



Copernicus Atmosphere Monitoring Service



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

**System evolution
and performance statistics
Status up to 1 June 2016**

Issued by: KNMI

Date: 21/09/2016

REF.: CAMS84_2015SC1_D.84.1.4_201609





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.

System evolution and performance statistics
Status up to 1 June 2016

Editors:

V. Huijnen (KNMI), 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), Y. Christophe (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 (CAMS-84).

Date:

9 September 2016.

Status:

Final

Citation:

Validation report of the CAMS near-real-time global atmospheric composition service.
System evolution and performance statistics; Status up to 1 June 2016.

V. Huijnen, 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_D.84.1.4_2016Q3_201609, September 2016.

Available at:

http://atmosphere.copernicus.eu/quarterly_validation_reports



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 June 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 delivery for the period March-May 2016 was very good, with 98% of the forecasts delivered before 22:00 UTC. Since December 2012 on average 96% of the forecasts were delivered on time.

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 around $\pm 10\%$ 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 March-May 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. Gradients in the UTLS are better represented by the control run.

The o-suite shows an overestimation of surface ozone for Europe during March and May 2016 with MNMBs of up to 10%. For USA the o-suite shows an overestimation of surface ozone up to 15%. For Asia, the o-suite shows a good agreement with surface ozone MNMBs within 12%. For the tropics, the surface ozone is overestimated between 15% and 30%. For Southern hemispheric stations, the o-suite shows a good correspondence with observed surface ozone mixing ratios. The data assimilation corrects the negative offset visible for the control run.

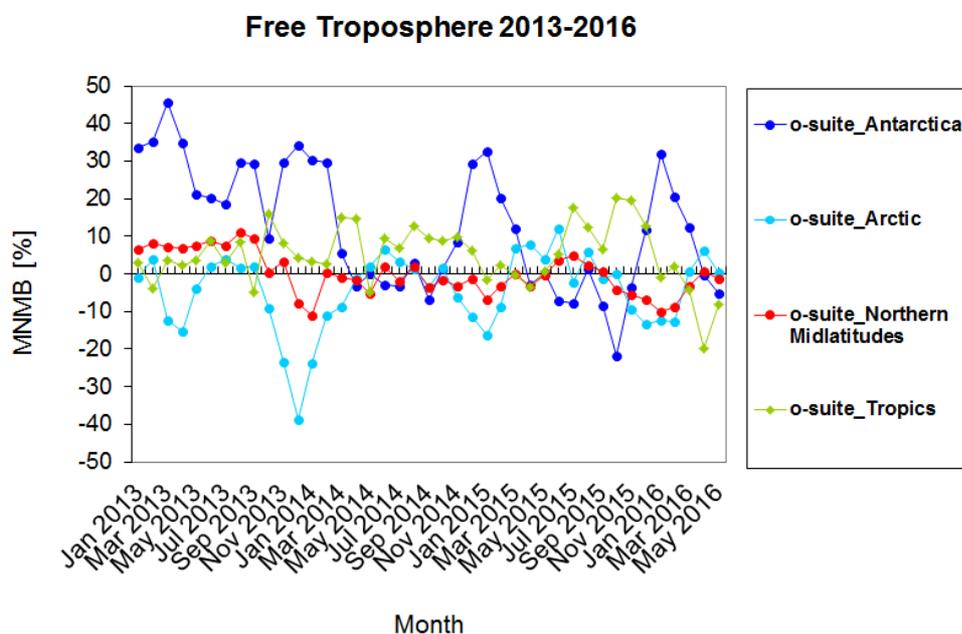


Figure S1: Time series of MNMB of ozone in the o-suite, compared against ozone sondes, averaged over different latitude bands.

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. 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 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, although it must be noted that significant differences between MOPITT and IASI are observed. Also compared to IAGOS aircraft observations over Europe and Asia the modelled free tropospheric CO mixing ratios are well in line with measurements, while the o-suite shows a negative bias in the boundary layer. This is confirmed with comparison against GAW surface observations (MNMBs between -6%

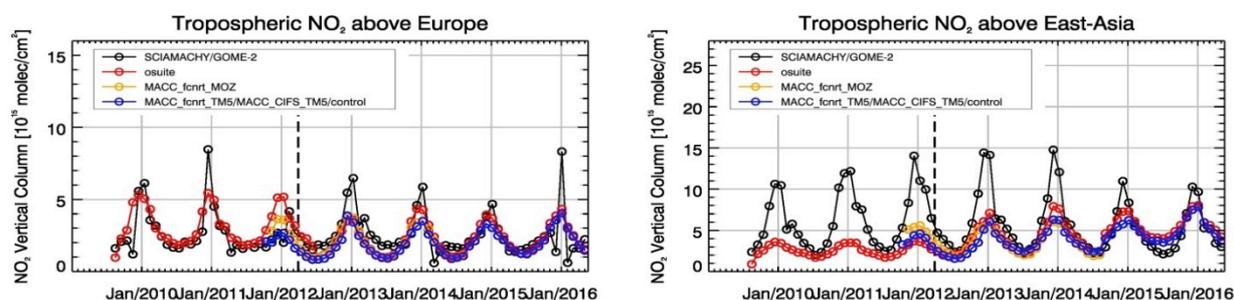


Figure S2: Time series of tropospheric NO₂ columns from SCIAMACHY (up to March 2012), GOME-2 (from April 2012 onwards) compared to model results for Europe and East-Asia. The o-suite is in red, control is in blue (before Sept 2014 blue and yellow represent older model configurations).

and -17%). During the fire season over Alaska and Siberia negative biases are only up to -5%.

Especially the control run shows an overestimation of CO total columns in the tropics and SH. 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, 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 winter 2015/2016 of +25%, measured as modified normalized mean bias against daily Aeronet sun photometer data. The +3 day forecasted aerosol distributions, since July

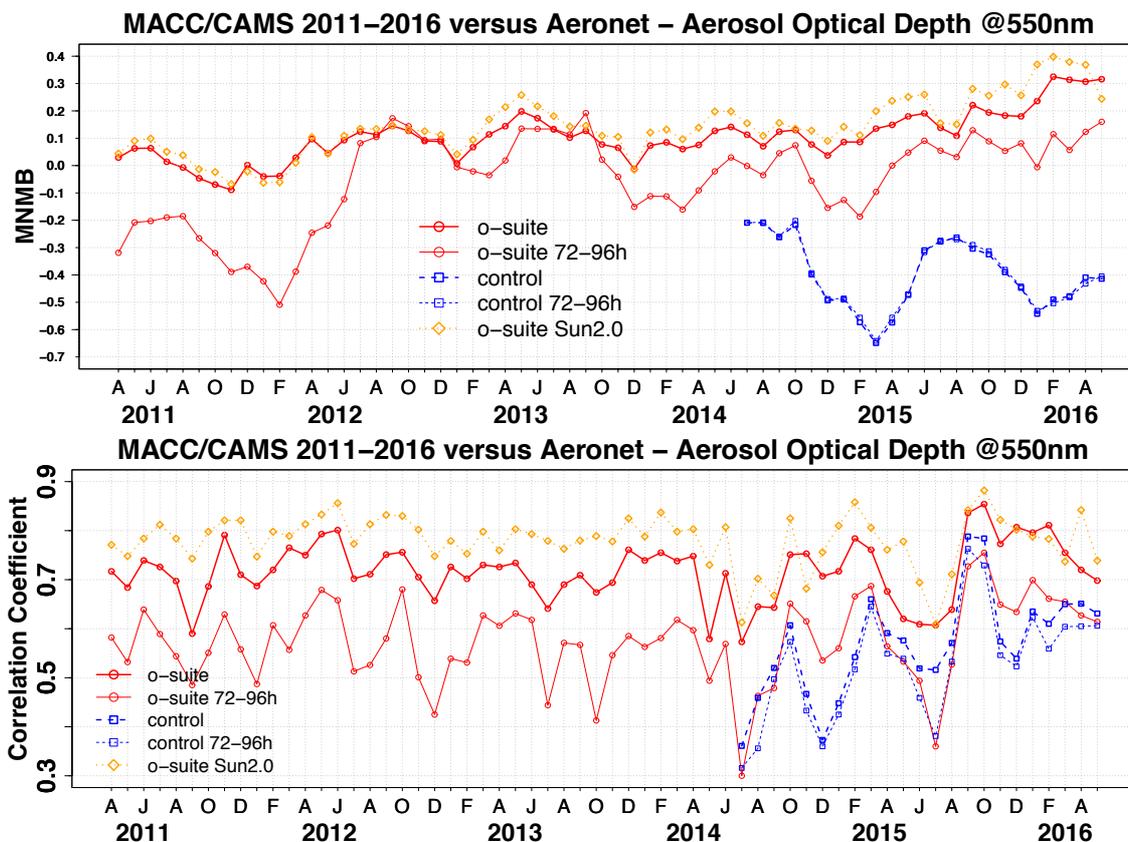


Figure S3. Aerosol optical depth at 550nm in IFS model simulations for April 2011 - May 2016 against daily matching Aeronet NRT level 1.5 and level 2.0 data a) Modified normalized mean bias (MNMB); o-suite (thick red curve); o-suite at last forecast day (light red curve); Control (blue dashed); Control 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.

2012, show 10-30% less aerosol optical depth (AOD) than those from the initial day, as shown all in Figure S3a. 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 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 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 47% of global mean AOD. Correlation is lower in autumn and winter.

An evaluation of the PM₁₀ surface concentration against a climatological average (2000-2009) at 155 remote sites in North America and Europe indicates overestimations of a factor 2 in regions closer to the coast, possibly due to high simulated sea salt concentrations, confirmed by the comparison to Airbase data in the Mediterranean, see below. PM₁₀ concentrations more inland exhibit a negative bias of 20-40%.

Modelled dust optical depth (DOD) over twelve different sub-regions of North Africa, Middle East and southern Europe was compared with AERONET sun photometers observations of coarse and total aerosol optical depths in springtime. The o-suite is able to reproduce the most important dust events over the region. The quality of the CAMS o-suite is very similar to the multi-model Median ensemble established at the WMO-SDS Regional Center but varies from region to region. In terms of correlation, o-suite improves the observed variability over Middle East and reduces the correlation over the Sahara during spring. Over the long-range transport regions, maximum dust activity is observed in Central and Eastern Mediterranean. Poor correlations observed in the Iberian Peninsula are mainly associated to a low number of events and low AOD values during all the period of study over this area.

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 assessed, using ICOS surface (15 sites) and TCCON total column (3 sites) measurements. Some surface sites appear to be more challenging due to the difficulty for the model either to reproduce local transport (e.g. mountain or coastal sites), or local emissions. Each outlier is analysed in detail in order to understand the reason for the higher differences. The station of La Réunion, which is hosting both ICOS and TCCON measurements, remains a challenging site. The model overestimates the synoptic variability of total columns for CO₂ but not at the surface level, whereas the reverse is observed for CH₄. More investigation will be needed for this station.

At the annual scale, both surface and total column measurements show a significant seasonal variability of CO₂ biases in the northern hemisphere, with maximum positive biases observed in winter/spring and a better agreement in summertime. The amplitude of the synoptic spikes is relatively well simulated at most European stations.

For CH₄ the underestimation of the seasonal amplitude for the TCCON European sites, is not observed at the surface stations. For the surface sites the model underestimates the concentrations by 10-20 ppb at most sites, but we notice an improvement with the new experiment started in 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 (4 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.

For the March 2016 – May 2016 period the ozone depletion events are seen in the surface measurements at the three coastal sites (Barrow, Alert and the Villum Research station) and the statistical parameters for these sites are offset by the events. The mean surface ozone concentrations as well as the ozone variability over Summit are well reproduced by the model. In the free troposphere the o-suite shows lower MNMBs than the control run.

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 (MNMB values down to -20%) and Eastern Mediterranean (MNMBs≈0). Temporal correlation coefficients between simulated and observed surface ozone mixing ratios range generally between 0.4 and 0.85.

The CAMS aerosol model run performance is evaluated over the Mediterranean using AOD observations from 37 AERONET sites. From March to May 2016, CAMS o-suite reproduces better the daily variability of AERONET observations than the control experiment, which tends to underestimate the background levels over Southern European sites (more influenced by urban/industrial aerosol sources). Observations at 19 Airbase PM background sites show that the CAMS experiments reproduce the daily variability of the most intense aerosol events observed, which are associated with long-range desert dust transport. However, the magnitude of CAMS simulated PM (control and o-suite) is still overestimated, being specially high over maritime stations.

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%. The o-suite shows a positive bias in surface ozone concentrations in Europe, with MNMBs for GAW and ESRL stations ranging between 0% and 10% between March and May 2016, and a small positive bias over North American stations. The o-suite underestimates surface CO concentration in Europe and Asia with MNMBs with respect to GAW of around -15%. Evaluation of the PM₁₀ surface concentration against a climatology in western coastal regions of North America and Europe, against inland vertical profiles of aerosol backscatter coefficient in Germany and NRT-Airbase data at coastal sites in Spain indicate a significant overestimation of sea salt concentrations and thus too high PM₁₀ boundary concentrations. These are difficult to quantify further



because short lived coarse sea salt fractions may or may not influence regional models, depending on their interpretation of the PM₁₀ boundary concentrations and because high sea salt episodes are sporadic as dust storms.

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, <http://www.ndacc.org>); 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%). In the Arctic and the Tropics, the o-suite total O₃ columns give results nearly identical to the reference TM3DAM system (a slight underestimation of 2-3 DU is noted in the Tropics but this is well within the observational uncertainties).

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 1 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 has decreased during summer months but since September 2015, 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 (Figure 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.

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 been evaluated but the results are not presented here.

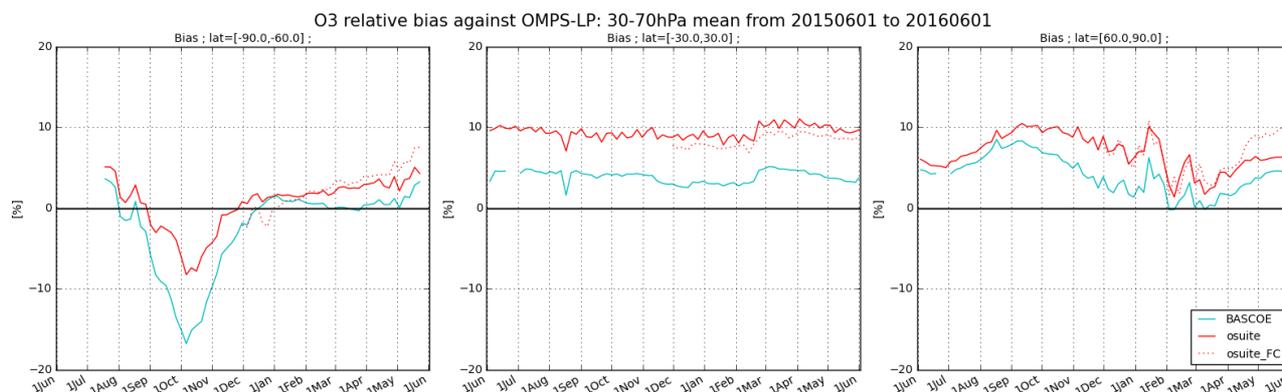


Figure S4: Time series of the normalized mean bias (%) between ozone from o-suite analyses (red, solid) or 4th day forecasts (red, dotted), or BASCOE analyses (cyan) and OMPS-LP satellite observations, in the middle stratosphere (30-70hPa averages).

Events

The Far East of Russia was affected by many *large fires* in May 2016. Analysis of IASI data for one of these fires shows a clear plume of CO, increasing with time, over this region and south-east transport towards the North Pacific Ocean. The CAMS forecasts captured the location and transport of the CO emissions, although the modelled CO values are underestimated over the easternmost area of Russia and over the Ocean while overestimated over the Baikal Lake area with the control run very similar to the o-suite. Also modeled NO₂ and HCHO columns due to fire emissions are over-estimated compared to the GOME-2 satellite data.

A *stratospheric ozone depletion event* developed over the Arctic during the focus period of the previous report (December 2015 - February 2016) and reached Northern Europe in February. During this depletion episode the o-suite analyses overestimated stratospheric ozone above Ny-Alesund as observed by ozone sondes and Microwave Radiometers, but according to ozone sondes this disagreement became much less severe after 20 January. Above the Northern European stations of Bern and Uccle the biases were much smaller - even when the polar vortex reached these latitudes (see Figure S5). From a seasonally averaged point of view (DJF 2015-2016), the relative bias between MWR instruments and the o-suite reached 14.5% above Ny-Alesund but only 0.7% above Bern.

Around the 9th May, a large *plume of dust* was exported from Algeria covering the area around the Central Mediterranean, moving later through the Eastern Mediterranean. The o-suite is able to timely reproduce the spatial distribution of the dust plume as observed by MODIS over the Mediterranean. Also the comparison of the modelled surface concentration shows how the model fits fairly well the reduction in visibility.



Table of Contents

1	Introduction	11
2	System summary and model background information.....	14
2.1	System based on the ECMWF IFS model	14
2.2	Evolution of the IFS-based system	18
2.3	Other systems.....	18
2.4	CAMS products.....	20
2.5	Availability and timing of CAMS products.....	20
3	Validation results for reactive gases and aerosol.....	21
3.1	Tropospheric Ozone	21
3.2	Tropospheric nitrogen dioxide.....	37
3.3	Carbon monoxide	42
3.4	Formaldehyde	58
3.5	Aerosol.....	62
3.6	Stratospheric ozone	84
3.7	Stratospheric NO ₂	90
4	Validation results for greenhouse gases.....	92
4.1	CH ₄ and CO ₂ validation against ICOS observations	92
4.2	CH ₄ and CO ₂ validation against TCCON observations	97
5	Events	101
5.1	Fire case in the Far East of Russia in May 2016	101
5.2	Ozone depletion above the Arctic and Northern Europe, Jan-Feb. 2016	102
5.3	A dust event over the Western Mediterranean in May 2016.....	107
6	References	111
	Annex 1: Acknowledgements.....	115



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).



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

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 1.2: 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_fcirt_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	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 Sept. 2014 onwards. CO: Top three rows; O₃ columns and profiles: rows 4, 8-11; SO₂: rows 5-7, Aerosol Optical Depth: rows 12-13. The SO₂ concentrations are only used above a threshold, typically related to volcano eruptions.

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=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.

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 - SO2 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 NO2 and O3 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

2.2 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.

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 NOA) 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° x 0.5°. 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

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.

Between December 2012 and February 2015 on average about 97% of the forecasts were delivered on time. For the period September November 2015, 95% of the forecasts were delivered before 21.14.

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



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	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 May 2015 to May 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 the MNMB is within the range -20 to +30%, for all months, see Fig. 3.1.1-3.1.3. Over the Arctic, the o-suite shows slightly positive MNMBs during summer and spring (MNMBs up to 12%), while during the winter season the MNMBs gets negative, with MNMBs mostly within -13%, Fig. 3.6.1. Especially for the last three months (March-May 2016) the o-suite shows lower MNMBs than the control run. Over the NH mid-latitudes MNMBs for the o-suite are on average close to zero all year round (maxima are -10% to +5%), which is a clear improvement compared to the control run, which shows larger positive MNMBs (up to $\pm 13\%$) during this period. MNMBs are larger ($\pm 30\%$) 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 between -20% and 25% for the control run.

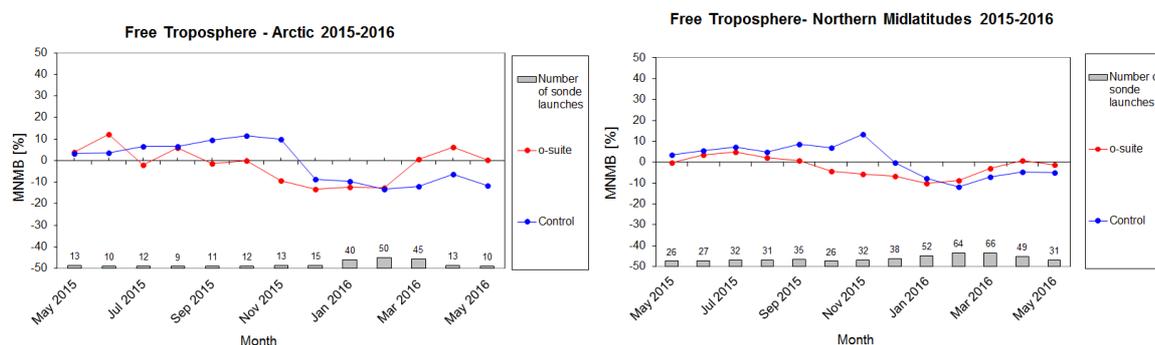


Figure 3.1.1: MNMBs (%) of ozone in the free troposphere (between 750 and 300 hPa) from the IFS model runs against aggregated sonde data over the Arctic (left) and the Northern midlatitudes (right). The numbers indicate the amount of individual number of sondes.

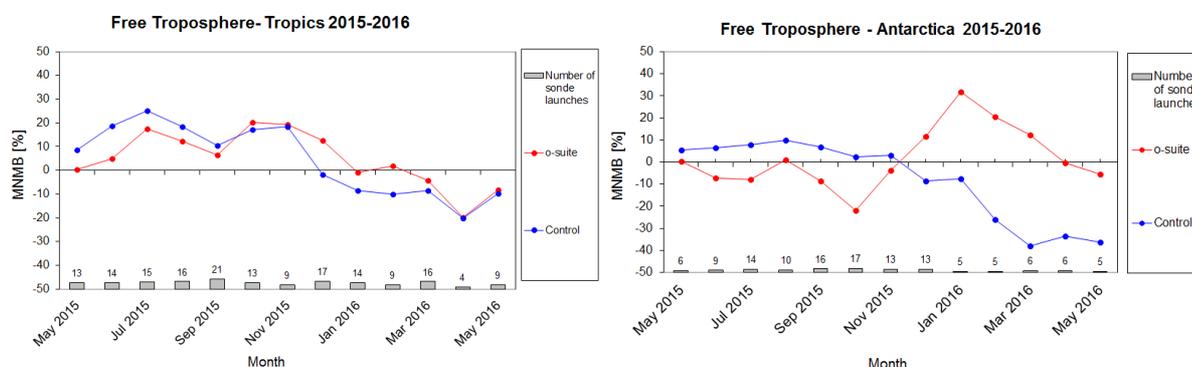


Figure 3.1.2: MNMBs (%) of ozone in the free troposphere (between 750 and 200 hPa (Tropics) / 300 hPa) from the IFS model runs against aggregated sonde data over the Tropics (left) and Antarctica (right). The numbers indicate the amount of individual number of sondes.

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 March 2015 to May 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.

With the whole fleet of 6 aircraft, operating fully over the three month period, we can expect a total of about 1260 flights. The actual number of profiles within the period was 799 (400 flights) giving a performance of 31%. The actual number of profiles with usable data was 261 (65% of the total possible). These flights are shown in Fig. 3.1.3 (top). Fifty eight percent (58%) (466 profiles) of the operational flights had usable measurements of ozone and 60% of flights had usable CO. Delivering these O₃ and CO data are an aircraft from China Airlines based in Taipei and an aircraft operated by Air France based in Paris. This report therefore displays profiles recorded by these aircraft, covering mainly the routes served by Air France to North America and West Africa 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).

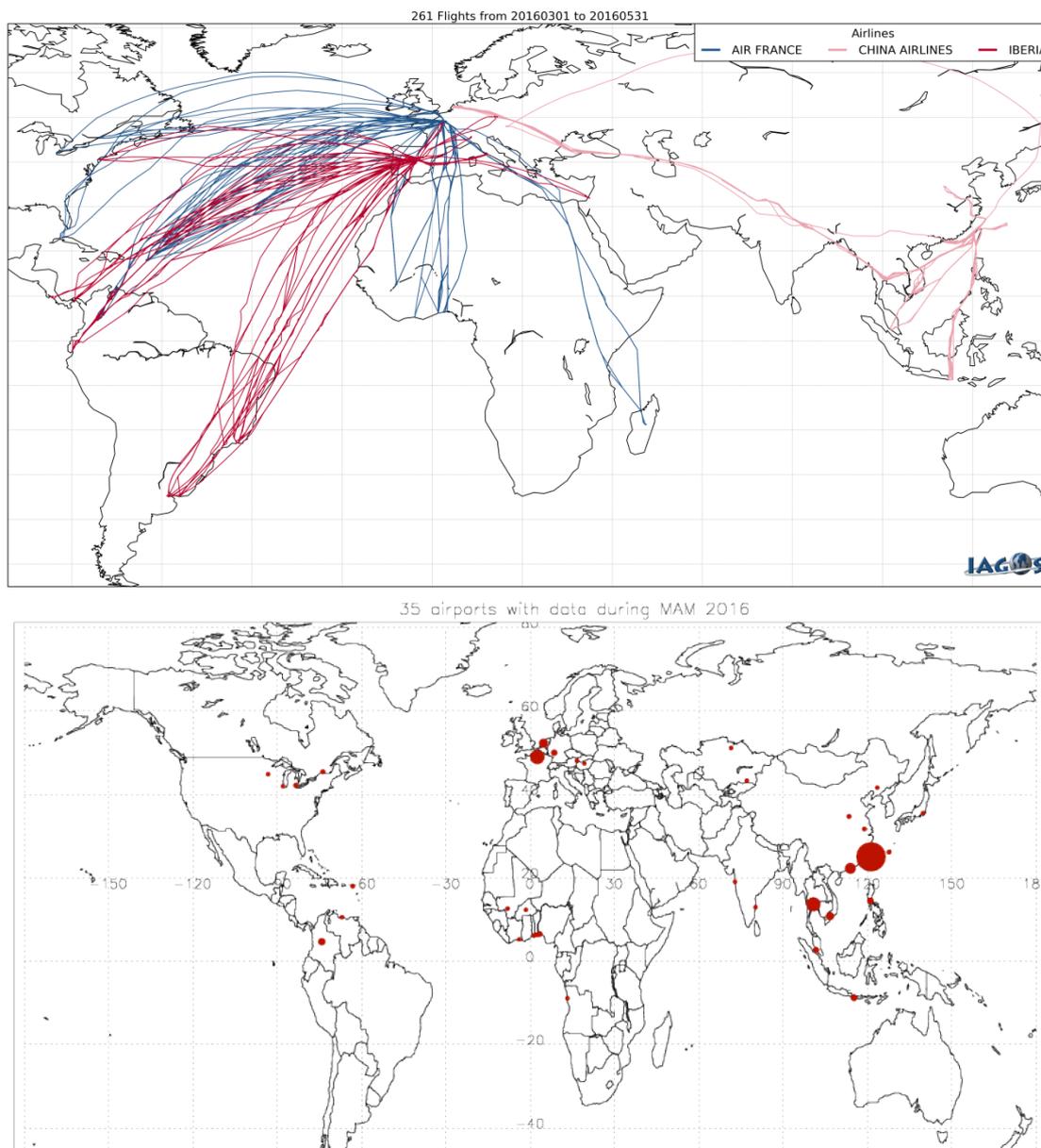


Figure 3.1.3: Map of the flights (top) and the visited airports (bottom) during the period December – February 2015/6, by the IAGOS equipped aircraft. The size of the plotting circle represents the number of profiles available.

Figure 3.1.4 takes advantage of the Air France aircraft operating from Paris and the China Airlines aircraft operating from Taipei. Generally the observations match better with the models at the Asian airports than at the European airports, however in April in the surface and boundary layers over Taipei the ozone is overestimated by the two runs.

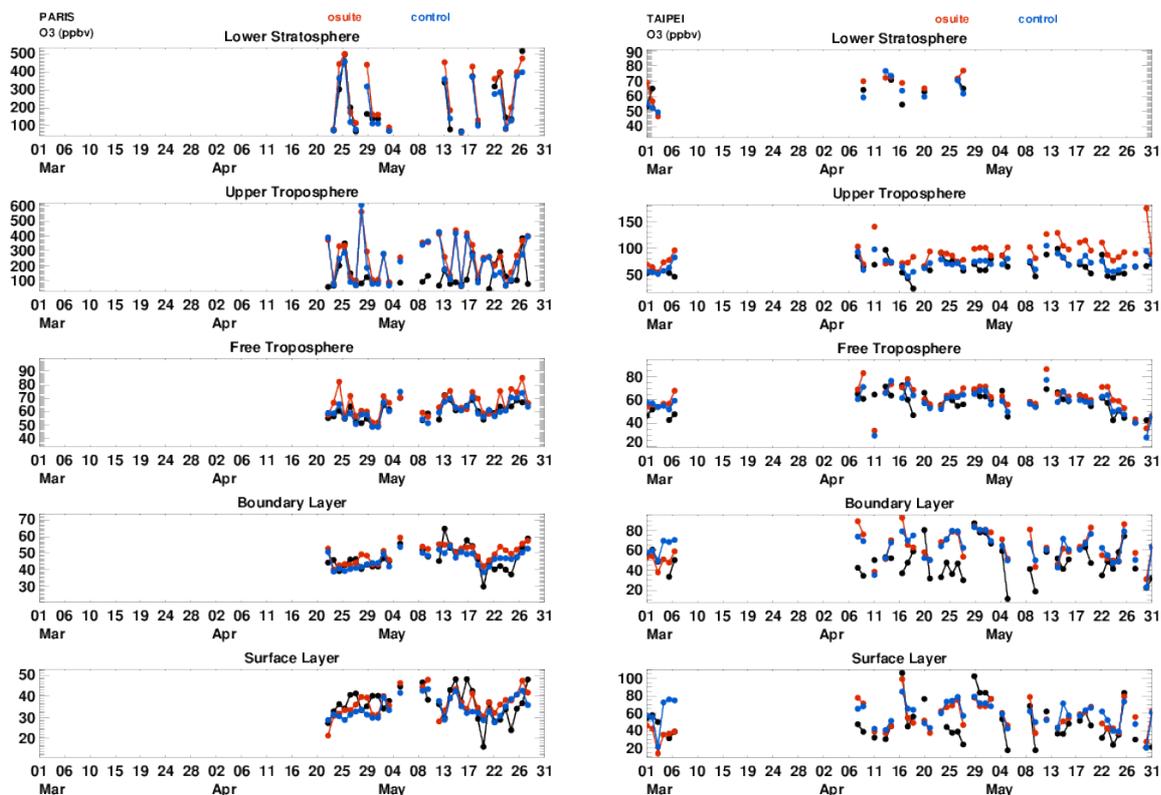


Figure 3.1.4: Time series of daily mean ozone over Paris (left) and Taipei (right) during March, April, May 2016 for 5 layers, Surface, Boundary layer, Free Troposphere, Upper Troposphere and Lower Stratosphere.

Europe

The examples in Fig. 3.1.5 show profiles over Amsterdam, Paris and Frankfurt. On 12th April there are profiles at Paris by Air France and at Amsterdam by China Airlines. There is a striking similarity between the two profiles measured by the two different aircraft and between the model output in the two locations. The concentrations of ozone in the free troposphere are around 50ppbv, and begin to increase at around 8000m. In both cases the control better captures the gradient of ozone in the UTLS, and the o-suite places the tropopause a little too low in altitude. Also in both cases the control does quite well but the o-suite overestimates the amount of ozone in the troposphere. Similar results are seen at Frankfurt on 6th April and at Paris on 10th 18th May.

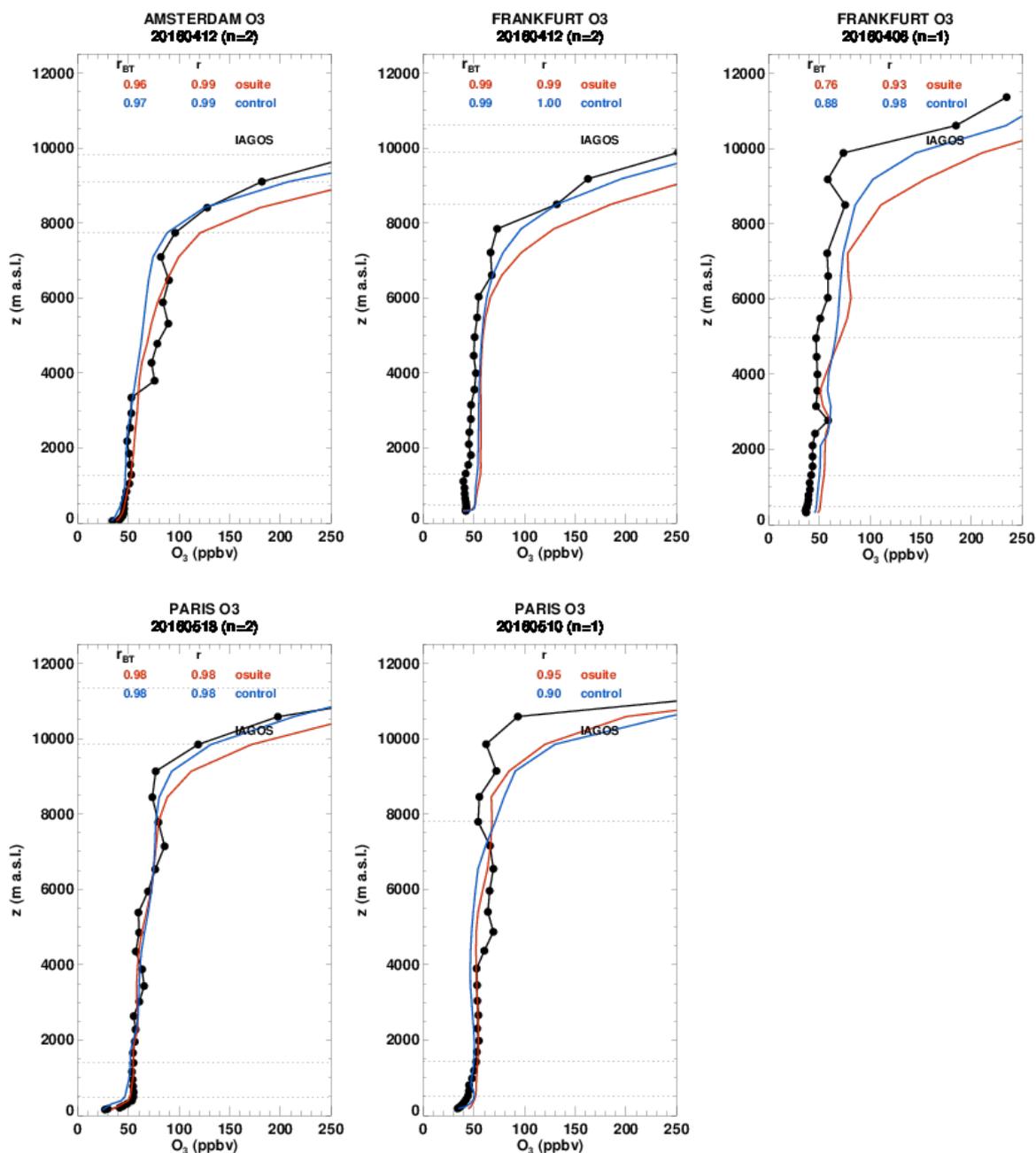


Figure 3.1.5: Selection of daily profiles of ozone from IAGOS (black) and the o-suite (red) and control (blue) over Europe (Paris, Frankfurt, Amsterdam) over the period March April May 2016.

Asia

As the timeseries in Fig. 3.1.4 also showed, the profiles in Fig. 3.1.6 for Taipei illustrate that the o-suite slightly overestimates ozone in the surface layer and boundary layer. This is the opposite of that observed in DJF 2015/6. In the free troposphere the runs compare well with the observations and in the UTLS there is a slight tendency to overestimate ozone. Behaviour at Hong Kong and Kuala Lumpur also follows this pattern. The scatter plots for Taipei and Hong Kong (Fig 3.1.6) also show that at the



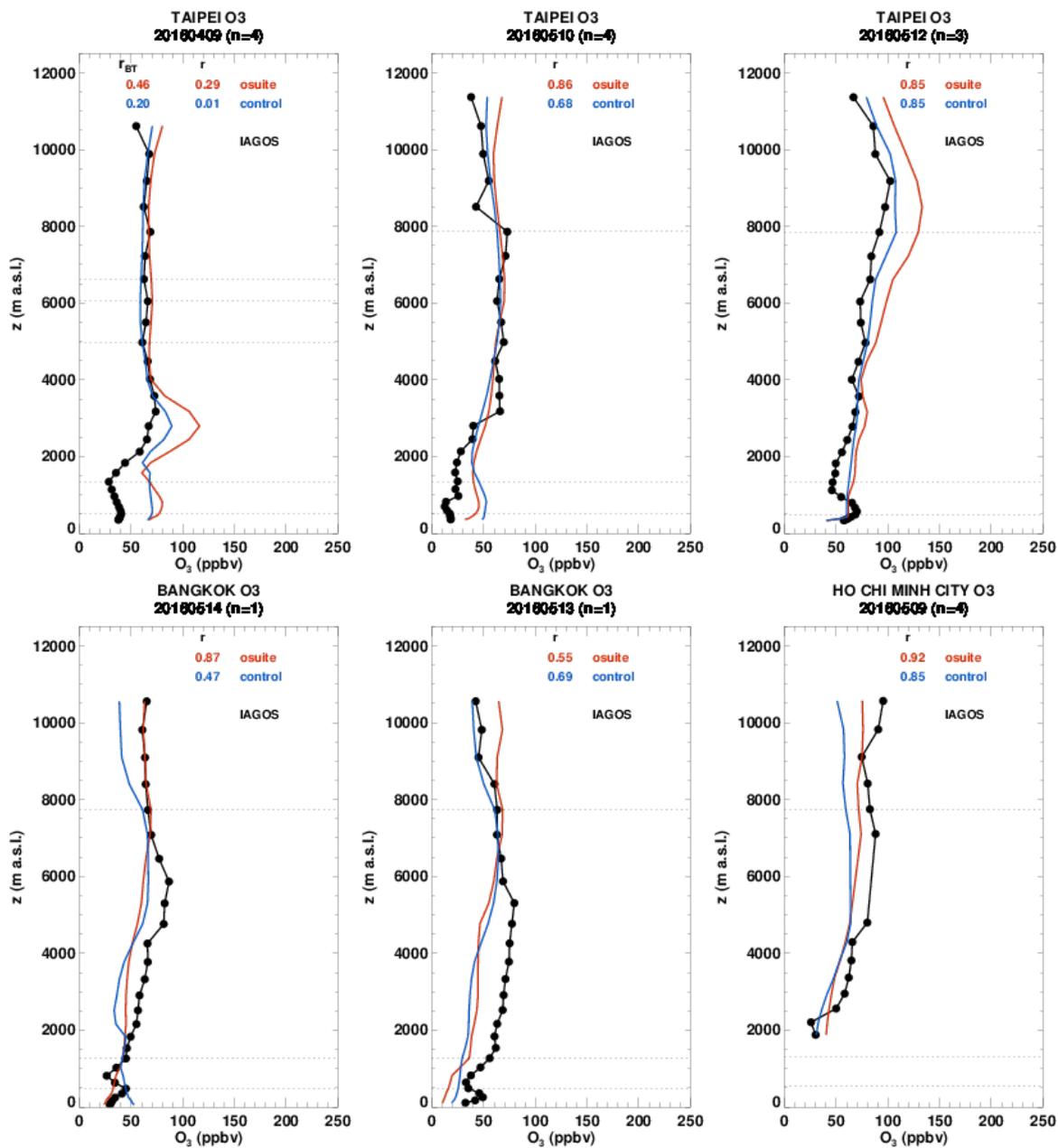
low ozone mixing ratios as found in the boundary layer, the o-suite has slightly underestimated the ozone. At other destinations across South-East Asia (Bangkok, Ho Chi Minh City, Manila) ozone is underestimated throughout the troposphere. Two profiles at Hong Kong show an increase in ozone around 2000m which is captured by the two runs although overestimated by the osuite.

Equatorial West-Africa

Air France visited several destinations in West Africa (Bamako, Lagos, Abijan, Cotonou, Ouagadougou). These cities are influenced by anthropogenic emissions from vehicles, and biomass burning from December to March. We do not therefore expect to see a big influence of biomass burning on the profiles shown here for MAM. Intense pollution from oil industries, all year long, has effects on the profiles at Lagos. In addition, intense pollution from oil industries all year long has effects on the profiles at Lagos. The Ozone is elevated in the surface and boundary layers at Bamako and Lagos, and at altitudes below 4000m at Abijan, likely related local pollution. Often the two runs fail to capture the increase in ozone at these altitudes whereas in the free troposphere the model does better at matching the observed ozone concentrations. The two runs also manage to capture an increase in ozone around 5000m at Ouagadougou.

Great Lakes Region

As over Europe, these profiles over the Great Lakes Region of North America (Fig. 3.1.8) show that the control run does better than the osuite at representing the gradients of ozone in the UTLS. The most significant event affecting this region was the fires in the Fort Mc Murray region of Alberta, beginning around 1 May, and being the worst fires ever recorded in Alberta. The wildfires were intense over Alberta, to the west of the Great Lakes Region, so we might expect to find transport of polluted plumes to this area. The peak in ozone at Montreal on 28 April at 8000m is possibly caused by this, however since the CO remains low, it is probably a stratospheric intrusion bringing ozone rich air down from the stratosphere.



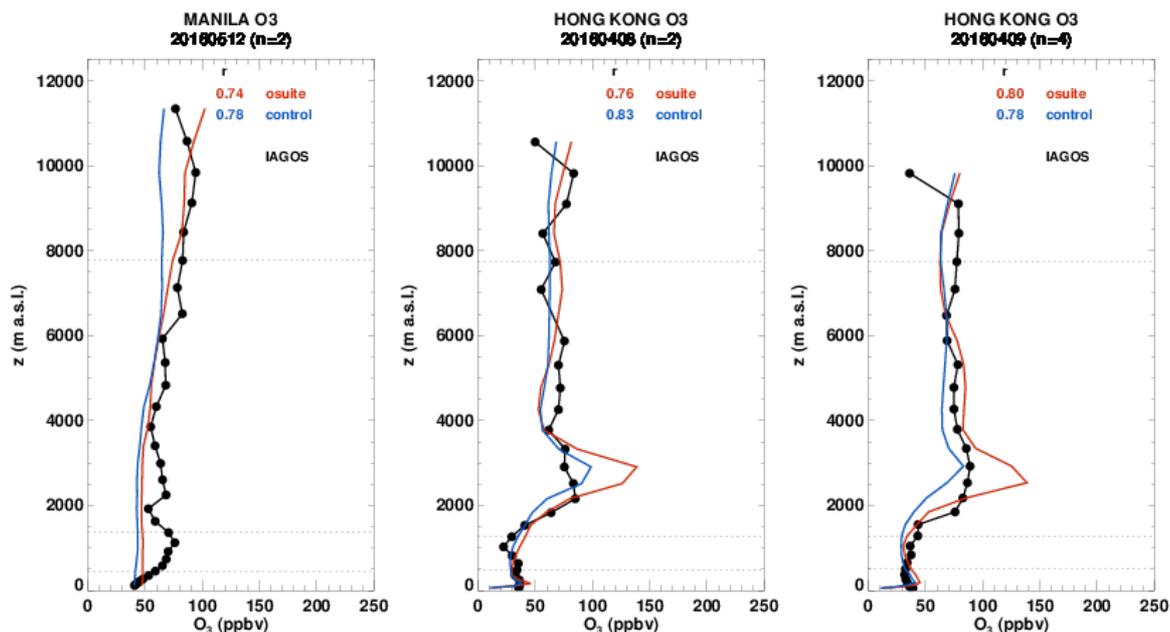
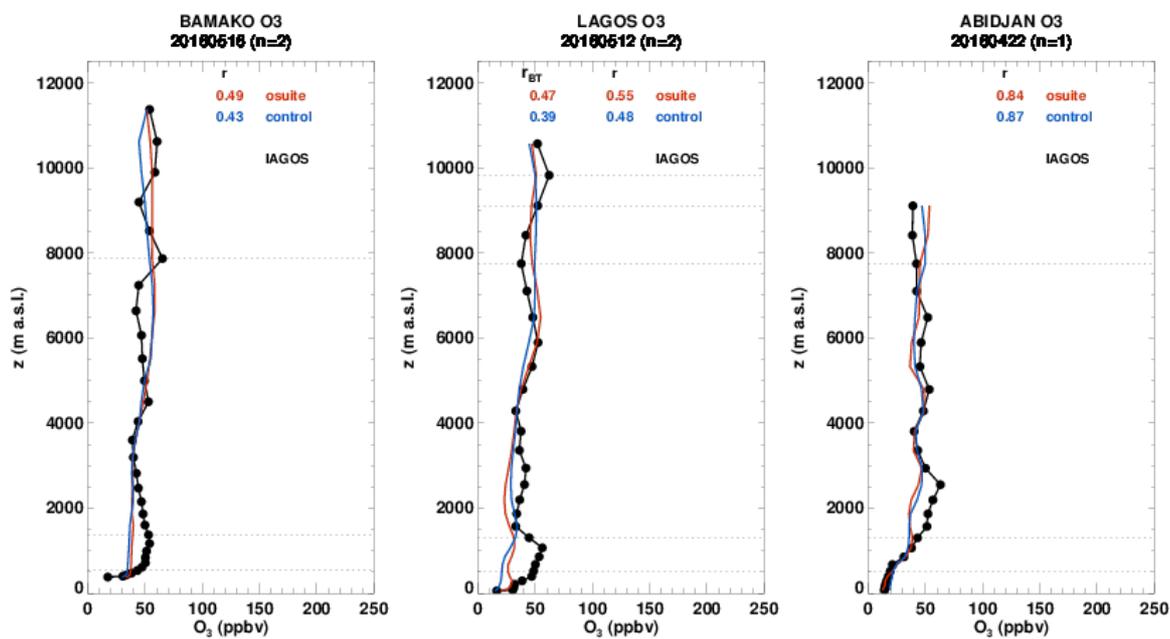


Figure 3.1.6: Selection of daily profiles of ozone from IAGOS (black) and the o-suite (red) and control (blue) over Asia (Taipei, Bangkok, Ho Chi Minh City, Manila) over the period March 2016-May 2016.



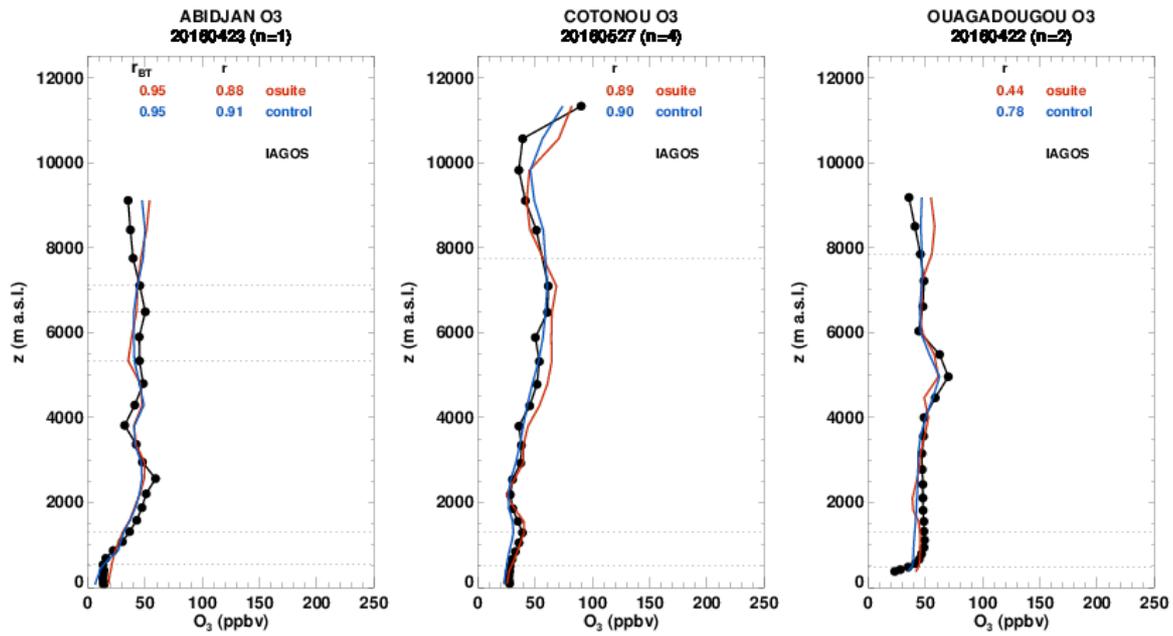


Figure 3.1.7: Profiles of ozone from IAGOS (black) and the two NRT runs over Port Harcourt and Abuja in January 2016.

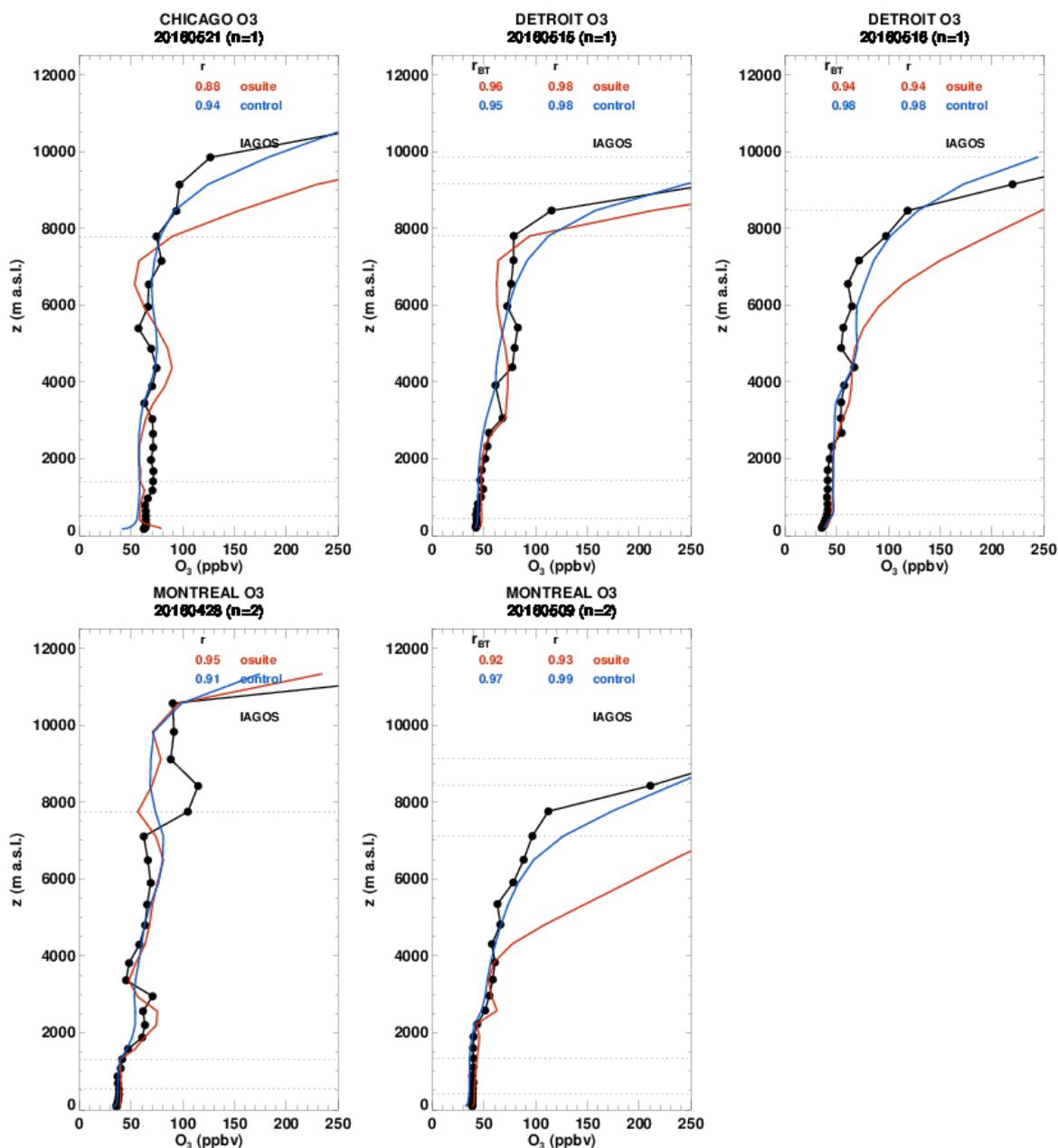


Figure 3.1.8: Profiles of ozone from IAGOS (black) and the two NRT runs over in the Great Lakes region.

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 March to May 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/>, and based on ESRL on <http://www.academyofathens.gr/kefak/cams/index.html>. Results are summarized in Figs 3.1.9 and 3.1.10.

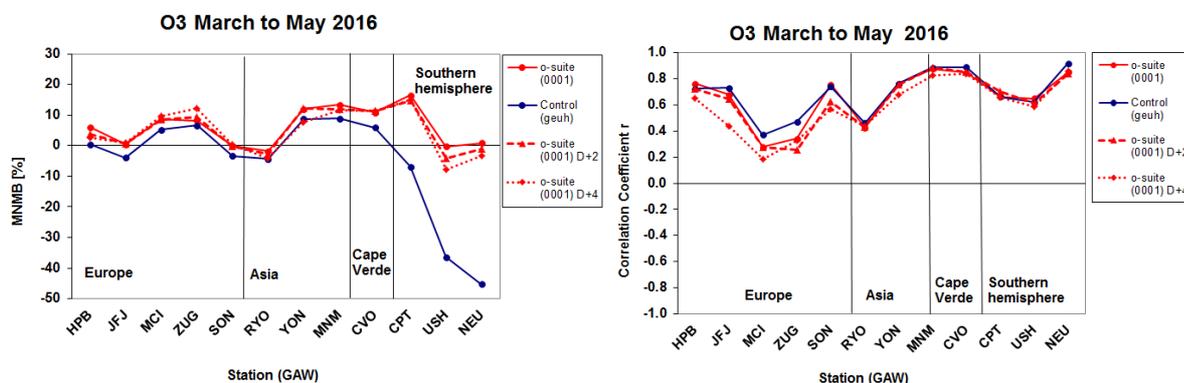


Figure 3.1.9: Modified normalized mean bias in % (left) and correlation coefficient (bottom right) of the NRT model runs compared to observational GAW data in the period March to May 2016. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.

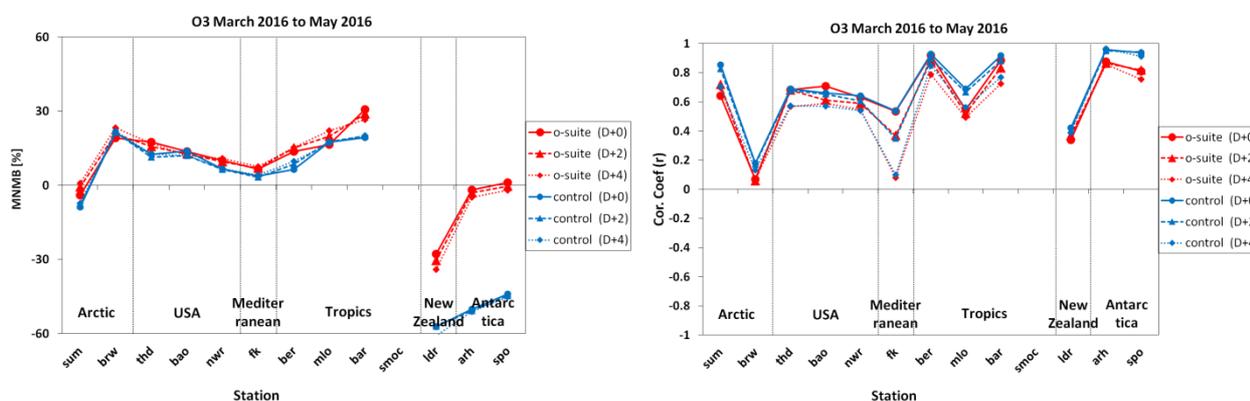


Figure 3.1.10: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT forecast runs compared to observational ESRL data in the period March to May 2016. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.

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. 3.1.9. 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. 3.1.10). Correlations between simulated and observed surface ozone values remain almost stable till D+2 (forecast run from 48h to 72h), but then drop in all regions (correlations for D+4 are lower than correlations for D+2 and D+0) with the only exception of Antarctica (see Fig. 3.1.10, right graph).

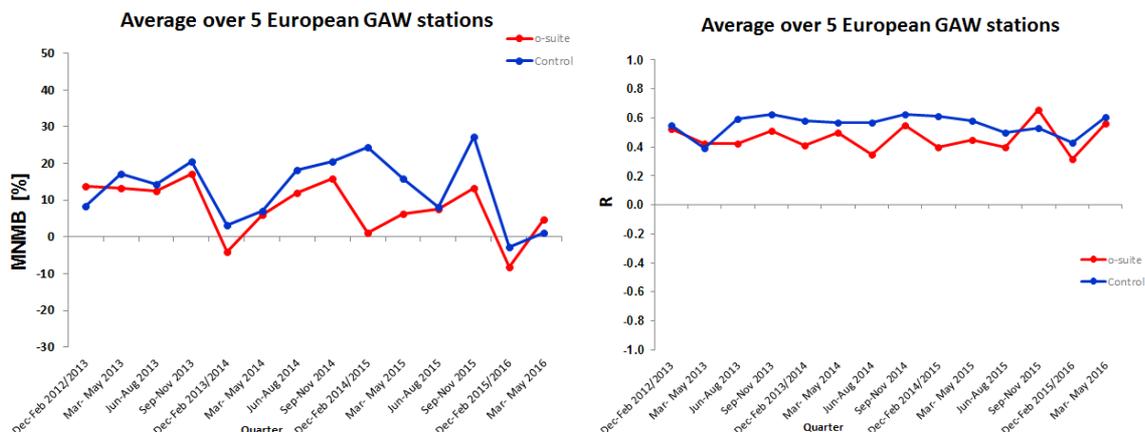


Figure 3.1.11: Long term (Dec. 2012 – May 2016) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

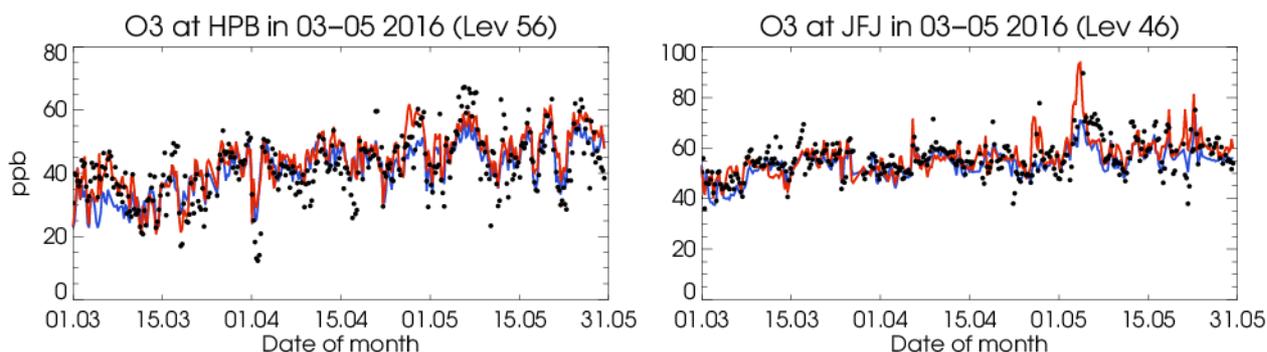


Figure 3.1.12: Time series for the o-suite (red) and Control (blue) compared to GAW observations at Hohenpeissenberg (47.8°N, 11.0°E) and Jungfrauoch (46.5°N, 7.9°E).

A comparison of the seasonal-mean MNMB over Europe (Fig. 3.1.11) 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.

Looking at different regions, for European stations (HPB, JFJ, ZUG, SON, MCI, FK), observed O₃ surface mixing ratios are mostly slightly overestimated by both, the o-suite and the control run, with MNMBs between 0 to 10% for the o-suite and between -4 and 7% for the control run (see also Fig. 3.1.12). Correlations for the European stations are between 0.28 to 0.76 for the o-suite and between 0.37 and 0.74 for the control run. The time series plots show that maximum values are captured by the o-suite, however, minimum concentrations are partly not resolved by the model, see Fig. 3.1.12.

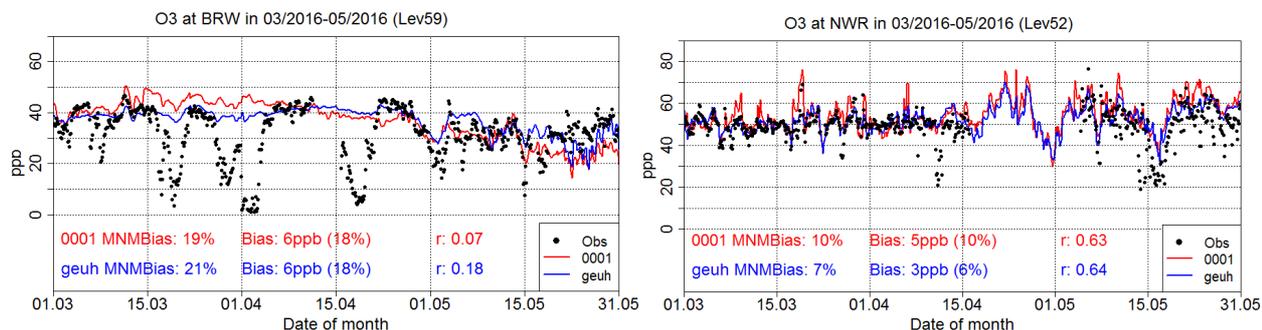


Figure 3.1.13: Time series for the o-suite (red) and control (blue) compared to ESRL observations at Point Barrow station (71.32°N, 156.61°W, left) and Niwot Ridge, Colorado station (40.04°N, 105.54°W, right).

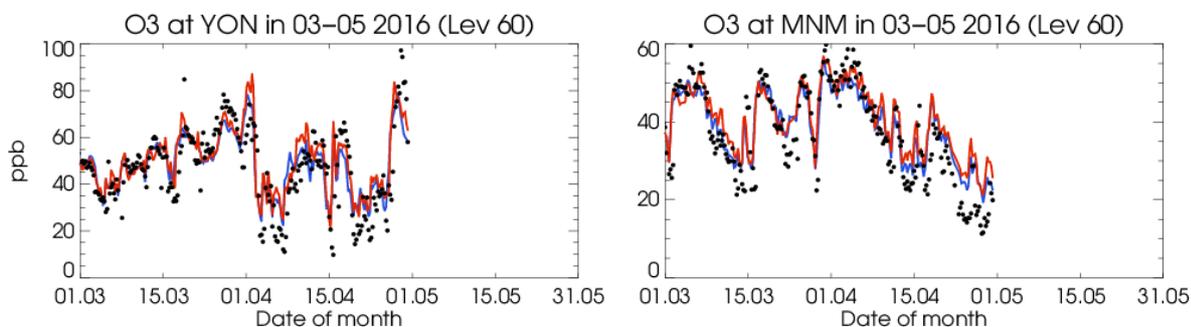


Figure 3.1.14: Time series for the o-suite (red) and Control (blue) compared to GAW observations at Yonagunishima (24.4°N, 123.02°E) and Minamitorishima (24.3°N, 153.9°E).

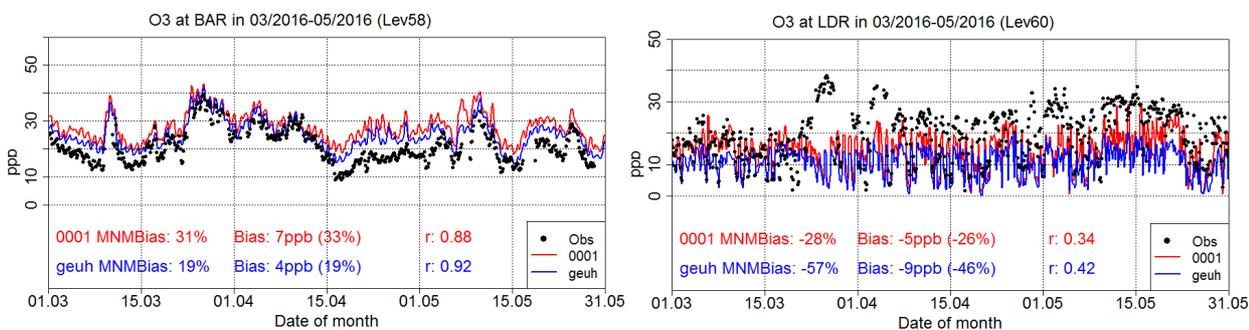


Figure 3.1.15: Time series for the o-suite (red) and control (blue) compared to ESRL observations (black dots) at Ragged Point, Barbados station (13.17°N, 59.46°W) and at Lauder (45.04°S, 169.68°E).

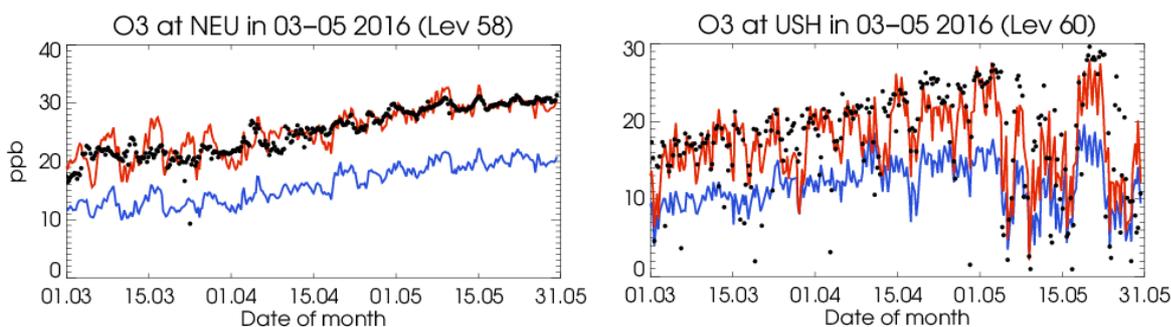


Figure 3.1.16: Time series for the o-suite (red) and control (blue) compared to GAW observations (black dots) at Neumayer (70.7°S, 8.3°W) and GAW observations at Ushuaia (54.9°S, 68.2°W).



The o-suite reproduces well surface ozone mean concentrations over Summit station (SUM) in the Arctic while the control run underestimates it by -10%. Both runs reproduce well also surface ozone variability ($r > 0.65$). On the contrary both runs strongly overestimate surface ozone mean concentrations over Point Barrow by 30% and fail to reproduce the surface ozone variability ($r \approx 0.0$). The time series plot shows that 3 ozone depletion events appear in the ozone observations at BRW during the period 20 March to 20 April, which are not reproduced by the model (Figure 3.1.13, right graph).

For USA stations (THD, BAO, NWR) both runs overestimate surface ozone mean concentrations between 17% at THD and 10% at NWR. Correlations between simulated and observed surface ozone for both runs are high in all 3 USA stations ($r > 0.6$).

For Asian stations (RYO, YON, MNM), both runs show a positive offset in surface O_3 mixing ratios for the stations YON and MNM, with MNMBs are around 12% for the o-suite and around 8% for the control run. The overestimation mostly concerns the minimum concentrations in the model. Concentration peaks are well reproduced, as can be seen in Figure 3.1.14. For RYO station, both runs show a small negative bias. Correlations are between 0.42 and 0.87 for both runs.

For the stations in the Southern Hemisphere (CPT, NEU, USH) the o-suite reproduces ozone mixing ratios well with MNMBs between 0% and 16%. The data assimilation corrects the negative offset in the control run, see Figure 3.1.16 (right panel). Correlation coefficients are between 0.67 and 0.92 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/>) 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, with the exception of the stations Cirat and Hospital Joan March (Majorca Island) in Spain. 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).



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 MAM 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	control	o-suite	control
Al Cornocales	ES1648A	-5.66	36.23	189	59	16	13.6	9.1	0.34	0.43
Caravaka	ES1882A	-1.87	38.12	1	60	73	-21.6	-26.5	0.44	0.44
Zarra	ES0012R	-1.10	39.08	885	55	70	0.0	-5.7	0.75	0.83
Villar Del Arzobispo	ES1671A	-0.83	39.71	430	60	48	-7.3	-12.5	0.52	0.44
Cirat	ES1689A	-0.47	40.05	466	60	37	-7.8	-13.3	0.09	0.09
Bujaraloz	ES1400A	-0.15	41.51	327	60	60	-10.8	-16.3	0.81	0.83
Morella	ES1441A	-0.09	40.64	1150	52	51	1.7	-4.3	0.64	0.70
Bc-La Senia	ES1754A	0.29	40.64	428	60	21	-20.4	-25.0	0.41	0.43
Ay-Gandesa	ES1379A	0.44	41.06	368	60	15	-14.6	-20.2	0.56	0.57
Ak-Pardines	ES1310A	2.21	42.31	1226	51	81	20.1	14.1	0.45	0.42
Hospital Joan March	ES1827A	2.69	39.68	172	56	3	15.7	9.0	-0.14	-0.07
Al-Agullana	ES1201A	2.84	42.39	214	60	25	1.8	-8.5	0.43	0.47
Av-Begur	ES1311A	3.21	41.96	200	58	9	0.3	-5.4	0.41	0.35
Plan Aups/Ste Baume	FR03027	5.73	43.34	675	54	21	1.3	-3.9	0.58	0.65
Gharb	MT00007	14.20	36.07	114	58	31	4.1	-2.4	0.51	0.61
Finokalia	GR0002R	25.67	35.32	250	55	4	6.6	3.9	0.53	0.54
Ineia	-	32.37	34.96	672	52	5	0.9	-1.7	0.58	0.67
Oros Troodos	-	32.86	34.95	1819	47	11	1.6	0.4	0.66	0.75
Agia Marina	CY0002R	33.06	35.04	532	53	14	3.0	0.0	0.70	0.73

In terms of biases, both model runs mostly underestimate surface ozone mean concentrations over Spain. Table 3.1.1 and Fig. 3.1.18, top right, indicate that over Spain the MNMBs for both model versions exceed -20% in the case that the elevation of station corresponds to the surface model level. Over the Mediterranean shore of Spain data assimilation seems to improve slightly the biases. In all other Mediterranean stations (Plan Aups/Ste Baume in France, Gharb in Malta, Finokalia in Crete, Ineia, Oros Troodos and Agia Marina in Cyprus) both o-suite and control runs reproduce very well surface ozone mean concentrations ($-4% < \text{MNMBs} < 6%$; see also Figure 3.1.17, central and lower graphs).

The spatial distribution of MNMBs and correlations of the o-suite over the Mediterranean is shown in 3.1.18, where the contrast in the model MNMBs between Mediterranean shore of Spain (strong negative signal) and Central and Eastern Mediterranean (MNMBs close to zero) is evident. On the other hand it clearly shows the generally high correlations between simulated and observed surface ozone values in all regions throughout de Mediterranean.

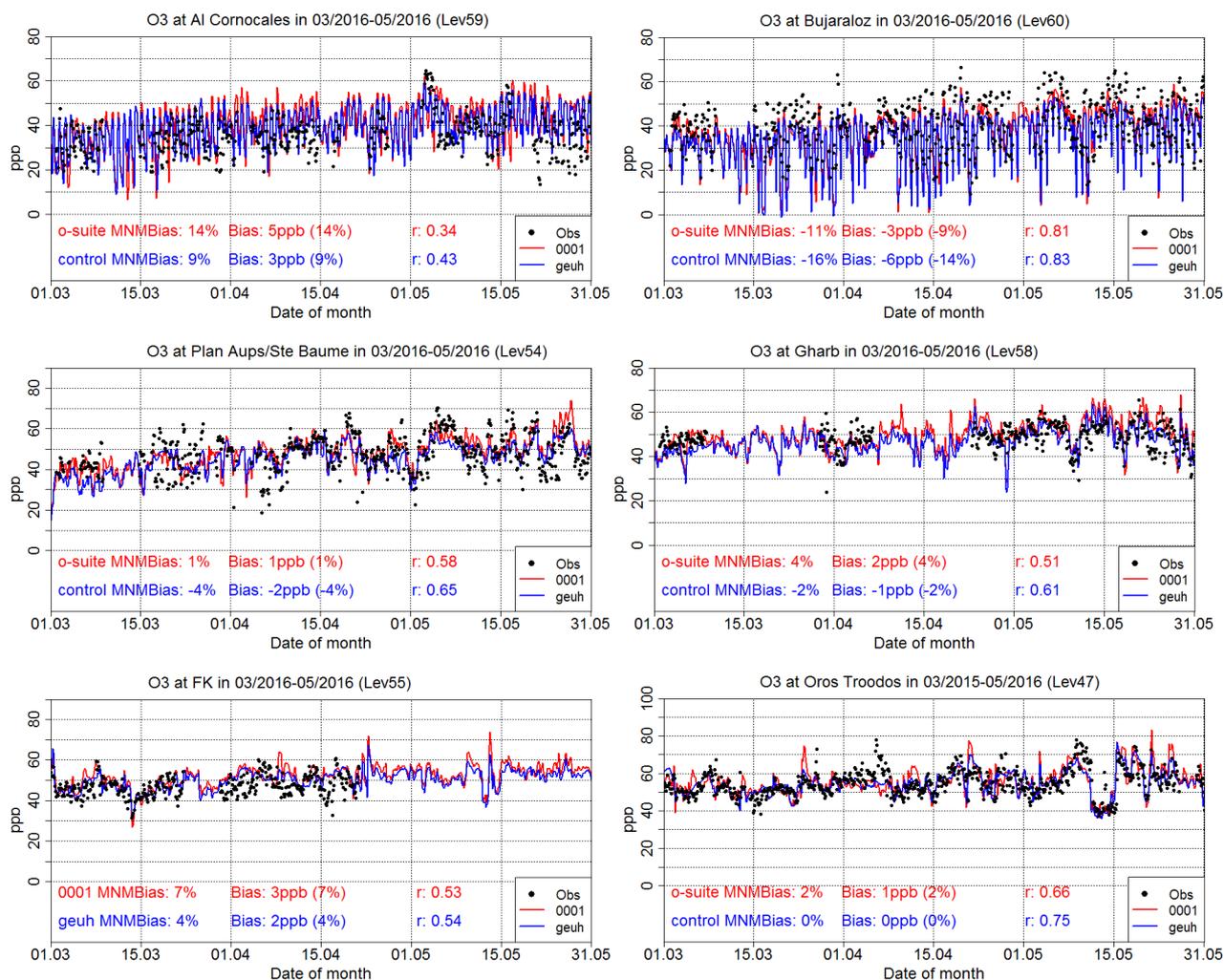


Figure 3.1.17: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Cornocales, Spain station (36.23°N, 5.66 °W, top left), at Bujaraloz, Spain station (41.51°N, 0.15°W, top right), at Plan Aups/Ste Baume, France station (43.34°N, 5.73°E, center left), at Gharb, 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 FK of Cyprus, at Troodos Mountain station (34.95°N, 32.86 °E, low right).

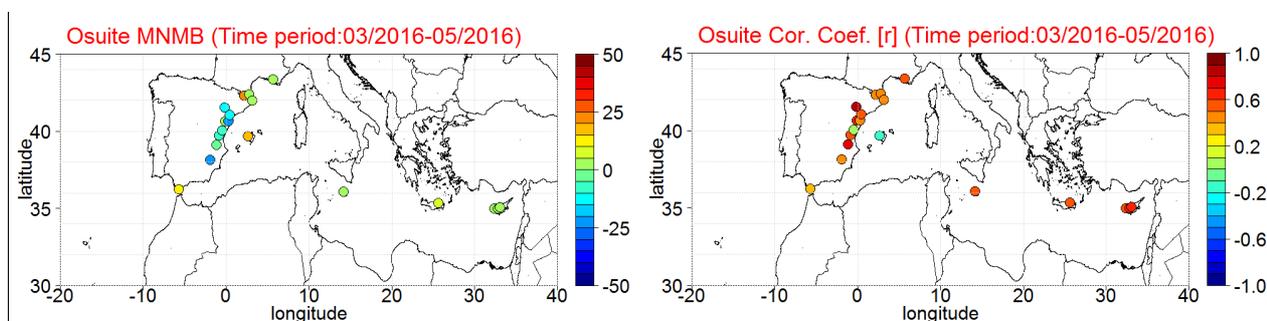


Figure 3.1.18. Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 March 2016 to 31 May 2016.

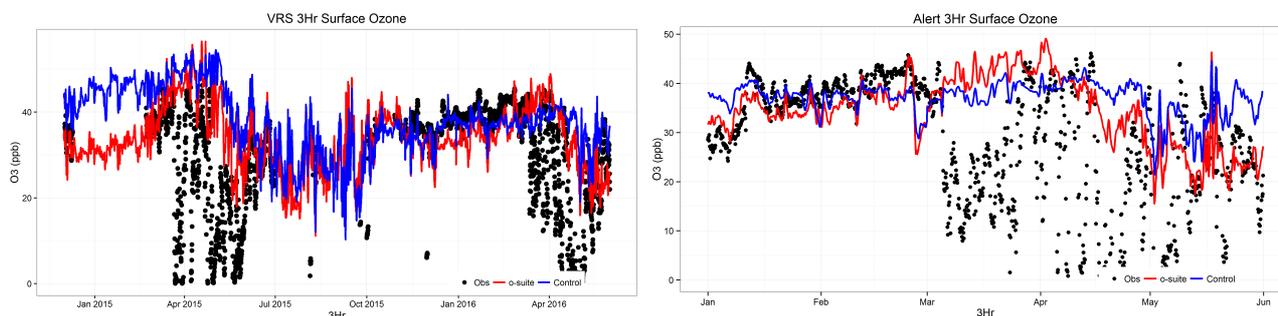


Figure 3.1.19: Time series for o-suite (red) and Control (blue) compared to observations (black dots) at the Villum Research Station, Station Nord, Greenland (left) and Alert, Nunavut (right).

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 and from Alert, Nunavut from the IASOA network, Figure 3.1.19.

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 – May 2016. Ozone depletion events in March – June in 2015 and 2016 are not captured by the model simulations during spring for both 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 March 2016 – May 2016 the measurements are not quality controlled. Due to the ozone depletion events the model simulations overestimate measured concentrations, except for the few days without depletion events, where the predicted model levels are in a fair agreement with observations for at the two sites. This results in a strong positive bias and low correlation coefficients (Table 3.1.2).

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 March – May 2016.

		NMB	r
Alert	o-suite	0.55	0.24
	control	0.66	0.18
VRS	o-suite	0.37	0.42
	control	0.46	0.28

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

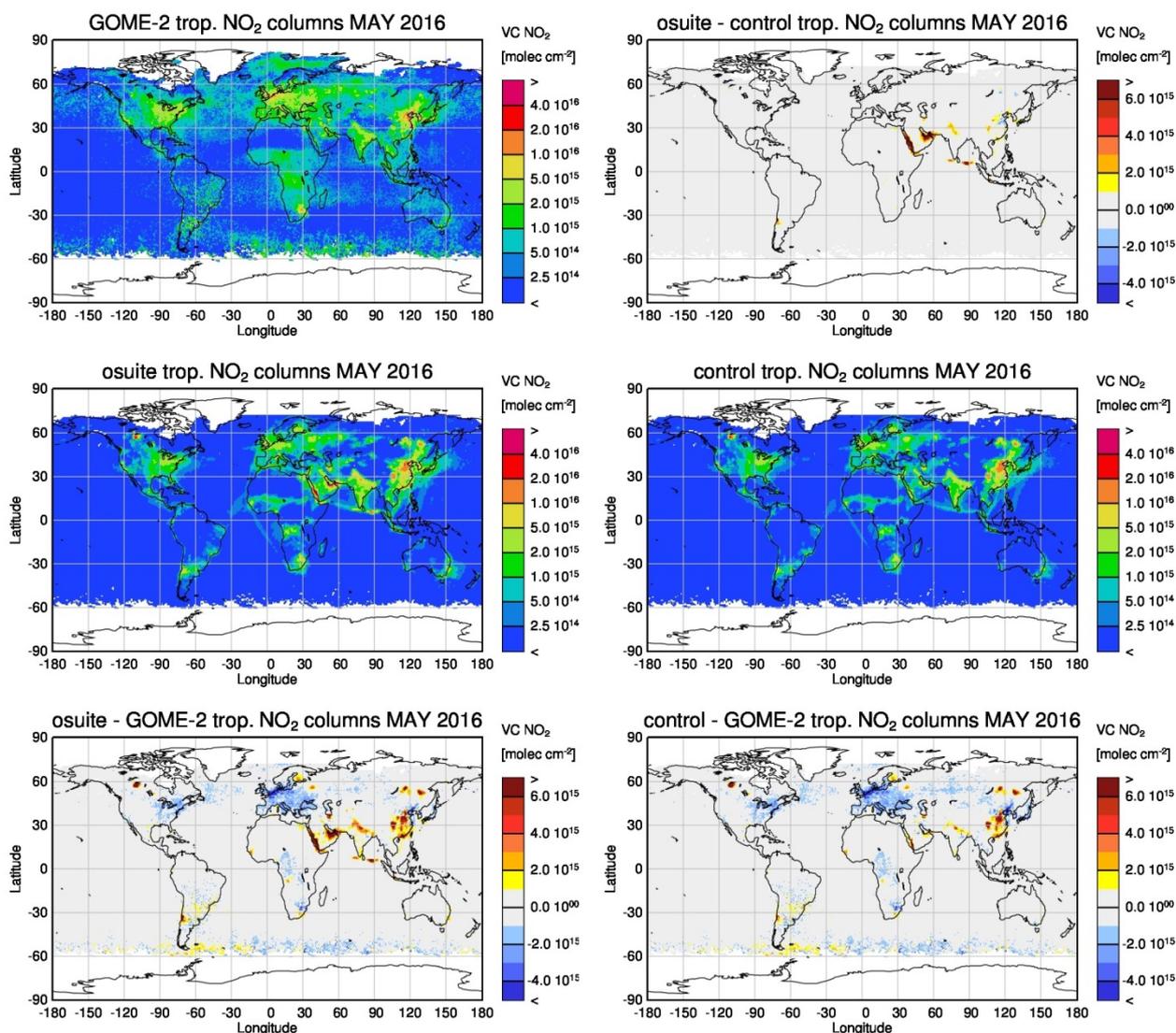


Figure 3.2.1: Global map comparisons of satellite retrieved and model simulated tropospheric NO₂ columns [molec cm⁻²] for May 2016. The top row shows monthly mean tropospheric NO₂ columns retrieved by GOME-2 as well as the difference between osuite and control, the second row shows the corresponding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

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 of 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 May 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. Moreover, the models overestimate values over the Red Sea, but the reasons for this require further investigation. The control run performs much better than the o-suite here, indicating that data assimilation may contribute to the overestimation of values in this region. Compared to satellite data, all model runs underestimate tropospheric background values over Africa, Europe and Eastern 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, Moscow and Helsinki are overestimated. Moreover, both runs overestimate fire emissions over Central Africa.

Moreover, both runs show local maxima over Siberia (around 120° E, 55° N) and Canada (around 125° W, 55° 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.

Closer inspection of the seasonal variation of tropospheric NO₂ in some selected regions (Fig. 3.2.2) 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 the NO₂ column retrievals decreased since 2014, model simulated values are since then in better agreement with the satellite retrieved ones. However, this decrease in values is not reproduced by 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 summer 2015 are by more than 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

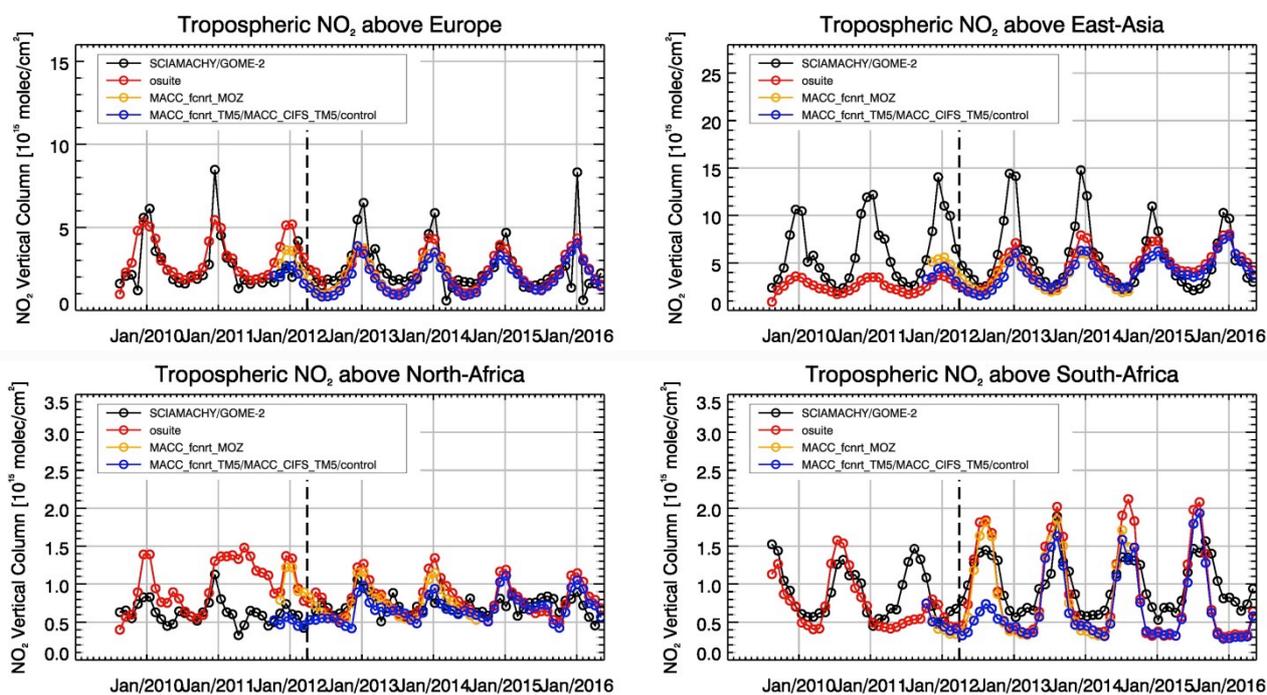


Figure 3.2.2: Time series of average tropospheric NO₂ columns [10^{15} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions (see Annex 2 for definition of regions). Upper panels represent regions dominated by anthropogenic emissions, lower panels represent those dominated by biomass burning. The blue line shows MACC_fcfrnt_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.

models. Summertime NO₂ columns over North-Africa are underestimated compared to the satellite data for 2015. The o-suite 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 1.6 larger compared to GOME-2 retrievals in August 2014). 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

¹ No contribution from UCLE in MAM 2016 due to instrument failure.

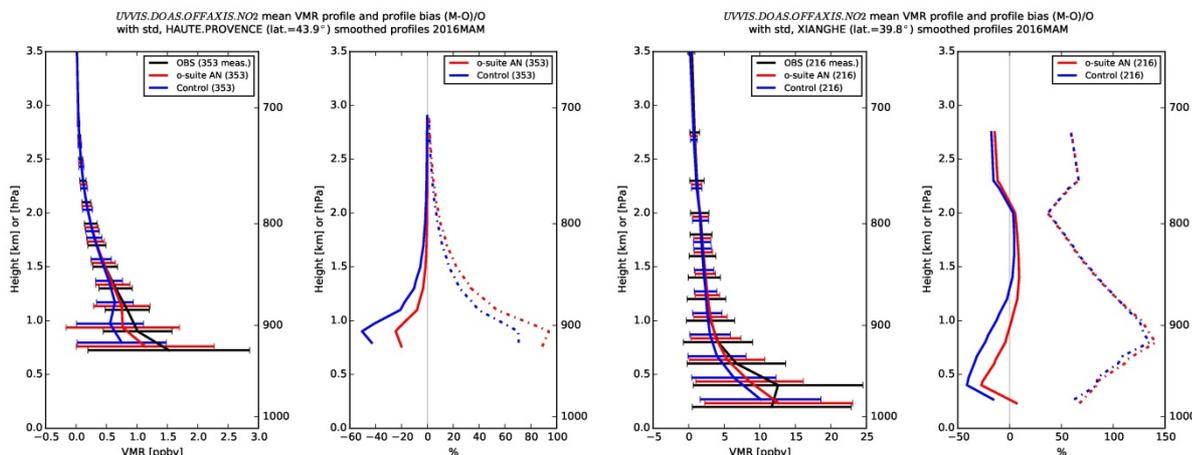


Figure 3.2.3: Seasonal mean tropospheric NO₂ profiles by o-suite (red) and Control (blue) compared to NDACC UVVIS DOAS data at Haute Provence (43.9°N, 5.71°E, left top) and Xianghe (39.8°N, 117°E, right) for March-April-May 2016.

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. 3.2.3). 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.

From Figs. 3.2.3 and the below table, we see the assimilation has a negative effect at the rural site of OHP. The positive effect is observed at the more polluted site in Uccle and in Xianghe.

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, Uccle 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).

		JJA			SON			DJF			MAM		
		MB	stddev	nobs									
osuite	ohp	11.68	56.81	228	-5.72	40.84	223	35.13	96.22	115	-6.08	48.27	353
control	ohp	6.61	51.59	228	-12.48	59.38	223	21.16	84.90	115	-20.36	33.14	353
osuite	uccle	-22.40	49.74	70	-20.35	34.65	53	-5.70	47.18	44	-	-	-
control	uccle	-41.93	32.34	70	-31.68	25.87	53	-20.89	56.31	44	-	-	-
osuite	xianghe	78.32	93.34	220	91.40	141.9	64	11.73	77.38	103	23.50	74.65	216
control	xianghe	65.35	92.65	220	69.98	104.1	64	24.63	105.5	103	5.21	63.77	216



3.3 Carbon monoxide

3.3.1 Validation with Global Atmosphere Watch (GAW) Surface Observations

For the Near-Real-Time (NRT) validation, 7 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 March – May 2016. The latest validation results can be found on the CAMS website:

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

On average for the study period the control run shows higher values for the southern hemisphere (Figure 3.3.1), which lead to a positive offset corrected by data assimilation for the o-suite. A comparison of the seasonal-mean MNMB over Europe (Figure 3.3.2) from December 2012 to present shows a slowly improving MNMB from about -20% in 2013 to -15% for more recent periods. Temporal correlation remains relatively constant at $r \sim 0.5$ on average.

For European stations, both runs show an underestimation of observed CO mixing ratios, with MNMBs between -6% and -17%. Correlation coefficients are between 0.34 and 0.70 for the o-suite and between 0.43 and 0.73 for the control run.

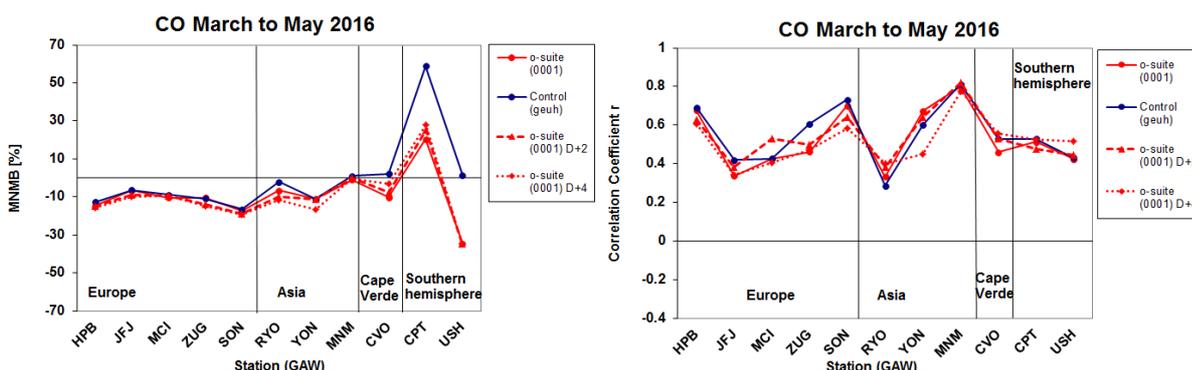


Figure 3.3.1: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT model runs compared to observational GAW data in the period December to February 2016. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.

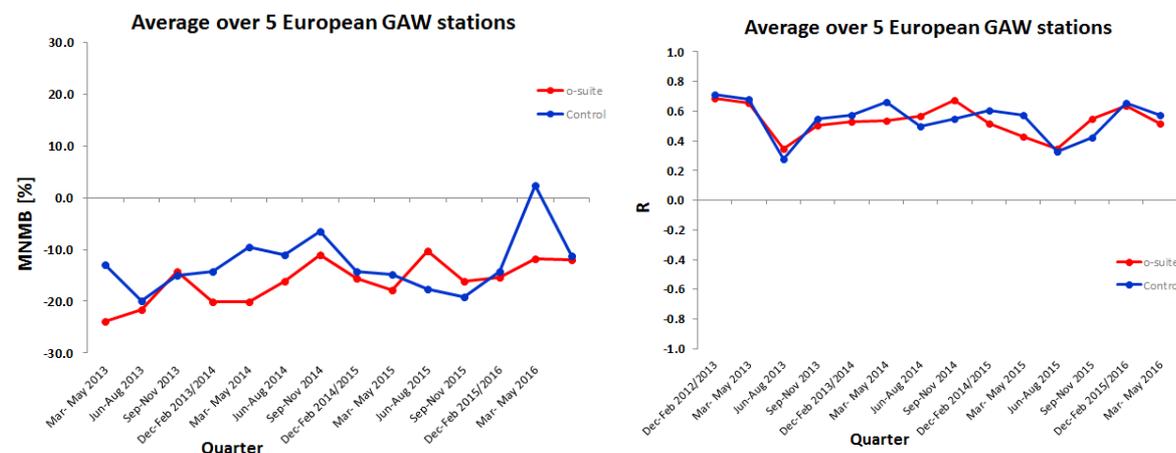


Figure 3.3.2: Long term (Dec. 2012 – May 2016) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

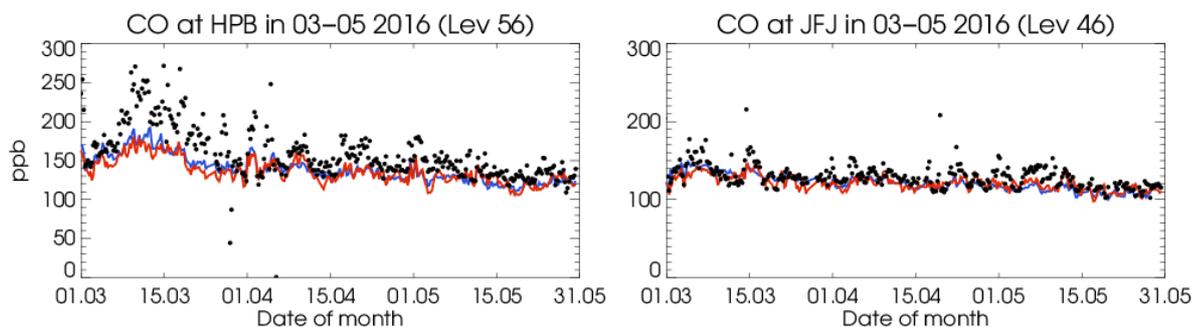


Figure 3.3.3: Time series for the o-suite (red) and control (blue) compared to GAW observations at Hohenpeissenberg (47.8°N, 11.0°E) and Jungfrauoch (46.5°N, 7.9°E).

For Asian stations, both runs correspond well to the observations, except for one peak in April for RYO, which does not correspond to the observations, with MNMBs between -1 and -12% for both runs, see Figure 3.3.4. Correlation coefficients are between 0.28 and 0.81.

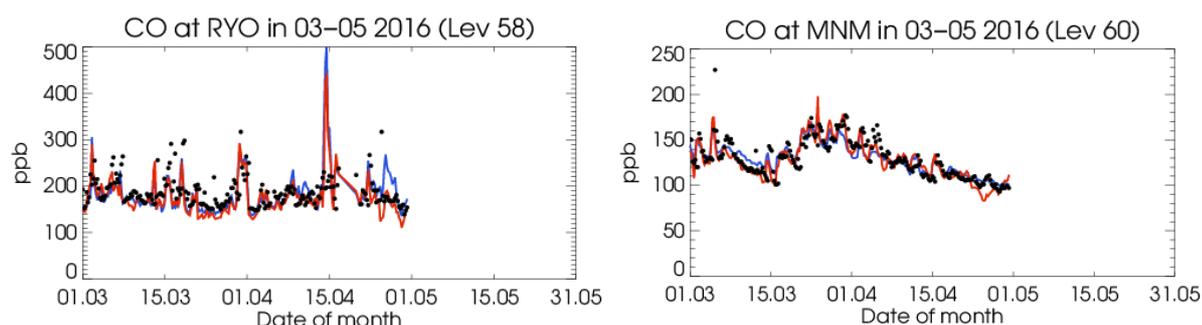


Figure 3.3.4: Time series for the o-suite (red) and control (blue) compared to GAW observations at Ryori (39.0°N, 1°141.8°E) and Minamitorishima (24.3°N, 123.9°E).

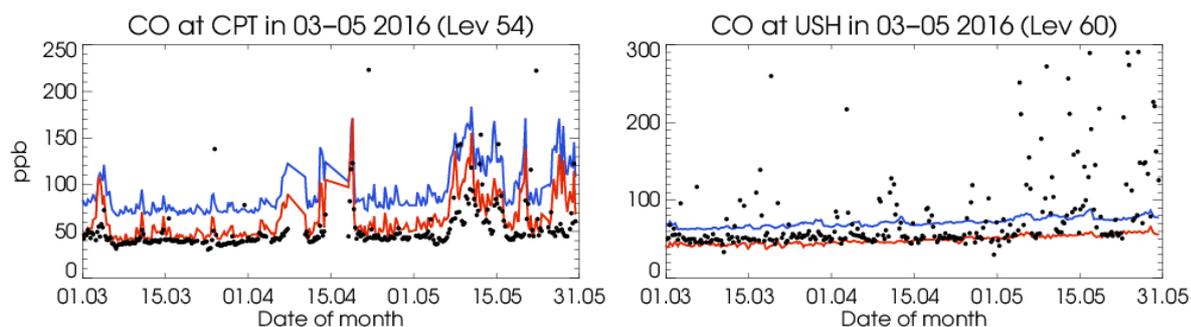


Figure 3.3.5: Time series for the o-suite (red) and control (blue) compared to GAW observations at Cape Point (34.35°S, 18.5°E) and Cape Verde (16.9°N, 24.9°W).

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 CO mixing ratios, see Figure 3.3.5.



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 March-April-May 2016, data from three aircraft have been validated, as discussed in Sec. 3.1.2. Figure 3.3.6 shows the time series of CO over Paris and Taipei for the 5 different layers throughout the troposphere. Over Taipei, the models do well at matching the concentrations of CO in all layers (see later profiles) however, in Europe there is a quasi-systematic underestimation of the CO surface and boundary layers in the winter months as also shown by Stein et al. (2014). The underestimation can be seen clearly at Paris throughout April/May particularly in the boundary layer.

Europe

Figure 3.3.7 gives examples of the CO profiles over Frankfurt, Amsterdam and Paris. In general we find that CO in the mid-troposphere is well estimated by the models. However, CO in the boundary and surface layers is frequently underestimated. At the same time, the models overestimate the abundance of CO in the UTLS. The profiles for the 12th April at Amsterdam and Frankfurt illustrate this well, being from two different aircraft.

Asia

The time series at Taipei (Fig. 3.3.6) illustrates 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 diverse airports show (Fig. 3.3.8).

At Bangkok, Ho Chi Minh City, Hong Kong and Manila, the profile of CO is well captured in the free troposphere with the o-suite making an improvement over the control run throughout the free troposphere where the control run often overestimates the amount CO. The maxima in CO seen at Taipei (8th April at 4000m), Bangkok (14th April at 2000m) and Hong Kong (8th, 9th April at 3000m), are well captured by the two runs in terms of altitude, and are generally improved in the osuite. Manila and Ho Chi Minh City both show a big increase in CO in the boundary and surface layers typical of large Asian cities and underestimated by both runs.

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. The MAM season is outside this period of intense burning with the profiles at Oagadougou and Cotonou showing, as expected, no notable peaks in CO. Some biomass burning was present in the countries neighbouring the Ivory Coast and is probably the cause of the peak in CO at Abijan on 22 and 23 April at around 2500m in altitude which is quite well captured by the control run. At Lagos pollution in the boundary layer is significant, with levels reaching 600ppbv. This is likely to be a combination of the pollution from oil industries all year long and the fact that Lagos is one of the world's most populous cities. The two runs estimate the CO in the free troposphere fairly well but underestimate the CO in the boundary layer at Lagos.

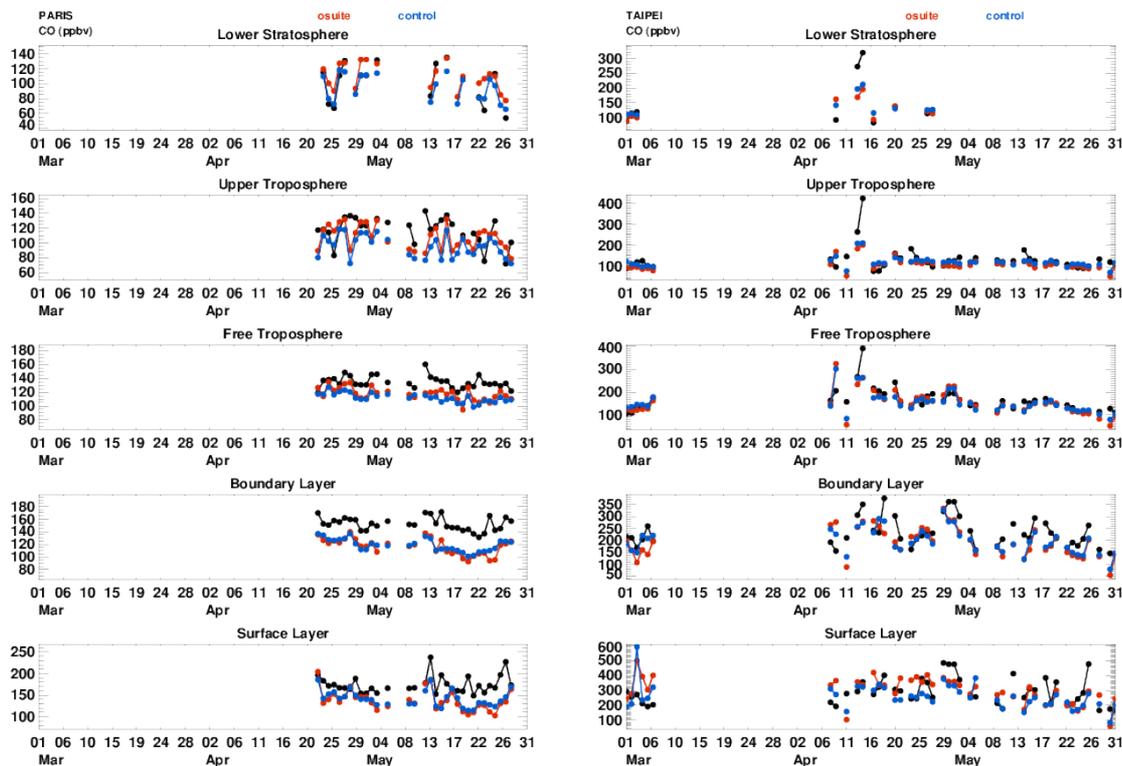


Figure 3.3.6: Time series of daily mean CO Taipei and Frankfurt during MAM 2016 for 5 layers, Surface, Boundary layer, Free Troposphere, Upper Troposphere and Lower Stratosphere.

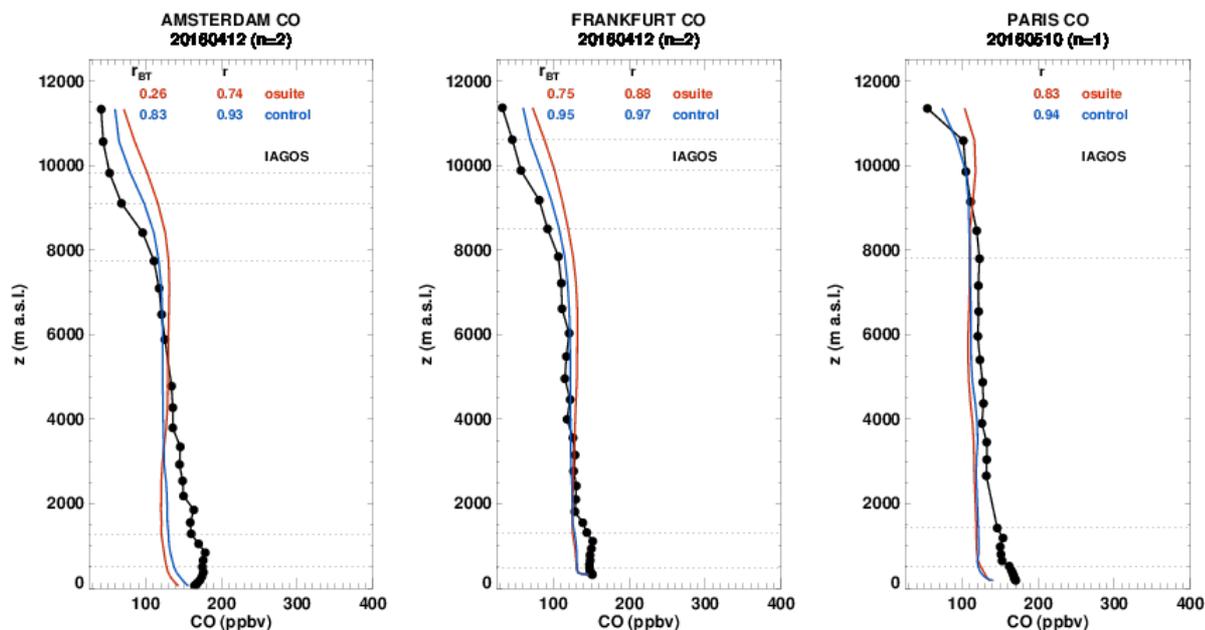
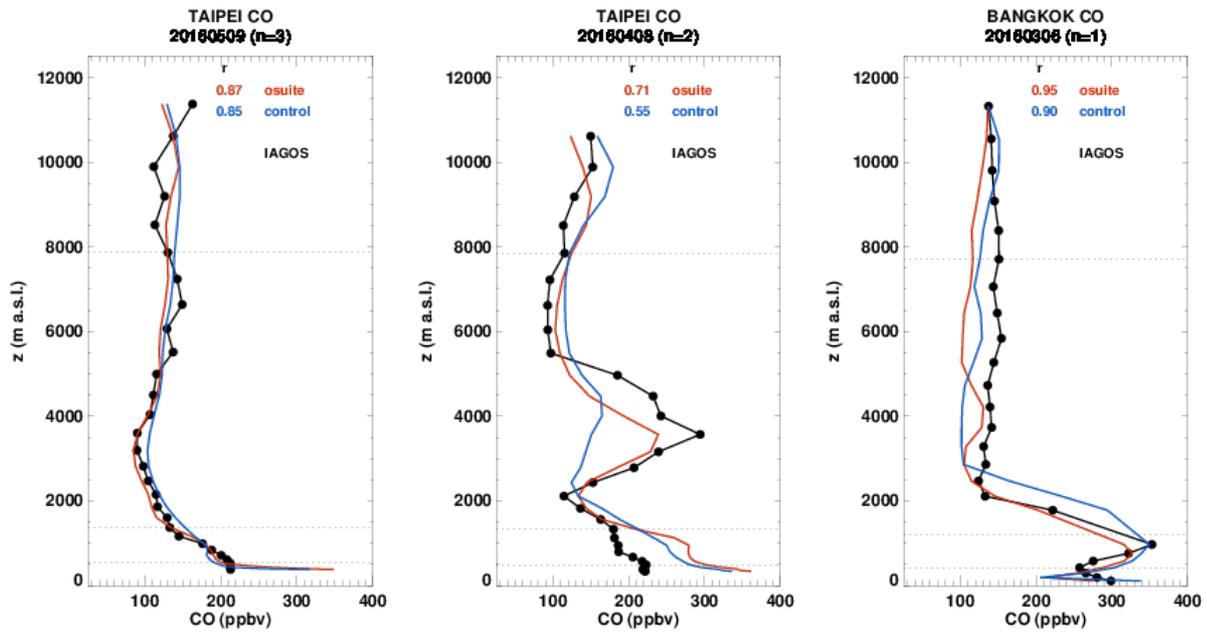


Figure 3.3.7: Selection of profiles of CO from IAGOS (black) and the two NRT runs over Europe in March-April-May 2016.



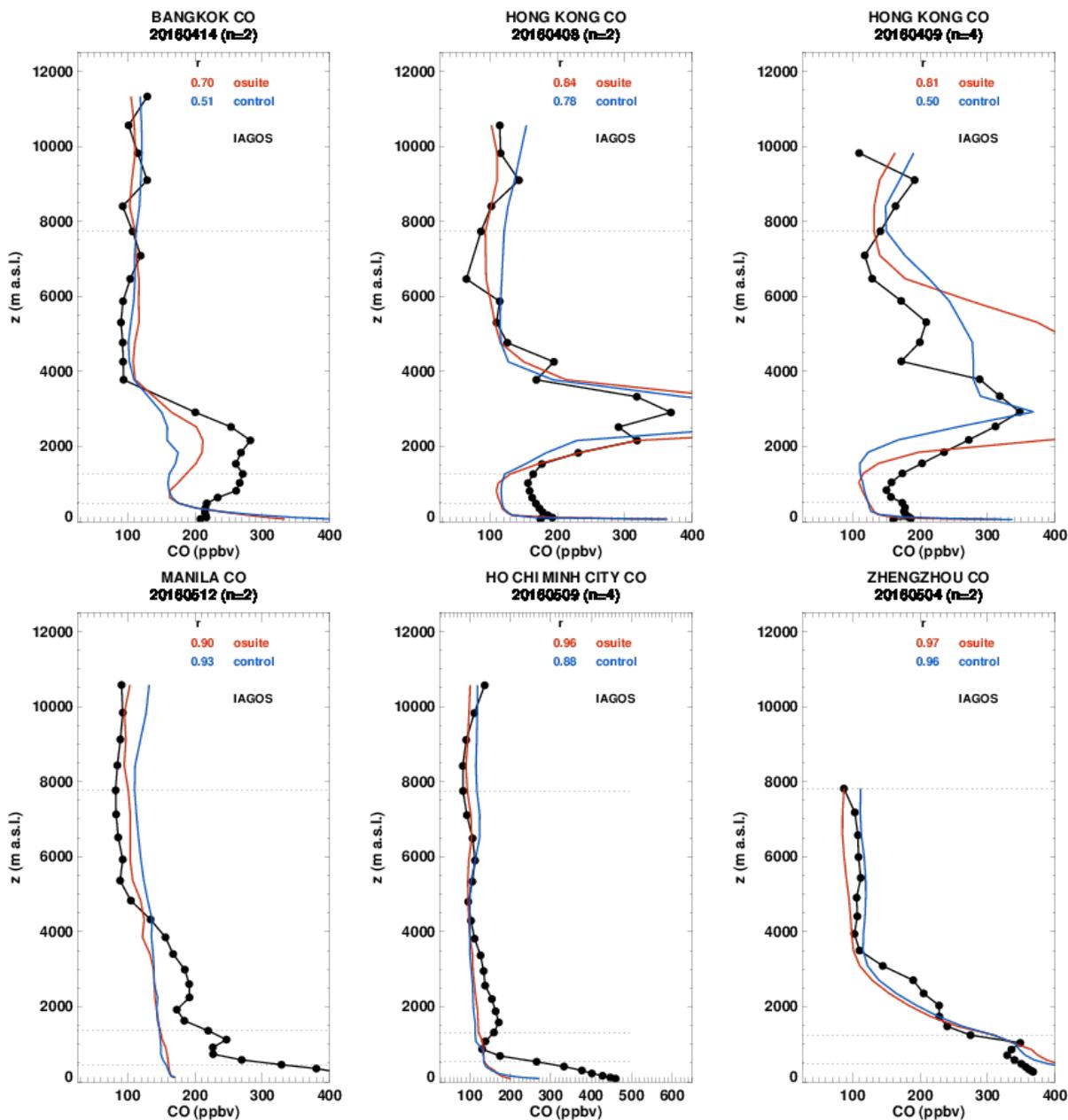


Figure 3.3.8: Profiles of CO from IAGOS (black) and the two NRT runs in China and South East Asia during the period March-May 2016.

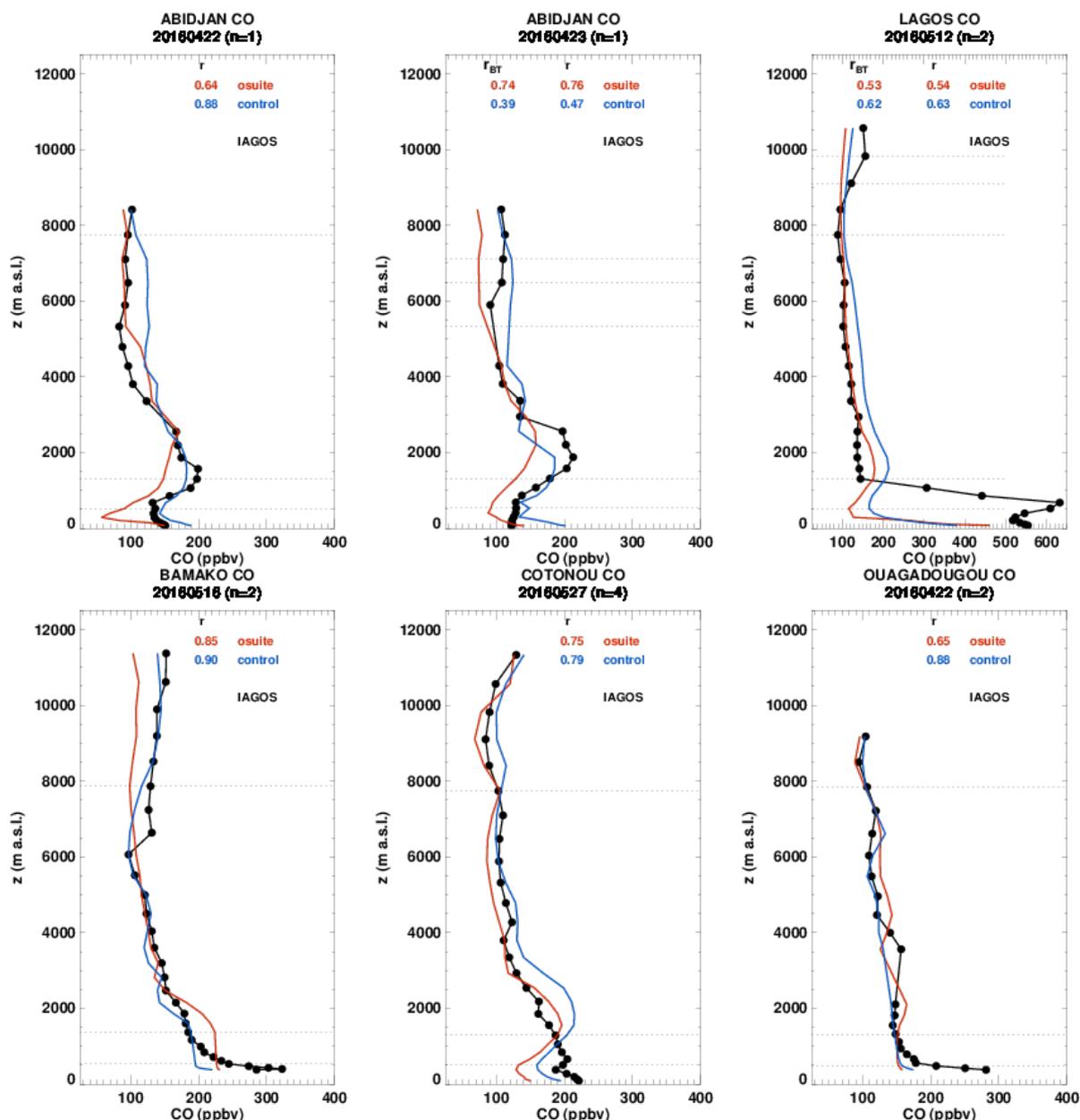


Figure 3.3.9: Profiles of CO from IAGOS (black) and the two NRT runs in China and South East Asia during the period March-May 2016

Great Lakes Region

The main event for this period was the intense forest fires that burned across Alberta particularly during the first two weeks in May 2016 leading to devastation in the Fort McMurray area. Downstream of Alberta during this time, we see peaks of carbon monoxide at altitudes from 4000-5000m which are possibly related to these fires. These are not captured by the two runs, however a peak is produced by both runs over Detroit on 16 May of the right magnitude but at the wrong altitude (2000m too high).

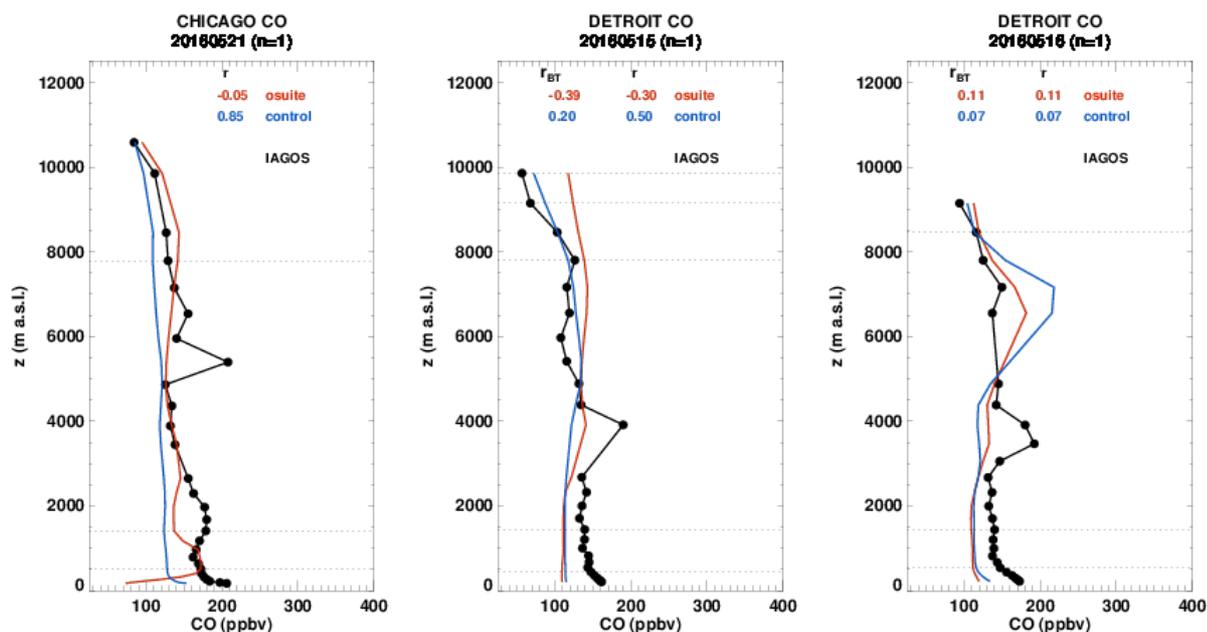


Figure 3.3.10: Profiles of CO from IAGOS (black) and the two NRT runs over Montreal, Chicago and Detroit, during the period MAM 2016.

3.3.3 Validation against FTIR observations from the NDACC network

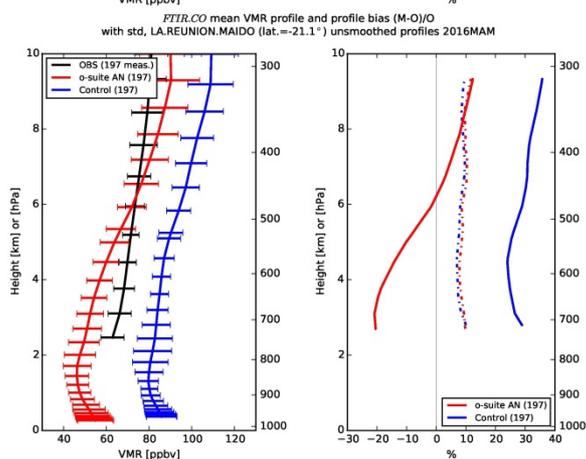
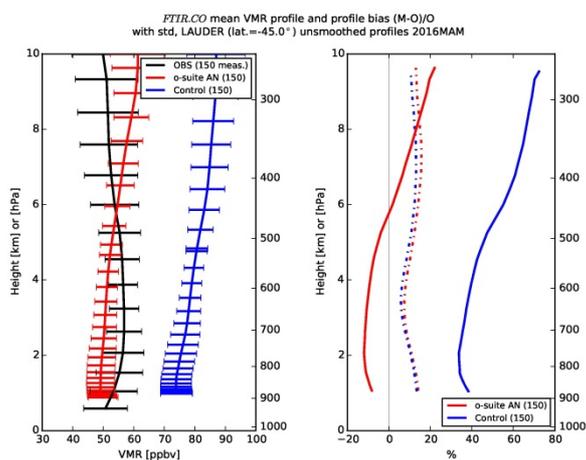
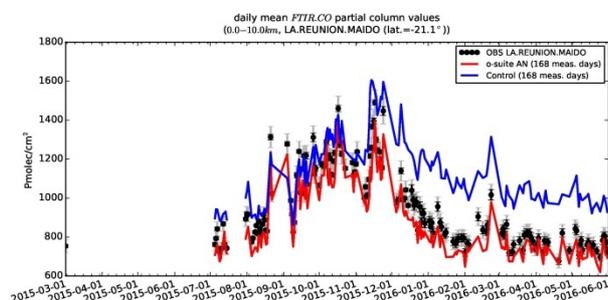
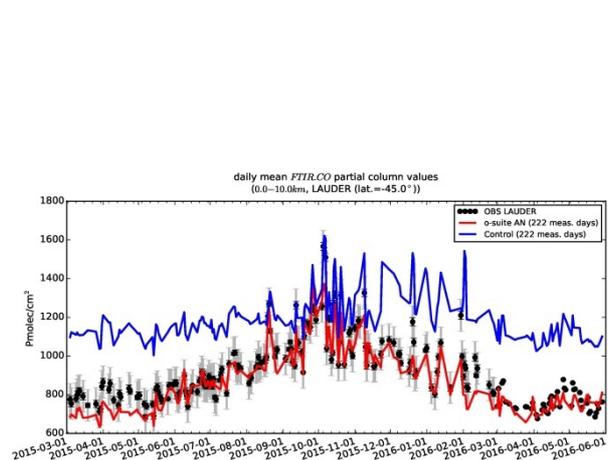
In this section, we compare the CO profiles of the CAMS models with FTIR measurements at Maido (21°S, 55°E, i.e. southern tropics, altitude 2.2km), Lauder (46°S, 169.7°E, altitude 370m) and Jungfraujoch (47°N, 8°E, Alpine station, altitude 3.6km). These ground-based, remote-sensing instruments are sensitive 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 and Fig. 3.3.11 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 Maido 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 30%-36%, clearly showing the positive effect of assimilation. At Jungfraujoch there is little difference between osuite and Control and the observed bias is comparable to the measurement's uncertainty.



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, Maito and Jungfraujoch (mean bias and stddev in %). The overall uncertainty for the CO measurements at Lauder and Maito is approximately 5%.

		JJA			SON			DJF			MAM		
		MB	stddev	nobs	MB	stddev	nobs	MB	stddev	nobs	MB	stddev	nobs
o-suite	Lauder	-1.51	4.57	120	-2.36	5.01	148	-8.53	5.27	93	-1.71	5.66	150
control	Lauder	28.24	10.45	120	14.13	7.92	148	36.82	12.2	93	45.87	5.17	150
o-suite	Maito	-7.37	4.06	125	-6.92	4.04	304	-9.07	3.25	117	-5.84	3.34	197
control	Maito	13.51	6.46	125	9.68	10.0	304	32.62	5.56	117	29.74	5.25	197
o-suite	Jungfraujoch	-3.22	6.22	98	-5.53	7.28	87	-10.23	4.69	57	0.17	6.16	91
control	Jungfraujoch	-6.72	8.86	98	-6.82	6.19	87	5.93	3.13	57	0.80	5.73	91



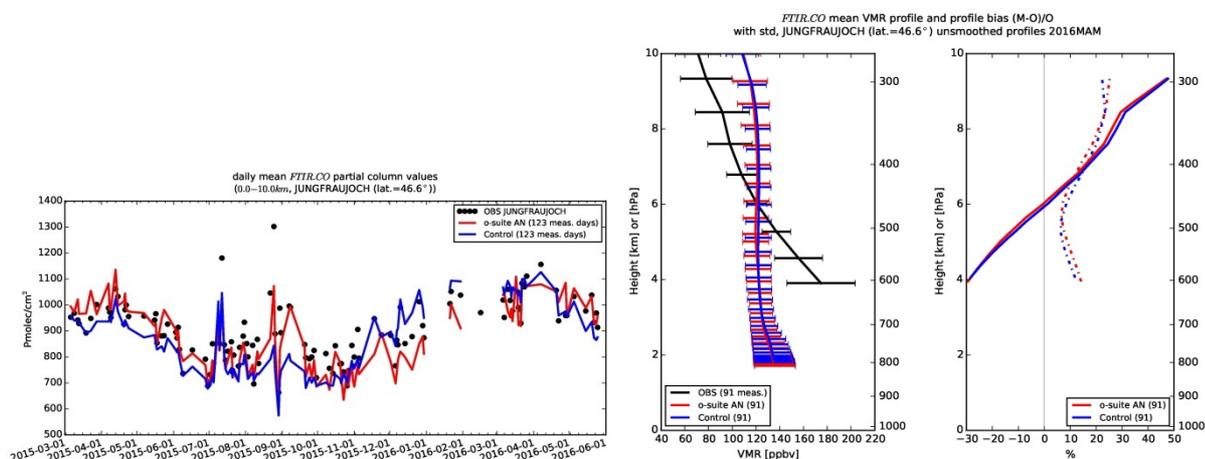


Figure 3.3.11: Daily mean values of tropospheric CO columns (till 10km) by the o-suite (red) and the Control run (blue) compared to NDACC FTIR data at Lauder, New Zealand (45°S, 169.7°E) (top), Mado (21°S, 55°E) (middle) and Jungfraujoch (47°N, 8°E) (bottom) for the period March 2015-May 2016. The unsmoothed profiles are averaged over March-April-May 2016. The number of measurement days is indicated in the legend.

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 the relative bias of model runs with respect to MOPITT V5, averaged for May 2016. MOPITT and IASI show relatively high values over Nepal and East Asia. Some difference between observations can be seen over the North Pacific Ocean and biomass burning area in Africa, indicating higher values in IASI compared to MOPITT.

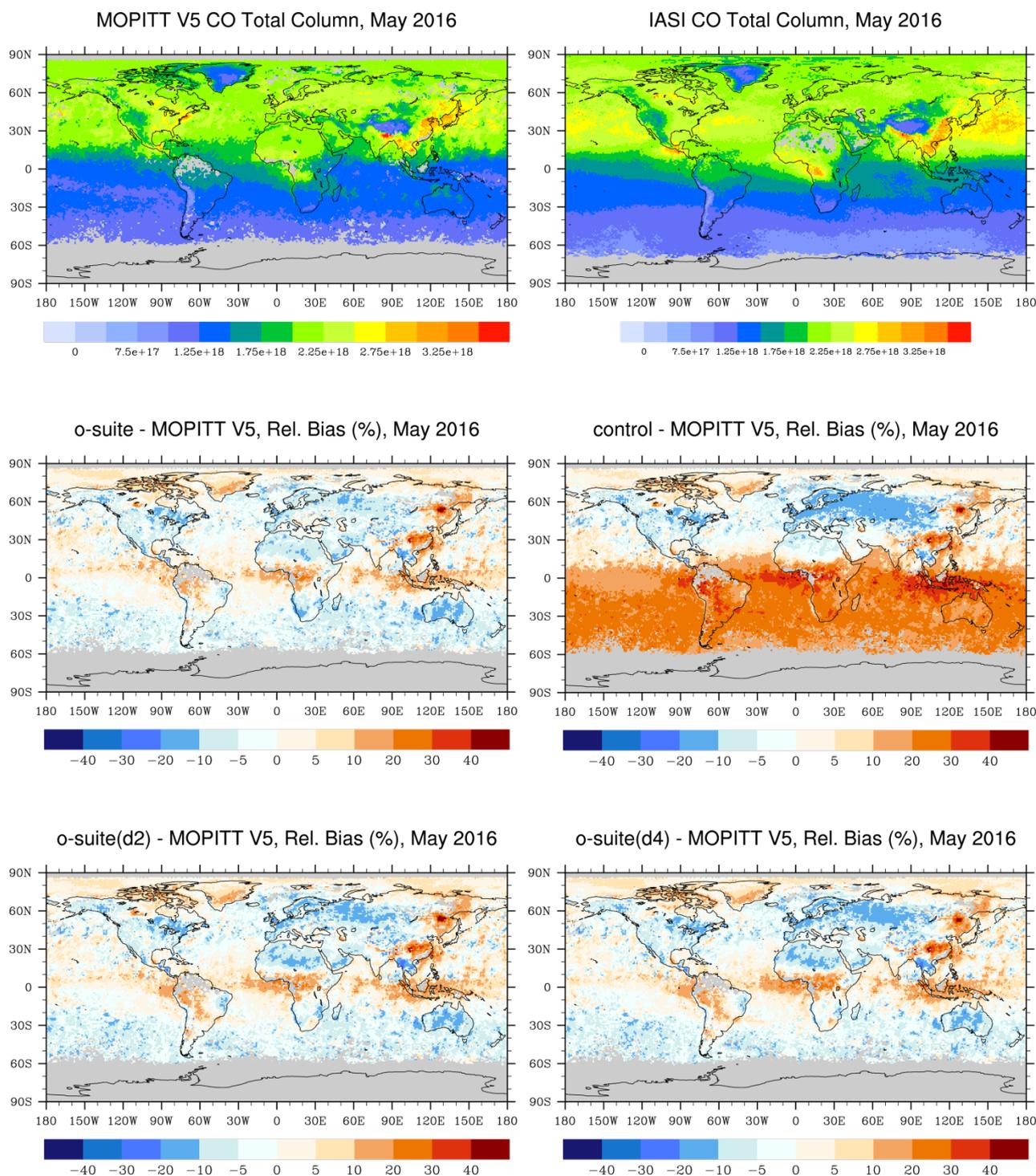


Fig. 3.3.12: CO total column for MOPITT V5 (top left) and IASI (top right) satellite retrievals and relative difference between the model runs and MOPITT for May 2016: o-suite (middle left), control run (middle right), o-suite 2nd forecast day (bottom left), o-suite 4th forecast day (bottom right). Grey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

The modeled CO geographical distribution and magnitude 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 south-west part of Russian Far East up to 40 % and also over the North of China up to 30%. In general, o-suite performs better than the control run, with some overestimation in the tropics and Northern Polar region and underestimations in the mid-latitudes of up to 20 %. The control run overestimates CO total columns over the Southern Hemisphere with stronger overestimation over Indonesia and biomass burning area in Africa by about 30%. Figure 3.3.12 shows no significant difference between the o-suite analysis and 2nd and 4th forecast days.

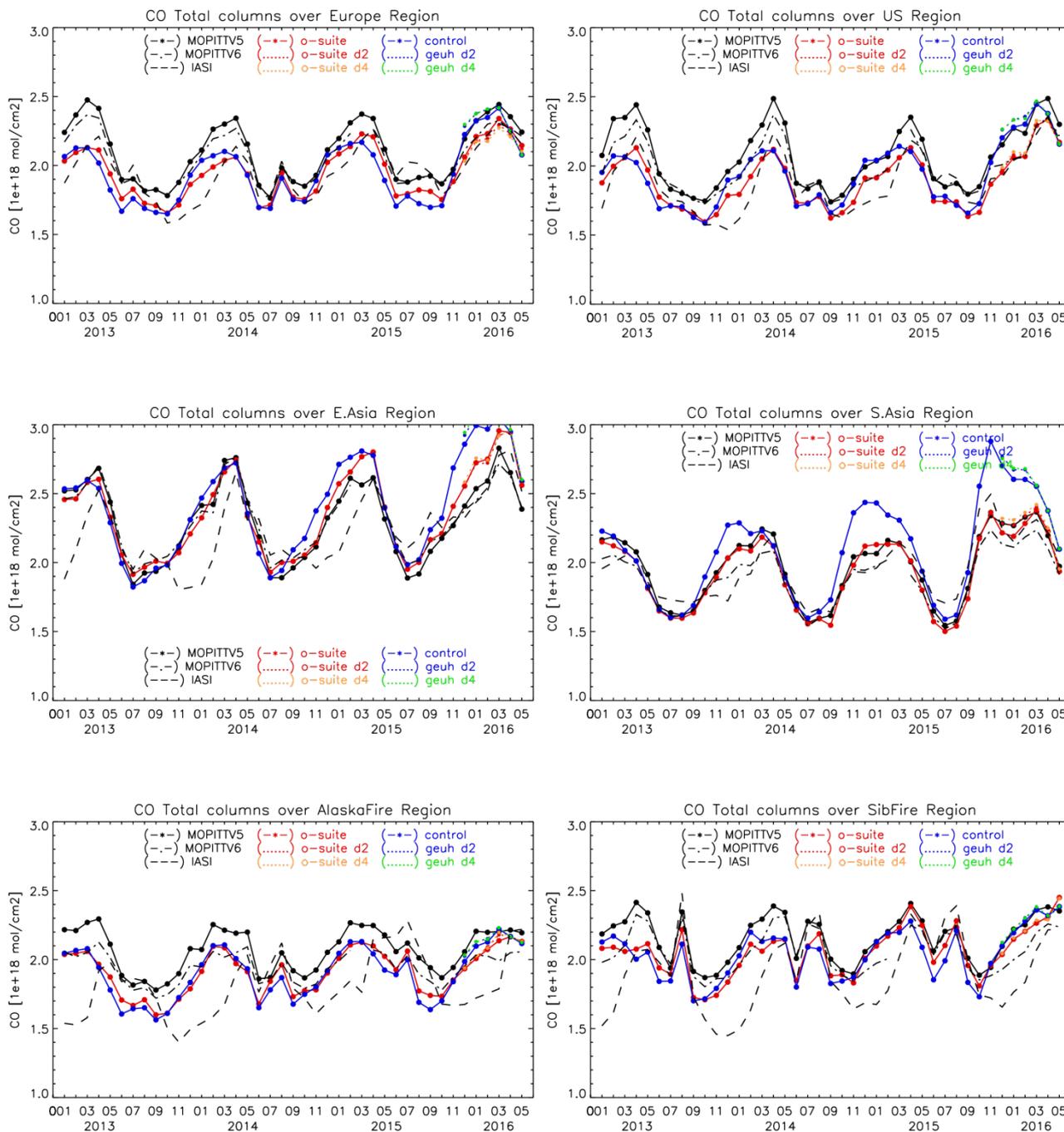
Figure 3.3.13 shows time series of CO total columns for MOPITT V5 and V6, IASI and the model runs over selected regions. For the comparison with MOPITT, the modelled CO concentrations were transformed using MOPITT V5 averaging kernels (Deeter, 2004). Both, MOPITT and IASI CO total column are assimilated 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 o-suite shows better agreement with the satellite retrievals over Europe and US regions, especially in the seasonal maximum in spring. Improvements can also be seen in Siberian fire region and North Africa.

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, 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 the South Africa the control run overestimates total column (satellite retrieved values) up to 25% in the second maximum in winter and in seasonal minimum in spring 2015, 2016. Compared to the last spring, both model runs show better agreement with the satellite observations over Alaska Fire region with bias within +/-5%.

The o-suite 2nd and 4th forecast days were also verified in this report. In Europe the o-suite 2nd and 4th forecast days show up to 2% higher negative bias compared to the analysis. In the biomass burning region in North Africa a rapidly increasing negative bias can be found in the 2nd and especially 4th forecast days and shows up to 5% difference with the analysis in winter 2016. In South Asia the 2nd and 4th days of forecast show up to 3% higher positive bias compared to analysis in winter season. In other regions the forecasts are almost identical to analysis (within 1% difference).



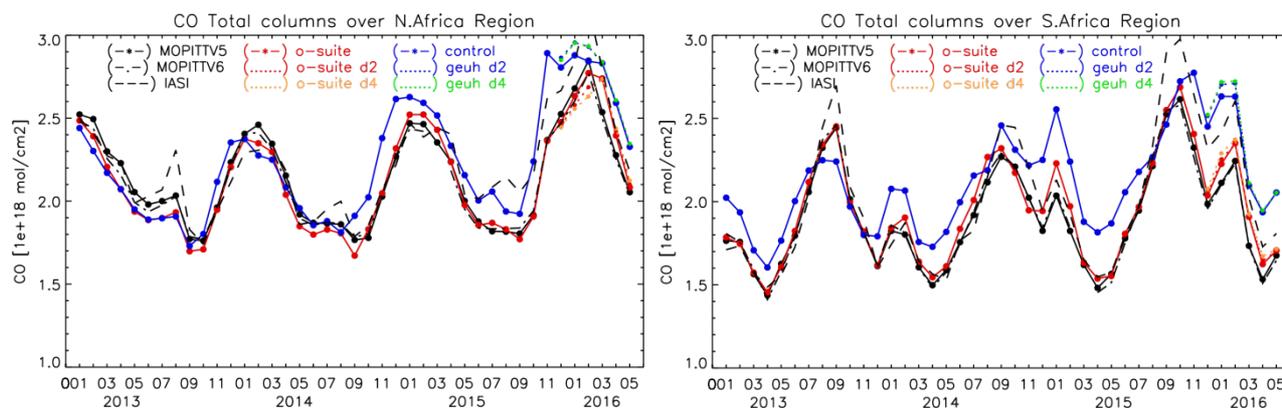
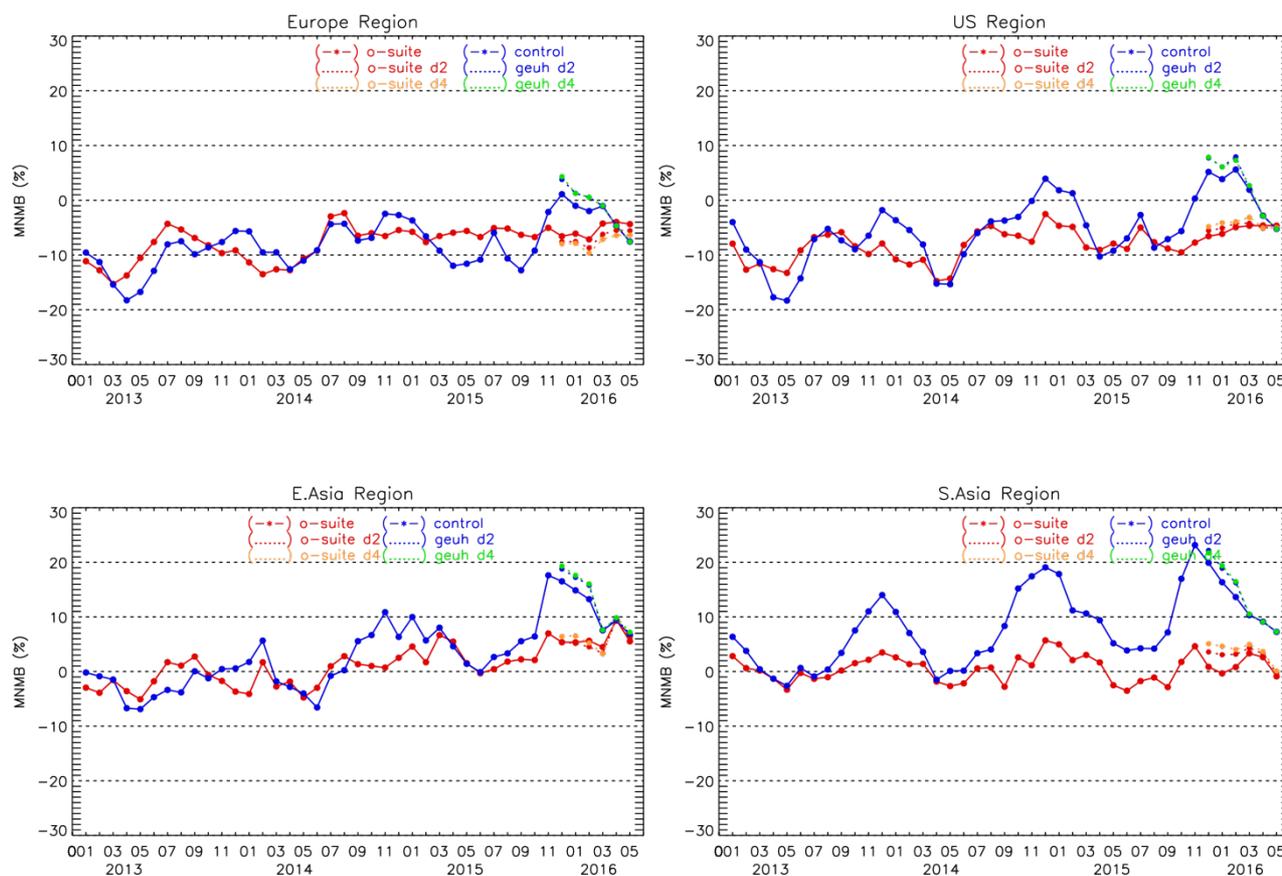


Fig. 3.3.13: Time series of CO total column for satellite retrievals MOPIT V5 and V6, IASI (black) and the model runs over the selected regions: o-suite (red, solid), control (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue,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



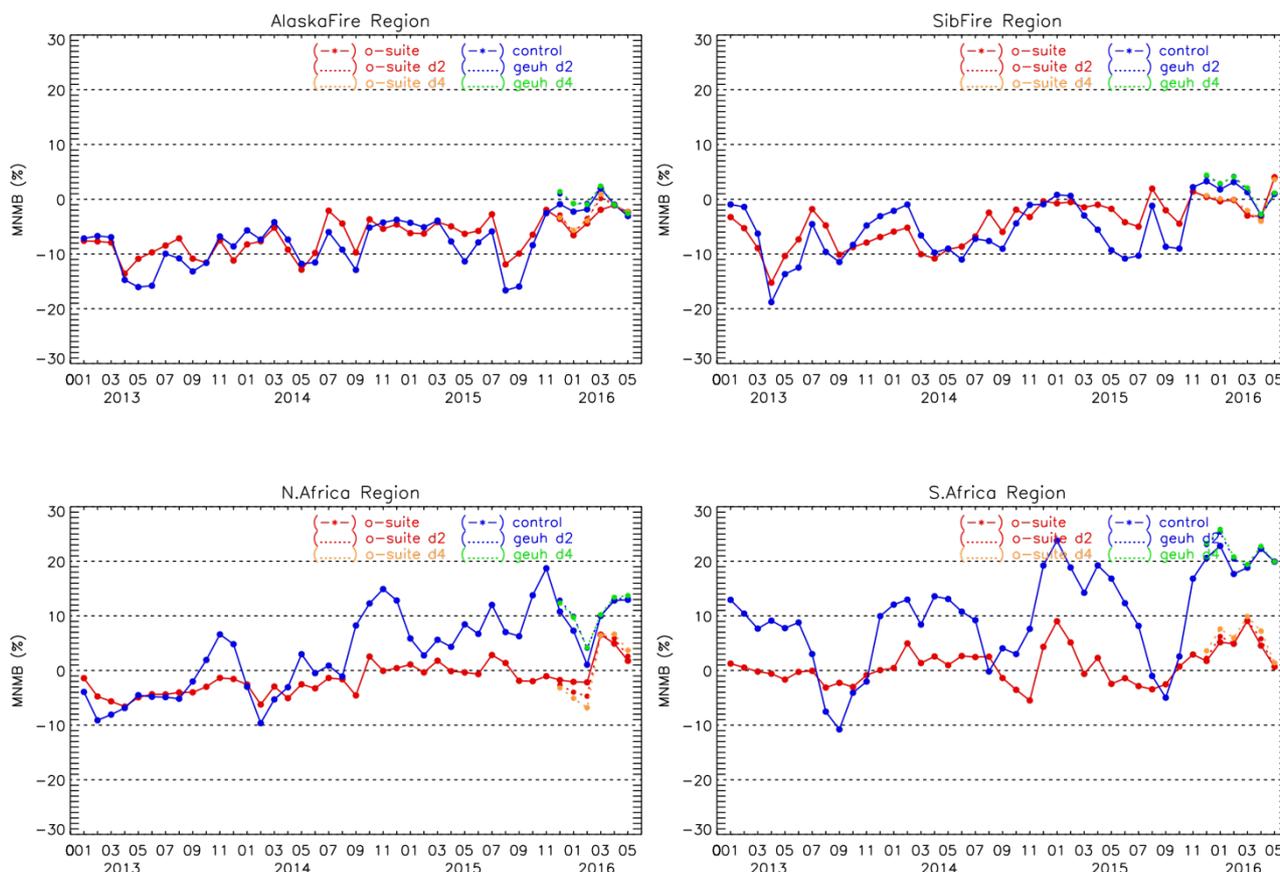


Fig. 3.3.14: Modified normalized mean bias (%) for CO total column from the model simulations vs MOPITT V5 retrievals over selected regions. O-suite (red, solid), control run (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue, 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 high-resolution CO simulations are compared to the o-suite, the control run as well as to the observations. At Bialystok and Orleans the control and high resolution FC model simulations agree within 10% with the measurements (Fig. 3.3.15 and 3.3.16). The seasonality is in general well represented by these models. The o-suite model shows the strongest discrepancies to the measurement for the period March-June 2015, where the modeled values are up to 25% too high.

At Reunion (3.3.17) the o-suite captures the seasonality and agrees with the measurements within 5-10%. 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 worsen again and the offset is similar to the period March 2015 – July 2015.

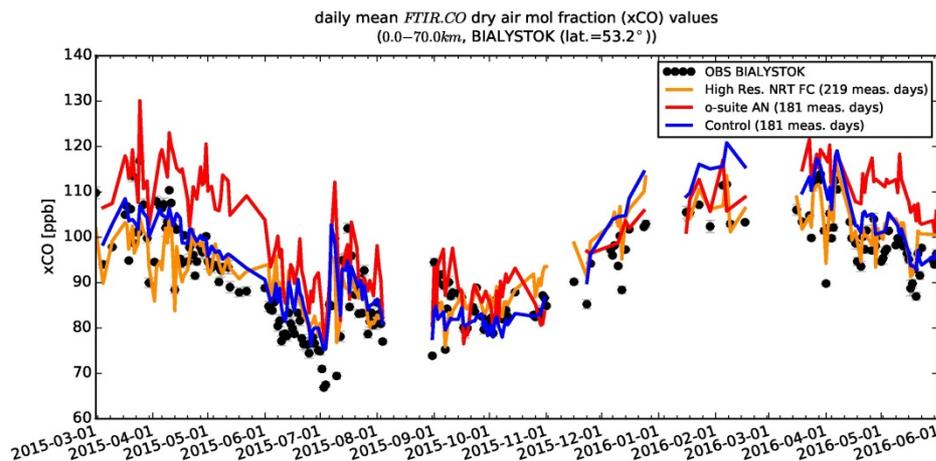


Figure 3.3.15: Time series of column averaged mole fractions of carbon monoxide (CO) at the TCCON site Bialystok compared to the o-suite (red), control (blue) and the high resolution NRT FC model (yellow).

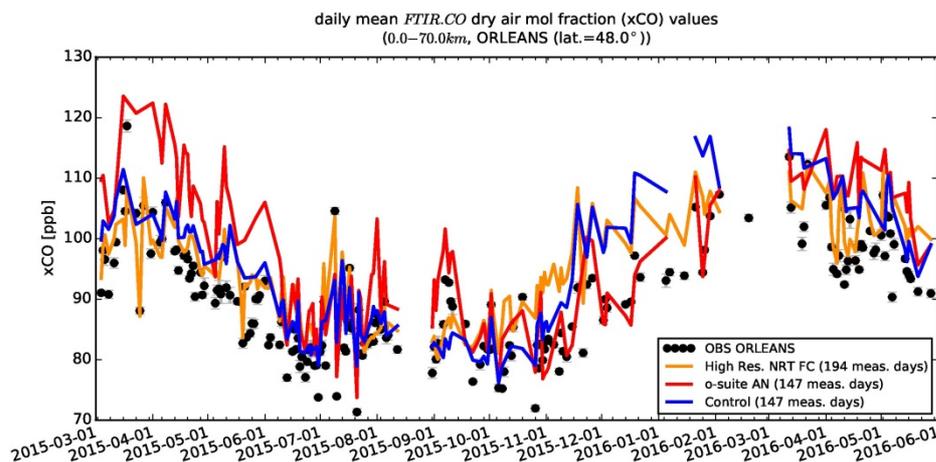


Figure 3.3.16: Time series of column averaged mole fractions of carbon monoxide (CO) at the TCCON site Orleans compared to the o-suite (red), control (blue) and the high resolution NRT FC model (yellow).

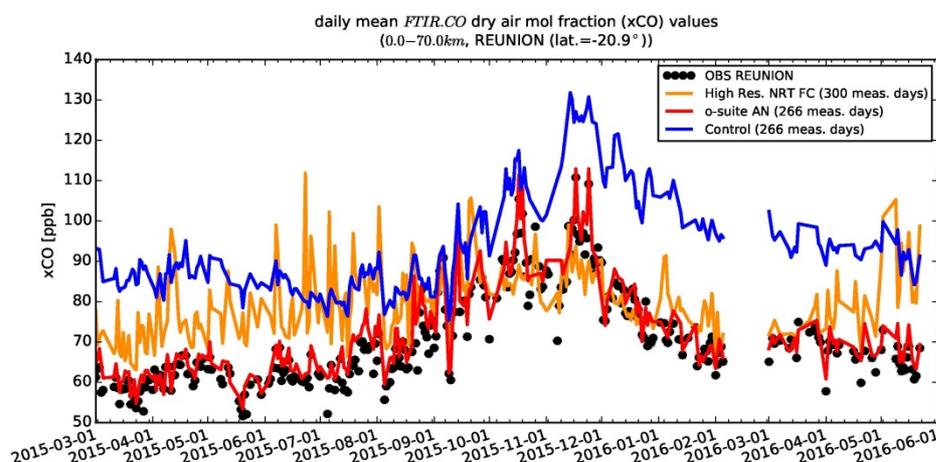


Figure 3.3.17: Time series of column averaged mole fractions of carbon monoxide (CO) at the TCCON site Reunion compared to the o-suite (red), control (blue) and the high resolution NRT FC model (yellow).



3.4 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 , monthly mean satellite HCHO columns are compared to model results for May 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.

Time series in Fig. 3.4.2 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 tend to 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. Over East-Asia, 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.
- 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 runs in the latest o-suite. MOZART-based simulations and observations agree reasonably well since 2012.
- 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.

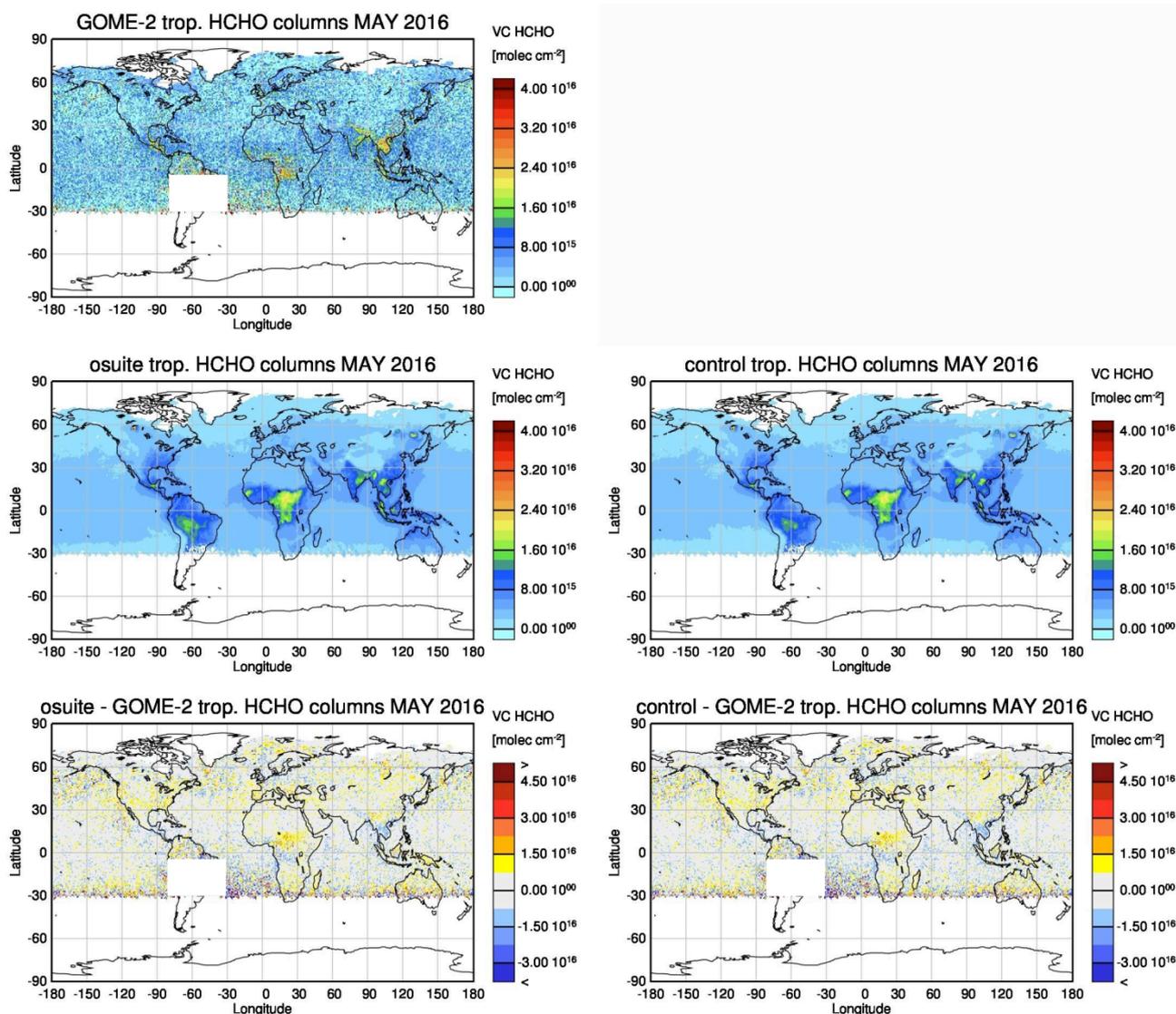


Figure 3.4.1: Global map comparisons of satellite retrieved and model simulated tropospheric HCHO columns [molec cm^{-2}] for May 2016. The top row shows monthly mean tropospheric HCHO columns retrieved by GOME-2, the second row shows the same but for model simulated averages. The third row shows differences of monthly means between models and GOME-2. GOME-2 data were gridded to model resolution (i.e. $0.75^\circ \text{ deg} \times 0.75^\circ \text{ deg}$). 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).

Details on the HCHO evaluation can be found at:

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

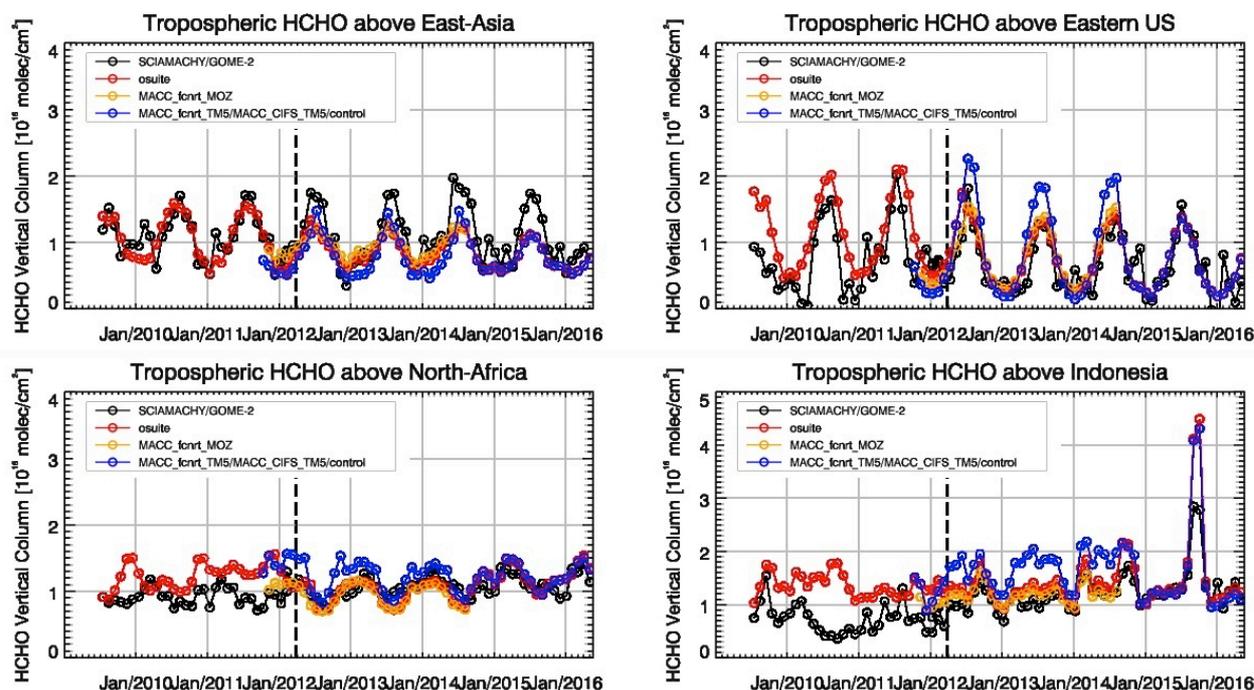


Figure 3.4.2: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions. The blue line shows MACC_fcprt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results 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 Uccle (50.8°N, 4.36°E, urban). This ground-based, remote-sensing instrument is 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 Figure 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 a priori. 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).

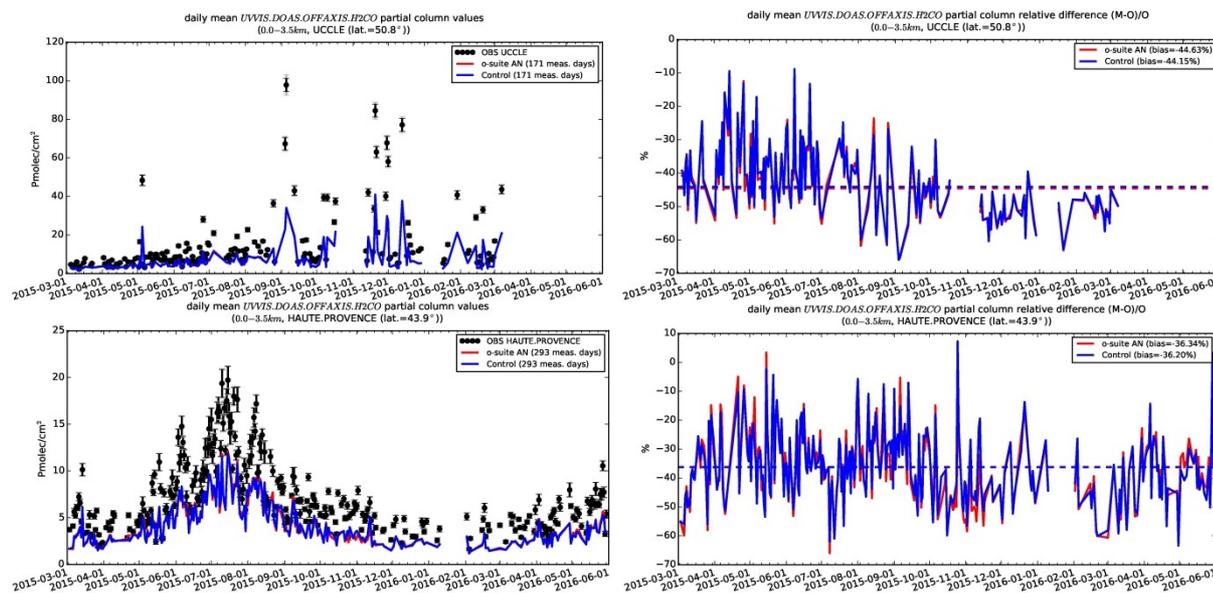


Figure 3.4.3: Daily mean relative differences of tropospheric HCHO columns (till 3.5km) by the o-suite (red) and the control run (blue) compared to NDACC UVVIS DOAS data at Uccle (50.8°N, 4.36°E, middle) and Haute Provence (43.9°N, 5.71°E, rural station, altitude 650m, bottom) for the period March. 2015 – March 2016. The number of measurements and median of differences is indicated in the legend (the overall measurement uncertainty is 10%).

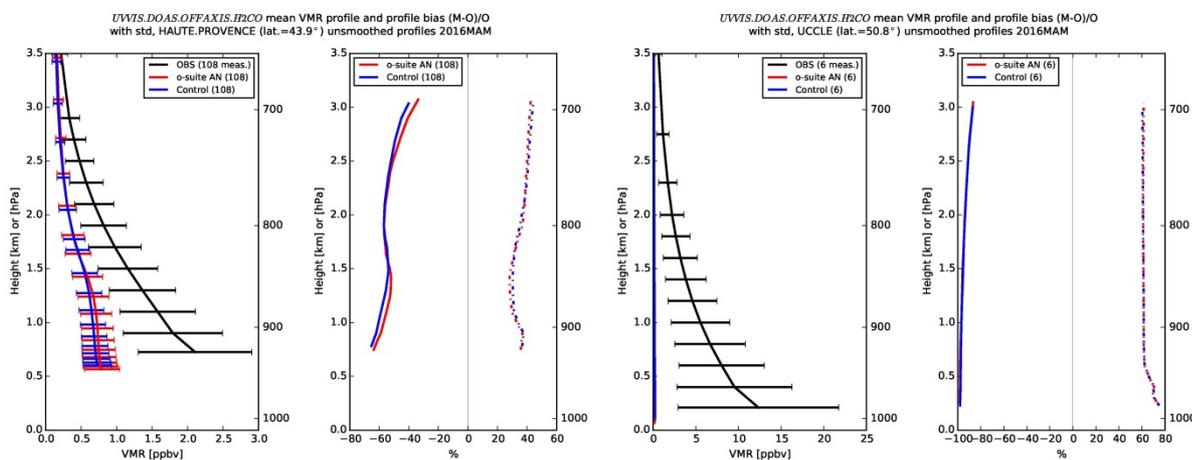


Figure 3.4.4: Mean tropospheric HCHO profiles by the o-suite (red) and the control run (blue) compared to NDACC UVVIS DOAS data at Haute Provence (43.9°N, 5.71°E, left) and Uccle (50.8°N, 4.36°E, right) for the period March-May 2016.

From Figure 3.4.3 and Figure 3.4.4 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 Figure 3.4.4). Due to instrument failure at UCCLE, the number of measurement in MAM 2016 is limited.



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 have a 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. In winter 2014/2015 the CAMS model MNMB bias against level 2.0 data was +5% higher than that against level 1.5 data. Figure S3 shows the evaluation against level 2.0 data for the whole time period. Note that the establishment of such correction of bias in the last months is rather difficult because of few level 2.0 data being available.

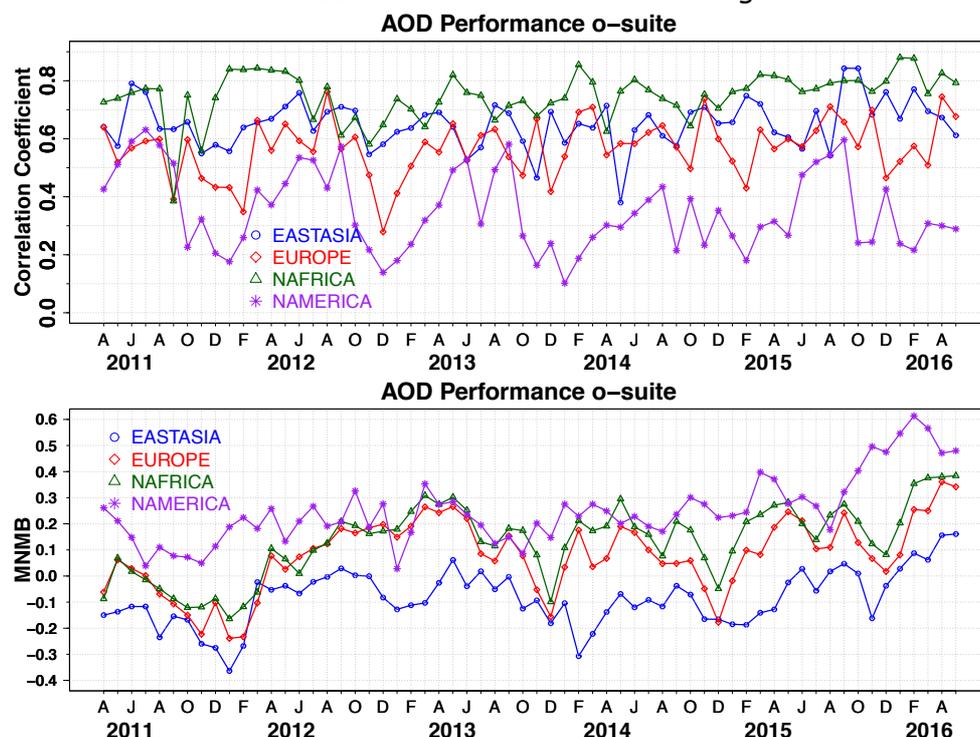


Figure 3.5.1: a) Correlation coefficient and b) modified normalized mean bias (MNMB) in AOD, since 2011, based on daily AOD comparison in four world regions [Eastasia(blue); Europe(red); NAfrica(green); NAMercia(purple)] for the o-suite.

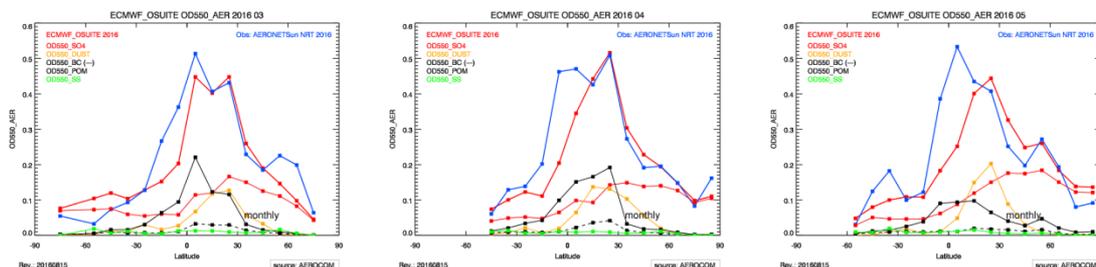


Figure 3.5.2: Aerosol optical depth of o-suite (red) compared to latitudinally aggregated NRT Aeronet level 1.5 data (blue) for the three months covered by this report.

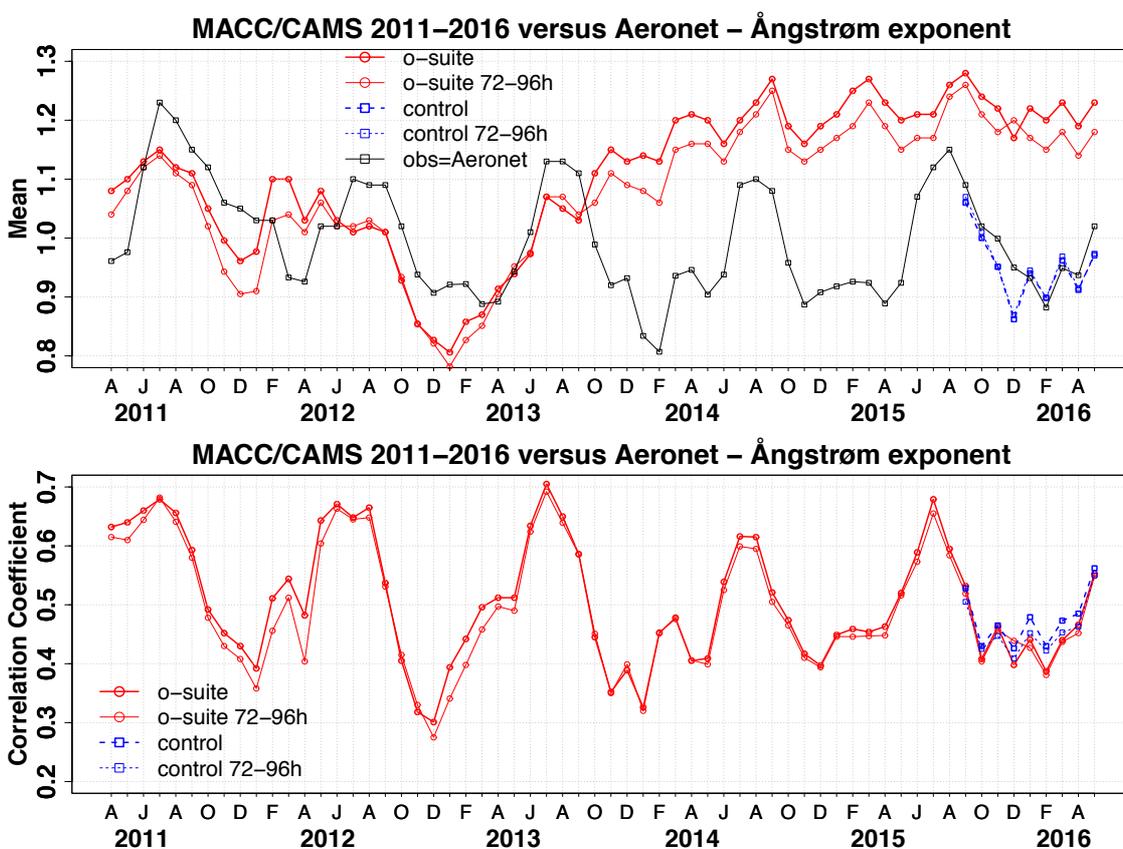


Figure 3.5.3 a) Evolution of mean Ångström exponent in o-suite and control at Aeronet sites, based on matching monthly mean values. o-suite (thick red curve); o-suite at last forecast day (light red curve); control (blue dashed curve); control at last forecast day (light blue dashed curve). b) Correlation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

The regional performance of the o-suite model exhibits some seasonal cycle in AOD depending on region (3.5.1a). 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 America (3.5.1b). The latitudinal display of model and Aeronet AOD in the period investigated here (3.5.2) shows the negative bias against Aeronet NRT in tropical and sub-tropical regions.

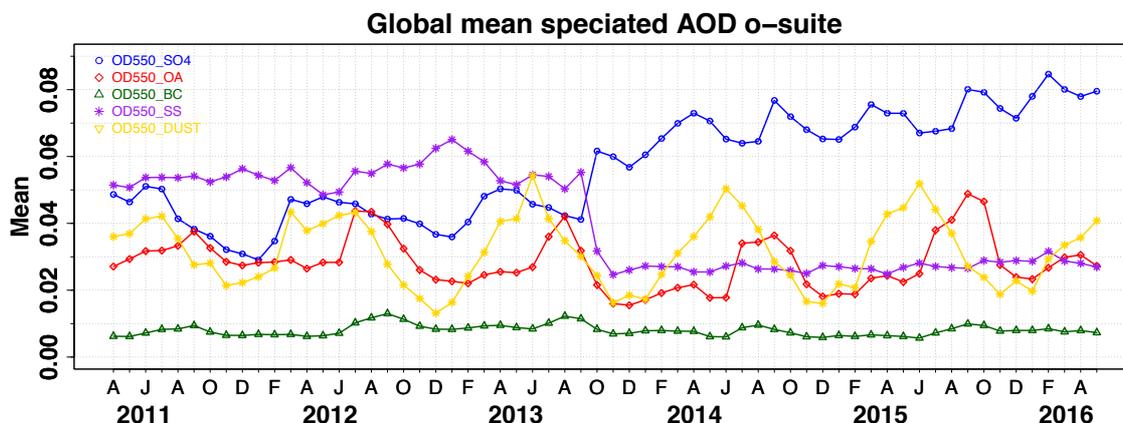


Figure 3.5.4: Evolution of aerosol component’s AOD@550nm [OD550_SO4 = sulphate(blue); OD550_OA = organics(red); OD550_BC = black carbon(green); OD550_SS = sea salt(purple); OD550_DUST = dust(yellow)].

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. 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 a positive bias of +40% (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 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%.

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 DJF 2015/16 0-24h	Change wrt to first day on day 4	Mean MAM 2016 0-24h	Change wrt to first day on day 4
AOD@550	0.165	-16%	0.181	-15%
BC-OD@550	0.008	-25%	0.008	-14%
Dust-OD@550	0.024	8%	0.037	-11%
OA-OD@550	0.025	-16%	0.029	-10%
S04-OD@550	0.078	-24%	0.079	-26%
SS-OD@550	0.030	-6%	0.028	0%

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 repeatedly use the latest GFAS values) determine more and more the shape and amplitude of the aerosol fields. The forecasted AOD fields have been used to establish global mean aerosol optical depth and forecast performance after three days (see comparison to first guess in Figure S3 in summary of this report) at Aeronet sites. Table 3.5.1 shows an average global decrease in total aerosol optical depth of 15% during the first four forecast days. The contributions to this reduction stem from almost all anthropogenic aerosol components, less so from sea salt and dust. Against Aeronet the o-suite forecast for day three has little overall bias in AOD (see figure S3). The control run with no assimilation shows significant less AOD (-50% compared to o-suite, see figure S3), 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 155 remote IMPROVE and EMEP stations (Figure 3.5.5). 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. The bias maps show that both in North America and Europe high bias appears at stations located in regions close to the coastlines. This is an indication that simulated PM₁₀ concentrations are high due to sea salt aerosols. Inner-continental sites have a small to moderate negative bias.

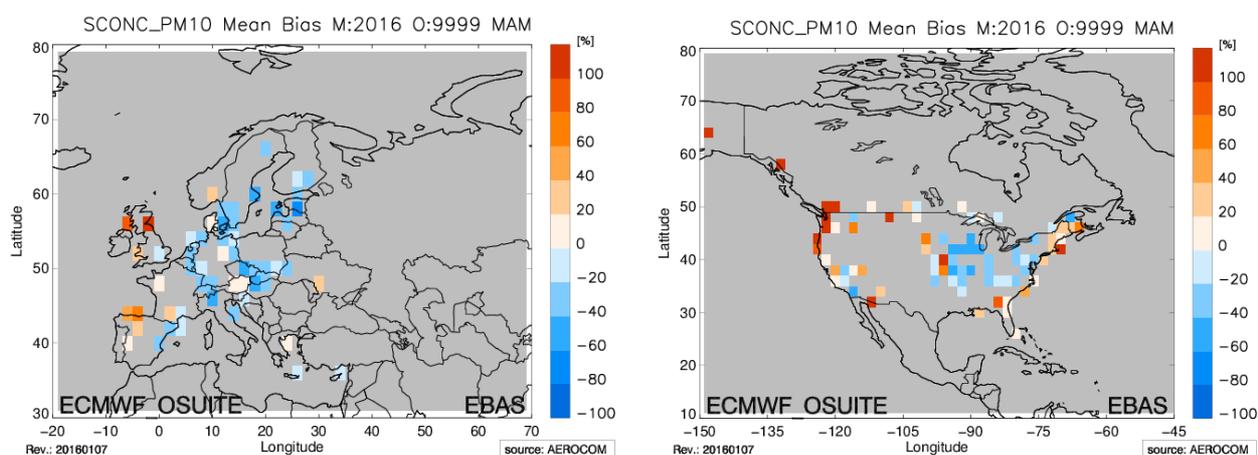


Figure 3.5.5: Bias [%] map of March/April/May mean PM 10 concentrations at EMEP (Europe) and IMPROVE sites (North America); simulated o-suite versus climatological average (2000-2009).

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

Daily dust aerosol optical depth (DOD) from CAMS o-suite and the control experiment have been validated against 71 AERONET stations grouped in twelve regions for the period 1 March – 31 May 2016 (Figure 3.5.6) to evaluate their performances over the dust source and transport regions in spring season. In this report, we validate the forecasts using AERONET (Holben et al., 1998) as reference, and compare with daily Multi-model DOD (at 550 nm) Median from nine dust prediction models. The Multi-model Median product is processed at 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/>).

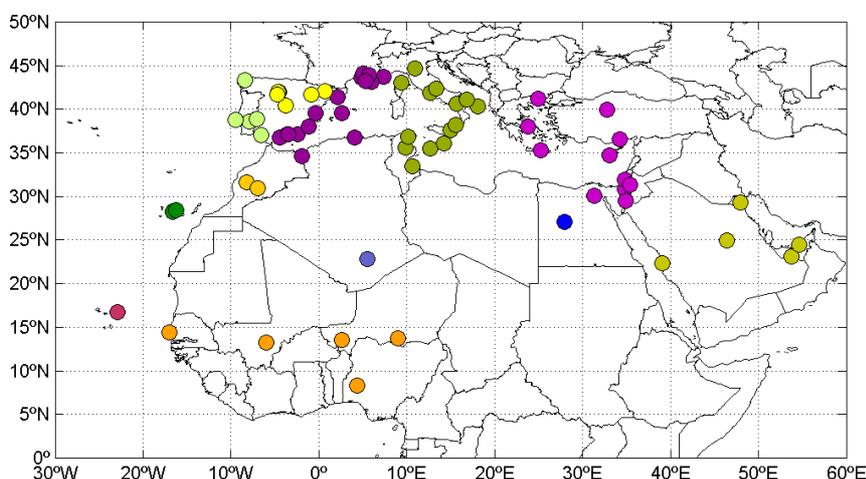


Figure 3.5.6: Map of 71 AERONET level-1.5 stations used in this analysis. The twelve regions considered in the analysis are shown by different colours.

We used AOD observations at 550 nm from 61 AERONET whose locations are depicted in Figure 3.5.6. Cloud-screened direct-sun data (Level 1.5) between 440 and 870nm, which contain an uncertainty about 0.02 for AOD under cloud-free conditions, are used. Quantitative evaluations of the modelled dust AOD are conducted for dust-dominated conditions; i.e. when the Angström exponent (AE) is ≤ 0.75 . All data with AE larger than 1.2 are considered to be free of dust (DOD = 0 is assumed). Values of AE between 0.75 and 1.2 are associated with mixed aerosols and are not included in the analysis. The AOD at 550 nm is derived from data between 440 and 870 nm following the Ångström's law. Because AERONET data are acquired at 15-min intervals on average, all measurements within ± 90 min of the models' outputs are used for the 3-hourly evaluation. Finally, DOD has been qualitatively compared with AOD from 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.

72 hour forecasts (on 3-hourly basis) of DOD from AERONET, o-suite, control and SDS-WAS Multi-model Median for the period 1 March 2016 – 31 May 2016 have been computed for twelve study regions shown in Figure 3.5.6. Mean Bias (MB), Fractional Gross Error (FGE), Root Mean Square Error (RMSE), Person correlation coefficient (r),

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	NDATA	MB	FGE	RMSE	r
Western Mediterranean	2181	-0.03	1.76	0.21	0.49	-0.04	1.78	0.22	0.50	2181	-0.04	1.77	0.22	0.48
Tropical North Atlantic	173	-0.05	0.44	0.15	0.74	-0.11	0.59	0.18	0.74	173	-0.14	0.73	0.21	0.72
Eastern Mediterranean	1483	0.01	1.05	0.24	0.62	-0.06	1.05	0.22	0.66	1481	-0.07	1.01	0.22	0.68
Sahel	1630	-0.13	0.37	0.38	0.55	-0.20	0.41	0.41	0.57	1628	-0.21	0.39	0.41	0.60
Subtropical North Atlantic	466	-0.01	1.79	0.09	0.65	-0.01	1.78	0.08	0.68	466	-0.01	1.79	0.08	0.69
Central Mediterranean	2100	0.00	1.14	0.29	0.64	-0.05	1.16	0.27	0.66	2099	-0.08	1.09	0.27	0.67
Middle East	1056	0.01	0.46	0.25	0.53	-0.04	0.39	0.24	0.58	1056	-0.08	0.39	0.25	0.56
Iberian Peninsula	926	-0.02	1.98	0.19	0.07	-0.02	1.98	0.19	0.06	926	-0.03	1.97	0.19	0.11
Western Iberian Peninsula	592	-0.09	1.96	0.20	0.12	-0.09	1.96	0.20	0.09	592	-0.09	1.97	0.20	0.13
North Western Maghreb	460	0.00	1.47	0.19	0.49	-0.02	1.43	0.16	0.55	458	-0.04	1.41	0.16	0.54
Sahara	357	0.02	0.39	0.31	0.67	-0.04	0.39	0.33	0.63	355	-0.11	0.40	0.34	0.66
Eastern Sahara	302	0.08	0.46	0.25	0.66	0.00	0.38	0.20	0.67	298	-0.01	0.35	0.18	0.73

and the number of data (NDATA), averaged over the study period for o-suite, control and SDS-WAS Multi-model Median, and for the twelve regions of study are shown in Table 3.5.2 and Table 3.5.3.

During the period of analysis, MODIS (see Figure 3.5.7) shows that major dust activity is concentrated over the Sahara (in the Bodelé Basin and the Mali/Mauritania border) and the Arabian Peninsula. CAMS o-suite is able to simulate the main areas of dust activity although it overestimates the observed AOD over southern Saudi Arabia and Libya, and underestimates the dust outflow over Dakar and Capo Verde and the high AOD values observed over Iraq and Syria.

From March to 31 May 2016, CAMS o-suite is the model that best reproduces the daily variability of AERONET observations (see correlation in Table 3.5.2). CAMS o-suite shows lower MB values than CAMS control over Middle East and Sahara achieving values similar to the SDS-WAS Median Multimodel (see Table 3.5.2). In terms of correlation, o-suite improves the observed variability over Middle (from 0.53 for control to 0.58 for o-suite, see Table 3.5.2) and reduces the correlation over the Sahara (from 0.67 for control to 0.63 for o-suite, see Table 3.5.2). It is noteworthy that in general, CAMS experiments show a better performance reducing the systematic overestimations observed in the old MACC model over desert dust sources (Sahara and Middle East).

Over long-range transport regions, o-suite shows the best correlations in most of the regions (with correlations above 0.50, see Table 3.5.2) in comparison with CAMS control and the SDS-WAS Median Multimodel except in the Iberian Peninsula ($r < 0.12$, see Table 3.5.2). Maximum dust activity is observed in Central and Eastern Mediterranean (see IMS-METU-ERDEMLI and Ben Salem in Figure 3.5.8). On the other hand, in the Western Mediterranean (see Palma de Mallorca in Figure 3.5.9), the moderate correlations are consistent with the presence of dust mobilized by mesoscale

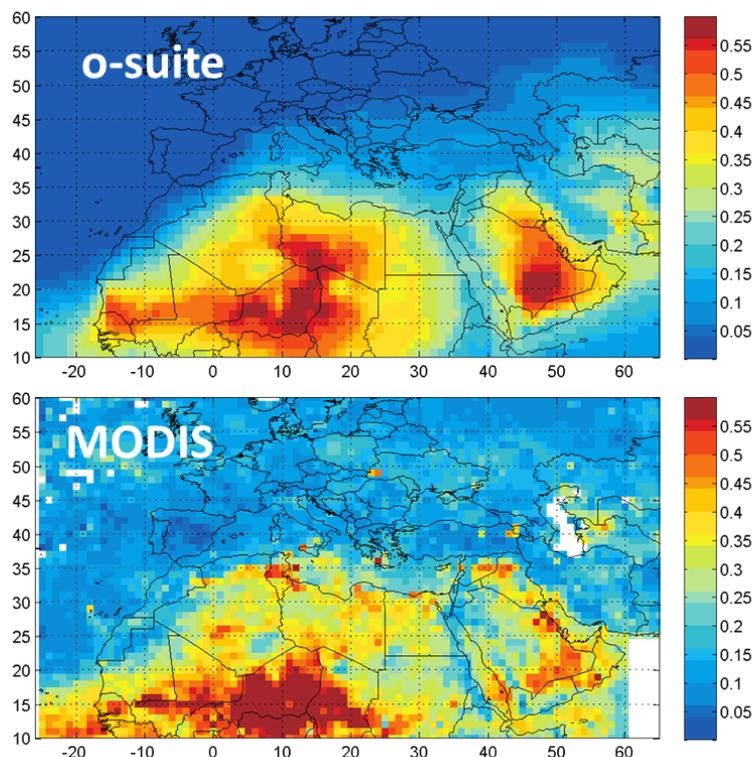


Figure 3.5.7: Averaged DOD from o-suite (top) and AOD from MODIS combined Dark target and Deep Blue product (bottom) from March 1 to May 31, 2016.

convective systems (MCSs) associated to wet events. Tropical North Atlantic (see Dakar in Figure 3.5.8) and sub-Tropical North Atlantic are the regions where the highest correlations are achieved ($r > 0.68$). Poor correlations in the Iberian Peninsula is mainly associated to low number of dust events and low AOD values (~ 0) observed in this area during all the period of study (see Cabo da Roca in Portugal in Figure 3.5.9).

Over long-range transport regions, o-suite shows the best correlations in most of the regions (with correlations above 0.50, see Table 5.3.2) in comparison with CAMS control and the SDS-WAS Median Multimodel except in the Iberian Peninsula ($r < 0.12$, see Table 5.3.2). Maximum dust activity is observed in Central and Eastern Mediterranean (see IMS-METU-ERDEMLI and Ben Salem in Figure 3.5.8). On the other hand, in the Western Mediterranean (see Palma de Mallorca in Figure 3.5.9), the moderate correlations are consistent with the presence of dust mobilized by mesoscale convective systems (MCSs) associated to wet events. Tropical North Atlantic (see Dakar in Figure 3.5.8) and sub-Tropical North Atlantic are the regions where the highest correlations are achieved ($r > 0.68$). Poor correlations in the Iberian Peninsula is mainly associated to low number of dust events and low AOD values (~ 0) observed in this area during all the period of study (see Cabo da Roca in Portugal in Figure 3.5.9).

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 (see Table 5.3.2 and Table 5.3.3), the correlation is reduced drastically from 48h to 72h in all the regions except Sahara and the North Atlantic regions (see Table 5.3.2)



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	MB	FGE	RMSE	r	MB	FGE	RMSE	r	MB	FGE	RMSE	r
Western Mediterranean	2181	-0.03	1.81	0.24	0.32	-0.04	1.83	0.24	0.31	-0.03	1.85	0.25	0.17	-0.04	1.86	0.25	0.18
Tropical North Atlantic	173	-0.05	0.57	0.19	0.59	-0.11	0.69	0.21	0.55	-0.08	0.64	0.22	0.47	-0.12	0.78	0.24	0.41
Eastern Mediterranean	1483	0.01	1.25	0.32	0.38	-0.04	1.26	0.29	0.38	0.03	1.36	0.40	0.12	0.00	1.37	0.38	0.12
Sahel	1630	-0.14	0.41	0.41	0.49	-0.20	0.46	0.43	0.45	-0.16	0.48	0.44	0.36	-0.21	0.53	0.47	0.30
Subtropical North Atlantic	466	-0.02	1.82	0.10	0.48	-0.02	1.82	0.10	0.50	-0.02	1.84	0.12	0.15	-0.02	1.83	0.12	0.11
Central Mediterranean	2100	0.02	1.25	0.38	0.43	-0.01	1.25	0.37	0.42	0.03	1.35	0.46	0.18	0.00	1.34	0.45	0.18
Middle East	1056	0.01	0.54	0.29	0.38	-0.03	0.49	0.28	0.42	0.02	0.63	0.35	0.13	-0.01	0.61	0.34	0.15
Iberian Peninsula	926	-0.02	1.97	0.19	0.13	-0.02	1.97	0.19	0.12	-0.02	1.98	0.19	0.01	-0.02	1.98	0.19	0.01
Western Iberian Peninsula	592	-0.09	1.96	0.20	0.03	-0.09	1.95	0.20	0.03	-0.09	1.97	0.21	-0.08	-0.09	1.96	0.21	-0.08
North Western Maghreb	460	0.00	1.58	0.21	0.37	-0.01	1.56	0.20	0.36	-0.01	1.64	0.23	0.23	-0.01	1.63	0.22	0.20
Sahara	357	0.04	0.46	0.35	0.58	-0.01	0.45	0.35	0.56	0.05	0.48	0.36	0.55	0.01	0.47	0.36	0.53
Eastern Sahara	302	0.10	0.66	0.32	0.43	0.03	0.58	0.26	0.50	0.13	0.85	0.41	0.14	0.07	0.79	0.37	0.17

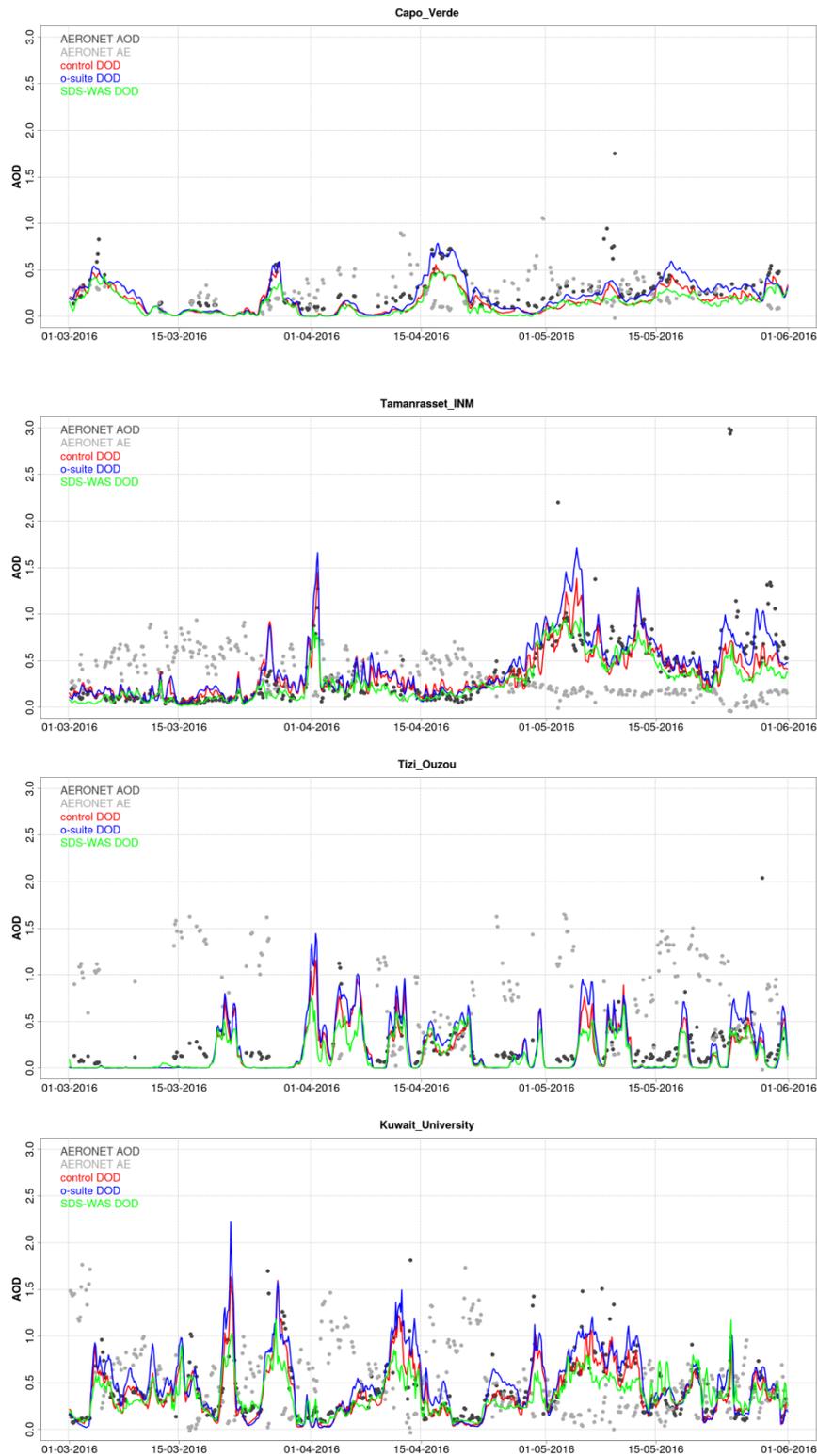


Figure 3.5.8: AOD from AERONET (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multimodel SDS-WAS Median (green line) for the period March 1st to May 31st, 2016 over Cape Verde (Tropical North Atlantic), Tamanrasset INM (Sahara), Tizi-Ouzou (NW Magrebh) and Kuwait University (Middle East).

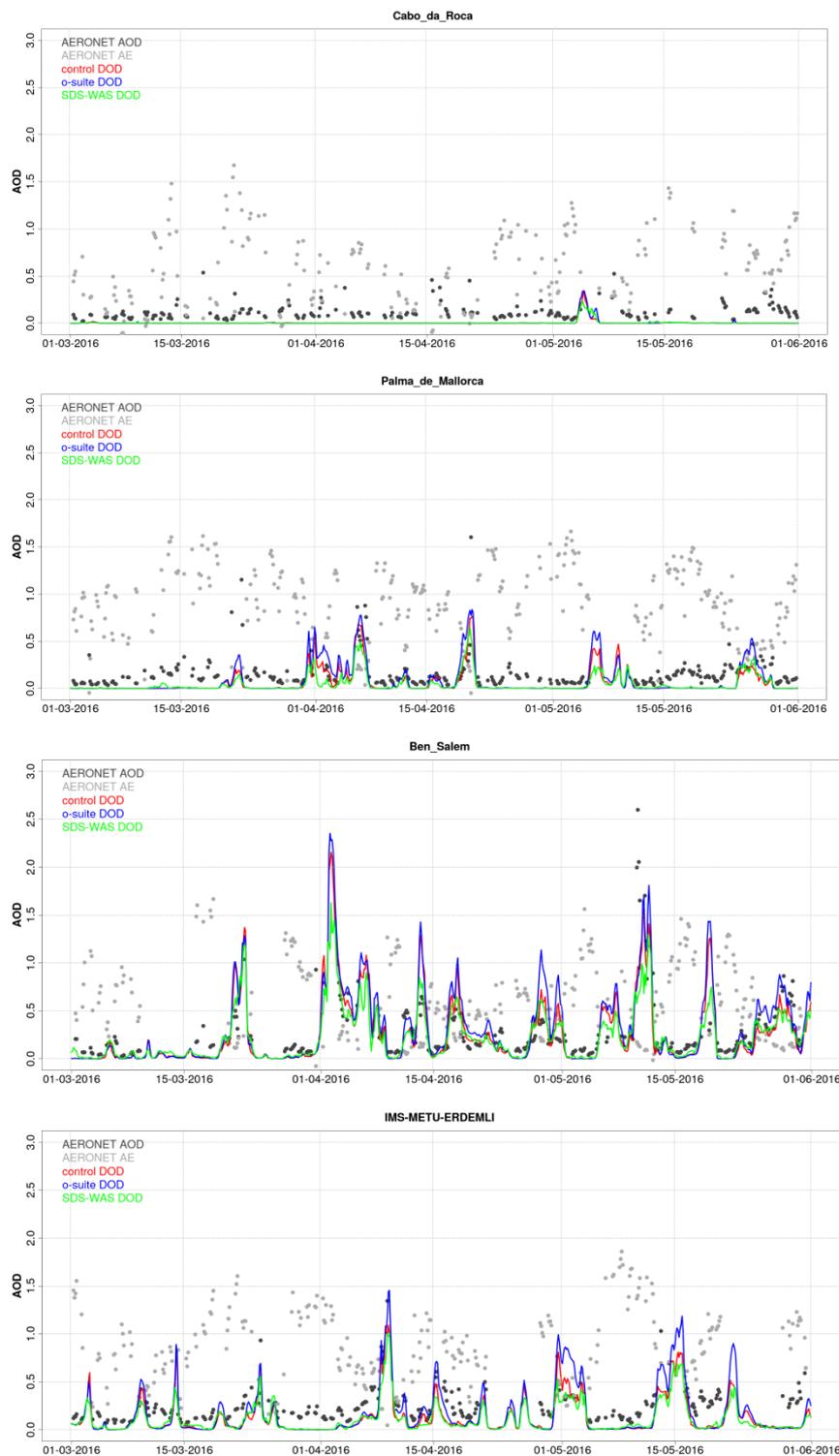


Figure 3.5.9: AOD from AERONET (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multimodel SDS-WAS Median (green line) for the period March 1st to May 31st, 2016 over Cabo da Roca (Portugal), Palma de Mallorca (Western Mediterranean), Ben Salem (Central Mediterranean), IMS-METU-ERDEMLI (Eastern 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 and its control run have been validated against 37 AERONET (Holben et al., 1998) and 18 Airbase stations in the Mediterranean region for the period 1 March – 31 May 2016. The main goal is to evaluate the ability of CAMS o-suite and control to reproduce AOD and surface concentration (PM_{2.5} and PM₁₀) over the Mediterranean in spring season when the maximum dust activity is observed in the Eastern Mediterranean.

Aerosol optical depth

The evaluation of the three-hourly values of AOD from the CAMS forecasts against AERONET is presented in Figure 3.5.10 and Figure 3.5.11, similar to the evaluation methodology presented in Section 3.5.2. Mean Bias (MB), Fractional Gross Error (FGE), Root Mean Square Error (RMSE), Person correlation coefficient (r), and the number of data (NDATA), averaged over the study period for o-suite and control, are computed for the 37 available stations.

From March to May 2016, CAMS o-suite best reproduces the daily variability of AERONET observations (see correlation in Figure 3.5.10). On average for all the sites, AOD MB increases from -0.01 for control to 0.03 for o-suite. In general, the control run tends to slightly underestimate the background levels compared with AERONET observations over Southern European sites (more influenced by urban/industrial aerosol sources), as indicated by systematic lower MB values in Figure 3.5.10. Otherwise, CAMS o-suite tends to overestimate the AOD background levels as is shown in Figure 3.5.11. The highest peaks of CAMS AOD forecasts are linked to natural sources (see Figure 3.5.11). During spring, maximum dust activity is observed in Central and Eastern Mediterranean.

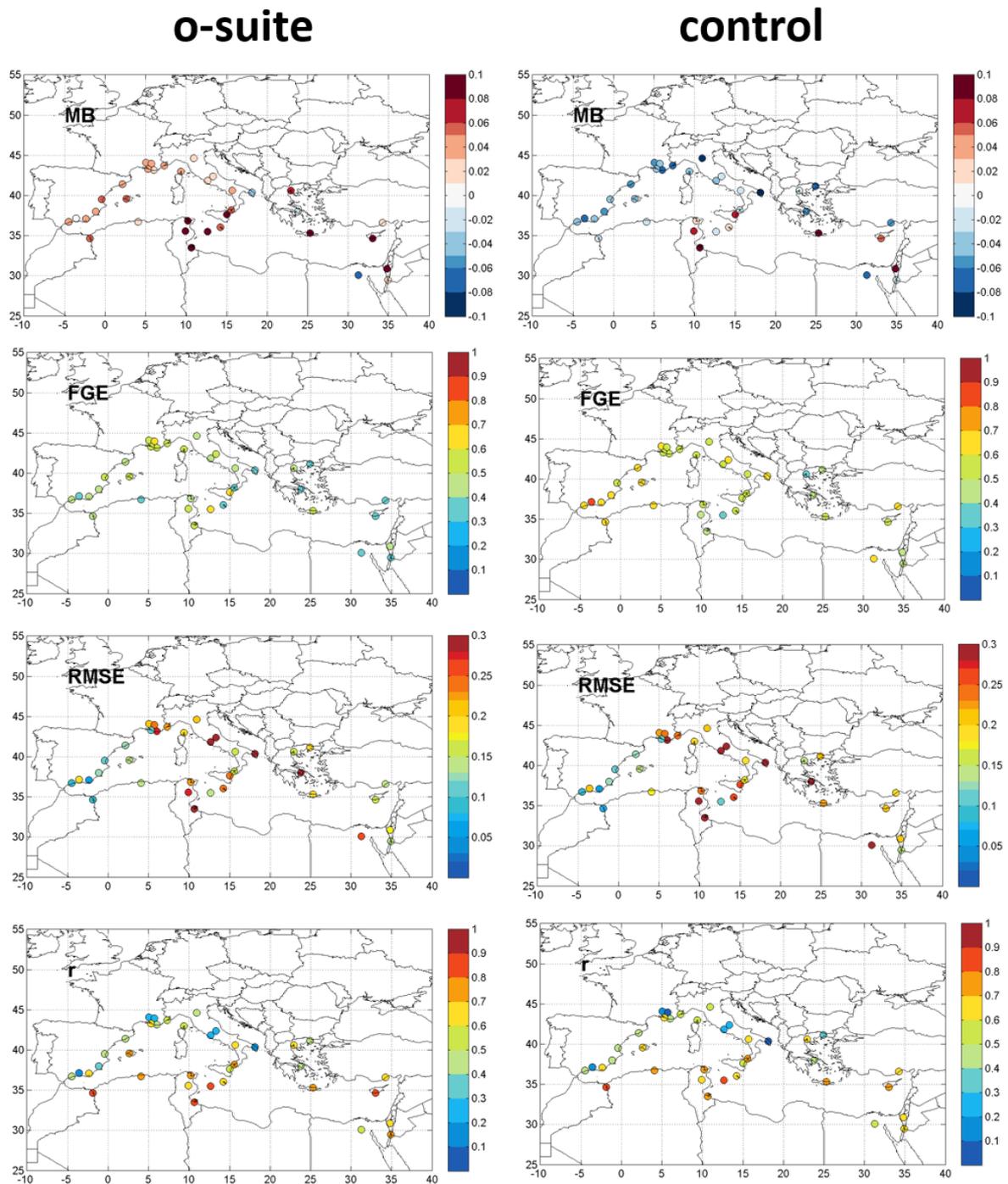


Figure 3.5.10: Skill scores (MB, FGE, RMSE and r) for 24-hour forecasts of CAMS o-suite and control for the study period from March 1st to May 31st, 2016. AOD from AERONET is the reference.

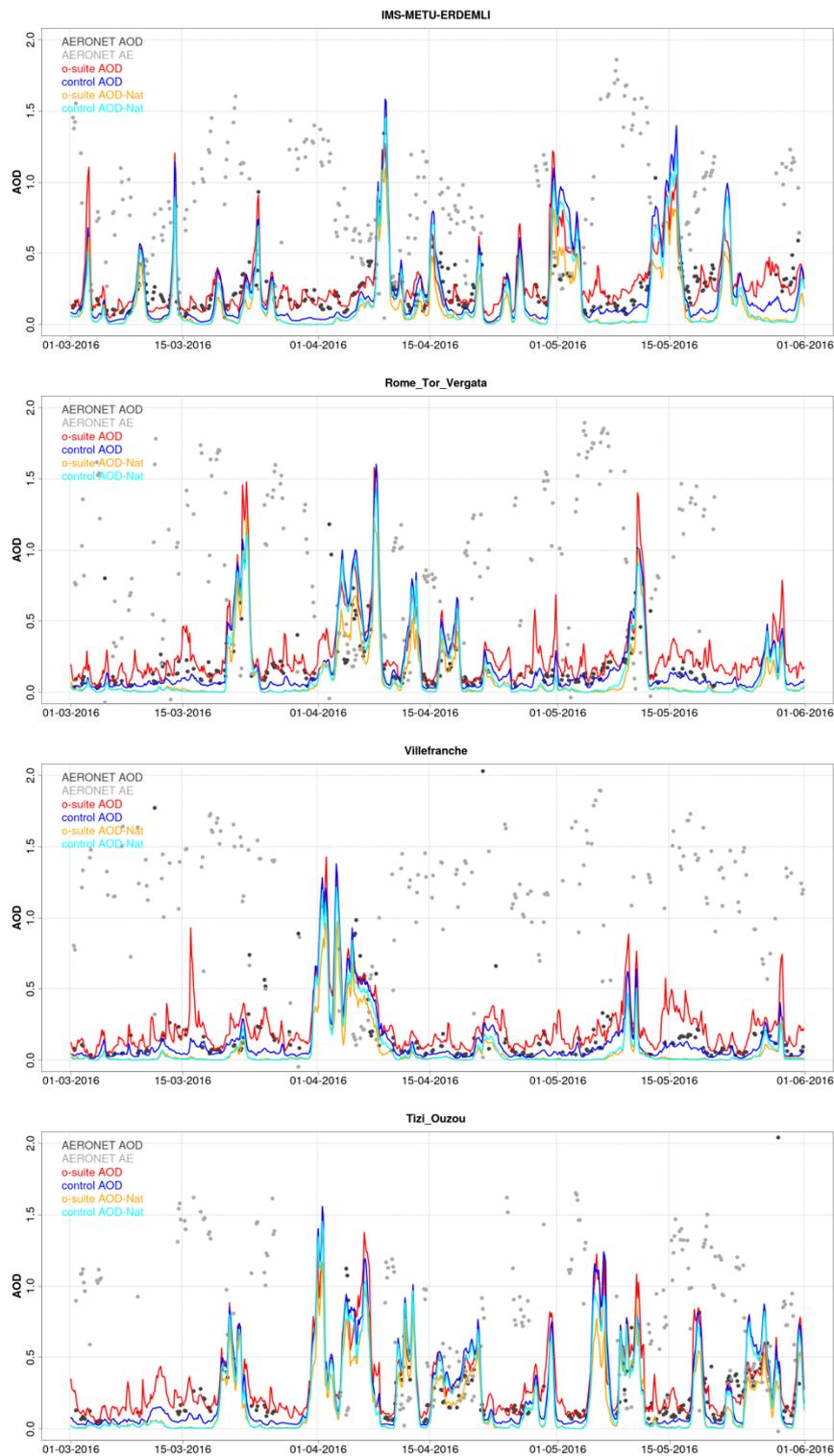


Figure 3.5.11: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), for the period March 1st to May 31st, 2016 over IMS-METU-ERDEMLI (Turkey), Rome Tor Vergata (Italy), Villefranche (France) and Tizi-Ouzou (Algeria). AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt.



Surface aerosol concentrations

For ground-level concentrations, we have used observations from the European Air quality database (AirBase; <http://acm.eionet.europa.eu/databases/airbase/>) which is the public air quality database system of the EEA. It contains air quality monitoring data and information submitted by the participating countries throughout Europe. The air quality database consists of multi-annual time series of air quality measurement data and their statistics for a representative selection of stations and for a number of pollutants. It also contains meta-information on the involved monitoring networks, their stations and their measurements. Only those stations considered as background sites in the Airbase catalogue in the Mediterranean region are considered in the present validation exercise. All measurements available in NRT (i.e non-validated observations) within ± 60 min of the models' outputs are used for the 3-hourly evaluation.

3-hourly values of PM₁₀/PM_{2.5} from Airbase, o-suite and control for the period 1 March – 31 May 2016 over 19 available selected sites are shown in Figure 3.5.12. Mean Bias (MB) and Fractional Gross Error (FGE) averaged over the study period for o-suite and control, and for the Airbase stations are shown in Figure 3.5.12.

In general, the CAMS o-suite shows better results in terms of load concentrations reducing the MB for both PM₁₀ and PM_{2.5} (see Figure 3.5.12). From March to May 2016, the CAMS experiments reproduce the daily variability of the most intense aerosol events observed by Airbase sites (see Figure 3.5.13) although both CAMS experiments tend to overestimate the observed values. On average for all the sites, PM₁₀ MB increases from 3.2 $\mu\text{g}/\text{m}^3$ for control to 5.6 $\mu\text{g}/\text{m}^3$ for o-suite, while PM_{2.5} MB increases 1.8 $\mu\text{g}/\text{m}^3$ for control to 3.4 $\mu\text{g}/\text{m}^3$ for o-suite. Furthermore, as in autumn and winter, the model shows extreme peaks particularly over maritime sites (see Hospital Joan March and Agullana on April 1-3 in Figure 3.5.13). These overestimated events are associated to sea-salt.

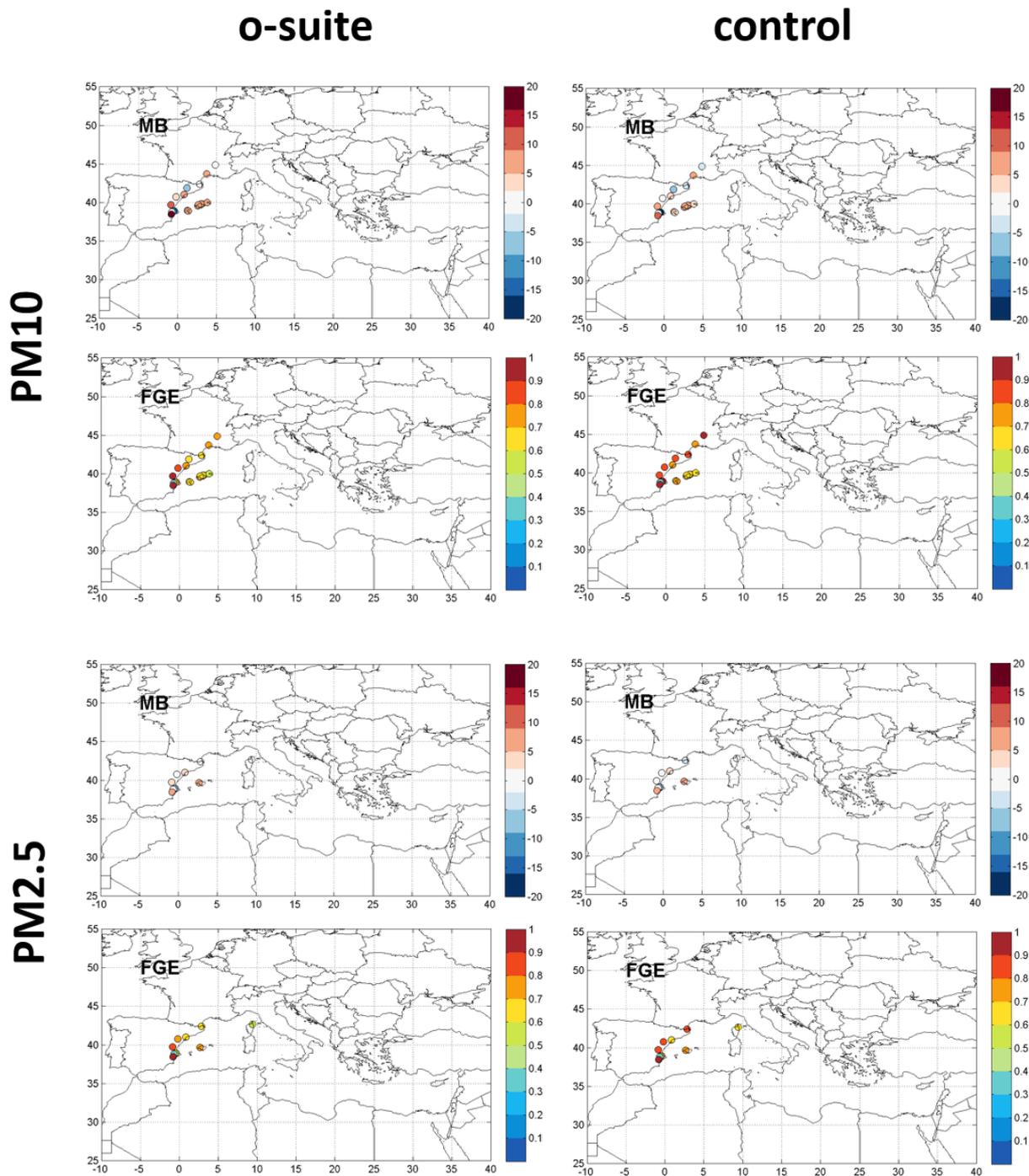


Figure 3.5.12: Skill scores (MB and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. PM10 and PM2.5 from Airbase is the reference. Only background suburban and rural stations are displayed.

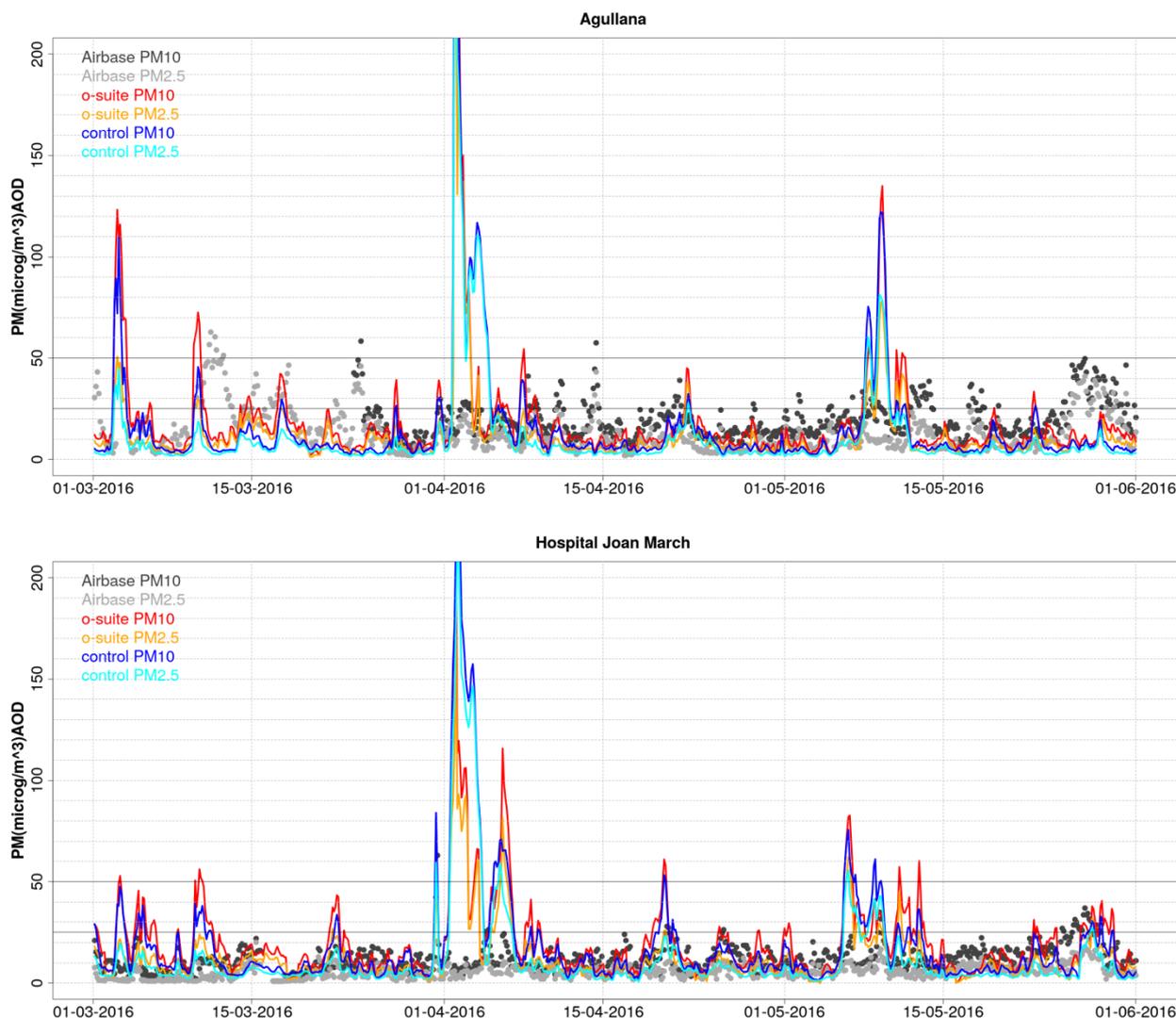


Figure 3.5.13: PM10 and PM2.5 Airbase observations (black and grey dots, respectively), PM10 and PM2.5 o-suite (red and orange lines, respectively) and PM10 and PM2.5 control (blue and cyan lines, respectively) for the period from March 1st to May 31st, 2016 over Agullana (40.73°N; 0.17°W, Spain) and 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 Eskes et al. (2016). In this section, the vertical variation of the backscatter coefficient (bsc) profiles, i.e. correlation and standard deviation, of o-suite 'g9rr' and control run 'geuh' (ctrl) vs ceilometers are evaluated and summarized in Taylor plots. The bias is not the focus, because it reflects the vertical integral of the backscatter/extinction coefficient and corresponds in its information content to AOD, which is evaluated with AERONET radiometer data. Also a skill measure for the horizontal extension of plumes/layers is not a primary goal of this evaluation. In order to avoid repetitions, issues revealed and discussed in former reports are only revised after significant model upgrades. These are:

- overestimated sea-salt due to emission at 80% humidity, underestimated sink

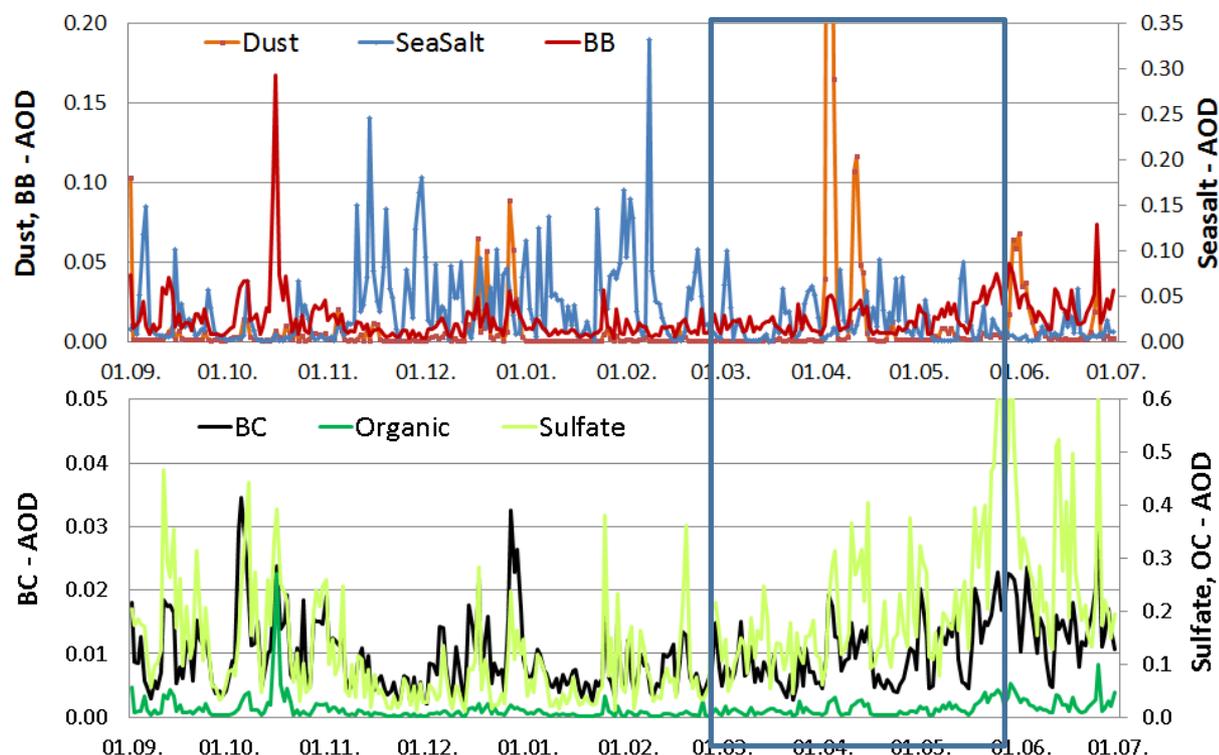


Figure 3.5.14: Maximum daily AOD around Soltau ($\pm 1^\circ$ lat/lon) for aerosols in the o-suite, from September 2015-June 2016: sea salt (blue), dust (orange), biomass burning (=OC+BC - red), BC (black), organic (green), and sulfate (light green). The reporting period is marked by a frame.

- missing sub-grid-scale structures (filaments, shallow layers,...)

The model aerosol optical depth (AOD) is used to select periods with significant aerosol plumes predicted over Germany. Figure 3.5.14 show 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) in early April and late May 2016 are confirmed by ceilometers. Sea-salt levels are seasonally low and there is no update to conclusions w.r.t. last report. Model updates are announced for autumn 2016 and spring 2017. The biomass burning aerosol concentration, the sum of black and organic carbon, was small over Germany during the reporting period. Elevated contributions of sulfate in late May/early June 2016 are associated with Canadian forest fires.

Saharan Dust

A SD plume swayed over Germany from 1-6 Apr 2016 (Figure 3.5.15, top panel) and reverse from 10-14 April 2016, thereafter moved towards NE. The ceilometer plots from Bamberg (Figure 3.5.15 middle panel) and the network overview below show the arrival of the plume on 2 April, moving from SE to NW (note that the x-axes are time!). The internal structure and the formation of clouds at the top of the plume are typical for SDE.

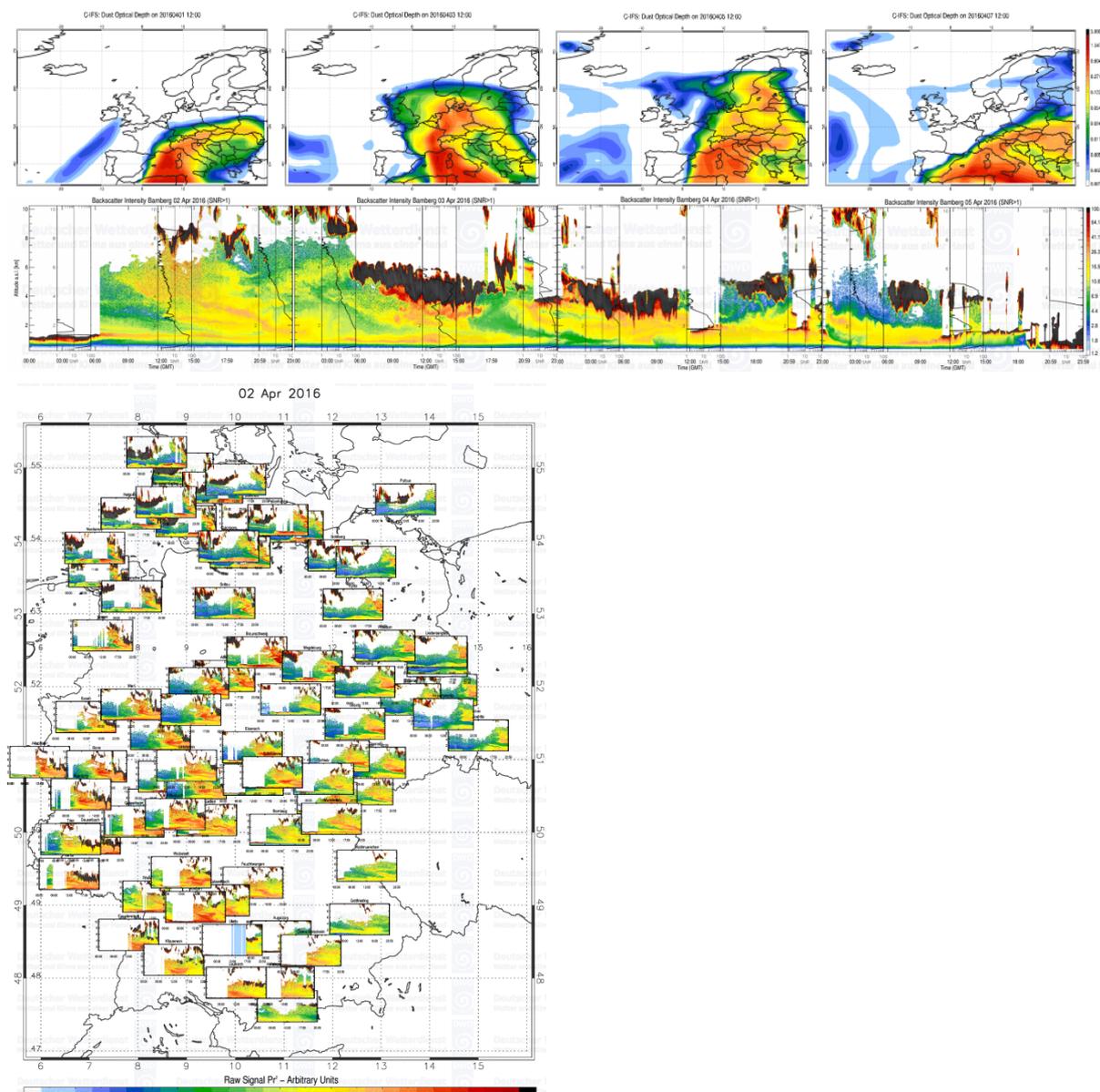


Figure 3.5.15: Top panel: Dust AOD from the o-suite on 1, 3, 5, 7 April 2016. Mid panel: Time-height sections of backscatter signal (Pr^2) over Bamberg on 2-5 April, showing the passage of the SD plume. Bottom panel: German network overview on 2 April 2016.

The model aerosol is made up mostly of dust and traces of sulfate during this period. The dust plume is generally captured, but the model tends to produce a bell-shaped vertical profile (Figure 3.5.16) while this is more flat in reality. The dust load (mass or bsc) is too large in the model, an issue generally related to the time step and resolution. The internal structure is largely missing according to the model resolution. The assimilation has a positive impact by reducing the dust load in the o-suite, but not enough. As before, individual layers are sometimes captured remarkably well (e.g. in 1 km on 2 and 4 April), probably a matter of scales, but there is no obvious improvement to the profile shape by the assimilation. This is not expected, too, since AOD but no profile information is assimilated.

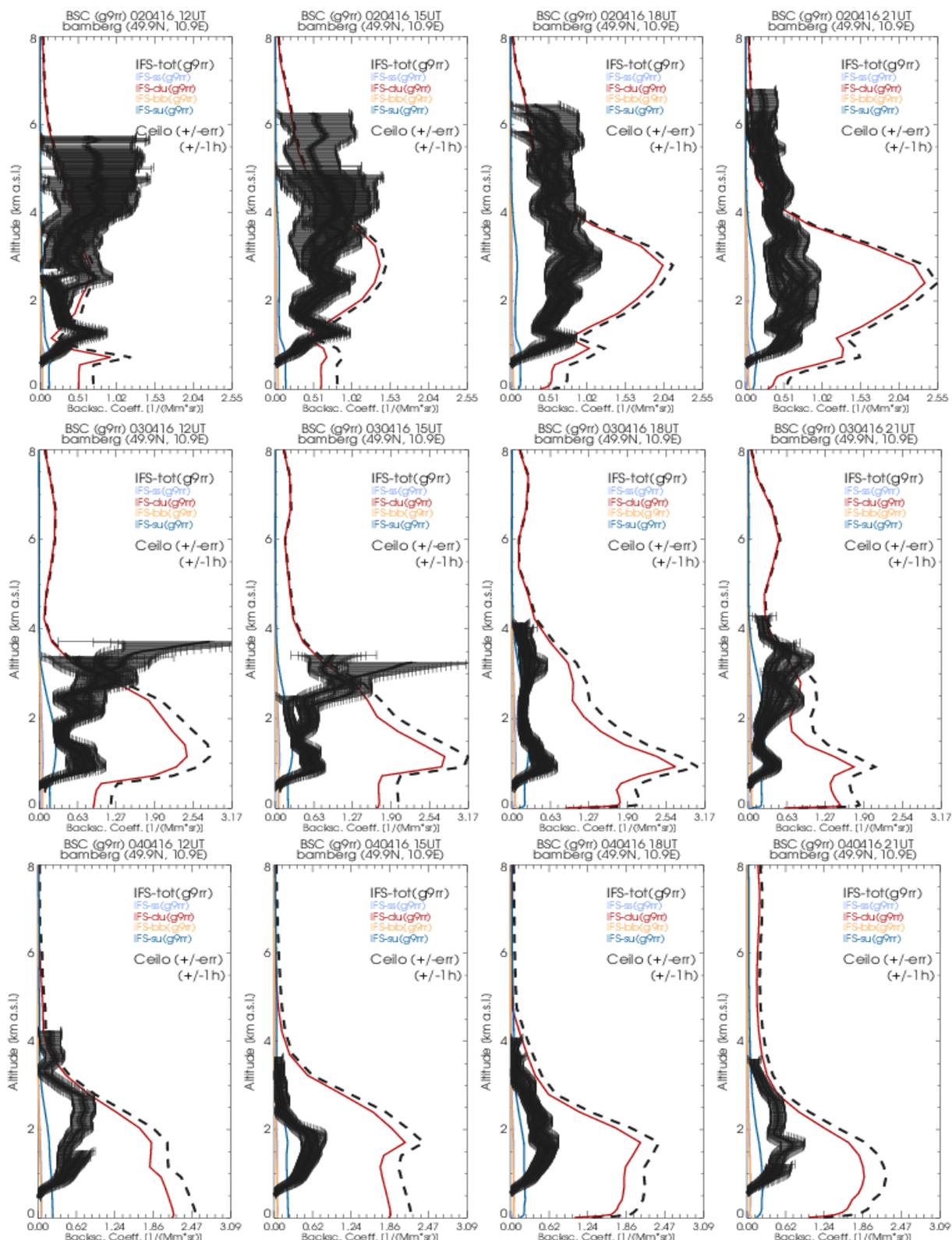


Figure 3.5.16: Model profiles of backscatter coefficients at Bamberg on 2,3,4 Apr 2016 (o-suite), splitted for contributions from sea-salt (light blue), dust (red), biomass burning (orange) and sulfate (dark blue), as well as the total aerosol (dashed black). Solid black lines: Ceilometer data with estimated error bars, plotted for -1h, +/-0h, +1h around the time of the 3-hly model profile.

With the excessive dust load, also the variability (in absolute terms) is too high. This reflects in standard deviations much larger than the reference (green) in the Taylor plots in Figure 3.5.17. The correlations are relatively low, because the internal structure is not captured by the smooth bell-shape model profile and further there are vertical and horizontal displacements.

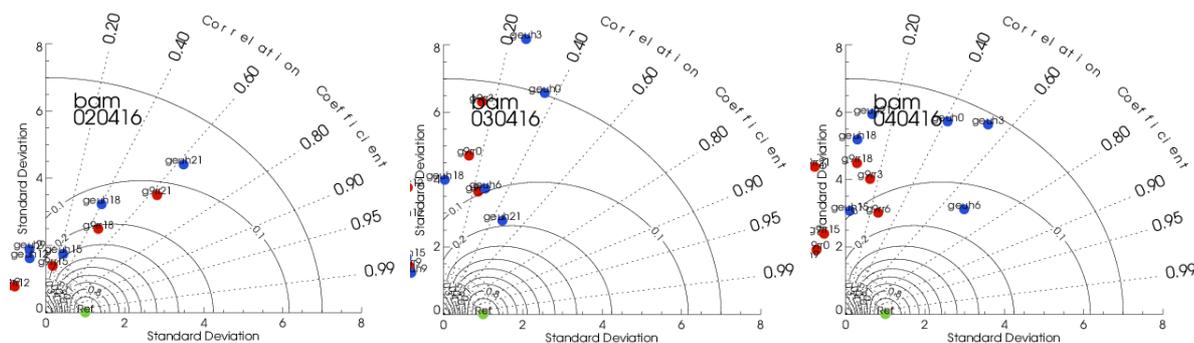


Figure 3.5.17: Taylor polar plots with standard deviation vs correlation coefficient of the 3-hourly profiles for Bamberg on 2, 3, and 4 April 2016. O-suite and control run are distinguished by red respective blue colors.

Sulfate

Sulfate concentrations increase throughout the spring and early summer (Figure 3.5.14) due to increasing $\text{SO}_2 \rightarrow \text{SO}_4$ oxidation, which is calculated by the model with a fixed time step and simple seasonal/meridional dependence (solar zenith angle). This leads to shortcomings in the representation of ammonia (NH_3) and nitrate (NO_3 – not present in the current model). The sulfate AOD is known to be 3-4 times too high due to unbalanced initial conditions, independent of the DMS source.

From 29 May 2016 onward a SD plume crossed Europe from the south and resided over northern Germany for some days. At the same time, biomass burning plumes were repeatedly transported from Canada across the Atlantic to Europe. During this event the plume shape and the total aerosol load have been captured better than in the SDE in April, evident as higher correlations and closer-to-reference variabilities (standard deviations) in Figures Figure 3.5.18 and Figure 3.5.19.

The total aerosol contains larger contribution of sulfate than of dust (Figures Figure 3.5.14, Figure 3.5.19), which cannot be verified by the ceilometer measurements. Owing to the co-variance with BC and OC (biomass burning), the sulfate can be assigned to boreal fires in Canada, where the burning season peaks in May/June. The peak SO_4 AOD of 0.74, however, seems rather high. At the Global Atmosphere Watch (surface) station Hohenpeissenberg, for example, model SO_4 on 27-29 May has a 3-daily average concentration of $4.2 \mu\text{g}/\text{m}^3$ ($0.6\text{-}12 \mu\text{g}/\text{m}^3$ - L60) (Figure 3.5.20) compared to an observed value of $0.65 \mu\text{g}/\text{m}^3$. This may be an indication that the good agreement of the profiles in Figure 3.5.19 could result from overestimated model sulfate, which would mean that dust concentrations are too low.

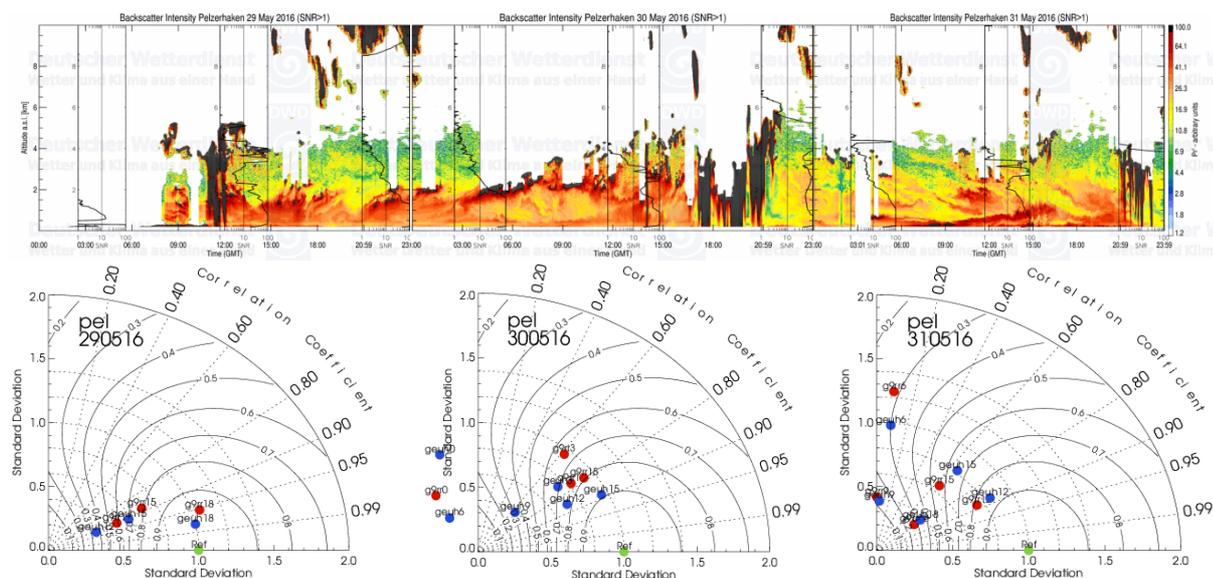


Figure 3.5.18: Top panel: Time-height sections of backscatter signal (Pr^2) over Pelzerhaken on 20-31 April 2016. Bottom panel: Taylor polar plots with standard deviation vs correlation coefficient of the 3-hourly profiles for Pelzerhaken on 29, 30, and 31 May 2016.

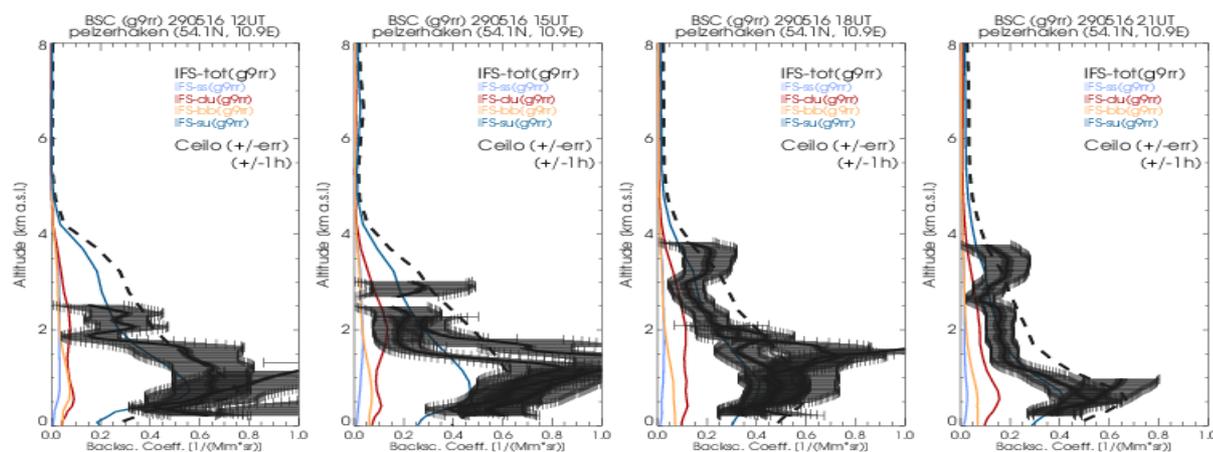


Figure 3.5.19: Model and ceilometer profiles of backscatter coefficients at Pelzerhaken on 29 Apr 2016 (o-suite) as in Fig. 3.5.16.

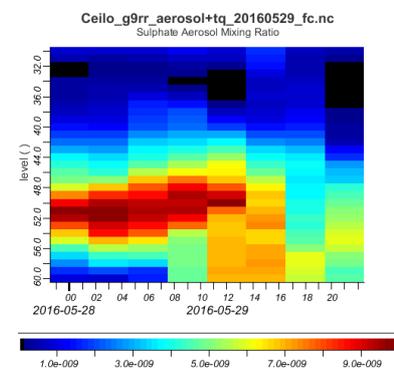


Figure 3.5.20: Model SO_4 mixing ratios (kg/kg) at $48^\circ N$, $11^\circ E$ (Hohenpeissenberg station) at surface levels (L60)

Summary



The reproduction of detailed aerosol profiles is a serious challenge for a global model. Keeping this in mind, the IFS aerosol model is already quite successive in simulating details of the aerosol distribution and partitioning over Germany (the domain of this evaluation), whereby spatial and temporal shifts occur as well as some issues with apportioning of correct aerosol concentrations. For daily averages in the months March to May 2016, this is shown in Figure 3.5.21.

- Dust events are forecasted with high reliability
- Profiles during dust events often well associated with observations, but shifts occur
- Small scale structures in dust plumes not resolved due to model resolution
- Sulfate concentrations are much too high in the model
- Fractionation between the different aerosol types may lead to error cancellation

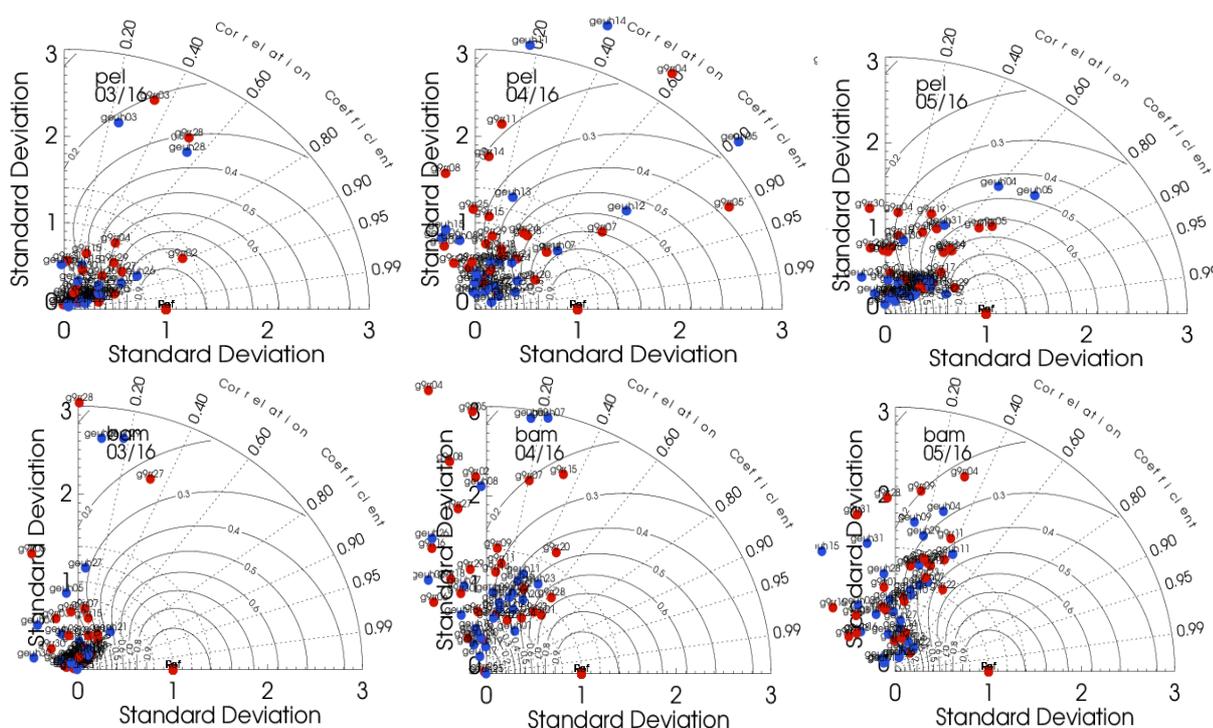


Figure 3.5.21: Taylor polar plots with daily average standard deviation vs correlation coefficient for two German sites Pelzerhaken (north-Germany - top panel) and Bamberg (mid-Germany - bottom panel) for the months March, April and May 2016. O-suite and control run are distinguished by red respective blue colors.



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% ($\sim 10\%$ 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 May 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 are 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 -5 to +10%, for all regions and months (some exceptions with MNMBs of up to 15% for single months in the high latitude regions), see Figure 3.6.1 and Figure 3.6.2.

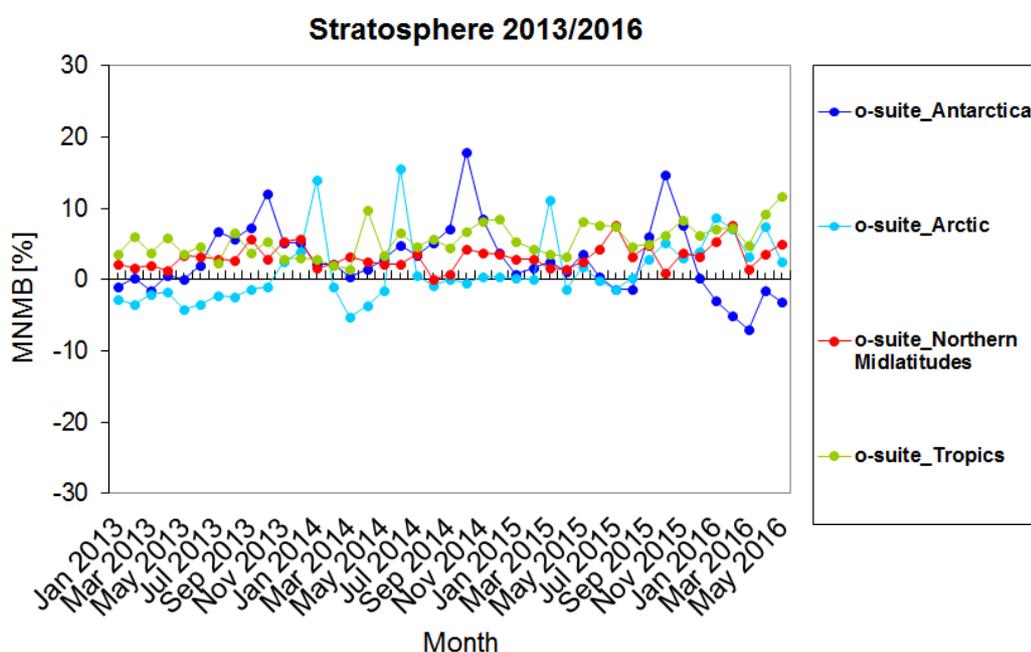


Figure 3.6.1: MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated sonde data in the Arctic (light blue), Antarctic (dark blue) northern midlatitudes (red) and tropics (green).

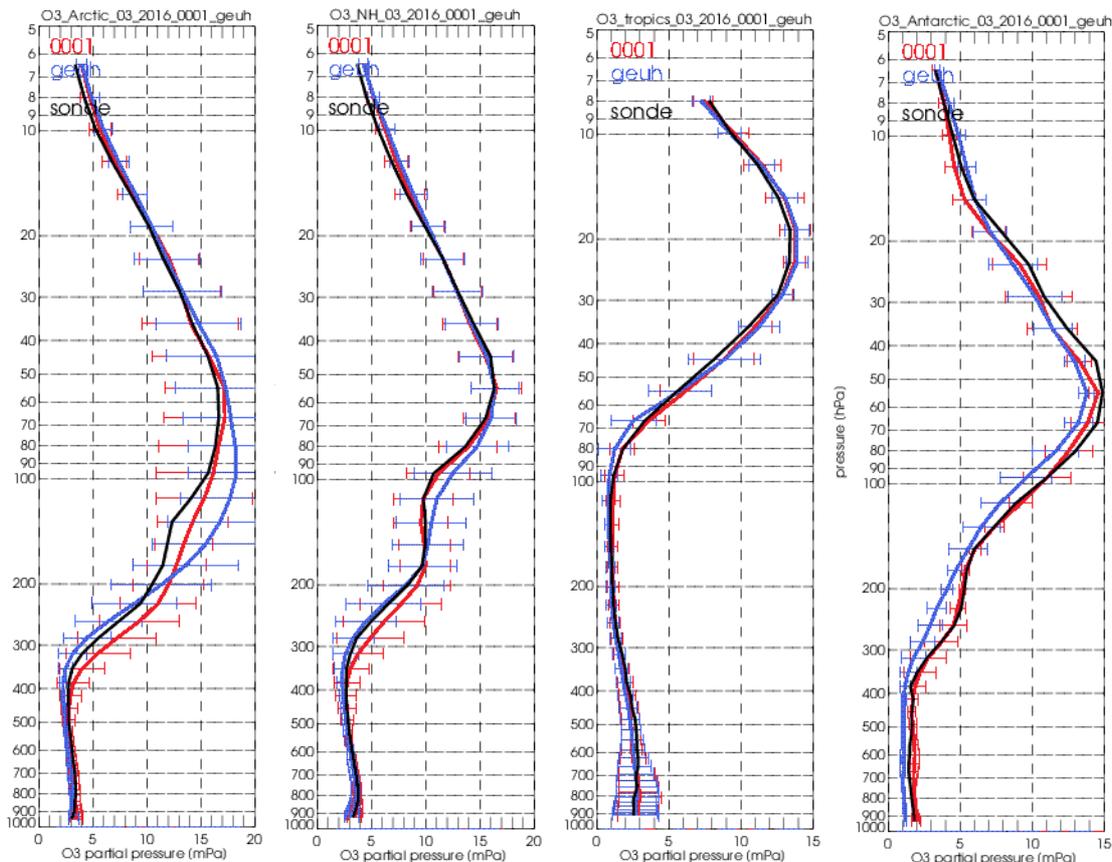


Figure 3.6.2: Comparison between mean O₃ profiles (units: mPa) of o-suite (red), and control (blue) in comparison with observed O₃ sonde profiles (black) for March 2016 for the various latitude bands: Arctic, NH-mid latitudes, Tropics and Antarctic.

O₃ partial pressures in the stratosphere are mostly slightly overestimated in all latitude bands, except for the Arctic and Antarctic summer season. MNMBs in Antarctica during the ozone hole season, from August to November, remain below $\pm 15\%$ for the o-suite.

For the Northern midlatitudes, the control run partly shows lower MNMBs than the o-suite, which results from a combination of over- and underestimation in different pressure levels, cancelling each other out in total.

Figure 3.6.2 compares the averaged profiles in each region during March 2016. The vertical distribution of stratospheric ozone is quite well represented for the Northern midlatitudes and the Tropics by the o-suite. For Antarctica, ozone partial pressure is slightly underestimated and for the Arctic, ozone partial pressure of the lower stratosphere is overestimated.

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 in the Annex 2 and 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)



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 %).

		JJA			SON			DJF			MAM		
		MB	stddev	nobs	MB	stddev	nobs	MB	stddev	nobs	MB	stddev	nobs
o-suite	Ny.Ale	-2.05	5.33	224	12.84	7.39	273	14.50	6.51	213	6.35	5.42	229
	Bern	-4.93	2.50	513	-0.54	2.32	687	0.74	3.43	527	0.97	2.82	609

From Table 3.6.1, the upper stratospheric partial column bias at Bern during Sept. 2015 –June 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%, which decreased to 6% during MAM. 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 Figure 3.6.5. At BERN the difference between osuite and MWR at 25-35km is negligible since Sept 2015 (compared to the MWR profile uncertainty).

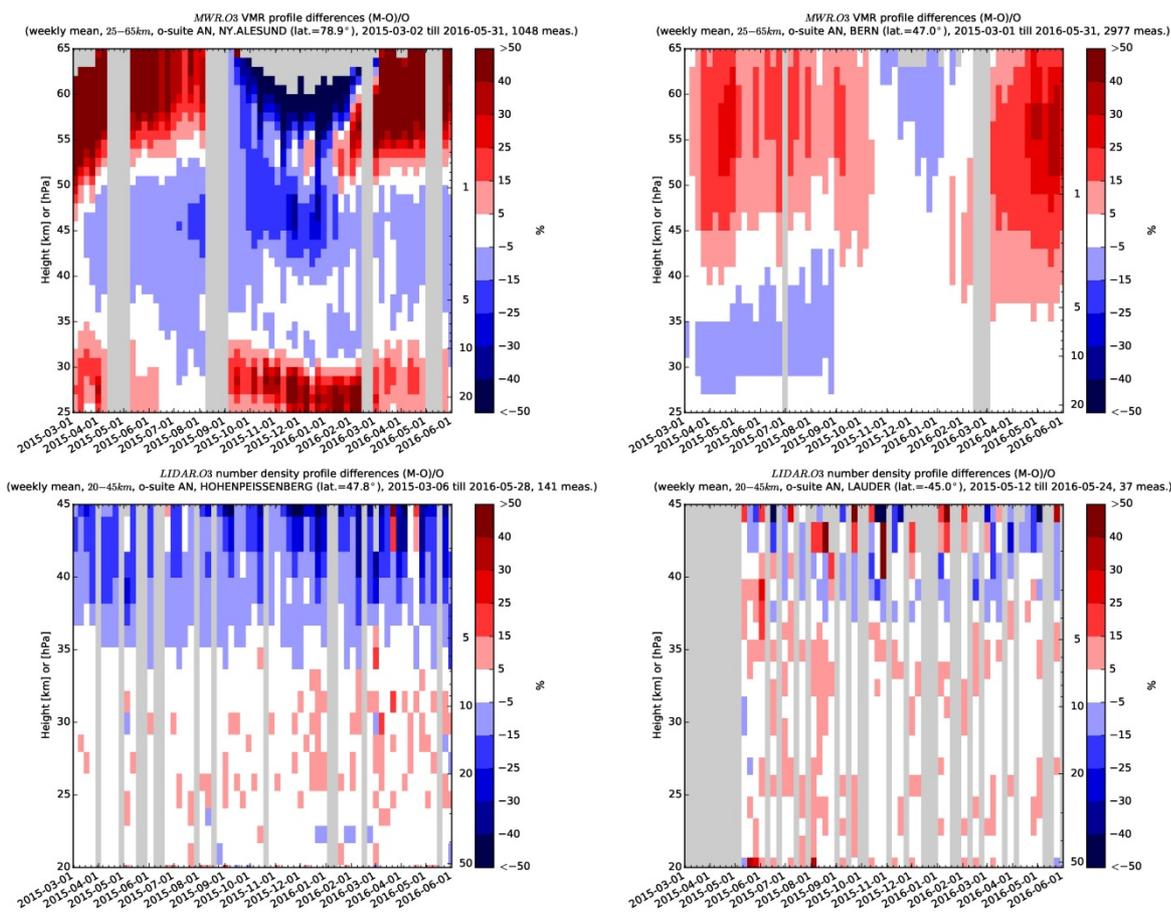


Figure 3.6.3: Comparison of the weekly mean profile bias between the O₃ mixing ratios of o-suite and the NDACC station at Ny Alesund, Bern, Hohenpeissenberg and Lauder. For the LIDAR stations, the measurement uncertainty above 35km is comparable to the observed profile bias.

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 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)

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.

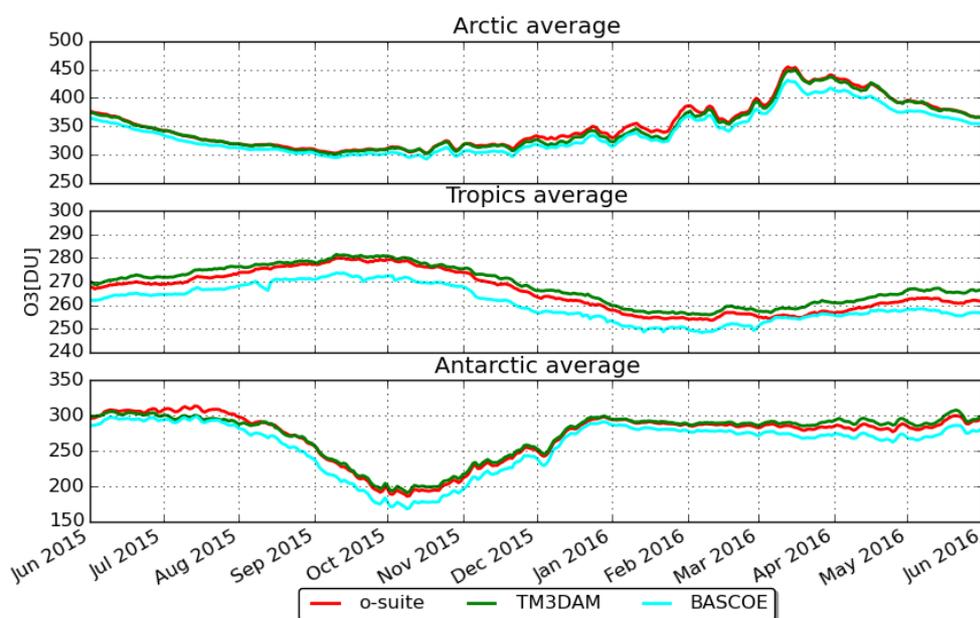


Figure 3.6.4: Zonally averaged ozone total column (Dobson Units) in the Arctic (60°N-90°N), Tropics (30°S-30°N) and Antarctic (90°S-60°S) during the period 2015/06/01-2016/06/01.



System intercomparison for total columns

Figure 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) 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. $\sim 3\%$. This disagreement is related to the ozone depletion event discussed in Section 5.1.
- 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 2-3 DU, i.e. $\sim 1\%$. For the last trimester, the underestimation seems to increase slightly.
- 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). Figure 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 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 Figure 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 10%, compared to approximately 6% for the analyses).

The bottom row of Figure 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 10% in the Tropics and varies in the polar regions from 5% (summer and fall) to 10% (winter and spring).



O3 relative bias against OMPS-LP: 30-70hPa mean from 20150601 to 20160601

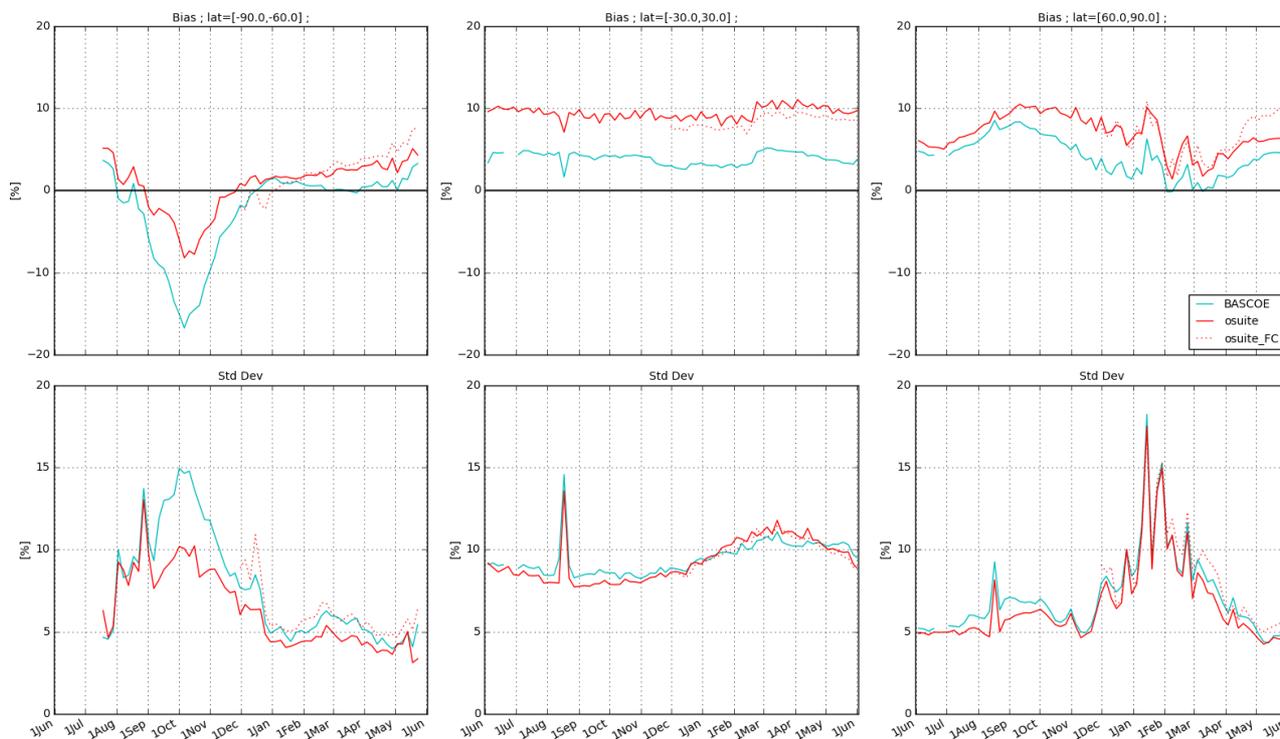


Figure 3.6.5: Time series comparing ozone from o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted), and BASCOE (cyan) with OMPS-LP satellite observations for the period 2015-06-01 to 2016-06-01 in the middle stratosphere (30-70hPa averages): top row, normalized mean bias (model-obs)/obs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

Figure 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. March-May 2016. In the northern hemisphere, a vertical discontinuity of the relative 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.

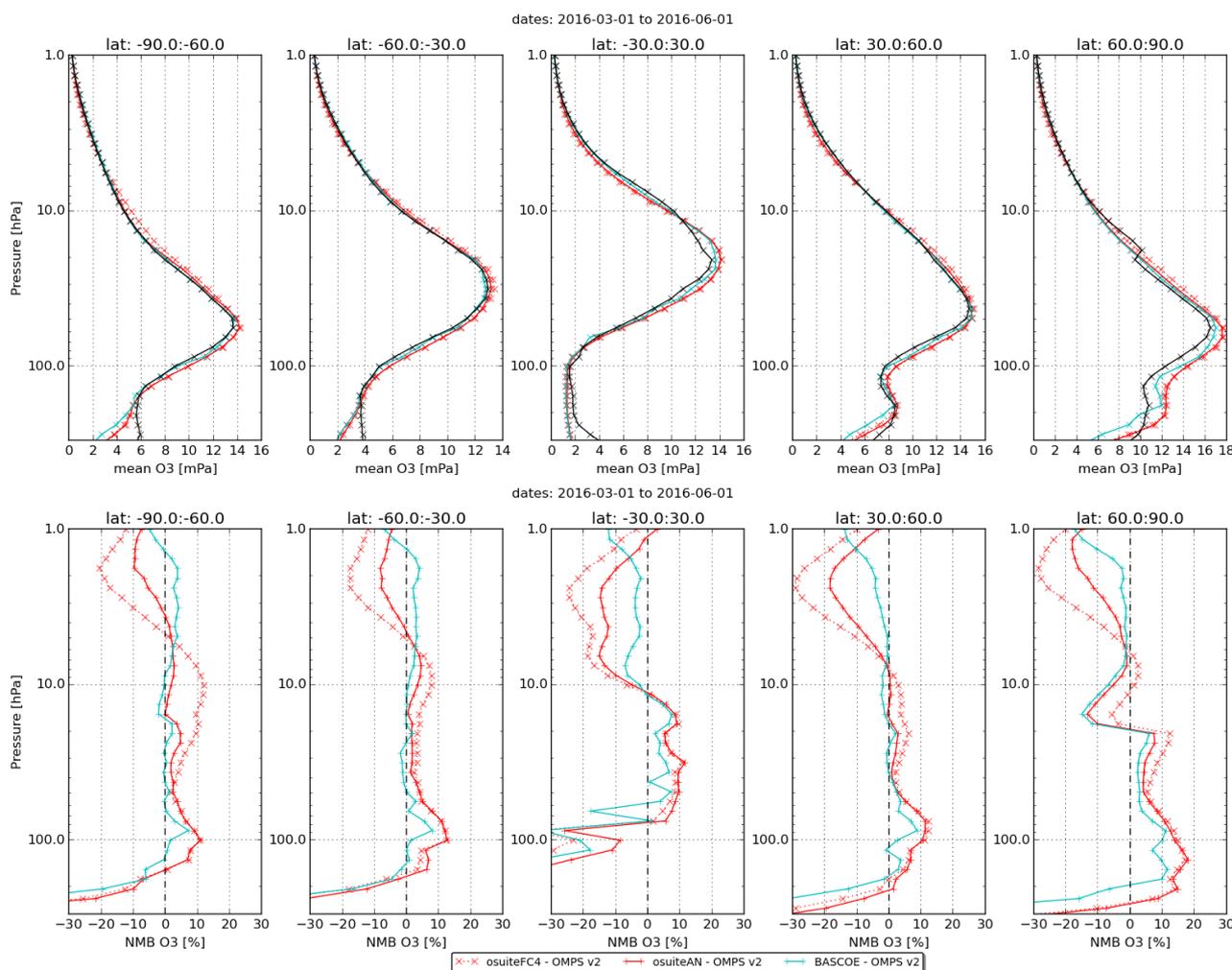


Figure 3.6.6: Mean value expressed in partial pressure (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with OMPS-LP v2 observations for the period March-April-May 2016.

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 (see methodology described in Annex 2). 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 NO₂ 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 NO₂ 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

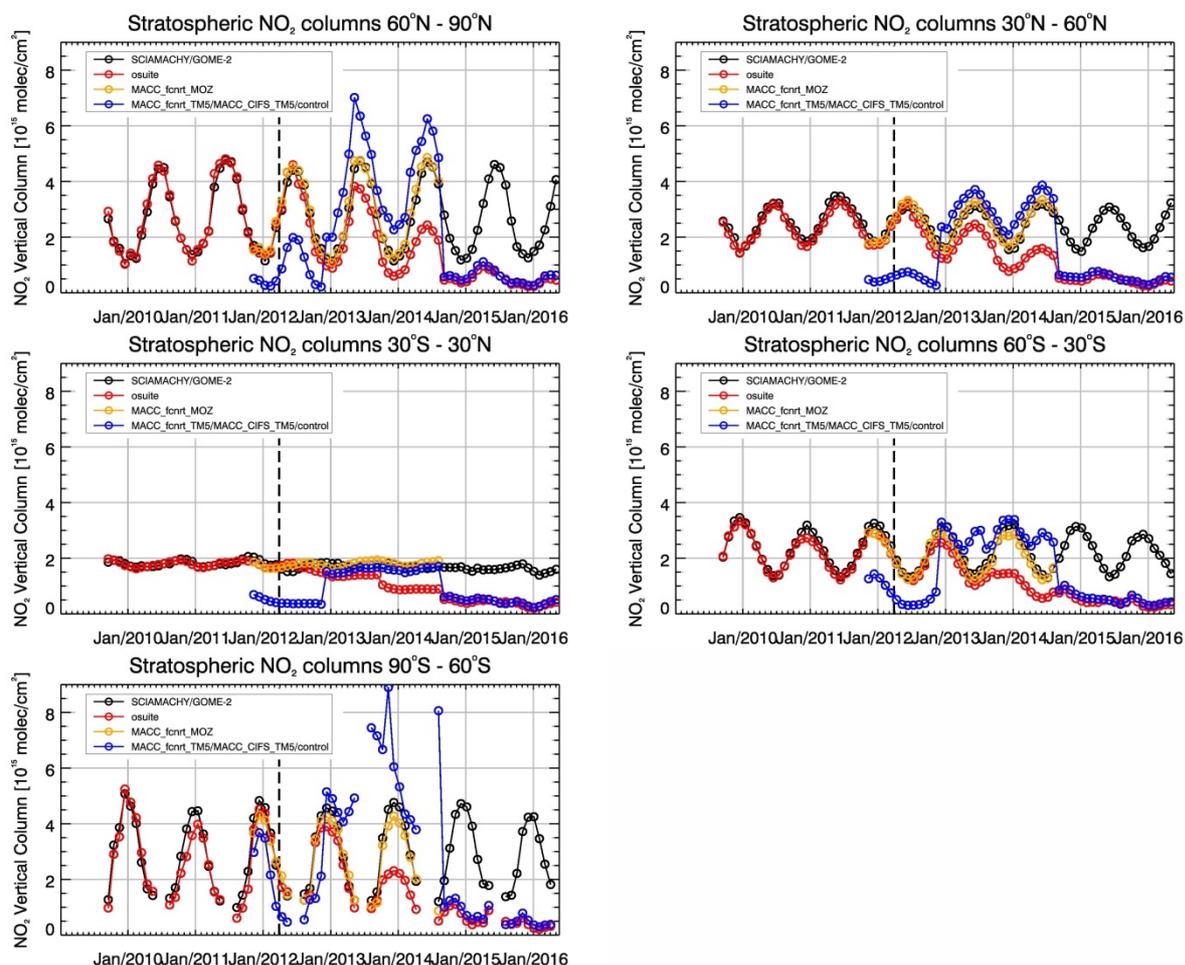


Figure 3.7.1: Time series of average stratospheric NO_2 columns [10^{15} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012) compared to model results for different latitude bands. See text for details. The blue line shows MACC_fcrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 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.

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

This section describes the NRT validation of the pre-operational, high resolution forecast of CO₂ and CH₄ from July 2015 to June 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 and TCCON dataset.

4.1 CH₄ and CO₂ validation against ICOS observations

4.1.1 Assessment of the diurnal cycle

We have compared the diurnal cycles at the 15 surface stations, using hourly means observations (Figure 4.1.1). To do so we have subtracted the daily averages to each hourly value, both in the observations and simulations over the one year period. For most of the stations the model underestimates the CO₂ diurnal amplitude, and overestimates the CH₄ amplitude (Figure 4.1.2). For CO₂ we also observe a better agreement for the tower sites located in plains, compared to mountain and coastal stations. On average the model underestimates the amplitude of the diurnal cycle by 1.1 ppm at the tall tower (≥ 100 m) sites in Europe.

The results of the comparison are much more scattered for CH₄. The higher dispersion of the model/data comparisons is probably due to the higher dispersion of CH₄ sources, compared to CO₂. For example, the model overestimates the amplitude of the diurnal cycle at the Trainou tower by 43 to 12 ppb when using sampling heights from 5 to 180m above the ground. The performance of the model is better at two other tall towers, also located in rural areas in France: 5.5 to 6.5 ppb overestimation at OHP (10 to 120m agl), and -1 to -3 ppb underestimation at OPE (10 to 100m agl). The specific issue at Trainou tower, compared to OHP and OPE, is likely due to the surface emissions rather than the parametrization of the atmospheric transport. Figure 4.1.3 shows the map of annual mean CH₄ fluxes around each station. Clearly the TRN tower is surrounded by a much larger range of surface fluxes, especially with the emission of Paris area 90 km to the North (3 grid boxes away) and the city of Orleans 20 km South west of the site (2 grid boxes away). As a consequence the spatial variability of atmospheric concentrations around TRN tower is much larger than the two other sites. Such situation represents a challenge for the model, even at high resolution, since any error in the regional distribution of the sources will have a huge impact on the simulations. The assessment of the transport model performance should focus as much as possible on the stations located in area with homogeneous flux distribution.

On the Figure 4.1.2 the Finokalia station (FKL) appears as an outlier for the CO₂ diurnal cycle, with simulated amplitude of 23 ppm when the observed one is only 0.5 ppm. The location of the model grid box used for FKL is shown in the Figure 4.1.4. The default rule used so far was to extract the closest continental grid box, but in the case of FKL this corresponds to an inland location with diurnal amplitudes of CO₂ and CH₄ largely overestimated by the high resolution model. When switching to the closest oceanic box we reach a much better agreement with observations (Figure 4.1.4). This location will be used by default in the future evaluations, and other sites are under review to optimize the choice of the grid box.

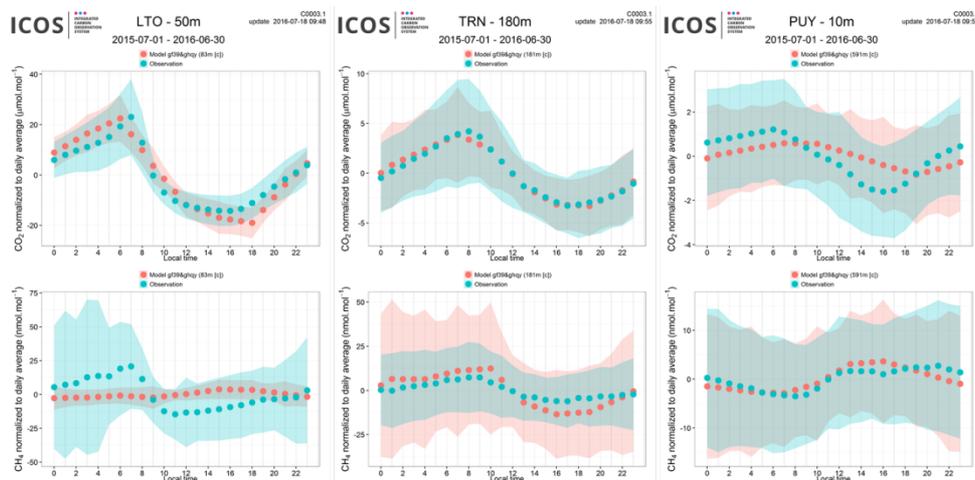


Figure 4.1.1: Mean diurnal cycle averaged over one year (July - June 2016) for CO₂ (above) and CH₄ (below) at three stations from left to right: Lamto (LTO), Trainou tower at the highest level (TRN) and Puy de dome (PUY). Observations are shown in blue and model results in red.

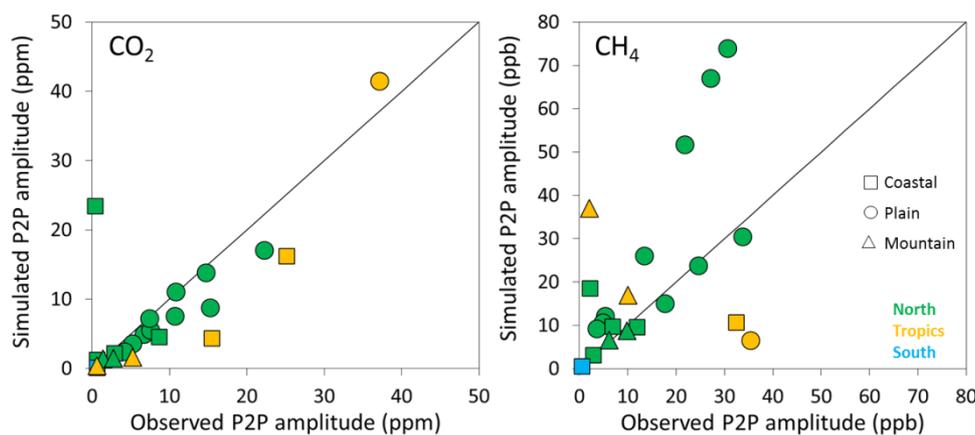


Figure 4.1.2: Comparison of the diurnal peak to peak amplitudes observed at surface stations compared to the model results.

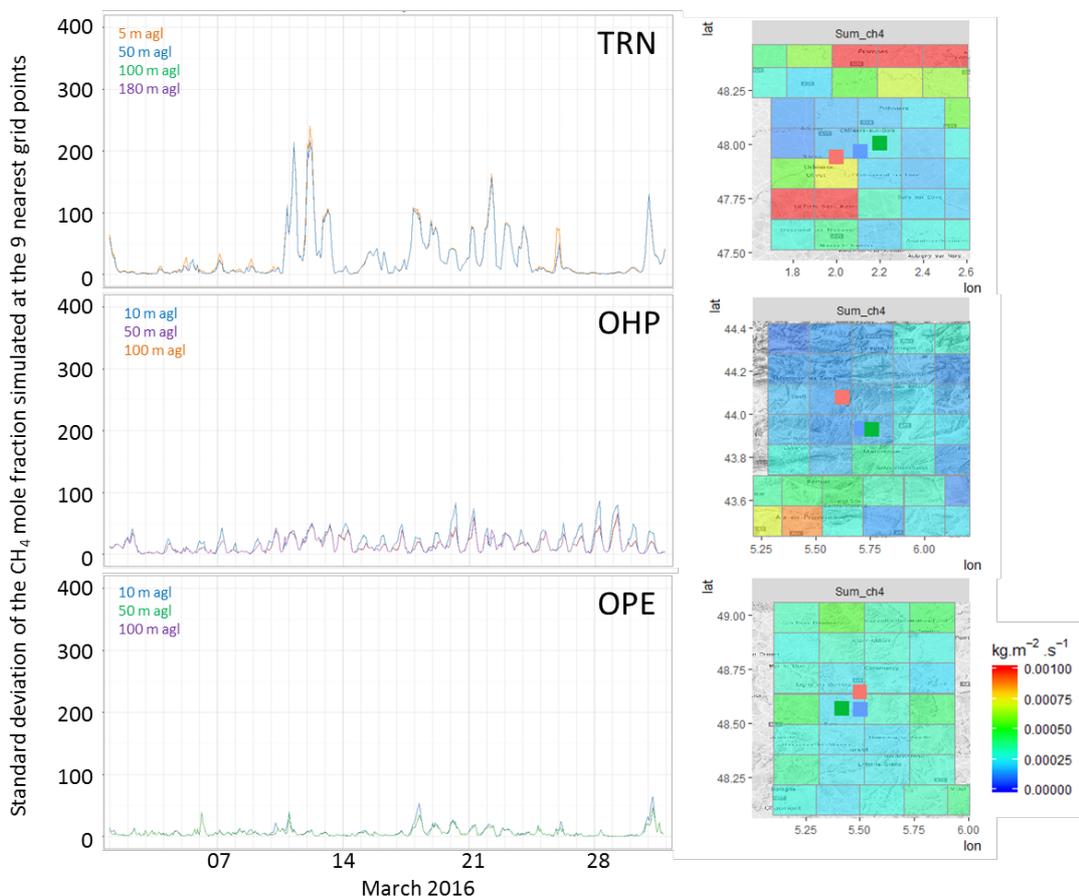


Figure 4.1.3: Right: Location of the stations (blue square), compared to the extracted grid points in high (green square) and low resolution (red square), on the CH₄ emission map (in kg CH₄.m⁻².s⁻¹) at high resolution (N640) Left: time series of the standard deviation of the CH₄ mole fractions simulated in March 2016 in the nine nearest grid point (N640) around each station (TRN, OHP and OPE towers).

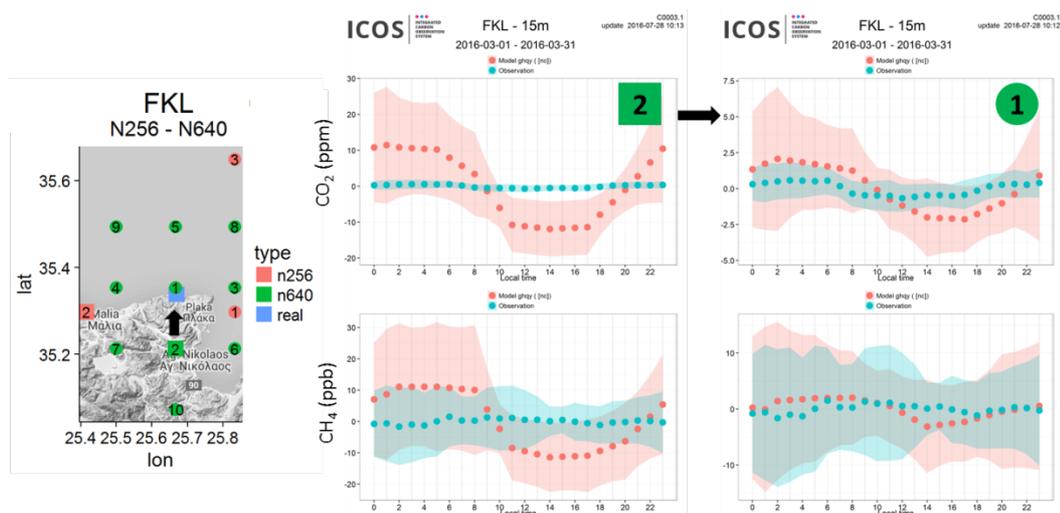


Figure 4.1.4: Left: Location of the FKL station (blue square), compared to the extracted grid points in high (green square) and low resolution (red square). Middle and Right: Mean diurnal cycles of CO₂ (above) and CH₄ (below) observed (blue) and simulated (red) in March 2016. The middle plot corresponds to the extracted grid box by default (closest continental box), whereas the right plot show the proposed grid box (closest oceanic box).



4.1.2 Evaluation of the synoptic to seasonal scale

The second step of the analysis aims to validate the representation by the model of variabilities ranging from the synoptic to the seasonal scale. In this 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. In this section it is important to note the change of experiment on March 1st 2016, leading to an abrupt change of CO₂/CH₄ concentrations at some sites. Figures 4.1.5 and 4.1.6 show the CO₂ and CH₄ comparisons at three selected stations in South hemisphere (Amsterdam I.), North hemisphere (Mace Head) and in the tropics (Lamto). For example we see a reduction of the bias at Amsterdam I. starting in March 2016 with the new experiment (ghqy). This abrupt change in the simulation for some stations (especially Amsterdam I.) biases the yearly mean metrics calculated from July 2015 to June 2016 (Figures 4.1.8). The mean CH₄ bias at Amsterdam I. is -28 ppb but it was -40 ppb in the first experiment (gf39) and converging to zero in the second experiment (ghqy).

Mace Head is the site where we have the best correlation coefficients (0.93) for both CO₂ and CH₄. At this station like for many other ones in Europe, the CO₂ correlation decreases in spring 2016, probably due to the representation of the biospheric carbon uptake. Overall for European stations, excluding FKL for the same reasons explained in the validation of the diurnal cycles, the yearly mean biases are 1.4 ± 1.8 ppm (RMSE: 5.4 ppm) for CO₂, and -4.1 ± 10 ppb (RMSE: 27 ppb) for CH₄. The CO₂ biases show significant seasonal pattern at most European sites, with maximum overestimation in winter time (from November 2015 to March 2016). It should be noted that the RMSE calculated for CH₄ at TRN are significantly higher compared to other European sites (Figure 4.1.8). This is due to an overestimation of synoptic variability by the model, probably explained by the vicinity of urban areas (Figure 4.1.3). This specific point will be further explored in the next report.

The performances of the model are more difficult to evaluate in the tropics where the available stations (RUN, STD, LTO, GUY, CHC) show more variability, but overall the agreement is generally lower than for European sites. 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, and underestimates the CH₄ ones.

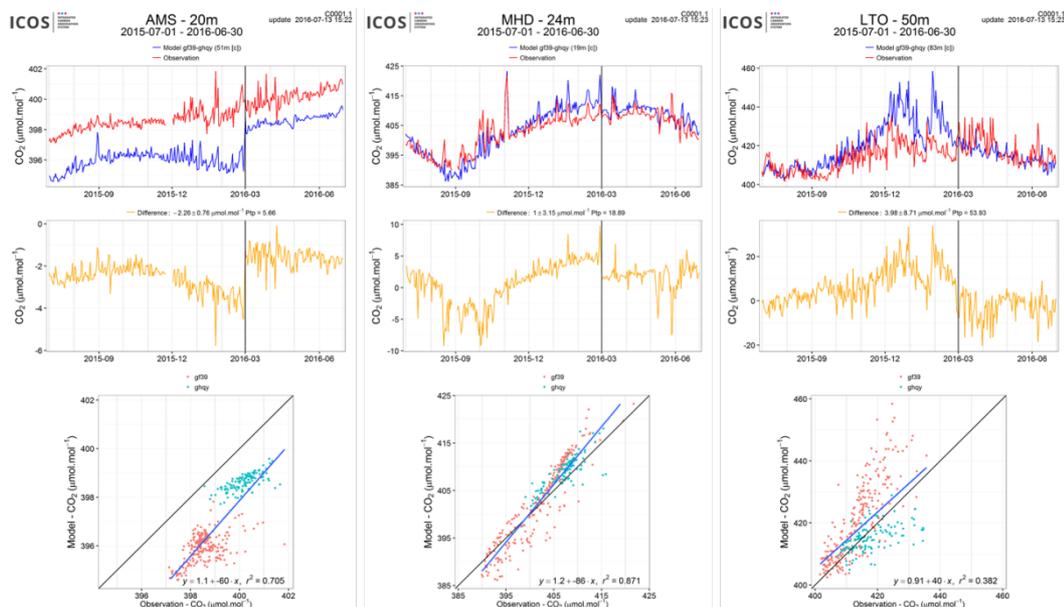


Figure 4.1.5: Above: Comparison of CO₂ daily means observed (red) and simulated (blue) at three stations (Amsterdam I., Mace Head and Lamto). Middle: differences of the observations minus the simulations. Below: Linear fit between observations and simulations. Blue points correspond to the ghqy experiment, initialized on March 1st, 2016.

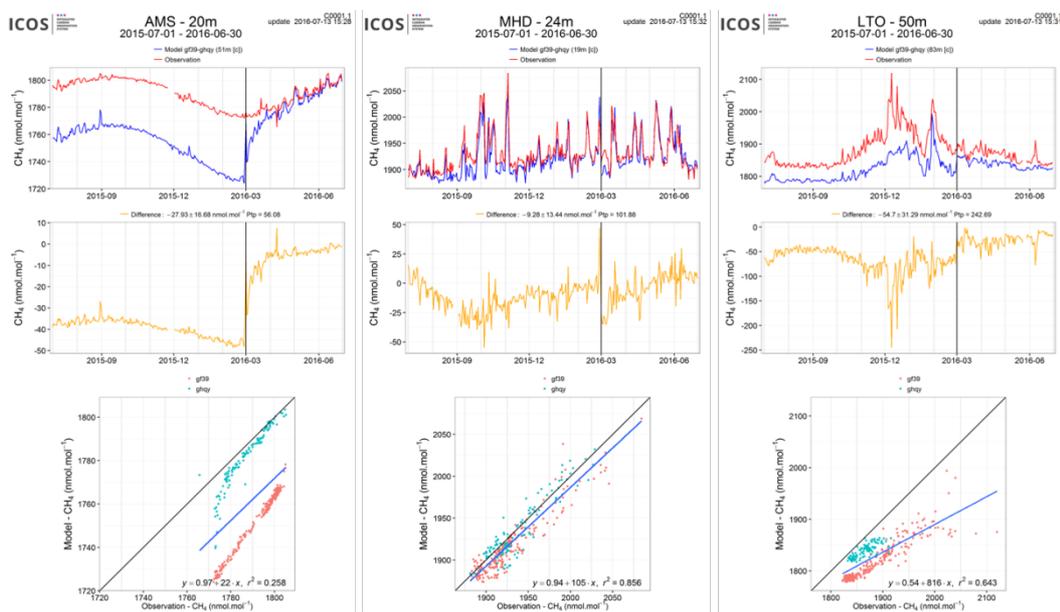


Figure 4.1.6: Same as figure 4.1.5 for CH₄

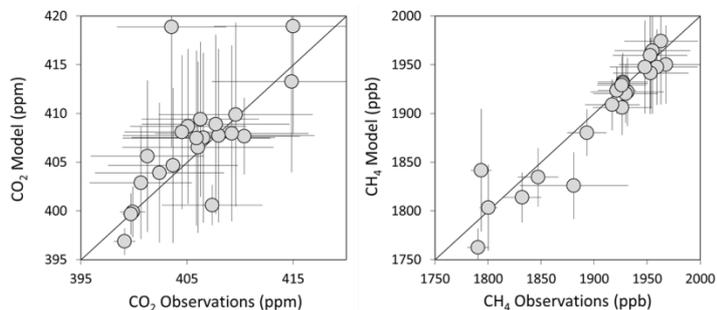


Figure 4.1.7: Comparison of the a observed and simulated annual means of CO₂ (left) and CH₄ (right) concentrations at the 15 sites, calculated from the daily averages over the period July 2015- June 2016.

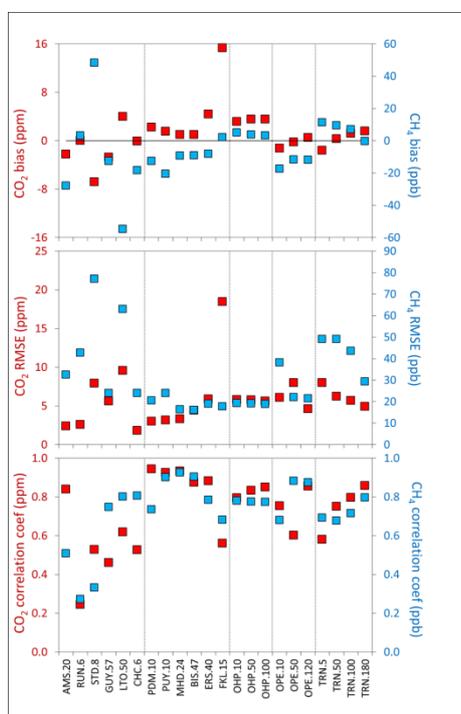


Figure 4.1.8: Annual metrics (bias, RMSE and coefficient correlation) calculated from one year of model-data comparison for daily means of CO₂ (red) and CH₄ (bleu) at the 15 sites (with multiple sampling heights at the last three sites).

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, which is often the case, 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. The largest overestimation occurs in 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 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 onwards the model agrees better with the measurements, especially the overestimation of about 0.5% in May 2015 is not present in May 2016. 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.

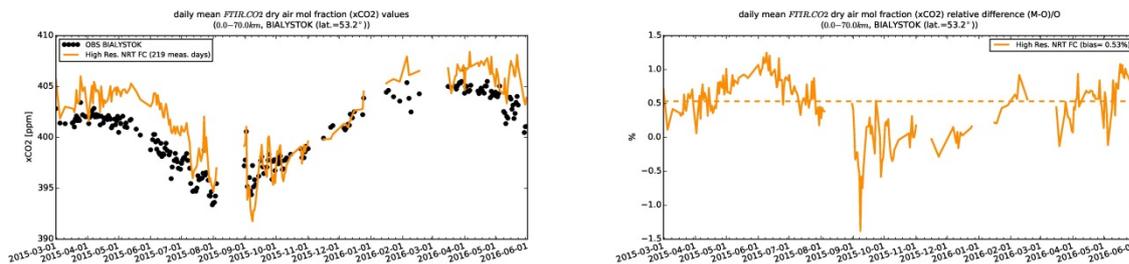


Figure 4.2.1: Time series of column averaged mole fractions (left) and relative difference (right) of carbon dioxide (CO_2) at the TCCON site Bialystok compared to high resolution NRT FC data (yellow)

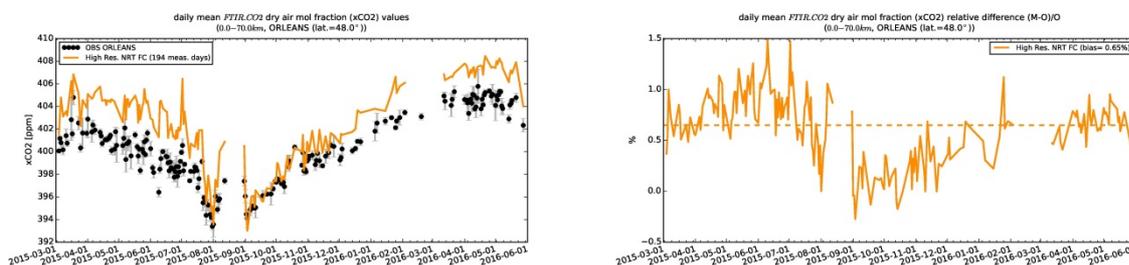


Figure 4.2.2: Time series of column averaged mole fractions (left) and relative difference (right) of carbon dioxide (CO_2) at the TCCON site Orleans compared to high resolution NRT FC data (yellow)

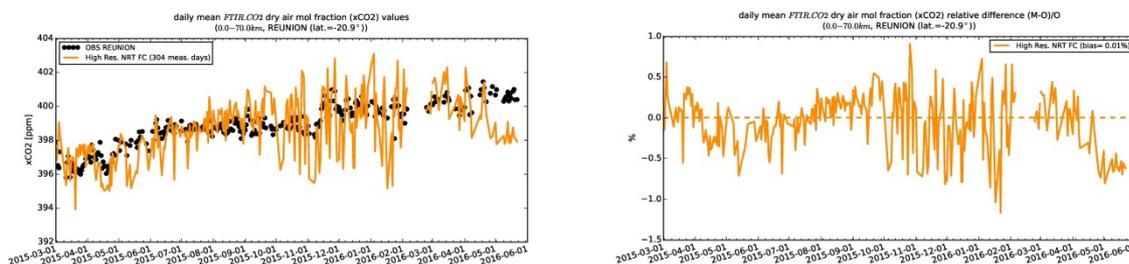


Figure 4.2.3 Time series of column averaged mole fractions (left) and relative difference (right) of carbon dioxide (CO_2) at the TCCON site Reunion compared to high resolution NRT FC data (yellow)

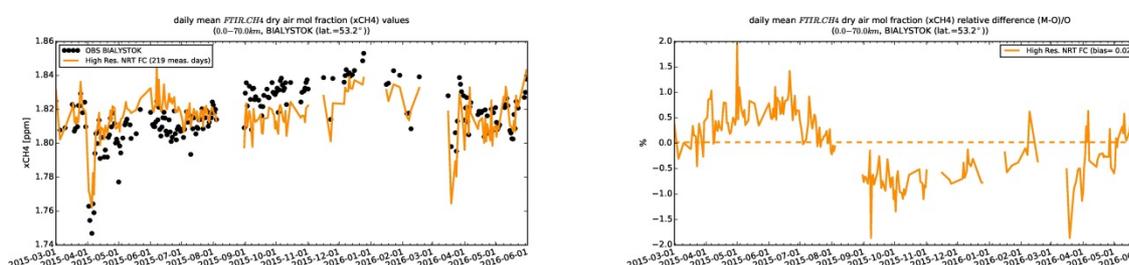


Figure 4.2.4: Time series of column averaged mole fractions (left) and relative difference (right) of methane (CH_4) at the TCCON site Bialystok compared to high resolution NRT FC data (yellow)

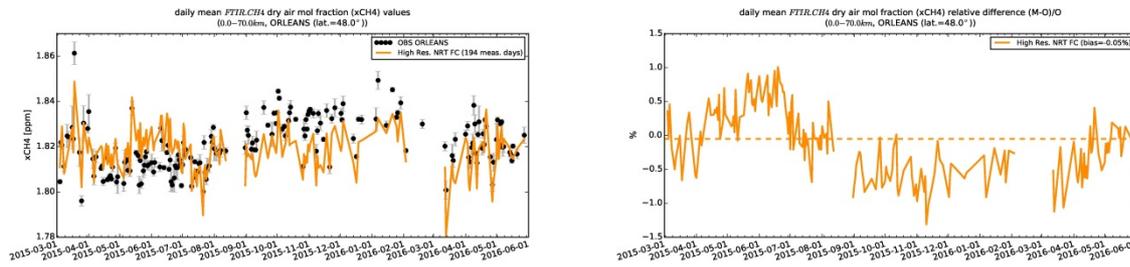


Figure 4.2.5: Time series of column averaged mole fractions (left) and relative difference (right) of methane (CH₄) at the TCCON site Orleans compared to high resolution NRT FC data (yellow)

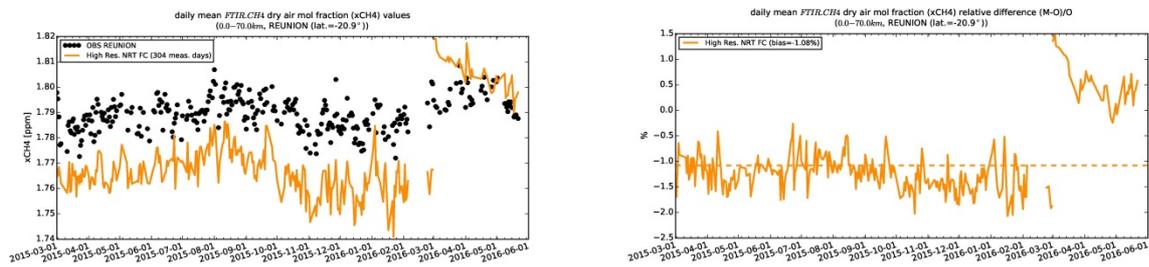


Figure 4.2.6: Time series of column averaged mole fractions (left) and relative difference (right) of methane (CH₄) at the TCCON site Reunion compared to high resolution NRT FC data (yellow).

5 Events

This section describes the validation results of the CAMS NRT global system for events that took place up to May 2016.

5.1 Fire case in the Far East of Russia in May 2016

The Far East of Russia was affected by fires in May 2016, see also <http://atmosphere.copernicus.eu/news-and-media/news/cams-monitors-siberian-wildfires-associated-warmest-month-record>. Analysis of IASI data for one of these fires shows a clear plume of CO, increasing with time, over this region and south-east transport towards the North Pacific Ocean (Fig. 5.1.1). Both model runs, o-suite and control run, captured the location of the plume and the south-east transport of CO. Modelled CO values underestimate IASI data over the easternmost area of Russia and over the Ocean and overestimate satellite data over the Baikal Lake area. Two model runs are very similar. The difference can be seen over the very Far East and over the Ocean where the run without data assimilation shows slightly higher values which fits better with IASI data. Figure 5.1.2 shows a strong overestimation of NO₂ and HCHO fire emissions over east of Russia by the both model runs compared to the GOME-2 satellite data.

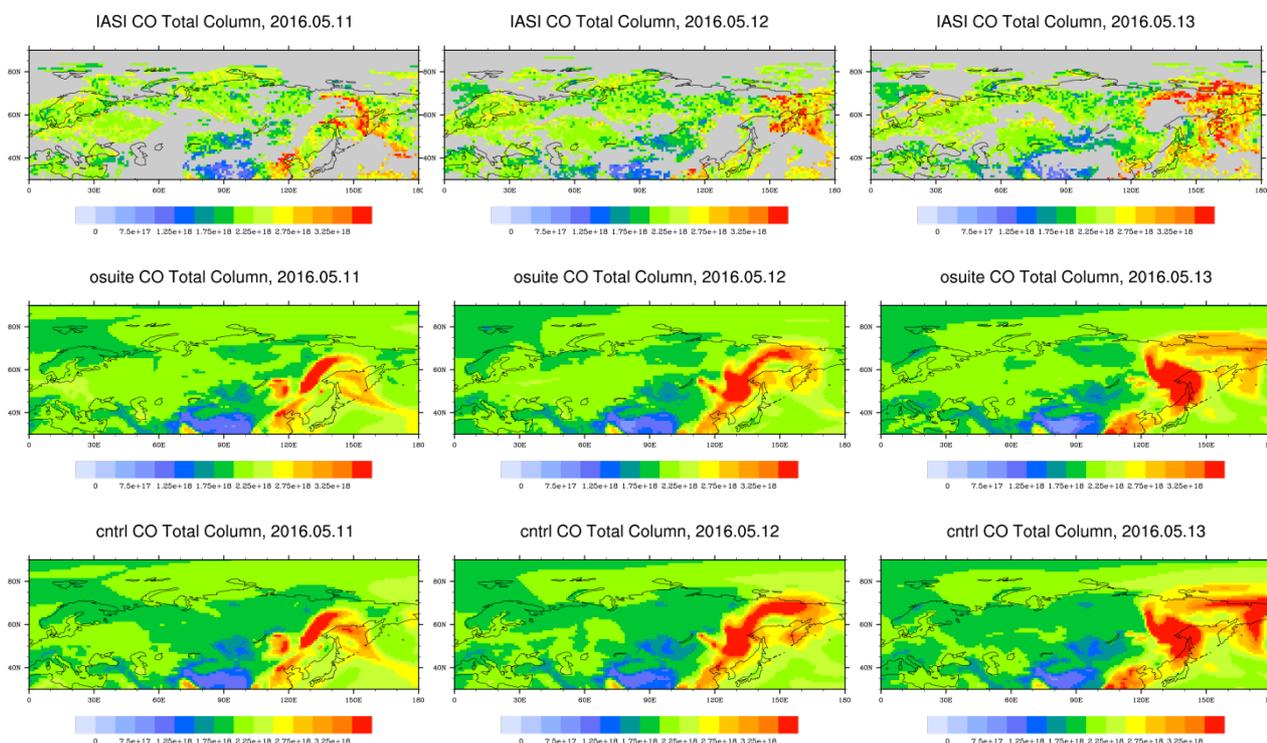


Fig. 5.1.1: CO total column from IASI (top), o-suite (middle) and control runs (bottom) for 11, 12 and 13 May 2016 over the selected region.

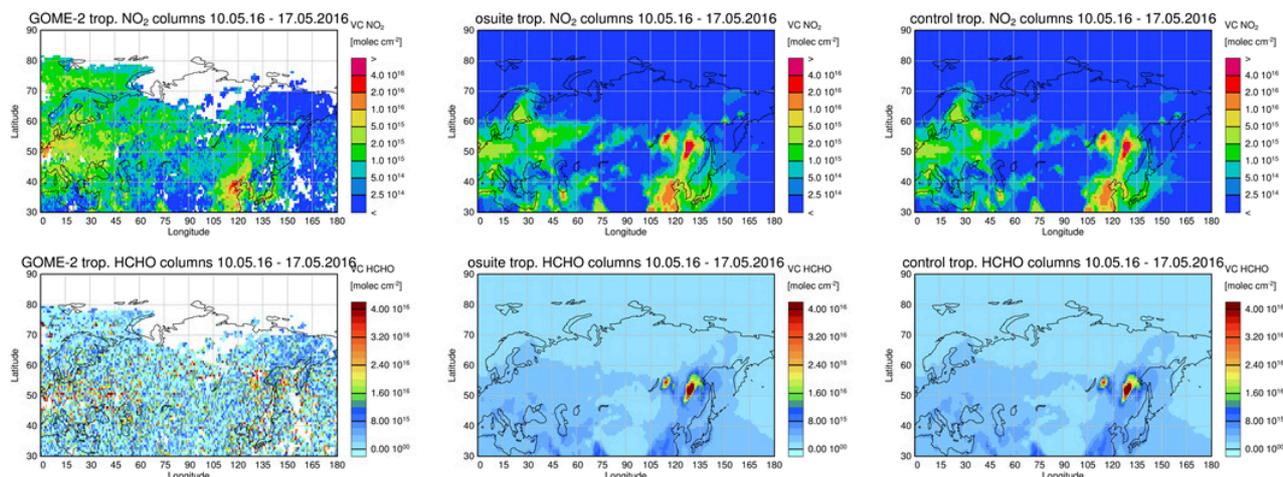


Fig. 5.1.2: Tropospheric NO₂ (top) and HCHO (bottom) columns from GOME-2 (left), o-suite (middle) and control runs (right) averaged for the time period 10-17 May 2016 over the selected region. GOME-2 data were gridded to model resolution (i.e. 0.75° deg x 0.75° deg). Model data were treated with the same reference sector subtraction approach as the satellite data.

5.2 Ozone depletion above the Arctic and Northern Europe, Jan-Feb. 2016

An ozone depletion event developed over the Arctic During the focus period of this report (December 2015 - February 2016) and reached Northern Europe in February. The meteorological conditions leading to this event are described by Braathen (2016). Figure 5.2.1 is copied from this WMO report. It shows, for four dates in the first half of the considered period and on the 40.5 hPa isobaric level, the BASCOE analyses of HNO₃, HCl and ClO+2Cl₂O₂; (based on Aura-MLS observations).

Figure 5.2.2 shows the corresponding ozone analyses by the CAMS o-suite, at the same level as well as in total column. The o-suite analyses show lowered ozone abundances in the Arctic polar vortex in January. Due to the weaker winter polar vortex in the Northern Hemisphere, there are important differences between ozone depletion episodes in the Arctic and in the Antarctic (Solomon et al., 2014). Yet the lower ozone abundance is due to the same processes in both cases: dynamical (downdraft and isolation of ozone-poor air masses in the vortex) and chemical (i.e. catalytic destruction of ozone by chlorine once the vortex is exposed to sunlight and Cl₂O₂ is photolyzed into ClO).

The quantitative attribution of ozone depletion to each process requires satellite observations of ozone chlorine compounds and dedicated model runs (see e.g. Strahan et al., 2013; Hommel et al., 2014). Here we address two more basic questions:

- How well do the o-suite analyses of ozone represent the event observed in January and February 2016?
- Did this event reach European stations in the middle northern latitudes and can the o-suite analyses be used in such an assessment?

Figure 5.2.3 shows eight individual balloon soundings above Ny-Alesund, which is one of the northernmost stations used in this report (79°N). These soundings, which were also used for the averaged Arctic profiles shown in fig. 3.6.2, show that the abundance and vertical distribution provided by the CAMS o-suite agree very well with the ozone

soundings. The only disagreements are ozone overestimations by the o-suite in the upper part (10–60 hPa) of the first sounding shown (2016/01/20) and in the lowermost stratosphere (~ 130 hPa) in the three last soundings (2016/02/10, 2016/02/13, 2016/02/21).

This underestimation of ozone depletion in the leftmost plot of Figure 5.2.3 (2016/01/20, 10–60 hPa) is confirmed by the time series of profiles retrieved from the ground-based Microwave Radiometer (MWR) at the same station, with positive biases reaching ~ 1 ppmv (Figure 5.1.4) i.e. 30 to 40% in the height range 25–30 km (see figure 3.6.3).

During the first days of February, Arctic air masses with depleted ozone and enhanced amounts of chlorine oxides reached as far south as the NDACC station of Bern (47°N). Figure 5.2.5 shows that the agreement between the o-suite and MWR observations is excellent, including during the ozone depletion episode which is clearly associated with elevated chlorine amounts in the BASCOE analyses.

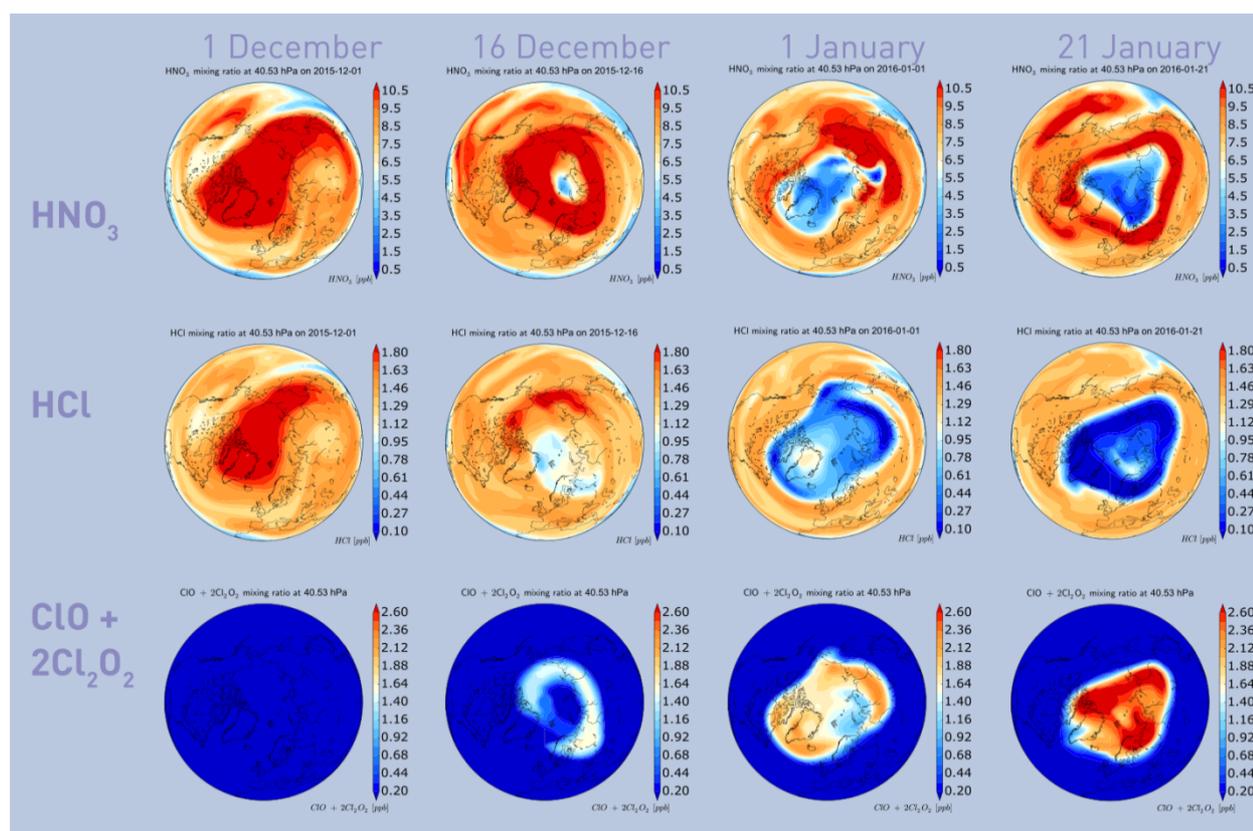


Figure 5.2.1: Results from BASCOE at the level of 40.5 hPa. The upper row shows the mixing ratio of nitric acid, the second row shows the mixing ratio of hydrochloric acid, and the third row shows the sum of chlorine monoxide and its dimer ($\text{ClO} + 2\text{Cl}_2\text{O}_2$). All three rows show the temporal development from 1 December to 21 January with intermediate frames shown for 16 December and 1 January. Copied from figure 8 in Braathen (2016).

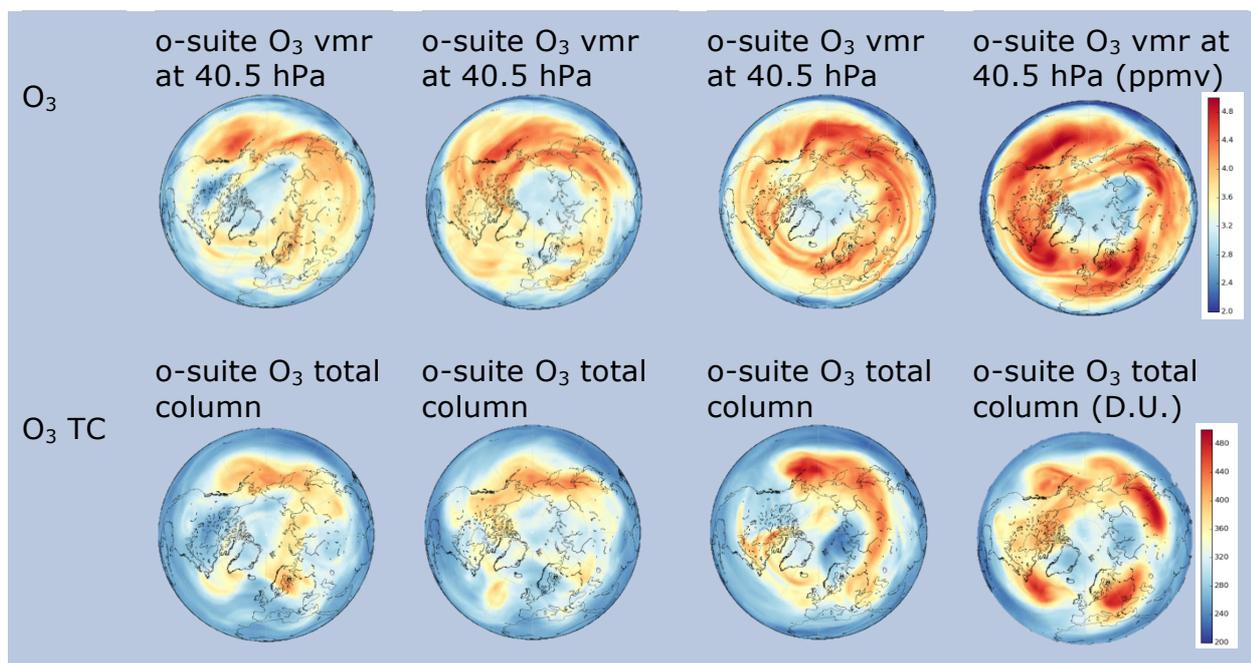


Figure 5.2.2: Results from the CAMS o-suite, on the same dates as the previous figure. The upper row shows the analyses of ozone mixing ratio at 40.5 hPa and the lower row shows the total column of ozone.

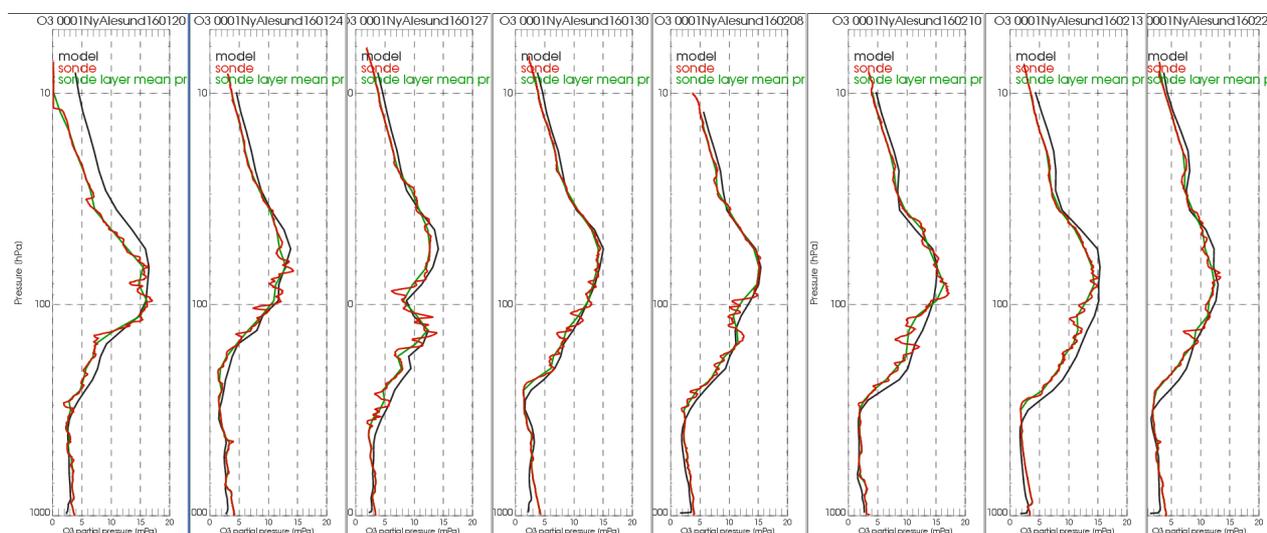


Figure 5.2.3: Eight individual ozone soundings above Ny-Alesund from 2016/01/20 until 2016/02/21. Note the non-standard color scheme: black lines, o-suite analyses; red lines, sonde observations; green lines, sonde observations smoothed to allow quantitative comparison with the o-suite.

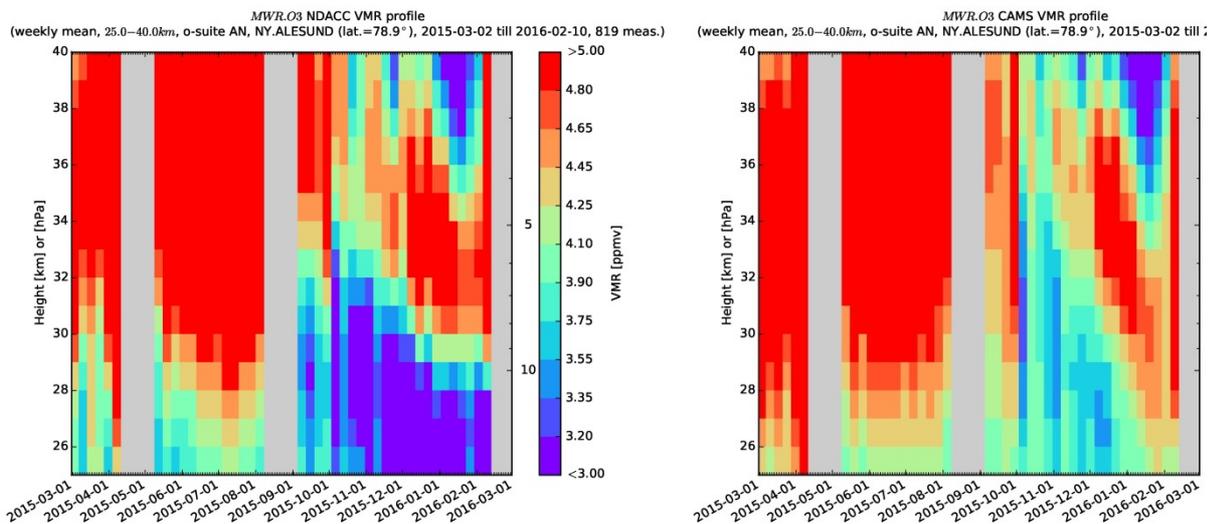


Figure 5.2.4: Ozone volume mixing ratio above Ny-Alesund as a function of time and height: NDACC MWR observations (left) versus o-suite analyses (right)

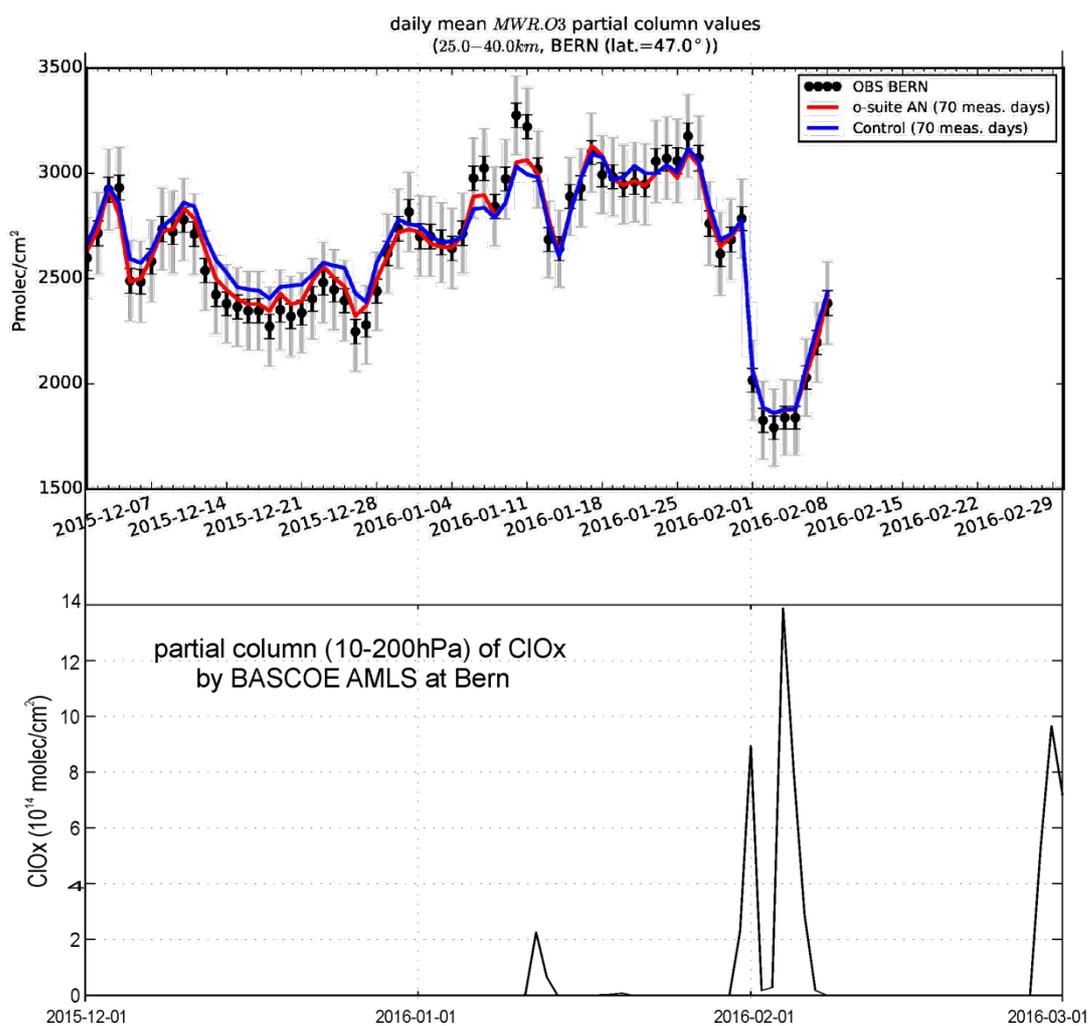


Figure 5.2.5: Time series) above Bern (Switzerland) of the ozone partial column (25-40km; red line shows o-suite analyses, symbols show the MWR observations) and of $\text{ClO} + 2 \cdot \text{Cl}_2\text{O}_2$ partial column (10-200 hPa; BASCOE analyses of Aura-MLS observations of ClO).

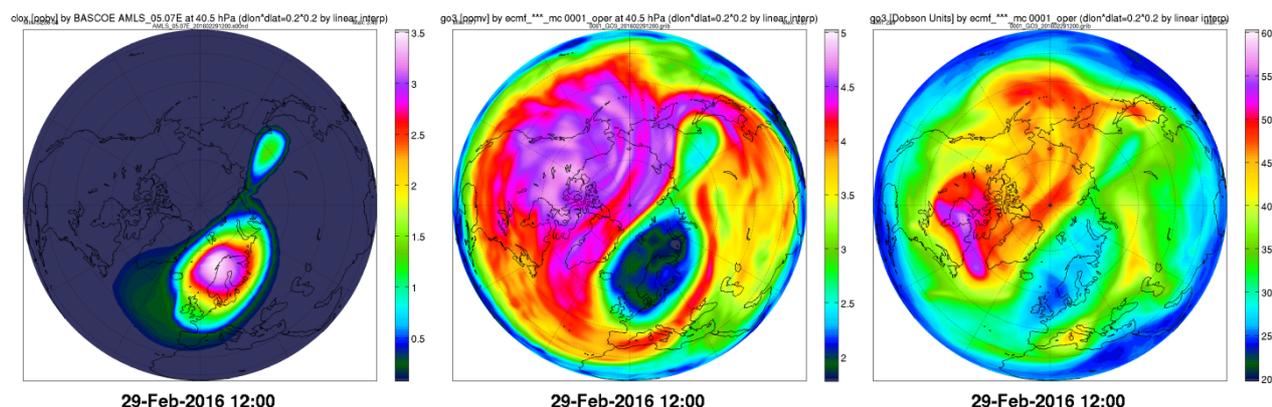


Figure 5.2.6: BASCOE analysis of ClO_x at 40.5 hPa (left), CAMS o-suite analysis of ozone at 40.5 hPa (middle) and CAMS o-suite analysis of ozone total column (right), on 29 February 2016. The ozone depletion event clearly covers Northern Europe.

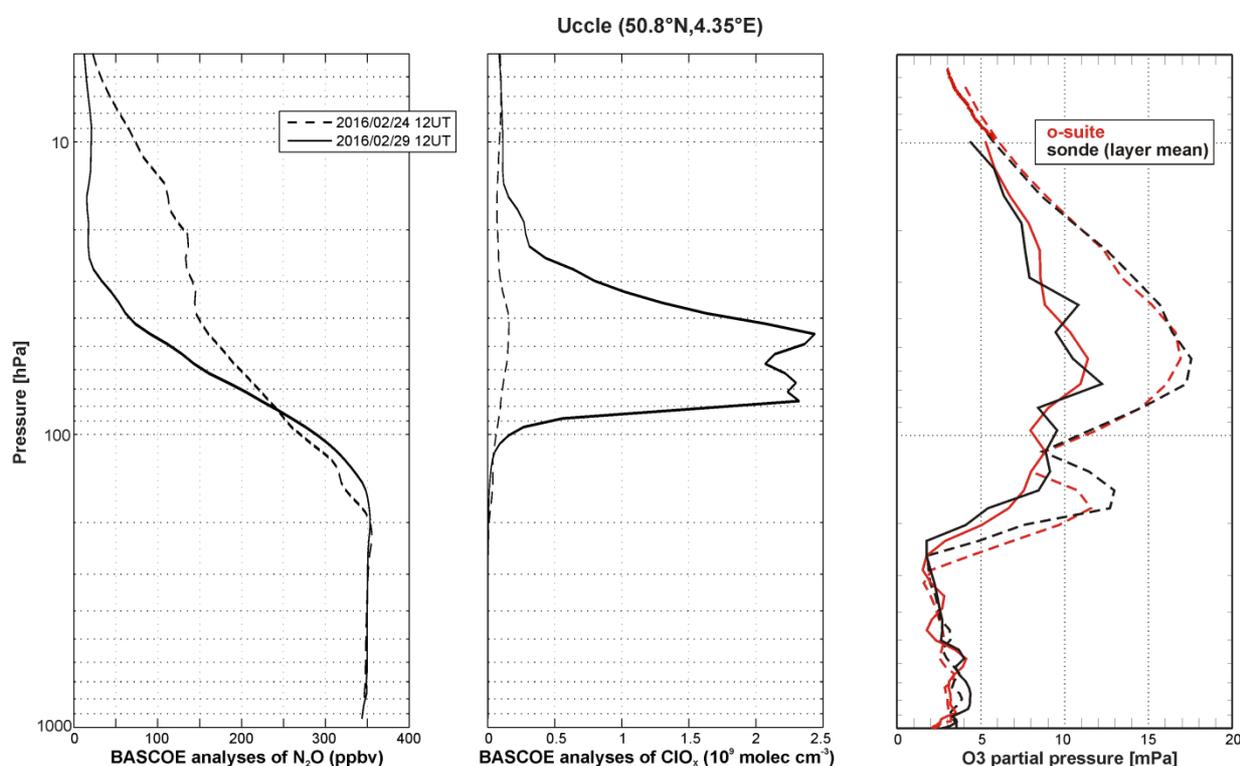


Figure 5.2.7: Vertical profiles above Uccle on 2016/02/24 (dashed lines) and 2016/02/29 (solid lines): BASCOE analyses of N_2O (left plot) and ClO_x (middle plot); ozone partial pressure (right plot) in the o-suite analyses (red lines) and as observed by ozonesondes (black lines)

On 29 February 2016, the last day studied in this report, the Arctic vortex reached the European mid-latitudes again, which could be seen even on a map of total column ozone (Figure 5.2.6). An ozone sonde was launched from Uccle on that day. It shows significantly lower values than 5 days earlier, both profiles agreeing very well with the corresponding o-suite analyses (Figure 5.2.7). The correlation with vortex air masses can be verified with the BASCOE analyses of the tracer N_2O and of active chlorine ClO_x which are shown on the same figure.

To summarize, during this depletion episode the o-suite analyses overestimated stratospheric ozone above Ny-Alesund as observed by ozone sondes and Microwave

Radiometers but according to ozone sondes this disagreement became much less severe after 20 January. The comparison with OMPS-LP in the Arctic latitude band (Figure 3.6.5) confirms this assessment with a bias varying quickly, from 10% until 1 February to 2-6% afterwards. Above the Northern European stations of Bern and Uccle the biases were much smaller - even when the polar vortex reached these latitudes. From a seasonally averaged point of view (DJF 2015-2016), the relative bias between MWR instruments and the o-suite reached 14.5% above Ny-Alesund but only 0.7% above Bern (Table 3.6.1).

5.3 A dust event over the Western Mediterranean in May 2016

The selected dust event corresponds to the period from May 9th to 13th, 2016. Around the 9th May, a layer dust plume moved from Algeria covering the area around the Central Mediterranean and moving later eastwards Eastern Mediterranean.

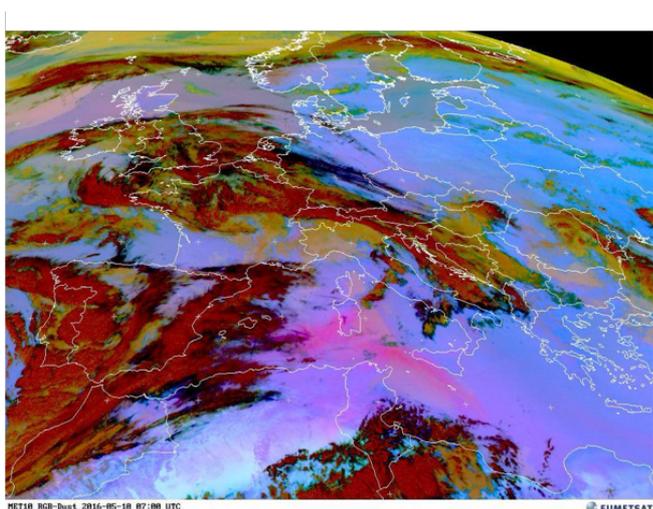


Figure 5.3.1. MSG/SEVIRI RGB Dust image (red is clouds and pink is dust), 10 May 7:00 UTC.

Three-hourly dust aerosol optical depth (DOD) from CAMS o-suite (Benedetti et al., 2009; Morcrette et al., 2009) has been compared with AOD from MODIS in order to see the skill of CAMS o-suite to mimic the spatio-temporal evolution of the dust plume (Figure 5.3.1). 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 9th to 13th May 2016 at 12UTC is shown in Figure 5.3.2. Moreover, DOD values from CAMS o-suite have been compared with those from CAMS control, Multi-Median model and AERONET AOD (Level 1.5) in four AERONET stations strategically located along the path of the dust plume over at Tizi-Ouzou (Algeria), Ben Salem (Tunisia) and Etna (Italy). Results are shown in Figure 5.3.3. Finally, visibility from METAR or SYNOP stations processed by the SDS-WAS NAMEE Regional Node is used to qualitatively evaluate the modelled CAMS surface concentration.

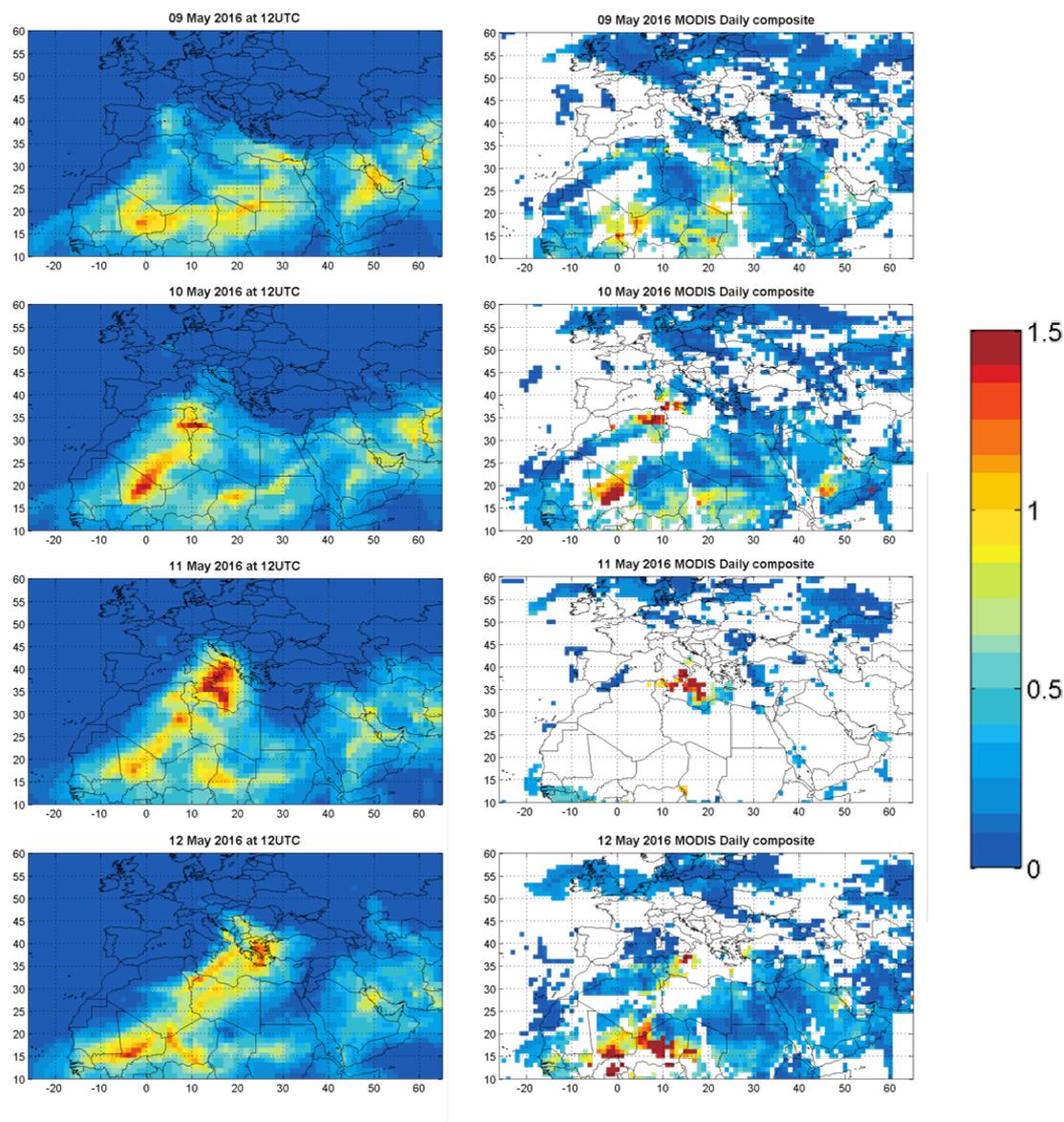


Figure 5.3.2. AOD from MODIS combined product form Dark Target and Deep Blue (left column), and DOD from o-suite (right column), for January 16th-20th, 2016 at 12UTC.

Dust mobilization is confined over Algeria on 9th May caused by a low pressure over North Africa. Low AOD and high AE are observed over Tunisia (see Ben Salem in Figure 4.2.3) on May 8th. The next days, dust is blown over Central Mediterranean affecting Tunisia and Italy from 10 to 12 May (see Figure 5.3.2). The Ben Salem and Etna AERONET sites show consecutively an increase in AOD (AOD up to 2.5 in Ben Salem) associated to a sharp decrease of AE associated with the arrival of the dust plume.

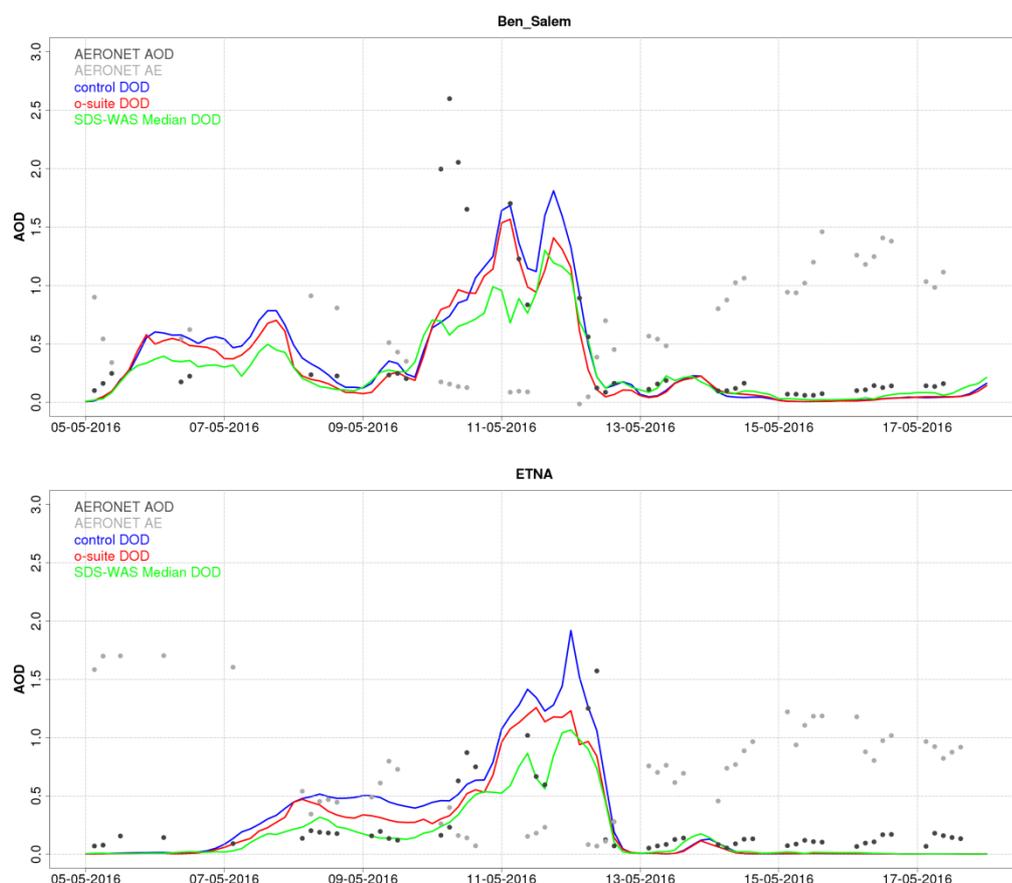


Figure 5.3.3 AOD at 550 nm from AERONET (black), DOD at 550 nm from the o-suite (blue), DOD at 550 nm from the control run (red), and DOD at 550 nm from SDS-WAS Multi-model Median (green) at Ben Salem (Tunisia) and Etna (Italy) AERONET sites during the case analysis from 5th to 18th May 2016.

The o-suite is able to timely reproduce the spatial distribution of the dust plume as observed by MODIS (Figure 5.3.2) over the Mediterranean. We can see how CAMS o-suite tracks fairly well the changes in both shape and size of the dust layer throughout the episode. The whole episode is well simulated by CAMS o-suite and CAMS control, although in Ben Salem none of the models could reproduce the AOD peak observed on May 10th (see Figure 5.3.3). The observed peak can be associated to cloud contamination because dense clouds were present during all the event (see a long cumulus line over Central Mediterranean in Figure 5.3.1).

Since weather records have an excellent 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 is able to track the reduction of visibility of this event (see Figure 5.3.4) localizing the origin of the event over Algeria on May 9th, moving northwards Tunisia on May 10th, crossing the Mediterranean and achieving Italy on May 11th and Greece and Turkey on May 12th.

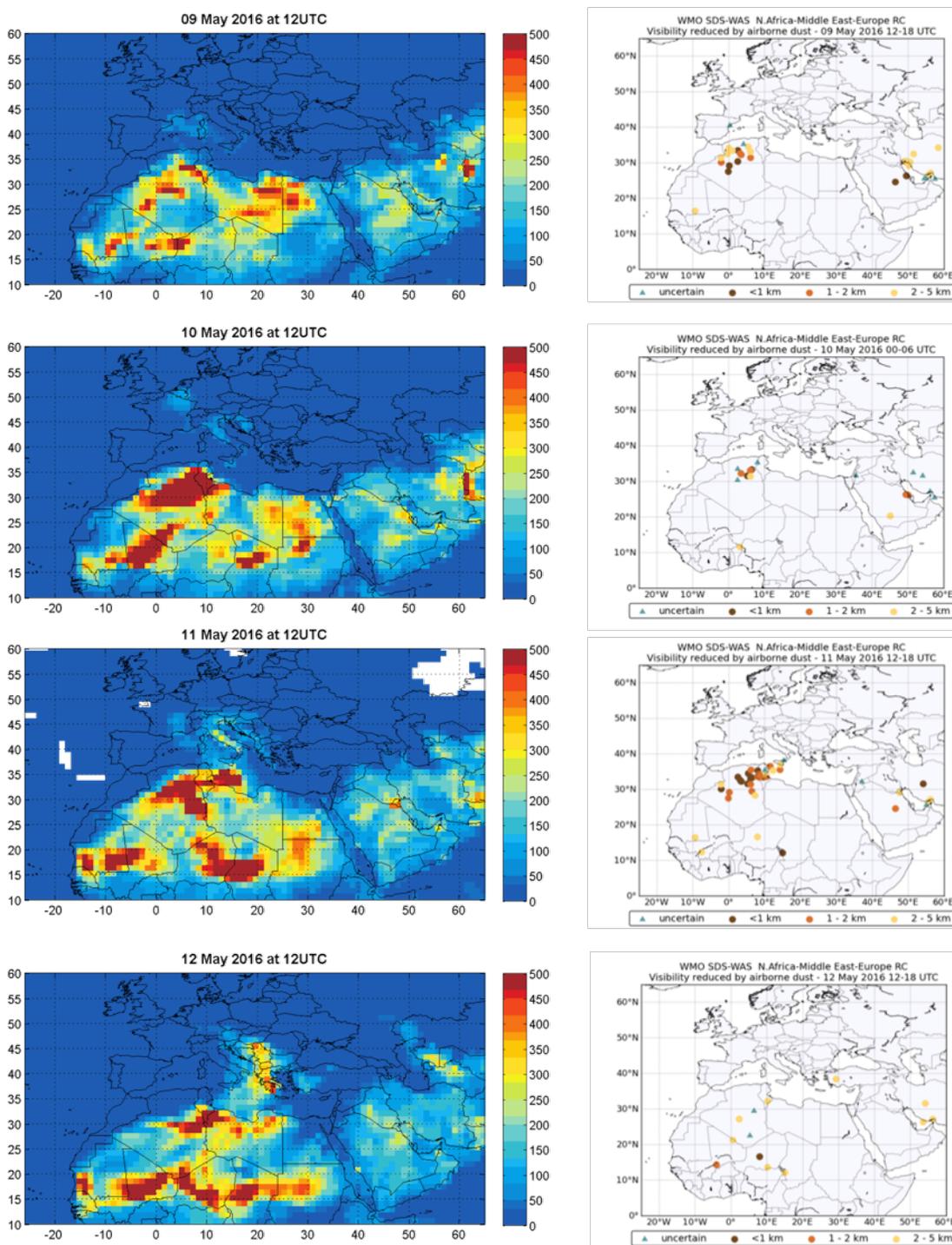


Figure 5.3.4. Dust surface concentration o-suite (left column) and visibility (right column) from METAR or SYNOP stations from SDS-WAS Regional Center from 10th to 12th May 2016. The maps show cases of horizontal visibility reduction by sand or dust to less than 5 km reported in METAR or SYNOP bulletins. 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 Teyssède, 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., Chabrilat, 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., Seftor, 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., Viscardy, 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/grg>

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 Chabrillat, 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 Slutker 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 Maito, 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.

We would like to thank the operators of the IASOA Network for access to surface ozone measurements.

The authors acknowledge Environment and Climate Change Canada for the provision of Alert ozone data.

We acknowledge the National Aeronautics and Space Administration (NASA), USA for providing the 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