link=o3_msr_intro.html.

The corresponding validation report can be found at http://www.copernicus-atmosphere.eu/services/gac/global_verification/validation_reports/.

The NRT TM3DAM product used for the validation of the CAMS NRT streams is the ozone analysis of Envisat/SCIAMACHY (until April 2012), AURA/OMI, and MetOp-A/GOME-2, run in the following configuration:

- total O₃ columns are assimilated
- Global horizontal grid with a 3° longitude by 2° latitude resolution.
- Vertical grid is hybrid-pressure and consists in 44 levels extending from 0.1 hPa to 100 hPa.
- Dynamical fields from ECMWF operational 6-hourly analysis.

An update of the MSR (MSR-2) was presented in van der A et al. (2015), which extended the record to 43 years based on ERA-interim reanalysis meteo and with an improved resolution of 1x1 degree.



2.3.3 SDS-WAS multimodel ensemble

The World Meteorological Organization's Sand and Dust Storm Warning Advisory and Assessment System (WMO SDS-WAS) for Northern Africa, Middle East and Europe (NAMEE) Regional Center (<http://sds-was.aemet.es/>) has established a protocol to routinely exchange products from dust forecast models as the basis for both near-real-time and delayed common model evaluation. Currently, nine (BSC-DREAM8b, MACC-ECMWF, DREAM-NMME-MACC, NMMB/BSC-Dust, NASE GEOS-5, NCEP NGAC, EMA_RegCM, DREAMABOL and NOAA) provides daily operational dust forecasts (i.e. dust optical depth, DOD, and dust surface concentration).

Different multi-model products are generated from the different prediction models. Two products describing centrality (multi-model median and mean) and two products describing spread (standard deviation and range of variation) are daily computed. In order to generate them, the model outputs are bi-linearly interpolated to a common grid mesh of $0.5^\circ \times 0.5^\circ$. The multimodel DOD (at 550 nm) Median from nine dust prediction models participating in the SDS-WAS Regional Center is used for the validation of the CAMS NRT streams.

2.4 CAMS products

An extended list of output products from the NRT stream o-suite are available as 3-hourly instantaneous values up to five forecast days. These are available from ECMWF (through ftp in grib2 and netcdf format, <http://atmosphere.copernicus.eu/global-near-real-time-data-access>).

2.5 Availability and timing of CAMS products

Table 2.6: Timeliness of the o-suite from March 2013 – to February 2016

Months	On time, 22 utc	80th perc	90th perc	95th perc
March-May 2013	97%	D+0, 17:54	D+0, 18:36	D+0, 18:49
June-August 2013	97%	D+0, 18:34	D+0, 18:46	D+0, 19:23
Sept-Nov 2013	99%	D+0, 19:14	D+0, 19:22	D+0, 19:29
Dec-Feb '13-'14	94%	D+0, 19:45	D+0, 20:40	D+0, 21:55
Mar-May 2014	98%	D+0, 19:44	D+0, 19:57	D+0, 20:03
Jun-Aug 2014	95%	D+0, 20:03	D+0, 20:57	D+0, 22:43
Sept-Nov 2014	96%	D+0, 19:24	D+0, 20:31	D+0, 21:14
Dec-Feb '14-'15	97%	D+0, 19:43	D+0, 20:28	D+0, 21:13
Mar-May 2015	96%	D+0, 19:38	D+0, 21:03	D+0, 21:40
Jun-Aug 2015	95%	D+0, 20:24	D+0, 20:53	D+0, 21:54
Sept-Nov 2015	95%	D+0, 19:44	D+0, 20:55	D+0, 21:51
Dec-Feb '15-'16	100%	D+0, 18:39	D+0, 18:57	D+0, 19:43
Mar-May 2016	98%	D+0, 19:32	D+0, 19:47	D+0, 20:00
Jun-Aug 2016 (00 and 12 cycle)	100%	D+0, 08:53 D+0, 20:55	D+0, 09:04 D+0, 21:01	D+0, 09:18 D+0, 21:18



The availability statistics provided in Table 2.6 are computed for the end of the 5-day forecast run, and are obtained from July 2012 onwards. A forecast is labeled "on time", if everything is archived on MARS before 22UTC. This is based on requirements from the regional models. We note that at present most regional models can still provide their forecasts even if the global forecast is available a bit later.

For the period June-August 2016, 100% of the forecasts were delivered before 22:00.



3. Validation results for reactive gases and aerosol

This section describes the validation results of the CAMS NRT global system (the o-suite) for reactive gases and aerosol up to February 2016. The validation focuses on the results from the NRT analysis (or D+0 FC) stream. For a selection of instances 2-4 day forecasts issued from them have been explicitly considered. Naming and color-coding conventions predominantly follow the scheme as given in Table 3.1.

Table 3.1 Naming and color conventions as adopted in this report.

Name in figs	experiment	Color
{obs name}	{obs}	black
o-suite D+0 FC	0001	red
Control	geuh, gjjh	blue

3.1 Tropospheric Ozone

3.1.1 Validation with sonde data in the free troposphere

Model profiles of the CAMS runs were compared to free tropospheric balloon sonde measurement data of 38 stations taken from the NDACC, WOUDC, NILU and SHADOZ databases for August 2015 to August 2016 (see Fig. 3.1.1 - 3.1.3). Towards the end of the period, the number of available soundings decreases, which implies that the evaluation results may become less representative. The figures contain the number of profiles in each month that are available for the evaluation. The methodology for model comparison against the observations is described in Annex 2 in CAMS VAL report #1. The free troposphere is defined as the altitude range between 750 and 200 hPa in the tropics and between 750 and 300 hPa elsewhere.

In all zonal bands MNMBs for the o-suite are within the range -20 to +30%, for all months, see Fig. 3.1.1-3.1.3. The control run shows larger negative MNMBs for Antarctica (up to -38%). Over the Arctic, the o-suite mostly shows slightly positive MNMBs during summer and spring (MNMBs up to 6%), while during the winter season the MNMBs get negative (within -13%) see, Fig. 3.1.1. The o-suite shows lower MNMBs than the control run during the last six months. Over the NH mid-latitudes MNMBs for the o-suite are on average close to zero all year round (maxima are -10% to +3%), which is a clear improvement compared to the control run, which shows larger MNMBs (up to $\pm 13\%$) during the respective period. MNMBs for the o-suite are generally larger (up to 30%) over Antarctica, where tropospheric O₃ values are comparatively lower than over the polluted NH. For the Tropics, MNMBs are between $\pm 20\%$ for the o-suite and the control run.

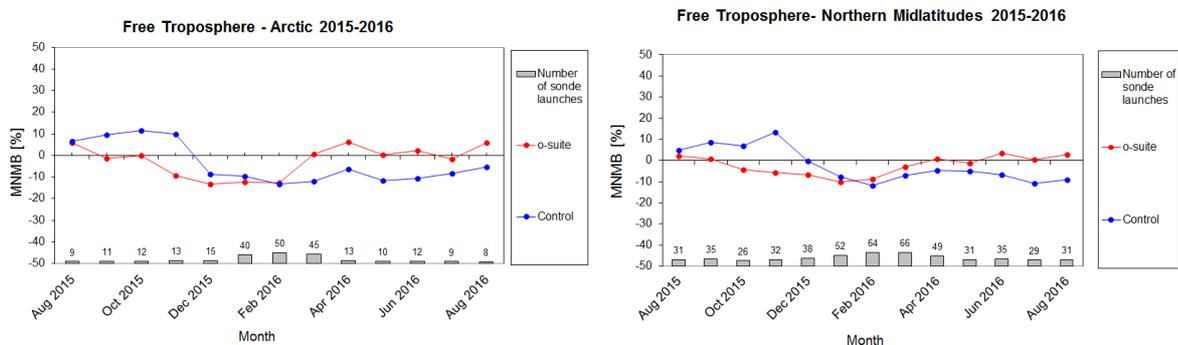


Figure 3.1.1: MNMBs (%) of ozone in the free troposphere (between 750 and 300 hPa) from the IFS model runs against aggregated sonde data over the Arctic (left) and the Northern midlatitudes (right). The numbers indicate the amount of individual number of sondes.

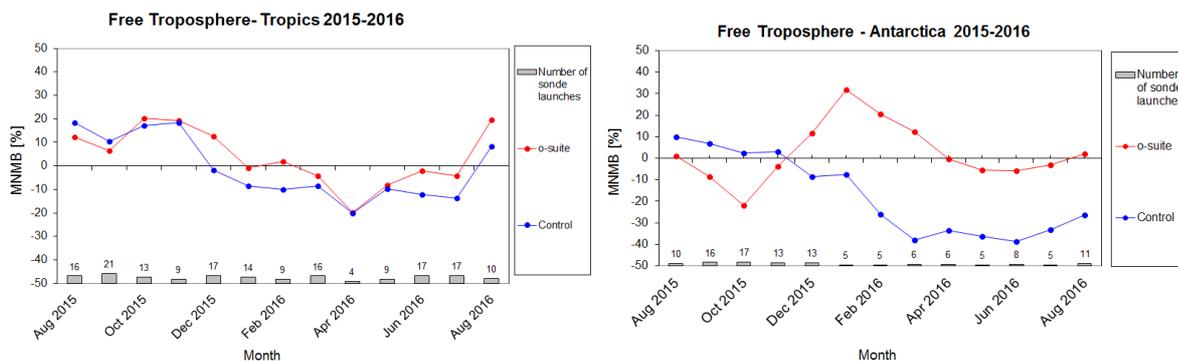


Figure 3.1.2: MNMBs (%) of ozone in the free troposphere (between 750 and 200 hPa (Tropics) / 300 hPa) from the IFS model runs against aggregated sonde data over the Tropics (left) and Antarctica (right). The numbers indicate the amount of individual number of sondes.

3.1.2 Ozone validation with IAGOS data

The daily profiles of ozone measured at airports around the world, are shown on the website at http://www.iagos.fr/macc/nrt_day_profiles.php. For the period from June 2016 to August 2016, the data displayed on the web pages and in this report include only the data as validated by the instrument PI. The available flights and available airports are shown in Fig. 3.1.3 top and bottom respectively. Performance indicators have been calculated for different parts of the IAGOS operations.

A new aircraft was equipped during this period, taking the fleet to 7. However since one aircraft has not been providing data for some time, we will consider the fleet to remain at 6. With the 6 aircraft, operating fully over the three month period, we can expect a total of about 1260 flights. The actual number of flights within the period was 859 (1718 profiles) giving a performance of 68%. The actual number of flights with usable data was 417 (49% of the total possible). These flights are shown in Fig. 3.1.3 (top). Fifty percent (50%) (864 profiles) of the operational flights had usable measurements of ozone and 30% of flights had usable CO. Delivering these O₃ and CO data are two aircraft from China Airlines based in Taipei, an aircraft operated by Air France based in Paris and an aircraft operated by Lufthansa based in Frankfurt. This report therefore displays profiles recorded by these aircraft, covering mainly the routes served by Air France to North America and West Africa

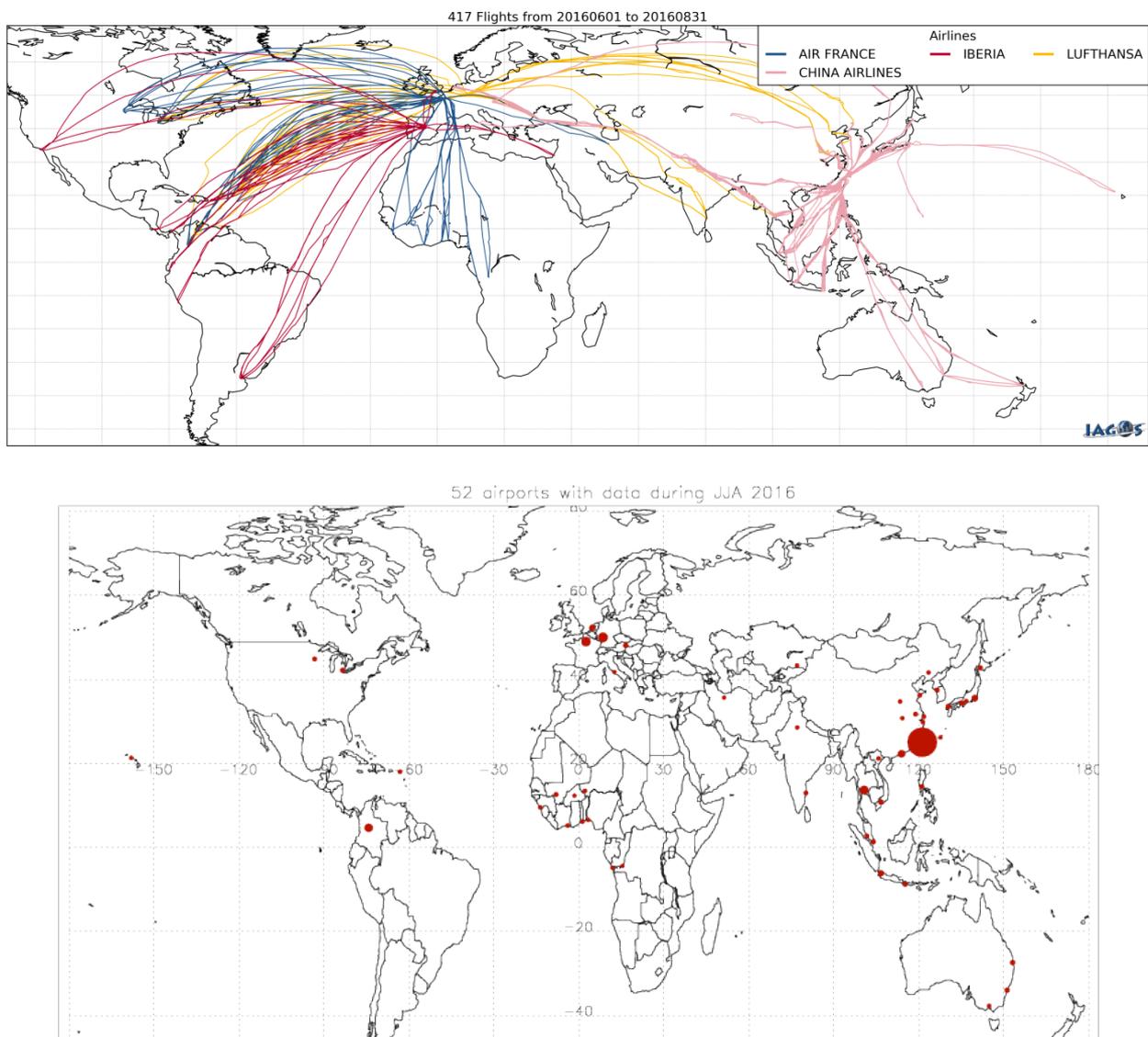


Figure 3.1.3: Map of the flights (top) and the visited airports (bottom) during the period JJA 2016, by the IAGOS equipped aircraft. The size of the plotting circle represents the number of profiles available.

and by China Airlines across South-East Asia as shown on the map in Fig. 3.1.3 (with a plotting circle scaled to the highest number of flights at an airport). Data are also available in Australia and in New Zealand.

Europe

Figure 3.1.4 presents ozone at Frankfurt and Paris during June, July and August 2016. Ozone is overestimated in the upper troposphere over both Paris and Amsterdam but in the other atmospheric layers the models perform well. During a brief heatwave around July 21st the ozone is underestimated at Frankfurt. This is discussed in more detail below.

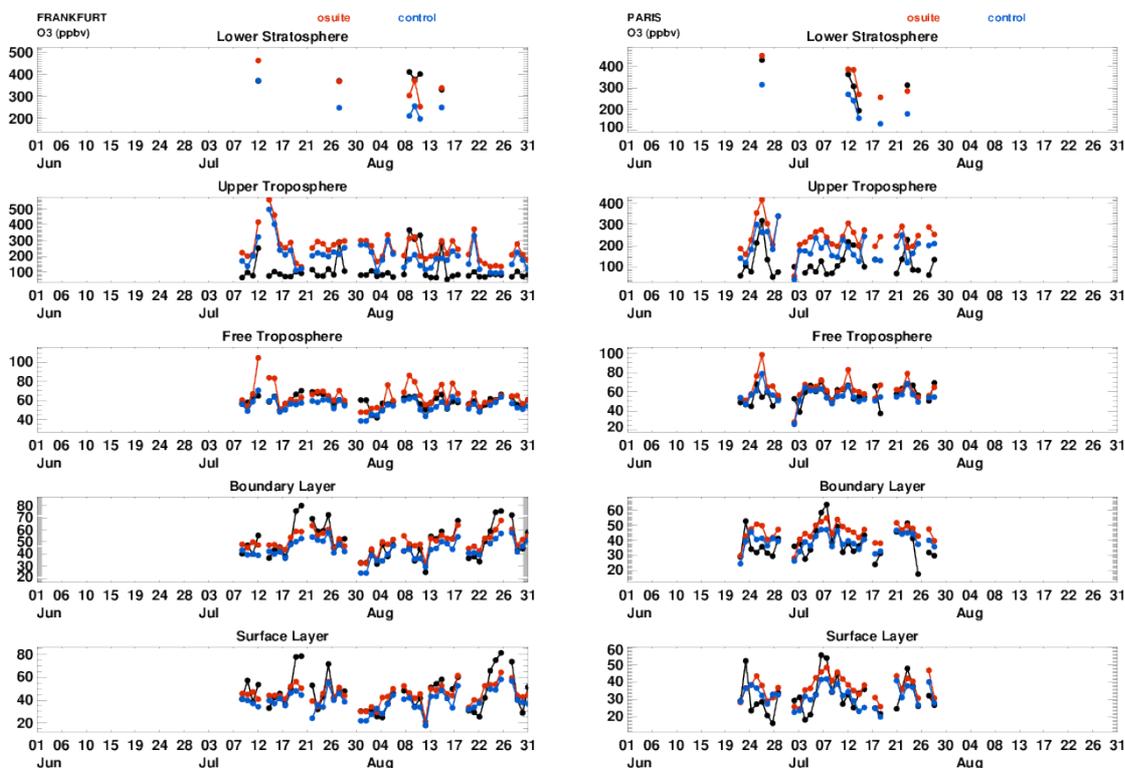


Figure 3.1.4: Time series of daily mean ozone over Frankfurt (left) and Paris (right) during June, July, August 2016 for 5 layers, Surface, Boundary layer, Free Troposphere, Upper Troposphere and Lower Stratosphere.

The examples in Fig. 3.1.5 show profiles over Amsterdam, Paris and Vienna. On 5th July there are profiles at Paris by Air France and at Amsterdam by China Airlines. The profiles from the two aircraft are quite similar. In both cases the control run does better than the o-suite throughout the profile. In the surface layer, ozone is overestimated by the two runs but the control run does slightly better than the o-suite. In both cases the control run does much better than the o-suite at capturing the profile around the tropopause. At about 3000m there is a small peak in the concentration of ozone seen at both airports but missed by the two runs. At Paris on 15th July, Amsterdam on 7th July and Vienna on 26th July the o-suite performs better than the control run in the UTLS, which is the opposite to that which is usually observed.

The two runs underestimate the peak in ozone seen at Frankfurt around 20th July (Fig. 3.1.6) where ozone levels reached almost double their mean value of 40ppbv calculated from 10 years of MOZAIC measurements. At the same time surface temperatures were also about 10 degrees higher than the 10 year mean. A similar underestimation of ozone was seen at Amsterdam on 21st July throughout the boundary and surface layers linked with the mini heatwave (Fig. 3.1.6; see also regional report).

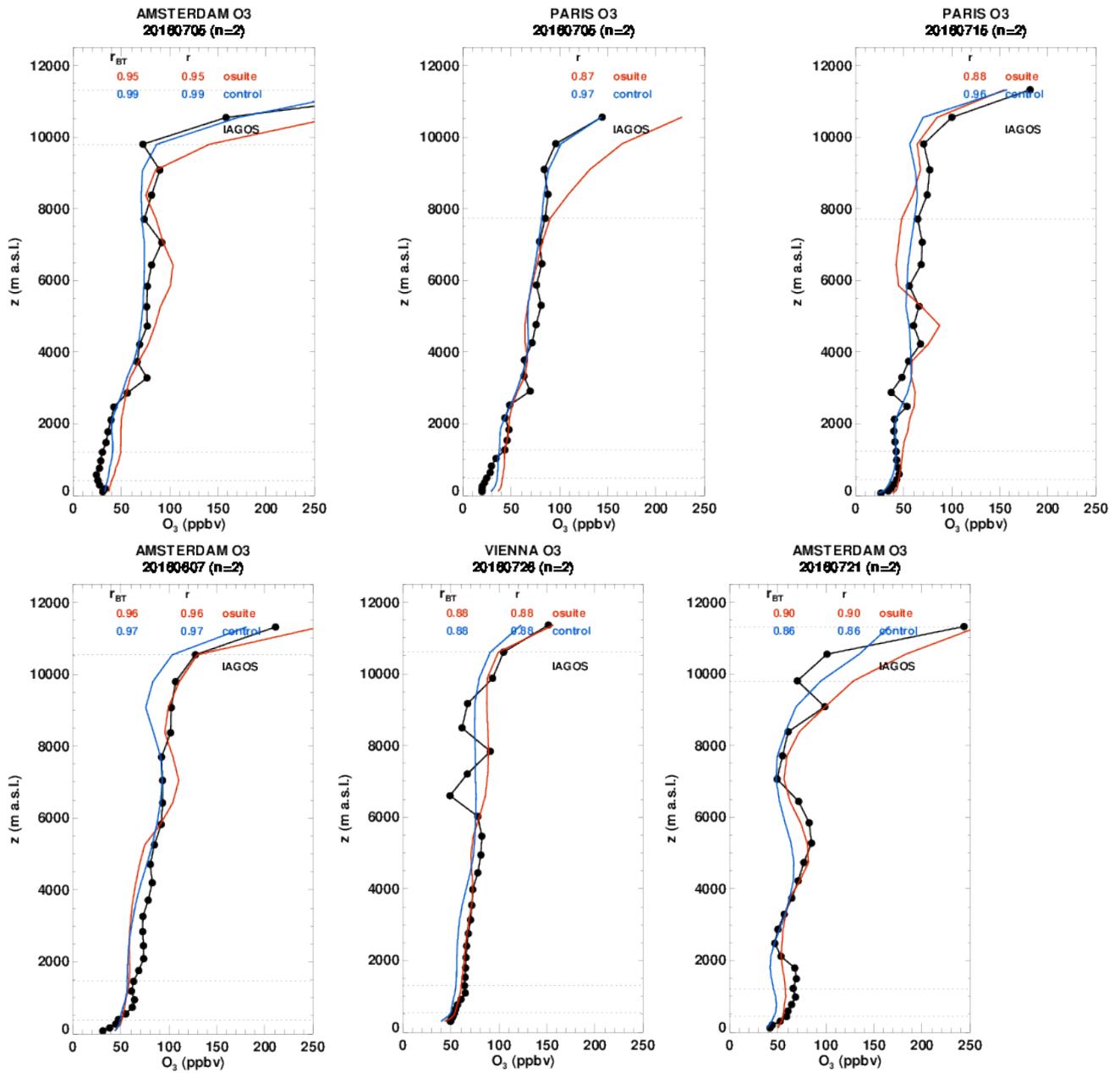


Figure 3.1.5: Selection of daily profiles of ozone from IAGOS (black) and the o-suite (red) and control (blue) over Europe (Paris, Amsterdam and Vienna) over the period June-August 2016.

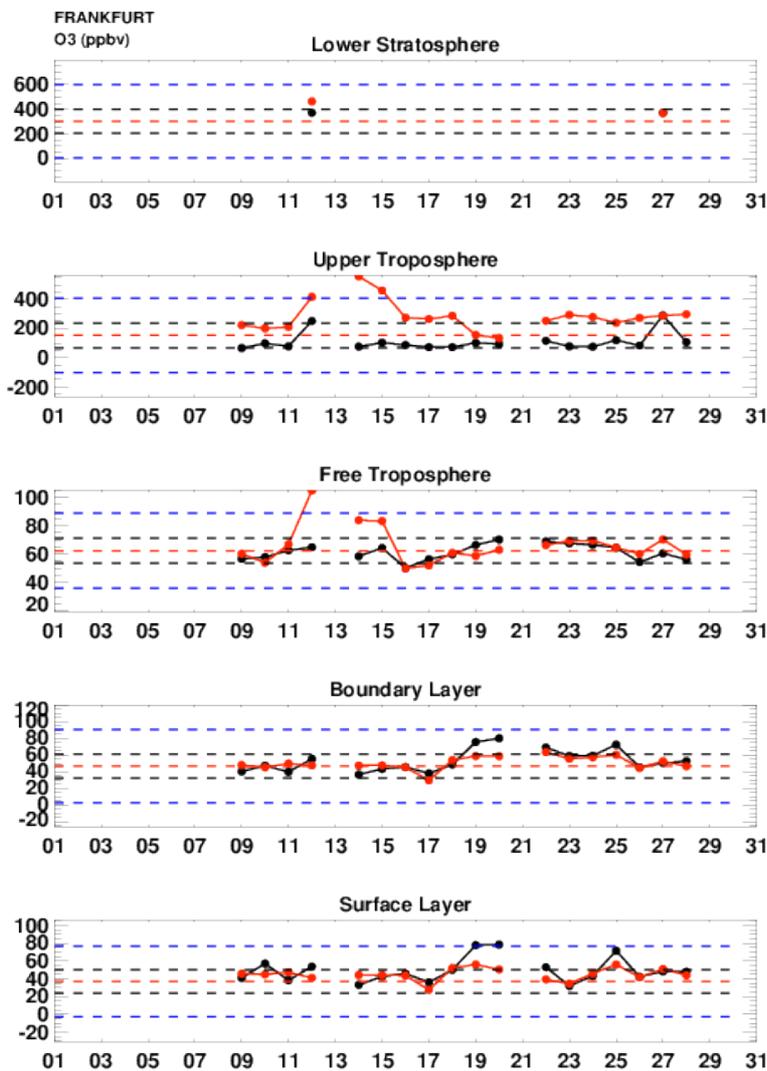


Figure 3.1.6: Time series for July 2016 of IAGOS observations (Black) and osuite (red) along with the mean Ozone calculated from 10 years of MOZAIC data (2002-2012) (red dashed line) and the standard deviation (black dashed line) and the three sigma line (blue dashed).

Asia

As the timeseries in Fig. 3.1.7 shows, the o-suite and the control overestimate ozone in all atmospheric layers (the lower stratosphere is not reached by the aircraft during this period). The profiles in Fig. 3.1.8 for Taipei show this in more detail, and clearly show that the o-suite improves upon the control throughout the profile. In the free troposphere, the bias between the runs and the observations is much reduced compared with the boundary and surface layers where the models are predicting three times the concentration that is observed. At Singapore, Ho Chi Minh City and Kuala Lumpur the models compare better with the observations.

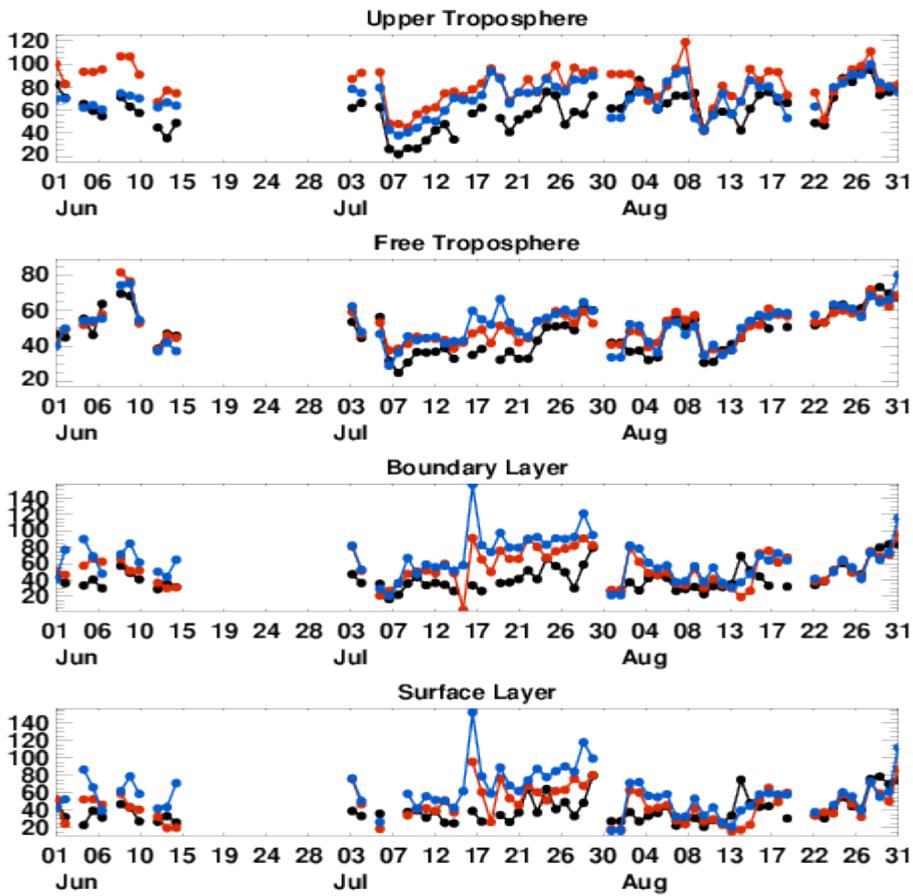
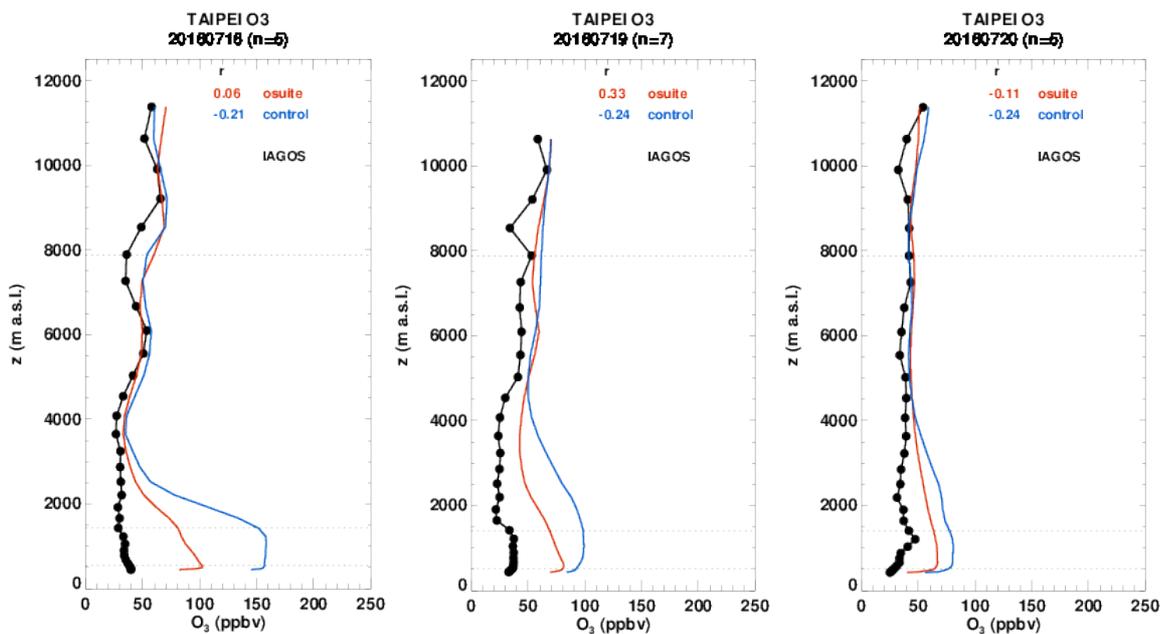


Figure 3.1.7: Time series of daily mean ozone over Taipei during JJA 2016 for 4 layers, Surface, Boundary layer, Free Troposphere, Upper Troposphere.



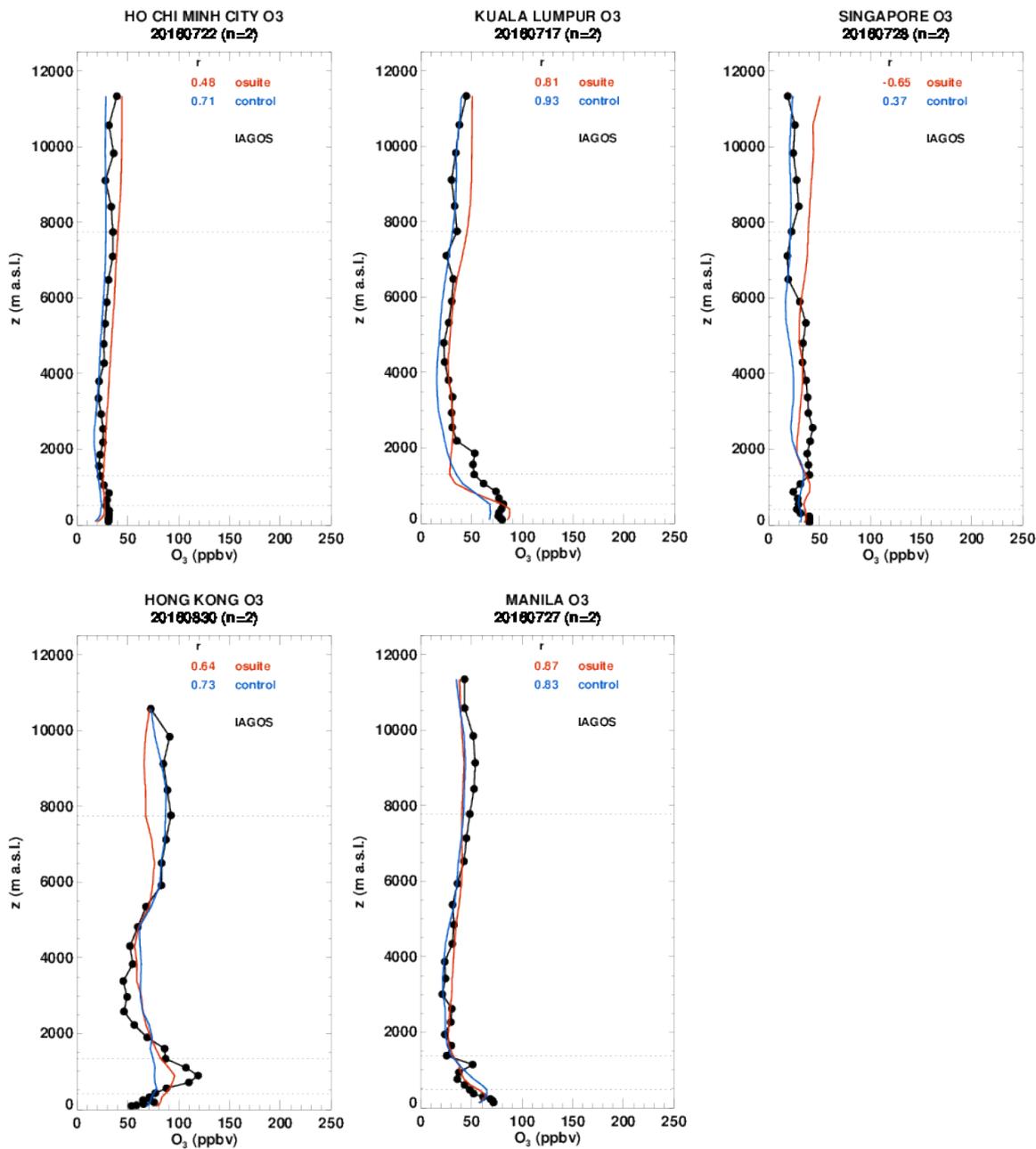


Figure 3.1.8: Selection of daily profiles of ozone from IAGOS (black) and the o-suite (red) and control (blue) over Asia (Taipei, Hong Kong, and Kuala Lumpur) over the period JJA 2016.

Typhoons Nepartak and Nida

Typhoons have contrasting effect on the profiles observed at Taipei. Typhoon Nepartak hit Taiwan directly on the 8th July. Extremely low ozone mixing ratios (<20ppbv) were observed throughout the profile (Fig. 3.1.9), along with very low CO (<80ppbv; Fig. 3.3.8) corresponding with the oceanic trajectory of the typhoon bringing clean maritime air from the Pacific Ocean. The mixing ratios in the clean air are overestimated by the models. In contrast, from 1st -4th August, Typhoon Nida passed to the south of Taipei, having already passed over the Philippines on 30th July. An ozone anomaly (Fig. 3.1.9) was observed at altitudes around 8000 which is really well captured by the two

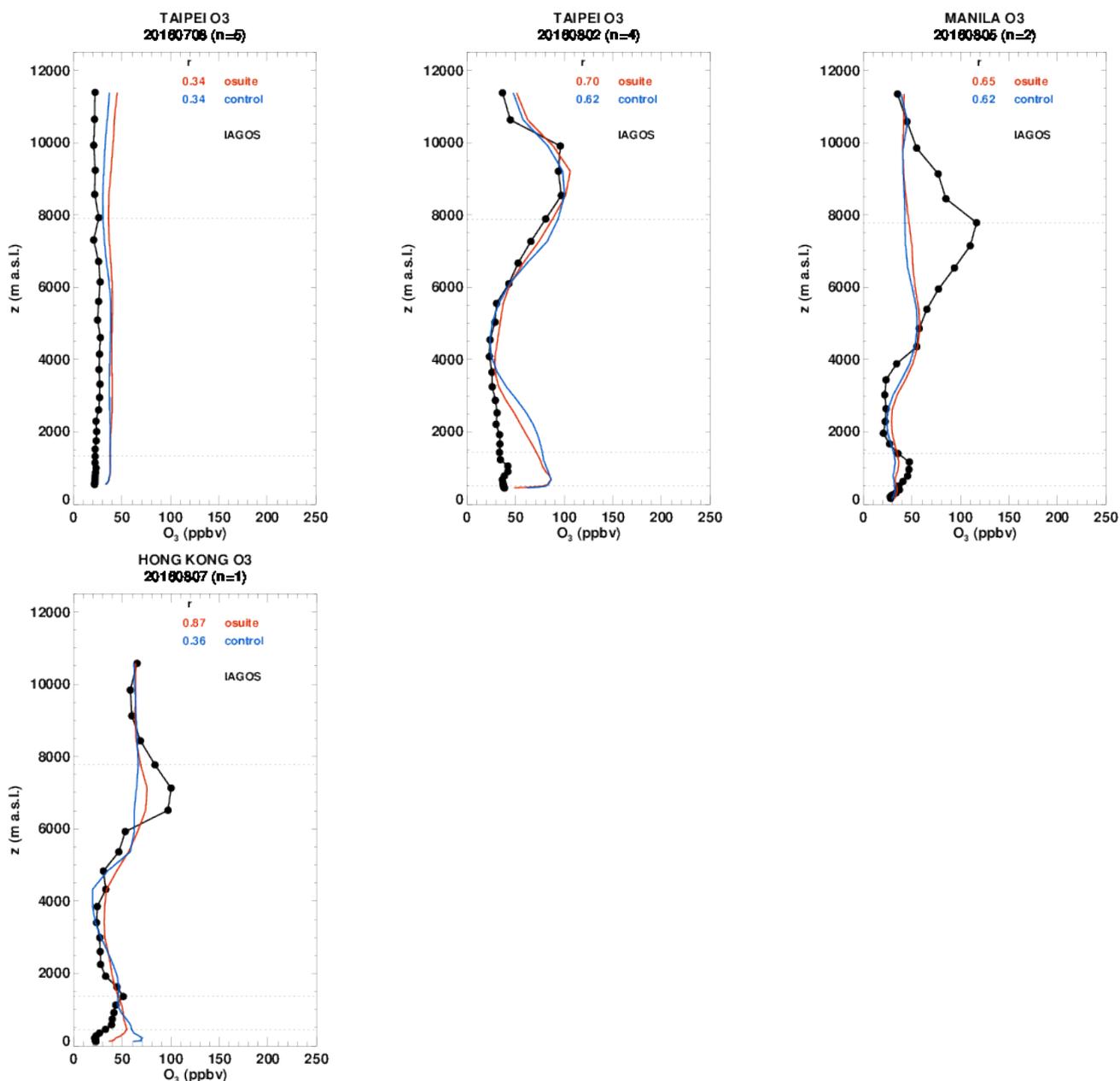


Figure 3.1.9: Selection of daily profiles of ozone from IAGOS during and following the passage of typhoon Nida, which made landfall on 2nd August 2016.

models. It coincided with an increase of CO (Fig.3.3.8). The ozone anomaly was also present over Manila and over Hong Kong, where it was less well simulated by the models. The signature of this anomaly is more likely related to the outflow of Typhoon Nida, which has picked up polluted boundary layer air from the coast of China and the Philippines. Typhoon Nida made landfall near Hong Kong on the 2nd August.

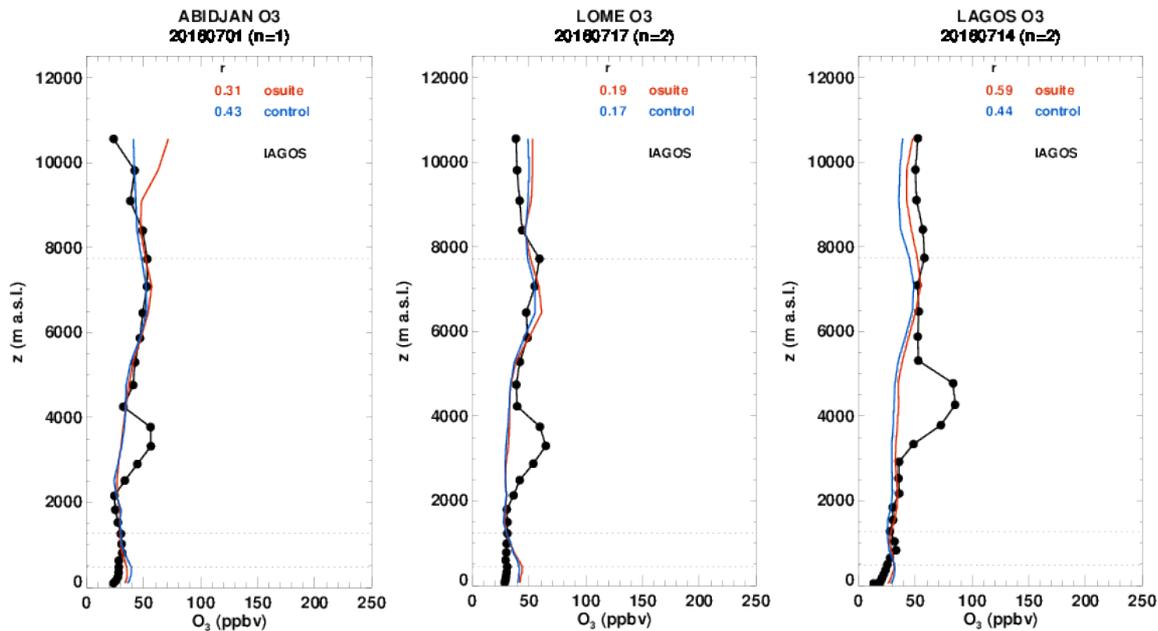


Figure 3.1.10: Profiles of ozone from IAGOS (black) and the two NRT runs over Abidjan, Lome and Lagos in July 2016.

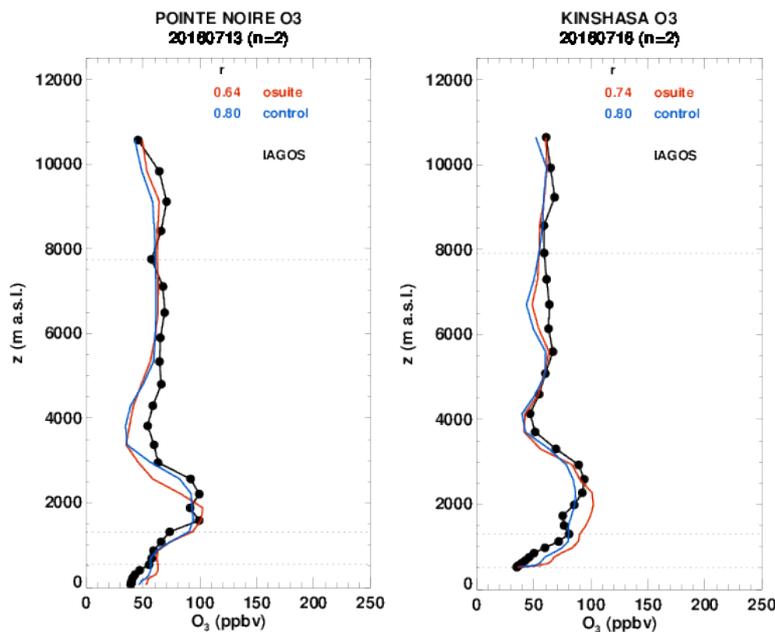


Figure 3.1.11: Profiles of ozone from IAGOS (black) and the two NRT runs over Pointe Noire and Kinshasa in July 2016.

West-Africa

Air France visited several destinations in West Africa (Bamako, Lagos, Abijan, Lome, Conakry and Ouagadougou). These cities are influenced by anthropogenic emissions from vehicles, and biomass burning from December to March. Over Abijan, Lome and Lagos, there is a peak in ozone at around 3000m related to enhanced CO at the same altitude and probably due to the transport of polluted

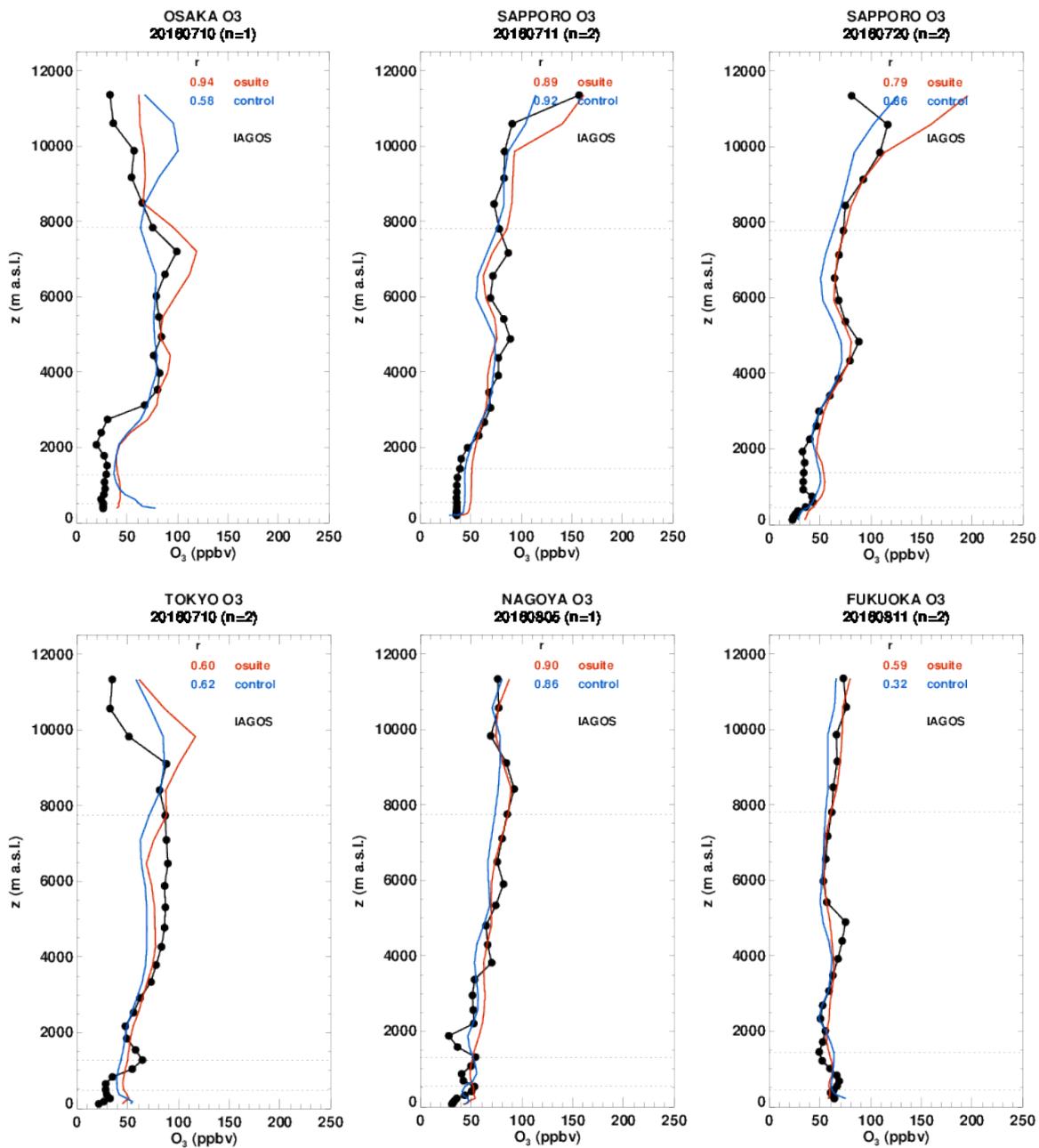


Figure 3.1.11: Profiles of ozone from IAGOS (black) and the two NRT runs over Japan.

air masses from south of the equator where biomass burning is very active. We see the effects of this intense biomass burning on the profiles at Pointe-Noire and Kinshasa in the Southern Hemisphere where the models do well at capturing enhanced ozone at around 2000m. The fact that this plume of ozone is underestimated in the northern hemisphere suggests that the interhemispheric transport of the biomass burning plumes is therefore not so well captured by the models.

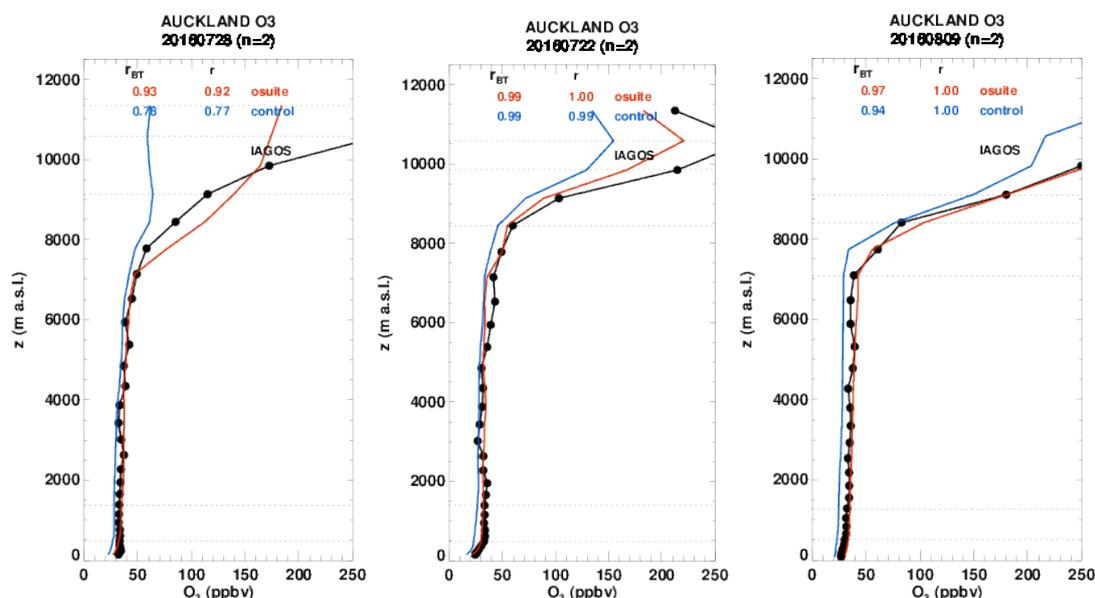


Figure 3.1.12: Profiles of ozone from IAGOS (black) and the two NRT runs over Auckland, New Zealand.

Japan

Thanks to the new China Airlines equipped in July, we now have profiles at a variety of cities across Japan, from Sapporo at 43°N North to Fukuoka and 33°N. The cities shown are all densely populated urban areas. Whilst the cities span a latitude range of 10°, all the cities are downstream of pollution emanating from China. The model runs compare well with the observed profiles, particularly at Sapporo, though in all cases the models overestimate the observed ozone concentrations in the surface layer. The ozone profile at Osaka seems to be more difficult for the models to capture.

New Zealand

The new China Airlines aircraft has also been providing profiles in Auckland, New Zealand (36°S, 174°E) for the first time in IAGOS. Here we would expect the concentrations of ozone to be the closest to background concentrations due to the low population density of New Zealand, its remoteness from other continents, and the arrival of clean maritime air from the Tasman Sea and the Pacific Ocean. The profiles show that the models do well at capturing the profile of ozone until the UTLS region. In the UTLS, the gradients are more difficult to capture, but the behaviour of the two runs are generally better than at Frankfurt and the o-suite seems to be better than the control.

3.1.3 Validation with GAW and ESRL-GMD surface observations

For the Near Real Time (NRT) validation, 13 GAW stations and 11 ESRL stations are currently delivering O₃ surface concentrations in NRT, and the data are compared to model results. In the following, a seasonal evaluation of model performance for the 2 NRT runs (o-suite and control) has been carried out for the period from June to August 2016. The latest validation results based on GAW stations can be found on the CAMS website,

<http://www.copernicus-atmosphere.eu/d/services/gac/verif/grg/gaw/>,

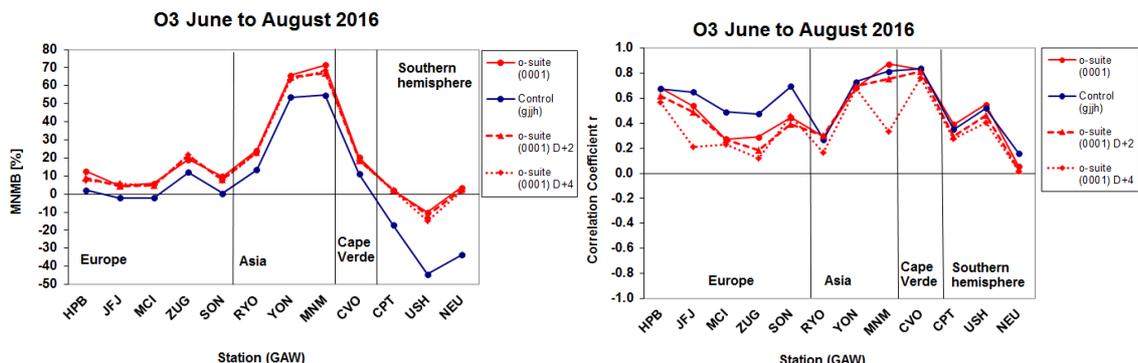


Figure 3.1.13: Modified normalized mean bias in % (left) and correlation coefficient (bottom right) of the NRT model runs compared to observational GAW data in the period March to May 2016. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.

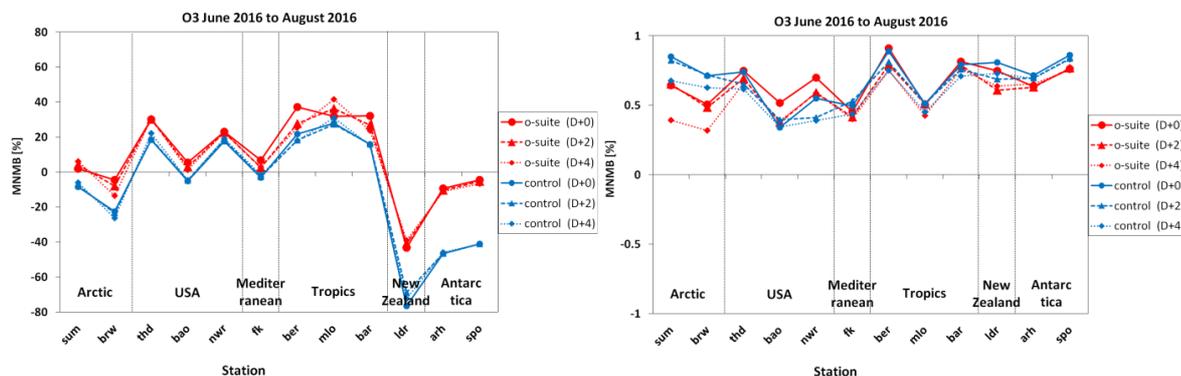


Figure 3.1.14: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT forecast runs compared to observational ESRL data in the period June to August 2016. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.

and based on ESRL on <http://www.academyofathens.gr/kefak/cams/index.html>. Results are summarized in Figs 3.1.13 and 3.1.14.

Modified normalized mean biases in % (left, panel) and correlation coefficients (right, panel) for different forecasts days (D+2, triangles and D+4, rhombs) with respect to GAW observations are shown in Fig. 3.1.13 (left). It indicates that MNMBs for both o-suite and control run remain stable till the D+4 (forecast run from 96h to 120h). Similar results concerning MNMBs stability are found for ESRL observations (Fig. 3.1.14). Correlations between simulated and observed surface ozone values remain almost stable till D+2 (forecast run from 48h to 72h), but then drop (correlations for D+4 are lower than correlations for D+2 and D+0), except for stations in Antarctica and the Tropics see Fig. 3.1.13 and 3.1.14, right graph).

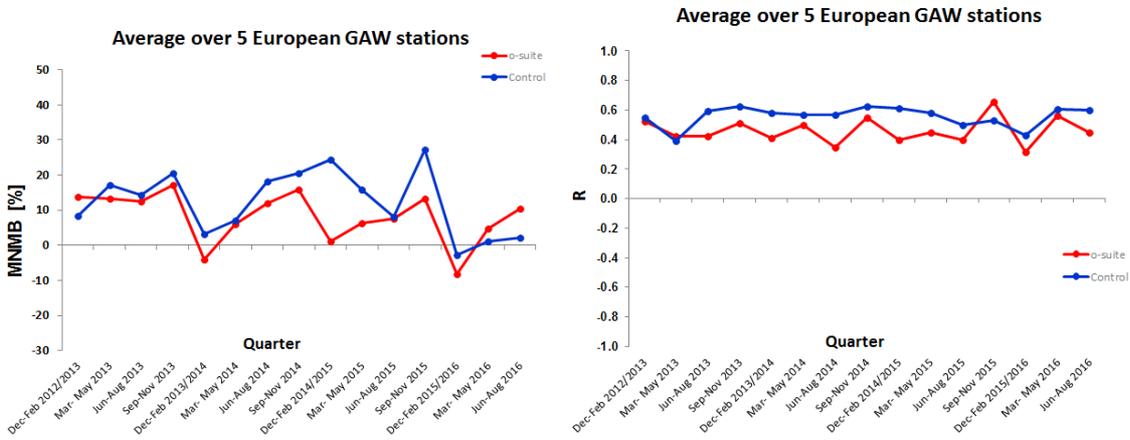


Figure 3.1.15: Long term (Dec. 2012 – Aug. 2016) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

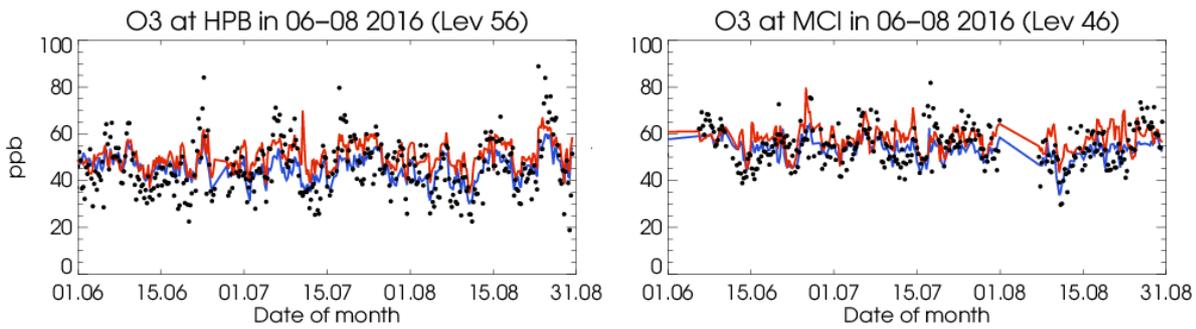


Fig. 3.1.16: Time series for the o-suite (red) and Control (blue) compared to GAW observations at Hohenpeissenberg (44.2°N, 10.7°E) and Monte Cimone (44.18°N, 10.7°E).

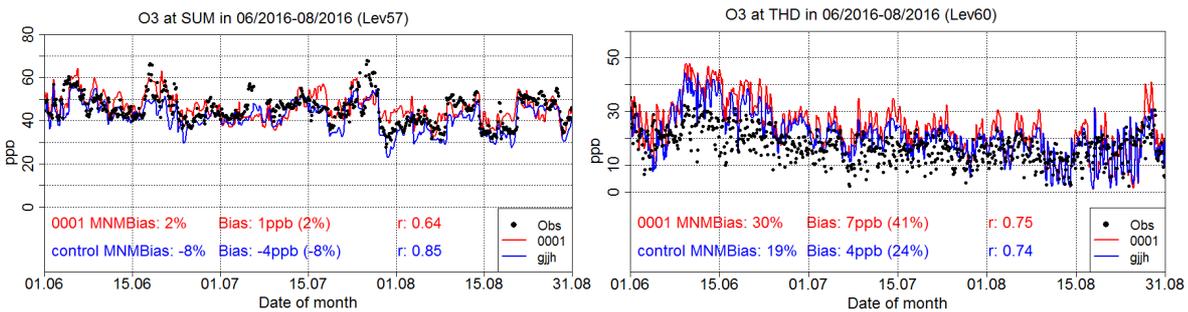


Figure 3.1.17: Time series for the o-suite (red) and control (blue) compared to ESRL observations at Summit, Greenland station (72.57°N, 38.38°W, left) and Trinidad Head, California station (41.05°N, 124.15°W, right).

A comparison of the seasonal-mean MNMB over Europe (Fig. 3.1.15) from December 2012 to present shows that the MNMB over European GAW stations is minimal during the winter season, and tends to increase in other months. Also on average the MNMB for the o-suite shows a slight improvement over the years, while it remains higher, and more variable for the consecutive control runs. Temporal correlation is consistently better for control than for the o-suite.

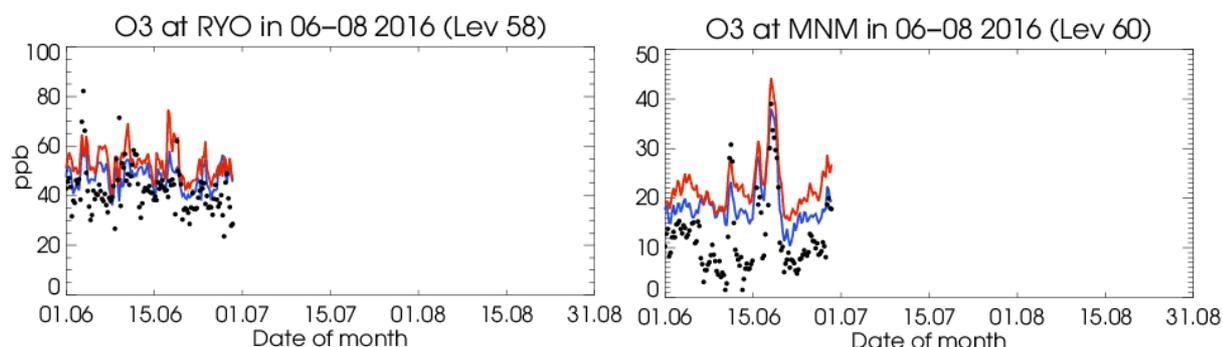


Figure 3.1.18: Time series for the o-suite (red) and Control (blue) compared to GAW observations at Ryori (39.03°N, 141.8°E) and Minamitorishima (24.3°N, 123.9°E).

Looking at different regions, for European stations (HPB, JFJ, ZUG, SON, MCI), observed O₃ surface mixing ratios are mostly slightly overestimated by the o-suite, and partly underestimated by the control run, with MNMBs between 4 to 20% for the o-suite and between -2 and 12% for the control run (see also Fig. 3.1.19). Correlations for the European stations are between 0.27 to 0.67 for the o-suite and between 0.47 and 0.69 for the control run. The time series plots show that especially minimum concentrations are partly not resolved by the model, see Fig. 3.1.16.

In the Arctic, the o-suite reproduces well surface ozone mean concentrations over Summit (SUM, MNMBs≈5%) and Point Barrow (BRW, MNMBs≈-5%) while the control run underestimates it by -10% and -25% respectively. Correlations between simulated and observed surface ozone in both stations are high for the o-suite ($r>0.5$) and even higher for the control run ($r>0.7$).

For USA stations (THD, BAO, NWR) both runs overestimate surface ozone mean concentrations at THD (o-suite MNMBs≈30%, control MNMBs≈20%) and at NWR (both runs MNMBs≈20%) while at BAO both runs reproduce well surface ozone mean. The o-suite run reproduces better the day to day surface ozone variability ($r>0.7$ at THD and at NWR and $r>0.5$ at BAO) than the control run ($r>0.7$ at THD, $r\approx 0.55$ at NWR and $r\approx 0.35$ at BAO).

For Asian stations (RYO, YON, MNM), both runs overestimate the low observed ozone concentrations with MNMBs of up to 70%. Concentration peaks are well reproduced, however, as can be seen in Fig. 3.1.18. Correlations are between 0.3 and 0.8 for both runs.

Over the tropical stations (BAR, BER, MLO) both runs overestimate ozone mixing ratios and the o-suite has a higher positive offset than the control run (o-suite MNMBs≈30% and control MNMBs≈20%). The overestimation mostly concerns the minimum concentrations in the model see also Fig. 3.1.19). Correlations between simulated and observed surface ozone concentrations are high in all stations ($r>0.8$ at BAR and BER and $r>0.5$ at MLO).

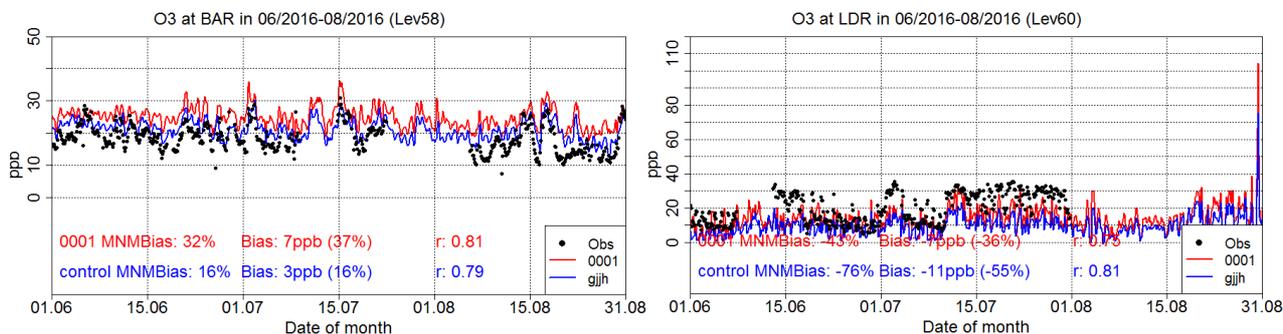


Figure 3.1.19: Time series for the o-suite (red) and control (blue) compared to ESRL observations (black dots) at Ragged Point, Barbados station (13.17°N, 59.46°W) and at Lauder (45.04°S, 169.68°E).

For the stations in the Southern Hemisphere (CPT, LDR, USH) the o-suite reproduces ozone mixing ratios well for CPT and USH with MNMBs between -10% and 4% and -40% for LDR. The data assimilation corrects the negative offset in the control run, see Fig. 3.1.19 and 3.1.20 (right panel). Correlation coefficients are between 0.4 and 0.7 for both runs.

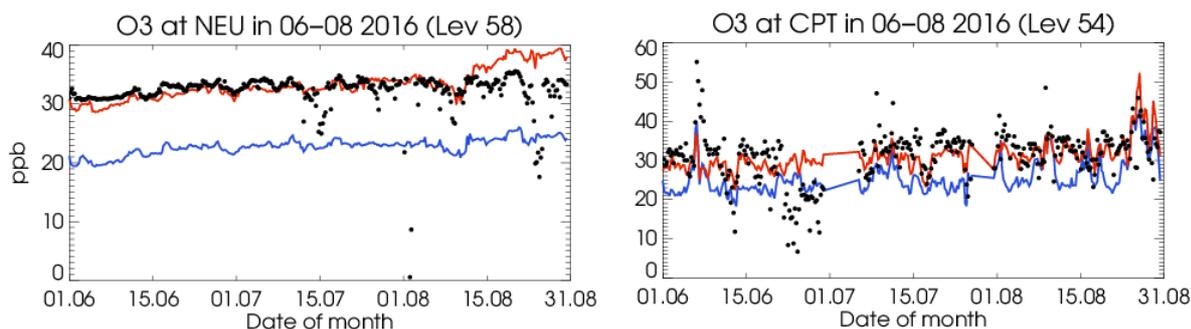


Figure 3.1.20: Time series for the o-suite (red) and control (blue) compared to GAW observations (black dots) at Neumayer (70.7°S, 8.3°W) and GAW observations at Cape Point (34.35°S, 18.48°E).

Finally over Antarctica stations (SPO, ARH, NEU) the o-suite reproduces well ozone mixing ratios with MNMBs between -10% and 4%, see Fig. 3.1.20 (left). The data assimilation corrects the negative offset in the control run (with MNMBs of up to -50%). Correlations between simulated and observed surface ozone over SPO and ARH stations are high for both runs ($r > 0.65$), but for NEU correlations are very low for both runs.

3.1.4 Validation with AirBase observations in Mediterranean

The surface ozone validation analysis over the Mediterranean is based on an evaluation against station observations from the Airbase Network (<http://acm.eionet.europa.eu/databases/airbase/>). In addition, 3 stations from the Department of Labour Inspection - Ministry of Labour and Social Insurance, of Cyprus (<http://www.airquality.dli.mlsi.gov.cy/>) as well as the Navarino Environmental Observatory (<http://www.navarinoneo.gr/index.php/en/>) station in Messene Greece are used in the validation analysis. For the validation analysis, stations in the Mediterranean located within about 100 km from the shoreline of the Mediterranean shore are used. Table 3.1.1 shows the names, coordinates, elevation and the MNMBs and correlations obtained with the 2 forecast runs (o-suite and control). It indicates that the variance explained by each station of both the o-suite and



control is high and correlations are highly significant in the West and Central Mediterranean, with the exception of Bc-La Senia station in Spain. On the contrary over stations in Greece and Cyprus in the eastern Mediterranean, correlations between modelled and simulated surface ozone values are low particularly for the o-suite run with the exception of Finokalia Station in Crete where correlations for both runs are high. It should be noted that the control run mostly reproduces slightly better the day to day variability than the o-suite run (see Table 3.1.1).

In terms of biases, both runs' MNMBs vary between -20% and 20% over Spain (for Stations Ak-pardines and Hospital Joan March o-suite MNMB exceed 25%). Over the Mediterranean shore of Spain data assimilation seems to improve the MNMBs. In all other Mediterranean stations (Plan Aups/Ste Baume in France, Gharb in Malta, Aliartos, NEO and Finokalia in Greece, Ineia, Oros Troodos and Agia Marina in Cyprus) both the o-suite and the control reproduce well surface ozone mean concentrations (-10%<MNMBs<10%; see also Fig. 3.1.22, central and lower graphs).

The spatial distribution of MNMBs and correlations of the o-suite over the Mediterranean is shown in 3.1.22, where the contrast in the model MNMBs between Mediterranean shore of Spain (higher deviations from 0) and Central and Eastern Mediterranean (MNMBs close to zero) is evident. On the other hand it clearly shows the contrast between Western and Central Mediterranean where higher correlations between modelled and simulated surface ozone values were observed, and Eastern Mediterranean where lower correlations were observed during summer 2016.

Table 3.1.1: Coordinates, elevation, corresponding model level (level 60 is the surface level), as well as validation scores (MNMBs and correlations for the period JJA 2016) obtained with the 2 forecast runs (o-suite and control), for each one of the selected Mediterranean stations. MNMBs and correlations with blue denote stations where control run performs better while with red are denoted stations where o-suite performs better.

Station Name	Stat_ID	Lon	Lat	Alt (m)	Level	Distance from the shore (km)	MNMB		Cor. Coef	
							o-suite	contro	o-suite	control
Al Cornocales	ES1648A	-5.66	36.23	189	59	16	19.6	10.6	0.68	0.70
Caravaka	ES1882A	-1.87	38.12	1	60	73	-12.2	-19.9	0.43	0.52
Zarra	ES0012R	-1.10	39.08	885	55	70	-3.4	-12.2	0.76	0.77
Villar Del Arzobispo	ES1671A	-0.83	39.71	430	60	48	-4.5	-11.9	0.52	0.69
Cirat	ES1689A	-0.47	40.05	466	60	37	13.1	6.1	0.52	0.66
Bujaraloz	ES1400A	-0.15	41.51	327	60	60	-4.8	-14.2	0.46	0.45
Morella	ES1441A	-0.09	40.64	1150	52	51	4.6	-4.0	0.64	0.70
Bc-La Senia	ES1754A	0.29	40.64	428	60	21	-5.0	-12.4	0.21	0.24
Ay-Gandesa	ES1379A	0.44	41.06	368	60	15	2.2	-5.7	0.49	0.63
Ak-Pardines	ES1310A	2.21	42.31	1226	51	81	25.2	16.0	0.42	0.26
Hospital Joan March	ES1827A	2.69	39.68	172	56	3	25.7	16.9	0.51	0.50
Al-Agullana	ES1201A	2.84	42.39	214	60	25	2.5	-4.6	0.39	0.39
Av-Begur	ES1311A	3.21	41.96	200	58	9	9.8	2.0	0.61	0.57
Plan Aups/Ste Baume	FR03027	5.73	43.34	675	54	21	11.5	3.0	0.67	0.65
Gharb	MT00007	14.20	36.07	114	58	31	10.2	3.2	0.55	0.58
Aliartos	GR0001R	23.11	38.37	110	60	18	-1.3	-10.5	-0.22	-0.19
NEO	-	21.67	37.00	50	60	2	5.2	-8.1	0.14	0.38
Finokalia	GR0002R	25.67	35.32	250	55	4	6.8	-3.1	0.45	0.49
Ineia	-	32.37	34.96	672	52	5	7.4	0.1	0.26	0.37
Oros Troodos	-	32.86	34.95	1819	47	11	2.2	-4.0	0.03	0.04
Agia Marina	CY0002R	33.06	35.04	532	53	14	10.2	5.3	0.27	0.38

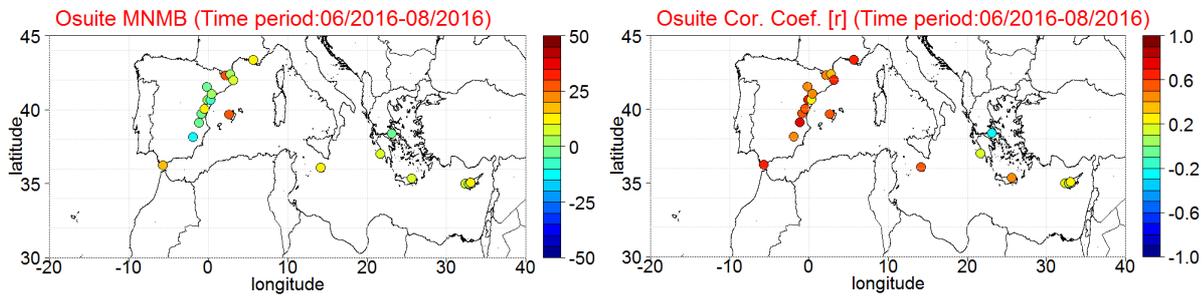


Figure 3.1.21: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 June 2016 to 31 August 2016.

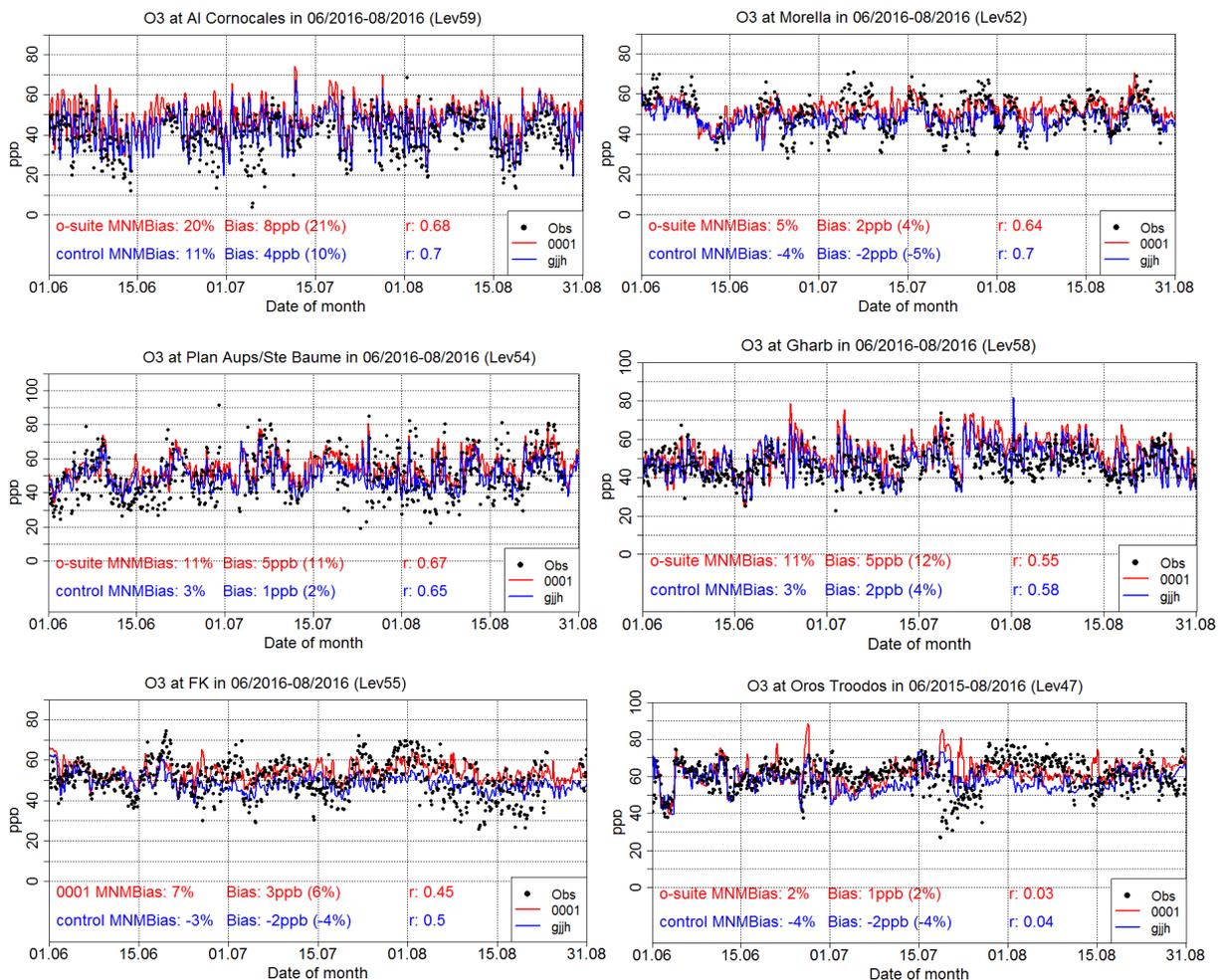


Figure 3.1.22: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Cornocales, Spain station (36.23°N, 5.66 °W, top left), at Morella, Spain station (40.64°N, 0.09°W, top right), at Plan Aups/Ste Baume, France station (43.34°N, 5.73°E, center left), at Gharb, Malta station (36.07°N, 14.20°E, center right) and at Finokalia, Crete station (35.32°N, 25.67°E, low left), and compared to observations provided by the Department of Labour Inspection - Ministry of Labour and Social Insurance of Cyprus, at Troodos Mountain station (34.95°N, 32.86 °E, low right).



Table 3.1.2. Normalised Mean Bias (NMB) and correlation coefficient (r) of the Control and the o-suite simulations for the two sites Alert, Nunavut and Villum Research Station, Greenland (VRS) for the period June – August 2016.

		NMB	R
Alert	o-suite	-0.10	0.31
	control	-0.24	0.45
VRS	o-suite	-0.15	0.18
	control	-0.29	0.26
Tiksi	o-suite	-0.16	0.50
	control	-0.36	0.47

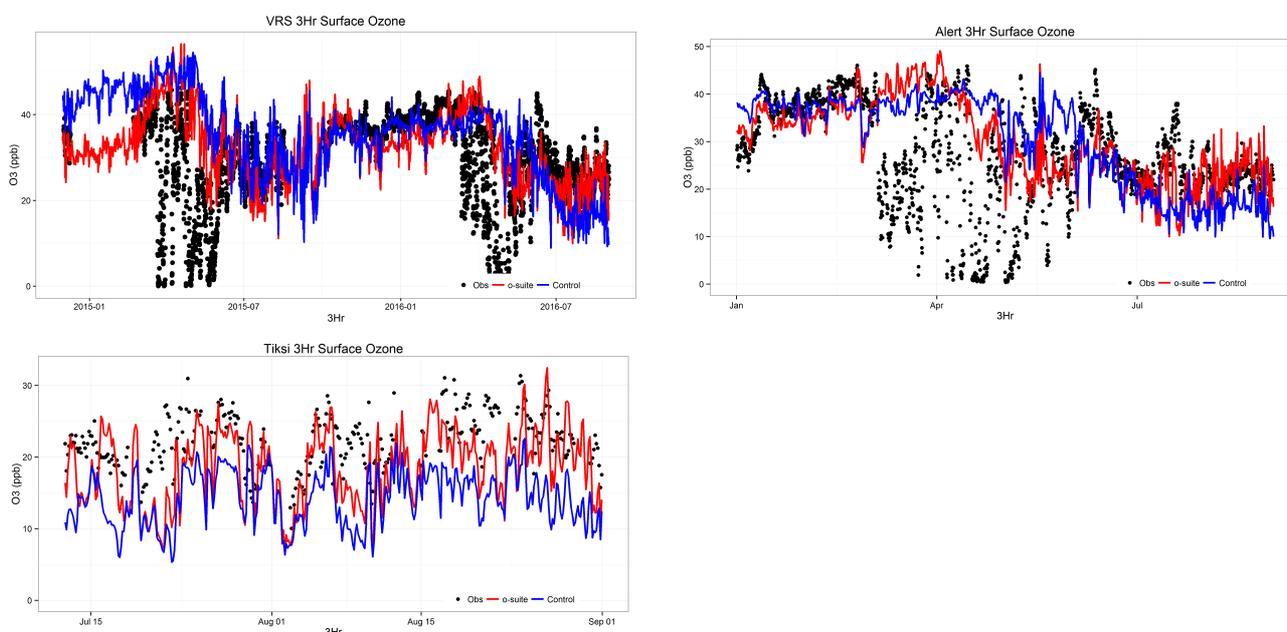


Figure 3.1.23: Time series for o-suite (red) and Control (blue) compared to observations (black dots) at the Villum Research Station, Station Nord, Greenland (top left), Alert, Nunavut (top right) and Tiksi, Russia (bottom left).

3.1.5 Validation with IASOA surface observations

Model results were compared to O₃ observations from the Villum Research Station, Station Nord in north Greenland (81.6°N 16.7°W), from Alert, Nunavut (82.5°N 62.5°W) and Tiksi, Russia (71.6°N 128.9°E) from the IASOA network, Fig. 3.1.23.

There are large gaps in the measurement time series for VRS covering the period from December 2014 to February 2016. Data from Alert covers the period January – August 2016 and data from Tiksi covers the period July – August 2016. Ozone depletion events in March – June in 2015 and 2016 are not captured by the model simulations during spring for the sites. These events are related to halogen chemistry reactions that are not represented in the model simulations. The simulations are on average in good agreement with the observations apart from the spring depletion events.



For the period June 2016 – August 2016 the measurements are not quality controlled. The model simulations underestimate the observed concentrations at all three sites. The levels predicted by the o-suite run is in better agreement with the observations with a normalized mean bias of -10 – -16% for the period compared to the normalized mean bias of -24 – -36% for the control run. The short-term variability is captured slightly better by the control run with $r = 0.26 – 0.47$, while it is $r = 0.18 – 0.50$ for the o-suite (Table 3.1.2).

3.2 Tropospheric nitrogen dioxide

3.2.1 Evaluation against GOME-2 retrievals

In this section, model columns of tropospheric NO₂ are compared to SCIAMACHY/Envisat NO₂ satellite retrievals (IUP-UB v0.7) [Richter et al., 2005] for model data before April 2012, and to GOME-2/MetOp-A NO₂ satellite retrievals (IUP-UB v1.0) [Richter et al., 2011] for more recent simulations. This satellite data provides excellent coverage in space and time and very good statistics. However, only integrated tropospheric columns are available and the satellite data is always taken at the same local time, roughly 10:00 LT for SCIAMACHY and 09:30 LT for GOME-2, and at clear sky only. Therefore, model data are vertically integrated, interpolated in time and then sampled to match the satellite data. Specifically, GOME-2 data were gridded to model resolution (i.e. 0.75° deg x 0.75° deg). Model data were treated with the same reference sector subtraction approach as the satellite data. Uncertainties in NO₂ satellite retrievals are large and depend on the region and season. Winter values in mid and high latitudes are usually associated with larger error margins. As a rough estimate, systematic uncertainties in regions with significant pollution are on the order of 20% – 30%.

Figure 3.2.1 shows global maps of GOME-2 and model monthly mean tropospheric NO₂ columns as well as differences between retrievals and simulations for July 2016. The overall spatial distribution and magnitude of tropospheric NO₂ is well reproduced by both model runs, indicating that emission patterns and NO_x photochemistry are reasonably represented. Some differences are apparent between observations and simulations, with generally larger shipping signals simulated by the models. For example, shipping signals are largely overestimated to the south of India. Compared to satellite data, all model runs underestimate tropospheric background values over Africa, Europe and the US. Local maxima of values observed over anthropogenic emission hotspots in East Asia (e.g. over the heavily populated Sichuan Basin; 30°N, 105°E), India and Moscow are overestimated. Both runs overestimate fire emissions over Central Africa, while fire emissions over South Eastern Africa seem to be partly overestimated but also partly underestimated.

Moreover, both runs show local maxima over Siberia (around 100° E, 65° N) and Canada (around 120° W, 65° N), which do not or only very weakly show up in the satellite retrievals. A reasonable explanation is an overestimation of NO_x fire emissions from boreal forests in these areas (the presence of fires is confirmed by global fire maps, e.g. <http://rapidfire.sci.gsfc.nasa.gov/firemaps/>). Note that this issue has occurred over biomass burning regions before and as such was already reported in previous MACC and CAMS near real time reports.

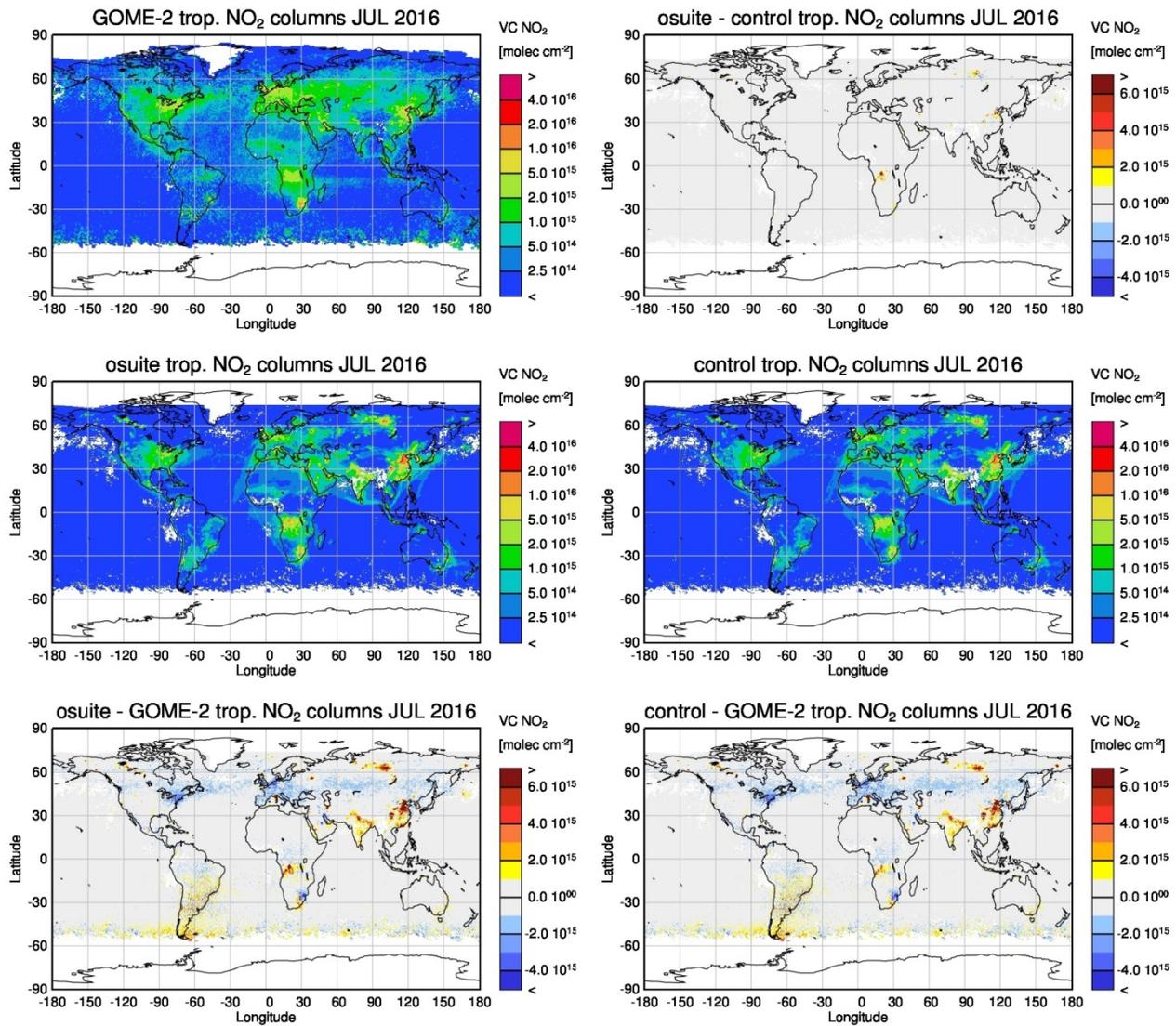


Figure 3.2.1: Global map comparisons of satellite retrieved and model simulated tropospheric NO₂ columns [molec cm⁻²] for July 2016. The top row shows monthly mean tropospheric NO₂ columns retrieved by GOME-2 as well as the difference between osuite and control, the second row shows the corresponding tropospheric NO₂ columns for model simulated averages. The third row shows differences of monthly means between models and GOME-2.

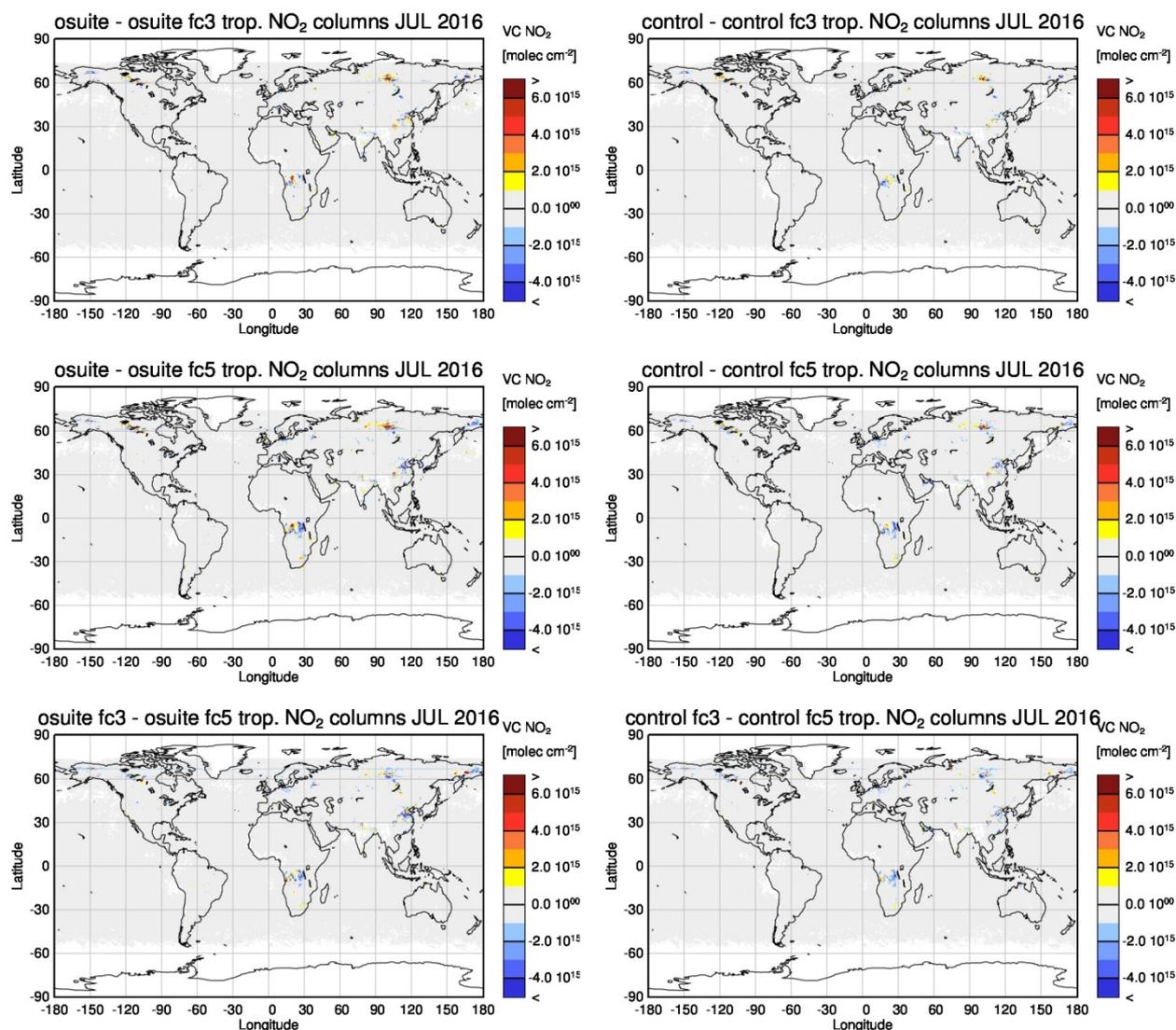


Figure 3.2.2: As in Figure 3.2.1, but for (top) differences between model runs for forecast day 1 and 3, (middle) forecast day 1 and 5, (bottom) forecast day 3 and 5.

Figure 3.2.2 shows differences between some selected forecast days of the models for July 2016. Overall, differences between individual forecast steps are small apart from specific regions, which can be explained by the fact that NO_2 has a short lifetime in the troposphere, so that data assimilation does not have a strong impact on model results. Differences over Moscow get larger from forecast day 3 to 5. This is also true for the compared to satellite retrievals overestimated boreal forest fire emissions over Siberia, which as such agree better to GOME-2 observations for forecast day 5 compared to previous forecast days. Note however that forecast day 5 simulations are closer to forecast day 1 simulations than forecast day 3 simulations for biomass burning regions over South Africa for August 2016 (not shown). Therefore, the development of simulations from one forecast day to another depends largely on region and season. However, as for previous reports (Eskes et al. 2016, quarterly validation report Dec 2015 – Feb 2016, https://atmosphere.copernicus.eu/quarterly_validation_reports; Richter and Blechschmidt 2015, MACC III deliverable D20.6), model performance does not change much from forecast day 1 to later

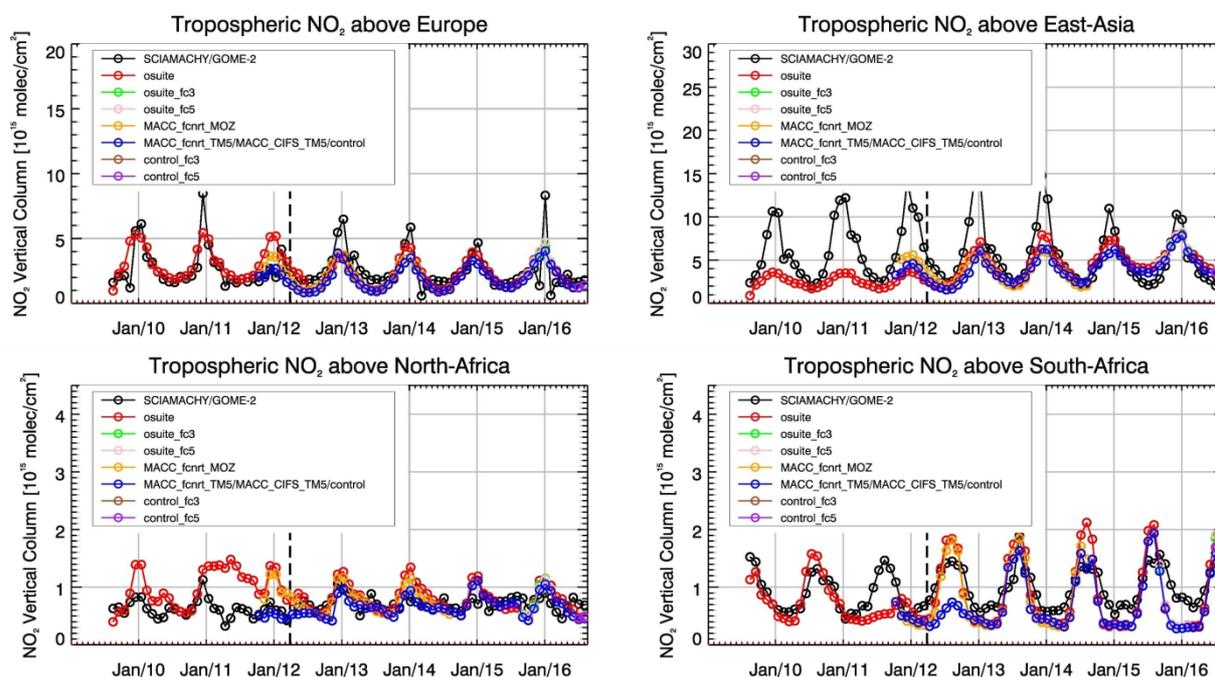


Figure 3.2.3: Time series of average tropospheric NO₂ columns [10^{15} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions (see Annex 2 for definition of regions). Upper panels represent regions dominated by anthropogenic emissions, lower panels represent those dominated by biomass burning. The blue line shows MACC_fcnrnt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 to August 2014 and control results from September 2014 onwards. Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

steps compared to GOME-2 satellite retrievals. The latter is also true for time series comparisons of tropospheric NO₂ discussed in the next paragraphs as well as for tropospheric HCHO (section 3.4) and will as such not be discussed any further within in the present validation report.

Closer inspection of the seasonal variation of tropospheric NO₂ in some selected regions (Fig. 3.2.3) reveals significant differences between the models and points to some simulation problems. Over regions where anthropogenic emissions are major contributors to NO_x emissions, models catch the shape of the satellite time series rather well. However, over East-Asia absolute values and seasonality are in general strongly underestimated by all model runs (most likely due to an underestimation of anthropogenic emissions), with the o-suite showing the best results since the upgrade in July 2012. As NO₂ column retrievals decreased since 2014, model simulated values are in better agreement with the satellite retrieved ones for recent years. However, this decrease in values is not reproduced by the simulations and as such, the better agreement for more recent years cannot be attributed to an improvement of the simulations. Springtime and summertime model values increased in 2015 compared to previous years, which is in contrast to the satellite retrievals, so that the simulated values for the summers 2015 and 2016 are about 50% larger than satellite retrieved ones. As for East-Asia, a decrease in satellite retrieved values also occurs for Europe where a peak is usually found around January, which is, as a result, only slightly underestimated by the models for January 2015. The underestimation of tropospheric NO₂ columns over Europe may be caused to some extent by a change of emission inventories in 2012. However,



the situation changed for winter 2015/2016, for which GOME-2 shows (compared to previous years) a strong increase in January peak values, combined with a decrease in values for December 2015 and February 2016, which is not reproduced by the models. It is not clear if the GOME-2 observations are realistic here, although a first inspection of daily GOME-2 satellite images did not point to any problems regarding the retrieval.

Over regions where biomass burning is the major contributor to NO_x emissions, seasonality and amplitude of model columns are determined by fire emissions. The seasonality for the two regions in Africa is simulated reasonably well for 2010 and after October 2011. In the time period in between, a bug in reading fire emissions lead to simulation errors for all MOZART runs. Over North-Africa, the o-suite shows improved results since the update in July 2012 and the change to CIFS-CB05 in September 2014. However, tropospheric NO₂ columns around December are still overestimated by the

models. Summertime NO₂ columns over North-Africa are underestimated compared to the satellite data for 2015 and 2016. The models strongly overestimates the seasonal cycle for South-Africa since 2014 with an overestimation of the seasonal maximum which usually occurs around August of each year (e.g. by a factor of 1.4 larger compared to GOME-2 retrievals in August 2016). For 2014 model runs without data assimilation agree much better with satellite observations, in contrast to more recent CB05-based o-suite runs since 2015. For November 2015, satellite retrieved values over South-Africa do not decrease below 1×10^{15} molec/cm², a feature which did not show up in the time series before. While wintertime values over South-Africa were also underestimated by the models for previous years, the underestimation is now even stronger given the comparatively large satellite retrieved NO₂ columns since November 2015.

Details on the NO₂ evaluation can be found at: http://www.doas-bremen.de/macc/macc_veri_iup_home.html.

3.2.2 Evaluation against ground-based DOAS observations

In this section, we compare the NO₂ profiles of the CAMS models with UVVIS DOAS measurements at Xianghe (39.8°N, 117°E, station near Beijing, altitude 92m) and Haute Provence (43.9°N, 5.71°E, rural station, altitude 650m).¹ This ground-based, remote-sensing instrument is sensitive to the NO₂ abundance in the lower troposphere, up to 1km altitude with an estimated uncertainty of 8%. Tropospheric NO₂ profiles and columns are validated (up to 3.5km). A description of the instruments and applied methodologies is the same all DOAS OFFAXIS measurements, see <http://nors.aeronomie.be>. It is important to mention here that the model partial column values between the surface and 3.5 km are calculated for the smoothed model profiles (see Fig. 3.2.4). This guarantees that the model levels where the measurement is not sensitive do not contribute to the observed bias. We should mention that the measurement data is still catalogued as rapid delivery and not in the consolidated NDACC database.

¹ No contribution from UCCLE due to instrument failure.

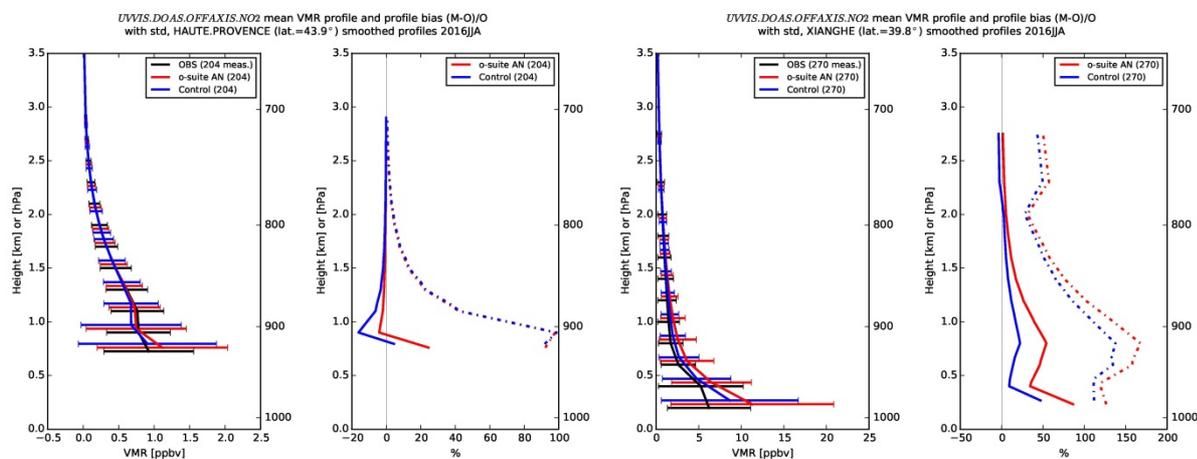


Figure 3.2.4: Seasonal mean tropospheric NO₂ profiles by o-suite (red) and Control (blue) compared to NDACC UVVIS DOAS data at Haute Provence (43.9°N, 5.71°E, left top) and Xianghe (39.8°N, 117°E, right) for JJA 2016.

From Figs. 3.2.4 and the below table, we see the assimilation has a negative effect at both sites during JJA 2016.

3.3 Carbon monoxide

3.3.1 Validation with Global Atmosphere Watch (GAW) Surface Observations

For the Near-Real-Time (NRT) validation, 11 GAW stations have delivered CO surface mixing ratios in NRT and data is compared to model results as described in Eskes et al (2016) and is used for CAMS model evaluation for June – August 2016. The latest validation results can be found on the CAMS website: <http://www.copernicus-atmosphere.eu/d/services/gac/verif/grg/gaw/>

For stations in Europe and the Southern hemisphere, the MNMBs and correlation coefficients indicate that the forecast remains stable till the D+4 (forecast run from 96h to 120h). For stations in Japan, results show lower MNMBs but also lower correlation for the forecasts D+2 and D+4.

Table 3.2.1: Seasonal relative mean bias (MB, %), standard deviation (STD, %) for the considered period and number of observations used (NOBS), compared to NDACC UVVIS OFFAXIS observations at Haute Provence and Xianghe (mean bias and stddev in %). The overall mean uncertainty for the NO₂ measurements is 5%. Colored numbers indicate best performance (osuite or control).

		SON			DJF			MAM			JJA		
		MB	stddev	nobs	MB	stddev	nobs	MB	stddev	nobs	MB	stddev	nobs
osuite	ohp	-5.72	40.84	223	35.13	96.22	115	-6.08	48.27	353	13.28	55.57	204
control	ohp	-12.48	59.38	223	21.16	84.9	115	-20.36	33.14	353	2.46	51.24	204
osuite	xianghe	91.40	141.89	64	11.73	77.38	103	23.50	74.65	216	76.55	105.5	270
control	xianghe	69.98	104.11	64	24.63	105.5	103	5.21	63.77	216	42.74	74.95	270

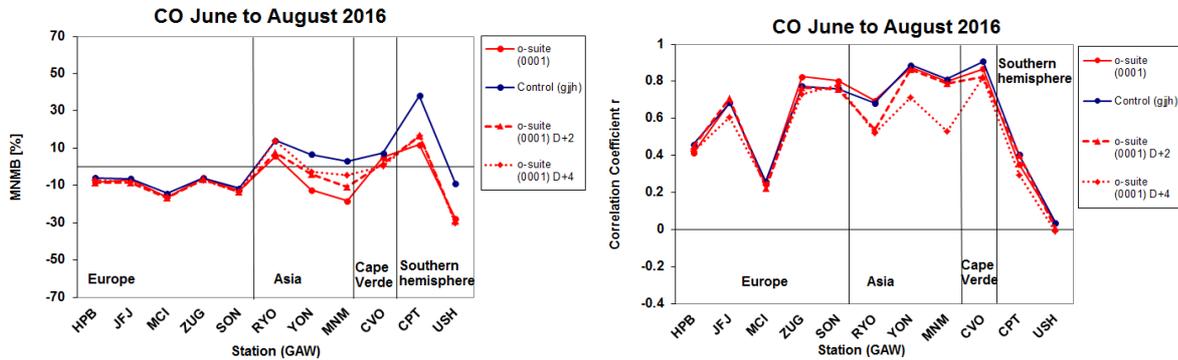


Figure 3.3.1: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT model runs compared to observational GAW data in the period June to August 2016. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.

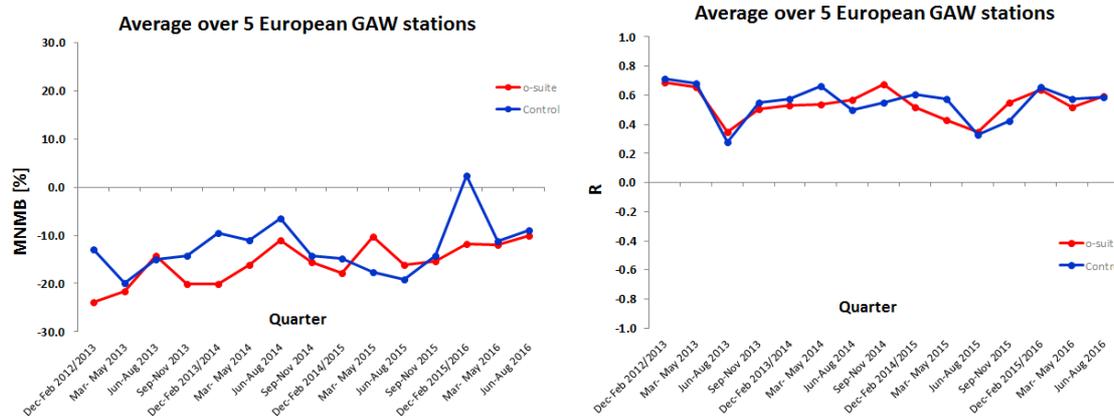


Figure 3.3.2: Long term (Dec. 2012 – August 2016) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

A comparison of the seasonal-mean MNMB over Europe (Fig. 3.3.2) from December 2012 to present shows a slowly improving MNMB from about -20% in 2013 to -10% for more recent periods. Temporal correlation remains relatively constant at $r=0.5$ on average.

For European stations, both analyses runs show an underestimation of observed CO mixing ratios, with MNMBs between -6% and -16%. Correlation coefficients are between 0.24 and 0.82 for the o-suite and between 0.25 and 0.77 for the control run.

For Asian stations, the control run runs correspond well to the observations, with MNMBs between 3 and 14% whereas the o-suite shows partly larger negative MNMBs for two stations (up to -18%), see Fig. 3.3.4. Correlation coefficients are high between 0.7 and 0.89 for both runs.

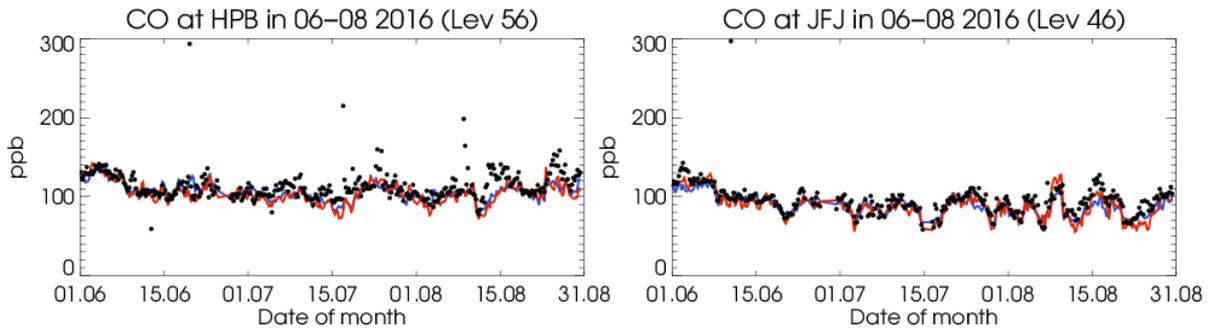


Figure 3.3.3: Time series for the o-suite (red) and control (blue) compared to GAW observations at Hohenpeissenberg (47.8°N, 11.0°E) and Jungfraujoch (46.5°N, 7.9°E).

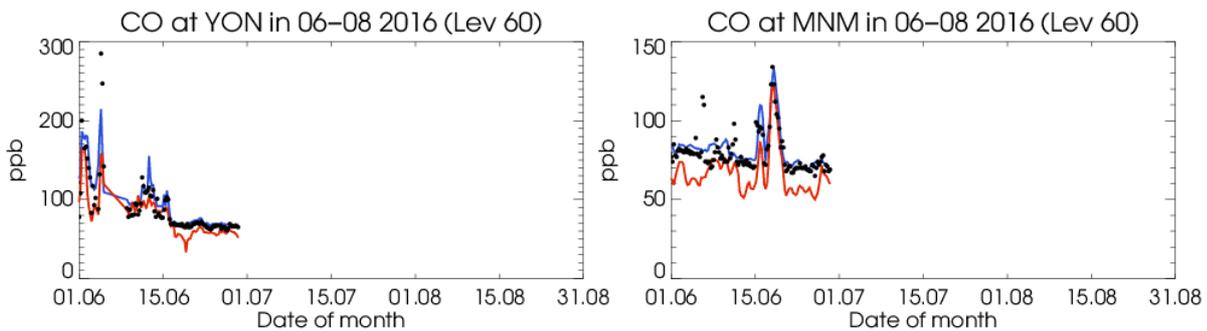


Figure 3.3.4: Time series for the o-suite (red) and control (blue) compared to GAW observations at Yonagunijima (24.47°N, 123.0°E) and Minamitorishima (24.3°N, 123.9°E).

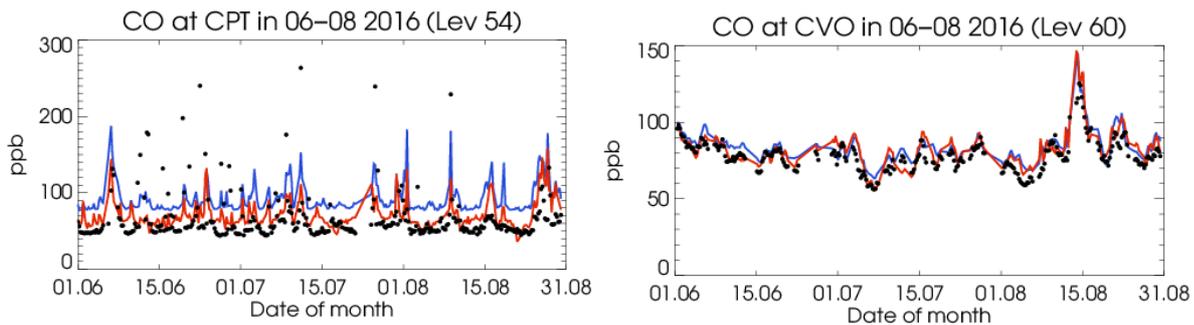


Figure 3.3.5: Time series for the o-suite (red) and control (blue) compared to GAW observations at Cape Point (34.35°S, 18.5°E) and Cape Verde (16.9°N, 24.9°W).

For the two stations in the Southern Hemisphere (CPT, USH), the positive offset visible for the control run is corrected by the data assimilation for the o-suite, see Fig. 3.3.5 (left panel).

Both runs show good results in reproducing the CO mixing ratios for Cape Verde station, see Fig.3.3.5 (right panel).



3.3.2 Validation with IAGOS Data

The daily profiles of ozone and CO measured at airports around the world are shown on the website at http://www.iagos.fr/macc/nrt_day_profiles.php . For the period June-July-August 2016, data from several aircraft have been validated, as discussed in Sec. 3.1.2.

Figure 3.3.6 shows the time series of CO over Frankfurt and Taipei for the 5 different layers throughout the troposphere. Over Taipei the models reproduce CO quite well in all atmospheric layers. However over Frankfurt, we can see an underestimation of CO in the surface and boundary layers as seen in previous seasons and as described by Stein et al. (2014).

Europe

Figure 3.3.7 gives examples of the CO profiles over Frankfurt, Amsterdam and Paris. CO in the boundary and surface layers is systematically underestimated. In general we find that CO in the free-troposphere is better estimated by the models. In JJA, there is little data considered to be in the lower stratosphere and the performance of the model in the upper troposphere is better as there is no gradient to capture.

Asia

The time series at Taipei (Fig. 3.3.6) showed that the CO from the model versions showed good correspondence to the observations in the free troposphere and upper troposphere. In general this is the case in many locations across Asia and South-east Asia as the profiles from a range of airports show (Fig 3.3.8). The underestimation of CO in the boundary and surface layers which is so prominent in Europe is less pronounced on profiles over Asia.

The varied effects of typhoons are also visible in the profiles of CO over Taipei. In Fig. 3.3.8 the profile of CO on 8th July shows low CO mixing ratios consistent with the arrival of clean maritime air brought by typhoon Nepartak which hit Taiwan directly after following an oceanic trajectory. In contrast, from 1st -4th August over Taipei an anomaly in CO was observed at altitudes around 8000m. The corresponding anomaly in ozone was really well captured, but in CO the magnitude of the anomaly is underestimated. This anomaly is due to the outflow of Typhoon Nida dumping polluted boundary layer air, which was picked up over the Philippines, into the upper troposphere.

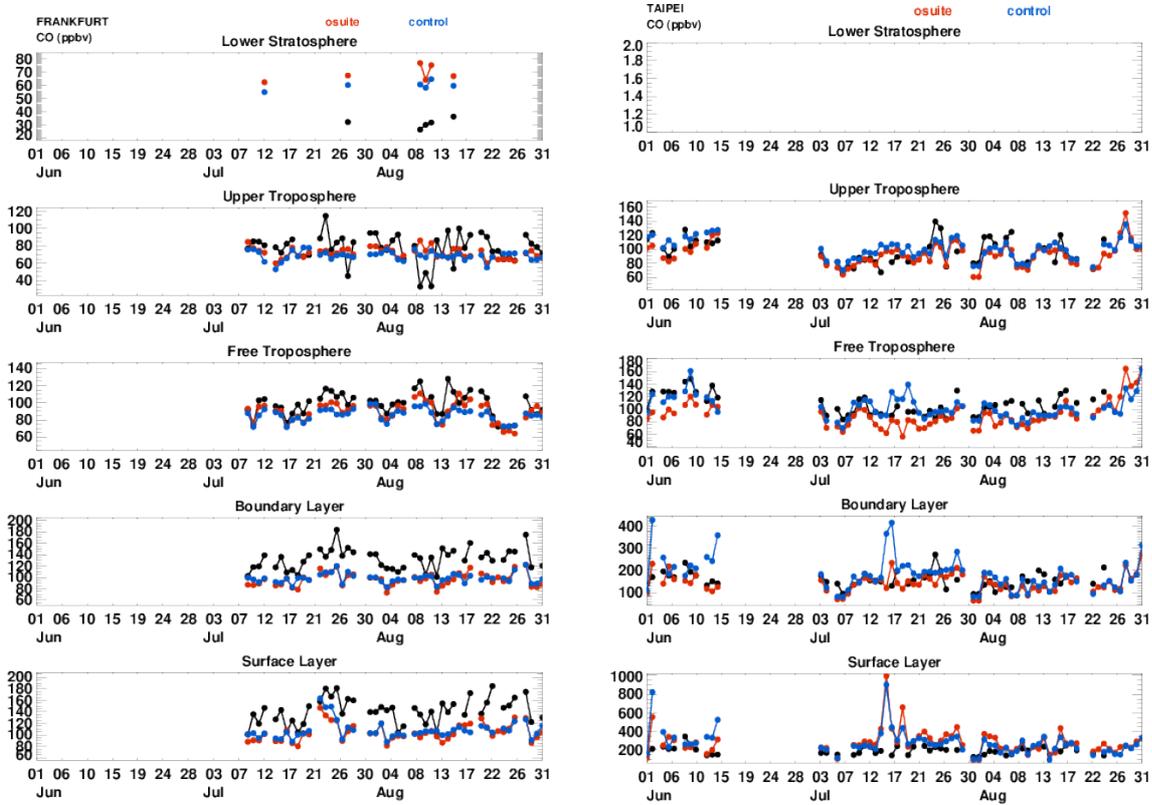


Figure 3.3.6: Time series of daily mean CO at Frankfurt and Taipei during JJA 2016 for 5 layers, Surface, Boundary layer, Free Troposphere, Upper Troposphere and Lower Stratosphere.

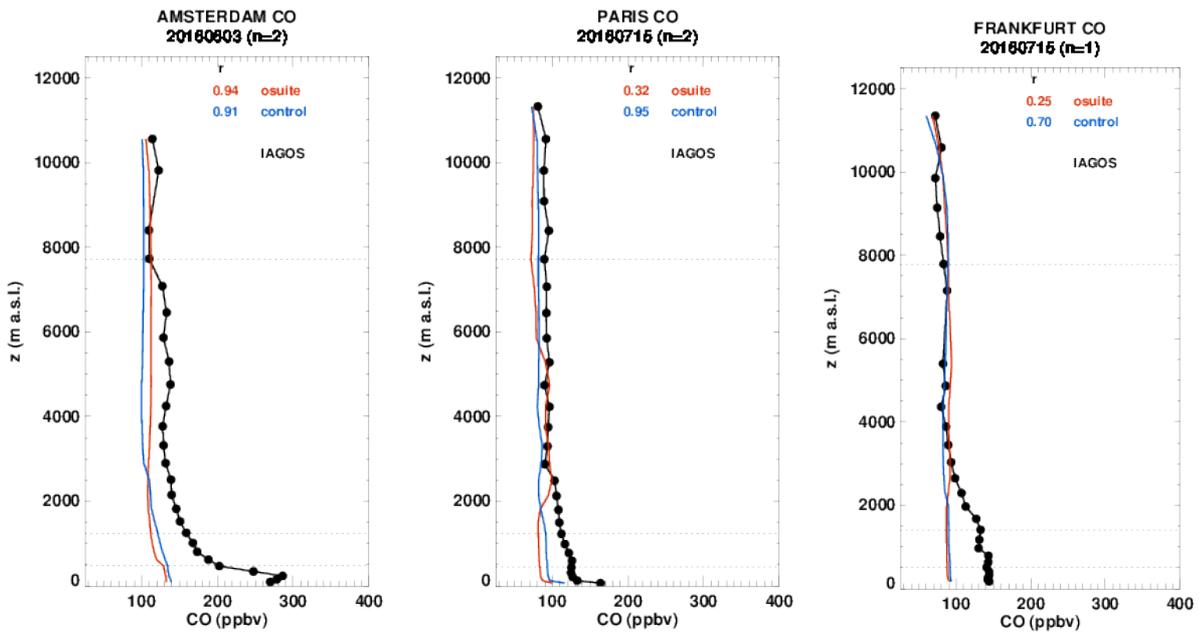


Figure 3.3.7: Selection of profiles of CO from IAGOS (black) and the two NRT runs over Europe in JJA 2016.

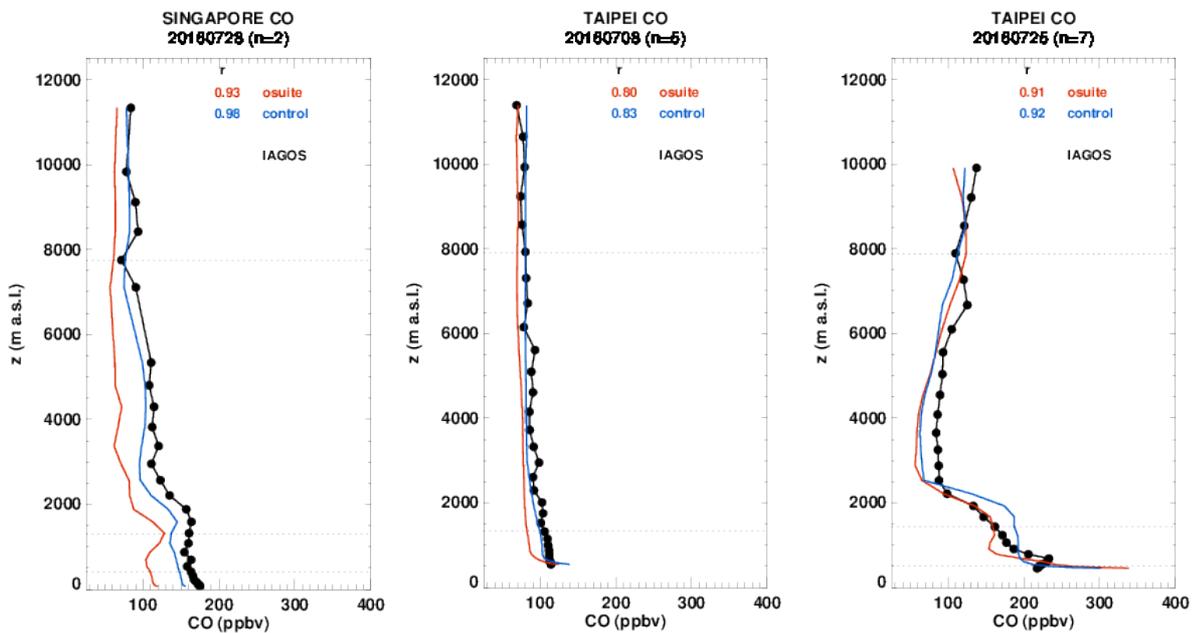


Figure 3.3.8: Profiles of CO from IAGOS (black) and the two NRT runs in China and South East Asia during the period June to August 2016.

Equatorial West-Africa

Figure 3.3.9 highlights some examples of CO profiles over Equatorial Africa as regularly sampled by Air France. For this period, there are profiles at Abijan Lagos and Bamako, Cotonou and Ouagadougou.

These cities are influenced by anthropogenic emissions from vehicles, and from biomass burning which stretches across Africa just north of the equator from December to March. Peaks in CO are evident (along with peaks in ozone) at Lome and Lagos, which are likely the result of cross equatorial transport of plumes from intense burning taking place around Kinshasa and Pointe Noire. The local biomass plumes are well captured at Pointe Noire and Kinshasa, but the transported plumes are missed at Lagos and Lome, suggesting that the cross equatorial transport of plumes is not so well captured by the models.

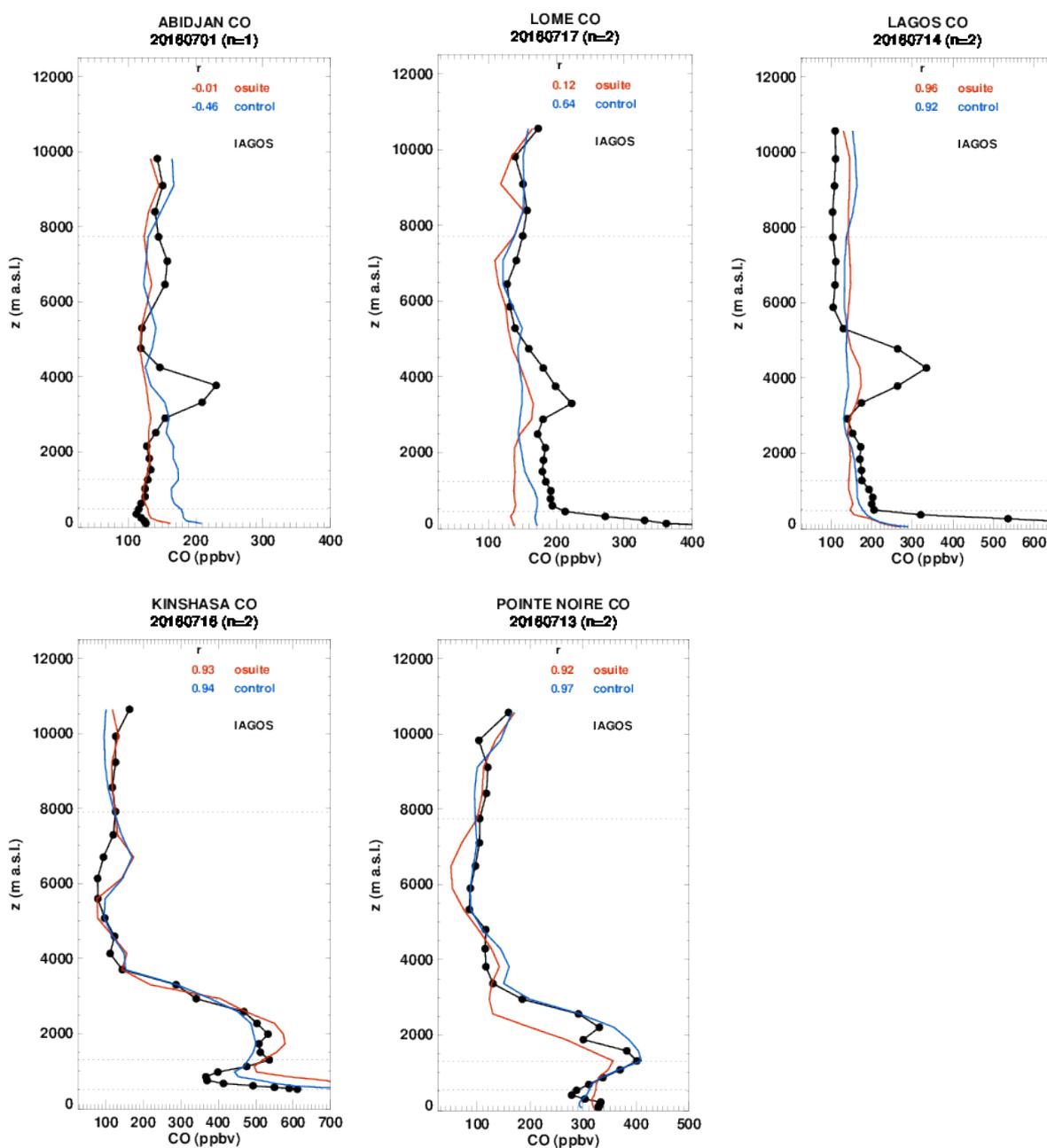


Figure 3.3.9: Profiles of CO from IAGOS (black) and the two NRT runs in Equatorial West-Africa for June-August 2016

3.3.3 Validation against FTIR observations from the NDACC network

In this section, we compare the CO profiles of the CAMS models with FTIR measurements at Maido (21°S, 55°E, i.e. southern tropics, altitude 2.2km) and Lauder (46°S, 169.7°E, altitude 370m). These ground-based, remote-sensing instruments are sensitive to the CO abundance in the troposphere and lower stratosphere, i.e. between the surface and up to 20 km altitude. Tropospheric CO profiles and columns are validated (up to 10km). A description of the instruments and applied methodologies can be found at <http://nors.aeronomie.be>.



Table 3.3.1: Seasonal relative mean bias (MB, %), standard deviation (STD, %) for the considered period and number of observations used (NOBS), compared to NDACC FTIR observations at Lauder and Maido (mean bias and stddev in %). The overall uncertainty for the CO measurements at Lauder and Maido is approximately 5%.

		SON			DJF			MAM			JJA		
		MB	stddev	nobs									
o-suite	Lauder	-2.36	5.01	148	-8.53	5.27	93	-1.71	5.66	150	6.65	23.35	150
control	Lauder	14.13	7.92	148	36.82	12.21	93	45.87	5.17	150	33.25	12.23	148
o-suite	Maido	-6.32	4.25	593	-8.60	3.34	290	-5.30	3.34	527	-6.59	3.55	822
control	Maido	10.16	10.19	593	34.00	6.14	290	30.70	5.29	527	21.77	8.28	822

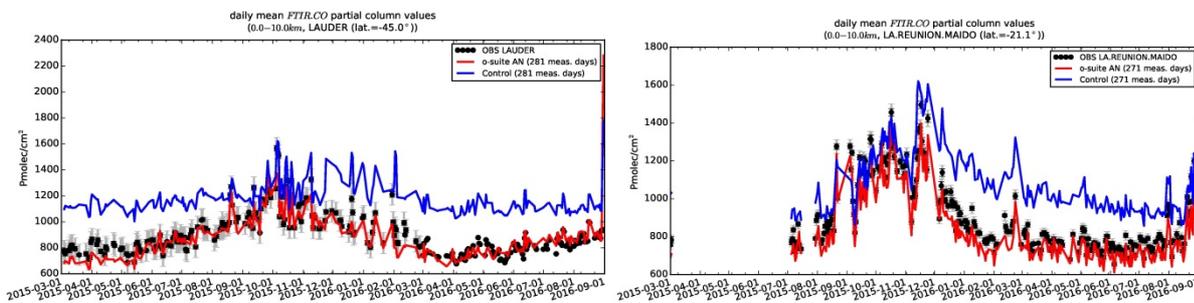


Figure 3.3.10: Daily mean values of tropospheric CO columns (till 10km) by the o-suite (red) and the Control run (blue) compared to NDACC FTIR data at Lauder, New Zealand (45°S, 169.7°E) (left) and Maido (21°S, 55°E) (right) for the period March 2015–September 2016. The number of measurement days is indicated in the legend. In Lauder very high model values are seen at the end of August 2016 inconsistent with the observations.

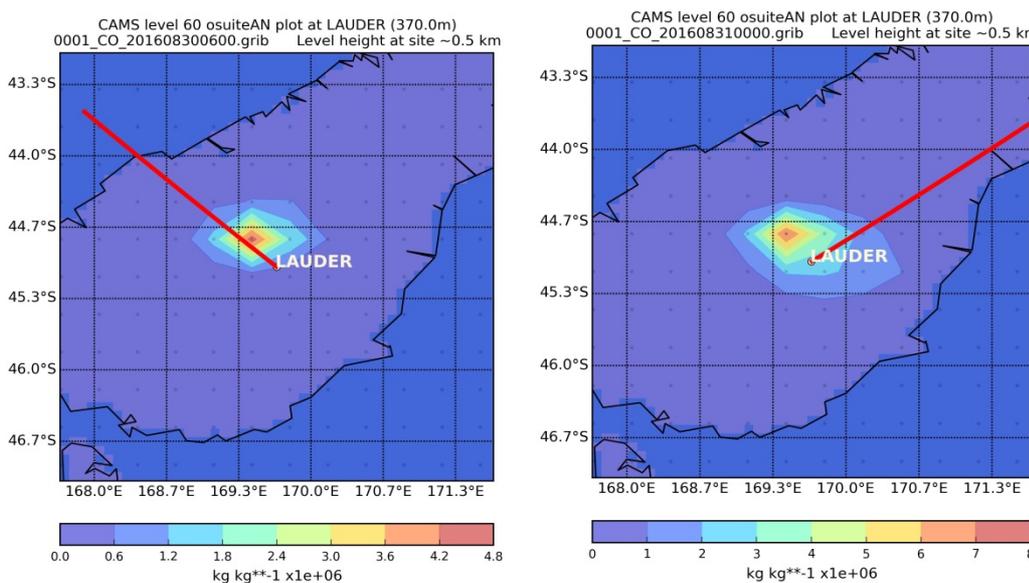


Figure 3.3.11: Example of CO source in the o-suite lowest level at Lauder for 30 (left) and 31 (right) August 2016. The red line is the line of sight of the FTIR measurement. The high CO abundances from the model are not observed by the FTIR measurements.



Table 3.3.1 and Fig. 3.3.10 show that the tropospheric columns of CO agree well. The o-suite underestimates CO at Lauder with values around 2%, which is within the measurements uncertainty range (6%). At Mado the o-suite underestimates the CO abundance (approx. -8%). The mean uncertainty on these measurements is 5%, so the observed o-suite biases are significant. For both stations, the control run shows an overestimation of CO with MBs between 20%-30%, clearly showing the positive effect of assimilation.

During the last days of August, the models see a strong increased concentration of CO near the surface at Lauder. Figure 3.3.11 shows that the high CO values in the model are not observed in the FTIR measurements.

3.3.4 Evaluation with MOPITT and IASI data

In this section, model CO total columns are compared to MOPITT versions 5 and 6 (thermal infrared radiances) (Emmons et al., 2009, Deeter et al., 2010) and IASI satellite retrievals (Clerbaux et al., 2009). Figure 3.3.12 shows the global distribution of CO total columns retrieved from MOPITT (top left) and IASI (top right) and relative bias of model runs with respect to MOPITT V5, averaged for August 2016. MOPITT and IASI show relatively high values over biomass burning areas in Africa, the central part of South America, and East of China. The difference between observations can be seen over the above mentioned regions, indicating higher values in IASI compared to MOPITT. The modeled CO geographical distribution and magnitude of values show that the model performs reasonably (not shown). The relative difference between the model runs and MOPITT shows that both model runs overestimate CO total column over the central part of South America and Indonesia up to 40%. In general, the o-suite performs better than the run without data assimilation, with some overestimation in the tropics and underestimations in the mid-latitudes of up to 20%. The control run shows an overestimation over the tropical and subtropical regions and Southern Hemisphere by about 20-30% with stronger overestimation over the central part of South America, Indonesia and biomass burning area in Africa. Figure 3.3.12 shows no significant difference between o-suite analysis and 2nd and 4th forecast days.

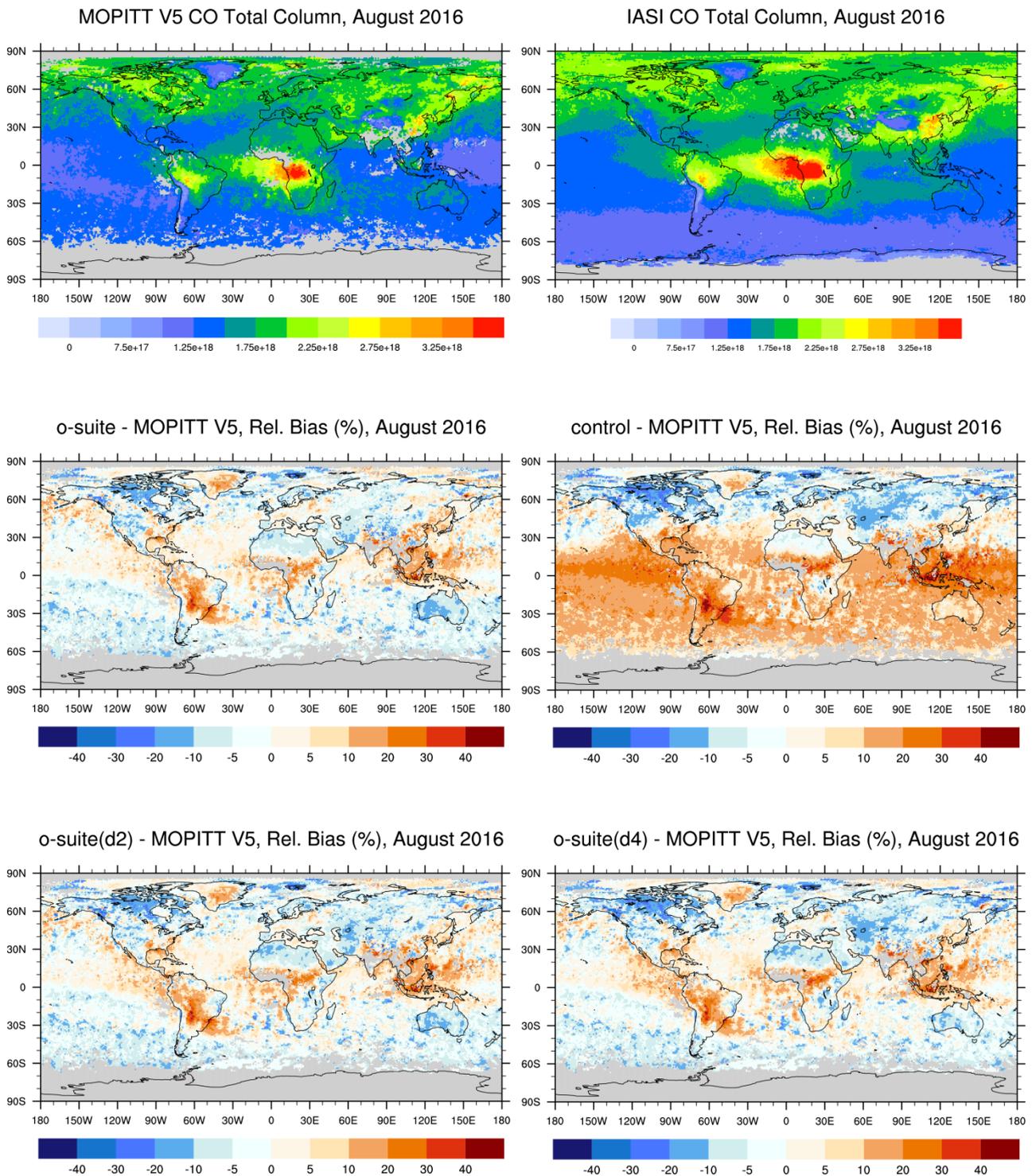


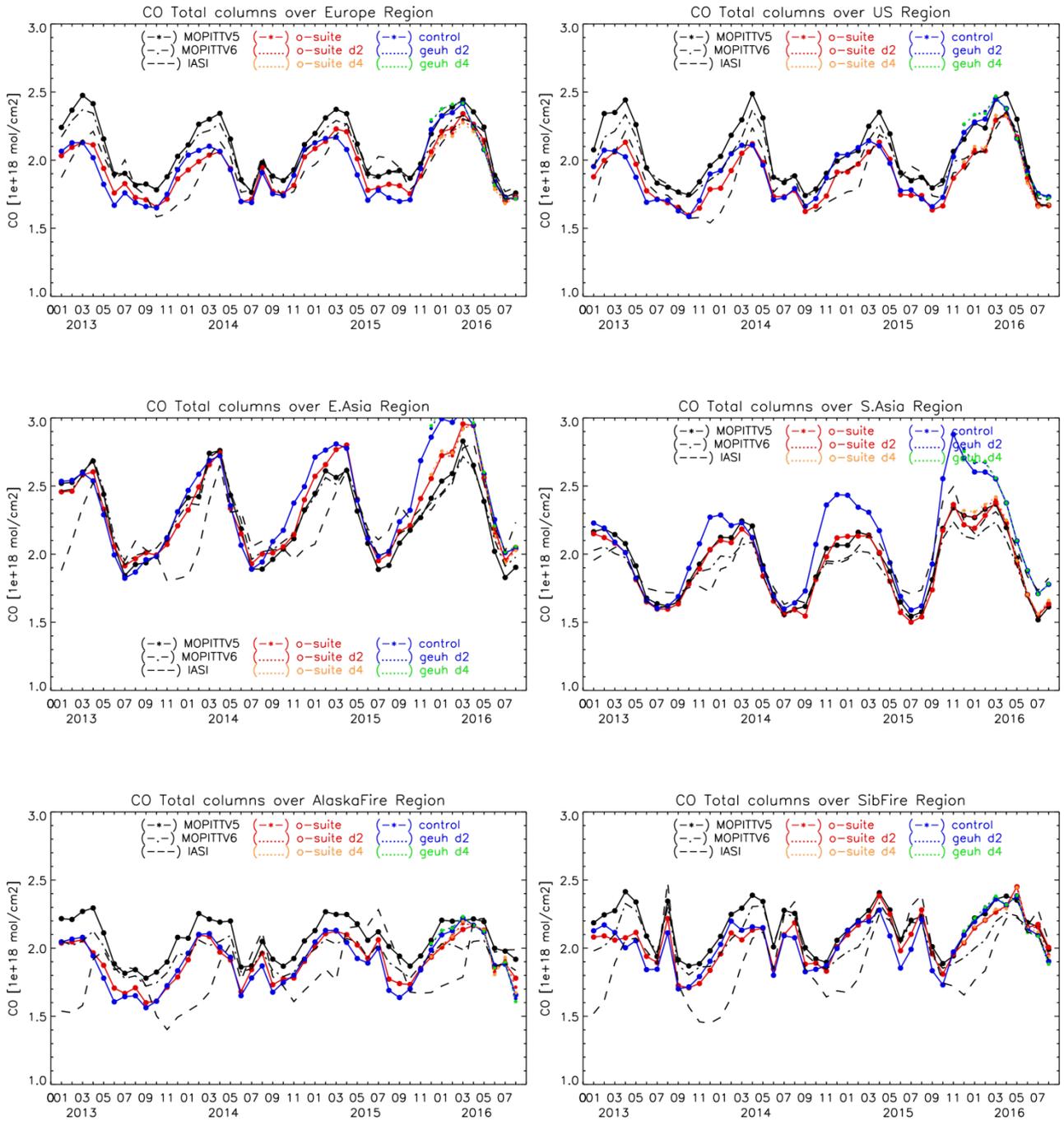
Fig. 3.3.12: CO total column for MOPITT V5 (top left) and IASI (top right) satellite retrievals and relative difference between the model runs and MOPITT for August 2016: o-suite (middle left), control run (middle right), o-suite 2nd forecast day (bottom left), o-suite 4th forecast day (bottom right). Grey color indicates missing values.



Figure 3.3.13 shows time series of CO total column for MOPITT V5 and V6, IASI, and the model runs over selected regions. For the comparison with MOPITT, the modelled CO concentrations were transformed using MOPITT V5 averaging kernels (Deeter, 2004). Both, MOPITT and IASI CO total column are assimilated in the o-suite run, while a bias correction scheme is applied to IASI data to bring it in line with MOPITT. MOPITT and IASI CO total columns show a relatively similar variability over different regions. In general, IASI CO values are lower compared to MOPITT over most regions with some seasonal exceptions. Significant difference between MOPITT and IASI are observed over the Alaskan and Siberian fire regions in winter seasons, with IASI CO total column values lower up to 30 %. Modelled seasonality of CO total columns is in relatively good agreement with the retrievals. In general, the comparison between o-suite and control run shows that assimilation of satellite CO has more positive, pronounced impact on model results over East and South Asia and North and South Africa and smaller impact over other regions.

Since September 2014 the o-suite shows better agreement with the satellite retrievals over Europe and US regions, especially in seasonal maximum in spring. Improvements can also be seen in Siberian fire region and North Africa. Since 2013, observations and models show growing CO over North African region as both seasonal minimums and seasonal maximums. In Siberian fire region, second CO minimum in summer 2016 is shifted to the later month compared to previous summers, where the minimum appeared in June.

The modified normalized mean bias (MNMB) of the model runs compared to MOPITT V5 (Fig. 3.3.14) allows quantifying the impact of assimilation on the model performance. All model runs show negative biases over Europe, the US region, and the Alaskan and Siberian fire regions with some seasonal exceptions. The control run shows a systematic positive bias of up to 20% over South Asia in November-December 2014, 2015. Over South Africa the control run overestimates satellite retrieved values of up to 25% in the second maximum in winter and in seasonal minimum in spring 2015, 2016. Compared to the last summer, the o-suite shows better agreement with the satellite observations over Europe, the US and the Siberian fire regions with biases just within 5%. The o-suite 2nd and 4th forecast days show growing negative biases up to 5% compared to the analysis in the Siberian and Alaskan fire regions, while positive biases appear in the North African region, which can be seen also in Fig.3.3.13. In Europe we can see a slightly higher negative bias compared to analysis. In other regions the forecasts are almost identical to analysis (within 1% difference).



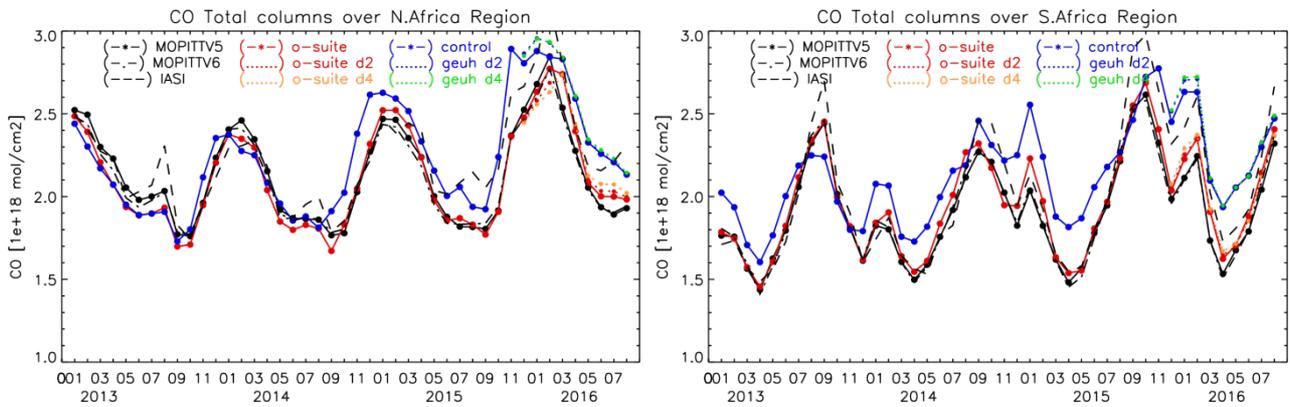
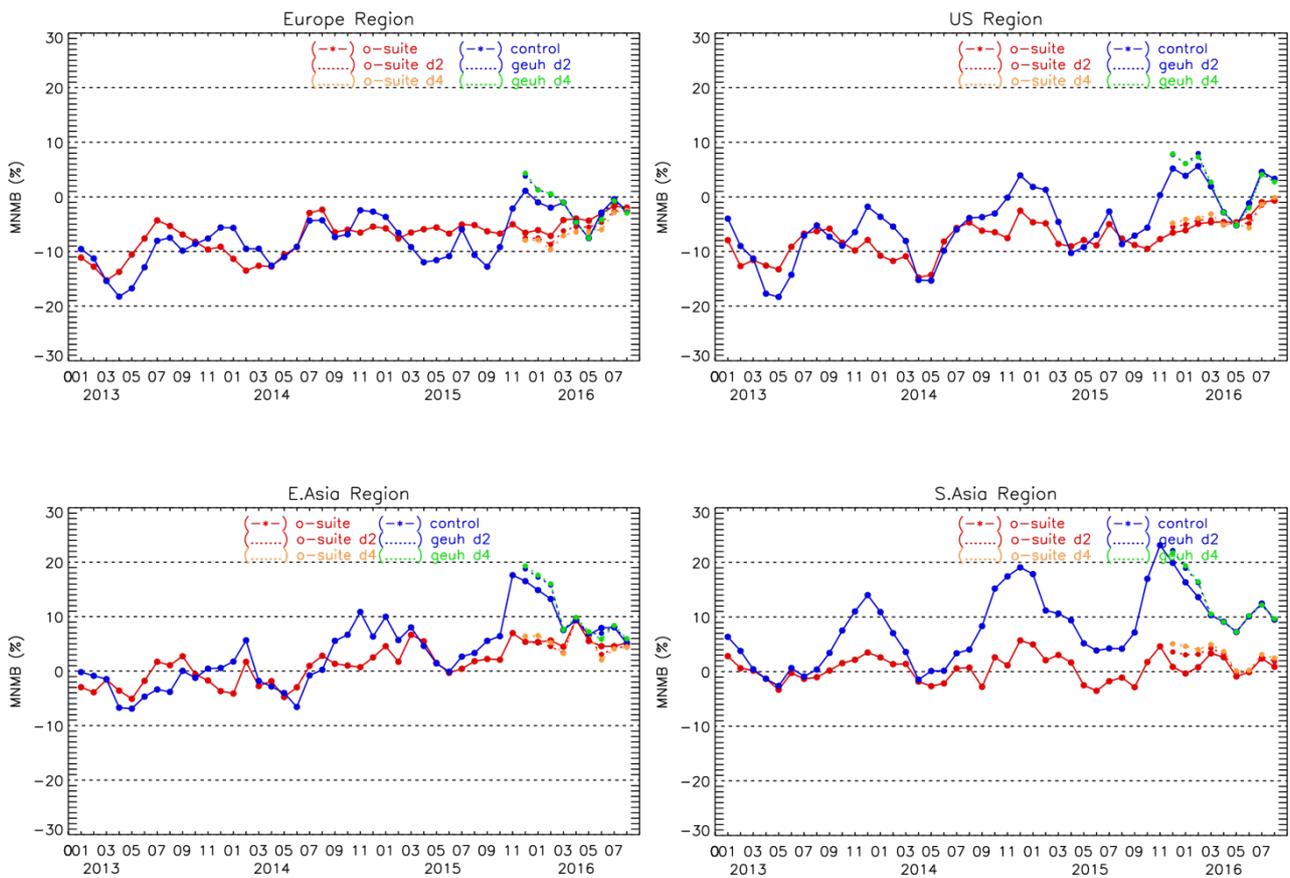


Fig. 3.3.13: Time series of CO total column for satellite retrievals MOPIT V5 and V6, IASI (black) and the model runs over the selected regions: o-suite (red, solid), control (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue, dotted), control 4th forecast day (green, dotted).



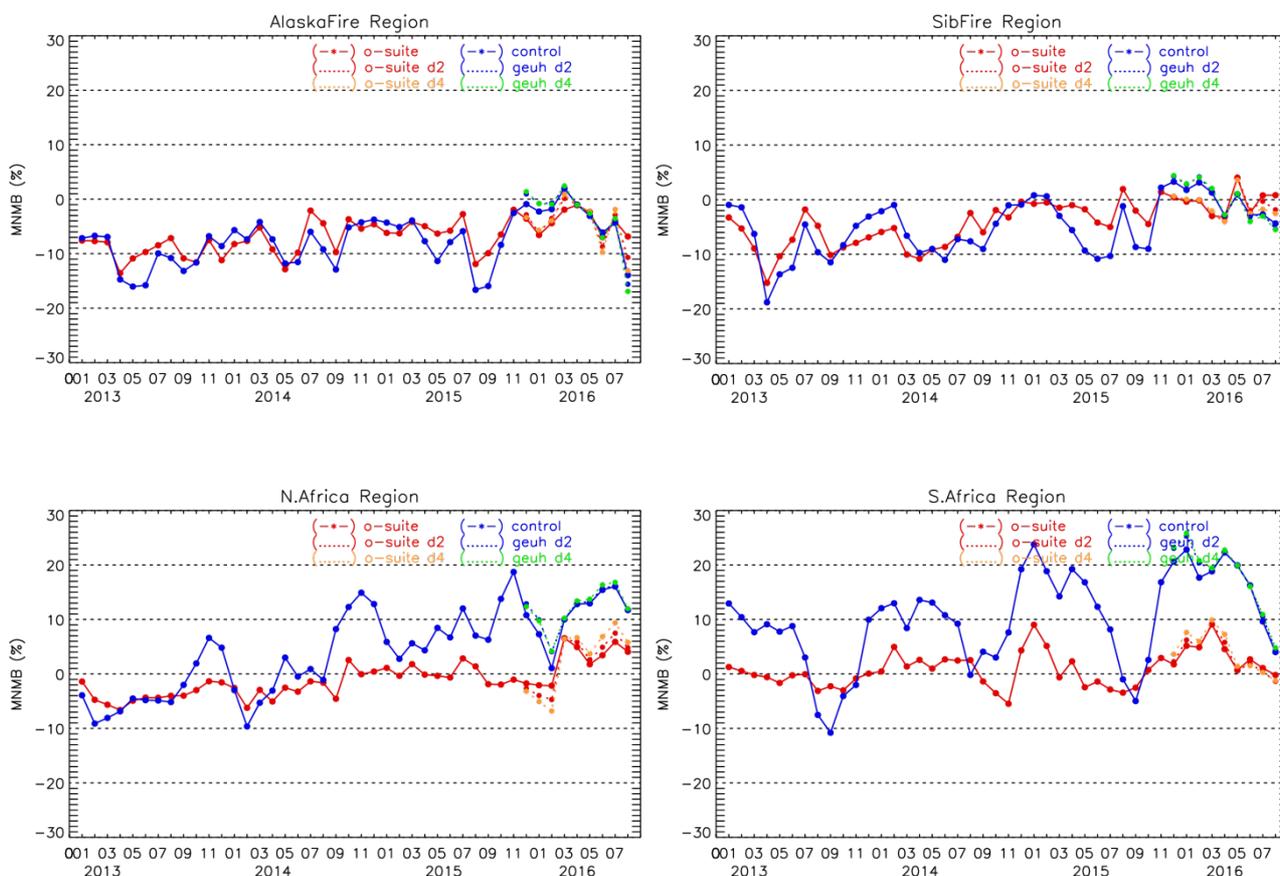


Fig. 3.3.14: Modified normalized mean bias (%) for CO total column from the model simulations vs MOPITT V5 retrievals over selected regions. O-suite (red, solid), control run (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue, dotted), control 4th forecast day (green, dotted).

3.3.5 Evaluation against TCCON CO

For the validation column averaged mole fractions of CO (denoted as XCO) from the Total Carbon Column Observing Network (TCCON) are used. Column averaged mole fractions provide a different information content than the in situ measurements and are therefore complementary to the in situ data. The observations are compared with the high-resolution CO simulations, the o-suite, as well as the control run. At Bialystok and Orleans all model simulations overestimate the XCO (Fig. 3.3.15 and 3.3.16). The seasonality is in general well represented by these models.

At Reunion (3.3.17) the o-suite captures the seasonality and agrees with the measurements within 5%. The control model shows unreasonable high CO concentration for the whole period of the comparison. The high resolution FC CO model simulations show a good agreement with the measurements for the period August 2015 to March 2016. Starting in April 2016 the agreement to the measurements worsens again and the offset is similar to the period March 2015 – July 2015.

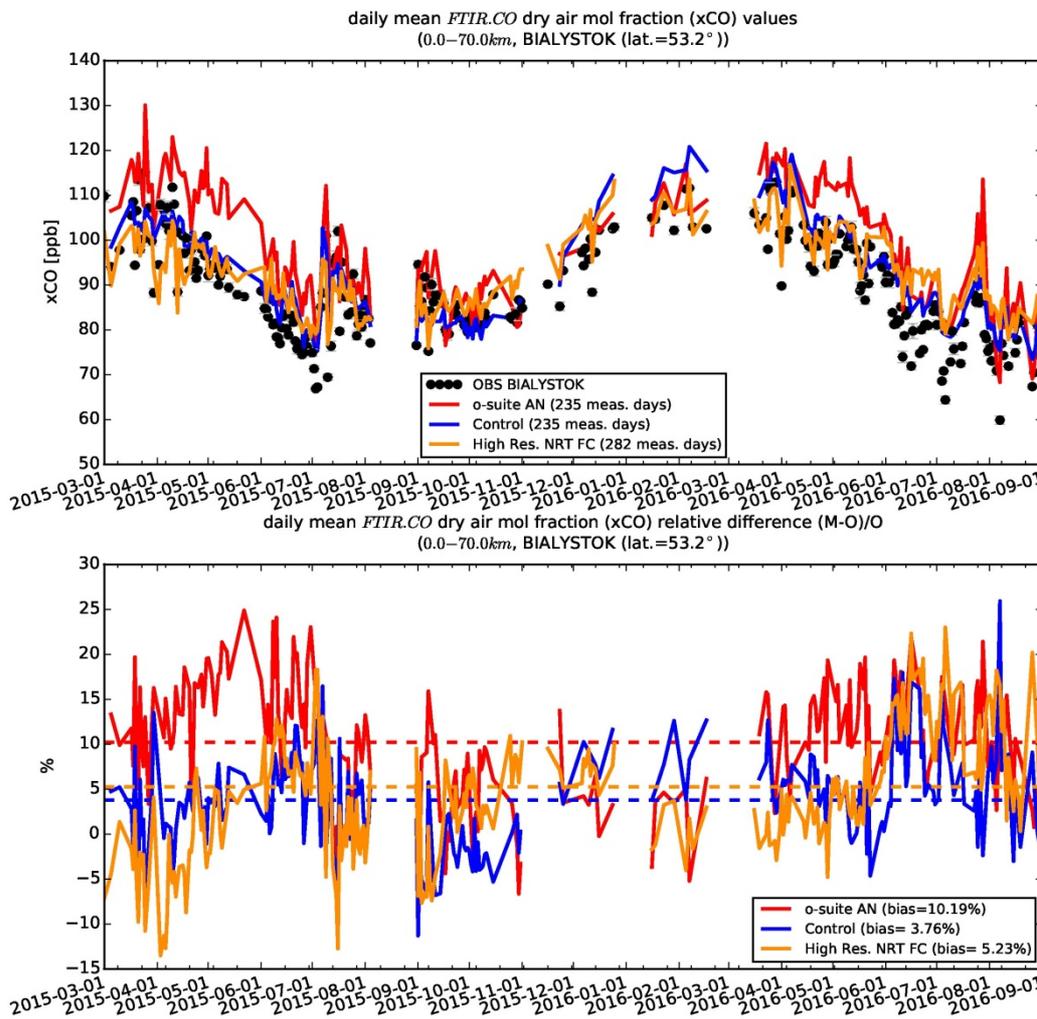


Figure 3.3.15: Time series and relative difference of column averaged mole fractions of carbon monoxide (CO) at the TCCON site Bialystok compared to the o-suite (red), control (blue) and the high resolution NRT FC model (yellow).

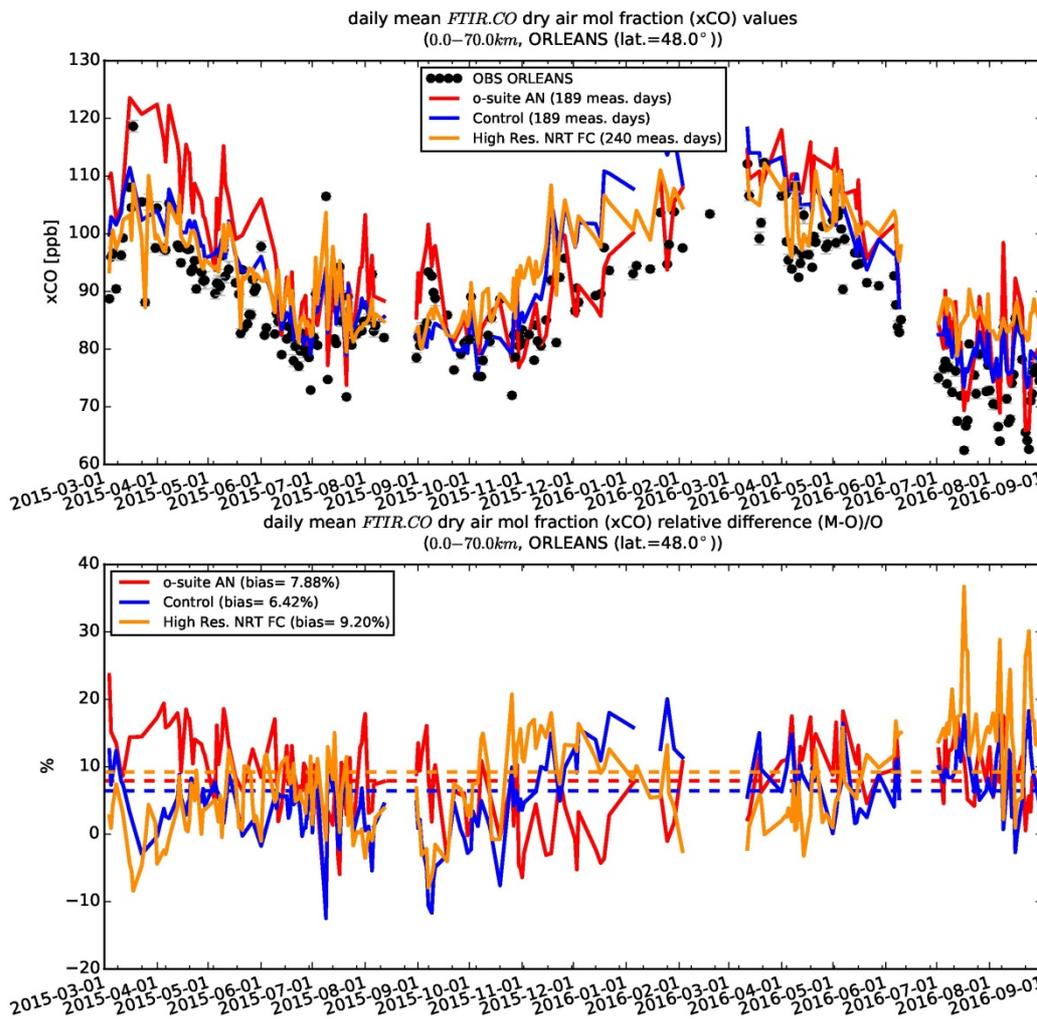


Figure 3.3.16: Time series and relative difference of column averaged mole fractions of carbon monoxide (CO) at the TCCON site Orleans compared to the o-suite (red), control (blue) and the high resolution NRT FC model (yellow).

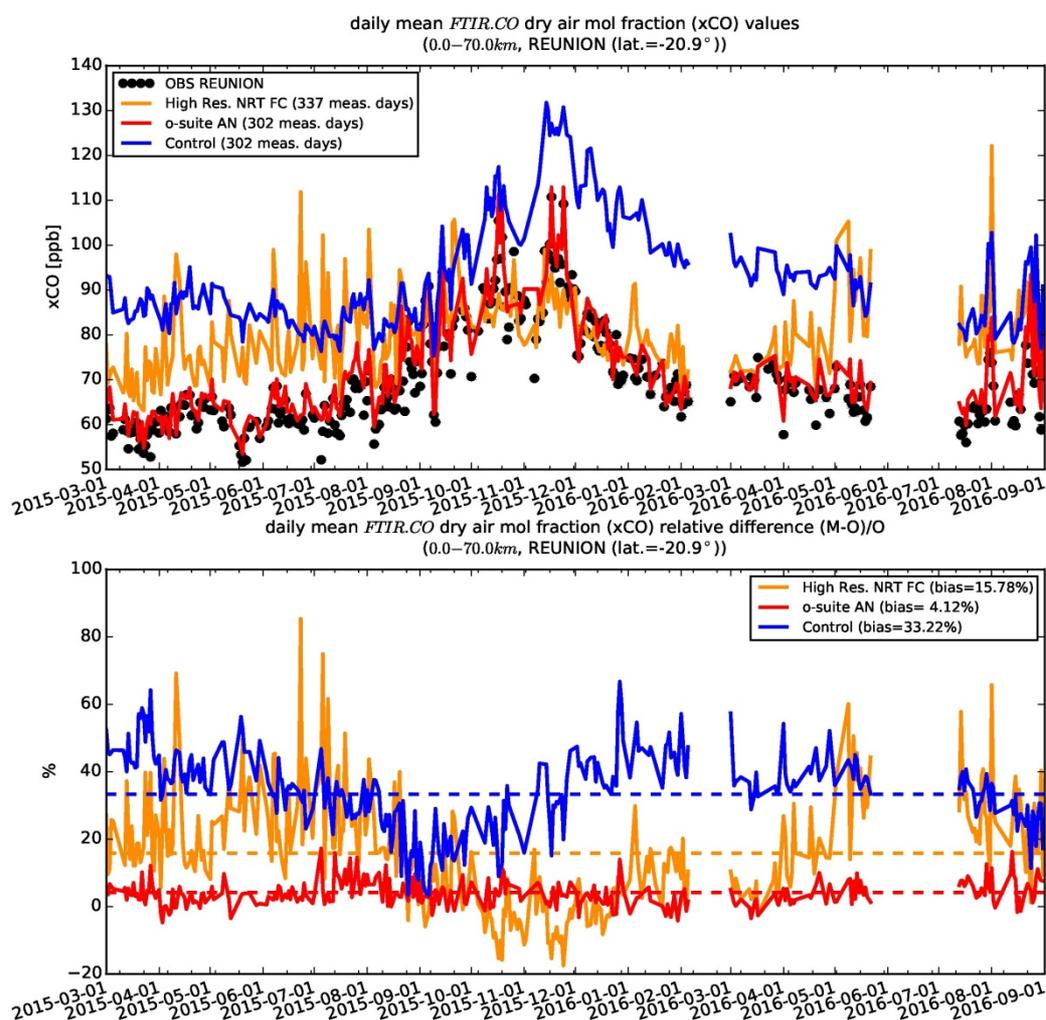


Figure 3.3.17: Time series and relative difference of column averaged mole fractions of carbon monoxide (CO) at the TCCON site Reunion compared to the o-suite (red), control (blue) and the high resolution NRT FC model (yellow).

3.3.6 Validation with AirBase observations in Mediterranean

The surface Carbon Monoxide validation over the Mediterranean is based on an evaluation against station observations from the Airbase Network (<http://acm.eionet.europa.eu/databases/airbase/>). In addition, observations from Agia Marina station in Cyprus (<http://www.airquality.dli.mlsi.gov.cy/>, provided by the Department of Labour Inspection - Ministry of Labour and Social Insurance, of Cyprus), the Navarino Environmental Observatory (<http://www.navarinoneo.gr/index.php/en/>) station in Messene Greece, as well as Finokalia station in Crete, Greece are used. For the validation analysis, only stations in the Mediterranean located within about 100 km from the shoreline of the Mediterranean shore are used. Table 3.3.2 shows the station names, coordinates, elevation and the MNMBs and correlations obtained with the 2 forecast runs (o-suite and control). It indicates that the variance explained by each station of both the o-suite and control is high and correlations are highly significant over Eastern Mediterranean (note the exceptional high correlation over Finokalia). On the contrary over stations Vandellòs l'Hospitalet de l'Infant in the Mediterranean shore of Spain and Gharb in Malta, correlations are lower.

Table 3.3.2: Coordinates, elevation, corresponding model level (level 60 is the surface level), as well as validation scores (MNMBs and correlations for the period JJA 2016) obtained with the 2 forecast runs (o-suite and control), for each one of the selected Mediterranean stations. MNMBs and correlations with blue denote stations where control run performs better while with red are denoted stations where o-suite performs better.

Station Name	Stat_ID	Lon	Lat	Alt (m)	Level	Distance from the shore (km)	MNMB		Cor. Coef	
							o-suite	contro	o-suite	control
Vandellòs i l'Hospitalet de l'Infant	ES1854A	0.83	41.01	189	54	7	-27.8	-27.2	0.32	0.07
Gharb	MT00007	14.20	36.07	114	58	31	-25.6	-26.8	0.21	0.25
NEO	-	21.67	37.00	50	60	2	-1.9	-5.8	0.44	0.41
Finokalia	GR0002R	25.67	35.32	250	55	4	14.5	11.7	0.93	0.86
Agia Marina	CY0002R	33.06	35.04	532	53	14	-19.7	-20.3	0.56	0.60

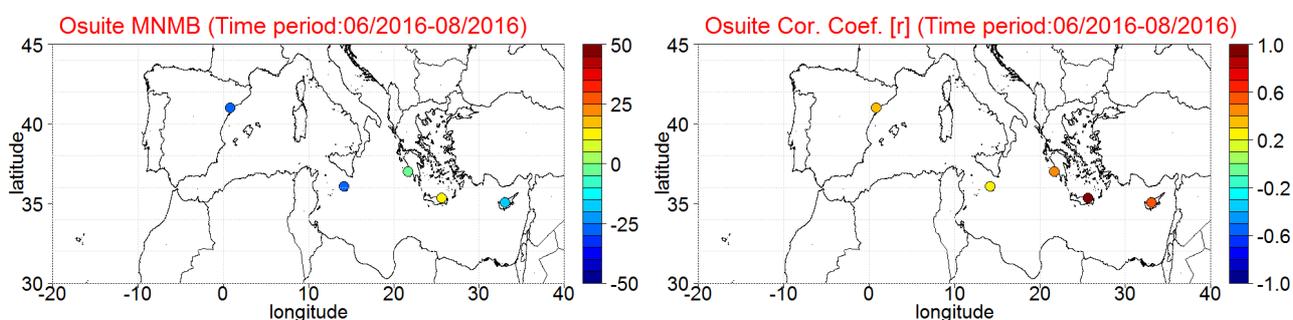


Figure 3.3.18: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 June 2016 to 31 August 2016.

In terms of biases, both runs underestimate CO values over Spain and Malta (MNMBs exceed -25%) as well as over Agia Marina station in Cyprus (MNMB=-20%). On the contrary both runs overestimates CO values over Finokalia by about 10% and finally over NEO station both the o-suite and the control reproduce well surface CO mean concentrations.

The spatial distribution of MNMBs and correlations of the o-suite over the Mediterranean is shown in 3.3.18, where the contrast in the model performance between Mediterranean Western and Central Mediterranean and the Eastern Mediterranean is evident in both correlations and biases (higher correlations and MNMBs closer to zero over the Eastern Mediterranean).

The time series in 3.3.19 show considerable day-to-day variability in the stations in Spain and Malta, where the model reproduces the lower values observed and has an overall negative bias. In Finokalia the short-term variability is much less, and the model does a good job to reproduce the variability.

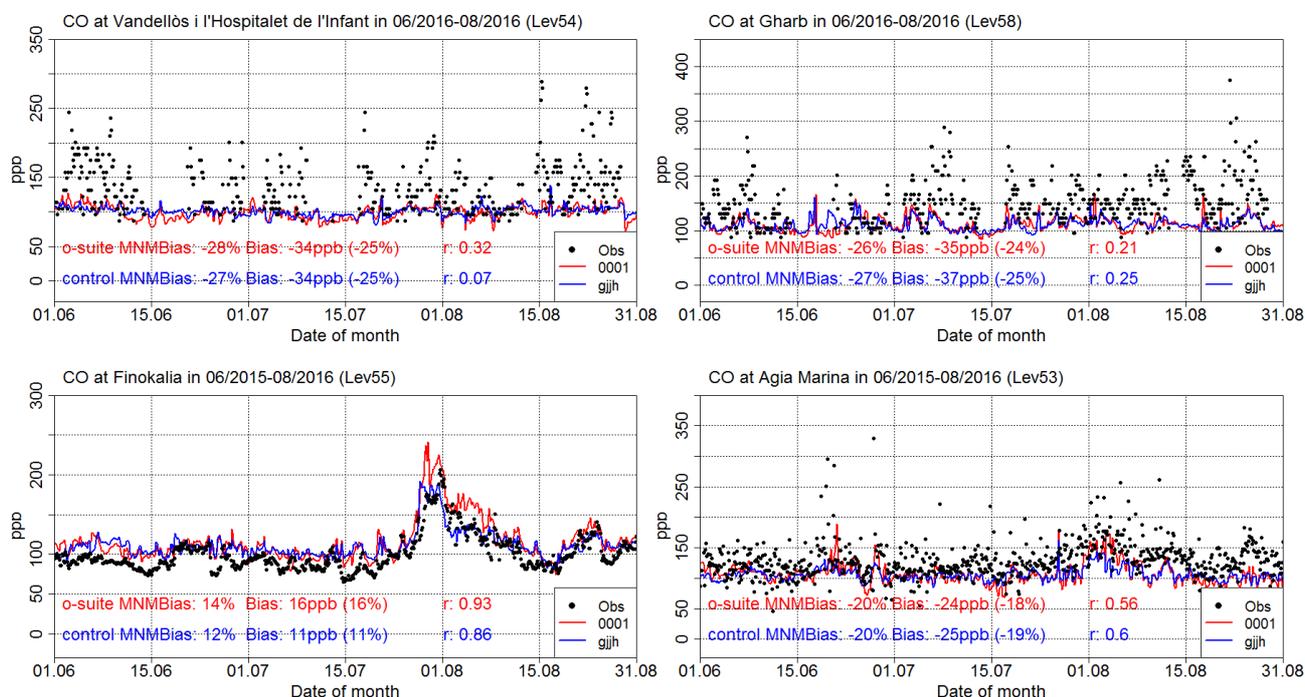


Figure 3.3.19: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Vandellòs l' Hospitalet de l' Infant, Spain station (41.01°N, 0.89 °E, top left), at Gharb, Malta station (36.07°N, 14.20°E, top right), at Finokalia, Crete station (35.32°N, 25.67°E, low left), and compared to observations provided by the Department of Labour Inspection - Ministry of Labour and Social Insurance of Cyprus, at Agia Marina station (35.04°N, 33.06 °E, low right).

3.4 Formaldehyde

3.4.1 Validation against satellite data

In this section, simulations of tropospheric formaldehyde are compared to SCIAMACHY/Envisat HCHO satellite retrievals (IUP-UB v1.0) [Wittrock et al., 2006] for model data before April 2012 and to GOME-2/MetOp-A HCHO data (IUP-UB v1.0) [Vrekoussis et al., 2010] afterwards. As the retrieval is performed in the UV part of the spectrum where less light is available and the HCHO absorption signal is smaller than that of NO₂, the uncertainty of monthly mean HCHO columns is relatively large (20% – 40%) and both noise and systematic offsets have an influence on the results. However, absolute values and seasonality are retrieved more accurately over HCHO hotspots.

In Figure 3.4.1, monthly mean satellite HCHO columns are compared to model results for July 2016. The magnitude of oceanic and continental background values and the overall spatial distribution are well represented by o-suite and control. Compared to GOME-2 satellite retrievals, there is an overestimation of values for Central Africa which may be due to an overestimation of fire emissions in this region. As for tropospheric NO₂ (section 3.2), boreal fire emissions over Siberia are in some parts overestimated, but the agreement for the (at least for NO₂) weaker boreal fire emissions over Canada is good.



Figure 3.4.2 shows differences between some selected forecast days of the models for July 2016. As described in section 3.2, model performance does not change much from forecast day 1 to later steps compared to GOME-2 satellite retrievals which is also true for time series comparisons described in the next paragraphs.

Time series in Fig. 3.4.3 highlight three cases:

- East-Asia and the Eastern US, where HCHO is dominated by biogenic emissions. Model results and measurements generally agree rather well. However, all model runs underestimate the yearly cycle over East-Asia since 2012. In contrast to MOZART runs, MACC_CIFS_TM5 overestimates satellite values for the Eastern US since the middle of 2013. However, the newer CIFS-CB05 runs perform well for Eastern US since 2015. For recent years and both regions, there is virtually no difference between the most recent o-suite run with CIFS-CB05 chemistry and the corresponding control runs without data assimilation. The variability or “ups and downs” in HCHO columns observed by GOME-2 since December 2014 is due to the lack of data (caused by instrument degradation) for these regions during Northern Hemisphere winter months (see Figure 3.4.1 for an example). This also explains the negative values in the GOME-2 time series for Eastern US in December 2015 and January 2016. Summertime maxima are still overestimated by the now, higher resolution runs for both regions in 2016.
- North-Africa, where biomass burning as well as biogenic sources largely contribute to HCHO and its precursors. Satellite observations over North-Africa are generally overestimated by CIFS-CB05 chemistry model runs but are in good agreement with the retrievals for the latest higher resolution model versions for July and August 2016..
- Indonesia, where HCHO is also dominated by biogenic sources and biomass burning. Models generally overestimate satellite values here (by a factor of 3 – 4 in the second half of 2010) and fail to reproduce the observed seasonality. This may be due to the use of fire emissions including El Nino years which experience much larger fire activities. MOZART simulations and observations agree much better since late 2012. CIFS-CB05 runs agree very well with satellite retrieved ones for December 2014 to August 2015. For September and October 2015, satellite retrieved HCHO columns show a pronounced maximum. 2015 was a strong El Nino year, which caused droughts and higher fire activity in Indonesia. As for previous El Nino years, fire emissions used by CIFS-CB05 seem to be largely overestimated, resulting in model simulated HCHO columns which are almost twice as large as those retrieved by GOME-2. Further investigations (see previous reports) show that this is not caused by cloud flagging applied to the satellite and model data.

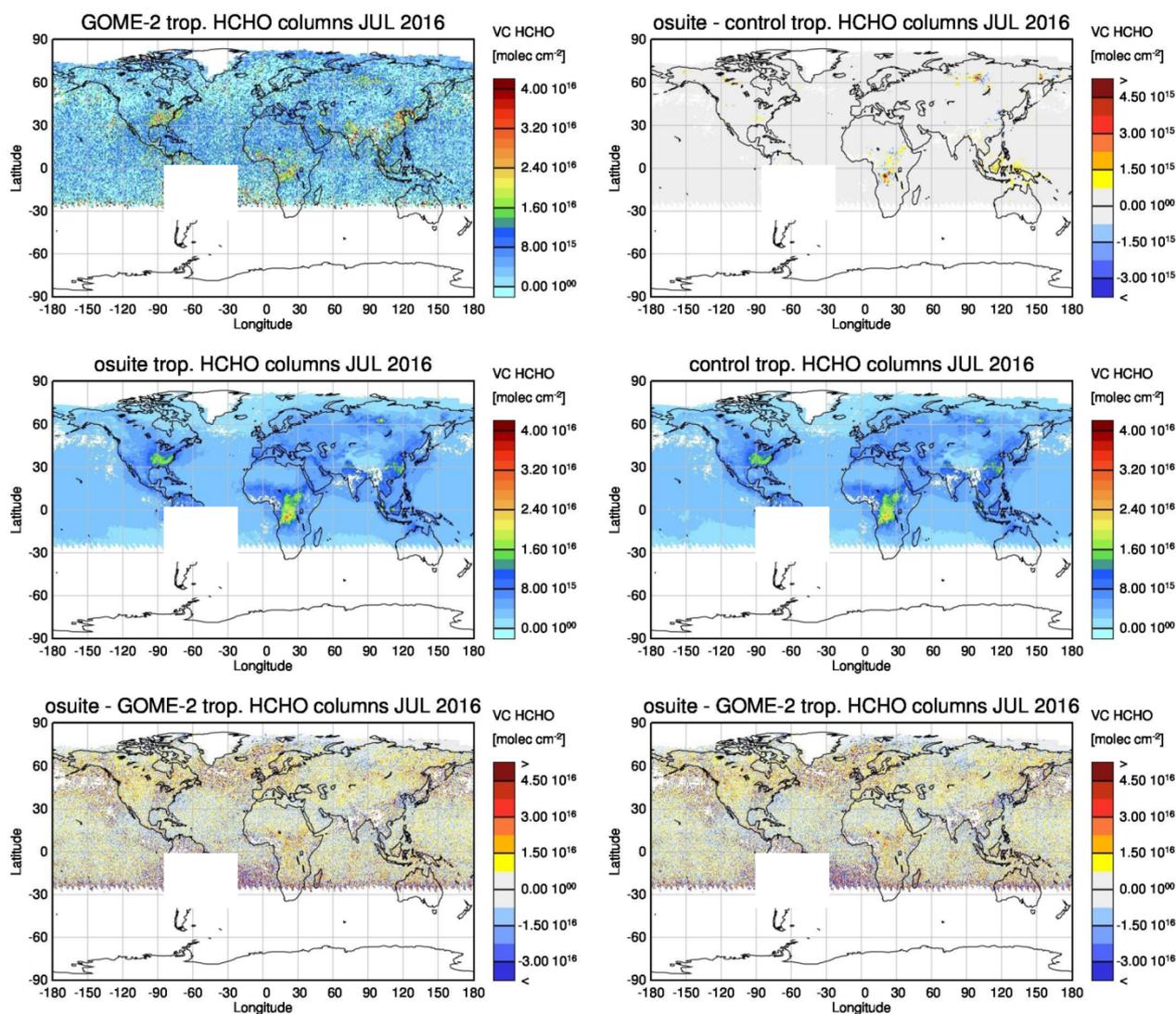


Figure 3.4.1: Global map comparisons of satellite retrieved and model simulated tropospheric HCHO columns [molec cm⁻²] for July 2016. The top row shows monthly mean tropospheric HCHO columns retrieved by GOME-2, the second row shows the same but for model simulated averages. The third row shows differences of monthly means between models and GOME-2. GOME-2 data were gridded to model resolution (i.e. 0.75° deg x 0.75° deg). Model data were treated with the same reference sector subtraction approach as the satellite data. Satellite retrieved values in the region of the South Atlantic anomaly are not valid and therefore masked out (white boxes in all images except those which show model results only).

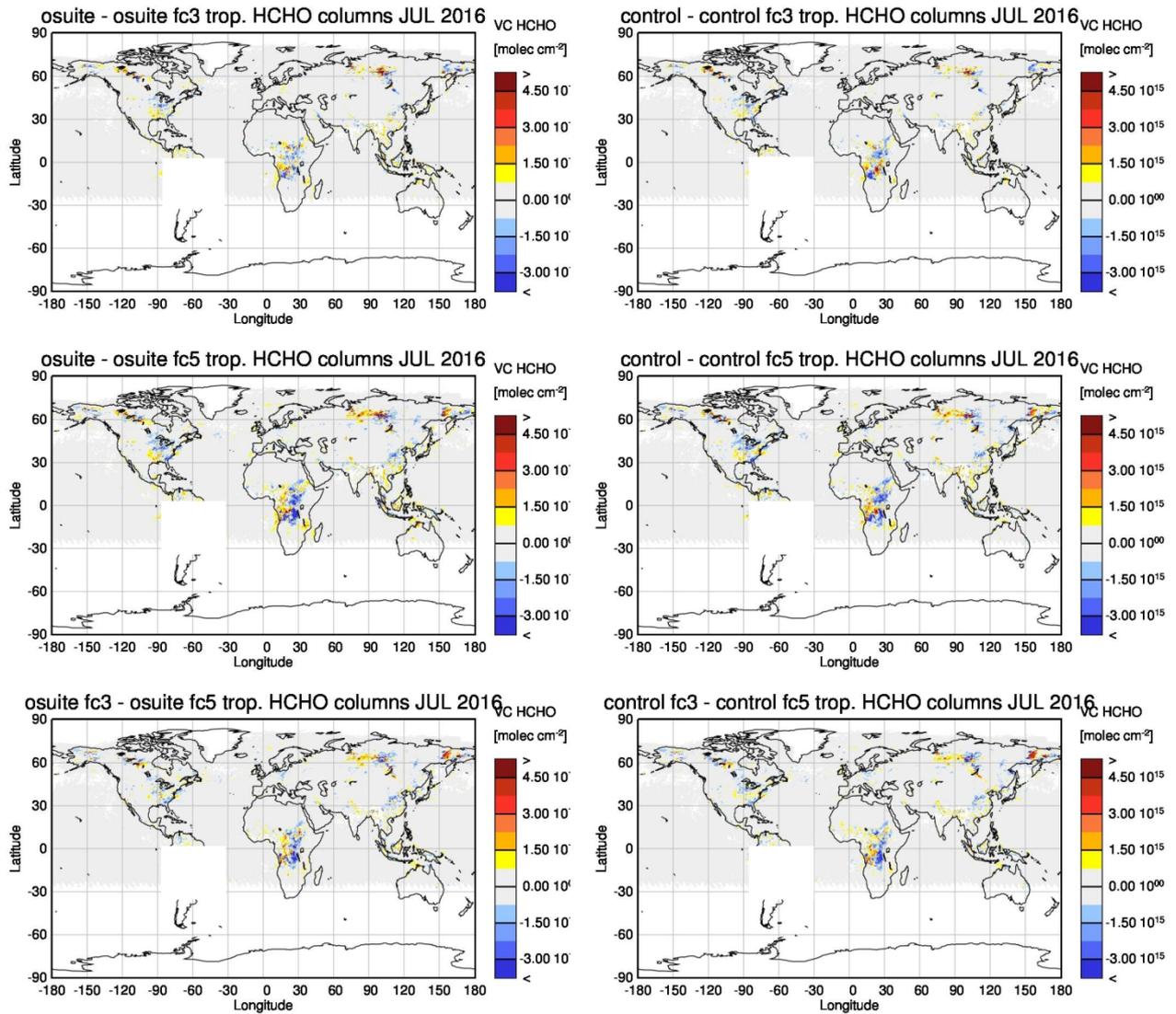


Figure 3.4.2: As in Figure 3.2.1, but for (top) differences between model runs for forecast day 1 and 3, (middle) forecast day 1 and 5, (bottom) forecast day 3 and 5.

Details on the HCHO evaluation can be found at:

http://www.doas-bremen.de/macc/macc_veri_iup_home.html.

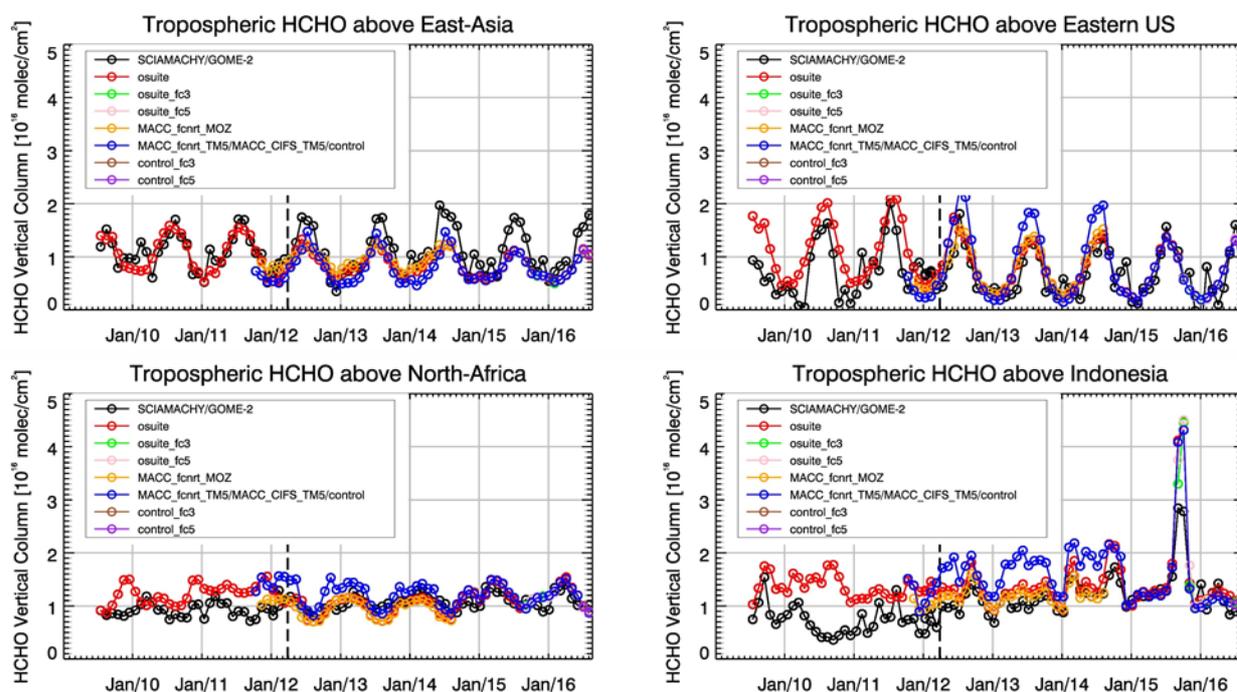


Figure 3.4.3: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions. The blue line shows MACC_fcncr_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 to August 2014 and control results from September 2014 onwards. The regions differ from those used for NO₂ to better focus on HCHO hotspots: East-Asia (25–40°N, 110–125°E), Eastern US (30–40°N, 75–90°W), Northern Africa (0–15°N, 15°W–25°E) and Indonesia (5°S–5°N, 100–120°E). Negative satellite retrieved values over Eastern US are due to a lack of data (caused by instrument degradation) during Northern Hemisphere winter months for this region. Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

3.4.2 Validation against UVVIS DOAS observations from the NDACC network

In this section, we compare the HCHO profiles of the CAMS models with UVVIS DOAS measurements at Haute Provence (43.9°N, 5.71°E, rural station, altitude 650m) and Xianghe (39.8°N, 117°E, station near Beijing, altitude 92m). Due to instrument failure, the Uccle (50.8°N, 4.36°E, urban) measurements are not displayed. These ground-based, remote-sensing instruments are sensitive to the HCHO abundance in the lower troposphere, up to 1km altitude. Tropospheric HCHO profiles and columns are validated (up to 3.5km). A description of the instruments and applied methodologies is the same as for the MWR O₃ and FTIR O₃ and CO validations see <http://nors.aeronomie.be>.

It is important to mention here that the model partial column values between the surface and 3.5 km are calculated for the smoothed model profiles (see Fig. 3.4.4, left). This guarantees that the model levels, where the measurement is not sensitive, do not contribute to the observed bias. In this specific situation the smoothing of the model profiles implies a strong increase of the model column data by the MAXDOAS apriori (and only the relative difference plots should be considered). We should mention that the measurement data is still catalogued as rapid delivery and not in the consolidated NDACC database. The measurements have been quality filtered on cloud conditions: only measurements under “clear sky” and “thin clouds” are used (see Gielen et al., 2014).

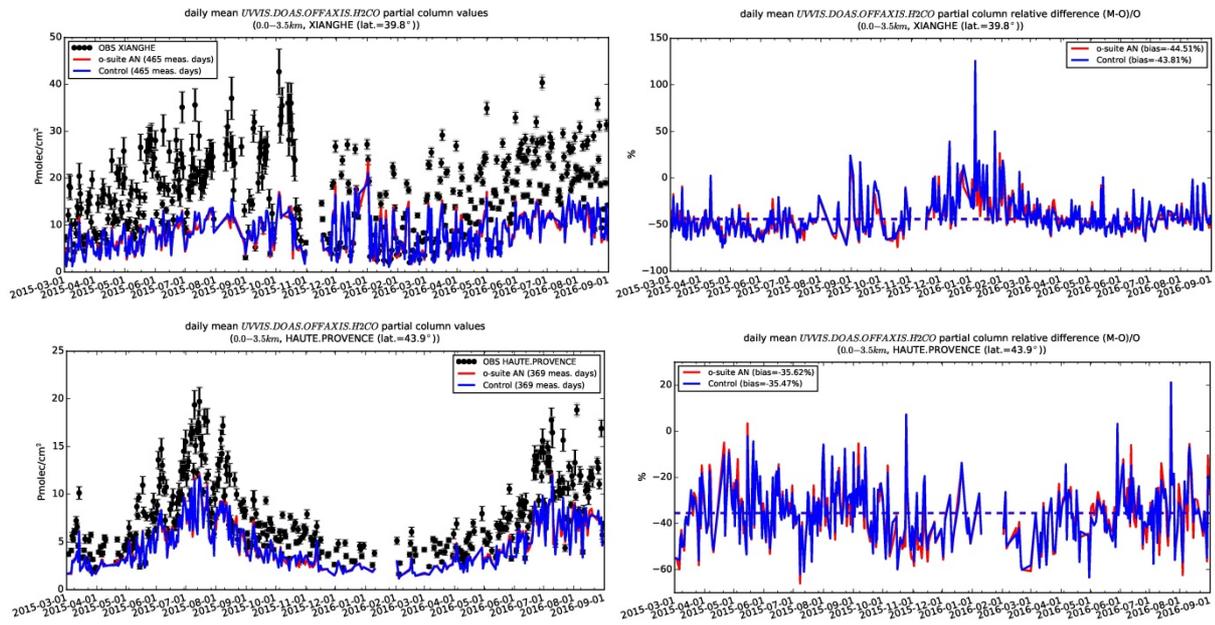


Figure 3.4.4: Daily mean relative differences of tropospheric HCHO columns (till 3.5km) by the o-suite (red) and the control run (blue) compared to NDACC UVVIS DOAS data at at Xianghe (39.8°N, 117°E, station near Beijing, altitude 92m) and Haute Provence (43.9°N, 5.71°E, rural station, altitude 650m, bottom) for the period March. 2015 –August 2016. The number of measurements and median of differences is indicated in the legend (the overall measurement uncertainty is 10%).

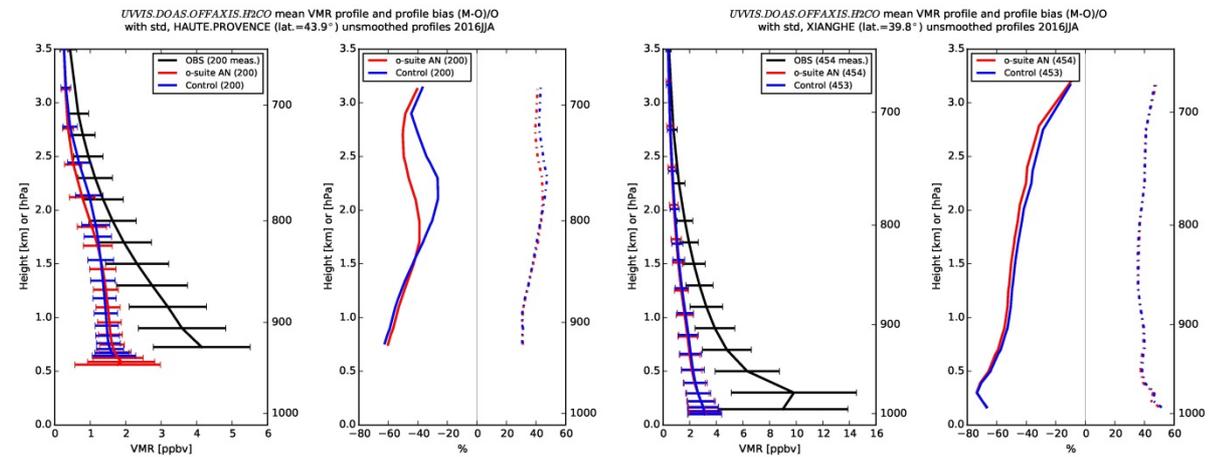


Figure 3.4.5: Mean tropospheric HCHO profiles by the o-suite (red) and the control run (blue) compared to NDACC UVVIS DOAS data at Haute Provence (43.9°N, 5.71°E, left) and at Xianghe (39.8°N, 117°E, right) for the period June-August 2016.

From Fig. 3.4.4 and Fig.3.4.5 we see little difference between the o-suite and the control run. Both models underestimate the observations below 1km. Although the background column values are well captured by the models, the high emission events are not (see Fig. 3.4.5).

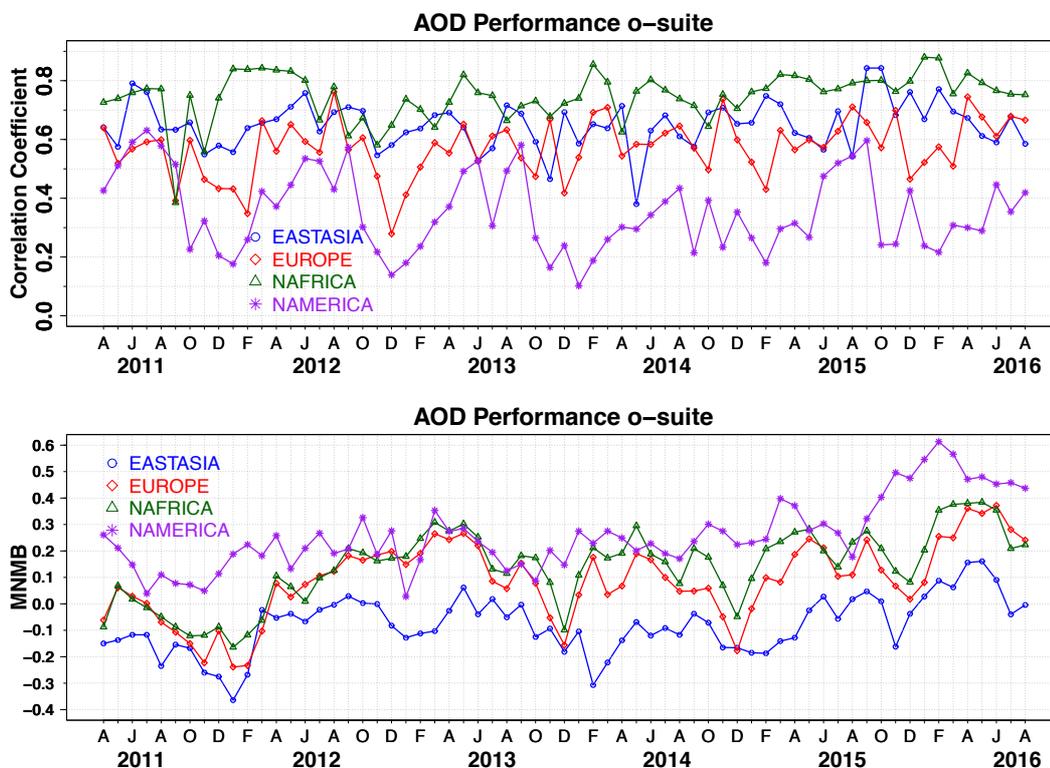


Figure 3.5.1 a) Correlation coefficient and b) modified normalized mean bias (MNMB) in AOD, since 2011, based on daily AOD comparison in four world regions [Eastasia(blue); Europe(red); NAfrica(green); NAmercia(purple)] for the o-suite.

3.5 Aerosol

3.5.1 Global comparisons with Aeronet and PM

Standard scores, maps, scatterplots, bias maps, time series comparison and histograms illustrating the performance of the aerosol simulation in the IFS system are made available through the AeroCom web interface:

http://aerocom.met.no/cgi-bin/aerocom/surfobs_annualrs.pl?PROJECT=CAMS&MODELLIST=CAMS-VALreports . The model run can be compared to the MACC reanalysis (available until Dec 2012) and the AeroCom Median model. A daily updated comparison against 30 selected Aeronet stations is available via the ECMWF CAMS service website:

<http://www.copernicus-atmosphere.eu/d/services/gac/verif/aer/nrt/>.

Correlation, based on daily aerosol optical depth and NRT Aeronet observations, is rather stable since 2011, exhibits significant variation and seems to have increased recently. The o-suite forecast at +3 days shows slightly lower correlation, as expected. See figure S3. Part of the month-to-month variation in correlation is due to the limited quality of the NRT Aeronet data, which are of preliminary nature. Retrospective analysis since the year 2011 shows that this level 1.5 NRT AOD Aeronet data, due to undetected cloud contamination and any uncorrected drift, are on global average +20% higher than quality assured level 2.0 data. However, using the MNMB bias score such bias is not as visible, because outliers have less impact. Since 2014 the CAMS model MNMB type of bias against level 2.0 data was +5-10% higher than that against level 1.5 data (see figure S3). Figure

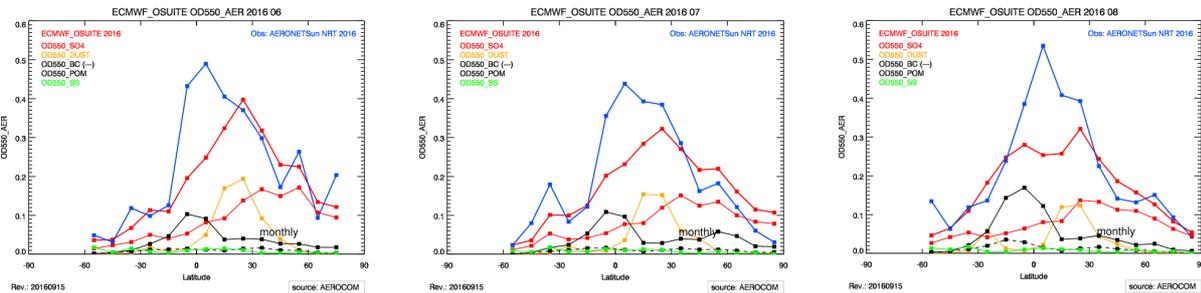


Figure 3.5.2: Aerosol optical depth of o-suite (red) compared to latitudinally aggregated NRT Aeronet level 1.5 data (blue) for the three months covered by this report.

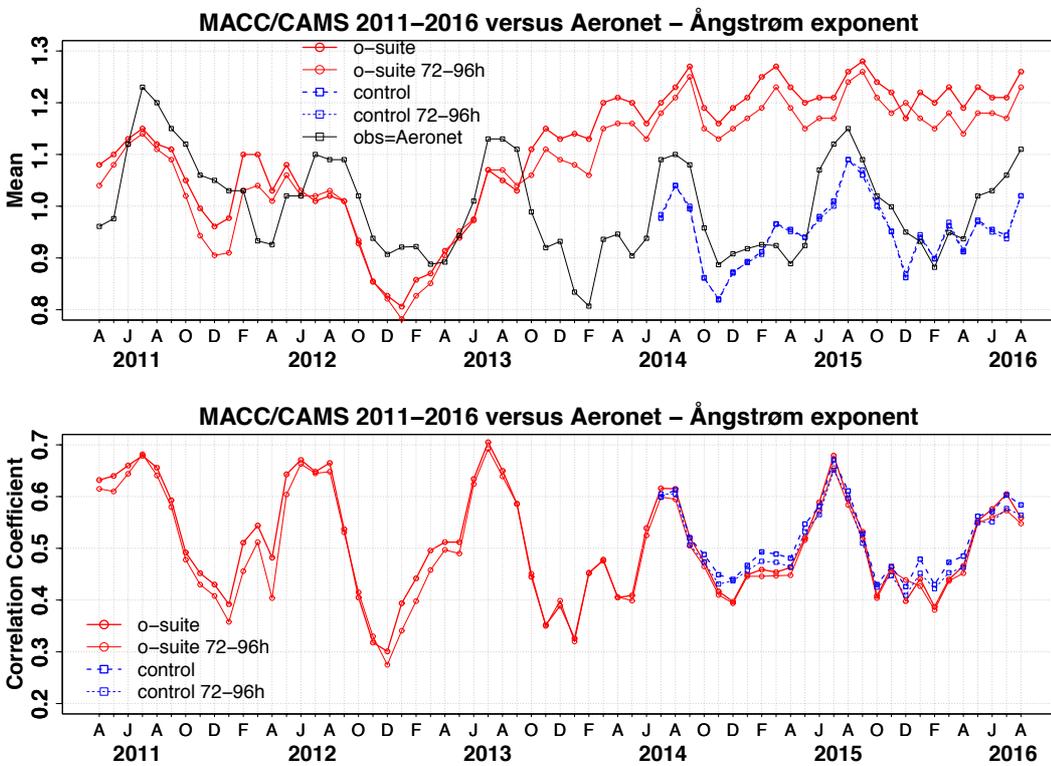


Figure 3.5.3 a) Evolution of mean Ångström exponent in o-suite and control at Aeronet sites, based on matching monthly mean values. o-suite (thick red curve); o-suite at last forecast day (light red curve); control (blue dashed curve); control at last forecast day (light blue dashed curve). b) Correlation using daily matching Ångström exponent.

S3 also shows the evaluation against level 2.0 data for the whole time period. Note that an establishment of a more precise correction of bias in the last months is rather difficult because of few level 2.0 data being available.

The regional performance of the o-suite model exhibits some seasonal cycle in AOD depending on region (Fig. 3.5.1 a). For instance, the model performance in the North American winter season with respect to correlation seems to be worst. In North America the low correlation in winter increasing into spring may be due to large uncertainties in satellite observations over bright land targets, which may not provide enough guidance to the IFS assimilation system, or missing model components such as nitrate. Noteworthy is also the persistent AOD overestimation over North

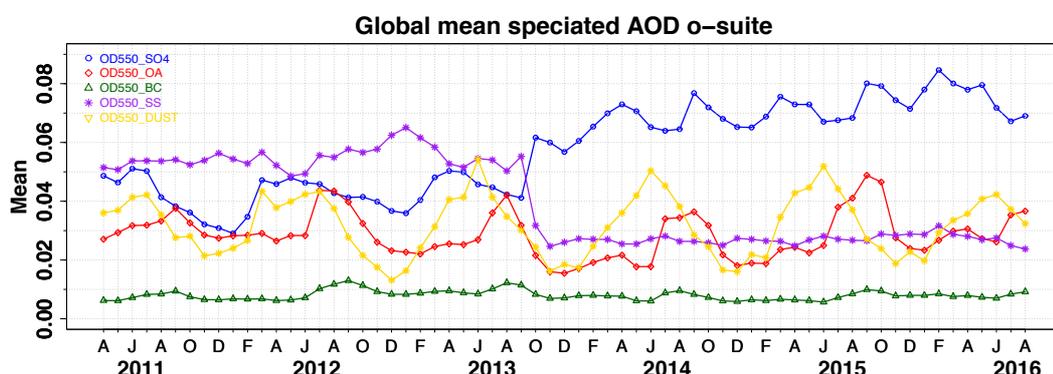


Figure 3.5.4: Evolution of aerosol component’s AOD@550nm [OD550_SO4 = sulphate(blue); OD550_OA = organics(red); OD550_BC = black carbon(green); OD550_SS = sea salt(purple); OD550_DUST = dust(yellow)].

America (Fig. 3.5.1 b). The latitudinal display of model and Aeronet AOD in the period investigated here (Fig. 3.5.2) shows the negative bias against Aeronet NRT in tropical and sub-tropical regions.

The simulated aerosol size distribution may be validated to first order using the wavelength dependent variation in AOD, computed as Ångström exponent, with higher Ångström exponents indicative of smaller particles. Figure 3.5.3 a) shows the temporal evolution of simulated and observed mean Ångström exponent, while the correlation is found in figure 3.5.3 b). We find in JJA 2016 a positive bias of +20% (against -5% before October 2013). Temporal and spatial variability is rather high and correlation is lower than for AOD (Figure 3.5.3 b). Figure 3.5.4 shows that the Oct 2013 model changes are responsible for this shift in Ångström exponent. Less sea salt and more

Table 3.5.1: Mean global total and speciated AOD in the o-suite for the last two periods covered by the VAL report and change after 3 forecast days.

	o-suite		o-suite	
	Mean	Change wrt to	Mean	Change wrt to
	MAM 2016	first day	JJA 2016	first day
	0-24h	on day 4	0-24h	on day 4
AOD@550	0.181	-15%	0.173	-12%
BC-OD@550	0.008	-14%	0.008	-21%
Dust-OD@550	0.037	-11%	0.037	+8%
OA-OD@550	0.029	-10%	0.033	-15%
SO4-OD@550	0.079	-26%	0.069	-22%
SS-OD@550	0.028	0%	0.025	-4%

sulphate shift the size distribution to smaller sizes. AOD due to sea salt decreased by 50%, that to due organics decreased by 25%, while that of sulphate increased by 40%.

The o-suite uses data assimilation to obtain a first guess aerosol field. In the forecast period, however, a-priori model parameterisations and emissions (except fire emissions, which are kept in the forecast equal to the latest GFAS emission values) determine more and more the shape and amplitude of the aerosol fields. The performance of the day three forecasted AOD fields as compared to the first guess is shown in Figure S3 in the summary of this report. Against Aeronet the o-suite forecast for day three has little overall positive bias in AOD while the control run with no



assimilation shows significant less AOD (-50% compared to o-suite, see figure S3). Table 3.5.1 shows an average global decrease in total aerosol optical depth of 12% during the first four forecast days, dominated by sulphate and organics, supporting the conclusion that either a-priori sources are too small or sinks are too effective in the IFS model.

Surface concentration of particulate matter below 10 µm (PM10) from the o-suite experiment have been validated against data from 150 background IMPROVE and EMEP stations (Figure 3.5.1). A climatological average has been constructed from data in the period 2000-2009 as available in the EBAS database hold at NILU. The data coverage is not the same at all stations, and sometimes covers only a few years. All used time series used are documented via the CAMS-AeroCom web interface.

In contrast to earlier validation reports have we taken since 2016 the PM10 concentrations as diagnosed by the IFS model in the mars archive, while before we have constructed a PM10 concentration from all available aerosol mass. This changes the bias evaluation considerably. The bias maps show that both in North America and Europe high bias appears at few stations located in regions close to the coastlines. This is an indication that simulated PM10 concentrations may be high due to sea salt aerosols. Regional models using the sea salt concentrations as boundary condition should take over the PM10 definition as used in IFS. Inner-continental sites indicate a negative MNMB bias of -30% both in Europe and North America.

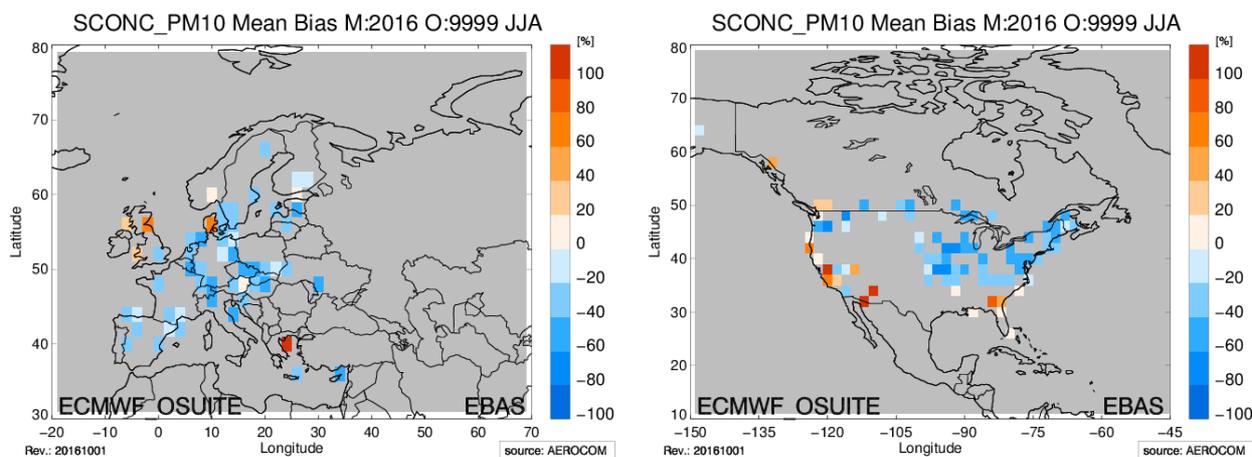


Figure 3.5.1: Bias [%] map of June/July/August mean PM 10 concentrations at EMEP (Europe) and IMPROVE sites (North America); simulated o-suite versus climatological average (2000-2009).

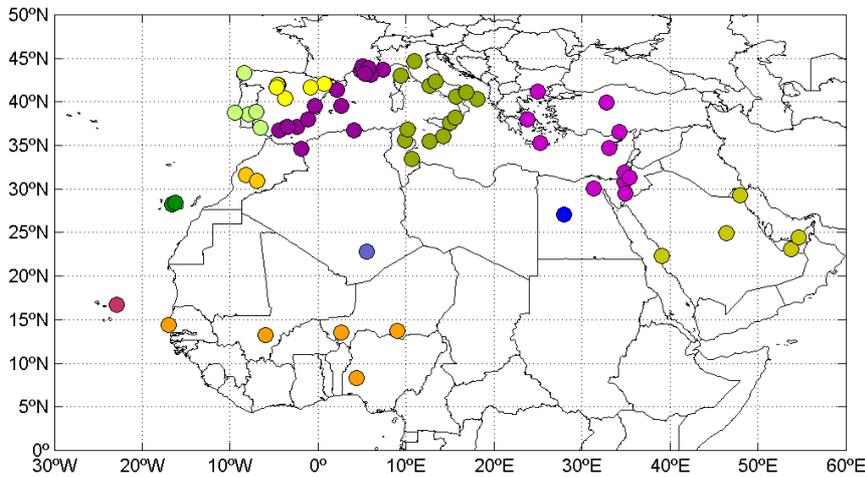


Figure 3.5.2: Map of 71 AERONET level-1.5 stations used in this analysis. The twelve regions considered in the analysis are shown by different colours

3.5.2 Dust forecast model intercomparison: Validation of DOD against AERONET, and comparisons with Multimodel Median from SDS-WAS

72 hour forecasts (on 3-hourly basis) dust aerosol optical depth (DOD) from CAMS o-suite and control experiments have been validated for the period 1 June – 31 August 2016 against 71 AERONET stations grouped in twelve regions (Figure 3.5.2), MODIS aerosol product available through the NASA’s EOSDIS system (MCDAODHD files) and compare with the SDS-WAS Multi-model Median DOD. The SDS-WAS Multi-model Median DOD is obtained from eleven dust prediction models participating in the the Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS) Regional Center for Northern Africa, Middle East and Europe (<http://sds-was.aemet.es/>).

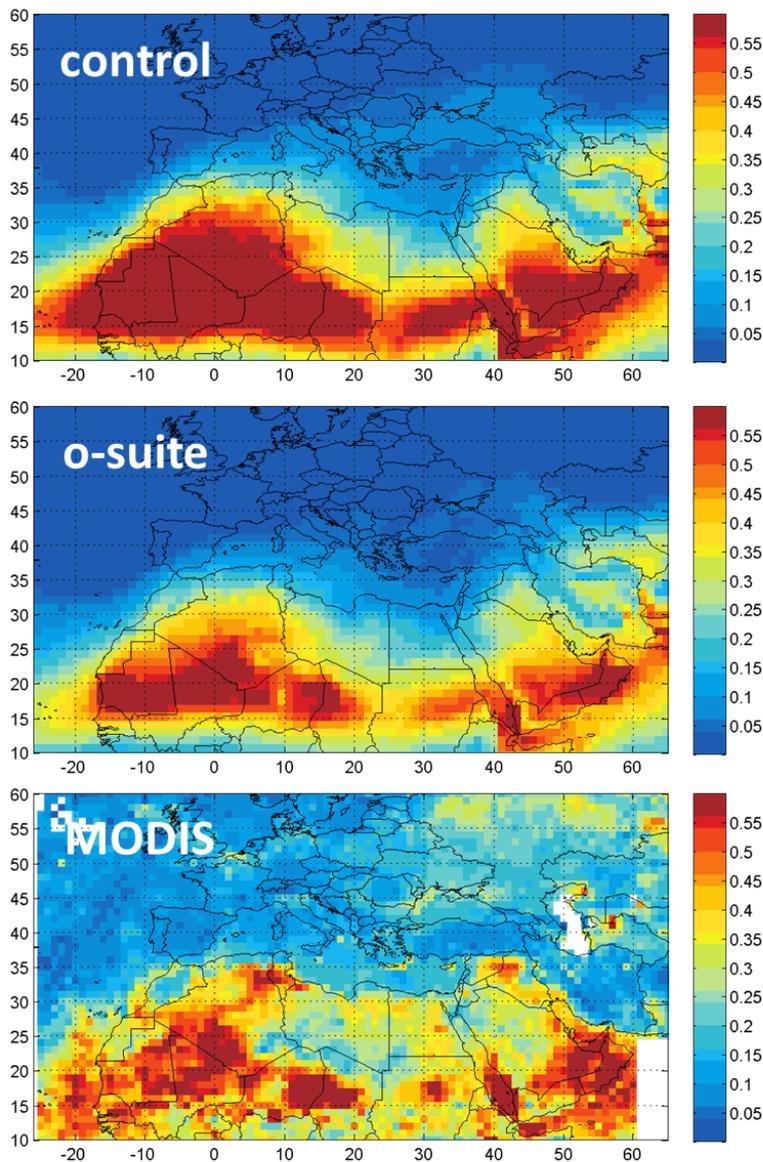


Figure 3.5.3: Averaged DOD 24h forecast from control (top) and o-suite (central) as well as AOD from MODIS combined Dark target and Deep Blue product (bottom) from June 1 to August 31, 2016.

During the period of analysis, satellites (see Figure 3.5.3) show that major dust activity is concentrated over the Sahara (in the Bodelé Basin and the Mali/Mauritania border) and the dust corridor of North Western Maghreb. CAMS model can simulate the main areas of dust activity in comparison with MODIS, although CAMS o-suite reduces the strong overestimations observed in CAMS control.

From June to September, CAMS o-suite is the model that best reproduces the daily variability of AERONET observations particularly over desert dust source regions achieving close values similar to the SDS-WAS Median Multimodel (see correlation of Sahara and Middle East in Table 3.5.2). CAMS o-suite improves the observed variability over the Middle East (the correlation values from 0.47 for



Table 3.5.2: Skill scores (MB, FGE, RMSE and r) of 24h forecasts for CAMS o-suite, CAMS control and SDS-WAS Multi-model Median for the study period, and the number of data (NDATA) used. Dust AOD (DOD) from AERONET is the reference.

	NDATA	control				o-suite DOD				SDS-WAS Median DOD			
		MB	FGE	RMSE	r	MB	FGE	RMSE	r	MB	FGE	RMSE	r
Western Mediterranean	3143	0.01	1.41	0.24	0.51	-0.04	1.43	0.23	0.52	-0.05	1.41	0.23	0.54
Tropical North Atlantic	221	0.07	0.32	0.22	0.55	-0.11	0.32	0.20	0.62	-0.12	0.31	0.20	0.64
Eastern Mediterranean	1676	0.03	1.62	0.13	0.73	-0.01	1.64	0.12	0.74	0.00	1.63	0.12	0.75
Sahel	1178	-0.07	0.48	0.50	0.31	-0.23	0.53	0.54	0.32	-0.21	0.47	0.53	0.33
Subtropical North Atlantic	982	0.11	1.12	0.25	0.63	0.03	1.06	0.18	0.63	0.03	1.06	0.19	0.62
Central Mediterranean	1918	0.05	1.39	0.18	0.73	-0.02	1.40	0.13	0.76	-0.02	1.37	0.11	0.80
Middle East	1076	-0.04	0.36	0.33	0.47	-0.08	0.34	0.31	0.56	-0.11	0.35	0.34	0.47
Iberian Peninsula	1013	0.00	1.79	0.16	0.38	-0.02	1.82	0.16	0.41	-0.02	1.82	0.16	0.41
Western Iberian Peninsula	653	-0.01	1.70	0.08	0.68	-0.02	1.72	0.08	0.70	-0.03	1.76	0.08	0.67
North Western Maghreb	714	0.05	0.46	0.23	0.61	-0.08	0.48	0.22	0.62	-0.11	0.47	0.22	0.65
Sahara	328	0.04	0.45	0.59	0.27	-0.15	0.35	0.57	0.41	-0.25	0.39	0.61	0.43
Eastern Sahara	232	0.10	0.94	0.22	0.58	0.03	0.86	0.17	0.60	0.02	0.83	0.17	0.63

control to 0.56 for o-suite, see Table 3.5.2 and Mezaira in Fig. 3.5.9) and the Sahara (correlation values from 0.27 for control to 0.41 for o-suite, see Table 3.5.2). The correlation values in Sahara are slightly lower than the previous spring season (which o-suite achieved values of 0.63). During summer is when is detected the maximum occurrence of strong and fast dust outbreaks associated with mesoscale convective systems, some of them haboobs, in the Sahara. Although the DOD trend is well reproduced by the models, both CAMS experiments and the SDS-WAS multimedian product, underestimate these dust episodes in the Sahara (see Tamanrasset in Fig. 3.5.8) since they are not capable of capturing them. These dust events are linked to the poor correlations observed in Sahel in the CAMS experiments and SDS-WAS Median Multimodel ($r \sim 0.3$, see Table 3.5.2). Therefore, despite CAMS o-suite reduces the overestimations observed in CAMS control (MB in Sahara decreases from 0.04 to -0.15 from control to o-suite in Table 3.5.2), CAMS o-suite still overestimates the observed DOD over Sahara (see Tamanrasset in Fig. 3.5.8).

Over long-range transport regions, CAMS o-suite shows the best correlations in all the regions (with correlations between 0.32 in Sahel to 0.76 in Central Mediterranean, see Table 3.5.2) in comparison with CAMS control. The skill scores of the CAMS o-suite are close to those obtained by the SDS-WAS Median Multimodel (see Table 3.5.2). Maximum dust activity is observed in Central and Western Mediterranean (see Saada in Fig. 3.5.8) and subtropical North Atlantic (see Santa Cruz de Tenerife in Fig. 3.5.8). The Mediterranean is the region where the highest correlations are achieved ($r > 0.70$).



Table 3.5.3: Skill scores (MB, FGE, RMSE and r) of 48h and 72h forecasts for CAMS o-suite and CAMS control for the study period, and the number of data (NDATA) used. Dust AOD (DOD) from AERONET is the reference.

	NDATA	48h control				48h o-suite				72h control				72h o-suite			
		MB	FGE	RMSE	r												
Western Mediterranean	3143	0.02	1.46	0.27	0.39	-0.03	1.48	0.25	0.38	0.02	1.53	0.30	0.25	-0.03	1.56	0.27	0.24
Tropical North Atlantic	221	0.06	0.42	0.27	0.32	-0.11	0.44	0.24	0.36	0.02	1.53	0.30	0.25	-0.12	0.57	0.31	-0.06
Eastern Mediterranean	1676	0.02	1.67	0.16	0.59	-0.01	1.69	0.14	0.59	0.02	1.71	0.17	0.51	-0.01	1.73	0.16	0.49
Sahel	1178	-0.07	0.55	0.54	0.17	-0.20	0.59	0.56	0.18	-0.08	0.65	0.57	0.00	-0.19	0.69	0.58	0.02
Subtropical North Atlantic	982	0.10	1.21	0.28	0.51	0.02	1.16	0.21	0.50	0.08	1.31	0.32	0.24	0.01	1.30	0.27	0.20
Central Mediterranean	1918	0.05	1.45	0.21	0.55	-0.01	1.50	0.18	0.54	0.03	1.54	0.24	0.31	-0.01	1.60	0.22	0.30
Middle East	1076	-0.03	0.39	0.35	0.37	-0.08	0.40	0.34	0.43	-0.02	0.46	0.40	0.21	-0.07	0.47	0.39	0.24
Iberian Peninsula	1013	0.01	1.83	0.18	0.19	-0.01	1.85	0.17	0.20	0.01	1.87	0.21	0.00	0.00	1.89	0.19	-0.01
Western Iberian Peninsula	653	0.00	1.79	0.13	0.23	-0.02	1.80	0.12	0.21	0.01	1.79	0.16	0.05	-0.01	1.82	0.14	0.02
North Western Maghreb	714	0.06	0.52	0.25	0.55	-0.06	0.55	0.25	0.49	0.04	0.61	0.28	0.44	-0.06	0.66	0.29	0.33
Sahara	328	0.03	0.48	0.59	0.27	-0.11	0.40	0.58	0.33	0.04	0.52	0.62	0.14	-0.07	0.46	0.61	0.17
Eastern Sahara	232	0.08	0.97	0.25	0.44	0.02	0.91	0.20	0.43	0.07	1.00	0.25	0.40	0.02	0.94	0.22	0.37

Finally, the comparison of 48h and 72h forecasts for both CAMS experiments shows that meanwhile the MB, RMSE and FGE is stable during the 3-days forecasts, the correlation is reduced drastically from 24h to 72h in all the regions (see Table 3.5.2 and Table 3.5.3). It is worthy to highlight the slightly lower correlation values of CAMS o-suite respect CAMS control for 48h and 72h forecast (see Table 3.5.3)

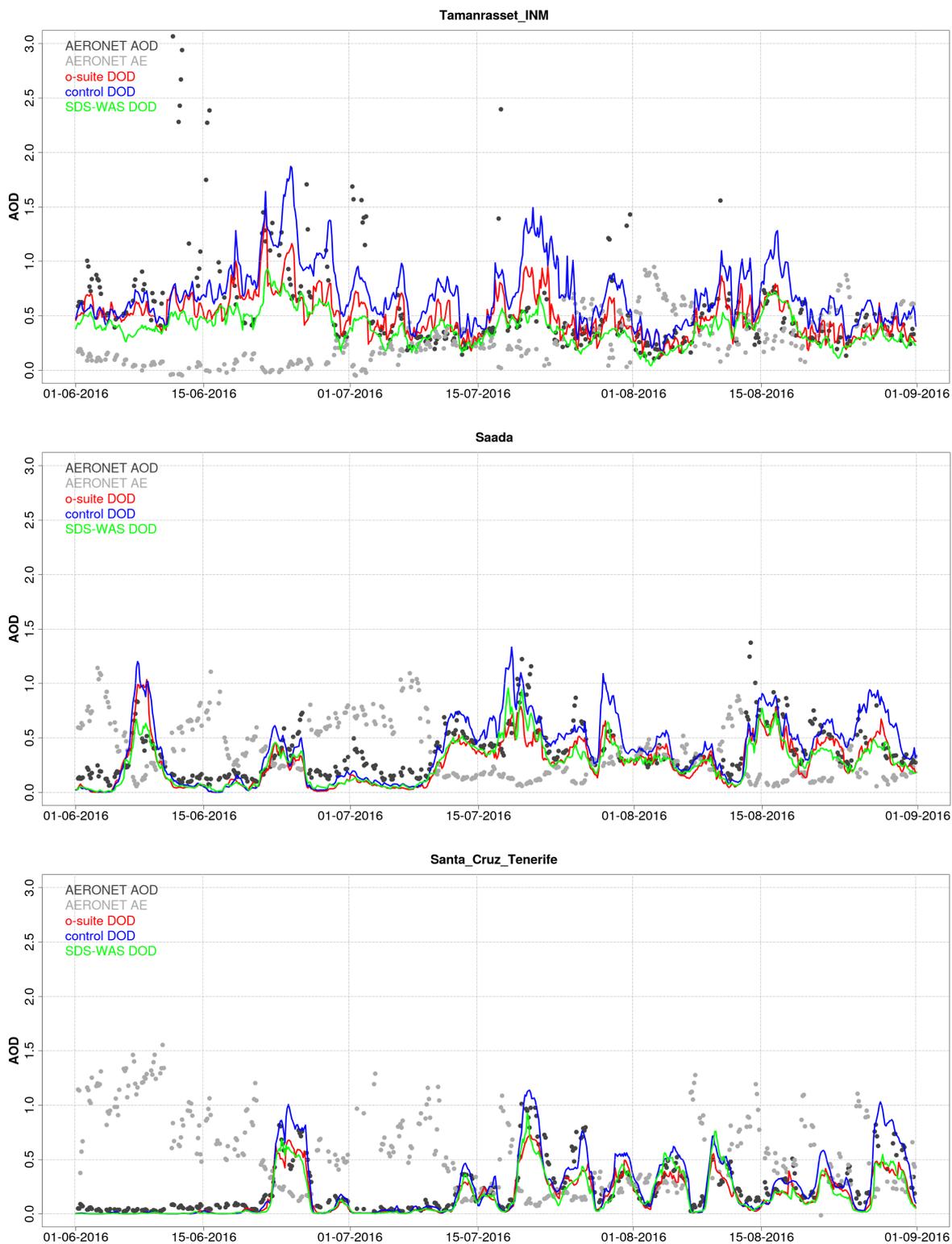


Figure 3.5.4: AOD from AERONET (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multimodel SDS-WAS Median (green line) for the period period June 1st to August 31st, 2016 over Tamanrasset (Sahara), Saada (NW Maggreb) and Santa Cruz de Tenerife (Subtropical North Atlantic).

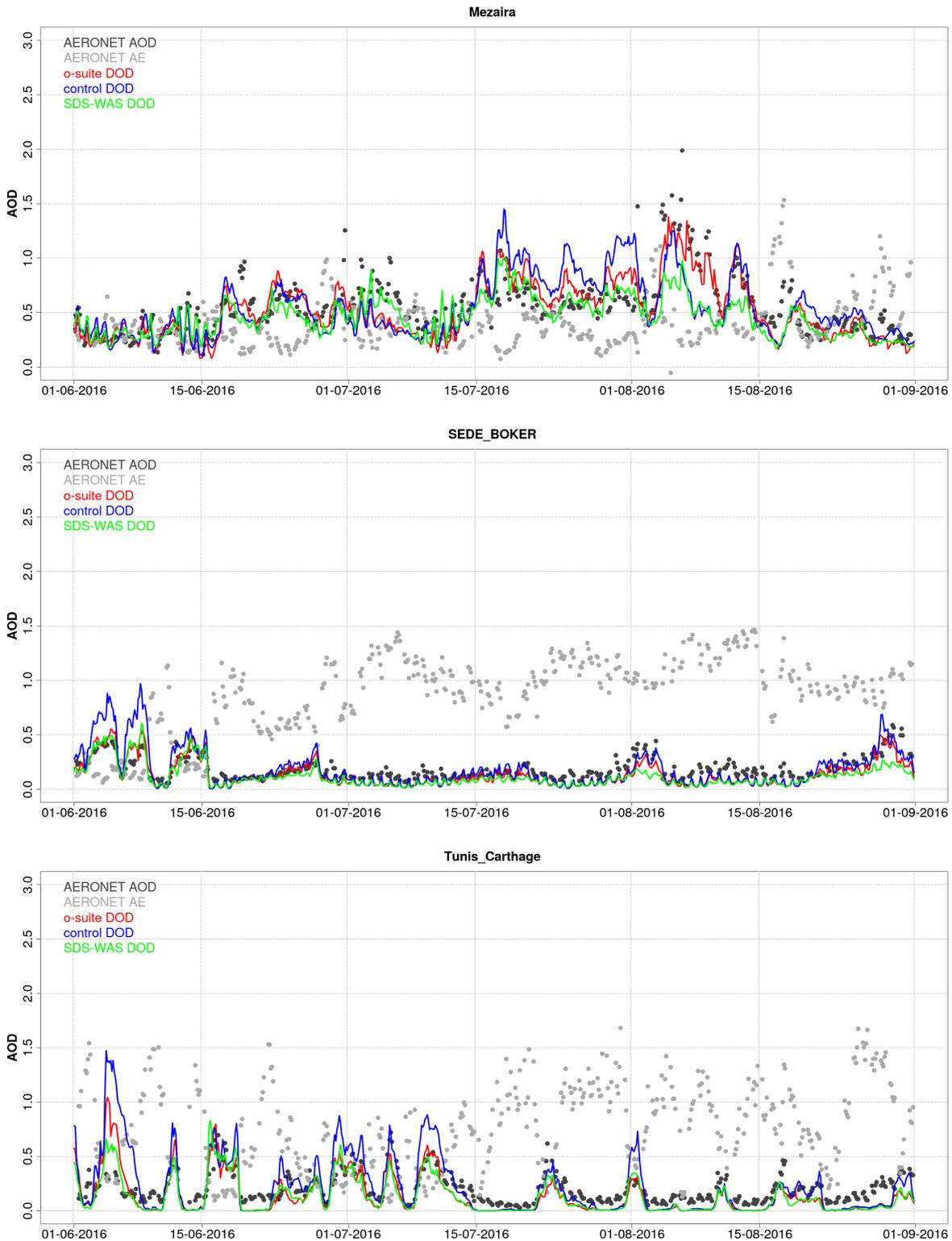


Figure 3.5.5: AOD from AERONET (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multimodel SDS-WAS Median (green line) for the period June 1st to August 31st, 2016 over Mezaira (Middle East), Sede Boker (E. Mediterranean) and Tunis-Carthage (Central Mediterranean).



3.5.3 Aerosol validation over the Mediterranean

Daily aerosol optical depth (AOD) and surface concentration (PM₁₀ and PM_{2.5}) from CAMS o-suite experiment (Morcrette et al., 2009; Benedetti et al., 2009) and CAMS control experiment have been validated against 37 AERONET and 18 Airbase stations in the Mediterranean region for the period 1 June – 31 August 2016. The main goal is to know the behaviour of AOD and surface concentration (PM_{2.5} and PM₁₀) from CAMS o-suite and control over the Mediterranean. In this report, we validate CAMS o-suite and control experiments using AERONET and Airbase observations as the reference.

Aerosol optical depth

3-hourly values of AOD from AERONET, o-suite and control for the period 1 June – 31 August 2016 over selected sites are shown in Figure 3.5.6 and Figure 3.5.7. From June to September 2016, CAMS o-suite is the model that best reproduces the daily variability of AERONET observations (see correlation values in Figure 3.5.6). In average for all the sites, MB increases from 0.02 for control to 0.03 for o-suite; and correlation increases from 0.60 for control to 0.64 for o-suite. The highest peaks on CAMS AOD simulations are linked to natural sources (see Figure 3.5.7). During summer, maximum dust activity is observed in Central and Western Mediterranean (see Figure 3.5.7). Poor scores are observed in Northwestern Mediterranean where correlations are under 0.4 and o-suite overestimates the observed AERONET values in comparison with CAMS control (see Figure 3.5.6 and Toulon in Figure 3.5.7). During summer, the photochemical processes are enhanced favouring the formation of secondary aerosols which may not be represented well in the model.

Surface aerosol concentrations

3-hourly values of PM₁₀/PM_{2.5} from Airbase, o-suite and control for the period 1 June – 31 August 2016 over selected sites are shown in Figure 3.5.8. In general, CAMS o-suite presents better results regarding load concentrations reducing the observed overestimations both PM₁₀ and PM_{2.5} (see Figure 3.5.8). In average for all the sites, PM₁₀ MB decreases from 0.8 µg/m³ for control to -3.3 µg/m³ for o-suite, while PM_{2.5} MB decreases 7.2 µg/m³ for control to 2.7 µg/m³ for o-suite. From June to September 2016, CAMS experiments reproduce the daily variability of the most intense aerosol events observed by Airbase sites (see Hospital Joan March in Figure 3.5.9) although both CAMS experiments tend to overestimate the observed values. During summer, desert dust intrusions over Western Mediterranean are frequent as it is observed in the AOD analysis. Although, the dust transport is localized at high altitudes and the contribution to surface levels is limited. Like previous seasons, CAMS model reproduces extreme peaks particularly over maritime sites (see Hospital Joan March on June 28 in Figure 3.5.9) that are not observed in the Airbase stations. These peaks are reduced in CAMS o-suite providing better results than control.

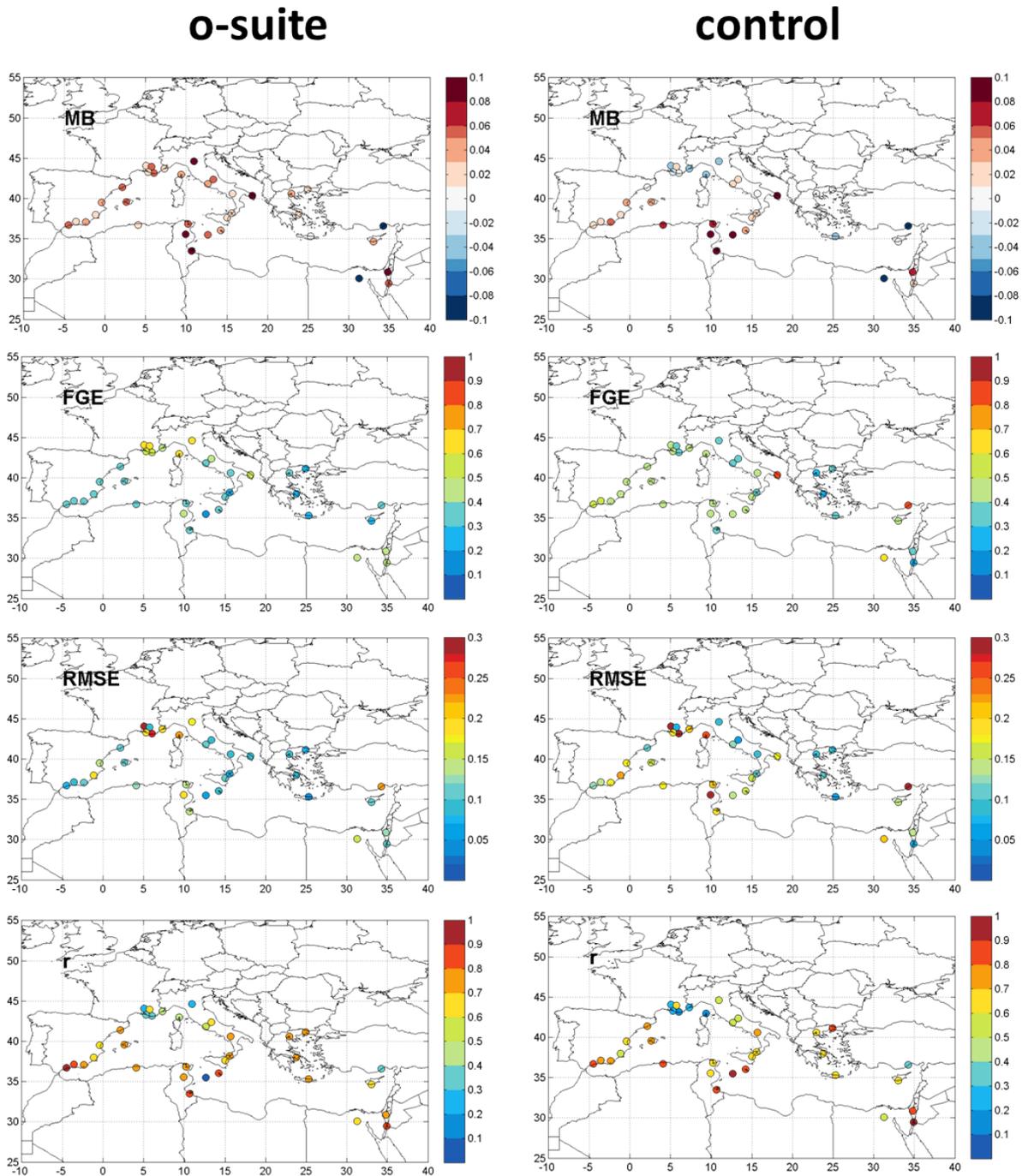


Figure 3.5.6: Skill scores (MB, FGE, RMSE and r) for 24-hour forecasts of CAMS o-suite and control for the study period from June 1st to August 31st, 2016. AOD from AERONET is the reference.

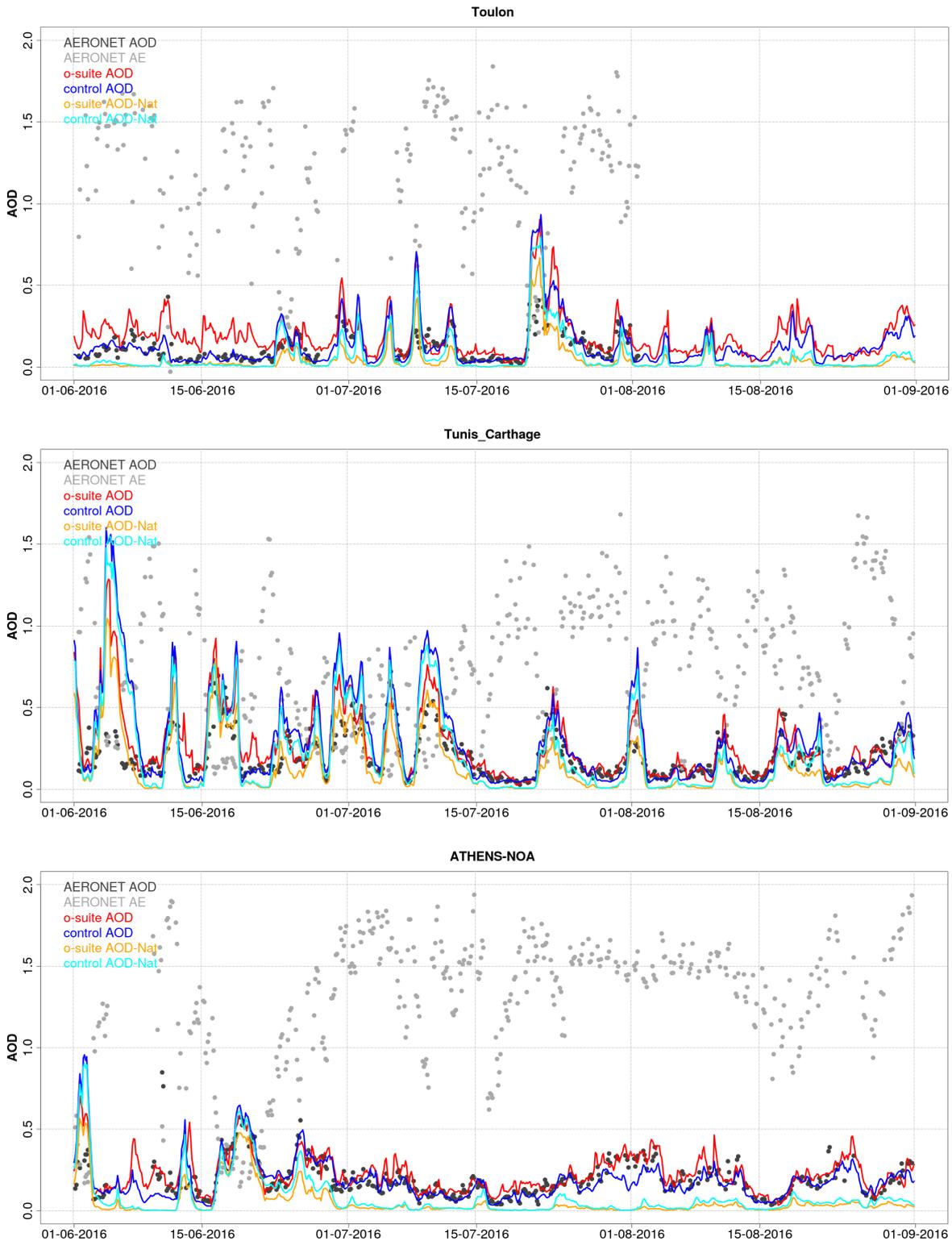


Figure 3.5.7: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), for the period June 1st to August 31st, 2016 over Toulon (France), Tunis-Carthage (Tunisia) and Athens (Greece). AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt.

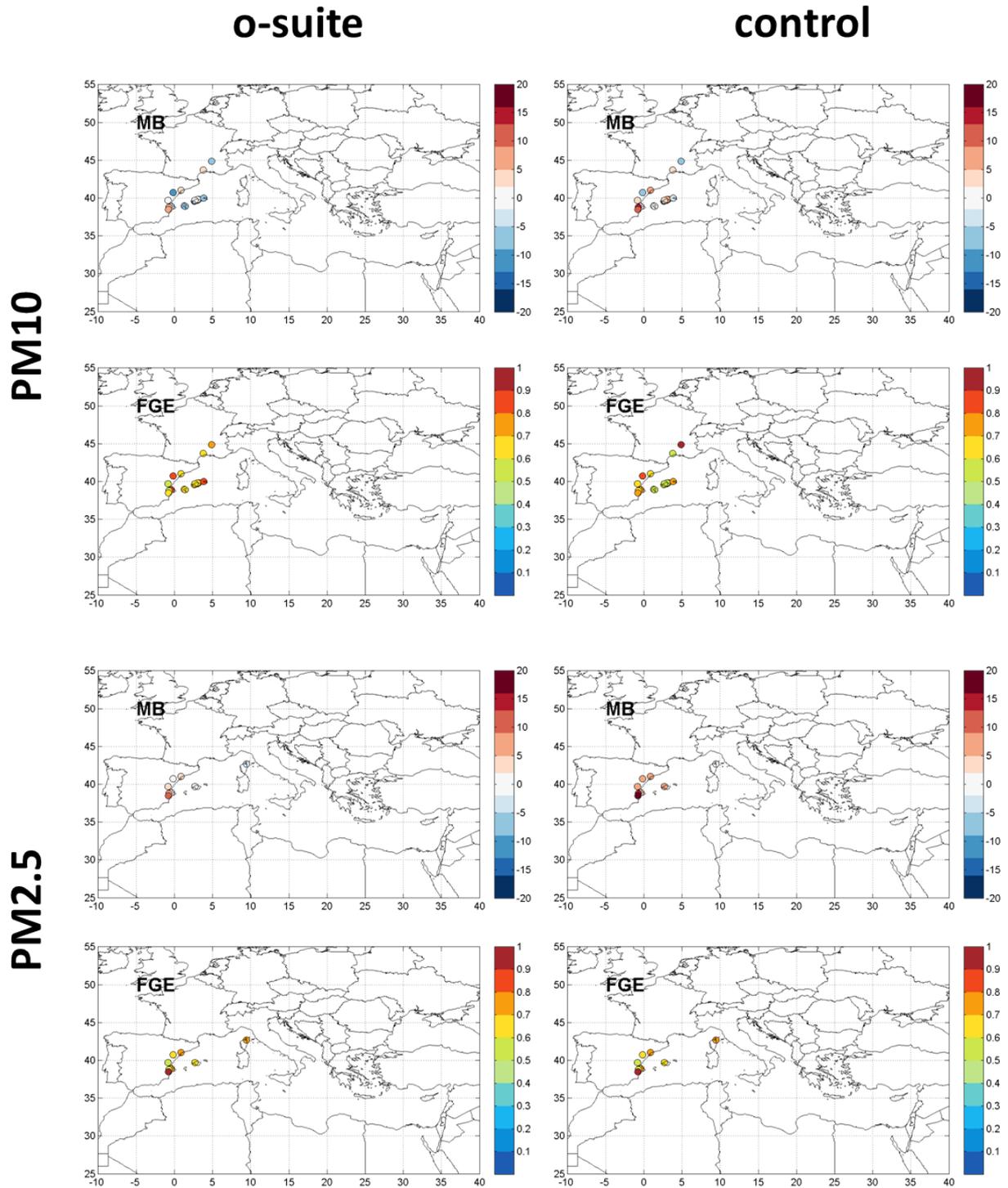


Figure 3.5.8: Skill scores (MB and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. PM10 and PM2.5 from Airbase are the reference. Only background suburban and rural available stations are displayed.

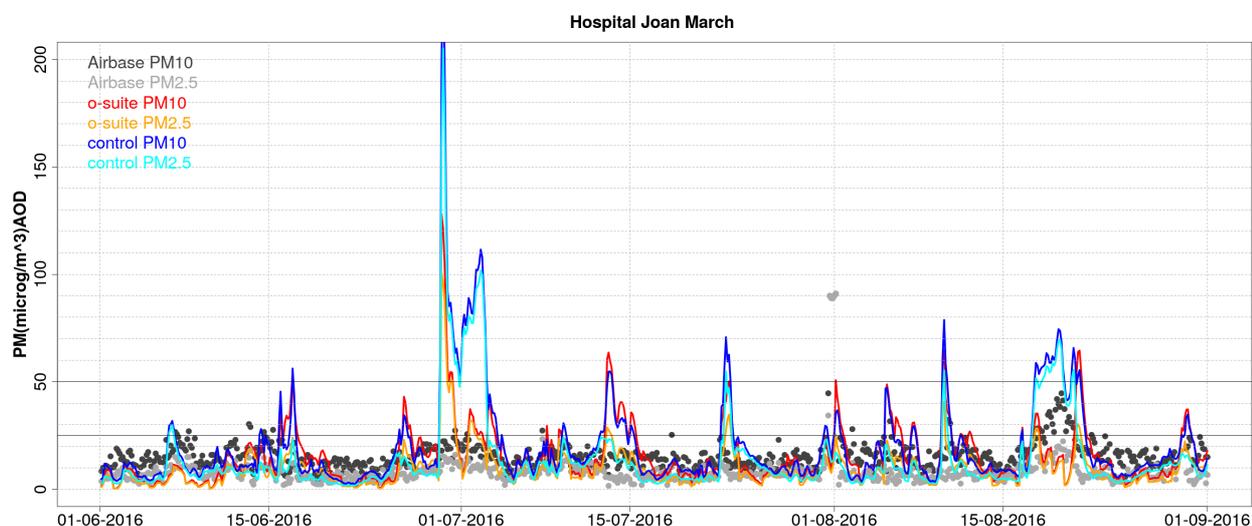


Figure 3.5.9: PM10 and PM2.5 Airbase observations (black and grey dots, respectively), PM10 and PM2.5 o-suite (red and orange lines, respectively) and PM10 and PM2.5 control (blue and cyan lines, respectively) for the period June 1st to August 31st, 2016 over Hospital Joan March (39.68°N; 2.69°E, Spain).

3.5.4 Backscatter profiles

The technical specifications of the data sources, evaluation parameters and methods are described in the report CAMS-84 D8.1. In this section, the vertical variation of the backscatter coefficient (bsc) profiles, i.e. correlation and standard deviation, of o-suite '0001' and control run 'gjjh' (ctrl) vs ceilometers are evaluated and summarized in Taylor plots. The vertically integrated bsc bias is not the focus, because it corresponds in its information content to AOD. Likewise, a skill measure for the horizontal extension/positioning of plumes/layers is not a primary goal of this evaluation. We focus to representative case studies, because this shows the typical behavior of the model most clearly. A statistical summary is given in Taylor plots. Issues with sea salt and sulfate were discussed in former reports and will be revised after significant model upgrades which are announced for autumn 2016 and spring 2017.

Period Overview

The model aerosol optical depth (AOD) is used to select periods with significant aerosol plumes predicted over Germany. Figure 3.5.14 shows this for Soltau (53N, 10E), central Germany, separately for contributions of mineral dust (SD), sea salt (SS), biomass burning (BB), black (BC) and organic carbon (OC), as well as sulfate (SU). Saharan dust events (SDE) around 30/05-05/06/16, 21-23/07/2016 and 26-28/08/16 are confirmed by ceilometers. A mixed SD/BB plume covered Europe in late July. Elevated contributions of sulfate in early summer 2016 are associated with Canadian forest fires. Sulfate and OC dominate AOD due to their large specific extinctions, but their mass mixing ratios are comparable to the other species. SO₄ and OC show annual cycles with summer maxima.

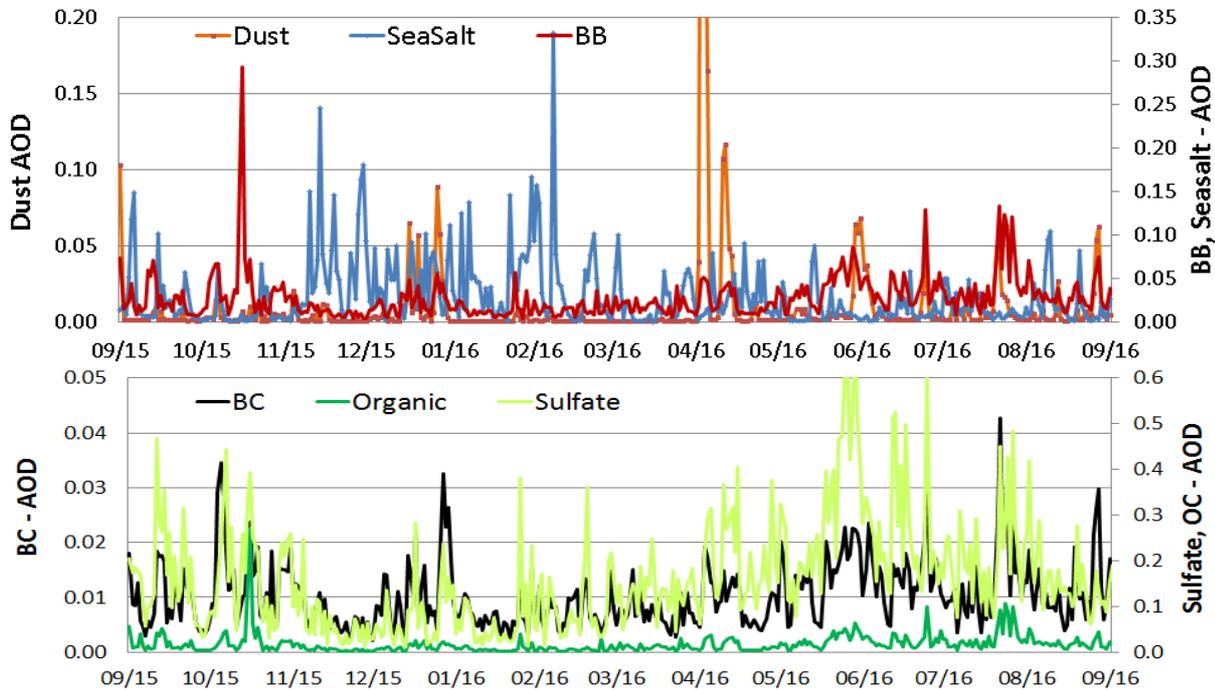


Figure 3.5.14: Maximum daily AOD around Soltau ($\pm 1^\circ$ lat/lon) for aerosols included in the IFS model from 09/15-09/16: sea salt (blue), dust (orange), biomass burning (=OC+BC - red), BC (black), organic (green), and sulfate (light green). Note the different vertical axis for the different aerosol species.

Saharan Dust, 01-07 June 2016:

A SD plume reached Germany via eastern Europe on 1 June and thinned out till 7 June (Fig. 3.5.15). Its top was around 4 km and appeared intermittent in the ceilometer 2-D plots, due to movements and dispersion over Europe.

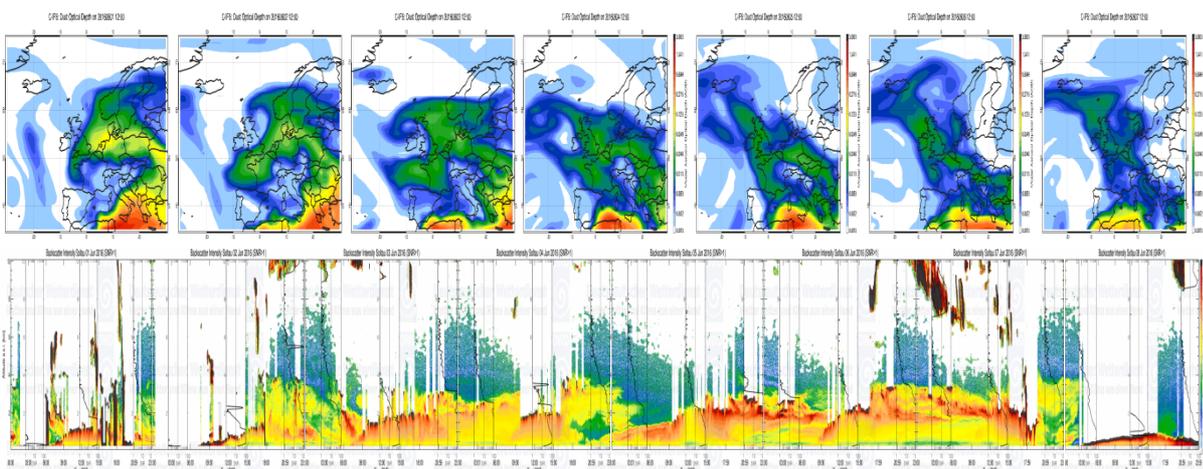


Figure 3.5.15: Upper panel: Dust AOD from g9rr on 1-8 June 2016. Lower panel: Time-height sections of backscatter signal (Pr^2) over Soltau on 1-7 June 2016, showing the passage of the SD plume.

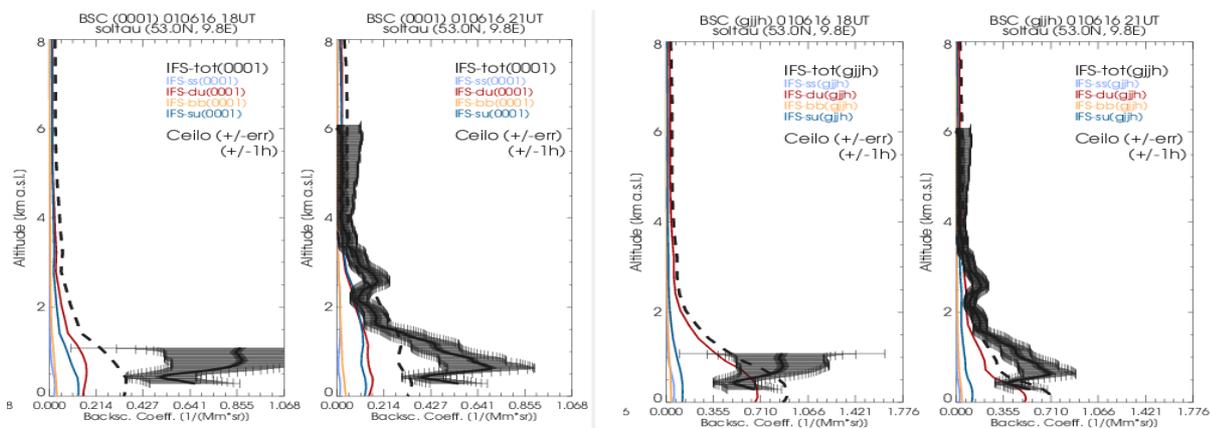


Figure 3.5.16: Profiles of backscatter coefficients at Soltau on 1 June 2016 (left o-suite, right control), split up for contributions from sea-salt (light blue), dust (red), biomass burning (orange) and sulfate (dark blue), as well as the total aerosol (dashed black). Solid black lines: Ceilometer total bsc with estimated error bars, plotted for -1h, +/-0h, +1h around the time of the 3-hly model profile.

The passage of both dust plumes is captured by the model. CONTROL RUN (gjjh): Peak height in 6 km on 1 June with max bsc of $1.5 \text{ Mm}^{-1}\text{sr}^{-1}$ at 06UT, decreasing to $0.7 \text{ Mm}^{-1}\text{sr}^{-1}$ at 18 UT, where good agreement to observations is found. Plume settling and bsc decreasing to $<0.4 \text{ Mm}^{-1}\text{sr}^{-1}$ on 2 June, at 21 UT only 20% of observed bsc. Next model plume on 3 June 6UT-15 UT with max of $0.3 \text{ Mm}^{-1}\text{sr}^{-1}$ is roughly as observed. Thereafter dust bsc is strongly underestimated by the model and the total mass is dominated by sea salt. OSUITE (0001): On 1-3 June similar shape of vertical profiles as gjjh, but bsc much lower than control run (20-40% of gjjh) and observations. After 4 June even stronger underestimation of bsc than gjjh and again dominance of sea salt.

Saharan Dust, 26-29 August 2016:

A SD plume crossed Germany from NW to SE from 26-29 Aug (Fig. 3.5.18). It reached up to at least 6 km and was split into layers near the surface and around 2 km. In the upper layer, clouds formed due to condensation at the dust particles. The plume position over N-Germany was correct in the model for 27 Aug (cf. Figs 3.5.16 and 3.5.17, upper panel). The double-layer structure was captured as well, though the rejoin was a bit too early in the model (Fig. 3.5.19 top panel). On 28 Aug midday a dust layer near 2km is reproduced, but too weak, by the o-suite, but not by the control run. The Taylor plots (Fig 3.5.19) reflect the too low model bsc values/variance with a bullet cloud near the axis origin.

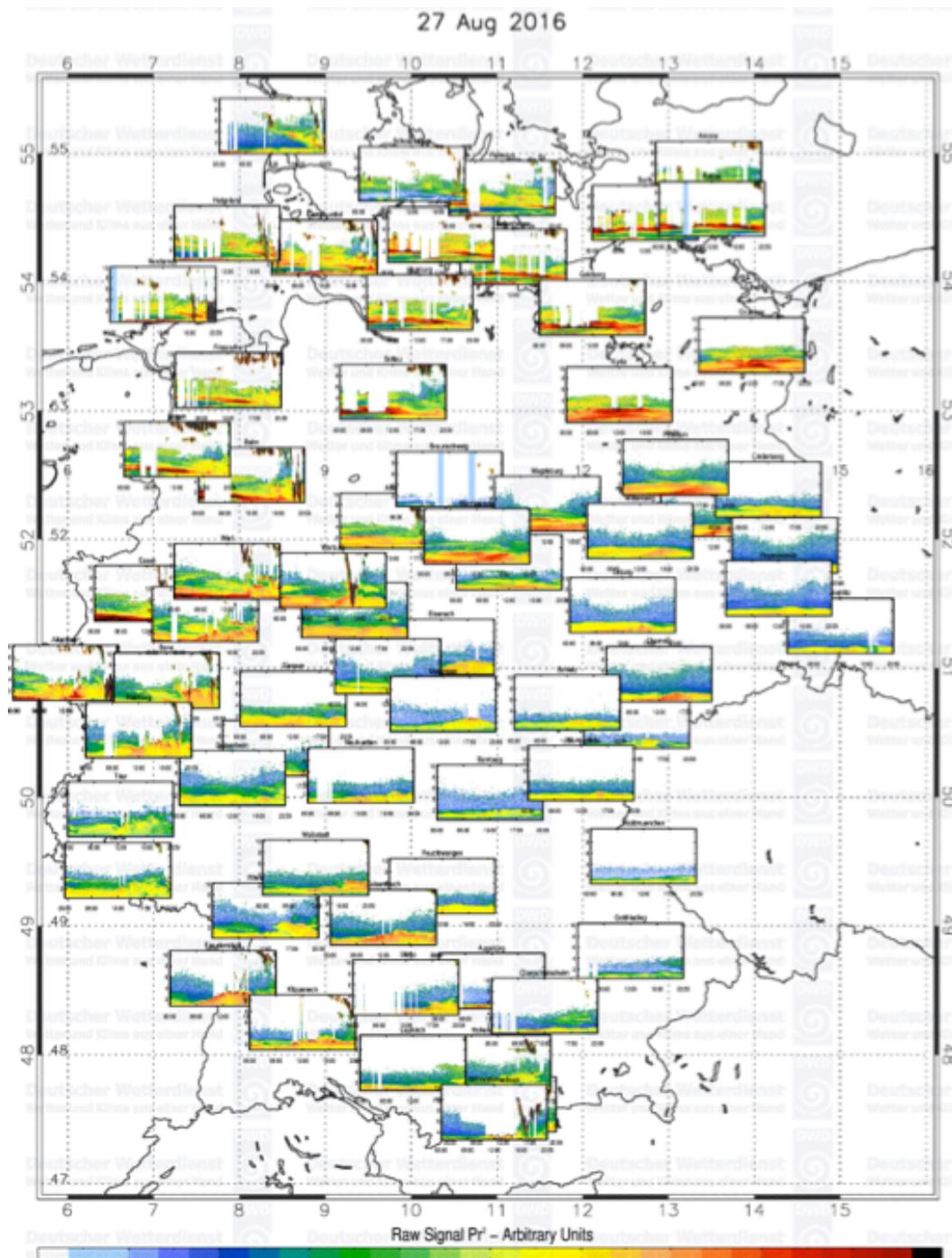


Fig. 3.5.17: German ceilometer network: 2-D time-height sections of Pr^2 for 27 August 2016.

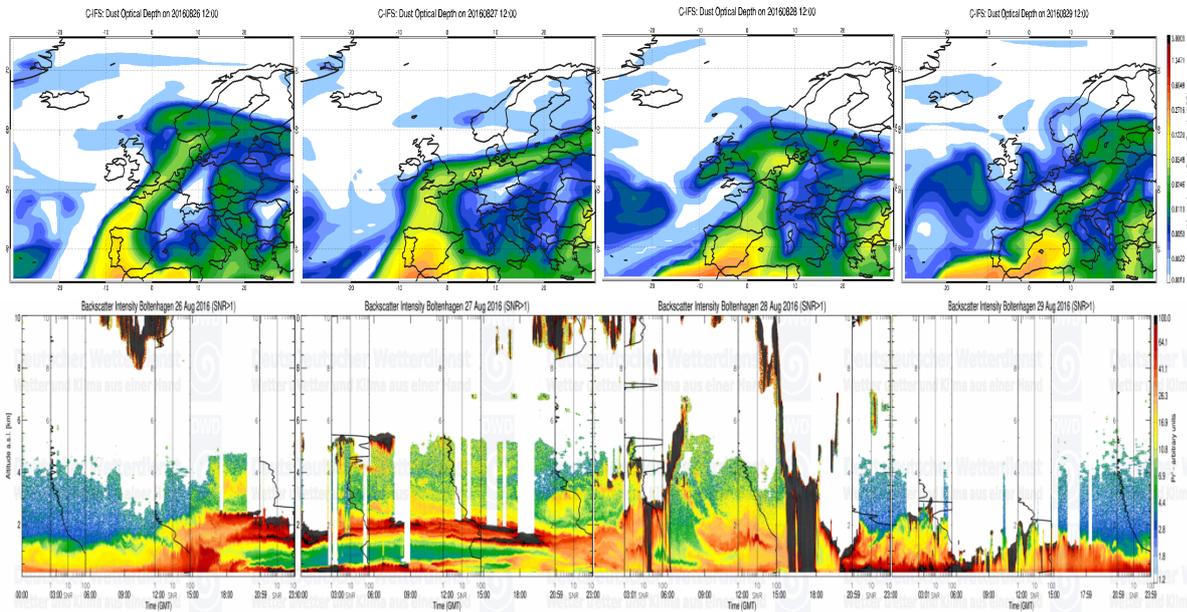


Figure 3.5.18: Upper panel: Dust AOD from g9rr on 26-29 August 2016. Lower panel: Time-height sections of backscatter signal (Pr^2) over Boltenhagen on 26-29 August 2016, showing the passage of the SD plume.

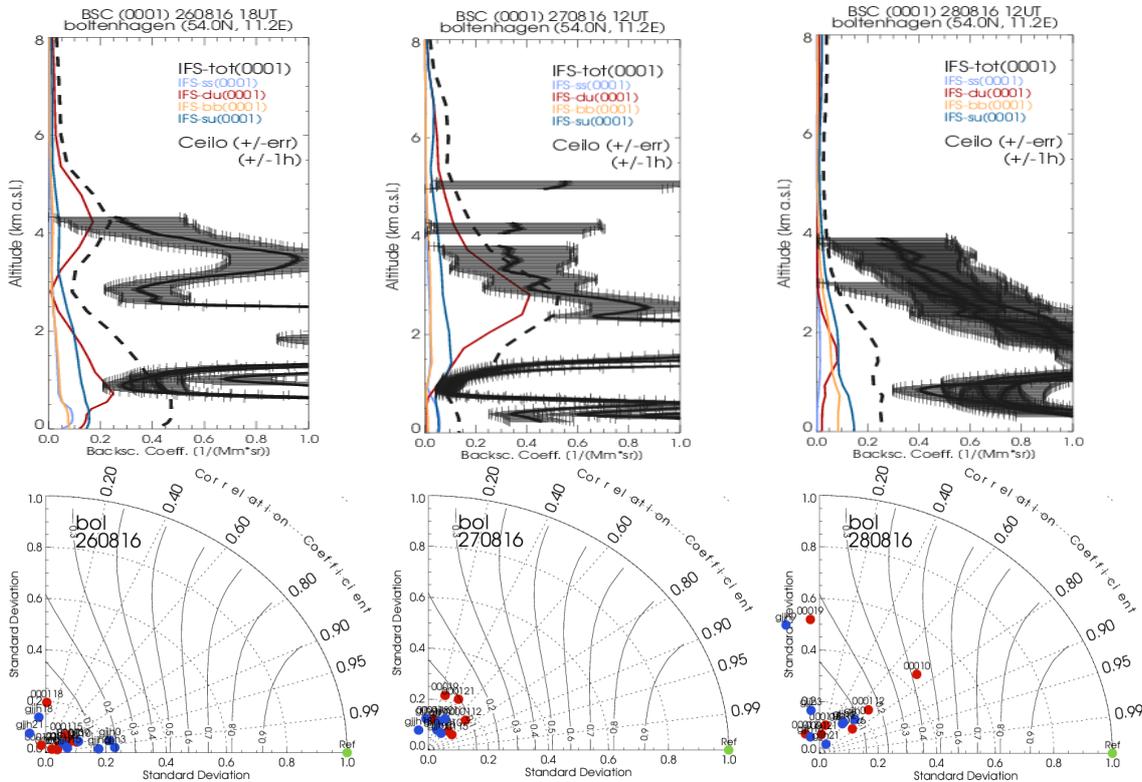


Figure 3.5.19: Top panel: Profiles of backscatter coefficients at Boltenhagen (54°N) on 26-28 Aug 2016 (o-suite), as in Fig 3.5.16. Lower panel: Taylor polar plots with standard deviation vs correlation coefficient of the 3-hourly profiles for Boltenhagen on 26, 27 and 28 August 2016. O-suite red, control blue.

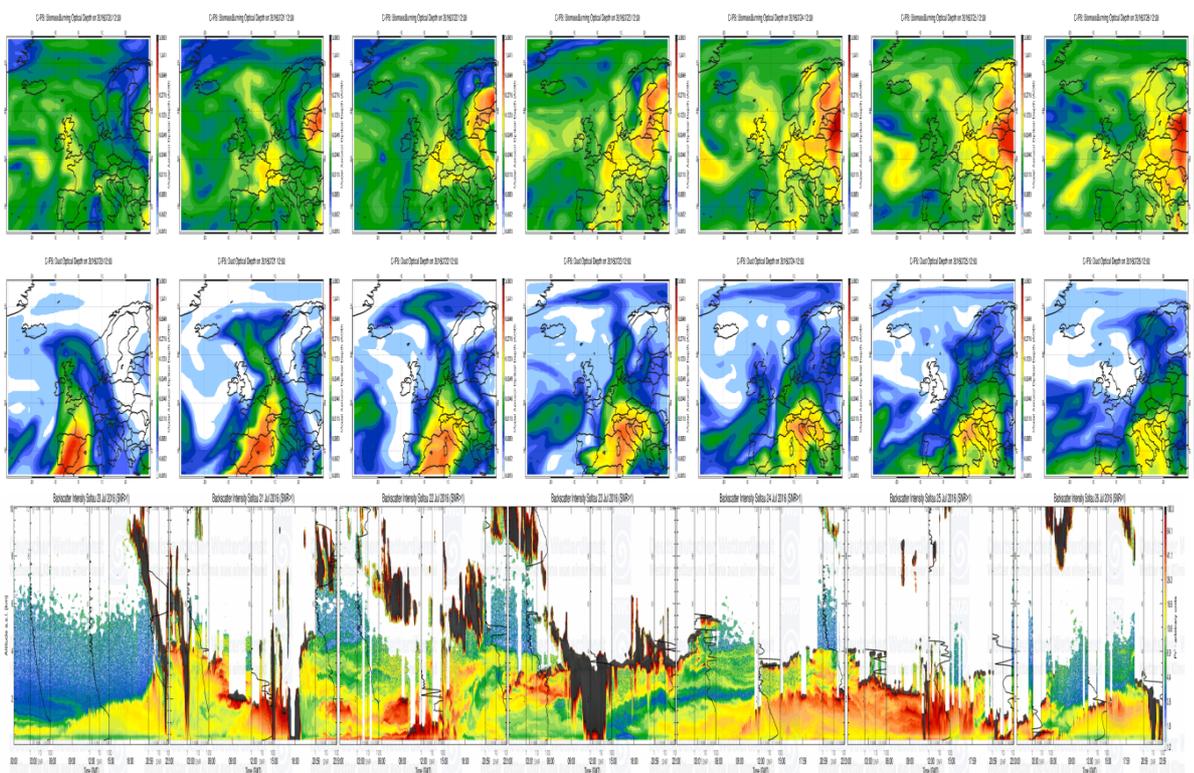


Figure 3.5.20: Upper panel: BB (top) and SD AOD from g9rr on 20-26 July 2016. Lower panel: Time-height sections of backscatter signal (Pr^2) over Soltau on 20-26 July 2016, showing the passage of the mixed SD-BB plume plume.

Mixed Saharan Dust and Biomass Burning, 20-26 July 2016:

The model indicates transport of SD, BB and SS aerosol to Europe in the period 20-26 July 2016. The BB aerosol is more concentrated in the tip of the SD air-mass. While higher model dust loads reach only to south Germany, the model BB aerosol covers the whole country. The partitioning between SD and BB mass cannot be verified by the ceilometer data. The arrival of the dust plume is 12 h late in the model, but during the following period the model bsc agree well with the observations as shown by the profiles and the Taylor plots in Fig. 8.

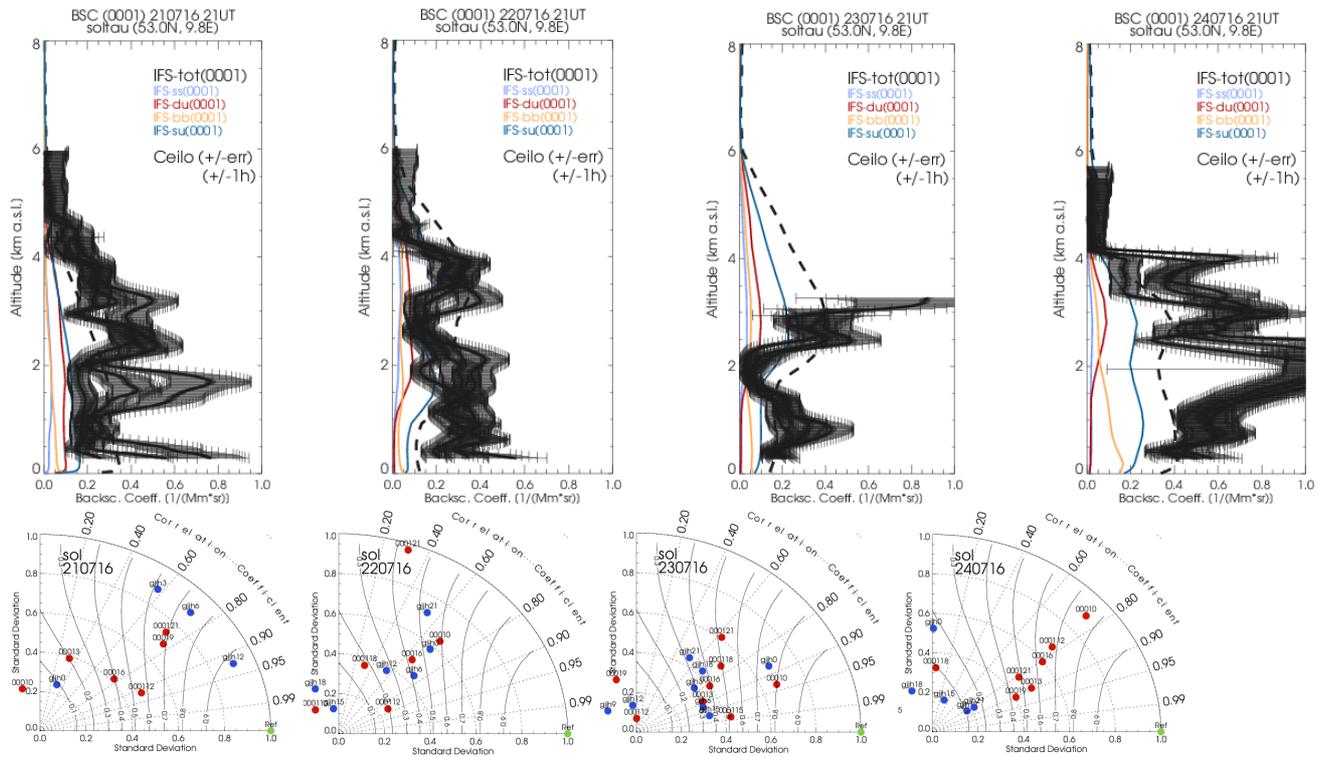


Figure 3.5.21: Profiles of backscatter coefficients at Soltau on 21-24 July 2016, each 21UT (o-suite), as in Fig 3. Lower panel: Taylor polar plots with standard deviation vs correlation coefficient of the 3-hourly profiles for Soltau (54°N) on 21-24 July 2016. O-suite red, control blue.

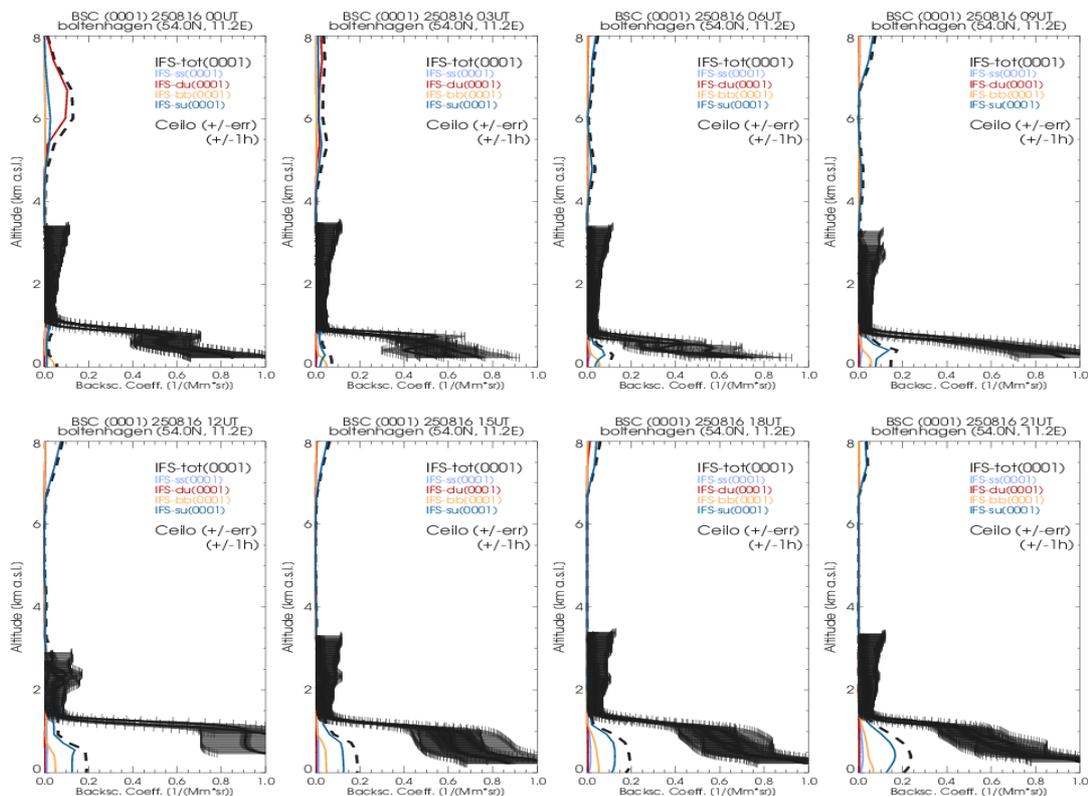


Fig. 3.5.22: Profiles of backscatter coefficients at Boltenhagen on 25 Aug 2016 (o-suite), as in Fig 3.5.16.

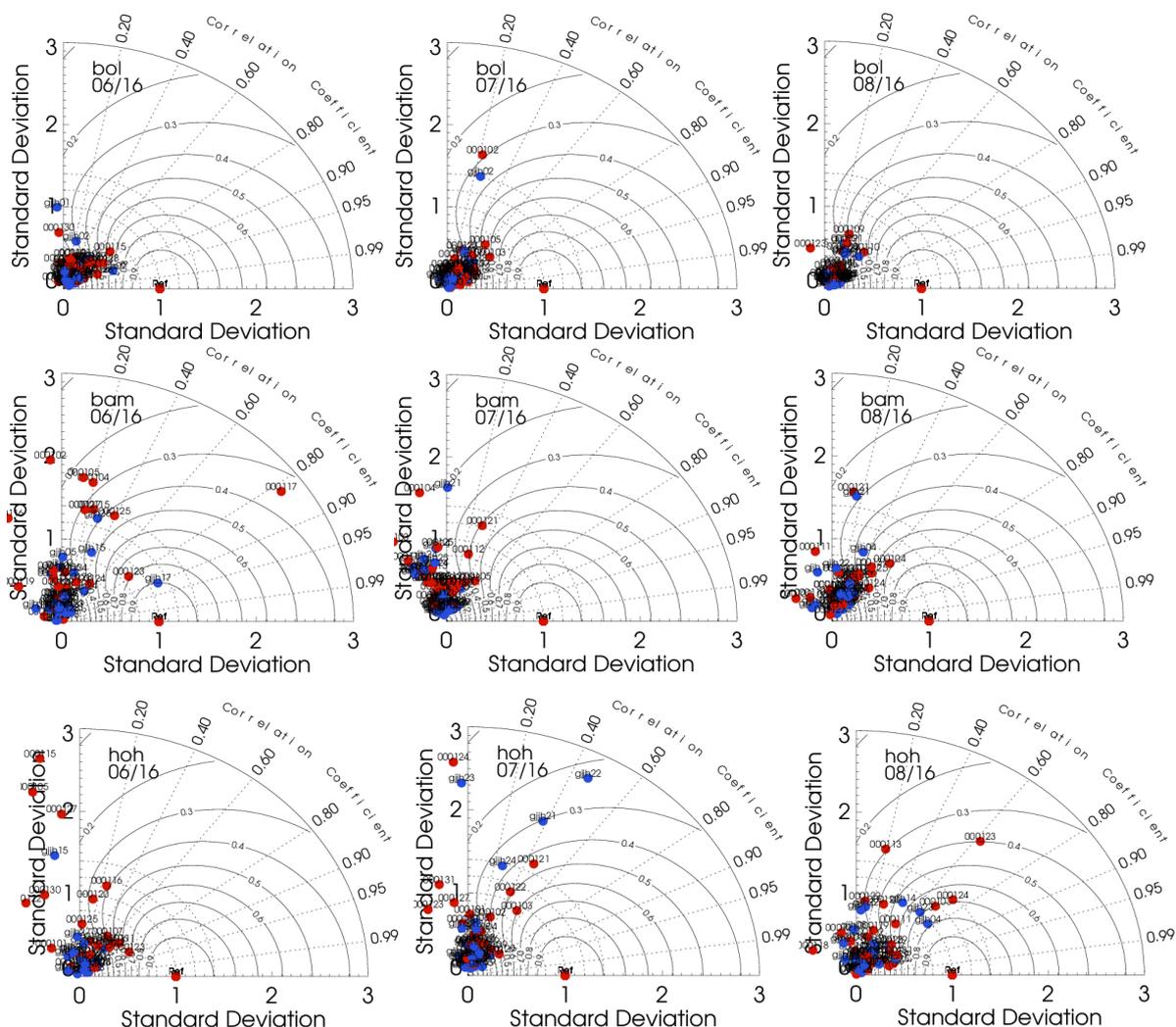


Fig. 3.5.23: Taylor polar plots with daily average standard deviation vs correlation coefficient for three German sites Boltenhagen (54°N), Bamberg (50°N) and Hohenpeissenberg (48°N) for June, July August 2016. O-suite red, control blue.

Planetary boundary Layer

The height of the PBL is generally reproduced within a vertical range of few 100m, but mass mixing ratios are generally too small. A typical example from an undisturbed summer day without PBL top clouds is shown in Figure 3.5.22.

Statistic Summary:

The reproduction of detailed aerosol profiles is a serious challenge, as for all global models. Strong gradients and fast changes are smoothed, structures often spatially and temporally shifted. The variability (standard deviation) of the model is much smaller than observed, reflecting in the bullet clouds near the axis origin in the Taylor plots (Fig. 3.5.23). This is also true during particle events like Saharan dust or fire plumes. Strong particle layers are generally reproduced, but mostly with spatiotemporal shifts. The planetary boundary layer height is mostly reproduced within few 100m vertical range but aerosol loads in the PBL is generally underestimated.

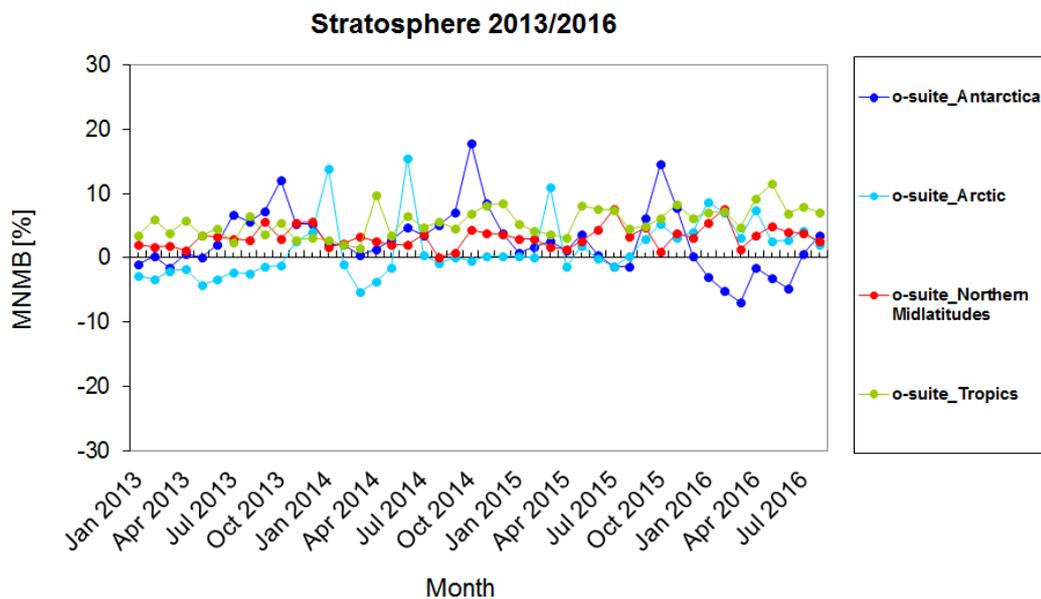


Figure 3.6.1: MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated sonde data in the Arctic (light blue), Antarctic (dark blue) northern midlatitudes (red) and tropics (green).

3.6 Stratospheric ozone

3.6.1 Validation against ozone sondes

In what follows, we present the results of the stratospheric ozone evaluation against ozone soundings from the NDACC, WOUDC, NILU and SHADOZ databases. The sondes have a precision of 3-5% (~10% in the troposphere for Brewer Mast) and an uncertainty of 5-10%. For further details see Cammas et al. (2009), Deshler et al. (2008) and Smit et al (2007). Model profiles of the o-suite are compared to balloon sondes measurement data of 44 stations for the period January 2013 to August 2016 (please note that towards the end of the validation period fewer soundings are available). As C-IFS-CB05 stratospheric composition products beyond O_3 in the o-suite is not useful we provide only a very limited evaluation of the control experiment. A description of the applied methodologies and a map with the sounding stations can be found in Eskes et al. (2016). Both runs, the o-suite and the control run, show MNMBs mostly within the range -7 to +10%, for all regions and months (some exceptions with MNMBs of up to 15% for single months in the high latitude regions), see Fig. 3.6.1-3.6.2.

O_3 partial pressures in the stratosphere are mostly slightly overestimated (MNMBs between 2-8%) in all latitude bands, except for the Arctic (before 2015) and Antarctic summer season. MNMBs in Antarctica during the ozone hole season, from August to November, remain below $\pm 15\%$ for the o-suite.

Fig. 3.6.2 compares the averaged profiles in each region during August 2016. The vertical distribution of stratospheric ozone is quite well represented for all regions by the o-suite, with little overestimation above 20hPa.

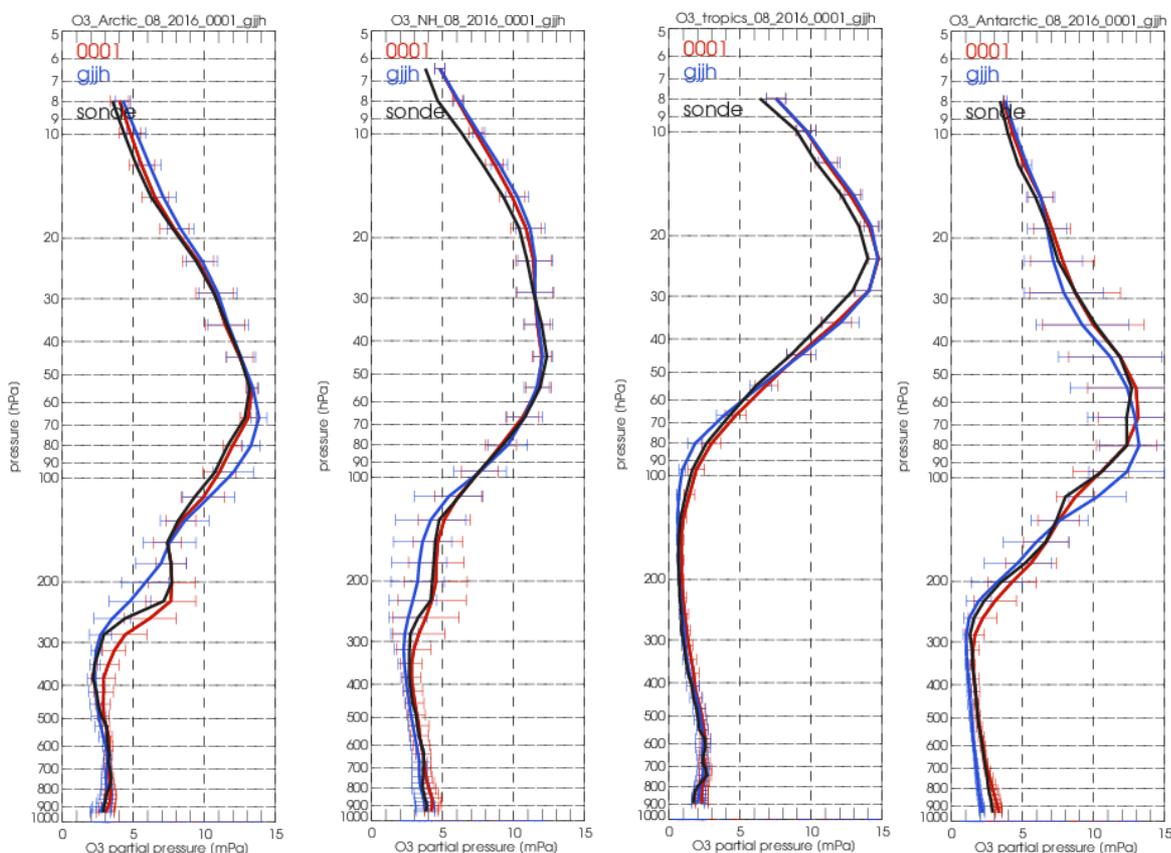


Figure 3.6.2: Comparison between mean O₃ profiles (units: mPa) of o-suite (red), and control (blue) in comparison with observed O₃ sonde profiles (black) for August 2016 for the various latitude bands: Arctic, NH-mid latitudes, Tropics and Antarctic.

3.6.2 Validation against observations from the NDACC network (MWR, LIDAR)

In this section we present a comparison between the CAMS o-suite and control run models against MWR and LIDAR observations from the NDACC network. A detailed description of the instruments and applied methodologies for all NDACC instruments can be found at <http://nors.aeronomie.be>. MWR (microwave) at Ny Alesund (79°N, 12°E, Arctic station) and Bern (47°N, 7°E, northern midlatitude station). LIDAR at Lauder, New Zealand (46°S, 169.7°E, altitude 370m) and Hohenpeissenberg, Germany (47°N, 11°E, altitude 1km).

From Table 3.6.1, the upper stratospheric partial column bias at Bern during Sept. 2015 –August 2016 is nearly vanishing (uncertainty on the partial column is 6%). At Ny Alesund, the o-suite overestimates the stratospheric ozone concentration with more than 10%, and this vanished during JJA. In MAM-JJA, both MWR stations observe a significant (i.e. comparable to the measurement uncertainty) overestimation of the upper stratosphere/mesosphere ozone content, and this evolved to an underestimation in SON-DJF, reaching values up to -30% (Ny Alesund), see also Fig. 3.6.3. At BERN the difference between osuite and MWR at 25-35km is negligible since Sept 2015 (compared to the MWR profile uncertainty).

At Lauder and Hohenpeissenberg (LIDAR), the o-suite slightly overestimates the observed ozone (<10%) between 25km and 35km. The uncertainty on the LIDAR concentration increases with



Table 3.6.1: Seasonal relative mean bias (MB, %), standard deviation (STD, %) of the partial (upper stratospheric 25km – 65km) ozone column for the considered period and number of observations used (NOBS), compared to NDACC microwave observations at Ny Alesund and Bern (mean bias and stddev in %).

		SON			DJF			MAM			JJA		
		MB	stddev	nobs	MB	stddev	nobs	MB	stddev	nobs	MB	stddev	nobs
o-suite	Ny.Ale	12.84	7.39	273	14.50	6.51	213	6.35	5.42	229	-0.70	5.74	132
	Bern	-0.54	2.32	687	0.74	3.43	527	0.97	2.82	609	0.09	2.78	571

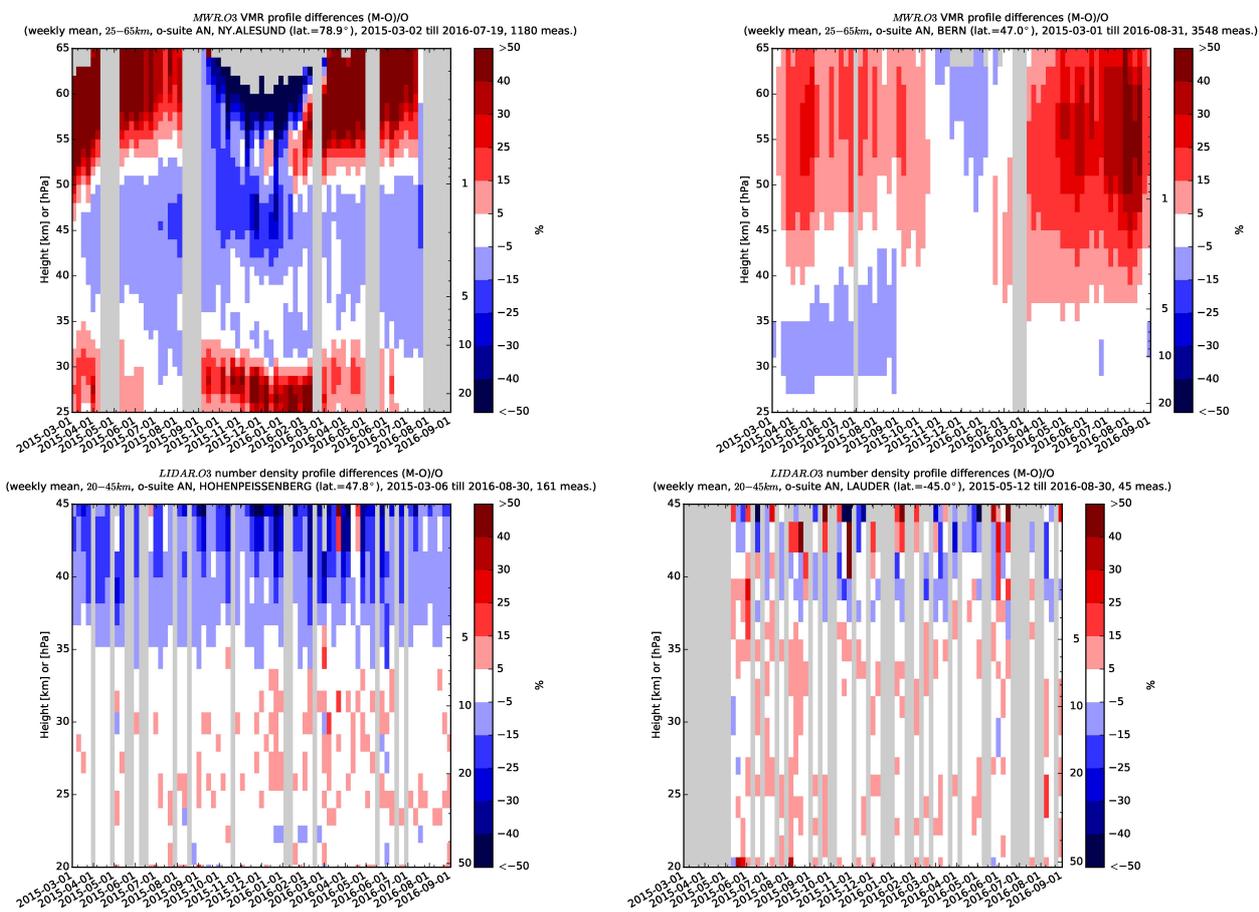


Figure 3.6.3. Comparison of the weekly mean profile bias between the O₃ mixing ratios of o-suite and the NDACC station at Ny Alesund, Bern, Hohenpeissenberg and Lauder. For the LIDAR stations, the measurement uncertainty above 35km is comparable to the observed profile bias.

altitude and above 35km the observed differences are comparable to the measurement uncertainty (>10%, see http://nors.aeronomie.be/projectdir/PDF/NORS_D4.2_DUG.pdf)

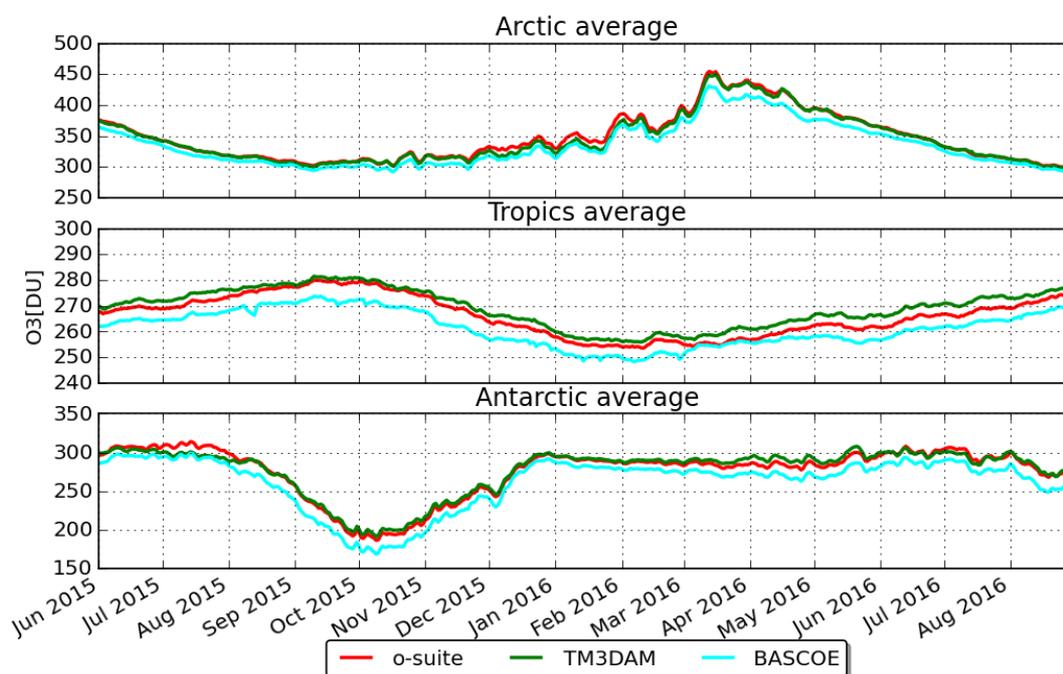


Figure 3.6.4: Zonally averaged ozone total column (Dobson Units) in the Arctic (60°N-90°N), Tropics (30°S-30°N) and Antarctic (90°S-60°S) during the period 2015/06/01-2016/09/01.

3.6.3 Comparison with dedicated systems and with observations by limb-scanning satellites

This section compares the output of the o-suite for the last period, based on the methodology described by Lefever et al. (2015). It also compares the model output with observations by two limb-scanning satellite instruments: Aura-MLS and OMPS-LP. The comparisons with Aura-MLS are only a verification since that dataset is assimilated in both the o-suite and BASCOE. The combination of these comparisons delivers a good picture of the performance of the CAMS o-suite analyses w.r.t. stratospheric ozone. We also include the comparisons for the o-suite forecasts of stratospheric ozone. These forecasts have a lead time of 4 to 5 days and are represented by red dotted lines in the figures.

All datasets are averaged over all longitudes and over the three most interesting latitude bands for stratospheric ozone: Antarctic (90°S-60°S), Tropics (30°S-30°N) and Arctic (60°N-90°N). In order to provide global coverage, the two mid-latitude bands (60°S-90°S and 60°N-90°N) are also included in some comparisons with satellite observations.

System intercomparison for total columns

Fig. 3.6.4 shows the ozone total column over the polar and tropical latitude bands, including results from TM3DAM (green lines) and BASCOE (cyan lines). Since TM3DAM applies bias corrections to the GOME-2 data based on the surface Brewer-Dobson measurements, we use the results from TM3DAM as a “reference” for the ground-truth.

Everywhere there is an underestimation for BASCOE of about 10-20 DU. This is due to the fact that BASCOE does not assimilate any observations of the total ozone column (only Aura-MLS profiles)

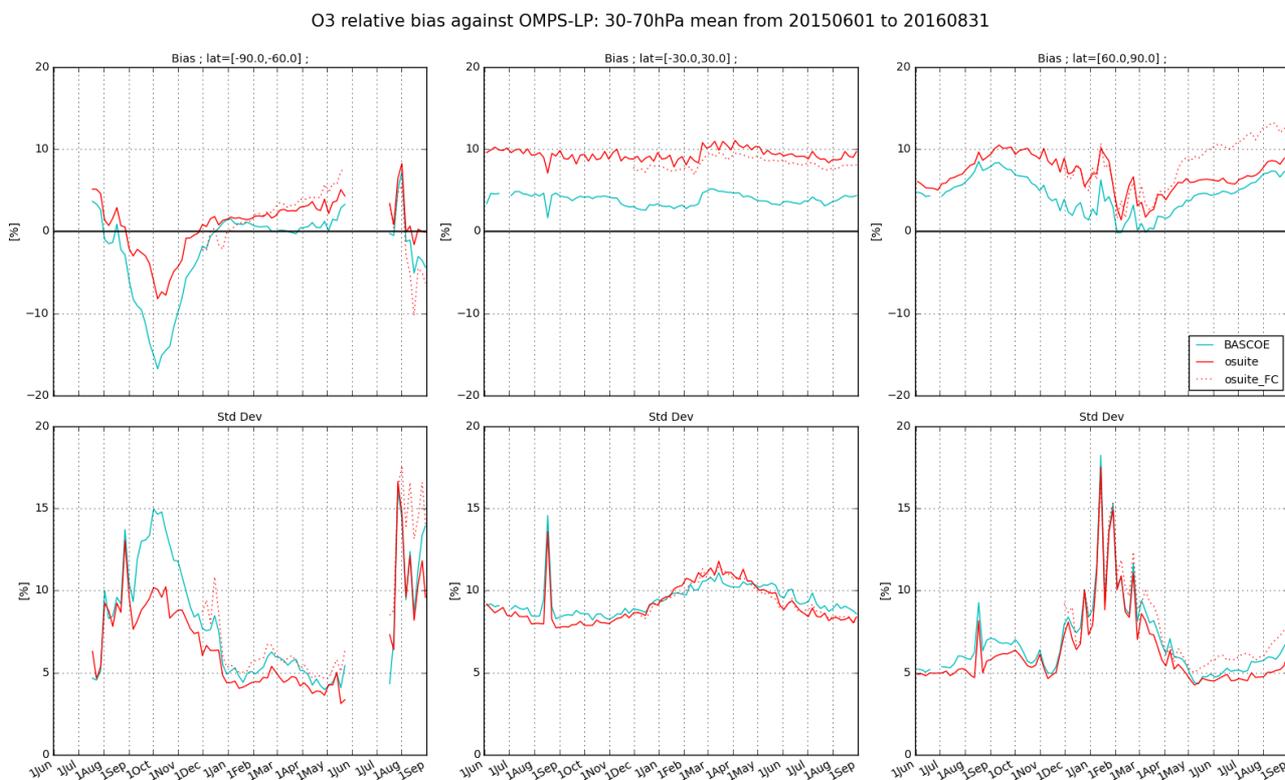


Figure 3.6.5: Time series comparing ozone from o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted), and BASCOE (cyan) with OMPS-LP satellite observations for the period 2015-06-01 to 2016-09-01 in the middle stratosphere (30-70hPa averages): top row, normalized mean bias (model-obs)/obs (%); bottom row, standard deviation of relative differences (%).

while the BASCOE model does not account for tropospheric sources of ozone. The o-suite results are much closer to those by TM3DAM:

- In the Arctic, the o-suite gives similar results to TM3DAM, except for the period of mid November 2015 to mid-February 2016, where it presents a slight overestimation of about 10 DU, i.e. ~3%.
- In the Tropics, the seasonal maximum of ozone, ranging from 270 to 290 DU, is reached in September. The o-suite presents slight underestimations w.r.t. TM3DAM of about 2-6 DU, i.e. ~2%.
- In the Antarctic, the o-suite matches TM3DAM during the whole period except for the month of July 2015 (overestimation reaching 15DU in mid-July).

Comparison with independent limb satellite datasets: OMPS-LP

In this section, we use the version 2 of OMPS-LP (i.e. the Limb Profiler) for comparison with the o-suite and BASCOE; note that it should not be confused with the nadir profiler (Kramarova et al., 2014; Taha et al., 2014). Fig. 3.6.5 shows that in the lower stratosphere (30-70hPa) there is a systematic overestimation by the o-suite (5 to 10%) and to a lesser extent by BASCOE, except over the Antarctic in September-November (i.e. ozone hole season) where the o-suite underestimates

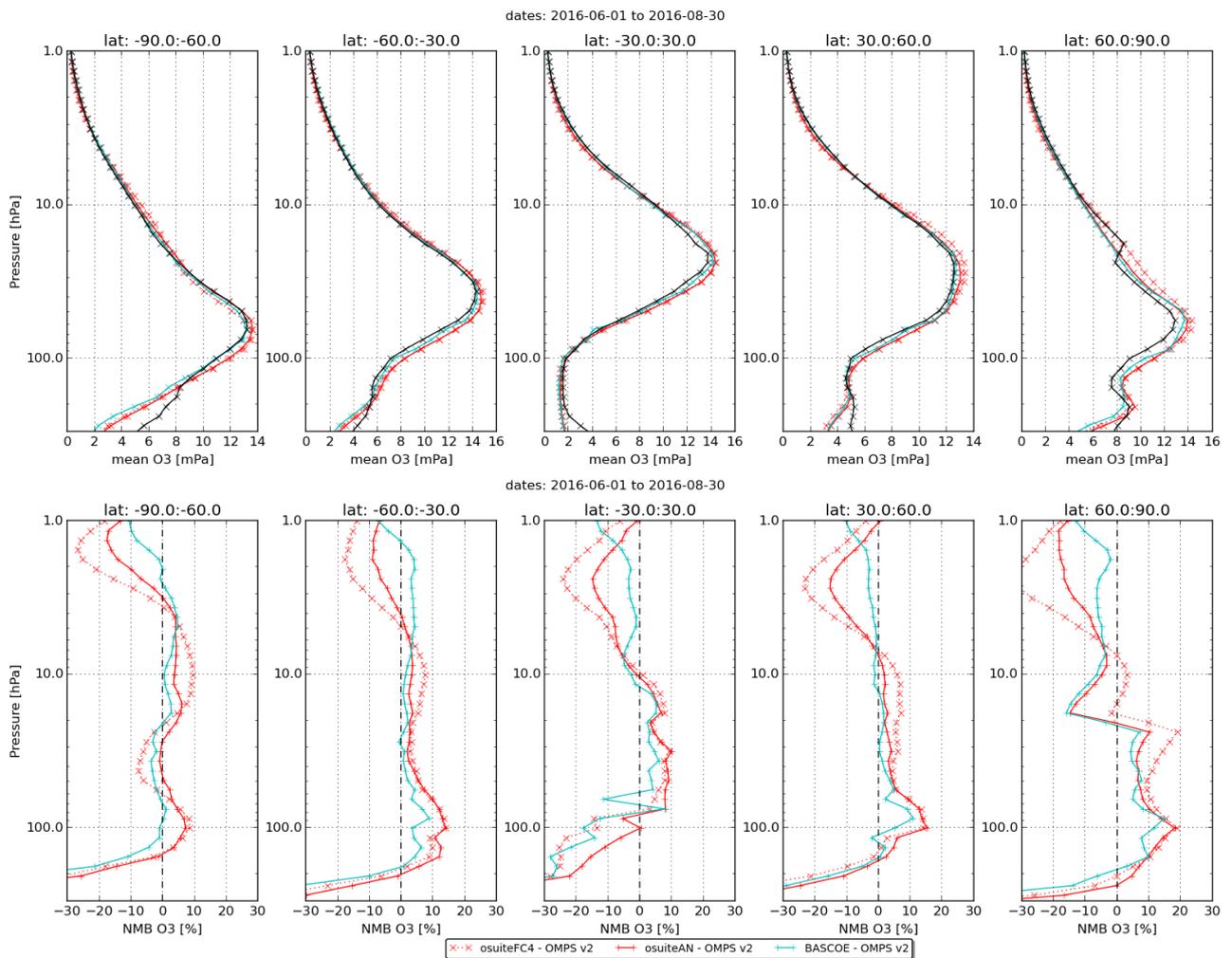


Figure 3.6.6. Mean value expressed in partial pressure (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with OMPS-LP v2 observations for the period June-July-August 2016.

ozone by up to 8%. Hence the polar ozone depletion described by the o-suite analyses is stronger than observed by OMPS-LP.

The 4th day forecasts of o-suite (since December 2015) are also depicted in Fig. 3.6.5. While the bias of the forecasts relative to OMPS-LP in the arctic region was similar to the analysis up to mid-April 2016, it has significantly increased since then (up to 13%, compared to approximately 8% for the analyses).

The bottom row of Fig. 3.6.5 shows the standard deviation of the differences and can be used to evaluate the random error in the analyses. Hence in the lower stratosphere, the random error of the o-suite is evaluated at 7% to 12% in the Tropics and varies in the polar regions from 5% (summer and fall) to 15% (winter and spring).

Fig. 3.6.6 displays vertical profiles of the relative biases between the o-suite or BASCOE and OMPS-LP. The difference is averaged over the most recent 3-month period considered in this validation report, i.e. June-August 2016. In the northern hemisphere, a vertical discontinuity of the relative

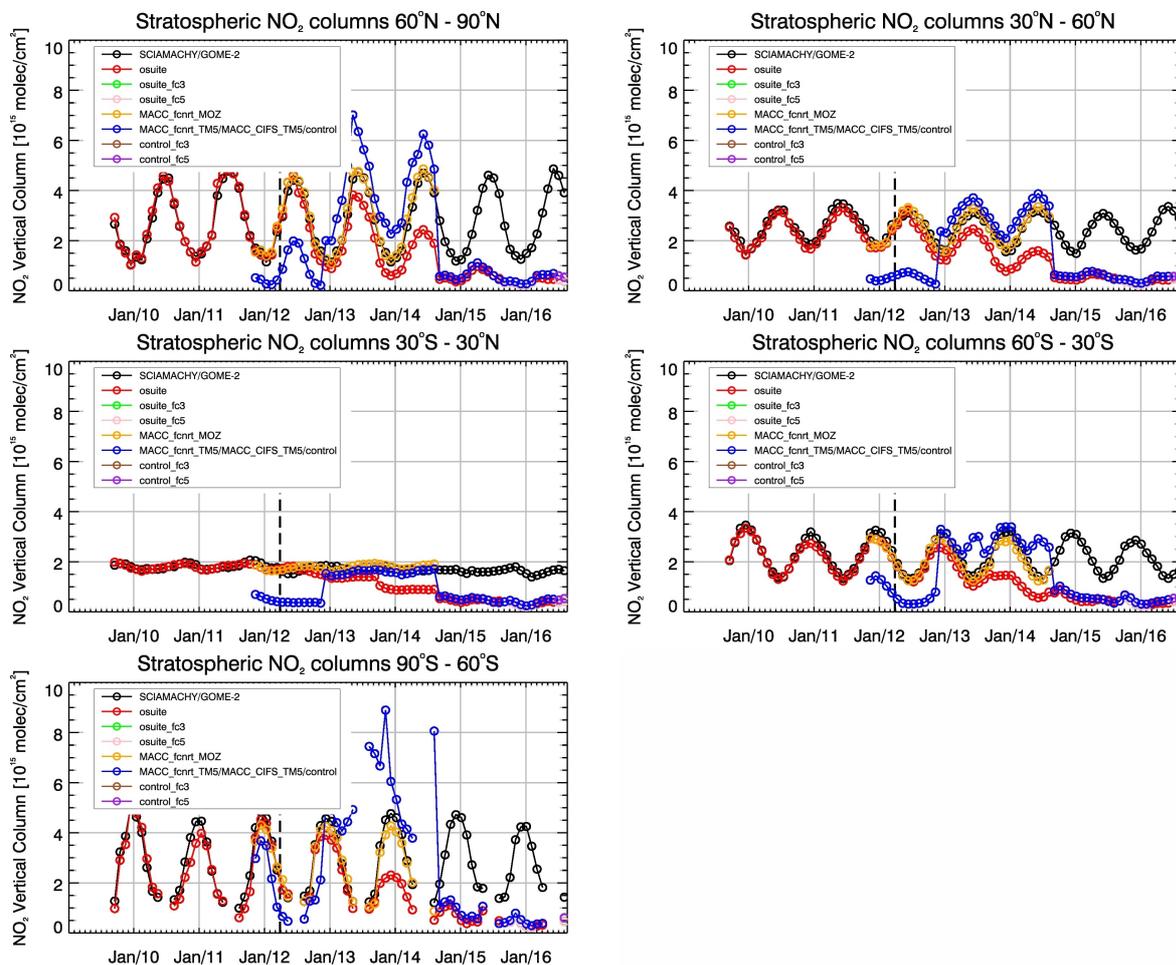


Figure 3.7.1: Time series of average stratospheric NO₂ columns [10^{15} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012) compared to model results for different latitude bands. See text for details. The blue line shows MACC_fcnrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 until August 2014 and control results from September 2014 onwards. The vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

differences is noted at 20 hPa, but this is a spurious feature due to a vertical discontinuity in the OMPS retrievals used here (transition from UV to visible detector).

This quantitative comparison with OMPS-LP confirms the good agreement in the middle stratosphere while the lower stratosphere (< 70hPa) reveals stronger discrepancies. The comparison with BASCOE (which assimilates the offline Aura-MLS dataset) confirms that the lower stratospheric vertical oscillations seen against Aura-MLS in the Tropical band (not shown) are an artifact.

3.7 Stratospheric NO₂

In this section, nitrogen dioxide from SCIAMACHY/Envisat satellite retrievals (IUP-UB v0.7) and GOME-2/MetOp-A satellite retrievals (IUP-UB v1.0) are used to validate modelled stratospheric NO₂ columns. Monthly mean stratospheric NO₂ columns from SCIAMACHY and GOME-2 have relatively small errors of the order of 20% in the tropics and in mid-latitudes in summer and even lower errors



at mid-latitudes in winter. As the time resolution of the saved model files is rather coarse and NO_x photochemistry in the stratosphere has a large impact on the NO_2 columns at low sun, some uncertainty is introduced by the time interpolation at high latitudes in winter.

As shown in Fig. 3.7.1, amplitude and seasonality of satellite stratospheric NO_2 columns are poorly modelled with CB05-based chemistry runs including the most recent version of the o-suite. There are no significant differences between o-suite and its control experiment. The significant differences between observations and CB05 chemistry runs, i.e. a strong underestimation of satellite retrievals by models, can be explained by the missing stratospheric chemistry for these model versions. The only constraint on stratospheric NO_x is implicitly made by fixing the HNO_3/O_3 ratio at the 10 hPa level. This assumption, in combination with the changing model settings for stratospheric O_3 for control compared to MACC_CIFS_TM5, may explain some of the jumps we see in stratospheric NO_2 . In any of these runs the stratospheric NO_2 is poorly constrained. It clearly indicates that stratospheric NO_2 in the latest version of the o-suite is not a useful product and should be disregarded.

Comparison of the o-suite from July 2012 until August 2014 with the other model runs and satellite observations shows that the previous version of the o-suite stratospheric NO_2 columns have a systematic low bias relative to those from MACC_fcrt_MOZ and satellite observations for all latitude bands. For example, o-suite values are a factor of 2 smaller than satellite values between 60°S to 90°S for October 2013. Best performance was achieved with the MOZART chemistry experiments without data assimilation (MACC_fcrt_MOZ, running until September 2014), especially northwards of 30°S. Details on the NO_2 evaluation can be found at: http://www.doas-bremen.de/macc/macc_veri_iup_home.html



4. Validation results for greenhouse gases

4.1 CH₄ and CO₂ validation against ICOS observations

This section describes the NRT validation of the pre-operational, high resolution forecast of CO₂ and CH₄ from 1st September 2015 to 31 August 2016 based on observations from 15 surface stations. Over this period the high resolution forecast corresponds to two experiments: gf39 and ghqy, see also Sec. 2.1.3. The same experiments are used for the validation with ICOS surface data and TCCON total column data.

In the present comparison we consider 24 hours daily means, without distinction of nighttime and daytime data although they may have very different footprints which could justify a separate analysis. It is important to note the change of experiment on March 1st 2016, leading to an abrupt change of CO₂/CH₄ mole fractions at some sites like Amsterdam Island. We should also note that compared to the previous reports we have changed the model gridbox used to compare with observations at three coastal stations: Finokalia (FKL), Biscarrosse (BIS) and S^t Denis in Reunion Island (STD). Initially we were extracting the closest continental gridbox for each site, but due to the influence of local emissions (anthropogenic and/or biospheric) this was resulting in an overestimated variability at diurnal and synoptic scales. For those three stations we are now using the closest marine gridbox. This change has greatly improved the comparison with observations at FKL, and has little impact for BIS. For STD the situation is more complicated since we have to choose between two coastal gridboxes (figure 4.1). The eastern and closest box (#1) improved the CH₄ comparison; whereas the western coastal box (#2) improved the CO₂ comparison (figure 4.1.1). By default the gridbox #1 will be used for the model evaluation.

The figures 4.1.1 and 4.1.2 present the annual metrics for all stations (including the new locations described in the previous paragraph) based on the comparison of the daily means mole fractions. For the three stations located in the southern Indian Ocean (AMS, RUN and STD) the mean annual biases would not be meaningful due to the change of experiment occurring on 1st March 2016 (gf39 to ghqy) and leading to an abrupt change in the simulated molar fractions (see AMS on figures 4.1.4 and 4.1.5). Indeed for those three stations the mean CH₄ biases calculated before and after the change of experiment are [-40.6/-3.2], [-33.8/4.1], [-21.1/36.2] for AMS, STD and RUN respectively. For CO₂ the same figures are [-2.7/-1.5], [-8.2/-7.2], [-1.1/1.6]. For the background station AMS the bias has been significantly reduced in the latest experiment (ghqy). For the coastal site STD it has been greatly improved for CH₄ and slightly improved for CO₂. For the mountain site RUN there is no improvement, but we are still facing an issue of spatial representativeness for this station. Due to the abrupt change of the simulations between the two experiments we have focused the statistics only on the latest one (ghqy) starting on 1st March 2016. Averaged over the 15 stations, the mean CO₂ and CH₄ biases are respectively -0.3 ± 2.8 ppm and -3.1 ± 13.8 ppb. For European stations the yearly mean biases are 0.3 ± 2.5 ppm (RMSE: 5.0 ppm) and -5.4 ± 11.1 ppb (RMSE: 25.7 ppb). The CO₂ biases show a significant seasonal pattern at most European sites (see for example MHD and TRN in figure 4.1.4) due to an overestimate of the seasonal cycle amplitude with too high concentrations in winter and too low in summer.

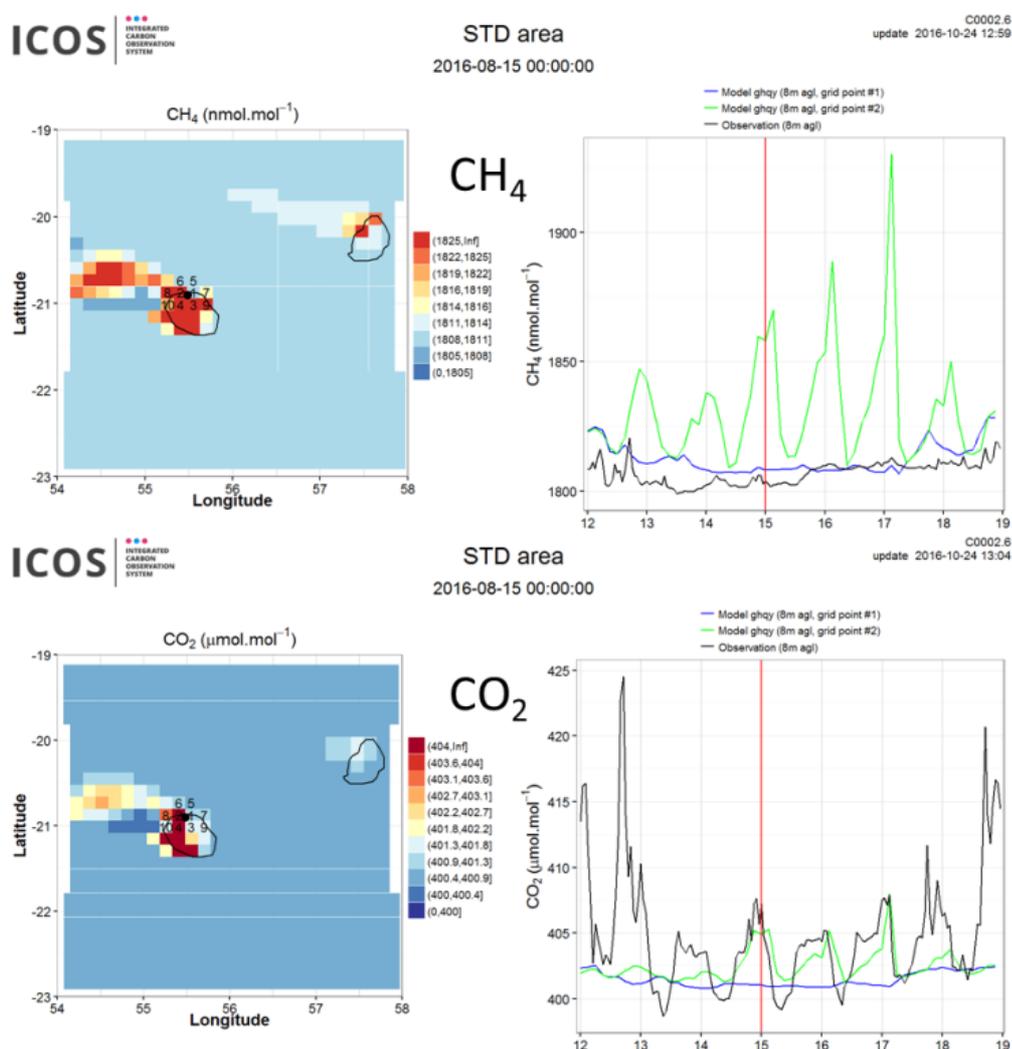


Figure 4.1.1: Left: CH₄ (above) and CO₂ (below) molar fraction distribution at the surface around La Réunion Island on 15-08-2016 (0hr). Right: Comparison of the simulation of CH₄ (above) and CO₂ (below) simulations at two coastal grid boxes (#1 in blue, #2 in green) to the observations (in black) during a 7 days period.

Figures 4.1.4 and 4.1.5 show the CO₂ and CH₄ comparisons at four selected stations in South hemisphere (Amsterdam I.), North hemisphere (Mace Head, Trainou tall tower) and in the tropics (Lamto). When considering only the latest experiment (ghqy) the remote station of Amsterdam Island is the site where we have the best correlation coefficients for CH₄ (0.97). The correlation is not as good for CO₂ (0.85) and we obtain better results at background site in the Northern hemisphere, like MHD, PDM and PUY where the correlation coefficients are greater than 0.9 ppm. The figure 4.6 shows an example of synoptic event at Amsterdam Island which occurred on June 2016. During this season we observed sometimes an increase of the observed trace gases, like Radon, CO₂, CH₄, CO, black carbon. Such events cannot be explained by local emissions and are due to rapid advectations (2-3 days) of pollutants from the southern Africa (Figure 4.1.7). This example demonstrates the capacity of the model to simulate this transportation over Indian Ocean, since both the amplitude and timing of the CH₄ spike is perfectly reproduced. For CO₂, the model underestimates the increase of concentrations and fails to reproduce the daily variation, indicating that the biospheric surface flux should be improved.

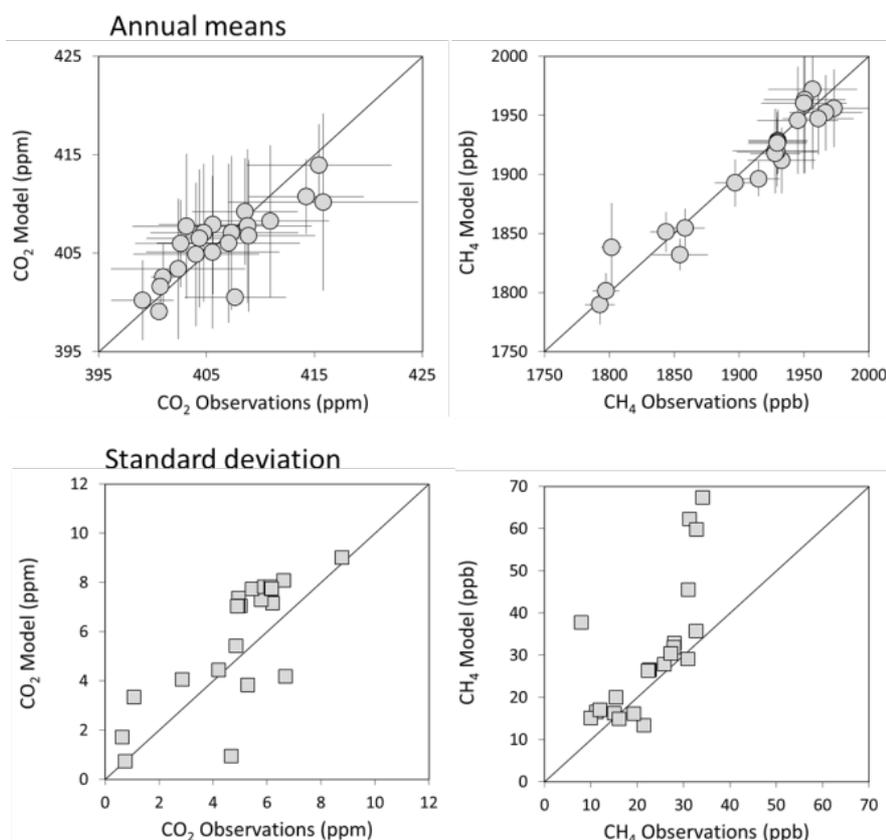


Figure 4.1.2: Above: Comparison of the observed and simulated means of CO₂ (left) and CH₄ (right) molar fractions at the 15 sites, calculated from the daily averages over the period 1st March to 31 August 2016 (experiment GHQY). Below: standard deviation of daily means CO₂ (left) and CH₄ (right) concentrations over the year.

In Northern hemisphere it is interesting to note at the tall towers (TRN, OPE, OHP) an improvement of the model performances in terms of bias, RMSE and coefficient correlations when looking at 100m or more above ground level (Figure 4.1.3). This is confirming the wise choice of the ICOS project to invest in tall towers for continental stations, since the model performances at continental sites above 100mh high are generally comparable to background sites.

In the tropics, even if the performances of the model are more difficult to evaluate due to the low number of monitoring sites (RUN, STD, LTO, GUY, CHC) and to the problem of spatial representativeness at some sites, it appears that the correlation coefficients are lower than in Europe. At the tropical site of Lamto, the phasing of the seasonal cycle is relatively good with maximum CO₂ and CH₄ concentrations from December to February due to biomass burnings in Western Africa. The model captures pretty well the double structure of this high concentration period including a temporary decrease in mid-January. However, the model overestimates the CO₂ peak (Figure 4.1.4), and underestimates the CH₄ ones (Figure 4.1.5).



Metrics comparison

2016-03-01 - 2016-08-31

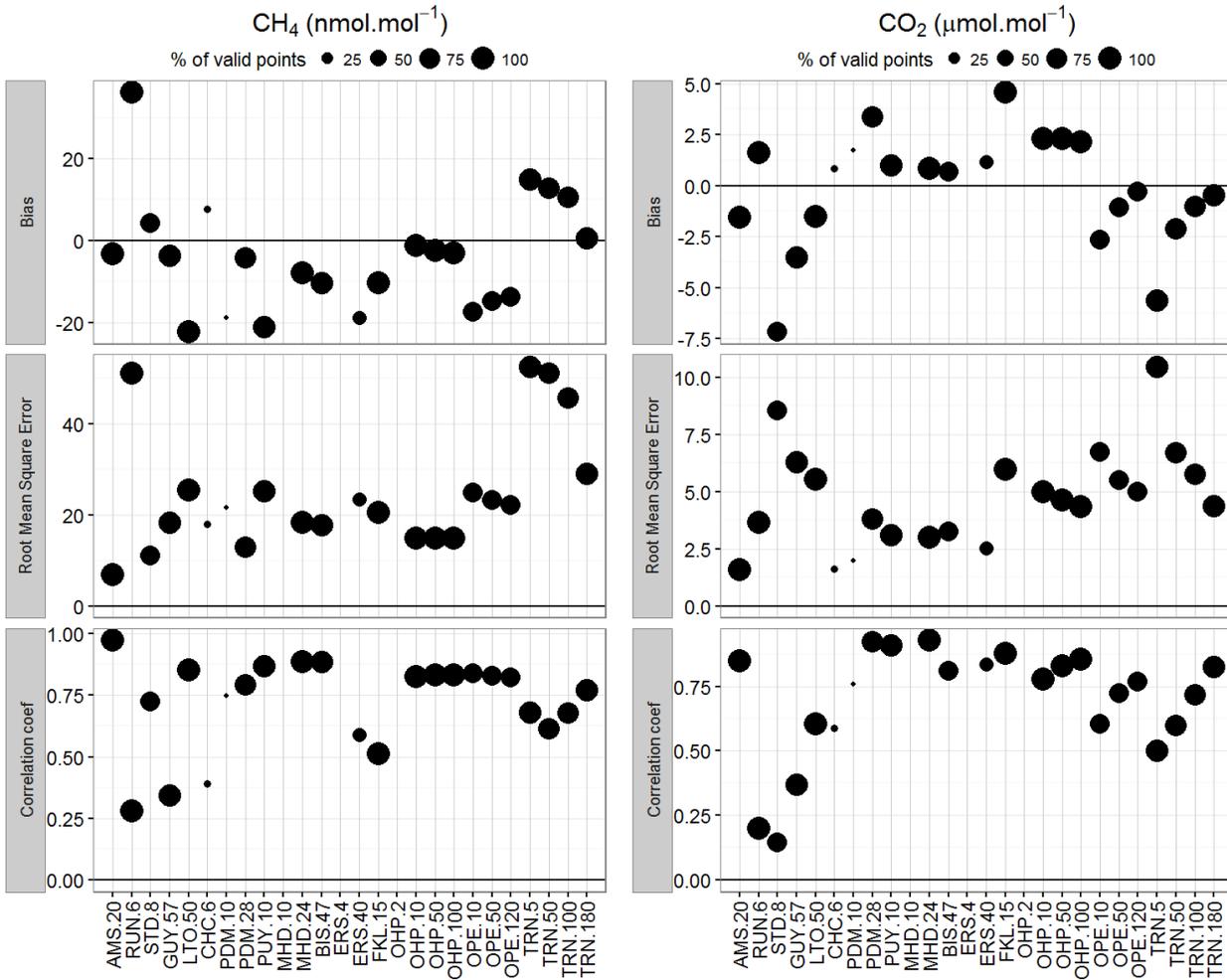


Figure 4.1.3: Annual metrics (bias, RMSE and coefficient correlation) calculated from the model-data comparison for daily means of CH₄ (left) and CO₂ (right) at the 15 sites (with multiple sampling heights at the last three sites). The size of each points is proportional to the percentage of available days over the considered period (1st March to 31 August 2016, experiment ghqy).

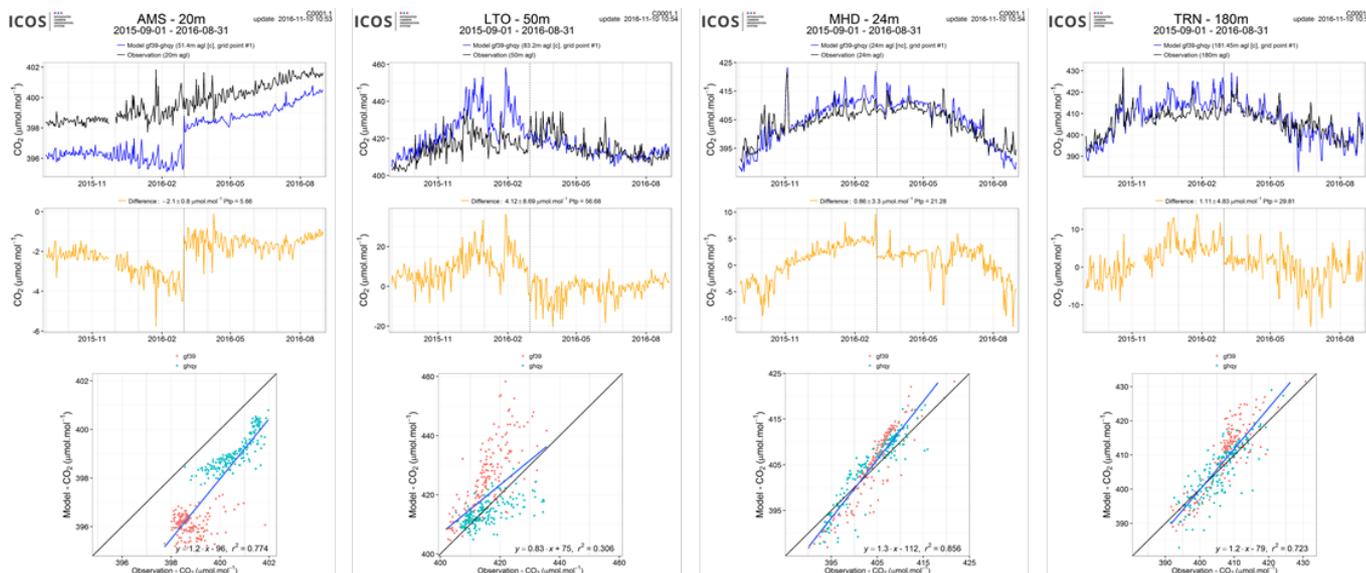


Figure 4.1.4: Above: Comparison of CO₂ daily means observed (red) and simulated (blue) at four stations (Amsterdam I., Mace Head, Lamto and Trainou tall tower). Middle: differences of the observations minus the simulations. Below: Linear fit between observations and simulations. The blue points correspond to the ghcy experiment, initialized on March 1st, 2016.

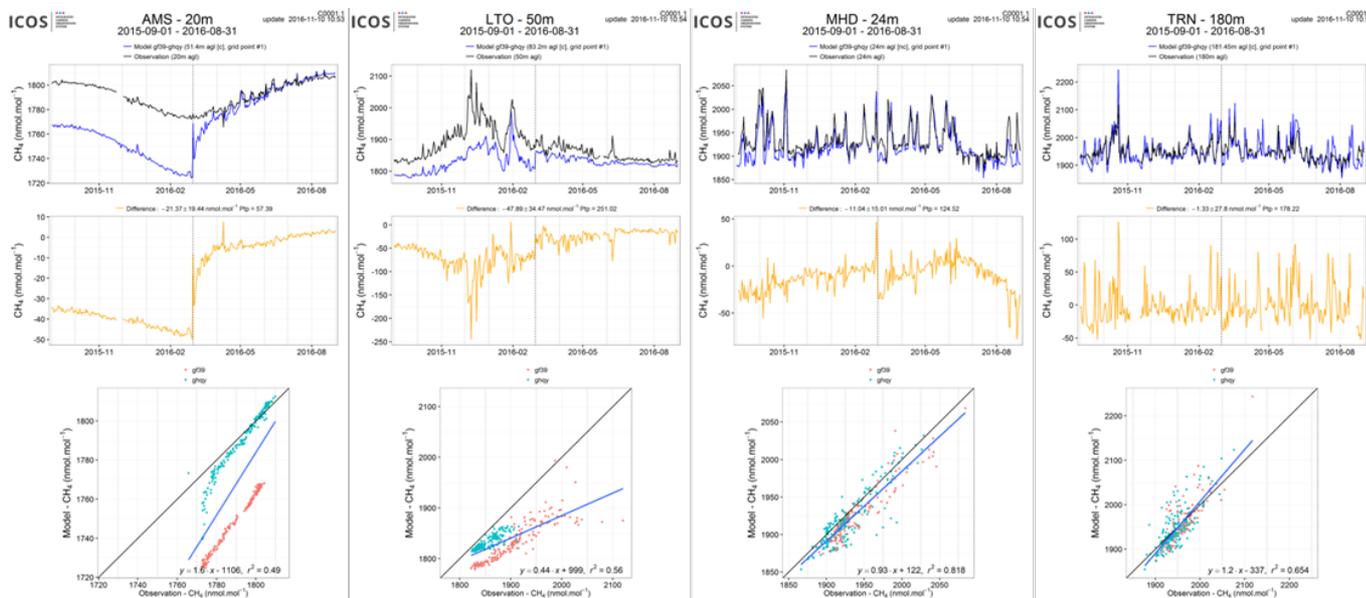


Figure 4.1.5: Same as figure 4.4 for CH₄

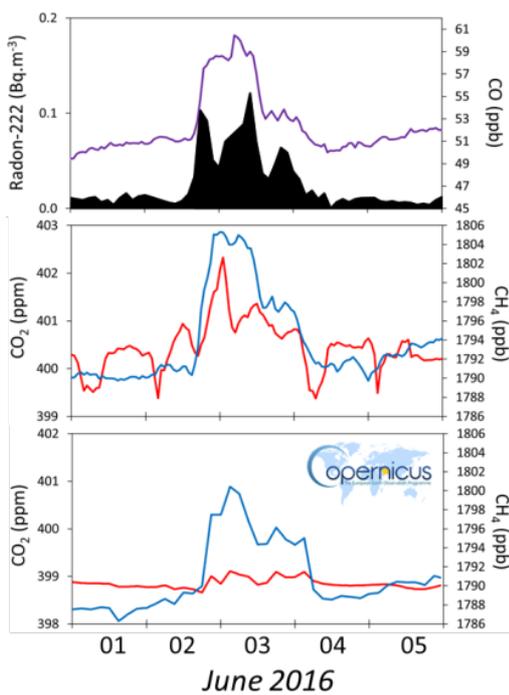


Figure 4.1.6: Short term variabilities of trace gases at Amsterdam Island during a radonic storm (01-05 June 2016) originating from Southern Africa. The above figure shows mole fractions of Radon-222 (black) and CO (purple). The middle figure shows mole fractions of CO₂ (red) and CH₄ (blue). The below figure shows simulations of CO₂ (red) and CH₄ (blue).

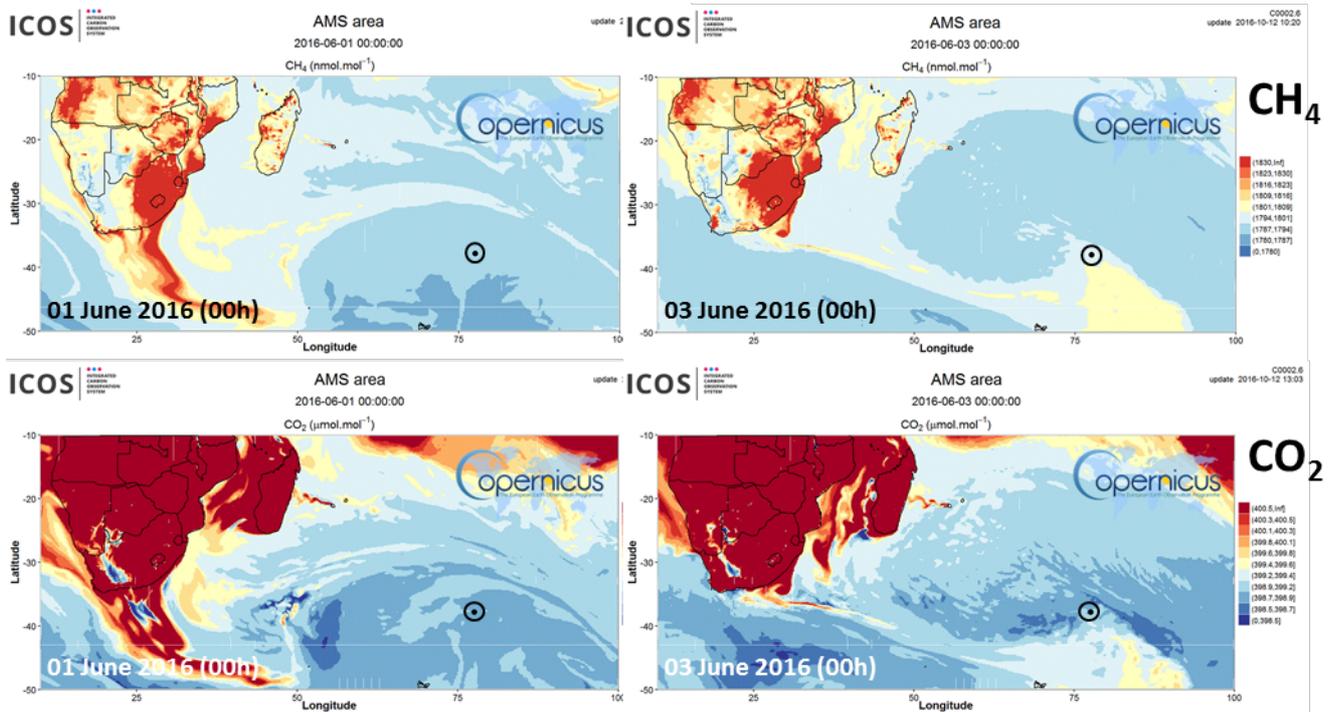


Figure 4.1.7: Simulated distribution of CH₄ (above) and CO₂ (below) mole fractions at the surface level on 1st (left) and 3rd (right) June 2016. The black circle indicates the position of Amsterdam Island.



4.2 CH₄ and CO₂ validation against TCCON observations

For the validation column averaged mole fractions of CO₂ and CH₄ (denoted as XCO₂ and XCH₄) from the Total Carbon Column Observing Network (TCCON) are used. Column averaged mole fractions provide a different information content than the in situ measurements and are therefore complementary to the in situ data. For example if models suffer from problems in vertical transport, the combination of TCCON and surface in situ measurements will provide a means to detect this.

For the model validation the official TCCON data cannot be used due to its availability of typically one year after the measurement. Some TCCON sites are providing rapid delivery data (RD-TCCON data), which is available at least one month after the measurement. TCCON sites that deliver RD-TCCON data currently include Trainou (France), Bialystok (Poland) and Reunion (France). Over the course of the project more TCCON sites might contribute. This largely depends on funding for the fast data product.

The validation routines used for TCCON data are the same as used for the NDACC network and are documented in Langerock et al. (2015). The routines have been adapted to use the TCCON data format.

4.2.1 Evaluation against TCCON CO₂

The data presented in the Figures 4.2.1-4.2.3 show a comparison for a full seasonal cycle from March 2015 – June 2016. At Bialystok (Fig. 4.2.1) and Orleans (Fig. 4.2.2) the difference between the model and the measurement shows a very similar seasonal pattern. The model overestimates the XCO₂ at both sites for the period January to July, where the modelled XCO₂ is in some cases more than 1% too high. During the minimum in August/September 2015 and August 2015 the overall agreement between the model and the measurements is reasonably good. This good agreement is still present during the XCO₂ increase from September 2015 to November 2015. At Orleans the model overestimates the XCO₂ from December 2015 to June 2016. At Bialystok the overestimation starts later (January 2016), but is then similar to Orleans in 2016.

At Reunion (Fig. 4.2.3) the overall agreement of the annual means is good but the model shows short-term fluctuations of $\pm 1\%$, which are not seen in the measurements. These short-term variations of several ppm are not reasonable and mostly occur between October 2015 and January 2016.

4.2.2 Evaluation against TCCON CH₄

At Bialystok and Orleans (Figs. 4.2.4 and 4.2.5) the model underestimates the seasonal amplitude. From April 2015 – July 2015 the modelled XCH₄ is up to 1% higher than the measurements and from September 2015 – January 2016 the modelled XCH₄ is up to 1% lower than the measurements. From April 2016 – July 2016 the agreement between model and measurement is better than during the same period in 2015. However, the data for August 2016 shows again modelled values that are up to 1% too low. At Reunion (Fig. 4.2.6) the modelled values were systematically too low between March 2015 and February 2016. This problem in the model seems to be resolved and from April 2016 onwards a relatively good agreement between the measurements and the model exist.

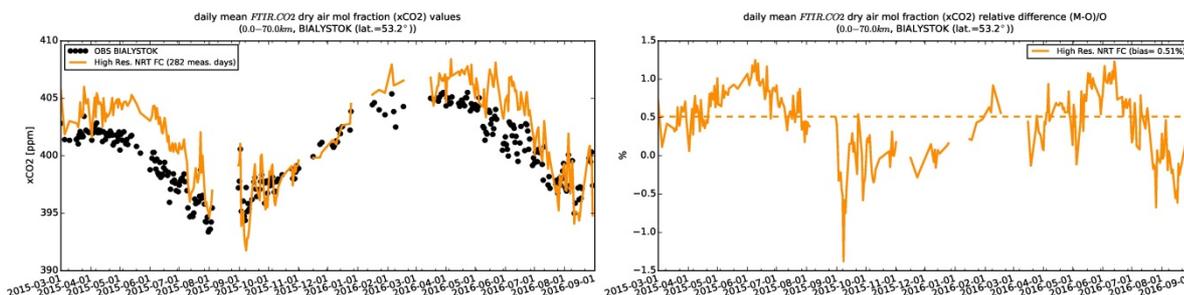


Figure 4.2.1: Time series of column averaged mole fractions (left) and relative difference (right) of carbon dioxide (CO₂) at the TCCON site Bialystok compared to high resolution NRT FC data (yellow).

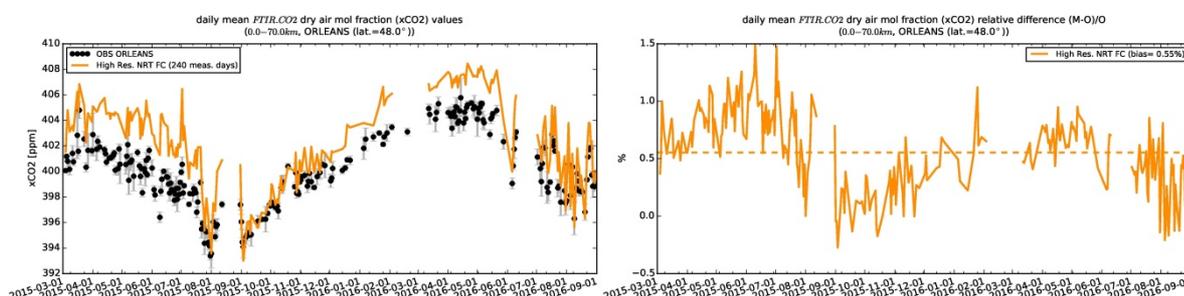


Figure 4.2.2: Time series of column averaged mole fractions (left) and relative difference (right) of carbon dioxide (CO₂) at the TCCON site Orleans compared to high resolution NRT FC data (yellow)

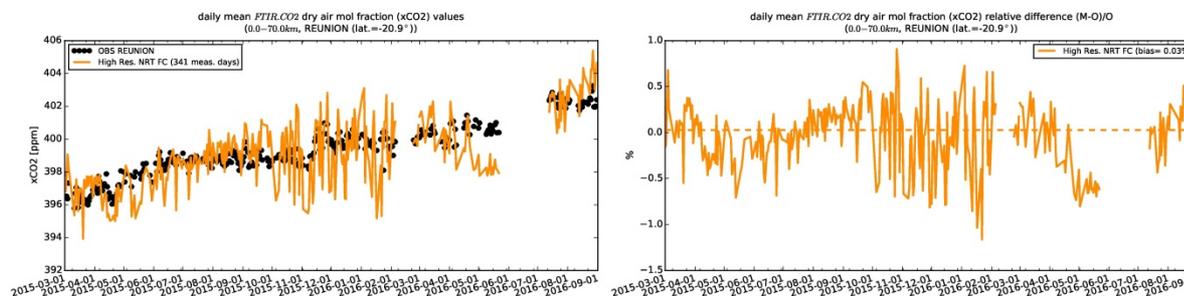


Figure 4.2.3 Time series of column averaged mole fractions (left) and relative difference (right) of carbon dioxide (CO₂) at the TCCON site Reunion compared to high resolution NRT FC data (yellow)

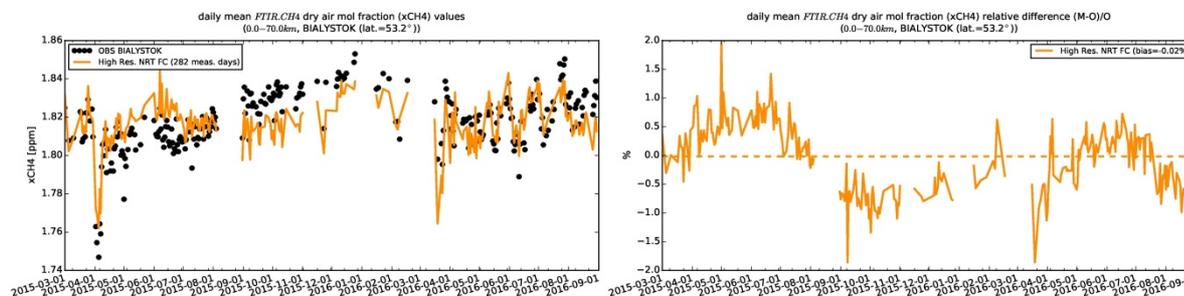


Figure 4.2.4: Time series of column averaged mole fractions (left) and relative difference (right) of methane (CH₄) at the TCCON site Bialystok compared to high resolution NRT FC data (yellow).

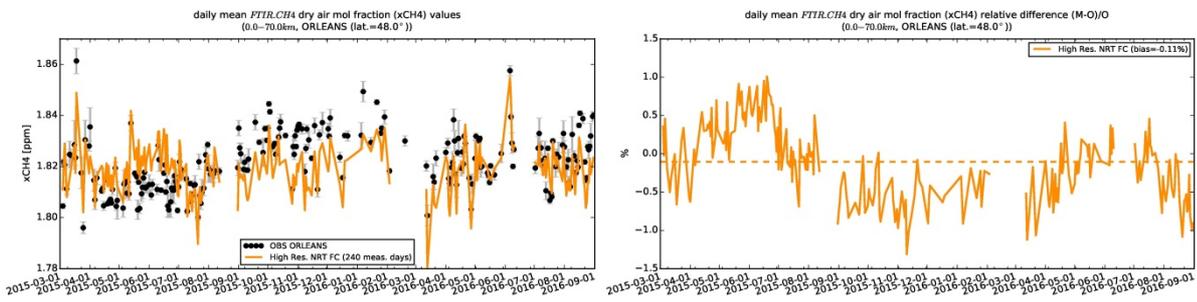


Figure 4.2.5: Time series of column averaged mole fractions (left) and relative difference (right) of methane (CH₄) at the TCCON site Orleans compared to high resolution NRT FC data (yellow)

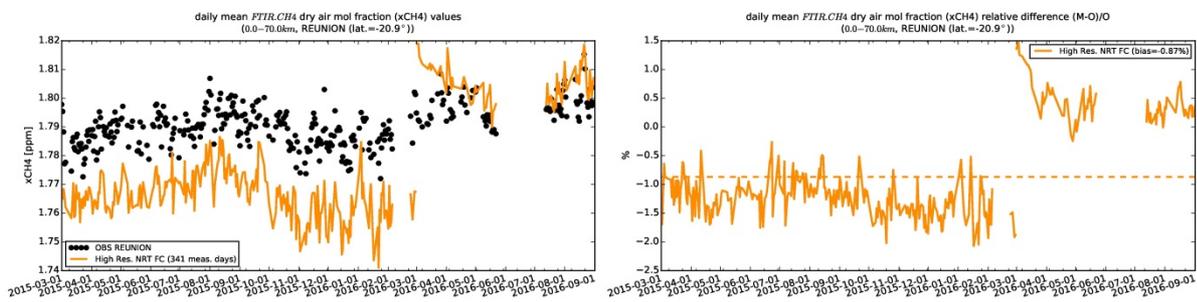


Figure 4.2.6: Time series of column averaged mole fractions (left) and relative difference (right) of methane (CH₄) at the TCCON site Reunion compared to high resolution NRT FC data (yellow).

5. Events

5.1 Fire case in the central part of South America in late August 2016

A fire event took place in the central part of South America during the 20th of August 2016. IASI data show a plume of CO over this region with south-east transport towards the Atlantic Ocean later on (Fig. 5.1.1). It is difficult to compare model results with IASI data for this case due to missing values, but it still can be seen that both model runs captured the location of the plume on 25th and 26th of August and south-east transport of CO on the 29th and the 30th of August. CO values from both model runs are very similar and seem to be in good agreement with the satellite data over the region of fire case and the transportation pathway in the grid boxes where data are present. Over the oceans the control run shows larger values compared to satellite data.

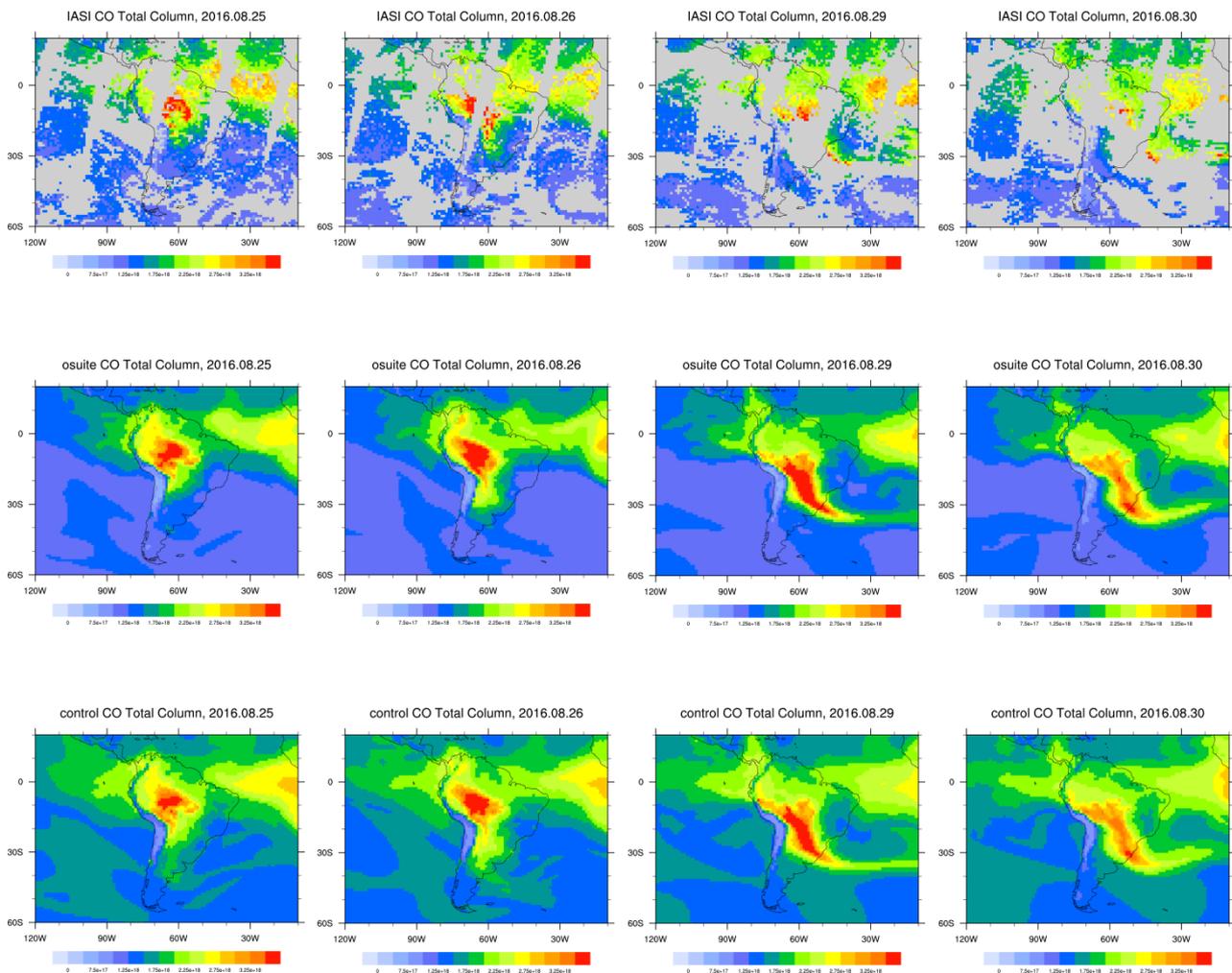


Fig. 5.1.1: CO total column from IASI (top), o-suite (middle) and control runs (bottom) for 25, 26, 29 and 30 of August 2016 over the selected region.

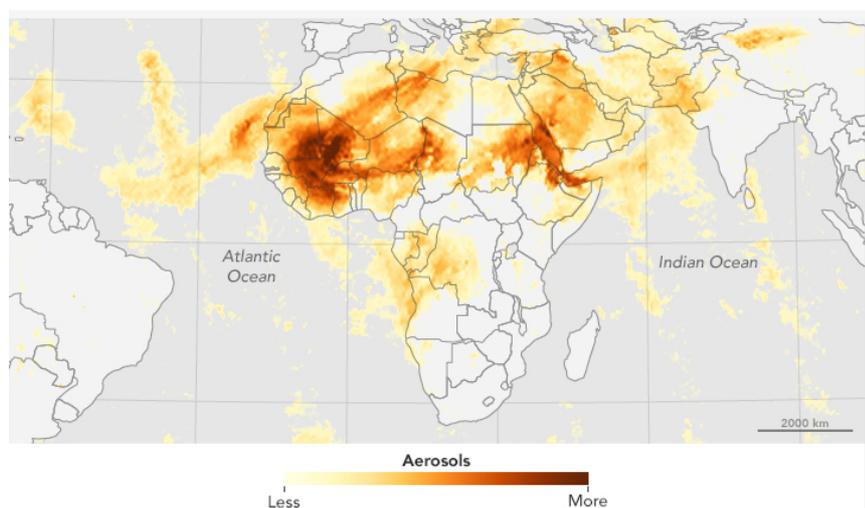


Figure 5.2.1. Aerosol content on June 19, 2016 from the Ozone Mapping Profiler Suite (OMPS) on the Suomi-NPP satellite. High concentrations are represented with shades of deep red; the lowest concentrations are shades of light yellow. The largest, thickest plume appears to stem from the Sahara in western Africa. Source: NASA Ozone Mapping and Profiler Suite.

5.2 A dusty period over North Africa, Middle East and Europe: mid-June 2016

In mid-June 2016, winds lofted thick plumes of dust from northern Africa's deserts high into the air. On June 19, winds had already swept a plume of dust westward over the Atlantic Ocean and the archipelago of Cabo Verde (Cape Verde) that reached the Canary Islands by June 22. Otherwise, a thinner plume of Saharan dust also spread north toward Europe starting on June 16. This dust outbreak from the Western Sahara was coincident with a plume of dust from Africa's northeast, carried eastward over the Red Sea. Figure 5.2.1 shows the concentration of aerosols on June 19, 2016, produced with data from the Ozone Mapping Profiler Suite (OMPS) on the Suomi-NPP satellite.

Dust aerosol optical depth (DOD) from CAMS has been compared with AOD from MODIS to see the skill of CAMS to track the spatiotemporal evolution of the different dust plumes analysed in the present section (Figure 5.2.1). The near-real-time MODIS aerosol product available through the NASA's EOSDIS system (MCDAODHD files), is used for this purpose. It is a level 3 gridded product specifically designed for quantitative applications including data assimilation and model validation. DOD simulated by CAMS o-suite and observed AOD by MODIS from 16th to 22nd June 2016 at 12UTC is shown in Figure 5.2.2. Moreover, DOD values from CAMS o-suite have been compared with those from CAMS control; Multi-Median model generated from the models participating in the WMO SDS-WAS NAMEE Regional Node (<http://sds-was.aemet.es/>) and AERONET AOD in four AERONET stations strategically located along the path of the different dust plumes.

CAMS o-suite can timely reproduce the spatial distribution of the different dust plumes over the North Atlantic on 19-21 June, affecting Europe on 17-19 June and the Red Sea on 17-20 June as observed by MODIS (Figure 5.2.2). We can see how CAMS o-suite tracks fairly well the changes in both shape and size of the dust layer throughout the dusty period. The whole episode is well simulated by CAMS o-suite and CAMS control, in the North Atlantic (Dakar and Santa Cruz de Tenerife in Figure 5.2.3) and Europe (see IMAA Potenza in Figure 5.2.3). In the Kaust Campus



AERONET site, all the models underestimate the maximum AOD peak observed on June 18th (see Figure 5.2.3). The dust plume affecting the Red Sea has its origin in the Tokar Delta on June 15th and out over the Red Sea and toward the Arabian Peninsula the next days arriving at the AERONET Kaust Campus (in Saudi Arabia) on June 18th (see Figure 5.2.3). The wind gusts that caused the dust outbreak on June 15 were due to a cold front moving south-east. The front was related to a cyclone centred near the Persian Gulf, and it caused turbulent mixing of air and a series of associated haboobs that any of the models can reproduce.

Since weather records have a good spatial and temporal coverage, horizontal visibility observations included in meteorological reports can be used as an alternative way to monitor dust events. Visibility is mainly affected at ground by the presence of aerosols and water in the atmosphere. On surface level, o-suite can track the reduction of visibility of the study period (see Figure 5.2.4) localising the origin of the event over Northern Algeria and the Tokar Delta on June 17th and in Mauritania and Mali on June 19-20.

Otherwise, visibility records also show an intense dust event with origin in Iraq on June 16 moving southwards crossing the Persian Gulf the next days. This dust outbreak is reproduced by CAMS o-suite (see Figure 5.2.2 and see Figure 5.2.4) but, its intensity is underestimated as it is partly shown in the AOD comparison with MODIS (see Figure 5.2.2).

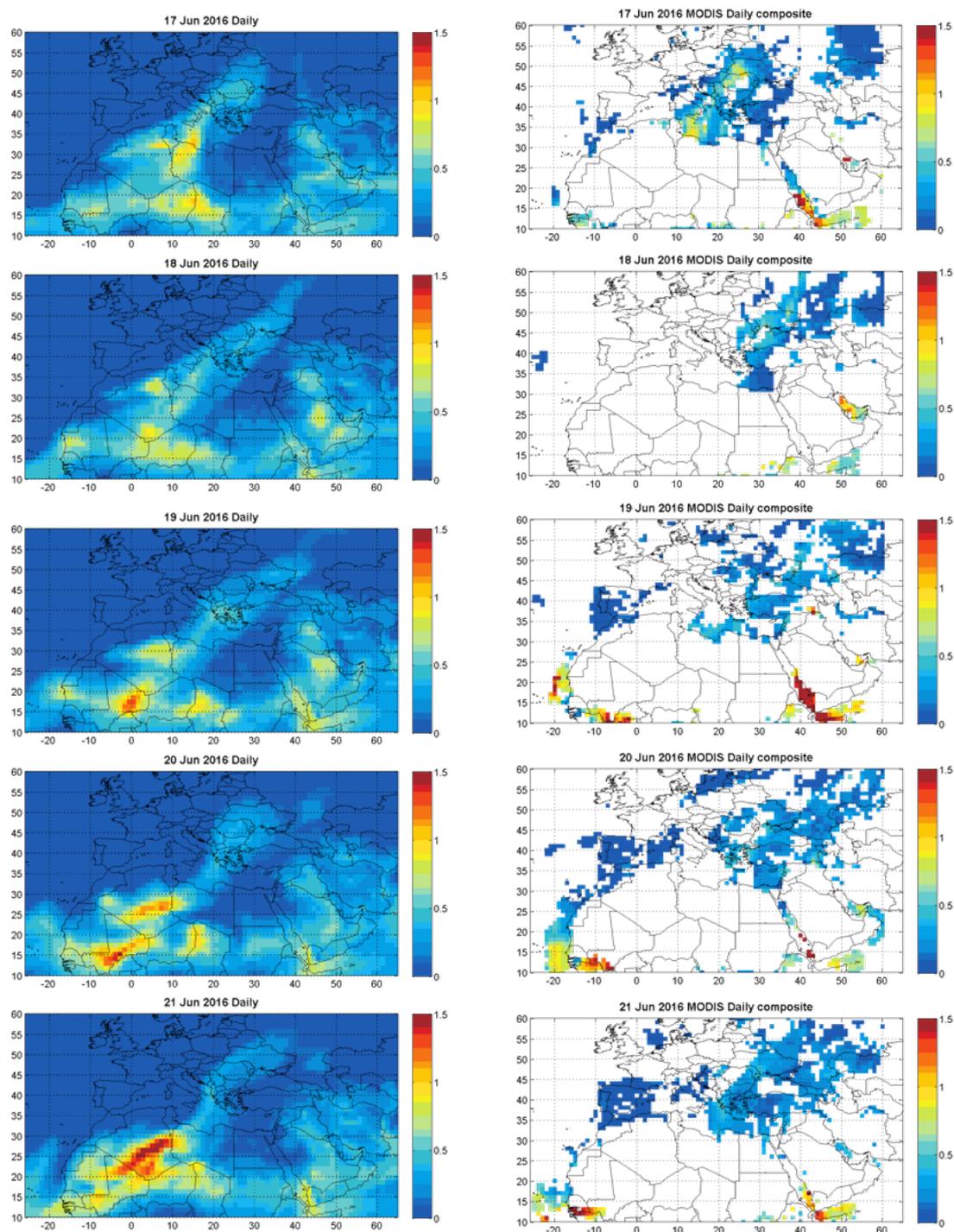


Figure 5.2.2. DOD from o-suite (right column) and AOD from MODIS combined Dark Target and Deep Blue aerosol products (left column), for June 17th- 21st, 2016 at 12UTC.

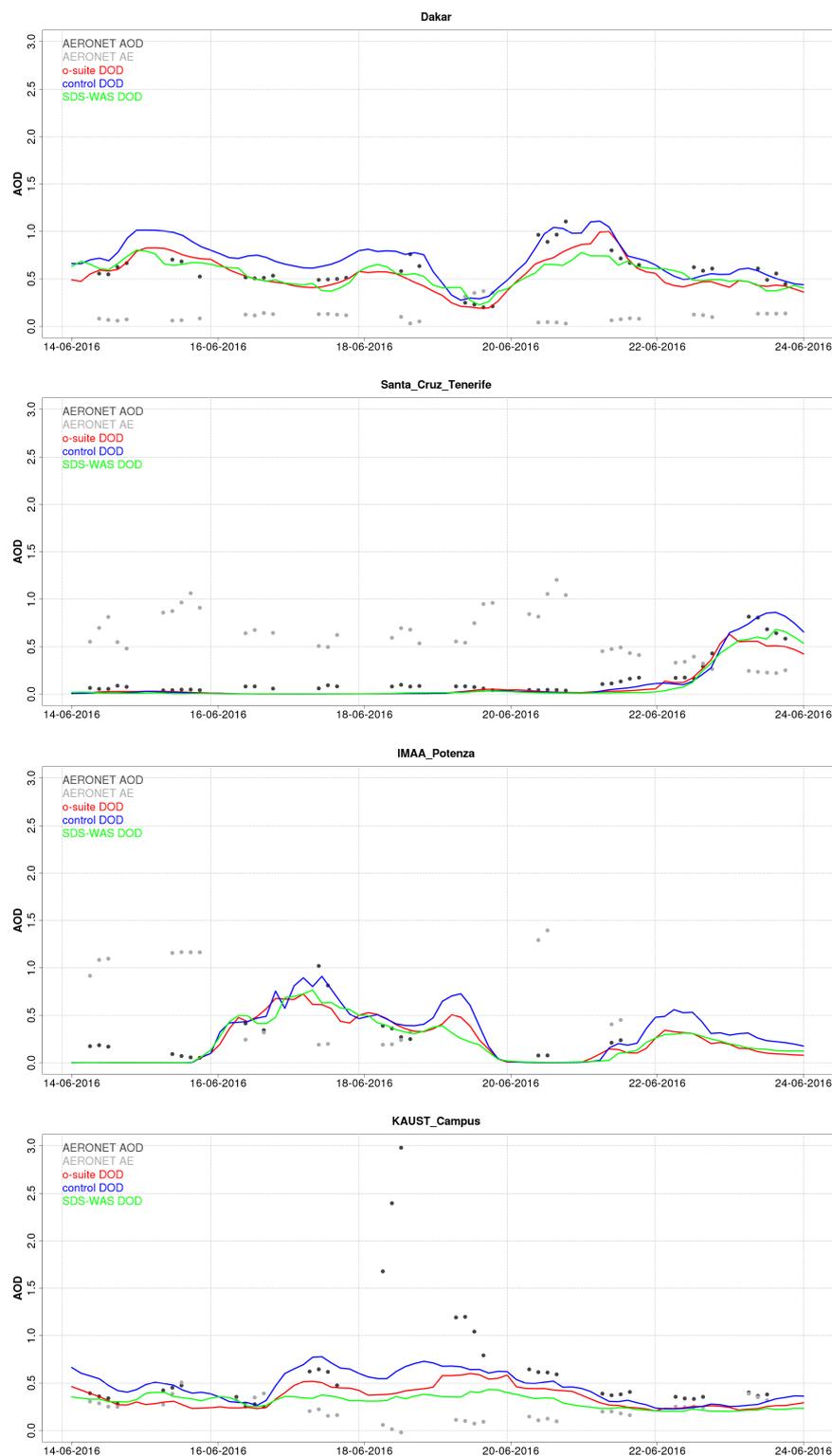


Figure 5.2.3. AOD at 550 nm from AERONET (black), DOD at 550 nm from the o-suite (blue), DOD at 550 nm from the control run (red), and DOD at 550 nm from SDS-WAS Multi-model Median (green) at Ben Salem (Tunisia) and Etna (Italy) AERONET sites during the case analysis from 5th to 18th May 2016.

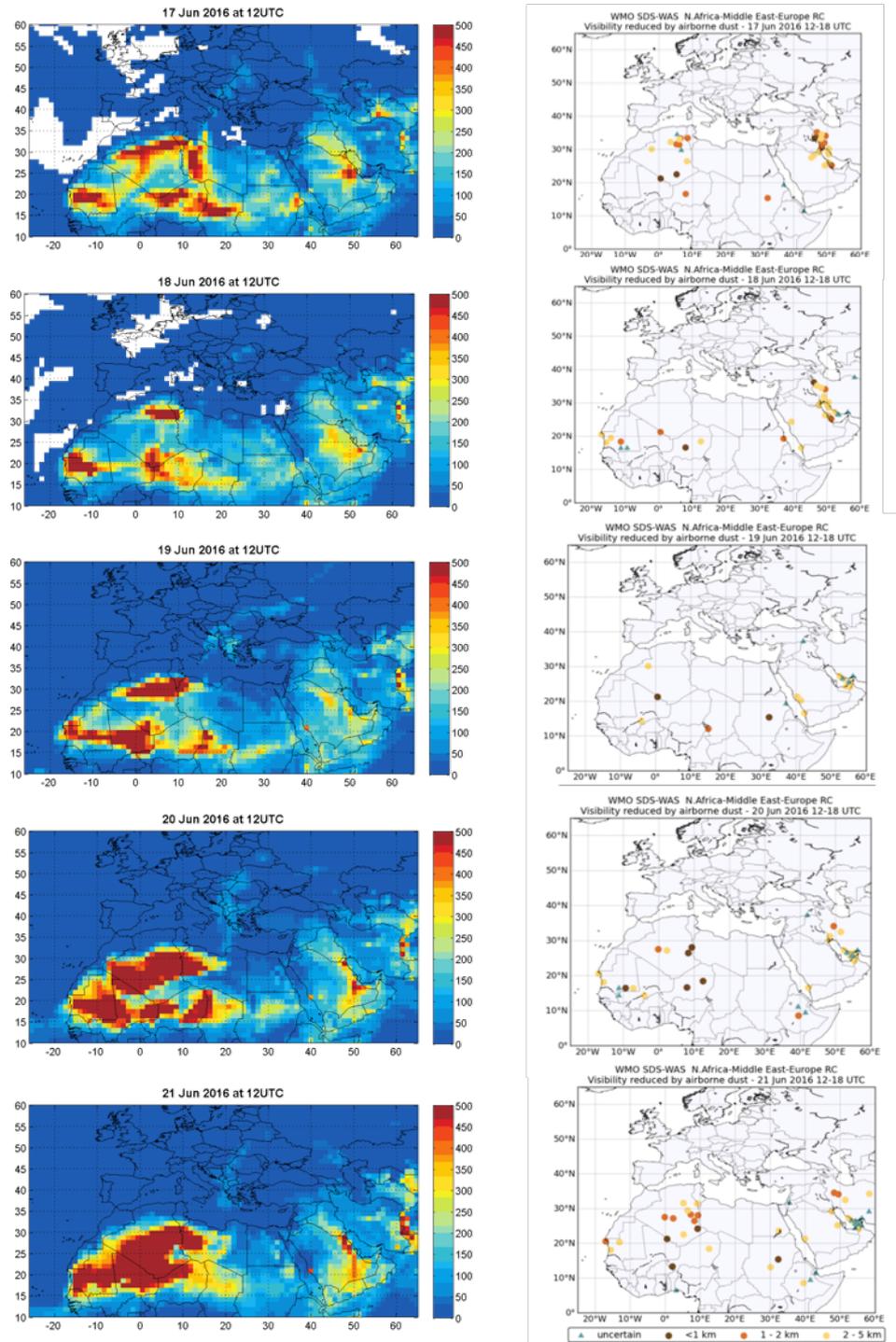


Figure 5.2.4. Dust surface concentration o-suite (right column) and visibility (left column) from METAR or SYNOP stations from SDS-WAS Regional Center during the case analysis from 17th to 21st June 2016. The maps show cases of visibility reduction by sand or dust to less than 5 km reported in METAR or SYNOP bulletins. More than 1,500 stations are checked every 6 hours. Brownish circles indicate stations where 'sand' or 'dust' has been explicitly reported. Triangles indicate stations where the present weather has been reported as 'haze', meaning that the visibility is reduced by particles of unspecified origin.



6. References

Agusti-Panareda, A., *Monitoring upgrades of analysis/forecast system, MACC-III Deliverable D44.04, June 2015.*

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A., and Levin, I.: *Inverse modeling of global and regional CH₄ emissions using SCIAMACHY satellite retrievals, J. Geophys. Res., 114, D22301, doi:10.1029/2009JD012287, 2009.*

Benedetti, A., J.-J. Morcrette, O. Boucher, A. Dethof, R. J. Engelen, M. Fisher, H. Flentjes, N. Huneus, L. Jones, J. W. Kaiser, S. Kinne, A. Mangold, M. Razinger, A. J. Simmons, M. Suttie, and the GEMS-AER team: *Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part II : Data assimilation, J. Geophys. Res., 114, D13205, doi:10.1029/2008JD011115, 2009.*

Boussetta, S., Balsamo, G., Beljaars, A., Agusti-Panareda, A., Calvet, J.-C., Jacobs, C., van den Hurk, B., Viterbo, P., Lafont, S., Dutra, E., Jarlan, L., Balzarolo, M., Papale, D., and van der Werf, G.: *Natural carbon dioxide exchanges in the ECMWF Integrated Forecasting System: implementation and offline validation, J. Geophys. Res.-Atmos., 118, 1–24, doi: 10.1002/jgrd.50488, 2013.*

Braathen, *WMO Arctic Ozone Bulletin No 1/2016, DOI:10.13140/RG.2.1.4929.6403, 2016.*

Cariolle, D. and Teysse re, H.: *A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations, Atmos. Chem. Phys., 7, 2183-2196, doi:10.5194/acp-7-2183-2007, 2007.*

Dee, D. P. and S. Uppala, *Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. Quart. J. Roy. Meteor. Soc., 135, 1830-1841, 2009.*

Deeter, M. N., Emmons, L. K., Edwards, D. P., Gille, J. C., and Drummond, J. R.: *Vertical resolution and information content of CO profiles retrieved by MOPITT, Geophys. Res. Lett., 31, L15112, doi:10.1029/2004GL020235, 2004.*

Deeter, M. N., et al. (2010), *The MOPITT version 4 CO product: Algorithm enhancements, validation, and long-term stability, J. Geophys. Res., 115, D07306, doi:10.1029/2009JD013005.*

Deshler, T., J.L. Mercer, H.G.J. Smit, R. Stubi, G. Levrat, B.J. Johnson, S.J. Oltmans, R. Kivi, A.M. Thompson, J. Witte, J. Davies, F.J. Schmidlin, G. Brothers, T. Sasaki (2008) *Atmospheric comparison of electrochemical cell ozonesondes from different manufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonesondes. J. Geophys. Res. 113, D04307, doi:10.1029/2007JD008975*

Dupuy, E., et al.: *Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), Atmos. Chem. Phys., 9, 287-343, doi:10.5194/acp-9-287-2009, 2009.*

Elbern, H., Schwinger, J., Botchorishvili, R.: *Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: System configuration. Journal of Geophysical Research (Atmospheres) 115, 6302, 2010.*



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

Errera, Q., Daerden, F., Chabrillat, S., Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D., *4D-Var Assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses*, *Atmos. Chem. Phys.*, 8, 6169-6187, 2008.

Errera, Q. and Ménard, R.: *Technical Note: Spectral representation of spatial correlations in variational assimilation with grid point models and application to the belgian assimilation system for chemical observations (BASCOE)*, *Atmos. Chem. Phys. Discuss.*, 12, 16763-16809, doi:10.5194/acpd-12-16763-2012, 2012.

Eskes, H., Huijnen, V., Arola, A., Benedictow, A., Blechschmidt, A.-M., Botek, E., Boucher, O., Bouarar, I., Chabrillat, S., Cuevas, E., Engelen, R., Flentje, H., Gaudel, A., Griesfeller, J., Jones, L., Kapsomenakis, J., Katragkou, E., Kinne, S., Langerock, B., Razinger, M., Richter, A., Schultz, M., Schulz, M., Sudarchikova, N., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: *Validation of reactive gases and aerosols in the MACC global analysis and forecast system*, *Geosci. Model Dev.*, 8, 3523-3543, doi:10.5194/gmd-8-3523-2015, 2015.

Eskes, H.J., V. Huijnen, S. Basart, A. Benedictow, A.-M. Blechschmidt, S. Chabrillat, H. Clark, Y. Christophe, E. Cuevas, H. Flentje, K. M. Hansen, J. Kapsomenakis, B. Langerock, M. Ramonet, A. Richter, M. Schulz, A. Wagner, T. Warneke, C. Zerefos: *Observations characterisation and validation methods document. Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2015SC1_D84.8.1_2016Q2_201603, March 2016. Available from: <http://atmosphere.copernicus.eu/user-support/validation/verification-global-services>*

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: *Tropospheric chemistry in the Integrated Forecasting System of ECMWF*, *Geosci. Model Dev.*, 8, 975-1003, doi:10.5194/gmd-8-975-2015, 2015.

Franco, B., et al., *Retrievals of formaldehyde from ground-based FTIR and MAX-DOAS observations at the Jungfraujoch station and comparisons with GEOS-Chem and IMAGES model simulations*, *Atmos. Meas. Tech.*, 8, 1733-1756, 2015

Gielen, C., Van Roozendaal, M., Hendrick, F., Pinardi, G., Vlemmix, T., De Bock, V., De Backer, H., Fayt, C., Hermans, C., Gillotay, D., and Wang, P.: *A simple and versatile cloud-screening method for MAX-DOAS retrievals*, *Atmos. Meas. Tech.*, 7, 3509-3527, doi:10.5194/amt-7-3509-2014, 2014.

Granier, C. et al.: *Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980–2010 period*. *Climatic Change* (109), 2011

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov A.: *AERONET – a federated instrument network and data archive for aerosol characterization*, *Remote Sens. Environ.*, 66, 1–16, 5529, 5533, 5537, 5544, 1998.

Hommel, R., Eichmann, K.-U., Aschmann, J., Bramstedt, K., Weber, M., von Savigny, C., Richter, A., Rozanov, A., Wittrock, F., Khosrawi, F., Bauer, R., and Burrows, J. P.: *Chemical ozone loss and ozone mini-hole event*



during the Arctic winter 2010/2011 as observed by SCIAMACHY and GOME-2, *Atmos. Chem. Phys.*, **14**, 3247-3276, doi:10.5194/acp-14-3247-2014, 2014.

Huijnen, V., et al.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, *Geosci. Model Dev.*, **3**, 445-473, doi:10.5194/gmd-3-445-2010, 2010.

Inness, A., Blechschmidt, A.-M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., Eskes, H., Flemming, J., Gaudel, A., Hendrick, F., Huijnen, V., Jones, L., Kapsomenakis, J., Katragkou, E., Keppens, A., Langerock, B., de Mazière, M., Melas, D., Parrington, M., Peuch, V. H., Razinger, M., Richter, A., Schultz, M. G., Suttie, M., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS, *Atmos. Chem. Phys.*, **15**, 5275-5303, doi:10.5194/acp-15-5275-2015, 2015.

Janssens-Maenhout, G., Dentener, F., Aardenne, J. V., Monni, S., Pagliari, V., Orlandini, L., Klimont, Z., Kurokawa, J., Akimoto, H., Ohara, T., Wankmueller, R., Battye, B., Grano, D., Zuber, A., and Keating, T.: EDGAR-HTAP: a Harmonized Gridded Air Pollution Emission Dataset Based on National Inventories, JRC68434, EUR report No EUR 25 299–2012, ISBN 978-92-79- 23122-0, ISSN 1831-9424, European Commission Publications Office, Ispra (Italy), 2012.

Jaross, G., Bhartia, P.K., Chen, G., Kowitt, M., Haken, M., Chen, Z., Xu, Ph., Warner, J., Kelly, T. : OMPS Limb Profiler instrument performance assessment, *J. Geophys. Res. Atmos* **119**, 2169-8996, 2014.

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

Kramarova, N. A., Nash, E. R., Newman, P. A., Bhartia, P. K., McPeters, R. D., Rault, D. F., Sefstor, C. J., Xu, P. Q., and Labow, G. J.: Measuring the Antarctic ozone hole with the new Ozone Mapping and Profiler Suite (OMPS), *Atmos. Chem. Phys.*, **14**, 2353-2361, doi:10.5194/acp-14-2353-2014, 2014.

Lahoz, W. A., Errera, Q., Viscardi, S., and Manney G. L., The 2009 stratospheric major warming described from synergistic use of BASCOE water vapour analyses and MLS observations, *Atmos. Chem. Phys.* **11**, 4689-4703, 2011

Lambert, A, et al., Aura Microwave Limb Sounder Version 3.4 Level-2 near real-time data user guide, <http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/MLS/documents/NRT-user-guide-v34.pdf>

Langerock, B., De Mazière, M., Hendrick, F., Vigouroux, C., Desmet, F., Dils, B., and Niemeijer, S.: Description of algorithms for co-locating and comparing gridded model data with remote-sensing observations, *Geosci. Model Dev.*, **8**, 911-921, doi:10.5194/gmd-8-911-2015, 2015.

Lefever, K., van der A, R., Baier, F., Christophe, Y., Errera, Q., Eskes, H., Flemming, J., Inness, A., Jones, L., Lambert, J.-C., Langerock, B., Schultz, M. G., Stein, O., Wagner, A., and Chabrillat, S.: Copernicus stratospheric ozone service, 2009–2012: validation, system intercomparison and roles of input data sets, *Atmos. Chem. Phys.*, **15**, 2269-2293, doi:10.5194/acp-15-2269-2015, 2015.

Liu, Z., et al., Exploring the missing source of glyoxal (CHOCHO) over China, *Geophys. Res. Lett.*, **39**, L10812, doi: 10.1029/2012GL051645, 2012



- Massart, S., Flemming, J., Cariolle, D., Jones, L., High resolution CO tracer forecasts, MACC-III Deliverable D22.04, May 2015, available from <http://www.gmes-atmosphere.eu/documents/macciii/deliverables/grq>
- Morcrette, J.-J., O. Boucher, L. Jones, D. Salmond, P. Bechtold, A. Beljaars, A. Benedetti, A. Bonet, J. W. Kaiser, M. Razinger, M. Schulz, S. Serrar, A. J. Simmons, M. Sofiev, M. Suttie, A. M. Tompkins, and A. Untch: Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part I: Forward modelling, *J. Geophys. Res.*, 114, D06206, doi:10.1029/2008JD011235, 2009.
- Richter, A., Burrows, J. P., Nüß, H., Granier, C., Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, *Nature*, 437, 129-132, doi: 10.1038/nature04092, 2005
- Richter, A., Begoin, M., Hilboll, A., and Burrows, J. P.: An improved NO₂ retrieval for the GOME-2 satellite instrument, *Atmos. Meas. Tech.*, 4, 1147-1159, doi:10.5194/amt-4-1147-2011, 2011
- Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrou, T., Müller, J.-F., Kuhn, U., Stefani, P., and Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, *Atmos. Chem. Phys.*, 14, 9317-9341, doi:10.5194/acp-14-9317-2014, 2014.
- Smit, H.G.J., W. Straeter, B.J. Johnson, S.J. Oltmans, J. Davies, D.W. Tarasick, B. Hoegger, R. Stubi, F.J. Schmidlin, T. Northam, A.M. Thompson, J.C. Witte, I. Boyd: Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), *J. Geophys. Res.* 112, D19306, doi:10.1029/2006JD007308, 2007.
- Solomon, S., Haskins, J., Ivy, D. J. and Min, F.: Fundamental differences between Arctic and Antarctic ozone depletion, *PNAS* 2014 111 (17) 6220-6225, doi:10.1073/pnas.1319307111, 2014.
- Stavrou, T., First space-based derivation of the global atmospheric methanol fluxes, *Atm. Chem. Phys.*, 11, 4873-4898, 2013.
- Strahan, S.E., A.R. Douglass, and P.A. Newman, The contributions of chemistry and transport to low arctic ozone in March 2011 derived from Aura MLS observations, *J. Geophys. Res. Atmos.*, 118, 1563–1576, doi:10.1002/jgrd.50181, 2013.
- Taha, G.; Jaross, G. R.; Bhartia, P. K.: Validation of OMPS LP Ozone Profiles Version 2.0 with MLS, Ozone Sondes and Lidar Measurements, American Geophysical Union, Fall Meeting 2014, abstract #A33J-3322, 2014.
- Taylor, K.E.: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, 106, 7183-7192, 2001.
- van der A, R. J. , M. A. F. Allaart, and H. J. Eskes, Multi sensor reanalysis of total ozone, *Atmos. Chem. Phys.*, 10, 11277–11294, doi:10.5194/acp-10-11277-2010, www.atmos-chem-phys.net/10/11277/2010/, 2010
- van der A, R., M. Allaart, H. Eskes, K. Lefever, Validation report of the MACC 30-year multi-sensor reanalysis of ozone columns Period 1979-2008, MACC-II report, Jan 2013, [MACCII_VAL_DEL_D_83.3_OzoneMSRv1_20130130.docx/pdf](#).



van der A, R. J., Allaart, M. A. F., and Eskes, H. J.: *Extended and refined multi sensor reanalysis of total ozone for the period 1970–2012*, *Atmos. Meas. Tech.*, **8**, 3021–3035, doi:10.5194/amt-8-3021-2015, 2015.

Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: *GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?*, *Atmos. Chem. Phys.*, **10**, 10145–10160, doi:10.5194/acp-10-10145-2010, 2010

Wittrock, F., A. Richter, H. Oetjen, J. P. Burrows, M. Kanakidou, S. Myriokefalitakis, R. Volkamer, S. Beirle, U. Platt, and T. Wagner, *Simultaneous global observations of glyoxal and formaldehyde from space*, *Geophys. Res. Lett.*, **33**, L16804, doi:10.1029/2006GL026310, 2006

WMO (2010), *Guidelines for the Measurement of Atmospheric Carbon Monoxide*, GAW Report No. 192, World Meteorological Organization, Geneva, Switzerland, 2010.

WMO (2013), *Guidelines for the Continuous Measurements of Ozone in the Troposphere*, GAW Report No. 209, World Meteorological Organization, Geneva, Switzerland, 2013.



Annex 1: Acknowledgements

Listed below are the authors contributing to the sections in this report. The authors contributing to the model description are also provided, as well as acknowledgements to the validation datasets.

Tropospheric reactive gases reactive gases

Annette Wagner, MPG (editor, O3 sondes, GAW data)
Hannah Clark, Valerie Thouret, CNRS-LA (IAGOS)
Harald Flentje, DWD (O3 sondes, GAW data)
Anne Blechschmidt and Andreas Richter, IUB Bremen (GOME-2 NO2, HCHO)
John Kapsomenakis, Christos Zerefos, AA (ESRL)
N. Sudarchikova, satellite IR observations (MPG)
Kaj Hansen, Ulas Im, AU (Arctic theme)
Bavo Langerock, BIRA (NDACC)

Tropospheric aerosol

Michael Schulz, MetNo (editor, AeroCom, Aeronet)
Anna Benedictow, Jan Griesfeller, MetNo (AeroCom, Aeronet)
Emilio Cuevas, Carlos Camino, AEMET (AERONET, MODIS/Aqua-DeepBlue)
Enric Terradellas and Francesco Benincasa, AEMET/SDS WAS RC NAMEE
José María Baldasano and Sara Basart, BSC-CNS (BSC-DREAM8b)
Harald Flentje, DWD (Backscatter profiles)

Stratospheric reactive gases

Yves Christophe, BIRA (editor, model intercomparisons)
Simon Chabrilat, BIRA (model intercomparisons)
Annette Wagner, MPI-M (O3 sondes)
Bavo Langerock, BIRA (NDACC FTIR, MWR, UVVIS DOAS, LIDAR)
Anne Blechschmidt and Andreas Richter, IUB-UB Bremen (SCIAMACHY/GOME-2 NO2)

Greenhouse gases

Michel Ramonet, IPSL (ICOS)
Olivier Jossoud and Leonard Rivier, LSCE (ICOS)
Thorsten Warneke, UBC (TCCON)
Bavo Langerock, BIRA (TCCON)

Reactive gases and aerosol modeling

Johannes Flemming (ECMWF), Antje Inness (ECMWF), Angela Benedetti (ECMWF), Sebastien Massart (ECMWF), Anna Agusti-Panareda (ECMWF), Johannes Kaiser (KCL/MPIC/ECMWF), Samuel Remy (LMD), Olivier Boucher (LMD), Vincent Huijnen (KNMI), Richard Engelen (ECMWF)



Acknowledgements for the validation datasets used

We wish to acknowledge the provision of NRT GAW observational data by: Institute of Atmospheric Sciences and Climate (ISAC) of the Italian National Research Council (CNR), South African Weather Service, National Centre for Atmospheric Science (NCAS, Cape Verde), National Air Pollution Monitoring Network (NABEL) (Federal Office for the Environment FOEN and Swiss Federal Laboratories for Materials Testing and Research EMPA), Atmospheric Environment Division Global Environment and Marine Department Japan Meteorological Agency, Chinese Academy of Meteorological Sciences (CAMS), Alfred Wegener Institut, Umweltbundesamt (Austria), National Meteorological Service (Argentina), Umweltbundesamt (UBA, Germany)

We are grateful to the numerous operators of the Aeronet network and to the central data processing facility at NASA Goddard Space Flight Center for providing the NRT sun photometer data, especially Ilya Slutsker and Brent Holben for sending the data.

We wish to acknowledge the provision of ozone sonde data by the World Ozone and Ultraviolet Radiation Data Centre established at EC in Toronto (<http://woudc.org>), by the Data Host Facility of the Network for the Detection of Atmospheric Composition Change established at NOAA (<http://ndacc.org>), by the Norwegian Institute for Air Research and by the National Aeronautics and Space Administration (NASA).

We wish to thank the NDACC investigators for the provision of observations at Ny Alesund, Bern, Jungfraujoch, Izaña, Xianghe, Harestua, Reunion Mado, Uccle, Hohenpeissen, Mauna Loa, Lauder and Haute Provence. Special thanks to the data providers at Ny-Alesund, Bern and Uccle for their support of the Arctic ozone depletion case study.

The authors acknowledge the help of Geir Braathen (WMO) for his support of the Arctic ozone depletion case study.

The authors acknowledge the NOAA Earth System Research Laboratory (ESRL) Global Monitoring Division (GMD) for the provision of ground based ozone concentrations.

The MOPITT CO data were obtained from the NASA Langley Research Center ASDC. We acknowledge the LATMOS IASI group for providing IASI CO data.

SCIAMACHY lv1 radiances were provided to IUP-UB by ESA through DLR/DFD.

GOME-2 lv1 radiances were provided to IUP-UB by EUMETSAT.

Index database and especially to the contributors to that database which provided the measurements for comparisons.

The authors acknowledge Environment and Climate Change Canada for the provision of Alert ozone data and Sara Crepinsek – NOAA for the provision of Tiksi ozone data. Surface ozone data from the Villum Research Station, Station Nord (VRS) was financially supported by “The Danish Environmental Protection Agency” with means from the MIKA/DANCEA funds for Environmental Support to the Arctic Region. The Villum Foundation is acknowledged for the large grant making it possible to build VRS in North Greenland.



We acknowledge the National Aeronautics and Space Administration (NASA), USA for providing the <http://osirus.usask.ca/OMPS> limb sounder data (<http://npp.gsfc.nasa.gov/omps.html>) and the Aura-MLS offline data (<http://mls.jpl.nasa.gov/index-eos-mls.php>).

The TCCON site at Orleans is operated by the University of Bremen and the RAMCES team at LSCE (Gif-sur-Yvette, France). The TCCON site at Bialystok is operated by the University of Bremen. Funding for the two sites was provided by the EU-project ICOS-INWIRE and the University of Bremen. The TCCON site at Réunion is operated by BIRA-IASB, in cooperation with UReunion and is funded by BELSPO in the framework of the Belgian ICOS program.

We acknowledge the provision of CO₂/CH₄ data from SNO-RAMCES/ICOS network coordinated by LSCE/OVSQ (CEA-CNRS-UVSQ, Université Paris-Saclay), as well as Laboratorio de Física de la Atmósfera (UMSA, Bolivia), Environmental Chemical Processes Laboratory (ECPL/UoC, Greece), Station Géophysique de LAMTO, Ivory Coast), C-CAPS/NUIG/EPA (Ireland), BIRA-IASB (Belgium) and the following research institutes in France: LaMP/OPGC, P2OA/LA/OMP, OPE/ANDRA, OHP/PYTHEAS, OPAR/LACY/OSUR, UMR EcoFoG, IPEV, IRD.



ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-atmosphere.eu