



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 March 2016

Issued by: KNMI

Date: 30/05/2016

REF.: CAMS84_2015SC1_D.84.1.3-2016Q2_201605





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 March 2016

Editors:

V. Huijnen (KNMI), H.J. Eskes (KNMI), A. Wagner (DWD), M. Schulz (MetNo), S. Chabrillat (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), M. Ramonet (LSCE), A. Richter (IUP-UB), M. Schulz (MetNo), N. Sudarchikova (MPI), V. Thouret (CNRS-LA), A. Wagner (MPI), T. Warneke (UBC), C. Zerefos (AA)

Report of the Copernicus Atmosphere Monitoring Service,
Validation Subproject (CAMS-84).

Date:

30 May 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 March 2016. V. Huijnen, H.J. Eskes, A. Wagner, M. Schulz, S. Chabrillat, M. Ramonet, S. Basart, A. Benedictow, A.-M. Blechschmidt, Y. Christophe, 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.84.1.3_2016Q2_201605, May 2016.

Available at:

http://atmosphere.copernicus.eu/quarterly_validation_reports



Executive Summary

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 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 March 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 December 2015-February 2016 was excellent, with 100% 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 -5% and 20% for stations in the Tropics (Fig. S1). This is an improvement compared to the control experiment without the assimilation of composition observations. For December-February 2016 good agreement is found over the NH mid latitudes, which is confirmed with IAGOS evaluations over Frankfurt and Dusseldorf, although here the gradients in the UTLS are better represented by the control run. Larger biases are found over Antarctica by up to $+30\%$ in January 2016.

The o-suite shows an underestimation of surface ozone for Europe during December and February 2016 with MNMBs of up to -20% . For Asia, the o-suite shows a good agreement with surface ozone MNMBs within $\pm 10\%$. For Southern hemispheric stations, the o-suite overestimates surface ozone mixing ratios with MNMBs between 10% and 30% , whereas the control shows strong negative MNMBs (-20% - 30%) at the surface as well as in the free troposphere. At none of the surface stations the o-suite

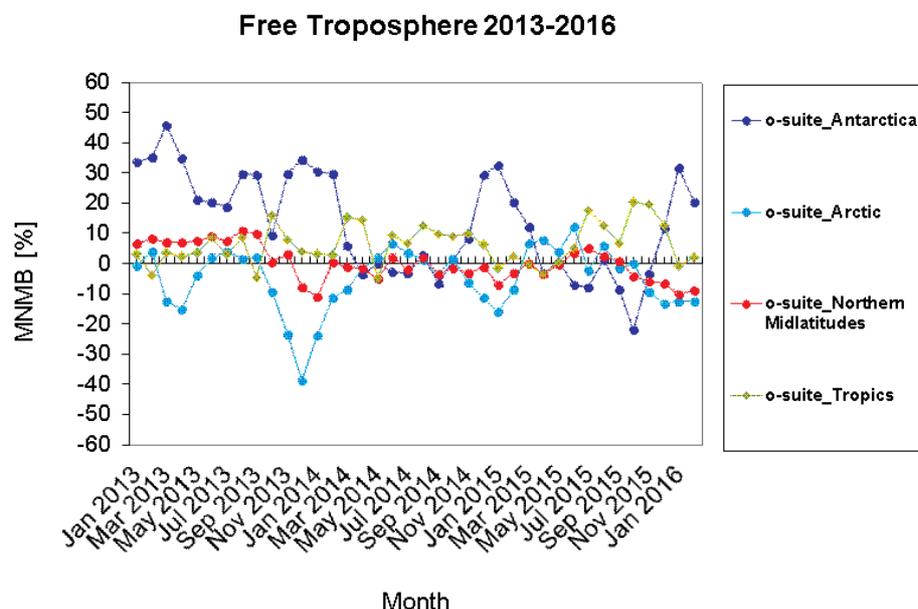


Figure S1: Time series of MNMB of ozone in the o-suite, compared against ozone sondes, averaged over different latitude bands.

does not show a significant drift during the 5-day forecast period, except for a few specific locations near mountainous regions.

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 compared to 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 -8%

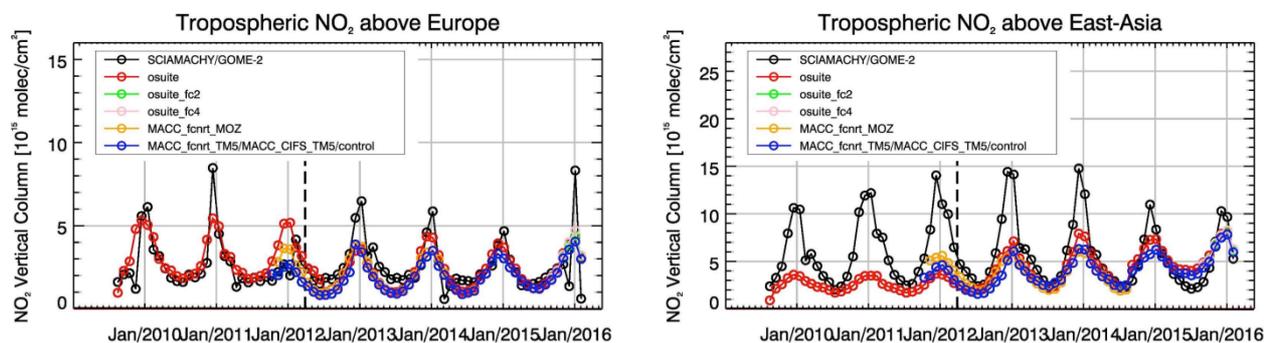


Figure S2: Time series of tropospheric NO_2 columns from SCIAMACHY (up to March 2012), GOME-2 (from April 2012 onwards) compared to model results for Europe and East-Asia. The o-suite is in red, control is in blue (before Sept 2014 blue and yellow represent older model configurations). Forecast results are given in green (D+2) and pink (D+4).

and -18%). During the fire season over Alaska and Siberia negative biases are only up to -5%.

The o-suite as well as the control overestimate CO total columns in the tropics and SH. Significant positive biases up to 50% for control are found over the central part of South America and for Indonesia. The 4-day forecasts over Northern equatorial Africa show furthermore a rapidly increasing negative bias with respect to MOPITT CO columns, likely associated with too low fire CO emissions, as confirmed by IAGOS measurements. Contrary, over Maritime south-east Asia a positive bias grows quickly, suggesting a too large CO source. In other regions the forecasts are almost 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 NO_2 , 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

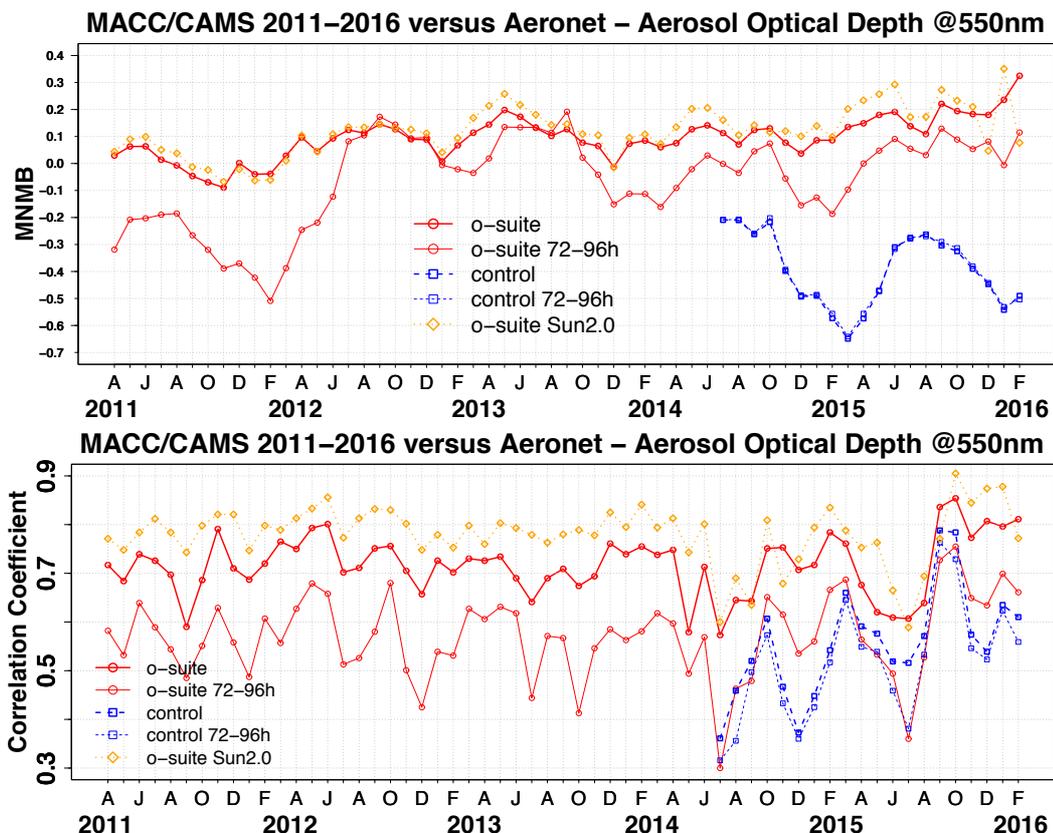


Figure S3. Aerosol optical depth at 550nm in IFS model simulations for April 2011 - February 2016 against daily matching Aeronet NRT level 1.5 and level 2.0 data a) Modified normalized mean bias (MNMB); o-suite (thick red curve); o-suite at last forecast day (light red curve); Control (blue dashed); 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 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. 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 different sub-regions of North Africa are compared with AERONET sun photometers observations and the SDS-WAS multi-model ensemble. During the period of analysis, dust activity over desert dust sources (Middle East, the Sahara and dust corridor of North-western Maghreb) was relatively low. The o-suite reproduces the dust peaks in areas remote from the sources, but underestimates DOD observations over the Sahel.

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. The model version with experiment id 'gcbt' compared to TCCON data shows several outliers in early 2015 and an abrupt jump in CH₄ levels associated to a change in the GHG initial conditions on the 24 September 2015. Because of these problems, a new high resolution experiment was run for 2015 (gf39) which is the one that is now made available to users. Overall, in comparison to ICOS station data the 'gf39' model version underestimates the amplitudes of the CO₂ diurnal cycle, and overestimates that for CH₄. For CH₄ the disagreement is larger, probably due to the higher variability of the emissions. The model validation suggests a better agreement with tall towers located in flat terrain, compared to mountain or coastal sites. Few surface stations emerge as outliers, due to the difficulty for the model either to reproduce local transport (e.g. mountain or coastal sites), or local emissions.

At the annual scale, both surface and total column measurements show a significant seasonal variability of the biases in the northern hemisphere. The bias indicates a one month shift in the simulated growing season, resulting in too high CO₂ concentrations in spring, and too low concentrations in summer. However a similar seasonality of the biases is observed for CH₄ whose emissions are not driven by the growing season. This feature will need more analysis, based on simulations for longer periods.

System performance in the Arctic

The CAMS model runs are validated using surface ozone measurements from the ESRL-GMD and the IASOA networks (3 sites) and ozone concentrations in the free troposphere are evaluated using balloon sonde measurement data. More data sets will be added to the validation in the next report.



For the period from December 2014 to February 2016 the simulations of the surface ozone concentrations are on average in good agreement with the observations apart from ozone depletion events in March – June 2015, 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 December 2015 – February 2016 period the surface ozone concentrations are underestimated at the three sites by between -1% and -17% (o-suite) and 0% and -10% (control). It seems that data assimilation reduces the model performance in terms of bias at the Arctic stations. The short term variability is well captured at the three sites with correlation coefficients between 0.55 and 0.75 for the o-suite and 0.5 and 0.7 for the control run. In the free troposphere the MNMB of ozone is mostly within -15% for both the o-suite and the control run.

System performance in the Mediterranean

The CAMS model run performance in the Mediterranean is evaluated using AOD observations from 42 AERONET sites. In the Mediterranean the CAMS o-suite reproduces the daily variability of AERONET observations, but tends to underestimate the background levels over Southern European sites (more influenced by urban/industrial aerosol sources). CAMS experiments reproduce the daily PM10 and PM2.5 variability of the most intense aerosol events observed by Airbase background stations in the Western and Central Mediterranean, which are associated with long-range desert dust transport. However, the PM10 and PM2.5 magnitude predicted by the CAMS simulations (control and o-suite) is significantly overestimated.

The model is further 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 up to -50%) and Eastern Mediterranean (MNMBs close to zero). Temporal correlation coefficients between simulated and observed surface ozone mixing ratios range generally between 0.4 and 0.8.

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 $\pm 10\%$. The o-suite shows a negative bias in surface ozone concentrations in Europe, with MNMBs for GAW and ESRL stations ranging between -20% and 0% between December 2015 and February 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 -10%. 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.

O3 relative bias against OMPS-LP: 30-70hPa mean from 20150301 to 20160301

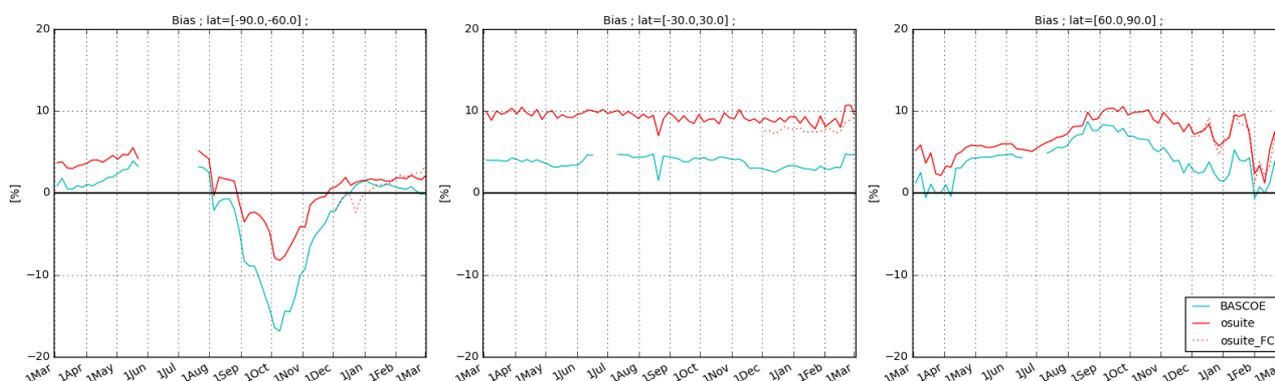


Figure S4: Time series of the normalized mean bias (%) between ozone from o-suite analyses (red, solid) or 4th day forecasts (red, dotted), or BASCOE analyses (cyan) and OMPS-LP satellite observations, in the middle stratosphere (30-70hPa averages).

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. 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. Above the Arctic the bias varied rapidly during the three last months, i.e. while an ozone depletion event took place (see below).

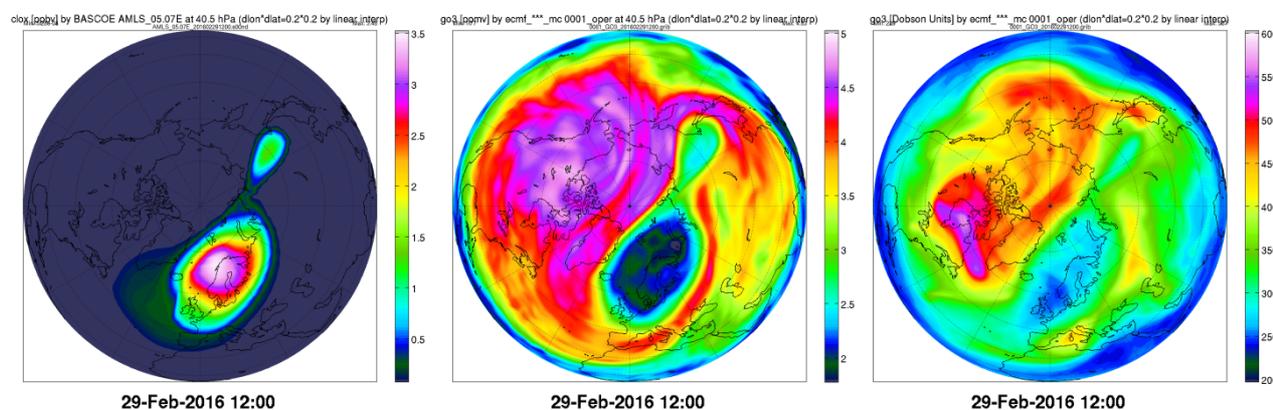


Figure S5: BASCOE analysis of ClO_x at 40.5 hPa (left), CAMS o-suite analysis of ozone at 40.5 hPa (middle) and CAMS o-suite analysis of ozone total column (right), on 29 February 2016. The ozone depletion event clearly covers Northern Europe.

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 NO_2 , have been evaluated but the results are not presented here.

Events

A *stratospheric ozone depletion event* developed over the Arctic during the focus period of this 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 17th January, a *strong dust plume* originated in Lybia moving eastwards and covering part of the Eastern Mediterranean and Middle East. Both CAMS o-suite and control captured quite well the passage of the dust plume on the AERONET stations in accordance with the SDS-WAS Multi-Median model. The DOD spatio-temporal evolution given by CAMS o-suite closely resembled the MODIS retrievals.

On February 26 the o-suite and the control run forecasted a *biomass burning plume* with high CO values over Western US, which was not observed by IASI retrievals. It turned out that between 26 February and 1 March inaccurate data was used from the MODIS instrument, resulting in erroneous GFAS fire emissions for these dates.



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.....	36
3.3	Carbon monoxide	40
3.4	Formaldehyde	56
3.5	Aerosol.....	60
3.6	Stratospheric ozone	76
3.7	Stratospheric NO ₂	82
4	Validation results for greenhouse gases.....	84
4.1	CH ₄ and CO ₂ validation against ICOS observations	84
4.2	CH ₄ and CO ₂ validation against TCCON observations	88
5	Events	92
5.1	Ozone depletion above the Arctic and Northern Europe, January-February 2016 92	
5.2	A dust event over Eastern Mediterranean-Middle East in January 2016.....	97
5.3	A false CO plume over North America, end of February 2016	99
6	References	101
	Annex 1: Acknowledgements.....	105



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) until 1 March 2016.

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>. Also dedicated netcdf files for use as boundary conditions in regional AQ models world-wide can be downloaded from Forschungszentrum Jülich, <http://join.iek.fz-juelich.de/macc>. 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,



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

Reactive gases – Troposphere
<p>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/kefak/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</p>
Reactive gases - Stratosphere
<p>Stratospheric composition: http://www.copernicus-stratosphere.eu NDACC evaluation in stratosphere and troposphere (the NORS server) http://nors-server.aeronomie.be</p>
Aerosol
<p>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</p>
Satellite data monitoring
<p>Monitoring of satellite data usage in the Reanalysis and Near-Real-Time production: http://copernicus-atmosphere.eu/d/services/gac/monitor/</p>

<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.

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_fcrrt_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

Starting from 18 September 2014 the o-suite consists of the C-IFS-CB05 chemistry combined with the MACC aerosol model. The chemistry is described in Flemming et al. (2015), aerosol is described by the bulk aerosol scheme (Morcrette et al., 2009). Dissemination of o-suite forecasts is at 22:00UTC. The forecast length is 120 h. The o-suite data is stored under expver '0001' of class 'MC'. On 3 September 2015 an update has been taken place, where the meteorological model has changed significantly, moving from cy40r2 to cy41r1. 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

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	20150903-present (g9rr) 20140918-20150902 (g4e2) 20120705-20140917 (fnyp)
Control	geuh g4o2	control FC run for g9rr / g4e2, without DA	20150901-present (geuh) 20140701-20150902 (g4o2)
GHG run	gcbt gf39 ghqy	High resolution, NRT CO ₂ and CH ₄ runs without DA	20150101-20151231 (gcbt) 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	V02 V3.4	O3 Profiles	20090901 – 20130107 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 6 layer profiles O3 21 layer profiles	20090901 - 20121006 20121007 -
IASI	MetOp-A	LATMOS/ULB	-	CO Total column	20090901 -
IASI	MetOp-B	LATMOS/ULB	-	CO Total column	20140918 -
MOPITT	TERRA	NCAR	V4 V5-TIR	CO Total column	20120705-20130128 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 -

- The 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 the 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).



Figure 2.1: Satellite observation usage in the real-time analysis, from Aug 2012 onwards. CO: Top three rows; O3 columns and profiles: rows 4, 8-14; SO2: rows 5-7, Aerosol Optical Depth: rows 15-16. 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 MACC 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 April 2016 analyzed in this report correspond to three experiments:

- "gcbt" from 1 January 2015 till February 2016. In this experiment fields were re-initialized on 24 September 2015. Spurious jumps were found in this experiment (See Sec. 4)
- "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.



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 (TBC)	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 via <http://atmosphere.copernicus.eu/user-support/operational-info>. 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 04.03):

- The following species are assimilated: O₃, H₂O, HNO₃, HCl, HOCl, N₂O. ClO has been added on 20130819.
- 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 137 levels extending from 0.01 hPa to the surface.



- Winds, temperature and surface pressure are interpolated in the ECMWF operational 6-hourly analyses.
- Timesteps 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, SACADA 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 (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>) and on the Jülich server (<http://join.iwk.fz-juelich.de/macc>, in netcdf format).

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



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 February 2015 to February 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 from during summer and spring (MNMBs up to 12%), while during the winter season the MNMBs gets negative, with MNMBs mostly within -15 %, Fig. 3.6.1. Over the NH mid-latitudes MNMBs for the o-suite are on average close to zero all year round (maxima are -10% to +10%), which is a clear improvement compared to the control run, which shows larger positive MNMBs (up to 20%) 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 -5% and 20% for the o-suite, and between -10% and 25% for the control run.

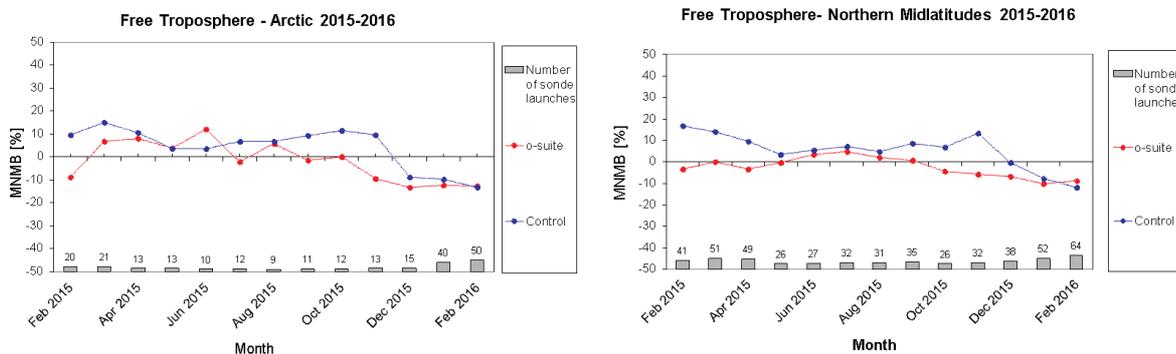


Figure 3.1.1: MNMBs (%) of ozone in the free troposphere (between 750 and 300 hPa) from the IFS model runs against aggregated sonde data over the Arctic (left) and the Northern midlatitudes (right). The numbers indicate the amount of individual number of sondes.

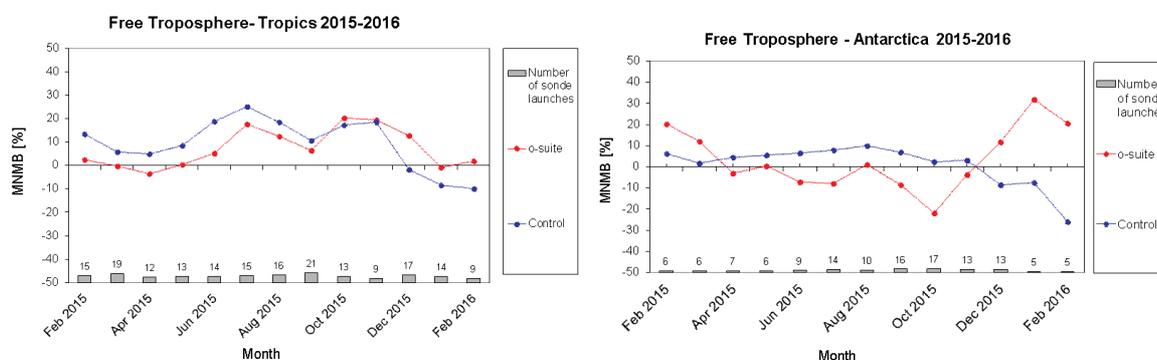


Figure 3.1.2: MNMBs (%) of ozone in the free troposphere (between 750 and 200 hPa (Tropics) / 300 hPa) from the IFS model runs against aggregated sonde data over the Tropics (left) and Antarctica (right). The numbers indicate the amount of individual number of sondes.

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 December 2015 to February 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 Figure 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 flights within the period was 720 (1440 profiles) giving a performance of 57%. The actual number of profiles with usable data was 512 (36% of the total possible). These flights are shown in Figure 3.1.3 (top). Fifty eight percent (58%) (298 profiles) of the operational flights had usable measurements of ozone and 60% of flights had usable CO.

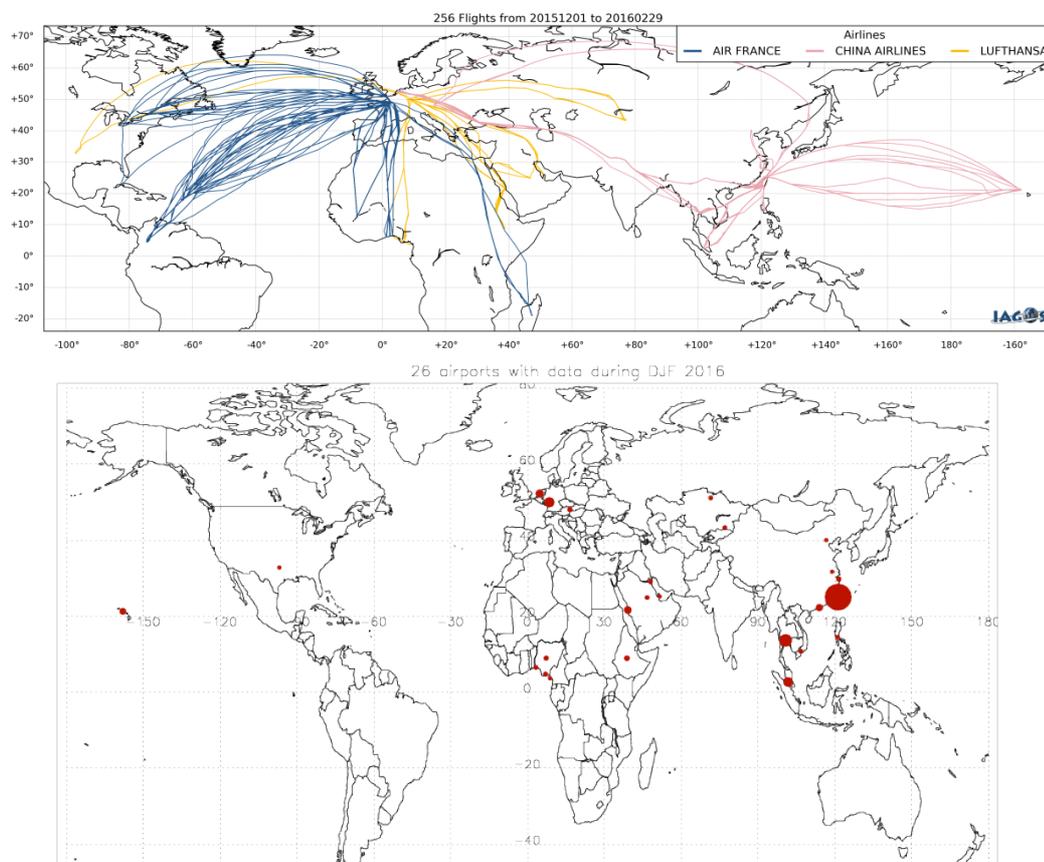


Figure 3.1.3: Map of the flights (top) and the visited airports (bottom) during the period December – February 2015/6, by the IAGOS equipped aircraft. The size of the plotting circle represents the number of profiles available.

Delivering these O_3 and CO data are an aircraft from Lufthansa (based in Frankfurt), and an aircraft from China Airlines based in Taipei. The Airfrance aircraft has no valid O_3 or CO measurements at the moment. This report therefore displays profiles recorded by these aircraft, covering mainly the routes served by China Airlines across South-East Asia as shown on the map in Figure 3.1.3 (with a plotting circle scaled to the highest number of flights at an airport). Apart from the home-base airport of Taipei, most of the airports are only visited at low frequency during this period.

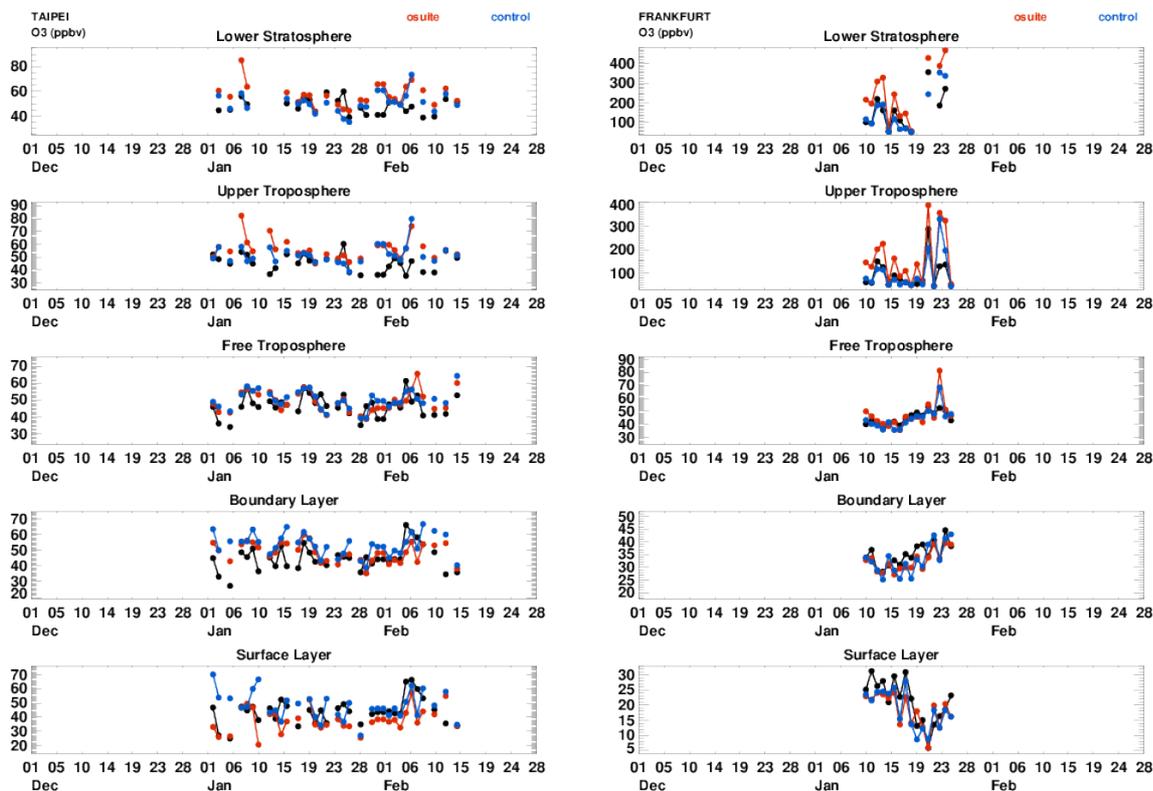


Figure 3.1.4: Time series of daily mean ozone over Taipei (left) and Frankfurt (right) from Dec. 2015 to Feb. 2016 for 5 layers, Surface, Boundary layer, Free Troposphere, Upper Troposphere and Lower Stratosphere.

Figure 3.1.4 takes advantage of the China Airlines aircraft operating from Taipei. Generally the observations match better with the models at the Asian airports than at the European airports. Also shown is the ozone at Frankfurt. In the surface layer at Taipei, the ozone is slightly underestimated by the o-suite whereas in the boundary layer, the o-suite matches better, and the control slightly overestimates the ozone. Otherwise there is not much difference between the observations and the two runs.

Europe

Examples in Figure 3.1.5 have been chosen to show the ability of the control run to capture the gradient in the UTLS. The control run does systematically better than the o-suite both at times when a strong gradient is encountered and at times when little gradient is encountered. On 11th January, there were very similar profiles at Frankfurt and Amsterdam taken by two different aircraft, showing that the tropopause was encountered at around 9000m. In both cases, the control does quite well but the o-suite overestimates the amount of ozone, and places the tropopause a little too low in altitude. On 16th January these two aircraft again visit Frankfurt and Amsterdam. The profiles are similar to those

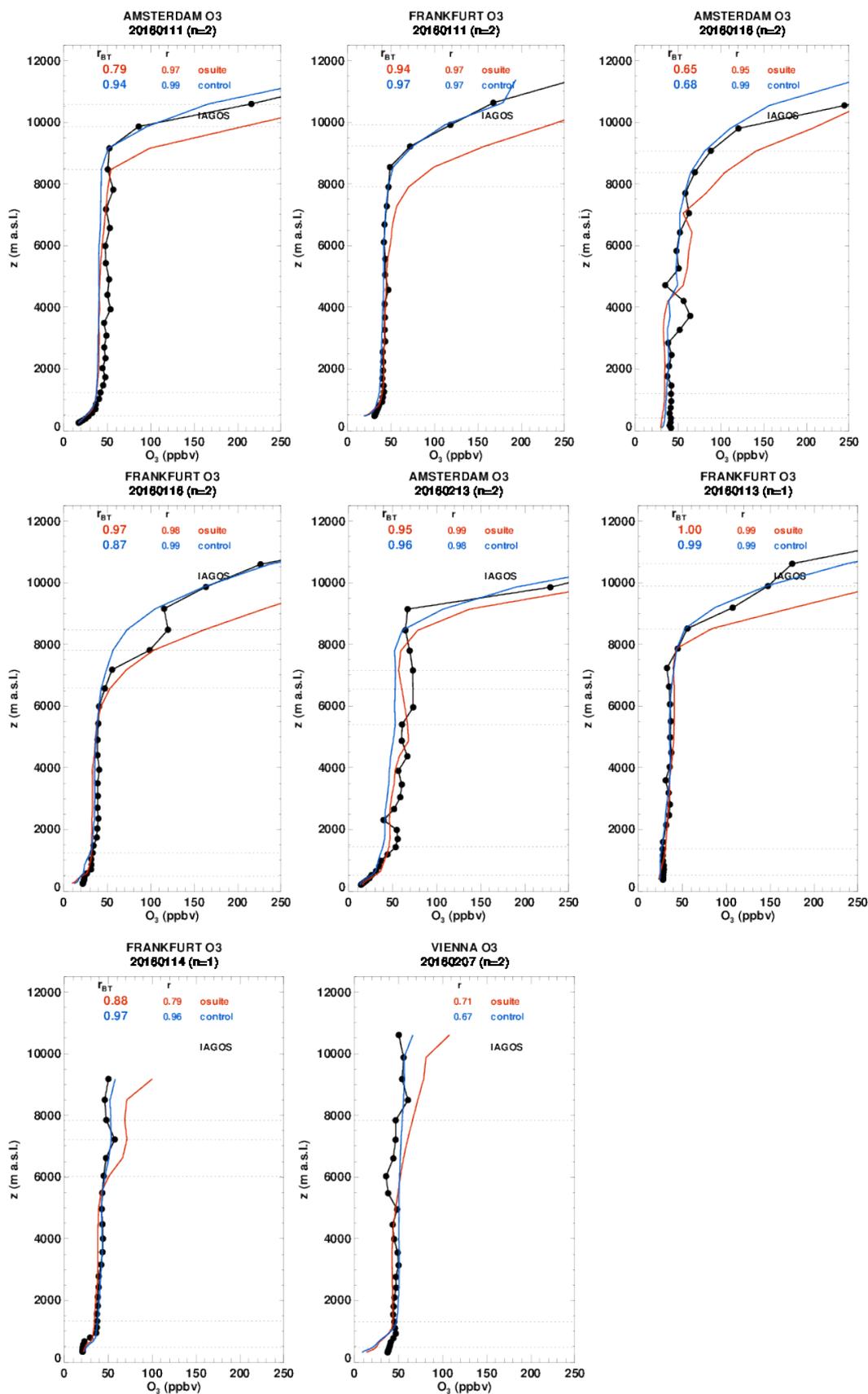


Figure 3.1.5: Selection of daily profiles of ozone from IAGOS (black) and the o-suite (red) and control (blue) over Europe (Frankfurt, Amsterdam, and Vienna) over the period December 2015-February 2016.

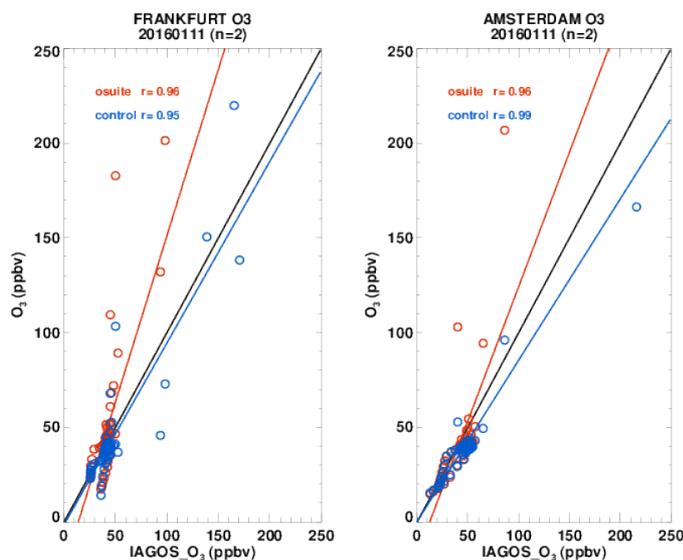


Figure 3.1.6: Scatter-plot osuite (red) and control(blue) against IAGOS showing all observations on the given day at Frankfurt and Amsterdam. Perfect correlation (IAGOS-IAGOS) is represented by the black line, with the best fit lines for osuite and control in red and blue respectively

on the 11th January, with the o-suite overestimating the amount of ozone. As the tropopause is a little lower this time, the o-suite does a better job at capturing the position of the tropopause.

In figure 3.1.6 we present scatter plots for Amsterdam and Frankfurt on the 11th January. The scatter plots represent all the data collected on the given day (i.e. the two flights are not averaged as they were for the profiles). The best-fit lines for the o-suite and control clearly show the difference between the two runs, with the o-suite overestimating ozone at high concentrations (near the tropopause) and the control underestimating ozone at high concentrations.

Asia

As the timeseries in Fig. 3.1.4 also showed, the profiles in Fig. 3.1.7 for Taipei illustrate that the o-suite slightly underestimates ozone in the surface layer and overestimates it in the upper troposphere, whilst in the free troposphere the runs compare very well with the observations. Behaviour at Hong Kong and Kuala Lumpur also follows this pattern. The scatter plots for Taipei and Hong Kong (Fig 3.1.8) also show that at the low ozone mixing ratios as found in the boundary layer, the o-suite has slightly underestimated the ozone.

Equatorial West-Africa

Thanks to the second Lufthansa aircraft equipped with IAGOS we have regular data over Equatorial and West Africa, especially over Nigeria (Lagos and Port-Harcourt). This region is characterized by an intense pollution from oil industries all year long, and is also under the influence of biomass burning over north-equatorial Africa from December to April. Ozone is elevated at altitudes around 2000m at Port Harcourt and Abuja and is underestimated by the two runs with the o-suite performing slightly better. In the free troposphere the two runs slightly underestimate the observed ozone.

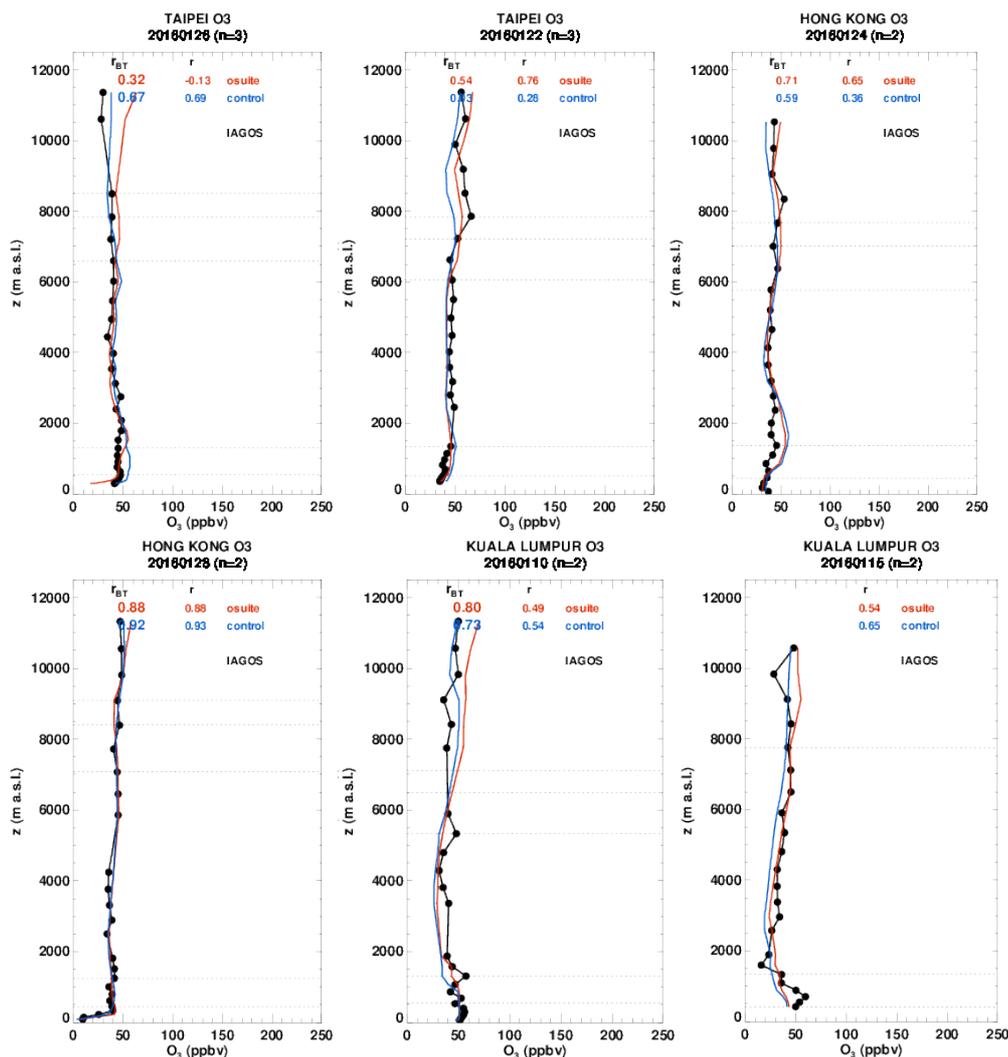


Figure 3.1.7: Selection of daily profiles of ozone from IAGOS (black) and the o-suite (red) and control (blue) over Asia (Taipei, Hong Kong, and Kuala Lumpur) over the period December 2015-February 2016.

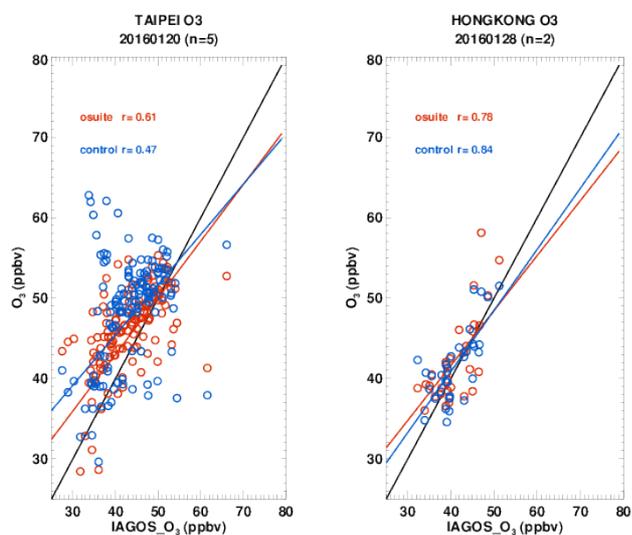


Figure 3.1.8: Scatter-plot o-suite (red) and control (blue) against IAGOS showing all observations on the given day at Taipei and Hong Kong. Perfect correlation (IAGOS-IAGOS) is represented by the black line, with the best fit lines for o-suite and control in red and blue respectively

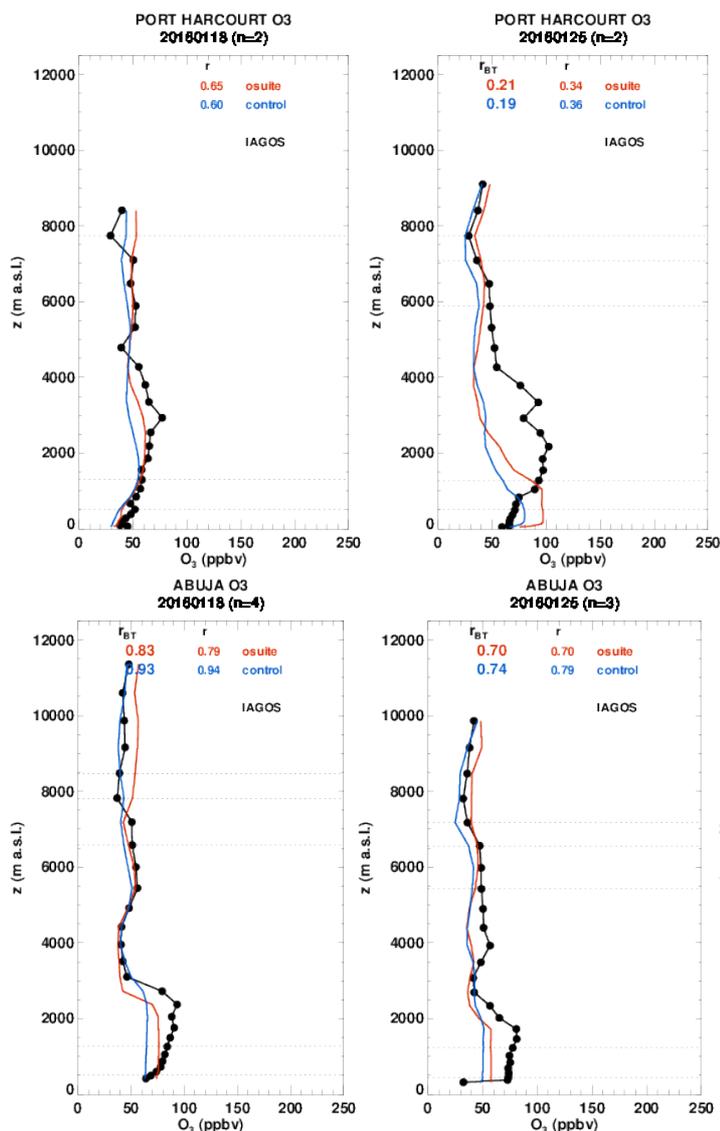


Figure 3.1.9: Profiles of ozone from IAGOS (black) and the two NRT runs over Port Harcourt and Abuja in January 2016.

Middle-East and East Africa

In previous seasons we have noted a quasi-systematic overestimation by the model throughout the troposphere. Here the models perform well over Jeddah and Kuwait in the free troposphere, Figure 3.1.10. At Addis Ababa (the airport is at 2300m above sea level) the models underestimate the increased ozone between the surface and 6000m but thereafter perform well.

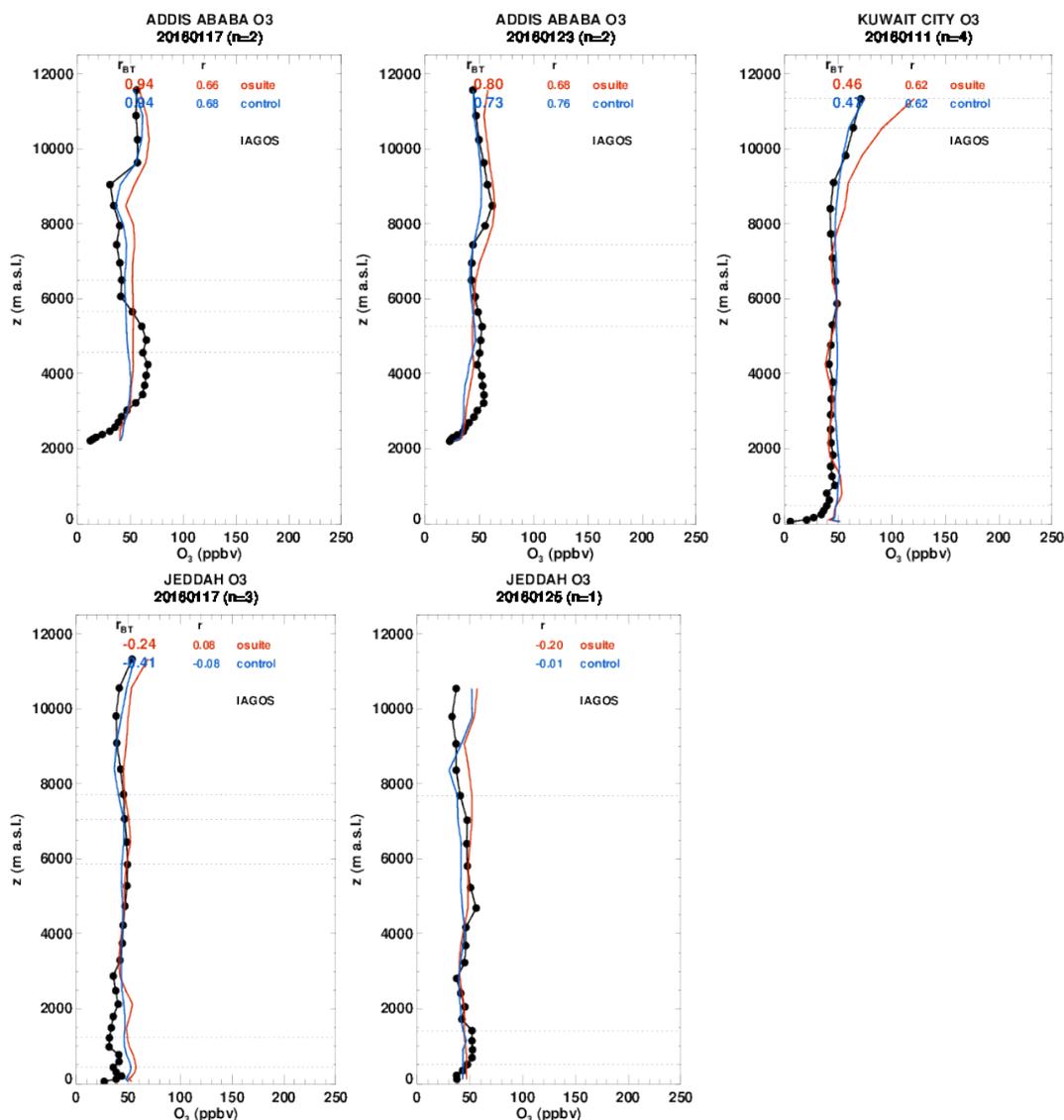


Figure 3.1.10: Profiles of ozone from IAGOS (black) and the two NRT runs over Doha, Dubai, Almaty, Kuwait City, Jeddah, and Muscat, over the period December 2015 – February 2016.

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 December 2015 to February 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 3.1.11 and 3.1.12.

Modified normalized mean biases in % (left, panel) and correlation coefficients (right, panel) between 3 different forecast days (D+0, circles; D+2, triangles and D+4, rhombs) and ESRL observations are shown in Fig. 3.1.12. It indicates that MNMBs for both o-suite and control run remain stable till the D+4 (forecast run from 96h to 120h)

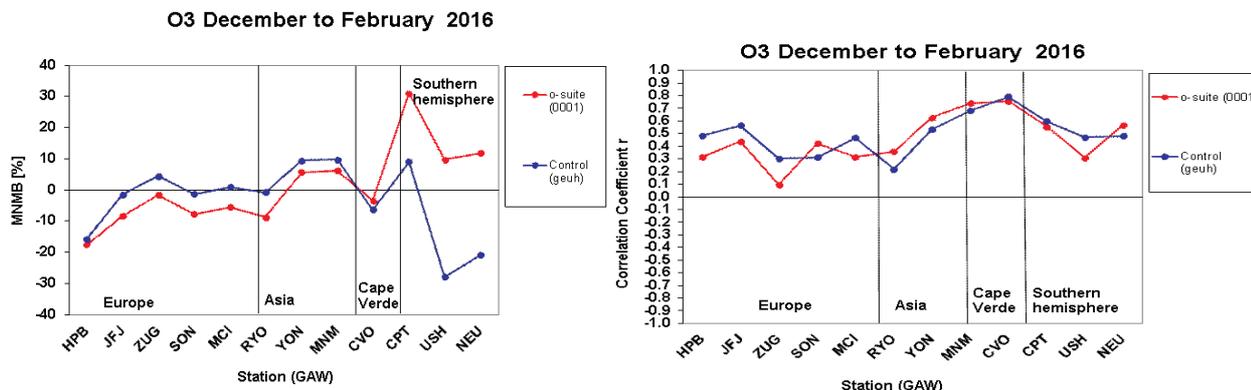


Figure 3.1.11: Modified normalized mean bias in % (left) and correlation coefficient (bottom right) of the NRT model runs compared to observational GAW data in the period December to February 2016.

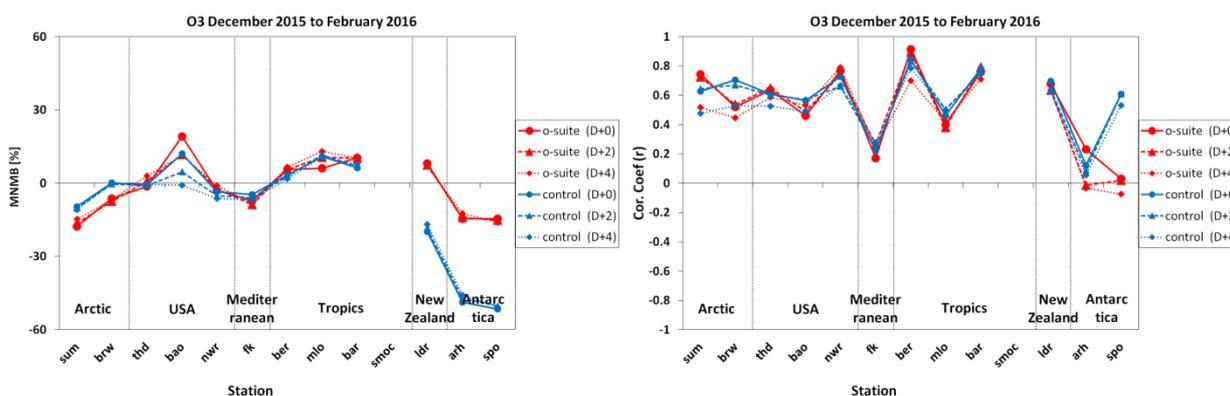


Figure 3.1.12: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT forecast runs compared to observational ESRL data in the period December 2015 to February 2016. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.

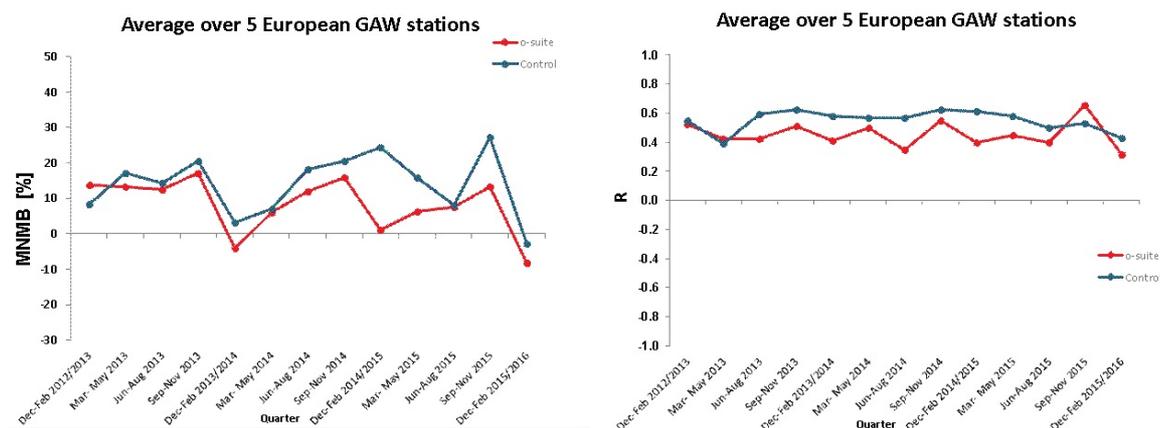


Figure 3.1.13: Long term (Dec. 2012 – February 2016) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

with the only exception for the Boulder, Colorado station (BAO), located in a mountainous area with complex topography which seems to affect the model stability. Correlations between simulated and observed surface ozone values remain almost stable till D+2 (forecast run from 48h to 72h), but then drop particularly in the polar regions (correlations for D+4 are lower than correlations for D+2 and D+0).

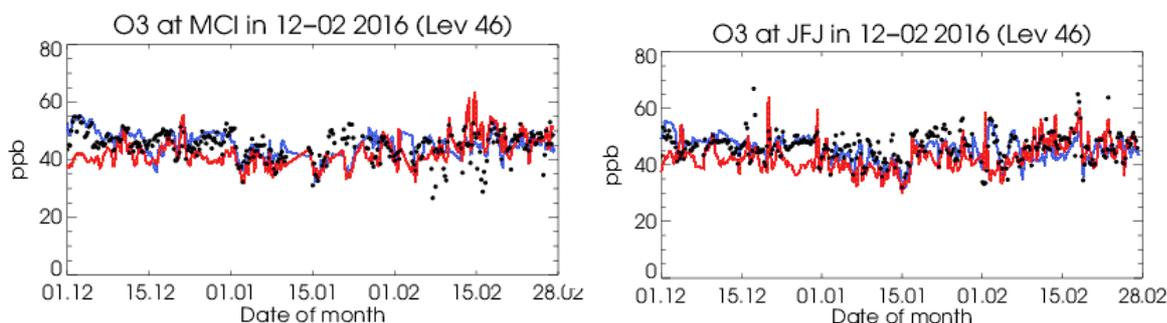


Fig. 3.1.13: Time series for the o-suite (red) and Control (blue) compared to GAW observations at Monte Cimone (44.2°N, 10.7°E) and Jungfraujoch (46.5°N, 7.9°E).

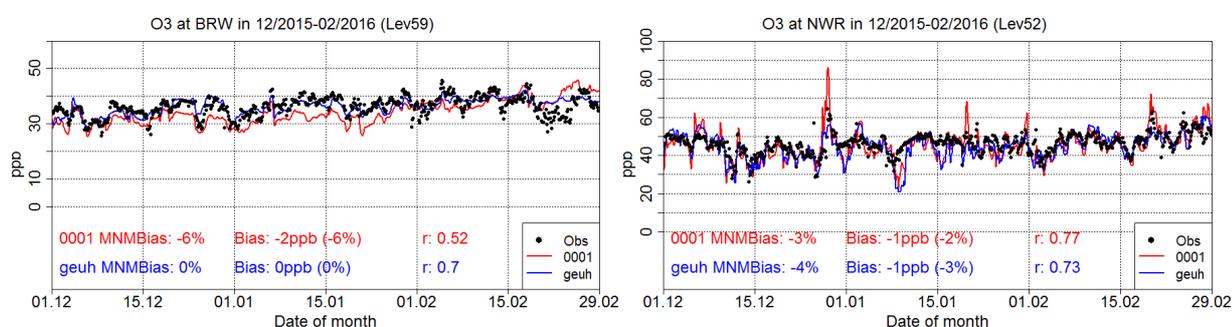


Figure 3.1.14: Time series for the o-suite (red) and control (blue) compared to ESRL observations at Point Barrow station (71.32°N, 156.61°W, left) and Niwot Ridge, Colorado station (40.04°N, 105.54°W, right).

A comparison of the seasonal-mean MNMB over Europe (Fig. 3.1.13) 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), observed O₃ surface mixing ratios are underestimated by the o-suite, with MNMBs between -2 to -18% for the o-suite (see also Figure 3.1.14). The control run shows lower MNMBs between 5 and -16%. Correlations for most stations in Europe are also lower for the o-suite (0.1-0.4) than for the Control run (0.3-0.6). In the time series plots larger deviations for the o-suite appear especially in December.

For stations in the Arctic (SUM and BRW) the o-suite underestimates surface ozone values by -17% (SUM) and -6% (BRW) respectively (see also Figure 3.1.15). The control run underestimates surface ozone values by -10% over SUM while over Point Barrow reproduces well surface ozone mean concentrations. It seems that at Arctic ESRL stations data assimilation reduces the model performance in terms of biases. Correlations between simulated and observed surface ozone at Summit and Point Barrow stations are high for both models runs ($r > 0.60$).

For USA stations both runs reproduce well surface ozone mean concentrations at THD and NWR (MNMBs $\approx 0\%$) and overestimate it at BAO (o-suite MNMBs $\approx 20\%$ and control MNMBs 12%). Correlations between simulated and observed surface ozone are high at NWR and at THD stations ($r > 0.6$) and lower at BAO (o-suite $r \approx 0.45$ and control

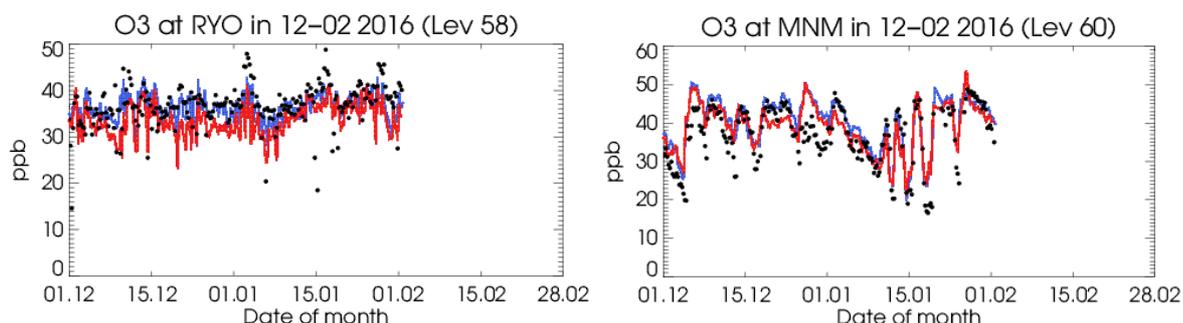


Figure 3.1.15: Time series for the o-suite (red) and Control (blue) compared to GAW observations at Ryori (39.0°N, 141.8°E) and Minamitorishima (24.3°N, 123.9°E).

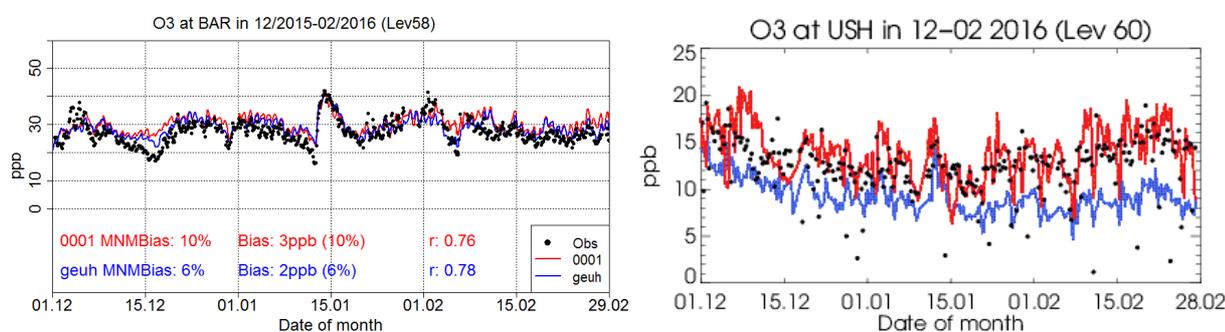


Figure 3.1.16: Time series for the o-suite (red) and control (blue) compared to ESRL observations (black dots) at Ragged Point, Barbados station (13.17°N, 59.46°W) and GAW observations at Ushuaia (54.9°S, 68.2°W).

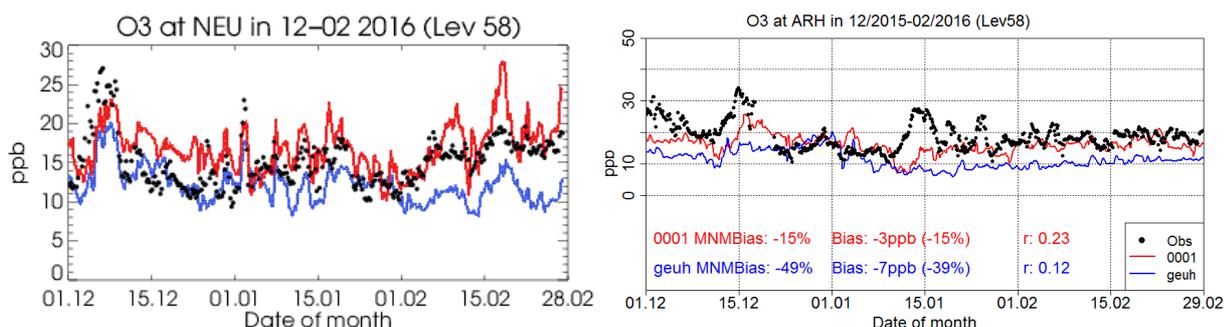


Figure 3.1.17: Time series compared to GAW observations (black dots) at Neumayer (70.7°S, 8.3°W) and ESRL observations at Arrival Heights, Antarctica (-77.80°N, 166.78°W).

$r \approx 0.55$). It should be noted that data assimilation seems to reduce the model performs over BAO station.

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 of around 6% for the o-suite and around 10% 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 3.1.16. For RYO station, the o-suite shows a negative offset with MNMB of -9% whereas the control run could reproduce background concentrations, however, but does not show the same linearity as the observations, which results in low correlation



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 DJF 2015-2016) 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	-20.9	-16.6	0.17	0.23
Caravaka	ES1882A	-1.87	38.12	1	60	73	-47.7	-44.3	0.56	0.55
Zarra	ES0012R	-1.10	39.08	885	55	70	-14.5	-11.2	0.27	0.17
Viillar Del Arzobispo	ES1671A	-0.83	39.71	430	60	48	-48.7	-53.3	0.80	0.80
Cirat	ES1689A	-0.47	40.05	466	60	37	-38.5	-39.1	0.87	0.83
Bujaraloz	ES1400A	-0.15	41.51	327	60	60	-51.4	-42.2	0.83	0.82
Morella	ES1441A	-0.09	40.64	1150	52	51	-9.4	-11.6	0.57	0.52
Bc-La Senia	ES1754A	0.29	40.64	428	60	21	-81.3	-74.0	0.75	0.71
Ay-Gandesa	ES1379A	0.44	41.06	368	60	15	-76.2	-75.9	0.80	0.81
Ak-Pardines	ES1310A	2.21	42.31	1226	51	81	19.8	22.7	0.40	0.39
Hospital Joan March	ES1827A	2.69	39.68	172	56	3	-18.6	-10.6	0.65	0.57
Al-Agullana	ES1201A	2.84	42.39	214	60	25	-57.8	-79.4	0.86	0.85
Av-Begur	ES1311A	3.21	41.96	200	58	9	-41.2	-38.3	0.70	0.67
Plan Aups/Ste Baume	FR03027	5.73	43.34	675	54	21	-15.7	-4.5	0.45	0.26
Finokalia	GR0002R	25.67	35.32	250	55	4	-7.1	-4.6	0.17	0.22
Ineia	-	32.37	34.96	672	52	5	-5.8	-4.7	0.72	0.73
Oros Troodos	-	32.86	34.95	1819	47	11	-1.8	1.6	0.46	0.63
Agia Marina	CY0002R	33.06	35.04	532	53	14	-0.9	-0.2	0.59	0.67

(0.2) for this station. The o-suite shows higher correlation for all Asian stations, with R between 0.4 and 0.7.

Over the tropical stations (BER, MLO, BAR) both runs slightly overestimate ozone mixing ratios, less than 10%, see Figure 3.1.17(left). Correlations are high at BER, and BAR ($r > 0.7$) and lower at MLO ($r > 0.4$).

For the three stations in the Southern Hemisphere, the o-suite reproduces ozone mixing ratios with exception of a few periods in December. The data assimilation mostly corrects the negative offset in the control run, see Figure 3.1.17(right panel). Correlation coefficients are between 0.3 and 0.6 for the o-suite and between 0.5 and 0.6 for the control run. Also at Antarctica stations (NEU, ARH and SPO) both the o-suite and the control underestimate surface ozone mean concentrations (by -15% and -50% respectively, see Figure 3.1.18. Data assimilation clearly improves the model performance in terms of biases. Correlations between simulated and observed surface ozone values are low for both o-suite and control ($r > 0.20$ for o-suite) at ARH. At SPO correlations between simulated and observed surface ozone values are good for the control run ($r = 0.6$) but absent for o-suite ($r \approx 0.0$).

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 this, 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

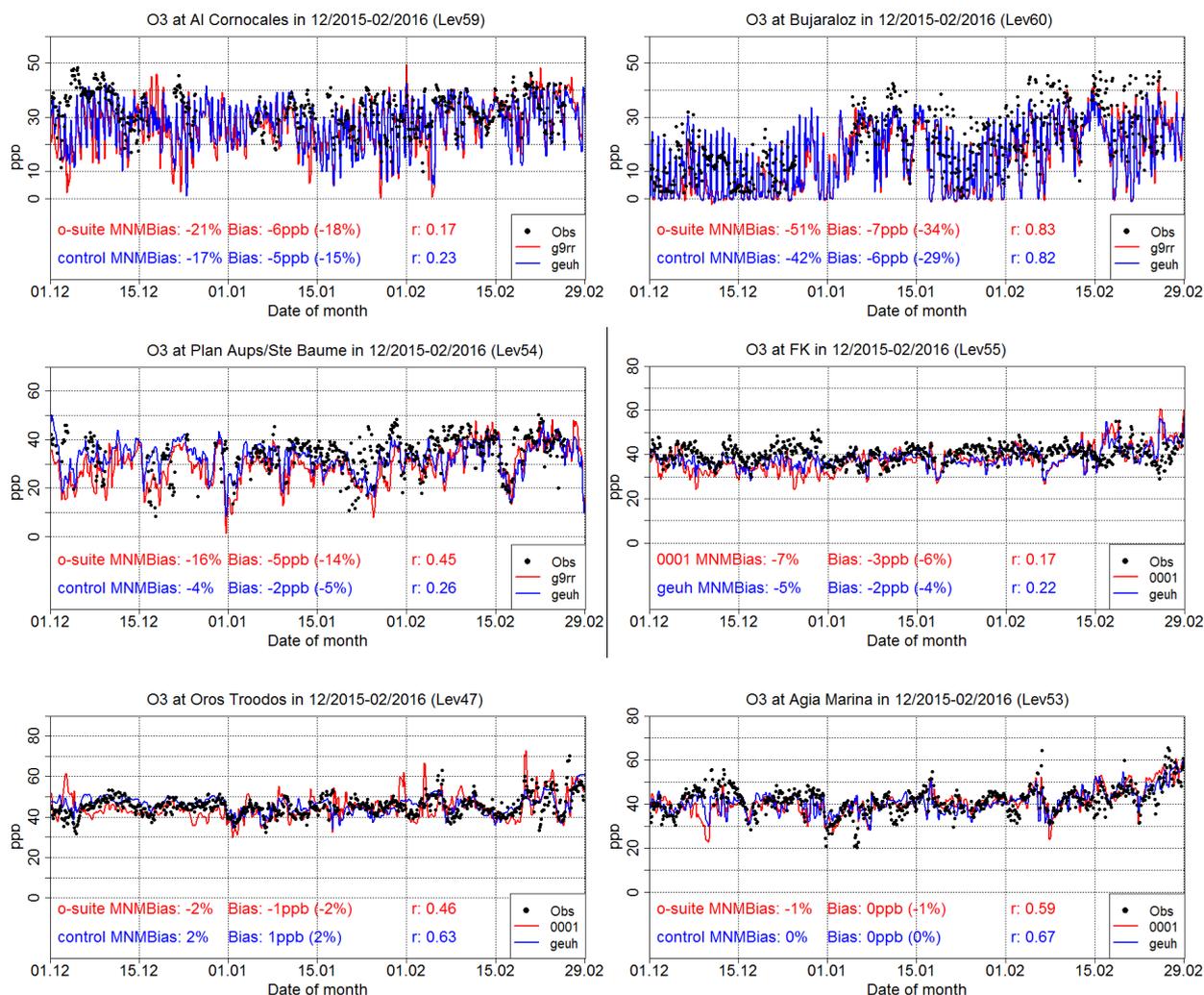


Figure 3.1.18: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Cornocales, Spain station (36.23°N, 5.66 °W, top left), at Bujaraloz, Spain station (41.51°N, 0.15°W, top right), at Plan Aups/Ste Baume, France station (43.34°N, 5.73°E, center left) and Finokalia, Crete station (35.32°N, 25.67°E, center right), and compared to observations provided by the Department of Labour Inspection - Ministry of Labour and Social Insurance of Cyprus, at Troodos Mountain station (34.95°N, 32.86 °E, low left), and Agia Marina station (35.04°N, 33.06°E, low right).

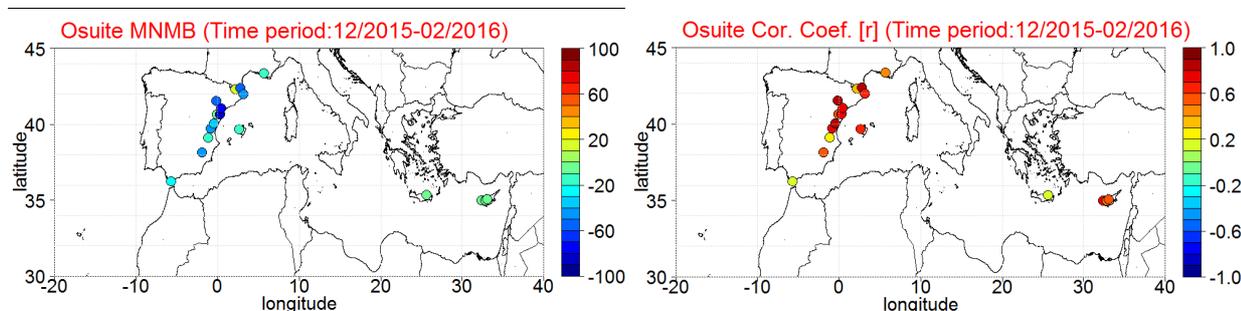


Figure 3.1.19: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 December 2015 to 29 February 2016.

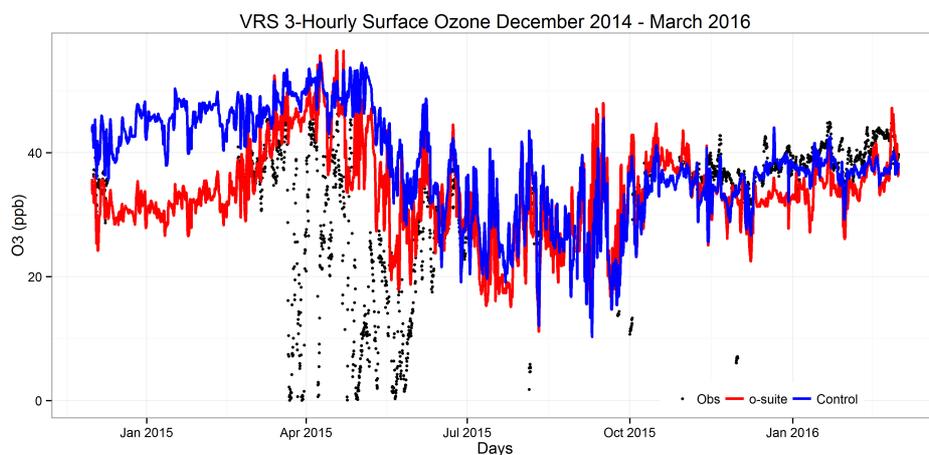


Figure 3.1.20: Time series for o-suite (red) and Control (blue) compared to observations (black dots) at the Villum Research Station, Station Nord, Greenland.

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 Al Cornocales and Zarra in Spain and Finokalia station in Crete. Over Spain and France the o-suite mostly reproduces better the day to day variability than the control run. In contrast, in the eastern Mediterranean the control run is mostly equal or better than the o-suite.

In terms of biases, both model runs mostly underestimate surface ozone mean concentrations over Spain and France. Table 3.1.1 and Figure 3.1.19, top right, indicate that over Spain the MNMBs for both model versions exceed -50% in the case that the elevation of station corresponds to the surface model level. Both o-suite and control runs underestimate slightly surface ozone mean concentrations at Finokalia station in Crete, Greece (MNMBs \approx -5%). Finally both runs reproduce very well surface ozone mean concentrations at the Ineia, Troodos mountain and Agia Marina stations in Cyprus, (MNMBs \approx 0; see also Figure 3.1.18, lower graphs).

The spatial distribution of MNMBs and correlations of the o-suite over the Mediterranean is shown in Figure 3.1.19, where the contrast in the model MNMBs between Mediterranean shore of Spain (strong negative signal) 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.

3.1.5 Validation with IASOA surface observations

O₃ observations from the Villum Research Station, Station Nord in north Greenland from the IASOA network were compared to model results, Figure 3.1.20. There are large gaps in the measurement time series covering the period from December 2014 to February 2016. Ozone depletion events in March – June 2015 are not captured by the model simulations during spring. 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 December 2015 – February 2016 the measurements are not quality controlled. The model simulations are on average in good agreement with the observations although they underestimate the observed concentrations, resulting in a monthly negative normalized mean bias between -10 and -13% for the o-suite and between -1 and -9% for the control run, respectively, for the period. The short-term variability is well captured with a correlation coefficient of $r = 0.37-0.59$ for the o-suite and $r = 0.12-0.69$ for the control run for the three months.

3.2 Tropospheric nitrogen dioxide

3.2.1 Evaluation against GOME-2 retrievals

In this section, model columns of tropospheric NO_2 are compared to SCIAMACHY/Envisat NO_2 satellite retrievals (IUP-UB v0.7, Richter et al., 2005) for model data before April 2012, and to GOME-2/MetOp-A NO_2 satellite retrievals (IUP-UB v1.0, Richter et al., 2011) for more recent forecasts, as described in Eskes et al., (2016). Satellite overpass is roughly at 10:00 local time (LT) for SCIAMACHY and 09:30 LT for GOME-2, and only clear sky data is used. Model data are vertically integrated, interpolated in time and then sampled to match the satellite data. As a rough estimate, systematic uncertainties in regions with significant pollution are of the order of 20% – 30%.

Fig. 3.2.1 shows global maps of GOME-2 and model monthly mean tropospheric NO_2 columns for January 2016 together with differences between retrievals and simulations. The overall spatial distribution and magnitude of tropospheric NO_2 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 a bit better here. Compared to satellite data, all model runs underestimate tropospheric background values over Africa, South America and Australia, while background values over North America are overestimated. Local maxima of values observed over anthropogenic emission hotspots in Central Europe are strongly underestimated by the models, while they are overestimated over East Asia (e.g. over the heavily populated Sichuan Basin; 30°N , 105°E), parts of India and North America. Moreover, both runs overestimate fire emissions over Central Africa. There is visually no difference between the o-suite simulations for forecast day 0, 2 and 4, which can be explained by the fact that NO_2 has a short lifetime in the troposphere, so that data assimilation does not have a strong impact on model results (independent on analysis or forecast).

Closer inspection of the seasonal variation of tropospheric NO_2 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 NO_2 column retrievals decreased since 2014, model simulated values are since then in better agreement with the satellite retrieved ones. However,

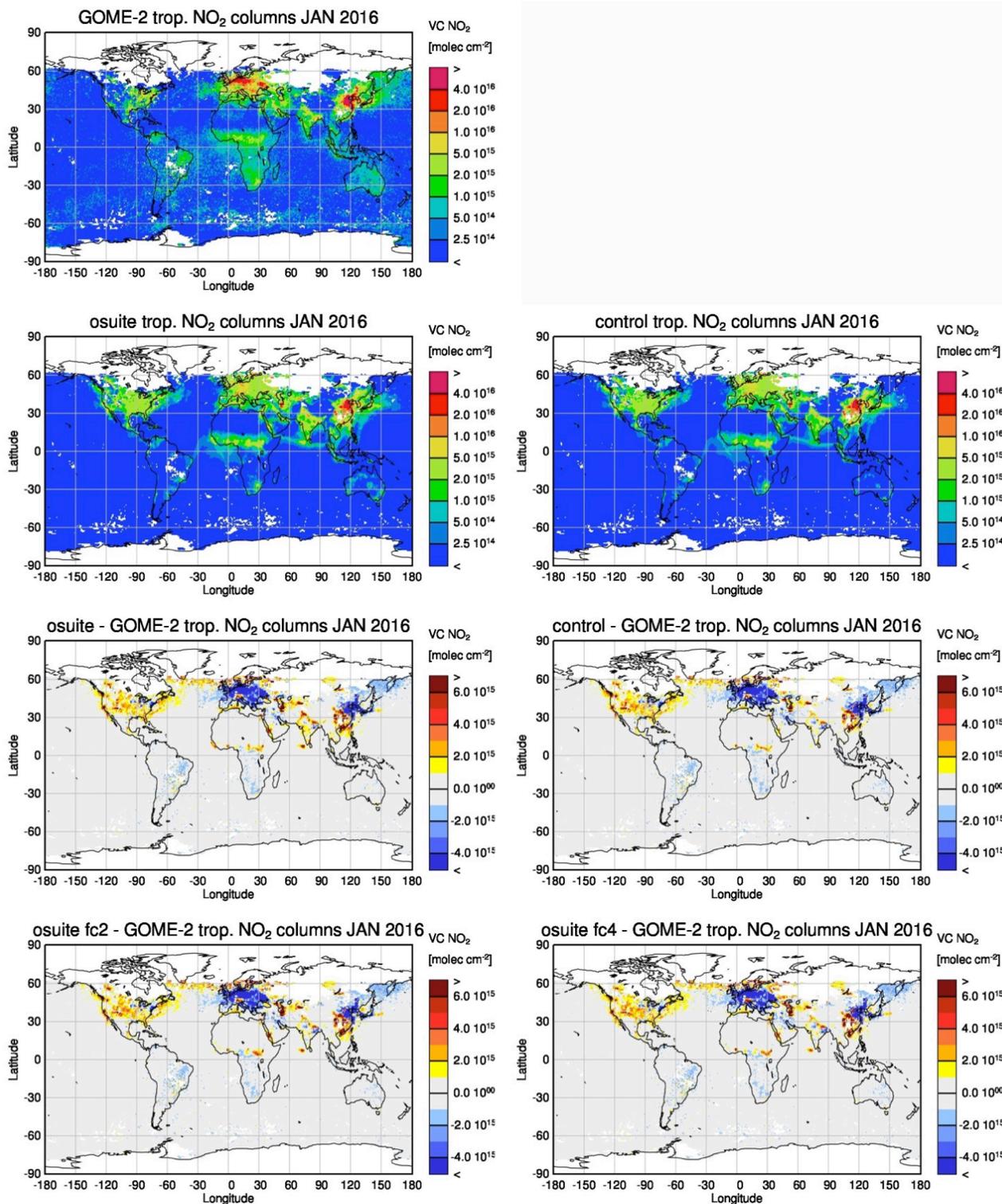


Figure 3.2.1: Global map comparisons of satellite retrieved and model simulated tropospheric NO₂ columns [molec cm⁻²] for January 2016. The top row shows monthly mean tropospheric NO₂ columns retrieved by GOME-2, the second row shows the same but for model simulated averages. The third row shows differences of monthly means from models and GOME-2. Images at the bottom show corresponding differences to o-suite results but for the (left) second day of forecast and (right) fourth day of forecast. GOME-2 data are gridded to model resolution (i.e. 0.75° deg x 0.75° deg). Model data are treated with the same reference sector subtraction approach as the satellite data.

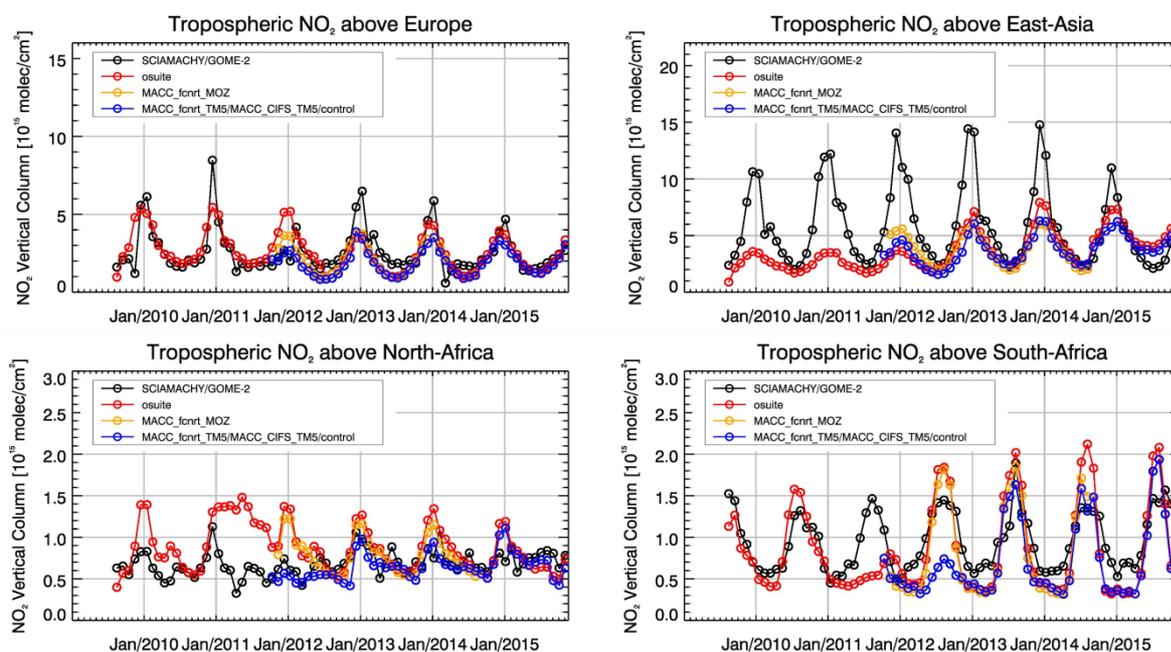


Figure 3.2.2: Time series of average tropospheric NO₂ columns [10^{15} molec cm⁻²] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions (see Annex 2 for definition of regions). Upper panels represent regions dominated by anthropogenic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

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 spring/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 models. Summertime NO₂ columns over North-Africa are underestimated compared to the satellite data for 2015. The o-suite overestimates the seasonal cycle for South-Africa for the years 2014/2015 with an overestimate 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.

As for the global maps displayed by the previous Figure, there is visually no difference between the o-suite simulations for forecast day 0, 2 and 4 as NO₂ has a short lifetime in the troposphere.

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 model versions with UVVIS DOAS measurements at Xianghe (a station near Beijing), Haute Provence (a rural station) and Uccle (near Bruxelles). This ground-based, remote-sensing instrument is sensitive to the NO₂ abundance in the lower troposphere, up to 1km altitude with an estimated uncertainty of 8%. Tropospheric NO₂ profiles and columns are validated (up to 3.5km). Smoothing using the averaging kernel of the observations has been applied to allow for a fair comparison to the observations, see also Eskes et al. (2016). We should mention that the measurement data is still catalogued as rapid delivery and not in the consolidated NDACC database.

From Figure 3.2.3 and the table 3.2.1 we see the assimilation has a negative effect at the rural site of OHP. The positive effect is observed at the more polluted sites 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).

		MAM			JJA			SON			DJF		
		MB	stddev	nobs									
osuite	ohp	4.99	62.44	239	11.68	56.81	228	-5.72	40.84	223	35.13	96.22	115
control	ohp	-10.02	45.56	239	6.61	51.59	228	-12.48	59.38	223	21.16	84.90	115
osuite	uccle	-27.17	32.14	155	-22.40	49.74	70	-20.35	34.65	53	-5.70	47.18	44
control	uccle	-39.32	29.98	155	-41.93	32.34	70	-31.68	25.87	53	-20.89	56.31	44
osuite	xianghe	17.34	59.89	311	78.32	93.34	220	91.40	141.8	64	11.73	77.38	103
control	xianghe	7.34	51.96	311	65.35	92.65	220	69.98	104.1	64	24.63	105.4	103

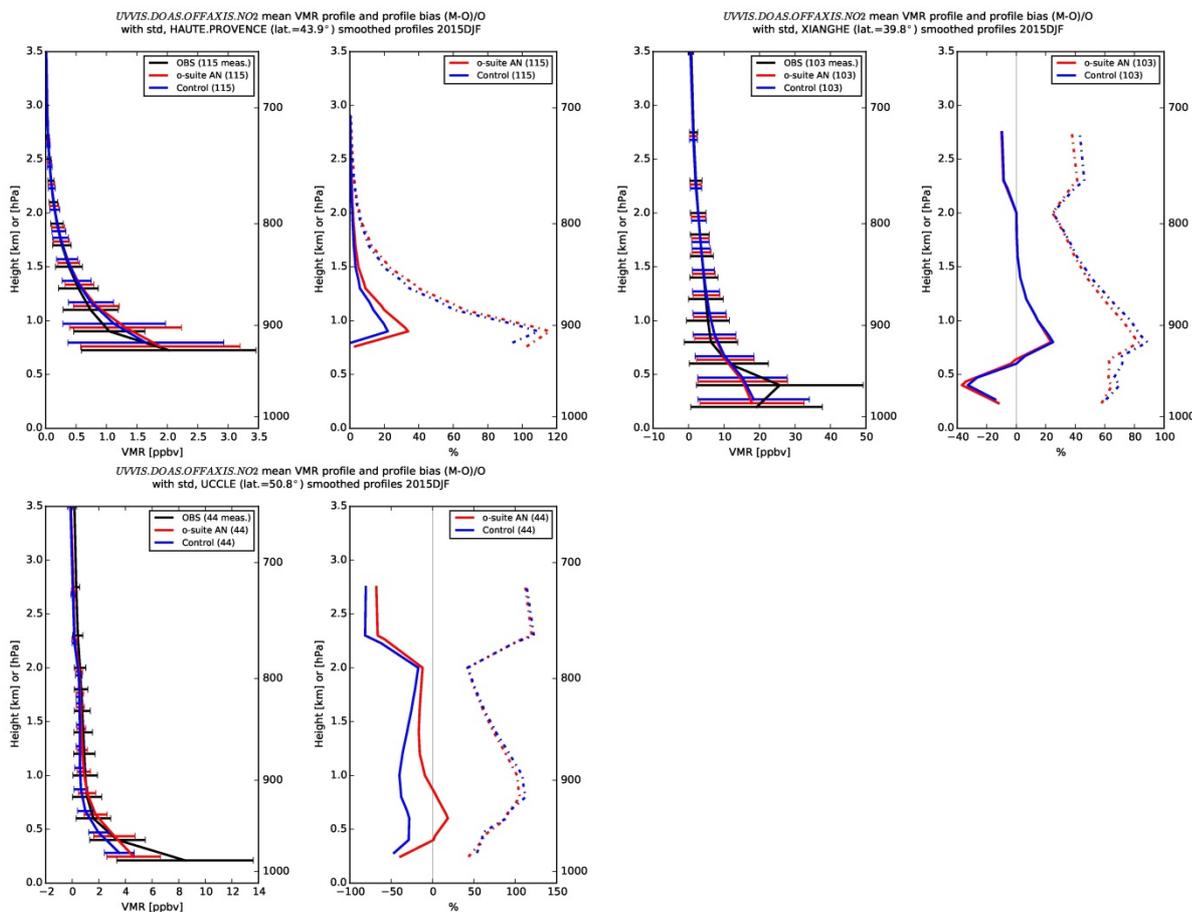


Figure 3.2.3: Seasonal mean tropospheric NO₂ profiles by o-suite (red) and control (blue) compared to NDACC UVVIS DOAS data at Haute Provence (43.9°N, 5.71°E, left top), Xianghe (39.8°N, 117°E, right) and Uccle (50.8°N, 4.36°E, bottom) for Dec-Jan-Feb 2016.

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 December - February 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 than the o-suite, which is an improvement over Europe but a strong degradation over the southern hemisphere (Figure 3.3.1). 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, the o-suite shows an underestimation of observed CO mixing ratios, with MNMBs between -8% and -18%. The control run shows smaller MNMBs between -3% and 6%. Data assimilation seems to cause a negative offset for surface

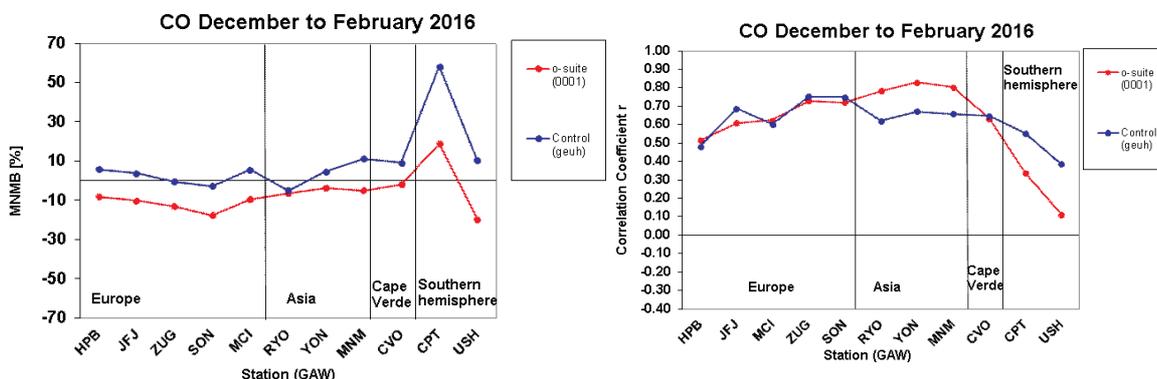


Figure 3.3.1: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT model runs compared to observational GAW data in the period December to February 2016.

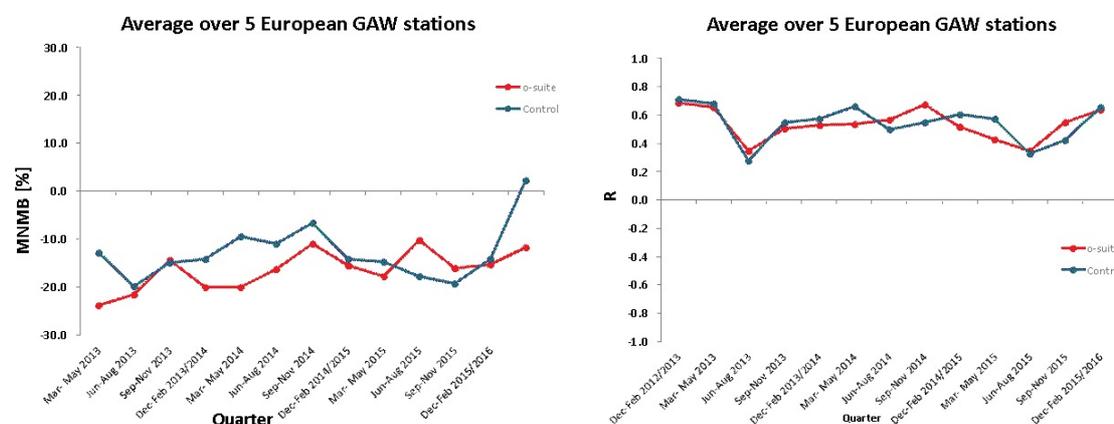


Figure 3.3.2: Long term (Dec. 2012 – February 2016) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

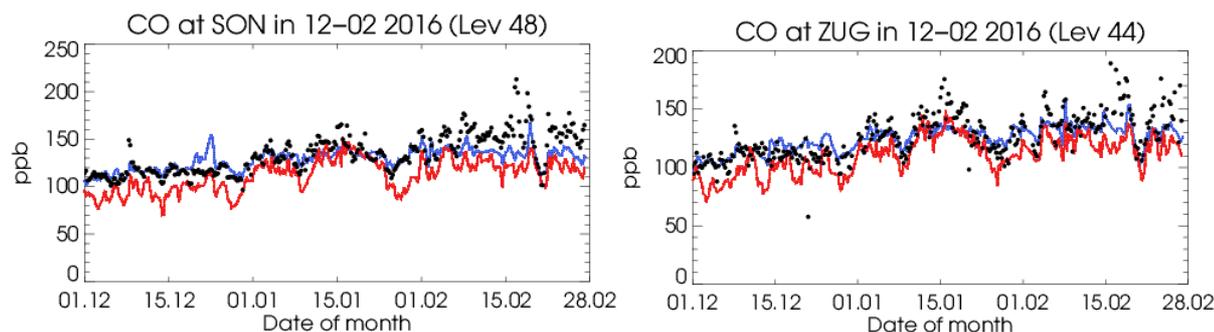


Figure 3.3.3: Time series for the o-suite (red) and control (blue) compared to GAW observations at Sonnblick (47.0°N, 12.9°E) and Zugspitze (47.1°N, 12.9°E).

CO in the o-suite for European stations. Correlation coefficients are slightly higher for the control run and lie between 0.5 and 0.8.

For Asian stations, both runs correspond well to the observations. The control run shows a positive offset for YON and MNM, whereas the o-suite has a slight negative offset, see Figure 3.3.4. Correlation coefficients are higher for the o-suite (around 0.8) than for the control run (0.6-0.7). Concentration peaks are resolved well by both runs.

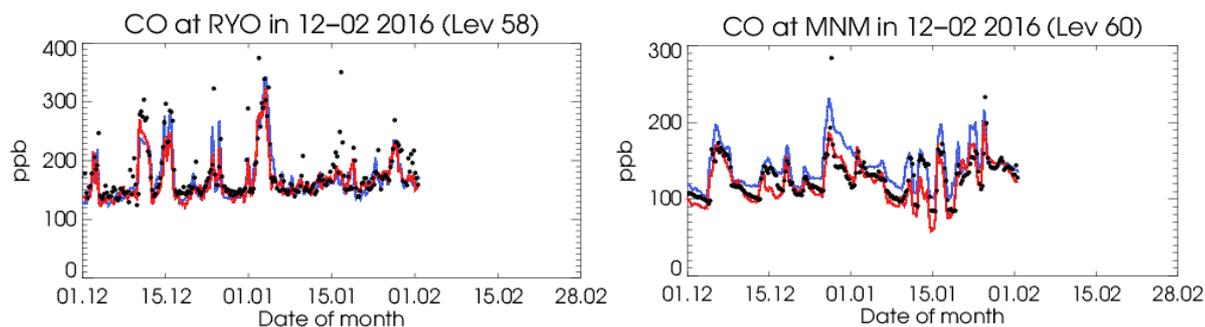


Figure 3.3.4: Time series for the o-suite (red) and control (blue) compared to GAW observations at Ryori (39.0°N, 141.8°E) and Minamitorishima (24.3°N, 123.9°E).

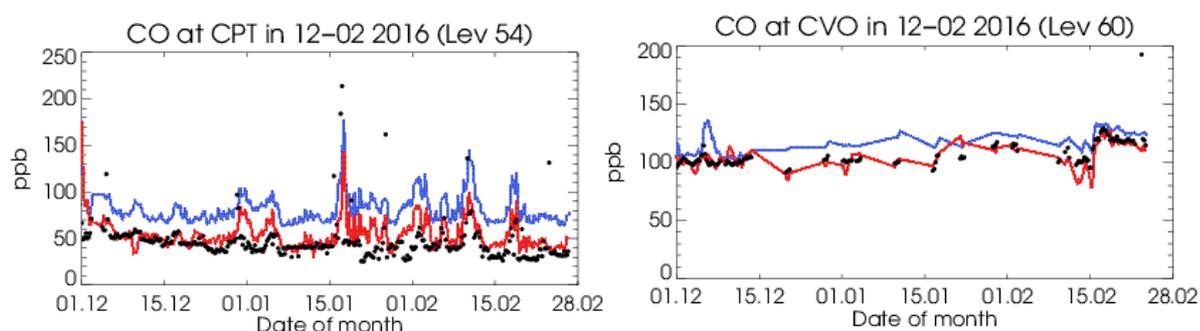


Figure 3.3.5: Time series for the o-suite (red) and control (blue) compared to GAW observations at Cape Point (34.35°S, 18.5°E) and Cape Verde (16.9°N, 24.9°W).

For the two stations in the Southern Hemisphere and for Cape Verde Observatory station (CVO), the data assimilation in the o-suite corrects the positive offset in CO mixing ratios visible for the control run, 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 December 2015–January 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 Taipei and Frankfurt 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).

Europe

Figure 3.3.7 gives examples of the CO profiles over Frankfurt, Amsterdam and Vienna. 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 region. The scatter-plot of Figure 3.3.8 also shows clearly that the model underestimates the highest CO mixing ratios (found in the boundary layer), and overestimates the lowest

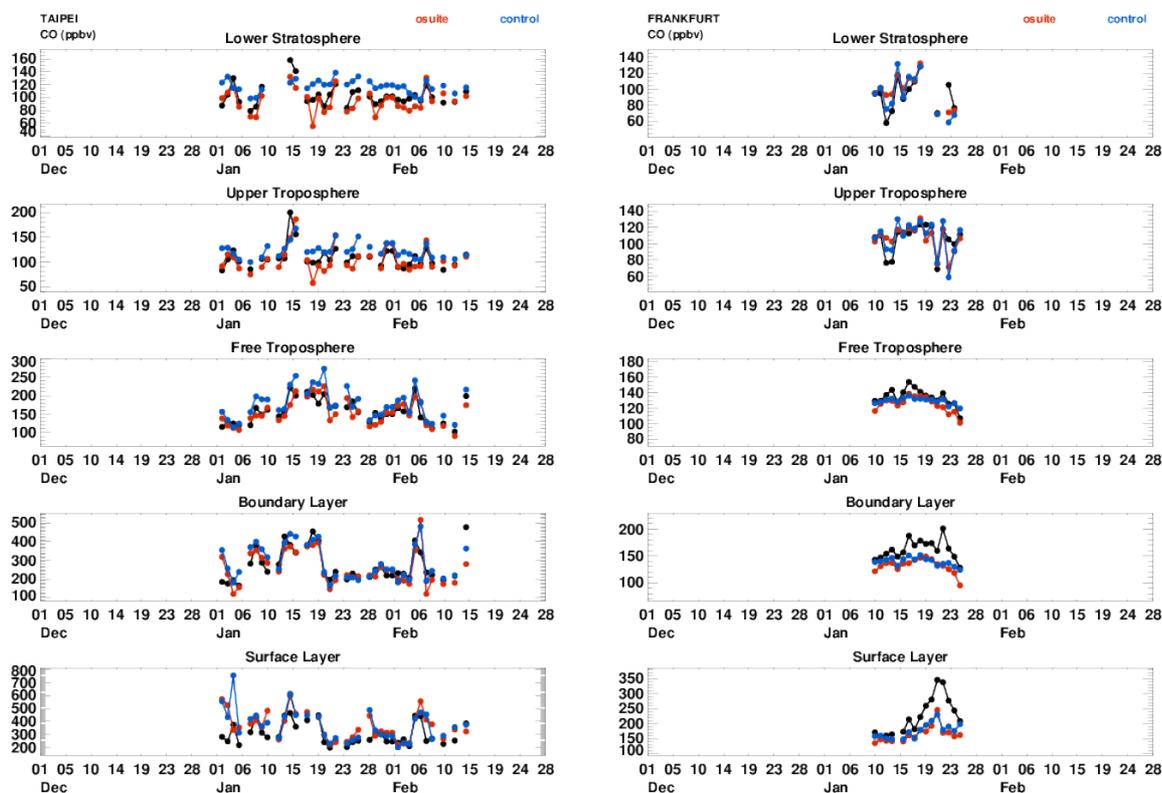


Figure 3.3.6: Time series of daily mean CO Taipei and Frankfurt from December 2016 to February 2015 for 5 layers, Surface, Boundary layer, Free Troposphere, Upper Troposphere and Lower Stratosphere.

CO mixing ratios (as found near the tropopause) with very little difference between the control run and the o-suite.

Asia

The time series at Taipei (Figure 3.3.6) showed that the CO from the model versions showed good correspondence to the observations. 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 Ningbo, Nanjing, Xiamen, Taipei and Hong Kong the profile of CO is generally well captured with the o-suite making an improvement over the control run throughout the free troposphere where the control run often overestimates the amount CO (see also scatter plot 3.3.9). Frequently there is a maximum in CO at around 2000m which is overestimated by the control run and improved in the o-suite after assimilation. At Manila, Bangkok, Ho Chi Minh City and Kuala Lumpur, the free troposphere is again well represented. If there is a maximum in the boundary layer as we see at Manila 14th February, Bangkok 22nd January, Ho Chi Minh City (7th January) and Kuala Lumpur on 10th January, this is usually missed by the models. However, if this maximum is not present the models do a good job at capturing the high concentrations in the boundary layer. The maxima seen at these locations around South East Asia and particularly at Ho Chi Minh City are the result of burning which was particularly intense across the Indochinese peninsula during the first week in January as seen by MODIS.

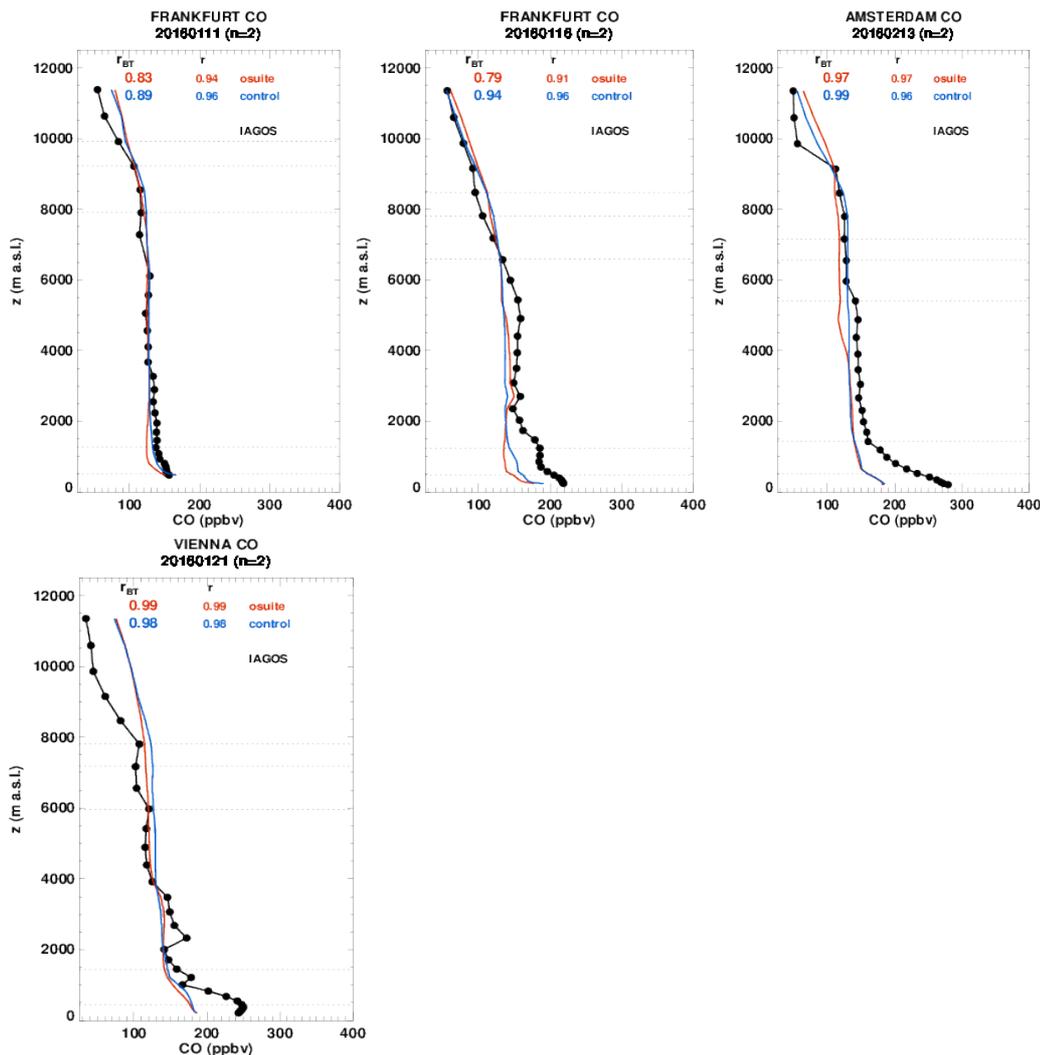


Figure 3.3.7: Selection of profiles of CO from IAGOS (black) and the two NRT runs over Europe in December 2015-January 2016.

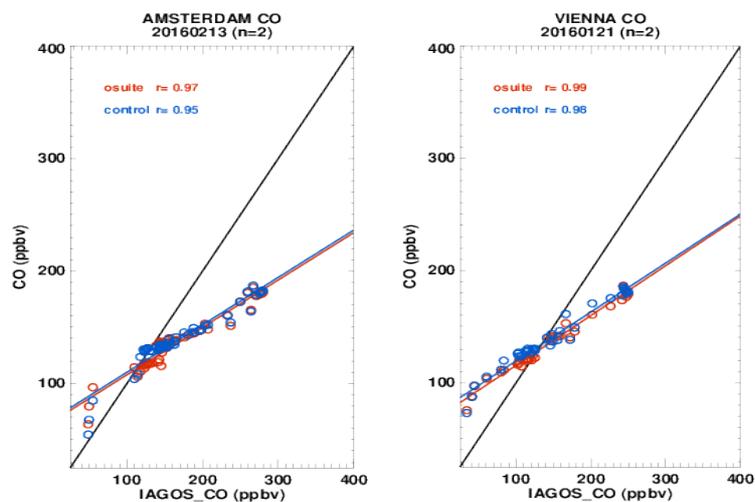


Figure 3.3.8: Scatter-plot o-suite (red) and control(blue) against IAGOS showing all observations on the given day at Amsterdam and Vienna. Perfect correlation (IAGOS-IAGOS) is represented by the black line, with the best fit lines for o-suite and control in red and blue respectively

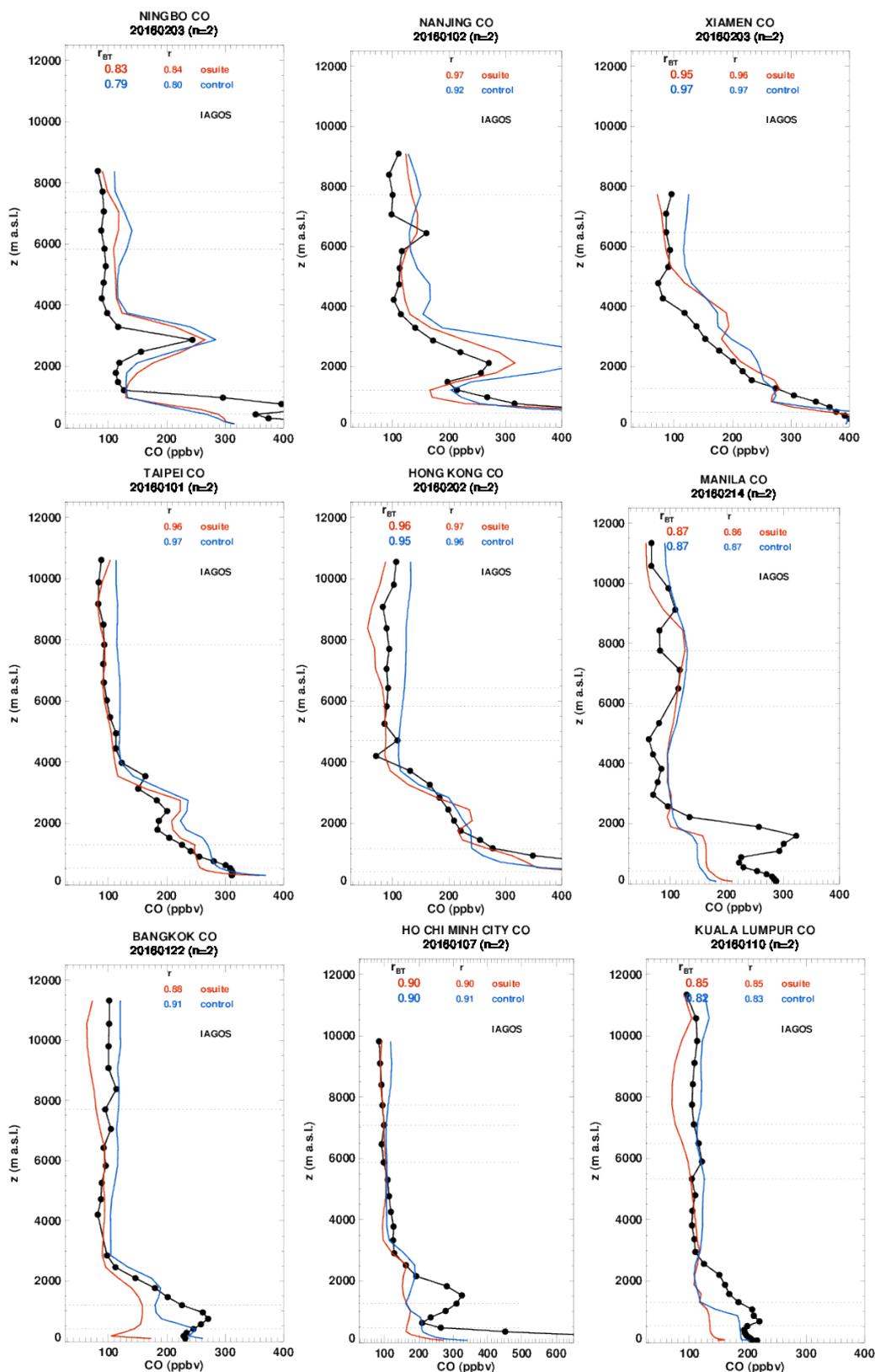


Figure 3.3.9: Profiles of CO from IAGOS (black) and the two NRT runs in China and South East Asia during the period December 2015-February 2016.

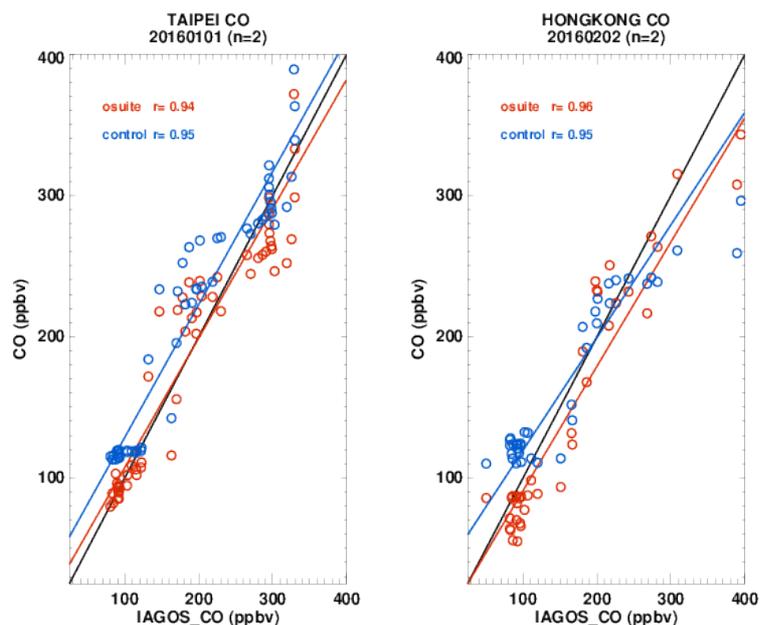


Figure 3.3.10: Scatter-plot osuite (red) and control(blue) against IAGOS showing all observations on the given day at Taipei and Hong Kong. Perfect correlation (IAGOS-IAGOS) is represented by the black line, with the best fit lines for osuite and control in red and blue respectively

Equatorial West-Africa

Figure 3.3.11 highlights some examples of CO profiles over Equatorial Africa as regularly sampled by the second Lufthansa IAGOS equipped aircraft since March 2015. In DJF, we see a layer of enhanced CO between the surface and about 4000m. This is due to the biomass burning across the northern part of Equatorial Africa which peaks during this season. Sometimes the models do well at capturing this increase in CO such as at Malabo or Lagos on 10th January, but the profiles at Port Harcourt and Abuja show that the models underestimate the amount of CO.

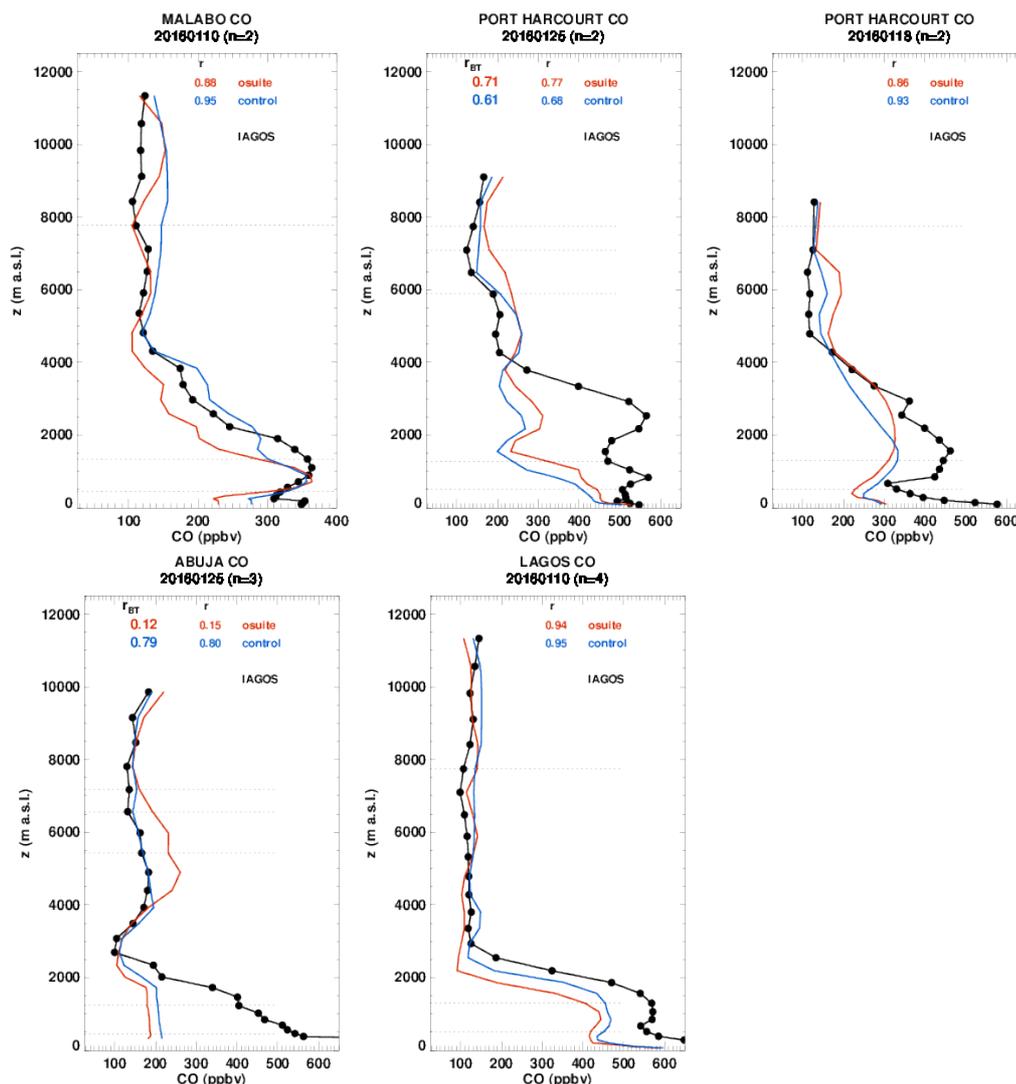


Figure 3.3.11: Profiles of CO from IAGOS (black) and the two NRT runs over West Africa during the period December 2015-February 2016

Middle East

CO is well captured by the two runs throughout the troposphere, particularly at Kuwait and Doha. We see an underestimation of CO in the boundary and surface layers as noted in most regions (Figure 3.3.12).

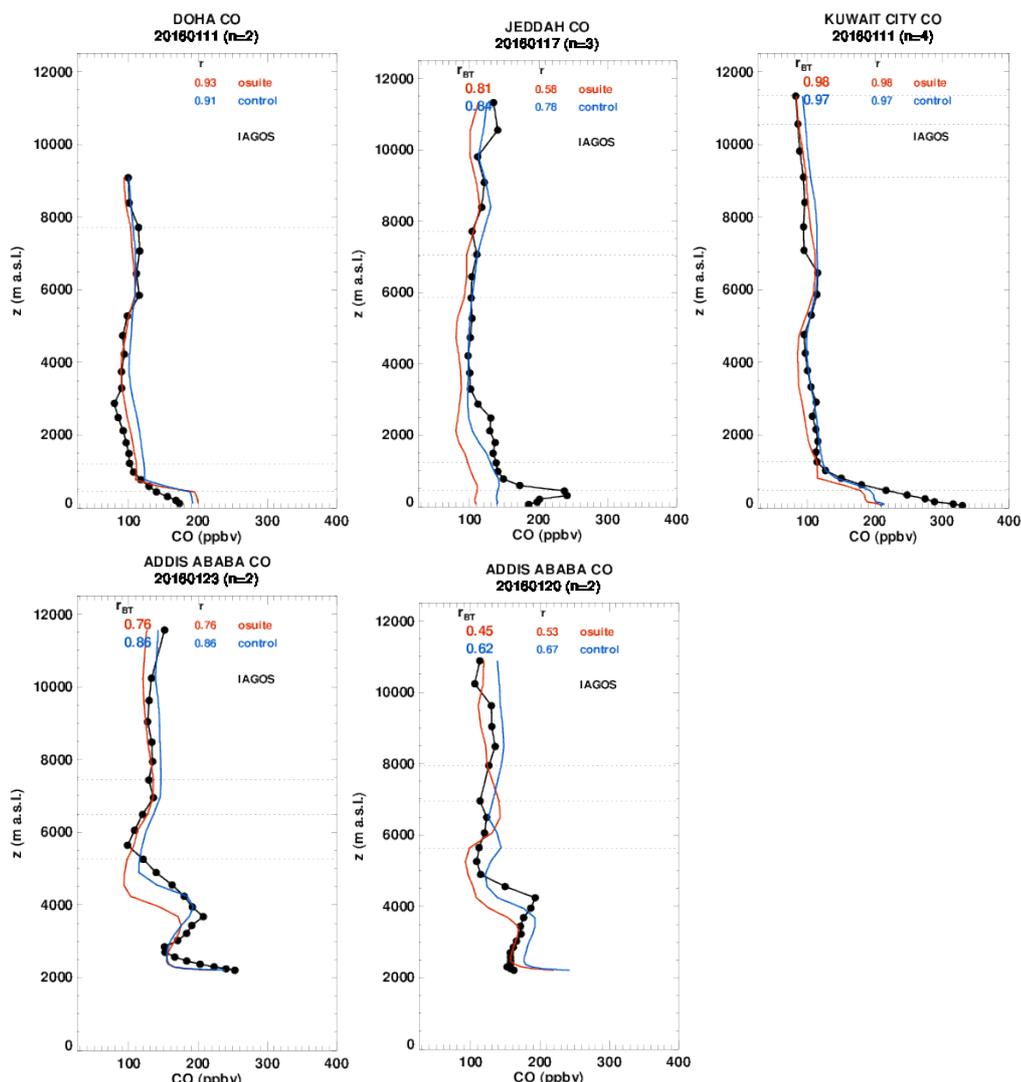


Figure 3.3.12: Profiles of CO from IAGOS (black) and the two NRT runs over Doha, Jeddah, Kuwait and Addis Ababa during the period December 2015-February 2016

3.3.3 Validation against FTIR observations from the NDACC network

In this section, we compare the CO profiles of the CAMS models with FTIR measurements at Maido (21°S, 55°E, i.e. southern tropics, altitude 2.2km) and Lauder (46°S, 169.7°E, altitude 370m). These ground-based, remote-sensing instruments are sensitive to the CO abundance in the troposphere and lower stratosphere, i.e. between the surface and up to 20 km altitude. Tropospheric CO profiles and columns are validated (up to 10km). A description of the instruments and applied methodologies can be found in Eskes et al. (2016).



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, Altzomoni and Maido (mean bias and stddev in %). The overall uncertainty for the CO measurements at Lauder and Maido is approximately 9% and at Maido 5%.

		MAM			JJA			SON			DJF		
		MB	stddev	nobs									
o-suite	Lauder	-8.44	5.07	92	-1.51	4.57	120	-2.36	5.01	148	-8.53	5.27	93
control	Lauder	41.04	5.35	92	28.24	10.4	120	14.13	7.92	148	36.82	12.2	93
o-suite	Maido	-	-	4	-7.37	4.06	125	-6.92	4.04	304	-9.07	3.25	117
control	Maido	-	-	4	13.51	6.46	125	9.68	10.0	304	32.62	5.56	117

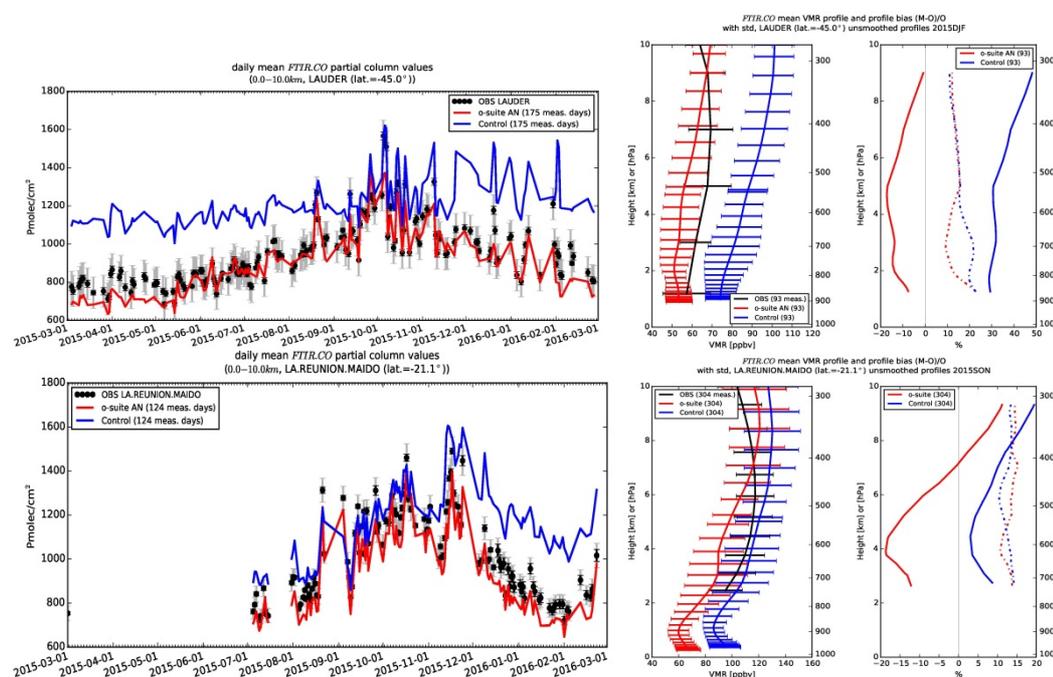


Figure 3.3.13: Daily mean values of tropospheric CO columns (till 10km) by the o-suite (red) and the Control run (blue) compared to NDACC FTIR data at Lauder, New Zealand (45°S, 169.7°E) (top) and Maido (21°S, 55°E) (bottom) for the period March 2015-March 2016. The unsmoothed profile are averaged over Dec-Jan-Feb 2015-2016. The number of measurement days is indicated in the legend.

Table 3.3.1 and Figure 3.3.13 show that the tropospheric columns of CO agree well. Since the beginning of 2015, the o-suite underestimates CO at Lauder with values around -9%, reducing to approx. -2% in JJA and SON 2015, and increasing again to -8% in DJF (which is of the same size as the measurement’s mean uncertainty). 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%.



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 3.3.14 shows the global distribution of CO total columns retrieved from MOPITT (top left) and IASI (top right) and relative biases of model runs with respect to MOPITT V5, averaged for January 2016. MOPITT and IASI show high values over biomass burning areas in Africa, South America and Indonesia. Some differences between observations can be seen in Mid-latitudes over Eurasia and East Asia, indicating higher values in MOPITT compared to IASI, while IASI shows higher values in the north of South America. The modelled geographical distribution indicates a reasonable model performance. The relative difference between the model runs and MOPITT shows that both model runs overestimate CO total column over the central part of South America up to 40% and also over the Indonesia and central Africa. In general, the o-suite performs better than the control, with some overestimation in the tropics over the land and underestimations in the mid-latitudes of up to 30%. The control run also shows an overestimation of CO over the Tropics and the Southern Hemisphere, as well as a large overestimation of CO over the central parts of South America and Australia (up to 40%). Figure 3.3.14 shows no significant difference between the o-suite analysis and 2nd and 4th forecast days.

Figure 3.3.15 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 columns 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 the 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 season, with lower IASI CO total column values (up to 30%). Modelled seasonality of CO total columns is in relatively good agreement with the retrievals. In general, the comparison between the o-suite and the control run shows that assimilation of satellite CO has more positive, pronounced impact on model results over East and South Asia and Africa and smaller impact over other regions.

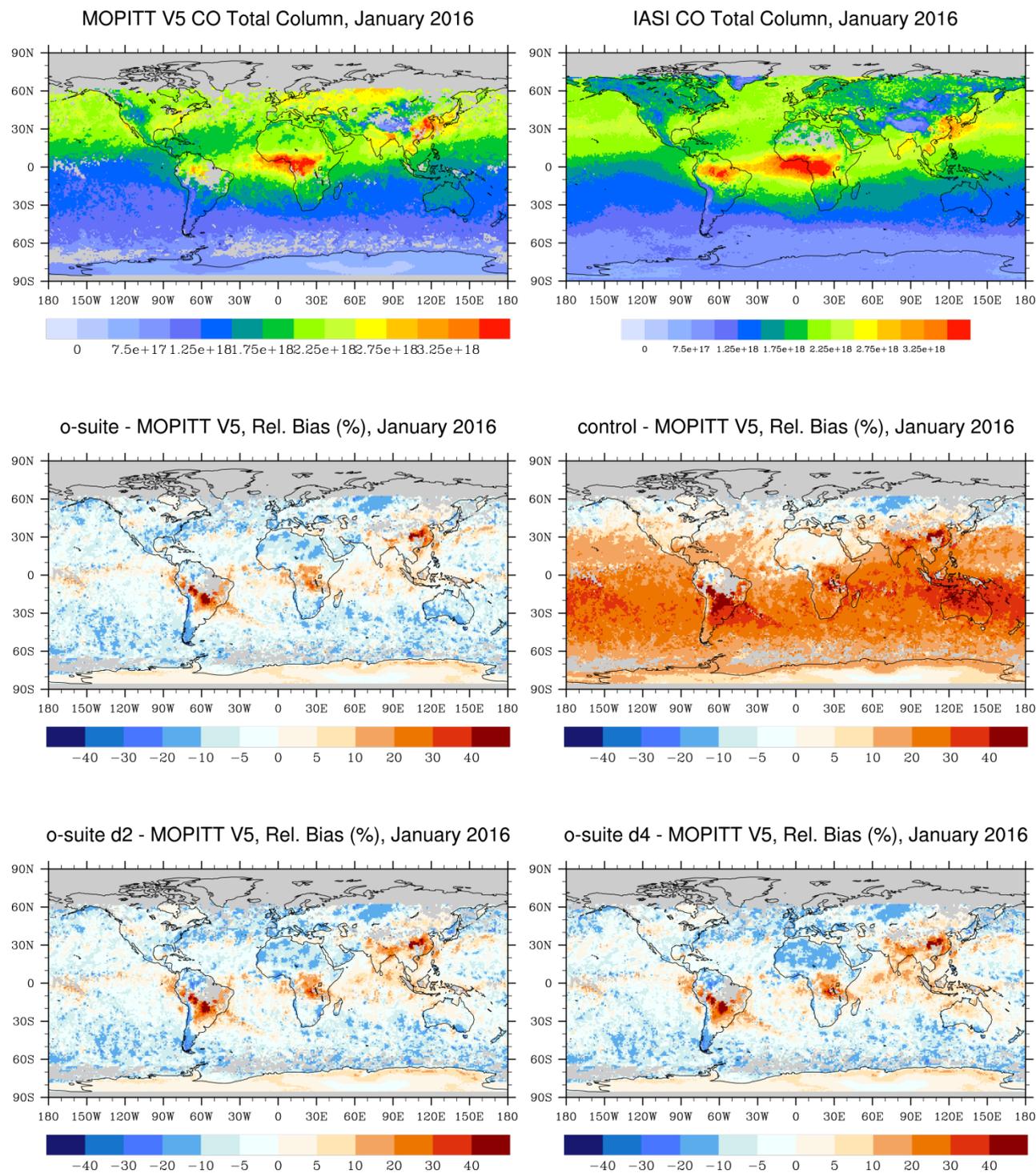


Figure 3.3.14: CO total column for MOPITT V5 (top left) and IASI (top right) satellite retrievals and relative difference between the model runs and MOPITT for January 2016: o-suite (middle left), control run (middle right), o-suite 2nd forecast day (bottom left), o-suite 4th forecast day (bottom right). Grey color indicates missing values.

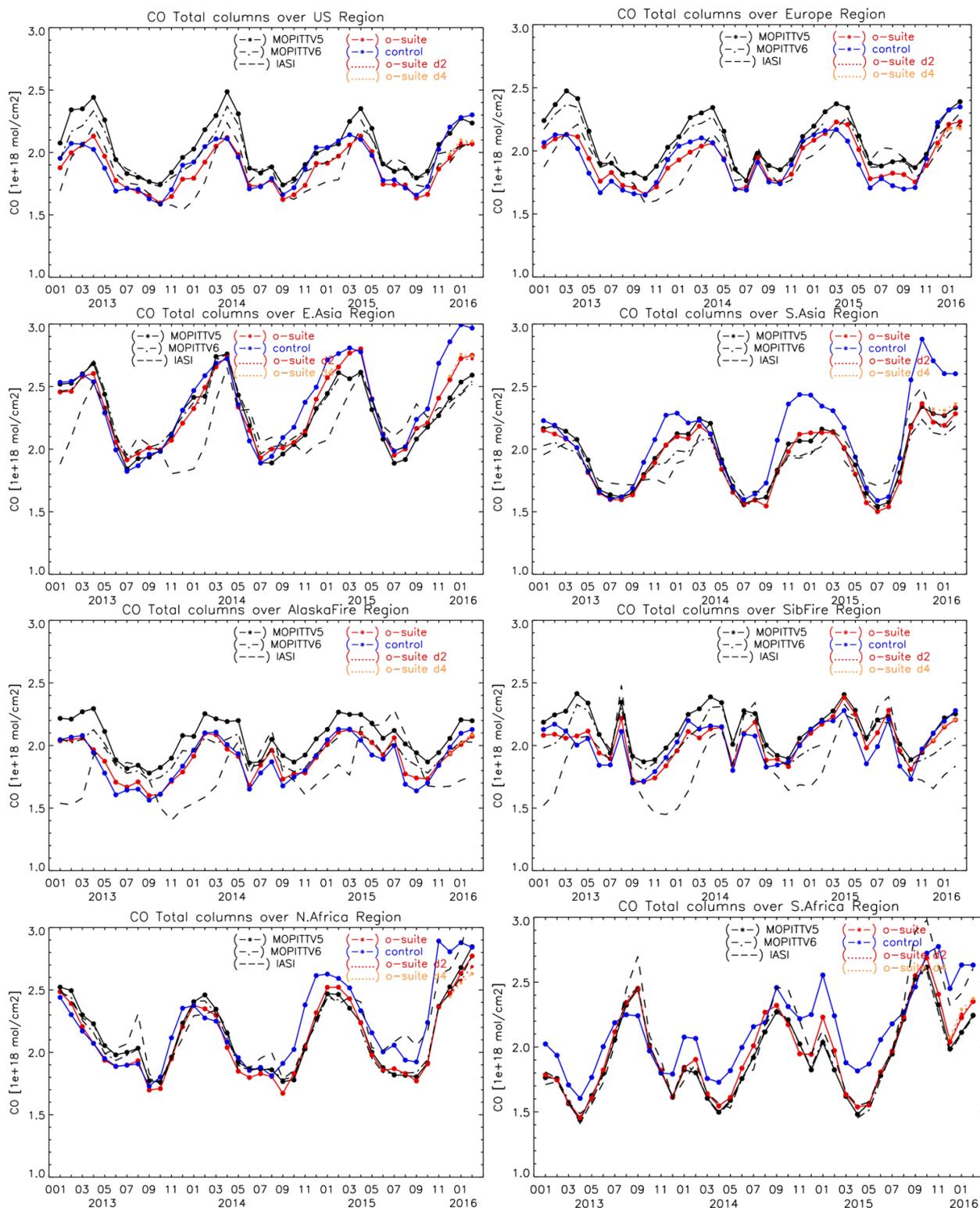


Figure 3.3.15: Time series of CO total column for satellite retrievals MOPIT V5 and V6, IASI (black) and the model runs over the selected regions: o-suite (red, solid), control (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted).

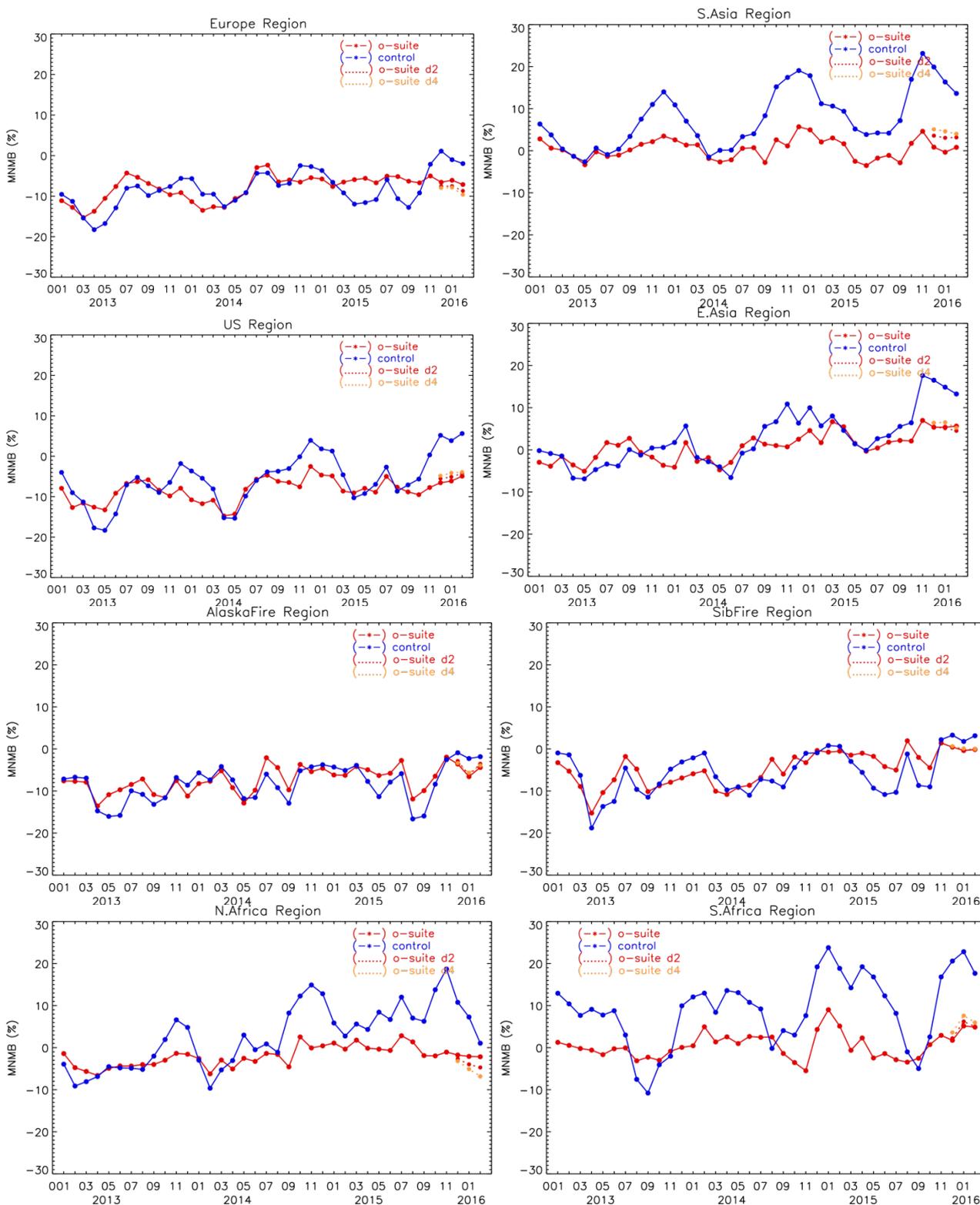


Figure 3.3.16: Modified normalized mean bias (%) for CO total column from the model simulations vs MOPITT V5 retrievals over selected regions; the o-suite (red, solid), and the control run (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted).

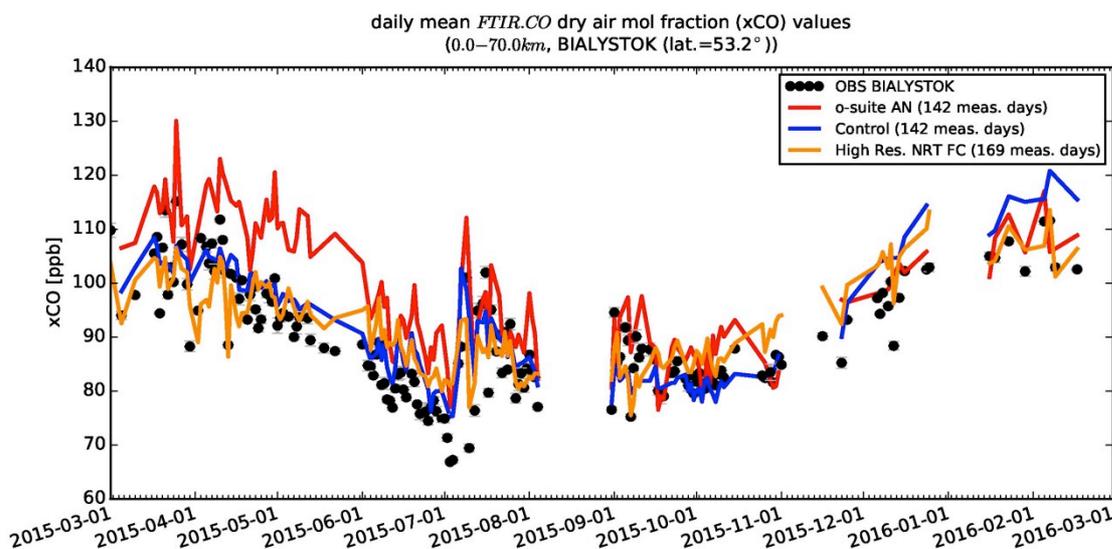


Figure 3.3.17: Time series of column averaged mole fractions of carbon monoxide (CO) at the TCCON site Bialystok compared to the o-suite (red), control (blue) and the high resolution NRT FC model (yellow).

The modified normalized mean bias (MNMB) of the model runs compared to MOPITT V5 (Figure 3.3.16) allows quantifying the impact of the assimilation on the model performance. All model runs show negative biases over Europe, the US region, and Alaskan and Siberian fire regions with some seasonal exceptions. In winter 2016, the o-suite shows good agreement with observations over Asia and Africa with MNMBs within 5%, while the control run shows positive MNMBs of up to 20%. Reasonable model performance for both runs can be seen in Alaskan and Siberian fire regions with MNMBs in the range of 0-5%. The control run shows systematic significant positive biases of up to 20% for South Asia and South Africa (with some seasonal exceptions). The o-suite 2nd and 4th forecast days were also verified in this report. 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 South Asia, the 2nd and 4th days of forecast show up to 3% higher positive biases compared to the analysis. In other regions the forecasts are almost identical to analysis (within 1% difference).

3.3.5 Evaluation against TCCON CO

The high-resolution linear CO model can also be evaluated and compared against CO from the o-suite and control run. At Bialystok and Orleans the control and linear CO model simulations agree within 10% with the measurements (Fig. 3.3.17-3.3.18). 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 (Fig. 3.3.19) the o-suite captures the seasonality and agrees with the measurements within 5-10%. The control and linear CO model simulations show too high values for the period March-August 2015. From September 2015 to February 2016 the linear CO model agrees with the measurements within 20%, but the control run is still 40-60% too high.

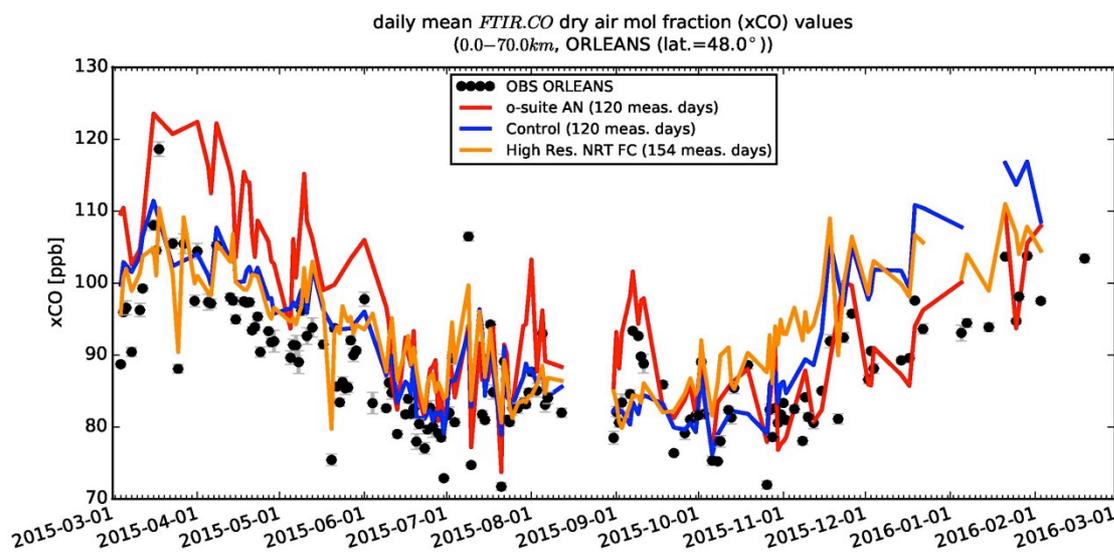


Figure 3.3.18: Time series of column averaged mole fractions of carbon monoxide (CO) at the TCCON site Orleans compared to the o-suite (red), control (blue) and the high resolution NRT FC model (yellow).

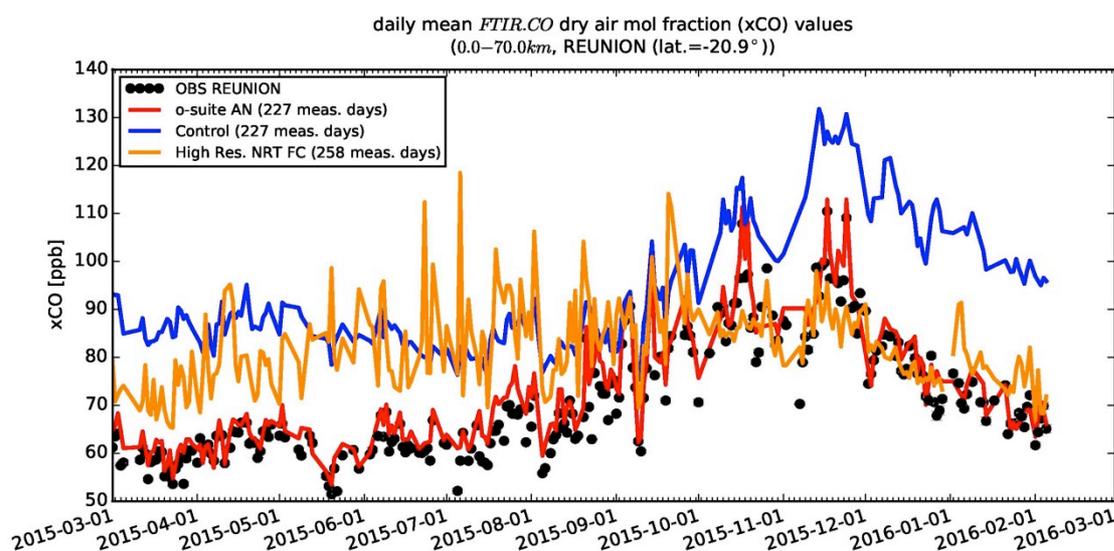


Figure 3.3.19: Time series of column averaged mole fractions of carbon monoxide (CO) at the TCCON site Reunion compared to the o-suite (red), control (blue) and the high resolution NRT FC model (yellow).



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. The uncertainty of monthly mean HCHO columns is relatively large (20% – 40%) and both noise and systematic offsets have an influence on the results, as also discussed in Eskes et al. (2016). However, absolute values and seasonality are retrieved more accurately over HCHO hotspots.

In Fig. 3.4.1, monthly mean satellite HCHO columns are compared to model results for January 2016. The magnitude of oceanic and continental background values and the overall spatial distribution are well represented by the o-suite and the control. Compared to GOME-2 satellite retrievals, there is a strong overestimation of values for Australia and Central Africa. As for tropospheric NO₂, the latter 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 reasonable since 2012.

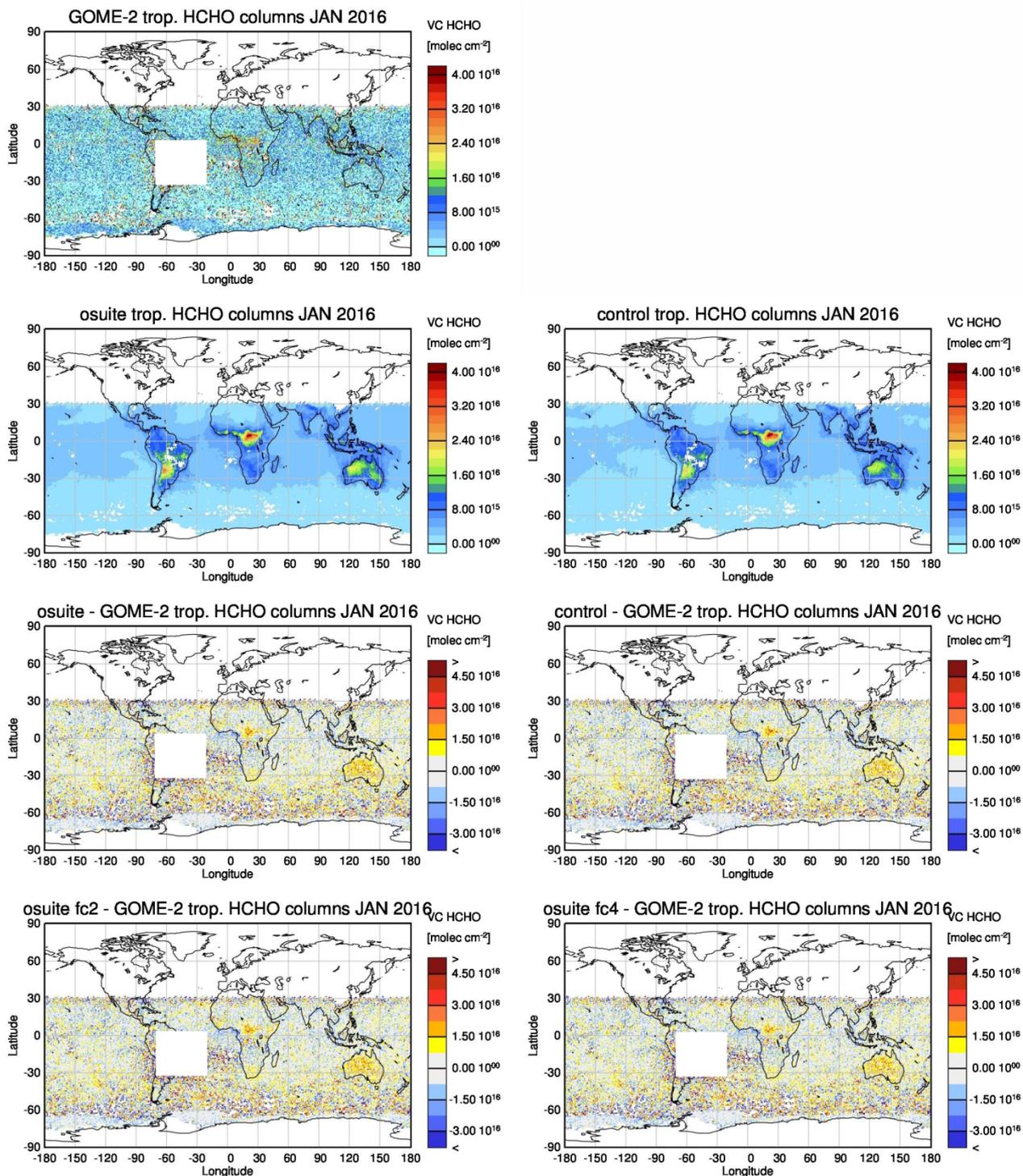


Figure 3.4.1: Global map comparisons of satellite retrieved and model simulated tropospheric HCHO columns [molec cm^{-2}] for January 2016. The top row shows monthly mean tropospheric HCHO columns retrieved by GOME-2; the second row shows the same but for model simulated averages. The third row shows differences of monthly means from models and GOME-2. Images at the bottom show corresponding differences to the o-suite results but for the (left) second day of forecast and (right) fourth day of forecast. GOME-2 data are gridded to model resolution (i.e. $0.75^\circ \text{ deg} \times 0.75^\circ \text{ deg}$). Model data are treated with the same reference sector subtraction approach as the satellite data. Satellite retrieved values in the region of the South Atlantic anomaly are not valid and therefore masked out (white boxes in all images except those which show model results only).

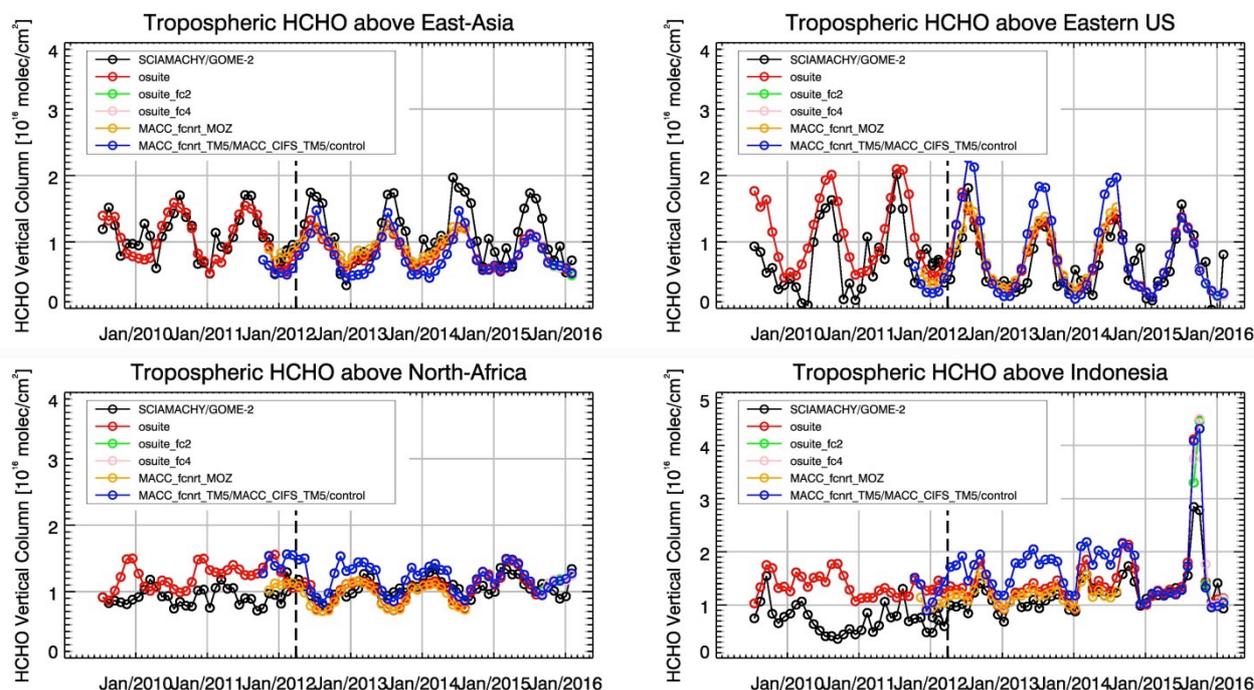


Figure 3.4.2: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions. The blue line shows MACC_fcnrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 to August 2014 and control results from September 2014 onwards. The regions differ from those used for NO₂ (Eskes et al., 2016) 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). osuite_fc2 and osuite_fc4 correspond to the second and fourth day of o-suite forecasts, respectively. Negative values over Eastern US are due to a lack of data (caused by instrument degradation) during Northern Hemisphere winter months for this region. The vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 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 report) show that this is not caused by cloud flagging applied to the satellite and model data.

Same as for tropospheric NO₂, there is virtually no difference between the o-suite results for forecast day 0 and 2, and for all HCHO results shown in this section. Details on the HCHO evaluation can be found at:

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

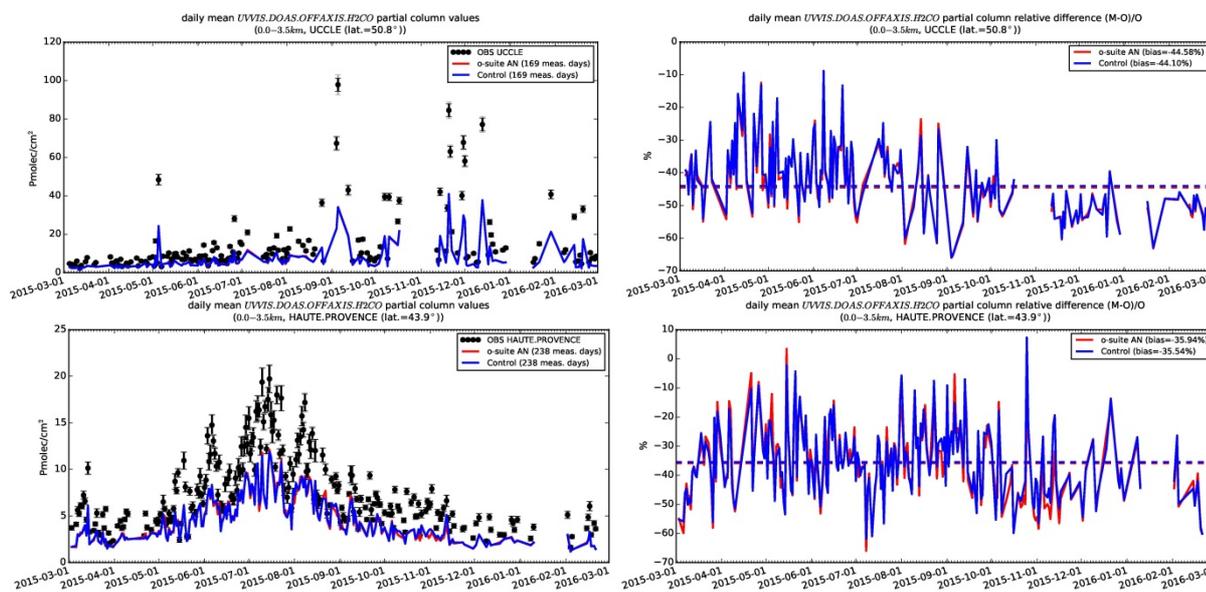


Figure 3.4.3: Daily mean relative differences of tropospheric HCHO columns (till 3.5km) by the o-suite (red) and the control run (blue) compared to NDACC UUVIS DOAS data at Uccle (50.8°N, 4.36°E, middle) and Haute Provence (43.9°N, 5.71°E, rural station, altitude 650m, bottom) for the period March. 2015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

3.4.2 Validation against UUVIS DOAS observations from the NDACC network

In this section, we compare the HCHO profiles of the CAMS models with UUVIS 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 also Eskes et al. (2016).

Smoothing using the averaging kernel of the observations has been applied to allow for a fair comparison to the observations. 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).

From Figs. 3.4.3 and 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. Sensitivity tests using the tropospheric 3D-CTM model IMAGES (Stavrakou et al., 2013) showed that this underestimation could be related to the underestimation of aromatics (benzene, toluene, xylene) in current anthropogenic emission inventories over China (see Liu et al., 2012). Regarding OHP, which is a mostly remote site, the strong underestimation by the model in summer seem to indicate an underestimation of HCHO source gases in biogenic emission inventories. A possible underestimation of regional anthropogenic sources transported to the station should be also investigated (Franco et al., 2015).

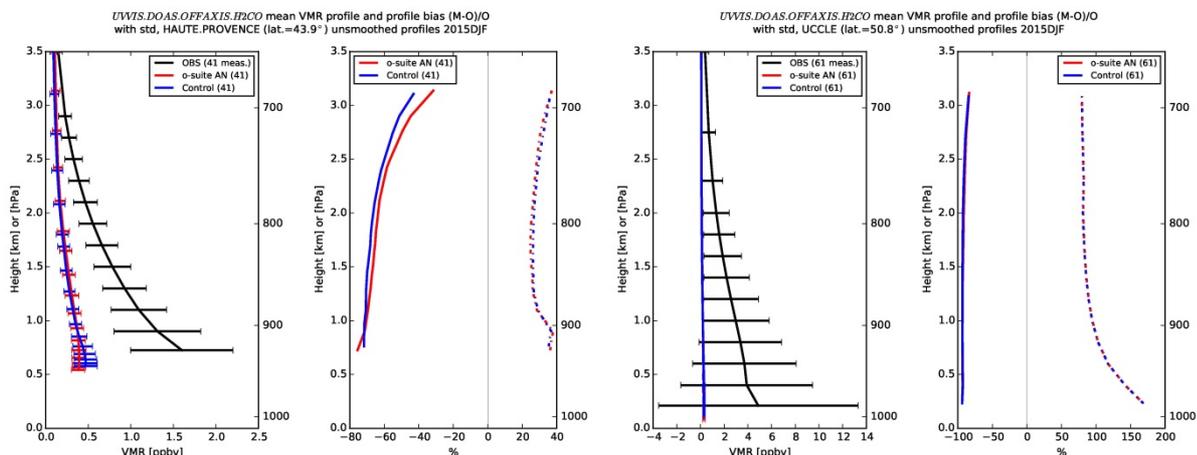


Figure 3.4.4: Mean tropospheric HCHO profiles by the o-suite (red) and the control run (blue) compared to NDACC UVVIS DOAS data at Haute Provence (43.9°N, 5.71°E, left) and Uccle (50.8°N, 4.36°E, right) for the period Dec 2015-Feb 2016.

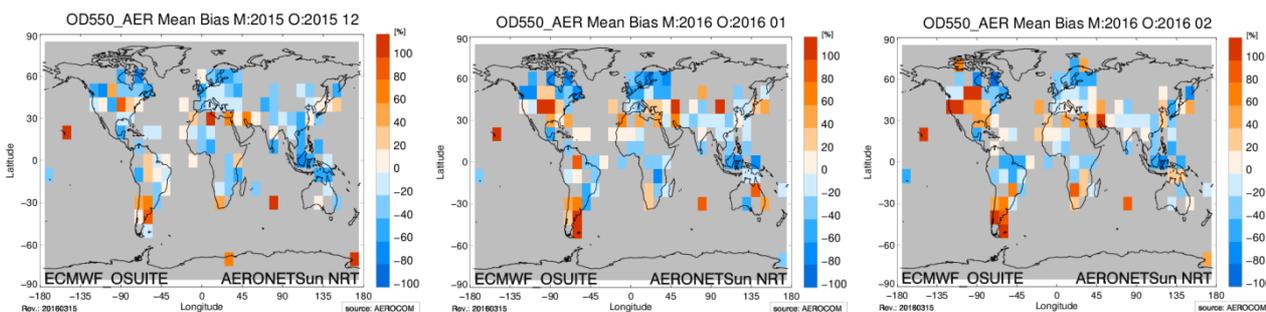


Figure 3.5.1: Aerosol optical depth bias of o-suite in %, against aggregated (10°x10°) NRT Aeronet level 1.5 data for the months of December, January and February 2015/16.

3.5 Aerosol

3.5.1 Global comparisons with Aeronet

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

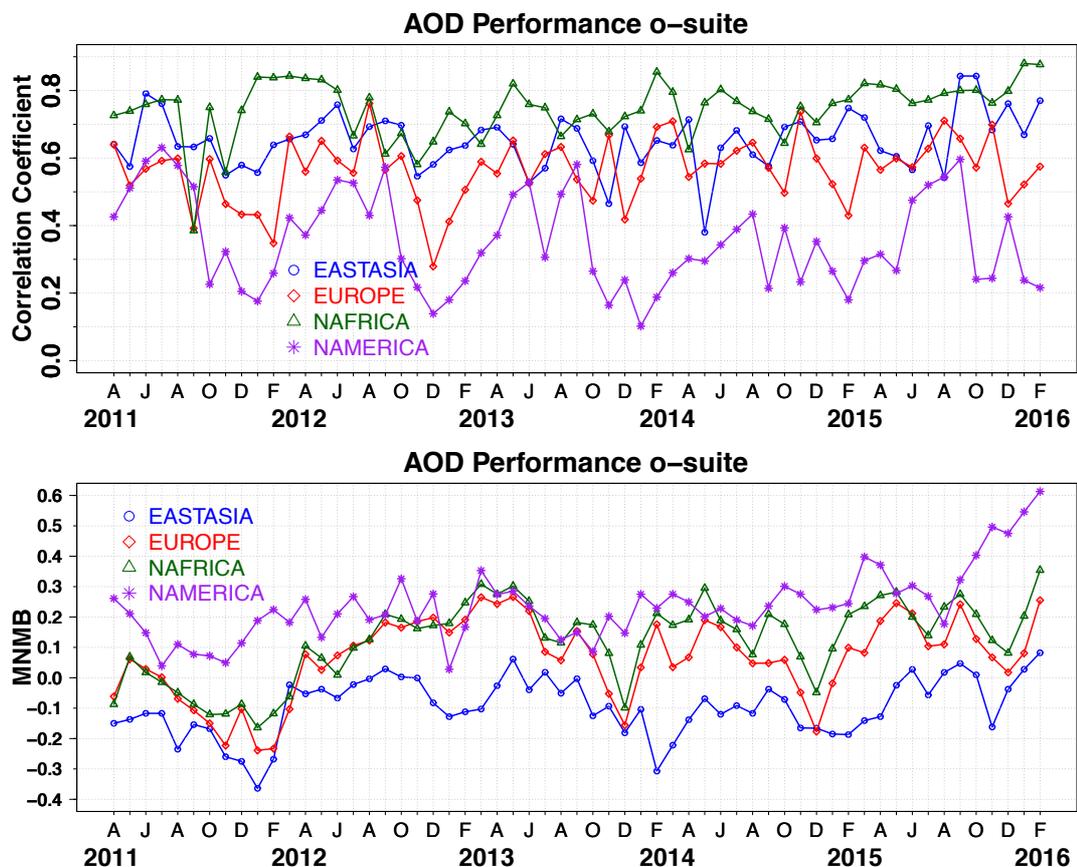


Figure 3.5.1: a) Correlation coefficient and b) modified normalized mean bias (MNMB) in AOD, 2011-2016, based on daily AOD comparison in four world regions [Eastasia(blue); Europe(red); NAfrica(green); NAMercia(purple)] for the o-suite.

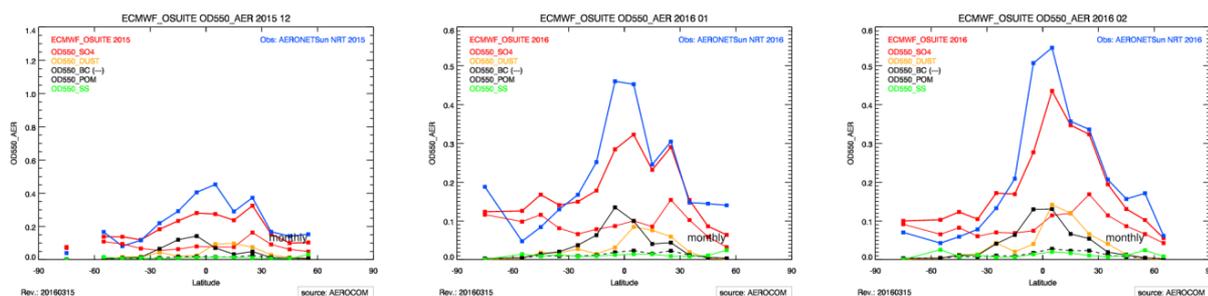


Figure 3.5.2: Aerosol optical depth of o-suite (red) compared to latitudinally aggregated NRT Aeronet level 1.5 data (blue) for the three months covered by this report.

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.

The spatial distributions of the AOD relative bias are shown in Figure 3.5.1. The bias pattern is spatially correlated in between months, with some high simulated aerosol optical depth in North and South America. Note the negative bias in some tropical regions and especially South East Asia, but also Northern Europe and Northern Canada.

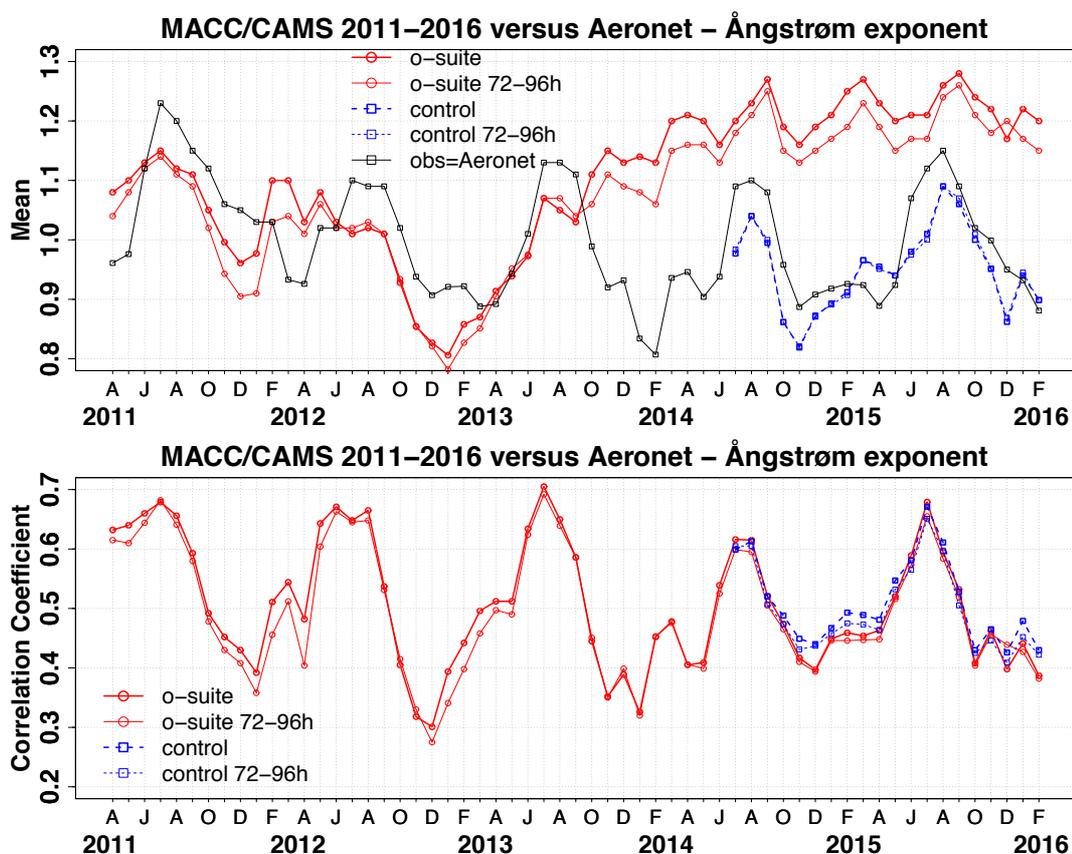


Figure 3.5.3a) Evolution of mean Ångström exponent in o-suite and control at Aeronet sites, based on matching monthly mean values. o-suite (thick red curve); o-suite at last forecast day (light red curve); control (blue dashed curve); control at last forecast day (light blue dashed curve). b) Correlation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

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

The simulated aerosol size distribution may be validated to first order using the wavelength dependent variation in AOD, computed as Ångström exponent, with higher Ångström exponents indicative of smaller particles. a) shows the temporal evolution of simulated and observed mean Ångström exponent, while the correlation is found in figure 3.5.4 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 3.5.4 b). Figure 3.5.5 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%.

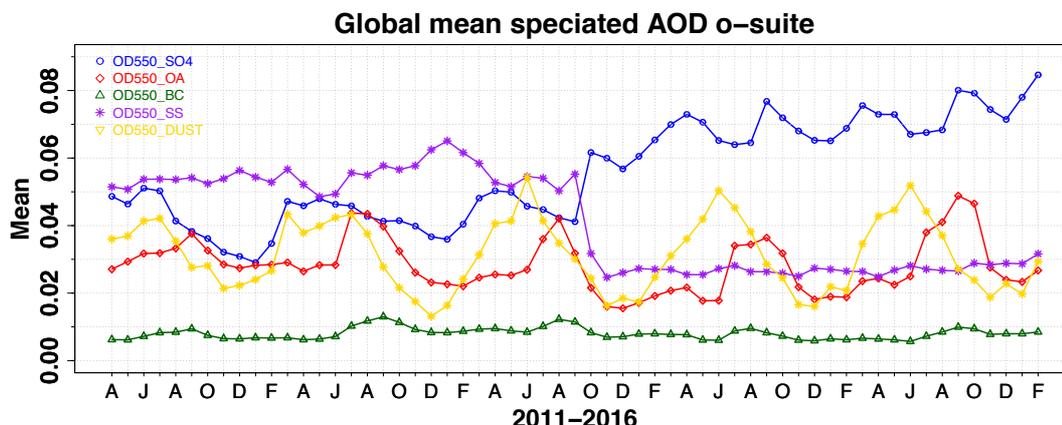


Figure 3.5.4: Evolution of aerosol component’s AOD@550nm [OD550_SO4 = sulphate(blue); OD550_OA = organics(red); OD550_BC = black carbon(green); OD550_SS = sea salt(purple); OD550_DUST = dust(yellow)].

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

	MACC_osuite		MACC_osuite	
	Mean SON 2015 0-24h	Change wrt to first day on day 4	Mean DJF 2015/16 0-24h	Change wrt to first day on day 4
AOD@550	0.179	-15%	0.165	-16%
BC-OD@550	0.009	-22%	0.008	-25%
Dust-OD@550	0.023	13%	0.024	8%
OA-OD@550	0.048	-17%	0.025	-16%
SO4-OD@550	0.078	-25%	0.078	-24%
SS-OD@550	0.028	-7%	0.030	-6%

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) at Aeronet sites. Table 3.5.1 shows an average global decrease in total aerosol optical depth of 16% during the first four forecast days. The contributions to this reduction stem from almost all aerosol components, except 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.

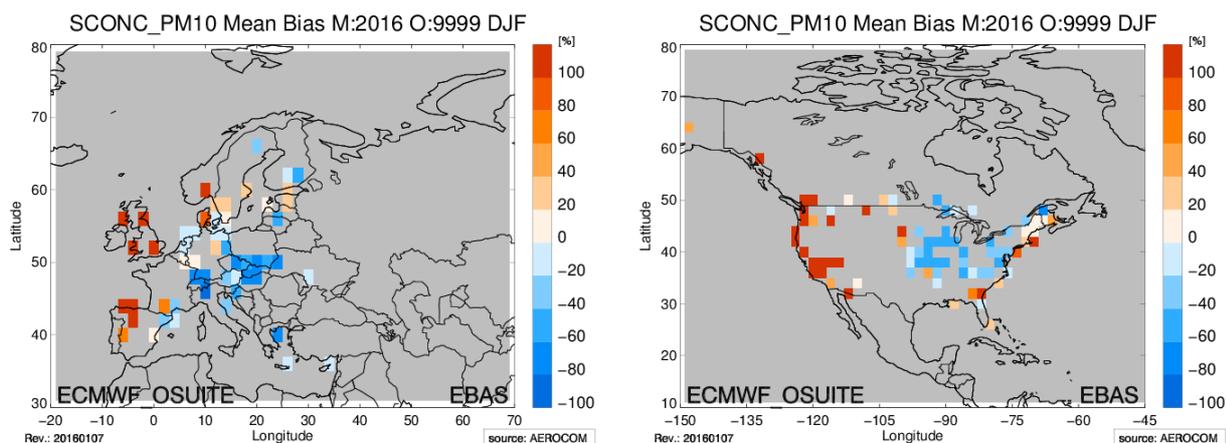


Figure 3.5.5: Bias [%] map of January/February mean PM 10 concentrations at EMEP (Europe) and IMPROVE sites (North America); simulated o-suite versus climatological average (2000-2009).

3.5.2 PM10 evaluation against a climatology over Europe and North America

Surface concentration of particulate matter below 10 μm (PM10) from the o-suite experiment has 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.

3.5.3 Dust forecast validation

Daily dust aerosol optical depth (DOD) from the o-suite and its control run have been validated against 71 AERONET stations grouped in twelve regions for the period 1 December 2015 – 29 February 2016 (Figure 3.5.6), which are dust source and transport regions. In the winter season the minimum dust activity is observed over the entire region. The AERONET (Holben et al., 1998) is used as reference (see Eskes et al. 2016 for details), and also compare with daily SDS-WAS Multi-model DOD Median at 550 nm. 3-hourly values of DOD from AERONET, o-suite, control and SDS-WAS Multi-model Median have been computed for twelve regions shown in Figure 3.5.7 and Figure 3.5.8.

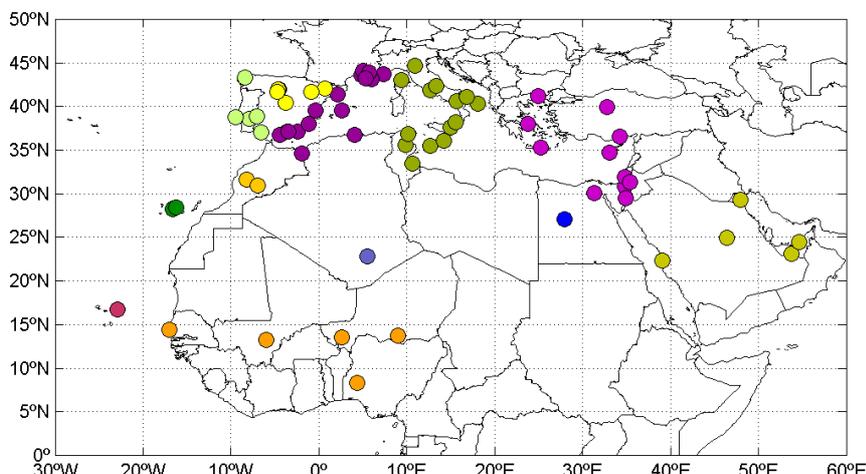


Figure 3.5.6: Map of 71 AERONET level-1.5 stations used in this analysis. The twelve regions considered in the analysis are shown by different colours.

Table 3.5.2: Skill scores (MB, FGE, RMSE and r) for control, o-suite 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. In blue the regions are indicated where o-suite is significantly better than control.

	NDATA	control				o-suite DOD				SDS-WAS Median DOD				
		MB	FGE	RMSE	r	MB	FGE	RMSE	r	NDATA	MB	FGE	RMSE	r
Western Mediterranean	1302	-0.08	1.73	0.25	0.35	-0.09	1.78	0.25	0.36	909	-0.10	1.77	0.27	0.33
Tropical North Atlantic	246	0.11	0.56	0.28	0.59	0.01	0.37	0.24	0.66	195	-0.09	0.40	0.28	0.65
Eastern Mediterranean	1287	-0.02	1.66	0.15	0.55	-0.02	1.67	0.14	0.58	622	-0.03	1.69	0.17	0.52
Sahel	1047	-0.28	0.60	0.54	0.38	-0.29	0.53	0.48	0.72	702	-0.29	0.51	0.48	0.74
Subtropical North Atlantic	359	0.02	1.32	0.28	0.30	-0.04	1.35	0.26	0.32	255	-0.06	1.34	0.27	0.34
Central Mediterranean	1804	-0.04	1.70	0.19	0.37	-0.05	1.72	0.19	0.40	1098	-0.04	1.70	0.16	0.39
Middle East	611	-0.03	1.03	0.17	0.60	-0.04	1.01	0.16	0.68	219	-0.06	1.01	0.19	0.68
Iberian Peninsula	472	-0.10	1.72	0.31	0.44	-0.12	1.78	0.32	0.44	341	-0.13	1.78	0.34	0.45
Western Iberian Peninsula	410	-0.13	1.60	0.26	0.30	-0.15	1.64	0.27	0.30	282	-0.16	1.64	0.28	0.29
North Western Maghreb	312	-0.07	1.48	0.28	0.16	-0.08	1.55	0.28	0.17	224	-0.08	1.52	0.29	0.18
Central Sahara	223	0.05	0.70	0.15	0.28	0.04	0.66	0.14	0.29	145	-0.04	0.57	0.12	0.33
Eastern Sahara	256	-0.02	0.76	0.10	0.76	-0.02	0.74	0.10	0.76	120	-0.03	0.55	0.09	0.82

Both CAMS runs similarly underestimate DOD in almost all regions, but especially in the Sahel, Northwestern Maghreb and the Iberian Peninsula, quite similar to the SDS-WAS Multimodel Median. In the Sahel, o-suite shows better results in terms of correlation than control (from 0.38 for control to 0.72 for o-suite) although DOD is highly underestimated (MB from -0.28 for control to -0.29 for o-suite). On the contrary, in Tropical North Atlantic o-suite reduces overestimations observed in the control experiment (see Cape Verde in Figure 3.5.7). In Middle East, we observe good results with negative MB in both experiments CAMS (-0.03 for control and -0.04 for o-suite, see Table 3.5.2) although the correlation is better for o-suite (from 0.60 for control to 0.68 for o-suite). In the Mediterranean, both CAMS experiments show a

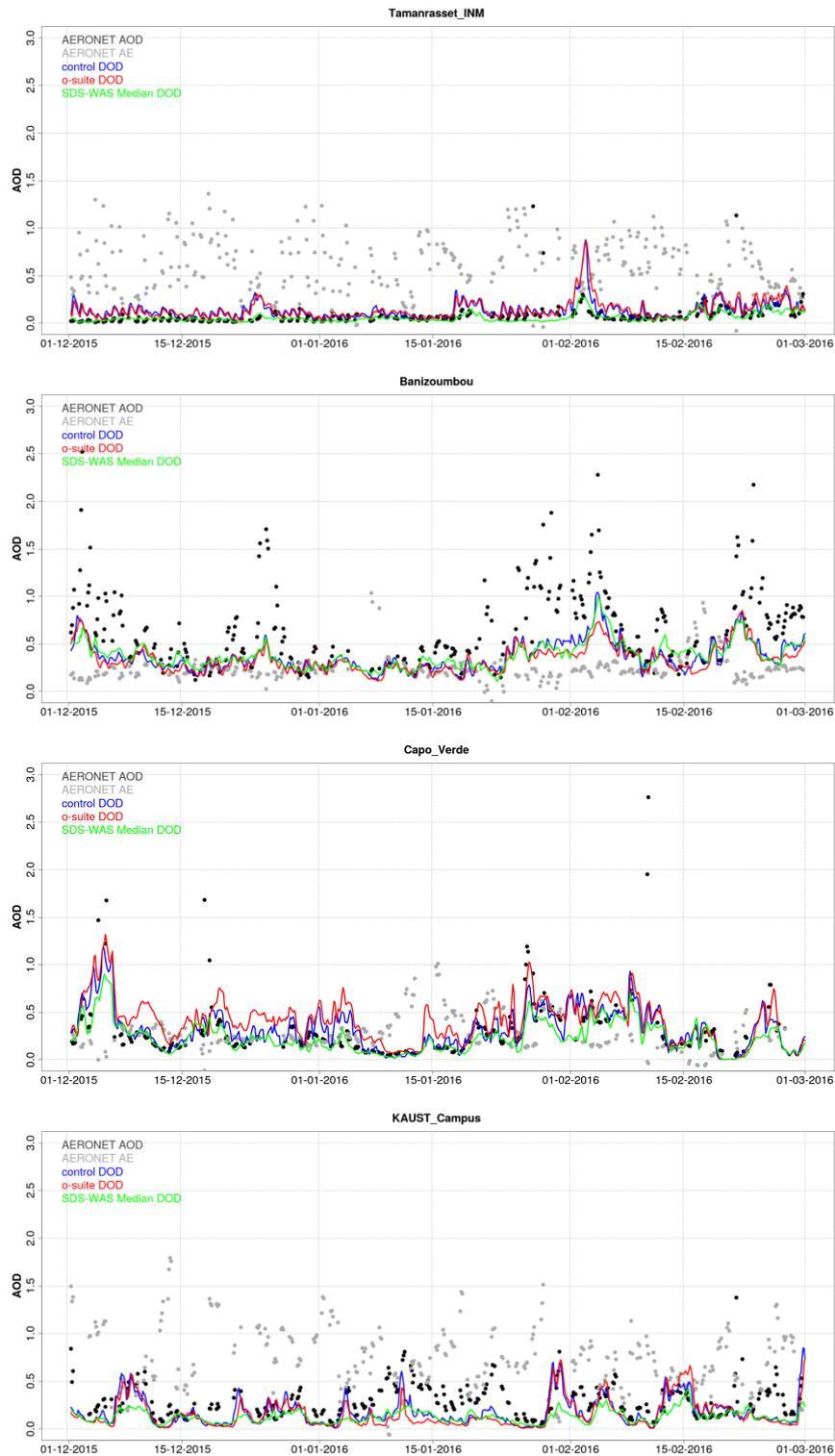


Figure 3.5.7: AOD from AERONET (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multimodel SDS-WAS Median (green line) for the period December 1st, 2015 to February 29th, 2016 over Tamanrasset INM (Sahara), Banizoumbou (Sahel), Cape Verde (Tropical North Atlantic) and Kaust Campus (Middle East).

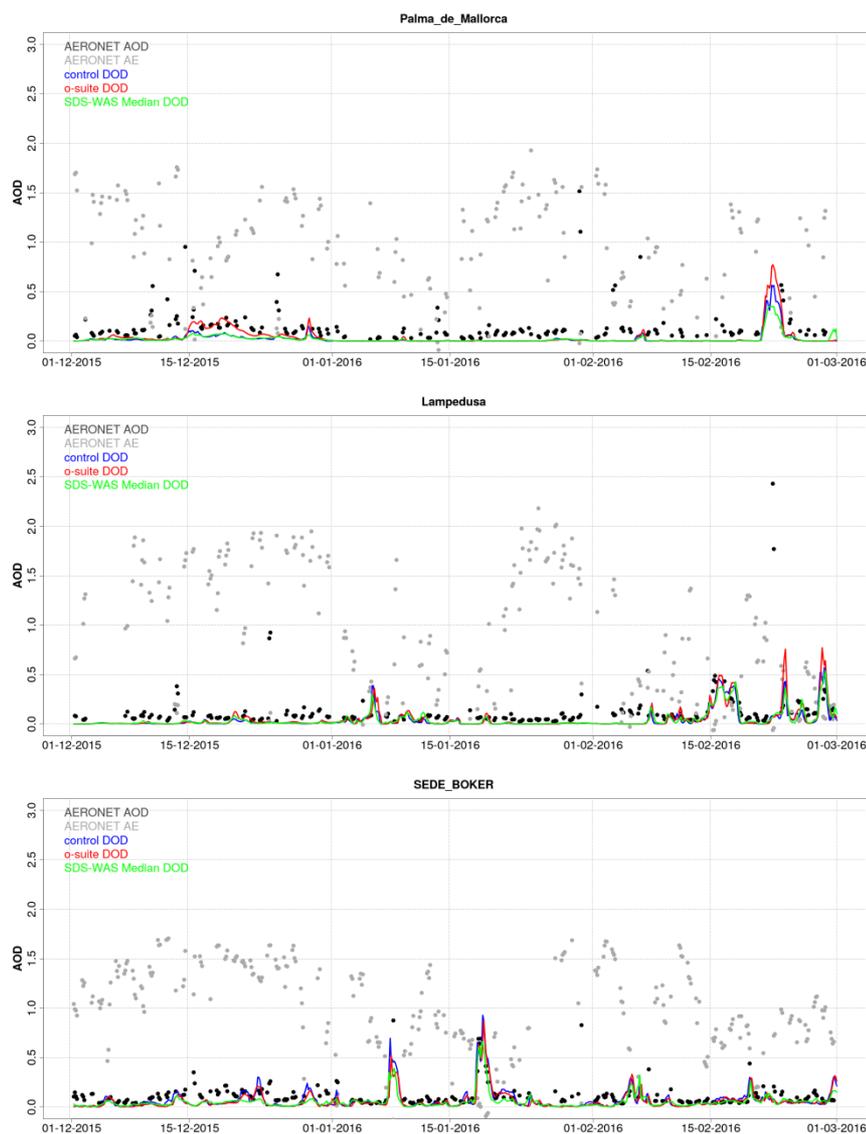


Figure 3.5.8: AOD from AERONET (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multimodel SDS-WAS Median (green line) for the period December 1st, 2015 to February 29th, 2016 over Palma de Mallorca (Western Mediterranean), Lampedusa (Central Mediterranean), Sede Boker (Eastern Mediterranean).

similar behavior (see Table 3.5.2) with good agreement with AERONET capturing all dust outbreaks (see Figure 3.5.7 and Figure 3.5.8).

During the December-February season the dust activity over desert dust sources (Middle East, the Sahara and dust corridor of North Western Maghreb) is relatively low. In this period the o-suite reproduces the daily variability of AERONET observations better than the control experiment, see Table 3.5.2.

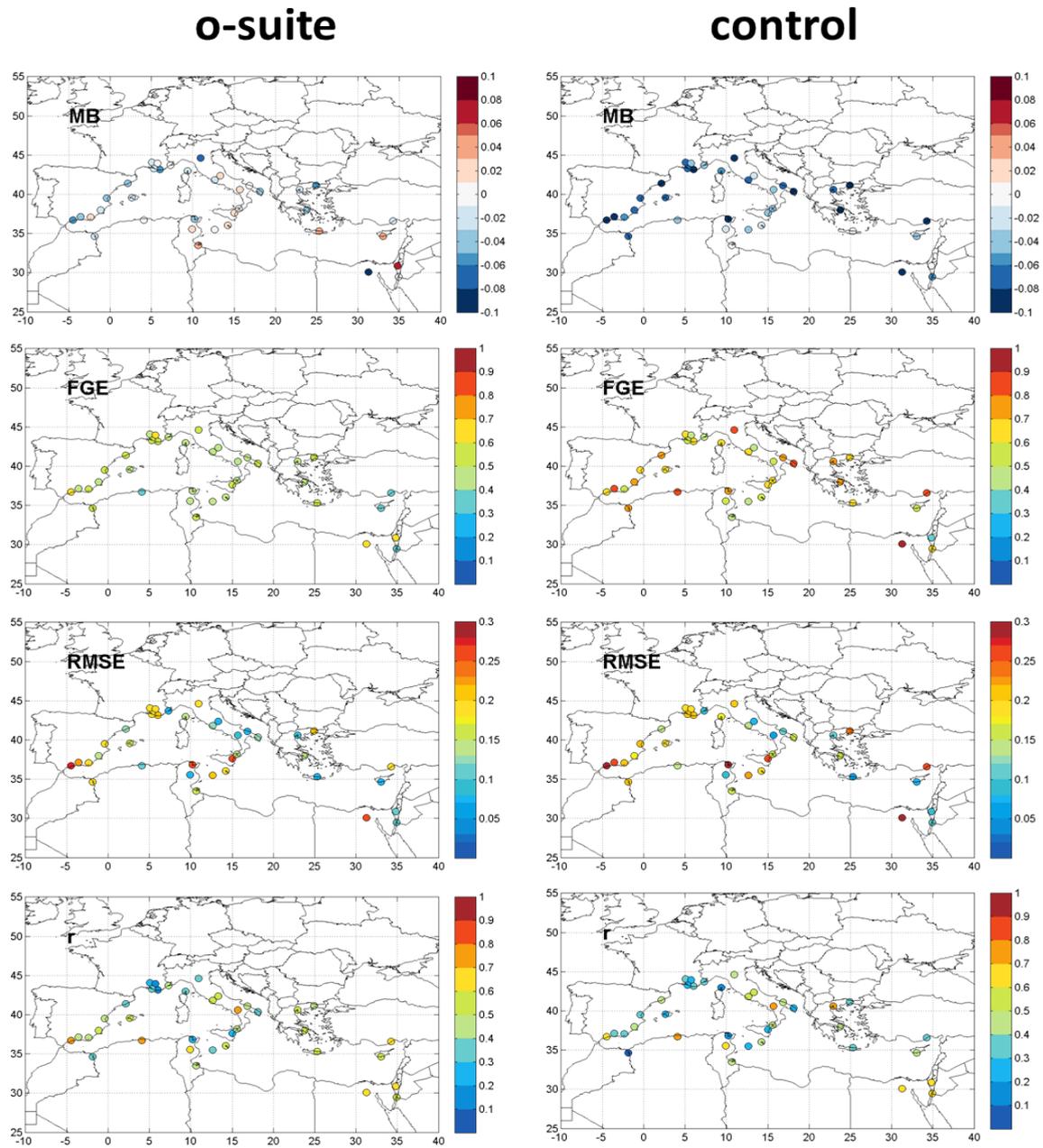


Figure 3.5.9: Skill scores (MB, FGE, RMSE and r) for 24-hour forecasts of CAMS o-suite and control for the study period. AOD from AERONET is the reference.

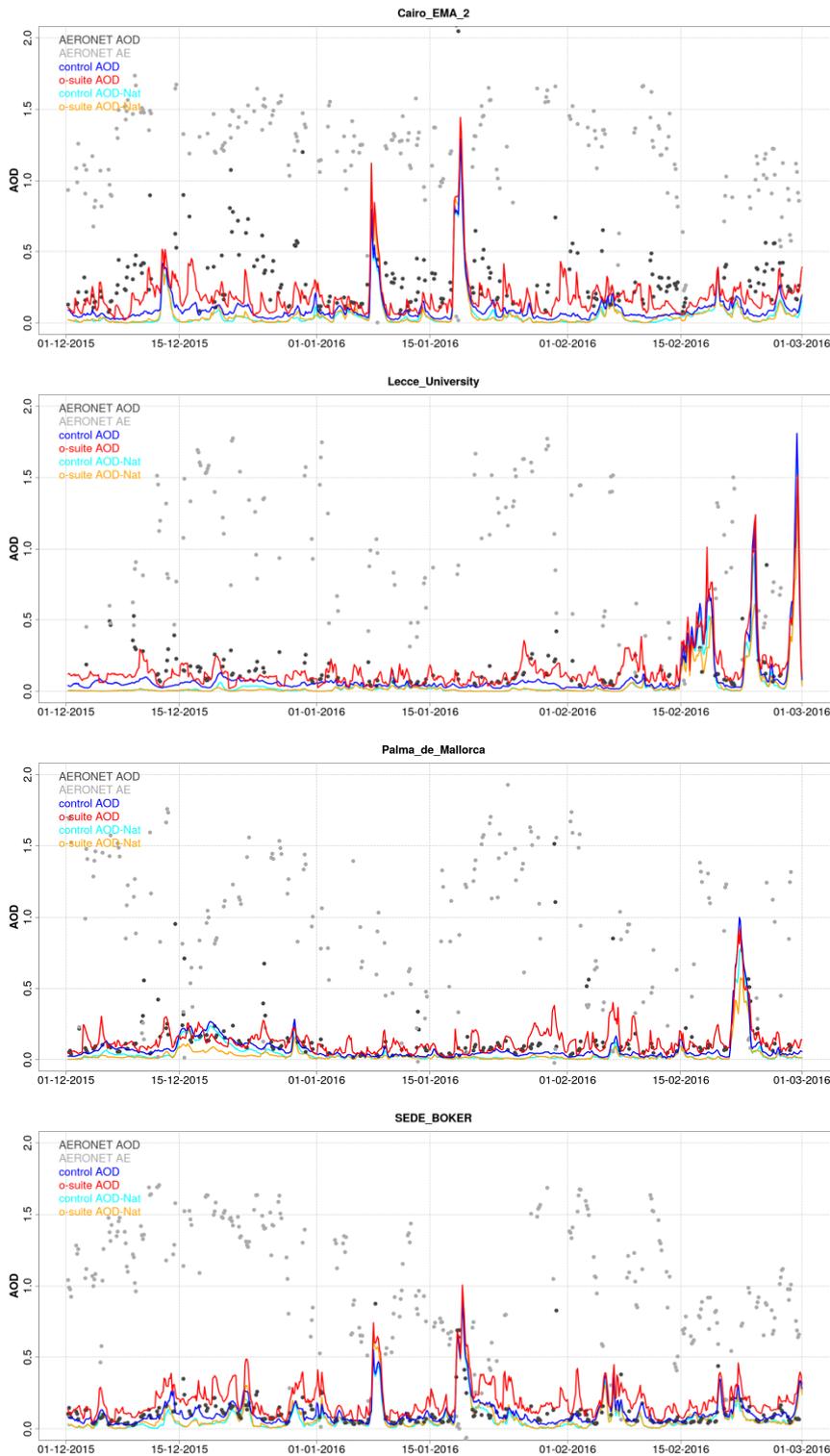


Figure 3.5.10: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), for the period September 1st to November 30th, 2015 over December, 2015 to 29th February, 2016 over Cairo EMA, Lecce University, Palma de Mallorca and Sede Boker. AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt.



3.5.4 Aerosol validation over the Mediterranean

Daily aerosol optical depth (AOD) and surface PM₁₀ and PM_{2.5} concentration from CAMS o-suite 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 December 2015 – 29 February 2016. Here we assess the model performance over the Mediterranean in the winter season, when the minimum dust activity is observed over the entire region.

Aerosol optical depth

From December 2015 to February 2016, the CAMS o-suite better reproduces the daily variability of AERONET observations than the control run (see correlation in Figure 3.5.9), and also provides a better representation of the AOD background levels, while the control run tends to underestimate compared with AERONET observations (see MB in Figure 3.5.9). The highest peaks in the AOD model runs are linked to natural sources (mainly desert dust, see Figure 3.5.10) that are well reproduced by both model runs.

Surface aerosol concentrations

From December 2015 to February 2016, CAMS control run and o-suite are able to reproduce the daily variability of Airbase observations associated to dust events as that at end-February, in which the ratio PM_{2.5}/PM₁₀ is close to 1. This event is well predicted in terms of AOD by both CAMS experiments (see Palma de Mallorca/Hospital Joan March in Figure 3.5.11). However, the CAMS simulations (control and o-suite) of this high aerosol event strongly overestimate PM₁₀/PM_{2.5} air quality measurements over Spain as shown in Zorita station in Figure 3.5.11.

Other events, with PM₁₀>50 µg/m³ and PM_{2.5}>25 µg/m³, associated to non-dust aerosols are predicted by both CAMS experiments during the period analysed, which are not observed in the Airbase stations (see Hospital Joan March station in Figure 3.5.11). In average for all the sites, PM₁₀ MB increases from 2.5 µg/m³ for control to 6.4 µg/m³ for o-suite, while PM_{2.5} MB increases from 1 µg/m³ for control to 2.4 µg/m³ for o-suite. Only in end-December in Western Mediterranean, control provides higher PM₁₀/PM_{2.5} values than o-suite (see Figure 3.5.11). Unfortunately, there are no observations available during this period to validate the experiments.

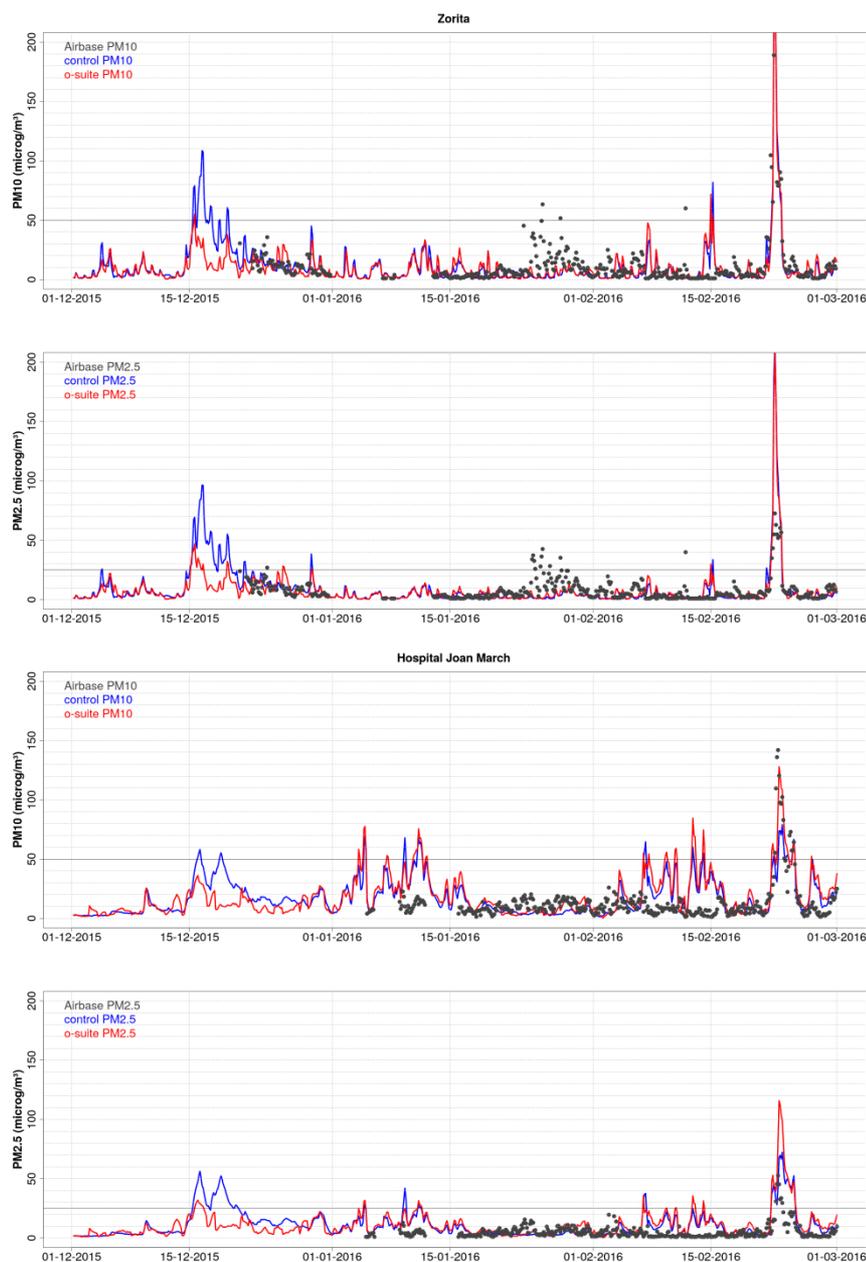


Figure 3.5.11: PM10 and PM2.5 Airbase observations (black dot), PM10 and PM2.5 o-suite (red line) and PM10 and PM2.5 control (blue line) for the period December 1st to February 29th, 2015 over Zorita (40.73°N; 0.17°W, Spain) and Hospital Joan March (39.68°N; 2.69°E, Spain).

3.5.5 Aerosol backscatter profiles

The technical specifications of the data sources, evaluation parameters and methods are described in Eskes et al (2016). Tools and concept of IFS aerosol profile evaluation with ceilometer backscatter data are still in development. In this section, the vertical variation of the backscatter coefficient (abbreviated bsc) profiles, i.e. correlation and standard deviation, of o-suite 'g9rr' and control run 'geuh' 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. A skill measure for the horizontal extension is not yet included.

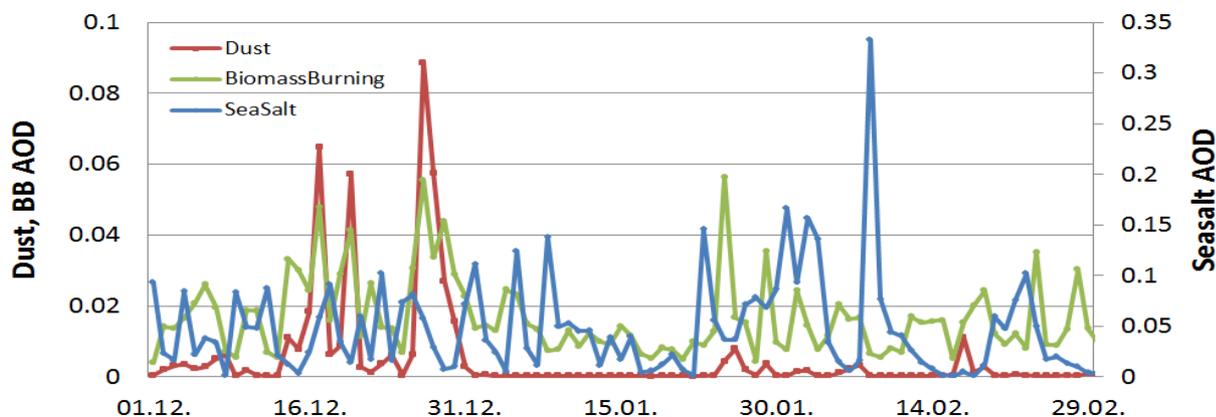


Figure 3.5.12: Maximum AOD around Soltau station ($\pm 1^\circ$ lat/lon) for sea salt (blue), dust (red) and biomass burning (=OC+BC) aerosol.

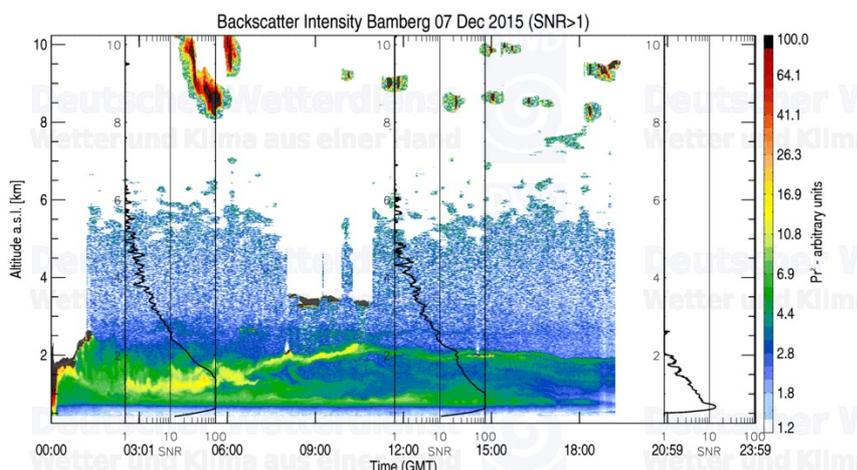


Figure 3.5.13: Backscatter signal (Pr²) at Bamberg on 7 Dec 2015, including 1-hour average signal-to-noise-ratio profiles at 3, 12, and 21 UT.

The model aerosol optical depth (AOD) is used to select cases or averaging periods when prominent production/transport of mineral (Saharan) dust (SD), sea salt (SS) or biomass burning (BB) plumes to Germany were predicted. In Figure 3.5.12 this is shown for Soltau (53N, 10E). Accordingly, SD is studied for 7/8 Dec 2015 and the periods 14-20, 24-30 Dec 2015, elevated SS levels for 8 February 2016 and for 20-24 February 16. The biomass burning aerosol contribution was small over Germany during the reporting period. Elevated contributions were either associated with Saharan dust or sea salt (see discussion below), which cannot be distinguished by ceilometers.

Saharan Dust: 7/8 Dec 2015

We analyse an eastward swaying plume, as observed over Bamberg (Southern Germany) on 7 December 2015, Figure 3.5.13. The plume is generally reproduced, but the thin filament around 2km height and internal structure are equally unresolved in osuite and control. The three-hourly profile shapes of

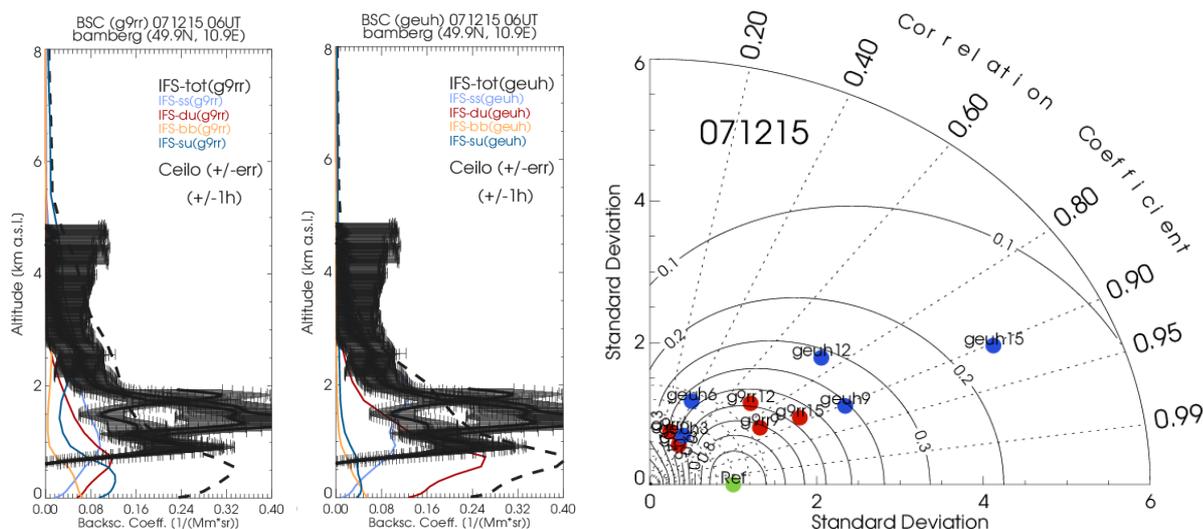


Figure 3.5.14: Left: profiles of backscatter coefficients for Bamberg on 7 Dec 15, 06 UT from o-suite (far left) and control (middle), splitted for contributions from sea-salt (light blue), dust (red), biomass burning (orange) and sulfate (dark blue), as well as the total (dashed black). Solid black lines: Ceilometer data with estimated error bars, plotted for -1h, 0h, +1h around the time of the 3-hly model profile. Right: Taylor polar plot of standard deviation vs correlation coefficient, evaluating the magnitude variability and shape within each of the 3-hourly profiles.

control and osuite are very similar (Figure 3.5.14) and the peak tends to be some 100m too low, control has two times larger bsc and larger diurnal variation than osuite (cf. Taylor plot in Fig. 3.5.14). Control has less overall bias in the morning but factor 2-3 in the afternoon and vice versa. The conspicuous (and often observed) large contribution from sea salt at this station (approx. 400 km inland) indicates shortcomings in the source or sink representation (see below).

Saharan Dust: 17 Dec 2015

Eastward swaying plume on 17 December 2015, observed over N-Germany (in Figure 3.5.15 data from Soltau station is shown). The initially flat profile shape within the plume is neither captured by the osuite nor control. Nevertheless, the 21 UT profile is precisely captured by osuite. Again the overestimation by control is larger. This time the dominant dust contribution indicates a largely unmixed SD air-mass. Unfortunately, a rain shower within the averaging period limited the height range of the ceilometer profile around 21 UT. As shown in the Taylor polar diagram, the correlation is high for both runs but the variability, due to the larger values, is larger than in the observation.

Saharan Dust: 27-29 December 2015

An SD plume moving over the North-Sea, with filaments crossing Germany. As for other German stations, the layers in Geisenheim (50°N, 8°E) are forecasted around the observed altitude, but not resolved vertically and lagged in time by several hours (Figure 3.5.16). Thus, the variability in the model is too small (r-coordinate in the Taylor diagram) and the correlation quite low. The integral over the plume maximum seems to be met quite well, though.

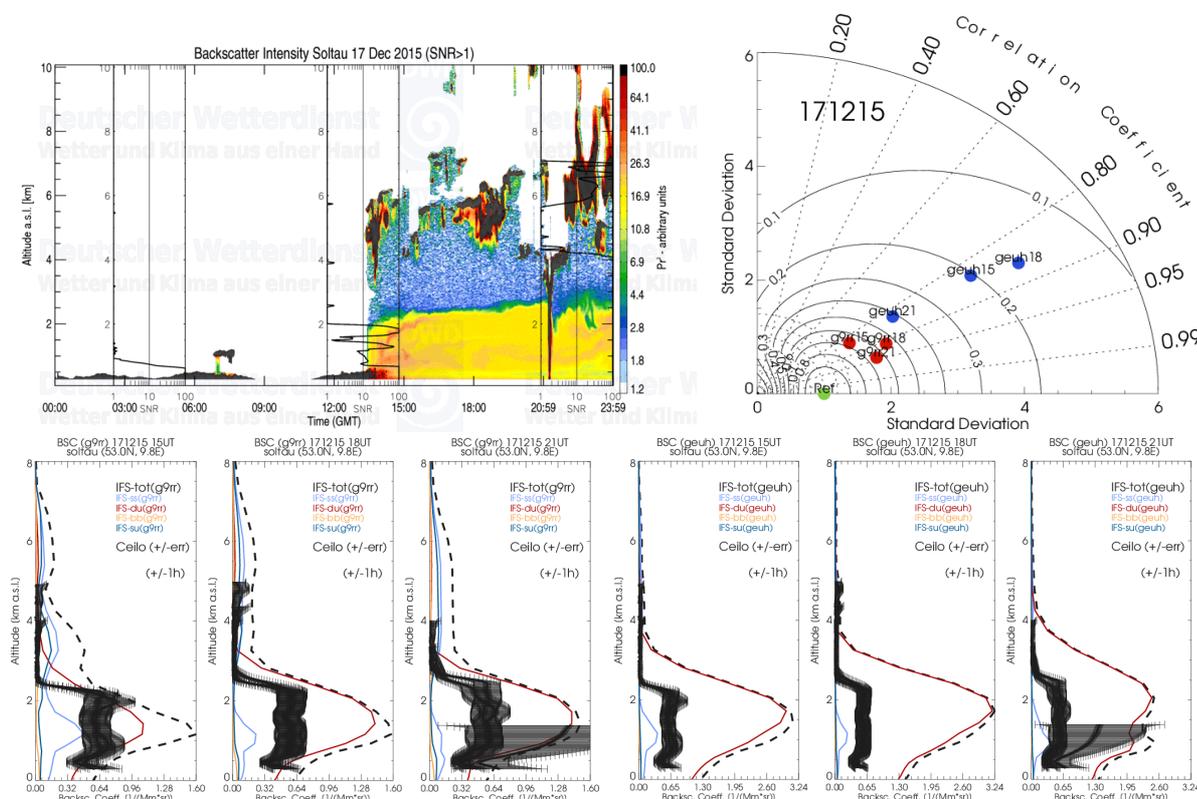


Figure 3.5.15: Top panel: Backscatter signal (Pr^2) at Soltau and Taylor diagram for 17 Dec 15. Bottom panel: Profiles as in Figure 3.5.14 but for 15/18/21 UT on 17 Dec 2015. Left three plots: o-suite, right three plots: control.

Sea salt

Elevated sea salt (SS) abundance was predicted for 1, 2, 3, 8, 9, 20, 21, 22, 23, 24 February 2016, particularly on 8 February associated with a storm over the English Channel. Generally, the model captures much of the observed profile properties as shown by the Taylor plots in Figure 3.5.17 for the inland stations Soltau and Lindenberg, and the left two panels in Figure 3.5.18 except during the storm period on 8 February. On this day the model aerosol consists almost entirely of SS. However, note that the aerosol speciation cannot be verified by the ceilometers. Further, the data coverage of the ceilometers is strongly limited by many clouds and rain. While the SS concentration before the storm is off with the observations by a factor of 1.5 - 2, it is overestimated during the storm, increasingly from the coast (e.g. Pelzerhaken) to several hundred km inland (e.g. Lindenberg) by a factor of 6 and more, as shown in Figure 3.5.17 and Figure 3.5.18 C and D. This indicates a too strong sea salt source as well as underestimated deposition/sedimentation velocity. The latter is most sensitive to the size distribution, thus indicating too small sizes. Production is calculated assuming 80% relative humidity (Morcrette et al, 2009), but only dry mass/particles are transported in the model without transfer between size-bins, while salt grows considerably at humidities >80%. Though wet density and radius are considered for deposition/sedimentation, this seems to be too slow. The representation by the 3 size-bin setup and associated modal splits adds further uncertainty. Note that, although the largest contribution to the model backscatter coefficient is from the 0.5-5 μm sizes (\sim factor 3 larger than from the coarse bin), the mass in the model is still dominated by the coarse particles with diameter 5-20 μm due to their smaller backscatter/mass

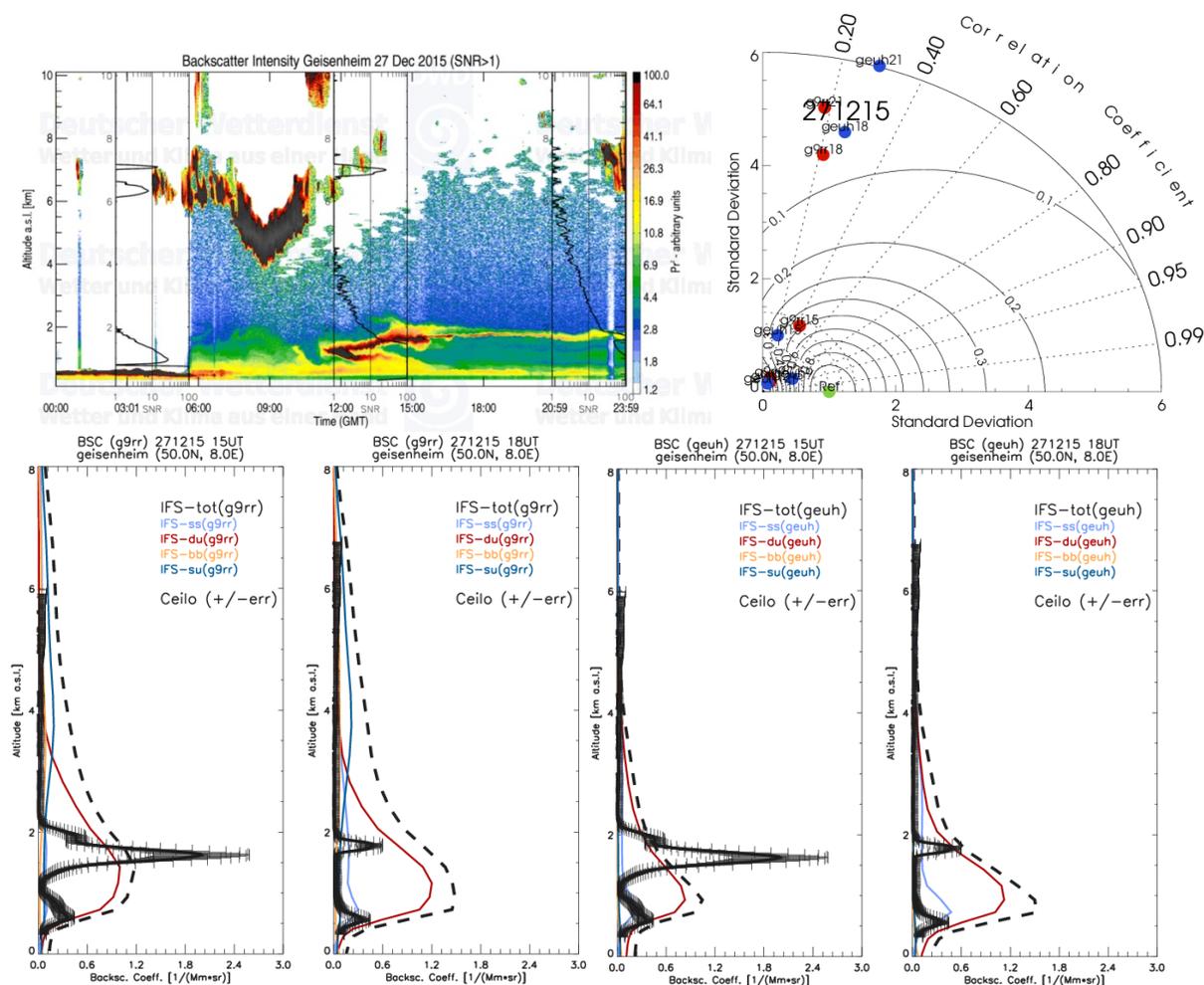


Figure 3.5.16: As Figure 3.5.14, but for Geisenheim on 27 Dec 2015, 15UT and 18UT. Left: o-suite (g9rr), right control (geuh) (please note the none-standard color-coding).

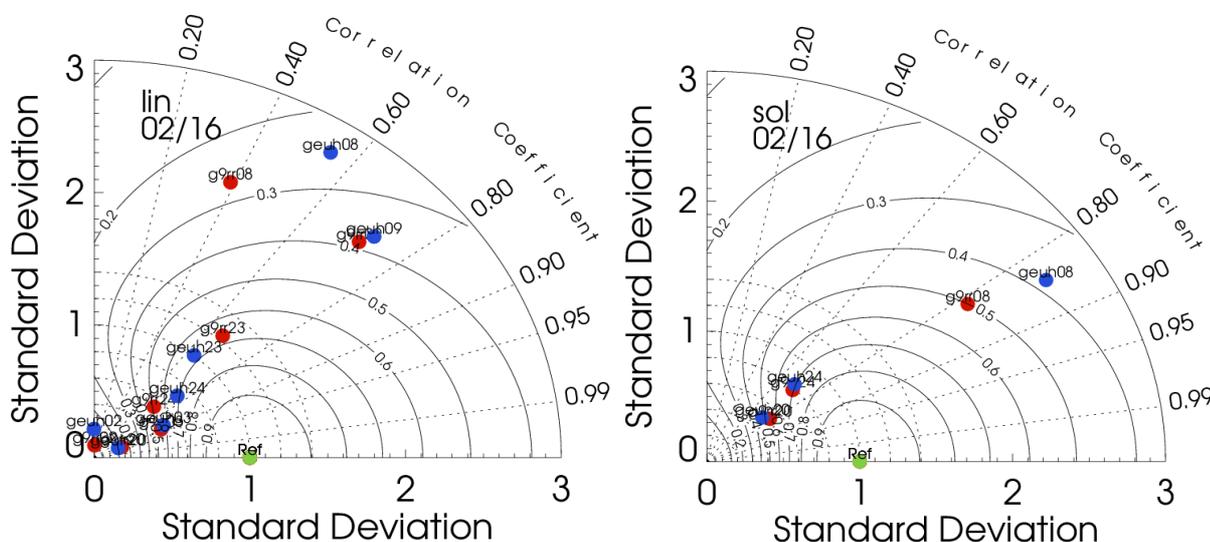


Figure 3.5.17: Taylor plots of correlation and variability for days with large sea salt abundance in Feb 2016 at the stations Lindenberg (left) and Soltau (right). The numbers associated with the red (g9rr, osuite) and blue (geuh, control) points refer to the day of the profile in February 2016.

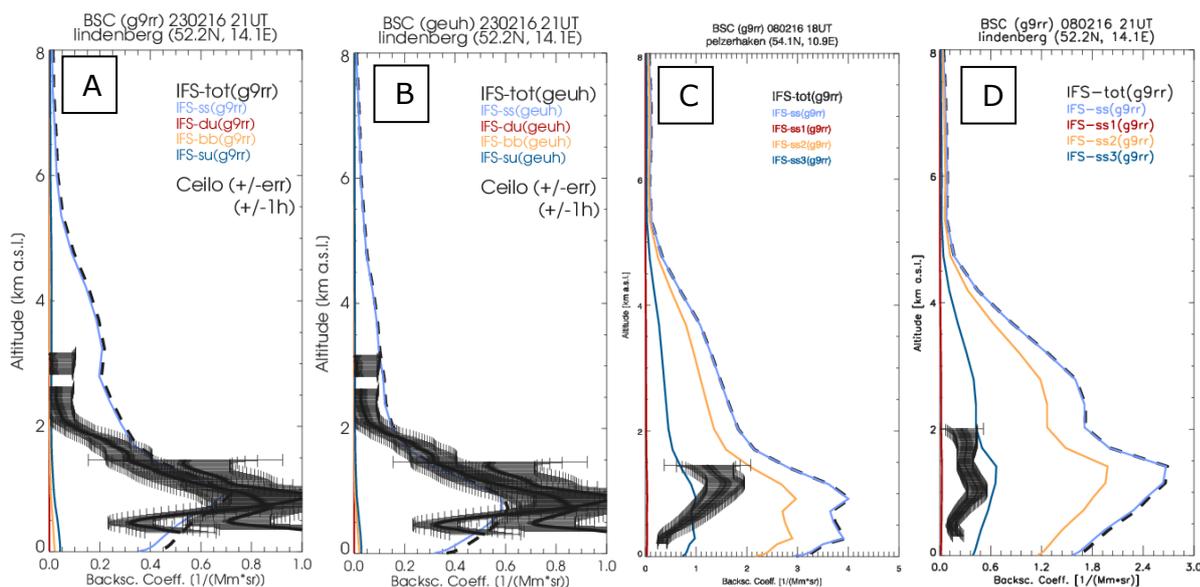


Figure 3.5.18: Bsc profiles at Lindenberg for osuite (A) and control (B) on 23 Feb 2016 splitted for the different aerosol types. Note that here the largest contribution to total modelled bsc is from sulfur aerosol. C,D: Bsc profiles at Pelzerhaken (C) and Lindenberg (D), for the osuite on 8 Feb 2016, splitted for contributions from the different size bins, ss1 (0.03-0.5 μm - red), ss2 (0.5-5 μm - orange) and ss3 (5-20 μm - blue).

efficiency. The inland increase of the over-estimation seems somewhat stronger in the o-suite than in the control run (not shown).

In summary, this preliminary evaluation of vertical profiles of aerosol backscatter coefficient 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.

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 February 2016 (please note that towards the end of the validation period fewer soundings are available). As C-IFS-CB05 stratospheric composition products beyond O_3 in the o-suite is not useful we provide only a very limited evaluation of the control experiment. A description of the applied methodologies and a map with the sounding stations can be found in Eskes et al. (2016). Both runs, the o-suite and the control run, show MNMBs mostly within the range -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 Fig. 3.6.1-3.6.2.

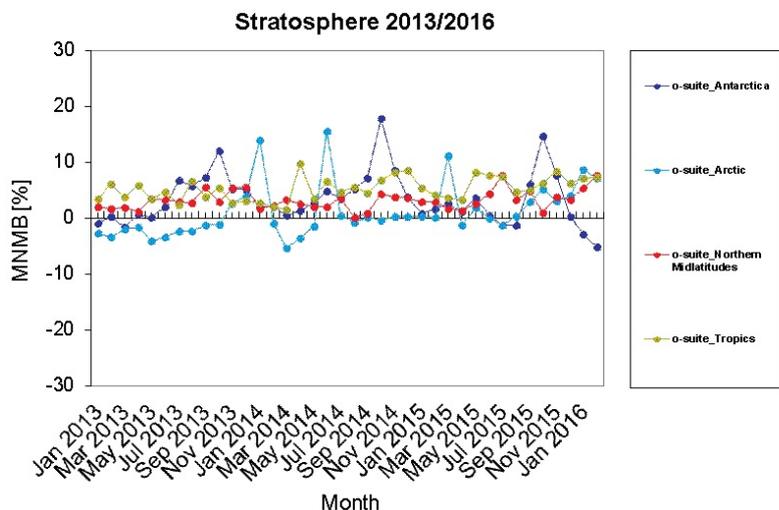


Figure 3.6.1: MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated sonde data in the Arctic (light blue), Antarctic (dark blue) northern midlatitudes (red) and tropics (green).

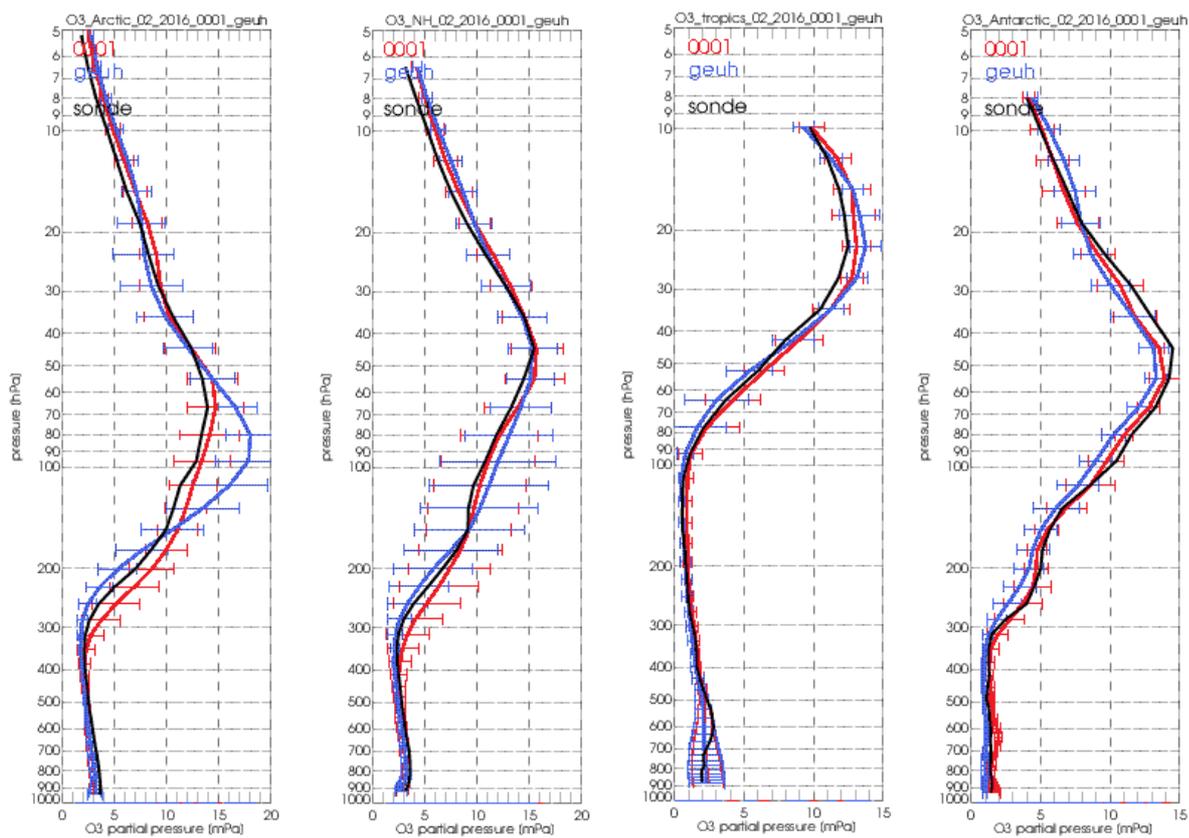


Figure 3.6.2: Comparison between mean O₃ profiles (units: mPa) of o-suite (red), and control (blue) in comparison with observed O₃ sonde profiles (black) for February 2016 for the various latitude bands: Arctic, NH-mid latitudes, Tropics and Antarctic.

O₃ partial pressures in the stratosphere are mostly slightly overestimated in all latitude bands. MNMBs in Antarctica during the ozone hole season, from August to November, remains 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.

Fig. 3.6.2 compares the averaged profiles in each region during DJF 2015-2016. The vertical distribution of stratospheric ozone is quite well represented by the o-suite, but it overestimates ozone in the Arctic and northern mid-latitudes of the lower stratosphere. This issue will be discussed in Section 5.1.

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 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 Eskes et al. (2016) and at <http://nors.aeronomie.be>. We use MWR (microwave) instruments located at Ny Alesund (79°N, 12°E, Arctic station) and Bern (47°N, 7°E, northern midlatitude station) and LIDAR instruments located at Lauder, New Zealand (46°S, 169.7°E, altitude 370m) and Hohenpeissenberg, Germany (47°N, 11°E, altitude 1km).

From Table 3.6.1, the stratospheric partial column bias at Bern during Sept. 2015 – Feb. 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%. This confirms the finding with ozone sondes in the Arctic and will be discussed further in Section 4.1. In JJA, both MWR stations overestimated the upper stratosphere ozone content by a significant amount, i.e. the disagreement was comparable to the measurement uncertainty. In SON the disagreement became an underestimation reaching values up to -30% at Ny Alesund (see also Fig. 3.6.3). At Bern the differences are negligible since Sept 2015 (i.e. they are smaller than the observational uncertainty).

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

Table 3.6.1: Seasonal relative mean bias (MB, %), standard deviation (STD, %) of the partial (stratospheric) 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 %).

		MAM			JJA			SON			DJF		
		MB	stddev	nobs									
o-suite	Ny.Ale	5.76	5.82	109	-2.05	5.33	224	12.84	7.39	273	14.50	6.51	213
	Bern	-4.05	2.17	641	-4.93	2.50	513	-0.54	2.32	687	0.74	3.43	527

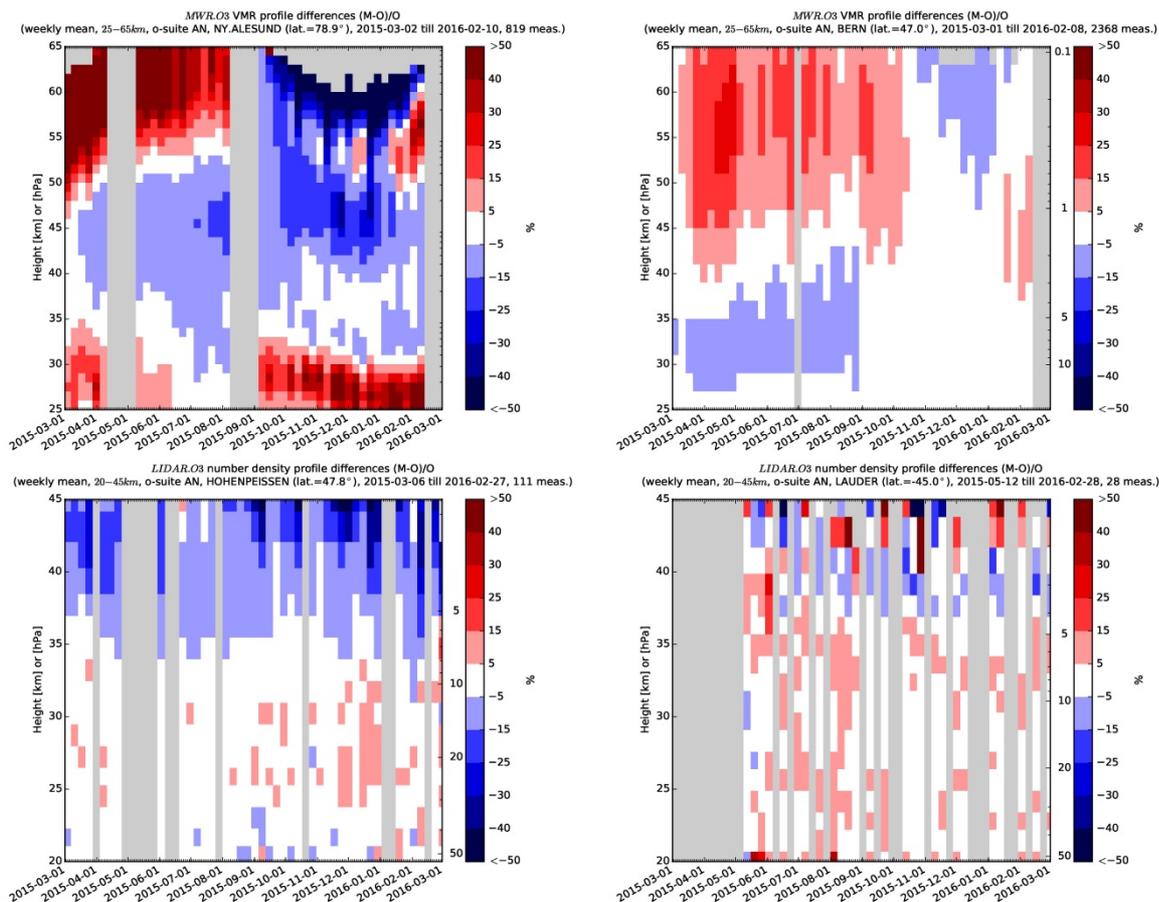


Figure 3.6.3: Normalized density profile weekly mean profile bias for the O₃ mixing ratios of o-suite/O and the NDACC station at Ny Alesund (MWR, upper left), Bern (MWR, upper right), Hohenpeissenberg (lidar, lower left) and Lauder (lidar, lower right). For the LIDAR stations, the measurement uncertainty above 35km is comparable to the observed profile bias.

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. For the first time in these validation reports, the comparisons also include 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.

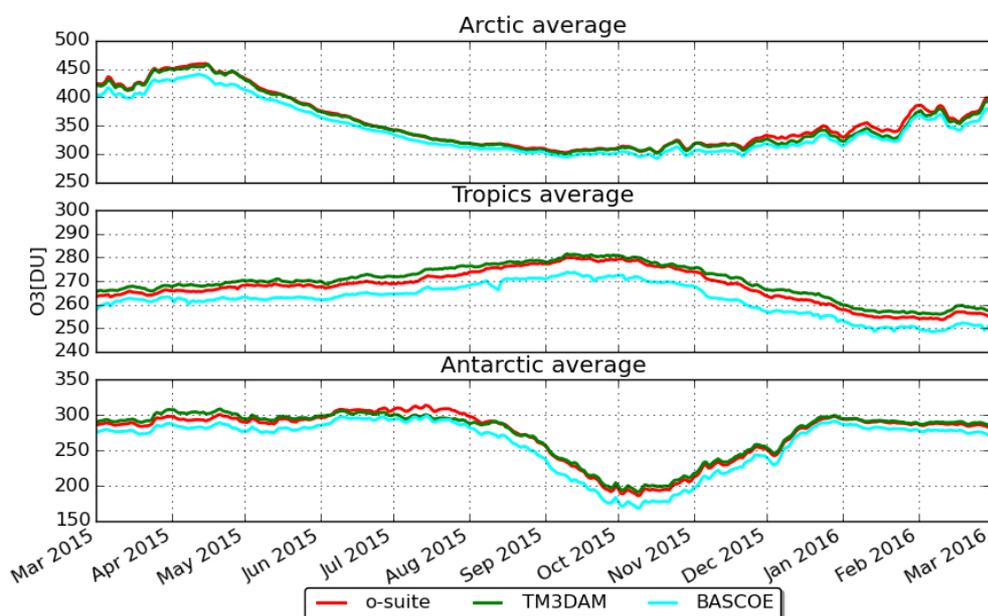


Figure 3.6.4: Zonally averaged ozone total column (Dobson Units) in the Arctic(60°N-90°N), Tropics(30°S-30°N) and Antarctic (90°S-60°S) during the period 2015/03/01-2016/03/01.

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 and regular underestimations w.r.t. TM3DAM of about 2-3 DU, i.e. $\sim 1\%$.
- In the Antarctic, the o-suite matches TM3DAM during the whole period except for the months of March-April 2015 (underestimation of up to 10DU) and 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 3.6.7 shows that in the

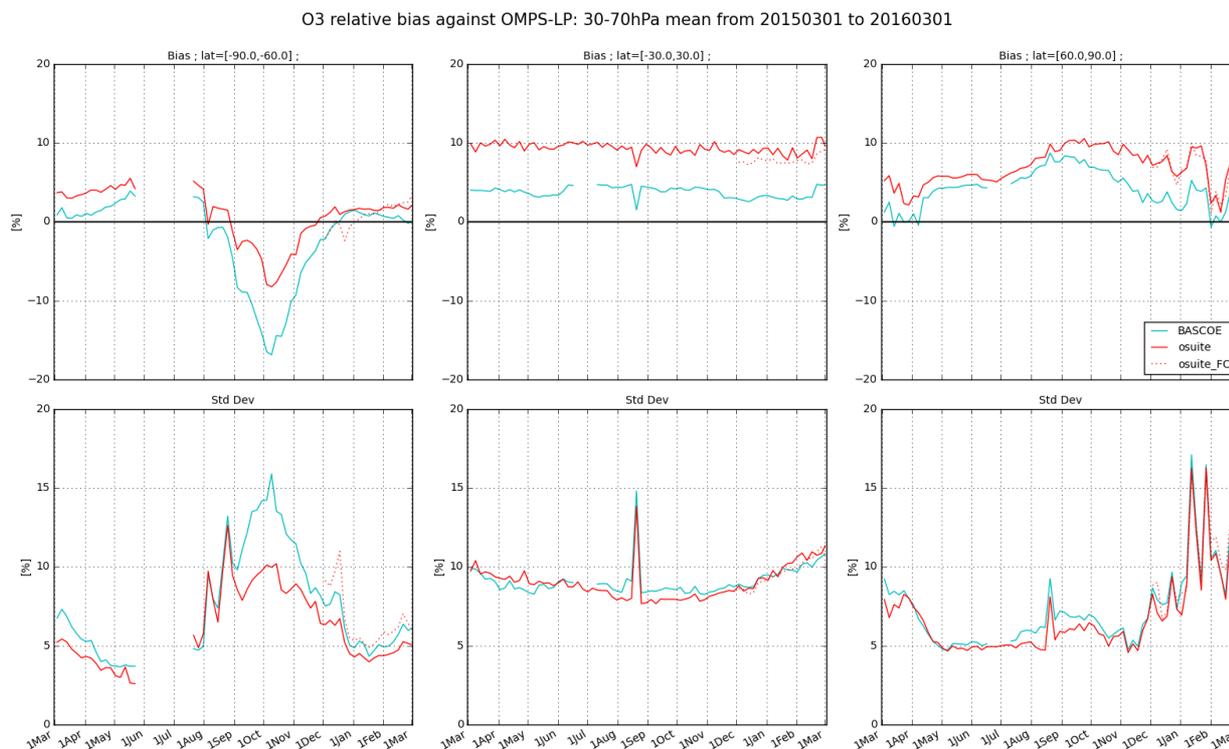


Figure 3.6.5: Time series comparing ozone from o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted), and BASCOE (cyan) with OMPS-LP satellite observations for the period 2015-03-01 to 2016-03-01 in the middle stratosphere (30-70hPa averages): top row, normalized mean bias (model-obs)/obs (%); bottom row, standard deviation of relative differences (%).

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 bottom row of fig. 3.6.5 shows the standard deviation of the differences and can be used to evaluate the random error in the analyses. Hence in the lower stratosphere, the random error of the o-suite is evaluated at 7% to 10% in the Tropics and varies in the polar regions from 5% (summer and fall) to 10% (winter and spring).

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. December 2015-February 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.

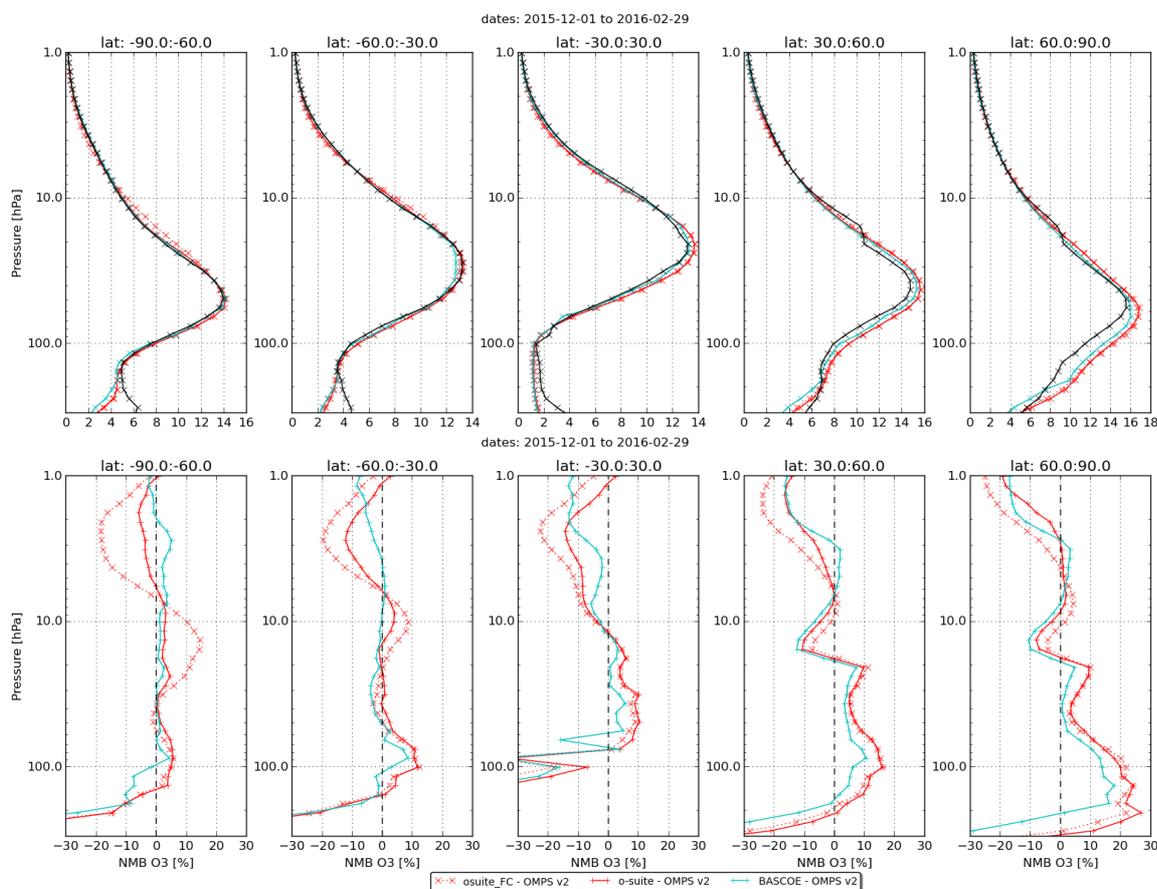


Figure 3.6.6: Mean value expressed in partial pressure (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with OMPS-LP v2 observations for the period December-January-February 2016.

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, its control experiment, and 2 to 4-day forecast results. The significant differences between the observations and the CB05 chemistry runs can be explained by the missing stratospheric chemistry for these model runs. The only constraint on stratospheric NO_x is implicitly made by fixing the HNO₃/O₃ ratio at the 10 hPa level. This assumption, in combination with the changing model settings for stratospheric O₃ for control compared

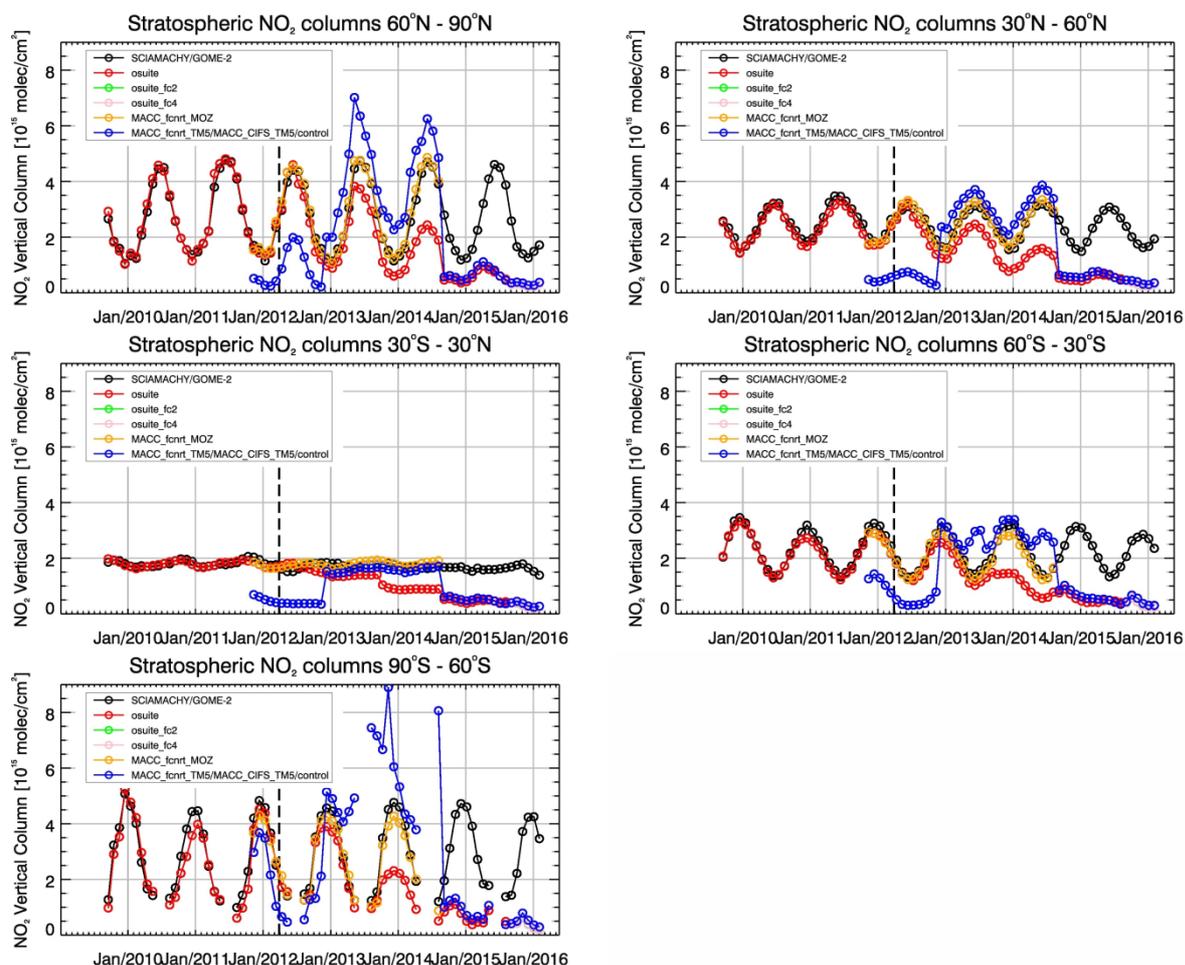


Figure 3.7.1: Time series of average stratospheric NO₂ columns [10^{15} molec cm⁻²] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012) compared to model results for different latitude bands. See text for details. The blue line shows MACC_fcrrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 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.

to MACC_CIFS_TM5, may explain some of the jumps we see in stratospheric NO₂. In any of these runs the stratospheric NO₂ is poorly constrained. It clearly indicates that stratospheric NO₂ 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 NO₂ columns have a systematic low bias relative to those from MACC_fcrrt_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_fcrrt_MOZ, running until September 2014), especially northwards of 30°S. Details on the NO₂ 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 March 2015 to April 2016 based on observations from 15 surface stations. Over this period the high resolution forecast corresponds to three experiments: gcbt, gf39 and ghqy, see also Sec. 2.1.3. Unfortunately two different experiments were used in the 2015-period of the comparison ("gf39" for ICOS, Sec. 4.1 and "gcbt" for TCCON, Sec. 4.2), but the same experiments are used since 1st January 2016 ("gf39" from 01 January 2016 to 29 February 2016 and "ghqy" since 01 March 2016).

4.1 CH₄ and CO₂ validation against ICOS observations

4.1.1 Assessment of the diurnal cycle

We have first compared the diurnal cycles, using hourly means observations (Fig. 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 when looking at 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 53 to 13 ppb when using sampling heights from 5 to 180m above the ground.

For high towers, also located in rural areas in France, the overestimation ranges from 8 to 5 ppb at OHP (10 to 120m agl), and is only ± 1 ppb at OPE (10 to 100m agl). The specific problem at Trainou is probably related to a local source located in the grid box of the tower, rather than a problem with the atmospheric transport since the diurnal cycle of CO₂ is properly represented (Figure 4.1.1). The problem could have been due to the city of Orléans, 17 km southwest from the TRN tall tower, but the location of the grid box used in the high resolution experiments for this site is located further northeast (Figure 4.1.3). Stations which also correspond to outliers in the daily scale comparison, probably due to representation of the surface fluxes, are Finokalia (FKL) and La Réunion Island (RUN at high altitude and STD at the sea level). In order to better understand the spatial representativeness of the model, in future we will evaluate the simulated variability at the four closest grid boxes around each monitoring sites, and based on that decide about possible change in the location of the extracted grid box, for the stations showing clear systematic problems. The two tropical stations of Lamto (LTO) and French Guyana (GUY) present similar and interesting characteristics. At both stations the model strongly underestimates the amplitude of the CH₄ diurnal cycle whereas it represents much better the CO₂ cycle (Figure 4.1.1). This could be explained either by a missing source by plants or soils in tropical

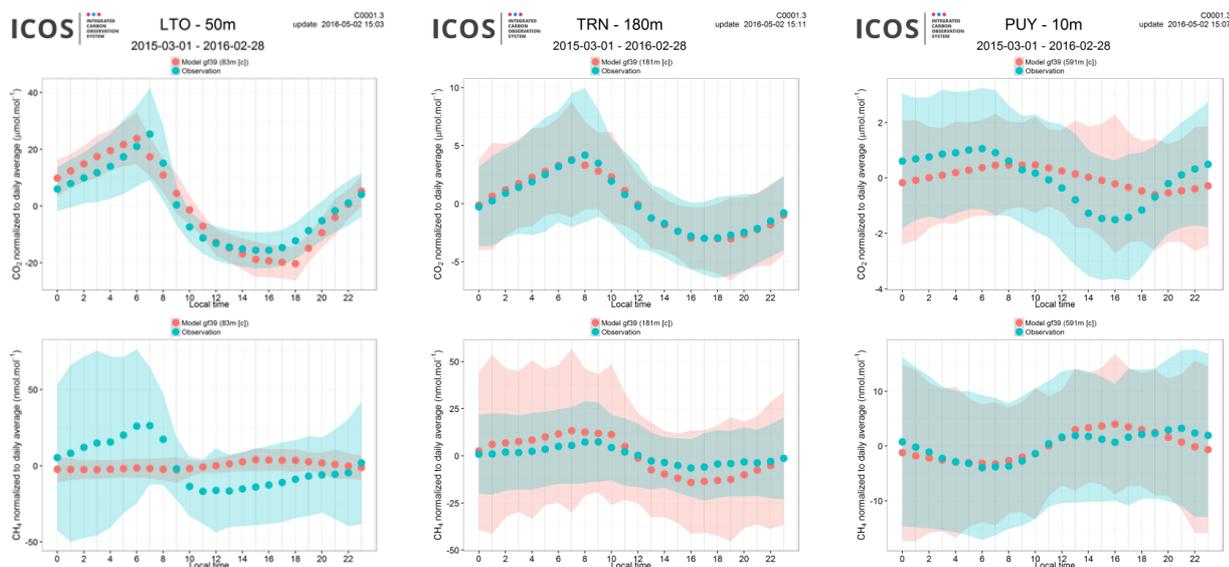


Figure 4.1.1: Mean diurnal cycle averaged over one year (March 2015-March 2016) for CO₂ (above) and CH₄ (below) at three stations from left to right: Lamto (LTO), Trainou tower at the highest level (TRN) and Puy de dome (PUY). Observations are shown in blue and model results in red.

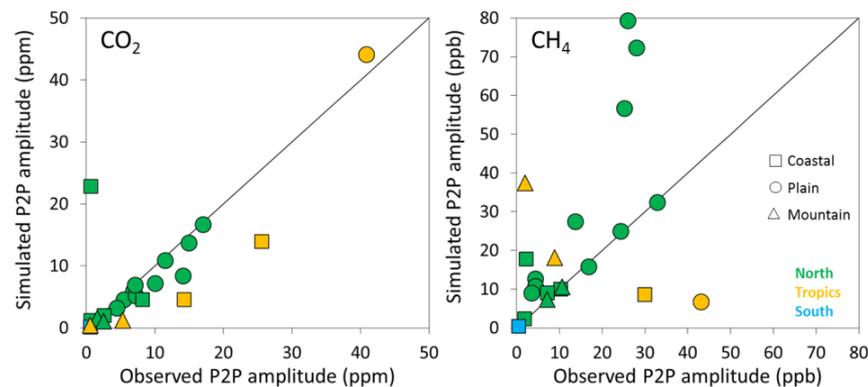


Figure 4.1.2: Comparison of the diurnal peak to peak amplitudes observed at surface stations compared to the model results.

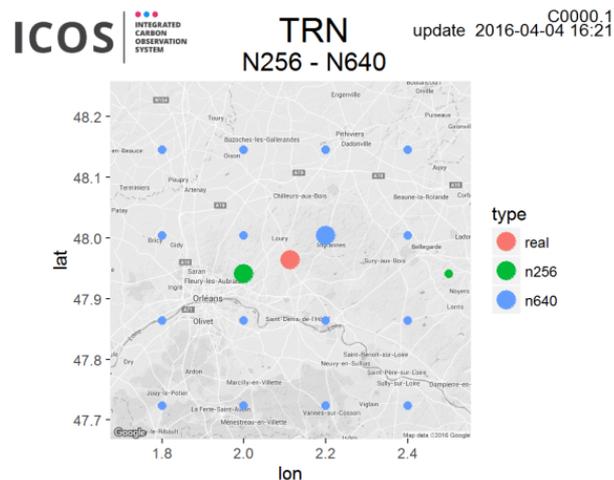


Figure 4.1.3 Location of the Trainou tower (red dot) compared to the location of the grid box representing this station in low (green dot) and high (blue dot) resolutions experiments.

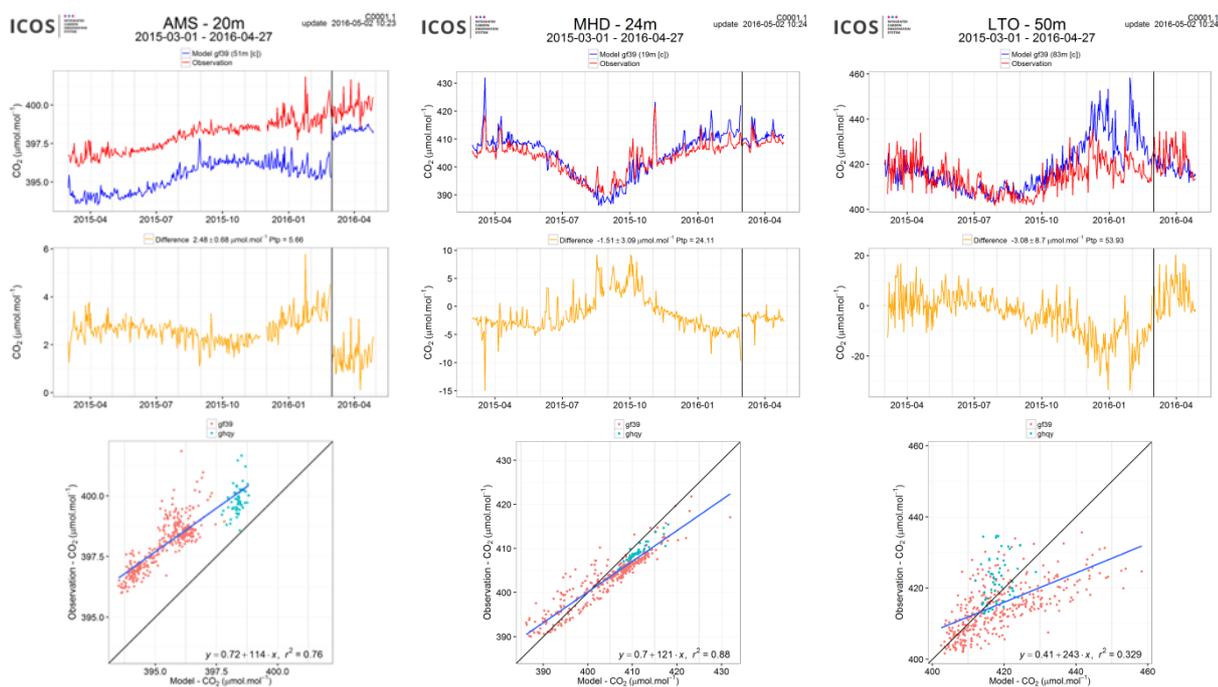


Figure 4.1.4: Above: Comparison of CO₂ daily means observed (red) and simulated (blue) at three stations (Amsterdam I., Mace Head and Lamto). Middle: differences of the observations minus the simulations. Below: Linear fit between observations and simulations. Blue points correspond to the ghqy experiment,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

environments, since these processes are still largely uncertain, or an effect of the CH₄ sink by the atmospheric oxidation with the OH radical.

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 24hrs 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 and CH₄ concentrations at some sites. This latest experiment has not been used in the current analysis. Figures 4.1.4 and 4.1.5 show the CO₂ and CH₄ comparisons respectively at three selected stations in South hemisphere (Amsterdam I.), North hemisphere (Mace Head) and in the tropics (Lamto). Mace Head is representative of the biases observed in the North hemisphere.

The model presents a systematic positive bias from March to July 2015, switching to a negative bias from August to end of October, and back to a positive bias in winter 2015/2016 (Figure 4.1.4 and 4.1.5). This seasonal behaviour is observed both for CO₂ and CH₄ at most stations in North hemisphere. Overall the mean annual bias at Northern stations is -3.1 ± 3.2 ppm and -0.5 ± 9.7 ppb for CO₂ and CH₄ respectively. When excluding Finokalia, which appear as clear outlier the mean biases are -2.4 ± 1.3 ppm and -0.2 ± 10 ppb. Similarly to the results for the diurnal cycles, CH₄ comparisons appear more scattered than CO₂ (Figure 4.1.6). Focusing on the three tall towers in France the mean CO₂ bias increases from -1.4 to -3 ppm when looking at lower (5 to 10m) to higher (100 to 180m) elevations above the ground. At the tropical site of Lamto, the biases from March to October 2015 are close to 0 ppm and 50ppb for CO₂

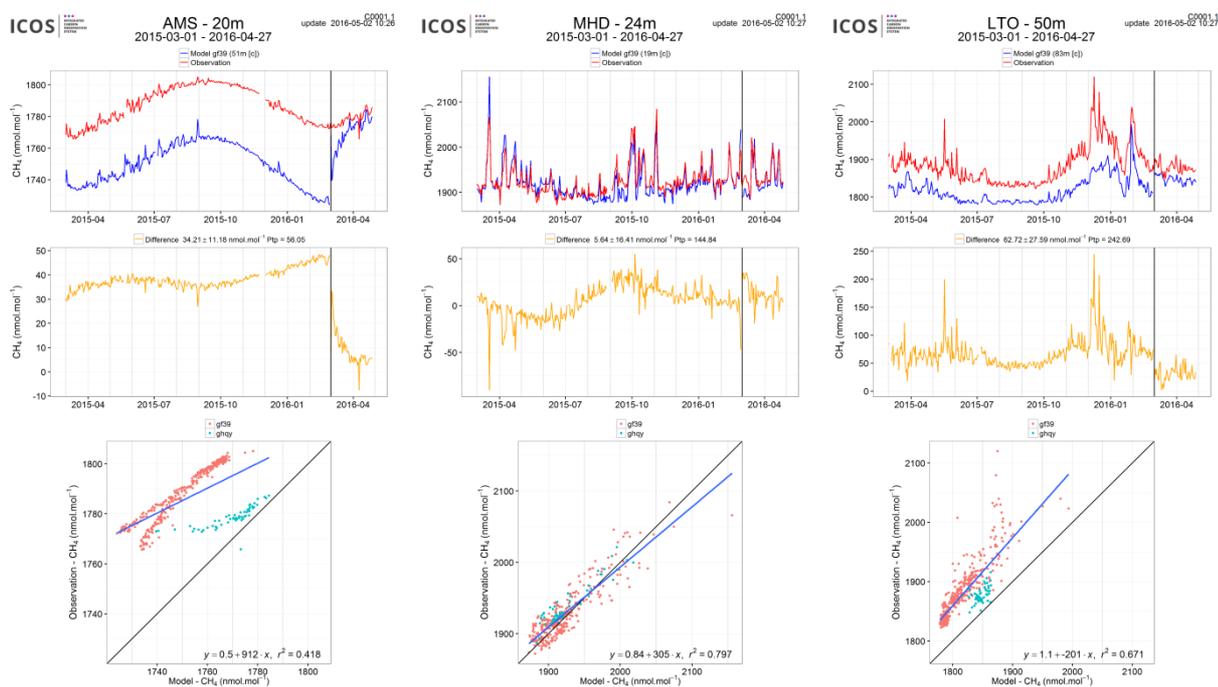


Figure 4.1.5: Same as figure 4.1.4 for CH₄

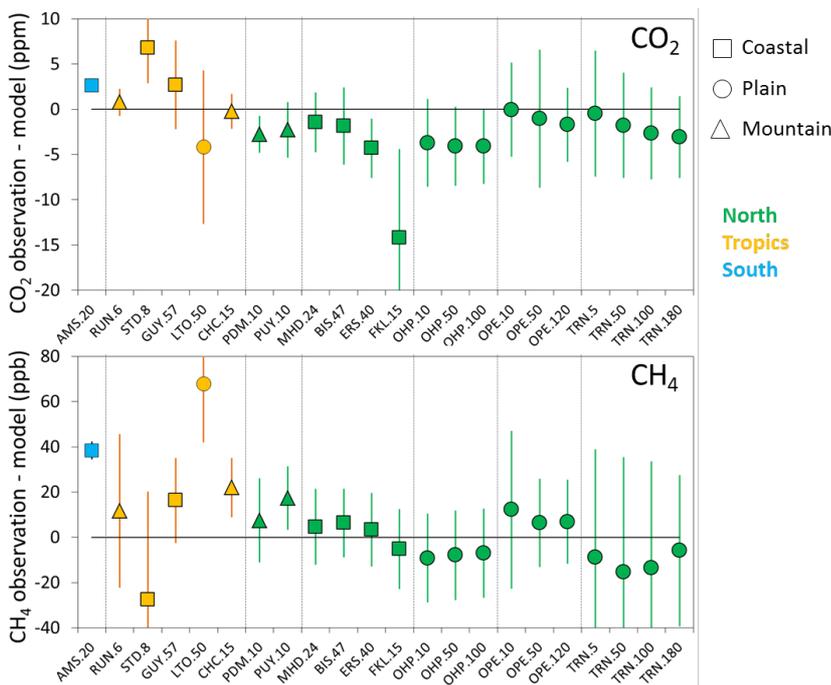


Figure 4.1.6: Mean differences of CO₂ (top) and CH₄ (bottom) concentrations at the 15 sites, calculated from the daily averages over the period March 2015-March 2016. Only the gf39 experiment is considered in this figure.

and CH₄ respectively. In wintertime during the Harmattan period when the station is strongly influenced by biomass burnings in Western Africa, the biases changed significantly with an overestimation of the CO₂ increase by about 20 ppm (Figure 4.1.4), and an underestimation of the CH₄ increase by 50 to 100 ppb (Figure 4.1.5)

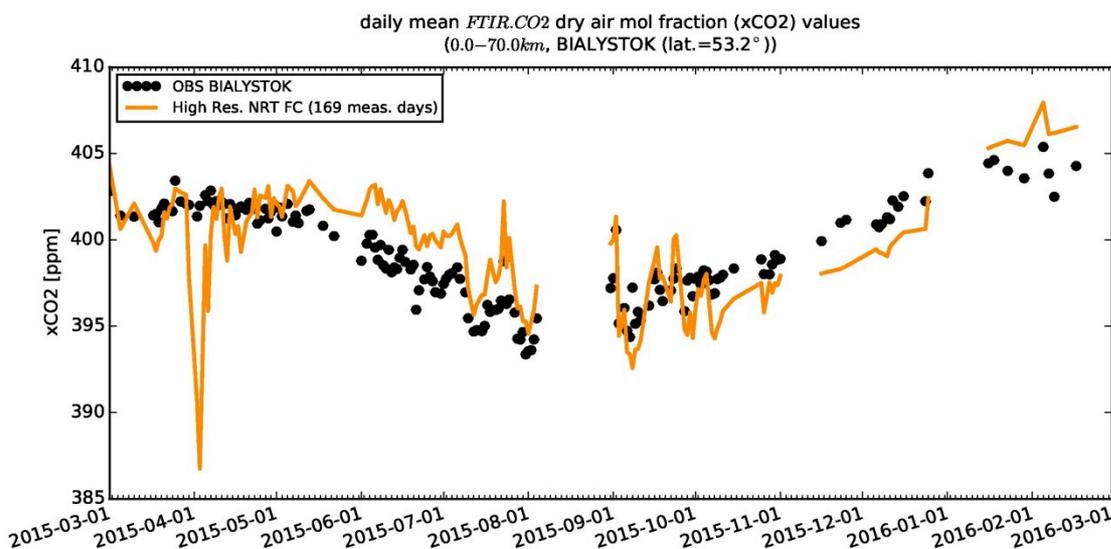


Figure 4.2.1: Time series of column averaged mole fractions of carbon dioxide (CO₂) at the TCCON site Bialystok compared to high resolution NRT FC data (yellow)

4.2 CH₄ and CO₂ validation against TCCON observations

For the validation of the 'gcbt' model version, column averaged mole fractions of CO₂, CH₄ and CO 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 and from the next NRT-report onwards comparisons with RD-TCCON data can be included.

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 – February 2016. At Bialystok (Fig. 4.2.1) the seasonality of the model seems to be roughly one month shifted, resulting in too high modelled values during the growing season (May 2015 – August 2015) and too low modeled values from October 2015 to December 2015. In line with Bialystok the onset of the growing season is about one month too late at Orleans (Fig. 4.2.2). However, in contrast to Bialystok the increase due to respiration is well captured by the model at

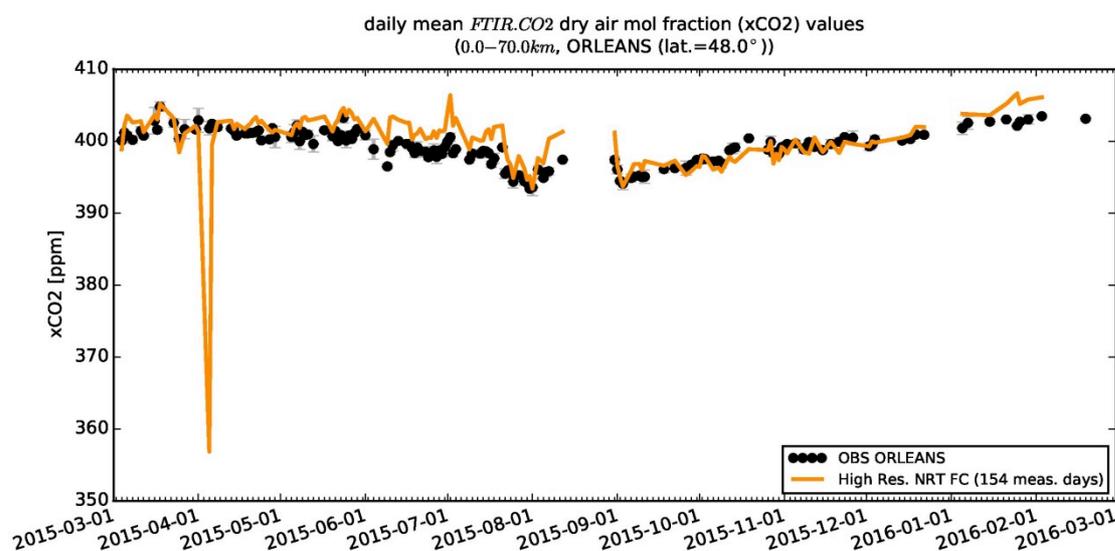


Figure 4.2.2: Time series of column averaged mole fractions of carbon dioxide (CO₂) at the TCCON site Orleans

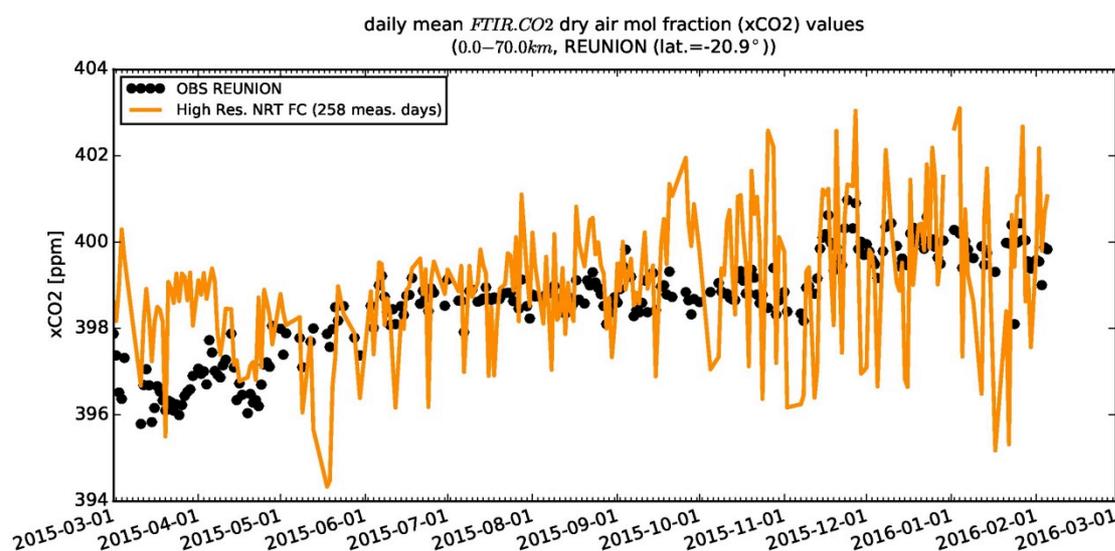


Figure 4.2.3: Time series of column averaged mole fractions of carbon dioxide (CO₂) at the TCCON site Reunion compared to high resolution NRT FC data (yellow)

Orleans. The overall modeled growing season net flux seems to be in agreement with the measurements. However this comparison is limited due to a lack of measurements at Bialystok in August 2015. At both sites the model shows unreasonable low values of in early April 2015.

Between December 2015 and January 2016 the model values are abruptly increasing. At Reunion (Fig. 4.2.3) the model shows strong short-term variations of several ppm, which are not reasonable. Besides these variations the model captures the seasonal variation for the period mid April 2015 to February 2016. Before mid April 2015 the model values are too high.

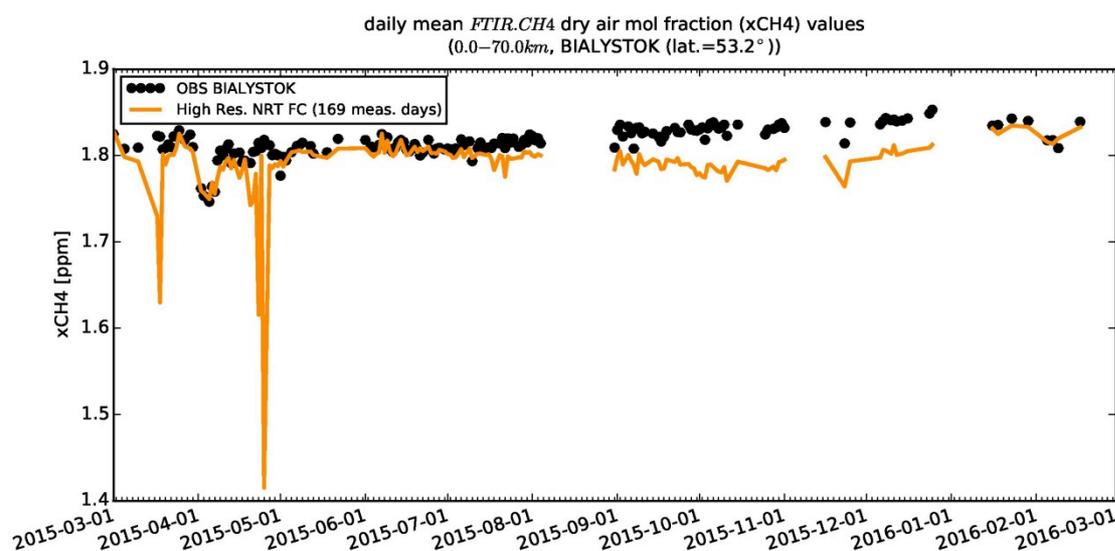


Figure 4.2.4: Time series of column averaged mole fractions of methane (CH_4) at the TCCON site Bialystok compared to high resolution NRT FC data (yellow)

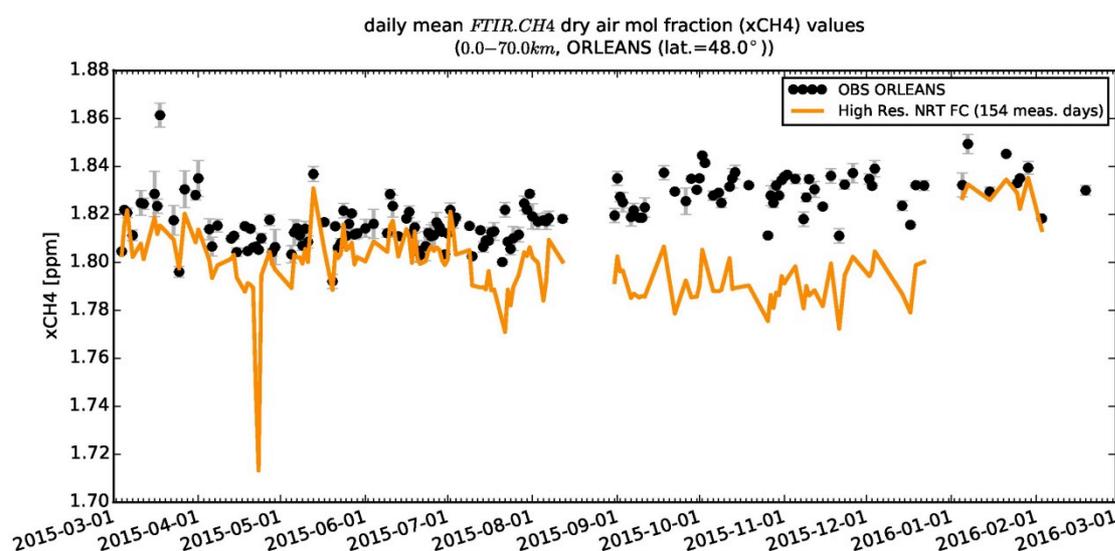


Figure 4.2.5: Time series of column averaged mole fractions of methane (CH_4) at the TCCON site Orleans compared to high resolution NRT FC data (yellow)

4.2.2 Evaluation against TCCON CH_4

At Bialystok the model shows two periods with unreasonable low values, one in March 2015 and the other one end of April 2015 (Fig 4.2.4). The latter period with unreasonable values is also seen at Orleans. Besides these outliers there is a good agreement between measurement and model at Bialystok and Orleans (Fig. 4.2.5) for the period March 2015 – July 2015. From August 2015 – December 2015 the modeled values are too low. Whereas the measurements show the expected seasonal variation, the model has no distinct seasonality. Between December 2015 and January 2016 the model values are abruptly increase to the measured values and are in good agreement with the measurements for January 2016 – March 2016. At Reunion (Fig. 4.2.6) the model is about 0.5% too low for the period March 2015 – mid September 2015. In mid-September 2015 the modeled values are suddenly decreasing resulting in 1-2% too low values for the period mid-September 2015 to February 2016.

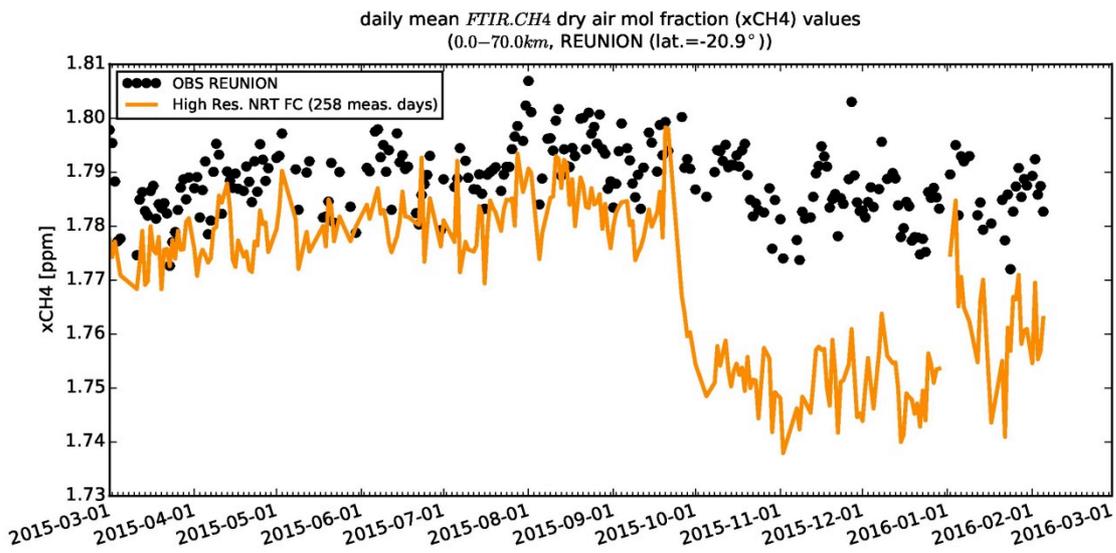


Figure 4.2.6: Time series of column averaged mole fractions of methane (CH₄) at the TCCON site Reunion compared to high resolution NRT FC data (yellow).



5 Events

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

5.1 Ozone depletion above the Arctic and Northern Europe, January-February 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 5.1.1 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 HNO_3 , HCl and $\text{ClO} + 2\text{Cl}_2\text{O}_2$; (based on Aura-MLS observations).

Figure 5.1.2 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 Cl_2O_2 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 5.1.3 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).

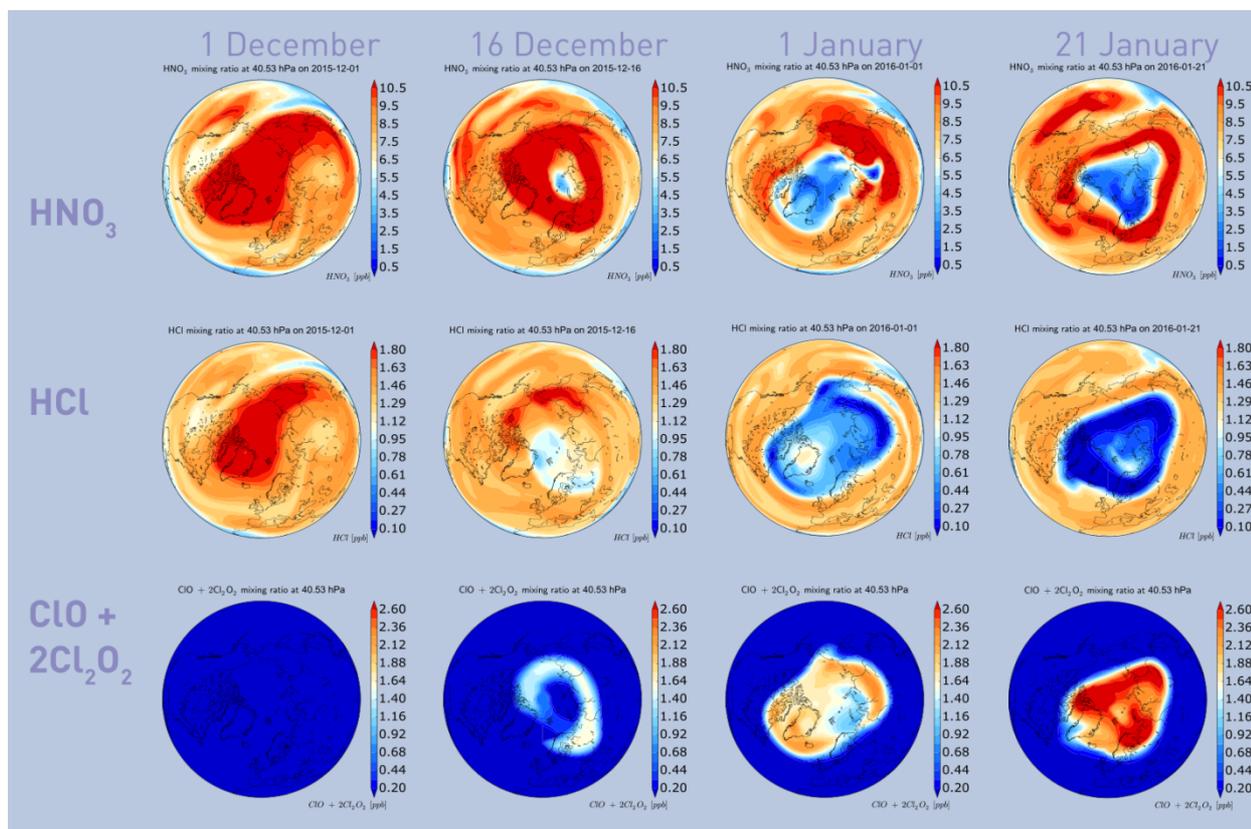


Figure 5.1.1: Results from BASCOE at the level of 40.5 hPa. The upper row shows the mixing ratio of nitric acid, the second row shows the mixing ratio of hydrochloric acid, and the third row shows the sum of chlorine monoxide and its dimer ($\text{ClO} + 2\text{Cl}_2\text{O}_2$). All three rows show 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).

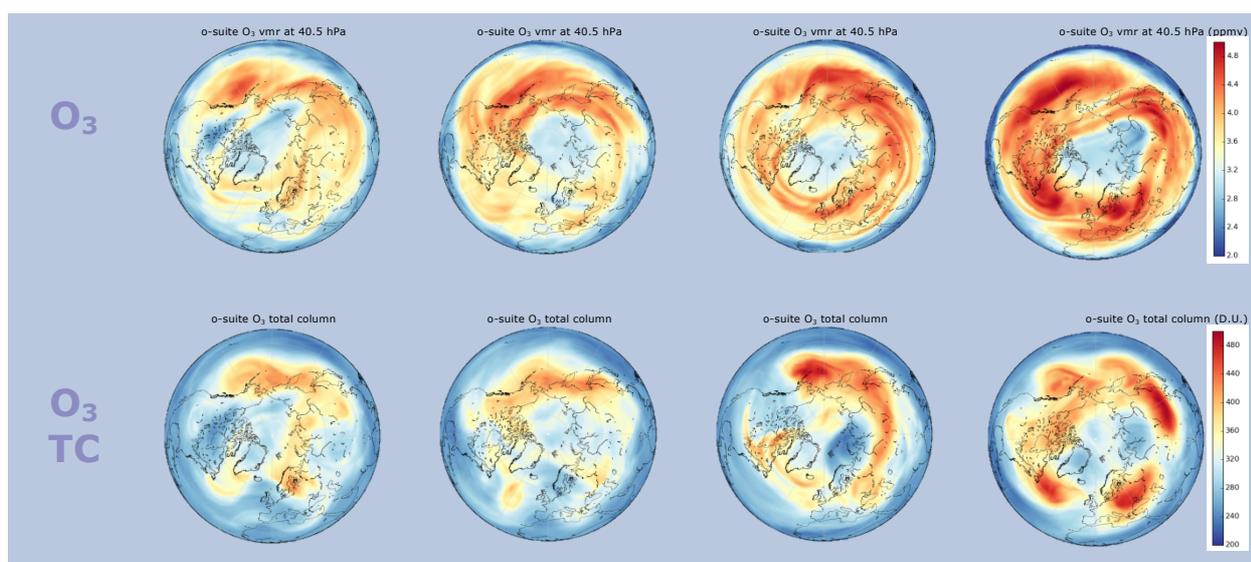


Figure 5.1.2: Results from the CAMS o-suite, on the same dates as the previous figure. The upper row shows the analyses of ozone mixing ratio at 40.5 hPa and the lower row shows the total column of ozone.

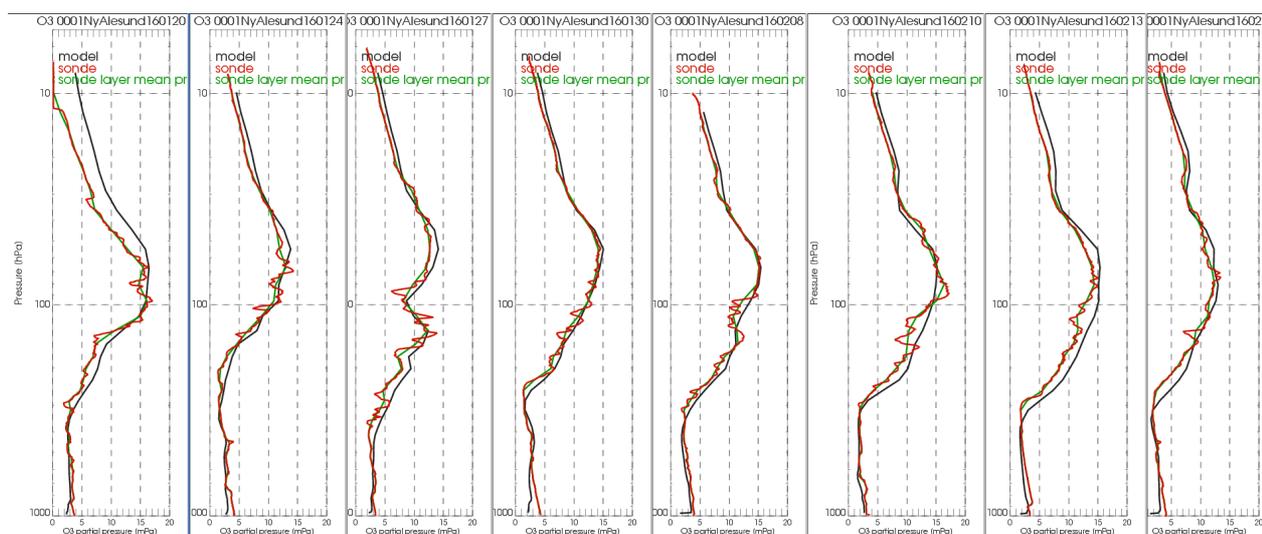


Figure 5.1.3: Eight individual ozone soundings above Ny-Alesund from 2016/01/20 until 2016/02/21. Note the non-standard color scheme: black lines, o-suite analyses; red lines, sonde observations; green lines, sonde observations smoothed to allow quantitative comparison with the o-suite.

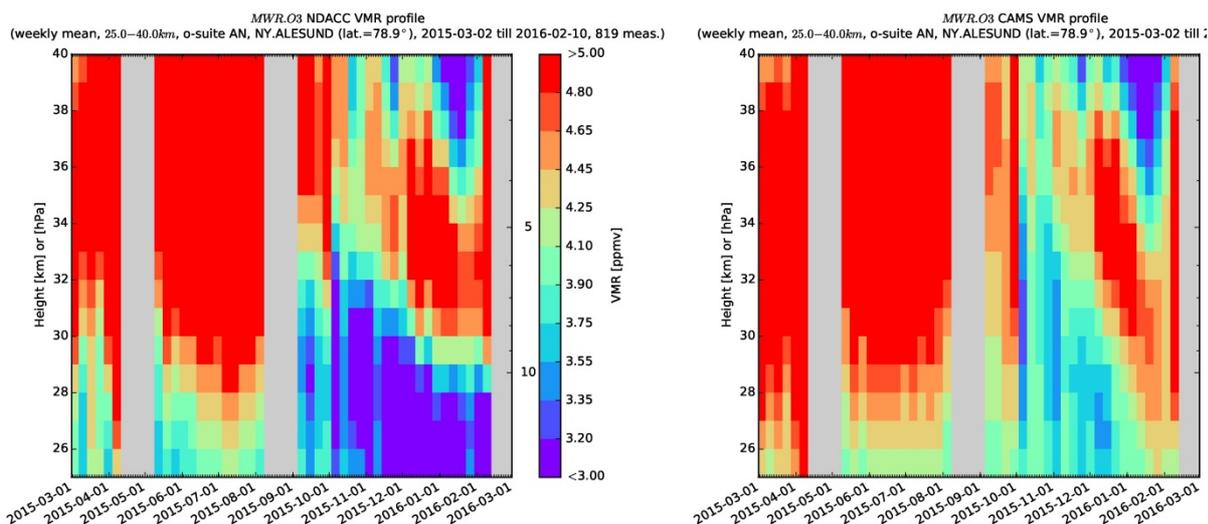


Figure 5.1.4: Ozone volume mixing ratio above Ny-Alesund as a function of time and height: NDACC MWR observations (left) versus o-suite analyses (right)

This underestimation of ozone depletion in the leftmost plot of Figure 5.1.3 (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 5.1.5 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.

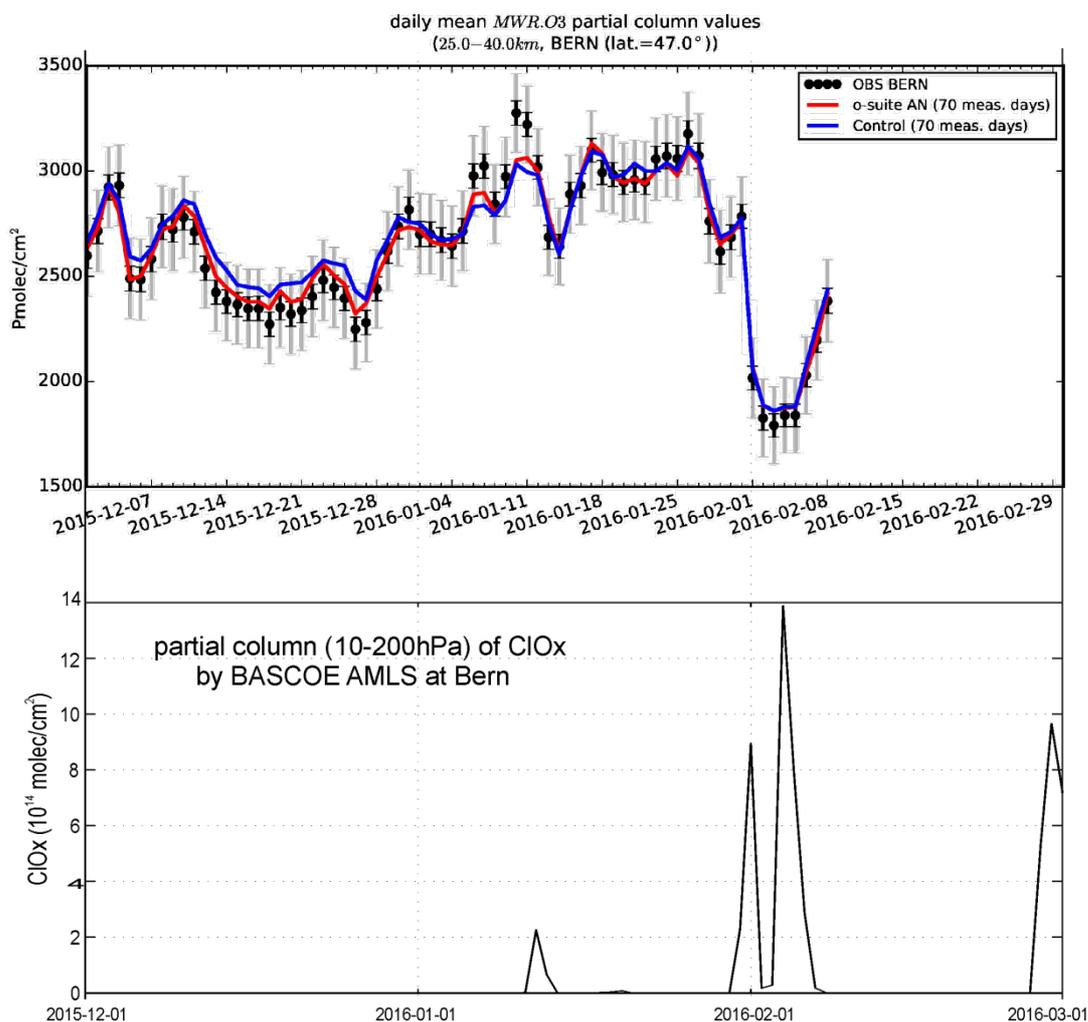


Figure 5.1.5: Time series) above Bern (Switzerland) of the ozone partial column (25-40km; red line shows o-suite analyses, symbols show the MWR observations) and of ClO+2*Cl₂O₂ partial column (10-200 hPa; BASCOE analyses of Aura-MLS observations of ClO).

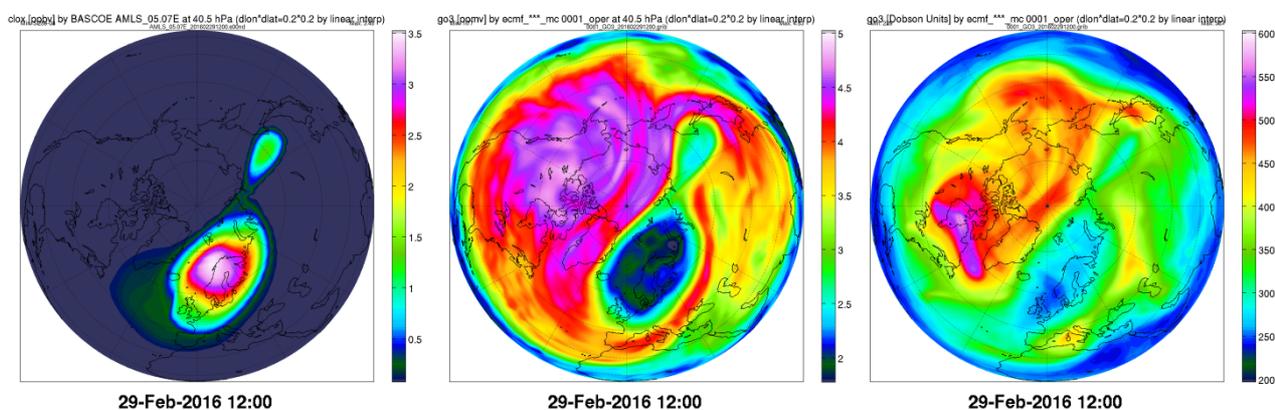


Figure 5.1.6: BASCOE analysis of ClO_x at 40.5 hPa (left), CAMS o-suite analysis of ozone at 40.5 hPa (middle) and CAMS o-suite analysis of ozone total column (right), on 29 February 2016. The ozone depletion event clearly covers Northern Europe.

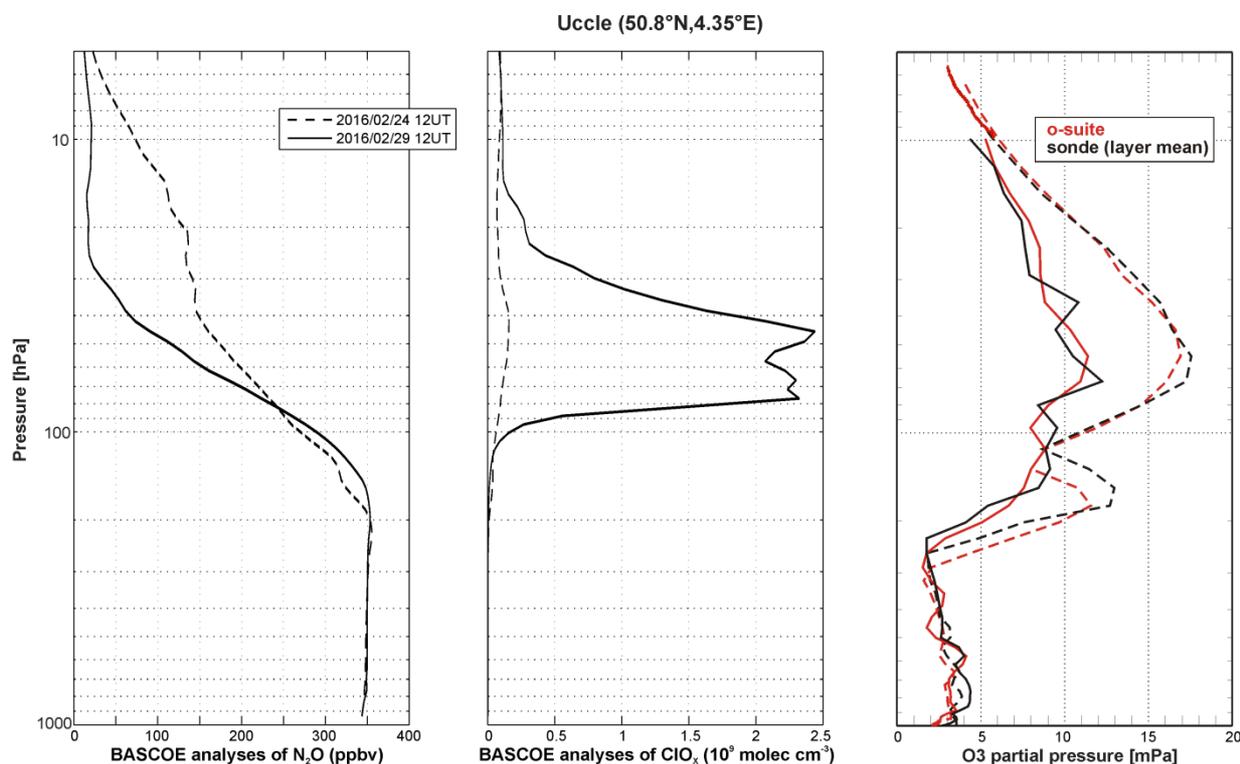


Figure 5.1.7: Vertical profiles above Uccle on 2016/02/24 (dashed lines) and 2016/02/29 (solid lines): BASCOE analyses of N₂O (left plot) and ClO_x (middle plot); ozone partial pressure (right plot) in the o-suite analyses (red lines) and as observed by ozonesondes (black lines)

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 5.1.6). 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 5.1.7). The correlation with vortex air masses can be verified with the BASCOE analyses of the tracer N₂O 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).

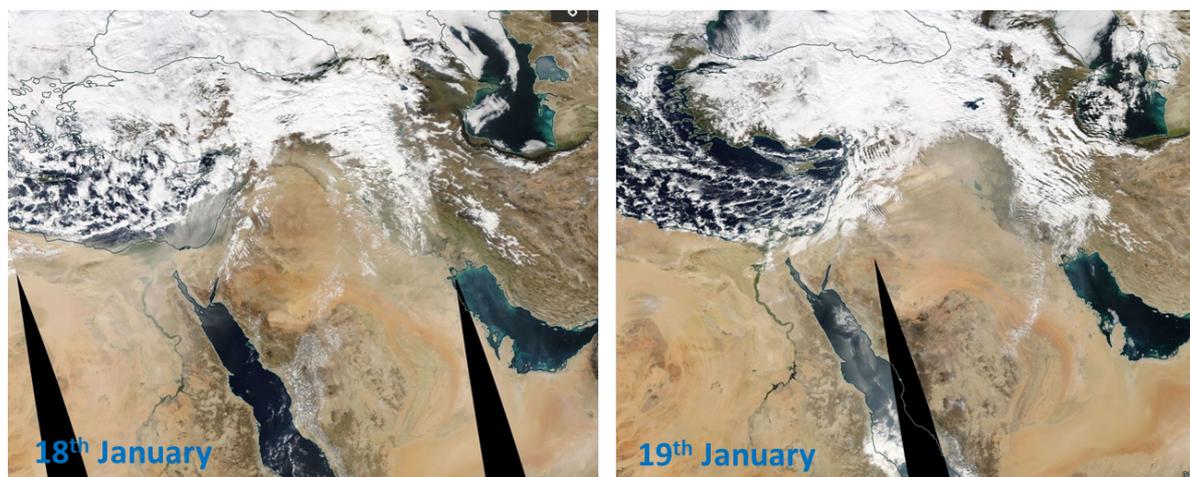


Figure 5.2.1. MODIS-Aqua image, daily composite for 18 and 19 January 2016 (overpass between 9-12 UTC), showing a large dust plume moving in eastward direction over Egypt and other countries in the Middle East.

5.2 A dust event over Eastern Mediterranean-Middle East in January 2016

Around the 17th January, a strong dust plume originated in Libya moved eastwards affecting part of the Eastern Mediterranean and Middle East. Three-hourly dust aerosol optical depth (DOD) from the o-suite has been compared with AOD from MODIS (level 3 gridded product) to evaluate the skill of the o-suite to track the spatio-temporal evolution of the dust plume (Figure 5.2.1). DOD simulated by the o-suite and observed AOD from MODIS from 16th to 20th January 2016 at 12UTC is shown in Figure 5.2.2. Moreover, DOD values from the o-suite, the control and from the Multi-Median model generated from the models participating in the WMO SDS-WAS NAMEE Regional Node (<http://sds-was.aemet.es/>) have been compared with observations from AERONET AOD (Level 1.5) at two stations strategically located along the path of the dust plume over the Cairo EMA (Egypt) and Sede Boker (Israel). Results are shown in Figure 5.2.3.

Both the o-suite and control experiment captured quite well the passage of the dust plume on AERONET stations, as well as the SDS-WAS Multi-Median model. The AOD spatiotemporal evolution given by the o-suite closely resembles the MODIS retrievals. Still the CAMS models underestimate the maximum observed AOD over Libya and Egypt (maximum DOD ~ 1 for o-suite and maximum AOD up to 1.5 for MODIS).

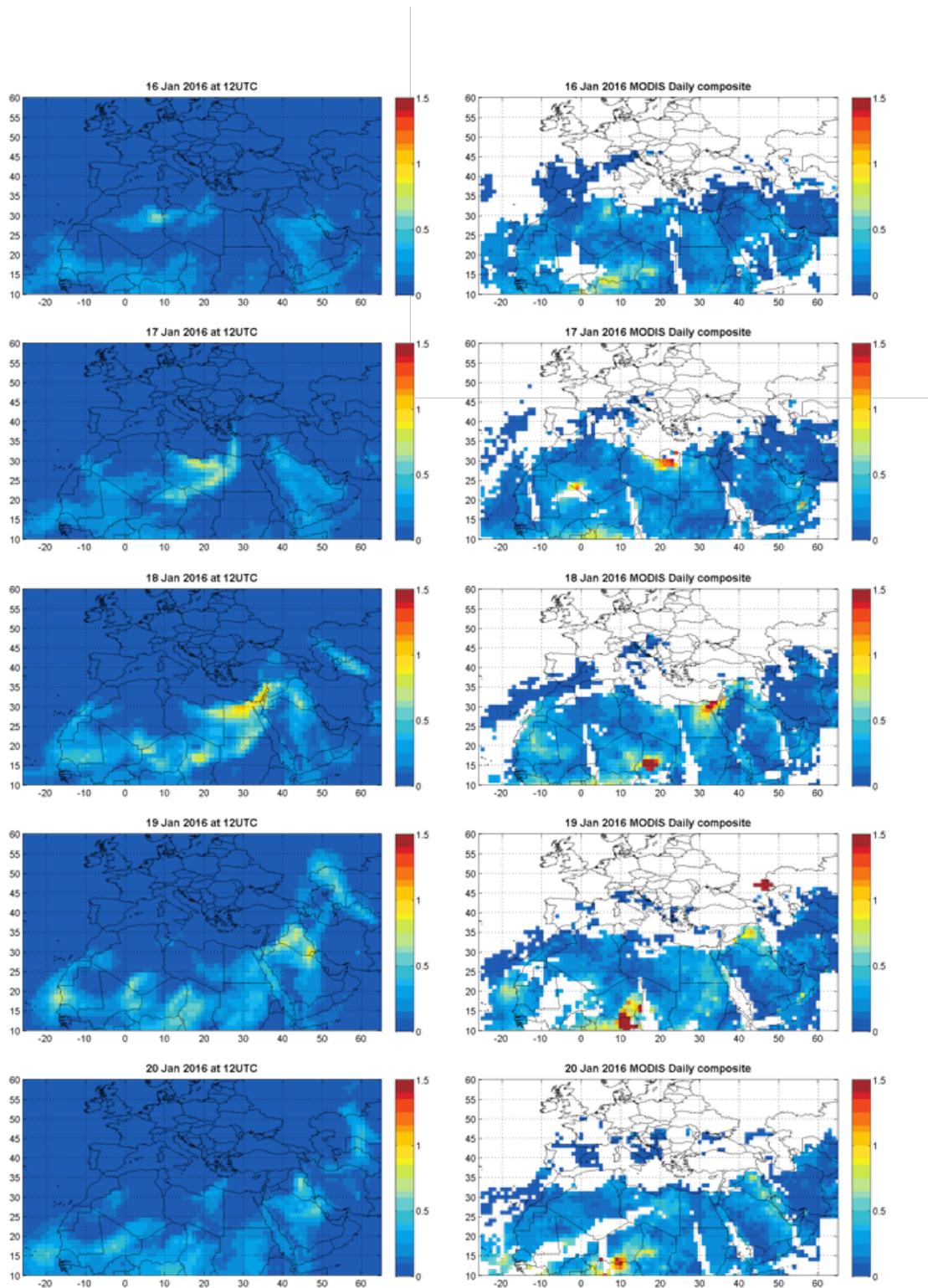


Figure 5.2.2. DOD from o-suite (right column) and AOD from MODIS (left column), for January 16th-20th, 2016 at 12UTC.

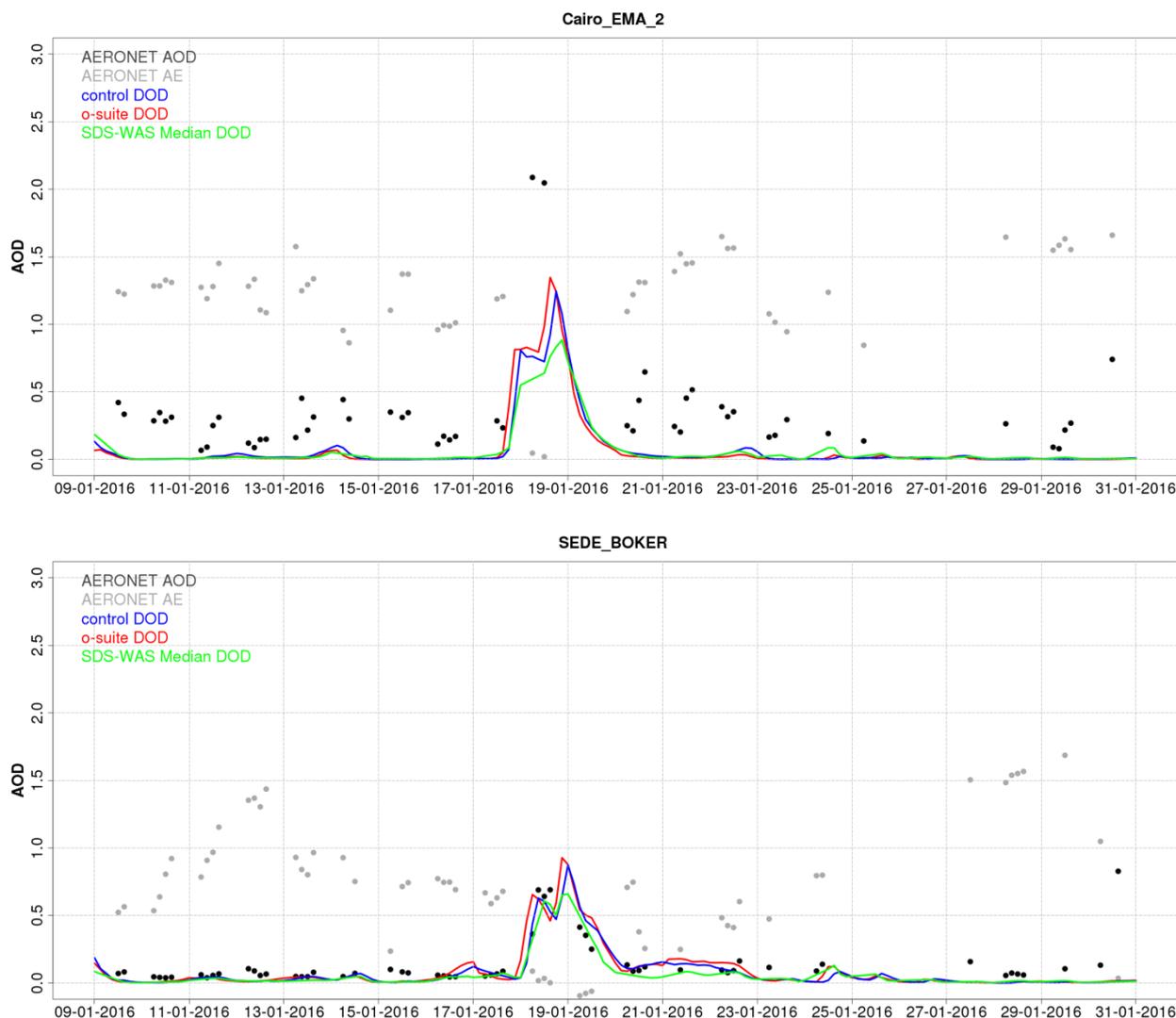


Figure 5.2.3. AOD at 550 nm from AERONET (black), DOD at 550 nm from the o-suite (blue), DOD at 550 nm from the control run (red), DOD at 550 nm from SDS-WAS Multi-model Median (green) at Cairo EMA and Sede Boker AERONET sites during the case analysis from 9th to 31th January 2016.

5.3 A false CO plume over North America, end of February 2016

On February 26 both model runs indicated a plume with high CO values over Western part of US (Fig. 5.3.1). In the following days the air parcels with high CO concentrations spread to the east, over the American continent. As can be seen from the upper row of the Figure 5.3.1, IASI retrievals does not show this high CO event. Also no fire events were found in this area for this time period. It is known that there is no data available from the MOPITT and MODIS from 18 February till 2 March 2016 due to an operational issue with the Terra satellite (<https://www2.acom.ucar.edu/mopitt/news>, see also Table 2.3 in Sec 2.2). The large modelled CO plume are an artifact of the erroneous GFAS fire emissions, caused by inaccurate TERRA-MODIS observations between 26 February and 1 March.

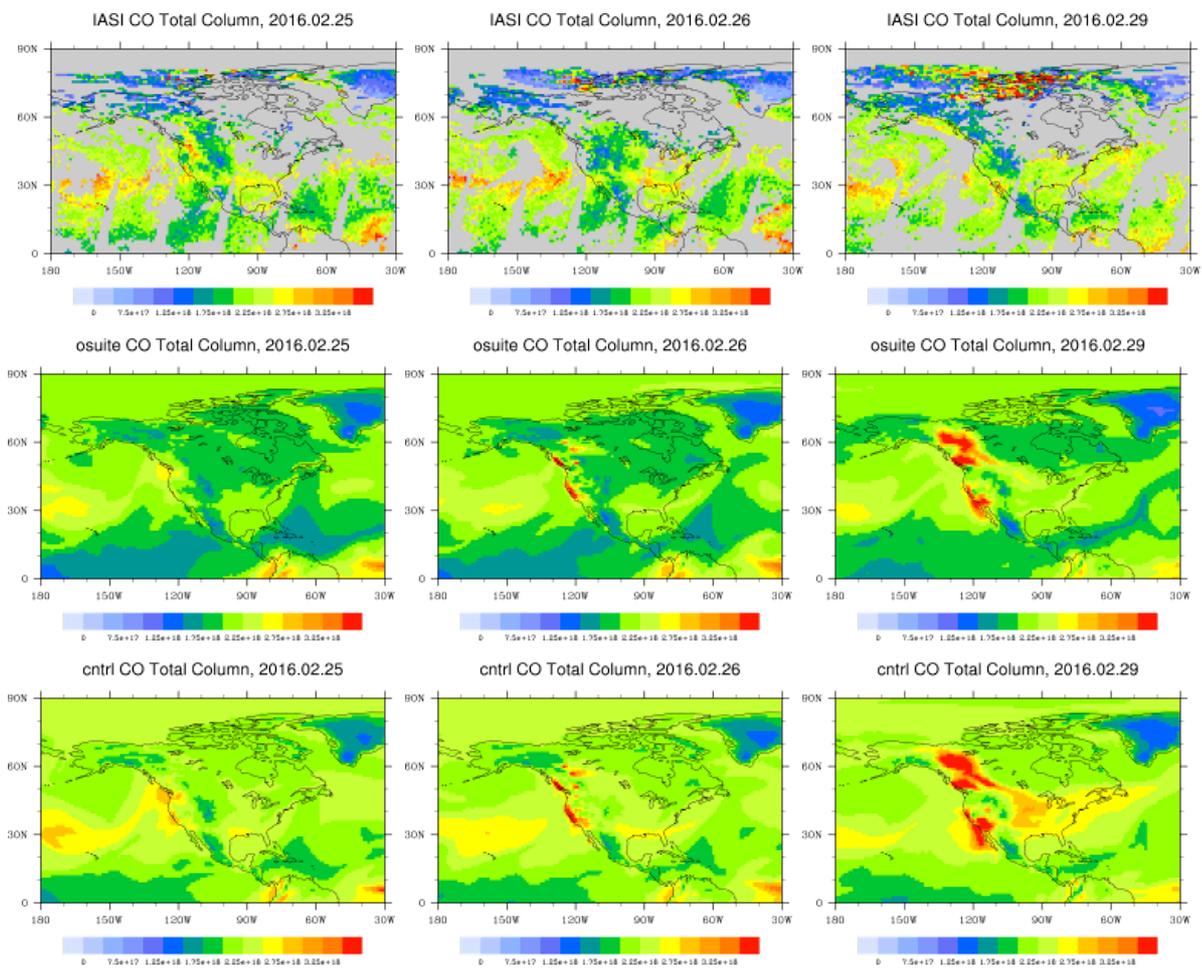


Fig. 5.3.1: CO total column from IASI (top), o-suite (middle) and control runs (bottom) for 25, 26, and 29 February 2016 over selected region.



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 Chabrilat, 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, MPI-M (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)
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)

Stratospheric reactive gases

Simon Chabrillat, BIRA (editor)
Yves Christophe, BIRA (model intercomparisons)
Bavo Langerock, BIRA (NDACC FTIR, MWR, UVVIS DOAS, LIDAR)
Andreas Richter and Anne Blechschmidt, IUB Bremen (SCIAMACHY 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 by ESA through DLR/DFD.

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.

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