ECMWF COPERNICUS REPORT

Copernicus Atmosphere Monitoring Service

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

September - November 2018

Issued by: KNMI

Date: 12 March 2019

Ref: CAMS84_2018SC1_D1.1.1_SON2018_v1

IMPLEMENTED BY ECMWF

Validation report of the CAMS near-real-time global atmospheric composition service. Period September - November 2018

EDITORS:
A. Wagner (MPG), M. Schulz (MetNo), Y. Christophe (BIRA-IASB), M. Ramonet (LSCE), H.J. Eskes (KNMI)

AUTHORS:
S. Basart (BSC), A. Benedictow (MetNo), Y. Bennouna (CNRS-LA), A.-M. Blechschmidt (IUP-UB), S. Chabrillat (BIRA-IASB), H. Clark (CNRS-LA), E. Cuevas (AEMET), H. Flentje (DWD), K. M. Hansen (AU), U. Im (AU), J. Kapsomenakis (AA), B. Langerock (BIRA-IASB), A. Richter (IUP-UB), N. Sudarchikova (MPG), V. Thouret (CNRS-LA), T. Warneke (UBC), C. Zerefos (AA)

REPORT OF THE COPERNICUS ATMOSPHERE MONITORING SERVICE, VALIDATION SUBPROJECT.

AVAILABLE AT:
http://atmosphere.copernicus.eu/quarterly_validation_reports

CITATION:
Wagner, A., M. Schulz, Y. Christophe, M. Ramonet, H.J. Eskes, S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, H. Clark, E. Cuevas, H. Flentje, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, A. Richter, N. Sudarchikova, V. Thouret, T. Warneke, C. Zerefos, Validation report of the CAMS near-real-time global atmospheric composition service: Period September-November 2018, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC1_D1.1.1_SON2018_v1.pdf, March 2019, doi:10.24380/dg9c-pm41.

STATUS:
Version 1, Final

DATE:
12 March 2019

Executive Summary

The Copernicus Atmosphere Monitoring Service (http://atmosphere.copernicus.eu, CAMS) is a component of the European Earth Observation programme Copernicus. The CAMS global near-real time (NRT) service provides daily analyses and forecasts of reactive trace gases, greenhouse gases and aerosol concentrations. This document presents the validation statistics and system evolution of the CAMS NRT service for the period up to 1 December 2018. Updates of this document appear every 3 months.

This summary is split according to service themes as introduced on the CAMS website: air quality & atmospheric composition, climate forcing, ozone layer and UV. Specific attention is given to the ability of the CAMS system to capture recent events. We focus on the 'o-suite' composition fields, which are the daily analyses and forecasts produced by the IFS (Integrated Forecast System) modelling system at ECMWF, using the available meteorological and atmospheric composition observations which are ingested in the ECMWF 4D-Var assimilation system. The model and assimilation configuration is summarised in section 2. 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 delayed-mode analyses and high-resolution forecasts of CO_2 and CH_4 are assessed in this report.

The o-suite data delivery for the period September-November 2018 (SON-2018) was excellent, with an availability of 100% at 10 and 22 UTC (two forecasts per day).

Air quality and atmospheric composition

Tropospheric ozone (O_3)

CAMS o-suite ozone is validated with surface and free tropospheric ozone observations from the GAW and ESRL networks, IAGOS airborne data and ozone sondes. For free tropospheric ozone against ozone sondes the o-suite modified normalized mean biases (MNMBs) are on average small, ±10% over the Northern Hemisphere (NH), between ±30% for stations in the Tropics, and between ±15% for the Arctic in more recent years (Fig. S.1). Over Antarctica o-suite biases are observed between 0% and +20% for recent years, whereas the control run shows negative biases. For SON 2018 good agreement is found over the NH mid latitudes in the free troposphere, which is confirmed with IAGOS evaluations over Frankfurt. Ozone is well represented throughout the profiles until the UTLS region when ozone is generally overestimated by the o-suite and overestimated or underestimated by control run.

In comparison with surface observations we find a steady improvement of the o-suite over the past 5 years over European GAW stations. Biases are generally around ±10%, and within ±20% (the Arctic is discussed below). The o-suite slightly overestimates surface ozone for Europe and the US during September to November 2018 with MNMBs up to 12%. Both runs overestimate O_3 observations for Asia with MNMBs up to 30%. For the tropics, surface ozone is overestimated with MNMBs within 25%. For Antarctica negative biases are observed for the control run, which is mostly efficiently

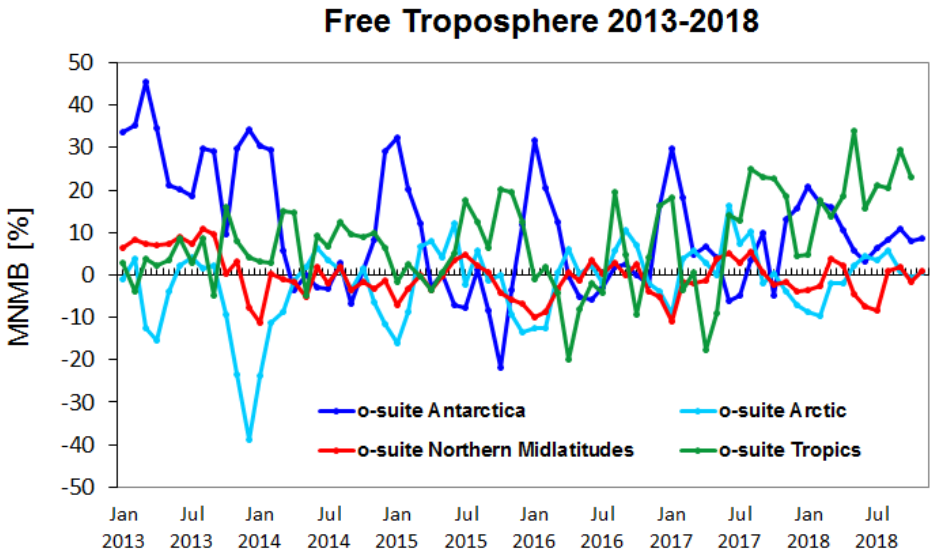


Figure S.1: Time series of MNMB of ozone in the o-suite, compared against ozone sondes, averaged over different latitude bands. The free troposphere is defined here as the layer between 750 and 300 hPa.

corrected by the assimilation in the o-suite (MNMBS < 10% for South Pole station). For Neumayer station, however, the o-suite slightly overestimates O_3 mixing rations (MNMBS 25%).

Tropospheric Nitrogen dioxide (NO_2)

Model validation, with respect to SCIAMACHY/Envisat NO_2 data before April 2012 and GOME-2/MetOp-A NO_2 data afterwards, shows that tropospheric NO_2 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 more pronounced than in the satellite retrievals. Tropospheric NO_2 columns over some local emission hotspots (e.g. Moscow, and Red Basin in China) are overestimated, while wintertime and springtime values over Europe around Benelux are underestimated. 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. S.2), as observed columns of NO_2 decreased recently, likely associated with reduced emissions, and (in contrast to the observations) simulated values show an increase over the whole timeseries available. Mainly in summer and autumn the models regularly show an overestimation over several regions with fire activity (e.g. Western Australia).

Tropospheric Carbon Monoxide (CO)

Model validation with respect to GAW network surface observations, IAGOS airborne data, FTIR observations (NDACC and TCCON) and MOPITT / IASI satellite retrievals reveals that the absolute values, latitude dependence and seasonality, as well as day-to-day variability of CO can be reproduced well by the CAMS-global analyses and forecasts. Biases are between 0% and -30% for European stations, and between -10 and -24% in Asia. A similar bias in Europe in the lower layers is observed by IAGOS. For stations in the southern hemisphere the comparison with NDACC (Fig. S.3) and GAW measurements shows that data assimilation reduces the large positive MNMBs in the control run.

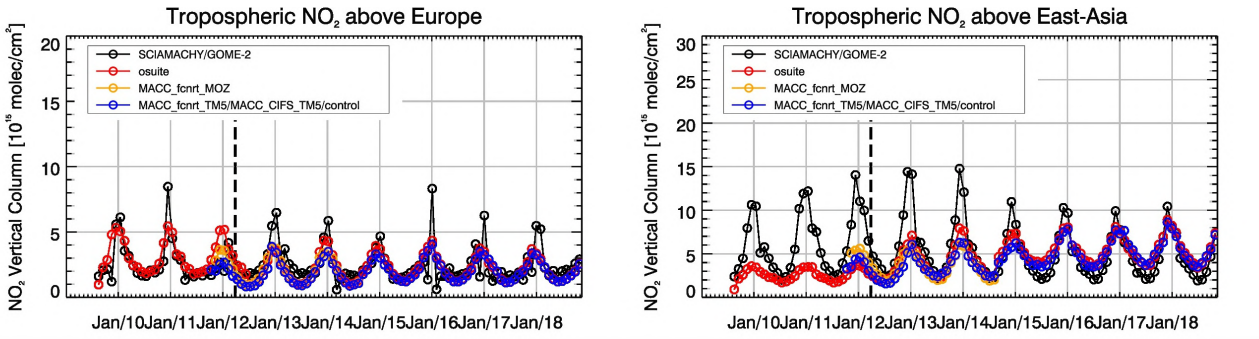


Figure S.2: Time series of tropospheric NO_2 columns from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for Europe and East-Asia. The o-suite is in red, control is in blue (the model run without data assimilation is termed control since Sep 2014).

Compared to IAGOS aircraft data over Frankfurt, the models generally underestimate CO from the surface to the upper troposphere. The best agreement between models and observations is found in the free and upper troposphere. Over North American airports similarly to Europe, CO is usually underestimated in the surface and boundary layer by the two runs, while there is a good agreement in the free troposphere.

Over Northern East Asia the profiles with complex shapes are in general well reproduced by the models with a usually better performance of the o-suite than of control run. In the surface layer mixing ratios are overestimated for observed values higher than 300 ppbv. Over Southern East Asia, surface values are largely overestimated by the models at both locations (Singapore and Surabaya), while the agreement in the boundary layer is better and CO is underestimated.

Formaldehyde

Model validation, with respect to SCIAMACHY/Envisat HCHO data before April 2012 and GOME-2/MetOp-A HCHO data afterwards (Fig. S.4), 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, an overestimation of values regularly occurs over Australia and Central Africa, which could be both related to biogenic emissions or fire emissions. 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.

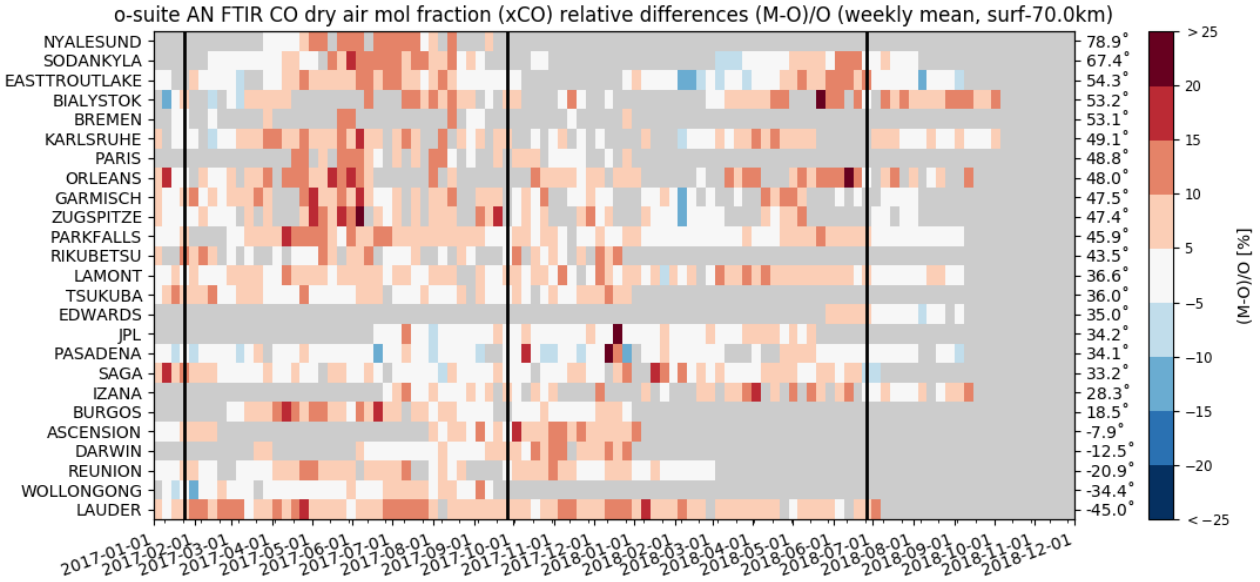


Figure S.3: Monthly mean relative CO bias (o-suite – observation)/observation for the last 2 years. Comparisons are made against FTIR CO remote sensing observations. Model upgrades are indicated with the black lines. The stations are sorted with decreasing latitude (northern to southern hemisphere). A similar plot for the control run shows biases up to 40% in the southern hemisphere.

Aerosol

We estimate that the o-suite aerosol optical depth showed an average positive bias in the latest three months of +17%, measured as modified normalized mean bias against daily Aeronet (V3 level 1.5) sun photometer data. The +3 day forecasted aerosol distribution shows 14% less aerosol optical depth (AOD) than that from the initial forecast day, as shown in Fig. S.5-a. The spatio-temporal correlation, shown in Fig. S.5-b, shows month-to-month variation in SON 2018 similar to autumn 2017, indicating the simulation reproduces approximately 60% of the day to day AOD variability across all Aeronet stations. We find a high AOD bias in Southern Latitudes. 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 o-suite forecast running each day at 12UTC shows almost identical performance as the forecast starting at 00UTC.

The AOD performance of the o-suite with respect to the AERONET data exhibits no pronounced seasonal cycle but somewhat less correlation in late summer. Since October 2017, global AOD is dominated by organics and sea salt. Sea salt AOD increased further due to the latest model upgrade in June 2018 with the new sea salt emission scheme activated, while dust AOD became lower compared to earlier years.

The aerosol Ångström exponent (AE) contains information about the size distribution of the aerosol, and implicitly about composition. In the last 6 months the o-suite AE became more positive indicating a change to slightly more fine particles since the latest model upgrade to version 45R1 in June 2018.

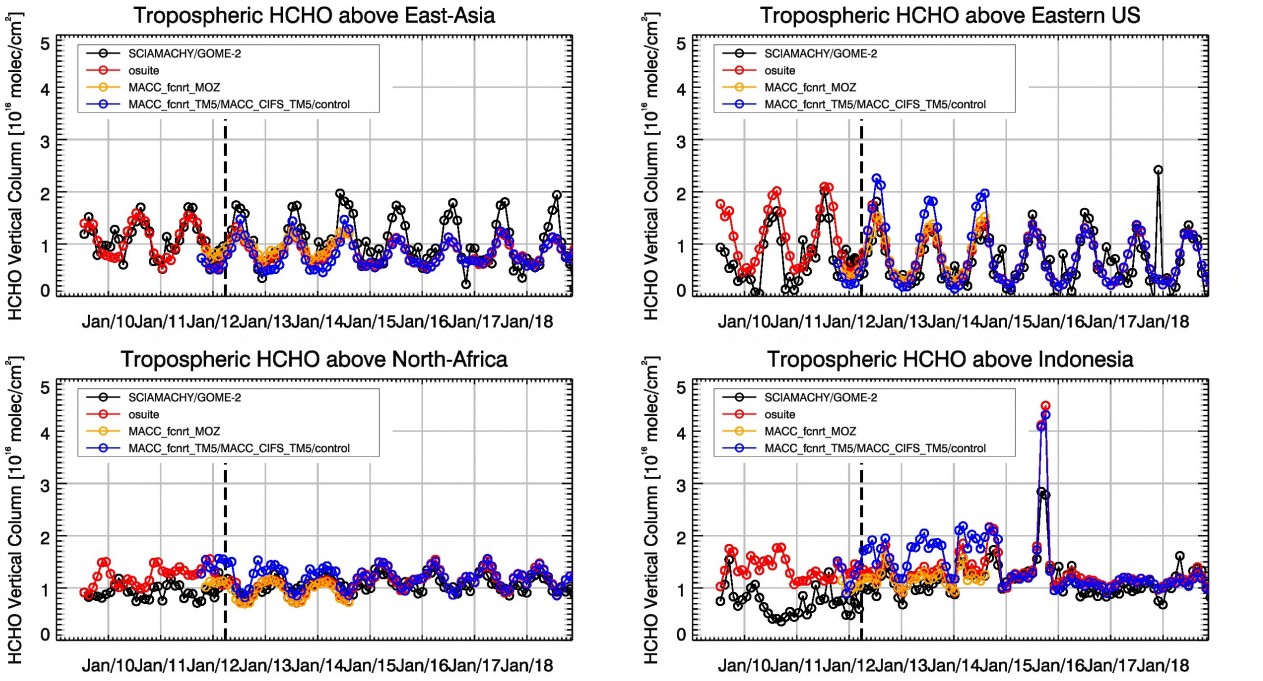


Figure S.4: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions. The blue line shows MACC_fcnrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 to August 2014 and control results from September 2014 onwards (the model run without data assimilation is termed control since Sep 2014). The regions differ from those used for NO_2 to better focus on HCHO hotspots: East-Asia (25-40°N, 110-125°E), Eastern US (30-40°N, 75-90°W), Northern Africa (0-15°N, 15°W-25°E) and Indonesia (5°S-5°N, 100-120°E). Negative satellite retrieved values over Eastern US are due to a lack of data (caused by instrument degradation) during Northern Hemisphere winter months for this region. Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

PM data are evaluated as defined by the IFS aerosol model. An evaluation of these PM10 and PM25 surface concentrations against an average from data in the period 2000-2009 at 160 background sites in North America and Europe indicate that PM10 concentrations exhibit on average in the latest period an underestimation with MNMB bias of -32% in Europe and -1% in North America. PM25 concentrations are underestimated -4% in Europe and overestimated 47% in North America. A higher positive bias is also found for AOD in North America than in Europe. The fraction of data within a factor 2 of observed values has increased for PM10 and decreased for PM25 compared to earlier years.

For autumn, dust activity is low in comparison with the previous season. The CAMS global o-suite reproduces the main locations of dust activity in the Sahara at latitude around 15ºN (in the Bodélé Basin and the Mali/Mauritania border). Otherwise, dust concentration in the Middle East as well as dust transport over the North Atlantic region and the Red Sea appears underestimated by the model in comparison with the satellites. For September to November, the o-suite reproduces the daily variability of AERONET observations with a correlation coefficient of 0.83, averaged over all the AERONET sites (as in the case of the SDS-WAS multi-model product), which is close to the control experiment that shows a correlation coefficient of 0.81.

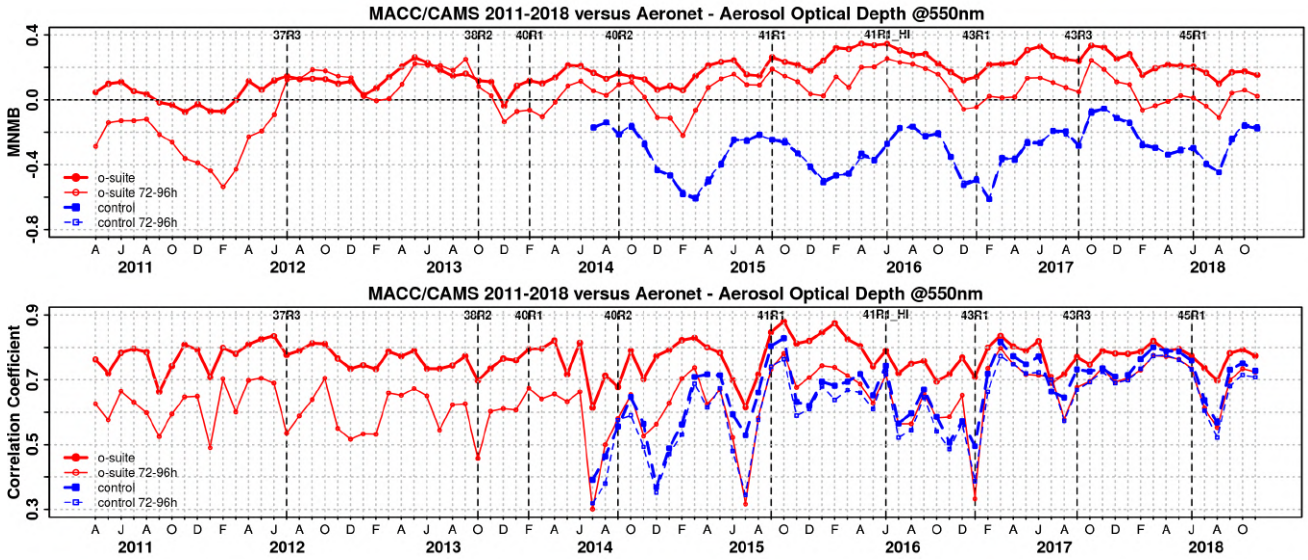


Figure S.5. Aerosol optical depth at 550nm in IFS 00Z model simulations for April 2011 – September 2018 against daily matching Aeronet Version3 level 1.5 data. a) Modified normalized mean bias (MNMB); o-suite (thick red curve); o-suite at last forecast day (light red curve); Control (blue dashed); Control at last forecast day (light blue dashed); b) Corresponding correlation coefficient. Model version changes are marked as vertical bars.

In the Sahel, the o-suite shows strong underestimations (MB of -0.18, slightly higher than control with MB of -0.03) although the o-suite better reproduces the observed daily variability (with a correlation value of 0.67 for o-suite in comparison to control that has a correlation of 0.59). Also, in the Tropical North Atlantic region the o-suite is underestimated (MB of -0.06) in agreement with the underestimation shown in the comparison with MODIS. In the Tropical North Atlantic region, the o-suite presents lower correlation coefficient (0.77) than the control experiment (0.88). The decrease in the skill score of the o-suite is associated with the prediction of the most extreme events during the period of 20-31 October in which the control experiment better captures the maximum DOD peaks. Over the Iberian Peninsula and the Mediterranean, the CAMS model presents high correlations between 0.72 and 0.84 and slight underestimations (MB between -0.01 and 0.02). However, in the Eastern Mediterranean region the o-suite shows an improvement with a correlation coefficient that goes from 0.68 for control to 0.78 for o-suite, probably connected to the assimilation of observations, including the most extreme events observed on 1st and 20th October.

Otherwise, the comparison of the 48h and 72h forecasts for both CAMS experiments shows that the prediction is stable during the 3-days forecasts with correlation coefficients 0.83, 0.81, and 0.80 for the 24, 48 and 72h forecasts respectively, for the o-suite. In Tropical North Atlantic, the correlation coefficient (r) for control increases from 0.86 to 0.90 in going from the 48h to 72h forecast. These scores are better than in the case of the o-suite that presents correlation coefficient (r) of 0.70 at 48h to 0.67 at 72h. Otherwise, in sub-Tropical North Atlantic, the o-suite presents better scores than control with correlation coefficients of 0.74 at 48h and 0.63 at 72h for the o-suite.

Backscatter coefficients are low-biased in the planetary boundary layer (PBL). Possible reasons are absence of ammonia and nitrate in the model (foreseen to be activated soon), assumption of too high particle densities (too compact materials) in the mass to backscatter conversion, and the lack of a vertical transport barrier at the top of the PBL, causing dilution with free troposphere air. Free

troposphere (FT) background backscatter coefficients are biased high, probably due to wrong re-distribution between PBL and FT. This is not fixed by the assimilation, which instead adds aerosol to the whole profile. The backscatter bias on a specific level thus depends on its relative position w.r.t. to the boundary layer height (BLH). The model BLH agrees reasonably (within a few 100 m) with observations under favourable measurement conditions. Very often, however, meteorological conditions prevent formation, unambiguousness or detectability of the BLH or make the latter a challenge.

System performance in the Arctic

The CAMS model runs are evaluated using surface ozone measurements from 7 sites within the IASOA networks and ozone concentrations in the free troposphere and the stratosphere are evaluated using ozone soundings.

For the period from December 2014 to November 2018 the simulations of the surface ozone concentrations are on average in good agreement with the observations apart from ozone depletion events in spring (March to June). During September – November 2018 there is generally an underestimation of the surface ozone concentrations in the Arctic for the o-suite (NMB = -9% to +2%) and an overestimation for the control run (NMB = 2% to 6%). The control run performs better with higher correlation coefficients (r = 0.27 – 0.84 for the o-suite, compared to r = 0.51 – 0.86 for the control run).

Ozone concentrations in the free troposphere are in good agreement with observations with low relative bias (-1% to 1% for the o-suite and 1% to 7% for the control run). In the UTLS, the o-suite has a slightly higher positive bias for the Arctic sites. Ozone concentrations in the stratosphere are in good agreement with observations with low relative bias (3% to 8% for the o-suite). The vertical distribution of stratospheric ozone is well represented in the Arctic.

Comparison with FTIR observations from the NDACC network shows that the CO tropospheric columns are well simulated at the two Arctic sites (Eureka and Kiruna) with low bias (-4.7% and 1.5%), while there is a large negative bias for the control (7-12%). The assimilation has a small positive effect on the correlation coefficient: r = 0.90-0.96 for the o-suite and 0.84-0.94 for the control run. Comparison MOPITT versions 7 shows that modeled CO total columns are in good agreement with the satellite retrievals with low bias in the Arctic (±10%).

System performance in the Mediterranean

The model is compared to surface O_3 observations from the AirBase network. Our analysis shows that model MNMBs vary between -20% and 30% depending on the station. Temporal correlation coefficients between simulated and observed surface ozone for both the o-suite and control runs are highly significant over the entire Mediterranean from Gibraltar to Cyprus.

The CAMS o-suite reproduces the daily variability of AERONET observations (with correlation coefficients of 0.75, 0.67 and 0.62, respectively for Western, Central and Eastern Mediterranean). Although, the operational CAMS model tends to overestimate the observed AOD values by AERONET sites resulting in an MB of 0.01, 0.01 and 0.02, respectively for Western, Central and Eastern Mediterranean. The highest peaks on CAMS AOD simulations are linked to desert dust intrusions. The second half of October was particularly intense in terms of dust outbreak in Eastern

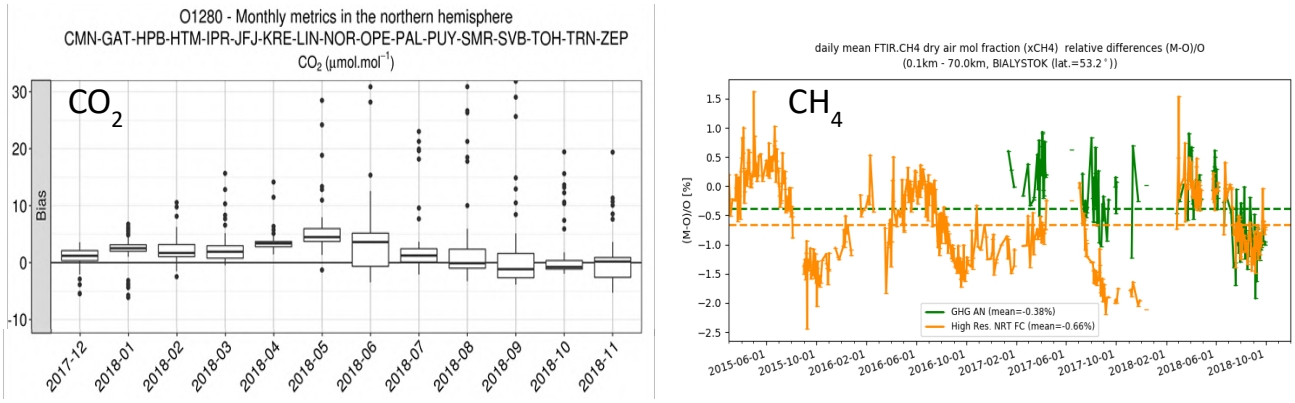


Figure S.6. Left: CO_2 monthly differences (in ppm) of the high-resolution model with surface observations for a composite of 17 European sites (Dec.2017-Dec.2018). Right: CH_4 differences (in %) of the high-resolution model (orange) and the analysis (green) with total column measurements at Bialystok, Poland (2015-2018).

Mediterranean achieving AOD values up to 1.4. Meanwhile, the Western Mediterranean sites received the most intense dust outbreaks during September. At surface levels, the o-suite tends to overestimate the PM10 and PM2.5 EIONET-Airbase observations in north-western Mediterranean sites meanwhile tends to underestimate the PM10 and PM2.5 observed values at Spanish sites.

Climate forcing

Greenhouse gases

CO_2 and CH_4 surface concentrations from ICOS network, and total or partial columns from TCCON and NDACC stations have been used to validate the analysis and high-resolution forecast experiments.

The surface and total column measurements indicate an overestimation of the amplitude of the CO_2 seasonal cycle in the northern hemisphere by 1 to 2% (Fig. S.6). The drought anomaly in spring/summer 2018 has an additional effect on the comparison with surface sites from May to July 2018, with an overestimated impact of the drought on the CO_2 concentrations.

The highest underestimation of CH4 by the model is observed between July and November 2018, both at ICOS, NDACC and TCCON sites. According to the long term TCCON measurement this feature occurs every year (Fig. S.6).

The model overestimates CH4 concentrations at surface and total columns sites located in the vicinity of megacities like Paris and Los Angeles, indicating an issue with the emission inventories, and the difficulty to represent the pollution plumes emitted from megacities.

Only small differences appear between the greenhouse gas analysis and the high-resolution simulation from 2018 onwards. However, for CH_4 the correlation coefficients of the analysis time series are slightly larger for most sites, compared to the high-resolution results, both for total column and surface measurements.

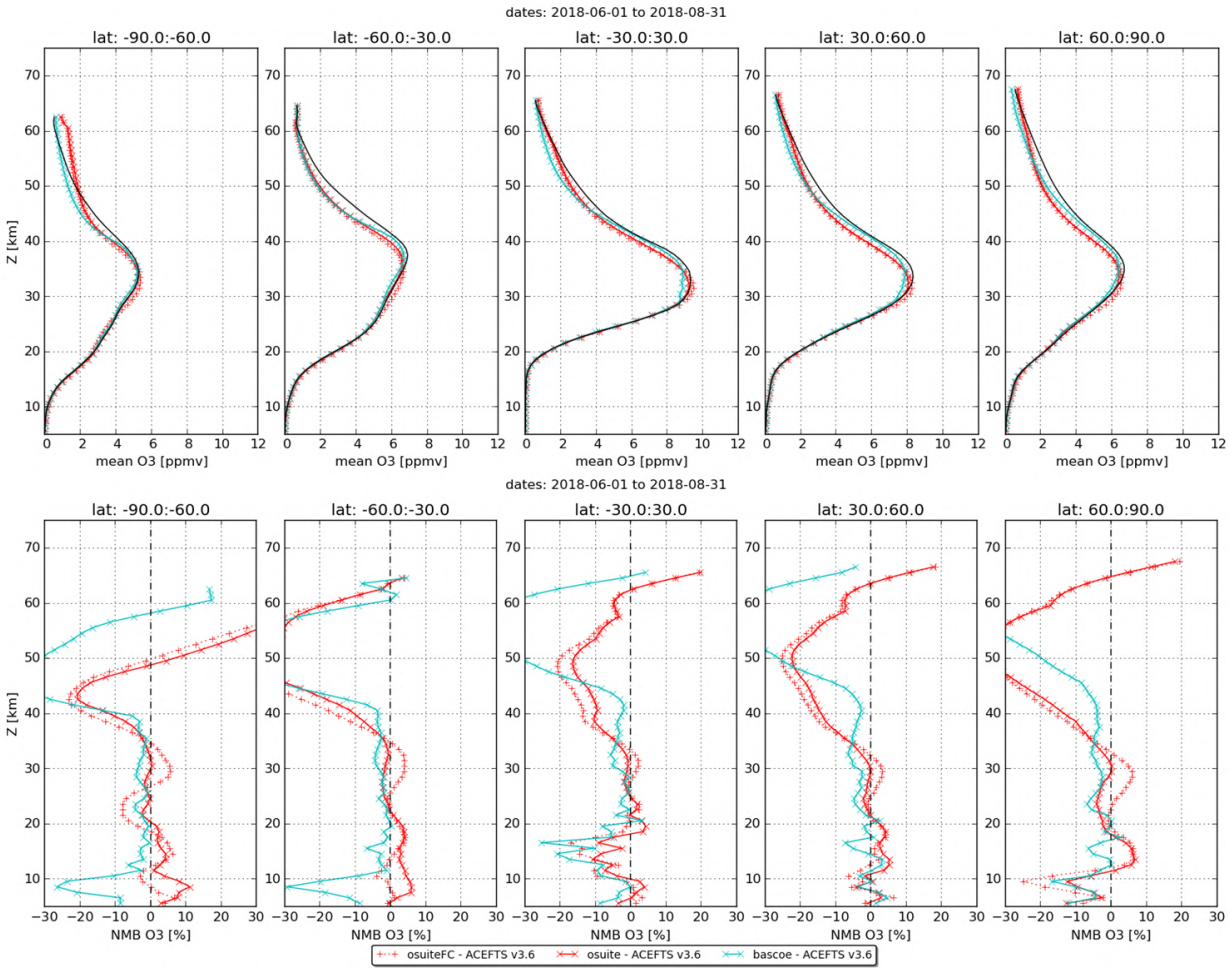


Figure S.7: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with ACE-FTS observations for the period September to November 2018.

Ozone layer and UV

Ozone partial columns and vertical profiles

Ozone columns and profiles have been compared with the following observations: vertical profiles from balloon-borne ozone sondes; ground-based remote-sensing observations from the NDACC (Network for the Detection of Atmospheric Composition Change, http://www.ndacc.org); and satellite observations by two instrument (OMPS-LP, ACE-FTS). Furthermore, the o-suite analyses are compared with those delivered by the independent assimilation system BASCOE.

Compared to ozone sondes the model O_3 partial pressures are slightly overestimated in all latitude bands (MNMB between 4 and +12%) except above the Antarctic.

Comparisons with the NDACC network include 17 stations for UVVIS and FTIR stratospheric columns, microwave profiles for Ny Alesund (78.9°N) and Bern (47°N) and LIDAR profiles at Hohenpeissenberg (47.8°N) and Observatoire Haute Provence (OHP), France (43°N). The comparison with the UVVIS stations are generally in agreement with the o-suite, while it indicates a

latitudinal dependence of the biases for the control run. The result from MWR and LIDAR comparisons for the current period are in line with those of previous reports.

The comparison with independent satellite observations is generally in good agreement for the considered period: for ACE-FTS (Fig. S.7), the NMB is mainly within 10% between 5km and 40km, and mostly within 5% between 15km and 35km except in the tropics. OMPS-LP has less regular profiles, but the NMB still remain within 15% for most parts of the 20-40 km range.

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 also been evaluated but the results are only indicative.

Events

Dusty Mediterranean during October 2018: In second half of October 2018, MODIS satellite detected two dust outbreaks over Central and Eastern Mediterranean Basin. The first outbreak on 20th October came in from Syria, moving towards Turkey. A second dust event was observed on 16th October with origin in Libya, affecting the Central Mediterranean. The CAMS o-suite AOD did timely reproduce the spatial distribution of the two dust plumes despite the model tendency to underestimate the observed maximum values, in particular for the event which originated in Libya.

Alaska fire events, July 2018: Several fire events were detected during July 2018 in Alaska region. The CO peak detected by MOPITT and IASI between July 7 and 9 was well captured by CAMS-global with similar CO concentrations. A second fire event detected around July 22 was also well captured, but somewhat underestimated by about 20%.

Table of Contents

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 near-real time services consist of daily analysis and forecasts with the ECMWF 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 recent status of the near-real time service. The 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.

This validation report is accompanied by the "Observations characterization and validation methods" report, Douros et al. (2017), 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.

Key CAMS NRT products and their users are: Boundary conditions for regional air quality models (e.g. AQMEII, air quality models not participating in CAMS); Long range transport of air pollution (e.g. LRTAP); Stratospheric ozone column and UV (e.g. WMO, DWD); 3D ozone fields (e.g. SPARC). As outlined in the MACC-II Atmospheric Service Validation Protocol (2013) and MACC O-INT document (2011), relevant user requirements are quick looks of validation scores, and quality flags and uncertainty information along with the actual data. This is further stimulated by QA4EO (Quality Assurance Framework for Earth Observation, http://www.qa4eo.org) who write that “all earth observation data and derived products is associated with it a documented and fully traceable quality indicator (QI)”. It is our long-term aim to provide such background information. The user is seen as the driver for any specific quality requirements and should assess if any supplied information, as characterised by its associated QI, are "fit for purpose" (QA4EO task team, 2010).

CAMS data are made available to users as data products (grib or netcdf files) and graphical products from ECMWF, accessible through the catalogue on http://atmosphere.copernicus.eu/.

A summary of the system and its recent changes is given in section 2. Subsequent sections gives an overview of the performance of the system for various species, and during recent events. Routine validation results can be found online via regularly updated verification pages, http://atmosphere.copernicus.eu/user-support/validation/verification-global-services.
Table 1.2 lists all specific validation websites that can also be found through this link.

Table 1.1: Overview of the trace gas species and aerosol aspects discussed in this CAMS near-real time validation report. Shown are the datasets assimilated in the CAMS analysis (second column) and the datasets used for validation, as shown in this report (third column). Green colours 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 PMAp AOD	AOD, Ångström: AERONET, GAW, Skynet, MISR, OMI, lidar, ceilometer
Aerosol mass (PM10, PM2.5)	MODIS Aqua/Terra	European AirBase stations
O_3, stratosphere	MLS, GOME-2A, GOME-2B, OMI, SBUV-2, OMPS	Sonde, lidar, MWR, FTIR, OMPS, ACE-FTS, OSIRIS, BASCOE and MSR analyses
O_3, UT/LS	MLS	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	IASI, MOPITT	IAGOS
CO, free troposphere	IASI, MOPITT	IAGOS, MOPITT, IASI, TCCON
CO, PBL / surface	IASI, MOPITT	Surface CO: WMO/GAW, NOAA/ESRL
NO_2, troposphere	OMI, GOME-2, 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	GOSAT	TCCON
CH_4, surface, PBL		ICOS
CH_4, column	GOSAT, IASI	TCCON

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

Reactive gases – Troposphere
IAGOS tropospheric ozone and carbon monoxide: http://www.iagos.fr/cams/ Surface ozone from EMEP (Europe) and NOAA-ESRL (USA): http://www.academyofathens.gr/cams Tropospheric nitrogen dioxide and formaldehyde columns against satellite retrievals: http://www.doas-bremen.de/macc/macc_veri_iup_home.html Tropospheric CO columns against satellite retrievals: http://cams.mpimet.mpg.de GAW surface ozone and carbon monoxide: https://atmosphere.copernicus.eu/charts/cams_gaw_ver/v0d_gaw_oper_operfc_nrt_sites?facets=undefined&time=2018060100,0,2018060100&fieldpair=CO&site=cmn644n00
Reactive gases - Stratosphere
Stratospheric composition: http://www.copernicus-stratosphere.eu NDACC evaluation in stratosphere and troposphere (the NORS server) http://nors-server.aeronomie.be
Aerosol
Evaluation against Aeronet stations: http://aerocom.met.no/cams-aerocom-evaluation/ More in-depth evaluations are available from the Aerocom website. WMO Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS) model intercomparison and evaluation: http://sds-was.aemet.es/forecast-products/models
Satellite data monitoring
Monitoring of satellite data usage in the Near-Real-Time production: https://atmosphere.copernicus.eu/charts/cams/cams_satmon?facets=undefined&time=2016071800&Parameter=AURA_MLS_profile_Ozone_1_GLOBE

Naming and color-coding conventions in this report follow the scheme as given in Table 1.3.

Table 1.3. Naming and colour conventions as adopted in this report.

Name in figs	**experiment**	**Colour**
{obs name}	{obs}	black
o-suite D+0 FC	0001	red
control	gsyg	blue
GHG high-resolution run	gqpe / ghqy	orange
GHG global analysis	gqiq	green

2. System summary and model background information

The specifics of the different CAMS model versions are given below (section 2.1) including an overview of model changes. Other systems used in CAMS are listed in section 2.2. An overview of products derived from this system is given in section 2.3. Timeliness and availability of the CAMS products is given in section 2.4.

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

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

Forecast system	Exp. ID	Brief description	Upgrades (e-suite ID)
o-suite	0001	Operational CAMS DA/FC run	20180626-present 20170926-20180625 20170124-20170926 20160621-20170124 20150903-20160620 20140918-20150902
Control	gzhy gsyg gnhb gjjh geuh g4o2	control FC run without DA	20180626-present (gzhy) 20170926-20180625 (gsyg) 20170124-20170926 (gnhb) 20160621-20170124 (gjjh) 20150901-20160620 (geuh) 20140701-20150902 (g4o2)
GHG run	ghqy gf39	High resolution T1279, NRT CO_2 and CH_4 without DA	20160301-20170621 (ghqy) 20150101-20160229 (gf39)
	gqpe	High resolution Tco1279 (~9km) NRT CO_2, CH_4 and linCO forecast, initialized from GHG analysis gqiq and CAMS operational CO analysis	20170101-present
	gqiq	GHG analysis Tco399 (~25km)	20170101-present

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	V4	O3 Profiles	20130107 -
OMI	AURA	NASA	V883	O3 Total column	20090901 -
GOME-2A	Metop-A	Eumetsat	GDP 4.8	O3 Total column	20131007 -
GOME-2B	Metop-B	Eumetsat	GDP 4.8	O3 Total column	20140512 -
SBUV-2	NOAA-19	NOAA	V8	O3 21 layer profiles	20121007 -
OMPS	Suomi-NPP	NOAA / EUMETSAT		O3 Profiles	20170124 -
IASI	MetOp-A	LATMOS/ULB Eumetsat	-	CO Total column	20090901 - 20180621 20180622 -
IASI	MetOp-B	LATMOS/ULB Eumetsat	-	CO Total column	20140918 - 20180621 20180622 -
MOPITT	TERRA	NCAR	V5-TIR V7-TIR V7-TIR Lance	CO Total column	20130129 - 20160124 - 20180626 20180626
OMI	AURA	KNMI	DOMINO V2.0	NO2 Tropospheric column	20120705 -
GOME-2A/2B	METOP A/B	Eumetsat	GDP 4.8	NO2 Tropospheric column	20180626 -
OMI	AURA	NASA	v003	SO2 Tropospheric column	20120705-20150901
GOME-2A/2B	METOP A/B	Eumetsat	GDP 4.8	SO2 Tropospheric column	20150902 -
MODIS	AQUA / TERRA	NASA	Col. 5 Deep Blue Col. 6, 6.1	Aerosol total optical depth, fire radiative power	20090901 - 20150902 - 20170124 -
PMAp	METOP-A METOP-B	EUMETSAT		AOD	20170124 - 20170926 -

2.1.1 o-suite

The o-suite consists of the IFS-CB05 chemistry combined with the CAMS bulk aerosol model. The chemistry is described in Flemming et al. (2015) and Flemming et al. (2017), aerosol is described in Morcrette et al. (2009). The forecast length is 120 h. The o-suite data is stored under **expver '0001'** of **class 'MC'**. On 21 June 2016 the model resolution has seen an upgrade from T255 to T511, and forecasts are produced twice per day. The latest upgrade of the system is based on IFS version cy45r1_CAMS (https://confluence.ecmwf.int/display/COPSRV/Current+global+production+suites) and took place on 26 June 2018. The validation for this upgrade is described in Eskes et al., 2018. The update relevant for this report (for SON-2018) is this 26 June 2018 upgrade. A short summary of the main specifications:

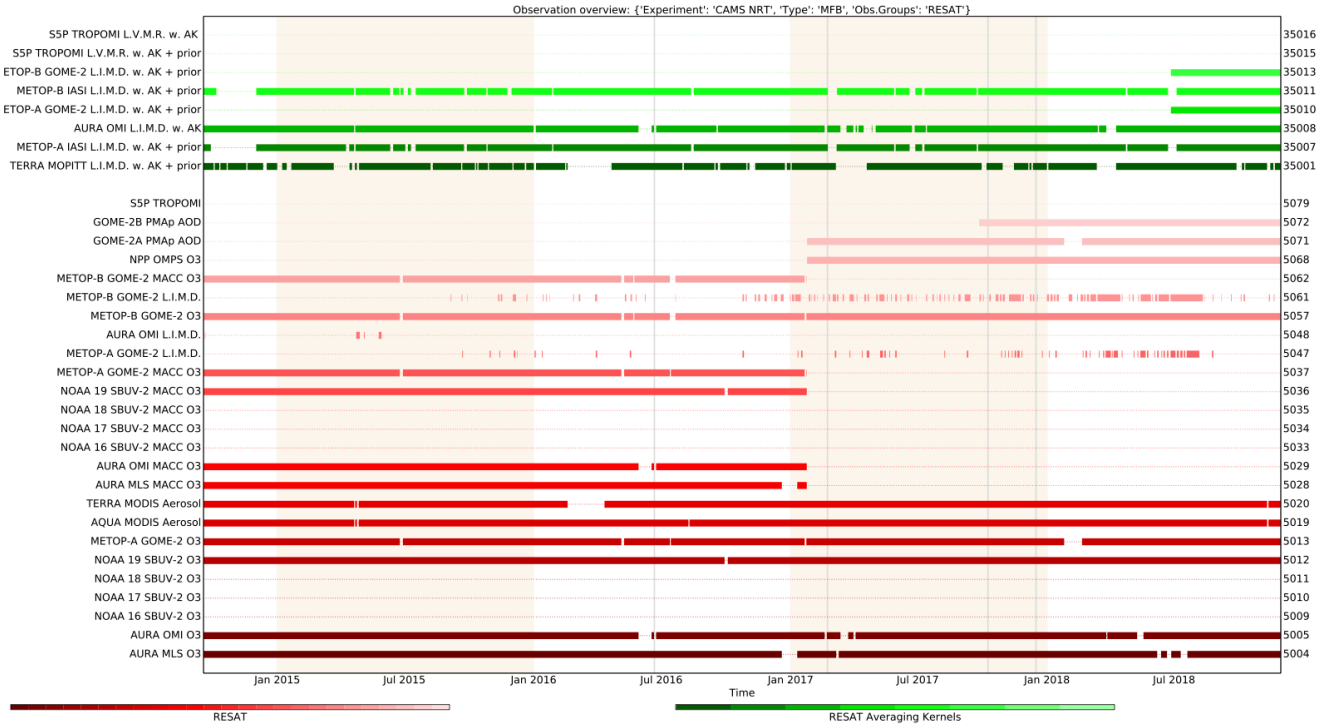


Figure 2.1: Satellite observation usage in the real-time analysis, for ozone, CO, aerosol AOD, from October 2014 onwards. Top eight rows: products assimilated with averaging kernels. New assimilated products since the 24 January 2017 upgrade are the PMaP AOD including GOME-2B and OMPS ozone profile observations. Note that the lines mentioning "MACC O3" should be discarded.

- The modified CB05 tropospheric chemistry is used (Williams et al., 2013), originally taken from the TM5 chemistry transport model (Huijnen et al., 2010)
- Stratospheric ozone during the forecast is computed from the Cariolle scheme (Cariolle and Teyssèdre, 2007) as already available in IFS, while stratospheric NOx is constrained through a climatological ratio of HNO_3/O_3 at 10 hPa.
- Monthly mean dry deposition velocities are based on the SUMO model provided by the MOCAGE team.
- Data assimilation is described in Inness et al. (2015) and Benedetti et al. (2009) for chemical trace gases and aerosol, respectively. Satellite data assimilated is listed in Table 2.2 and Fig. 2.1.
- Anthropogenic and biogenic emissions are based on MACCity (Granier et al., 2011) and a climatology of the MEGAN-MACC emission inventories (Sindelarova et al., 2014)
- NRT fire emissions are taken from GFASv1.2 (Kaiser et al. 2012).

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 SO_2 (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).

New source scheme for Secondary Organic Aerosols (part of the Organic Matter), based on scaled CO emissions. This is a change form the current AEROCOM-based emissions. The impact is an

increase of organic matter aerosol concentrations. The upgrade of 24 January 2017 introduced the following adjustments: 1. Reduced dust emissions over Taklamakan desert and India. 2. Dust emissions adjusted towards more larger particles. 3. Reduction in sulphate aerosol. 4. Mass fixer for aerosols.

A history of updates of the o-suite is given in Table 2.4, and is documented in earlier MACC-VAL reports: https://atmosphere.copernicus.eu/node/326. This includes a list with changes concerning the assimilation system.

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.

On 26 June 2018 the system has been upgraded to cy45r1. A validation report for this upgrade (Eskes et al., 2018) is available here: https://atmosphere.copernicus.eu/sites/default/files/2018-06/CAMS84_2015SC2_D84.3.1.5_201802_esuite_v1_0.pdf The addendum to this document is called CAMS84_2015SC3_D84.3.1.5_201802_esuite_addendum.pdf.

2.1.2 Control

The control run (relevant expver = **gzhy,** since 26/06/2018) applies the same settings as the respective o-suites, based on the coupled IFS-CB05 system with CAMS aerosol for cy54r1, 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 considered to be a useful product of this run. 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 CO_2 and CH_4 forecasts and delayed-mode analyses

The pre-operational forecasts of CO_2 and CH_4 use an independent setup of the IFS as the o-suite, 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 CO_2 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,

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 IFS-CB05 is introduced to model atmospheric chemistry.
2015.09.03	Update to experiment g9rr: based on meteo cy41r1.
2016.06.21	Update to experiment 0067: based on meteo cy41r1, but a resolution increase from T255 to T511, and two production runs per day
2017.01.24	Update to cycle 43R1_CAMS, T511L60
2017.09.26	Update to cycle 43R3_CAMS, T511L60
2018.06.26	Update to cycle 45R1_CAMS, T511L60

consistent with those used as prior information in the CH_4 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 also included a linear CO scheme (Massart et al., 2015).

The experiments analysed in this report are:

- "**ghqy**" from March 2016. 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. Note that the CO_2, CH_4 and linear CO tracers are free-running.
- "**gqpe**" from January 2017 to present. It runs with a TCO1279 Gaussian cubic octahedral grid (equivalent to approximately 9km horizontal resolution). Note that the CO_2, CH_4 and linear CO tracers are initialized with the GHG analysis (gqiq) for CO_2 and CH_4 and the CAMS operational analysis for CO.
- The greenhouse gas analysis experiment "**gqiq**" runs on a TCO399 grid (equivalent to around 25km) and 137 vertical levels and is available from January 2017. This experiment runs in

delayed mode (4 days behind real time) and makes use of observations from TANSO-GOSAT (methane and CO_2) and MetOp-IASI (methane).

2.2 Other systems

2.2.1 BASCOE

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

- The following species are assimilated: O_3, H_2O, HNO_3, HCl, HOCl, N_2O and ClO.
- It lags by typically 4 days, due to latency time of 4 days for arrival of non-ozone data from Aura/MLS-SCI (i.e. the scientific offline Aura/MLS dataset).
- Global horizontal grid with a 3.75° longitude by 2.5° latitude resolution.
- Vertical grid is hybrid-pressure and consists in 86 levels extending from 0.01 hPa to the surface.
- Winds, temperature and surface pressure are interpolated in the ECMWF operational 6-hourly analyses.
- Time steps of 20 minutes, output every 3 hours

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

2.2.2 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 O_3 columns are assimilated
- Global horizontal grid with a 3° longitude by 2° latitude resolution.
- Vertical grid is hybrid-pressure and consists in 44 levels extending from 0.1 hPa to 100 hPa.
- Dynamical fields from ECMWF operational 6-hourly analysis.

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

2.2.3 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, twelve regional and global models (see the complete list in the following link https://sds-was.aemet.es/forecast-products/forecast-evaluation/model-inter-comparison-and-forecast-evaluation/at_download/file) 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 0.5º x 0.5º. 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.

2.3 CAMS products

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

2.4 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. The CAMS production KPI is defined as the percentage of cycles in which all the general data dissemination tasks are completed before the deadlines: 10 UTC for the 00:00 and 22 UTC for the 12:00 UTC run. This was in part 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. Note that since 21 June 2016 two CAMS forecasts are produced each day.

For the period September-November 2018, 100% of the forecasts were delivered on time.

Table 2.6: Timeliness of the o-suite from Dec 2014 to the end of May 2018. From June 2016 onwards CAMS has produced two forecasts per day.

Months	**On time, 10 & 22 utc**	**80th perc**	**90th perc**	**95th perc**
Dec-Feb '14-'15	97%	D+0, 19:43	D+0, 20:28	D+0, 21:13
Mar-May 2015	96%	D+0, 19:38	D+0, 21:03	D+0, 21:40
Jun-Aug 2015	95%	D+0, 20:24	D+0, 20:53	D+0, 21:54
Sept-Nov 2015	95%	D+0, 19:44	D+0, 20:55	D+0, 21:51
Dec-Feb '15-'16	100%	D+0, 18:39	D+0, 18:57	D+0, 19:43
Mar-May 2016	98%	D+0, 19:32	D+0, 19:47	D+0, 20:00
Jun-Aug 2016 (00 and 12 cycle)	100%	D+0, 08:53 D+0, 20:55	D+0, 09:04 D+0, 21:01	D+0, 09:18 D+0, 21:18
Sep-Nov 2016	98.9%	D+0, 08:44 D+0, 20:44	D+0, 08:51 D+0, 20:48	D+0, 08:52 D+0, 20:51
Dec 2016 - Feb 2017	99.4%	D+0, 09:02 D+0, 21:01	D+0, 09:11 D+0, 21:02	D+0, 09:18 D+0, 21:04
Mar-May 2017	100%	D+0, 09:08 D+0, 21:07	D+0, 09:14 D+0, 21:09	D+0, 09:19 D+0, 21:11
Jun-Aug 2017	100%	D+0, 09:05 D+0, 21:05	D+0, 09:07 D+0, 21:08	D+0, 9:09 D+0, 21:10
Sep-Nov 2017	100%	D+0, 09:02 D+0, 21:00	D+0, 09:05 D+0, 21:04	D+0, 9:09 D+0, 21:07
Dec 2017 - Feb 2018	98.33%	D+0, 08:55 D+0, 20:54	D+0, 08:59 D+0, 20:59	D+0, 09:01 D+0, 21:02
Mar-May 2018	98.9%	D+0, 09:00 D+0, 21:00	D+0, 09:06 D+0, 21:03	D+0, 09:08 D+0, 21:06
Jun-Aug 2018	100%	D+0, 09:11 D+0, 21:07	D+0, 09:14 D+0, 21:09	D+0, 09:20 D+0, 21:11
Sep-Nov 2018	100%	D+0, 09:05 D+0, 21:03	D+0, 09:09 D+0, 21:07	D+0, 09:13 D+0, 21:10

3. Tropospheric Ozone

3.1 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 November 2017 to November 2018 (see Fig. 3.1.1 - 3.1.2). 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 Douros et al., 2017. The free troposphere is defined as the altitude range between 750 and 200hPa in the tropics and between 750 and 300hPa elsewhere.

MNMBs for the o-suite are mostly within the range ±20%, for all months, in all zonal bands, except for the Tropics, where larger positive MNMBs up to 34% appear, see Fig. 3.1.1.-3.1.4. The control run generally shows larger negative MNMBs, (up to -40%).

Over the Arctic, the o-suite mostly shows slightly positive MNMBs during summer and spring (MNMBs up to 10%), while during the winter season the MNMBs get negative (within -10%) see, Fig. 3.1.1.

Over the NH mid-latitudes MNMBs for the o-suite are on average close to zero all year round (maxima are -8% to +6%), which is generally a clear improvement compared to the control run, which shows larger negative MNMBs during the respective period.

Over the Tropics and over Antarctica, ozone mixing ratios are mostly overestimated by the o-suite (up to 34%) by the o-suite, see Fig. 3.1.2. The control run shows large negative MNMBs for Antarctica.

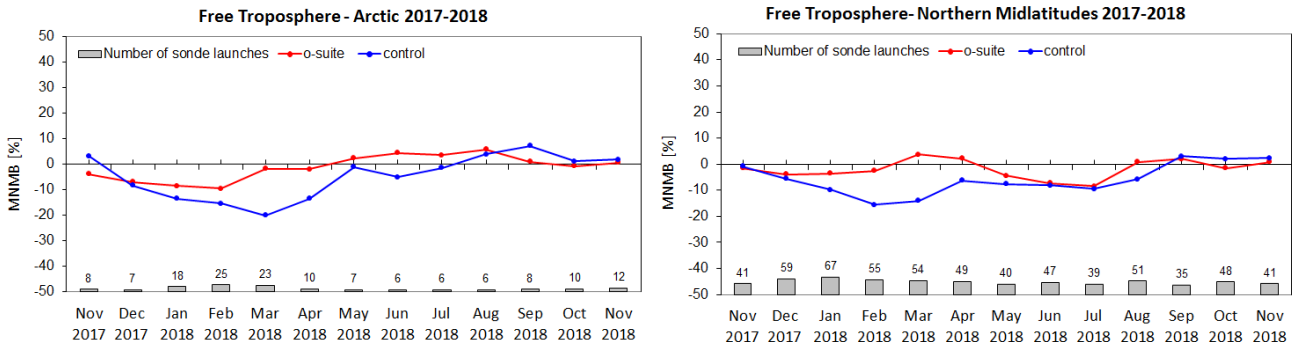


Figure 3.1.1: MNMBs (%) of ozone in the free troposphere (between 750 and 300 hPa) from the IFS model runs against aggregated sonde data over the Arctic (left) and the Northern mid latitudes (right). The numbers indicate the amount of individual number of sondes.

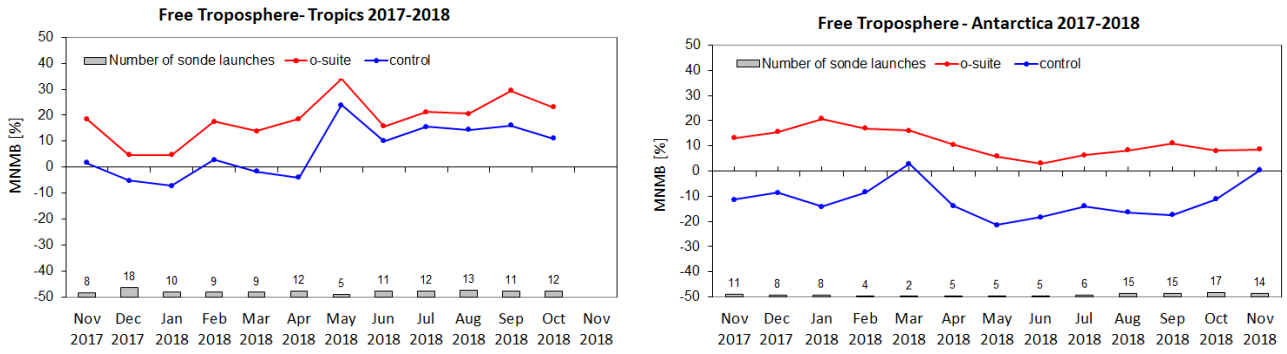


Figure 3.1.2: MNMBs (%) of ozone in the free troposphere (between 750 and 200hPa (Tropics) / 300hPa) from the IFS model runs against aggregated sonde data over the Tropics (left) and Antarctica (right). The numbers indicate the amount of individual number of sondes.

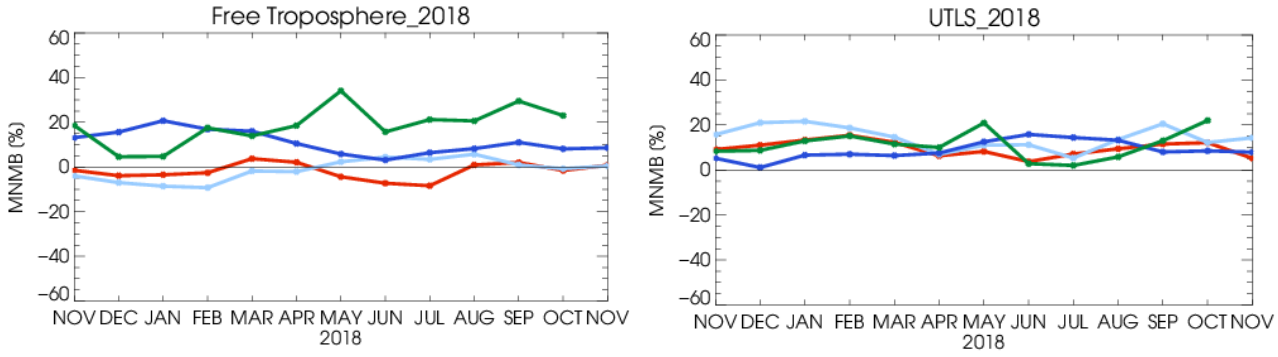


Figure 3.1.3: MNMBs (%) of ozone in the free troposphere (left, between 750 and 200hPa (Tropics) / 300hPa) and UTLS (right, between 300 and 100hPa (Tropics) / 60hPa) from the IFS model runs against aggregated sonde data over the Tropics (green) and Antarctica (blue), Arctic (light blue) and Northern Midlatitudes (red).

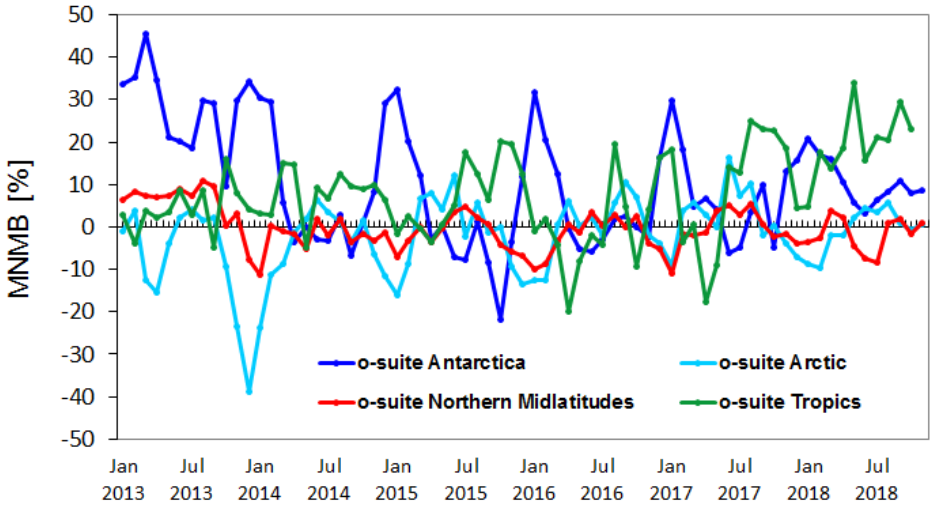


Figure 3.1.4: Time series of MNMB of ozone in the o-suite, compared against ozone sondes, averaged over different latitude bands. The free troposphere is defined here as the layer between 750 and 300 hPa.

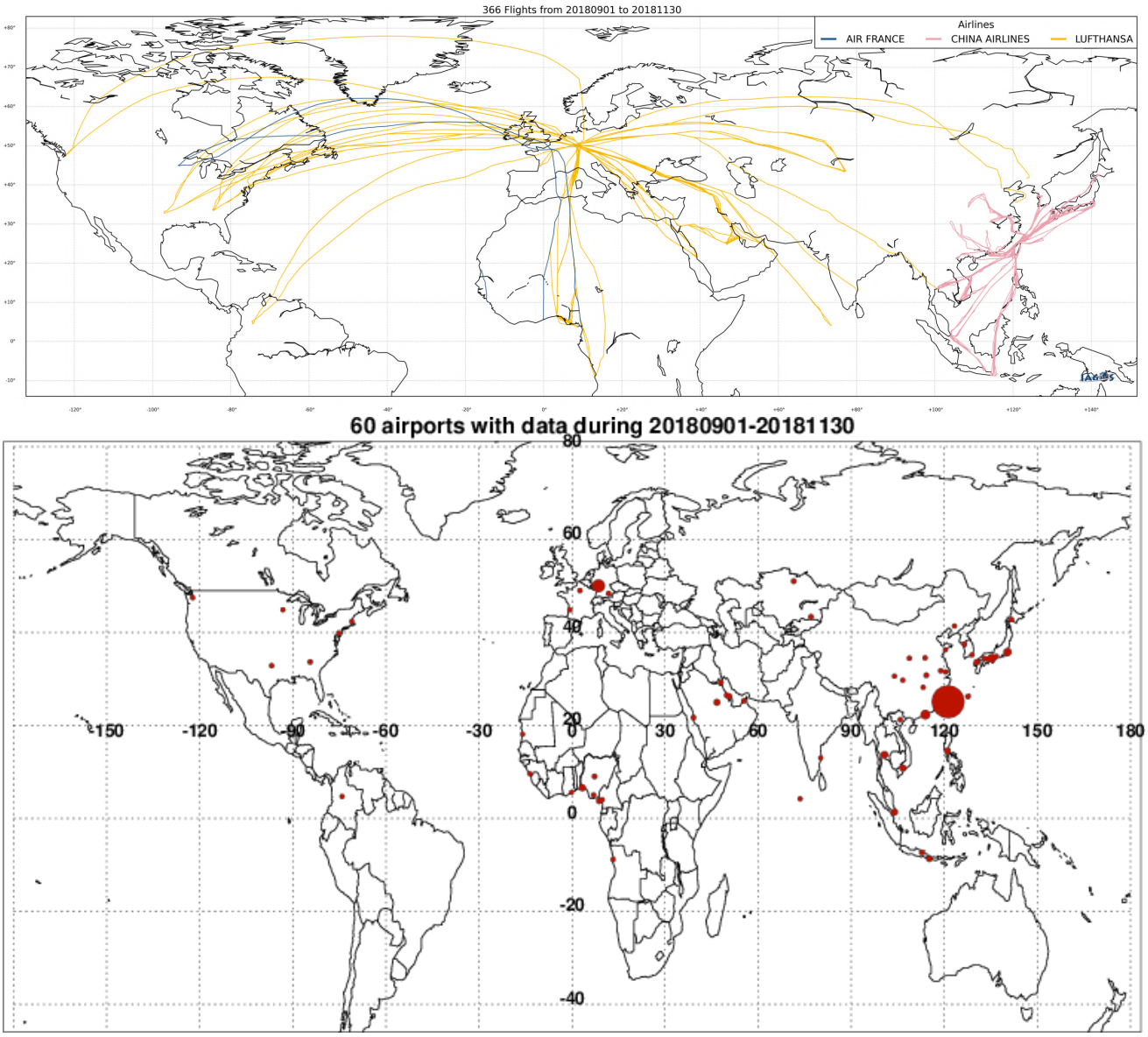


Figure 3.2.1. Map of the flights (top) and the visited airports (bottom) during the period September - November 2018, by the IAGOS-equipped aircraft. The size of the plotting circle represents the number of profiles available.

3.2 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 September - November 2018, the data displayed on the web pages and in this report include only the data as validated by the instrument PI. The available flights and available airports are shown in Fig. 3.2.1 top and bottom respectively. Performance indicators have been calculated for different parts of the IAGOS operations.

Six aircrafts were operating during this period. With these aircrafts, when operating continuously over the three-month period, we can expect a total of about 1260 flights. The actual number of

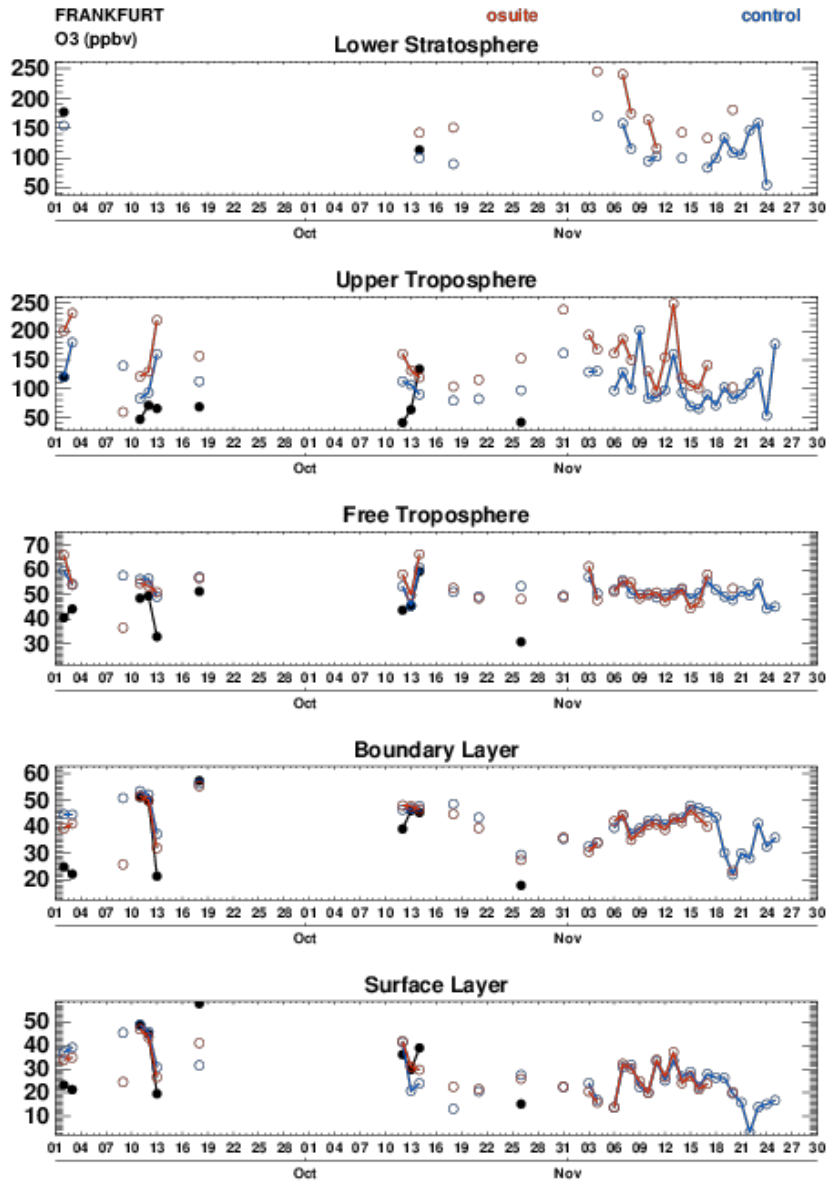


Figure 3.2.2. Time series of daily mean ozone over Frankfurt during SON 2018 for 5 layers: Surface, Boundary layer, Free Troposphere, Upper Troposphere and Lower Stratosphere.

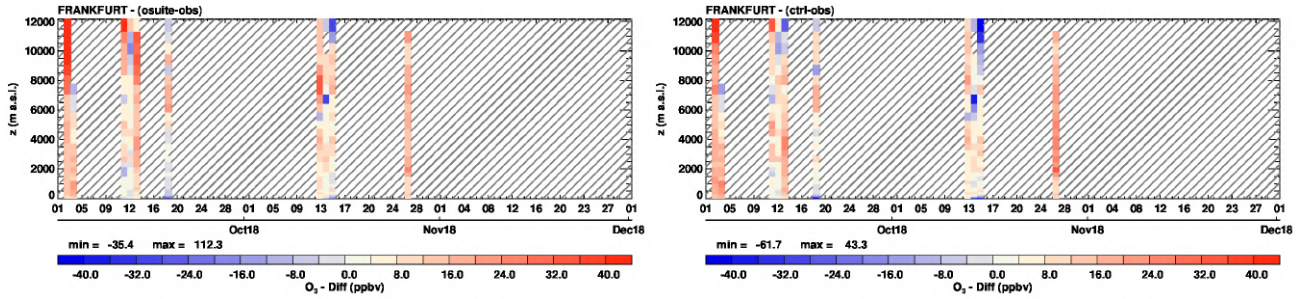


Figure 3.2.3 Time series of the absolute differences (model - observations) in daily profiles for ozone over Frankfurt during SON 2018. Left panel correspond to o-suite and right panel to control run.

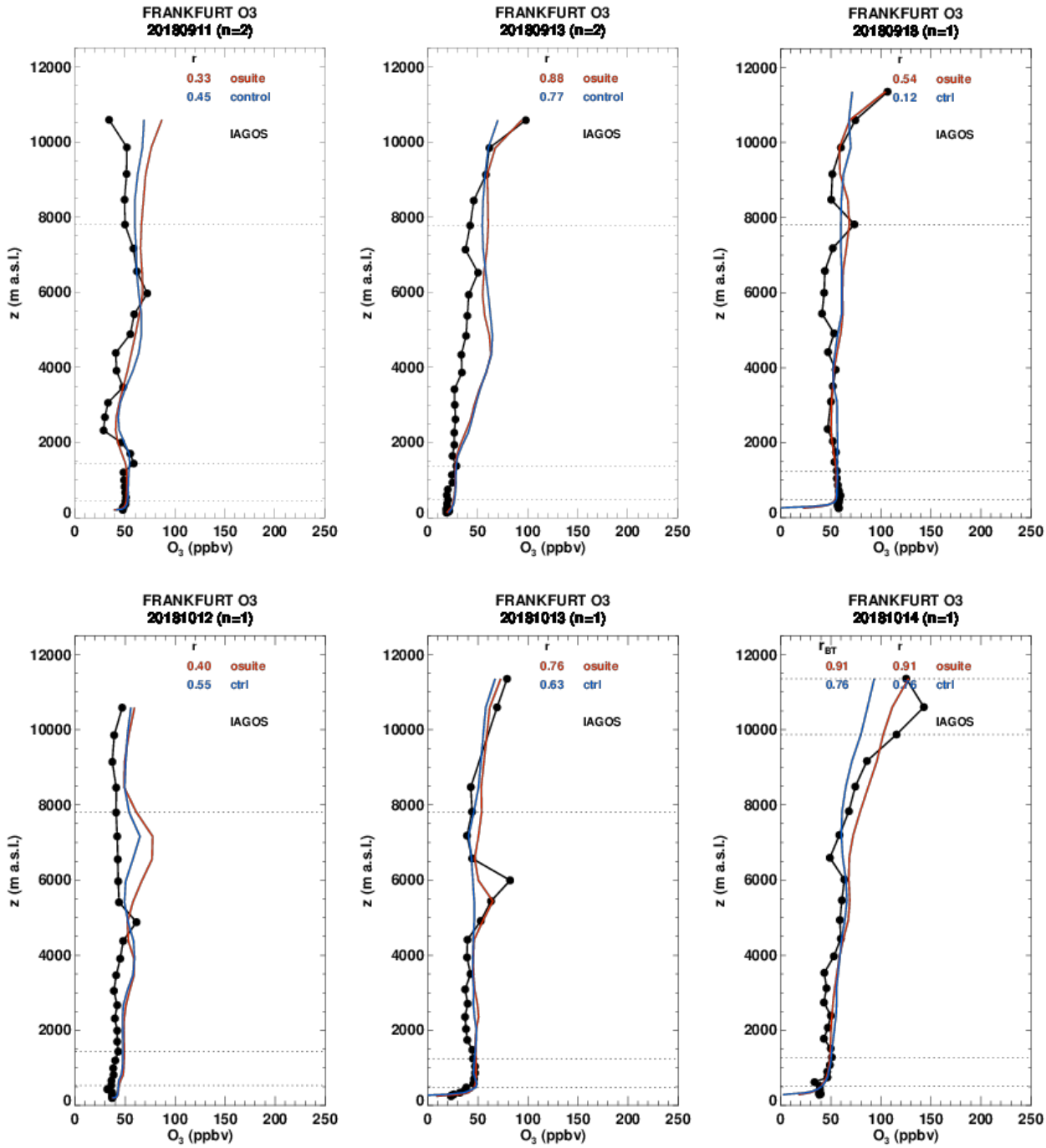


Figure 3.2.4 Selection of daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Europe during SON 2018.

flights within the period was 366 (732 profiles) giving a performance of 29 %. These flights are shown in Fig. 3.2.1 (top). Five percent (5%) of the operational flights had usable measurements of ozone and 38% of flights had usable CO. Delivering these O_3 and CO data were two aircraft from Lufthansa operating from Frankfurt, two aircraft operated by Air France based in Paris, two from China Airlines based in Taipei. Fig. 3.2.1 (bottom) shows the available airports, with a plotting circle scaled to the highest number of flights at an airport.

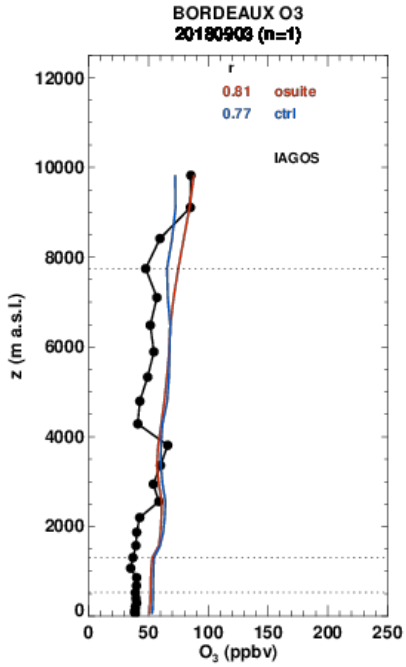


Figure 3.2.5 Profile for ozone at Bordeaux on 3 September 2018, from IAGOS (black) and the two NRT runs (o-suite: red, control: blue).

It is important to mention that during SON 2018, ozone sensors are working only on the Lufthansa aircrafts which explains partly the poor availability of ozone data. Moreover, due to transmission problems on one of these two packages many profiles are not available, which also affects the availability of CO data. The other ozone instrument started to present measurement problems, and a number of measurements could not be validated as Level 1. For all these reasons, there is only a small number of ozone profiles for this autumn season, mainly at Frankfurt but also a few at the other airports visited by Lufthansa.

Europe

Fig. 3.2.2 presents ozone time series at Frankfurt during September – November 2018. Despite the limited time coverage of IAGOS, ozone is well represented throughout the lower/middle troposphere. In the UTLS region ozone a few profiles show an overestimation by the o-suite and better results for the control run, consistent with findings in the period before September. Some examples of individual profiles are presented in Fig. 3.2.4.

In the time series at Frankfurt (Fig 3.2.2), an increase in ozone is observed in the free troposphere with values reaching more than 50 ppbv. In Fig. 3.2.4 profiles for 11 and 13 September are presented. On 11 September, the profile presents a maximum value in ozone near 6000 m with about 80 ppbv. This profile is well reproduced by both o-suite and control run with similar results. On 13 September, ozone values decrease and observations present a nearly constant profile from the surface to about 8000 m with ozone of about 20 ppb. However, in the modeled profiles ozone values remain high in the free troposphere with about 70 ppbv between 4000 and 8000 m. On 18 September, high ozone is again observed in the boundary layer and free troposphere with a nearly constant profile of approximately 50 ppbv up to 8000 m. Both models detect well this increase with similar performance and a slight underestimation in the upper part of the free troposphere.

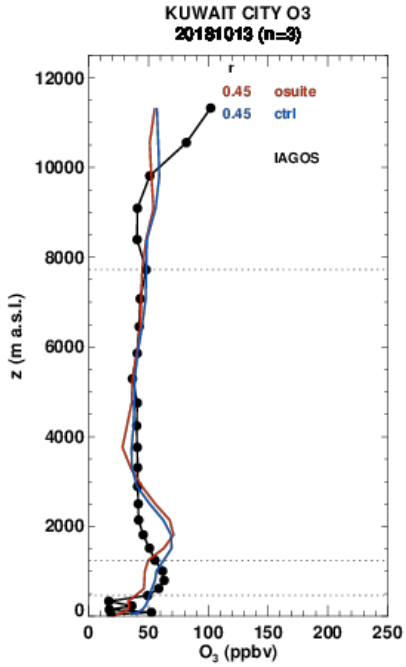


Figure 3.2.6 Profile for ozone at Kuwait City on 13 October 2018, from IAGOS (black) and the two NRT runs (o-suite: red, control: blue).

Around 13 October, a peak in ozone is observed in the mid-troposphere between 4000 and 6000 m (Fig. 3.2.4). On the profile of 12 October ozone maximum is observed near 5000 m with 60 ppbv, and at a higher altitude near 6000 m on 13 October. On 14 October, the sharp maximum of the two previous days is not observed and ozone values are of 60 ppbv in the range of altitudes between 4000 and 6000 m. For these days, the models in general detect the presence of these peaks although sometimes at a higher altitude as it is the case on 12 October and with a slightly better performance from the o-suite compared to control run.

A new French airport has been visited during SON 2018, with two flights involving Bordeaux. The ozone profile observed on 3 September (Fig. 3.2.5) presents a maximum in ozone between 2500 m and 3500 m with about 70 ppbv while outside of this range of altitudes values are of about 50 ppbv. This pattern appears correlated with high CO values at the same altitudes of about 230 ppv (see section 4.2). The shape of the ozone profile with the existence of maxima is not reproduced by the models which present a nearly constant profile of about 60 ppbv.

Middle East: On 13 October 2018, a profile is available at the airport of Kuwait city (Fig. 3.2.6). This profile shows a maximum in ozone in the boundary layer with values around 50ppbv. This feature is reproduced with a similar magnitude by both models but at higher altitude near 2000 m.

3.3 Validation with GAW and ESRL-GMD surface observations

For the Near Real Time (NRT) validation, 11 GAW stations and 14 ESRL stations are currently delivering O_3 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 September to November 2018. The latest validation results based on GAW stations and based on ESRL observations can be found on the CAMS website, see section 1, table 1.2.

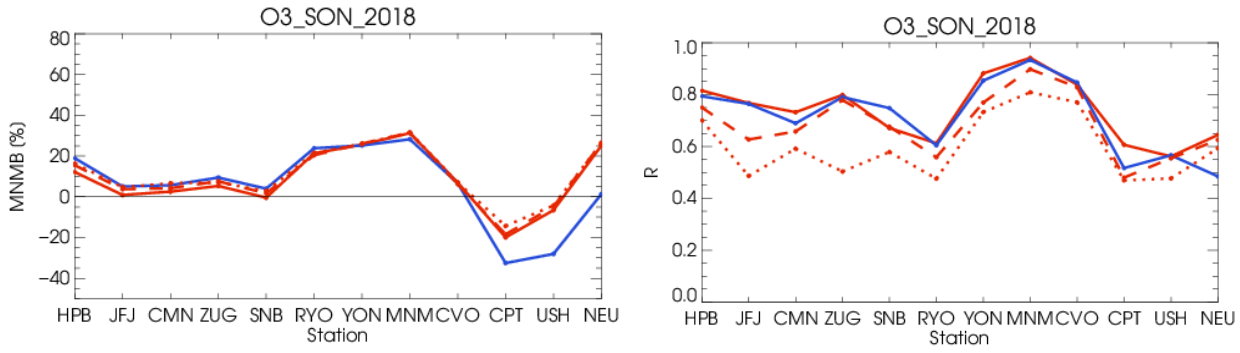


Figure 3.3.1: Modified normalized mean bias in % (left) and correlation coefficient (bottom right) of the NRT model runs compared to observational GAW data in the period September 2018 to November 2018 (o-suite: solid red, D+2: red-dashed, D+4: red-points, and control: blue).

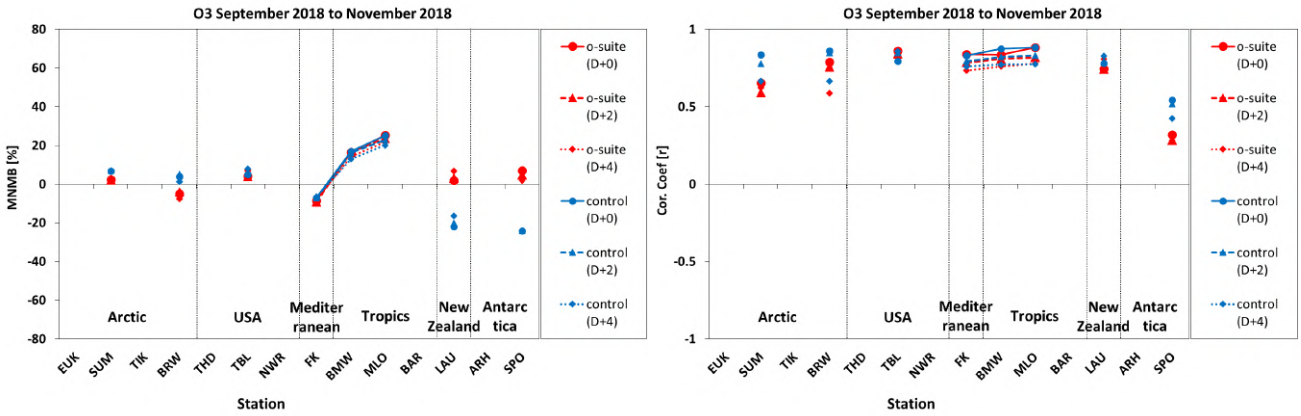


Figure 3.3.2: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT forecast runs compared to observational ESRL data in the period September to November 2018. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.

Modified normalized mean biases in % (left, panel) and correlation coefficients (right, panel) for different forecasts days (D+2, red-dashed and D+4, red- pointed) with respect to GAW and ESRL observations are shown in Figs. 3.3.1 and 3.3.2. It indicates that MNMBs for both o-suite and control run mostly remain stable up to D+4 (forecast run from 96h to 120h). Correlations between simulated and observed surface ozone values remain almost stable up to D+2 (forecast run from 48h to 72h), but then drop (correlations for D+4 are lower than correlations for D+2 and D+0), see Fig. 3.3.1 and 3.3.2, right graph).

A comparison of the seasonal-mean MNMB over Europe (Fig. 3.3.3) 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 control run. The temporal correlation is consistently better for control than for the o-suite, but the o-suite shows strong improvements recently. The GAW results are summarized in Figs 3.3.1 and 3.3.3.

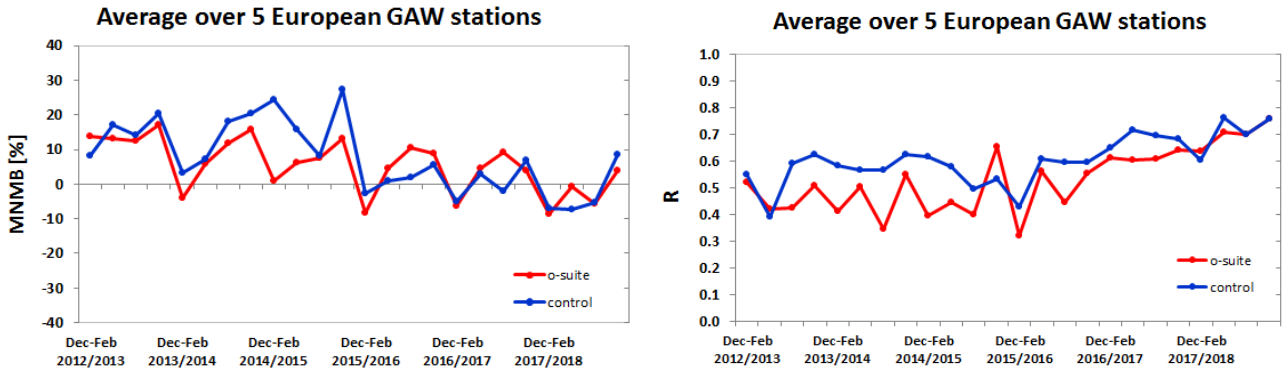


Figure 3.3.3: Long term (Dec. 2012 – November 2018) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

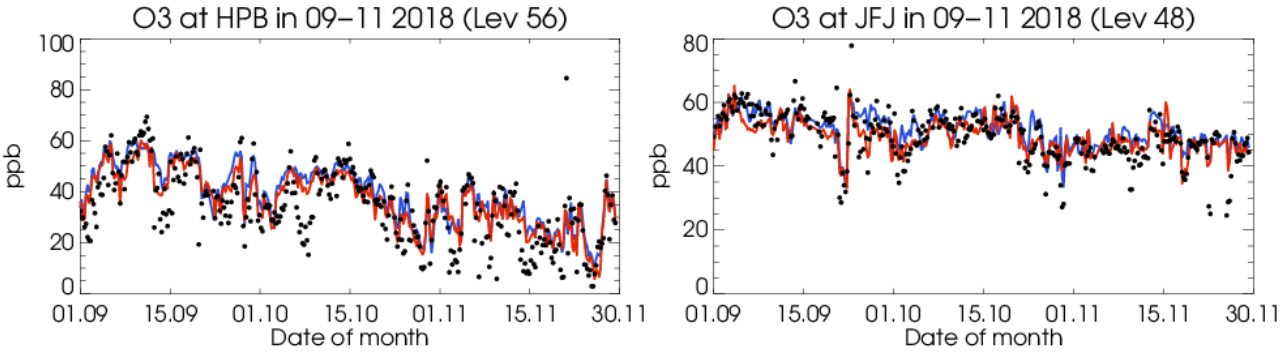


Figure 3.3.4: Time series for the o-suite (red) and control (blue) compared to GAW observations at Hohenpeissenberg (47.8°N, 11.02°E) and Jungfraujoch (39.03°N, 141.8°E)

Looking at different regions, for European stations (HPB, JFJ, ZUG, SNB, CMN, PRS), observed O_3 surface mixing ratios are very close to the observations. MNMBs are between 0 and 12% for the o-suite and between 5% and 18% for the control run, see Fig. 3.3.1. Correlations for the European stations are between 0.67 and 0.81 for the o-suite and between 0.68 and 0.79 for the control run, see Fig. 3.3.1.

Over Arctic stations (BRW and SUM), the o-suite overestimate surface ozone values by 2% at Summit and underestimates it at Point Barrow by -5%. On the other hand the control run overestimates surface ozone by 6% at Summit and by 4% at Point Barrow. Notable is that the control run reproduces better than the o-suite the day to day surface ozone variability over the Arctic stations ($r_{o\text{-}suite}\approx0.79$ and $r_{control}\approx0.86$ at BRW and $r_{o\text{-}suite}\approx0.65$ and $r_{control}\approx0.84$ at SUM), see Fig. 3.3.5.

For stations located in Asia (RYO, YON, MNM) both runs strongly overestimate O_3 mixing ratios with MNMBs between 21% and 31% for the o-suite and between 23% and 28% for the control run, see Fig 3.3.6. Correlation coefficients range between 0.60 and 0.94.

For TBL USA station, the observed ozone mixing ratios are slightly overestimated (MBMB≈5%) by both the o-suite and the control run. Correlations between modelled and observed surface ozone values are 0.85 for the o-suite and 0.80 for the control run.

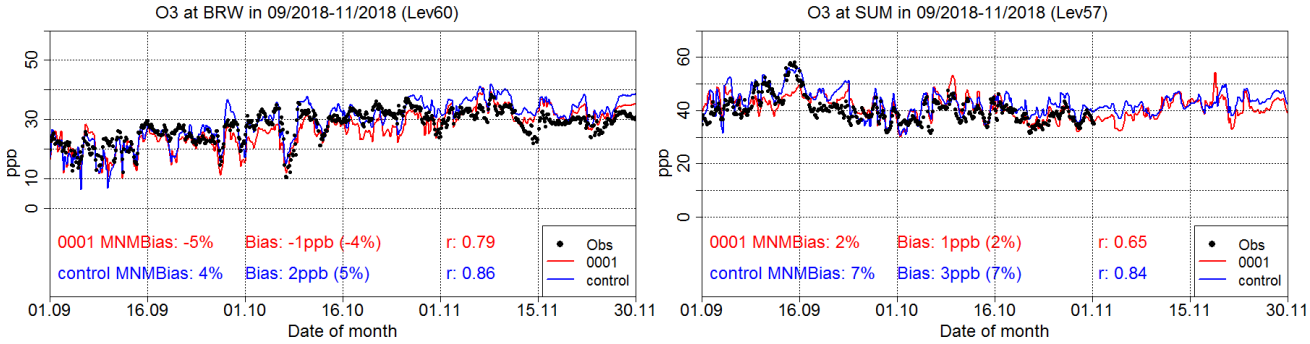


Figure 3.3.5: Time series for the o-suite (red) and control (blue) compared to ESRL observations at Summit, Greenland station (72.57°N, 38.48°W, left) and at Point Barrow, Alaska station (71.32°N, 156.51°W, right)

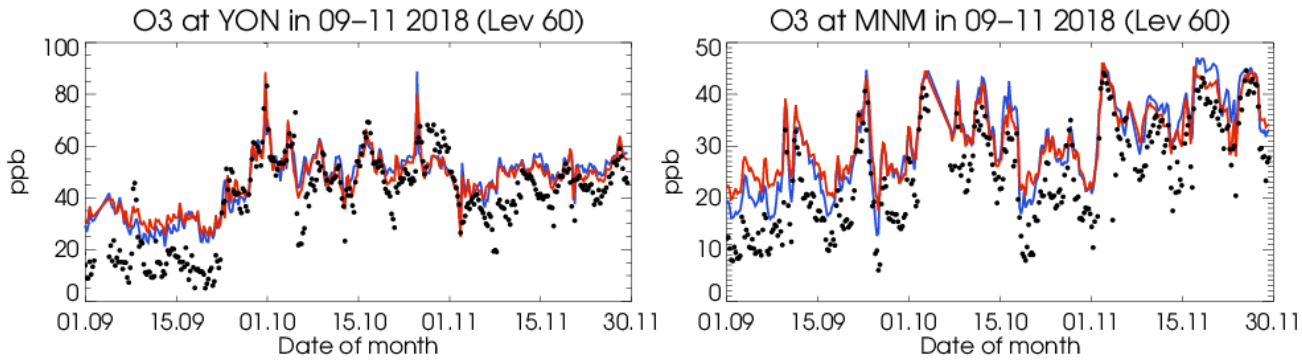


Figure 3.3.6: Time series for the o-suite (red) and control (blue) compared to GAW observations for Yonagunijima (24.47°N, 123.02°E) and Minamitorishima (24.29°N, 153.98°E)

The observed ozone mixing ratios are strongly overestimated by both runs over Mauna Loa (MLO) station in the Tropics ($MNMB_{o\text{-}suite}$ and $MNMB_{control}$≈25%). Both runs also overestimate surface ozone values by 15% at Bermuda (BMW) station. However correlations between simulated and observed surface ozone are high for both the o-suite and the control run over both Bermuda and Mauna Loa stations (r>0.85).

O_3 mixing ratios of the southern hemispheric stations (CPT, USH) are underestimated with MNMBs between -6 and -19% by the o-suite. The control run shows larger underestimations up to -32%, see Fig 3.3.8. Correlation coefficients range around 0.51 and 0.6. At Lauder (LDR) station in New Zealand the o-suite reproduces well O_3 mixing ratios (MNMB≈0%) while the control run underestimates it by -20%. Correlations between simulated and observed surface ozone values are 0.75 for both runs.

Finally, for South Pole station in Antarctica (SPO), the data assimilation corrects the negative offset in the control run (o-suite MNMBs= 6%, control MNMB= -20%), see Fig 3.3.7 (right panel). Correlations between simulated and observed surface ozone at SPO station are 0.32 for the o-suite and 0.54 for the control run.

For Neumayer station (NEU) the MNMB is 25% for the o-suite. The control run is closer to the observations for this period with MNMB 1% but shows lower correlation, Fig. 3.3.8.

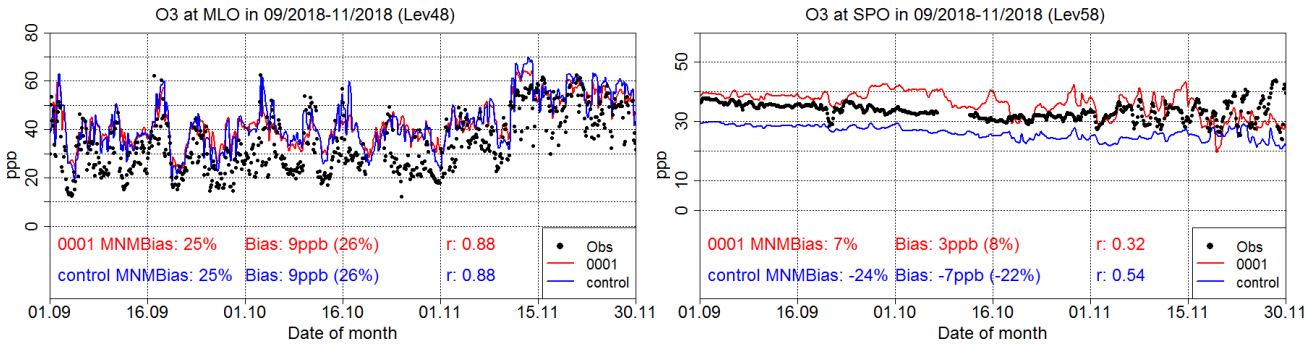


Figure 3.3.7: Time series for the o-suite (red) and control (blue) compared to ESRL observations (black dots) at Maouna Loa, Hawaii station (19.54°N, 155.58°W) and at South Pole, Antarctica station (90.00°S, 24.80°W).

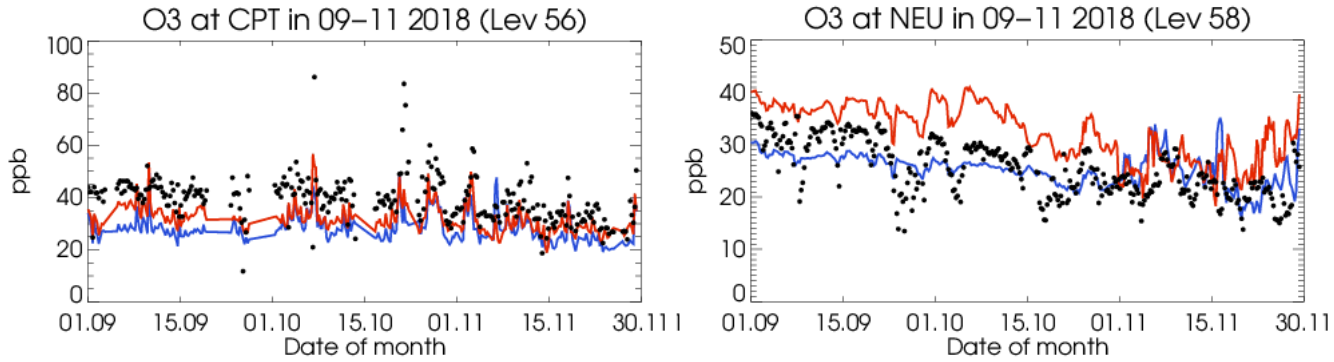


Figure 3.3.8: Time series for the o-suite (red) and control (blue) compared to GAW observations (black dots) at Cape Point (34.55°S, 18.48°W) and GAW observations at Neumayer (70.65°S, 8.25°W).

3.4 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, 2 stations from the Department of Labour Inspection - Ministry of Labour and Social Insurance, of Cyprus (http://www.airquality.dli.mlsi.gov.cy/) are used in the validation analysis. For the validation analysis, stations in the Mediterranean located within about 100 km from the shoreline of the Mediterranean shore are used. Table 3.4.1 shows the names, coordinates, elevation and the MNMBs and correlations obtained with the 2 forecast runs (o-suite and control). It indicates that the variance explained by each station of both the o-suite and control is high and correlations are highly significant over Western, Central and Eastern Mediterranean. It should be noted that the control run reproduces slightly better than the o-suite the surface ozone day to day variability over the Eastern Mediterranean (see Table 3.4.1).

The time series are shown in Fig. 3.4.1. In terms of biases, o-suite MNMBs vary between -20% and 30% depending on the stations over the Mediterranean shore of Spain. For the control run MNMBs are on average 3.0% higher than o-suite MNMBs. Over the stations Plan Aups/Ste Baume in France and Gharb in Malta the o-suite overestimate surface ozone concentrations by 10% and 2% respectively. Again, for the control run MNMBs are slightly higher ≈2% than for the o-suite. Over

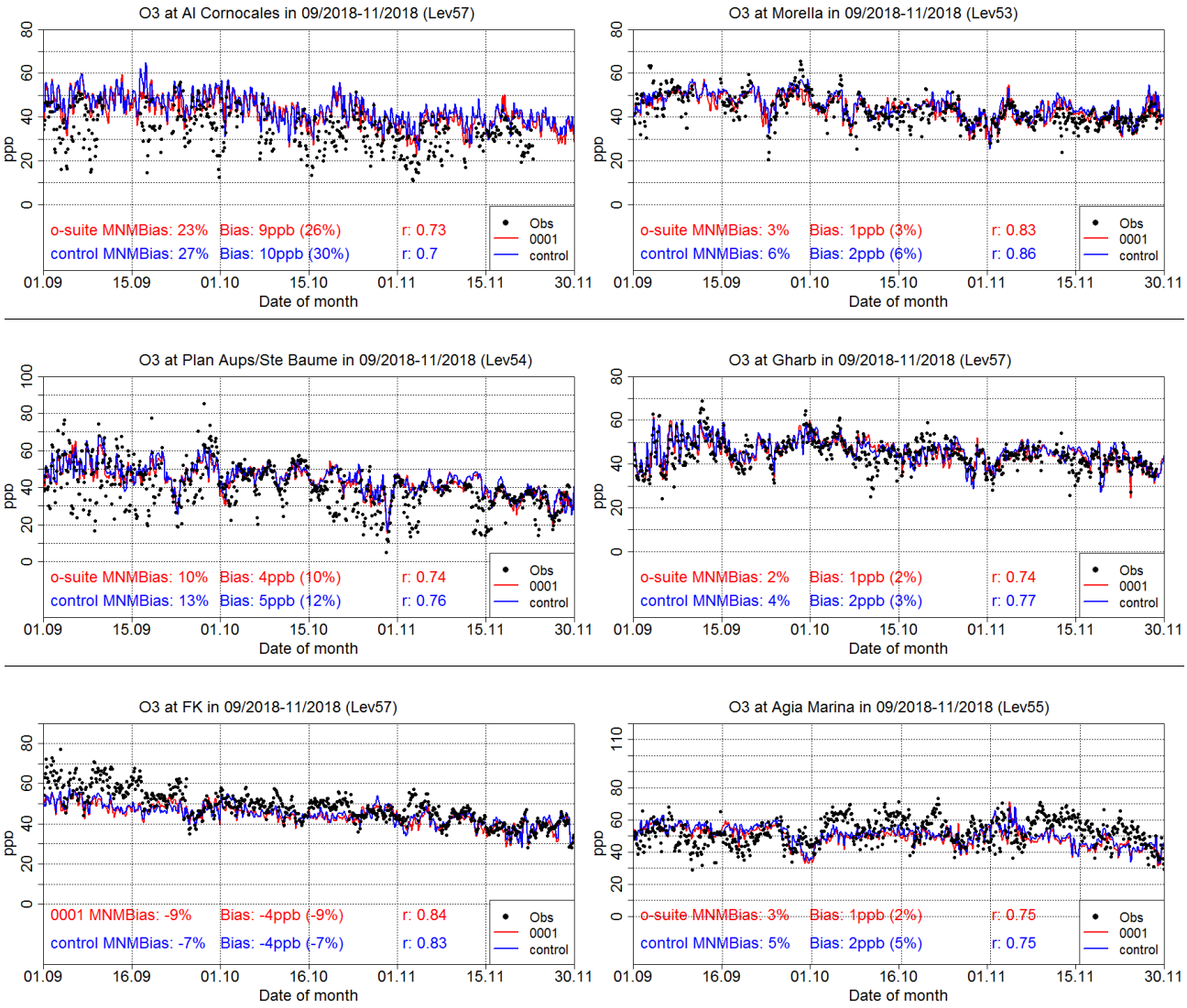


Figure 3.4.1: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Cornocales, Spain station (36.23°N, 5.66 °W, top left), at Morrela, Spain station (40.64°N, 0.09°W, top right), at Plan Aups/Ste Baume, France station (43.34°N, 5.73°E, center left), at Gharb, Malta station (36.07°N, 14.20°E, centre right at Finokalia, Crete Greece station (35.32°N, 25.67°E, bottom left) and compared to observations provided by the Department of Labour Inspection - Ministry of Labour and Social Insurance of Cyprus) at Agia Marina, Cyprus station (35.04°N, 33.06 °E, low right).

Finokalia station in Crete both the o-suite and the control run underestimate surface ozone by -8.5% and -7% respectively. Finally, over Agia Marina and Oros Troodos stations in Cyprus the o-suite has almost zero bias. The Control's MNMBs are on average 3% higher than o-suite MNMBs.

The spatial distribution of MNMBs and the correlation coefficients of the o-suite over the Mediterranean are shown in Fig. 3.4.2, where it is evident that correlations over the entire Mediterranean from Gibraltar to Cyprus are highly significant. It is also evident that the CAMS NRT runs have a better performance over Central and eastern Mediterranean compared to the Mediterranean shore of Spain in terms of biases.

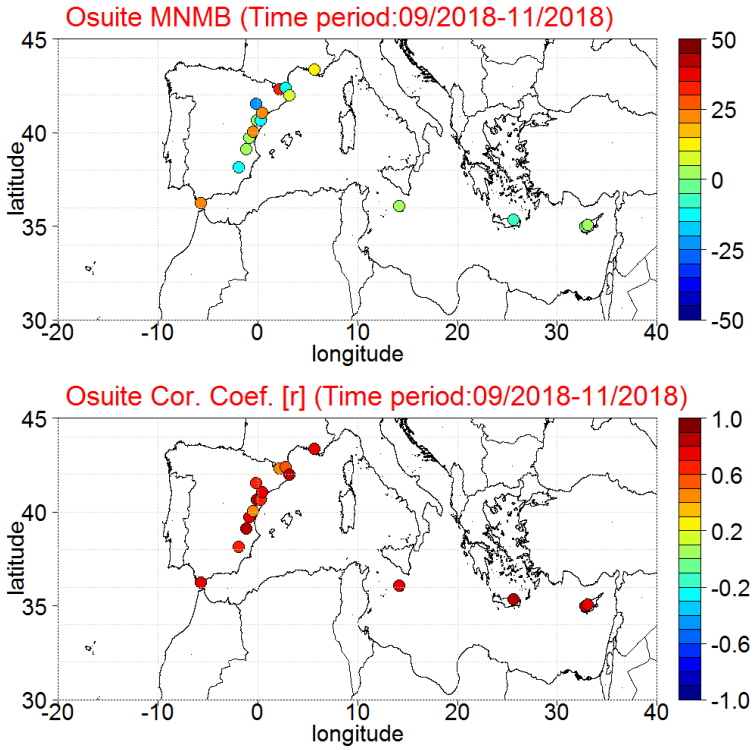


Figure 3.4.2: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 September 2018 to 30 November 2018.

Table 3.4.1: Coordinates, elevation, corresponding model level (level 60 is the surface level), as well as validation scores (MNMBs and correlations for the period SON 2018) 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.

						Distance from the	MNMB		Cor. Coef	
Station Name	Stat_ID	Lon	Lat	Alt (m)	Level	shore (km)	o-suite	control	o-suite	control
Al Cornocales	ES1648A	-5.66	36.23	189	57	16	23.5	27.3	0.73	0.70
Caravaka	ES1882A	-1.87	38.12	1	60	73	-11.2	-8.8	0.62	0.62
Zarra	ES0012R	-1.10	39.08	885	56	70	4.1	6.4	0.82	0.81
VIillar Del Arzobispo	ES1671A	-0.83	39.71	430	60	48	1.4	5.3	0.76	0.76
Cirat	ES1689A	-0.47	40.05	466	60	37	24.7	28.1	0.45	0.46
Bujaraloz	ES1400A	-0.15	41.51	327	60	60	-21.5	-2.2	0.67	0.76
Morella	ES1441A	-0.09	40.64	1150	53	51	3.1	5.8	0.83	0.86
Bc-La Senia	ES1754A	0.29	40.64	428	59	21	-12.3	-7.1	0.63	0.65
Ay-Gandesa	ES1379A	0.44	41.06	368	58	15	23.5	26.4	0.73	0.76
Ak-Pardines	ES1310A	2.21	42.31	1226	57	81	32.3	35.1	0.48	0.53
Hospital Joan March	ES1827A	2.69	39.68	172	57	3	NA	NA	NA	NA
Al-Agullana	ES1201A	2.84	42.39	214	60	25	-14.8	-12.7	0.58	0.61
Av-Begur	ES1311A	3.21	41.96	200	56	9	7.9	11.3	0.81	0.81
Plan Aups/Ste Baume	FR03027	5.73	43.34	675	54	21	10.3	12.7	0.74	0.76
Gharb	MT00007	14.20	36.07	114	57	31	2.1	3.5	0.74	0.77
Aliartos	GR0001R	23.11	38.37	110	59	18	NA	NA	NA	NA
NEO	-	21.67	37.00	50	60	2	NA	NA	NA	NA
Finokalia	GR0002R	25.67	35.32	250	57	4	-8.6	-6.9	0.84	0.83
Oros Troodos	-	32.86	34.95	1819	49	11	-2.1	1.3	0.88	0.89
Agia Marina	CY0002R	33.06	35.04	532	55	14	2.1	5.4	0.75	0.75

3.5 Validation with IASOA surface observations

Model results were compared to surface O_3 observations from the Villum Research Station, Station Nord in north Greenland (81.6°N 16.7°W), Alert Nunavut, Canada (82.5°N 62.5°W), Eureka, Nunavut, Canada (80.1°N 86.4°W), Zeppelin Mountain, Svalbard (78.9°N 11.9°E), and Tiksi, Russia (71.6°N 128.9°E) from the IASOA network, Fig. 3.5.1.

The data from Svalbard and VRS are covering the period from December 2014 to November 2018. Data from Eureka covers the period August 2016 – November 2018, data from Alert covers the period January 2016 – November 2018, and data from Tiksi covers the period September 2016 – mid-October 2018. Ozone depletion events in March – June in 2015 – 2018 are not captured by the model simulations during spring at any of the sites. These events are related to halogen chemistry reactions that are not represented in the model simulations. The simulations are on average in good agreement with the observations apart from the spring depletion events.

For the period September – November 2018 the measurements are not quality controlled. The bias is low with a tendency for underestimation for the o-suite for the four stations with full coverage (-9% - -1%) and overestimation for the control run (2% - 6%) (Table 3.5.1). The bias is generally lower for this autumn season compared to previous years (Fig. 3.5.1). The control run performs better than the o-suite in terms of the correlation; r = 0.27 – 0.80 for the o-suite compared to r = 0.45 – 0.86 for the control run.

Table 3.5.1. Normalised Mean Bias (NMB) and correlation coefficient (r) of the Control and the O-suite simulations for the sites Alert, Eureka, Svalbard, Tiksi and Villum Research Station (VRS) for the period September – November 2018.

		NMB	R
Alert	o-suite	-0.05	0.80
	control	0.02	0.86
Eureka	o-suite	-0.01	0.27
	control	0.03	0.51
Svalbard	o-suite	-0.09	0.64
	control	0.06	0.73
Tiksi	o-suite	0.04*	0.40*
	control	0.07*	0.45*
VRS	o-suite	-0.05	0.63
	control	0.02	0.76

* Only data for 1.5 months

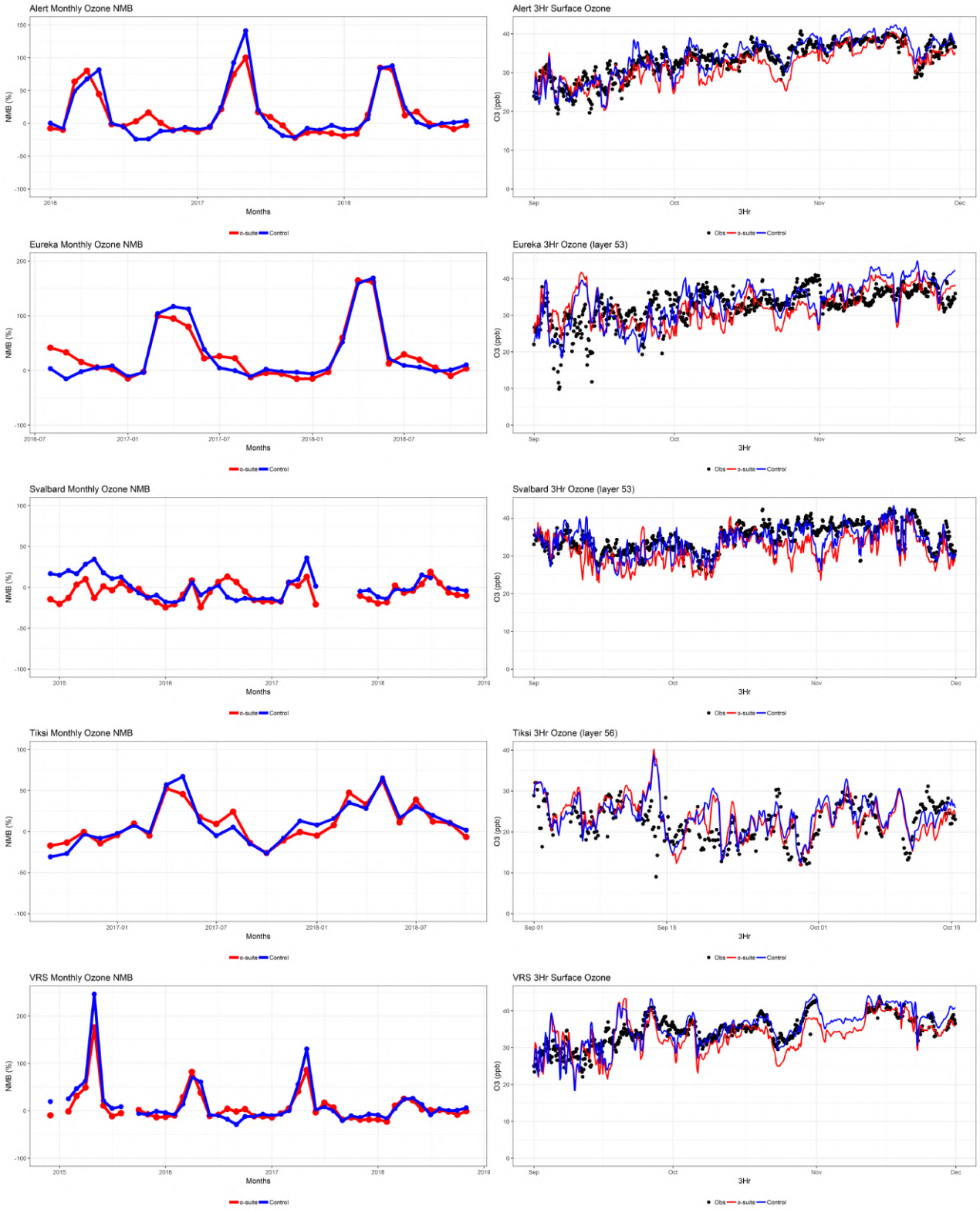


Figure 3.5.1: Time series for o-suite (red) and Control (blue) compared to observations (black dots) at Alert, Nunavut, Canada (Top row) Eureka, Nunavut, Canada (second row), Svalbard (third row), Tiksi, Russia (fourth row) and the Villum Research Station, Station Nord, Greenland (bottom row) NMB for the full period (left) and concentrations for September-November (right).

4. Carbon monoxide

4.1 Validation with Global Atmosphere Watch (GAW) Surface Observations

For the Near-Real-Time (NRT) validation, 10 GAW stations have delivered CO surface mixing ratios in NRT and data is compared to model results as described in Douros et al (2017) and is used for CAMS model evaluation for September – November 2018. The latest validation results can be found on the CAMS website, see section 1.

For stations in the Northern Hemisphere, both runs mostly show slightly negative MNMBs. For the stations located in the Southern Hemisphere, both runs show a strong positive offset, which is reduced by the data assimilation for the o-suite.

For most stations, the MNMBs and correlation coefficients indicate that the forecast remains stable till the D+4 (forecast run from 96h to 120h). However, for the stations in the Southern Midlatitudes, MNMBs and Rs increase already for D+2.

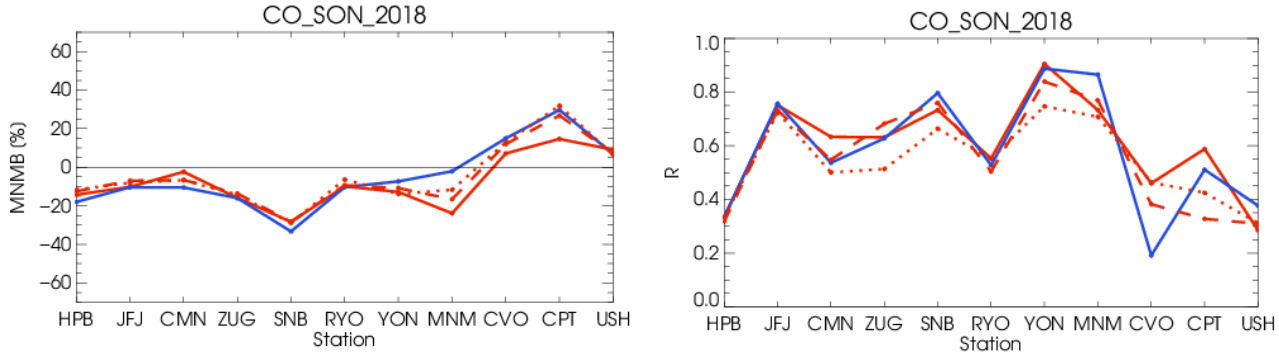


Figure 4.1.1: Modified normalized mean bias in % (left) and correlation coefficient (bottom right) of the NRT model runs compared to observational GAW data in the period September 2018 to November 2018 (o-suite: solid red, D+2: red-dashed, D+4: red-points, and control: blue).

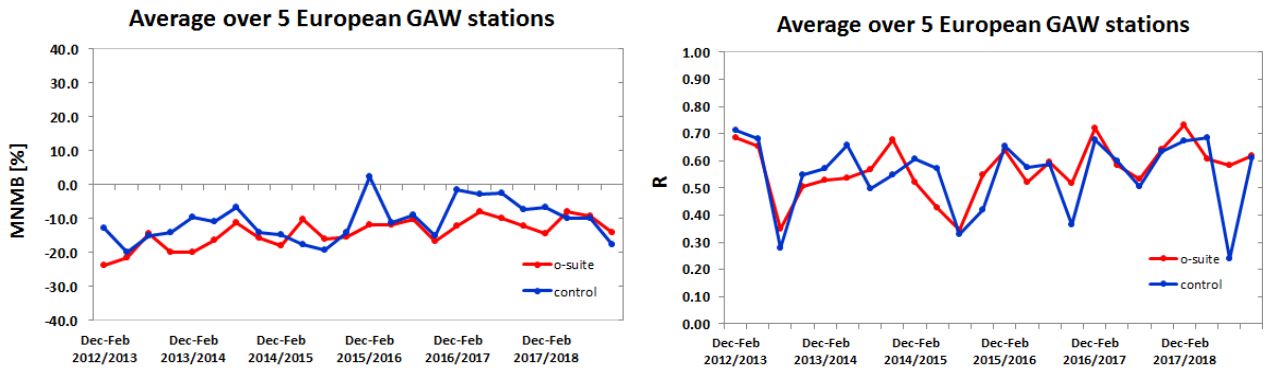


Figure 4.1.2: Long term (Dec. 2012 – November 2018) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

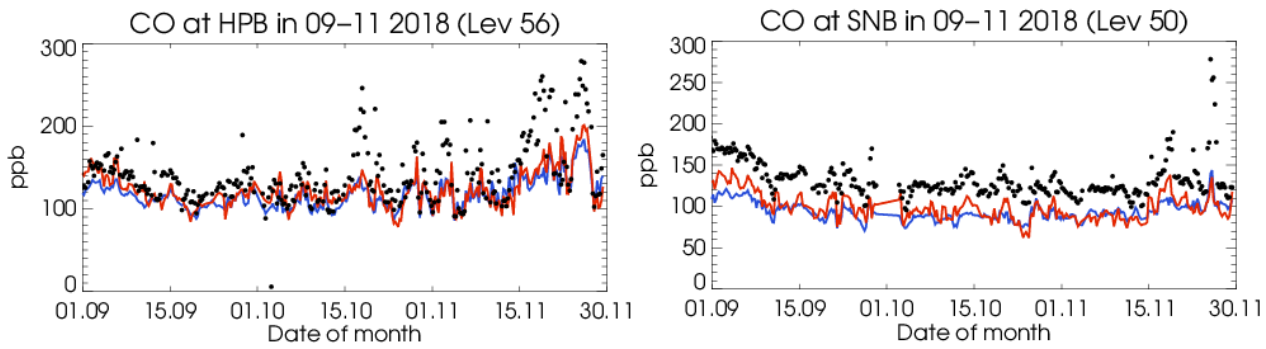


Figure 4.1.3: Time series for the o-suite (red) and control (blue) compared to GAW observations at Hohenpeissenberg (47.8°N, 11.02°E) and Sonnblick (47.05°N, 12.96°E)

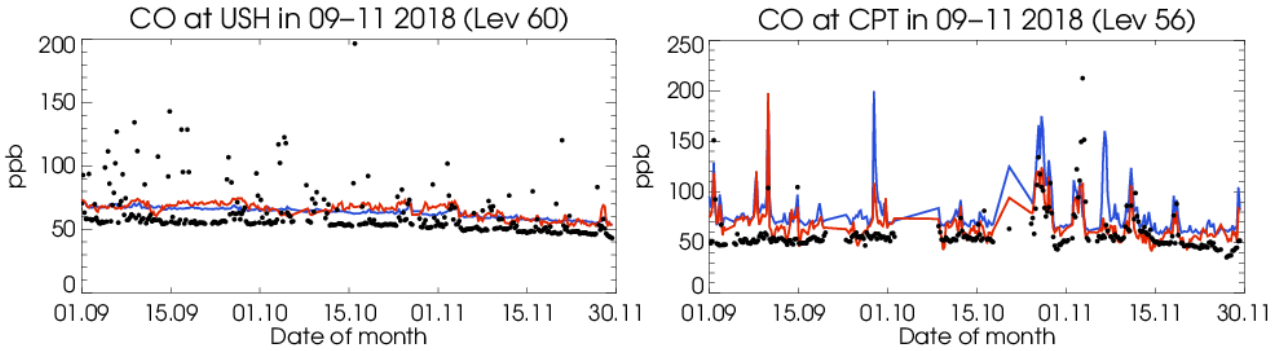


Figure 4.1.4: Time series for the o-suite (red) and control (blue) compared to GAW observations at Ushuaia (-54.85°N, -68.32°W) and Cape Point (34.35°S, 18.5°E).

A comparison of the seasonal-mean MNMB over Europe (Fig. 4.1.2) from December 2012 to present shows a slowly improving MNMB from about -20% in 2013 to -10% for more recent periods. Temporal correlation remains relatively constant at r=0.6 on average, except for the quarter JJA in 2018, where the correlation of the control run drops to 0.24.

For European stations, the o-suite shows an underestimation of observed CO mixing ratios, with MNMBs between -2% and -28%. The control shows slightly higher CO mixing ratios for European stations with MNMBs between -10% and -33%. Correlation coefficients are between 0.33 and 0.75 for the o-suite and between 0.33 and 0.79 for the control run.

For stations in Asia (RYO, YON, MNM) the control run shows a good correspondence with the observations with MNMBs between -9 and -23% for the o-suite and between -2 and -10% for the control run. The data assimilation thus reduces the CO mixing ratios and introduces a negative bias see Fig. 4.1.5. Correlation coefficients range between 0.55 and 0.93.

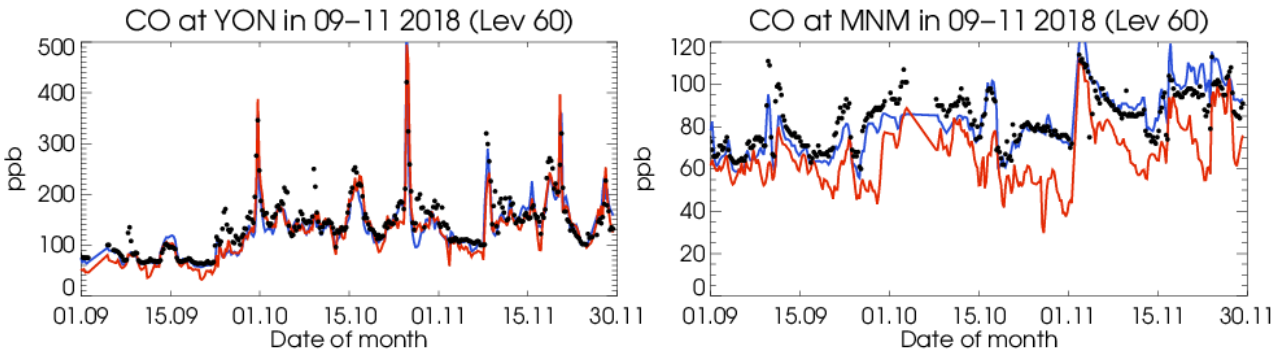


Figure 4.1.5: Time series for the o-suite (red) and control (blue) compared to GAW observations for Yonagunijima (24.47°N, 132.02°E) and Minamitorishima (24.29°N, 123.02°E).

4.2 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 September-November 2018, data from several aircraft have been validated, as discussed in Sec. 3.2.

Figure 4.2.1 shows the time series of daily mean values in different atmospheric layers over Frankfurt. At Frankfurt, the models underestimate CO from the surface to the upper troposphere (Fig. 4.2.2). The best agreement between models and observations is found in the free and upper troposphere. In the low stratosphere, models underestimate the amount of CO during SON 2018.

A profile of CO over Bordeaux is available on 3 September (Fig. 4.2.3), which presents a complex shape with two CO maxima in the low and mid-troposphere of respectively 230 ppbv at about 3000 m and 150 ppbv at 4500 m. CAMS global only simulates the major peak in the low troposphere but at a lower altitude and with an underestimated magnitude with 180 ppbv. The control run does not detect any maximum and show a nearly constant CO profiles with about 110 ppbv.

At the beginning of September, the time series in Frankfurt show high CO mixing ratios in the boundary layer and free troposphere. On the 2nd of September, the profile at Frankfurt shows a slow steady decrease from 200 ppbv to 150 ppbv up to an altitude of 11000 m (Fig. 4.2.4). On the same day a similar profile is observed at Munich, as well as on the 3rd of September at Frankfurt. CAMS underestimates the mixing ratios especially in the surface and boundary layer, with a better performance from the o-suite as compared to control run which shows underestimations as large as 40% in the lowest layers.

Around the 10 of November there is a peak of CO in the UTLS with values reaching 120 ppbv as shown on the time series (Fig. 4.2.1). The profile on 12 November shows nearly constant profile of CO with mixing ratio of 100 ppbv, which is well represented by the models.

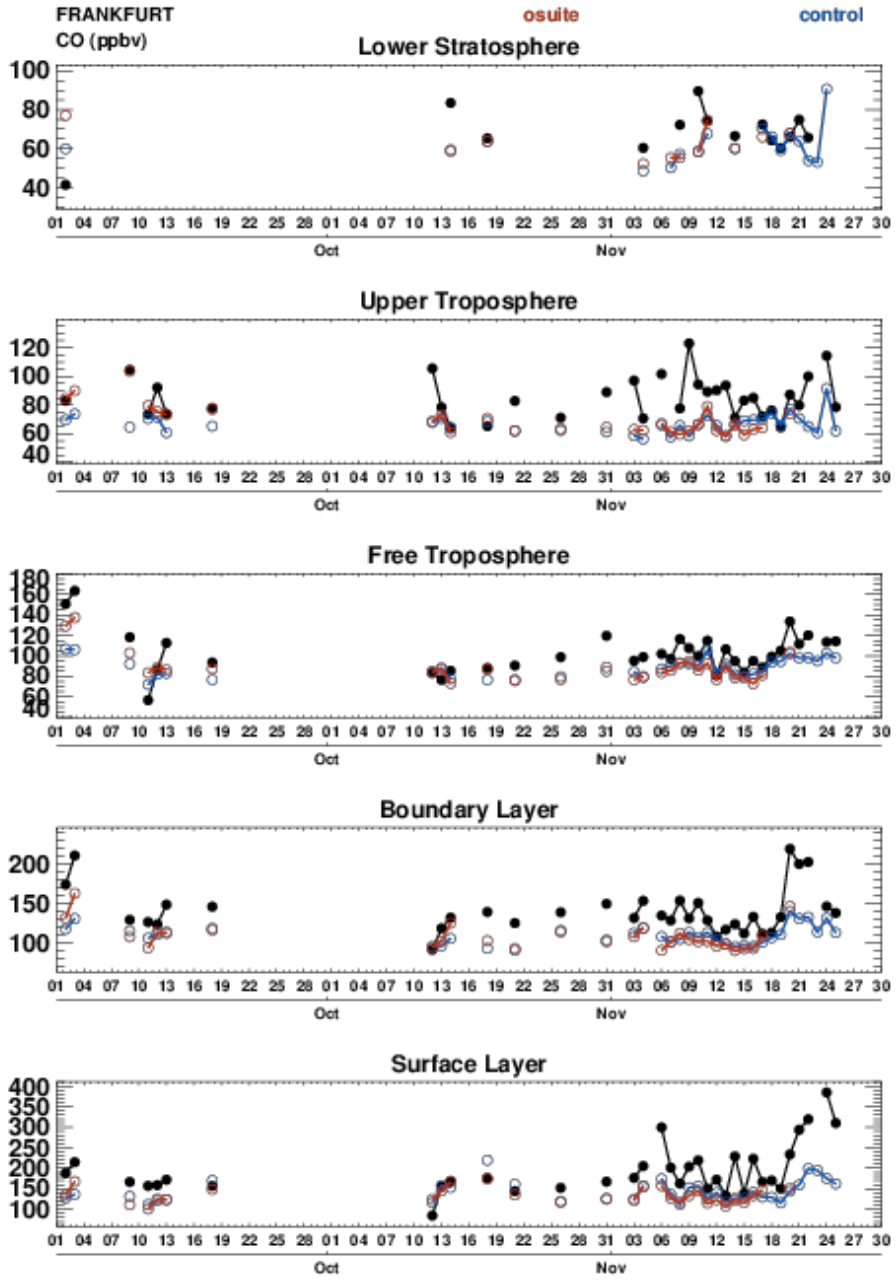


Figure 4.2.1. Time series of daily mean CO over Frankfurt during SON 2018 for 5 layers: Surface, Boundary layer, Free Troposphere, Upper Troposphere and Lower Stratosphere.

On 20 and 21 November high CO values are observed from the surface to the free troposphere especially in the surface and boundary layer with 300 ppbv near the surface and about 200 ppbv in the boundary layer on 21 November (Fig. 4.2.4). The models behave similarly and detect an increase but much smaller than observations and the values are largely underestimate with nearly a factor 2 near the surface and about 30% in the boundary layer. In the free troposphere and UTLS the agreement is better.

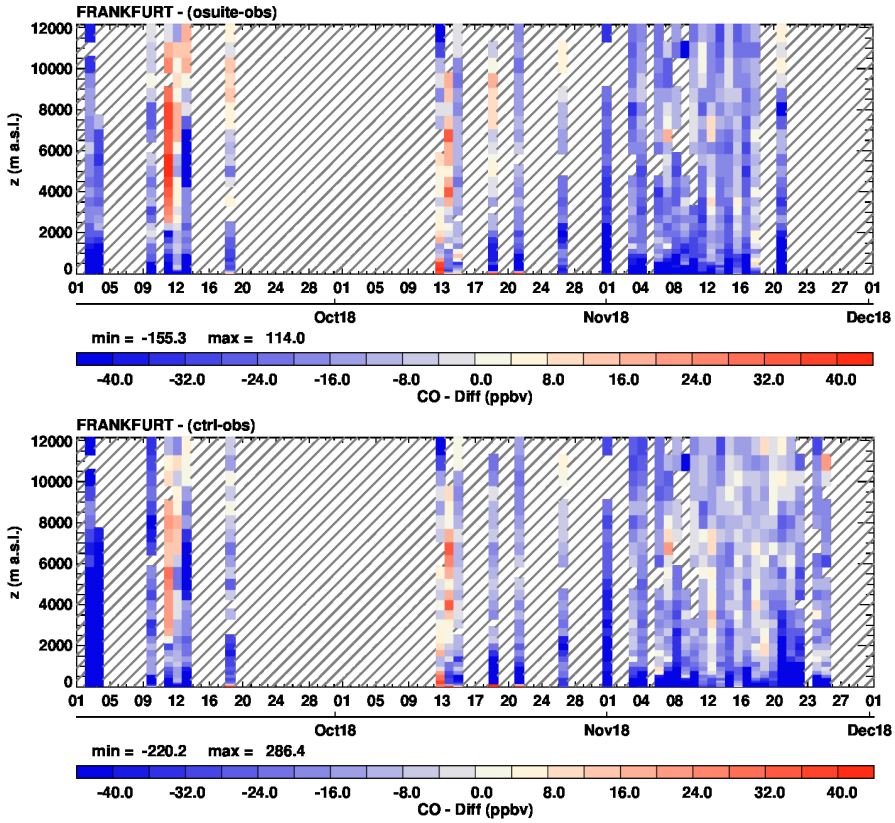


Figure 4.2.2. Time series of the absolute differences (model - observations) in daily profiles for CO over Frankfurt during SON 2018. Left panel corresponds to o-suite and right panel to control run.

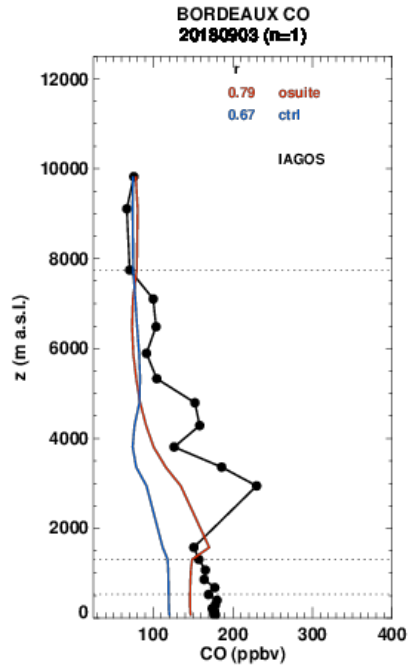


Figure 4.2.3. Profile for CO at Bordeaux on 3 September 2018, from IAGOS (black) and the two NRT runs (o-suite: red, control: blue).

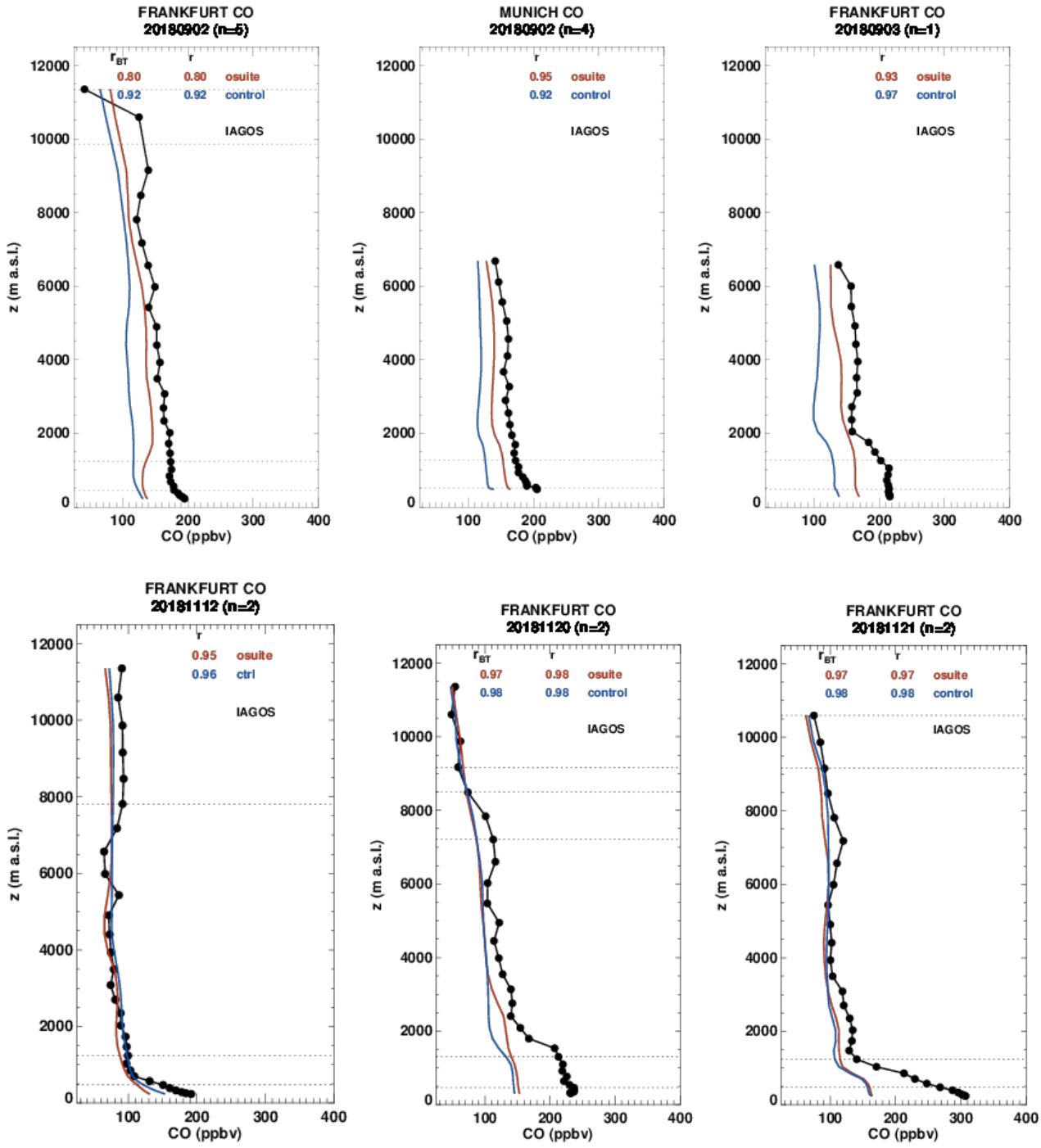


Figure 4.2.4. Selection of daily profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Europe during SON 2018.

West Africa

Fig. 4.2.5 highlights some examples of profiles for CO over West Africa sampled by Air France aircraft during SON 2018. The profiles presented here are for airports of Lagos (coastal location) and Malabo (on the island of Bioko) situated in the Gulf of Guinea. The cities in this region are affected by anthropogenic emissions from vehicles, oil industries, and from biomass burning which stretches

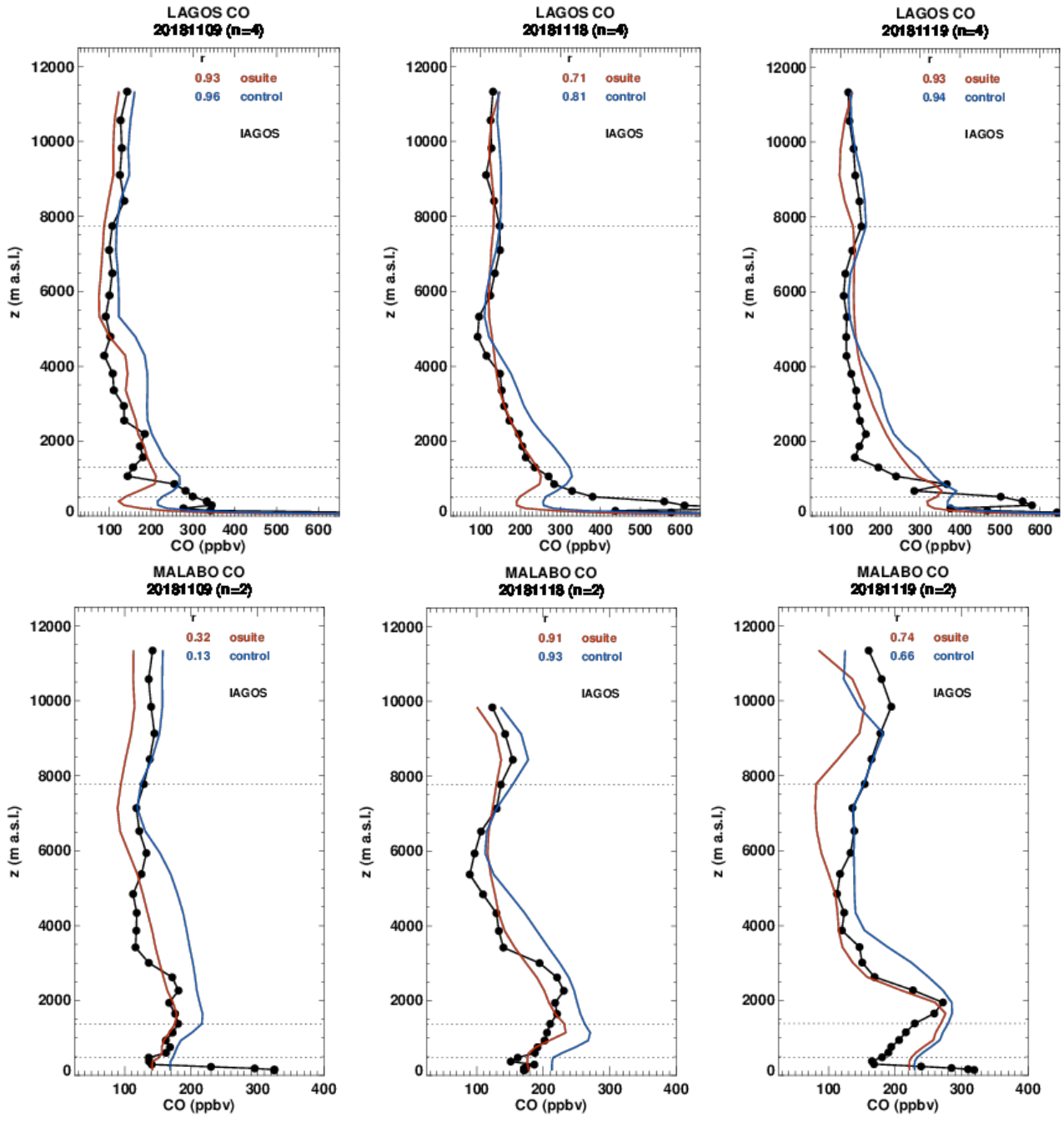


Figure 4.2.5. Selection of daily profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over West Africa during SON 2018.

across Africa just north of the equator from December to March and south of the equator from June to October. At the airports of Lagos and Malabo the profiles show that CO increases are generally seen at the correct altitudes by both o-suite and control run, which reproduces reasonably the shapes of the profiles. For both locations, in the boundary layer and free troposphere CAMS global performs better than control run which generally overestimates. Near the surface the results of both runs are rather similar in general. At Lagos, high surface CO (>300 ppbv) is largely overestimated by the models, while at Malabo it is underestimated by about 30%. In the UTLS the

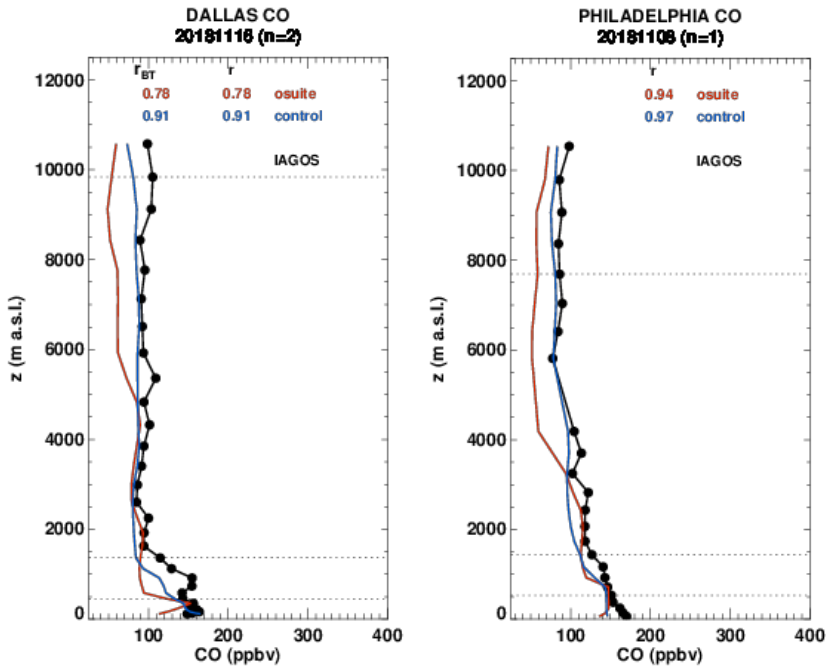


Figure 4.2.6. Selection of daily profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during SON 2018.

results of both models are similar at Lagos with a reasonable agreement. At Malabo, the shape of the profile in the UTLS present maxima reaching 200 ppbv near 9000 m which are well reproduced by the models at a correct altitude, but with some differences in the performances with in most cases a slightly better performance from the control run as o-suite present large underestimations (see 9 and 19 November).

North America

In North American airports similarly to Europe, it has been shown in previous reports that CO is usually underestimated in the surface and boundary layer by the two runs, while there is a good agreement in the free troposphere. In Fig. 4.2.6, two profiles are presented, one at Philadelphia and one profile at Dallas on 6 November and 16 November respectively. These profiles show low CO in the surface and boundary layer with about 150 ppbv, and nearly constant profile above the boundary layer of 100 ppbv. For these two profiles the models behave similarly in the low troposphere and show a rather good agreement. Above the mid-troposphere the control run shows better performance than the o-suite which underestimates.

Northeastern Asia

Fig. 4.2.7 and 4.2.8 show profiles at several locations in China (Shenyang, Wuhan and Taipei), Japan (Fukuoka). The profiles with complex shapes are in general well reproduced by the models with a usually better performance from the o-suite than from control run. In the surface layer mixing ratios are in general overestimated for observed values higher than 300 ppbv. Conversely in the boundary layer, the results from the models present large underestimations with a better performance from the o-suite than control run (with the exception of the 22 September at Taipei). In the mid-

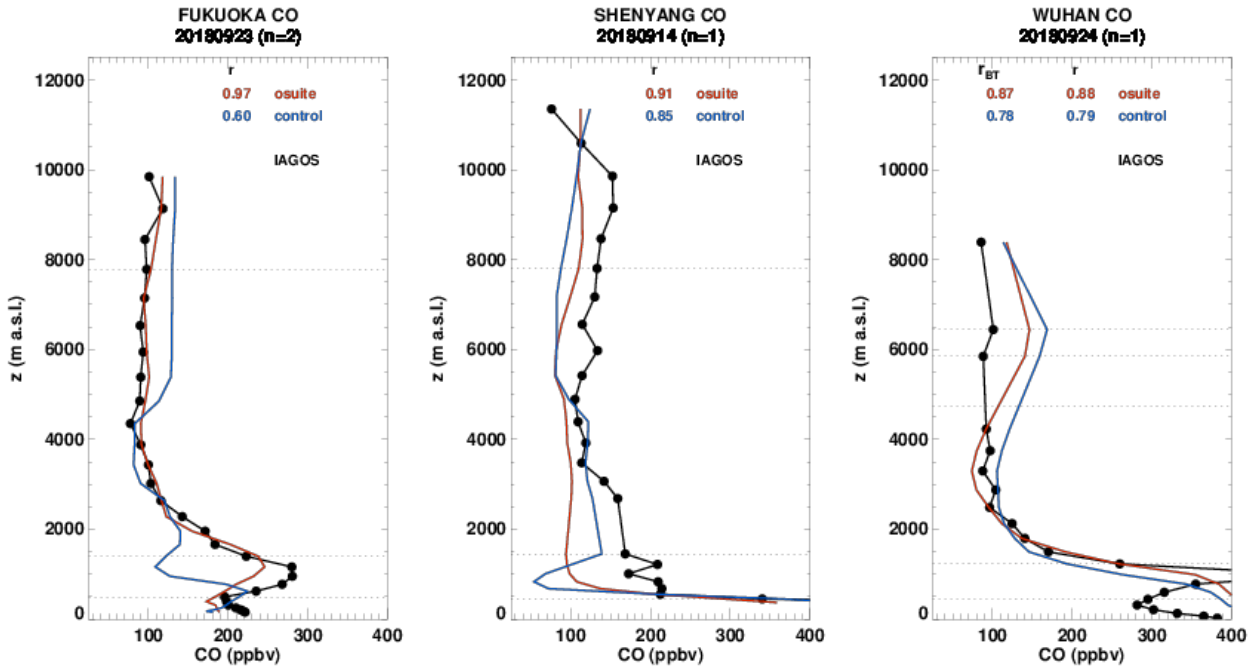


Figure 4.2.7. Selection of daily profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over north-east Asia during SON 2018.

troposphere the results from the two models are more similar and show large underestimations (about 30%) when CO values are high (>100 ppbv) such as for Taipei on 7 and 9 September (Fig. 4.2.7). In the UTLS the results are of the two runs are also similar and the agreement is the rather good for all these profiles except at Shenyang on 14 September (Fig. 4.2.6) where CO mixing ratio reaches 150 ppbv in this region of the atmosphere.

Southeastern Asia

In Fig. 4.2.8, two profiles located in south-east Asia are presented: one at the airport of Singapore on 15 September and the other at the airport of Surabaya on 9 September. For these profiles, the results from both runs are rather similar with some discrepancies in the boundary layer for the profile at Singapore. Surface values are largely overestimated by the models at both locations, while the agreement in the boundary layer is better and CO is underestimated. In the case of Singapore the profiles deviate substantially from observations below 4 km. In the lower part of the free troposphere high CO values are largely underestimated as it is the case in Singapore on 15 September with CO reaching more than 150 ppbv near 3000 m. The best agreement is found in the UTLS for both models with CO mixing ratios of about 100 ppbv.

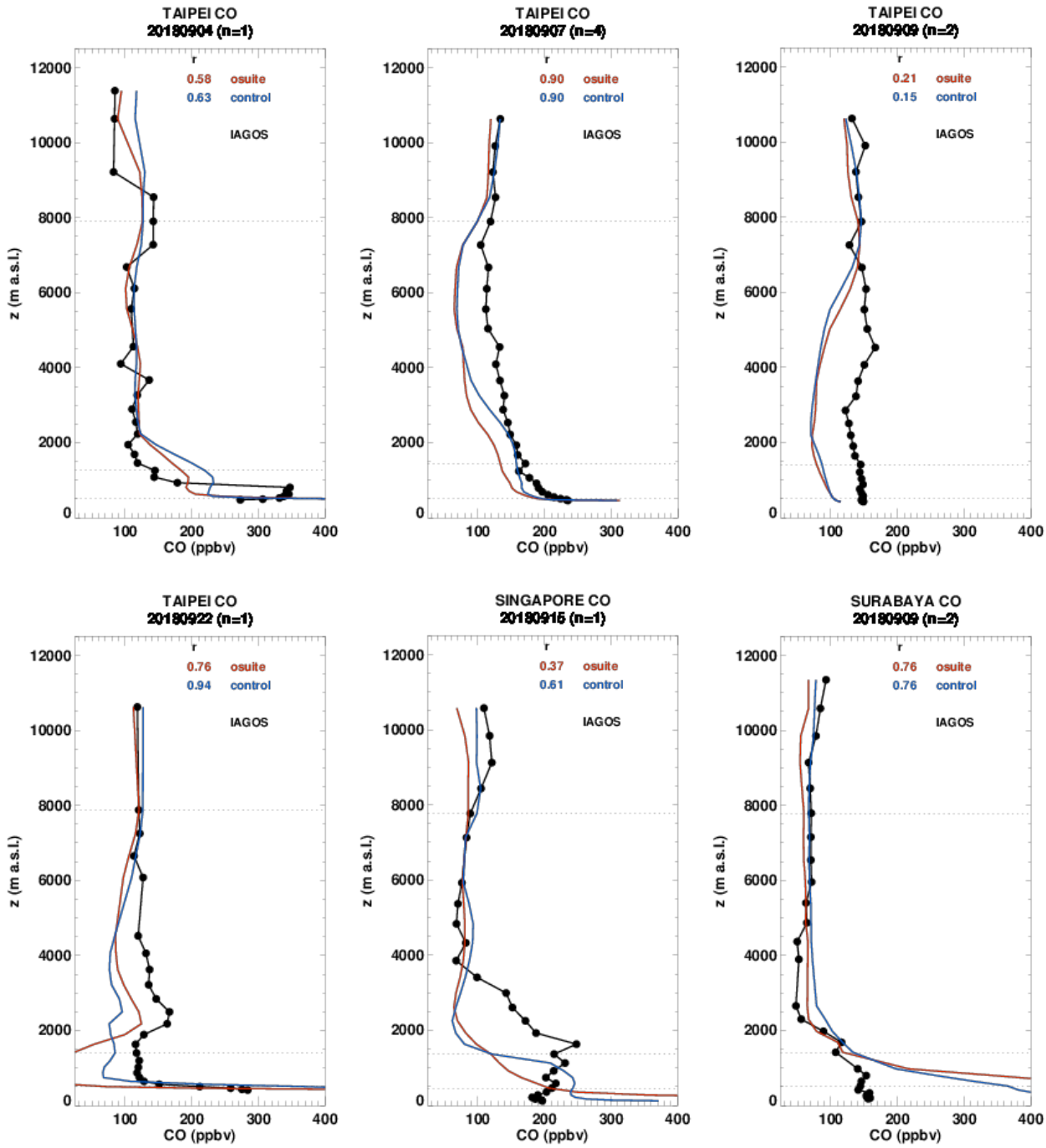


Figure 4.2.8. Selection of daily profiles for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over north-east (Taipei) and south-east Asia (Singapore, Surabaya) during SON 2018.

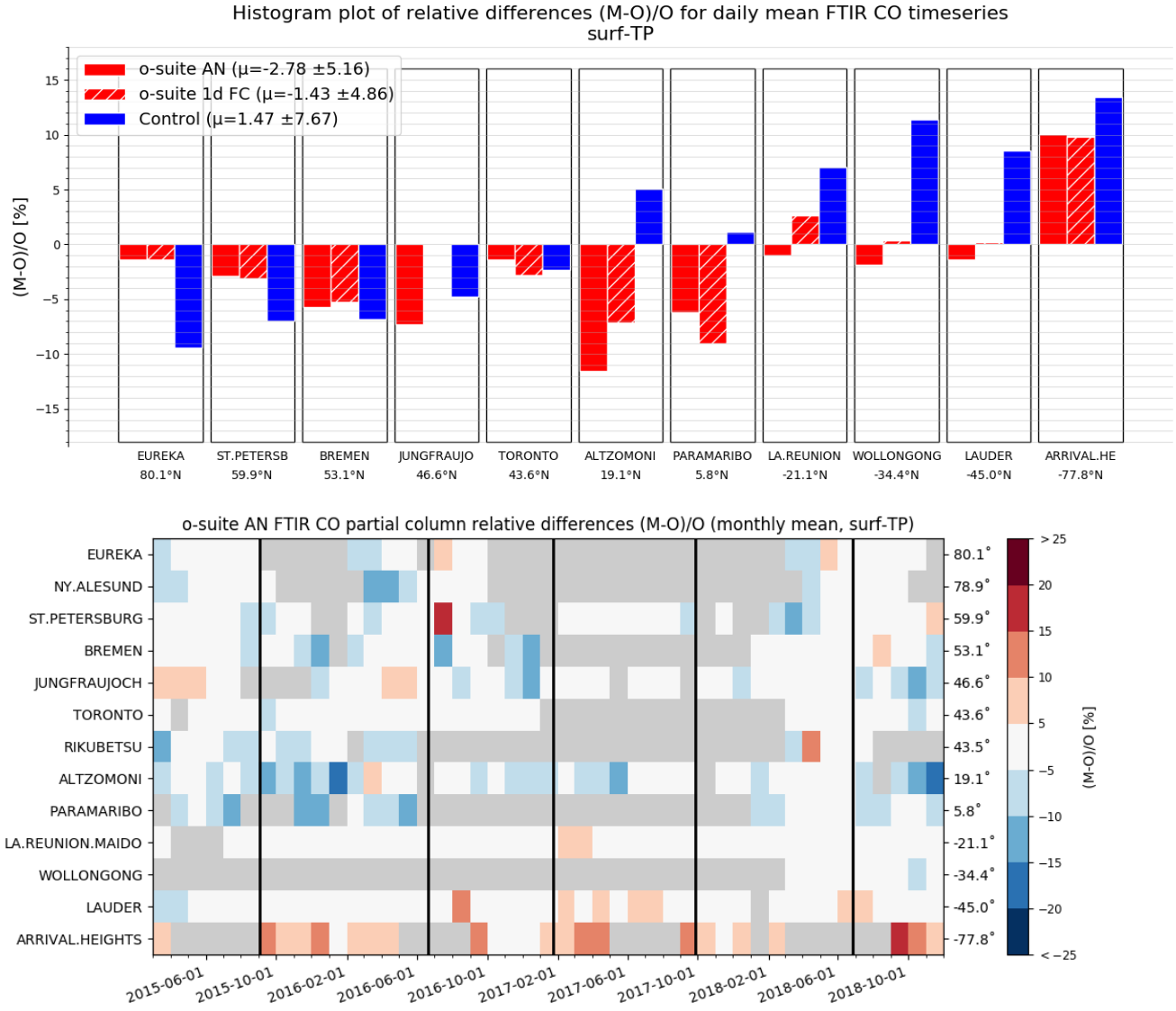


Figure 4.3.1: Seasonal relative mean bias for tropospheric CO columns (MB, %), standard deviation (STD, %) and number of observations used for the considered period SON 2018 (top) and monthly mean biases for a longer time period (bottom, model upgrades are indicated in black).The overall uncertainty for the CO measurements is approximately 5% and the o-suite averaged bias for all stations is -2.8% for SON. Stations are sorted with decreasing latitude (northern to southern hemisphere).

4.3 Validation against FTIR observations from the NDACC network

In this section, we compare the CO profiles of the CAMS models with FTIR measurements at different FTIR stations within the NDACC network. 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 columns are validated. A description of the instruments and applied methodologies can be found at http://nors.aeronomie.be.

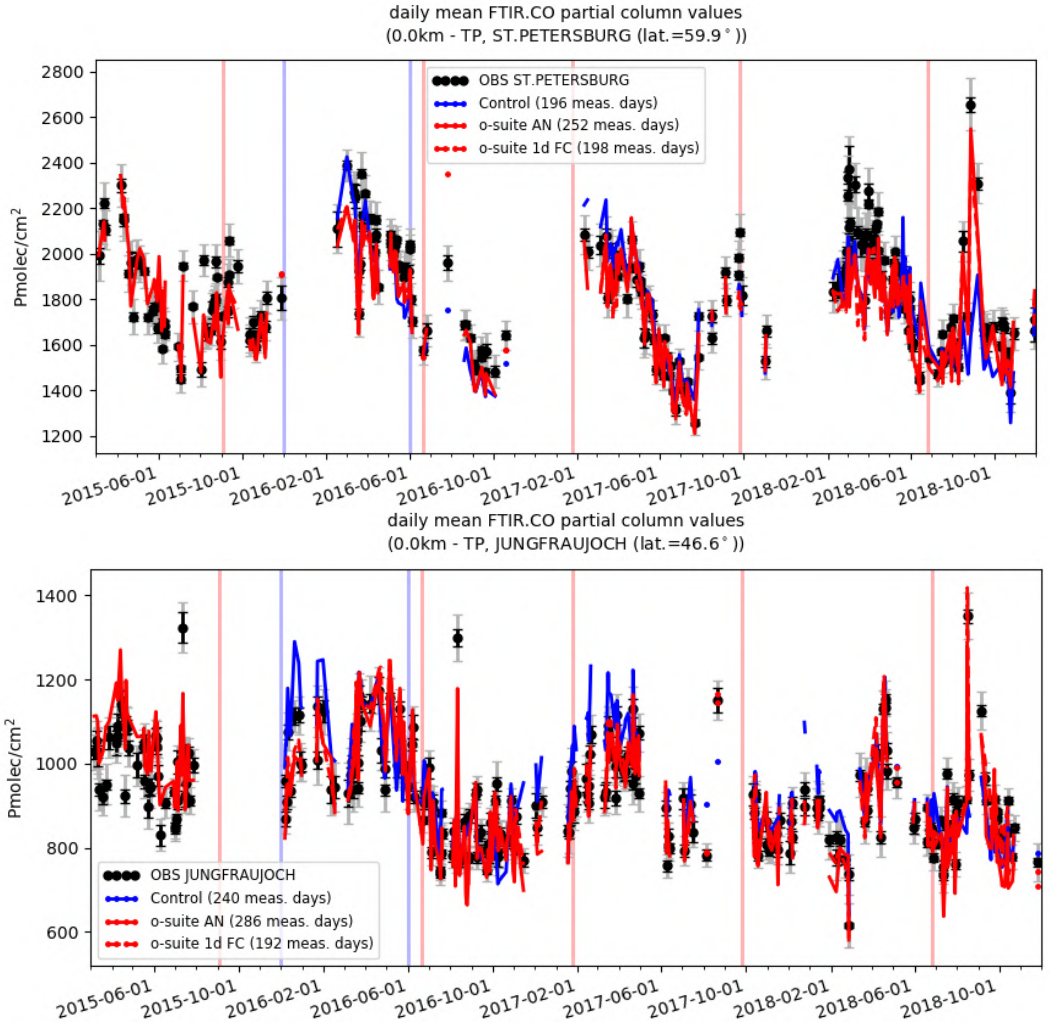


Figure 4.3.2: Daily mean values of tropospheric CO columns by the o-suite (AN and 1d FC, red) and the Control run (blue) compared to NDACC FTIR data at St Petersburg and Jungfraujoch for the period March 2015-November 2018. During March 2018 the o-suite underestimated the CO columns at St. Petersburg.

Figure 4.3.1 show that the o-suite tropospheric columns of CO agree well. All biases for the o-suite AN and 1d FC are within the measurement uncertainty. For the NH mid latitude stations, the bias becomes slightly negative (<-5%) in JJA-SON compared to the reporting period MAM 2018, which is related to a seasonal cycle observed in the relative differences, see Fig. 4.3.2 and Fig. 4.3.3. In the southern hemisphere, the bias is positive but remains within the 5% range.For all stations in the southern hemisphere, the control run overestimates the CO with MBs up to 13%.

The Taylor diagrams in Figure 4.3.4 provide information on the correlation between three CAMS configurations (o-suite analysis, o-suite forecast and control) and the FTIR time series. Leaving out the sites with only few measurements, the assimilation has a positive effect on the correlation coefficient. Looking at the correlation values for the period SON 2018, the o-suite 1d FC (averaged correlation for all sites is 0.83) seems to perform slightly worse compared the o-suite AN (averaged correlation for all sites is 0.86). The variability in the FTIR tropospheric columns is in general smaller compared the variability of the corresponding model columns (approximately by a factor of 0.9).

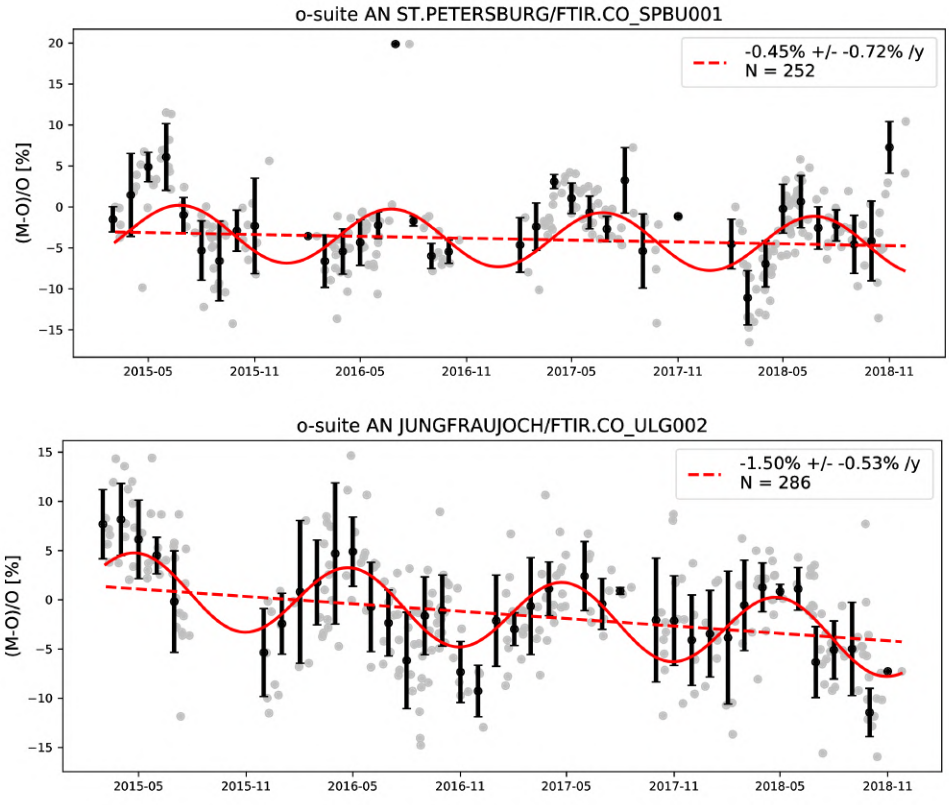


Figure 4.3.3: Fitted seasonal cycle through the monthly mean relative differences for the o-suite AN for the time series shown in Fig. 4.3.2. An underestimation is observed during the local autumn/winter months.

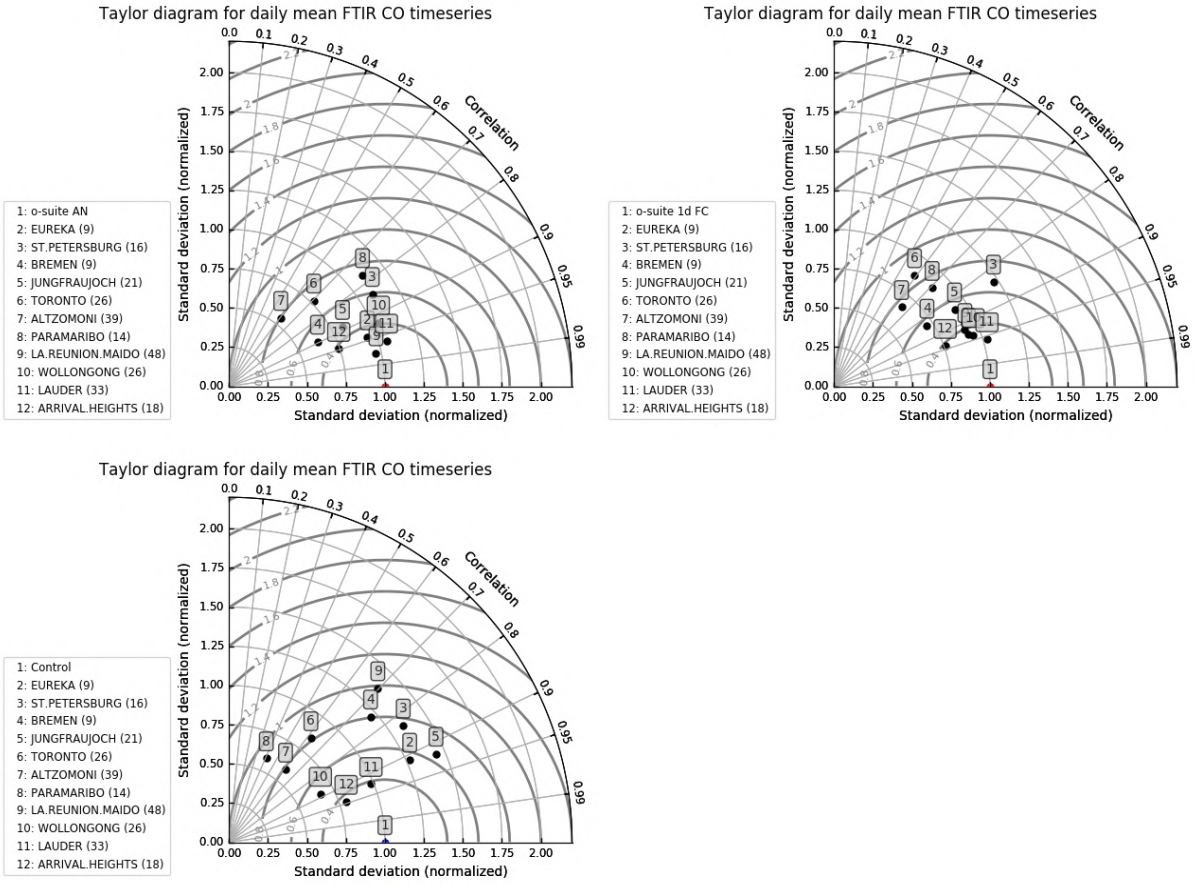


Figure 4.3.4: Taylor diagrams relating the standard deviations for the model versus ground-based time series of tropospheric CO column data and their correlation. All timeseries are normalized such that the standard deviation of the model is 1. The o-suite FC has higher variability in the CO columns compared to the FTIR variability. Top-left: o-suite analysis; top-right: o-suite 1-day forecast; bottom: control run.

4.4 Validation against FTIR observations from the TCCON network

CO column averaged mole fractions of the CAMS models is compared with data from the Total Carbon Column Observing Network (TCCON). For the validation column averaged mole fractions of CO (denoted as XCO) are used. Column averaged mole fractions provide different information content than the in-situ measurements and are therefore complementary to the in-situ data. Only measurements within 2.5h around local noon have been used for the comparison. The reason is that at high solar zenith angles the comparisons worsen due to the averaging kernels.

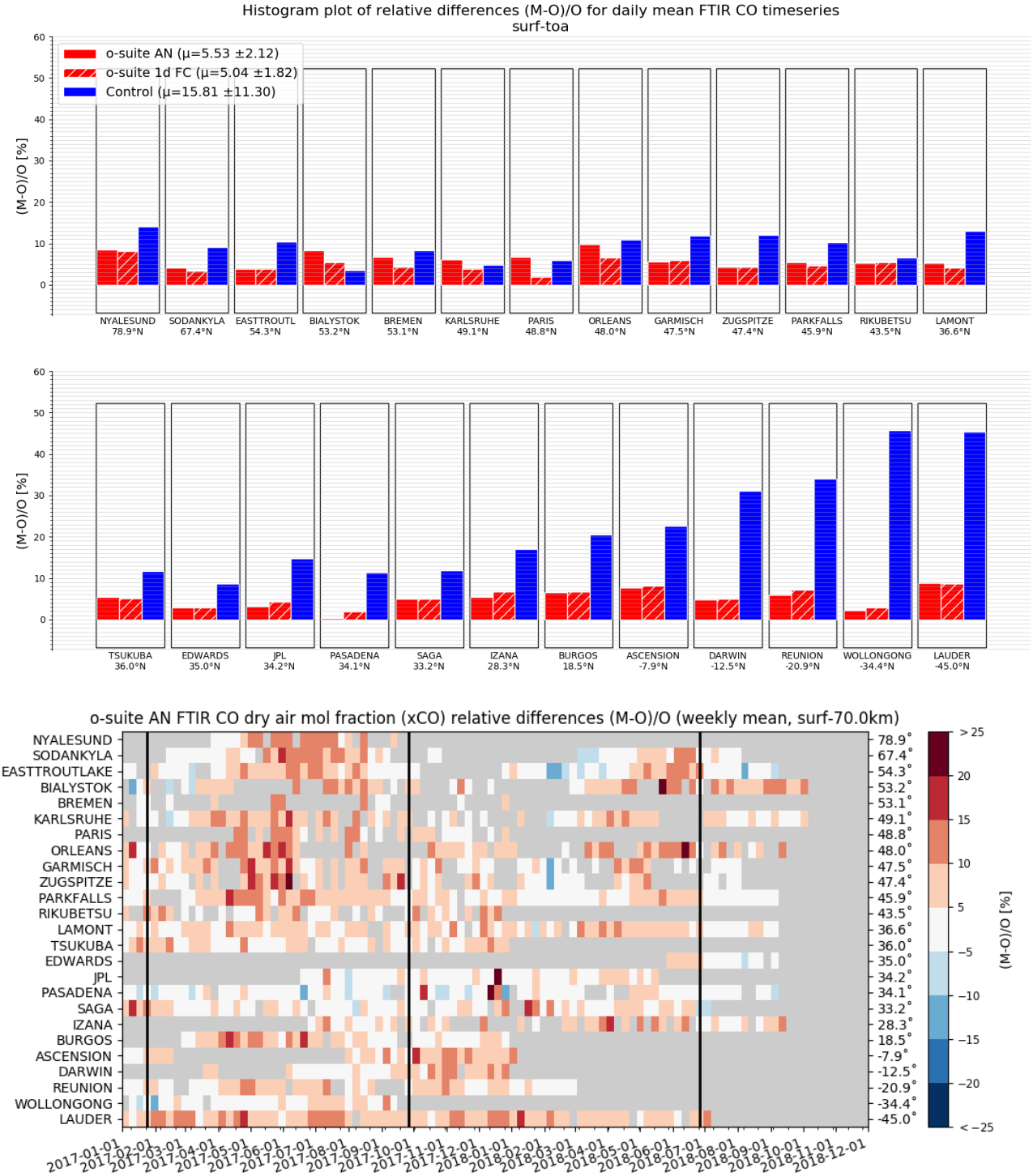


Figure 4.4.1: Seasonal relative mean bias for CO columns (MB, %), standard deviation (STD, %) and number of observations used for the considered period 2017-2018 (top) and weekly-mean biases for the last 2 years (bottom, model upgrades are indicated in black). The stations are sorted with decreasing latitude (northern to southern hemisphere).

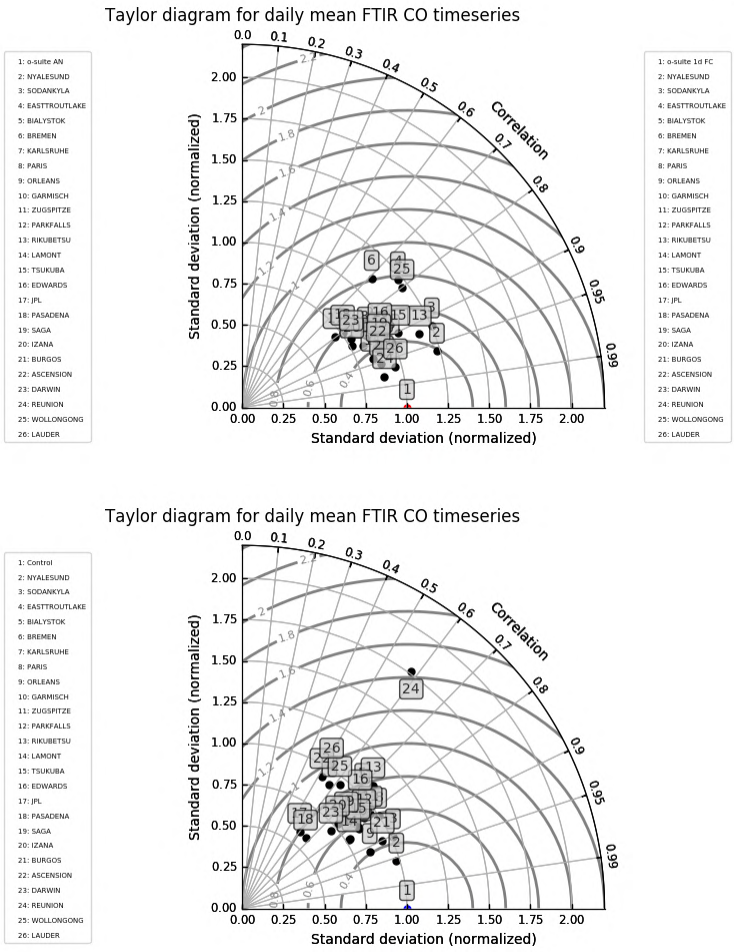


Figure 4.4.2: Taylor diagrams relating the standard deviations for the model /GB time series of column averaged CO mole fractions and their correlation. All timeseries are normalized such that the standard deviation of the model is 1. The Taylor diagrams provide information on the correlation of all three models under consideration with the FTIR time series. Top-left: o-suite analysis; top-right: o-suite 1-day forecast; bottom: control run.

Figure 4.4.1 show that the o-suite data slightly overestimates the CO for all stations. All biases for the o-suite analysis and 1d forecast are between 2% and 8%. This is above the measurement uncertainty, which is about 1%. No latitudinal pattern can be seen in the relative differences.

The control run also shows a positive bias, which is significantly higher than for the o-suite (up to 40%). This bias is increasing in the southern hemisphere towards the South Pole.

The Taylor diagrams in Fig. 4.4.2. show the positive effect of the assimilation: the o-suite has higher correlations compared to the control run.

4.5 Evaluation with MOPITT and IASI data

In this section, modeled CO total columns are compared to MOPITT versions 6 and 7 (thermal infrared radiances) (Emmons et. al., 2009, Deeter et al., 2010) and IASI satellite retrievals (Clerbaux et al., 2009). Figure 4.5.1 shows the global distribution of CO total columns retrieved from MOPITT V7 (top left) and IASI (top right) and the relative biases of the model runs with respect to MOPITT V7, averaged for September 2018.

In September 2018, both, MOPITT and IASI show high CO values over the biomass burning areas in Central Africa and South America, and over North-East China. IASI observations show somewhat higher values than MOPITT. The modeled CO geographical distribution and magnitude of values show that the model performs reasonably well. The relative difference between the model runs and MOPITT shows that both CAMS runs have a positive bias over South America, Central Africa and Indonesia up to 40 % with the control run more biased than the o-suite. The control run also shows negative biases over the Northern mid- and high-latitudes up to 30%. In general, the o-suite performs better than the run without data assimilation, with some regional biases by about +/-10% (except for the above-mentioned biomass burning areas). The control run shows overestimations in the Southern Hemisphere by about 20-30% and in the low Northern latitudes and underestimation over middle and high northern latitudes. Figure 4.5.1 shows no significant difference between the o-suite analysis and 2nd and 4th forecast days, but slightly growing positive biases over the fire active regions, especially over Indonesian islands.

Figure 4.5.2 shows time series of CO total column for MOPITT V6 and V7, IASI and the model runs over the eight selected regions. For the comparison with MOPITT, the modelled CO concentrations were transformed using MOPITT V7 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 different regions. IASI CO values are lower than MOPITT over most regions with some seasonal exceptions till the year 2016. Since then IASI and MOPITT are more consistent with each other over Europe, the US and East Asia. Significant difference between MOPITT and IASI are observed over the Alaskan and Siberian fire regions in winter seasons, with IASI CO total column values lower up to 30 %. In North and South Africa, deviations become larger since 2016 with IASI values being higher than MOPITT by up to 20%. The modelled seasonality of CO total columns is in relatively good agreement with the retrievals. In general, the comparison between the o-suite and control runs shows that assimilation of satellite CO has more positive, pronounced impact on model results over East and South Asia, South Africa, and since the end of 2016 over the US in winter and spring seasons, and smaller impact over the other regions. Since June 2016 the o-suite shows very good agreement with the satellite retrievals over Europe and the US with bias less than 5 %. In late summer and early autumn 2018 over Europe, the control run has larger negative biases compared to the satellite data then early in 2018 and the two previous autumn seasons. In general, in autumn 2018, the CO total columns lower compared to previous autumns over Europe and the US. The Siberian fire region shows a decrease in seasonal minimums compared to previous autumns from both, satellite and modeled data. The seasonal maximum in South Africa also showing decrease of CO values compared to previous autumns.

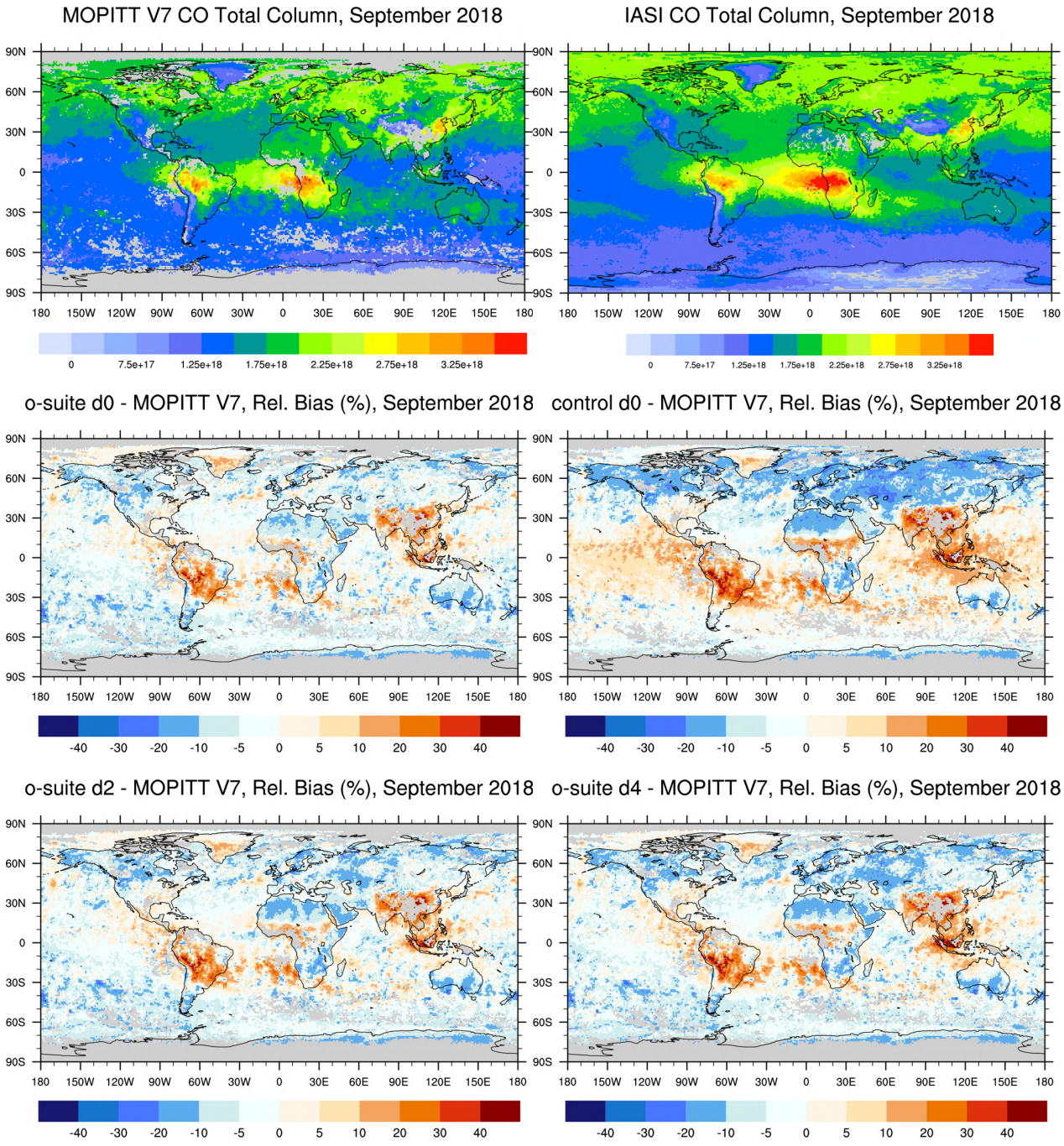


Fig. 4.5.1: CO total columns for MOPITT V7 (top left) and IASI (top right) satellite retrievals and relative difference between the model runs and MOPITT for September 2018: o-suite (middle left), control run (middle right), o-suite 2nd forecast day (bottom left), o-suite 4th forecast day (bottom right). Grey color indicates missing values.

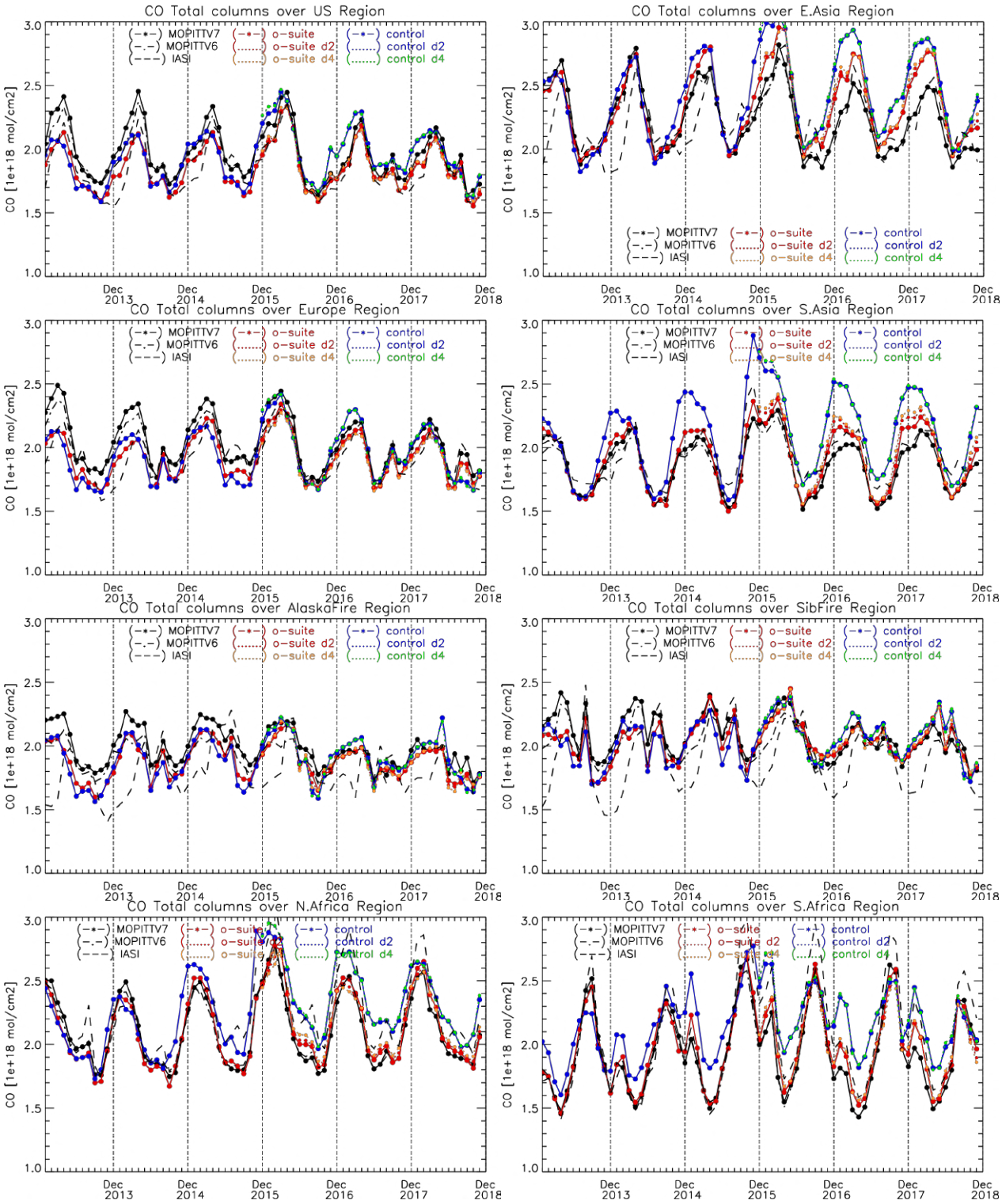


Fig. 4.5.2: Time series of CO total columns for satellite retrievals MOPIT V6 and V7, IASI (black) and the model runs over the selected regions: o-suite (red, solid), control (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue, dotted), control 4th forecast day (green, dotted).

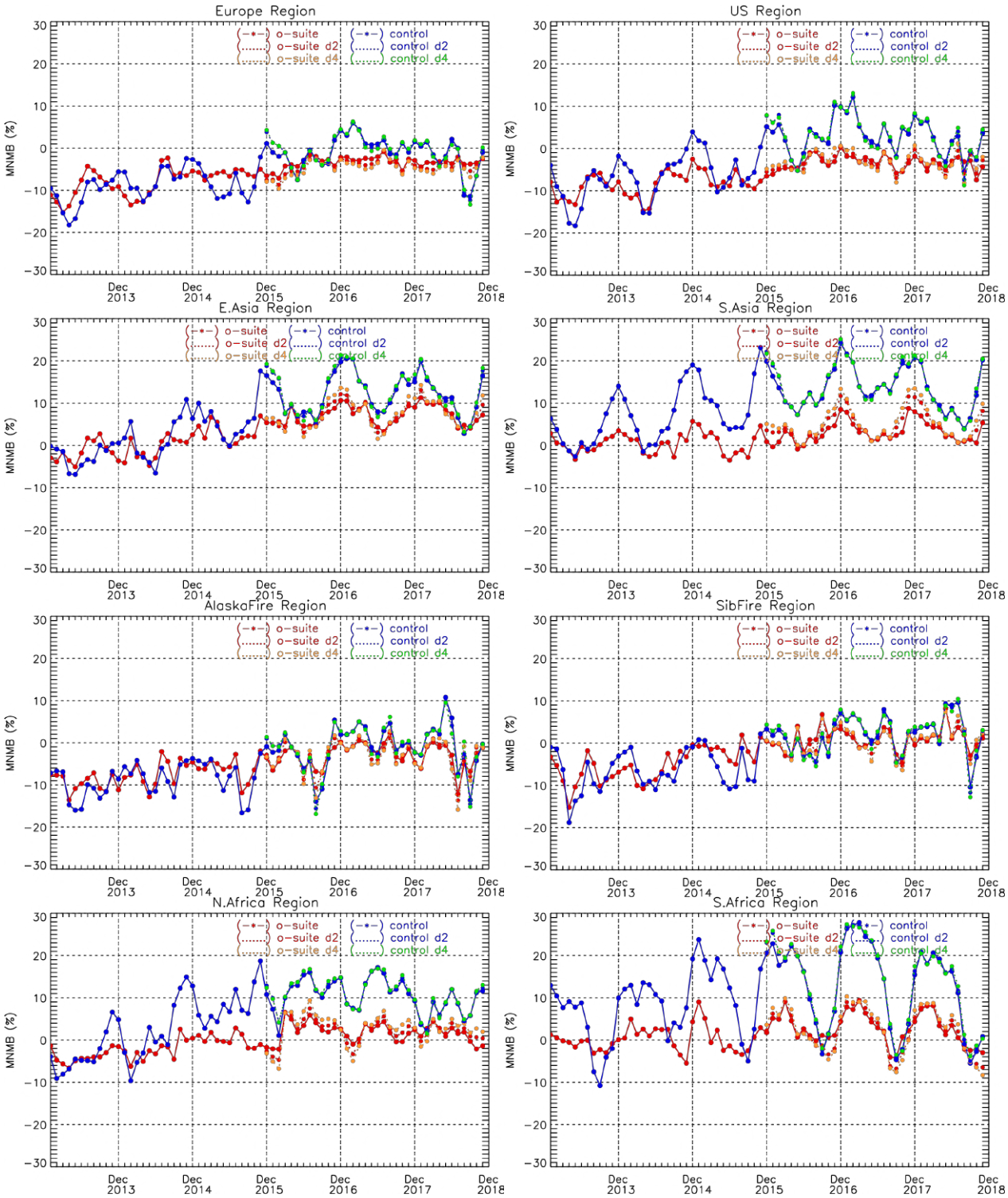


Fig. 4.5.3: Modified normalized mean bias (%) for CO total columns from the model simulations vs MOPITT V7 retrievals over selected regions. O-suite (red, solid), control run (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue, dotted), control 4th forecast day (green, dotted).

The modified normalized mean bias (MNMB) of the model runs compared to MOPITT V7 (Fig. 4.5.3) allows quantifying the impact of the assimilation on the model performance. The o-suite model run shows negative biases over Europe, the US and Alaskan fire regions with some seasonal exceptions. The control run shows a systematic positive bias up to 20% over South Asia in November-December 2014, 2015, 2016, and 2017. Over southern Africa the control run overestimates satellite retrieved values by up to 25% in winter and spring 2015, 2016, and 2017. In general, the o-suite is within +/-10% in all regions, while the control run shows larger biases over East and South Asia and North and South Africa, as well as stronger seasonal cycles.

In August and September 2018 the control run shows an untypical negative bias of about 10% over Europe, while the o-suite follows its typical value by about -5%. In September, the o-suite run shows a slightly growing negative bias for the forecast day 4 over Europe. Over East and South Asia in November the control run biases reached +20%, which is systematic over these areas in the late autumn, early winter. In September the control run shows negative biases of about 10-14% over Alaskan and Siberian fire regions, which is untypical for these regions. In October and November the o-suite shows a growing negative bias for the forecast days 2 and 4 of about 5% over South Africa.

5. Tropospheric nitrogen dioxide

5.1 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 NO2 satellite retrievals (IUP-UB v1.0) [Richter et al., 2011] for more recent simulations. This satellite data provides excellent coverage in space and time and very good statistics. However, only integrated tropospheric columns are available, and the satellite data is always taken at the same local time, roughly 10:00 LT for SCIAMACHY and 09:30 LT for GOME-2, and at clear sky only. Therefore, model data are vertically integrated, interpolated in time and then sampled to match the satellite data. GOME-2 data were gridded to model resolution (i.e. 0.4° deg x 0.4° deg). Model data were treated with the same reference sector subtraction approach as the satellite data. Uncertainties in NO_2 satellite retrievals are large and depend on the region and season. Winter values in mid and high latitudes are usually associated with larger error margins. As a rough estimate, systematic uncertainties in regions with significant pollution are on the order of 20% – 30%.

Figure 5.1.1 shows global maps of GOME-2 and model monthly mean tropospheric NO_2 columns as well as differences between retrievals and simulations for October 2018 as an example of the maps for autumn 2018. 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 much more pronounced in model simulations to the south of India. Emissions over Europe, and especially the pollution hotspots around the Benelux countries are underestimated. However, other local maxima of values observed over anthropogenic emission hotspots in East Asia (e.g. over the heavily populated Sichuan Basin; 30°N, 105°E), India and others such as Teheran, Mecca, around Lebanon/Israel and Moscow are overestimated. Values over western Australia are overestimated by both model runs, fires were active in this region and time of the year (see the CAMS data catalogue to compare to satellite based fire radiative power), meaning that an overestimation of NOx fire emissions could be a reason for this.

Closer inspection of the seasonal variation of tropospheric NO_2 in some selected regions (Fig. 5.1.2) 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 were strongly underestimated before 2014 by all model runs (most likely due to an underestimation of anthropogenic emissions) for all seasons apart from summertime minima, with the o-suite showing the best results since an upgrade in July 2012. As wintertime NO_2 column retrievals decreased since 2014, model simulated wintertime maxima are in better agreement with the satellite retrieved ones for recent years. However, the observed NO_2 decrease is not reproduced by the simulations and therefore the better agreement for more recent years cannot be attributed to model improvements. Moreover, summertime model minima increased in 2015 compared

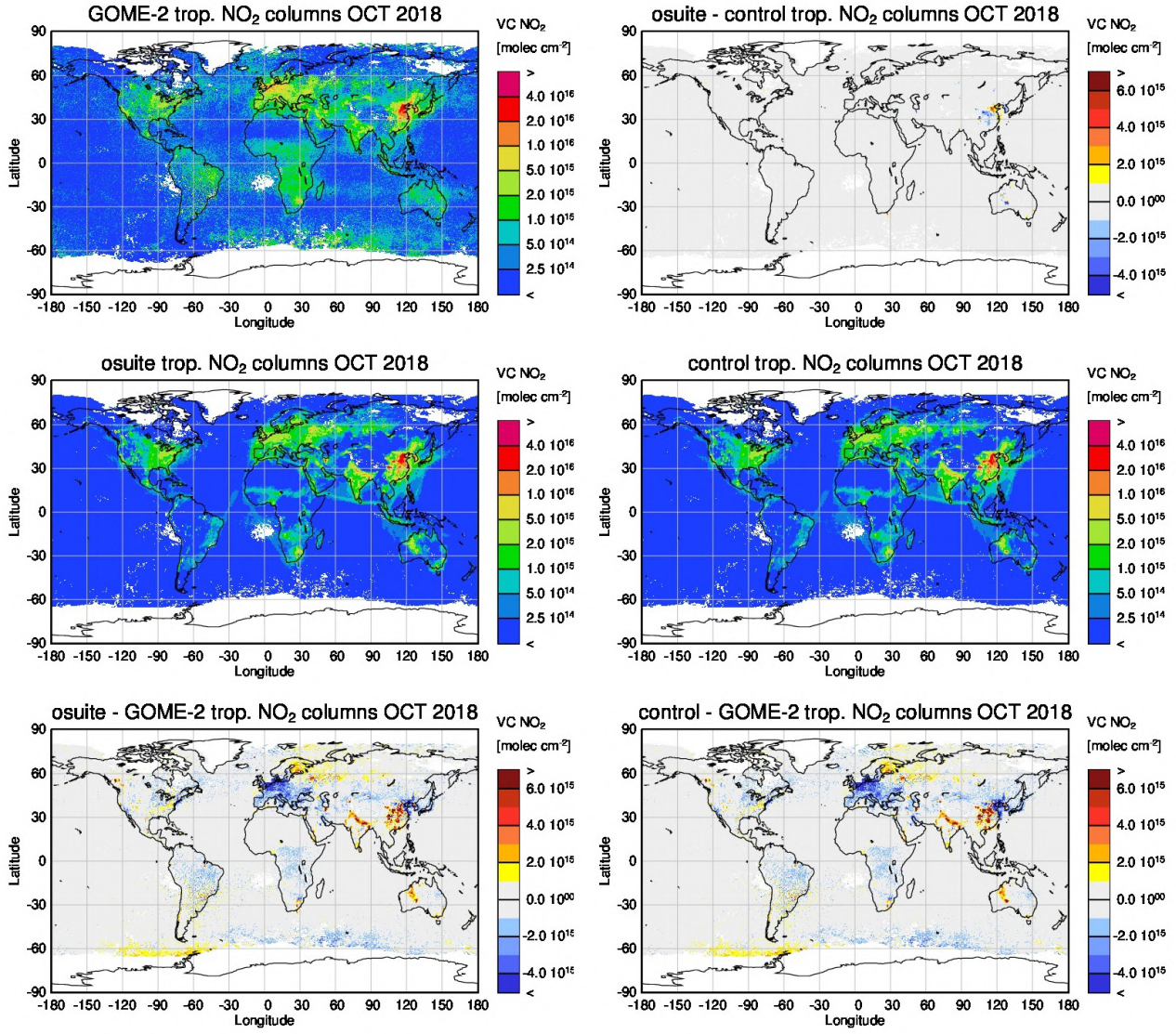


Figure 5.1.1: Global map comparisons of satellite retrieved, and model simulated tropospheric NO2 columns [molec cm-2] for October 2018. The top row shows monthly mean tropospheric NO2 columns retrieved by GOME-2 as well as the difference between o-suite and control, the second row shows the corresponding tropospheric NO2 columns for model simulated averages. The third row shows differences of monthly means between models and GOME-2. GOME-2 data were gridded to model resolution (i.e. 0.4° deg x 0.4° deg). Model data were treated with the same reference sector subtraction approach as the satellite data.

to previous years, which is in contrast to the satellite retrievals, so that the simulated values for the summers since 2015 are about 50% larger than satellite retrieved ones. The observed July and August means of 2018 over East-Asia are a bit lower compared to previous years, but it has to be seen if this trend continues over the next summers. As for East-Asia, a decrease in satellite retrieved values also occurred in 2015 over Europe where a peak is usually found around January, which was, as a result, only slightly underestimated by the models for January 2015. The underestimation of tropospheric NO_2 columns over Europe may be caused to some extent by a change of emission inventories in 2012. However, the situation changed over the last three winter periods, for which

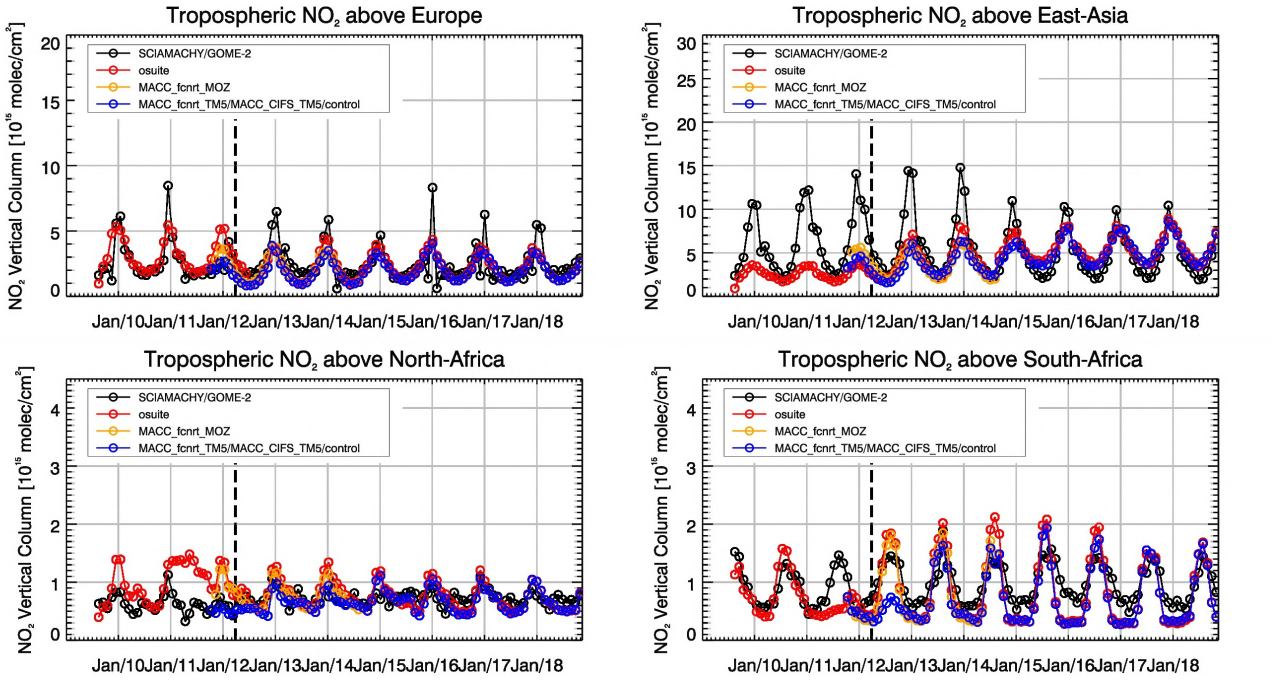


Figure 5.1.2: Time series of average tropospheric NO_2 columns [10^{15} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions (see Annex 2 for definition of regions). Upper panels represent regions dominated by anthropogenic emissions, lower panels represent those dominated by biomass burning. The blue line shows MACC_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 model run without data assimilation is termed control since Sep 2014). Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

GOME-2 shows (compared to previous years) a strong increase in January peak values, combined with a decrease in values for December and February, that 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 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 was 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 an update in July 2012 and the change to IFS-CB05 in September 2014. However, tropospheric NO_2 columns around December are still overestimated by the models. Summertime NO_2 columns over North-Africa are underestimated compared to the satellite data from 2015 onwards. The models (especially the o-suite) generally overestimate the seasonal cycle for South-Africa, particularly for 2014 -2016 with an overestimation of the seasonal maximum which usually occurs around August (e.g. by a factor of 1.4 larger compared to GOME-2 retrievals in 2016). However, summertime values are in better agreement since the upgrade of the o-suite in 2017, but summertime minima remain underestimated.

More NO_2 evaluation plots can be found on the CAMS website, see table 1.2.

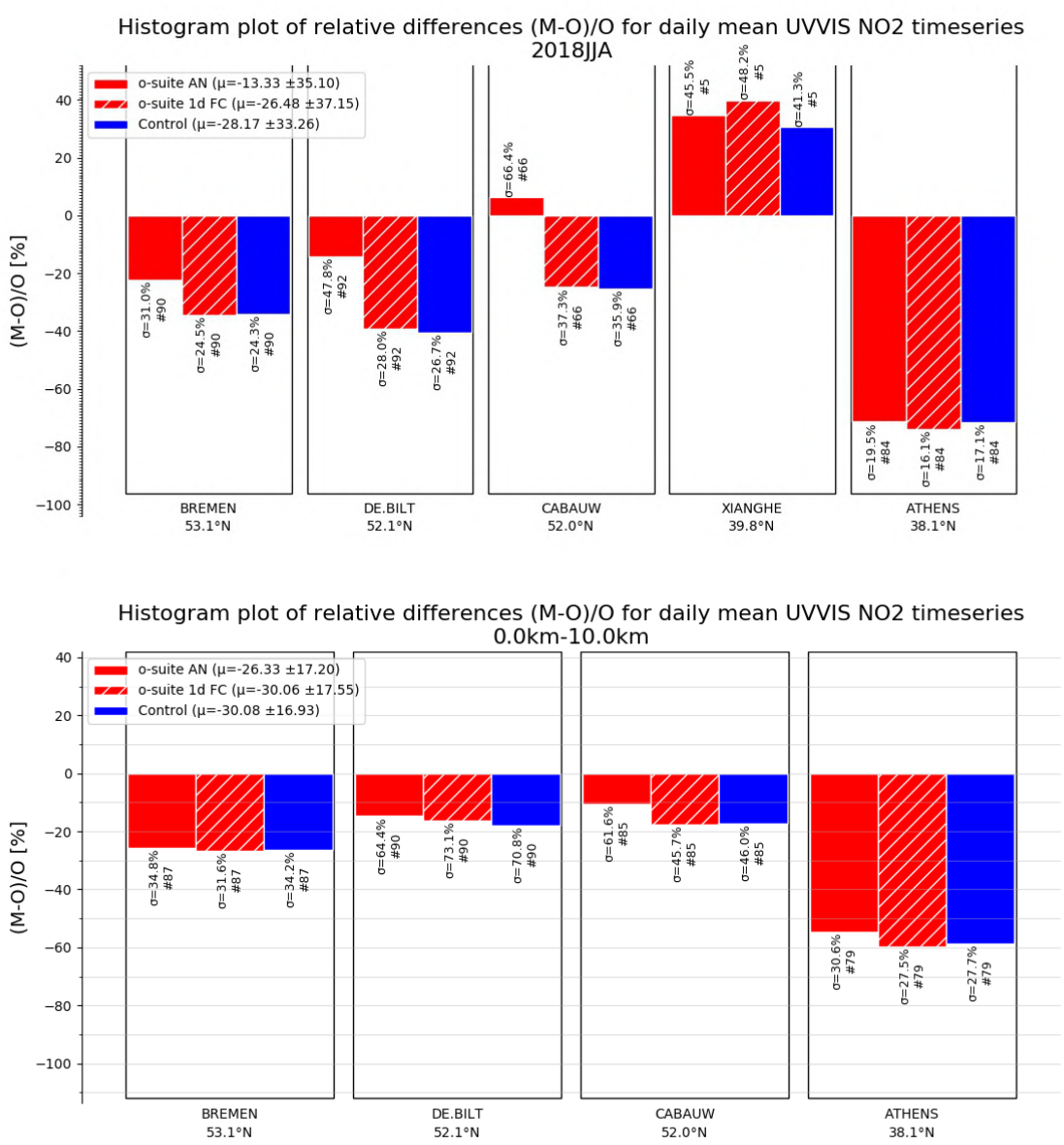


Figure 5.2.1: Table diagram showing the seasonal bias June –August 2018 (top) and Sept-Nov 2018 (bottom) for five stations, sorted by latitude. Compared to the previous validation period JJA, the relative biases in SON have not changed significantly. Due to instrument failure the instrument at Xianghe did not generate any measurements.

5.2 Evaluation against ground-based DOAS observations

In this section, we compare the NO_2 columns of the CAMS models with UVVIS DOAS profile measurements at Xianghe (39.8°N, 117°E, station near Beijing, altitude 92m) and column data from the other stations.[1] This ground-based, remote-sensing instrument is sensitive to the NO_2 abundance in the lower troposphere, up to 1km altitude with an estimated uncertainty of 8%.

[1] No contribution from Reunion and OHP due to instrument failure.

Tropospheric NO_2 profiles and columns are validated (up to 3.5km or 10km). A description of the instruments and applied methodologies is the same all DOAS OFFAXIS measurements, see

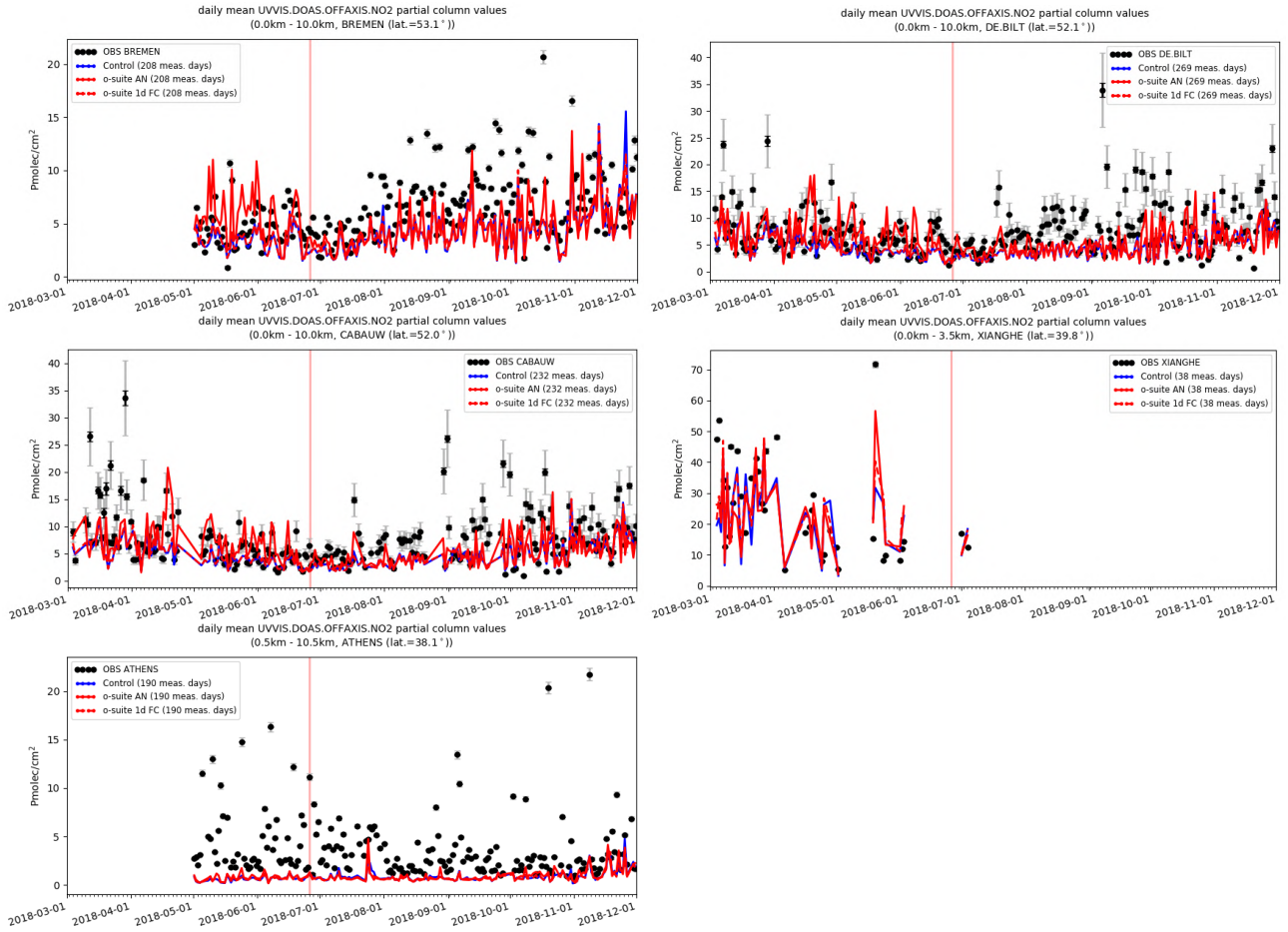


Figure 5.2.2: Time series of NO_2 partial columns at the five different sites. For all sites except Athens, background concentrations are well captured by the model and the o-suite AN seems to miss the high pollution days since the model update in June 2018.

http://nors.aeronomie.be. It is important to mention here that the model partial column values are calculated from the smoothed model profiles. This guarantees that the model levels where the measurement is not sensitive do not contribute to the observed bias. We should mention that the measurement data is still catalogued as rapid delivery and not in the consolidated NDACC database.

Fig. 5.2.1 shows the biases for the latest validation periods June-August 2018 and Sept-Nov 2018 at the different sites. Both Bremen and Athens are outliers (both stations cover the time period only partially Fig 5.2.2). At Athens a strong underestimation is observed. The o-suite AN was able to capture some of the high pollution events in the previous validation period, but for the SON period, the high pollution events at Bremen, De Bilt and Cabauw are not well captured.

6. Formaldehyde

6.1 Validation against satellite data

In this section, simulations of tropospheric formaldehyde are compared to SCIAMACHY/Envisat HCHO satellite retrievals (IUP-UB v1.0) [Wittrock et al., 2006] for model data before April 2012 and to GOME-2/MetOp-A HCHO data (IUP-UB v1.0) [Vrekoussis et al., 2010] afterwards. As the retrieval is performed in the UV part of the spectrum where less light is available and the HCHO absorption signal is smaller than that of NO_2, the uncertainty of monthly mean HCHO columns is relatively large (20% – 40%) and both noise and systematic offsets have an influence on the results. However, absolute values and seasonality are retrieved more accurately over HCHO hotspots.

In Figure 6.1.1, monthly mean satellite HCHO columns are compared to model results for October 2018. The magnitude of oceanic and continental background values and the overall spatial distribution are well represented by the o-suite and control. The models overestimate values over Northern Australia and Central Africa which could be due to fire or biogenic emissions.

Time series in Fig. 6.1.2 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 underestimate the yearly cycle over East-Asia since 2012. In contrast to MOZART runs, MACC_CIFS_TM5 overestimated satellite values for the Eastern US since the middle of 2013. However, the newer IFS-CB05 runs perform well for Eastern US since 2015. For recent years and both regions, there is virtually no difference between the most recent o-suite run with IFS-CB05 chemistry and the corresponding control run 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 winter in the Northern Hemisphere, leading to e.g. the negative values in the GOME-2 time series for Eastern US since December 2015. Summertime maxima are still underestimated by the now higher resolution runs over East-Asia since 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 slightly overestimated by IFS-CB05 chemistry model runs since 2014 and also the latest higher resolution model versions since July 2016.
- Indonesia, where HCHO is also dominated by biogenic sources and biomass burning. Old MOZART based model versions generally overestimated satellite values here (by a factor of 3 – 4 in the second half of 2010) and failed 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 agreed much better since late 2012. IFS-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

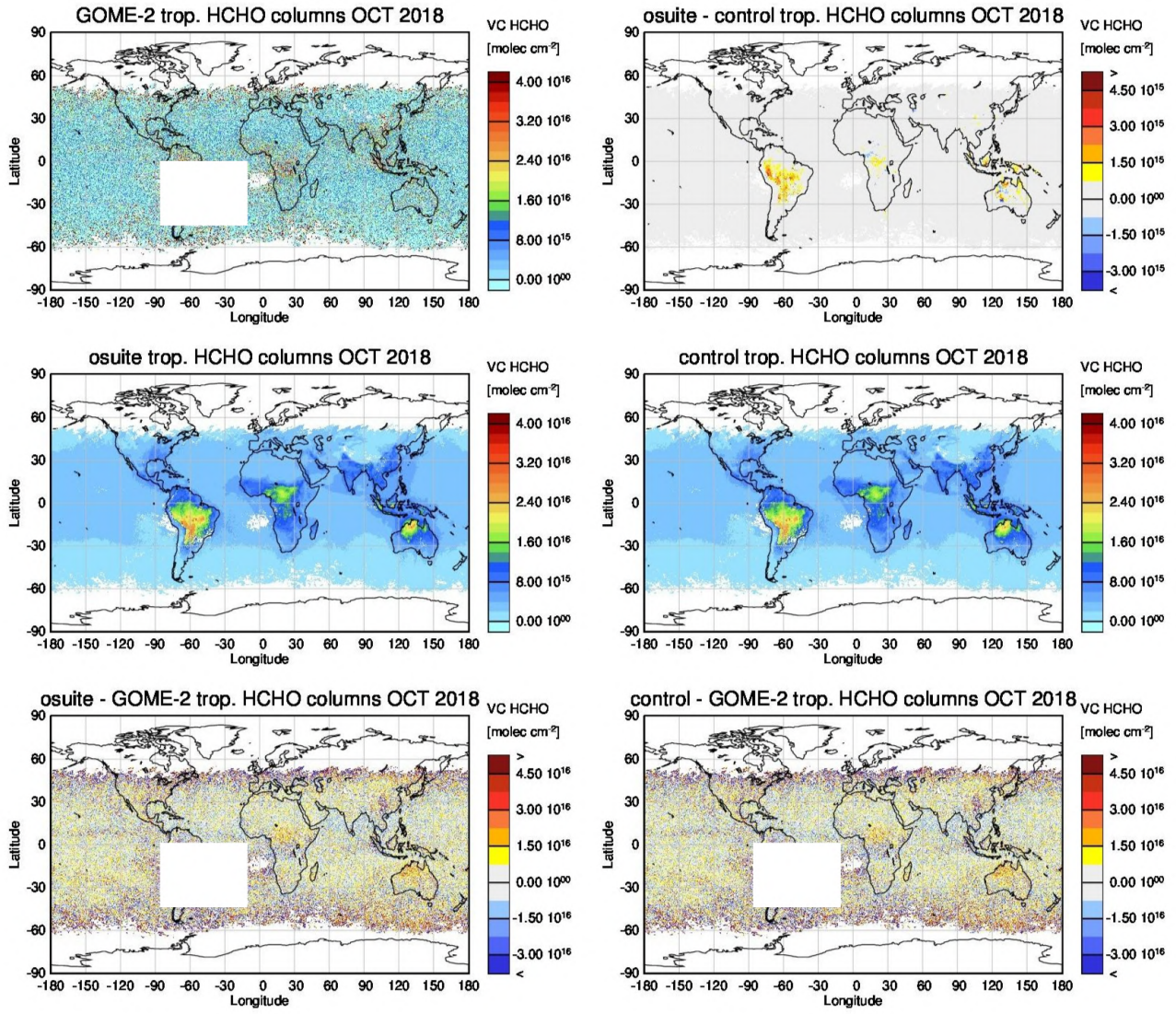


Figure 6.1.1: Global map comparisons of satellite retrieved and model simulated tropospheric HCHO columns [molec cm-2] for October 2018. The top row shows monthly mean tropospheric HCHO columns retrieved by GOME-2, the second row shows the same but for model simulated averages. The third row shows differences of monthly means between models and GOME-2. GOME-2 data were gridded to model resolution (i.e. 0.4° deg x 0.4° deg). Model data were treated with the same reference sector subtraction approach as the satellite data. Satellite retrieved values in the region of the South Atlantic anomaly are not valid and therefore masked out (white boxes in all images except those which show model results only).

El Nino years, fire emissions used by IFS-CB05 seem to be largely overestimated, resulting in model simulated HCHO columns which are almost twice as large as those retrieved by GOME-2. Further investigations (see previous reports) show that this is not caused by cloud flagging applied to the satellite and model data. The recent higher resolution runs in general overestimate values over Indonesia as well. There is mainly little variation from one month to another in both, satellite observations and model simulations since middle of 2016, the decrease in retrieved HCHO columns for Dec 17/ Jan 18 and an increase in May 2018 is not reproduced by the simulations.

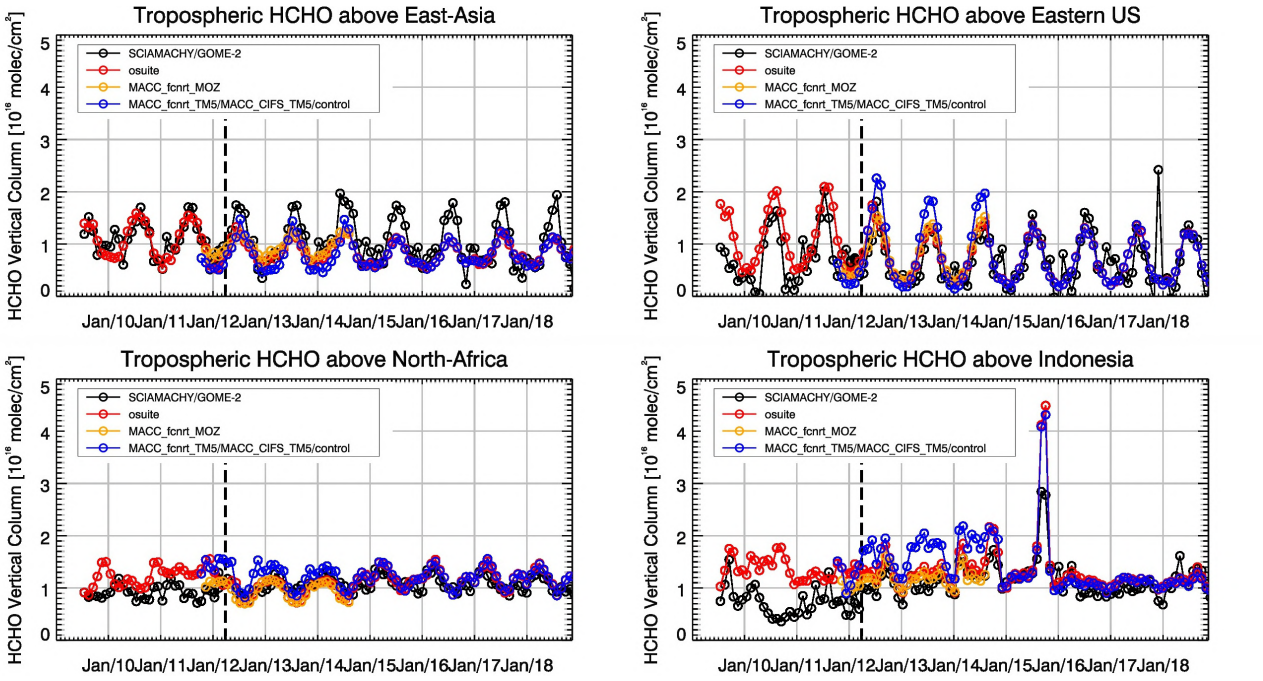


Figure 6.1.2: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions. The blue line shows MACC_fcnrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 to August 2014 and control results from September 2014 onwards (the model run without data assimilation is termed control since Sep 2014). The regions differ from those used for NO_2 to better focus on HCHO hotspots: East-Asia (25-40°N, 110-125°E), Eastern US (30-40°N, 75-90°W), Northern Africa (0-15°N, 15°W-25°E) and Indonesia (5°S-5°N, 100-120°E). Negative satellite retrieved values over Eastern US are due to a lack of data (caused by instrument degradation) during Northern Hemisphere winter months for this region. Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

For details on the HCHO evaluation: http://www.doas-bremen.de/macc/macc_veri_iup_home.html

6.2 Evaluation against ground-based DOAS observations

In this section, we compare the H2CO columns of the CAMS models with UVVIS DOAS measurements at Xianghe, Cabauw and De Bilt.[2] These ground-based, remote-sensing instruments are sensitive to the HCHO abundance in the lower troposphere. Tropospheric HCHO profiles and columns are validated (up to 3.5km (Xianghe) or 10km (Cabauw and De Bilt)). A description of the instruments and applied methodologies is the same as for the MWR O_3 and FTIR O_3 and CO validations see http://nors.aeronomie.be. It is important to mention here that the model partial column values are calculated for the smoothed model profiles. This guarantees that the model levels where the measurement is not sensitive do not contribute to the observed bias. We should mention that the measurement data is catalogued as rapid delivery and not in the consolidated NDACC database.

[2] No contribution from Reunion and OHP due to instrument failure.

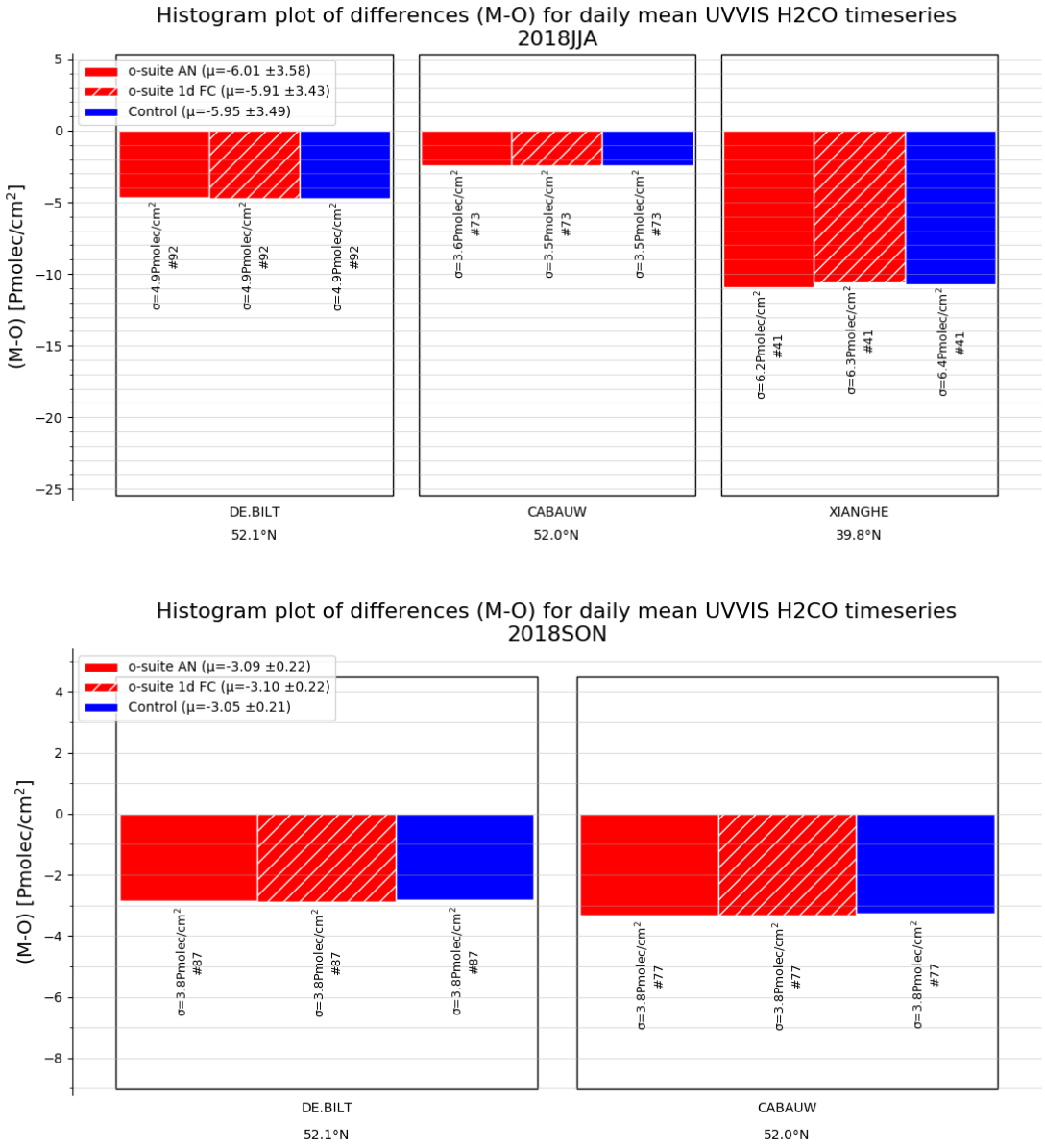


Figure 6.2.1: Table diagram showing the seasonal absolute bias (June-August 2018 and Sept-Nov 2018) for three stations, sorted by latitude.

Fig. 6.2.1 shows the absolute biases for June-August and Sept-Nov 2018 at the different sites, which shows an underestimation of all models at all sites. The instrument at Xianghe is down. At Cabauw and De Bilt the underestimation has decreased during the SON period and during November 2018 the bias has changed sign: the HCHO abundance decreased due to less production and this seasonal variation is not fully captured by the model. From Fig. 6.2.1 and 6.2.2 we see little difference between the o-suite and the control run. Although the background column values are well captured by the models, the high emission events are not.

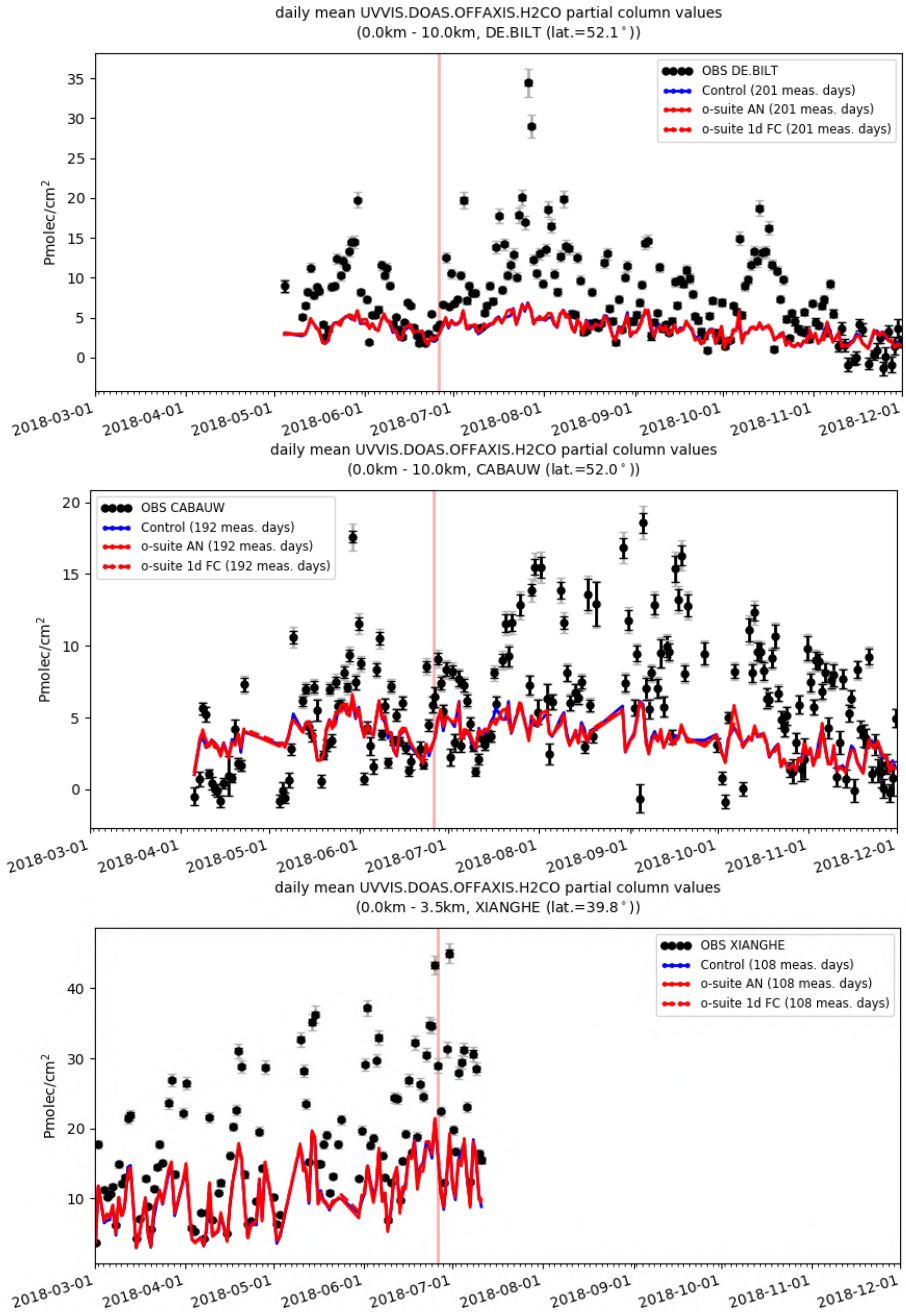


Figure 6.2.2: Time series of H2CO partial columns at the five different sites. All models underestimate the H2CO concentrations. During November, the bias changes at Cabauw and De Bilt and the model overestimates the abundance of H2CO.

7. Aerosol

7.1 Global comparisons with Aeronet and EMEP

The comparison of the CAMS simulation of time series of aerosol optical depth can be found for all Aeronet stations at: http://aerocom.met.no/cams-aerocom-evaluation/

More detailed evaluation including scores, maps, scatterplots, bias maps and histograms illustrating the performance of the aerosol simulation in the IFS system are made available through the AeroCom web interface. The model run can be compared here to eg. the CAMS interim reanalysis and other models, such as the AeroCom Median model.

Correlation, based on daily aerosol optical depth and NRT Aeronet observations, has been rather stable recently. The o-suite forecast at +3 days shows only slightly lower correlation. See figure S3.

Part of the month-to-month variation in correlation is due to the varying quality and coverage of the Aeronet network. This has been improved by the version 3 from Aeronet. We use therefore version 3 level 1.5 for all global comparison to Aeronet.

The performance of the o-suite model exhibits some seasonal variation in AOD depending on region (Fig. 7.1.1). Noteworthy is the persistent AOD overestimation over North America (Fig. 7.1.1-bottom), but also a long-term trend to overestimation in East Asia. The latitudinal display of model and Aeronet AOD in the period investigated here (Fig. 7.1.2) shows a specific positive bias against Aeronet in the Southern Hemisphere.

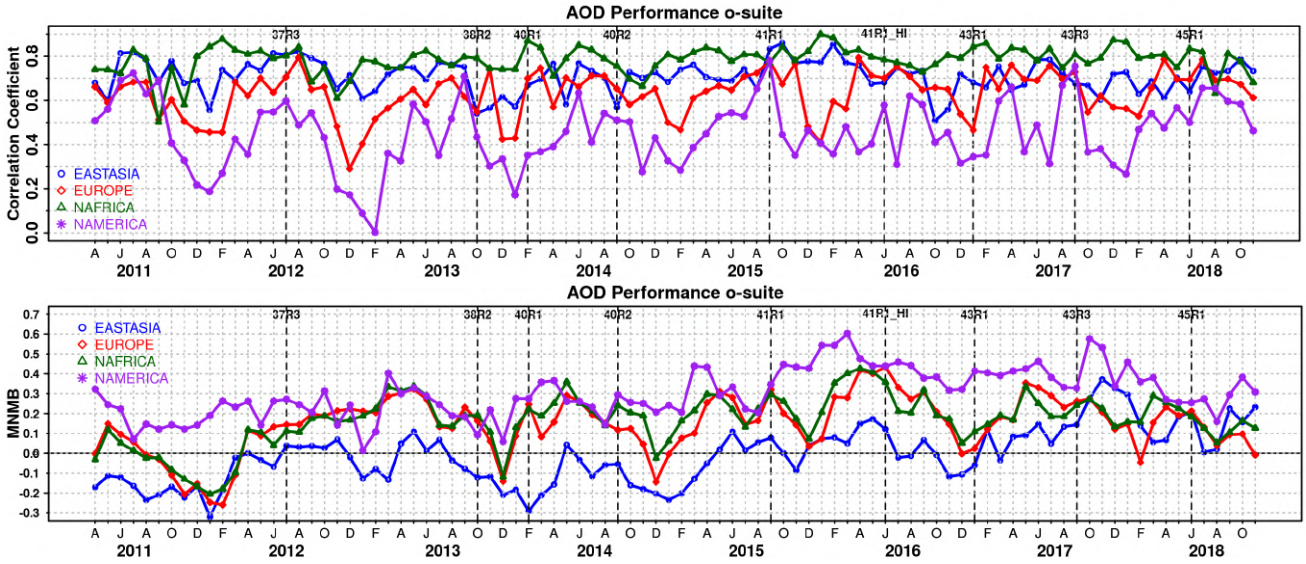


Figure 7.1.1. (top) Correlation coefficient and (bottom) modified normalized mean bias (MNMB) in AOD, since 2011, based on daily AOD comparison (Aeronet V3 level 1.5 data) in four world regions [East-asia (blue); Europe (red); North Africa (green); North America (purple)] for the o-suite.

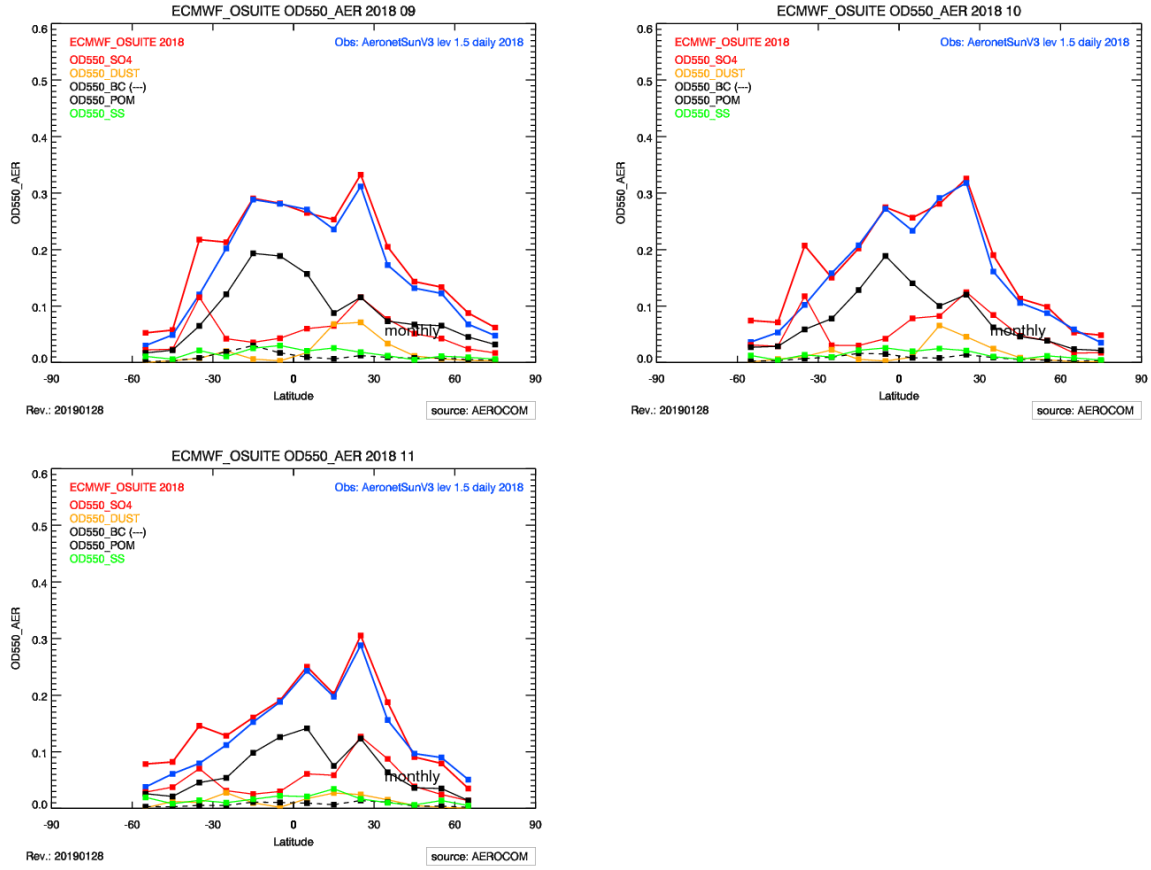


Figure 7.1.2. Aerosol optical depth of o-suite (red) compared to latitudinally aggregated Aeronet V3 level 1.5 data (blue) for the three months covered by this report.

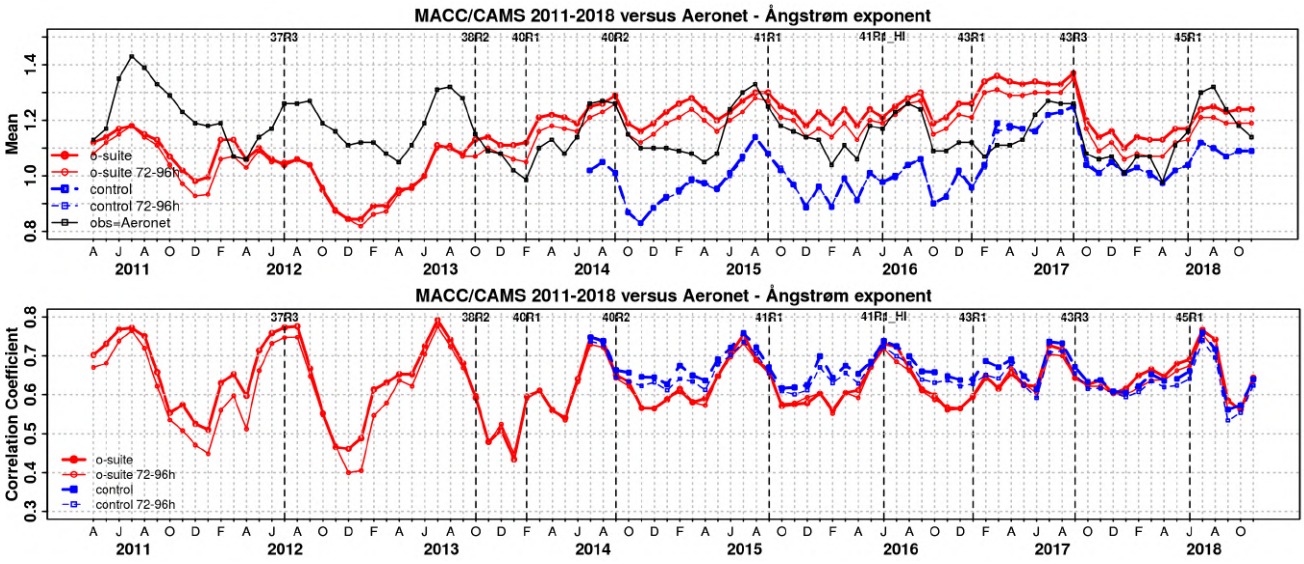


Figure 7.1.3. a) (top) Evolution of mean Ångström exponent in o-suite and control at Aeronet sites (Aeronet V3 level 1.5 data), based on matching monthly mean values. o-suite (thick red curve); o-suite at last forecast day (light red curve); control (blue dashed curve); control at last forecast day (light blue dashed curve). b) (bottom) Correlation using daily matching Ångström exponent.

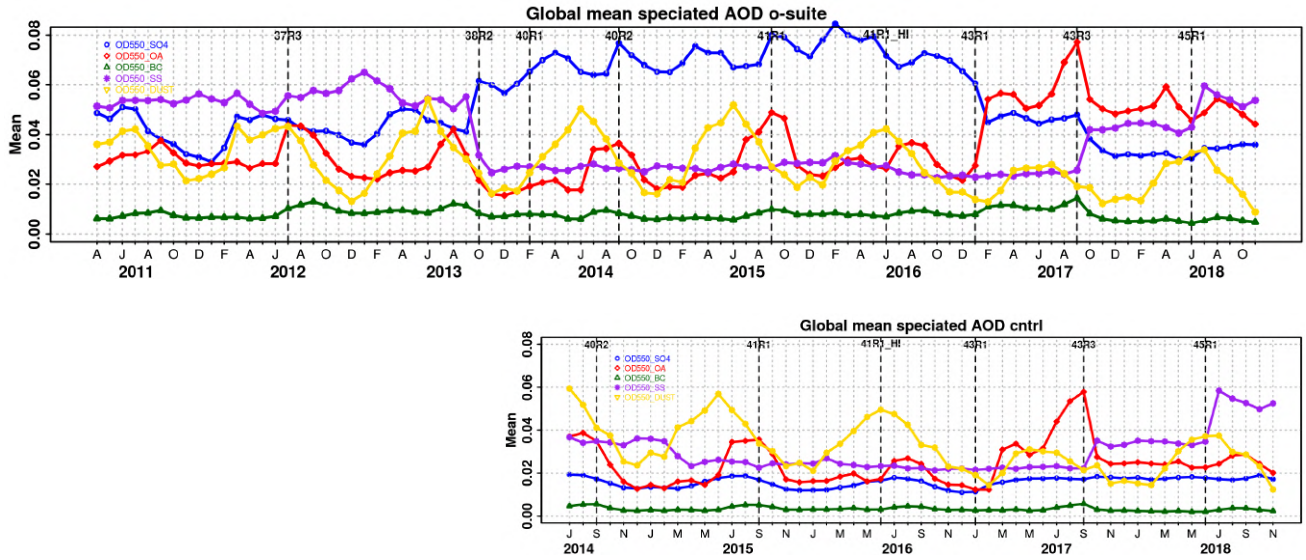


Figure 7.1.4. Evolution of the aerosol components of total AOD@550nm [OD550_SO4 = sulphate(blue); OD550_OA = organics(red); OD550_BC = black carbon(green); OD550_SS = sea salt(purple); OD550_DUST = dust(yellow)] in o-suite and control simulation.

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. We find in SON 2018 a small bias (Figure 7.1.3-a). Temporal and spatial variability is difficult to capture, but correlation from all daily data is lower than for AOD (Figure 7.1.3-b and S3). Figure 7.1.4 shows that the Sep 2017 and Jun 2018 model changes are responsible for a shift in Ångström exponent. More organic matter seems to shift the size distribution to smaller sizes. The model upgrade in Feb 2017 with a bugfix for sea salt and improved parameterisations for SO4 lead to sea salt increased with 45% while sulphate further decreased a bit. Sea salt has increased further due to a new sea salt emission scheme implemented in the latest model upgrade and is back to earlier 2011-2013 levels.

The o-suite uses data assimilation to obtain a first guess aerosol field. In the forecast period, however, a-priori model parameterisations and emissions (except fire emissions, which are kept in the forecast equal to the latest GFAS emission values) determine increasingly the aerosol fields. The performance of the day three forecasted AOD fields as compared to the first guess is shown in Figure S3 in the summary of this report. Table 7.1.1 shows an average global decrease in total aerosol optical depth during the first four forecast days, dominated by sulphate and organics. The control run with no assimilation shows significant less AOD (-25% compared to o-suite, see figure S3). All this supports the conclusion that either a-priori IFS aerosol and aerosol precursor sources are too small or sinks are too effective in the IFS model.

Surface concentration of particulate matter below 10 µm (PM10) and below 2.5 µm (PM25) from the o-suite experiment have been validated against data from 160 background IMPROVE and EMEP stations. A climatological average has been constructed from data in the period 2000-2009 as available in the EBAS database hold at NILU. The data availability is not the same at all stations, and sometimes covers only a few years.

Table 7.1.1. Mean global total and speciated AOD in the o-suite for the last two periods covered by the VAL report and change after 3 forecast days.

o-suite				
	Mean JJA 2018 0-24h	Change wrt to first day on day 4	Mean SON 2018 0-24h	Change wrt to first day on day 4
AOD@550	0.172	-14%	0.158	-14%
BC-OD@550	0.005	-21%	0.005	-19%
Dust-OD@550	0.031	3%	0.015	13%
OA-OD@550	0.050	-24%	0.048	-24%
SO4-OD@550	0.033	-22%	0.035	-22%
SS-OD@550	0.053	-7%	0.053	-6%

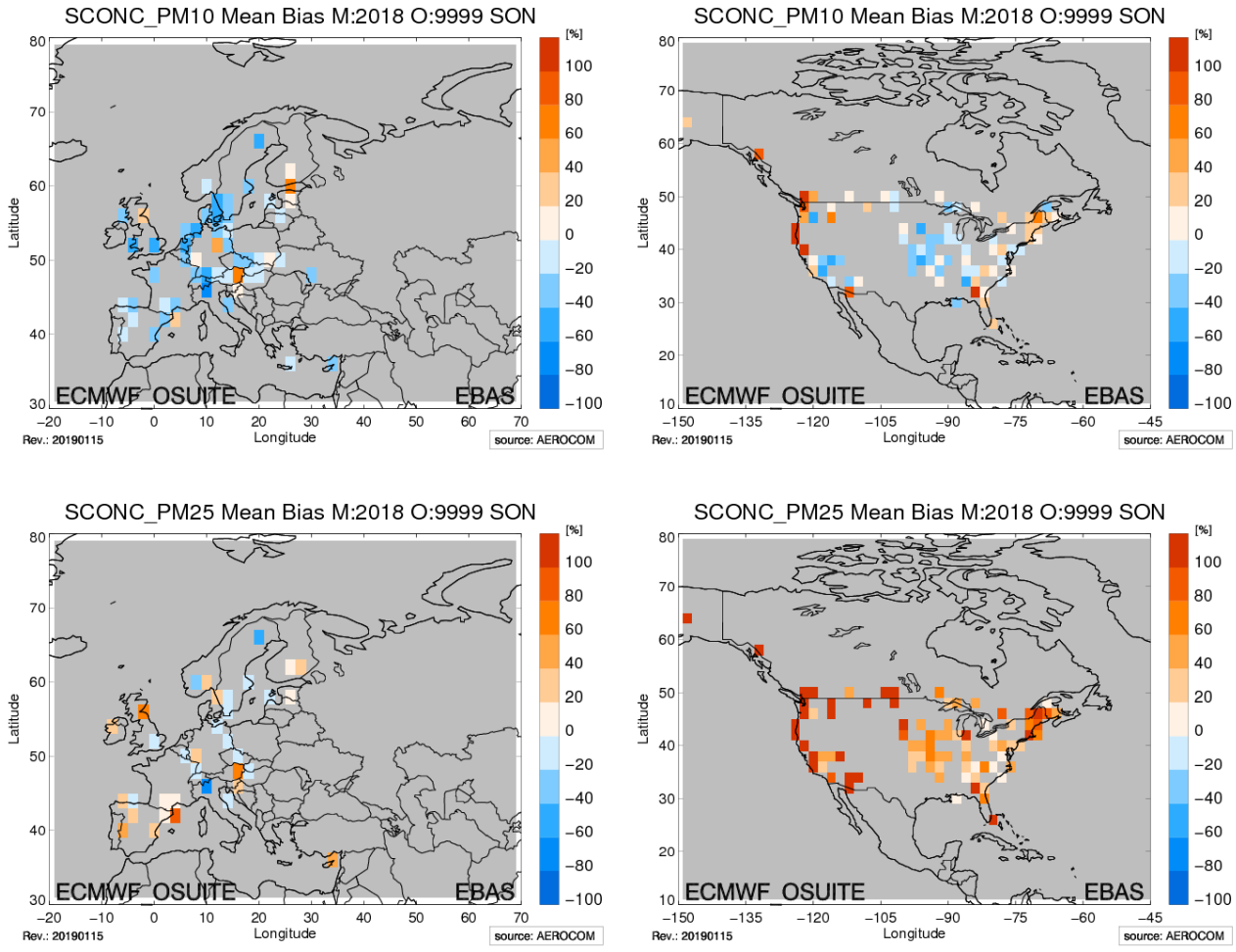


Figure 7.1.5. Bias [%] map of monthly mean PM10 and PM2.5 concentrations at EMEP (Europe) and IMPROVE sites (North America); simulated o-suite versus EMEP/IMPROVE derived climatological average (2000-2009).

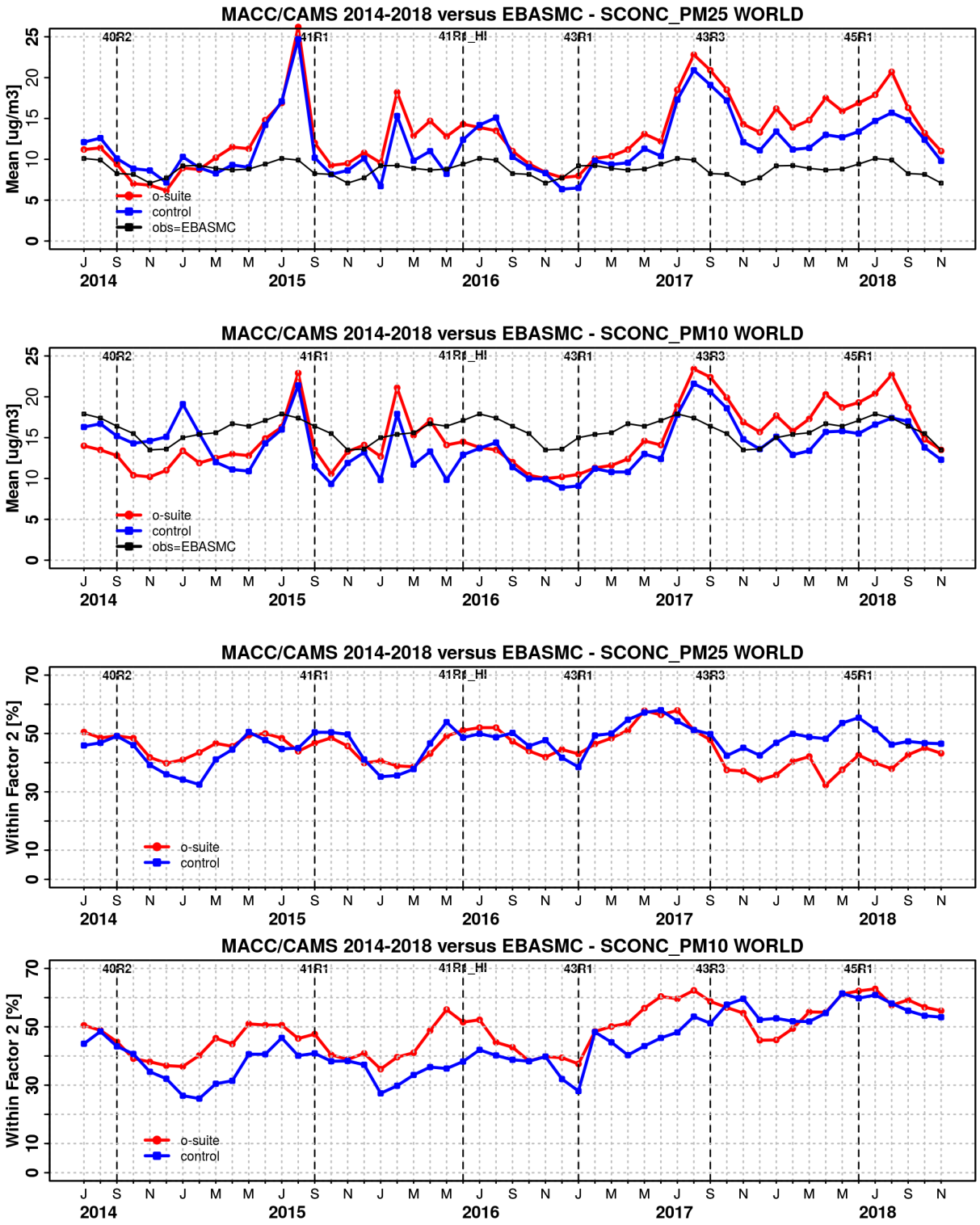


Figure 7.1.6. Temporal evolution of monthly mean average PM25 and PM 10 concentrations at EMEP (Europe) and IMPROVE sites (North America) and data fraction within a factor 2 of observed; ca 160 sites, observed data averaged from data available in EBAS from 2000-2009.

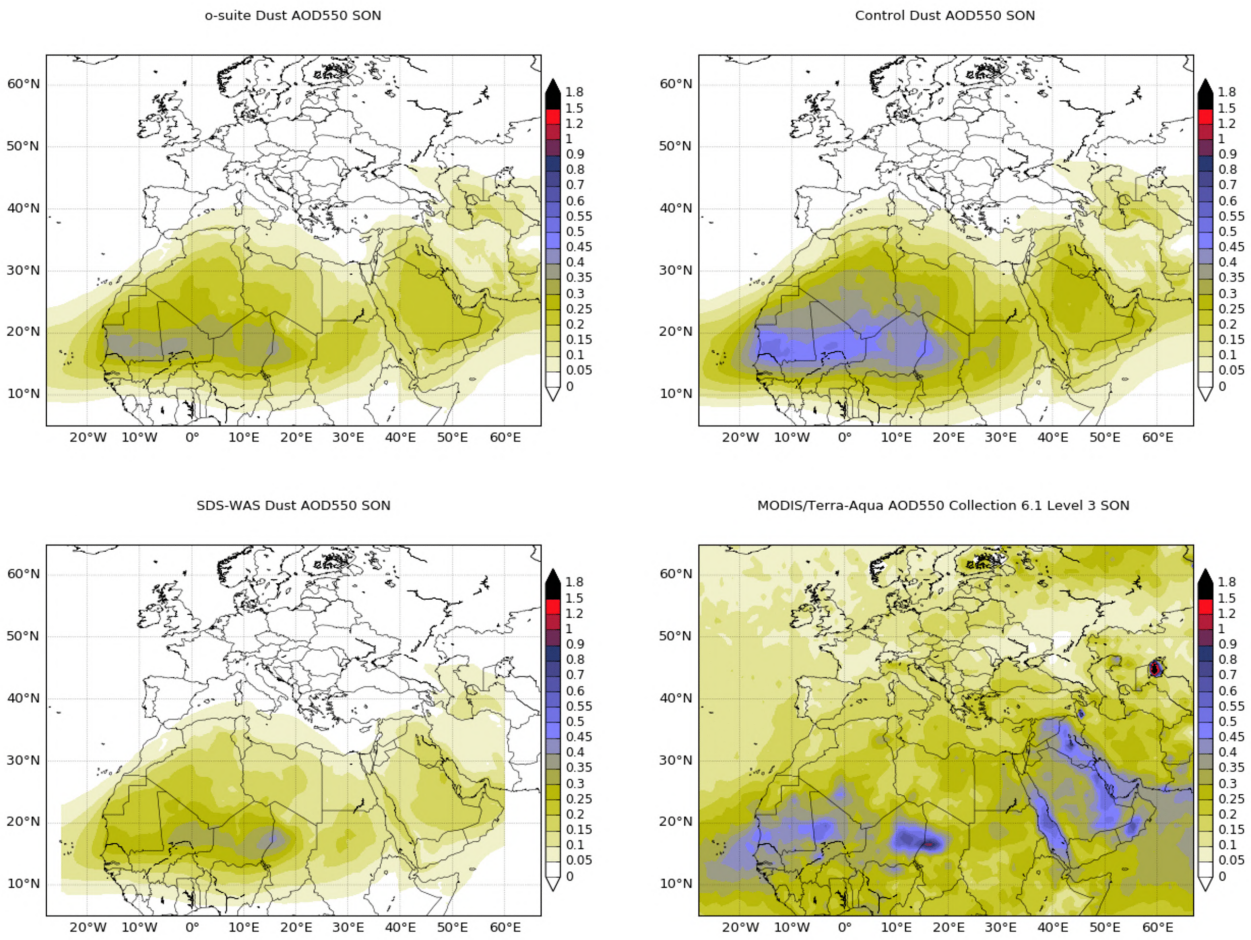


Figure 7.2.1: Averaged DOD 24h forecast from o-suite (top left) and control (top right), DOD of the multi-model SDS-WAS Median product (bottom left) as well as AOD from MODIS/Terra-Aqua Collection 6.1 Level 3 combined Dark target and Deep Blue product (bottom right) for the study period.

A negative MNMB bias of PM10 in Europe and an overestimate in North America PM2.5 appears (Fig. 7.1.5), consistent with the AOD bias in the two regions. Figure 7.1.6 shows the evolution of mean observed and simulated PM10 and PM2.5. The biggest change appeared in July 2017 with the bias of o-suite now becoming positive overall. Shown is also the statistics of being within factor 2, a more robust metrics for a comparison to climatological data. This statistical indicator has clearly improved over time, indicating best PM10 and PM25 performance in summer months for the o-suite. O-suite is also better most of the times than the control simulation for PM10. For PM25 the difference is less clear, but since September 2017 (upgrade to 43R3) the control is performing better than the o-suite.

7.2 Dust forecast model inter-comparison: Validation of DOD against AERONET, and comparisons with Multi-model Median from SDS-WAS

The 72-hour forecasts (on a 3-hourly basis) dust aerosol optical depth (DOD) from CAMS o-suite and control have been validated for the period 1 September 2018 – 30 November 2018 against AERONET direct-sun cloud-screened observations, MODIS/Terra and Aqua Collection 6.1 Level 3 (1º

x 1º) and SDS-WAS Multi-model Median DOD. The SDS-WAS Multi-model Median DOD is obtained from (currently) twelve dust prediction models participating in the Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS) Regional Center for Northern Africa, Middle East and Europe (http://sds-was.aemet.es/).

For autumn, dust activity is low in comparison with the previous season. Satellites (see MODIS in Figure 7.2.1) show that major dust activity in Northern Africa is concentrated over the Sahara in latitude around 15ºN (in the Bodélé Basin and the Mali/Mauritania border) and in Iraq and south-eastern Arabia Peninsula. In North Africa, both CAMS experiments can simulate the main areas of dust activity in the North Africa in comparison with MODIS (see Figure 7.2.1) although o-suite presents lower season values (seasonal DOD up to 0.4) than control (seasonal DOD up to 0.6) which are closer to the SDS-WAS multi-median product. Otherwise, dust transport over the North Atlantic region and the Red Sea appears underestimated in comparison with the satellites (particularly in o-suite). Neither both CAMS experiments, nor the SDS-WAS multi-model product are capturing the maximum AOD values observed in the Middle East, indicating missing dust emissions in all the experiments.

For September to November, the o-suite reproduces the daily variability of AERONET observations with a correlation coefficient of 0.83, averaged over all the AERONET sites (as in the case of the SDS-WAS multi-model product), this is a close to the control experiment that shows a correlation coefficient of 0.81. Regarding mean bias (MB), both CAMS experiments (o-suite and control), as well as the SDS-WAS Multi-model, underestimate the AERONET observations resulting in an MB of 0 for control and -0.03 for o-suite and-0.04 the SDS-WAS multi-model. The central Mediterranean (see Tunis Carthage in Figure 7.2.2 and Table 7.2.1) region has the best results in the AERONET comparison in terms of correlation. Both experiments can reproduce the daily variability with correlation coefficient of 0.84 both for o-suite and control.

Over the Sahara both CAMS experiments show closer results with a correlation coefficient of 0.67 for control and 0.69 for o-suite although o-suite shows slightly higher overestimation (MB of 0.02) than the control experiment which tends to overestimate the AERONET observations (MB of 0.13). o-suite presents better results than the SDS-WAS Median Multi-model that shows a correlation coefficient of 0.62 for the Sahara, although it underestimates the observations with an MB of -0.03. In the Sahel (see Banizoumbou in Figure 7.2.2 and Table 7.2.1), the o-suite shows strong underestimations (MB of -0.18, slightly higher than control with MB of -0.03) although the o-suite better reproduces the observed daily variability (with a correlation value of 0.67 for o-suite in comparison of for control that has a correlation of 0.59). The underestimations observed in o-suite in the Sahel are also spread to the Tropical North Atlantic (MB of -0.06 for o-suite) in agreement with the underestimation also observed in the comparison with MODIS (Figure 7.2.1). In the Tropical North Atlantic region (see Dakar in Figure 7.2.2), o-suite presents lower correlation coefficient (0.77) than in the case of the control experiment (0.88). The decrease in the skill score of the o-control is associated to the prediction of the most extreme events during the period of 20-31 October in which the control experiment better captures the maximum DOD peaks. In the North-Western Maghreb, the CAMS experiments show underestimations (MB of -0.08 for control and -0.12 for o-suite) in comparison with the AERONET observations and correlation coefficients of 0.79 for o-suite and 0.73 for control which are close to the results of the SDS-WAS Median Multi-model (0.79).

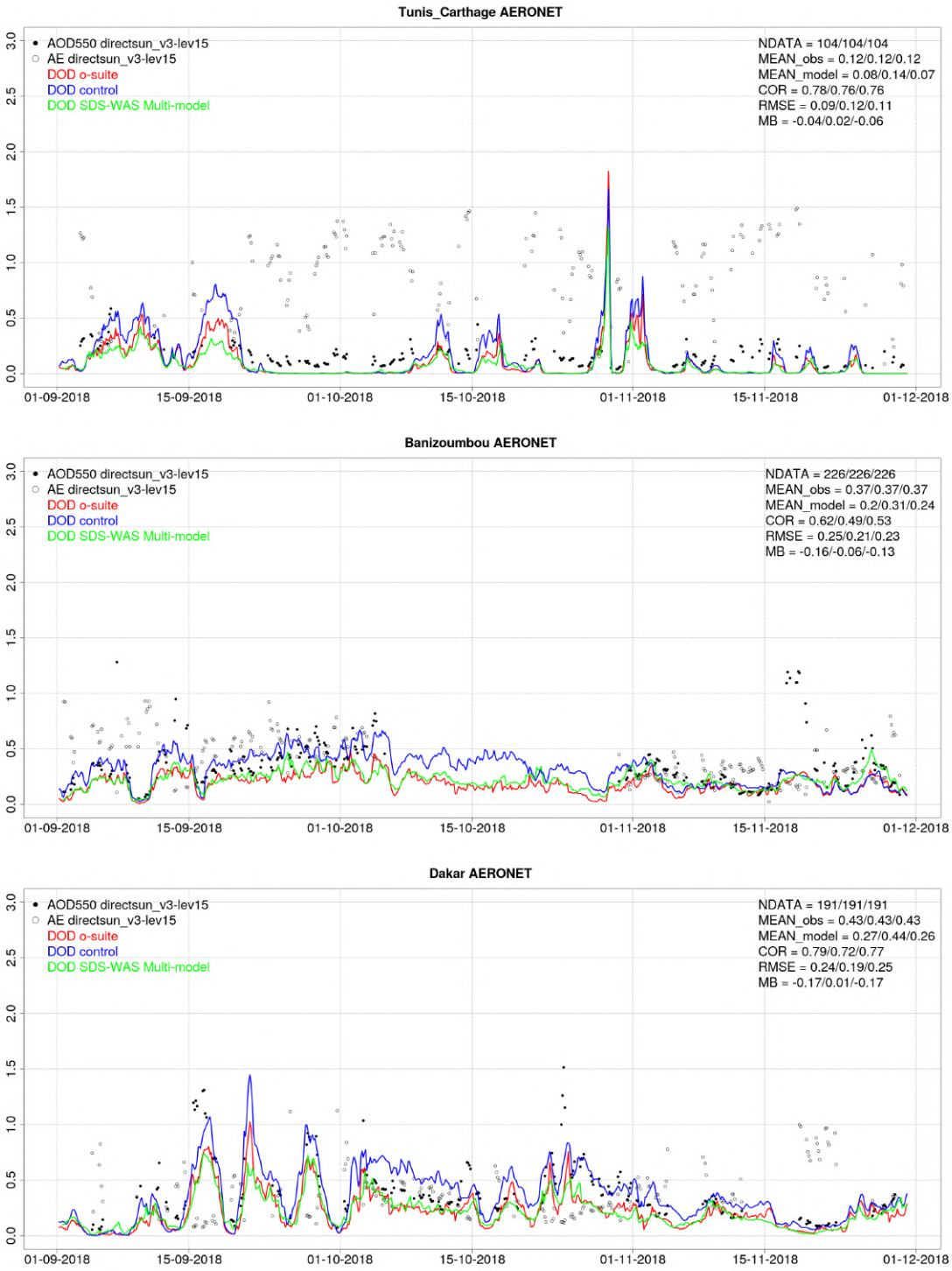
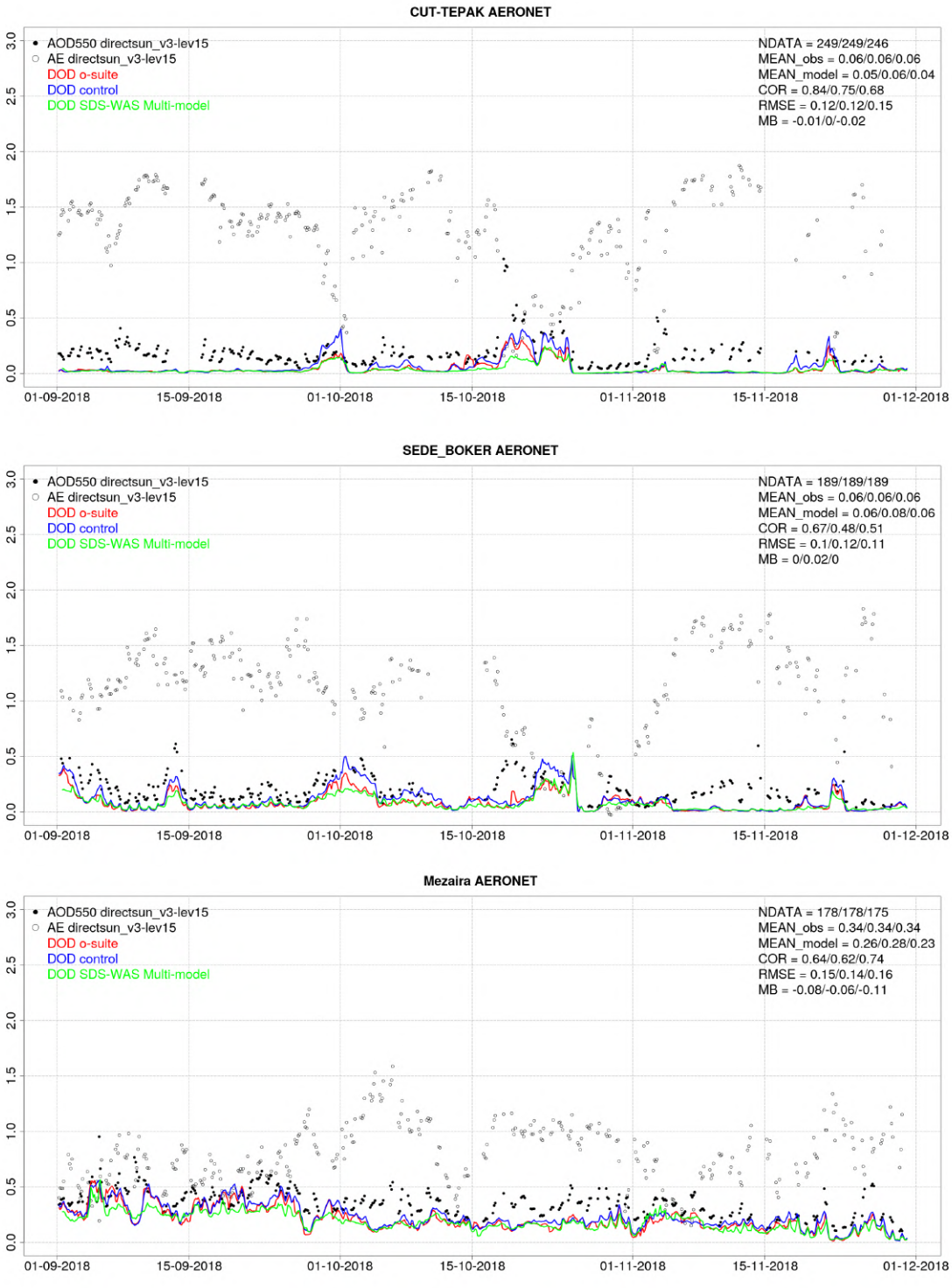


Figure 7.2.2: AOD from AERONET (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multimodel SDS-WAS Median (green line) for the study period over Tunis Carthage (North Western Magrebh), Banizoumbou (Sahel) and Dakar (Tropical North Atlantic). Skill scores per each individual site and model (o—suite/control/SDS-WAS Multi-model) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN_model, COR, RMSE, MB).



S

Figure 7.2.3: AOD from AERONET (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multimodel SDS-WAS Median (green line) for the study period over CUT-TEPAK and SEDE BOKER (Eastern Mediterranean) and Mezaira (Middle East). Skill scores per each individual site and model (o—suite/control/ SDS-WAS Multi-model) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN_model, COR, RMSE, MB).

Table 7.2.1: Skill scores (MB, FGE, RMSE and r) of 24h forecasts for CAMS o-suite, CAMS control and SDS-WAS Multi-model Median for the study period, and the number of data (NDATA) used. Dust AOD (DOD) from AERONET is the reference.

		control				o-suite DOD				SDS-WAS Median DOD			
	NDATA	MB	FGE	RMSE	r	MB	FGE	RMSE	r	MB	FGE	RMSE	r
Sahara	271	0.13	0.56	0.17	0.67	0.02	0.15	0.09	0.69	-0.03	-0.08	0.09	0.62
Sahel	676	-0.03	-0.06	0.21	0.59	-0.18	-0.50	0.26	0.67	-0.16	-0.39	0.25	0.65
Tropical North Atlantic	23	-0.04	-0.67	0.04	0.88	-0.06	-1.03	0.06	0.77	-0.06	-1.10	0.06	0.79
Subtropical North Atlantic	258	0.00	0.22	0.09	0.70	-0.02	0.11	0.08	0.78	-0.05	-0.09	0.09	0.74
North Western Maghreb	129	-0.08	-0.18	0.15	0.73	-0.12	-0.39	0.17	0.79	-0.12	-0.40	0.16	0.79
Western Iberian Peninsula	578	-0.02	0.67	0.06	0.62	-0.03	0.63	0.06	0.67	-0.02	0.64	0.06	0.67
Iberian Peninsula	828	0.00	1.40	0.04	0.76	-0.01	1.38	0.05	0.79	-0.01	1.38	0.05	0.77
Western Mediterranean	1351	-0.01	1.35	0.05	0.72	-0.01	1.32	0.06	0.72	-0.01	1.31	0.06	0.74
Central Mediterranean	1444	0.02	1.34	0.08	0.84	-0.01	1.27	0.06	0.84	-0.01	1.25	0.06	0.85
Eastern Mediterranean	1283	0.00	1.48	0.11	0.68	-0.01	1.44	0.11	0.78	-0.02	1.42	0.12	0.66
Eastern Sahara	-	-	-	-	-	-	-	-	-	-	-	-	-
Middle East	309	-0.05	0.33	0.18	0.76	-0.08	0.27	0.19	0.77	-0.11	0.18	0.23	0.73
All sites	**7150**	**0.00**	**1.04**	**0.11**	**0.81**	**-0.03**	**0.94**	**0.11**	**0.83**	**-0.04**	**0.92**	**0.12**	**0.83**

In the Middle East (see Table 7.2.1 and the Mezaira AERONET site in Figure 7.2.3), o-suite better reproduces the daily variability than control (with correlation coefficient of 0.77 for o-suite and 0.76 for control) and the SDS-WAS Multi-model presents lower correlations (0.73). Underestimations are observed in both CAMS experiments (MB of -0.05 for control and -0.08 for o-suite).

The CAMS o-suite model results are better than control over sub-Tropical North Atlantic region, characterized by long-range transport, with correlation values from 0.70 for control to 0.78 for o-suite (see Table 7.2.1). Both CAMS models present high correlations between 0.72 and 0.84 and slight underestimations (MB between -0.01 and 0.02) over the Iberian Peninsula and the Mediterranean. The Eastern Mediterranean region (see SEDE BOKER and CUT TEPAK in Figure 7.2.3) is the region where o-suite shows an improvement with respect the control experiment in term of variability with a correlation coefficient that goes from 0.68 for control to 0.78 for o-suite shows. This is connected the improvement that represents the assimilation of observations to correct the most extreme events observed on 1st and 20th October.

The comparison of 1 to 3-day forecasts (see Table 7.2.2) shows that the prediction is stable during the 3-days forecasts with correlation coefficients of 0.83 (0.81), 0.81 (0.80), and 0.80 (0.79) respectively for 24, 48 and 72h forecasts for all the sites for o-suite (control). In Tropical North Atlantic, the correlation coefficient (r) for control increases from 0.86 to 0.90 in going from the 48h to 72h forecast. These scores are better than in the case of the o-suite that presents correlation coefficient (r) of 0.70 at 48h to 0.67 at 72h. Otherwise, in sub-Tropical North Atlantic, o-suite presents better scores that control with correlation coefficients of 0.74 (0.65) at 48h and 0.63 (0.56) at 72h for o-suite (control).

Table 7.2.2: Skill scores (MB, FGE, RMSE and r) of 48h and 72h forecasts for CAMS o-suite and CAMS control for the study period, and the number of data (NDATA) used. Dust AOD (DOD) from AERONET is the reference.

		48h control				48h o-suite				72h control				72h o-suite			
	NDATA	MB	FGE	RMSE	r	MB	FGE	RMSE	r	MB	FGE	RMSE	r	MB	FGE	RMSE	r
Sahara	268	0.13	0.57	0.17	0.69	0.05	0.26	0.10	0.69	0.13	0.55	0.17	0.59	0.06	0.31	0.12	0.55
Sahel	676	-0.03	-0.06	0.21	0.59	-0.15	-0.40	0.25	0.61	-0.03	-0.05	0.21	0.59	-0.13	-0.34	0.24	0.60
Tropical North Atlantic	23	-0.04	-0.71	0.04	0.86	-0.05	-0.93	0.06	0.70	-0.03	-0.71	0.04	0.90	-0.05	-0.91	0.06	0.67
Subtropical North Atlantic	258	0.00	0.17	0.10	0.65	-0.03	0.05	0.08	0.74	-0.01	0.11	0.11	0.56	-0.03	0.00	0.10	0.63
North Western Maghreb	125	-0.07	-0.16	0.14	0.69	-0.11	-0.36	0.15	0.71	-0.05	-0.09	0.15	0.57	-0.09	-0.29	0.15	0.62
Western Iberian Peninsula	568	-0.02	0.67	0.06	0.59	-0.02	0.64	0.06	0.61	-0.01	0.70	0.06	0.40	-0.02	0.68	0.05	0.46
Iberian Peninsula	807	0.00	1.40	0.04	0.72	-0.01	1.38	0.05	0.73	0.00	1.42	0.05	0.50	0.00	1.40	0.04	0.53
Western Mediterranean	1321	-0.01	1.35	0.05	0.71	-0.01	1.33	0.06	0.71	0.00	1.37	0.06	0.66	-0.01	1.35	0.05	0.65
Central Mediterranean	1417	0.02	1.33	0.08	0.83	0.00	1.28	0.06	0.85	0.02	1.32	0.08	0.80	0.00	1.28	0.06	0.82
Eastern Mediterranean	1272	0.00	1.48	0.11	0.66	-0.01	1.44	0.11	0.72	0.00	1.47	0.11	0.65	-0.02	1.44	0.12	0.69
Eastern Sahara	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Middle East	304	-0.05	0.34	0.18	0.76	-0.07	0.29	0.20	0.73	-0.04	0.38	0.18	0.73	-0.06	0.34	0.19	0.71
All sites	**7039**	**0.00**	**1.04**	**0.11**	**0.80**	**-0.03**	**0.95**	**0.11**	**0.81**	**0.00**	**1.04**	**0.11**	**0.79**	**-0.02**	**0.97**	**0.11**	**0.80**

7.3 Backscatter profiles

Technical specifications of data sources, evaluated parameters and methods are described in Douros et al. (2017). In this section, the temporal and vertical variation of the backscatter coefficient (BSC) profiles are evaluated, statistically as bias, correlation, and standard deviation of o-suite '0001' and control run 'gzyh' versus ceilometers, and summarized in Taylor plots. Second topic is the reproduction of the planetary boundary layer (PBL), and thirdly, the representation of individual aerosol types is considered. Covariance plots of daily 2-D time-height sections serve for case studies and for development of a fractional skill score. The vertically integrated BSC is not in focus, because it reveals similar but less accurate information like AOD. As several case studies have confirmed the reliability of the methodology, we focus towards evaluation metrics and interpretations of aerosols representation in the model versions.

Period Overview

The model aerosol optical depth (AOD) and ceilometer overviews exhibit periods with significant aerosol plumes over Germany. Figure 7.3.1 shows the maximum AOD over Germany, separately for contributions of mineral dust (SD), sea salt (SS), carbonaceous matter (CM), black (BC) and organic carbon (OC), as well as sulfate (SU). Weak SD events occurred in the SON 2018 period around 15, 29 Oct 2018. The other components follow their usual seasonality.

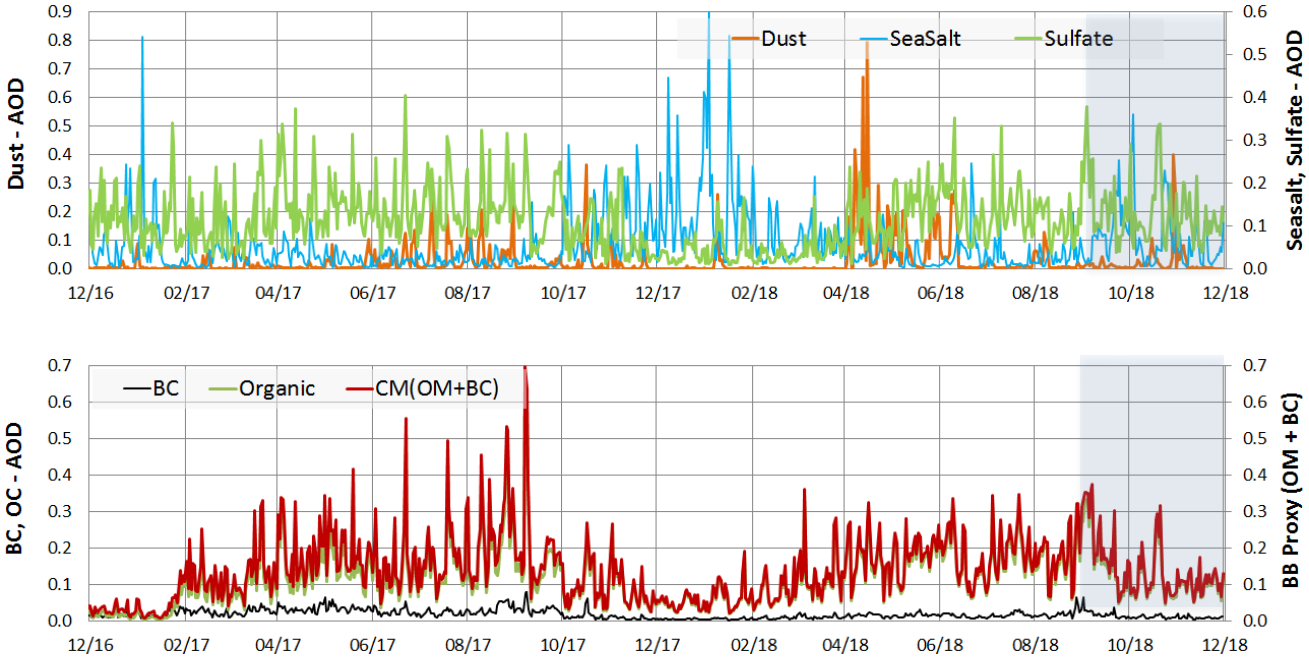


Figure 7.3.1: Maximum CAMS daily AOD over Germany for aerosol types included in the IFS model from 12/2016 - 11/2018: sea salt (blue), dust (orange), sulphate (light green), black carbon (BC, black), organic matter (green), proxy for 'biomass burning' (as OC+BC - red). Note the different y-axes for the aerosol species.

Mean profiles:

Model BSC in the PBL are on average lower than observed (cf. Tab. 3.5.1). The annual variation of the PBL height is rudimentary reproduced but is strongly smoothed. While enhanced emissions of organic matter (OM) have been introduced in Jan 2017, and parametrizations of SO2/SO4 conversion/deposition were improved, nitrate and ammonia are still missing in the current model version, which contribute roughly 10-30% of aerosol mass (as NO3NH4 or (NH4)2SO4) in the rural central European PBL. (According to pers. communication - S. Remy/Z. Kipling - nitrates and ammonium are ready in the current model and possibly get activated in the next cycle). Technically, this could be compensated by assimilating observational data in the o-suite, but without profile information this results in a vertical redistribution with a positive bias in the free troposphere, a too smooth transition at the PBL top (Fig. 7.3.2) and a negative bias in the PBL. Thus, the assimilation increases aerosol loads in the o-suite relative to the control run, but does not introduce a realistic step at the top of the PBL to lower values in the free troposphere (FT) as in the ceilometer profiles. Therefore, the amplitude of the vertical profile, coded in the standard deviation in the Taylor plots, is too low of the model compared to observations (reference).

Secondly, our forward operator (including mass→volume conversion) presently uses particle densities of the pure materials, not taking into account that dry atmospheric particles are often porous (sponge-like, even fractal) with entrapped air owing to coagulation and variable internal mixing, and thus exhibit reduced bulk density. A high-biased particle density results in low-biased equivalent volume, and a corresponding underestimation of optical properties, because these depend strongly on the particle size. Density reductions for accumulation mode particles, composed of hydrophilic and hydrophobic materials may be as high as a factor 1.5 (~1.3 for surface). Thirdly,

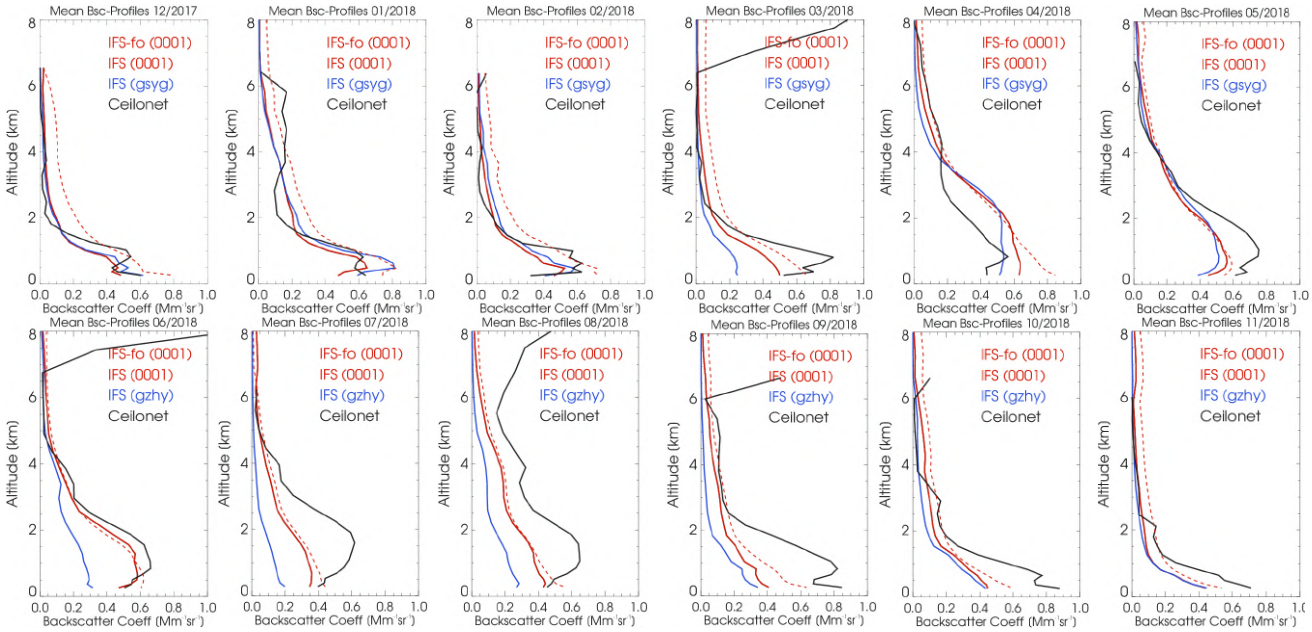


Figure 7.3.2: Monthly mean profiles of backscatter coefficients from o-suite (red), control run (blue), and ceilometers (black) combined from 21 German stations in Dec 2017 to Nov 2018. Two profiles are given for the o-suite each, calculated with forward operators from DWD (IFS-fo) and ECMWF (IFS), respectively. The profiles are partly contaminated by remaining cloud artefacts.

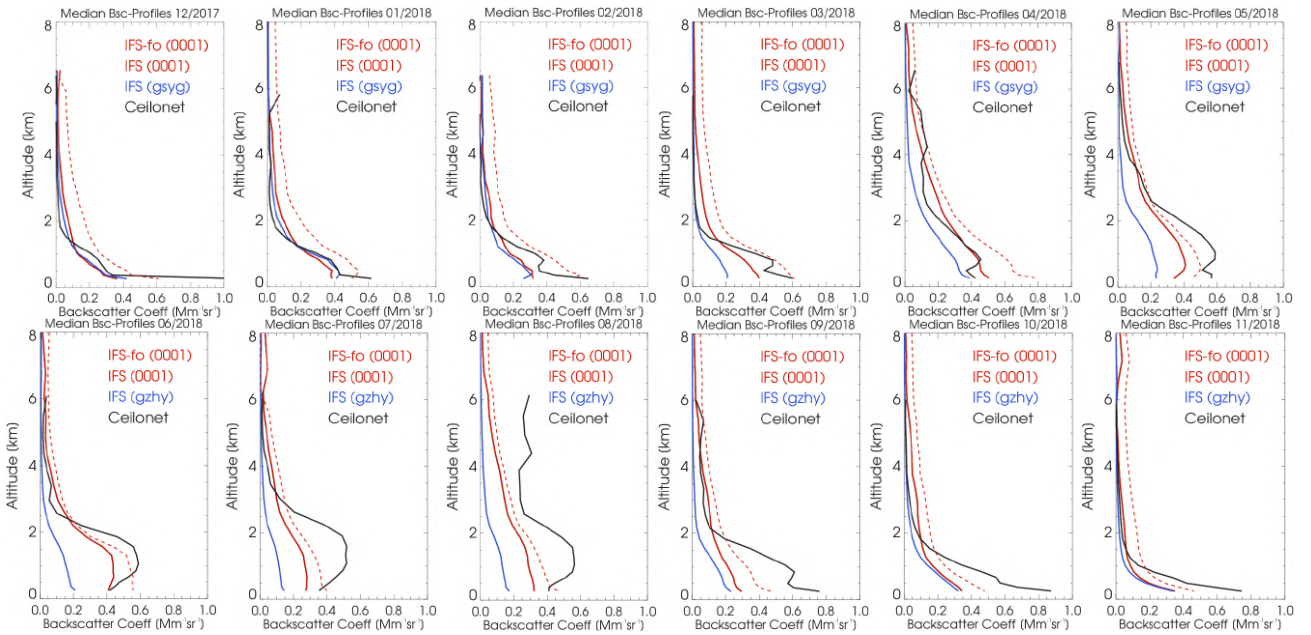


Figure 7.3.3: Monthly median profiles of backscatter coefficients from o-suite (red), control run (blue), and ceilometers (black) combined from 21 German stations in Dec 2017 to Nov 2018. Two profiles are given for the osuite each, calculated with forward operators from DWD and ECMWF, respectively. The profiles are partly contaminated by remaining cloud artefacts.

the capping transport barrier at the PBL top is less effective in the model, diluting high PBL concentrations with clean FT air. On average, however, the PBL height seems reasonable (this has been discussed in an earlier report).

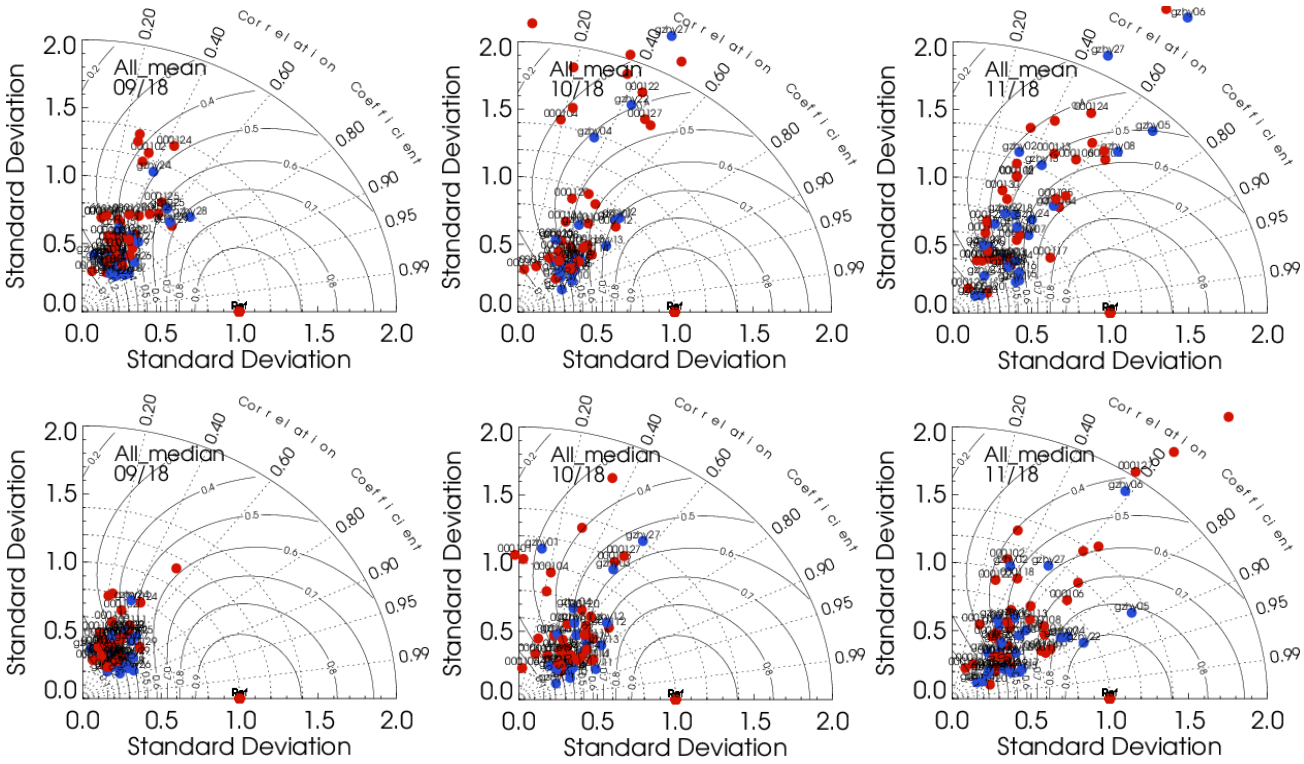


Figure 7.3.4: Taylor polar plots with daily average standard deviation of vertical profiles vs correlation coefficient, averaged over 21 German ceilometer sites for Sept, Oct, Nov 2018. O-suite red, control blue. Top: mean values, Bottom: median values.

As before, Figs. 7.3.2/3 show for the o-suite each two BSC profiles, calculated from the model mass mixing ratios, one using a forward-operator (foFO calculated by ourselves (DWD) and the other with the one implemented in the IFS. Notably, the difference between the monthly mean model BSC profiles calculated from our (DWD) forward operator and the one calculated within the IFS (IFS-FO), again shows larger values in autumn than in summer where it was few % small. Though some uncertainties are involved in calculating relative humidity and handling (strongly non-linear) hygroscopic growth, these do not seem to affect the nearly particle-free troposphere, as calculating the profiles with forced high (90%) relative humidity by the DWD-FO does hardly change the FT profile where the bias is largest. In individual cases, the IFS-FO profile is smoother. Other possible reasons like consistency of mass → backscatter conversion coefficients, particle densities, handling of the molecular portion etc. are currently checked.

Taylor Plots

The average coefficient of correlation between modelled and observed vertical backscatter profiles clusters around r = 0.4-0.6. The absolute standard deviation (SD) are normalized to the SD of observations per day, as reference value at SD ≡ 1.0. The better performance of the o-suite (red dots) compared to the control run (blue dots) in terms of SD, observed during summer, becomes less marked during autumn. There is a very large day-to-day and also a seasonal variation of the performance.

Summary

Backscatter coefficients are low-biased in the planetary boundary layer (PBL). Possible reasons are missing of ammonia and nitrate in the model, assumption of too high particle densities (for pure compact materials) in the mass to backscatter conversion, and the lack of vertical transport barrier at the top of the PBL, causing dilution with free troposphere air. Free troposphere (FT) background backscatter coefficients are biased high, probably due to wrong re-distribution between PBL and FT. This is not fixed by the assimilation, which instead adds aerosol to the whole profile. The bsc bias on a specific level thus depends on its relative position w.r.t. to the BLH.

Monthly average vertical backscatter profiles show a high low bias in the extraordinary warm summer 2018. Interestingly, the model profiles, based on DWD's forward operator and the one implemented in the IFS (IFS-FO), differ depending on season and are few % small only in summer. The reason for this discrepancy is currently investigated. The monthly Taylor plots over 21 German stations indicate that daily averages of Pearsons correlation coefficients cluster around 0.4-0.6, and that the o-suite still (though less than in summer) on average has a lower bias than control.

7.4 Aerosol validation over the Mediterranean

Three-hourly aerosol optical depth (AOD) and surface concentration (PM10 and PM2.5) from o-suite experiment and control experiment have been validated for the period 1 September 2018 – 30 November 2018 against AERONET direct-sun cloud-screened observations.

Aerosol optical depth

CAMS o-suite can reproduce the daily variability of AERONET observations. In Western, Central and Eastern Mediterranean, the correlation coefficient increases from 0.67, 0.60 and 0.55, to 0.75, 0.67 and 0.62, respectively for control and o-suite during autumn (see the correlation coefficient by sites in Figure 7.4.1). Underestimations observed in the Mediterranean Basin in control (MB of -0.03, 0 and -0.01 for Western, Central and Eastern Mediterranean regions respectively) are corrected in o-suite introducing overestimations in the Basin (MB of 0.01, 0.01 and 0.02 for Western, Central and Eastern Mediterranean regions respectively) as it is shown in Figure 7.4.1. The highest peaks on CAMS AOD simulations are linked to desert dust intrusions. Second-half of October was particularly intense in terms of dust outbreak in Eastern Mediterranean achieving AOD values up to 1.4 (see IMS-METU-ERDEMLI AERONET sites in Figure 7.4.2). Meanwhile, Western Mediterranean sites received the most intense dust outbreaks during September (see Palma de Mallorca in Figure 7.4.2).

Surface aerosol concentrations

For autumn, PM10 and PM2.5 results of CAMS o-suite and control show similar skill scores in comparison with EIONET-Airbase observations (see Figure 7.4.3). CAMS model tends to overestimate the PM10 and PM2.5 EIONET-Airbase observations in north-western Mediterranean sites (see Figure 7.4.3 and Figure 7.4.4) meanwhile tends to underestimate the PM10 and PM2.5 observed values at Spanish sites.

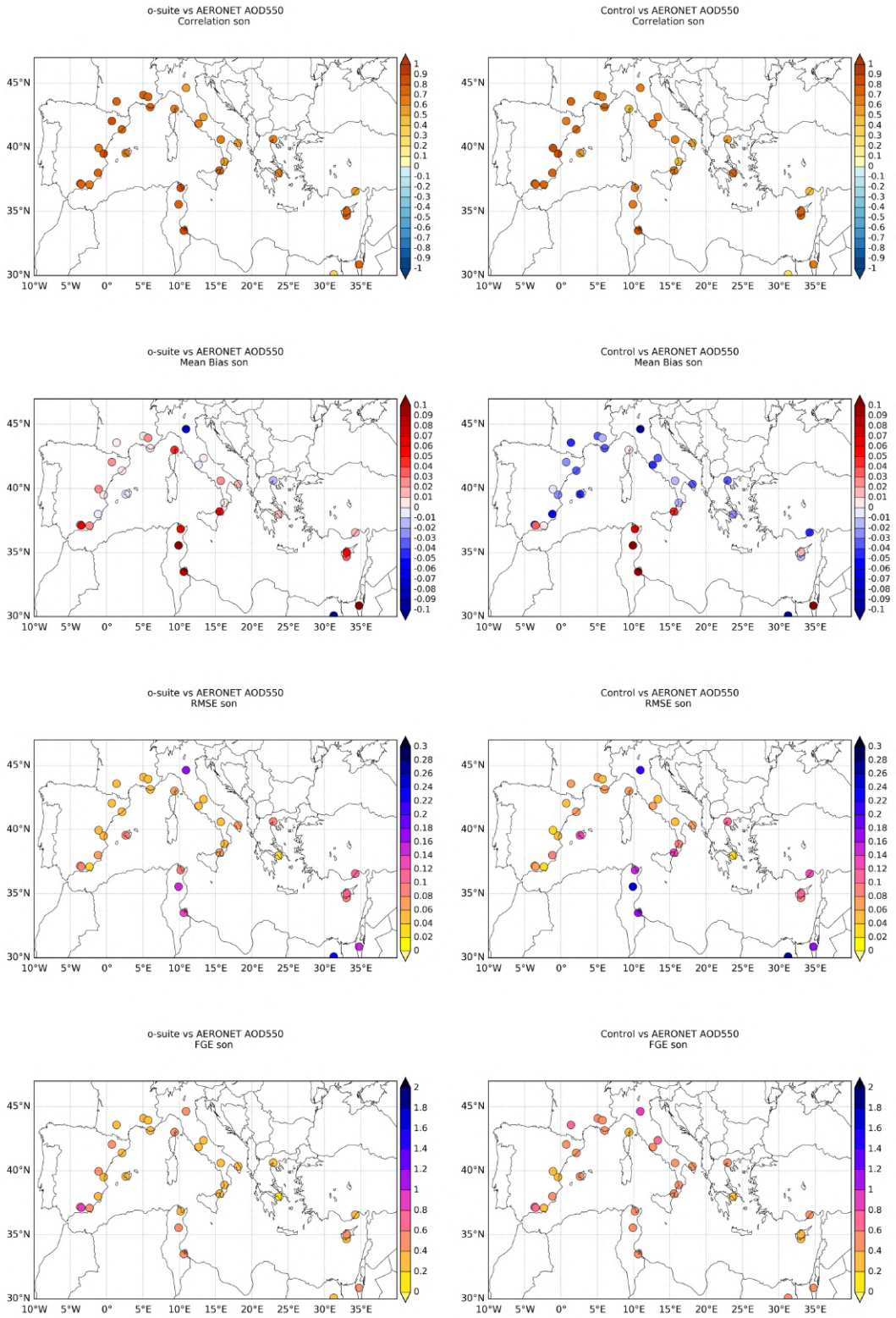


Figure 7.4.1: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. AOD from AERONET is the reference.

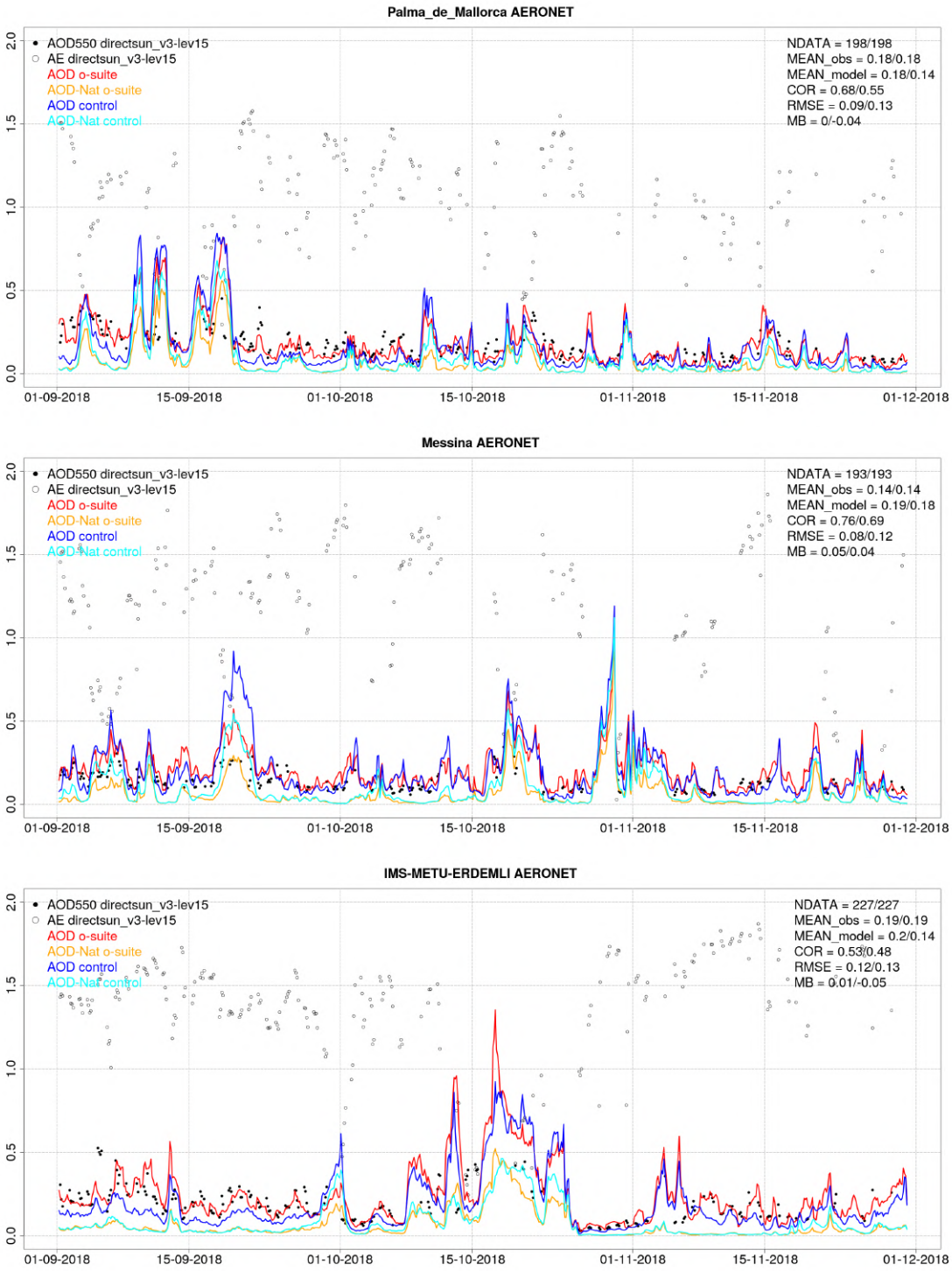


Figure 7.4.2: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), for the study period over Palma de Mallorca (Balearic Islands, Spain), Messina (Italy) and IMS-METU-ERDEMLI (Israel). AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt. Skill scores per each individual site and model (o—suite/control) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN_model, COR, RMSE, MB).

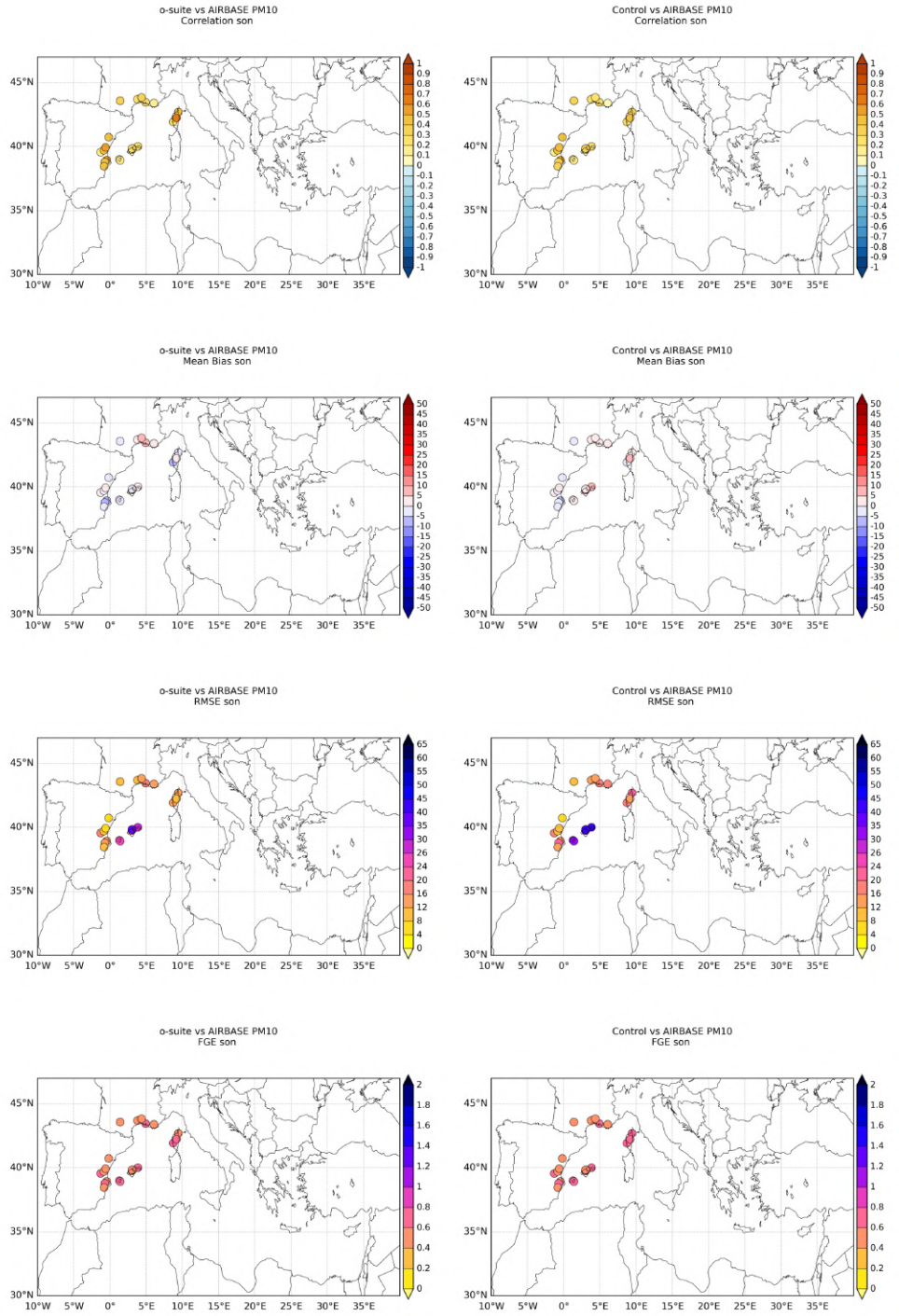


Figure 7.4.3: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. PM10 from EIONET are the reference. Only background suburban and rural available stations are displayed.

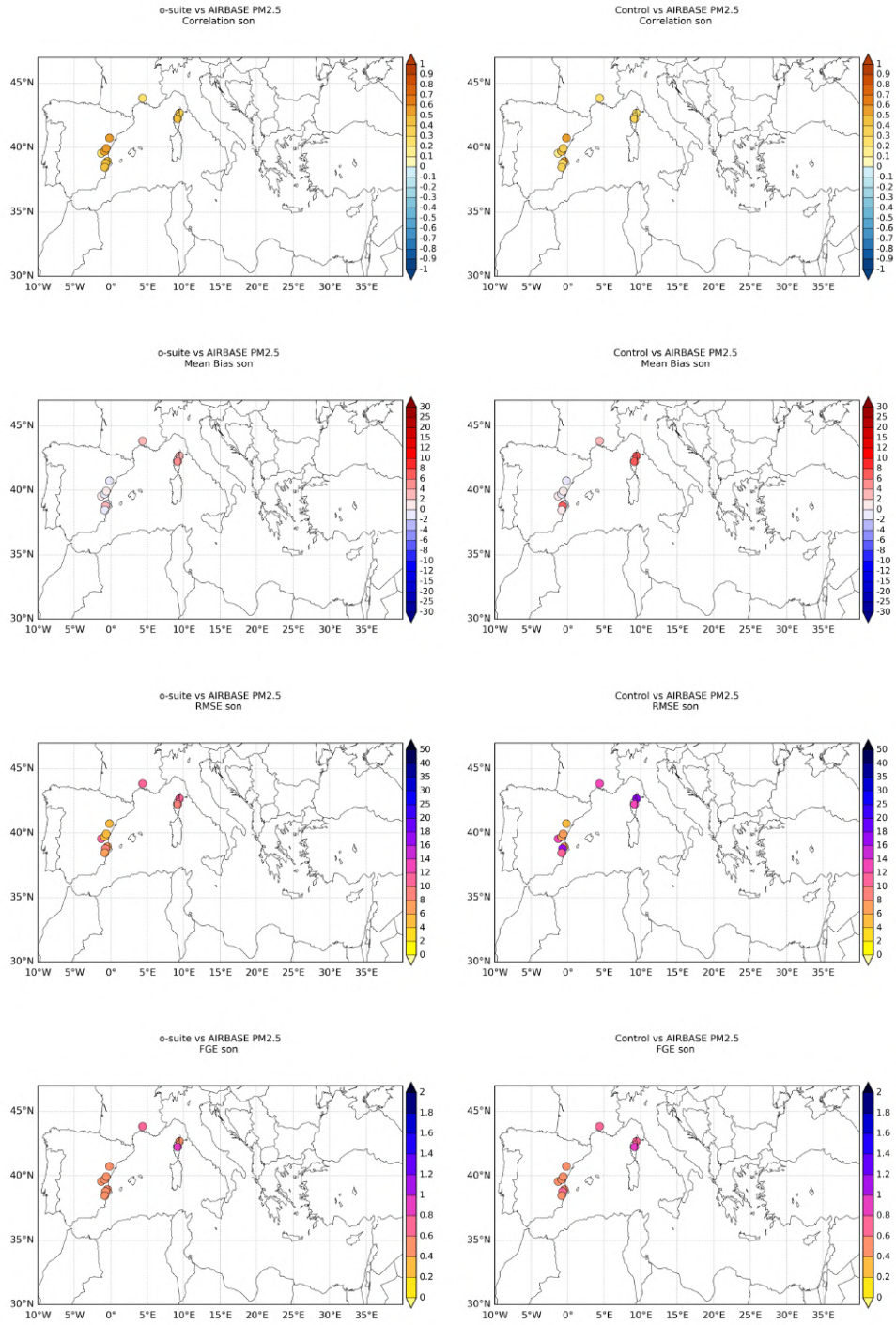


Figure 7.4.4: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. PM2.5 from EIONET are the reference. Only background suburban and rural available stations are displayed.

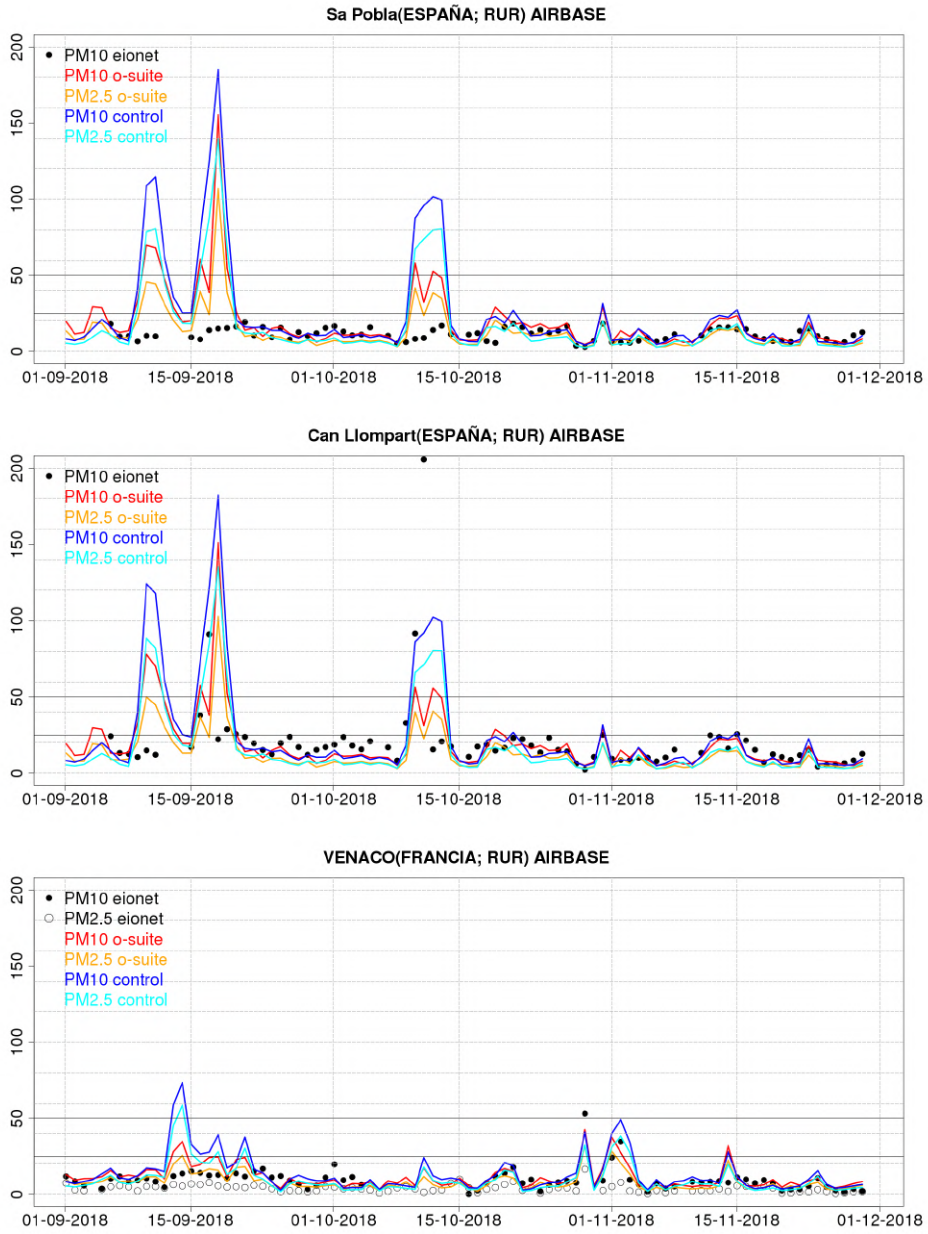


Figure 7.4.5: PM10 and PM2.5 Airbase observations (black and grey dots, respectively), PM10 and PM2.5 o-suite (red and orange lines, respectively) and PM10 and PM2.5 control (blue and cyan lines, respectively) for the study period over Sa Pobla and Can Llompart (Balearic Islands, Spain) and Venaco (Corse, France).

Overestimations at north-western Mediterranean sites are linked to the overestimation of the CAMS model of the dust outbreak in mid-September that it is better predicted by o-suite than control (see Venaco in Figure 7.4.5 in which control experiment shows concentration above 50 $\mu g/m^3$ and around 35 $\mu g/m^3$ o-suite). Meanwhile, in western Mediterranean the results are quite heterogeneous between a close clusters of sites. For example, Sa Pobla site is detecting any aerosol event above 50 50 $\mu g/m^3$ (which is the EU PM10 daily threshold) and Can Llompart is observed two aerosol event on 15-18 September and 12-14 October. In both sites, o-suite and control predicts similar results in better agreement with Can Llompart than Sa Pobla EIONET-Airbase sites.

8. Stratosphere

8.1 Validation against ozone sondes

In this section, we present the results of the stratospheric ozone evaluation against ozone soundings from the NDACC, WOUDC, NILU and SHADOZ databases. The sondes have a precision of 3-5% (~10% in the troposphere for Brewer Mast) and an uncertainty of 5-10%. For further details see Cammas et al. (2009), Deshler et al. (2008) and Smit et al (2007). Model profiles of the o-suite are compared to balloon sondes measurement data of 44 stations for the period January 2013 to November 2018 (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). The o-suite shows MNMBs within the range ±12%, for all regions and months (some exceptions with MNMBs of up to ±18% for single months in the high latitude regions). Figure 8.1.1. shows the results for the past year, and Figure 8.1.2 for the past 6 years.

Fig. 8.1.3 compares the averaged profiles in each region during March 2018. The vertical distribution of stratospheric ozone is quite well represented for all regions by the o-suite, with little overestimation in all latitude bands (MNMBs between 2 to 8% for SON).

The control run shows a strong overestimation of stratospheric ozone in the upper stratosphere, and an underestimation between 40hPa (Arctic 60 hPa) and 300 hPa in the Arctic and the Northern Midlatitudes. In the Tropics, the underestimation is between 80 and 30 hPa. The Antarctic profile shows an underestimation between 40 and 80 hPa. Above, O_3 partial pressures are overestimated.

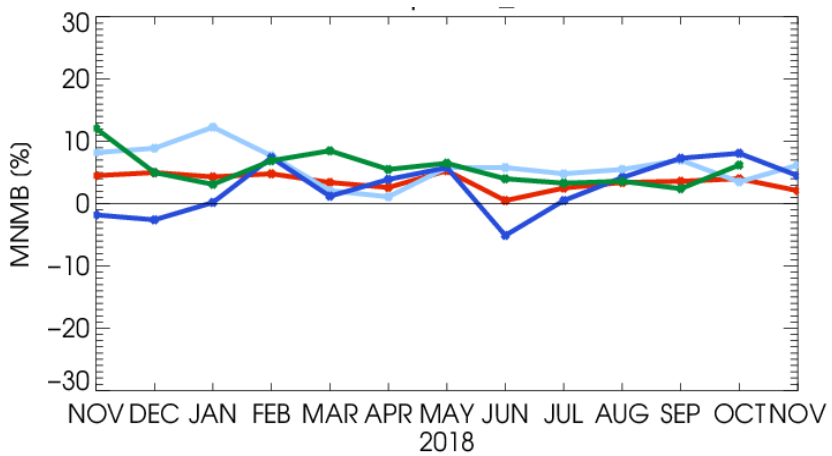


Figure 8.1.1: MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated sonde data in the Arctic (light blue), Antarctic (dark blue) northern midlatitudes (red) and tropics (green). Period November 2017 to November 2018. The stratosphere is defined as the altitude region between 60 and 10 hPa in the tropics and between 90 and 10 hPa elsewhere.

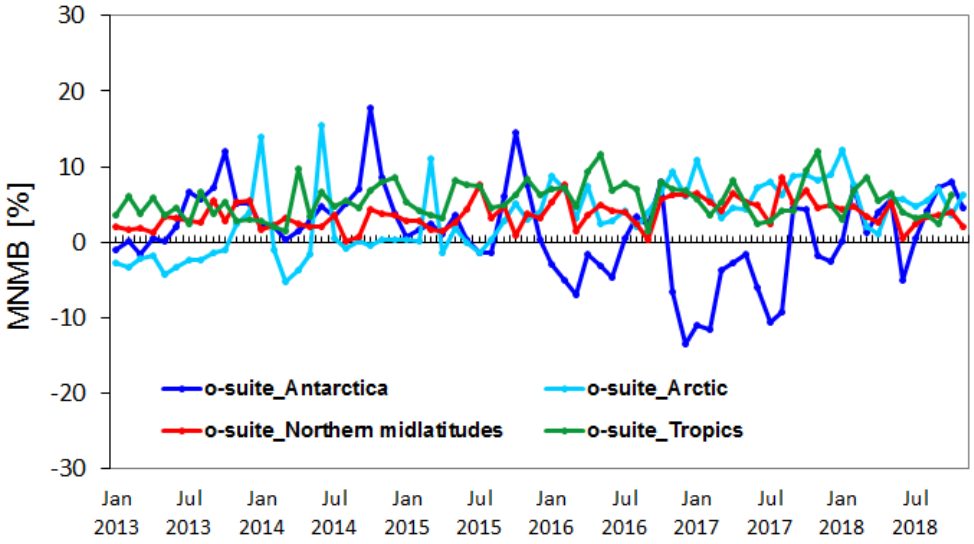


Figure 8.1.2: Same as Fig. 8.1.1, but for the extended time period January 2013 to November 2018.

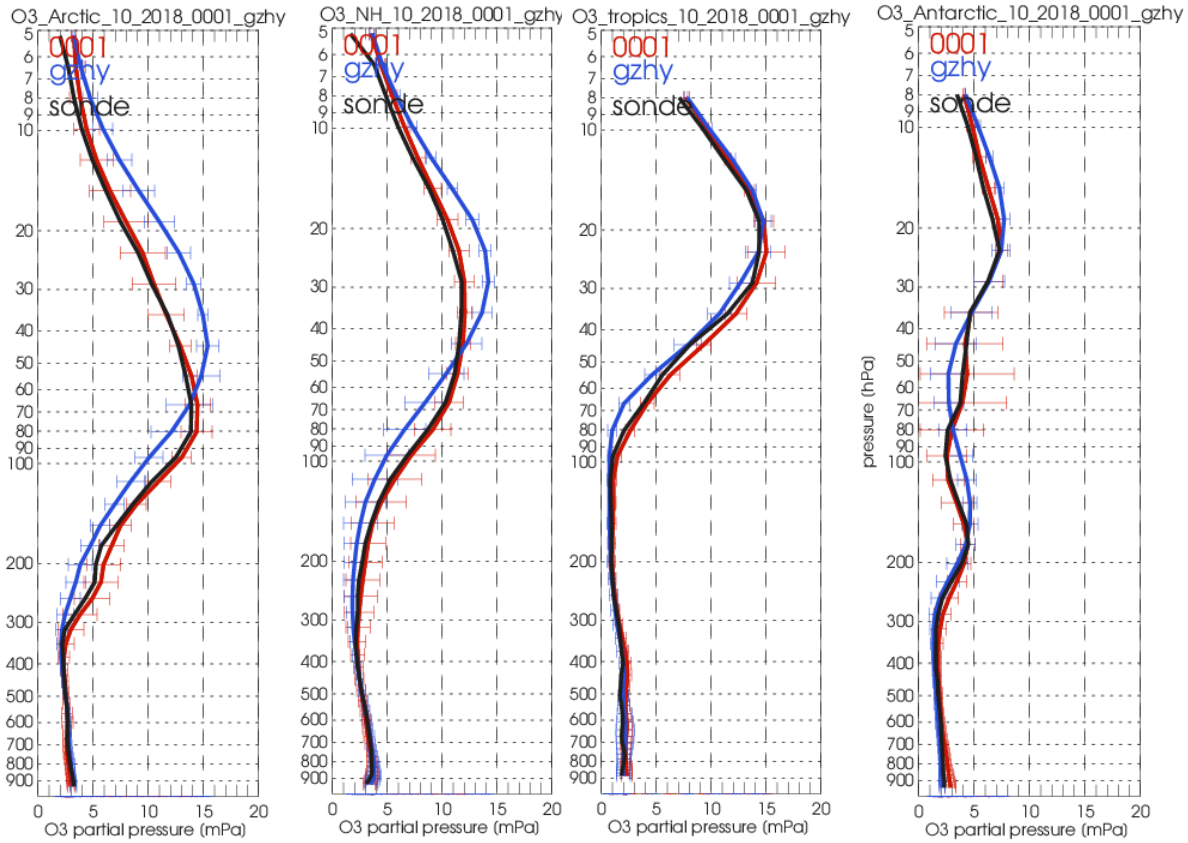


Figure 8.1.3. Comparison between mean O_3 profiles (units: mPa) of o-suite (red), and control (blue) in comparison with observed O_3 sonde profiles (black) for October 2018 for the various latitude bands: Arctic, NH-mid latitudes, Tropics and Antarctic.

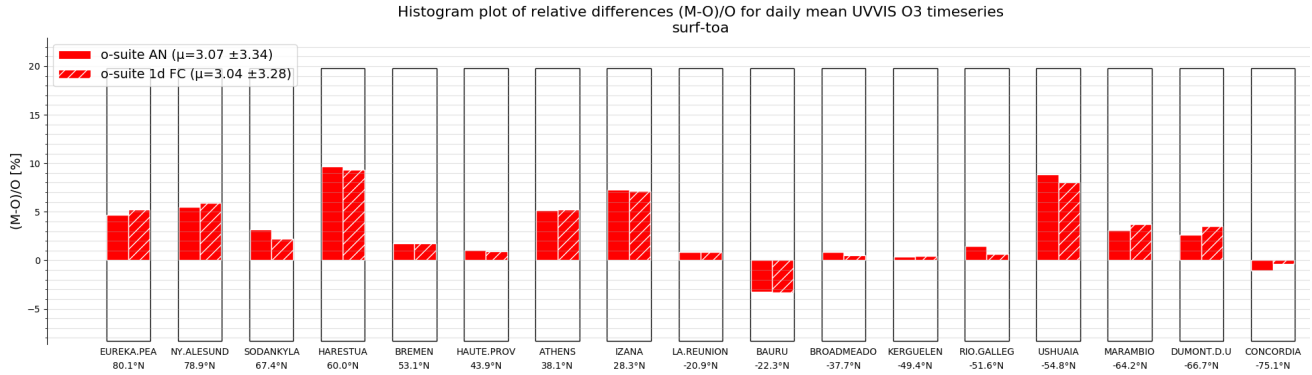


Figure 8.2.1. Relative biases during quarter SON 2018 for 17 UVVIS stations measuring stratospheric ozone columns with ZENITH measurement geometry (stations sorted with decreasing latitude). The overall relative bias is positive for all latitudes and within the typical measurement uncertainty of 5% for most of the sites.

8.2 Validation against observations from the NDACC network (UVVIS, FTIR, MWR and LIDAR)

UVVIS and FTIR stratospheric columns

Since the start of the CAMS27 project, the number of UVVIS Zenith ozone measurements have increased at NDACC. Currently fifteen sites provided data in the recent quarter allowing for a representative picture on the latitude dependence of the model data.

The systematic uncertainty of the UVVIS measurements is typically 5%, hence the relative biases for most sites for both the AN and 1d FC of the o-suite are very close to each other and within the uncertainty ranges, see Figure 8.2.1. The averaged bias for the 17 UVVIS sites is 4.5% and the averaged correlation is 0.93.

Figure 8.2.3 depicts the FTIR stratospheric columns showing a discontinuity in the o-suite 1d FC model for the tropical sites (Mauna Loa, Altzomoni and Reunion) in the June 2016 model update.

The correlations between the sites and the model are presented in the Taylor diagrams in Figure 8.2.2. Again, the o-suite AN and 1d FC perform very similarly in correlation coefficients. The FTIR Altzomoni (close to Mexico City) only has three co-located measurements days and no conclusions can be drawn from the apparent outlier in the analysis plot.

Profile comparison using LIDAR and MWR

In this section we present a comparison between the CAMS o-suite and control run models against MWR and LIDAR observations from the NDACC network. A detailed description of the instruments and applied methodologies for all NDACC instruments can be found at http://nors.aeronomie.be. MWR (microwave) at Ny Alesund (79°N, 12°E, Arctic station) and Bern (47°N, 7°E, northern midlatitude station). LIDAR at Observatoire Haute Provence (OHP), France (43°N, 5.7°E, altitude 650m) and Hohenpeissenberg, Germany (47°N, 11°E, altitude 1km)

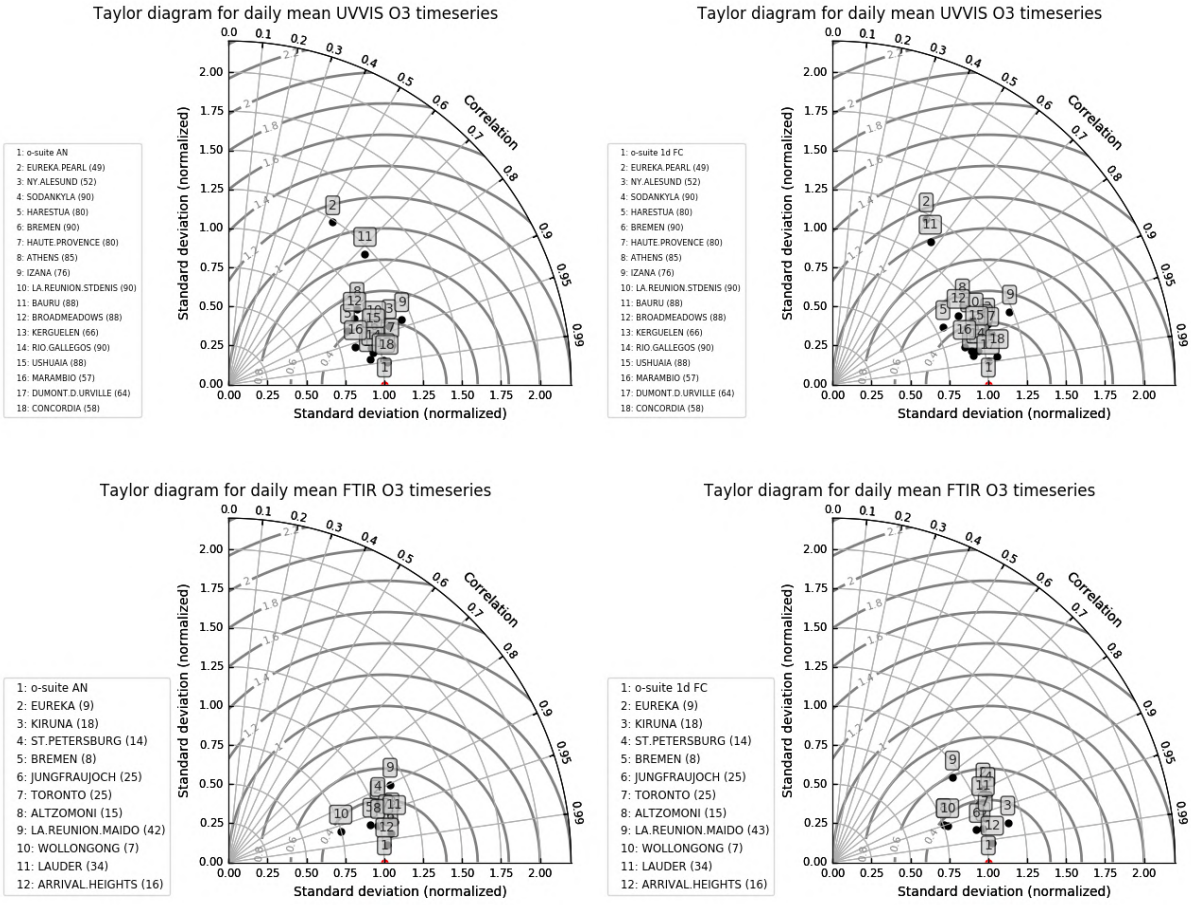


Figure 8.2.2 Taylor diagrams relating the standard deviations for the o-suite and ground-based stratospheric column time series and their correlation for the time period SON 2018. All timeseries are normalized such that the standard deviation of the model is 1. The performance for the o-suite analysis (two panels on the left) is slightly better (averaged correlation is 0.96 for FTIR and 0.91 for UVVIS) compared to the 1d FC (two panels on the right; averaged correlation is 0.93 for FTIR and 0.89 for UVVIS). Top panels: results for the UV-VIS instruments. Bottom panels: results for O3 measured by FTIR.

From Figure 8.2.3 at Ny Alesund the o-suite overestimates the stratospheric ozone concentration with more than 10% during SON/DJF/MAM and the bias vanishes during summer JJA. Between the model upgrades from September 2015 and January 2017 the relative bias at Bern vanishes (i.e. is comparable to the measurement uncertainty). In Figure 8.2.3 a discontinuity is observed for the mesospheric ozone column measured by the MWR at Bern at the Jan 2017 update which is due to an underestimation of the upper stratospheric ozone column (see Figure 8.2.4).

In MAM-JJA 2015-2018, both MWR stations observe a significant (i.e. comparable to the measurement uncertainty) overestimation of the upper stratospheric/mesospheric ozone content, and the converse is seen during autumn and winter SON-DJF, underestimating up to -30% at Ny Alesund, see also Figure 8.2.4Figure .

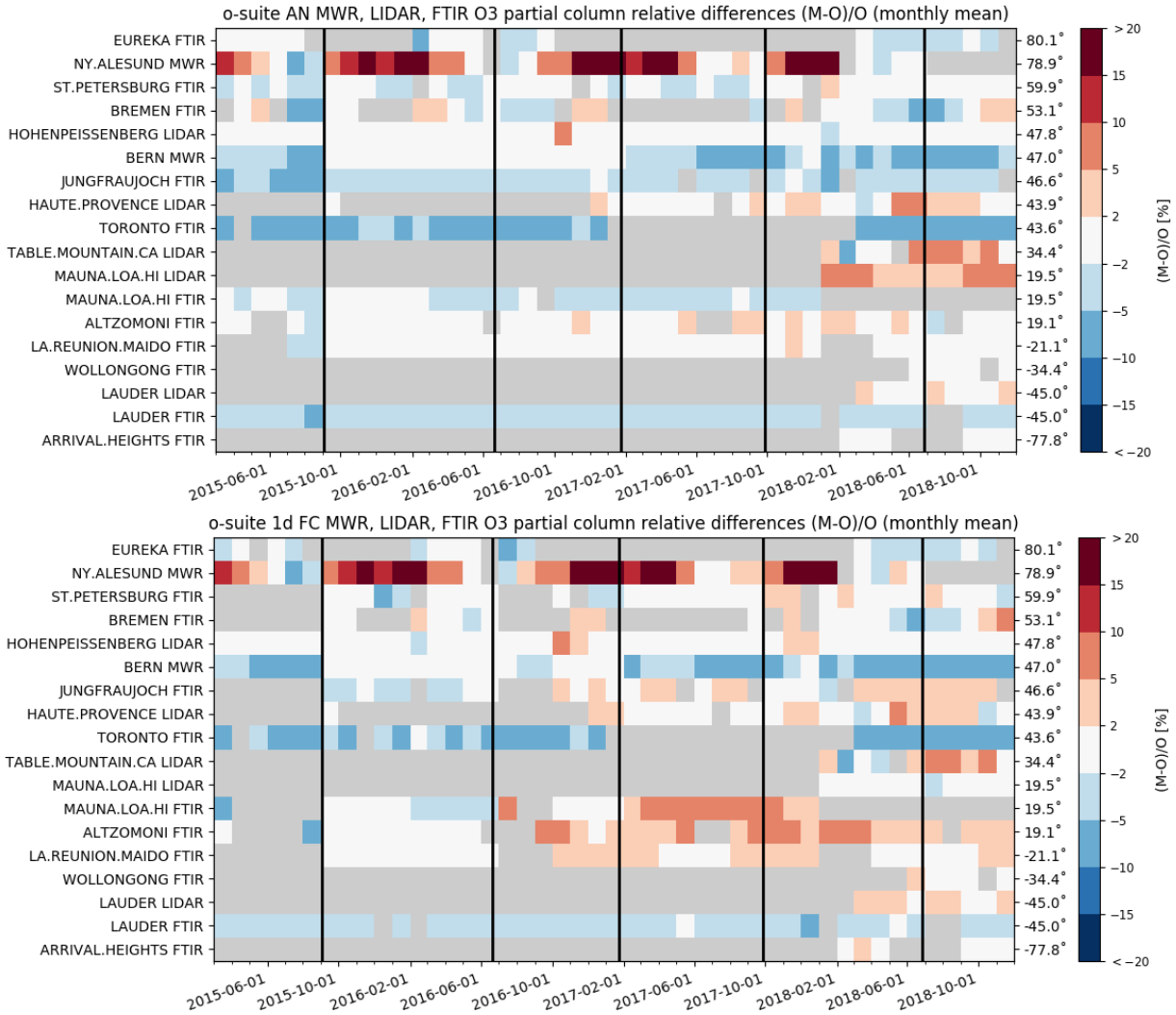


Figure 8.2.3 Time series of monthly mean relative differences for stratospheric columns (FTIR and LIDAR) and mesospheric (MWR) columns along with model cycle updates (black vertical lines) (o-suite AN top, o-suite 1d FC bottom). The stratospheric FTIR columns at St Peterburg, Jungfraujoch, Altzomoni and La Reunion show a slight positive trend in the relative differences, which is probably introduced by model updates.

At OHP 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)

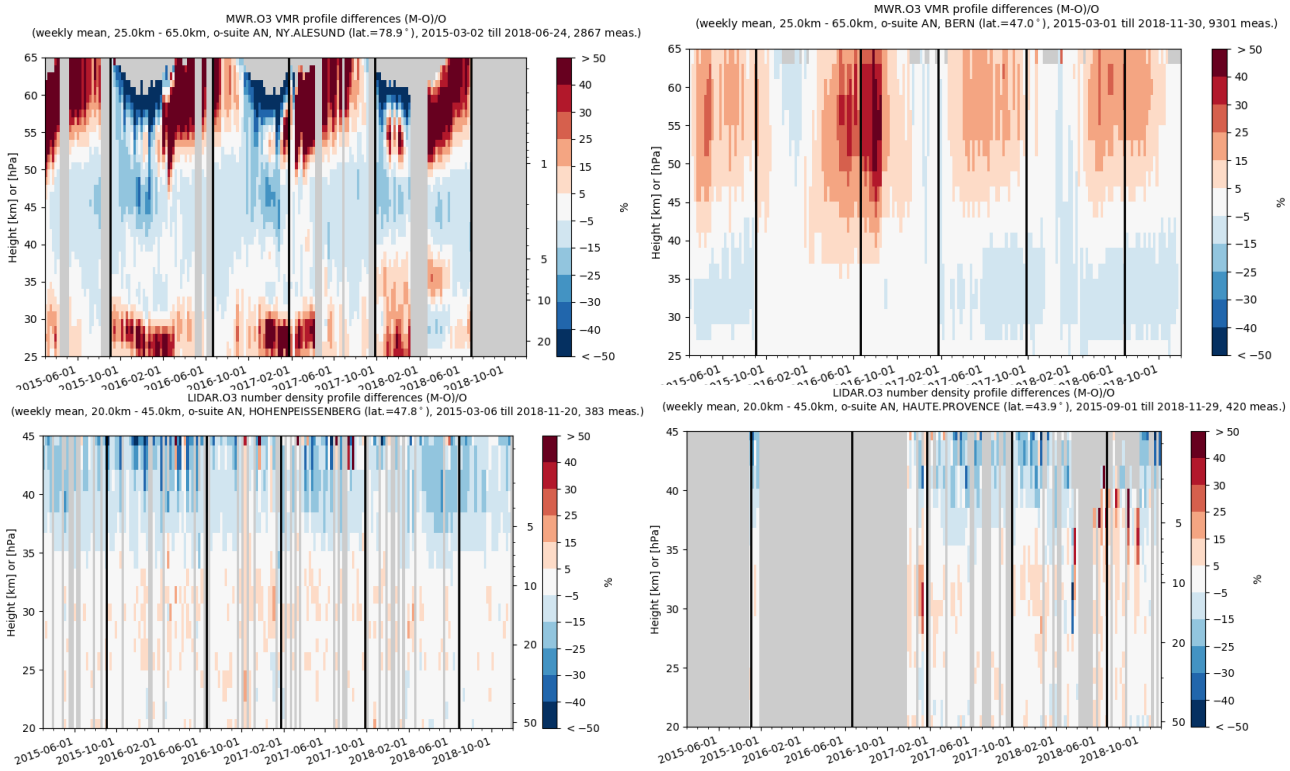


Figure 8.2.4: Comparison of the weekly mean profile bias between the O_3 mixing ratios of o-suite and the NDACC station at Ny Alesund, Bern, Hohenpeissenberg and OHP. For the LIDAR stations, the measurement uncertainty above 35km is comparable to the observed profile bias.

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

This section compares the output of the o-suite for the last period with observations by limb-scanning satellite instruments, using the methodology described by Lefever et al. (2015). We also include the comparisons for the o-suite 4th day forecasts (96h to 120h) of stratospheric ozone. These forecasts are represented by 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.

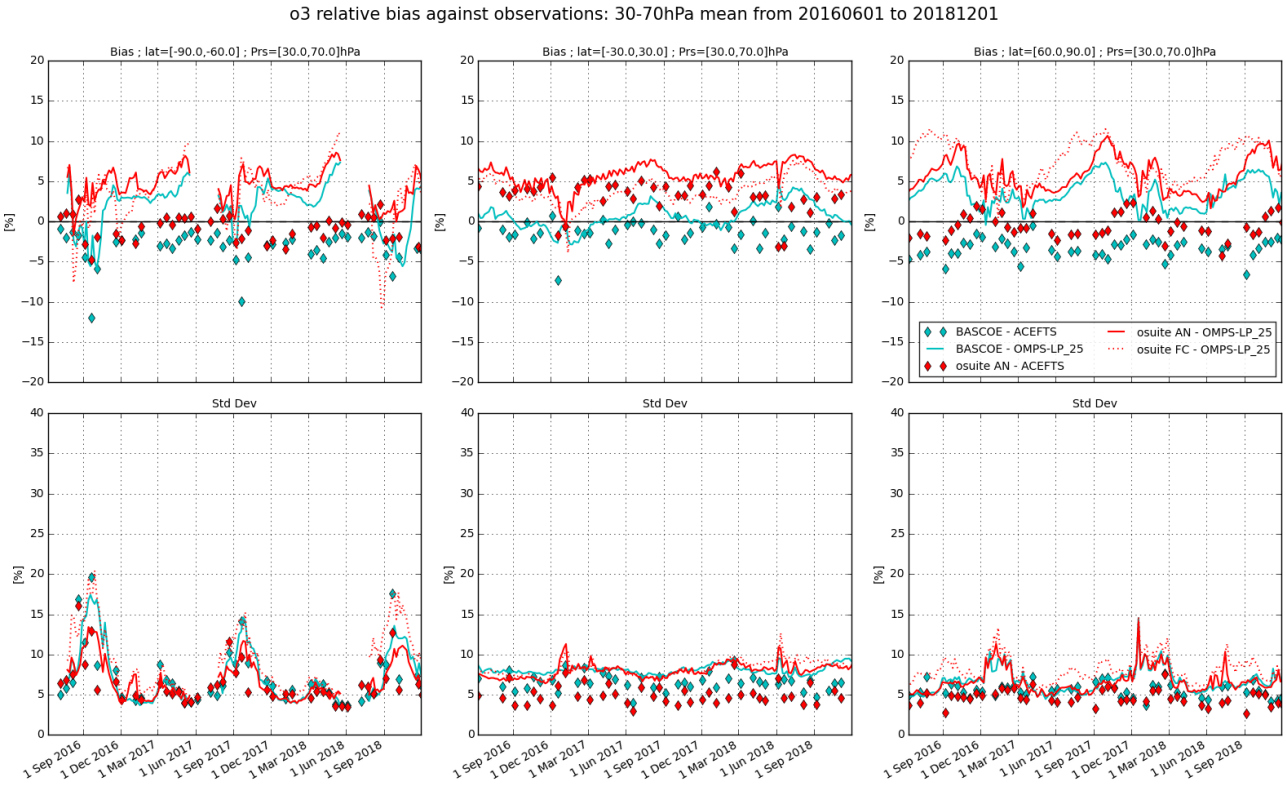


Figure 8.3.1: Time series comparing models to observations for the period 2016-06-01 to 2018-12-01 in the middle stratosphere (30-70hPa averages): o-suite analyses vs OMPS-LP (red, solid), o-suite forecasts 4th day vs OMPS-LP (red, dotted), o-suite analyses vs ACE-FTS (red markers), BASCOE vs OMPS-LP (cyan, solid) and BASCOE vs ACE-FTS(cyan markers). Top row, normalized mean bias (model-obs)/obs (%); bottom row, standard deviation of relative differences (%).

In this section, we use on one hand the version 2.5 of OMPS-LP (i.e. the Limb Profiler) and the version 3.6 of ACE-FTS. For reference, we include also the BASCOE analyses which are very constrained by the AURA MLS offline profiles.

Figure 8.3.1 and Figure 8.3.2 present, in the upper row, the timeseries over the last 30 months of the bias of the o-suite against the two satellite measurements for respectively two regions of the lower stratosphere and UTLS (30-70hPa and 70-100hPa); the bottom row of the figures shows the standard deviation of the differences and can be used to evaluate the random error in the analyses.

Compared to OMPS-LP in the 30hPa to 70hPa region, there is a systematic overestimation by the o-suite: up to 8% in the South polar region and in the tropics and up to 12% in the North polar region. Compared to OMPS-LP in the 70hPa to 100hPa region, the North polar bias increases up to 20% at various periods, while the variability of the bias is much stronger in the South polar region; the tropics exhibits a strong seasonal variation for the bias, with a high variability indicated by the standard deviation.

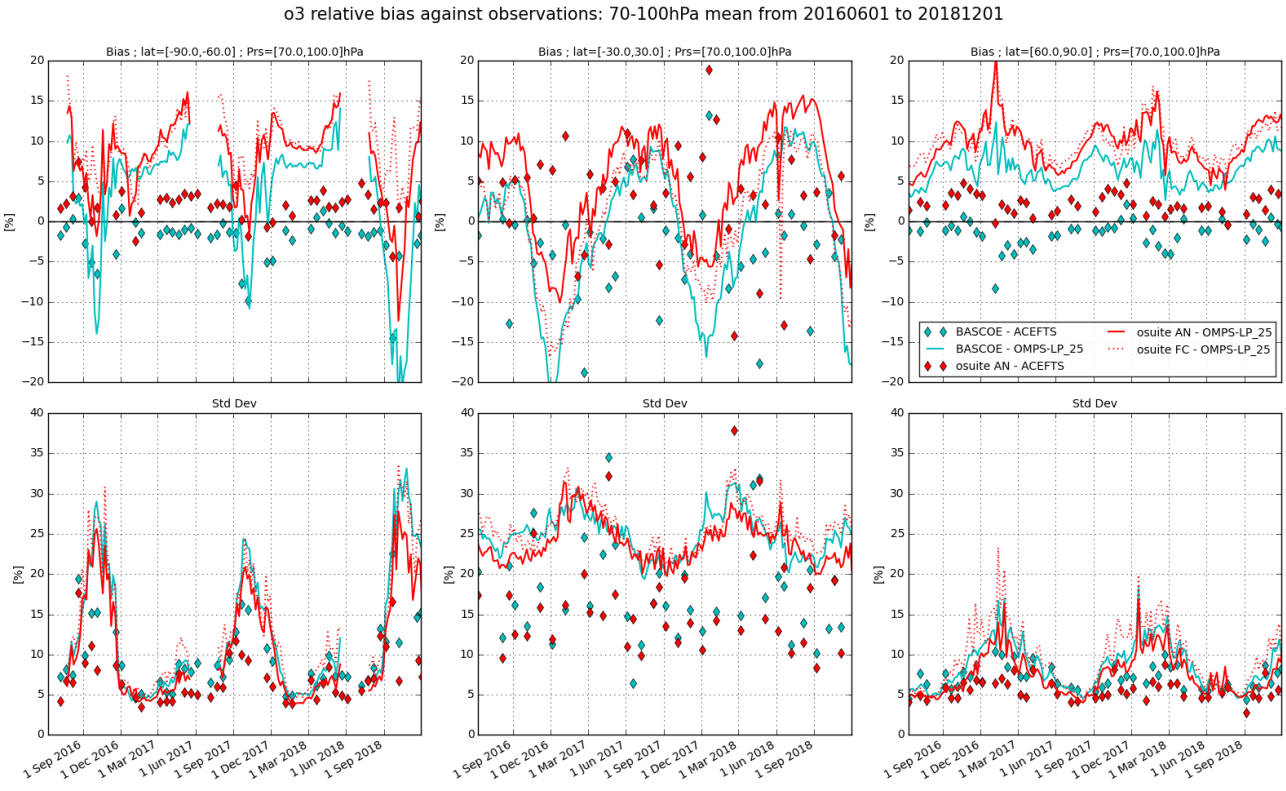


Figure 8.3.2: Time series comparing models to observations for the period 2016-06-01 to 2018-12-01 in the lower stratosphere (70-100hPa averages): o-suite analyses vs OMPS-LP (red, solid), o-suite forecasts 4th day vs OMPS-LP (red, dotted), o-suite analyses vs ACE-FTS (red markers), BASCOE vs OMPS-LP (cyan,solid) and BASCOE vs ACE-FTS(cyan markers). Top row, normalized mean bias (model-obs)/obs (%);bottom row, standard deviation of relative differences (%).

The agreement with ACE-FTS is much better: the bias is generally within ±5%, except in the tropics for 70hPa to 100hPa region, where the standard deviations indicate less reliable results.

The bias of BASCOE against the satellite observations for the considered regions is systematically lower, but follows a similar evolution as the o-suite.

Figure 8.3.3 and Figure 8.3.4 display vertical profiles of the relative biases between the o-suite or BASCOE and the satellite measurements. The difference is averaged over the most recent 3-month period considered in this validation report, i.e. June to August 2018.

The OMPS-LP profiles are much more irregular than the ACEFTS or MLS profiles, but the relative bias between o-suite and OMPS-LP is mostly within ±10% between 20km and 35km.

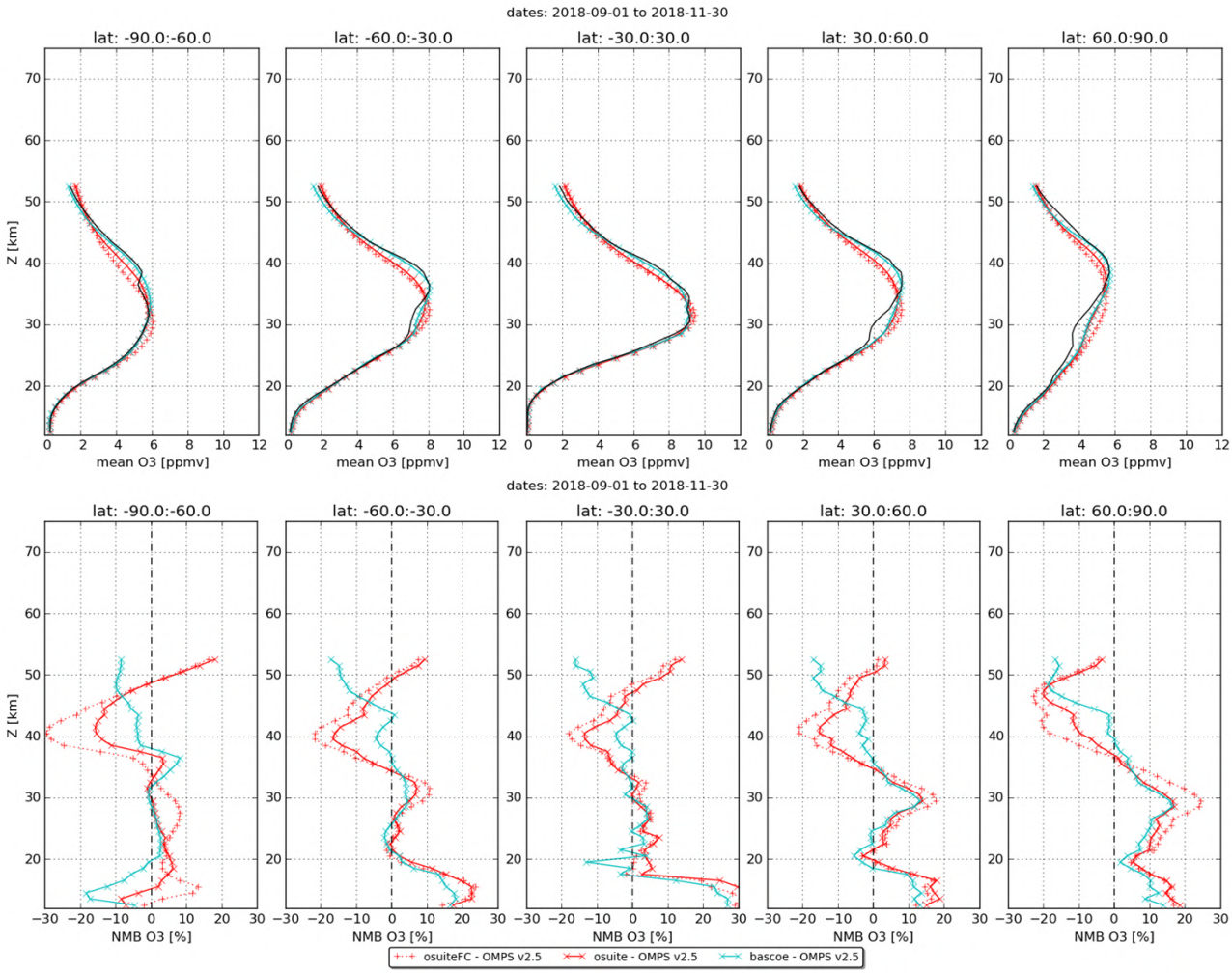


Figure 8.3.3: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with OMPS-LP v2.5 observations for the period September to November 2018.

The negative bias above 40km is confirmed by the ACEFTS profiles, otherwise there is a good agreement in the middle and lower stratosphere.

It must be noted that the different instruments have a variety of spatial and temporal coverage: for a 3-month period and over the latitude bands considered, OMPS and Aura MLS (not shown) provide daily data with more than 40000 valid profiles, while ACE-FTS provides around 300 profiles in the polar region and 90 profiles in the tropics.

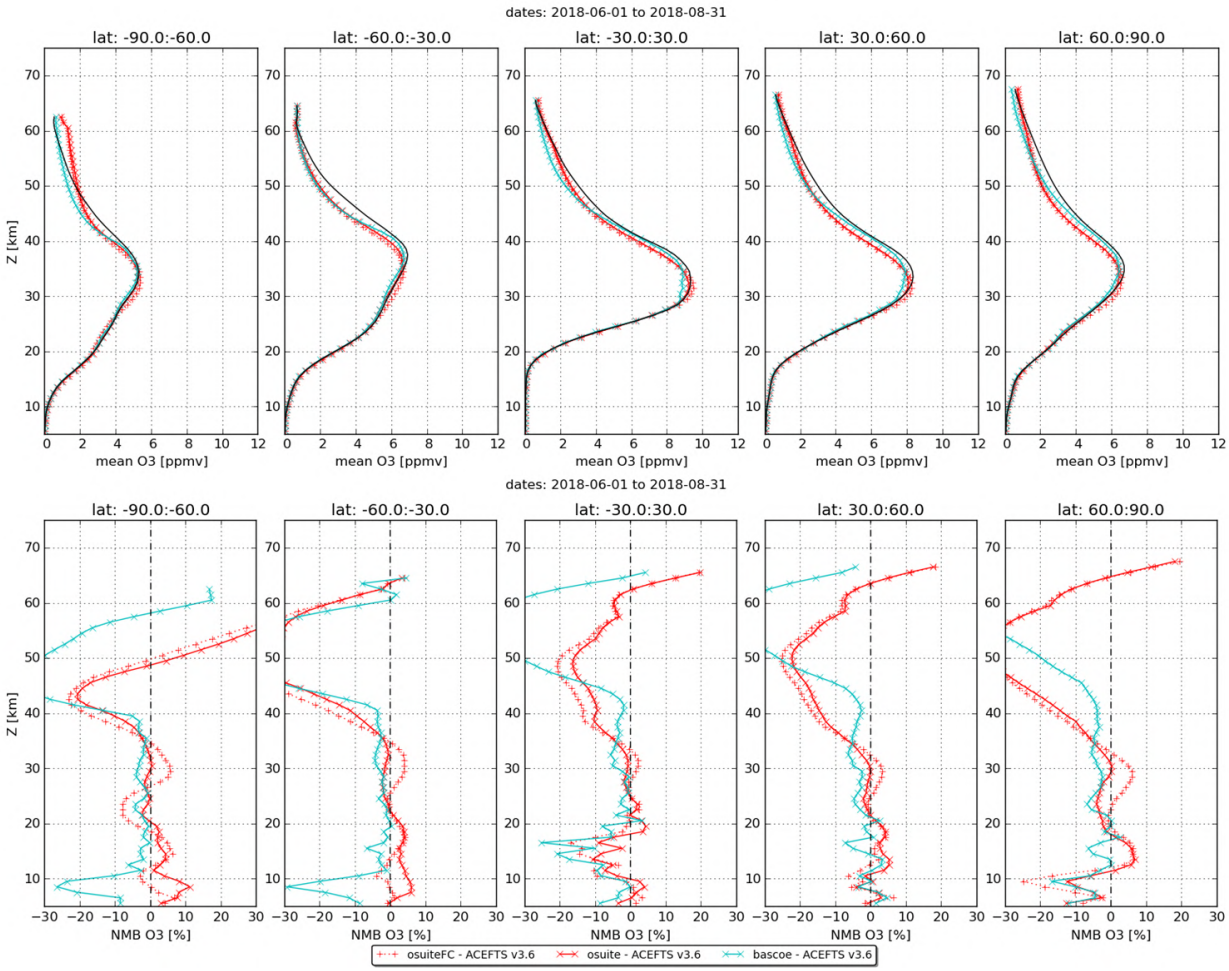


Figure 8.3.4: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with ACE-FTS observations for the period September to November 2018.

8.4 Stratospheric NO_2

The CAMS model uses a tropospheric chemistry scheme in combination with a parameterization for stratospheric ozone. Stratospheric ozone is also well constrained by satellite observations. Therefore, the only useful product in the stratosphere is ozone, and all other compounds, including NO2, should not be used, as demonstrated by the validation results presented here.

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 NO_2 columns. Monthly mean stratospheric NO_2 columns from SCIAMACHY and GOME-2 have relatively small errors on the order of 20% in the tropics and in mid-latitudes in summer and even lower errors at mid-latitudes in winter. As the time resolution of the saved model files is rather coarse and NO_x photochemistry in the stratosphere has a large impact on the NO_2 columns at low sun, some uncertainty is introduced by the time interpolation at high latitudes in winter.

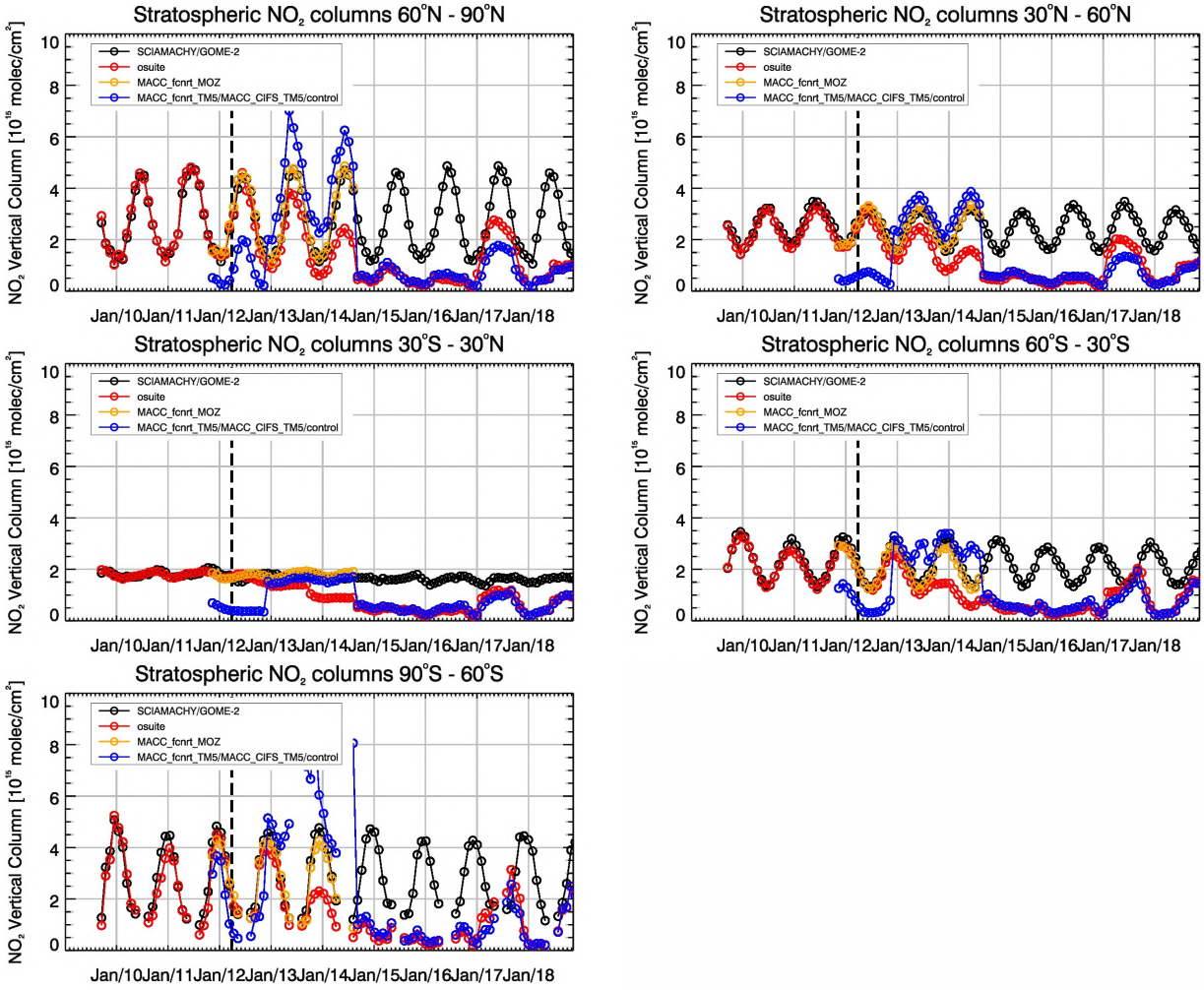


Figure 8.4.1: Time series of average stratospheric NO2 columns [10^{15} molec cm-2] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012) compared to model results for different latitude bands. See text for details. The blue line shows MACC_fcnrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 until August 2014 and control results from September 2014 onwards (the model run without data assimilation is termed control since Sep 2014). The vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

As shown in Figure 8.4.1, amplitude and seasonality of satellite stratospheric NO_2 columns are poorly modelled with CB05-based chemistry runs including the more recent versions of the o-suite. The significant differences between observations and CB05 chemistry runs, i.e. a strong underestimation of satellite retrievals by models, can be explained by the missing stratospheric chemistry for these model versions. The only constraint on stratospheric NO_x is implicitly made by fixing the HNO_3/O_3 ratio at the 10 hPa level. This assumption, in combination with the changing model settings for stratospheric O_3 for control compared to MACC_CIFS_TM5, may explain some of the jumps we see in stratospheric NO_2. In any of these runs the stratospheric NO_2 is poorly constrained. It clearly indicates that stratospheric NO_2 in the latest versions of the o-suite is not a useful product and should be disregarded. However, model simulated values increased with an upgrade of the osuite in February 2017, so that simulations are closer to the satellite observations for 2017 only, especially for northern hemisphere latitude bands where seasonality seems to have

been reproduced (in contrast to the Southern Hemisphere) by the o-suite apart from the pronounced underestimation. O-suite values are larger than the control in 2017 at all latitude bands. However, the better agreement found for 2017 does not continue for 2018 and values decreased again to the magnitude of 2015-2016 runs at all latitude bands.

Comparison of the o-suite from July 2012 until August 2014 with the other model runs and satellite observations shows that the previous version of the o-suite stratospheric NO_2 columns had a systematic low bias relative to those from MACC_fcnrt_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_fcnrt_MOZ, running until September 2014), especially northwards of 30°S. Details on the NO_2 evaluation can be found at: http://www.doas-bremen.de/macc/macc_veri_iup_home.html.

9. Validation results for greenhouse gases

This section describes the NRT validation of the pre-operational, high resolution forecast of CO_2 and CH_4 from 1st December 2017 to 1st December 2018 based on observations from 17 surface stations, located in Western Europe; 15 TCCON stations measuring xCO_2 and xCH_4 total columns, and 13 NDACC stations measuring partial and total CH_4 columns. We compare the observations to the high-resolution forecast experiment (*gqpe, Tco1279; 9x9 km*), coupled to the analysis experiment (*gqiq, Tco399, 25x25 km*). This forecast experiment is using the IFS model cycle CY43R1, and has been officially implemented on 1st Nov. 2017.

9.1 CH4 and CO2 validation against ICOS observations

The CO_2 and CH_4 simulations from the analysis and high-resolution forecast have been compared to the 17 ICOS stations. The near–real time data processing of the in-situ measurements is ensured by the Atmospheric Thematic Center (Hazan *et al.*, 2016). Among the 17 stations we can distinguish three sites located on top of mountain (PUY, JFJ, CMN), two background sites (PAL, ZEP) and 12 tall towers. For the later we consider only in this report the highest sampling level which is at least at 100m above the ground.

For CO_2, the best correlations are obtained for the background stations, but the variability is overestimated (Figure 9.1.1). For CH_4, there is not much difference in the correlation coefficients between the background sites and the tall towers, which are close to 0.8 with two exceptions at Monte Cimone and Ispra where correlations are below 0.6. Overall, we notice a small degradation of the CH_4 correlations with the high-resolution forecast experiment compared to the analysis (Figure 9.1.1). This is particularly true for a station like the Trainou tall tower as shown on figure 9.1.2. In this case we see that several spikes are wrongly simulated by the models in summer, and their amplitude is getting worse with the high-resolution forecast. This is probably due to the vicinity of hot spot emissions in Paris area, either mis-located or overestimated in the emission inventory, whose influence is amplified when using higher resolution. The two outliers (Ispra, and Monte Cimone for CH_4) are located respectively in the Po valley and at the top of a mountain, so the difficulty of the model to properly reproduce the observed variabilities probably results from the complex orography.

We carried out an evaluation of the model performance based on a composite of all European stations (Figure 9.1.3). Considering the results found at few stations like Ispra and Monte Cimone, we will have to consider for the next report to remove those outliers. Overall we observe an overestimation of CO_2 and CH_4 concentrations most of the time, with a seasonal dependence. For CO_2, the overestimation is maximal in Spring (Feb-June), and the model concentrations are decreasing below observations in late Summer and Fall. This overestimation of the CO_2 seasonal amplitude by ±1% was already described in previous reports. For CH_4 the results at the European surface stations is quite similar, with the simulations being too high by 1% (20 ppb) most of the times, and getting lower between July and October 2018. We also observe a decrease of the correlation coefficients in summertime for both CO_2 and CH_4. The worst correlation for CO_2 is observed in June 2018. It seems to be related to the influence of the drought 2018. This event

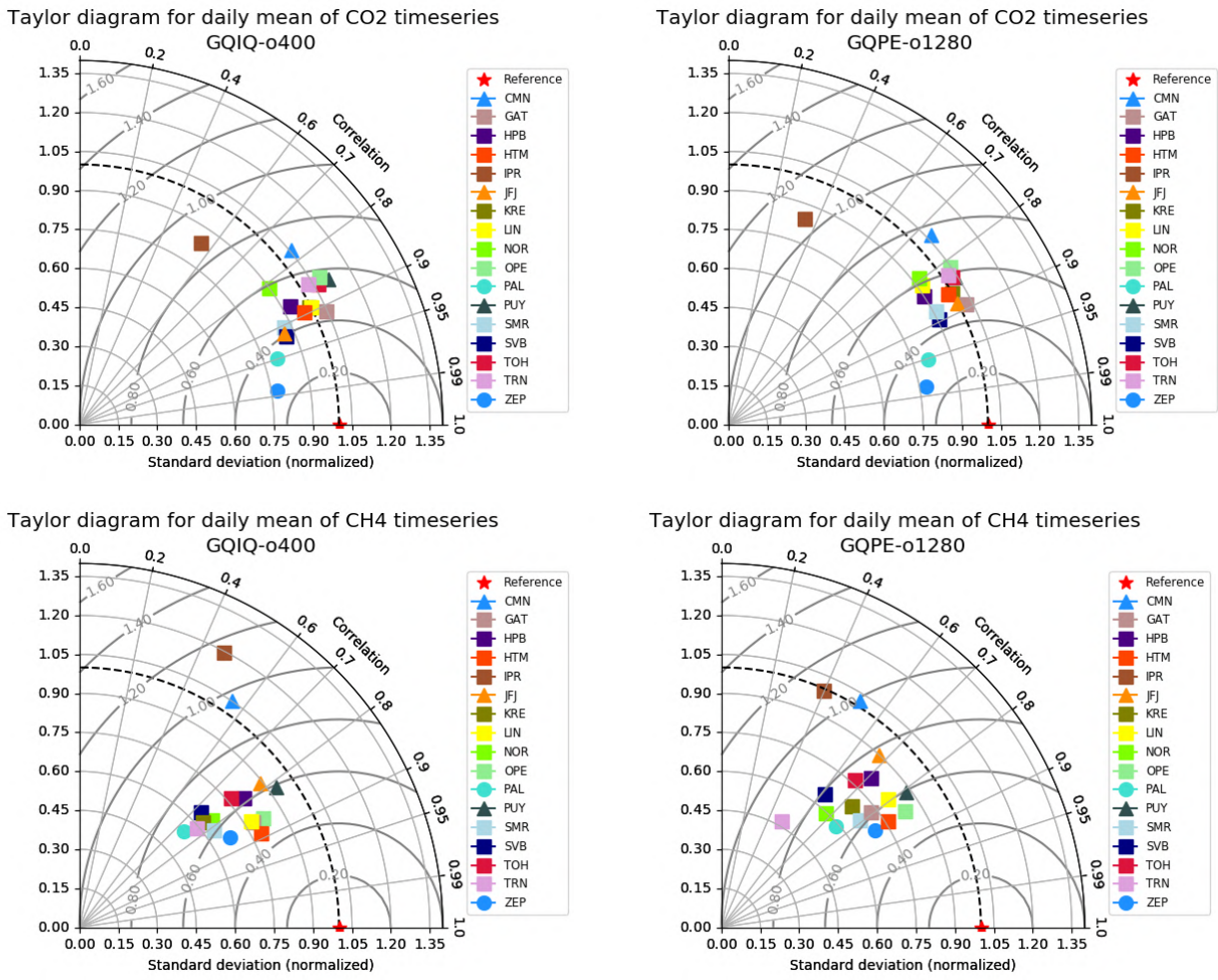


Figure 9.1.1: Taylor diagrams relating the standard deviations for the model and time series of CO2 (above) and CH4 (below) mole fractions and their correlation. The left panels show the analysis, and the right panels the high-resolution forecast. All timeseries are normalized such that the standard deviation of the model is equal to one.

reduces by 2-4 ppm the CO_2 drawdown at ICOS stations located in Northern Europe in summer 2018 compared to 2017. It seems that the resulting CO_2 anomaly is overestimated by the model, which display an abrupt increase of concentrations between May and July resulting in a 2% bias (Figure 9.1.4).

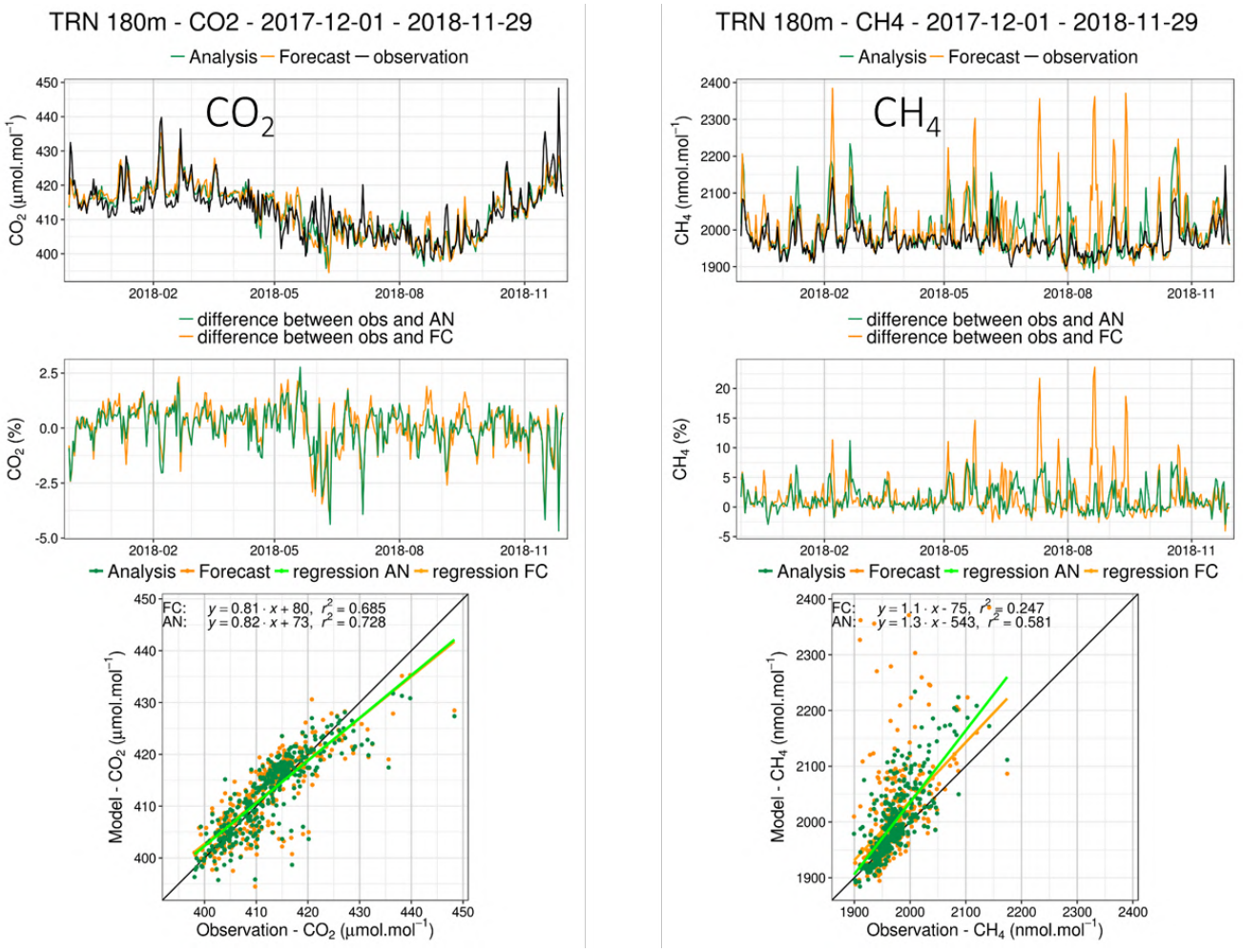


Figure 9.1.2: Comparison of CO2 (left) and CH4 (right) daily means observed (black) with the analysis model (red) and the high-resolution forecast (blue) at the Trainou tall tower. Middle: differences of the observations minus the simulations. Below: Linear fit between observations and simulations.

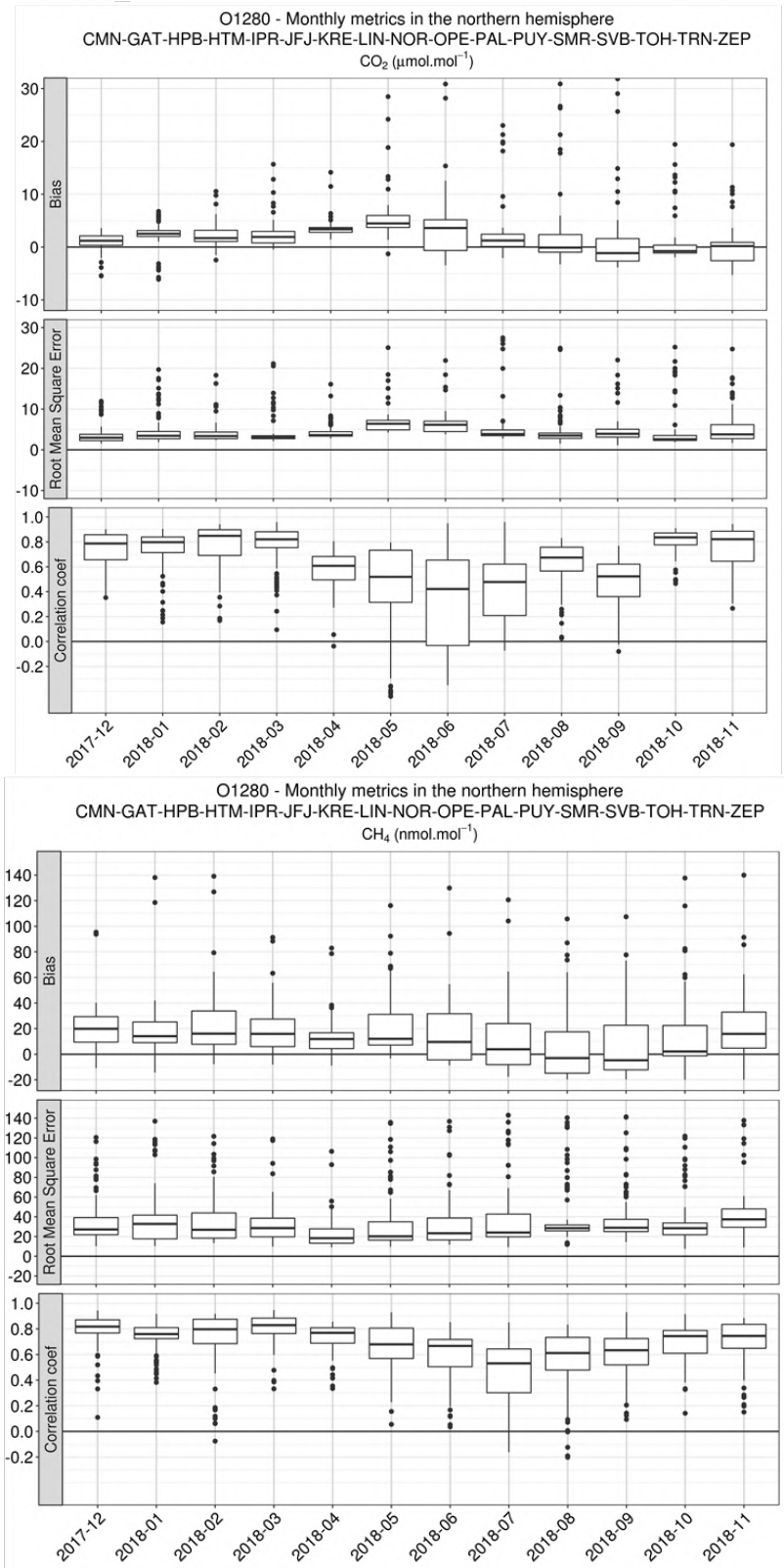


Figure 9.1.3: Monthly statistics (bias, RMSE, correlation coefficients) of the high-resolution forecast experiment compared to ICOS surface measurements. The results obtained for all European sites (see the list of sites in the title) are averaged. Top: CO_2; Bottom: CH_4.

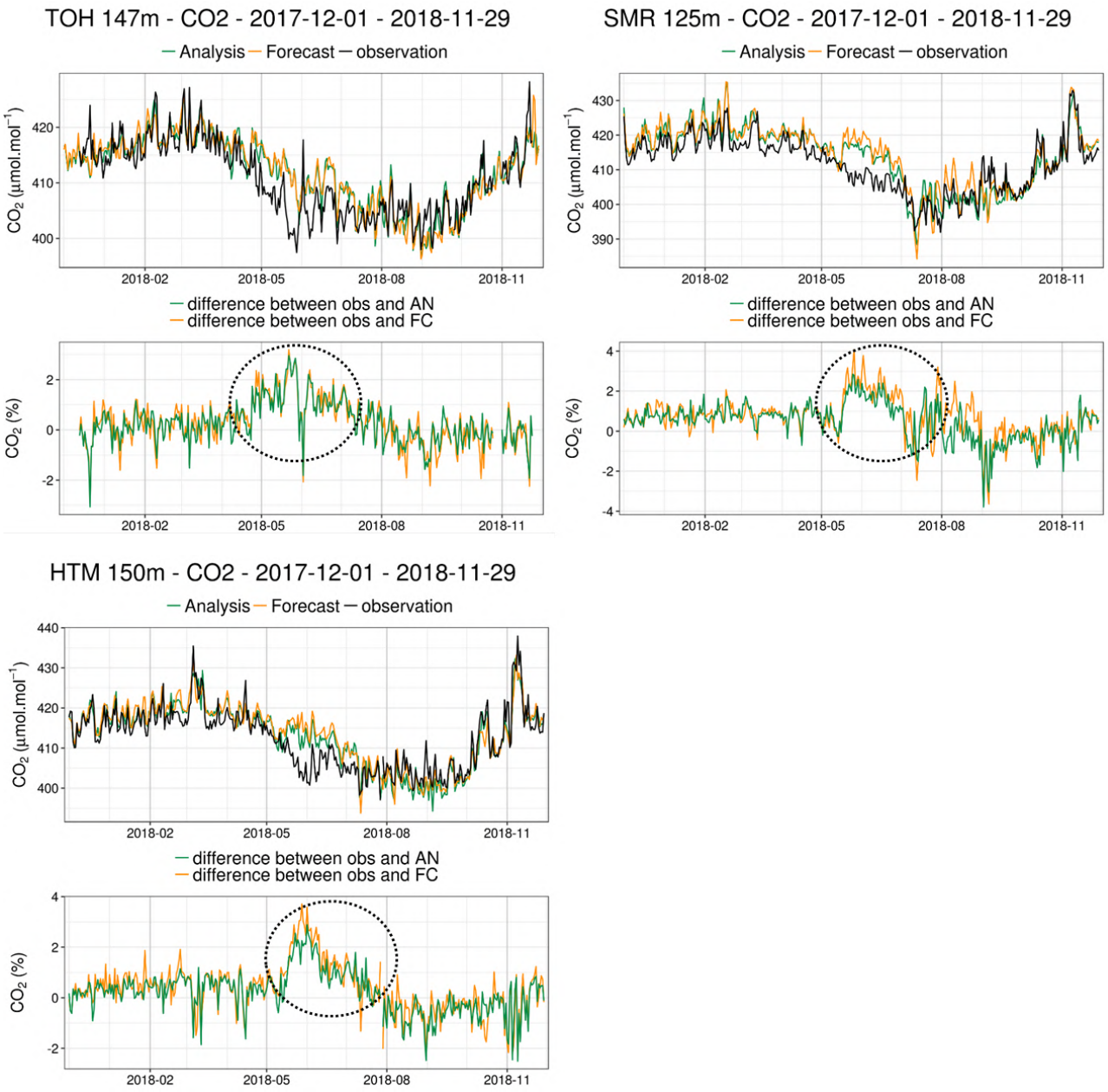


Figure 9.1.4: Comparison of CO2 daily means observed (black) with the analysis model (red) and the high-resolution forecast (blue) at the Torfhaus (TOH, Germany), Hyttyala (SMR, Finland) and Hyltemossa (HTM, Sweden) tall towers. Low panels: differences of the observations minus the simulations.

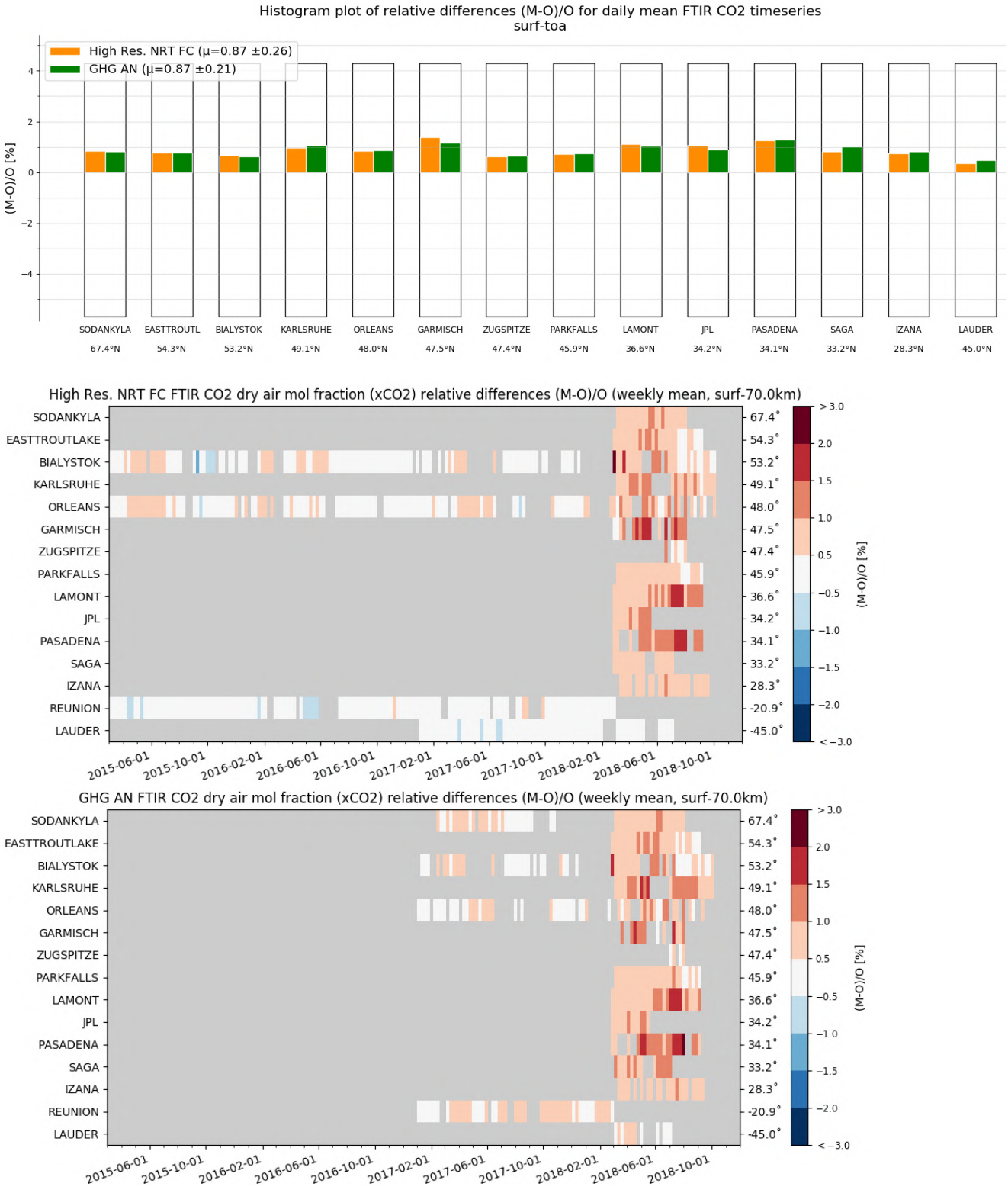


Figure 9.2.1. Relative mean bias for CO_2 column mole fractions (MB, %), standard deviation (STD, %) and number of observations used (top) for the period March-Nov 2018 and weekly mean biases for the last 2 years (bottom, model upgrades are indicated in black). The stations are sorted with decreasing latitude (northern to southern hemisphere).

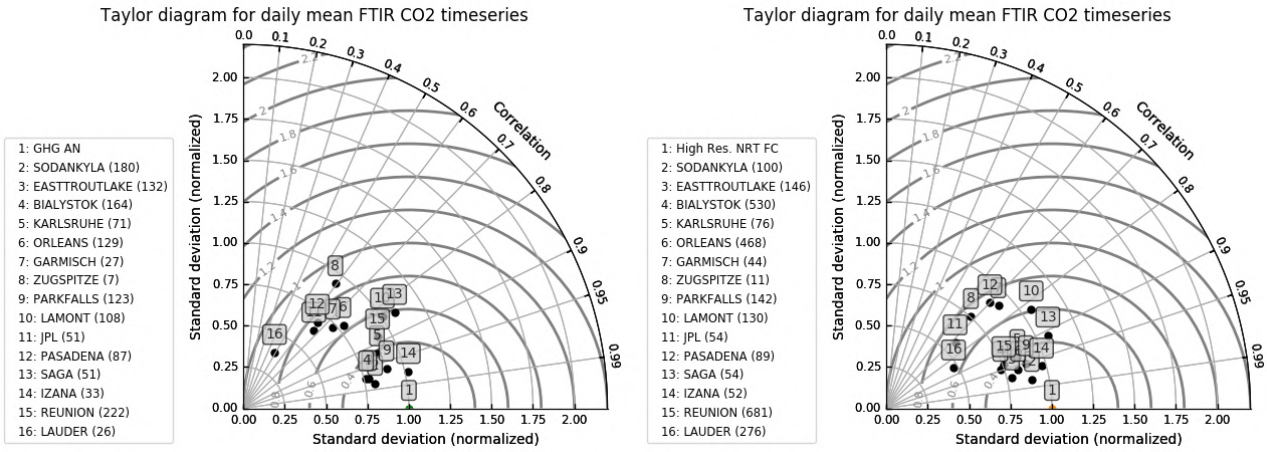


Figure 9.2.2: Taylor diagrams relating the standard deviations for the model /GB time series of column averaged CO2 mole fractions and their correlation. All timeseries are normalized such that the standard deviation of the model is 1. Left: GHG analysis; Right: GHG high-resolution forecast.

9.2 CH_4 and CO_2 validation against TCCON observations

For the validation column averaged mole fractions of CO_2 and CH_4 (denoted as XCO_2 and XCH_4) from the Total Carbon Column Observing Network (TCCON) are used. Column averaged mole fractions provide different information than the in-situ measurements and are therefore complementary to the in situ data. For example, if models suffer from problems in vertical transport, the combination of TCCON and surface in situ measurements will provide a means to detect this.

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. Only measurements within 2.5h around local noon have been used for the comparison.

Carbon dioxide (CO_2)

Figure 9.2.1 shows that model data overestimates the CO_2 at all stations. For most stations the relative difference is about 1% or less. Exceptions are a) the urban sites closed to Los Angeles (JPL and Pasadena) and b) the Lamont site in Oklahoma. At Lamont the drawdown is significantly stronger in the observations. The Taylor plots are given in Figure 9.2.2.

Methane (CH_4)

Figure 9.2.3 shows that during the reporting period the models agree or underestimate the CH_4 for all sites besides the urban sites JPL and Pasadena. At the latter two sites the model overestimates the CH_4. It can also be seen that the very last month of the comparison shows the highest underestimation by the model. Fig 9.2.4 shows that this underestimation is a feature that occurred also during previous years for this time of the year.

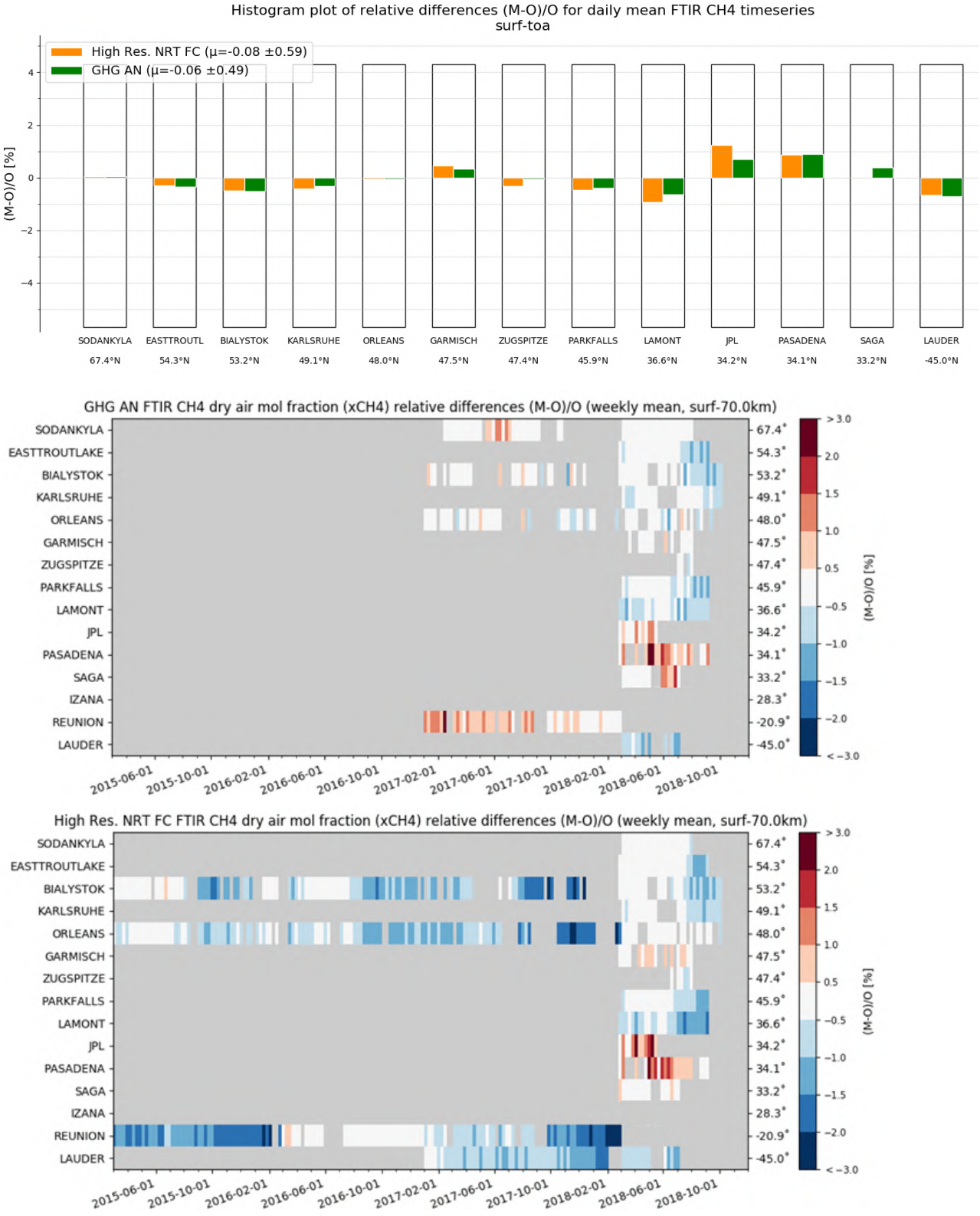


Figure 9.2.3: Relative mean bias for CH_4 column mole fractions (MB, %), standard deviation (STD, %) and number of observations used for the period March-Nov 2018 (top) and weekly mean biases for the last 2 years (bottom, model upgrades are indicated in black). The stations are sorted with decreasing latitude (northern to southern hemisphere).

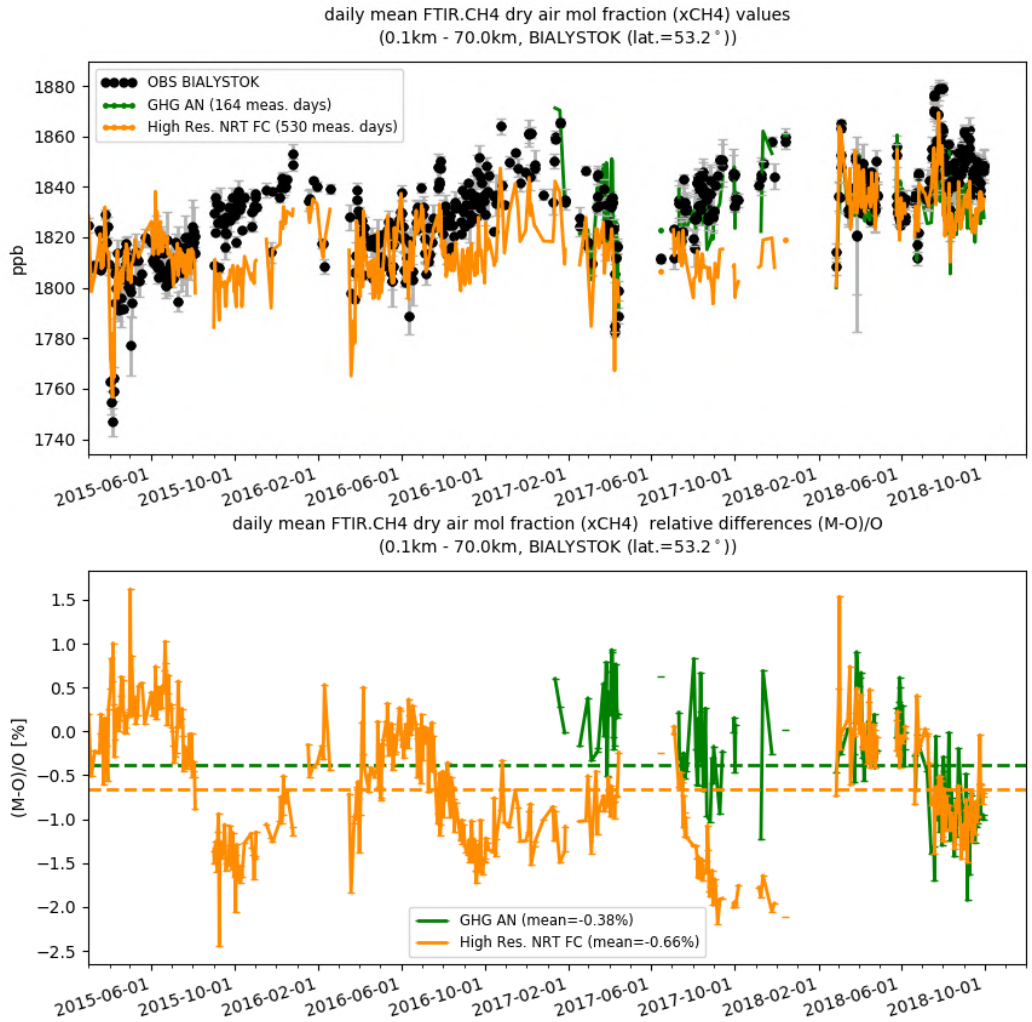


Figure 9.2.4: Seasonality of CH4 at Bialystok. The upper plot shows the models simulations and the measurements. The lower plot shows the relative differences.

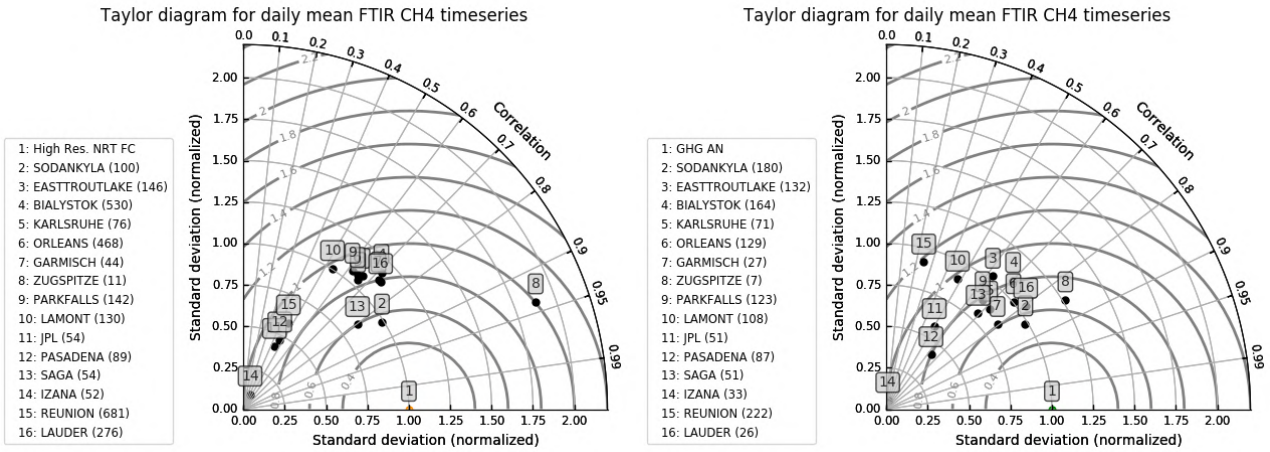


Figure 9.2.5: Taylor diagrams relating the standard deviations for the model /GB time series of column averaged CH4 mole fractions and their correlation. All timeseries are normalized such that the std of the model is 1. Left: high-resolution forecast. Right: analysis.

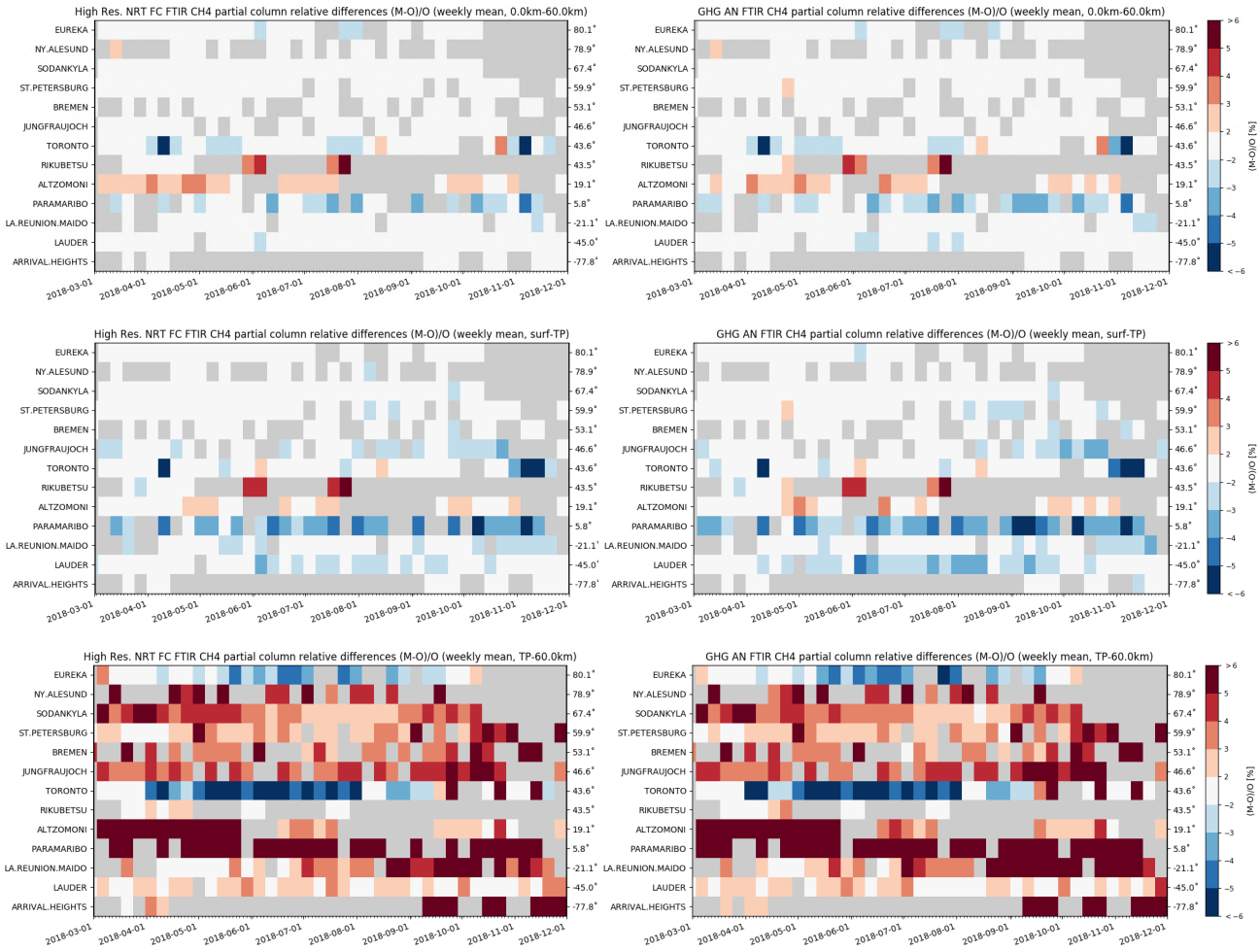


Figure 9.3.1: Weekly mean relative bias for total (top row), tropospheric (middle row) and stratospheric CH_4 columns (bottom row) for the period March 2018 – November 2018 for high resolution forecast (left) and the analysis (right). The mid latitude stations at Rikubetsu and Altzomoni show a strong overestimation for both models of the CH_4 column. The overall uncertainty for the CH_4 total column measurements is approximately 4%.

9.3 Validation against FTIR observations from the NDACC network

In this section, we compare the CH_4 profiles of the CAMS GHG models with FTIR measurements at different FTIR stations within the NDACC network. These ground-based, remote-sensing instruments are sensitive to the CH_4 abundance in the troposphere and lower stratosphere, i.e. between the surface and up to 25 km altitude. Tropospheric and stratospheric CH_4 columns are calculated from the FTIR profile data and used to validate corresponding columns obtained from the model data. A description of the instruments and applied methodologies can be found at http://nors.aeronomie.be. The typical uncertainty on the FTIR tropospheric column is 3.5%, while the uncertainty on the stratospheric column is 7.5%, adding together to a 4% uncertainty on the total column. The systematic uncertainty is large for the NDACC methane product mostly due to higher spectroscopic uncertainties.

Figure 9.3.2: Daily mean of relative differences for tropospheric CH4 columns (left) and stratospheric CH4 columns (right) at Eureka (top) and St Petersburg (bottom). The performance change at Eureka changes from a positive bias in March-April for the stratospheric column to a negative bias in July-September, while at St Petersburg the tropospheric column performs worse during June-October.

Figure 9.3.1 shows that the tropospheric columns of CH_4 agree well and only small differences appear between the analysis and the high-resolution model. The total column validation results correspond to the TCCON observations in Figure 9.2.3, where an overestimation is seen at the mid latitude stations in the North Hemisphere (Rikubetsu and Altzomoni) and an underestimation in Lauder (New Zealand).

At some sites a clear trend is observed in either the tropospheric or stratospheric concentrations. Due to the limited time period, it is unclear if this is related to seasonal dependent model performance or a long-term trend. In Figure 9.3.2 the tropospheric and stratospheric relative difference time series are plotted at Eureka and St. Petersburg.

Figure 9.3.3 shows Taylor diagrams for the SON time period and for a selected number of sites: stations with limited observations are left out. The standard deviation of the Toronto FTIR CH_4 columns is approx. 3 times higher compared to the GHG model data and falls out of the plot. Assimilation has a small effect on the correlation coefficients.

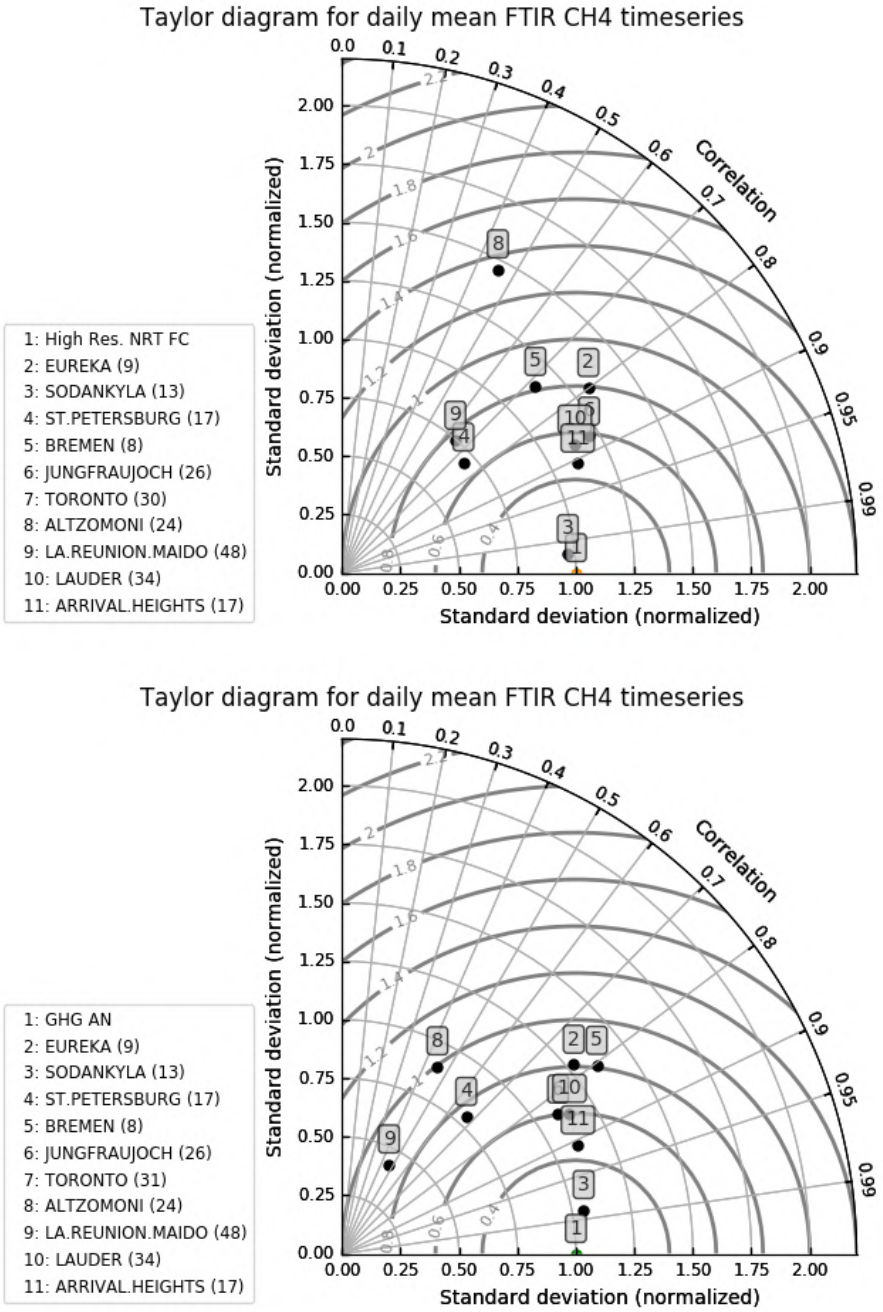


Figure 9.3.3: Taylor diagrams relating the standard deviations for the model /GB time series of total CH4 column data and their correlation for the period SON 2018 (the stations with a limited number of measurements are left out). All timeseries are normalized such that the std of the model is 1. The correlation for the analysis time series is larger for most sites. For Altzomoni and Reunion (Maido) the assimilation decreases the ratio of the standard deviations of both time series significantly: the analysis methane columns are more variable compared to the high-resolution forecast. Top: high-resolution forecast. Bottom: analysis.

10. Event studies

10.1 Dusty Mediterranean during October 2018

In second half of October 2018, MODIS satellite detected two dust outbreaks over Central and Eastern Mediterranean Basin (see Figure 10.1.1). One outbreak occurred on 20th October, originating from Syria and transported towards Turkey. A second dust event was observed on 16th October which originated in Libya and affected the Central Mediterranean. Both events were nicely tracked by the AERONET sun photometers in the Eastern Mediterranean (see Figure 7.4.2).

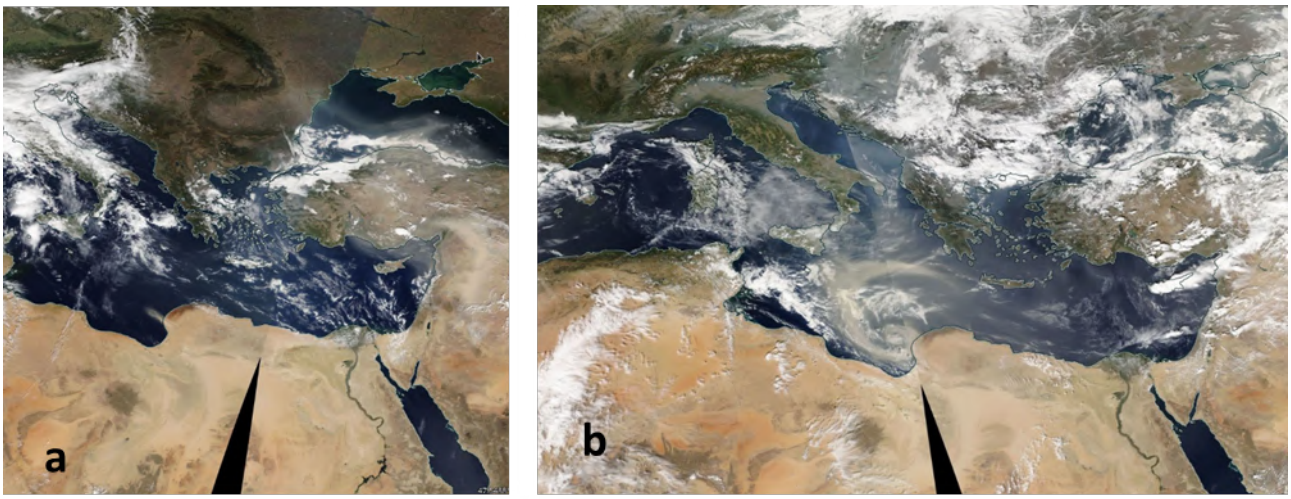


Figure 10.1.1: Daily composite of NASA MODIS/Terra on 16th October (left image) and 20th October (right image) 2018 over Central-Eastern Mediterranean.

CAMS AOD o-suite did timely reproduce the spatial distribution of the two dust plumes as shown the comparison with MODIS/Aqua AOD comparison (see Figure 10.1.2) despite the model tends to underestimate the observed maximum values particularly the second event with origin in Libya in which MODIS shows values of AOD up to 1.5 and the o-suite predicts AOD values around 0.7 on 20th October. This second dust outbreak was associated to a deep depression.

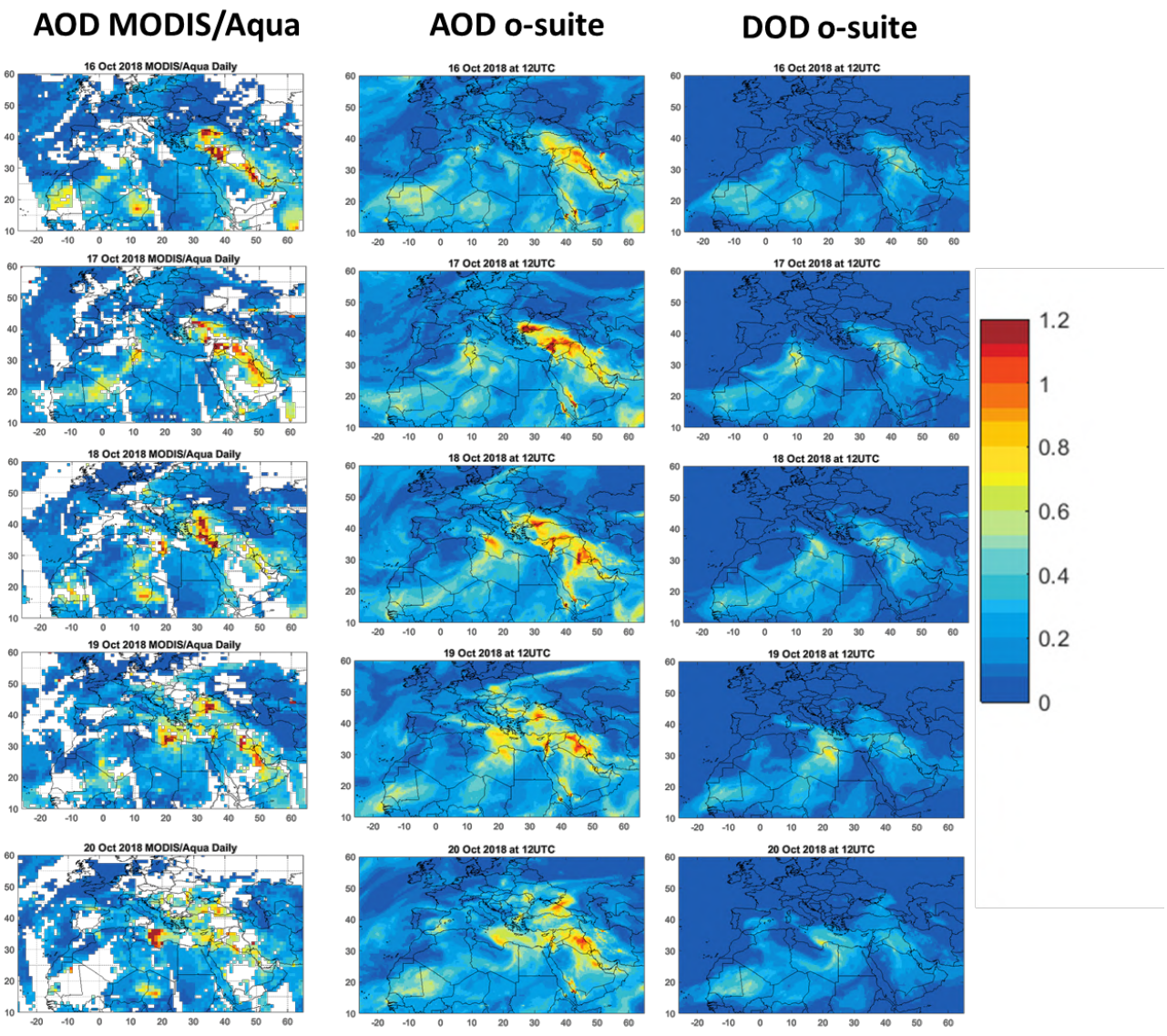


Figure 10.1.2. AOD from MODIS/Aqua Collection 6.1 Daily Level 3 product as well as AOD and DOD at 12UTC from o-suite (central and left columns) for 16-20 October 2018.

10.2 Fire events in Alaska in July 2018

Several fire events were detected during July 2018 in Alaska region. The first CO peak detected by MOPITT and IASI from July 7 to 9 was well captured by both model runs. A second fire event detected around July 22 was underestimated by both model runs by about 20%. Apart from the days with high CO values, the control run shows slightly better agreement with the satellite data compared to o-suite.

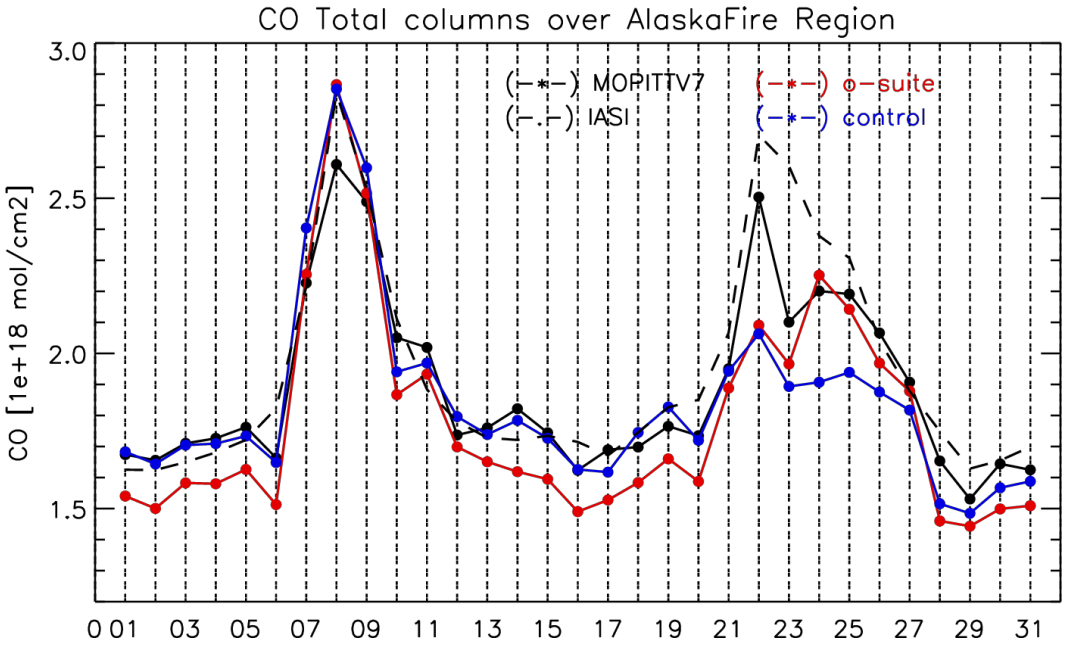


Fig. 10.2.1. Time series of CO total columns during July 2018 over Alaska regions from: MOPITT v7 (black), IASI (black), o-suite (red) and control (blue) runs.

11. 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. Huneeus, 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.

Cammas, J.P., Brioude J., Chaboureau J.-P., Duron J., Mari C., Mascart P., Nédélec P., Smit H., Pätz H.-W., Volz-Thomas A., Stohl A., and Fromm M., Injection in the lower stratosphere of biomass fire emissions followed by long-range transport: a MOZAIC case study. Atmos. Chem. Phys., 9, 5829-5846, 2009

Cariolle, D. and Teyssèdre, 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.

Dentener, F., et al., 2006: Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom, Atmos. Chem. Phys., 6, 4321 – 4344.

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 maufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonsondes. J. Geophys. Res.113, D04307, doi:10.1029/2007JD008975

Douros, J., S. Basart, A. Benedictow, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, H. Clark, E. Cuevas, H.J. Eskes, H. Flentje, K. M. Hansen, J. Kapsomenakis, B. Langerock, K. Petersen, 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_2015SC2_D.84.8.1.1-2017_observations_v2.pdf, October 2017. Available from: http://atmosphere.copernicus.eu/user-support/validation/verification-global-services.

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

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

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

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

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

Eskes, H., Wagner, A., Schulz, M., Christophe, Y., Ramonet, M., Basart, S., Benedictow, A., Blechschmidt, A.-M., Chabrillat, S., Clark, H., Cuevas, E., Flentje, H., Hansen, K. M., Im, U., Kapsomenakis, J., Langerock, B., Petersen, K., Richter, A., Sudarchikova, N., Thouret, V., Warneke, T., and Zerefos, C.: Validation report of the CAMS near-real time global atmospheric composition service December 2016 - February 2017, CAMS84_2015SC2_D84.1.1.7_2017DJF_v1.1, June 2017. Available from: http://atmosphere.copernicus.eu/user-support/validation/verification-global-services

Eskes, H. J., T. Antonakaki, S. Basart, A. Benedictow, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, H. Clark, E. Cuevas, K. M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, K. Petersen, A. Richter, M. Schulz, N. Sudarchikova, V. Thouret, A. Wagner, C. Zerefos, Upgrade verification note for the CAMS near-real time global atmospheric composition service, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2015SC2_D84.3.1.3_201706_esuite_v1.pdf, July 2017.

Eskes, H. J., S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, H. Clark, E. Cuevas, K. M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, K. Petersen, M. Schulz, A. Wagner, C. Zerefos, Upgrade verification note for the CAMS near-real time global atmospheric composition service, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2015SC3_D84.3.1.5_201802_esuite_v1.pdf, February 2018; Eskes et al., Upgrade verification note for the CAMS near-real time global atmospheric composition service, Addendum July 2018, CAMS84_2015SC3_D84.3.1.5_201802_esuite_v1.pdf.

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.

Flemming, J., Benedetti, A., Inness, A., Engelen, R. J., Jones, L., Huijnen, V., Remy, S., Parrington, M., Suttie, M., Bozzo, A., Peuch, V.-H., Akritidis, D., and Katragkou, E.: The CAMS interim Reanalysis of Carbon Monoxide, Ozone and Aerosol for 2003–2015, Atmos. Chem. Phys., 17, 1945-1983, doi:10.5194/acp-17-1945-2017, 2017.

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

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

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

Hommel, R., Eichmann, K.-U., Aschmann, J., Bramstedt, K., Weber, M., von Savigny, C., Richter, A., Rozanov, A., Wittrock, F., Khosrawi, F., Bauer, R., and Burrows, J. P.: Chemical ozone loss and ozone mini-hole event during the Arctic winter 2010/2011 as observed by SCIAMACHY and GOME-2, Atmos. Chem. Phys., 14, 3247-3276, doi:10.5194/acp-14-3247-2014, 2014.

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

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

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

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

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

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

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

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

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

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

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

Massart, S., Flemming, J., Cariolle, D., Jones, L., High resolution CO tracer forecasts, MACC-III Deliverable D22.04, May 2015, available from http://www.gmes-atmosphere.eu/documents/macciii/deliverables/grg

Morcrette, J.-J., O. Boucher, L. Jones, D. Salmond, P. Bechtold, A. Beljaars, A. Benedetti, A. Bonet, J. W. Kaiser, M. Razinger, M. Schulz, S. Serrar, A. J. Simmons, M. Sofiev, M. Suttie, A. M. Tompkins, and A. Untch: Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part I: Forward modelling, J. Geophys. Res., 114, D06206, doi:10.1029/2008JD011235, 2009.

Richter, A., Burrows, J. P., Nüß, H., Granier, C, Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, Nature, 437, 129-132, doi: 10.1038/nature04092, 2005

*Richter, A., Begoin, M., Hilboll, A., and Burrows, J. P.: An improved NO2 retrieval for the GOME-2 satellite instrument, Atmos. Meas. Tech. , **4**, 1147-1159, doi:10.5194/amt-4-1147-2011, 2011*

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, 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.

Stavrakou,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*

Wennberg, P. O., Mui, W., Wunch, D., Kort, E. A., Blake, D. R., Atlas, E. L., Santoni, G. W., Wofsy, S. C., Diskin, G. S., Jeong, S., and Fischer, M. L.: On the sources of methane to the Los Angeles atmosphere, Environ. Sci. Technol., 46, 9282–9289, https://doi.org/10.1021/es301138y, 2012

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.

Wunch, D., Wennberg, P. O., Toon, G. C., Keppel-Aleks, G., and Yavin, Y. G.: Emissions of greenhouse gases from a North American megacity, Geophys. Res. Lett., 36, 1–5, https://doi.org/10.1029/2009GL039825, 2009.

Annex 1: Acknowledgements

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

Tropospheric reactive gases reactive gases

Annette Wagner, MPG (editor, O3 sondes, GAW data)
Yasmine Bennouna, 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)
Natalia Sudarchikova, satellite IR observations (MPG)
Kaj Hansen, Ulas Im, AU (Arctic theme)
Bavo Langerock, BIRA (NDACC)

Tropospheric aerosol

Michael Schulz, MetNo (editor, Aerocom, Aeronet)
Anna Benedictow, Jan Griesfeller, MetNo (Aerocom, Aeronet)
Sara Basart, MTeresa Pay, Oriol Jorba, BSC-CNS (Aeronet, MODIS, AirBase, SDS-WAS NAMEE RC)
Emilio Cuevas, AEMET (Aeronet, MODIS, AirBase, SDS-WAS NAMEE RC)
Harald Flentje, DWD (Backscatter profiles)

Stratospheric reactive gases

Yves Christophe, BIRA (editor, model-satellite intercomparisons)
Simon Chabrillat, BIRA (model intercomparisons)
Annette Wagner, MPI-M (O_3 sondes)
Bavo Langerock, BIRA (NDACC FTIR, MWR, UVVIS DOAS, LIDAR)
Anne Blechschmidt and Andreas Richter, IUB-UB Bremen (SCIAMACHY/GOME-2 NO_2)

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.

The authors thank to all researchers, data providers and collaborators of 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 Node. Also special thank to Canary Government as well as AERONET, MODIS, U.K. Met Office MSG, MSG Eumetsat and EOSDIS World Viewer principal investigators and scientists for establishing and maintaining data used in the activities of the WMO SDS-WAS NAMEE Regional Center (http://sds-was.aemet.es/).

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 Maido, Uccle, Hohenpeissen, Mauna Loa, Lauder and Haute Provence.

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

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

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

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

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

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

We thank the Canadian Space Agency and ACE science team for providing level 2 data retrieved from ACE-FTS on the Canadian satellite SCISAT-1.

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 CO_2/CH_4 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.

The European Environment Information and Observation Network (Eionet) Air Quality portal provides details relevant for the reporting of air quality information from EU Member States and other EEA member and co-operating countries. This information is submitted according to Directives 2004/107/EC and 2008/50/EC of the European Parliament and of the Council.

Copernicus Atmosphere Monitoring Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-atmosphere.eu